International Journal of Engineering and Advanced Technology (IJEAT)
ISSN: 2249-8958 (Online), Volume-13 Issue-5, June 2024

Domestic Cats Facial Expression Recognition Based on Convolutional Neural Networks

Abubakar Ali, Crista Lucia Nchama Onana Oyana, Othman S. Salum

Abstract: ***Despite extensive research on Facial Expression Recognition (FER) in humans using deep learning technology, significantly less focus has been placed on applying these advancements to recognize facial expressions in domestic animals. Recognizing this gap, our research aims to extend FER techniques specifically to domestic cats, one of the most popular domestic pets. In this paper, we present a real-time system model that employs deep learning to identify and classify cat facial expressions into four categories: Pleased, Angry, Alarmed, and Calm. This innovative model not only helps cat owners understand their pets' behavior more accurately but also holds substantial potential for applications in domestic animal health services. By identifying and interpreting the emotional states of cats, we can address a critical need for improved communication between humans and their pets, fostering better care and well-being for these animals. To develop this system, we conducted extensive experiments and training using a diverse dataset of cat images annotated with corresponding facial expressions. Our approach involved using convolutional neural networks (CNNs) to analyze and learn from the subtleties in feline facial features by investigating the models' robustness considering metrics such as accuracy, precision, recall, confusion matrix, and f1-score. The experimental results demonstrate the high recognition accuracy and practicality of our model, underscoring its effectiveness. This research aims to empower pet owners, veterinarians, and researchers with advanced tools and insights, ensuring the well-being and happiness of domestic cats. Ultimately, our work highlights the potential of FER technology to significantly enhance the quality of life for cats by enabling better understanding and more responsive care from their human companions.***

Keywords: ***Facial Expression Recognition, Domestic Cats, CNN, Haar Cascade Classifier, Deep learning.***

I. INTRODUCTION

In recent years, interest has surged in using machine learning techniques, especially Convolutional Neural Networks (CNNs), for human facial expression recognition [1][33][34]35]. The findings from these studies show great promise in accurately identifying and categorizing facial expressions, opening up new possibilities for applications in veterinary science and animal behavior research. However, the majority of these studies have focused on species other than domestic cats, leaving a notable gap in the literature regarding automated facial expression recognition specifically tailored to felines [2]. Domestic cats, beloved companions to millions of people worldwide, possess a rich array of facial expressions that communicate various emotions and states of being (L. Dawson et al.) [3]. Understanding these expressions is not only fascinating from a behavioral standpoint but also crucial for enhancing the human-feline bond and potentially improving veterinary care. However, accurately interpreting feline facial expressions poses a significant challenge due to the subtle nuances and complexities involved [4].

Moreover, Despite the acknowledged importance of understanding feline facial expressions, the area of automated recognition and interpretation of these expressions is still largely uncharted [5]. Traditional methods for studying facial expressions in cats rely heavily on subjective human interpretation, which can be inconsistent and prone to biases. Additionally, the lack of standardized facial expression databases for cats further complicates the development of accurate recognition systems. One significant area requiring improvement in the field of domestic cat facial expression recognition is the development of robust CNN models trained on annotated datasets of feline facial expressions. Existing datasets for other species, such as humans and dogs, may not adequately capture the diverse range of facial expressions exhibited by cats [6]. Furthermore, Standardized protocols for capturing and annotating facial expressions are necessary to ensure consistency and reliability in training and evaluating CNN models. Overcoming these challenges is crucial for advancing the automated recognition of domestic cat facial expressions and realizing its full potential in various practical applications. This paper seeks to advance the growing body of research in this area by presenting a method for recognizing domestic cat facial expressions using Convolutional Neural Networks. We will detail the methodology for collecting and annotating a dataset of feline facial expressions, along with the implementation of the CNN model on our customized domestic cat dataset. Additionally, we will present experimental results demonstrating the effectiveness of our approach and discuss potential applications and future directions for research in this exciting area. So, our contributions to this paper are as follows:

- This study intends to contribute insights into the effectiveness of Recognizing Facial Expressions of Domestic Cats.

Manuscript received on 22 May 2024 | Revised Manuscript received on 28 May 2024 | Manuscript Accepted on 15 June 2024 | Manuscript published on 30 June 2024.

**Correspondence Author(s)*

Abubakar Ali*, School of Electrical Automation and Information Engineering, Tianjin University, Tianjin, China. Email: moumin0089@gmail.com, ORCID ID: 0009-0009-7153-4177

Crista Lucia Nchama Onana Oyana, School of Electrical Automation and Information Engineering, Tianjin University, Tianjin, China. Email: jessyluz04@gmail.com., ORCID ID: 0009-0008-2408-6669

Othman S. Salum, College of Software, Nankai University, Tianjin, China. Email: Salteensaleh@gmail.com, ORCID ID: 0009-0004-6496-1326

- Develop a Customized dataset of domestic Cats' facial expressions representing four emotional states which are angry, calm, pleased, and alarmed.
- Evaluate the performance of the proposed model and assess its potential applications in the facial expression recognition of domestic cats considering metrics such as accuracy, recall, precision, f1-score, Bar chart, and confusion matrices.
- Additionally, this research aims to highlight potential areas for improvement and future directions in the development of emotion detection in domestic animals and potential applications in animal welfare, behavior analysis, and human-animal interaction.

II. RELATED WORK

Lin and Kuo et al. [7] focus on individual cat identification. They train a CNN to detect the facial features of cats but for identification, they use conventional machine learning methods (SVMs and PCA). A tiny data set of 150 cats' 1,500 images is also used by them. From the publication, it is unclear whether the 94.1% claimed identification accuracy includes or excludes people or training images.

L. Xingxing et al. [8] offered an experimental evaluation of the system model's only on Cat face detection. They applied a Haar rectangular eigenvalue-integral graph and extracted features of a cat face. Then screened features for classification and recognition. Nevertheless, they applied the AdaBoost to alter a feeble classifier into a robust classifier that can effectually recognize cat faces. The use of the Gaussian Mixture Model with Mel-Frequency Cepstrum Coefficients for cat face recognition is highlighted by (Yu Fan and Chen et al., 2021) [9].

They used these approaches to identify the distinctive characteristics of cats and were utilized as an innovation to extract cat facial features after the Gaussian Mixture Model was created for each cat. For assessing the model's parameters, the maximum likelihood estimate is utilized. Although there are just 30 Cat faces in the dataset used for their study. An approach for cat recognition and identification that uses autoencoders combined with convolutional neural networks (CNN) was proposed by (P.Chen et al.,2021) [10].

They also produced a brand-new dataset with 1,994 images of 17 cats. Additionally, they provided a thorough explanation of how to use Autoencoder to denoise cat image data and combine it with CNN to produce a powerful model for the same cat recognition for future study.

III. METHODS

The proposed system models consist of a Convolutional Neural Network (CNN) by classifying domestic cat facial expression recognition. Figure 1 below, summarize the process of classifying the cat's emotions from the input image. It comprises the following steps, image processing, feature extraction, and classification. all of the steps are very significant for analyzing models for Cat facial expression recognition.

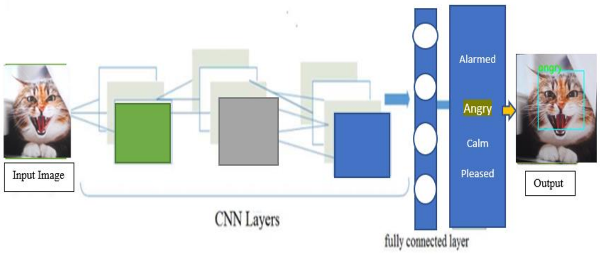


Figure 1. The Process Diagram Representation of Classifying Domestic Cats' Facial Expression Recognition

A.Dataset Preparation

In this study, we collected the domestic cats faces images with an input of a resolution of 64x64. The Dataset is a meticulously organized collection of cat facial expressions, categorized into four distinct emotional states: 'calm', 'alarmed', 'angry', and 'pleased' as illustrated by Figure 2 below. This dataset serves as a valuable resource for researchers and developers interested in exploring the nuances of feline emotional expression and behavior. Each image in the dataset captures the unique facial features and expressions exhibited by domestic cats in various contexts.

In addition, before training the deep learning model, the dataset undergoes several pre-processing steps to enhance model performance and facilitate efficient training [11]. These pre-processing steps ensure robust model training the steps are as follow. Image resizing: Resizing images to a uniform resolution to ensure consistency and reduce computational complexity during training, in this study, we resize the images to 64 x 64. Data augmentation: we applied various augmentation techniques such as rotation, flipping, cropping, and brightness adjustment to increase the diversity and size of the dataset, thereby improving model generalization. Flip: Horizontal, Vertical, Crop: 0% Minimum Zoom, 20% Maximum Zoom, Rotation: Between -15° and +15°, Grayscale: Apply to 15% of images, Brightness: Between -15% and +15%, and Noise: Up to 0.1% of pixels . Furthermore, as part of the preprocessing pipeline, all images in the dataset have undergone normalization to enhance model convergence and performance. Normalization ensures that the input data is standardized, reducing the impact of variations in pixel intensity and improving the overall stability of the training process [12]. Moreover, a Domestic cats customized dataset of a total of experimental numbers for each expression is shown in **Table 1** below.

Figure 2. Illustration of the Proposed Facial Expression Distribution, the Pictures Convey A Variety of Feelings

Retrieval Number: 100.1/ijeat.E448413050624
DOI: 10.35940/ijeat.E4484.13050624
Journal Website: www.ijeat.org

Published By:
Blue Eyes Intelligence Engineering
and Sciences Publication (BEIESP)

Table 1. The Experimental Number of Customized cat Dataset

Expression Label	Cat dataset
Alarmed	199
Angry	200
Calm	200
Pleased	200

B. Proposed Model Architecture

Our proposed model's structure is a pure convolutional network. The architecture of the model employed in our project is shown in Figure 3 below.

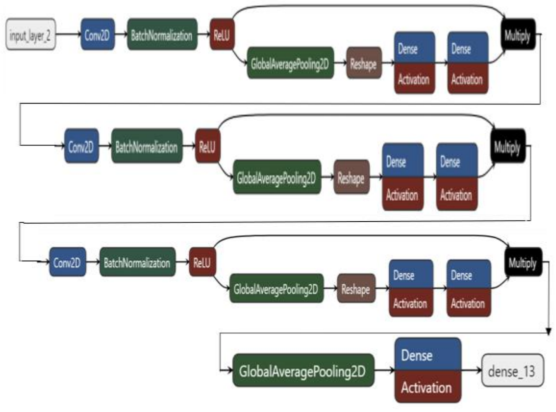


Figure 3. The Proposed Model Architecture is a Series of Convolutional Layers Connected with the Squeeze-and-Excitation Block, Leading to a SoftMax Classifier that Outputs the Probability Value

The first layer of the Proposed model architecture is the Convolution Layer [13], it extracts the significant features from the input image which is denoted as pixel values in the form of a matrix. In the convolution operation, the filter is the first part involved and the stride parameter controls how the filter moves across the image. In addition, the filter goes over the image one pixel at a time when the step is 1, and two pixels at a time when the step is 2 [14]. Additionally, the convolution process is calculated by multiplying the two matrices, where the first matrix is the input image and the second matrix is the filter or kernel (Zeiler & Fergus, 2014) [15]. The input image is the result of adding each element of the input image to its neighbor's weights. Moreover, the activation/feature map is the name of the convolution operation's output. Mathematically, the convolution operation for a 2D input can be expressed as:

$$(I * K)(i,j) = \sum_m \sum_n I(i+m, j+n) \cdot K(m,n) \quad (1)$$

Here, ***I*** is the input data. ***K*** is the convolutional filter (kernel). (i, j) represents the position in the output feature map.(m, n) iterates over the filter dimensions.

Furthermore, the output of the convolution process is known as the activation or feature map and the most frequently practical activation function is the ReLU layer (Rectified Linear Unit) which is applied to nonlinearity into the output(Krizhevsky et al., 2017) [16]. The ReLU function returns to 0 if all of the input values are negative, and to (x) if all of the input values are positive. The result is a rectified feature map. The ReLU activation function is provided by equation (2).

$$f(x) = \max(0, x) \quad (2)$$

Moreover, the Spatial pooling layer is applied to diminish the number of parameters in a big image and it is also known as up-sampling and down-sampling. Nevertheless, the purpose of this layer is to reduce the size of the feature map while maintaining the important portions of the image (S. Lawrence et al., 1997) [17]. There are three (3) main types of spatial pooling, the first type is max pooling which is used to take the biggest component of the corrected activation map. The mathematical formula for Max Pooling can be defined as follows:

$$\text{Max Poling}(x,s)_{i,j} = \max_{m,n}\left(x_{i\cdot s+m,\, j\cdot s+n}\right) \quad (3)$$

Here: **x** is the input volume. ***s*** is the stride of the pooling operation, representing the step size with which the pooling window moves. ***i*** and ***j*** are the indices of the output feature map.***m*** and ***n*** iterate over the spatial dimensions of the pooling window. In simpler terms, for each position (i, j) in the output feature map, the Max Pooling operation looks at the region in the input volume defined by the size of the pooling window and the stride. It then selects the maximum value from that region

The second type is average pooling (Szegedy et al., 2015) [18], which is practical for selecting the average values of the element in the rectified feature map. The third type is Sum pooling which is used to sum all available elements in the rectified feature map. Additionally, a crucial technique for lowering the measurement after the convolution layer is this layer. Average Pooling is another commonly used pooling operation, and it calculates the average value of the elements in the pooling window. The formula for Average Pooling is:

$$\text{Average Pooling}\,(x,s)_{i,j} = \frac{1}{m\cdot n}\sum_{m,n}\left(x_{i\cdot s+m,\, j\cdot s+n}\right) \quad (4)$$

Here, the average is taken over all the elements in the pooling window. In both cases, the pooling operation helps reduce the spatial dimensions of the input volume, making it computationally more efficient while retaining important features. The choice between Max Pooling and Average Pooling often depends on the specific requirements of the task and the characteristics of the data.

Also, Convolution layer expansion can exacerbate the issue of a generic pooling technique that explores feature maps (M. Jogin et al.,2018) [19]. In addition to the above, after completing the technique of spatial pooling. After the 2-dimensional matrix is flattened into a linear-continuous long vector in the following layer, it is passed to a fully connected layer for classification.

Finally, fully connected layer denotes every neuron being linked to every other neuron. This completely linked layer receives the output from the flattened layer, which transforms the two-dimensional matrix into a long vector (S. Lawrence et al., 1997) [20]. After training, the feature vector from this layer is subsequently used to divide images into distinct classes. Moreover, in this layer, all the inputs are linked to every activation part of the next layer. Since all the parameters are occupied in the fully connected layer, it causes overfitting. So, dropout is one of the methods that diminishes overfitting.

Retrieval Number: 100.1/ijeat.E448413050624
DOI: 10.35940/ijeat.E4484.13050624
Journal Website: www.ijeat.org

Published By:
Blue Eyes Intelligence Engineering
and Sciences Publication (BEIESP)

After passing through the stage of this layer, the SoftMax activation function is the final layer which is applied to obtain the object's probability based on input that falls into particular classes (S. Shalev-Shwartz et al.) [21].

C.Implementation and Training Model Parameters

a. Implementation

We innovated our model by adding The Squeeze-and-Excitation (SE) block. it is a novel architectural unit designed to improve the representational power of convolutional neural networks by enabling them to perform dynamic channel-wise feature recalibration. Introduced by (Jie Hu, and Gang Sun, 2018) [22], in their paper "Squeeze-and-Excitation Networks," the SE block aims to enhance the quality of the feature maps generated by a CNN. Figure 4 provides an illustration of the fundamental structure of the SE block.

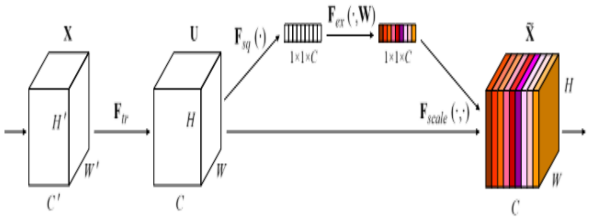


Figure 4. Squeeze-and-Excitation (SE) Block [22]

The core idea is to explicitly model the interdependencies between the channels of the convolutional features. This process involves two main operations: squeeze and excitation. In the "squeeze" step, global average pooling is applied to each feature map independently, producing a channel descriptor that captures global spatial information. This step reduces each 2D feature map into a single numerical value, summarizing the entire spatial extent. Following this, the "excitation" step consists of two fully connected (dense) layers that form a self-gating mechanism. The first dense layer reduces the number of channels by a reduction ratio (typically set to 16), capturing the channel interdependencies with a non-linear transformation. The second dense layer restores the original number of channels. The output of this second layer is passed through a sigmoid activation function to generate a set of weights between 0 and 1. These weights act as adaptive recalibration parameters that scale the original feature maps channel-wise [22]. By emphasizing important features and diminishing less useful ones, the SE block allows the network to focus more on informative features, thereby improving the overall performance. The entire SE block operation can be summarized as:

$$\mathbf{X'}_{i,j,c}=\mathbf{X}_{i,j,c}\cdot\sigma2(\mathbf{W}2\sigma1(\mathbf{W}1\mathbf{z}))c \qquad (5)$$

where $\mathbf{z}$ is the squeezed vector obtained from global average pooling, $\mathbf{W}1$ and $\mathbf{W}2$ are the weight matrices of the two fully connected layers, and $\sigma1$ and $\sigma2$ are the ReLU and sigmoid activation functions, respectively. The equation 5 describes how the SE block adaptively recalibrates the channel-wise feature responses by leveraging global information, thus enhancing the representational capacity of the network.

b. Training Model Parameters

The model parameters in the context of neural networks refer to the settings that define the architecture and behavior of the network, such as learning rate, batch size, optimizer, etc. [23]. The following hyperparameters were applied. a learning rate of 0.001 was used during training. Correspondingly, the batch size of 64 was applied, Batch size refers to the number of samples processed in one forward and backward pass during training. Also, the Categorical Cross-Entropy Loss Function was applied during training. It measures the dissimilarity between the predicted probability distribution and the true distribution of class labels [24]. In addition, for the model optimization, The Adam optimizer was applied in this study, it combines the benefits of both momentum and RMSprop optimization techniques. The algorithm is known for its efficiency, fast convergence, and robustness to different types of data [25]. Table 2. Display the experimental parameters applied in this study on Customized Cats datasets.

Table 2. Experimental Model Parameters on Cats Dataset

Learning Rate	Batch Size	Optimizer	Image Size
0.001	64	Adam	64x64

IV. RESULT AND DISCUSSION

A.Result

The Experimental Result of the proposed model for domestic cats' facial expression recognition was evaluated using a set of evaluation metrics, including precision, recall, F1 score, training accuracy, training loss, and a confusion matrix. We can observe that the model achieved good results in the different facial expressions. As can be observed from Table 3 which reveals the performance metrics results values in all four domestic cats' facial expressions with their support values.

Table 3. The Proposed Model Evaluation Results on Domestic Cats Dataset

Metrics / Cats Data	Precision	Recall	F1-Score	Support
Alarmed	0.99	0.95	0.97	199
Angry	0.98	0.99	0.99	200
Calm	0.97	0.98	0.97	200
Pleased	0.99	0.99	0.99	200
Accuracy			0.98	799
Macro avg	0.98	0.98	0.98	799
Weighted avg	0.98	0.98	0.98	799

Precision metric measures the proportion of correctly predicted positive cases out of all cases predicted as positive as expressed in equation 1 below [26].

$$\text{Precision} = \frac{\text{True Positives}}{\text{True Positives} + \text{False Positives}} \qquad (1)$$

For the Alarmed expression, the precision is very high at 0.99, indicating that when the model predicts an instance as Alarmed, it is correct 99% of the time.

Retrieval Number: 100.1/ijeat.E448413050624
DOI: 10.35940/ijeat.E4484.13050624
Journal Website: www.ijeat.org

Published By:
Blue Eyes Intelligence Engineering
and Sciences Publication (BEIESP)

Similarly, for Angry and Pleased expressions, the precision values are also high, at 0.98 and 0.99 respectively, indicating high accuracy in identifying these expressions. Moreover, the recall metric measures the proportion of correctly predicted positive cases out of all actual positive cases which is expressed in Equation 2 [27].

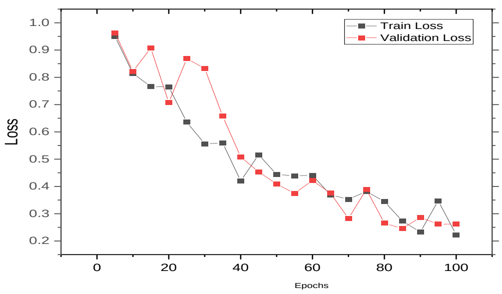
$$\text{Recall} = \frac{\text{True Positives}}{\text{True Positives} + \text{False Negatives}} \quad (2)$$

For the Alarmed expression, the recall is 0.95, implying that the model correctly identifies 95% of all actual Alarmed instances. For Angry, Calm, and Pleased expressions, the recall values are notably higher, indicating that the model performs very well in capturing instances of these expressions. Furthermore, the f1-score is the harmonic mean of precision and recall, providing a balanced measure that considers both false positives and false negatives as given in equation 3 below [28].

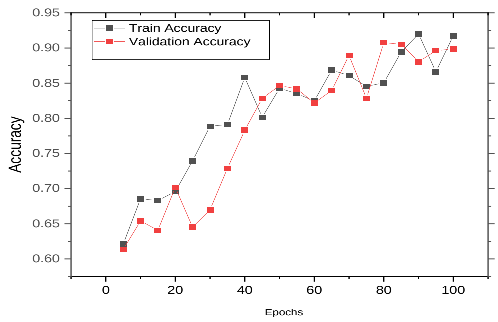
$$f1 - score = \frac{2\times \text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \quad (3)$$

Across all classes, the f1-scores are consistently high, ranging from 0.97 to 0.99, indicating overall strong performance of the model in terms of both precision and recall. Lastly, the accuracy metric measures the overall correctness of the model's predictions across all classes. With an accuracy of 0.98, it indicates that the model correctly predicts the facial expressions in the customized dataset with a high degree of accuracy.

The training accuracy and loss graphs illustrate the model's learning progression over epochs. They show how the accuracy of the model improves and the loss decreases as the training proceeds. Figure 5 reveals the training loss and accuracy of the proposed model in the customized domestic cats' dataset, the suggested model's validation accuracy is shown alongside its training accuracy.



(a)



(b)

Figure 5. The Proposed Model Loss (a) and Accuracy (b) Training Graphs in the Customized Cat's Dataset

Furthermore, we used the bar chart to present a comparison of the model's performance across different facial expression categories. It visualizes the distribution of correct and incorrect predictions for each expression category, allowing for a comparative analysis of the model's accuracy and misclassification rates [29]. Figure 6 illustrates the bar chart for achieving the proposed model performance metrics in four domestic cats' expressions. The model achieved excellent results in all four expressions.

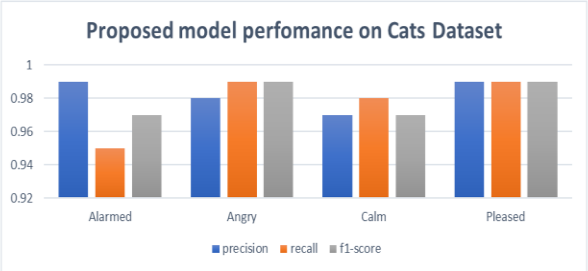


Figure 6. The Bar chart of Emotions in the Customized cat's Dataset and their Precision, Recall, and f1-score by the Proposed Model

Finally, we assessed the proposed model performance by using a confusion matrix. It displays the distribution of true positive, false positive, true negative, and false negative predictions across different expression categories, offering insights into the model's classification performance [30]. Figure 7 displays the normalized confusion matrix results evaluation of the proposed model on the Customized Cat's dataset.

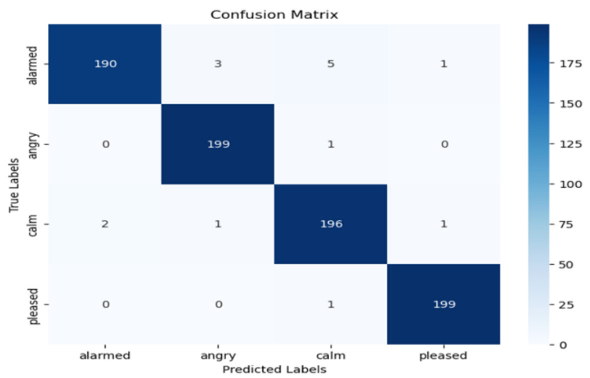


Figure 7. The Heat Map of the Confusion Matrix on the Customized cat's Dataset by the Proposed Model

B. Domestic Cats Facial Expression Recognition Real-time Results by Proposed Model

Real-time Domestic cats' facial expression recognition result is portrayed in Figure 8, the model showcases its agility by accurately identifying and categorizing a spectrum of feline expressions, spanning from 'pleased' and 'happy' to 'alarmed' and 'calm'. This adaptation of the proposed model highlights its versatility, extending beyond conventional applications to tackle the complex domain of feline emotional expression recognition.

Retrieval Number: 100.1/ijeat.E448413050624
DOI: 10.35940/ijeat.E4484.13050624
Journal Website: www.ijeat.org

Published By:
Blue Eyes Intelligence Engineering
and Sciences Publication (BEIESP)

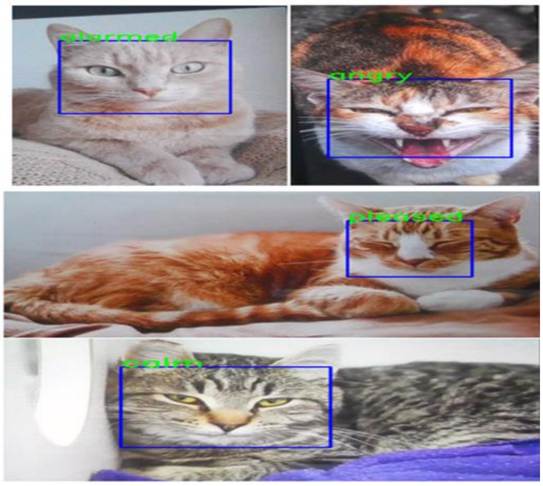


Figure 8. Domestic Cats Real-time Facial Expressions Recognition Results by Proposed model

C. Discussion

The evaluation results of the proposed model on domestic cat's dataset offer valuable insights into the model's performance in classifying emotions within each animal category. The accuracy of the model demonstrates its overall effectiveness in correctly classifying emotions for domestic cats, with accuracies of 98%. While this accuracyy indicate solid performance, it also suggests room for refinement, especially considering the complexities of interpreting animal emotions. Our proposed method may benefit from the introduction of Squeeze-and-Excitation (SE) blocks in our model architecture which enhanced the representational power of a network by enabling it to perform dynamic channel-wise feature recalibration. Table 4. Shows the accuracy results comparison of our customized network with existing studies.

Table 4. The Results Comparison of our Proposed Method with Existing Studies

Method	Optimizer	Accuracy
Proposed Method	Adam	98%
Inception-v3 [31]	Adam	80.42%
Resnet50[32]	Adam	91%

V. CONCLUSION AND FUTURE WORK

A. Conclusion

Our study provides contribution insights into deep learning applications in understanding animal behavior, especially in the context of facial expressions in domestic cats, which is relatively unexplored. These results suggest that the Proposed model is well-suited for the task of recognizing domestic facial expressions, with a particularly strong aptitude for identifying expressions of anger, calmness, and pleasure. Such performance metrics indicate the model's potential applicability in various real-world scenarios where understanding and interpreting human emotions from facial expressions are crucial, such as in human-computer interaction, emotion recognition systems, and mental health assessment tools. Also, this research contributes to the growing field of animal emotion recognition and opens avenues for developing applications focused on enhancing human-animal interaction and welfare.

However, it's important to note that while these results are promising, further validation and testing in diverse and real-world contexts would be necessary to assess the model's robustness and generalizability. Nonetheless, based on the evaluation report, the Proposed model demonstrates a high level of proficiency in recognizing domestic facial expressions, laying a solid foundation for its potential deployment in practical applications aimed at understanding and responding to human emotions effectively.

B. Future Work

One of the primary areas for future work involves expanding and diversifying the datasets used for training and evaluation. While the current datasets provide valuable insights, they may not fully represent the diversity of emotional expressions and variations within different breeds of cats and dogs. Future work could involve collecting larger and more diverse datasets, encompassing a wider range of breeds, ages, and environmental conditions. Additionally, including more subtle variations in emotional states and expressions would enable the models to generalize better across different scenarios and populations.

Also, by combining different patterns using deep learning algorithms to Investigate the potential benefits of combining Squeeze-and-Excitation (SE) blocks and Exploration of other machine learning techniques or data sources to improve the accuracy and robustness of cat facial expression recognition in a hybrid approach to leverage the strengths.

ACKNOWLEDGMENT

We would like to acknowledge our main supervisor for this project, Prof. Yang Jiachen. He led us to the amazing world of deep learning projects, specifically image processing, guided our research project, and helped us mature. He is the mentor we need, and he provides guidance and advice wherever and whenever we desire. Thank you, Prof. Yang Jiachen, for giving us the opportunity to embark on this great journey we have taken as your students. Also, we would like to thank our fellow laboratory members Xiao Shuai, Lan Guipeng, and Darous K. Joboe for their mentoring during this research project, and last but not least, we would like to acknowledge the School of Electrical Automation and Information Engineering of Tianjin University and College of Software of Nankai University for their continuous support.

DECLARATION STATEMENT

Funding	No, We did not receive any financial support for this article.
Conflicts of Interest	No conflicts of interest to the best of our knowledge.
Ethical Approval and Consent to Participate	No, the article does not require ethical approval and consent to participate with evidence.
Availability of Data and Material	Not relevant.
Authors Contributions	All authors have equal participation in this article.

Retrieval Number: 100.1/ijeat.E448413050624
DOI: 10.35940/ijeat.E4484.13050624
Journal Website: www.ijeat.org

Published By:
Blue Eyes Intelligence Engineering
and Sciences Publication (BEIESP)

REFERENCES

1. C. Dalvi, M. Rathod, S. Patil, S. Gite, and K. Kotecha, “A Survey of AI-Based Facial Emotion Recognition: Features, ML & DL Techniques, Age-Wise Datasets and Future Directions,” IEEE Access, vol. 9, pp. 165806–165840, 2021, doi: 10.1109/ACCESS.2021.3131733. https://doi.org/10.1109/ACCESS.2021.3131733
2. S. Li and W. Deng, “Deep Facial Expression Recognition: A Survey,” IEEE Trans. Affective Comput., vol. 13, no. 3, pp. 1195–1215, Jul. 2022, doi: 10.1109/TAFFC.2020.2981446. https://doi.org/10.1109/TAFFC.2020.2981446
3. L. Dawson, “Are you a cat whisperer? How to read Fluffy’s facial expressions,” The Conversation. Accessed: Oct. 10, 2023. [Online]. Available: http://theconversation.com/are-you-a-cat-whisperer-how-to-read-fluffys-facial-expressions-128686
4. “Emotions Body Language,” Purina.com.au, 2023. https://www.purina.com.au/cats/behaviour/emotions-body-language (accessed Aug. 26, 2023).
5. C. P. Udeh, L. Chen, S. Du, M. Li, and M. Wu, “Multimodal Facial Emotion Recognition Using Improved Convolution Neural Networks Model,” Journal of Advanced Computational Intelligence and Intelligent Informatics, vol. 27, no. 4, pp. 710–719, Jul. 2023, doi: 10.20965/jaciii.2023.p0710. https://doi.org/10.20965/jaciii.2023.p0710
6. Azizi, E., & Zaman, L. (2023). Deep learning pet identification using face and body. *Information*, 14(5), 278. https://doi.org/10.3390/info14050278
7. T.-Y. Lin and Y.-F. Kuo, “<i> Cat face recognition using deep learning</i>,” in 2018 Detroit, Michigan July 29 - August 1, 2018, American Society of Agricultural and Biological Engineers, 2018. https://doi.org/10.13031/aim.201800316
8. L. Xingxing, “Cat Face Detection Based on Haar Cascade Classifier,” in 2019 2nd, CSP, 2019.
9. Y. Fan, C.-C. Yang, and C.-T. Chen, “Cat Face Recognition Based on MFCC and GMM,” in 2021 6th International Conference on Image, Vision and Computing (ICIVC), Qingdao, China: IEEE, Jul. 2021, pp. 81–84. https://doi.org/10.1109/ICIVC52351.2021.9527024
10. P. Chen, A. X. Qin, and J. Lu, “Cat Recognition Based on Deep Learning,” HAL Archives Ouvertes, Dec. 01, 2021. https://hal.science/hal-03501010/ (accessed Aug. 26, 2023).
11. W. Setiawan, Moh. I. Utoyo, and R. Rulaningtyas, “Classification of neovascularization using convolutional neural network model,” TELKOMNIKA (Telecommunication Computing Electronics and Control), vol. 17, no. 1, p. 463, Feb. 2019, doi: https://doi.org/10.12928/telkomnika.v17i1.11604
12. T. N. Tran, B. M. Lam, A. T. Nguyen, and Q. B. Le, “Load forecasting with support vector regression: influence of data normalization on grid search algorithm,” International Journal of Electrical and Computer Engineering (IJECE), vol. 12, no. 4, p. 3410, Aug. 2022, doi: https://doi.org/10.11591/ijece.v12i4.pp3410-3420
13. C. P. Udeh, L. Chen, S. Du, M. Li, and M. Wu, "Multimodal Facial Emotion Recognition Using Improved Convolution Neural Networks Model," Journal of Advanced Computational Intelligence and Intelligent Informatics, vol. 27, no. 4, pp. 710–719, Jul. 2023, https://doi.org/10.20965/jaciii.2023.p0710
14. A.-L. Cîrneanu, D. Popescu, and D. Iordache, “New Trends in Emotion Recognition Using Image Analysis by Neural Networks, a Systematic Review,” Sensors, vol. 23, no. 16, p. 7092, Aug. 2023, https://doi.org/10.3390/s23167092
15. M. D. Zeiler and R. Fergus, “Visualizing and Understanding Convolutional Networks,” Computer Vision – ECCV 2014, pp. 818–833, 2014, https://doi.org/10.1007/978-3-319-10590-1_53
16. A. Krizhevsky, I. Sutskever, and G. E. Hinton, “ImageNet classification with deep convolutional neural networks,” Communications of the ACM, vol. 60, no. 6, pp. 84–90, May 2017, https://doi.org/10.1145/3065386
17. S. Lawrence, C. L. Giles, Ah Chung Tsoi, and A. D. Back, "Face recognition: a convolutional neural-network approach," in IEEE Transactions on Neural Networks, vol. 8, no. 1, pp. 98-113, Jan. 1997, https://doi.org/10.1109/72.554195
18. C. Szegedy et al., "Going deeper with convolutions," 2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Boston, MA, USA, 2015, pp. 1-9, https://doi.org/10.1109/CVPR.2015.7298594
19. M. Jogin, Mohana, M. S. Madhulika, G. D. Divya, R. K. Meghana and S. Apoorva, "Feature Extraction using Convolution Neural Networks (CNN) and Deep Learning," 2018 3rd IEEE International Conference on Recent Trends in Electronics, Information & Communication Technology (RTEICT), Bangalore, India, 2018, pp. 2319-2323, https://doi.org/10.1109/RTEICT42901.2018.9012507
20. S. Lawrence, C. L. Giles, Ah Chung Tsoi, and A. D. Back, "Face recognition: a convolutional neural-network approach," in IEEE Transactions on Neural Networks, vol. 8, no. 1, pp. 98-113, Jan. 1997, https://doi.org/10.1109/72.554195
21. S. Shalev-Shwartz and S. Ben-David, Understanding Machine Learning. Cambridge University Press, 2014. https://doi.org/10.1017/CBO9781107298019
22. Hu, Jie, Li Shen, and Gang Sun. "Squeeze-and-excitation networks." In Proceedings of the IEEE conference on computer vision and pattern recognition, pp. 7132-7141. 2018. https://doi.org/10.1109/CVPR.2018.00745
23. K. Thavasimani and N. Kasturirangan Srinath, “Hyperparameter optimization using custom genetic algorithm for classification of benign and malicious traffic on internet of things–23 dataset,” International Journal of Electrical and Computer Engineering (IJECE), vol. 12, no. 4, p. 4031, Aug. 2022, https://doi.org/10.11591/ijece.v12i4.pp4031-4041
24. S. Gupta, P. Kumar, and R. K. Tekchandani, “Facial emotion recognition based real-time learner engagement detection system in online learning context using deep learning models,” Multimedia Tools and Applications, vol. 82, no. 8, pp. 11365–11394, Sep. 2022, doi: https://doi.org/10.1007/s11042-022-13558-9
25. D. P. Kingma and J. Ba, "Adam: A method for stochastic optimization," in International Conference on Learning Representations (ICLR), 2015.
26. J. Davis and M. Goadrich, "The relationship between Precision-Recall and ROC curves," in Proceedings of the 23rd international conference on Machine learning, 2006, pp. 233-240. https://doi.org/10.1145/1143844.1143874
27. S. Ravidas and M. A. Ansari, “Deep learning for pose-invariant face detection in unconstrained environment,” International Journal of Electrical and Computer Engineering (IJECE), vol. 9, no. 1, p. 577, Feb. 2019, https://doi.org/10.11591/ijece.v9i1.pp577-584
28. M. Sokolova, N. Japkowicz, and S. Szpakowicz, "Beyond accuracy, F-score and ROC: a family of discriminant measures for performance evaluation," in Proceedings of the Twenty-Second International Conference on Machine Learning, 2005, pp. 947-954. https://doi.org/10.1007/11941439_114
29. L. Rasheed, U. Khadam, S. Majeed, S. Ramzan, M. S. Bashir, and M. M. Iqbal, “Face Recognition Emotions Detection Using Haar Cascade Classifier and Convolutional Neural Network,” In Review, preprint, Nov. 2022. https://doi.org/10.21203/rs.3.rs-2048290/v1
30. R. B. Kala, N. S. Gill, and A. K. Verma, "A comparative analysis of classification algorithms using confusion matrix for intrusion detection system," in 2016 2nd International Conference on Applied and Theoretical Computing and Communication Technology (iCATccT), 2016, pp. 547-552.
31. R. Zhang, "Classification and Identification of Domestic Cats based on Deep Learning," in 2021 2nd International Conference on Artificial Intelligence and Computer Engineering (ICAICE), pp. 106-110, IEEE, 2021. https://doi.org/10.1109/ICAICE54393.2021.00029
32. Wu, Yujie. "Emotion Detection of Dogs and Cats Using Classification Models and Object Detection Model." 2023.
33. Kumari, J., Patidar, K., Saxena, Mr. G., & Kushwaha, Mr. R. (2021). A Hybrid Enhanced Real-Time Face Recognition Model using Machine Learning Method with Dimension Reduction. In Indian Journal of Artificial Intelligence and Neural Networking (Vol. 1, Issue 3, pp. 12–16). https://doi.org/10.54105/ijainn.b1027.061321
34. P A, J., & N, A. (2022). Faceium–Face Tracking. In Indian Journal of Data Communication and Networking (Vol. 2, Issue 5, pp. 1–4). https://doi.org/10.54105/ijdcn.b3923.082522
35. Maddileti, T., Rao, G. S., Madhav, V. S., & Sharan, G. (2019). Home Security using Face Recognition Technology. In International Journal of Engineering and Advanced Technology (Vol. 9, Issue 2, pp. 678–682). https://doi.org/10.35940/ijeat.b3917.129219

AUTHORS PROFILE

Abubakar Ali, Received B.Eng. degree in Electronics and Communication Engineering from St. Joseph University in Tanzania, Dar es Salaam Tanzania in 2014. He is currently pursuing an M.S. degree in Information and Communication Engineering at the School of Electrical Automation and Information Engineering, Tianjin University, Tianjin, China.

Published By:
Blue Eyes Intelligence Engineering and Sciences Publication (BEIESP)

His research interests include image processing and pattern recognition, Computer Vision, Deep learning, and wireless communications.

Crista Lucia Nchama Onana Oyana, Received a B.S. in Measurement Technology and Control Instrumentation from Hebei University of Technology, Tianjin, China in 2020. She is currently pursuing an M.S. degree at the School of Electrical Automation and Information Engineering, Tianjin University, Tianjin, China. Her research interest includes facial emotion detection, computer vision and pattern recognition, and wireless communications.

Othman Saleh Salum, Received a B.S. in Information Technology from the Institute of Finance Management (IFM), Dar es Salaam Tanzania in 2012. He is currently pursuing a M.S. degree in Software Engineering at Nankai University, Tianjin, China. His research interests include face analysis, computer vision, and image processing & pattern recognition.

Retrieval Number: 100.1/ijeat.E448413050624
DOI: 10.35940/ijeat.E4484.13050624
Journal Website: www.ijeat.org

Published By:
Blue Eyes Intelligence Engineering
and Sciences Publication (BEIESP)