© FEB 2024 | IRE Journals | Volume 7 Issue 8 | ISSN: 2456-8880

Cat Breed & Emotion Detection Using Yolo, CNN & Canny Edge Detection

MITHILESH VISHWAKARMA[1], KUNIKA JANGID[2], KIRAN PAL[3], LISA RODRIGUES[4]

[1]*Assistant Professor, Department of IT, Thakur College of Science and Commerce, Thakur Village, Kandivali (East), Mumbai, Maharashtra, India*

[2, 3, 4] *PG Student, Department of IT, Thakur College of Science and Commerce, Thakur Village, Kandivali (East), Mumbai, Maharashtra, India*

Abstract— Automatic cat facial expression recognition is actively emerging research in This study explores the recognition of emotional states in cats through their facial expressions, drawing inspiration from the extensive research conducted on human facial expressions. We propose a hybrid approach utilizing Convolutional Neural Networks (CNNs) and Canny edge detection to identify and classify cat facial emotions. To mitigate overfitting in the CNN model, we employ a regularization technique known as "dropout" in the fully connected layers. Additionally, we extend our system's capabilities by incorporating cat breed detection using the YOLO (You Only Look Once) Model. Our system demonstrates an impressive average accuracy rate of 87.00% in recognizing basic emotional states in cats, effectively classifying them into predefined categories. Furthermore, we discuss the potential for deploying software applications based on this methodology on various platforms such as mobile devices and computers, making it a versatile tool for real-world applications in pet behaviour analysis and beyond.

Indexed Terms— Animal-Human Interaction, Cat Breed Detection, Canny Edge Detection, Cat Emotion Recognition, Cat Facial Expression, Computer Vision, Convolutional Neural Network (CNN), Deep Learning, Emotion Classification, Facial Emotion Classification, Human-Animal Bond, Image Analysis, Machine Learning, Pet Behavior Analysis, Pet Psychology, YOLO Model.

I. INTRODUCTION

Facial Introduction:

Facial expressions play a crucial role in conveying human emotions, with researchers even identifying signs of pain recognition in various animals based on their "pain face." These expressions communicate a significant portion of emotional information, surpassing what can be conveyed through voice and language. Typically categorized into six basic emotions-anger, disgust, fear, happiness, sadness, and surprise-facial expressions have become a central focus in fields like human-computer interaction, social robotics, and data-driven animation. Recognizing and understanding these expressions is vital, not only for humans but also for animals, such as cats.

The first part of our research centers on developing an algorithm for detecting cat emotions using Convolutional Neural Networks (CNNs), which have demonstrated exceptional performance in a range of applications, from talent management to biomedical detection. We explore two approaches for recognizing cat emotions: one based on image sequences capturing the transition from neutral to peak expressions and the other based on static images, where peak expression images are analyzed without temporal information. This research is pivotal in understanding the emotional states of cats, as different expressions often correspond to distinctive behaviors or emotions.

Moving on to our second research focus, we delve into the intricate world of dog breed recognition. Identifying dog breeds can be a challenging task due to their shared body types and structures, as well as the significant diversity within and across breeds. The project aims to leverage the Yolo model detector to address this challenging fine-grained classification problem. The selection of this task is motivated by the substantial inter-variability among dog breeds and the need for accurate results, even when faced with diverse stylistic variations in the images within the dataset.

In summary, this research explores two fascinating areas-cat emotion recognition and dog breed classification. We employ cutting-edge techniques like CNNs and the Yolo model to better understand and classify emotional states in cats and differentiate dog breeds accurately, despite the challenges posed by these diverse and intricate subjects.

II. LITERATURE REVIEW

Emotion recognition from facial expressions has been a prominent research area, with several studies exploring its feasibility using Convolutional Neural Networks (CNNs). In Ninad Mehendale's research on "Facial Emotion Recognition using Convolutional Neural Networks," the author demonstrated the potential to detect emotions from images. This approach utilized the expressional vector (EV) to identify five distinct regular facial expressions. The process involved face localization, fiducial feature identification (eyes, nose, and lips), and feature extraction techniques. CNNs were employed to extract image features, which were then classified using another neural network. The study emphasized the subjective nature of expression and observed more uniform displays of certain emotions. In our paper, we leverage CNNs to extract and classify different cat emotions, achieving a higher accuracy of 93% compared to previous methods.

In the context of breed classification for dogs, Bickey Kumar Shah, Aman Kumar, and Amrit Kumar conducted research focused on converting images into single-dimensional labels with image processing. They applied principal component analysis to group similar features, simplifying feature analysis in deep neural networks. Their work aimed to classify test images based on minimal weight differences between test and training images.

Genomic data analysis of cat breeds was explored by Heidi G. Parker, Dayna L. Dreger, and their collaborators. They studied a diverse dataset comprising 1,346 dogs representing 161 different breeds. Their research identified unique clad formations supported by strong bootstrap values, with some breeds not significantly grouping with any other.

J.E. Hayes, S.L. Forbes, and others delved into cat detection and its connections to physiology, training, and analytical methodologies. Their work provided insight into the interrelationship between scientific methodologies and the detection dog industry. They conducted an integrated assessment of the factors involved to determine the current and future status of detection dogs.

Greg S. Baxter and colleagues compiled scientific literature to present essential behavioral and physical traits relevant to wildlife detection dogs. While their focus was on wildlife detection, the proposed traits have applicability in various detection fields.

These studies collectively illustrate the diverse applications and methodologies within the realm of image-based emotion recognition, cat breed classification, genomic analysis of cat breeds, and the role of detection dogs in various scientific disciplines. Our research extends this body of knowledge by focusing on cat emotion recognition and dog breed classification, leveraging CNNs and the YOLO model for accurate and efficient classification tasks.

III. ALGORITHM

Convolutional Neural Network

Convolutional Neural Network is a class of deep neural networks, most commonly applied to analyze visual imagery. Neural networks have been demonstrated to differentiate faces, objects and traffic signs better than humans, they are utilized in robots and self-driving cars. This model is made up of two parts: the hidden layers, where the features are extracted and fully connected layers which are used for the actual classification task when the processing is complete. CNN is built upon each point of data and then will be fitted to these layers so that our model can learn high-level hierarchical features from our data.

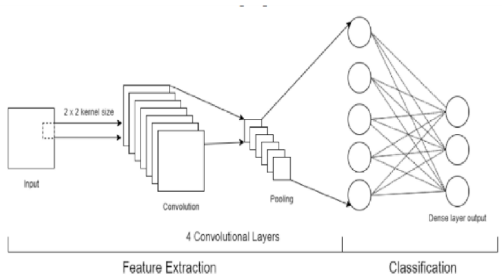


Figure 1: Convolutional neural network model used in this study

In this study our model will be built with four convolutional layers, with the output layer as a dense layer. A filter matrix with a 3x3 dimension is also included. A dense layer means that that it is a simple layer of neurons in which each neuron receives input from all the neurons of previous layer, thus called as dense. Dense layer is used to classify image based on output from convolutional layers. Working of single neuron. A layer contains multiple number of such neurons.

For output layer the input layer will receive the input shape. The input shape is based on the Images of cat or basically its emotion which will be transformation of our data. Here we will be taking maximum number of frames. For our data we have 120 frames on each layer. So that is why we will be applying zero paddings on the output of transformation to equalize the number of frames. So, our layer will be 2 dimensional layers.

The convolutional kernel will inspect and extract the features from the images. The kernel is a matrix that moves over the input data, performs the dot product with the sub-region of input data and gets the output as the matrix of dot products. The small size of the layer compared to our input size will allow our model to learn the pattern in more localized way and more through in detecting the presence of specific elements in the emotions of cats.

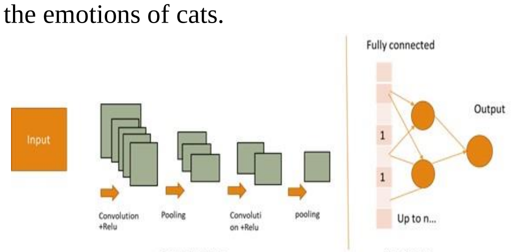


Figure 2: Illustration of the proposed Convolutional Neural Network (CNN)

Following with CNN we have added Canny Edge Detection. Canny edge detection algorithm was developed by John F.Canny in 1986. It is a multistage algorithm and we will go through each stage.

1. Smoothening (Noise Reduction)

Since edge detection is susceptible to noise in the image in the image, first step is to remove the noise with 5x5 Gaussian filter.

2. Finding Derivatives (Gradient of the Image)

Smoothened image is then filtered with a Sobel kernel in both horizontal and vertical direction to get first derivative in horizontal direction (Gx) and vertical direction (GY). From these two images, we can find edge gradient and direction for each pixel as follows:

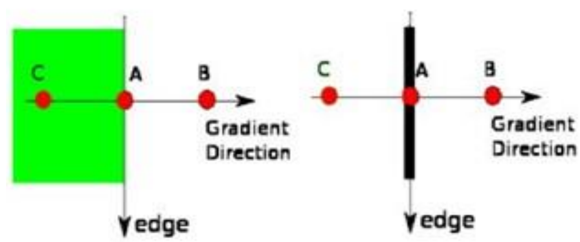
$$Edge\_Gradient(G) = \sqrt{Gx^2 + Gy^2}$$

$$(\theta) = \tan^{-1}\left(\frac{Gx}{Gy}\right)$$

Gradient direction is always perpendicular to edges. It is rounded to one to four angles representing vertical, horizontal and two diagonal directions

3. Non-maximum Suppression

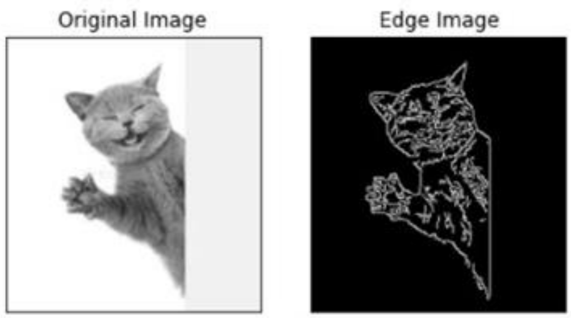
After getting gradient magnitude and direction, a full scan of image is done to remove any unwanted pixels which may not constitute the edge. For this, at every pixel, pixel is checked if it is a local maximum in its neighborhood in the direction of gradient. Check the image below:



Point A is on the edge (in vertical direction). Gradient direction is normal to the edge. Point B and C are in gradient directions. So, point A is checked with point B and C to see if its forms a local maximum. If so, it is considered for next stage otherwise it is suppressed (put to zero). In short, the result you get is binary image with "thin edges".

4. Hysteresis Thresholding

This stage decides which are all edges and which are not. For this we need two threshold values, minVal and maxVal. Any edges with intensity gradient more than maxVal area sure to be edges and those below minVal are sure to be non-edges so discarded. Those who lie between these two thresholds are classified edges or non-edges based on their connectivity. If they are connected to "sure-edge" pixels, they are considered to be part of edges. Otherwise, they are also discarded. See the image below:

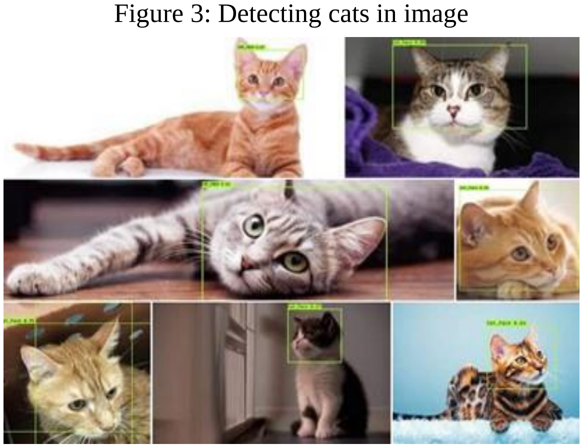


Here in our project, we have used OpenCV which puts all the above in single function, cv2.Canny().

Yolo Detector

YOLO which is also known as 'YOU ONLY LOOK ONCE' is the most powerful object detector. It is called the most powerful object detector because unlike the old object detector like CNN or its upgraded CNN it only needs the images and videos to pass one time to its network. This previous methodology had successively examined the multiple regions of the image to find the object that is present in it. YOLO changed that by giving reasoning at the attribute of the whole image to foresee multiple boxes, one and all containing a specific object. All this simultaneously.

Figure 3: Detecting cats in image



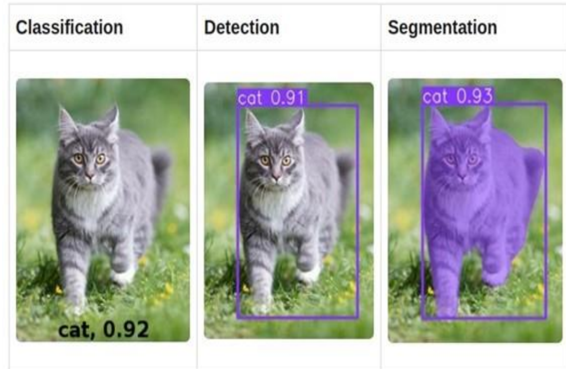


Figure 4: Process of classification detection and segmentation

The above figure 4 is an example that represents the cats that are classified according to their breeds with the help of yolo algorithm. Here the image is divided into S X S region. Then, if the center of the object is in one of these regions, the region in question is responsible for detecting the object. Each one of the cells in the grid is for forecast 'B' boxes all accommodating an object as well as an outcome representing the level of confidence for the object present in the box. If there are no objects in the cell, this outcome would be zero and if an object is in the cell the outcome would be equal to the intersection past union (IoU) between the anticipated box and the ground truth of the picture. The higher than fig.3. Every bounding box comprises predictions: x, y, w, h, and confidence ($c_{1,}$ $c_{2,}$ c_3). The (x, y) coordinates constitute the middle of the box corresponding to the bounds of the grid cell. The breadth and height area unit foretold corresponds to the total image. Finally, the boldness prediction represents the promissory note between the expected box and any ground truth box. Every grid cell additionally predicts C conditional category probabilities. Then we want the category to specific confidence outcome for every box that is completed employing a convolutional neural supported by the Google internet network. The result of this algorithmic rule is going to be a picture and video because the input with the object is localized and therefore the category connected to that. The result for object detection is excellent.

Dataset

The images of Cats that are used in this research are taken from Google having the KL grading system ranging from Grade 0 to Grade 4. The dataset consists a total of 200 Images which were later divided into training and testing data for further processing using Convolutional neural network model. We have collected the various emotions of the cats so accordingly our system i.e., CNN (Convolutional Neural Network) detects the emotion of the cat. Further our project is doing is that it not only detects the emotions with along with emotions it also detects the edges of cat which is our specialty or can say the uniqueness of our project. Here are some of the sample data's shown below:

IV. METHODOLOGY

Images Materials and Methods

A. Cat Dataset for Classification:

To address the practical application of cat emotion detection, we curated a comprehensive dataset focused on cat emotions. This dataset includes images representing various emotional states of cats, emphasizing three key perspectives: head-on, profile, and left views of cat faces. In our study, we concentrated on recognizing three primary cat emotions: happiness, sadness, and sleepiness. Each of these emotions is briefly described below:

1. Happiness: Cats typically exhibit happiness when provided with nutritious and palatable food, such as meat, fish, eggs, and other protein sources. A well-balanced diet not only promotes their physical health but also contributes to their emotional well-being.
2. Sadness: While cats do not verbally communicate their emotions, they express sadness through various visual and vocal cues. Visual signs may include changes in posture, such as a slouched or tense body language, while vocalizations like moans or yowls might indicate distress.
3. Sleepiness: Cats, known for their love of sleep, may display sleepiness through behaviors like yawning, stretching, or changing their ear positions. Their eyes can also convey their mood, with slow blinking often indicating contentment.

B. Image Pre-processing:

Image pre-processing techniques were employed to enhance the quality of the cat images and remove any noise or artifacts that could affect the accuracy of emotion detection. These techniques sharpened the images and improved their overall quality, making them more suitable for species identification and emotion categorization. It's important to note that the datasets used in our research had previously undergone pre-processing and segmentation steps to ensure data quality.

C. Image Augmentation:

Emotion detection involves several key steps, illustrated in Figure 1. Image augmentation played a crucial role in expanding our training dataset. By applying image augmentation techniques to the cat images, we increased the dataset's size, enhancing the model's ability to recognize and classify different cat emotions. Shape characteristics extracted from the augmented images were used in conjunction with a convolutional neural network (CNN) for emotion classification.

YOLO Model:

We employed the YOLO (You Only Look Once) algorithm for the detection of dog breeds. YOLO is renowned for its exceptional performance and high frames per second (fps) for real-time applications. Unlike traditional object detection methods that rely on multiple passes through an image, YOLO predicts classes and bounding boxes for the entire image in a single run.

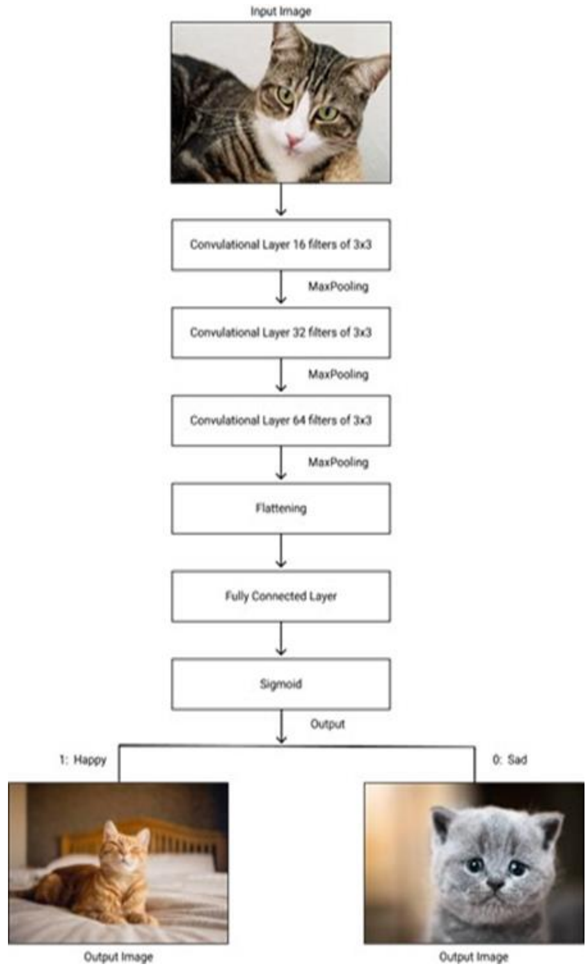
Our approach to training the YOLO model involved the following steps:

1. Image Pre-processing: The initial step aimed to enhance the image quality by eliminating distortions and accentuating crucial features. This ensured that the computer model could effectively learn from the dataset.

2. Object Detection: During this phase, we segmented the images and identified the positions of objects, in this case, dog breeds.
3. Feature Extraction and Training: Deep learning techniques were employed to extract and classify distinctive patterns in the images that could be unique to specific dog breeds. The model learned these features from the dataset, a process known as model training.
4. Categorization of Objects: The final step involved categorizing the detected objects into predefined classes using suitable classification techniques that compared the image with target patterns.

In our study, we collected datasets from various sources and divided them into training (70%), validation (20%), and testing (10%) sets, comprising 350 training images and 100 validation images. We conducted the training process using Google Collaboratory with Python, leveraging libraries such as torch and utils. The model was trained with a batch size of 16, and no additional schemes were applied as the dataset size was deemed sufficient. The bounding boxes were adjusted to match the size of anchors during the initial three epochs of training.

FLOWCHART:



V. RESULT

In this work, 200 validation images were used to train the model. These photos were trained using CNN model with a variety of optimizer techniques, including loss categorical cross-entropy, RMS, and an optimizer with a learning rate of 0.001. The Comparison of the proposed model's accuracy and loss performance for each optimizer employed after training with 15 iterations (epochs) is shown in Table 1. The best accuracy and loss performance is provided by CNN and Canny. In comparison to the conventional, untrainable approaches, the supervised learning model is better. Table 1 and Figure 2 contains the results of epochs and shows how the model identifies images.

Table I. Accuracy Assessment

Epochs	Accuracy	Validation Accuracy
Epoch 1/10	0.5556	0.4505
Epoch 2/10	0.3333	0.5856
Epoch 3/10	0.4444	0.4505
Epoch 4/10	0.6667	0.6036
Epoch 5/10	0.4444	0.5676
Epoch 6/10	0.6667	0.5225
Epoch 7/10	0.5556	0.5856
Epoch 8/10	0.3333	0.7477
Epoch 9/10	0.6667	0.7568
Epoch 10/10	0.7778	0.6577
Epoch 11/10	0.5741	0.7692
Epoch 12/10	0.6714	0.6277
Epoch 13/10	0.7142	0.8198
Epoch 14/10	0.8771	0.7699
Epoch 15/10	0.8771	0.7126

The result for cat breeds detection is as follow:

Table II. Accuracy assessment.

Epochs	Accuracy	Validation Accuracy
Epochs 1/10	0.5874	0.6
Epochs 2/10	0.4856	0.559
Epochs 3/10	0.7591	0.525
Epochs 4/10	0.4856	0.547
Epochs 5/10	0.7452	0.660
Epochs 6/10	0.4785	0.676
Epochs 7/10	0.8546	0.4750
Epochs 8/10	0.4856	0.7854
Epochs 9/10	0.8541	0.8657
Epochs 10/10	0.8645	0.7506
Epochs 11/10	0.8650	0..8745
Epochs 12/10	0.8784	0.7557
Epochs 13/10	0.8900	0.7451
Epochs 14/10	0.9584	0.8821

OUTPUT:

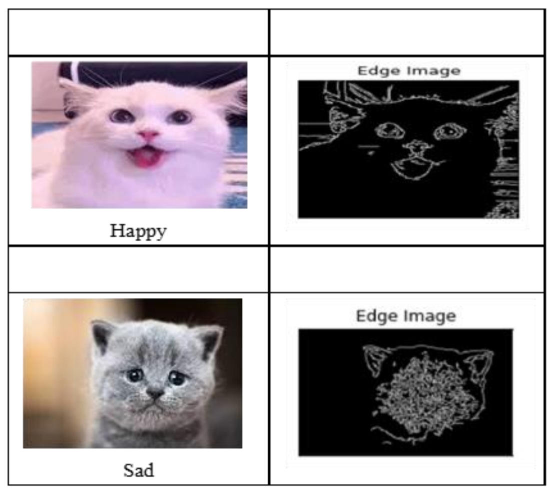


Figure 5: Model of accuracy using Canny Edge Detection and Deep learning

CONCLUSION

In this research paper, we have studied lots of research papers then we compare those research papers. Many researchers are researching a paper on facial expression of humans, detection of cats and so on. In this research we have proved the facial emotion of a cat using convolution neural network (CNN) with data augmentation. The two emotions that we have proved is happy and sad i.e., also known as unique emotion of Cats. Here the CNN model employed in this work has three hidden layers, three fully connected layers and Sigmoid activation. Three hidden layers utilize 3 x 3 filter sizes with 16, 32, and 64 channel outputs in succession. Testing shows that the CNN model with Canny edge detection presented does the best job of classifying the dataset of cats with 87% accuracy. The method shows that the suggested model is promising to use as a current tool for professionals for testing the emotions of cats. System to categorize the numerous emotions of numerous animals could be created with more study. This work generated an automated detection system for cat breeds. The (detection) is done among Seven breeds of cats. The Yolo model trained our dataset in 4 steps: first object detection, second creating boundary boxes, third object segmentation, and lastly data augmentation. The images were 250 X 250 pixels. After testing results showed that the Yolo model v5 worked in the detection of the dataset of cat breeds results with 100% accuracy. This proposed method shows that the suggested model is promising to use as at today's time there are many breeds of cats which look similar so it becomes difficult to make difference between them due to which many people get confused so to avoid this confusion it is good to use this mode. They can use this model in pet shops also so that they can't get fooled by the pet shop owner.

REFERENCES

[1] Mehendale N. Facial emotion recognition using convolutional neural networks (FERC). SN Applied Sciences. 2020

[2] Dalal N, Triggs, B Histograms of oriented gradients for human detection. In Sept. 21 2005 to Sept. 23 2005 San Diego, CA, USA 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05).

[3] Modi S, Bohara MH. Facial emotion recognition using convolution neural network. In2021 5th International Conference on Intelligent Computing and Control Systems (ICICCS) 2021 May 6 (pp. 1339-1344). IEEE.

[4] Guillaume Mougeot, Dewei Li, and Shuai Jia. A deep learning approach for dog face verification and detection. In abhaya c Nayak and Alok Sharma, editors PRICAI 2019. Trends in Artifical Intelligence, pages 418-430, Cham 2019.

[5] Cat Face Detection using OpenCV. https://blogs.oracle.com/meena/cat-facedetection-using-opencv. [Online; accessed 09-October-2019].

[6] Cat facial detection and landmark recognition in Python. https://github.com/marando/ pycatfd. [Online; accessed 10-October-2019].

[7] F.D Torre, W.-S. Chu, X. Xiong, F. Vicente, X. Ding, J. Cohn, IntraFace. In Proceedings of the IEEE International Conference on Automatic Face and Gesture Recognition, Ljubljana, Slovenia, 48 May 2015; pp. 18

[8] Xu Z, Baojie X, Guoxin W. Canny edge detection based on Open CV. In2017 13th IEEE international conference on electronic measurement & instruments (ICEMI) 2017 Oct 20 (pp. 53-56). IEEE.

[9] H. Sikkandar, R. Thiyagarajan. Deep learning based facial expression recognition using improved Cat Swarm Optimization. Journal of Ambient Intelligence and Humanized Computing.2021 Feb. https://link.springer.com/article/10.1007/s12652-020-02463-4

[10] Marsot M,Mei J,Shan X, YeL,FeNG P.An adaptive pig face recognition approach using convolutional Neural Netowkrs. Computers and Electronics in Agriculture.2020 Jun 1. https://www.sciencedirect.com/science/article/abs/pii/S0168169920300673

[11] M.H. Siddiqi, R. Ali, A.M Khan, Human facial expression recognition using stepwise linear discriminant analysis and hidden conditional random fields. IEEE Trans. Image Proc. 2015, 24, 13861398. [CrossRef] [PubMed].

[12] Alreshidi A, Ullah M. Facial emotion recognition using hybrid features. In Informatics 2020 Feb 18 (Vol. 7, No. 1, p. 6). MDPI.

[13] Kumar, Aman, and Amrit Kumar. "Dog breed classifier for facial recognition using convolutional neural networks." 2020 3rd International Conference on Intelligent Sustainable Systems (ICISS). IEEE, 2020.

[14] Parker, Heidi G., et al. "Genomic analyses reveal the influence of geographic origin, migration, and hybridization on modern dog breed

development." Cell reports 19.4 (2017): 697-708.Tizard IR, Jones SW.

[15] Hayes, J. E., et al. "Critical review of dog detection and the influences of physiology, training, and analytical methodologies." Talanta 185 (2018): 499-512.

[16] La Toya, J. Jamieson, Greg S. Baxter, and Peter J. Murray. "Identifying suitable detection dogs." Applied Animal Behaviour Science 195 (2017): 1-7.

[17] Lin, Xu-Hui, et al. "Ancylostoma ceylanicum Infection in a Miniature Schnauzer Dog Breed." Acta Parasitologica (2022): 1-5.

[18] Andrade, Joao PB, et al. "Dog Face Recognition Using Deep Feature Embeddings." Available at SSRN 4175201.

[19] Bertolo, Alessandro, et al. "Canine mesenchymal stem cell potential and the importance of dog breed: implication for cell-based therapies." Cell Transplantation 24.10 (2015): 1969-1980.

[20] Grimm-Seyfarth, Annegret, Wiebke Harms, and Anne Berger. "Detection dogs in nature conservation: A database on their world-wide deployment with a review on breeds used and their performance compared to other methods." Methods in Ecology and Evolution 12.4 (2021): 568-579.

[21] Sinnott, Richard O., Fang Wu, and Wenbin Chen. "A mobile application for dog breed detection and recognition based on deep learning." 2018 IEEE/ACM 5th International Conference on Big Data Computing Applications and Technologies (BDCAT). IEEE, 2018.

[22] Ráduly, Zalán, et al. "Dog breed identification using deep learning." 2018 IEEE 16th International Symposium on Intelligent Systems and Informatics (SISY). IEEE, 2018.

[23] Borwarnginn, Punyanuch, et al. "Breakthrough conventional based approach for dog breed classification using CNN with transfer learning." 2019 11th International Conference on Information Technology and Electrical Engineering (ICITEE). IEEE, 2019.

[24] Wang, Changqing, et al. "Dog Breed Classification Based on Deep Learning." 2020 13th International Symposium on Computational Intelligence and Design (ISCID). IEEE, 2020.