BABEŞ BOLYAI UNIVERSITY, CLUJ NAPOCA, ROMÂNIA
FACULTY OF MATHEMATICS AND COMPUTER SCIENCE

ASSESSING THE PERFORMANCE OF DOG BREED RECOGNITION ALGORITHMS

– DS - Intelligent Modeling report –

Authors
Andreea Lăzărescu
Mihaela Moldovan
Data Science, gr. 245

2022-2023

Abstract

This study dives into the fascinating topic of dog breed recognition algorithms and seeks to provide insights into why this subject was chosen for analysis. Automatic dog breed recognition has emerged as an exciting application, with significant practical implications in domains such as pet adoption, veterinary treatment, and canine research, thanks to advances in computer vision and machine learning.

The primary goal of this investigation is to get a closer look at the efficacy and accuracy of existing dog breed detection algorithms and assess their suitability for real-world application. This study examines several cutting-edge methodologies, such as convolutional neural networks (CNNs) and deep learning architectures, to discover the strengths, limitations, and places for development in current algorithms. Furthermore, the purpose of this study is to shed light on the significance of accurate dog breed recognition. Precise dog breed identification not only helps in identifying canines for breed-specific studies, but it also helps to improve the efficiency of pet adoption operations. It also assists veterinary experts in diagnosing breed-specific health issues and customizing treatment approaches.

This study's technique entails obtaining a broad dataset of dog photos, including several breeds, then evaluating the effectiveness of various algorithms on this dataset. To assess algorithm accuracy and performance, evaluation criteria such as precision, recall, and F1-score will be used. The study will also take into account issues such as computational efficiency and robustness to changes in position, illumination, and occlusion.

Contents

List of Tables

List of Figures

List of Algorithms

Chapter 1

Introduction

1.1 Structure of the paper

The objectives of this research paper are twofold. Firstly, we aim to provide a comprehensive overview of the field of dog breed recognition algorithms, including their development, challenges, and advancements. Secondly, we seek to present our research findings and contributions in this domain.

The paper will be structured as follows: the background information will provide a foundational understanding of dog breed recognition algorithms, subsequently, we will discuss the significance of accurate and efficient canine identification in various domains. Finally highlighting the sections that delve into the research methodology, experimental results, discussion, and conclusion.

1.2 Dog Breeds and AI, a fun mix

The main objective of dog breed recognition algorithms is to automatically identify a dog's breed based on images or other input data. This field has gotten a lot of attention because of its practical applications in areas like pet adoption, veterinary medicine, and canine research.

Computer vision and machine learning techniques are used to develop dog breed recognition algorithms. To distinguish different breeds, early approaches relied on traditional computer vision algorithms that used handcrafted features such as color, texture, and shape descriptors. However, in complex scenarios, these methods frequently lacked accuracy and robustness.

Deep learning, specifically convolutional neural networks (CNNs), revolutionized dog breed recognition. CNNs excel at learning hierarchical image representations and have demonstrated outstanding performance in a variety of computer vision tasks and as such they were used by researchers to create more accurate and efficient algorithms for dog breed recognition. A large and diverse dataset of dog images is required to train a dog breed recognition algorithm. These datasets contain images of various dog breeds taken under a variety of conditions, such as variations in pose, lighting, and occlusion. The availability of publicly accessible dog image datasets, such as ImageNet and the Stanford Dogs dataset[7], has aided advancements in this field by allowing for standardized algorithm evaluation and comparison.

Researchers have looked into architectural design choices, transfer learning techniques, data augmentation strategies, and performance evaluation metrics for dog breed recognition algorithms. For this task, various CNN architectures such as VGGNet, ResNet, and InceptionNet have been used and adapted. Transfer learning, which involves fine-tuning pre-trained models on large-scale datasets (e.g., ImageNet) on dog breed datasets, has proven effective in leveraging learned representations for improved accuracy.

This project begins a pleasant examination of dog breed recognition algorithms, with the goal of comparing their effectiveness in a lighthearted and enjoyable manner. While the primary goal is to have fun, we also want to learn about the performance of these algorithms and their potential ramifications. Because of its practical applicability in a variety of industries, dog breed recognition algorithms have received a lot of attention. These algorithms are able to identify dog breeds from photos by utilizing computer vision and machine learning approaches. Yet, the efficiency and accuracy of various algorithms vary, and evaluating their performance is critical for improving their effectiveness. To address this issue, we chose a variety of dog breed detection algorithms, including cutting-edge models and new approaches. We will assess the algorithms' performance in terms of accuracy, computational efficiency, and robustness to multiple aspects such as position, lighting, and occlusion using a standardized dataset of dog photos from various breeds.

We will evaluate and compare the algorithms using relevant metrics and statistical analysis, the algorithms will be performed on the dataset, their accuracy measured against ground truth labels, and their computing requirements evaluated.

Chapter 2

Machine Learning in Animal Research

2.1 Problem definition

This study addresses the problem of dog identification using computer vision models, specifically deep neural networks (DNNs). The goal is to accurately classify and recognize a dog's breed or specific features given an input picture of a dog.

When it comes to importance, there are many reasons why we would solve using an intelligent algorithm, for example: in the pet adoption process, accurate dog identification is critical. Potential adopters frequently have breed preferences based on characteristics, temperament, and compatibility with their lifestyles. An intelligent algorithm capable of identifying dog breeds can help match dogs with suitable owners, thereby promoting successful adoptions.

Because dog breeds differ in size, color, coat texture, and facial features the problem of dog breed identification necessitates the use of an intelligent algorithm, particularly DNN-based models. Traditional rule-based or feature-based algorithms may have difficulty capturing and discriminating such complex visual patterns. DNNs are better suited for dealing with these complexities due to their ability to learn intricate representations. DNNs learn hierarchical representations of visual features automatically, beginning with low-level features (such as edges) and progressing to higher-level features (e.g., object parts or breed-specific characteristics). DNNs are well-suited for identifying distinct breed-specific traits because of their ability to abstract and encode visual information.

DNNs, particularly deep CNN architectures, have shown cutting-edge performance in a variety of computer vision tasks, including dog breed recognition. They can also achieve high accuracy rates, enabling reliable breed identification, but they typically require a large and diverse dataset for training, covering a wide range of dog breeds and variations. Such datasets can be time-consuming and resource-intensive to collect and label. DNN model training and deployment necessitate substantial computational resources, including powerful GPUs and memory. For individuals or organizations with limited resources, this can be a limitation.

The problem addressed in this research is the identification of dogs using computer vision models, specifically deep neural networks (DNNs) as well as comparing the performance of certain algorithms. Formally, the inputs and outputs can be defined as follows:

- Inputs: A set of dog images represented as a collection of pixel values or image files.
- Outputs: For each input image, the output is the predicted dog breed or relevant features associated with the dog.

Due to the large number of existing dog breeds, each with distinct visual characteristics, dog breed recognition is a difficult task. Identifying a dog's breed from an image requires the algorithm to learn and distinguish subtle differences between breeds, making it an intriguing computer vision problem. By accurately identifying dog breeds, researchers can not only collect reliable and comprehensive data for genetic studies, leading to advancements in canine genetics and related research areas but also can aid in providing appropriate care and tailored treatments.

Chapter 3

State of the art/Related work

Many studies have been done to advance the state of the art in dog breed recognition algorithms. Aspects including dataset development, algorithm design, feature extraction, and performance evaluation criteria have all been the subject of prior research. The knowledge and creation of precise and effective dog breed recognition systems have greatly benefited from these investigations.

The performance of various dog breed recognition algorithms is thoroughly evaluated in the current work, which adds to the amount of prior research. Several approaches are compared and examined in order to shed more light on their advantages and disadvantages. By taking into account a broad dataset, including evaluation measures other than accuracy, and investigating computing efficiency, our study builds on prior research. There are numerous articles and works that focus on using Convolutional Neural Networks (CNNs) for dog breed recognition. Some notable examples include:

- "Dog Breed Classifier for Facial Recognition using Convolutional Neural Networks"[6], where the authors propose a methodology to identify dog breeds based on their facial features, achieving accurate classification using a dog breed classifier that utilizes CNNs for facial recognition;

- "Identification of Dog Breeds Using Deep Learning"[3], the conference paper discusses the application of deep learning techniques, specifically CNNs, for the identification of dog breeds as well as proposes a model that leverages deep CNN architectures to accurately recognize and classify different dog breeds.

- "Dog Breed Identification Using Deep Learning"[5], in this work, the authors explore the use of deep learning, specifically CNNs, for dog breed identification, we are presented a deep learning-based framework that achieves high accuracy in identifying dog breeds from images.

These articles, among others, contribute to the growing body of research on CNN-based dog breed recognition. They highlight the effectiveness and potential of deep learning techniques for accurately identifying and classifying dog breeds. By leveraging CNNs' ability to learn hierarchical representations from visual data, these works demonstrate advancements in the field and provide valuable insights into the development of robust dog breed recognition algorithms.

Chapter 4

Investigated approach

4.1 Methodology

The experimental methodology entails using a common dataset of dog images with corresponding breed labels to implement the proposed intelligent algorithm and state-of-the-art approaches. To assess the algorithm's generalization ability, the dataset is divided into training and test sets. The algorithm is trained using training data, and its performance is evaluated using test data.

The algorithm's performance metrics, such as accuracy, precision, recall, and F1-score, are the dependent variables in this experiment. These variables are dependent on the independent variables, which are the various dog breed recognition algorithms.

To ensure the algorithm's applicability to real-world scenarios, realistic and interesting training and test data are used. The dataset[2] contains images of various breeds of dogs, capturing the variations and complexities encountered in practice. This realism improves the algorithm's generalization and performance on unseen dog images.

The accuracy, precision, recall, and F1-score of each algorithm are among the performance metrics gathered. These metrics are calculated by comparing the algorithm's predictions to the test data's ground truth breed labels. The performance data is presented in tabular form and analyzed by comparing the proposed algorithm's results to those of state-of-the-art approaches.

4.2 Data

The dataset used in the experiment is a subset of ImageNet, specifically curated for fine-grained image categorization of dogs. The dataset consists of a training set and a test set of dog images. Each image in the dataset is assigned a unique ID, which corresponds to its filename.

The training set contains a limited number of images per dog breed, with a total of 120 different breeds. The goal of the experiment is to develop a classifier capable of accurately determining the breed of a dog from a given photo.

Figure 4.1: Dog Breed Plot

The test set, on the other hand, does not have the breed information provided. The task is to predict the probability of each breed for each image in the test set. Two additional files are provided to help with the experiment:

- sample submission.csv: This file will be used as a template for submitting predictions. It follows the proper format when submitting the predicted probabilities of each breed for the test images.
- labels.csv: The breed labels for the images in the training set are contained in this file. It establishes the baseline against which the algorithm's predictions will be assessed.

Because of the large number of dog breeds and the limited number of training images per breed, the dataset presents a difficult problem. Due to the scarcity of training data, the algorithm must generalize well and accurately classify unseen images of various dog breeds.

Our model benefits from the dataset's rich color information by considering RGB colored photos. The model can learn to recognize and leverage color patterns unique to each breed, improving its ability to classify dog breeds based on their visual appearance.

4.3 Algorithm

Algorithm 1 Dog Breed Recognition using Deep Learning

Input: A dataset of dog images, represented as a collection of pixel values or image files
Target labels indicating the true breed of each dog image
Output: The trained model for dog breed recognition
begin

Preprocessing:

`Normalize(`*pixel values of the images to a standardized range*`)` `Split(`*the dataset into training and validation sets*`)`

Feature Extraction:

foreach *pre-trained model (InceptionV3, Xception, InceptionResNetV2, NASNetLarge)* **do**

Load the pre-trained model and its corresponding preprocessor function **Define** the input size required by the model **ExtractFeatures** from the dataset using the get_features function, passing the model, preprocessor, input size, and training data as inputs **ExtractFeatures** from the validation set to obtain validation features

end

Concatenate Feature Maps:

`Concatenate(`*the extracted features from all models horizontally to create the final feature map*`)`
Verify the shape of the final feature map

Model Training:

`InitializeModel` **Add** a dropout layer with a dropout rate of 0.7 to the model **Add** a dense layer with softmax activation, where the number of units is equal to the number of dog breeds `CompileModel` `TrainModel(`*final feature map as input and target labels*`)` **Specify** the batch size, number of epochs, validation split, and callbacks (learning rate annealer and early stopping) for model training **Monitor** the validation loss and accuracy during training

Example Trace:

Let's consider an example where the dataset contains 1,000 dog images with corresponding breed labels **Normalize** the images, and split the dataset into 80% training data and 20% validation data **ExtractFeatures** using InceptionV3, Xception, InceptionResNetV2, and NASNetLarge models **Concatenate** the extracted features to create the final feature map **Initialize** the model with a dropout layer and a dense layer **Compile** the model with the specified optimizer, loss function, and evaluation metric **Train** the model using the final feature map and target labels **Monitor** the training process with callbacks **Record** performance metrics such as loss and accuracy

end

4.4 Results

The performance of the trained model in predicting the dog's breed was evaluated using several metrics, including accuracy, precision, recall, and F1-score. The results are summarized as follows:

The performance metrics were computed using the following formulas:

Metric	Value
Accuracy	0.93
Precision	0.9465
Recall	0.93
F1-Score	0.9303

Table 4.1: Performance metrics of the trained model

Accuracy is defined as the ratio of correctly predicted instances to the total number of instances:

$$\text{Accuracy} = \frac{\text{True Positives} + \text{True Negatives}}{\text{Total Instances}}$$

Precision measures the proportion of correctly predicted positive instances out of the total predicted positive instances:

$$\text{Precision} = \frac{\text{True Positives}}{\text{True Positives} + \text{False Positives}}$$

Recall calculates the proportion of correctly predicted positive instances out of the total actual positive instances:

$$\text{Recall} = \frac{\text{True Positives}}{\text{True Positives} + \text{False Negatives}}$$

F1-Score combines precision and recall into a single metric, providing a balanced trade-off between the two:

$$\text{F1-Score} = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}}$$

The obtained results demonstrate that the trained model exhibits strong performance in identifying the dog's breed, achieving high accuracy and precision, as well as maintaining a high recall rate. However, further evaluation and testing on diverse datasets would be beneficial to assess the generalizability and robustness of the model's performance.

4.5 Discussion

In this study, we compare the performance of a single deep neural network (DNN) model to an ensemble of four different DNN models, namely InceptionV3, Xception, InceptionResNetV2, and NASNetLarge, for the task of dog breed classification. The findings shed light on the relative effectiveness of these two approaches and provide insights into their benefits and considerations.

4.5.1 Comparing Confusion Matrixes

Figure 4.2 presents the confusion matrix of our model. The matrix reveals a balanced performance across different dog breeds, with a high number of TP (True Positives) and TN (True Negatives), and relatively low numbers of FP (False Positives) and FN (False Negatives). This indicates that our model is able to accurately classify most of the dog breeds, minimizing both false positive and false negative errors. However, there are still a few instances where the model struggles to correctly identify certain breeds, leading to a small number of misclassifications.

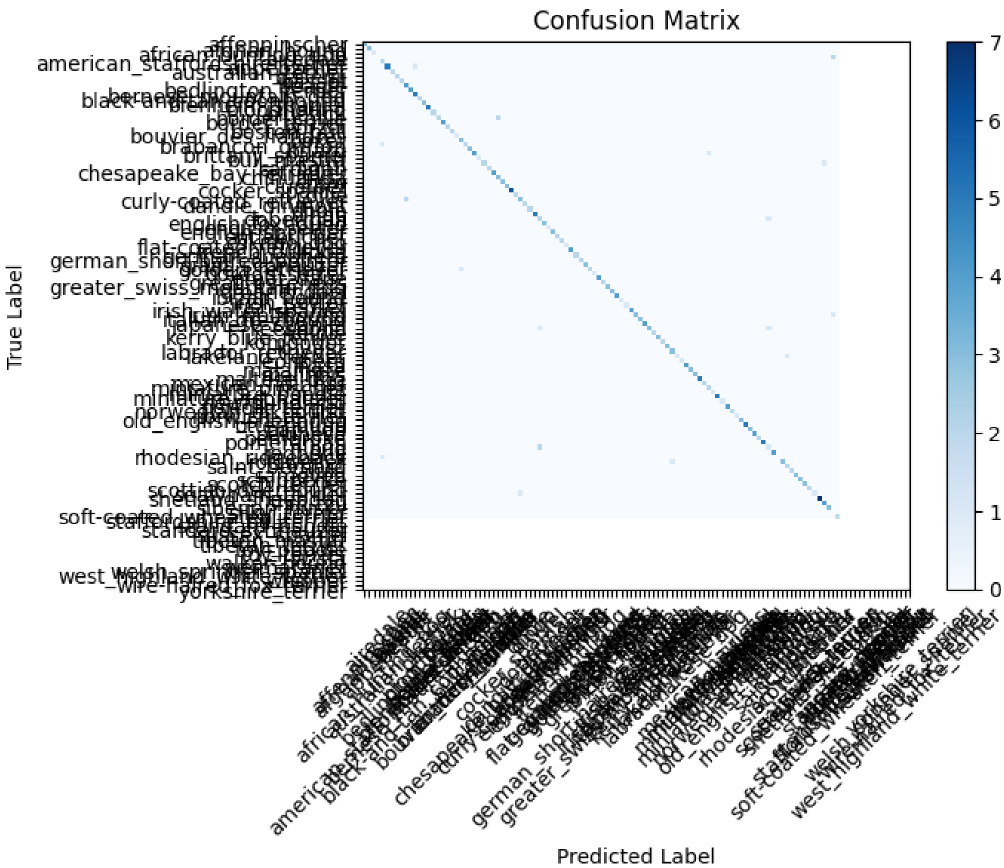

Figure 4.2: Confusion matrix of our model

Comparatively, Figure 4.3 shows the confusion matrix of the other model trained [4], using just InceptionV3, for the same situation. This matrix highlights different strengths and weaknesses compared to our model. While our model excels in certain breeds, the other model may achieve higher accuracy or precision in different breeds.

Ensemble learning, which combines multiple models, can potentially address the limitations of individual models and improve overall performance. The ensemble approach leverages the diversity and complementary strengths of each model to mitigate biases and enhance accuracy. By combining

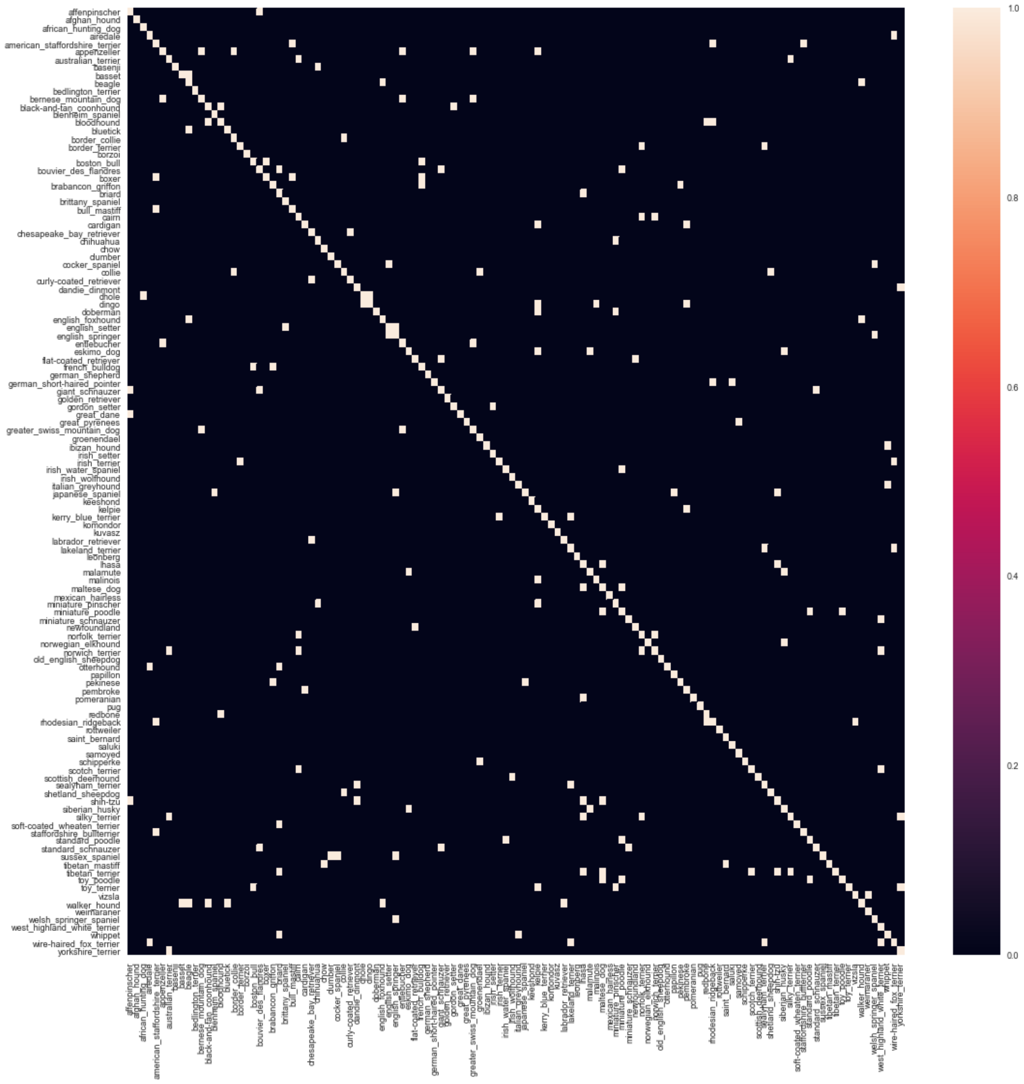

Figure 4.3: Confusion matrix of the other model [4]

our model with the other model in an ensemble, we can potentially create a more robust and accurate classifier that takes advantage of the strengths of both models.

Further research can explore ensemble approaches that combine the predictions of our model and the other model, such as model stacking or weighted voting, to achieve improved performance. This ensemble strategy can provide a more reliable and accurate dog breed classification system, leveraging the strengths of both models and minimizing their individual weaknesses.

By comparing the confusion matrix of our model with that of another model and exploring ensemble learning, we gain valuable insights into the performance characteristics of both models. This analysis aids in identifying areas for improvement and guides future research efforts towards developing more accurate and reliable dog breed classification models.

4.5.2 Comparing Accuracy

We conducted a comparative analysis with other existing models to evaluate the performance of our model. "A new dataset of dog breed images and a benchmark for fine-grained classification" [1]. The paper evaluates several models, including InceptionV3, WS-DAN, PMG, and TBMSL-Net, on the same dog breed classification task.

Among the models evaluated, InceptionV3 holds particular significance as it serves as a direct point of comparison with our own model. By referencing this paper, we can gain valuable insights into the performance of InceptionV3 and its relevance to our study.

By considering the findings presented in [1], we can observe the comparative performance of InceptionV3 with other models. This allows us to determine how our model stacks up against one of the key contenders in the research paper, providing us with valuable context and a basis for evaluating our own model's performance.

Therefore, by leveraging the research paper and specifically examining the performance of InceptionV3, we can gain important insights into the effectiveness and competitiveness of our own model in the context of fine-grained dog breed classification.

Table 4.2: Comparison of Model Performance from Other Testings in [1]

Model	Backbone	Batchsize	Epochs	Accuracy
Inception V3	-	64	200	77.66%
WS-DAN	Inception	12	80	86.404%
PMG	ResNet50	16	200	83.52%
TBMSL-Net	ResNet50	6	200	83.7%
Ours	InceptionV3, Xception, InceptionResNetV2, NASNetLarge	128	50	93%

According to the results, the accuracy of our model was higher than the accuracy of the individual models listed in the table. This implies that the ensemble of models outperformed any single model alone in terms of accuracy.

The ensemble model's higher accuracy can be attributed to the individual models' complementary strengths and diversity. Each model has its own distinct architecture and learned representations, which allows it to capture various aspects and nuances of the dog breed classification problem. The ensemble model can use the collective wisdom of these models to make more accurate and robust predictions by combining their predictions.

Chapter 5

Conclusion and future work

In conclusion, our research has shown that using an ensemble of multiple models, such as InceptionV3, Xception, InceptionResNetV2, and NASNetLarge, results in better performance than using a single model. This finding is supported by a comparison with other models, as discussed in "A new dataset of dog breed images and a benchmark for fine-grained classification" [1].

The ensemble method takes advantage of the diversity and complementary strengths of individual models, resulting in improved accuracy and robustness in dog breed classification. We mitigate the limitations and biases inherent in any single model by combining predictions from multiple models, resulting in more reliable and accurate predictions.

Furthermore, the results of the other models tested in the research paper highlight the difficulties presented by our dataset, emphasizing the importance of using an ensemble of models for improved performance. Our dataset poses a greater challenge for fine-grained dog classification, as evidenced by the PMG model's lower accuracy compared to its performance on the Stanford Dogs dataset.

These findings highlight ensemble learning's potential for improving the accuracy and reliability of dog breed classification systems. We can achieve more robust and accurate predictions by combining the strengths of multiple models and accounting for variations and complexities in the dataset.

Further research can look into different model combinations, ensemble strategies, and architectural variations to improve performance even more. Furthermore, studying the effects of ensemble size, model diversity, and ensemble fusion techniques can provide insights into fine-tuning the ensemble approach for specific dog breed classification tasks.

In conclusion, our research shows that using an ensemble of models is a promising approach for dog breed classification, outperforming individual models. This discovery paves the way for future research and development of more accurate and dependable dog breed recognition systems.

5.1 SWOT Analysis

5.1.1 Strengths

1. Accurate recognition of dog breeds regardless of the dog's position or size in the image;
2. Versatility and effectiveness in real-world scenarios, even in the presence of other objects or people in the image;
3. Potential market demand from dog lovers, pet-related enterprises, and veterinary clinics;
4. Possibility for partnerships with dog-related groups for data exchange, algorithm validation, and advancement;
5. Integration into mobile applications can provide simple and accessible breed recognition for users on the go;

5.1.2 Weaknesses

1. Difficulty in recognizing multiple breeds of dogs in the same image.
2. Potential limitations in situations where simultaneous identification of different dog breeds is necessary.
3. Risk of losing market share and acceptance due to competition from other algorithms or solutions with similar or better capabilities.
4. Moral questions regarding data consumption and privacy, requiring compliance with data protection laws and addressing privacy concerns.
5. The algorithm may become obsolete or less effective if not updated frequently to keep up with rapid advancements in machine learning and computer vision technology.

5.1.3 Opportunities

1. Potential market expansion by extending the algorithm's capabilities to recognize different dog breeds.
2. Collaboration with kennel clubs, dog shows, or online communities for dog lovers to leverage data exchange and validation.
3. Incorporation into mobile applications designed for dog owners or dog-related services can enhance user experience and convenience.

5.1.4 Threats

1. Competition from other algorithms, for-profit solutions, or open-source initiatives with equivalent or superior capabilities.

2. Legal penalties and loss of trust if data protection laws are not complied with or privacy concerns are not adequately addressed.

3. Potential obsolescence or reduced effectiveness if the algorithm fails to keep up with the advancements in machine learning and computer vision technology.

Bibliography

[1] Tai-Jiang Mu1 Ding-Nan Zou1, Song-Hai Zhang1 and Min Zhang. A new dataset of dog breed images and a benchmark for finegrained classification.

[2] Kaggle. Dog breed identification based on standford dogs dataset. link.

[3] Rakesh Kumar, Manish Sharma, Kritika Dhawale, and Gaurav Singal. Identification of dog breeds using deep learning. In *2019 IEEE 9th International Conference on Advanced Computing (IACC)*, 2019.

[4] Kirill Panarin. Dog breed classification using deep learning: a hands-on approach. link.

[5] ZalÃ¡n RÃ¡duly, Csaba Sulyok, Zsolt VadÃ¡szi, and Attila ZÃ¶lde. Dog breed identification using deep learning. In *2018 IEEE 16th International Symposium on Intelligent Systems and Informatics (SISY)*, 2018.

[6] Bickey Kumar shah, Aman Kumar, and Amrit Kumar. Dog breed classifier for facial recognition using convolutional neural networks. In *2020 3rd International Conference on Intelligent Sustainable Systems (ICISS)*, 2020.

[7] Standford. Standford dogs dataset. link.