ISSN: 2454-9940

INTERNATIONAL JOURNAL OF APPLIED SCIENCE ENGINEERING AND MANAGEMENT

E-Mail :
editor.ijasem@gmail.com
editor@ijasem.org

www.ijasem.org

INTERNATIONAL JOURNAL OF APPLIED SCIENCE ENGINEERING AND MANAGEMENT

ISSN2454-9940

www.ijasem.org

Vol 18, Issue 2, 2024

FACE DETECTION AND FACE DETECTION AND RECOGNITION USING LSTM

[1] K. LAKSHMAN REDDY SIR, [2] CHINTAPALLI NIKHITHA

[1](Assistant Professor), MCA, S.V.K.P & Dr K.S. Raju Arts & Science College

[2]MCA, scholar, S.V.K.P & Dr K.S. Raju Arts & Science College

ABSTRACT

The project has entitled as **"FACE DETECTION AND RECOGNITION USING LSTM"**, and developed by using Python as front end. Face detection is a computer vision technology that helps to locate/visualize human faces in digital images. This technique is a specific use case of object detection technology that deals with detecting instances of semantic objects of a certain class in digital images and videos. With the advent of technology, face detection has gained a lot of importance especially in fields like photography, security, and marketing.

This project is designed to detect the human face appearing in front of webcam. It recognizes the face and compares the same face with the dataset. If the face trained before appearing on screen the accuracy will be high (> 90%) otherwise the accuracy will be low for new faces (< 40%). This face detection process is done by using LSTM technique. It helps to increases the accuracy that existing system. The training time required is very less when compared to exiting system. Before applying LSTM it will take more time to train the face. Here the face detection accuracy level is increases gradually and time requirement is reduced more.

An image is nothing but a standard Numpy array containing pixels of data points. More the number of pixels in an image, the better is its resolution. You can think of pixels to be tiny blocks of information arranged in the form of a 2 D grid, and the depth of a pixel refers to the color information present in it. In order to be processed by a computer, an image needs to be converted into a binary form.

1.INTRODUCTION

In this project, the proposed lightweight LSTM (Long short-term memory) architecture is designed to detect faces of people in front of computer and laptops when their face might not be very visible over surrounding background. We have designed a algorithm light architecture cooperating with the proposed sliding window procedure. The designed model works with maximum simplicity to support all desktop versions. An output from processing presents a box on face location in the screen of device. The model was trained by using Adam and tested on various images. Results show that proposed lightweight LSTM detects human faces over various textures with accuracy above 99% and precision above 98% what proves the efficiency of our proposed model.

Our objective is to learn a deep face representation from large-scale data with massive noisy and occluded face. Besides, we add an adaptive fusion of softmax loss and center loss as supervision signals, which are helpful to improve the performance and to conduct the final classification. The experiment results show that the suggested system achieves comparable performances with other state-of-the-art methods on the Labeled Faces in the Wild and YouTube face verification tasks.

Deep learning is a trendy term these days, and it refers to a modern age in machine learning in which algorithms are taught to identify patterns in vast amounts of data. It mostly refers to studying various layers of representation, which assists in the understanding of pictures. To interact with the objects in a video series, many researchers use a form of deep learning called a LSTM. Face detection involves several face-related technologies, such as face authentication, facial recognition, and face clustering. For identification and understanding, effective preparation must be carried out. The standard technique did not produce a positive outcome in terms of face recognition precision. The objectives of this research are by using a deep learning model to enhance the accuracy of face detection. For recognizing faces from datasets, the proposed model utilizes a deep learning technique named LSTM. The proposed work is applied using the same technique, a well-known deep learning process. Our model is trained and validated using the face dataset, which includes 30 photos for single face collected from dataset code. The training accuracy of the model was 95.72% percent, and the validation accuracy was 96.27%.

OBJECTIVES

The main objectives of the research are to

- To initiate the webcam for image processing.
- To access webcam present in system as well as with face detection technique can be used for finding the trained faces.
- To convert the RGB image of face into binary format for getting more accuracy.
- To highlight the end user face.
- To apply the training face images set to find the authorized persons.
- To provide accurate result.

2.LITERATURE SURVEY

SUGGESTION MINING FROM OPINIONATED TEXT OF BIG SOCIAL MEDIA DATA

AUTHORS : YouseefAlotaibi[1, *] , Muhammad Noman Malik[2] , Huma Hayat Khan [3] , AnabBatool[2] , Saiful Islam[4] ,AbdulmajeedAlsufyani[5] and Saleh Alghamdi Social media data are rapidly increasing and constitute a source of user opinions and tips on a wide range of products and services. The increasing availability of such big data on biased reviews and blogs creates challenges for customers and businesses in reviewing all content in their decision-making process. To overcome this challenge, extracting suggestions from opinionated text is a possible solution. In this study, the characteristics of suggestions are analyzed and a suggestion mining extraction process is presented for classifying suggestive sentences from online customers' reviews. A classification using a word-embedding approach is used via the XGBoost classifier. The two datasets used in this experiment relate to online hotel reviews and Microsoft Windows App Studio discussion reviews. F1, precision, recall, and accuracy scores are calculated. The results demonstrated that the XGBoost classifier outperforms—with an accuracy of more than 80%. Moreover, the results revealed that suggestion keywords and phrases are the predominant features for suggestion extraction. Thus, this study contributes to knowledge and practice by comparing feature extraction classifiers and identifying XGBoost as a

better suggestion mining process for identifying online reviews. Online texts of reviews and blogs are continuously increasing and constitute public opinions regarding products, services, individuals, organizations, or events.

The expression of sentences in available online text can be related to sentiments and emotions [1], and generally referred to as opinions, ecommendations, instructions, advice, and tips for others regarding any entity. Such opinions can be collectively termed as suggestions [2]. Studies have described suggestion mining as sentence classification, which is based on predicting opinionated text into the binary forms of suggestions and non-suggestions [3–5]. The literature has generally defined suggestion mining as the "extraction of suggestions from the opinionated text, where suggestions keyword denotes the recommendation, advice, and tips" [3]. These sugges- tions are valuable to customers and business organizations [6] if extracted comprehensively from opinionated text [7]. Suggestions must be extracted using computers because online reviews, blogs, and forums that contain suggestions are continuously increasing, resulting in large datasets [6]. The high data volume makes it challenging to extract suggestions [8]; therefore, automatic suggestion mining has emerged as a new research area [1]. Suggestion mining is an approach that largely emphasizes analyzing and identifying sentences to explore explicitly the suggestions they contain [2]. Identifying opinions about products and services that are discussed on social media is useful to organizations' management and to consumers. These opinions offer suggestions that assist management in deciding on improvements to products and services [6]. In addition, consumers can benefit from these suggestions by using them to decide whether to buy a particular product or service. Such increased opinionated text has constituted the major dataset in the majority of recent research [9–11]. Some studies have focused

on product reviews [4,5,12] related to tourism (e.g., hotel service) [10,11] and on social media data (e.g., Twitter) [13]. Moreover, several challenges in suggestion mining approaches relate to analyzing the senti- ments of the sentence, identifying the relationship between suggestions, and

selecting annotators for supervised and unsupervised learning [14]. Suggestion mining is a recent research area, and thus, studies on extracting suggestions involving different classifiers and algorithms are relatively limited [15]. Studies related to support vector machines (SVMs) [16], long short-term memory (LSTM) [8], hidden Markov [17], Random Forest [18,19], Naïve Bayes [20,21], and other areas [22] have also contributed to improvements in suggestion mining. Thus, the present study is among the few such studies that are aimed at improving suggestion mining results by experimenting with the word-embedding approach and the XG Boost classifier. This study is aimed to capture context and similarity with other words. Furthermore this study contributes by improving the classifier performance through the XGBoost classifier, as compared with Naïve Bayes and Random Forest. Moreover, variations in the proposed suggestion mining extraction process casting improved suggestion mining results. The remainder of the paper isstructured as follows. Section 2 describes related work regarding suggestion mining and Section 3 explains the proposed suggestion mining extraction process. Section 4 describes the detailed exper- iment results and Section 5 presents a results analysis and discussion. Last, Section 6 describes the conclusion and future work. The availability of opinionated text regarding social media data is increasing, which can assist in decision-making if extracted and analyzed carefully. The extracted suggestions, tips, and advice must be carefully analyzed to improve the business and subsequently benefit customers. Recent studies have explored suggestions from online reviews through different classifiers, such as Random Forest and Naïve Bayes. The results of these studies are not mature enough and require further improvements. Therefore, this study proposed a suggestion mining process to improve the results further.

A REAL TIME FACE EMOTION CLASSIFICATION ANDRECOGNITION USING DEEP LEARNING MODEL

AUTHORS: Shaik Asif Hussain and Ahlam Salim Abdallah Al Balushi 2020 J. Phys.: Conf. Ser.1432 012087.

Facial Detection and recognition research has been widely studied in recent years. The facial recognition applications plays an important role in many areas such as security,

camera surveillance, identity verification in modern electronic devices, criminal investigations,database management systems and smart card applications etc. This work presents deeplearning algorithms used in facial recognition for accurate identification and detection. Themain objective of facial recognition is to authenticate and identify the facial features. However,the facial features are captured in real time and processed using haar cascade detection. These quential process of the work is defined in three different phases where in the first phase human face is detected from the camera and in the second phase, the captured input is analyzedbased on the features and database used with support of keras convolutional neural network model. In the last phase human face is authenticated to classify the emotions of human as happy, neutral, angry, sad, disgust and surprise. The proposed work presented is simplified in three objectives as face detection, recognition and emotion classification. In support of this work Open CV library, dataset and python programming is used for computer vision techniques involved. In order to prove real time efficacy, an experiment was conducted for multiple students to identify their inner emotions and find physiological changes for each face.The results of the experiments demonstrates the perfections in face analysis system. Finally, theperformance of automatic face detection and recognition is measured with Accuracy. Human computer interaction is a common trend and innate ability to distinguish among multiple faces. Until recent past computer vision problems were quite challenging but advent of modern technologies has trivially improved from the problems of varying light, changed by age, hair and other accessories[1]. However, face recognition applications are used improve access to identify and verify the peopleby their face features. Hence interpreting the facial features and their actions is much required. As these features and expressions helps in classify the emotions of human face. Recent advances in technology has resulted in the use of Artificial intelligence system as these systems are capable to understand and realize the emotion

recognition through facial features. Hence this is an attempt toprove the existence of latest technological developments for human-computer interaction using deep learning or Convolution neural network models [2]. To recognize and classify the human face various methods are required but deep learning technique outperforms other methods by its large capabilities of different datasets and fast computation capabilities. Usually the process of face recognition and classification involves various steps such as pre processing, detection, orientation, extraction offeatures and classification of emotion. The proposed work carries out in three sequential steps as Face Detection, Face Recognition and Face Classification. In the first step a video camera is used to capture the human face and detect the exact location of face by a bounding box coordinates for the face detected in real-time. This stepinvolves face detection using Haar cascade detection with open CV library. Viola jones algorithm andhaar cascade features are combined to detect human face. The images detected have shapes, objectsand landscapes etc. In this phase human face is detected and face features are extracted and stored in the database for face recognition. The CNN model as shown in figure 4 uses VGG 16 to match theface from the database and recognize with the name for the face detected. Faces are recognized from the database and are compared to identify or detect the face through embedding vectors. The

distribution platform use Anaconda and python 3.5 software in processing face detection, recognition and classification. The image features in the database dlib and other libraries. First face is detected and then recognized with the database features and matching using CNN model training and testingdatabase. Finally the recognized human face is classified based on the expression in real time asAngry, fear, disgust, happy, neutral and surprise. The network architecture VGG 16 is built with CNNmodel for large database recognition and classification. The designed network model has honey comb3 x 3 layers where the two connected layers have 4096 nodes with Softmax classification. The localbinary model histogram is used as open CV library for detecting the human faces. The image pixels

are identified by setting a threshold and the

end result is represented in form of a binary number.

SPATIAL FEATURE FUSION FOR BIOMEDICAL IMAGECLASSIFICATION BASED ON ENSEMBLE DEEP CNN ANDTRANSFER LEARNING

AUTHORS : Sanskruti Patel, Rachana Patel, NilayGanatra, Atul PatelSmt. Chandaben Mohanbhai Patel Institute of Computer Applications Charotar University of Science and Technology, Changa, India

Biomedical imaging is a rapidly evolving field thatcovers different types of imaging techniques which are used for diagnostic and therapeutic purposes. It plays a vital role in diagnosis and treating health conditions of human body. Classification of different imaging modalities plays a vital role interms of providing better care and treatment options to the patients. Advancements in technology open up the new doors formedical professionals and this involves deep learning methods for automatic image classification. Convolutional neural network(CNN) is a special class of deep learning that is applied to visual imagery. In this paper, a novel spatial feature fusion based deep CNN is proposed for classification of microscopic peripheral blood cell images. In this work, multiple transfer learning features are extracted through four pre-trained CNN architectures namely VGG19, ResNet50, MobileNetV2 andDenseNet169. These features are fused into a generalized feature space that increases the classification accuracy. The dataset considered for the experiment contains 17902 microscopic images that are categorized into 8 distinct classes. The result shows that the proposed CNN model with fusion of multiple transfer learning features outperforms the individual pre-trained CNNmodel. The proposed model achieved 96.10% accuracy, 96.55%F1-score, 96.40% Precision and 96.70% Recall values. Biomedical imaging refers to capturing of an organ ortissue for diagnostic purpose. The field is very broad andrapidly evolving that covers different types of imaging modalities like ultrasound, magnetic resonance imaging (MRI),computerized tomography (CT), positron emission tomography(PET), etc. [1]. Biomedical and medical imaging plays asignificant role in diagnosis and treating health conditions

ofhuman body. It helps to identify problematic health conditionsin their early stages that certainly lead for providing better treatment to the patients [2]. The structural and functional changes in biological tissues of the human body normally cause the possible health problems. Biomedical imaging

provides a way to view inside the human body that helps toreveal such changes [3]. Classification of different imaging modalities plays a vitalrole in terms of providing better care and treatment options tothe patients. The traditional way to classify these different modalities of images is the naked eye classification that is performed by medical professional or subject expert. This isthe cumbersome and sometimes time-consuming method. Advancements in technology open up the new doors formedical professionals that involve computer-aided diagnosis(CAD) methods for automatic image classification [4]. The increasing advancement in the field of medical imaging technology, medical research and diagnosis become easy. Various types of imaging modalities and procedures are included in medical imaging technology that helps in diagnosis and treatment of the patients. Hence, it plays a dominant role indeciding the actions for the benefit of the patient's health. In past few years, artificial intelligence (AI) brings a new way for analysing and interpreting data that also called predictive analysis that helps to identify the early signs of any of the health conditions [5]. Deep learning, an approach of AI,emerged as an outperformer in interpreting and analyzing image data. Significant advancement has been made in the field of medical image diagnosis that improves disease diagnosis process significantly. Deep learning uses the architecture of artificial neural network that mimics the working of human brain. The complex computer vision tasksare effectively solved by the deep learning algorithms such asimage recognition, classification and segmentation. The special class of deep leaning algorithms is known as a convolutional neural network (CNN) is widely used to solve image classification problems and achieved a significant performance on benchmark datasets. The reason behind popularity of CNNis the large availabilities of datasets and support of powerful Graphics Processing Units (GPU)

that makes the integration ofdeep learning methods with computer vision popular [6].There are several distinct imaging modalities in whichbiomedical images are generated. They are different in shapeand types. Due to the diverse data distribution patterns, it mayhappen that the same CNN model may show differentperformance on different datasets [7]. CNN models are sensitive to the particulars of the training data. This makes possible that each time they are trained; they may find a different set of weights. These different predictions generate high variance [8]. Moreover, these deep features face the problem of small intra-class variance and large inter-class variance [9].

LEARNING AFFECTIVE VIDEO FEATURES FORFACIAL EXPRESSION RECOGNITIONVIA HYBRID DEEP LEARNING

AUTHORS :SHIQING ZHANG 1, XIANZHANG PAN1, YUELI CUI1, XIAOMING ZHAO1, AND LIMEI LIU

One key challenging issues of facial expression recognition (FER) in video sequences isto extract discriminative spatiotemporal video features from facial expression images in video sequences.In this paper, we propose a new method of FER in video sequences via a hybrid deep learning model.The proposed method first employs two individual deep convolutional neural networks (CNNs), includinga spatial CNN processing static facial images and a temporal CN network processing optical flow images,to separately learn high-level spatial and temporal features on the divided video segments. These two CNNsare fine-tuned on target video facial expression datasets from a pre-trained CNN model. Then, the obtainedsegment-level spatial and temporal features are integrated into a deep fusion network built with a deep beliefnetwork (DBN) model. This deep fusion network is used to jointly learn discriminative spatiotemporalfeatures. Finally, an average pooling is performed on the learned DBN segment-level features in a videosequence, to produce a fixed-length global video feature representation. Based on the global video feature representations, a linear support vector machine (SVM) is employed for facial expression classification tasks.The extensive experiments on three public video-based

facial expression datasets, i.e., BAUM-1s, RML,and MMI, show the effectiveness of our proposed method, outperforming the state-of-the-arts. Facial expression is one of the most natural nonverbal waysfor expressing human emotions and intentions. In recentyears, automatic facial expression recognition (FER), whichaims to analyze and understand human facial behavior, hasbecome an increasingly active research topic in the domainsof computer vision, artificial intelligence, pattern recognition,etc. This is because FER has many potential applicationssuch as human emotion perception, social robotics, human-computer interaction and healthcare [1]–[5]. Inspired by the strong feature learning ability of deep neural networks, this paper proposes a new deep neural network-based FER method in video sequences by using a hybrid deeplearning model. Our hybrid deep learning model containsthree deep models. The first two deep models are deep Con-volutional Neural Networks (CNNs) [16], including a spatialCNN network processing static facial images and a tem-poral CNN network processing optical flow images. Thesetwo CNNs are separately used to learn high-level spatialfeatures and temporal features on the divided video segments.The third deep model is a deep fusion network built with aDeep Belief Network (DBN) [17] model, which is trained tojointly learn a discriminative spatio-temporal segment-levelfeature representation. When finishing the joint training ofa DBN, an average-pooling is applied on all the divided videosegments to produce a fixed-length global video feature representation. Then, a linear Support Vector Machine (SVM)is adopted to perform facial expression classification tasks invideo sequences.

This paper proposes a hybrid deep learning model, whichconsists of the spatial CNN network, the temporal CNNnetwork, and the DBN fusion network, to apply for FER invideo sequences. We implement our proposed method in twostages. (1) We employ the existing VGG16 model pre-tainedon ImageNet data to individually fine-tune the spatial CNNnetwork and the temporal CNN network on target video-based facial expression data. (2) To deeply fuse the learnedspatio-temporal CNN features, we train a deep DBN model to jointly learn discriminative spatio-temporal features. Exper-iment results on three public

video-based facial expressiondatasets, i.e., BAUM-1s RML, and MMI, demonstrate the advantages of our proposed method.In future, we will extend our work to practical applications.For instance, it is challenging to develop a real-time FERsystem based on our proposed method. In addition, it is alsointeresting to explore deep compression of deep models so asto reduce the large network parameters of deep models.

CROWDSOURCED REQUIREMENTS ENGINEERING CHALLENGES AND SOLUTIONS:A SOFTWARE INDUSTRY PERSPECTIVE

AUTHORS : Huma Hayat Khan 1,*, Muhammad Noman Malik 2, Youseef Alotaibi3 , AbdulmajeedAlsufyani 4 andSaleh Alghamdi

Software crowdsourcing (SW CS) is an evolving software development paradigm, in which crowds of people are asked to solve various problems through an open call (with the encouragement of prizes for the top solutions). Because of its dynamic nature, SW CS has been progressively accepted and adopted in the software industry. However, issues pertinent to the understanding of requirements among crowds of people and requirements engineers are yet to be clarified and explained. If the requirements are not clear to the development team, it has a significant effect on the quality of the software product. This study aims to identify the potential challenges faced by requirements engineers when conducting the SW–CS based requirements engineering (RE) process. Moreover, solutions to overcome these challenges are also identified. Qualitative data analysis is performed on the interview data collected from software industry professionals. Consequently, 20 SW–CS based RE challenges and their subsequent proposed solutions are devised, which are further grouped under seven categories. This study is beneficial for academicians, researchers and practitioners by providing detailed SW–CS based RE challenges and subsequent solutions that could eventually guide them to understand and effectively implement RE in SW CS. SW CS is the assignation of a worldwide pool of online developers who can be appointed on-demand tocontribute to several types of software development tasks [5–7]. The

process is facilitated by platforms thatlink requesters and online developers. On this platform, the requester disseminates tasks to volunteer online developers, who solve the tasks based on their motivation for reward (e.g., money or respect). The SW CSplatform has significant importance because it provides guidelines for managing and coordinating theprocesses and people in business and at technical levels [5,8]. Moreover, the platform permits requestersto discover talent outside their borders and earn the benefits of cost, quality, time and expertise [3,5,6]. Inessence, this reveals that the software development team encounters an extensive user audience, which isreferred to as the crowd of people. It is important to involve the crowd to satisfy crowd requirements,which demonstrates the importance of crowdsourcing [9]. Involving an enormous number of users inrequirements engineering (RE) has always been challenging with customary RE methods [10,11]; specifically, this is true when RE would include an enormous number of users (a crowd) who are beyondan organisation's reach [10,12].Traditional approaches to conducting the RE process typically involve an inadequate number of representatives in requirement-gathering sessions through interviews or focus groups [13]. Conversely, recent RE approaches, which are functional in market-driven RE, permit organisations to directly interact with main stakeholders who employ ad hoc feedback-gathering networks [14]. However, these recent RE approaches miss the prospect of uninterruptedly involving diverse groups of users who share their feedback through a variety of media [12,15,16]. Consequently, it is believed that the software evelopment team cannot consider the varied circumstances of user subgroups when evolving the product[17,18]. So, valued resources for RE continue to be unexploited and software products might not fulfillusers' needs [2,11].CS-based RE is an umbrella term used for computerised or semi-computerised approaches to elicit andanalyse information from a crowd to stem authenticated user requirements [19]. Usually, the crowd is anundefined set of people [5]; however, in the context of CS-based RE, a crowd is usually a large group ofcurrent or prospective software users who act together or with the software organisation's representatives(e.g., product owner or

development team) [11]. These users are involved at the run time for requirement elicitation through users' feedback [20]. CS is considered a promising paradigm for eliciting requirements of a software system in a dynamic, possibly unknown context with a large pool of users [21]. CS-based RE is likely to enhance the quality, inclusiveness and even monetary viability of requirements elicitation. It allows software development teams to receive more updated knowledge about users' perceptions of the system's role to fulfil their requirements [9].

EMOTIONET CHALLENGE: RECOGNITION OF FACIAL EXPRESSIONS OF EMOTION IN THE WILD

AUTHORS :C. Fabian Benitez-Quiroz, Ramprakash Srinivasan, Qianli Feng, Yan Wang, Aleix M. MartinezDept. Electrical and Computer Engineering
The Ohio State University

This paper details the methodology and results of the EmotioNet challenge. This challenge is the first to test the ability of computer vision algorithms in the automatic anal-ysis of a large number of images of facial expressions of emotion in the wild. The challenge was divided into two tracks. The first track tested the ability of current computer vision algorithms in the automatic detection of action units(AUs). Specifically, we tested the detection of 11 AUs. Thesecond track tested the algorithms' ability to recognize emo-tion categories in images of facial expressions. Specifically, we tested the recognition of 16 basic and compound emo-tion categories. The results of the challenge suggest that current computer vision and machine learning algorithms are unable to reliably solve these two tasks. The limitations of current algorithms are more apparent when trying to rec-ognize emotion. We also show that current algorithms arenot affected by mild resolution changes, small occluders,gender or age, but that 3D pose is a major limiting factor
on performance. We provide an in-depth discussion of thepoints that need special attention moving forward. Much progress has been made in computer vision inthe last couple decades. Of late, tremendous efforts havebeen made to design algorithms that can

detect and rec-ognize generic and specific object categories in images, with tremendous achievements recorded just in the last few years. ImageNet, PASCAL-VOC and COCO are three ex-ample databases and challenges [16, 6, 12] that have helped fuel the progress and success of these efforts. In the present paper, we ask if these achievements ex-tend to the detection and recognition of facial expressionsof emotion in the wild. Previous studies have reported good results in the automatic analysis of facial expressions of emotion [2]. However, these results were obtained with the analysis of images and videos taken in the laboratory [23, 3]. That is, even when the expressions were sponta-neous, the filming was done in controlled conditions withthe full awareness of the participants. Also, researchers had access to both the training and testing data, allowing prac-titioners to modify their algorithms until these work on the known testing data.To evaluate whether current computer vision algorithm scan fully automatically analyze images of facial expressions

of emotion in the wild, we collected a large database of 1million image [1]. This datase is called EmotioNet. We also developed a computational model based on the latest knowledge of how the human brain analyzes facial expres-sions [15]. This algorithm was used to annotate a subsetof 950,000 images of the EmotioNet set. The remaining 50,000 images were manually labelled by expert annotators.The 950k images that were automatically annotated by our algorithm defined the training set. The training set was made available to participants of the challenge. 25k of the manually annotated images were also made available to par-ticipants as a validation set. This was necessary because the images in the training set are not accurately annotated, i.e.,the labels are unreliable. The accuracy of annotations in the training set is about 81%. The validation set can thus be used to determine how well one's algorithm adapts to this unreliability. The EmotioNet Challenge was divided into two tracks.The first track tested computer vision algorithms' ability indetecting 11 AUs. This is a task successfully completed by human experts only. The second track assessed com-puter vision algorithms' ability to recognize 16 emotion categories. This is a task successfully completed by most peo-

ple, experts and non-experts.The images in the EmotioNet challenge are of facial expressions of emotion in the wild, Figures 1-3. This is amajor departure from previous studies and challenges were images of facial expressions of emotion had been collected in controlled conditions in the lab. The results of the EmotioNet Challenge illustrate the dif-ficulty of these two task in images of facial expressions of emotion in the wild. Nonetheless, a surprising and inter-esting outcome of the challenge is the extra difficulty algorithms have in successfully completing the task of the second track.

3. EXISTING SYSTEM

In existing system, neural network architecture has utilized to detect the face that captured by the webcam present in system. Then the face image only can take for face recognition process. Detecting the face on live video is not implemented. When we try to attain the same result by applying live video, the accuracy got decreased. So theexisting system was applicable only for an image input.

DRAWBACKS OF EXISTING SYSTEM

- Accuracy is less when applying to live video.
- Time taken to process the image is high.
- Not applicable for live video.
- Missed to process high pixel images.

3.1 PROPOSED SYSTEM

In this system, LSTM - **Long short-term memory**algorithm is implemented to detect the face on live video taking by the webcam present in system. Then the face present in front of webcam will be highlighted and proceeds to check with dataset. This dataset contains trained face images. Each face captured with 30 sets of image for increasing the accuracy level. Before applying the recognition process the dataset should be trained by authorized faces. The algorithm is designed to train more than one face. Detecting the face on live video is difficult. By using the LSTM architecture the object can be attained with more accuracy.

INTERNATIONAL JOURNAL OF APPLIED SCIENCE ENGINEERING AND MANAGEMENT

ISSN2454-9940

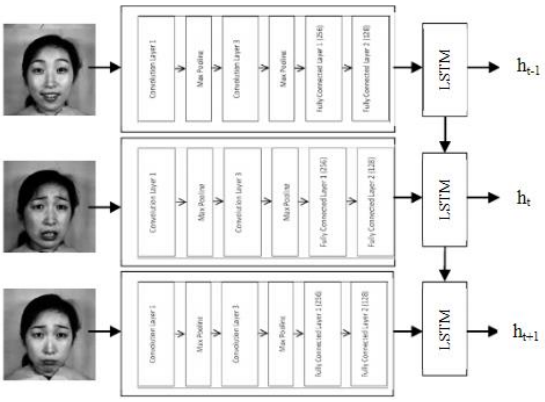
www.ijasem.org

Vol 18, Issue 2, 2024

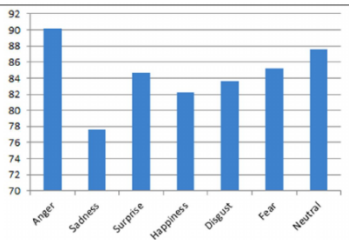
ADVANTAGE OF PROPOSED SYSTEM

- Accuracy is high.
- Enhancing the structure of face is attained accurately.
- It takes few seconds to provide exact result.
- Applicable for live video taking by webcam.
- Applicable for both low and high pixel images.

4. OUTPUTSCREENS



	Angry	Sadness	Surprise	Happiness	Disgust	Fear	Neutral
Angry	**90.38**	3.85	0.96	0	4.81	0	0
Sadness	1.63	**78.05**	2.44	7.32	2.03	5.28	3.25
Surprise	0	0.88	**84.58**	0	0	9.25	5.29
Happiness	2.02	4.03	0.81	**81.85**	0	1.61	9.68
Disgust	6.33	4.08	0	1.81	**83.26**	2.71	1.81
Fear	1.63	6.10	1.22	0.81	1.63	**84.96**	3.65
Neutral	0	5.07	2.30	4.61	0	0	**88.02**



5. CONCLUSION

We use a sequence of recurring neuronal networks and feedforward neural nets in an over-to-finish architecture to identify the problem to deal with nonlinear deformations in frontal shapes of face imagery. At first CNN-RNN network maps the face image's input directly to the estimated face form in low resolution. The subsequent local network then takes the components of faces as inputs based on the initial network output to obtain more accurate, higher resolution landmarks. Finally, by removing a small

patch of each landmark and refining all the sites together, the local network adjusts results. Our method obtains impressive results when dealing with face images in videos through the full use of the time series. Moreover, our method also achieves good results on still images by using more images from RGB channels or enhanced images. In comparison with state-of - the-art methods, such as CFAN, SDM, TCDN, RCPR, the effectiveness of our method is validated on three databases in wild as well as interior conditions.

6. REFERENCE

[1] Y. Alotaibi, M. Noman Malik, H. Hayat Khan, A. Batool,A. Alsufyani, and S. Alghamdi, "Suggestion mining from opinionated text of big social media data," Computers, Ma-terials& Continua, vol. 68, no. 3, pp. 3323–3338, 2021.

[2] S. A. Hussain and A. S. A. Al Balushi, A Real Time FaceEmotion Classification and Recognition Using Deep LearningModel, IOP Publishing Ltd, Bristol, UK, 2019.

[3] H. Jiang and E. Learned-Miller, "Face detection with the fasterR-CNN," in Proceedings of the 2017 12th IEEE International Conference on Automatic Face & Gesture Recognition (FG2017), Washington, DC, USA, May 2017.

[4] S. Zhang, X. Pan, Y. Cui, X. Zhao, and L. Liu, "Learningaffective video features for facial expression recognition viahybrid deep learning," IEEE Transactions on Multimedia,vol. 7, 2019.

[5] H. Hayat Khan, M. Noman Malik, Y. Alotaibi, A. Alsufyani,and S. Alghamdi, "Crowdsourced requirements engineeringchallenges and solutions: a software industry perspective,"Computer Systems Science and Engineering, vol. 39, no. 2,pp. 221–236, 2021

[6] C. Fabian Benitez-Quiroz, R. Srinivasan, Q. Feng, Y. Wang,and A. M. Martinez, "Emotionet challenge: recognition offacial expressions of emotion in the wild," 2017, org/abs/1703.01210.