Research Objectives

The goal of the research is to develop an automated system that can precisely identify breeds of dogs from images. Among its many useful applications include assisting owners in finding missing dogs, monitoring the health records of their pets, and offering personalized veterinary care. The work specifically focuses on creating a convolutional neural network (CNN) model that employs deep learning and transfer learning techniques to effectively classify dog photographs into the 133 different breeds included in the dataset.

Proposed Solutions and Research Methodologies/Network Architecture

There are five main steps in the methodology:

1. Importing datasets.

2. Utilizing a trained ResNet50 model to identify dogs in pictures.

3. Creating a CNN architecture with three convolutional layers followed by a max pooling layer each. A global average pooling layer was used at the end to convert each feature map from the max pooling layer to scalar.

4. Accuracy of 40%+ was attained by training the CNN from scratch for 250 epochs.

5. Making the most of transfer learning by incorporating global average pooling, dense layers, and bottleneck features from ResNet50 as model input.

With only 20 epochs, the model achieves 93.53% training accuracy and 90.89% validation accuracy enabled by the transfer learning approach. This shows how transfer learning can be used to retrain models efficiently, even with limited datasets.

Results

93.53% training accuracy and 90.86% testing accuracy are attained by the algorithm. For the majority of dog breeds, the confusion matrix and classification reports provide accurate classification. Results on test images that haven't been viewed before also indicate strong generalization ability. In the paper, training the CNN from scratch and transfer learning are compared. When trained from the start, the model only reached 40% accuracy after 250 epochs, but with transfer learning, it obtained over 90% accuracy in just 20 epochs. This demonstrates the benefits of using pre-trained models instead of training on small amounts of data.

Comparative Analysis

The research paper compares its CNN model’s performance with several models. Using the Stanfoed Dogs Dataset, Attention-based Model has achieved 76.80% accuracy and Depth-wise Seperable Colcolution yield 83.30%. On the other hand, Combination of Grey-scale SIFT Descriptors and Color Histogram Features approach, using the Columbia Dogs Dataset, achieved a breed classification rate of 67.00%. Deformable Part Model and Bag-of-Words Model reported 69.00% accuracy for a dataset of 37 breeds of cats and dogs. Using face geometries of different dog breeds Labelled Landmark Metadata with Grassmann Manifold model gave the accuracy of 96.50%. Also, Fisher Linear Projection and Preservation with One-Shot Similarity model using dog face identification reported a 96.87% rank 4 accuracy.

The suggested CNN model in the study paper is thought to be better than the models that were previously explored for a variety of reasons. Firstly, the model showed steady and adaptable performance over a variety of image sets, achieving excellent classification precision (93.53% for training and 90.86% for testing) across multiple datasets. Secondly, using pre-trained networks on large datasets, the model applies transfer learning techniques to increase accuracy and efficiency. This is especially useful in situations where training data is limited. Furthermore, the CNN model's architecture enables effective feature extraction and generalization, which boosts its adaptability in handling the diverse characteristics of different dog breeds. In contrast to some of the other complicated models discussed, the CNN model maintains a comparatively simpler and more efficient design while maintaining a high level of performance.

Conclusion

The study describes a successful CNN-based method that employs transfer learning to categorize 133 dog breeds with an accuracy rate of above 90%. The approach and findings have beneficial real-world implications. Accuracy can be further increased in the future by employing techniques like data augmentation, layer addition, model integration, dataset expansion, and multiple dog detection.