International Journal of Scientific Research in Engineering and Management (IJSREM)
Volume: 09 Issue: 05 | May - 2025 SJIF Rating: 8.586 ISSN: 2582-3930

Human Emotion Detection Using Thermal Images

Krishna Swaroop A, Tejaswini A S, Siri K S, T S Saketh Puranik, Vidhwan H D Malnad College of Engineering,Hassan

Email: ksa@mcehassan.ac.in, tejaswiniashwini1134, kssiri256@gmail.com, tssakethpuranik@gmail.com,vidhwanhd@gmail.com

Abstract-**In healthcare, security, and even in human-computer interactions, the application of emotion detection technology is becoming increasingly popular. Exiting methods employing light pictures have problems with occlusion and illumination. This project uses thermal imaging to identify the changes in facial temperatures due to emotions, which guarantees functionality in all lighting conditions. It also uses real-time emotion detection from thermal face data using an optimized YOLOv5 model. The model is trained using labeled thermal datasets and is modified using transfer learning to streamline the model for specific features of the domain. This approach ensures emotion detection is done fast, accurately, and stealthily. This is critical in low-light situations or where privacy is paramount. Examples of such include access control, adaptive learning, and mental health care. The technology enhances emotional computing by integrating deep learning with infrared imaging.**

I. INTRODUCTION

In emotion detection, which recognises human emotions using biometric data, facial expressions are a key indicator. Traditional visible-light image-based systems suffer from poor lighting and occlusions. This work uses thermal imaging to detect emotional variations in facial temperature, ensuring constant performance under challenging conditions. Temperature variations in regions like the forehead and eyes are indicative of emotional states. YOLOv5, a fast and accurate object detection model, is adapted to scan thermal pictures and classify emotions. It is trained on annotated thermal datasets using transfer learning. The system offers real-time, non-intrusive emotion recognition. Applications include surveillance, healthcare, and human-computer interaction.

II. Existing System

RGB-based emotion identification uses machine learning, although it struggles with obstructed faces and low light levels. Thermal photography addresses these issues and ensures accurate results in any lighting situation by capturing heat signatures.However, because thermal images lack texture, feature extraction is more challenging. Deep learning improves accuracy, especially CNNs that use transfer learning. While some systems use traditional classifiers for emotion classification and YOLO for facial recognition, they often perform poorly when applied to a variety of emotions or in real-time. Effective thermal emotion detection necessitates a tailored approach.

III. Literature Survey

A significant progress was observed in the field of image analysis of ten people using deep learning, and researchers used various architecture and pre -treatment methods to increase the classification and accuracy of emotion. In the heat classification of images using the deep learning method 2023, the Prosenjit Chatterjee and ANK Zaman, together with the Calmane filter, used a CNN model such as the Calmane filter to enhance the classification performance

The study discovered a restriction on small -scale data sets and risks of experience, but it has been found to be effective in extracting signs from the thermal image. Similarly, modified Resnet152 (2021) Aisware K. Prabhakaran et al. Resnet152 Use depth to classify emotions from the thermal image of the NVIE data set. While the model worked well, the dependence on one data set caused concern about generalization under various visualization conditions. Deep BLTZMANN MACHINE (2014) S. WANG et al. Evaluation of the basic stage by applying DBMS for emotion recognition. This model emphasized thermal transfer training on characteristics without manual production and was innovative at the time. Nevertheless, the lack of high calculation requirements and more new deep learning models limited the practical relationship. Recently, Densenet (2023) brand Rajendra Pracad and Dr. Using B. SAY CHANDAN, the perception of human emotions in the heat image integrated the effect of extracting signs. This multi-mode pipeline has reached an impressive accuracy (95.97%) in the RGB-D-T data set, but has introduced difficulties to prevent distribution with limited real-time environments

Finally, voluntary recognition of emotions in the thermal image 2020 of the face with Chirag Kjal is mainly focused on the localization of the face by using the machine learning method trained by the USSTC-NVIE data set. This study aims to replace traditional approaches based on rules based on educational models for tasks such as face and expression analysis. Despite the fact that he provided valuable information on facial structure analysis in ten images, he proposed a limited discussion on direct classification of emotions, and there was no evaluation in the actual scenario. This study emphasizes the need for a larger and more diverse data set, emphasizing the promise of deep training on thermal image processing, and optimizes the model in real time and improves generalization under various conditions.

IV. About THERMAL IMAGES

The thermal image is effective in all lighting conditions by capturing the heat released by the human body. They visualize temperature changes that can represent emotional state. The main areas of the face, such as foreheads and cheeks, show noticeable heat changes. Unlike RGB images, colors and texture details are not enough for column images. This makes the extract of function more complicated. However, there is little effect on occlusion, such as glasses and hair. Ten visualization provides a non -engineer method that monitors physiological reactions. This is especially useful in safety, medical and emotional calculations.

METHODOLOGY

IV.I Problem Description

The perception of emotions is important in areas such as computers and human health care, observation and interaction. Traditional systems rely heavily on the image of RGB, which is very sensitive to lighting conditions and depends on obstruction. These imitations reduce the accuracy of low lighting conditions or the real world. In addition, the emotions subtly expressed on the human face may not be completely captured in the visible image of the spectrum. As a result, traditional methods are struggling to guarantee confidence that can trust emotions under various conditions. Ten visualization provides promising solutions to find infrared radiation released from the skin to identify physiological changes related to emotions. It works effectively in a complete darkness and affects less external obstacles. However, the treatment of thermal images to recognize emotions is a difficult task due to low texture and contrast. In real time, you need an effective system that can trust your emotions and use heat data to recognize it.

IV.II Objective

The main goal of this project is to develop and implement a system that detects feelings in real time using heat assembly and deep learning YOLOV5 models. This system must accurately classify emotional state that analyzes the heat signature of happiness, sad, evil or neutral, and the heat signature of the face. By guaranteeing the functions of low lines and obsolete media, we aim to overcome the limitations of traditional methods based on RGB. This project detects a heater using transmission training in a thin settings in YOLOV5. He will also develop an emotional classifier that can interpret the thermal pattern The pipeline must handle both the static image and the live video pipeline in the column chamber. Additional goals include system optimization for convenient web interface creation and real -time performanceModels are scalable and need to be distributed to local or cloud platforms. The continuous improvement of models using retraining is also part of long -term goals.

IV.III . System Architecture

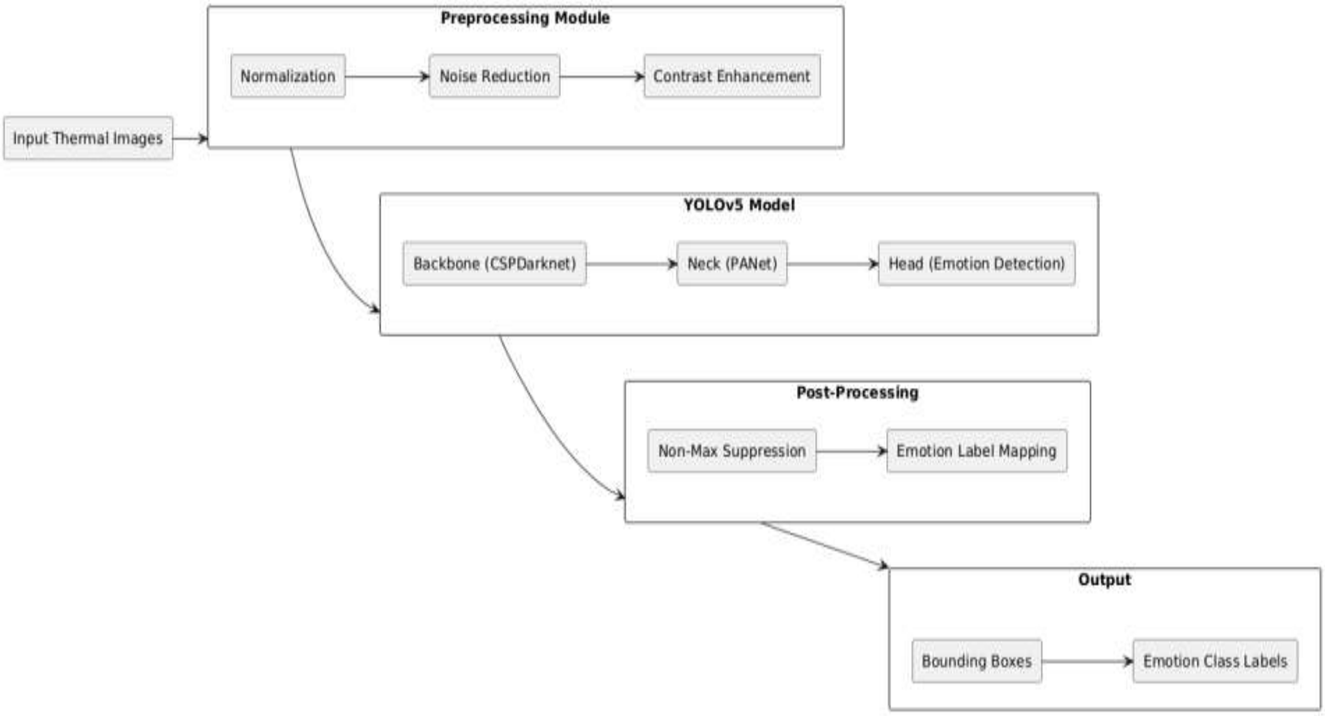


Fig 1. System Architecture

This system starts with data collection using a dataset of a thermal image or column chamber. The pre -processing steps normalize and change the size of the image to match the input format YOLOV5. YOLOV5 is subtly adjusted to detect the face area in the column image and creates a limited box around the face found. Then this crop is submitted to emotional classification models such as pre -trained models such as user CNN or Resnet or Mobilenet. The classifier predicts the emotional state of the heat man. For real -time functions, the thermal frame is constantly treated in the classification cycle of detection. Optimization, such as quantification, pruning and acceleration of the graphic processor, provides smooth performance. Designed Internet interface using Flask or Streamlit allows users to download images or stream heat videos. The interface indicates the found face and predicted feelings in real time. Finally, the system supports distribution of devices or cloud platforms through monitoring tools and mechanisms for retraining the model.

V. IMPLEMENTATION

V.I Selecting Thermal Images Dataset

Select column image data set The first step in implementation was the collection of thermal image sets to train the model training and emotional classification models of the YOLOV5 face. Since the public set of heat emotions is limited, this project provides a visible and infrared image of the face using a NVIE data set (natural visible and infrared expression). This image is marked with appropriate emotional tags such as happiness, sadness, evil, unexpected, neutrality. If the data set is insufficient, you can use the Flir Lepton column chamber to get additional colon images to capture the appropriate lighting and conditions for various emotional conditions

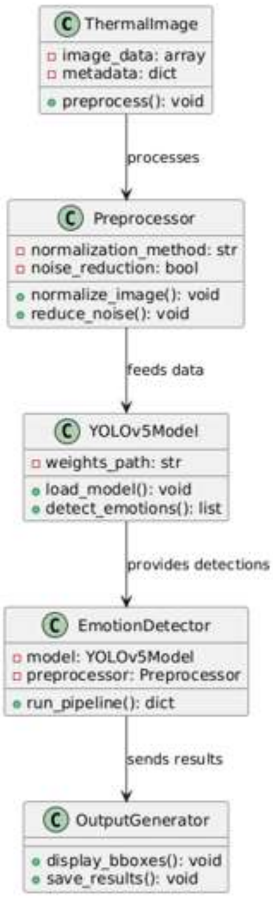
Fig 2. CNN Model Architecture

VI.II. Data Augmentation

Data increase Considering a relatively small data set, the increase in data was important to increase the likelihood of generalizing the model. We have created greater training from limited data using methods such as random pruning, coup, rotation and scaling. It also helps to simulate the actual change of the camera's expression and angle, which may vary depending on the actual placement. The data set is divided into education, verification and test sets to evaluate the performance of the model.

Why YOLOV5? YOLOV5 was selected as a facial detection model due to the effect, speed and high accuracy when detecting objects in real time. The YOLOV5 is an improved version of the YOLO series model (only once in case) and is famous for detecting objects at direct intersections of the

VI.III. CNN Model Architecture



The gesture recognition model is deployed through a Convolutional Neural Network (CNN) trained to effectively extract spatial features from hand gesture images. The model takes a preprocessed grayscale image of 120×120×1 as input. It has three consecutive blocks of convolutional layers with ReLU activation followed by max pooling, which decrease the spatial dimensions step by step while extracting the most important features. The output is flattened and fed into a dense layer with ReLU activation and dropout regularization to avoid overfitting. The network finally ends in a dense output layer of seven neurons, one for each gesture class, followed by Softmax activation to handle the multi-class classification.

VI.Face Dection using YOLOv5

that no unwanted actions are initiated. This transparent correspondence between gesture classes and media commands supports usability and responsiveness in real-time applications.

VI.V Implementation Steps

VI.V.I Preprocessing Face Regions

Interface design The user interface (user interface) was developed using Flask, Light Python and HTML5/CSS3. The user interface displays Real -time video in column chamber There are emotional marks by limiting boxes around the found face.Performance metrics such as the accuracy of the classification found on the dashboard

VI.V.II Choosing the Classification Model

Model deployment and testing As soon as the system was developed, it was placed for testing in a real environment. This system has tested accuracy, speed and reliability under various conditions such as various facial directions, obstruction and lighting conditions (despite the fact that the thermal image does not affect lighting). The system showed the classification of emotions in reliable detection and uncontrollable media on the face..

VI.V.III Model Training and Fine-Turning

In order to process every input frame for model inference, the system employs a specialized image processing function. This function initially resizes the entire captured frame to a specified resolution and crops a predetermined Region of Interest (ROI) according to preset coordinates. The cropped ROI, which is centered on the hand gesture region, is resized to 120×120 pixels to conform to the model's input size. It is then converted into grayscale and binarized via thresholding, which emphasizes the contours of the gesture. The processed image returned as a result along with the original frame enables both to be predicted and displayed in the interface.

VI.V.IV Model Evaluation

Model architecture and education The YOLOV5 architecture consists of a bone network extracted from the input image (indicating a spanning neuro network or CNN), and a detection head (person) for predicting a drawer and class mark for the detected object is followed YOLOV5 training to find people
We started with a pre -trained YOLOV5 model and was trained as a coco data set. COCO contains various categories of objects, which helps to study low levels such as edges and textures

The model is accurately configured in the user set of the thermal image. We changed the last layer to predict the class of the face and adjusted the in size to handle the resolution of the thermal image

Fig 3. Example for Thermal Recognition

VI.VI Real time Processing

Data increase Considering a relatively small data set, the increase in data was important to increase the likelihood of generalizing the model. We have created greater training from limited data using methods such as random pruning, coup, rotation and scaling. It also helps to simulate the actual change of the camera's expression and angle, which may vary depending on the actual placement. The data set is divided into education, verification and test sets to evaluate the performance of the model.

VI.VII System Inference

Real -time conclusion In order to achieve real -time processing, the system was optimized by accelerating the output using the acceleration of the graphics processor (using CUDA for GPU NVIDIA). YOLOV5 was used to detect people in each frame, and after the face was found, it was transmitted as a classification of emotions. Then the result (a box limit with the emotions) was displayed in the graphic user interface (GUI). This system works in real time by processing about 30 frames (FPS) per second and suitable for applications such as live observation and monitoring.

VI.VIII. Model Deployment and Testing

Since the public set of heat emotions is limited, this project provides a visible and infrared image of the face using a NVIE data set (natural visible and infrared expression). This image is marked with appropriate emotional tags such as happiness, sadness, evil, unexpected, neutrality. If the data set is insufficient, you can use the Flir Lepton column chamber to get additional colon images to capture the appropriate lighting and conditions for various emotional conditions

To teach the facial detection model, the data set has a variety of skin shades, facial directions and environmental images with a variety of column images, which ensures that the model can be trusted.

VI. Conclusion

Face Emotion Detection in Thermal Images is the title of this study. The goal was to create a reliable, real-time facial emotion identification system that works well in low-light and privacy-sensitive settings utilising YOLOv5. This method uses thermal imaging, which records heat signatures from human skin, especially the face, to provide more reliable physiological indications associated with emotions than traditional emotion detection systems that rely on RGB images and are sensitive to illumination fluctuations.

Future Scope

The project holds huge potential for future development and expansion. The areas of major focus for future improvement are:

• Health monitoring detection of heat emotions can help stress, anxiety or mental health problems in real time

• Driver's safety system You can detect fatigue or emotional stress between drivers to prevent traffic accidents.

• Adaptation Education Emotional educational tools can be adjusted to provide content according to students' participation.

• Confidential maintenance preservation Heat visualization can recognize emotions without disclosing useful identity for sensitive environments.

• Wearable integration future wearable devices can track emotions in real time using small thermal sensors.

These developments enable the widespread use of gesture recognition technologies in various industries, such as entertainment, healthcare, automotive, and assistive use. The presented method lays a solid foundation for future research and development of gesture-based user interfaces.

VII. REFERENCES

[1] An Automated Eye Disease Detection System Using Convolutional Neural Network (2023), Authors Mayu Sirsat, Akanksha Sathe, Shreya Ladhe

[2] A Study on Retinal Disease Classification and Filteration Approaches (2015), Authors: Parul, Neetu Sharma

[3] Diverse hand gesture recognition dataset by Zahra Mohammadi, Alireza Akhavanpour, Razieh Rastgoo ,Mohammad Sabokrou. (2022)

[4] Review of Image Processing and Machine Learning Techniques for Eye Disease Detection and Classification (2016), Authors: Mr. Langade Umesh, Ms. Malkar Mrunalini, Dr. Swati Shinde

[5] Detection of Retinal Hemorrhage in Color Fundus Image (2017), Authors: Priyanka Kale, Prof. N Janwe

[6] Real-Time Detection of Retinal Hemorrhage in Color Fundus Image

[7] (2017), Authors: Priyanka Kale, Prof. N Janwe

[8] Gesture Recognition System based on Millimeter Wave and Thermal Imager By Wen-Hsiang Yeh, Yi- Lin Cheng and Yu-Ping Liao (2023)

[9] Resource Multi categorical of common eye disease detect using convolutional neural network: a transfer learning approach (2022), Authors: Abu Kowshir Bitto, Imran Mahm

[10] Robust Hand Gestures Recognition Using a Deep CNN and Thermal Images By Daniel Skomedal Breland , Aveen Dayal , Ajit Jha , Phaneendra K. Yalavarthy, Om Jee Pandey and Linga Reddy Cenkeramaddi (2021)

[11] Machine Learning-Based Approach for Hand Gesture Recognition By M. Deepika, Shilpa Choudhary, Sandeep Kumar and Kodi Srinivas (2023).