

НАЦІОНАЛЬНИЙ УНІВЕРСИТЕТ «ОДЕСЬКА ПОЛІТЕХНІКА»
МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ

НАЦІОНАЛЬНИЙ УНІВЕРСИТЕТ «ОДЕСЬКА ПОЛІТЕХНІКА»
МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ

Кваліфікаційна наукова
праця на правах рукопису

ДОБРИШЕВ РУСЛАН ЄВГЕНОВИЧ

УДК 004.932.2

ДИСЕРТАЦІЯ

**МОДЕЛІ ТА МЕТОДИ ВІЗУАЛЬНОГО АНАЛІЗУ НАТОВПУ В СИСТЕМАХ
ІНТЕЛЕКТУАЛЬНОГО ВІДЕОСПОСТЕРЕЖЕННЯ**

Спеціальність: 122 – «Комп'ютерні науки»

Галузь знань: 12 – «Інформаційні технології»

Подається на здобуття наукового ступеня доктора філософії

Дисертація містить результати власних досліджень. Використання ідей, результатів і текстів інших авторів мають посилання на відповідне джерело.

_____ Р. Є. Добришев

Науковий керівник:

Максимов Максим Віталійович, доктор технічних наук, професор

Одеса – 2025

АНОТАЦІЯ

Добришев Р.Є. Моделі та методи візуального аналізу натовпу в системах інтелектуального відеоспостереження. – Кваліфікаційна наукова праця на правах рукопису.

Дисертація на здобуття наукового ступеня доктора філософії за спеціальністю 122 – Комп'ютерні науки. – Національний університет «Одеська політехніка», МОН України, Одеса, 2025.

У **вступі** показана актуальність вирішення задач візуального аналізу натовпу сучасними інтелектуальними системами відеоспостереження. Показана недосконалість існуючих моделей та методів візуального аналізу натовпу за умов великої щільності натовпу. Визначено об'єкт, предмет, задачі та методи дослідження; наведено наукову новизну та практичне значення отриманих результатів; висвітлено особистий внесок здобувача.

В **першому розділі** дисертаційної роботи був проведений аналіз проблеми візуального аналізу натовпу в системах інтелектуального відеоспостереження.

Проведений аналіз призначення та структури інтелектуальних систем відеоспостереження, сформульована задача візуального аналізу натовпу, а також виділені окремі підзадачі в даній області. Для кожної підзадачі проаналізовані наукові роботи, визначені виклики, проблеми та сучасний стан вирішення цієї підзадачі. Крім того, були визначені перспективи подальших досліджень в області візуального аналізу натовпу.

Визначено, що більшість сучасних методів підрахунку людей у натовпі не враховують геометричну структуру сцени, зокрема перспективні спотворення та варіації масштабу, які можуть суттєво впливати на точність оцінки щільності.

Для подолання даного недоліку був запропонований метод інтеграції геометричних ознак сцени, зокрема карти глибини, до глибинної архітектури згорткової нейронної мережі з метою адаптивного масштабування ознак і підвищення точності оцінки щільності натовпу.

Подальший аналіз задачі візуального аналізу натовпу показав, що існуючі методи моделювання руху в щільному натовпі мають низьку стійкість до оклюзій, не враховують довготривалі просторово-часові залежності та включають нерелевантні області сцени, що знижує точність і підвищує обчислювальні витрати. У складних динамічних сценах із невизначеними потоками сегментація ускладнюється ще більше.

Для подолання цього недоліку запропонований представлений метод сегментації потоків натовпу у сценах з високою щільністю. Метод використовує кластеризацію траєкторій на основі активних контурів, яка акцентує увагу на відокремленні популярних патернів руху в локальних регіонах натовпу.

Під час розробки методу сегментації потоків натовпу у сценах з високою щільністю було встановлено, що більшість підходів до моделювання траєкторій руху натовпу не враховують форму, напрямок, положення та щільність руху, ігнорують часову динаміку та локальний контекст, а також застосовують одномасштабний аналіз, що знижує точність у складних умовах.

Для подолання цього недоліку була запропонована модель представлення траєкторій, яка включає багатовимірні ознаки – форму, положення, напрямок і багатомасштабну щільність – з метою забезпечення стійкого та ефективного групування траєкторій і виявлення потоків натовпу у складних сценах.

Аналіз наукових робіт в області вирішення задачі виявлення аномальних ситуацій у сценах з натовпом високої щільності показав, що раннє виявлення аномалій у натовпі є критично важливим для громадської безпеки, проте ця задача ускладнена через оклюзії, змінну щільність та складне тло. Існують два основні підходи: об'єктні, які відстежують окремих осіб, але малоефективні в переповнених сценах; та холістичні, що розглядають натовп як єдине ціле, але часто не помічають локальних порушень. Методи короткочасного аналізу руху не відображають довготривалі просторово-часові закономірності. Підходи на основі частинок частково це компенсують, однак мають проблеми з обробкою щільних областей через перекриття векторів руху.

Для подолання цього недоліку був запропонований метод виявлення аномальних сцен натовпу за допомогою оцінки ентропії спрямованих міні-треків, побудованих з просторово-часових точок інтересу. Запропонований метод відкидає короткі та зашумлені траєкторії для збереження геометричної узгодженості напрямку та інтегрує оцінку ентропії з аналізом часового відхилення заповнюваності, щоб виявити різкі зміни в динаміці сцени.

За результатами проведеного аналізу сформульовано мету та завдання досліджень.

В **другому розділі** дисертаційної роботи запропонований метод підвищення точності оцінки щільності натовпу шляхом інтеграції геометричних ознак сцени. В рамках запропонованого методу, пропонується використовувати інформацію про глибину сцени як допоміжну інформацію для адаптивного формування масштабованих зображень, що дозволяє краще відображати реальні розміри об'єктів у просторі та підвищити точність прогнозу щільності.

Впровадження інформації стосовно глибини сцени у нейромережеву модель відбувається завдяки додаванню спеціального модуля інтеграції інформації про глибину сцени, який додає інформацію про глибину сцени та змінює величини окремих ознак в окремих ділянках зображення для отримання представлення зображення з урахуванням глибини.

Робота запропонованого методу складається з наступних етапів:

1. Блок впровадження глибини в складі модуля модифікує простір ознак за допомогою карти глибини.

2. Блок корекції глибини уточнює карту глибини для генерації ваг з урахуванням масштабу. Генерація ваг відбувається за допомогою механізму просторової уваги, що дозволяє надати дані про розташування об'єктів в певних ділянках зображення.

3. Блок вбудовування застосовує отримані ваги для повторної зміни величини оригінальних ознак у конкретній ділянці зображення. Модифіковані таким чином ознаки демонструватимуть геометричні відмінності в масштабах між об'єктами на

передньому плані в різних положеннях, що безпосередньо сприятиме кращій оцінці значень щільності з урахуванням масштабу.

Таким чином, сформульовано **перший пункт наукової новизни:** *удосконалено* метод інтеграції інформації про геометричні ознаки сцени у глибинні нейронні мережі, які здійснюють оцінку щільності натовпу. Використання запропонованого методу дозволяє краще відображати реальні розміри об'єктів у просторі та підвищити достовірність оцінки щільності.

В **третьому розділі** дисертаційної роботи запропоновані моделі та методи обробки та аналізу сцен з високою щільністю натовпу для виявлення аномалій.

Для **підвищення точності сегментації потоків натовпу в сценах з високою щільністю натовпу** був запропонований метод сегментації потоків натовпу на основі кластеризації траєкторій з використанням активних контурів, що дозволяє ефективно виділяти локальні патерни колективного руху та підвищити якість аналізу складних сцен.

Робота запропонованого методу складається з наступних кроків:

1. Сегментація на основі активних контурів – відокремлення області натовпу від усього кадру.
2. Виділення траєкторій – виявлення точок інтересу в блочній структурі і відстеження цих блоків на кадрах для виділення траєкторій.
3. Кластеризація траєкторій – кластеризація вилучених траєкторій за допомогою алгоритму кластеризації, який використовує форму, розташування, напрямок і щільність сусідства шаблонів траєкторій.
4. Сегментація потоку натовпу – позначення окремих пікселів залежно від їхньої належності до певного патерну потоку для фінальної сегментації потоків натовпу.

Таким чином, сформульовано **другий пункт наукової новизни:** *удосконалено* метод сегментації потоків натовпу з високою щільністю на основі кластеризації траєкторій з використанням методу активних контурів. Використання запропонованого методу дозволяє ефективно виділяти локальні

патерни колективного руху та підвищити точність сегментації потоків натовпу у сценах з високою щільністю натовпу.

Для забезпечення більш точної та стійкої сегментації потоків натовпу в процесі сегментації потоків натовпу була запропонована модель представлення траєкторії руху в натовпі, яка включає багатовимірні ознаки – форму, положення, напрямок і багатомасштабну щільність.

Модель представлення траєкторії руху в натовпі описується кортежем:

$$F_j = (fs, location, fd, fu)$$

де fs – форма траєкторії (поліном третього порядку);

$location$ – середнє розташування траєкторії;

fd – напрямок руху;

fu – щільність сусідніх траєкторій на трьох масштабах.

Отриманий кортеж ознак далі використовується для кластеризації траєкторій.

Таким чином, сформульовано **третій пункт наукової новизни**: *набула подальшого розвитку* модель представлення траєкторії руху в натовпі, що враховує комплексні ознаки руху. Використання даної моделі забезпечує більш точну і стійку сегментацію потоків натовпу.

Для підвищення точності виявлення аномальних ситуацій у сценах з натовпом високої щільності був запропонований метод виявлення аномальних сцен натовпу за допомогою оцінки ентропії спрямованих міні-треків, побудованих з просторово-часових точок інтересу, а також аналізу часового відхилення заповнюваності, щоб виявити різкі зміни в динаміці сцени.

Використання запропонованого методу забезпечує підвищену чутливість до нетипової поведінки в складних сценах спостереження та, таким чином, сприяє підвищенню точності виявлення аномальних ситуацій.

Робота запропонованого методу складається з наступних кроків:

1. Сегментація області активного контуру – відокремлення області переднього плану від усієї відеосцени.

2. Генерація міні-треків – виявлення просторово-часових блоків точок інтересу та відстеження їх на кадрах.

3. Обчислення напрямку потоку – обчислення напрямків руху міні-треків та генерація гістограми орієнтованих міні-треків.

4. Виявлення аномальних сцен – обчислення ентропії та часової зайнятості для орієнтованих міні-треків.

Таким чином, сформульовано **четвертий пункт наукової новизни:** *удосконалено* метод виявлення аномальних ситуацій у сценах з натовпом високої щільності, що ґрунтується на аналізі ентропії напрямлених міні-треків, сформованих із просторово-часових точок інтересу та оцінки часового відхилення заповнюваності сцени. Використання запропонованого методу дозволяє збільшити чутливість систем візуального аналізу натовпу до нетипової поведінки людей в складних сценах, що дозволяє підвищити точність розпізнавання аномальних подій в сценах з високою щільністю натовпу.

В **четвертому розділі** дисертаційної роботи розроблено підсистему візуального аналізу натовпу у складі системи інтелектуального відеоспостереження.

Сформульовано мету, задачі та визначено роль цієї підсистеми у контексті забезпечення громадської безпеки, оперативного виявлення та реагування на критичні ситуації в місцях масового скупчення людей. Побудовано модель інформаційних потоків, що деталізує вхідні та вихідні дані, а також взаємодію з іншими компонентами системи і зовнішніми користувачами.

Визначено чіткий перелік функціональних і нефункціональних вимог, які включають аспекти продуктивності, безпеки, надійності та сумісності. Розроблено структурні моделі, такі як діаграми логічного уявлення, розгортання, пакетів і класів, що дозволило формалізувати архітектуру підсистеми. Проведено комплексне функціональне тестування підсистеми, за результатами якого підтверджено коректність її роботи та готовність до практичного використання.

В п'ятому розділі дисертаційної роботи проведено дослідження компонентів підсистеми візуального аналізу натовпу в складі системи інтелектуального відеоспостереження.

В процесі дослідження методу сегментації потоків натовпу в сценах з високою щільністю натовпу, було проведено тестування на чотирьох наборах даних: ShanghaiTech-B, WorldExpo, UCF_CC_50 та Mall. На більшості з них запропонований метод показав покращення. Зокрема, на наборі ShanghaiTech-B MAE зменшився з 8,9 до 8,1 (покращення на 8,99%), на WorldExpo – з 8,86 до 7,84 (на 11,54%), а на UCF_CC_50 – з 271,8 до 253,5 (на 6,73%). Лише на наборі Mall модель трохи поступалася TAFNet: MAE становив 2,10 проти 2,06 (зниження точності на 1,94%). Результати підтверджують ефективність урахування глибини сцени для покращення точності підрахунку в умовах складного фону та варіативної щільності.

В процесі дослідження методу сегментації потоків натовпу в сценах з високою щільністю натовпу було вказано, що метод базується на поєднанні сегментації переднього плану та кластеризації траєкторій, що дозволяє виділяти потоки людей у складних динамічних умовах.

Проведено емпіричний підбір параметрів на основі ROC-кривих. На наборі UCF CSD запропонований підхід досяг середнього значення міри Жаккара 0,86, що перевищує результати аналогів на 10,26 – 50,88%, а також F-міри до 1,0. На складнішому наборі Collective Motion отримано MAE = 0,72 (на 15,29 – 59,78% краще), міру Жаккара 0,83 і F-міру 0,86 (покращення на 7,50–120,51%). Результати підтверджують ефективність методу в задачах сегментації руху та виділення групової динаміки в натовпі.

В процесі дослідження методу виявлення аномалій у сценах з високою щільністю натовпу було проведено тестування на наборах UMN, UCF CSD та Violent Flows Crowd, яке показало стабільне покращення результатів. Зокрема, на UMN отримано AUC = 0,995 (+1,53%), на UCF CSD – 0,850 (+3,66%), на Violent Flows Crowd – 0,950 (+2,70%) порівняно з найкращими з існуючих методів. Таким

чином, запропонований підхід демонструє високу точність і придатність для застосування в системах інтелектуального відеоспостереження.

Таким чином, підтверджено спроможність та правильність технічних рішень, заснованих на результатах дисертаційного дослідження.

Розроблені в роботі методи та інструментальні засоби отримали впровадження у діяльності ТОВ «Провектус ІТ» та ТОВ «СФ ІНВЕСТ» та знайшли відображення у навчальному процесі та науково-дослідницькій діяльності Національного університету «Одеська політехніка».

Ключові слова: інтелектуальне відеоспостереження, візуальний аналіз натовпу, комп'ютерний зір, аналіз траєкторій, виявлення аномалій, кластеризація руху, моделювання поведінки натовпу, щільність натовпу, дескриптори траєкторій, громадська безпека.

Список публікацій здобувача за темою дисертації

1. Добришев Р. Є. Візуальний аналіз натовпу в сучасних системах інтелектуального відеоспостереження // Наука і техніка сьогодні (Серія «Техніка»), 2024, Vol. 37, № 9. Р. 639 – 651. (Index Copernicus)

<https://perspectives.pp.ua/index.php/nts/article/view/15110>

2. Dobryshev R. Y. & Maksymov M. V. “Accurate crowd counting for intelligent video surveillance systems”. Herald of Advanced Information Technology. 2024; Vol. 7, No. 3: 253–261. (Index Copernicus)

<https://hait.op.edu.ua/index.php/journal/article/view/20>

3. Dobryshev R. Y. “Anomaly detection in crowded scenes: technologies, challenges and opportunities”. Applied Aspects of Information Technology. 2024; Vol. 7, No. 3: 219–230. (Index Copernicus)

<https://aait.op.edu.ua/index.php/journal/article/view/252>

4. Добришев Р. Є. Виявлення аномалій в багатолюдних сценах // Наука і техніка сьогодні (Серія «Техніка»), 2025, Vol. 46, № 5. Р. 1437 – 1449. (Index Copernicus)

<https://perspectives.pp.ua/index.php/nts/article/view/24510>

5. Dobryshev R., Purish S., Lobachev M., Hodovychenko M. Crowd counting in intelligent video surveillance systems. *Radioelectronic and Computer Systems*, 2025, №. 2, P. 132–142. (Scopus)

<http://nti.khai.edu/ojs/index.php/reks/article/view/reks.2025.2.08>

6. Добришев Р.Є., Лобачев М.В. Підрахунок людей у натовпі з урахуванням глибини сцени в задачах оцінки щільності // Proceedings of the 15th International Scientific Conference of students and young researchers «Modern Information Technology», MIT 2025, Odesa, Ukraine. May 15-16, 2025. С. 144–146.

https://drive.google.com/file/d/1zi1TabIH7iso1-SW__9ntZ_LG0glE4Wa/view?usp=drive_link

7. Добришев Р. Є. Інтелектуальне виявлення аномалій у натовпі: методи, проблеми та перспективи // The 7th International scientific and practical conference

«Future of science: innovations and perspectives». SSPG Publish, Stockholm, Sweden, May 19 – 21, 2025. Pp. 172-176.

<https://sci-conf.com.ua/wp-content/uploads/2025/05/FUTURE-OF-SCIENCE-INNOVATIONS-AND-PERSPECTIVES-19-21.05.25.pdf>

8. Добришев Р. Є. Технології візуального аналізу натовпу в інтелектуальних системах відеоспостереження // Proceedings of the The 20th International scientific and practical conference «Innovative scientific research: latest theories, modern methods and practices». International Science Group, Seville, Spain, May 20 – 23, 2025. Pp. 387-389.

<https://isg-konf.com/wp-content/uploads/2025/05/INNOVATIVE-SCIENTIFIC-RESEARCH-LATEST-THEORIES-MODERN-METHODS-AND-PRACTICES.pdf>

9. Добришев Р. Є. Сучасні методи підрахунку людей у натовпі в системах інтелектуального відеоспостереження // Proceedings of the 2nd International Scientific and Practical Conference «Science and Information Technologies in the Modern World». International Scientific Unity, Athens, Greece, May 21 – 23, 2025. Pp. 694 –698.

https://isu-conference.com/wp-content/uploads/2025/05/Athens_Greece_21.05.25.pdf

10. Добришев Р. Є. Автоматизоване виявлення глобальних аномалій у динаміці натовпу // Proceedings of the 11th International scientific and practical conference «Scientific achievements of contemporary society». Cognum Publishing House. London, United Kingdom, May 29 – 31, 2025. Pp. 255–259.

<https://sci-conf.com.ua/wp-content/uploads/2025/05/SCIENTIFIC-ACHIEVEMENTS-OF-CONTEMPORARY-SOCIETY-29-31.05.25.pdf>

11. Добришев Р. Є. Сегментація потоків у високощільних натовпах за допомогою активного контуру та кластеризації траєкторій // Proceedings of the 1st International Scientific and Practical Conference «Modern Perspectives on Science and Economic Progress». International Scientific Unity, Vilnius, Lithuania, June 4 – 6 , 2025. Pp. 510 – 513.

https://isu-conference.com/wp-content/uploads/2025/06/Vilnius_Lithuania_4.06.25.pdf

ABSTRACT

Dobryshev R. Ye. Models and methods of visual crowd analysis in intelligent video surveillance systems. – Qualification scientific work in the form of manuscript.

Thesis for the PhD degree in specialty 122 Computer science. – Odessa Polytechnic National University, Ministry of Education and Science of Ukraine, Odesa, 2025.

The **introduction** emphasizes the importance of addressing the challenges faced in applying visual crowd analysis in existing intelligent video surveillance. It demonstrates how current theories and methods are inadequate, particularly when there are many people in a small space. The study object and subject, goals, objectives, methods are formulated explicitly. The section also addresses scientific contribution, practical applicability of results as well as the author's own contribution.

Chapter one of the thesis presents in detail the issues that arise working with visually analysis of crowds in a smart surveillance environment.

The chapter discusses the components and the roles of intelligent video surveillance systems, defines the fundamental problem of crowd analysis and provides an overview of its most prominent tasks. We scrutinize each of these sub-tasks in their respective academia in order to discover what the ongoing research challenges are, what problems have yet to be solved, and what the status of the art is in the discipline. The section also includes interesting topics that future works on the topic of visual crowd analytics should address.

The study found that the majority of modern crowd counting methods do not factor in the geometric structure of the scene being observed, for example perspective distortion and variations in size. These conditions can allow for density estimates that are significantly in error.

The remedy to this problem according to the research, is to incorporate depth maps, which are geometric scene features, in the deep structure of the CNNs. This approach allows you to adjust the size of features, which improve crowd density estimation.

To mitigate this, a new manner for the discovery of the intricate crowd scene in unusual scenes was developed using the entropy measurement of directed mini-tracklets

that are composed of spatio-temporal interest points. This method also removes short and noisy trajectories to ensure the directionality, and utilizes entropy analysis as well as the variance of temporal occupancy to detect sudden change in crowd behaviors. This analysis illuminated what the research questions were.

Chapter two of the dissertation discusses a method in which geometric information inside the scene is utilized in order to improve crowd density estimates. The approach leverages depth as an additional input in order to build scale-aware representations, which better represent the actual size of the objects leading to more plausible density predictions.

A special module, that processes and encodes the depth information, provides the scene depth to the neural network. This module enhances feature representations by altering some parts of the image according to the local depth values. It makes the scene feel more as it is.

The key components of the proposed method include:

1. The depth injection module fuses the feature map and the depth picture to bring spatial geometry to the feature space.
2. The depth refinement block makes the depth map more complete and learns scale-sensitive weights through a spatial attention mechanism, which reveals where objects exist in different regions of picture.
3. The weights are then used by the embedding unit to alter the original features in some locations. The new features consider the variation of size of foreground objects with depth, enabling the density estimation to be more robust and scale-aware.

It is clear that this constitutes the **first contribution to scientific innovation**: a methodology for incorporating geometric scene features, in the form of depth maps, into deep neural network architectures has been refined. This enhancement now allows accurate display of the size of items in space and contributes to reliable prediction about crowd density.

Chapter Three. This work and developed methods and computational tools are presented here in Chapter 3 with the focus on analysis and interpretation of high-density crowd scenarios to detect anomalous behavior.

For the purpose of improving the accuracy of the motion flows' segmentation in the dense scene, vertical crowded flows were segmented based on the clustering of motion trajectories by deformable contour model was proposed. This approach further facilitates isolating local motion patterns and increasing the interpretability of challenging dynamic scenes.

The proposed approach's implementation proceeds in the following steps:

1. Contour-based region segmentation – separation of the region covered by the crowd from the context of the rest of the visual stream using flexible boundary models.
2. Pathway identification – finding motion anchors in a grid structure and following the anchors across successive frames to form movement paths.
3. Trajectory grouping – grouping formed motion structures based on its characteristics such as the motion path or the coordinates of the center or also the motion vector as well as the neighborhood density.
4. Flow pattern mapping – pixel-wise labeling according to the proximity to a given motion cluster in order to create the final segmentation of movement streams.

According to this, the **second major scientific contribution** is formulated: a method to refine a method of segmenting densely crowded crowd motion using trajectory aggregation with the help of deformable boundary modeling. This approach can accurately locate local behavior trend patterns and enhances the segmentation performance on crowd scenes.

To enhance the robustness and the coherence of crowd flow analysis, a model for individual flow paths of crowds was proposed. This model combines several descriptors such as the curvature of the path, its spatial center, direction vector and multi-scale spatial concentration of neighboring trajectories as follows.

The trajectory is represented by the following set of attributes:

$$F_j = (f_s, \text{location}, f_d, f_u)$$

where:

- f_s denotes form of the movement curve structure (third-polynomial model);

- location represents the midpoint of the motion path;
- fd indicates the current direction of movement;
- fu captures the local density distribution of neighboring trajectories over a range of spatial scales.

The cluster process uses more clusters from each body part, but it is tricky- how to define what movement pattern is "similar" to what.

Thus, **third scientific novelty**: A three-dimensional representation of trajectory based on movement within a crowd scene has been created and updated. The use of this enhanced representation enables a more accurate and robust motion flow demarcation.

An advanced approach was proposed to identify the anomalous behavioral patterns in crowd scenes with high crowd density based on entropy analysis of directionally coherent short-term motion segments which are learned from temporal and spatial salient points. Furthermore, it integrates time-based occupancy deviation to consider sudden changes that also appear in scene dynamics.

The proposed approach improves remarkably the sensitivity to on-line unusual activities in crowded surroundings and the precision in detecting anomalous behavioral events.

The fundamental steps of this technique can be summarized as follows:

1. Foreground region isolation - detecting and enclosing the dynamic activity area by means of an elastic contour shaping.
2. Mini-path synthesis – recognizing important spatiotemporal blocks and follow their evolution constructing short, directed movement segments.
3. Motion orientation profiling – computes the principal direction of mini-trajectories and constructs directional distribution profile.
4. Irregularity detection – computing motion entropy as well as checking for spatial occupancy temporal displacement to signal such an abnormal scene state.

Thus, the **fourth scientific novelty** is the improvement of anomaly detection mechanism in dense crowd conditions. This is achieved by performing entropy based analysis over directional mini-trajectories built using motion sensitive points of interest, while monitoring the temporal variation of the occupancy. The introduction of this

framework results in increased sensitivity of the crowd behavior anomalies and better reliability for detecting abnormal events in visually crowded environments.

Chapter four of the thesis describes the engineering and deployment of a visual analytics module for crowd monitoring within a generic intelligent surveillance environment.

The overall goal, specific aims, and strategic significance of this module were set in the context of preserving civic security and detecting and mitigating emergencies in public heavily visited spaces. A data flow diagram was designed, in which were defined the semantic structure of incoming and outgoing data and the protocols of communication with edge components and external operators.

A detailed statement of operational and non-operational requirements was provided that covered performance requirements, cybersecurity, system reliability, and interoperability. The subsystem's design was formally documented with a variety of architectural diagrams such as logic diagrams, deployment diagrams, or package and class diagrams. Significant validation tests were performed to ensure functional integrity of the design indicating that the system is ready for deployment and its performance is stable under operational loads.

A focused analysis on the constituting elements Propagation, Aggregation and Global and Local Crowd Description of the visual crowd analysis system are described in **Chapter Five**.

In the process of studying the method of crowd flow segmentation in scenes with high crowd density, we tested it on four datasets: ShanghaiTech-B, WorldExpo, UCF_CC_50, and Mall. On most of them, the proposed method showed improvements. In particular, on the ShanghaiTech-B dataset, MAE decreased from 8.9 to 8.1 (an improvement of 8.99%), on WorldExpo - from 8.86 to 7.84 (11.54%), and on UCF_CC_50 - from 271.8 to 253.5 (6.73%). Only on the Mall set was the model slightly inferior to TAFNet: MAE was 2.10 vs. 2.06 (a decrease in accuracy of 1.94%). The results confirm the effectiveness of taking into account the depth of the scene to improve the counting accuracy in conditions of complex backgrounds and variable densities.

In the process of studying the method of segmenting crowd flows in scenes with high crowd density, it was indicated that the method is based on a combination of foreground segmentation and trajectory clustering, which allows to distinguish flows of people in complex dynamic conditions.

The empirical selection of parameters based on ROC curves was carried out. On the UCF CSD dataset, the proposed approach achieved an average Jaccard measure of 0.86, which exceeds the results of analogs by 10.26-50.88%, as well as an F-measure of up to 1.0. On the more complex Collective Motion set, we obtained MAE = 0.72 (15.29 - 59.78% better), Jaccard measure 0.83, and F-measure 0.86 (7.50-120.51% improvement). The results confirm the effectiveness of the method in the tasks of motion segmentation and group dynamics detection in a crowd.

In the process of studying the method for detecting anomalies in scenes with high crowd density, testing was conducted on the UMN, UCF CSD and Violent Flows Crowd datasets, which showed a steady improvement in results. In particular, on UMN, we obtained AUC = 0.995 (+1.53%), on UCF CSD - 0.850 (+3.66%), on Violent Flows Crowd - 0.950 (+2.70%) compared to the best existing methods. Thus, the proposed approach demonstrates high accuracy and suitability for use in intelligent video surveillance systems.

Thus, the feasibility and correctness of the technical solutions based on the results of the dissertation research have been confirmed.

The methods and tools developed in this work have been implemented in the activities of Provectus IT LLC and SF INVEST LLC and are reflected in the educational process and research activities of the Odessa National Polytechnic University.

Keywords: vision architectures and models for robotics, probabilistic and statistical systems, mobile robotic systems, smart camera networks, dynamic scene analysis, surveillance, localization, tracking, multi-target tracking, motion and tracking segmentation.

List of publications of the applicant on the topic of the dissertation

1. Добришев Р. Є. Візуальний аналіз натовпу в сучасних системах інтелектуального відеоспостереження // Наука і техніка сьогодні (Серія «Техніка»), 2024, Vol. 37, № 9. Р. 639 – 651. (Index Copernicus)

<https://perspectives.pp.ua/index.php/nts/article/view/15110>

2. Dobryshev R. Y. & Maksymov M. V. “Accurate crowd counting for intelligent video surveillance systems”. Herald of Advanced Information Technology. 2024; Vol. 7, No. 3: 253–261. (Index Copernicus)

<https://hait.op.edu.ua/index.php/journal/article/view/20>

3. Dobryshev R. Y. “Anomaly detection in crowded scenes: technologies, challenges and opportunities”. Applied Aspects of Information Technology. 2024; Vol. 7, No. 3: 219–230. (Index Copernicus)

<https://aait.op.edu.ua/index.php/journal/article/view/252>

4. Добришев Р. Є. Виявлення аномалій в багатолюдних сценах // Наука і техніка сьогодні (Серія «Техніка»), 2025, Vol. 46, № 5. Р. 1437 – 1449. (Index Copernicus)

<https://perspectives.pp.ua/index.php/nts/article/view/24510>

5. Dobryshev R., Purish S., Lobachev M., Hodovychenko M. Crowd counting in intelligent video surveillance systems. *Radioelectronic and Computer Systems*, 2025, №. 2, P. 132–142. (Scopus)

<http://nti.khai.edu/ojs/index.php/reks/article/view/reks.2025.2.08>

6. Добришев Р.Є., Лобачев М.В. Підрахунок людей у натовпі з урахуванням глибини сцени в задачах оцінки щільності // Proceedings of the 15th International Scientific Conference of students and young researchers «Modern Information Technology», MIT 2025, Odesa, Ukraine. May 15-16, 2025. С. 144–146.

https://drive.google.com/file/d/1zi1TabIH7iso1-SW__9ntZ_LG0glE4Wa/view?usp=drive_link

7. Добришев Р. Є. Інтелектуальне виявлення аномалій у натовпі: методи, проблеми та перспективи // The 7th International scientific and practical conference

«Future of science: innovations and perspectives». SSPG Publish, Stockholm, Sweden, May 19 – 21, 2025. Pp. 172-176.

<https://sci-conf.com.ua/wp-content/uploads/2025/05/FUTURE-OF-SCIENCE-INNOVATIONS-AND-PERSPECTIVES-19-21.05.25.pdf>

8. Добришев Р. Є. Технології візуального аналізу натовпу в інтелектуальних системах відеоспостереження // Proceedings of the The 20th International scientific and practical conference «Innovative scientific research: latest theories, modern methods and practices». International Science Group, Seville, Spain, May 20 – 23, 2025. Pp. 387-389.

<https://isg-konf.com/wp-content/uploads/2025/05/INNOVATIVE-SCIENTIFIC-RESEARCH-LATEST-THEORIES-MODERN-METHODS-AND-PRACTICES.pdf>

9. Добришев Р. Є. Сучасні методи підрахунку людей у натовпі в системах інтелектуального відеоспостереження // Proceedings of the 2nd International Scientific and Practical Conference «Science and Information Technologies in the Modern World». International Scientific Unity, Athens, Greece, May 21 – 23, 2025. Pp. 694 –698.

https://isu-conference.com/wp-content/uploads/2025/05/Athens_Greece_21.05.25.pdf

10. Добришев Р. Є. Автоматизоване виявлення глобальних аномалій у динаміці натовпу // Proceedings of the 11th International scientific and practical conference «Scientific achievements of contemporary society». Cognum Publishing House. London, United Kingdom, May 29 – 31, 2025. Pp. 255–259.

<https://sci-conf.com.ua/wp-content/uploads/2025/05/SCIENTIFIC-ACHIEVEMENTS-OF-CONTEMPORARY-SOCIETY-29-31.05.25.pdf>

11. Добришев Р. Є. Сегментація потоків у високощільних натовпах за допомогою активного контуру та кластеризації траєкторій // Proceedings of the 1st International Scientific and Practical Conference «Modern Perspectives on Science and Economic Progress». International Scientific Unity, Vilnius, Lithuania, June 4 – 6 , 2025. Pp. 510 – 513.

https://isu-conference.com/wp-content/uploads/2025/06/Vilnius_Lithuania_4.06.25.pdf

ЗМІСТ

Вступ.....	26
1 Аналіз задачі візуального аналізу натовпу в системах відеоспостереження.....	31
1.1 Визначення інтелектуальних систем відеоспостереження	31
1.2 Визначення задачі візуального аналізу натовпу	33
1.2.1 Задача підрахунку натовпу та оцінки щільності	38
1.2.2 Задача виявлення та відстеження об'єктів	40
1.2.3 Задача аналізу руху натовпу	42
1.2.4 Задача розпізнавання поведінки, активності та контексту	44
1.2.5 Задача виявлення аномалій	46
1.2.6 Задача прогнозування натовпу	49
1.2.7 Перспективи досліджень в області візуального аналізу натовпу	51
1.3 Постановка задачі дисертаційного дослідження	54
1.3.1 Інтеграція геометричних ознак у моделі підрахунку щільності натовпу на основі глибинного навчання.....	54
1.3.2 Задача сегментації потоків натовпу у сценах з високою щільністю	54
1.3.3 Представлення траєкторій для сегментації руху у натовпі.....	55
1.3.4 Виявлення аномальної поведінки у високощільному натовпі.....	56
1.4 Висновки до першого розділу	57
2 Метод підвищення точності оцінки щільності натовпу шляхом інтеграції геометричних ознак сцени.....	60
2.1 Вступ	60
2.2 Вплив інформації про глибину сцени на підрахунок людей у натовпі	61
2.3 Загальна схема роботи методу.....	63
2.4 Отримання карти глибини	65
2.5 Модуль інтеграції інформації про глибину сцени.....	66
2.5.1 Блок впровадження глибини.....	67
2.5.2 Блок корекції глибини.....	68
2.5.3 Блок вбудовування	71
2.6 Нейромережева модель SAWNet.....	72

2.6.1 Навчання моделі SAWNet.....	73
2.6.2 Дослідження роботи блоків модулю інтеграції глибини	74
2.7 Висновки до другого розділу.....	77
3 Моделі та методи обробки та аналізу сцен з високою щільністю натовпу для виявлення аномалій.....	79
3.1 Метод сегментації потоків натовпу в сценах з високою щільністю натовпу ...	79
3.1.1 Загальний принцип роботи запропонованого методу	80
3.1.1.1 Сегментація на основі активних контурів.....	81
3.1.1.2 Обчислення траєкторій	83
3.1.1.3 Модель представлення траєкторій руху натовпу	87
3.1.1.4 Кластеризація траєкторій.....	89
3.1.1.5 Сегментація потоків натовпу людей	92
3.2 Метод виявлення аномалій у сценах з високою щільністю натовпу.....	95
3.2.1 Загальна ідея роботи запропонованого методу	97
3.2.1.1 Сегментація переднього плану	98
3.2.1.2 Генерація міні-треків.....	99
3.2.1.3 Визначення напрямку руху міні-треків.....	99
3.2.1.4 Побудова гістограми орієнтованих міні-треків	100
3.2.2 Виявлення аномальних сцен.....	101
3.2.2.1 Ентропія орієнтованих міні-треків	103
3.2.2.2 Класифікація сцени на базі ентропії.....	104
3.2.2.3 Класифікація сцени на основі показника часової зайнятості	105
3.2.2.4 Вибір порогу прийняття рішення	106
3.3 Висновки до третього розділу	107
4 Реалізація підсистеми візуального аналізу натовпу в складі системи інтелектуального відеоспостереження.....	110
4.1 Мета та задачі підсистеми візуального аналізу натовпу.....	110
4.2 Визначення потоків інформації підсистеми візуального аналізу натовпу.....	110
4.3 Визначення функціональних вимог підсистеми візуального аналізу натовпу.....	113

4.4 Визначення нефункціональних вимог до підсистеми візуального аналізу натовпу.....	118
4.5 Проектування діаграми логічного уявлення підсистеми візуального аналізу натовпу.....	120
4.6 Проектування розгортання підсистеми візуального аналізу натовпу	121
4.7 Визначення структури програмного проєкту підсистеми візуального аналізу натовпу.....	122
4.8 Розробка діаграми класів підсистеми візуального аналізу натовпу	123
4.9 Управління вихідним кодом підсистеми візуального аналізу натовпу.....	125
4.10 Аналіз вихідного коду підсистеми візуального аналізу натовпу	126
4.11 Функціональне тестування підсистеми візуального аналізу натовпу	127
4.12 Висновки до четвертого розділу.....	135
5 Дослідження підсистеми візуального аналізу натовпу в складі системи інтелектуального відеоспостереження.....	136
5.1 Дослідження методу сегментації потоків натовпу в сценах з високою щільністю натовпу	136
5.1.1 Тестове оточення.....	136
5.1.2 Використані набори даних.....	137
5.1.3 Порівняння запропонованого методу з аналогами.....	139
5.1.4 Обговорення результатів дослідження	143
5.2 Дослідження методу сегментації потоків натовпу в сценах з високою щільністю натовпу	144
5.2.1 Дослідження значення параметрів методу.....	144
5.2.2 Використані метрики оцінки роботи методу	145
5.2.3 Використані набори даних.....	147
5.2.4 Порівняння запропонованого методу з аналогами.....	147
5.2.4.1 Дослідження на наборі даних UCF Crowd Segmentation dataset	147
5.2.4.2 Дослідження на наборі даних Collective Motion.....	149
5.2.5 Обговорення результатів дослідження	150

5.3 Дослідження методу виявлення аномалій у сценах з високою щільністю натовпу.....	151
5.3.1 Налаштування параметрів методу	152
5.3.2 Використані метрики оцінки ефективності методу.....	152
5.3.3 Використані набори даних.....	153
5.3.4 Порівняння запропонованого методу з аналогами.....	154
5.3.4.1 Оцінка роботи методу на наборі UMN	154
5.3.4.2 Оцінка роботи методу на наборі UCF Crowd Segmentation dataset.....	156
5.3.4.3 Оцінка роботи методу на наборі Violent Flows Crowd	157
5.3.5 Обговорення результатів дослідження	159
5.4 Висновки до п'ятого розділу	160
Висновки.....	162
Список використаних джерел.....	165
Додаток А Список публікацій здобувача за темою дисертації	182
Додаток Б Акт впровадження результатів у ТОВ «Провектус ІТ»	185
додаток В Акт впровадження результатів у ТОВ «СФ ІНВЕСТ».....	187
Додаток Г Довідка про використання результатів дисертаційної роботи у навчальному процесі Національного університету «Одеська Політехніка»	188
Додаток Д Довідка про використання результатів дисертаційної роботи у науково-дослідницькій діяльності Національного університету «Одеська Політехніка»	189
Додаток Е Вихідний код програмних інструментальних засобів	190

ВСТУП

У XXI столітті забезпечення громадської безпеки в умовах зростаючої урбанізації, інтенсивного розвитку транспортної інфраструктури та постійного зростання чисельності населення у великих містах постає як одне з ключових завдань державних та муніципальних органів. Значна частина небезпек, що виникають у публічних просторах, пов'язана з масовим скупченням людей, де навіть незначні інциденти можуть швидко призводити до паніки, тисняви або порушення правопорядку. У цьому контексті візуальний аналіз натовпу, як складова систем інтелектуального відеоспостереження, набуває стратегічного значення. Застосування інструментів візуального аналізу натовпу дозволяє виявляти аномалії поведінки, визначати небезпечні тенденції в русі натовпу, оперативно реагувати на кризові ситуації, запобігати конфліктам або загрозам, пов'язаним із терористичними актами, раптовими техногенними аваріями чи порушеннями громадського порядку.

Особливу актуальність ця задача має для критично важливих об'єктів: аеропортів, вокзалів, стадіонів, торговельних центрів, площ, маршрутів евакуації тощо. У таких середовищах ручний моніторинг поведінки людей є обмеженим через великий обсяг інформації та складність обробки відеопотоків у режимі реального часу. Саме тому виникає потреба в автоматизованих, адаптивних, інтелектуальних технологіях, здатних аналізувати динаміку натовпу, моделювати поведінкові патерни, розпізнавати потенційно небезпечні ситуації та формувати своєчасні рекомендації для операторів систем відеоспостереження чи екстрених служб. Такі системи мають бути не лише функціональними, але й масштабованими, стійкими до шумів, здатними до самонавчання та адаптації до нових сценаріїв, що лише підсилює потребу у нових моделях та методах візуального аналізу.

Разом із тим, візуальний аналіз натовпу в умовах високої щільності пов'язаний з низкою серйозних технічних і методологічних перешкод. Зокрема, істотними викликами є сильне перекриття об'єктів, низька роздільна здатність окремих фрагментів зображення, неоднорідність освітлення, постійні зміни

масштабу та кута огляду, а також складність диференціації індивідуальних траєкторій. У таких обставинах традиційні методи комп'ютерного зору та машинного навчання виявляються недостатньо стійкими та чутливими до поведінкових патернів, що виникають у складних просторово-часових контекстах.

Таким чином, існує протиріччя між запитом ринку на впровадження систем візуального аналізу натовпу, з однієї сторони, та недосконалістю існуючих моделей та методів візуального аналізу натовпу, з іншої сторони.

Тому актуальними є дослідження, що направлені на усунення цього протиріччя та присвячені вирішенню важливої науково-практичної задачі підвищення ефективності систем візуального аналізу натовпу.

Метою дослідження є підвищення ефективності систем візуального аналізу натовпу шляхом удосконалення моделей та методів оцінки щільності натовпу за рахунок використання інформації про глибину сцени та виявлення аномальних ситуацій у сценах з натовпом високої щільності за рахунок оцінки ентропії напрямлених міні-треків та часового відхилення заповнюваності сцени.

Для досягнення мети дослідження необхідно вирішити наступні завдання:

- провести аналіз основних проблем, обмежень та викликів, які виникають в процесі побудови систем візуального аналізу натовпу та намітити шляхи їх вирішення;
- удосконалити метод інтеграції інформації про геометричні ознаки сцени у глибинні нейронні мережі, які здійснюють оцінку щільності натовпу;
- удосконалити метод сегментації потоків натовпу з високою щільністю на основі кластеризації траєкторій з використанням методу активних контурів;
- удосконалити модель представлення траєкторії руху в натовпі, що враховує комплексні ознаки руху;
- удосконалити метод виявлення аномальних ситуацій у сценах з натовпом високої щільності;
- розробити інструментальні засоби, що реалізують запропоновані рішення;

– провести експериментальне дослідження отриманих теоретичних результатів.

Об'єктом дослідження є процеси візуального аналізу натовпу в системах інтелектуального відеоспостереження, що здійснюють автоматизовану обробку відеоданих з метою моніторингу поведінки людей у публічних просторах.

Предметом дослідження є моделі та методи візуального аналізу натовпу, зокрема алгоритми виявлення, класифікації та інтерпретації групової поведінки людей в умовах високої щільності, динамічної сцени та обмежених візуальних спостережень.

Методи дослідження. Для вирішення поставлених задач були використані методи математичної статистики для оцінки точності моделей, оптимізаційні підходи для налаштування параметрів, а також методи глибинного навчання, зокрема згорткові нейронні мережі та механізм просторової самоуваги для аналізу контексту у відеокадрах. Використано сегментацію сцен для виділення зон активності, кластеризацію траєкторій для виявлення поведінкових патернів, а також математичне моделювання для опису динаміки й взаємодій у натовпі.

Наукова новизна отриманих результатів. Основний науковий результат полягає у розв'язанні важливої науково-практичної задачі підвищення ефективності систем візуального аналізу натовпу шляхом розробки моделей та методів оцінки щільності натовпу та виявлення аномалій у сценах з натовпом.

У рамках виконаних досліджень отримано такі наукові результати:

– *удосконалено* метод інтеграції інформації про геометричні ознаки сцени у глибинні нейронні мережі, які здійснюють оцінку щільності натовпу. Використання запропонованого методу дозволяє краще відображати реальні розміри об'єктів у просторі та підвищити достовірність оцінки щільності;

– *удосконалено* метод сегментації потоків натовпу з високою щільністю на основі кластеризації траєкторій з використанням методу активних контурів. Використання запропонованого методу дозволяє ефективно виділяти локальні патерни колективного руху та підвищити точність сегментації потоків натовпу у сценах з високою щільністю натовпу;

– *набула подальшого розвитку* модель представлення траєкторії руху в натовпі, що враховує комплексні ознаки руху. Використання даної моделі забезпечує більш точну і стійку сегментацію потоків натовпу;

– *удосконалено* метод виявлення аномальних ситуацій у сценах з натовпом високої щільності, що ґрунтується на аналізі ентропії напрямлених міні-треків, сформованих із просторово-часових точок інтересу та оцінки часового відхилення заповнюваності сцени. Використання запропонованого методу дозволяє збільшити чутливість систем візуального аналізу натовпу до нетипової поведінки людей в складних сценах, що дозволяє підвищити точність розпізнавання аномальних подій в сценах з високою щільністю натовпу.

Практичне значення одержаних результатів. Результати дисертаційного дослідження мають прикладне значення для підвищення ефективності систем інтелектуального відеоспостереження в умовах високої щільності натовпу. Розроблені моделі та методи оцінки щільності, представлення траєкторій руху та виявлення аномальної поведінки натовпу дозволяють забезпечити точніше і своєчасніше реагування на потенційно критичні ситуації у місцях масового скупчення людей.

Запропоновані моделі та методи були інтегровані у прототип підсистеми візуального аналізу, що дозволяє, в подальшому, адаптувати їх для застосування в існуючих відеоаналітичних платформах. Практичні компоненти дослідження можуть бути використані при побудові модулів безпеки для транспортних вузлів, публічних заходів, а також у смарт-системах міського моніторингу.

Дисертацію виконано згідно тематичних планів НДР Національного університету «Одеська політехніка» за період 2024 – 2025 рр. в рамках держбюджетної теми № 237-177 – «Програмні системи машинного навчання та їх застосування в галузі інтелектуального аналізу даних».

Розроблені в дисертаційній роботі моделі та методи отримали впровадження у діяльності товариства з обмеженою відповідальністю «Провектус ІТ» та товариства з обмеженою відповідальністю «СФ ІНВЕСТ», а також знайшли

відображення у навчальному процесі та науково-дослідницькій діяльності Національного університету «Одеська політехніка».

Особистий внесок здобувача. Основні наукові положення, висновки і рекомендації, що містяться в дисертації та виносяться на захист, отримано здобувачем особисто в період з 2024 по 2025 рік. У статті [1] та роботах [8] внесок здобувача полягає в проведенні аналізу предметної області, видів та задач візуального аналізу натовпу, у статтях [2, 5] та роботах [6, 9] – запропоновано метод інтеграції геометричних ознак сцени, зокрема карти глибини, до архітектур глибинних нейронних мереж, у статті [3] – аналіз предметної області, наукових робіт та сучасного стану вирішення задачі виявлення аномалій у сценах з натовпом, у статті [4] та роботах [7, 8, 10, 11] – запропонований метод метод сегментації потоків натовпу, модель представлення траєкторії руху в натовпі та метод виявлення аномальних ситуацій у сценах з натовпом високої щільності

Апробація результатів дисертації. Основні положення і практичні результати дисертаційної роботи доповідалися та одержали схвалення на таких конференціях: міжнародна наукова конференція МІТ-2025 (м. Одеса, 2025 рік), а також на різних міжнародних науково-практичних онлайн-конференціях, які проходили протягом 2025 року у Швеції, Іспанії, Греції, Великобританії та Литві.

Публікації. Основні результати дисертаційної роботи викладено в 11 публікаціях з них: 5 статей у наукових фахових виданнях України з технічних наук, з них 1 стаття у фаховому виданні України, яке індексується у наукометричній базі Scopus, 6 публікацій у працях і матеріалах наукових конференцій.

Структура і обсяг роботи. Дисертація складається зі вступу, п'яти розділів, висновків, списку використаних джерел і додатків. Повний обсяг дисертації складає 201 сторінки, в тому числі: 139 сторінок основного тексту, 26 рисунків і 17 таблиць, список використаних джерел зі 135 найменувань і 6 додатків.

1 АНАЛІЗ ЗАДАЧІ ВІЗУАЛЬНОГО АНАЛІЗУ НАТОВПУ В СИСТЕМАХ ВІДЕОСПОСТЕРЕЖЕННЯ

1.1 Визначення інтелектуальних систем відеоспостереження

Інтелектуальні системи відеоспостереження (ICBC) – це складні апаратні та прикладні системи для автоматичного збору, обробки, аналізу та розуміння відео на основі комп'ютерного зору, штучного інтелекту, статистичного моделювання та глибокого навчання. Основне застосування цих систем – динамічний моніторинг просторово-часової активності в різних умовах середовища (громадського, промислового, транспортного тощо) з метою підвищення безпеки, часу реагування та оцінки ситуації [1].

На відміну від традиційних систем відеоспостереження, які просто фіксують і зберігають відео, ICBC включають аналітичні можливості, що забезпечують розпізнавання та аналіз просторової природи об'єктів, відстеження руху, виявлення подій і прийняття рішень «в реальному часі». Вони не є пасивними інструментами спостереження, навпаки, це активна система інтерпретації подій, яка дозволяє автоматизувати багато процесів управління, діагностики та прогнозування.

З технічної точки зору, ICBC складається з декількох модулів (рис. 1.1):

- модуль захоплення зображення (включає камери відеоспостереження з різною роздільною здатністю і кутом огляду);
- модуль попередньої обробки відео (фільтрація шуму, нормалізація зображення, аналіз кадрів тощо);
- аналітичні модулі (наприклад, детектори об'єктів, треки, класифікатори або моделі розпізнавання поведінки);
- модуль прийняття рішень (логічні правила, евристики, предиктивні моделі);
- інтерфейс для зв'язку з оператором терміналу або з іншими інформаційними системами.

В рамках розвитку сучасних технологій, ICBC використовують переваги використання моделей глибокого навчання (CNN, RNN, GCN, трансформатори), які навчаються на великих візуальних корпусах для вирішення складних завдань комп'ютерного зору, таких як семантична сегментація, розпізнавання дій, аналіз натовпу та виявлення аномалій. Водночас ці системи є не лише дзеркалом стану справ, але й можуть передбачати можливу небезпеку на основі певних закономірностей.

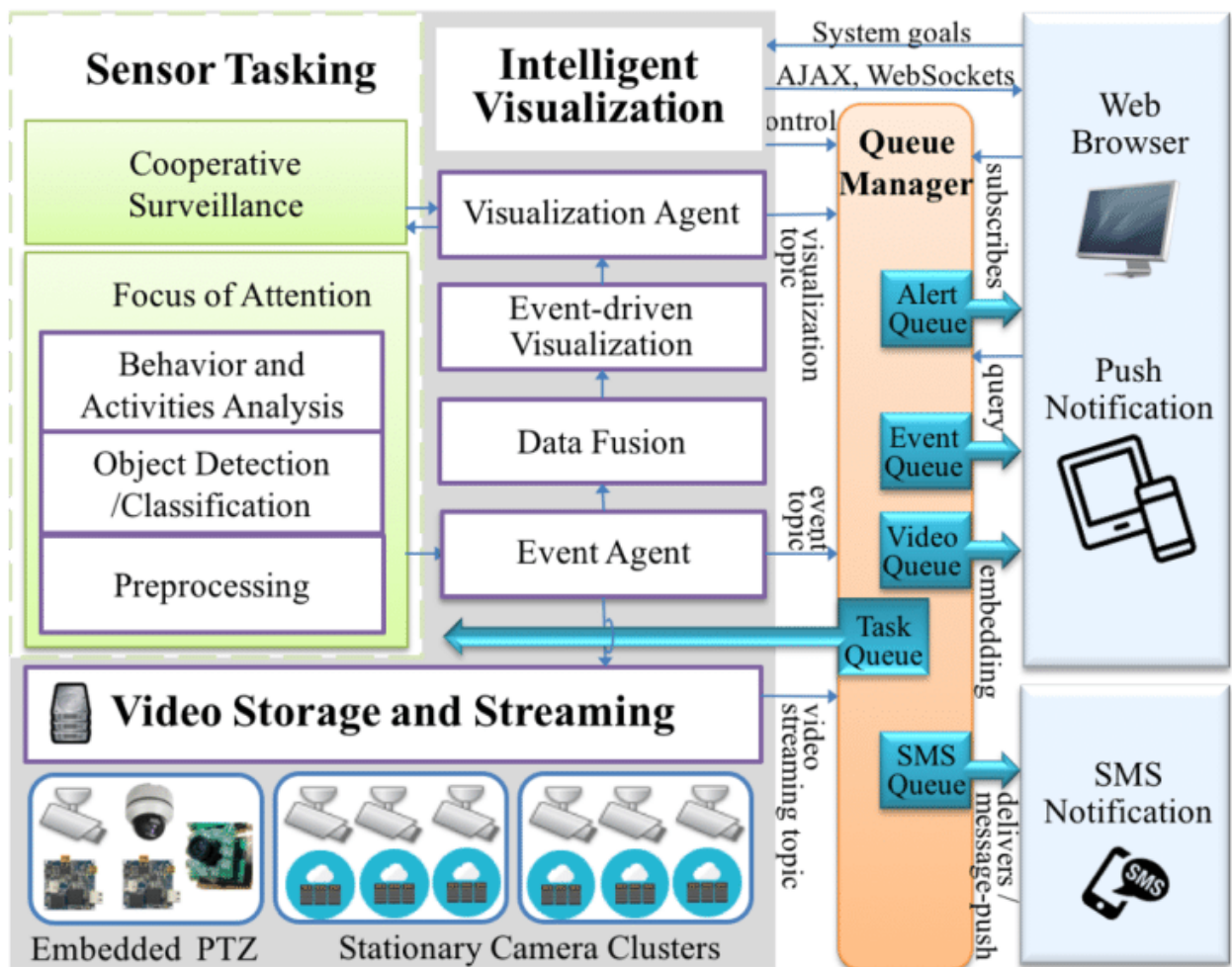


Рисунок 1.1 – Типова структура інтелектуальної системи відеоспостереження [2]

У загальному розумінні інтелектуальна система відеоспостереження – це інформаційна система з ситуаційним аналізом візуального середовища, яка повинна бути здатна самостійно генерувати оцінку ризиків, реагувати на зміни і допомагати (без безпосередньої участі людини) у прийнятті рішень.

Застосування ICBC для візуального аналізу натовпу є ключовим аспектом її використання в сфері громадської безпеки та управління масовими заходами. Оскільки відеопотік може оброблятися автоматично в режимі реального часу, ICBC дозволяють не тільки зберегти факт присутності великої кількості людей, але і вирішити такі нагальні завдання, як:

- підрахунок людей у щільному натовпі, з урахуванням часткового перекриття між людьми;
- оцінка щільності натовпу для зон потенційного ризику виникнення надзвичайних ситуацій (скупчення, тиснява);
- аналіз траєкторій руху для виявлення незвичної або аномальної поведінки чи відхилень від потоків;
- виявлення аномальної поведінки, наприклад, паніки, раптового руху, переміщення натовпу, індикаторів агресії тощо.

ICBC сприяють ранньому виявленню підозрілих загроз, щоб полегшити повідомлення відповідних служб, допомогти особам, які приймають рішення, вирішувати проблеми на основі просторових і часових моделей. Це ефективні базові технології для моніторингу та управління масами людей у чутливих або критично важливих системах, таких як транспорт, культурні заходи та громадські місця [3].

Таким чином, інтелектуальні системи відеоспостереження це не тільки технологія, але й «розумний» інструмент для прийняття рішень, який перетворює відеодані на структуровану інформацію, придатну для аналізу.

1.2 Визначення задачі візуального аналізу натовпу

Зростання міського населення є однією з причин більш частого скупчення людей у громадських місцях, що створює складну проблему забезпечення громадської безпеки. Для ефективного управління натовпом необхідно правильно і вчасно аналізувати візуальні дані. Ручний моніторинг є трудомістким і

суб'єктивним, але автоматична система відеоаналізу пропонує новий підхід до моніторингу натовпу в реальному часі [4].

За останнє десятиліття глибинне навчання досягло значного прогресу і виявилось дуже ефективним для вирішення широкого кола завдань, включаючи підрахунок кількості осіб у натовпі, виявлення об'єктів, аналіз траєкторій, прогнозування поведінки, виявлення аномалій, прогнозування динаміки натовпу і так далі.

Однак, все ще залишаються відкриті питання, особливо у випадках високої щільності натовпу, зміни масштабу, сильних оклюзій і більш складної поведінки людини в натовпі.

Проблему візуального аналізу натовпу можна розділити на такі задачі [5]:

- підрахунок кількості осіб у натовпі – оцінка кількості людей або щільності на певній сцені;
- виявлення та відстеження об'єктів – виявлення та визначення місцезнаходження людей і предметів, що становлять інтерес;
- аналіз руху – визначення напрямку, швидкості та поведінки потоку;
- визначення поведінки – визначення групової поведінки (наприклад, агресії, паніки) натовпу;
- виявлення аномалій – виявлення відхилень від стандартної поведінки (індивідуальної або групової);
- прогнозування – передбачення майбутніх змін щільності натовпу та тенденції руху натовпу.

Глибоке навчання може автоматично виявляти моделі поведінки, що відрізняються, у візуальних зразках, а навчання з перенесенням може доопрацьовувати попередньо навчені моделі з меншою потребою в перенавчанні.

Об'єднання цих методів в автоматизованому аналізі натовпу є основою сучасних інтелектуальних систем відеоспостереження, що є популярною темою досліджень, особливо в сферах громадської безпеки, транспорту та міського планування.

З формальної точки зору, відеопотік можна представити як послідовність кадрів RGB-зображення, отримана протягом інтервалу часу $t \in \{1, \dots, T\}$:

$$V = \{I_1, I_2, \dots, I_T\}, \quad I_t \in \mathbb{R}^{H \times W \times 3}, \quad (1.1)$$

де V – відеопотік як скінченна впорядкована послідовність кадрів;

$T \in \mathbb{N}$ – кількість кадрів у відеофрагменті;

$t \in \{1, \dots, T\}$ – дискретний час\індекс кадру;

I_t – t -ий кадр відео.

Нехай у спостережуваній сцені є динамічний натовп людей або об'єктів. Мета полягає у тому, щоб розмістити, зобразити та згенерувати інформацію про структуру за допомогою шести підзадач.

Задача 1 – підрахунок натовпу (оцінка регресії/щільності). Різні області інтересу та їх особливості виділяються шляхом згортки безпосередньо з вхідного зображення для того, щоб захопити різномасштабну інформацію. Позначимо через $N_t \in \mathbb{N}$ кількість людей у кадрі I_t . Пряма регресія обчислюється за наступною формулою:

$$\hat{N}_t = f_{count}(I_t; \Theta). \quad (1.2)$$

де $\hat{N}_t$ – оцінка моделі (прогноз кількості);

$f(\cdot)$ – параметризована функція підрахунку;

Θ – всі навчані параметри.

Карта щільності $D_t \in \mathbb{R}^{H \times W}$ обчислюється за наступною формулою:

$$N_t = \sum_{x=1}^H \sum_{y=1}^W D_t(x, y), \quad D_t = f_{density}(I_t). \quad (1.3)$$

де D_t – карта щільності для кадру I_t ;

N_t – оцінка кількості людей у кадрі;

$f(\cdot)$ – модель оцінки щільності.

Моделі f_{count} або $f_{density}$ часто ґрунтуються на згорткових нейронних мережах (CNN) або на основі трансформерних мереж.

Задача 2 – виявлення об'єктів та багатооб'єктне відстеження. За вхідним відео треба визначити n об'єктів у кожному кадрі і відстежувати їх у часі:

$$O_t = \left\{ \left(b_1^{(t)}, c_1 \right), \dots, \left(b_n^{(t)}, c_n \right) \right\}, \quad b_t^{(t)} \in \mathbb{R}^4, \quad (1.4)$$

де $b_i^{(t)}$ – координати обмежувальної рамки;

c_i – мітка класу.

Відстеження базується на часовому співвідношенні:

$$\tau_i = \left\{ b_i^{(1)}, b_i^{(2)}, \dots, b_i^{(T)} \right\}. \quad (1.5)$$

де τ_i – траєкторія i -го об'єкта у часі;

k – ідентифікатор треку, який зберігається, поки об'єкт видимий.

Задача 3 – оцінка потоку руху. Аналіз просторового поля руху між двома послідовними кадрами:

$$M_t = f_{flow}(I_{t-1}, I_t) \in \mathbb{R}^{H \times W \times 2}. \quad (1.6)$$

де M_t – щільне поле векторів руху між кадрами I_{t-1} та I_t ;

$f(\cdot)$ – прямий потік від I_{t-1} до I_t .

Кожен вектор $M_t(x, y) = (u, v)$ кодує піксельне зміщення. У якості альтернативи можна обчислити глобальну статистику:

$$v_t = \frac{1}{HW} \sum_{x,y} \|M_t(x,y)\|, \quad \theta = \text{домінуючий напрямок } M_t. \quad (1.7)$$

де v_t – середня швидкість руху по кадру;

H та W – висота та ширина кадру.

Задача 4 – розпізнавання поведінки. З послідовності кадрів $\{I_{t-k}, \dots, I_t\}$ зробити висновок про поведінковий стан натовпу:

$$B_t = f_{behavior}(I_{t-k}, \dots, I_t) \in C. \quad (1.8)$$

де C – скінченна множина колективних станів, таких як спокій, агресія, паніка, збудження тощо, які зазвичай моделюються за допомогою рекурентних або просторово-часових графових мереж.

Задача 5 – виявлення аномалій. Виявлення незвичайних тенденцій у русі або структурі натовпу:

$$A_t = f_{anomaly}(I_{t-k}, \dots, I_t) \in \{0,1\}. \quad (1.9)$$

де A_t – ознака аномалії в кадрі;

$f(\cdot)$ – модель виявлення аномалій.

Оцінка аномалії обчислюється за формулою:

$$s_t = \|I_t - \hat{I}_t\|^2, \quad \hat{I}_t = f_{recon}(I_{t-k}, \dots, I_{t-1}). \quad (1.10)$$

де s_t – помилка реконструкції;

$\hat{I}_t$ – реконструкція кадру;

$f(\cdot)$ – модель реконструкції кадру.

Застосування порогового значення $s_t > \delta$ призводить до появи флагу аномалії.

Задача 6 – прогнозування натовпу. За заданою послідовністю $\{I_{t-k}, \dots, I_t\}$ передбачити майбутній стан натовпу $\hat{S}_{t+\Delta}$ (наприклад, кількість, рух, щільність):

$$\hat{D}_{t+\Delta} = f_{pred}(I_{t-k}, \dots, I_t), \quad \hat{N}_{t+\Delta} = \sum_{x,y} D_{t+\Delta}(x, y). \quad (1.11)$$

Прогнозування може використовувати просторово-часові моделі кодера-декодера або предиктори на основі уваги.

1.2.1 Задача підрахунку натовпу та оцінки щільності

Підрахунок натовпу та оцінка щільності є двома важливими завданнями в автоматичному моніторингу натовпу. Такі завдання шукають кількість та/або розподіл людей на сцені. Хоча кількість людей є популярною метрикою, карта щільності робить більш локалізовані висновки, що є корисним для закритих і перевантажених сцен (рис. 1.2).

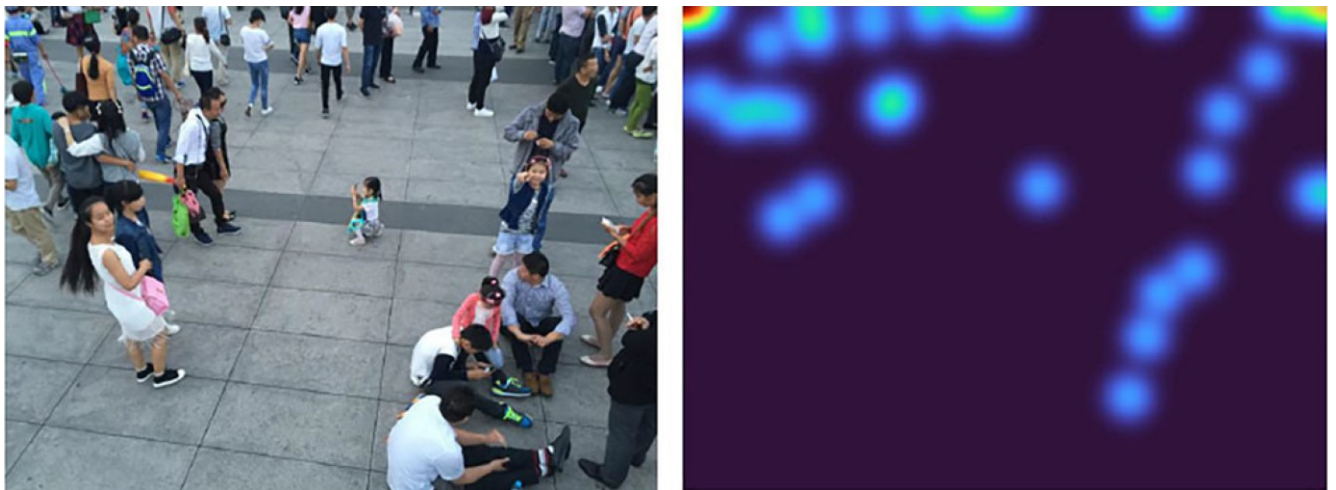


Рисунок 1.2 – Підрахунок натовпу за допомогою оцінки щільності. Прогнозована карта щільності показує щільність натовпу по всій сцені зображення. Сума значень усіх пікселів карти щільності дорівнює прогнозованому значенню кількості на зображенні [7]

Домінуючим методом підрахунку натовпу є оцінка щільності за допомогою згорткових нейронних мереж, в якій моделі прогнозують карту щільності з ядрами Гаусса, зосередженими на виявлених головах. Зазвичай навчання полягає в мінімізації середньоквадратичної похибки.

Ранні моделі, такі як CrowdCNN [7], використовували одностовпчикову архітектуру, але страждали від нездатності впоратися з варіаціями масштабу. Багатостовпчикові штучні нейронні мережі вирішили цю проблему, обробляючи вхід на кількох масштабах (різних роздільностях) у паралельних гілках; це покращує роботу з об'єктами різного розміру, але збільшує споживання пам'яті через одночасне зберігання кількох карт ознак і параметрів.

Наступні архітектури, такі як мережі кодерів-декодерів та багатомасштабні моделі, покращили вирішення задачі у багатьох аспектах, підвищивши надійність і зберігши просторову деталізацію – всупереч тому, що можна було б припустити на основі наборів даних для виявлення корисності. Ієрархічне вилучення ознак, звичайно, можна досягти за допомогою таких моделей, але це пов'язано з довшим часом навчання та виведення [8].

В останні роки були розроблені моделі на основі трансформаторів, які краще враховують глобальний контекст, використовуючи самоуважність. Але вони важкі в обчислювальному плані і менш застосовні для сценаріїв реального часу або граничних ситуацій.

Підрахунок натовпу також є одним з найпопулярніших напрямків комп'ютерного зору в останні роки, з різноманітними наборами даних, такими як ShanghaiTech [9], UCF-QNRF [10], JHU-Crowd [11] і NWPU-Crowd [12]. Ці набори даних надають можливість для порівняльного аналізу за різних умов. Хоча ранні розробки були значно кращими, останнім часом покращення MAE відбувається повільніше.

Розглянемо відкрити проблеми та виклики задачі підрахунку людей у натовпі.

Незважаючи на ці вдосконалення штучних нейронних мереж, такі проблеми, як оклюзія, спотворення перспективи і складний фон залишаються. Використання

попередньої обробки сегментації все ще дає більше переваг, але це пов'язано з більшими обчислювальними вимогами. Узагальнення також залишається проблемою: моделі, які добре працюють на одному наборі даних, не працюють у нових середовищах [13].

Ризики – надмірне використання методів найближчого сусіда (KNN) може приховати погані прогнози на складних прикладах, якщо загальний середній показник виглядає добре.

Крім того, більш глибокі моделі також переходять до ще більш неефективних висновків, прокладаючи шлях до непрактичного розгортання з точки зору граничних пристроїв з обмеженими ресурсами. Полегшені моделі розробляються, але вони, як правило, неефективні в щільних середовищах.

Із вищезазначеного можна зробити висновок, що хоча оцінка щільності і підрахунок значно просунулися вперед, тепер необхідно звернути увагу на навчання для узагальнення, створення специфічних наборів даних і розробку обчислювально ефективних моделей для реальних додатків, що працюють в режимі реального часу.

1.2.2 Задача виявлення та відстеження об'єктів

Методи виявлення та відстеження об'єктів покладаються на різні ознаки і тому шукають ознаки під час розгортання. Виявлення та відстеження об'єктів є основними проблемами комп'ютерного зору і можуть бути використані для аналізу натовпу.

Виявлення об'єктів – це процес знаходження класів об'єктів (наприклад, пішоходів, транспортних засобів) в одному кадрі, тоді як відстеження - це процес відстеження виявлених об'єктів у наступних кадрах (рис. 1.3).

Сучасні методи виявлення об'єктів можна розділити на два типи: якірні та без'якірні.

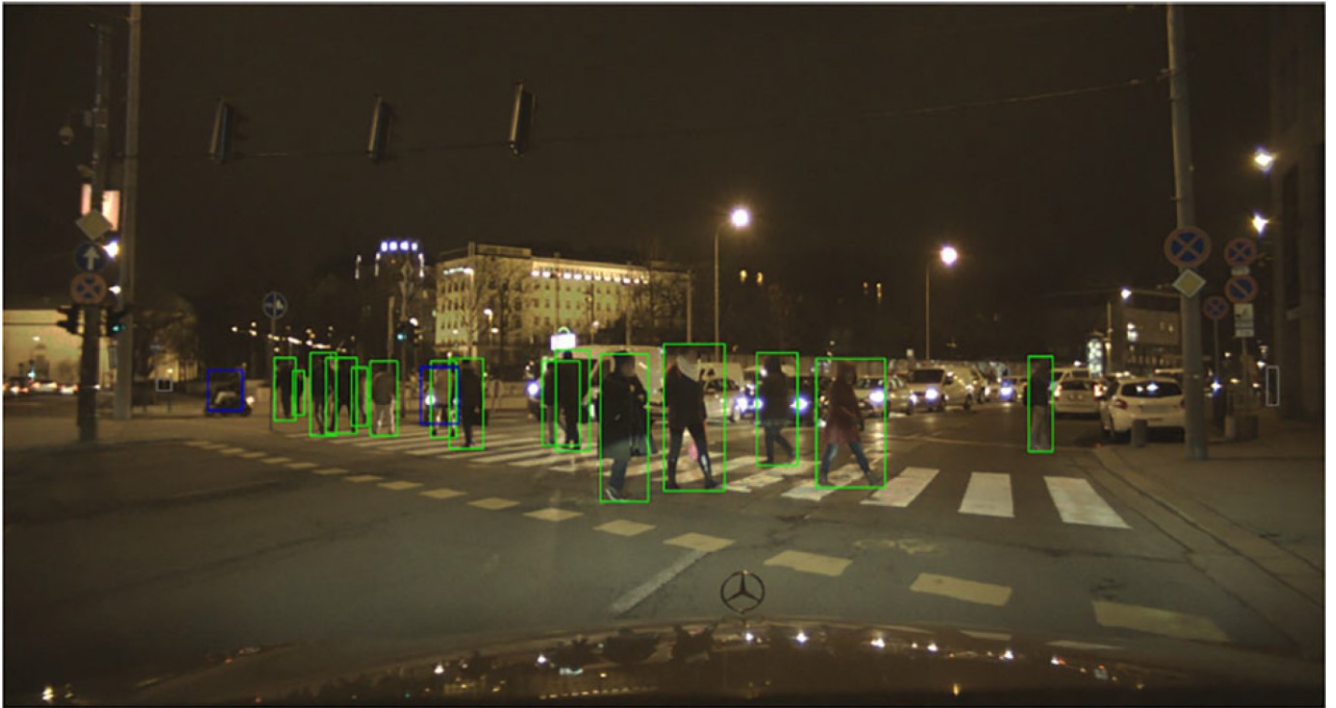


Рисунок 1.3 – Ілюстрація задачі відстеження об'єктів [14]

Методи на основі якорів, такі як Faster R-CNN [15], SSD [16] і YOLOv3 [17], передбачають місцезнаходження об'єктів, використовуючи попередньо визначені опорні точки. Вони дають дуже точні прогнози і добре узагальнюють складні сцени, але чутливі до конструкції опорних точок, а також до розбіжностей масштабу і форми об'єктів.

Без'якірні методи стали простішими та ефективнішими варіантами. Такі моделі, як CornerNet [18] і CenterNet [19] визначають ключові точки (наприклад, кути або центри) замість опорних точок.

Метод FCOS [20] виконує попиксельне прогнозування і здатен досягати конкурентних результатів. Хоча складність цих методів зменшується, але об'єкти, що перекриваються, або малі об'єкти нелегко виявити, і ці методи вимагають більшої кількості навчальних даних.

Детектори на основі трансформерів ViT [21] або DETR [22] використовують самоуважність для фіксації довготривалих залежностей і глобального контексту. Хоча вони сильно конкурують за точністю з CNN [23], CNN все ще переважають завдяки швидшому навчанню і кращій придатності до розгортання.

Набори PASCAL VOC [24] і MS-COCO [25] є одними з найпоширеніших еталонних наборів даних, що використовуються для аналізу сучасних моделей виявлення.

Незважаючи на те, що за допомогою таких підходів було досягнуто багатообіцяючих результатів, у контексті аналізу зображень натовпу все ще існують певні проблеми, які необхідно вирішити:

- інваріантність до точок зору – один і той самий об'єкт виглядає по-різному з різних точок зору;
- зміна форми – все деформується в межах кадру (наприклад, люди, що рухаються);
- екстремальна оклюзія – об'єкт ледь помітний, адже затуляється накладеними один на одного об'єктами;
- зміна освітлення та захаращення – якщо освітлення не контролюється, точність виявлення також обмежується фоновим шумом.

Ці фактори мають великий вплив на поведінку моделі в реальних ситуаціях моніторингу. Хоча ці методи підвищують достовірність розпізнавання, жоден з них не зміг вирішити ці проблеми в умовах щільного натовпу. Як наслідок, більшість сучасних детекторів погано працюють в таких умовах, що обмежує їх використання для досягнення безпеки, наприклад, у правоохоронних органах [26].

1.2.3 Задача аналізу руху натовпу

Знання, отримані з динаміки натовпу, виходять за рамки підрахунку або оцінки щільності. Цікавим може бути визначення того, чи є натовп статичним або динамічним, а також вивчення таких показників, як шляхи, кути та швидкість [27]. Аналіз руху також включає в себе: визначення підгруп, що відпочивають, і виявлення аномального руху. Застосування включають контроль доступу, аналіз трафіку, розпізнавання дій і спостереження за великими територіями (рис. 1.4).

Аналіз руху натовпу в основному спрямований на розпізнавання та аналіз нормальних моделей руху в натовпі, а також на виведення динамічної поведінки.

Це стосується просторово-часового моделювання та аналізу зображень у класичному та сучасному розумінні. Традиційні методи включають сегментацію руху, часову різницю та оптичний потік, які обчислюють попиксельний рух між кадрами [28].

Дослідниками запропоновано різні типи представлення руху:

- оптичний потік – вектори руху в кожному пікселі;
- потоки частинок/штрихових ліній – моделювання полів потоку шляхом інтегрування;
- треклети – короткі таймінги у впорядкованих серіях подій;
- просторово-часові особливості – патерни у просторі та часі.

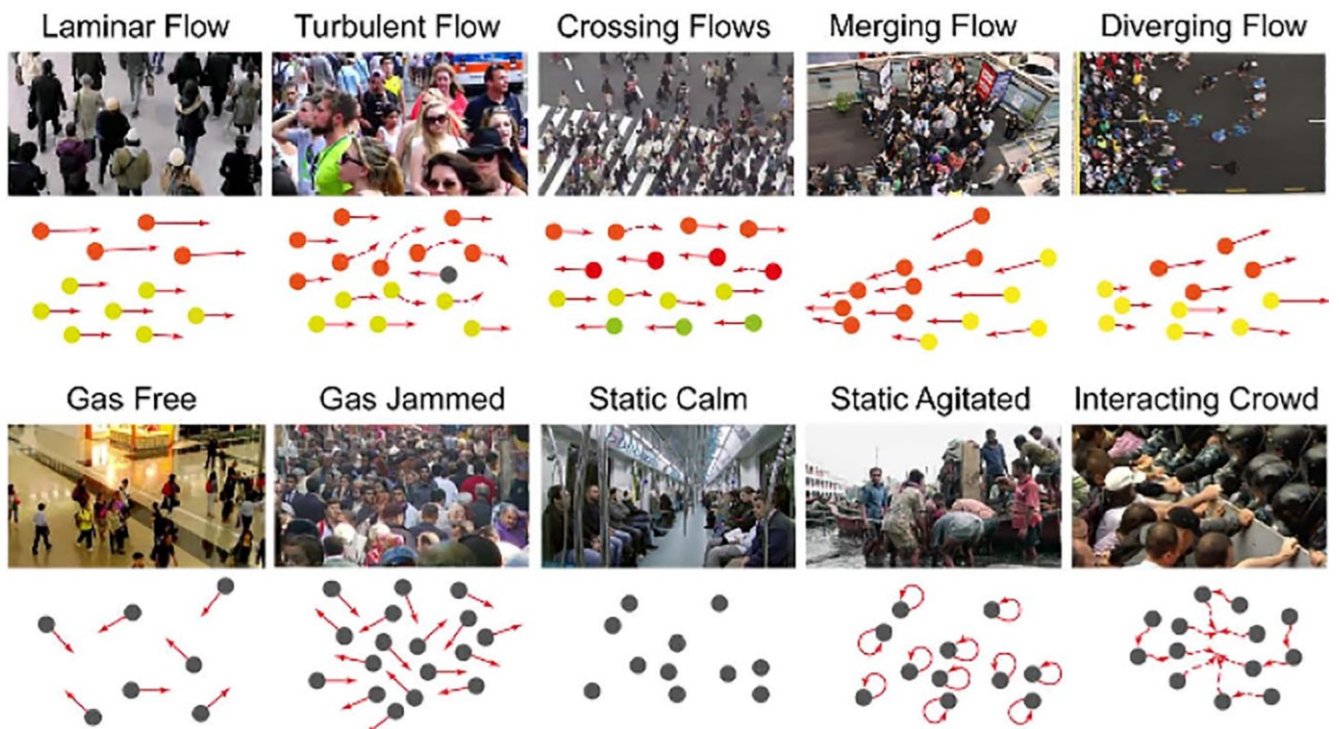


Рисунок 1.4 – Моделі руху натовпу в наборі даних Crowd 11 [29]

Сегментація руху за допомогою віднімання фону або кластеризації є дуже чутливим способом до шуму, змін освітлення та оклюзій.

Окремо слід розглянути методи на основі оптичного потоку. Вони швидкі і тому застосовуються у реальному часі, але мають труднощі з невеликими рухами або текстурами, що повторюються. Останнім часом досліджуються глобальні

моделі руху, такі як течії і хвилі [30], моделюються і сегментуються потоки натовпу.

Традиційні методи здебільшого використовують низькорівневі візуальні елементи, такі як колір або текстура, для невеликої дискримінативної семантичної інформації. Щоб полегшити семантично значуще розуміння сцени, високорівневі представлення повинні пов'язувати локальну динаміку руху з поведінкою натовпу, яка може бути пов'язана з семантикою. Основні обмеження включають:

- відсутність великомасштабних маркованих наборів відеоданих;
- відсутність можливості виявлення макропатернів або взаємозв'язків;
- низька продуктивність у реальних сценаріях, коли присутні оклюзії та складна організація потоків.

Системи повинні включати в себе моделі, які можуть не тільки навчитися прогнозувати семантику руху, але й навчитися діяти в щільних, динамічних та стохастичних середовищах [31].

1.2.4 Задача розпізнавання поведінки, активності та контексту

Розпізнавання поведінки передбачає аналіз колективних рухів і соціальних взаємодій протягом тривалого часу (від хвилин до годин) у натовпі. На цьому рівні інтерпретації метою є виявлення глобальних станів, таких як спокій, хвилювання, агресія і паніка. З іншого боку, розпізнавання активності стосується набагато коротших часових проміжків і має на меті ідентифікувати певні групові або індивідуальні дії, включаючи ходьбу, жестикуляцію, протест і біг (рис. 1.5).

Розуміння контексту – це термін, який стосується просторових, часових і семантичних аспектів знання ролі та розташування спостережуваних складних дій відповідно до конкретних умов навколишнього середовища [32].

Системи розпізнавання активності, які широко використовуються сьогодні, базуються як на традиційних імовірнісних моделях (наприклад, HMM [33] або CRF [34]), так і на DL-моделях (наприклад, 3D CNN [35]/RNN [36]/GCN [37]). Дії на мікрорівні часто локалізуються шляхом покадрового або просторово-часового

вилучення ознак, а моделювання поведінки на макрорівні передбачає узагальнення тенденцій у просторі та часі. На груповому рівні деякі методи визначають дії в термінах групових дескрипторів (наприклад, відносне положення, швидкість, згуртованість, орієнтація траєкторії) і застосовують контрольовану класифікацію (наприклад, метод групової контекстної активності) для класифікації типів поведінки [38].



Рисунок 1.5 – Приклади розпізнавання активності за кількома видами діяльності натовпу [39]

Такі набори даних, як Collective Activity [40], UCLA Courtyard [41] та Volleyball Dataset [42], також можна використовувати для навчання розпізнавання індивідуальної та групової активності під контролем. Вони пропонують позначені послідовності для базових дій (наприклад, ходьба, стояння, очікування) і складної групової поведінки, що вимагає координації та взаємодії. Деякі моделі також використовують теорії соціальних сил або біологічні симуляції для опису динаміки натовпу в громадських місцях.

Основною проблемою даної задачі є «семантична невизначеність» дій – візуальний вхід може кодувати різні дії в залежності від контексту. Крім того, оклюзія, зміна масштабу та щільність реальних сцен обмежують точне відстеження дій та збереження ідентичності. Ситуація ускладнюється, коли для розпізнавання контексту об'єднуються кілька вхідних потоків (наприклад, відео, аудіо, датчики параметрів навколишнього середовища).

Існуючі сучасні методи не здатні інтегрувати мультимодальні дані в режимі реального часу. Повністю автономні рішення потребують архітектури, яка здатна міркувати в різних часових і просторових масштабах і навчатися репрезентаціям, які узагальнюють різні середовища. Крім того, потужне контекстне моделювання все ще недостатньо вивчене [43].

Нещодавні вдосконалення глибоких нейронних мереж покращили продуктивність розпізнавання дій, але в умовах високих ризиків, коли потрібна інтерпретація і впевненість (наприклад, в управлінні натовпом або реагуванні на стихійні лиха), інтерпретованість і надійність ще не є достатніми.

Майбутні архітектури повинні бути спрямовані на гібридні моделі, які поєднують символічні міркування з вивченими функціями і які можуть використовувати і інтегрувати зовнішні джерела знань (наприклад, сигнали з соціальних мереж, метадані подій). Саме це призведе до більш адаптивного, зрозумілого і масштабованого розуміння поведінки у великих громадських середовищах, що, в свою чергу, сприятиме превентивному нагляду і заходам безпеки [44].

1.2.5 Задача виявлення аномалій

Виявлення аномалій – це процедура, яка ідентифікує виняткову або неочікувану поведінку, що суперечить нормальній поведінці. З точки зору моніторингу натовпу, ці аномалії можуть проявлятися у вигляді потенційних загроз, наприклад, панічної поведінки, ажіотажу, масової паніки/тисняви, актів публічного насильства [45]. Автоматичне виявлення аномалій є важливим

аспектом громадської безпеки, який має широкий спектр застосувань у таких сферах, як запобігання злочинності, моніторинг дорожнього руху, захист критичної інфраструктури та управління подіями (рис. 1.6).

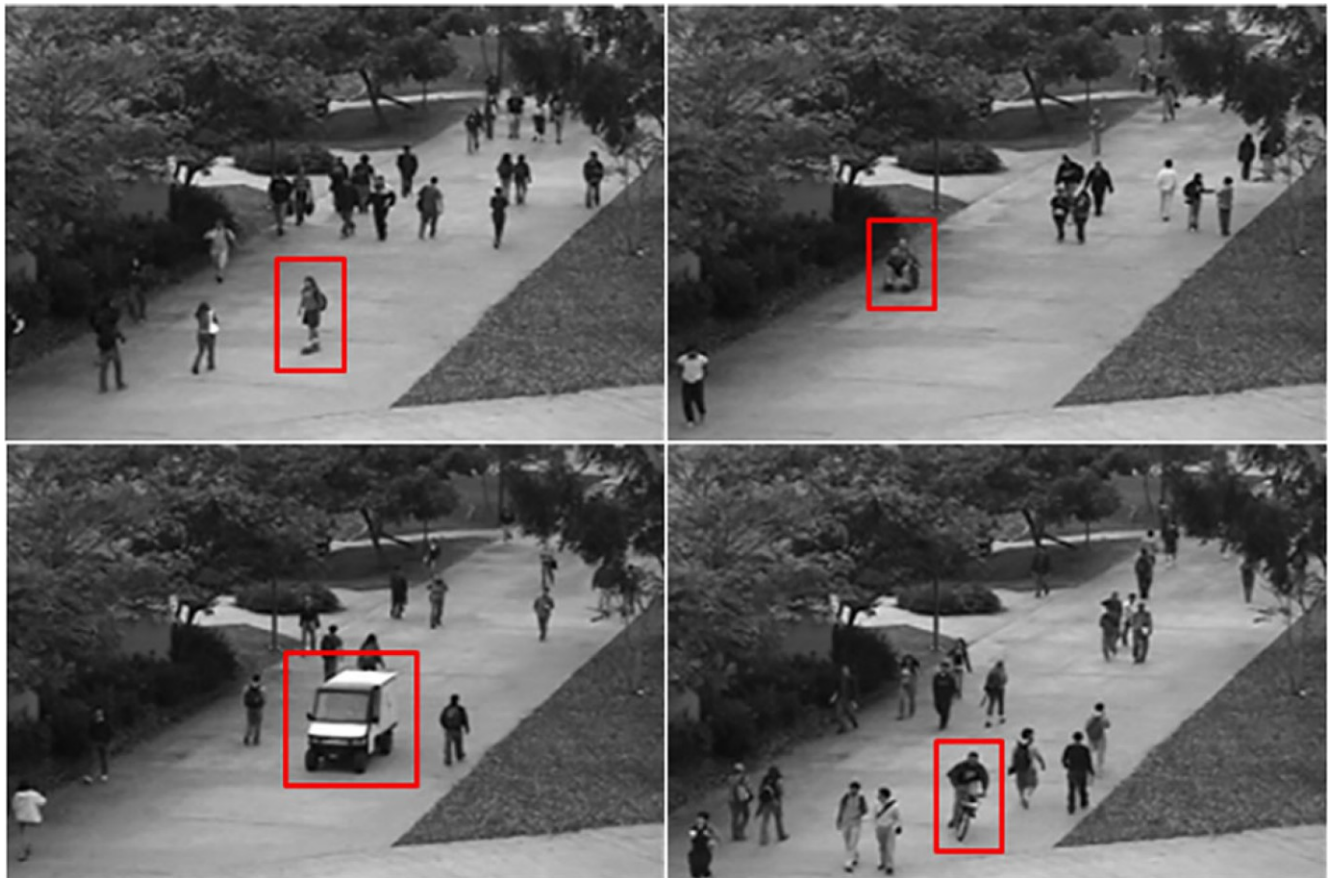


Рисунок 1.6 – Приклади аномалій у пішохідних просторах, наприклад, інвалідний візок, скейтер, велосипедист та візок [46]

Виділення аномалій є складним завданням, особливо в умовах значної розрідженості, широкого спектру варіацій і залежності від контексту. Зазвичай аномалії класифікують як:

- точкові аномалії – розглядаються як поодинокі випадки певного явища або класу об'єктів;
- нормальні речі, що відбуваються в аномальних контекстах та/або в моменти масової аномалії в часі;
- колективні аномалії – аномальна поведінка сукупностей об'єктів.

Підходи до виявлення аномалій варіюються від статистичного навчання до методів глибокого навчання:

- класичне машинне навчання – K-середні [47], SVM [48], моделі гауссових сумішей [49], однокласові SVM [50];
- глибоке навчання – CNNs [23], LSTMs [51], GANs [52], автокодери [53];
- моделі на основі фізики [54] та BoW-кодування з просторово-часовими дескрипторами [55].

Сучасні системи використовують різні рівні аналізу, такі як вектори руху на рівні об'єктів, патерни потоку в масштабі сцени або контекстно-чутливе семантичне міркування. Такі методи, як автокодери, навчаються лише на нормальних даних, щоб вивчити нормальну поведінку, щоб великі відхилення в реконструкції вказували на аномалії. Методи на основі GAN намагаються імітувати нормальний розподіл натовпу, а відхилення від нормального розподілу прогножуються як аномалії.

У дослідженнях широко використовуються 5 еталонних наборів даних: UCSD [56], PETS2009 [57], CUHK-Avenue [58], UMN [59], ShanghaiTech Campus [9]. Ці набори даних надають різні типи сцен і аномалій, але вони зазвичай страждають від розрідженого масштабу, різноманітності та якості анотацій.

До ключових викликів задачі виявлення аномалій можна віднести:

- контекст – нормальний рух в одній ситуації може бути ненормальним в іншому контексті;
- обмеженість даних – наявні на сьогоднішній день набори даних мають невеликий розмір, їм бракує екологічного різноманіття, а визначення міток не є уніфікованими;
- збір даних – справжні аномалії важко змоделювати або запрограмувати на виникнення в контрольованому середовищі;
- обчислювальні вимоги – через великі обсяги відеоданих моделі глибокого навчання повинні навчатися на великих обсягах відеоданих, а отже, потребують значної кількості графічних/теплових процесорів.

Для ефективного рішення задачі виявлення аномалій на практиці, наступні майбутні виклики можуть бути спрямовані на:

- зростання масштабів даних, різноманітність і добре анотовані дані;
- переносимі (трансферні) та адаптивні до галузі моделі в різних галузях і середовищах;
- інтеграцію мультимодальних джерел інформації, таких як аудіо, сенсорні дані та соціальні мережі;
- інтерпретовані методи ШІ для пояснення причини виявлення аномалій;
- легкі моделі для периферійного розгортання.

Підводячи підсумок, можна сказати, що хоча виявлення аномалій у натовпі все ще залишається складним завданням, це необхідна робота в ICBC. Досягнення в цій галузі залежатимуть від розвитку архітектури моделей, кращої інфраструктури даних і підвищення надійності інтерпретації дій в умовах невизначеності або динаміки [60].

1.2.6 Задача прогнозування натовпу

Задача прогнозування натовпу – це задача заздалегідь передбачити скупчення людей у певному регіоні. Прогнозування натовпу заздалегідь має ключове значення в різних сценаріях і додатках, таких як заходи, виявлення інцидентів, туристичні атракції і так далі. Прогнозування натовпу, як правило, є складним і часто використовує мультимодальні дані з різних джерел, таких як камери спостереження, соціальні мережі, сигнали мобільних телефонів, датчики тощо.

Для вивчення та прогнозування майбутнього натовпу використовується машинне навчання та методи статистичного моделювання. Цей процес може стосуватися конкретних масових заходів або зібрань, оцінки тривалості заходу, прогнозування пікових годин скупчення людей або передбачення популярності певних районів чи пам'яток під час заходу [61].

Натовп у певних адміністративних районах може виникати, як правило, по-різному, наприклад:

- «верхівка фонтану» – натовп з'являється з одного напрямку і розтікається в різні боки, подібно до фонтану;
- «вузьке місце» – натовп з'являється з різних напрямків і збирається в одній точці;
- «провулок» – натовп з'являється з рівномірною швидкістю і рухається в одному напрямку.

Прогнозування натовпу передбачає прогнозування сукупних потоків (вхідних і вихідних) у визначеному регіоні. В рамках проблеми прогнозування натовпу також вивчається додаткова інформація, така як походження і призначення потоків натовпу [62].

Прогнозування натовпу можна просто змодельовати як задачу прогнозування часових рядів для регресії збільшення щільності натовпу з часом, починаючи зі стану відсутності натовпу, або використовувати просторову статистику натовпу для прогнозування натовпу в просторових регіонах [63].

Традиційно він використовує методи ковзного усереднення, такі як авторегресійне інтегроване ковзне середнє (ARIMA [64]). Однак ці методи не здатні врахувати складні часові та просторові залежності, незважаючи на низку методів інженерії ознак.

Щоб впоратися зі складними залежностями, були запропоновані методи глибокого навчання, включаючи ШНМ [65], LSTM [51] та графові нейронні мережі (GNN) [66].

Хоча мультимодальні дані можуть бути використані для прогнозування натовпів, достатньої кількості таких наборів даних не існує.

Використання LSTM [51] та CNN [23] разом є більш перспективним методом для прогнозування натовпів на короткі проміжки часу. Однак для більш тривалих часових проміжків застосовуються методи прогнозування (наприклад, ARIMA [64]).

Виклики та відкриті проблеми. Прогнозування натовпу можна або логічно вивести з аналізу моделей руху натовпу, або спрогнозувати безпосередньо за допомогою моделей прогнозування часових рядів. Однак дослідження в цій галузі дуже обмежені, і дуже мало робіт, які часто використовують методи глибокого навчання, такі як CNN [23] і LSTM [51].

Прогнозування натовпу має як просторову, так і часову залежність, і, як правило, досить складно передбачити скупчення натовпу в регіоні протягом тривалих періодів.

Прогнозування у витоках і вузьких місцях є більш складним, ніж у схемі руху по смугах.

Через просторову залежність прогнозування потоку натовпу в регіонах неправильної форми може бути складнішим, ніж у регіонах правильної форми [67].

1.2.7 Перспективи досліджень в області візуального аналізу натовпу

В таблиці 1.1 наведено стислий опис вищезгаданих шести сфер візуального аналізу натовпу, деякі загальні приклади завдань у кожній з них, а також відкриті дослідницькі виклики.

В результаті аналізу кожної задачі в області візуального аналізу було визначено, що прогрес у вирішенні кожної задачі суттєво відрізняється. Наприклад, і підрахунок натовпу, і виявлення натовпу були значно покращені останнім часом завдяки великим еталонним наборам даних [68].

На практиці ці завдання зазвичай вирішуються за допомогою розгорнутих систем моніторингу натовпу в реальних сценаріях.

Існує також дуже чіткий бенчмаркінг у підрахунку та виявленні, який можна легко порівняти. Але, безумовно, є прогалини, які нам потрібно усунути.

У літературі з підрахунку та оцінки щільності використовується широко прийнята регресійна функція втрат (евклідова втрата або середньоквадратична похибка (СКП)).

Таблиця 1.1 – Задачі та проблеми в області візуального аналізу натовпу

Область	Приклад задач	Відкриті проблеми
Підрахунок натовпу	Знайти загальну кількість людей у зображенні, локалізація натовпу у великому регіоні	Робота з важкими оклюзіями, варіаціями масштабу у зображеннях натовпу високої щільності, спотворення перспективи, підрахунок із багатьох камер
Виявлення об'єктів	Виявлення та локалізація об'єктів інтересу, відстеження об'єктів на відео	Варіації точки зору, деформація об'єкта, сильна оклюзія, захаращеність, недостатня освітленість тощо
Аналіз руху натовпу	Оцінка колективності натовпу. Чи натовп повільно/швидко рухається? Куди рухається натовп? З якою швидкістю? Знайти траєкторії та патерни руху	Складні семантичні ознаки, складні взаємозв'язки між низькорівневими зображувальними ознаками з високорівневими семантичними ознаками високого рівня.
Вставлення поведінки	Натовп спокійний чи активний? Чи люди протестують, танцюють, б'ються тощо?	Комплексне відображення низькорівневих ознак на види діяльності, семантичний зв'язок та віддзеркалення виявлених видів діяльності на поведінку людини, стандартні визначення діяльності для семантичної поведінки.
Виявлення аномалій	Виявлення покинутих предметів. Виявлення будь-якої зброї або злочину? Виявити вторгнення	Відсутність єдиного визначення аномалій, відсутність реалістичних наборів даних, величезний розмір відеоданих
Прогнозування натовпу	Визначити, чи очікується натовп у регіоні? Як натовп накопичується з часом?	Важко зафіксувати дуальність (просторову та часову) залежності, складні патерни руху.

Однак було виявлено, що останнім часом було введено ряд інших функцій втрат (наприклад, втрата оптимального транспорту, втрата середньої точності, втрата площі під PR-кривою тощо) з заявленою кращою продуктивністю. Однак, в подальших роботах не було можливості зробити це з цими функціями втрат, і

багато останніх робіт продовжують використовувати оригінальну функцію втрат MSE. Дослідження, які розглядають статистику мобільності на рівні натовпу в аналізі руху, як правило, використовують потокові та просторово-часові характеристики. Порівняно більш вивченими є моделі «потік-поле». З іншого боку, моделі кластеризації починають привертати до себе увагу, завдяки чому вони працюють з більш складними сценаріями [69]. Завдання і дії в аналізі руху бувають різного роду, такі як виявлення джерела/поглинача натовпу, знаходження траєкторії, швидкості тощо. Ця область застосування ще не була ефективно порівняна, оскільки завдання є складним і враховує багато цілей.

Аналіз і виявлення поведінки (які іноді переходять одна в одну) є більш складним завданням, і поточний стан техніки в цій області все ще дуже поганий і фрагментований як з точки зору методів, підходів, припущень, так і цілей, хоча він є найбільш важливим у багатьох випадках використання. Нечітке визначення поведінки, діяльності та аномалії підштовхує дослідників до розгляду різних цілей і заходів, що призводить до несправедливого бенчмаркінгу.

Рекомендується використовувати загальні визначення як діяльності, так і аномалій, оскільки контекст може відігравати суттєву роль, розширюючи або покращуючи існуючі набори даних, а також створюючи більші та збалансованіші набори даних [70].

Методи, натхненні фізикою (наприклад, енергетичні моделі), є цікавими шляхами для виявлення аномалій. Насправді, зараз очікується значний прогрес в найближчому майбутньому в менш досліджених областях завдяки нещодавнім розробкам у галузі генеративного ШІ, що знижуює залежність від великих обсягів навчальних даних.

Крім того, аналіз натовпу для ICBC зазвичай вимагає швидких висновків. Це пов'язано з тим, що при високошвидкісному дротовому з'єднанні (відеоспостереження) може бути простіше виконувати всі завдання (обробку і висновки в нашому випадку) на сервері; в той час як при спостереженні з повітря (дрон) обробка і висновки на пристрої в певних випадках можуть бути більш практичними. Це означає, що легкі моделі аналізу натовпу будуть у виграші [54].

1.3 Постановка задачі дисертаційного дослідження

1.3.1 Інтеграція геометричних ознак у моделі підрахунку щільності натовпу на основі глибинного навчання

Проведений аналіз сучасних методів підрахунку людей у натовпі на основі глибинного навчання засвідчив, що більшість з них не враховують геометричну структуру сцени, зокрема перспективні спотворення та масштабні варіації, які можуть суттєво впливати на точність оцінки щільності:

- в класичних згорткових мережах просторові деформації, зумовлені перспективою, часто ігноруються, що призводить до локальних похибок, особливо у випадках сильних змін масштабу в межах одного зображення;
- відсутність урахування глибини сцени як додаткового джерела інформації обмежує здатність моделей до узагальнення в умовах складної геометрії сцени;
- сучасні моделі здебільшого ґрунтуються на єдиній універсальній шкалі, не адаптуючи рівень деталізації до локальної структури простору, що обмежує їхню точність у випадках високої щільності натовпу.

Враховуючи вказані обставини, був запропонований метод інтеграції геометричних ознак сцени, зокрема карти глибини, до глибинних нейронних мереж з метою адаптивного масштабування ознак і підвищення точності оцінки щільності натовпу.

1.3.2 Задача сегментації потоків натовпу у сценах з високою щільністю

Аналіз моделей руху в натовпі високої щільності показав, що в контексті реального відеоспостереження існує цілий ряд обмежень для існуючих методів захоплення та сегментації колективних моделей руху:

- традиційні методи оцінки руху (наприклад, оптичний потік) не здатні моделювати довгострокові просторово-часові залежності і є некогерентними протягом тривалих проміжків часу. Це особливо складно в переповнених сценах,

де оклюзії та складність сцени перешкоджають точній інтерпретації руху на рівні пікселів;

– алгоритми відстеження, засновані на виявленні об'єктів, страждають від обмеженої продуктивності на сценах з високою щільністю натовпу внаслідок перекриття об'єктів, складних патернів руху і зменшеного розміру людей в кадрі. Такі проблеми призводять до фрагментованих або недосконалих траєкторій, що перешкоджає ефективному відображенню їхніх рухів;

– існуючі методи на основі траєкторій обробляють відстеження руху у великому кадрі зображення, який містить нерелевантний фон (наприклад, будівлі, дерева), і, таким чином, абсурдно збільшують обчислювальні витрати і вносять шум у сегментацію руху;

– у динамічних сценах натовпу, таких як масові збори на переповнених міських територіях (таких як тротуари, мости, алеї), існує додаткова складність, пов'язана з тим, що кількість та орієнтація потоків руху апріорі невідомі, що робить завдання сегментації ще складнішим.

Для вирішення цих проблем у даній роботі представлений метод сегментації потоків натовпу у сценах з високою щільністю. Метод використовує кластеризацію траєкторій на основі активних контурів, яка акцентує увагу на відокремленні популярних патернів руху в локальних регіонах натовпу.

1.3.3 Представлення траєкторій для сегментації руху у натовпі

Проведений аналіз методів сегментації руху натовпу на основі траєкторій показав, що існуючі підходи мають низку обмежень:

– більшість методів не враховують комплексну багатовимірну природу траєкторій, зокрема такі характеристики, як форма руху, його напрямок, положення в просторі та локальна щільність траєкторій, що обмежує точність сегментації потоків;

– частина методів спирається виключно на просторові координати траєкторій, ігноруючи їхню часову динаміку, що знижує здатність моделі розпізнавати складні патерни руху в умовах високої щільності натовпу;

– у багатьох існуючих реалізаціях не враховується локальне оточення траєкторії, що є важливим для відокремлення схожих, але функціонально відмінних рухів, які перетинаються в просторі;

– використання одномасштабного аналізу щільності руху не дозволяє коректно враховувати вплив різних зон кадру на локальну динаміку руху.

У зв'язку з цим у даній роботі запропонована модель представлення траєкторій, яка включає багатовимірні ознаки – форму, положення, напрямок і багатомасштабну щільність – з метою забезпечення стійкого та ефективного групування траєкторій і виявлення потоків натовпу у складних сценах.

1.3.4 Виявлення аномальної поведінки у високощільному натовпі

Раннє виявлення аномальних подій у натовпі є важливим для громадської безпеки, оскільки це може призвести до виявлення і, можливо, запобігання насильницьким інцидентам з великою кількістю жертв. Однак динаміка натовпу дуже складна через загородження, різну щільність, низьку роздільну здатність і захаращеність навколишнього середовища, а отже, проблема є дуже складною:

– методи виявлення аномальних подій для класифікації дій з об'єктним та холістичним фоном. Традиційні методи виявлення аномальних подій поділяються на дві загальні парадигми: об'єктні та холістичні. Об'єктні методи намагаються представити динаміку натовпу на основі вивчення окремих елементів. Однак, у переповненому натовпі сильна ізоляція, погана видимість і малий масштаб окремих осіб ускладнюють надійне відстеження на рівні окремих осіб, що призводить до часткового або помилкового розуміння події;

– щоб вирішити цю проблему, холістичні підходи розглядають натовп як єдине ціле і обробляють інформацію про загальну динаміку і загальні закономірності руху, не відокремлюючи окремих осіб. Такі методи зазвичай

використовують дескриптори оптичного потоку та поля для обчислення макрорівневих характеристик колективної поведінки. При всій своїй корисності вони часто не враховують тонкі відхилення/переходи, які призводять до явищ (аномалій);

– попередні методи, засновані на оцінці руху від кадру до кадру, не дуже добре працюють у часових міркуваннях. Для кращого представлення основних просторово-часових закономірностей були запропоновані такі методи, як адвекція частинок і аналіз траєкторій на тривалих часових інтервалах;

– однак існує обмеження в зображенні щільних областей руху, оскільки моделі на основі частинок мають тенденцію перекривати вектори руху, і частинки більше не можуть бути ідентифіковані, що впливає на застосовність підходу до довгострокового виявлення аномалій.

Для вирішення цих проблем в роботі представлено ентропійний метод виявлення аномальних сцен натовпу за допомогою спрямованих міні-треків, побудованих з просторово-часових точок інтересу. Запропонований метод відкидає короткі та зашумлені траєкторії для збереження геометричної узгодженості напрямку та інтегрує оцінку ентропії з аналізом часового відхилення заповнюваності, щоб виявити різкі зміни в динаміці сцени.

1.4 Висновки до першого розділу

В першому розділі дисертаційного дослідження виконано детальний та системний аналіз сучасних концепцій та методів моделювання руху щільних натовпів, визначення їх структурних та динамічних особливостей. Проаналізовано існуючі роботи, серед недоліків яких особливо відзначено неефективність використання геометричної інформації сцен, нездатність враховувати довготривалі часові зв'язки в складних середовищах та відсутність врахування просторового локального контексту траєкторій. Ці недоліки значно обмежують ефективність аналізу поведінки натовпу в (погано контрольованих) сценах натовпу в реальних умовах спостереження.

Зважаючи на ці проблеми, в розділі обґрунтовується необхідність вдосконалення методів, які б вбудовували геометрію сцени в глибинні нейронні мережі. Це уможливить адаптивне масштабування об'єктів і покращить оцінку щільності.

Крім того, представлено цілісну модель представлення траєкторії, яка інтегрує багатовимірні дескриптори, включаючи форму, просторове положення, напрямок руху та різномасштабну локальну щільність, для кодування інформації про деталізацію транспортних потоків. Такі дескриптори використовуються для реалізації нового процесу кластеризації даних про рух, щоб мати змогу виявляти патерни руху в переповнених потоках.

Крім того, у розділі представлено вступні концепції для розробки ентропійного методу виявлення аномалій у динаміці натовпу. Використовуючи тимчасові стабільні, просторово когерентні представлення міні-треків і поєднуючи оцінку ентропії з локальними флуктуаціями заповнюваності, запропонований підхід покращує інтерпретованість і надійно виконує виявлення аномалій.

Сформульовано мету та задачі дослідження:

Метою дослідження є підвищення ефективності систем візуального аналізу натовпу шляхом удосконалення моделей та методів оцінки щільності натовпу за рахунок використання інформації про глибину сцени та виявлення аномальних ситуацій у сценах з натовпом високої щільності за рахунок оцінки ентропії напрямлених міні-треків та часового відхилення заповнюваності сцени.

Для досягнення мети дослідження необхідно вирішити наступні завдання:

- провести аналіз основних проблем, обмежень та викликів, які виникають в процесі побудови систем візуального аналізу натовпу та намітити шляхи їх вирішення;
- удосконалити метод інтеграції інформації про геометричні ознаки сцени у глибинні нейронні мережі, які здійснюють оцінку щільності натовпу;
- удосконалити метод сегментації потоків натовпу з високою щільністю на основі кластеризації траєкторій з використанням методу активних контурів;

- удосконалити модель представлення траєкторії руху в натовпі, що враховує комплексні ознаки руху;
- удосконалити метод виявлення аномальних ситуацій у сценах з натовпом високої щільності;
- розробити інструментальні засоби, що реалізують запропоновані рішення;
- провести експериментальне дослідження отриманих теоретичних результатів.

2 МЕТОД ПІДВИЩЕННЯ ТОЧНОСТІ ОЦІНКИ ЩІЛЬНОСТІ НАТОВПУ ШЛЯХОМ ІНТЕГРАЦІЇ ГЕОМЕТРИЧНИХ ОЗНАК СЦЕНИ

2.1 Вступ

Не зважаючи на те, що усунення викривлень перспективи традиційно розглядається як ключовий етап у класичних ручних підходах до підрахунку людей у натовпі на основі ознак [71], більшість сучасних методів, що використовують згорткові нейронні мережі, не акцентують на цьому аспекті належної уваги. Постає логічне питання: чи може врахування просторової структури сцени, зокрема її перспективних характеристик, підвищити ефективність нейромережевих моделей при оцінюванні кількості людей у натовпі?

Виходячи з цього припущення, у даному розділі пропонується метод інтеграції геометричних особливостей сцени до глибинних моделей. Крім того, в даному розділі проводиться аналіз впливу цих даних на точність оцінки щільності натовпу та демонструється, яким чином варіації масштабу в межах одного кадру можуть спричиняти суттєві відхилення у результатах. В рамках запропонованого методу, пропонується використовувати інформацію про глибину сцени як допоміжну інформацію для адаптивного формування масштабованих зображень, що дозволяє краще відображати реальні розміри об'єктів у просторі та підвищити достовірність прогнозу щільності.

Для реалізації цієї ідеї було запропоновано спеціальний модуль, який виконує інтеграцію інформації про глибину сцени поверх існуючих архітектур нейронних мереж. Цей модуль відповідає за обробку карт глибини і виконує просторову зміну значень, що характеризують індивідуальні об'єкти на зображенні. В результаті, було реалізовано дві архітектури – на базі глибокої та менш глибокої мережі.

З метою об'єктивного оцінювання ефективності розробленого методу, в п'ятому розділі дисертаційного дослідження було проведено порівняльний аналіз модифікованих мереж з низкою провідних сучасних методик підрахунку людей у

натовпі. Отримані результати підтверджують, що явне врахування геометричної інформації, зокрема перспективних трансформацій, сприяє суттєвому зростанню точності оцінки кількості осіб у складних сценах із вираженою зміною масштабу.

2.2 Вплив інформації про глибину сцени на підрахунок людей у натовпі

Коли на зображенні багато людей, зазвичай виникає проблема: об'єкти, розташовані ближче до камери, виглядають набагато більшими, тоді як ті, що знаходяться далі, – набагато меншими. Таке спотворення масштабу дуже ускладнює роботу автоматичних систем підрахунку натовпу, особливо якщо камери ще не налаштовані. На рисунку 2.1 показано, що різниця в розмірах між людьми, які знаходяться близько і далеко від камери на одному знімку, може бути дуже великою. Спрощено кажучи, чим далі людина знаходиться від камери, тим меншою вона здається.

Карти щільності використовуються в найновіших методах підрахунку натовпу, щоб здогадатися про кількість людей шляхом додавання чисел у різних частинах зображення. Але для того, щоб ці оцінки були правильними, модель повинна враховувати зміни масштабу, які відбуваються через перспективу, і змінювати числа щільності відповідно до них.

Приклад на рисунку 2.1 показує, як це працює на практиці. Три обрані ділянки зображення мають однаковий розмір, тобто однакову кількість пікселів. Однак кількість людей в кожній області дуже відрізняється: 3 в лівій області, 20 в центральній області і 1 в правій області. Це означає, що навіть якщо ці локації мають однаковий розмір, кількість людей в них дуже відрізняється. Отже, щоб модель отримала правильну кількість людей, вона повинна «знати», що предмети однакового розміру на зображенні можуть означати різну кількість людей, залежно від того, на якій відстані вони знаходяться від камери.

Це означає, що система повинна самостійно змінювати значення щільності залежно від того, наскільки великими є об'єкти в кадрі. Наприклад, значення у віддалених регіонах мають бути вищими, ніж у ближніх, оскільки на обмеженій

площі може бути більше людей. Цей метод враховує спотворення перспективи і робить підрахунок набагато точнішим, особливо в складних сценаріях з великою кількістю об'єктів, розкиданих на різній глибині.



Рисунок 2.1 – Три області, які займають однакову кількість пікселів, показують різну чисельність населення [72]

Перш ніж використовувати ознаки, значення яких залежить від перспективи, для отримання регресії щільності, класичні методи підрахунку людей у натовпі [73] часто нормалізують їх. Це робиться для того, щоб переконатися, що характеристики є правильними та надійними, а таку процедуру часто називають «нормалізацією перспективи».

Не зважаючи на те, що методи, засновані на згорткових нейронних мережах, показують більшу ефективність в процесі вирішення задачі підрахунку натовпу, ніж методи, які підбирають ознаки вручну, механізм нормалізації перспективи, який використовується в «ручних» методах, є корисним, оскільки дозволяє врахувати важливу побічну інформацію.

Звідси виникає питання, чи можливо адаптувати механізм нормалізації перспективи до використання в методах на базі глибокого навчання.

В рамках запропонованого методу пропонується використання механізму нормалізації перспективи для отримання інформації із показника глибини для формування ознак з урахуванням масштабу, а також формування оцінок щільності з урахуванням масштабу для методів підрахунку людей у натовпі на базі глибинного навчання.

Завдяки тому, що, як правило, показник масштабу буде обернено пропорційним до показника глибини, в рамках даного дослідження висувається гіпотеза, що нейронна мережа, яка буде мати інформацію стосовно глибини сцени, буде краще сприймати зміни масштабу в рамках одного зображення.

Впровадження інформації стосовно глибини сцени до нейромережевої моделі відбувається завдяки додаванню спеціального модуля врахування глибини, який додає інформацію про глибину сцени та змінює величини окремих ознак в окремих ділянках зображення для отримання представлення зображення з урахуванням глибини.

На першому етапі, блок впровадження глибини в складі модуля врахування глибини модифікує простір ознак за допомогою карти глибини. На наступному кроці, блок корекції глибини уточнює карту глибини для генерації ваг з урахуванням масштабу. Генерація ваг відбувається за допомогою механізму просторової уваги, що дозволяє надати дані про розташування об'єктів в певних ділянках зображення.

Нарешті, блок вбудовування застосовує отримані ваги для повторної зміни величини оригінальних ознак у конкретній ділянці зображення. Модифіковані таким чином ознаки демонструватимуть геометричні відмінності в масштабах між об'єктами на передньому плані в різних положеннях, що безпосередньо сприятиме оцінці значень щільності з урахуванням масштабу.

2.3 Загальна схема роботи методу

У даному дослідженні використовується популярна архітектура «кодер-декодер», яка продемонструвала свою ефективність у задачі оцінки щільності

натовпу [74]. В даному випадку роль кодера виконує згорткова нейронна мережа (CNN), яка отримує вхідне зображення і повертає ієрархію високорівневих представлень ознак, які зберігають найбільш релевантну семантичну та структурну інформацію зі сцени. Згодом декодер 2D-згорткової нейронної мережі (CNN) обробляє ці багатошарові карти ознак і генерує карту щільності, яка забезпечує приблизну оцінку кількості людей у будь-якій області зображення.

Як показано на рисунку 2.2, запропонована модифікація архітектури (позначимо її як SAWNet) намагається модифікувати оригінальні ознаки для врахування важливих геометричних властивостей за допомогою модулю інтеграції інформації про глибину, який продукує коефіцієнти масштабування для окремих ділянок на карті ознак.

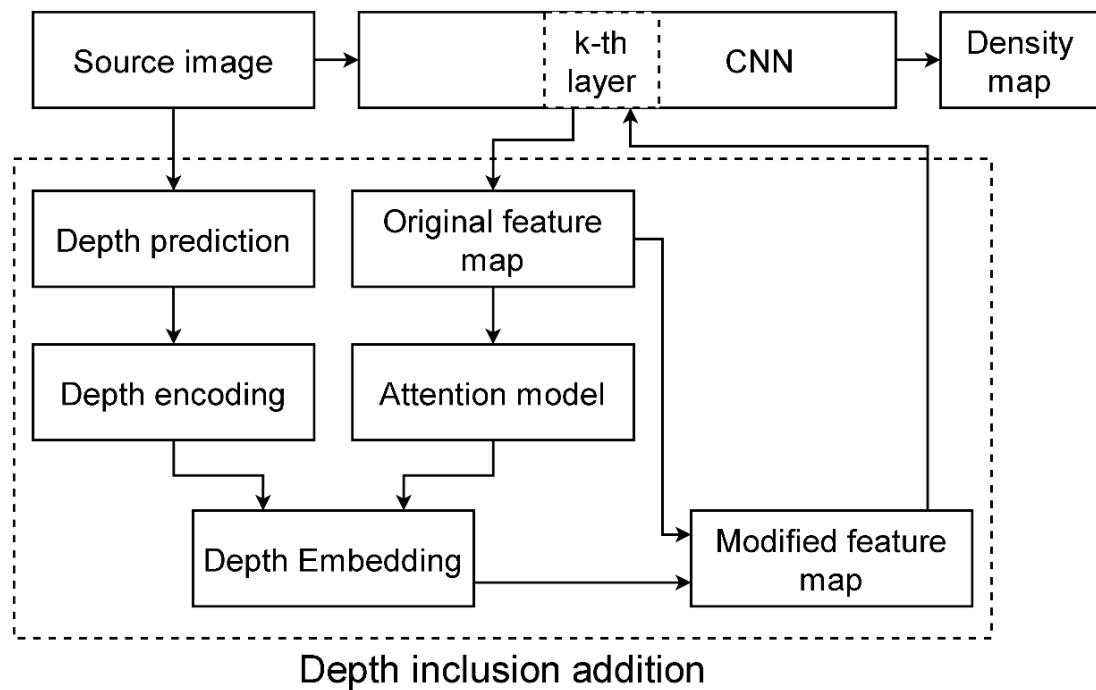


Рисунок 2.2 – Загальна схема архітектури із використанням запропонованого модулю

З формальної точки зору, припустимо що існує вхідне зображення P та відповідна карта глибини M . На i -ому шарі енкодера, коефіцієнти масштабування, які можна позначити як ваги з урахуванням масштабу ω^i , це функція від M та поточних ознак згорткової мережі W^i на шарі i .

Таким чином, запропонована мережа SAWNet модифікує поточні ознаки T^i за допомогою ваг з урахуванням масштабу ω^i у рекурентний спосіб наступним чином:

$$\begin{aligned} T^i &= CNN(W^{i-1}) \\ \omega^i &= \Psi^i(T^i, M) \\ W^i &= f(T^i, \omega^i), \end{aligned} \tag{2.1}$$

де T^i – вихід попереднього шару згортки в мережі;

M – передбачена карта глибини, отримана на попередньо навчених моделях;

Ψ^i – функція перетворення, яка продукує ваги з урахуванням масштабу в модулі інтеграції інформації про глибину сцени;

$f(\cdot)$ – вагова функція, яка модифікує ознаки за допомогою отриманих ваг з урахуванням масштабу.

Ознаки на виході шару будуть використані як вхід наступного шару до останнього шару енкодера, після чого декодер перетворить вихідні ознаки у значення щільності з урахування масштабу.

2.4 Отримання карти глибини

Так як масштаб об'єкту в сцені напряму залежить від того, на якій відстані він знаходиться від камери, для моделювання зміни масштабу об'єктів на різних ділянках сцени необхідно використати інформацію про глибину, яку можна отримати при аналізі зображення.

Нажаль, майже всі набори даних для задачі підрахунку людей у натовпі використовують окремі зображення у якості даних, що робить задачу отримання даних про глибину сцени нетривіальною.

Для отримання інформації про глибину сцени була використана модель прогнозування глибини, яка представлена в роботі [75]. Модель представляє собою

модель глибоких згорткових нейронних полів (ГЗНП), яка навчена на наборах даних та дозволяє отримувати прогноз глибини для певної ділянки сцени. Модель навчена окремо для сцен в приміщенні та для вуличних сцен.

Навчена модель ГЗНП була використана без змін та додаткового тонкого налаштування та дозволила отримати досить надійні результати прогнозів. Візуалізація прогнозів показника глибини разом з вихідним сценами наведена на рисунку 2.3. Аналіз візуальних результатів показує, що використана модель має здатність до адаптації до різних конфігурацій сцен та положення камери відносно площини сцени.

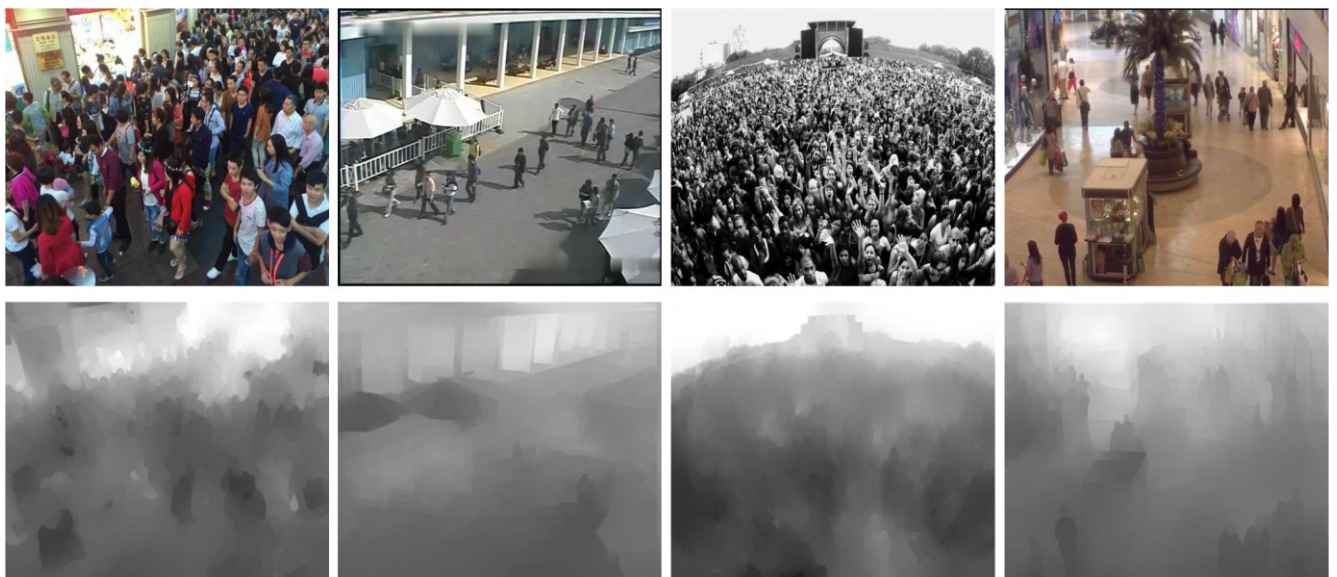


Рисунок 2.3 – Візуалізація карт глибин з попередньо навченої моделі ГЗНП для прогнозування глибини. У першому рядку показані зразки зображень з чотирьох наборів даних підрахунку натовпу. Другий ряд візуалізує прогнозовану карту глибини для кожного зображення [76]

2.5 Модуль інтеграції інформації про глибину сцени

Як було зображено на рисунку 2.2, модуль інтеграції інформації про глибину сцени складається з трьох блоків: блок впровадження глибини, блок корекції глибини та блок вбудовування. Для більш детального розуміння роботи модулю, необхідно детально розглянути роботу кожного блоку.

2.5.1 Блок впровадження глибини

Припустимо, що для вхідного зображення P , $P \in \mathbb{R}^{H \times W \times 3}$, наявна його карта глибини, отримана за допомогою моделі прогнозування глибини [75], яка позначається як M , $M \in \mathbb{R}^{H \times W}$. В рамках функціонування модулю інтеграції інформації про глибину сцени, для шару нейронної мережі i , карта глибини масштабується для того, щоб її розмір співпадав з розміром карти ознак T^i відповідного шару.

Для того, щоб отримати значення ознак з урахуванням глибини, для дальніх об'єктів, які мають малий масштаб, повинні бути призначені більші значення ваг. З огляду на те, що потрібна для цього інформація знаходиться в карті глибини, застосовується нелінійна функція кодування за допомогою параметризованої сигмоїдної функції [77], яка дозволяє нормалізувати значення глибини в діапазоні (0,1):

$$v^i = g(M) = \frac{1}{1 + e^{-\eta^i M + \mu^i}}, \quad (2.2)$$

де η^i та μ^i – навчені параметри, які використовуються для налаштування функції енкодера.

Ця функція є диференційованою, що означає, що для її навчання можна використовувати стандартні алгоритми пошуку статичного градієнтного спуску. Часткові похідні цільової функції L відносно параметрів η^i можна виразити наступним чином:

$$\frac{\delta L}{\delta \eta^1} = \frac{\delta L}{\delta v^i} \frac{\delta v^i}{\delta \eta^i} = \sum_j \frac{\delta L}{\delta v_j^i} \frac{\delta v_j^i}{\delta \eta^i} = \sum_i \frac{\delta L}{\delta v_j^i} M_j g(M_j) (1 - g(M_j)) \quad (2.3)$$

де v_j^i та D_j – j -ий елемент v^i та M .

Подібним чином, часткові похідні цільової функції L з урахуванням μ^i можуть бути визначені як:

$$\frac{\delta L}{\delta \mu^1} = - \sum_i \frac{\delta L}{\delta v_j^i} g(M_j) (1 - g(M_j)). \quad (2.4)$$

2.5.2 Блок корекції глибини

Незважаючи на те, що карта глибини надає інформацію для масштабування об'єктів на сцені, вона не розпізнає об'єкти переднього плану та фон. Використовуючи тільки карту глибини, ознаки на фонових ділянках також будуть модифіковані, що є небажаним та може зіпсувати вивченні представлення об'єктів. Наприклад, ознакам на віддалених ділянках неба будуть надані дуже великі значення ваг, що не буде впливати на оцінювання зміни масштабу між цільовими об'єктами, а також може призвести до появи додаткових шумів. З метою більш ефективного використання передбаченої карти глибини, в модуль інтеграції інформації про глибину сцени був доданий блок корекції глибини.

Для корекції глибини було б корисно мати інформацію про ділянки сцени із можливим скупченням людей. Однак, набори даних, які використовуються для рішення задачі підрахунку людей, мають лише точкові мітки людей, а додатково виділяти ділянки із скупченням людей є накладним. Тому для вирішення даної задачі був використаний механізм просторової уваги [78], який дозволяє відокремити об'єкти переднього плану за допомогою м'якої маски n^i , яка використовується в блоці корекції глибини в шарі i . Ця маска уваги допомагає зосередитися на оцінці глибини серед цільових об'єктів та не брати до уваги дані глибини фонових ділянок. Маска уваги $n^i \in \mathbb{R}^{M \times N}$ може бути представлена як функція карти ознак $T^i \in \mathbb{R}^{M \times N \times C}$:

$$n^i = \text{sigmoid}(\Omega_s(T^i)), \quad (2.5)$$

де Ω_s – модель уваги згорткової нейронної мережі, яка складається з двох шарів згортки з ядром 3×3 (перший шар має 512 фільтрів, другий шар має 1 фільтр).

Ваги моделі уваги знаходяться за допомогою поелементного використання сигмоїдної функції на вихідній карті від Ω_s для того, щоб підкреслити найбільш важливі ділянки на сцені. В рамках поставленої задачі, фокус буде зосереджений на ділянках з людьми на передньому плані.

На рисунку 2.4 наведені візуалізації масок уваги для декількох сцен із зазначенням вихідної сцени. У другому та третьому стовпчиках показані варіації карти уваги для різних показників глибини вихідної моделі.

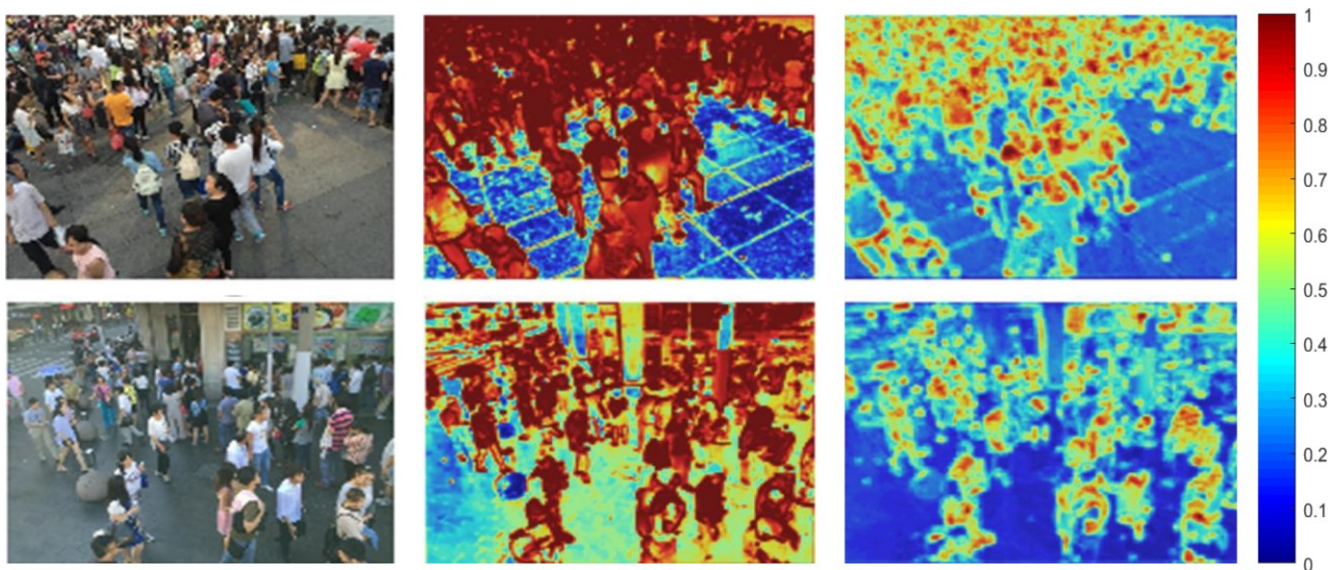


Рисунок 2.4 – Візуалізація масок уваги. У першому стовпчику наведено два приклади зображення. Другий і третій стовпчик відповідно візуалізують навчені маски уваги маски, коли модуль уваги встановлюється на все більший глибини моделі

Отримана візуалізація результатів дає можливість стверджувати, що механізм просторової уваги спроможний виділяти ділянки з людьми на передньому плані від ділянок фону. Крім того, завдяки ієрархічній природі моделі представлення ознак, яка реалізується в згорткових нейронних мережах, механізм просторової уваги має змогу генерувати маски уваги на різних семантичних рівнях – при збільшенні показника глибини, маска уваги концентрується на більш

абстрактних представленнях. Простіше кажучи, при збільшенні показника глибини, маска уваги переходить від підкреслення ділянок зі скупченням людей до підкреслення ділянок голів окремих суб'єктів.

На рисунку 2.5 проілюстрований результат роботи блоку корекції глибини. Візуалізація результатів дає змогу стверджувати що після роботи блоку корегування фонові ділянки були пригнічені, разом з тим відмінність в показниках глибини між об'єктами переднього плану зберігається, що говорить про ефективність процедури отримання скорегованих ваг.

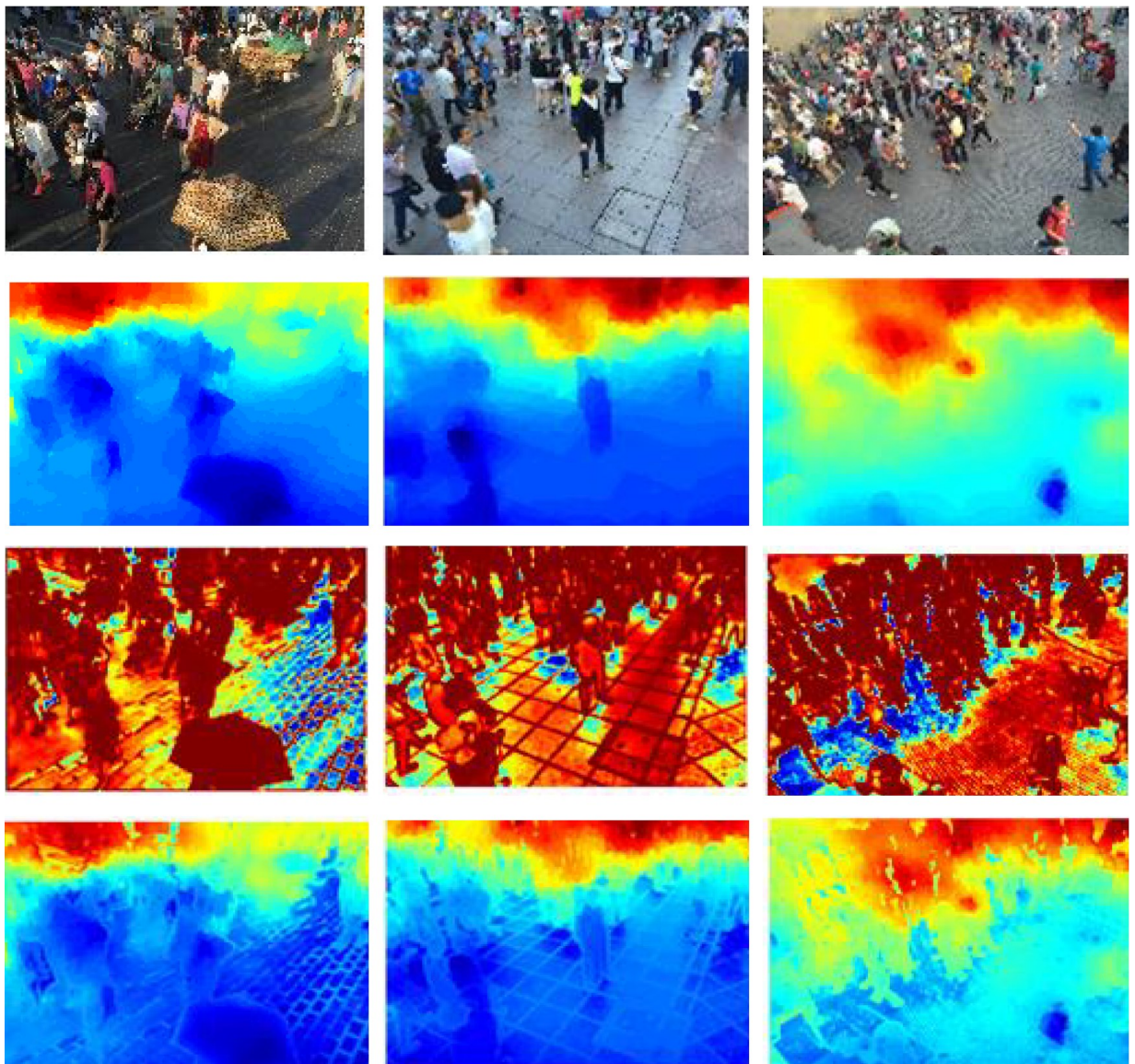


Рисунок 2.5 – Візуалізація зображень, масок уваги, карт глибини та згенерованих скорегованих карт ваг з урахуванням масштабу

Отримана маска уваги далі використовується для корекції початкової закодованої глибини для отримання ваг ω^i , які враховують масштаб та механізм уваги:

$$\omega^i = \Psi^i(T^i, M) = n^i \odot v^i, \quad (2.6)$$

де $\odot$ – означає поелементне множення двох матриць.

Завдяки перемноженню маски уваги та закодованих значень глибини можна скорегувати та пригнітити неважливі сигнали фонових ділянок.

2.5.3 Блок вбудовування

Отримані на виході роботи блоку корекції скореговані ваги, дозволяють модифікувати оригінальні ознаки T^i мережі. Це відбувається за допомогою лінійної вагової функції f , яка використовується у якості петлі зворотного зв'язку. Багато популярних стратегій модифікації ваг агрегують ознаки з різних ділянок на основі отриманих ваг.

В даному випадку відбувається поелементне множення. Внаслідок цього, ознаки змінюються з урахуванням як інформації про глибину, так і семантичної інформації в одній ділянці.

Отримані ознаки з урахуванням масштабу W^i з підкресленими змінами масштабу людей на передньому плані можуть бути обчислені за допомогою наступного рівняння:

$$W^i = f(T^i, \omega^i) = T^i \otimes \omega^i, \quad (2.7)$$

де $\otimes$ – означає операцію поелементного множення, яка виконується незалежно для кожного каналу.

2.6 Нейромережева модель SAWNet

Запропонований модуль інтеграції інформації про глибину сцени працює незалежно та має однакову розмірність вхідних та вихідних даних. Це означає, що він може бути використаний в стандартній архітектурі згорткової нейронної мережі для покращення представлення сцени без додаткових змін в оригінальній архітектурі.

Для дослідження роботи запропонованого модулю на різних нейромережевих моделях, була запропонована мережа SAWNet, яка представляє собою модель на базі архітектури енкодер-декодер [74]. В рамках експериментального дослідження були розроблені два варіанти мережі SAWNet: SAWNet-S (Shallow) та SAWNet-D (Deep).

Мережа SAWNet-S представляє собою полегшену нейромережеву модель, яка має три згорткові шари в енкодері та декодері. Модель побудована по принципу повнозв'язної мережі та приймає дані різних розмірів. Мережа SAWNet-D представляє собою глибинну нейронну мережу на базі архітектури VGG [79]. Архітектури представлених мереж наведені в таблиці 2.1.

Таблиця 2.1 – Архітектури обраних мереж

Мережа	SAWNet-S	SAWNet-D
Енкодер	7x7x32 згортка 7x7x64 згортка, зсув 2 5x5x128 згортка	(3x3x64 згортка)x2, зсув 2 (3x3x128 згортка)x2, зсув 2 (3x3x256 згортка)x2, зсув 2 (3x3x512 згортка)x2, зсув 2
Декодер	5x5x64 згортка 7x7x32 деконв, апсемпл 2 7x7x1 деконв, апсемпл 2	(3x3x512 згортка, дилат 2)x2 3x3x256 згортка, дилат 2 3x3x128 згортка, дилат 2 3x3x64 згортка, дилат 2 1x1x1 згортка

. За кожним згортковим блоком представлених мереж слідує блок ReLU [80], який доповнюється для підтримання роздільної здатності.

Слід зазначити, що запропонований модуль інтеграції інформації про глибину сцени може бути вбудований в мережу в декількох екземплярах на різних шарах моделі.

2.6.1 Навчання моделі SAWNet

Моделі SAWNet може бути навчена за допомогою попіксельної Евклідової функції втрат [81]:

$$L = \|\widehat{DM} - DM\|^2, \quad (2.8)$$

де $\widehat{DM}$ та DM – прогнозована та фактична карта щільності відповідно.

Для зображення P з точковими анотаціями A_p (анотації є частиною навчальної вибірки), фактична карта щільності представляє собою суму множини двовимірних Гауссових функцій, центрованих в кожній точці:

$$\forall p \in P, DM(p) = \sum_{\mu \in A_p} \mathbb{N}(p; t, \Sigma), \quad (2.9)$$

де $\mathbb{N}(p; t, \Sigma)$ – нормалізоване двовимірне Гауссове ядро, обчислене для p , з середнім t та коваріаційною матрицею Σ .

Процес навчання моделі складається з трьох етапів:

- 1) базова модель оптимізується за допомогою цільової функції L ;
- 2) відбувається навчання моделі уваги для проведення кращої ініціалізації;
- 3) додається модуль інтеграції інформації про глибину сцени та загальна модель навчається наскрізно з використанням цільової функції L .

2.6.2 Дослідження роботи блоків модулю інтеграції глибини

Для визначення ефекту запропонованого модулю інтеграції інформації про глибину сцени, були проведені детальні експерименти на наборі даних ShanghaiTech-part B, який є другою частиною набору ShanghaiTech [9]. Даний набір даних містить 716 анотованих зображень із різних камер відеоспостереження, що демонструють сцени з натовпом. Через викривлення перспективи на одному зображенні можуть бути люди з різними масштабами. У дослідженні використано рекомендовані параметри з оригінальної роботи: 400 зображень для навчання, 316 —, для тестування. Для тренування випадково обирались 20 фрагментів розміром 224×224 пікселі.

У якості метрик оцінки роботи нейромережевої моделі були обрані метрики середньої абсолютної похибки (Mean Absolute Error, MAE) [82] та середньої квадратичної похибки (Mean Squared Error, MSE) [83].

Аналіз роботи блоків модулю. Для аналізу ролі кожного блоку в складі модуля, були проведені експерименти з двома варіантами запропонованого модулю. Перший варіант модулю містить блок впровадження глибини та блок вбудовування, але блок корекції глибини відсутній (позначимо мережу з таким модулем як CNN-1). Другий варіант модулю містить блок корекції та блок вбудовування, але блок впровадження глибини відсутній (позначимо мережу з таким модулем як CNN-2).

Для більш детального вивчення ефектів від модифікації ознак моделі на різних шарах, модуль інтеграції інформації про глибину та його варіанти також були впроваджені на різних шарах базової нейромережевої моделі. Зокрема, впровадження блоку на шарі n -ної групи шарів conv-relu-pool як «етап n ». Для моделі SAWNet-S, етапи 1, 2 та 3 представляють собою шари pool1, pool2 та conv3 відповідно. Результати експериментів з різними блоками модулю та з використанням мережі SAWNet-S у якості базової, наведені в таблиці 2.2.

З результатів експериментів видно, що за допомогою інформації про глибину сцени, мережа CNN-1 показує за показниками MAE/MSE кращі результати ніж

базова модель без блоку інтеграції інформації про глибину сцени. Причому покращення спостерігаються на різних етапах впровадження модулю в базову модель. Це дозволяє підтвердити гіпотезу, що використання інформації про глибину сцени дозволяє покращити ефективність рішення задачі підрахунку людей у натовпі. Додавання блоку корекції глибини ще більш покращує результати порівняно з мережею CNN-1, що говорить про те, що використання механізму просторової уваги дозволяє уникнути шуму з боку фонових ділянок.

Таблиця 2.2 – Показники роботи блоків на наборі ShanghaiTech-B

	Етапи		
Мережа	1	2	3
SAWNet-S без модуля	13,03 / 21,90 (MAE / MSE)		
CNN-1	12,68 / 22,12	12,89 / 22,45	12,78 / 22,42
CNN-2	12,24 / 21,10	11,95 / 21,10	12,08 / 20,15
SAWNet-S	11,83 / 19,76	11,85 / 20,50	12,24 / 20,03

Слід зауважити, що мережа CNN-2 лише незначно покращує результати роботи порівняно з базовою моделлю без модулю та показує гірші результати порівняно з мережею CNN-1 та порівняно з використанням повноцінного модулю. Це свідчить про те, що модуль інтеграції інформації про глибину дозволяє покращити вирішення задачі підрахунку людей у натовпі в більшій мірі за рахунок роботи блоку впровадження глибини та блоку вбудовування.

Використання декількох екземплярів модулю. Для того, щоб визначити вплив використання модулів інтеграції інформації про глибину на декількох шарах мережі на точність підрахунку людей у натовпі, далі був проведений експеримент з послідовною модифікацією ознак на декількох етапах базових моделей SAWNet-S та SAWNet-D (табл. 2.3).

По-перше, можна помітити, що для мережі SAWNet-S, незважаючи на кількість інтегрованих модулів інтеграції інформації про глибину сцени, вона

продовжує показувати кращі результати порівняно з базовим варіантом, що підтверджує ефективність запропонованого методу.

Аналогічний висновок можна зробити і при використанні SAWNet-D у якості опорної моделі. Крім того, при використанні SAWNet-S у якості базової моделі, використання двох модулів відповідно на першому та другому етапах є кращим, ніж використання лише одного. Однак, при додаванні третього модулю інтеграції інформації про глибину сцени, можна зробити висновок, що робота моделі виходить на плато, так як показник MSE покращується, тоді як показник MAE стає гіршим.

Таблиця 2.3 – Показники роботи мереж з різною кількістю модулів

	Кількість модулів		
Мережа	1	2	3
SAWNet-S без модуля	13,03 / 21,90 (MAE / MSE)		
SAWNet-S	11,83 / 19,76	11,33 / 18,58	11,64 / 18,40
SAWNet-D без модуля	10,5 / 15,8 (MAE / MSE)		
SAWNet-D	8,10 / 13,01	8,06 / 13,50	8,25 / 14,39

При використанні мережі SAWNet-D продуктивність значно покращується при використанні одного модуля, але досягає плато при використанні двох модулів. При використанні трьох модулів, робота мережі стає гіршою.

Порівнюючи результати для SAWNet-S та SAWNet-D, було визначено, що для неглибоких мереж з низькою здатністю до представлення, послідовна корекція ознак для вбудовування необхідної інформації про глибину є корисним і дозволяє повністю використати потенціал самої базової моделі.

У той час як для більш глибоких мереж, які від початку є сильнішими, щоб генерувати надійні представлення, використання модулю інтеграції інформації про глибину на ранніх шарах мережі є достатнім, тоді як додавання додаткових модулів може привести до перенавчання моделі.

2.7 Висновки до другого розділу

У цьому розділі було розроблено метод підрахунку людей у натовпі з урахуванням глибини сцени як додаткового джерела просторової інформації. Метод спрямований на компенсацію викривлень перспективи, які призводять до суттєвих змін масштабу людей в межах одного зображення. Запропоновано інтегрувати карту глибини в обчислення щільності натовпу шляхом зміни просторових ознак відповідно до положення об'єктів у глибині сцени. Це дозволяє покращити узгодженість між фізичними розмірами об'єктів та їх цифровим представленням на рівні ознак.

На основі карти глибини виконується модифікація ознак згорткової мережі: дальні області підсилюються, ближні —, приглушуються. Для зменшення впливу фонових ділянок на результати було використано механізм просторової уваги, який автоматично виділяє регіони, що ймовірно містять скупчення людей. Це дозволяє більш точно визначати просторову структуру сцени та уникати спотворень у регіонах, де людей немає.

Особливу увагу в дослідженні приділено вивченню ролі кожного окремого блоку в складі запропонованого методу. За допомогою серії експериментів було оцінено внесок блоку впровадження глибини, блоку корекції та блоку вбудовування. Результати показали, що навіть часткова реалізація підходу (наприклад, лише впровадження глибини без корекції) дає покращення точності порівняно з базовими моделями. Повна комбінація трьох блоків демонструє найкращі результати, що підтверджує доцільність комплексного врахування просторових характеристик сцени.

Крім того, було проведено аналіз впливу кількості екземплярів модуля інтеграції інформації про глибину на різних шарах мережі. Показано, що для неглибоких моделей послідовне додавання модулів на декількох рівнях дає стаłe покращення. Натомість для глибших мереж надмірна кількість модулів може призвести до перенавчання й погіршення результатів. Це вказує на необхідність балансування між глибиною архітектури та кількістю просторових корекцій.

У підсумку, результати дослідження підтверджують ефективність методу, що враховує глибину сцени. Такий підхід дозволяє зменшити похибки, пов'язані з перспективними викривленнями, точніше відображати щільність натовпу та є перспективним для застосування у складних умовах відеоспостереження.

Отриманий наступний пункт наукової новизни:

– *удосконалено* метод інтеграції інформації про геометричні ознаки сцени у глибинні нейронні мережі, які здійснюють оцінку щільності натовпу. Використання запропонованого методу дозволяє краще відображати реальні розміри об'єктів у просторі та підвищити достовірність оцінки щільності.

3 МОДЕЛІ ТА МЕТОДИ ОБРОБКИ ТА АНАЛІЗУ СЦЕН З ВИСОКОЮ ЩІЛЬНІСТЮ НАТОВПУ ДЛЯ ВИЯВЛЕННЯ АНОМАЛІЙ

3.1 Метод сегментації потоків натовпу в сценах з високою щільністю натовпу

Виявлення структури та напрямку руху натовпу є важливим кроком у візуальному аналізі натовпу в системах інтелектуального відеоспостереження. Масові заходи, такі як мітинги, марафони, паради, протести, фестивалі, спортивні матчі тощо, передбачають рух натовпу в обмеженому просторі, наприклад, на міському тротуарі, мостових переходах або вузьких дорогах. Важливим кроком в аналізі цих зон скупчення людей є сегментування домінуючих моделей і напрямків руху.

Сцена натовпу може складатися з декількох потоків натовпу, при чому інформація про кількість потоків та їх знаходження відсутня. Це робить сегментацію натовпу складним завданням. Призначенням системи комп'ютерного зору є захоплення точної інформації про рух, яка базується на його представленні. Необхідне представлення руху повинно створювати довгі і надійні траєкторії. Використовуючи ці траєкторії, можна описати всі потоки на відео.

Традиційно метод оптичного потоку використовується для обчислення попиксельного потоку між двома кадрами [84]. Однак, оптичний потік не може бути безпосередньо використаний для представлення руху через деякі неявні недоліки. Наприклад, оптичний потік не здатен виділити просторово-часові патерни руху на великій часовій відстані, що було б корисними у великій кількості застосувань. У сценах з високою щільністю натовпу виділення повних траєкторій руху натовпу є складним завданням. Через вищезгадані недоліки оптичних потоків, багато робіт присвячених аналізу натовпу, використовують підходи на основі траєкторій або треків [85]. Ці методи дозволяють отримати довгі та надійні траєкторії для представлення руху, але, разом з тим, вони мають багато недоліків. Наприклад, в роботі [86] траєкторії виділяються у всьому кадрі з натовпом, що не є необхідним,

враховуючи, що частина зображення зайнята будівлями, стінами, деревами тощо. Для вирішення даної проблеми необхідно провести попередню обробку зображення для сегментування області натовпу для зменшення обчислень.

Більше того, ефективність більшості стандартних методів відстеження, заснованих на виявленні об'єктів, погіршується в щільному натовпі [87]. Це пов'язано з тим, що в сценах з високою щільністю натовпу спостерігається складна динаміка між окремими особами, а також присутність об'єктів дуже малого розміру з оклюзією, які непрактично відслідковувати одночасно.

В рамках даного дисертаційного дослідження пропонується метод сегментації натовпу з високою щільністю. Метод використовує алгоритм кластеризації траєкторій на основі активних контурів з метою сегментації патернів натовпу.

3.1.1 Загальний принцип роботи запропонованого методу

На рисунку 3.1 наведена загальна схема роботи методу. Запропонований метод має дві сильні сторони: по-перше, низька обчислювальна складність з процедурою оновлення дозволяє використовувати цей метод в системах інтелектуального відеоспостереження в реальному часі. По-друге, представлення ознак траєкторії створює надійний дескриптор. Застосовуючи метод активних контурів, метод дозволяє відстежувати лише область натовпу замість того, щоб відстежувати весь кадр відео. Траєкторії руху є ефективним способом отримання та представлення складного руху на сцені протягом часу. Тому ми вирішуємо задачу сегментації потоку натовпу за допомогою вилучення траєкторій та кластеризації, яка відбувається в області натовпу.

Запропонований метод складається з наступних кроків:

- 1) сегментація на основі активних контурів – відокремлення області натовпу від усього кадру;
- 2) виділення траєкторій – виявлення точок інтересу в блочній структурі і відстеження цих блоків на кадрах для виділення траєкторій;

3) кластеризація траєкторій – кластеризація вилучених траєкторій за допомогою алгоритму кластеризації, який використовує форму, розташування, напрямок і щільність сусідства шаблонів траєкторій;

4) сегментація потоку натовпу – позначення окремих пікселів залежно від їхньої належності до певного патерну потоку для фінальної сегментації потоків натовпу.



Рисунок 3.1 – Загальна схема роботи методу

3.1.1.1 Сегментація на основі активних контурів

Сегментація зображень відіграє важливу роль у виявленні, класифікації, розумінні сцени та інших завданнях візуального аналізу. Моделі активних контурів (АСМ) широко використовуються для сегментації зображень [88]. Для отримання точних результатів ця модель використовує попередні знання, пов'язані з розподілом інтенсивності зображення, формою границь і текстурою.

Моделі активних контурів можна класифікувати як моделі на основі країв [89] або моделі на основі областей [90]. У методах на основі країв градієнт зображення змушує активні контури зміщуватися до необхідних меж заданих об'єктів. Ці методи схильні до шуму і слабких країв, які мають малі значення градієнта, що може призвести до витоку країв.

У методах, заснованих на областях, використовується статистична інформація зображення для примусового переміщення активних контурів до межі

об'єкта. У деяких випадках цей метод є більш ефективним, ніж методи на основі країв. В рамках запропонованого методу використовується метод на основі областей, який ґрунтується на тому, що в сценах з високою щільністю натовпу область натовпу здається іманентно однорідною через те, що вона виглядає однаково на відстані. Виходячи з цього, однорідна область забезпечує невеликі зміни інтенсивності, тоді як неоднорідна має величезну варіацію інтенсивності. Таким чином, різниця між інтенсивністю є важливою ознакою для сегментації контуру переднього плану з усього кадру.

Для зменшення області трекінгу шляхом відстеження лише області з натовпом, використовується сегментація контуру за допомогою моделі CV [90]. В рамках процесу сегментації ми вважаємо, що відеокادر можна розбити на область переднього плану C_{in} і область фону C_{out} , а розподіл інтенсивності в різних частинах вважається незалежним. Контур генерується рекурсивно в неконтрольований спосіб для мінімізації енергетичної функції, і мінімум досягається, коли контур, що еволюціонує, досягає межі між областями переднього і заднього планів.

Функція енергії CV-моделі для скалярного зображення обчислюється наступним чином:

$$\eta \cdot Len(C) + \omega \cdot Area(C_{in}) + \gamma^F \int_{C_{in}} (P(x) - C^F(C))^2 + \gamma^B \int_{C_{out}} (P(x) - C^B(C))^2, \quad (3.1)$$

де C – контур;

$P(x) \in R$ – інтенсивність зображення в точці x ;

$\eta \geq 0$ – параметр регуляризації, який контролює гладкість контуру;

C_{in} та C_{out} – області переднього плану та фону відповідно;

$\omega \geq 0$ – параметр регуляризації, який штрафує за велику площу переднього плану;

$Len(C)$ – довжина контуру;

$Area(C_{in})$ – площа області всередині контуру (передній план).

Значення $C^F(C) = avg(P(x)|x \in C_{in})$ та $C^B(C) = avg(P(x)|x \in C_{out})$ представляють середні інтенсивності переднього плану та фону відповідно. Параметри $\omega^F, \omega^B \geq 0$ відповідають за вплив двох термів енергії зображення $\omega^F \int_{C_{in}} (P(x) - C^F(C))^2$ та $\omega^B \int_{C_{out}} (P(x) - C^B(C))^2$ відповідно всередині та ззовні контуру.

Контур C можна виразити як множина нульового рівня допоміжної функції $\beta: \Theta \rightarrow \mathbb{R}$:

$$C := \{x \in \Theta: \beta(x) = 0\}. \quad (3.2)$$

Виходячи з вищезазначеного рівняння, передній та задній плани контуру C можна окремо представити наступним чином:

$$C_{in} := \{x \in \Theta: \beta(x) > 0\}, \quad C_{out} := \{x \in \Theta: \beta(x) < 0\}. \quad (3.3)$$

Сегментована область переднього плану далі обробляється для виявлення просторово-часових точок інтересу, які потім відстежуються протягом наступних кадрів.

Згідно із запропонованим методом, вікно оновлюється через певну кількість кадрів t , після чого існуючий контур області повторно сегментується (ця процедура відбувається кожен раз після t кадрів).

3.1.1.2 Обчислення траєкторій

Зазвичай, відео зі скупченням людей щільно заповнені рухомими людьми, а кожна особа займає на зображенні малу кількість пікселів. За таких умов, методи основані на виявленні та відстеженні окремих об'єктів є неефективними із-за великої похибки та низької швидкості обробки кадру. Таким чином, замість

концентрації на окремих об'єктах, в запропонованому методі використовується відстеження на рівні блоків у сегментованому регіоні C_{in} .

Для реалізації даного механізму, сегмент C_{in} розбивається на блоки, які не перетинаються, розмір кожного блоку становить $s \times s$. У граничних точках деякі з блоків містять як передній, так і задній план. В рамках даного методу розглядаються лише ті блоки, які містять більше ϕ відсотків пікселів сегменту з натовпом:

$$B(k) = \begin{cases} \text{обраний, якщо } C_{in}(\text{countPixels}(B(i))) > \phi\% \\ \text{відкинутий, в іншому випадку} \end{cases}. \quad (3.4)$$

Основною ознакою для вивчення руху натовпу виступають точки інтересу. Точки інтересу демонструють чіткі сутності, які можна ефективно відстежувати у переповненій сцені.

Для знаходження точок інтересу був обраний Гаррісів виявляч кутів [91], який дозволяє ці точки знаходити. Нехай ϵ зображення у відтінках сірого, яке представлено як $G^2 \rightarrow G$, зміна простору зображення в точці (x, y) показана згорнутою матрицею N_s :

$$N_s = g(\sigma) \times \begin{pmatrix} P_x^2 & P_x P_y \\ P_x P_y & P_y^2 \end{pmatrix}. \quad (3.5)$$

Значення P_x та P_y представляють похідні інтенсивності зображення в напрямках x та y відповідно, $g(\sigma)$ – гаусове ядро з дисперсією σ :

$$g(\sigma) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{x^2+y^2}{2\sigma^2}}. \quad (3.6)$$

Точка інтересу визначається шляхом обчислення позитивних максимумів кутової функції, як показано в наступному рівнянні:

$$POI = \det(N_s) - K \times (tc(N_s))^2. \quad (3.7)$$

де $\det(N_s)$ – визначник N_s ;

$tc(N_s)$ – слід N_s б сума діагональних елементів;

K – емпіричний коефіцієнт у Harris-відгуку;

Оскільки нас цікавить просторово-часовий масштаб, ми слідуємо роботі [92] і узагальнюємо описаний вище кутовий детектор Гарріса на просторово-часову область. Згорнута матриця точок інтересу в просторово-часовій області визначається наступним рівнянням:

$$N_s = g(\sigma) \times \begin{pmatrix} P_x^2 & P_x P_y & P_x P_d \\ P_x P_y & P_y^2 & P_y P_d \\ P_d P_x & P_d P_y & I_d^2 \end{pmatrix}, \quad (3.8)$$

де P_x та P_y представляють похідні інтенсивності зображення за напрямками x та y відповідно;

P_d позначає похідну в часовій області або послідовні кадри в часі.

Просторово-часові точки інтересу вибираються за їхньою потужністю, яка визначається кутовою функцією:

$$POI = \det(N_s) - K \times (tc(N_s))^3. \quad (3.9)$$

де K представляє потужність просторово-часових точок інтересу (безрозмірна константа, зазвичай знаходиться в діапазоні від 0,04 до 0,06).

Чим більше варіація в просторі і часі, тим більше значення k має точка інтересу, як показано в наступному рівнянні:

$$k = \frac{\gamma_1 \gamma_2 \gamma_3}{(\gamma_1 + \gamma_2 + \gamma_3)^3}, \quad (3.10)$$

де γ_1, γ_2 та γ_3 – власні значення N_s , наведені в (3.8). Нас цікавлять лише додатні значення k , і в цьому випадку всі власні значення повинні бути більшими за 0 для просторово-часових точок інтересу.

Після визначення точок інтересу для кожного блоку обчислюється центроїд ключових точок. За допомогою алгоритму відстеження Канаде-Лукаса-Томасі [93] центроїд кожного блоку відстежується на наступних кадрах. Трекер фіксує траєкторії для кожного блоку, що рухається.

Нові об'єкти одночасно з'являються на відео натовпу з плином часу, деякі з них зникають менш ніж через 20 кадрів, а потім з'являються знову, що призводить до появи дуже коротких траєкторій (міні-треки). Для того, щоб врахувати нові особи, які з'являються для відстеження, ми оновлюємо кадри через кожні t кадрів. Таким чином, існуюча область контуру переднього плану і точки інтересу обчислюються заново.

Коли блоки виходять з кадру, вони відкидаються з трекера, а для подальшого відстеження розглядаються нові побудовані блоки. Ми зберігаємо лише ті траєкторії, довжина яких перевищує певний поріг. Покроковий процес виділення траєкторій в області активного контуру описано в Алгоритмі 3.1.

Алгоритм 3.1 – Вилучення міні-треків в області активного контуру

Алгоритм: Виділення міні-треків у зоні активного контуру

Вхідні дані:

- Відеокадри: $I_1, I_2, \dots, I_n$
- Частота оновлення кадрів: $r = 40$
- Порогова довжина міні-треку: $L_T = 30$

Вихідні дані:

- Виділені міні-треки

Для кожного кадру з індексом i від 1 до n :

Якщо $(i \bmod r == 0)$ або $(i == 1)$, тоді:

$C_{in} \leftarrow$ сегментація переднього плану (згідно з рівнянням 3.3)
 Розбити область C_{in} на m блоків: $b_1, b_2, \dots, b_m$ (кожен розміром $b \times b$)

```

Для кожного блоку  $b_j$  у  $C_{in}$ :
    Якщо кількість пікселів у  $b_j$  (в межах  $C_{in}$ )  $\geq$  порога:
         $I_{pts} \leftarrow$  виявлення ключових точок методом Harris
         $Centroid\_j \leftarrow$  обчислення центроїда множини точок  $I_{pts}$ 
        Зберегти  $Centroid\_j$ 
    Інакше:
        Відкинути блок  $b_j$ 
Інакше:
    Для кожного попередньо обчисленого  $Centroid\_k$  ( $k = 1$  до  $t$ ):
         $Minitrack\_k \leftarrow$  відслідковування  $Centroid\_k$  за допомогою
        трекара KLT
        Зберегти  $Minitrack\_k$ 

Для кожного  $Minitrack\_k$ :
     $L_k \leftarrow$  обчислити довжину  $Minitrack\_k$ 
    Якщо  $L_k \geq LT$ :
        Зберегти  $Minitrack\_k$ 
    Інакше:
        Відкинути  $Minitrack\_k$ 

```

3.1.1.3 Модель представлення траєкторій руху натовпу

Для аналізу поведінки натовпу в динамічному середовищі однією з ключових задач є побудова інформативного дескриптора траєкторії, що враховує її геометричні та контекстуальні властивості.

Модель представлення траєкторії використовує чотири характеристики, які включають форму, розташування, напрямок і щільність. Конкатенація цих ознак створює модель представлення траєкторії.

Припустимо, що траєкторія Z має вигляд $Z = f\{(x_p, y_p), (x_{p+1}, y_{p+1}), \dots, (x_k, y_k)\}$, де (x_j, y_j) позначає координати міні-треку на j -му кадрі, де p є початковим, а k – кінцевим. Форма, розташування, напрямок та щільність сусідства траєкторії визначаються наступним чином.

Форма. Ми фіксуємо просторово-часову форму траєкторії, виражаючи її як поліноміальну функцію третього порядку від часу. Ми окремо представляємо її форму у вимірах x та y наступним чином:

$$\begin{aligned} x(t) &= l_0 + l_1 v + l_2 v^2 + l_3 v^3, \\ y(t) &= m_0 + m_1 v + m_2 v^2 + m_3 v^3, \end{aligned} \quad (3.11)$$

де $v \in (1, b), b = e - s$;

e – кінцевий кадр;

s – початковий кадр;

l та m – коефіцієнти кубічних поліномів, які параметризують форму траєкторії окремо за осями.

Функція форми f_s позначає форму траєкторії, яку представлено у вигляді $f_s = [l_0, \dots, l_3, m_0, \dots, m_3]^T$.

Місцезнаходження. Місцезнаходження міні-треку визначається середнім значенням його координат x та y , представленим вектором ознаки місцезнаходження $location = [\bar{x}, \bar{y}]^T$.

Напрямок руху. Ця функція визначає напрямок руху траєкторій у відео. Тут ми визначаємо 8 напрямків від 0 до 360 градусів. Кожна траєкторія зіставляється зі скалярною змінною f_d , яка може приймати значення від 1 до 8 в залежності від її напрямку.

Щільність. Ми виділяємо багатомасштабну характеристику щільності з кожної траєкторії. Траєкторія отримує інформацію від сусідніх траєкторій, що є корисним для розділення траєкторій, які просторово перетинаються. Для траєкторії Z_j обчислюється її щільність в околі радіуса наступним чином:

$$\mu_{j\phi} = |\{Z^i | \forall i \neq j, d(f^j, f^i) < \phi\}|, \quad (3.12)$$

де $f^j = [f_p^j, f_b^j, f_d^j]$ представляє конкатенацію ознак форми, розташування та напрямку. Ми обчислюємо багатомасштабну функцію щільності для трьох значень ϕ для покращення робастності. Наша багатомасштабна функція щільності f_d обчислюється як $f_d = [\mu_{j,\phi_1}, \mu_{j,\phi_2}, \mu_{j,\phi_3}]^T$.

Остаточний вектор ознак F^j траєкторії T_j будується шляхом конкатенації всіх відповідних векторів ознак, витягнутих з траєкторії, що представлено наступним рівнянням:

$$F^j = \begin{bmatrix} f_p^j \\ f_b^j \\ f_m^j \\ f_d^j \end{bmatrix}. \quad (3.13)$$

Аналогічно, всі траєкторії формують свій вектор ознак, як описано в рівнянні вище, а потім відбувається кластеризація траєкторій.

3.1.1.4 Кластеризація траєкторій

Кластеризація траєкторій дає переваги для багатьох задач технічного зору, таких як сегментація руху, виявлення об'єктів, розпізнавання дій і моделювання сцен. Траєкторії часто лежать у низьковимірному просторі ознак. Метою кластеризації траєкторій у нашому підході є сегментація власних патернів у сцені з натовпом людей.

В даному випадку ми виконуємо кластеризацію, використовуючи міри подібності вектору ознак траєкторій F^j . Ми хочемо, щоб траєкторії, які належать до одних і тих же сегментів, були об'єднані в кластери.

Формування кластерів відбувається у два етапи: Спочатку всі вхідні траєкторії розбиваються на множину примітивних кластерів $C = \cup C_n$, де $n = 1 \dots K$, потім примітивні кластери, що належать до одного сегменту потоку, групуються в N кластерів.

Кінцеві кластери траєкторій C_i , отримані на другому етапі, трактуються як сегменти потоків натовпу S_i . Кожен сегмент S_i – це множина траєкторій, подібних

за вектором ознак і бінарним профілем моделей; формально ми ототожнюємо $S_i \equiv C_i$.

Розбиття траєкторій на K примітивних кластерів. Для створення набору примітивних кластерів траєкторій ми використовуємо стандартний алгоритм кластеризації K -середніх [47].

K -середнє вимагає значення параметра K апіорі. Зазвичай значення K є більшим, ніж прогнозована кількість сегментів у відео.

Однак визначення значення K не є критичним, оскільки примітивні кластери, отримані за допомогою алгоритму K -середнього, врешті-решт об'єднуються в кластеризовані траєкторії на основі мінімальної парної відстані.

Спочатку K траєкторій випадковим чином вибираються як початкові центри кластерів, а потім решта траєкторій призначаються до найближчих центроїдів кластерів відповідно до метрики евклідової відстані. K -середнє слідує ітераційній процедурі до тих пір, поки центри кластерів не стабілізуються.

Цей процес сходиться лише до локального мінімуму, оскільки при різних ініціалізаціях можуть виникати різні центри кластерів. Нарешті, відбувається формування примітивних кластерів, які використовуються як будівельні блоки для побудови остаточних власних кластерів.

Об'єднання K -примітивного кластера в N кластерів. Після отримання примітивних кластерів за допомогою кластеризації K -середніх ми випадковим чином вибираємо частину траєкторій з кожного примітивного кластера. Ці траєкторії розглядаються як представники всіх траєкторій, які називаються моделями. Потім кожна траєкторія порівнюється з усіма моделями для отримання набору пріоритетів для кожної траєкторії.

У наборі пріоритетів будується бінарна матриця, в якій зберігається відповідність (1 або 0) кожної траєкторії всім моделям на основі заданого порогу. Подібність моделі між траєкторіями обчислюється за наступною формулою:

$$P_{ji} = \begin{cases} 1, & \text{якщо } |Z_j - M_i| \leq \Delta \\ 0, & \text{в іншому випадку} \end{cases} . \quad (3.14)$$

Потім ми виконуємо агломеративну кластеризацію між матрицею наборів пріоритетів, що представляє попарну відстань між будь-якими двома траєкторіями $d(Z_i; Z_j)$, використовуючи міру відстані Жаккара, як у наступному рівнянні:

$$d(Z_i, Z_j) = \frac{|M_i \cup M_j| - |M_i \cap M_j|}{|M_i \cup M_j|}, \quad (3.15)$$

де M_i та M_j позначають дві множини моделей, що відповідають траєкторіям Z_i та Z_j .

Розглядаючи кожен траєкторію як кластер, ми безперервно об'єднуємо примітивні кластери, які мають мінімальну відстань Жаккара на кожній ітерації [94]. Після об'єднання кластерів з мінімальною відстанню на кожній ітерації ми оновлюємо матрицю пріоритетів і обчислюємо відстань Жаккара для решти кластерів.

Процес повторюється до тих пір, поки найменша відстань між будь-якими двома кластерами не стане більшою за δ . Після зупинки агломерації всі поточні кластери $C_1, \dots, C_N$ розглядаються як остаточні сегменти потоків $S_1, \dots, S_f$ (тобто $S = \{C_i\}$). Таким чином, ми кластеризуємо траєкторії, що належать до одних і тих самих шаблонів. Алгоритм кластеризації траєкторій наведений нижче.

Алгоритм 3.2 – Алгоритм сегментації потоків натовпу

Алгоритм: Сегментація потоків натовпу

Вхідні дані:

- Кластери траєкторій: $C_1, C_2, \dots, C_N$

Вихідні дані:

- Сегменти потоків: $S_1, S_2, \dots, S_f$

Ініціалізація:

- Позначити всі кластери як непереглянуті: $unv(C) = 0$
- $f = 0$ (лічильник сегментів потоків)

Для кожного кластеру траєкторій C_i :

- Знайти сусідні кластери: $N_i \leftarrow N_{bh}(C_i)$
- Збільшити лічильник сегментів: $f \leftarrow f + 1$
- Викликати функцію $ClExp(C_i, Sf, N_i)$
- Позначити кластер C_i як переглянутий

Функція $ClExp(C_i, Sf, N_i)$:

- Додати кластер C_i до сегменту Sf

Для кожного кластеру C_j в N_i :

- Якщо C_j ще не був відвіданий:
 - Знайти його сусідів: $N_j \leftarrow N_{bh}(C_j)$
 - Об'єднати множини сусідів: $N_i \leftarrow N_i \cup N_j$
 - Позначити C_j як відвіданий

Якщо C_j ще не входить до Sf :

- Додати C_j до сегменту Sf

Повернути сегменти: $S_1, S_2, \dots, S_f$

Очевидно, що згаданий метод кластеризації траєкторій перекласифікує заданий набір траєкторій.

Щоб вирішити цю проблему, ми виконуємо другий раунд кластеризації над попередньо кластеризованими траєкторіями.

3.1.1.5 Сегментація потоків натовпу людей

Деякі кластери траєкторій, які складають один сегмент, є структурно різними. Така ситуація виникає, коли люди з однаковим напрямком руху з'являються на сцені в різний час (тобто вони мають траєкторії з різними початковими та кінцевими точками).

Для вирішення цієї проблеми був використаний алгоритм кластеризації на основі щільності [95]. Примітивні кластери використовуються як вхідні дані для власного алгоритму сегментації.

Для початку ми випадковим чином вибираємо кластер C_i у якості початкового кластеру і визначаємо його околицю N_i , яка складається з $C_j (j \neq i)$ кластерів. Кластер C_i та його околиця N_i повинні задовольняти наступним трьом правилам:

- просторове перекриття між кластерами C_i та C_j повинно бути більшим за α ;
- варіація в напрямку потоку між кластерами C_i та C_j не повинна виходити за межі значення β ;
- евклідова відстань між центроїдами (середніми координатами) кластерів C_i та C_j має бути не більшою за γ .

На основі вищезазначеної парадигми обчислюється околиця кластера C_i та об'єднується з усіма кластерами околиці N_i з кластером C_i . Потім ми розширюємо кожен кластер околиці по одному і позначаємо його як відвіданий.

Далі процес повторюється для кожного нового кластера, доки околиця не буде оброблена. Необроблені кластери відновлюються і проходять ту ж саму процедуру. Алгоритм 3.3 пояснює покрокову процедуру остаточної сегментації потоків натовпу людей.

Результати сегментації потоку представлені в піксельній області (рис. 3.2, б). Пікселі траєкторій, які належать одному потоку, об'єднані в одну область. Далі виконується операція ковзного сусідства для згладжування меж сегментів потоків (рис. 3.2, в).

Алгоритм 3.3 – Алгоритм кластеризації траєкторій

Алгоритм: Кластеризація траєкторій

Вхідні дані:

- Набір ознак усіх траєкторій: $F_1, F_2, \dots, F_t$

Вихідні дані:

- Кластери траєкторій: $C_1, C_2, \dots, C_N$

Для кожної траєкторної ознаки F_t з $FeatAll(t)$:

Кластери_K $\leftarrow$ кластеризація методом K-середніх: $ClusterK(F_t, K)$

Моделі $\leftarrow$ вибір моделей на основі К-кластерів: `ModSel(Кластери_K, FeatAll(t))`

$P \leftarrow$ визначення пріоритетних множин моделей: `PSets(Моделі, FeatAll(t))` за рівнянням 3.6

`JacDist` $\leftarrow$ обчислення попарної відстані між елементами P за метрикою Жаккара: `JacDist(P, P, Jaccard)`

Кластери_траєкторій $\leftarrow$ об'єднання траєкторій: `CombTraj(JacDist, P)`
Кінець циклу

Для кожної траєкторії i від 1 до t :

Створити кластер $C_i = [i]$

Поки мінімальне значення у матриці `JacDist` < 1 :

$[i, j] \leftarrow$ знайти позицію мінімального значення у `JacDist`

Об'єднати кластери: $C_i \leftarrow C_i \cup C_j$ (агломеративна кластеризація)

Видалити кластер C_j

Оновити множину $P(i)$

Видалити $P(j)$

Перерахувати матрицю `JacDist`

Повернути: $C_1, C_2, \dots, C_N$

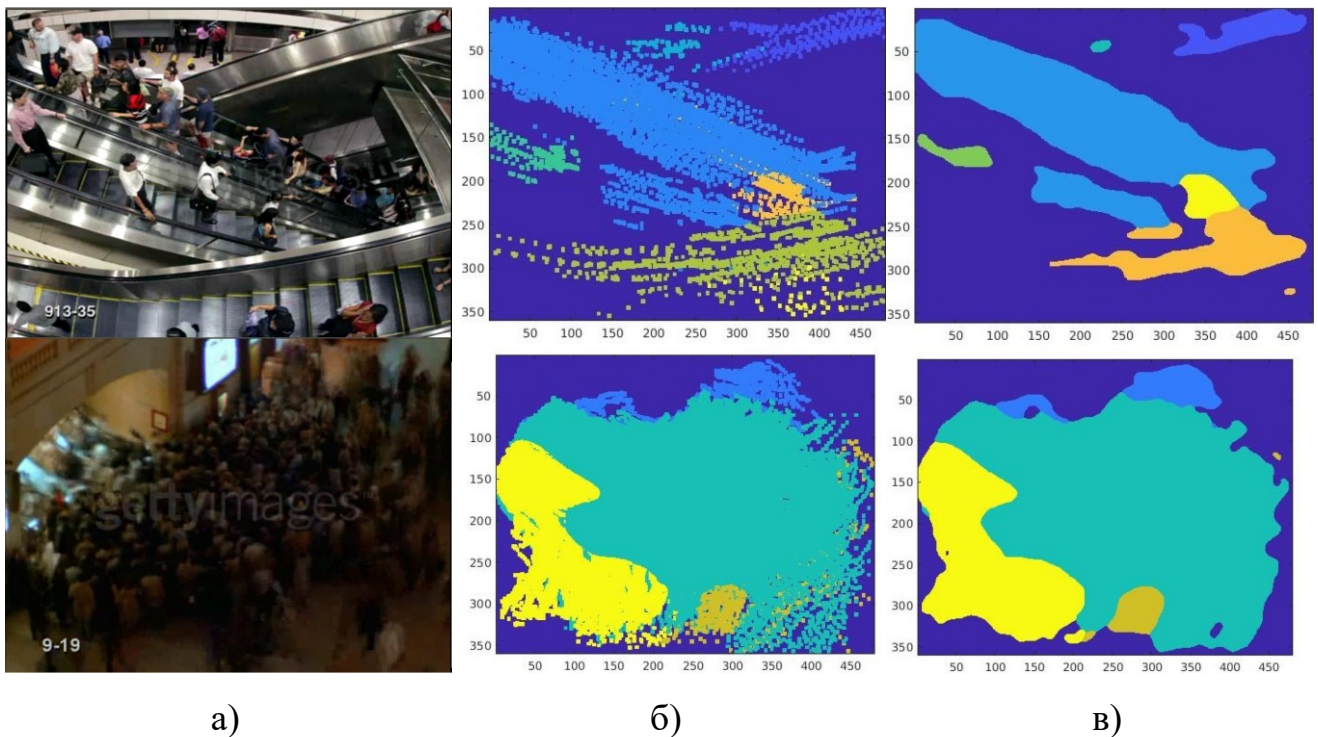


Рисунок 3.2 – Результат роботи алгоритму сегментації потоків натовпу: а) приклади кадрів з відео з натовпу; б) попиксельна сегментація потоків; в) згладжені сегменти потоків.

3.2 Метод виявлення аномалій у сценах з високою щільністю натовпу

Метою виявлення аномальних ситуацій у натовпі є раннє виявлення або прогнозування аномальних подій, які можуть призвести до небезпечних для життя людей умов. Завдяки ранньому виявленню аномальних сцен можна зменшити або запобігти потенційно небезпечним наслідкам.

Однак аналіз сцени натовпу є дуже складним завданням через численні оклюзії між людьми, випадкові зміни щільності натовпу в часі, відеопослідовності з низькою роздільною здатністю, складний фон, а також труднощі з точним моделюванням поведінки натовпу.

Моделі поведінки натовпу, які з'являються часто, називаються нормальними моделями, а ті, що з'являються рідко, вважаються ненормальними або аномальними моделями.

Існуючі підходи до аналізу поведінки натовпу можна умовно поділити на об'єктні та холістичні. Перший підхід розглядає натовп як сукупність індивідів і стикається з проблемами розпізнавання індивідів у сценах з високою щільністю натовпу через велику кількість об'єктів, низьку роздільну здатність, малий розмір об'єктів, сильні оклюзії тощо.

Такі обставини призводять до втрати інформації про цільові об'єкти в сценах великого скупчення людей. Щоб подолати цю проблему, деякі дослідники [96] застосували низькорівневі ознаки та ймовірнісні моделі для аналізу щільного натовпу замість того, щоб зосереджуватися на відстеженні окремих осіб.

Автори роботи [97] дослідили ієрархічну байєсівську модель для сегментації руху на різні види діяльності, використовуючи візуальні ознаки та атомарні види діяльності.

В роботі [96] автори класифікують різну поведінку пішоходів на основі змішаної моделі динамічного пішохода, яка вивчає колективні моделі поведінки пішоходів. Але модель суміші розглядає афінні перетворення і стикається з труднощами в представленні складних форм.

У холістичному підході натовп розглядається як глобальний об'єкт для оцінки поведінки на всій сцені. Ці методи оцінюють динаміку всього натовпу, а не зосереджуються на конкретній активності кожної окремої людини.

Таким чином, ці методи уникають індивідуального виявлення та відстеження людей, а досліджують особливості натовпу для аналізу поведінки всієї сцени. Цей клас зазвичай включає в себе оптичні та польові підходи.

Автори роботи [98] розробили модель натовпу, засновану на напрямку і величині, і запропонували алгоритм сегментації на основі регіонів, який виявляє події натовпу, вивчаючи різні патерни руху.

Автори роботи [99] представили глобальну модель руху за допомогою оптичної гістограми, яка використовується для виявлення ситуацій тисняви, таких як тиснява на параді. Автори роботи [100] розробили імовірнісний фреймворк на основі оптичної гістограми потоків з використанням ріманових просторів для виявлення активності натовпу.

На відміну від існуючих робіт, які залежать від сигналів руху лише між послідовними кадрами, розроблено більш просунуті підходи, які розширюють ці сигнали до більших часових інтервалів, відстежуючи деякі ключові [85] точки або адвекції частинок [101]. Ці методи дають змогу отримати траєкторії, які відображають більш суттєвий часовий рух, що допомагає аналізувати закономірності руху натовпу.

У недавньому минулому деякі дослідники зосередилися на підході адвекції частинок [102]. У цих підходах сітка частинок випадковим чином розподіляється на кадри для представлення індивідів натовпу і адвективно переміщується разом з оптичною хвилюю. Тут частинки нагадують пікселі, а пікселі важко відрізнити від своїх сусідів, що призводить до спотворення міні-треків.

В даному дисертаційному дослідженні відбувається спроба вирішити цю проблему шляхом використання просторово-часових точок інтересу, які потрібно відстежувати.

Відбувається відсіювання траєкторій короткої довжини, щоб не впливати на напрямок постійного натовпу. В даному дослідженні пропонується ентропійний

метод орієнтованих міні-треків для виявлення аномальних сцен. На додаток до ентропії, також відбувається обчислення тимчасового відхилення заповнення, яке вказує на раптові різкі зміни в сцені натовпу.

3.2.1 Загальна ідея роботи запропонованого методу

Запропонований метод направлений на виявлення аномальних сцен натовпу для подій з високою щільністю натовпу. Метод виконує орієнтоване виявлення міні-треків та їх статистичний аналіз в області активного контуру для виявлення аномальних сцен.

Метод складається з чотирьох основних кроків:

- 1) сегментація області активного контуру – відокремлення області переднього плану від усієї відеосцени;
- 2) генерація міні-треків – виявлення просторово-часових блоків точок інтересу та відстеження їх на кадрах;
- 3) обчислення напрямку потоку – обчислення напрямків руху міні-треків та генерація гістограми орієнтованих міні-треків;
- 4) виявлення аномальних сцен – обчислення ентропії та часової зайнятості для орієнтованих міні-треків.

Виявлення аномалій у сцені натовпу передбачає відхилення натовпу від нормальної поведінки з точки зору ентропії та часової зайнятості. Щоб класифікувати сцену як аномальну, вона повинна відповідати наступним двом умовам:

- ентропія сцени більша за поріг прийняття рішення;
- часова зайнятість відхиляється від заданої граничної межі.

Далі видається сигнал тривоги, щоб запобігти катастрофі, пов'язаній з натовпом. Загальна схема роботи методу зображення на рисунку 3.3, детальне пояснення кожного кроку наведено далі.

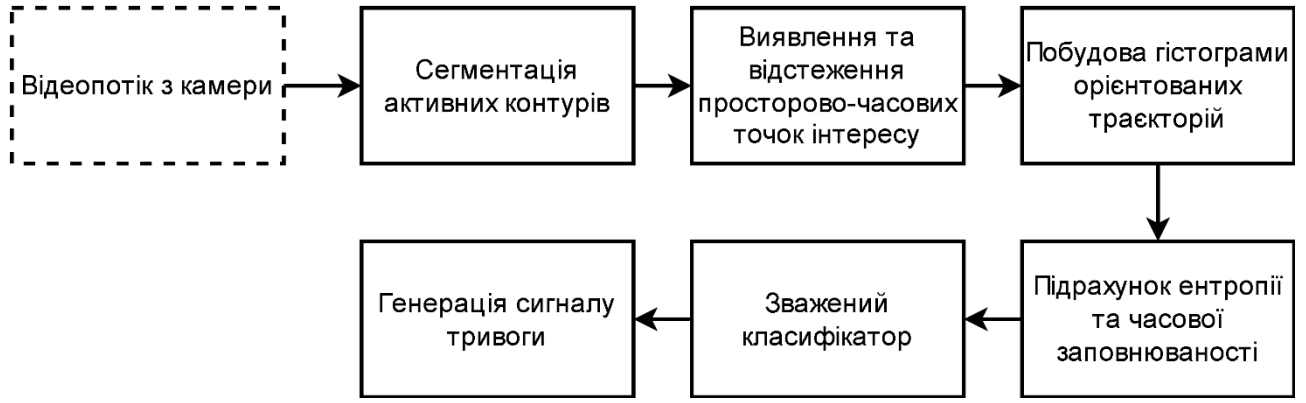


Рисунок 3.3 – Загальна схема роботи методу

3.2.1.1 Сегментація переднього плану

На даному кроці відбувається відокремлення переднього плану та фонових ділянок. Дана процедура відбувається за допомогою методу на основі активних контурів, як було показано раніше в даному дисертаційному дослідженні.

Для сегментації була використана модель CV [90], яка визначає два терми для виділення переднього плану від фонові ділянки:

$$\int_{C_{in}} |P(x) - v_1|^2 dx + \int_{C_{out}}^{F_1(C_{in})+F_2(C_{out})} |P(x) - v_2|^2 dx. \quad (3.16)$$

Перший терм F_1 зменшує контур, а другий F_2 – збільшує контур. Коли контур торкається межі між областями переднього і заднього плану, обидва терми стабілізуються. Тут v_1 і v_2 позначають середні інтенсивності зображення для областей відповідно всередині та поза контуром. Сегментований передній план C_{in} і фон C_{out} контуру C представлено окремо, як показано в наступному рівнянні:

$$C_{in} := \{x \in \Theta: \beta(x) > 0\}, \quad C_{out} := \{x \in \Theta: \beta(x) < 0\}. \quad (3.17)$$

Фонова область C_{out} відкидається, оскільки вона не містить натовпу, і подальша обробка буде виконуватися в області натовпу на передньому плані.

3.2.1.2 Генерація міні-треків

Міні-треки генеруються в просторово-часовій області, використовуючи просторово-часові точки інтересу, як було зазначено раніше в дисертаційному дослідженні. У запропонованому методі генерація міні-треків використовує структуру на рівні блоків у сегментованій області переднього плану.

Відстеження структури на рівні блоків є можливим у натовпі високої щільності, оскільки відстеження окремих осіб у такому натовпі є дуже громіздким процесом. Відстеження точок інтересу також є занадто зашумленим, щоб підтримувати треки. Тому для генерації міні-треків передній план C_{in} розбивається на блоки, що не перетинаються, розміром $b \times b$. Блоки, які складаються з менш ніж $\phi\%$ пікселів області переднього плану, відкидаються. Для кожного виділеного блоку визначається просторово-часові точки інтересу, узагальнюючи роботу [92] в просторово-часовій області. Після виявлення точок інтересу для кожного блоку обчислюється центроїд ключових точок, який потім відстежується на послідовних кадрах за допомогою алгоритму відстеження [93].

Трекер фіксує міні-треки для кожного блоку, що рухається. Повна процедура генерації міні-треків повторюється для кожної кількості кадрів r , щоб врахувати нові особи, які з'являються для відстеження. Метод зберігає лише ті траєкторії, довжина яких перевищує певний поріг LT .

3.2.1.3 Визначення напрямку руху міні-треків

Після видалення міні-треків у часовому вікні з r кадрів відеопослідовності, метод отримує характеристики напрямку руху міні-треків. Ознака напрямку потому визначає орієнтацію міні-треків. Метод розглядає напрямки міні-треків у 8 кутах з діапазоном $[1 : 8]$ кутів. Потім до кожного міні-треку додається змінна f_d , яка може приймати значення від 1 до 8, що відповідає напрямку міні-треку.

Напрямок руху міні-треків обчислюється за допомогою арктангенса, як показано у наступному рівнянні [93]:

$$A = [(\arctan 2(\Delta x, \Delta y)/\pi \times 4 + 4)], \Delta x = \frac{x_{t+r} - x_t}{r}, \Delta y = \frac{y_{t+r} - y_t}{r}, \quad (3.18)$$

де Δx та Δy – градієнти міні-трека.

Домінуючий напрямок потоку у відео можна визначити за медіанним напрямком міні-треків як репрезентативним напрямком, як показано у наступному рівнянні:

$$f_d = \text{median}(A(\text{TrecAll})), \quad (3.19)$$

Далі запропонований метод розподіляє напрямок руху міні-треків по комірках гістограми.

3.2.1.4 Побудова гістограми орієнтованих міні-треків

Гістограма розбивається на 8 рівномірних корзин, всі міні-треки розділяються по корзинам відповідно до напрямку міні-треку. Міні-треки, пов'язані з кутом 1, розподіляються в першій корзині, з кутом 2 – в другій і так далі. Кожна корзина на осі x гістограми складається з орієнтації міні-треків, а вісь y визначає частоту міні-треків при їх відповідній орієнтації. Значення восьми корзин визначають вісім напрямків руху міні-треків, як показано на рисунку 3.4.

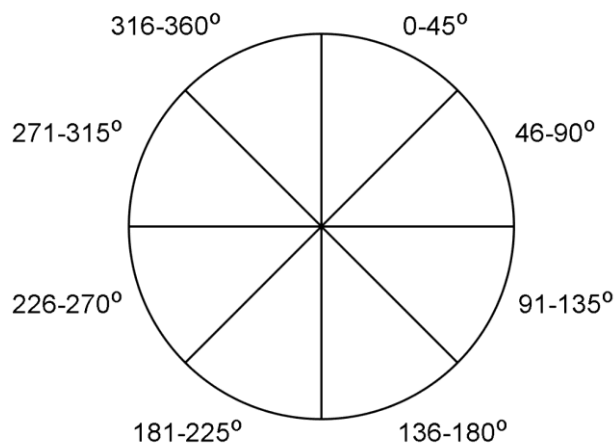


Рисунок 3.4 – Розподіл міні-треків по корзинам

Гістограма орієнтації обчислюється після кожного часового вікна з r кадрів. Далі обчислюється ентропія кожної корзини згідно з наступним рівнянням [93]:

$$\begin{aligned} Entropy_{Bin_k} &= -P(Bin_k) \log_2 P(Bin_k), \\ P(Bin_k) &= \frac{Mg(Bin_k)}{\sum_{k=1}^8 M(Bin_k)} \end{aligned} \quad (3.20)$$

де $Mg(Bin_k)$ та $P(Bin_k)$ – величина та ймовірність міні-треків в корзині k .

Потім все підсумовується для обчислення остаточної ентропії цього часового вікна, як показано в наступному рівнянні:

$$Entropy(all) = \sum_{k=1}^8 E_{Bin_k}. \quad (3.21)$$

3.2.2 Виявлення аномальних сцен

Аномальні дії зазвичай класифікуються як проблема виявлення відхилень. Визначення аномалії залежить від контексту. Наприклад, коли людина біжить у марафоні, це не вважається аномальним, але той самий сценарій стає аномальним на дорозі або в торговому центрі.

У відеопослідовності аномалії можна розділити на дві категорії:

- локальні аномалії;
- глобальні аномалії.

У локальній аномалії невеликі ділянки відео поведуться по-різному порівняно з найближчими сусідами. Наприклад, наявність візка на пішохідній доріжці є локальною аномалією, як показано на рисунку 3.5, а. У випадку глобальної аномалії весь кадр поводить інакше порівняно з наступними за ним кадрами, а не з сусідами в межах кадру. Страшна ситуація тисняви називається

глобальною аномалією, як показано на рисунку 3.5, б. Запропонований метод зосереджується на виявленні глобальних аномалій, тому що в сцені натовпу те, що відбувається, важливіше, ніж те, хто це робить.

В даному дисертаційному дослідженні розглядається виявлення аномальних сцен натовпу за допомогою статистичного аналізу орієнтованих міні-треків натовпу. Для виявлення аномалій використовуються два статистичні параметри: ентропія та часова зайнятість.

Ідея методу полягає в тому, щоб змодельовати нормальну сцену натовпу на основі цих параметрів, і будь-яке відхилення від звичайної поведінки натовпу виявляється як аномальне. Детальний опис цих параметрів наведено далі.



Рисунок 3.5 – Приклади локальних та глобальних аномалій: а) локальна аномалія – візок для гольфу у пішохідній зоні; б) глобальна аномалія – екстремальна ситуація у всій сцені [103]

3.2.2.1 Ентропія орієнтованих міні-треків

Ентропія відображає те, що нечасті події є більш інформативними, ніж часті. Ентропія забезпечує ступінь невизначеності або випадковості. Вищий рівень неупорядкованості або хаосу призводить до вищого рівня ентропії.

Тільки рухомий натовп може призвести до аномальної активності, тому розглядаються тільки рухомі блоки. Орієнтовані міні-треки використовуються для імітації напрямку руху натовпу. Запропонований метод має на меті виявити аномальну поведінку в сценах натовпу відповідно до ентропійної міри орієнтованих міні-треків. Тимчасове вікно вважатиметься аномальним, якщо ентропія його міні-треків перевищує певний поріг.

Для обчислення ентропії орієнтованих міні-треків, кожен міні-трек розподіляється по корзинам гістограми відповідно до їх орієнтації в напрямку потоку. Розподіл ймовірностей міні-треків у кожній корзині обчислюється наступним чином:

$$P_k(\Omega) = \frac{T_k(\Omega)}{T}, \quad k \in [1,8], \quad (3.22)$$

де $T_k(\Omega)$ – кількість орієнтованих міні-треків в комірці k під відповідним кутом;

T – загальна кількість міні-треків у всіх комірках. Ентропія $Entropy(t)$ всіх орієнтованих міні-треків обчислюється наступним чином:

$$Entropy(t) = \sum_{k=1}^8 P_k(\Omega) \log_2 \frac{1}{P_k(\Omega)}, \quad k \in [1,8]. \quad (3.23)$$

де $P_k(\Omega)$ – розподіл ймовірності потрапляння міні-треку в кошик k .

Ентропія отриманих міні-треків вимірюється для кожного часового вікна з r кадрів, і якщо ентропія міні-треків зростає понад певний поріг τ , це означає, що відбувається щось аномальне.

3.2.2.2 Класифікація сцени на базі ентропії

Після обчислення ентропії послідовності кадрів відбувається класифікація шляхом прогнозування класу сцени в реальному часі, мінімізуючи при цьому втрату «нуль-один» (тобто мінімізуючи кількість помилкових класифікацій події та максимізуючи узгодження).

Для бінарної класифікації (нормальна сцена або аномальна сцена) використовується зважений класифікатор KNN , якій використовує міру подібності. Значення ентропії, обчислене в (3.23) є змінною для прогнозування.

Припустимо, що $N_k(x)$ позначає k найближчих сусідів ентропії поточної сцени x і, $\eta(x, x_l)$, $l = 1, \dots, k$ позначає схожість між ентропією поточної сцени x і зразком найближчого сусіда x_l . Для обчислення міри подібності між x та x_l використовується Гауссове ядро:

$$\eta(x, x_l) = \frac{1}{\sqrt{2\pi}\sigma} \exp \left[-\frac{E^2(x, x_l)}{2\sigma^2} \right]. \quad (3.24)$$

де $E(x, x_l)$ – Евклідова відстань між x та x_l . Подібність між x та x_l нормується за формулою:

$$n_l = \frac{\eta(x, x_l)}{\sum_{x_l \in N_k(x)} \eta(x, x_l)}. \quad (3.25)$$

Віднесення події x до класу $C1$ виконується з використанням наступної міри ймовірності:

$$P(y = C1|x) = \sum_{x_l \in N_k(x)} \mathbb{I}_l n_l. \quad (3.26)$$

де $\mathbb{I}_l$ – індикаторна функція, представлена у вигляді:

$$\mathbb{I}_l = \begin{cases} 1, & \text{якщо } y = C1 \\ 0, & \text{в іншому випадку} \end{cases} \quad (3.27)$$

Зважений класифікатор *KNN* відносить подію x до аномальної (тобто до класу $C1$), якщо $P(y = C1|x) \geq \tau$, інакше сцена класифікується як нормальна, де τ використовується як поріг прийняття рішення.

3.2.2.3 Класифікація сцени на основі показника часової зайнятості

Часова зайнятість вимірює площу, зайняту рухомими блоками натовпу з плином часу. Щоб виміряти заповненість сцени, відбувається оцінка або підрахунок кількості міні-треків у сцені. Заповненість прямо пропорційна кількості міні-треків, тобто чим більше міні-треків, тим більшу площу вони займають. Заповненість сцени можна визначити шляхом підрахунку людей, що знаходяться на сцені.

Але звичайні методи виявлення та підрахунку натовпу є громіздкими і менш точними в умовах щільного натовпу через високу щільність, сильну оклюзію, низьку роздільну здатність і тощо.

Метод обчислює міні-треки малих вікон, що відповідають області рухомого натовпу, а потім підраховує їх у відповідній орієнтації.

Відхилення у кількості міні-треків порівняно з попереднім часовим вікном оцінюється наступним чином:

$$DEV = \frac{O(t + t_{int}) - O(t)}{100}, \quad (3.28)$$

$$DEV(t, t + t_{int}) = \begin{cases} true, & \text{якщо } DEV > \Psi \\ false, & \text{в іншому випадку} \end{cases}$$

де (O) – відсоток площі, які зайняті міні-треками протягом часу t_{int} ;

Ψ – поріг прийняття рішення для часової зайнятості.

Велике відхилення у значенні DEV вказує на раптову зміну сцени натовпу. Ми можемо судити про ненормальність сцени, якщо спостерігається масове відхилення часової зайнятості. Цей метод є швидким і показує ефективність порівняно з методами виявлення або оцінки натовпу на основі шаблонів.

Алгоритм 3.4 підсумовує крок за кроком процес виявлення аномальної сцени, розглядаючи обидва методи, згадані вище (тобто, ентропію та відхилення часового заповнення).

3.2.2.4 Вибір порогу прийняття рішення

Для оцінки порогу прийняття рішення для класифікації аномальних сцен використовується метрика ROC. ROC оцінює оптимальний поріг прийняття рішення для заданих еталонних даних (ground truth), який відповідає максимальному узгодженню. Регулювання порогу прийняття рішення використовується для покращення чутливості та специфічності класифікатора за заданим критерієм.

Було проаналізовано три стандартні набори даних для практичної реалізації та були апроксимовані різні пороги прийняття рішень, які задовольняють заданій чутливості та специфічності. Серед експериментальних порогів обрано найбільш підходящий поріг, що відповідає найбільшій площі під кривою.

Алгоритм 3.4 – Алгоритм виявлення аномальних сцен

Алгоритм: Виявлення аномальних сцен

Вхідні дані:

- Треклеті: $t_1, t_2, \dots, t_n$

Вихідні дані:

- Аномальна або нормальна сцена

Для кожного треклета t_i від 1 до n :

Визначити бін (діапазон) для t_i : $B_{in\phi} \in [1; 8]$

Кінець циклу

Для кожної корзини Bin_ϕ від 1 до 8:

$Pt(Bin_\phi) \leftarrow$ відносна ймовірність корзини: $Pt(Bin_\phi) = \frac{Mag(Bin_\phi)}{\sum_{i=1}^8 Mag(Bin_\phi)}$

Обчислити ентропію корзини: $E_{Bin_\phi} = -Pt(Bin_\phi) * \log_2(Pt(Bin_\phi))$

Обчислити загальну ентропію сцени:

$$E = \sum_{i=1}^8 E(Bin_\phi)$$

Кінець циклу

Для кожної сцени з інтервалом кадрів, що відповідає частоті оновлення (r):

Класифікувати сцену на основі умов з рівнянь 3.26 та 3.28:

Якщо $Entropy \leq$ порогового значення ξ

і $TOD(t, t + t_int) = true$, тоді:

Сцена є аномальною

Згенерувати сигнал тривоги

Інакше:

Сцена є нормальною

Кінець циклу

3.3 Висновки до третього розділу

У цьому розділі проаналізовано основні виклики, пов'язані з виявленням структури руху в натовпі високої щільності. Встановлено, що існуючі методи мають обмежену здатність до моделювання просторово-часових залежностей, нестійкі до оклюзій, не враховують складну динаміку взаємодії між людьми та часто залучають до обробки нерелевантні області сцени. Це призводить до зниження точності та збільшення обчислювальних витрат.

Для подолання цих недоліків у роботі запропоновано метод сегментації потоків натовпу на основі кластеризації траєкторій з використанням активних контурів, що дозволяє ефективно виділяти локальні патерни колективного руху та підвищити якість аналізу складних сцен.

Аналіз існуючих методів сегментації руху натовпу засвідчив їх обмежену ефективність у складних сценах. Зокрема, багато підходів не враховують багатовимірні характеристики траєкторій, такі як форма, напрямок, положення та локальна щільність, а також ігнорують часову динаміку та контекст локального

оточення. Крім того, одномасштабний аналіз щільності обмежує точність у змінних просторових умовах.

У відповідь на ці недоліки в роботі запропоновано модель представлення траєкторій, що враховує комплексні ознаки руху та забезпечує більш точну і стійку сегментацію потоків натовпу.

В процесі аналізу викликів, пов'язаних з виявленням аномальних ситуацій у високощільному натовпі було встановлено, що об'єктно-орієнтовані підходи обмежені через складність точного відстеження індивідів у переповнених сценах, а холістичні методи не завжди фіксують локальні порушення у структурі руху. Також виявлено, що традиційні методи, які базуються на короткочасному аналізі руху між кадрами, не дозволяють достовірно описати довготривалі просторово-часові залежності.

З метою подолання цих обмежень у дослідженні запропоновано метод виявлення аномалій, який ґрунтується на аналізі напрямлених міні-треків, сформованих із просторово-часових точок інтересу. Метод включає фільтрацію коротких і шумових траєкторій, а також інтеграцію оцінки варіативності заповнення сцени, що дозволяє виявляти раптові зміни в динаміці натовпу. Такий підхід забезпечує підвищену чутливість до нетипової поведінки в складних сценах спостереження.

Отримано наступні пункти наукової новизни:

- *удосконалено* метод сегментації потоків натовпу з високою щільністю на основі кластеризації траєкторій з використанням методу активних контурів. Використання запропонованого методу дозволяє ефективно виділяти локальні патерни колективного руху та підвищити точність сегментації потоків натовпу у сценах з високою щільністю натовпу;

- *набула подальшого розвитку* модель представлення траєкторії руху в натовпі, що враховує комплексні ознаки руху. Використання даної моделі забезпечує більш точну і стійку сегментацію потоків натовпу;

- *удосконалено* метод виявлення аномальних ситуацій у сценах з натовпом високої щільності, що ґрунтується на аналізі ентропії напрямлених міні-треків,

сформованих із просторово-часових точок інтересу та оцінки часового відхилення заповнюваності сцени. Використання запропонованого методу дозволяє збільшити чутливість систем візуального аналізу натовпу до нетипової поведінки людей в складних сценах, що дозволяє підвищити точність розпізнавання аномальних подій в сценах з високою щільністю натовпу.

4 РЕАЛІЗАЦІЯ ПІДСИСТЕМИ ВІЗУАЛЬНОГО АНАЛІЗУ НАТОВПУ В СКЛАДІ СИСТЕМИ ІНТЕЛЕКТУАЛЬНОГО ВІДЕОСПОСТЕРЕЖЕННЯ

4.1 Мета та задачі підсистеми візуального аналізу натовпу

Підсистема візуального аналізу натовпу дозволяє інтелектуальним системам відеоспостереження бути більш потужними, автоматично аналізуючи в режимі реального часу щільність натовпу і можливі аномальні ситуації. Вона призначена для забезпечення громадської безпеки в щільному середовищі, в тому числі в громадському транспорті, спортивних стадіонах, торгових центрах і під час масових заходів, а також для раннього виявлення незвичайних або небезпечних ситуацій.

Ця підсистема є важливою частиною узагальненої системи відеоспостереження на основі технічного зору, яка використовує методи комп'ютерного зору і машинного навчання для аналізу відеопотоків і отримання корисної інформації про групову діяльність людей. Вона може відстежувати переповненість, аномальні траєкторії руху, які можуть сигналізувати про заворушення, невідкладну медичну допомогу або загрозу безпеці.

Організації громадської безпеки, поліція, організатори заходів, оператори інфраструктури та містобудівники є цільовими клієнтами. Вони покладаються на ці типи технологій для прийняття ефективних рішень, ефективного застосування ресурсів безпеки та своєчасного реагування на інциденти. Крім того, підсистема може бути використана як платформа для розробки, тестування та розгортання нових алгоритмів у сфері штучного інтелекту та розумних міст.

4.2 Визначення потоків інформації підсистеми візуального аналізу натовпу

Інформаційні потоки в інформаційних системах – це схема руху інформації в інформаційній системі. Вони описують спосіб, у який дані збираються,

обробляються, транспортуються та зберігаються від одного кінця системи до іншого. Розуміння та моделювання інформаційних потоків є життєво важливим при побудові інформаційної системи. Вони допомагають визначити межі системи, кордони, області інтеграції та потенційні вузькі місця, що є важливими для продуктивності, масштабованості та надійності системи.

Вхідні інформаційні потоки – це дані, які система отримує з мережі або від користувача. Наприклад, це може бути форма користувача, сигнал датчика або файл, який імпортується в систему. До вхідних інформаційних потоків слід віднести:

- дані з камер відеоспостереження – безперервна одноразова передача зображення в режимі онлайн як основне джерело даних для аналізу поведінки натовпу;

- метадані відео – детальна інформація про джерело відео: ідентифікатор камери, мітка часу, дата та координати розташування. Це дозволяє правильно зрозуміти контекст, в якому відбуваються події;

- налаштування на рівні системи – налаштування оператора або адміністратора, наприклад, зони відстеження, налаштування чутливості, періоди активності. Використовуються для налаштування поведінки системи за певних умов;

- сценарії або правила реагування – формалізм того, коли система повинна вживати заходів (наприклад, тривога, коли щільність натовпу перетинає певний поріг);

- конфігураційні файли – файли налаштувань – структура зони моніторингу, формат відео, правила обробки та порогові значення. Зчитуються для перезавантаження або зміни системи;

- зовнішні API-виклики – виклики REST/ SOAP-інтерфейсу з інших систем або програмних компонентів. Це можуть бути інструкції, умови фільтрації або дані для об'єднання із зовнішньою системою (наприклад, загальноміською системою відеоспостереження).

Вихідні інформаційні потоки – це дані, що виробляються системою або передаються зовнішнім клієнтам. Це може бути згенерований звіт, повідомлення про помилку або результат обробки даних. До вихідних інформаційних потоків слід віднести:

- результати аналізу та звіти про події – дані, отримані системою після обробки відео: відформатовані та завантажені дані про щільність натовпу, виявлені аномалії та рух. Для прийняття подальших рішень або візуалізації;

- негайні сповіщення та попередження – автоматичне надсилання повідомлень у разі виникнення певного сценарію (наприклад, переповнення камер, нестандартна поведінка) операторам або стороннім системам. Повідомлення можуть відображатися в інтерфейсі користувача, надсилатися електронною поштою або надсилатися на інші платформи;

- журнальні та аудиторські записи – машинний запис інформації про дії користувача, реакції системи та результати обробки. Критично важливі для діагностики, аудиту та відстеження;

- візуалізація даних (інформаційні панелі, теплові карти, накладання відео);

- візуалізація результатів аналізу – наприклад, теплові карти розподілу щільності або накладання на відеопотоки. Робить моніторинг та огляд ситуації більш інтуїтивно зрозумілим;

- експорт даних та звіти – файли PDF, CSV або JSON, які система створює з узагальненими результатами аналізу, історичною статистикою або звітами про інциденти. Можуть надаватися стороннім додаткам або організаціям;

- відповіді на запити API від зовнішніх сторін – відформатовані відповіді на запити REST та SOAP API. Це можуть бути бажані дані для аналізу, стан системи або поточні спостереження для відправки в іншу систему.

На рисунку 4.1 наведена діаграма інформаційних потоків підсистеми візуального аналізу натовпу в складі системи інтелектуального відеоспостереження.



Рисунок 4.1 – Діаграма інформаційних потоків підсистеми

4.3 Визначення функціональних вимог підсистеми візуального аналізу натовпу

Функціональні вимоги визначають, що повинна робити підсистема – завантажувати відеопотоки, оцінювати щільність натовпу, виявляти аномальні події тощо. Комплексні функціональні вимоги забезпечують однозначність вимог, щоб підсистема відповідала вимогам місії, забезпечувала точну обробку даних і належним чином взаємодіяла з іншими елементами системи.

Перш ніж визначити функціональні вимоги, треба встановити акторів системи. Актори системи – це точки з інших систем або користувачі, які взаємодіють з підсистемою. Ідентифікація акторів є критично важливим кроком в аналізі вимог, оскільки вони дозволяють встановити межі системи і прояснити ролі, які різні зацікавлені сторони відіграють по відношенню до системи.

В результаті аналізу мети, призначення та інформаційних потоків підсистеми, були визначені наступні актори:

- оператор системи – користувач системи відповідає за визначення регіонів моніторингу, вибір порогових значень для запуску виявлення та перегляд вихідних даних, згенерованих модулем аналізу в реальному часі;

- камери спостереження – надсилають відеопотоки до підсистеми. Вони є основними джерелами, що надаються підсистемі, і повинні бути присутніми, щоб зробити аналіз натовпу можливим;

- платформа міського відеоспостереження – позамережева точка інтеграції, яка отримує тривоги, статуси та аналітичні дані, надіслані з підсистеми через API, для централізованого моніторингу та координації інцидентів;

- група реагування на інциденти – автоматично отримує сповіщення про аномальну поведінку натовпу та вживає заходів на місці відповідно до отриманого сповіщення;

- інженер з обслуговування системи – відповідає за забезпечення функціональної стабільності підсистеми, встановлення програмного забезпечення, ведення журналів продуктивності та усунення технічних несправностей.

Для того, щоб правильно визначити функціональні вимоги до підсистеми візуального аналізу натовпу, необхідно побудувати діаграму варіантів використання. Діаграма варіантів використання – це діаграма UML, яка показує взаємодію акторів з системою та варіанти використання, які роблять цю взаємодію можливою.

На рисунку 4.2 наведена діаграма варіантів використання підсистеми візуального аналізу натовпу в складі системи інтелектуального відеоспостереження.

Розробка діаграми використання підсистеми дозволило провести декомпозицію функціональних сценаріїв підсистеми візуального аналізу натовпу в складі системи інтелектуального відеоспостереження.

ФВ1 Визначення зони, що підлягає моніторингу. Для того, щоб реалізувати описаний вище сценарій, необхідно виконати наступні кроки:

- увійти в систему під обліковим записом оператора через інтерфейс автентифікації, ввівши дійсне ім'я користувача та пароль;

- з головного меню перейти до модуля конфігурації зон спостереження;

- обрати потрібну камеру для спостереження з доступних джерел відеосигналу;

- за допомогою вікна попереднього перегляду створити маску, нарисувавши або позначивши ділянку кадру, яка підлягає моніторингу;

– як варіант, можна позначити декілька зон або частин кадру, які потрібно виключити.

– натиснути «Зберегти», щоб зберегти цю зону моніторингу в конфігурації системи.

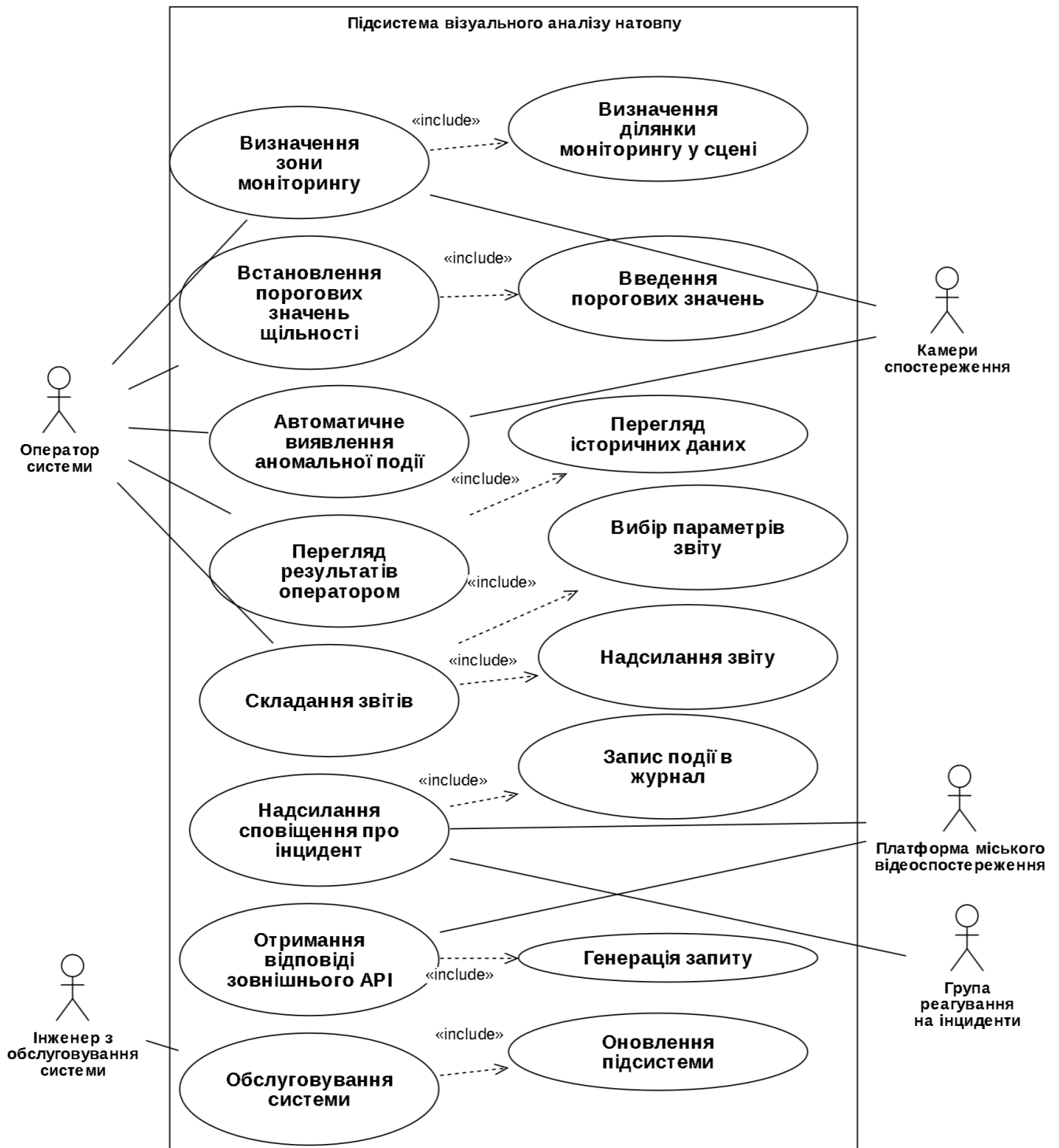


Рисунок 4.2 – Діаграма варіантів використання підсистеми

ФВ2 Встановлення порогових значень щільності натовпу. Для того, щоб реалізувати описаний вище сценарій, необхідно виконати наступні кроки:

- увійти в систему як оператор;
- перейти до налаштувань системи, відкрити вкладку «Параметри аналізу»;
- ввести число для порогу критичної щільності натовпу;
- встановити рівень попереджень в інформаційних повідомленнях менший за критичний;
- активувати прапорець «Активувати автоматичне визначення порогу»;
- натиснути кнопку «ОК», щоб зберегти змінені налаштування.

ФВ3 Автоматичне виявлення аномальної події. Для того, щоб реалізувати описаний вище сценарій, необхідно виконати наступні кроки:

- отримати стабільний відеопотік у реальному часі з підключеної камери спостереження;
- виявити одну або декілька заданих зон моніторингу та обробити кадр;
- провести аналіз руху людей за допомогою вбудованого модуля відстеження та моделі щільності натовпу;
- ідентифікувати аномальну поведінку (біг, розосередження та припинення потоку);
- класифікувати подію як «аномальну»;
- зробити запис про кадр у системному журналі подій із позначкою часу, ідентифікатором камери та зображенням.

ФВ4 Надсилання сповіщення про інцидент. Для того, щоб реалізувати описаний вище сценарій, необхідно виконати наступні кроки:

- при виявленні аномальної ситуації створюється об'єкт оповіщення, що містить усі необхідні метадані;
- до об'єкта додається короткий опис, координати зони впливу та зображення;
- об'єкт сповіщення надсилається на зовнішню платформу міського спостереження;

- підсистема надсилає оповіщення одночасно на інтерфейс оператора та пристрої групи реагування на інциденти (наприклад, планшети, настільні комп'ютери);

- аномальна ситуація записується з позначкою часу та результатом (надіслано чи не надіслано) до системного журналу;

ФВ5 Перегляд результатів оператором. Для того, щоб реалізувати описаний вище сценарій, необхідно виконати наступні кроки:

- з'єднатися з даними облікового запису оператора;
- перейти в головне меню в розділ «Аналітика» / «Історія подій»;
- обрати потрібну камеру та інтервал часу або певну дату;
- переглядати графіки щільності натовпу, теплові карти та список виявлених подій з класифікацією;
- експортувати дані у форматі CSV, PDF або знімка, а також для подальшого звітування або обміну.

ФВ6 Обслуговування системи. Для того, щоб реалізувати описаний вище сценарій, необхідно виконати наступні кроки:

- увійти у обліковий запис інженера з технічного обслуговування;
- відкрити вкладку «Діагностика», відстежити стабільність відеопотоку, перевірити завантаження процесора, наявність системних помилок;
- перевірити журнали на наявність важливих проблем або шаблонів збоїв в обробці;
- випускати оновлення або патчі програмного забезпечення інтерфейсу обслуговування системи;
- записувати результати процесу обслуговування в журнал (дата, час, виконані завдання, стан системи).

ФВ7 Періодичне складання аналітичних звітів. Для того, щоб реалізувати описаний вище сценарій, необхідно виконати наступні кроки:

- увійти в систему як оператор;
- перейти до розділу «Звіти» або «Зведення аналітики»;

- вибрати тип звіту (щоденний, щотижневий або обрати власний період);
- задати часові рамки, камери для розгляду та метрики (наприклад, середня щільність, кількість аномалій, кількість порушень);
- вибрати тип звіту (PDF, CSV або інтерактивна дошка);
- надіслати звіт на електронну пошту або експортувати його через API системи).

ФВ8 Отримання відповіді зовнішнього API. Для того, щоб реалізувати описаний вище сценарій, необхідно виконати наступні кроки:

- зовнішнє джерело надсилає один запит до API підсистеми;
- підсистема аналізує запит, аутентифікує та авторизує, перевіряє права доступу та форму параметрів;
- генерує запит (наприклад, запит аналітики для камери або часу);
- повертає відповідь у запитованому форматі JSON або XML з невеликою кількістю інформації (статистика, журнали, статус);
- повертає відповідь зовнішній системі та записує в журнал подію виконавця API з IP, міткою часу та типом запиту.

4.4 Визначення нефункціональних вимог до підсистеми візуального аналізу натовпу

Нефункціональні вимоги необхідні при проєктуванні підсистеми візуального аналізу натовпу, оскільки саме вони визначають стабільність, продуктивність, надійність, масштабованість і безпеку системи в умовах реального середовища. Чітко визначені та вимірювані нефункціональні вимоги дозволяють уникнути вузьких місць у продуктивності, мінімізувати простой системи та забезпечити відповідність якості системи очікуванням користувачів і зацікавлених сторін (особливо для критичних систем відеоспостереження та громадської безпеки).

До підсистеми візуального аналізу натовпу в складі системи інтелектуального відеоспостереження висуваються такі нефункціональні вимоги.

НФВ1. Система повинна підтримувати щонайменше 32 одночасні відеопотоки з роздільною здатністю 1080p;

НФВ2. Часова затримка між захопленням кадру та виявленням події не повинна перевищувати 750 мс;

НФВ3. Середній час обробки одного кадру не повинен перевищувати 150 мс при використанні GPU-прискорення;

НФВ4. Доступність системи повинна становити 99.7% часу в межах календарного місяця;

НФВ5. Результати будь-якого аналізу мають зберігатися в базі даних щонайменше 30 діб;

НФВ6. Оцінка щільності натовпу має оновлюватися не рідше ніж кожні 2 секунди модулем виявлення аномалій;

НФВ7. Автоматичне перепідключення після втрати відеопотоку має відбуватись не довше ніж за 5 секунд;

НФВ8. Система повинна підтримувати роботу з камерами, що відповідають стандарту ONVIF профілю S;

НФВ9. Встановлення підсистеми у попередньо підготовленому середовищі не повинно тривати більше ніж 30 хвилин;

НФВ10. На сервері з 16 ГБ оперативної пам'яті використання RAM системою не повинно перевищувати 75% при повному навантаженні;

НФВ11. Сервер має бути обладнаний принаймні 8 фізичними ядрами процесора (або 16 логічними потоками) для забезпечення обробки в реальному часі;

НФВ12. Середній час відповіді на API-запити не повинен перевищувати 800 мс при навантаженні до 1000 запитів на годину;

НФВ13. Дані під час передавання мають бути зашифровані за допомогою HTTPS з використанням TLS 1.2 або новішої версії;

НФВ14. Аутентифікація користувача повинна здійснюватися за допомогою токенів, які дійсні не більше ніж 12 годин;

НФВ15. Журнали подій повинні зберігатися щонайменше за останні 60 діб;

НФВ16. Повне резервне копіювання системи має виконуватися автоматично щонайменше раз на 24 години;

НФВ17. Робота системи не повинна погіршуватися при навантаженні до 10 000 зареєстрованих подій на добу;

НФВ18. У разі відмови однієї камери система повинна продовжувати роботу без збоїв (підсистема має бути відмовостійкою).

4.5 Проєктування діаграми логічного уявлення підсистеми візуального аналізу натовпу

На рисунку 4.3 наведена діаграма логічного уявлення підсистеми візуального аналізу натовпу в складі системи інтелектуального відеоспостереження.

Раціональна діаграма є важливим етапом у розробці підсистеми візуального аналізу натовпу і є відображенням структури програмної системи на концептуальному рівні. Вона ілюструє склад системи, спосіб взаємодії її частин, зв'язки між частинами та розподіл обов'язків між модулями. На відміну від фізичних уявлень або моделей розгортання, логічне уявлення стосується не того, як щось робиться (це робить програмне забезпечення), а того, що робиться.

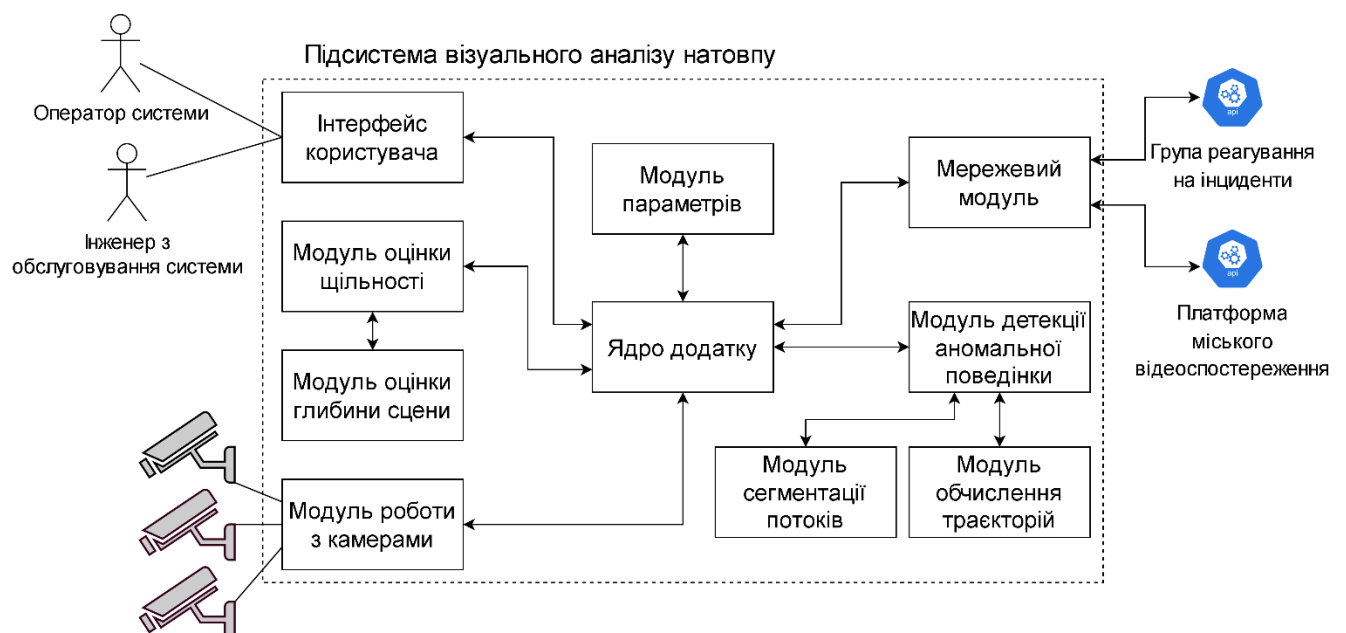


Рисунок 4.3 – Діаграма логічного уявлення підсистеми

4.6 Проектування розгортання підсистеми візуального аналізу натовпу

На рисунку 4.4 наведена діаграма розгортання підсистеми візуального аналізу натовпу в складі системи інтелектуального відеоспостереження. Розробка діаграми розгортання також є важливим елементом проектування підсистеми, щоб показати, як програмне забезпечення системи буде розгорнуто на апаратному забезпеченні. У той час як логічна модель пояснює, як структуровані модулі, діаграма розгортання розглядає їх як фізичні вузли.

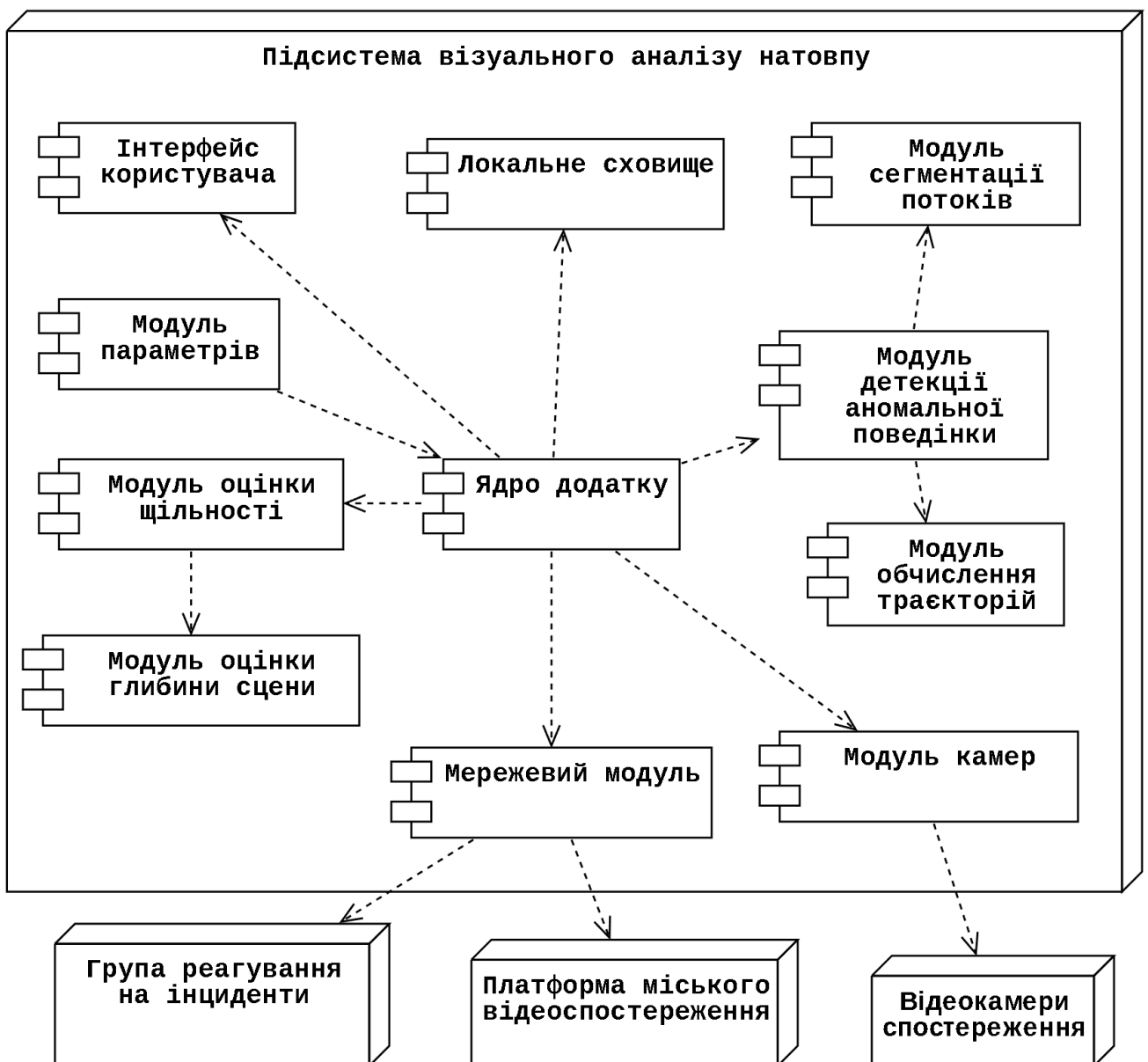


Рисунок 4.4 – Діаграма розгортання підсистеми

4.7 Визначення структури програмного проєкту підсистеми візуального аналізу натовпу

На рисунку 4.5 наведена діаграма пакетів підсистеми візуального аналізу натовпу в складі системи інтелектуального відеоспостереження.

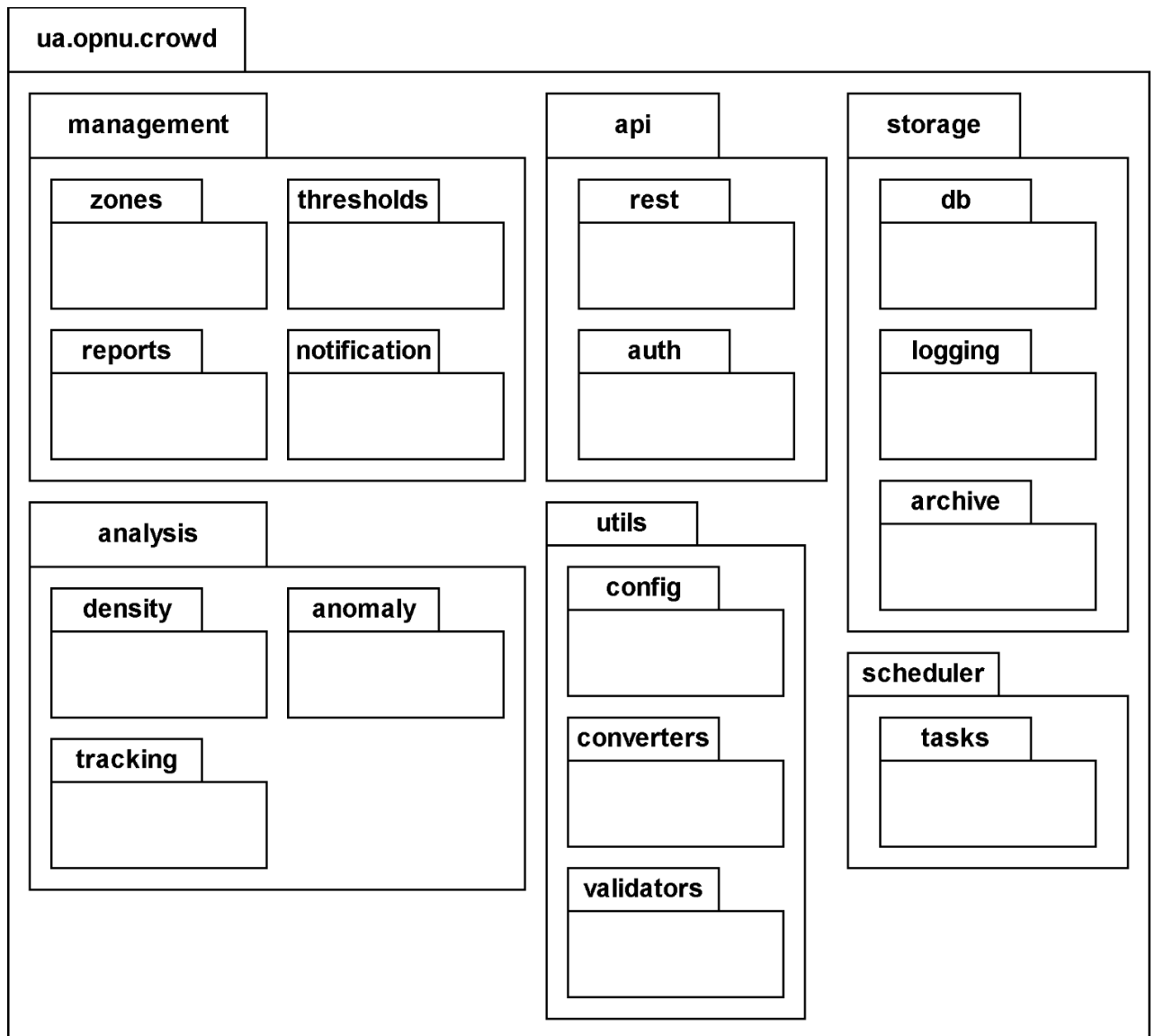


Рисунок 4.5 – Діаграма пакетів підсистеми

Для більш повного розуміння структури програмного проєкту підсистеми візуального аналізу натовпу в складі системи інтелектуального відеоспостереження, необхідно розглянути роль та призначення основних програмних пакетів підсистеми.

– video_input – взаємодіє з камерами спостереження та захоплює відеопотоки в реальному часі. Базується на пакеті SDK від NPC (NetSDKProtocol), містить драйвери, ONVIF та RTSP обгортки, а також нову концепцію "рівня" (внесена нами);

– video_preprocessing – виконує початкову підготовку кадру або масштабування, нормалізацію, фільтрацію шуму, обрізання та застосування масок, щоб зосередити аналіз лише на потрібних ділянках;

– frame_extraction – перевіряє кадри з відеопотоку. Підтримує обмеження частоти кадрів (наприклад, до 1 кадру в секунду) для оптимального завантаження системи;

– analysis – основний модуль обчислювальної аналітики. Координатор між виконанням моделі, збором результатів, послідовністю процесів і зв'язком з модулями визначення щільності натовпу та виявлення аномалій;

– crowd_density – складається з ML-моделей або алгоритмів, що обробляють кадри та обчислюють щільність натовпу;

– anomaly_detection – модуль для виявлення аномалій у подіях активності натовпу шляхом аналізу відмінностей у напрямку, швидкості та щільності. Сумісний з історією руху, відстеженням та евристичними правилами;

– motion_tracking – відповідає за виявлення та відстеження об'єктів (людей) у відеопотоці. Застосовує алгоритми SORT, DeepSORT, фільтр Калмана, оптичний потік;

– management – централізоване управління конфігурацією системи: моніторинг зон, порогові значення, сценарії та політики контролю доступу.

4.8 Розробка діаграми класів підсистеми візуального аналізу натовпу

На рисунку 4.6 наведена діаграма класів концептуального рівня підсистеми візуального аналізу натовпу в складі системи інтелектуального відеоспостереження.

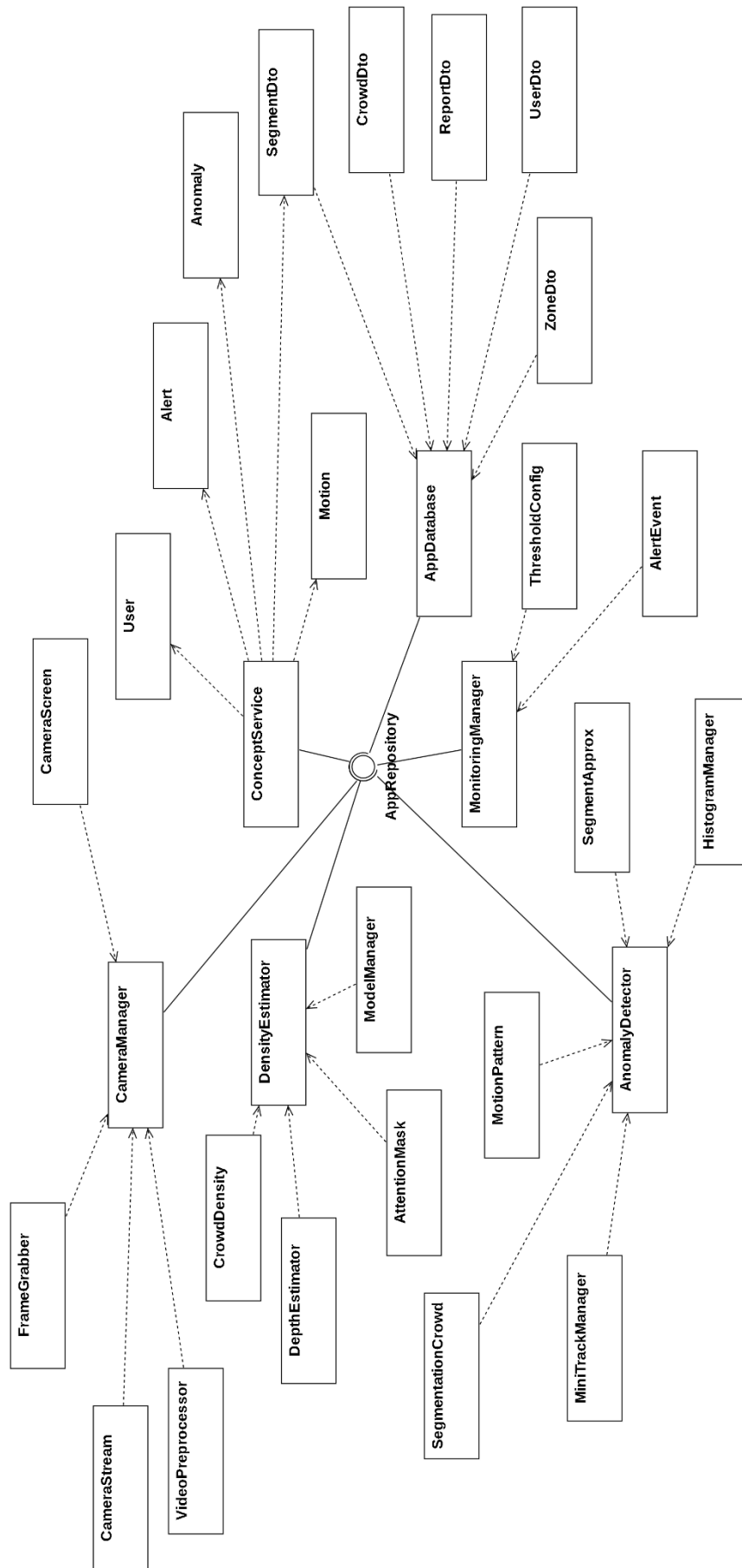


Рисунок 4.6 – Діаграма класів підсистеми

4.9 Управління вихідним кодом підсистеми візуального аналізу натовпу

Використання системи контролю версій GitHub під час створення підсистеми візуального аналізу натовпу забезпечує ефективне управління розробкою, командною співпрацею та відстеженням змін у кодовій базі. GitHub дозволяє зберігати історію змін, автоматизувати CI/CD процеси, створювати ізольовані гілки для різних функцій, а також проводити рев'ю коду та виправлення помилок у контрольованому середовищі.

Для розробки підсистеми використовувались такі практики:

- створення окремих гілок (feature/, bugfix/, release/) для паралельної роботи над компонентами;
- використання pull requests для перевірки та обговорення змін;
- інтеграція з GitHub Actions для запуску автоматичних тестів;
- ведення документації та технічних записів у файлах README.md і Wiki.

В таблиці 4.1 наведені метрики вихідного коду репозиторію підсистеми візуального аналізу натовпу.

Таблиця 4.1 – Метрики вихідного коду репозиторію

Назва метрики	Значення	Опис метрики
Кількість комітів	245	Загальна кількість комітів у репозиторії
Кількість учасників	5	Кількість унікальних розробників, що зробили внески
Кількість форків	12	Кількість форків репозиторію іншими користувачами
Кількість pull requests	30	Кількість pull requests, створених у проєкті
Кількість issues	18	Загальна кількість відкритих і закритих issues
Середній час вирішення issue	3.2	Середній час між створенням та закриттям issue
Середній розмір коміту	45	Середня кількість рядків коду, змінених у комітах
Кількість релізів	4	Кількість створених релізів програмного продукту
Кількість тестових сценаріїв	96	Загальна кількість юніт та інтеграційних тестів у проєкті

4.10 Аналіз вихідного коду підсистеми візуального аналізу натовпу

У процесі розробки програмних систем важливо не лише забезпечити правильну функціональність програмного забезпечення, а й досягти високої якості вихідного коду. Якість коду безпосередньо впливає на продуктивність, підтримуваність, масштабованість, стабільність і безпеку програмного продукту. Одним із практичних і об'єктивних підходів до оцінки якості коду є використання метрик програмного коду –, формалізованих числових показників, які дозволяють проаналізувати різні характеристики програмного забезпечення, зокрема його складність, структуру, модульність, ефективність і надійність.

Метрики дозволяють виявити «вузькі місця» в архітектурі, складні для підтримки компоненти, недокоментовані ділянки або потенційно небезпечні конструкції. Застосування метрик є особливо актуальним на етапах перевірки якості розробки, підготовки до релізу, рефакторингу або інтеграційного тестування. Практичне застосування метрик забезпечує можливість формалізованої оцінки якості без суб'єктивної експертизи.

В таблиці 4.2 наведені значення метрик вихідного коду підсистеми візуального аналізу натовпу.

Отримані показники метрик свідчать про високу якість реалізації підсистеми візуального аналізу натовпу. Обсяг коду (1836 рядків) свідчить про помірну складність проєкту. Кількість коментарів перевищує 28%, що забезпечує хорошу підтримуваність і розуміння коду.

Середня цикломатична складність не перевищує 5, що означає просту й тестовану логіку. Дублікатів коду менше 5%, що підтверджує використання принципу DRY. Довжина функцій у межах 12–15 рядків, а зв'язність між класами –, помірна, що сприяє гнучкій архітектурі.

Класи мають високу когезію, що вказує на дотримання принципу єдиної відповідальності. Тестове покриття перевищує 70%, усі збірки проходять успішно. Вразливостей не виявлено, критично складних функцій майже немає. Загалом, показники демонструють якісну, стабільну та підтримувану реалізацію.

Таблиця 4.2 – Значення метрик вихідного коду підсистеми

Назва метрики	Значення	Опис
Кількість рядків коду (LOC)	1836	Загальна кількість рядків коду, що вказує на помірну складність системи.
Кількість класів	38	Оптимальна модульність: кожен клас виконує чітко визначену функцію.
Кількість методів	142	Методи розбиті на вузькі функціональні блоки, що полегшує тестування і рефакторинг.
Середня кількість рядків на метод	12,9	Методи залишаються компактними і легко підтримуваними.
Співвідношення коментарів до коду (%)	22,5	Добре документований код, що підвищує читабельність та зрозумілість.
Кількість тестових сценаріїв	45	Наявність покриття основних функціональностей системи модульними тестами.
Кількість помилок компіляції	0	Відсутність синтаксичних або компіляційних помилок – свідчення стабільності коду.
Рівень покриття тестами (%)	81,4	Переважає більшість коду покрита тестами, що гарантує надійність поведінки.
Кількість зовнішніх залежностей	11	Використано лише критично необхідні бібліотеки, що зменшує ризики вразливостей.
Кількість виявлених вразливостей	0	Жодної критичної вразливості – код безпечний для розгортання.

4.11 Функціональне тестування підсистеми візуального аналізу натовпу

Актуальність проведення функціонального тестування для розробки програмного модуля полягає в необхідності перевірки правильності реалізації всіх функцій системи згідно з вимогами. Такий вид тестування дозволяє виявити помилки на рівні логіки обробки сценаріїв, забезпечити надійність та передбачуваність поведінки програми у типових ситуаціях, а також підтвердити відповідність очікуваному функціоналу для користувача.

Для проведення функціонального тестування були розроблені наступні функціональні тести.

ТВ1 Вхід в систему під обліковим записом оператора. Для проведення тестування необхідно виконати наступні кроки:

- ввести правильне ім'я користувача у поле логіну;
- ввести дійсний пароль;
- натиснути кнопку входу;
- дочекатися відповіді від серверу;
- перевірити, що користувач перенаправлений до головної сторінки системи.

ТВ2 Створення нової зони моніторингу. Для проведення тестування необхідно виконати наступні кроки:

- перейти до модуля «Конфігурація зон спостереження»;
- вибрати камеру з переліку доступних джерел;
- намалювати область спостереження у вікні попереднього перегляду;
- вказати назву нової зони;
- натиснути кнопку «Зберегти» та перевірити, що зона збереглась.

ТВ3 Редагування зони моніторингу. Для проведення тестування необхідно виконати наступні кроки:

- відкрити існуючу зону спостереження;
- внести зміни до форми або розміру маски;
- змінити назву або параметри;
- зберегти зміни;
- перевірити, що оновлена зона правильно відображається.

ТВ4 Встановлення порогу щільності натовпу. Для проведення тестування необхідно виконати наступні кроки:

- перейти до вкладки «Параметри аналізу»;
- встановити значення критичного порогу;
- задати значення для попереджувального рівня;
- встановити чекбокс для автоматичної адаптації;
- натиснути кнопку «ОК» та перевірити збереження параметрів.

ТВ5 Ініціація відеопотоку з камери. Для проведення тестування необхідно виконати наступні кроки:

- вибрати камеру в інтерфейсі;
- перевірити з'єднання з джерелом;
- ініціювати передачу потоку;
- переконатися в появі зображення у вікні перегляду;
- відстежити стабільність відеопотоку протягом 10 секунд.

ТВ6 Виявлення аномальної події. Для проведення тестування необхідно виконати наступні кроки:

- забезпечити наявність активної камери та зони моніторингу;
- імітувати зміну швидкості/напрямку руху людей;
- дочекатися реакції модуля виявлення аномалій;
- перевірити наявність запису події у журналі;
- перевірити класифікацію події як «аномальної».

ТВ7 Надсилання оповіщення про інцидент. Для проведення тестування необхідно виконати наступні кроки:

- дочекатися фіксації аномалії;
- перевірити створення об'єкта оповіщення;
- переглянути метадані (координати, опис, зображення);
- переконатися в надсиланні на зовнішню платформу;
- перевірити статус у журналі – має бути «надіслано».

ТВ8 Перегляд історії подій. Для проведення тестування необхідно виконати наступні кроки:

- увійти під обліковим записом оператора;
- перейти у вкладку «Історія подій»;
- вибрати часовий інтервал і камеру;
- перевірити відображення переліку подій;
- відкрити запис події та переглянути її деталі.

ТВ9 Генерація аналітичного звіту. Для проведення тестування необхідно виконати наступні кроки:

- перейти до вкладки «Аналітика» → «Звіти»;
- вибрати період звітності;
- вказати камери та метрики;
- обрати формат звіту (PDF/CSV);
- перевірити, що звіт згенеровано і доступний для завантаження.

ТВ10 Відповідь на запит від API. Для проведення тестування необхідно виконати наступні кроки:

- від імені сторонньої системи надіслати запит до API;
- перевірити авторизацію і формат параметрів;
- відстежити відповідь у форматі JSON або XML;
- перевірити коректність структури відповіді;
- знайти запис запиту в журналі подій.

ТВ11 Відображення теплової карти. Для проведення тестування необхідно виконати наступні кроки:

- відкрити модуль перегляду історії;
- вибрати конкретну дату та камеру;
- встановити тип візуалізації – «теплова карта»;
- перевірити відображення кольорових зон;
- зіставити колірні маркери з даними щільності.

ТВ12 Вихід користувача з системи. Для проведення тестування необхідно виконати наступні кроки:

- натиснути кнопку «Вийти» у головному меню;
- перевірити завершення сесії;
- спробувати відкрити внутрішні сторінки без авторизації;
- переконатися, що система повертає на екран входу;
- перевірити запис про вихід у журналі дій користувача.

ТВ13 Обробка втрати з'єднання з камерою. Для проведення тестування необхідно виконати наступні кроки:

- імітувати втрату сигналу з камери;
- зачекати на реакцію системи;
- перевірити повідомлення про помилку;
- переконатися в записі інциденту в журналі;
- перевірити автоматичне відновлення з'єднання.

ТВ14 Встановлення нової версії ПЗ. Для проведення тестування необхідно виконати наступні кроки:

- увійти як інженер з обслуговування;
- перейти до розділу «Оновлення»;
- завантажити останню версію;
- запустити інсталяцію;
- перевірити коректність функціонування системи після оновлення.

ТВ15 Автоматичне очищення архіву подій. Для проведення тестування необхідно виконати наступні кроки:

- встановити політику зберігання подій на 30 днів;
- запустити систему з фіктивними подіями старше 30 днів;
- перевірити, що старі події видалені;
- переглянути журнал очищення;
- переконатися, що актуальні події залишилися.

ТВ16 Обробка кількох камер одночасно. Для проведення тестування необхідно виконати наступні кроки:

- підключити щонайменше 3 камери;
- визначити зони для кожної з них;
- переконатися в стабільності потоку з усіх камер;
- перевірити реєстрацію аномалій паралельно;
- переглянути зведену аналітику по всіх потоках.

ТВ17 Відмова бази даних. Для проведення тестування необхідно виконати наступні кроки:

- імітувати втрату доступу до БД;
- виконати спробу запису події;
- переконатися у спрацьовуванні механізму черги/повтору;
- відновити підключення до БД;
- перевірити, що всі дані з черги збережено.

ТВ18 Авторизація користувача з невірними даними. Для проведення тестування необхідно виконати наступні кроки:

- ввести неправильний логін;
- ввести неправильний пароль;
- натиснути кнопку входу;
- перевірити, що доступ не надано;
- перевірити, що згенеровано повідомлення про помилку.

ТВ19 Вивантаження даних у форматі CSV. Для проведення тестування необхідно виконати наступні кроки:

- відкрити вкладку «Історія»;
- обрати події за останній тиждень;
- натиснути кнопку «Експорт CSV»;
- відкрити завантажений файл;
- перевірити наявність даних та правильність колонок.

ТВ20 Перевірка прав доступу. Для проведення тестування необхідно виконати наступні кроки:

- увійти як користувач без прав адміністратора;
- спробувати відкрити розділ «Конфігурація»;
- перевірити заборону доступу;
- переглянути повідомлення про обмеження;
- переконатися, що спроба доступу зафіксована в журналі.

В таблиці 4.3 наведений протокол результатів функціонального тестування підсистеми візуального аналізу натовпу.

Таблиця 4.3 – Протокол результатів функціонального тестування

Тест	Очікуваний результат	Фактичний результат	Результат
ТВ1	Користувач успішно входить у систему та перенаправляється на головну сторінку.	Користувач був успішно авторизований і перенаправлений до головної сторінки.	Успішно
ТВ2	Нова зона моніторингу створюється та відображається у списку.	Зона моніторингу збережена, її назва відображається у списку зон.	Успішно
ТВ3	Зміни до зони збережені та коректно відображаються.	Маска змінена, назва оновлена, зміни видно у графічному вікні.	Успішно
ТВ4	Нові порогові значення збережені у конфігурації системи.	Система прийняла зміни, відображаються нові параметри.	Успішно
ТВ5	Зображення з камери з'являється, відеопотік стабільний протягом 10 секунд.	Зображення з'явилося, з'єднання не переривалось впродовж 10 секунд.	Успішно
ТВ6	Система розпізнає зміну поведінки натовпу та фіксує аномалію в журналі.	Система виявила біг натовпу та позначила кадр як аномалію.	Успішно
ТВ7	Оповіщення надсилається до платформи моніторингу, статус 'надіслано'.	Повідомлення було надіслано з усіма необхідними метаданими.	Успішно
ТВ8	Інтерфейс показує повний список подій з деталями за вибраний період.	Всі події за обраний день відображаються, є доступ до деталей.	Успішно
ТВ9	Звіт з заданими параметрами успішно формується у PDF та доступний для завантаження.	PDF-файл створено, збережено на диск, містить потрібні метрики.	Успішно
ТВ10	Сторонній API отримує валідну відповідь, подія фіксується в журналі.	JSON-відповідь містить статистику, журнал містить запис про запит.	Успішно

Продовження таблиці 4.3

Тест	Очікуваний результат	Фактичний результат	Результат
ТВ11	Теплова карта генерується відповідно до щільності, кольори відображають інтенсивність.	Карта відображає інтенсивність у кольорах, дані відповідають обраній даті.	Успішно
ТВ12	Сесія завершується, спроба доступу блокується, подія записується в журнал.	Після натискання 'Вийти' користувача перенаправлено на сторінку входу.	Успішно
ТВ13	Система виявляє втрату сигналу, показує повідомлення, пробує відновити з'єднання.	Після втрати сигналу з'явилось попередження, зв'язок було відновлено.	Успішно
ТВ14	Оновлення проходить успішно, всі функції системи залишаються стабільними.	Після інсталяції оновлення система працює стабільно.	Успішно
ТВ15	Старі події автоматично видаляються, актуальні залишаються.	Архів очищено від подій старше 30 днів, інші залишилися.	Успішно
ТВ16	Система обробляє відео з кількох камер одночасно без втрати стабільності.	Усі три камери транслюють стабільно, відображаються в аналітиці.	Успішно
ТВ17	Механізм черги зберігає події під час втрати БД, дані успішно записуються після відновлення.	Система зберегла події після відновлення підключення до БД.	Успішно
ТВ18	Авторизація не відбувається, виводиться повідомлення про помилку.	Вхід заблоковано, повідомлення про помилку з'явилось.	Успішно
ТВ19	Файл CSV містить повні дані за обраний період з правильними колонками.	Файл CSV відкрився коректно, структура таблиці збережена.	Успішно
ТВ20	Користувач отримує повідомлення про заборону доступу, подія записується в журналі.	Повідомлення про відмову доступу показано, запис є в журналі.	Успішно

4.12 Висновки до четвертого розділу

У четвертому розділі дисертаційного дослідження було комплексно реалізовано підсистему візуального аналізу натовпу в складі системи інтелектуального відеоспостереження. Сформульовано мету, основні задачі та призначення цієї підсистеми в контексті забезпечення громадської безпеки, а також визначено її роль у виявленні критичних ситуацій у місцях масового скупчення людей.

Було побудовано модель інформаційних потоків підсистеми, що деталізує вхідні та вихідні дані, типи взаємодій з іншими компонентами системи та зовнішніми клієнтами. Визначено функціональні вимоги на основі сценаріїв використання, у тому числі варіанти роботи оператора, реагування на аномалії, перегляд аналітики та технічне обслуговування.

Сформульовано чіткі нефункціональні вимоги до системи, які охоплюють аспекти продуктивності, надійності, безпеки, масштабованості та сумісності з обладнанням. Розроблено структурні моделі підсистеми – діаграму логічного уявлення, діаграму розгортання, діаграму пакетів та діаграму класів, що дозволило формалізувати архітектуру рішення на концептуальному рівні.

Окрему увагу приділено організації процесу розробки та управлінню вихідним кодом за допомогою системи GitHub. Наведено відповідні метрики активності розробки, оцінено якість коду на основі формалізованих показників. Аналіз показав відповідність коду сучасним вимогам підтримуваності, модульності та безпеки.

Проведено функціональне тестування підсистеми, під час якого було реалізовано 20 ключових тестових сценаріїв. Усі функціональні вимоги підтверджено – тести пройдено успішно, що демонструє стабільність реалізованого рішення та готовність підсистеми до використання в реальному середовищі.

5 ДОСЛІДЖЕННЯ ПІДСИСТЕМИ ВІЗУАЛЬНОГО АНАЛІЗУ НАТОВПУ В СКЛАДІ СИСТЕМИ ІНТЕЛЕКТУАЛЬНОГО ВІДЕОСПОСТЕРЕЖЕННЯ

5.1 Дослідження методу сегментації потоків натовпу в сценах з високою щільністю натовпу

5.1.1 Тестове оточення

Тестове оточення реалізовано за допомогою мови Python з додатковими бібліотеками з використанням графічної відеокарти Nvidia RTX 3060 8 Gb. Початковий темп навчання встановлений у значення 10^{-5} та ділиться на 10 у випадку, коли функція втрат знаходить «плато». У якості заходів із аугментації навчальної вибірки використовувалися частини зображень, які випадковим чином вирізалися із зображень навчальної вибірки, а також деякі зображення були перевернуті. Остаточне значення кількості людей на зображенні обчислювалося шляхом підсумування значень щільності по всьому зображенню. У якості метрик оцінки якості роботи запропонованих моделей були обрані метрики, які дуже часто зустрічаються в роботах, присвячених вирішенню задачі підрахунку людей у натовпі – середня абсолютна похибка (Mean Absolute Error, MAE) [82] та середня квадратична похибка (Mean Squared Error, MSE) [83]. У випадку наявності M екземплярів тестових даних, значення MAE обчислюється як:

$$MAE = \frac{1}{M} \sum_{i=1}^M |\hat{C}^i - C^i|, \quad (5.1)$$

де $\hat{C}_i$ та C_i – прогнозоване та фактичне значення кількості людей для i -го зображення.

Для метрики MSE, яка виміряє надійність прогнозованої оцінки кількості людей, у випадку наявності M екземплярів тестових даних, значення MSE обчислюється як:

$$MSE = \sqrt{\frac{1}{M} \sum_{i=1}^M (\hat{C}^i - C^i)}. \quad (5.2)$$

5.1.2 Використані набори даних

Експериментальне дослідження запропонованого методу проводилося на чотирьох наборах даних: ShanghaiTech [9], Mall [104], UCF_CC_50 [105] та WorldExpo [76]. Ці набори даних часто використовуються в різних роботах присвячених темі підрахунку людей у натовпі [106-108] завдяки їх різноманітності та здатності оцінити різні методи підрахунку. На рисунку 5.1 наведені приклади зображень з вищезазначених наборів даних.



Рисунок 5.1 – Зразки зображень з чотирьох наборів даних

Розглянемо детальніше набір даних ShanghaiTech [9]. Це великий набір даних реальних вуличних сцен. Набір розділений на дві частини з різною скупченістю людей у натовпі. Перша частина (ShanghaiTech-part A) містить 482 зображення, які зібрані в мережі Інтернет, 300 з яких використовуються для навчання та 182 для тестування моделей. Зображення в даній частини містять багато людей з великими перешкодами у сценах. Друга частина (ShanghaiTech-part B) містить 716 анотованих зображення, які були отримані з різних камер відеоспостереження та містить різні сцени з натовпом. Кожне зображення має суттєві викривлення, викликані перспективою, тому на одному зображенні можуть бути присутні люди з великими змінами у фізичному розмірі на зображенні. В процесі дослідження запропонованого методу були використані параметри, які рекомендовані в оригінальній роботі, в якій було представлено набір даних. Зокрема, для другої

частини було використано 400 зображень для навчання та 316 зображень для тестування роботи нейронних мережей. Для тренування моделей було випадковим чином отримано 20 ділянок зображення розміром 224x224 пікселі.

Розглянемо набір даних WorldExpo. Даний набір даних складається з 1132 анотованих відеопослідовностей, які були отримані за допомогою 108 камер відеоспостереження. З цих 1132 послідовностей було обрано 3980 кадрів, які були розмічені точковими анотаціями (кожна точка розміщується по центру голови людини). Із розмічених 3980 кадрів, 3380 кадрів з 103 сцен використовуються як дані для навчання та 600 кадрів із 5 інших сцен використовуються для тестування. Для кожної сцени надані дані сегменту на зображенні, де відбувається скупчення людей, а також зображення перспективи. Як і у випадку попереднього набору даних, для навчання моделей з кожного зображення випадковим чином були обрані 20 ділянок розміром 224x224 пікселі. Дані сегменту із скупченням людей були використані як бінарна маска – дані про кількість людей враховувалися тільки якщо люди знаходилися у зазначеному сегменті.

Розглянемо набір UCF_CC_50. Даний набір містить 50 зображень, які були зібрані та анотовані із сцен, знайдених в мережі Інтернет. Особливістю даного набору є велика різниця між кількістю людей на зображенні – дані варіюються від 94 до 4543 людей на зображенні. Додатково, набір є невеликим за об'ємом, що суттєво ускладнює процес навчання моделей. Для навчання моделей була обрана стратегія з п'ятикратною перехресною перевіркою для оцінки роботи моделей.

Розглянемо набір Mall. Даний набір даних складається з 2000 зображень, які були зібрані в торговельному центрі, що означає, що даний набір містить сцени в приміщенні.

Ще однією особливістю даного набору є невелика кількість людей у сценах – від одного до шести десятків. Перевагою цього набору є наявність викривлень із-за перспективи та зміни в освітленні, що дає велику варіативність у розмірі та позах людей на зображеннях.

В процесі використання даного набору, були використані стандартні рекомендації – перші 800 зображень були використані для навчання, 1200

зображень були використані для тестування. З кожного зображення були випадковим чином обрані 12 ділянок розміром 160x160 пікселів.

5.1.3 Порівняння запропонованого методу з аналогами

Для підтвердження ефективності запропонованого методу, було проведено його порівняння з методами-аналогами на різних наборах даних (табл. 5.1 – 5.4). Так як датасет Mall має в своєму складі лише одну сцену, а також містить кілька зображень, було запропоновано використати мережу SAWNet-S з двома модулями інтеграції інформації про глибину, які впроваджуються на етапах 1 і 2, щоб порівняти продуктивність на цьому наборі даних (див. розділ 2.6). Для інших датасетів для порівняння результатів використовується SAWNet-D з одним модулем інтеграції інформації про глибину сцени, впровадженням на етапі 1 (див. розділ 2.6). Розглянемо результати дослідження на наборі даних ShanghaiTech-B. Як видно з таблиці 5.1, на наборі даних ShanghaiTech-B, запропонований метод перевершує інші розглянуті підходи.

Таблиця 5.1 – Порівняння результатів для набору даних ShanghaiTech-B

Метод	MAE	MSE
MATT [109]	21,5	33,6
SANet [110]	19,9	29,9
DeepCount [111]	20,9	29,6
PCCNet [112]	16,9	27,5
MFCC [113]	13,4	21,0
TAFNet [114]	8,9	14,2
SAWNet-D без модуля	10,4	15,8
SAWNet-D	8,1	12,9

Зокрема, порівняно з методами, які обробляють варіації масштабу, в основному використовуючи ознаки різного масштабу [111-113], мережа SAWNet-

D досягає кращої продуктивності, що демонструє ефективність використання глибини для явного моделювання варіацій масштабу для підрахунку людей у натовпі натовпу.

Розглянемо результати дослідження на наборі даних WorldExpo. У таблиці 5.2 наведено порівняння MAE з іншими методами на кожній з тестових сцен WorldExpo, а також середнє значення MAE для всіх сцен.

Таблиця 5.2 – Порівняння результатів для набору даних WorldExpo

Метод	Сцена 1	Сцена 2	Сцена 3	Сцена 4	Сцена 5	Середнє
MATT [109]	12,0	60,8	37,8	23,0	24,7	31,66
SANet [110]	11,6	60,0	36,9	21,9	23,6	30,8
DeepCount [111]	9,9	13,9	14,4	22,4	3,9	12,9
PCCNet [112]	3,5	20,8	13,0	12,9	7,9	11,62
MFCC [113]	4,5	15,6	9,8	11,1	6,1	9,42
TAFNet [114]	2,8	14,9	10,6	10,3	5,7	8,86
SAWNet-D б/м	3,6	15,2	12,2	11,7	5,2	9,58
SAWNet-D	2,1	15,2	9,1	9,5	3,3	7,84

З даних таблиці видно, що мережа SAWNet-D перевершує попередні методи із середнім значенням MAE 7,84, що демонструє ефективність запропонованого методу для підрахунку натовпу на різних сценах. Слід зазначити, що для сцени номер 2, модель SAWNet-D показує гірші результати. Даний факт зумовлений тим, що на даній сцені ділянка із скупченням людей знаходиться майже безпосередньо під камерою спостереження.

У цій ситуації спотворення перспективи та зміна масштабу не є основними факторами, що впливають на точність підрахунку, що, таким чином, обмежує ефективність запропонованого методу для обробки змін масштабу, представлених у загальних сценах спостереження.

Розглянемо результати дослідження на наборі даних UCF_CC_50. Як видно з таблиці 5.3, запропонований метод досягає найкращого значення MAE порівняно з іншими методами.

Таблиця 5.3 – Порівняння результатів для набору даних UCF_CC_50

Метод	MAE	MSE
MATT [109]	377,8	508,9
SANet [110]	361,5	493,6
DeepCount [111]	333,9	425,5
PCCNet [112]	317,9	440,3
MFCC [113]	295,9	321,1
TAFNet [114]	271,8	390,9
SAWNet-D без модуля	285,9	397,7
SAWNet-D	253,5	364,5

Такий результат говорить про ефективність запропонованого підходу на екстремально щільних сценах.

Розглянемо результати дослідження на наборі даних Mall. Як видно з таблиці 5.4, при додаванні двох модулів вбудовування глибини в базову модель, продуктивність покращується і стає порівнянною з іншими методами, що демонструє ефективність і робастність запропонованого підходу на невеликих наборах даних з меншою кількістю людей.

Метод TAFNet [114] додатково об'єднує результати на основі виявлення для коригування регресійної оцінки щільності, що особливо корисно для розріджених сцен, таких як набір даних Mall, і таким чином досягає найкращих результатів порівняно з методами на основі чистої регресії.

Таблиця 5.4 – Порівняння результатів для набору даних Mall

Метод	MAE	MSE
MATT [109]	3,57	18,5
SANet [110]	3,16	15,8
DeepCount [111]	2,42	9,14
PCCNet [112]	1,84	2,72
MFCC [113]	2,76	2,39
TAFNet [114]	2,06	2,57
SAWNet-S без модуля	3,15	3,89
SAWNet-S	2,10	2,65

На рисунку 5.2 якісно візуалізовано і порівняно карти щільності та оцінені кількості людей у натовпі з модулем вбудовування глибини і без нього.

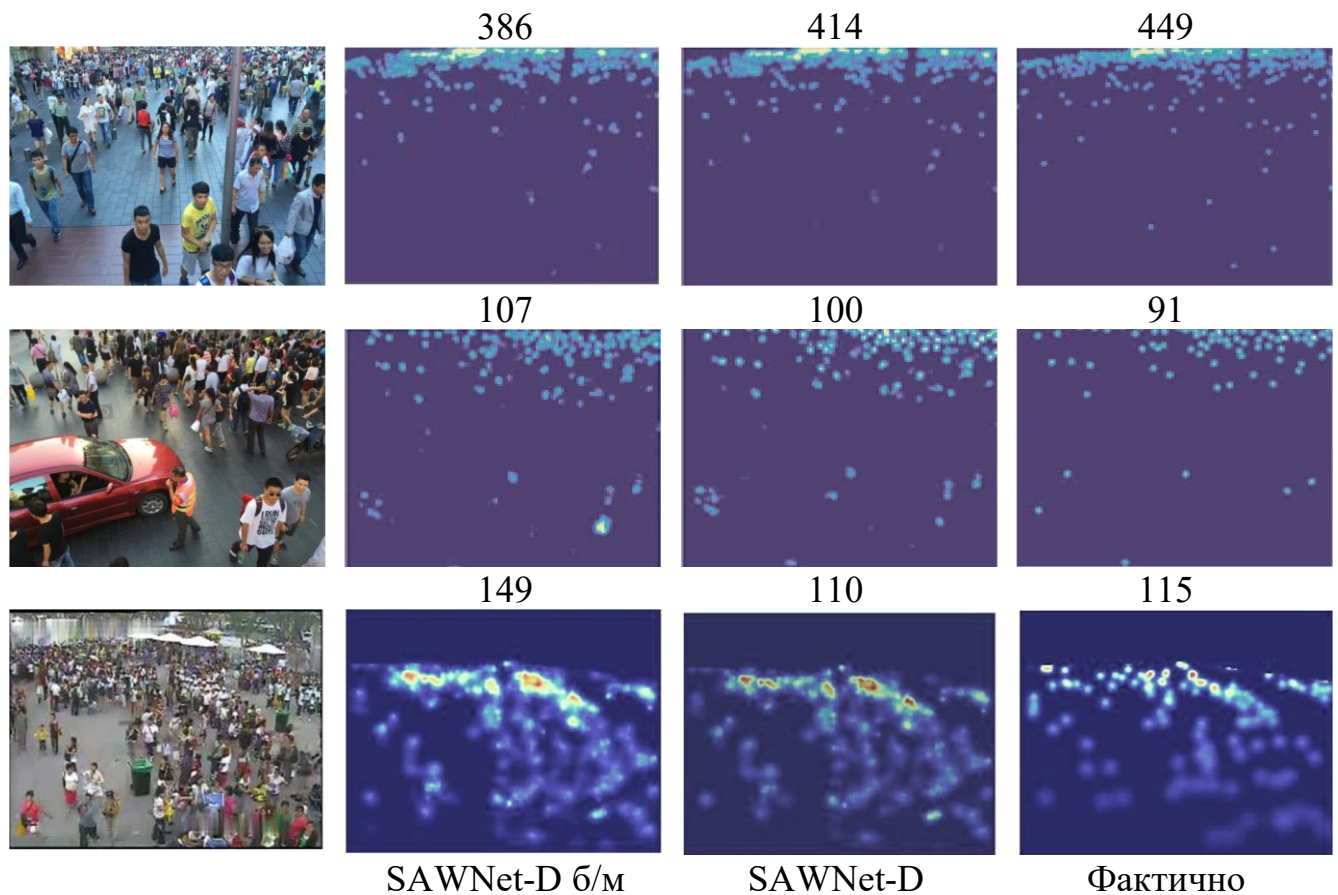


Рисунок 5.2 – Візуалізація карт щільностей

Як видно, з запропонованим модулем оцінена карта густини стає ближчою до фактичної, а оцінені кількості стають більш точними. Наприклад, у другому прикладі після вбудовування глибини відгук у ближніх, нижніх правих областях зменшився і став ближчим фактичного, що означає здатність мережі SAWNet-D генерувати значення щільності з урахуванням масштабу для людей у різних положеннях.

5.1.4 Обговорення результатів дослідження

Експериментальне дослідження підтверджує ефективність запропонованого методу. Ефективність глибинних нейронних мереж, які вбудовують геометричні властивості в архітектуру моделі, додатково демонструє нездатність оригінальних CNN-моделей як у сприйнятті масштабу, так і в моделюванні ознак, що враховують масштаб.

Для задач, які вимагають знання геометрії сцени, корисно використовувати знання про геометрію сцени, щоб допомогти мережі виявити ці геометричні зміни у сцені.

Зокрема, для набору ShanghaiTech-B модель SAWNet-D показала $MAE = 8,1$, що є покращенням на 8,99% порівняно з найкращим серед існуючих методів (TAFNet: 8,9). MSE знизилось з 14,2 до 12,9, що становить покращення на 9,15%.

На наборі WorldExpo середнє значення MAE для SAWNet-D становило 7,84, що на 11,54% краще за попередній найкращий результат (TAFNet: 8,86).

На складному наборі UCF_CC_50, де сцени мають високу густину, MAE зменшилось з 271,8 (TAFNet) до 253,5, тобто на 6,73%. Значення MSE знизилосся з 390,9 до 364,5, що відповідає покращенню на 6,77%.

Натомість для набору Mall запропонована модель SAWNet-D показала $MAE = 2,10$ та $MSE = 2,65$, що є дещо гіршими показниками у порівнянні з TAFNet (2,06 та 2,57 відповідно). Таким чином, на цьому наборі запропонований метод поступається найкращому з існуючих підходів.

У підсумку, розроблений модуль інтеграції інформації про глибину продемонстрував покращення точності на 7–12% на ключових тестових наборах, що підтверджує доцільність урахування просторової структури сцени при підрахунку людей у натовпі за допомогою глибоких моделей.

5.2 Дослідження методу сегментації потоків натовпу в сценах з високою щільністю натовпу

В третьому розділі був представлений метод сегментації натовпу, який було сформульовано з основною метою точної сегментації натовпу у сценах з великою скупченістю людей.

Сегментація контуру переднього плану та кластеризація траєкторій, що належать до однієї і тієї ж групи, дають комплексний, оптимальний результат, що підвищує ефективність сегментації натовпу. Щоб аналітично довести це твердження, треба застосувати запропонований метод до різних наборів даних і оцінили експериментальні результати.

5.2.1 Дослідження значення параметрів методу

За контроль результатів представленого методу відповідає низка параметрів. Для отримання кращих результатів дуже важливо розуміти і призначати відповідні значення цих параметрів.

Значення всіх згаданих параметрів були підібрані шляхом емпіричного аналізу. Для цього ми провели експеримент з різними значеннями порогів на різних відеопослідовностях двох наборів даних. Ми побудували ROC-криву та отримали оптимальне значення порогу.

Після активного контурування отримана область переднього плану розбивається на блоки розміром $b \times b$. Для того, щоб визначити розмір блоку, був проведений експеримент, щоб оцінити, як розмір блоку може вплинути на виділення траєкторії. Якщо розмір блоку занадто малий, це призведе до

відстеження точок інтересу на основі пікселів, що є занадто шумним для відстеження. Більший розмір блоку знову ж таки призводить до сильної оклюзії. У проведених експериментах розмір блоку був встановлений як 16×16 для отримання надійних траєкторій.

Далі, блоки, які містять більше $\phi = 25\%$ пікселів переднього плану, вибираються для виявлення точки інтересу і відстежуються по кадрах, тобто поріг відбору становить 25%.

Цей поріг був визначений шляхом емпіричного аналізу. Для цього був проведений експеримент з 20 різними значеннями ϕ на 40 відеопослідовностях двох наборів даних.

Після цього була побудована ROC-крива і було отримане оптимальне значення ϕ . Після виділення блоків вони залишаються незмінними для $r = 40$ кадрів, а потім весь процес оновлюється, щоб врахувати нові об'єкти, що з'являються в сцені.

Для вибору траєкторії був обраний поріг довжини $LT = 30$. Деякі об'єкти рухаються з дуже високою швидкістю і залишаються в кадрі дуже короткий проміжок часу. Ці тимчасові об'єкти не сприяють отриманню довгих і надійних траєкторій.

При кластеризації траєкторій поріг схожості $\Delta = 100$ (3.14), а поріг припинення кластеризації $\delta = 1$. Для сегментації потоку, значення α , β та γ становить 10%, 20 і 150 відповідно. Значення α , β та γ – поріг просторового перекриття, різниця в напрямку потоку та середнє розташування між кластерами відповідно (див. розділ 3.1.1.5).

5.2.2 Використані метрики оцінки роботи методу

Для оцінки ефективності представленого методу були використані три стандартні показники: міра Жаккара, F-міра та середня абсолютна похибка (MAE). Міра Жаккара вимірює схожість між двома наборами даних наступним чином:

$$J(G, E) = \frac{|G \cap E|}{|G \cup E|}, \quad (5.3)$$

де G – фактичне значення сегментації потоків натовпу для відеопослідовності,
 E – значення сегментації потоків натовпу, отримане за допомогою методу сегментації потоків натовпу.

Перетин підраховує кількість пікселів, які є спільними для обох маркованих наборів. Об'єднання підраховує загальну кількість сегментованих пікселів, зайнятих будь-якою з двох маркованих множин.

Далі обчислюється F-міра згідно наступного рівняння [115]:

$$F = 2 \times \frac{Precision \times Recall}{Precision + Recall} \quad (5.4)$$

Метрика F-міри визначає, наскільки добре працює запропонований метод. Було зроблено припущення, що область натовпу правильно сегментована, якщо сегментована область збігається з позначеною областю базової істини більш ніж на 50%.

Також було обчислено значення середньої абсолютної похибки (MAE) для всіх послідовностей у наборі даних за допомогою наступного рівняння:

$$MAE = \frac{1}{N} \sum_{i=1}^N |z_i - \hat{z}_i|. \quad (5.5)$$

де N – загальна кількість сегментів потоків у відеопослідовності,
 $\hat{z}_i$ – кількість сегментованих потоків натовпу за допомогою методу,
 z_i – фактична кількість сегментів потоків людей у натовпі.

5.2.3 Використані набори даних

Експерименти були проведені на загальнодоступних наборах даних про натовп високої щільності, які включають UCF Crowd Segmentation dataset [116], Collective Motion [117] та були використані деякі послідовності із набору даних Violent Flows [118]. Набори даних містять різноманітні графічні зображення та реальні сценарії з широким спектром динаміки, такі як дорожній рух, марафон, залізничні станції тощо. Для дослідження були обрані лише реальні сценарії з натовпом людей.

5.2.4 Порівняння запропонованого методу з аналогами

5.2.4.1 Дослідження на наборі даних UCF Crowd Segmentation dataset

Перший експеримент був проведений на наборі даних натовпу UCF Crowd Segmentation dataset, який складається з 20 відеопослідовностей графічних та реальних сценаріїв.

Фактичні дані сегментації натовпу були згенеровані вручну з усіх відеопослідовностей.

Запропонований метод був порівняний з методами-аналогами [116, 119, 120]. Оскільки ці методи залежать від відео, вони найкраще підходять для порівняльного аналізу. Якісне порівняння запропонованого підходу обмежується лише 4 відеороликами (Inside Mecca та New York Marathon), оскільки доступність результатів існуючих робіт [116, 119, 120] для поширених відеороликів є дуже незначною.

Більшість відповідних методів не розкривають деталей реалізації. Вони представляють лише якісні результати, які є недостатніми для представлення кількісного порівняння між найсучаснішими методами. Кількісне порівняння з іншими методами за значеннями міри Жаккара наведено в таблиці 5.5 для деяких

відео. Кількісні дані були отримані шляхом компіляції результатів робіт [116, 119, 120].

Запропонований метод перевершує метод [116] зі значенням 0,86 за мірою Жаккара.

Цей результат можна пояснити тим, що запропонований метод сегментує активну область із натовпом, тоді як метод [116] виконує сегментацію по всьому кадру, що знижує точність виділення.

Запропонований метод також був порівняний з методами, представленими в роботах [119, 120] та отримав кращі результати. Це зумовлено тим, що вищезазначені методи-аналоги обробляють потоки через середню статистику. Метод [120] використовує випадкове поле для моделювання потоків, що є більш повільним методом, порівняно із запропонованим.

Таблиця 5.5 – Результати роботи методів на наборі UCF за мірою Жаккара

Відео	Міра Жаккара			
	Метод [116]	Метод [119]	Метод [120]	Запроп. метод
Відео 1	0,67	0,68	0,68	0,73
Відео 2	0,63	0,60	0,90	0,95
Відео 3	0,57	0,74	0,75	0,86
Відео 4	0,41	0,62	0,81	0,92
Середнє	0,57	0,66	0,78	0,86

Виявлено, що запропонований метод може впоратися з потоками натовпу, що перетинаються, тоді як методи-аналоги не можуть ефективно обробляти такі ситуації. Серед недоліків запропонованого методу серед виділити те, що він працює гірше за умови наявності високошвидкісних відео, де об'єкти залишаються на відео менше, ніж 5-10 кадрів. Для таких випадків виділені траєкторії дуже короткі, тому вони не можуть бути визначені як потік.

В таблиці 5.6 наведені результати запропонованого методу за метрикою F-міри.

Таблиця 5.6 – Кількісні показники запропонованого методу на деяких відео з веб-набору даних UCF за мірою F-рахунку

Відео	Фактичних потоків	Вірно передбачених потоків	Помилково знайдених потоків	Пропущених потоків	F-міра
Відео 1	7	6	2	1	0.8
Відео 2	2	2	0	0	1
Відео 3	4	4	0	0	1
Відео 4	3	3	0	0	1

Нажаль, даних по метриці F-міри для інших методів були відсутні. Можна помітити, що запропонований метод працює надзвичайно добре, маючи лише кілька помилкових сегментацій, що пов'язано з неоднозначним або складним рухом об'єкта.

5.2.4.2 Дослідження на наборі даних Collective Motion

В другому експерименті, запропонований метод був оцінений на наборі даних Collective Motion. Цей набір даних містить 413 відеопослідовностей, які відбуваються на 62 різних сценах із натовпом людей, кожна відеопослідовність містить 100 кадрів.

Експериментальні дослідження були проведені на реальних відео із зображенням натовпу високої щільності. Запропонований метод був порівняний з шістьма методами сегментації руху:

- метод Лагранжевої динаміки частинок [116];
- метод сегментації області локальної трансляції [121];
- метод на базі когерентної фільтрації [122];
- метод на основі вимірювання колективності [123];
- метод загальної сегментації руху [124];
- метод сегментації зображень на основі анізотропної дифузії [125].

Кількісне порівняння запропонованого методу з іншими методами було взяти з роботи [126] та наведено в таблиці 5.7.

Кількісні результати порівнюваних методів подано в термінах середньої абсолютної похибки (MAE), міри Жаккара та F-міри. Результати роботи методів демонструють, що запропонований метод досягає менших значень MAE і вищих значень міри Жаккара та F-міри, ніж методи-аналоги.

Це свідчить про те, що запропонований метод може точно сегментувати натовп на відміну від інших підходів, які зазвичай надмірно або недостатньо сегментують регіони натовпу.

Таблиця 5.7 – Результати роботи методів на наборі Collective Motion

Метод	MAE	Міра Жаккара	F-міра
Метод Лагранжевої динаміки частинок [116]	1,25	0,74	0,80
Метод сегментації області локальної трансляції [121]	0,94	0,63	0,69
Метод на базі когерентної фільтрації [122]	1,04	0,44	0,50
Метод на основі вимірювання колективності [123]	0,97	0,47	0,51
Метод загальної сегментації руху [124]	1,79	0,38	0,39
Метод сегментації на основі анізотропної дифузії [125]	0,85	0,44	0,45
Запропонований метод	0,72	0,83	0,86

5.2.5 Обговорення результатів дослідження

Запропонований метод був протестований на множині еталонних наборів даних, а результати були порівняні з кількома найсучаснішими методами.

Запропонований метод показує хороші результати для сцен з високою щільністю натовпу завдяки відстеженню на рівні блоків, що є основною причиною зниження продуктивності стандартного відстеження за допомогою індивідуальних підходів до виявлення.

Оскільки представлений метод не вимагає великих обчислювальних ресурсів, він може бути використаний для сегментації патернів у відеопослідовностях натовпу в реальному часі. Особливістю запропонованого методу є сегментація натовпу з одночасним врахуванням нових об'єктів, що з'являються на сцені.

Це досягається шляхом застосування процедури оновлення через певний проміжок часу, що складається з певної кількості кадрів.

Крім того, важливим аспектом представленого методу є використання переваг сегментації переднього плану за допомогою використання методу активних контурів.

Зокрема, на наборі UCF CSD середнє значення міри Жаккара для запропонованого методу склало 0,86, що перевищує показники методів-аналогів на 10,26–50,88%. За F-мірою метод досяг значень у межах 0,8–1,0, причому у трьох із чотирьох відеофрагментів результат був 1,0.

На наборі Collective Motion запропонований метод досяг середньої абсолютної похибки (MAE) 0,72, що на 15,29–59,78% нижче, ніж у порівнюваних підходів. Значення міри Жаккара склало 0,83 (перевага від 12,16%), а F-міра – 0,86, що на 7,50–120,51% краще за інші методи.

5.3 Дослідження методу виявлення аномалій у сценах з високою щільністю натовпу

Метою запропонованого методу є аналіз подій у натовпі високої щільності для виявлення насильницької або аномальної поведінки. В рамках експериментального дослідження було проведено налаштування параметрів методу, визначені метрики оцінки ефективності методу, визначені набори даних, а також був проведений порівняльний аналіз роботи запропонованого методу з методами-аналогами для різних наборів даних.

5.3.1 Налаштування параметрів методу

В процесі контурної сегментації кадр розбивається на блоки розміром 16×16 . Блоки, які містять більше 25% пікселів переднього плану, обираються для виявлення точок інтересу і відстежуються по кадрах, тобто поріг відбору становить 25%. Після того, як блоки виділено, вони залишаються незмінними для $r = 40$ кадрів.

Для виділення траєкторій: поріг виділення міні-треків $t = 30$, частота оновлення кадрів $r = 40$. Для орієнтованих траєкторій ми розглядаємо напрямок руху міні-треків в діапазоні від 0 до 360 градусів. Орієнтовані міні-треки розподіляються на $n = 8$ корзин з інтервалом 45 градусів.

В аналізі аномальної поведінки обчислюється ентропія орієнтованих траєкторій для різних насильницьких, аномальних і нормальних сцен натовпу. Зважаючи на різноманітність сценаріїв, оптимальною ентропією для аномальної сцени є $Entropy = 0,7$.

Для класифікації використовується зважений KNN класифікатор з $k = 5$. При виявленні аномалій, якщо ентропія збільшується понад певний поріг τ , генерується попередження про аномальну подію, щоб запобігти катастрофам, пов'язаним з натовпом.

5.3.2 Використані метрики оцінки ефективності методу

Для кількісного аналізу було прийнято чотири показники ефективності: частота істинних спрацьовувань (TPR), частота хибних спрацьовувань (FPR), робоча характеристика приймача (ROC) та площа під кривою (AUC).

Частота правильних спрацьовувань відображає частоту правильного виявлення аномальної сцени, а FPR – частоту неправильного виявлення аномальної сцени.

5.3.3 Використані набори даних

Для оцінки запропонованого методу були використані три загальнодоступні набори даних про натовп, які включають UMN [127], UCF Crowd segmentation dataset [116], Violent Crowd Flows [118], а також деякі відеопослідовності, зібрані вручну. Набори даних містять різноманітні сценарії з широким діапазоном динаміки, такі як дорожній рух, марафоне, залізничні станції тощо. Ці відео є колекцією як синтетичних, так і реальних сцен (рис. 5.3).



Рисунок 5.3 – Приклади сцен з використаних наборів даних

В таблиці 5.7 наведена статистика використаних наборів даних. В рамках проведеного експериментального дослідження були використані тільки реальні сценарії натовпу. Деякі набори даних натовпу низької щільності також розглядаються для порівняльного дослідження, оскільки більшість літературних робіт обмежується натовпом низької щільності.

Таблиця 5.7 – Параметри використаних наборів даних

Набір даних	Кадрів	Розмір	FPS	Істинні дані	Сцена/Поведінка	Щільність
UMN Crowd [127]	7725	640x480	30	Так	Реальна/Штучна	Низька
Violent Flows [118]	221450	Різний	30	Ні	Реальна/Реальна	Висока
UCF CSD [116]	24500	640x480	24	Ні	Реальна/ Реальна	Висока
Зібрані відео	22000	Різний	30	Ні	Реальна/ Реальна	Висока

5.3.4 Порівняння запропонованого методу з аналогами

5.3.4.1 Оцінка роботи методу на наборі UMN

Перше дослідження було проведено на наборі даних UMN [127], який дозволяє оцінити продуктивність методів виявлення аномалій. Набір даних містить 11 відеокліпів з трьома різними сценаріями. Кожна відеопослідовність починається із загальної або нормальної поведінки і закінчується натовпом, що панічно тікає.

Він складається з 7725 кадрів з роздільною здатністю 320x240 пікселів. Хоча набір даних обмежений щодо щільності натовпу та його варіацій, він був використаний, так як для цього набору є дані роботи інших методів.

Порівняльні результати запропонованого методу з іншими сучасними методами, такими як модель соціальних сил [128], оптичний метод [128], метод розрідженої вартості реконструкції (SRC) [129], метод матричної апроксимації [130] та метод хаотичних інваріантів [131], наведені в таблиці 5.8. Всі кількісні результати порівняння підходів були взяті з роботи [129].

Таблиця 5.8 – Порівняння результатів для набору даних UMN

Метод	AUC	Швидкість (кадрів\сек.)
Оптичний потік [128]	0,840	10
Модель соціальних сил [128]	0,960	20
Метод розрідженої вартості реконструкції [129]	0,980	15
Апроксимація матриць [130]	0,980	15
Хаотичні інваріанти [131]	0,980	20
Запропонований метод	0,995	30

Ці методи найкраще підходять для порівняльного аналізу. Автори роботи [128] використовували модель соціальних сил для розрахунку взаємодії між натовпом. Вона виявляє аномалії на основі високого відхилення у величині сили

взаємодії між частинками. Їхній підхід досягає хороших результатів лише для розрідженого натовпу. Вони також використовували чистий оптичний метод, заснований на потоці, який досягає 0,840 AUC.

Автори роботи [130] обчислюють матрицю руху відеопослідовностей. Методи, засновані на русі, як правило, працюють гірше при наявності шуму, який завжди присутній в реальних сценах натовпу.

Автори роботи [131] визначають хаотичні інваріанти для аналізу сцени, а автор роботи [129] використовує розріджену вартість реконструкції поля руху.

Зображення ROC-кривої запропонованого методу для різних наборів даних наведена на рисунку 5.4.

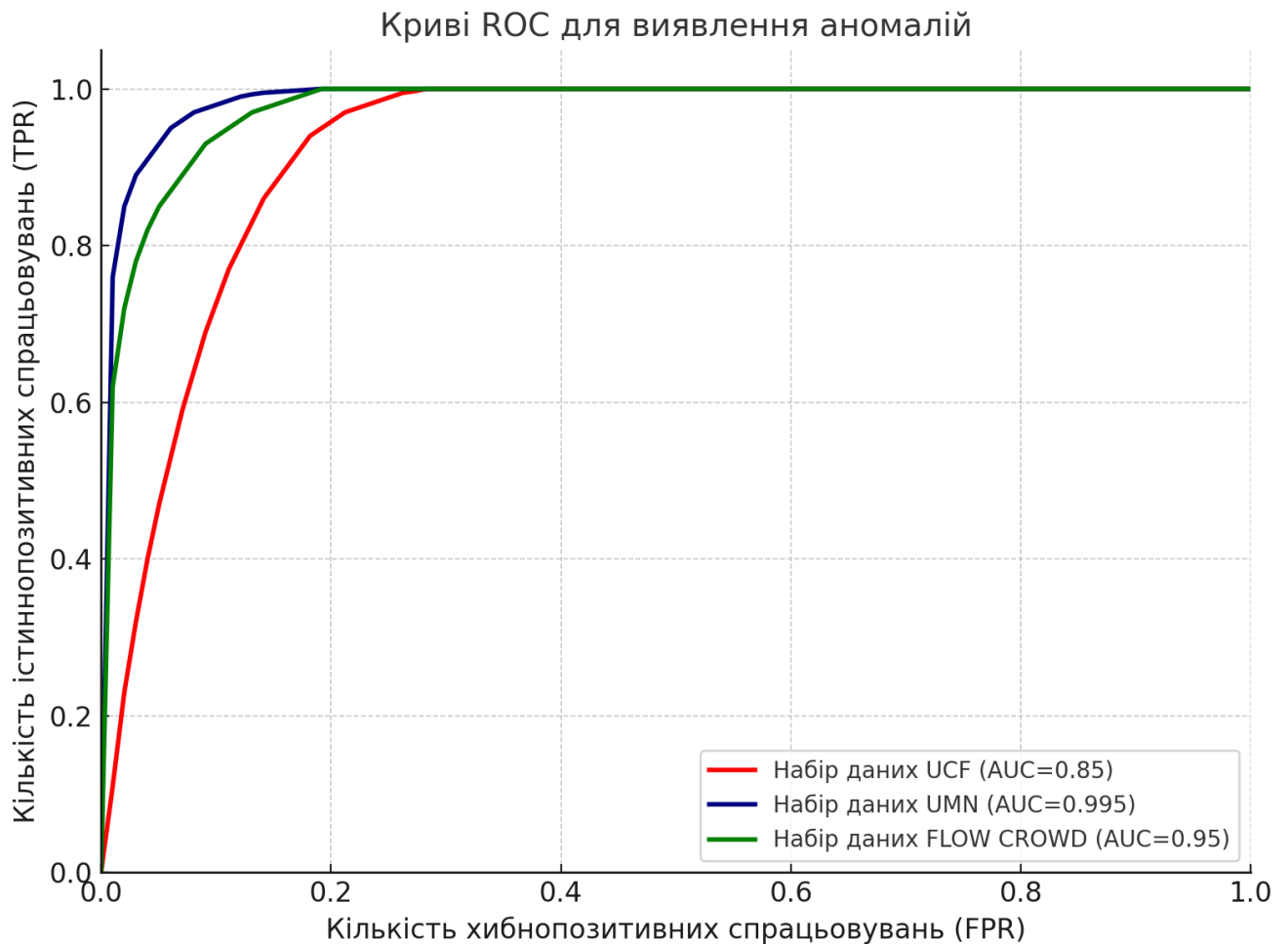


Рисунок 5.4 – Графік ROC-кривих для кожного з експериментальних наборів даних

За результатами огляду методів-аналогів можна зробити висновок, що вони виявляють аномалії на рівні кадру або пікселя, тоді як запропонований метод працює на рівні сцени, оскільки різниця між нормальною та аномальною сценою стає більш очевидною після певного інтервалу часу.

Цей набір даних призводить до значних змін у структурі руху сцени. Завдяки цьому факту, запропонований метод досягає майже ідеальних результатів, найбільш близьких до істини (значення площі під ROC-кривою становить 0,995). Порівняні методи також показують хороші результати, але запропонований метод перевершує їх, як показано в таблиці 5.8.

5.3.4.2 Оцінка роботи методу на наборі UCF Crowd Segmentation dataset

Наступне експериментальне дослідження було проведене на наборі даних UCF Crowd Segmentation dataset. Цей набір даних містить 20 відео з нормальними та аномальними сценами натовпу з роздільною здатністю 640x480.

Нормальною поведінкою натовпу вважається пішохідна прогулянка або марафонський біг, а ненормальною – тиснява, зіткнення протестувальників, бійки та біг у випадкових напрямках. З 20 відеозаписів 12 відеопослідовностей є нормальними, а решта 8 – ненормальними.

Для кожного відеоматеріалу, аномальні кадри були вручну позначені як істинні. Для оцінки ефективності запропонованого методу та методів-аналогів було обчислено TPR, FPR, точність, а також була побудована ROC-крива (рис. 5.4).

Значення площі під ROC-кривою було порівняно з моделлю соціальних сил [128], оптичним методом [128] та методом GLCM [132]. Результати порівняння наведені в таблиці 5.9.

Метод [128] аналізує аномальність, вводячи сітку частинок у сцену і переміщуючи їх за допомогою елементарного оптичного поля потоку. Через схожість пікселів, адвективні частинки важко розрізнити, що призводить до спотворення потоку.

Таблиця 5.9 – Порівняння результатів для набору даних UCF

Метод	AUC	Швидкість (кадрів\сек.)
Оптичний потік [128]	0,660	5
GLCM [132]	0,820	5
Модель соціальних сил [128]	0,730	10
Запропонований метод	0,850	30

Автори роботи [132] запропонували метод виявлення аномалій на основі текстур GLCM, який вимагає сильної структури, що є малодоступною в реальних сценах випадкового натовпу. Цей набір даних не був широко протестований, тому порівняння обмежується лише двома методами, як зазначено вище, і запропонований метод перевершує ці два методи за показником AUC. У роботі [128] значення AUC для методу оптичної залежності та методу соціальної сили становлять 0,65 та 0,73 відповідно. Значення AUC для [132] становить 0,78, що набагато краще, ніж обидва підходи, запропоновані в [128]. Запропонований метод досягає значення AUC 0,85, що є найкращим серед усіх порівнюваних підходів.

5.3.4.3 Оцінка роботи методу на наборі Violent Flows Crowd

В останньому експерименті був використаний дуже складний набір Violent Flows Crowd, який був представлений в роботі [118]. Єдиною метою використання цього набору даних є оцінка аномальності натовпу та насильства в сценах. Набір даних складається з 123 зразків насильницьких і ненасильницьких сцен натовпу, зібраних з відеозаписів, завантажених на різних веб-сайтах.

Відео з насильством складається з різних послідовностей, які дають подібне візуальне сприйняття, як у реальних сценах натовпу. Тому цей набір даних підходить для оцінки ефективності запропонованого методу для виявлення аномальних сцен у натовпі високої щільності.

Кадри, на яких сцена із звичайної перетворюється на насильницьку, були розмічені вручну. Для виявлення аномальних сцен була використана метрика AUC. Окрім запропонованого методу, в дослідженні були використані п'ять методів-аналогів:

- в методі [133] аномальна поведінка виявляється за допомогою обробки згустків. Даний метод розрахований на розпізнавання бойових дій та не підходить для загального аналізу насильства у натовпі;

- метод [134] також зосереджений на виявленні сцен насильства, покладаючись на величину оптичного поля. Даний підхід вимагає навченої моделі для класифікації насильницьких і ненасильницьких сцен;

- автор роботи [135] для розрізнення насильницьких і ненасильницьких дій фіксує характерні локальні форми і патерни руху активності. Локальна форма активності виявляється в сценах з високою щільністю натовпу;

- автори роботи [132] описують візуальну текстуру сцен, використовуючи статистичні властивості функції GLCM. Ключовою вимогою їхнього підходу є підтримка сильної структури сцен, що не є практичним у системах спостереження в реальному часі, оскільки на відзнятий матеріал можуть впливати ефекти освітлення, темряви, оклюзії тощо.

На відміну від вищезгаданих підходів, запропонований метод не вимагає попередньо навченої моделі і працює в реальному часі. В таблиці 5.10 наведене кількісне порівняння запропонованого методу з іншими сучасними методами.

Таблиця 5.10 – Порівняння результатів для набору даних UMN

Метод	AUC	Швидкість (кадрів\сек.)
FF-method [133]	0,750	10
OViF [134]	0,805	20
MoSIFT [135]	0,875	20
GLCM [132]	0,925	10
Запропонований метод	0,950	30

Видно, що запропонований метод перевершує сучасні методи зі значенням AUC 0,95 значенням AUC (рис. 5.4).

5.3.5 Обговорення результатів дослідження

Експерименти були проведені на трьох наборах даних: UMN, UCF Crowd Segmentation dataset та Violent Flow. Набори даних UCF Crowd Segmentation dataset та Violent Flows є компліментарними до запропонованого методу, оскільки вони складаються зі сцен щільного натовпу з реальних ситуацій, таких як марафон, стадіон, тиснява, політичні мітинги тощо.

Експериментальні результати дають багатообіцяючі результати у порівнянні з існуючими підходами. Запропонований метод має низьку обчислювальну складність і може бути застосований до реальних відеопослідовностей натовпу для виявлення потенційно аномальних сцен.

Зокрема, для набору UMN запропонований метод досяг $AUC = 0,995$, що на 1,53% краще, ніж у найкращого з конкурентів – методу хаотичних інваріантів ($AUC = 0,980$). Це покращення є найбільшим серед розглянутих наборів і свідчить про високу ефективність моделі в умовах зміни структури сцени.

На наборі UCF CSD запропонований метод досяг $AUC = 0,850$, що становить покращення на 3,66% порівняно з найкращим серед методів-аналогів – GLCM, який показав $AUC = 0,820$. Це свідчить про здатність моделі виявляти складні аномалії, пов'язані з поведінкою натовпу.

На складному наборі Violent Flows Crowd, який містить сцени насильства в реальних умовах, запропонований метод досяг $AUC = 0,950$, перевершивши найкращий з існуючих підходів (GLCM: $AUC = 0,925$) на 2,70%. Таким чином, навіть у складних і варіативних умовах модель демонструє високу точність.

Отже, запропонований підхід стабільно перевершує існуючі методи в межах 1,53–3,66% на всіх протестованих наборах даних.

5.4 Висновки до п'ятого розділу

В даному розділі дисертаційної роботи було проведено дослідження компонентів підсистеми візуального аналізу натовпу в складі системи інтелектуального відеоспостереження.

В процесі дослідження методу сегментації потоків натовпу в сценах з високою щільністю натовпу, було проведено тестування на чотирьох наборах даних: ShanghaiTech-B, WorldExpo, UCF_CC_50 та Mall. На більшості з них запропонований метод показав покращення. Зокрема, на ShanghaiTech-B MAE зменшився з 8,9 до 8,1 (покращення на 8,99%), на WorldExpo – з 8,86 до 7,84 (на 11,54%), а на UCF_CC_50 – з 271,8 до 253,5 (на 6,73%). Лише на наборі Mall модель трохи поступалася TAFNet: MAE становив 2,10 проти 2,06 (зниження точності на 1,94%). Результати підтверджують ефективність урахування глибини сцени для покращення точності підрахунку в умовах складного фону та варіативної щільності.

В процесі дослідження методу сегментації потоків натовпу в сценах з високою щільністю натовпу було вказано, що метод базується на поєднанні сегментації переднього плану та кластеризації траєкторій, що дозволяє виділяти потоки людей у складних динамічних умовах.

Проведено емпіричний підбір параметрів на основі ROC-кривих. На наборі UCF CSD запропонований підхід досяг середнього значення міри Жаккара 0,86, що перевищує результати аналогів на 10,26 – 50,88%, а також F-міри до 1,0. На складнішому наборі Collective Motion отримано MAE = 0,72 (на 15,29 – 59,78% краще), міру Жаккара 0,83 і F-міру 0,86 (покращення на 7,50–120,51%). Результати підтверджують ефективність методу в задачах сегментації руху та виділення групової динаміки в натовпі.

В процесі дослідження методу виявлення аномалій у сценах з високою щільністю натовпу було проведено тестування на наборах UMN, UCF CSD та Violent Flows Crowd, яке показало стабільне покращення результатів. Зокрема, на UMN отримано AUC = 0,995 (+1,53%), на UCF CSD – 0,850 (+3,66%), на Violent

Flows Crowd – 0,950 (+2,70%) порівняно з найкращими з існуючих методів. Таким чином, запропонований підхід демонструє високу точність і придатність для застосування в системах інтелектуального відеоспостереження.

ВИСНОВКИ

Сукупність отриманих в дисертації результатів вирішує актуальну науково-технічну задачу підвищення ефективності систем візуального аналізу натовпу шляхом розробки моделей та методів оцінки щільності натовпу за рахунок використання інформації про глибину сцени та аналізу сцен з високою щільністю натовпу для виявлення аномалій за рахунок оцінки ентропії напрямлених міні-треків та часового відхилення заповнюваності сцени.

На основі проведених досліджень отримані наукові результати, які складають істотний внесок у подальший розвиток систем візуального аналізу натовпу та систем інтелектуального відеоспостереження.

Проведено системний аналіз сучасних підходів до моделювання руху щільних натовпів, із акцентом на їх структурні та динамічні характеристики. Встановлено, що наявні методи часто ігнорують геометричну інформацію сцени, довготривалі часові зв'язки та локальний просторовий контекст, що знижує ефективність аналізу в реальних умовах. У зв'язку з цим обґрунтовано доцільність інтеграції геометрії сцени у глибинні нейронні мережі для покращення масштабування та оцінки щільності. Також представлено модель представлення траєкторій із багатовимірними дескрипторами, яка дає змогу виявляти патерни руху шляхом кластеризації. Окремо окреслено концепцію ентропійного підходу до виявлення аномалій, що базується на стабільних міні-треках та локальних флуктуаціях, що підвищує точність і інтерпретованість аналізу. Сформульовано мету та задачі дисертаційного дослідження.

Розроблено метод підрахунку людей у натовпі з урахуванням глибини сцени для компенсації перспективних викривлень. Запропоновано інтегрувати карту глибини у просторові ознаки моделі, що покращує відповідність між фізичними розмірами об'єктів і їх цифровим представленням. Метод змінює акценти в ознаках згорткової мережі залежно від глибини, а механізм просторової уваги дозволяє ігнорувати фонові області. Проведений аналіз окремих блоків показав, що навіть часткова реалізація покращує точність, а повна комбінація забезпечує найкращі

результати. Також вивчено вплив кількості модулів глибини на різних рівнях мережі: для неглибоких моделей додавання модулів дає покращення, тоді як для глибоких – може призводити до перенавчання. Отримані результати підтверджують ефективність урахування глибини сцени для точнішої оцінки щільності натовпу.

Ось розширений, але стислий за структурою варіант тексту з усіма ключовими акцентами на запропоновані методи та модель:

Запропоновано два методи, спрямовані на покращення аналізу натовпу високої щільності в умовах складного відеоспостереження. Перший метод – метод сегментації потоків натовпу на основі кластеризації траєкторій із використанням активних контурів. Такий підхід дозволяє ефективно виділяти локальні патерни колективного руху навіть у динамічно змінних умовах. Для реалізації цього методу розроблено модель представлення траєкторій, яка враховує ключові просторові та динамічні ознаки: форму траєкторії, напрямок руху, просторове положення об'єкта та локальну щільність. Завдяки врахуванню багатовимірних характеристик траєкторій та локального контексту сцени модель забезпечує підвищену точність і стійкість сегментації потоків у складних сценах, що відзначається як перевага порівняно з одномасштабними або узагальненими методами. Другий запропонований метод – метод виявлення аномалій у русі натовпу – базується на аналізі напрямлених міні-треків, сформованих із просторово-часових точок інтересу. Він враховує динамічну зміну структури заповнення сцени й включає фільтрацію коротких та шумових траєкторій. Додатково, інтеграція оцінки варіативності щільності дозволяє виявляти раптові зміни в русі, характерні для аномальних подій. Завдяки цьому метод демонструє високу чутливість до нетипової поведінки, що є важливим для практичних застосувань у реальному середовищі відеоспостереження.

Реалізовано підсистему візуального аналізу натовпу як компонент системи інтелектуального відеоспостереження. Визначено її мету, задачі та роль у виявленні критичних ситуацій у місцях масового скупчення людей. Розроблено модель інформаційних потоків, описано функціональні та нефункціональні вимоги,

сформовано архітектуру через UML-діаграми (логічну, розгортання, пакетів, класів). Особливу увагу приділено організації розробки в GitHub, аналізу якості коду та дотриманню принципів модульності й безпеки. Функціональне тестування за 20 сценаріями підтвердило повну відповідність вимогам і готовність підсистеми до практичного застосування.

Проведено дослідження компонентів підсистеми візуального аналізу натовпу в системі інтелектуального відеоспостереження. Метод підрахунку з урахуванням глибини сцени протестовано на наборах ShanghaiTech-B, WorldExpo, UCF_CC_50 і Mall. Отримано покращення: MAE на ShanghaiTech-B зменшився з 8,9 до 8,1 (–8,99%), на WorldExpo – з 8,86 до 7,84 (–11,54%), на UCF_CC_50 – з 271,8 до 253,5 (–6,73%). Лише на Mall результат трохи гірший: MAE = 2,10 проти 2,06 (–1,94%). Метод сегментації потоків на основі кластеризації траєкторій показав ефективність: на UCF CSD міра Жаккара – 0,86 (+10,26 – 50,88%), F-міра – до 1,0; на Collective Motion MAE = 0,72 (покращення на 15,29 – 59,78%), Жаккара – 0,83 від 12,16%), F-міра – 0,86 (+7,50 – 120,51%). Метод виявлення аномалій протестовано на UMN, UCF CSD та Violent Flows Crowd: AUC = 0,995 (+1,53%), 0,850 (+3,66%) та 0,950 (+2,70%) відповідно. Усі результати підтверджують ефективність підходів у складних умовах відеоспостереження.

Дисертацію виконано згідно тематичних планів НДР Національного університету «Одеська політехніка» за період 2024 – 2025 рр. в рамках держбюджетної теми № 237-177 – «Програмні системи машинного навчання та їх застосування в галузі інтелектуального аналізу даних».

Розроблені в дисертаційній роботі моделі та методи отримали впровадження у діяльності товариства з обмеженою відповідальністю «Провектус ІТ» та товариства з обмеженою відповідальністю «СФ ІНВЕСТ», а також знайшли відображення у навчальному процесі та науково-дослідницькій діяльності Національного університету «Одеська політехніка».

СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ

1. Sreenu G., Saleem Durai M. A. Intelligent video surveillance: a review through deep learning techniques for crowd analysis // *Journal of Big Data*. – 2019. – Vol. 6, No. 1. – Article 48. – DOI: <https://doi.org/10.1186/s40537-019-0212-5>.
2. Fan Ch.-T., Wang Y.-K., Huang C.-R. Heterogeneous information fusion and visualization for a large-scale intelligent video surveillance system // *IEEE Transactions on Systems, Man, and Cybernetics: Systems*. – 2016. – Vol. 47. – P. 1–12. – DOI: <https://doi.org/10.1109/TSMC.2016.2531671>.
3. Yangınlar G. Internet of Things (IoT) in Intelligent Transportation Systems: Benefits and Challenges of Implementation // *The Eurasia Proceedings of Science Technology Engineering and Mathematics*. – 2024. – Vol. 27. – P. 16–23. – DOI: <https://doi.org/10.55549/epstem.1517792>.
4. Cheong K. H., Poeschmann S., Lai J. W. Practical automated video analytics for crowd monitoring and counting // *IEEE Access*. – 2019. – Vol. 7. – P. 182608–182618. – DOI: <https://doi.org/10.1109/ACCESS.2019.2958255>.
5. Khan M. A., Menouar H., Hamila R. Visual crowd analysis: open research problems // *AI Magazine*. – 2023. – Vol. 44, No. 1. – P. 76–87. – DOI: <https://doi.org/10.1002/aaai.12117>.
6. Boominathan L., Kruthiventi S. S. S., Babu R. V. CrowdNet: A deep convolutional network for dense crowd counting // *arXiv preprint*. – 2016. – arXiv:1608.06197. – Режим доступу: <https://arxiv.org/abs/1608.06197>. – Дата звернення: 16 червня 2025 р.
7. Ma Zh., Wei X., Hong X., Gong Y. Bayesian Loss for Crowd Count Estimation with Point Supervision // *arXiv preprint*. – 2019. – arXiv:1908.03684. – Режим доступу: <https://arxiv.org/abs/1908.03684>. – Дата звернення: 16 червня 2025 р.
8. Thanasutives P., Fukui K., Numao M., Kijisirikul B. Encoder-Decoder Based Convolutional Neural Networks with Multi-Scale-Aware Modules for Crowd Counting // *arXiv preprint*. – 2020. – arXiv:2003.05586. – Режим доступу: <https://arxiv.org/abs/2003.05586>. – Дата звернення: 16 червня 2025 р.

9. Zhang Y., Zhou D., Chen S., Gao S., Ma Y. Single-Image Crowd Counting via Multi-Column Convolutional Neural Network // Proceedings of the 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). – 2016. – P. 589–597. – DOI: <https://doi.org/10.1109/CVPR.2016.70>.
10. Idrees H., Tayyab M., Athrey K., Zhang D., Al-Maadeed S., Rajpoot N., Shah M. Composition Loss for Counting, Density Map Estimation and Localization in Dense Crowds // arXiv preprint. – 2018. – arXiv:1808.01050. – Режим доступа: <https://arxiv.org/abs/1808.01050>. – Дата звернення: 16 червня 2025 р.
11. Sindagi V. A., Yasarla R., Patel V. M. JHU-CROWD++: Large-Scale Crowd Counting Dataset and A Benchmark Method // arXiv preprint. – 2020. – arXiv:2004.03597. – Режим доступа: <https://arxiv.org/abs/2004.03597>. – Дата звернення: 16 червня 2025 р.
12. Wang Q., Gao J., Lin W., Li X. NWPU-Crowd: A Large-Scale Benchmark for Crowd Counting and Localization // IEEE Transactions on Pattern Analysis and Machine Intelligence. – 2021. – Vol. 43, No. 6. – P. 2141–2149. – DOI: <https://doi.org/10.1109/TPAMI.2020.3013269>.
13. Khan K., Khan R. U., Albattah W., Nayab D., Qamar A. M., Habib S., Islam M. Crowd counting using end-to-end semantic image segmentation // Electronics. – 2021. – Vol. 10, No. 11. – Article 1293. – DOI: <https://doi.org/10.3390/electronics10111293>.
14. Braun M., Krebs S., Flohr F., Gavrilă D. M. EuroCity Persons: A Novel Benchmark for Person Detection in Traffic Scenes // IEEE Transactions on Pattern Analysis and Machine Intelligence. – 2019. – Vol. 41, No. 8. – P. 1844–1861. – DOI: <https://doi.org/10.1109/TPAMI.2019.2897684>.
15. Ren Sh., He K., Girshick R., Sun J. Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks // arXiv preprint. – 2016. – arXiv:1506.01497. – Режим доступа: <https://arxiv.org/abs/1506.01497>. – Дата звернення: 16 червня 2025 р.
16. Liu W., Anguelov D., Erhan D., Szegedy Ch., Reed S., Fu Ch.-Y., Berg A. C. SSD: Single Shot MultiBox Detector // Computer Vision – ECCV 2016. – Springer

International Publishing, 2016. – P. 21–37. – DOI: https://doi.org/10.1007/978-3-319-46448-0_2.

17. Redmon J., Farhadi A. YOLOv3: An Incremental Improvement // arXiv preprint. – 2018. – arXiv:1804.02767. – Режим доступу: <https://arxiv.org/abs/1804.02767>. – Дата звернення: 16 червня 2025 р.

18. Law H., Deng J. CornerNet: Detecting Objects as Paired Keypoints // International Journal of Computer Vision. – 2020. – Vol. 128, No. 3. – P. 642–656. – DOI: <https://doi.org/10.1007/s11263-019-01204-1>.

19. Duan K., Bai S., Xie L., Qi H., Huang Q., Tian Q. CenterNet: Keypoint Triplets for Object Detection // Proceedings of the 2019 IEEE/CVF International Conference on Computer Vision (ICCV). – 2019. – P. 6568–6577. – DOI: <https://doi.org/10.1109/ICCV.2019.00667>.

20. Tian Z., Shen Ch., Chen H., He T. FCOS: Fully Convolutional One-Stage Object Detection // arXiv preprint. – 2019. – arXiv:1904.01355. – Режим доступу: <https://arxiv.org/abs/1904.01355>. – Дата звернення: 16 червня 2025 р.

21. Li Y., Mao H., Girshick R., He K. Exploring Plain Vision Transformer Backbones for Object Detection // arXiv preprint. – 2022. – arXiv:2203.16527. – Режим доступу: <https://arxiv.org/abs/2203.16527>. – Дата звернення: 16 червня 2025 р.

22. Carion N., Massa F., Synnaeve G., Usunier N., Kirillov A., Zagoruyko S. End-to-End Object Detection with Transformers // arXiv preprint. – 2020. – arXiv:2005.12872. – Режим доступу: <https://arxiv.org/abs/2005.12872>. – Дата звернення: 16 червня 2025 р.

23. LeCun Y., Bottou L., Bengio Y., Haffner P. Gradient-based learning applied to document recognition // Proceedings of the IEEE. – 1998. – Vol. 86, No. 11. – P. 2278–2324. – DOI: <https://doi.org/10.1109/5.726791>.

24. Everingham M., Van Gool L., Williams C. K. I., Winn J., Zisserman A. The Pascal Visual Object Classes (VOC) Challenge // International Journal of Computer Vision. – 2010. – Vol. 88, No. 2. – P. 303–338. – DOI: <https://doi.org/10.1007/s11263-009-0275-4>.

25. Lin T.-Y., Maire M., Belongie S., Hays J., Perona P., Ramanan D., Dollár P., Zitnick C. L. Microsoft COCO: Common Objects in Context // Computer Vision – ECCV 2014. – Springer, 2014. – P. 740–755. – DOI: https://doi.org/10.1007/978-3-319-10602-1_48.
26. Idrees H., Soomro K., Shah M. Detecting humans in dense crowds using locally-consistent scale prior and global occlusion reasoning // IEEE Transactions on Pattern Analysis and Machine Intelligence. – 2015. – Vol. 37, No. 10. – P. 1986–1998. – DOI: <https://doi.org/10.1109/TPAMI.2015.2396051>.
27. Wang J., Lv W., Jiang H., Fang Z., Ma J. Exploring crowd persistent dynamism from pedestrian crossing perspective: An empirical study // arXiv preprint. – 2023. – arXiv:2311.04827. – Режим доступа: <https://arxiv.org/abs/2311.04827>. – Дата звернення: 16 червня 2025 р.
28. Teng Li, Huan Chang, Meng Wang, Bingbing Ni, Richang Hong, Shuicheng Yan. Crowded Scene Analysis: A Survey // arXiv preprint. – 2015. – arXiv:1502.01812. – Режим доступа: <https://arxiv.org/abs/1502.01812>. – Дата звернення: 16 червня 2025 р.
29. Dupont C., Tobías L., Luvison B. Crowd-11: A Dataset for Fine Grained Crowd Behaviour Analysis // Proceedings of the 2017 IEEE Conference on Computer Vision and Pattern Recognition Workshops (CVPRW). – 2017. – P. 2184–2191. – DOI: <https://doi.org/10.1109/CVPRW.2017.271>.
30. Behera S., Dogra D. P., Bandyopadhyay M. K., Roy P. P. Understanding Crowd Flow Movements Using Active-Langevin Model // arXiv preprint. – 2020. – arXiv:2003.05626. – Режим доступа: <https://arxiv.org/abs/2003.05626>. – Дата звернення: 16 червня 2025 р.
31. Luque Sánchez F., Dupont I., Tabik S., Herrera F. Revisiting crowd behaviour analysis through deep learning: Taxonomy, anomaly detection, crowd emotions, datasets, opportunities and prospects // Information Fusion. – 2020. – Vol. 64. – P. 318–335. – DOI: <https://doi.org/10.1016/j.inffus.2020.07.008>.

32. Xuan Wang, Zhigang Zhu. Context Understanding in Computer Vision: A Survey // arXiv preprint. – 2023. – arXiv:2302.05011. – Режим доступу: <https://arxiv.org/abs/2302.05011>. – Дата звернення: 16 червня 2025 р.
33. Rabiner L. R. A tutorial on Hidden Markov Models and selected applications in speech recognition // Proceedings of the IEEE. – 1989. – Vol. 77, No. 2. – P. 257–286. – DOI: <https://doi.org/10.1109/5.18626>.
34. Lafferty J., McCallum A., Pereira F. Conditional Random Fields: Probabilistic Models for Segmenting and Labeling Sequence Data // Proceedings of the Eighteenth International Conference on Machine Learning (ICML'01). – San Francisco, CA, USA: Morgan Kaufmann Publishers Inc., 2001. – P. 282–289.
35. Ji S., Xu W., Yang M., Yu K. 3D Convolutional Neural Networks for Human Action Recognition // IEEE Transactions on Pattern Analysis and Machine Intelligence. – 2013. – Vol. 35, No. 1. – P. 221–231. – DOI: <https://doi.org/10.1109/TPAMI.2012.59>.
36. Mikolov T., Karafiát M., Burget L., Černocký J., Khudanpur S. Recurrent neural network based language model // Interspeech. – 2010. – P. 1045–1048. – DOI: <https://doi.org/10.21437/Interspeech.2010-343>.
37. Kipf T. N., Welling M. Semi-Supervised Classification with Graph Convolutional Networks // arXiv preprint. – 2017. – arXiv:1609.02907. – Режим доступу: <https://arxiv.org/abs/1609.02907>. – Дата звернення: 16 червня 2025 р.
38. Gustavo Fuhr, Claudio Rosito Jung. Collective behavior recognition using compact descriptors // arXiv preprint. – 2018. – arXiv:1809.10499. – Режим доступу: <https://arxiv.org/abs/1809.10499>. – Дата звернення: 16 червня 2025 р.
39. Wang H., Liao J., Cheng T., Gao Z., Liu H., Ren B., Bai X., Liu W. Knowledge Mining with Scene Text for Fine-Grained Recognition // 2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). – 2022. – P. 4614–4623. – DOI: <https://doi.org/10.1109/CVPR52688.2022.00458>.
40. Choi W., Shahid K., Savarese S. What are they doing? : Collective activity classification using spatio-temporal relationship among people // 2009 IEEE 12th International Conference on Computer Vision Workshops (ICCV Workshops). – 2009. – P. 1282–1289. – DOI: <https://doi.org/10.1109/ICCVW.2009.5457461>.

41. Wu L.-F., Wang Q., Jian M., Qiao Y., Zhao B.-X. A Comprehensive Review of Group Activity Recognition in Videos // *International Journal of Automation and Computing*. – 2021. – Vol. 18, No. 3. – P. 334–350. – DOI: <https://doi.org/10.1007/s11633-020-1258-8>.
42. Zalluhoglu C., Ikizler-Cinbis N. Collective Sports: A multi-task dataset for collective activity recognition // *Image and Vision Computing*. – 2020. – Vol. 94. – P. 103870. – DOI: <https://doi.org/10.1016/j.imavis.2020.103870>.
43. Vrigkas M., Nikou C., Kakadiaris I. A. A Review of Human Activity Recognition Methods // *Frontiers in Robotics and AI*. – 2015. – Vol. 2. – Article 28. – DOI: <https://doi.org/10.3389/frobt.2015.00028>.
44. Oltramari A., Francis J., Henson C., Ma K., Wickramarachchi R. Neuro-symbolic architectures for context understanding // *arXiv preprint*. – 2020. – arXiv:2003.04707. – Режим доступу: <https://arxiv.org/abs/2003.04707>. – Дата звернення: 16 червня 2025 р.
45. Sabokrou M., Fathy M., Hosseini M., Klette R. Real-Time Anomaly Detection and Localization in Crowded Scenes // *arXiv preprint*. – 2015. – arXiv:1511.06936. – Режим доступу: <https://arxiv.org/abs/1511.06936>. – Дата звернення: 16 червня 2025 р.
46. Mahadevan V., Li W., Bhalodia V., Vasconcelos N. Anomaly detection in crowded scenes // *2010 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR)*. – 2010. – P. 1975–1981. – DOI: <https://doi.org/10.1109/CVPR.2010.5539872>.
47. Hartigan J. A., Wong M. A. Algorithm AS 136: A K-Means Clustering Algorithm // *Journal of the Royal Statistical Society. Series C (Applied Statistics)*. – 1979. – Vol. 28, No. 1. – P. 100–108. – Режим доступу: <http://www.jstor.org/stable/2346830>. – Дата звернення: 16 червня 2025 р.
48. Cortes C., Vapnik V. Support-vector networks // *Machine Learning*. – 1995. – Vol. 20, No. 3. – P. 273–297. – DOI: <https://doi.org/10.1007/BF00994018>.
49. Arthur P. Dempster, Nan M. Laird, Donald B. Rubin. Maximum likelihood from incomplete data via the EM algorithm // *Journal of the Royal Statistical Society*,

Series B (Methodological). – 1977. – Vol. 39, No. 1. – P. 1–38. – DOI: <https://doi.org/10.1111/j.2517-6161.1977.tb01600.x>.

50. Manevitz L. M., Yousef M. One-class SVMs for document classification // *Journal of Machine Learning Research*. – 2002. – Vol. 2. – P. 139–154. – ISSN 1532-4435. – Режим доступу: <http://www.jmlr.org/papers/v2/manevitz01a.html>. – Дата звернення: 16 червня 2025 р.

51. Hochreiter S., Schmidhuber J. Long short-term memory // *Neural Computation*. – 1997. – Vol. 9, No. 8. – P. 1735–1780. – DOI: <https://doi.org/10.1162/neco.1997.9.8.1735>.

52. Goodfellow I., Pouget-Abadie J., Mirza M., Xu B., Warde-Farley D., Ozair S., Courville A., Bengio Y. Generative adversarial nets // *Advances in Neural Information Processing Systems*. – 2014. – Vol. 27. – Режим доступу: https://proceedings.neurips.cc/paper_files/paper/2014/file/f033ed80deb0234979a61f95710dbe25-Paper.pdf. – Дата звернення: 16 червня 2025 р.

53. Makhzani A., Shlens J., Jaitly N., Goodfellow I., Frey B. Adversarial autoencoders // *arXiv preprint*. – 2016. – arXiv:1511.05644. – Режим доступу: <https://arxiv.org/abs/1511.05644>. – Дата звернення: 16 червня 2025 р.

54. Zhang X., Yu Q., Yu H. Physics Inspired Methods for Crowd Video Surveillance and Analysis: A Survey // *IEEE Access*. – 2018. – Vol. 6. – P. 66816–66830. – DOI: <https://doi.org/10.1109/ACCESS.2018.2878733>.

55. Govender D., Tapamo J.-R. Spatio-Temporal Scale Coded Bag-of-Words // *Sensors*. – 2020. – Vol. 20, No. 21. – Article number: 6380. – DOI: <https://doi.org/10.3390/s20216380>.

56. Berroukham A., Housni K., Lahraichi M., Boulfrifi I. Deep learning-based methods for anomaly detection in video surveillance: a review // *Bulletin of Electrical Engineering and Informatics*. – 2023. – Vol. 12, No. 1. – P. 314–327. – DOI: 10.11591/eei.v12i1.3944. – Режим доступу: <https://beei.org/index.php/EEI/article/view/3944>. – Дата звернення: 16 червня 2025 р.

57. Ferryman J. M., Shahrokni A. PETS2009: Dataset and challenge // 2009 Twelfth IEEE International Workshop on Performance Evaluation of Tracking and

Surveillance. – 2009. – P. 1–6. – Режим доступу: <https://api.semanticscholar.org/CorpusID:34646357>. – Дата звернення: 16 червня 2025 р.

58. Li A., Miao Z., Cen Y., Zhang Xiao-Ping, Zhang L., Chen S. Abnormal event detection in surveillance videos based on low-rank and compact coefficient dictionary learning // Pattern Recognition. – 2020. – Vol. 108. – Art. № 107355. – DOI: <https://doi.org/10.1016/j.patcog.2020.107355>.

59. Sikdar A., Chowdhury A. S. An adaptive training-less framework for anomaly detection in crowd scenes // Neurocomputing. – 2020. – Vol. 415. – P. 317–331. – DOI: <https://doi.org/10.1016/j.neucom.2020.07.058>.

60. Sharif M. H., Jiao L., Omlin C. W. Deep Crowd Anomaly Detection: State-of-the-Art, Challenges, and Future Research Directions [Електронний ресурс] // arXiv preprint. – 2022. – Режим доступу: <https://arxiv.org/abs/2210.13927>. – Дата звернення: 16 червня 2025 р.

61. Cecaј A., Lippi M., Mamei M., Zambonelli F. Sensing and Forecasting Crowd Distribution in Smart Cities: Potentials and Approaches // IoT. – 2021. – Vol. 2, No. 1. – P. 33–49. – DOI: 10.3390/iot2010003. – Режим доступу: <https://www.mdpi.com/2624-831X/2/1/3>. – Дата звернення: 16 червня 2025 р.

62. Niu Y., Shi W., Liu W., He S., Pan J., Chan A. B. Over-crowdedness Alert! Forecasting the Future Crowd Distribution // arXiv preprint. – 2020. – Vol. abs/2006.05127. – Режим доступу: <https://arxiv.org/abs/2006.05127>. – Дата звернення: 16 червня 2025 р.

63. Cardia M., Luca M., Pappalardo L. Enhancing crowd flow prediction in various spatial and temporal granularities [Електронний ресурс] // arXiv preprint. – 2022. – Режим доступу: <https://arxiv.org/abs/2203.07372>. – Дата звернення: 16 червня 2025 р.

64. Lai Y., Dzombak D. A. Use of the Autoregressive Integrated Moving Average (ARIMA) Model to Forecast Near-Term Regional Temperature and Precipitation // Weather and Forecasting. – 2020. – Vol. 35, No. 3. – P. 959–976. – DOI: 10.1175/WAF-D-19-0158.1. – Режим доступу:

<https://journals.ametsoc.org/view/journals/wefo/35/3/waf-d-19-0158.1.xml>. – Дата звернення: 16 червня 2025 р.

65. Rumelhart D. E., Hinton G. E., Williams R. J. Learning representations by back-propagating errors // *Nature*. – 1986. – Vol. 323, No. 6088. – P. 533–536. – DOI: <https://doi.org/10.1038/323533a0>.

66. Scarselli F., Gori M., Tsoi A. C., Hagenbuchner M., Monfardini G. The Graph Neural Network Model // *IEEE Transactions on Neural Networks*. – 2009. – Vol. 20, No. 1. – P. 61–80. – DOI: 10.1109/TNN.2008.2005605.

67. Junkai Sun, Junbo Zhang, Qiaofei Li, Xiuwen Yi, Yuxuan Liang, Yu Zheng. Predicting Citywide Crowd Flows in Irregular Regions Using Multi-View Graph Convolutional Networks // *arXiv preprint*. – 2019. – arXiv:1903.07789. – Режим доступу: <https://arxiv.org/abs/1903.07789>. – Дата звернення: 16 червня 2025 р.

68. Fan Z., Zhang H., Zhang Z., Lu G., Zhang Y., Wang Y. A survey of crowd counting and density estimation based on convolutional neural network // *Neurocomputing*. – 2022. – Vol. 472. – P. 224–251. – DOI: <https://doi.org/10.1016/j.neucom.2021.02.103>. – Дата звернення: 16 червня 2025 р.

69. Wang B., Liu H., Samaras D., Hoai M. Distribution Matching for Crowd Counting [Електронний ресурс]. – 2020. – Режим доступу: <https://arxiv.org/abs/2009.13077>. – Дата звернення: 16 червня 2025 р.

70. Zhao R., Hua F., Wei B., Li C., Ma Y., Wong E. S. W., Liu F. A Review of Abnormal Crowd Behavior Recognition Technology Based on Computer Vision // *Applied Sciences*. – 2024. – Vol. 14, No. 21. – Article No. 9758. – DOI: 10.3390/app14219758. – Дата звернення: 16 червня 2025 р.

71. Loy C. C., Chen K., Gong S., Xiang T. Crowd Counting and Profiling: Methodology and Evaluation // *Modeling, Simulation and Visual Analysis of Crowds: A Multidisciplinary Perspective* / ред. Ali S., Nishino K., Manocha D., Shah M. – New York, NY: Springer New York, 2013. – P. 347–382. – ISBN 978-1-4614-8483-7. – DOI: 10.1007/978-1-4614-8483-7_14. – Режим доступу: https://doi.org/10.1007/978-1-4614-8483-7_14. – Дата звернення: 16 червня 2025 р.

72. Pexels [Електронний ресурс]. – Режим доступу: <https://www.pexels.com>. – Дата звернення: 16 червня 2025 р.

73. Lempitsky, V., Zisserman, A. Learning To Count Objects in Images // Advances in Neural Information Processing Systems : [proc.]. – 2010. – Vol. 23. – Curran Associates, Inc. – Режим доступу: https://proceedings.neurips.cc/paper_files/paper/2010/file/fe73f687e5bc5280214e0486b273a5f9-Paper.pdf. – Дата звернення: 16 червня 2025 р.

74. Shen, Z., Xu, Y., Ni, B., Wang, M., Hu, J., Yang, X. Crowd Counting via Adversarial Cross-Scale Consistency Pursuit // 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition : [proc.]. – 2018. – P. 5245–5254. – DOI: 10.1109/CVPR.2018.00550.

75. Liu, F., Shen, C., Lin, G., Reid, I. Learning Depth from Single Monocular Images Using Deep Convolutional Neural Fields // IEEE Transactions on Pattern Analysis and Machine Intelligence. – 2016. – Vol. 38, No. 10. – P. 2024–2039. – DOI: 10.1109/TPAMI.2015.2505283.

76. Zhang, C., Li, H., Wang, X., Yang, X. Cross-scene crowd counting via deep convolutional neural networks // 2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). – 2015. – P. 833–841. – DOI: 10.1109/CVPR.2015.7298684.

77. Zhang, C., Woodland, P. C. Parameterised sigmoid and ReLU hidden activation functions for DNN acoustic modelling // Interspeech 2015. – 2015. – P. 3224–3228. – DOI: 10.21437/Interspeech.2015-649.

78. Xu, K., Ba, J., Kiros, R., Cho, K., Courville, A., Salakhutdinov, R., Zemel, R., Bengio, Y. Show, Attend and Tell: Neural Image Caption Generation with Visual Attention [Електронний ресурс] // arXiv preprint arXiv:1502.03044. – 2016. – Режим доступу: <https://arxiv.org/abs/1502.03044>. – Дата звернення: 16 червня 2025 р.

79. Simonyan, K., & Zisserman, A. Very deep convolutional networks for large-scale image recognition [Електронний ресурс] // arXiv preprint arXiv:1409.1556. – 2015. – Режим доступу: <https://arxiv.org/abs/1409.1556>. – Дата звернення: 16 червня 2025 р.

80. He K., Zhang X., Ren S., Sun J. Delving Deep into Rectifiers: Surpassing Human-Level Performance on ImageNet Classification // 2015 IEEE International Conference on Computer Vision (ICCV). – 2015. – P. 1026–1034. – DOI: <https://doi.org/10.1109/ICCV.2015.123>.

81. Sindagi, V. A., Patel, V. M. Generating High-Quality Crowd Density Maps Using Contextual Pyramid CNNs // 2017 IEEE International Conference on Computer Vision (ICCV). – 2017. – P. 1879–1888. – DOI: <https://doi.org/10.1109/ICCV.2017.206>.

82. Sam, D. B., Jiang, X., Xu, Z., Zhang, B., Zhang, X., Chen, X., Doermann, D., & Shao, L. Crowd Counting and Density Estimation by Trellis Encoder-Decoder Network [Електронний ресурс] // arXiv preprint. – 2019. – Режим доступу: <https://arxiv.org/abs/1903.00853>. – Дата звернення: 16 червня 2025 р.

83. Kim T., Oh J., Kim N., Cho S., Yun S.-Y. Comparing Kullback–Leibler Divergence and Mean Squared Error Loss in Knowledge Distillation [Електронний ресурс] // arXiv preprint. – 2021. – Режим доступу: <https://arxiv.org/abs/2105.08919>. – Дата звернення: 16 червня 2025 р.

84. Cho S., Huang J., Kim S., Lee J.-Y. FlowTrack: Revisiting Optical Flow for Long-Range Dense Tracking // 2024 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). – 2024. – P. 19268–19277. – DOI: <https://doi.org/10.1109/CVPR52733.2024.01823>.

85. Shao J., Loy C. C., Wang X. Scene-Independent Group Profiling in Crowd // 2014 IEEE Conference on Computer Vision and Pattern Recognition. – 2014. – P. 2227–2234. – DOI: <https://doi.org/10.1109/CVPR.2014.285>.

86. Sundararaman R., De Almeida Braga C., Marchand E., Pettre J. Tracking Pedestrian Heads in Dense Crowd [Електронний ресурс] // arXiv preprint. – 2021. – Режим доступу: <https://arxiv.org/abs/2103.13516>. – Дата звернення: 16 червня 2025 р.

87. Ali S., Shah M. Floor Fields for Tracking in High Density Crowd Scenes // Proceedings of the 10th European Conference on Computer Vision: Part II (ECCV'08, Marseille, France). – Berlin: Springer-Verlag, 2008. – С. 1–14. – DOI: https://doi.org/10.1007/978-3-540-88688-4_1.

88. Kass M., Witkin A., Terzopoulos D. Snakes: Active contour models // International Journal of Computer Vision. – 1988. – Vol. 1, No. 4. – P. 321–331. – DOI: <https://doi.org/10.1007/BF00133570>.
89. Vasilevskiy A., Siddiqi K. Flux maximizing geometric flows // IEEE Transactions on Pattern Analysis and Machine Intelligence. – 2002. – Vol. 24, No. 12. – P. 1565–1578. – DOI: <https://doi.org/10.1109/TPAMI.2002.1114849>.
90. Chan T.F., Vese L.A. Active contours without edges // IEEE Transactions on Image Processing. – 2001. – Vol. 10, No. 2. – P. 266–277. – DOI: <https://doi.org/10.1109/83.902291>.
91. Harris C.G., Stephens M.J. A combined corner and edge detector // Alvey Vision Conference. – 1988. – [Электронный ресурс]. – Режим доступа: <https://api.semanticscholar.org/CorpusID:1694378>. – Дата звернення: 16 червня 2025 р.
92. Laptev I., Lindeberg T. Interest point detection and scale selection in space-time // Scale Space Methods in Computer Vision / ed. Griffin L.D., Lillholm M. – Berlin, Heidelberg : Springer Berlin Heidelberg, 2003. – P. 372–387. – ISBN 978-3-540-44935-5. – DOI: https://doi.org/10.1007/3-540-44935-3_26.
93. Tomasi C. Detection and tracking of point features [Электронный ресурс] // 1991. – Режим доступа: <https://api.semanticscholar.org/CorpusID:238434334>. – Дата звернення: 16 червня 2025 р.
94. Gupta S., Thakar U., Tokekar S. A comprehensive survey on techniques for numerical similarity measurement // Expert Systems with Applications. – 2025. – Vol. 277. – P. 127235. – DOI: <https://doi.org/10.1016/j.eswa.2025.127235>.
95. Ester M., Kriegel H.-P., Sander J., Xu X. A density-based algorithm for discovering clusters in large spatial databases with noise // Proceedings of the Second International Conference on Knowledge Discovery and Data Mining (KDD'96), Portland, Oregon, USA. – AAAI Press, 1996. – P. 226–231.
96. Zhou B., Wang X., Tang X. Understanding collective crowd behaviors: Learning a Mixture model of Dynamic pedestrian-Agents // 2012 IEEE Conference on

Computer Vision and Pattern Recognition. – 2012. – P. 2871–2878. – DOI: <https://doi.org/10.1109/CVPR.2012.6248013>.

97. Wang X., Ma X., Grimson W.E.L. Unsupervised Activity Perception in Crowded and Complicated Scenes Using Hierarchical Bayesian Models // IEEE Transactions on Pattern Analysis and Machine Intelligence. – 2009. – Vol. 31, No. 3. – P. 539–555. – DOI: <https://doi.org/10.1109/TPAMI.2008.87>.

98. Benabbas Y., Ihaddadene N., Djeraba C. Motion Pattern Extraction and Event Detection for Automatic Visual Surveillance // EURASIP Journal on Image and Video Processing. – 2011. – Vol. 2011, No. 1. – P. 163682. – DOI: <https://doi.org/10.1155/2011/163682>.

99. Krausz B., Bauckhage C. Loveparade 2010: Automatic video analysis of a crowd disaster // Computer Vision and Image Understanding. – 2012. – Vol. 116, No. 3. – P. 307–319. – DOI: <https://doi.org/10.1016/j.cviu.2011.08.006>.

100. Rao A. S., Gubbi J., Marusic S., Palaniswami M. Crowd Event Detection on Optical Flow Manifolds // IEEE Transactions on Cybernetics. – 2016. – Vol. 46, No. 7. – P. 1524–1537. – DOI: <https://doi.org/10.1109/TCYB.2015.2451136>.

101. Mehran R., Moore B. E., Shah M. A streakline representation of flow in crowded scenes // Proceedings of the 11th European Conference on Computer Vision: Part III (ECCV'10), Heraklion, Crete, Greece, 2010. – Berlin : Springer-Verlag, 2010. – P. 439–452.

102. Gu X., Cui J., Zhu Q. Abnormal crowd behavior detection by using the particle entropy // Optik. – 2014. – Vol. 125, No. 14. – P. 3428–3433. – DOI: <https://doi.org/10.1016/j.ijleo.2014.01.041>.

103. Lalit R., Purwar R. K., Verma S., Jain A. Crowd abnormality detection in video sequences using supervised convolutional neural network // Multimedia Tools and Applications. – 2022. – Vol. 81, No. 4. – P. 5259–5277. – DOI: <https://doi.org/10.1007/s11042-021-11781-4>.

104. Loy, C. C., Gong, S., Xiang, T. From Semi-supervised to Transfer Counting of Crowds // 2013 IEEE International Conference on Computer Vision (ICCV). – 2013. – P. 2256–2263. – DOI: [10.1109/ICCV.2013.270](https://doi.org/10.1109/ICCV.2013.270).

105. Idrees H., Saleemi I., Seibert C., Shah M. Multi-source multi-scale counting in extremely dense crowd images. In: 2013 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2013. P. 2547–2554. DOI: 10.1109/CVPR.2013.329.
106. Sam D. B., Surya S., Babu R. V. Switching Convolutional Neural Network for Crowd Counting. arXiv preprint arXiv:1708.00199, 2017. Режим доступу: <https://arxiv.org/abs/1708.00199>. – Дата звернення: 16 червня 2025 р.
107. Liu J., Gao C., Meng D., Hauptmann A. G. DecideNet: Counting Varying Density Crowds Through Attention Guided Detection and Density Estimation. arXiv preprint arXiv:1712.06679, 2018. Режим доступу: <https://arxiv.org/abs/1712.06679>. – Дата звернення: 16 червня 2025 р.
108. Shi Z., Zhang L., Liu Y., Cao X., Ye Y., Cheng M.-M., Zheng G. Crowd Counting With Deep Negative Correlation Learning. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), June 2018.
109. Lei Y., Liu Y., Zhang P., Liu L. Towards using count-level weak supervision for crowd counting. Pattern Recognition, 2021, vol. 109, p. 107616. doi: <https://doi.org/10.1016/j.patcog.2020.107616>.
110. Cao X., Wang Z., Zhao Y., Su F. Scale Aggregation Network for Accurate and Efficient Crowd Counting. In: Ferrari V., Hebert M., Sminchisescu C., Weiss Y. (eds.) Computer Vision – ECCV 2018. Cham: Springer International Publishing, 2018. P. 757–773. doi: https://doi.org/10.1007/978-3-030-01228-1_45.
111. Chen Z., Cheng J., Yuan Y., Liao D., Li Y., Lv J. Deep Density-aware Count Regressor. arXiv preprint arXiv:1908.03314, 2020. Режим доступу: <https://arxiv.org/abs/1908.03314>. – Дата звернення: 16 червня 2025 р.
112. Gao J., Wang Q., Li X. PCC Net: Perspective Crowd Counting via Spatial Convolutional Network. arXiv preprint arXiv:1905.10085, 2019. Режим доступу: <https://arxiv.org/abs/1905.10085>. – Дата звернення: 16 червня 2025 р.
113. Gu S., Lian Z. A unified RGB-T crowd counting learning framework. Image and Vision Computing, 2023, vol. 131, p. 104631. doi: <https://doi.org/10.1016/j.imavis.2023.104631>.

114. Tang H., Wang Y., Chau L.-P. TAFNet: A Three-Stream Adaptive Fusion Network for RGB-T Crowd Counting. arXiv preprint arXiv:2202.08517, 2022. Режим доступу: <https://arxiv.org/abs/2202.08517>. – Дата звернення: 16 червня 2025 р.

115. Powers D. M. W. Evaluation: from precision, recall and F-measure to ROC, informedness, markedness and correlation. arXiv preprint arXiv:2010.16061, 2020. Режим доступу: <https://arxiv.org/abs/2010.16061>. – Дата звернення: 16 червня 2025 р.

116. Ali S., Shah M. A Lagrangian Particle Dynamics Approach for Crowd Flow Segmentation and Stability Analysis. In: 2007 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2007. P. 1–6. doi: <https://doi.org/10.1109/CVPR.2007.382977>.

117. Zou Y., Zhao X., Liu Y. Measuring Crowd Collectiveness by Macroscopic and Microscopic Motion Consistencies. IEEE Transactions on Multimedia, 2018, vol. 20, no. 12, pp. 3311–3323. doi: <https://doi.org/10.1109/TMM.2018.2832601>.

118. Hassner T., Itcher Y., Kliper-Gross O. Violent flows: Real-time detection of violent crowd behavior. In: 2012 IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops (CVPRW), 2012. P. 1–6. doi: <https://doi.org/10.1109/CVPRW.2012.6239348>.

119. Biswas S., Praveen R. G., Babu R. V. Super-pixel based crowd flow segmentation in H.264 compressed videos. In: 2014 IEEE International Conference on Image Processing (ICIP), 2014. P. 2319–2323. doi: <https://doi.org/10.1109/ICIP.2014.7025470>.

120. Kruthiventi S. S. S., Babu R. V. Crowd Flow Segmentation in Compressed Domain using CRF. arXiv preprint arXiv:1506.06006, 2015. Режим доступу: <https://arxiv.org/abs/1506.06006>. – Дата звернення: 16 червня 2025 р.

121. Wu S., Wong H. S. Crowd motion partitioning in a scattered motion field. IEEE Transactions on Systems, Man, and Cybernetics, Part B: Cybernetics, 2012, vol. 42, no. 5, pp. 1443–1454. doi: <https://doi.org/10.1109/TSMCB.2012.2192267>.

122. Zhou B., Tang X., Wang X. Coherent Filtering: Detecting Coherent Motions from Crowd Clutters. In: Fitzgibbon A., Lazebnik S., Perona P., Sato Y., Schmid C. (eds.)

Computer Vision – ECCV 2012. Berlin, Heidelberg: Springer Berlin Heidelberg, 2012. P. 857–871. doi: https://doi.org/10.1007/978-3-642-33709-3_61.

123. Zhou B., Tang X., Wang X. Measuring Crowd Collectiveness. In: 2013 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2013. P. 3049–3056. doi: <https://doi.org/10.1109/CVPR.2013.392>.

124. Brox T., Malik J. Object segmentation by long term analysis of point trajectories. In: Proceedings of the 11th European Conference on Computer Vision (ECCV 2010), Part V. Berlin, Heidelberg: Springer-Verlag, 2010. P. 282–295. DOI: <https://doi.org/10.5555/1888150.1888173>.

125. Wu Y., Wang Y., Jia Y. Adaptive diffusion flow active contours for image segmentation. Computer Vision and Image Understanding, 2013, vol. 117, no. 10, pp. 1421–1435. doi: <https://doi.org/10.1016/j.cviu.2013.05.003>.

126. Lin W., Mi Y., Wang W., Wu J., Wang J., Mei T. A diffusion and clustering-based approach for finding coherent motions and understanding crowd scenes. arXiv preprint arXiv:1602.04921, 2016. Режим доступа: <https://arxiv.org/abs/1602.04921>. – Дата звернення: 16 червня 2025 р.

127. Bird N., Atev S., Caramelli N., Martin R., Masoud O., Papanikolopoulos N. Real time, online detection of abandoned objects in public areas. In: Proceedings of the 2006 IEEE International Conference on Robotics and Automation (ICRA 2006), 2006. P. 3775–3780. DOI: <https://doi.org/10.1109/ROBOT.2006.1642279>.

128. Mehran R., Oyama A., Shah M. Abnormal crowd behavior detection using social force model. In: 2009 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2009. P. 935–942. doi: <https://doi.org/10.1109/CVPR.2009.5206641>.

129. Cong Y., Yuan J., Liu J. Abnormal event detection in crowded scenes using sparse representation. Pattern Recognition, 2013, vol. 46, no. 7, pp. 1851–1864. DOI: <https://doi.org/10.1016/j.patcog.2012.11.021>.

130. Wang L., Dong M. Real-time detection of abnormal crowd behavior using a matrix approximation-based approach. In: 2012 19th IEEE International Conference on

Image Processing (ICIP), 2012. P. 2701–2704. DOI: <https://doi.org/10.1109/ICIP.2012.6467456>.

131. Mahadevan V., Li W.-X., Bhalodia V., Vasconcelos N. Anomaly Detection in Crowded Scenes. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2010. Vol., pp.: doi: 10.1109/CVPR.2010.5539872.

132. Lloyd K., Rosin P. L., Marshall D., Moore S. C. Detecting violent and abnormal crowd activity using temporal analysis of grey level co-occurrence matrix (GLCM)-based texture measures. Machine Vision and Applications, 2017, vol. 28, no. 3–4, pp. 361–371. DOI: <https://doi.org/10.1007/s00138-017-0830-x>.

133. Serrano Gracia I., Deniz Suarez O., Bueno Garcia G., Kim T.-K. Fast Fight Detection. PLoS ONE, 2015, vol. 10, no. 4, e0120448. DOI: <https://doi.org/10.1371/journal.pone.0120448>.

134. Gao Y., Liu H., Sun X., Wang C., Liu Y. Violence detection using Oriented Violent Flows. Image and Vision Computing, 2016, vol. 48–49, pp. 37–41. DOI: <https://doi.org/10.1016/j.imavis.2016.01.006>.

135. Xu L., Gong C., Yang J., Wu Q., Yao L. Violent video detection based on MoSIFT feature and sparse coding. In: 2014 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), 2014. P. 3538–3542. DOI: <https://doi.org/10.1109/ICASSP.2014.6854259>.

ДОДАТОК А СПИСОК ПУБЛІКАЦІЙ ЗДОБУВАЧА ЗА ТЕМОЮ ДИСЕРТАЦІЇ

1. Добришев Р. Є. Візуальний аналіз натовпу в сучасних системах інтелектуального відеоспостереження // Наука і техніка сьогодні (Серія «Техніка»), 2024, Vol. 37, № 9. Р. 639 – 651. (Index Copernicus)

<https://perspectives.pp.ua/index.php/nts/article/view/15110>

2. Dobryshev R. Y. & Maksymov M. V. “Accurate crowd counting for intelligent video surveillance systems”. Herald of Advanced Information Technology. 2024; Vol. 7, No. 3: 253–261. (Index Copernicus)

<https://hait.op.edu.ua/index.php/journal/article/view/20>

3. Dobryshev R. Y. “Anomaly detection in crowded scenes: technologies, challenges and opportunities”. Applied Aspects of Information Technology. 2024; Vol. 7, No. 3: 219–230. (Index Copernicus)

<https://aait.op.edu.ua/index.php/journal/article/view/252>

4. Добришев Р. Є. Виявлення аномалій в багатолюдних сценах // Наука і техніка сьогодні (Серія «Техніка»), 2025, Vol. 46, № 5. Р. 1437 – 1449. (Index Copernicus)

<https://perspectives.pp.ua/index.php/nts/article/view/24510>

5. Dobryshev R., Purish S., Lobachev M., Hodovychenko M. Crowd counting in intelligent video surveillance systems. *Radioelectronic and Computer Systems*, 2025, №. 2, P. 132–142. (Scopus)

<http://nti.khai.edu/ojs/index.php/reks/article/view/reks.2025.2.08>

6. Добришев Р.Є., Лобачев М.В. Підрахунок людей у натовпі з урахуванням глибини сцени в задачах оцінки щільності // Proceedings of the 15th International Scientific Conference of students and young researchers «Modern Information Technology», MIT 2025, Odesa, Ukraine. May 15-16, 2025. С. 144–146.

https://drive.google.com/file/d/1zi1TabIH7iso1-SW__9ntZ_LG0glE4Wa/view?usp=drive_link

7. Добришев Р. Є. Інтелектуальне виявлення аномалій у натовпі: методи, проблеми та перспективи // The 7th International scientific and practical conference «Future of science: innovations and perspectives». SSPG Publish, Stockholm, Sweden, May 19 – 21, 2025. Pp. 172-176.

<https://sci-conf.com.ua/wp-content/uploads/2025/05/FUTURE-OF-SCIENCE-INNOVATIONS-AND-PERSPECTIVES-19-21.05.25.pdf>

8. Добришев Р. Є. Технології візуального аналізу натовпу в інтелектуальних системах відеоспостереження // Proceedings of the The 20th International scientific and practical conference «Innovative scientific research: latest theories, modern methods and practices». International Science Group, Seville, Spain, May 20 – 23, 2025. Pp. 387-389.

<https://isg-konf.com/wp-content/uploads/2025/05/INNOVATIVE-SCIENTIFIC-RESEARCH-LATEST-THEORIES-MODERN-METHODS-AND-PRACTICES.pdf>

9. Добришев Р. Є. Сучасні методи підрахунку людей у натовпі в системах інтелектуального відеоспостереження // Proceedings of the 2nd International Scientific and Practical Conference «Science and Information Technologies in the Modern World». International Scientific Unity, Athens, Greece, May 21 – 23, 2025. Pp. 694 –698.

https://isu-conference.com/wp-content/uploads/2025/05/Athens_Greece_21.05.25.pdf

10. Добришев Р. Є. Автоматизоване виявлення глобальних аномалій у динаміці натовпу // Proceedings of the 11th International scientific and practical conference «Scientific achievements of contemporary society». Cognum Publishing House. London, United Kingdom, May 29 – 31, 2025. Pp. 255–259.

<https://sci-conf.com.ua/wp-content/uploads/2025/05/SCIENTIFIC-ACHIEVEMENTS-OF-CONTEMPORARY-SOCIETY-29-31.05.25.pdf>

11. Добришев Р. Є. Сегментація потоків у високощільних натовпах за допомогою активного контуру та кластеризації траєкторій // Proceedings of the 1st International Scientific and Practical Conference «Modern Perspectives on Science and

Economic Progress». International Scientific Unity, Vilnius, Lithuania, June 4 – 6 , 2025.
Pp. 510 – 513.

https://isu-conference.com/wp-content/uploads/2025/06/Vilnius_Lithuania_4.06.25.pdf

ДОДАТОК Б АКТ ВПРОВАДЖЕННЯ РЕЗУЛЬТАТІВ У ТОВ «ПРОВЕКТУС ІТ»

Юридична та фактична адреса:
65014, м. Одеса, вул. Володимира Рутківського, 5, офіс 303Н
Код ЄДРПОУ 37281450
п/р UA763510050000026004446698800 в АТ "УкрСиббанк"
телефон +38 (067) 481-19-00

«ЗАТВЕРДЖУЮ»
Директор ТОВ «Провектус ІТ»
Ревенко Петро Георгійович
«12» червня 2025 року

АКТ випробування результатів дисертаційного дослідження Добришева Руслана Євгеновича на тему: «Моделі та методи візуального аналізу натовпу в системах інтелектуального відеоспостереження»

Розроблені в межах дисертаційного дослідження моделі та методи, спрямовані на вдосконалення аналізу натовпу в умовах високої щільності та складної динаміки сцени, були адаптовані до напрямів прикладної діяльності товариства з обмеженою відповідальністю «Провектус ІТ» (далі - Provectus) та пройшли успішне тестування в умовах реального середовища розробки інтелектуальних рішень у сфері відеоаналітики.

У межах випробування моделі аналізу щільності натовпу, траєкторій руху та аномальної поведінки стали частиною ініціатив компанії, орієнтованих на забезпечення безпеки в публічному просторі та розробку рішень на основі штучного інтелекту, які приносять практичну користь суспільству. Зокрема, протестовані методи демонструють потенціал для посилення функціональності систем, пов'язаних із моніторингом поведінки у місцях масового скупчення людей — як всередині інфраструктурних об'єктів, так і на відкритих територіях.

Випробуване використання моделей оцінки заповнюваності простору та інструментів виявлення аномальних ситуацій свідчить про можливість розширення потенціалу власних ML-рішень у напрямках smart city, охорони здоров'я, а також безпеки виробничих середовищ. В перспективі практичне впровадження результатів дисертації сприяє посиленню здатності компанії до швидкої адаптації передових наукових розробок до продуктів, які реально змінюють якість життя.

Таким чином, тестове випробування результатів дослідження в рамках технологічної практики Provestus підтвердила їхню прикладну значущість, ефективність, доцільність та відповідність високим вимогам сучасної індустрії штучного інтелекту. Разом з тим, компанія наразі розглядає можливість повноцінного та всебічного впровадження зазначених моделей та методів у свої продуктивні рішення.

Директор



Петро РЕВЕНКО

ДОДАТОК В АКТ ВПРОВАДЖЕННЯ РЕЗУЛЬТАТІВ У ТОВ «СФ ІНВЕСТ»**ТОВАРИСТВО З ОБМЕЖЕНОЮ ВІДПОВІДАЛЬНІСТЮ
«СФ ІНВЕСТ»**

65042, м. Одеса, вул. Балтська дорога, 6; ЄДРПОУ 44482912

Вих. № 12 від 10.06.25

«ЗАТВЕРДЖУЮ»
Директор ТОВ «СФ ІНВЕСТ»
«10» червня 2025 року**АКТ**
впровадження результатів дисертаційного дослідження
Добришева Руслана Євгеновича на тему:
«Моделі та методи візуального аналізу натовпу в системах інтелектуального
відеоспостереження»

Розроблені в межах дисертаційного дослідження моделі та методи візуального аналізу натовпу були адаптовані до потреб ТОВ "СФ Інвест" і впроваджені у практичну діяльність компанії з надання консалтингових та супровідних послуг у сферах комерційної діяльності, управління та безпеки бізнес-процесів.

Зокрема, результати дослідження були застосовані в рамках супроводу проектів, пов'язаних із організацією систем відеоаналітики на об'єктах з підвищеною відвідуваністю – бізнес-центрах, торгових площах, зонах очікування клієнтів. Моделі щільності натовпу та інструменти виявлення аномальної поведінки стали основою для формування рекомендацій щодо оптимізації просторового планування, зонування клієнтських потоків, а також для підвищення рівня безпеки в умовах обмеженого нагляду.

Інструментальні рішення, запропоновані в межах дослідження, використовуються компанією у консалтинговій діяльності для підготовки аналітичних висновків, оцінки ризиків у публічному просторі, а також для розробки технічних завдань до систем моніторингу у проектах корпоративних клієнтів. Їх використання дозволило підвищити обґрунтованість управлінських рішень та покращити рекомендації у сферах комерційної безпеки, управління об'єктами та клієнтського сервісу.

Таким чином, результати дисертаційного дослідження знайшли ефективне практичне застосування в комерційно-правовій діяльності ТОВ "СФ Інвест" як елемент сучасного інструментарію підтримки бізнесу.

Директор ТОВ «СФ ІНВЕСТ»



Євгеній ЧЕХОВСЬКИЙ

ДОДАТОК Г ДОВІДКА ПРО ВИКОРИСТАННЯ РЕЗУЛЬТАТІВ ДИСЕРТАЦІЙНОЇ РОБОТИ У НАВЧАЛЬНОМУ ПРОЦЕСІ НАЦІОНАЛЬНОГО УНІВЕРСИТЕТУ «ОДЕСЬКА ПОЛІТЕХНІКА»



ДОВІДКА

про використання результатів дисертаційної роботи
Добришева Руслана Євгеновича
«Моделі та методи візуального аналізу натовпу в системах
інтелектуального відеоспостереження»
у навчальному процесі
Національного університету «Одеська Політехніка»

Довідка видана стосовно того, що в програмах, навчально-методичних матеріалах та курсах лекцій з дисциплін, що вивчаються на кафедрі Штучного інтелекту та аналізу даних Інституту штучного інтелекту та робототехніки бакалаврами спеціальності 122 «Комп'ютерні науки» – «Інтелектуальний аналіз даних», використовуються наукові результати, отримані в дисертації Добришева Руслана Євгеновича:

– дисципліна «Загальні засади машинного навчання» – демонстрація процесу оцінки та валідації моделей на прикладі модифікованої нейромережевої моделі з використанням модулю інтеграції геометричних особливостей сцени. Тема «оцінка та валідація моделей: метрики, крос-валідація та вибір моделі»;

– дисципліна «Нейромережеві технології» – приклад використання нейромережевих моделей в задачах інтелектуального відеоспостереження та візуального аналізу натовпу. Тема «Перспективи розвитку та застосування нейронних мереж у різних галузях».

Використані результати дисертації Добришева Руслана Євгеновича «Моделі та методи візуального аналізу натовпу в системах інтелектуального відеоспостереження» свідчать про глибоке опрацювання наукової літератури за тематикою дисертації. Висновки та пропозиції відзначаються науковою новизною, можливістю реалізації в інформаційних системах, в яких потрібно здійснювати візуальний аналіз сцени із скупченням людей, тому істотно поліпшують матеріал означених дисциплін і сприяють підвищенню якості підготовки в галузі інтелектуального аналізу даних.

Перший проректор, проректор з наукової,
педагогічної та виховної роботи,
д.т.н., професор



С.А. Нестеренко

Голова методичної ради ІШІР,
к.т.н., доцент

П.В. Ступень

ДОДАТОК Д ДОВІДКА ПРО ВИКОРИСТАННЯ РЕЗУЛЬТАТІВ ДИСЕРТАЦІЙНОЇ РОБОТИ У НАУКОВО-ДОСЛІДНИЦЬКІЙ ДІЯЛЬНОСТІ НАЦІОНАЛЬНОГО УНІВЕРСИТЕТУ «ОДЕСЬКА ПОЛІТЕХНІКА»



ДОВІДКА

про використання результатів дисертаційної роботи
Добришева Руслана Євгеновича
«Моделі та методи візуального аналізу натовпу в системах інтелектуального
відеоспостереження»
у науково-дослідницькій діяльності
Національного університету «Одеська Політехніка»

Довідка видана в тому, що у науково-дослідницькій діяльності
Національного університету «Одеська Політехніка» використані наступні наукові
результати, отримані у дисертаційній роботі Добришева Руслана Євгеновича:

– метод інтеграції геометричних особливостей сцени до глибинних моделей
підрахунку людей у натовпі на основі використання карти глибини та механізму
просторової уваги;

– метод виявлення аномальних ситуацій у сценах з великою кількістю людей
шляхом аналізу поведінкових патернів натовпу з використанням ентропійних
характеристик орієнтованих напрямків та метрики часової зміни заповнюваності.

Дисертація є складовою частиною НДР № 237-177 «Програмні системи
машинного навчання та їх застосування в галузі інтелектуального аналізу даних»
(№ держреєстрації 0123U102594).

Проректор



Дмитро ДМИТРИШИН

ДОДАТОК Е ВИХІДНИЙ КОД ПРОГРАМНИХ ІНСТРУМЕНТАЛЬНИХ ЗАСОБІВ

```
# app/routes/monitor_zones.py

from fastapi import APIRouter, Depends, HTTPException, status
from sqlalchemy.orm import Session
from typing import List

from app.db import models, crud, database
from app.schemas.monitor_zone import MonitorZoneCreate, MonitorZoneOut,
MonitorZoneUpdate
from app.dependencies import get_current_operator

router = APIRouter(
    prefix="/monitor-zones",
    tags=["Monitor Zones"],
    responses={404: {"description": "Not found"}},
)

# Dependency to get DB session
def get_db():
    db = database.SessionLocal()
    try:
        yield db
    finally:
        db.close()

@router.get("/", response_model=List[MonitorZoneOut])
def get_zones(skip: int = 0, limit: int = 100, db: Session = Depends(get_db),
user=Depends(get_current_operator)):
    return crud.get_monitor_zones(db, skip=skip, limit=limit)

@router.post("/", response_model=MonitorZoneOut,
status_code=status.HTTP_201_CREATED)
def create_zone(zone: MonitorZoneCreate, db: Session = Depends(get_db),
user=Depends(get_current_operator)):
    return crud.create_monitor_zone(db=db, zone=zone)

@router.get("/{zone_id}", response_model=MonitorZoneOut)
def read_zone(zone_id: int, db: Session = Depends(get_db),
user=Depends(get_current_operator)):
    db_zone = crud.get_monitor_zone(db, zone_id=zone_id)
    if db_zone is None:
        raise HTTPException(status_code=404, detail="Zone not found")
    return db_zone

@router.put("/{zone_id}", response_model=MonitorZoneOut)
def update_zone(zone_id: int, zone: MonitorZoneUpdate, db: Session =
Depends(get_db), user=Depends(get_current_operator)):
    db_zone = crud.get_monitor_zone(db, zone_id=zone_id)
    if db_zone is None:
        raise HTTPException(status_code=404, detail="Zone not found")
    return crud.update_monitor_zone(db=db, zone_id=zone_id, zone=zone)
```

```

@router.delete("/{zone_id}", status_code=status.HTTP_204_NO_CONTENT)
def delete_zone(zone_id: int, db: Session = Depends(get_db),
user=Depends(get_current_operator)):
    db_zone = crud.get_monitor_zone(db, zone_id=zone_id)
    if db_zone is None:
        raise HTTPException(status_code=404, detail="Zone not found")
    crud.delete_monitor_zone(db=db, zone_id=zone_id)

# app/routes/thresholds.py

from fastapi import APIRouter, Depends, HTTPException, status
from sqlalchemy.orm import Session

from app.db import crud, database
from app.schemas.threshold import ThresholdConfigCreate, ThresholdConfigOut,
ThresholdConfigUpdate
from app.dependencies import get_current_operator

router = APIRouter(
    prefix="/thresholds",
    tags=["Density Thresholds"],
    responses={404: {"description": "Not found"}},
)

# Dependency to get DB session
def get_db():
    db = database.SessionLocal()
    try:
        yield db
    finally:
        db.close()

@router.get("/", response_model=ThresholdConfigOut)
def get_threshold(db: Session = Depends(get_db),
user=Depends(get_current_operator)):
    config = crud.get_threshold_config(db)
    if not config:
        raise HTTPException(status_code=404, detail="Threshold config not found")
    return config

@router.post("/", response_model=ThresholdConfigOut,
status_code=status.HTTP_201_CREATED)
def create_threshold(config: ThresholdConfigCreate, db: Session = Depends(get_db),
user=Depends(get_current_operator)):
    existing = crud.get_threshold_config(db)
    if existing:
        raise HTTPException(status_code=400, detail="Threshold config already
exists")
    return crud.create_threshold_config(db=db, config=config)

@router.put("/", response_model=ThresholdConfigOut)
def update_threshold(config: ThresholdConfigUpdate, db: Session = Depends(get_db),
user=Depends(get_current_operator)):
    existing = crud.get_threshold_config(db)
    if not existing:
        raise HTTPException(status_code=404, detail="Threshold config not found")
    return crud.update_threshold_config(db=db, config=config)

# app/routes/detection.py

```

```

from fastapi import APIRouter, Depends, HTTPException, status
from sqlalchemy.orm import Session

from app.db import database, crud
from app.schemas.detection import DetectionRequest, DetectionResult, IncidentAlert
from app.services import detection_service
from app.dependencies import get_current_operator

router = APIRouter(
    prefix="/detection",
    tags=["Anomaly Detection"],
    responses={404: {"description": "Not found"}},
)

def get_db():
    db = database.SessionLocal()
    try:
        yield db
    finally:
        db.close()

@router.post("/analyze", response_model=DetectionResult)
def analyze_video(request: DetectionRequest, db: Session = Depends(get_db),
user=Depends(get_current_operator)):
    """
    Аналіз поточного кадру чи потоку на наявність аномальної поведінки.
    """
    result = detection_service.analyze_stream(
        camera_id=request.camera_id,
        zones=request.zones,
        timestamp=request.timestamp,
        db=db
    )
    if not result:
        raise HTTPException(status_code=400, detail="Аналіз не виконано")
    return result

@router.post("/alert", response_model=IncidentAlert,
status_code=status.HTTP_201_CREATED)
def send_alert(alert: IncidentAlert, db: Session = Depends(get_db),
user=Depends(get_current_operator)):
    """
    Створення інциденту після виявлення аномалії та надсилання оповіщення.
    """
    try:
        record = detection_service.process_alert(alert=alert, db=db)
    except Exception as e:
        raise HTTPException(status_code=500, detail=str(e))
    return record

class DetectionRequest(BaseModel):
    camera_id: int
    zones: List[int]
    timestamp: Optional[datetime] = None

class DetectionResult(BaseModel):
    anomaly_detected: bool
    type: Optional[str]
    density_score: float
    heatmap_path: Optional[str]
    frame_image_path: Optional[str]

```

```

        time: datetime

class IncidentAlert(BaseModel):
    camera_id: int
    zone_id: int
    description: str
    image_path: str
    timestamp: datetime

# app/routes/reports.py

from fastapi import APIRouter, Depends, HTTPException, Query
from sqlalchemy.orm import Session
from datetime import datetime
from typing import List, Optional

from app.db import database, crud
from app.schemas.reports import EventLogOut, ReportRequest, ReportResponse
from app.services import reporting_service
from app.dependencies import get_current_operator

router = APIRouter(
    prefix="/reports",
    tags=["Reports and History"],
    responses={404: {"description": "Not found"}},
)

def get_db():
    db = database.SessionLocal()
    try:
        yield db
    finally:
        db.close()

@router.get("/events", response_model=List[EventLogOut])
def get_event_history(
    camera_id: Optional[int] = None,
    date_from: Optional[datetime] = Query(None),
    date_to: Optional[datetime] = Query(None),
    db: Session = Depends(get_db),
    user=Depends(get_current_operator),
):
    """
    Отримати журнал подій за камерою та часовим діапазоном.
    """
    events = crud.get_event_logs(db, camera_id=camera_id, date_from=date_from,
    date_to=date_to)
    return events

@router.post("/generate", response_model=ReportResponse)
def generate_report(request: ReportRequest, db: Session = Depends(get_db),
    user=Depends(get_current_operator)):
    """
    Згенерувати аналітичний звіт за вибраними параметрами.
    """
    try:
        report = reporting_service.generate_report(
            db=db,
            camera_ids=request.camera_ids,
            date_from=request.date_from,
            date_to=request.date_to,

```

```

        metrics=request.metrics,
        output_format=request.output_format
    )
    return report
except Exception as e:
    raise HTTPException(status_code=500, detail=str(e))

from pydantic import BaseModel
from typing import List, Optional
from datetime import datetime

class EventLogOut(BaseModel):
    id: int
    camera_id: int
    zone_id: int
    timestamp: datetime
    event_type: str
    description: Optional[str]
    image_path: Optional[str]

class ReportRequest(BaseModel):
    camera_ids: List[int]
    date_from: datetime
    date_to: datetime
    metrics: List[str] # нап. ["avg_density", "anomaly_count"]
    output_format: str # "pdf", "csv", "json"

class ReportResponse(BaseModel):
    report_path: str
    report_type: str
    created_at: datetime

# app/routes/external_api.py

from fastapi import APIRouter, Depends, HTTPException, Request, Header, Query
from sqlalchemy.orm import Session
from typing import Optional

from app.db import database, crud
from app.services.external_api_service import handle_external_query
from app.schemas.external import ExternalAPIResponse
from app.dependencies import verify_api_key

router = APIRouter(
    prefix="/external-api",
    tags=["External API"],
    responses={403: {"description": "Unauthorized"}},
)

def get_db():
    db = database.SessionLocal()
    try:
        yield db
    finally:
        db.close()

@router.get("/query", response_model=ExternalAPIResponse)
def external_query(
    request: Request,
    camera_id: Optional[int] = Query(None),
    date_from: Optional[str] = Query(None),
    date_to: Optional[str] = Query(None),

```

```

x_api_key: str = Header(...),
db: Session = Depends(get_db),
):
    """
    Обробка зовнішнього запиту до API.
    """
    if not verify_api_key(x_api_key):
        raise HTTPException(status_code=403, detail="Invalid API key")

    try:
        response = handle_external_query(
            db=db,
            camera_id=camera_id,
            date_from=date_from,
            date_to=date_to,
            client_ip=request.client.host,
            api_key=x_api_key
        )
        return response
    except Exception as e:
        raise HTTPException(status_code=500, detail=f"Server error: {str(e)}")

from pydantic import BaseModel
from typing import Optional, List
from datetime import datetime

class ExternalAPIResponse(BaseModel):
    status: str
    camera_id: Optional[int]
    statistics: Optional[dict]
    event_count: Optional[int]
    logs: Optional[List[str]]
    timestamp: datetime

import datetime
from app.db import crud
from app.utils.logger import log_external_call

def handle_external_query(db, camera_id, date_from, date_to, client_ip, api_key):
    # 1. Журналювання виклику
    log_external_call(ip=client_ip, key=api_key, time=datetime.datetime.utcnow())

    # 2. Підготовка даних
    logs = crud.get_event_logs(db, camera_id=camera_id, date_from=date_from,
date_to=date_to)

    stats = {
        "total_events": len(logs),
        "anomalies": sum(1 for log in logs if log.event_type == "anomaly"),
    }

    # 3. Повернення відповіді
    return {
        "status": "ok",
        "camera_id": camera_id,
        "statistics": stats,
        "event_count": len(logs),
        "logs": [f"{l.timestamp} - {l.event_type}" for l in logs],
        "timestamp": datetime.datetime.utcnow()
    }

# app/services/zone_service.py

```

```

from sqlalchemy.orm import Session
from typing import List, Optional

from app.db import crud
from app.schemas.monitor_zone import MonitorZoneCreate, MonitorZoneUpdate
from app.db.models import MonitorZone

def list_zones(db: Session, skip: int = 0, limit: int = 100) -> List[MonitorZone]:
    """
    Отримати список усіх зон моніторингу з пагінацією.
    """
    return crud.get_monitor_zones(db, skip=skip, limit=limit)

def get_zone_by_id(db: Session, zone_id: int) -> Optional[MonitorZone]:
    """
    Отримати зону моніторингу за її ідентифікатором.
    """
    return crud.get_monitor_zone(db, zone_id)

def create_zone(db: Session, zone_data: MonitorZoneCreate) -> MonitorZone:
    """
    Створити нову зону моніторингу.
    """
    return crud.create_monitor_zone(db=db, zone=zone_data)

def update_zone(db: Session, zone_id: int, zone_data: MonitorZoneUpdate) -> MonitorZone:
    """
    Оновити наявну зону моніторингу.
    """
    return crud.update_monitor_zone(db=db, zone_id=zone_id, zone=zone_data)

def delete_zone(db: Session, zone_id: int) -> None:
    """
    Видалити зону моніторингу.
    """
    crud.delete_monitor_zone(db=db, zone_id=zone_id)

# app/services/detection_service.py

from sqlalchemy.orm import Session
from datetime import datetime
from typing import Optional

from app.schemas.detection import DetectionRequest, DetectionResult, IncidentAlert
from app.db import crud
from app.core import crowd_density_model, anomaly_detector, heatmap_generator
from app.utils.logger import log_event

def analyze_stream(camera_id: int, zones: list, timestamp: Optional[datetime], db: Session) -> DetectionResult:
    """
    Аналіз відеопотоку камери в заданих зонах: оцінка щільності, виявлення аномалій.
    """
    # 1. Отримати кадр з камери (умовно, тут мав би бути відеопотік або API-кадр)
    frame = _get_latest_frame(camera_id)
    if frame is None:

```

```

        raise ValueError("Кадр не отримано")

# 2. Обробити кадр: оцінка щільності натовпу
density_score = crowd_density_model.predict_density(frame, zones)

# 3. Генерація теплової карти (для звітів)
heatmap_path = heatmap_generator.generate_heatmap(frame, zones)

# 4. Виявлення аномалій
anomaly_type = anomaly_detector.detect_anomaly(frame, zones)

# 5. Логування
log_event(
    db=db,
    camera_id=camera_id,
    zone_ids=zones,
    timestamp=timestamp or datetime.utcnow(),
    event_type="anomaly" if anomaly_type else "normal",
    description=anomaly_type or "no anomaly",
)

return DetectionResult(
    anomaly_detected=bool(anomaly_type),
    type=anomaly_type,
    density_score=density_score,
    heatmap_path=heatmap_path,
    frame_image_path=frame, # умовно, шлях до збереженого кадру
    time=timestamp or datetime.utcnow()
)

def process_alert(alert: IncidentAlert, db: Session):
    """
    Створення запису про інцидент після виявлення аномалії.
    """
    # 1. Зберегти інцидент у БД
    incident = crud.create_incident(
        db=db,
        camera_id=alert.camera_id,
        zone_id=alert.zone_id,
        description=alert.description,
        image_path=alert.image_path,
        timestamp=alert.timestamp
    )

    # 2. Надіслати сповіщення (псевдоінтеграція)
    _send_alert_to_operator_interface(incident)
    _send_alert_to_external_platform(incident)

    # 3. Запис у системний журнал
    log_event(
        db=db,
        camera_id=alert.camera_id,
        zone_ids=[alert.zone_id],
        timestamp=alert.timestamp,
        event_type="incident",
        description=alert.description,
    )

    return incident

# ----- Внутрішні функції (імітація) -----

```

```

def _get_latest_frame(camera_id: int) -> Optional[str]:
    # Псевдореалізація: повертає шлях до збереженого зображення
    return f"/data/frames/camera_{camera_id}_{datetime.utcnow().isoformat()}.jpg"

def _send_alert_to_operator_interface(incident):
    # Псевдореалізація надсилання на UI
    print(f"[UI] Alert: camera {incident.camera_id}, zone {incident.zone_id}")

def _send_alert_to_external_platform(incident):
    # Псевдореалізація інтеграції з міською системою безпеки
    print(f"[API] Incident pushed to external system: {incident.id}")

# app/services/reporting_service.py

import os
import csv
import pandas as pd
from datetime import datetime
from sqlalchemy.orm import Session
from typing import List

import matplotlib.pyplot as plt

from app.db import crud
from app.schemas.reports import ReportResponse

def generate_report(
    db: Session,
    camera_ids: List[int],
    date_from: datetime,
    date_to: datetime,
    metrics: List[str],
    output_format: str = "pdf"
) -> ReportResponse:
    """
    Генерація звіту за заданими параметрами: камера, час, метрики, формат.
    """

    # 1. Отримання подій із БД
    events = crud.get_event_logs(db, camera_ids=camera_ids, date_from=date_from,
    date_to=date_to)
    if not events:
        raise ValueError("Немає подій для звіту")

    # 2. Перетворення в DataFrame
    df = pd.DataFrame([
        "timestamp": e.timestamp,
        "camera_id": e.camera_id,
        "zone_id": e.zone_id,
        "event_type": e.event_type
    ] for e in events])

    # 3. Обчислення метрик
    results = {}
    if "avg_density" in metrics:
        results["avg_density"] = round(df[df["event_type"] == "anomaly"].shape[0] /
df.shape[0], 3)
    if "anomaly_count" in metrics:
        results["anomaly_count"] = df[df["event_type"] == "anomaly"].shape[0]

```

```

if "event_count" in metrics:
    results["event_count"] = df.shape[0]

# 4. Побудова графіка (для PDF або PNG)
chart_path = None
if output_format == "pdf":
    chart_path = _generate_plot(df)

# 5. Формування звіту
timestamp_str = datetime.utcnow().strftime("%Y%m%d_%H%M%S")
report_path = f"reports/report_{timestamp_str}.{output_format}"

if output_format == "csv":
    df.to_csv(report_path, index=False)
elif output_format == "json":
    df.to_json(report_path, orient="records", indent=2)
elif output_format == "pdf":
    _export_pdf(report_path, results, chart_path)
else:
    raise ValueError(f"Unsupported output format: {output_format}")

return ReportResponse(
    report_path=report_path,
    report_type=output_format,
    created_at=datetime.utcnow()
)

def _generate_plot(df: pd.DataFrame) -> str:
    """
    Побудова графіка частоти подій.
    """
    fig, ax = plt.subplots(figsize=(8, 4))
    df["timestamp"] = pd.to_datetime(df["timestamp"])
    df.set_index("timestamp", inplace=True)
    df.resample("1H").size().plot(ax=ax)

    ax.set_title("Щогодинна кількість подій")
    ax.set_xlabel("Час")
    ax.set_ylabel("Кількість подій")
    fig.tight_layout()

    chart_path =
f"reports/event_plot_{datetime.utcnow().strftime('%Y%m%d_%H%M%S')}.png"
    plt.savefig(chart_path)
    plt.close(fig)

    return chart_path

def _export_pdf(path: str, metrics: dict, chart_path: str):
    """
    Експортує звіт у PDF (простий варіант через matplotlib + text).
    """
    from matplotlib.backends.backend_pdf import PdfPages

    with PdfPages(path) as pdf:
        # Сторінка 1 - текст
        fig, ax = plt.subplots(figsize=(8.3, 11.7))
        ax.axis('off')
        lines = ["Аналітичний звіт", "", f"Дата створення: {datetime.utcnow().isoformat()}"]

```

```

for key, value in metrics.items():
    lines.append(f"{key}: {value}")

ax.text(0.1, 0.9, "\n".join(lines), fontsize=12, verticalalignment='top')
pdf.savefig(fig)
plt.close(fig)

# Сторінка 2 - графік
if chart_path and os.path.exists(chart_path):
    img = plt.imread(chart_path)
    fig, ax = plt.subplots(figsize=(8.3, 5))
    ax.imshow(img)
    ax.axis('off')
    pdf.savefig(fig)
    plt.close(fig)

# app/services/system_monitor.py

import os
import psutil
import datetime
from sqlalchemy.orm import Session

from app.db import crud
from app.schemas.system_status import SystemStatus, LogEntry

def get_system_status(db: Session) -> SystemStatus:
    """
    Отримати поточний статус системи: CPU, пам'ять, активні потоки, критичні логи.
    """

    # 1. Використання ресурсів
    cpu_percent = psutil.cpu_percent(interval=1)
    mem = psutil.virtual_memory()
    memory_percent = mem.percent
    uptime = datetime.datetime.utcnow() - \
datetime.datetime.fromtimestamp(psutil.boot_time())

    # 2. Перевірка відеопотоків (умовна логіка)
    stream_status = crud.check_all_streams_status(db)

    # 3. Аналіз логів (вибірка помилок за останні 24 год)
    error_logs = crud.get_recent_errors(db, hours=24)

    return SystemStatus(
        cpu_load=cpu_percent,
        memory_usage=memory_percent,
        uptime=str(uptime),
        active_streams=stream_status["active"],
        inactive_streams=stream_status["inactive"],
        recent_errors=[
            LogEntry(timestamp=log.timestamp, level=log.level,
message=log.message)
            for log in error_logs
        ]
    )

def perform_maintenance_task(db: Session, task_name: str) -> str:
    """
    Виконання технічного завдання (наприклад, оновлення, очистка кешу).
    """

```

```

if task_name == "clear_cache":
    _clear_cache()
    message = "Кеш очищено"
elif task_name == "restart_streams":
    _restart_streams()
    message = "Відеопотоки перезапущено"
else:
    message = f"Невідома задача: {task_name}"

crud.log_maintenance_action(db, task=task_name, status=message)
return message

# ----- Внутрішні функції -----

def _clear_cache():
    cache_dir = "/tmp/crowdanalytics_cache"
    if os.path.exists(cache_dir):
        for f in os.listdir(cache_dir):
            try:
                os.remove(os.path.join(cache_dir, f))
            except Exception:
                continue

def _restart_streams():
    # Псевдореалізація: тут міг би бути виклик shell-скрипта, служби або API
    print("[INFO] Відеопотоки умовно перезапущено")

```