

ОДЕСЬКИЙ НАЦІОНАЛЬНИЙ ПОЛІТЕХНІЧНИЙ УНІВЕРСИТЕТ
ІНСТИТУТ КОМП'ЮТЕРНИХ СИСТЕМ

Кваліфікаційна наукова
праця на правах рукопису

ТИМЧЕНКО Борис Ігорович

УДК 004.932.72'1

ДИСЕРТАЦІЯ
НЕЙРОМЕРЕЖЕВІ МЕТОДИ АНАЛІЗУ ПЛАНАРНИХ
ЗОБРАЖЕНЬ В СИСТЕМАХ АВТОМАТИЗОВАНОГО
СКРИНІНГУ

122 — Комп'ютерні науки
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.

_____ Б. І. Тимченко

Науковий керівник: **Антощук Світлана Григорівна,**
доктор технічних наук, професор

Одеса — 2021

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.

Актуальність теми зумовлюється стрімким розвитком різних областей технологій і процесів, які потребують швидкого реагування на нетипові зміни для подальшого коригування, поки ціна реакції є низькою. Задача виявлення таких нетипових змін називається задачею скринінгу. Для підвищення продуктивності, створюються інтелектуальні системи, що дозволяють вирішувати задачу скринінгу автоматично без участі людини. В таких системах постає задача автоматизованого скринінгу.

Актуальним методом скринінгу є скринінг заснований на аналізі зображень, що дозволяє спростити забір проб, а також використати інструментальні засоби віддаленого спостереження, такі як рентгенографія, або супутникове зондування. В задачах скринінгу, що заснований на зображеннях можуть використовуватися як планарні (рентгенограми, двомірні дефектоскопи, та ін.), так і об'ємні зображення (томографії, тривимірні ультразвукові дослідження та ін.). Особливістю планарних зображень є відсутність перспективних спотворень та можливість знехтувати глибиною об'єктів.

Важливими складовими інтелектуальних систем автоматизованого скринінгу є класифікація зображень та локалізація регіонів інтересу.

Ефективність роботи систем автоматизованого скринінгу напряму залежить від достовірності класифікації та локалізації, оскільки хибно-позитивні

результати можуть призвести до передчасного реагування, а хибно-негативні результати - до пропуску наявних проблем та пізнього реагування. Традиційно, в таких системах використовуються класичні методи комп'ютерного зору, однак достовірність їхніх прогнозів суттєво обмежена через різноманіття можливих проблем та складність їх формалізації за допомогою правил.

Все частіше для аналізу зображень впроваджуються методи на основі машинного навчання, зокрема глибинного, через значне підвищення достовірності прогнозів. Однак, в наявні підходи на основі глибоких нейронних мереж можуть бути застосовані виключно за умови наявності великих тренувальних наборів розмічених даних. Часто, в задачах скринінгу, отримання великої кількості даних, а також якісної розмітки від професіоналів є дуже дорогим, та іноді не є можливим в принципі через малу кількість прикладів, їх різноманіття та неузгодженість розмітників під час формування наборів даних. Достовірність прогнозів визначається як міра Дайса-Соренсена, або F1-міра на валідаційному наборі з точною розміткою.

Таким чином, має місце **протиріччя** між підвищенням вимог щодо достовірності прогнозів нейромережових методів з одного боку, та необхідністю зниження витрат на отримання розмітки навчальних даних.

В дисертаційній роботі пропонується новий підхід до розв'язання цього протиріччя шляхом розробки моделей нейронних мереж та методів їх навчання для аналізу планарних зображень, що здатні використовувати неточно розмічені тренувальні набори даних.

Запропонований метод використовує багатозадачне навчання для підвищення достовірності класифікації та семантичної сегментації за рахунок навчання нейронних мереж на більш загальних та семантично-близьких задачах одночасно з основною задачею з подальшою комбінацією результатів виконання цих задач.

Тобто, підвищення достовірності прогнозів нейромережових методів в

задачах автоматизованого скринінгу через розробку та удосконалення методів багатозадачного навчання є **актуальним**.

Зв'язок роботи з науковими програмами, планами, темами.

Метою дослідження є підвищення достовірності класифікації та семантичної сегментації планарних зображень в системах автоматизованого скринінгу через розробку моделі та методів аналізу планарних зображень на основі штучних нейронних мереж.

Для досягнення мети дослідження поставлено і розв'язано такі **задачі**:

- Здійснено аналіз зашумлення розмітки в наборах даних для задач автоматизованого скринінгу, проведено аналіз існуючих моделей, методів, алгоритмів та засобів аналізу планарних зображень в системах автоматизованого скринінгу і обґрунтовано вибір напрямку досліджень.
- На основі проведеного аналізу розроблено параметричну модель набору даних, що складається з планарних зображень із зашумленою розміткою для задач класифікації та семантичної сегментації, а також метод його генерації.
- Розроблено модель нейронної мережі та методи навчання й прогнозування для аналізу планарних зображень, що за рахунок використання декількох семантично-близьких задачах одночасно дозволяють підвищити точність класифікації та семантичної сегментації в умовах зашумлення розмітки даних.
- Розроблено інструментальні засоби, що виконують розроблені моделі та методи, і провести їх випробування в рамках експерименту

Об'єктом дослідження є процес аналізу планарних зображень.

Предмет дослідження – методи класифікації та семантичної сегментації планарних зображень, структури та моделі згорткових нейронних мереж та методи їх навчання.

Методи дослідження:

Наукова новизна отриманих результатів.

- *Удосконалено* параметричну модель набору даних та метод його генерації, що за допомогою додавання контрольованого зашумлення розмітки, яке відповідає характеристикам зашумлення в реальних задачах автоматизованого скринінгу дає можливість виконувати тестування нейромережевих методів семантичної сегментації та класифікації через використання синтетичних наборів даних.
- *Удосконалено* модель згорткової нейронної мережі та метод багато-задачного навчання штучних нейронних мереж для роботи в умовах частково помилкової розмітки навчального набору даних в задачах автоматизованого скринінгу що за рахунок навчання на декількох задачах дозволяє підвищити достовірність семантичної сегментації на 8%.
- *Вперше* розроблено метод об'єднання семантично-близьких задач на етапі прогнозування, що за допомогою використання результатів декількох задач дозволило підвищити достовірність класифікації та сегментації планарних зображень на 12% без підвищення витрат часу.
- *Отримав подальший розвиток* метод локалізації важливих для класифікації ознак зображення в умовах відсутності розмітки для сегментації в навчальному наборі даних, що за рахунок

Практичне значення одержаних результатів полягає у створенні інструментальних засобів генерації наборів даних відповідно до розробленої параметричної моделі, а також навчання та прогнозування нейронних мереж для задач автоматизованого скринінгу.

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 автома-

тизованого скринінгу.

Розроблені в роботі методи отримали **впровадження** в програмний продукт SafetyRadar компанії *VITech Lab*.

Ключові слова: аналіз зображень, глибокі нейронні мережі, багатозадачне машинне навчання, семантична сегментація, функції втрат.

ЗМІСТ

Розділ 1. Інструментальні засоби та експериментальні дослідження	9
1.1. Інструментальні засоби	9
1.1.1. Узагальнена структура системи автоматизованого скринінгу	9
1.1.2. Узагальнена структура системи навчання глибоких нейронних мереж	9
1.1.3. Узагальнена структура прогнозування за допомогою глибоких нейронних мереж	9
1.2. Експерименти на синтетичних даних	9
1.2.1. Метод оцінки прогнозів нейронної мережі	9
1.2.2. Структура нейронних мереж	10
1.2.3. Параметри експериментів	11
1.3. Експерименти на реальних даних	27
1.3.1. Класифікація стадії осередків діабетичної ретинопатії	27
1.3.2. Детальний опис процедури навчання	28
1.4. Класифікація раку шкіри та сегментація родимок	30
1.4.1. Опис набору даних SIIM-ISIC Melanoma Classification	30
1.4.2. Детальний опис процедури навчання	32
1.4.3. Результати експерименту	32
1.5. Класифікація та сегментація формацій хмар	32
1.5.1. Опис набору даних Understanding Clouds from Satellite Images	32
1.5.2. Детальний опис структури нейронної мережі	33

1.5.3. Результати експерименту	33
--	----

РОЗДІЛ 1

ІНСТРУМЕНТАЛЬНІ ЗАСОБИ ТА ЕКСПЕРИМЕНТАЛЬНІ ДОСЛІДЖЕННЯ

В даному розділі розглянуто експерименти на синтетичних даних, що були згенеровані за допомогою моделі, а також експерименти в реальних задачах скринінгу.

1.1. Інструментальні засоби

1.1.1. Узагальнена структура системи автоматизованого скринінгу.

1.1.2. Узагальнена структура системи навчання глибоких нейронних мереж.

1.1.3. Узагальнена структура прогнозування за допомогою глибоких нейронних мереж.

1.2. Експерименти на синтетичних даних

1.2.1. Метод оцінки прогнозів нейронної мережі.

Основні метрики для оцінки задач. В наведених нижче експериментах використовуються наступні метрики:

- Метрика F1 для задачі класифікації
- Коефіцієнт Дайса для задачі семантичної сегментації
- Коефіцієнт роздільності внутрішніх представлень K_{sep}

В експериментах поставлена задача багатокласової класифікації та семантичної сегментації, для отримання фінального значення метрик, вони усереднені по класах.

Метод кількісної оцінки роздільності внутрішніх представлень нейронної мережі. Для оцінки роздільності внутрішніх представлень, використовується метод К-середніх. Оскільки в задачах класифікації та семантичної сегментації відома кількість класів, можна заздалегідь визначити необхідну кількість кластерів ознак у внутрішньому представленні нейронної мережі.

Для набору даних над ознаками з внутрішнього представлення обчислюється кількість кластерів, що відповідає кількості класів в основній задачі. Після чого, обчислюється середня Евклідова відстань між центрами кластерів. Таким чином, більшій відстані між кластерами відповідає краща роздільність внутрішніх представлень.

Також, для наочної демонстрації складу внутрішніх представлень, ознаки візуалізовано за допомогою алгоритму нелінійного відображення Uniform Manifold Approximation and Projection (UMAP) [?] на двомірну площину.

1.2.2. Структура нейронних мереж. У всіх експериментах використана однакова структура нейронних мереж якщо не зазначено інакше.

Структура базової моделі. Базова модель являє собою екземпляр архітектури UNet. В ролі енкодера використана архітектура нейронної мережі ResNet34, що ініціалізована вагами претренованої на наборі даних Imagenet нейронної мережі.

Декодер виконано відповідно до класичної архітектури UNet, що складається з п'яти стадій (рисунок ??). Кількість каналів в згортках стадій від найглибшої: 256, 128, 64, 32, 16. Функція активації декодера - ReLU. На кожній стадії використовується пакетна нормалізація ознак. Додаткові

з'єднання від енкодера передаються за допомогою операції конкатенації. Функція активації останнього шару декодера - логістична сигмоїда.

Структура запропонованої моделі. Енкодер та декодер сегментації запропонованої моделі ідентичні до базової моделі. Додано декодер класифікації, архітектура якого зображена на рисунку. Входом декодера класифікації є вихід останнього шару енкодера.

При навчанні запропонованої моделі використовується обмежена функція втрат для задачі сегментації, а при прогнозуванні використовується метод об'єднання задач класифікації та сегментації.

1.2.3. Параметри експериментів.

Загальна процедура навчання. Складається з одного циклу навчання та валідації.

Використано наступні параметри навчання:

- Розмір зображення: 224×224 пікселя
- Розмір пакету навчання: 32
- Кількість тренувальних зображень: 15000
- Кількість тестових зображень: 5000
- Кількість епох навчання: 20
- Оптимізатор: RAdam [?]
- Темп навчання: 10^{-3}
- Коефіцієнт L2 регуляризації параметрів мережі: 10^{-4}
- Закон зміни темпу навчання: косинус

Експеримент 1. Порівняння базового та запропонованого методів на простому наборі даних без зашумлення розмітки. Параметри генерації набору даних зазначено в таблиці 1.1.

Таблиця 1.1: Параметри набору даних експеримента 1

Об'єкти	MNIST
Текстура фону	константна
Текстура об'єктів	константна
P_e	0
P_d	0
N_{obj}	10
S_{obj}	36
Δ_{max}	0.2

Результати експерименту наведено в таблиці 1.2.

Таблиця 1.2: Результати експеримента 1

Модель	Міра Дайса	F1-міра	K_{sep}
Базова модель	0.86	-	14.33
Запропонований метод	0.87	0.98	17.51

На рисунку 1.1 зображено приклади прогнозів нейронної мережі для тестових зображень.

На рисунках 1.2 та 1.3 зображено UMAP-проекцію ознак останнього шару енкодера для набору тестових зображень.

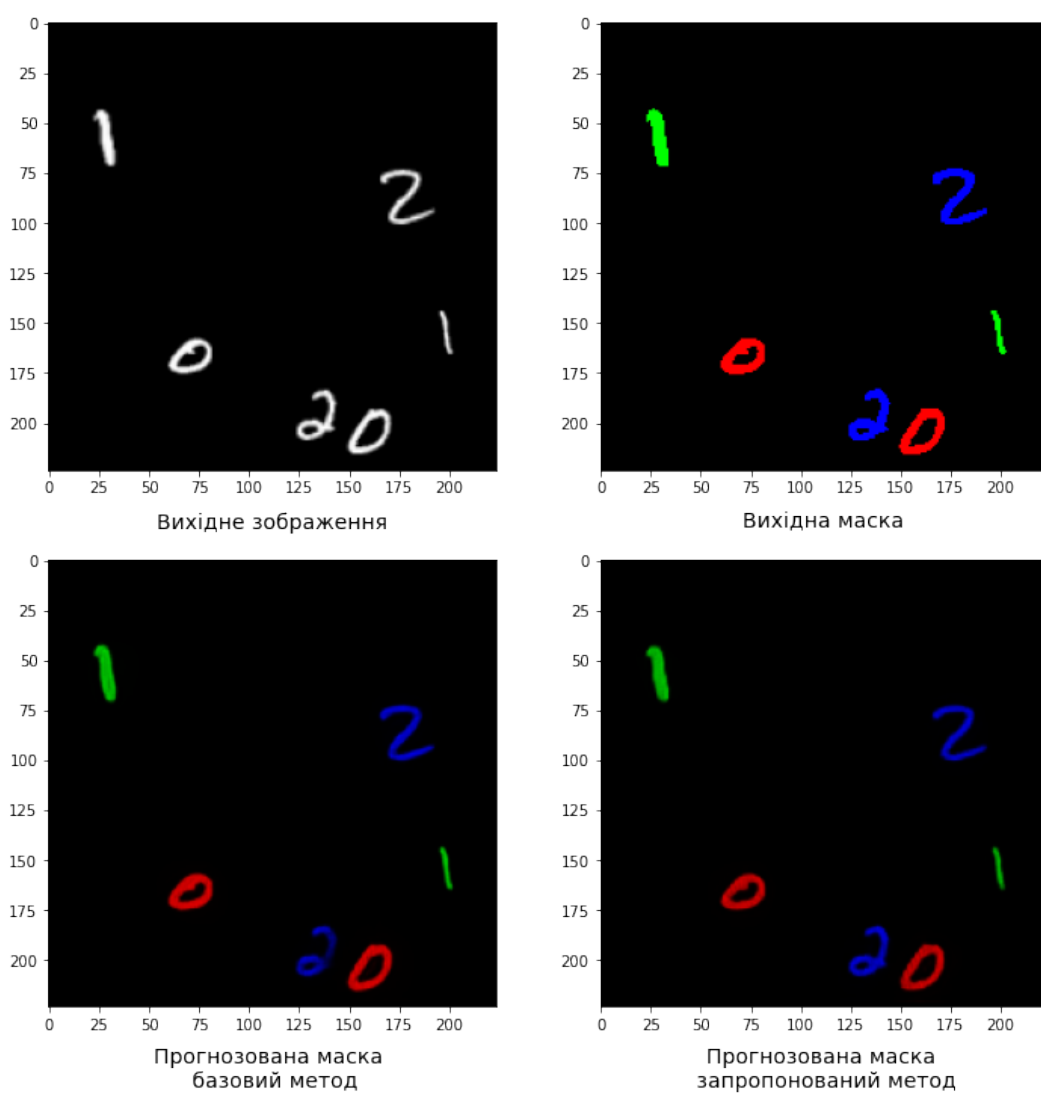


Рис. 1.1: Структура декодера класифікації

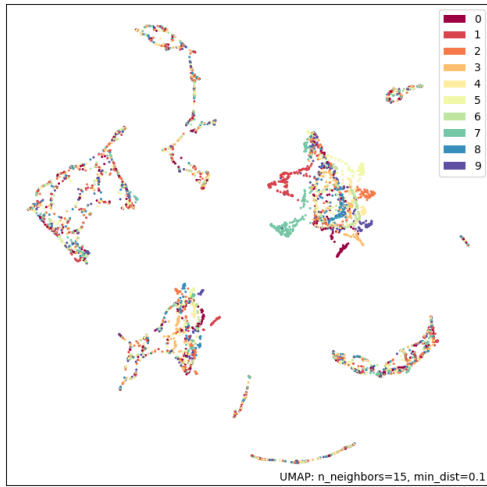


Рис. 1.2: UMAP ознак
базової моделі

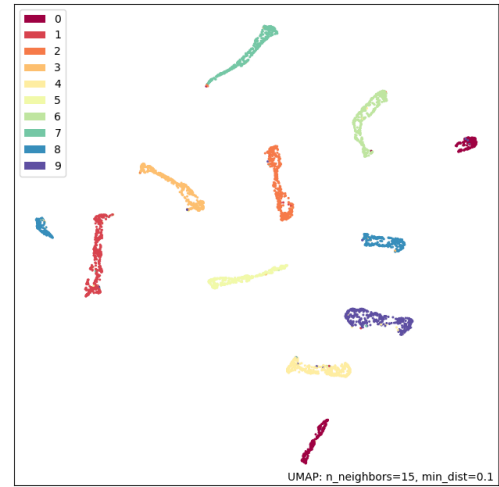


Рис. 1.3: UMAP ознак
запропонованої моделі

Експеримент 2. Порівняння базового та запропонованого методів на простому наборі даних із зашумленням розмітки. Параметри генерації набору даних зазначено в таблиці 1.3.

Таблиця 1.3: Параметри набору даних експеримента

Об'єкти	MNIST
Текстура фону	константна
Текстура об'єктів	константна
P_e	0.1
P_d	0.25
N_{obj}	10
S_{obj}	36
Δ_{max}	0.2

Результати експерименту наведено в таблиці 1.4.

Таблиця 1.4: Результати експеримента

Модель	Міра Дайса	F1-міра	K_{sep}
Базова модель	0.81	-	12.10
Запропонований метод	0.91	0.93	15.73

На рисунку 1.4 зображено приклади прогнозів нейронної мережі для тестових зображень.

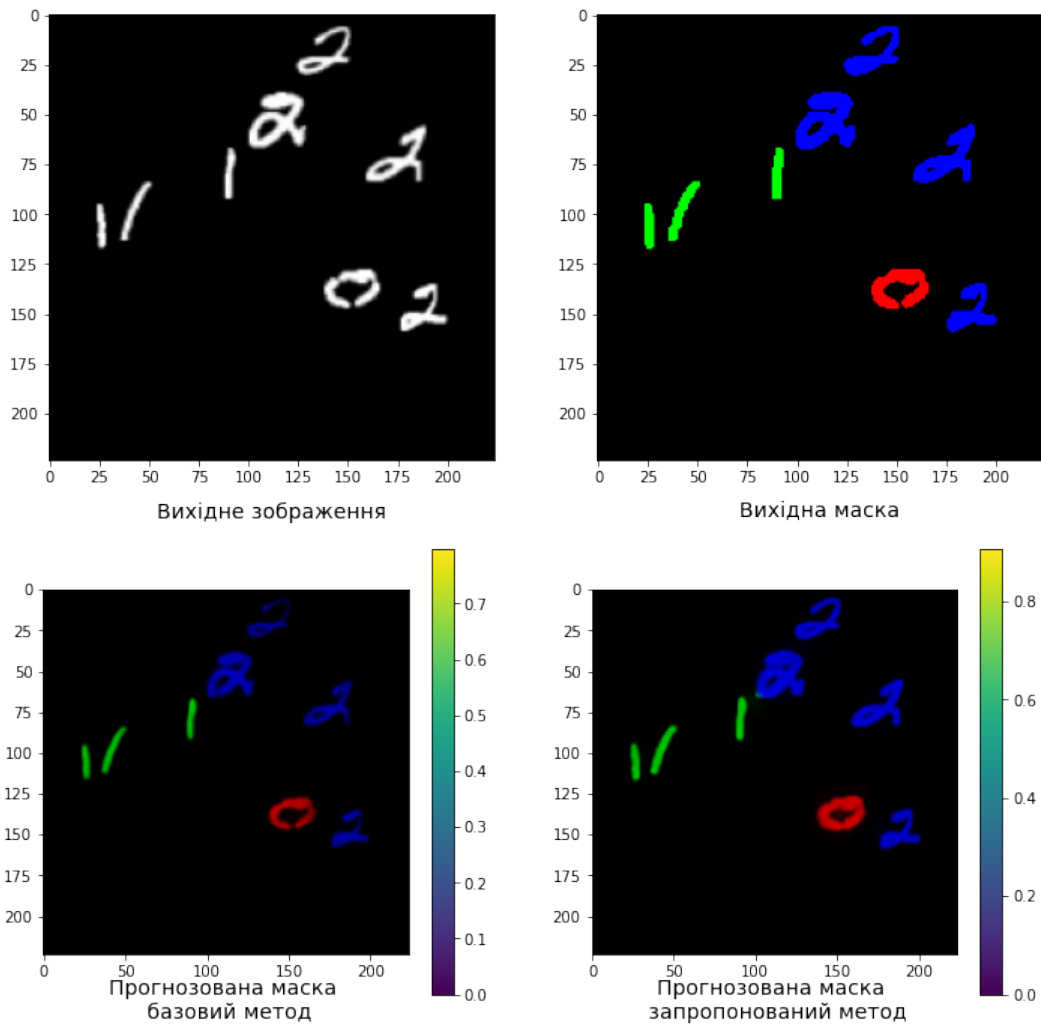


Рис. 1.4

На рисунках 1.5 та 1.6 зображено UMAP-проекцію ознак останнього шару енкодера для набору тестових зображень.

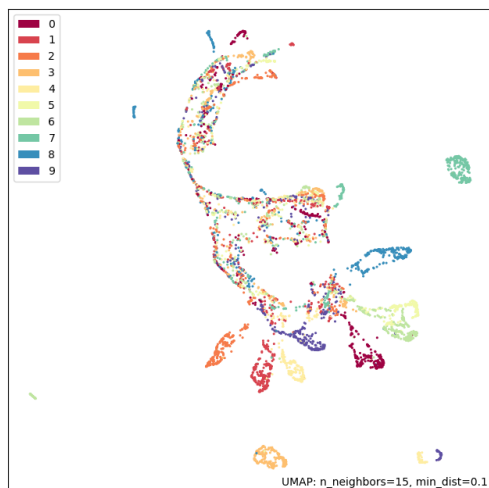


Рис. 1.5: UMAP ознак
базової моделі

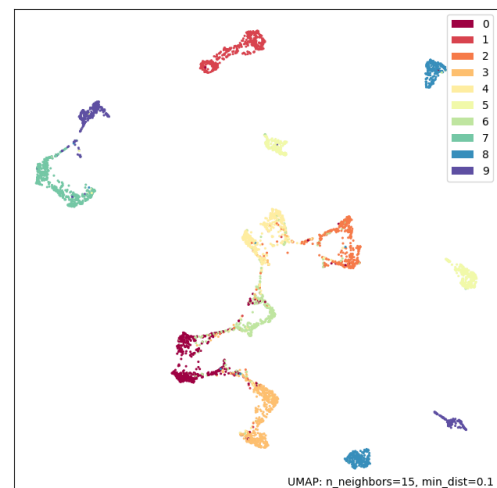


Рис. 1.6: UMAP ознак
запропонованої моделі

В умовах зашумленої розмітки, запропонований метод дає більш впевнені прогнози, також зберігаючи тонкі лінії, що порушуються при виконанні операції ерозії над розміткою.

Експеримент 3. Порівняння базового та запропонованого методів на простому наборі даних з ускладненими об'єктами та зашумленням розмітки. Параметри генерації набору даних зазначено в таблиці 1.5.

Таблиця 1.5: Параметри набору даних експеримента

Об'єкти	FashionMNIST
Текстура фону	константна
Текстура об'єктів	константна
P_e	0.1
P_d	0.25
N_{obj}	10
S_{obj}	36
Δ_{max}	0.2

Результати експерименту наведено в таблиці 1.6.

Таблиця 1.6: Результати експеримента

Модель	Міра Дайса	F1-міра	K_{sep}
Базова модель	0.83	-	27.49
Запропонований метод	0.87	0.88	32.12

На рисунку 1.7 зображено приклади прогнозів нейронної мережі для тестових зображень.

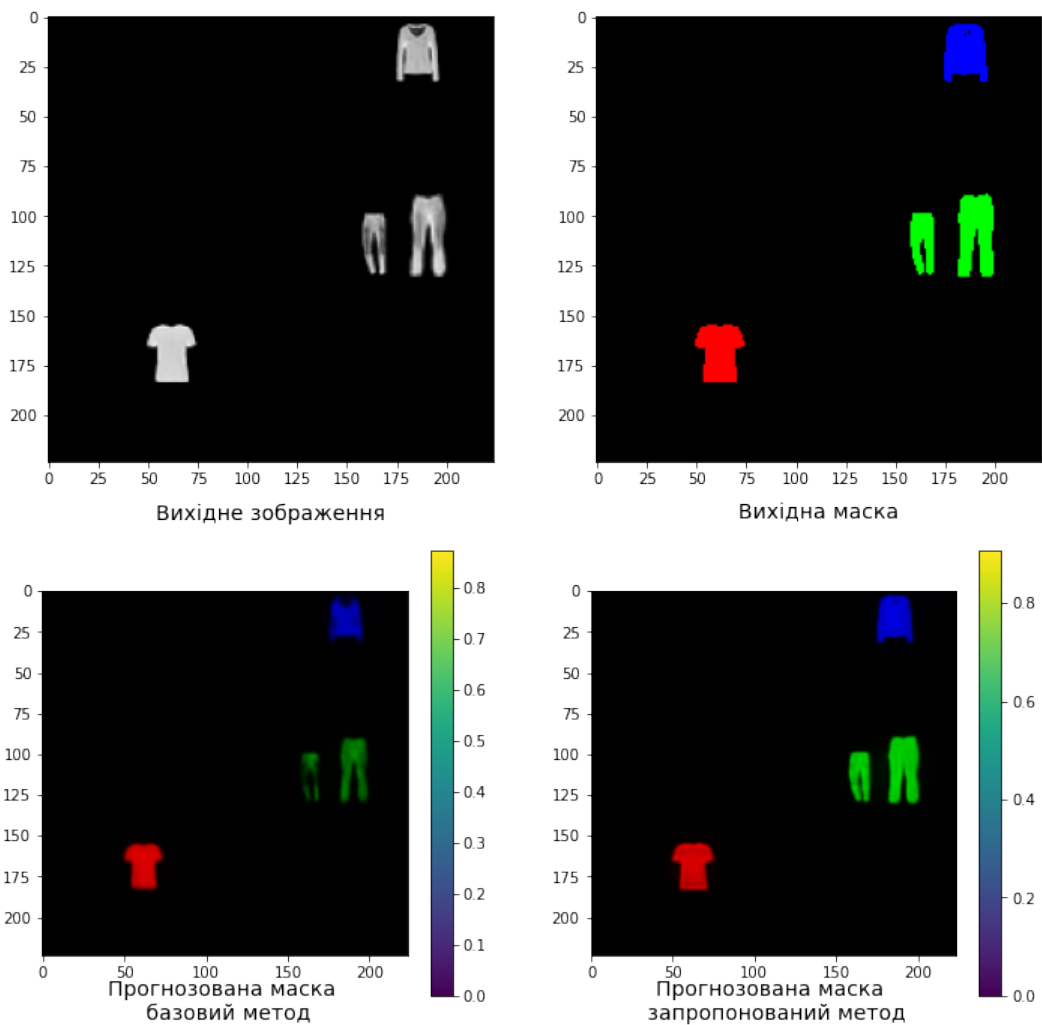


Рис. 1.7

На рисунках 1.8 та 1.9 зображено UMAP-проекцію ознак останнього

шару енкодера для набору тестових зображень.

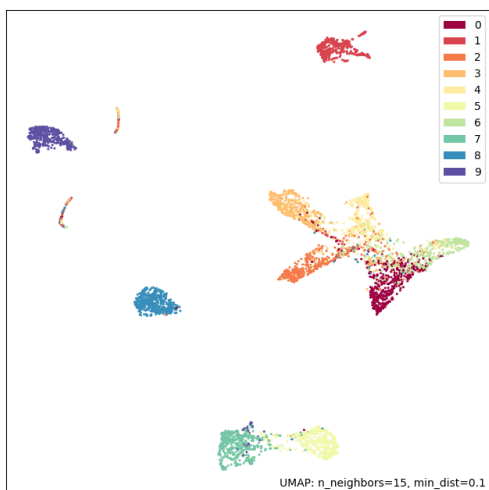


Рис. 1.8: UMAP ознак
базової моделі

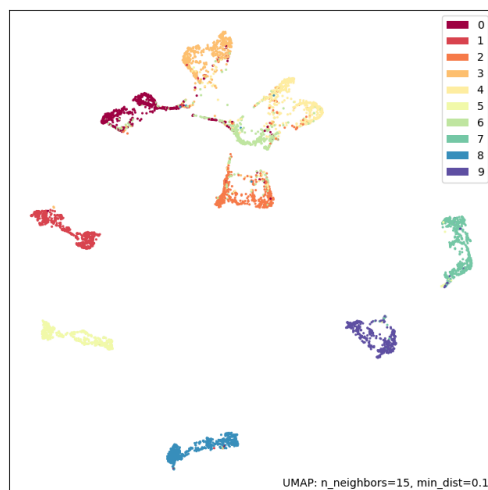


Рис. 1.9: UMAP ознак
запропонованої моделі

Як і в попередньому експерименті, в умовах зашумленої розмітки, запропонований метод дає більш впевнені прогнози, також зберігаючи більше малих деталей.

Експеримент 4. Порівняння базового та запропонованого методів на складному наборі даних з простими об'єктами та зашумленням розмітки. Параметри генерації набору даних зазначено в таблиці 1.7.

Таблиця 1.7: Параметри набору даних експеримента

Об'єкти	MNIST
Текстура фону	Imagenette
Текстура об'єктів	Imagenette
P_e	0.1
P_d	0.25
N_{obj}	10
S_{obj}	36
Δ_{max}	0.2

Результати експерименту наведено в таблиці 1.8.

Таблиця 1.8: Результати експеримента

Модель	Міра Дайса	F1-міра	K_{sep}
Базова модель	0.70	-	7.02
Запропонований метод	0.75	0.78	10.6

На рисунку 1.10 зображено приклади прогнозів нейронної мережі для тестових зображень.

На рисунках 1.11 та 1.12 зображено UMAP-проекцію ознак останнього шару енкодера для набору тестових зображень.

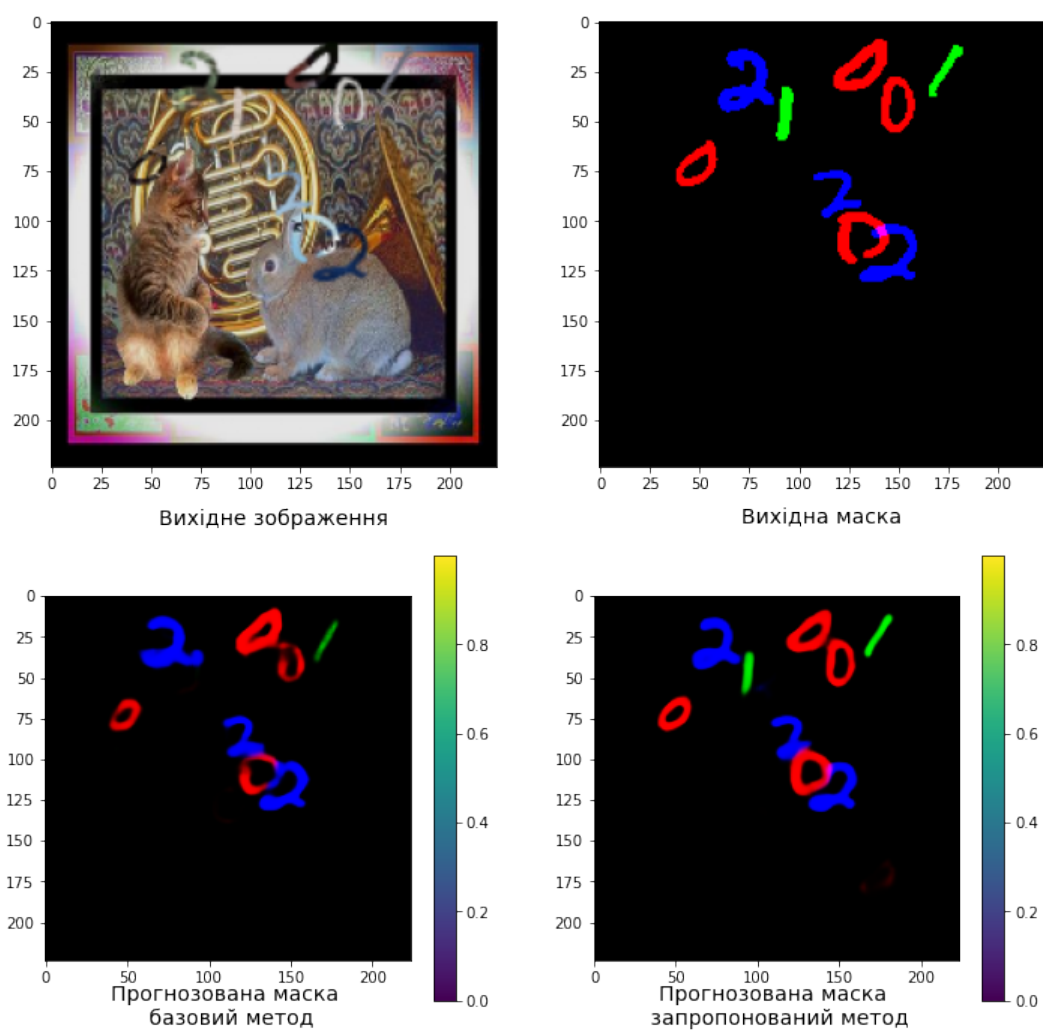


Рис. 1.10

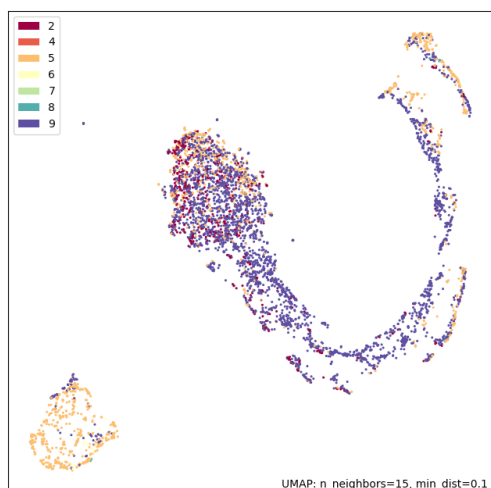


Рис. 1.11: UMAP ознак
базової моделі

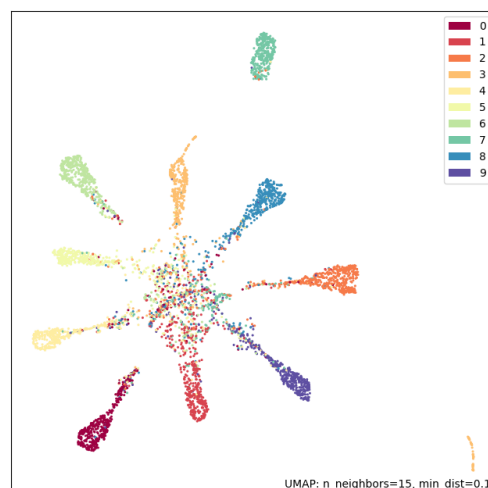


Рис. 1.12: UMAP ознак
запропонованої моделі

Як і в попередньому експерименті, в умовах зашумленої розмітки, запропонований метод дає більш впевнені прогнози, також зберігаючи більше малих деталей. Однак, внутрішні представлення мають гіршу роздільність через більш складний тип вхідних даних.

Експеримент 5. Порівняння базового та запропонованого методів на складному наборі даних з ускладненими об'єктами та зашумленням розмітки. Параметри генерації набору даних зазначено в таблиці 1.9.

Таблиця 1.9: Параметри набору даних експеримента

Об'єкти	FashionMNIST
Текстура фону	Imagenette
Текстура об'єктів	Imagenette
P_e	0.1
P_d	0.25
N_{obj}	10
S_{obj}	36
Δ_{max}	0.2

Результати експерименту наведено в таблиці 1.10.

Таблиця 1.10: Результати експеримента

Модель	Міра Дайса	F1-міра	K_{sep}
Базова модель	0.70	-	7.02
Запропонований метод	0.75	0.78	10.6

На рисунку 1.13 зображено приклади прогнозів нейронної мережі для тестових зображень.

На рисунках 1.14 та 1.15 зображено UMAP-проекцію ознак останнього шару енкодера для набору тестових зображень.

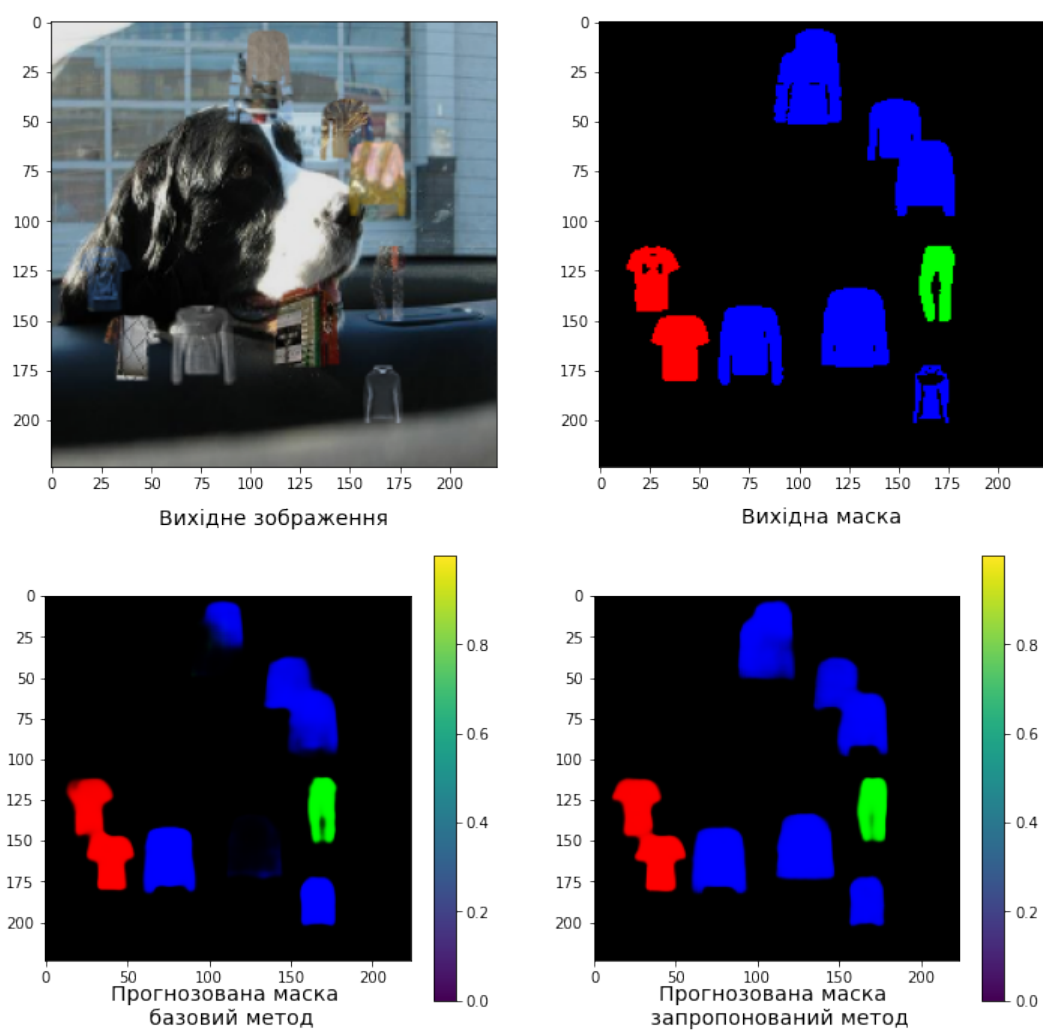


Рис. 1.13

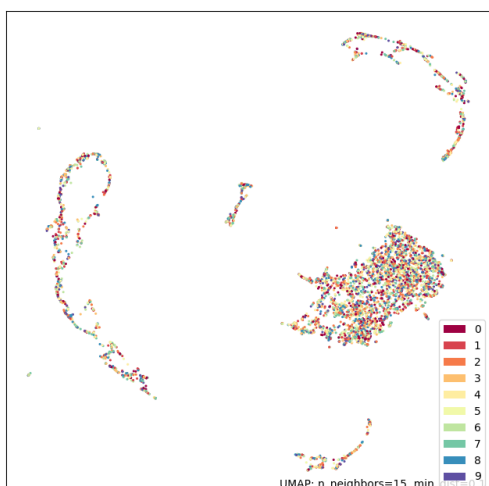


Рис. 1.14: UMAP ознак
базової моделі

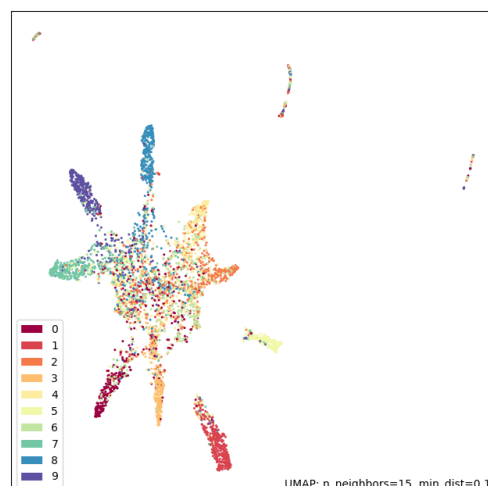


Рис. 1.15: UMAP ознак
запропонованої моделі

Як і в попередньому експерименті, в умовах зашумленої розмітки, запропонований метод дає більш впевнені прогнози, також зберігаючи більше малих деталей. Однак, внутрішні представлення мають гіршу роздільність через більш складний тип вхідних даних.

Експеримент 6. Стійкість методів до зашумлення розмітки на простих даних.

Для оцінки стійкості базового та запропонованого методів до сили зашумлення розмітки, проведено серію експериментів зі зміною вірогідності зашумлення для кожного з об'єктів в наборі даних.

Результати експерименту наведено в таблиці 1.11

Таблиця 1.11: Порівняння стійкості моделей до шуму

P_e	P_d	Міра Дайса Базовий метод	Міра Дайса Запропонований метод
0.1	0.1	0.84	0.96
0.1	0.25	0.81	0.91
0.25	0.1	0.65	0.83
0.5	0.25	0.51	0.62
0.25	0.5	0.57	0.69
0.5	0.5	0.3	0.47
0.75	0.75	0.12	0.23

З таблиці видно, що запропонований метод є більш стійким до більшого рівня зашумлення розмітки. Також, як запропонований, так і базовий методи менш стійкі до підвищеної ймовірності ерозії.

Експеримент 7. Стійкість методів до зашумлення розмітки на складних даних.

Для перевірки стійкості до шуму розмітки в умовах наближених до реальних, попередній експеримент проведено з використанням фону та текстур об'єктів з набору даних Imagenette. Результати експерименту наведено в таблиці 1.12.

Таблиця 1.12: Порівняння стійкості моделей до шуму

P_e	P_d	Міра Дайса Базовий метод	Міра Дайса Запропонований метод
0.1	0.1	0.72	0.78
0.1	0.25	0.70	0.75
0.25	0.1	0.61	0.69
0.5	0.25	0.48	0.53
0.25	0.5	0.55	0.60
0.5	0.5	0.21	0.41
0.75	0.75	0.08	0.15

Результати експерименту збігаються з результатами попереднього експерименту, хоча мають тенденцію до зменшення середнього значення метрики через підвищену складність даних.

Дослідження внеску окремих компонентів. Параметри генерації набору даних зазначено в таблиці 1.13.

Таблиця 1.13: Параметри набору даних для проведення дослідження внеску окремих компонентів

Текстура фону	Imagenette
Текстура об'єктів	Imagenette
P_e	0.2
P_d	0.2
N_{obj}	10
S_{obj}	36
Δ_{max}	0.2

В таблиці 1.14 наведено значення метрики при послідовному додаванні складових частин запропонованого метода до базової моделі. Використані скорочення:

- MN - MNIST
- FM - FashionMNIST
- IN - ImageNette
- C - константний колір

Таблиця 1.14: Дослідження внеску окремих компонентів

Змінна	MN/C/C	FM/C/C	MN/C/IN	MN/IN/IN
Базова модель		-		
+ декодер класифікації				
+ обмеження функції втрат				
+ комбінація прогнозів				

З зазначеної таблиці видно, що найбільший приріст дають обмеження функції втрат та комбінація прогнозів, хоча для отримання повного ефекту необхідні всі складові частини запропонованого методу.

1.3. Експерименти на реальних даних

1.3.1. Класифікація стадії осередків діабетичної ретинопатії.

1.3.1.1. Опис наборів даних. Використані в цьому експерименті зображень, були взяті з декількох наборів даних. Для претренування нейронних мереж було використано відкритий набір даних від Kaggle: “Diabetic Retinopathy Detection Challenge 2015” dr2015.

Цей набір даних є найбільшим із загальнодоступних. Він складається з 35126 фотографій очного дна для лівого та правого ока американських громадян, з розміченими стадіями діабетичної ретинопатії за стандартним протоколом:

- Відсутність діабетичної ретинопатії (мітка 0)
- Легка діабетична ретинопатія (мітка 1)
- Помірна діабетична ретинопатія (мітка 2)
- Важка діабетична ретинопатія (мітка 3)
- Проліферативна діабетична ретинопатія (мітка 4)

Крім того, були використані менші набори даних: набір зображень індійської діабетичної ретинопатії (IDRiD) idrid, з якого використано 413 фотографій очного дна, та MESSIDOR (Методи оцінки методів сегментації та

індексації в області офтальмології сітківки) messidor, з якого використано 1200 фотографій очного дна. Оскільки розмітка оригінального набору даних MESSIDOR відрізняється від інших наборів даних, ми використали версію, яку група офтальмологів messidor_relabeled помаркувала відповідно до стандартного протоколу.

Усі зазначені набори даних мають однаковий розподіл міток класів, що є фундаментальною властивістю для цієї задачі.

Оцінка проводиться на наборі даних Kaggle APTOS2019 aptos2019, дослідники мають доступ лише до тренувальної та валідаційної частин. Повний набір даних складається з 18590 фотографій очного дна, які розділені на набори з 3662 тренувальних, 1928 валідаційних та 13000 тестових зображень, що були розділені організаторами змагань Kaggle.

1.3.1.2. Детальний опис структури нейронної мережі.

1.3.2. Детальний опис процедури навчання.

1.3.2.1. Ініціалізація нейронної мережі. Відпочатку, енкодер ініціалізується параметрами нейронної мережі, що була натренована на наборі даних ImageNet. Параметри декодерів ініціалізуються випадково (ініціалізація Хе).

1.3.2.2. Трансферне навчання. Використання трансферного навчання можливо, оскільки природні особливості діабетичної ретинопатії узгоджуються між різними людьми і не залежать від набору даних. Крім того, різні набори даних збираються на різному обладнанні. Включення цих знань у модель підвищує її здатність до узагальнення та підвищує важливість природних ознак за рахунок зменшення чутливості до особливостей обладнання.

Для попереднього навчання, ініціалізована нейронна мережа навчається на протязі 20 епох на наборі даних DRDC2015 за допомогою стохастич-

чного градієнтного спуску. Основна мета попереднього навчання - створити ініціалізацію параметрів на розподілі даних, що є близьким до цільового. Після попереднього навчання, параметри нейронної мережі використовуються як ініціалізація для основного навчання.

Під час попереднього навчання, кожен декодер мінімізує свою функцію втрат: перехресну ентропію для класифікаційного декодера, бінарну перехресну ентропію для декодера порядкової регресії та середню абсолютну похибку для декодера регресії.

Основне тренування проводиться на наборах даних APTOS2019, IDR1D та MESSIDOR разом. Починаючи з ваг, отриманих на етапі попереднього тренування, виконується 5-кратна перехресна перевірка та оцінка моделі на відкладеному наборі даних. На цьому етапі функції втрат для декодерів змінено: Focal loss focalloss для класифікаційного декодера, бінарна Focal loss focalloss для декодера порядкової регресії та середньоквадратична помилка для декодера регресії.

Зміна функцій втрат на даному етапі пов'язана з необхідністю знизити значення втрат для зразків, що вже коректно класифіковані, щоб зменшити вплив ознак, що не відносяться до захворювання.

Для кожної з підвиборок було натреновано одну модель, протягом 75 епох, використовуючи алгоритм оптимізації Rectified Adam, radam, зі змінною темпу навчання за косинусним законом. Щоб зберегти параметри енкодера, поки декодери перебувають у випадковому стані, оновлення параметрів енкодера не відбувається на протязі перших п'яти епох навчання.

1.3.2.3. Комбінування результатів декількох задач. На етапі після основного навчання, для кожної з моделей, навчається модель лінійної регресії з виходів декодерів в єдине значення.

Лінійна регресія навчається після основного навчання, оскільки в іншому випадку, сходиться до неоптимальних локальних мінімумів з вагами

двох декодерів, близьких до нуля. Ці нульові ваги запобігають оновленню відповідних ваг декодерів і, відповідно, запобігають навчанню. Початкові ваги для кожного з виходів декодерів були встановлені рівними $1/3$, а потім тренувались протягом п'яти епох, щоб мінімізувати середньоквадратичну функцію помилки.

1.3.2.4. Результати експерименту.

1.4. Класифікація раку шкіри та сегментація родимок

1.4.1. Опис набору даних SIIM-ISIC Melanoma Classification.

Дані зображень, використані в цьому дослідженні, були взяті з декількох наборів даних з однаковою структурою: SIIM&ISIC з 2017, 2018, 2019 та 2020 років. Ці набори даних були створені Міжнародною співпрацею з обробки зображень шкіри (англ. International Skin Imaging Collaboration - ISIC), а зображення отримані з наступних джерел:

- лікарня Клінік де Барселона
- Віденський медичний університет
- Центр раку Меморіал Слоун Кеттерінг
- Австралійський інститут меланоми
- Квінслендський університет
- Афінська медична школа

Загалом, ці набори даних складаються із приблизно 50000 RGB-зображень, з яких близько 3000 мають зображення злоякісних уражень. Набір даних містить 434 повторюваних зображення. Окрім даних про зображення, були надані метадані про пацієнтів. Один пацієнт має декілька зображень різних родинок.

Зображення та метадані надані у форматі DICOM, який є загальнозживаним форматом даних медичних зображень. Крім того, набір даних доступний у форматі JPEG із розмірами зображень, змінених до 1024x1024.

Метадані також надаються за межами формату DICOM, у файлах CSV [12].

Метадані містять наступну інформацію:

- `image_name` - унікальний ідентифікатор, вказує на ім'я файлу відповідного зображення DICOM;
- `patient_id` - унікальний ідентифікатор пацієнта;
- `sex` - стать пацієнта;
- `age_approx` - приблизний вік пацієнта на момент фотографії (ціле число);
- `anatom_site_general_challenge` - розташування родинки на тілі;
- `diagnosis` - детальна інформація про діагностику (рядок);
- `benign_malignant` - показник злоякісності зображеного ураження (`cnhjrf`, одна з "benign" і "malignant");
- `target` - бінарізована версія цільової змінної.

Значення для `anatom_site_general_challenge` та `diagnosis` беруться із заздалегідь визначеного кінцевого набору.

Метадані доступні для кожного пацієнта, тому різні зображення можуть мати однаковий набір функцій рівня пацієнта. Ми використовуємо всі доступні метадані, крім ідентифікатора пацієнта та діагностики, оскільки вони доступні лише в навчальних наборах даних.

Набір даних має дисбаланс високого класу. Розподіл діагнозів показано на рис. 1. Для значень `unknown` в полі `diagnosis`, автори набору даних гарантують, що новоутворення не є злоякісним [13].

Через те, що зображення в наборі даних з різних джерел та мають різні стандарти візуалізації вони мають структурований шум у вигляді лінійних смуг, областей, позначених ручкою, центруючих ліній тощо. Залежно від місця новоутворення та статі, також спостерігається волосся. Всі ці додаткові ознаки можуть значно вплинути на тренувальний процес і можуть призвести до перенавчання.

1.4.2. Детальний опис процедури навчання. В даному експерименті використано одноступеневий навчальний процес, який включає одночасно трансферне навчання та багатозадачне навчання.

Навчання проводиться у 3-кратній схемі перехресної перевірки. Енкодер відпочатку ініціалізується за допомогою вагів, отриманих за допомогою алгоритму [25]. У наших експериментах ми спостерігали, що така ініціалізація призводить до стабільно кращих результатів, ніж ініціалізація Imagenet.

1.4.3. Результати експерименту.

1.5. Класифікація та сегментація формацій хмар

1.5.1. Опис набору даних Understanding Clouds from Satellite Images. Набір даних Understanding Clouds from Satellite Images (UCSID) складається з 10000 RGB зображень, зроблених з двох супутників TERRA та AQUA, що знаходяться на полярній орбіті [6]. Кожен з цих супутників проходить над певну область один раз на день. Через обмежене поле зору камер, встановлених на цих супутниках, кожне зображення зшито з двох супутників, що знаходяться над однією й тою самою областю одночасно, але на різних орбітах. Решта зображення, для якої не було знято даних під час прольоту супутників (між орбітами) заповнена чорним кольором. Оскільки зображення підлягають компресії, є артефакти стиснення, та чорний колір має деякі незначні аномалії.

На знімках є регіони, які містять певні хмарні утворення, та помічені дослідниками відповідно: Риба, Квітка, Гравій, Цукор (англ. Fish, Flower, Gravel, Sugar). Кожне зображення має принаймні одне хмарне утворення і може містити до всіх чотирьох одночасно.

Розмітка для цих регіонів була створена під час краудсорсингу в Інституті метеорології імені Макса Планка в Гамбурзі, Німеччина, та Laboratoire de météorologie Dynamique у Парижі, Франція. Команда з 68 вчених

виявила ділянки з хмарами на кожному зображенні, і кожне зображення було розмічене, в середньому, трьома різними вченими. Кожен вчений мав виділити хмари за допомогою прямокутних областей на власний розсуд. Основна розмітка була створена об'єднанням областей, що були розічені всіма вченими для цього зображення, після видалення чорної смуги з цих областей [7].

Через те, як був зібраний UCSID, він має значну кількість пікселів шуму масках для хмар. Оскільки маски складаються з прямокутників, які повністю перекривають хмару, існує багато пікселів, які позначені як хмари, але, насправді, відповідають фону. Також, не всі хмари позначені масками. Через об'єднання масок від різних анотаторів, класи можуть суттєво перекриватися, включаючи випадки, коли всі чотири класи присвоюються одним і тим самим пікселям.

Усі зображення мають вихідну роздільну здатність 2100x1400 пікселів. Більшість зображень містять більше ніж один клас хмар. Немає зображень без хмар. Розподіл класів хмар на зображеннях показано на рисунку.

Різні типи хмар зазвичай зустрічаються разом, розподіл комбінацій різних типів хмар зазначено на рисунку. UCSID розділено на 5546 тренувальних, та 3698 валідаційних зображень. Крім того, немає доступу до міток валідаційного набору даних. Значення метрики можливо отримати через систему валідації.

До розподілу міток класів в наборі даних не було виконано змін (недискретизація, передискретизація тощо).

1.5.2. Детальний опис структури нейронної мережі.

1.5.3. Результати експерименту.