

## АНОТАЦІЯ

*Тимченко Б.І.* Нейромережеві методи аналізу планарних зображень в системах автоматизованого скринінгу. – Кваліфікаційна наукова праця на правах рукопису.

Дисертація на здобуття наукового ступеня доктора філософії за спеціальністю 122 – Комп'ютерні науки. – Одеський національний політехнічний університет МОНУ, Україна, Одеса, 2021.

У **вступі** обґрунтовано актуальність удосконалення систем автоматизованого скринінгу (виявлення та класифікації нетипових об'єктів або процесів) шляхом створення моделей наборів даних та методів їх генерації, а також моделей нейронних мереж та методів багатозадачного навчання з метою підвищення достовірності класифікації та сегментації планарних зображень без підвищення витрат часу. Визначено об'єкт, предмет, задачі і методи дослідження; показано зв'язок з науково-дослідними роботами; наведено наукову новизну та практичне значення одержаних результатів; висвітлено особистий внесок здобувача.

В **першому розділі** дисертаційної роботи проведено аналіз проблем аналізу планарних зображень (які не мають виміру глибини, та в яких можна знехтувати масштабом об'єктів) в системах автоматизованого скринінгу на основі нейромережевих технологій.

Проаналізовано існуючі системи автоматизованого скринінгу в предметних областях медицини, метеорології та віддаленого зондування та особливості їх функціонування шляхом аналізу планарних зображень. Показано, що ефективність роботи таких систем безпосередньо залежить від достовірності класифікації та сегментації зображень, оскільки хибно-позитивні результати можуть призвести до передчасного реагування, а хибно-негативні

результати - до пропуску наявних проблем та пізнього реагування.

Крім того, не менш важливим фактором, що впливає на ефективність скринінгових систем є витрати часу на реалізацію процесу навчання, що в свою чергу пов'язано з якістю попередньо створених навчальних вибірок даних. Але, в задачах скринінгу, отримання достатньої кількості даних та проведення якісного анотування професіоналами є, як правило, затратним, а іноді неможливим через малу кількість та різний вигляд навчальних прикладів, суб'єктивність під час формування анотацій в наборах даних. Ці фактори негативно впливають на достовірність класифікації та сегментації при навчанні глибоких нейронних мереж стандартними методами. Для усунення такого впливу запропоновано при побудові моделей класифікації та сегментації вирішити задачу навчання з урахуванням частково-помилкових анотацій.

Показано, що в таких умовах доцільне використання глибоких нейронних мереж, через можливість автоматизації процесу навчання та застосування методів багатозадачного навчання, які підвищують достовірність класифікації та сегментації.

Проведений аналіз наявних моделей наборів даних та методів їх генерації показав обмеженість моделей наборів даних, які б дозволяли формувати частково-помилкові анотації, які характерні для реальних задач автоматизованого скринінгу.

Таким чином, вирішення важливої науково-практичної задачі підвищення достовірності класифікації та сегментації в задачах аналізу планарних зображень для систем автоматизованого скринінгу шляхом удосконалення моделей нейронних мереж та методів їх багатозадачного навчання, особливо за умови частково-помилково анотованих даних навчальної вибірки.

У **другому розділі** розроблено метод генерації наборів даних для забезпечення можливості тестування нейронних мереж при навчанні з використанням частково-помилкових анотацій та запропоновано параметричну

формалізацію моделі набору даних із частково-помилковими анотаціями, які характерні для реальних задач автоматизованого скринінгу.

Оскільки одержання достовірних анотацій класифікації та сегментації для тестувальних вибірок часто є неможливим або дуже затратним, запропоновано використання модельних наборів даних для оцінки достовірності моделей та методів навчання глибоких нейронних мереж. Такі набори даних дають можливість проводити навчання та тестування в контрольованих умовах шляхом штучного додавання помилок до анотацій навчальної вибірки даних (як правило, тестова вибірка залишається з достовірними анотаціями).

Запропонована модель набору даних  $\mathcal{M}$  має наступне представлення:

$$\mathcal{M} \in \{\mathcal{X}_b, \mathcal{X}_{tex}, N, S_{img}, S_{obj}, \Delta_{max}, N_{obj}, P_e, P_d, S_e, S_d\} \quad (1)$$

де  $\mathcal{X}_b$  - набір зображень фону,  $\mathcal{X}_f$  - набір об'єктів,  $\mathcal{X}_{tex}$  - набір зображень текстур об'єктів,  $N$  - кількість зображень в генерованому наборі даних,  $S_{img}$  - розмір генерованих зображень в пікселях,  $S_{obj}$  - середній розмір об'єкта в пікселях,  $\Delta_{max}$  - максимальне відхилення розміру об'єкта в відсотках,  $N_{obj}$  - максимальна кількість об'єктів на зображенні,  $P_e$  та  $P_d$  - ймовірності зменшення та збільшення маски кожного з об'єктів,  $S_e$  та  $S_d$  - допустимі масштаби збільшення та зменшення масок всіх об'єктів.

Останні чотири параметри введено для контрольованого створення помилок в анотаціях.

Сформульовано **перший пункт наукової новизни**: вперше запропоновано параметричну формалізацію моделі набору даних із частково-помилковими анотаціями, які характерні для реальних задач автоматизованого скринінгу, що дозволило розробити метод генерації навчальних, тестових та валідаційних наборів даних.

На основі запропонованої параметричної моделі (1) розроблено метод генерації наборів даних із частково-помилковими анотаціями, який містить

наступні кроки:

1. Вибрати випадкове зображення фону:  $x_{bg} \sim \mathcal{X}_b$
2. Вибрати кількість об'єктів на зображенні:  $N_{obj} \sim \mathcal{U}(1, N_{obj})$
3. Провести ініціалізацію маски сегментації:  $M = 0$  так що  $M \in \mathcal{R}^{C \times S_{img} \times S_{img}}$
4. Виконати наступні кроки  $n_{obj}$  разів:
  - 4.1 Вибрати розміри об'єкта:  $s \sim \mathcal{U}(S_{obj} - \Delta_{max}, S_{obj} + \Delta_{max})$
  - 4.2 Вибрати координати розміщення об'єкта:

$$i_f \sim \mathcal{U}(0, S_{img} - s)$$

$$j_f \sim \mathcal{U}(0, S_{img} - s)$$

4.3 Вибрати зображення об'єкта  $x_{fg} \sim \mathcal{X}_f$  та відповідний клас об'єкта  $c_{fg} \sim \mathcal{Y}_f$

4.4 Змінити розмір зображення об'єкта за допомогою білінійної інтерполяції:

$$\hat{x}_{fg} = R_{bilinear}(x_{fg})$$

4.5 Вибрати зображення текстури  $x_{tex} \sim \mathcal{X}_{tex}$

4.6 Модифікувати зображення об'єкта за допомогою текстури:

$$\hat{x}_{fg} = x_{fg} \circ x_{tex}[i_f : i_f + s, j_f : j_f + s]$$

4.7 Розмістити зображення об'єкта на зображенні фону:

$$x_{bg}[i_f : i_f + s, j_f : j_f + s] = (1 - x_{fg}) \circ x_{bg} + \hat{x}_{fg}$$

4.8 Сформувати маску сегментації об'єкта:

$$M_{seg} = x_{fg} > \theta_{seg}$$

де  $\theta_{seg}$  - поріг бінаризації вихідного зображення об'єкта. Для набору даних MNIST  $\theta_{seg} = 0.2$ , для набору даних FashionMNIST  $\theta_{seg} = 0.1$ .

4.9 Модифікувати маску сегментації відповідно до необхідного рівня помилок:

$$M_{seg} = \begin{cases} M_{seg} \oplus K^{S_d \times S_d} & \text{якщо } p_d \sim \mathcal{U}(0, 1) < P_d \\ M_{seg} \ominus K^{S_e \times S_e} & \text{якщо } p_e \sim \mathcal{U}(0, 1) < P_e \end{cases}$$

де  $K^{S_d \times S_d}$  - матриця ядра  $K^{S_e \times S_e}$  - матриця ядра ерозії.

4.10 Розмістити модифіковану маску сегментації об'єкта на загальному зображенні маски сегментації:

$$M[c_{fg}, i_f : i_f + s, j_f : j_f + s] = \max \{M[c_{fg}, i_f : i_f + s, j_f : j_f + s], M_{seg}\}$$

4.11 Зберегти зображення  $x_{bg}$  та маску  $M$

## 5. Завершити генерацію

Таким чином, удосконалено метод генерації наборів даних із частково-помилковими анотаціями на основі параметричної моделі, що за рахунок генерації анотацій: частково-помилкових для тренувальної вибірки та достовірних - для тестової, дало можливість виконувати тестування впливу рівня помилок анотацій на роботу нейромережових методів сегментації та класифікації.

Метод дозволяє отримати безліч наборів даних із схожими характеристиками та використовувати непараметричні статистичні методи (бутстрепінг) для оцінки моделей в умовах відсутності реальних тренувальних даних. Метод становить **другий пункт наукової новизни**.

В **третьому розділі** дисертаційної роботи розроблено модель нейронної мережі та метод багатозадачного навчання для одночасного підвищення достовірності класифікації та сегментації без зниження оперативності.

Для реалізації методів багатозадачного навчання, запропоновано удосконалити модель глибоких нейронної мережі з використанням архітектури енкодер-декодер (UNet, LinkNet) введенням додаткового декодера з шаром нормалізації. Таким чином, вдосконалена модель складається з

енкодера та двох декодерів (для задач сегментації та класифікації відповідно).

Модель глибокої нейронної мережі представлено наступним виразом:

$$v_1, v_2 \dots v_n = F_{encoder}(x, \theta_{enc}) \quad (2)$$

$$M_{seg} = F_{seg}((v_1, v_2 \dots v_n), \theta_{seg}) \quad (3)$$

$$C_{cls} = F_{cls}((v_n), \theta_{cls}) \quad (4)$$

де  $\theta_{enc}$  - набір параметрів енкодера,  $\theta_{seg}$  та  $\theta_{cls}$  - набори параметрів декодерів сегментації та класифікації відповідно,  $F_{seg}$  та  $F_{cls}$  - нейронні мережі декодера сегментації та класифікації відповідно. Запропонована модель складає **третій пункт наукової новизни**.

Завдяки введенню додаткового декодера класифікації з шаром нормалізації, з'явилася можливість реалізації методів багатозадачного навчання глибоких нейронних мереж та передбачення результатів.

Запропоновано метод багатозадачного навчання нейронних мереж в умовах частково-помилкових анотацій навчальних даних, який спирається на використання задач, пов'язаних з оригінальною. Показано, що для задачі сегментації існує близька задача класифікації, для якої анотації навчальних даних є якіснішими, ніж для вихідної задачі сегментації.

Метод багатозадачного навчання складається з двох етапів.

**Етап 1: генерація похідної задачі.** В контексті багатозадачного навчання, задача класифікації зводиться до визначення набору сегментованих зразків (*англ.* multiple instance learning) на планарному зображенні, замість маркування кожного з об'єктів для всіх класів. При цьому анотація вихідного планарного зображення являє собою множину з одним, чи декількома визначеними об'єктами.

Для кожної з задач (класифікація та сегментації) окремо обчислюється функція втрат. Для навчання декодера сегментації використовується обме-

жена зверху функція втрат  $L_{seg}]$ , в той час як для декодера класифікації - звичайна  $L_{cls}$ , а загальне значення функції втрат визначається як середнє арифметичне між індивідуальними значеннями:

$$L_{total} = \frac{L_{seg}] + L_{cls}}{2} \quad (5)$$

Відповідно, загальний градієнт функції втрат буде також середнім арифметичним градієнтів складових частин  $\nabla L_{seg}]$  і  $\nabla L_{cls}$ :

$$\nabla L_{total} = \frac{\nabla L_{seg}] + \nabla L_{cls}}{2} \quad (6)$$

Таким чином, забезпечується наявність ненульових градієнтів для оновлення параметрів від хоча б однієї функції втрат для кожного вхідного прикладу.

**Етап 2: Введення обмежень.** Для зменшення впливу помилкової частини анотацій запропоновано ввести обмеження другого роду (зверху) до функції втрат для задачі сегментації з менш якісними анотаціями даних. Це дозволило при навчанні на декількох задачах зменшити вплив градієнтів функції втрат на прикладах з помилковими анотаціями:

$$\mathcal{L}] = \min(L, \theta)$$

де  $\theta$  - поріг обмеження функції втрат.

Введені обмеження представлені наступним виразом:

$$\nabla \min(L, \theta) = \begin{cases} 1, & \text{якщо } L \in (-\infty, \theta] \\ 0, & \text{якщо } L \in (\theta, \infty) \end{cases}$$

Метод прогнозування містить наступні етапи:

**Етап 1. Нормалізація.** Нехай  $C_{cls} \in \mathcal{R}^C$  та  $M_{seg} \in \mathcal{R}^{C \times H \times W}$  - результати декодерів класифікації та сегментації відповідно, значення яких знаходяться на проміжку  $(-\infty, +\infty)$  (логіти).

Для отримання результатів на проміжку  $[0, 1]$  використовується логістична сигмоїдна функція активації:

$$\sigma(x) = \frac{1}{1 + e^{-x}}$$

**Етап 2. Ренормалізація.** Ренормалізація полягає у зважуванні карти сегментації за допомогою нормованих логітів класифікатора. Першим кроком є трансформація логітів сегментації та класифікації в некалібровані оцінки на проміжку  $[0, 1]$ :

$$\begin{aligned}\hat{M}_{seg} &= \sigma(M_{seg}) \\ \hat{C}_{cls} &= \sigma(C_{cls})\end{aligned}$$

Ці оцінки мають ті самі розмірності, що й оригінальні маска та класи, для зручності репрезентації операцій додано додаткові розмірності до вектору класів:  $\hat{M}_{seg} \in \mathcal{R}^{C \times H \times W}$  та  $\hat{C}_{cls} \in \mathcal{R}^{C \times 1 \times 1}$

Зважування карти сегментації відбувається за допомогою добутку Адамара між матрицями  $\hat{M}_{seg}$  та  $\hat{C}_{cls}$

$$M_{refined} = \hat{M}_{seg} \circ \hat{C}_{cls}$$

В умовах відсутності розмітки для прикладів в задачах сегментації, можливе ітеративне уточнення цієї розмітки. Для цього обчислюється уточнена карта ознак класифікації шляхом добутку Адамара між нормованим за допомогою сигмоїдної функції виходом декодера сегментації та логітами класифікації:

$$M_{unsup} = \hat{M}_{seg} \circ C_{cls} \quad (7)$$

Далі, для отримання результату класифікації виконується сумація елементів  $M_{unsup}$  з нормалізацією за сумою елементів оригінальної ненормалізованої карти локалізації:

$$C_{unsup} = \frac{\sum_{h=0}^H \sum_{w=0}^W M_{unsup}(h,w)}{\sum_{h=0}^H \sum_{w=0}^W M_{seg}(h,w) + c} \quad (8)$$

На основі запропонованих моделі нейронної мережі та методу її навчання, розроблено метод багатозадачного передбачення. Це дозволило підвищити достовірність класифікації та сегментації планарних зображень без зниження оперативності.

За допомогою моделі (1) протестовано запропоновані методи в різних умовах, виконаний аналіз внеску окремих компонентів та проведено аналіз стійкості запропонованого методу до різних рівнів помилок в анотаціях. В середньому, підвищення коефіцієнта Дайса відносно базової моделі склало 13%.

Сформульовано **четвертий пункт наукової новизни**: удосконалено методи багатозадачного навчання та передбачення результатів на основі удосконаленої моделі згорткових нейронних мереж шляхом об'єднання класифікації та сегментації і введення обмеження другого роду (зверху) при обчисленні функції втрат сегментації, що дозволило підвищити достовірність сегментації та класифікації в задачах автоматизованого скринінгу.

У **четвертому розділі** розроблені інструментальні засоби, що реалізують запропоновані рішення. Проведено випробування розробленого методу в рамках експериментів як на синтетичних даних, що були згенеровані за допомогою запропонованої моделі, а також експерименти в реальних задачах: скринінг діабетичної ретинопатії, скринінг меланоми, та скринінг хмарних утворень.

Інструментальні засоби розроблено мовою програмування Python з використанням фреймворку автоматичного диференціювання PyTorch. На основі розроблених інструментальних засобів створено ефективні програмні модулі, які інтегровано з “хмарними” сервісами для вирішення ресурсомістких задач навчання нейронних мереж, що забезпечує високу обчислювальну потужність та швидкість прогнозування в задачах автоматизованого скринінгу.

Для задачі автоматизованого скринінгу при сегментації патернів орга-

нізації хмар на супутникових знімках (в рамках проекту “Understanding Clouds from Satellite Images” на платформі для змагань з наук про дані Kaggle) було використано запропоновані моделі нейронних мереж, а також методи їх навчання та прогнозування результатів. Підвищення достовірності (міра Дайса) відносно базової моделі склало 3.9%.

Для задачі автоматизованого скринінгу при класифікації стадій діабетичної ретинопатії (в рамках проекту “APTOS 2019 Blindness Detection” на платформі для змагань з наук про дані Kaggle) було використано запропоновані методи багатозадачного навчання та прогнозування результатів. Підвищення достовірності (F1-міра) відносно базової моделі склало 2.1%.

Для задачі розпізнавання уражень шкіри при скринінгу меланоми (у рамках проекту SIIM-ISIC Melanoma Classification на платформі для змагань з наук про дані Kaggle) було використано запропонований локалізації важливих для класифікації ознак зображення. Використання запропонованого методу дозволило спростити процес контролю навчання нейронних мереж, що допомогло попередити перенавчання і підвищити достовірність класифікації на 3.5%.

Розроблені в роботі методи та інструментальні засоби отримали впровадження в навчальний процес ОНПУ та програмний продукт SafetyRadar компанії VITech Lab, основним призначенням якого є скринінг наявності елементів засобів індивідуального захисту на людях в умовах будівельних майданчиків, або лікарень та лабораторій.

*Ключові слова:* аналіз зображень, модель даних, глибинні нейронні мережі, багатозадачне машинне навчання, сегментація, класифікація, функції втрат.

## Список публікацій здобувача за темою дисертації.

1. Tymchenko B., Neural network methods for planar image analysis in automated screening systems // Applied aspects information technology. 2021. Том 4, № 1. С. 71-79. (Index Copernicus)  
<https://ait.opu.ua/?fetch=articles&with=info&id=73>
2. Tymchenko B., Khvedchenya E., Marchenko P., Spodarets D. Classification of skin lesions using multi-task deep neural networks // Herald of advanced information technology. 2020. Том 3, № 3. С. 136–148. (Index Copernicus)  
<https://hait.opu.ua/?fetch=articles&with=info&id=62>
3. Tymchenko B., Marchenko P., Spodarets D. Segmentation of cloud organization patterns from satellite images using deep neural networks // Herald of advanced information technology. 2020. Том 3, № 1. С. 352-361 (Index Copernicus)  
<https://hait.opu.ua/%EF%BF%BD%EF%BF%BD%EF%BF%BD%EF%>
4. Tymchenko B., Marchenko P., Spodarets D. Deep learning approach to diabetic retinopathy detection // Proceedings of the 9th International Conference on Pattern Recognition Applications and Methods, ICPRAM 2020, Valletta, Malta, February 22-24, 2020 / Ed. by Maria De Marsico, Gabriella Sanniti di Baja, Ana L. N. Fred. — SCITEPRESS, 2020. С. 501–509. (Scopus)  
<https://www.scitepress.org/Papers/2020/89708/>
5. Tymchenko B., Hramatik A., Tulchiy H. Classifying mixed patterns of proteins in microscopic images with deep neural networks // Herald of advanced information technology. 2019. Том 2, № 1. С. 29-36 (Index Copernicus)  
<https://hait.opu.ua/?fetch=articles&with=info&id=18>
6. Tymchenko B., Antoshchuk S. Race from pixels: Evolving neural network controller for vision-based car driving // XVIII International Conference

- on Data Science and Intelligent Analysis of Information / Springer. 2018. С. 20–29. (Index Copernicus)  
<https://www.springerprofessional.de/en/race-from-pixels-evolving-neural-network-controller-for-vision-b/16003024>
7. Tymchenko B, Hramatik A., Tulchyi H., Antoshchuk S. Making money: Evolving neural network for stock prediction // VI українсько-німецька конференція «Інформатика. Культура. Техніка». 2018. С. 34-35
  8. Tymchenko B., Antoshchuk S. Evolution strategy for policy search in robotics // Сучасні Інформаційні Технології / Одеський Національний Політехнічний Університет. 2018.  
<http://dspace.opu.ua/jspui/handle/123456789/8056>
  9. Tymchenko B., Halchonkov O. Online lane detection algorithm for line scan camera // Сучасні Інформаційні Технології / Одеський Національний Політехнічний Університет. 2017. С. 86-89  
<http://dspace.opu.ua/jspui/handle/123456789/3351>
  10. Tymchenko B. Global position system sensor model for robotics simulator // Праці Одеського політехнічного університету. 2017. № 3. С. 88–93.  
<http://dspace.opu.ua/jspui/handle/123456789/7939>
  11. Tymchenko B., Samodelok V., Putilina D., Galchonkov O. The robust control system for skid elimination in dynamic road environments // Електротехнічні та комп'ютерні системи. 2016. № 23. С. 107–112.  
<http://dspace.opu.ua/jspui/handle/123456789/1176>
  12. Komleva N., Cherneha K., Tymchenko B., Komlevoy O. Intellectual approach application for pulmonary diagnosis 2016 IEEE First International Conference on Data Stream Mining Processing (DSMP). 2016. С. 48–52.  
<https://ieeexplore.ieee.org/document/7583505/>
  13. Cherneha K., Tymchenko B., Komleva N. Decision support system for

automated medical diagnostics // Електротехнічні та комп'ютерні системи. 2016. № 23. С. 65–72. (Index Copernicus)

*<http://dspace.opu.ua/xmlui/handle/123456789/1140>*

## ABSTRACT

*Tymchenko B.I.* Neural Network Methods for Planar Image Analysis – Qualification scientific work in the form of the manuscript.

Thesis for the PhD degree in specialty 122 "Computer science" (12 – Information Technology). – Odessa National Polytechnic University, Ukraine, Odessa, 2021.

In **introduction**, the relevance of improving automated screening systems (screening is defined as detection and classification of atypical objects or processes) by creating models of datasets and methods for their generation, as well as models of neural networks and multi-tasking learning methods to increase the accuracy of classification and segmentation of planar images without increasing time. The object, subject, tasks and methods of research are defined; the connection with research works is shown; the scientific novelty and practical significance of the obtained results are given; the personal contribution of the applicant is covered.

In the **first section** of the dissertation the analysis of problems of analysis of planar images (which do not have depth measurement, and in which it is possible to discard scale of objects) in automated screening systems on the basis of neural network technologies is carried out.

The existing systems of automated screening in the subject areas of medicine, meteorology and remote sensing and features of their functioning by analyzing planar images are analyzed. It is shown that the effectiveness of such systems directly depends on the accuracy of image classification and segmentation, as false-positive results can lead to premature response, and false-negative to the omission of existing problems and late response. In addition, no less important factor influencing the effectiveness of screening systems is the time spent

on the implementation of the educational process, which, in turn, is associated with the quality of pre-created samples of training data. However, in the tasks of verifying the receipt of sufficient data and conducting high-quality annotations by professionals is usually expensive, and sometimes impossible due to the small number and variety of training examples, subjectivity in the formation of annotations in data sets. These factors negatively affect the accuracy of classification and segmentation in the training of deep neural networks by standard methods. To eliminate such influence, it is proposed to solve the problem taking into account partially erroneous annotations when constructing models of classification and segmentation.

It is shown that in such conditions it is expedient to use deep neural networks, due to the possibility of automating the learning process and the use of multi-tasking learning that increase the accuracy of classification and segmentation.

The analysis of the existing models of datasets and methods of their generation showed the limitations of these models, which would allow to form partially erroneous annotations, which have characteristic of real automated screening tasks.

Thus, solving an important scientific and practical problem of increasing the accuracy of classification and segmentation in the analysis of planar images for automated screening systems by improving models of neural networks and methods of their multitasking, especially in terms of partially erroneously annotated data samples.

In the **second section** method for generating datasets to enable testing of neural networks in learning using partially erroneous annotations and proposed a parametric formalization of the data set model with partially erroneous annotations, which are characteristic of real automated screening tasks were developed.

Because obtaining reliable classification and segmentation annotations for

test samples is often impossible or very costly, it is proposed to use model datasets to assess the reliability of models and methods of learning deep neural networks. Such datasets make it possible to conduct training and testing in controlled conditions by artificially adding errors to the annotations of the training data sample (test samples remains with precise annotations).

Proposed model of datasets  $\mathcal{M}$  has the following definition:

$$\mathcal{M} \in \{\mathcal{X}_b, \mathcal{X}_{tex}, N, S_{img}, S_{obj}, \Delta_{max}, N_{obj}, P_e, P_d, S_e, S_d\} \quad (9)$$

where  $\mathcal{X}_b$  - bacground images set,  $\mathcal{X}_f$  - objects set,  $\mathcal{X}_{tex}$  - object texturing set,  $N$  - number of images in the generated dataset,  $S_{img}$  - pixel size of generated images,  $S_{obj}$  - average size of the objects (pixels),  $\Delta_{max}$  - maximum deviation of objects size,  $N_{obj}$  - maximum number of objects in the image,  $P_e$  and  $P_d$  - probabilities of mask erosion and dilation for every object,  $S_e$  and  $S_d$  - kernel sizesfor erosion and dilation.

The last four parameters are introduced for controlled annotation errors.

Formulated **first point of scientific novelty**: for the first time a parametric formalization of the dataset model with partially erroneous annotations, which are characteristic of real problems of automated screening, was proposed, which allowed the development of a method of generating training, test and validation data sets.

Based on the proposed parametric model (9), a method for generating data sets with partly-erroneous annotations has been developed, which contains the following steps:

1. Choose random background:  $x_{bg} \sim \mathcal{X}_b$
2. Choose number of objects in the image:  $N_{obj} \sim \mathcal{U}(1, N_{obj})$
3. Initialize the segmentation mask:  $M = 0$  sos  $M \in \mathcal{R}^{C \times S_{img} \times S_{img}}$
4. For  $n_{obj}$  times:
  - 4.1 Sample object size:  $s \sim \mathcal{U}(S_{obj} - \Delta_{max}, S_{obj} + \Delta_{max})$

4.2 Sample object coordinates:

$$i_f \sim \mathcal{U}(0, S_{img} - s)$$

$$j_f \sim \mathcal{U}(0, S_{img} - s)$$

4.3 Sample object image  $x_{fg} \sim \mathcal{X}_f$  and the corresponding object class  $c_{fg} \sim \mathcal{Y}_f$

4.4 Resample the image with bilinear interpolation:

$$\hat{x}_{fg} = R_{bilinear}(x_{fg})$$

4.5 Sample the texture image  $x_{tex} \sim \mathcal{X}_{tex}$

4.6 Modify object image with the object texture:

$$\hat{x}_{fg} = x_{fg} \circ x_{tex}[i_f : i_f + s, j_f : j_f + s]$$

4.7 Place object image into the background image:

$$x_{bg}[i_f : i_f + s, j_f : j_f + s] = (1 - x_{fg}) \circ x_{bg} + \hat{x}_{fg}$$

4.8 Form the segmentation mask:

$$M_{seg} = x_{fg} > \theta_{seg}$$

where  $\theta_{seg}$  - is the source object binarization threshold. For the MNIST dataset  $\theta_{seg} = 0.2$ , and for FashionMNIST  $\theta_{seg} = 0.1$ .

4.9 Modify the segmentation mask according to the desired error level:

$$M_{seg} = \begin{cases} M_{seg} \oplus K^{S_d \times S_d} & \text{iff } p_d \sim \mathcal{U}(0, 1) < P_d \\ M_{seg} \ominus K^{S_e \times S_e} & \text{iff } p_e \sim \mathcal{U}(0, 1) < P_e \end{cases}$$

where  $K^{S_d \times S_d}$  - dilation kernel matrix and  $K^{S_e \times S_e}$  - erosion kernel matrix.

4.10 Place modified object mask into the common mask:

$$M[c_{fg}, i_f : i_f + s, j_f : j_f + s] = \max \{M[c_{fg}, i_f : i_f + s, j_f : j_f + s], M_{seg}\}$$

#### 4.11 Save image $x_{bg}$ and mask $M$

### 5. Finish generation

Thus, the method of generating datasets with partially erroneous annotations based on the parametric model was improved, which due to the generation of annotations: partially erroneous for the training sample and accurate - for the test, made it possible to test the impact of annotation errors on neural networks.

The method allows to obtain many data sets with similar characteristics and to use non-parametric statistical methods (such as bootstrapping) to evaluate models in the absence of real training data. The method is **the second point of scientific novelty**.

In the **third section** of the dissertation neural network models and methods of multi-tasking learning were developed to simultaneous increase in accuracy of classification and segmentation without decrease in efficiency.

To implement multi-task learning methods, it is proposed to improve the models of deep neural networks using the encoder-decoder architectures (UNet, LinkNet) by introducing an additional decoder with the normalization layer. Thus, the advanced model consists of an encoder and two decoders (for segmentation and classification tasks, respectively).

The model of the deep neural network is represented by the following expression:

$$v_1, v_2 \dots v_n = F_{encoder}(x, \theta_{enc}) \quad (10)$$

$$M_{seg} = F_{seg}((v_1, v_2 \dots v_n), \theta_{seg}) \quad (11)$$

$$C_{cls} = F_{cls}((v_n), \theta_{cls}) \quad (12)$$

where  $\theta_{enc}$  - are encoder parameters,  $\theta_{seg}$  and  $\theta_{cls}$  - parameters of encoder and decoder, respectively,  $F_{seg}$  and  $F_{cls}$  - decoder neural networks.

The proposed model is **the third point of scientific novelty**.

With the introduction of an additional classification decoder with a normalization layer, it is possible to implement methods of multi-tasking learning and inference for deep neural networks.

During the practical use of neural network methods in screening tasks, it is established that partially erroneous annotations of training data prevent obtaining high accuracy of classification and segmentation.

The method of multi-task learning in the conditions of partially-erroneous annotations of training data, which is based on use of tasks that are connected with original task is offered. It is shown that for the segmentation problem there is a similar classification problem, for which the annotations of training data are of better quality than for the original segmentation problem.

The proposed method of multi-task learning consists of two stages.

**Stage 1. Generating the derivative task.** In the context of multi-task learning, the task of classification is defined as a multiple instance learning task, instead of marking each of the objects for all classes. In this case, the annotation of the original planar image is a set with one or more objects.

The loss function is calculated separately for each of the tasks (classification and segmentation). The top-limited loss function  $L_{seg}]$  is used to train the segmentation decoder, while the classification decoder uses the normal  $L_{cls}$  loss function, and the total value of the loss function is defined as the arithmetic mean between the individual values:

$$L_{total} = \frac{L_{seg}] + L_{cls}}{2} \quad (13)$$

Accordingly, the total gradient of the loss function will also be the arithmetic mean of the gradients of the components  $\nabla L_{seg}]$  and  $\nabla L_{cls}$ :

$$\nabla L_{total} = \frac{\nabla L_{seg}] + \nabla L_{cls}}{2} \quad (14)$$

Thus, the presence of non-zero gradients is provided to update the parameters from at least one loss function for each input example.

**Stage 2. Loss function bounding.** To reduce the impact of the erroneous part of the annotations, it is proposed to introduce a constraint to the loss function for the segmentation problem with lower quality data annotations. That allowed to reduce the influence of gradients of the loss function on examples with erroneous annotations when learning on several tasks:

$$\mathcal{L}] = \min(L, \theta)$$

where  $\theta$  - loss function threshold.

Thus, the introduced boundaries are represented by the following expression:

$$\nabla \min(L, \theta) = \begin{cases} 1, & \text{iff } L \in (-\infty, \theta] \\ 0, & \text{iff } L \in (\theta, \infty) \end{cases}$$

In the absence of markup for examples in segmentation problems, iterative refinement of this markup is possible. To do this, a refined map of classification features is calculated by the Hadamard product between the output of the segmentation decoder normalized by the sigmoid function and the classification logs:

$$M_{unsup} = \hat{M}_{seg} \circ C_{cls} \quad (15)$$

Next, to obtain the classification result, the summation of the elements  $M_{unsup}$  is performed with normalization by the sum of the elements of the original non-normalized localization map:

$$C_{unsup} = \frac{\sum_{h=0}^H \sum_{w=0}^W M_{unsup}(h,w)}{\sum_{h=0}^H \sum_{w=0}^W M_{seg}(h,w) + c} \quad (16)$$

Based on the proposed model of the neural network and the method of its training, a method of multi-taski prediction has been developed. This

allowed to increase the accuracy of classification and segmentation of planar images without increasing time needed.

Method consists of the following stages:

**Stage 1. Normalization.** Let  $C_{cls} \in \mathcal{R}^C$  and  $M_{seg} \in \mathcal{R}^{C \times H \times W}$  - logits of classification and segmentation decoders, respectively and lie in the interval  $(-\infty, +\infty)$ .

To obtain normalized to the interval  $[0, 1]$ , sigmoid activation function is used:

$$\sigma(x) = \frac{1}{1 + e^{-x}}$$

**Stage 2. Renormalization.** Renormalization is the weighing of a segmentation map using normalized classifier logits. The first step is the transformation of segmentation and classification logs into uncalibrated estimates between  $[0, 1]$ :

$$\begin{aligned}\hat{M}_{seg} &= \sigma(M_{seg}) \\ \hat{C}_{cls} &= \sigma(C_{cls})\end{aligned}$$

These estimates have the same dimensions as original segmentation mask and classification vector, additional dimensions are added for simplicity:  $\hat{M}_{seg} \in \mathcal{R}^{C \times H \times W}$  and  $\hat{C}_{cls} \in \mathcal{R}^{C \times 1 \times 1}$

Segmentation map is weighted using the Hadamard product:  $\hat{M}_{seg}$  and  $\hat{C}_{cls}$

$$M_{refined} = \hat{M}_{seg} \circ \hat{C}_{cls}$$

Using the model, the proposed methods were tested in different conditions, ablation studies were performed and the analysis of the resistance of the proposed method to different levels of errors in the annotations was performed. On average, the increase in the Dice coefficient relative to the base method was 13 %.

Thus, **fourth point of scientific novelty** is formulated, namely, improved methods of multi-task learning and inference based on an improved model of convolutional neural networks by combining classification and segmentation and introducing a constraint when calculating the segmentation loss function, which allowed to increase the accuracy of segmentation and classification in automated screening tasks.

In **fourth section** developed tools that implement the proposed solutions. The developed method was tested as part of experiments on synthetic data generated using the proposed model, as well as in experiments in real problems: screening of diabetic retinopathy, melanoma screening, and screening of cloud formations.

The tools are developed in the Python programming language using the PyTorch automatic differentiation framework. Based on the developed tools, effective software modules have been created, which are integrated with cloud services for solving resource-intensive learning tasks of neural networks, which provides high computing power and speed of prediction in automated screening tasks.

For the task of automated screening in the cloud organization patterns segmentation in satellite images (within the project "Understanding Clouds from Satellite Images" on the platform for competitions in data science Kaggle) the proposed models of neural networks were used, as well as methods for training and inference. The increase in Dice measure relative to the base model was 3.9%.

For the task of automated screening in the diabetic retinopathy classification (within the project "APTOS 2019 Blindness Detection" on the platform for competitions in data science Kaggle) used the proposed methods of multi-tasking and inference. The increase in F1-measure relative to the base model was 2.1 %.

For the task of recognizing skin lesions in melanoma screening (within

the SIIM-ISIC Melanoma Classification project on the Kaggle Data Science Competition Platform), the proposed localization of image-relevant image features was used. The use of the proposed method simplified the process of monitoring the learning of neural networks, which helped prevent retraining and increase the reliability of the classification by 3.5%.

The methods and tools developed in the work were introduced into the educational process of ONPU and the SafetyRadar software product of VITech Lab, the main purpose of which is to screen the presence of elements of personal protective equipment on people in construction sites, or hospitals and laboratories.

*Key words:* image analysis, dataset models, deep neural networks, multi-tasking machine learning, segmentation, classification, loss functions.