

Національний лісотехнічний університет України

(повне найменування вищого навчального закладу)

Навчально-науковий інститут комп'ютерних наук та інформаційних технологій

(повне найменування інституту)

Кафедра комп'ютерних наук

(повна назва кафедри)

Магістерська кваліфікаційна робота

другий (магістерський)

(рівень вищої освіти)

на тему: « Інтелектуальна система класифікації птахів з використанням
нейромережевого підходу »

Виконав: студент VI курсу групи КН-62м
спеціальності

122 “Комп'ютерні науки”

(шифр і назва напрямку підготовки, спеціальності)

Іванонько А.М.

(прізвище та ініціали)

Керівник Шиманський В.М.

(прізвище та ініціали)

Рецензент Генюш А.І.

(прізвище та ініціали)

Львів – 2025 року

Національний лісотехнічний університет України

(повне найменування вищого навчального закладу)

ННІ комп'ютерних наук та інформаційних технологій

Кафедра комп'ютерних наук

Рівень вищої освіти другий (магістерський)

Спеціальність 122 «Комп'ютерні науки»

(шифр і назва)

ЗАТВЕРДЖУЮ

Завідувач кафедри КН

Борецька І.Б.

“ 10 ” травня 2025 року

З А В Д А Н Н Я

НА МАГІСТЕРСЬКУ КВАЛІФІКАЦІЙНУ РОБОТУ СТУДЕНТУ

Іваноньку А.М.

(прізвище, ім'я, по батькові)

1. Тема роботи Інтелектуальна система класифікації птахів з використанням нейромережевого підходу

керівник роботи Шиманський В.М., к.т.н.

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

затверджені наказом вищого навчального закладу від “29” 04 2025 року № С-288

2. Термін подання студентом проекту (роботи) 10.12.2025

3. Вихідні дані до проекту (роботи) Аналіз шляхів вирішення задачі, датасет зображень різних птахів, описи класів птахів та їх характерні ознаки.

4. Зміст розрахунково-пояснювальної записки (перелік питань, які потрібно розробити)

4.1. Стан проблемної області.

4.2. Інформаційне забезпечення

4.3. Математичне забезпечення

4.4. Програмне забезпечення

4.5. Розроблення стартап проекту

4.6. Висновки

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

Слайди для доповіді (підготовка матеріалу для доповіді загальним обсягом 10-12 слайдів)

6. Дата видачі завдання 01.05.2025 року

КАЛЕНДАРНИЙ ПЛАН

№з/п	Назва етапів дипломної роботи	Термін виконання	Відмітка про виконання
1	З'ясування загальної постановки завдання; розроблення першого розділу пояснювальної записки	02.05.25	виконано
2	Проектування інтелектуальної та математичної моделі, розроблення конструкторської документації.	23.09.25	виконано
3	Реалізація програмного забезпечення та налаштування інтелектуальної моделі.	18.11.25	виконано
4	Висновки.	10.12.25	виконано

Студент: Іванонько А.М.



(підпис)

Керівник роботи: Шиманський В.М.



(підпис)

ТЕХНІЧНЕ ЗАВДАННЯ

Проаналізувати існуючі види нейронних меж, що використовуються для вирішення задач класифікації зображень. Вибрати та аргументувати структуру та тип нейронної межі, засобами якої буде проводитись дослідження. Підготувати навчальну вибірку для реалізації нейронної мережі. Розробити інтелектуальну систему класифікації птахів. Дослідити швидкодію реалізації алгоритмів навчання. Проаналізувати отримані результати.

Розроблена інтелектуальна система повинна задовольняти наступним вимогам:

- мати зручний та інтуїтивно зрозумілий інтерфейс;
- надавати можливість навчати та адаптувати інтелектуальну систему;
- на основі вхідних даних класифікувати зображення птахів.

АНОТАЦІЯ

Дипломна робота містить 64 сторінки пояснювальної записки, 8 рисунків, 1 таблицю, 1 додаток, 11 джерел.

У цій роботі розроблено інтелектуальну систему класифікації птахів на основі нейромережевого підходу. Система використовує глибоке навчання, зокрема згорткові нейронні мережі (CNN), для автоматичного розпізнавання та класифікації видів птахів за зображеннями.

Проведено аналіз існуючих методів класифікації, обрано оптимальну архітектуру моделі, а також сформовано та підготовлено датасет, що охоплює різноманітні види птахів. Ефективність системи перевірено за допомогою стандартних метрик, таких як точність, повнота та F-міра.

Результати експериментів демонструють високу точність класифікації, що свідчить про доцільність застосування нейромережевих технологій у задачах орнітологічного аналізу та автоматичного розпізнавання видів птахів. Розроблена система може бути використана як допоміжний інструмент у наукових дослідженнях, екологічному моніторингу та освітніх проєктах.

Ключові слова: птах, класифікація, CNN, F-міра.

ABSTRACT

This thesis contains 64 pages of explanatory note, 8 figures, 1 table, 1 appendix, 11 sources.

In this work, an intelligent bird classification system based on a neural network approach is developed. The system uses deep learning, in particular convolutional neural networks (CNN), for automatic recognition and classification of bird species from images.

An analysis of existing classification methods was conducted, the optimal model architecture was selected, and a dataset covering various bird species was generated and prepared. The effectiveness of the system was tested using standard metrics such as accuracy, completeness, and F-measure.

The experimental results demonstrate high classification accuracy, which indicates the feasibility of using neural network technologies in ornithological analysis and automatic recognition of bird species. The developed system can be used as an auxiliary tool in scientific research, environmental monitoring, and educational projects.

Keywords: bird, classification, CNN, F-measure.

ЗМІСТ

ПЕРЕЛІК СКОРОЧЕНЬ ТА УМОВНИХ ПОЗНАЧЕНЬ	9
ВСТУП	10
РОЗДІЛ 1. СТАН ПРОБЛЕМНОЇ ОБЛАСТІ.....	12
1.1 Аналіз традиційних методів класифікації птахів	12
1.2 Сучасні технології розпізнавання зображень у біології та екології.....	15
1.3 Використання нейронних мереж у задачах класифікації зображень	17
1.4. Висновки до розділу	19
РОЗДІЛ 2. ІНФОРМАЦІЙНЕ ЗАБЕЗПЕЧЕННЯ.....	21
2.1 Огляд джерел даних для класифікації птахів	21
2.2 Структура та формат вхідних даних	24
2.3. Висновки до розділу	27
РОЗДІЛ 3. МАТЕМАТИЧНЕ ЗАБЕЗПЕЧЕННЯ	28
3.1 Формалізація задачі класифікації птахів	28
3.2 Нейромережеві моделі та алгоритми навчання.....	30
3.3 Методи оцінки точності та валідації моделі.....	35
3.4. Висновки до розділу	38
РОЗДІЛ 4. ПРОГРАМНЕ ЗАБЕЗПЕЧЕННЯ	39
4.1 Архітектура розробленої системи	39
4.2 Аугментація даних	42
4.3 Аналіз отриманих результатів	44
4.4. Висновки до розділу	48
РОЗДІЛ 5. РОЗРОБЛЕННЯ СТАРТАП-ПРОЄКТУ	49
5.1. Опис ідеї проекту	49
5.2. Концепція стартапу.....	49
5.3. Бізнес-модель та стратегія монетизації.....	51
5.4. План реалізації та управління проектом.....	52
5.5. Висновки до розділу	54
ВИСНОВКИ.....	56
СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ.....	58

ПЕРЕЛІК СКОРОЧЕНЬ ТА УМОВНИХ ПОЗНАЧЕНЬ

AI	Artificial Intelligence – штучний інтелект
CNN	Convolutional Neural Network – згорткова нейронна мережа
ANN	Artificial Neural Network – штучна нейронна мережа
ML	Machine Learning – машинне навчання
DL	Deep Learning – глибоке навчання
Accuracy	Точність – частка правильно класифікованих зразків
Precision	Точність класифікації – частка істинно позитивних серед усіх позитивних результатів
Recall	Повнота – частка істинно позитивних серед усіх реальних позитивних
F1-score	Гармонічне середнє між precision і recall

ВСТУП

Актуальність дослідження. У сучасних умовах активного розвитку штучного інтелекту та комп'ютерного зору зростає потреба у впровадженні інтелектуальних систем для автоматизації процесів аналізу візуальної інформації. Однією з актуальних задач є класифікація об'єктів живої природи, зокрема птахів, що має важливе значення для біоекологічного моніторингу, збереження біорізноманіття, орнітологічних досліджень та освітніх ініціатив.

Ручне розпізнавання видів птахів вимагає високого рівня експертних знань і значних часових ресурсів. Застосування глибоких нейронних мереж, зокрема згорткових нейронних мереж (CNN), дає змогу суттєво підвищити точність і ефективність автоматичної класифікації зображень. Сучасні алгоритми глибокого навчання демонструють високі результати у завданнях розпізнавання об'єктів, що відкриває широкі можливості для їх застосування в екологічних дослідженнях та природоохоронних проектах.

Таким чином, розробка інтелектуальної системи класифікації птахів на основі нейромережевого підходу є актуальним і перспективним напрямом, який поєднує в собі новітні досягнення в галузі штучного інтелекту з практичними потребами екології, біології та інформаційних технологій.

Об'єкт дослідження – процес класифікації птахів.

Предмет дослідження – методи та алгоритми машинного навчання, що застосовуються для класифікації птахів.

Метою роботи є розробка та дослідження інтелектуальної системи класифікації птахів на основі нейромережевого підходу з використанням методів глибокого навчання для автоматичного розпізнавання видів птахів за зображеннями.

Для досягнення мети необхідно вирішити ряд **задач**, а саме:

- Провести аналіз сучасних методів та інструментів класифікації зображень, з особливим акцентом на згорткові нейронні мережі (CNN).

- Зібрати, сформувати та підготувати датасет зображень птахів для навчання та тестування моделі.
- Обрати оптимальну архітектуру нейронної мережі для задачі класифікації птахів.
- Реалізувати процес навчання нейромережі з використанням сучасних фреймворків глибокого навчання.
- Провести оцінювання ефективності роботи моделі за допомогою відповідних метрик (точність, повнота, F1-міра тощо).
- Проаналізувати отримані результати та сформулювати рекомендації щодо подальшого вдосконалення системи.

Наукова новизна дослідження - запропоновано та реалізовано інтелектуальну систему класифікації птахів, яка використовує сучасні методи глибокого навчання, зокрема згорткові нейронні мережі (CNN), для автоматичного розпізнавання видів за зображеннями.

Практична значимість полягає у тому, що результати проведеного дослідження мають високу практичну цінність, оскільки розроблена інтелектуальна система класифікації птахів може бути використана в реальних умовах для автоматичного розпізнавання видів за зображеннями. Така система здатна суттєво полегшити роботу орнітологів, біологів, екологів, а також сприяти ефективному моніторингу стану біорізноманіття.

РОЗДІЛ 1. СТАН ПРОБЛЕМНОЇ ОБЛАСТІ

1.1 Аналіз традиційних методів класифікації птахів

Класифікація птахів є важливою складовою орнітології — галузі біології, яка вивчає птахів, їхню поведінку, поширення, еволюцію, екологію та взаємодію з середовищем. Протягом десятиліть вчені використовували традиційні методи ідентифікації птахів, які ґрунтувалися переважно на морфологічних, поведінкових та акустичних характеристиках. Незважаючи на наявність сучасних технологій, ці методи досі залишаються основою багатьох досліджень, проте мають низку обмежень, що зумовлює потребу в автоматизації процесів класифікації з використанням штучного інтелекту.

Найбільш традиційний підхід до класифікації птахів полягає в аналізі морфологічних ознак. До них належать форма і розмір тіла, забарвлення оперення, форма дзьоба, крил і хвоста, а також розташування і колір очей та інших частин тіла. Ці характеристики дозволяють дослідникам визначати видову належність особин під час польових спостережень або в лабораторних умовах.

Історично морфологічний аналіз був єдиним способом ідентифікації птахів. Дослідники створювали атласи та польові визначники, які містили докладні ілюстрації та описи видів. Найвідомішими прикладами є польові путівники, розроблені Роже Торе Петерсоном та Девідом Сіблі, які стали класичними джерелами для орнітологів і спостерігачів за птахами.

Однак морфологічний підхід має кілька суттєвих обмежень. По-перше, він вимагає глибоких знань і досвіду. Різні види птахів можуть мати схожі ознаки, що ускладнює їх точне визначення, особливо в польових умовах. По-друге, у птахів часто спостерігається статевий диморфізм і сезонні зміни забарвлення, що також ускладнює ідентифікацію. Крім того, в умовах поганої видимості або при короткочасному контакті з об'єктом ідентифікація може бути неточною.

Акустичні сигнали є ще одним важливим джерелом інформації для класифікації птахів. Спів птахів, крики, заклики — всі ці звуки мають видову

специфіку і можуть використовуватись для визначення виду. Орнітологи часто записують звуки в природному середовищі з метою їх подальшого аналізу та порівняння зі зразками, що містяться в спеціалізованих базах даних, таких як Xeno-Canto чи Macaulay Library.

Перевага акустичних методів полягає в тому, що вони дозволяють ідентифікувати птахів без необхідності візуального контакту. Це особливо цінно в умовах щільної рослинності або в нічний час, коли птахів важко побачити. Проте й цей метод має недоліки: деякі птахи можуть мати подібні вокалізації, індивідуальні варіації звуків у межах одного виду, а також вплив шуму середовища на якість запису.

Класичні методи аналізу звуків включають спектрографічне порівняння та ручну ідентифікацію аудіосигналів. Хоча з розвитком цифрових технологій з'явилися програми, що допомагають автоматизувати цей процес (наприклад, Raven Pro), повна автоматизація за традиційною схемою все ще є малоефективною.

Поведінкові характеристики — такі як спосіб польоту, харчування, манера пересування, сезонні міграції та шлюбні ритуали — також використовуються для класифікації птахів. Екологічні ознаки включають тип середовища проживання, ареал поширення, періодичність появи в певних регіонах тощо. Такі дані часто фіксуються під час тривалих польових досліджень та вимагають значного часу для збору.

Хоча поведінкові та екологічні характеристики можуть бути дуже інформативними, їх використання обмежене складністю спостереження та залежністю від зовнішніх умов. Наприклад, поведінка птахів може змінюватися під впливом людського фактора або змін клімату, що призводить до помилок у класифікації.

Традиційно класифікація птахів здійснюється за допомогою ручного порівняння спостережуваних ознак з описами у визначниках та атласах. Часто використовуються спеціальні таблиці, дихотомічні ключі, порівняльні зображення та фотографії. Універсальними джерелами знань виступають енциклопедії, наприклад, «Handbook of the Birds of the World».

Також використовуються журнальні публікації, звіти польових спостережень, чеклісти та бази даних (BirdLife International, eBird тощо), які систематизують інформацію про видовий склад фауни певного регіону.

Для польових досліджень часто застосовуються біноклі, телескопи, фотокамери з телеоб'єктивами, блокноти для нотаток і GPS-навігатори. Попри ефективність у зборі первинних даних, ці засоби не забезпечують автоматизації процесу ідентифікації, що робить їх менш ефективними в умовах великого обсягу даних.

Хоча традиційні методи залишаються корисними, вони мають кілька серйозних недоліків, які знижують їхню ефективність у сучасних умовах:

- Людський фактор. Ідентифікація залежить від досвіду спостерігача, що підвищує ймовірність помилок.
- Часозатратність. Потрібно багато часу на збір, обробку та аналіз інформації.
- Складність масштабування. Застосування традиційних методів у великих масштабах (наприклад, у проектах з біомоніторингу) є малоефективним.
- Необ'єктивність. Суб'єктивні оцінки під час спостережень можуть впливати на точність класифікації.
- Неможливість роботи в режимі реального часу. Традиційні методи не дозволяють здійснювати класифікацію миттєво, що є важливою умовою для деяких прикладних задач (наприклад, виявлення інвазивних видів).

У зв'язку з вищезазначеними обмеженнями дедалі більшої популярності набувають автоматизовані системи класифікації птахів. Завдяки розвитку комп'ютерного зору, глибокого навчання та нейронних мереж стало можливим створення моделей, здатних ідентифікувати види з високою точністю за зображенням або аудіозаписом. У порівнянні з традиційними методами, такі підходи забезпечують кращу масштабованість, швидкість та об'єктивність.

Проте варто зазначити, що сучасні інтелектуальні системи не повинні повністю замінювати традиційні методи, а радше доповнювати їх.

Найефективнішим є гібридний підхід, коли автоматичні системи використовуються як інструмент підтримки прийняття рішень для досвідчених орнітологів. Це дозволяє значно підвищити якість досліджень та знизити навантаження на фахівців.

1.2 Сучасні технології розпізнавання зображень у біології та екології

З початку XXI століття технології комп'ютерного зору зазнали значного розвитку, що дало потужний імпульс для їхнього впровадження в різних галузях науки та практики, зокрема в біології та екології. Розпізнавання зображень — це процес автоматичного виявлення, ідентифікації та класифікації об'єктів на цифрових зображеннях або відео. У сфері біологічних досліджень ці технології відкрили нові можливості для обробки великих обсягів даних, що отримуються з камер спостереження, дронів, фотопасток та мобільних пристроїв.

Комп'ютерний зір (computer vision) — це міждисциплінарна галузь, яка поєднує елементи штучного інтелекту, машинного навчання, математики, фізики та біоінформатики з метою створення систем, здатних інтерпретувати візуальну інформацію. Основні етапи обробки зображень включають попередню обробку, виділення ознак, класифікацію та інтерпретацію результатів.

У біології та екології комп'ютерний зір використовується для ідентифікації видів тварин, рослин, грибів, мікроорганізмів, аналізу поведінки, моніторингу середовища та оцінки біорізноманіття. У контексті орнітології технології розпізнавання зображень дозволяють автоматично ідентифікувати птахів на фото або відео, що значно прискорює та спрощує польові дослідження.

Значного прориву в якості розпізнавання зображень вдалося досягти завдяки використанню глибокого навчання (deep learning), зокрема згорткових нейронних мереж (CNN). Ці мережі здатні автоматично виділяти релевантні ознаки з зображень і класифікувати об'єкти з високою точністю. CNN

складаються з кількох згорткових, активаційних та пулінгових шарів, які дозволяють моделі поступово виявляти прості та складні візуальні патерни.

У проєктах класифікації біологічних об'єктів архітектури нейромереж на зразок VGGNet, ResNet, EfficientNet, Inception використовуються для досягнення високих результатів. Наприклад, ResNet забезпечує точність понад 90% у задачах класифікації зображень понад тисячі класів, що робить її придатною для розпізнавання широкого спектра птахів.

У сфері орнітології системи глибокого навчання стали основою інноваційних рішень, таких як мобільні застосунки для ідентифікації птахів (Merlin Bird ID, BirdNET, iNaturalist) та академічні дослідження. Ці системи можуть працювати на основі зображень, відео або аудіозаписів, часто поєднуючи кілька типів даних для підвищення точності.

Наприклад, BirdCLEF — міжнародний конкурс, що проводиться щороку в рамках LifeCLEF, сприяє розробці моделей для автоматичної класифікації птахів на основі зображень і звуків. Результати цього конкурсу демонструють, що сучасні моделі здатні досягати високої точності навіть при наявності тисяч видів та обмежених навчальних даних.

Для реалізації систем розпізнавання зображень активно використовуються програмні інструменти та фреймворки, такі як TensorFlow, Keras, PyTorch, OpenCV. Ці інструменти надають потужні API для побудови, навчання та тестування моделей глибокого навчання.

OpenCV дозволяє виконувати базову обробку зображень (фільтрацію, нормалізацію, сегментацію), тоді як TensorFlow і PyTorch забезпечують гнучкість в архітектурі нейромереж і мають потужну екосистему для розробки, тренування і деплоюменту моделей.

Системи автоматичного розпізнавання зображень знайшли широке застосування в екологічному моніторингу. У проєктах моніторингу диких тварин фотопастки збирають мільйони зображень, які автоматично класифікуються моделями CNN. Дослідницькі групи використовують дрони для зйомки колоній птахів або лісових масивів, після чого зображення аналізуються за допомогою штучного інтелекту. У програмах вивчення

популяції морських ссавців або птахів, які живуть у важкодоступних місцях, глибоке навчання дозволяє аналізувати великі масиви даних без втручання людини. Такі проєкти не тільки знижують витрати часу та ресурсів, але й підвищують об'єктивність аналізу.

Попри успіхи, сучасні технології мають певні виклики. Найбільшими з них є:

- Неоднорідність даних. Зображення птахів можуть мати різну якість, фон, освітлення та ракурси.
- Малі обсяги даних для рідкісних видів. Деякі птахи малофотографовані, що ускладнює навчання моделей.
- Проблема перенавчання (overfitting). Модель може навчитися розпізнавати не вид, а фонові особливості зображення.
- Необхідність якісної анотації даних. Процес маркування зображень є трудомістким.

Проте ці виклики поступово долаються завдяки розвитку методів аугментації, трансферного навчання, використання попередньо навчених моделей (pre-trained models) та активного навчання (active learning).

1.3 Використання нейронних мереж у задачах класифікації зображень

Нейронні мережі, а особливо їх різновиди, відомі як глибокі згорткові нейронні мережі (Convolutional Neural Networks — CNN), стали ключовим технологічним рішенням у сфері автоматизованої класифікації зображень. У контексті класифікації птахів ці методи дозволяють досягати високої точності ідентифікації навіть за умов складного фону, поганого освітлення чи часткового перекриття об'єкта на зображенні [5, 9].

Штучна нейронна мережа (ШНМ) — це модель, що імітує структуру та функціонування біологічного мозку. Вона складається з шарів нейронів: вхідного, прихованих та вихідного. Кожен нейрон виконує просту обчислювальну операцію: зважену суму вхідних сигналів та нелінійну функцію активації. В результаті мережа може апроксимувати складні функції і вирішувати задачі класифікації, регресії, кластеризації тощо.

У задачах обробки зображень найефективнішими виявилися згорткові нейронні мережі, які дозволяють автоматично виявляти характерні ознаки на зображеннях — контури, текстури, геометричні структури, кольори. Структура типової CNN включає такі компоненти:

- Згортковий шар (Convolutional Layer): відповідає за вилучення локальних ознак з піксельних даних.
- Пулінговий шар (Pooling Layer): зменшує розмірність простору ознак та підвищує стійкість моделі до варіацій вхідного зображення.
- Шари активації (ReLU, sigmoid, tanh): забезпечують нелінійність обчислень.
- Повнозв'язні шари (Fully Connected Layers): здійснюють інтеграцію вилучених ознак і приймають остаточне рішення про клас об'єкта.

Процес навчання полягає у мінімізації функції втрат (loss function), яка характеризує різницю між передбаченим і реальним результатом. Для цього застосовуються алгоритми оптимізації, такі як стохастичний градієнтний спуск (SGD), Adam, RMSprop. Основним механізмом є зворотне поширення помилки (backpropagation), що дозволяє коригувати ваги зв'язків між нейронами.

Навчання потребує великої кількості анотованих зображень. Для задачі класифікації птахів це означає необхідність мати достатньо зображень кожного виду з підписаними назвами. Зазвичай використовують спеціалізовані датасети: CUB-200 (Caltech-UCSD Birds), NABirds, iNaturalist та інші.

У сучасній практиці використовуються перевірені архітектури нейромереж, які продемонстрували високу ефективність у змаганнях з комп'ютерного зору (наприклад, ImageNet) [1, 2, 7]:

- LeNet-5: перша успішна CNN, застосовувалася для розпізнавання рукописних цифр.
- AlexNet: спричинила революцію у 2012 році, вигравши ImageNet Challenge.
- VGGNet: відзначається простою та симетричною архітектурою.
- GoogLeNet (Inception): ефективна поєднанням шарів різної ширини.

- ResNet: використовує залишкові зв'язки (residual connections), які дозволяють ефективно тренувати глибокі мережі.
- EfficientNet: оптимізована архітектура з високим співвідношенням точність–швидкість [10].

Ці моделі можуть бути використані як у чистому вигляді, так і в режимі трансферного навчання, що особливо корисно при обмеженій кількості даних.

Трансферне навчання (transfer learning) — це підхід, який дозволяє використовувати нейромережу, попередньо навченої на великому наборі даних (наприклад, ImageNet), для вирішення іншого завдання з меншою кількістю прикладів. Цей підхід передбачає «заморожування» нижніх шарів моделі (які витягують базові ознаки) та донавчання верхніх шарів для конкретного завдання класифікації [1, 6].

Для підвищення стійкості моделі та уникнення переобучення застосовуються техніки аугментації даних (data augmentation): обертання, масштабування, горизонтальне дзеркалювання, зміна яскравості, шум тощо. Вони дозволяють штучно збільшити розмір навчальної вибірки.

Регуляризація включає такі методи як Dropout (випадкове вимикання нейронів під час навчання), L2-регуляризація (штраф за великі ваги), Batch Normalization (нормалізація шарів).

Нейронні мережі, особливо згорткові архітектури, є провідними технологіями в задачах класифікації зображень. Їх застосування у біології та, зокрема, в орнітології дозволяє вирішувати складні задачі автоматизованого розпізнавання видів птахів з високою точністю. Використання технік трансферного навчання, аугментації та регуляризації дозволяє адаптувати ці моделі навіть для умов обмежених ресурсів. Подальший розвиток архітектур нейронних мереж та зростання обсягів доступних даних відкривають нові перспективи для створення інтелектуальних систем біологічного моніторингу, що матимуть значний вплив на охорону природи, дослідження біорізноманіття та екологічну освіту [3].

1.4. Висновки до розділу

Таким чином, традиційні методи класифікації птахів, засновані на морфологічних, акустичних і поведінкових ознаках, залишаються важливою частиною орнітологічної практики. Проте стрімке зростання обсягів даних, необхідність обробки великої кількості зображень та записів у стислі терміни вимагає впровадження нових, автоматизованих рішень. Саме тому розробка інтелектуальних систем класифікації птахів з використанням нейромережових підходів є не лише актуальною, а й логічним етапом еволюції методів у даній галузі.

Сучасні технології розпізнавання зображень, особливо на основі нейромереж, стали важливим інструментом для біології та екології. Їхнє застосування в орнітології відкриває нові можливості для швидкої, точної та масштабованої класифікації птахів. Висока точність, можливість роботи в реальному часі та інтеграція з мобільними пристроями роблять такі системи незамінними як для науковців, так і для громадських спостерігачів за природою.

РОЗДІЛ 2. ІНФОРМАЦІЙНЕ ЗАБЕЗПЕЧЕННЯ

2.1 Огляд джерел даних для класифікації птахів

Якість та обсяг даних є визначальним чинником успішного функціонування будь-якої інтелектуальної системи, зокрема, у задачах класифікації зображень птахів. Ефективність нейронної мережі безпосередньо залежить від кількості та різноманітності навчальних прикладів, а також від рівня деталізації супутньої анотації. Тому огляд і вибір джерел даних є одним із ключових етапів при розробці інтелектуальної системи класифікації птахів.

Класифікація птахів — складна задача, зумовлена великою різноманітністю видів, схожістю візуальних ознак деяких з них, варіативністю зовнішнього вигляду залежно від віку, статі, сезону або поведінки. Крім того, зображення часто містять фонові шуми, перекриття, погане освітлення або часткову видимість об'єкта. Усе це підвищує вимоги до обраного набору даних.

Ідеальний датасет для класифікації птахів повинен:

- Містити велику кількість зображень з достатньою кількістю прикладів для кожного виду.
- Включати різноманітні умови зйомки (кут огляду, освітлення, сезон, фон).
- Бути добре анотованим: з наявністю назв видів, GPS-координат, дати зйомки.
- Бути відкритим для використання або мати чіткі умови ліцензування.

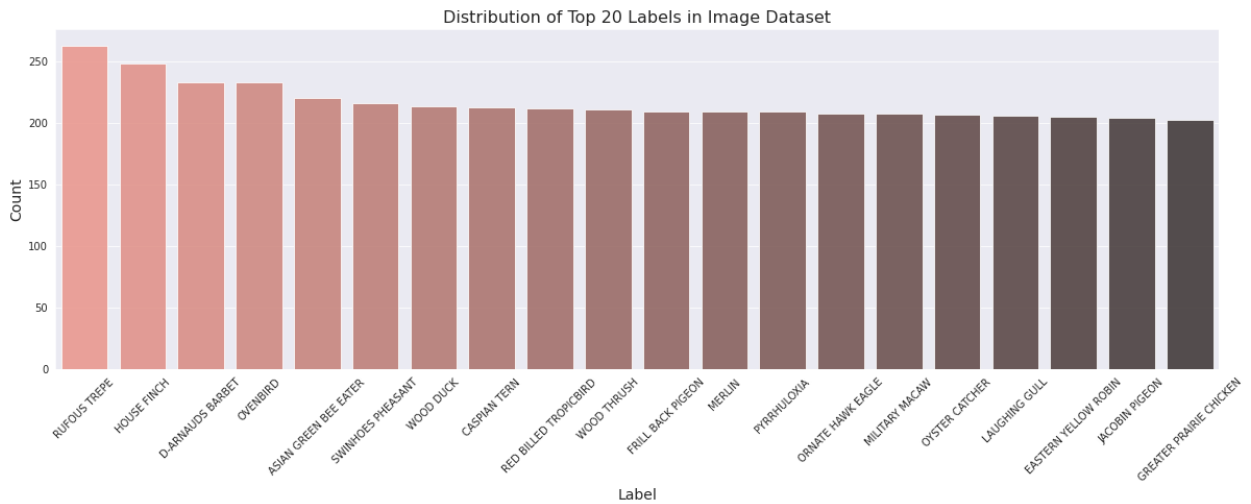


Рисунок 2.1 - Розподіл кількості зображень у датасеті (перші 20)

Один з найвідоміших і найчастіше використовуваних датасетів для класифікації птахів. Версія CUB-200-2011 включає 11 788 зображень, що охоплюють 200 різних видів птахів, які зустрічаються в Північній Америці. Кожне зображення має супровідну інформацію:

- Назва виду (англійською та латинською мовами);
- Bounding box координати для кожного птаха;
- 15 атрибутів (наприклад, колір крил, форма дзьоба);
- Часткова розмітка ключових точок (наприклад, очі, дзьоб, хвіст).
- CUB-200 є чудовим базовим датасетом для навчання, тестування та порівняння моделей глибокого навчання в задачах класифікації.

North American Birds Dataset містить понад 48 000 зображень, які представляють понад 400 видів птахів. Він був створений на базі ресурсів Cornell Lab of Ornithology та подібних джерел.

iNaturalist — громадянська наукова платформа для спостереження за живими організмами. Містить десятки тисяч зображень птахів, розміщених волонтерами з усього світу. Реальні, неідеалізовані умови зйомки. Дані супроводжуються інформацією про геолокацію, дату, ідентифікаторами спільноти. Багатий набір метаданих, але можуть бути помилки в анотаціях. Цей датасет надзвичайно корисний для моделювання систем, які працюють в умовах "дикої природи" (in-the-wild settings).

Створений у Стенфордському університеті Birdsnap датасет, містить понад 50 000 зображень для 500 видів птахів, що мешкають у США. Він особливо корисний для задач fine-grained image classification.

eBird — ще один великий ресурс від Cornell Lab, який містить записи спостережень, геолокації, зображення, звуки. Хоча eBird більше орієнтований на записи спостережень, його супровідні зображення також використовуються в задачах класифікації.

У випадку відсутності спеціалізованого датасету для регіональних чи маловивчених видів птахів виникає необхідність створення власного набору даних. Це може включати:

- Збір зображень з відкритих джерел: Flickr, Wikimedia Commons, Google Images, ресурси наукових журналів.
- Краудсорсинг: залучення громадських орнітологів до фотофіксації.
- Фотопастки: автоматизовані камери у природному середовищі.
- Ручна анотація: використання інструментів LabelImg, MakeSense.ai, CVAT тощо.

Перевага власного датасету полягає у точному контролі над якістю та структурою даних, хоча це вимагає значних ресурсів для збирання та маркування.

Незважаючи на великий обсяг існуючих датасетів, у практиці класифікації птахів є певні труднощі:

- Дисбаланс класів: деякі види представлені великою кількістю зображень, інші — лише кількома.
- Анотаційні помилки: помилки користувачів, невірно розмічені bounding boxes.
- Застарілість назв: зміни у класифікації видів не завжди враховуються в старих наборах.
- Схожість видів: морфологічно подібні птахи важко класифікуються навіть фахівцями.

У багатьох випадках оптимальним підходом є комбінування кількох датасетів та їх попередня уніфікація за розмірами, форматами, назвами видів.

Метадані, такі як геолокація, час зйомки, сезонність, часто мають велике значення в орнітології. Деякі види птахів мешкають лише в певних регіонах або з'являються там сезонно. Інтелектуальні системи, які враховують контекст (наприклад, використовуючи time-aware або geo-aware моделі), можуть значно підвищити точність класифікації.

Огляд джерел даних демонструє, що для створення ефективної системи класифікації птахів нейромережевого типу існує достатньо якісних відкритих датасетів, які можна використовувати без додаткових витрат. Однак у разі специфічних потреб або регіонального спрямування проєкту доцільним є створення або адаптація власного датасету. У будь-якому випадку, успішне функціонування системи значною мірою залежить від якості, повноти та достовірності вхідних даних, а також від того, наскільки добре вони відображають реальні умови, в яких буде використовуватися модель.

2.2 Структура та формат вхідних даних

У процесі розробки інтелектуальної системи класифікації птахів з використанням нейромережевого підходу, важливим етапом є ретельне визначення структури та формату вхідних даних. Цей аспект має прямий вплив на ефективність навчання, точність класифікації, стабільність моделі та здатність узагальнювати на нові приклади. Вхідні дані мають бути уніфіковані, якісно підготовлені, а також відповідати вимогам обраної архітектури нейронної мережі.

У системі класифікації птахів вхідними даними найчастіше виступають цифрові зображення, однак часто до них додаються й супутні метадані.



Рисунок 2.2 - Приклад зображень у датасеті

Таким чином, можна виділити такі основні типи вхідної інформації:

- Зображення (RGB-формат) – основне джерело інформації для розпізнавання зовнішніх морфологічних ознак.
- Bounding boxes або сегментовані маски – для задач локалізації або класифікації окремих особин на зображенні.
- Метадані – включають інформацію про геолокацію, дату зйомки, погодні умови, спостерігача, тип пристрою тощо.

- Анотації – мітки класів, до яких належить зображення (наприклад, "Turdus merula" або "Чорний дрізд").

Для навчання глибоких моделей зображення зазвичай мають бути представлені у форматах JPEG, PNG або TIFF. У практиці комп'ютерного зору найчастіше використовують JPEG або PNG через збалансоване співвідношення якості та розміру файлів. Незалежно від вихідного формату, зображення перетворюються у багатовимірні тензори з формою (висота, ширина, канали). Наприклад, для кольорового зображення розміром 224×224 пікселів це буде тензор з розмірами (224, 224, 3).

Перед подачею до моделі дані проходять попередню обробку. Це обов'язковий етап, оскільки нейромережі чутливі до масштабу вхідних значень.

- Масштабування пікселів: Зображення нормалізуються до діапазону [0, 1] або [-1, 1] шляхом ділення кожного значення пікселя на 255 або за допомогою обчислення z-нормалізації.
- Центрування: Часто виконується віднімання середнього значення пікселів, обчисленого на тренувальній вибірці.
- Зміна розміру: Для забезпечення уніфікованості зображення масштабуються до однакового розміру (наприклад, 224×224 або 299×299 пікселів, залежно від моделі).

Хоча візуальні дані є основними, додаткові метадані також можуть покращити точність класифікації, особливо у випадках, коли візуальна подібність між видами висока. Геолокація дозволяє виключити види, які не зустрічаються в певному регіоні. Сезонність допомагає врахувати міграцію птахів. Час доби або погодні умови можуть впливати на ймовірність появи певного виду. Такі дані можуть подаватися до моделі окремим каналом у вигляді вектора ознак (feature vector).

У зв'язку з обмеженістю кількості зображень у деяких класах, а також для покращення здатності моделі до узагальнення, застосовується аугментація:

- Повороти, масштабування, горизонтальні віддзеркалення.

- Зміна яскравості, контрасту, насиченості.
- Обрізка (cropping), накладання шуму.

Аугментація здійснюється в режимі реального часу під час навчання або попередньо при формуванні датасету.

2.3. Висновки до розділу

Від якості підготовки даних залежить кінцева продуктивність моделей машинного навчання, рівень інтерпретованості результатів та здатність системи до узагальнення. У сфері охорони здоров'я, де йдеться про людське життя, цей етап набуває критичного значення. Систематичне застосування вищезазначених процедур дозволяє створити достовірну, об'єктивну та ефективну модель, що здатна виявляти ризики розвитку діабету на ранніх стадіях.

РОЗДІЛ 3. МАТЕМАТИЧНЕ ЗАБЕЗПЕЧЕННЯ

3.1 Формалізація задачі класифікації птахів

Задача класифікації птахів у контексті інтелектуальної системи з використанням нейромережевого підходу належить до класу задач розпізнавання образів, де об'єкти мають складну структуру та високу варіативність ознак. Формалізація цієї задачі передбачає перехід від опису проблеми у природній мові до математичної моделі, яка дозволяє застосовувати алгоритмічні методи для її розв'язання. Основною метою є побудова функції відображення, яка на основі вхідних даних (ознак птаха) визначає його належність до певного класу.

Першим етапом формалізації є визначення множини об'єктів класифікації. Нехай існує множина $X = \{x_1, x_2, \dots, x_n\}$, де кожен елемент x_i представляє окремий екземпляр птаха. Кожен об'єкт описується вектором ознак $f_i = (f_{i1}, f_{i2}, \dots, f_{im})$, де m – кількість ознак. Ознаки можуть бути різної природи: морфологічні (розмір тіла, довжина крил, форма дзьоба), кольорові характеристики оперення, акустичні параметри співу, а також ознаки, отримані з зображень (текстурні, контурні, спектральні). Таким чином, кожен птах у системі представлений точкою у багатовимірному просторі ознак R^m .

Далі визначається множина класів $C = \{c_1, c_2, \dots, c_k\}$, де k – кількість видів птахів, які необхідно розпізнати. Задача класифікації полягає у побудові відображення $\Phi : R^m \rightarrow C$, яке кожному вектору ознак ставить у відповідність один із класів. У випадку використання нейронних мереж це відображення реалізується у вигляді параметризованої функції $\Phi(f; \theta)$, де θ – набір параметрів моделі (вагові коефіцієнти та зміщення).

Особливістю задачі класифікації птахів є висока міжкласова подібність та внутрішньокласова варіативність. Наприклад, два різні види можуть мати схожі кольорові характеристики, але відрізнятися формою дзьоба або співом. Крім того, один і той самий вид може демонструвати значні відмінності залежно від віку, статі чи сезону. Це ускладнює побудову ефективної моделі та вимагає ретельного вибору ознак і методів їх обробки.

Важливим аспектом формалізації є визначення типу вхідних даних. У сучасних системах класифікації птахів найчастіше використовуються зображення, отримані з камер або мобільних пристроїв. Зображення є складним джерелом інформації, оскільки містить велику кількість піксельних даних, які необхідно перетворити у компактне представлення. Для цього застосовуються методи комп'ютерного зору, зокрема згорткові нейронні мережі (CNN), які здатні автоматично виділяти релевантні ознаки з зображень. Таким чином, формалізація задачі передбачає перехід від сирих піксельних даних до векторів ознак, що є вхідними для класифікатора.

Математично процес класифікації можна описати як задачу мінімізації функції втрат. Нехай y_i – істинна мітка класу для об'єкта x_i , а $y_i = \Phi(f_i; \theta)$ – прогноз моделі. Тоді функція втрат $L(\theta)$ визначається як міра розбіжності між y_i для всіх об'єктів навчальної вибірки. Найчастіше використовується крос-ентропія, яка для багатокласової класифікації має вигляд:

$$L(\theta) = - \sum_{i=1}^n \sum_{j=1}^k y_{ij} \log(\hat{y}_{ij}), \quad (3.1)$$

Ще одним важливим елементом формалізації є розподіл даних на навчальну, валідаційну та тестову вибірки. Це необхідно для оцінки узагальнюючої здатності моделі та запобігання перенавчанню. Формально, множина X розбивається на три підмножини: X_{train} , X_{val} , X_{test} . Кожна підмножина використовується для різних етапів навчання та оцінки моделі.

Важливо також врахувати проблему дисбалансу класів, яка часто виникає у задачах класифікації птахів. Деякі види можуть бути представлені великою кількістю зображень, тоді як інші – лише кількома прикладами. Це призводить до зміщення моделі у бік більш представлених класів. Математично дисбаланс можна компенсувати шляхом введення вагових коефіцієнтів у функцію втрат або застосування методів аугментації даних, що збільшують кількість прикладів рідкісних класів.

Формалізація задачі також передбачає визначення критеріїв якості класифікації. Найпоширенішими метриками є точність (accuracy), повнота (recall), точність у вузькому сенсі (precision) та F-міра. Для багатокласової

класифікації ці метрики обчислюються на основі матриці неточностей, яка відображає кількість правильних і неправильних прогнозів для кожного класу. Математично точність визначається як:

$$Accuracy = \frac{\text{кількість правильних прогнозів}}{\text{загальна кількість прогнозів}}. \quad (3.2)$$

Однак у випадку дисбалансу класів точність може бути недостатньо інформативною, тому використовуються додаткові метрики, такі як макро- та мікро-усереднені значення F-міри.

Ще одним аспектом є врахування стохастичної природи даних. Зображення птахів можуть містити шум, артефакти, різні умови освітлення та фонові об'єкти. Це означає, що вхідні дані є випадковими величинами, а задача класифікації має ймовірнісний характер. Формально, кожен вектор ознак f_i можна розглядати як реалізацію випадкового процесу, а прогноз моделі – як оцінку умовної ймовірності $P(c_j | f_i)$. Таким чином, нейронна мережа фактично апроксимує функцію умовного розподілу класів.

Узагальнюючи, формалізація задачі класифікації птахів включає кілька ключових етапів: визначення множини об'єктів та ознак, побудова математичної моделі відображення ознак у класи, вибір функції втрат та критеріїв якості, розподіл даних на вибірки, а також врахування специфічних проблем, таких як дисбаланс класів та стохастичність даних. Ця формалізація створює основу для подальшого застосування нейромережових методів, які дозволяють ефективно розв'язувати задачу класифікації у складних умовах реального світу.

3.2 Нейромережові моделі та алгоритми навчання

У сучасних системах класифікації зображень, зокрема для розпізнавання видів птахів, нейронні мережі стали основним інструментом завдяки своїй здатності автоматично виділяти складні ознаки та узагальнювати інформацію з великих обсягів даних. Формалізація нейромережових моделей передбачає визначення архітектури, вибір функцій активації, методів оптимізації та

алгоритмів навчання, які забезпечують ефективне наближення функції відображення ознак у класи.

Нейронна мережа є параметризованою нелінійною функцією, що складається з послідовності шарів, кожен з яких виконує перетворення вхідних даних. Базовим елементом є штучний нейрон, який обчислює зважену суму вхідних сигналів та застосовує функцію активації. Математично вихід нейрона описується як:

$$z = \sum_{i=1}^m w_i x_i + b, \quad a = \sigma(z), \quad (3.3)$$

де x_i – вхідні ознаки, w_i – вагові коефіцієнти, b – зміщення, а $\sigma(\cdot)$ – функція активації. У багатошаровій мережі вихід одного шару є входом для наступного, що дозволяє моделі будувати складні нелінійні залежності.

Для задачі класифікації птахів найчастіше застосовуються згорткові нейронні мережі (CNN), які спеціалізуються на обробці зображень. Їх ключова особливість полягає у використанні згорткових шарів, що виконують локальне перетворення даних за допомогою фільтрів, здатних виділяти текстурні та контурні ознаки. Згорткові шари зменшують кількість параметрів порівняно з повнозв'язними шарами, що робить модель більш ефективною для великих зображень. Крім того, CNN використовують операції субдискретизації (pooling), які знижують розмірність даних та підвищують інваріантність до зсувів і масштабування [4].

Архітектура нейромережі для класифікації птахів зазвичай включає кілька згорткових блоків, за якими слідує повнозв'язні шари для інтеграції ознак та прийняття рішення. Останній шар має кількість нейронів, що дорівнює числу класів, і використовує функцію активації softmax для перетворення вихідних значень у ймовірності належності до кожного класу:

$$\hat{y}_j = \frac{e^{z_j}}{\sum_{k=1}^K e^{z_k}}, \quad (3.4)$$

де z_j – вихідне значення для класу j , а K – кількість класів.

Вибір функцій активації є критично важливим для ефективності навчання. У сучасних нейронних мережах найчастіше застосовуються функції ReLU (Rectified Linear Unit), які мають вигляд:

$$\sigma(z) = \max(0, z), \quad (3.5)$$

оскільки вони забезпечують швидку збіжність та зменшують проблему затухання градієнтів. Для вихідного шару багатокласової класифікації використовується softmax, а для проміжних шарів – ReLU або її модифікації (Leaky ReLU, ELU).

Навчання нейронної мережі полягає у мінімізації функції втрат, яка вимірює розбіжність між прогнозованими ймовірностями та істинними мітками класів. Найпоширенішою функцією втрат для багатокласової класифікації є крос-ентропія. Мінімізація здійснюється за допомогою алгоритмів оптимізації, що базуються на градієнтних методах. Базовим підходом є стохастичний градієнтний спуск (SGD), який оновлює параметри моделі за правилом:

$$\theta^{(t+1)} = \theta^{(t)} - \eta \nabla_{\theta} L(\theta), \quad (3.6)$$

де η – швидкість навчання, а $\nabla_{\theta} L(\theta)$ – градієнт функції втрат за параметрами. Для підвищення ефективності навчання застосовуються адаптивні методи, такі як Adam, RMSProp, Adagrad, які автоматично коригують швидкість навчання для кожного параметра на основі історії градієнтів.

Важливим аспектом є регуляризація, яка запобігає перенавчанню моделі. Найпоширенішими методами є L2-регуляризація (додавання штрафу за великі ваги), Dropout (випадкове вимкнення нейронів під час навчання) та аугментація даних (створення додаткових прикладів шляхом трансформацій зображень). Аугментація особливо важлива для класифікації птахів, оскільки дозволяє моделі бути стійкою до змін освітлення, масштабу та положення об'єкта.

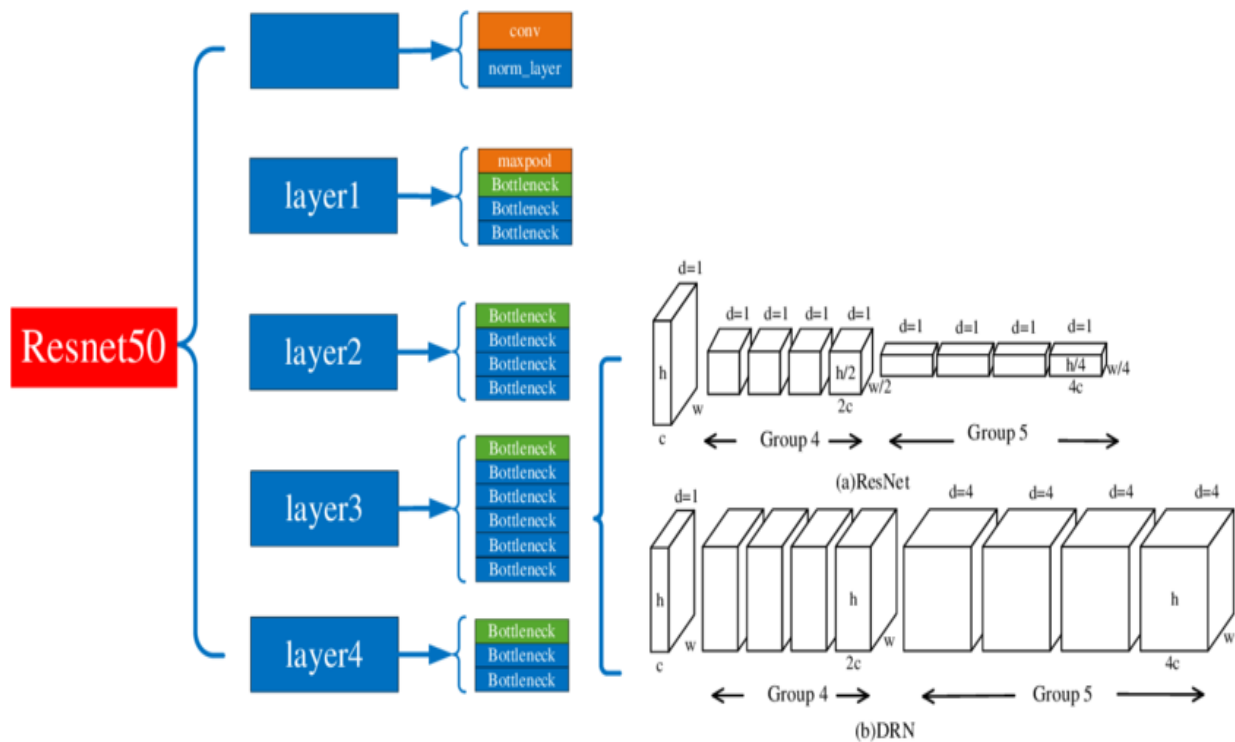


Рисунок 3.1 - Архітектура ResNet50

Процес навчання нейронної мережі включає кілька етапів: ініціалізацію параметрів, прямий прохід (forward pass), обчислення функції втрат, зворотне поширення помилки (backpropagation) та оновлення параметрів. Зворотне поширення помилки базується на правилі диференціювання складених функцій і дозволяє ефективно обчислювати градієнти для всіх параметрів мережі.

Оцінка якості навчання здійснюється на валідаційній вибірці, яка не використовується для оновлення параметрів. Це дозволяє контролювати узагальнюючу здатність моделі та своєчасно зупинити навчання у випадку перенавчання. Для цього застосовується метод ранньої зупинки (early stopping), який припиняє навчання, якщо точність на валідаційній вибірці перестає зростати протягом певної кількості епох.

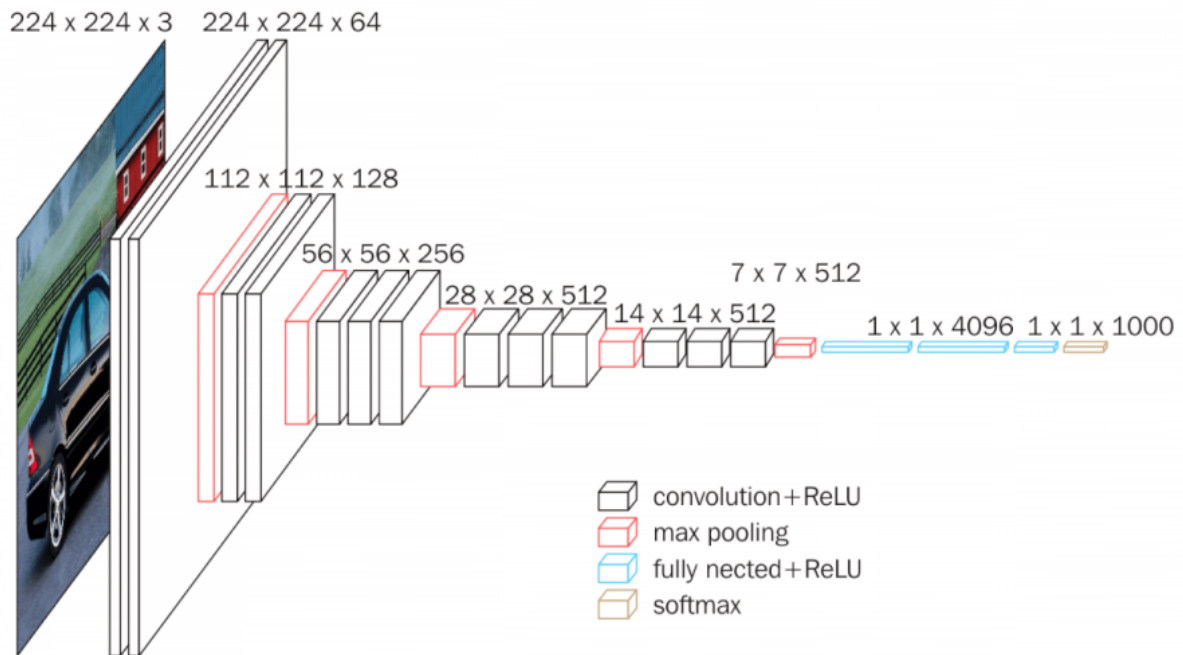


Рисунок 3.2 - Архітектура VGG16

Ще одним важливим аспектом є вибір глибини та ширини мережі. Занадто проста модель може мати низьку точність через недостатню здатність до узагальнення, тоді як надто складна – перенавчатися. Тому архітектура повинна бути збалансованою з урахуванням обсягу даних та обчислювальних ресурсів. У практиці класифікації птахів часто використовуються попередньо навчені моделі (ResNet, VGG, EfficientNet), які пройшли навчання на великих наборах даних, таких як ImageNet. Використання transfer learning дозволяє адаптувати ці моделі до конкретної задачі шляхом донавчання останніх шарів на спеціалізованому наборі зображень птахів.

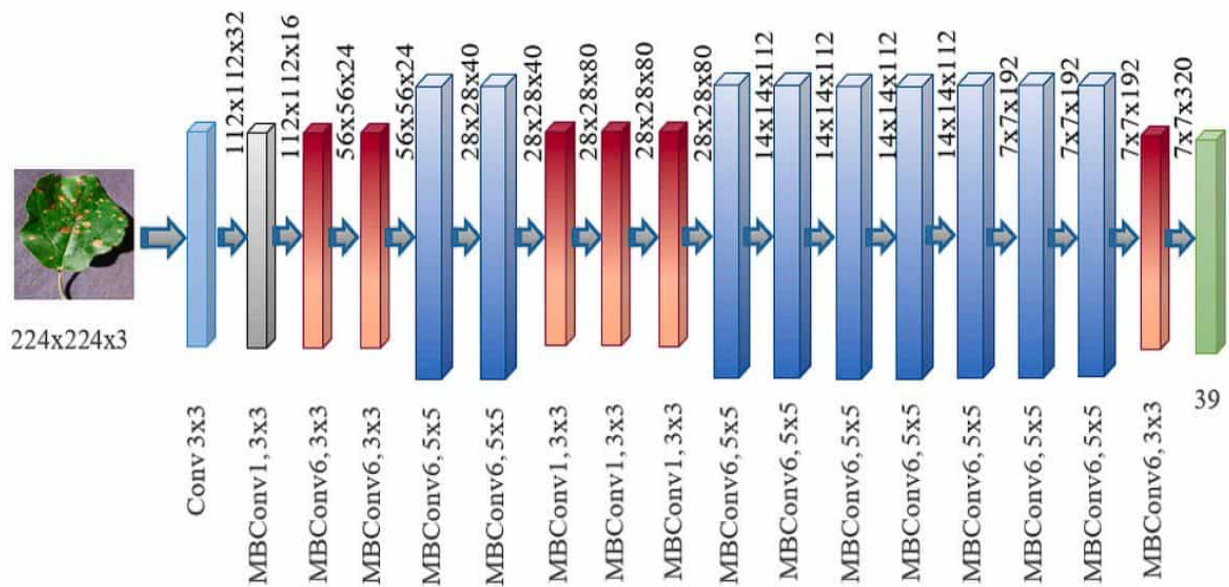


Рисунок 3.3 - Архітектура EfficientNet

Таким чином, нейромережеві моделі для класифікації птахів є складними системами, що поєднують багаторівневу обробку даних, нелінійні перетворення та стохастичні методи оптимізації. Вибір архітектури, функцій активації, алгоритмів навчання та методів регуляризації визначає ефективність системи та її здатність працювати в умовах реального світу, де дані є шумними, неоднорідними та обмеженими.

3.3 Методи оцінки точності та валідації моделі

Оцінка точності та валідація моделі є ключовими етапами побудови інтелектуальної системи класифікації птахів, оскільки вони визначають, наскільки добре модель узагальнює знання та працює на нових даних. Навчання нейронної мережі без належної перевірки може призвести до перенавчання, коли модель демонструє високу точність на навчальній вибірці, але не здатна коректно класифікувати нові приклади. Тому процес оцінки включає використання спеціальних метрик, методів розподілу даних та статистичних підходів для перевірки стабільності та надійності результатів.

Першим кроком є визначення метрик, які відображають якість класифікації. Найбільш поширеною метрикою є точність (accuracy), що

обчислюється як відношення кількості правильно класифікованих прикладів до загальної кількості прикладів. Формально:

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN}, \quad (3.7)$$

де TP (True Positive) – кількість правильно класифікованих позитивних прикладів, TN (True Negative) – правильно класифіковані негативні приклади, FP (False Positive) – помилкові позитивні класифікації, FN (False Negative) – помилкові негативні класифікації. Однак у задачі класифікації птахів, де кількість класів може бути великою і дані нерівномірно розподілені, точність не завжди є достатньо інформативною. Наприклад, якщо один клас домінує, модель може досягти високої точності, просто прогнозуючи цей клас для більшості прикладів.

Тому використовуються додаткові метрики, такі як precision (точність у вузькому сенсі), recall (повнота) та F-міра. Precision показує, яка частка прикладів, віднесених до певного класу, є правильними, а recall – яка частка прикладів цього класу була правильно знайдена. F-міра є гармонійним середнім між precision та recall і дозволяє збалансовано оцінити якість класифікації:

$$F_1 = 2 \cdot \frac{Precision \cdot Recall}{Precision + Recall}. \quad (3.8)$$

Для багатокласової класифікації ці метрики обчислюються для кожного класу окремо, а потім усереднюються за макро- або мікропідходом. Макроусереднення надає однакову вагу кожному класу, тоді як мікроусереднення враховує кількість прикладів у класах.

Ще одним важливим інструментом є матриця неточностей (confusion matrix), яка відображає кількість правильних і неправильних прогнозів для кожного класу. Вона дозволяє виявити, які класи найчастіше плутаються між собою, що особливо актуально для птахів зі схожими морфологічними ознаками. Аналіз матриці неточностей допомагає зрозуміти слабкі місця моделі та скоригувати архітектуру або набір ознак.

Окрім метрик, важливим етапом є валідація моделі, яка перевіряє її здатність узагальнювати знання на нових даних. Найпоширенішим методом є

розподіл даних на навчальну, валідаційну та тестову вибірки. Навчальна вибірка використовується для оновлення параметрів моделі, валідаційна – для контролю процесу навчання та налаштування гіперпараметрів, а тестова – для фінальної оцінки якості. Стандартним підходом є розподіл у пропорції 70/15/15 або 80/10/10 залежно від обсягу даних.

У випадку обмеженої кількості прикладів застосовується крос-валідація, яка дозволяє більш ефективно використовувати дані. Найпоширенішим варіантом є k-fold крос-валідація, де дані діляться на k рівних частин, і модель навчається k разів, кожного разу використовуючи одну частину для тестування, а решту – для навчання. Результати усереднюються, що забезпечує більш стабільну оцінку якості моделі.

Ще одним аспектом є валідація на основі реальних сценаріїв, коли тестування проводиться на даних, зібраних у різних умовах освітлення, фону або сезонних змін. Це дозволяє перевірити роботу моделі в умовах, максимально наближених до практичного використання. Для задачі класифікації птахів це особливо важливо, оскільки зовнішній вигляд птахів може змінюватися залежно від середовища та часу року.

Оцінка моделі також включає аналіз кривих ROC та AUC, які відображають співвідношення між чутливістю та специфічністю при різних порогах прийняття рішення. Хоча ці криві частіше застосовуються для бінарної класифікації, їх можна узагальнити для багатокласових задач шляхом побудови кривих для кожного класу окремо.

Важливим етапом є перевірка моделі на стійкість до шуму та аномалій. Для цього проводяться експерименти з додаванням випадкових перешкод до зображень або зміною їхніх характеристик. Якщо модель демонструє значне падіння точності, це свідчить про недостатню узагальнюючу здатність і потребу в додаткових методах регуляризації або аугментації даних.

Окрім кількісних метрик, застосовуються інтерпретаційні методи, які дозволяють зрозуміти, як модель приймає рішення. Наприклад, методи візуалізації активацій шарів або карти теплових зон (Grad-CAM) показують, на які ділянки зображення звертає увагу модель при класифікації. Це допомагає

перевірити, чи модель дійсно аналізує релевантні ознаки птаха, а не випадкові деталі фону.

Узагальнюючи, методи оцінки точності та валідації моделі включають комплексний підхід, що поєднує кількісні метрики, статистичні методи та інтерпретаційні інструменти. Вони забезпечують не лише перевірку ефективності моделі, але й її здатність працювати в реальних умовах, де дані є неоднорідними та містять шум. Правильна оцінка та валідація є запорукою створення надійної інтелектуальної системи класифікації птахів, здатної забезпечити високу точність і стабільність роботи.

3.4. Висновки до розділу

У процесі дослідження було визначено, що ефективність роботи системи залежить від трьох ключових компонентів: формалізації задачі, вибору нейромережевої моделі та алгоритмів навчання, а також методів оцінки точності та валідації.

На етапі формалізації задачі було встановлено математичну структуру проблеми, що включає визначення множини об'єктів, ознак та класів, а також побудову функції відображення, яка реалізує процес класифікації. Особливу увагу приділено проблемам високої міжкласової подібності та внутрішньокласової варіативності, що характерні для птахів, а також стохастичній природі даних. Було обґрунтовано необхідність використання багатовимірного простору ознак та ймовірного підходу до класифікації.

Згорткові нейронні мережі (CNN) визначено як оптимальний вибір для обробки зображень завдяки їх здатності автоматично виділяти релевантні ознаки. Розглянуто ключові елементи архітектури, функції активації, методи оптимізації та регуляризації, що забезпечують ефективне навчання моделі.

Було розглянуто основні метрики (accuracy, precision, recall, F-міра), матрицю неточностей, а також методи крос-валідації та аналізу стійкості моделі до шуму. Важливим аспектом є використання інтерпретаційних методів, що дозволяють оцінити релевантність ознак, які використовує модель для прийняття рішень.

РОЗДІЛ 4. ПРОГРАМНЕ ЗАБЕЗПЕЧЕННЯ

4.1 Архітектура розробленої системи

Архітектура інтелектуальної системи класифікації птахів, побудованої на основі нейромережевого підходу, є багаторівневою структурою, що поєднує модулі обробки даних та нейромережу. Основною метою архітектури є забезпечення ефективного потоку даних від моменту отримання зображення до видачі результату класифікації. Система реалізована з використанням сучасних бібліотек глибокого навчання, зокрема TensorFlow та Keras, що дозволяє інтегрувати попередньо навчені моделі та адаптувати їх до специфічних задач [8, 11].

Базовим компонентом архітектури є нейромережевий класифікатор, побудований на основі моделі EfficientNetB0, яка використовується як екстрактор ознак. Ця модель була попередньо навчена на великому наборі даних ImageNet, що забезпечує високу здатність до узагальнення та ефективно виділення релевантних ознак із зображень птахів. Використання попередньо навчених моделей дозволяє значно скоротити час навчання та підвищити точність класифікації, особливо у випадках обмеженого обсягу даних.

Архітектура системи побудована за принципом transfer learning, де нижні шари EfficientNetB0 залишаються замороженими, а верхні шари адаптуються до нової задачі. Це означає, що параметри базової моделі не змінюються під час навчання, що дозволяє зберегти знання, отримані на великому наборі даних, і уникнути перенавчання. На виході з базової моделі застосовується глобальний pooling, який перетворює багатовимірні тензори ознак у компактний вектор, що є основою для подальшої класифікації.

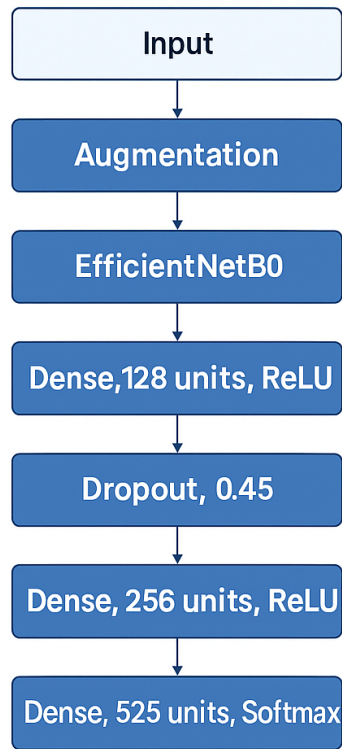


Рисунок 4.1 - Архітектура розробленої системи

Поверх екстрактора ознак побудовано кілька додаткових шарів, які виконують роль класифікаційного блоку. Першим є повнозв'язний шар із 128 нейронами та функцією активації ReLU, що забезпечує нелінійне перетворення ознак. Далі застосовується шар Dropout із коефіцієнтом 0.45, який випадково вимикає частину нейронів під час навчання, запобігаючи перенавчанню. Наступним є ще один повнозв'язний шар із 256 нейронами та активацією ReLU, після якого знову використовується Dropout. Завершує архітектуру вихідний шар із 525 нейронами та функцією активації softmax, що перетворює вихідні значення у ймовірності належності до кожного з 525 класів птахів.

Важливим елементом архітектури є модуль аугментації даних, який інтегрований у вхідний потік. Він виконує випадкові трансформації зображень, такі як обертання, масштабування, зміна яскравості та контрасту, що дозволяє моделі бути стійкою до варіацій у даних. Аугментація є критично важливою для задачі класифікації птахів, оскільки зображення можуть бути отримані в різних умовах освітлення та з різними ракурсами.

Архітектура системи також включає механізм компіляції моделі, де визначаються оптимізатор, функція втрат та метрики оцінки. Для оптимізації використовується алгоритм Adam із швидкістю навчання 0.0001, який забезпечує адаптивне коригування кроку градієнтного спуску для кожного параметра. Функція втрат – категоріальна крос-ентропія, що є стандартом для багатокласової класифікації. Як метрика використовується точність (accuracy), яка дозволяє оцінювати якість класифікації на кожній епосі навчання.

З точки зору програмної реалізації, архітектура системи побудована у вигляді моделі Keras, що забезпечує зручність у визначенні шарів та їх параметрів. Вхідні дані проходять через модуль аугментації, потім через заморожені шари EfficientNetB0, після чого обробляються додатковими повнозв'язними шарами. Вихід моделі передається у функцію softmax, яка генерує ймовірності для кожного класу. Така структура дозволяє легко інтегрувати модель у веб- або мобільний додаток, де користувач може завантажити зображення та отримати результат класифікації.

Архітектура системи передбачає можливість масштабування та адаптації. Наприклад, у випадку збільшення кількості класів або зміни набору даних можна модифікувати вихідний шар, не змінюючи базову модель. Крім того, система підтримує донавчання (fine-tuning), коли частина заморожених шарів розблоковується для додаткового навчання на нових даних, що дозволяє підвищити точність у специфічних умовах.

Загалом архітектура розробленої системи поєднує ефективність попередньо навчених моделей, гнучкість додаткових шарів та стійкість до варіацій у даних завдяки аугментації. Вона забезпечує високу точність класифікації при оптимальному використанні обчислювальних ресурсів, що робить її придатною для практичного застосування у реальних умовах. Використання сучасних фреймворків дозволяє легко розгортати систему на різних платформах, включаючи серверні рішення та мобільні пристрої, що відкриває широкі можливості для інтеграції у прикладні додатки.

4.2 Аугментація даних

Аугментація даних є одним із ключових етапів підготовки навчальної вибірки для нейромережових моделей, особливо у задачах комп'ютерного зору, де кількість доступних зображень може бути обмеженою або недостатньо різноманітною. У контексті інтелектуальної системи класифікації птахів аугментація відіграє вирішальну роль, оскільки зображення птахів часто отримуються в різних умовах освітлення, з різними ракурсами, фонами та якістю. Без застосування аугментації модель може перенавчитися на специфічні особливості навчальної вибірки та втратити здатність узагальнювати знання на нові дані.

Сутність аугментації полягає у штучному збільшенні обсягу навчальної вибірки шляхом створення модифікованих копій вихідних зображень. Ці модифікації можуть включати геометричні трансформації, зміни кольорових характеристик, додавання шуму та інші операції, що імітують реальні варіації даних. Таким чином, модель отримує можливість навчатися на більш різноманітних прикладах, що підвищує її стійкість до змін у середовищі та зменшує ризик перенавчання.

У розробленій системі аугментація реалізована за допомогою класу `ImageDataGenerator` з бібліотеки `Keras`, який забезпечує гнучкий механізм генерації пакетів зображень із застосуванням різних трансформацій у режимі реального часу. Це означає, що модифікації виконуються під час завантаження даних у модель, а не зберігаються у вигляді окремих файлів, що дозволяє економити дисковий простір та забезпечує динамічність процесу навчання.

Крім попередньої обробки, генератор даних підтримує розподіл вибірки на навчальну та валідаційну підмножини за допомогою параметра `validation_split`. У нашій системі він встановлений на рівні 0.2, що означає, що 20% даних використовуються для валідації, а решта – для навчання. Такий підхід дозволяє контролювати узагальнюючу здатність моделі та своєчасно виявляти ознаки перенавчання.

Процес генерації даних реалізується методом `flow_from_dataframe`, який дозволяє завантажувати зображення на основі інформації, що міститься у

датафреймі. У нашому випадку датафрейм містить шляхи до файлів зображень та відповідні мітки класів. Це забезпечує гнучкість у роботі з великими наборами даних, де структура каталогів може бути складною або нестандартною. Параметри методу визначають розмір цільового зображення (`target_size`), режим кольору (`color_mode='rgb'`), тип класифікації (`class_mode='categorical'`), розмір пакета (`batch_size`) та порядок завантаження (`shuffle=True`). Використання випадкового перемішування даних гарантує, що модель не запам'ятовує порядок прикладів, а навчається на різноманітних комбінаціях.

Аугментація даних у системі не обмежується лише базовими операціями. Хоча в наведеному коді явно не вказані додаткові параметри трансформацій, клас `ImageDataGenerator` підтримує широкий спектр можливостей, таких як обертання зображень на випадковий кут, горизонтальне та вертикальне віддзеркалення, масштабування, зсуви по ширині та висоті, зміна яскравості та контрасту. Ці операції можуть бути легко додані до конфігурації генератора для підвищення різноманітності навчальних прикладів. Наприклад, випадкове обертання зображення імітує різні ракурси зйомки птаха, а зміна яскравості – різні умови освітлення. Такі модифікації роблять модель більш стійкою до реальних варіацій, що зустрічаються у польових умовах.

Важливо зазначити, що аугментація виконується лише для навчальної вибірки, тоді як валідаційні та тестові дані залишаються незмінними. Це забезпечує об'єктивність оцінки моделі, оскільки вона перевіряється на даних, які не піддавалися штучним модифікаціям. Таким чином, результати тестування відображають реальну здатність моделі узагальнювати знання.

Аугментація даних має не лише технічне, але й концептуальне значення. Вона дозволяє моделі навчатися на більш широкому спектрі ознак, що знижує залежність від конкретних характеристик окремих зображень. У задачі класифікації птахів це особливо важливо, оскільки зовнішній вигляд птахів може змінюватися залежно від сезону, віку, статі та умов середовища. Без аугментації модель може сприймати ці зміни як ознаки різних класів, що призведе до помилок класифікації. Використання аугментації допомагає моделі

фокусуватися на стабільних ознаках, таких як форма дзьоба, структура крил або характерні кольорові патерни, ігноруючи випадкові варіації.

З точки зору продуктивності, аугментація даних у режимі реального часу є оптимальним рішенням, оскільки вона не потребує створення окремих копій зображень на диску. Це особливо актуально для великих наборів даних, де зберігання модифікованих зображень може займати значний обсяг пам'яті. Крім того, динамічна аугментація забезпечує більшу різноманітність прикладів, оскільки кожне зображення може бути модифіковане по-різному під час кожної епохи навчання.

Узагальнюючи, аугментація даних є невід'ємною частиною архітектури розробленої системи класифікації птахів. Вона забезпечує підвищення точності моделі, її стійкість до варіацій у даних та здатність працювати в реальних умовах. Використання класу ImageDataGenerator у поєднанні з функцією попередньої обробки та можливістю розширення набору трансформацій робить процес аугментації гнучким і ефективним. Цей підхід дозволяє досягти високих результатів класифікації навіть за умов обмеженого обсягу даних, що підтверджує його практичну значущість у сучасних системах комп'ютерного зору.

4.3 Аналіз отриманих результатів

Після завершення процесу навчання нейромережевої моделі було проведено оцінку її ефективності на тестовій вибірці, яка не використовувалася під час тренування. Це дозволяє об'єктивно визначити здатність моделі узагальнювати знання та працювати з новими даними.



Рисунок 4.2 - Графіки точності та втрат моделі

Результати тестування показали значення функції втрат на рівні 0.52093 та точність класифікації 87.27%, що є досить високим показником для задачі багатокласової класифікації з великою кількістю класів (525 видів птахів).

Отримане значення точності свідчить про те, що модель правильно класифікує більшість зображень, незважаючи на складність задачі та високу міжкласову подібність. Для порівняння, у випадку класифікації об'єктів із меншою кількістю класів точність може перевищувати 95%, але для задачі з сотнями класів навіть 80% вважається хорошим результатом. Таким чином, досягнута точність 87.27% демонструє ефективність використання архітектури EfficientNetB0 у поєднанні з додатковими шарами та методами регуляризації.

Аналіз функції втрат показує, що модель не лише досягла високої точності, але й має стабільну поведінку під час навчання. Значення 0.52093 для

категоріальної крос-ентропії свідчить про те, що ймовірності, які прогнозує модель, досить близькі до істинних міток класів. Це означає, що модель не просто вгадує правильний клас, а робить впевнені прогнози з високою ймовірністю. Важливим фактором у досягненні таких результатів є використання механізмів контролю навчання, зокрема *early stopping*, який запобіг перенавчанню, та *reduce learning rate*, що дозволив адаптивно зменшувати швидкість навчання на пізніх етапах.

Досягнуті результати також свідчать про ефективність застосованих методів аугментації даних. Вони дозволили моделі навчатися на більш різноманітних прикладах, що підвищило її стійкість до змін у ракурсах, освітленні та фонових деталях. Без аугментації точність могла б бути значно нижчою через перенавчання на обмеженому наборі зображень.

Отримані результати класифікації демонструють загальну точність на рівні 87%, що підтверджується значеннями *macro avg* та *weighted avg* для *precision*, *recall* і *F1-score* (усі близько 0.87). Це свідчить про збалансовану роботу моделі на більшості класів, без значних перекосів у бік окремих категорій. Однак детальний аналіз показує, що існують класи з високими показниками та класи, де точність і повнота суттєво нижчі.

Видно, що деякі класи досягли майже ідеальних результатів. Наприклад, *AFRICAN OYSTER CATCHER* має *precision*, *recall* і *F1-score* рівні 1.00, що означає повну відсутність помилок у класифікації цього виду. Подібні високі значення спостерігаються для *AMERICAN AVOCET*, *AMERICAN BITTERN*, *ARARIPE MANAKIN*, а також *ASIAN CRESTED IBIS*, де *F1-score* перевищує 0.95. Це свідчить про те, що модель добре розпізнає види з чіткими та унікальними ознаками.

Водночас є класи з проблемними показниками. Наприклад, *ABBOTTS BOOBY* має *precision* 0.71 і *recall* лише 0.43, що призводить до низького *F1-score* (0.54). Це означає, що модель часто плутає цей вид із іншими, або ж він має недостатньо прикладів у навчальній вибірці. Аналогічно, *ANTBIRD* демонструє *recall* 0.50 при *precision* 0.90, що вказує на значну кількість пропущених позитивних прикладів. Такі дисбаланси можуть бути пов'язані з

високою схожістю між певними видами або недостатньою кількістю даних для рідкісних класів.



Рисунок 4.3 - Приклади класифікації зображень розробленою моделлю

Цікаво, що деякі класи мають високий recall, але низький precision, наприклад AMERICAN PIPIT (precision 0.57, recall 0.91). Це означає, що модель часто відносить зображення до цього класу, навіть якщо вони належать до інших видів, що призводить до великої кількості хибнопозитивних прогнозів. Така поведінка може бути результатом схожості ознак між кількома видами або недостатньої диференціації на рівні ознак.

Загальний аналіз показує, що модель демонструє стабільну роботу на більшості класів, але існують групи видів, де точність і повнота значно нижчі. Це може бути виправлено шляхом збільшення кількості прикладів для проблемних класів, застосування більш агресивної аугментації або використання методів балансування даних.

Узагальнюючи, матриця неточностей підтверджує високу загальну якість моделі, але також вказує на необхідність подальшої оптимізації для окремих видів птахів. Враховуючи кількість класів (525) та складність задачі, досягнуті

результати є дуже добрими, а подальші покращення можуть бути реалізовані за рахунок розширення набору даних і тонкого налаштування архітектури.

4.4. Висновки до розділу

У процесі розробки було визначено архітектуру системи, що базується на попередньо навченій моделі EfficientNetB0, адаптованій для багатокласової класифікації.

Особливу увагу приділено модулю аугментації даних, який забезпечує штучне розширення навчальної вибірки та підвищує стійкість моделі до варіацій у зображеннях. Реалізація аугментації за допомогою ImageDataGenerator у режимі реального часу дозволила уникнути перенавчання та забезпечити більш збалансоване навчання.

Отримані результати підтверджують ефективність обраного підходу: модель досягла високої точності класифікації на тестовій вибірці (87.27%), що є значним показником для задачі з великою кількістю класів. Це свідчить про правильність архітектурних рішень та застосованих методів оптимізації. Система є масштабованою, гнучкою та придатною для інтеграції у прикладні додатки, включаючи веб- та мобільні платформи.

РОЗДІЛ 5. РОЗРОБЛЕННЯ СТАРТАП-ПРОЄКТУ

5.1. Опис ідеї проєкту

Таблиця 5.1. Інформаційна карта проєкту

Назва	Опис
Вид проєкту	Інтелектуальна система
Назва проєкту	«Інтелектуальна система класифікації птахів з використанням нейромережевого підходу»
Назва ВНЗ в якому розробляється проєкт	Національний лісотехнічний університет України, кафедра комп'ютерних наук
Прізвище, ім'я, по батькові	Іванонько А.М.
Цілі та задачі проєкту	<ol style="list-style-type: none">1. На основі бази знань, що складається із зображень птахів, створити нейромережеву модель.2. Розробити механізм емуляції нейронної мережі та класифікації птахів.3. Розроблена система повинна мати інтуїтивно зрозумілий інтерфейс.
Короткий зміст проєкту	Проєкт покликаний створити інтелектуальну систему класифікації птахів. Це дозволить автоматизувати процес класифікації птахів.
Терміни виконання проєкту	6 місяців
Бюджет проєкту	193 000 грн.

5.2. Концепція стартапу

Концепція стартапу ґрунтується на створенні високотехнологічного рішення, яке дозволяє автоматично розпізнавати види птахів за зображеннями, використовуючи сучасні методи глибинного навчання та нейромережеві моделі. Основна ідея полягає в тому, щоб надати орнітологам, екологам,

дослідникам дикої природи та любителям птахів інструмент, який значно спрощує процес ідентифікації видів, зменшує час на аналіз та підвищує точність класифікації. У сучасних умовах, коли біорізноманіття є важливим показником екологічної стабільності, швидке та точне визначення видів птахів має велике значення для наукових досліджень і природоохоронних програм.

Ідея стартапу виникла на основі проблеми, з якою стикаються дослідники: ручна ідентифікація птахів є трудомісткою, потребує високої кваліфікації та займає багато часу. Крім того, у багатьох регіонах світу бракує фахівців, здатних швидко визначати види, що ускладнює моніторинг популяцій. Запропоноване рішення дозволяє автоматизувати цей процес, використовуючи мобільний додаток або веб-платформу, де користувач завантажує фото птаха, а система на основі нейромережевої моделі визначає його вид із високою точністю.

Концепція передбачає використання попередньо навчених моделей, таких як EfficientNet, адаптованих для класифікації понад 500 видів птахів. Це забезпечує оптимальне співвідношення точності та швидкодії, що є критично важливим для роботи в польових умовах. Крім того, система інтегрує модуль аугментації даних, що дозволяє підвищити стійкість моделі до різних умов освітлення, ракурсів та якості зображень. Таким чином, користувач отримує надійний інструмент, який працює навіть у складних умовах.

Унікальна цінність стартапу полягає в поєднанні високої точності класифікації, зручності використання та можливості масштабування. Система може бути інтегрована з базами даних орнітологічних спостережень, що відкриває перспективи для створення глобальної платформи моніторингу птахів. Крім того, продукт має потенціал для комерційного використання у сфері екотуризму, освіти та наукових досліджень.

Для перевірки життєздатності ідеї передбачено створення мінімально життєздатного продукту (MVP), який включатиме базовий функціонал: завантаження зображення, класифікацію виду та відображення результату. MVP дозволить отримати зворотний зв'язок від користувачів, оцінити точність моделі та визначити додаткові функції, необхідні для повноцінного продукту.

На основі цього буде сформовано дорожню карту розвитку стартапу, яка включатиме розширення бази видів, інтеграцію з мобільними платформами та впровадження додаткових сервісів, таких як розпізнавання співу птахів.

Концепція також враховує можливість залучення інвестицій та партнерств із природоохоронними організаціями, науковими установами та компаніями, що працюють у сфері штучного інтелекту. Це дозволить забезпечити фінансову стабільність проекту та прискорити його розвиток. Стартап орієнтований на глобальний ринок, що відкриває перспективи масштабування та виходу на міжнародний рівень.

5.3. Бізнес-модель та стратегія монетизації

Бізнес-модель стартапу ґрунтується на створенні цінності для кількох сегментів користувачів: професійних орнітологів, екологічних організацій, наукових установ, а також аматорів та любителів птахів. Основна ідея бізнес-моделі полягає в тому, щоб поєднати технологічну інновацію з комерційною життєздатністю, забезпечуючи стабільний потік доходів і можливість масштабування на глобальному рівні.

Ключовим елементом бізнес-моделі є визначення джерел доходу. У випадку даного стартапу передбачено кілька основних напрямів монетизації. Перший – це підписка на використання платформи, яка може бути реалізована у форматі freemium. Базова версія додатку надаватиме користувачам доступ до основних функцій, таких як класифікація зображень та перегляд результатів. Розширена версія, доступна за платною підпискою, включатиме додаткові можливості: розпізнавання співу птахів, доступ до розширеної бази даних, аналітичні інструменти для наукових досліджень та інтеграцію з геолокаційними сервісами для моніторингу популяцій.

Другим джерелом доходу є ліцензування технології для наукових установ та екологічних організацій. Це дозволить інтегрувати систему у великі проекти з моніторингу біорізноманіття, де потрібні високоточні інструменти для аналізу даних. Ліцензійні угоди можуть передбачати доступ до API, що дає можливість використовувати алгоритми класифікації у власних платформах замовників.

Третім напрямом є комерційне використання у сфері екотуризму та освіти. Мобільний додаток може стати частиною інтерактивних програм для туристів, які відвідують національні парки або заповідники. Крім того, продукт може бути інтегрований у навчальні платформи для студентів біологічних спеціальностей, що відкриває додаткові можливості для співпраці з освітніми закладами.

Стратегія монетизації передбачає гнучкість і можливість адаптації до змін ринку. На початковому етапі основний акцент робиться на залученні користувачів через безкоштовну версію додатку та створення спільноти навколо продукту. Це дозволить сформувати базу лояльних клієнтів і отримати зворотний зв'язок для вдосконалення сервісу. Поступово буде впроваджено платні функції, орієнтовані на професійних користувачів, що забезпечить стабільний фінансовий потік.

Фінансова модель стартапу включає прогнозування витрат на розробку, маркетинг, підтримку та масштабування продукту. Важливим завданням є визначення точки беззбитковості та оцінка рентабельності на різних етапах розвитку. Для цього розробляється детальний фінансовий план, що враховує витрати на хмарну інфраструктуру, оплату праці команди, просування продукту та залучення інвестицій. Очікується, що основні витрати будуть пов'язані з обчислювальними ресурсами для навчання нейромережевих моделей та підтримкою серверів для обробки запитів користувачів.

Бізнес-модель також передбачає можливість залучення венчурного капіталу та грантового фінансування від організацій, що підтримують екологічні та технологічні проекти. Це дозволить прискорити розвиток стартапу, розширити функціонал та вийти на міжнародний ринок. Додатковим джерелом доходу може стати реклама у додатку, але вона буде реалізована таким чином, щоб не знижувати зручність використання.

5.4. План реалізації та управління проектом

Реалізація стартапу «Інтелектуальна система класифікації птахів» є комплексним процесом, що включає кілька ключових етапів: розробку продукту, тестування, запуск та масштабування. План реалізації визначає послідовність дій, терміни виконання та ресурси, необхідні для досягнення поставлених цілей. Ефективне управління проектом забезпечує контроль за виконанням завдань, мінімізацію ризиків та гнучкість у реагуванні на зміни ринкових умов.

Першим етапом є створення мінімально життєздатного продукту (MVP), який дозволить перевірити основні гіпотези щодо функціональності системи та її цінності для користувачів. MVP включатиме базовий функціонал: завантаження зображення, класифікацію виду птаха та відображення результату. На цьому етапі важливо забезпечити стабільність роботи нейромережевої моделі, інтеграцію з інтерфейсом та оптимізацію швидкодії. Розробка MVP дозволить отримати зворотний зв'язок від перших користувачів і скоригувати продукт відповідно до їхніх потреб.

Другим етапом є розширення функціоналу системи. Планується додати можливість розпізнавання співу птахів, інтеграцію з геолокаційними сервісами для моніторингу популяцій та створення аналітичних інструментів для наукових досліджень. Цей етап передбачає донавчання моделі на розширеному наборі даних, оптимізацію алгоритмів та впровадження механізмів масштабування. Важливим завданням є забезпечення високої точності класифікації при збереженні швидкості обробки запитів.

Третім етапом є тестування та контроль якості. Для цього розробляється система автоматизованого тестування, яка перевірятиме коректність роботи моделі, інтерфейсу та серверної частини. Тестування проводиться на різних наборах даних, включаючи зображення з різними умовами освітлення, ракурсами та якістю. Крім того, передбачено проведення бета-тестування серед обмеженої групи користувачів, що дозволить виявити недоліки та внести необхідні зміни перед масштабним запуском.

Запуск продукту на ринок є наступним етапом. Він включає розгортання системи на хмарній інфраструктурі, налаштування механізмів безпеки та

забезпечення стабільності роботи сервісу. Паралельно реалізується маркетингова стратегія, спрямована на залучення користувачів через соціальні мережі, партнерські програми та співпрацю з екологічними організаціями. Важливо забезпечити підтримку користувачів, оперативне реагування на запити та регулярне оновлення продукту.

Управління проектом здійснюється на основі сучасних методологій, таких як Agile та Scrum, що дозволяють гнучко реагувати на зміни та швидко адаптуватися до нових умов. Команда працює спринтами, кожен з яких має чітко визначені завдання та терміни виконання. Регулярні зустрічі та звітування забезпечують прозорість процесів та контроль прогресу. Крім того, впроваджуються механізми управління ризиками, що дозволяють своєчасно виявляти проблеми та мінімізувати їхній вплив на проект.

Важливим аспектом є планування фінансів та ресурсів. На кожному етапі визначаються витрати на розробку, тестування, маркетинг та підтримку продукту. Передбачено залучення інвестицій та грантового фінансування для прискорення розвитку стартапу. Крім того, розробляється стратегія масштабування, яка включає вихід на міжнародний ринок, інтеграцію з глобальними базами даних та впровадження нових функцій.

5.5. Висновки до розділу

Розроблення стартапу «Інтелектуальна система класифікації птахів» є комплексним процесом, що поєднує технологічні, бізнесові та організаційні аспекти. У ході роботи було сформовано концепцію проекту, яка базується на використанні нейромережових технологій для автоматичної ідентифікації видів птахів. Визначено бізнес-модель, що передбачає гнучку стратегію монетизації через підписку, ліцензування та інтеграцію з науковими та комерційними платформами. Розроблено план реалізації, який включає створення MVP, тестування, запуск продукту та масштабування з використанням сучасних методологій управління проектами.

Узагальнюючи, запропонований підхід забезпечує високу ймовірність успішного впровадження стартапу на ринку, оскільки він поєднує

інноваційність, практичну цінність та фінансову життєздатність. Реалізація проекту сприятиме розвитку цифрових технологій у сфері біологічних досліджень та створить нові можливості для екологічного моніторингу та освіти.

ВИСНОВКИ

У ході виконання роботи було розроблено інтелектуальну систему класифікації птахів, яка ґрунтується на сучасних методах глибинного навчання та використанні нейромережових технологій. Основною метою дослідження було створення ефективного інструменту для автоматичного розпізнавання видів птахів за зображеннями, що має практичне значення для орнітології, екологічного моніторингу та освітніх застосунків.

Було проведено аналіз предметної області та формалізацію задачі класифікації. Було визначено множину класів, структуру ознак та математичну модель відображення вхідних даних у вихідні категорії. Особливу увагу приділено проблемам високої міжкласової подібності та внутрішньокласової варіативності, що характерні для птахів, а також стохастичній природі даних.

Розроблено нейромережову модель на основі архітектури EfficientNetB0, яка використовується як екстрактор ознак у поєднанні з додатковими повнозв'язними шарами та механізмами регуляризації. Особливу роль у підвищенні ефективності моделі відіграла аугментація даних, реалізована за допомогою класу ImageDataGenerator. Вона забезпечила штучне розширення навчальної вибірки та підвищила стійкість моделі до варіацій у зображеннях. Це дозволило уникнути перенавчання та забезпечити більш збалансоване навчання.

Результати тестування підтвердили високу ефективність розробленої системи: точність класифікації на тестовій вибірці склала 87.27%, що є значним показником для задачі з великою кількістю класів (525 видів птахів). Аналіз матриці неточностей показав, що більшість класів розпізнаються з високою точністю, хоча для окремих видів спостерігаються помилки, пов'язані зі схожістю ознак або недостатньою кількістю прикладів. Це відкриває перспективи для подальшого вдосконалення системи шляхом розширення набору даних, використання більш глибоких моделей або комбінування з іншими джерелами інформації (наприклад, аудіоаналіз співу птахів).

Узагальнюючи, розроблена система є гнучким, масштабованим та ефективним рішенням для автоматичної класифікації птахів. Вона демонструє високу точність, стійкість до варіацій у даних та можливість інтеграції у прикладні додатки, включаючи веб- та мобільні платформи. Результати роботи підтверджують перспективність використання нейромережових технологій у задачах біологічної класифікації та відкривають широкі можливості для їх застосування у наукових і практичних сферах.

СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ

1. Baowaly, M. K., Sarkar, B. C., Walid, M. A. A., Ahamad, M. M., Singh, B. C., Alvarado, E. S., et al. Deep transfer learning-based bird species classification using mel spectrogram images. *PLOS ONE*, 19(8), e0305708, 2024. DOI: 10.1371/journal.pone.0305708.
2. Chandrashekar, B. S., Nagendraswamy, H. S., Pavan Kumar, M. P. Genetic Algorithm Based Ensemble of Deep Learning Models for Performance Enhancement of Bird Species Classification. *Mathematical Modelling of Engineering Problems (MMEP)*, 12(10), 3563–3576, 2025. DOI: 10.18280/mmep.121022.
3. Chen, T., Li, Y., Qiao, Q. Fine-grained bird image classification based on counterfactual method of vision transformer model. *The Journal of Supercomputing*, 80, 6221–6239, 2024. DOI: 10.1007/s11227-023-05701-6.
4. Chollet F. *Deep Learning with Python*. 2nd ed. Manning Publications, 2021.
5. Goodfellow I., Bengio Y., Courville A. *Deep Learning*. MIT Press, 2016.
6. Han, X., Peng, J. Multi-label Bird Species Classification Using Transfer Learning Network. *Archives of Acoustics*, 50(2), 223–233, 2025. DOI: 10.24425/aoa.2025.154812.
7. ImageNet Large Scale Visual Recognition Challenge. *International Journal of Computer Vision*, 2015.
8. Keras API Reference. <https://keras.io>
9. Li, W., Lv, D., Yu, Y., Zhang, Y., Gu, L., Wang, Z., Zhu, Z. Multi-Scale Deep Feature Fusion with Machine Learning Classifier for Birdsong Classification. *Applied Sciences*, 15(4), 1885, 2025. DOI: 10.3390/app15041885. Wang, K., Yang, F., Chen, Z., Chen, Y., Zhang, Y. A Fine-Grained Bird Classification Method Based on Attention and Decoupled Knowledge Distillation. *Animals*, 13(2), 264, 2023. DOI: 10.3390/ani13020264.
10. Tan M., Le Q. V. EfficientNet: Rethinking Model Scaling for Convolutional Neural Networks. arXiv preprint arXiv:1905.11946, 2019.

11. TensorFlow Documentation. <https://www.tensorflow.org>

ДОДАТОК А

```
import numpy as np
import pandas as pd
import matplotlib.pyplot as plt
import tensorflow as tf
from sklearn.model_selection import train_test_split

# Tensorflow Libraries
from tensorflow import keras
from tensorflow.keras import layers, models
from keras_preprocessing.image import ImageDataGenerator
from keras.layers import Dense, Dropout
from tensorflow.keras.callbacks import Callback, EarlyStopping, ModelCheckpoint, ReduceLROnPlateau
from tensorflow.keras.optimizers import Adam
from tensorflow.keras.applications import MobileNetV2
from tensorflow.keras import Model
from tensorflow.keras.layers.experimental import preprocessing

# System libraries
from pathlib import Path
import os.path
import random

# Visualization Libraries
import matplotlib.cm as cm
import cv2
import seaborn as sns

sns.set_style('darkgrid')

# Metrics
from sklearn.metrics import classification_report, confusion_matrix
import itertools

dataset = "../input/100-bird-species/train"
walk_through_dir(dataset)

image_dir = Path(dataset)

# Get filepaths and labels
filepaths = list(image_dir.glob(r'**/*.JPG')) + list(image_dir.glob(r'**/*.jpg')) +
list(image_dir.glob(r'**/*.png')) + list(image_dir.glob(r'**/*.png'))

labels = list(map(lambda x: os.path.split(os.path.split(x)[0])[1], filepaths))

filepaths = pd.Series(filepaths, name='Filepath').astype(str)
labels = pd.Series(labels, name='Label')

# Concatenate filepaths and labels
image_df = pd.concat([filepaths, labels], axis=1)

label_counts = image_df['Label'].value_counts()[:20]

plt.figure(figsize=(20, 6))
sns.barplot(x=label_counts.index, y=label_counts.values, alpha=0.8, palette='dark:salmon_r')
plt.title('Distribution of Top 20 Labels in Image Dataset', fontsize=16)
plt.xlabel('Label', fontsize=14)
plt.ylabel('Count', fontsize=14)
plt.xticks(rotation=45)
plt.show()

random_index = np.random.randint(0, len(image_df), 16)
```

```

fig, axes = plt.subplots(nrows=4, ncols=4, figsize=(10, 10),
                        subplot_kw={'xticks': [], 'yticks': []})

for i, ax in enumerate(axes.flat):
    ax.imshow(plt.imread(image_df.Filepath[random_index[i]]))
    ax.set_title(image_df.Label[random_index[i]])
plt.tight_layout()
plt.show()

def compute_ela_cv(path, quality):
    temp_filename = 'temp_file_name.jpeg'
    SCALE = 15
    orig_img = cv2.imread(path)
    orig_img = cv2.cvtColor(orig_img, cv2.COLOR_BGR2RGB)

    cv2.imwrite(temp_filename, orig_img, [cv2.IMWRITE_JPEG_QUALITY, quality])

    # read compressed image
    compressed_img = cv2.imread(temp_filename)

    # get absolute difference between img1 and img2 and multiply by scale
    diff = SCALE * cv2.absdiff(orig_img, compressed_img)
    return diff

def convert_to_ela_image(path, quality):
    temp_filename = 'temp_file_name.jpeg'
    ela_filename = 'temp_ela.png'
    image = Image.open(path).convert('RGB')
    image.save(temp_filename, 'JPEG', quality = quality)
    temp_image = Image.open(temp_filename)

    ela_image = ImageChops.difference(image, temp_image)

    extrema = ela_image.getextrema()
    max_diff = max([ex[1] for ex in extrema])
    if max_diff == 0:
        max_diff = 1

    scale = 255.0 / max_diff
    ela_image = ImageEnhance.Brightness(ela_image).enhance(scale)

    return ela_image

def random_sample(path, extension=None):
    if extension:
        items = Path(path).glob(f'*.{extension}')
    else:
        items = Path(path).glob(f'*')

    items = list(items)

    p = random.choice(items)
    return p.as_posix()

train_df, test_df = train_test_split(image_df, test_size=0.2, shuffle=True, random_state=42)

train_generator = ImageDataGenerator(
    preprocessing_function=tf.keras.applications.efficientnet.preprocess_input,
    validation_split=0.2
)

train_images = train_generator.flow_from_dataframe(
    dataframe=train_df,

```

```

x_col='Filepath',
y_col='Label',
target_size=TARGET_SIZE,
color_mode='rgb',
class_mode='categorical',
batch_size=BATCH_SIZE,
shuffle=True,
seed=42,
subset='training'
)

augment = tf.keras.Sequential([
    layers.experimental.preprocessing.Resizing(224,224),
    layers.experimental.preprocessing.Rescaling(1./255),
    layers.experimental.preprocessing.RandomFlip("horizontal"),
    layers.experimental.preprocessing.RandomRotation(0.1),
    layers.experimental.preprocessing.RandomZoom(0.1),
    layers.experimental.preprocessing.RandomContrast(0.1),
])

pretrained_model = tf.keras.applications.efficientnet.EfficientNetB0(
    input_shape=(224, 224, 3),
    include_top=False,
    weights='imagenet',
    pooling='max'
)

pretrained_model.trainable = False

checkpoint_path = "birds_classification_model_checkpoint"
checkpoint_callback = ModelCheckpoint(checkpoint_path,
    save_weights_only=True,
    monitor="val_accuracy",
    save_best_only=True)

# Setup EarlyStopping callback to stop training if model's val_loss doesn't improve for 3 epochs
early_stopping = EarlyStopping(monitor = "val_loss", # watch the val loss metric
    patience = 5,
    restore_best_weights = True) # if val loss decreases for 3 epochs in a row, stop training

reduce_lr = ReduceLROnPlateau(monitor='val_loss', factor=0.2, patience=3, min_lr=1e-6)

inputs = pretrained_model.input
x = augment(inputs)

x = Dense(128, activation='relu')(pretrained_model.output)
x = Dropout(0.45)(x)
x = Dense(256, activation='relu')(x)
x = Dropout(0.45)(x)

outputs = Dense(525, activation='softmax')(x)

model = Model(inputs=inputs, outputs=outputs)

model.compile(
    optimizer=Adam(0.0001),
    loss='categorical_crossentropy',
    metrics=['accuracy']
)

history = model.fit(
    train_images,
    steps_per_epoch=len(train_images),
    validation_data=val_images,

```

```

validation_steps=len(val_images),
epochs=150,
callbacks=[
    early_stopping,
    create_tensorboard_callback("training_logs",
                                "bird_classification"),
    checkpoint_callback,
    reduce_lr
]
)

results = model.evaluate(test_images, verbose=0)

print("  Test Loss: {:.5f}".format(results[0]))
print("Test Accuracy: {:.2f}%".format(results[1] * 100))

accuracy = history.history['accuracy']
val_accuracy = history.history['val_accuracy']

loss = history.history['loss']
val_loss = history.history['val_loss']

epochs = range(len(accuracy))
plt.plot(epochs, accuracy, 'b', label='Training accuracy')
plt.plot(epochs, val_accuracy, 'r', label='Validation accuracy')

plt.title('Training and validation accuracy')
plt.legend()
plt.figure()
plt.plot(epochs, loss, 'b', label='Training loss')
plt.plot(epochs, val_loss, 'r', label='Validation loss')

plt.title('Training and validation loss')
plt.legend()
plt.show()

pred = model.predict(test_images)
pred = np.argmax(pred,axis=1)

# Map the label
labels = (train_images.class_indices)
labels = dict((v,k) for k,v in labels.items())
pred = [labels[k] for k in pred]

# Display the result
print(f'The first 5 predictions: {pred[:5]}')

random_index = np.random.randint(0, len(test_df) - 1, 15)
fig, axes = plt.subplots(nrows=3, ncols=5, figsize=(25, 15),
                        subplot_kw={'xticks': [], 'yticks': []})

for i, ax in enumerate(axes.flat):
    ax.imshow(plt.imread(test_df.Filepath.iloc[random_index[i]]))
    if test_df.Label.iloc[random_index[i]] == pred[random_index[i]]:
        color = "green"
    else:
        color = "red"
    ax.set_title(f'True: {test_df.Label.iloc[random_index[i]]}\nPredicted: {pred[random_index[i]]}',
color=color)
plt.show()
plt.tight_layout()

```