Human Computer Interaction using CNN and Fuzzy Inference System (FIS)

Swati A. Atone[1]. Dr. A.S.Bhalchandra[2], Methil Muley[3]

[1,2,3]*Department of Electronics and Telecommunication Government Engineering College, Aurangabad, Maharashtra, swati.atone@gmail.com*

In the evolutionary world of technology, humanoid robot has gained increased attention in different areas such as security systems, medical systems (psychiatric clinic), entertainment, military applications, home appliances, education etc. Companion robot is one of the major challenges in humanoid robots. Companion robot can play role of companion for children or senior citizens rather than machine assistance. Such robots are expected to exhibit capabilities primarily includes propensity to recognize human emotions with a skill to express its own expressions as sensible as human. This paper aims to present system to recognize human facial expression and respond by generating an expressive robot face. Convolutional Neural Network (CNN) approach is used to recognize human facial expressions. This is accomplished by detecting the occurrence of facial Action Units (AUs) as a subpart of Facial Action Coding System (FACS) which represents human expressions. Real-World Affective Face Database (RAF-DB) has been employed for training the algorithm. Fuzzy inference system is adopted to decide responsive expressions on robotic face simulator.

Keywords: Human Computer Interaction (HCI), CNN, Facial expression detection, facial expression generation, fuzzy-based response emotion intensity generator (FBREIG).

1. Introduction

In last few decades, interactive robots are getting attention a lot. Interactive robots which has capability to communicate with people in human centric terms, which can mimic the human expressions, recognize human facial expression (expressive face).Now a days, researchers in HCI have focussed on humanoid social Intelligent robots. It is a field dedicated to understanding, designing and evaluating robotic system. HRI may take multimodal forms like computer vision to process facial expression, gesture and speech recognition for language understanding and dialogue. Facial expressions are prominent characteristic of human conversation and emotions. Human Facial expression analysis has pioneered with C. Darwin’s (1872) principles . According to him, expression due to certain emotion is not culture-specific. Facial expressions are authoritative means to recognize expression recognition. Next part of HRI is to use these recognised expressions for interaction from robot. In last decade, many researchers have taken efforts to use recognised expressions. Efforts are successful for single expression detection; neutral happy, sad, surprise, angry, fear, disgust. But challenge is; single expression does not indicate a single emotion, usually it consists of mixed emotions. Interactive robots have capacity to recognize intensities of all emotions in facial expression and also characteristics, positive and negative. It responds accordingly using these recognised expressions.

Concept is extended for senior citizens, who choose to live alone. They are physically fit but need emotional support to combat social isolation. Nothing can replace human touch but with a prediction that there could be shortage of people who will be available as caregiver for growing senior population. Companion robots or interactive robots are proposed as solution to the situation. For interaction, it is necessary to recognize and understand human facial expression and interact accordingly. Robot acts as companion not as machine, This paper presents human expression recognition and generation of expression on simulated robotic face. Human expressions are recognized and accordingly response expressions are created on simulated robotic face. CNN algorithm is used to detect human facial expression and fuzzy based system is used for emotion intensity generation on Robotic face simulator. Paper is organised as, section 2 introduced some previously done research work in area of human emotion analysis. Section 3 presents proposed system wherein 3.1, describes proposed convolutional neural network to recognize human facial expressions and 3.2 describes fuzzy- based emotion intensity generator (FBREIG) to decide response intensities for robot simulator. Section 4 presents results and discussion with some simulations of robotic face. Lastly conclusion is presented.

2. LITERATURE SURVEY

Automatic facial expression recognition is the framework to detect and classify every input facial expression into respective classes of basic facial expressions; Conventional machine learning methods used for classification. These methods are compared based on classification accuracy and presented in Table 1.

Table 1 Comparison of accuracy achieved by various machine learning approaches

Author	Algorithm	Dataset	Accuracy	Disadvantages
Li et al., (2015)	ADB along with Dynamic Bayesian network (DBN)	Denver Intensity of Spontaneous Facial Action(DISF A)	72.77%(DBN) 69.31%(ADB)	Less Accuracy
Kakadiaris et al. (2007)	Haar feature selection Techniques Geometric – Based approach	Face Recognition Grand Challenge v2 (FRGC V2)	97%	Does not suit good in case of real-time application

Huang et al. (2011)	Stochastic neighbor embedding, Gabor wavelet, SVM	Japanese Female Facial Expression (JAFFE)	58.70%	insignificant accuracy, Reduced database
Tang et al. (2008)	Feature Distribution entropy, Euclidian distance between 83 3D facial feature points, ADB	Binghamton University 3D Facial Expression (BU-3DFE)	95.1%	Not suitable for real-time application
Pal et al. (2016)	k- means clustering including back propagation	Cohn-Kanade (CK)	98%	Do not perform well in real-time application
Zhang et al. (2013)	Local Binary Pattern (LBP)	Video data of 10 volunteers	89.60%	Less accuracy, Reduced database
Soyel et al. (2007)	Distance Vector , Neural network	BU-3DFE	91.30%	insignificant accuracy
Savran et al. (2013)	SVM along with Gaussian Curvature, Shape index, Gabor wavelet	Bosphorus (2902 images)	63.10%	25 Action units are considered only from steady images
Mahoor et al. (2009)	SVM along with Spectral Regression	Video stream of 18 minutes	92%	Insufficient data
Ahmed et al. (2013)	SVM includ ing Gradient-based ternary texture patterns	Ck	97.10%	-

Pantic et al. (2000)	Herculer engine along with Facial data generator	Data capture using webcam	89.60%	Low accuracy lagging
Song et al. (2011)	Active Appearance Model (AAM), Lucas-Kanade, Back Propagation Neural Network (BPNN)	BU-3DFE	83.80%	Poor Accuracy
Chang et al. (2005)	Iterative closest point(ICP), Principal component analysis(PCA)	355 images	92%	Poor Accuracy

With the enhancement of research concerning psychology and neuroscience, Carl-Herman developed a facial action coding system which was further improved by Ekman & Friesen (1978). Emotion recognition is also identified as 'Affection computing' which plays a vital role in human-machine interaction. Affective computing which detects the emotions along with intensity is introduced by Picard. Mase et al. (1991) obtained motion direction of facial muscle using Optical Flow and classified into four basic facial expressions.

These machine learning algorithms perform well for facial expressions under controlled circumstances but not in the wild. It also performs well in case of the smaller datasets of around few hundred features but underperforms with large data, almost in terms of millions of images with maximum resolution. To overcome these limitations; Deep Learning approach is opted which extracts the features statistically and classify them.

Hinton et al. (2006) pioneered the deep learning era. Prediction accuracy along with classification accuracy is being boosted by extracting distinct measures of abstract description from actual data. Hinton et al. (2006) employed a deep convolutional neural network ImageNet giving best results. Tarik A. Rashid (2016) employed three different neural networks to recognize facial expression; Multilayer perceptron, CNN and Decision tree. All of them were tested on two different datasets, Bosphorus and JAFFE. Kukla et al. (2015) opted cascade neural network approach to recognize facial expression. A Cascaded network with single perceptron with three layer network and middle layer with five hidden neurons is utilised. It could recognize happiness and surprise correctly, but not fear emotion.

Networks resulted in 76% accuracy with Karolinska Directed Emotional Faces (KDEF dataset).

Existing frameworks generally have not discussed the mixture of emotions; intensities more than one emotion in a single expression. Many frameworks have provided the result with a single expression that possesses maximum intensity. However many times single expression may possess a combination of more than one emotion. This paper presents expression recognition including respective intensities involved in that particular expression and in continuation with recognition of emotion, efforts are made to develop a robot which can respond accordingly

Developing a robot which can respond to detected emotions had mechanical orientation. Hyung et al. (2019) proposed a Genetic Algorithm technique for autonomously producing face expressions. 16 control points of face were selected with one motor for each control point. Deformations of all 16 motors combined together to generate facial expression which made it highly effective for surprised and sad faces. Habib et al. (2014) adopted combination of neural network and genetic algorithm to generate facial expression on robotic face. Control signal generator in addition with genetic algorithm is adopted to replicate facial features as identical as human face. Kim et al. (2010) developed an emotion model based on theoretical approach called cognitive appraisal theory. It used probabilistic description predicted by computational algorithms. System used observable Markov process (POMDP). It is a revolutionary technique that combines the Markov model with observability. This technique is to enable robots behave more naturally and effectively, like humans.

Above techniques have successfully developed variety of interactive robots. But such robots are unable to create emotions in continuous manner. There is less discussion on robotic reaction which possesses mixture of emotion. To deal with response mixed emotion intensity generation, fuzzy logic based approach has been considered known as fuzzy-nets. In crisp logic, classes of expressions possess sharp boundaries, only broad categorization is possible. But, emotion doesn't always belongs to broad category like; only happy. There might be ambiguous boundaries in particular expression. Instead of only happy, degree of expression may vary to more happy or most happy. Such situations need flexibility of emotion instead of rapid changes. It is provided by fuzzy logic. Considering the existing scenario, work focuses on expression recognition including respective intensities involved in that particular expression using CNN and generates an expressive robotic face as a response to the recognized emotion using fuzzy logic

3. PROPOSED WORK

Figure 1, shows block diagram of proposed system. Proposed work encompasses three major blocks; facial emotion recognition, Fuzzy based response emotion, intensity generator and robotic face simulator. Algorithm pioneer with Facial emotion recognition wherein Proposed CNN algorithm has been trained utilizing two different databases; Extended Cohn-Kanade dataset (CK+) and Real-World Affective Face Database (RAF-DB) with three different optimizers; Stochastic Gradient descent momentum (SGDM), Root mean square propagation

(RMSprop) and Adam optimizer. Trained model has used to recognize facial expressions in real time.

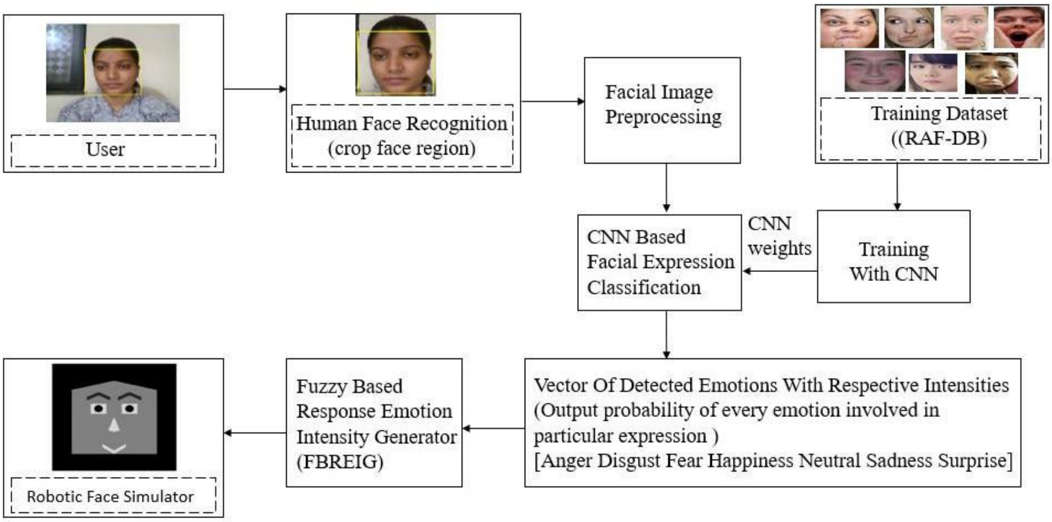


Fig. 1 Block Diagram of Proposed System

In real time, initially Camera captures user's face. Viola jones algorithm is used to detect human face in terms of bounding box. Detected face undergoes image preprocessing. Moreover it goes through proposed CNN which recognize user's expressions along with intensities of each emotion involved in that particular expression. During conversation, Degree of expressions may vary. There might be ambiguous boundaries in particular class of expression. For example, Instead of 'happy', person can be normal happy, more happy or most happy. Instead of immediate transition, flexibility of emotion is needed for engaging conversation and fuzzy logic provides the same. Fuzzy acquire input intensities of expression and use membership function to generate corresponding degree of expression. FBREIG is designed using fuzzy inference system to generate response emotion intensities autonomously by Atone & Bhalchandra (2022). A survey based on rule table is conducted to predict expression on robotic face. Fuzzy rules are created by conducting and analysing the survey. Intensities recognized using CNN further used as input to FBREIG to decide response emotion intensities. That Intensities further pass to robotic face simulator which generates response face to interact with the user. A robotic facial simulator is employed to produce a responsive behaviour for the desired human-robot interaction.

Figure 2 shows the multi-layer perceptron model with 11 layers. The increment in layers improves the performance of model. The increment after total 11 layers is stopped as performance of model after 11 layers decreases due to over fitting problem.

4. PROPOSED CONVOLUTIONAL NEURAL NETWORK

4.1 Dataset Preprocessing:

The architecture of proposed CNN is represented in Figure 2. As depicted in Figure 2, the net contains five convolutional layers, four max-pooling layers and one fully-connected layer. The output of the fully-connected layer is supplied to the 7-way softmax, which produces a distribution over 7 class labels which are anger, disgust, fear, happiness, neutral, sadness, and surprise. Max pooling layer follows each convolutional layer. The ReLU (Rectified Linear Unit) non-linearity is applied to the output of every convolutional layer. First convolutional layer filters 100×100 input image with 12 kernels of size 3×3 with a stride of 1 pixel. Output of the first max-pooling layer is considered as input to the second convolutional layer and filters it with 12 kernels of size 3×3×24. The third, fourth and fifth convolutional layers are connected to one another with an intervening max-pooling layer. Third, fourth and fifth convolutional layer has 48 kernels of size 3×3×48 connected to each convolutional layer. The fully-connected layer has 2352 neurons. To prevent overfitting, the dropout technique is applied to the fully-connected layers. Softmax function plays vital role here. It provides decimal probabilities to each class. It turns each input between 0 and 1, regardless of whether the input is zero, positive, negative, or greater than 1. It assigns lower probability to the input which is negative or little in value. In contrast, it assigns higher probability to the large input value. Sum of decimal probabilities of all the class will be equal to 1. Formula of softmax function is as follow

$\sigma(\vec{Z}) = \; e^{Z}i$

Where σ represents softmax function, $\vec{Z}$ denotes input vector, e^{Z_i} *and* e^{Z_j}denote standard exponential function of the input and output vector respectively, K specifies number of classes considered in multiclass classifier. As shown in figure 2, output of 7 way softmax function is probability of each emotion involved in that particular expression.

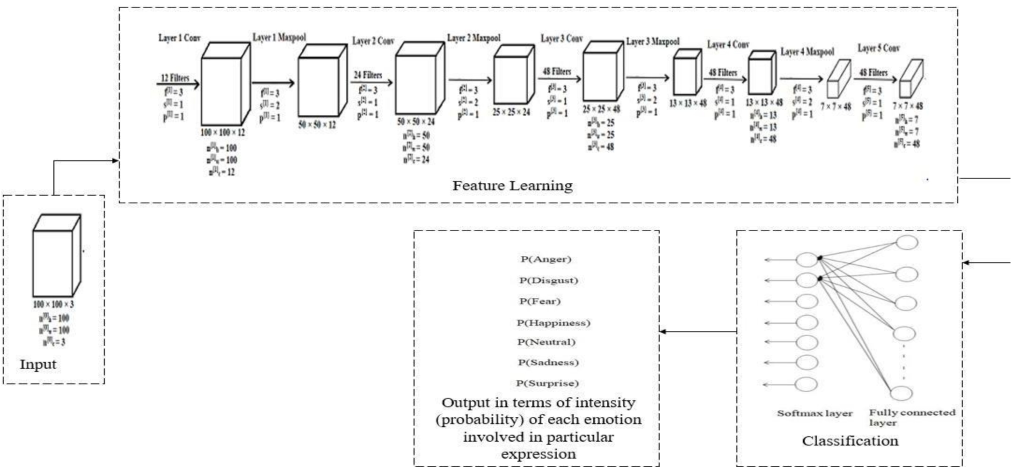


Figure 2. Proposed CNN architecture

The Network structure of proposed CNN structure is shown in Table 2.

Table 2 Network structure of proposed CNN architecture 4.1.Fuzzy-based response emotion intensity generator

Characteristics	Layer 1		Layer 2		Layer 3		Layer 4		Layer 5
	Conv	Maxpool	Conv	Maxpool	Conv	Maxpool	Conv	Maxpool	Conv
Filter Size	3 × 3	3 × 3	3 × 3	3 × 3	3 × 3	3 × 3	3 × 3	3 × 3	3 × 3
Stride	1	2	1	2	1	2	1	2	1
Number of Filters	12	12	24	24	48	48	48	48	48
Output Size	100 × 100	50 × 50	50 × 50	25 × 25	25 × 25	13 × 13	13 × 13	7 × 7	7 × 7

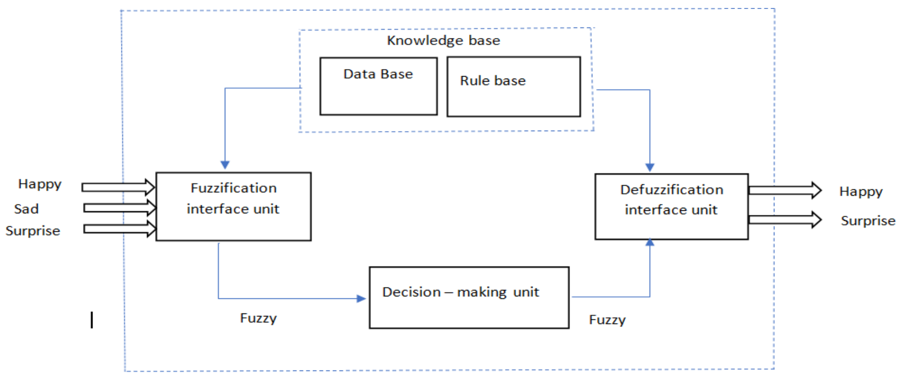


Fig. 3. Structure of fuzzy based emotion intensity generation system

System has input of mixed emotions in form of respective intensity. These intensities are used to generate expression on robotic face simulator and Fuzzy Inference system is adopted for this task. Mamdani fuzzy interface system is used to decide the response emotion intensity of robotic face simulator. Figure 3, depicts the framework of the FBREIG (Fuzzy-Based Emotion Intensity Generator). It encompasses four main blocks named as fuzzification interface unit, knowledge base, decision-making unit, and defuzzification interface unit. In the framework, the intensities of individual's sentimental state are regarded as inputs of FBREIG. In fuzzy inference system, no. of possible rules are depends upon number of input variable and number of linguistic variables. 3^7 = 2187 no. of rule will be possible, if 7 input intensities are taken into account. System will become more complex. As an illustration, three input intensities from the aforementioned vector; happiness, sadness, and surprise are taken into account and provided to the FBREIG system. Being an optimistic robot, it produces responses that are intensely positive and surprising. Therefore, output intensities such as happiness and surprise are taken into consideration. System pioneer with determination of linguistic variable for each input and output variable followed by assigning the descriptor to each linguistic variable. However, fuzzy logic works on fuzzy values, so crisp values are converted into fuzzy values. Fuzzification unit converts crisp input intensities into fuzzy sets. Knowledge base is composed of two blocks, database followed by rule base. Database involves description of membership functions (MFs) related to fuzzy sets linked with linguistic variables. System opted for a triangular membership function. Input variables along with their linguistic variables determine the number of maximum possible rules to be generated. Fuzzy rules are designed by conducting and analyzing the survey. Prototype fuzzy rules consists of response of basic output behaviours corresponding to primary sentimental states. FBREIG can generalize these prototype rules to all situations that can happen to the robot by evaluating the proper fuzzy rules. Input variables along with their linguistic variables determine the number of maximum possible rules to be generated. Systems consist of three input variables along with three linguistic regions contributes 33 =27 possible number of rules. Decision-Making Unit provides the degree along with every fuzzy rule which affects the result. Instructions regarding every rule are combined together after every rule is examined independently. Final result is achieved by abstracting the responses identified by every particular rule. Resultant fuzzy values are converted into crisp by adopting defuzzification block. Illustration to evaluate FBREIG is as depicted in Table 4.

Table 3 Input, output variables with their linguistic variable, descriptor, and ranges

Input				Output			
Emotions	Linguistic Variables	Descriptor	Range	Emotions	Linguistic Variables	Descriptor	Range
Happiness	Happy	HA	[0 30 50]	Happiness	Happy	HA	[0 30 50]
	More happy	REHA	[30 50 70]		More happy	REHA	[30 50 70]
	More happy	STHA	[50 75 100]		Most happy	STHA	[50 75 100]
Sadness	Sad	SA	[0 30 50]				
	More sad	RESA	[30 50 70]	surprise	Surprise	SU	[0 30 50]
	Most sad	STSA	[50		More	RESU	[30

			75 100]	Surprise		50 70]
surprise	Surprise	SU	[0 30 50]	most Surprise	STSU	[50 75 100]
	More	RESU	[30 50 70]			

Table 4: Input intensities for fuzzy interface model

happy (%)	Surprise (%)	Sad (%)
55	30	15

As depicted in Table 3, membership degree for Happy contributes in two regions as shown in Figure 4, which are more happy (REHA) and most happy (STHA), Therefore, membership degree of the happy intensity in each of the two regions is calculated as follows,

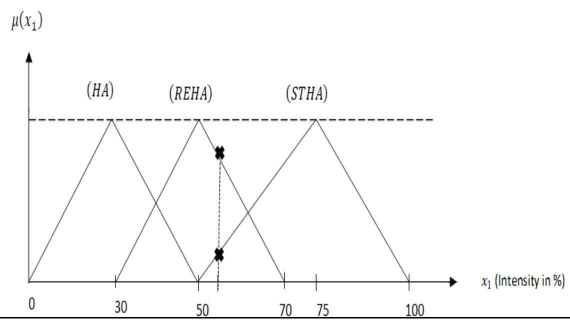


Figure 4. Happy membership degree determination

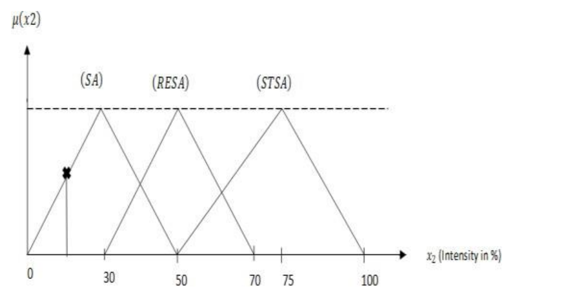


Figure 5. Sad membership degree determination

As the surprise lies on the center of the normal surprise region as shown in Table 3, membership degree is 1.

$(x_3) = 1$ ….(4

Above 4 equations which are (1), (2), (3), (4) leads to 2 rules, Which corresponds to rule no. 10 and rule 19. As shown in fig 9 c)

$\mu(x_1)$and$\mu_{NSA}(x_2)$ and $\mu_{NSU}(x_3)$ => Rule no. 10

$\mu(x_1)$and$\mu_{NSA}(x_2)$ and $\mu_{NSU}(x_3)$ => Rule no. 19

Therefore two fuzzy rules are fired, which are rule no. 10 and Rule no. 19. As shown in Figure 9 c), antecedent part of each of the above rules is connected by "and" operator, min operator is being used in order to evaluate the strength of each rule.

R10 = min ($_{A}(x_1)\mu_{SA}(x_2)\mu_{SU}(x_3)$)

$= \min (\frac{3}{4}, 1, \frac{1}{2}) = \frac{1}{2}$

R19 =min ($_{A}(x_1)\mu_{NSA}(x_2)\mu_{SU}(x_3)$)

$= \min (\frac{1}{5}, 1, \frac{1}{2}) = \frac{1}{5}$

To find value of output membership function, Mean of max (Mom) Defuzzification technique is used.

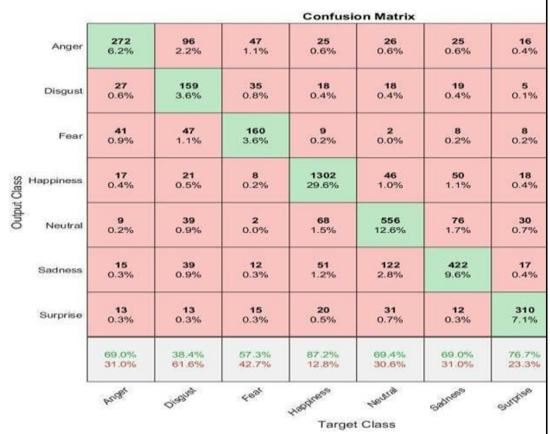
Maximum Strength = max {R10, R19} = max (12 , 15)=12 corresponds to rule no. 10. After solving the equations, output provides 50% happy intensity and 27.5% surprise intensity.

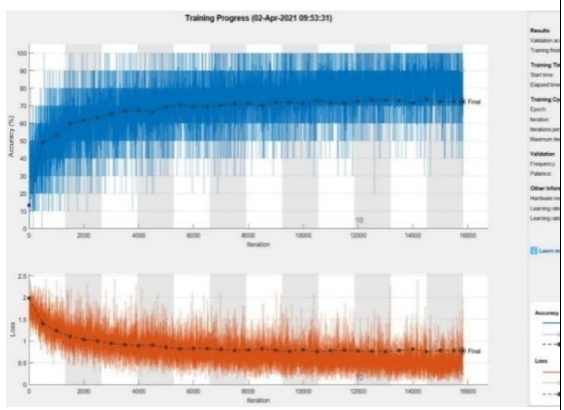
5. Results and Discussion

A system to recognize facial emotion expression and respond the same by generating expression with a robotic face simulator is developed using software MATLAB/SIMULINK '2020a' version. Algorithm has been trained using two different databases ;CK+ database introduced by Lucey et al. (2010) and RAF- DB introduced by Shan et al. (2017) with three different optimizers. Accuracy of each different optimizer for both the dataset is as shown in Table 5. CK+ dataset with the RMSProp optimizer has comparatively better accuracy; 99.18% and required time; 1 minute 8 seconds. RAF dataset achieved 72.34% accuracy with Adam optimizer and required time is 46 minutes 53 seconds. CK+ dataset has 736 samples, which are very few to train CNN, leading to less accuracy with real- time testing on the CK+ dataset. RAF dataset has better accuracy in real-time, whereas RAF dataset has better performance with Adam optimizer as compared to SGDM optimizer. Accuracy with CNN proved to be better as compared to other two classifiers. Shan Ke et al. (2017) experienced the same performance by CNN. They suggested two comparative algorithms for recognition of facial expressions; KNN and CNN with two databases, JAFFE and CK+ giving 76.74% and 83.80% accuracy respectively.

Considering above result, Adam optimizer is best fit for the learning data. Adam parameters utilized the adaptive learning method which reduces the computational power and speeds up learning during the training. From Figure 6, it can be observed that accuracy for the expression 'happiness' is more. Target class and output class in the confusion matrix correspond to the predicted class and true class respectively. The confusion matrix is created from the 75:25 split. 75% samples from each expression are utilized for training purpose whereas, 25% samples are utilized for validation purpose. Figure 7 represents the training progress of RAF dataset using Adam optimizer.

Table 5: Accuracy of two different datasets with three different optimizers

<table>
<tr>
<td>


</td>
<td>


</td>
</tr>
<tr>
<td>Figure 6. Confusion matrix for the proposed model using RAF dataset and Adam optimizer</td>
<td>Figure 7. Training progress of RAF dataset using Adam optimizer</td>
</tr>
</table>

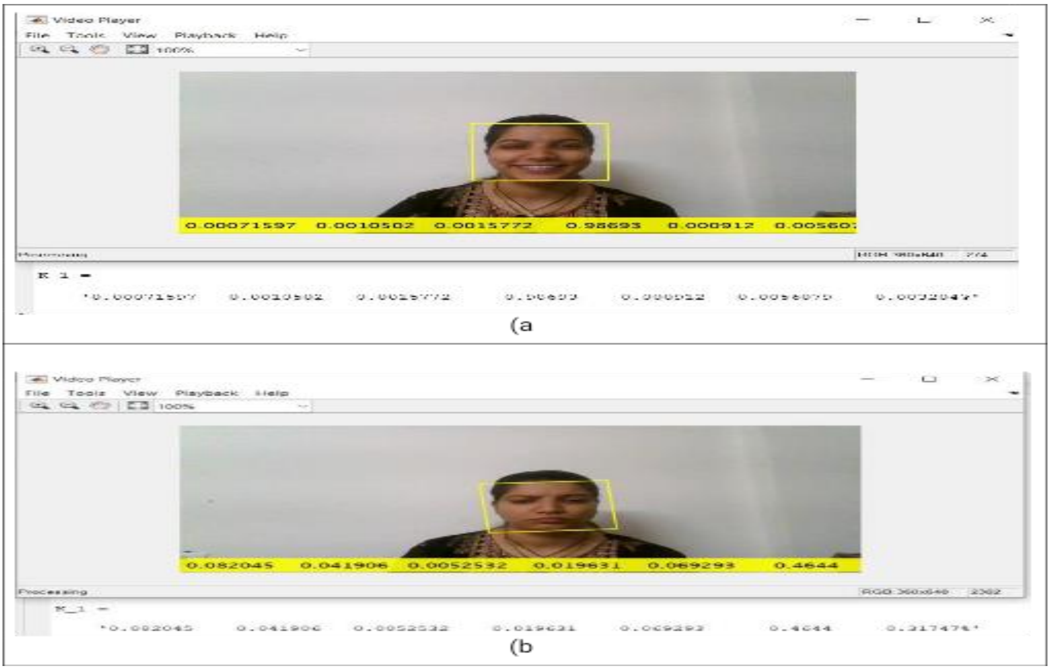


Figure 8. Real time testing model to detect facial expressions

As shown in Figure 8, intensity of various emotions in a single expression is represented in form of probability for an image. Figure 8, represents the % probability vector of emotions;

anger, disgust, fear, happiness, neutral, sadness, and surprise in sequence. In Figure 8 a), the emotion vector has more probability of happiness as the person in the image is extremely happy. Figure 8 b) shows that the emotion vector has intensity value 0.46 for sad, 0.31 for surprise and all other are subtle. This is because the person is not extremely sad or happy, rather has many emotions in a single expression. Based on facial emotion recognition, robotic simulator is developed. Basically, two types of responses can be generated; optimistic and pessimistic. As an example, expression generation for three inputs and two outputs is explained below.

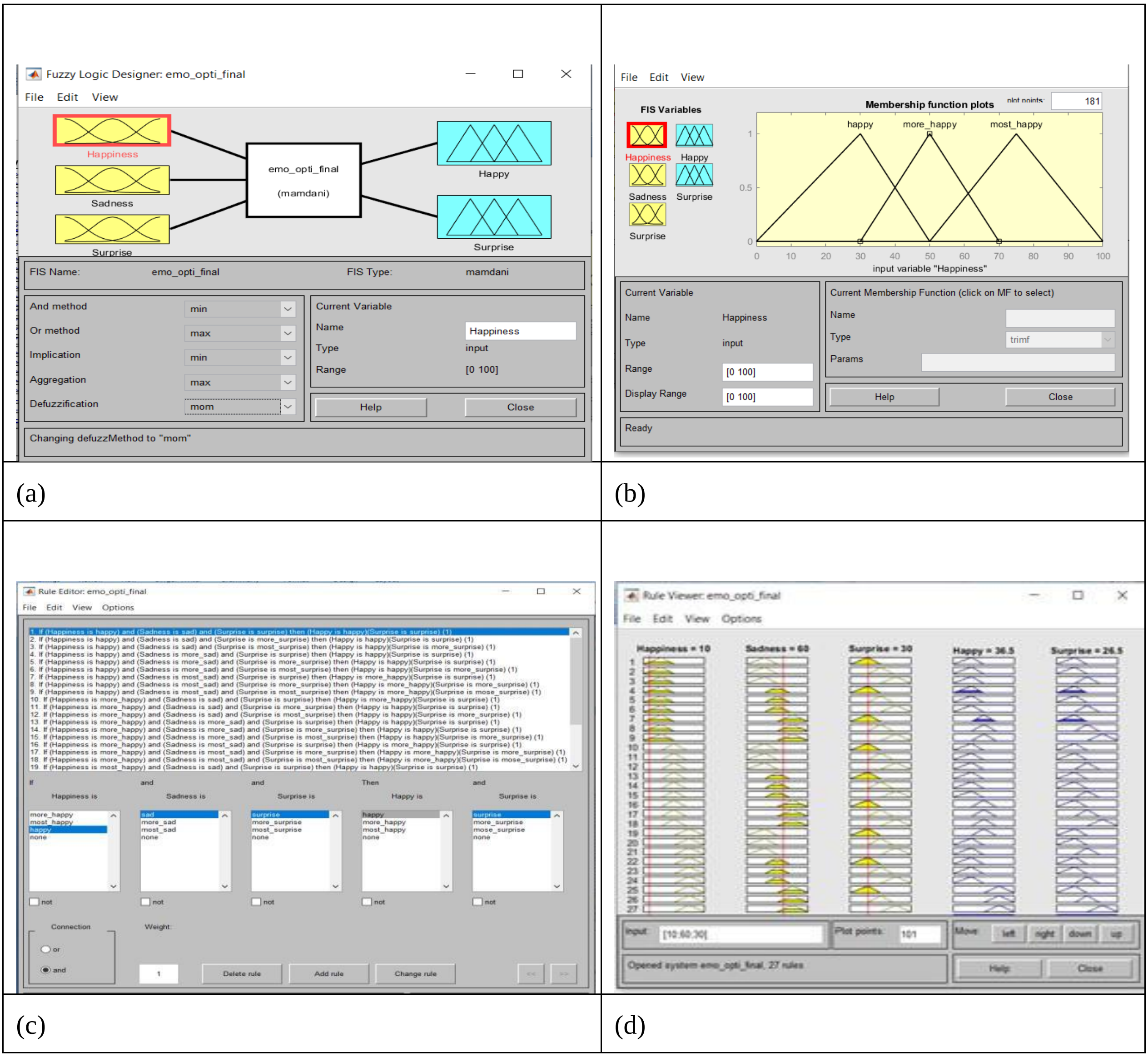
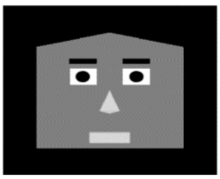
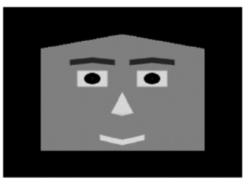
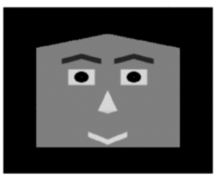
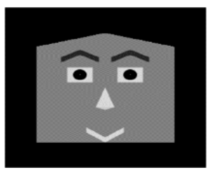
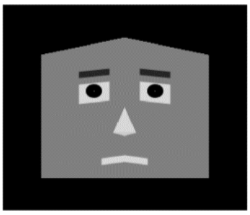
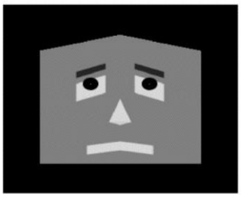
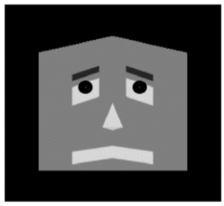
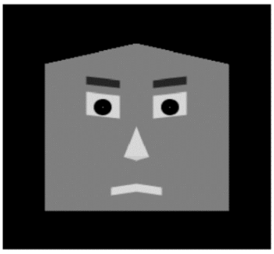
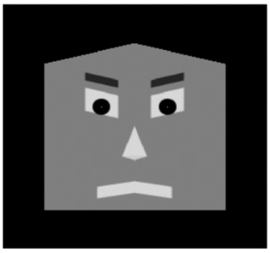
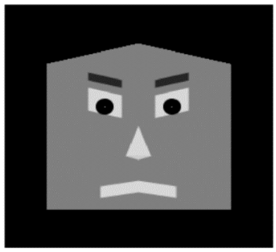


Figure 9. a) Mapping of input output variables b) parameters selected for input c) designed 27 rules d) Illustration of three input intensities with their resultant output

FBEIG is developed for expression generation on robotic face simulator. As shown in fig. 9 a, the system has three input intensities; happiness, sadness, and surprise, and two output intensities happiness and surprise. Fig 9b shows parameters selected for input. Fig 9c, shows 27 rules for FBREIG system. Fig.9d) shows an illustration of three input intensities and their resultant output intensities. Based on these rules, various expressions are generated using robotic simulator as shown in figure 10. Basic face is neutral face as shown in Figure 10 a Three different responses, happy, sad, and anger are shown with different ranges in figure 10b, 10c, and 10d. These three responses are representing three different single emotions but in real-time situation, mixed emotion response is expected. The primary action is to decide whether an optimistic response or pessimistic response is to be generated. Decision is based on detected emotions of human being. If detected emotion is happiness and its intensity is greater than any other emotion then an optimistic response is generated and if detected emotion is sadness or anger and its intensity is greater than any other emotion then pessimistic response is generated. Once response type is fixed, based on intensity value of detected emotions, FBREIG decides response emotion intensities for robotic face. Expression generation is done using variation in eyebrow positions and mouth area. Six emotions are detected using CNN. Each emotion has three ranges. For example, three ranges for happy response are; happy, more happy and most happy. Similarly for sad, surprise, anger, disgust, and fear responses. With six emotions and three gradations, 37=2187 combinations are possible. Rule table representation for these many combinations is difficult here so as an exemplification, a system of three emotions with three ranges i.e. 33= 27 combinations are presented here. As an example, the optimistic response is depicted. For optimistic type, happy and surprise are main contributing responses for robotic face. Some of the combinations of happy and surprise are given below in Figures 10e,

Neutral			
Figure 10 a)			
Happiness			
	Happy	More Happy	Most Happy
Figure 10 b)			

Sadness			
	Sad	More Sad	Most Sad
Figure 10 c)			
Angry			
	Angry	More Angry	Most Angry
Figure 10 d)			

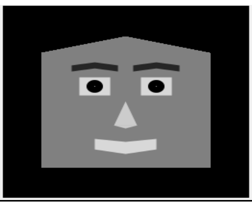
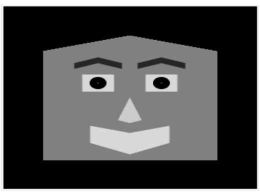
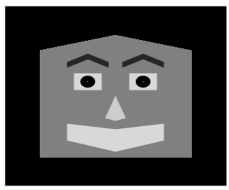
Mixture of Happy and Surprise			
	Happy and Surprise	More Happy and More Surprise	Most Happy and Most Surprise
Figure 10 e)			

Figure10. Different responses generated by robotic simulator for various input combinations.

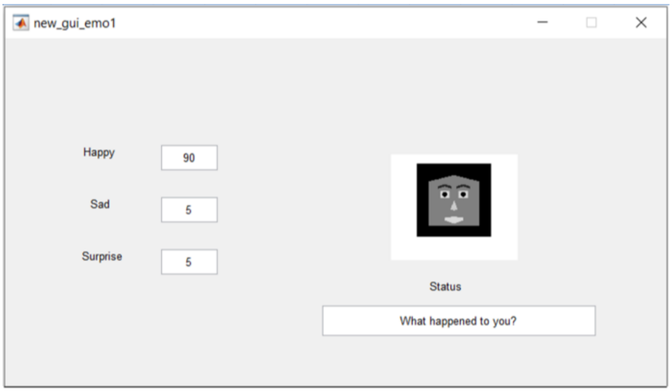


Figure 11. Recognized facial expression intensity with response of the robotic face simulator and

Figure 11, shows recognized facial expression intensity with response of the robot face simulator and interaction by robot with the user. System is developed for three input intensities, 33= 27 rules are generated along with interactive messages. As an example, one of the situations is, if user has 90 % happiness, 5% sadness and 5% surprise on user's, robotic simulator responds with more happy and surprise emotion including interaction, "What happened to you''? with usen.

6. CONCLUSION

A system is developed to detect the human facial expression. Proposed system developed a robotic face simulator to communicate with people as a companion. Primarily, the emotions of user are detected, classified, and displayed emotion vector possessing probability of each emotion out of the six basic emotions, happy, sad, surprise, anger, disgust, and fear in real-time. System trained on two datasets with three different optimizers and reached to the conclusion that RAF dataset with Adam optimizer has given best accuracy which is 82.34%. Training accuracy of the CK+ dataset is much greater but the testing in real-time has poor performance because of an insufficient number of images. The achievement in the accuracy was nearly state-of-the-art on the RAF dataset as compared to another dataset. After emotion detection by CNN, response emotion intensity generation is carried out by considering three input intensities and two output intensities. FBREIG has been employed in order to generate response emotion intensities for the robotic face simulator. Robotic simulator is developedto display response emotion intensities generated by FBREIG. Model has the ability to work in real world.However if for six input intensities, 37= 2187 rules have to be generated which will make the system more complex. Generative adversarial network can be adopted in order to give companion reactions in more flexible way.

References

1. Darwin, C.: The Expression of the Emotions in Man and Animals. London, 1872. DOI:
2. https://doi.org/10.1017/CBO9780511694110
3. Atone S.A., Bhalchandra A.S. (2022) Autonomous Facial Expression Detection and Human-Robot-Interaction: A Survey. In: Second International Conference on Image

Processing and Capsule Networks. ICIPCN. Lecture Notes in Networks and Systems, (300). Springer.DOI: https://doi.org/10.1007/978-3-030-84760-9_43
4. Li, Y.; Mavadati, S.M.; Mahoor, M.H.; Zhao, Y.; Ji, Q. Measuring the intensity of spontaneous facial action units with dynamic Bayesian network. Pattern Recognit., 48, 3417–3427, 2015. DOI: https://doi.org/10.1016/j.patcog.2015.04.022.
5. Kakadiaris, I.A.; Passalis, G.; Toderici, G.; Murtuza, M.N.; Lu, Y.; Karampatziakis, N.; Theoharis, T. Three- dimensional face recognition in the presence of facial expressions: An annotated deformable model approach. IEEE Trans. Pattern Anal. Mach. Intell., 29, 640–649, 2007. DOI:10.1109/TPAMI.2007.1017
6. Huang, M.; Wang, Z.; Ying, Z. Facial expression recognition using stochastic neighbor embedding and SVMs. In Proceedings of the 2011 International Conference on System Science and Engineering (ICSSE), Macao, China, 8–10, 671–674, June 2011. DOI: 10.1109/ICSSE.2011.5961987
7. Tang, H.; Huang, T.S. 3D facial expression recognition based on automatically selected features. In Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops, CVPRW'08, Anchorage, AK, USA, 23–28, 1–8, June 2008. DOI:10.1109/CVPRW.2008.4563052
8. Pal, R.; Satsangi, C. Facial Expression Recognition Based on Basic Expressions and Intensities Using K-Means Clustering. Int. J. Sci. Res. 5, 1949–1952, 2016.
9. Zhang, H.; Luo, S.; Yoshie, O. Facial expression recognition by analyzing features of conceptual regions. In Proceedings of the IEEE/ACIS 12th International Conference on Computer and Information Science (ICIS), Niigata, Japan, 16–20, 529–534, June 2013. DOI:10.1109/ICIS.2013.6607893.
10. Soyel, H.; Demirel, H. Facial expression recognition using 3D facial feature distances. In International Conference Image Analysis and Recognition; Springer: Berlin/Heidelberg, Germany, pp. 831– 838, 2007. DOI: https://doi.org/10.1007/978-3-540-74260-9_74.
11. Savran, A.; Sankur, B.; Bilge, M.T. Regression-based intensity estimation of facial action units. Image Vis. Comput. 30, 774–784, June 2013. DOI:https://doi.org/10.1016/j.imavis.2011.11.008.
12. Mahoor, M.H.; Cadavid, S.; Messinger, D.S.; Cohn, J.F. A framework for automated measurement of the intensity of non-posed facial action units. In Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops, Miami, FL, USA, 74–80, June 2009.DOI: 10.1109/CVPRW.2009.5204259
13. Ahmed, F.; Hossain, E. Automated facial expression recognition using gradient-based ternary texture patterns. Chin. J. Eng. , 2013.DOI: https://doi.org/10.1155/2013/831747
14. Pantic, M.; Rothkrantz, L.J. An expert system for recognition of facial actions and their intensity. In Proceedings of the 17th National Conference on Artificial Intelligence and 12th Conference on on Innovative Applications of Artificial Intelligence, Austin, TX, USA, 30, 1026–1033. August 2000;
15. Song, K.T.; Chen, Y.W. A design for integrated face and facial expression recognition. In Proceedings of the IECon 2011-37th Annual Conference on IEEE Industrial Electronics Society, Melbourne, Australia, 7–10, 4306–4311, November 2011.DOI: 10.1109/IECON.2011.6120016
16. Chang, K.J.; Bowyer, K.W.; Flynn, P.J. Effects on facial expression in 3D face recognition. In Proceedings of the Biometric Technology for Human Identification II, Orlando, FL, USA, 5779, 132–144, 28 March 2005.
17. Ekman, P., Friesen W.V.: Facial Action Coding System. A Technique for the Measurement of Facial Movement, 1978.DOI: https://doi.org/10.1037/t27734-000
18. Mase K. Recognition of facial expression from optical flow. IEICE Trans E[J]. Ieice Transactions on Information & Systems, 74(10), 1991.

19. Hinton G E; Salakhutdinov R R. Reducing the dimensionality of data with neural networks [J]. Science, 313, 504-507, 2006. DOI: 10.1126/science.1127647.
20. Tarik A. Rashid, Convolutional Neural Networks based Method for Improving Facial Expression Recognition, Second International Symposium on Intelligent Systems Technologies and Applications (ISTA'16), Jaipur, 2016. DOI:10.1007/978-3-319-47952-1_6.
21. Kukla E., & Nowak P. Facial Emotion Recognition Based on Cascade of Neural Networks. New Research in Multimedia and Internet Systems, 67–78, 2015. DOI:10.1007/978-3-319-10383-9_7.
22. Hyung, H.-J., Yoon, H.U., Choi, D., Lee, D.-Y., Lee, D.-W.: Optimizing android facial expressions using genetic algorithms. Appl. Sci. 9, 3379 , 2019. DOI: https://doi.org/10.3390/app9163379.
23. Habib, A., Das, S.K., Bogdan, I., Hanson, D., Popa, D.O.: Learning human-like facial expressions for Android Phillip K. Dick. In: 2014 IEEE International Conference on Automation Science and Engineering (CASE), New Taipei, Taiwan, pp. 1159–1165, 2014. DOI: 10.1109/CoASE.2014.6899473.
24. Kim, H.R., Kwon, D.S.: Computational model of emotion generation for human-robot interaction based on the cognitive appraisal theory. J. Intell. Robot. Syst. 60, 263–283, 2010. DOI:10.1007/s10846-010-9418-7
25. P. Lucey, J. F. Cohn, T. Kanade, J. Saragih, Z. Ambadar and I. Matthews.:The Extended Cohn-Kanade Dataset (CK+): A complete dataset for action unit and emotion-specified expression, 2010 IEEE Computer Society Conference on Computer Vision and Pattern Recognition - Workshops, San Francisco, CA, USA, 94-101, doi: 10.1109/CVPRW.2010.5543262.
26. Li, Shan and Deng, Weihong and Du, JunPing, Reliable Crowdsourcing and Deep Locality-Preserving Learning for Expression Recognition in the Wild, 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2584—2593. doi: 10.1109/TIP.2018.2868382.
27. Shan Ke, GuoJunqi, Youwenwan,Lu Di And Bierongfang. Automatic facial expression recognition based on a deep convolutional-neural-network structure.IEEE15th International Conference on Software Engineering Research, Management and Applications (SERA), 123-128, 2017. DOI: 10.1109/SERA.2017.7965717.