

Міністерство освіти і науки України
Тернопільський національний технічний університет імені Івана Пулюя

Факультет комп'ютерно-інформаційних систем і програмної інженерії
(повна назва факультету)

Кафедра програмної інженерії
(повна назва кафедри)

КВАЛІФІКАЦІЙНА РОБОТА

на здобуття освітнього ступеня

(назва освітнього ступеня)
на тему: Розробка архітектури базової нейронної мережі для сегментації меж та
границь на зображеннях

Виконав(ла): студент(ка) IV курсу, групи СП-42
спеціальності

121 Інженерія програмного забезпечення
(шифр і назва спеціальності)

	(підпис)	Мельник О.О. (прізвище та ініціали)
Керівник	(підпис)	Бойко І.В. (прізвище та ініціали)
Нормоконтроль	(підпис)	Стоянов Ю.М. (прізвище та ініціали)
Завідувач кафедри	(підпис)	Петрик М.Р.. (прізвище та ініціали)
Рецензент	(підпис)	(прізвище та ініціали)

Тернопіль 2025

Міністерство освіти і науки України
Тернопільський національний технічний університет імені Івана Пулюя

Факультет Комп'ютерно інформаційних систем і програмної інженерії
(повна назва факультету)

Кафедра Програмної інженерії
(повна назва кафедри)

ЗАТВЕРДЖУЮ

Завідувач кафедри

Петрик М. Р.
(підпис) (прізвище та ініціали)
« » 2025 р.

З А В Д А Н Н Я НА КВАЛІФІКАЦІЙНУ РОБОТУ

на здобуття освітнього ступеня бакалавр
(назва освітнього ступеня)

за спеціальністю 121 Інженерія програмного забезпечення
(шифр і назва спеціальності)

студенту Мельник Олександр Олександрович
(прізвище, ім'я, по батькові)

1. Тема роботи Розробка архітектури базової нейронної мережі для сегментації меж та границь на зображеннях

Керівник роботи д.ф-м.н., доц. Бойко І.В.
(прізвище, ім'я, по батькові, науковий ступінь, вчене звання)

Затверджені наказом ректора від « » 2025 року №

2. Термін подання студентом завершеної роботи

3. Вихідні дані до роботи Предметна область, технічне завдання, вимоги та специфікація, програмне рішення

4. Зміст роботи (перелік питань, які потрібно розробити)

Вступ. Розділ 1. Аналіз предметної області та задач. Розділ 2. Розробка структури нейромережі із функціоналом для опрацювання зображень та виявлення країв та меж у них. Розділ 3.

Навчання розробленої нейронної мережі. вхідні тренувальні датасети. Розділ 4. Безпека життєдіяльності та основи охорони праці. Висновок.

Список використаних джерел. Додатки.

5. Перелік графічного матеріалу (з точним зазначенням обов'язкових креслень, слайдів)

Діаграми, знімки екрану з проробленою роботою, ілюстративні зображення, інформативні зображення для доповнення тексту.

6. Консультанти розділів роботи

Розділ	Прізвище, ініціали та посада консультанта	Підпис, дата	
		завдання видав	завдання прийняв
Безпека життєдіяльності, основи охорони праці			

7. Дата видачі завдання

КАЛЕНДАРНИЙ ПЛАН

№ з/п	Назва етапів роботи	Термін виконання етапів роботи	Примітка
1	Ознайомлення з завданням до кваліфікаційної роботи		
2	Підбір джерел по темі кваліфікаційної роботи		
3	Опрацювання джерел інформації по темі кваліфікаційної роботи		
4	Розробка архітектури програмної системи		
5	Розробка структури нейромережі із функціоналом для опрацювання зображень та виявлення країв та меж у них		
6	Навчання розробленої нейронної мережі та тестування її роботи		
7	Виконання звіту по проробленій роботі		
8	Виконання завдання до підрозділу «Безпека життєдіяльності»		
9	Виконання завдання до підрозділу «Основи охорони праці»		
10	Оформлення кваліфікаційної роботи		
11	Нормоконтроль		
12	Перевірка на плагіат		
13	Попередній захист кваліфікаційної роботи		
14	Захист кваліфікаційної роботи		

Студент

(підпис)

Мельник О. О.

(прізвище та ініціали)

Керівник роботи

(підпис)

Бойко І.В.

(прізвище та ініціали)

АНОТАЦІЯ

Розробка архітектури базової нейронної мережі для сегментації меж та границь на зображеннях // Кваліфікаційна робота освітнього рівня «Бакалавр» // Мельник Олександр Олександрович // Тернопільський національний технічний університет імені Івана Пулюя, факультет комп'ютерно-інформаційних систем і програмної інженерії, кафедра програмної інженерії, група СП-42 // Тернопіль, 2025 // Ст. – 79, рис. – 18, табл. – 5, кресл. – 0, додат. – , бібліогр. – 42.

Ключові слова: нейронна мережа, сегментація зображень, обробка зображень, межі та границі, комп'ютерний зір, глибоке навчання, архітектура моделі, класифікація пікселів, конволюційна нейронна мережа, семантична сегментація.

Дана кваліфікаційна робота на здобуття освітнього ступеня бакалавра присвячена питанням обробки, сегментації та аналізу зображень за допомогою нейронних мереж. Виявлення контурів є фундаментальним завданням комп'ютерного зору, яке відіграє важливу роль у таких застосуваннях, як розпізнавання об'єктів, розуміння сцени та сегментація зображень. На відміну від сучасних методів, що використовують попередньо навчені ваги та складні налаштування гіперпараметрів, наша запропонована модель пропонує новий підхід, який усуває ці залежності, зберігаючи при цьому високу точність та ефективність.

Однією з ключових переваг нашої моделі є її ефективність у навчанні. У той час як більшість моделей на основі глибокого навчання для виявлення контурів потребують приблизно 20 епох для збіжності, наша модель навчається з нуля та досягає збіжності лише за 9 епох. Це суттєве скорочення часу навчання не впливає на узагальнюючу здатність моделі, оскільки вона демонструє високу продуктивність у різноманітних і раніше невідомих сценах.

Крім того, у дипломній роботі представлено ретельно відібраний набір даних, спеціально розроблений для завдання виявлення контурів. Цей набір є одним із найбільш якісних серед доступних, що дозволяє значно покращити узагальнення

алгоритмів, заснованих на навчанні. Унікальні характеристики цього набору даних забезпечують ефективну генералізацію моделей, що підвищує надійність та застосовність методів виявлення контурів.

Таким чином, розвинений підхід пропонує більш ефективне та адаптивне рішення для виявлення контурів. Усунення залежності від попередньо навчених ваг та складних налаштувань гіперпараметрів, а також значне скорочення часу навчання дозволяє нам створити метод, який поєднує обчислювальну ефективність із високою узагальнюючою здатністю. Крім того, представлення спеціалізованого набору даних сприяє подальшому розвитку цієї галузі, забезпечуючи міцну основу для майбутніх досліджень у сфері виявлення контурів та суміжних напрямків.

ABSTRACT

Development of a Basic Neural Network Architecture for Image Edge and Boundary Segmentation // Qualification work for the educational level ‘Bachelor’ / Melnyk Oleksandr Oleksandrovykh // Ternopil Ivan Puluj National Technical University, Faculty of Computer Information Systems and Software Engineering, Department of Software Engineering, group SP-42 // Ternopil, 2025 // Pp. – 79, Fig. – 18, Table – 5, Drawings – 0, Supplement – , Bibliography – 42.

Keywords: neural network, image segmentation, image processing, edges and boundaries, computer vision, deep learning, model architecture, pixel classification, convolutional neural network, semantic segmentation.

This qualification work for obtaining a Bachelor's degree is dedicated to the issues of image processing, segmentation, and analysis using neural network. Edge detection is a fundamental task in computer vision, playing a crucial role in applications such as object recognition, scene understanding, and image segmentation. Unlike state-of-the-art methods that rely on pretrained weights and complex hyperparameter tuning, our proposed model introduces a novel approach that eliminates these dependencies while maintaining high accuracy and efficiency.

One of the key advantages of our model is its training efficiency. While most deep learning-based edge detection models require approximately 20 epochs to converge, our model is trained from scratch and achieves convergence in just 9 epochs. This significant reduction in training time does not compromise the generalization capability of the model, as it demonstrates high performance across diverse and previously unseen scenes.

Furthermore, we present a highly curated dataset specifically designed for edge detection tasks. This dataset is among the most refined collections available, significantly enhancing the generalization capabilities of learning-based algorithms. The unique characteristics of this dataset ensure that models trained on it can effectively generalize to new environments, thereby improving the reliability and applicability of edge detection techniques.

Our approach provides a more efficient and adaptable solution for edge detection. By removing the dependency on pretrained weights and complex hyperparameter settings while significantly reducing training time, we offer a method that balances computational efficiency with strong generalization capability. Additionally, the introduction of our specialized dataset contributes to advancements in the field, providing a solid foundation for future research in edge detection and related domains.

Beyond efficiency and accuracy, our model is designed to be versatile and scalable, allowing it to be seamlessly integrated into various real-world applications. Whether for medical imaging, autonomous navigation, or industrial inspection, our approach enables robust edge detection with minimal computational overhead. By eliminating the reliance on extensive pretraining, we make edge detection more accessible and adaptable to diverse conditions, ensuring that our method remains effective even in challenging environments.

In conclusion, our work not only advances the state of edge detection technology but also sets a new standard for efficient model training and dataset curation. By addressing key challenges in the field, we pave the way for more reliable, generalizable, and computationally efficient edge detection techniques that can be readily applied across a wide range of domains.

ЗМІСТ

ВСТУП.....	10
РОЗДІЛ 1. ДОСЛІДЖЕННЯ ТА АНАЛІЗ ПРЕДМЕТНОЇ ОБЛАСТІ.	
АКТУАЛІЗАЦІЯ ТА ФОРМАЛЬНИЙ АНАЛІЗ ПРОБЛЕМИ	12
1.1. Значення візуальної обробки інформації в конкурентних нейронних мережах при обробці зображень.....	12
1.2. Поняття про краї, контури та межі в зображеннях та методи їх встановлення.....	16
1.3. Значення структури даних для графічної інформації.....	17
1.4. Існуючі суміжні розробки та підходи.	20
РОЗДІЛ 2 РОЗРОБКА СТРУКТУРИ НЕЙРОМЕРЕЖІ ІЗ ФУНКЦІОНАЛОМ ДЛЯ ОПРАЦЮВАННЯ ЗОБРАЖЕНЬ ТА ВИЯВЛЕННЯ КРАЇВ ТА МЕЖ У НИХ.....	
2.1. Розробка щільної екстремальної початкова мережа для виявлення країв у зображеннях.....	24
2.2. Розробка щільної екстремальної початкової мережі для виявлення країв у зображеннях.....	26
2.3. Покрокова розробка нейромережі USNet для збільшення роздільної здатності.....	27
2.4. Визначення функції втрат при роботі з розробленою нейронною мережею та її розрахунок.	29
РОЗДІЛ 3 НАВЧАННЯ РОЗРОБЛЕНОЇ НЕЙРОННОЇ МЕРЕЖІ. ВХІДНІ ТРЕНУВАЛЬНІ ДАТАСЕТИ.	
3.1. Навчання нейронної мережі з використанням набору даних VIPED.....	31
3.2. Тестування із еталонними контрольними наборами даних.....	33
3.3. Аналіз отриманих результатів, верифікація та тестування.....	37

3.4. Оцінка та основні результати щодо верифікації. Налаштування архітектури розробленого програмного забезпечення	39
3.5. Виконання кількісного порівняння методів	44
3.6. Якісне порівняння методів	47
РОЗДІЛ 4 БЕЗПЕКА ЖИТТЄДІЯЛЬНОСТІ, ОСНОВИ ОХОРОНИ ПРАЦІ.....	51
4.1. Принципи, способи та засоби захисту населення	51
4.2. Заходи, що запобігають впливу на людину агресивних та токсичних речовин, які використовуються в технологічному процесі.	54
ВИСНОВКИ	58
СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ.....	60
ДОДАТКИ	64
Додаток А. Диск із кваліфікаційною роботою бакалавра	65

ВСТУП

Виявлення країв є ключовим завданням у багатьох застосуваннях комп'ютерного зору, оскільки саме вони формують основу для аналізу зображень та подальшої обробки вищого рівня. Виявлення країв використовується в розпізнаванні об'єктів, аналізі сцени, тривимірному реконструюванні та багатьох інших завданнях. Незважаючи на значні досягнення в цій галузі, проблема точного та надійного виділення країв залишається відкритою.

Сучасні методи здебільшого ґрунтуються на глибокому навчанні, де вирішальну роль відіграють два основні чинники: зміст навчального набору даних та архітектура нейронної мережі. Проте більшість загальнодоступних наборів даних не призначені спеціально для задач виявлення країв, що створює певні обмеження та ускладнює розвиток цієї галузі. Наприклад, найбільш поширені набори, такі як BSDS, були створені для задач сегментації сцен і містять розмічені контури об'єктів, а не детальні краї. Це ускладнює навчання моделей, які потребують точного розпізнавання країв, а не лише меж між об'єктами.

У рамках даного дослідження у дипломній роботі розглянуто проблему нестачі спеціалізованих наборів даних для задач виявлення країв. По-перше, проведено аналіз відмінностей між краями, контурами та межами, оскільки, попри їхню схожість, вони є різними візуальними характеристиками, які потребують окремих еталонних наборів даних. На основі цього підходу запропоновано новий набір даних, спеціально орієнтований на краї, який містить високоточні анотації для навчання моделей. Даний набір даних створено з урахуванням важливості деталізації та коректного розподілу країв по всій сцені, що дозволяє отримувати якісніші результати порівняно з існуючими рішеннями.

По-друге, розроблено нову архітектуру нейронної мережі, що отримала назву Dense Extreme Inception Network for Edge Detection (DexiNed). Особливістю цієї архітектури є можливість навчання з нуля без використання попередньо натренованих вагових коефіцієнтів. Це дозволяє моделі гнучко адаптуватися до

нового набору даних та досягати високої продуктивності навіть у складних умовах, таких як різні освітлення, текстури та складні сцени.

Проведені порівняння підтвердили, що DexiNed демонструє вищу продуктивність у порівнянні з іншими існуючими алгоритмами на запропонованому наборі даних. Крім того, модель проявляє здатність до узагальнення на інших наборах даних без необхідності додаткового налаштування. Висока якість результатів DexiNed також підтверджується візуально завдяки чіткішим та більш деталізованим краям, що генеруються мережею. Дослідження також демонструє перевагу запропонованої архітектури перед класичними підходами та сучасними моделями, які використовують попередньо навчені ваги, що підтверджується експериментальними результатами та метриками оцінювання.

Таким чином, результати дипломної роботи роблять значний внесок у розвиток методів виявлення країв, пропонуючи якісний еталонний набір даних та ефективну архітектуру нейронної мережі. Це, у свою чергу, сприяє підвищенню точності та продуктивності алгоритмів у різних сценаріях застосування, що є важливим кроком у вдосконаленні технологій комп'ютерного зору. Запропоновані методи можуть бути корисними для широкого спектру застосувань, зокрема в автоматичному аналізі зображень, робототехніці, системах автономного водіння та інших сферах, що потребують високоточних алгоритмів обробки візуальної інформації.

РОЗДІЛ 1. ДОСЛІДЖЕННЯ ТА АНАЛІЗ ПРЕДМЕТНОЇ ОБЛАСТІ. АКТУАЛІЗАЦІЯ ТА ФОРМАЛЬНИЙ АНАЛІЗ ПРОБЛЕМИ

1.1. Значення візуальної обробки інформації в конкурентних нейронних мережах при обробці зображень.

Краї забезпечують важливі підказки у візуальній обробці інформації, від класичних алгоритмів комп'ютерного зору в розпізнаванні зображень [2] до сучасних методів у генеративних конкурентних мережах (GAN) [3]. Виявлення країв відіграє ключову роль у багатьох застосуваннях, включаючи розпізнавання об'єктів, сегментацію зображень і розуміння сцен. Незважаючи на їхню важливість, надійне виявлення країв залишається відкритою проблемою в комп'ютерному зорі. Традиційні підходи часто покладаються на методи, що базуються на градієнтних методах, зокрема таких і як детектор країв Кенні, який використовує заздалегідь визначені пороги інтенсивності для виявлення країв. Однак ці методи зазнають труднощів у задачах із наявним шумом, змінами освітлення та областями, багатими на текстуру, що обмежує їхню ефективність у застосуванні до складних зображень.

Щоб продемонструвати труднощі виявлення країв, розглянемо чотири приклади зображень, які подані на Рис. 1. Ці зображення містять різні структури, включаючи об'єкти з складними внутрішніми деталями та різними рівнями контрасту. Третій-п'ятий ряди ілюструють результати трьох сучасних моделей (RCF [4], BDCN [5] і CATS [6]), навчених на наборі даних BSDS [1]. Візуально помітно, що жодна з цих моделей не відображає краї дрібних деталей належним чином. Наприклад, внутрішні краї дзвоноподібної скульптури або гелікоптера повністю ановані в еталонному зображенні, однак результати цих моделей не відображають ці деталі. Крім того, продуктивність цих моделей значно варіюється залежно від змісту сцени та характеристик об'єкта, що підкреслює їх обмеження у генералізації за межами навчального набору даних.

Подібне явище можна спостерігати у смугах будівлі та зебри. У цьому випадку, хоча ці деталі також виключені з еталонного зображення, краї чітко видно

на зображенні. Невідповідність між результатами моделей і людськими анотаціями свідчить про те, що існуючі набори даних можуть не повністю відображати всю складність реальних країв. Крім того, моделі мають труднощі з дрібними деталями

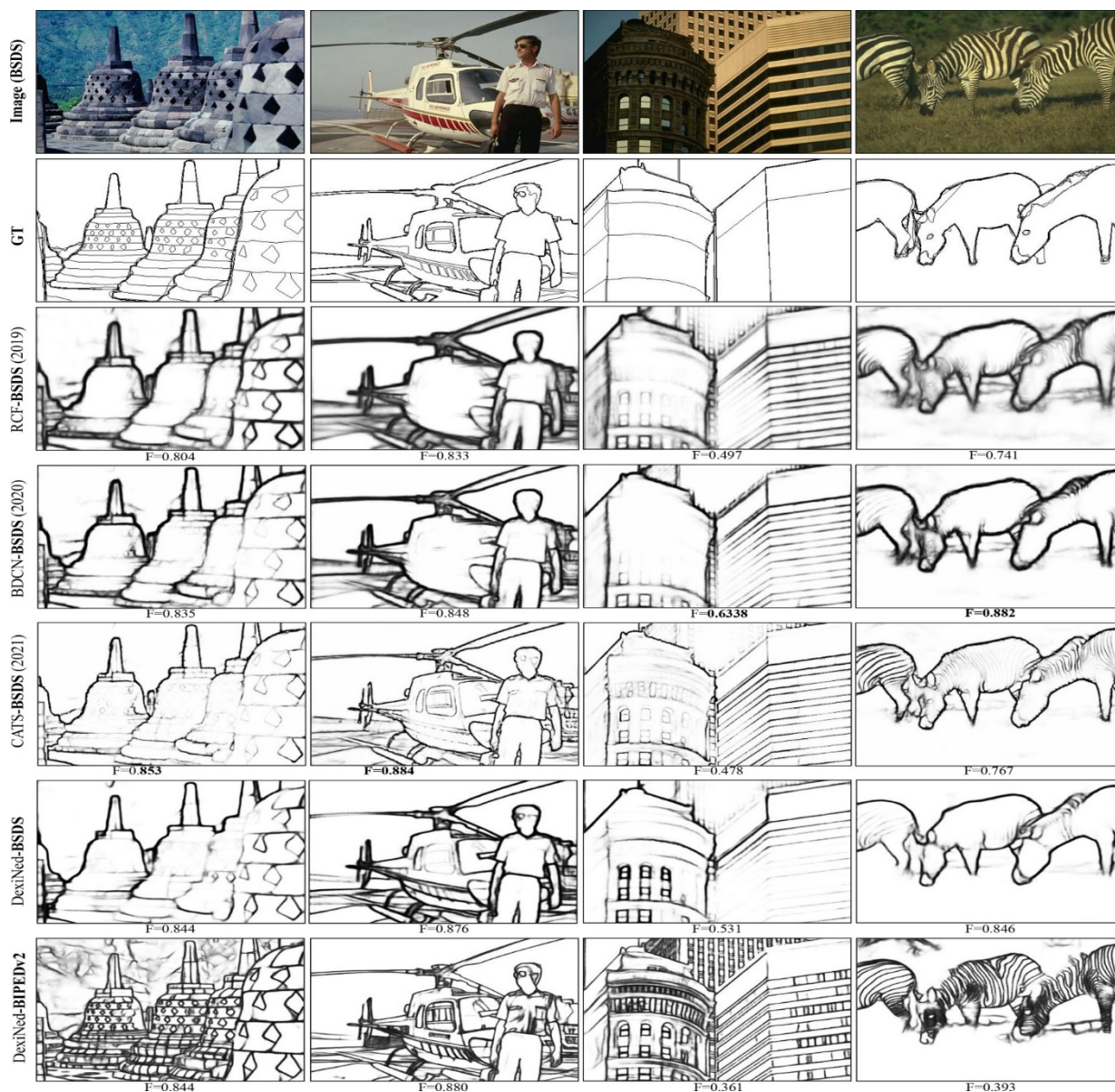


Рис. 1. Результати використання кількох алгоритмів на наборі даних BSDS [1]. Суфікс у назвах моделей (наприклад, BSDS і BIPEDv2) вказує на навчальний набір цієї моделі.

та варіаціями текстур, що призводить або до надмірного виявлення (появи хибнопозитивних результатів), або до недостатнього виявлення (пропуску важливих країв). Ця невідповідність створює труднощі в застосуваннях, де необхідне точне виявлення країв, наприклад, у медичній візуалізації та автономній навігації.

На основі цього слід зробити висновок, що поточні моделі виявлення країв потребують подальшого вдосконалення для надійного виявлення країв і кращої генералізації до нових сцен незалежно від їхнього навчального набору. Майбутні покращення у виявленні країв можуть отримати вигоду від гібридних підходів, що поєднують глибоке навчання з традиційними методами виявлення країв. Наприклад, інтеграція втрат, чутливих до країв, у навчання нейронних мереж, використання масштабних представлень ознак і включення контекстної інформації з навколишніх пікселів можуть підвищити надійність моделі. Крім того, розширення еталонних наборів даних для включення більшої кількості категорій зображень та анотування дрібних структур країв може покращити генералізацію моделей. Усунення цих викликів буде важливим для розробки надійних систем виявлення країв, які зможуть стабільно працювати у різних реальних застосуваннях.

Таким чином, можна ствердно сказати, що одним з основних обмежень сучасних підходів глибокого навчання (DL) є анотована еталонна розмітка у навчальному наборі даних. BSDS широко визнаний у спільноті як еталонний набір даних для навчання та оцінки алгоритмів виявлення країв. Однак цей набір даних спочатку був розроблений та анотований для сегментації сцен. Таким чином, еталонна розмітка відповідає в основному межах об'єктів високого рівня, а не низькорівневих країв. Оскільки моделі, засновані на глибокому навчанні, значною мірою залежать від навчальних даних, необхідно отримати незалежні набори даних для трьох завдань: виявлення країв, контурів та меж [7, 8]. Це має велике значення як для навчання, так і для оцінки. Мережі складно навчитися цим трьома концепціям з набору даних, який поєднує краї, контури та межі. Крім того, важко оцінити, наскільки добре або погано працює мережа в одному з цих завдань.

Щоб продемонструвати цю важливість на перший погляд, звернемо увагу на дві нижні рядки у Рис. 1. Єдина різниця між цими двома рядками – це навчальний набір. Коли наша мережа – Dense Extreme Inception Network for Edge Detection (DexiNed) – тренується на BSDS, вона стикається з тими ж проблемами, що й інші моделі. Однак, коли ми навчаємо DexiNed на нашому наборі даних (VIPEDv2), вона добре узагальнює результати на зображення BSDS у дрібномасштабних краях.

Зміст зображень цих двох наборів даних якісно відрізняється. Набір VIPEDv2 головним чином містить зображення міського середовища. Точно визначені краї на зображеннях із зброєю та скульптурою дзвона вказують на те, що модель DexiNed-VIPEDv2 є стійкою до нових сцен. Ця стійкість пояснюється ретельною ручною анотацією країв у запропонованому наборі даних. Ми забезпечили, щоб усі краї, як усередині, так і між об'єктами, були відображені в еталонних даних.

Для подальшої демонстрації значення правильного набору навчальних даних можна проаналізувати вплив навчання на якість виявлення країв у різних умовах освітлення та середовищах. У нашому дослідженні ми протестували DexiNed на різних категоріях зображень, включаючи природні ландшафти, архітектурні структури та штучні об'єкти. Результати показують, що при тренуванні на BSDS модель має тенденцію втрачати тонкі деталі, а при тренуванні на VIPEDv2 вона демонструє підвищену точність у визначенні складних контурів.

Ще одним ключовим аспектом є те, що VIPEDv2 включає широкий спектр варіацій у текстурі та освітленні, що покращує здатність моделі адаптуватися до нових середовищ. Наприклад, тестування на сценах із поганим освітленням показало, що модель, навчена на VIPEDv2, краще відтворює краї об'єктів, ніж її аналог, навчений на BSDS. Це підкреслює важливість не лише глибини моделі, а й якості її навчальних даних.

1.2. Поняття про краї, контури та межі в зображеннях та методи їх встановлення.

Виявлення країв є ключовим завданням у багатьох застосуваннях комп'ютерного зору. Сучасні методи здебільшого ґрунтуються на глибокому навчанні, де вирішальну роль відіграють два основні чинники: зміст навчального набору даних та архітектура нейронної мережі. Однак більшість загальнодоступних наборів даних не призначені спеціально для задач виявлення країв, що створює певні обмеження та ускладнює розвиток цієї галузі.

У рамках даної дипломної роботи буде розглянуто проблему нестачі спеціалізованих наборів даних для задач виявлення країв. По-перше, проведено аналіз відмінностей між краями, контурами та межами, оскільки, попри їхню схожість, вони є різними візуальними характеристиками, які потребують окремих еталонних наборів даних. Контури зазвичай описують межі всередині об'єктів, тоді як межі визначають зовнішні границі об'єктів у сцені. Визначення цих понять є критично важливим для побудови ефективних алгоритмів обробки зображень. На основі цього підходу запропоновано новий набір даних, спеціально орієнтований на краї. Відмінність цього набору від традиційних датасетів полягає в тому, що він включає більш точне анотоване представлення меж і контурів, що дозволяє нейронним мережам краще навчатися та узагальнювати отримані знання. Це є особливо важливим для різних практичних застосувань, таких як сегментація сцен, розпізнавання об'єктів, покращення якості зображень та інші завдання комп'ютерного зору.

По-друге, розроблено нову архітектуру нейронної мережі, що отримала назву Dense Extreme Inception Network for Edge Detection (DexiNed). Особливістю цієї архітектури є можливість навчання з нуля без використання попередньо натренованих вагових коефіцієнтів. На відміну від класичних алгоритмів, які покладаються на певні припущення щодо зображень, DexiNed здатна адаптуватися

до різних умов освітлення, текстур та складності об'єктів. Завдяки цьому покращується точність визначення меж і контурів навіть у складних сценах.

Проведені експерименти підтвердили, що DexiNed демонструє вищу продуктивність у порівнянні з іншими існуючими алгоритмами на запропонованому наборі даних. Крім того, модель проявляє здатність до узагальнення на інших наборах даних без необхідності додаткового налаштування. Це робить DexiNed універсальним інструментом, який можна застосовувати в різних сферах, включаючи медичну візуалізацію, автономні транспортні засоби та відеоспостереження. Висока якість результатів DexiNed також підтверджується візуально завдяки чіткішим та більш деталізованим краям, що генеруються мережею. Це дозволяє отримувати високоточні результати навіть у випадках, коли традиційні алгоритми виявлення країв демонструють обмежену ефективність.

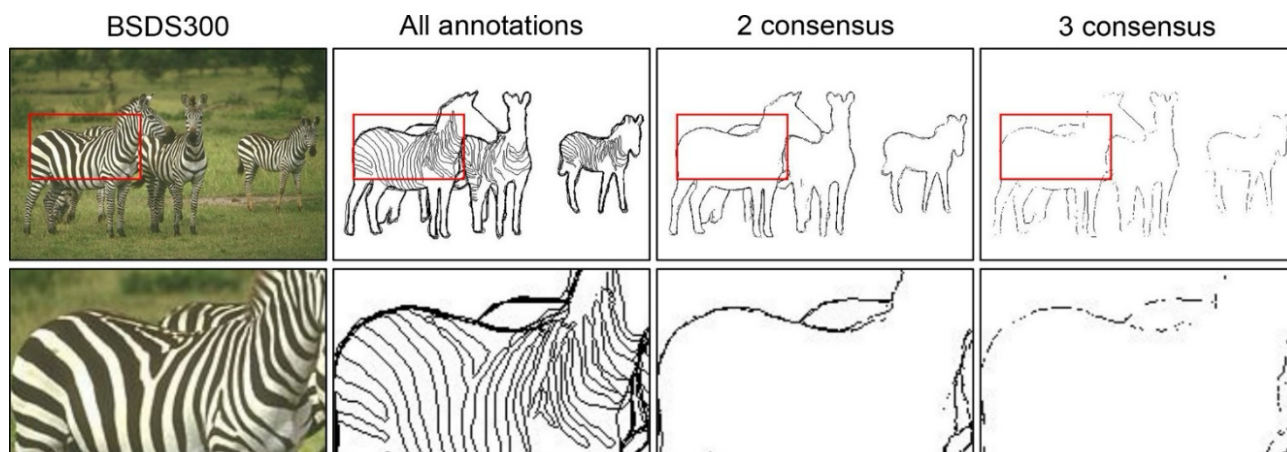
Таким чином, результати дослідження роблять значний внесок у розвиток методів виявлення країв, пропонуючи якісний еталонний набір даних та ефективну архітектуру нейронної мережі. Це, у свою чергу, сприяє підвищенню точності та продуктивності алгоритмів у різних сценаріях застосування, що є важливим кроком у вдосконаленні технологій комп'ютерного зору. Подальші дослідження можуть зосередитися на вдосконаленні архітектури нейронних мереж, зокрема шляхом інтеграції багаторівневих характеристик зображень, використання самонавчальних механізмів та глибшого аналізу семантичного розмежування країв, контурів і меж.

1.3. Значення структури даних для графічної інформації.

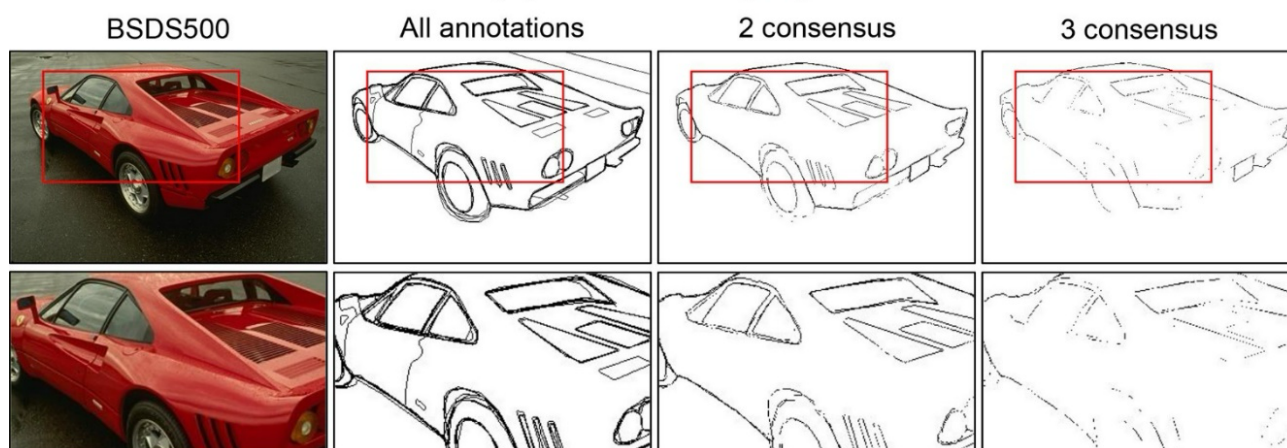
На сьогодні глибоке навчання є основним підходом до задачі виявлення країв. Відомо, що якість навчального набору даних відіграє вирішальну роль у цій парадигмі [18]. З цією метою сучасні дослідження значною мірою спираються на набір даних Berkeley Segmentation Dataset, зокрема BSDS300 [14] та BSDS500 [1]. Первісно цей набір був створений для задач сегментації зображень, однак пізніше

він став загальноприйнятим еталоном для виявлення меж об'єктів. Більшість досліджень використовують BSDS300 для навчання моделей і BSDS500 для їхньої валідації.

Кожне зображення в наборі BSDS містить анотації меж, нанесені щонайменше п'ятьма різними учасниками. В багатьох випадках ці анотації значно відрізняються між собою, що можна побачити на Рис. 2. Відмінності у маркуванні спричиняють складнощі в навчанні моделей, оскільки визначення точного контуру об'єкта стає неоднозначним. Для пом'якшення цієї проблеми використовується функція втрат, що враховує узгоджене значення істинних меж [4, 5, 19]. Під час етапу узгодження пікселі, позначені більше ніж двома учасниками, вважаються істинними краями, тоді як решта відкидається. Це сприяє покращенню збіжності процесу навчання [19]. Однак етап узгодження має й негативні наслідки, оскільки він виключає значну частину реальних країв. Це можна чітко спостерігати на Рис. 2: наприклад, смуги на зебрі повністю зникають навіть за помірного узгодження двох учасників. Подібна ситуація спостерігається і у випадку червоного автомобіля. Аналогічна проблема характерна для багатьох зображень у BSDS, що свідчить про необхідність створення більш точних і детальних наборів даних.



(a) BSDS300 [14]



(b) BSDS500 [1]

Рис. 2. Ретельний аналіз еталонних наборів даних BSDS для двох зразкових зображень та їх відповідних еталонних значень виявляє значні розбіжності між намальованими людьми «краями». Вимога узгодженості між двома або трьома анотаціями призводить до усунення справжніх країв. Ця проблема демонструє, що BSDS є придатним набором даних для оцінки меж, але не країв.

Окрім того, використання набору даних, спочатку призначеного для сегментації, ускладнює задачі навчання алгоритмів глибокого навчання. Для ефективного виявлення країв необхідно мати анотації, які б відповідали реальним межам об'єктів, включаючи тонкі деталі та текстури. Відсутність таких даних може призводити до неточностей у прогнозах моделей і обмежувати їхню здатність узагальнювати на нових зображеннях.

Зважаючи на ці виклики, у дослідженнях останніх років приділяється значна увага розробці нових наборів даних, що орієнтовані спеціально на виявлення країв.

Наприклад, у деяких роботах пропонується використання більш детальних розміток або застосування альтернативних підходів до анотування даних, таких як комбіновані методи глибокого навчання та ручного маркування. Це дозволяє зберегти більше інформації про реальні краї та покращити продуктивність моделей.

Отже, хоча BSDS500 залишається стандартним набором даних у цій галузі, його обмеження вимагають пошуку альтернативних рішень. Удосконалення методів маркування та розширення навчальних наборів можуть сприяти значному покращенню точності та якості виявлення країв у комп'ютерному зорі.

Отже, нейронні мережі, навчені на наборі даних BSDS, фактично стають детекторами меж, а не детекторами країв. Така відмінність має велике значення, особливо в тих випадках, коли саме краї відіграють ключову роль у подальших задачах комп'ютерного зору, наприклад, як у [20]. Помилкова інтерпретація країв як меж може призводити до значних помилок у процесі обробки зображень, що впливає на точність алгоритмів у таких застосуваннях, як підвищення роздільної здатності зображень, реконструкція чи генерація зображень.

Дана дипломна робота фактично покликана розв'язати зазначену проблему, пропонуючи новий еталонний набір даних, орієнтований саме на краї, що дозволяє проводити більш точне оцінювання алгоритмів виявлення країв. Завдяки ретельному визначенню країв у запропонованому наборі даних стає можливим навчання нейронних мереж, здатних ідентифікувати краї з високою точністю без помилкового розпізнавання меж. Це, у свою чергу, сприяє поліпшенню роботи систем комп'ютерного зору, які використовують інформацію про краї як фундаментальний елемент подальшої обробки.

1.4. Існуючі суміжні розробки та підходи.

Існує величезний масив літератури, присвяченої виявленню країв (див. огляди в [10] і [11]). У цьому розділі детально розглядається набір

репрезентативних алгоритмів. Вони можуть бути умовно поділені на чотири основні групи: i) алгоритми, що базуються на низькорівневих ознаках; ii) алгоритми, натхненні біологічними механізмами зору; iii) класичні методи на основі навчання; iv) алгоритми глибокого навчання.

Алгоритми, що базуються на низькорівневих ознаках: ранні алгоритми виявлення країв, такі як оператори Собеля [22], Робертса та Прюїтта [23], ґрунтуються на першій похідній. Вхідне зображення згладжується за допомогою лінійних локальних фільтрів, зазвичай двох орієнтацій (горизонтальної та вертикальної). На завершальному етапі краї виявляються шляхом порогової обробки. Подальший розвиток цих методів передбачав використання другої похідної, де краї визначаються через знаходження нульових перетинів. У середині 1980-х років Кенні [9] запропонував метод виявлення країв, який об'єднував три ключові процеси. Незважаючи на застарілість, ці методи все ще використовуються у сучасних додатках комп'ютерного зору.

Алгоритми, натхненні біологічними механізмами зору: ця група алгоритмів спирається на відомі механізми біологічної зорової системи. Дослідження з 1960-х років на мавпах і котах суттєво розширили наше розуміння первинної зорової кори (V1). Наприклад, було виявлено, що група простих клітин реагує на краї [24]. У [24] автори розробили математичну модель, що імітує людський зір через похідні Гауса. Відтоді було реалізовано безліч алгоритмів, що відтворюють механізм біологічних нейронів при обробці реальних зображень. У [25] було представлено спеціальну функцію ваги лінії для посилення країв, яка поєднує нульові та другорядні функції Ерміта. Пізніше в [13] було запропоновано енергетичні карти Габора для моделювання некласичних рецептивних полів V1, що дозволило виявляти контури. Цей підхід було перевірено на 40 зображеннях, що зробило його одним із перших анованих наборів даних для верифікації алгоритмів виявлення контурів. Пізніші роботи, такі як [26], вдосконалили цей підхід, моделюючи просторове сприяння та оточуюче гальмування через локальні групувальні функції та фільтри Габора. У [27] було впроваджено модуляцію подвійних опонентних клітин для вилучення більш складних властивостей країв, що було доповнено інтеграцією другої зорової зони

та зворотними зв'язками [28]. Наступні експерименти показали, що моделювання різних варіантів оточуючої модуляції, характерних для нейронів зорової кори, підвищує ефективність виявлення країв [28]. У [7] було запропоновано поєднання трьох фільтрів Габора на різних масштабах, за якими слідує PCA та класифікатор машинного навчання. Це приводить до наступної групи методів — навчальних.

Класичні методи на основі машинного навчання: завдання виявлення країв у природних сценах спонукало до розробки алгоритмів на основі навчання. Одним із перших підходів у цьому напрямку був Pb (Probabilistic boundary), запропонований у [14], що використовував різні фільтри для витягування градієнтів яскравості, кольору та текстури. Автори навчили логістичний регресійний класифікатор для генерації ймовірного контуру. Подальші роботи, такі як [29], розширили цей підхід, додаючи глобальну інформацію, що привело до створення gPb (globalized Probability of boundary). У [30] було представлено метод, що базується на умовному випадковому полі, яке навчається на основі анотацій країв. Інший підхід у [31] використовує рідкісне кодування градієнтів, де патчі класифікуються за допомогою методу опорних векторів (SVM). Також у [32] розглядається байєсівська модель, яка прогнозує карти країв за допомогою послідовного підходу Монте-Карло. Крім того, у [33] представлено алгоритм виявлення країв на основі випадкового лісу, де кожне дерево навчається незалежно для захоплення локальних особливостей.

Глибокі нейронні мережі: останнє десятиліття відзначилося бурхливим розвитком згорткових нейронних мереж (CNN), які стали стандартом у комп'ютерному зорі. Було показано, що внутрішні представлення мереж класифікації об'єктів значною мірою спираються на краї [34]. Перший CNN-метод виявлення країв було запропоновано у [35], де мережа навчалася на анотованих картах країв. Відтоді з'явилися численні методи, серед яких основою багатьох архітектур є VGG16 [36], що містить 13 згорткових шарів. Ця архітектура зазвичай попередньо навчається на ImageNet для класифікації, а потім донавчається на наборах даних для виявлення контурів. До таких підходів належать HED [19], RCF [4], BDCN [5], CATS [6], Chrnet [37] та інші. У HED [19] використовується багатомасштабна мережа з глибоким контролем, де кожен блок VGG генерує

прогноз країв. У RCF [4] цей підхід розширено шляхом отримання країв з кожного шару в блоках. Chrnet [37] використовує кілька блоків уточнення для об'єднання виходів третього, четвертого та п'ятого блоків VGG, що дозволяє отримати чистіші карти країв.

Хоча кількісні результати зазначених методів перевершують традиційні підходи, прогнозовані карти країв часто залишаються недостатньо чіткими. Для вирішення цієї проблеми було запропоновано Context-Aware Tracing Strategy (CATS) [6], що покращує різкість країв. Останні досягнення трансформерних моделей призвели до появи EDTER (Edge Detection with Transformer) [38], але його головним недоліком є понад 400 млн параметрів, що вимагає значних обчислювальних ресурсів. Хоча більшість моделей для виявлення країв базується на VGG16, деякі використовують ResNet50 [39], як у [4]. Попри це, покращення становить лише близько 1%, тому VGG16 залишається стандартом для подібних задач.

РОЗДІЛ 2 РОЗРОБКА СТРУКТУРИ НЕЙРОМЕРЕЖІ ІЗ ФУНКЦІОНАЛОМ ДЛЯ ОПРАЦЮВАННЯ ЗОБРАЖЕНЬ ТА ВИЯВЛЕННЯ КРАЇВ ТА МЕЖ У НИХ.

2.1. Розробка щільної екстремальної початкова мережа для виявлення країв у зображеннях.

У нашій роботі результатом є нейронна мережа DexiNed. Її структурний вигляд ретельно спроектований для полегшення наскрізного навчання без необхідності ініціалізації ваг із попередньо навчених моделей виявлення об'єктів, що є типовим для більшості детекторів країв на основі глибокого навчання. Для практичної реалізації ми використали результати дослідження [33]. Далі ми зробили критично важливе спостереження: ознаки країв, які обчислюються на дрібних шарах, мають тенденцію до зникнення на глибших шарах. Цей ключовий висновок стимулював нас на створення інноваційно іншої архітектури, яка нагадує Xception [34], але містить дві паралельні пропускні з'єднання, що ефективно зберігають і використовують усю інформацію про краї, обчислену на різних рівнях мережі.

DexiNed найкраще розглядати як структурну інтеграцію двох взаємопов'язаних підмереж, як показано на Рис. 3: Dense Extreme Inception Network (Dexi) і Upsampling Network (USNet). Dexi виконує функцію основного компонента, що обробляє вхідне RGB-зображення, систематично розбиваючи його на різні обчислювальні блоки. Кожен із цих блоків витягує основні карти ознак, які потім передаються в USNet, забезпечуючи точне захоплення та уточнення високоякісних деталей країв.

Інноваційність розробленої нейромережі DexiNed полягає в її здатності підвищувати ефективність виявлення країв, систематично зберігаючи пов'язані з краями ознаки протягом усього проходження мережі. На відміну від традиційних підходів, що покладаються на попередньо навчені моделі, DexiNed автономно навчається розпізнавати контури, що суттєво покращує

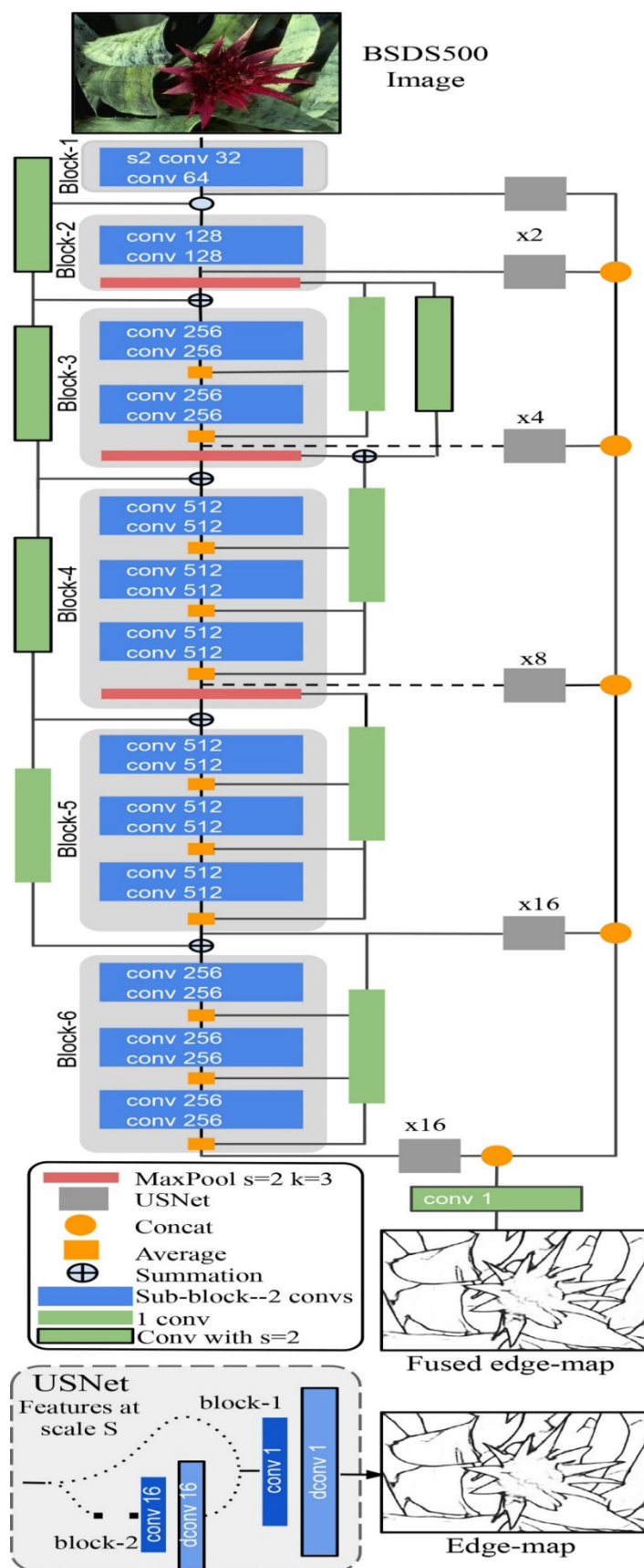


Рис. 2.1. Блок-схема запропонованої архітектури (DexiNed), яка складається з двох конструктивних блоків: Dense Extreme Inception Network і мережі підвищення роздільної здатності (USNet).

його адаптивність до різних контекстів зображень. Унікальні паралельні пропуски з'єднання відіграють ключову роль у збереженні цілісності країв, дозволяючи мережі підтримувати високу точність і стійкість навіть у складних візуальних сценаріях.

Крім того, USNet розроблений для покращення та уточнення виявлених країв за допомогою вдосконаленого механізму підвищення роздільної здатності. Він обробляє ієрархічні карти ознак, отримані від Dexi, поступово реконструюючи дрібні контури таким чином, щоб забезпечити чіткість і узгодженість. Завдяки цій структурованій двокомпонентній системі DexiNed досягає вищої продуктивності виявлення країв, долаючи обмеження традиційних методів глибокого навчання.

Загалом, комбінована ефективність Dexi і USNet утверджує DexiNed як новаторську модель у сфері виявлення країв на основі глибокого навчання, що поєднує точність і обчислювальну ефективність.

2.2. Розробка щільної екстремальної початкової мережі для виявлення країв у зображеннях.

Розроблена нами архітектура нейромережі Dexi складається з шести блоків, які діють подібно до енкодера. Кожен блок являє собою набір менших підблоків із групою згорткових шарів. Пропуски з'єднання об'єднують блоки, а також їхні підблоки (зображені світло-сірими та блакитними прямокутниками на Рис. 3). Карти ознак, створені на кожному з блоків, передаються до окремої USNet для створення проміжних карт контурів. Ці проміжні карти контурів конкатенуються, утворюючи стек навчених фільтрів. У самому кінці мережі ці ознаки об'єднуються для формування єдиної карти контурів.

В свою чергу, кожен підблок (блакитні прямокутники на Рис. 3) містить два згорткових шари (кількість ядер вказана праворуч від блакитних прямокутників). Усі ядра мають розмір 3×3 . У першому блоці згортки виконуються зі зсувом 2, що

позначається як s_2 у його назві. Кожен згортковий шар супроводжується нормалізацією пакетів (Batch Normalization) та виправленою лінійною активацією (ReLU). Починаючи з третього блоку (світло-сірі прямокутники), останній згортковий шар останнього підблоку не містить функції ReLU. Прямокутники червоного кольору є операторами максимального тренду з ядром розміру 3×3 і кроком 2. Велика кількість згорткових операцій у глибині обробки спричиняє зникнення важливих ознак контурів [43], [45]. Щоб уникнути цієї проблеми, ми впроваджуємо паралельні пропускні з'єднання, натхненні He et al. [39] та Zhang [46]. Починаючи з третього блоку (Block-3), вихід кожного підблоку усереднюється з іншим пропускним з'єднанням, яке називається другим пропускним з'єднанням — SSC (зелені прямокутники праворуч на Рис. 3). Після операції макс-пулінгу ці SSC усереднюють вихід пов'язаних підблоків перед підсумовуванням із першим пропускним з'єднанням — FSC (зелені прямокутники ліворуч на Рис. 3). Паралельно з цим вихід шарів макс-пулінгу безпосередньо передається наступним підблокам. Наприклад, підблоки блоку-3 отримують вхідні дані з першого шару макс-пулінгу, а підблоки блоку-4 отримують вхідні дані, підсумовані з першого та другого шарів макс-пулінгу.

2.3. Покрокова розробка нейромережі USNet для збільшення роздільної здатності.

Мережа збільшення роздільної здатності (USNet) є спеціалізованою структурою, що складається з двох основних блоків, які працюють у тісній взаємодії. Ці блоки відіграють ключову роль у процесі покращення роздільної здатності ознак, дозволяючи системі точно відновлювати та масштабувати отримані карти ознак.

Кожен із блоків містить у собі спеціально підібрану послідовність операцій, яка включає згортковий і дезгортковий шари. Завдяки цим операціям мережа

здатна покроково покращувати якість ознак, поступово збільшуючи їхню роздільну здатність на кожному етапі обчислень. Така організація дозволяє USNet ефективно працювати з отриманими ознаками, зберігаючи важливу інформацію та мінімізуючи втрати деталей.

Другий блок, позначений як Блок-2, активується в тому випадку, коли виникає необхідність у масштабуванні вхідних карт ознак, отриманих із мережі Dexi. Він повторює обчислення в кілька ітерацій, доки карта ознак не досягне потрібного рівня збільшення, а саме масштабу, що вдвічі перевищує розмір GT (еталонної карти ознак). Тільки після цього оброблена карта ознак передається в перший блок — Блок-1, який виконує подальші перетворення. Блок-1 є завершальним етапом у процесі обробки ознак. Він починає свою роботу із застосування згорткового шару з ядром розміром 1×1 , що дозволяє оптимізувати отримані карти ознак, зменшуючи кількість непотрібної інформації. Після цього застосовується функція активації ReLU, яка підсилює суттєві особливості країв і робить модель більш чутливою до деталей. Наступним кроком є виконання транспонованої згортки (дезгортки) з ядром розміром $s \times s$, де s — рівень масштабу поточної карти ознак. Завдяки цьому етапу модель поступово відновлює карту ознак до потрібного рівня деталізації.

Завершальний шар у Блоці-1 виконує фінальне збільшення роздільної здатності, приводячи карту ознак до розміру, що відповідає GT. Важливо зазначити, що останній згортковий шар у цьому процесі не використовує функцію активації, оскільки її застосування на цьому етапі може спотворити отриману карту ознак.

Для покращення продуктивності та адаптації до різних умов ми розглянули три підходи до реалізації процесу збільшення роздільної здатності:

- Білінійна інтерполяція — підхід, що дозволяє згладжувати краї та покращувати загальну чіткість об'єктів у зображенні.
- Субпіксельна згортка — метод, який забезпечує точніший рівень деталізації, розподіляючи ознаки більш рівномірно.

- Транспонована згортка — підхід, що дозволяє найбільш ефективно зберігати точність і геометрію країв у складних сценах.

Вибір конкретної стратегії відіграє вирішальну роль у процесі створення точних і реалістичних карт країв. Особливо важливо, що ці методи дозволяють моделі формувати тонкі та виразні краї, що значно покращує якість візуалізації. Це один із ключових аспектів, який було враховано при розробці DexiNed. Детальний аналіз ефективності цих стратегій буде представлений далі у дипломній роботі.

2.4. Визначення функції втрат при роботі з розробленою нейронною мережею та її розрахунок.

Модель DexiNed можна формально описати як регресійну функцію відображення $\partial(\cdot)$. Іншими словами, вихідні передбачені карти країв $\hat{Y}$ можна виразити як $\hat{Y} = \partial(x, y)$, де:

$x \in R^{m \times d \times c}$ — вхідне зображення, яке має розмірність $m \times d \times c$, де c дорівнює 3 через використання RGB-каналів під час навчання моделі;

$y \in \{0, 1\}^{m \times d \times 1}$ — відповідна мапа країв (ground truth), яка містить бінарні значення (0 або 1).

Вихідний датасет $\hat{Y}$ містить набір передбачених мап країв, представлених у вигляді $\hat{Y} = [\hat{y}_1, \hat{y}_2, \dots, \hat{y}_N]$, де кожне значення величини $\hat{y}_n$ відповідає передбаченій мапі країв на певному рівні обчислень у DexiNed. Загальна кількість передбачень позначена як N (відповідає сірим прямокутникам на Рис. 3). Остання мапа країв, $\hat{y}_N$, отримується після фінального згорткового шару. Вона є результатом об'єднання всіх карт $\hat{y}_n$, де початкові ваги встановлюються за формулою $1 / (N - 1)$. Важливо зазначити, що розмір $\hat{y}_N$ повністю збігається з розміром y , тобто відповідає розмірам еталонної мапи країв.

В переважній кількості робіт, що пов'язані із розпізнаванням меж у зображеннях функція втрат базувалася на підході HED (Holistically-Nested Edge Detection), що використовує зважену крос-ентропію. Проте в цьому дослідженні ми провели оцінку ефективності функцій втрат із трьох підходів: HED, RCF (Richer Convolutional Features) і BDCN (Bi-Directional Cascade Network). Після порівняльного аналізу було обрано функцію втрат BDCN із невеликою модифікацією. Формалізований вираз функції втрат подано нижче:

$$\begin{aligned} w(y > 0) &= \frac{1}{y^- + y^+}; \quad w(y \leq 0) = \frac{1.1}{y^- + y^+}; \\ l_n &= \frac{1}{m \cdot d} \sum_n \left(-w \left\{ y \log \theta(\hat{y}_n) + (1 - y) \log [1 - \theta(\hat{y}_n)] \right\} \right), \end{aligned} \quad (2.1)$$

де

w — ваговий коефіцієнт функції втрат;

y^- і y^+ — кількість негативних та позитивних зразків країв у GT відповідно,

θ — сигмоїдна функція активації,

λ_n — набір гіперпараметрів, які використовуються для балансування кількості позитивних і негативних зразків.

В результаті остаточно функція втрат для всієї нейромережі визначається як:

$$L = \sum_{n=1}^N \lambda_n l_n. \quad (2.2)$$

Ця функція втрат відіграє ключову роль у навчанні моделі, дозволяючи їй ефективніше обробляти та прогнозувати тонкі краї об'єктів у зображеннях. Основна ідея такої побудови втрат полягає в тому, щоб збалансувати внесок кожного рівня DexiNed у процес навчання. Гіперпараметри λ_n забезпечують гнучкість у налаштуванні ваги кожного рівня передбачень, що покращує узагальнюючу здатність моделі.

Завдяки використанню адаптивної вагової функції та багаторівневого підходу до прогнозування країв, DexiNed демонструє високу точність виявлення навіть найтонших контурів у складних зображеннях. Більш детальний опис вибору значень параметрів наведено у відповідному розділі, присвяченому реалізації та експериментальним результатам.

РОЗДІЛ 3 НАВЧАННЯ РОЗРОБЛЕНОЇ НЕЙРОННОЇ МЕРЕЖІ. ВХІДНІ ТРЕНУВАЛЬНІ ДАТАСЕТИ.

3.1. Навчання нейронної мережі з використанням набору даних VIPED.

Другим важливим внеском цієї дипломної роботи є оновлення початкового набору даних, відомого як Barcelona Images for Perceptual Edge Detection (VIPED), який був вперше представлений у більш загальному дослідженні [21]. В оновленій версії, представлений у цьому дослідженні, були здійснені значні покращення, спрямовані на підвищення якості зображень та їх розмітки. Усі зображення, які входили до початкового набору даних, були оброблені повторно, з використанням удосконалених методів анотації, що забезпечило вищу точність розмічених контурів. Набір VIPED містить 250 реальних зображень міських середовищ, кожне з яких має роздільну здатність 1280×720 пікселів.

Для створення точних GT-карт контурів було використано платформу Labelbox6, що дозволило здійснити детальну ручну анотацію контурів у зображеннях. Як зазначено в [8], існує три основні методи створення GT для наборів даних із виявлення контурів: перший метод передбачає створення синтетичних зображень на основі попередньо згенерованих контурів, другий полягає в ручному маркуванні контурів безпосередньо на реальних зображеннях, а третій базується на поєднанні карт контурів, отриманих за допомогою різних алгоритмів виявлення країв, із подальшим аналізом їх узгодженості.

Оновлення набору VIPED у цьому дослідженні базується на використанні комбінації другого та третього методів, що дозволило отримати максимально реалістичні та точні GT-карти. На ілюстраціях у Рис. 3.1 представлено приклади зображень із набору VIPED разом із відповідними GT-контурними картами. У Рис. 3.1 (а) показано оригінальне зображення, у Рис. 3.1(б) рожевими лініями позначено відкриті контури (полігональні криві), а синіми лініями – закриті контури (замкнені об'єкти). Рис. 3.1(с) демонструє остаточну GT-карту контурів, отриману після перевірки й коригування.

Щоб отримати високоточні GT-карти контурів, ми розробили багаторівневий процес маркування, що включав кілька етапів: початкове ручне позначення контурів, експертну перевірку та коригування, автоматичне порівняння з результатами алгоритмів виявлення контурів і остаточну верифікацію отриманих GT-карт. Кожне зображення аналізувалося з використанням сучасних алгоритмів, таких як HED (Holistically-Nested Edge Detection), що дозволило оцінити відповідність вручну позначених контурів реальним межах об'єктів.



Рис. 3.1. Зразок зображення з набору даних VIPED.

Це оновлення значно покращує точність та відповідність GT-контурів реальним об'єктам, роблячи набір VIPED одним із найдетальніших і найякісніших у цій галузі. Він забезпечує більш точну інформацію для навчання моделей машинного навчання, які використовуються у завданнях виявлення країв, і відкриває можливості для подальших досліджень у комп'ютерному зорі.



Рис. 3.2. Різниця між попередньою існуючою анотацією GT у VIPED і запропонованою версією, BIPED-GTv2.

Хід тестування роботи є такий:

- Експерти з комп'ютерного зору позначають всі краї на кожному зображенні, роблячи лише одну анотацію на зображення.
- Адміністратор (також експерт з комп'ютерного зору) перевіряє результат попереднього етапу.
- Отримані GT використовуються для навчання моделі HED [19] і перевірки коректності передбачених контурів.
- Враховуючи результати HED, адміністратор повторно переглядає весь набір даних, виправляючи помилки та додаючи нові анотації. Ця версія також була представлена у праці [21].
- Завдяки оновленій версії LabelBox покращений інструмент масштабування дозволяє побачити додаткові деталі, які не враховувалися в попередній версії через недостатню деталізацію зображень VIPED. Це змусило нас додати більше анотацій майже на всіх зображеннях.
- Різницю між початковими анотаціями (VIPEDv1) і поточною версією (VIPEDv2) можна побачити на Рис. 3.2; нові інструменти анотації дозволяють малювати дуже тонкі краї на наданих зображеннях.



Рис. 3.3. Приклад виділення мапи ребер у розробленому методі.

3.2. Тестування із еталонними контрольними наборами даних.

Детальне та всебічне якісне оцінювання запропонованої моделі було ретельно проведене на широкому наборі датасетів, які широко використовуються в науковій

літературі для вирішення задач виявлення контурів, границь та країв. Зокрема, ми протестували наш підхід на таких датасетах: MDBD [7], BSDS500 [1], BSDS300 [14], NYUD [40], PASCAL-CONTEXT [41] (далі просто PASCAL), CID [13] та нашому новому датасеті BIPED. Кожен із цих наборів даних був обраний для того, щоб забезпечити всебічне оцінювання здатності нашої моделі працювати з різними типами структур країв і контурів, які зустрічаються в природних зображеннях.

Серед наведених датасетів MDBD є особливо важливим для нашого дослідження, оскільки його анотації точно відповідають реальним краям, що робить його найбільш релевантним тестовим набором для оцінки ефективності нашої моделі у виявленні країв. Решта датасетів, хоча й не були спеціально створені для цієї задачі, є цінними орієнтирами для оцінювання продуктивності моделі в задачах виявлення контурів та границь. Використовуючи широкий спектр датасетів для тестування, ми прагнемо забезпечити справедливе та об'єктивне порівняння нашого методу із сучасними передовими підходами, демонструючи його гнучкість і надійність у різних контекстах.

Крім того, використання декількох датасетів дозволяє нам проаналізувати, наскільки добре наша модель узагальнює свої результати на різних типах зображень та стилях анотації. Оскільки різні датасети мають різні критерії розмітки, оцінювання на широкому наборі тестових даних забезпечує перевірку того, що наш метод не просто адаптується до одного конкретного набору, а залишається ефективним і універсальним у широкому діапазоні сценаріїв виявлення країв та контурів. Отримані результати підкреслюють сильні сторони запропонованої моделі та підтверджують її конкурентоспроможність у порівнянні з існуючими методами в цій галузі.

Мультифакторний датасет для виявлення границь (MDBD) був спеціально зібраний для підтримки психофізичних досліджень, спрямованих на розуміння того, як люди сприймають границі об'єктів у складних візуальних сценах. Дизайн цього датасету був ретельно продуманий з урахуванням різних візуальних факторів, таких як яскравість, колірні варіації, рухові характеристики та

бінокулярна диспаратність, які відіграють ключову роль у процесі сприйняття границь [7]. Завдяки інтеграції цих різноманітних особливостей, MDBD забезпечує складний і насичений тестовий набір для оцінки ефективності алгоритмів виявлення контурів і границь.

Датасет MDBD складається з коротких бінокулярних відеопослідовностей, що охоплюють як природні зовнішні, так і внутрішні середовища. Кожна сцена містить 10 послідовних кадрів, що дозволяє дослідникам аналізувати процес виявлення контурів у динамічних умовах. Із цих відеопослідовностей було ретельно відібрано 100 високоякісних зображень з роздільною здатністю 1280×720 пікселів. Ці зображення є основним набором даних для вирішення задач виявлення контурів та границь. Щоб забезпечити високу надійність розмітки, кожне зображення було анотоване декількома учасниками: шість разів для задачі виявлення країв та п'ять разів для задачі визначення границь. Такий ретельний процес розмітки створює якісний еталон, за яким можуть оцінюватися комп'ютерні моделі.

Для оцінювання продуктивності запропонованої архітектури DexiNed ми використали анотації країв, надані в датасеті MDBD. Як правило, у дослідженнях, що включають MDBD, 80% доступних зображень використовуються для навчання моделей, тоді як решта 20% виділяється для тестування та оцінки продуктивності алгоритмів на основі машинного навчання. Дотримуючись цього стандартного підходу, ми випадковим чином відібрали 20% зображень із датасету для проведення оцінювання продуктивності DexiNed. Крім того, щоб забезпечити чесне та неупереджене порівняння з іншими передовими моделями, усі алгоритми, які використовувалися в нашому кількісному аналізі, були повторно навчені на тих самих 80% навчального набору. Такий підхід гарантує узгодженість у порівнянні моделей і підвищує достовірність отриманих результатів.

Дотримуючись цієї методології, ми створюємо комплексну та об'єктивну систему оцінки DexiNed, демонструючи його здатність ефективно виконувати виявлення контурів у реальних, складних сценаріях. Використання MDBD як

тестового датасету підкреслює надійність і гнучкість запропонованої моделі у вирішенні завдань, пов'язаних із виявленням границь.

CID: База даних контурних зображень (Contour Image Database) містить 40 градаційно-сірих зображень із відповідними контурами, що слугують еталоном [13]. Розмір зображень у цьому наборі становить 512×512 пікселів. Основним обмеженням цього датасету є невелика кількість анотованих зображень. Архітектуру DexiNed було оцінено на всьому наборі даних, аналогічно до попередніх досліджень у літературі (наприклад, [26], [28]). Цей набір даних є складним для підходів на основі глибокого навчання через його градієнтно-сіру природу та відсутні анотації в деяких наданих зображеннях. BSDS500: Набір даних сегментації Берклі (Berkeley Segmentation Data Set, BSDS) було вперше опубліковано у 2001 році [14] і він складався з 300 зображень, розділених на 200 для навчання та 100 для валідації (цей набір відомий як BSDS300). Остання версія [1] додала ще 200 нових зображень для тестування, розширивши набір до BSDS500. Кожне зображення у BSDS було анотоване щонайменше 5 людьми. Зображення у цьому датасеті мають розмір 481×321 пікселів. Цей набір даних призначений переважно для сегментації зображень та виявлення меж, проте, як буде показано в наступних розділах, деякі зображення не мають належної анотації для задач детекції контурів. Зазвичай, для оцінки продуктивності моделі глибокого навчання на BSDS500, нові 200 зображень використовують для тестування, тоді як BSDS300 застосовують для навчання моделі. Оскільки DexiNed тренувався лише на наборі даних VIPED, у якісній оцінці, що представлена у цьому розділі, результати поділено на дві частини: BSDS300 та BSDS500. Зображення для якісного порівняння взято з тестової частини BSDS500 та валідаційної частини BSDS300.

NYUD: Набір даних Нью-Йоркського університету (New York University Dataset) містить 1449 зображень у форматі RGBD (колір+глибина) з 464 різних внутрішніх приміщень. Він призначений для задач сегментації та поділений на три підмножини: навчальну, валідаційну та тестову. Тестова множина містить 654 зображення, тоді як решта використовується для навчання та валідації. У цій роботі

для оцінки запропонованої моделі використано лише тестову частину, оскільки DexiNed було навчено лише на VIPED. Хоча більшість зображень у NYUD мають повну анотацію для сегментації, деякі містять неточні анотації. Відсутність деяких контурів у зображеннях негативно впливає на якість методів глибокого навчання для виявлення контурів.

PASCAL: набір даних PASCAL [41] є популярним для задач сегментації, оскільки містить велику кількість категорій об'єктів. Наразі більшість основних методів глибокого навчання для виявлення контурів використовують PASCAL для навчання та тестування (наприклад, [19], [4]), оскільки його еталонні анотації охоплюють сцени, які відрізняються від зображень у BSDS. Цей набір даних містить 11 530 анотованих зображень, однак лише близько 5 % випадково вибраних зображень (505) було використано для тестування DexiNed. Незважаючи на те, що в PASCAL міститься більше різноманітних анотованих даних, більшість зображень позначені лише для кількох об'єктів, навіть якщо сцена містить багато інших візуальних особливостей.

3.3. Аналіз отриманих результатів, верифікація та тестування.

В даному розділі дипломної роботи спочатку опишемо метрики, що використовуються для оцінювання; далі надаються деталі щодо реалізації та налаштувань запропонованого підходу. Нарешті, представлено великий набір експериментальних результатів разом із порівнянням із найсучаснішими підходами. Знову ж таки, більшість датасетів, які не призначені для виявлення контурів, використовуються лише для якісної оцінки, а для задачі виявлення контурів підходи VIPED і MDBD застосовуються для кількісних оцінок і порівнянь із DexiNed.

Алгоритми виявлення контурів можуть бути оцінені двома основними підходами, кожен з яких має свої переваги та застосування. Перший підхід

передбачає непряме оцінювання через вплив алгоритмів на продуктивність інших завдань комп'ютерного зору, таких як сегментація зображень, розпізнавання об'єктів або відстеження руху. У такому випадку ефективність алгоритму виявлення контурів вимірюється через його внесок у покращення результатів цих завдань. Наприклад, більш точні контури можуть сприяти кращому розпізнаванню меж між об'єктами, що, у свою чергу, покращує якість сегментації. Такий підхід часто використовується у випадках, коли виявлення контурів є лише проміжним етапом у складнішій системі комп'ютерного зору.

Другий підхід є безпосереднім і ґрунтується на порівнянні отриманих алгоритмом контурів із контурами, нанесеними вручну людьми. Це більш об'єктивний і широко застосовуваний метод оцінки, оскільки дозволяє безпосередньо порівняти, наскільки добре алгоритм відтворює межі об'єктів, які бачить людина. Цей підхід особливо важливий у випадках, коли контур є ключовою характеристикою зображення, наприклад, у медичній діагностиці, де точне виявлення меж пухлини має критичне значення. У цій роботі ми зупинилися саме на другому підході, оскільки він є стандартним для оцінки алгоритмів у загальноприйнятих тестових наборах даних

.Для оцінки якості роботи алгоритму використовуються такі метрики: Optimal Dataset Scale (ODS) – це метрика, яка визначає загальну продуктивність алгоритму на всьому датасеті, застосовуючи єдиний глобальний поріг для всіх зображень. Вона дозволяє оцінити, наскільки добре алгоритм може узагальнювати свої результати на різних зображеннях.

Optimal Image Scale (OIS) – ця метрика працює дещо інакше, оскільки передбачає підбір оптимального порогу для кожного окремого зображення. Таким чином, вона оцінює найкращу продуктивність алгоритму на кожному зображенні окремо, що дозволяє зрозуміти його максимальні можливості.

Average Precision (AP) – середня точність, що відображає загальну ефективність алгоритму у виявленні контурів, враховуючи співвідношення правильних і хибних позитивних спрацьовувань.

Окрім цих метрик, широко використовується F-міра (F-measure), яка є збалансованою оцінкою точності та повноти алгоритму. Вона обчислюється за наступною формулою:

$$F = \frac{2 \times \text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}}. \quad (3.1)$$

У формулі (3.1) Precision (точність) визначає частку правильно знайдених контурів серед усіх передбачених алгоритмом, а Recall (повнота) показує, яку частку всіх істинних контурів вдалося виявити. Високе значення F-міри вказує на те, що алгоритм добре збалансований між точністю та повнотою, що є важливим критерієм для якісного виявлення контурів.

Завдяки використанню цих метрик можна зробити детальний аналіз ефективності алгоритму, порівняти його з іншими сучасними підходами та зрозуміти, в яких аспектах його можна покращити.

3.4. Оцінка та основні результати щодо верифікації. Налаштування архітектури розробленого програмного забезпечення

Запропонована архітектура DexiNed була навчена з нуля, не використовуючи попередньо навчені ваги, що є її унікальною особливістю. Це суттєво відрізняє її від більшості сучасних нейронних мереж, які зазвичай використовують попередньо навчені ваги з датасету ImageNet. Використання навчання з нуля дозволяє моделі адаптуватися саме до завдання виявлення контурів, без впливу особливостей інших задач, на яких були натреновані стандартні попередньо навчені моделі. У середньому DexiNed досягає збіжності після 9 епох (у той час як у попередньому дослідженні [21] було потрібно 15 епох). Навчання здійснюється з розміром пакета 8, використовуючи оптимізатор Adam та навчальну точність (learning rate) 10^{-4} , яка зменшується у 10 і 15 епохах на коефіцієнт 0.1. Додатково використовується зменшення ваги (weight decay) із значенням 10^{-8} , що допомагає уникнути

перенавчання моделі. Процес навчання триває приблизно 1 день на графічному процесорі TITAN X. Вхідні зображення мають розмір 352×352 пікселів, а для датасету MDBD використовується розмір 480×480 пікселів. Щодо початкових ваг для шару злиття (fusion layer), вони ініціалізуються за формулою $1/(N-1)$ (див. Розділ 3.4 для детальної інформації про параметр N). Після проведення пошуку оптимальних гіперпараметрів для DexiNed, найкраща продуктивність була досягнута за таких умов:

Використання ядер 3×3 , 1×1 та $s \times s$ у різних згорткових (conv) і зворотно-згорткових (deconv) шарах. Ініціалізація ваг у Dexi здійснювалася за допомогою методу Xavier [47], тоді як для останніх згорткових і зворотно-згорткових шарів мережі USNet використовувалася нормальна розподілена ініціалізація. Коефіцієнти λ у функції втрат були налаштовані як $[0.7, 0.7, 1.1, 1.1, 0.3, 0.3, 1.3]$, що дозволяє моделі адаптуватися до різних рівнів контурної деталізації. Для навчання і валідації DexiNed було випадково відібрано 200 зображень із датасету VIPED, а решта 50 зображень використовувалися для тестування. Щоб збільшити розмір вибірки навчальних даних, був застосований процес сегментації даних, який включав наступні кроки:

- оскільки зображення у VIPED мають високу роздільну здатність, кожне зображення розділялося навпіл уздовж його ширини;
- аналогічно до підходу, що застосовувався в HED, отримані зображення оберталися на 15 різних кутів і обрізалися за внутрішнім прямокутником, орієнтованим уздовж осей;
- виконувалося горизонтальне віддзеркалення зображень, що допомагає навчити модель розпізнавати контури незалежно від напрямку об'єкта;
- було застосовано дві корекції гамми (з коефіцієнтами 0.3030 і 0.6060), що дозволяє адаптувати модель до варіацій яскравості та контрастності зображень.

У результаті кожне вихідне зображення було розширене до 288 варіантів, що значно збільшило обсяг тренувального набору і дозволило покращити узагальнюючу здатність моделі.

Щодо вибору найкращого методу підвищення роздільної здатності (upsampling), нами було проведено оцінку трьох підходів, які могли б забезпечити оптимальні результати. У ході цього дослідження було виявлено, що використання транспонованої згортки (transpose convolution) з ядрами, що навчаються демонструє найкращу ефективність у порівнянні з іншими методами. Саме тому цей підхід було обрано для поточної роботи як основний метод підвищення роздільної здатності.

З іншого боку, важливим аспектом є стратегія об'єднання вихідних результатів, отриманих на різних рівнях моделі. Розглядалися різні варіанти, серед яких – вихідні результати рівнів Output1, Output2, Output3, Output4, Output5, Output6, а також результати, отримані після застосування фінальних операцій моделі DexiNed-f та DexiNed-a. Архітектура DexiNed (див. рис. 3.1 вище в дипломній роботі) була емпірично оцінена відповідно до підходів, використаних у [21], що дозволило визначити оптимальні варіанти об'єднання результатів. Згідно з отриманими експериментальними даними, найкращі результати були досягнуті при використанні стратегій DexiNed-f та DexiNed-a.

DexiNed-f відповідає результату, отриманому після застосування процесу об'єднання на фінальному етапі роботи архітектури DexiNed (див. рис. 3.1). У свою чергу, DexiNed-a представляє собою середньозважене значення усіх прогнозованих edge-map (мапи контурів), включаючи також результат DexiNed-f. Дослідження, наведене у [21], демонструє, що обидві стратегії злиття забезпечують подібні кількісні показники точності. Саме тому в подальшому в цій роботі будуть представлені результати для обох стратегій. Проте, у деяких випадках, через обмеження простору, буде представлено лише результати DexiNed-f. У таких ситуаціях вони будуть позначені узагальнено як DexiNed.

Далі подано кількісне дослідження критичних частин та налаштувань архітектури DexiNed. Оскільки наша модель містить два типи пропускних з'єднань (skip connections), які позначені зеленими прямокутниками з лівого та правого боків на Рис. 3.1, було проведено аналіз ефективності різних варіантів їх використання.

Метою цього дослідження було визначення оптимальної кількості пропускових з'єднань для покращення точності моделі.

Результати цього дослідження наведені на Рис. 3.4 (ліворуч). Було розглянуто три різні конфігурації: DexiNed0C – конфігурація, яка взагалі не використовує пропускові з'єднання. У цьому випадку модель працює без можливості передачі проміжної інформації між шарами, що може вплинути на якість виявлення контурів. DexiNed1C – конфігурація, яка використовує лише одне пропускове з'єднання, розташоване з лівого боку схеми (відповідно до Рис. 3.1). Це дозволяє частково передавати інформацію між рівнями архітектури, але ефективність цього підходу обмежена. DexiNed2C – конфігурація, що містить два пропускові з'єднання: одне на лівому боці і друге на правому боці архітектури. Завдяки цьому забезпечується більш ефективний обмін інформацією між різними рівнями моделі, що сприяє покращенню продуктивності.

Аналіз результатів показав, що використання двох пропускових з'єднань (DexiNed2C) суттєво покращує ефективність запропонованої архітектури у порівнянні з іншими конфігураціями. Це пояснюється тим, що додаткові зв'язки між шарами дозволяють зберігати більше інформації про структуру зображення та краще передавати її на наступні рівні нейромережі. Отже, у подальшому для отримання найкращих результатів рекомендується використовувати саме конфігурацію DexiNed2C, яка забезпечує найвищу точність у задачі виявлення контурів.

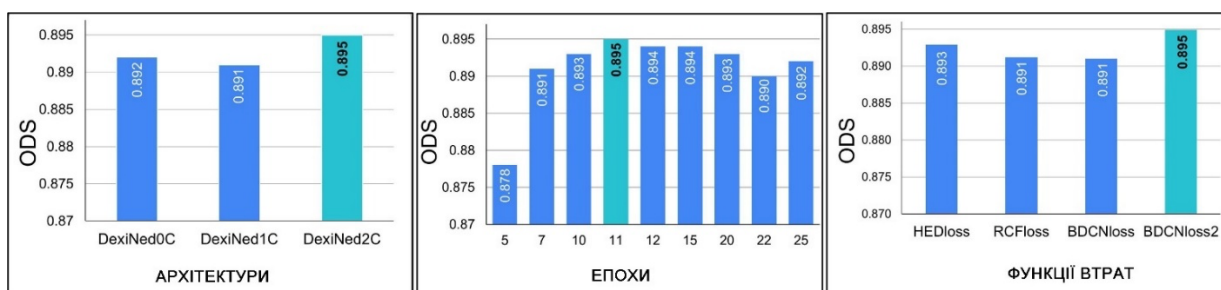


Рис. 3.4. (зліва) Оцінка архітектури DexiNed з пропусками з'єднань (DexiNed1C: 1 з'єднання; DexiNed2C: 2 з'єднання) та без них (DexiNed0C). (в середині) Зміна продуктивності DexiNed під час навчання. (праворуч) Оцінка з різними функціями втрат.

Запропонована архітектура також була навчена з використанням різних функцій втрат для оцінки її продуктивності. Було розглянуто функції втрат, що використовуються в наступних підходах: HED [19], RCF [4], BDCN [5], а також BDCNloss2 – це незначно модифікована функція втрат, заснована на BDCN [5]. Основна мета цього дослідження полягала в тому, щоб визначити, яка саме функція втрат забезпечує найкращу якість виявлення контурів при використанні архітектури DexiNed.

Як показано на Рис. 3.4 (праворуч), модифікована функція втрат BDCNloss2 демонструє покращені результати порівняно з іншими функціями втрат. Перевага цієї функції складає приблизно 1% у точності, що є вагомим показником у задачах детектування контурів. Отримані результати підтверджують ефективність запропонованої модифікації функції втрат у контексті навчання мережі DexiNed.

Крім того, для визначення найкращої продуктивності моделі було проведено оцінку її передбачень на різних етапах навчання – від 1 до 25 епох. Як можна побачити на рис. 6 (посередині), найкращі результати були досягнуті після 11 епох навчання. Це означає, що подальше збільшення кількості епох не дає суттєвого покращення продуктивності, а лише підвищує витрати обчислювальних ресурсів.

З огляду на отримані результати, у подальших розділах цього дослідження порівняння продуктивності DexiNed на різних наборах даних для виявлення контурів будуть проводитися з використанням такої конфігурації:

- архітектура DexiNed2C (із двома пропускними з'єднаннями).
- використання 11 епох навчання;
- використання модифікованої функції втрат BDCNloss2;

Для зручності в тексті ця конфігурація буде просто позначатися як DexiNed.

3.5. Виконання кількісного порівняння методів

Далі ми розглянемо порівняння DexiNed із сучасними підходами до виявлення контурів на спеціалізованих наборах даних. Окрім того, було проведено дослідження загальної здатності до узагальнення на двох наборах даних (MDBD і VIPED). Іншими словами, у цьому дослідженні аналізуються результати перенесення навченої моделі з одного набору даних на інший, що є важливим фактором для оцінки реальної ефективності нейромережі.

Як зазначалося в Розділі 1, набір даних BSDS [1] не розглядається, оскільки він не може бути використаний для узагальнення результатів у домені виявлення контурів. Це пояснюється його особливостями, що обмежують можливість створення універсальних моделей для детектування країв.

Для порівняння DexiNed із найкращими сучасними підходами було обрано передові методи виявлення контурів, меж і границь [11]. Зокрема, розглянуто такі методи:

- RCF [4] (Richer Convolutional Features)
- BDCN [5] (Bi-Directional Cascade Network)
- CATS [6] (Contour-Aware Transformer for Segmentation)

Щоб забезпечити чесне порівняння, усі ці методи були навчені на однакових наборах даних, призначених для детектування контурів, а саме MDBD і VIPED. При цьому важливо зазначити, що:

- використано однакові методи аугментації зображень для всіх моделей, що дозволяє мінімізувати вплив відмінностей у попередній обробці даних.
- у випадку MDBD, всі моделі було навчено на однаковому навчальному наборі, а не на випадково відібраних зображеннях, як це часто робиться в інших дослідженнях. Це гарантує максимальну об'єктивність отриманих результатів і підвищує точність оцінки продуктивності моделей.

Далі будуть представлені результати порівняння DexiNed із вказаними методами, а також детальний аналіз узагальнюючої здатності моделі на різних наборах даних.

Результат, що подано на Табл. 1 та Рис. 3.5 демонструють різні оцінки для DexiNed (DexiNed-f та DexiNed-a) та підходів з останніх досягнень, згаданих вище. Представлені результати для обох сценаріїв — навчання та тестування на одному й тому ж наборі даних, а також крос-оцінки (тобто навчання на одному наборі даних та оцінювання на іншому). Можна побачити, що DexiNed досягає найкращих показників (за всіма трьома метриками ODS, OIS та AP) при навчанні та тестуванні на одному й тому ж наборі даних. Крім того, DexiNed також показує найкращий результат при навчанні на VIPED, але оцінюванні на наборі даних MDBD. З іншого боку, щодо здатності до узагальнення набору даних, враховуючи метрику оцінювання ODS, ми знаємо, що найкращі показники досягаються при навчанні та оцінюванні на одному й тому ж наборі даних.

Табл. 1. Кількісні результати: продуктивність моделі DexiNed та набору даних VIPED у порівнянні з останніми розробленими моделями на основі DL та набором даних на основі виявлення меж, MDBD [7].

Методи	Навчання на основі	Тестовано на основі	ODS	OIS	AP	Навчання на основі	ODS	OIS	AP
RCF [4] (2019)	MDBD [7]	MDBD [7]	0.879	0.888	0.926	VIPED	0.814	0.828	0.871
BDCN [5] (2020)			0.887	0.891	0.793		0.854	0.863	0.768
CATS [6] (2022)			0.891	0.899	0.809		0.813	0.831	0.791
DexiNed-f (наш)			0.891	0.896	0.930		0.787	0.807	0.844
DexiNed-a (наш)			0.894	0.902	0.951		0.789	0.813	0.875
RCF [4] (2019)	VIPED	VIPED	0.849	0.861	0.906	MDBD [7]	0.839	0.854	0.865
BDCN [5] (2020)			0.890	0.899	0.934		0.855	0.864	0.692
CATS [6] (2022)			0.887	0.892	0.817		0.837	0.840	0.496
DexiNed-f (наш)			0.895	0.900	0.927		0.863	0.871	0.867
DexiNed-a (наш)			0.893	0.897	0.940		0.862	0.874	0.919

Отже, ми пропонуємо проаналізувати втрату продуктивності при оцінюванні на іншому наборі даних, відмінному від того, що використовувався для навчання.

Ця втрата продуктивності менша, якщо різні підходи навчаються на VIPED, в середньому лише 3,74% зниження продуктивності спостерігається при оцінюванні на MDBD; навпаки, ця втрата продуктивності досягає в середньому 7,14%, якщо навчання проводиться на MDBD та оцінювання на VIPED. Наше емпіричне спостереження показує, що погана продуктивність моделі, навченої на MDBD, значною мірою залежить від наборів даних, що використовуються для навчання — тобто обсягу даних, їхньої мінливості та точності анотацій.

У випадку набору даних MDBD основна проблема полягає у відсутності консенсусу наданих анотацій, як детально описано в розділі 1; та ж проблема виникає під час анотації BSDS — див. ілюстрацію на Рис. 3.5. Цей консенсус необхідний для впевненості в анотаціях. Для отримання точної інформації набори даних з кількома анотаціями вимагають 2 або 3 консенсусних анотацій, що призводить до видалення деяких країв. Ці пропущені краї впливають на кількісні оцінки навчених моделей — оскільки деякі справжні краї не містяться в еталонних даних, але, навпаки, вони могли бути оцінені мережами. Ця погана продуктивність спостерігається не лише у DexiNed, навченого на MDBD, але й в інших архітектурах.

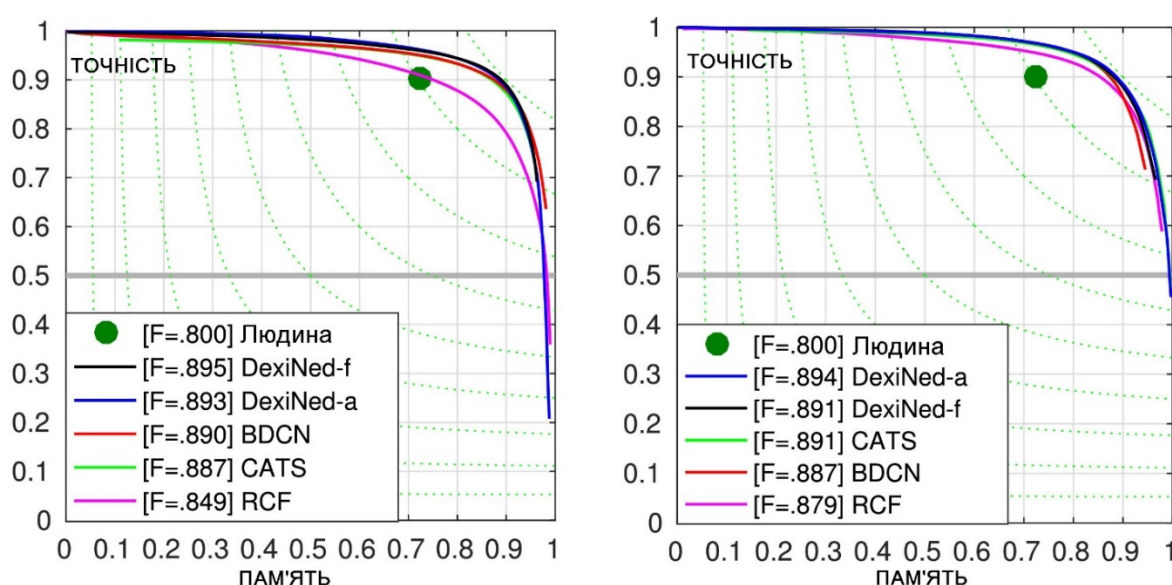


Рис. 3.5 (ліворуч) Крива «точність-повнота» тестового набору VIPED. (праворуч)

Крива "точність-повнота" MDBD [48].

3.6. Якісне порівняння методів

Далі розглянемо основні ілюстрації як якісні порівняння мап країв, передбачених усіма моделями, розглянутими під час кількісного порівняння в попередньому розділі. Зауважимо, що всі ці моделі були навчені на VIPED; Рис. 3.6 показує мапи країв оцінюваних архітектур (навчених на VIPED) при використанні на наборах даних, детально описаних у в попередньому розділі дипломної роботи. Якісні результати, подібні до представлених на Рис. 3.6, але навчені на MDBD [7], представлені як додатковий матеріал. Нижче наведено деякі коментарі та висновки з аналізу отриманих результатів:

- VIPED: Для цього набору даних представлено дві ілюстрації, оскільки цей набір даних використовується для навчання вибраних моделей, розглянутих у порівняннях. Ці дві ілюстрації відповідають найкращим і найгіршим результатам ODS від DexiNed. Як показано на Рис. 3.6, карти країв, передбачені BDCN, CATS та DexiNed, є чистими та точними представленнями. З точки зору сприйняття, ми можемо сказати, що карти країв, передбачені BDCN та DexiNed, містять більше правильно виявлених країв. Навпаки, карти країв, передбачені CATS, є чистішими, але містять менше передбачень. Як висновок, при пошуку моделі, яка передбачає більше країв, але також менше артефактів, DexiNed надає нам ці вимоги без компромісу.
- Інші набори даних: MDBD, CID, BSDS300, BSDS500, NYUD та PASCAL. Як показано на Рис. 3.6, результати на цих наборах даних також подібні до передбачень з тестового набору на VIPED. Модель CATS показує чистіші карти країв, але з меншою кількістю країв, ніж DexiNed. Більш помітну різницю в сприйнятті цього твердження можна спостерігати в передбаченнях MDBD, CID, NYUD та PASCAL. Дивлячись на останній рядок на Рис. 3.6, який відповідає набору даних PASCAL, DexiNed виявив кілька країв на підлозі сцени, але жодна з інших моделей не здатна їх виявити.

Таким чином ключові зміни та адаптації, що отримані нами такі:

- Термінологія: Замінено англійські терміни на їхні українські відповідники (наприклад, "edge-maps" на "карти мап", "dataset" на "набір даних").
- Стиль: Зроблено текст більш зрозумілим та читабельним для української аудиторії.
- Послідовність: Уточнено послідовність деяких речень для кращої логічної зв'язності. Додано пояснення: Додано деякі пояснення для кращого розуміння технічних термінів.

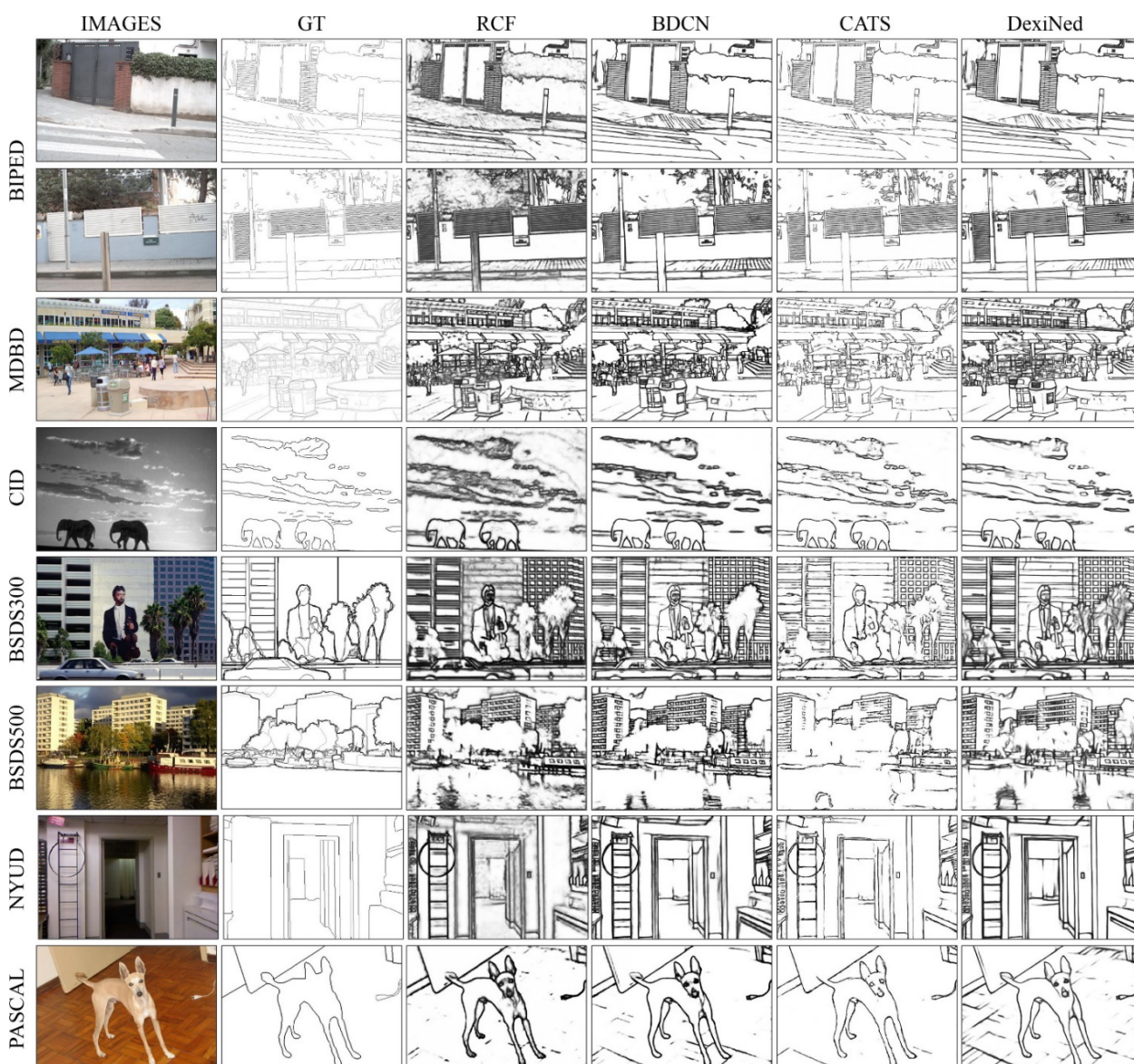


Рис. 3.6. Результати кількох сучасних архітектур, навчених на запропонованому наборі VIPED даних. Слід звернути увагу, що істинність цих наборів даних, за винятком VIPED та MDBD, відповідає завданням визначення меж та сегментації.

Підсумовуючи, друга версія VIPED забезпечує стійкість до узагальнення для моделей, розглянутих для порівняння. Загалом, на всіх наборах даних, представлених на Рис. 3.6, завдяки унікальним характеристикам VIPED, в зображеннях передбачається найбільше візуально сприйнятих країв. Щодо архітектури DexiNed, передбачення, надані цією моделлю, є чистішими, ніж у її аналогів, і без втрати справжніх країв, процедура навчання є простішою, ніж у будь-яких інших моделей на основі глибокого навчання, розглянутих для порівняння. DexiNed не потребує попередньо навчених ваг, він може збігатися за менший час, ніж інші моделі, але при цьому досягає результатів на рівні сучасних досягнень.

Розроблена нами нейромережа DNet — це архітектура, спеціально розроблена для виявлення країв. Головна мета цієї дипломної роботи — продемонструвати, що при використанні належного набору даних немає потреби в попередньо навчених вагах або налаштуванні гіперпараметрів на рівні шарів мережі, як це роблять більшість моделей сучасного рівня (SOTA).

Результати, що подані в Табл. 1 показують, що DNet досягає найкращих результатів, навчаючись з нуля за короткий період часу. З меншим налаштуванням гіперпараметрів DNet досягає пікової продуктивності на 11-й епосі, тоді як попередні дослідження повідомляють про пікову продуктивність лише після 20-ї епохи. Хоча наша пропозиція має більш ніж удвічі більше параметрів, ніж BDCN [5] або CATS [6], її обчислювальний час менший як під час навчання, так і під час тестування. Це можливо завдяки значному скороченню пошуку гіперпараметрів та специфічному дизайну архітектури DNet. DNet містить 35 мільйонів параметрів, і, незважаючи на свою глибоку архітектуру, краї не зникають на картах ознак глибоких шарів (поширена проблема для глибоких детекторів країв), а завдяки внеску USNet вихід DNet містить велику кількість справжніх країв з чистішими контурами порівняно з моделями SOTA.

Ще одним внеском цієї виконаної роботи є злегка модифікована функція втрат з BDCN [5] та HED [19]. Ця модифікація допомагає адаптуватися та отримати кращу вартість процесу навчання при використанні розміру міні-пакету,

враховуючи середню втрату замість суми. Остання версія дозволяє досягти найкращих результатів на MDBD та VIPED. Як зазначено в попередніх розділах, ми не можемо кількісно порівнювати з BSDS або іншими подібними наборами даних, що використовуються спільноту, оскільки ці набори даних не мають анотацій на рівні країв, оскільки вони призначені для виявлення контурів або меж. Як згадувалося вище, DNet демонструє хорошу продуктивність при навчанні на VIPED, але оцінюванні на MDBD, що певним чином показує хороше узагальнення запропонованої моделі. На жаль, немає інших наборів даних для виявлення країв, щоб оцінити цю функцію узагальнення. Вхідний колірний простір DNet (як і більшості інших глибоких мереж) — RGB, канали якого сильно корелюють у природних зображеннях (середня кореляція для каналів R, G та B в ImageNet становить близько 0,85 [49]). Через цю високу внутрішньоосову кореляцію краї, виявлені в кожному каналі, значною мірою перекриваються. На відміну від цього, наша зорова система декорелює вхідний простір у колірну опозицію, що полегшує виявлення інформації (наприклад, країв) як у хроматичних, так і в яскравих каналах. Ми розглядаємо для себе це як майбутню роботу для дослідження впливу вхідного колірного простору на продуктивність DNet або інших моделей виявлення країв.

РОЗДІЛ 4 БЕЗПЕКА ЖИТТЄДІЯЛЬНОСТІ, ОСНОВИ ОХОРОНИ ПРАЦІ

4.1. Принципи, способи та засоби захисту населення

Державна політика України у сфері захисту населення і територій базується на Конституції України, відповідних законах, актах Президента України і урядових рішеннях. «Людина, її життя і здоров'я, честь і гідність, недоторканість і безпека визначається в Україні найвищою соціальною цінністю», - декларує стаття 3 Конституції України. Закон України «Про захист населення і територій від надзвичайних ситуацій техногенного та природного характеру» визначає організаційні та правові основи захисту громадян України, захисту об'єктів виробничого і соціального призначення, довкілля від надзвичайних ситуацій техногенного та природного характеру [43]. Готовність цивільного захисту до виконання покладених на неї завдань, у кінцевому рахунку, визначається її здатністю до підготовки і проведенню комплексу заходів, спрямованих на захист населення та території всієї нашої країни.

Для зменшення втрат та шкоди проводиться комплекс заходів:

- Оповіщення та інформування – завчасне створення систем оповіщення для оперативного повідомлення населення.
- Спостереження та контроль – моніторинг довкілля, продуктів, води на наявність шкідливих речовин.
- Укриття – створення фонду захисних споруд для тимчасового укриття людей.
- Інженерний захист – врахування небезпек під час забудови, будівництво безпечних об'єктів, захисні інженерні споруди, плани ліквідації наслідків надзвичайних ситуацій.
- Евакуація – планується заздалегідь у районах з об'єктами підвищеної небезпеки та у воєнний час.
- Медичний захист – профілактика, лікування, контроль епідемій, створення медичних формувань, забезпечення засобами захисту.

- Біозахист – виявлення біологічних загроз, ізоляція, знезараження, профілактика, карантинні заходи.
- Радіаційний і хімічний захист – контроль рівнів забруднення, дозиметрія, спеціальна обробка, засоби індивідуального захисту.

Окремо виділяють, протирадіаційний і протихімічний захист – це комплекс заходів, спрямованих на запобігання або послаблення дії іонізуючих випромінювань, отруйних речовин і сильнодіючих отруйних речовин. Радіаційний і хімічний захист включає такі основні заходи:

- виявлення і оцінка радіаційної й хімічної обстановки;
- організація і проведення дозиметричного і хімічного контролю;
- розробка і введення в дію типових режимів радіаційного захисту;
- забезпечення населення і невоєнізованих формувань засобами індивідуального захисту;
- ліквідація наслідків радіаційного і хімічного зараження.

На сьогодні розроблені і рекомендуються 8 типових режимів захисту для різних категорій населення:

- 1-3-й режими – для непрацюючого населення;
- 4-7-й режими – для робітників і службовців;
- 8-й режим – для невоєнізованих формувань:

Дотримання режимів забезпечить протирадіаційний захист населення, основна мета якого не допустити опромінення людей понад встановлені допустимі дози.

Радіоактивне зараження місцевості – підступний і небезпечний фактор ураження. Воно розповсюджується навіть на ті райони, котрі не піддавалися дії надзвичайної ситуації, тому що утворена при аварії на АЕС радіоактивна хмара може переміщуватись на великі відстані.

Основними способами захисту населення при хімічному зараженні є:

- оповіщення про небезпеку хімічного зараження;
- укриття в захисних спорудах;

- використання засобів індивідуального захисту (протигазів, засобів захисту шкіри);
- застосування антидотів і ППП-8;
- дотримання режимів (правил) поведінки на зараженій території;
- евакуація людей із зон зараження;
- санітарна обробка людей, одягу, території, техніки, споруд, майна.

Укриття населення в захисних спорудах – це комплекс заходів із завчасним будівництвом захисних споруд, а також пристосуванням наявних приміщень для захисту населення та підтримання їх у готовності до використання. Укриттю в захисних спорудах підлягає все населення України. Фонд захисних споруд створюється шляхом:

обстеження та обліку підземних та надземних будівель і споруд, що відповідають вимогам захисту населення;

- дообладнання з урахуванням реальної обстановки підвалів, погребів та інших заглиблених приміщень;
- обстеження і взяття на облік підземних і надземних будівель і споруд, гірничих виробок і природних порожнин, що відповідають вимогам захисту, у разі необхідності – переобладнання цих приміщень;
- будівництво заглиблених споруд пристосованих для захисту;
- масового будівництва в період загрози надзвичайних ситуацій найпростіших сховищ та укриттів;
- будівництва окремих сховищ та протирадіаційних укриттів.

Потреби в захисних спорудах визначають, виходячи з необхідності укриття всіх працюючих за місцем роботи і проживання, та всього непрацюючого населення за місцем проживання. Укриття населення в захисних спорудах є надійним способом захисту від уражаючих факторів ядерної, хімічної, бактеріологічної, звичайної зброї, при аваріях і деяких стихійних лихах (ураганах, снігових заметах). Захисні споруди є засобами колективного захисту від уражаючих факторів. До них відносяться: сховища, протирадіаційні укриття,

найпростіші укриття.

4.2. Заходи, що запобігають впливу на людину агресивних та токсичних речовин, які використовуються в технологічному процесі.

У сучасному технологічному виробництві широко використовуються речовини, що мають агресивні або токсичні властивості. Їх наявність у виробничому середовищі створює потенційну небезпеку для здоров'я та життя працівників. Тому надзвичайно важливим завданням є розроблення і впровадження комплексу заходів, спрямованих на запобігання шкідливому впливу таких речовин. Ці заходи повинні бути інтегрованими у систему управління охороною праці підприємства та включати технічні, організаційні, адміністративні, інформаційні та медико-профілактичні складові. Поряд з цим важливою є роль відповідного нормативного регулювання, оскільки вимоги щодо безпеки праці, роботи з хімічними речовинами, оцінювання ризиків та впровадження системи контролю повинні відповідати сучасним стандартам, зокрема ДСТУ. Виконання вимог стандартів забезпечує єдиний підхід до організації безпеки та охорони здоров'я працівників.

Ідентифікація речовин, які є потенційно небезпечними, є першочерговим етапом у формуванні системи безпеки. Відповідно до ДСТУ ISO 14698-1:2008 [44], підприємства мають здійснювати контроль повітряного середовища на робочих місцях з метою оцінки рівнів забруднення та встановлення джерел викиду. Аналіз хімічних речовин у технологічному процесі дозволяє визначити рівень ризику для працівника та необхідні заходи контролю. До типових небезпечних речовин, що використовуються в технологічному процесі, належать: кислоти (сірчана, азотна, хлоридна), які можуть викликати опіки, подразнення дихальних шляхів; луги (гідроксид натрію, калію), що становлять загрозу при потраплянні на шкіру або в очі; леткі органічні сполуки (бензол, толуол), що мають канцерогенні та мутагенні

властивості; пари важких металів (ртуть, свинець), які накопичуються в організмі та мають хронічний токсичний ефект; гази (аміак, хлор, сірководень), які викликають гостре отруєння при перевищенні допустимих концентрацій. На основі переліку використаних речовин формується реєстр небезпек, визначаються гігієнічні нормативи та розробляються карти ризиків.

Технічний захист базується на мінімізації або усуненні контактів працівника з токсичними речовинами за рахунок застосування безпечних технологій. До основних принципів належать: герметизація технологічного обладнання, що виключає витік шкідливих речовин (ДСТУ EN ISO 14123-1:2016) [45]; заміна небезпечної речовини на менш токсичну, за можливості; використання замкнених технологічних циклів, де робоче середовище повністю ізольоване від навколишнього простору; впровадження автоматизованих систем подачі, змішування та транспортування реагентів; застосування систем аварійного відключення, автоматичних клапанів та сповіщувачів витіку. Крім того, важливим аспектом технічного захисту є регулярна діагностика обладнання та перевірка його герметичності, проведення профілактичного технічного обслуговування та модернізація систем з урахуванням новітніх досягнень у галузі безпеки праці.

Колективні засоби захисту передбачають створення інженерних систем, що ізолюють або нейтралізують небезпечні речовини до їхнього контакту з людиною. До таких засобів належать: вентиляційні системи з локальними витяжками, які забезпечують виведення шкідливих парів безпосередньо з місця утворення (ДБН В.2.5-67:2013) [46]; витяжні шафи, які встановлюються в лабораторіях та на робочих місцях для роботи з леткими речовинами; фільтраційні установки з багатоетапними системами очищення повітря (вуглецеві, хімічні, НЕРА-фільтри); фізичні бар'єри, перегородки та герметичні камери для відокремлення працівника від джерела небезпеки; системи рециркуляції повітря з виведенням очищених газів за межі робочої зони. Ефективність колективного захисту залежить від правильного проектування, монтажу та обслуговування обладнання. Необхідно забезпечувати регулярну перевірку систем, зміну фільтрів, очищення повітропроводів.

Засоби індивідуального захисту застосовуються тоді, коли інженерні заходи недостатні або не повністю виключають контакт із шкідливою речовиною. До них відносяться: респіратори, фільтруючі півмаски та протигази (ДСТУ EN 149:2017) [47], вибір яких залежить від типу речовини (аерозоль, газ, пара); захисні окуляри, екрани, маски для захисту органів зору від агресивних парів (ДСТУ EN 166:2017) [48]; захисний одяг, комбінезони з кислотостійких і лугостійких матеріалів (ДСТУ EN 14605:2017) [49]; рукавиці з гуми, латексу, неопрена, нітрилу залежно від речовини; захисне взуття з покриттям, стійким до дії агресивних речовин. Працівники мають проходити навчання щодо правильного використання, зберігання та догляду засоби індивідуального захисту. Усі засоби повинні регулярно перевірятись, а несправні — замінюватись.

Організаційні заходи включають розробку нормативної документації, проведення навчань та моніторинг дотримання вимог безпеки. Основні складові: розроблення та впровадження інструкцій з охорони праці для кожної операції з токсичними речовинами; регулярне проведення вступного, первинного, повторного та позапланового інструктажу; організація роботи відповідальних осіб за безпеку хімічних речовин на підприємстві; ведення документації щодо обліку, зберігання та використання небезпечних речовин; забезпечення наочності: плакати, сигнальні таблички, евакуаційні схеми (ДСТУ EN ISO 7010:2019) [50]; проведення періодичних аудитів стану охорони праці. Ці заходи сприяють формуванню культури безпеки, що є запорукою зниження ризиків у повсякденній роботі.

Контроль рівня концентрації токсичних речовин у повітрі виконується згідно з методиками, визначеними в ДСТУ ISO 16000 [51]. Він включає: автоматизовані системи виявлення перевищення ГДК; індикатори кольорової зміни для виявлення наявності газів; портативні прилади для експрес-аналізу повітря; лабораторне дослідження повітря, води, поверхонь обладнання; аналіз та документування результатів з подальшим коригуванням заходів безпеки. Контрольні заходи мають виконуватись систематично, згідно з графіками та у випадках надзвичайних ситуацій (виток, аварія).

Для підтримання здоров'я працівників необхідно: проводити попередній медичний огляд при прийомі на роботу; здійснювати періодичні огляди згідно з чинними нормативами; створити медичний пункт з доступом до первинної медичної допомоги; мати засоби надання допомоги при хімічних ураженнях (антидоти, нейтралізатори); забезпечити працівників інформацією про симптоми отруєння, методи самодопомоги, профілактику хімічних уражень. У разі виникнення нещасного випадку необхідно вжити заходів евакуації, ізоляції джерела небезпеки та викликати медичну допомогу.

Основні нормативні документи, що регулюють роботу з токсичними речовинами: ДСТУ ISO 45001:2019 — Системи управління охороною здоров'я та безпекою праці [52]; ДСТУ EN 60068-2-1:2022 — Оцінювання впливу хімічних речовин, що вдихаються [53]; ДСТУ Б В.2.5-67:2013 — Вентиляція та кондиціонування повітря [46]; ДСТУ EN ISO 14123-1:2018 — Безпечність машин [54]. Захист від шкідливих речовин; ДСТУ EN 149:2017 — Респіратори фільтрувальні [47]; ДСТУ EN 943-1:2021— Захисний одяг для газів [55]; ДСТУ EN ISO 7010:2019— Знаки безпеки [50]. Також враховуються інші нормативно-правові акти, включаючи Накази МОЗ України, санітарні правила, технічні регламенти.

Комплексна система запобігання впливу токсичних речовин у технологічному процесі повинна включати технічні, організаційні, медичні та нормативні компоненти. Виконання вимог чинних ДСТУ дозволяє ефективно виявляти та нейтралізувати загрози, що виникають при використанні небезпечних речовин. Рекомендується: регулярно оновлювати та переглядати програми безпеки; проводити аналіз ризиків при впровадженні нових технологій; впроваджувати сучасні системи автоматичного контролю; забезпечувати ефективну комунікацію між службами безпеки, керівництвом та працівниками; інвестувати в навчання персоналу та модернізацію ЗІЗ. Тільки системний підхід дозволить створити умови праці, що забезпечують здоров'я, безпеку та ефективність роботи у технологічному середовищі, де присутні агресивні або токсичні речовини.

ВИСНОВКИ

У даній дипломній роботі бакалавра представлено ефективну модель виявлення країв, яка демонструє високий рівень узагальнення на нових зображеннях і сценах. Основна мета дослідження – розробка надійного підходу до виявлення країв, який не потребує попередньої підготовки на великих загальних наборах даних, таких як ImageNet.

Для досягнення цієї мети ми виконали два ключові кроки. По-перше, створено еталонний набір даних спеціально для задачі виявлення країв – Barcelona Images for Perceptual Edge Detection (BIPEDv2). Цей набір містить ретельно анотовані зображення, які забезпечують точне навчання моделей глибокого навчання. Він був створений із урахуванням специфіки реальних сценаріїв та містить зображення високої якості з чітко визначеними краями.

По-друге, нами розроблена нейромережева архітектура DeNet, яка містить паралельні пропуски з'єднань (skip-connections). Вона дозволяє покращити розпізнавання структурних особливостей зображень, забезпечуючи точне відтворення контурів об'єктів. На відміну від багатьох сучасних моделей, DeNet не вимагає попереднього навчання на великих загальних наборах даних, таких як ImageNet. Це означає, що модель може бути навчена з нуля на вузькоспеціалізованих наборах даних без втрати продуктивності.

Щоб оцінити здатність DeNet до узагальнення, ми провели серію експериментів, навчаючи модель лише на BIPEDv2 і перевіряючи її ефективність на інших наборах даних. Найбільш значущим результатом стало те, що навіть без додаткового навчання на нових даних, модель показала високу точність на MDBD, що свідчить про її ефективність у широкому спектрі застосувань.

Окремо слід відзначити, що хоча DeNet містить удвічі більше параметрів, ніж популярні моделі на основі VGG16, вона зберігає подібну продуктивність у кадрах за секунду (FPS). Крім того, наша архітектура не втрачає мапі ознак (feature maps) на глибоких шарах нейромережі, що дозволяє отримувати більш детальні та точні

контури об'єктів на зображенні.

Загалом, результати дослідження підтверджують можливість створення потужної моделі виявлення країв, яка може бути навчена з нуля у повністю автоматизованому (end-to-end) режимі. У перспективі ці результати відкривають нові можливості для розробки компактніших нейронних мереж, які забезпечать високу якість виявлення країв при меншій обчислювальній складності. Подальші дослідження будуть зосереджені на оптимізації гіперпараметрів, скороченні обсягу обчислень та підвищенні ефективності моделі для використання у реальних застосуваннях, зокрема в комп'ютерному зорі, автономних системах та обробці зображень у медицині.

СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ

1. Goodfellow, I., Bengio, Y., & Courville, A. (2016). Deep Learning. MIT Press.
2. Bishop, C. M. (2006). Pattern Recognition and Machine Learning. Springer.
3. LeCun, Y., & Hinton, G. (2015). Deep Learning. *Nature*, 521(7553), 436–444.
4. Schmidhuber, J. (2015). Deep Learning in Neural Networks: An Overview. *Neural Networks*, 61, 85–117.
5. Chollet, F. (2017). Deep Learning with Python. Manning.
6. Nielsen, M. (2015). Neural Networks and Deep Learning. Determination Press.
7. Aggarwal, C. C. (2018). Neural Networks and Deep Learning: A Textbook. Springer.
8. Li, Y., & Zhang, L. (2018). Deep Learning for Image Processing Applications. Springer.
9. Krizhevsky, A., Sutskever, I., & Hinton, G. E. (2012). ImageNet Classification with Deep Convolutional Neural Networks. *Advances in Neural Information Processing Systems (NeurIPS)*.
10. He, K., Zhang, X., Ren, S., & Sun, J. (2016). Deep Residual Learning for Image Recognition. *IEEE CVPR*.
11. Simonyan, K., & Zisserman, A. (2015). Very Deep Convolutional Networks for Large-Scale Image Recognition. *ICLR*.
12. Szegedy, C., et al. (2015). Going Deeper with Convolutions. *IEEE CVPR*.
13. Redmon, J., & Farhadi, A. (2018). YOLOv3: An Incremental Improvement. *arXiv:1804.02767*.
14. Dosovitskiy, A., et al. (2020). An Image is Worth 16x16 Words: Transformers for Image Recognition at Scale. *ICLR*.
15. Vaswani, A., et al. (2017). Attention Is All You Need. *NeurIPS*.

16. Ronneberger, O., Fischer, P., & Brox, T. (2015). U-Net: Convolutional Networks for Biomedical Image Segmentation. MICCAI.
17. Long, J., Shelhamer, E., & Darrell, T. (2015). Fully Convolutional Networks for Semantic Segmentation. IEEE CVPR.
18. Girshick, R. (2015). Fast R-CNN. IEEE ICCV.
19. Ren, S., He, K., Girshick, R., & Sun, J. (2015). Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks.
20. S.Lin, T. Y., et al. (2017). Feature Pyramid Networks for Object Detection. IEEE CVPR.
21. Zhou, Z., Siddiquee, M. M. R., Tajbakhsh, N., & Liang, J. (2018). UNet++: A Nested U-Net Architecture for Medical Image Segmentation. arXiv:1807.10165.
22. Kingma, D. P., & Welling, M. (2014). Auto-Encoding Variational Bayes. ICLR.
23. Goodfellow, I., et al. (2014). Generative Adversarial Networks.
24. S.Radford, A., Metz, L., & Chintala, S. (2016). Unsupervised Representation Learning with Deep Convolutional Generative Adversarial Networks. ICLR.
25. Karras, T., Laine, S., & Aila, T. (2019). A Style-Based Generator Architecture for Generative Adversarial Networks. IEEE CVPR.
26. Brock, A., Donahue, J., & Simonyan, K. (2019). Large Scale GAN Training for High Fidelity Natural Image Synthesis. ICLR.
27. Paszke, A., et al. (2019). PyTorch: An Imperative Style, High-Performance Deep Learning Library. NeurIPS.
28. Abadi, M., et al. (2016). TensorFlow: Large-Scale Machine Learning on Heterogeneous Systems. arXiv:1603.04467.
29. Chollet, F. (2015). Keras: Deep Learning Library for Theano and TensorFlow. GitHub.
30. Esteva, A., et al. (2017). Dermatologist-Level Classification of Skin Cancer with Deep Neural Networks. *Nature*, 542, 115–118.
31. Lundervold, A. S., & Lundervold, A. (2019). An Overview of Deep Learning in Medical Imaging. *Frontiers in Medicine*.

32. Litjens, G., et al. (2017). A Survey on Deep Learning in Medical Image Analysis. *Medical Image Analysis*, 42, 60–88.
33. Karpathy, A. (2019). A Recipe for Training Neural Networks. Blog.
34. Zeiler, M. D., & Fergus, R. (2014). Visualizing and Understanding Convolutional Networks. *ECCV*.
35. He, K., et al. (2017). Mask R-CNN. *IEEE ICCV*.
36. Sabour, S., Frosst, N., & Hinton, G. (2017). Dynamic Routing Between Capsules. *NeurIPS*.
37. Lin, T. Y., et al. (2014). Microsoft COCO: Common Objects in Context. *ECCV*.
38. Bengio, Y. (2009). Learning Deep Architectures for AI. *Foundations and Trends in Machine Learning*, 2(1), 1–127.
39. Jaderberg, M., et al. (2015). Spatial Transformer Networks. *NeurIPS*. Xie, S., et al. (2017). Aggregated Residual Transformations for Deep Neural Networks. *IEEE CVPR*.
40. Han, S., et al. (2015). Deep Compression: Compressing Deep Neural Networks with Pruning, Trained Quantization and Huffman Coding. *ICLR*.
41. Courbariaux, M., et al. (2016). Binarized Neural Networks: Training Deep Neural Networks with Weights and Activations Constrained to +1 or -1. *NeurIPS*.
42. Дистанційний курс «Кваліфікаційні роботи бакалаврів» сайту дистанційного навчання ТНТУ [Електронний ресурс]. – Режим доступу: URL: <https://dl.tntu.edu.ua/bounce.php?course=5329>.
43. Стищенко Т.Є., Пронюк Г.В., Сердюк Н.М., Хондак І.І. «Безпека життєдіяльності»: навч. посібник / Т.Є Стищенко, Г.В. Пронюк, Н.М. Сердюк, І.І. Хондак. – Харків: ХНУРЕ, 2018. – 336 с.
44. ДСТУ ISO 14698-1:2008 Якість повітря. Чисті приміщення та відповідні контрольовані середовища. Контролювання біозабруднень. Частина 1. Загальні принципи та методи (ISO 14698-1:2003, IDT).
45. ДСТУ EN ISO 14123-1:2016 Безпечність машин. Зниження ризику для здоров'я, спричинюваного небезпечними речовинами, виділюваними машинами.

46. ДБН В.2.5-67:2013 Опалення, вентиляція та кондиціонування
47. ДСТУ EN 149:2017 Засоби індивідуального захисту органів дихання. Фільтрувальні півмаски для захисту від аерозолів. Вимоги, випробування, маркування (EN 149:2001+A1:2009, IDT).
48. ДСТУ EN 166:2017 Засоби індивідуального захисту очей. Технічні умови (EN 166:2001, IDT).
49. ДСТУ EN 14605:2017 Одяг захисний. Захист від рідких хімічних речовин.
50. ДСТУ EN ISO 7010:2019 Графічні символи. Кольори та знаки безпеки. Зареєстровані знаки безпеки.
51. ДСТУ ISO 16000. Повітря в приміщенні.
52. ДСТУ ISO 45001:2019 Системи управління охороною здоров'я та безпекою праці. Вимоги та настанови щодо застосування (ISO 45001:2018, IDT).
53. ДСТУ EN 60068-2-1:2022 Випробування на вплив зовнішніх чинників.
54. ДСТУ EN ISO 14123-1:2018 Безпечність машин. Зниження ризику для здоров'я від небезпечних речовин, що виділяють машини.
55. ДСТУ EN 943-1:2021 Одяг захисний. Захист від небезпечних твердих, рідких та газоподібних хімічних речовин, зокрема рідких та твердих аерозолів.

ДОДАТКИ

Додаток А. Диск з роботою