Comparative analysis of Identification and Classification of Face Emotions Using Different Machine Learning and Deep Learning Algorithms

Dasharath.K.Bhadangkar[1]
S.D.M.College of Engg. &Tech, Dharwar – 580 008, INDIA,
dashrathb@gmail.com
Jagadeesh D. Pujari[2]
S.D.M.College of Engg. &Tech, Dharwar – 580 008, INDIA,
jaggudp@yahoo.com

ABSTRACT: Sentiments are significant and profound features of individual conduct. Examining facial looks and acknowledging their expressive condition is stimulating job by extensive reaching functions. Human Face expression Recognition is one of very influential and stimulating chores in public interaction. Usually, face expressions are usual and straight ways for human beings for communicate their emotions and intentions. Face expressions are the key characteristics of non-verbal communication. Here, we introduce available dataset i.e., CK+, JAFFE and FED dataset that are widely used in this work. This paper focuses upon facial expression recognition technique founded on machine learning algorithms pair and also deep learning algorithms that will assist in precise recognition and organization of human emotion.
KEYWORDS: FED, SVM, KNN, RNN, CNN, DBN;

I. Introduction

Humans are expressive beings. Our expressive condition notifies how we perform by very important procedures, to compound activities and hard choices. Our lives are in numerous paths directed from our feelings, so understanding additional regarding feelings lets us to see more regarding human conduct extra usually. It's apparent that comprehending expressive condition of human could be valuable for variety of uses by improving good comprehension of human psychology, for examining behavior for better consumer practices, for improving useful publicity canvasses, and outside.

Human emotions are categorized like: fear, contempt, disgust, and anger, surprise, sad, happy, and neutral. These feelings are extremely delicate. Facial musculus frowns are extremely slight and sensing these changes is extremely daring as even minor change consequences in changed expressions [1]. Expressions of dissimilar or even same human may differ to same emotion as well, as emotions are immensely situation supported [2]. Although one could emphasize on only those parts of face that show most of feelings such as about mouth and eyes [3], how we obtain these signs and classify them is yet an significant query. Neural networks and machine learning are utilized to these chores and contain gotten decent outcomes. Machine learning algorithms have established that they are extremely useful in pattern detection and categorization. Highly significant features for any machine learning algorithm are characteristics.

Facial expression detection utilizes algorithm for detecting faces, codes facial expression and spot expressive conditions. It performs this from examining faces in images or videos by cameras fixed in laptops, phones, or computers. This recognizes diverse feelings on human face, business images plus videos in present for examining video supplies. Facial expression detection discovers uses in animated cinemas, check person's pressure stage, mining demonstrations of psychiatric patients. This is related in motorist's sleepiness discovery as well. Intelligent automobiles can aware driver when he is sensing sleep from initially noticing his face and later eyes. This is appropriate in emotion recognition in meeting for determining whether applicant's character is decent appropriate for work as well. This is utilized in analysis in video games as well. Throughout examining stage, consumers are requested to compete in game for specified time and its response is combined for making last creation.

FED (Face Expression Dataset) possesses significant phase is feature extraction and cataloguing. Feature extraction comprises 2 kinds and these are symmetrical founded and look founded. Cataloging is one of significant procedures as well where aforementioned expressions like smirk, gloomy, fury, repugnance, astonishment, and distress are characterized. Geometrically founded characteristic abstraction includes eye, mouth, nose, eyebrow, other facial parts plus look-founded characteristic abstraction includes precise unit of face [4].

The overview of the Face expression recognition structure is demonstrated in Fig. 1. Face expression recognition scheme comprises main phases like face image pre-treating, feature abstraction and cataloging.

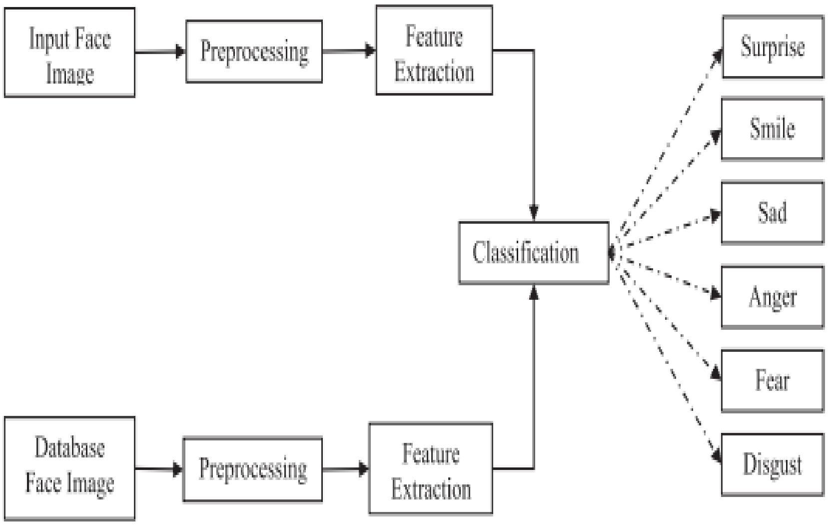


Fig. 1. Architecture of face expression detection structure

II. RELATED WORK

Problem of detection and fortitude of human feelings is always a significant study part from the period of Charles Darwin who initially pointed connection amid induced expressive conditions and typical expressions of human face. Darwin's model is established from numerous various surveys in following periods [5–7]. He presumed that emotional expressions are multimodal behavioural outlines of person, and therefore shaped his personal full depictions of 40+ emotional conditions [8]. Over last century, numerous diverse replicas for emotion arrangements, fluctuating from generally felt elementary emotions to exclusive and compound ones were sensitively demarcated. Two replicas investigated in arena of emotion detection [9, 10] are mainly utilized in preceding period: elementary cataloguing of 6 emotional conditions from Ekman and Russell's circumflex prototype of emotions.

For present feeling detection structure, numerous methods are predictable. Nominal number of AUs are used for detecting facial expressions [11, 12], even though changing numbers of AUs utilized in evolving sharp facial expression detection structure. Delaunay triangulation technique is utilized for connecting 68 facial activity parts in focussed face for detecting 7 facial expressions like contentment, fury, terrified, astonishment, grief, neutral, plus repugnance. Multi class support vector machine (SVM) classifier executes fine having highest mean facial expression detection degree of 84% founded on spatiotemporal traits in categorizing emotional expressions [13].

Kartali et al. have described outcomes of orthodox (SVM, Multi-Layer Perceptron (MLP)) and deep studying approaches (Convolutional Neural Network (CNN), Alexnet CNN, Affdex CNN) founded facial expression detection of 4 emotions (contentment, grief, fury, and distress) plus attained highest detection precision of 85.05% utilizing Affdex CNN [14]. New vectored emotion detection prototype is projected for identifying 3 main emotions: livid, pleased, plus neutral, utilizing seventy facial vectors and deep neural network (DNN), and attained average precision of 84.33% [15]. In new works, researchers have utilized altitudinal and chronological data by input video arrangements for classifying diverse facial expressions utilizing CNN, Collaborative Multi-level CNN, and Long Short-term Memory (LSTM) [16–19]. Few usual subjects stated in previous literatures because of absence of examples or information groups, minimal precision in categorizing facial expressions, advanced computational difficulty (extra memory and power needed to process information), not appropriate for present uses, and not easy method (limitations in utilizing structure for diversity of uses) [20].

III.DATASET

In Present scenario Facial Expressions acts a significant part in all fields like Scientific, Medical, Social, business, gaming, child psychology etc. In this work we have used CK+, JAFFE and FED Dataset.

CK+: Lengthy Cohn-Kanade (known as CK+) facial expression folder [22] is community folder for action part and emotion detection. This comprises posed and non-posed (impulsive) expressions as well. CK+ includes whole of 593 series thru 123 subjects. In majority of preceding compositions, final frame of these orders is selected and utilized for image

JAFFE: Japanese Female Facial Expression (JAFFE) folder is lab- regulated image folder which comprises 213 models of modelled expressions by 10 Japanese ladies. Every person possesses 3˜4 images giving all 6 elementary facial expressions (fury, repugnance, distress, pleasure, grief, and astonishment) and 1 image having neutral expression. Folder is difficult since this comprises some instances per subject/expression. Classically, every image is utilized for leave-one-subject-out test.

FED: Facial expressions dataset with respect to perfect lighting conditions and lab environment. Subjects photographed are between the age 5 to 70, school children, college students, professionals and stage artists have participated in this process with the consent. Photography is done in various places such as school, college and drama stages.

Method

Development of the FED Database. The participants were 49 Children (M : 12.65 years, SD : 0.57; age range, 12–16; 57% Females), 41 Young (M : 20.24 years, SD : 1.15; age span, 19–31; 58% women), middle aged 30 (M: 41.06 years, SD : 4.26; age span, 38–55; 46% women) Old Aged 58 and above (M: 61.33 Years , SD:0.16 ;age span 58-80) extras, specials, or performers who only had certain exclusions, Every contributors are from school, college and drama school.

Before to photo-shoot gathering, contributors were educated through mobile, e-mail, or fax that goal of scheme was depositing record including portrayals of child, young, middle-aged, and older adults showing 6 various facial expressions to utilize in technical studies. Contenders who are proficient of convey every 6 various faces (by aid of face teaching and by backing of skilled practiced associate). Applicants were requested to approach with no cosmetics and attire regular dress on photo-shooting gathering day.

Procedure and Materials.
Photo-shooting shifts happened at KSS amid FEB 2019 and May 2019 in photo studio precisely arranged for this use. Skilled expert associate notified participants regarding overall goal of scheme plus of day's shift, and also regarding specific process in photo studio. Contributors were expressed that they would be snapped numerous instances, displaying every 6 facial expressions. Photography professional and photography associate aided contributors display these facial expressions by afresh industrialized process which included 3 (partially mixed).

Validation of the Facial expression Database
Validation of Facial Expression Database is done to help the research people with help of assessors, all the 600 images are perceived as child, young, middle aged and old aged.

Participants in the validation study many contributors were requested for validating face data, they have asked for identifying the different emotions by rating them. Totally 600 images were selected.

IV. Machine Learning Algorithms To Recognize And Classify Face Emotions

Support Vector Machines (SVM)
Support Vector Machine (SVM) was initially perceived in 1992, presented from Boser, Guyon, and Vapnik in COLT-92. These are group of linked regulated learning approaches utilized to classify and regression. They are of group of comprehensive linear classifiers. In another words, this is cataloguing and regression estimate device which utilizes machine learning model for maximizing prognostic precision although unconsciously evading over-fit to info. These is outlined as structures that utilize theory space of linear functions in high dimensional feature space, directed by learning algorithm by optimization model which applies learning preconception resultant by statistical learning model.

SVM are very influential cataloguing algorithms. The outline is finding ideal hyper plane that splits 2 groups precisely. There is notion of boundary as well, that is meant to greatest among both groups for avoiding further imbrication amid 2 groups [21]. Info that is not linearly divisible is charted into advanced dimension for achieving healthier cataloguing outcomes. Kernel functions like radial basis function (rbf) and polynomial are utilized for non-linear info [22].

In the event of emotion recognition, typically multi-class SVM is utilized in place of binary for detecting emotions like fury, disdain, repugnance, distress, pleased, dejection and astonishment. K-fold cross-authentication is utilized for removing somewhat alterations in folder and for comparing diverse machine learning algorithms [23]. In k-fold cross authentication, folder is split k times to k shares, and forecast outcomes are averaged over every iteration. Loconsole et al. [24] utilized Principal component analysis (PCA) for characteristic set decrease and later nourishing condensed characteristic fit to SVM. In PCA algorithm, image characteristic distance is altered to eigen distance utilizing eigen matrix [25].

Alongside kernel arrangement, SVM possess approaches for changing limits such as C and γ [26]. In this, C is penalty function for misclassification and gamma aids in optimising decision boundary. Both limits disturb precision of classifiers and is altered for getting best consequences in both binary and multi-class cataloguing.

K nearest neighbour (KNN)

KNN is plain nonlinear classifier utilized in numerous purposes, counting epilepsy recognition, driver sleepiness recognition, emotion detection, seizure recognition, and numerous extra difficulties. KNN is non-probabilistic learning algorithm utilized for classifying unknown trial info founded on bulk of alike info amongst k-nearest neighbours nearby trial/unidentified info. Diverse distance quantities can quantify space among trial info and every training info, like Manhattan, Euclidean, Minkowski, and Chebyshev. Here, overhead 4 space capacities are utilized for distinguishing facial emotional demonstrations, and average Precision of every space measure is stated.

In here, k-Nearest Neighbours algorithm (or k-NN) is non-parametric technique that is utilized to classify and regression. Input contains of k neighbouring training instances in characteristics space. Output hangs on whether k-NN is utilized for regression or cataloguing. In k-NN cataloguing, output is group associate. Object is categorised from bulk support of their neighbour, with object being allocated to group very usual amongst their k adjoining neighbour (k is +ve integer, k<1). If k=1, nevertheless entity is only allocated to group of that solitary adjoining neighbour.

In K-NN regression, output is property quantity for the object. This value is mean of values of their k adjoining neighbors. K-NN originates below example-founded learning, or lazy learning, in which function is simply estimated locally and every assessment is delayed till cataloguing. KNN algorithm is amongst modest of every machine learning algorithm regarding to cataloguing and regression, this could be valuable to weight aids of neighbours, therefore close neighbours donate extra to mean compared to additional far ones. For instance, regular weighing system contains in providing every neighbour weight of 1/d, in which d is distance to neighbour. Neighbours are occupied by group of items for which object property amount (for K-NN regression) and class (for K-NN classification) is identified. This could be believed of as training group for algorithm; however no clear training phase is obligatory.

In K-NN cataloguing, training designs are planned in d dimensional period, in which d is quantity of characteristics existing. These designs are planned corresponding to their experiential characteristics quantities plus are labelled corresponding to their recognized group. An unlabelled test design is planned inside similar space and is categorised corresponding to extremely commonly happening group amid its k extremely alike training designs; its adjoining neighbours. Very usual resemblance amount for K-NN organization is Euclidian distance metric, outlined among feature vectors $\vec{x}$ and $\vec{y}$ as:

$$euc(\vec{x}, \vec{y}) = \sqrt{\sum_{i=1}^{f} (x_i - y_i)^2}$$

In which, f signifies quantity of characteristics utilized for representing every form. Minor space amounts signify bigger resemblance. Cataloguing happens subsequently classifying k utmost alike training points to inquiry point. Relatively than utilizing normal voting system, algorithm utilized here allocates class tags to query points utilizing prejudiced system founded upon every neighbour's immediacy to query point. Let d be distance amount, and x1, x2, x,...,xk be k adjoining neighbours of x decided in escalating sequence of d(xi,x). So x1 is primary adjoining neighbour of x. They suggest for assigning weight wi to i-th adjacent neighbor x_i characterised as:

$$w_i = \begin{cases} \frac{d(x_k,x)-d(x_i,x)}{d(x_k,x)-d(x_1,x)}, & \text{if } d(x_k,x) \neq d(x_1,x) \\ 1, & \text{if } d(x_k,x) = d(x_1,x) \end{cases}$$

Pattern x is allotted to group having weights of envois amongst k adjoining neighbours amount to highest quantity. This regulation was displayed to produce lesser fault degrees compared to those gotten utilizing voting K-NN regulation.

V. Deep Learning Algorithms To Recognize And Classify Face Emotions

Convolutional neural network (CNN)

In neural networks, ConvNets or CNNs is a chief group to perform images detection, images guiding. Items recognitions, detecting faces etc., are parts inwhich CNNs are extensively utilized. CNN were stimulated from biotic procedure in which link relationship amid neurons looks like configuration of human pictorial cortex. Computers perceives input image as collection of pixels and this hangs on image resolution.

Founded on image resolution, this will perceive h x w x d (h = Height, w = Width, d = Dimension). Elementary section illustration of CNN is exhibited in fig. 2. It contains 3 levels: Convolution layer, Pooling Layer and Fully connected layer.

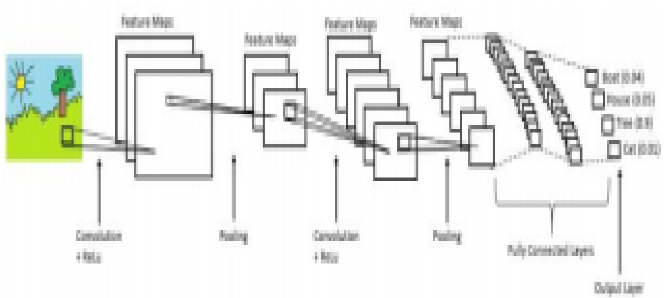


Fig. 2. Block diagram of CNN

1) Convolution Layer

This is primary layer for extracting characteristics by input image. Convolution conserves association amid pixels from learning image characteristic utilizing little squares of input info. This is scientific process which uses 2 inputs like image matrix plus sieve or kernel. Sizes are as ensues:

Image matrix; h ×w ×d

Filter; $fh \times fw \times d$

Output; $(h-fh+1) \times (w-fw+1) \times 1$

This output is denoted as characteristics map. ReLU situates for Rectified Linear Unit for non-linear process. Output is $f(x) = \max(0,x)$.

2) Pooling Layer

This unit will decrease quantity of limits when images are excessively big. Spatial pooling a.k.a subsampling or down sampling that declines dimensionality of every map nevertheless possesses significant info.

Spatial pooling has numerous classes:

- Max Pooling
- Average Pooling
- Sum Pooling

Max. Pooling uses major section by corrected characteristic map. Captivating main component might captivate mean pooling as well. Totality of every element in characteristic map is called as sum pooling.

3) Fully Connected Layer

Layer a.k.a FC layer, we compressed our matrix as vector and supply that as completely linked level as neural network. This performs like human neuron, that interlinks along one another for broadcast of data. Every characteristic maps by pooling unit are interlinked for providing trained output. Hereafter, output level will distinguish trained images. This trained image is utilized to compare to novel image. Subsequently contrast appropriate expression is documented.

CNN is significantly utilized in varied computer vision uses, comprising Fer. At start of 21st period, numerous lessons in FER literature [27], [28] discovered that CNN is strong for facing position variations and scale differences and acts well compared to multilayer perceptron (MLP) in circumstance of formerly hidden confront posture differences. [29] Working CNN for addressing difficulties of topic freedom and conversion, revolution, and scale invariance in detection of facial expressions also. CNN has 3 kinds of varied levels: convolutional levels, pooling levels, and completely linked levels. Convolutional level contain group of learnable sieves for convolving throughout entire input image and give numerous precise kinds of start characteristic maps. Convolution process is related to 3 chief advantages: local connectivity, that studies associations amongst neighbouring pixels; weight sharing in similar characteristic map, that importantly decreases amount of limits to be studied; plus shift-invariance to position of item. Pooling level trails convolutional level and is utilized for reducing latitudinal volume of characteristic maps and computational price of system. Regular pooling and max pooling are 2 utmost usually utilized nonlinear down-sampling plans to conversion invariance. Completely linked level is regularly comprised at close of system for ensuring that every neuron in level are completely linked to starts in preceding level and for enabling 2D characteristic maps to be transformed to 1D characteristic maps for more characteristic illustration and cataloguing. We catalogue formations and features of few famous CNN prototypes which are used for FER in Table 3. Also these systems, numerous famous resultant structures subsist as well. In [30], [31], region-based CNN (R-CNN) [32] was used to learning characteristics for FER. In [33], Faster R-CNN [34] was utilized for identifying facial expressions from producing great excellence area offers. Furthermore, Ji et al. projected 3D CNN [35] for capturing motion info programmed in manifold nearby limits for action detection through 3D convolutions. Tran et al. [36] projected well-made C3D, that uses 3D convolutions on important managed training folders for learning spatio-temporal characteristics.

CNN is distinctive and extensively utilized prototype for deep learning. Deep learning goals for inevitably study and abstract multilevel characteristic illustration by unprocessed info. Features of CNN, like local assembly, weight distribution, and down selection process, making it promising to efficiently reducing difficulty of system, decrease number of teaching limits, and give benefits of tough sturdiness and error lenience, and being simple for training and optimizing also.

Deep belief network (DBN)

DBN projected from Hinton et al. [37] is graphical prototype which absorbs for extracting deep graded illustration of training info. Conventional DBN is constructed using load of restricted Boltzmann machines (RBMs) [38], that are 2 level reproductive stochastic prototypes comprised of visible-unit level plus concealed component level. These 2 levels in RBM should shape two-part graph with no side networks. In DBN, components in advanced levels are coached for learning provisional dependences amongst components in adjoining inferior levels, bar upper 2 levels, that contain directionless networks. Guidance of DBN comprises 2 stages: pre-training and fine-tuning [39]. Initially, well-organized level by level avaricious learning strategy [40] is utilized for initializing deep network in unsupported method, that may stop meagre limited best outcomes to some degree deprived of obligation of huge quantity of branded info. Throughout this process, contrastive divergence [41] is utilized for training RBMs in DBN for estimating estimate gradient of log-likelihood. Before, limits of network plus anticipated output are perfected having humble incline lineage below management.

Hinton projected quick and avaricious algorithm which learns profound, focused belief networks; 1 layer at 1 period. Upper 2 levels shape directionless associative memory [42]. DBN is probabilistic reproductive prototype comprised of numerous levels of stochastic plus latent variables [43]. 2 utmost important characteristics of DBN are applying efficient, layer-by-layer studying process plus implication obligatory for starting percept is together quick plus precise. Figure 3 displays reproductive prototype of DBN having 1 noticeable level and 3 hidden levels' network. Reproductive network produces contenders by input info whereas discriminative network assesses them. This is approved for generating exclusive and truthful facial imageries and additional. Studying probability sharing of input info differentiates DBNs. In DBNs, Restricted Boltzmann Machine (RBM) possess capability of representing info characteristics, therefore this is utilized for building its elementary shape. RBM contains 2 levels; noticeable level and concealed level. Figure 4 displays elementary RBM prototype.

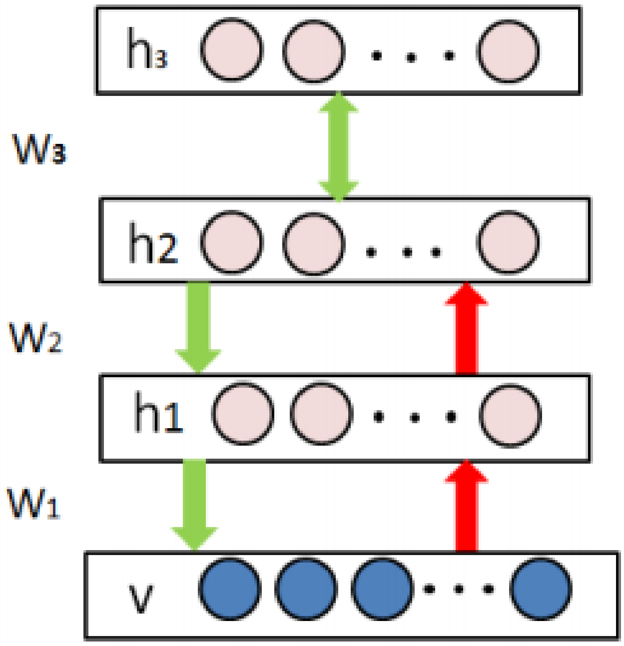


Fig 3: Generative model of DBN having one noticeable and three concealed levels

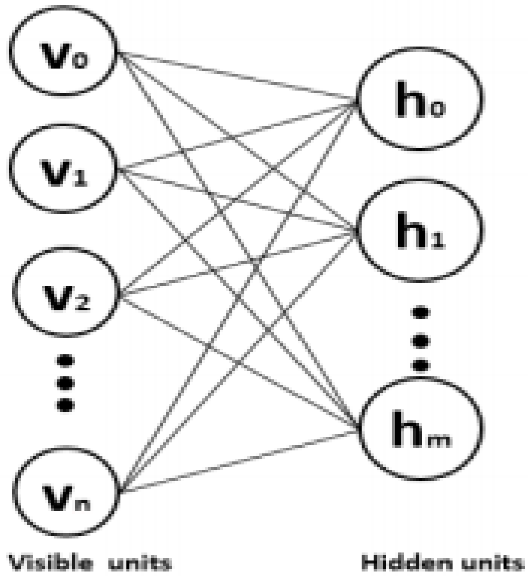


Fig 4: Basic model of Restricted Boltzmann machine

Recurrent neural network (RNN)

RNN is connectionist prototype which imprisons temporal data plus is highly appropriate to consecutive information forecast by random distances. In supplement to coaching deep neural network in solitary feed-forward way, RNNs comprise repeated ends which cover adjoining period phases and distribute similar limits throughout every stages. Classic back transmission via time (BPTT) is utilized for training RNN. Long-short term memory (LSTM), presented from Hochreiter & Schmidhuber, is distinctive method of conventional RNN which is utilized for addressing gradient disappearing and blast complications which are normal in guiding RNNs. Cell status in LSTM is planned and structured from 3 gates: an input gate which lets or obstructs change of cell status from input signal, an output gate which allows or stops cell status for affecting additional neurons, and forget gate which controls cell's self-repeated linking for accumulating or forgetting their preceding status. By uniting these 3 gates, LSTM can mould long-term addictions in order and is extensively used in video-founded expression detection duties.

We presumed that Euclidean metric, normal space amid two spots in frame sequence archives order of events. For example, in state of contentment or pleasure, there are swift eye flashing, crow's feet crinkles in side edge of eyes thrust up cheeks, association by muscle which ranges eye (eye, cheek, chin…). Nevertheless, in situation of dejection, eyes have extremely sluggish eye flashing, sagging higher eyelids, trailing emphasis in eyes, slender dragging low of lip corners. In contrast to situation of repulse, there is distorting of face leftward or rightward, nose is warped up to crumpled nose bridge, tapering eyes, dropped brows. Also this, in dread condition, there are elevated eyebrows, strained reduce eyelids, eyebrows pulled collected, lips pulled parallel. Finally, in rage condition, eyebrows are dragged low collected, eyes are wide-opened and evident, upper eyelids are elevated in gaze, lips are fully unlocked to develop rectangle, and firmly shut by red margins of lips developing thinner, and lips developing thinner. Exact repeated neural network classifier, known as Long Short-Term Memory, was presumed to be utilized so as to obtain benefit of their aptitude to utilize dynamic temporal behaviour of order for arrangement [44], symmetrical descriptor of every edge was nourished to network face recognition phase later.

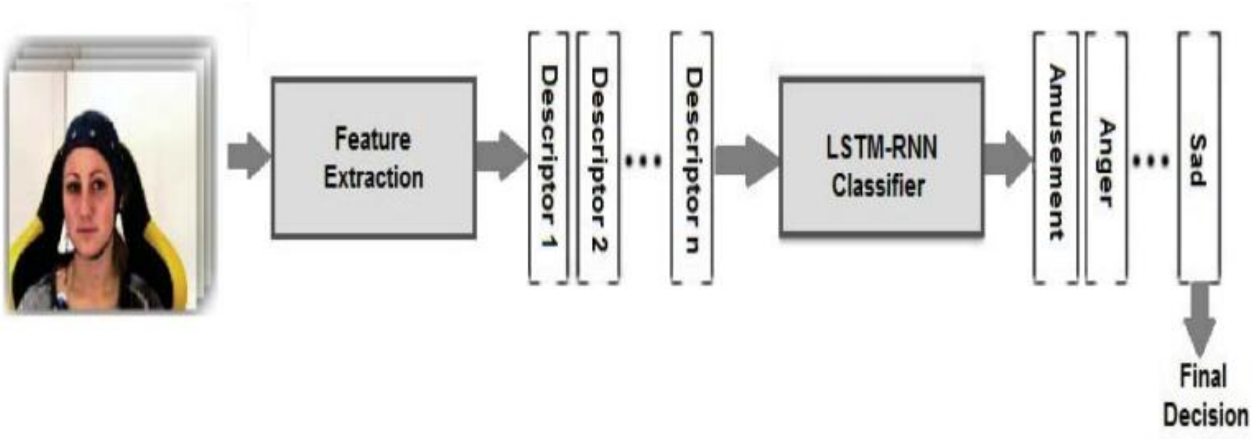
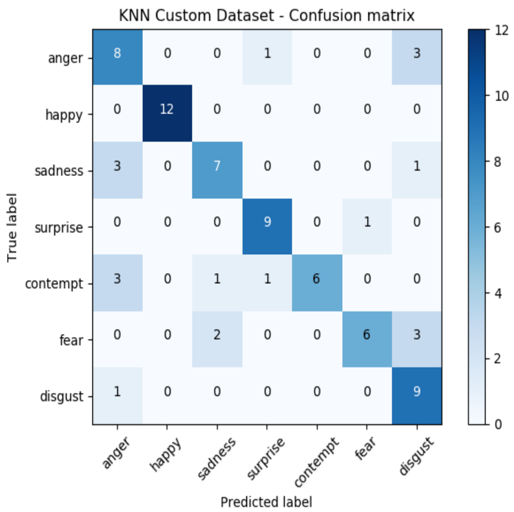
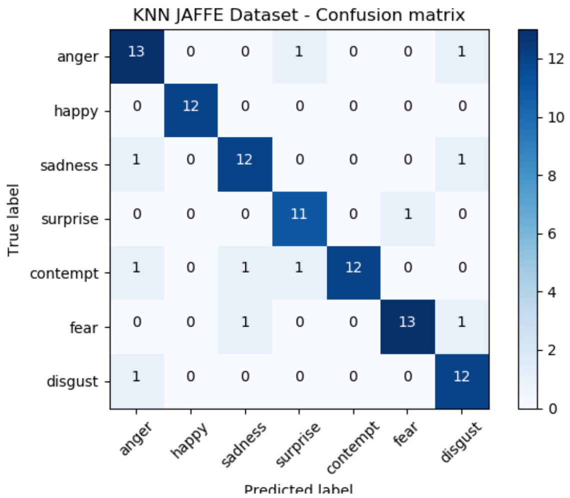


Fig 5: Sorting utilizing LSTM-RNN Classifier

VI. RESULTS

The results of this work are shown below:





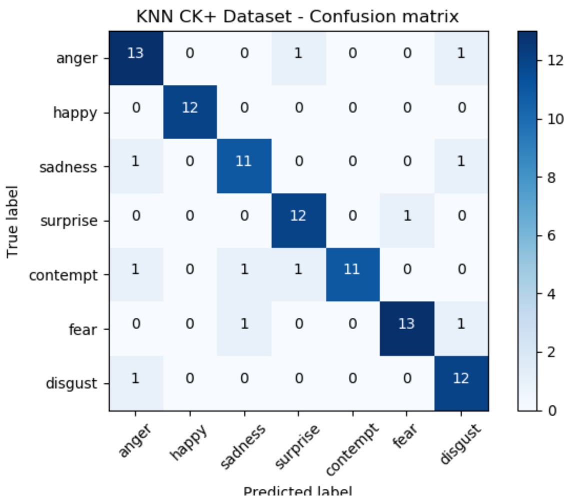
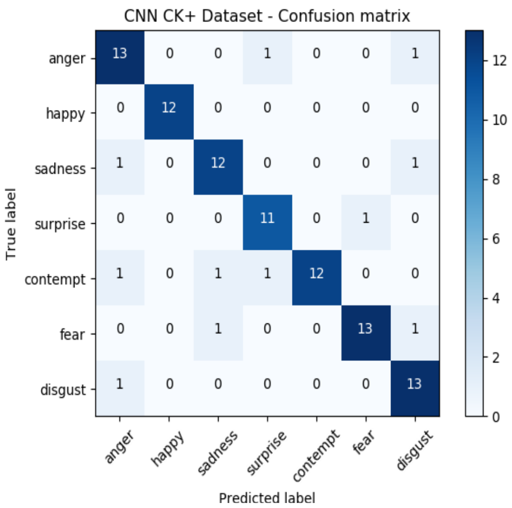
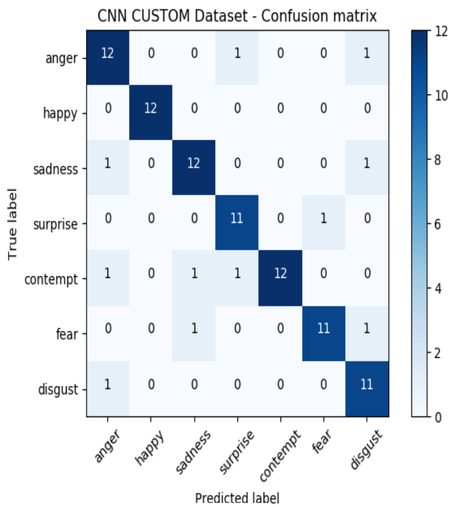


Table 1: Comparison between machine learning algorithms i.e. KNN and SVM

Sl.No	Algorithm	Accuracy
1	KNN	95.5%
2	SVM	96.5%

SVM gives better accuracy in identification and classification of face emotions compared to KNN as shown in above Table 1





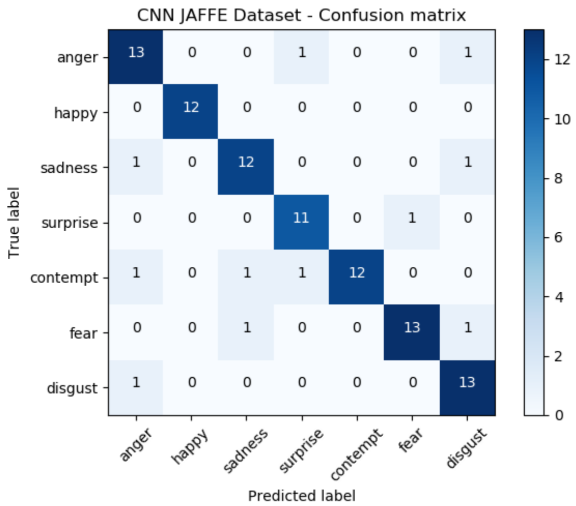
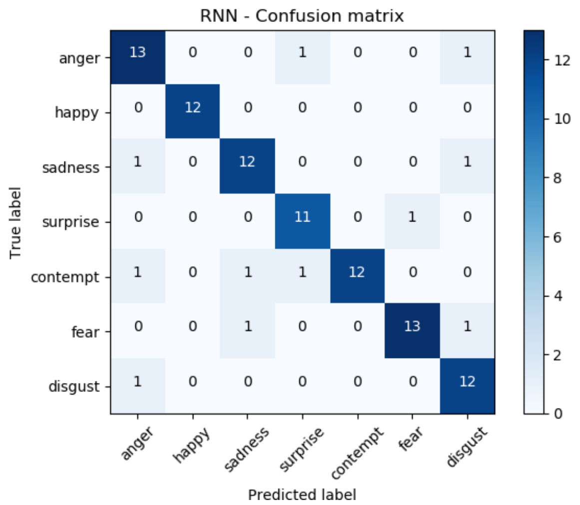
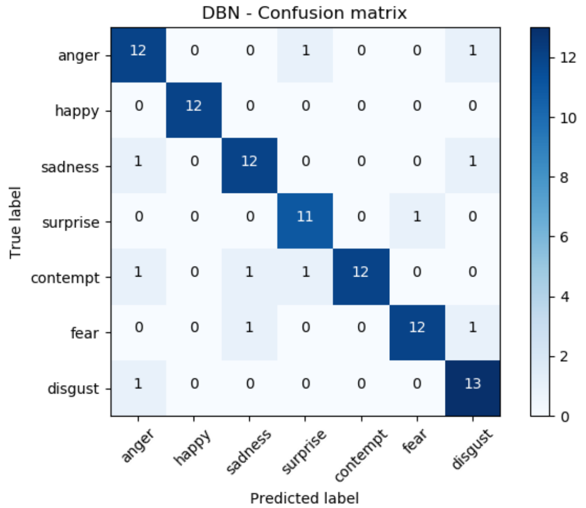


Table 2: Comparison between deep learning algorithms i.e. CNN, RNN and DBN

Sl.No	Algorithm	Accuracy
1	CNN	97%
2	RNN	96%
3	DBN	95%

CNN gives better accuracy in identification and classification of face emotions compared to other algorithms as shown in the above Table 2





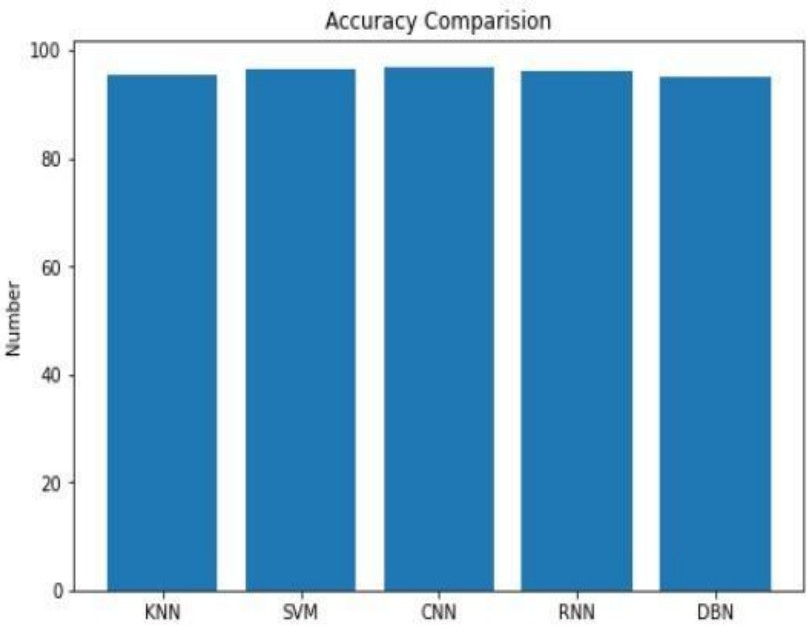


Fig 3: Accuracy Graph

VII. Conclusion

Here, we presented accessible dataset i.e., CK+, JAFFE and FED dataset which extensively are utilized in here. This paper also focused on facial expression detection technique founded on pair of machine learning algorithms and also deep learning algorithms that aid us in precise recognizing and cataloguing of human feeling. Conferring to numerous classifiers SVM classifier provides improved detection precision and this delivers healthier cataloguing. In FER, SVM classifier is extra usable comparison to further classifiers to recognize expression. Neural network founded classifier CNN provides improved precision compared to various deep neural network founded classifiers. CNN classifier is extra usable comparison to various classifiers for improved cataloguing.

References

[1]. A. C. Le Ngo, Y. H. Oh, R. C. W. Phan and J. See, "Eulerian emotion magnification for subtle expression recognition," 2016 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), Shanghai, 2016

[2]. K. M. Rajesh and M. Naveenkumar, "A robust method for face recognition and face emotion detection system using support vector machines," 2016 International Conference on Electrical, Electronics, Communication, Computer and Optimization Techniques (ICEECCOT), Mysuru, 2016

[3]. H. Ebine, Y. Shiga, M. Ikeda and O. Nakamura, "The recognition of facial expressions with automatic detection of the reference face," 2000 Canadian Conference on Electrical and Computer Engineering. Conference Proceedings. Navigating to a New Era (Cat. No.00TH8492), Halifax, NS, 2000, pp. 1091-1099 vol.2.

[4]. Zhao, X., Zhang, S., 2016. A review on facial expression recognition: feature extraction and classification. IETE Tech. Rev. 33, 505–517. https://doi.org/ 10.1080/02564602.2015.1117403.

[5]. Matsumoto, D.; Yoo, S.H.; Fontaine, J. Mapping expressive differences around the world. J. Cross Cult. Psychol. 2008, 39, 55–74. [CrossRef]

[6]. Røysamb, E.; Nes, R.B.; Czajkowski, N.O.; Vassend, O. Genetics, personality and wellbeing. A twin study of traits, facets and life satisfaction. Sci. Rep. 2018, 8, 12298. [CrossRef]

[7]. Songa, G.; Slabbinck, H.; Vermeir, I. How do implicit/explicit attitudes and emotional reactions to sustainable logo relate? A neurophysiological study. Food Qual. Prefer. 2019, 71, 485–496. [CrossRef]

[8]. Keltner, D. Born to Be Good: The Science of a Meaningful Life; WW Norton & Company: New York, NY, USA, 2009.

[9]. Lövheim, H. A new three-dimensional model for emotions and monoamine neurotransmitters. Med. Hypotheses 2012, 78, 341–348. [CrossRef]

[10]. Wundt, W.M. Outlines of Psychology (Classic Reprint); Fb&c Limited: London, UK, 2017; ISBN 9780331795738.

[11]. Suk, M. and B. Prabhakaran, Real-time Mobile Facial Expression Recognition System—A Case Study, in The IEEE Conference on Computer Vision and Pattern Recognition (CVPR) Workshops. 2014, IEEE. p. 132-137

[12]. Ryan, A., et al., Automated Facial Expression Recognition System in 43rd Annual 2009 International Carnahan Conference on Security Technologies. 2009. p. 172-177.

[13]. Golzadeh, D. R. Faria, L. J. Manso, A. Eka´rt and C. D. Buckingham, "Emotion Recognition using Spatiotemporal Features from Facial Expression Landmarks," 2018 International Conference on Intelligent Systems (IS), Funchal—Madeira, Portugal, 2018, pp. 789-794, https://doi.org/10.1109/IS.2018. 8710573

[14]. A. Kartali, M. Roglić, M. Barjaktarović, M. Đurić-Jovičić, M. M. Janković, Real-time Algorithms for Facial Emotion Recognition: A Comparison of Different Approaches, 2018 14th Symposium on Neural Networks and Applications (NEUREL), Belgrade, 2018, pp. 1-4.

[15]. G. Yang, J. S. Y. Ortoneda and J. Saniie, "Emotion Recognition Using Deep Neural Network with Vectorized Facial Features," 2018 IEEE International Conference on Electro/Information

Technology (EIT), Rochester, MI, 2018, pp. 0318-0322.
[16]. Hai-Duong Nguyen, Soonja Yeom, Guee-Sang Lee, Hyung-Jeong Yang, In-Seop Na, Soo-Hyung Kim, Facial Emotion Recognition Using an Ensemble of Multi-Level Convolutional Neural Networks, International Journal of Pattern Recognition and Artificial Intelligence, 33(11), 1940015, 2019 https://doi.org/ 10.1142/S0218001419400159.
[17]. Liliana, Emotion recognition from facial expression using deep convolutional neural network, Journal of Physics: Conference Series,1193, 012004, 2019.
[18]. M. Mohammadpour, H. Khaliliardali, S. M. R. Hashemi and M. M. AlyanNezhadi, “Facial emotion recognition using deep convolutional networks,” 2017 IEEE 4th International Conference on Knowledge Based Engineering and Innovation (KBEI), Tehran, 2017, pp. 0017-0021.
[19]. Ko BC. A Brief Review of Facial Emotion Recognition Based on Visual Information. Sensors (Basel). 2018 Jan 30; 18(2):401. https://doi.org/10.3390/s18020401 PMID: 29385749
[20]. Mehta D., Siddiqui M., Javaid A. Y. (2019). Recognition of Emotion Intensities Using Machine Learning Algorithms: A Comparative Study. Sensors (Basel, Switzerland), 19(8), 1897. https://doi.org/10.3390/ s19081897 PMID: 31010081
[21]. X. Jiang, "A facial expression recognition model based on HMM," Proceedings of 2011 International Conference on Electronic & Mechanical Engineering and Information Technology, Harbin, Heilongjiang, China, 2011
[22]. W. Swinkels, L. Claesen, F. Xiao and H. Shen, "SVM point-based real-time emotion detection," 2017 IEEE Conference on Dependable and Secure Computing, Taipei, 2017.
[23]. J. M. Saragih, S. Lucey and J. F. Cohn, "Real-time avatar animation from a single image," Face and Gesture 2011, Santa Barbara, CA, USA, 2011
[24]. C. Loconsole, C. R. Miranda, G. Augusto, A. Frisoli and V. Orvalho, "Real-time emotion recognition novel method for geometrical facial features extraction," 2014 International Conference on Computer Vision Theory and Applications (VISAPP), Lisbon, Portugal, 2014
[25]. Neerja and E. Walia, "Face Recognition Using Improved Fast PCA Algorithm," 2008 Congress on Image and Signal Processing, Sanya, Hainan, 2008
[26]. K. M. Rajesh and M. Naveenkumar, "A robust method for face recognition and face emotion detection system using support vector machines," 2016 International Conference on Electrical, Electronics, Communication, Computer and Optimization Techniques (ICEECCOT), Mysuru, 2016 50
[27]. B. Fasel, “Robust face analysis using convolutional neural networks,” in Pattern Recognition, 2002. Proceedings. 16th International Conference on, vol. 2. IEEE, 2002, pp. 40–43.
[28]. ——, “Head-pose invariant facial expression recognition using convo lutional neural networks,” in Proceedings of the 4th IEEE International Conference on Multimodal Interfaces. IEEE Computer Society, 2002. p. 529.
[29]. M. Matsugu, K. Mori, Y. Mitari, and Y. Kaneda, “Subject independent facial expression recognition with robust face detection using a convolu tional neural network,” Neural Networks, vol. 16, no. 5-6, pp. 555–559, 2003.
[30]. B. Sun, L. Li, G. Zhou, X. Wu, J. He, L. Yu, D. Li, and Q. Wei, “Combining multimodal features within a fusion network for emotion recognition in the wild,” in Proceedings of the 2015 ACM on Interna tional Conference on Multimodal Interaction. ACM, 2015, pp. 497–502.
[31]. B. Sun, L. Li, G. Zhou, and J. He, “Facial expression recognition in the wild based on multimodal texture features,” Journal of Electronic Imaging, vol. 25, no. 6, p. 061407, 2016.
[32]. R. Girshick, J. Donahue, T. Darrell, and J. Malik, “Rich feature hierarchies for accurate object detection and semantic segmentation,” in Proceedings of the IEEE conference on computer vision and pattern recognition, 2014, pp. 580–587.
[33]. J. Li, D. Zhang, J. Zhang, J. Zhang, T. Li, Y. Xia, Q. Yan, and L. Xun, “Facial expression recognition with faster r-cnn,” Procedia Computer Science, vol. 107, pp. 135–140, 2017.
[34]. S. Ren, K. He, R. Girshick, and J. Sun, “Faster r-cnn: Towards real-time object detection with region

proposal networks," in Advances in neural information processing systems, 2015, pp. 91–99.
[35]. S. Ji, W. Xu, M. Yang, and K. Yu, "3d convolutional neural networks for human action recognition," IEEE transactions on pattern analysis and machine intelligence, vol. 35, no. 1, pp. 221–231, 2013.
[36]. D. Tran, L. Bourdev, R. Fergus, L. Torresani, and M. Paluri, "Learning spatiotemporal features with 3d convolutional networks," in Computer Vision (ICCV), 2015 IEEE International Conference on. IEEE, 2015, pp. 4489–4497.
[37]. G. E. Hinton, S. Osindero, and Y.-W. Teh, "A fast learning algorithm for deep belief nets," Neural computation, vol. 18, no. 7, pp. 1527–1554, 2006.
[38]. G. E. Hinton and T. J. Sejnowski, "Learning and releaming in boltz mann machines," Parallel distributed processing: Explorations in the microstructure of cognition, vol. 1, no. 282-317, p. 2, 1986.
[39]. G. E. Hinton, "A practical guide to training restricted Boltzmann machines," in Neural networks: Tricks of the trade. Springer, 2012, pp. 599–619.
[40]. Y. Bengio, P. Lamblin, D. Popovici, and H. Larochelle, "Greedy layer wise training of deep networks," Advances in neural information processing systems, 2007, pp. 153–160.
[41]. G. E. Hinton, "Training products of experts by minimizing contrastive divergence," Neural computation, vol. 14, no. 8, pp. 1771–1800, 2002.
[42]. G. E. Hinton, S. Osindero, and Y. W. Teh, "A fast learning algorithm for deep belief nets", Neural Computation, Vol. 18, No. 7, pp. 1527-1554, 2006.
[43]. G. E. Hinton, "Deep belief networks" Scholarpedia, Vol. 4, No. 5, p. 5947, 2009
[44]. F. A. Gers, N. N. Schraudolph, and J. Schmidhuber, "Learning precise timing with lstm recurrent networks," Journal of machine learning research, vol. 3, no. Aug, pp. 115–143, 2002.