*Nanotechnology Perceptions*
*ISSN 1660-6795*
*www.nano-ntp.com*

# Dog Breed Prediction using Deep Learning

**S. Rubia Parveen, J. Iswarya, A. Bhavana, A. Vamsi Madhav, K. N. Jaya Sekhar**

*Department of CSE RGMCET, Nandyal, India*
*Email: ruby0561@gmail.com*

Fine-grained picture recognition, or the identification of a dog breed in a given image, is one of the multi-magnificence classification problems in the current paper. The device in question uses convolutional neural networks, one of the deep learning methods. Two outstanding networks are trained and evaluated using the Stanford Dogs dataset, enabling a comprehensive analysis of performance metrics. An application that enables user interaction and model deployment provides access to the use and assessment of a convolutional neural network, Libraries and components from a neural community comparison are incorporated into this cell buyer and critical server, which can be used in an offline or online setting. In addition to increasing the accuracy of breed identification, this ground-breaking method advances fine- grained classification research across numerous other fields.

**Keywords:** inception resnet v2, mobile trained model, fine- grained image recognition, image classification, convolutional neural networks, dog breed identification.

## 1. Introduction

CNN is a new technology that comes from the non- deep learning field. In the field of image recognition, it has attained particular prominence. CNN automatically extracts pertinent resources from raw pixel data, in contrast to two conventional methods that rely on resources. These resources are excellent for a wide range of uses Object detection and error recognition may be on that list. These capabilities are now widely accepted in many fields.

In this paper we hone in on fine-tuning the CNN architecture using Stanford Dogs dice set in regard to dog breed classification. This set of dice brings upon challenges because of name difference due to races One could refer to this as an advanced image recognition issue.CNN's architecture was modeled after the structure of the human brain. However, a layer known as convolutional, fully connected, and clustering layers actually processes the image through a number of steps. These networks are ideal for handling massive volumes of data. Even in the areas of image classification and extensive video analysis, it can produce remarkable outcomes.

Our approach not only includes training these models. However, it also involved the development of easy- to use software that allows people to upload images to identify their species. This system covers all recognized breeds in detail. This illustrates the practical application of deep learning to animal classification. It is anticipated that the projects outcomes will provide important new information about CNNs' potential for better image classification tasks. Advances in deep learning, particularly with Convolutional Neural Networks (CNNs), have revolutionized computer vision in recent years. CNNs are extremely powerful tools for a wide range of tasks, such as object detection, image classification, and even video analysis, since learning the spatial hierarchies of features from images is their primary function, they do so automatically and adaptively.

Their unique architecture, which consists of layers that apply filters to the input image, pooling layers that lower the dimensionality, and fully connected layers that perform the classification, is responsible do this.

What is compelling about the application of CNNs to dog breed identification is the intricate variations between breeds. Traditional methods for image recognition have usually struggled with such subtleties, whereas CNNs do better at picking up on complex patterns and features that characterize one breed as opposed to another. The finetuning of already established architectures like Inception- ResNet V2 and NASNet-A mobile on a comprehensive dataset such as Stanford Dogs will further boost classification accuracy.

More than this, our project emphasizes the practical application and the user's engagement in the technical aspects of model training. The software system developed can upload images by the users and will immediately provide them with feedback regarding breed identification along with information details about the characteristics of each breed. This deep learning integration with the user- friendly technology highlights the scope of Al applications in day-to-day life and further encourages more appreciation and awareness for canine diversity Ultimately, this project is meant to show how CNNs can be used to solve significant real-world prokiems while opening up avenues for future research into fine-grained image recognition.

## 2. Literature Review

Image classification has significantly improved in recent years thanks to deep learning techniques, particularly Convolutional Neural Networks (CNN). This includes keeping dog breeds in mind. Numerous studies on different approaches have been conducted with the goal of improving the dog breed classification system's precision and effectiveness.

One method is by researchers who can use CNN to classify dog breeds based on images. Their research suggests a significant application of big data sets. The labeled photos in the Stanford dog dataset follow a variety of patterns. Thousands distributed among various species. Therefore, a high accuracy rate was obtained in the CNN model when it was trained on this dataset. This demonstrates the capacity to handle deep learning the authors claim that using conventional methods to accurately identify species is difficult.

Because they all look the same even though they have slight differences Learn to recognize what makes a dog different from other dogs with CNN's extraction capabilities without human intervention.

Transfer learning methods that make use of pre-trained VGG16 and ResNet50 architectures also make a significant contribution. by modifying the model to fit the dataset on dog breeds. An accuracy of over 90% was attained by the researchers. Transfer learning makes use of features from big datasets that have already been learned. This increases the model's efficiency and decreases the amount of training data needed for particular tasks. This method works really well. This is especially true when it comes to classifying dog breeds, which might show the least amount of variation across species. Methods other than CNN blend traditional image processing methods with deeplearning.

Apart from CNN, several researches have implemented hybrid approaches that combine traditional image processing techniques with deep learning. For example, the integration of methods such as LBP and HOG with a CNN improves the classification accuracy further. As a result, hybrid models of these approaches benefit from handcrafted features besides learned representations, thus giving better comprehension of the visual characteristics that define different breeds of dogs. Besides, improvements in mobile technology have made possible the creation of applications using these deep learning models for identifying dog breeds in real-time Implementation of CNNs within mobile apps enables a user to upload an image of a dog and obtain immediate breed identification. This is an easy application which improves user experience but is also helpful for veterinarians and breeders seeking the actual breed.

The literature also mentions the difficulties with fine- grained image recognition tasks such as dog breed classification. It has been observed that researchers require dataset quality and diversity along with effective preprocessing techniques such as data augmentation to achieve a high accuracy level. These methods enhance the robustness of the model against real-world variations in lighting conditions, angles, & backgrounds.

Some research has worked to address the challenges that face Using detection techniques to recognize fine-grained images Zhang et al., for example, extended the R-CNN architecture to recognize different image components.

Duan and associates found localizable attributes that assist in distinguishing among related classes. In his approach. Angelova et al addressed these by fusing segmentation and object detection methods.

Chen and ai additionally proposed a selective pooling vector that is specifically made for the recognition of fine-grained images. To put it shortly, this corpus of dog breed identification via CNN shows great improvements in the implementation of deep learning for an image classification problem. Through large- scale data, advance architectures, and transfer techniques, plus hybrids, accuracy increases and becomes much more efficient. Further developing these tools will enable some very pragmatic applications, ranging from the veterinary clinics and pet adoptions services. Further explorations of innovative methodologies will undoubtedly continue to contribute to progress in this very exciting field.

## 3. Methodology

In response to the increasing demand for accurate dog breed identification, we propose a comprehensive classification system named PawPrint. This system is designed to utilize advanced Convolutional Neural Networks (CNNs) and leverage the Stanford Dogs dataset to improve dog breed classification's accuracy and effectiveness. The suggested system seeks to address critical challenges such as fine-grained image recognition, data imbalance, and the need for real-time processing in various applications, including veterinary services and pet adoption platforms.

A. DATASET

The Stanford Dogs dataset, which comprises 120 different dog breeds, serves as the basis for our project. The training set consists of approximately 12,000 images, with around 100 images per breed, while the testing set contains 8,580 images that are unevenly distributed across the breeds. Supervised learning is made possible by labeling each image with the appropriate breed.

To prepare the data for model training, we first resize all images to standardized dimensions-299x299 pixels for Inception-ResNet V2, and 256x256 pixels for NASNet-A mobile. All input images are guaranteed to meet the specific tons of the CNN architectures being used thanks to this resizing.

The training dataset is then split into training folds and a validation fold to facilitate hyperparameter tuning Specifically, we employ a 5- fold cross validation technique, which generates five separate training and validation datasets, each containing 9,600 training images and 2,400 validation images.

Fine-tuning hyperparameters is necessary to maximize model performance. After identifying the best hyperparameters through cross- validation, we train our CNN models on the complete training dataset for Stanford Dogs. To guarantee an objective evaluation of model performance, the test dataset is left unaltered until evaluation.

Our approach leverages state-of-the-art CNN architectures are well-known for their efficiency in tasks involving image classification. By employing transfer learning techniques with models such as ResNet50 and Inception- ResNet V2, we can capitalize on pre-trained weights that have been optimized on large-scale datasets. This strategy not only speeds up convergence but also enhances accuracy by allowing our models to learn from previously acquired features relevant to dog breed classification. In addition to improving classification accuracy, pawprint aims to provide practical.

B. FEATURE EXTRACTION

Since feature extraction is used in your dog breed identification project to enhance the performance of convolutional neural networks (CNNs), the image data must be pre-processed, including resizing, normalizing, and augmenting, before the meaningful features that could significantly distinguish different dog breeds are obtained. This chapter elaborates on feature extraction methods and techniques specifically designed for dog breed classification.

Key Techniques in Feature Extraction

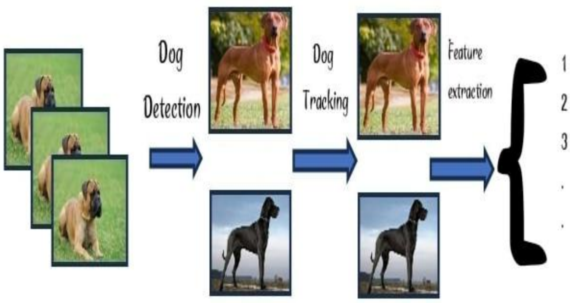
Convolutional Layers These make up the core of CNNs. Important components like shapes, edges, and textures are captured by applying filters to the input images. Certain filters can be adjusted for specific dog breeds to detect distinguishing characteristics like ear shape, muzzle length, and coat patterns.

Transfer Learning: ResNet and VGG16 are examples of pre-trained models that can be used to extract features effectively. These models are already able to identify many features from large datasets. You can benefit from these models' powerful feature extraction capabilities by optimizing them on your specific dog breed dataset and utilizing less training data.

Pooling Layers: Max pooling layers, which reduce dimensionality while preserving important features, are used to down sample feature maps. This aids the network in concentrating on dominant characteristics that are crucial for breed classification. Data Augmentation: Rotation, flipping, and scaling are examples of data augmentation techniques used to reduce overfitting and boost the model's resilience. This improves the model's ability to generalize across breeds and diversifies the training data.

Attention Mechanisms: By including attention mechanisms, the model will be able to focus on specific areas of an image.

In order to identify the location of the distinguishing features in this case, body structure or facial markings-a model that makes distinctions between similar breeds employ attention layers, like Alaskan Malamutes and Siberian Huskies.




Example Images for Feature Extraction

Use these illustrations to provide better visual evidence, including both Siberian Huskies and Alaskan Malamutes side-by-side. Points of comparison would be:

Coat Patterns: These dogs have double thick coats but the texture and length differ from each other. A Malamute has longer, thicker fur, where a Husky has shorter fur, a streamlined version of this double coat.

Facial Features: The muzzle shape and ear placement are different between the two breeds. The Husky has a narrower muzzle and erect ears, while the Malamute has a broader muzzle with ears that can tilt slightly forward.

Body Structure. The Malamute is usually larger and heavier than the Husky, which is a size difference that can be depicted in side-by-side comparison images.

With these advanced feature extraction techniques and the use of illustrative examples, your project on dog breed classification can achieve higher accuracy and robustness in identifying various breeds based on subtle visual cues.

## C. MACHINE LEARNING MODELS

It applies advanced machine learning techniques to dog breed classification using CNNs for images with their breeds.

Transfer learning is in use, letting the classifiers exploit a well-trained image classification model created elsewhere, saving time and resources on not having to develop our own model This approach benefits our model in that it uses learned features from these pre- trained networks instead of learning them from scratch.

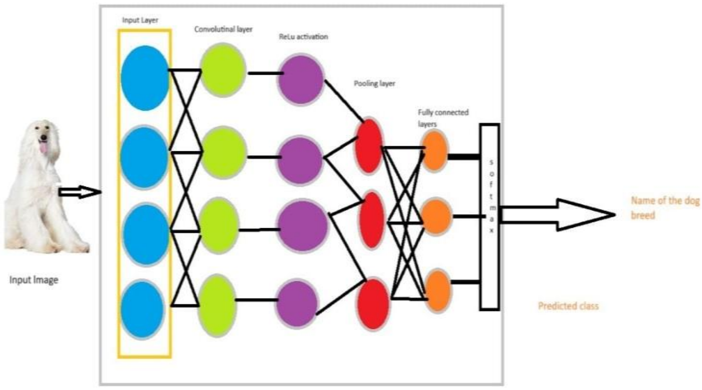



Fig. : Proposed Methodology

We consider two architectures for pre-trained CNNs: NASNet-A mobile and Inception-ResNet V2 NASNet- A is constructed using a framework of Neural Architecture Search, a Google AutoML concept that optimizes the structure of the network to get better performance. On the contrary, Inception-ResNet V2 is one of the deeper architectures with more than 300 layers and was developed by Google developers. The Stanford Dogs dataset, which includes pictures of more than 120 breeds, is used to refine both models.

Data augmentation is a key component of sur methodology to avoid overfitting during training. This technique essentially perturbs the input training images with various transformations before feeding them into the model for our CNNs, we use a particular set of preprocessing routines during training, including random reflection as well as cropping images using TensorFlow's distorted bounding box algorithm.

Moreover, during evaluation, we also apply a 87.5% central crop and further preprocess so as to have consistency in the input data.

We train our CNN models in-house on a personal computer with a GeForce GTX 1080 GPU and an Intel Core i5-6400 CPU. In this phase, we fine-tune the NASNet-A and Inception-ResNet V2 models with a SoftMax Cross- Entropy loss function and Nesterov momentum optimizers In this fine-tuning phase, we unfreeze the last fully connected layer, or logits, and freeze all the earlier layers to keep the learned features intact.

Hyperparameter tuning is an optimization of model performance, where we use cross-validation to empirically select good hyperparameters before training on the entire Stanford Dogs dataset. Common hyperparameters include a batch size of 64 and an exponentially decreasing learning rate, which starts at different initial values for each model: 0.029 for NASNetA and 0.1 for Inception-ResNet V2. The rate of learning drops by 10% every three epochs.

To guarantee efficient model accuracy assessment, performance metrics are monitored throughout the training and testing stages. After training, we will analyze the resultant through confusion matrices to visualize how well each breed is classified and any common misclassifications.

Last but not least, once the models are trained and evaluated, we prepare them for deployment by freezing the trained networks into a single binary file that contains the neural networks' parameters and structure. It will enable integration with the applications where users can upload images of dogs to identify breed.

In summary, our machine learning models rely on state of-the-art CNN architectures and advanced strategies like data augmentation and transfer learning to develop a successful dog breed classification system. This approach not only enhances accuracy but also ensures that our models are robust and ready for practical application in real-world scenarios involving dog identification.

D. EVOLUTION OF MODEL

We assess the performance of our dog breed classification models using a variety of evaluation metrics, including accuracy, precision, recall, and F1 score. Accuracy is a measure of the model's overall correctness in classifying dog breeds.

While precision displays the proportion of true positive predictions among all positive predictions, recall assesses the model's ability to identify all relevant cases. The F1 score provides a balanced measure of precision and recall, ensuring comprehensive evaluation of model performance.

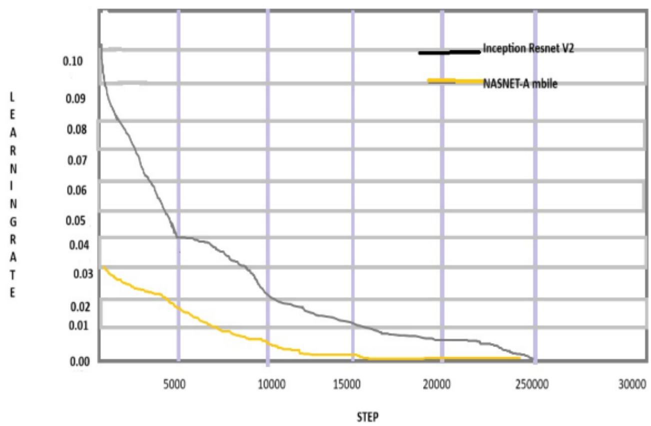



Fig:1. Adaptive Learning Rate During Training Sessions

## 4. Experimental Results

In the validation of our project's created machine learning models for dog breed classification, we used accuracy as the primary metric. Accuracy is widely recognized in classification tasks as it is easy to interpret and calculate; it even gives clear-cut indications of the performance a model is doing by determining the ratio of correct predictions made by the model.

Two distinct pre-trained convolutional neural networks were used in our project to train models:

1. NASNet-A mobile and Inception-ResNet V2. Each model was evaluated based on its ability to accurately classify images of dogs into their respective breeds. The experimental evaluation yielded the following results.

2. NASNet-A Mobile Model: This model achieved an impressive accuracy of 95% in identifying various dog breeds from images.

3. Inception-ResNet V2 Model: Using this architecture, the model correctly classified dog breeds with 97% accuracy, differentiating between similar- looking breeds.

These results show the efficacy of the architectures used in identifying the breeds, thus showing that our approach is strong and dependable for applying deep learning methods to tasks involving the classification of fine-grained images.

To provide a thorough evaluation of model performance, we took into account not only accuracy but also precision, recall, and F1 score.

These metrics help understand not only how many predictions were correct but also how well the models performed in identifying true positives and minimizing false positives and negatives.

All things considered, the experimental findings show that it is feasible to classify dog breeds using sophisticated CNN architectures, paving the way for additional study and use in real-world animal identification & welfare situations.

## 5. Conclusion and Future Work

Two distinct convolutional neural network architectures the deep InceptionResNet V2 and the mobile NASNet-A-have been proposed. Dog breed recognition was the particular image classification task for which these models were evaluated. The Stanford Dogs dataset was used to refine the pre- trained networks, and the outcomes are encouraging The NASNet-A mobile model demonstrated its efficiency, even as a smaller, mobile-friendly version, only 10% worse than the deep Inception-ResNet V2 model in terms of accuracy.

We have also developed a mobile application called Sniff! that uses these fine-tuned CNNs to classify dog breeds from images without needing an internet connection. Because of this feature, users can access it in a variety of circumstances.

A number of strategies could improve these convolutional neural networks' performance even more. Such methods involve generating more training images through the usage of Generative Adversarial Networks, alternative loss functions center loss, other architectures like multiple

CNNs, and involving more popular dog breeds to the dataset. The additional suggestions may include employing detectors for more than one dog within one single image and enhancing server classification and mobile classification on servers for better performance.

## References

1. "ImageNet Classification with Deep Convolutional Neural Networks," by A. Krizhevsky, I. Sutskever, and G. E. Hinton, in Advances in Neural Information Processing Systems 25. Curran Associates, Inc.,2012, pp. 1097-1105.
2. S. Ren, K. He, R. B. Girshick, and J. Sun, "Faster R-CNN: towards real-time object detection with region proposal networks," CORR, volabs/1506.01497.2015.
3. "Convolutional neural networks for speech recognition," IEEE/ACM Transactions on Audio, Speech, and Language Processing, vol. 22, no. 10, pp. 1533-1545, October 2014, by O. Abdel-Hamid, A. r. Mohamed, H. Jiang, L. Deng, G. Penn, and D. Yu.
4. I. Goodfellow, J. Pouget-Abadie, M.Mirza, B. Xu, D. Warde-Farley, S. Ozair, A. Courville, and Y. Bengio,"Generative adversarial nets," in Advances in neural information processing systems, 2014, pp. 2672-2680.
5. T. Lindeberg, "Scale Invariant Feature Transform," Scholarpedia, vol. 7, no. 5, p. 10491, 2012, revision #153939.
6. N. Dalal and B. Triggs, "Histograms of oriented gradients for human detection," in Computer Vision and Pattern Recognition, 2005. CVPR 2005. IEEE Computer Society Conference on, vol. 1. IEEE, 2005, pp. 886-893.
7. A. Khosla, N. Jayadevaprakash, B. Yao, and L. Fei- Fei, "Novel dataset for fine- grained image categorization," in First Workshop on Fine-Grained Visual Categorization, IEEE Conference on Computer Vision and Pattern Recognition, Colorado Springs, CO, June 2011.
8. R. J. Schalkoff, Artificial neural networks. McGraw- Hill New York,1997, vol. 1.
9. A. Krizhevsky. 1. Sutskever, and G. E. Hinton, "Imagenet classification with deep convolutional neural networks" [10] A Karpathy, G. Toderici, S. Shetty, Leung, R. Sukthankar, and L. Fei-Fei, "Large- scale video classification.