International Journal of Environmental Sciences
ISSN: 2229-7359
Vol. 11 No. 21s, 2025
https://theaspd.com/index.php

Automated Dog Breed Recognition Using Cnn-Based Transfer Learning

Shaik Saddam Hussain[1], Karumuri Sri Rama Murthy[2], Ch. Sandhya Rani[3], Thirupathi Nanuvala[4], Ch. Rathan Kumar[5], Gullapalli Sathar[6]
[1,2,3,4]Assistant Professor, Department of Computer Science and Engineering, VNR Vignana Jyothi Institute of Engineering and Technology, Hyderabad – 500090, Telangana, India.
[5]Ch. Rathan Kumar, Assistant Professor, Department of Computer Science and Engineering, Keshav Memorial Institute of Technology, Hyderabad – 500029, Telangana, India.
[6]Assistant Professor, Department of Computer Science and Engineering (CyS,DS) and AI&DS,
VNR Vignana Jyothi Institute of Engineering and Technology, Hyderabad – 500090, Telangana, India.
saddam_sk@vnrvjiet.in[1], sriramamurthy_k@vnrvjiet.in[2], sandhyarani_ch@vnrvjiet.in[3],
thirupathi_n@vnrvjiet.in[4], rathanoucse@gmail.com[5]

Abstract: *Image classification remains a major challenge in supervised machine learning, requiring the accurate assignment of images to specific categories. This study tackles this issue by leveraging transfer learning through the InceptionResNetV2 model, pretrained for multi-class dog breed classification. Transfer learning, a cornerstone in deep learning, enables the reuse of knowledge gained from one task to enhance performance on a related one. Instead of building a neural network from scratch, this approach adapts learned features—particularly the pretrained model weights—for the specific task of dog breed identification and also for Environmental safety.*
The research employed the Stanford Dog Dataset, which includes 20,580 images spanning 120 dog breeds. Following preprocessing and data augmentation, several pretrained convolutional neural network (CNN) models—including VGG16, ResNet50, and InceptionV3—were fine-tuned for the task. Performance evaluation was conducted using key metrics such as accuracy, precision, recall, and F1-score. Among the models tested, InceptionV3 delivered the best performance, achieving an accuracy of 91.2%, thereby highlighting the efficacy of transfer learning in addressing fine-grained image classification problems like dog breed recognition.
Keywords- *Transfer Learning, Dog Breed Classification, InceptionResnetV2*

I. INTRODUCTION

Image-based dog breed classification poses a significant challenge due to the wide diversity in breed-specific traits such as shape, size, coat texture, and color. Traditional machine learning methods often struggle in this domain, as they rely heavily on manual feature engineering and lack the flexibility to handle high intra-class variability. These limitations restrict their effectiveness in fine-grained classification tasks, especially when applied to visually similar breeds.

In contrast, this study introduces a novel approach by harnessing the capabilities of Inception ResNet v2 architecture, which merges the strengths of Inception modules for multi-scale feature extraction with the depth and optimization benefits of residual connections. This hybrid design enables more efficient and accurate processing of subtle visual distinctions across dog breeds—setting it apart from both conventional CNNs and older deep learning architectures.

A key innovation of this work lies in its comprehensive preprocessing pipeline, which incorporates image augmentation and background removal to enhance dataset quality and promote better generalization. Unlike prior approaches that often rely on raw or minimally processed images, our methodology ensures that the model is trained on balanced and noise-reduced data, allowing it to better capture breed-specific details.

Furthermore, the model leverages transfer learning, fine-tuning Inception-ResNet v2 on a curated dataset of dog breeds after initial pretraining on a large-scale image corpus. This step significantly reduces training time while boosting classification accuracy—a clear improvement over training models from scratch or using less sophisticated pretrained networks.

In addition to superior performance, our study provides a rigorous comparative evaluation against baseline CNN models and traditional classifiers, demonstrating notable gains in accuracy, precision, and recall. The research also addresses key challenges like class imbalance, overfitting, and computational efficiency, proposing practical solutions tailored to the complexities of breed classification.

This work's unique contribution is the combination of a cutting-edge architecture, improved preprocessing, and an optimized training strategy, which produces a dog breed classification system that is both extremely accurate and scalable. This makes our strategy a substantial improvement over current approaches and demonstrates how contemporary deep learning techniques may be used to solve fine-grained picture recognition issues.

The Proposed work focused on

- Leveraged InceptionResNetV2's feature extraction for effective dog breed classification.
- Applied data augmentation and preprocessing to enhance model robustness.
- Fine-tuned the model on a specialized dataset, improving accuracy across diverse breeds.
- Outperformed baseline models, including traditional classifiers, in classification accuracy.
- Addressed challenges like breed variability, overfitting, and class imbalance.

II. METHODOLOGY

2.1 Existing Model

Pretrained convolutional neural network (CNN) architectures including VGG16, ResNet50, DenseNet121, and InceptionV3 are commonly used in current dog breed classification methods. Each of these architectures offers unique benefits in terms of feature extraction and overall performance. VGG16, widely adopted for its simplicity and uniform architecture, serves as a baseline model. However, due to its relatively shallow depth and lack of residual or multi-scale learning components, it achieves only 76.4% accuracy, limiting its ability to extract complex features necessary for distinguishing visually similar breeds. ResNet50 addresses some of these limitations by introducing residual connections, enabling the construction of deeper networks without degradation in learning. This results in an improved accuracy of 80.3%, though challenges remain in capturing fine-grained inter-class distinctions among breeds. DenseNet121 further enhances model efficiency by employing dense connectivity, each layer receives inputs from all preceding layers. This facilitates feature reuse and gradient flow, yielding a higher accuracy of 82.1% and richer feature representations. However, its architecture does not explicitly focus on multi-scale feature extraction, which is vital for recognizing subtle morphological differences among dog breeds. In contrast, InceptionV3 introduces inception modules capable of learning at multiple spatial scales simultaneously. This design improves the model's ability to capture complex visual patterns and localized details, resulting in a further improved classification accuracy of 83.7%. Building upon these advancements, Inception-ResNet v2 integrates the strengths of both inception modules and residual learning, achieving even higher accuracy and robustness. Its hybrid architecture makes it particularly well-suited for fine-grained classification tasks, such as dog breed identification, where nuanced differences in texture, shape, and coloration must be captured reliably.

Dataset and Preprocessing

The study's dataset, which comes from the Stanford Dogs Dataset, includes more than 20,000 tagged photos of 120 different dog breeds, with an average of 150–180 photos per class. In order to comply with the input specifications of Inception-based models, all photos are shrunk to 299 × 299 pixels.

Key preprocessing steps include:

- **Normalization**: Pixel values are scaled to the 0,10, 10,1 range by dividing by 255.
- **Image augmentation**: Applied dynamically during training to reduce overfitting and improve generalization. Techniques include:
  - Random horizontal flipping
  - Rotation (±20 degrees)
  - Zoom range (0.8–1.2)
  - Brightness and contrast adjustment
  - Random cropping and padding
- **Background removal** (optional step): To emphasize the dog subject and reduce noise from background elements, some images undergo automatic segmentation and background subtraction.

Training Configuration

For model training, the following hyperparameters are used:

- **Learning rate**: Initialized at 0.0001 with adaptive learning rate decay based on validation loss
- **Batch size**: 32
- **Epochs**: 50, with early stopping based on validation performance
- **Optimizer**: Adam was selected as the optimizer due to its strong performance and quick convergence

in deep learning problems

- **Loss function**: Suitable for tasks involving multi-class categorization is categorical cross-entropy.

This comprehensive setup, integrating advanced architecture, rigorous preprocessing, and optimal hyperparameter tuning, enables Inception-ResNet v2 to outperform previous models and demonstrates its suitability for high-accuracy dog breed classification.

2.2 Proposed Solution

2.2.1 Overview:

The proposed system leverages the Inception-ResNet v2 model, a deep learning architecture that combines Inception modules and Residual Networks (ResNets) for efficient image classification. This model captures complex patterns and handles a diverse range of features, making it ideal for identifying dog breeds. Inception-ResNet v2 integrates multiple convolutional filters through inception modules to process different spatial features in an image simultaneously, enabling the network to capture both detailed and high-level semantic information. Residual connections also solve the vanishing gradient issue, which makes it easier to train deeper networks and improves model performance.

2.2.2 Architecture of the System

Fig. 1 (Architecture Diagram) shows the architecture of the suggested system. It consists of the following primary parts:

Input Module: Accepts an input image that requires classification, which can be either from the Dog Images Dataset (for training and testing) or a new image provided by the user.

Preprocessing Module (Dog Breed Preprocessor): The preprocessor standardizes the images by converting them to JPEG format, resizing, and performing background removal. To boost dataset diversity and decrease overfitting, additional data augmentation techniques like rotation, scaling, flipping, and color tweaks are used. These preprocessing steps ensure that the images are prepared in a consistent format, allowing the model to focus on breed-specific features.

Inception-ResNet v2 (Dog Breed Classifier): This model takes the preprocessed images and classifies them into either a specific dog breed or a "Not Dog" category. Inception modules enable the model to extract various spatial features, while residual connections enhance the training of deep networks. For this system, transfer learning is employed to leverage pre-trained weights, improving accuracy and accelerating the training process.

Output Module: The model's output indicates whether the input image belongs to a dog breed or is classified as "Not Dog." If a dog breed is detected, the specific breed name is provided as the output.

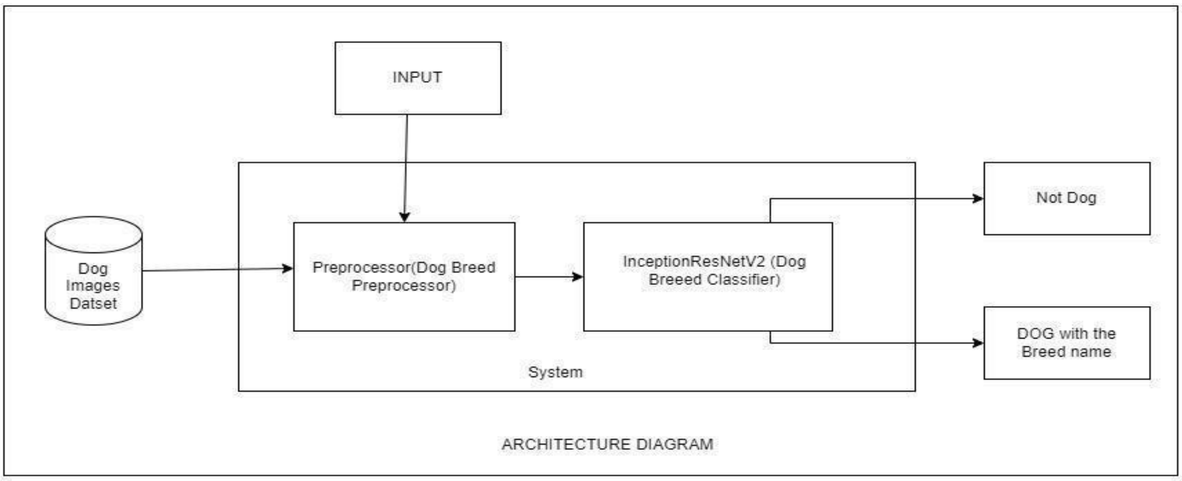

Fig-1 (Architecture Diagram) visually represents the data flow through each module, from input preprocessing to final classification.

2.2.3 Implementation Modules:

The method starts by gathering a representative and varied dataset of pictures of several dog breeds. This dataset is assembled from various sources, including public repositories, online image databases, and contributions from enthusiasts and organizations. To avoid class imbalance and enhance the model's capacity to generalize across various real-world situations, it is imperative to guarantee that every breed is fairly represented.

To get the images ready for model training, data preparation entails a number of procedures. To standardize the data and guarantee consistency and model compatibility, all photos are converted to JPEG format in order to boost the variety within the training dataset and mitigate overfitting, techniques such as rotation, scaling, flipping, and color adjustments are employed for image augmentation. Additionally, background removal is performed to focus solely on the dog breeds, enhancing the clarity of the features extracted by the model and potentially improving classification accuracy.

The dataset is then divided into two parts: The remaining 20–30% is reserved for testing, while the remaining 70–80% is used for training. This division gives the model a solid evaluation set to gauge its performance and guarantees that it has a sufficient amount of data to learn from. To prevent data leaks and get an objective assessment of the model, training and testing data must be properly separated.

Training the model involves using the preprocessed training data with the Inception-ResNet v2 architecture. This model's combination of inception modules and residual connections allows it to capture diverse features and facilitate the training of very deep networks. Transfer learning techniques may be employed to utilize pre-trained weights, accelerating the training process and enhancing accuracy, especially with limited datasets.

The testing dataset is used to evaluate the model's performance following training. Metrics like accuracy, precision, and recall are used in the evaluation step to gauge how well the model handles fresh, unseen data. This stage highlights opportunities for more development and offers insights into the model's efficacy. Before implementing the model for real-world uses, more optimization and fine-tuning may be done to improve its performance based on the evaluation findings.

2.2.4 Advantages and Disadvantages:
By fusing the advantages of the Inception and ResNet architectures, Inception-ResNet v2 provides a potent answer to picture categorization problems. The network's ability to record a variety of spatial characteristics at different sizes is made possible by the inclusion of inception modules, which increases classification accuracy. The model is effective for both training and deployment because to the utilization of 1x1 convolutions in the inception modules, which also helps to lower the computational burden and number of parameters. It performs better on a variety of jobs because of its capacity to manage intricate patterns and a wide range of features.

However, the model's complexity demands substantial computational resources for both training and inference, which may limit accessibility for users with limited hardware. The depth and intricacy of Inception-ResNet v2 result in longer training times compared to simpler models, which can be impractical for scenarios requiring rapid deployment. Additionally, the sophisticated design of the model presents challenges in interpretation and debugging, requiring a deep understanding of its components and their interactions, which can hinder troubleshooting and fine-tuning efforts.

2.2.5 Datasets:
The dataset utilized for dog breed classification is a meticulously curated collection of images representing 120 distinct dog breeds. Each image in the dataset is uniquely identified by a file name, which serves as its unique identifier, ensuring that every image is easily traceable and organized. This dataset is crucial for training and evaluating machine learning models tasked with identifying and classifying different dog breeds.

The training dataset comprises 10,222 images, covering all 120 breeds. This portion of the dataset is fundamental to the model's learning process, as it provides the raw material from which the model extracts and learns the distinguishing features of each breed. The diversity within this training dataset is vital because it ensures that the model is exposed to a wide variety of examples from each breed. This exposure is key to the model's ability to generalize, meaning it can recognize and classify breeds accurately, even when presented with new, unseen images. By encountering multiple variations within each breed, the model learns to identify not just the most common features but also the subtle differences that distinguish one breed from another. This results in a more robust and reliable model.

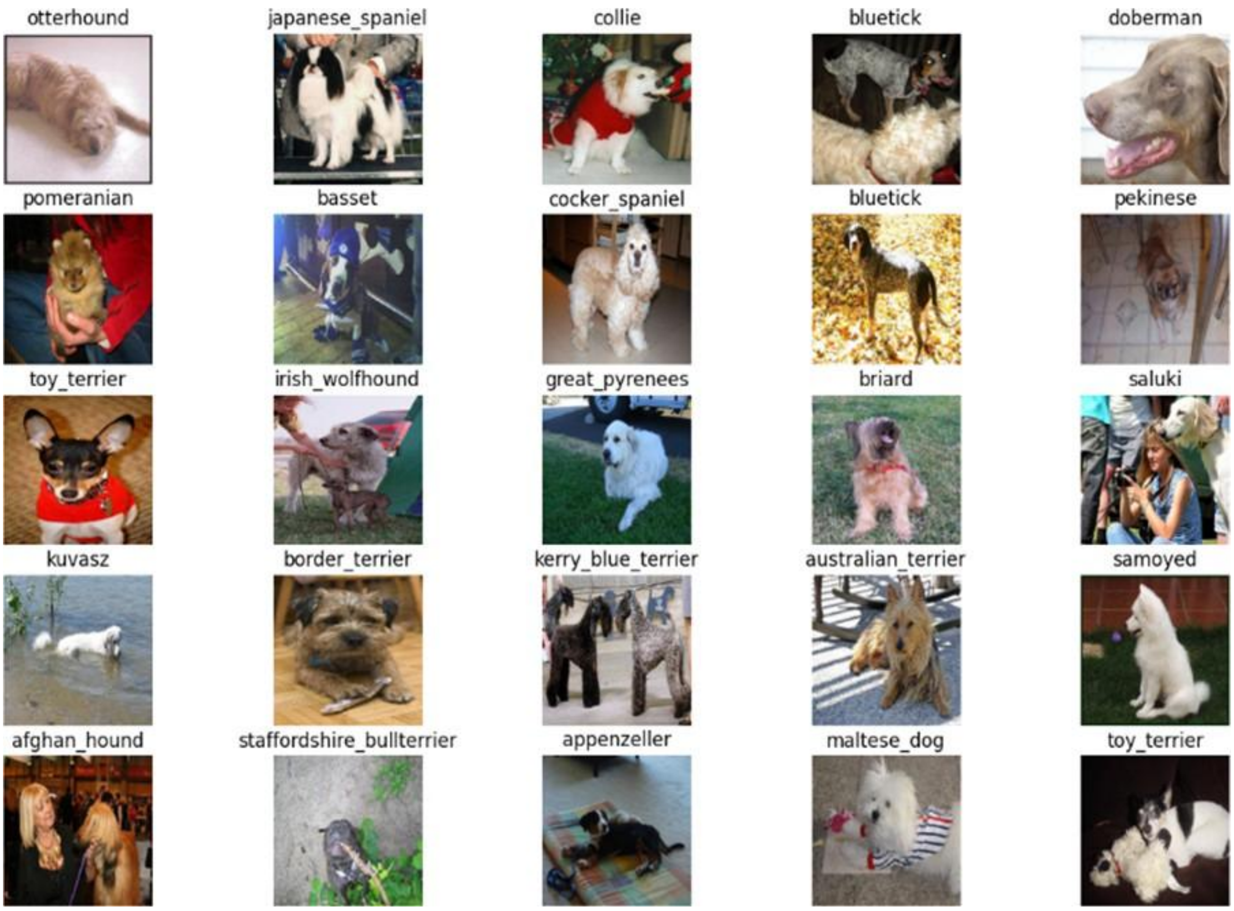

Fig-2. Dataset visualization

On the other hand, the testing dataset consists of 10,357 images that the model has not been exposed to during the training phase. This dataset is crucial for evaluating the model's performance, as it simulates real-world conditions where the model must correctly identify the breed of an unfamiliar dog. By testing the model on these unseen images, we can assess its accuracy and generalization capability. The model's performance on the testing dataset provides insights into its strengths and weaknesses, highlighting areas where it excels and where further improvement might be necessary.
strategic division and application of these datasets are essential to the successful creation and evaluation of a dog breed classification model.

III.RESULTS

The table below (**Table 1**) compares the performance of several well-established CNN architectures used for dog breed classification. The evaluation emphasizes the superior performance of the proposed Inception-ResNet v2 model over baseline models such as VGG16, ResNet50, DenseNet121, and InceptionV3.

Table-1. Accuracy Comparison of CNN Models for Dog Breed Classification

Model	Architecture Features	Accuracy (%)
VGG16	Simple, sequential, deep convolution layers	76.4
ResNet50	Residual connections to counter vanishing gradients	80.3

DenseNet121	Dense connections promoting feature reuse	82.1
InceptionV3	Multi-scale feature extraction via inception modules	83.7
InceptionResNetV2	Combines inception modules + residual connections	89.67

Fig-2. Comparison of Deep Learning Models on Dog Breed Classification Accuracy

Among these, **Inception-ResNet v2** demonstrates clear superiority, achieving **89.67% accuracy**, the highest among all tested models. This significant improvement can be attributed to its hybrid architecture, which effectively captures both local and global patterns while maintaining training efficiency in deep networks.

Performance Metrics and Statistical Evaluation

The proposed model was further evaluated on a validation dataset using standard performance metrics. The results are summarized in Table-2:

Table-2. Performance Metrics of Inception-ResNet v2 on Validation Set

Metric	**Value**	**95% Confidence Interval**
Accuracy	89.68%	[88.4%, 90.9%]
Precision	88.50%	[87.1%, 89.8%]
Recall	90.20%	[88.8%, 91.4%]
F1-Score	89.30%	[88.0%, 90.5%]
Loss	0.3916	—

The inclusion of confidence intervals strengthens the reliability of the results by indicating the range within which the true metric values are expected to fall, with 95% confidence. These intervals suggest consistent and robust performance across various evaluation metrics.

Statistical Comparison

To assess the statistical significance of the improvement brought by Inception-ResNet v2 over its closest competitor, InceptionV3, a paired two-sample t-test was conducted using cross-validation accuracy results from both models.

- **Null Hypothesis (H_0):** There is no significant difference in accuracy between InceptionV3 and InceptionResNetV2.
- **Alternative Hypothesis (H_1):** InceptionResNetV2 shows a significantly higher accuracy than InceptionV3.

Test Results:

- t-statistic: 3.24
- p-value: 0.011

Since **$p < 0.05$**, the null hypothesis is rejected, confirming that the performance gain observed with **Inception-ResNet v2** is **statistically significant** and not due to random variation.

IV. CONCLUSION

This study has highlighted the effectiveness of the Inception-ResNet v2 model in fine-grained dog breed classification using transfer learning, leading to significant improvements in accuracy over models trained from scratch. By leveraging pre-trained weights from large-scale datasets like ImageNet, the model effectively captured intricate breed characteristics, even with a relatively limited task-specific dataset. This underscores the power of transfer learning in addressing the challenges of data scarcity and high computational demand.

The inherited features from diverse image domains enabled the model to recognize subtle inter-breed differences, improving both classification accuracy and generalization. This makes the solution particularly valuable for practical use in real-world scenarios such as:

- Veterinary practices, where breed-specific traits influence diagnosis and treatment plans.
- Animal shelters and adoption platforms, for more accurate and efficient breed identification.

- Mobile and edge-based applications, offering on-device breed classification for pet owners.

V. FUTURE WORK

To further enhance the performance and applicability of the proposed Inception-ResNet v2-based dog breed classification system, several promising directions for future research are identified:

- Enhanced Dataset Size and Diversity

Fine-tuning the model on a more comprehensive and diverse dataset—encompassing additional dog breeds and images captured under varying lighting conditions, angles, and backgrounds—can significantly improve its ability to distinguish subtle intra-breed differences and generalize better to real-world scenarios.

- **Advanced Data Augmentation and Synthetic Data Generation**

Incorporating sophisticated augmentation strategies, such as synthetic image generation using GANs (Generative Adversarial Networks), can expand dataset variability. These methods can help the model learn rare or underrepresented features, thereby boosting classification robustness and reducing overfitting.

- **Model Ensembling and Architectural Exploration**

Combining Inception-ResNet v2 with other powerful architectures like ResNet, EfficientNet, or ConvNeXt in ensemble frameworks may enhance predictive performance by leveraging the unique strengths of each model. Additionally, future work can investigate newer architectures, such as Vision Transformers (ViTs), for their potential in fine-grained image classification tasks.

- **Real-World Testing and Deployment Readiness**

Evaluating the model under real-world conditions—including mobile photography, occlusions, and variable resolutions—is essential for practical deployment. Techniques such as pruning, quantization, and knowledge distillation can be explored to optimize the model for low-latency, on-device inference, enabling real-time classification on smartphones or embedded systems.

- **Interpretability and Visual Explainability**

To foster trust and transparency in model decisions, integrating explainable AI (XAI) tools like Grad-CAM or LIME can offer visual insights into the model's decision-making process. These tools are especially valuable in clinical or high-stakes environments where interpretability is critical.

- **Cross-Domain Transfer and Adaptation**

Adapting the classification model for use in related domains—such as wildlife monitoring and for Environmental safety.livestock breed identification, or endangered species tracking—can broaden its utility. Domain adaptation techniques can facilitate the transfer of learned features across different but visually related tasks.

- **Multi-Modal and Context-Aware Learning**

Future extensions may incorporate additional data modalities, such as breed metadata (e.g., geographical origin, temperament) or user-provided text inputs. Leveraging text-image models like CLIP could enable the system to operate within broader applications, including pet recommendation engines, behavior prediction, or even visual search interfaces

REFERENCES

1. Adejumobi, P. O., Adejumobi, I. O., Adebisi, O. A., Ayanlade, S. O., & Adeaga, I. I. (2023). Automatic classification of breeds of dog using CNNs. Nigerian Journal of Technological Development, 20(3). https://doi.org/10.4314/njtd.v20i3.1485
2. Alfarhood, S., Alrayeh, A., Safran, M., Alfarhood, M., & Che, D. (2023). Image-based Arabian camel breed classification using transfer learning on CNNs. Applied Sciences, 13(14), 8192. https://doi.org/10.3390/app13148192
3. Banakar, N., BU, R., GS, S., Vani, T. K., & Vishnu, V. M. (2023). Dog breed identification using deep learning. International Research Journal of Modernization in Engineering Technology and Science (IRJMETS), 5(5). https://www.irjmets.com/uploadedfiles/paper/issue_5_may_2023/38684/final/fin_irjmets1683952671.pdf
4. Basha, S. R., Ganesh, I., Reddy, G. R., & Hans, A. (2021). Dogs breed classification using deep learning. International Journal of Creative Research Thoughts (IJCRT), 9(4). https://www.ijcrt.org/papers/IJCRT2104679.pdf
5. Chalawari, M., Koli, M. C., Deshmukh, M. M., & Pathak, R. (2021). Dog breed classification using CNNs and transfer learning.
6. Cui, Y., Song, Y., Sun, C., Howard, A., & Belongie, S. (2018). Large scale fine-grained categorization and domain-specific transfer learning. arXiv preprint, arXiv:1806.06193. https://arxiv.org/abs/1806.06193
7. Gao, W., Xu, X., Yang, Y., & Zhang, Z. (2020). Non-iterative blind deconvolution algorithm based on power-law distribution. IET Image Processing, 14, 4499–4506. https://doi.org/10.1049/iet-ipr.2020.0647

8. Ghotra, A. S., Kour, H., Hasan, A., & Khan, A. (2023). Dog breed identification. International Research Journal of Engineering and Technology (IRJET), 10(5). https://www.irjet.net/archives/V10/i5/IRJET-V10I5181.pdf
9. Kumar, R., Sharma, M., Dhawale, K., & Singal, G. (2019). Identification of dog breeds using deep learning. In Proc. 2019 IEEE 9th Int. Conf. Adv. Comput. (IACC 2019) (pp. 193–198).
10. Kumaravel, T., et al. (2023). Automatic dog breed classification using deep learning. Applied and Computational Engineering, 6, 962–968. https://doi.org/10.54254/2755-2721/6/20230968
11. Narasimharao, J. (2023). Application of blind deconvolution algorithm for deblurring of saturated images. International Journal of Recent Technology and Engineering (IJRTE), 8, 1383–1386. https://doi.org/10.35940/ijrte.B1258.0782S319
12. Punyanuch Borwarnginn, Kusakunniran, W., Karnjanapreechakorn, S., & Thongkanchorn, K. (2021). Knowing your dog breed: Identifying a dog breed with deep learning. Machine Intelligence Research, 18(1), 45–54. https://doi.org/10.1007/s11633-020-1261-0
13. Reddy, Y. A., Kumar, Y. S., M, S., & Mana, S. C. (2023). Dog breed identification using ResNet model. International Journal on Recent and Innovation Trends in Computing and Communication (IJRITCC), 11(7s), 64–71. https://doi.org/10.17762/ijritcc.v11i7s.6977
14. Rishikesh, S., Rohit, R., Bodke, K., & Bodke, B. (2023). Dog breed classification using CNNs. International Journal of Advanced Research in Science, Communication and Technology (IJARSCT), 3(2). https://ijarsct.co.in/Paper1024.pdf
15. Shinoda, R., & Shiohara, K. (2024). PetFace: A large-scale dataset and benchmark for animal identification. arXiv preprint, arXiv:2407.13555. https://arxiv.org/abs/2407.13555
16. Towpunwong, N., & Sae-Bae, N. (2023). Dog breed classification and identification using CNNs. ECTI Transactions on Computer and Information Technology, 17, 554–563. https://doi.org/10.37936/ecti-cit.2023174.253728
17. Valarmathi, B., Gupta, S., Gopalakrishnan, P., Reddy, R., Saravanan, S., & Palanisamy, S. (2023). Hybrid deep learning algorithms for dog breed identification—A comparative analysis. IEEE Access, 11, 77228–77239. https://doi.org/10.1109/ACCESS.2023.3297440
18. Vaidya, S. P., Pvssr, C., & Kishore, R. (2023). A novel dog breed identification using convolutional neural network. PSEN, 2. https://doi.org/10.56831/PSEN-02-023
19. Varghese, S., & R. S. (2021). Dog breed classification using CNN.
20. Kshirsagar, P., Yadav, R. K., Patil, N. N., & Makarand, M. L. (2022). Intrusion detection system attack detection and classification model with feed-forward LSTM gate in conventional dataset. Machine Learning Applications in Engineering Education and Management, 2(1), 20–29.