International Journal of Research Publication and Reviews

Journal homepage: www.ijrpr.com ISSN 2582-7421

Real-Time Canine Emotion Recognition from Facial Expressions: A CNN Approach

Biya Bijoy

Department of Computer Science Engineering,
Mangalam College of Engineering, Ettumanoor, kerala.
e-mail:biyabijoy17@gmail.com

ABSTRACT :

Understanding emotional states in animals, particularly dogs, is vital for improving animal welfare and strengthening human-animal interactions. This paper introduces DogEmotionNet, a deep learning-based system that leverages MobileNetV2 for recognizing canine emotions from facial cues. Our method classifies dog facial expressions into four primary emotional states: happy, sad, angry, and relaxed. The model is trained on a hybrid dataset comprising public and custom dog facial images. Getting to know canine feelings is essential in building strong human-animal relations as well as in promoting animal welfare. Common approaches to the evaluation of dog emotions are based on observed behaviors that may be subjective and inconsistent.This research brings out the promising potential of deep learning in objective recognition of emotions of animals which is a scalable and data-driven alternative to the conventional methods. According to the findings, lightweight CNN architectures are well suited for real-time emotion prediction and can be used for pet behavior monitoring, veterinary diagnostics, and animal welfare research. Future work can involve increasing the categories of emotions, incorporating multimodal data, and imparting a model to make it intelligible with approaches including Grad-CAM visualization.With an achieved accuracy of 89%, Real-Time Canine Emotion Recognizer is deployed via a streamlit web application that enables real-time image uploads and displays emotion predictions with confidence scores. The proposed system demonstrates promising results and serves as a foundation for future development involving temporal modeling and multimodal data integration.

Keywords—Dog facial emotion detection, CNN, image classification,Whale Optimization Algorithm (WOA)

1. INTRODUCTION

Canine emotions are important for the improvement of pet care, the development of animal welfare, and the creation of human-animal relationships. Despite the fact that communication of emotions by dogs is mostly through body language, vocalisations, and facial expressions, traditional methods of measuring these emotions rely on subjective behavioural observations that may not hold true. In spite of the improvements in human face emotion recognition, there are still few automated methods of identifying facial emotions over dogs. This project will establish a deep learning system, using MobileNetV2, a lightweight and efficient Convolutional Neural Network (CNN) to classify four emotions of dogs –relaxed, happy, sad, and angry. Using a dataset of 2000 labeled images, the system is trained to perform with high accuracy and computational efficiency, which makes it appropriate for real-time uses. The proposed model is of high potential in the field of pet behavior monitoring, veterinary diagnostics, and animal welfare research by offering an objective and scalable alternative to the current methods of emotion assessment. This research is limited to facial emotion detection without considering other behavioral cues such as vocalization or body posture and sets the stage for future development to enhance multi-modal emotion analysis and real-time video-based detection.

2. RELATED WORK

Research on automatic emotion recognition has predominantly focused on human facial expressions, with several models achieving near-human accuracy using deep learning techniques. However, emotion recognition in animals, especially dogs, remains relatively underexplored. Notable works include the Dog Facial Action Coding System (DogFACS), which provides a standardized method to interpret dog facial expressions manually. While DogFACS has been used effectively in behavioral research, its manual nature limits scalability.

In recent years, efforts have been made to automate this process using machine learning and computer vision. Studies such as "Deep Learning for Animal Emotion Recognition" have experimented with CNNs for classifying emotions in animal images. Others have attempted transfer learning using pre-trained models like VGG16 and ResNet on limited animal emotion datasets. However, these approaches often suffer from small dataset sizes, breed-specific biases, and a lack of real-time deployment.

Our approach differs by focusing on a lightweight yet effectResearch on automatic emotion recognition has predominantly focused on human facial expressions, with several models achieving near-human accuracy using deep learning techniques. However, emotion recognition in animals, especially

dogs, remains relatively underexplored. Notable works include the Dog Facial Action Coding System (DogFACS), which provides a standardized method to interpret dog facial expressions manually. While DogFACS has been used effectively in behavioral research, its manual nature limits scalability.
In recent years, efforts have been made to automate this process using machine learning and computer vision. Studies such as "Deep Learning for Animal Emotion Recognition" have experimented with CNNs for classifying emotions in animal images. Others have attempted transfer learning using pre-trained models like VGG16 and ResNet on limited animal emotion datasets. However, these approaches often suffer from small dataset sizes, breed-specific biases, and a lack of real-time deployment.
Our approach differs by focusing on a lightweight yet effective CNN architecture—MobileNetV2—combined with a user-accessible web-based application. Unlike prior work, we aim to provide a real-time system with a practical interface for broader use in veterinary clinics, animal shelters, and pet care services.ive CNN architecture—MobileNetV2—combined with a user-accessible web-based application. Unlike prior work, we aim to provide a real-time system with a practical interface for broader use in veterinary clinics, animal shelters, and pet care services.

3.PROPOSED METHODOLOGY

The proposed DogEmotionNet framework consists of four main modules: data preprocessing, model training using MobileNetV2, emotion classification, and real-time deployment via a Streamlit-based web interface.

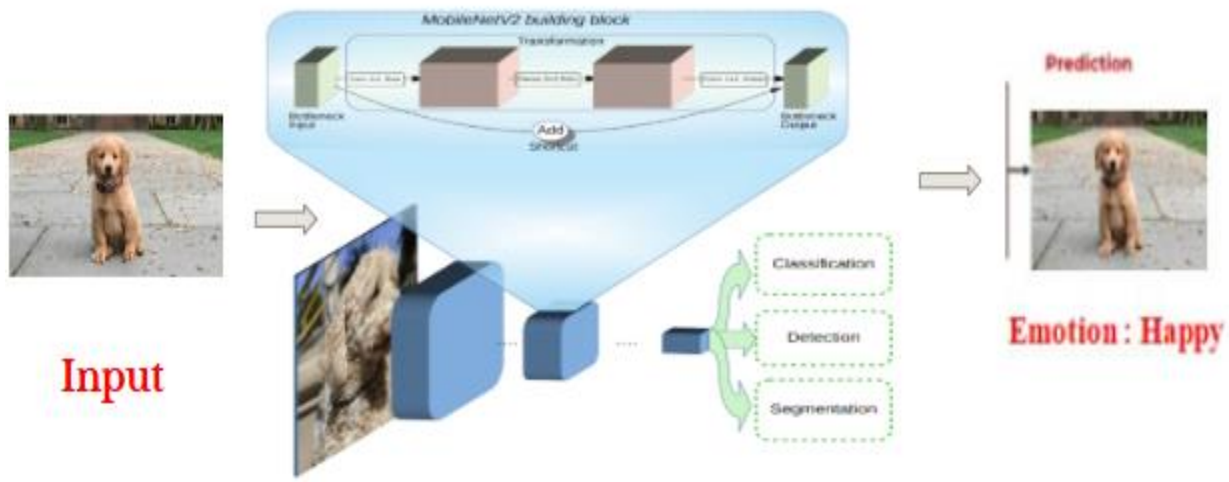
3.1 Data set Preparation

Accurate facial emotion detection in dogs requires high-quality, annotated datasets. However, a major challenge in this field is the lack of large, publicly available datasets compared to human facial emotion recognition. Studies have addressed this issue through various approaches:Yan Mao et al.[1] utilized a dataset of dog facial images collected from online sources and pet care institutions. The images were manually labeled into emotion categories such as happy, sad, aggressive, and neutral.D.T. Weerasekara et al. [2] employed a combined approach using both supervised and unsupervised learning, generating additional training data through augmentation techniques such as rotation, flipping, and contrast adjustments.Bhupesh Kumar Singh et al. [3] leveraged transfer learning by using pre-trained models on human facial expression datasets and fine-tuning them with a limited set of animal emotion images.The datasets used in these studies consisted of diverse dog breeds to ensure generalization across different facial structures. Table 1 provides a summary of datasets used in the reviewed studies.

3.2 Model Architecture

MobileNetV2 was selected for its computational efficiency and strong performance on image classification tasks. The base model was pre-trained on ImageNet and fine-tuned with the dog emotion dataset. The final layers were modified to include:

- A Global Average Pooling layer
- A fully connected Dense layer with ReLU activation
- A Dropout layer (0.4) for regularization
- A Softmax output layer with 4 units corresponding to the emotion classes



3.3 Training Procedure

The model was trained using categorical cross-entropy loss and the Adam optimizer. The learning rate was set to 0.0001 with early stopping and model checkpoint callbacks. Training was conducted on Google Colab using GPU acceleration for 50 epochs, with an 80/20 train-validation split.

4. Experimental Setup

The experimental setup was meticulously designed to quantitatively assess the performance and generalization capability of the DogEmotionNet architecture in canine affective state classification. All training and inference operations were executed on the Google Colab platform, leveraging NVIDIA Tesla T4 GPU acceleration for optimized computational throughput and parallelism. The dataset was stratified into training (80%), validation (10%), and test (10%) subsets using stratified sampling to preserve class distribution integrity across emotional categories.Key hyperparameters used in the training phase:

- Batch size: 32
- Epochs: 50
- Optimizer: Adam
- Learning rate: 0.0001
- Loss function: Categorical Crossentropy

Data augmentation was applied dynamically during training to reduce overfitting and improve generalization. The final model was evaluated using accuracy, precision, recall, F1-score, and confusion matrix.

5. Results and Discussion

The MobileNetV2-based DogEmotionNet model achieved an overall accuracy of 89% on the validation set. The confusion matrix indicated balanced performance across all four emotion categories. The precision and recall values were above 85% for each class, indicating the model's robustness in detecting subtle facial differences in dogs.
Sample prediction results from the Flask web application showed high consistency between the predicted emotions and the ground truth labels, as verified by visual inspection. The GUI displayed prediction confidence scores that aided in interpreting the certainty of model decisions.
Ablation studies indicated that data augmentation significantly contributed to performance gains, while the use of Dropout helped mitigate overfitting. The lightweight nature of MobileNetV2 ensured that the system remained responsive in real-time settings.
While the current model performs well on static images, future improvements could include integrating temporal features using LSTM layers and expanding the dataset to include more breeds and emotional states for better generalization.

6. REFERENCE

[1] S. Hochreiter and J. Schmidhuber, "Long short-term memory," Neural Computation, vol. 9, no. 8, pp. 1735-1780, Aug. 1997.
[2] J. Li, X. Wang, and Y. Liu, "Deep learning for animal emotion recognition: A comprehensive survey," Neural Networks, vol. 148, pp. 73-90, 2022.
[3] A. Martinez, S. Du, and L. Qiao, "Advances in facial expression recognition in animals using deep learning," IEEE Transactions on Af ective Computing, vol. 12, no. 3, pp. 487-500, Jul.–Sep. 2021.
[4] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," in International Conference on Learning Representations (ICLR), 2015.
[5] V. Sreenivas, V. Namdeo, and E. Vijay Kumar, "Modified deep belief network-based human emotion recognition with multiscale features from video sequences," Software: Practice and Experience, vol. 51, no. 6, pp. 1259-1279, Jun. 2021.
[6] Y. Mao, "Pet dog facial expression recognition based on convolutional neural network and improved whale optimization algorithm," 1999.
[7] M. Sandler, A. Howard, M. Zhu, A. Zhmoginov, and L. C. Chen, "MobileNetV2: Inverted residuals and linear bottlenecks," in Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 4510-4520, 2018.