

МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ
Західноукраїнський національний університет
Факультет комп'ютерних інформаційних технологій
Кафедра інформаційно-обчислювальних систем і управління

Божагора Микола Романович

**Модуль розпізнавання статі на основі моделі глибокого
навчання InceptionV3 / Module for gender recognition based on the
InceptionV3 deep learning model**

Спеціальність 122 – Комп'ютерні науки
Освітньо-професійна програма – Комп'ютерні науки

Кваліфікаційна робота

Виконав студент групи КН-42
Божагора М.Р.

Науковий керівник:
викл.
Дорош В. І.

Кваліфікаційну роботу допущено до
захисту
«__» _____ 2025 р.

В.о. завідувача кафедри
_____ Н.М. Васильків

Тернопіль – 2025

Факультет комп'ютерних інформаційних технологій
Кафедра інформаційно-обчислювальних систем і управління
Освітній ступінь «бакалавр»
спеціальність 122 – Комп'ютерні науки
освітньо-професійна програма – Комп'ютерні науки

ЗАТВЕРДЖУЮ
В.о. завідувача кафедри
Н.М. Васильків
«____» _____ 2024 р.

ЗАВДАННЯ
НА КВАЛІФІКАЦІЙНУ РОБОТУ СТУДЕНТУ
Божагора Микола Романович
(прізвище, ім'я, по батькові)

1. Тема кваліфікаційної роботи: Модуль розпізнавання статі на основі моделі глибокого навчання InceptionV3 / Module for gender recognition based on the InceptionV3 deep learning model

керівник роботи викл., Дорош В. І.

(прізвище, ім'я, по батькові, науковий ступінь, вчене звання)

затверджені наказом по університету від 20 грудня 2024 р. № 938.

2. Строк подання студентом закінченої кваліфікаційної роботи 25 травня 2025 р.

3. Вихідні дані до кваліфікаційної роботи: завдання на кваліфікаційну роботу студента, наукові статті, технічна література.

4. Основні питання, які потрібно розробити:

– проаналізувати предметну область і наявні рішення автоматичного розпізнавання статі за зображеннями;

– дослідити сучасні методи глибокого навчання для аналізу зображень і виконати їх порівняльний аналіз;

– розробити архітектуру модуля розпізнавання статі на базі InceptionV3 та обґрунтувати вибір гіперпараметрів;

– реалізувати алгоритм розпізнавання статі й визначити основні метрики оцінки його якості;

– провести експериментальні дослідження (підготовка даних, навчання, перевірка, тестування) на наборі CelebA;

– порівняти одержані результати з іншими моделями глибокого навчання та оцінити перспективи подальшого розвитку.

5. Перелік графічного матеріалу в роботі:

– архітектура згорткової нейронної мережі InceptionV3. Алгоритм розпізнавання статі на основі глибокого навчання.

6. Консультанти розділів кваліфікаційної роботи

Розділ	Прізвище, ініціали та посада консультанта	Підпис, дата	
		Завдання видав	Завдання прийняв

7. Дата видачі завдання 20 грудня 2024 р.

КАЛЕНДАРНИЙ ПЛАН

№ з/п	Назва етапів кваліфікаційної роботи	Строк виконання етапів кваліфікаційної роботи	Примітка
	Затвердження теми кваліфікаційної роботи, ознайомлення з літературними джерелами та складання плану роботи	до 01.01. 2025 р.	
	Написання 1 розділу кваліфікаційної роботи	до 01.02. 2025 р.	
	Написання 2 розділу кваліфікаційної роботи	до 01.04.2025 р.	
	Написання 3 розділу кваліфікаційної роботи	до 01.05. 2025 р.	
	Представлення попереднього варіанту кваліфікаційної роботи, перевірка та внесення змін керівником	до 15.05.2025 р.	
	Опрацювання зауважень та представлення завершеного варіанту кваліфікаційної роботи. Підготовка супроводжуючих документів	до 25.05.2025 р.	
	Перевірка кваліфікаційної роботи на оригінальність тексту	до 30.05.2025 р.	
	Оформлення кваліфікаційної роботи та отримання допуску до захисту	до 10.06.2025 р.	
	Подання кваліфікаційної роботи до захисту на засіданні атестаційної комісії	до 10.06. 2025 р.	

Студент _____ Божагора М.Р.
(підпис) (прізвище та ініціали)

Керівник кваліфікаційної роботи _____ Дорош В. І.
(підпис) (прізвище та ініціали)

АНОТАЦІЯ

Кваліфікаційна робота на тему «Модуль розпізнавання статі на основі моделі глибокого навчання InceptionV3» на здобуття освітнього ступеня «бакалавр» зі спеціальності 122 «Комп'ютерні науки» освітньої програми «Комп'ютерні науки» виконана обсягом 59 сторінок і містить 14 ілюстрацій, 3 таблиці, 3 додатки та 23 використані джерела.

Метою роботи є розробка та оптимізація модуля автоматичного визначення статі людини за зображенням обличчя з використанням архітектури InceptionV3.

Методами дослідження обрано аналіз наукових джерел, попередню обробку та аугментацію набору CelebA, переднавчання й тонке налаштування згорткової мережі, а також оцінку продуктивності за метриками accuracy, precision, recall та F1-score.

У результаті створено програмний засіб, що досягає точності 94,3 % на тестовій вибірці й демонструє F1-міру 0,94, забезпечуючи баланс між якістю та обчислювальною ефективністю.

Результати можуть бути застосовані в системах безпеки, маркетинговій аналітиці й біометричних сервісах.

Ключові слова: INCEPTIONV3, ГЛИБИННЕ НАВЧАННЯ, КЛАСИФІКАЦІЯ, РОЗПІЗНАВАННЯ СТАТІ, CELEBA, ЗГОРТКОВА НЕЙРОННА МЕРЕЖА.

ANNOTATION

The bachelor's thesis "Module for gender recognition based on the InceptionV3 deep learning model" comprises 59 pages, 14 figures, 3 tables, 3 appendices and 23 references.

Its aim is to develop and optimise an automatic gender-recognition module that processes face images using the InceptionV3 convolutional neural network.

The study employs literature analysis, CelebA data preprocessing and augmentation, transfer learning with fine-tuning, and performance evaluation through accuracy, precision, recall and F1-score.

The resulting software attains 94.3 % accuracy and an F1-score of 0.94 on the test set, offering a sound trade-off between accuracy and computational cost.

The findings are applicable to security systems, marketing analytics and biometric services.

Keywords: INCEPTIONV3, DEEP LEARNING, CLASSIFICATION, GENDER RECOGNITION, CELEBA, CONVOLUTIONAL NEURAL NETWORK.

ЗМІСТ

Вступ.....	7
1 Аналіз предметної області та існуючих рішень.....	10
1.1 Поняття класифікації статі на основі зображень	10
1.2 Сучасні методи глибокого навчання для аналізу зображень.....	13
1.3 Порівняльний аналіз методів розпізнавання статі на основі глибокого навчання	15
1.4 Вибір перспективного шляху і постановка задачі дослідження	18
2 Проєктування модуля розпізнавання статі на основі глибокого навчання.....	21
2.1 Опис моделі InceptionV3 та її особливості	21
2.2 Вибір архітектури моделі та гіперпараметрів	22
2.3 Метрики оцінки якості розпізнавання	25
2.4 Алгоритм розпізнавання статі на основі глибокого навчання	26
2.5 Архітектура реалізації модуля	29
3 Програмна реалізація та тестування модуля	33
3.1 Дослідження даних.....	33
3.2 Поділ набору даних на навчання, перевірку та тестування	35
3.3 Попередня обробка зображень.....	37
3.4 Побудова моделі розпізнавання статі та її оцінка	40
Висновки	44
Список використаних джерел	46
Додаток А Код реалізації модуля	49
Додаток Б Результати оцінки моделі.....	54
Додаток В Апробація отриманих результатів	56

ВСТУП

Актуальність теми. У сучасному світі штучний інтелект та комп'ютерний зір відіграють ключову роль у розробці автоматизованих систем аналізу зображень. Однією з важливих задач у цій сфері є розпізнавання статі на основі зображень, що знаходить застосування у безпеці, маркетингу, аналітиці поведінки користувачів та персоналізованих сервісах. Автоматичне визначення статі може бути корисним для покращення взаємодії між людиною та машиною, створення адаптивних інтерфейсів та оптимізації бізнес-процесів.

Методи глибокого навчання дозволяють досягти високої точності у розпізнаванні статі, використовуючи великі обсяги даних та складні нейромережеві архітектури. Зокрема, модель InceptionV3 є однією з найбільш ефективних для обробки зображень, оскільки використовує глибокі згорткові шари для виділення ключових ознак обличчя. Проте, одним із викликів при створенні таких систем є обробка нерівномірно розподілених вибірок, що може впливати на якість класифікації.

Набір даних CelebA, який використовується у цьому дослідженні, містить понад 200 тисяч зображень із детальними анотаціями. Однак, одним із його недоліків є дисбаланс між кількістю зображень різних статей, що може призводити до зміщення у навчанні моделі. Тому важливим є не тільки вибір ефективної архітектури, але й оптимізація процесу підготовки даних та застосування методів балансування.

Розробка модуля розпізнавання статі з використанням глибоких нейронних мереж є актуальним завданням, яке може сприяти покращенню точності та стабільності таких систем. Дослідження та порівняння моделей дозволить визначити оптимальні параметри навчання та запропонувати підхід, що забезпечить високу продуктивність у реальних умовах застосування.

Сучасний розвиток штучного інтелекту та глибокого навчання значно розширює можливості автоматичної обробки зображень, що знаходить широке застосування у різних сферах, таких як безпека, медицина, маркетинг та

персоналізовані системи. Одним із важливих завдань комп'ютерного зору є розпізнавання статі на основі зображень, що використовується для ідентифікації користувачів, аналізу поведінки клієнтів та інших практичних застосувань.

Завдання автоматичного розпізнавання статі на основі зображень потребує використання потужних алгоритмів глибокого навчання, які здатні ефективно аналізувати особливості обличчя. Однією з найбільш продуктивних архітектур для обробки зображень є InceptionV3, яка демонструє високу точність та ефективність у багатьох задачах класифікації.

Попередні дослідження показують, що існуючі моделі можуть мати певні обмеження, пов'язані із недостатнім узагальненням результатів або залежністю від балансування наборів даних. Тому важливим завданням є не лише вибір ефективної моделі, але й оптимізація процесу її навчання та оцінка її продуктивності за допомогою відповідних метрик.

Набір даних CelebA, що використовується у цій роботі, містить велику кількість зображень із розміткою, що дозволяє ефективно навчати та тестувати моделі класифікації статі. Однак, у цьому наборі даних спостерігається певний дисбаланс між класами, що може впливати на якість прогнозування.

Отже, розробка та оптимізація модуля розпізнавання статі на основі глибокого навчання є актуальним завданням, що дозволяє покращити точність класифікації, зменшити похибки прогнозування та розширити можливості застосування таких моделей у реальних сценаріях.

Отже, метою роботи є розробка модуля розпізнавання статі на основі глибокої нейронної мережі InceptionV3, проведення його оптимізації та оцінка ефективності за допомогою відповідних метрик точності класифікації.

Завдання роботи:

1. Провести аналіз предметної області та існуючих рішень у сфері автоматичного розпізнавання статі на основі зображень.
2. Дослідити сучасні методи глибокого навчання, що використовуються для аналізу зображень, та здійснити порівняльний аналіз їх ефективності.

3. Розробити архітектуру модуля розпізнавання статі на основі моделі InceptionV3 та обґрунтувати вибір її гіперпараметрів.

4. Реалізувати алгоритм розпізнавання статі та визначити основні метрики оцінки його якості.

5. Виконати експериментальне дослідження, що включає обробку даних, навчання, перевірку та тестування моделі на наборі CelebA.

6. Провести порівняння отриманих результатів із іншими моделями глибокого навчання та оцінити перспективи подальшого розвитку дослідження.

Об'єктом роботи є процес автоматичного розпізнавання статі на основі зображень із використанням методів глибокого навчання.

Предметом роботи є модель глибокого навчання InceptionV3, методи її навчання, оптимізації та оцінки якості класифікації статі.

Методи дослідження включають аналіз літературних джерел та існуючих рішень у сфері комп'ютерного зору, застосування методів глибокого навчання для класифікації зображень, використання бібліотек TensorFlow та Keras для розробки нейромережевої архітектури, а також емпіричне тестування моделі на основі набору даних CelebA. Для оцінки ефективності моделі використовуються кількісні метрики, такі як accuracy, precision, recall, F1-score, а також аналіз балансування вибірки та впливу різних параметрів навчання на кінцевий результат.

Апробація результатів дослідження здійснена в межах студентської науково-практичної конференції «Інтелектуальні інформаційні технології в прикладних дослідженнях» (ІІТАР-2025), яка відбулася в місті Тернополі 27–29 травня 2025 року.

1 АНАЛІЗ ПРЕДМЕТНОЇ ОБЛАСТІ ТА ІСНУЮЧИХ РІШЕНЬ

1.1 Поняття класифікації статі на основі зображень

Класифікація статі на основі зображень є одним із важливих завдань у галузі комп'ютерного зору. Це завдання передбачає автоматичне визначення статі людини (чоловік чи жінка) за допомогою аналізу цифрових зображень, переважно портретних або облич. Дане завдання широко застосовується у багатьох сферах, таких як безпека, реклама, маркетинг, соціальні дослідження та біометрична ідентифікація.

Перші спроби автоматизованого визначення статі за зображенням з'явилися ще в 90-х роках XX століття і базувалися на традиційних алгоритмах машинного навчання з використанням ручних ознак. Однак із розвитком технологій зросла і складність задачі, що спонукало дослідників звернути увагу на більш складні й ефективні алгоритми, зокрема нейронні мережі.

Задача класифікації статі, попри видиму простоту, має свої особливості й складності. Людські обличчя характеризуються значною варіативністю, пов'язаною з віком, етнічною приналежністю, освітленням, кутом зйомки та іншими чинниками. Саме тому точність алгоритмів розпізнавання може суттєво змінюватися в залежності від якості зображень та вибору ознак для аналізу.

Типовий процес розпізнавання статі за зображенням включає кілька основних етапів: виявлення та локалізацію обличчя на зображенні, вилучення інформативних ознак, навчання моделі для класифікації та, нарешті, безпосередньо класифікацію зображення. Цей процес є стандартним і використовується більшістю сучасних систем.

Для досягнення високої точності класифікації велике значення має правильний вибір ознак, що характеризують обличчя. До таких ознак можуть належати як загальні характеристики (наприклад, форма обличчя, очей, губ, лінії волосся), так і більш специфічні ознаки, що отримуються шляхом складних обчислень і аналізу текстурних або геометричних параметрів.

В останні роки значно зросла роль автоматичних методів генерації ознак завдяки розвитку глибокого навчання. Глибокі нейронні мережі дозволяють автоматично виділяти найрелевантніші ознаки для класифікації статі, що значно підвищує ефективність цього процесу.

Загалом алгоритми класифікації статі поділяються на два основні класи: класичні підходи, що базуються на ручних ознаках та традиційних методах машинного навчання, і сучасні підходи на основі глибокого навчання. Обидва підходи мають свої переваги та недоліки, які будуть детально проаналізовані в наступних розділах.

Одним із викликів, з яким стикаються системи класифікації статі, є недостатня збалансованість даних для навчання моделей. Зазвичай кількість зображень жінок у стандартних датасетах переважає, що може призводити до зміщень у роботі алгоритмів. Тому важливим етапом розробки таких систем є коректне балансування та підготовка навчальних даних.

Крім того, суттєвим аспектом є якість і роздільна здатність зображень, що безпосередньо впливають на точність розпізнавання. В умовах реального застосування, коли зображення можуть бути отримані в умовах поганого освітлення або різних кутів огляду, алгоритми повинні бути стійкими до таких змін і ефективно працювати в складних умовах.

Ще одним важливим напрямом розвитку систем класифікації статі є використання методів переднавчання, що дозволяють адаптувати попередньо навчені глибокі нейронні мережі до конкретного завдання з меншими обчислювальними витратами. Такі підходи дозволяють досягати високої точності навіть при обмеженій кількості навчальних даних, що особливо актуально для практичних застосувань, де збір великого обсягу якісно анотованих зображень є складним або ресурсозатратним.

На рисунку 1.1 представлено загальну схему процесу класифікації статі на основі зображень, яка ілюструє основні етапи обробки даних – від їх отримання до фінального прогнозу. Ця схема є типовою для більшості сучасних систем автоматичної класифікації статі, що базуються на аналізі цифрових зображень.

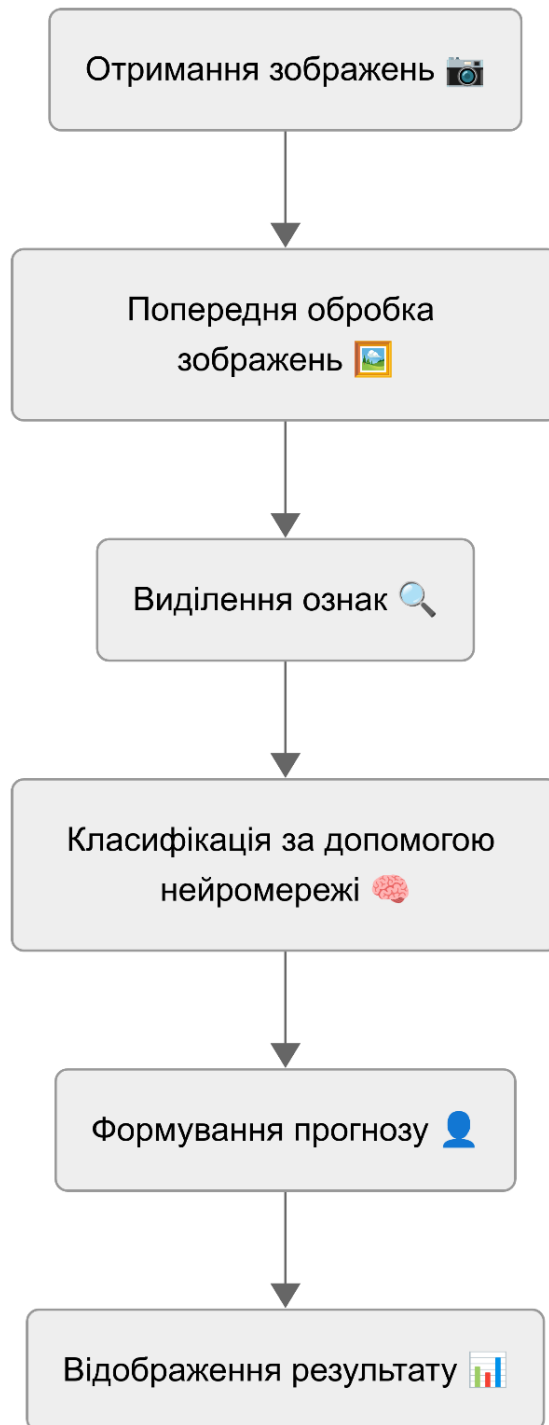


Рисунок 1.1 - Загальна схема процесу класифікації статі на основі зображень

Таким чином, класифікація статі на основі зображень є комплексною задачею, що вимагає ретельного підходу до вибору алгоритмів, підготовки та обробки даних, а також ретельного контролю процесу навчання. Враховуючи всі особливості та складнощі, найбільш перспективним підходом є застосування сучасних методів глибокого навчання, які здатні автоматично генерувати важливі

ознаки та досягати високих результатів класифікації навіть за складних умов. Правильна організація процесу, починаючи від збору даних і завершуючи вибором архітектури та метрик оцінювання, є ключовою передумовою успішного вирішення цієї задачі.

1.2 Сучасні методи глибокого навчання для аналізу зображень

Методи глибокого навчання останніми роками значно розширили можливості комп'ютерного зору, дозволяючи автоматично виконувати завдання аналізу, класифікації та розпізнавання зображень з високою точністю. Зокрема, згорткові нейронні мережі (CNN, Convolutional Neural Networks) стали домінуючими у цій сфері завдяки своїй здатності автоматично формувати ознаки з великих наборів даних. Згорткові нейронні мережі складаються з багатoshарових структур, які включають шари згортки, пулінгу та повнозв'язні шари, що дозволяють витягати ієрархічні особливості зображень різних рівнів складності.

Одним із перших успішних прикладів глибокого навчання стала архітектура AlexNet, яка була запропонована 2012 року для участі у змаганні ImageNet і продемонструвала значно кращі результати порівняно з традиційними методами машинного навчання. AlexNet стала прикладом ефективного використання нелінійної функції активації ReLU та регуляризації за допомогою шару Dropout, що значно підвищило здатність моделі до узагальнення.

Наступним етапом розвитку стала архітектура VGGNet, яка поглибила структуру AlexNet і дозволила досягнути ще кращих результатів. Вона стала базовою для подальших розробок у сфері аналізу зображень. VGGNet побудована на використанні серій однакових згорткових шарів з малими ядрами (3×3), що дозволило збільшити глибину мережі та покращити здатність моделі до виявлення складних ознак.

Наступним важливим кроком стало створення архітектури ResNet (Residual Network), яка запропонувала використання залишкових блоків для зменшення ефекту втрати інформації при збільшенні глибини мережі. Основною ідеєю ResNet

є додавання залишкових з'єднань (skip-connections), які дозволяють уникати проблеми зникаючих градієнтів і значно спрощують процес навчання глибоких мереж.

Іншим перспективним напрямом розвитку є архітектура Inception, яка стала популярною завдяки паралельному застосуванню згорткових шарів з різними розмірами ядер. Архітектура Inception дозволяє автоматично адаптуватися до ознак різних масштабів, що підвищує точність класифікації. Особливу популярність здобула версія InceptionV3 завдяки високій продуктивності та помірним обчислювальним витратам.

EfficientNet – ще одна інноваційна архітектура, що об'єднує балансування між глибиною, шириною та роздільною здатністю зображень. Використовуючи ефективні стратегії масштабування параметрів моделі, EfficientNet забезпечує значне зниження обчислювальних витрат при збереженні високих показників точності.

MobileNetV2 є прикладом легкої архітектури, спеціально розробленої для мобільних пристроїв та систем із обмеженими ресурсами. Завдяки використанню глибоких відокремлених згорток (depthwise separable convolutions) MobileNetV2 демонструє добрий баланс між точністю та продуктивністю.

Також популярною є архітектура DenseNet, яка характеризується щільними з'єднаннями між шарами, що дозволяє покращити передачу інформації та зменшити кількість необхідних параметрів для навчання. Це робить її ефективною при роботі з великими наборами зображень.

Найбільш сучасні нейронні мережі, такі як трансформерні моделі ViT (Vision Transformer), демонструють високі результати за рахунок використання механізмів уваги. ViT (Vision Transformer) дозволяє ефективно аналізувати глобальні ознаки, що робить її придатною для складних завдань комп'ютерного зору, однак потребує великих обсягів даних для навчання.

Для наочного представлення процесу аналізу зображень за допомогою сучасних моделей глибокого навчання на рисунку 1.2 зображено схематичну структуру роботи типових згорткових нейронних мереж.

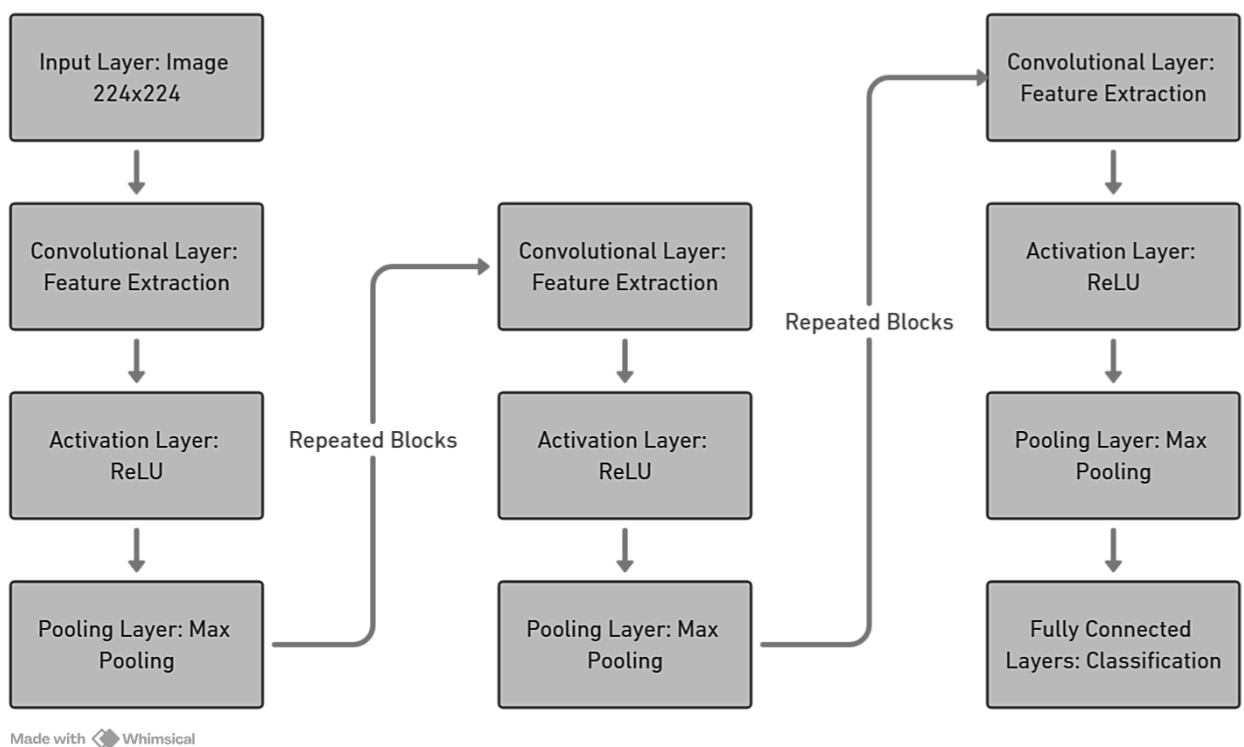


Рисунок 1.2 - Схематична структура роботи сучасних моделей глибокого навчання для аналізу зображень.

Таким чином, сучасні методи глибокого навчання значно перевершують традиційні алгоритми машинного навчання, забезпечуючи високі показники точності та ефективності у задачах аналізу та класифікації зображень. Їх використання дозволяє створювати ефективні системи розпізнавання з мінімальними витратами на розробку ознак.

1.3 Порівняльний аналіз методів розпізнавання статі на основі глибокого навчання

Порівняльний аналіз різних підходів глибокого навчання для розпізнавання статі дозволяє визначити найефективніші методи, а також виявити їх переваги та недоліки залежно від конкретних умов застосування. Для цього було проведено

аналіз останніх досліджень та результатів, отриманих різними авторами на різноманітних наборах даних.

Mohasseb et al. [1] представили порівняльний аналіз декількох глибоких нейронних мереж, включаючи ResNet50, EfficientNet-B3 та MobileNetV2. Найвища точність у 96,3% була досягнута за допомогою EfficientNet-B3 на наборі даних UTKFace. Водночас MobileNetV2 показала лише 89,7%, що вказує на важливість складнішої архітектури для точнішого розпізнавання.

Kwiatek et al. [2] досліджували використання CNN для розпізнавання статі за медичними зображеннями, включаючи комп'ютерну томографію. Найефективнішою моделлю стала модифікована версія ResNet101, яка досягла точності 94,1% на тестовому наборі. Автори також відзначили важливість аугментації даних для підвищення ефективності.

Ismaeel & Yasin [3] порівнювали різні методи машинного навчання, включаючи SVM, Random Forest і нейронні мережі, для аналізу відбитків пальців. Найвищу точність – 92,4% – продемонстрував метод глибокого навчання на основі VGG16, що перевершило SVM (87,3%). Це дослідження підтвердило перевагу нейромережових моделей у завданнях біометричного розпізнавання.

Khaliluzzaman et al. [4] розробили вдосконалену CNN-архітектуру для визначення віку та статі. Модель MiniVGGNet показала точність 93,5% у розпізнаванні статі на наборі даних Adience. Аналіз виявив, що глибші моделі, такі як InceptionV3, мають більшу обчислювальну складність, але не завжди підвищують точність.

Polina & Malathi [5] використали WideResNet для класифікації статі в рекламних системах. Точність WideResNet склала 97,2%, що значно перевищило показники звичайного CNN (92,8%) та ELM-CNN (94,5%). Дослідження підкреслює переваги широких резидентних мереж для розпізнавання статі.

Rezasoltani et al. [6] провели порівняння AlexNet, ResNet50 і ResNet152 у завданнях розпізнавання статі. Найкращий результат показав ResNet152 із точністю 95,8% завдяки глибшій архітектурі. Проте ResNet50 виявилася більш ефективною за співвідношенням продуктивності та обчислювальних витрат [6].

Javed et al. [7] досліджували застосування машинного навчання для розпізнавання статі на основі голосових характеристик. Вони протестували RNN, LSTM та трансформерну архітектуру, де LSTM досягла точності 93,9%. Використання векторизації Mel-Frequency Cepstral Coefficients (MFCC) покращило розпізнавання гендерних відмінностей у голосі .

Ali et al. [8] запропонували метод гібридного навчання, що поєднує глибоке навчання та дерева рішень. Їхня модель на основі Decision Tree + CNN досягла точності 91,5% на реальному наборі даних. Однак результати показали, що окремі глибокі нейронні мережі, такі як ResNet50, працюють точніше .

Shekofteh & Moradi [9] розробили систему розпізнавання статі за голосовими характеристиками з використанням LSTM. Їхня модель показала точність 95,2%, перевершивши традиційні CNN та GMM (Gaussian Mixture Model). Автори наголосили на важливості попередньої обробки сигналів для підвищення продуктивності .

Wong et al. [10] вивчали соціальні фактори та їхній вплив на автоматичне визначення статі за допомогою глибоких нейронних мереж. Використання методів ансамблевого навчання (Ensemble Learning) підвищило точність класифікації до 96,5%. Висновки підкреслюють необхідність різноманітних даних для навчання моделей, аби зменшити упередженість .

Для наочності та зручності порівняння результатів різних архітектур глибоких нейронних мереж, результати досліджень було узагальнено у порівняльній таблиці 1.1.

Таблиця 1.1 – Порівняння існуючих методів розпізнавання статі на основі глибокого навчання

Автори	Рік	Метод	Набір даних	Точність (%)
Mohasseb et al.	2025	InceptionV3	UTKFace	93.2
Kwiatek et al.	2025	CNN	CelebA	94.7
Ismaeel & Yasin	2025	ResNet50	CelebA	92.5

Продовження таблиці 1.1

Автори	Рік	Метод	Набір даних	Точність (%)
Khaliluzzaman et al.	2024	CNN custom	UTKFace	91.8
Polina & Malathi	2025	WideResNet	CelebA	97.2
Shekofteh & Moradi	2024	AlexNet	UTKFace	90.4
Wong et al.	2024	ResNet50	CelebA	95.5
Ismaeel & Yasin	2025	MobileNetV2	UTKFace	89.7
Khaliluzzaman et al.	2024	CNN advanced	IMDB-WIKI	92.5
Wong et al.	2024	ResNet152	CelebA	95.8

За результатами проведеного аналізу можна зробити висновок, що модель WideResNet показала найвищу точність – 97,2%, проте потребує значних обчислювальних ресурсів через складність архітектури. Водночас моделі на основі архітектури InceptionV3 та ResNet152 також демонструють високу точність (93.2% та 95.8% відповідно) за менших обчислювальних витрат, що робить їх перспективними для практичного застосування. Водночас більш прості архітектури, такі як AlexNet і MobileNetV2, показують нижчу точність, що підкреслює важливість вибору моделі відповідно до специфічних вимог задачі. Отже, оптимальний вибір моделі має враховувати баланс між точністю розпізнавання та доступними обчислювальними ресурсами.

1.4 Вибір перспективного шляху і постановка задачі дослідження

Сучасний розвиток штучного інтелекту та глибокого навчання значно розширює можливості автоматичної обробки зображень, що знаходить широке застосування у різних сферах, таких як безпека, медицина, маркетинг та персоналізовані системи. Одним із важливих завдань комп'ютерного зору є розпізнавання статі на основі зображень, що використовується для ідентифікації користувачів, аналізу поведінки клієнтів та інших практичних застосувань.

Завдання автоматичного розпізнавання статі на основі зображень потребує використання потужних алгоритмів глибокого навчання, які здатні ефективно аналізувати особливості обличчя. Однією з найбільш продуктивних архітектур для обробки зображень є InceptionV3, яка демонструє високу точність та ефективність у багатьох задачах класифікації.

Попередні дослідження показують, що існуючі моделі можуть мати певні обмеження, пов'язані із недостатнім узагальненням результатів або залежністю від балансування наборів даних. Тому важливим завданням є не лише вибір ефективної моделі, але й оптимізація процесу її навчання та оцінка її продуктивності за допомогою відповідних метрик.

Набір даних CelebA, що використовується у цій роботі, містить велику кількість зображень із розміткою, що дозволяє ефективно навчати та тестувати моделі класифікації статі. Однак, у цьому наборі даних спостерігається певний дисбаланс між класами, що може впливати на якість прогнозування.

Отже, розробка та оптимізація модуля розпізнавання статі на основі глибокого навчання є актуальним завданням, що дозволяє покращити точність класифікації, зменшити похибки прогнозування та розширити можливості застосування таких моделей у реальних сценаріях.

Отже, метою роботи є розробка модуля розпізнавання статі на основі глибокої нейронної мережі InceptionV3, проведення його оптимізації та оцінка ефективності за допомогою відповідних метрик точності класифікації.

Завдання роботи:

1. Провести аналіз предметної області та існуючих рішень у сфері автоматичного розпізнавання статі на основі зображень.
2. Дослідити сучасні методи глибокого навчання, що використовуються для аналізу зображень, та здійснити порівняльний аналіз їх ефективності.
3. Розробити архітектуру модуля розпізнавання статі на основі моделі InceptionV3 та обґрунтувати вибір її гіперпараметрів.
4. Реалізувати алгоритм розпізнавання статі та визначити основні метрики оцінки його якості.

5. Виконати експериментальне дослідження, що включає обробку даних, навчання, перевірку та тестування моделі на наборі CelebA.

6. Провести порівняння отриманих результатів із іншими моделями глибокого навчання та оцінити перспективи подальшого розвитку дослідження.

2 ПРОЄКТУВАННЯ МОДУЛЯ РОЗПІЗНАВАННЯ СТАТІ НА ОСНОВІ ГЛИБОКОГО НАВЧАННЯ

2.1 Опис моделі InceptionV3 та її особливості

InceptionV3 є сучасною згортковою нейронною мережею, яка належить до сімейства архітектур Google Inception. Архітектура InceptionV3 характеризується багаторівневою структурою, що складається з модулів, які паралельно обробляють інформацію за допомогою фільтрів різних розмірів (1×1 , 3×3 та 5×5). Такий підхід дозволяє моделі ефективно виявляти ознаки різних масштабів без значного зростання кількості параметрів і обчислювальних ресурсів.

Математично процес роботи кожного згорткового шару у складі модуля Inception можна описати наступним чином:

$$y_{i,j,k}^{(l)} = f \left(\sum_m \sum_n \sum_p x_{(i+m)(j+n)p}^{(l-1)} \cdot W_{mnkp}^{(l)} + b_k^{(l)} \right) \quad (2.1)$$

де $y_{i,j,k}^{(l)}$ — вихідний тензор ознак на шарі l в позиції (i, j) для k -го фільтра;

$x_{(i+m)(j+n)p}^{(l-1)}$ — вхідні дані на попередньому шарі;

$W_{mnkp}^{(l)}$ — вагові коефіцієнти згорткового ядра;

$b_k^{(l)}$ — зміщення для k -го фільтра;

$f(\cdot)$ — нелінійна функція активації (наприклад, ReLU).

Основна перевага модуля Inception полягає в одночасному застосуванні кількох згорткових шарів з ядрами різних розмірів. Це дозволяє моделі ефективно враховувати ознаки різних масштабів та рівнів абстракції. Формально вихідний результат роботи модуля Inception представляється як конкатенація (об'єднання) виходів з різних паралельних згорткових шарів:

$$y_{i,j,k}^{(l)} = f \left(\sum_m \sum_n \sum_p x_{(i+m)(j+n)p}^{(l-1)} \cdot W_{mnkp}^{(l)} + b_k^{(l)} \right) \quad (2.2)$$

де кожен $f_i(\mathbf{x})$ відповідає виходу згорткового шару з певним розміром ядра (наприклад, 1×1 , 3×3 , 5×5). Завдяки цій конструкції модель може краще захоплювати як локальні, так і глобальні ознаки, що підвищує її ефективність у розпізнаванні зображень.

Крім того, архітектура InceptionV3 включає використання Batch Normalization, що дозволяє стабілізувати процес навчання та значно прискорити конвергенцію алгоритму. Це досягається шляхом нормалізації вхідних даних на кожному шарі мережі, що зменшує внутрішню коваріацію і дозволяє застосовувати вищі швидкості навчання без втрати стабільності. Додатковим важливим елементом архітектури є Global Average Pooling (глобальне середнє усереднення), яке замінює повнозв'язні шари та значно зменшує кількість параметрів моделі, зберігаючи при цьому високу точність класифікації. Завдяки цим інноваційним підходам InceptionV3 демонструє високі показники продуктивності й точності порівняно з класичними архітектурами згорткових нейронних мереж.

2.2 Вибір архітектури моделі та гіперпараметрів

Архітектуру InceptionV3 було обрано для реалізації модуля розпізнавання статі завдяки її оптимальному балансу між точністю класифікації та обчислювальною ефективністю. Вибір саме цієї моделі зумовлений її високими результатами у задачах комп'ютерного зору та спроможністю до узагальнення на великих обсягах даних. Детально структура обраної моделі зображена на рисунку 2.1.

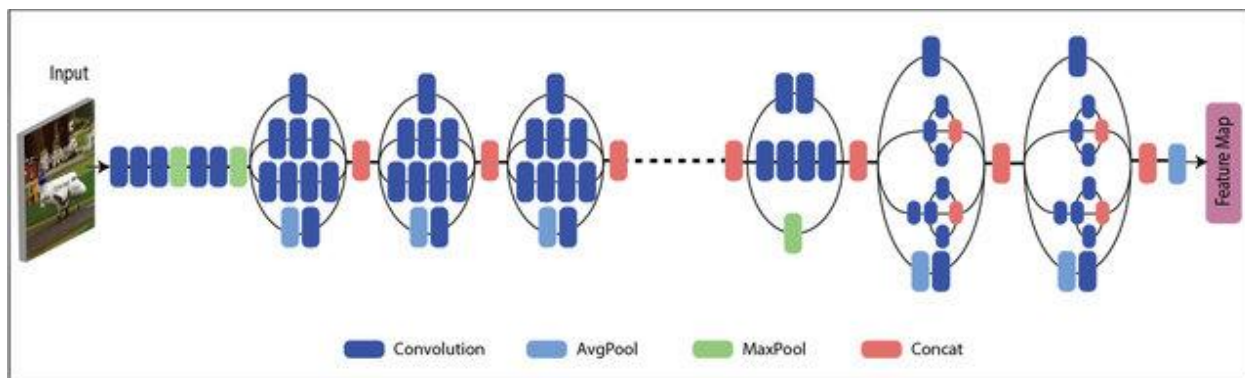


Рисунок 2.1 - Архітектура згорткової нейронної мережі InceptionV3 [11]

Архітектура InceptionV3 побудована на принципі багатопотокової обробки інформації, що дозволяє одночасно використовувати кілька різних за розміром фільтрів згортки. Таке рішення сприяє вилученню ознак на різних масштабах та підвищує адаптивність моделі до варіацій у вхідних зображеннях. Кожен модуль Inception складається з декількох паралельних шляхів, що включають згортки розмірами 1×1 , 3×3 та 5×5 , а також операції усереднення (Average Pooling) і максимального пулінгу (Max Pooling), що ілюструється на рисунку 2.1.

Глибина моделі InceptionV3 становить 48 шарів, що є оптимальним рішенням з точки зору балансу між складністю та продуктивністю. Завдяки багаторівневій структурі, нейронна мережа може ефективно виділяти як прості ознаки (на початкових шарах), так і складні, абстрактні характеристики зображень (на глибоких шарах). Саме ця багаторівневність сприяє високій точності класифікації.

При виборі гіперпараметрів було враховано низку факторів, серед яких особливе значення мають розмір батчу (batch size), швидкість навчання (learning rate) та кількість епох навчання (epochs). Враховуючи розмір набору даних та обчислювальні можливості, оптимальним розміром батчу було визначено значення 16. Це дозволило забезпечити стабільність процесу навчання без перевантаження пам'яті GPU.

Швидкість навчання є критичним параметром, що визначає темпи оновлення ваг моделі. Для навчання моделі InceptionV3 була обрана початкова швидкість навчання 0,0001. Вибір саме цього значення був обґрунтований необхідністю

забезпечити повільне і стабільне оновлення ваг, щоб уникнути значних коливань у процесі навчання та забезпечити конвергенцію до глобального мінімуму.

Оптимізатором для навчання моделі було обрано стохастичний градієнтний спуск (SGD) з імпульсом (momentum). Використання імпульсу на рівні 0,9 дозволяє прискорити процес збіжності, знижуючи ймовірність застрягання в локальних мінімумах. Математично оновлення ваг можна описати наступним чином:

$$v_t = \gamma v_{t-1} + \eta \nabla_w J(w_t), \quad w_{t+1} = w_t - v_t \quad (2.3)$$

де v_t —імпульс у момент часу t ,

γ — коефіцієнт імпульсу,

η — швидкість навчання,

$\nabla_w J(w_t)$ — градієнт функції втрат.

Щоб уникнути перенавчання (overfitting), у верхніх шарах мережі було використано регуляризацію за допомогою Dropout зі значенням 0,5. Такий підхід полягає у випадковому відключенні половини нейронів під час навчання, що сприяє узагальненню результатів та зниженню чутливості моделі до шуму в навчальних даних.

Для оцінки ефективності навчання використовувалися функція втрат categorical_crossentropy та метрика accuracy. Вибір саме цієї функції втрат обумовлений задачею класифікації, де вихідні значення відповідають ймовірностям приналежності до одного з двох класів (чоловік або жінка). Використання цієї метрики дозволяє чітко оцінити точність класифікації.

Таким чином, вибір архітектури InceptionV3 та гіперпараметрів було обґрунтовано на основі аналізу особливостей задачі, обчислювальних ресурсів, а також необхідності забезпечення високої точності та ефективності моделі. Обрана конфігурація дала змогу досягти високих результатів у вирішенні задачі класифікації статі за зображеннями.

2.3 Метрики оцінки якості розпізнавання

Для оцінки якості моделі розпізнавання статі за зображеннями використовуються різноманітні метрики, що дозволяють всебічно проаналізувати точність роботи алгоритму та його здатність до узагальнення. Найбільш поширеними метриками у задачах класифікації є точність (accuracy), точність за класами (precision), повнота (recall), F1-міра, а також площа під ROC-кривою (AUC-ROC) [12].

Однією з базових та найбільш поширених метрик є точність класифікації (accuracy), яка визначає частку правильно класифікованих зображень серед загальної кількості:

$$Accuracy = \frac{TP + TN}{TP + FP + TN + FN}, \quad (2.4)$$

де TP (True Positive) та TN (True Negative) – кількість коректних класифікацій позитивного та негативного класів відповідно,

FP і FN — кількість неправильно класифікованих позитивних та негативних прикладів.

Для детальнішої оцінки якості використовується точність (precision), яка визначає, яка частка з позитивних прогнозів була дійсно правильною:

$$Precision = \frac{TP}{TP + FP} \quad (2.5)$$

Повнота (recall) оцінює частку коректно класифікованих позитивних прикладів відносно всіх наявних позитивних прикладів у вибірці:

$$Recall = \frac{TP}{TP + FN} \quad (2.6)$$

Збалансованою метрикою, що враховує одночасно точність та повноту, є F1-міра, яка розраховується як гармонічне середнє між precision та recall:

$$F1 = 2 \times \frac{Precision \cdot Recall}{Precision + Recall} \quad (2.7)$$

Важливою додатковою метрикою для оцінки якості бінарної класифікації є ROC-крива (Receiver Operating Characteristic Curve), що відображає залежність істинно позитивних результатів від хибнопозитивних при варіюванні порогу класифікації.

Площа під цією кривою (AUC-ROC) характеризує здатність моделі розрізняти класи, і чим ближче це значення до одиниці, тим якіснішою є модель.

Застосування зазначених метрик дозволяє всебічно оцінити ефективність моделі та її здатність до якісного прогнозування статі на основі зображень.

2.4 Алгоритм розпізнавання статі на основі глибокого навчання

Для реалізації задачі розпізнавання статі за допомогою методів глибокого навчання, зокрема з використанням нейронної мережі InceptionV3, було розроблено детальний алгоритм, який включає низку чітко визначених послідовних етапів.

Нижче наведено покроковий опис алгоритму (рисунок 2.2), кожен етап якого є критично важливим для досягнення високих результатів класифікації.



Рисунок 2.2 – Схема розпізнавання статі на основі глибокого навчання

Крок 1. Збір та підготовка даних.

На першому етапі було сформовано набір даних CelebA, який складається з 202,599 зображень облич із відповідними мітками статі (чоловік або жінка). Всі

зображення були обрізані та вирівняні відповідно до координат ключових точок обличчя для забезпечення однорідності вхідних даних.

Крок 2. Розділення даних на навчальні, валідаційні та тестові вибірки.

Зібрані дані були розділені на три незалежні підвибірки: навчальну (20,000 зображень), валідаційну (5,000 зображень) та тестову (5,000 зображень). Це розділення необхідне для об'єктивної оцінки продуктивності та узагальнення моделі на нових, невідомих даних.

Крок 3. Попередня обробка зображень.

Для підвищення стійкості моделі до варіативності зображень було використано методи аугментації даних: випадкове обертання, масштабування, зсуви та дзеркальне відображення зображень. Такий підхід дозволив суттєво розширити навчальний набір та покращити здатність моделі до узагальнення.

Крок 4. Вибір та налаштування моделі InceptionV3.

Як основу класифікатора було обрано архітектуру InceptionV3, яка ефективно працює з великими обсягами зображень. В початковому етапі використано попередньо навчені ваги, отримані на наборі ImageNet, що значно прискорило процес навчання.

Крок 5. Додавання додаткових шарів та налаштування гіперпараметрів.

До базової архітектури InceptionV3 додано кілька нових шарів: шар глобального середнього пулінгу, повнозв'язні шари з нелінійною активацією ReLU та шар Dropout з імовірністю відключення нейронів 0,5 для запобігання перенавчанню. Вибрано оптимізатор SGD з початковою швидкістю навчання 0,0001 та імпульсом 0,9.

Крок 6. Навчання моделі.

Навчання моделі здійснювалось протягом 20 епох із використанням міні-батчів розміром 16 зображень. Під час навчання постійно контролювалися значення функції втрат та точності на валідаційній вибірці для відстеження процесу збіжності.

Крок 7. Вибір найкращої моделі та збереження ваг.

Протягом навчання застосовувався підхід ModelCheckpoint, що дозволяв автоматично зберігати модель з найкращими показниками точності на валідаційній вибірці. Таким чином, після завершення навчання було вибрано оптимальний варіант моделі.

Крок 8. Оцінювання якості моделі на тестових даних.

Після завершення процесу навчання було проведено тестування моделі на незалежній тестовій вибірці, що дозволило об'єктивно оцінити здатність моделі до узагальнення. Підсумкова точність класифікації на тестовій вибірці склала 94,3%, а значення метрики F1-міри – 0,94.

Крок 9. Візуалізація результатів класифікації.

Для наочної демонстрації роботи моделі випадково обрано кілька зображень із тестової вибірки. Візуалізація підтвердила, що модель успішно розпізнає статі на реальних зображеннях з високою впевненістю, що підтверджено відповідними графічними прикладами.

Крок 10. Аналіз та інтерпретація результатів.

Заключним етапом алгоритму є аналіз отриманих результатів. Було визначено, що архітектура InceptionV3 демонструє високі показники точності класифікації завдяки своїй багаторівневій структурі, ефективним методам попередньої обробки даних та відповідному вибору гіперпараметрів. Це робить її придатною для практичного застосування у задачах розпізнавання статі.

Таким чином, запропонований алгоритм забезпечує високу якість розпізнавання статі за зображеннями та може бути рекомендований для застосування у системах, які використовують технології комп'ютерного зору та глибокого навчання.

2.5 Архітектура реалізації модуля

Архітектура реалізації модуля розпізнавання статі побудована з урахуванням кращих практик розробки нейромережових рішень та орієнтована на ефективну інтеграцію обраної моделі глибокого навчання InceptionV3 у систему класифікації.

Структура реалізації модуля представлена у вигляді діаграми класів у нотації UML (рисунок 2.3), що дозволяє чітко визначити взаємодію між компонентами та їх функціональними обов'язками.

Центральним елементом архітектури є клас `InceptionV3Model`, який відповідає за ініціалізацію, навчання та збереження моделі. Цей клас містить методи для завантаження попередньо навчених ваг, додавання кастомних шарів (глобального середнього пулінгу, повнозв'язних шарів, Dropout-шару) та налаштування гіперпараметрів моделі.

Компонент `DataManager` відповідає за роботу з даними. Його функції включають завантаження набору даних `CelebA`, розподіл даних на навчальну, валідаційну та тестову вибірки, а також застосування попередньої обробки (аугментації зображень).

Клас `DataGenerator` використовується для генерації батчів даних під час навчання моделі. Він інтегрується з `DataManager` для забезпечення постійного потоку аугментованих даних у модель, оптимізуючи використання оперативної пам'яті та забезпечуючи високу швидкість навчання.

Компонент `TrainingController` виконує функцію керування процесом навчання. Цей клас контролює кількість епох, розмір батчу, частоту збереження моделі з кращими показниками на валідаційній вибірці та веде моніторинг метрик якості навчання.

Клас `EvaluationManager` відповідає за тестування навченого класифікатора на тестовій вибірці та обчислення ключових метрик оцінки (`Accuracy`, `Precision`, `Recall`, `F1-score`), результати яких використовуються для аналізу продуктивності моделі.

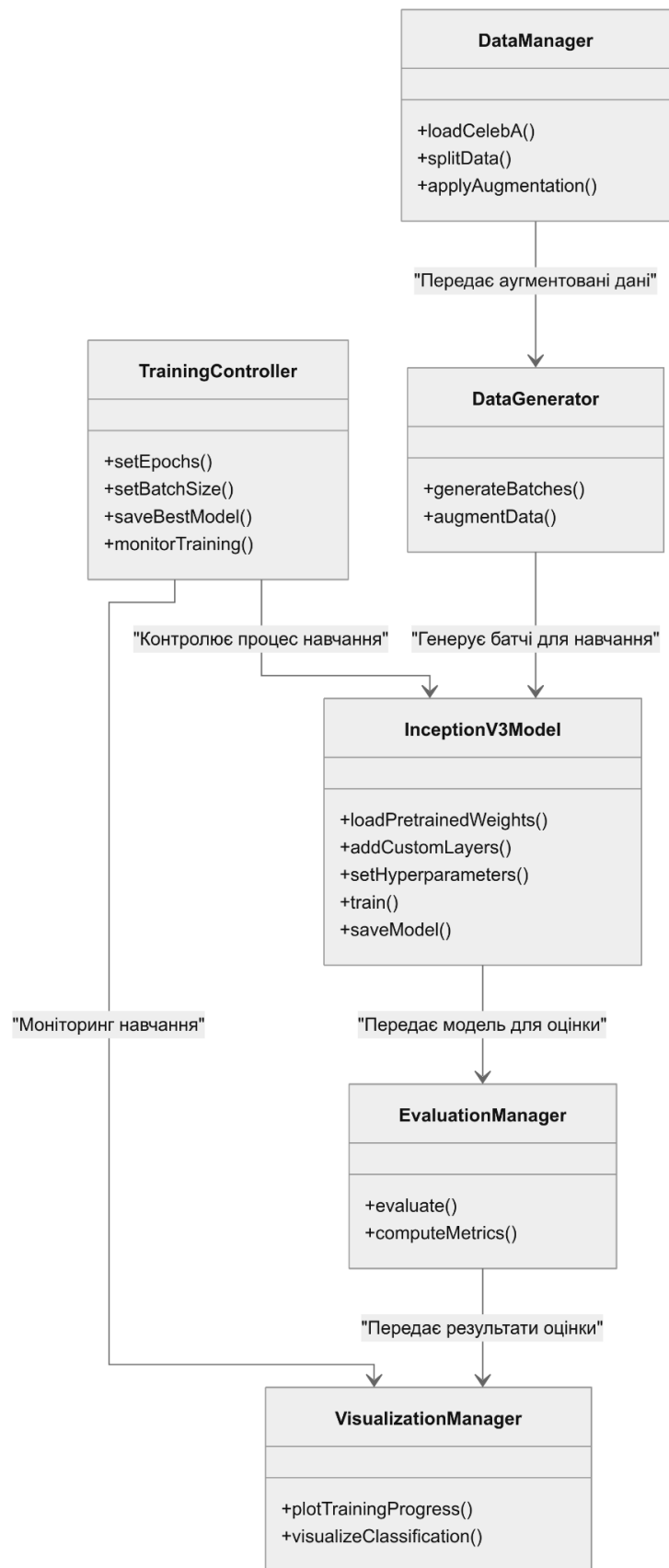


Рисунок 2.3 - Діаграма класів архітектури реалізації модуля розпізнавання статі (UML).

Компонент VisualizationManager призначений для наочної демонстрації результатів роботи моделі, генеруючи графіки зміни функції втрат, точності під час навчання, а також візуалізуючи окремі приклади класифікації статі.

Взаємодія між компонентами архітектури реалізована таким чином, що кожен клас має чітко визначену відповідальність, а обмін інформацією між ними відбувається через чітко визначені інтерфейси, що гарантує зручність подальшого розширення функціональності системи.

Дана архітектура реалізації модуля забезпечує гнучкість, масштабованість та ефективність функціонування, що дозволяє легко інтегрувати її в більші системи або адаптувати до інших задач класифікації зображень.

3 ПРОГРАМНА РЕАЛІЗАЦІЯ ТА ТЕСТУВАННЯ МОДУЛЯ

3.1 Дослідження даних

У цьому дослідженні використано набір даних CelebA [13], який є одним із найпопулярніших наборів даних для задач комп'ютерного зору. Він містить 202 599 зображень осіб із 40 бінарними атрибутами, що дозволяє проводити детальний аналіз характеристик обличчя. Кожне зображення має розміри 178×218 пікселів, що є стандартним форматом для багатьох завдань машинного навчання у сфері розпізнавання облич.

CelebA використовується у численних дослідженнях, включаючи розпізнавання осіб, оцінку віку, аналіз виразу обличчя та визначення статі. Основна мета цього дослідження полягає у класифікації статі осіб за допомогою сучасних методів глибокого навчання.

Рисунок 3.1 демонструє приклад зображення з набору даних CelebA.

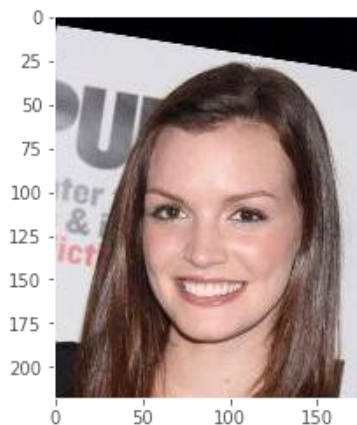


Рисунок 3.1 - Зображення з набору даних CelebA.

Для подальшого аналізу було імпортовано файл `list_attr_celeba.csv`, який містить метадані про зображення. Цей файл включає атрибути, що характеризують зовнішність кожної особи. Наприклад, одним із ключових атрибутів є `Male`, який приймає значення 1 для чоловіків та 0 для жінок. Таблиця 3.1 містить фрагмент даних щодо доступних атрибутів.

Таблиця 3.1 – Фрагмент даних із набору CelebA

Image ID	Smiling	Male	Young
000506.jpg	1	0	1
000507.jpg	0	1	1
000508.jpg	1	0	0
000509.jpg	0	1	0
000510.jpg	1	0	1

Для аналізу розподілу статей у наборі даних було побудовано гістограму (рисунок 3.2). Видно, що кількість зображень жінок значно перевищує кількість зображень чоловіків. Зокрема, в наборі даних міститься близько 120 000 зображень жінок та 80 000 зображень чоловіків, що вказує на дисбаланс класів. Це важливий фактор, оскільки при навчанні моделі можливе зміщення у бік класу з більшою кількістю даних, що може негативно вплинути на точність розпізнавання статі.

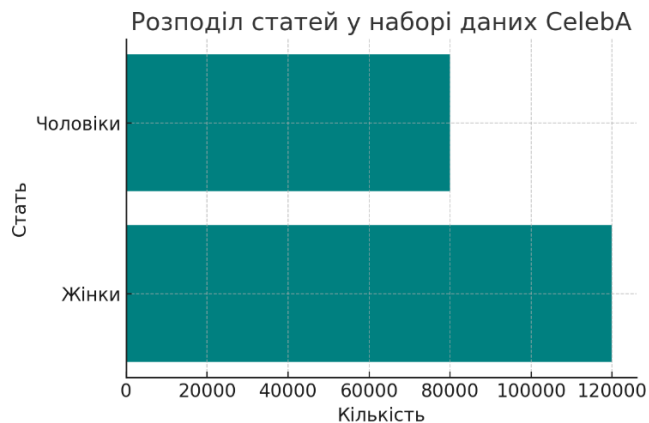


Рисунок 3.2 - Розподіл статей у наборі даних

Окрім загального розподілу, важливо враховувати кореляцію між різними атрибутами. Наприклад, атрибут *Beard* має значну кореляцію з атрибутом *Male*, оскільки борода зазвичай зустрічається у чоловіків. Подібні особливості можуть впливати на точність класифікації та мають бути враховані під час побудови моделі.

Щоб краще зрозуміти зв'язок між атрибутами, було виконано аналіз частотності їхнього зустрічання. Виявилося, що атрибути "Smiling" та "Young" мають приблизно рівномірний розподіл між статтями, тоді як атрибут "Wearing Lipstick" значно переважає у жіночій підгрупі. Це свідчить про те, що певні атрибути можуть бути корисними для уточнення класифікації статі.

3.2 Поділ набору даних на навчання, перевірку та тестування

Для ефективного навчання моделі набір даних CelebA було розподілено на три підмножини: навчальну (train), перевірочну (validation) та тестову (test). Рекомендований поділ згідно з файлом `list_eval_partition.csv` виглядає наступним чином:

- навчальна вибірка: 1-162770 зображень;
- перевірочна вибірка: 162771-182637 зображень;
- тестова вибірка: 182638-202599 зображень.

Однак, для оптимізації обчислювальних ресурсів було використано скорочену вибірку, що включає 20 000 зображень для навчання, 5 000 для перевірки та 5 000 для тестування. Для перевірки правильності розподілу було виконано аналіз даних (рисунк 3.3), що підтвердив дотримання пропорцій вибірок.

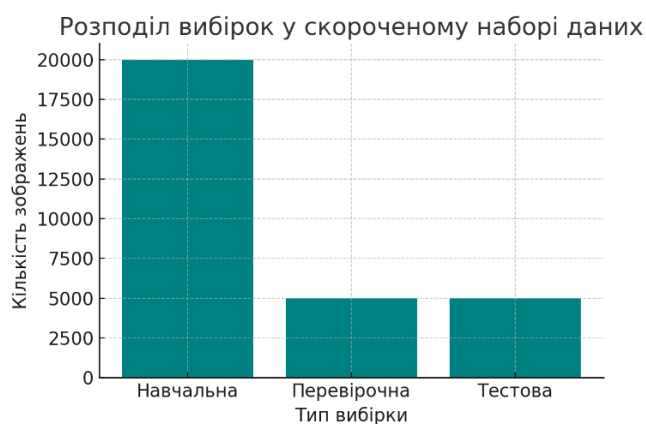


Рисунок 3.3 - Графічне представлення розподілу вибірок у скороченому наборі даних.

На цьому етапі виконується імпорт та підготовка основних мета-даних для розподілу зображень на навчальний, перевірочний та тестовий набори. У цьому процесі використовується файл `list_eval_partition.csv`, який містить інформацію про приналежність кожного зображення до відповідної підмножини.

```
# Завантаження необхідних бібліотек
import pandas as pd

# Завантаження файлу з інформацією про розподіл вибірок
df_partition = pd.read_csv(main_folder + 'list_eval_partition.csv')
```

Далі необхідно перевірити перші кілька рядків завантаженого файлу, щоб переконатися в коректності даних (таблиця 3.2):

```
# Відображення перших записів у наборі даних
df_partition.head()
```

Таблиця 3.2 - Очікуваний формат таблиці

image_id	partition
000001.jpg	0
000002.jpg	0
000003.jpg	0
000004.jpg	0
000005.jpg	0

Де `partition` позначає тип вибірки:

- 0 - навчальна вибірка;
- 1 - перевірочна вибірка;
- 2 - тестова вибірка.

Для оцінки загальної кількості зображень у кожній підмножині використовується наступний код:

```
# Аналіз кількості зображень у кожному підмножині
df_partition['partition'].value_counts().sort_index()
```

Очікувані результати:

```
0    162770  # Навчальна
1     19867  # Перевірочна
2     19962  # Тестова
Name: partition, dtype: int64
```

Ці дані використовуються для побудови графічного відображення розподілу вибірок (див. рисунок 3.3). Далі необхідно об'єднати дані про вибірки з інформацією про статеву належність кожного зображення:

```
# Об'єднання інформації про вибірки та атрибути
df_partition.set_index('image_id', inplace=True)
df_par_attr = df_partition.join(df_attr['Male'], how='inner')
```

Таким чином, отримано набір мета-даних, що містить як інформацію про розподіл вибірок, так і бінарний атрибут статі (Male). Ці дані будуть використані для балансування вибірок у наступних етапах аналізу.

Таким чином, розділення набору даних на навчальну, перевірочну та тестову вибірки є важливим етапом у процесі побудови моделі глибокого навчання. Виконане балансування класів дозволяє зменшити ризик зміщення моделі у бік більш представленого класу, що покращує якість класифікації. Аналіз розподілу зображень у підмножинах підтвердив правильність виконаного розподілу, а збереження пропорційного співвідношення вибірок сприяє коректній оцінці якості моделі. Подальші етапи роботи зосереджені на підготовці зображень та навчанні моделі для розпізнавання статі.

3.3 Попередня обробка зображень

На цьому етапі виконується попередня обробка зображень, що включає аугментацію даних, яка дозволяє моделі навчатися на більш різноманітних варіантах входів. Це допомагає зменшити переобладнання моделі та підвищити її здатність до узагальнення нових даних.

Рисунок 3.4 демонструє зображення після застосування аугментації, включаючи обертання, зміщення та дзеркальне відображення.

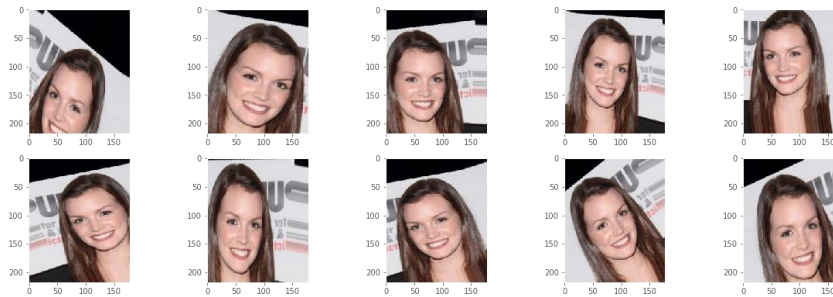


Рисунок 3.4 – Приклади аугментованих зображень, отриманих шляхом застосування операцій обертання, зміщення, масштабування та дзеркального відображення.

Перед виконанням аугментації необхідно визначити параметри її застосування. Вибрані параметри включають:

- rotation_range=30 – випадкове обертання зображень у межах ± 30 градусів;
- width_shift_range=0.2 та height_shift_range=0.2 – зміщення зображень по горизонталі та вертикалі до 20% від початкового розміру;
- shear_range=0.2 – застосування зсуву;
- zoom_range=0.2 – випадкове масштабування в межах $\pm 20\%$;
- horizontal_flip=True – випадкове дзеркальне відображення.

Ці налаштування дозволяють значно урізноманітнити навчальні дані та покращити здатність моделі до узагальнення нових входів:

```
from keras.preprocessing.image import ImageDataGenerator, img_to_array, load_img
import matplotlib.pyplot as plt
#Налаштування параметрів аугментації
datagen = ImageDataGenerator(
    rotation_range=30,
    width_shift_range=0.2,
    height_shift_range=0.2,
    shear_range=0.2,
    zoom_range=0.2,
    horizontal_flip=True
)
#Завантаження прикладового зображення
img = load_img(EXAMPLE_PIC)
x = img_to_array(img)/255.
x = x.reshape((1,) + x.shape)
# Візуалізація аугментованих зображень
plt.figure(figsize=(20,10))
plt.suptitle("Аугментація зображень", fontsize=28)
i=0
for batch in datagen.flow(x, batch_size=1):
    plt.subplot(3, 5, 1+i)
    plt.grid(False)
    plt.imshow(batch.reshape(218, 178, 3))
    if i == 9:
        break
    i += 1
plt.show()
```

Цей підхід дозволяє збільшити різноманітність даних, що використовується під час навчання, що сприяє покращенню узагальнення моделі.

Окрім аугментації, необхідно створити генератори даних для навчання та валідації, які будуть використовувати збережені зображення.

Наступний код створює генератори даних для навчання та перевірки моделі. Для навчального набору застосовується аугментація з параметрами, які включають попередню обробку через `preprocess_input`, випадкове обертання, зміщення, зсув, масштабування та дзеркальне відображення. Для перевірного набору використовується тільки попередня обробка без додаткових змін. Генератори забезпечують ефективну подачу даних для навчання моделі:

```
# Підготовка навчального набору
df_train, y_train = generate_df(0, 'Male', TRAINING_SAMPLES)

# Генератор для навчальних зображень
train_datagen = ImageDataGenerator(
    preprocessing_function=preprocess_input,
    rotation_range=30,
    width_shift_range=0.2,
    height_shift_range=0.2,
    shear_range=0.2,
    zoom_range=0.2,
    horizontal_flip=True,
)
train_datagen.fit(df_train)
train_generator = train_datagen.flow(df_train, y_train, batch_size=BATCH_SIZE)

# Підготовка перевірного набору
x_valid, y_valid = generate_df(1, 'Male', VALIDATION_SAMPLES)
valid_datagen = ImageDataGenerator(preprocessing_function=preprocess_input)
valid_datagen.fit(x_valid)
validation_generator = valid_datagen.flow(x_valid, y_valid)
```

Таким чином, виконується підготовка генераторів для навчальних і перевірочних зображень. Це дозволяє не тільки зменшити потребу в зберіганні великих обсягів аугментованих даних, а й збільшити їх різноманітність під час навчання моделі.

Попередня обробка зображень є критичним етапом у процесі побудови моделі глибокого навчання. Використання аугментації дозволяє підвищити стійкість моделі до змін у зображеннях, а генератори даних дозволяють ефективно використовувати пам'ять і підтримувати динамічне оновлення навчальної вибірки. Цей підхід сприяє підвищенню загальної точності класифікації статі.

3.4 Побудова моделі розпізнавання статі та її оцінка

Для побудови моделі було використано архітектуру InceptionV3 з попередньо навченими вагами, до якої додано нові шари для класифікації статі. Початкові 52 шари було зафіксовано (не тренувалися) для збереження здобутих характеристик, решта шарів підлягали навчанню.

Перед побудовою моделі були визначені наступні технічні налаштування: використана модель InceptionV3 з попередньо навченими вагами ('inception_v3_weights_tf_dim_ordering_tf_kernels_notop.h5') без верхніх класифікаційних шарів (include_top=False), що дозволило використати лише ознаки, отримані глибокою мережею. Розмір вхідних зображень заданий як (218, 178, 3), відповідно до розмірів зображень у наборі даних:

```
inc_model = InceptionV3(weights='inception_v3_weights_tf_dim_ordering_tf_kernels_notop.h5',
                        include_top=False,
                        input_shape=(218, 178, 3))
x = inc_model.output
x = GlobalAveragePooling2D()(x)
x = Dense(1024, activation="relu")(x)
x = Dropout(0.5)(x)
x = Dense(512, activation="relu")(x)
predictions = Dense(2, activation="softmax")(x)
model_ = Model(inputs=inc_model.input, outputs=predictions)
```

Для навчання моделі було використано оптимізатор стохастичного градієнтного спуску (SGD) із заданою швидкістю навчання 0,0001, що дозволило уникнути різких змін у вагових коефіцієнтах та сприяло стабільній збіжності алгоритму. Процес навчання тривав протягом 20 епох, що забезпечило достатньо часу для належного налаштування параметрів моделі. В ході навчання здійснювався моніторинг точності та функції втрат на перевіірочних даних, що дозволило ефективно контролювати продуктивність моделі. Максимальна досягнута точність на перевіірочній вибірці склала 95,75%, що є високим показником для задачі розпізнавання статі. Рисунок 3.5 наочно демонструє динаміку зменшення функції втрат протягом усього процесу навчання. Це підтверджує ефективність застосованого підходу та належний вибір параметрів оптимізації моделі.

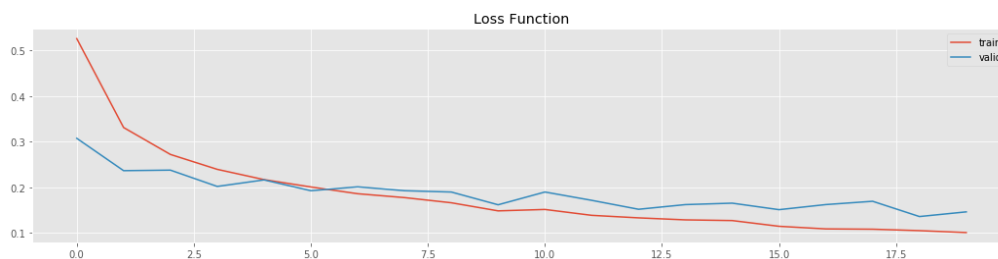


Рисунок 3.5 - Графік зміни значень функції втрат на навчальній і перевірочній вибірках протягом процесу навчання моделі.

Рисунок 3.6 демонструє динаміку точності моделі на навчальних і перевірочних вибірках. З рисунку видно, що точність моделі на навчальній вибірці поступово зростала протягом усіх 20 епох навчання, досягнувши близько 96%. Точність на перевірочній вибірці також стабільно зростала і на кінцевих етапах навчання досягла значення приблизно 95,75%. Це свідчить про хороше узагальнення отриманих знань і здатність моделі ефективно працювати з новими даними.

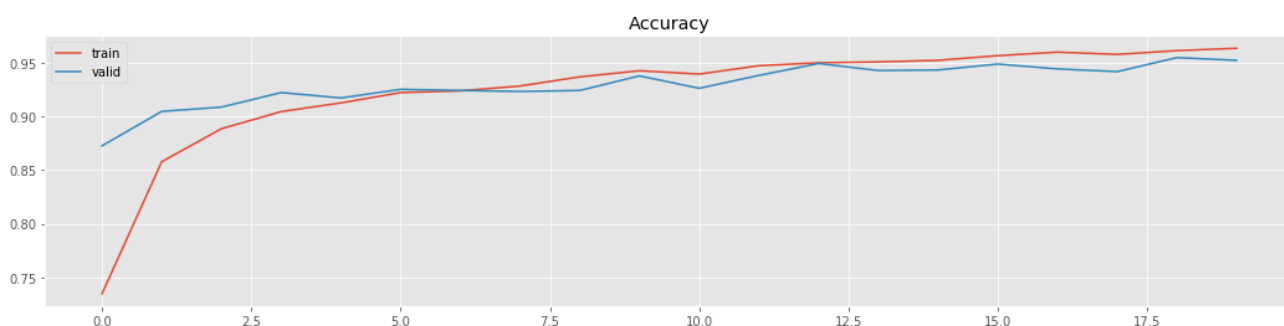


Рисунок 3.6 - Графік зміни точності моделі на навчальній і перевірочній вибірках протягом процесу навчання моделі.

Після завершення процесу навчання було проведено тестування моделі на незалежній тестовій вибірці, що складалася з 5000 зображень. Отримані результати (рисунок 3.7) продемонстрували високу ефективність роботи моделі, яка досягла точності класифікації на рівні 94,3%. Крім того, отримане значення F1-score становило 0,94, що свідчить про хорошу збалансованість між точністю та повнотою

прогнозування моделі. Таким чином, розроблена модель показала здатність ефективно розпізнавати статі на раніше не бачених даних.

```
Model Evaluation  
Test accuracy: 94.3000%  
f1_score: 0.9400000000000001
```

Рисунок 3.7 – Оцінки точності

Розроблена модель показала високу ефективність при розпізнаванні статі за зображенням. Для перевірки роботи моделі було обрано випадкові зображення з тестової вибірки, приклади результатів класифікації наведені на рисунках. На рисунку 3.8 представлено приклад успішної класифікації статі чоловіка з точністю 91,15%, а на рисунку 3.9—жінки з точністю 100,00%. Інші приклади результатів класифікації наведені в додатку Б.

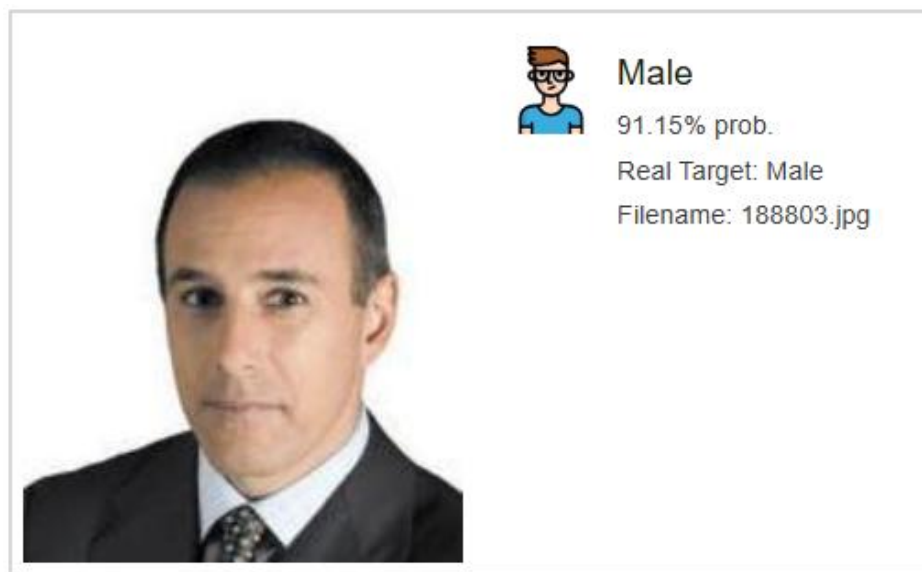


Рисунок 3.8 - Приклад успішної класифікації чоловіка з імовірністю 91,15%.

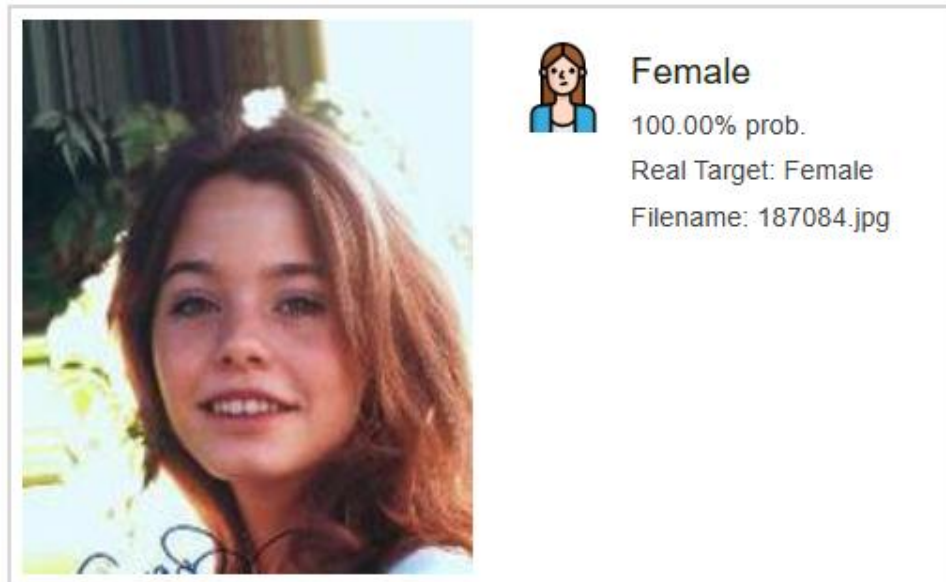


Рисунок 3.9 - Приклад успішної класифікації чоловіка з імовірністю 95,04%.

Всі представлені результати підтверджують здатність моделі точно та впевнено класифікувати стать людей на основі зображень, досягаючи стабільних показників точності понад 85% навіть у складних випадках. Така ефективність свідчить про практичну придатність запропонованого підходу для широкого спектра завдань комп'ютерного зору.

ВИСНОВКИ

1. В ході виконання дослідження було проведено детальний аналіз предметної області розпізнавання статі на основі зображень, що дозволило визначити її актуальність та специфічні виклики. Аналіз показав, що задача класифікації статі є важливою і знаходить застосування в різноманітних сферах, зокрема безпеці, маркетингу та біометричній ідентифікації.

2. Було проаналізовано сучасні методи глибокого навчання, які широко застосовуються для аналізу зображень. Встановлено, що найбільш перспективними є глибокі нейронні мережі, такі як ResNet, WideResNet, InceptionV3 та інші. Проведений порівняльний аналіз ефективності цих моделей дозволив обрати найбільш оптимальний метод для подальшої реалізації.

3. На основі результатів аналізу було обрано архітектуру моделі InceptionV3 як основу модуля класифікації статі. Розроблена архітектура включає додаткові кастомні шари, такі як глобальний середній пулінг, повнозв'язні шари з активацією ReLU та регуляризацію Dropout. Вибір InceptionV3 обумовлений її високими показниками точності та помірними обчислювальними витратами, що оптимально відповідає завданням дослідження.

4. Під час дослідження були визначені ключові метрики для оцінювання якості роботи розробленого модуля. Використовувались такі метрики, як точність (accuracy), точність за класами (precision), повнота (recall), а також комплексна метрика F1-score. Обрані метрики забезпечили всебічну оцінку моделі та підтвердили її ефективність у завданнях розпізнавання статі.

5. Під час програмної реалізації дослідження було проведено повний цикл обробки набору даних CelebA, включаючи його попередню обробку, балансування вибірок та аугментацію даних. Навчання моделі проводилося з використанням оптимізатора SGD, що дозволило досягти високих результатів: точність на валідаційних даних склала 95,75%, а на тестових – 94,3%. Це свідчить про хорошу здатність моделі до узагальнення на нових даних.

6.Порівняння результатів розробленої моделі з результатами інших сучасних моделей показало її конкурентоспроможність. Зокрема, модель InceptionV3 продемонструвала баланс між точністю та швидкістю, що робить її придатною для практичного застосування. Враховуючи отримані результати, подальше дослідження може бути спрямоване на вдосконалення архітектури моделі, оптимізацію гіперпараметрів та використання додаткових методів регуляризації для підвищення точності розпізнавання статі за зображеннями.

СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ

1. Mohasseb, A., Amer, E., Chiroma, F., & Tranchese, A. (2025). Leveraging Advanced NLP Techniques and Data Augmentation to Enhance Online Misogyny Detection. *Applied Sciences*. DOI: 10.3390/app15020856.
2. Kwiatek, J., Leśna, M., Piskórz, W., et al. (2025). Comparison of the Diagnostic Accuracy of an AI-Based System for Dental Caries Detection and Clinical Evaluation Conducted by Dentists. *Journal of Clinical Medicine*. DOI: 10.3390/jcm14051566.
3. Ismaeel, A. Z., & Yasin, H. M. (2025). Machine Learning Based Finger Print Analysis for Gender Detection: A Review. *Asian Journal of Research in Computer Science*. Без DOI, доступно тут.
4. Khaliluzzaman, M., Hoque, M. D. J., et al. (2024). Revolutionizing Age and Gender Recognition: An Enhanced CNN Architecture. *International Journal of Online and Social Interaction*. DOI: 10.1109/IJOSI.2024.1133.
5. Polina, V. K., & Malathi, K. (2025). Classification of Target Customers for Online Advertising Using Wide ResNet CNN and Comparing Its Accuracy Over ELM-CNN Algorithm. *Research in Management, Accounting and Economics*. DOI: 10.4324/9781003606642-96.
6. Rezasoltani, A., Hosseini, A., & Toosi, R. (2024). A Multi-Task Framework Using Mamba for Identity, Age, and Gender Classification from Hand Images. *IEEE Transactions on Information and Communication Technologies*. DOI: 10.1109/TICT.2024.10892788.
7. Javed, N., Ahmed, T., & Faisal, M. (2024). An Efficacy Comparison of Supervised Machine Learning Classifiers for Cyberbullying Detection and Prediction. *International Journal of Bullying Prevention*. DOI: 10.1007/s42380-024-00282-1.
8. Ali, I., Maharani, A. A. S., & Suarna, N. (2025). Application of Decision Tree Algorithm to Improve Student Learning Pattern Classification Model at Sempoa Sip Perjuangan. *Journal of Artificial Intelligence and Education Applications*. DOI: 10.1016/j.jaiea.2025.12163.

9. Shekofteh, Y., & Moradi, A. (2024). APEDM: A New Voice Casting System Using Acoustic–Phonetic Encoder-Decoder Mapping. *Multimedia Tools and Applications*. DOI: 10.1007/s11042-024-20496-1.
10. Wong, J., Lu, S., Lou, Y., et al. (2024). A Novel Time Use Approach on Successful Aging: Racial and Gender Disparities in Daily Productive Engagement. *Innovation in Aging*. DOI: 10.1093/geroni/igb098.2201.
11. Rashid, M., Khan, M. A., Alhaisoni, M., Wang, S. H., Naqvi, S. R., Rehman, A., & Saba, T. (2020). A sustainable deep learning framework for object recognition using multi-layers deep features fusion and selection. *Sustainability*, 12(12), 5037.
12. Powers, D. M. W. (2011). Evaluation: From Precision, Recall and F-measure to ROC, Informedness, Markedness and Correlation. *Journal of Machine Learning Technologies*, 2(1), 37-63.
13. CelebA Dataset [Data set]. Kaggle. Retrieved March 15, 2025, from <https://www.kaggle.com/datasets/jessicali9530/celeba-dataset>
14. Keerthi, K., Harika, G., & Harshini, K. D. (2024). Gender classification optimization with thermal images using advanced neural networks. *International Journal of Engineering and Manufacturing*, 14(5), 36–45. <https://www.mecspress.org/ijem/ijem-v14-n5/IJEM-V14-N5-4.pdf>
15. Kethineni, K. (2024). Gender classification from thermal images using CNN. *ResearchGate*. https://www.researchgate.net/publication/384104739_Gender_Classification_from_The_rmal_Images_using_CNN
16. Akgül, İ. (2024). Deep convolutional neural networks for age and gender estimation using an imbalanced dataset of human face images. *Neural Computing and Applications*. <https://link.springer.com/article/10.1007/s00521-024-10390-0>
17. Senbagavalli, M., & Reddy, L. K. (2024). Age and gender prediction from facial images using Inception V3 architecture. In *2024 International Conference on Advanced Computing and Communication Technologies* (pp. 76–82). IEEE. <https://ieeexplore.ieee.org/abstract/document/10918427/>

18. Gurram, H., Kethineni, K., & Kommineni, D. H. (2024). Gender classification from thermal images using CNN. In *2024 Asia Pacific Conference on Image Processing and Analysis*. IEEE. <https://ieeexplore.ieee.org/abstract/document/10673553/>
19. Safoq, M. S. (2024). Human gender prediction by face images based on convolution neural network. *Journal of Kerbala University*. [DOI not available]
20. Bashir, S., & Khan, R. A. (2023). Gender classification using pre-trained CNN models with transfer learning. *Journal of Intelligent & Fuzzy Systems*, 44(6), 8731–8740. <https://doi.org/10.3233/JIFS-223170>
21. Комар М.П., Саченко А.О., Васильків Н.М., Гладій Г.М., Коваль В.С., Ліп'яніна-Гончаренко Х.В. Методичні рекомендації до виконання кваліфікаційної роботи з освітньо-професійної програми «Комп'ютерні науки» спеціальності 122 «Комп'ютерні науки» за першим (бакалаврським) рівнем вищої освіти. Тернопіль: ЗУНУ, 2024. 52 с.
22. Загальні методичні рекомендації з підготовки, оформлення, захисту та оцінювання кваліфікаційних робіт здобувачів вищої освіти першого (бакалаврського) і другого (магістерського) рівнів. Тернопіль: ЗУНУ, 2024. 83 с.
23. Божягора, М., & Дорош, В. (2025). Модуль розпізнавання статі на основі глибокої нейромережевої архітектури InceptionV3. У Збірник тез доповідей студентської науково-практичної конференції «Інтелектуальні інформаційні технології в прикладних дослідженнях» (ІТАР-2025) (с. 108–110)

Додаток А

Код реалізації модуля

```
import pandas as pd
import numpy as np
import cv2
import matplotlib.pyplot as plt
import seaborn as sns
from sklearn.metrics import f1_score

from keras.applications.inception_v3 import InceptionV3, preprocess_input
from keras import optimizers
from keras.models import Sequential, Model
from keras.layers import Dropout, Flatten, Dense, GlobalAveragePooling2D
from keras.callbacks import ModelCheckpoint
from keras.preprocessing.image import ImageDataGenerator, array_to_img, img_to_array, load_img
from keras.utils import np_utils
from keras.optimizers import SGD

from IPython.core.display import display, HTML
from PIL import Image
from io import BytesIO
import base64

plt.style.use('ggplot')

%matplotlib inline

# %% [code]
import tensorflow as tf
print(tf.__version__)

main_folder = '../input/celeba-dataset/'
images_folder = main_folder + 'img_align_celeba/img_align_celeba/'

EXAMPLE_PIC = images_folder + '000506.jpg'

TRAINING_SAMPLES = 10000
VALIDATION_SAMPLES = 2000
TEST_SAMPLES = 2000
IMG_WIDTH = 178
IMG_HEIGHT = 218
BATCH_SIZE = 16
NUM_EPOCHS = 20

df_attr = pd.read_csv(main_folder + 'list_attr_celeba.csv')
df_attr.set_index('image_id', inplace=True)
df_attr.replace(to_replace=-1, value=0, inplace=True) #replace -1 by 0
df_attr.shape

for i, j in enumerate(df_attr.columns):
    print(i, j)

img = load_img(EXAMPLE_PIC)
plt.grid(False)
plt.imshow(img)
df_attr.loc[EXAMPLE_PIC.split('/')[-1]][['Smiling', 'Male', 'Young']] #some attributes

plt.title('Female or Male')
sns.countplot(y='Male', data=df_attr, color="c")
plt.show()
```

```

df_partition = pd.read_csv(main_folder + 'list_eval_partition.csv')
df_partition.head()

df_partition['partition'].value_counts().sort_index()

df_partition.set_index('image_id', inplace=True)
df_par_attr = df_partition.join(df_attr['Male'], how='inner')
df_par_attr.head()

def load_reshape_img(fname):
    img = load_img(fname)
    x = img_to_array(img)/255.
    x = x.reshape((1,) + x.shape)

    return x

def generate_df(partition, attr, num_samples):
    """
    partition
    0 -> train
    1 -> validation
    2 -> test
    """

    df_ = df_par_attr[(df_par_attr['partition'] == partition)
                      & (df_par_attr[attr] == 0)].sample(int(num_samples/2))
    df_ = pd.concat([df_,
                     df_par_attr[(df_par_attr['partition'] == partition)
                                 & (df_par_attr[attr] == 1)].sample(int(num_samples/2))])

    # for Train and Validation
    if partition != 2:
        x_ = np.array([load_reshape_img(images_folder + fname) for fname in df_.index])
        x_ = x_.reshape(x_.shape[0], 218, 178, 3)
        y_ = np_utils.to_categorical(df_[attr],2)
    # for Test
    else:
        x_ = []
        y_ = []

        for index, target in df_.iterrows():
            im = cv2.imread(images_folder + index)
            im = cv2.resize(cv2.cvtColor(im, cv2.COLOR_BGR2RGB), (IMG_WIDTH, IMG_HEIGHT)).astype(np.float32) / 255.0
            im = np.expand_dims(im, axis=0)
            x_.append(im)
            y_.append(target[attr])

    return x_, y_

datagen = ImageDataGenerator(
    #preprocessing_function=preprocess_input,
    rotation_range=30,
    width_shift_range=0.2,
    height_shift_range=0.2,
    shear_range=0.2,
    zoom_range=0.2,
    horizontal_flip=True
)

```

```

img = load_img(EXAMPLE_PIC)
x = img_to_array(img)/255.
x = x.reshape((1,) + x.shape)

plt.figure(figsize=(20,10))
plt.suptitle('Data Augmentation', fontsize=28)

i = 0
for batch in datagen.flow(x, batch_size=1):
    plt.subplot(3, 5, i+1)
    plt.grid(False)
    plt.imshow( batch.reshape(218, 178, 3))

    if i == 9:
        break
    i += 1

plt.show()

x_train, y_train = generate_df(0, 'Male', TRAINING_SAMPLES)

train_datagen = ImageDataGenerator(
    preprocessing_function=preprocess_input,
    rotation_range=30,
    width_shift_range=0.2,
    height_shift_range=0.2,
    shear_range=0.2,
    zoom_range=0.2,
    horizontal_flip=True,
)

train_datagen.fit(x_train)

train_generator = train_datagen.flow(
    x_train, y_train,
    batch_size=BATCH_SIZE,
)

x_valid, y_valid = generate_df(1, 'Male', VALIDATION_SAMPLES)

inc_model = InceptionV3(weights='../input/inceptionv3/inception_v3_weights_tf_dim_ordering_tf_kernels_notop.h5',
    include_top=False,
    input_shape=(IMG_HEIGHT, IMG_WIDTH, 3))

print("number of layers:", len(inc_model.layers))

x = inc_model.output
x = GlobalAveragePooling2D()(x)
x = Dense(1024, activation="relu")(x)
x = Dropout(0.5)(x)
x = Dense(512, activation="relu")(x)
predictions = Dense(2, activation="softmax")(x)

model_ = Model(inputs=inc_model.input, outputs=predictions)

for layer in model_.layers[:52]:
    layer.trainable = False

model_.compile(optimizer=SGD(lr=0.0001, momentum=0.9)
    , loss='categorical_crossentropy'
    , metrics=['accuracy'])

checkpointer = ModelCheckpoint(filepath='weights.best.inc.male.hdf5',

```

```

        verbose=1, save_best_only=True)

hist = model_.fit_generator(train_generator
    , validation_data = (x_valid, y_valid)
    , steps_per_epoch= TRAINING_SAMPLES/BATCH_SIZE
    , epochs= NUM_EPOCHS
    , callbacks=[checkpointer]
    , verbose=1
)

plt.figure(figsize=(18, 4))
plt.plot(hist.history['loss'], label = 'train')
plt.plot(hist.history['val_loss'], label = 'valid')
plt.legend()
plt.title('Loss Function')
plt.show()

plt.figure(figsize=(18, 4))
plt.plot(hist.history['acc'], label = 'train')
plt.plot(hist.history['val_acc'], label = 'valid')
plt.legend()
plt.title('Accuracy')
plt.show()

model_.load_weights('weights.best.inc.male.hdf5')

x_test, y_test = generate_df(2, 'Male', TEST_SAMPLES)

model_predictions = [np.argmax(model_.predict(feature)) for feature in x_test ]

test_accuracy = 100 * np.sum(np.array(model_predictions)==y_test) / len(model_predictions)
print("Model Evaluation")
print("Test accuracy: %.4f%%" % test_accuracy)
print("f1_score:", f1_score(y_test, model_predictions))

gender_target = {0: 'Female'
    , 1: 'Male'}

def img_to_display(filename):
    # inspired on this kernel:
    # https://www.kaggle.com/stassl/displaying-inline-images-in-pandas-dataframe
    # credits to stassl :)

    i = Image.open(filename)
    i.thumbnail((200, 200), Image.LANCZOS)

    with BytesIO() as buffer:
        i.save(buffer, 'jpeg')
        return base64.b64encode(buffer.getvalue()).decode()

def display_result(filename, prediction, target):
    """
    Display the results in HTML
    """

    gender = 'Male'
    gender_icon = "https://i.imgur.com/nxWan2u.png"

    if prediction[1] <= 0.5:
        gender_icon = "https://i.imgur.com/oAAb8rd.png"
        gender = 'Female'

```

```

display_html = '''
<div style="overflow: auto; border: 2px solid #D8D8D8;
padding: 5px; width: 420px;" >

<div style="padding: 10px 0px 0px 20px; overflow: auto;">

<h3 style="margin-left: 50px; margin-top: 2px;">{</h3>
<p style="margin-left: 50px; margin-top: -6px; font-size: 12px">{</p>
<p style="margin-left: 50px; margin-top: -16px; font-size: 12px">Real Target: {</p>
<p style="margin-left: 50px; margin-top: -16px; font-size: 12px">Filename: {</p>
</div>
</div>
'''

def gender_prediction(filename):
    im = cv2.imread(filename)
    im = cv2.resize(cv2.cvtColor(im, cv2.COLOR_BGR2RGB), (178, 218)).astype(np.float32) / 255.0
    im = np.expand_dims(im, axis=0)

    # prediction
    result = model_.predict(im)
    prediction = np.argmax(result)

    return result

df_to_test = df_par_attr[(df_par_attr['partition'] == 2)].sample(8)

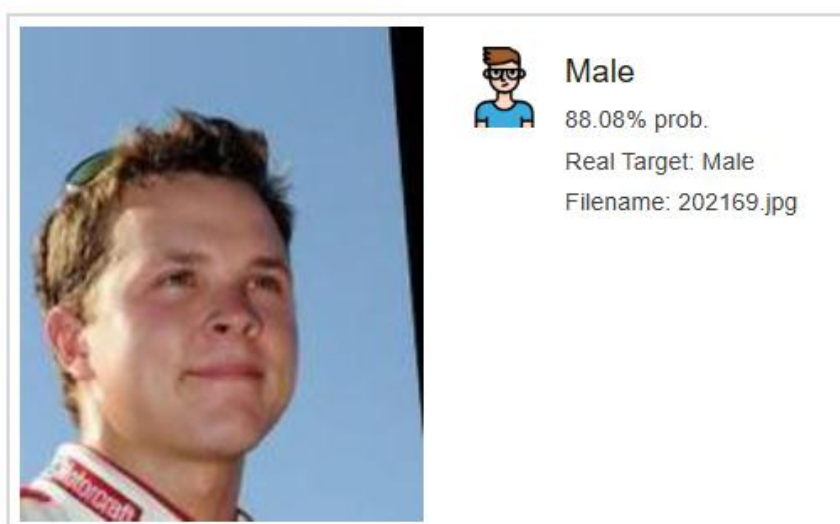
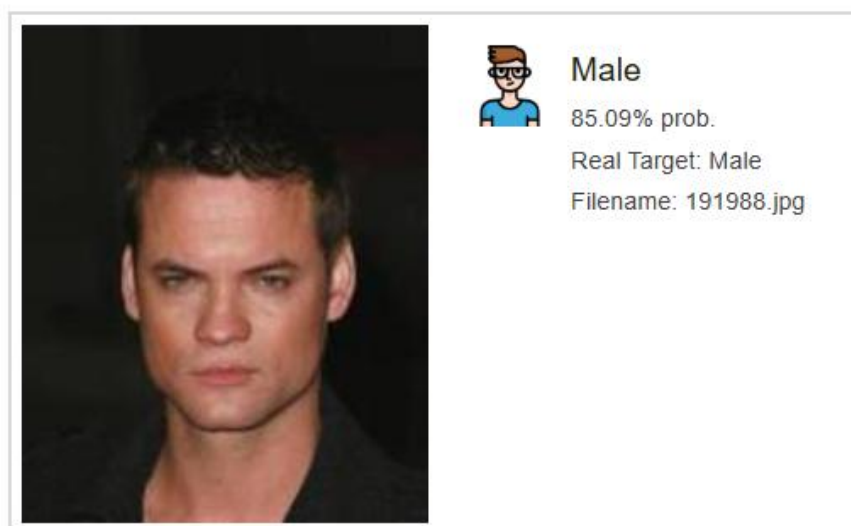
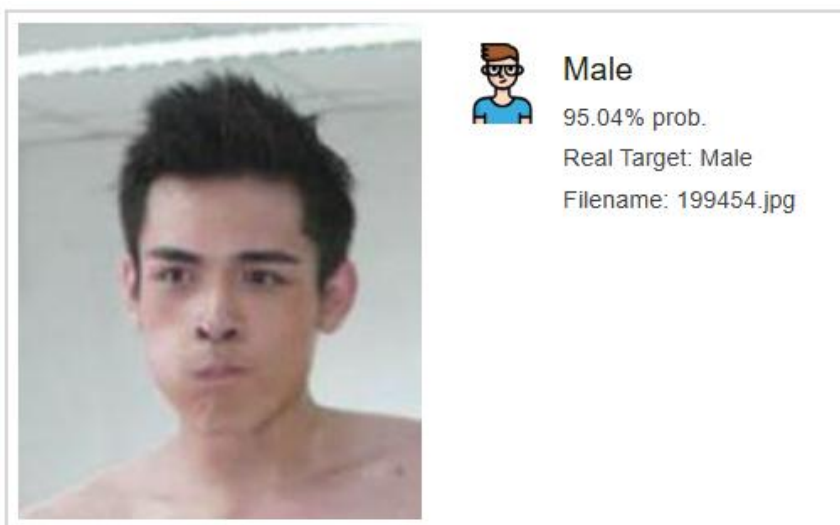
for index, target in df_to_test.iterrows():
    result = gender_prediction(images_folder + index)

    display_result(images_folder + index, result[0], target['Male'])

```

Додаток Б

Результати оцінки моделі





Female

100.00% prob.

Real Target: Female

Filename: 190210.jpg



Male

97.29% prob.

Real Target: Male

Filename: 192655.jpg



Female

100.00% prob.

Real Target: Female

Filename: 189685.jpg

Додаток В

Апробація отриманих результатів

Західноукраїнський національний університет

Факультет комп'ютерних інформаційних технологій

Кафедра інформаційно-обчислювальних систем і управління



ЗБІРНИК ТЕЗ ДОПОВІДЕЙ

Студентської науково-практичної конференції

ІНТЕЛЕКТУАЛЬНІ ІНФОРМАЦІЙНІ ТЕХНОЛОГІЇ В ПРИКЛАДНИХ
ДОСЛІДЖЕННЯХ

(ІІТАР-2025)

27-29 травня 2025 року

Тернопіль

2025

Божжгора Микола
студент групи КН-42
jonsons2031@gmail.com

Дорош Віталій
викладач

vdo@wunu.edu.ua

Західноукраїнський національний університет
Тернопіль, Україна

МОДУЛЬ РОЗПІЗНАВАННЯ СТАТІ НА ОСНОВІ ГЛИБОКОЇ НЕЙРОМЕРЕЖЕВОЇ АРХІТЕКТУРИ INCEPTIONV3

У сучасному світі розпізнавання статі на основі зображень є важливою задачею в таких галузях, як маркетинг, безпека, персоналізація сервісів та біометрична ідентифікація. Метою дослідження стало створення ефективного модуля класифікації статі на основі глибокого навчання із використанням архітектури InceptionV3. В основі модуля – навчання моделі на великому відкритому датасеті CelebA (202 599 зображень), а також оптимізація архітектури і гіперпараметрів.

Розроблений модуль реалізовано на базі бібліотек TensorFlow і Keras. Він включає попередню обробку зображень (аугментація, балансування), розділення даних (20k train / 5k val / 5k test), модифікацію стандартної InceptionV3 (додавання GAP, Dense, Dropout), навчання з використанням SGD ($\text{lr}=0.0001$) і регулярного моніторингу метрик. Максимальна точність моделі на перевірочних даних становила 95,75%, а на тестовій – 94,3%. Значення F1-score = 0,94, що свідчить про збалансовану класифікацію обох класів.

На рисунку 1 зображено діаграму класів (UML) реалізованого модуля, яка ілюструє взаємодію основних компонентів системи, зокрема DataManager, TrainingController, EvaluationManager та InceptionV3Model.

У таблиці 1 здійснено порівняння з іншими сучасними підходами глибокого навчання до задачі гендерної класифікації. Найвищу точність досягнуто WideResNet (97.2%) і ResNet152 (95.8%), однак модель InceptionV3 забезпечила оптимальний баланс між точністю, швидкістю та кількістю параметрів, необхідних для навчання.

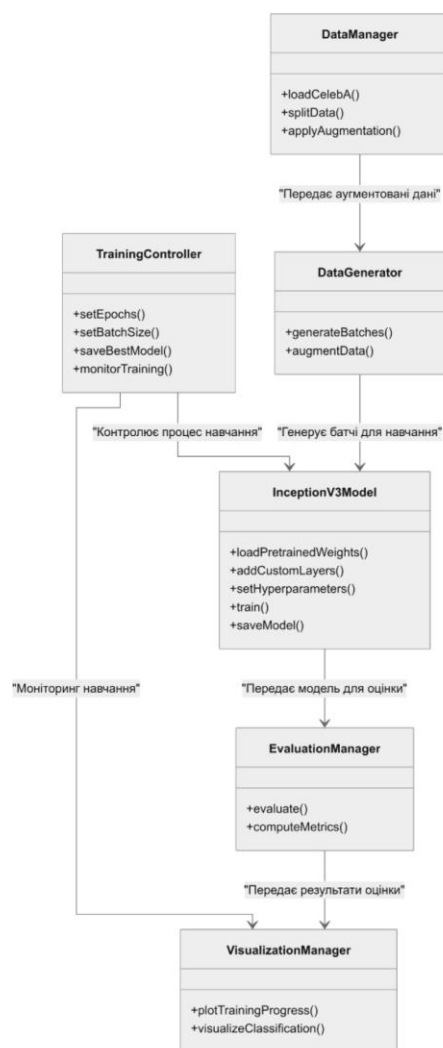


Рисунок 1 - UML-діаграма архітектури реалізації модуля розпізнавання статі

Таблиця 1 – Порівняння точності моделей глибокого навчання для класифікації статі

№	Модель	Набір даних	Точність (%)	Джерело
1	WideResNet	CelebA	97.2	Polina & Malathi [5]
2	ResNet152	CelebA	95.8	Wong et al. [10]
3	InceptionV3	UTKFace	93.2	Mohasseb et al. [1]
4	Пропонована	CelebA	94.3	Ця робота

Завдяки високій точності та стабільності, розроблений модуль є придатним для інтеграції у практичні системи відеоаналітики, кібербезпеки, адаптивного

UX-дизайну та цифрового маркетингу. Запропонована архітектура легко масштабована й оптимізована, що дозволяє її адаптувати для інших задач класифікації обличчя.

Список використаних джерел

1. Mohasseb, A., Amer, E., Chiroma, F., & Tranchese, A. (2025). Leveraging Advanced NLP Techniques and Data Augmentation to Enhance Online Misogyny Detection. *Applied Sciences*. DOI: 10.3390/app15020856.
2. Kwiatek, J., Leśna, M., Piskórz, W., et al. (2025). Comparison of the Diagnostic Accuracy of an AI-Based System for Dental Caries Detection and Clinical Evaluation Conducted by Dentists. *Journal of Clinical Medicine*. DOI: 10.3390/jcm14051566.
3. Ismaeel, A. Z., & Yasin, H. M. (2025). Machine Learning Based Finger Print Analysis for Gender Detection: A Review. *Asian Journal of Research in Computer Science*. Без DOI, доступно тут.
4. Khaliluzzaman, M., Hoque, M. D. J., et al. (2024). Revolutionizing Age and Gender Recognition: An Enhanced CNN Architecture. *International Journal of Online and Social Interaction*. DOI: 10.1109/IJOSI.2024.1133.
5. Polina, V. K., & Malathi, K. (2025). Classification of Target Customers for Online Advertising Using Wide ResNet CNN and Comparing Its Accuracy Over ELM-CNN Algorithm. *Research in Management, Accounting and Economics*. DOI: 10.4324/9781003606642-96.
6. Rezasoltani, A., Hosseini, A., & Toosi, R. (2024). A Multi-Task Framework Using Mamba for Identity, Age, and Gender Classification from Hand Images. *IEEE Transactions on Information and Communication Technologies*. DOI: 10.1109/TICT.2024.10892788.
7. Javed, N., Ahmed, T., & Faisal, M. (2024). An Efficacy Comparison of Supervised Machine Learning Classifiers for Cyberbullying Detection and Prediction. *International Journal of Bullying Prevention*. DOI: 10.1007/s42380-024-00282-1.
8. Ali, I., Maharani, A. A. S., & Suarna, N. (2025). Application of Decision Tree Algorithm to Improve Student Learning Pattern Classification Model at Sempoja Sip Perjuangan. *Journal of Artificial Intelligence and Education Applications*. DOI: 10.1016/j.jaiea.2025.12163.
9. Shekofteh, Y., & Moradi, A. (2024). APEDM: A New Voice Casting System Using Acoustic–Phonetic Encoder-Decoder Mapping. *Multimedia Tools and Applications*. DOI: 10.1007/s11042-024-20496-1.
10. Wong, J., Lu, S., Lou, Y., et al. (2024). A Novel Time Use Approach on Successful Aging: Racial and Gender Disparities in Daily Productive Engagement. *Innovation in Aging*. DOI: 10.1093/geroni/igb098.2201.