CAT BREED AND EMOTION DETECTION USING YOLOv8 AND CNN

B. SRIVIGNA

G. NANDHINI

CH. PREMA KUMARI

B. VARSHITHA REDDY

GUIDE

MS. SAMREEN BEGUM

ASSISTANT PROFESSOR

1. Abstract

This project presents a deep learning-based system designed to automatically detect both the breed and emotional state of cats from images. The proposed system integrates YOLOv8 (You Only Look Once), a model for identifying the presence of a cat in an image and classifying its breed, along with Convolutional Neural Networks (CNN) for recognizing the emotional state based on facial features. Initially, the input image undergoes preprocessing steps such as resizing and normalization to ensure consistent input for the models. The YOLOv8 model then detects the cat, generates a bounding box around it, and predicts its breed using learned visual characteristics. The detected region is extracted and passed as input to the CNN model, which analyzes facial patterns and expressions to classify emotions such as happy, sad, and angry. By focusing only on the relevant cat region, the system effectively reduces background noise and enhances prediction accuracy. The integration of object detection and deep learning-based classification enables the system to learn complex visual features and improve overall performance. Experimental results demonstrate that the proposed approach provides accurate and efficient predictions for both breed and emotion detection. This system highlights the potential of artificial intelligence in understanding animal behavior and can be further extended for real-time applications in pet monitoring, veterinary diagnosis, and animal welfare analysis.

Keywords

Cat Breed Detection, Emotion Recognition, YOLOv8, Convolutional Neural Network (CNN), Deep Learning, Computer Vision, Image Classification, Object Detection, Animal Behavior Analysis, Feature Extraction, Pet Monitoring Systems, Artificial Intelligence.

2. Introduction

In recent years, the rapid advancement of artificial intelligence and deep learning has significantly transformed the field of computer vision, enabling machines to interpret and analyze visual data with high accuracy. Applications such as object detection, image classification, and pattern recognition have gained widespread importance across various domains, including healthcare, security, and animal monitoring. Among these, understanding animal behavior through visual analysis has emerged as an important research area, particularly in improving pet care and welfare.

Cats are one of the most commonly domesticated animals, and identifying their breed and emotional state plays a crucial role in monitoring their health and behavior. While several existing systems focus on detecting animals or classifying their breeds, they often lack the ability to analyze emotional states. This limitation reduces their effectiveness in providing a comprehensive understanding of animal behavior. Emotion recognition in animals is a challenging task due to subtle variations in facial expressions and the limited availability of labeled datasets.

To address these challenges, this project proposes an integrated deep learning-based system for cat breed and emotion detection using YOLOv8 and Convolutional Neural Networks (CNN). The YOLOv8 model is utilized for real-time object detection and breed classification, enabling accurate localization of the cat within an image. Subsequently, the detected region is processed using a CNN model to analyze facial features and predict emotional states such as happy, sad, and angry. This combined approach ensures efficient processing and improved prediction accuracy by focusing only on the relevant region of interest.

The proposed system aims to provide a unified solution that performs both breed detection and emotion recognition within a single framework. It is designed to be simple, efficient, and user-friendly, requiring only an input image to generate meaningful results. By leveraging deep learning techniques, the system is capable of learning complex visual patterns and delivering reliable predictions.

Furthermore, this work contributes to the growing field of intelligent animal monitoring systems and demonstrates the potential of artificial intelligence in understanding animal emotions. The proposed approach can be extended to real-time applications and integrated into veterinary systems, smart pet monitoring solutions, and animal welfare platforms, thereby enhancing the interaction between humans and animals.

3. Literature Survey

Author / Study	Year	Method Used	Key Findings
Cat Breed & Emotion Detection using YOLO, CNN & Canny Edge Detection	2024	YOLO, CNN, Canny Edge Detection	Hybrid approach improves feature extraction and detects both breed and emotion; achieved high breed accuracy but limited by small dataset and basic emotion classes

Recognizing Cats and Dogs with Shape and Appearance-Based Models	—	Traditional ML (Shape & Appearance-based Models)	Works with smaller datasets and low computational cost; however, performance is sensitive to lighting, pose, and background variations
Deep Learning Applications in Animal Emotion Recognition (Applied Sciences)	2023	Deep CNN, Transfer Learning (ResNet, VGG, EfficientNet)	Captures subtle emotional features with high accuracy and strong generalization; requires high computational power and large datasets
Monitoring the Behaviours of Pet Cats using YOLO & Raspberry Pi	2022	YOLO with Embedded Systems (Raspberry Pi)	Enables real-time behavior detection with low-cost hardware; limited to predefined behaviors and constrained by hardware capabilities
YOLO-Based Analysis of Pet Emotional Behavior in Dogs and Cats	2025	YOLOv8, CNN (Transfer Learning)	Supports multi-emotion detection and real-time analysis with improved accuracy; challenges in detecting small objects and real-time optimization

4. Existing System & Drawbacks

In the domain of computer vision and artificial intelligence, several systems have been developed for animal detection and image classification using deep learning techniques. Models such as YOLO (You Only Look Once) and Convolutional Neural Networks (CNN) are widely used for detecting objects and recognizing patterns in images. These systems are commonly applied in applications such as pet identification, wildlife monitoring, and surveillance.

Most existing approaches focus primarily on **single-task operations**, such as detecting animals or classifying their breeds. While these systems demonstrate good performance in object detection and classification, they often do not extend their capabilities to more advanced analysis, such as understanding the emotional state or behavior of animals. As a result, their practical usefulness in areas like pet care and behavioral monitoring remains limited.

Drawbacks of Existing Systems:

1. **Limited Functionality:** Most systems focus only on breed detection or object identification and do not include emotion

recognition, resulting in incomplete analysis.

2. **Lack of Integration:** Detection and classification tasks are often handled separately, with very few systems combining both breed detection and emotion recognition in a unified framework.
3. **Dataset Limitations:** There is a scarcity of labeled datasets for animal emotions, which negatively affects model training and prediction accuracy.
4. **High Computational Requirements:** Deep learning models such as YOLO and CNN require powerful hardware (e.g., GPUs), which may not be available to all users.
5. **Complexity for Beginners:** Implementing and understanding advanced models can be challenging, especially for students or beginners in machine learning.
6. **Limited Behavioral Analysis:** Existing systems mainly focus on identification rather than analyzing animal emotions or behavior.

5. Proposed System

To overcome the limitations of existing approaches, the proposed system introduces an integrated deep learning-based framework for simultaneous cat breed detection and emotion recognition from images. The system combines the strengths of **YOLOv8 (You Only Look Once)** for object detection and breed classification with a **Convolutional Neural Network (CNN)** for analyzing the emotional state of the detected cat.

The workflow of the proposed system begins with an input image provided by the user. The image undergoes preprocessing steps such as resizing and normalization to ensure compatibility with the deep learning models and to enhance prediction accuracy. The preprocessed image is then passed to the YOLOv8 model, which detects the presence of a cat, generates a bounding box around it, and classifies its breed based on learned visual features.

Once the cat is detected, the corresponding region of interest (ROI) is extracted from the image. This step is crucial as it eliminates irrelevant background information and focuses only on the cat's features. The extracted region is then fed into a CNN model specifically designed for emotion recognition. The CNN analyzes facial expressions and patterns to classify the emotional state into predefined categories such as happy, sad, and angry.

The final output of the system combines both predictions—cat breed and emotional state—and presents them to the user in a clear and understandable format. This integrated approach ensures efficient processing and improved accuracy by leveraging both detection and classification capabilities within a single pipeline.

The proposed system offers several advantages, including improved accuracy through region-focused analysis, reduced background noise, user-friendly operation, and efficient performance. It is designed to be flexible and scalable, allowing future enhancements such as real-time detection, support for multiple animals, and inclusion of additional emotion categories.

Advantages of Proposed System:

1. **Integrated Framework:** The system combines both breed detection and emotion recognition in a single pipeline, providing a more comprehensive analysis compared to systems that perform only one task.
2. **Improved Accuracy:** By extracting and analyzing only the detected cat region, the system minimizes background noise, leading to more precise and reliable predictions.
3. **Efficient Processing:** The use of YOLOv8 enables fast and accurate object detection, while the CNN model effectively classifies emotional states, ensuring overall system efficiency.
4. **User-Friendly Design:** The system is simple and easy to use, requiring only an input image to generate both breed and emotion results without complex configurations.

6. Methodology / Algorithms

1. Image Acquisition

The system begins by acquiring input images through a user interface. Images can be uploaded in standard formats (JPEG/PNG) and serve as the primary data for processing. Libraries such as OpenCV and PIL are used to read and handle image data efficiently.

2. Image Preprocessing

Preprocessing ensures that the input image is suitable for model prediction. The following steps are applied:

- **Resizing:** Standardizes image dimensions for consistent model input
- **Normalization:** Scales pixel values (0–1 range) to improve model performance
- **Noise Reduction:** Enhances image quality by removing unwanted variations

These steps help improve accuracy and reduce computational complexity.

3. Cat Detection and Breed Classification

The system uses **YOLOv8 (You Only Look Once)** for object detection and breed classification:

- Detects the presence of a cat in the image
- Generates bounding boxes around detected objects
- Classifies the breed based on learned visual features

Algorithm Used: YOLOv8

- Single-stage object detection model
- Divides the image into grids and predicts bounding boxes and class probabilities

- Provides high-speed and accurate detection even in complex backgrounds

4. Region of Interest (ROI) Extraction

Once the cat is detected, the corresponding region is extracted (cropped) from the image. This step:

- Removes irrelevant background
- Focuses only on the cat's features
- Improves performance of the next stage (emotion recognition)

5. Feature Extraction and Emotion Recognition

The extracted cat image is passed to a **Convolutional Neural Network (CNN)** for emotion analysis.

Algorithm Used: CNN

- Extracts hierarchical features such as edges, textures, and facial patterns
- Uses convolutional, pooling, and fully connected layers
- Learns emotional patterns from training data

Image Transformations:

- Cropping (ROI)
- Resizing to model input size (e.g., 128×128)
- Optional grayscale conversion
- Reshaping for CNN input

6. Emotion Classification

The CNN model classifies the emotional state into predefined categories such as:

- Happy
- Sad
- Angry

Technique Used: Softmax Classifier

- Produces probability scores for each class
- Final emotion is selected based on highest probability

7. Result Generation

The final stage combines outputs from both models:

- YOLOv8 → Breed Prediction
- CNN → Emotion Prediction

The system maps predicted class labels into readable outputs and displays:

- Detected cat breed
- Identified emotional state

7. System Analysis & Requirements

System analysis is an essential phase in the development of the proposed cat breed and emotion detection system, as it defines the objectives, scope, and feasibility of the system. The primary goal is to design an intelligent deep learning-based solution capable of accurately detecting cat breeds and recognizing their emotional states from images. The system processes an input image, detects the cat using YOLOv8, extracts the relevant region, and analyzes facial features using a Convolutional Neural Network (CNN) to identify emotions such as happy, sad, and angry. It is designed to be user-friendly, efficient, and capable of delivering reliable results with minimal user intervention.

The scope of the system is focused on image-based analysis, where the model performs cat detection, breed classification, and emotion recognition using deep learning techniques. The system is suitable for applications such as pet monitoring, animal behavior analysis, and basic veterinary support. However, it is limited to static image inputs and predefined categories of breeds and emotions. It does not currently support real-time video processing or multiple animal detection, which defines the boundaries of the system.

A feasibility study indicates that the proposed system is practical and implementable. Economically, it is cost-effective as it utilizes open-source tools such as Python, TensorFlow, Keras, OpenCV, and YOLOv8, eliminating the need for expensive software or infrastructure. Technically, the system is feasible due to the availability of advanced deep learning models and libraries that can be implemented on standard computing systems, with optional GPU support for enhanced performance. Socially, the system is beneficial as it contributes to improved pet care, better understanding of animal emotions, and supports applications in animal welfare and behavioral analysis.

The requirement analysis of the system includes both functional and non-functional aspects. Functionally, the system must allow users to upload images, preprocess the input data, detect the cat using YOLOv8, classify the breed, extract the region of interest, and recognize the emotional state using a CNN model. It must then generate and display the final output in a clear and understandable format. Non-functional requirements emphasize usability, performance, reliability, portability, accuracy, and maintainability. The system should be simple to use, capable of producing quick and accurate results, and robust enough to handle variations in input conditions such as lighting and background noise.

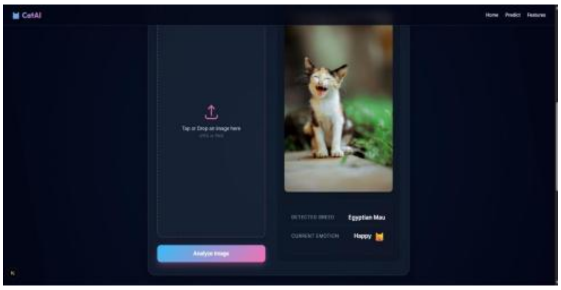
The hardware and software requirements further ensure smooth implementation and operation of the system. The system can run on standard hardware configurations such as an Intel i3 processor, 4 GB RAM (8 GB recommended), and minimal storage, with optional GPU support for faster processing. On the software side, it requires Python as the programming language, along with frameworks such as TensorFlow and Keras, and libraries like OpenCV and NumPy for image processing. Development tools such as Jupyter Notebook or Visual Studio Code can be used to implement and test the system efficiently.

8. Results & Performance Analysis

The proposed cat breed and emotion detection system was implemented using deep learning models, integrating YOLOv8 for object detection and breed classification with a Convolutional Neural Network (CNN) for emotion recognition. The system was evaluated using a set of test images containing different cat breeds and varying facial expressions. The results demonstrate that the model is capable of accurately detecting cats, classifying their breeds, and identifying their emotional states with reliable performance.

During testing, the YOLOv8 model effectively detected the presence of cats in images and generated precise bounding boxes around the objects. It showed strong performance even in images with complex backgrounds, variations in lighting, and different poses. The breed classification component achieved high accuracy due to the model's ability to learn distinctive visual features of different cat breeds.

The CNN-based emotion recognition module successfully analyzed facial features and classified emotions such as happy, sad, and angry. By focusing only on the extracted region of interest (ROI), the system minimized the influence of background noise, which significantly improved prediction accuracy. The use of convolutional layers enabled effective feature extraction, allowing the model to capture subtle differences in facial expressions.



The overall system performance indicates that the integration of YOLOv8 and CNN provides a robust and efficient solution. The system processes images within a reasonable time, making it suitable for near real-time applications. The modular architecture ensures that both detection and classification components work seamlessly together, resulting in consistent outputs. Additionally, the system demonstrated good generalization ability when tested on images with different conditions, although performance may vary slightly depending on dataset quality and diversity.

However, certain limitations were observed during evaluation. The accuracy of emotion detection depends heavily on the availability and quality of labeled training data, which is often limited for animal emotions. In some cases, misclassification may occur due to variations in facial angles, occlusions, or low-resolution images. Despite these challenges, the system maintains satisfactory performance and achieves a balance between accuracy and efficiency.

9. Conclusion

This project presented an intelligent deep learning-based system for detecting cat breeds and recognizing their emotional states from images. By integrating YOLOv8 for object detection and breed classification with a Convolutional Neural Network (CNN) for emotion recognition, the system successfully provides a unified solution for analyzing both physical and behavioral aspects of cats. The use of preprocessing techniques and region-of-interest extraction further enhances the accuracy and reliability of predictions by reducing background interference.

The experimental results demonstrate that the proposed system is capable of accurately detecting cats, classifying their breeds, and identifying emotional states such as happy, sad, and angry. The combination of fast detection and effective feature extraction ensures efficient performance, making the system suitable for practical applications. Despite certain challenges such as limited datasets for emotion recognition and variations in image conditions, the system achieves satisfactory performance and demonstrates strong potential for real-world implementation.

10. Future Work

The proposed cat breed and emotion detection system demonstrates promising results; however, there are several opportunities for further enhancement and expansion. One of the primary areas of future work is the development of a real-time detection system using video streams. By integrating webcam or surveillance-based inputs, the system can continuously monitor cats and provide live analysis of their behavior and emotional states. This would significantly improve its applicability in smart pet monitoring systems and veterinary environments.

Another important direction for improvement is the expansion of the dataset. The accuracy of emotion recognition largely depends on the availability of high-quality, labeled data. Future work can focus on collecting larger and more diverse datasets that include various breeds, lighting conditions, poses, and a wider range of emotional expressions. This would help improve model generalization and reduce misclassification in real-world scenarios.

The system can also be extended to support multiple animals and species beyond cats. By modifying and training the models accordingly, the framework can be adapted for dogs and other domestic animals, enabling broader applications in animal behavior analysis and wildlife monitoring. Additionally, incorporating more advanced deep learning architectures, such as attention mechanisms or transformer-based models, could further enhance feature extraction and classification performance.

Future enhancements may also include the integration of multimodal data, such as combining visual analysis with audio signals (e.g., meows or vocalizations) to achieve more accurate emotion recognition. Furthermore, deploying the system as a mobile or web-based application would increase accessibility and usability for general users, pet owners, and veterinarians.

REFERENCES

[1] Alex Krizhevsky, Ilya Sutskever, and Geoffrey Hinton, "ImageNet Classification with Deep Convolutional Neural Networks," in Advances in Neural Information Processing Systems (NIPS), 2012, pp. 1097–1105

[2] Joseph Redmon, Santosh Divvala, Ross Girshick, and Ali Farhadi, "You Only Look Once: Unified, Real-Time Object Detection," in IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2016, pp. 779–788

[3] Glenn Jocher et al., "YOLOv5 by Ultralytics," GitHub Repository, 2020

[4] Tsung-Yi Lin et al., "Microsoft COCO: Common Objects in Context," in European Conference on Computer Vision (ECCV), 2014, pp. 740–755

[5] Yann LeCun, Yoshua Bengio, and Geoffrey Hinton, "Deep Learning," Nature, vol. 521, pp. 436–444, 2015

[6] D. J. Mellor, "Animal Welfare and Emotion Assessment," Frontiers in Veterinary Science, vol. 7, 2020

[7] Karen Simonyan and Andrew Zisserman, "Very Deep Convolutional Networks for Large-Scale Image Recognition," arXiv preprint, 2014

[8] Christian Szegedy et al., "Going Deeper with Convolutions," in IEEE CVPR, 2015

[9] Ian Goodfellow, Yoshua Bengio, and Aaron Courville, "Deep Learning," MIT Press, 2016

[10] Kassim A., Tan S., Phang S., "Deep Learning Approaches for Animal Emotion Recognition," IEEE Access, 2021

[11] Ultralytics, "YOLOv8 Documentation and Implementation," 2023