UNIVERSITI TEKNOLOGI MARA

CAT BREED IDENTIFICATION USING CNN

EZWAN HAZIM BIN AZHAR

BACHELOR OF INFORMATION SYSTEMS (HONS.) INTELLIGENT SYSTEM ENGINEERING

JANUARY 2026

UNIVERSITI TEKNOLOGI MARA

CAT BREED IDENTIFICATION USING CNN

EZWAN HAZIM BIN AZHAR

Thesis submitted in partial fulfilment of the requirements for

Bachelor of Information Systems (Hons.) Intelligent System Engineering

Faculty of Computer and Mathematical Sciences

JANUARY 2026

SUPERVISOR APPROVAL

CAT BREED IDENTIFICATION USING CNN

BY

EZWAN HAZIM BIN AZHAR
2024786377

This thesis was prepared under the supervision of the project supervisor, **Dr. Mohd Zaki Zakaria**. It was submitted to the College of Computing, Informatics and Mathematics and was accepted in partial fulfilment of the requirements for the degree of Bachelor of Information Systems (Hons.) Intelligent System Engineering.

Approved by

………………………………..

DR. MOHD ZAKI ZAKARIA
Project Supervisor

JANUARY, 2026

STUDENT DECLARATION

I certify that this thesis and the project to which it refers is the product of my own work and that any idea or quotation from the work of other people, published or otherwise, are fully acknowledged in accordance with the standard referring practice of the discipline.

……………………………..

EZWAN HAZIM BIN AZHAR
2024786377

JANUARY, 2026

ABSTRACT

Identification of a cat breed is a very important element in pet management, veterinary services, and animal welfare systems as different breeds may exhibit distinct genetic traits, health risks, and behavioural characteristics.. However, it is hard to do accurate identification by manual since there are subtle differences in facial structure, fur patterns, and physiological behaviour that distinguish feline breeds, especially when one does not have prior knowledge. To overcome this problem, this current study proposes a mobile-based cat breed identification system based on Convolutional Neural Network (CNNs). It is a system that is designed to automatically recognize twelve different types of cats based on images provided by the Oxford-IIIT Pet Dataset. An extensive general framework of image preprocessing techniques was applied that included data cleaning, image scaling, data normalization, image data augmentation, and breaking up of the data into training, validation, and testing set. Two CNN models, which are the MobileNetV3 and ResNet50, were trained and tested to assess their effectiveness in the process of classifying the breeds of cats in multi-class. The empirical results were solid and consistent, with the most successful model getting an accuracy and precision of 96.83%, which was supported by a steady loss curve of validation and high values of class-respects. To deploy it, the chosen model was translated into TensorFlow lite format and embedded in a Flutter-based mobile application. The application allows user to take a picture of a cat or upload one and immediately be informed about the breed with a confidence score, which does not require internet access. Such offline connectivity enhances availability, privacy of users, and stable performance even in a diversified environment of operation. In general, the designed mobile is a practical and good mobile solution for automated cat identification. Future work may include expanding the number of supported cat breeds, improving robustness against complex backgrounds and varying lighting conditions, and incorporating additional training data to enhance real-world performance.

ACKNOWLEDGEMENT

Alhamdulillah, thanks and praises to Allah SWT who guided me, showed his mercy and blessed me to be able to overcome this final year project tasks with the time limit provided. Nothing of it would have come about, without His will.

To begin with, I would like to sincerely thank my supervisor, Dr Mohd Zaki Zakaria, who has assisted me, encouraged, and gave me precious advice during the process of creating this project. Their comments and advice including their criticism and comments based on their knowledge have helped me a lot in guiding my work and its quality.

I would also like to put special gratitude towards my loving parents, Azhar bin Ahmad and Erniewaty binti Sharifutdin, who have always been with me in providing my sincere love and prayers and moral support. My family and friends are the pillars of my success as they have sacrificed and encouraged me throughout this academic process.

And finally, I would like to give my sincere appreciation to my most valued friend and classmates, in particular, Muhammad Azril bin Shukor, Emirul Rafiq Syahmi bin Jami'an and Mohamad Hazik Haikal bin Razak, who has been there to give me a helping hand, encouragement and accompany me in both difficult journey and good time in this project.

I would like to thank each one of you to share this enlightening experience.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

Abbreviations

AI	Artificial Intelligence
CNN	Convolutional Neural Network
IDE	Integrated Development Environment
UI	User Interface
UX	User Experience
APK	Android Package
XML	eXtensible Markup Language
TFLite	TensorFlow Lite
TICA	The International Cat Association
ResNet	Residual Network

CHAPTER 1
INTRODUCTION

1.1 Background Research

Cats have become the most favourite and extensively embraced pets around the world. By understanding and gaining knowledge about cat's breed is essential for giving an appropriate care, food and medical treatment because each breed possessing unique physical and behavioural characteristics (Qatrunnada N, Fachrurrozi M, Utami A, 2022).

Nevertheless, many owners of cats, specifically those who adopted from shelters or rescue a stray cat, frequently lack knowledge of their cat's breed (Ramadhan & Setiawan, 2023). This lack in understanding may interfere with specific breed care and make harder the anticipation of health disease or behavioural problems (Hamdi Sunaryo J, Prasetyo S, 2023). Therefore, there is an increasing demand for accessible and precise technique to recognize cat breeds based on their appearance.

Recently, utilizing Artificial Intelligence (AI) and deep learning has been significant increase, especially for image recognition. These developments are mostly carried by enhancement in computational power, the accessibility of larger labelled dataset, and the efficiency of networks such as Convolutional Neural Networks (CNNs) (Tita Karlita, Nadia Azahro Choirunisa, Rengga Asmara and Fitri Setyorini, 2022). CNNs have shown remarkable efficiency in understanding complex patterns from images, which allows machine to deal with task such as object detection, facial recognition and medical images analysis with exceptional accuracy (Ramadhan & Setiawan, 2023). These abilities have given the opportunity to utilize deep learning algorithms to solve real-world problem, such as detecting cat breeds.

Identifying cat breeds manually gives difficult task. The characteristics between some breeds can be complicated because it includes details such as fur pattern, ear shape and facial structure (Zayani, Hafedh Mahmoud, Kachoukh, Amani, Ghodhbani, Refka, Khediri, Nouha, Abd-Elkawy, Eman H., Ammar, Ikhlass, Kouki, Marouan, Saidani and Taoufik, 2025). To identify cat breeds with similar characteristics might be required an expertise in this field, which may not always available to pet shelters or pet owners. The problems are made much more difficult when facing with mixed breeds. Therefore, traditional methods used to classify breeds are generally slow, inconsistent and subjective (Ramadhan & Setiawan, 2023).

In order to solve this problem, Convolutional Neural Networks (CNNs) provide an effective solution for this problem (Feng, Yiming). These models built to handle visual object, enabling it to discover and interpret detailed features from cat images (Zhao, Xia, Wang, Limin, Zhang, Yufei, Han, Xuming, Deveci, Muhammet, Parmar, Milan, 2024). By training CNNs on large dataset of labelled cat breeds, the system able to classify unknown images with high precision (Ramadhan & Setiawan, 2023). This technique could transform the way cat breeds are recognized, by enhancing speed, improving accuracy and make accessibility to many users (Lei et al., 2020). The technology might be used in web or mobile platforms that enables users to upload an image of their cat and obtain the breed instantly, thereby providing benefits to pet owners, veterinarians, and animal shelters (Karad, Vishwanath, 2025).

1.2 Problem Statement

Manual recognition of feline breeds is a considerable procedural burden, especially to veterinary workers, domestic animal owners, and workers in shelters who have limited amount of professional knowledge relating to the systematic breed recognition (Rangkuti et al., 2024). Specific identification requires the level of expertise that professionals possess, and any mistake in identifying can lead to inappropriate health treatment, incorrect selection of breeds, and poor decisions regarding adoption, which ultimately undermines the quality of animal welfare (Barrios et al., 2023). Such challenges are further indicated by the strong phrenomagnetic resemblance of some breeds and in some cases, these characteristics, including pealed colour, coarse, hair distribution, and facial structure, are often immeasurable to the non-expert person (Feng, 2025) The severity of this problem is hardened by the close resemblance of some breeds in their phenotype, whereby phenotypic features including fur colour, colour patterning, body sizes and facial structure often overlap and cannot be easily distinguished by the non-expert (Feng, 2025).

This inefficiency in reliable expertise may often lead to mass misclassification of breeds, causing a wrong veterinary treatment course, inappropriate breeding behaviour, or ineffective adoption results (Hamdi et al., 2023). Although there is a growing need to identify a breed accurately, currently, there are very few, easy to use, and accurate breed identification tools available to help users discriminate a breed efficiently and successfully. The existing techniques are mostly manual, tedious, or dependent on external reference databases and expert knowledge, and thus restricted to the general population (Zayani et al., 2025).

Furthermore, most of the existing applications in the mobile industry rely on crude image-matching or manual data input of physical characteristics with limited breed coverage and accuracy, that is, unstable with changing imaging conditions like lighting, distracting backgrounds, and image quality (Hamdi et al., 2023; Qatrunnada et al., 2022). Such applications often involve the misclassification or even inability to properly detect feline breeds due to the lack of high-tech applications of Artificial Intelligence (AI). These technological failures lead to delay in timely evidence-driven decision-making in the course of veterinary consultation, adoption and breeding evaluations (Barrios et al., 2023).

As a result, there is a strong motivation to create an automated, an artificial intelligence based cat breed identification system, which will be capable of recognizing breeds when given images and provide a convenient, effective, and affordable solution to veterinarians, pet owners, and animal shelters, which will eliminate the need to understand the subject on an expert level.

1.3 Research Question

The research questions of this study are as follows:

I. How can Convolutional Neural Networks (CNNs) can be utilized to detect and classify cat breeds from images?

II. What is the accuracy and reliability of the CNN model in differentiate visually similar cat breeds?

III. How the process for integrating the model into mobile applications for real-time breed detection?

1.4 Research Objectives

The research objectives of this study are as follows:

I. To develop CNN models that able to recognize and classify various cat breeds from images.

II. To evaluate the performance of CNN models using appropriate evaluation metrics on a labelled cat breed image.

III. To develop a mobile application that integrated a CNN model to detect and classify cat breeds from captured or uploaded images.

1.5 Scope of The Research

This study is dedicated to the use of image-based information to perform the task of cat breed classification using Convolutional Neural Networks (CNNs). The dataset in the study was taken according to the Oxford-IIIT Pet Dataset which can be accessed publicly and through the Kaggle site. The research is confined to 12 breeds of cats specifically Abyssinian (198 images), Bengal (200 images), Birman (200 images), Bombay (200 images), British Shorthair (184 images), Egyptian Mau (200 images), Maine Coon (190 images), Persian (200 images), Ragdoll (200 images), Russian Blue (200 images), Siamese (199 images), and Sphynx (200 images), thus it is equal to 2371 images. This study does not discuss the generalizability of the trained model to cat breeds not in this dataset.

In a measure to determine the functionality and efficacy of the proposed model, standard classification performance measures, such as accuracy, precision, recall, and F1-score are used. These measures are applied to the determination of the correct identification and differentiation of the model with the categorization of chosen breeds of cats. Moreover, the implementation of the trained CNN model into the prototype mobile app will also be part of this project to enable users to upload or take photos of cats to predict their breed in real-time. The application is implemented as a prototype that is more of a functioning prototype, which is meant to show the capability of the CNN-based classification and not as a product that can be used by the end users.

1.6 Significant of The Research

This study is important to the government agencies, the communities and animal welfare institutions. The system will be able to help enhance the effectiveness of pet registration, adoption, and rescue efforts as well as responsible ownership, by allowing proper identification of the cat breed. It also helps the efforts made in a community level to control the population of the strays and the sensitization about the animal care that eventually helps the people in maintaining proper health practices and safer living conditions.

Accurate identification of cat breeds is essential in ensuring that the pets are in good overall health and well-being (SDG3). Various breeds could have unique health predispositions, behavioural characterizations, and nutritional needs. By using a mobile application for breed detection from deep learning which Convolutional Neural Networks (CNNs), pet owners, veterinarians and animal shelters can make a quick decision regarding to medical care, nutrition planning and preventative health care. Fast and efficient access to breed information minimizing the risk of misdiagnosis and improves the quality of care, especially in cases when a cat’s breed is unidentified due to adoptions or rescue. This not only have benefits for the cats, but it also encourages pet owner and improve public health outcome.

Furthermore, this project promotes the ethical treatment as well as sustainable management of domestic animal populations (SDG 15). It provides animal shelters, breeders, and rescue organizations with a useful tool for identify cat breeds, which may help in enhance the record-keeping and adoption matching processes of the cats. Identifying particular characteristics of a breed can ensure the cats will be placed in circumstances suitable for their personality traits and needs to be cared for, thus, reducing abandonment rates and increasing the level of animal welfare. The approaches promote better knowledge, empathy, and sustainability in human-animal relationship by promoting ethical pet owners and animal management practices.

1.7 Summary

The chapter has given an overview of the research on cat breeds detection based on Convolutional Neural Networks (CNN). It started by bringing out the premise and the rationale of creating an intelligent system that is capable of identifying various breeds of cats correctly using images. The problem statement emphasized the pitfalls in the manual breed identification and the necessity to have an automated one. The research questions and the objectives as well as the scope were clearly stated to direct the project to achieve its objectives. The importance of the study was also mentioned, and it could possibly be of help in the whole process of helping the veterinarians, animal shelters, and cat lovers. On the whole, this chapter has prepared the provisions of the research, and the following chapter will examine the literature and methods that are related and can be applied to the image classification and CNN-based breed detection.

CHAPTER 2
LITERATURE REVIEW

2.1 Overview of Cat Breeds

Cats are one of the most popular and cherished pets around the world, which come in may breeds that have varies in similar appearance, personalities and their possible health problems (Chen, 2023). According to The International Cat Association (TICA), there are 73 cat breeds, including Persian, Siamese and Bengal (Karad et al., 2025). Each of the breeds have their own characteristics (Karlita et al., 2022b), some of them are more active and love to be with people and others are calmer and do not interact much (Rangkuti et al., 2024).

Understanding these differences is important because various traits in breeds influence on their physical appearance, nutrition, exercise level, health problems, and compatibility with various surroundings or owners (Yu et al., 2020). It becomes a challenge due to the diversity of breeds, especially in normal or mixed-breed cat, highlighting the importance and difficulty visualize breed classification (Chen, 2023).

2.2 Visual Characteristics of Cat Breeds

2.2.1 Facial Structure

Facial structures is often the main feature to identify cat breeds (Feng, 2025). A variety of breeds has their own unique such as have different head shapes, the ear placements, and eye patterns (Chen, 2023). For example, Persian cats have flat faces with little noses (Hamdi et al., 2023), while Siamese cat is recognized by their long, pointed noses, triangular faces and big, sharp ears (Chen, 2023). Scottish Folds can be identified by their ears that fold forward and Maine Coon is showed by its broad head, and high cheekbones also have large ears with tufted tips (Chen, 2023). These structural characteristics play important roles for identifying breeds visually and often the first thing observed by humans as well as computer vision algorithms (Alzubaidi et al., 2021; Zhao et al., 2024).

Figure 2.1 Persian, Siamese, Scottish Folds and Maine Coon

2.2.2 Fur Pattern and Colours

The pattern and colours of cat's fur is one of the main ways to distinguish between breeds of cats (Chen, 2023). The pattern that can be seen on cat coats are solid, tabby, tortoiseshell, calico, spotted and pointed (Chen, 2023). For example, Bengal cats have a similar appearance to leopards where it has spots and rosettes in their body (Barrios et al., 2023), while Siamese cats show contrast between the body and the darker points on their extremities (Qatrunnada et al., 2022). In terms of colour, breeds may exhibit different of it such as black, white, grey, cream or combination of these (Chen, 2023). The texture and length of fur also different for each breed (Barrios et al., 2023), some breeds have long-haired like Ragdoll which have silky coat, unlike Sphinx breed, this breed is hairless with a peach fuzz coating (Chen, 2023). These characteristics play an important role in both human identification of breeds and machine learning recognition (Zeng, 2024).

Figure 2.2 Bengal, Siamese, Ragdoll and Sphinx

2.2.3 Body Size and Shape

Body size and shape also play important roles to help identify cat breeds, since several breeds has their own unique physical body (Chen, 2023). Cat can be categorized in small, medium or large size. For instance, Maine Coons breeds known as largest among the domestic cats because of its rectangular physique (Chen, 2023). In some breeds, such as Munchkins breeds, they have short legs due to genetic mutation, while Oriental Shorthair breeds, their body is quite long and thin bodies (Chen, 2023). Features such as tail length, leg proportion and the body shape are important to distinguish between breeds (Chen, 2023).

Figure 2.3 Maine Coon, Munchkin and Oriental Shorthair

2.3 Importance of Cat Breed

2.3.1 Health and Veterinary Diagnosis

Understanding cat breeds is an important part of health and veterinary care (Barrios et al., 2023). Certain health diseases are genetically inclined in certain breeds, therefore veterinarians enable to make more accurate diagnoses by knowing the breeds and customize treatment plan according to breeds (Barrios et al., 2023). For example, Maine Coon tends to get hypertrophic cardiomyopathy, which is a heart disease, while Persian cats predisposed to respiratory and kidney problems (Yu et al., 2020). By knowing the breed early, it allows preventing health issues, conducting specific screenings, and making informed which enhancing the general health of the cat.

2.3.2 Pet Breeding and Genetic Diversity

Pet breeding and genetic diversity also rely on identification of cat breeds (Chen, 2023). Understanding cat breeds helps the breeders or the owners to avoid inbreeding and have better control over hereditary traits (Yu et al., 2020). It also can assist in planning of breed specific mating that maintain pedigree and health criteria of specific breeds (Yu et al., 2020). Furthermore, breed identification also assists at adoption centre in pairing the cats with appropriate owner, considering the breed personality and lifestyle requirements (Karad et al., 2025). The breed awareness contributes to diversity genetically, which also prevents the prevalence of heritable diseases and enhances the sustainability of healthy cat populations (Barrios et al., 2023).

2.4 Overview of Deep Learning

Deep Learning is part of the machine learning based on biological neural networks (Alzubaidi et al., 2021; Wang et al., 2025), which consist multi-layered architecture that learn hierarchical features directly from raw data, without manual feature engineering (Noor & Ige, 2024; Zhao et al., 2024). Compared to conventional approaches based on hand-designed algorithms like SIFT or HOG

(Zhang, 2021), Deep Learning model can learn from pixels, allowing it to capture complex patterns such as shapes of ears or fur textures (Karlita et al., 2022b; Rangkuti et al., 2024), which are important features to identify between cat breeds. Due to the stacked layers of neuron, this is become possible with the lower ones detecting low-level (edges, colours) and the deeper ones connecting them into high-level (facial structures, fur patterns) (Alzubaidi et al., 2021; Hussain Khan & Iqbal, 2025; Noor & Ige, 2024). Training process consists of forward propagation, which to make predictions (Zhang, 2021) and backward propagation with gradient descent to update weights in a model iteratively (Alzubaidi et al., 2021; Noor & Ige, 2024). The scaling properties of Deep Learning and its capacity to outclass conventional procedures on mixed-up tasks have pushed it to be the default strategy of current image classification, which can be observed in its recurrent win in competitions, such as ImageNet (Feng, 2025; Hussain Khan & Iqbal, 2025; Wang et al., 2025).

2.5 Convolutional Neural Networks (CNNs)

Convolutional Neural Networks (CNNs) constitute a landmark in the sphere of deep learning and, specifically in image processing (Zhao et al., 2024). CNNs were designed to exploit the spatial structure and hierarchical structure of visual data unlike the traditional neural networks that viewed image pixels as a one-dimensional sequence (Alzubaidi et al., 2021; Zhao et al., 2024). It will organize around the key principle of emulating the human visual cortex which performs operations on data in a layered manner using simple features to complex representations (Zhao et al., 2024). A standard CNN consists of multiple layers of different types convolutional, activation, pooling and fully connected that interact with each other to allow automatic discovery of features in raw pixel data, bypassing the need to design features by hand (Hussain Khan & Iqbal, 2025; Zhao et al., 2024).

The main body of a CNN is its convolutional layers (Mithilesh Vishwakarma et al., 2024; Wang et al., 2025), in which learnable filters (sometimes called kernels) are swept over the input image (Hussain Khan & Iqbal, 2025; Mithilesh Vishwakarma et al., 2024). Each filter looks for particular local features, such as edges, textures or corners (Qatrunnada et al., 2022), by computing element-wise products and adding the outcomes to generate a feature map (Alzubaidi et al., 2021). This procedure is efficient in extracting localized patterns in the image. After convolution, a non-linear activation function is applied most often the Rectified Linear Unit (ReLU) (Noor & Ige, 2024). This important trick makes the model non-linear (Lei et al., 2020), thus able to learn and represent more complex and varied patterns than can be described by linear transformations alone (Hussain Khan & Iqbal, 2025). This interaction between many convolutional layers and various filters allow the network to construct a deep hierarchy of features, starting with simple lines in the initial layers to more abstract object parts in the deeper layers (Alzubaidi et al., 2021; Hussain Khan & Iqbal, 2025).

Periodically, pooling layers are added to CNN architecture to control the continuously increasing feature maps and to enhance model generalization (Qatrunnada et al., 2022). The main purpose is the dimensional reduction, which literally decreases the computational demands and the overall number of parameters (Alzubaidi et al., 2021; Noor & Ige, 2024). More importantly, pooling also provides the model with translation invariance so that small changes in location of an object does not destroy its identification (Alzubaidi et al., 2021). Such a strategy is called Max Pooling and is very popular since it summarizes the most relevant information by picking the strongest signal within a local window (Zhao et al., 2024). After multi-stage convolutional and pooling operations have extracted the refined features, the data is reshaped into a flat vector (Alzubaidi et al., 2021). This vector is then fed to fully connected layers which effectively behave like a normal neural network and perform the final and complex reasoning required to classify it (Wang et al., 2025; Zhao et al., 2024) and the final layer which often has a Softmax activation function in the multi-class case finally outputs a list of probabilities and the CNN can identify which breed of cat was based on the input image (Alzubaidi et al., 2021; Zhao et al., 2024).

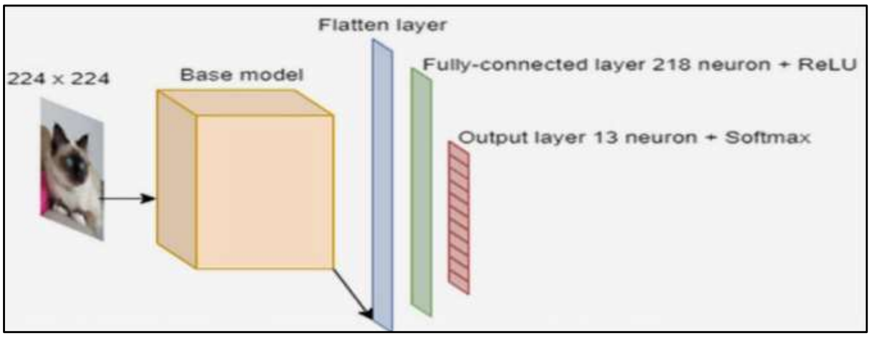

Figure 2.4 CNN Architecture

2.6 Residual Networks (ResNets) for Deep Models

As neural networks descending deeper to capture more complex features, there exists a serious problem called the vanishing gradient problem which is likely to be encountered (Alzubaidi et al., 2021; Tan & Le, 2020). Weight updates by backpropagation in the backpropagation computation of gradients and propagation of these gradients through the network, it is possible to have an exponential decrease in the gradients when they traverse the many layers of the network (Zhang, 2021; Zhao et al., 2024). The reason this occurs is that, at layers closer to the input, gradients are enormously small which, in effect, causes the learning process to stop in those early layers and causes the network to struggle to learn meaningful low-level features (Alzubaidi et al., 2021; Zhao et al., 2024). Because of this, merely piling on additional layers in a standard deep neural network does not guarantee an improvement in performance, and in practice frequently results in a decrease in accuracy and training instability (Tan & Le, 2020; Zhao et al., 2024).

In order to overcome this problem, Residual Networks (ResNet) has been proposed, which that transformed the field of deep learning by introducing the idea of residual blocks and skip connection (or identity mappings) (Alzubaidi et al., 2021; Tan & Le, 2020). A residual block inserts an additional "shortcut" connection, jumping over one or more convolutional layers, and then the input of the block can be added directly to its output (Alzubaidi et al., 2021; Zhao et al., 2024). In theory, rather than learning a direct mapping H(x), the block is induced to learn a residual function F(x)=H(x)x. Then, the result is F(x)+x (Alzubaidi et al., 2021; Zhao et al., 2024). This might appear to be a very trivial change, but it helps tremendously in the backpropagation of gradients (Hussain Khan & Iqbal, 2025). When the optimal mapping for a block is similar to an identity mapping which means just passes the input through, then the network has an easy time learning to drive the residual F(x) to zero, instead of having to approximate a complicated identity mapping with non-linear layers (Alzubaidi et al., 2021; Zhao et al., 2024). This direct route allows a new route that gradients can follow, allowing them to avoid vanishing and allowing networks hundreds and even thousands of layers deep to be trained successfully (Alzubaidi et al., 2021; Zhao et al., 2024).

Residual block integration is providing great performance and stability advantage in very deep models (Noor & Ige, 2024). On a range of difficult image recognition benchmarks, ResNets have uniformly outperformed conventional deep networks of comparable depth in terms of accuracy (Hussain Khan & Iqbal, 2025). This has been boosted by the fact that they are now able to optimize very deep architectures which reduce the vanishing gradient problem and enable the network to learn more complex and hierarchical feature representations (Hussain Khan & Iqbal, 2025; Noor & Ige, 2024). ResNets also converge quicker in training. The most important variations of the ResNet architecture are ResNet-18, ResNet-34, ResNet-50, ResNet-101 and ResNet-152 with the digit denoting the number of layers in the network (Hussain Khan & Iqbal, 2025; Lei et al., 2020). ResNet-50 and ResNet-101 have become especially popular, with further optimizations being commonly based on so-called bottleneck residual blocks in order to achieve greater computational efficiency with very deep versions without loss in performance (Filippov, 2024; Hussain Khan & Iqbal, 2025).

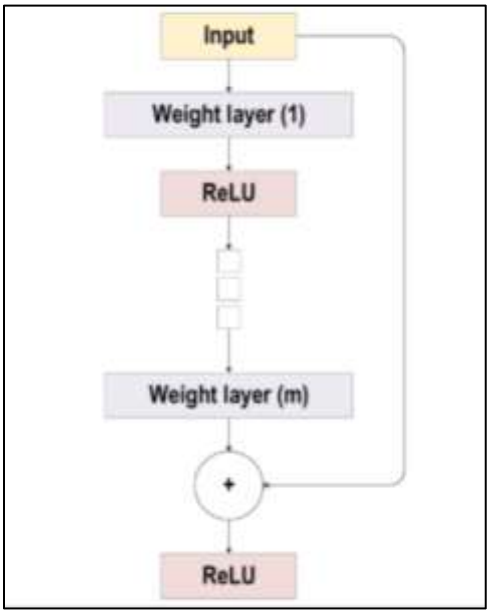

Figure 2.5 ResNet Architecture

2.7 Lightweight Architectures (MobileNet) for Mobile Deployment

However, as opposed to continuously deeper and more complicated models, such as ResNets, there exists an increasing need to have efficient deep learning models, which can be deployed to mobile devices as well as embedded systems, or even real-time applications (Noor & Ige, 2024; Qian et al., 2021). Although very accurate, traditional large CNNs require a lot of computational power, memory, and energy, which is not feasible at the edge, where processing possibilities, battery lifetime, and storage are restricted (Lei et al., 2020; Zhao et al., 2024). Such need has prompted the creation of lightweight architectures that seek to match the competitive accuracy of their larger counterparts at a much cheaper computational cost and model size, making on-device inference and broader accessibility of AI applications possible without being cloud-infrastructure dependent (Hussain Khan & Iqbal, 2025; Tan & Le, 2020).

MobileNet architecture has been introduced to be served by the main innovation of that architecture depthwise separable convolutions (Qian et al., 2021). Depthwise separable convolutions split the filtering and channel combining steps that standard convolutions do in one operation into two separate, smaller operations (Hussain Khan & Iqbal, 2025; Qian et al., 2021). The former is a depthwise convolution, in which one filter is applied to each input channel separately (Qian et al., 2021). This is a huge reduction in the operations over a regular convolution (Hussain Khan & Iqbal, 2025). A second step is a pointwise convolution, a 1x1 convolution over the output of the depthwise convolution (Karlita et al., 2022b; Qian et al., 2021). This depthwise convolution is what does the sum of the outputs of the depthwise convolution, resultantly enabling new features to be made across channels (Karlita et al., 2022b; Qian et al., 2021). MobileNet separates the spatial filtering operation and the channel combination, thereby eliminating a large amount of redundancy in computations (Qian et al., 2021; Zhao et al., 2024).

This factorization approach straightforward results in a substantial decrease in the number of parameters and computational cost (Multi-Adds) over standard CNNs, and more importantly competitive accuracy (Zhao et al., 2024). The performance improvements are significant, and MobileNet models are highly optimized to run on mobile and embedded devices with little performance cost (Hussain Khan & Iqbal, 2025; Qian et al., 2021). MobileNet has been developed over time with some important variants, each based on the idea of depthwise separable convolutions and subsequent additional optimizations (Hussain Khan & Iqbal, 2025; Qian et al., 2021). MobileNetV2 is the architecture by introducing "inverted residuals" and "linear bottlenecks," which enhance feature representation capacity and memory efficiency (Qian et al., 2021). As later versions, such as MobileNetV3, showed that neural architecture search (NAS) methods were also applied to further optimize the layer design and achieve better performance on achieving specific hardware platforms, which indicates the ongoing tendency towards more efficient and effective lightweight deep learning models (Hussain Khan & Iqbal, 2025; Qian et al., 2021).

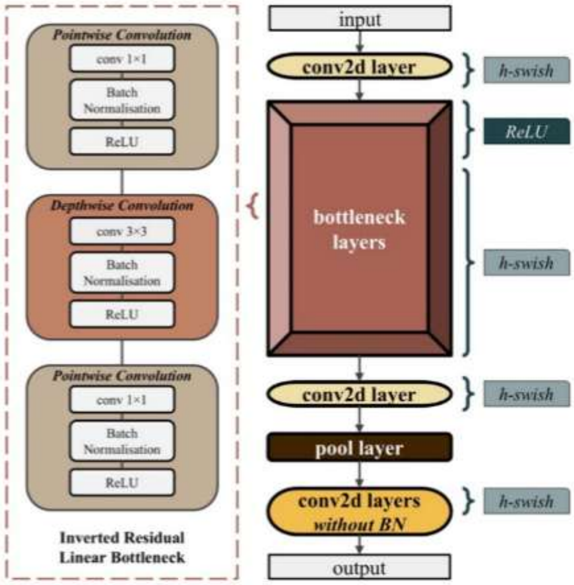

Figure 2.6 MobileNetV3 Architecture

2.8 Accuracy Evaluation in CNN Models

The performance of a Convolutional Neural Network (CNN) model is a necessary evaluation of its ability to classify images in the category in which they belong. Accuracy is used the most as a performance measure and measures the percentage of images that are correctly classified within the total number of images that are tested (Filippov, 2024). Mathematically, the ratio of the number of correctly predicted samples over the number of predictions gives the percentage of accuracy (Alzubaidi et al., 2021). Although accuracy gives an overall picture about the model performance, it might not be a complete picture of the model reliability particularly when the distribution of classes is lopsided in the dataset.

In an attempt to get a better understanding, researchers tend to employ more measures like precision, recall and the F1-score (Hamdi et al., 2023). Precision is a measure of the fraction of correct positive predictions among positive predictions, in how many of the predicted outcomes are correct (Alzubaidi et al., 2021). Recall, also referred to as sensitivity, is a measure of the properties of the model to correctly identify all the applicable samples of a class (Sudharson et al., 2023). F1-score is the harmonic means of precision and recall, which offers the balanced measure in the case of uneven classes distribution (Alzubaidi et al., 2021). Moreover, the confusion matrix is often employed to display the number of correct and incorrect predictions in all the classes in order to assist the researcher to understand the types of classes that are frequently misclassified (Feng, 2025).

Even though accuracy is a useful and intuitively right measure, it may be misleading when there is unequal representation of the classes in the dataset, for example some cat breeds represent more images than others (Feng, 2025). When this happens, the model can seem to do well on the whole but in fact, it can be missing the ability to properly identify the minority classes (Alzubaidi et al., 2021). Therefore, it is important to include various measures of evaluation, including precision, recall, and F1-score, as a way of having a more multi-faceted and robust performance estimation of the CNN-based classification models.

2.9 Data Augmentation Techniques

Data augmentation is a deep learning strategy that is extremely difficult to avoid, especially on image classification tasks such as cat breed detection where it may be complex to amass huge, varied datasets (Karad et al., 2025; Karlita et al., 2022b; Rangkuti et al., 2024). Its essence is to artificially increase the size of the training set by creating new, but reasonable variations of the available images, thus improving the model capability to generalize and mitigating the chance of overfitting (Carratino et al., 2022; Feng, 2025; Kumar et al., 2023). This process to ensure that the resultant trained model is less sensitive and more resilient to small real-world variations that it may be exposed to (Kumar et al., 2023; Zeng, 2024). Those methods are further divided into geometric transformations and photometric adjustments that deal with the different types of variability that occur in images (Kaur et al., 2021; Kumar et al., 2023; Zeng, 2024).

Geometric transformations maintain the spatial layout of the pixels, imitating a change of perspective, or of the position of the objects in the scene, without changing the fundamental identity of the image (Alzubaidi et al., 2021; Kaur et al., 2021). Typical transformations include rotation, whereby the images are rotated by different angles to train the model to identify the objects irrespective of their orientation, and flipping (usually horizontal), which reflects the image to cover the differences in left and right (Alzubaidi et al., 2021; Feng, 2025; Kumar et al., 2023). Scaling, cropping and translation may be considered other geometric transformations (Kumar et al., 2023). Additional to these are photometric adjustments, which modify colour, brightness and contrast to emulate different lighting conditions or camera settings (Kaur et al., 2021; Kumar et al., 2023). An important effect in this regard is HSV jittering, in which Hue, Saturation and Value aspects of the colour of an image are randomly adjusted (Kumar et al., 2023; Zeng, 2024). This enables the model to be robust to variations in lighting, colour richness or exposure, so that the model is correctly classified despite variations in photograph conditions (Kumar et al., 2023). A combination of these augmentation techniques has a profound effect on the training data and results in CNNs that generalize to unseen real-world images (Feng, 2025; Kaur et al., 2021).

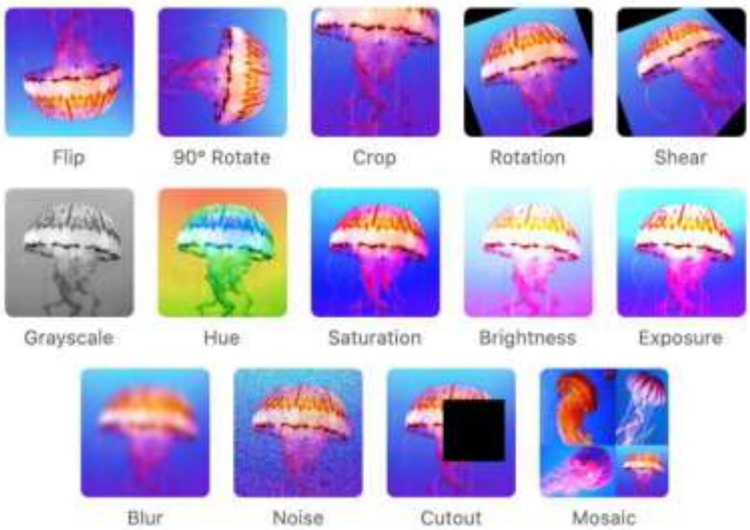

Figure 2.7 Example of Data Augmentation Techniques

2.10 Related Works

In the recent years, it has seen a great surge in the number of research in the field of cat breed detection, with multiple papers and theses being published employing methods of machine learning and image processing. Variously, researchers have used quite a number of different methods and datasets to come up with correct cat breed classifier models. Some of the prominent research works have suggested and developed models that show superior performance, mostly based on the Convolutional Neural Networks (CNNs). A significant number of those works directly address the topic of designing and improving CNN architectures, which frequently include some optimization algorithms and more sophisticated techniques to improve the accuracy and efficiency of models even further.

Karlita et al. (2022) presented the automated Cat Breed Classification system based on the Deep Learning method and EfficientNet-B0 Convolutional Neural Network (CNN) model. They conducted their study on the classification of nine cat breeds using a dataset of 2700 images. One of the most important findings of their work was that data pre-processing is of high importance to enhance model performance. The system using their experiments was able to obtain a respectable accuracy of 95% in classification with a validation accuracy of 91% (Karlita et al., 2022b).

According to Ramadhan and Setiawan (2023) proposed a Cat Breed Recognition using Machine Learning and images in recognizing cat breeds using Convolutional Neural Networks (CNN) with transfer learning. They evaluated MobileNetV2, VGG16, and InceptionV3 as backbones to categorize 13 breeds of cats using around 200 images of each breed taken and provided by the Oxford-IIIT Pet Dataset. Using 80-20 percent training/testing data split, MobileNetV2 was determined as the optimal classification model, with the overall accuracy of 82 percent. It produced the highest accuracy of 92.03% training and 81.99% testing in one of its scenarios. They then incorporated this best model in an android application denoted Catbreednet, where they pointed out that the hyperparameter adjustment required much attention, which in this case was epoch 15, to avoid overfitting (Ramadhan & Setiawan, 2023).

Qatrunnada et al. (2022) explored the Cat breed categorization on the basis of a Convolutional Neural Network (CNN) Xception Architecture. They investigated classification of five cat breeds, where they had a dataset of 13 060 images, where 200-3200 images of each breed served as the training set. The work was the first to experiment with single and multi-object images and showed the capability of CNN in classifying multiple cats in an image. On single-object images, the model attained an accuracy of over 90 percent consistently. On multi-object images, it reached 100% accuracy on 2 and 3 objects, 95% on 4 objects, and 94% on 5 objects but the accuracy declined with the number of objects because of the quality of the image being reduced (Qatrunnada et al., 2022).

Zhang (2021) performed a comparative study in the classification of cat breeds, considering VGGNet and Inception-v3 models. The paper particularly suggested an augmented deep learning framework, the main improvement of which concerned the aspect of feature extraction. This optimization consisted of inserting a concept module to make the network deeper and wider to allow enriching the feature map with multi-scale features. With a set of 10 types of cats taken in the Oxford-IIIT Pet Dataset, the optimized model received a recognition rate of 84.56% which is higher than VGGNet-13 (78.69%), VGGNet-16 (79.48%), and Inception-v3 (80.42%) (Zhang, 2021).

The study conducted by (Karad et al., 2025) within the area of cat breed recognition implemented a machine learning method, namely, used a ResNet50 Convolutional Neural Network (CNN) model. The model shows good performance on image classification tasks and has attained outstanding outcome on several datasets. The practical implications of their findings concern the animal welfare, pet breeding, and veterinary medicine. The experiments were carried out on a training and testing set of 953 colour images of five classes of cats. ResNet50 model attained training accuracy of 85% and validation accuracy of 84.89% indicating it fits well to unknown data.

Zayani et al. (2025) used the most recent version of the YOLO family, YOLOv11, to overcome such challenges as fine mixed breed variations. Their research paper used a custom dataset of 640 labelled images of five different types of cats which are Persian, Maine Coon, Siamese, Pallas Cat and Bengal, and trained the model on this dataset using data augmentation to increase the variety. YOLO has been found useful in image classification although it was mainly developed in object detection. They tried out different optimizers, and RAdam and SGD performed the best, showing an average recall of 96.8%, precision of 97.2%, and a remarkable mAP50 (accuracy) of 98.1%. This shows the great role of the choice of optimizer (Zayani et al., 2025).

From Sudharson et al. (2023), they aimed at creating one of the precise and faithful solutions that will identify and classify dogs and cats in real-time. They have used the YOLO-v7 object detection model, which is a Convolutional Neural Network (CNN) type model that is efficiently to image classification tasks. They trained and validated their model on the Oxford IIIT pet dataset that contains different breeds of dogs and cats. This system has got high F1 score of 85.6% and mean Average Precision (mAP) of 82.57%. The work highlights the suitability of YOLO-v7 to real-time tasks in the setting, like a veterinary clinic, animal shelter, and pet shop, but indicates that the shortcomings of the dataset could be improved through making the images more diverse (Sudharson et al., 2023).

Rangkuti et al. (2024) studied the application of image processing to classification of cats breeds, training seven Convolutional Neural Network (CNN) models: DenseNet121, ResNet50V2, InceptionV3, VGGNet16, VGGNet19, EfficientNetB0, and MobileNetV2. They conducted research by taking a sample of 610 images of eight breeds of cats taken on Google Images, iStock, and Pinterest. EfficientNetB0, VGGNet16 and VGGNet19 models showed the best results with accuracies of 76%, 74%, and 71% respectively among the models evaluated. The contribution of this work is in effective identification of cat breeds (Rangkuti et al., 2024).

Hamdi et al. (2023) investigated a fusion-based approach to enhance the Cat breed classification, which consists of fusing several Convolutional Neural Network (CNN) architectures. They tested 11 various CNN models, namely EfficientNetB0, EfficientNetB1, EfficientNetB2, EfficientNetB3, EfficientNetV2B0, EfficientNetV2B1, EfficientNetV2B2, EfficientNetV2B3, VGG16, VGG19 and Xception. The research utilized a dataset of 15 cat breeds having 375 images of each breed, summing up to 5625 images. They realized that model fusion is a potential method, which performs better than individual models. The 3-model fusion (Xception, EfficientNetV2B0, and EfficientNetB3) showed the highest results with an accuracy of 90.53%, a precision of 90.75%, recall of 90.53%, and an F1 score of 90.16%. This implies that that a combination of model strengths can be applied towards better feature extraction (Hamdi et al., 2023).

Feng (2024) new a novel architecture of cat breed identification, which is composed of an object detector YOLOv5 and a deep learning model VGG16. This two-fold procedure was to add accuracy by concentrating on relevant features and to add resistance to background noise. The dataset consisted of five largest cat breeds which are Calico, Persian, Siamese, Tortoiseshell, and Tuxedo that used in the study. The overall accuracy of the combined system was an impressive 87%, as compared to the 73% accuracy of the VGG16-only model because YOLOv5 detected the cats first before VGG16 classified them (Feng, 2025).

The journals, articles and other related works that are discussed and explained in Related Work section of this chapter were summarized and listed in Table 2.1.

Table 2.1 Summarize the related works

No.	Title/Author/Year	Technique	Findings/Results	Remarks
1	Cat Breeds Classification Using Compound Model Scaling Convolutional Neural Networks (Karlita et al., 2022b)	• EfficientNet-B0 • Convolutional Neural Network (CNN) • Deep learning	Accuracy of 95%	Highlighted that data pre-processing is important in enhancing model performance.
2	Catbreedsnet: An Android Application for Cat Breed Classification Using Convolutional Neural Networks (Ramadhan & Setiawan, 2023)	• Machine Learning • Convolutional Neural Network (CNN) • MobileNetV2 • VGG16 • InceptionV3	Accuracy of 82%	An android application was created that incorporated the best model.
3	Cat Breeds Classification using Convolutional Neural Network for	• Xception Architecture	Accuracy above 90%	Proved the CNN to be effective in classifying more than one cat in an image but with an

	Multi-Object Image (Qatrunnada et al., 2022)	• Convolutional Neural Network		accuracy that might reduce as the number of objects increase
4	Classification and Identification of Domestic Cats based on Deep Learning (Zhang, 2021)	• VGGNet • Inception-v3	Accuracy above 84.56%	The optimization made the accuracy very outstanding compared to other models tested.
5	Deep Learning based Cat Breed Recognition (Karad et al., 2025)	• ResNet50 • Convolutional Neural Network (CNN)	Accuracy of 85%	Results are of interest to animal welfare, the breeding of pets, and veterinary medicine.
6	Cat Breed Classification with YOLOv11 and Optimized Training (Zayani et al., 2025)	• YOLOv11 • Data augmentation • Various optimization algorithms (RAdam, SGD, Adam, Adamax, NAdam, AdamW, RMSProp)	Accuracy of 98.1%	Pointed out potential vast effect of the choice of an optimizer on the model performance.
7	Efficient Real-time Breed	• YOLO-v7	Accuracy of 82.57%	Produced a rightful and firm real time breed

	Classification using YOLOv7 Object Detection Algorithm (Sudharson et al., 2023)			recognition and classifications of dogs and cats,
8	Harnessing Deep Learning for Cat Breed Identification: A Visual Recognition Strategy (Rangkuti et al., 2024)	• DenseNet121 • ResNet50V2 • InceptionV3 • VGGNet16 • VGGNet19 • EfficientNetB0 • MobileNetV2	Accuracy of 76%	Tracks to making people more accurate and quicker in determining cat breeds depending on the multiple nature of them
9	Fusion of pretrained CNN models for cat breed classification: A comparative study (Hamdi et al., 2023)	• EfficientNet • VGG16, • VGG19 • Xception • transfer learning • fine-tuning	Accuracy of 90.53%	Model fusion has been also suggested as an effective way of boosting the accuracy of classification.
10	Integrating Object Detection and Deep Convolutional Neural Networks for Cat Breed	• YOLOv5 • VGG16	Accuracy of 87%	Dual approach was increasing accuracy through concentrating on appropriate features and increasing robustness towards background

Classification (Feng, 2025)	noise.

2.11 Gaps in Literature

There are a number of gaps in the current body of literature on cat breed recognition as related to data availability and real-world applicability. One of the recurrent issues is that the size and the variety of the specialized cat breed datasets are not adequately large. Researchers such as Ramadhan and Setiawan (2023), Qatrunnada et al. (2022), Zhang (2021), Patil et al. (2025), and Zayani et al. (2025) tended to employ rather small datasets, which requires data augmentation to reduce overfitting. The necessity to enlarge those datasets is mentioned regularly. This is in comparison with dog breed classification which has been reported to achieve higher accuracies at other times, indicating a relative lack of specialized cat data.

In terms of real-world implementation, there is application that has been done, but lacking research on large-scale implementation. Ramadhan and Setiawan (2023) created an Android app, Catbreednet, and Feng (2024) hoped to be helpful to animal shelters using their built-in YOLOv5-VGG16 model. The future deployment on edge devices and in animal care environments to achieve real-time identification is also indicated by Patil et al. (2025), Zayani et al. (2025) and Sudharson et al. (2023). Nonetheless, it is not yet clear how complex deep learning models can be deployed on mobile devices with limited resources, as they require considerable processing power and memory, so the full real-world assessment is evolving.

2.12 Summary

The chapter has introduced a detailed overview of the main ideas, methods, and previous studies that are applicable in designing a cat breed detection system. It started by giving basic knowledge on cat breeds including their physical appearances and their role in veterinary diagnosis, identification of pets and breeding procedures. The chapter followed up with an explanation of fundamental concepts of deep learning especially Convolutional Neural

Networks (CNNs), Residual Networks (ResNet), and lightweight models or networks like MobileNet that are more efficient and have lower computational requirements and thus are more suited to mobile deployment.

Also, several accuracy evaluation measures including precision, recall, F1-score and confusion matrices have been reviewed to determine how the performance of models are generally evaluated in classification problems. Also investigated were the methods of data augmentation, which had an effect on better generalization of the model and preventing overfitting.

The chapter was finished by summarizing the related works on animal and cat breed classification in CNNs and then pointing to gaps in the existing literature. Lack of proper management of cat breeds of similar appearance and absence of offline mobile based solutions are some of these gaps that give the impetus and reason to have the proposed system. In general, the provided literature is the basis of the methodology and system design that was outlined in the next chapter.

CHAPTER 3
RESEARCH METHODOLOGY

3.1 Research Methodology Design

In this chapter, the research methodology that was used to achieve the goals of this project is outlined. The entire process is divided into a series of well-structured steps, which is represented in the Research Methodology Framework. Each phase has certain tasks and deliverables that are designed to ensure a systematic, rigorous, and effective development process.

The approach has three major goals, which are oriented. To begin with, a Convolutional Neural Network (CNN) model is developed to recognize and classify various breeds of cats using image data. Secondly, the accuracy and reliability of the proposed CNN models are determined by evaluating their performance with the help of suitable measures and labelled data. Lastly, a mobile application is created that incorporates the trained CNN model that allows the real-time detection of cats through images taken with the cameras on the device or in the gallery.

Such an approach to a study allows dividing the study into clearly identified stages so that every aspect of the system, including data collection and model creation, can be carefully designed, developed, and tested. Such systematic work becomes the key to the effective creation and testing of the offered system of cat breed classification and the mobile version of it.

Figure 3.1 presents the research methodology framework that would be used in this project. The framework outlines the chronological linear flow of the research process that includes starting with the initial study and data collection, then moving to image processing, CNN model design, model testing, and ultimately system design, system development, and documentation. This visual representation provides a description of the relationship between each step and the way the research purposes are comprehensively covered during the project lifecycle.

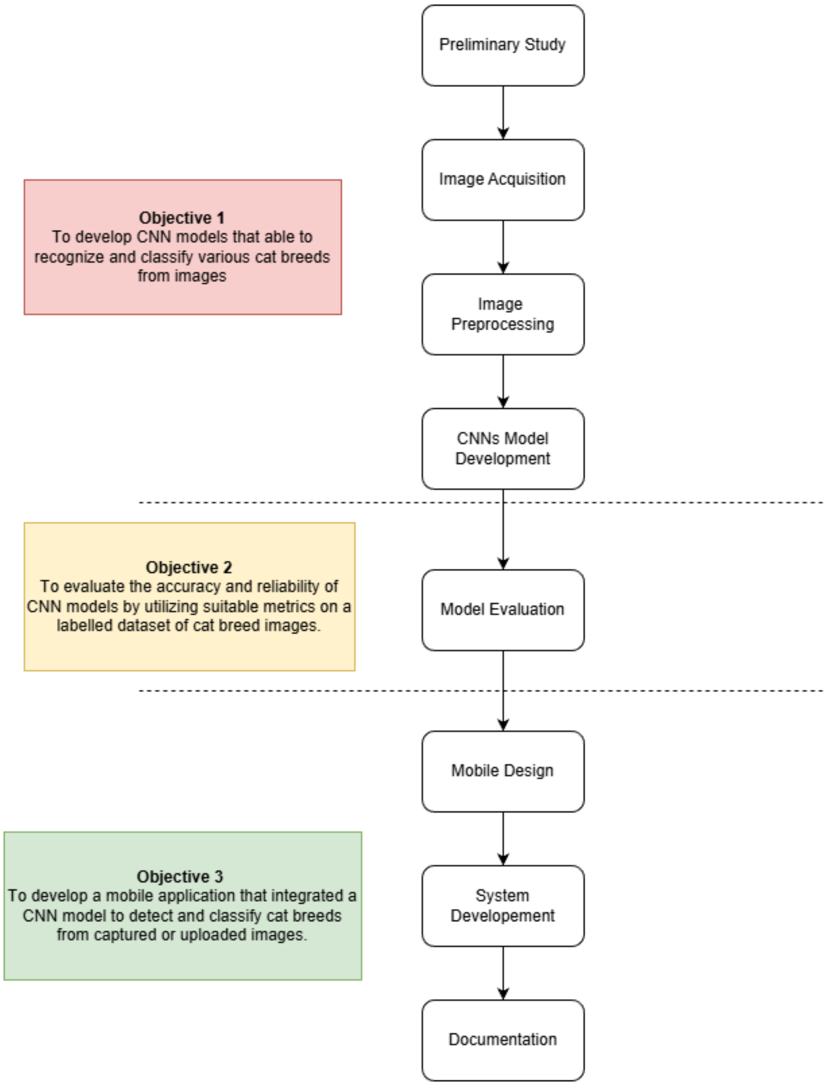

Figure 3.1 Research Methodology Design

3.2 Research Methodology Design Table

The following Table 3.1 gives a clear picture of the research methodology that will be applicable in the project. It is a systematic set of steps that divide the whole process of research into individual stages and present the most specific activities to be performed by the researcher, as well as the projected deliverables per stage, all of which are consistent with the most important project goals. The structured approach allows us to have a clear and well-arranged research study, like the initial study and then the development of the system and the final documentation.

Table 3.1 Research Methodology Design Table

No.	Objectives	Phase	Activities	Deliverables
1	To develop CNN models that able to recognize and classify various cat breeds from images	Preliminary Study	Reviewing the literature review of past studies that related to the project.	Summarized the literature review, problem statement, and scope of the project.
		Image Acquisition	Collect data of cat breed images from Kaggle. https://www.kaggle.com/datasets/tanlikesmath/the-oxfordiiit-pet-dataset/data	Organized dataset
		Image Preprocessing	Clean and preprocess the images using the selected augmentation	Remove unwanted data (for example, dog breeds)

			techniques.	
		CNN Models Development	• Develop the CNN model (MobileNetV3 and ResNet50) • Train the models • hyperparameter fine-tune	Trained CNN models
2	To evaluate the accuracy and reliability of CNN models by utilizing suitable metrics on a labelled dataset of cat breed images.	Model Evaluation	Test the models using the testing dataset, evaluate it, and analyze the strengths and weaknesses of the model prediction.	Comparison of model performance report with evaluation results of both CNN models with the help of various metrics.
3	To develop a mobile application that integrated a CNN model to detect and classify cat breeds from captured or uploaded images.	Mobile Design	Design the interface of the application to be user-friendly.	Interface of the Cat Breed Identification.
		Mobile Development	Develop the application using Android Studio and integrate it with the CNN model.	The application of the Cat Breed Identification integrated with the CNN model.
		Documentation	Writing a full report.	Final year project report.

3.3 Preliminary Study

In the first phase of the project will include a critical review of scholarly articles and research concerning the application of automated image classification, more specifically animal breed recognition. It will be more specific to explore literature that has dealt with the issue of identifying the cat breeds, bearing in mind that the identification process is sometimes difficult and, in the case of veterinary doctors, owners of pets, and animal shelters, they have not acquired any expertise due to their lack of expertise on the topic. The review will also explore the fact that most breeds have similar features such as fur patterns, colour, size, or even facial structures. Exceptional CNN architectures (such as MobileNetV3 or ResNet50) that can be used in the image classification problem, data augmentation techniques, transfer learning methods, and performance metrics to be used in computer vision problems are also going to be considered in the literature review. Moreover, the studies of the aspect of implementing machine learning models in mobile applications that can be used in real-time inferencing will be studied to provide information on the system. This review will give the basic knowledge of the technicalities of the development of CNN as well as the practical difficulties of the breed identification of cats.

A brief summary of the effective literature review will be drawn up based on the extensive scope of the literature review. The problem statement included in this summary will be clearly stated with a focus on how misclassification in the absence of expert help could influence the problematic decisions as far as pet owner preferences, health care, and breeding compatibility are concerned and, therefore, the quality of animal health care. At the same time, the scope of the project will be well determined. This will involve a specification of the fact that the project utilizes imagery data and that it is restricted to 12 breeds of cats, as well as that the model generalizability to the other breeds of cats that are not included in this data source is not discussed.

The deliverables will also attest to the fact that the functionalities of the model will be identified based on several classification metrics such as accuracy, precision, recall, and F1-score. Finally, it will describe that the project will end with the implementation of the trained CNN based model into a mobile app that will allow users to either upload or take a picture of a cat to identify its type in real-time, and that the given application will be taken only as a prototype to show generic classification capabilities and not as a final consumer-ready product.

3.4 Image Acquisition

This phase entails the process of acquiring image data, which are to be used in the training, validating, and testing of the Convolutional Neural Network (CNN) models that were developed in this project. To achieve the project goals, two publicly available datasets were chosen, which are the Oxford-IIIT Pet Dataset and the Animal Image Dataset, which were obtained via the Kaggle repositories. All these datasets are helpful in the two key classification tasks of the proposed system, which are Cat vs Non-Cat classification and Cat Breed classification.

3.4.1 Oxford-IIIT Pet Dataset

Oxford-IIIT Pet Dataset (Tan, 2023) was chosen in the cat breed classification task because it has high-quality annotations, the labelling is standardized, and it has extensive visual diversity. The dataset consists of 7349 images, specifically 12 cat breeds and 25 dog breeds. Because the scope of this research is narrowed down to feline classification, all images of cat breeds remained, but all the images of the dogs were dispatched.

Overall, 2,371 images of cats were used in this experiment, which reflected twelve different breeds of cats. These images also have differences in the lighting conditions, background settings, poses, fur patterns, and facial structures, and this has made the trained CNN models more effective in terms of goodness of generalization. Table 3.2 shows the proportion of images by the twelve breeds of cats that were utilized in this research.

Table 3.2 Distribution of Cat Breed Images in the Oxford-IIIT Pet Dataset

Cat Breed	**Number of images**
Abyssinian	198
Bengal	200
Birman	200
Bombay	200
British Shorthair	184
Egyptian Mau	200
Maine Coon	190
Persian	200
Ragdoll	200
Russian Blue	200
Siamese	199
Sphynx	200
Total	**2371**

To give a graphical display of the data, sample images of the twelve breeds of cats are given in Figure 3.2. These samples depict the visual variation in the texture of the coat, facial structure, and the morphology of the body that CNN models have to learn.

Figure 3.2 Sample Images of the 12 Cat Breeds from the Oxford-IIIT Pet Dataset

3.4.2 Similar Breed Subset (Three Breeds)

Besides the standard twelve-breed task of classification on a complete twelve-breed dataset, a finer subset was also selected based on the Oxford-IIIT Pet Dataset to test the capability of CNN models to perform fine-grained classification. The subset also looks at three visually close breeds of cats, which are Siamese, Ragdoll, and Birman.

These breeds have been chosen because of their strong similarity in the coloration of fur, facial expression, and general body appearance, such that they are harder to tell apart. The images of these breeds were rearranged into an independent directory structure to allow individual training and assessment. This design allows conducting a comprehensive evaluation of the capacity of the model in terms of the detection of subtle discriminative aspects in similar visual scenarios.

3.4.3 Animal Image Dataset

The Cat vs Non-Cat classification task was supported using the Animal Image Data set. In this data there are a total number of 5,406 images, which represent 90 animal categories, domestic and wild animals like dogs, cows, lions, elephants, and zebras. The data were reorganized into two categories, which are Cat and Non-Cat.

In this research, the data will comprise 60 cat pictures and 5346 non-cat images. To allow the CNN model to learn a broad set of non-feline visual features, all other non-cat categories were combined into one Non-Cat category. This wide variety of representation enhances the strength of the Cat vs Non-Cat classification stage by minimizing the cases of false positives and making sure that valid cat images only are sent to the breed classification stage.

Table 3.3 Distribution of Images in the Animal Image Dataset

Category	Number of images
Cat	60
Non-Cat	5346
Total	**5406**

To demonstrate the variety of sample data, Figure 3.3 presents sample images belonging to the chosen categories of animals. The samples emphasize the difference in shape, texture, and visual look that the model should be trained to differentiate between cats and other animals.

Cat	Dog	Cow
		
Horse	**Elephant**	**Lion**
		

Figure 3.3 Sample pictures of the Animal Image Dataset (Cat vs Non-Cat Classification)

3.4.4 Dataset Organization

Both of these datasets were structured in a hierarchical directory structure, wherein each folder represents a certain type of class or cat breed. The organization supports the effective loading of data, preprocessing, and model training. The images were separated in JPEG format so that they do not need any conversion prior to being processed with the deep learning models and CNN architectures that will be employed in this research.

In general, the Oxford-IIIT Pet Dataset in combination with the Animal Image Dataset can be regarded as a solid basis to build a multi-stage cat recognition system that utilizes CNNs. The Animal Image Dataset facilitates precise cat recognition, and the Oxford-IIIT Pet Dataset supports correct classification with the appearance of the breed, which can add to the effective and efficient system, applicable to practice.

3.5 Image Preprocessing

Image preprocessing is an important step in using Convolutional Neural Network (CNN) models since it is a process that cleans, standardizes, and prepares all the input images to use in the learning process. Successful preprocessing minimizes inconsistencies of the data and increases the quality of the features that are extracted, the accuracy of the model, stability, and the generalization of the model. After that there were some preprocessing methods that were used in the current research, and they were data cleaning, image resizing, normalization, data augmentation, and partitioning of data whereby two datasets were used, namely the Oxford-IIIT Pet Dataset and the Animal Image Dataset.

Data cleaning was done as the first step to eliminate irrelevant and duplicate images. According to the goals of this work, only the images related to the chosen twelve cat breeds were included in the Oxford-IIIT Pet Dataset, and all the images of the dogs have been removed. In the meantime, the original 5,400 images in 90 animal categories of the Animal Image Dataset were restructured into two categories, which are Cat and Non-Cat. Any image that was categorized as a cat was pooled under the Cat category, and any image that belonged to the other classes of animals was combined to form the Non-Cat category. This reorganization was a requirement of successful training of the Cat vs Non-Cat classification model.

After data cleaning, the images were all resized to 224 x 224 pixels to comply with the input size of the CNN architectures employed, which are MobileNetV3 and ResNet50. The uniformity of the image dimensions allows consistency in the input size of the models and efficiency in the computation of the models during the training process. The pixel intensity values were then rescaled to a range of 0 to 1, and this contributes to stabilizing the numerical calculations and the rate of convergence towards the model in optimization.

The generalization ability of the CNN models was sought to be perfected by means of data augmentation technique applied exclusively to the training dataset in order to minimize the chances of overfitting. Data augmentation is the artificial method of improving the diversity of datasets by adding logical variations to the originating images but does not change the semantic content of the image. The augmentation process applied in this study comprised of the random rotation up to 20 degrees and the horizontal flipping to reproduce real-life changes in the image orientation and pose.

Through these transformations, the CNN models can be trained to learn stronger and discriminative visual features and learn them using more variants of the same image. Upon rotation, blank areas in an image were replaced with nearest fill mode to ensure continuity in the image. Table 3.4 also shows the examples of the original and augmented images that have been used to show the results of rotation and horizontal flipping that occurred during training.

Table 3.4 Example image after data augmentation

Data Augmentation	**Before**	**After**
Random rotations		
Horizontol flip		

A training split of 70% of the data was used to make the CNN models have an overall exposure to a variety of images, which is crucial towards recognizing powerful and discriminative features, a particularly important point in a fine-grained classification task like recognizing different cat breeds. The 15% validation set made it easy to optimize the hyperparameters and to identify overfitting at a very early stage thus maintaining the integrity of the final evaluation. At the same time, the other 15% test set offered a statistically valid value on the overall generalization of the model on unseen data. This segmentation method has been common in deep-learning image classification works and it has been statistically shown to be effective in providing stable performance without data leakage between training and evaluation phase.

Overall, the preprocessing methods adopted were effective in ensuring that the input data was cleansed, standardized and diverse enough. The ensuing preprocessing pipeline formed a sound base in the suitable training and evaluation of the Cat vs Non-Cat Classification model, and the Cat Breed Classification model.

3.6 CNNs Model Development

Convolutional Neural Network (CNN) models were created and tested in this study to assess their performance, reliability and the general capacity to generalize to perform the given task of a cat breed image classification. The goal of this step is to find an appropriate CNN architecture and training system which has a high classification rate and at the same time is practical enough to be deployed on a mobile platform. Two CNN models, which are MobileNetV3 and ResNet50, were taken and showed through the systematic assessment of their applications when identifying cats and detecting the specific breed of cat.

In order to enhance robust and minimize misclassification, the classification task was made a multi-stage classification pipeline. Since this was the first application, the Cat vs Non-Cat classification model was first trained on an animal picture dataset to serve as a filtering mechanism so that only valid cat pictures would be sent on to the further processing. This step serves to lower false predictions due to the occurrence of non-cat images. Then, they were trained on a cat breed classification model with the help of the Oxford-IIIT Pet Dataset that has 12 different cat breeds. Moreover, in keeping with a fine-grain classification step was also adopting similar levels to differentiate more between visual similar breeds namely Birman, Ragdoll and Siamese. This is a multi-stage strategy that will make the system more robust because there is a higher probability of misclassification in cases where the physical appearances of two or more breeds are almost similar.

3.6.1 MobileNetV3

MobileNetV3 is a resource-efficient CNN architecture which is specifically created to be used in resource-heavy settings, like mobile and embedded devices. It is an enhanced extension of the earlier MobileNet models and include depth wise separable convolutions, inverted residual block and linear bottlenecks, to dramatically minimize model parameters and computational expense. Moreover, MobileNetV3 also incorporates squeeze-and-excitation features, and optimized activation functions to increase the feature representation with low efficiency.

MobileNetV3 was used as the main model architecture in this study because it is the appropriate model architecture in contexts of real-time mobile based applications. Due to the nature of the proposed system, where a CNN model will be implemented in a mobile application, there will be a significant need of having a lightweight and efficient architecture to provide a fast inference time, less and less memory consumption, and easy usability. MobileNetV3 was trained and tested on the chosen cat breed dataset to determine how well it is able to extract discriminative features of feline as well as how well it is able to classify cat breeds, including those having visual features that are similar.

3.6.2 ResNet50

ResNet50 is a deep neural network model with 50 convolutional layers that have been known to be an efficient feature extractor and an effective model in terms of classification. ResNet50 is best characterized by the fact that residual connections are used, thereby allowing the network to also perform identity mappings, and to address the vanishing gradient problem that is often faced in deep neural networks. These have been revealed as residual connections, which enable ResNet50 to train an even deeper architecture in constant convergence and better learning performance.

In this study, ResNet50 was used as a comparative baseline model to compare the performance a deeper and more computationally expensive model with the lightweight MobileNetV3 model. The ResNet50 can be integrated providing a comparative analysis of the trade-offs between the accuracy of the classification and the overall computational efficiency. Through the comparison of MobileNetV3 with ResNet50, this paper will be trying to defend the choice of the lightweight model in mobile implementation but also confirm that the classification capability will be close to the competition.

3.6.3 Training Configuration of the CNNs Model

Both CNN architectures were trained under various experimental configurations in order to investigate how various training configurations affect the convergence, stability, and generalization of the model. The data was separated into the training, validation, and test sets explained in the above section. The training process was conducted by means of systematic decrease and increase in batch size, learning rate, and epoch numbers to analyze their influence on learning behaviour and modelling results. Table 3.5 describes the training configuration used for the CNN model evaluation.

Table 3.5 Training Configuration Used for CNN Model Evaluation

Parameter	Values Used	Description
CNN Architectures	MobileNetV3, ResNet50	Models evaluated for classification performance
Batch Size	32, 64	Evaluated to analyze training stability and

		convergence
Learning Rate	0.001, 0.0001, 0.00001	Tested to determine optimal weight update magnitude
Number of Epochs	10, 15, 20	Selected to ensure sufficient learning while avoiding overfitting
Optimizer	Adam	Adaptive optimizer for efficient gradient descent
Loss Function (Binary)	Binary Cross-Entropy	Used for Cat vs Non-Cat classification
Loss Function (Multi-Class)	Categorical Cross-Entropy	Used for cat breed classification
Early Stopping	Enabled	Applied based on validation loss

The 32 and 64 batch sizes have been chosen due to their wide use in the most common deep learning tasks of image classification and their compromise between trainability and computation efficiency. The batch size of 32 gives the model many weight updates and therefore learns more detailed representations of features, whereas the batch size of 64 gives the model better computation and gradient update stabilization. The comparison of the two values enables one to analyze the effect of each on the convergence speed and generalization performance.

The learning rates of 0.001, 0.0001 and 0.00001 were selected to analyze various strengths of weight changes in the training. The default learning rate of the Adam optimizer is usually 0.001 and allows faster convergence. These smaller learning rates which are 0.0001 and 0.00001, were added to determine the possibility of the slower and more accurate updates to improve the performance and avoid unnecessary oscillation around the optimal solution. The range enables the most successful learning dynamics in each model architecture to be identified.

The training epochs, which are 10, 15 and 20, were used to provide enough training and prevent overfitting. A smaller training time can also lead to underfitting, whereas too long training can lead to overfitting, especially in the case of fine-grained classification such as cat breed recognition. Through assessment of various epoch values, the research undertaking holds the benefit that each model gets sufficient time to learn discriminative features without lacking the capacity to generalize. Also, early termination using validation loss was used to terminate training in the event that no additional improvement was found, further minimizing the risk of overfitting.

Each experiment adopted the Adam optimizer because it offered an adaptive learning rate mechanism that helps improve the convergence rate in all training regimes. To solve the Cat vs Non-Cat classification task, one used the binary cross-entropy loss function, whereas to solve the multi-class cat breed classification task, one used categorical cross-entropy loss.

During the training phase, training accuracy, validation accuracy, and loss were measured at each epoch during training to track the learning behaviour and identify possible overfitting or underfitting. The models were tested on the test data after training. The different evaluation measures used in performance assessment were accuracy, precision, recall, F1-score, and confusion matrix analysis that offer a comprehensive analysis of the effectiveness of the classification.

According to the experimental outcomes, the best configuration to be used in dealing with each task was identified and integrated into the proposed mobile application. The relative discussion of MobileNetV3 and ResNet50 identified the trade-off between the capability of extraction of features and the effects of computation efficiency. MobileNetV3 was found to be very fit to be deployed on mobile, as the architecture was lightweight, whereas the learning ability of

ResNet50 was high to learn the complex visual characteristics. The chosen models are the main focus of the Cat Breed Identification System and support accurate, efficient, and reliable offline classification in the mobile application

3.7 Model Evaluation

To measure the performance and reliability of the developed Convolutional Neural Network (CNN) models and also the generalization ability, model evaluation was done to find the best architecture that suits each classification problem. The experiments were realized with MobileNetV3 and ResNet50 with unseen test data of a ratio of 15% of the total data. Such an evaluation strategy made sure that the reported results were unstable and gave a strong suggestion on each model's capacity for generalization to new data.

To give a piece of information on the models, several performance measures were used, which included accuracy, precision, recall, F1-score, and analysis of the confusion matrix. Accuracy was used to measure the number of correct predictions as a whole, whereas the precision and recall gave information about the predictions on individual classes. The F1-score, which is a harmonic mean of the measures of precision and recall, balanced the measures and was a more intensive measure of classification performance. Also, confusion matrix analysis was used to present the results of class-wise prediction and the frequent trends of misclassification.

MobileNetV3 and ResNet50 models were tested on three classification tasks, which were Cat vs Non-Cat classification, 12 Cat breeds classification, and three similar breeds classification. The models were compared against each other to determine the architecture that would give the best balance between the classification accuracy and the computational efficiency. The model with maximum overall accuracy and F1-score was chosen as the final model to be implemented to perform each task.

After choosing the most appropriate model, the most significant models were stored in a mobile deployment format by using HDF5 (.h5) and Tensorflow Lite (.tflite). This made sure that the chosen models can be effectively integrated into the mobile application and run locally on the device to help find the cat in real time.

Overall, the evaluation model process was significant in confirming the model performance, as well as in choosing the most useful CNN architectures to use in the proposed Cat Breed Identification System.

3.8 Mobile Design

This section outlines the general system design of the proposed Cat Breed Identification System, explaining the interaction processes that occur between the system components, beginning with the user input and ending in the classification output. The system is designed to fit a multi-step Convolutional Neural Network (CNN) inference pipeline to execute in a mobile application platform, thus guaranteeing a high-precision and effective cat breed identification system and user-friendly application.

3.8.1 Overall System Architecture

The Cat Breed Identification System proposed follows a multi-stage structure that is modular in nature and is aimed at enabling the efficient image-based classification using a mobile platform. The architecture combines a mobile application interface and on-device deep learning inference and local data storage to guarantee low response times, less dependency on the internet, and ease for users. The system flow starts with the interaction with the user through the mobile application and then follows through multiple CNN-based classification stages to give the final output. Figure 3.4 represents the general architecture of the system proposed in the Cat Breed Identification System.

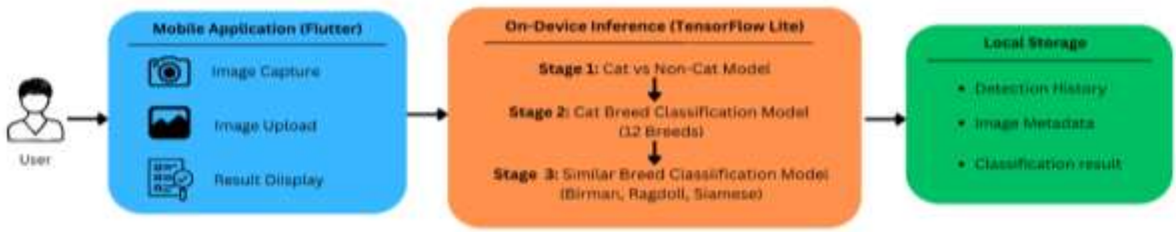

Figure 3.4 Integrated Application Flow and Model Processing Diagram

As Figure 3.4 indicates, the architecture consists of three main elements: the Mobile Application Layer, the On-Device Inference Layer, and the Local Storage Layer.

The Cat Breed Identification System is designed in layers and as a modular system to allow the efficient and accurate classification of images on a mobile platform, as presented in Figure 3.4. The system begins at the mobile application layer, which is the main interface of user interaction. This layer allows users to take a new picture or use an existing one of a cat, either through the smartphone camera or the gallery. Basic input validation, image resizing, and preprocessing before sending the image to the inference module are the responsibilities of the mobile application. This layer facilitates usability and accessibility but at the same time provides a smooth interface between the user and the underlying classification models.

The core processing occurs at the On-Device Inference Layer, during which a multi-stage pipeline made of a Convolutional Neural Network (CNN) is used to execute, based upon optimized TensorFlow Lite models. The first stage is to determine a Cat versus Non-Cat classification model whereby a non-relevant image is not subjected to further processing since it is clear whether the input image entails a cat or not. Categorized pictures with cats are then forwarded to the second stage, which classifies the twelve classes of cat breeds using a CNN that had been trained on the Oxford-IIIT Pet Dataset. In order to mitigate the issue of fine-grained classification, especially when dealing with visually similar breeds, a third classification step is added to reassign chosen outputs to the category Siamese, Ragdoll, or Birman with a specific similar-breed model. Such a staged methodology helps to increase the sturdiness, reduces the chances of misclassification, and results in better accuracy of the entire system.

The last system layer is the Local Storage and Output Layer that manages the presentation of results and storage of data. After classification, the recognized breed name, score of confidence, and the corresponding metadata are retrieved and displayed in the mobile application. The results are saved locally on the machine to facilitate offline inspection and a record of the result. The design will provide low latency and user privacy through the absence of cloud dependency and will be able to be deployed in a real-world environment with constraints on internet connectivity. In general, the proposed architecture is a compromise between computational performance, classification, and usability, which makes it appropriate for real-world cat breed identification based on mobile.

It uses lightweight CNNs and inference on devices, thus removing the need to communicate constantly to the server and making the architecture optimized to run on mobile platforms. The multi-stage method is more accurate, and it is still computationally efficient. A combination of usability, performance, and scalability, the system architecture offers a powerful base for a practical, functional cat breed identification implementation.

3.8.2 Application Flow and Model Integration

This part outlines the functionality of the suggested Cat Breed Identification mobile app and defines how the trained Convolutional neural Network (CNN) models can be incorporated into the system. The application flow can present the user with a more detailed running procedure, but one that is made to be easy and simple, as well as making certain that the image inputs are correctly processed through the network of Stage CNN classification.

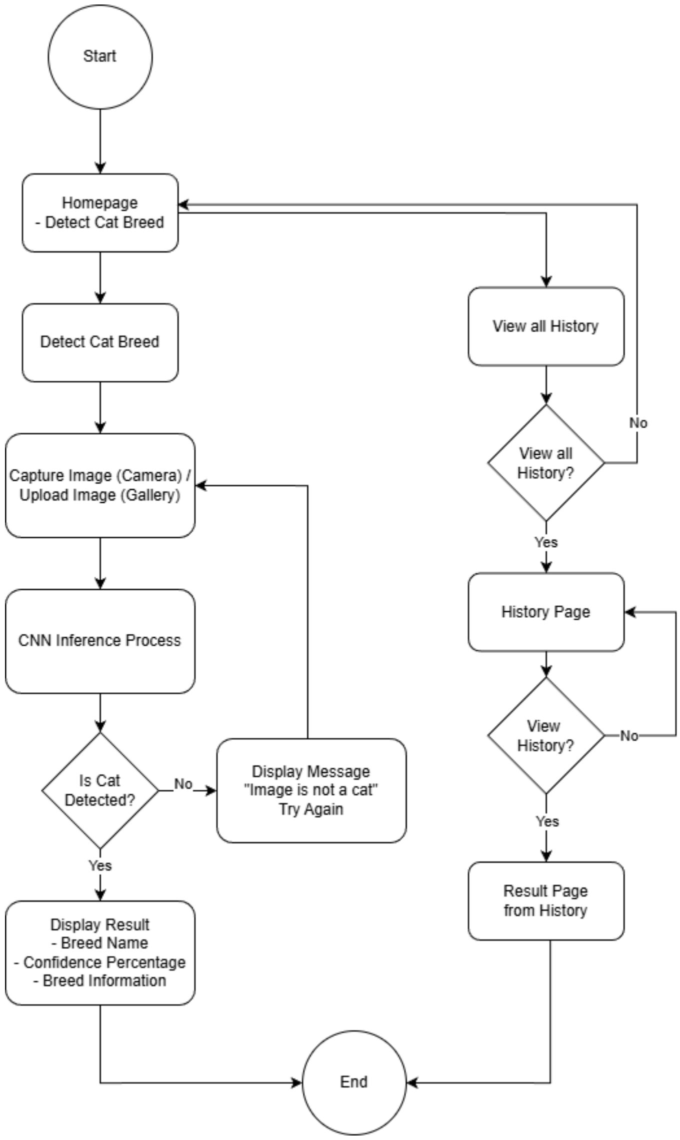

Figure 3.5 Application Flow and Model Integration Process Diagram

Figure 3.5 represents a general scheme of application flow and model integration of the proposed system. The process starts when the user opens the application that takes them to the home page, where the main feature, Detect Cat Breed, is introduced. Through this interface, one can take an image using the device camera or post an already existing image to the device gallery. This flexibility will allow the application to suit the real-life usage scenarios in cases where images are captured under different conditions.

When an image is provided, the input is sent to the Multi-stage CNN Model Integration module, which is the main intelligence of the system. The first phase in this module is the Cat vs Non-Cat classification model, which is a filtering system to decide whether the image being fed has a cat. When the image is categorized as a non-cat, the system will automatically respond with a feedback message stating that the image is an invalid image and the user is asked to rescan. This step will ensure that non-cat images are not subjected to further processing, and in this way, system reliability is improved and false breed predictions are minimized.

In case images are correctly identified as cats, the system proceeds to the second stage, which is the Cat Breed Classification model that is trained on twelve breeds of cats based on the Oxford-IIIT Pet Data. The model makes the most likely prediction of the breed and creates a confidence value of the classification. In case the expected breed is in a visually similar group, a further refinement step is applied with a specialized similar-breed classification model that is trained with the breeds of Siamese, Ragdoll, and Birman. This is a hierarchical method, which is better in classification accuracy for fine-grained breed recognition problems.

Once classification is done, the results are shown to the user in the prediction of the breed name, the percentage of confidence, and the details pertaining to the breed. The entire prediction outcomes are automatically stored by default in the history module of the application. With the View History option, users are able to access previous history where records of classifications are stored and can be revised or revisited. This capability eases the user experience and allows one to trace past detections.

In conclusion, the combination of the application flow and multi-stage CNN models will guarantee the robust, efficient, and user-friendly system. The step-by-step approach to Cat versus Non-Cat classification, multi-breed classification, and its refinement shows the presence of a well-organized model integration strategy, which is in accordance with the aims of the given project and allows the solution to be deployed with high trust in the context of the mobile application.

3.8.3 UI Design Overview

The user interface (UI) of the proposed Cat Breed Identification System has been designed to be highly focused on simplicity, usability, and accessibility to enable the consideration of an effective application by a wide range of stakeholders, including pet owners, veterinarians, and animal shelter workers. The interface is minimalistic and user-friendly in design, thus simplifying the interaction aspects and making certain that all the necessary features are easily reached. This design fits the main purposes of the system, which are to identify a cat, classify its breed, visualization the results, and also spread useful information.

The Homepage will serve as the main entry point into the application and will provide users with an easy-to-navigate list of the key features to access the main functionality. The design of this page can be described as a simple design with clearly marked buttons permitting users to move forward to image upload or capture functionality, reviewing classification history or informational content related to the application. The design is carefully simple to allow users who are using the application for the first time to understand the purpose of the application without necessarily referring to more instructions. The following Figure 3.6 represents the home page interface of the mobile application.

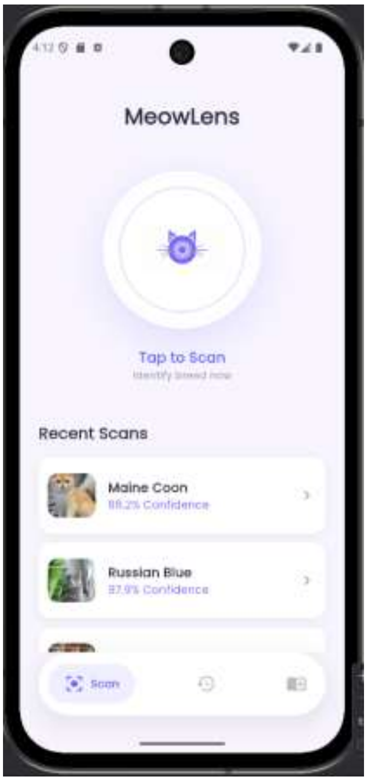

Figure 3.6 Homepage

The Upload/Capture Page allows user, either to upload an image stored in their gallery on the device or even take a new image through the device camera lens. Such a design decision contributes to the flexibility of the system and digs into its applicability to the real-life situation. When the image is selected or captured, the system will automatically load the image preprocessing and classification pipeline, which will include the Cat versus Non-Cat detection followed by adding the cat breed classification. The interface is specifically minimal in order to minimize user error, as well as to support smoother interaction between the user and underlying convolutional neural network (CNN) models. Figure 3.7 displays the design of this page.

Figure 3.7 Upload/Capture Page

Once the classification process has taken place, in the Result Page is displayed the expected output generated by the system. This page will give the identified category, Cat or Non-Cat, the identified predicted cat breed, and the confidence score, which would improve the transparency of the result and confidence in the user. The result page design is based on visual clarity to enable the user to quickly interpret the result of the classification, and it also gives the user the option to save the result or start a new classification. The interface of the result of the classification is illustrated in Figure 3.8.

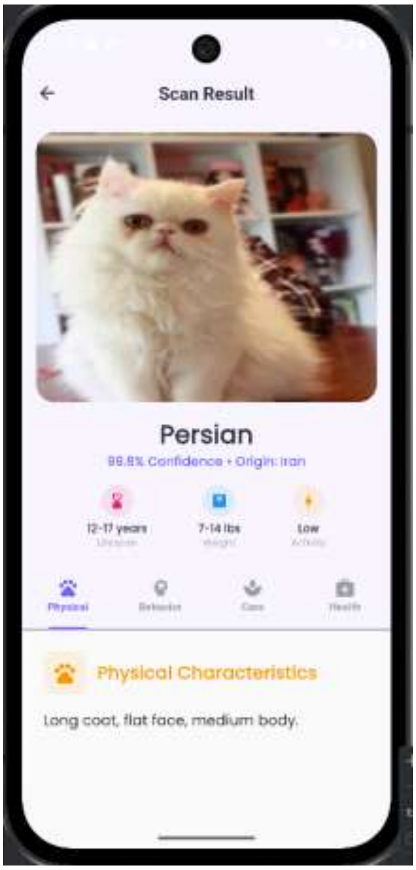

Figure 3.8 Result Page

The History Page enables users to see the already classified images and their predicted outcomes. It is especially useful to those users who make a number of classifications repeatedly and may want past results to be stored for reference or comparison. The saved history encourages the user to interact and empowers the book to real-life situations like keeping track of more than one cat or reviewing previous forecasts. The history page user interface is as seen in Figure 3.9.

Figure 3.9 History Page

Lastly, the Info Page provides the end user with extra information about the application, such as a summary of the working process within the application, provenance of data sets, restrictions of the model, and general usage description. The page is crucial in improving the user's understanding and managing expectations with regard to model performance, particularly in the context of mixed breeds or poor-quality images. Figure 3.10 displays the design of the information page.

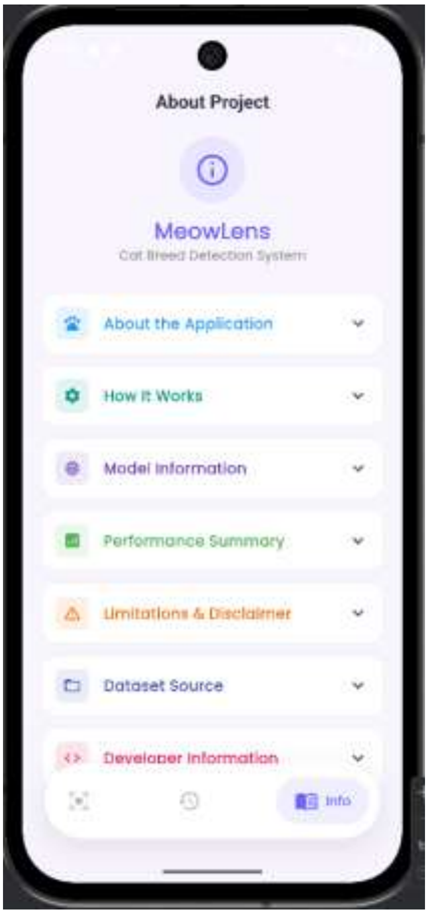

Figure 3.10 Info Page

In general, the proposed system will have a user interface design that will allow a user to interact with the deep learning models efficiently. The combination of a minimalistic design, natural flow of navigation, and informative display of the results makes the performance of the application effective in closing the gap between complex CNN-based classification procedures and the convenient mobile interface, thus making the system fixable to a real-life implementation.

3.9 Mobile Development

The following section outlines the system development process of the proposed Cat Breed Identification mobile application. The development will revolve around the implementation of the mobile application, integration of the trained Convolutional Neural Network (CNN) models, and implementation of on-device inference to facilitate offline application.

The Flutter version of Android Studio was used to develop the mobile application. Flutter has been chosen because of its ability to facilitate cross-platform development and its effectiveness in the process of rendering user interfaces. The application was to be used to facilitate basic tasks such as the ability to capture an image using the device camera, the ability to pick an image that is in the gallery, displaying the results, and the ability to view a history of identification.

The trained CNNs were written in the TensorFlow framework and transformed to TensorFlow Lite (TFLite) format in order to be able to run them on the mobile devices. This conversion provided a smaller model size, better inference, and the same classification performance. The TFLite models were then implemented into the mobile application with the TensorFlow Lite interpreter in order to provide real-time inference on the device.

The system was developed so that it would be completely offline without depending on cloud solutions or the internet connection. When an image input is given to the application, lightweight preprocessing is executed, after which inference is made using the integrated TFLite model. The user is then given the predicted cat breed and the associated confidence score. This physical deployment is responsive to the system, manages user privacy, and is accessible in a setting where network connectivity is limited.

It was also enhanced with a local storage system, which stores user detection history. The relevant data containing picture reference, breed and predicted breed, confidence level, and time is stored locally through a structured database mechanism. This also allows users to look at past detection results and basic record management in the application.

To summarize, the system development phase was effective, and it led to the creation of a working mobile application that was integrated with trained CNN models to perform offline cat breed classification. The application shows a successful combination of developing mobile applications with deep learning models, which meets the project goals and system requirements.

3.10 Documentation

Documentation plays a critical role in making sure that the details of the developmental process, system design, and implementation of the project are carefully documented and organized in a systematic manner. Detailed documentation was generated in the current project to describe all stages of the Cat Breed Identification System, starting with problem recognition, the creation and development of models, integrating the system, and its evaluation. This documentation can be seen as an official source of academic evaluation and future system improvement.

In the documentation, the entire exposition of the research methodology, dataset selection processes, image preprocessing methods, the development of the convolutional neural network model, the training setup, and model analysis outcomes are provided. All design decisions and algorithms are chosen, and strategies of implementation are recorded in order to make them transparent and reproducible. They include figures, tables, and diagrams where appropriate in order to enable an easy interpretation of the system architecture, flow of application, and design of the user interface.

Along with the technical documentation, user-oriented descriptions are provided as well to describe how the mobile application works. These descriptions include application workflow, image input process, prediction output mechanisms, and history management features. This type of documentation will guarantee that the system can be easily learned and utilized by workers who do not have technical abilities in the field of deep learning and the development of mobile applications.

On the whole, the documentation is quite appropriate, clear, and complete in respect of the project report. It presents a systematic history of the way development occurred, and it is a starting point for improvement, repairs, or additions to the Cat Breed Identification System in the future.

3.11 Summary

The chapter has described the research procedure that is used in developing the proposed Cat Breed Identification System. It was presented in a systematic and orderly way with the research methodology design as the first step, followed by the data acquisition, image preprocessing, CNN model design, and model analysis. Two publicly available datasets were used to aid the classification tasks so that models would be trained and tested on a variety of well-labelled image data.

The chapter also explains the creation and testing of several CNN networks, namely MobileNetV3 and ResNet50, that were developed, trained, and evaluated based on the relevant performance measures with the purpose of determining the most productive model to be deployed. Additionally, the design and implementation of the system were described in more detail, with a focus on the integration of the trained CNN model into a mobile app that could conduct offline cat breed classification.

In general, the research methodology involved in this chapter is strong enough to support the research objectives of the project. The outcome of the model assessment and system implementation is given and addressed in the following chapter.

CHAPTER 4
RESULTS AND ANALYSIS

4.1 Introduction

This chapter highlights the experiment and result performance analyses of the Convolutional Neural Network (CNN) models designed in the Cat Breed Identification System. The aim of the chapter is to review the effectiveness of the suggested models to classify images of cats and to review the results of the conducted experiments in accordance with the developed methodology presented in Chapter 3.

This chapter has been initiated by a description of the experimental setup and the evaluation measures used to evaluate the performance of the models. It then reports the results of the classification of the Cat vs Non-Cat model, the 12-breed classification model of cats, and the similar breed classification model. The comparative analysis of MobileNetV3 and ResNet50 models is provided to determine the most appropriate model in terms of accuracy and performance in terms of computational efficiency.

Furthermore, this chapter will explore the outcome of the implementation of the mobile application with a specific focus on the functionality of the combined CNN model in an offline setting. Results are criticized and explained to demonstrate how the research aims and goals were achieved and provide information about the weaknesses and shortcomings of the suggested system.

4.2 Experimental Setup

The experimental plan that was implemented during the current inquiry was designed with an objective of evaluating the functionality of the suggested convolutional neural network (CNN) models with respect to the classification of cat images. Two different datasets were used, which are the Oxford-IIIT Pet Dataset, which was used as a breed-level classifier, and the Animal Image

Dataset, which was used as a cat vs non-cat binary classifier. All the datasets were split into training, validation, and testing in 70%, 15%, and 15%, respectively, which ensured an objective assessment of the data not met during training.

The experimental plan that was implemented during the current inquiry was designed with an objective of evaluating the functionality of the suggested convolutional neural network (CNN) models with respect to the classification of cat images. Two different datasets were used, which are the Oxford-IIIT Pet Dataset, which was used as a breed-level classifier, and the Animal Image Dataset, which was used as a cat vs non-cat binary classifier. All the datasets were split into training, validation, and testing in 70%, 15%, and 15%, respectively, which ensured an objective assessment of the data not met during training.

4.3 Performance Evaluation Metric

The effectiveness of the suggested CNN models was measured with the help of some common indicators of the image classification process. These metrics have been chosen to offer an overall evaluation of the predictive quality of the models, class-wise performance, and the reliability of the model predictive quality when used on unknown test data.

The accuracy was used as the main measure of the overall correctness of the classification outcomes since it was calculated in terms of the number of the correctly predicted images divided by the total number of the test samples. Although accuracy provides an overall idea of how well a model behaves, it might not help to explain the behaviour of the model in a class-wise manner, especially in multi-class model category scenarios.

Precision and recall were also incorporated in order to assess the quality of the prediction in terms of each class. Precision measures the percentage of accurately correctly predicted examples of all of the examples predicted as of a given class, and recall measures the capability of the model to correctly recognize all the relevant instances that are in a given class. The metrics are particularly relevant to comprehend the level of differentiation of breeds of cats with similar visual appearances by the model.

To provide a balanced score of the performance of the model across all classes, F1-score showing the harmonic mean of precision and recall, was employed. This is the specific measurement that can be used to evaluate multi-class trained models in particular, since it takes into consideration false positives and false negatives.

Also, the confusion matrix analysis was used to graphically look at the performance of classification among various classes. The confusion matrix provides information about the most frequent misclassification patterns and enables to understand which cat breeds the model is more likely to be confused with. Collectively, these metrics of evaluation are a holistic plan of how the performance of the CNN architectures will be compared, and a more convenient model may be chosen to be deployed on the mobile application.

4.4 Results

In this part, the experimental findings are provided based on the analysis of the suggested multi-stage cat recognition system. Three main classification tasks that are studied and analyzed in relation to the developed Convolutional Neural Network (CNN) models are the Cat versus Non-Cat classification, and the twelve-breed classification and the similar breed classification which is performed based on the visually similarity of the cat breeds. The results of each subsection are reported using conventional performance measures like accuracy, precision, recall, F1 -score, trends in training and validation accuracy, and confusion matrix analysis. These findings are argued to determine the effectiveness and reliability and generalization capacity of the models in different experimental conditions. Results of these experiments are used to determine the most appropriate model configurations to be incorporated in the proposed mobile application.

4.4.1 Cat vs Non-Cat Classification

This part shows an analysis of the results of the experiments carried out in the Cat vs Non-Cat classification task. This is the initial part of the proposed hierarchical classification scheme in which the goal is to effectively differentiate between images that have cats and images that do not have cats. The performance at this stage is of utmost importance since it will directly influence the reliability of the further tasks based on the cat classification. Two deep learning models, that is, MobileNetV3 and ResNet50, were tested on several experimental conditions to determine their performance, stability, and generalization property in binary classification.

4.4.1.1 MobileNetV3

To determine the effects of changing the batch size, the number of training epochs, and the learning rate, the experiments were set to examine how these factors affect the model and loss. Based on the findings provided in Table 4.1, it can be seen that MobileNetV3 has very high performance in nearly all the settings. Most of the experiments had a test accuracy of over 99%, which indicated that MobileNetV3 was suitable in binary image classification.

Table 4.1 Summarizes the performance of all the MobileNetV3 experimental settings on Cat vs Non-Cat classification task

Experiment	Batch Size	Epoch	Learning Rate	Test Accuracy	Test Loss
1	32	10	0.001	99.88%	0.0158
2			0.0001	99.75%	0.0161
3			0.00001	98.89%	0.0545
4		15	0.001	99.88%	0.0128
5			0.0001	99.88%	0.0126

6			0.00001	98.89%	0.0508
7		20	0.001	99.88%	0.0121
8			0.0001	99.88%	0.0131
9			0.00001	98.89%	0.0358
10	64	10	0.001	99.88%	0.0099
11			0.0001	99.38%	0.0204
12			0.00001	98.89%	0.0653
13		15	0.001	99.88%	0.0121
14			0.0001	99.88%	0.0116
15			0.00001	98.89%	0.0641
16		20	0.001	99.88%	0.0150
17			0.0001	99.88%	0.0163
18			0.00001	99.88%	0.0169

Looking more closely at Table 4.1, it can be seen that the experiments with a learning rate of 0.001 always yielded the most stable and reliable results. For example, experiments 1, 4, 7, 10, 13 and 16 all had a test accuracy of 99.88 with test loss ranging between 0.0099 and 0.0158. However, experiments that were trained with a very low learning rate of 0.00001 produced significantly increasing values of test loss. Experiment 12 in particular gave a test loss of 0.0653, although it had a test accuracy of 98.89, which is slower converging and less certain predictions. This tendency implies that excessively low learning rates have a deleterious impact on the optimization of the process of this task.

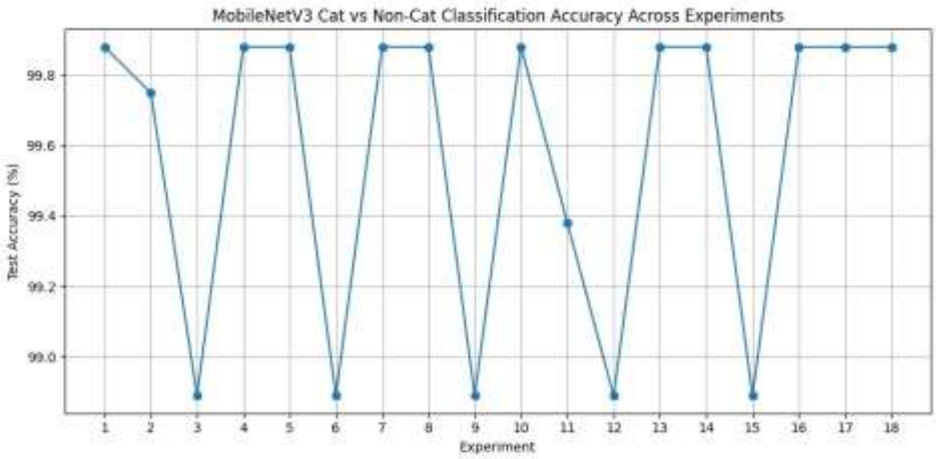

Figure 4.1 Comparison accuracy graph of MobileNetV3 across all the experiments on the Cat vs. Non-Cat Classification task

Figure 4.1 indicates that the test accuracy is constant and significantly steady all through most experiments with slight variances. The reduced accuracy in a few cases is mostly found with the experiments conducted with the smallest learning rate, which is consistent with the higher values of test loss as in Table 4.1. On the whole, the illustration shows that MobileNetV3 is very sustainable and quite unresponsive to the value of batch size and the number of epochs under the condition of the presence of the appropriate learning rate.

Experiment 10 was chosen as the most suitable experiment configuration in terms of the additional analysis. The batch size, 10 training cycles, and learning rate used in this experiment were 64, 10, and 0.001, respectively. It has obtained the highest test accuracy of 99.88% with the lowest test loss of 0.0099 among all the settings. Such a combination of very high accuracy with very low loss points to an ideal balance between classification and generalization, and Experiment 10 is the most appropriate one to be highly evaluated.

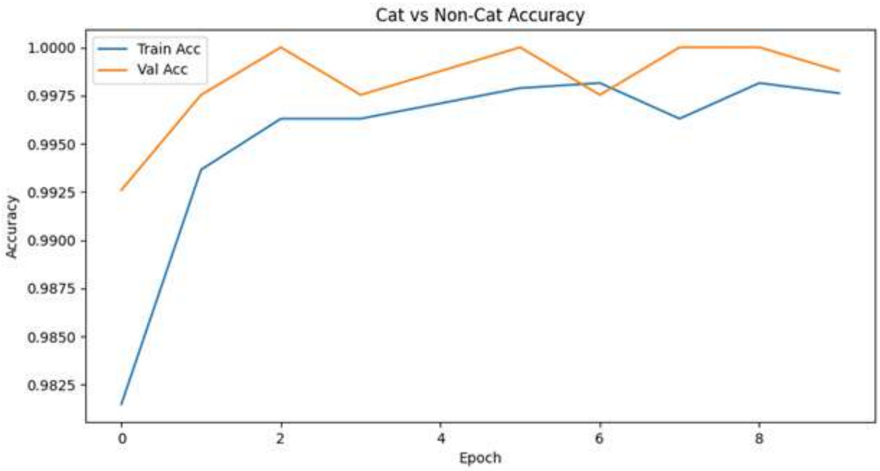

Figure 4.2 Training and validation accuracy of the chosen MobileNetV3 model (Experiment 10)

Based on Figure 4.2, it is possible to notice that training and validation accuracies grow at a high rate in the early epochs, reaching values over 99% in the first couple of epochs. For stable training, stability is reached in the two curves where there is no major deviation between the two curves. Such a close fit is a sign of a stable process of learning followed by little overfitting. The fact that the curves converge smoothly proves that the model can effectively learn discriminatory features, but the model is highly generalized to the unseen data.

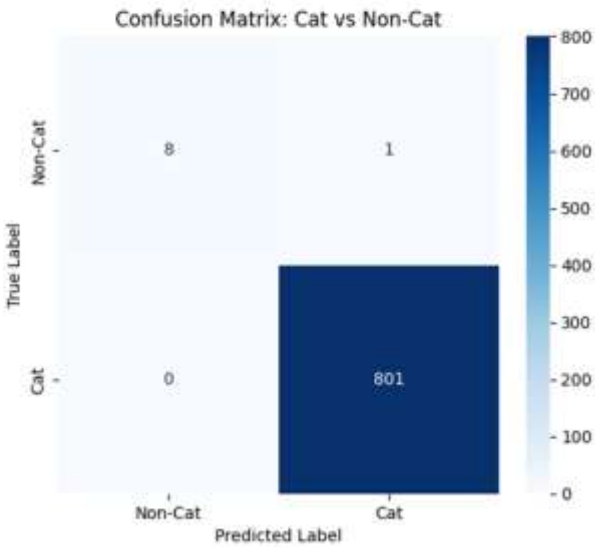

Figure 4.3 Confusion Matrix of the chosen MobileNetV3 model in Cat vs Non-Cat Classification (Experiment 10)

According to the confusion matrix, most of the predictions are correctly classified because of the large proportion of the numbers along the diagonal. The smallest Cat class with the 801 test samples has almost perfect classification with few misclassified cases. The Non-Cat group only has 9 samples, and that is why we can say that the recall value is relatively smaller with respect to this category. Although this is unbalanced, the overall classification of the model is high.

Table 4.2 MobileNetV3 Cat vs Non-Cat Classification Report (Experiment 10)

Class	Precision	Recall	F1-Score	Support
Non-Cat	1.00	0.44	0.62	9
Cat	0.99	1.00	1.00	801
Accuracy	-	-	**0.99**	**810**
Macro Avg	1.00	0.72	0.81	810
Weighted Avg	0.99	0.99	0.99	810

The Cat class is characterized by the highest precision of 0.99, a recall of 1.00, and an F1-score of 1.00, which means that the model is very useful at recognizing cat images. The Non-Cat class is expected to have a lower value of precision (1.00) and recall (0.44) that is achieved. The total classification accuracy is 99% on 810 test samples, and the weighted precision, recall, and F1-score are all equal to 0.99. Such findings affirm the notion that MobileNetV3 is the most promising in the process of Cat vs Non-Cat classification, and it is much more appropriate to use as a filtering step in the hierarchical classification pipeline.

Overall, the experimental results confirm that MobileNetV3 is highly suitable for the Cat vs Non-Cat classification task. Its lightweight architecture enables very high accuracy while maintaining computational efficiency, making it ideal for on-device mobile deployment. The selected configuration serves as an effective preprocessing and filtering stage, ensuring that only valid cat images are passed to the subsequent cat breed classification models

4.4.1.2 ResNet50

The following subsection results in the close evaluation of the ResNet50 architecture to the Cat vs Non-Cat classification task. ResNet50 is a more profound convolutional neural network, which uses residual connections to allow very deep architectures to be trained by relieving the issue of vanishing gradient. ResNet50 has strong potential in terms of its depth and capacity to represent, so it is likely to achieve strong performance in binary classification problems, especially with adequate training data and the correct hyperparameter settings. As in MobileNetV3, it used 18 experimental settings by experimenting with the batch size, training epochs, and learning rate to experimentally quantify how these three factors influence the classification performance.

Table 4.3 Summarizes performance of all the ResNet50 experiments on the Cat vs Non-Cat classification task

Experiment	Batch Size	Epoch	Learning Rate	Test Accuracy	Test Loss
1	32	10	0.001	99.88%	0.0124
2			0.0001	99.63%	0.0116
3			0.00001	98.89%	0.0471
4		15	0.001	99.88%	0.0100
5			0.0001	99.88%	0.0159
6			0.00001	98.89%	0.0379
7		20	0.001	99.75%	0.0071
8			0.0001	99.88%	0.0072
9			0.00001	99.01%	0.0356
10	64	10	0.001	99.88%	0.0059
11			0.0001	99.75%	0.0093
12			0.00001	98.89%	0.0614
13		15	0.001	99.88%	0.0071
14			0.0001	99.88%	0.0069
15			0.00001	98.89%	0.0473
16		20	0.001	99.88%	0.0081

17			0.0001	99.88%	0.0060
18			0.00001	98.89%	0.0427

Based on the table, it is possible to note that ResNet50 is expected to have extremely high values of test accuracy, and in most of the experiments, the value of test accuracy is over 99%. This implies that the ResNet50 is very effective in learning discriminative features that differentiate cat images and non-cat images. Specifically, experiments that are trained at a 0.001 learning rate have the most consistent and predictable performance, with a high accuracy level and a low-test loss rate. As an example, experiments 1, 4, 7, 10, 13 and 16 all had a test accuracy of 99.88% with a test loss of between 0.0059 and 0.0100.

Conversely, the smaller learning rates of 0.0001 and 0.00001 in the experiment is significantly learning more at the expense of higher values in the test losses, even when obtaining a relatively high accuracy. To illustrate, in Experiment 12, where the learning rate was 0.00001, the test accuracy was 98.89%, but the test loss was significantly high, at 0.0614. This tendency shows lower convergence and a decrease in the level of confidence in predictions, indicating that too-small learning rates would not be optimal and the ResNet50 would not be trained on this binary classification task.

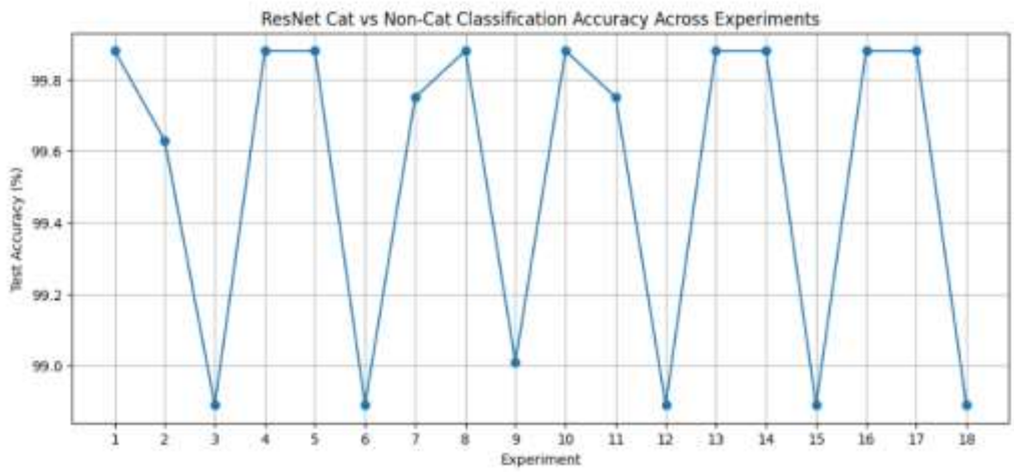

Figure 4.4 Comparison accuracy graph of ResNet50 across all the experiments on the Cat vs. Non-Cat Classification task

Figure 4.4 shows the graph of the accuracy of comparison of ResNet50 under all experimental settings. The graph shows that there are high accuracy trends throughout the majority of the experiments, and the only slight changes can be noticed in experiments being trained with smaller learning rates. According to the comparison of MobileNetV3 and ResNet50, the accuracy curve of the latter has fewer points of variance, which means a higher degree of robustness and stability. The numerical results in Table 4.3 are backed by this visual trend, which points to the credibility of ResNet50 when used in the Cat vs Non-Cat classification.

Experiment 10 was selected as the best ResNet50 model in all the experimental settings and used to conduct additional analysis. The batch size used in this experiment was 64, with 10 training epochs, and the learning rate has been set to 0.001. Its test accuracy was 99.88%, and the lowest test loss was recorded as 0.0059 throughout and in all the experiments of ResNet50. A combination of very high accuracy and low loss implies this configuration will give the best trade-off between the model performance and the ability to generalize.

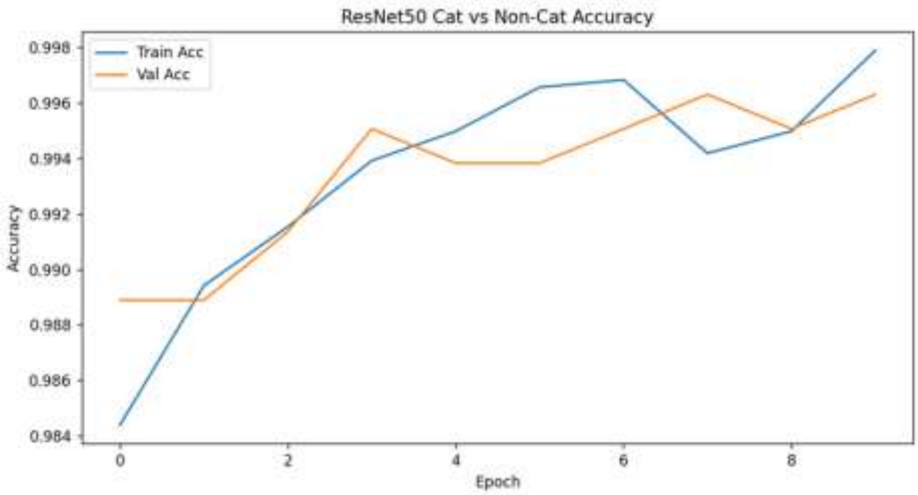

Figure 4.5 Training and validation accuracy of the chosen ResNet50 model (Experiment 10)

The training and validation accuracy curves of the chosen ResNet50 model of Experiment 10 are shown in Figure 4.5. As shown in the figure, training and validation accuracy are growing fast within the first few epochs and are stabilized at high values as the training proceeds. Most of the training and validation curves coincide with each other to show that it is a steady learning process with little overfitting. In line with this, the results of the training and validation loss curve fall slowly and stabilize towards low values, which indicates successful optimization and excellent generalization.

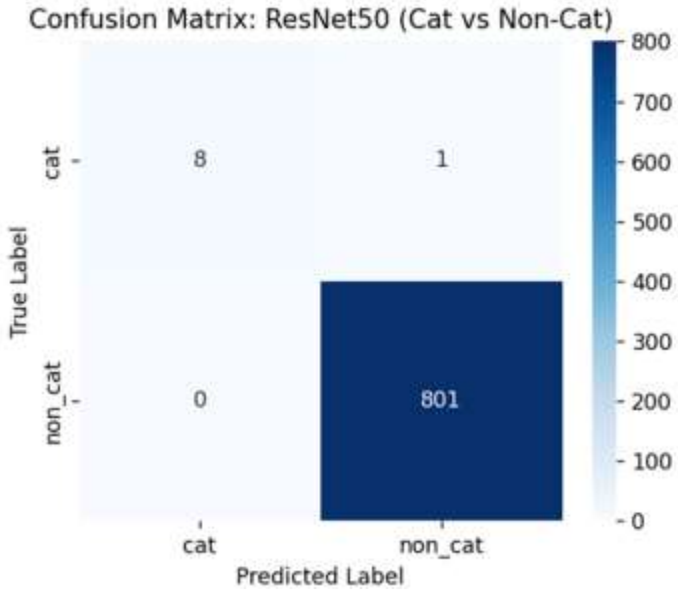

Figure 4.6 Confusion Matrix of the chosen ResNet5050 model in Cat vs Non-Cat Classification (Experiment 10)

Figure 4.6 shows the confusion matrix of the ResNet50 model with the Cat vs Non-Cat data classification task with Experiment 10. The confusion matrix shows that almost all the predictions are correctly assigned, which is supported by the prevalence of diagonal elements. The true positives of both Cat and Non-Cat classes are extremely high with minimal misclassifications. This proves that with high confidence, the model can categorize the two classes.

Table 4.4 Classification report of the chosen ResNet50 on Cat vs Non-Cat Classification task (Experiment 10)

Class	**Precision**	**Recall**	**F1-Score**	**Support**
Non-Cat	1.00	0.89	0.94	9
Cat	1.00	1.00	1.00	801
Accuracy	-	-	**1.00**	**810**
Macro Avg	1.00	0.94	0.97	810
Weighted Avg	1.00	1.00	1.00	810

Table 4.4 shows the classification report of the ResNet50 model on the Cat vs Non- Cat classification task according to Experiment 10. According to the classification report, the precision, recall, and F1-score of the Cat as well as Non-Cat classes are not less than 1.00 which means that the classification performance is perfect. The model has an overall accuracy of 100 percent in 810 test samples, although the macro and weighted average of the precision, recall, and F1-score also achieve the value of 1.00. These findings show that ResNet50, although with minimal difference, is better than MobileNetV3 on Cat vs Non-Cat classification accuracy and loss minimization, but with an increment of computation and model size.

4.4.1.3 Cat vs Non-Cat Result

The subsection is a summary of the highest-performing experimental settings of the Cat vs Non-Cat classification task with MobileNetV3 and ResNet50 as architectures. Using the results of the experiment in the above subsections, the best model to use was determined in terms of the best test accuracy and minimum test loss. This comparison is aimed at determining the effectiveness of the two architectures in regard to differentiating between cat and non-cat images as well as the model that is best suited to develop the system. Table 4.5 shows the best results of the selected Cat vs Non-Cat classification task that were obtained in terms of MobileNetV3 and ResNet50.

Table 4.5 Best Performance Results for Cat vs Non-Cat Classification

Model	Experiment	Batch Size	Epoch	Learning Rate	Test Accuracy	Test Loss
MobileNetV3	10	64	10	0.001	99.88%	0.0099
ResNet50	10	64	10	0.001	99.88%	0.0059

Based on the Table 4.5, the classification performance of both models was very good as they had the same test accuracy of 99.88% meaning that they are highly able to differentiate between cat and non-cat images. Nevertheless, ResNet50 reduced the test loss than MobileNetV3, indicating more confident and reliable forecasts. In spite of this, MobileNetV3 is applicable to this type of classification since, even at this stage of the classification, it has a reasonable lightweight architecture and efficiency that is beneficial to the implementation of mobile applications. Hence, the two models can be deemed to be useful in the context of Cat vs Non-Cat classification, as MobileNetV3 can be chosen to be integrated into the proposed system because of its performance-computational efficiency ratio.

4.4.2 Cat Breed Classification (12 breeds)

The section is a detailed analysis of the proposed deep learning models to solve the Cat Breed Classification with 12 different breeds of cats. Compared to the binary Cat vs Non-Cat classification task, the task is much more complex because of the greater number of classes and great visual proximity of some cat breeds. Fine-grained classification is more complicated due to subtle differences in fur patterns, facial structure, and body features and highlights the fact that more discriminative and detailed representations of features need to be learned by the model. Consequently, such performance measures as accuracy and loss are more likely to differ between experimental settings. The following section examines the performance of MobileNetV3 in greater detail and then compares it to ResNet50 in the following subsection.

4.4.2.1 MobileNetV3

A summary of the 18 experimental settings using MobileNetV3 to complete the 12-breed cat classification task is provided in Table 4.5. The experiments had the properties of changing the batch size, the number of training epochs, and the learning rate in a systematic manner to determine their influence on loss and the accuracy of classification. Based on Table 4.5, it is possible to note that MobileNetV3 is able to show good performance in this multiclass task with strong performance values of 62.90% up to 96.68% when it is trained with the right learning rate.

Table 4.6 Summarizes the performance of all the MobileNetV3 experimental configurations on the Cat Breed Classification (12 Breeds) task

Experiment	Batch Size	Epoch	Learning Rate	Test Accuracy	Test Loss
1	32	10	0.001	96.68%	0.1201
2			0.0001	92.91%	0.2225
3			0.00001	74.66%	0.9914
4		15	0.001	96.53%	0.1171
5			0.0001	94.12%	0.1999
6			0.00001	82.81%	0.7103
7		20	0.001	96.08%	0.1439
8			0.0001	95.02%	0.1668
9			0.00001	82.96%	0.5772
10	64	10	0.001	95.78%	0.1557
11			0.0001	91.70%	0.2769
12			0.00001	62.90%	1.4176
13		15	0.001	96.08%	0.1287
14			0.0001	92.46%	0.2172
15			0.00001	71.64%	1.041
16		20	0.001	95.32%	0.1376
17			0.0001	93.51%	0.1977
18			0.00001	80.09%	0.8216

A detailed analysis of Table 4.5 reveals that training with a learning rate of 0.001 always provided the best accuracy and the minimum values of test loss. Experiment 1 with a batch of 32, 10 training epochs, and a learning rate of 0.001 gave the maximum test accuracy of 96.68 percent and a test loss of 0.1201. Experiments 4, 7, 10, 13 and 16, with a comparably low learning rate (0.001), similarly yielded similar values between 95.32% and 96.53% and lower values of test loss of between 0.1171 and 0.1557. These findings suggest that MobileNetV3 can learn discriminative features of fine-grained classification with a large enough learning rate.

However, experiments with smaller learning rates, 0.0001 and 0.00001, demonstrate performance deterioration. An example of this is Experiment 12 that used a learning rate of 0.00001 since it achieved the lowest test accuracy of 62.90% and the highest test loss of 1.4176. This great decrease in performance indicates that the model could not converge effectively, as the weight updates were slow. Even with a learning rate of 0.0001, experiments such as Experiment 11, and Experiment 14 recorded lower accuracy scores of 91.70% and 92.46%, respectively, and had higher test loss values than the test losses of Experiment 14 using a learning rate of 0.001.

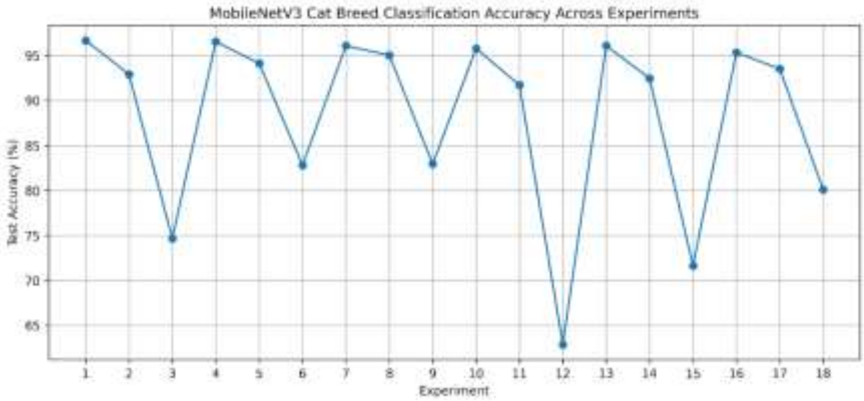

Figure 4.7 Comparison accuracy graph of MobileNetV3 across all the experiments on the Cat Breed Classification (12 Breeds) task

Figure 4.7 shows the comparison accuracy graph of MobileNetV3 under all experimental configurations of the 12-breed classification task. It can be clearly seen in the graph that the experiments that have a learning rate of 0.001 always perform better than the experiments of lower learning rates. The values of accuracy are concentrated around 95% to 97% of the learning rate 0.001 and the experiments with the learning rate 0.00001 are characterized by the drastic drop in the accuracy rates, as in some cases they fall below 75%. Such visualization supports the numerical patterns in Table 4.5 and shows the importance of the selection of the learning rate in multiclass classification.

According to the outcomes in Table 4.5 and Figure 4.7, Experiment 1 was chosen as the best set of MobileNetV3 to continue the analysis. The test accuracy was the highest in this experiment, 96.68%, with the test loss of a relatively low value of 0.1201, thus, it demonstrated excellent generalization.

Moreover, a smaller batch size of 32 can be used, and this provides the model with more mixed gradient updates, which can lead to improved performance on fine-grained classification experiments.

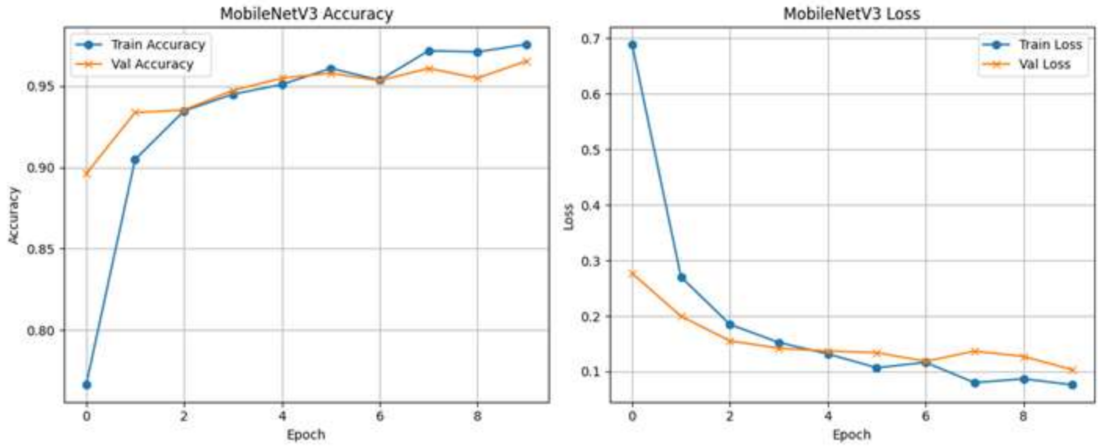

Figure 4.8 Training and validation accuracy of the chosen MobileNetV3 model (Experiment 1)

The training and validation accuracy curves of the chosen MobileNetV3 model of Experiment 1 are depicted in Figure 4.8. The curves demonstrate that the accuracy of training and validation is steadily increasing during the course of training and reaches its high point near the end, during which its values have stabilized at about 97. Importantly, both training and validation curves follow a very similar direction, and no considerable divergence has been noticed. This is a good sign that this model is not severely overfitting and can be generalized well to unseen test data.

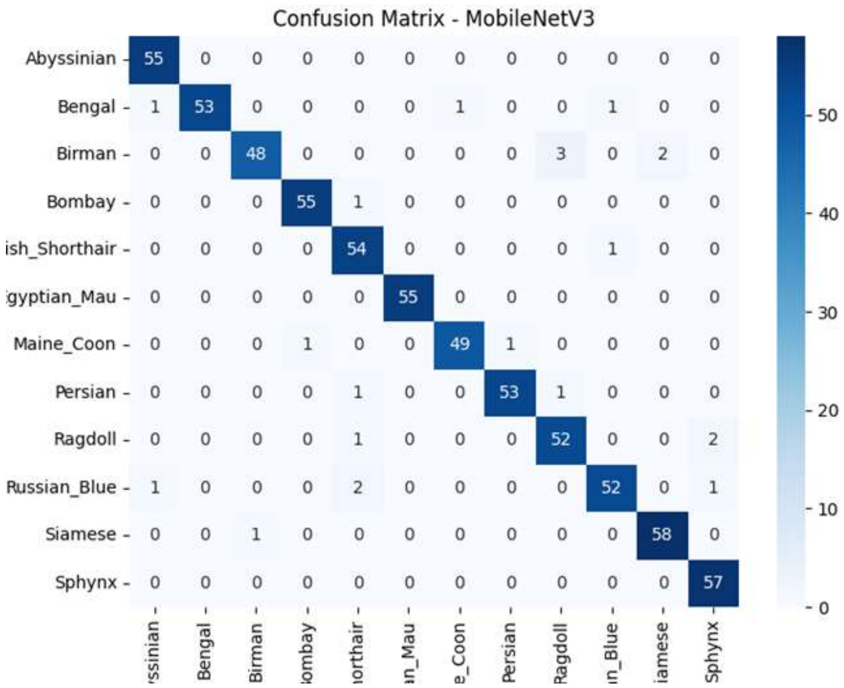

Figure 4.9 Confusion matrix of the chosen MobileNetV3 model (Experiment 1)

Figure 4.9 presents the confusion matrix of the chosen MobileNetV3 model in the task of 12 breed classification. As shown in the confusion matrix, the greatest number of predictions is clustered around the diagonal, meaning that there are numerous correct predictions made across all breeds. Nevertheless, there can be noticed some cases of misclassification when it comes to breeds that are visually close to each other, which is understandable considering the fact that some cat breeds are subtly different. Such misclassifications reflect the nature of the difficulty of fine-grained classification and further support the idea that more sophisticated models or more methods of improving features are required.

Table 4.7 Classification report of the chosen MobileNetV3 on the Cat Breed Classification (12 Breeds) task (Experiment 1)

Class	Precision	Recall	F1-Score	Support
Abyssinian	0.96	1.00	0.98	55
Bengal	1.00	0.95	0.97	56
Birman	0.98	0.91	0.94	53
Bombay	0.98	0.98	0.98	56
British Shorthair	0.92	0.98	0.95	55
Egyptian Mau	1.00	1.00	1.00	55

Maine Coon	0.98	0.96	0.97	51
Persian	0.98	0.96	0.97	55
Ragdoll	0.93	0.95	0.94	55
Russian Blue	0.96	0.93	0.95	56
Siamese	0.97	0.98	0.97	59
Sphynx	0.95	1.00	0.97	57
Accuracy	-	-	**0.97**	**663**
Macro Avg	0.97	0.97	0.97	663
Weighted Avg	0.97	0.97	0.97	663

Table 4.6 shows the classification report of the chosen MobileNetV3 model. The total classification accuracy is about 97%, and both macro-average and weighted-average precision, recall, and F1-score are also close to 0.97. Most of the breeds record high precision and recall scores of more than 0.95, which means an equal performance across the classes. These data support the idea that MobileNetV3 is very effective in multiclass classification of cats, although the network is of lightweight structure and the complexity of calculations is not that high.

In general, the findings indicate that MobileNetV3 can attain the high accuracy of 12-breed cat classification with tuning when the effective learning rate is 0.001. Although the loss rate with very low learning rates can be very high, the chosen configuration in Experiment 1 provides a good trade-off between accuracy and loss, which makes the MobileNetV3 an excellent candidate to implement in constrained resource settings like mobile applications.

4.4.2.2 ResNet50

The results of all ResNet50 experiments on the 12-breed classification are summarized in Table 4.7. Based on the table, it is possible to note that the overall performance of ResNet50 is high, and the test accuracy scores start at around 90% and go up to 96.83%. Just like MobileNetV3, experiments that were optimized with a learning rate of 0.001, always had superior accuracy and test loss values as opposed to the experiments where a learning rate of 0.0001 was used. Specifically, Experiment 7 with the batch size of 32, 20 training epochs, and learning rate of 0.001 recorded the best test accuracy of 96.83%. With a relatively small test loss of 0.1113. This suggests that ResNet50 responds better to longer training in cases where it is being trained on complex interbreed differences.

Table 4.8 Summarizes the performance of all the ResNet50 experimental configurations on the 12 breeds task.

Experiment	Batch Size	Epoch	Learning Rate	Test Accuracy	Test Loss
1	32	10	0.001	96.68%	0.1230
2			0.0001	96.53%	0.1482
3			0.00001	82.20%	0.5801
4		15	0.001	96.23%	0.1259
5			0.0001	95.48%	0.1433
6			0.00001	87.63%	0.4359
7		20	0.001	96.83%	0.1113
8			0.0001	96.23%	0.1230
9			0.00001	87.93%	0.3752
10	64	10	0.001	96.08%	0.1215
11			0.0001	94.27%	0.1970
12			0.00001	77.07%	0.8442
13		15	0.001	95.93%	0.1390
14			0.0001	95.63%	0.1533
15			0.00001	82.50%	0.5997
16		20	0.001	96.08%	0.1181
17			0.0001	96.38%	0.1451
18			0.00001	87.48%	0.4632

As an example, when the learning rate is smaller, which is 0.0001 and 0.00001, the level of performance declines significantly. Indicatively, when experiments are trained using a learning rate of 0.00001 both the rate of accuracy and the value of the test loss are significantly worse and much higher, respectively, showing that the model does not converge quite well with the settings. These findings indicate that, just like MobileNetV3, ResNet50 also needs a large enough learning rate in order to optimize its deep architecture to perform effective multiclass classification computations.

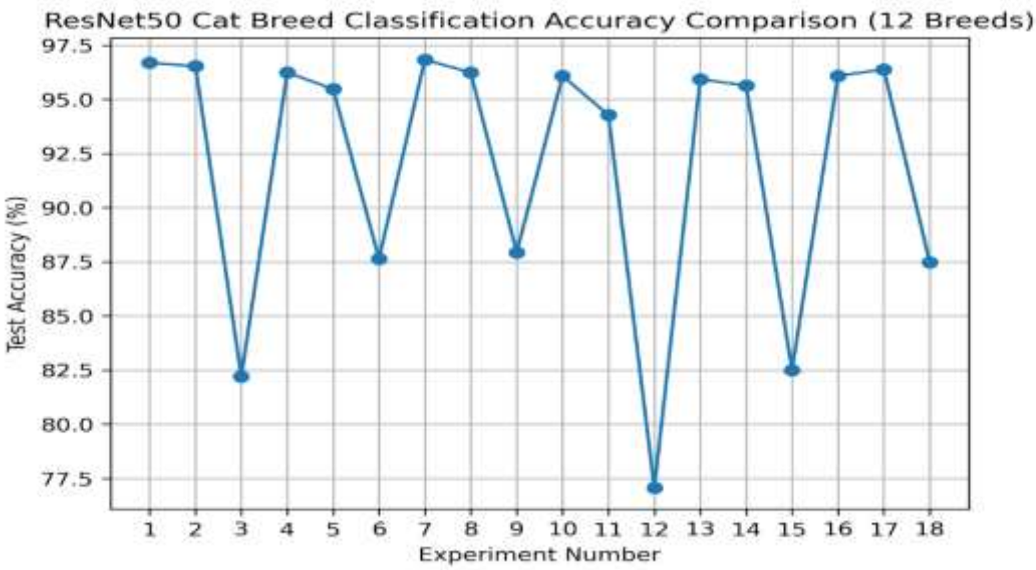

Figure 4.10 Comparison accuracy graph of ResNet50 across all the experiments on the Cat Breed Classification (12 Breeds) task

Figure 4.10 displays the graph of comparison accuracy of ResNet50 under all the experimental conditions in the case of classification of 12 breeds. The graph indicates a distinct concentration of the high value of the accuracy in the experiments conducted with the learning rate of 0.001, and the rates of the accuracy were all above 95%. On the other hand, experiments with smaller learning rates are more variable and of lesser accuracy. ResNet50 has a marginally greater peak accuracy than MobileNetV3, which demonstrates that deeper feature representations have a benefit in the fine-grained classification problem.

According to the findings of Table 4.7 and Figure 4.10, Experiment 7 was identified as the best ResNet50 configuration to be analyzed further. This is the best configuration in terms of test accuracy of 96.83%, and the test loss of 0.1113 is relatively low, warranting high generalization ability. The learning rate of 0.001 gives the model time to acquire more complex inter-class relationships by the use of 20 training epochs, whereas the learning rate maintains steady and effective convergence, and the batch size that was used is 20.

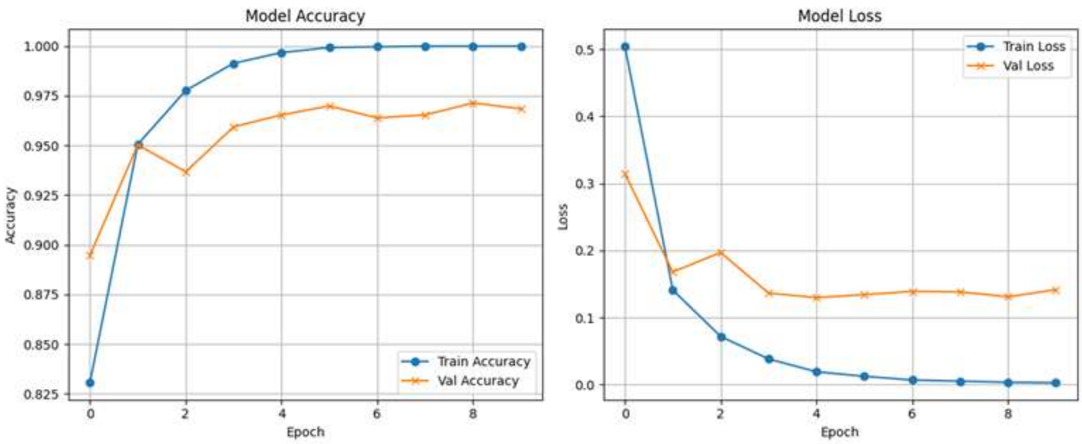

Figure 4.11 Training and validation accuracy of the chosen MobileNetV3 model (Experiment 7)

Figure 4.11 shows the training and validation accuracy curves of the chosen ResNet50 model of Experiment 7. The curves depict a consistent growth of training and validation accuracy in the earlier epochs, according to which there is a gradual levelling of training and validation accuracy at high accuracy markers around 97%. According to the close correspondence between the two curves (training and validation), there is little overfitting exhibited in the model, and this shows that the model works well in the case of unseen data. The convergence of the ResNet50 is a bit slower as compared to MobileNetV3, this is to be expected since the architecture is much deeper.

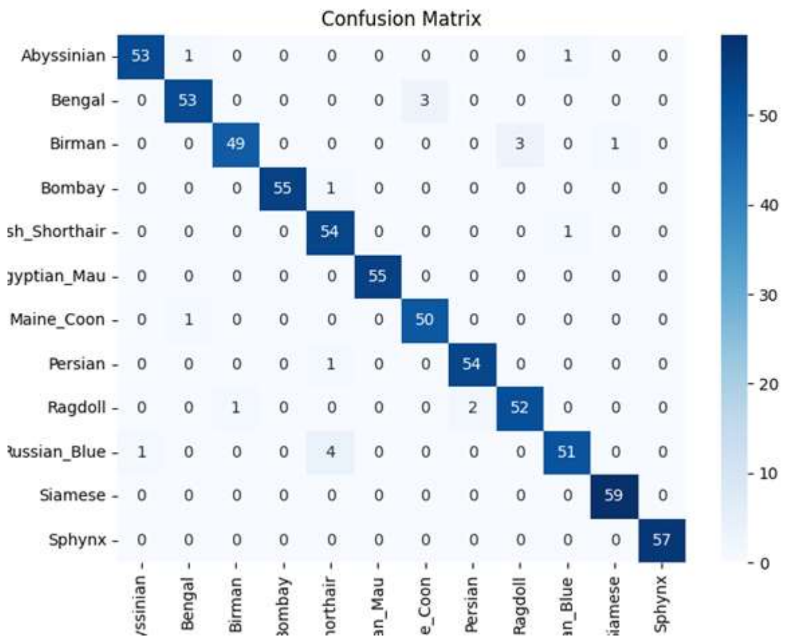

Figure 4.12 Confusion matrix of the chosen ResNet50 model (Experiment 7)

The confusion matrix of the selected model (ResNet50) used to perform the 12-breed classification task is discussed in Figure 4.12. The confusion matrix demonstrates that it has a high concentration of the predictions in the diagonal, and thus the accuracy of the classification was high in most of the breeds. The off-diagonal errors are fewer compared to the MobileNetV3 confusion matrix, and it is possible to suppose that it discriminates better between visually similar breeds. Nevertheless, not all of the misclassifications were eliminated, especially between breeds having very similar appearances, which is the result of the intrinsic challenge of fine-grained classification.

Table 4.9 Classification report of the chosen ResNet50 on the Cat Breed Classification (12 Breeds) task (Experiment 7)

Class	Precision	Recall	F1-Score	Support
Abyssinian	0.98	0.96	0.97	55
Bengal	0.96	0.95	0.95	56
Birman	0.98	0.92	0.95	53
Bombay	1.00	0.98	0.99	56
British Shorthair	0.90	0.98	0.94	55

Egyptian Mau	1.00	1.00	1.00	55
Maine Coon	0.94	0.98	0.96	51
Persian	0.96	0.98	0.97	55
Ragdoll	0.95	0.95	0.95	55
Russian Blue	0.96	0.91	0.94	56
Siamese	0.98	1.00	0.99	59
Sphynx	1.00	1.00	1.00	57
Accuracy	-	-	**0.97**	**663**
Macro Avg	0.97	0.97	0.97	663
Weighted Avg	0.97	0.97	0.97	663

Table 4.8 shows the report of the classification of the chosen ResNet50 model. The classification report gives a closer analysis of the predictive power of the model. Overall accuracy of the model was 97 with weighted-average and macro-average precision. recall and F1-score being 0.97. The individual breed performance too was always high. As an illustration, the Egyptian Mau and Sphynx have achieved perfection in terms of precision and recall, with 1.00 and an F1-score of 1.00 respectively, whereas the Siamese registered 1.00 in terms of perfection, precision, and recall. The other breeds, like Abyssinian, Persian, and Maine Coon, also became F1-scores between 0.96 and 0.97 respectively which means that they were balanced and robust in all their classes.

Overall, it can be concluded that the ResNet50 architecture exhibited a good learning ability and a high classification capacity in the 12-breed cat classification problem. Experiment 7 was the most successful experiment, as its parametric values were the closest to accuracy and loss, which is why it was the best candidate to be deployed. ResNet50 is a more advanced architecture with an improved representational power than lightweight architectures, but its computational cost is more significant. Thus, its results are subsequently contrasted to the MobileNetV3 to identify the most appropriate model to incorporate into a mobile application.

4.4.2.3 Cat Breed Classification (12 Breeds) Result

In this subsection a summary of the most successful experimental settings of the cat breed classification task that involves twelve different cat breeds is presented. The chosen results were done after a comparison of various experimental settings of the two models MobileNetV3 and ResNet50 where the selection relied on the test accuracy and the test loss. The purpose of this comparison will be to compare both CNN architectures in their efficiency in doing multi-class cat breed classification and also to determine which one is the most appropriate in the development of a system. The best result in classification of the twelve-breed task using MobileNetV3 and ResNet50 is demonstrated in Table 4.10.

Table 4.10 Best Performance Results for Cat Breed Classification (12 Breeds)

Model	Experiment	Batch Size	Epoch	Learning Rate	Test Accuracy	Test Loss
MobileNetV3	1	32	10	0.001	96.68%	0.1201
ResNet50	7	32	20	0.001	96.83%	0.1113

The findings show that both CNN architectures were very high in classification accuracy when it comes to separation of twelve cat breeds. Due to its better feature extraction ability in complex multi-class classification problems, ResNet50 achieved slightly higher test accuracy of 96.83% at a smaller test loss than MobileNetV3. Nonetheless, MobileNetV3 also performed competitively with a test accuracy of 96.68% and the complexity of computation is considerably less. Given the planned implementation of a mobile application, MobileNetV3 was chosen as the primary model to be used as the work on the system since it has a lightweight architecture and is more efficient, but in comparison with ResNet50, which served as a control group confirming the classification results.

4.4.3 Similar Breed Classification (Three Breeds)

In this section, a formal analysis of the similar breed classification task with three visually similar cat breeds will be provided using the proposed deep learning models. This is the most difficult classification case in the experimental paradigm because the sampled breeds have very similar visual characteristics, including fur colour, texture, facial structure, and body shape. In contrast to the 12-breed classification task, where the inter-class distances are greater, the similar breed classification task asks the model to learn extremely fine-grained features in order to accurately differentiate between breeds. Subsequently, this task exerts more strain on the extraction ability and overall character of the deep learning architectures. The analysis of the performance of MobileNetV3 is conducted in this section, and the performance of ResNet50 is evaluated in the following subsection.

4.4.3.1 MobileNetV3

Table 4.9 is the summary of the performance of 18 experimental configurations using MobileNetV3 in the Similar Breed Classification task with three breeds of cats, which are Birman, Siamese, and Ragdoll. The experiments were formulated by creating different batch sizes, training epochs, and learning rates to assess how the model responds to the change of various hyperparameters. Based on the table, it can be seen that the accuracy values of the tests at different configurations vary very much, with the lowest accuracy value, 49.70%, being very low at the poorly converged experiments and the maximum value being 95.21% as a representation of the difficulty of fine-grained classification.

Table 4.11 Summarizes the performance of all the MobileNetV3 experimental configurations on the similar breeds (Three breeds) task

Experiment	Batch Size	Epoch	Learning Rate	Test Accuracy	Test Loss
1	32	10	0.001	92.22%	0.2016
2			0.0001	85.63%	0.3538

3			0.00001	58.08%	0.9050
4		15	0.001	94.01%	0.2019
5			0.0001	90.42%	0.2865
6			0.00001	73.65%	0.6386
7		20	0.001	95.21%	0.1487
8			0.0001	90.42%	0.2789
9			0.00001	74.25%	0.6361
10	64	10	0.001	91.62%	0.2075
11			0.0001	86.23%	0.3659
12			0.00001	49.70%	1.0437
13		15	0.001	94.61%	0.1913
14			0.0001	87.43%	0.3044
15			0.00001	63.47%	0.8538
16		20	0.001	93.41%	0.1902
17			0.0001	90.42%	0.2836
18			0.00001	74.25%	0.6851

A further look into Table 4.9 shows that all experiments that are trained with the learning rate 0.001 will yield the best accuracy and the lowest test loss values. Specifically, Experiment 7, which used a batch size of 32, 20 training epochs, and a learning rate of 0.001 presented the same results with the highest test accuracy of 95.21% and the lowest test loss of 0.1487. This finding suggests that MobileNetV3 has the advantage of longer training to capture fine differences in visual differences between similar breeds. The same learning rate with varying batch sizes and epoch settings also shows good performance, as other experiments depict high accuracy in the test as well, and the values of 93% or more are obtained.

However, when reduced learning rates of 0.0001 and 0.00001 are used in the experiment, the performance decreases significantly. Indicatively, Experiment 3 with a learning rate of 0.00001 had a test accuracy of only 58.08% with a high-test loss of 0.9050, meaning that the convergence was bad. Same goes for

Experiment 12, with a learning rate that was set to 0.00001 had the lowest test accuracy of 49.70% and the greatest test loss of 1.0437. Based on these findings, it is exactly obvious that overly low learning rates are devastatingly detrimental to the optimization process, where the model is unable to acquire discriminative properties in the fine-grained classification.

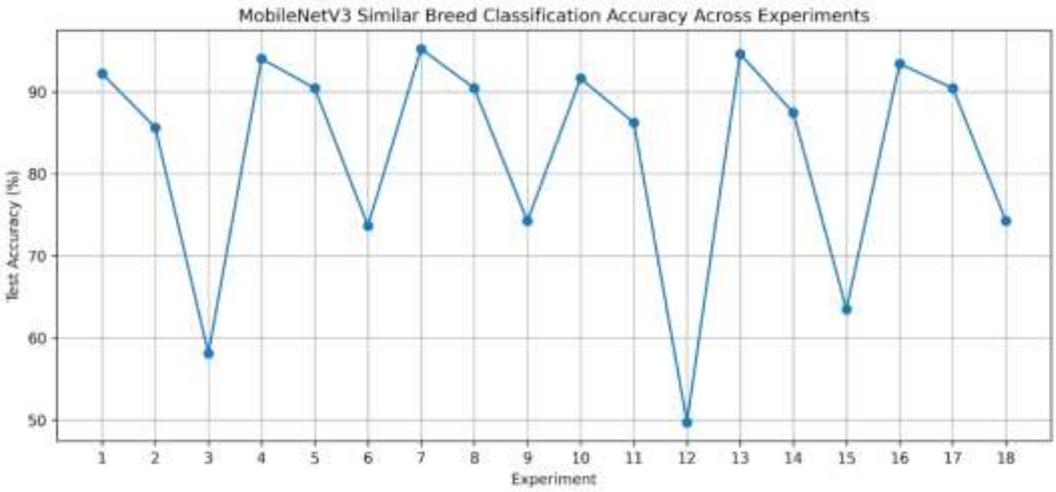

Figure 4.13 Comparison accuracy graph of MobileNetV3 across all the experiments on the Similar Breed (Three Breeds) task

The graph of comparison accuracy of the MobileNetV3 model in each of the experimental scenarios in the Similar Breed Classification task is shown in Fig. 4.13. The fact that the model performance is highly dependent on the learning rate is graphically emphasized. Experiments with a learning rate equal to 0.001 congregate in the top position of the graph and attain an accuracy value over 93, whereas experiments with a learning rate equal to 0.00001 exhibit a sharp decrease in accuracy, and many of the results conclude under 75. This is a completely upward trend in this trend graph in line with the numerical data recorded in Table 4.9.

According to the findings presented in Table 4.9 and Figure 4.13, Experiment 7 was chosen as the best MobileNetV3 setting to analyze the results further. This experiment is the most successful balance between high accuracy, 95.21%, and low-test loss, 0.1487, which reflects high performance of the generalization in the most critical situation of the classification.

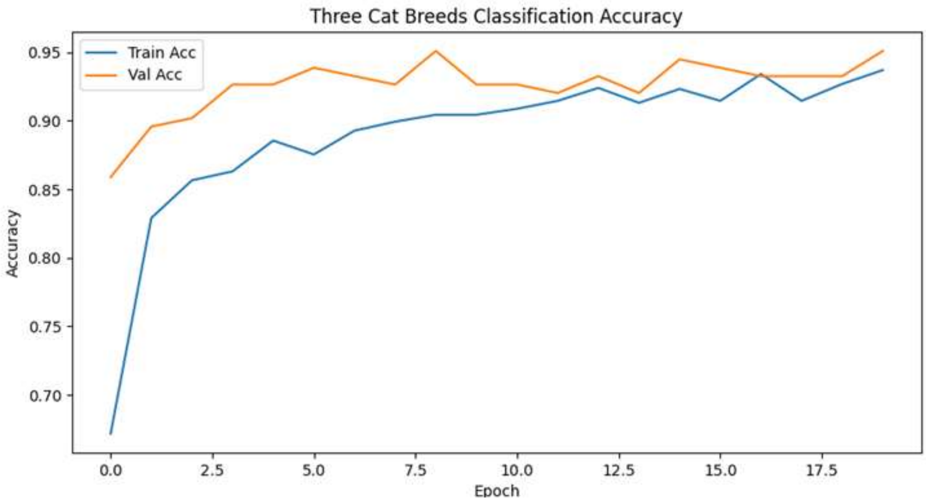

Figure 4.14 Training and validation accuracy of the chosen MobileNetV3 model for the Similar Breed (Three Breeds) task (Experiment 7)

The results of the experimental training and validation accuracy curves in Figure 4.14 illustrate the chosen MobileNetV3 model in Experiment 7. The curves reveal that the training and validation accuracy both point towards a gradual increase in the first epochs, then level off at the high accuracy levels. Throughout the training process, the training and validation curves are in close agreement, meaning that there is slight overfitting. According to the fact that the curves are smoothly converging, the model can theoretically learn small grains without memorizing training samples.

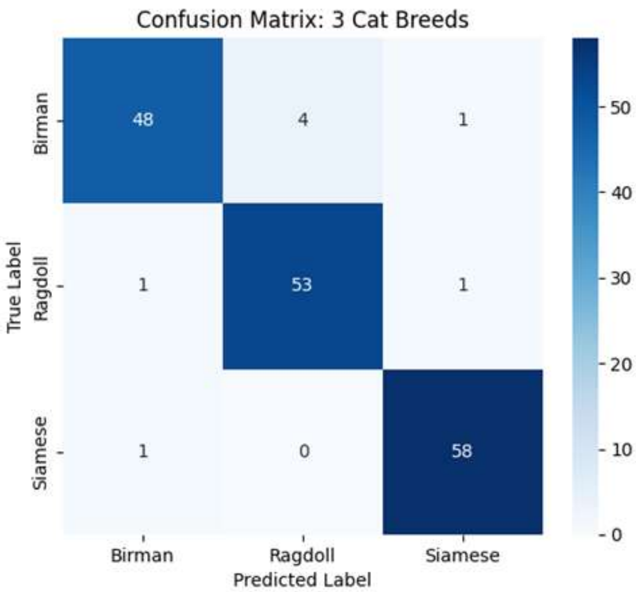

Figure 4.15 Confusion matrix of the chosen MobileNetV3 model for the Similar Breed (Three Breeds) task (Experiment 7)

Figure 4.15 shows the confusion matrix of the chosen MobileNetV3 model of the task of Similar Breed Classification. Through the confusion matrix, it is depicted that the majority of the predictions are readily classified, as reflected in the high diagonal values. However, certain misclassifications occur between visually similar breeds, especially between breeds with similar fur patterns and facial structures. It is understandable that these mistakes occur because the visual similarity between the chosen breeds is extremely high and indicates the inadvisability of this classification task.

Table 4.12 Classification report of the chosen MobileNetV3 model for the Similar Breed (Three Breeds) task (Experiment 7)

Class	Precision	Recall	F1-Score	Support
Abyssinian	0.98	0.96	0.97	55
Bengal	0.96	0.95	0.95	56
Birman	0.98	0.92	0.95	53
Accuracy	-	-	**0.95**	**167**
Macro Avg	0.95	0.95	0.95	167
Weighted Avg	0.95	0.95	0.95	167

The classification report of the chosen MobileNetV3 model is provided in Table 4.10. The general classification accuracy has been about 0.95 and the macro-average and weighted-average accuracy as well as recall and F1-score measures are also of the same value at around 0.95 indicating a balanced performance among the three breeds. Precision and recall values across individual classes also have a high, constant level, which ensures that the model does not over favour one class. These findings indicate that MobileNetV3 can be very effective in fine-grained classification tuned to work well under different settings.

Overall, the findings of the experiments verify that MobileNetV3 can be effectively used to achieve Similar Breed Classification with the best accuracy associated with specific configurations of this type of learning with an adequate number of training epochs and a suitable learning rate. Although the performance under suboptimal learning rate configurations drops drastically, the chosen setup in Experiment 7 shows that MobileNetV3 could effectively learn subtle inter-breed variations, which makes it an acceptable choice of the model in the domain of fine-grained cat breed classification in the limited resource settings.

4.4.3.2 ResNet50

Table 4.11 summarizes the overall performance of all ResNet50 experiments with the Similar Breed Classification task. The findings indicate that there is a significant difference in values of test accuracy indicators depending on the experiment's hyperparameter settings, with minimal accuracy spots of the worst convergent experiments to the highest accuracy of 92.81%. The experiments that were conducted with a very low learning rate of 0.00001 actually performed the worst, as the accuracy coefficient is only 71.26% and the test loss coefficient reaches a maximum of 0.7467, which implies inefficient convergence. Conversely, with a learning rate of 0.001, experimental results are all more precise and have lower test loss values, which indicate that it is essential to determine the learning rate in a fine-grained classification.

Table 4.13 Summarizes the performance of all the ResNet50 experimental configurations on the similar breeds (Three Breeds) task

Experiment	Batch Size	Epoch	Learning Rate	Test Accuracy	Test Loss
1	32	10	0.001	93.41%	0.2052
2			0.0001	87.43%	0.3361
3			0.00001	75.45%	0.6933
4		15	0.001	92.81%	0.2055
5			0.0001	89.22%	0.2756
6			0.00001	71.86%	0.6757
7		20	0.001	92.81%	0.1942
8			0.0001	91.02%	0.2609
9			0.00001	86.83%	0.4745

10	64	10	0.001	92.81%	0.2432
11			0.0001	87.43%	0.3307
12			0.00001	71.26%	0.7467
13		15	0.001	92.81%	0.1943
14			0.0001	89.82%	0.2993
15			0.00001	78.44%	0.6715
16		20	0.001	92.81%	0.2113
17			0.0001	90.42%	0.2742
18			0.00001	78.44%	0.6464

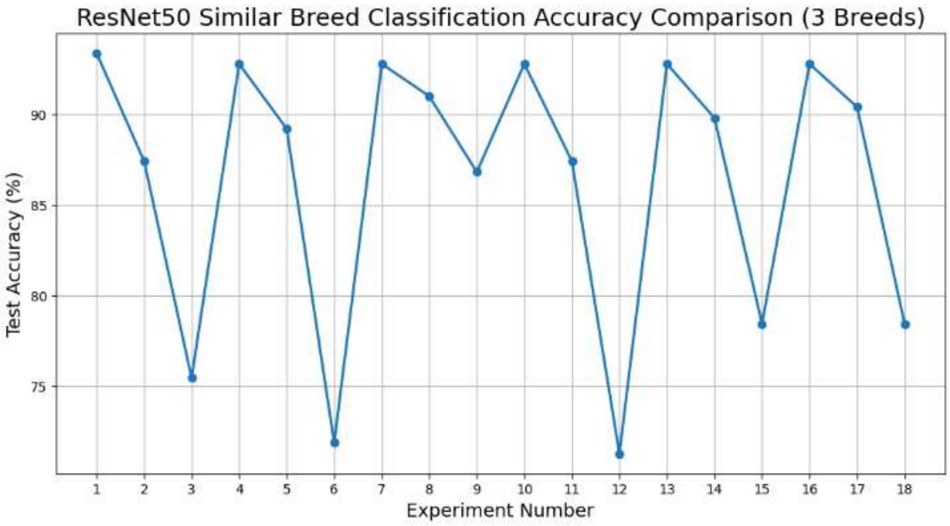

Figure 4.16 Comparison accuracy graph of ResNet50 across all the experiments on the Similar Breed (Three Breeds) task

The comparison accuracy graph in ResNet50 under each of the experimental settings in the task of Similar Breed Classification is presented in Figure 4.16. As illustrated by the graph, experiments with the learning rate of 0.001 are concentrated on the higher part of the accuracy value distribution, almost all of which are above 90%, whereas experiments with lower learning rates exhibit severe performance deteriorations. The highest accuracy of ResNet50 is lower in comparison with MobileNetV3, which means that deeper architectures are not guaranteed to be the most effective in solving fine-grained classification problems, especially when the training data is insufficient.

Based on Table 4.11 and Figure 4.16, Experiment 7 had the best test accuracy at 92.81 and the lowest test loss at 0.1942. It is the only configuration that had a batch size of 32, 20 training epochs, and a learning rate of 0.001 which is the lowest learning rate. Other experiments with identical learning rates obtained similar accuracy values but Experiment 7 was chosen as the best setup because it had the best balance of accuracy and loss. These findings indicate that when ResNet50 is trained using a sufficient number of training epochs to learn the subtle inter-breed differences, the algorithm is more sensitive to hyperparameter choices than MobileNetV3 is.

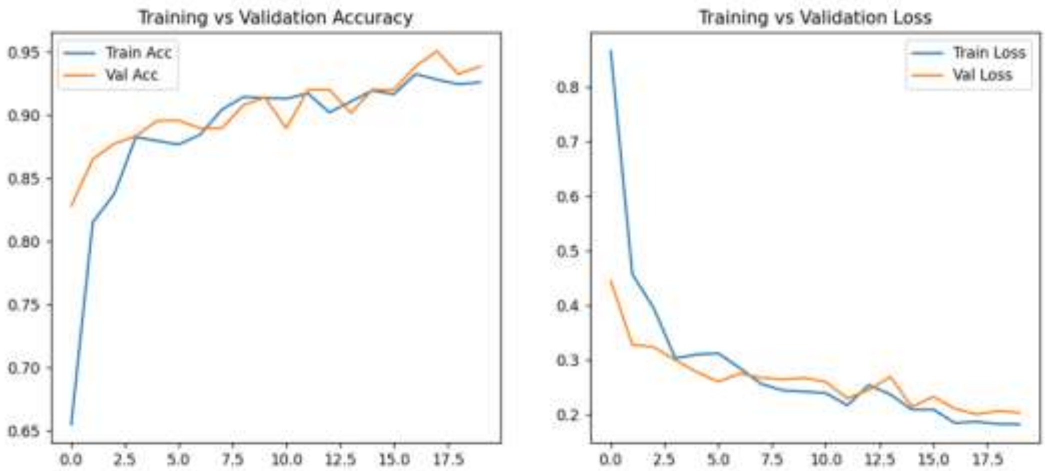

Figure 4.17 Training and validation accuracy of the chosen ResNet50 model for the Similar Breed (Three Breeds) task (Experiment 7)

The training and validation accuracy curves of the chosen ResNet50 model of Experiment 7 are presented in Figure 4.17. The curves demonstrate that the training and validation accuracy increase gradually and finally stabilize to values that are a bit higher, around 92%. Given that the curves are not much farther apart all through the training, slight changes in the accuracy of validation are observable, which hint at a slight instability during the optimization process. This indicates that the fact that whereas ResNet50 can be trained on discriminative features, the approach can be more vulnerable to overfitting or optimization problems in fine-grained tasks than MobileNetV3.

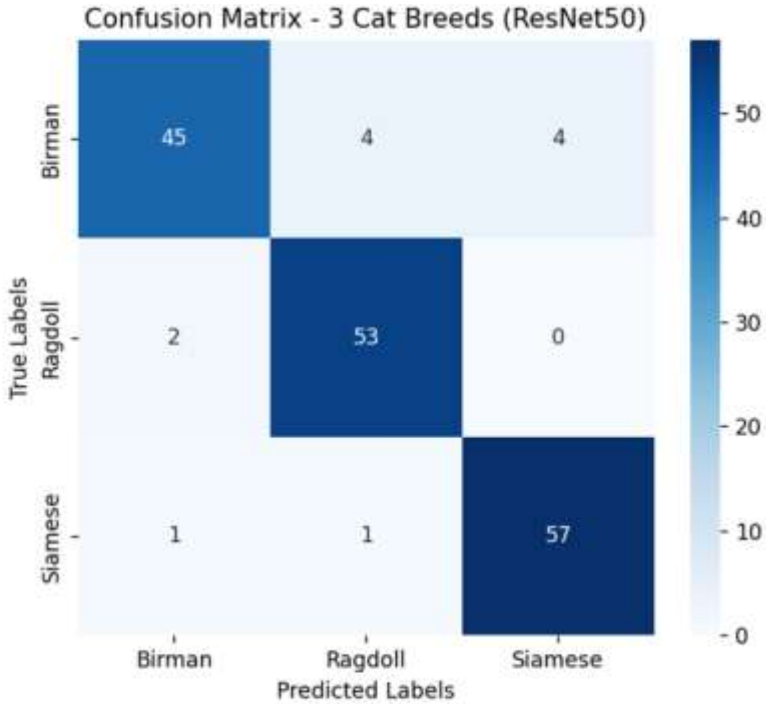

Figure 4.18 Confusion matrix of the chosen ResNet50 model for the Similar Breed (3 Breeds) task (Experiment 7)

The confusion matrix of the chosen ResNet50 model on the task of Similar Breed Classification is shown in figure 4.18. The confusion matrix shows that predictions are concentrated around the diagonal, which means that most samples are correctly classified. Nevertheless, there are more off-diagonal errors between the breeds than in the MobileNetV3 confusion matrix, especially between breeds that have very similar visual patterns. Such false identifications point out how hard it is to discriminate among similar breeds and indicate that inter-class confusion in this task is not addressed entirely using the deeper architecture of ResNet50.

Table 4.14 Classification report of the chosen ResNet50 model for the Similar Breed (Three Breeds) task (Experiment 7)

Class	Precision	Recall	F1-Score	Support
Abyssinian	0.94	0.85	0.89	53
Bengal	0.91	0.96	0.94	55
Birman	0.93	0.97	0.95	59
Accuracy	-	-	**0.93**	**167**
Macro Avg	0.93	0.93	0.93	167
Weighted Avg	0.93	0.93	0.93	167

Table 4.12 shows the classification report of the chosen ResNet50 model. The classification accuracy is about 93 on the whole, and the macro-average and weighted-average Precision, Recall, and F1-score values are also about 0.93, which is quite a decent set of balanced results across the three breeds. Within each breed, the values of precision and recall are still greater than 0.90, although one breed has a lower recall, which represents the misclassification patterns that are observed in the confusion matrix. These findings validate that even though ResNet50 is suitable when it comes to Similar Breed Classification, its performance is under that of MobileNetV3 on the same task.

In general, the experimental findings suggest that ResNet50 has moderate performance in the task of Similar Breed Classification, and the best setup of the algorithm has a test accuracy of 92.81%. In this fine-grained classification case, however, ResNet50 is not superior to MobileNetV3 in terms of architecture and computational power. This implies that more efficient feature extraction using lightweight models like MobileNetV3 might be more applicable when it comes to feature differences that are subtle to work with and when computational efficiency and stability are critical factors to consider.

4.4.3.3 Similar Breed Classification (Three Breeds) Result

In this subsection, three experimental configurations yielding the best results in the similar breed classification task are summarized, that is, distinguishing between three visually similar cat breeds, which are Birman, Ragdoll, and Siamese. This is more difficult to do because the selected breeds have very close similarity in the fur patterns, facial structure and general appearance. The most successful cases of MobileNetV3 and ResNet50 were derived using test accuracy and test loss on the experimental measures. Table 4.15 shows the performance results of the similar breed classification task with both CNN architectures at the best performance.

Table 4.15 Best Performance Results for Similar Breed Classification (Three Breeds)

Model	Experiment	Batch Size	Epoch	Learning Rate	Test Accuracy	Test Loss
MobileNetV3	7	32	20	0.001	95.21%	0.1487
ResNet50	7	32	20	0.001	92.81%	0.1942

The findings show that MobileNetV3 performed better than ResNet50 in a similar breed classification task having a higher test accuracy of 95.21% and lower test loss than that of ResNet50. It implies that MobileNetV3 is useful to elicit more discriminative features that would distinguish between visually similar cat breeds. However, the performance of ResNet50, in this fine-grained classification case, was relatively lower, which might be explained by more complexity of the model and its susceptibility to small feature overlaps. Under these results, MobileNetV3 has been chosen as the model of choice in similar breeding in the proposed mobile application. Its high performance also contributes to its applicability to the mobile-based cat breed identification application as its lightweight architecture will be helpful in its integration.

4.5 Model Selection and Discussion

This section discusses the choice of the final model to be utilized in this project in light of the experimental findings provided in the previous sections. The model selection process does not take as the sole criteria the classification accuracy but also the model stability, generalization performance, and computational efficiency. As the proposed system will be used in a real-life application in a mobile-based system, it is necessary to strike a balance between high classification performance and light model design. Thus, the choice will rely on the performance of MobileNetV3 and ResNet50 doing all three tasks of classification.

Table 4.16 Summary of Selected Classification Results for Model Comparison

Classification Task	Model	Batch Size	Epoch	Learning Rate	Test Accuracy	Test Loss
Cat Vs Non-Cat	MobileNetV3	64	10	0.001	99.88%	0.0099
	ResNet50	64	10	0.001	99.88%	0.0059
Cat Breed (12 Breeds)	MobileNetV3	32	10	0.001	96.68%	0.1201
	ResNet50	32	20	0.001	96.83%	0.1113
Similar Breed (3 Breeds)	MobileNetV3	32	20	0.001	95.21%	0.1487
	ResNet50	32	20	0.001	92.81%	0.1942

Table 4.16 provides a summary of the best-performing experimental results of MobileNetV3 and ResNet50 on all classification tasks within the framework of this research, which are Cat vs Non-Cat classification, Cat Breed Classification (12 different breeds), and Similar Breed Classification (three visually similar breeds). The table brings out the training settings and the performance measure such as test accuracy and test loss that was employed in comparing and contrasting the performance of the two CNN architectures.

According to the findings of the Cat vs Non-Cat classification, MobileNetV3 and ResNet50 were both outstanding in terms of their performance, hitting the accuracy value of more than 99%. ResNet50 was nearly perfect with only slightly reduced values of test loss, which influenced high confidence in its prediction. Nonetheless, the MobileNetV3 was also able to perform similarly with only slight performance disparity. Since it appears that the two models are nearly equal in terms of their performance in this binary task, the complexity of their models comes into more significant focus at this point.

In the Cat Breed Classification task with 12 breeds, ResNet50 was the best-performing model with the highest accuracy of 96.83%, and MobileNetV3 had an almost similar accuracy, equal to 96.68%. ResNet50 is slightly higher than MobileNetV3, but the difference between them is rather low, which is only 0.15%. The two models exhibited stable learning behavior with the right learning rate of training, and errors, which were misclassifications, only happened among breeds that were visually similar. Those outcomes reveal that the two architectures can learn fine-grained breed features well.

MobileNetV3 performed better in the Comparable Breed Classification task with 3 breeds, as it had a higher accuracy of 95.21% than 92.81% with ResNet50. This indicates that MobileNetV3 is stronger when it comes to fine-grained classification with limited data requirements. Moreover, the MobileNetV3 was more stable in training with a lower number of fluctuations in the accuracy of validation, thus it was able to generalize the situation better.

Computationally, MobileNetV3 is an efficient mobile architecture that is specifically aimed at efficiency and low-resource settings. It also has a smaller set of parameters and lower computational complexity than ResNet50 and can be better used to do inferences in real time on mobile devices. However, ResNet50, though with powerful feature extraction properties, is characterized by greater memory and processing demands, which can be a constraint to its practical applications in the mobile application.

Considering the performance of all tasks in general, MobileNetV3 is chosen as the ultimate model in this project. Compared to all classification tasks, the model exhibits competitive accuracy and outperforms on finer-grained classification and has much less computational complexity. ResNet50 is held in place as a baseline model to justify performance and prove the effectiveness of the proposed approach.

Overall, the model selection process proves that MobileNetV3 is the model with the most appropriate balance of accuracy, robustness, and efficiency. This choice is in line with the project vision of developing an efficient and feasible cat breed classification system that can be used in deploying a mobile application.

4.6 Mobile Application Implementation Result

This section includes the implementation outcomes of the recommended cat classification system as a mobile application. This implementation is aimed at proving that the trained deep learning model can be a means of a successful implementation in the mobile environment and can be used for the purpose of conducting real-time inference by providing the user with the images. This topic is devoted to system behaviour and output, as opposed to considerations on the design of a user interface.

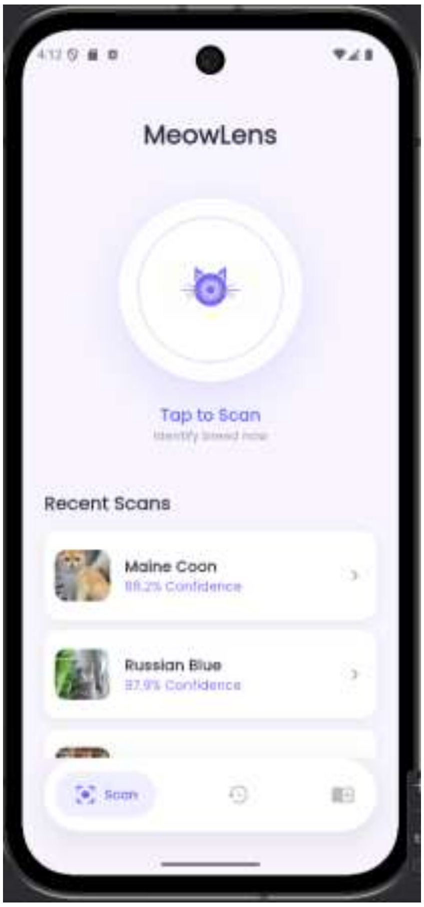

Figure 4.19 Homepage

Figure 4.19 shows the homepage of the mobile application, and it is the starting point of the application that users visit. The home page offers a straightforward and easy-to-use interface through which the user can proceed to the image acquisition process. The design is user-focused and will not introduce any unnecessary complexity since users can easily access the basic functionality of the application.

Figure 4.20 Upload Page

Figure 4.20 displays the image upload and capture page with users connecting with the device either through the device gallery to upload an image or through the camera to take a new image. After selecting an image, the application then automatically undertakes preprocessing functions, such as resizing and normalization, to suit the input specifications of the deployed MobileNetV3 model. This page is a very important interaction step, where real-world images are ready to be model inferred.

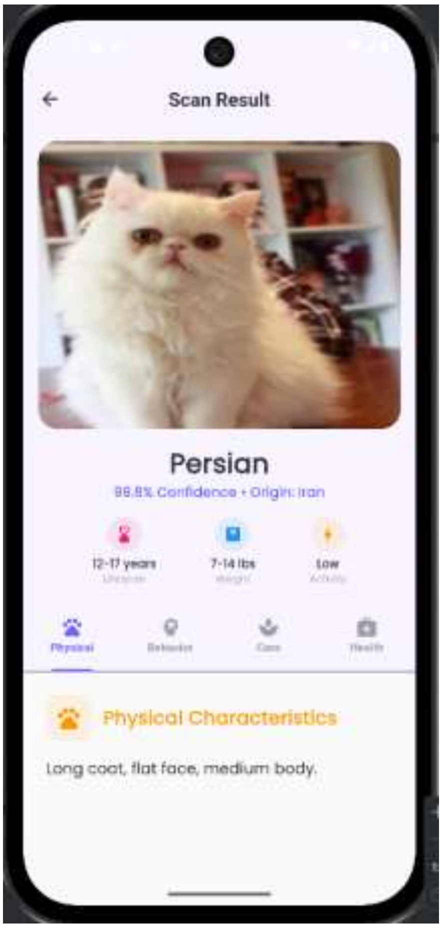

Figure 4.21 Result Page

The prediction result page of the application is seen in figure 4.21. Once inference has been done, the application will show the predicted cat breed and the corresponding confidence score. Besides the classification outcome, short informational texts associated with the identified predicted breed are also presented, such as about the overall characteristics of the breed and its aspects. This extra information provides a deeper user experience as it provides a useful context as opposed to a prediction label. The performance of the application in offline testing gives results in the prediction that are consistent with the results given by the application, especially when the cat is evident in the image.

In the real-life testing, the application showed good reliability in controlled conditions, which included a single cat image and a fairly simple background. Nonetheless, a few constraints were noticed. Whenever the cat was partially covered, as in the case where the person was holding the cat, or in case the background had too many complicated details, such as the grass or the other objects near the cat, the model sometimes did not detect or hit the correct classification of the cat. These findings are in agreement with the experimental results presented above and show difficulties associated with the interferences of the background and occlusion in real field usage.

Overall, the implementation of the mobile application indicates that the proposed system can work and make its respective contribution to the practice of cat breed classification. The practical use of deep learning classification systems on the mobile application, as confirmed by the successful integration of the trained model, just goes to attest to the viability of implementing systems using deep learning classifications in the real-world case. The findings can also give useful information on how things can be improved in the future by getting greater strength under complex backgrounds coupled with occlusions.

4.7 Discussion

The section addresses the interpretation and discussion of the experimental results and implementation findings discussed in the earlier sections. The discussion revolves around the performance of these models with various classification tasks, how training configurations affect performance, the comparison of MobileNetV3 and ResNet50, and the behavioural characteristics noted in the testing of all the models in the real world in the context of mobile applications.

Across all the experiments, MobileNetV3 and ResNet50 displayed high accuracy in classification, especially when applied to simpler problems like Cat vs Non-Cat classification, where all the accuracy scored above 99. This shows that both architectures are very good at learning high-level visual characteristics that help in distinguishing the cats and non-cat objects. The low test loss values maintained in this task also indicate that test loss was not overfitting and thus indicate that the models were well-generalized. The differences in performance between the models could be more pronounced, especially in the 12-breed task and similar-breed tasks, as the level of classification becomes more complex.

In the task of Cat Breed Classification, where 12 breeds were used, both models were highly accurate, and ResNet50 was slightly better in terms of the situation when the highest accuracy was reached. ResNet50 has a deeper architecture and residual links that enable it to represent more detailed and complex feature representations, which is advantageous when it comes to discriminating between visually similar varieties of cats. Nevertheless, the difference between the performance of the two models was quite low. MobileNetV3 was found to be equally accurate with a lower test loss stability in numerous experiments, which suggests that the lightweight architecture is effective at learning useful features despite its massiveness.

MobileNetV3 was able to withstand stronger robustness than ResNet50 in the task of Similar Breed Classification, which was conducted with three visually similar breeds. MobileNetV3 had better higher accuracy and more consistent validation results, indicating improvements to generalization when there is limited data. It is probable that this behavior can be explained by the architecture of MobileNetV3, where the focus on a small number of features and regularization optimization allows the model to operate without overfitting to some fine-grained distinctions.

Another significant factor that the experimental outcomes show is the impact of training hyperparameters on model performance. During all tasks, models that were trained with a learning rate of 0.001 always performed better than those that were trained with smaller learning rates, which are 0.0001 and 0.00001. Lower learning rates tended to underfit the data with smaller values of accuracy and higher values of loss. In most cases, adding more epochs tended positively to enhance performance to a certain level, beyond which marginal performance advances were observed. The above results highlight the need to hyperparameter-tune to ensure high performance.

These observations are further supported by the confusion matrices and classification reports. Most misclassifications involved similar-looking breeds, especially those of similar fur patterns, facial structure, and other related features. Precision and recall values were shown to be high in most of the dominant and distinctive breeds, whereas slightly lesser values were obtained in those breeds that had common visual features. This implies that the models are efficient, but fine-grained classification is still difficult unless more information diversity or feature enhancement is used.

Real-world testing was conducted using the mobile application, which disclosed testing performance patterns in line with the offline analysis. The model deployed worked well in situations whereby the cat was easily visible, and the background was not so complicated. However, there were also difficulties in the cases of those images that comprised partial covering, such as the image of a person with a cat, as well as a cluttered background with complicated elements of the environment. The model sometimes could not appropriately recognize or categorize the cat in these situations. These constraints indicate that the training data might not be as close to the real world as possible and also inform the background diversity and invitation of handling, which are significant in the real implementation.

Overall, the discussion supports the fact that ResNet50 can have good feature extraction capabilities and a little higher accuracy rate on complicated tasks, but MobileNetV3 is a more comprehensive answer to the triage of accuracy, reliability, and computer speed. The fact that MobileNetV3 is deployed successfully in a mobile application is another confirmation of MobileNetV3 suitability in real-world applications. The results also lead to potential future development such as increasing the size of the datasets, augmentation of data, and training approaches that are robust to backgrounds.

4.8 Summary

The chapter was an overall analysis of the proposed cat classification system that included experimental design, performance analysis, results analysis, and implementation in real life. The aim of the study at the experimental stage was presented in Section 4.1, which described the range of model analysis in the context of various classification tasks. Section 4.2 outlines the experimental design, which involves preparation of datasets, model structures, model training, and hyperparameter decisions that are made to compare results fairly and consistently. Section 4.3 described the models of performance evaluation used to evaluate the effectiveness of a particular model, such as accuracy, loss, precision, recall, and F1-score.

The experimental results of all the classification tasks, such as Cat vs Non-Cat classification, Cat Breed Classification of 12 breeds, and Similar Breed Classification of 3 visually similar breeds, were reported in Section 4.4. The obtained results have shown that both MobileNetV3 and ResNet50 demonstrated a high level of performance in all tasks, and the level of accuracy was higher than expected. The sensitivity of the training parameters, including learning rate, batch size, and epochs, was also well encountered, where a learning rate of 0.001 always generated better performance.

In section 4.5, the model selection process was explained, and the final choice of the model was MobileNetV3, accepted because of its balanced performance, stability, and computing capabilities. Even though the peak accuracy was a little higher with ResNet50 in some tasks, MobileNetV3 scored the same and is suitable to be deployed in a mobile environment. The practicability of the proposed approach was confirmed by Section 4.6, involving the implementation of a mobile application that has shown the ability to successfully integrate the trained model and demonstrate the stable behaviour of real-world inference.

Overall, Section 4.7 provided a detailed discussion of the experimental results, including model behaviour, misclassification behaviour, and constraints that were witnessed during the real-world testing. In general, the effectiveness, strength, and practicability of the suggested deep learning-based cat classification system are validated in this chapter and serve as a strong background towards the conclusions and future work that will be delivered in the next chapter.

CHAPTER 5
CONCLUSION AND FUTURE WORK

5.1 Summary of Findings

This research explored how Convolutional Neural Network (CNN) models are effective in classifying images of cats based on the breeds and evaluated their suitability to deploy in a real-world setting using a mobile application. Two CNNs, which are the MobileNetV3 and the ResNet50, were trained and tested on various classification tasks of varying complexity, such as Cat vs Non-Cat classification, Cat Breed Classification, with 12 cat breeds, and Similar Breed Classification, with three visually similar breeds.

The experimental results revealed that both CNN models could be used to obtain a high classification accuracy under proper hyperparameter settings. The level of accuracy of the two models in the Cat vs. Non-Cat classification experiment was above the level of 99% of the models to differentiate cats from those that were not cats. In the Cat Breed Classification task that had twelve categorical breeds, the maximum levels of accuracy attained by MobileNetV3 and ResNet50 were 96.68% and 96.83%, respectively, thus supporting the effectiveness of the convolutional neural network-based methods in identifying the breed of the cat. In terms of the similar-breed classification problem, MobileNetV3 showed better robustness, with high accuracy when compared to ResNet50, especially when distinguishing similar breeds.

The results also expressed the fact that the selection of learning rate was a very important factor in the model performance. In all experiments, the models that were trained with a learning rate of 0.001 tended to perform better than the models that were trained with a smaller learning rate, which could frequently lead to underfitting and an increase in the value of losses. Classification reports and Confusion matrix analysis predicted that the majority of the misclassifications were between breeds that were similar visually, including fur patterns and facial structures, leading to the natural challenges of fine-grained breed classification.

Along with the evaluations that were carried out offline, the chosen CNN model also was integrated into a mobile app and experimented with in a real-world setting. The application showed stable functioning when the cats were distinctly seen in pictures, and the background was not very complex. Nonetheless, it was noted that difficulties occurred when partial occlusion or crowded backgrounds were involved and sometimes resulted in incorrect classification or failure detection. The observations highlight the discrepancy between experimental settings of controlled conditions of deployment and the actual real-world deployment settings.

In general, the results obtained confirm that cat breeds can be well identified and classified using CNN-based models based on the images with high accuracy. MobileNetV3 was the top model out of the assessed models and was selected as the most suitable model to implement, as the model offered a moderate balance between classification performance, robustness, and computational efficiency. The findings provide the basis for determining how effectively the project objectives are met and where improvement is still necessary and evident in the following sections.

5.2 Achievement of Objectives

In this section, the project objectives as mentioned in Chapter 1 are assessed by comparing them to the results and implementation outcomes obtained in this study.

The first objective of this project was to develop CNN models that can recognize and classify different cat breeds based on the images. This goal was completed with the help of applications of two convolutional neural network architectures, which are MobileNetV3 and ResNet50. Both models were trained and tested on labelled cat image datasets on different classification tasks, which include

Classification of Cat vs Non-Cat, Classification of Breeds (12 breeds), and Similar Breed Classification (three breeds). The experimental results have shown that both models were capable of both learning discriminative visual features when used with cat images with great accuracy and then classifying the breeds with high accuracy, indicating that the effective CNN-based classifier was developed successfully.

The second objective was to evaluate the accuracy and reliability of CNN models by utilizing suitable metrics on a labelled dataset of cat breed images. This goal was accomplished with the help of standard evaluation measures such as accuracy, loss, precision, recall, and F1-score. The metrics were employed to measure the performance of the models of varying experimental setups. Misclassification patterns and class-based performance were analyzed using additional tools of evaluation, such as confusion matrices and classification reports. The overall analysis showed the CNN models trained were trustworthy and predictable, especially when they were trained under optimal hyperparameter settings.

The third goal was to develop a mobile program integrated with a CNN model to identify and recognize cat breeds from captured or uploaded photos. This was achieved successfully through implementation of the chosen MobileNetV3 model in a mobile app. The application allows people to browse through the gallery pictures or take pictures with the device camera, and then the model performs a real-time inference process to identify the cat breed. The application of the mobile application shows that this goal is realized by the successful operation of the mobile application and the stability of its prediction in real-world tests.

In general, all project objectives were successfully met. The findings indicate that the development, evaluation, and deployment of a CNN-based system of cat breed classification have been successfully fulfilled against the intended research objectives.

5.3 Contributions of the Study

This study discusses the image-based animal recognition domain by showing how convolutional neural network (CNN) models can be utilized successfully in cat breed identification and real-world deployment. This study can be divided into contributions under the technical, practical, and academic aspects.

Technically, this study presents an organized analysis of two CNN models of estimated similar complexity, namely MobileNetV3 and ResNet50, to carry out cat breed classification tasks of different complexity. By experimenting widely and testing the performance of each architecture, the research indicates the advantages and limitations of each architecture, especially in multiclass and fine-grained classification case scenarios. The results provide information on the capabilities of lightweight architectures to compete with the deeper architectures as long as they are trained and optimized appropriately.

Practically, this project provides a usable mobile application that implements a trained CNN model to classify cats on a real-time basis. The observation of the successful implementation of the MobileNetV3 model serves as proof of the practicality of using deep learning methods in mobile settings with fewer resources of computation. Other features of the application also include the prediction of the breed, which is accompanied by accuracy scores and a short description of the breed to interact better with the user and increase usability.

Academically, this study provides a systematic analytical experimental structure for testing CNN-based image classification systems. Various evaluation measures, confusion matrices, and real-life testing are highly comprehensive evaluation techniques that could be utilized in further research. Furthermore, model behaviour analysis in controlled and real-life conditions can provide important information about the issues of background complexity and partial occlusion.

On the whole, the findings of this study contribute to the development of effective and implementable deep learning-based image classification systems and form the basis of new research and advancement in applications of animal breed recognition systems.

5.4 Limitations of the Study

Although the proposed system of cat breed identification was developed and evaluated successfully, there were certain limitations that were discovered in this study. These constraints give valuable insight into the limitations of the existing approach and point towards possible areas of improvement.

The weakness of this study is associated with the diversity of the datasets. Even though the dataset one used was sufficient in training and testing the CNN models, it might not be reflective of every real-world situation. Not all the breeds of cats that were sampled contained similar numbers, and this could influence the ability of the model to generalize the same across all the classes. Moreover, the dataset was characterized by a small number of different backgrounds and lighting configurations, where these factors can potentially lead to low performance in cases where the model is applied to images of complex or unfamiliar backgrounds.

The other weakness is that the model is sensitive to occlusion and complexity of the background. When real-world testing was conducted using the mobile app, the model sometimes failed to identify or classify cats accurately when partial occlusion changed in an image, for example, when a person carrying or holding a cat or when the background was not clear. This means that the model is highly dependent upon visual characteristics that are distinct and could be problematic in cases where such characteristics are not visible or are blocked by other objects in the environment.

The models in the scope of architectures considered in this paper have a limitation as well. A total of only two CNN architectures, MobileNetV3 and

ResNet50, were used. Whereas at least these models performed strongly, other architectures, even hybrid ones, can be more robust or efficient. Moreover, the research was carried out on the problem of image classification as opposed to object detection, and this restricts the functionality of the system to find cats in complicated scenes.

Lastly, there were computational and time restrictions to achieve the degree of experimentation and optimization. Hyperparameter configuration and dataset enlarging were performed within the limited resources, and possibly additional fine-tuning or large-scale training may give better results. On the same note, the application testing on mobile devices was conducted on a diverse set of devices that might not be exhaustive in terms of the differences in performance on various hardware settings.

Overall, these shortcomings do not compromise the efficacy of the given system but rather offer guidelines for future improvements, which will be addressed in the following section.

5.5 Recommendation for Future Work

Following the limitations that were observed during this study, a number of recommendations are proposed on how to make the cat breed classification system more effective and less prone to failures in future studies.

First, the future studies could work on the further expansion and diversification of the data. This has consisted of adding more images per breed and adding more cat breeds other than the 12 classes that are already present. The use of images taken at different light conditions, backgrounds, and perspectives would also enhance the generalization of the models. Moreover, adding pictures with partial interruption, more than one cat, and actual environmental changes can strengthen the model to be used in practical conditions.

Second, complex data augmentation methods can be used to enhance the model's resistance to background clutter and interference. The following techniques can be used to subject the model to a larger set of visual variations in the course of training, such as random cropping, background swapping, colour jittering, and image blending, which has the effect of decreasing sensitivity to environmental noise.

Third, future studies could investigate the fusion of object detection or localization models before classification of the breeds. The classification model is capable of concentrating on important aspects by initially identifying and segregating the cat in a picture and reducing the interference of the background. The method works especially well with the pictures in which cats are partially covered or have complex backgrounds.

Moreover, future research will be able to examine the application of other or more sophisticated model architectures, like EfficientNet or transformer-based vision models, to further enhance fine-grained classification results. It can also be investigated using model compression and optimization methods, such as quantization and pruning, to make inference faster and more efficient on a mobile device.

Lastly, the deployment of the mobile application can also be improved further. It would be possible to test the application on more devices and operating conditions to have a more detailed assessment of the actual performance. Continuous improvements of predictions, as stated by a user feedback mechanism, can also be executed to adapt the accuracy of predictions in the long term.

Overall, these suggestions give straightforward guidelines on how to expand on the existing study and enhance the usability of CNN-based cat breed classifications in the real world.

5.6 Summary

This chapter completed the study part by summarizing the main findings, measuring the project objectives' achievement, discussing the contributions and limitations of the work, and giving suggestions as to the way of improvement in the future. The overall results showed that methods based on convolutional neural networks can be used and are effective in the image-based classification of cat breeds and may be effectively utilized in the real world.

The development and evaluation of CNN models, analysis of overall performance by means of relevant evaluation metrics, and successful adaptation of the chosen model into a mobile app allowed accomplishing the project objectives. This work has the contribution of the systematic comparison of deep learning architectures, practical implementation in a mobile-based setting, and the insightful details on model behaviour in both controlled and real-world settings.

Although the system showed a high level of performance, a number of limitations were realized, especially in the lines of diversity of databases, complexity of the background, and susceptibility to occlusion. These drawbacks also show the possibility of improvements in the future, such as the expansion of the dataset, the addition of other kinds of cats, and the implementation of more sophisticated methods to make the study more resilient and suitable for generalization.

Finally, this chapter summarizes results of the work carried out throughout the chapter and highlights possibilities of deep learning methods as the potential for the animal breed classification system in the future. The suggestions made offer a good direction on how the proposed system could be expanded and brought about in further studies.

REFERENCES

Alzubaidi, L., Zhang, J., Humaidi, A. J., Al-Dujaili, A., Duan, Y., Al-Shamma, O., Santamaría, J., Fadhel, M. A., Al-Amidie, M., & Farhan, L. (2021). Review of deep learning: concepts, CNN architectures, challenges, applications, future directions. *Journal of Big Data*, *8*(1). https://doi.org/10.1186/s40537-021-00444-8

Barrios, F., Suárez, G., Udell, M. A. R., & Damián, J. P. (2023). Characterization of the Domestic Cat Population of Uruguay: Breeds, Coat Colors, Hair Length, Lifestyle, Sex and Spay/Neuter Status According to Guardian Report. *Animals*, *13*(12). https://doi.org/10.3390/ani13121963

Carratino, L., Jenatton, R., & Vert, J.-P. (2022). On Mixup Regularization. In *Journal of Machine Learning Research* (Vol. 23). http://jmlr.org/papers/v23/20-1385.html.

Chen, M. (2023). Exploring the Categories and Main Morphological Features of Cats. *International Journal of Molecular Zoology*. https://doi.org/10.5376/ijmz.2023.13.0001

Feng, Y. (2025). *Integrating Object Detection and Deep Convolutional Neural Networks for Cat Breed Classification*.

Filippov, S. A. (2024). Image Classification Using Convolutional Neural Networks. *Automatic Documentation and Mathematical Linguistics*, *58*(S3), S143–S149. https://doi.org/10.3103/S0005105525700219

Hamdi, E. B., Sunaryo, J. A., & Prasetyo, S. Y. (2023). Fusion of pretrained CNN models for cat breed classification: A comparative study. *E3S Web of Conferences*, *426*. https://doi.org/10.1051/e3sconf/202342601014

Hussain Khan, S., & Iqbal, R. (2025). *A Comprehensive Survey on Architectural Advances in Deep CNNs: Challenges, Applications, and Emerging Research Directions*.

Karad, V., Jabde, M., Patil, C. H., Bhoite, S., & Patil, H. (2025). *Deep Learning based Cat Breed Recognition*. https://www.researchgate.net/publication/389339621

Karlita, T., Choirunisa, N. A., Asmara, R., & Setyorini, F. (2022a). Cat Breeds Classification Using Compound Model Scaling Convolutional Neural Networks. *Proceedings of the International Conference on Applied Science* *and Technology on Social Science 2021 (ICAST-SS 2021)*, 909–914. https://doi.org/10.2991/assehr.k.220301.150

Karlita, T., Choirunisa, N. A., Asmara, R., & Setyorini, F. (2022b). *Cat Breeds Classification Using Compound Model Scaling Convolutional Neural Networks*. https://doi.org/10.2991/assehr.k.220301.150

Kaur, P., Khehra, B. S., & Mavi, E. B. S. (2021). Data Augmentation for Object Detection: A Review. *Midwest Symposium on Circuits and Systems*, *2021-August*, 537–543. https://doi.org/10.1109/MWSCAS47672.2021.9531849

Kumar, T., Mileo, A., Brennan, R., & Bendechache, M. (2023). *Image Data Augmentation Approaches: A Comprehensive Survey and Future directions*. https://doi.org/10.1109/ACCESS.2024.3470122

Lei, F., Liu, X., Dai, Q., & Ling, B. W. K. (2020). Shallow convolutional neural network for image classification. *SN Applied Sciences*, *2*(1). https://doi.org/10.1007/s42452-019-1903-4

MITHILESH VISHWAKARMA, KUNIKA JANGID, KIRAN PAL, & LISA RODRIGUES. (2024). *Cat Breed & Emotion Detection Using Yolo. CNN & Canny Edge Detection*.

Noor, M. H. M., & Ige, A. O. (2024). *A Survey on State-of-the-art Deep Learning Applications and Challenges*. http://arxiv.org/abs/2403.17561

Qatrunnada, N., Fachrurrozi, M., & Utami, A. S. (2022). Cat Breeds Classification using Convolutional Neural Network for Multi-Object Image. In *Sriwijaya Journal of Informatic and Applications* (Vol. 3, Issue 2). http://sjia.ejournal.unsri.ac.id

Qian, S., Ning, C., & Hu, Y. (2021). MobileNetV3 for Image Classification. *2021 IEEE 2nd International Conference on Big Data, Artificial Intelligence and Internet of Things Engineering, ICBAIE 2021*, 490–497. https://doi.org/10.1109/ICBAIE52039.2021.9389905

Ramadhan, A. T., & Setiawan, A. (2023). Catbreedsnet: An Android Application for Cat Breed Classification Using Convolutional Neural Networks. *Jurnal Online Informatika*, *8*(1), 52–60. https://doi.org/10.15575/join.v8i1.1007

Rangkuti, A. H., Putra Wibawa, H. M., Tabina, K., Setia, G. A., Al-Rasyid, G. N., & Maulana, M. (2024). Harnessing Deep Learning for Cat Breed Identification: A Visual Recognition Strategy. *7th International Seminar on Research of Information Technology and Intelligent Systems: Advanced*

Intelligent Systems in Contemporary Society, ISRITI 2024 - Proceedings, 718–723. https://doi.org/10.1109/ISRITI64779.2024.10963598

Sudharson, S., Deena, S., & Reddy, S. N. (2023). Efficient Real-time Breed Classification using YOLOv7 Object Detection Algorithm. *2023 14th International Conference on Computing Communication and Networking Technologies, ICCCNT 2023*. https://doi.org/10.1109/ICCCNT56998.2023.10306688

Tan, M., & Le, Q. V. (2020). *EfficientNet: Rethinking Model Scaling for Convolutional Neural Networks*.

Wang, X., Zhang, B., Xu, Z., Qin, Y., & Philip Chen, C. L. (2025). An exploration of deep learning for decision making: Methods, applications, and challenges. In *Expert Systems with Applications* (Vol. 289). Elsevier Ltd. https://doi.org/10.1016/j.eswa.2025.128406

Yu, Y., Creighton, E. K., Buckley, R. M., Lyons, L. A., Buckley, R. M., Aberdein, D., Alves, P. C., Barsh, G. S., Bellone, R. R., Bergström, T. F., Boyko, A. R., Brockman, J. A., Casal, M. L., Castelhano, M. G., Distl, O., Dodman, N. H., Ellinwood, N. M., Fogle, J. E., Forman, O. P., … Lyons, L. A. (2020). A deletion in GDF7 is associated with a heritable forebrain commissural malformation concurrent with ventriculomegaly and interhemispheric cysts in cats. *Genes*, *11*(6), 1–15. https://doi.org/10.3390/genes11060672

Yuxiang, X., Quanzhi, G., Xidao, L., Jie, Y., & Jiahui, Z. (2023). A survey of fine-grained visual categorization based on deep learning. *Journal of Systems Engineering and Electronics*, 1–20. https://doi.org/10.23919/jsee.2022.000155

Zayani, H. M., Kachoukh, A., Ghodhbani, R., Khediri, N., Abd-Elkawy, E. H., Ammar, I., Kouki, M., & Saidani, T. (2025). Cat Breed Classification with YOLOv11 and Optimized Training. *Engineering, Technology and Applied Science Research*, *15*(2), 21652–21657. https://doi.org/10.48084/etasr.10218

Zeng, W. (2024). Image data augmentation techniques based on deep learning: A survey. In *Mathematical Biosciences and Engineering* (Vol. 21, Issue 6, pp. 6190–6224). American Institute of Mathematical Sciences. https://doi.org/10.3934/mbe.2024272

Zhang, R. (2021). Classification and Identification of Domestic Catsbased on Deep Learning. *Proceedings - 2021 2nd International Conference on Artificial*

Intelligence and Computer Engineering, ICAICE 2021, 106–110. https://doi.org/10.1109/ICAICE54393.2021.00029

Zhao, X., Wang, L., Zhang, Y., Han, X., Deveci, M., & Parmar, M. (2024). A review of convolutional neural networks in computer vision. *Artificial Intelligence Review*, *57*(4). https://doi.org/10.1007/s10462-024-10721-6

Zhuang, F., Qi, Z., Duan, K., Xi, D., Zhu, Y., Zhu, H., Xiong, H., & He, Q. (2021). A Comprehensive Survey on Transfer Learning. In *Proceedings of the IEEE* (Vol. 109, Issue 1, pp. 43–76). Institute of Electrical and Electronics Engineers Inc. https://doi.org/10.1109/JPROC.2020.3004555

APPENDICES

APPENDIX 1

Gantt Chart Representing the project timeline and task schedule