iJRASET
International Journal For Research in Applied Science and Engineering Technology



INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 10 **Issue:** VIII **Month of publication:** August 2022

DOI: https://doi.org/10.22214/ijraset.2022.46387

www.ijraset.com

Call: 08813907089 | E-mail ID: ijraset@gmail.com

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 10 Issue VIII Aug 2022- Available at www.ijraset.com

Patient Monitoring Using Emotion Recognition

Ashish Suraj[1], Amarnath S. Kaushik[2], Kannika Bai[3]
[1, 2, 3]*Department of Computer Science & Engineering, BNM Institute of Technology, Banashankari-560070, Karnataka, India*

Abstract: Due to technology improvements, recognising the patient's emotions using deep learning algorithms has gotten a lot of interest recently. Automatically detecting emotions can aid in the development of smart healthcare centers that can detect pain and tiredness in patients so that medication can be started sooner. One of the most fascinating themes is the use of advanced technology to discern emotions, as it defines the human-machine interaction. Various strategies were used to teach machines how to predict emotions. Recent research in the field of employing neural networks to recognise emotions will be used in our system. Our system focuses on recognising emotions from facial expressions and demonstrating several approaches for implementing these algorithms in the real world. Techniques for recognising emotions can be utilized as a surveillance system in healthcare canters to monitor patients.

Keywords: Pain, Pain Expression, CNN, Emotion, Neural Networks

I. INTRODUCTION

With the advancement of deep learning technology, deployment possibilities in smart health systems are becoming increasingly important. The healthcare sector uses a plethora of machine learning-based technology consisting of a range of surveillance systems and remote disease diagnostics, in healthcare and elderly care establishments, and so forth, to comprehend patients' feelings and emotions. These systems of technology are used to detect early emotion in order for quick interventions to be put in place to reduce the symptoms of stress and depression.

To identify patients' emotions, we use a deep neural network. These techniques can be utilized as a surveillance system, capturing photos and videos using various equipment such as cameras. The applications of patient emotion detection are numerous, and they involve many important aspects of daily life, such as safe driving, mental health monitoring, social security, and so on. Many polls go into great detail about this topic from various angles. The writers focused on body gestures, voice expressions, and audio-visual expressions when researching emotion recognition.

These surveys' authors concentrated on multi-modal approaches that combine facial or speech with body motions to improve emotion recognition.In the experiment, they employed the device as a sensor to recognise emotions. They then compared their method to a webcam-based method for recognising emotion. The results of the trial showed that utilizing MHL with AR provided better accuracy than using a camera. Examining Patients Emotion recognition for a better health system has been vital in the previous decade because it aids in a variety of disciplines, including medicine. It aids doctors in recognising their patients' psychological issues and, as a result, initiating treatment early. Many hospitals around the world have begun to use AI to treat patients, and many researchers are working on developing neural networks that can understand a patient's emotions. One prominent AI strategy uses three separate modalities to discern emotions: spoken, facial, and audio-visual methods. To give readers a comprehensive understanding of how neural networks are utilized in the medical field, we present the most frequent strategies for recognising emotions.

II. LITERATURE REVIEW

A. Data Preprocessing

According to Karamitsos and others [6], grayscale keeps one channel for each image and simplifies the convolution operation. Color is not effective in this matter. However, it can be important in other issues (for example, grading fruits to different quality grades based on many factors such as color). Histogram equalization unifies and enhances the contrast of each image to improve edge detection. As a result, the image is never too bright or too dark.

Ilyas & Ur Rehman [7], note, in implementing an offline video with any random duration, the initial step in processing input videos is to extract frames. The frame sizes vary, thus we resize them to 324 x 240, the smallest frame size in the collection, in order to bring them into uniform dimensions. Additionally, all frames are turned into grayscale in order to reduce the computational overhead.

©IJRASET: All Rights are Reserved

Muhammad Fahim and Denis Rangulov [8] states, sequences from 118 participants in the CK+ dataset were used in the trials. For prototypical expression recognition for each sequence, the initial image i.e., neutral face and three peak frames were utilized. Face detection using Haar cascades for each image is followed by cropping and resizing to 96 96. They employ grayscale pictures.
Kim, Byung-Gyu, and Chilamkurti, Naveen [9] states to prevent the overfitting issue, the pre-processed dataset has been increased. Since CNN requires input with a defined size, each network then receives those datasets as input. So that the network may extract useful features effectively, superfluous sequence sections were eliminated, and crucial features were emphasized.
According to Palanivel S. and Sujanaa J. [10] a straightforward preprocessing phase where the RGB photos are changed into grayscale images is first applied to the photographs. The input frames were pre-processed using the Scikit-library. To feed the CNN, the frames are transformed into a NumPy array format.

B. Feature Extraction

Benjula Anbu Malar M.B., Praveen.R. [11] states the item for which it has been prepared is recognised from the source using the Haar Cascade classifier. The positive image is prepared by layering it over several negative images to create the Haar Cascade. Commonly, the training is carried out on a server and at several levels. The mouth and eyebrows are the key areas of the face that are taken into account for emotion identification. Region splitting is the term for the division of the lips and brows.
According to Zhao et al., [3] employing a depth-cascading multitasking MTCNN framework, face detection, and alignment can be completed at the same time. This leverages the internal relationship between the two to improve performance and extract global facial features. As a result, the positions of the face, left and right eyes, nose, and the left and right corners of the mouth can thus be obtained.
Using the cascade classifier, an algorithm devised by Paul Viola and Michael Jones that employs machine learning techniques, [10] note that these cascade classifiers are trained with samples containing facial and non-facial images. With the aid of the Haar-based cascade classifier, the mouth and face of each video frame is detected.
By using the Gabor filter, an image processing tool, as noted by [12], when applied broadly to feature extraction, stores information about digital images and uses a new algorithm with the help of a neural network that has been trained by the extracted features of the Gabor filters. As a result, this new technique helps to scale RMS contrast and present fuzzily skewed filtering.

C. Prediction

Anthony Gachagan, and Stephen Marshall [12] Out of the many activation functions discussed the soft max function offers good accuracy, it gives the output in the range of 0-1 which can be easily used to predict the emotion.

III. METHODOLOGY

A. CNN Architecture

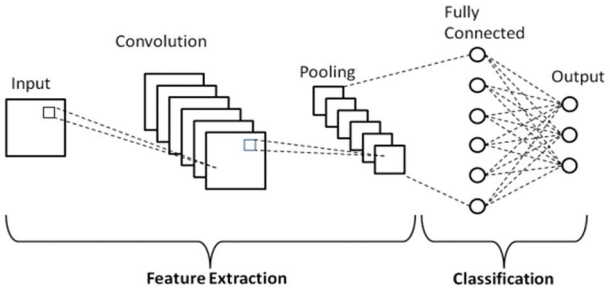


Fig. 1 CNN Architecture

The CNN architecture shown in Fig. 1 employs convolutional land pooling layers. After convolutional layers, pooling, a sort of non-linear down-sampling, is widely used to eliminate superfluous features and reduce the number of parameters during training to prevent overfitting.
Max pooling is specifically utilized. After convolutional and max pooling layers, the multidimensional array is flattened into a two-dimensional array using the fully connected (FC) and flattening (FL) layers. In the output layer, the probability of each class is calculated using the SoftMax function in order to recommend the most likely categorization.

The photos go through a simple preprocessing step that converts RGB images to grayscale images. Equalization of histograms to unify and improve image contrast for improved edge identification. The Haar cascade classifier is a machine learning-based algorithm. These cascade classifiers are trained on data that include both face and non-facial images.The Haar-based cascade classifier is utilized to recognise the mouth and face in each video frame.

B. System Architecture

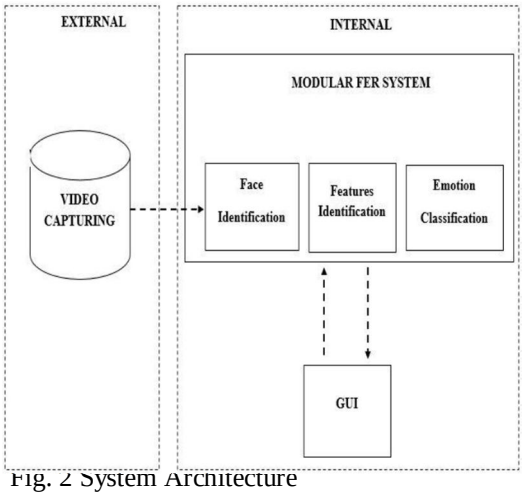


Fig. 2 System Architecture

The purpose of this project is to use CNN architecture and transfer learning to detect a person's emotion in real time. The main system procedures include face and mouth detection, training, validation, and real-time testing.The Pain Recognition and Facial Expression dataset will be used to train the model.

IV. MODULE DESIGN

A. Face Extraction

The human face is captured using the webcam on the computer or an external webcam. The face is retrieved from the live stream, and all other undesired components are ignored.

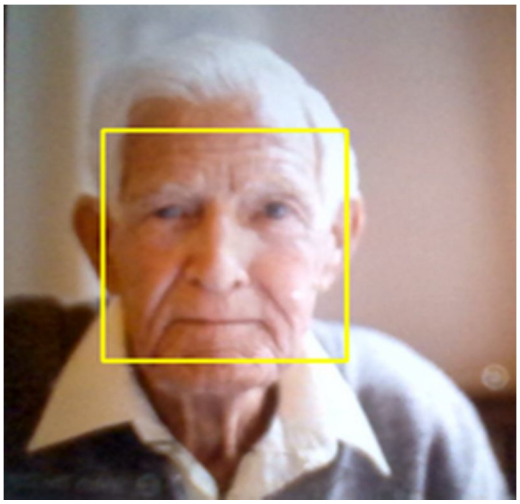


Fig. 3 Face Detection

B. Data Preprocessing

The images go through a simple preprocessing phase that converts RGB images to grayscale images using openCV libraries.

©IJRASET: All Rights are Reserved | SJ Impact Factor 7.538 | ISRA Journal Impact Factor 7.894 |

C. Classifier

The Haar cascade classifier is a machine learning technique-based algorithm. These cascade classifiers are trained on data that include both face and non-facial images. The Haar-based cascade classifier is utilized to recognise the mouth and face in each video frame.

D. Emotion Classification

Following the completion of the feature extraction subtask, the person's reaction is generated in real time based on the softmax scores generated at the last layer of the cnn architecture.

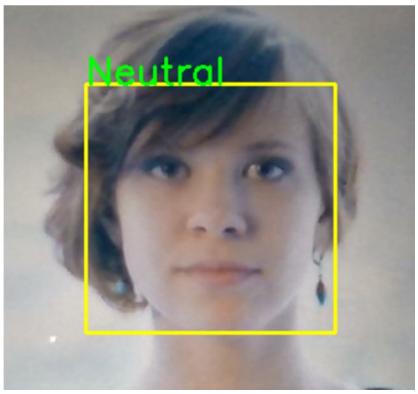


Fig. 4 Resting Face

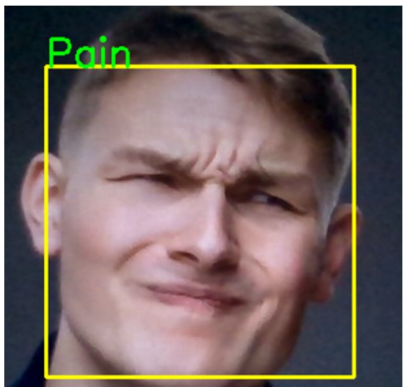


Fig. 5 Face in Pain

V. RESULTS

Using a web camera in an open laboratory setting, the face photos are retrieved. The experimental findings demonstrate that the proposed CNN system accurately (79.71%) detects the two moods utilizing mouth and eye data.

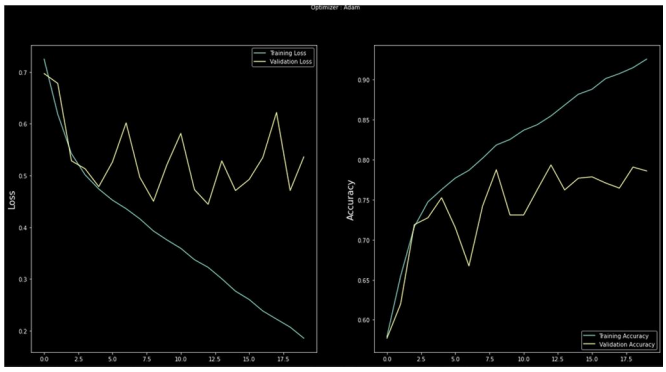


Fig. 6 Loss and Accuracy by Epoch

VI. CONCLUSION

This paper proposes a real-time emotion detection system that uses CNN architecture to recognise the emotions normal, pain, and tiredness in real-time video feeds. A web camera is used to extract the face photographs. Our proposed approach can provide medical personnel with information about a patient's status based on their facial expressions.

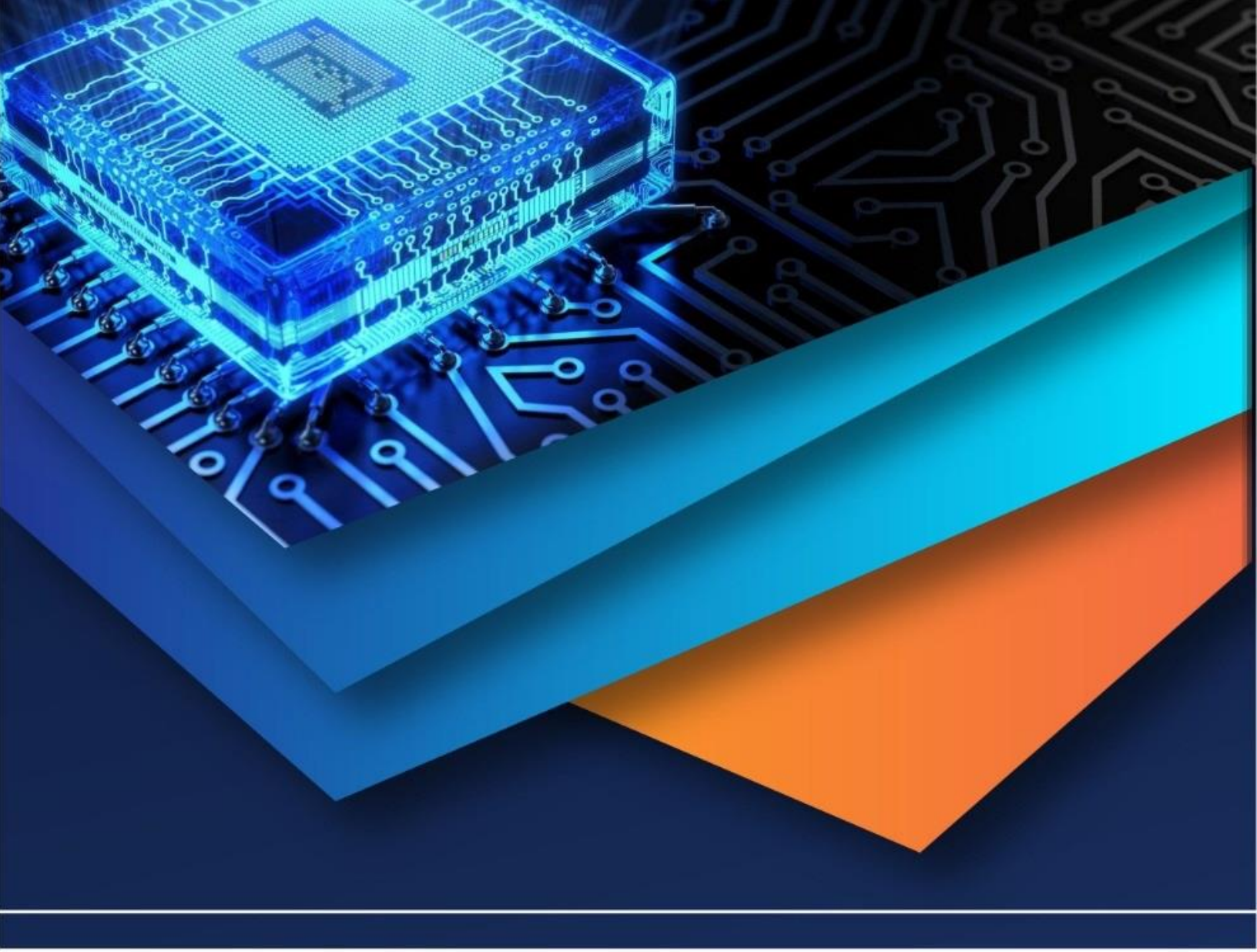
VII. ACKNOWLEDGMENT

The authors are thankful to Prof. Jayashree, Assistant professor, BNMIT, Bangalore for providing a good environment and well-equipped laboratory in the Department of Computer Science and Engineering to carry out our project in an efficient manner.

REFERENCES

[1] Karamitsos, I., Seladji, I. and Modak, S. (2021) “A Modified CNN Network for Automatic Pain Identification Using Facial Expressions”. Journal of Software Engineering and Applications,14, 400-417.
[2] Ninad Mehendale1,2. (2020) “Facial emotion recognition using convolutional neural networks (FERC)”.
[3] Zuopeng Zhao, Nana Zhou, Lan Zhang, Hualin Yan, Yi Xu, and Zhongxin Zhang(2020)” Driver Fatigue Detection Based on Convolutional Neural Networks Using EM-CNN”.
[4] Xu, X.J., Huang, S.R., Children’s Hospital and De Sa, V.R (2019) “Pain Evaluation in Video Using Extended Multitask Learning from Multidimensional Measurements”.
[5] Roy, S.D., Bhowmik, M.K., Saha, P. and Ghosh, A.K (2016) “An Approach for Automatic Pain Detection through Facial Expression”.
[6] Karamitsos, I., Seladji, I. and Modak, S (2021)"A Modified CNN Network for Automatic Pain Identification Using Facial Expressions".
[7] Sadia Ilyas, Hafeez Ur Rehman (2020)"A Deep Learning based Approach for Precise Video Tagging".
[8] Denis Rangulov, Muhammad Fahim (2020)"Emotion Recognition on large video dataset based on Convolutional Feature Extractor and Recurrent Neural Network".
[9] Seo-Jeon Park, Byung-Gyu Kim, and Naveen Chilamkurti (2021)"A Robust Facial Expression Recognition Algorithm Based on Multi-Rate Feature Fusion Scheme".
[10] Sujanaa J, Palanivel S (2020)"Real-time video-based emotion recognition using convolutional neural network and transfer learning".
[11] Praveen.R, Benjula Anbu Malar M.B (2020)"Emotion Recognition using Convolutional Neural Network".
[12] Rahimeh Rouhi, Mehran Amiri and Behzad Irannejad (2012)"A review on feature extraction techniques in face recognition".
[13] Chigozie Enyinna Nwankpa, Winifred Ijomah, Anthony Gachagan,and StephenMarshall(2018)"Activation Functions:Comparison of Trends in Practice and Research for Deep Learning".

©IJRASET: All Rights are Reserved



10.22214/IJRASET

45.98

IMPACT FACTOR: 7.129

IMPACT FACTOR: 7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Call : 08813907089 (24*7 Support on Whatsapp)