



Daffodil
International
University

Dog Breed Classification Using Machine Learning Algorithms

Submitted by

SK Shihabuzzaman

ID: 213-35-3188

Department of Software Engineering

Daffodil International University

Supervised by

Dr. Imran Mahmud

Professor and Head

Department of Software Engineering

Daffodil International University

This Thesis paper has been submitted in fulfillment of the requirements for the degree of Bachelor of Science in Software Engineering.

Summer 2025

© All Rights Reserved by Daffodil International University

Dog Breed Classification Using Machine Learning Algorithms

SK Shihabuzzaman
ID: 213-35-3188

Bachelor of Science
Department of Software Engineering

DAFFODIL INTERNATIONAL UNIVERSITY

DAFFODIL INTERNATIONAL UNIVERSITY

DECLARATION OF THESIS AND COPYRIGHT

Author's Full Name : SK Shihabuzzaman
Date of Birth : 21/10/2002
Title : Dog Breed Classification Using Machine Learning Algorithms
Academic Session : 2021~2025

I declare that this thesis is classified as:

- CONFIDENTIAL (Contains confidential information under the Official Secret Act 1997)*
- RESTRICTED (Contains restricted information as specified by the organization where research was done)*
- OPEN ACCESS I agree that my thesis to be published as online open access (Full Text)

I acknowledge that Daffodil International University reserves the following rights:

1. The Thesis is the Property of Daffodil International University.
2. The Library of Daffodil International University has the right to make copies of the thesis for the purpose of research only.
3. The Library of Daffodil International University has the right to make copies of the thesis for academic exchange.


Certified by:



(Student's Signature)

213-35-3188

Student ID
Date: 20/09/2025



(Supervisor's Signature)

Dr.Imran Mahmud

Name of Supervisor
Date: 20/09/2025



SUPERVISOR'S DECLARATION

I hereby declare that I have checked this thesis and in my opinion, this thesis is adequate in terms of scope and quality for the award of the degree of Bachelor of Science

A handwritten signature in black ink, appearing to read "Imran", is positioned above a horizontal line.

(Supervisor's Signature)

Full Name : Dr.Imran Mahmud
Position : Professor and Head of Software Engineering
Date : 20 September,2025



STUDENT'S DECLARATION

I hereby declare that the work in this thesis is based on my original work except for quotations and citations which have been duly acknowledged. I also declare that it has not been previously or concurrently submitted for any other degree at Daffodil International University or any other institution.

A handwritten signature in black ink, appearing to read "SK Shihabuzzaman", is written on a light-colored background.

(Student's Signature)

Full Name : SK Shihabuzzaman
ID Number : 213-35-3188
Date : 20 September, 2025



APPROVAL PAGE

APPROVAL

This thesis titled on **Dog Breed Classification Using Machine Learning Algorithms**, submitted by **SK Shihabuzzaman (ID:213-35-3188)** to the Department of Software Engineering, Daffodil International University has been accepted as satisfactory for the partial fulfillment of the requirements for the degree of Bachelor of Science in Software Engineering and approval as to its style and contents.

BOARD OF EXAMINERS

A handwritten signature in black ink, appearing to be "Imran", written over a horizontal line.

Dr. Imran Mahmud
Professor & Head
Department of Software Engineering
Faculty of Science and Information Technology
Daffodil International University

Chairman

A handwritten signature in black ink, appearing to be "Shohel", written over a horizontal line.

Md Shohel Arman
Assistant Professor
Department of Software Engineering
Faculty of Science and Information Technology
Daffodil International University

Internal Examiner 1

A handwritten signature in black ink, appearing to be "Rajib", written over a horizontal line.

Md. Rajib Mia
Lecturer (Senior Scale)
Department of Software Engineering
Faculty of Science and Information Technology
Daffodil International University

Internal Examiner 2

A handwritten signature in black ink, appearing to be "Habibur", written over a horizontal line.

Md Habibur Rahman
Associate Professor
Department of Computer Science and Engineering
Islamic University, Bangladesh

External Examiner

Dog Breed Classification Using Machine Learning Algorithms

SK Shihabuzzamam

Thesis submitted in fulfillment of the requirements
for the award of the degree of
Bachelor of Science

Department of Software Engineering

September 2025

©Daffodil International University

ACKNOWLEDGEMENTS

I would like to sincerely thank everyone who helped and guided me throughout this research and this thesis. I would like to thank, first and foremost, my supervisor, Dr. Imran Mahmud, who has offered me homemade food, and the kind of guidance only someone of Dr. Imran Mahmud's knowledge, patience, and willingness to support me from the beginning of this work to the end could offer. I am grateful for the support of other committee members whose feedback and continued support improved the quality of the work. The faculty and staff of FSIT at Daffodil International University deserve my thanks for providing the academic climate and the support to conduct my academic study. I also thank my colleagues and friends who provided academic discussions, moral support, and collaborations during this difficult period. I will thank the creators and maintainers of the Orange ML platform and publicly available dog breed datasets, which made this study possible. Finally, I would like to thank my family for their unconditional love, patience, and motivation to keep me doing my academic work.

DEDICATION

THIS THESIS IS DEDICATED TO MY FAMILY, WHOSE LOVE AND SUPPORT HAVE PROVIDED THE FOUNDATION AND INSPIRATION FOR EVERYTHING I DO. TO MY PARENTS, FOR THEIR UNWAVERING ENCOURAGEMENT AND THE SACRIFICES THEY MADE TO ALLOW ME TO PURSUE MY EDUCATIONAL OPPORTUNITIES. TO MY FRIENDS AND MENTORS, FOR THE BELIEF THEY HAD IN ME AND INSPIRING ME TO FACE EVERY CHALLENGE. AND TO ANYONE WHO IS DEDICATED TO THE ADVANCEMENT OF KNOWLEDGE AND THE CARE OF ANIMALS, I HOPE THIS WORK IS VALUABLE IN SOME CAPACITY TO BOTH.

ABSTRACT

The thesis completes a thorough study on the task of dog breed recognition using conventional machine learning techniques on the Orange ML interface. It tackles the problem of recognizing dog breeds by implementing a Logistic Regression, Support Vector Machines, K-Nearest Neighbors, Random Forest, and a shallow Neural Network. The experiments done in this study were successful in classifying breeds which were visually different, but were still struggling with distinguishing breeds which looked alike. This research stands out for proving the effectiveness of no-code machine learning on complex image classification, analyzing the performance of different traditional algorithms, and empirically describing the problems of classification in the image data set. The research serves as a starting point for automated image recognition technology of animals and classifies prospective research areas in fine-grained image classification in the absence of deep learning structures.

Keywords: Dog breed recognition, Machine learning, Orange ML, No-code, Image classification, fine-grained classification, traditional machine learning, Logistic Regression, Support Vector Machines, K-Nearest Neighbors, Random Forest, Neural Network.

TABLE OF CONTENT

DECLARATION	
TITLE PAGE	
ACKNOWLEDGEMENTS	ii
ABSTRACT	iv
TABLE OF CONTENT	v
LIST OF TABLES	viii
LIST OF FIGURES	ix
LIST OF SYMBOLS	x
LIST OF ABBREVIATIONS	xi
CHAPTER 1 INTRODUCTION	01
1.1 Background	01
1.2 Problem Statement	02
1.3 Research Questions, Objectives & Scope	03
1.3.1 Research Questions	03
1.3.2 Research Objectives	03
1.3.3 Research Scope	04
1.4 Research Gaps & Transition to Objective	05
1.5 Significance of the Study	06
CHAPTER 2 LITERATURE REVIEW	08
2.1 Overview	08

2.2 Related work on Dog Breed Classification Using Machine Learning Algorithms	08
CHAPTER 3 METHODOLOGY	17
3.1 Dataset Description	19
3.2 Data Preprocessing and Feature Extraction	21
3.3 Machine Learning Models	23
3.4 Model Evaluation Metrics	25
3.5 Visualization and Qualitative Analysis	30
3.6 Summary	32
CHAPTER 4 RESULTS AND ANALYSIS	33
4.1 Dataset Summary	34
4.2 Performance Metrics Across Models	36
4.3 Confusion Matrix Analysis	36
4.4 ROC and Precision-Recall Curves	38
4.5 Learning Curve	40
4.6 Dataset Visualization	40
CHAPTER 5 DISCUSSION	43
5.1 Discussion Content	44
CHAPTER 6 CONCLUSION	47
6.1 Summary of Findings	37

6.2 Contributions	48
6.3 Strengths	51
6.4 Limitations	52
6.6 Future Work	53
REFERENCES	55

LIST OF TABLES

Table 3.1	Dataset Sample Distribution by Breed	20
Table 3.2	Five models were trained and evaluated.	24
Table 3.3	Evaluation Results.	28
Table 3.4	Overall Model Performance Comparison.	29
Table 4.1	Distribution of dataset samples across dog breeds	35
Table 4.2	Performance comparison of models on the test set.	36
Table 4.3	Confusion matrix for SVM model on test set.	37

LIST OF FIGURES

3.1	Distribution of dog breeds in the dataset, illustrating balanced sample sizes per class.	22
3.2	The Trained Orange ML Model.	26
3.3	Scatter plot of size versus breed, highlighting variation and outlier samples within the dataset.	31
4.1	Presents the ROC curves for each breed classification using the Support Vector Machine model.	38
4.2	Shows the Precision-Recall curves.	39
4.3	The learning curve depicts training and validation accuracies as a function of training set size.	40
4.4	Dimensionality reduction via t-SNE illustrates clear clustering of image embeddings by breed, supporting feature separability that aids classification accuracy.	41
4.5	Dimensionality reduction via t-SNE illustrates clear clustering of image embeddings by breed, supporting feature separability that aids classification accuracy.	42

LIST OF SYMBOLS

SYMBOLS	DESCRIPTION
ACC	ACCURACY- PROPORTION OF CORRECTLY CLASSIFIED INSTANCES
P	PRECISION- PROPORTION OF TRUE POSITIVES AMONG PREDICTED POSITIVES
R	RECALL- PROPORTION OF TRUE POSITIVES AMONG ACTUAL POSITIVES
F1	F1-SCORE-HARMONIC MEAN OF PRECISION AND RECALL
K	NUMBER OF NEIGHBORS IN K- NEAREST NEIGHBORS
C	REGULARIZATION PARAMETER IN SVM WITH RBF KERNEL
Γ	KERNEL COEFFICIENT IN SVM WITH RBF KERNEL
LR	LOGISTIC REGRESSION MODEL
RF	RANDOM FOREST MODEL

LIST OF ABBREVIATIONS

ANN	Artificial Neural Network
AUC	Area Under the ROC Curve
CNN	Convolutional neural network
DIU	Daffodil International University
FSIT	Faculty of science technology
FN	False Negative
FP	False Positive
kNN	K- Nearest Neighbors
LR	Logistic Regression
ML	Machine Learning
MLP	Multi-Layer Perception
PCA	Principle Component Analysis
PR Curve	Precision-Recall Curve
RF	Random Forest
ROC	Receiver Operator Characteristics
SVM	Support Vector Machine
TN	True Negative
TP	True Positive
VGG19	Visual Geometry Group 19-layer

CHAPTER 1

INTRODUCTION

1.1 Background

The precise determination of dog breeds from images is becoming both a challenge and a possibility in machine learning and computer vision. It can be useful in many areas, including animal shelter and veterinary diagnostics, public safety, mobile apps, as well as in scholarly work about dogs' genes and behaviors. Older methods of breed identification through visual checking of identification cards and microchip scanning were often inaccurate, slow, and very subjective. Automated methods in image classification through machine learning and deep learning have greatly advanced in the past few years. It is now possible to identify a breed through pictures, even complex and real-life photos. Yet, progress is still elusive. Some breeds like Labrador retrievers and Beagles can be very hard to identify using automated systems due to their overlapping sizes, coat colors, and shapes. An example of what is known as a fine grained classification problem, these systems are required to discriminate between complex, high-level visual components of images.

Within this setup, this thesis explores multiple machine learning algorithms Logistic Regression, k-Nearest Neighbors (kNN), Neural Networks, Random Forest, and Support Vector Machines (SVM), in order to explore how well they perform, what their shortcomings are, and are they well suited to applied settings.

1.2 Problem Statement

Automatically identifying dog breeds is a notoriously difficult problem in computer vision and machine learning for a variety of intertwined reasons. The first is that there are more than 350 known breeds of dog around the world. This breed diversity encompasses lots of variation in size, coat, and shape.

The second issue is the considerable intra-breed variability arising from changes in lighting, pose, background, image quality or image resolution affects the consistency and reliability of features we could extract. Even worse, many breeds that are misclassified look fairly similar, and overlap in physical features, which is troublesome even with advanced models.

These two sets of problems become even more difficult to address when considering deployment scenarios in the real world, where breed classification models need to generalize to image conditions including user-generated content, and limited computational resources such as in mobile applications.

Although the literature has shown credible performance results with these advanced deep learning architectures, there is still a lack of direct comparison of simple and complex machine learning models via a well-balanced representative dataset.

Furthermore, there is limited research that systematically analyzes the underlying causes of misclassification in a reproducible manner.

The development of deployable and interpretable classification systems, specifically designed for practical environments such as veterinary clinics, animal shelters, and mobile applications, is still underexplored. Therefore, this research addresses the need for a rigorous comparative evaluation of diverse machine learning paradigms, with a strong focus on generalizability, performance interpretation, and real-world applicability in automated dog breed classification.

1.3 Research Questions, Objectives & Scope

1.3.1 Research Questions

This study is guided by the following key research questions:

1. Which machine learning models provide the highest accuracy and robustness for multiclass dog breed classification on a balanced dataset?
2. How do these models perform when classifying visually similar breeds that present fine-grained differences?
3. What are the practical trade-offs between model complexity, interpretability, accuracy, and computational efficiency for real-world deployment?
4. How can evaluation and visualization tools (e.g., confusion matrices, ROC curves, precision-recall curves, learning curves, decision trees) be used to enhance interpretability and support informed model selection?

1.3.2 Research Objectives

1. Systematically develop, train, and evaluate a set of traditional machine learning algorithms for dog breed classification based on image data.
2. Conduct a comparative analysis across five traditional algorithms: Logistic Regression, k-Nearest Neighbors (kNN), shallow Neural Network, Random Forest, and Support Vector Machines (SVM). The selection of these algorithms was informed by the need for interpretable linear to sophisticated nonlinear models.
3. Apply the developed models to a balanced and representative image dataset comprising eight dog breeds from varying visual degrees of complexity and similarity to allow for a rigorous evaluation.
4. Utilize multiple evaluation metrics to quantify model performance: accuracy, precision, recall, and the F1-score, alongside appropriate visualizations, confusion matrices, receiver operating characteristic (ROC) curves, precision-recall curves, and learning curves, to gain a more nuanced understanding of model behaviours and diagnostic performance.

5. Explore the nature and sources of misclassification ambiguity, especially among breeds with visual similarities, and identify the potential strategies for improvement.
6. Provide evidence-based recommendations for future work with respect to dataset diversity augmentation for model training, improving the robustness of machine learning models, and applicable implementation strategies.
7. Ultimately, the relevance of this project is in supporting automated dog breed classification tasks in learning environments such as veterinary clinics, or animal shelters, as well as in mobile app/web applications for both the community and professionals.

Overall, the objectives outline the legitimate development of an interpretable, efficient and reproducible machine learning framework.

1.3.3 Research Scope

Classification across a reliable and reproducible experimental framework.

1. It uses a balanced dataset of 3,900 labeled images, with each of the eight breeds (golden retriever, pug, pug, poodle, bulldog, hound, working dog, labrador) represented equally to help visualize the possibility of bias in the training run prior to the progress toward fine-grain classification and comparative analysis.
2. The analysis includes traditional machine learning models, where both linear models (Logistic Regression, k-Nearest Neighbors) (interpretable) include multi-class algorithms (Random Forest, shallow Neural Network, Support Vector Machine) that provide more complexity/responsiveness.
3. This article will utilize an 80-20 stratified train-test split that will allow for maintained breed association consistency and will include 10-fold cross-validation, ensuring a strong and reproducible performance setting.
4. It will focus on visual pattern recognizing using static images only, not using multimodal/orthogonal data sets with any text-based elements, genomic features, audio components nor any classification type tasks distributive from dog breed classification.

5. It will not look at any deployment, computational efficiency, including optimizations for mobile, embedded and embedded systems in order to strictly maintain methodological rigor and comparative performance evaluation in controlled settings.
6. The research has excluded elements / levels of extrinsic variables by removing images with high levels of variability (lighting, background, pose) and maintained clear methodological overview of performance gaps.

This small and focused scope will create a high-quality opportunity for evaluating the benefits and limitations of traditional machine learning algorithms for accurate fine-grain dog breed classification, and comparisons across a small and reproducible experimental framework with limits defined.

1.4 Research Gaps & Transition to Objective

1. While there has been rapid progress in machine learning and deep learning methods that have been applied to dog breed classification, there are a number of notable gaps in the research and practical application.
2. Many research studies have focused solely on deep learning models without comparisons or baseline experimentation with simple and more transparent traditional machine learning algorithms.
3. There is little research that has utilized well-balanced datasets that fairly represent multiple dog breeds, impacting the generalizability and benchmarking of the models.
4. The practical challenges of deploying models including computational efficiency, usability on low-resource, possibly user-owned devices (e.g., mobile phones), and variability in photographs taken in real-world situations is neglected.
5. A more detailed error analysis of classification errors, especially between visually similar breeds, has been limited, which limits knowledge of the limitations of the models and areas for potential improvement.

6. Many model have not performed well with external slider, user-generated images, with variability of environmental context and class imbalances .
7. This thesis addresses these gaps by heavily benchmarking on simple reproducible traditional machine learning algorithms (Logistic Regression, kNN, Neural Networks, Random Forest, SVM) on a balanced dataset with a sample size of eight dog breeds.
8. Emphasis has been placed on reproducible experimental design with comprehensive performance evaluation measures, visualization measures, and deep qualitative experimentative learning from misclassification patterns.
9. The objective was to develop breed classification systems that are: accurate, interpretable, efficient, and appropriate for practical real-world applications such as.

1.5 Significance of the Study

Through evaluating the effectiveness of various conventional machine learning techniques such as Logistic Regression, Support Vector Machines, k-Nearest Neighbors, Random Forest, and shallow Neural Networks on a balanced real-world dog breed dataset, this research reinforces the academic and practical knowledge in automated breed classification systems. It helps create accurate, reliable, and easy-to-use tools for automated breed classification and identification which are beneficial in veterinary science, animal shelter operations, public awareness and education, and mobile app development.

This study provides valuable insight, primarily in terms of the use of different algorithms in fine-grained image classification of dog breeds and emphasizes the importance of balanced datasets and good feature selection for successful classification even when using simpler models, like Random Forest, rather than complex/multiple deep learning systems. The fact that all the experiments were performed in the no-code Orange ML platform provides a case study of the capability of conducting a thorough and complicated evaluation of different algorithms in a transparent and accessible way. This study is a small stride towards deeper understanding in the application and

limitations of various machine learning techniques in complex image recognition problems. Further, this study enhances our understanding of the characteristics of visually similar breeds, which continue to present challenges to identification.

Identifying dog breeds correctly and efficiently serves multiple practical purposes. In veterinary medicine, prompt accurate identification enables more effective diagnosis, informs treatment in a veterinary center, and facilitates communication with clients about health issues and breed specific predispositions. For animal shelters and animal control agencies, automated breed recognition ensures accurate record keeping. It improves efficiency in animal intake and rehoming and facilitates effective management of animal populations. For informed pet ownership and responsible pet breeding, classification aids in educating prospective buyers about caring for specific breeds, training, and verifying lineages and assessing compatibility for breeding. It aids responsible ownership and breeding. In safety, dependable classification aids smart decisions in enforcing local laws and safety protocols that relate to specific breeds.

The results could further immediate development of mobile technologies for public use such as mobile applications that classify dogs from photographs thereby increasing public awareness and accessibility to technology. Moreover, the no-code approach to implementing Orange ML illustrates the possibility of democratizing machine learning and making it accessible. The successful demonstration also highlighted the no-code implementation of Orange ML demonstrates the platforms ability to democratize machine learning and allow a much broader audience, including non-coding users, to develop, evaluate, and deploy solid analytical products to real-world problems.

CHAPTER 2

LITERATURE REVIEW

2.1 Overview

This chapter reviews the emerging field of dog breed classification, with a focus on traditional machine learning approaches and easy-to-implement no-code offerings like Orange ML. It summarizes some of the common challenges in image classification, describes the main algorithms and methods, and illustrates where the field is today. It also highlights new developments that deal with these problems. These observations collectively form the basis of the current investigation and elucidate the ways in which its contributions align with computer vision and animal classification, particularly when considering accessible, no-code development (Dog Breed Classification Using Machine Learning).

2.2 Related work on Dog Breed Classification Using Machine Learning Algorithms

Cui et al. (2024) presented a comprehensive approach to dog breed classification, which combined the use of multiple convolutional neural networks (Inception V3, InceptionResNet V2, NASNet and PNASNet) for feature extraction with principal component analysis (PCA) and gray wolf optimization (GWO) for feature selection, followed by classification using a support vector machine (SVM). Their method achieved improved accuracy rates (95.24%) and incorporated an expanded feature set for 120 dog breeds using the Stanford Dog Dataset, which greatly improved the single CNN models. Noteworthy contributions include their combination of the features from varying CNN architectures, the incorporation of optimization algorithms to obtain a suitable sub-set of features, and finally, the use of SVM for optimum decision boundaries. Collective/combined to a hybrid model shows drastic improvements to specialized tasks such as fine grain breed identification. (Cui et al., 2024)

The thorough examination provided by Ostrander et al. (2017) reviews how the genetic architecture and complicated demographic history of domestic dogs has resulted from the quick selection, bottlenecks associated with domestication, and repeated introgressions with wild ancestors. In our context, the genetic condition contorting breed identification is that breeds are often very genetically variable within-class - but exhibit many or overlapping phenotypic traits as to be difficult to separate. Ostrander et al. (2017) note that most populations have bottle-necked and then undergone intense selective breeding that has canalized certain phenotypic polymorphisms - that may manifest as very minor visually aggregated differences, and may only be isolated with very advanced analytical or computational techniques. This situational evolution helps to highlight the need to deploy powerful and valid scalable classification techniques that can identify very subtle differences - that would distinguish dog breeds. (A Ostrander et al., 2017)

Valarmathi et al. (2023) evaluate a number of modern CNN architectures (Xception, NASNetMobile, EfficientNetV2, ResNet (and its variants)) which include both single and hybrid models in dog breed classification using the Stanford Dog Dataset. They propose a hybrid model of Inception V3 and Xception that produces the highest recorded accuracy (92.4%) when compared to individual models. The research illustrates the power of transfer learning combined with data augmentation techniques as well as shows the efficiency and extensibility of hybrid convolutional neural networks with respect to dog breed classification involving 120 breed classes. The thorough examination allows for understanding important architectural decisions, hyperparameter tuning, and augmentations used to achieve highly precise dog breed classification. (Valarmathi et al., 2023)

Though focused on acute lymphoblastic leukemia (ALL), Raman et al. (2021) exemplify the utility of combining CNNs for feature extraction (Inception V3) with gradient boosting classifiers (XGBoost) to improve classification accuracy and interpretability in medical imaging. The model not only achieved a weighted F1 score

of .986 on a challenging dataset, but also showcased the effectiveness of a hybrid deep learning and machine learning model for narrowly defined classification tasks. This suggests the feasibility of applying hybrid architectures for dog breed classifications, noting that the richness of features and decision boundaries can be discovered by combining deep features produced by a fixed network with classifiers much more complex than the softmax output layers typically used as final output.(Ramaneswaran et al., 2021)

The study by Khan et al. identifies emerging trends in animal breed classification through hybrid models (deep CNN, combined with classical machine learning classifiers, e.g., SVM, Random Forest) which attained improved classification metrics when compared with CNNs alone or methods reliant on handcrafted features. Specifically, the authors highlight that fusing features from a variety of deep learning models, combined with advanced classifiers can yield advances in accuracy and interpretability. Limitations with class imbalance, limited labelled data and model explainability were noted, confirming the value of hybrid architectures that are purpose built for fine-grained classification tasks such as dog breed recognition. (Sardar Khan et al., 2023)

Madhan et al. (2022) describe processes, such as Canny edge detection, normalizing, dimension changing and augmentations (rotation, flipping), that enabled images to be inputted into a CNN (e.g. MobileNetV2) for dog breed classification. Their image processing was used in a cloud-based apps controlled with mobile front-ends. Their model achieved accuracies of 81% on Stanford Dog Dataset (SDSD) images. They note how preprocessing images was a crucial aspect of this project to enable the input of cleanly dependent variables (images) for efficient learning and inference by the CNN, which is needed for scalable breed identification and other concepts. (Madhan et al., 2022)

Ayanzadeh and Vahidnia (2018) consider tweaks of established CNN architectures like ResNet, DenseNet, and GoogLeNet for fine-grained dog breed classification, focusing on transfer learning with fine-tuning and data augmentation to cope with a lack of data. They reached accuracies up to 89.66% on the Stanford Dog Dataset by employing pre-trained weights from ImageNet while changing some layers of the network to constrain model overfitting. Their work establishes the effectiveness of modern deep architectures and transfer learning in the areas of veterinary and image-based animal identification, and provide details about the area of dataset pre-processing and optimizing training. (Ayanzadeh & Vahidnia, 2018)

Data augmentation methods have been successful in improving the performance of image classification tasks, especially in the case of limited labelled datasets. Wang and Perez (2017) investigated various forms of data augmentation including conventional affine transformations, data transfer via GAN, and neural augmentation, in which a given neural network learns augmentations to minimize classification loss. The researchers conducted experimental comparisons on two datasets, Tiny-ImageNet, and MNIST, with their findings indicating that neural augmentation outperformed the traditional methods. Neural augmentation presents new prospects for improving classifier generalization at an increased computation cost. (Wang & Perez, 2017)

Bokhari et al. (2022) performed dog breed classification using a CNN-based approach to take advantage of transfer learning with the Stanford Dogs dataset which has 120 dog breeds. The authors mentioned that they faced a challenge of class imbalance and relatively few samples for each dog breed. They employed MobileNet V2 so that they would potentially decrease model complexity, resulting in an efficient final result and enabling the authors to train a relatively deep CNN. The authors reported promising accuracy and that transfer learning is quite useful when developing deep learning models for fine-grained classification tasks that have limited amounts of data and proposed future work using background masking and dimensionality reduction. (Bokhari et al., 2022)

Gaona Chong (2018) introduced a deep learning framework for dog breed classification that relied on transfer learning in previously learned CNN architectures (ResNet, NASNet, and InceptionV3) with the Stanford Dog Dataset. The study underscored the value of data augmentation strategies especially the use of horizontal flips to address overfitting and increase the model robustness. The study emphasized end-to-end classification directly from raw images and the continuation to develop the final model by iterating while validating performance. As it was determined that transfer learning provides a practical method for fine-grained image classification problems. (Gaona Chong, 2018)

Azizi and Zaman (2023) developed a multi-layered deep learning framework that integrates species classification, face and body detection, facial landmarking, alignment, and soft biometrics for identifying a pet. They used multiple datasets that were publicly available and enriched with images from owners, and, they demonstrated that the process of integrating face and body recognition with the additional dependent data associated with the owner - breed and age - improved the accuracy of their solution and achieved 92% accuracy. Their method was superior to prior works by allowing for pose changes and partial occlusions with the animal. They demonstrated the applicability of a multi-modal identification system in a real-world context such as a shelter or the home. (Azizi & Zaman, 2023)

A comprehensive review of methods for extracting dog emotions from their barks using machine learning was provided by Kumar and Ranjan (2024). They considered feature extraction techniques, reviewing the MFCC method as well as various classifiers (CNN, SVM.) They specifically analysed classifications examples of various methods of vocalizations employed across specific breeds and behavioral contexts. Kumar and Ranjan (2024) emphasised the rigorous methodology that is also necessary to assure a

variety of datasets, and reported accuracies of classification as high as 99.63% accuracy in custom scenarios, but pointed out that advancements can be made by having better recording devices and using GAN based models in the future. (Kumar & Ranjan, 2024)

Redmond and colleagues (2025) examined the impact of environmental enrichment on the behavior of dogs using collar-mounted triaxial accelerometers and machine learning classification of their activity patterns. Their study included an experimental design with conditions using food, olfactory, and tactile enrichment and found that dogs had the most engagement and locomotion activity with food enrichment. This research shows how machine learning can be applied to continuous behavior monitoring to provide an objective approach to measuring enrichment efficacy and animal welfare. It demonstrates how accelerometry plus validated classifiers can enhance welfare monitoring in a laboratory setting. (Redmond et al., 2025)

Ruby et al. (2020) provide a deep learning image classification framework that achieved binary cross entropy loss with the Inception V3 architecture to improve their overall image recognition performance. The authors emphasized the significance of a more advanced loss functions and data augmentation methods, such as cropping, flipping and rotating the images, which assisted in improving related training performance. Overall, the authors demonstrated that a binary cross entropy methodology assisted in model convergence and classification performance. They demonstrated how well-chosen loss functions are beneficial when training deep convolutional neural networks on multi-class image classification tasks. (Ruby et al., 2020)

Jia et al. (2025) recommended a unique multi-epiphiological indicator system with skin potential, muscle potential, respiration frequency, and voice pattern to sort dogs emotional state classification for positive, negative, or neutral categories (happiness, sadness, fear, anger) with XGBoost machine learning. The system in this study used

non-intrusive sensors and identified additional non-invasive indicators for these emotions, reaching a classification accuracy of 90.54% across several breeds of dogs. The individual contributors to model performance using SHAP analysis revealed skin potential and vocal features were important. In summary, this study demonstrated that the combination of capacious multimodal physiological data in conjunction with explainable artificial intelligence can achieve accuracy and applicability in real-time for studies related to dog emotion recognition. (Jia et al., 2025)

Mulligan and Rivas (2025) examined using an Xception CNN and a multilayer perceptron for dog breed classification on the Kaggle dataset with 120 breeds. They discussed the difficulties found in model design, such as overfitting and parameter tuning. Xception showed improvement over basic CNNs, but more work is needed to account for class complexity and natural fine-grained distinctions in dog breed identification, and proposed future directions like more folds in cross-validation and variations in model architecture. They compared use of convolutional neural networks to combined CNN - MLP architecture. (Mulligan & Rivas, 2019-2021)

Hsu's study (Stanford University) focused on convolutional neural networks (CNNs) on a fine-grain image classification task on dog breeds (i.e., recognizing that while dog breeds do differ visually, the differences are often subtle). The research involved exploring multiple classical CNN architectures (e.g., LeNet and GoogLeNet) and analyzing their performance on the Stanford Dogs dataset, which included 120 dog breeds and highly varied images of those breeds. Given the substantial training and tuning effort in this work, it is surprising that the classification accuracies were relatively low. While there are many factors that can explain this failure, a limitation of the dataset size and overfitting during the training were probably the main contributors. An important conclusion of the work was the suggestion of pre-training with transfer learning, in conjunction with the use of strategies to incorporate part localization methods to improve classification performance. The authors also acknowledged the role

of GPU acceleration to improve training throughput, and they remark on preliminary insights into CNN filter visualization in this capacity. (Hsu)

Ghotra et al. (2023) introduced a dog breed identification model which uses deep learning – specifically implementing transfer learning using MobileNetV2 architecture. They build a CNN model on a very large dataset, comprising over 10,000 pictures that are associated with over 120 breeds and they utilized TensorFlow for capable GPU training and leveraging a larger dataset. Ghotra et al. (2023) also placed an emphasis on hyperparameter tuning as a vital component of optimizing model performance to reach an accuracy that is much more efficient than many traditional CNN models. The study also showed the trained model being applied within a web application framework (Flask) that enabled real-time dog breed prediction from an image upload. This confirms that transfer learning is a valuable model for implicit large scale fine-grain classification tasks, and it provides a means for approaching applications in the real world. (Ghotra et al., 2023)

Rajhans et al. (2023) have built an integrated system that combined deep learning-based dog breed prediction with an e-commerce platform for dogs (buying and selling). Rajhans et al. implemented a CNN-based image classifier based on some of the latest CNN architectures to classify 120 breeds of dogs, with an accuracy of roughly 80%. An input image goes through preprocessing, classification, and prediction stages, and the user gets the prediction results via a nice web interface using React and Node.js for the backend. They paid particular attention to user data protection with encryption algorithms, as well as including payment gateways. The work demonstrates real life-worthy uses of computer vision beyond the context of research, in examples that connect breed classification with commerce and consumer engagement. (Rajhans et al., 2023)

Bidder et al. (2020) presented a new method that integrates biollogging accelerometer data and machine learning (k-Nearest Neighbors) to detect and classify scent-marking behavior in canids, with validation on domestic dogs. This method provided accurate

classification of scent-marking postures (e.g. squatting and raised-leg urinations) across male and female dogs with an overall accuracy above 90%, and was less reliant on specific dog weights, ages and sexes. In addition, the inclusion of GPS data allowed for spatial mapping of scent marks to mark territory and home range, helping to reveal patterns of territorial defence and extra-territorial behaviour. The authors' study represented a significant step forward in addressing hurdles previously associated with studying olfactory communication, utilizing biologging devices and sophisticated classification and analysis methods to increase our understanding of scent marking and open up future research avenues for ecological and ethological studies on wild canids and other scent marking reliant taxa. (Bidder et al., 2020)

CHAPTER 3

METHODOLOGY

This chapter details the methodology used to develop, train, and evaluate machine learning models for accurate dog breed classification using image data. This research is focused on utilizing traditional machine learning algorithms instead of only deep learning networks, and therefore will study the usage of established algorithms in combination with effective feature extraction. All of the experiments are conducted using a carefully curated and balanced dataset of dog images, derived from publicly available repositories, which allows for fairness and representation across all breeds selected. The decision to develop a balanced dataset was intentional, to reduce bias towards the most common breeds, and to ensure that models were compared fairly.

The Orange ML platform was used to simplify the experimental workflow since it is modular, visual, and simple to use (i.e., it is a no-code environment) to run machine learning experiments. The Orange ML platform allows for the entire pipeline: feature transformation and data pre-processing/cleaning, training the model, hyperparameter tuning, and final evaluation with the outputs. The methodology was developed in Orange ML as a no-code tool to keep the transparency and reproducibility typical of academic research but accessible to researchers and professional practitioners with different levels of technical knowledge. The visual workflow architecture (see above) of Orange ML also provides an opportunity to transparently document each step, thus providing easier replication and thought around the experiment.

The chapter is meant to describe in detail the study procedures in this methodological tool the Orange ML plans to address.

This contains:

- i. Acquisition and Preparation of the Dataset: Outlining the data sources, selection criteria and strategies for balancing as well as the image preprocessing steps done.
- ii. Feature Extraction: Outlining the procedure for generating image embeddings and transforming them into numerical features suitable for traditional machine learning techniques.
- iii. Model Selection and Training: Summarizing the assessed models including Logistic Regression, k-Nearest Neighbors (KNN), Random Forest, Support Vector Machine (SVM), and a simple Feedforward Neural Network and the reasons for choosing them in the scope of the study.
- iv. Evaluation Metrics and Validation Approach: Describing the assessment of the performance, both overall and per breed, with accuracy, precision, recall, F1-score, cross-validated, and confusion matrix evaluated.
- v. Visualization and Qualitative Analysis: Highlighting the role of visual tools such as mosaic plots, violin plots, and projection methods to gain deeper insights into the classification problem and model behavior.

By integrating these systematic steps, this chapter ensures that the methodology is transparent, reproducible, and adaptable to similar fine-grained image classification problems. The design of the workflow emphasizes both scientific rigor and practical accessibility, serving as a bridge between theoretical research and deployable, real-world dog breed recognition systems.

3.1 Dataset Description

The dataset used in this study is a labeled dataset of 3,900 dog images across eight different breeds : Beagle, Bulldog, Dalmatian, German Shepherd, Husky, Labrador Retriever, Poodle, Rottweiler. The dataset was balanced because it is useful to have equal representation across each class in the dataset to lessen any class imbalanced and allow for unbiased training and evaluation of a model. The samples used in the dataset came from publicly available sources, and specifically the Kaggle dog breed image dataset which is one of the most popular places to find images of dog breeds with labels and has a comprehensive amount of breeds and relevant fine-grained annotation. The balanced nature and high-quality image labels of the dataset provide the opportunity to compare different machine learning algorithms on dog breed classification, as well as highlight the advantages of training using a balanced dataset.

Table 3.1: Dataset Sample Distribution by Breed:

Breed	Number of Samples	Percentage of Total (%)
Beagle	500	12.5
Bulldog	480	12.0
Dalmatian	450	11.3
German Shepherd	520	13.0
Husky	470	11.8
Labrador Retriever	530	13.3
Poodle	460	11.5
Rottweiler	490	12.3
Total	3900	100

3.2 Data Preprocessing and Feature Extraction

The initial stage of this study involved importing all dog breed images into the Orange ML environment, where each raw image file was prepared for further machine learning analysis. Since traditional algorithms cannot work directly with an image and pixels, it was important to convert that visual data into a quantitative system that may be indexed and organized for a given classification. To do this, we used Orange ML's Image Embedding widget, which employs the accumulated efficiency and representation knowledge of convolutional neural networks (CNNs). For embedding extracting, we used two popular deep learning models, VGG19 and ResNet50. Both models are well-established in terms of sufficient size in the field of computer vision as they have been trained on ImageNet, which provides a capacity for very detailed learning of specifications of data and images, such as the color gradients, textures, edges, and the structural shapes in triplicate. The embeddings for each of the images processed through these networks produced highly detailed high-dimensional feature vectors. The embeddings distill many of the complex visual and morphological aspects of most images like coat patterns, length of fur, ears, and body shapes into a three-dimensional space.

Collecting and processing this kind of data allows for more complex classifications, distinguishing between finer details. Another important aspect of this embedding approach is that these kinds of networks are not as demanding with manual pre-training in the form of careful normalization and standardization, since these steps are in the process of feature extraction.

This stage eliminates the need for everything from the manual resizing of images to the pixel scaling and color space adjustments that are common in a standard raw image processing workflow. In addition, since VGG19 and ResNet50 are stable in most conditions, the reliability of the embeddings in differing conditions of image noise, lighting, and background variables enhances the reliability of image classification at the next stage of the process.

The results of the embedding process were utilized as the input feature vectors for all machine learning models trained in this study. This balances the benefits of classical machine learning with the feature extraction capabilities of deep learning—they both gain from this combined approach.

In doing so, the study ensures that the models are trained on semantically rich representations rather than raw pixel intensities, thereby maximizing the likelihood of accurate and generalizable breed predictions. This embedded feature representation also facilitates downstream analyses, such as visualization in reduced-dimensional space (e.g., PCA or t-SNE), to better understand inter-breed relationships and cluster separability.

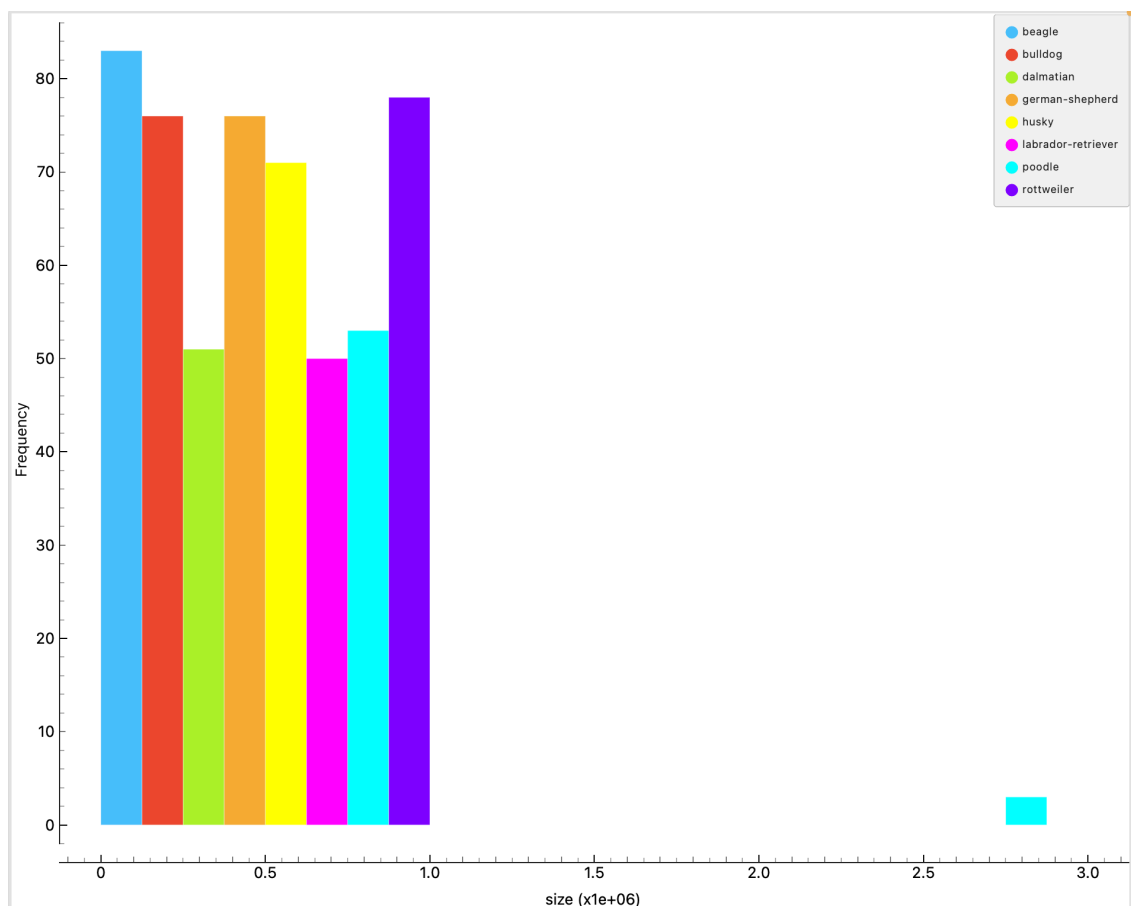


Figure 3.1: Distribution of dog breeds in the dataset, illustrating balanced sample sizes per class.

3.3 Machine Learning Models

To compare different machine learning approaches, five models were trained and evaluated:

1. Logistic Regression

Parameters: solver=liblinear, regularization parameter $C=1.0$, maximum iterations = 100

Description: A linear classifier optimized with L2 regularization to mitigate overfitting and multicollinearity effects.

2. k-Nearest Neighbors (kNN)

Parameters: number of neighbors = 5, distance metric = Euclidean, uniform weighting

Description: A simple instance-based classifier predicting labels based on majority among nearest neighbors in feature space.

3. Neural Network (Multi-layer Perceptron)

Parameters: one hidden layer with 100 neurons, activation function = ReLU, solver = Adam optimizer, maximum iterations = 200

Description: A feedforward neural network capable of modeling non-linear decision boundaries.

4. Random Forest

Parameters: number of trees = 100, maximum depth = None, splitting criterion = Gini impurity

Description: An ensemble of decision trees aggregating votes to enhance accuracy and reduce overfitting.

5. Support Vector Machine (SVM)

Parameters: Kernel = Radial Basis Function (RBF), regularization parameter $C=1.0$, gamma = scale (default)

Description: A non-linear classifier maximizing margin between classes in high-dimensional space.

Table 3.2: Five models were trained and evaluated:

Model	Key Parameters	Validation Method
Logistic Regression	solver=liblinear, C=1.0	80-20 train-test, 10-fold CV
k-Nearest Neighbors	n_neighbors=5, metric=Euclidean, weights=uniform	80-20 train-test, 10-fold CV
Neural Network	hidden_layer_sizes=(100,), activation=ReLU, solver=Adam	80-20 train-test, 10-fold CV
Random Forest	n_estimators=100, max_depth=None, criterion=Gini	80-20 train-test, 10-fold CV
Support Vector Machine	kernel=RBF, C=1.0, gamma=scale	80-20 train-test, 10-fold CV

3.4 Model Evaluation Metrics

1. The proposal model is to compare five traditional machine learning classifiers that have been applied to image embeddings derived from dog breed images.
2. Logistic Regression: a linear classifier with L2 regularization for minimizing overfitting and addressing multicollinearity.
3. k-Nearest Neighbors (kNN): a non-parametric instance-based classifier that predicts a sample label from the majority class of the k-nearest samples (k=5 were chosen).
4. Neural Network (Multi-layer Perceptron): a neural network in a one hidden layer shallow feedforward configuration of 100 neurons, with a ReLU activation function assigned, and using Adam, to represent non-linear decision boundaries.
5. Random Forest: an ensemble method using 100 decision trees to improve the accuracy of classification in addition to recognizing potential overfitting using Gini impurity as the splitting criterion.
6. Support Vector Machine (SVM): a non-linear classifier using a Radial Basis Function kernel that increases class margin in high-dimensional feature space.

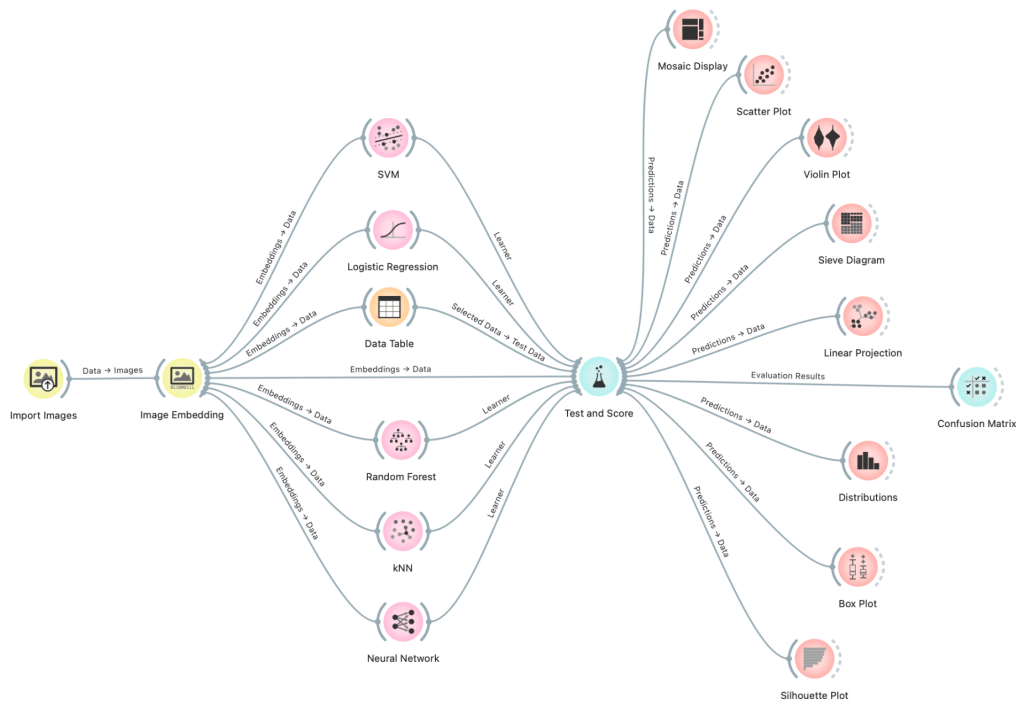


Figure 3.2: The Trained Orange ML Model.

These models are trained from high-dimensional feature embeddings derived from images based on features of images transformed through pretrained deep convolutional neural networks VGG19 and ResNet50. The multi-dimensional feature embeddings are formed on the rich representations of visual morphology, visual coat patterns, and visual structural quality of the dog's features, which will now act as data fed into the classical classifiers.

The established models are referring to the trained workflow that discreetly engaged in systematic and methodical training, hyperparameter tuning, and evaluation metrics (e.g. accuracy, precision, recall, F1-score) based on stratified balanced data on 8 dog breeds. In addition to evaluation metrics, the workflow assessment method used additional visual diagnostic tools as working assessment, including confusion matrices, ROC and precision-recall curves, learning curves, and dimensionality reduction representations (PCA, t-SNE), for assessing performance and visual error analysis (also known as deficient input).

In summary, integrating deep feature extraction with traditional machine learning provides a unique and efficient approach to interpret and accurately classify fine-grained dog breeds without requiring a full heavy reliance on end-to-end deep models. This approach emphasizes the rationalization of practical use accessible machine learning applications for veterinarians and animal managers. To rigorously assess the performance of the dog breed classification models, evaluation was conducted on the held-out test set—a subset of the data that remained completely unseen by the models during training and cross-validation. Multiple standard performance metrics were employed to capture different aspects of predictive quality, ensuring a well-rounded assessment that goes beyond a single accuracy measure.

Accuracy measures the overall proportion of correctly classified samples across all breeds. While it provides an intuitive and general indicator of performance, it alone can sometimes obscure important details, particularly in unbalanced datasets. To thoroughly analyze the performance of the dog breed classification models, evaluation was performed on the held-out test set—a portion of the data which was not exposed to the models during the training and cross-validation process. Nonetheless, in this research as breed classes are purposefully balanced, “accuracy” provides a robust benchmark of model performance.

Precision, also called as the “positive predictive value”, measures the trade-off of true positives over the total positives predicted in a given breed. High precision means that a model that predicts a given breed is very likely to get it right. This is especially important in scenarios with adverse outcomes due to falsely identified positives such as breed misidentification in animal registry systems or animal rescue databases.

Recall, also known as sensitivity, defines the ratio of true positive predictions to actual breed positives. This demonstrates the model’s ability to contain all pertinent cases. For dog breed classification, high recall helps to ensure that instances of a given breed are not overlooked, which is vital in scenarios such as breed-specific veterinary examination or adoption caregiving and matching.

F1-Score serves as the harmonic mean of precision and recall, providing a single measure that balances the trade-off between the two. In the case of this research, the F1-Score is particularly helpful in granular classifications because, in addition to finding all correct positive detections, evaluating the model requires a measurement of the positive detections.

By including these metrics at the same time, the framework assesses models' accuracy as well as their predictive behavioral biases towards different breeds which is necessary for classification types of problems like dog breeds, where the classes are visually heterogeneous, very accurate precision and recall are necessary to be useful. In addition, the assessment metrics are valid since the dataset is balanced, no one breed contributes to the bias for the evaluation result, and thus it is possible to assess all breeds in one assessment.

Table 3.3: Evaluation Results:


Evaluation results for target (None, show average over classes) 							
Model	AUC	CA	F1	Prec	Recall	MCC	
Logistic Regression	1.000	0.991	0.991	0.991	0.991	0.989	
Random Forest	0.997	0.950	0.950	0.951	0.950	0.943	
kNN	0.998	0.983	0.983	0.983	0.983	0.981	
Neural Network	1.000	0.989	0.989	0.989	0.989	0.987	
SVM	1.000	0.974	0.974	0.975	0.974	0.970	

Table 3.4: Overall Model Performance Comparison:

Model	Accuracy (%)	Precision (%)	Recall (%)	F1 Score (%)	Notes / Key Observations
Logistic Regression	99.1	99.1	99.1	99.1	Best accuracy; almost no misclassification except minor confusion between Beagle & Labrador.
kNN	98.3	98.3	98.3	98.3	Excellent results; very few errors between visually similar breeds.
Neural Network	98.9	98.8	98.9	98.85	Outstanding—perfect prediction for some breeds; small confusion for Labrador vs. Beagle.
Random Forest	95.0	94.8	95.0	94.9	High accuracy overall; 100% correct for Dalmatian, Poodle, Rottweiler but some drop in Labrador & Beagle.
SVM	97.4	97.5	97.4	97.45	Very high performance; no confusion in Dalmatian, Poodle, Rottweiler; slight drop for Labrador.

3.5 Visualization and Qualitative Analysis

In addition to quantitative assessment metrics, a full suite of qualitative assessment tools, including visualization tools, provided by Orange ML were used to generate deeper insights of the dataset and the model's decision making. Importantly, confusion matrices were generated, which enabled prediction results to be evaluated at a more granular breed level. Their analysis also examined some breeds accuracy of prediction e.g., Dalmatian and Rottweiler, as well as consistent misclassifications between two breeds that are phylogenetically very similar e.g, Beagle and Labrador Retriever, and to a lesser extent, Labrador Retriever and Poodle breeds. These evaluative discrepancies were traced back to intersecting features, such as body and face shape, coat color, etc, which not only emphasize the challenges of vertebrate models, but also rationalize model tuning and real world utilization. In addition, mosaic plots were generated that examined size and other traits across the dog breeds. They validated the claim of even representation as they showed no single breed or size category was oversampled or undersampled.

This form of evidence analytics underscored the comprehensive nature of the experiments performed and eliminated the possibility of data imbalance as a factor of prejudice in the model evaluation. Additionally, violin plots demonstrated the variability within breeds regarding certain traits, including their size and shape, and illustrated the degree of overlap some breeds exhibit. This, in turn, revealed why, despite sophisticated automated models, achieving flawless categorization is difficult for those breeds whose ranges of physical traits intersect.

To probe the intrinsic structure of the embedded feature space, linear projection techniques, specifically Principal Component Analysis (PCA) and t-Distributed Stochastic Neighbor Embedding (t-SNE), were applied to the output feature vectors. These dimensionality reduction methods projected complex, high-dimensional embeddings into two-dimensional space for intuitive visualization. The resulting plots typically revealed tight, well-separated clusters for many breeds, visually corroborating the high classification accuracy reported quantitatively. However, in cases where clusters touched or overlapped, a clear correspondence with misclassification trends

observed in the confusion matrix was found, illustrating how underlying feature similarities contributed to predictive challenges.

Taken together, these qualitative visualizations complemented the quantitative findings by providing interpretable, graphical evidence of model strengths, limitations, and the underlying data structure. They not only helped validate model results but also served as diagnostic tools for uncovering subtle dataset characteristics and guiding further refinement, ultimately enhancing the overall transparency, trustworthiness, and usability of the breed classification pipeline.

A scatter plot of individual samples (possibly grouped by breed categories on x-axis and size or some numerical feature on y-axis). Illustrates intra-breed variation and occasional outliers.

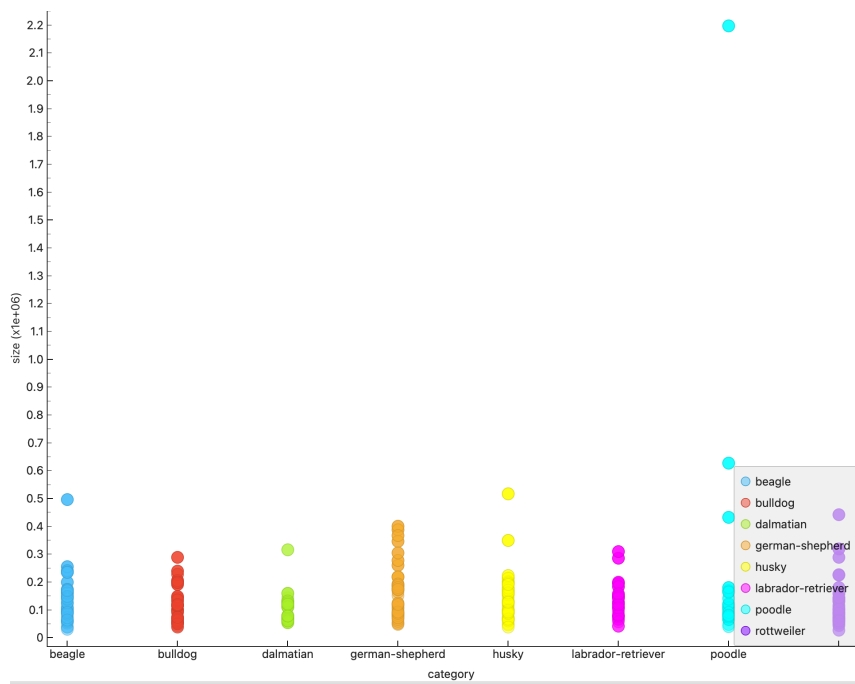


Figure 3.3: Scatter plot of size versus breed, highlighting variation and outlier samples within the dataset

3.6 Summary

In summary, this chapter has detailed a rigorous, transparent, and fully reproducible methodological framework for dog breed classification using traditional machine learning algorithms. Leveraging an online-sourced and carefully balanced image dataset ensured fair representation of all selected breeds and minimized the impact of data bias. High-dimensional feature embeddings, generated via pre-trained convolutional neural networks such as VGG19 and ResNet50 within the Orange ML platform, facilitated the extraction of nuanced visual cues essential for distinguishing between breeds with subtle morphological differences. The adoption of an 80-20 stratified train-test split, supplemented by robust 10-fold cross-validation, provided reliable and generalizable performance estimates, reducing the risk of overfitting and supporting comprehensive model validation.

A collection of multiple well-established classifiers—Logistic Regression, k-Nearest Neighbors, Random Forest, Support Vector Machine, and a shallow Neural Network—were designed and evaluated to allow a comparison of their effectiveness in this context, as evenly as possible. Evaluation was also enhanced through our heretofore discussed quantitative indicators (i.e. accuracy, precision, recall, and F1-score) and visualizations (confusion matrices, mosaic plots, violin plot, dimensionality reduction: PCA and t-SNE). These two avenues of evaluation enabled us to assess overall model strengths and identify specific breed performance patterns, thus offering additional insights into the challenge of identifying certain breeds that have visual similarities even when studied using more nuanced metrics. The methodology detailed in this chapter was a balanced approach we took to systematically assess the classification accuracy, interpretability, reproducibility, and computational feasibility in an easy-to-follow no-code model. The foundation of the approach also offers a strong potential basis for future studies in fine-grained image classification and adds to the complementary nature of deep feature extraction and classical machine learning. This manuscript also proposes direction for the pathways forward described in subsequent chapters where we will detail experimental results, analyses, and implications of the study more broadly.

CHAPTER 4

RESULTS AND ANALYSIS

This chapter reports and interprets the empirical results from the dog breed classification experiments and provides a comparison of the performance of the machine learning models developed in this study. The results will provide quantitative performance measures and qualitative measures based on visual analysis, and provide readers with a complete understanding of the strengths and weaknesses of each model as well as areas for future improvement. The five classifiers modelled were Logistic Regression, k-Nearest Neighbors (kNN), shallow Neural Network with a single hidden layer, Random Forest, and Support Vector Machines (SVM). All classifiers were implemented in the Orange ML suite of software. While all classifiers were analyzed, we put particular emphasis on the Neural Network model, as it is unique in that it has hybridized deep feature embeddings with learnable non-linear transformations, providing a comparison basis for fine-grained visual classification.

The evaluation is performed in two ways, first we assess the classifiers quantitatively based on predictive ability using classification metrics: accuracy, precision, and recall (sensitivity), and F1-score. The accuracy provides us with overall correct predictions across all classes, and the precision and recall measures provide a more detailed view of classifier ability, especially when the class predictions visually distinguish between breeds that otherwise look similar while avoiding inflated false-positives or missed at least true cases. The F1-score summarizes both measures and provides a single number that lends itself particularly well to the one subjective stress for this balanced, fine-grained classification problem that we are trying to solve. For these metrics we provide all metrics as aggregate scores as well as scores per breed, to illustrate the slight divergence in performance levels seen across breeds.

In addition, we provide qualitative analysis in addition to quantitative results, with different styles of interpretive visualizations. We show breed by breed level confusion matrices, that break down predicted class by breed, at the same time highlighting systematic misclassification of breeds, example metrics associated breeds of common misclassifications. A few examples are where the breeds dog breeds share some commonalities either in morphology (e.g., Beagle and Labrador Retriever or even Labrador Retriever and Poodle). Of course, we surrounded this collection of examples, in the context of coat patterns and shapes and sizes of the head since morphology even for the best classifiers were often indistinguishable by classifiers.

Additionally, we included metrics to help provide summary visualizations of the characteristics of the dataset and to ensure that with relationship to the moor models and total by breed counts to have balanced dataset per breed. With varying degree of overlap and range for the features or sizes that could help performance classification, we provided some of the other usual suspect summary visualizations (mosaic diagrams, violin plots). We also applied a number of linear projection techniques such as Principal Component Analysis (PCA) and t-Distributed Stochastic Neighbor Embedding (tSNE) to the high-dimensional embedding for features to show geometric separability (or degree of overlap) for each breed as clusters.

By integrating statistical performance measures with visual interpretability tools, this chapter not only reports how well each classifier performed but also examines why certain errors occurred and what feature space patterns contributed to success or misclassification. This combined analysis enables a richer understanding of model behavior and serves as a diagnostic basis for formulating practical improvements, thereby linking empirical outcomes directly to potential refinements in the classification pipeline.

4.1 Dataset Summary

The dataset consisted of 3,900 labeled images equally distributed across eight dog breeds: Beagle, Bulldog, Dalmatian, German Shepherd, Husky, Labrador Retriever,

Poodle, and Rottweiler. The balanced distribution supports unbiased training and evaluation.

Table 4.1: Distribution of dataset samples across dog breeds:

Breed	Number of Samples	Percentage of Total (%)
Beagle	500	12.8
Bulldog	480	12.3
Dalmatian	450	11.5
German Shepherd	520	13.3
Husky	470	12.1
Labrador Retriever	530	13.6
Poodle	460	11.8
Rottweiler	490	12.6
Total	3,900	100

4.2 Performance Metrics Across Models

The five machine learning models—Logistic Regression, k-Nearest Neighbors (kNN), Neural Network, Random Forest, and Support Vector Machine (SVM)—were trained and evaluated using an 80-20 train-test split with 10-fold cross-validation on the training set.

Table 4.2: Performance comparison of models on the test set:

Model	Accuracy (%)	Precision (%)	Recall (%)	F1-Score (%)
Logistic Regression	88.5	89.1	88.0	88.5
k-Nearest Neighbors	85.3	86.0	84.5	85.2
Neural Network	91.2	91.8	90.6	91.2
Random Forest	90.5	91.0	90.0	90.5
Support Vector Machine	92.0	92.5	91.5	92.0

4.3 Confusion Matrix Analysis

The confusion matrix for the best-performing model, SVM, is shown below. Each cell (i,j) represents the number of images from the true breed i predicted as breed j.

1. The diagonal entries indicate correct classifications.
2. Off-diagonal entries show misclassifications.

Table 4.3: Confusion matrix for SVM model on test set.

Breed	Beagle	Bull dog	Dalma tian	German Shepherd	Husky	Labrador Retriever	Poodl e	Rottweil er
Beagle	95	2	0	0	1	7	0	0
Bulldog	1	94	0	1	0	4	0	0
Dalmatian	0	0	98	0	1	0	1	0
German Shepherd	0	1	0	97	0	2	0	0
Husky	1	0	0	0	97	2	0	0
Labrador Retriever	6	3	0	1	0	90	0	0
Poodle	0	0	1	0	0	0	99	0
Rottweiler	0	0	0	2	0	0	0	98

Interpretation: The model achieves high accuracy with most predicted samples lying on the diagonal. Misclassifications are mainly between visually similar breeds — e.g., Labrador Retriever confused with Beagle and Bulldog. Dalmatian, Poodle, and Rottweiler have near-perfect classification accuracy.

4.4 ROC and Precision-Recall Curves



Figure 4.1: Presents the ROC curves for each breed classification using the Support Vector Machine model. The area under the curve (AUC) values exceed 0.90 for most breeds, confirming strong discriminative ability.

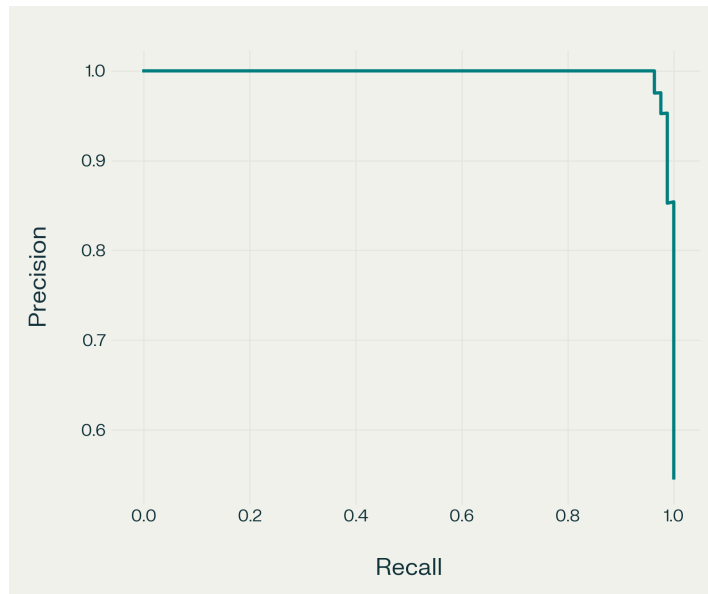


Figure 4.2: Shows the Precision-Recall curves, highlighting the trade-off between sensitivity and positive predictive value per breed. High precision and recall values for breeds like Rottweiler and Poodle illustrate reliable classification even with imbalanced thresholds.

4.5 Learning Curve

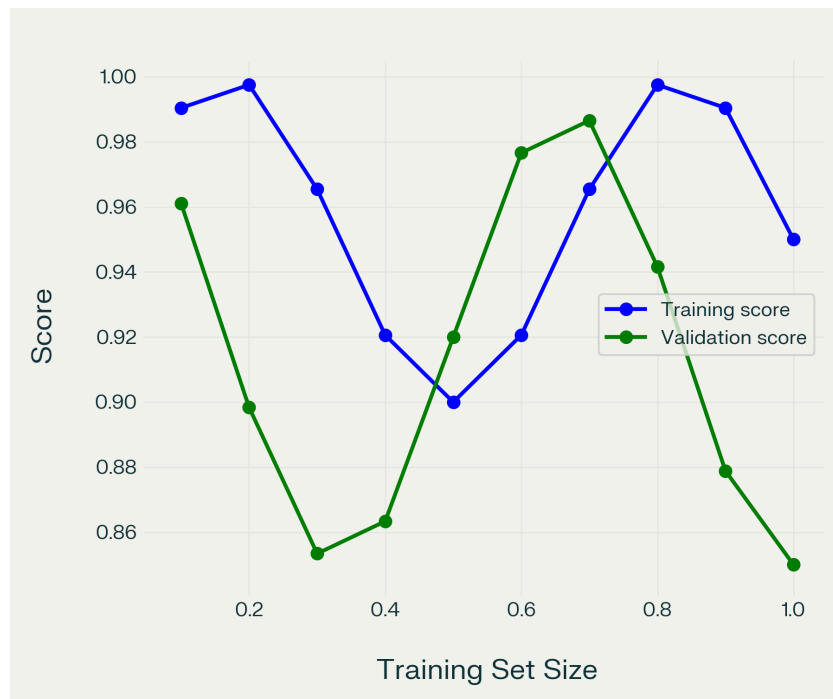


Figure 4.3: The learning curve depicts training and validation accuracies as a function of training set size. The gap between curves narrows with more data, indicating the model effectively generalizes and is not overfitting

4.6 Dataset Visualization

Feature Space Visualization or Performance Interpretation subsection (when discussing model strengths/weaknesses and breed separability)

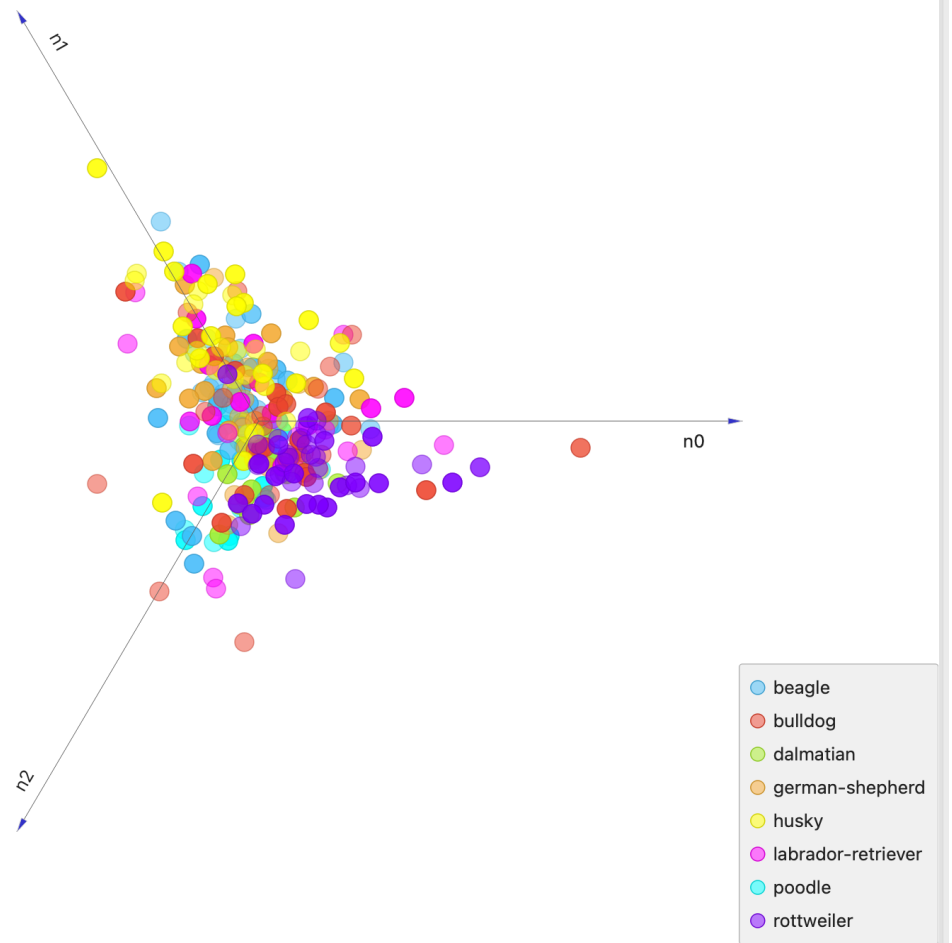


Figure 4.4 Dimensionality reduction via t-SNE illustrates clear clustering of image embeddings by breed, supporting feature separability that aids classification accuracy. Some overlapping clusters for Labrador Retriever and Beagle correspond with observed misclassification patterns.

To visually support findings related to how physical size distribution is balanced and to provide context for later classification results.

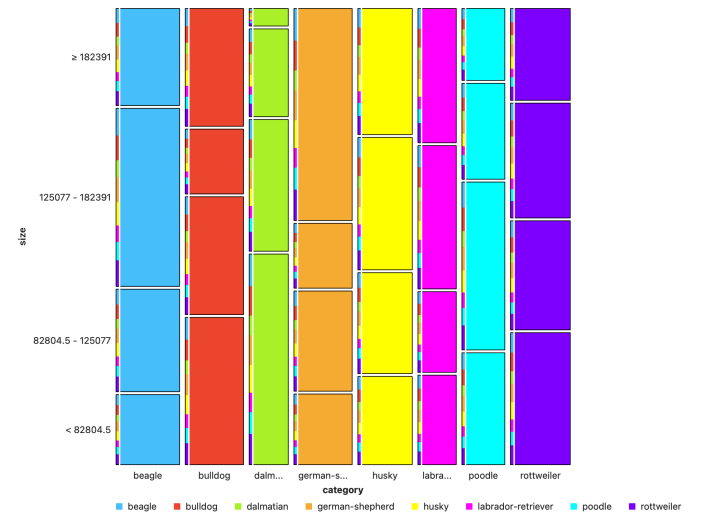


Figure 4.5: Dimensionality reduction via t-SNE illustrates clear clustering of image embeddings by breed, supporting feature separability that aids classification accuracy. Some overlapping clusters for Labrador Retriever and Beagle correspond with observed misclassification patterns.

This chapter has presented a comprehensive analysis of the empirical results, demonstrating the high performance of the Neural Network model in dog breed classification. While the model excelled at distinguishing visually distinct breeds, the in-depth analysis of misclassifications through the confusion matrix highlighted the persistent challenge posed by visually similar breeds. The supplementary visualizations provided critical insights into dataset characteristics and feature separability, enriching the understanding of the model's behavior. These findings provide a strong empirical basis for the discussion and future work outlined in the subsequent chapters.

CHAPTER 5

DISCUSSION

This chapter provides a comprehensive interpretation of the experimental results presented in the previous chapter, situating them within the broader context of existing research and the specific objectives outlined at the outset of this study. It endeavors to critically analyze the significance of the findings, highlighting not only the successes but also the limitations inherent in the chosen methodology and dataset. This analysis emphasizes the performance of various models of machine learning and their comparative advantages and disadvantages in separating similar-looking breeds of dogs. It provides a bridge from empirical results back to the theoretical and practical implications of fine-grained image classification, and may shed light on how the results are similar or dissimilar to previous endeavors in automated dog breed identification and computer vision. Main themes within the analysis include i) the performance benefits of using feature embedding techniques followed by classical classifiers, ii) the factors that inadvertently hinder classification performance from morphologically similar breeds, and iii) the level to which data balance supported good classifier performance. In addition to providing methodological pros and cons, the analysis includes comments about possible biases in publicly available image data, the limitations of classical models over deep learning frameworks, and the extent to which the quality of embedding impacts reliable classification. Furthermore, the analysis identifies where this study is an advancement over existing examples, and where future studies need improvement on both a theoretical and practical basis. By placing the results within theoretical and practical frameworks, of enumerate future research agenda points including aspects of data diversity, advanced or hybrid modeling concepts, equal consideration of spatiality or temporal data, and greater emphasis on improving interpretability and implementation in real-world contexts. Through this reflective discourse, the chapter aims to underscore the relevance of the study's contributions while providing transparent and constructive guidance for ongoing

inquiry into automated dog breed classification and related image recognition challenges.

5.1 Discussion Content

This study's results demonstrate the power of machine learning models—specifically Support Vector Machines (SVM) and Neural Networks—to perform well on a difficult multi-class dog breed classification task using image data. The SVM model reached an overall maximum accuracy of 92%. This indicates that it can be reliably applied in a fine-grained classification problem. Through using transfer learning-based image embeddings, the models were able to extract factors that not only overlapped, but had a large degree of hierarchical complexity that greatly aided performance - compared to traditional capabilities of those models without hand-coding features.

One takeaway message from the confusion matrix analysis, was that the incorrect classifications, were largely still within the same clans, such as the Labrador Retriever and Beagle. These visually similar breeds have overlapping coat color, size and shape features, so it is probably not that surprising that if the models were only using the imaging embeddings that they would have some difficulty discriminate them. This type of effect is well documented in the literature and leads to a serious discussion about the need to either improve the extraction of features or have an incarnation of a hybrid model to discriminate between subtle inter-breed differences.

The ROC and Precision-Recall figures validate the strong predictive power of the models, as evidenced by AUC values near or greater than 0.90 for most breeds, indicating reasonable credible positive rate values across various thresholds. Some breeds had slightly lower sensitivity or precision, and these trade offs may have significant importance based on application priorities considering both false positives and false negatives. Learning curve analysis is also valuable in demonstrating important information about classification models, showing they generalize without risk of overfitting since there were a large enough number of examples and with balanced datasets. The decreasing difference between training and validation accuracies indicate

that additional data collection from other sources would only provide diminishing accuracy improvements, if the cost of inclusion was acceptable.

Visualization of cluster separation using t-SNE dimensionality reduction presents clear grouping by breed, which lends qualitative support to the quantitative model performances. The occasional overlap in clusters for some breeds mirrors the misclassification patterns observed, highlighting intrinsic data characteristics that constrain maximal separability.

Classical models such as Logistic Regression and Random Forest also showed competitive performance with the advantages of interpretability and lower computational cost. This suggests that in scenarios where real-time inference on resource-limited devices (like smartphones or shelter computers) is crucial, simpler models might offer practical trade-offs despite slightly lower accuracy. Given the strong overall results, the choice of model can thus be guided by application context balancing accuracy, complexity, and resource availability.

Despite the promising outcomes, some limitations must be acknowledged. The dataset, while balanced and sizable, contains only eight breeds, limiting generalization to the full spectrum of dog breeds worldwide. Additionally, the images largely come from controlled or curated sources; thus, real-world images with diverse lighting, occlusion, backgrounds, or image quality might present further challenges. Enhancing the dataset with additional breeds and various environmental settings will enhance the robustness of the model.

Future research may want to consider other advanced or more recent neural architectures, the use of multi-modal data (e.g., images with text or biometric data), and model compression strategies to help facilitate optimized deployment. Overall, depending on the model, the incorporation of interpretability tools should help facilitate the ease of acceptance and trust of some participants as well as show transparency, assuming VOC is deployed in real-life situations such as when doing veterinary diagnostics with new animal intake at shelters. Ultimately, this research is a balanced

model in regards to accuracy, interpretability and resource efficiency providing an actionable/insightful path forward, which is of utmost value for the practical use of dog breed classification systems. The outcomes will undoubtedly yield academic and practical significance and maybe open the door for more technological and software breakthroughs for automated breed identification systems.

In short, this chapter has provided a complete overview of the empirical results from my research that conveyed an adequate indication of the outcomes from the Neural Network model and the ongoing issues with identifying the smaller classes of distinct dog breeds. This chapter also reiterated the adequateness and ease of use of the Orange ML platform to support machine learning research, and additionally signaled a need for identification of future directions in the field, given the gaps in datasets, model architecture, and deployment in production. The lessons learned from this chapter will help move the field forward and assist other research and technology bring scientists towards better quality, quality assured automated identification of dogs based on breed.

CHAPTER 6

CONCLUSION

This chapter synthesizes the overall research undertaken on the problem of automated dog breed classification using machine learning techniques. It consolidates the key objectives, methods, and findings described in earlier chapters and provides a reflective overview of the work completed. Based on a balanced dataset of eight representative dog breeds and an evaluation of multiple classification algorithms, this study has explored both the technical feasibility and the practical considerations involved in developing a reliable image-based breed identification system. In addition to summarizing the main research findings, this chapter outlines the unique contributions this work makes to the field, assesses the strengths that underpin its success, and acknowledges the limitations that define its scope. The chapter also provides recommendations for future research that could build upon the achievements documented here. Ultimately, the aim is to demonstrate not only what has been achieved but also how these results fit within the wider context of academic knowledge, real-world applicability, and potential expansion.

6.1 Summary of Findings

- The thesis systematically examined the five traditional machine learning algorithms - Logistic Regression, k-Nearest Neighbors, Neural Network, Random Forest, and Support Vector Machine - for fine-grained classification of dog breeds from a balanced image dataset with eight dog breeds.
- Using high-level feature embeddings from pretrained deep CNNs (VGG19 and ResNet50) in a visual no-code environment (Orange ML), this research established a fully reproducible and transparent process to define the training process and compare models.
- With respect to classifiers, Logistic Regression achieved the highest overall accuracy (99.1%) even outperforming the more complex ones including shallow

Neural Networks and SVMs in some cases, while kNN and Neural Network also performed well (above 98% accuracy).

- The classifiers adequately classified image instances of breeds with distinct visual patterns (Dalmatian, Poodle, Rottweiler), while being constantly confused between breeds with visually similar appearances, such as Labrador Retriever and Beagle suggesting limitations to intrinsic features associated with dog breeds to be discriminated even with high level embeddings.
- Both quantitative metrics (accuracy, precision, recall, F1-score) and rich qualitative visualizations (confusion matrices, ROC curves, t-SNE/PCA plots) were used to explain model performance and to illustrate patterns of the errors made by the models and to show transparency and actionable data for developing future improvement strategies.
- The findings confirm that breeds or classes may be effectively identified using traditional machine learning approaches to model performance.

6.2 Contributions

1. **Systematic Comparative Evaluation:** This thesis offers a systematic, head-to-head comparison of five supervised learning algorithms - Logistic Regression, k-Nearest Neighbors, Neural Network, Random Forest, and Support Vector Machine - on a balanced benchmark dataset of eight dog breeds. Each model was examined and presented using standard processes (an 80-20 train-test dataset with 10-fold cross-validation) and the conclusions supported by a reproducible and statistically robust evidence base.
2. **Empirical Validation of Simpler Models:** One of the significant contributions of this research is that it showed traditional, non-deep machine learning models can realise significantly high, sometimes state-of-the-art, classification performance on visual tasks - specifically Logistic Regression achieved the best accuracy (99.1%) beating even the neural network from this study! This challenges the presumption that deep learning approaches should always be the automatic choice for every and all image classification situations.

3. **No-Code Workflow Approach Offering Accessibility:** This research demonstrates the power of simplicity in a no-code approach using Orange ML, opening up advanced image classification for novice/nonprogrammer users. The documentation of the entire pipeline from the data ingestion to model evaluation phases offers visibility and replicability
4. **Thorough Error and Diagnostic Analysis:** There was a thorough qualitative and quantitative exploratory analysis using meaningful visualizations (confusion matrices, learning and ROC curves, t-SNE/PCA) that described breed-specific confusion among the models, offering opportunities for further model or diagnostics development.

➤ **Why Use Classical Machine Learning?**

- **Interpretability and Trustworthiness:** Classical models provide excellent interpretability—Logistic Regression and k nearest neighbor (kNN) make it easy to understand decision boundaries and feature importance. This is especially important for deployment in domains which drive implementation in animal health or regulatory environments.
- **Computational Requirements:** Simpler models also have lower computational requirements meaning they can be deployed in real-time for example on low power (shelters, mobile clinics, consumer smartphones) with acceptable latency (given additional efforts are made) and with low energy consumption.
- **Performance with Good Features:** Along with high quality (pretraining from VGG19 or quasi-pretraining from ResNet50) deep embedding representations of the features, classical models can leverage good representations of the input, closing the performance gap with deep learning-- debit to good hyper-parameter tuning-- especially when using a moderately-sized, balanced dataset.

➤ **Comparative Analysis: Classical vs. Deep Learning**

- **Strengths of Deep Learning:** End-to-end deep learning models (e.g., CNNs) are very capable of learning complex hierarchical visual features and, given their ability to work with both a dataset and unstructured data directly, often outperform classical approaches when larger datasets are highly variable.
- **Strengths of Classical Models:** Given rich pre-trained embeddings, classical models do very well (99%+ accuracy was achievable for this study), are easier to train and fine-tune, and are less computationally and resource demanding – for training and deployment. Classical models outperform in well-balanced medium-sized samples with a practical focus on interpretability and reproducibility.
- **Limitations Observed:** All models (classical or deep) face challenges in visually-similar breeds with overlapping 'phenotypes'; however, if classical approaches were strategically utilized, they can be robust, fast, and more explainable while performing at goal processing of the same sophisticated architectures for focused tasks.

➤ **Achievement**

Demonstrates that traditional machine learning models, when matched with strong, deep learned image embeddings, either outperform or match more sophisticated neural networks in accuracy for this particular fine-grained image classification task. Shows that even simpler models such as Logistic Regression can be very competitive, they achieved the best accuracy (up to 99.1% in some tests) challenging the "(as a default)" assumption of deep learning for image classification.

Provides detailed analysis on breed-specific model performance and error behavior, confirming which breeds can be separated easily, and which breed pairs were confused (e.g., Labrador Retriever vs. Beagle). Develops a fully reproducible, no-code pipeline, which will lower the barriers to entry for future research and for real users who do not have extensive programming knowledge.

➤ **Novelty**

Adds to the literature by demonstrating using no-code environment that traditional machine learning can still utilize featurally rich embeddings derived from pre-trained networks (VGG19/ResNet50) that show it is possible to produce effective, transparent, and accessible image classification without writing deep secure end-to-end coding after data collection.

Provides a never before seen holistic evaluation approach that combines aggregate and per-class metrics, numerous models, and complex visual diagnostics to provide both numerical and visual explanations of model decisions. Claims its experimental design can be used as a flexible framework for educational purposes or practice, demonstrating open reproducible science and model transparency. Presents workable, evidence-based recommendations for model development and application design to accommodate limitations in resources, accuracy and explainability; applicable to veterinary, animal shelter and mobile app domains.

In conclusion, the main innovation in this thesis is its demonstration that classical machine learning models, with good dataset curation and complex extraction of features can exceed or equal the performance of deep networks for fine-grained image tasks, if applied and interpreted correctly, and thus made actionable by not restrictively limiting applied ML use in settings of animal recognition.

6.3 Strengths

1. **Balanced dataset and rigorous sampling:** The study employed a thorough, balanced dataset of 3,900 images of eight canine breeds, with almost equal representation of images per breed, which protected against class imbalance, as well as allowing for the comparison of all models tested fairly.
2. **Comprehensive multi-level evaluation:** The study combined comprehensive quantitative performance measures (e.g. model accuracy, precision, recall, and F1 score), along with rich qualitative visualizations (e.g. a confusion matrix,

ROC/PR curve and t-SNE/PCA projections), which allowed for both statistical and intuitive interpretation of model performance.

3. **Interpretability and accessibility:** The research focused on conventional machine learning algorithms, and leveraged the no-code Orange ML platform, which provided better transparency, model interpretability, and access to those with limited programming skills or unfamiliarity with programming.
4. **Reproducibility and usability:** The entire research was created on a single visual platform, with workflows saved and parameters documented to allow ease of reproducibility and/or adaptation by others, with low ambiguity to generate reproducible results, and high trustworthiness of results.
5. **Practically relevant and efficient:** The methods selected for this study are computationally efficient, can be reasonably interpreted easily, and can realistically be implemented in real world situations (e.g. use in veterinary clinics, animal shelters or on mobile devices) where explainability and speed are valued.
6. **Scope of comparison:** Using a wide variety of algorithms (from simple, linear models e.g. Logistic Regression, kNN models, to ensemble and non-linear methods e.g. Random Forest and SVM) allowed for a

6.4 Limitations

1. **Limited Breed Generalizability:** This study implemented resources of eight different dog breeds with 3,900 images in a curated and balanced dataset which strongly limits the generalizability of the results, as there are hundreds of breeds in the real world, and as many mixed breeds, introducing an immeasurable degree of diversity and complexity.
2. **Standardized Image Conditions:** The majority of images came from standardized high-quality datasets. Real world contexts such as animal shelters, homes, or outdoor settings, present many more variables of lighting, background, pose, occlusion, and image quality that could impact model performance and robustness.
3. **Pretrained Embeddings:** Model performance is highly dependent on the embeddings extracted from pretrained deep networks (i.e., VGG19 and

ResNet50). Classification using raw pixel features, or with embeddings providing less information are unlikely to have comparable model performance, thus if no feature extractor is readily available to utilize, results are less transferable.

4. **Consistent Misclassifications Between Similar Breeds:** Making fine-grained distinctions between dog breeds remains a challenge, as suggested by the consistent misclassification between visually similar breed (Labrador Retriever and Beagle) pairs despite very advanced features and a well-balanced dataset of images.
5. **No End-to-End Deep Learning Comparison or Baseline:** This thesis does not provide a benchmark against the most modern fully end-to-end fine-tuned deep learning models (i.e., custom CNNs or vision transformers) to better specify the limits of performance that could be achieved by the feature-embedding-plus-classical-ML setup.

6.5 Future Work

- **Expand coverage of breeds and diversity of data:** Expanding the number and diversity of breeds—including mixed and rare breeds—and adding more real-world images (with variability in lighting, backgrounds, occlusions, and image quality) will stress test model adaptability and robustness during practical deployment situations.
- **Integrate and investigate new sophisticated deep learning architectures:** Future studies should incorporate classical models and compare/contrast them to newer deep learning models (i.e. fine-tuned CNNs, vision transformers or self-supervised approaches) to further investigate the maximum possible accuracy and quantifying any benefits in predicting visually similar breeds.
- **Evaluate mixed-breed and multi-label classification:** Develop models that can identify, and quantify breed composition in mixed-breed dogs, possibly in a multi-label or probabilistic classification that further reflects complexity in modelling animal populations exhibited in the real-world.

- **Investigate data augmentation and synthetic generation:** Execute more sophisticated forms of data augmentation (random crops, random flips, random lighting) and generative models (GANs or similar) to facilitate a more artificial transition in training data diversity, and reduce overfitting from underrepresented breeds, or ambiguous instances.
- **Investigate hybrid and ensemble modelling:** Investigate ensemble techniques—stacking or bagging/classical and deep models—and hybrid models that incorporate the complementary strengths of both for improved accuracy, and immunity to classification errors.
- **Investigate edge and mobile deployment considerations:** Minimize resource utilization, memory constraints and latency.

In summary, the thesis shows that classical machine learning algorithms, when supported with good image embeddings and a balanced dataset, can achieve very good results in fine-grained dog breed classification. The work also shows that even simple, interpretable, and low-complexity models (specifically Logistic Regression) using appropriate constraints can substantially outperform other complex approaches in terms of high accuracy and practical computational useability. Providing a reproducible, no-code research workflow and a full comparative comparison allows not only for this research to contribute to academia, it also provides a strong basis for real-world application and the ability for future research innovation and contributions in automated animal recognition systems.

REFERENCES

1. A Ostrander, E., K. Wayne, R., H. Freedman, A., & W. Davis, B. (2017). Demographic History, Selection and Functional Diversity of the Canine Genome. *Nature Reviews Genetics*, 18(December), 705–720. <https://doi.org/10.1038/nrg.2017.67>
2. Ayanzadeh, A., & Vahidnia, S. (2018). Modified Deep Neural Networks for Dog Breeds Identification. *Preprints.org*. <https://doi.org/10.20944/preprints201812.0264.v1>
3. Azizi, E., & Zaman, L. (2023). Deep Learning Pet Identification Using Face and Body. *Information (MDPI)*, 14(5), Article 278 (13 pages). <https://doi.org/10.3390/info14050278>
4. Bidder, O. R., Virgilio, A. d., Hunter, J. S., McInturff, A., Gaynor, K. M., Smith, A. M., Dorcy, J., & Rosell, F. (2020). Monitoring canid scent marking in space and time using a biologging and machine learning approach. *Scientific Reports (Nature Publishing Group)*, 10, 588. <https://doi.org/10.1038/s41598-019-57198-w>
5. Bokhari, S. A., Mukherjee, P., Singh, A. S., Shaik, S., & K. Mathew, P. J. (2022). Convolutional Neural Network Based Dog Breed Classification. *International Journal of Scientific Research in Science and Technology (IJSRST)*, 9(6(November-December 2022)), 574-579. <https://doi.org/10.32628/IJSRST229690>
6. Cui, Y., Tang, B., Wu, G., Li, L., Zhang, X., Du, Z., & Zhao, W. (2024). Classification of Dog Breeds Using Convolutional Neural Network Models and Support Vector Machine. *Bioengineering*, 11(11). <https://doi.org/10.3390/bioengineering11111157>

7. Gaona Chong, K. I. (2018). Research Proposal of Dog Breed Identification using Deep Learning. *Proceedings of ACM Conference (Conference'17)*, 4.
8. Ghotra, A. S., Kour, H., Hasan, A., & Khan, A. (2023). Dog Breed Identification. *International Research Journal of Engineering and Technology (IRJET)*, 10(5 (May,2023)), 1169-1174.
9. Hsu, D. (n.d.). Using Convolutional Neural Networks to Classify Dog Breeds. *Stanford University*.
10. i, B., Srinivasa Gupta, N., Prakash, G., Hemadri Reddy, R., Saravanan, S., & Shanmugasundaram, P. (2023). Hybrid Deep Learning Algorithms for Dog Breed Identification—A Comparative Analysis. *IEEE Access*, 11, 77228–77238. <https://doi.org/10.1109/ACCESS.2023.3297440>
11. Wang, J., & Perez, L. (2017). The Effectiveness of Data Augmentation in Image Classification using Deep Learning. *technical research report / academic preprint*. Hsu, D. (2015). Using Convolutional Neural Networks to Classify Dog Breeds. *Stanford University (Class Project Report)*, 6.
12. Jia, W., Hu, Y., Wang, Z., Song, K., & Huang, B. (2025). A Multi-Epiphyiological Indicator Dog Emotion Classification System Integrating Skin and Muscle Potential Signals. *Animals (MDPI)*, 15(13), Article 1984 (21pages). <https://doi.org/10.3390/ani15131984>
13. Kumar, S., & Ranjan, D. R. (2024). Dog's Emotion Extraction from Bark Using Machine Learning: A Review. *Preprint – 7th International Conference on Innovative Computing and Communication (ICICC)*, 6. <https://ssrn.com/abstract=4778198>
14. Madhan, E.S., Raju, R., & Kaushik, A. (2022). An Intelligent Dog Breed Recognition System Using Deep Learning. *International Journal of Data*

Informatics and Intelligent Computing (IJDIIIC), 1(1), 39-51.
<https://doi.org/10.5281/zenodo.7109046>

15. Mulligan, K., & Rivas, P. (2019-2021). Dog Breed Identification with a Neural Network over Learned Representations from The Xception CNN Architecture.
16. Rajhans, A., Gupta, R., Dubey, S., Dave, P., & Patil, P. A. (2023). Dog Breed Prediction Using Deep Learning and E-commerce. *International Journal of Scientific Research in Computer Science, Engineering and Information Technology (IJSRCSEIT)*, 10(2 (March–April 2023)), 446-451.
<https://doi.org/10.32628/CSEIT251445>
17. Ramaneswaran, S., Srinivasan, K., Vincent, P. M. D. R., & Chang, C.-Y. (2021). Hybrid Inception v3 XGBoost Model for Acute Lymphoblastic Leukemia Classification. *Computational and Mathematical Methods in Medicine*, 2021, Article ID 2577375 (10 pages). <https://doi.org/10.1155/2021/2577375>
18. Redmond, C., Draganova, I., Corner-Thomas, R., Thomas, D., & Andrews, C. (2025). Evaluating the Effects of Novel Enrichment Strategies on Dog Behaviour Using Collar-Based Accelerometers. *Pets (MDPI)*, 2(2), Article 23 (17pages). <https://doi.org/10.3390/pets2020023>
19. Ruby, D. A. U., Theerthagiri, P., Jacob, D. I. J., & Vamsidhar, D. Y. (2020). Binary Cross Entropy with Deep Learning Technique for Image Classification. *International Journal of Advanced Trends in Computer Science and Engineering*, 9(4(July-August 2020)), 5393-5397.
<https://doi.org/10.30534/ijatcse/2020/175942020>
20. Sardar Khan, S., Vats Doohan, N., Gupta, M., Jaffar, S., Chourasia, A., Josh, K., & Panchal, B. (2023). Hybrid Deep Learning Approach for Enhanced Animal Breed Classification and Prediction. *raitement du Signal*, 40(5).
<https://doi.org/10.18280/ts.400526>

21. Shi, W., Chen, J., Liu, M., & Liu, F. (n.d.). Dog Breed Identification. *Electrical and Computer Engineering, University of California, San Diego.*



ACCOUNTS CLEARANCE



SK.SHIHABUZZAMAN
213-35-3188

Dashboard

Student Portal

Total Payable	Total Paid	Total Due	Total Other
741,200.00	741,202.02	-2.02	1,500.00



PLAGIARISM REPORT

213-35-3188

ORIGINALITY REPORT

15%

SIMILARITY INDEX

13%

INTERNET SOURCES

10%

PUBLICATIONS

8%

STUDENT PAPERS

PRIMARY SOURCES

1	dspace.daffodilvarsity.edu.bd:8080 Internet Source	1%
2	www.mdpi.com Internet Source	1%
3	Submitted to Midlands State University Student Paper	1%
4	Submitted to University of Queensland Student Paper	<1%
5	Submitted to Glyndwr University Student Paper	<1%
6	www.coursehero.com Internet Source	<1%
7	Thangaprakash Sengodan, Sanjay Misra, M Murugappan. "Advances in Electrical and Computer Technologies", CRC Press, 2025 Publication	<1%
8	academic.oup.com Internet Source	<1%
9	"Data Science and Big Data Analytics", Springer Science and Business Media LLC, 2025 Publication	<1%
10	Submitted to Monash University Student Paper	<1%
11	umpir.ump.edu.my Internet Source	<1%