

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. — ІКС ОНПУ, 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 науковими програмами та планами; 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 функціонування систем автоматизованого скринінгу на основі 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 для вирішення задач аналізу планарних зображень в скринінгових системах, отримання великої кількості попередньо маркованих (розмічених) якісних даних навчальних вибірок є дуже дорогим, та іноді не є можливим в принципі через малу кількість прикладів, різноманіття зовнішнього вигляду та неузгодженість анотаторів (через складність процесу розмітки, непрофесіоналізм, тощо) під час формування анотацій в наборах даних. Ці фактори негативно відображаються на достовірності класифікації та сегментації при навчанні глибоких нейронних мереж стандартними методами. Для її розв'язання 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. Також, використання глибоких нейронних мереж дозволяє організацію багатозадачного глибокого навчання, що підвищує достовірність класифікації.

Проведений аналіз наявних моделей наборів даних та методів їх генерації показав обмеженість моделей наборів даних, які б дозволяли змінювати рівень похибки в анотаціях відповідно до характеристик помилок в реальних наборах даних.

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 за-

безпечення 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, Δ_{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.

Введення цих параметрів дозволяє ввести контрольоване створення помилок в анотаціях відповідно до реальних характеристик.

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 зашумленими 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 θ_{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 що за рахунок генерації анотацій: частково-помилкових для тренувальної вибірки та достовірних - для тестової, дало можливість виконувати тестування впливу рівня помилок анотацій на роботу нейромережевих 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 моделі нейронних мереж та методи багатозадачного навчання для одночасного підвищення достовірності класифікації та сегментації без зниження оперативності.

Для реалізації методів багатозадачного навчання, запропоновано удосконалити моделі глибоких нейронної мережі з використанням архітектури енкодер-декодер (UNet, LinkNet) введенням додаткового декодера з додатковим 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 θ_{enc} - набір параметрів енкодера, θ_{seg} та θ_{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'явилася можливість реалізації методів багатозадачного навчання глибоких нейронних мереж та передбачення результатів.

При практичному використанні нейромережових методів в задачах скринінгу встановлено, що частково-помилкові анотації тренувальних даних заважають отриманню високої достовірності в задачах класифікації та сегментації. Запропоновано метод багатозадачного навчання нейронних мереж в умовах частково-помилкових анотацій навчальних даних спирається на використання задач, похідних від оригінальної. Показано, що для задачі сегментації існує близька більш загальна задача класифікації, для якої анотації навчальних даних є більш влучними, ніж для вихідної задачі сегментації.

В такому контексті, задача класифікації зводиться до задачі навчання

за набором зразків: замість маркування кожного з об'єктів для всіх класів на зображенні, зображення являє собою мішок з одним, чи декількома об'єктами та відповідним маркуванням, чи є об'єкти заданих класів на зображенні.

На відміну від попередніх методів, що спираються на вивчення більш детальних близьких задач, в запропонованому методі використання більш точних даних, для загальніших задач дозволяє покращити роздільність внутрішніх представлень нейронної мережі, що, в свою чергу, покращує результати на вихідній задачі. Також, оскільки задачі є близькими, не відбувається конфлікту градієнтів, що є типовим при навчанні різнорідних задач.

Для кожної з задач окремо обчислюється функція втрат. Для навчання декодера сегментації використовується обмежена зверху функція втрат, в той час як для декодера класифікації - звичайна. Загальне значення функції втрат визначається як арифметичне середнє між індивідуальними значеннями:

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

Відповідно, загальний градієнт функції втрат буде сумою градієнтів складових частин:

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

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

Для зменшення впливу помилкової частини анотацій, запропоновано зміну до методу оцінки при навчанні нейронних мереж: введено обмеження другого роду (зверху) на функцію втрат для задачі з менш точною маркуванням, таким чином, при навчанні на декількох задачах, градієнти

в прикладах з неправильними анотаціями не впливають на процес навчання:

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

де θ - поріг обмеження функції втрат.

Для функції обмеження зверху градієнт визначено лише на проміжку $(-\infty, \theta]$, тому для проміжку (θ, ∞) градієнт встановлено рівним нулю:

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

На основі запропонованих моделі нейронної мережі та методу її навчання, розроблено метод об'єднання близьких задач на етапі прогнозування, для того, щоб підвищити достовірність класифікації та сегментації планарних зображень без підвищення витрат часу.

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:

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

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

$$\hat{M}_{seg} = \sigma(M_{seg})$$

$$\hat{C}_{cls} = \sigma(C_{cls})$$

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}

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

Для покращення можливостей інтерпретації прогнозів моделей, було вдосконалено метод локалізації важливих для класифікації ознак зображення за рахунок використання методів багатозадачного навчання, що дозволило його використання в умовах відсутності анотованих даних для сегментації в навчальному наборі даних.

В основі запропонованого методу лежить ітеративне уточнення карти ознак локалізації за допомогою направлення градієнтів від задачі класифікації:

Обчислюється уточнена 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:

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

За допомогою моделі 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%.

Таким чином, можна сформулювати **четвертий пункт наукової новизни**, а саме 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 організації 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.