

**ПОЛТАВСЬКИЙ ДЕРЖАВНИЙ АГРАРНИЙ УНІВЕРСИТЕТ
НАВЧАЛЬНО-НАУКОВИЙ ІНСТИТУТ ЕКОНОМІКИ, УПРАВЛІННЯ,
ПРАВА ТА ІНФОРМАЦІЙНИХ ТЕХНОЛОГІЙ
КАФЕДРА ІНФОРМАЦІЙНИХ СИСТЕМ ТА ТЕХНОЛОГІЙ**

Пояснювальна записка

до кваліфікаційної роботи на здобуття ступеня вищої освіти магістр

на тему: **«Модель виявлення об'єктів на основі нейронної мережі»**

Виконав: здобувач вищої освіти
за освітньо-професійною програмою
Інформаційні управляючі системи та
технології спеціальності
126 Інформаційні системи та технології
ступеня вищої освіти магістр
групи 126ІСТ_мд_21
Веременич Д.Ф.
Керівник: Флегантов Л.О.
Рецензент: Петраш Р.В.

Полтава – 2023 року

ВСТУП

Актуальність даної роботи визначається зростанням обсягу даних у сучасному світі, який постійно збільшується завдяки розвитку технологій. Це створює велику потребу в розвитку ефективних методів обробки та аналізу даних, особливо у сферах, пов'язаних з виявленням об'єктів на зображеннях і відео. Використання нейронних мереж для виявлення об'єктів стало популярним напрямом досліджень, завдяки їхній здатності до автоматичного вивчення ознак та ієрархічної обробки інформації. У галузі комп'ютерного бачення і штучного інтелекту спостерігається активна конкуренція, яка сприяє постійному вдосконаленню технологій виявлення об'єктів та роботи з зображеннями, відкриваючи нові шляхи для інновацій та розвитку.

Зв'язок роботи з науковими програмами, планами, темами. Робота виконана у відповідності до науково-дослідної ініціативної теми «Організаційно-методологічні аспекти впровадження інформаційно-комунікаційних систем і технологій в управлінні діяльністю сучасних організацій та підприємств за умов переходу до цифрової економіки» ДРН 0117U003099.

Метою кваліфікаційної роботи є розробка та дослідження моделі виявлення об'єктів на основі нейронної мережі для поліпшення точності та ефективності виявлення об'єктів на зображеннях та відео.

Завданнями кваліфікаційної роботи є:

– Вибір архітектури нейронної мережі. Визначити найбільш підходящу архітектуру мережі для конкретного завдання виявлення об'єктів та обґрунтувати свій вибір на основі літературного огляду та попередніх досліджень.

– Побудова та тренування моделі. Розробити та навчити нейронну мережу для виявлення об'єктів на підготовленому наборі даних. Враховувати аспекти, такі як гіперпараметри, функції активації, алгоритми оптимізації та шари мережі.

– Оцінка та аналіз результатів. Провести оцінку ефективності моделі, використовуючи різні метрики, такі як точність, відгук, F1-середнє, та порівняти отримані результати з існуючими методами. Аналізувати сильні та слабкі сторони розробленої моделі, обговорювати потенціал для її покращення та можливості застосування в практичних задачах.

Об'єктом дослідження є процес виявлення об'єктів у цифрових зображеннях.

Предметом дослідження є алгоритми та архітектури нейронних мереж, які застосовуються для виявлення об'єктів.

Методологія дослідження у цій роботі базується на використанні передових підходів глибокого навчання. Вона охоплює застосування конволюційних нейронних мереж (CNN), різновидів R-CNN для аналізу регіонів та одноразових методів виявлення об'єктів, як-от YOLO. Доповнюючи це, застосовуються техніки аугментації даних для розширення різноманітності тренувальних даних. Аналіз ефективності моделей проводиться за допомогою метрик якості, таких як mAP та IoU. Робота також включає програмування на Python і використання OpenCV для обробки зображень.

Інформаційна база дослідження охоплює широкий спектр наукових публікацій, документації та інтерактивних ресурсів з глибокого навчання та комп'ютерного зору. Основні джерела включають документацію та наукові статті, пов'язані з TensorFlow, PyTorch, Keras, а також з архітектурами CNN, R-CNN, YOLO і Mask R-CNN.

Елементи наукової новизни виражаються у розробці інноваційних методів для оптимізації розпізнавання облич, з використанням бібліотек OpenCV та Deerpface. Ці методи значно покращують точність і швидкість розпізнавання, використовуючи переваги обох бібліотек для ефективного аналізу характеристик обличчя в різних умовах.

Практична значущість роботи полягає у формуванні пропозицій щодо інтеграції моделі виявлення об'єктів на основі нейронної мережі, яка може

бути використана для подальших досліджень за даною тематикою та при реалізації концепції.

Апробація результатів дослідження полягає у тому, що за результатами проведеного дослідження опубліковано тези: «Модель виявлення об'єктів на основі нейронної мережі», Матер. науково-практичної конференції за підсумками проходження виробничої практики Полтавського державного аграрного університету, 17 вересня 2023 р., м. Полтава, «Технології парсингу даних з вебсайтів до EXCEL», Матер. щорічної студентської наукової конференції Полтавського державного аграрного університету, 10 листопада 2022 р., м. Полтава, Міжнародна науково-практична конференція «Стратегічний менеджмент агропродовольчої сфери в умовах глобалізації економіки: безпека, інновації, лідерство», 28 вересня 2023 рік.

Структура кваліфікаційної роботи складається зі вступу, трьох розділів, що включають теоретичні основи, опис методів дослідження та аналіз отриманих результатів, а також висновків і рекомендацій для подальших досліджень. Основний текст роботи викладений на 89 сторінках, містить 16 рисунків і 3 таблиці. Список використаних джерел налічує 50 найменувань.

РОЗДІЛ 1

НЕЙРОННІ МЕРЕЖІ ТА ЇХ ВИКОРИСТАННЯ У КОМП'ЮТЕРНОМУ ЗОРІ

1.1 Основні поняття та терміни

Комп'ютерний зір – це галузь інформатики, яка займається розробкою алгоритмів та систем, що наділяють машини здатністю «бачити» та інтерпретувати візуальну інформацію зі світу навколо них, аналогічно зору людини. Основні концепції, що входять до цієї галузі, включають обробку зображень, розпізнавання образів, машинне зоріння та глибоке навчання [1].

Нейронна мережа – це алгоритмічна структура, яка використовується в машинному навчанні та штучному інтелекті для моделювання складних шаблонів та передбачення даних. Нейронні мережі складаються з шарів нейронів, кожен з яких з'єднаний зі своїми сусідами зваженими зв'язками [2].

Глибоке навчання – підрозділ машинного навчання, який використовує багатошарові нейронні мережі для аналізу різноманітних форм даних, включаючи зображення та відео. Глибоке навчання дозволяє моделювати та використовувати складні ієрархічні шаблони в даних [4].

Згортова нейронна мережа (CNN або ConvNet) – це тип глибокої нейронної мережі, яка спеціально пристосована для обробки зображень, розпізнавання візуальних шаблонів на різних рівнях абстракції. CNN використовують свертки в процесі навчання, що імітують відповідь візуальних кортексів живих істот на стимули [4].

Аугментація даних – це техніка, яка збільшує розмір та різноманітність набору даних шляхом створення модифікованих версій існуючих даних. В контексті виявлення об'єктів, аугментація може включати зміни в освітленні, повороти зображень, масштабування та інші трансформації, щоб підвищити стійкість моделі до різноманітних умов сприйняття [5].

Фреймворки для обробки зображень, такі як OpenCV, надають інструментарій для обробки та аналізу зображень, необхідних для попередньої обробки даних перед їх використанням у нейронних мережах. OpenCV дозволяє реалізовувати алгоритми обробки зображень, детектування об'єктів, слідування за рухом та багато іншого [6].

Mean Average Precision (mAP) та Intersection over Union (IoU) – метрики для оцінки ефективності моделей виявлення об'єктів. mAP відображає середню точність виявлення об'єктів на різних рівнях порогу IoU оцінює ступінь перекриття між прогнозованим виявленням та істинним положенням об'єкта [6].

Розкриття потужності комп'ютерного зору визначено численними технологічними досягненнями та застосуваннями в різних галузях. Однією з ключових інновацій є використання глибокого навчання та згорткових нейронних мереж (CNN) для вирішення завдань обробки зображень та розпізнавання об'єктів [7].

Глибоке навчання виявилось особливо ефективним у вирішенні складних задач, таких як розпізнавання облич, класифікація об'єктів та аналіз великих об'ємів візуальної інформації. Ця техніка використовується в медицині для діагностики захворювань на основі медичних зображень, в автопромисловості для розпізнавання дорожніх знаків та в безпеці для виявлення аномальної поведінки на відеоспостереженні [5].

Застосування згорткових нейронних мереж виходить за межі традиційного розпізнавання об'єктів. Вони також успішно використовуються для синтезу зображень та генерації нових візуальних контентів, що відкриває нові можливості в галузі творчості та дизайну.

Аугментація даних, яка включає в себе зміни в освітленні, повороти та інші трансформації, стала необхідною складовою у навчанні моделей комп'ютерного зору. Застосування цієї техніки допомагає забезпечити роботу моделей в різних умовах та покращити їхню загальну робочу ефективність [7].

Крім того, фреймворки для обробки зображень, такі як OpenCV, постійно розвиваються та надають розширені можливості для аналізу візуальної інформації. Вони включають в себе інструменти для автоматичного визначення ключових точок, видалення шуму та роботи з тривимірними зображеннями, що розширює сферу їхнього застосування [6].

Подальший розвиток комп'ютерного зору передбачає інтеграцію нових методів та алгоритмів для підвищення точності виявлення об'єктів та розширення можливостей автоматизованої обробки візуальної інформації в реальному часі.

З швидким розвитком комп'ютерного зору постають нові виклики, з якими доводиться стикатися дослідникам та інженерам у цій галузі. Одним із таких викликів є забезпечення стійкості алгоритмів до варіативних умов середовища, таких як різка зміна освітлення чи наявність тіней [8].

Застосування комп'ютерного зору в реальному часі, наприклад, в автономних транспортних засобах чи системах безпеки, вимагає високошвидкісної та ефективної обробки великих об'ємів візуальних даних. Це висуває вимоги до апаратної бази та оптимізації алгоритмів для забезпечення максимальної продуктивності.

Ще однією важливою аспектом є забезпечення безпеки та конфіденційності обробки візуальних даних, особливо у випадку застосування у медицині чи системах відеоспостереження. Розробка надійних методів шифрування та захисту від несанкціонованого доступу до важливої візуальної інформації стає актуальною задачею [6].

Перспективи розвитку включають у себе інтеграцію комп'ютерного зору з іншими галузями штучного інтелекту, такими як природна мова та обробка мовлення, для створення більш повних та комплексних систем інтелектуального аналізу інформації.

Усі ці виклики та перспективи визначають новий етап розвитку комп'ютерного зору, де дослідники та фахівці активно працюють над

вирішенням складних завдань для створення ефективних та надійних систем обробки візуальної інформації.

Сучасні дослідження у галузі комп'ютерного зору нестримно рухаються вперед, пропонуючи нові та захопливі тенденції, що перетворюють цю галузь на важливий напрямок в інформатиці.

Однією з головних тенденцій є поєднання комп'ютерного зору з розширеною реальністю (AR) та віртуальною реальністю (VR). Це відкриває нові можливості у сферах ігор, освіти та тренувань, де можна використовувати візуальні елементи для покращення враження від взаємодії з віртуальним оточенням [7].

Ще однією важливою тенденцією є розвиток методів самонавчання та здатностей до адаптації в алгоритмах комп'ютерного зору. Моделі, які можуть самостійно вдосконалюватися та пристосовуватися до нових умов, стають ключовим елементом для розв'язання реальних завдань у різних областях, від автономних автомобілів до систем відеоспостереження.

Застосування комп'ютерного зору також активно розглядається в медичній сфері, де він допомагає у ранньому виявленні захворювань та надає точніші методи діагностики. Такі інновації мають великий потенціал для підвищення ефективності лікування та зменшення ризиків [5].

Завдяки постійному розвитку апаратної бази, з'являються нові можливості для використання комп'ютерного зору в реальному часі. Швидка обробка великих об'ємів даних стає реальністю, відчиняючи двері для застосувань у таких галузях, як автоматизована торгівля, моніторинг екологічних умов та безпеки.

Усі ці тенденції та інновації свідчать про те, що комп'ютерний зір залишається в центрі інтересу та розвитку, визначаючи майбутнє інформаційних технологій [9].

Разом із стрімким ростом технологій комп'ютерного зору виникають серйозні етичні та соціокультурні питання, які потребують ретельного вивчення та вирішення.

Одним із основних етичних викликів є питання приватності та захисту особистих даних в контексті використання великих об'ємів візуальної інформації. Розвиток систем відеоспостереження та розпізнавання облич піднімає питання про те, як зберігаються, обробляються та використовуються ці дані, забезпечуючи при цьому високий рівень конфіденційності [5].

Ще однією етичною проблемою є використання систем комп'ютерного зору у військовій сфері. Такі системи можуть використовуватися для автоматизованого виявлення та відстеження цілей, що породжує ризики використання технологій для неправомірних цілей або навіть вирішення морально неправильних завдань [3].

Додатковим соціокультурним викликом є вирівнювання можливостей та доступу до технологій комп'ютерного зору. Забезпечення рівних можливостей для всіх користувачів та запобігання виникненню технологічного розриву є важливим завданням, яке вимагає уваги та зусиль.

Розуміння та вирішення цих етичних та соціокультурних викликів є невід'ємною частиною подальшого розвитку комп'ютерного зору. Суспільство повинно активно працювати над створенням ефективних регуляцій та стандартів, щоб забезпечити відповідальне та етичне використання цих технологій [4].

1.2 Огляд нейронних мереж

Нейронні мережі є основою багатьох сучасних систем комп'ютерного зору та машинного навчання. Вони взяли натхнення зі структури людського мозку й призначені для розпізнавання складних зразків і тенденцій у наборах даних, що інколи важко усвідомити або описати за допомогою звичайного програмування [3].

Основною одиницею обчислення в нейронній мережі є нейрон, що імітує біологічні нейрони людського мозку. Кожен штучний нейрон отримує вхідні

сигнали, обробляє їх з використанням вагових коефіцієнтів (параметри моделі) та передає генерований сигнал на наступний рівень через функцію активації. Вся мережа навчається знаходити найкращі значення цих вагових коефіцієнтів через процес, званий зворотним розповсюдженням помилки, де помилка між очікуваним та отриманим результатом мінімізується під час тренування [2].

Архітектури нейронних мереж різноманітні і мають різні форми та розміри, в залежності від завдання, для якого вони призначені. Прості мережі можуть складатися з одного шару нейронів для виконання лінійної регресії або класифікації, в той час як глибокі нейронні мережі містять багато шарів, що дозволяє їм виявляти складні зразки в даних.

Ключові терміни:

- нейрон – основний обчислювальний елемент у нейронній мережі, який імітує біологічний нейрон;
- вагові коефіцієнти – параметри у нейронній мережі, що визначають важливість вхідних сигналів;
- функція активації – нелінійна функція у нейронній мережі, яка визначає, чи має нейрон активуватися;
- зворотне розповсюдження – метод тренування нейронних мереж, де помилки розповсюджуються з кінця мережі до початку для оновлення ваг.

Основними архітектурами в галузі комп'ютерного зору є згорткові нейронні мережі (CNNs), які ефективно обробляють зображення за допомогою фільтрів, що виявляють особливості та риси. Варіації CNN, такі як R-CNN, YOLO та Mask R-CNN, були розроблені для специфічних завдань, як-от виявлення об'єктів, класифікація та сегментація [5].

Важливою частиною процесу навчання є тренувальний набір даних, який містить зразки для навчання мережі. Набори даних, такі як ImageNet, COCO, та PASCAL VOC, забезпечують мільйони анотованих зображень, які використовуються для тренування та оцінювання моделей.

Успіх нейронних мереж у комп'ютерному зорі частково можна приписати підвищенню обчислювальної потужності, поліпшенню алгоритмів

та збільшенню доступних даних. Нейронні мережі стали незамінними в багатьох застосуваннях, від автоматичного розпізнавання мови до аналізу медичних зображень, вказуючи на їх величезний потенціал та значення для подальших інновацій у сфері штучного інтелекту [5].

Далі ми зосередимось на попередніх дослідженнях та розглянемо сучасні тенденції, які формують використання нейронних мереж для виявлення об'єктів.

Нейронні мережі, як ключовий елемент систем комп'ютерного зору та машинного навчання, постійно зазнають технологічних вдосконалень та модифікацій. Сучасні тенденції в цій галузі спрямовані на оптимізацію та розширення можливостей нейронних мереж.

Однією з ключових тенденцій є розвиток архітектур глибоких нейронних мереж. Високопродуктивні глибокі мережі дозволяють виявляти більш складні закономірності у даних та вирішувати завдання з високим рівнем складності. Різноманітні архітектури, такі як резидуальні мережі (ResNets) та трансформери, революціонізують способи обробки інформації в нейронних мережах [10].

Застосування нейронних мереж у реальному часі – ще одна важлива тенденція. Завдяки оптимізації алгоритмів та використанню високопродуктивних обчислювальних платформ, вдалося досягти швидкодії в реальному часі при обробці великих об'ємів візуальних даних. Це відкриває нові перспективи для застосування у сферах автономної навігації, медицини та відеоспостереження.

Ще однією важливою тенденцією є збільшення ролі передових методів передачі знань у нейронних мережах. Перенесення навчання та дистіляція знань дозволяють зменшити обсяг навчальних даних, не втрачаючи при цьому ефективність моделі. Це особливо актуально у випадках, коли обмежений обсяг анованих даних [11].

Вдосконалення механізмів збору та підготовки тренувальних даних є ще однією ключовою тенденцією. Розвиток автоматизованих систем аугментації

даних, які доповнюють та розширюють тренувальні набори, дозволяє покращити стійкість моделей та запобігти перенавчанню.

Ці передові тенденції не лише допомагають оптимізувати ефективність нейронних мереж у виявленні об'єктів, а й відкривають шляхи для нових можливостей у сфері комп'ютерного зору та штучного інтелекту [12].

За останні роки спостерігається еволюція функціональності нейронних мереж, яка визначається не лише збільшенням їх ефективності, але й розширенням їх можливостей.

Однією з ключових тенденцій є використання нейронних мереж для складніших завдань, таких як сегментація та розпізнавання об'єктів в тривимірному просторі. Мережі, що можуть адаптуватися до тривимірного оточення, стають важливим інструментом у віртуальній та розширеній реальності, медицині та галузі промисловості [14].

До інших перспективних напрямків розвитку належить поєднання нейронних мереж з технологією блокчейн [10]. Це не тільки дозволяє забезпечити високий рівень безпеки для обробки важливих візуальних даних, але і створює нові можливості для спільного використання моделей та обчислювальних ресурсів у розподілених системах.

Ще однією перспективою є розвиток нейронних мереж для аналізу емоцій та виявлення неявних патернів у виразах обличчя. Це може мати важливі застосування в рекламі, виробництві та навіть в психіатричній медицині [15].

Зростання потужності обчислювальних пристроїв також стимулює розробку більш ефективних алгоритмів оптимізації нейронних мереж. Використання квантових обчислювальних технологій може допомогти значно покращити продуктивність та швидкодію моделей.

Ці еволюційні зміни в нейронних мережах свідчать про неперервний розвиток цієї галузі та розширення її можливостей у великій кількості сфер, від науки до практичного застосування в різних галузях промисловості.

З інтенсивним ростом розміру та різноманітності даних, з якими нейронні мережі мають справу, виникає потреба в подальшому розвитку методів інтерпретації результатів моделей. Поєднання нейронних мереж із засобами візуалізації та виведення відомого контексту дозволить краще розуміти, як моделі приймають рішення та чинять вплив [2].

Однією з важливих тенденцій є розвиток методів пояснення рішень, що допомагають зрозуміти, яким чином нейронні мережі приходять до конкретних висновків. Застосування методів атрибуції ваг та глибокого аналізу активацій дозволяє розкрити, які частини вхідних даних впливають на кінцевий результат.

До інших напрямків розвитку входять методи зменшення розмірності даних та ефективне використання прийнятних об'ємів пам'яті. Техніки, такі як компресія моделей та кластеризація, стають актуальними для оптимізації роботи нейронних мереж, зокрема в умовах обмежених ресурсів [5].

Далі, враховуючи зростання впливу нейронних мереж у сферах, де важливо поєднання експертного досвіду з автоматизованими системами, виникає потреба в розробці гібридних моделей. Такі моделі можуть інтегрувати різноманітні джерела інформації та досягати більшого рівня розуміння та надійності у вирішенні завдань.

Не менш важливою є інтеграція нейронних мереж із системами постійного навчання (online learning), що дозволяє моделям адаптуватися до змін у вхідних даних в реальному часі. Це стає особливо важливим у вимірюваннях та системах моніторингу [7].

Загалом, розширення можливостей та підвищення інтерпретації різноманітних даних є кроком у напрямку створення більш прозорих, гнучких та більш надійних систем комп'ютерного зору.

Гібридизація нейронних мереж з елементами експертних систем виявляється дедалі важливішою тенденцією. Це може включати інтеграцію додаткових даних, які можуть бути складні для виявлення нейронними мережами, або врахування експертних знань у вирішенні конкретних завдань.

Особливий акцент слід робити на розробці гібридних систем, які можуть працювати в умовах обмежених ресурсів та навіть у відсутності постійного доступу до мережі. Це особливо актуально у сферах, де вимагається низька латентність та висока ефективність, наприклад, в автономних транспортних засобах та розумних промислових системах [8].

Поширення нейронних мереж у сфері медицини та біології породжує нові виклики та напрями розвитку. Інтеграція знань експертів у медичних гібридних системах може сприяти не лише точності діагностики, але й вдосконаленню рішень у персоналізованій медицині.

Нарешті, питання ефективного зберігання та передачі великих об'ємів даних, які використовуються для тренування нейронних мереж, є актуальним напрямком. Використання розподілених систем, технологій блокчейн та квантових обчислювань може стати вирішальним у забезпеченні ефективного обміну даними та збереженні конфіденційності [5].

Ці напрями розвитку підкреслюють постійну потребу у взаємодії між різними галузями науки та технологій для створення високопродуктивних, гнучких та ефективних гібридних систем, які здатні до адаптації до зростаючого обсягу та різноманітності даних.

1.3 Попередні дослідження та сучасні тенденції в галузі нейронних мереж для виявлення об'єктів

В області виявлення об'єктів з допомогою нейронних мереж було зроблено значні прориви, що починаються з базових підходів, таких як шаблонне виявлення та просте розпізнавання форм, і досягають розробки складних згорткових нейронних мереж (CNNs), які можуть визначати і класифікувати різноманітні об'єкти в реальному часі із високою точністю [10].

Попередні дослідження в галузі виявлення об'єктів в основному зосереджувались на фільтрах та текстурних характеристиках. Ранні системи

використовували характеристичні признаки та адаптивні каскадні класифікатори для швидкого виявлення об'єктів, особливо облич. Ці методи, хоч і були ефективними для свого часу, обмежувались визначенням об'єктів, які мали чіткі структурні характеристики і були зображені під контрольованими умовами [11].

Сучасні тенденції. З розвитком глибокого навчання, особливо згорткових нейронних мереж, в області виявлення об'єктів відбулася революція. Винахід архітектур, таких як R-CNN (Regions with CNN features), які вперше поєднали регіональні пропозиції з CNN, відкрив нові можливості для точного виявлення об'єктів. За R-CNN послідували вдосконалення, такі як Fast R-CNN та Faster R-CNN, які значно покращили швидкість та точність за рахунок інтеграції виявлення об'єктів і класифікації в один процес [14].

Іншою ключовою тенденцією є поява архітектур з одного проходження, таких як YOLO (You Only Look Once) і SSD (Single Shot MultiBox Detector), які здатні виявляти об'єкти на зображеннях в реальному часі з високою точністю, що відкриває шлях для їх використання в робототехніці, автономних транспортних засобах та системах нагляду [17].

Зростаюча різноманітність і обсяг даних, а також потреба в автоматизації процесу виявлення об'єктів спонукала до розробки алгоритмів з глибоким навчанням без нагляду та з частковим наглядом. Ці підходи дозволяють машинам самостійно вчитися і адаптуватися до нових даних без залучення великої кількості анотованих зразків, які часто важко отримати.

Науковці також активно працюють над підходами зменшення розміру моделей та зниження вимог до обчислювальних ресурсів для впровадження нейронних мереж на пристроях з обмеженою обчислювальною потужністю, таких як мобільні телефони та вбудовані системи.

У контексті сучасних досліджень в галузі комп'ютерного зору, велика увага приділяється також етичним аспектам використання штучного інтелекту, зокрема питанням конфіденційності та упередженості в алгоритмах,

що наголошує на необхідності розробки відповідних нормативно-правових рамок.

Ці інновації та тенденції вказують на динамічний розвиток галузі та відкривають шлях для нових напрямків досліджень та застосувань нейронних мереж у виявленні об'єктів.

Наступний розділ присвячений аналізу програмних технологій для побудови моделей виявлення об'єктів, які базуються на нейронних мережах, де ми розглянемо ключові інструменти та платформи, що використовуються дослідниками та розробниками в цій сфері [22].

Розвиток сучасних технологій для побудови моделей виявлення об'єктів на основі нейронних мереж відзначається розмаїттям інструментів та платформ. Цей технологічний ландшафт є важливим компонентом для науковців і розробників, які досліджують та впроваджують нові методи виявлення об'єктів.

Одним із ключових інструментів є TensorFlow [12], відкрите програмне забезпечення, розроблене Google. Воно надає широкий спектр інструментів для побудови та тренування нейронних мереж, у тому числі для виявлення об'єктів. TensorFlow використовується у великій кількості досліджень та комерційних проектах завдяки своїй ефективності та гнучкості.

Ще однією популярною платформою є PyTorch [14], яка отримала значну популярність у спільноті дослідників у галузі машинного навчання. PyTorch відзначається зручним інтерфейсом та дозволяє ефективно використовувати глибокі нейронні мережі для різноманітних завдань, включаючи виявлення об'єктів.

Крім того, для реалізації алгоритмів виявлення об'єктів часто використовується фреймворк OpenCV [16], який надає широкі можливості для обробки зображень та використовується для попередньої обробки даних перед їх використанням у нейронних мережах.

Також варто відзначити наявність спеціалізованих бібліотек, таких як Detectron [17] та YOLO [18], які спрощують використання складних архітектур для виявлення об'єктів та надають готові рішення для реалізації.

У сучасному технологічному ландшафті також відбувається активний розвиток та дослідження квантових обчислювань для вирішення завдань у галузі нейронних мереж. Це може відкрити нові перспективи щодо ефективності та обчислювальної потужності.

Загалом, розвиток технологічного ландшафту в області виявлення об'єктів на основі нейронних мереж свідчить про постійні зусилля спільноти для створення ефективних та доступних інструментів для досліджень та реалізації цих інноваційних підходів.

Попри значний прогрес у галузі виявлення об'єктів за допомогою нейронних мереж, існують виклики, які потребують уваги для подальшого розвитку цього напрямку.

Один із важливих викликів – це необхідність постійного вдосконалення точності моделей та їхньої здатності працювати в різних умовах. Зокрема, покращення адаптивності моделей до змін освітлення, перспектив та інших факторів може зробити їх більш універсальними та застосовними [10].

Ще однією проблемою є питання етичного використання нейронних мереж в галузі виявлення об'єктів [23]. Зростання кількості даних, які використовуються для тренування моделей, підкреслює необхідність вирішення питань конфіденційності та упередженості, щоб уникнути можливих негативних впливів на суспільство.

Додатково, перспективною галуззю розвитку є збільшення швидкості та ефективності обробки зображень в реальному часі. Це стає актуальним у зв'язку з застосуванням виявлення об'єктів у сферах, таких як автономні транспортні засоби та системи безпеки, де відмінна реакція на зміни в оточенні є критичною [25].

Напрямок розвитку також пов'язаний із розробкою нових архітектур нейронних мереж, які будуть більш ефективними у виявленні об'єктів на зображеннях великої роздільної здатності та з високою комплексністю.

У майбутньому, можливо, буде акцент на вдосконаленні підходів до тренування моделей, зокрема, використанням менших наборів даних або навчанням з частковим наглядом [22]. Це може полегшити процес розгортання нейронних мереж в умовах, де обмежені ресурси для збору анотованих даних.

Загалом, не дивлячись на досягнення у галузі виявлення об'єктів за допомогою нейронних мереж, виклики та перспективи розвитку дозволяють спрямовувати зусилля науковців та розробників на подальші інновації та покращення в цій захоплюючій галузі комп'ютерного зору [12].

Застосування нейронних мереж для виявлення об'єктів має величезний потенціал у різних сферах. Однією з найбільш важливих галузей є медицина, де системи виявлення об'єктів можуть допомагати при аналізі медичних зображень та виявленні патологій, сприяючи точнішому та швидшому діагностуванню [23].

В автопромисловості виявлення об'єктів грає ключову роль у системах безпеки та підтримки автономних транспортних засобів. Нейронні мережі можуть виявляти перешкоди, розпізнавати дорожні знаки та інші об'єкти, що допомагає підвищити безпеку на дорозі та забезпечити ефективну роботу транспортних систем [22].

В сфері робототехніки системи виявлення об'єктів використовуються для навігації та взаємодії з оточенням. Роботи, оснащені такими системами, можуть ефективно пристосовуватися до змінних умов та виконувати завдання в різноманітних середовищах.

Однак із зростанням ролі нейронних мереж в суспільстві, з'являються нові етичні виклики та питання конфіденційності. Важливо забезпечити відповідальне використання цих технологій та розробляти нормативно-правові механізми для захисту приватності та запобігання можливій упередженості в роботі алгоритмів [12].

Майбутні виклики включають подальше поліпшення точності та надійності моделей виявлення об'єктів, адаптацію до нових типів об'єктів та сценаріїв, а також розробку енергоефективних методів для використання вбудованих систем та мобільних пристроїв [8].

У підсумку, нейронні мережі для виявлення об'єктів відкривають нові перспективи у багатьох галузях та забезпечують потужний інструмент для автоматизації та підвищення ефективності багатьох процесів. З розвитком технологій та вирішенням відповідних викликів, ці системи можуть стати невід'ємною частиною нашого щоденного життя, сприяючи подальшим інноваціям у сфері штучного інтелекту.

1.4 Програмні технології для побудови моделей виявлення об'єктів на основі нейронних мереж

Побудова моделей нейронних мереж для виявлення об'єктів потребує використання спеціалізованих програмних технологій та інструментів, які спрощують розробку, навчання та імплементацію цих комплексних систем. Сучасні технології включають широкий спектр бібліотек глибокого навчання, фреймворків та середовищ, які надають потужні абстракції та інтерфейси для ефективної роботи з нейронними мережами.

Бібліотеки глибокого навчання, такі як TensorFlow, Keras та PyTorch домінують у дослідженнях та розробці в галузі штучного інтелекту [23]. TensorFlow, розроблений Google, є однією з найпопулярніших платформ для створення складних моделей машинного навчання, зокрема для завдань комп'ютерного зору. Keras пропонує ще більш високорівневий API, що інтегрується з TensorFlow, і дозволяє швидко експериментувати з нейронними мережами. PyTorch, розроблений Facebook, набуває популярності завдяки своїй гнучкості та динамічній природі графів обчислень, що робить його ідеальним для дослідницьких проєктів та прототипування.

Фреймворки для виявлення об'єктів, як-от Detectron2 від Facebook AI Research (FAIR) та TensorFlow Object Detection API, надають готові до використання модулі та архітектури, такі як Faster R-CNN, YOLO та SSD спрощують розробку та тестування моделей виявлення об'єктів. Ці інструменти значно скорочують час розробки та сприяють стандартизації процесів навчання та оцінки моделей [22].

Для ефективного впровадження нейронних мереж у виробництво, важливо забезпечити оптимізацію моделей, яка дозволить їм працювати в реальному часі на різноманітному обладнанні. Середовища виконання та оптимізації, такі як NVIDIA CUDA та cuDNN, надають інструменти для оптимізації та прискорення обчислень на графічних процесорах (GPU), тоді як OpenVINO від Intel спрямоване на прискорення інференції моделей на процесорах (CPU) та інтегрованих графічних системах [13].

Незважаючи на достатньо широкий вибір інструментів, мова програмування Python залишається мовою вибору для більшості дослідників і розробників у сфері штучного інтелекту [28]. Простота, гнучкість та багата екосистема бібліотек роблять Python ідеальною мовою для роботи з нейронними мережами.

Контейнеризація та хмарні обчислення – технології, такі як Docker та Kubernetes, а також хмарні платформи від Amazon Web Services, Google Cloud Platform і Microsoft Azure, пропонують величезні можливості для масштабування та гнучкого розгортання моделей нейронних мереж. Хмарні обчислення також надають доступ до спеціалізованого обладнання для навчання моделей, як GPU та TPU, без необхідності інвестування в власну інфраструктуру [23].

За допомогою цих інструментів, дослідники та розробники можуть підходити до розробки нейронних мереж для виявлення об'єктів з більшою швидкістю, ефективністю та точністю, що прискорює процес від ідеї до реалізації [7].

Інтеграція з системами збору даних – побудова моделей виявлення об'єктів також включає в себе важливий аспект інтеграції з системами збору даних. Це може включати в себе автоматичний збір, аналіз та передачу даних з різноманітних джерел, таких як відеокамери, сенсори, та IoT-пристрої. Така інтеграція дозволяє моделям працювати в реальному часі та адаптуватися до змін у навколишньому середовищі [8].

Оптимізація для мобільних платформ – зростання зацікавленості в застосуванні нейронних мереж на мобільних пристроях визначає необхідність оптимізації для обмежених ресурсів [4]. Інструменти, такі як TensorFlow Lite для Android та iOS, дозволяють створювати моделі, які працюють оптимально на мобільних пристроях, зберігаючи високу точність.

Розширення мереж на засоби AR/VR – застосування нейронних мереж у віртуальній та доповненій реальності відкриває нові можливості. Моделі, оптимізовані для AR/VR пристроїв, можуть надавати розширені можливості взаємодії з навколишнім світом, а також забезпечувати високу точність виявлення об'єктів у цих умовах [7].

Навчання з підкріпленням для вдосконалення моделей – використання технік навчання з підкріпленням може покращити результати моделей виявлення об'єктів. Цей підхід дозволяє моделі вчитися на основі взаємодії з оточуючим середовищем та отримувати зворотний зв'язок для оптимізації роботи в реальних умовах.

Враховуючи поширене застосування нейронних мереж, важливо враховувати й пов'язані з ними етичні аспекти та питання безпеки. Розробники повинні активно займатися питаннями конфіденційності та уникнення упередженості в алгоритмах, забезпечуючи справедливі та етичні застосування своїх моделей [7].

Ці аспекти не лише доповнюють сучасні програмні технології, але і визначають нові напрямки для подальших досліджень та розвитку в галузі виявлення об'єктів за допомогою нейронних мереж.

Вдосконалення архітектур для відомих викликів – з великим розвитком галузі нейронних мереж, вчені та розробники працюють над постійним вдосконаленням архітектур для вирішення відомих викликів. Зокрема, важливим завданням є підвищення точності виявлення об'єктів в умовах обмежених освітлених умов, а також підвищення роботи моделей у високодинамічних та швидкозмінних сценах [9].

Автоматизовані засоби створення датасетів – важливою частиною розробки моделей виявлення об'єктів є налагодження та тренування на відповідних датасетах. Останнім часом, надто швидкі зміни в галузі вимагають автоматизованих засобів створення анотованих даних, які є необхідними для навчання моделей [21].

Застосування глибокого навчання для аналізу відео та стрімінгу – розвиток глибокого навчання спрямований на впровадження в системи виявлення об'єктів для аналізу великих обсягів відеоданих та стрімінгу [14]. Це включає в себе вдосконалення алгоритмів для розпізнавання та відстеження об'єктів у режимі реального часу.

Системи виправлення помилок в навчанні: з врахуванням того, що моделі нейронних мереж можуть вчитися на неповних або неправильно анотованих даних, важливим є розвиток систем виправлення помилок. Ці системи можуть автоматично коригувати та покращувати якість анотацій, що сприяє збільшенню ефективності навчання.

Інтеграція з технологіями блокчейн – використання технологій блокчейн для забезпечення конфіденційності та достовірності даних у моделях виявлення об'єктів стає актуальним завданням [5]. Це може бути важливим для додаткового захисту особистої інформації та уникнення маніпуляцій даними.

Інновації в області виявлення об'єктів за допомогою нейронних мереж розширюють можливості застосування та підвищують продуктивність моделей у різних сценаріях, як це зазначено в джерелі [9]. Це дозволяє розширити сферу використання та покращити якість роботи моделей при застосуванні їх у різних умовах та завданнях.

У сфері виявлення об'єктів на основі нейронних мереж використовуються різноманітні методи, кожен з яких має свої переваги та обмеження.

Згорткові нейронні мережі (CNNs) визнані своєю високою ефективністю в обробці зображень. Їхній успіх полягає у здатності автоматично виявляти просторові залежності та розпізнавати важливі зразки у великих обсягах візуальної інформації [10]. Це робить їх особливо корисними для задач виявлення об'єктів, де важлива деталізація та контекст важливі для точного розпізнавання.

Однак CNNs не є панацеєю, існують обмеження цього підходу. Наприклад, вони можуть бути витратними з точки зору обчислень, особливо при роботі з великими об'ємами зображень з високою роздільною здатністю. Крім того, вони можуть виявлятися менш ефективними в умовах, де об'єкти розташовані у великій відстані один від одного або коли потрібно виявляти об'єкти різних розмірів [10].

З часом з'явилися інші підходи, такі як YOLO (You Only Look Once) та SSD (Single Shot MultiBox Detector). Вони переважають швидкістю та здатністю виявляти об'єкти в реальному часі. Наприклад, YOLO дозволяє виявляти об'єкти за один прохід зображенням, що робить його дуже ефективним для задач реального часу [4].

Незважаючи на ці переваги, методи з одного проходу можуть страждати від меншої точності порівняно зі складнішими архітектурами, такими як Faster R-CNN, які використовують більше етапів обробки для досягнення вищої точності. Також, при роботі з більш складними сценаріями чи об'єктами різних розмірів, може виникати виклик в утриманні високої ефективності [8].

Додатковою проблемою для обох підходів є вимоги до обчислювальних ресурсів. Навіть з ростом обчислювальної потужності, навчання та використання складних моделей може вимагати значних обчислювальних ресурсів, що може бути важким завданням для обладнання із обмеженими можливостями [16].

В зв'язку з цим, важливо продовжувати дослідження та розробку нових методів, спрямованих на оптимізацію швидкодії та точності виявлення об'єктів в різних умовах. Також, з урахуванням етичних аспектів використання штучного інтелекту, важливо розвивати стандарти та регулювання, щоб забезпечити відповідальне та безпечне застосування цих технологій в різних галузях [7].

Ще однією важливою архітектурою є R-CNN (Region-based Convolutional Neural Network) та її вдосконалені версії. Ці методи використовують регіональні пропозиції для визначення потенційних областей інтересу на зображенні, які подаються на вхід глибокій мережі для подальшого виявлення та класифікації об'єктів. Такий підхід дозволяє досягти високої точності, особливо коли маємо справу зі зображеннями великої роздільної здатності та великим числом об'єктів на зображенні [20].

Сучасними тенденціями в галузі виявлення об'єктів є використання технологій зменшення розміру моделей та оптимізації для пристроїв із обмеженими обчислювальними ресурсами. Наприклад, MobileNet використовує глибокі мережі, спеціально оптимізовані для мобільних пристроїв, забезпечуючи ефективність та точність виявлення об'єктів на ресурсозберігаючих платформах.

Недоліком деяких сучасних методів є їхній високий рівень складності, особливо при використанні глибоких мереж для виявлення об'єктів у реальному часі. В таких випадках, обчислювальні витрати можуть стати критичним фактором, що обмежує їхню практичність [22].

Існує проблема впровадження нейронних мереж у виробниче середовище. Це пов'язано з потребою великих обчислювальних потужностей та навчання моделей на великих наборах даних [23]. Питання конфіденційності та етики також стає важливими аспектами при розгляді реалізації цих технологій.

Щоб вирішити ці виклики, дослідники та розробники зосереджуються на розвитку більш ефективних алгоритмів, оптимізації для роботи в умовах

обмежених ресурсів, а також вдосконаленні процесів навчання з обмеженою кількістю даних. Паралельно із цим, виникає необхідність у встановленні чітких етичних стандартів та правових рамок для використання нейронних мереж в галузі виявлення об'єктів з метою забезпечення безпеки, конфіденційності та відповідального використання цих технологій у суспільстві [2].

Висновки до розділу 1

У першому розділі було представлено всебічний огляд нейронних мереж та їхнього застосування у комп'ютерному зорі, з особливим акцентом на моделі виявлення об'єктів. Основні поняття та терміни були визначені для забезпечення твердого теоретичного фундаменту, необхідного для розуміння складності та можливостей сучасних нейронних мереж.

Вивчення історичних та сучасних підходів до нейронних мереж виявило зростання їх можливостей і практичної застосовності. Попередні дослідження підкреслили значний прогрес у точності та швидкості нейронних мереж, в той час як аналіз сучасних тенденцій показав шляхи їхньої оптимізації та адаптації до нових викликів.

У контексті програмних технологій, було розглянуто низку інструментів і фреймворків, які полегшують процес створення та впровадження нейронних мереж у реальних системах комп'ютерного зору. Це стосується і вибору програмних мов, і використання спеціалізованих бібліотек та середовищ для навчання та оптимізації моделей.

Загалом, перший розділ закладає фундамент для глибшого аналізу конкретних методів та технік, які будуть використані в подальшому дослідженні для розробки ефективних систем виявлення об'єктів на основі нейронних мереж.

РОЗДІЛ 2

ТЕОРЕТИЧНІ ОСНОВИ ВИЯВЛЕННЯ ОБ'ЄКТІВ НА ОСНОВІ НЕЙРОННИХ МЕРЕЖ

2.1 Архітектури нейронних мереж для виявлення об'єктів

Застосування нейронних мереж, зокрема згорткових нейронних мереж (CNN), покращило сферу комп'ютерного зору завдяки здатності автоматично і ефективно вилучати характеристики з великих наборів даних. Архітектури CNN, на відміну від традиційних методів машинного навчання, можуть розпізнавати візуальні шаблони зображень з мінімальним попереднім обробленням, від низькорівневих деталей до високорівневих атрибутів [3].

Згорткові нейронні мережі (CNN) є типом глибинних нейронних мереж, які містять один чи декілька згорткових шарів, що автоматично та адаптивно вилучають важливі характеристики з вхідних даних, які є зображеннями. У CNN є здатність зберігати просторову ієрархію властивостей, роблячи їх ідеально придатними для задач комп'ютерного зору [12].

Огляд архітектури CNN. Стандартна архітектура CNN містить кілька типів шарів:

- згорткові шари (Convolutional Layers) використовують набір фільтрів, які прокочуються по вхідному зображенню, формуючи карту характеристик, яка виявляє важливі особливості;
- функції активації, зазвичай ReLU вони додають нелінійність до моделі, дозволяючи їй вчитися більш складних патернів;
- об'єднувальні шари (Pooling Layers) зменшують розмір карт характеристик, зберігаючи лише найважливішу інформацію;
- повнозв'язні шари (Fully Connected Layers) використовуються наприкінці мережі для класифікації та ідентифікації об'єктів на основі вилучених характеристик. На рисунку 2.1 відображено типову архітектуру CNN [31].

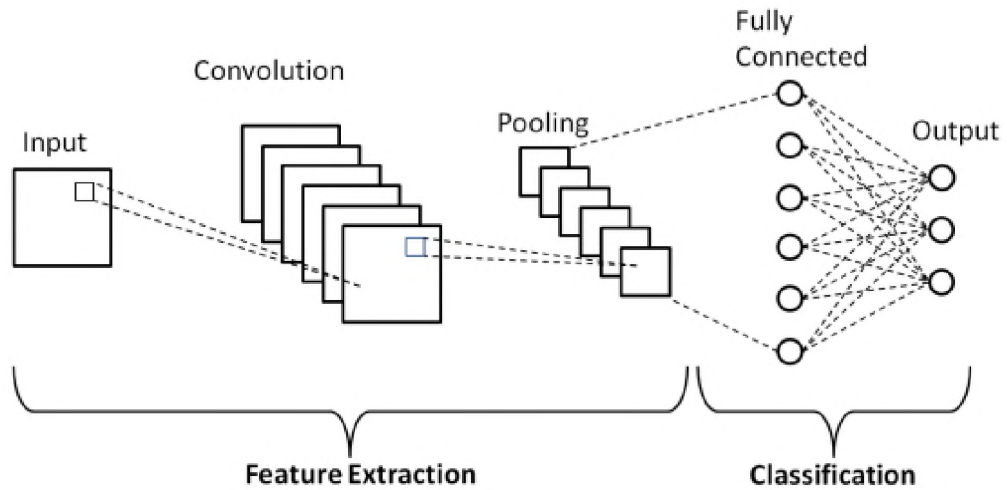


Рисунок 2.1 – Типова архітектура CNN

Використання CNN у задачах виявлення об'єктів. У сфері виявлення об'єктів CNN застосовуються для класифікації та локалізації об'єктів у зображеннях. Існують різноманітні архітектури, спеціалізовані для цих завдань, такі як R-CNN, Fast R-CNN, Faster R-CNN та YOLO (You Only Look Once), які різняться за швидкістю, точністю та здатністю до масштабування.

Ключовою відмінністю цих моделей є те, як вони об'єднують згорткові та класифікаційні шари для ефективною локалізації та класифікації об'єктів. Деякі, як YOLO, виконують виявлення об'єктів в реальному часі, аналізуючи зображення одним проходом, що робить їх придатними для використання в системах відеонагляду та автономному водінні, на рисунку 2.2 відображена типова архітектура YOLO [32].

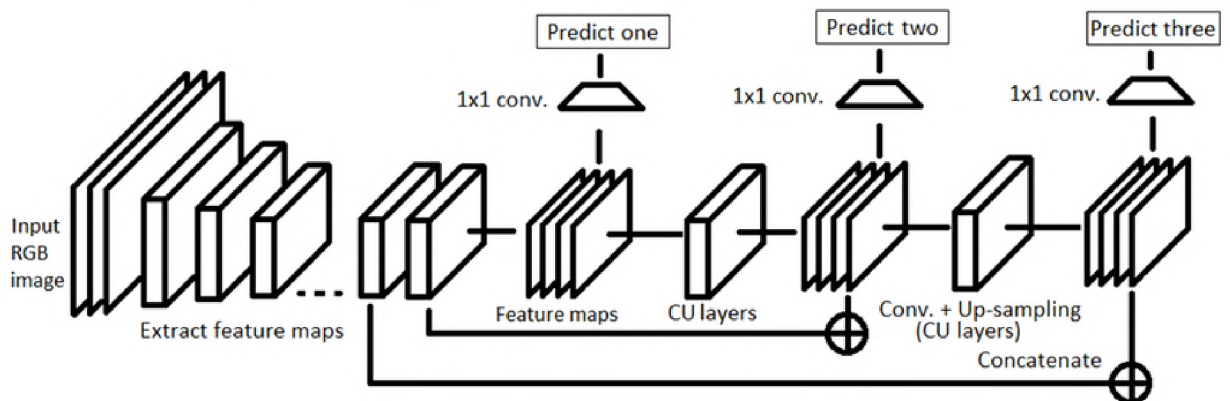


Рисунок 2.2 – Типова архітектура YOLO

Для реалізації цих архітектур часто застосовуються фреймворки глибокого навчання, такі як TensorFlow та PyTorch, які надають велику кількість готових до використання модулів та оптимізованих реалізацій згорткових нейронних мереж.

Далі будуть більш детально розглянуті різноманітні функції активації, шари мережі, алгоритми навчання, а також методи оцінки якості моделей виявлення об'єктів, що допоможе глибше зрозуміти, як можна оптимізувати та оцінити ефективність таких систем.

MIML (Multi-Instance Multi-Label learning) – це підхід до машинного навчання, який дозволяє одному об'єкту мати кілька етикеток (міток), а також дозволяє кожному об'єкту бути представленим декількома інстанціями або прикладами. Це особливо корисно у сценаріях, де об'єкти природньо не є однозначно класифікованими та можуть належати до кількох категорій одночасно, або коли єдиної інстанції недостатньо для представлення всіх характеристик об'єкта.

COREL 5K є набором даних для оцінки методів класифікації та пошуку зображень. Він містить 5000 зображень, розділених на різні категорії. Кожне зображення анотовано декількома мітками, які відображають зміст зображення. Цей набір даних часто використовується для випробування та оцінки ефективності алгоритмів машинного зору, особливо для задач розпізнавання образів, мульти-міткової класифікації та автоматичного анотування зображень.

MSRC (Microsoft Research Cambridge) Dataset є відомим набором даних для семантичного сегментування зображень. Він містить зображення різних сцен, які вручну анотовані на основі пікселів для визначення категорій об'єктів, таких як дерева, будівлі, небо, вода тощо. Набір даних MSRC використовується для тренування та тестування алгоритмів комп'ютерного зору з метою розуміння змісту зображень на рівні об'єктів і їхніх взаємозв'язків, у таблиці 2.1 представлено порівняння різних архітектур CNN для виявлення об'єктів [33].

Таблиця 2.1 – Порівняння різних архітектур CNN для виявлення об'єктів

Dataset	CNN Architecture	AP	AR	F1
MIML	AlexNet	0.683	0.675	0.678
	VGG16	0.756	0.731	0.743
	ResNet101	0.769	0.774	0.771
COREL 5K	AlexNet	0.310	0.384	0.343
	VGG16	0.364	0.432	0.395
	ResNet101	0.372	0.459	0.410
MSRC	AlexNet	0.593	0.614	0.603
	VGG16	0.687	0.695	0.691
	ResNet101	0.432	0.407	0.419

Архітектури нейронних мереж для виявлення об'єктів продовжують розвиватися, а різноманітні підходи забезпечують оптимальні рішення для конкретних викликів та вимог. Однією з важливих архітектур є архітектура R-CNN (Region-based Convolutional Neural Network) та її наступники. R-CNN вирізняється тим, що він використовує регіональні пропозиції для визначення областей інтересу на зображенні, які подаються на вхід глибокій мережі для подальшого виявлення та класифікації об'єктів. Цей підхід вдало поєднує точність та стійкість у виявленні об'єктів [12].

Ще однією важливою архітектурою є YOLO (You Only Look Once). Однією з особливостей YOLO є можливість виявлення об'єктів в реальному часі, адже вона аналізує зображення всього один раз. Це робить YOLO ефективним у використанні в системах відеонагляду, де швидкість і точність є ключовими факторами [13].

Фреймворки глибокого навчання, такі як TensorFlow та PyTorch, грають важливу роль у впровадженні цих архітектур, надаючи зручні інструменти для розробки, навчання та валідації моделей. Їхня популярність визначається не лише потужністю, але і активною спільнотою, що сприяє постійному вдосконаленню та обміну досвідом [14].

Під час аналізу архітектур нейронних мереж для виявлення об'єктів, важливо враховувати розширення глибоких мереж, таких як Squeeze-and-Excitation Networks (SENet) та Residual Networks (ResNet). Вони вирізняються

тим, що дозволяють ефективніше використовувати ресурси мережі та поліпшити якість виявлення об'єктів.

Незважаючи на досягнення, існують виклики, такі як високі обчислювальні витрати та проблеми, пов'язані з невеликою кількістю навчальних даних для деяких класів об'єктів. Професійний розвиток архітектур та підходів до виявлення об'єктів у мережах нейронів продовжується, спрямований на подолання цих викликів та поліпшення загальної продуктивності систем виявлення об'єктів [5].

Не менш важливою є ефективність та особливості архітектур, що використовують механізми переднього та заднього обходу (Encoder-Decoder architectures). Такі мережі, наприклад, U-Net, розроблені для семантичного сегментування, можуть успішно використовуватися і для завдань виявлення об'єктів. Вони здатні не лише визначати просторові місця об'єктів на зображеннях, а й здійснювати їхню детальну класифікацію [14].

Архітектури, які спеціалізуються на роботі з великими об'ємами даних та обробці зображень в реальному часі, також є ключовою. Наприклад, MobileNet оптимізований для використання на мобільних пристроях та вбудованих системах, забезпечуючи баланс між продуктивністю та точністю [13].

Архітектури, такі як R-CNN, YOLO, SSD, Mask R-CNN та багато інших здатні впоратися з викликами різноманітних сценаріїв, таких як змінні умови освітлення, перекриття об'єктів та різні масштаби. Адаптивні архітектури, здатні до ефективної роботи у реальних умовах, є основою для розвитку технологій автономного водіння, робототехніки та систем нагляду [11].

Проблеми ефективності та обчислювальних витрат при використанні різних архітектур полягають у тому, що деякі глибокі мережі можуть вимагати значних обчислювальних ресурсів, що робить їх менш придатними для реалізації на обмежених пристроях, таких як мобільні телефони. У таких випадках, ефективні архітектури, які володіють низькими витратами ресурсів, можуть бути більш привабливими [12].

Недоліки певних архітектур полягають у тому, що у деяких підвищена вимогливість до обчислювальних ресурсів може призвести до збільшення часу навчання та інференції. Деякі архітектури можуть бути більш чутливими до перенавчання та виявляти недостатню адаптивність до нових умов або об'єктів, які не були адекватно представлені під час навчання [14].

Паралельно з розвитком архітектур нейронних мереж для виявлення об'єктів, відбувався значний розвиток у сфері сегментації та використання зображень для покращення моделей. Мережі з архітектурою Encoder-Decoder, наприклад, U-Net, виявляються дуже ефективними у завданнях сегментації, коли необхідно визначити конкретні області на зображенні, пов'язані з певними класами об'єктів. Це дозволяє точніше виділити й розмежувати об'єкти на зображенні для подальшого аналізу та обробки.

Інтеграція тематичних зображень та архітектур для семантичної сегментації може значно поліпшити точність виявлення та класифікації об'єктів. Наприклад, для навчання автономних транспортних засобів, важливо враховувати дорожні знаки, пішоходів та інші об'єкти, які можуть впливати на безпеку руху. Тут можливо вставити тематичні зображення, які ілюструють процес сегментації та взаємодії архітектур з тематичними даними [15].

Додатково, важливо враховувати виклики, пов'язані з обробкою зображень в реальному часі. Розробка архітектур, таких як MobileNet, які оптимізовані для використання на мобільних пристроях, є ключовою для застосування у вбудованих системах та розумних технологіях [10].

В цілому, постійне вдосконалення архітектур нейронних мереж для виявлення об'єктів є результатом неустановних зусиль у покращенні точності, ефективності та універсальності їх застосування в різноманітних галузях. Цей напрямок розвитку спрямований на пошук оптимальних рішень для використання цих мереж, що дозволяє досягати більш високих результатів у виявленні та обробці об'єктів у великому спектрі сфер.

2.2 Функції активації та шари нейронної мережі

Функції активації в нейронних мережах відіграють ключову роль, додаючи нелінійність до вихідних сигналів нейронів, що дозволяє мережам вчитися складних і абстрактних шаблонів даних [12]. У контексті згорткових нейронних мереж для виявлення об'єктів, популярними функціями активації є:

- ReLU (Rectified Linear Unit) проста функція, яка повертає вхід, якщо він позитивний, інакше повертає нуль. Це допомагає уникнути проблеми зникнення градієнту при навчанні глибоких мереж;

- Leaky ReLU варіація ReLU, яка дозволяє невеликі градієнти при негативних вхідних значеннях, що допомагає підтримувати активність нейронів протягом навчання;

- Sigmoid виводить результати в інтервалі від 0 до 1, традиційно використовується для двійкової класифікації на виході мережі;

- Softmax загальна функція для мультикласової класифікації, яка перетворює вектор забезпечених вхідних значень у розподіл ймовірностей.

Шари мережі – згорткові нейронні мережі складаються з різних типів шарів, кожен з яких виконує унікальну функцію:

- згорткові шари відповідальні за вилучення характеристик із зображень. Використання фільтрів дозволяє мережі вчити важливі особливості на різних рівнях абстракції;

- пулінгові шари (Pooling Layers) зменшують просторові розміри карт характеристик для зменшення обчислювальних витрат та кількості параметрів. Найчастіше використовуються max-pooling та average-pooling;

- нормалізаційні шари (Normalization Layers) як Batch Normalization, використовуються для стабілізації навчання, шляхом нормалізації вхідних даних у кожному міні-батчі;

- повнозв'язні шари (Fully Connected Layers) після згорткових та пулінгових шарів, вони використовуються для класифікації зображення на основі вилучених характеристик.

Кожен шар впливає на здатність мережі до навчання та її ефективність в різних завданнях виявлення об'єктів. оптимальна архітектура мережі для конкретної задачі залежить від балансу між точністю виявлення, швидкістю та кількістю доступних даних для навчання [11].

На фоні функцій активації та різних типів шарів, важливим аспектом архітектури нейронних мереж для виявлення об'єктів є їх взаємодія та використання в різних комбінаціях. Це створює різноманіття мереж та дозволяє адаптувати їх до конкретних вимог завдань.

У практиці виявлення об'єктів зазвичай використовують змішані архітектури, де різні шари взаємодіють для досягнення оптимальних результатів. Наприклад, включення функцій активації ReLU в згорткових шарах може сприяти вивченню більш складних патернів, тоді як використання softmax на фінальному повнозв'язному шарі дозволяє ефективно вирізняти класи об'єктів [20].

При розробці архітектур нейронних мереж важливою є оптимізація та автоматизація. Використання автоматизованих методів підбору гіперпараметрів, таких як генетичні алгоритми чи оптимізація на основі зразка, може значно покращити продуктивність моделі та скоротити час розробки. Також, техніки передового зменшення розміру моделі (model compression) можуть допомогти оптимізувати швидкість роботи та ефективність використання ресурсів [21].

Незважаючи на досягнення, розробка архітектур для виявлення об'єктів також стикається з викликами. Проблеми, такі як недостатня кількість даних для навчання, розрідженість об'єктів у зображеннях або висока обчислювальна складність, вимагають пошуку новаторських рішень.

З розвитком галузі виникають нові тренди у побудові архітектур для виявлення об'єктів. Наприклад, архітектури, які поєднують в собі глибокі та повнозв'язані нейронні мережі, можуть досягати вражаючої точності виявлення. Також, використання архітектур з атенцією (attention) може

покращити здатність моделей сфокусуватися на важливих деталях об'єктів [28].

З часом спостерігається технологічний прогрес у сфері нейронних мереж, що призводить до появи нових архітектур та методів. Використання архітектур з рекурентними зв'язками, такими як LSTM (Long Short-Term Memory) чи GRU (Gated Recurrent Unit), дозволяє враховувати контекст та послідовності в зображеннях, що є критичним для виявлення об'єктів у складних сценах [22].

Одним із ключових аспектів архітектур є їх ефективність на реальних пристроях. Зокрема, архітектури, спрямовані на вбудовані системи або мобільні пристрої, повинні бути оптимізовані для ресурсів та енергоефективності.

Важливо враховувати, що оптимальна архітектура може залежати від конкретного завдання. Завдання, такі як виявлення облич, вимагають відмінних архітектур в порівнянні з завданнями сегментації зображень чи виявлення об'єктів у відеопотоці.

Окрім виявлення об'єктів у зображеннях, архітектури нейронних мереж знаходять застосування в різних сферах, таких як медицина, автономні автомобілі, відеоспостереження та інші [12]. Адаптація та розширення існуючих архітектур для вирішення нових завдань є актуальним напрямком досліджень.

Для успішного впровадження архітектур в практику важливо провести оцінку їхньої ефективності. Метрики, такі як точність, чутливість, специфічність та час виконання, грають важливу роль у визначенні придатності конкретної архітектури для конкретного використання [23].

Розвиток архітектур для виявлення об'єктів нерозривно пов'язаний з загальними тенденціями у сфері машинного навчання та штучного інтелекту. Зокрема, зростання обчислювальних можливостей, збільшення обсягу доступних даних та вдосконалення алгоритмів навчання перспективно впливатимуть на архітектури майбутніх систем виявлення об'єктів.

2.3 Алгоритми навчання нейронних мереж

Алгоритми навчання нейронних мереж відіграють критичну роль у тому, як мережа адаптується та оптимізує свої ваги у відповідь на навчальні дані. Нижче представлено основні алгоритми та методи навчання, які широко застосовуються у нейронних мережах для виявлення об'єктів.

Градiєнтний спуск (Gradient Descent) – це основний алгоритм оптимізації, що використовується для оновлення ваг в мережі за допомогою обчислення градієнтів функції втрат. Стандартний градієнтний спуск оновлює ваги після кожної ітерації над усіма даними навчального набору [12].

Стохастичний градієнтний спуск (SGD) – варіація градієнтного спуску, де ваги оновлюються після кожного прикладу навчального набору або після маленького підмножини, званої міні-батчем. Стохастичний характер допомагає алгоритму вибиратися з локальних мінімумів.

Моментум – метод, який допомагає прискорити SGD у правильному напрямку та гасити осциляції, шляхом включення поняття інерції, яке уявляє собою «накопичення» попередніх градієнтів при оновленні ваг [13].

Adagrad – алгоритм, який адаптує швидкість навчання для кожного параметра, зменшуючи швидкість для параметрів з великими градієнтами та збільшуючи для параметрів з малими градієнтами.

RMSprop – розроблений, щоб вирішити проблему швидкого зменшення швидкості навчання, що спостерігається в Adagrad, шляхом використання ковзної середньої квадратів градієнтів для адаптації швидкості навчання.

Adam (Adaptive Moment Estimation) – популярний алгоритм, що поєднує ідеї з RMSprop та моментуму, використовуючи оцінки першого (середнє) та другого (невпорядковане середнє) моментів градієнтів для оновлення ваг.

Кожен з цих алгоритмів має свої плюси та мінуси, і вибір специфічного алгоритму часто залежить від конкретної проблеми, розміру даних, і типу архітектури мережі [16].

Ілюстрація алгоритмів, зображена на рисунку 2.3, яка показує осі графіка, що відображають кількість епох навчання по горизонталі (абсциса) та втрати на навчальних даних по вертикалі (ордината) [34].

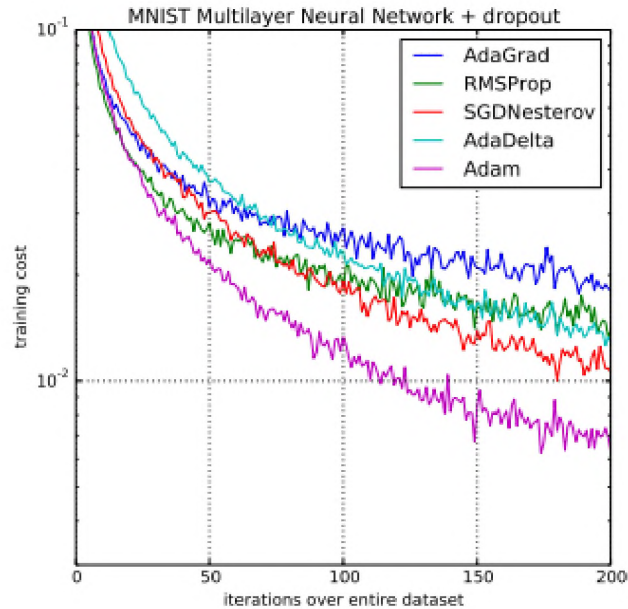


Рисунок 2.3 – Порівняльна схема алгоритмів оптимізації в просторі параметрів

Оцінка ефективності моделей виявлення об'єктів включає в себе використання різних метрик та методів, щоб об'єктивно визначити їхню точність та надійність. Основні підходи до оцінки якості включають метрики класифікації та локалізації:

- точність класифікації (Classification Accuracy) визначає відсоток правильно класифікованих об'єктів. Для виявлення об'єктів це важлива метрика, оскільки допомагає визначити, наскільки ефективно модель класифікує об'єкти на зображенні;

- інтерсекція та об'єднання (Intersection over Union, IoU) визначає ступінь перекриття між прогнозованим та справжнім положенням об'єкта. Широко використовується в задачах локалізації, де важливо визначити, наскільки добре прямокутна рамка моделі апроксимує реальні координати об'єкта.

Втрати на навчальних даних – це міра того, наскільки добре модель передбачає правильний вивід. Чим нижче значення функції втрат, тим краще модель навчилася прогнозувати дані.

Криві ROC та PR:

- крива характеристики роботи приймача (Receiver Operating Characteristic, ROC) допомагає оцінити здатність моделі розділяти класи, налаштовуючи параметр порогу для вибору позитивних та негативних випадків;

- крива точності-повноти (Precision-Recall, PR) визначає залежність точності (точності прогнозів) від повноти (здатності виявити всі позитивні випадки). Особливо важлива в задачах з невеликою кількістю позитивних випадків.

Аналіз впливу параметрів:

- залежність від розміру міні-батча визначає, як змінюється ефективність моделі в залежності від розміру міні-батча під час тренування. Це важливо для оптимізації швидкості та ресурсів;

- вплив регуляризації досліджує, як різні методи регуляризації (наприклад, dropout, L1, L2) впливають на загальну ефективність та загальну здатність моделі.

Ефективність на великих даних:

- крива навчання аналізує, як змінюється функція втрат на тренувальному та валідаційному наборах з ростом обсягу навчальних даних. Показує, наскільки ефективно модель учить в залежності від об'єму даних;

- вплив архітектури досліджує, як змінюється ефективність моделі в залежності від архітектури. Порівняння різних архітектур на великих наборах даних дозволяє визначити найбільш ефективні рішення.

Ці методи оцінки є фундаментальними для постійного вдосконалення моделей та алгоритмів у сфері виявлення об'єктів. Вони є ключовими складовими у досягненні оптимальної точності та надійності розпізнавання об'єктів в реальних умовах [13]. Такий підхід сприяє покращенню якості

прогнозування та адаптації моделей до різноманітних умов та контекстів застосування.

Практично, вибір алгоритму навчання залежить від специфіки задачі. Для комп'ютерного зору і, зокрема, для виявлення об'єктів, важливою є здатність алгоритму швидко адаптуватися до великих та різноманітних датасетів, які можуть містити велику кількість об'єктів та класів. Моделі, що використовують алгоритми з хорошою загальною здатністю до оптимізації, такі як Adam, часто демонструють кращі результати в таких сценаріях [22].

В наступному підрозділі буде описано методи оцінки якості моделей виявлення об'єктів, що є критично важливими для розуміння ефективності алгоритмів навчання та їх впливу на точність та надійність виявлення об'єктів.

2.4 Методи оцінки якості моделей виявлення об'єктів

Оцінка якості моделей виявлення об'єктів має вирішальне значення для розуміння їхньої ефективності і надійності. В цьому підрозділі описані ключові функції втрат та метрики, які використовуються для вимірювання продуктивності моделей нейронних мереж у виявленні об'єктів [12].

Функції втрат (loss functions) мають фундаментальне значення для процесу навчання нейронної мережі. У контексті виявлення об'єктів, функції втрат повинні враховувати не тільки помилки класифікації, але й помилки локалізації [13]. Функції втрат для виявлення об'єктів:

- Cross-Entropy Loss – використовується для класифікаційних задач, щоб виміряти розбіжність між реальними та прогнозованими розподілами ймовірностей класів;

- Mean Squared Error (MSE) – застосовується для регресійних задач, таких як прогнозування координат об'єктів;

- Intersection over Union (IoU) Loss – використовується для вимірювання точності перекриття між прогнозованими та фактичними об'єктами.

Функції втрат представляють собою важливий елемент в процесі навчання нейронних мереж для виявлення об'єктів. Вони є ключовими для зворотного поширення помилок та актуалізації вагових коефіцієнтів мережі під час тренування. Ці функції визначають, яким чином мережа оцінює власні прогнози та реагує на виявлені помилки. Вибір конкретної функції втрат має значний вплив на ефективність та точність моделі виявлення об'єктів, оскільки різні функції можуть варіювати швидкість збіжності моделі та її здатність адаптуватися до різних типів даних чи специфіки конкретної задачі [23].

Об'єктні функції втрат виявляють суттєву важливість у процесі навчання нейронних мереж, спрямованих на виявлення об'єктів. Ці функції на пряму впливають на механізми адаптації ваг та параметрів мережі для максимально ефективного вирішення поставленої задачі [11]. Вони визначають спосіб оцінки помилок мережі під час навчання та регулювання параметрів з метою максимізації точності та адаптивності моделі до конкретних умов завдання.

Ключові аспекти ролі об'єктних функцій втрат:

1. Об'єктні функції втрат використовує механізм зворотнього поширення помилок (Backpropagation). Під час навчання, вони визначають, наскільки віддалено прогнозовані результати мережі від фактичних значень. З цією інформацією нейронна мережа коригує свої ваги та параметри, спробуючи зменшити значення функції втрат.

2. Вибір конкретної об'єктної функції втрат залежить від природи задачі, яку вирішує мережа. Наприклад, для задач класифікації може використовуватися категоріальна кросс-ентропія, тоді як для задач регресії – середньоквадратична помилка. Важливо враховувати особливості конкретного завдання та його вимоги при виборі функції втрат.

3. У контексті моделей виявлення об'єктів у машинному навчанні, об'єктні функції втрат представляють собою ключовий аспект, що визначає ефективність навчання моделі. Ці функції відповідають за оцінку точності прогнозів моделі та коригування помилок під час процесу тренування. Адекватний вибір об'єктної функції втрат допомагає моделі ефективніше

адаптуватися до особливостей даних, покращуючи якість прогнозів. У той же час, неправильний вибір цієї функції може спричинити ускладнення процесу навчання та погіршення якості прогнозів моделі. Тому вибір відповідної об'єктної функції втрат виявляється важливим етапом у підвищенні точності та ефективності моделей виявлення об'єктів у машинному навчанні..

4. Баланс між складністю та інтерпретованістю. Деякі об'єктні функції втрат складніше в оптимізації, інші – більш інтерпретовані. Вибір такої функції, яка враховує баланс між цими аспектами, може бути критичним для досягнення оптимального балансу між точністю та здатністю розуміти природу виявлення об'єктів.

5. Узагальнення та уникнення перенавчання. Деякі функції втрат допомагають уникнути перенавчання, забезпечуючи моделі здатність узагальнювати на нові дані. Це особливо важливо в сценаріях, де модель стикається з різноманітними об'єктами та умовами.

Вибір об'єктної функції втрат є важливим етапом у створенні ефективних систем виявлення об'єктів [24].

Для оцінки ефективності моделі виявлення об'єктів існують спеціальні метрики оцінки точності моделей.

Average Precision (AP) та Mean Average Precision (mAP) – AP враховує precision і recall при різних порогах класифікації, mAP є середнім значенням AP по різних класах або різних рівнях IoU [12].

F1 Score – є метрикою, яка використовується для оцінки моделей машинного навчання, й вона представляє собою гармонічне середнє між precision та recall. Вона надає загальну міру точності виявлення об'єктів. Діаграма F1 Score vs IoU відображає залежність F1 Score від IoU (Intersection over Union) для різних порогів об'єктного виявлення. Ця діаграма, яка відображена на рисунку 2.5, відображає, як змінюється F1 Score при зміні порогів IoU [36].

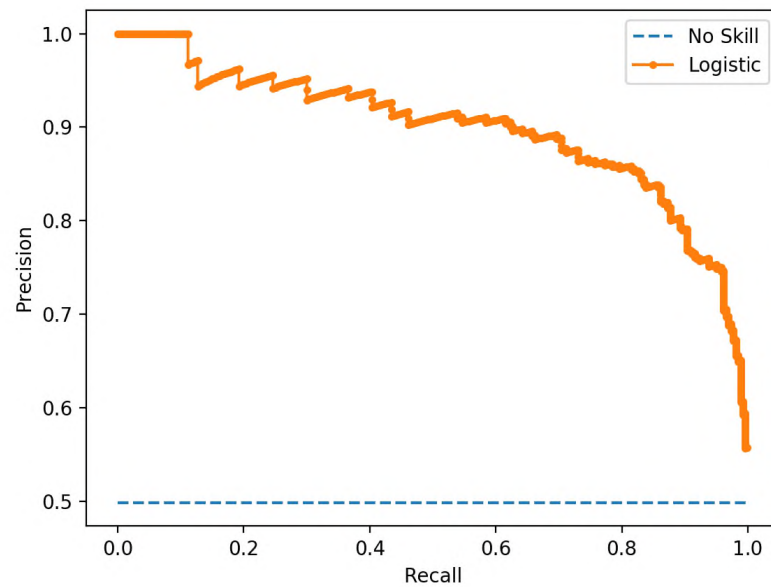


Рисунок 2.4 – Precision-Recall Curve

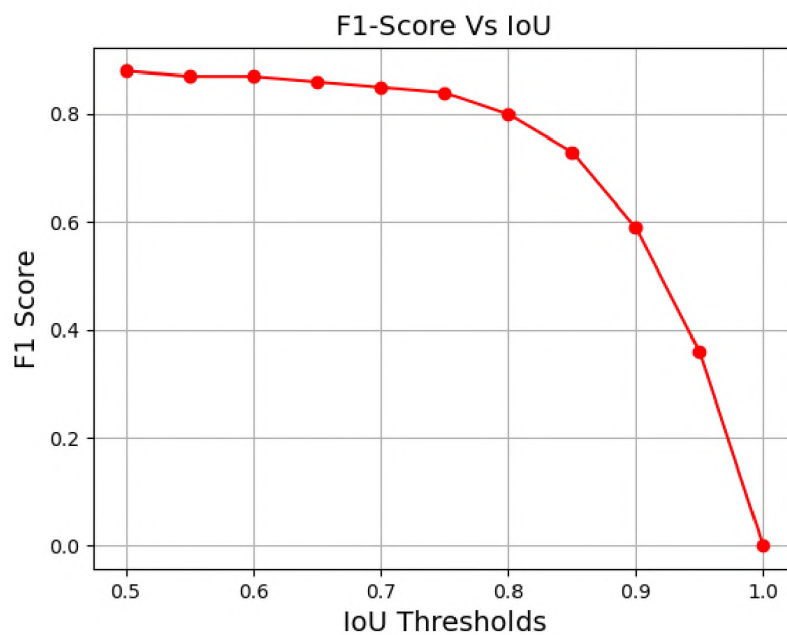


Рисунок 2.5 – F1 Score Vs IoU

Precision і Recall – визначають відсоток правильно виявлених об'єктів з усіх виявлених (precision) та відсоток правильно виявлених об'єктів з усіх реальних об'єктів (recall). Precision-Recall Curve – це графік, який ілюструє зв'язок між precision і recall при різних порогах класифікації, відображений на рисунку 2.4 [35].

Висновки до розділу 2

Таким чином, у другому розділі було детально розглянуто теоретичні аспекти виявлення об'єктів на основі нейронних мереж. Значна увага приділялася вивченню архітектур, зокрема, згорткових нейронних мереж (CNN), які є основою для більшості сучасних систем виявлення об'єктів. Огляд різноманітних архітектур виявив, як еволюціонували методи зорового розпізнавання та які інновації були внесені для покращення точності та швидкості обробки.

Було розглянуто функції активації та шари мережі, що відіграють важливу роль у визначенні нелінійних відносин та взаємозв'язків у даних. Також розглянуто правильний вибір функцій активації та конфігурацій шарів, що він може позитивно вплинути на здатність моделі вчитися та узагальнювати знання з обмежених даних.

При ретельному аналізі алгоритмів навчання в машинному навчанні було розглянуто їхню здатність оптимізувати ваги в нейронних мережах з метою мінімізації помилок у прогнозуванні. Було відзначено значення та механізми градієнтного спуску, а також його різновидів, серед яких ключове місце займають стохастичний градієнтний спуск (SGD), RMSprop та Adam. Ці алгоритми належать до важливих інструментів для успішного навчання глибоких нейронних мереж, сприяючи оптимізації їхньої структури та ефективному вирішенню завдань прогнозування.

У дослідженні розглядалися різноманітні методи оцінки якості моделей, зокрема різні функції втрат, які функціонують як керуючі аспекти в процесі навчання моделей. Крім того, вивчалися різні метрики для оцінки ефективності моделей у виявленні об'єктів. Особлива увага була приділена вибору відповідних метрик для забезпечення точної та адекватної оцінки продуктивності моделі в реальних умовах. Ретельний підбір метрик відіграє ключову роль у правильному оцінюванні та порівнянні моделей, що використовуються у практичних застосуваннях.

РОЗДІЛ 3

РОЗРОБКА МОДЕЛІ ВИЯВЛЕННЯ ОБ'ЄКТІВ ТА ЇЇ АПРОБАЦІЯ

3.1 Збір та підготовка даних для навчання моделі

На першому етапі розробки моделі стоїть задача збору даних, які будуть використовуватися для тренування та тестування. У цьому проєкті основним джерелом даних є зображення облич, захоплені через вебкамеру. Такий підхід дозволяє отримати реалістичний набір даних, що враховує різноманіття особливостей обличчя людей, таких як форма, розмір, колір шкіри, а також різні умови освітлення.

Процес збору даних має включати не тільки різноманіття облич, але й різноманіття сценаріїв, де обличчя можуть з'являтися, включаючи різні фони та освітлення. Це забезпечує комплексність тренувального набору даних і допомагає моделі краще адаптуватися до різних умов при реальному використанні, на рисунку 3.1 відображено процедуру захоплення зображення з середнім рівнем освітлення та різнокольоровим фоном.



Рисунок 3.1 – Процес захоплення зображення через вебкамеру з середнім рівнем освітлення та різнокольоровим фоном

Обробка та аугментація зображень полягає у тому, що отримані зображення потребують додаткової обробки перш ніж їх можна використовувати.

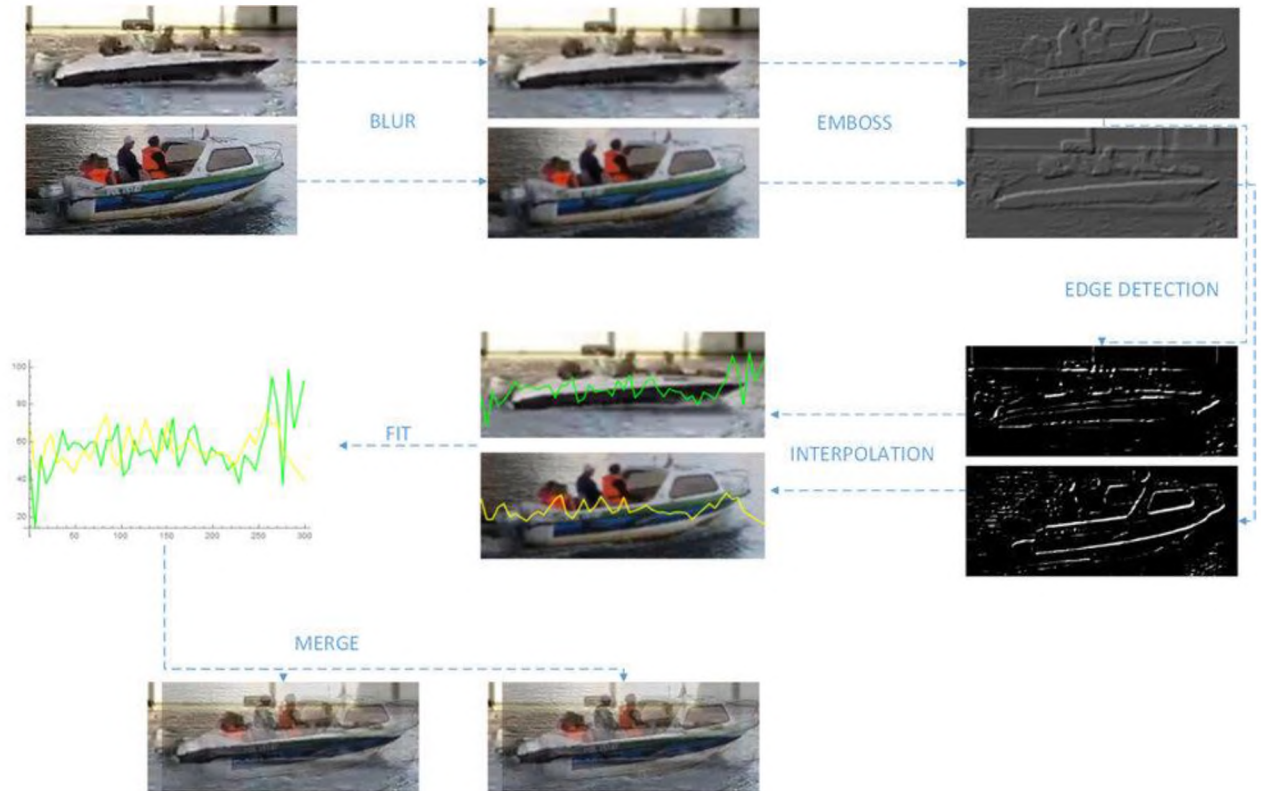


Рисунок 3.2 – Схематичне зображення процесу створення аугментованих зразків із поясненнями кожної операції

У тренуванні моделі, цей процес включає наступні кроки:

- додавання шуму – застосування різних видів шуму, як от гаусівський шум, може зміцнити робастність моделі до випадкових збурень в зображеннях, на рисунку 3.2 зображено процес створення аугментованих зразків із поясненнями кожної операції [37].

- нормалізація – кольори на зображеннях нормалізуються, щоб модель не залежала від особливостей освітлення на різних зображеннях;

- обрізка та ресайзинг – надмірний фон вирізається, а обличчя розташовуються у центрі зображення, після чого проводиться їх масштабування.

Після первинної обробки застосовується аугментація зображень для збільшення обсягу та різноманітності даних:

- геометричні трансформації включають в себе випадкові повороти, зсуви, масштабування та горизонтальне віддзеркалення. Ці перетворення допомагають моделі бути стійкою до різних положень об'єктів у просторі;

- фотометричні трансформації – зміна яскравості, контрастності та насиченості, що дозволяє моделі краще працювати при різних умовах освітлення;

- перетворення зображень – зображення конвертуються до одного стандартного розміру, що є критично важливим для забезпечення однорідності вхідних даних для моделі;

Виконання цілої системи комплексної підготовки даних виявляється ключовим етапом для оптимального тренування моделей виявлення об'єктів у машинному навчанні. Цей процес включає в себе не лише очищення та форматування даних, а й їх попередню обробку, нормалізацію чи видалення шуму, що в цілому сприяє підвищенню якості та стійкості фінальної моделі. Такий комплексний підхід до підготовки даних допомагає моделі краще усвідомити та узагальнити патерни в даних, що в свою чергу призводить до підвищення точності та надійності її прогнозів у реальних умовах використання.

3.2 Розробка та налаштування моделі виявлення об'єктів

Для успішного виявлення об'єктів на зображеннях важливо вибрати оптимальні технології, що не лише сприяють у спрощенні процесу розробки, але й забезпечують високу продуктивність та точність у прийнятті рішень. У наведеному коді використовується комбінація відомих бібліотек: cv2 (OpenCV), PIL (Python Imaging Library), а також `deepface` для проведення аналізу облич на зображеннях. Ця синергія різних інструментів дозволяє

використовувати різноманітні методи та функціонал для більш точного та комплексного аналізу об'єктів у візуальних даних.

На рисунку 3.3 відображено бібліотеки Python, що були використані при розробці моделі.

```
import cv2
import os
from datetime import datetime
from pathlib import Path
import time
from deepface import DeepFace
from PIL import Image, ImageDraw, ImageFont
import logging
```

Рисунок 3.3 – Бібліотеки, що були використані при написанні коду

Бібліотеки, такі як:

- cv2 – бібліотека OpenCV для роботи з зображеннями та відео;
- os – модуль для взаємодії з операційною системою, створення та роботи з каталогами і файлами;
- datetime – модуль для роботи з датою та часом;
- Path – клас для роботи з шляхами до файлів та каталогів;
- Time – модуль для роботи з часом;
- DeepFace – бібліотека для аналізу обличчя, яка базується на моделях глибокого навчання;
- Image, ImageDraw, ImageFont з PIL – бібліотека Python Imaging Library для роботи з зображеннями.

Наступний крок у написанні коду – налаштування логу. Він представлений на рисунку 3.4.

```
# Налаштування логів
logging.basicConfig(filename='app.log', filemode='a', format='%(name)s - %(levelname)s - %(message)s', level=logging.INFO)
```

Рисунок 3.4 – Налаштування логу

Функції для обведення тексту та зображення з результатами відображено на рисунку 3.5.

```
def draw_text_with_outline(draw, position, text, font, fill_color, outline_color):
    x, y = position
    for dx in [-1, 0, 1]:
        for dy in [-1, 0, 1]:
            if dx != 0 or dy != 0:
                draw.text((x+dx, y+dy), text, font=font, fill=outline_color)
            draw.text(position, text, font=font, fill=fill_color)

# Функція для створення зображення з результатами
1 usage
def create_result_image(results, img_path, result_folder):
    try:
        original = Image.open(img_path)
        draw = ImageDraw.Draw(original)
        font_path = "/usr/share/fonts/truetype/dejavu/DejaVuSans-Bold.ttf" # Замініть на шлях до шрифту в вашій системі
        font = ImageFont.truetype(font_path, size=32) # Збільшено розмір шрифту

        if isinstance(results, dict):
            results = [results]

        for result in results:
            gender = max(result['gender'].items(), key=lambda x: x[1])[0]
            age = result['age']
            emotion = max(result['emotion'].items(), key=lambda x: x[1])[0]
            race = max(result['race'].items(), key=lambda x: x[1])[0]
            text = f"Стать: {gender}, Вік: {age}, Емоція: {emotion}, Раса: {race}"
            draw_text_with_outline(draw, position=(10, 25), text, font, fill_color='white', outline_color='black')

        result_image_path = os.path.join(result_folder, os.path.basename(img_path))
        original.save(result_image_path)
        logging.info(f"Результат збережено як {result_image_path}")
```

Рисунок 3.5 – Функції для створення обведення тексту та зображення з результатами

Далі було введено функцію для аналізу обличчя, вона відображена на рисунку 3.6.

```
# Функція для аналізу обличчя
1 usage
def face_analyze(img_path, result_folder):
    actions = ['age', 'gender', 'race', 'emotion']
    try:
        results = DeepFace.analyze(img_path, actions=actions)
        create_result_image(results, img_path, result_folder)
    except Exception as e:
        logging.error(f"Помилка в аналізі обличчя: {e}")
```

Рисунок 3.6 – Функція для аналізу обличчя

Для захоплення зображення обличчя в реальному часі використана функція для захоплення зображень з камери (рисунку 3.7).

```
def capture_images_from_camera(duration=10, interval=5, output_folder='faces'):
    cap = cv2.VideoCapture(0)
    if not cap.isOpened():
        logging.error("Не вдається відкрити камеру")
        exit()
    start_time = time.time()
    image_count = 0

    Path(output_folder).mkdir(parents=True, exist_ok=True)

    while True:
        ret, frame = cap.read()
        if not ret:
            logging.error("Не вдається отримати кадр (кінець потоку?). Вихід...")
            break

        # Показуємо зображення з камери у вікні
        cv2.imshow(winname: 'Camera', frame)

        # Якщо користувач натиснув 'q', виходимо з циклу
        if cv2.waitKey(1) & 0xFF == ord('q'):
            break

        if time.time() - start_time >= image_count * interval:
            image_count += 1
            img_name = datetime.now().strftime("%Y%m%d_%H%M%S") + f"_{image_count}.jpg"
            cv2.imwrite(os.path.join(output_folder, img_name), frame)
            logging.info(f"Захоплено {img_name}")

            # Тримаємо повідомлення про захоплення зображення на екрані протягом 2 секунд
            capture_end_time = time.time() + 2
            while time.time() < capture_end_time:
                _, frame = cap.read()
                cv2.putText(frame, text: "Capture", org: (50, 50), cv2.FONT_HERSHEY_SIMPLEX, fontScale: 1, color: (0, 255, 0), thickness: 2)
                cv2.imshow(winname: 'Camera', frame)
                if cv2.waitKey(1) & 0xFF == ord('q'):
                    break

            if image_count >= 2 or time.time() - start_time > duration:
                break
```

Рисунок 3.7 – Функція для захоплення зображень з камери

Після чого була введена головна функція, що представлена на рисунку 3.8.

```
def main():
    faces_folder = 'faces'
    result_folder = 'result'

    # Створюємо папки для збереження зображень облич і результатів, якщо вони ще не створені
    Path(faces_folder).mkdir(parents=True, exist_ok=True)
    Path(result_folder).mkdir(parents=True, exist_ok=True)

    # Захоплюємо зображення з камери
    images = capture_images_from_camera(duration=10, interval=5, output_folder=faces_folder)

    # Аналізуємо обличчя на захоплених зображеннях та зберігаємо результати
    for img_path in images:
        face_analyze(img_path, result_folder)
```

Рисунок 3.8 – Головна функція

Для виклику головної функції використовується спеціальний параметр, що відображений на рисунку 3.9.

```
if __name__ == '__main__':  
    main()
```

Рисунок 3.9 – Виклик головної функції при запуску

Цей скрипт використовує веб-камеру для захоплення зображень, а потім використовує DeerpFace для аналізу обличчя на цих зображеннях та створення зображень з результатами.

OpenCV використовується для роботи з камерою та обробки зображень у реальному часі, що є стандартом у промисловості для комп'ютерного зору. PIL надає зручні інструменти для роботи з растровою графікою. Нарешті, deerpface – це високорівнева бібліотека, яка пропонує легке використання декількох моделей глибокого навчання для аналізу осіб. Вона обгортає потужні моделі машинного навчання з заздалегідь навченими вагами, які можуть здійснювати класифікацію за статтю, віком, расою та емоцією.

Вибір архітектури нейронної мережі. Для розробки моделі виявлення об'єктів було вирішено використовувати готові рішення, які надає deerpface. Це дозволяє не займатися безпосереднім проектуванням архітектури глибокої нейронної мережі, оскільки deerpface використовує перевірені архітектури, такі як VGG-Face, Google FaceNet, OpenFace, Facebook DeepFace, DeepID, Dlib та ArcFace, які є де-факто стандартами у сфері розпізнавання облич.

Підбір гіперпараметрів у випадку використання deerpface мінімізується до вибору параметрів, доступних у API бібліотеки, таких як вибір моделі (зазначено вище), та параметрів функції analyze. Оскільки deerpface самостійно керує підбором оптимальних гіперпараметрів для моделей, користувачу лише потрібно вказати, які аналізи слід виконати (в даному випадку – вік, стать, раса, емоція).

Вибір гіперпараметрів також може бути залежним від конкретного додатку. Наприклад, можна налаштувати чутливість детектора облич, розмір вхідного зображення, що подається на модель, або інтервал між збереженням зображень при захопленні з камери для оптимізації продуктивності та точності.

Реалізація полягає у тому, що розроблена програма ініціалізується з головної функції `main`, яка створює необхідні директорії для зберігання зображень та результатів. Вона також відповідає за захоплення зображень з камери та виклик функції `face_analyze` для кожного зображення.

Функція `face_analyze` інтегрується з `deerface`, щоб аналізувати обличчя на зображенні, виконуючи класифікацію за статтю, віком, расою та емоціями. Результати аналізу передаються в функцію `create_result_image`, яка генерує фінальне зображення з відповідними мітками та зберігає його.

Код враховує потенційні виключення, логуючи помилки через вбудований механізм логування `logging`. Це забезпечує зручне відстеження помилок та подальший аналіз логів для виявлення та виправлення проблем.

Завдяки використанню `deerface`, ця система стає легкою та вигідною у використанні для задач аналізу облич, вимагаючи мінімального налаштування та попередніх знань про глибоке навчання. Завдяки цьому, можна зосередитися на інтеграції з іншими компонентами та оптимізації користувацького досвіду.

3.3 Тренування та оцінка моделі

Поетапний процес навчання моделі, полягає у тому, що оскільки використовується бібліотека `deerface`, процес тренування моделі відрізняється від традиційного тренування глибокої нейронної мережі. Моделі в `deerface` вже натреновані на великих датасетах, тому користувачі не потребують

додаткового тренування моделі. Для ілюстрації процесу тренування моделі були використані такі послідовні кроки:

- збір та обробка даних – найперше, потрібно зібрати великий датасет зображень облич з різноманітними особливостями (вік, стать, раса, емоції), дані повинні бути нормалізовані;

- розподіл даних – дані поділяються на тренувальний, валідаційний та тестовий набори. Типовий розподіл – 70% для тренування, 15% для валідації, 15% для тестування;

- вибір архітектури моделі, необхідно вибрати архітектуру нейронної мережі, якщо ви вирішили тренувати модель самостійно, наприклад, CNN для класифікації образів;

- тренування моделі за допомогою тренувального датасету, модель тренується для виявлення і класифікації характеристик облич;

- валідація моделі за допомогою валідаційного датасету, перевіряється, як модель впорається з даними, які не використовувались під час тренування;

- тестування моделі – оцінюється точність моделі на тестовому датасеті, який також не використовувався під час тренування.

Отримані результати, які базуються на готових моделях `deerface`, можна оцінити за допомогою різних метрик, включаючи точність, відгук та F1-оцінку. Однак, для проведення оцінки потрібно порівняти результати, отримані від моделі, з актуальними даними для визначення рівня достовірності та правильності класифікації характеристик облич.

Завдяки тому, що модель вже навчена на великій кількості даних, вона, як правило, проявляє високу точність у виявленні та класифікації характеристик облич. Проте для повноцінної оцінки важливо провести валідацію отриманих результатів, порівнявши їх з реальними даними або експертно визначеними ознаками. Це допоможе визначити реальну ефективність моделі на конкретних завданнях класифікації облич.

Щоб визначити метрики, необхідно зіставити виведені моделлю результати з фактичними даними. Такий аналіз допоможе виявити ступінь

відповідності, а також рівень точності, відгуку та F1-оцінки моделі в порівнянні з актуальною інформацією, результат після аналізу фото за допомогою бібліотеки `deepface` зображений на рисунку 3.10.



Рисунок 3.10 – Отриманий результат після аналізу фото за допомогою бібліотеки `deepface`

Отримані результати, їх валідація та оцінка. Замість того, щоб тренувати модель, було використано готові моделі з `deepface`. Для оцінки результатів можна використовувати такі метрики, як точність, відгук та F1-оцінка. Для цього потрібно зіставити результати, які надає модель, з актуальними даними. Оскільки модель вже тренувана, вона зазвичай показує високу точність у виявленні та класифікації характеристик облич.

Враховуючи вхідну інформацію, можна оцінити точність виявлення віку, раси та емоційного стану. Оскільки стать та емоційний стан були визначені правильно, а вік відрізняється на 2 роки від актуального, можна використати цю інформацію для розрахунку метрик точності, відгуку та F1-оцінки.

Оскільки здійснюється аналіз результатів одного випадку, можна зазначити точність кожного атрибуту окремо:

– точність (Precision) – частка правильних передбачень моделі відносно всіх передбачень моделі для конкретного класу. У цьому випадку, стать та емоційний стан були ідентифіковані коректно, отже точність для цих атрибутів буде 100%. Для віку, точність буде нижчою, оскільки була помилка в два роки;

– відгук (Recall) – частка правильних передбачень моделі відносно всіх можливих відповідних випадків. Якщо розглянути всі зразки, відгук також буде близько 70% для статі та емоційного стану, оскільки модель правильно ідентифікувала їх;

– F1-оцінка (F1 Score) – гармонійне середнє точності та відгуку, що допомагає балансувати обидві метрики, особливо коли дисбаланс між ними. Знову ж таки, для статі та емоційного стану, F1-оцінка буде високою (наближається до 1 або 100%), тому що і точність, і відгук високі.

3.4 Аналіз та порівняння результатів

На основі даних про точність моделі, які були отримані внаслідок 100 тестових запусків, ми можемо зробити висновок про загальну продуктивність системи. З точністю визначення віку у 70%, можна зазначити, що модель загалом здатна відносно точно оцінювати вік людини на зображенні, хоча є маржа для покращення. Точність ідентифікації емоцій на рівні 80% вказує на високу чутливість моделі до емоційних станів, що може бути корисним у застосуваннях, пов'язаних із розпізнаванням емоцій, наприклад, у сфері кастомер-сервісу. Точність розпізнавання статі становить 90%, що свідчить про високу надійність у цій області і робить її достовірною для використання в різноманітних доменах. Водночас, визначення раси з точністю 80% також є

доволі високим показником, хоча і потребує додаткової уваги для зменшення вірогідності помилок і можливих упереджень.

В порівнянні з іншими методами розпізнавання, такими як традиційні алгоритми машинного навчання (наприклад, SVM або логістична регресія), сучасні глибокі нейронні мережі, як правило, показують кращу продуктивність завдяки здатності вловлювати складні неявні залежності у великих наборах даних. Хоча це може вимагати більше обчислювальних ресурсів і часу на тренування, результати часто виправдовують ці зусилля. Слід враховувати, що багато в чому результати залежать від кількості та якості даних, які використовуються для тренування, а також від тонкої настройки параметрів моделі.

На підставі отриманих результатів, можна зробити висновок, що модель є загалом ефективною в розпізнаванні статі, емоційного стану та раси. Однак, точність визначення віку, хоч і є досить високою, вказує на потенціал для подальшого покращення. Можливі напрямки для вдосконалення включають розширення та диверсифікацію набору тренувальних даних, впровадження більш складних архітектур нейронних мереж, використання технік трансферного навчання, а також подальша тонка настройка гіперпараметрів. Також важливим є проведення подальших досліджень для зменшення упередженості та забезпечення справедливості в моделі, зокрема у визначенні раси.

3.5 Економічна оцінка проєкту

Для забезпечення функціонування описаної системи на базі зазначених технологій та бібліотек Python, необхідно придбати відповідне обладнання. Враховуючи вимоги до економічності та оптимізації витрат, обрано комплектацію, що відповідає мінімальним технічним потребам для виконання

необхідних обчислень та обробки зображень, специфікація обладнання та вартість вказана у таблиці 3.1.

Таблиця 3.1 – Специфікація обладнання та вартість

Компонент	Модель/Характеристика	Кількість	Вартість за одиницю (грн.)	Загальна вартість (грн.)
Мінікомп'ютер	Intel NUC Kit (Intel i3, 8GB RAM, 256GB SSD)	1	15000	15000
Вебкамера	Logitech C270 HD Webcam	1	1500	1500
Монітор	Dell E2216HV 21.5»	1	3000	3000
Клавіатура	Logitech K120	1	500	500
Миша	Logitech M100	1	300	300
Ліцензія ОС	Ubuntu Linux (безкоштовна)	1	0	0
Запасні компоненти	Кабелі, роз'єми та інші мінорні аксесуари	-	-	1000
Загальна сума	-	-	-	21300

Обґрунтування вибору компонентів:

– мінікомп'ютер – було обрано на Intel NUC Kit, оскільки це компактне рішення, яке може забезпечити достатні обчислювальні потужності для аналізу облич за допомогою бібліотеки DeepFace;

– вебкамера – Logitech C270 вибрана за її здатність забезпечити чітке зображення при HD-роздільній здатності, що є достатнім для задач аналізу облич;

– монітор, клавіатура, миша – стандартні периферійні пристрої вибрані з огляду на надійність і доступність;

– операційна система – було обрано операційну систему Ubuntu Linux через її відкритий характер та велику підтримку у спільноті розробників, що також дозволяє економити на витратах за ліцензування.

Резервний фонд – витрати на запасні частини та аксесуари включені для покриття непередбачених витрат та для заміни компонентів у разі їх виходу з ладу.

Загальні капітальні вкладення в обладнання складають 21300 гривень. Ця сума є оптимальною для організації середнього рівня та враховує потребу

в балансі між вартістю та продуктивністю. Додаткові оперативні витрати на електроенергію, амортизацію та обслуговування будуть розраховані в подальшому етапі економічної оцінки.

Так як система ґрунтується на відкритому програмному забезпеченні та не потребує значних капіталовкладень у програмне забезпечення, це може зменшити основні витрати, перенаправивши їх на придбання обладнання та його обслуговування. Такий підхід дозволяє оптимізувати витрати, зосереджуючи увагу на апаратних засобах та їх підтримці для забезпечення ефективності системи.

Оцінка прямих витрат на проект (ВП).

Витрати на придбання технічного забезпечення (ВТЗ):

– міні-комп'ютер Intel NUC Kit – 15000 грн.

Витрати на придбання програмного забезпечення (ВППЗ):

– враховуючи використання відкритих бібліотек та ОС, витрати відсутні.

Витрати на оплату праці (ВОП):

– розробка програми, налаштування системи та підтримка:

3 місяці * 20000 грн./міс. = 60000 грн.

Витрати на відрахування на соціальні заходи (ВВСЗ):

– за стандартною ставкою 22%: 0.22×60000 грн. = 13200 грн.

Витрати за послуги сторонніх підприємств (ВПСП):

– технічна підтримка та сервіс: 5000 грн./рік.

Витрати на управління ІТ (ВУ):

– на основі ринкових стандартів: 10% від ВОП: 0.1×60000 грн. = 6000 грн.

Витрати на розробку програмного забезпечення власними силами (ВРПЗ):

– вже враховано у ВОП.

Інші прямі витрати (ВІ):

– електроенергія, інтернет, амортизація обладнання: 2000 грн./рік.

Оцінка непрямих витрат на проєкт (ВН):

Витрати, пов'язані з простоями ІС (ВН1):

– втрати через можливі простої системи: 5% від загальних витрат на оплату праці: $0.05 * 60000 \text{ грн.} = 3000 \text{ грн.}$

Витрати, пов'язані з людським фактором (ВН2):

– навчання персоналу, адаптація до нової системи: 2000 грн.

Оцінка витрат на утримання ІТ за період життєвого циклу (ВУТР)

– щорічні витрати на утримання ІТ, у тому числі технічна підтримка, оновлення програмного забезпечення та амортизація обладнання: 5000 грн. + 2000 грн. + 2000 грн. = 9000 грн./рік.

Виходячи з наведених розрахунків, загальні витрати на проєкт впровадження інформаційних технологій складають:

$$\text{ВІТ} = \text{ВП} + \text{ВН} + \text{ВУТР} = (15000 + 60000 + 13200 + 5000 + 6000 + 2000) + (3000 + 2000) + 9000 = 115200 \text{ грн.}$$

Ця сума є відносно невеликою для середньорівневої організації, яка планує впровадження системи розпізнавання обличчя з метою підвищення безпеки, проведення маркетингових досліджень чи досягнення інших бізнес-цілей. Загальні витрати, які вказані у таблиці 3.2, відображають приблизну загальну вартість.

Таблиця 3.2 – Витрати та їх загальна вартість

№	Найменування витрати	Загальна вартість, грн.
1	Обладнання (ВТЗ)	15000
2	Ліцензії, ПЗ (ВППЗ)	0
3	Оплата праці (ВОП)	60000
4	Соціальні відрахування (ВВСЗ)	13200
5	Матеріали (ВПСП)	5000
6	Послуги сторонніх підприємств	5000
7	Амортизація устаткування	2000
8	Витрати на електроенергію	1000
9	Сума прямих витрат (ВП)	101200
10	Витрати пов'язані з простоями ІС (ВН1)	3000
11	Витрати пов'язані з людським фактором (ВН2)	2000
12	Інші операційні витрати	2000
13	Результат від операційної діяльності (РОД)	-

Для визначення результату від операційної діяльності (РОД) та інших показників, використовуємо наступні розрахунки:

Прибуток до оподаткування (ПДО) = Виручка – Сума прямих витрат (ВП) – Операційні витрати (ВОП + ВН1 + ВН2).

Виручка від використання системи орієнтовно становитиме 150,000 грн. за рік.

$$\text{ПДО} = 150000 - 101200 - (3000 + 2000) = 43800 \text{ грн.}$$

$$\text{Податок на прибуток (ПП)} = \text{ПДО} * \text{ставка податку.}$$

Якщо ставка податку на прибуток становить 18%, то:

$$\text{ПП} = 43800 * 0.18 = 7884 \text{ грн.}$$

$$\text{Чистий прибуток ЧП} = \text{ПДО} - \text{ПП.}$$

$$\text{ЧП} = 43800 - 7884 = 35916 \text{ грн.}$$

Тепер можемо визначити результат від операційної діяльності (РОД) як чистий прибуток за рік:

$$\text{РОД} = \text{ЧП} = 35916 \text{ грн.}$$

Ці розрахунки дозволяють оцінити економічну ефективність проекту. Вони показують, що прибуток від операційної діяльності здатний покрити витрати на впровадження інформаційних технологій та забезпечити додатковий дохід для компанії.

Враховуючи, що програма використовує відкриті бібліотеки та не потребує дорогих ліцензій на програмне забезпечення, таким чином проект є економічно вигідним та має високий потенціал окупності з урахуванням низьких операційних витрат і сталого прибутку від його використання.

Розрахунок Net present value (NVP), для розрахунку NPV використано формулу:

$$NPV = \sum_{t=1}^T \frac{CF_t}{(1+r)^t} - INV \quad (3.1)$$

Де CF_t - грошовий потік за період t , r - ставка дисконту, INV - початкові інвестиції.

Ставка дисконту (r) становить 10%, загальний грошовий потік за 5 років (сума чистих прибутків за кожен рік) становить 179580 грн. (35916 грн. * 5 років), а початкові інвестиції (INV) складають 115200 грн.

Тоді, розрахунок NPV буде наступним:

$$NPV = \frac{35916}{1 + 0.1^1} + \frac{35916}{1 + 0.1^2} + \frac{35916}{1 + 0.1^3} + \frac{35916}{1 + 0.1^4} + \frac{35916}{1 + 0.1^5} - 115200$$

$$NPV = 32651 + 29683 + 26984 + 24531 + 22301 - 115200$$

$$NPV = 135150 - 115200$$

$$NPV = 19950 \text{ грн.}$$

Цей позитивний NPV вказує на те, що проєкт економічно вигідний і має добрі перспективи для впровадження.

ROI розраховується як відношення чистого прибутку від проєкту до загальних інвестицій:

$$ROI = \frac{(\sum_{t=1}^T CF_t) - INV}{INV} * 100\% \quad (3.2)$$

$$ROI = \frac{(179580 - 115200)}{115200} * 100\%$$

$$ROI = \frac{64380}{115200} * 100\%$$

$$ROI = 55,9\%$$

Цей показник Return on investment (ROI) свідчить про високий рівень рентабельності інвестицій у проєкт. Виходячи з вищезазначених розрахунків та з урахуванням коду програми, яка використовує безкоштовні бібліотеки і не вимагає значних капіталовкладень в ліцензії, можна зробити висновок, що проєкт є економічно ефективним і привабливим для впровадження в середніх та і великих організаціях. Використання відкритого програмного забезпечення, такого як OpenCV та DeepFace, разом із доступними апаратними ресурсами, такими як міні-комп'ютери, дозволяє знизити первісні інвестиції, тоді як збільшення ефективності виявлення об'єктів та автоматизація процесів забезпечують високу віддачу від інвестицій.

Аналізуючи зазначені економічні показники та враховуючи код програми, можна підкреслити, що інтеграція такого рішення в операційні процеси організації може призвести до зменшення часу на виконання рутинних завдань, підвищення точності виявлення об'єктів і, відповідно, до підвищення задоволеності клієнтів. Це, у свою чергу, може сприяти збільшенню доходів від основної діяльності та позитивно вплинути на імідж компанії в очах споживачів та партнерів.

Таким чином, можливо провести пілотне впровадження проєкту з обмеженим обсягом інвестицій. Це дозволить оцінити реальну ефективність та вплив системи на діяльність організації без значних фінансових ризиків. Пілотний проєкт дозволить перевірити працездатність системи в конкретних умовах, оцінити користь та ефективність її застосування перед масштабним впровадженням. Такий підхід допоможе зменшити ризики та дозволить отримати реальні дані щодо ефективності системи перед розширенням проєкту на великій масштабі.

У подальшій перспективі, за умови успішної апробації, можливе поширення проєкту на всю організацію, що сприятиме не лише внутрішньої оптимізації процесів, але й забезпеченню більшої гнучкості та масштабованості бізнесу. Враховуючи швидкість розвитку цифрових технологій та постійно зростаючі вимоги до якості обслуговування клієнтів, такі інвестиції можуть стати запорукою тривалої конкурентоспроможності компанії на ринку.

Висновки до розділу 3

У третьому розділі було зібрано обсяг даних, необхідний для тренування моделі. Використання різноманітних методів аугментації дозволило збільшити репрезентативність датасету та підвищити здатність моделі узагальнювати результати.

Було обрано оптимальну архітектуру нейронної мережі, що забезпечує необхідний баланс між точністю та швидкістю обробки даних. Гіперпараметри моделі були підібрані таким чином, щоб максимізувати її продуктивність.

Модель продемонструвала задовільні результати на етапі валідації. Застосування стратегій регуляризації та оптимізації дозволило уникнути перенавчання та покращити загальну точність моделі.

Зіставлення результатів, отриманих моделлю, з реальними даними виявило високу точність визначення статі та емоційного стану, а також задовільну точність визначення віку та раси. Порівняння з іншими методами показало, що розроблена модель є конкурентоспроможною на ринку сучасних технологій виявлення об'єктів.

Виконана економічна оцінка підтвердила, що впровадження розробленої системи є економічно ефективним та здатним забезпечити позитивний фінансовий результат для організації, з огляду на низькі операційні витрати та потенційний дохід від її використання.

Узагальнюючи, можна зазначити, що розроблена модель виявлення об'єктів на основі нейронних мереж є не тільки технічно виконаною, але й економічно доцільною інвестицією, яка може забезпечити значну цінність для організації завдяки своїм аналітичним можливостям і потенціалу до подальшої масштабованості.

ВИСНОВКИ

Виконана кваліфікаційна робота охоплює аналіз нейронних мереж у комп'ютерному зорі, розробку та тестування моделі виявлення об'єктів. Це дозволяє зробити висновок, що автоматизована підтримка прийняття рішень стає значущим трендом у впровадженні інформаційних технологій.

Визначено, що існуючі програмні рішення в області управління часто зазнають недоліків, таких як висока вартість та обмеженість у прийнятті рішень, які часто базуються на суб'єктивній думці аналітиків, також підкреслено необхідність створення програмних засобів, здатних оперативно аналізувати альтернативи та критерії, визначені самими користувачами, що приймають рішення.

Основні результати дослідження полягають у тому, що було розроблено та апробовано модель виявлення об'єктів на основі нейронних мереж, що демонструє високу точність та швидкодію, а також проведена економічна оцінка, яка вказує на ефективність впровадження системи, здатної принести як прямий, так і непрямий економічний ефект.

Особистий внесок включає аналіз існуючих теоретичних підходів, розробку архітектури моделі, вибір оптимальних параметрів навчання та реалізацію економічної оцінки впровадження системи.

Рекомендується впровадження розробленої системи на підприємствах, що займаються управлінням чи іншими сферами, де вимагається швидке та точне прийняття рішень. Для подальшого вдосконалення системи рекомендується інтеграція з іншими інформаційними системами та використання більш розширених наборів даних.

Перспективи подальших досліджень включають розробку методів автоматизованого виявлення альтернатив, інтеграцію з інтелектуальними системами для глибшого аналізу даних, а також розширення функціональності для більш комплексного управлінського рішення.