International Journal of Disaster Recovery and Business Continuity
Vol.12, No. 1, (2021), pp. 930–944

Facial Expression Recognition with Fused Deep and Geometric Features

Garima Sharma, Dr. Rekha Vig, Dr. Latika Singh

garimasharma@ncuindia.edu, rekhavig@ncuindia.edu, latikaduhan@sushantuniversity.edu.in

Abstract

A novel Facial expression recognition(FER) systemproposed in this study uses a combinationof deep and handcrafted geometric facial features to automatically recognise expressions from input images. The system captures the subtle details in the input image by extracting deep features of the image using a Convolution neural network built from scratch. The handcrafted geometric features used in our work carry the spatial information and the relative distances between various facial landmark points like eyebrows, eyesand mouth etc. Finally ,the two different feature vectors obtained aremerged together in order to create the final features set which is fed to a multi-class SVM for classification. The hyper-parameters for the SVMare also optimised to achieve good recognition accuracy. The performance of the system is assessed by performing experiments on two publically available image databases.Experiments reveal that recognition accuracy of 81.39% and 86.01% is achieved for JAFFE and MUG database respectively.

1. Introduction

Facial expression recognition systems have gained attention of numerous researchers as well as psychologists over the last few decades due to its enormous applications in the field of human computer interaction systems[1], surveillance security systems[2], social robots[3], biometric recognition systems[4] and interactive gaming systems. Face is considered to be an important medium for non-verbal communication due to the subtle cues it encodes in form of expressions. This encoded information is utilised by various automatic facial expression recognition systems to classify emotions into one of the six basic emotion classes proposed by Ekman and Friesen[5]. The six basic emotions are disgust, anger, fear, sadness, happiness and surprise. Recognition of emotions from facial expression appears to be a simple task from human's perspective but building an automatic system to detect emotions from the facial gestures is still a challenge.

The process of automatic facial expression recognition(FER) in various traditional systems proposed is divided into 3 phases[6]: Face detection and pre-processing, extraction of relevant features and finally expression classification as shown in Fig.1. The first phase extracts the face region from the input image to remove the unnecessary background information and further performs the pre-processing of the extracted face region like resizing the images, denoising, contrast enhancement etc. The output pre-processed image is then passed on as input to the feature extraction phase which represents the image into a feature vector containing the meaningful information required for the classification. The commonly used handcrafted feature extraction techniques[7] used in the process of facial expression recognition are categorised into: Geometric and appearance based features. Geometric facial features carry the spatial information like the relative distances between various feature points like eyes and nose etc., whereas the appearance based features encode the texture or the pixel information in the image like

ISSN: 2005-4289 IJDRBC

intensity values. Some of the proposed work[7],[8],[9] uses a combination of both geometric as well as appearance based features to generate the feature vector. The final step is the classification of expressions into one of the emotion classes by training a classifier based on the extracted features in the second phase. FER'S based on traditional handcrafted features techniques include BoW[10][11], LDA[12], SiFT[13], Gabor filters[14], landmark techniques [15],[16]etc.

Figure 1 Traditional Facial Expression Recognition System

In the recent works, researchers have focussed on applying deep learning models [17], [18] for automatically extracting features from the images as well as performing classification due to their ability to overcome challenges like illumination changes, occlusions and relative camera angle pose. The features extracted from the deep learning models provides a distributed representation of low level as well as high level features[19]. An attentional convolution network based system [20]performs the expression classification on multiple databases like FER-2013, JAFFE, CK+ and FERG by focussing on the significant feature rich regions of the face and achieved good recognition accuracy.In [21],an illumination augmentation scheme used with compact CNN model also showed promising results.

This study proposes a novel FER system by forming a fusion of deep and handcrafted geometric features. The CNN model automatically generates the deep features and the geometric feature vector is formed by obtaining the co-ordinates of relevant facial feature pointsand the relative distances between them. Both the feature vectors are concatenated together and then a multi-class SVMclassifier is trained for classifying the images into one of the target expression classes. **The proposed** model is tested on two publically available facial expression databases namely: The Japanese Female Facial Expression (JAFFE) [41] and Multimedia Understanding Group (MUG)[42]. Promising classification accuracy is obtained for both the databases. The workflow of the proposed system include: (i) Extraction of deep features before the classification layer of a CNN model build from scratch (ii) Fusion of extracted deep features with geometric features, and finally (iii) Performing learning and expression classification of the fused features using a SVM model. The fused features result in capturing even the minute expression changes. The rest of the paper is organisedas follows: Section 2 focusses on the related works available in literature, Section 3 explains the proposed work in detail, and Section 4 presents the implementation details as well as description about the databases used in the study, followed by Section 5 which contains the conclusion and the future scope of the proposed work.

2. Related works

Several works have focussed on use of deep and handcrafted features for building efficient facial emotion recognition system. This section focusses on the existing systems proposed in this area. Most of the early systems proposed for automatic facial expression recognition utilised handcrafted features for emotion classification[22]. A study [23]presents a feature based method which finds out the descriptive features from the image and performsclassification using random forest and decision tree classifiers. The system was evaluated on five publically available databases and obtained a maximum accuracy of 99.7%.In [24], a facial expression recognition used discrete cosine transform, Gabor filtering technique along with angular radial transform features for recognising facial expressions and achieved promising results. Another FER system proposed in[25], performed feature extraction using complex wavelet transforms and implemented three different classifiers to evaluate the performance for the system to achieves desirable recognition accuracy. One of our earlier works [26]focussed on finding the dimensional attributes from the mouth region of the face image and used those attributes to categorize the input image into happy and neutral emotion classes. The system's performance was evaluated for two publically available databases and achieved recognition accuracy of 70% on JAFFE and NimStim database and 95% on MUG database.

Deep- learned features have gained attention of the researchers in recent years due to the rich semantic information they provide and also because of the availability of large training databases. A 3-D CNN with inception-ResNet layers along with LSTM[27]extractsthe spatio-temporal facial features. The facial landmark points are also fed into the network to improve the recognition accuracy.

Another study [28]presented a multi-task shallow CNN to capture the low level features details of the images along with a part-based framework to capture dynamic spatio-temporal features from local facial regions. A unified loop methodologyis presented in [29]which uses both unsupervised and supervised feature learning to obtain robust features for efficient expression classification. A study[30] focussed on extracting both the local as well as global facial features by using an ensemble of CNN for multiple facial regions.The expression recognition is done by combining the weighted prediction scores from each sub-network built for different facial regions. A study [31]proposed in 2019 experimented by varying the CNN parameters like size and the number of filters and optimizer type and finally achieved recognition accuracy of 65.23% and 65.77% by building two CNN models. Recently micro-expression recognition systems have also been developed using deep features[32], [33], [34]. Although, availability of micro-expression training datasets is still a problem.Some of the recent works[35],[36],[37]have attempted to use an ensemble of CNN forincreasing the performance of FER systems by improving the recognition rates.

In this paper, a blend of deep and handcrafted features is utilised to generate the final feature vector for performing facial expression classification. Although, deep and handcrafted features have been integrated together in some of the studies[38], [39], but we have extracted the deep features from a CNN network which is built from scratch and the variation hyper-parameters of the network are fine-tuned to attain better results. Also the handcrafted features used in the study are the geometric features proposed in one of

ISSN: 2005-4289 IJDRBC

our previous works[40] . Amulti-class SVM is trained using the combined features to perform expression classification. Theexpression recognition performance of the system is assessed by conductingexperiments on two publically availabledatabases **JAFFE** [41]**and MUG**[42]**.**

3. The Proposed model

The proposed approach uses two separate feature extraction system to form the complete framework for feature set generation shown in Figure 2. Finally when both the deep and handcrafted features are extracted separately, both the feature vectors are fused together to perform the final classification using SVM Classifier.

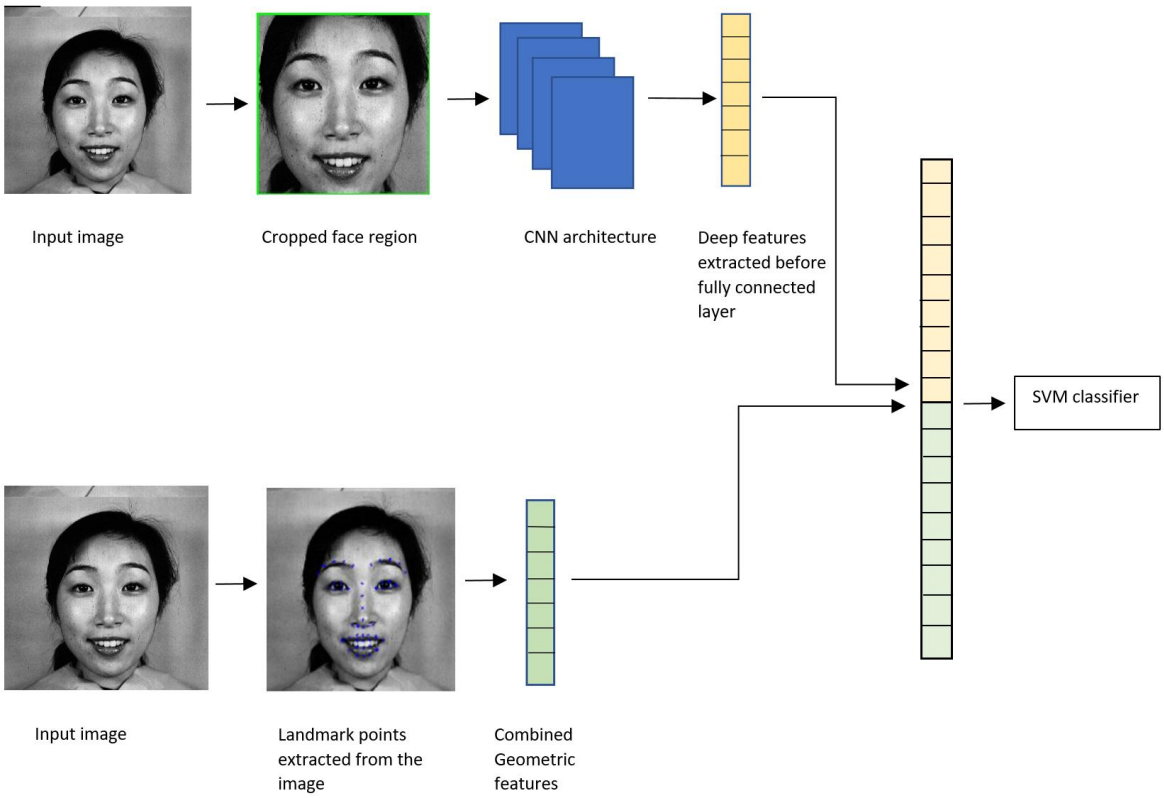

Figure 2 Overview of the proposed framework

A. Extraction of deep features

The process of extraction of deep features is performed using CNN. The input images are pre-processed and passed on to the CNN. In pre-processing step, face detection is performed to extract the facial region from the given image by using Viola-Jones algorithm[43]**.** The algorithm scans a small window over the entire input image and utilises the modified Adaboost algorithm which constructs a strong classifier by combining cascaded weak classifiers to select the best performing features in the image. The cascaded classifiers differentiate the windows containing the face and non-face region.

The cropped facial image is then resized to 128*128 and normalised to zero mean and unit variance before passing it to the CNN model for better and faster results. The

ISSN: 2005-4289 IJDRBC

proposed 6-layer CNN architecture as shown in Figure 3 is built from scratch. The CNN network contains three convolution layers, two max-pooling layers and one fully-connected layer. The purpose of convolution layers is to obtain the relevant features of the face image while conserving the spatial relationship between the pixels. The max pooling layers targets at reducing the dimensionality of the obtained feature maps by choosing the largest element within the window.

The first convolution layer filters the grayscale face image using 6 kernels of size 5*5 and obtains the output of size 62*62*6. The output obtained from the first convolution layer is passed to max-pooling layer followed by second convolution layer. The kernel size used in the second convolution and third convolution layer is also 5*5. The output of size 58*58*120 is obtained from the second convolution layer which is then fed into the max-pooling layer. Finally the third convolution layer is applied and output size of 25*25*120 is obtained. ReLU[44] activation function is used for all the convolution layers**.** Dropout layer[45]is added after the third convolution layer and also after the first fully connected layer in order to prevent overfitting. Loss is calculated using softmax function.

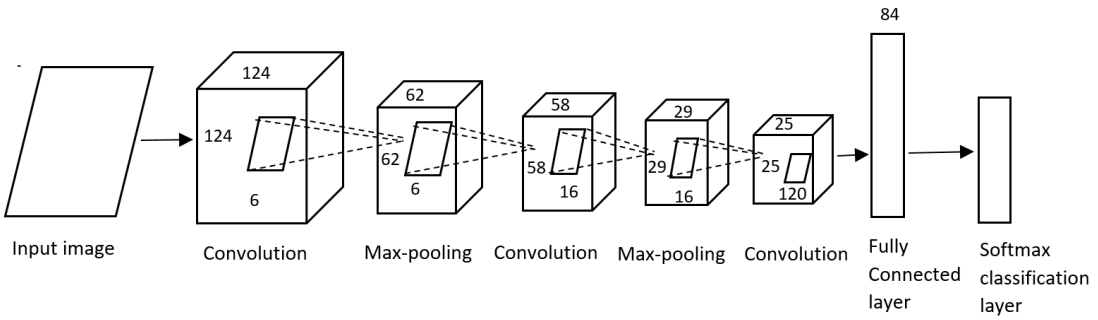

Figure 3. The proposed CNN architecture

84 deep features are extracted from CNN model just before the final classification layer to create the feature vector as shown in Figure 4. The deep features are actually the activation maps formed after the fully connected layer. A total of 84 values are extracted to form the feature vector which are then normalised using Batch Normalisation technique which reduces the internal co-variant shift to improve the model training process.

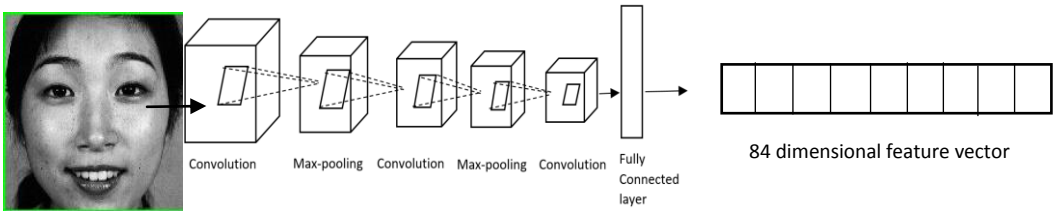

Figure 4 Deep features extraction

B. Extraction of Handcrafted Geometric feature

The geometric features extracted from the face images from one of our previous work[40] is used as handcrafted features in this study. The process includes pre-processing the input images to remove noise by using a Gaussian filter and then performing histogram equalisation to enhance the contrast of the images. Finally a landmark localisation feature extraction technique is employed to extract 49 fiducial points from the face image. This process gives a feature vector of size 98 containing spatial coordinates of the extracted facial points. Further the feature vector is enhanced to include relative distances between the feature points for attaining better recognition accuracy of the system. In total, 15 distances proposed in the previous study are added to construct the total feature vector of size 113. Euclidean distance measuring technique is deployed to find out the distance values between the coordinate points.

C. Fusing the deep and handcrafted features

The deep and handcrafted geometric features extracted are concatenated together to build the final feature set of 297 features as shown in Table 1. The combined features are deployed for the classification of images into 7 target emotion classes.

Feature sets	No. of attributes
Deep features	84
Handcrafted features	113
Combined features	297

Table 1 Description of features used in the proposed work

4. Implementation details

A. Databases Used

Two publically available databases, JAFFE and MUG are used to evaluate the recognition performance of the proposed framework into 7 target classes including six universal emotion classes i.e. disgust, anger, fear, sadness, happiness, surprise, and one neutral emotion class. The details of the databases used in the study is as follows:

I. Japanese Female Facial Expression (JAFFE) database

JAFFE database has 213 images posed by female subjects of Japanese origin in all 6 emotion states as well as in neutral state. The images are in 8-bit grayscale format and the resolution of each image is 256*256 pixels. Each image is stored in .Tiff format. There are several images of each emotion class for all the subjects.The number of images considered for angry and surprise emotion classes are 30 each. 31 images each are used for both happy and sad classes. 29 and 32 images are used for disgust and fear classes respectively.

ISSN: 2005-4289 IJDRBC

All the 213 images from the dataset are publically available and are used in the experiments performed in this study.

II. Multimedia Understanding Group (MUG) database

MUG database consists of facialimages of 86 subjects of Caucasian origin, out of which only 52 are available to authorised internet users. The subjects include 35 women and 51 men in age group of 20 years to 35 years. The resolution for each image in the database is 896*896 pixels. Images are saved in .JPG extension. All the images are in RGB format and size of the images is between 240kb to 340 KB. This study performed experiments on 712 images selected randomly from 1462 publically available image sequences of the dataset with 52 subjects. The number of images considered for angry, disgust, happy, neutral and surprise emotions is 104 images per emotion class. 94 and 98 images are used for fear and sad emotion classes respectively.

B. Results

In order to assess the performance of the proposed model, the process of facial expression recognition is implemented using three different classification models given below:

i) The first model classifies the input images using only deep features on the CNN model itself,
ii) In the second model, deep features are obtained from the CNN's fully connected layer and passed to SVM multi-class classifier for training and classification
iii) The third model uses the combined deep and handcrafted features and performs the classification on the SVM classifier.

I. Classification using deep features on CNN

In this model, the CNN architecture discussed in Section 3 of this paper for the extraction of deep features is utilised for the classification as well. The CNN model is trained for both the datasets i.e. JAFFE and MUG independently and the testing accuracies for both the datasets is obtained. The databases are split in the ratio of 80:20 to form the training and testing datasets.Table 4 contains the total number of images used for training and testing for each database.

Table 2 Number of images used for training and testing for both JAFFE and MUG database

Database	Images used for training	Images used for testing
JAFFE	170	43
MUG	569	143

Early stopping method is used to avoid over fittingof data due to increased number of epochs while training the network. The maximum number of epochs is fixed to 50 but

ISSN: 2005-4289 IJDRBC

the model stops training when the results does not improve. The model runs for 21 epochs in case of JAFFE dataset and achieves a validation accuracy of 65.12%. For MUG dataset, accuracy of 84.62% is achieved after 30 epochs as mentioned in Table 5.

Table 3 Recognition accuracy achieved for deep features on CNN model

Datasets	*JAFFE*	*MUG*
Recognition accuracy on CNN model	65.12%	84.62%

Figure 5 shows the graphs obtained for accuracy and loss after training the model for both the databases.

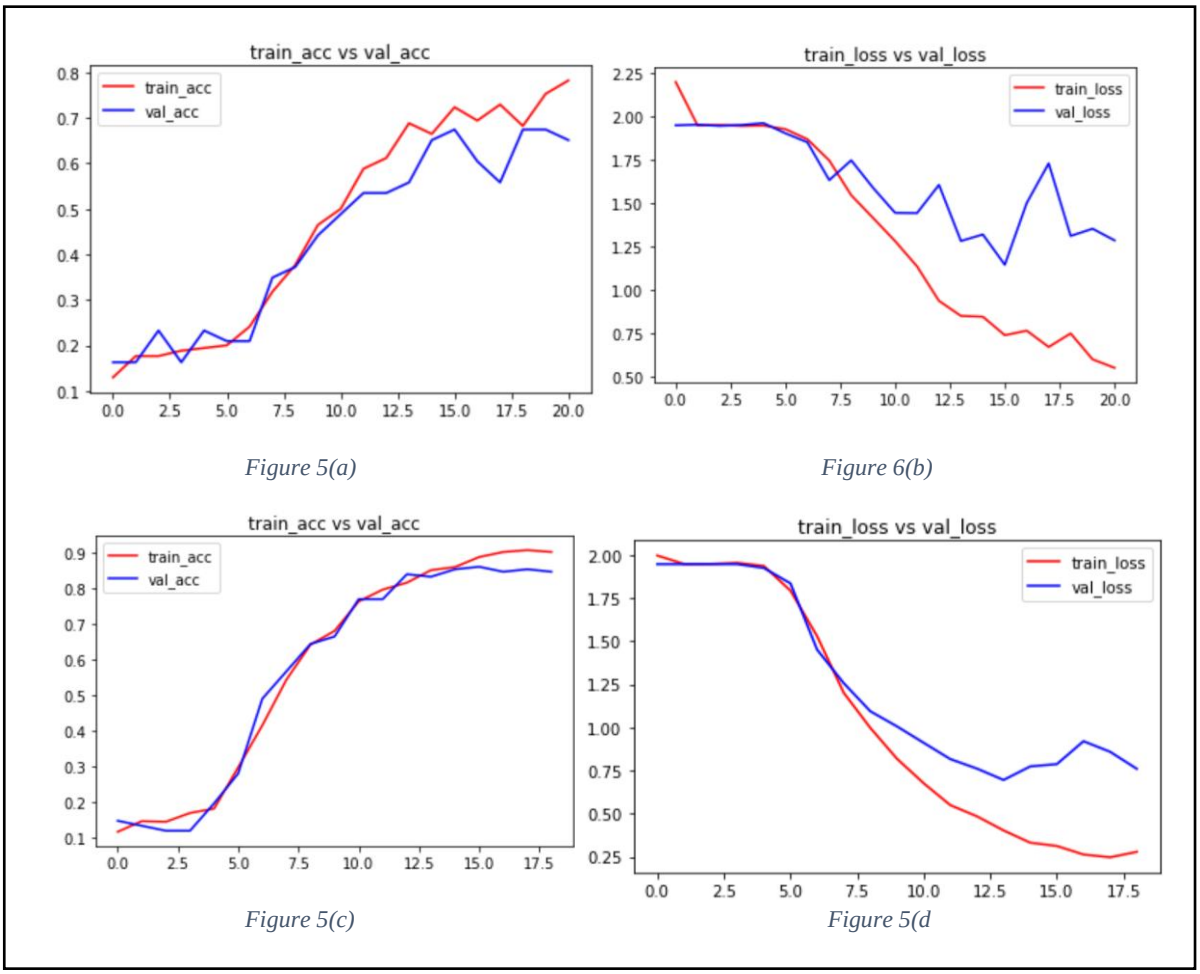

Figure 5(a) *Figure 6(b)*

Figure 5(c) *Figure 5(d*

Figure 5

Figure 5(a)&(b) Accuracy and loss graph obtained for JAFFE database; Figure 5(c) & (d)Accuracy and loss graph obtained for MUG database

Average recognition accuracy of 65.11% is obtained on JAFFE dataset and 82.52% on MUG database as shown in Table 6.

ISSN: 2005-4289 IJDRBC

Table 4 Recognition accuracy achieved for deep features on SVM classification model

Datasets	*JAFFE*	*MUG*
Recognition accuracy on SVC model	65.11%	82.52%

III. Classification of combined features on SVM

Finally the combined deep and handcrafted features are evaluated using Support vector classification(SVC) model[46] due to its high recognition accuracy as compared to other classifiers. SVC is based on statistical learning, which aims to minimise the uncertainty of the model and maximise the fitness.

T

Datasets	*JAFFE Database*	*MUG database*
Combined deep and handcrafted feature classification using SVM	81.39 %	86.01%

a b l e 5

Recognition accuracy achieved using combined features on SVM classification model

As Support vector machines(SVM) are originally suited for binary classification problem[47]**,**one-versus one strategy is utilised in the study to perform multi-class classification of the input features as one of the seven target classes of emotions. The total number of classifiers required to perform n-class(n=7) classification in case of one-versus-one technique is depicted by equation (1) below:

$$\boldsymbol{n = n * (n - 1)/2} \qquad \textbf{(1)}$$

The multi-class classifier is constructed by combining all one-versus-one classifiers generated. Further, the SVC model is optimised to find out the best performing parameters for the given datasets by using 'GridSearchCV' class in Python. The C, gamma and kernels are the hyper parameters we have experimented with, in this study.

The hyper-parameters chosen for the two datasets used in the experiments in given in Table 8 below:

Table 6 Optimal Hyper-parameter values chosen to perform classification for both databases

Dataset	Kernel	Gamma	C
JAFFE	Radial basis function(rbf)	0.0001	1000
MUG	Linear	0.001	1

The obtained result show that the combined features achieved better recognition accuracy percentage as compared to deep features as shown in Figure 6. For JAFFE database, using the combined features on SVM classifier yielded recognition accuracy of 81.39% which is considerably better than 65.12% and 65.11% achieved using just deep features on both CNN and SVM classifier respectively. Performing the same experiment on MUG database gave a maximum of 86.01% accuracy using the combined features whereas 82.52% and 85.31% of recognition accuracy is achieved using deep features on CNN and SVM classifiers respectively. Although the

ISSN: 2005-4289 IJDRBC

difference value in the accuracy for MUG database is not very large but the combined features have shown better results than using just deep features on both the classifiers.

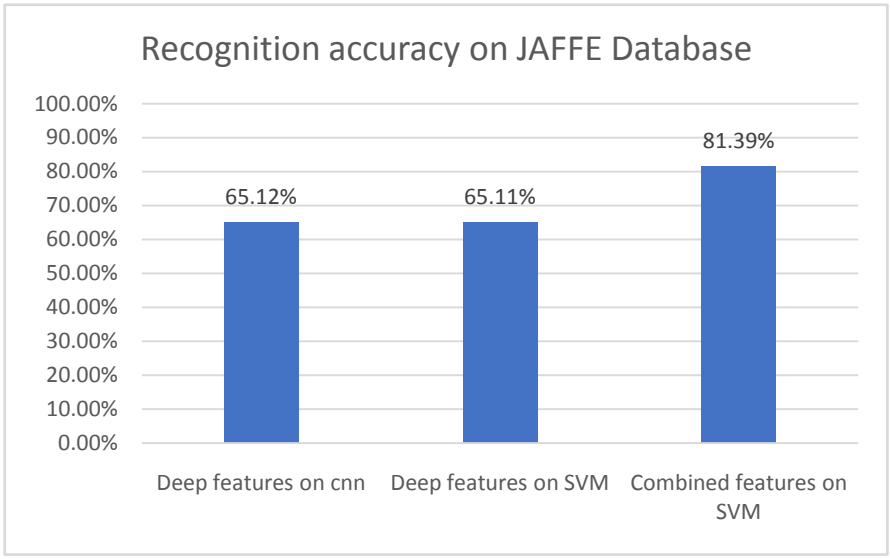

Figure 6(a) Comparison of recognition accuracies achieved for JAFFE database using all 3 classification models

Figure 6(b) Comparison of recognition accuracies achieved for MUG database using all 3 classification models

Figure 6.

The study further evaluated the classification done by the third classification model of combined features using precision, recall and f1-score as evaluation parameters. The values achieved for both JAFFE and MUG databases is summarised in Table 9 and 10 respectively.

ISSN: 2005-4289 IJDRBC

Table 7. Values of evaluation parameters obtained for JAFFE Database

JAFFE Database			
Emotion	Precision	Recall	f1-score
Angry	0.62	0.83	0.71
Disgust	0.75	0.67	0.71
Fear	1.00	0.88	0.93
Happy	1.00	1.00	1.00
Neutral	1.00	1.00	1.00
Sad	0.25	0.33	0.29
Surprise	1.00	0.86	0.92

Table 8 Values of evaluation parameters obtained for MUG Database

MUG Database			
Emotion	Precision	Recall	f1-score
Angry	0.78	1.00	0.88
Disgust	0.75	0.86	0.80
Fear	0.67	0.67	0.67
Happy	1.00	0.89	0.94
Neutral	0.88	0.85	0.86
Sad	0.90	0.86	0.88
Surprise	0.90	0.82	0.86

Table 11 shows a comparison of the proposed system with some state of the art methods which validates the efficiency of the proposed system. The results reveal thattraining a SVM classifier using both deep and the handcrafted geometric features achieved better recognition accuracies.

Table 9 Comparison of the proposed sytem with other state-of-the art methods

Related Works	**JAFFE**	**MUG**
[48]	70.48%	_
[49]	76.7442%	_
[50]	75%	_
[51]	_	82.5 -85.4 %
[52]	_	81%
[53]	75.5%	66.4
[54]	79.21%	_
Our Proposed system	**81.39%**	**86.01%**

5. Conclusion and Future work

The study presented a novel facial expression recognition system which utilised the deep features extracted from trained CNN model and combined them with handcrafted geometric features containing the spatial information of the input face image. The recognition accuracy achieved from the combined features is compared with deep features and it has been observed

ISSN: 2005-4289 IJDRBC

that the combined features gave better recognition rate for both the databases used in the proposed work. The recognition results are also comparable with some of the state of the art methods available in literature and the results indicate that the presentedsystem has performed well in most of the cases.

The system can be further improved by trying a fusion different geometric and appearance featuresalong with the extracted deep features, in order to incorporate more information about the image. Moreover different pre-trained networks like Alex-net, VGG, InceptionV3 can be experimented for extracting the deep features from the databases. Also a different combination of hyper-parameters of the CNN model can be tried for enhancing the recognition performance of the system.

REFRENCES

[1] F. Dornaika and B. Raducanu, "Facial Expression Recognition for HCI Applications," *Encyclopedia of Artificial Intelligence*, 2011.

[2] M. A. Butalia, M. Ingle, and P. Kulkarni, "Facial Expression Recognition for Security," *International Journal of Modern Engineering Research. www.ijmer.com*, vol. 2, no. 4, pp. 1449–1453, 2012.

[3] G. Littlewort, M.S. Bartlett, I.R. Fasel, J. Chenu,T. Kanda, H. Ishiguro, &J.R. Movellan, "Towards social robots: Automatic evaluation of human-robot interaction by face detection and expression classification," *Adv. Neural Inf. Process. Syst.*, 2004.

[4] S. Cheng, I. Kotsia, M. Pantic, and S. Zafeiriou, "4DFAB: A large scale 4d database for facial expression analysis and biometric applications," *Proceedings of the IEEE conference on computer vision and pattern recognition*. 2018.

[5] P. Ekman, W.V. Friesen,M. O'sullivan, A. Chan, I. Diacoyanni-Tarlatzis, K. Heider,R. Krause, W.A LeCompte,T. Pitcairn, P.E. Ricci-Bitti, and K. Scherer, "Universals and Cultural Differences in the Judgments of Facial Expressions of Emotion," *Journal of Personality and Social Psychology*, 1987.

[6] I. M. Revina and W. R. S. Emmanuel, "A Survey on Human Face Expression Recognition Techniques," *Journal of King Saud University - Computer and Information Sciences*, 2018.

[7] A. A. A. Youssif and W. A. A. Asker, "Automatic Facial Expression Recognition System Based on Geometric and Appearance Features," *Computer and Information Sciences*, vol. 4, no. 2, 2011.

[8] M. Rahul, N. Kohli, R. Agarwal, and S. Mishra, "Facial expression recognition using geometric features and modified hidden Markov model," *Int. J. Grid Util. Comput.*, vol. 10, no. 5, pp. 488–496, 2019.

[9] Z. Lin, Y. Huang, and Y. Liu, "Facial expression recognition based on displacement feature and random forest," *Guangxue Jishu/Optical Tech.*, vol. 44, no. 1, pp. 25–29, 2018.

[10] D. Al Chanti and A. Caplier, "Improving bag-of-Visual-Words towards effective facial expressive image classification," *VISIGRAPP 2018 - Proc. 13th Int. Jt. Conf. Comput. Vision, Imaging Comput. Graph. Theory Appl.*, vol. 5, no. Visigrapp, pp. 145–152, 2018.

ISSN: 2005-4289 IJDRBC

[11] R. T. Ionescu, M. Popescu, and C. Grozea, "Local Learning to Improve Bag of Visual Words Model for Facial Expression Recognition," *Work. challenges Represent. Learn. ICML*, pp. 1–6, 2013.

[12] J. H. Shah, M. Sharif, M. Yasmin, and S. L. Fernandes, "Facial expressions classification and false label reduction using LDA and threefold SVM," *Pattern Recognit. Lett.*, 2017.

[13] S. Berretti, B. Ben Amor, M. Daoudi, and A. Del Bimbo, "3D facial expression recognition using SIFT descriptors of automatically detected keypoints," *Vis. Comput.*, vol. 27, no. 11, pp. 1021–1036, 2011.

[14] W. F. Liu and Z. F. Wang, "Facial expression recognition based on fusion of multiple gabor features," *Proc. - Int. Conf. Pattern Recognit.*, vol. 3, no. January 2006, pp. 536–539, 2006.

[15] Fuzail Khan, "Facial expression recognition using facial landmark detection and feature extraction via neural networks,"arXiv preprint arXiv:1812.04510, 2018.

[16] S. L. Happy and A. Routray, "Automatic facial expression recognition using features of salient facial patches," *IEEE Trans. Affect. Comput.*, vol. 6, no. 1, pp. 1–12, 2015.

[17] E. Barsoum, C. Zhang, C. C. Ferrer, and Z. Zhang, "Training deep networks for facial expression recognition with crowd-sourced label distribution," *ICMI 2016 - Proc. 18th ACM Int. Conf. Multimodal Interact.*, pp. 279–283, 2016.

[18] A. A. Salah, H. Kaya, and F. Gurpnar, "Video-based emotion recognition in the wild," *Multimodal Behav. Anal. Wild Adv. Challenges.*, no. December, pp. 369–386, 2018.

[19] X. Zhao, X. Shi, and S. Zhang, "Facial expression recognition via deep learning," *IETE Tech. Rev. (Institution Electron. Telecommun. Eng. India)*, vol. 32, no. 5, pp. 347–355, 2015.

[20] S. Minaee and A. Abdolrashidi, "Deep-emotion: facial expression recognition using attentional convolutional network," *arXiv*, 2019.

[21] R. Reji, P. Sojan Lal, A. M. Philip, and V. Vishnu, "A compact deep learning model for robust facial expression recognition," *Int. J. Eng. Adv. Technol.*, vol. 8, no. 6, pp. 2956–2960, 2019.

[22] Y. Tian, T. Kanade, and J. F. Cohn, *Handbook of Face Recognition*, no. January 2017. 2011.

[23] S. Gupta, K. Thakur, and M. Kumar, "2D-human face recognition using SIFT and SURF descriptors of face's feature regions," *Vis. Comput.*, 2020.

[24] H. H. Tsai and Y. C. Chang, "Facial expression recognition using a combination of multiple facial features and support vector machine," *Soft Comput.*, vol. 22, no. 13, pp. 4389–4405, 2018.

[25] Alaa Eleyan, "Comparative Study on Facial Expression Recognition using Gabor and Dual-Tree Complex Wavelet Transforms," *International Journal of Engineering and Applied Sciences*, vol. 9, no. 1, pp. 1–1, 2017.

[26] G. Sharma, L. Singh, and S. Gautam, "Facial Feature Extraction for Emotion Classification using Fuzzy c-mean Clustering," *Recent Adv. Comput. Sci. Commun.*,

ISSN: 2005-4289 IJDRBC

vol. 13, pp. 1–10, 2020.

[27] B. Hasani and M. H. Mahoor, "Facial Expression Recognition Using Enhanced Deep 3D Convolutional Neural Networks," *IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit. Work.*, vol. 2017-July, pp. 2278–2288, 2017.

[28] M. Yu, H. Zheng, Z. Peng, J. Dong, and H. Du, "Facial expression recognition based on a multi-task global-local network," *Pattern Recognit. Lett.*, vol. 131, pp. 166–171, 2020.

[29] P. Liu, S. Han, Z. Meng, and Y. Tong, "Facial expression recognition via a boosted deep belief network," *Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit.*, pp. 1805–1812, 2014.

[30] Y. Fan, J. C. K. Lam, and V. O. K. Li, "Multi-region ensemble convolutional neural network for facial expression recognition," *arXiv*, 2018.

[31] A. Agrawal and N. Mittal, "Using CNN for facial expression recognition: a study of the effects of kernel size and number of filters on accuracy," *The Visual Computer*, vol. 36, no. 2, pp. 405–412, 2020.

[32] M. A. Takalkar and M. Xu, "Image Based Facial Micro-Expression Recognition Using Deep Learning on Small Datasets," *DICTA 2017 - 2017International conference on digital image computing: techniques and applications*, vol. 2017-December, pp. 1–7, 2017.

[33] J. Li, Y. Wang, J. See, and W. Liu, "Micro-expression recognition based on 3D flow convolutional neural network," *Pattern Anal. Appl.*, vol. 22, no. 4, pp. 1331–1339, 2019.

[34] S. P. T. Reddy, S. T. Karri, S. R. Dubey, and S. Mukherjee, "Spontaneous facial micro-expression recognition using 3d spatiotemporal convolutional neural networks," *arXiv*, pp. 1–8, 2019.

[35] K. Liu, M. Zhang, and Z. Pan, "Facial Expression Recognition with CNN Ensemble," *Proc. - 2016 Int. Conf. Cyberworlds, CW 2016*, pp. 163–166, 2016.

[36] Z. Yu and C. Zhang, "Image based static facial expression recognition with multiple deep network learning," *in Proceedings of ACM International Conference on Multimodal Interaction*, pp. 435–442, 2015.

[37] V. Rami, R. Chirra, S. R. Uyyala, V. Krishna, and K. Kolli, "Virtual facial expression recognition using deep CNN with ensemble learning," *J. Ambient Intell. Humaniz. Comput.*, no. 0123456789, 2021.

[38] M. I. Georgescu, R. T. Ionescu, and M. Popescu, "Local learning with deep and handcrafted features for facial expression recognition," *IEEE Access*, vol. 7, pp. 64827–64836, 2019.

[39] D. T. Nguyen, T. D. Pham, N. R. Baek, and K. R. Park, "Combining deep and handcrafted image features for presentation attack detection in face recognition systems using visible-light camera sensors," *Sensors*, vol. 18, no. 3, pp.699, 2018.

[40] G. Sharma, L. Singh, and S. Gautam, "Automatic Facial Expression Recognition Using Combined Geometric Features," *3D Res.*, vol. 10, no. 2, 2019.

ISSN: 2005-4289 IJDRBC

[41] M. J. Lyons, S. Akamatsu, M. Kamachi, J. Gyoba, and J. Budynek, "The Japanese female facial expression (JAFFE) database," *Proc. third Int. Conf. Autom. face gesture Recognit.*, 1998.

[42] A. Aifanti, N., Papachristou, C., & Delopoulos, "The MUG Facial Expression Database," in *Workshop Image Analysis for Multimedia Interactive Services (WIAMIS)*, 2010.

[43] Y.-Q. Wang, "An Analysis of the Viola-Jones Face Detection Algorithm," *Image Process. Line*, 2014.

[44] A. F. M. Agarap, "Deep Learning using Rectified Linear Units (ReLU)," *arXiv*, no. 1, pp. 2–8, 2018.

[45] H. Wu and X. Gu, "Max-pooling dropout for regularization of convolutional neural networks," *Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics)*, vol. 9489, pp. 46–54, 2015.

[46] J. Bi and T. Zhang, "Support vector classification with input data uncertainty," *Adv. Neural Inf. Process. Syst.*, 2005.

[47] E. Mayoraz and E. Alpaydm, "Support vector machines for multi-class classification," *Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics)*, vol. 1607, pp. 833–842, 1999.

[48] Z. Ying, L. Cai, J. Gan, and S. He, "Facial expression recognition with local binary pattern and laplacian eigenmaps," *Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics)*, vol. 5754 LNCS, pp. 228–235, 2009.

[49] K. Shan, J. Guo, W. You, D. Lu, and R. Bie, "Automatic facial expression recognition based on a deep convolutional-neural-network structure," *Proc. - 2017 15th IEEE/ACIS Int. Conf. Softw. Eng. Res. Manag. Appl. SERA 2017*, pp. 123–128, 2017.

[50] Noh, Sungkyu, et al. "Feature-adaptive motion energy analysis for facial expression recognition." *International Symposium on Visual Computing*. Springer, Berlin, Heidelberg, 2007.

[51] Aghamaleki, Javad Abbasi, and Vahid Ashkani Chenarlogh. "Multi-stream CNN for facial expression recognition in limited training data." *Multimedia Tools and Applications*, Vol. 78.16, pp. 22861-22882, 2019.

[52] Da Silva, Flávio Altinier Maximiano, and Helio Pedrini. "Effects of cultural characteristics on building an emotion classifier through facial expression analysis." *Journal of Electronic Imaging*, *Vol.* 24.2, pp.22861-82, 2015.

[53] Dufourq, Emmanuel, and Bruce A. Bassett. "Deep Evolution for Facial Emotion Recognition." *arXiv preprint arXiv:2009.14194* (2020).

[54] He, Lianghua, et al. "An enhanced LBP feature based on facial expression recognition." *2005 IEEE Engineering in Medicine and Biology 27th Annual Conference*. IEEE, 2006.

ISSN: 2005-4289 IJDRBC