International Journal of
INTELLIGENT SYSTEMS AND APPLICATIONS IN ENGINEERING
ISSN:2147-6799 www.ijisae.org Original Research Paper

2D Face Emotion Recognition and Prediction Using Labelled Selective Transfer Machine and CNN Transfer Learning Techniques for Unbalanced Datasets

Dipti Pandit *[1, 2], Dr. Sangeeta Jadhav [2]

Submitted: 10/09/2022 **Accepted:** 20/12/2022

Abstract: Emotion recognition and prediction using facial expressions is one of the most challenging activities in the computing arena. The most traditional approaches largely depend on pre-processing and feature extraction techniques. This paper factually represents the implementation and evaluation of learning algorithms like LSTM and different CNNs for recognition and predicting emotions of 2D facial expressions based on recognition rate, learning time as well as effect of unbalanced datasets. The proposed system has targeted two datasets CK+ and JAFFE for two different techniques, where LSTM includes histogram equalisation and LBP as pre-processing and feature extraction techniques respectively. A transfer learning technique is explored using Dense Net and VGG-19 algorithm. The recognition rate with Dense Net was 96.64% and 97.45% respectively on the CK+ and JAFFE datasets. LSTM also showed 98.43% of recognition rate on JAFFE where as 72.63% on CK+ dataset.

Keywords: *Face Emotion Recognition, LBP, LSTM, CNN, Transfer learning, Dense Net, VGG-19.*

1. Introduction

Face is an authoritative biometric aspect of humans, and to contribute to the research, face emotion prediction has fully-fledged a lot of interest from researchers in the computer vision, pattern recognition, and machine learning areas. Emotional face, or nonverbal communication, [1], [2], [3], is intercepted and interpreted in an assortment of contexts, straddling biology, neuroscience, commercial, website testing, sociology, security, computer science, psychology and plenty of others. Different models for emotion classification have been defined by emotion theorists and psychologists, ranging from commonly revealed basic emotions to customarily distinctive complicated emotions [4]. The most dominated facial expression research in psychology have been: Ekman's basic set of emotions [5], and Russell's circumflex model of affect [6]. Russell's circumflex model in figure 1 illustrates the emotions in four quadrants as intense, pleasant, mild and unpleasant.

Despite decades of research in Human Computer Interaction (HCI), many questions remain unanswered as to what should be the most relevant indications and expressions to be evaluated for meaning encoding of communications. The problem persists because it is critical to generalise classifiers to unknown subjects that differ in actions and facial structures such as brows, wrinkles due to ageing can miss or offer false expressions, and so on [7], [8]. Enough training data for individual classifiers is a potential solution that is not always available or feasible. Non-accurate computation, spontaneous affective behaviour, lighting variation due to head movements, registration techniques cause registration faults, and accessories or camera movements may cause occlusions are the other challenges in involuntary emotion recognition. Emotion classification always has been a measure concern, and it is handled using a variety of traditional and innovative techniques.

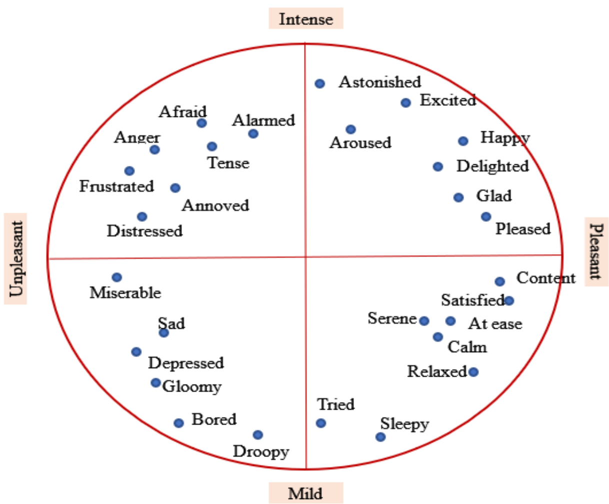

Fig. 1. Human Emotions in Russell's Circumflex Model

[1].*VIIT College of Eng. Pune – 411048, India*
ORCID ID : 0000-0003-2541-5248
[2] *D. Y Patil College of Eng. Akurdi –411044, India*
ORCID ID : 0000-0003-2541-5248
[3] *Army Institute of Technology, Dighi, Pune – 411015, India*
ORCID ID : 0000-0002-0610-0374
** Corresponding Author Email: author@email.com*

In this paper, the suggested technique is found to have impressive accuracy on both datasets with any pre-trained model. This study's primary contributions can be summarised as follows:

(i) Development of an efficient FER technique based on DCNN [9], [10] models that addresses the issues via Transfer Learning which can be used in real world.

(ii) Comparison of the proposed method's emotion recognition accuracy with existing methods and exploration of the method's competency, particularly with profile views, which is vital for practical usage.

The paper is organized as follows: Section 2 briefly reviews related to the proposed work. Section 3 includes the overview of LSTM, VGG19 and Dense-Net model for better of the proposed Face Emotion Prediction. Dataset overview is given in the Section 4. In section 5, Experimentations and evaluation related to the performance of all the models is discussed. Section 6 concludes the paper with a remark and possible future research directions.

2. Literature Survey

Face acquisition and registration are critical components for recognition of facial expression. Face detection and tracking can be accomplished using the Viola-Jones method [11], [12], [13] and Subspace Constrained Mean-shifts [14], [15]. The Viola-Jones [11] method for object and face detection, which employs the Harr technique, is the most well-known method for face detection. Several models [16], [17], [18], [19], [20], [21], [22] addressed issues such as modelling time-based variation of expressions, domain knowledge, expression dimensions, and personalization of existing models. STM encourages the configuration of a generic classifier without the test subjects' external labels by attenuating individual unique mismatches [23], [24]. STM achieves this effect by simultaneously learning a classifier and re-weighting the training samples that are most relevant to the test subject. Labelled STM is an extension of STM and a special kind of Recurrent Neural Network, that improves significantly when applied to available labelled test data [25].

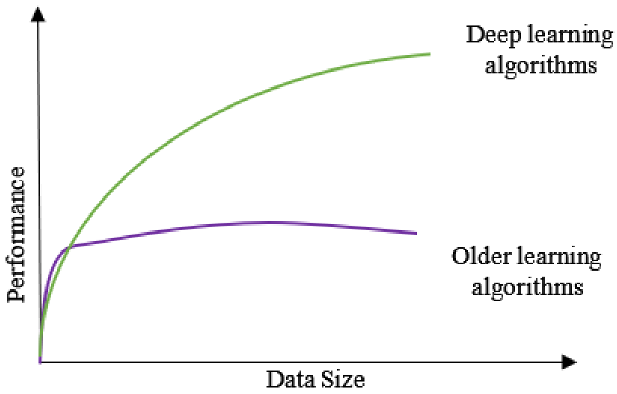

Fig. 2. The performance of deep learning regarding the size of data

Table 1. FER analysis for different algorithms

Pre-processing technique	*Feature Extraction technique*	*Classification method*	*Author*
Histogram equalization	--	CNN	[33]
Cropping, Normalization	DWT, HOG	SVM	[34]
Harr Face detection	Gabor filter and anisotropic texture, Local histogram	SVR	[35]
SQI, Sobel filter	Gabor filter	ART, DCT	[36]
Face detection	Gabor filter, Gentle boost algorithm	DWT, SVM	[37]
Face detection	Intensity normalization, Gaussian-weighted average	CNN	[38]
Haar-like features, histogram equalization	--	CNN, KNN	[39]
Viola & Jones, Intra-Face library Normalization	--	AlexNet, VGG	[40]

Hinton and Salakhtdinov, published an article [26], showcasing neural network with multiple hidden layers has better learning characteristics which was the beginning of new era 'deep learning'. By getting various degrees of abstract representation of the original data, it can increase the accuracy of prediction and categorization. According to Andrew Ng [27], the core of deep learning is that we now have fast enough computers and large data to train large neural networks.

A Deep Convolution Neural Network consists of many hidden convolutional layers and uses high-dimensional images, making input and training exceedingly difficult. Every different model holds significantly different connections and arrangement in the layers. Alex Net model trained on ImageNet dataset [28] with eight learned layers consisting of five convolution and three fully connected layers. VGG-16, VGG-19 and Conv Net [29], [30] architecture model were proposed with 13, 16 and 22 convolution layers respectively with improved accuracy level every time. DCNN [31], [32], model was then used for Face Emotion Recognition using Emotion Recognition in the Wild Challenge (EmotiW) 2015. Mostly researchers prefer CNN over deeper models for FER due to many reasons like large dataset, high resolution images required, and generalization of the model is difficult. Adding more to this the increase in no of layers are not the only solution to increase the accuracy as it leads to vanishing gradient problem.

Different techniques and algorithms are analysed for face emotion recognition in Table 1, according to the pre-processing, feature extraction techniques as well as classification methods. Histogram equalization, cropping and normalization, Harr for face detection, SQI and Sobel filters are few to mention methods used for pre-processing of a 2D FER. Gabor filter, PCA, HOG, DWT, LBP, LBP-TOP, Intra-face normalization are used with different classifier techniques like SVM, SVR, ART and DCT. Mostly KNN, CNN and Deep CNN are also used with large datasets. The overall conclusions on the bases of the papers above are due to local histogram the feature vector dimensionality lowers [36], SQI filter can overcome the insufficient light and shade light.

Designing a deep neural network for a specific task entail determining multiple configurations and parameters that can ensure the network is well suited for the task at hand. Complex models, such as CNNs, can easily overfit the data when confronted with a small dataset (CK+ and JAFFE images) based on static facial expression recognition challenge, which have mixed ie. posed and non-posed type data. To tackle the issue of training a high-capacity classifier on small datasets, previous work in this area has relied on transfer learning across tasks, in which the CNN's weights are configured with those from a network trained for related tasks before fine tuning them using the target dataset [40]. This method has consistently produced better results than directly training the network on the small dataset leading to the concept name called Transfer Learning [41], [42]. With reference to [43], it is observed that VGG-16, VGG-19 uses large no of parameters but compared to accuracy there are few algorithms using less parameters yet having good accuracy. Thus, we can say that, the no of parameters used for learning and the accuracy need not have any specific relation.

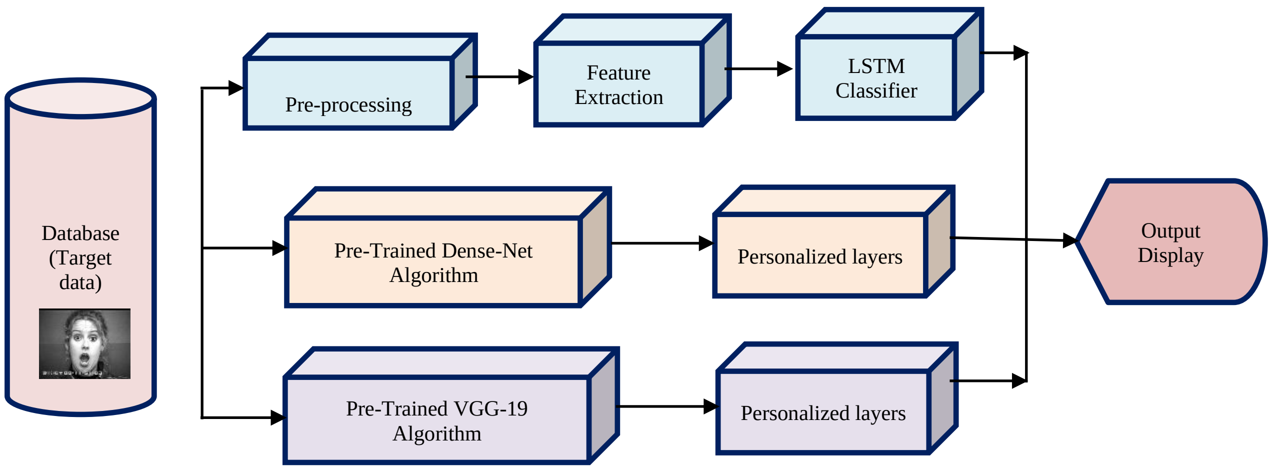

Fig. 3. Process schema for the proposed model for Face Emotion Prediction

3. Network Architectures

This research presents a FER system with the implementation of LSTM [23], [44], VGG19 [45], and Dense-Net [46] algorithms, as well as performance evaluation. VGG-19 and Dense-Net algorithms are implemented using transfer learning to reduce development efforts.

Labelled Selective Transfer Machine (L-STM) [23], uses labelled data for training where both the sampling weights and classifier parameters are jointly optimized. LSTM networks are a type of recurrent neural network that would make decisions based on learning order sequence prediction concerns. The key concept of STM [24] is to re-weight the training samples in order to form a dispersal which approximates the test dispersal.

$$S^{tr} = \{\, x_i, y_i \,\}_{i=1}^{n_{tr}} \qquad (1)$$

where, x_i = column vector of i^{th} column of the matrix X ; and $y_i \in \{+1, -1\}$

To compensate for the offset, stacking one in each data vector x_i denotes the column vector as $x_i \epsilon \mathbb{R}^{d+1}$. The STM is further formulated for minimizing the objective as

$$g(f,s) = \min_{f,s} R_f\,(S^{tr}, s) + \lambda\Omega_s(X^{tr}, X^{te}) \qquad (2)$$

where,

$R_f(S^{tr}, s)$ = On the decision function f, SVM experimental risk is specified. S^{tr} = training set in relation to time by selection coefficients $s \in \mathbb{R}^{n_{tr}}$, for training sample x_i, each s_i reacts to positive weight. $\Omega_s(X^{tr}, X^{te})$ = dispersal mismatch as feature of s. When Ω_s is smaller, training and test distribution are nearer and the trade-off between mismatch of risk and distribution is when λ >0. STM optimizes the decision feature f as well as the selective coefficient s. The linear penalized SVM in the form $f(x) = w^T x$ has the goal decision function and minimizes:

$$R_w(S^{tr}, s) = \frac{1}{2}\|w\|^2 + C\sum_{i=1}^{n_{tr}} s_i\, L^p(y_i, w^T x_i) \qquad (3)$$

where,

$$L^p(y, \cdot\,) = \max\,(0, 1 - y\,\cdot)^p \qquad (4)$$

Here, L can be any loss function as hinge loss and quadratic loss for p = 1 and 2 respectively. The second term in STM, $\Omega_s(X^{tr}, X^{te})$, copies mismatch region and the objectives are to find a re-weighting function that reduces the difference among the dispersals of train and test. Kernel mean matching is used to reweight the function, which reduces the empirical mean distance between t_r and t_e in Kernel Hilbert Space $\mathcal{H}$.

$$\Omega_s(X^{tr}, X^{te}) = \left\| \frac{1}{n_{tr}}\sum_{i=1}^{n_{tr}} s_i\varphi(x_i^{tr}) - \frac{1}{n_{te}}\sum_{j=1}^{n_{te}} \varphi(x_j^{te}) \right\|_{\mathcal{H}}^2 \qquad (5)$$

Let, $k_i := \frac{n_{tr}}{n_{te}} \sum_{j=1}^{n_{te}} k(x_i^{tr}, x_j^{te})$, where, i=1,……..,$n_{tr}$, which will collect the proximity among a training and each test sample. It is possible to accept such labels as the only reference to the target topic and to help in deciding the personalized classifier. In L-STM, the extension of STM, by adding regularization term $\lambda_L \Omega_L(S^L)$, a customized classifier is developed to adapt the target labels, where, $S^L = \{x_j^L, y_j^L\}_{j=1}^{n_L}$, as target data and their labels with, $y_j^L \in \{+1, -1\}, 0 \leq n_L \leq n_{te}$ and $\lambda_L > 0$, is used as a trade-off parameter. The objective of this additional parameter is to normalize the excellence of classification on the targeted labelled data, which is specified as, $\Omega_L(S^L) = \sum_{j=1}^{n_L} L^P\left(y_j^L, f(x_j^L)\right)$. The VGG-19 [45] Model was proposed by researchers at the Visual Geometry Group at Oxford and is known as VGG. This network was distinguished by its pyramidal structure, with the bottom layers closest to the image being wide and the top layers being deep.

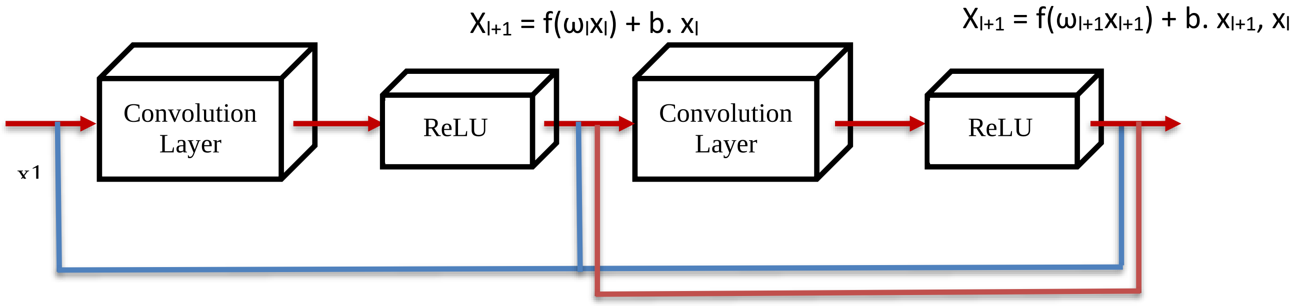

Fig. 4. Internal connections of layers in Dense Net

The 19 layers models input dimensions for VGG-19 are 224x224, ReLu activation function is used with all 3 fully connected layers and 1 Softmax layer. VGG has proven as a very good architecture for benchmarking on a particular task but is very slow to train due to the no of parameters are around 19.6 billion. The Dense-Net [46] model has a completely new connection mode that connects each current network layer to the previous network layer. This method alleviates/lightens the problem of gradient disappearance to some extent. Since each layer is connected with all previous layers, the previous features can be repeatedly used to generate more feature maps with less convolution kernel. This architecture uses hyperparameter growth rate to control the no of feature map channels in the network. The growth rate k indicates that the output feature maps of each network layer is k.

4. Dataset

The dataset for facial behaviour are created using few hardware synchronized cameras with good resolution, where the volunteers are instructed to pose a series of emotions or even the small video is displayed to capture the instant emotions [17]. Anger, happiness, fear, disgust and sadness are the 6 basic emotions whereas power, expectation, arousal, valence, contempt and relief are classified as non-basic emotions. The Extended Cohn-Kanade Dataset [47] and the Japanese Female Facial Expression Database [48] are posed and unposed datasets with data classified in six plus one basic sentiments: surprise, sad, happy, disgust, fear, angry and neutral. The CK dataset has been modified with numerous modified intensity destinations, such as frame-by-frame action unit intensity annotations and a total of 14 action units, and is now known as the CK+ [49] dataset. Spontaneous recording and original subjects, with remarks and labels for non-basic emotions, contempt, are included in CK+ dataset [49] . Japanese Female Facial Expression (JAFFE) has 7 expressions out of which 6 are basic and 1 neutral, posed by Japanese females [50]. Both the datasets differ in the distribution for different emotions, where JAFFE is equally distributed whereas CK+ is unevenly distributed for all different emotions. The CK+ and JAFFE database complies of more than 2000 image sequences from over 200 subjects ranging from 18 to 30 years together.

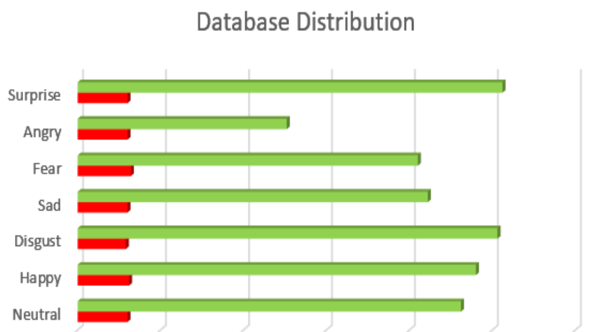

Fig. 5. Dataset Distribution for different emotions for extended CK and JAFFE datasets

5. Performance Evaluation

The classifier algorithms were evaluated for its facial emotion recognition on the JAFFE [50] and CK+ [47]. The dataset includes African-American, Japanese as well as Asian or Latino students where the distribution of emotions are unequal in CK+. Digitalized images from neutral to target in the dimension of 640 by 480-pixel arrays are arrayed with grayscale values of 8-bit. The database also includes color images and some images vary in size of 240 by 240. The performance of all the classifiers were evaluated on OpenCV platform on

(i) Core i5 9th generation 16GB memory device
(ii) 4xNVIDIA Professional Series Quadro P6000 RTX PCle 3.0-24GB

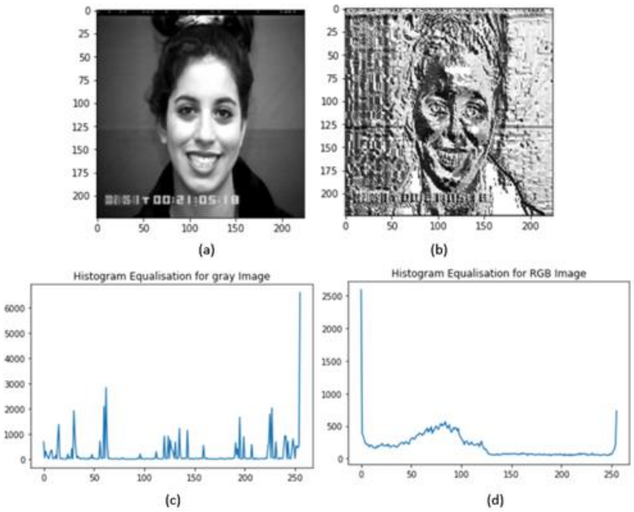

Fig. 6. LBP applied on images, (a) RGB, (b) Gray, Histogram equalization on (a) and (b) in figure (c and (d) respectively

The experimental paradigm was randomly divided into 7 groups of approximately equal size and cross-validated by 'leaving one group out' using the 7-fold cross validation method. The total sample set is divided into two sets: testing and training, with the training and test sets used to fit the models and evaluate the error of the chosen model.

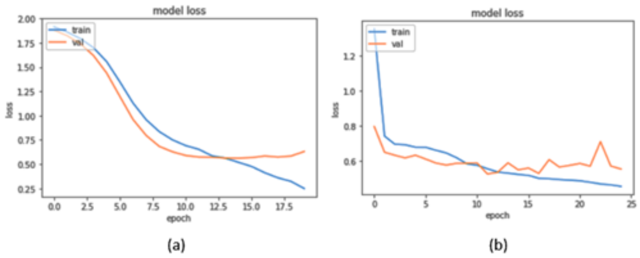

Fig. 7. Learning Curve on (a) JAFFE; (b) CK+, for LBP with LSTM

The validation set is used to assess prediction error for the model where validation score is calculated which should be high and loss value should be less. The loss function, evaluates the algorithm over the dataset. If prediction is less, loss function will be a high value and vice-versa.

Table 2. FER analysis for different algorithms

Methodology	Dataset	No of iterations	No of epochs	Accuracy
LBP with LSTM	JAFFE	10	20	98.43
	CK+	10	20	72.63

For the implementation of the LSTM model, the input image is pre-processed using the Viola Jones algorithm for face detection and reshaping the detected faces to a definite size. Detected faces are applied LBP for feature extraction and histogram to scale all with 256 features. The recognition accuracy, significantly improved with LSTM using 27495 learning parameters. Total learning loss of 0.2479 and 0.4536 for JAFFE and CK+ dataset respectively, the overall accuracy of 98.43 and 76.59 as well as validation accuracy of 67.06 and 72.12 respectively.

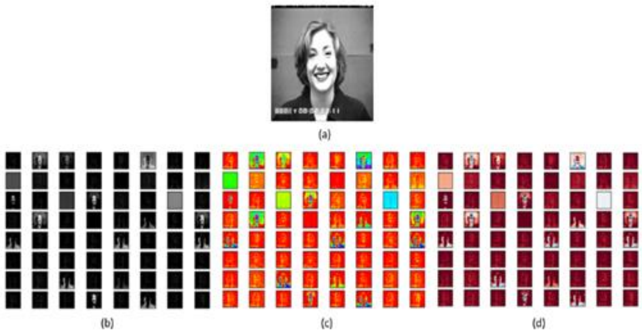

Fig. 8. Layer 1 models filter input images in three different color space as (a) Gray, (b) HSV and (c) RdBu image

The fine-tuning method is used for VGG19 and Dense-Net implementation, where some trained layers are freeze and a few layers are trained with personalised dataset. The VGG19 Original model used with 20 million parameters and other 6 million were added by modified layers. The integration of red, green, and blue intensities encoded in each colour plan at the pixel's location decides the colour of each pixel [51]. The HSV (Hue-Saturation-Value) colour space is more closely related to how humans perceive colours can be easy visualized in figure 8 and 9, where the filter layers for different colour space shows the features that are used by each layer for training.

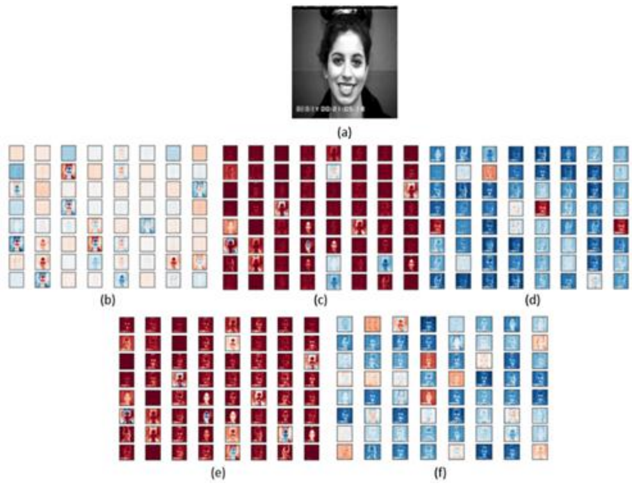

Fig. 9. Layer filter images for layer 2, 5, 9, 13, 17 related to Dense-Net model

The total parameters for Dense-Net 201 Model were almost 42 million, out of which 24 million parameters were trainable parameter and 18 million were non-trainable parameters. In the figure it is observed that the current layer can take the output feature maps of all the previous layers as input features.

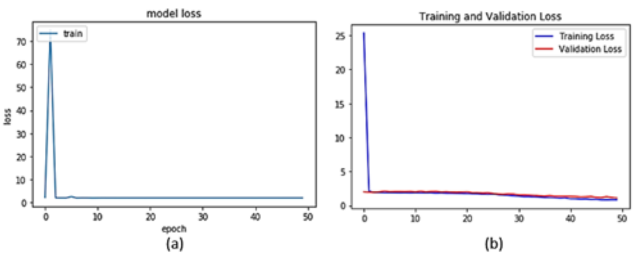

Fig. 10. Training and validation loss on CK+ for VGG19 Model

The Transfer learning model for Face emotion detection was built on VGG19 and Dens-Net with fixed iterations and epochs based on the dataset. The iteration of the JAFFE dataset was set to 20 with 15 epochs, while the iteration of the CK+ dataset was set to 256 with 50 epochs.

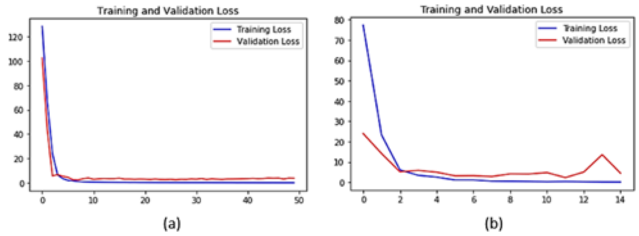

Fig. 11. Results on CK+ and JAFFE for Dense-Net201 model respectively

The loss curve exhibits a good learning rate, and the model behavior for the CK+ dataset exhibits well-fitting learning curves. However, the generalization gap is too small in this case. The model behaviour on both the dataset was good and shows good fit learning curve with good generalization gap. The learning loss of VGG16 was 1.9457 and 0.83 for the JAFFE and CK+ datasets, respectively, while it was 0.1228 and 1.1521 for Dense-Net.

Table 3. Performance Evaluation of Classifier Algorithms (Accuracy) on JAFFE and CK+ datasets

Methodology	Dataset	No of iterations	No of epochs	Accuracy	AUC ROC	Precision
VGG19	JAFFE	20	15	13.33	50.18	17.11
	CK+	256	50	68.4	78.1	75.64
Dense-Net	JAFFE	20	15	97.45	85.91	67.28
	CK+	256	50	96.64	92.99	84.46

The CPU training time for VGG19 was 22 hours, while the GPU training time was 6 hours. The CPU training time for Dense-Net was 26 hours, while the GPU training time was 7.5 hrs. The average accuracy of 97.05% on both different dataset of the Dense-Net model is pretty well compared to VGG19.

Table 4. Evaluation of our approaches (LBP+LSTM and Dense-Net for transfer learning), Gabor, LBP, LDP, DTP, PBP+NBP with SVM classifier

Methodology	CK+ Recognition rate %	JAFFE Recognition rate %
Gabor + SVM [52]	78.9 ± 4.8	75.5 ± 5.8
LBP + SVM [53]	79.1 ± 4.6	77.2 ± 7.6
LDP + SVM [54]	82.6 ± 4.1	86.9 ± 2.8
DTP + SVM [55], [56]	94.58 ± 1.80	84.09 ± 7.13
PBP + NBP + SVM [55], [56]	95.58 ± 1.42	87.77 ± 7.15
LBP + LSTM (our approach)	72.63 ± 0.51	98.43 ± 0.24
Dense Net (our approach for TL)	96.64 ± 0.15	97.45 ± 0.12

Based on the comprehensive study of different techniques [55], we used different approaches of feature extraction like Gabor, LBP, LDP, DTP, PBP and NBP with SVM as classifiers for evaluating the performance of one of our approaches which is based on LBP technique for feature extraction and LSTM for classification. The performance of all the approaches is done on CK+ dataset as well as JAFFE to avoid biased evaluation. The overall performance of LBP+LSTM approach outrated the recognition rate on JAFFE but seems to be low performer on CK+ as the dataset is not evenly divided. Other approach used for comparison was on transfer learning technique approach using Dense Net, which completely outperformed all the approaches on both the dataset with significant recognition rate.

6. Conclusion

In this paper, we evaluated the performance of facial emotion prediction using LSTM as well as two distinct deep learning models, VGG19 and Dense-Net on two different datasets CK+ and JAFFE. The model's performance can be generalised based on the training parameters, time required, accuracy and precision. A pre-processing and feature extraction method based on Viola-Jones and histogram for face detection and Local Binary Pattern (LBP) is used for LSTM. LSTM with 27495 learning parameters and 3 layers, VGG19 with 26 million parameters and 19 layers where as Dense-Net with 42 million parameters with 201 layers were used for experimentation. On both types of datasets, the accuracy level of the Dense-Net model (96.64% and 97.45%) appears to outperform the LSTM (98.43% and 72.63%) and VGG19 algorithm. The effect of unbalanced dataset is almost negligible on Dense-Net model. When compared to the CPU, the dedicated GPU significantly reduced the training time. The researchers have a good option after generalising the model for face emotion prediction using the transfer learning technique.

Acknowledgements

The authors would like to thank the research centre (D Y Patil college, Pune), and affiliated colleges (Vishwakarma Institute of Information Technology and Army Institute of Technology) for all the support. Special thanks to Army Institute of Technology for the GPU provided for our research.

Author contributions

Prof. Dipti P. Pandit received her Bachelor and Masters of Engineering from University of Pune (India), and is pursuing PhD in Electronics and Telecommunication Engineering. Presently she is working as Assistant Professor at Vishwakarma Institute of Information Technology, Pune. She has 16 years of teaching experience and 2 years of industrial experience. She has published 5 papers in reputed national and international conference/referred Journals. Her research interest are signal and image processing, embedded systems and deep machine learning.

Dr. (Prof) Sangeeta Jadhav has received BE (Electronics), ME (Electronics & Telecommunication) from Govt College of Engineering, Pune (India) and PhD (Electronics) from Government Engineering College, Aurangabad, Maharashtra, (India). She is in the field of Engineering Education for last 26 years. Currently she is working as Head of Information Technology at Army Institute of Technology, Pune, Maharashtra (India). Her research interests are in the area of image and signal processing with special focus on Blind signal processing and independent component analysis. She has more than 15 publications to her credit published at various international Journals, conferences and seminars.

Conflicts of interest

The authors declare no conflicts of interest.

References

[1]. M. Altamura et al., "Facial emotion recognition in bipolar disorder and healthy aging," J. Nerv. Ment. Dis., vol. 204, no. 3, pp. 188–193, 2016, doi: 10.1097/NMD.0000000000000453.

[2]. L. Ricciardi et al., "Facial emotion recognition and expression in Parkinson's disease: An emotional mirror mechanism?," PLoS One, vol. 12, no. 1, pp. 1–16, 2017, doi: 10.1371/journal.pone.0169110.

[3]. M. B. Harms, A. Martin, and G. L. Wallace, "Facial emotion recognition in autism spectrum disorders: A review of behavioral and neuroimaging studies," Neuropsychol. Rev., vol. 20, no. 3, pp. 290–322, 2010, doi: 10.1007/s11065-010-9138-6.

[4]. C. F. Benitez-Quiroz, R. B. Wilbur, and A. M. Martinez, "The not face: A grammaticalization of facial expressions of emotion," Cognition, vol. 150, pp. 77–84, May 2016, doi: 10.1016/j.cognition.2016.02.004.

[5]. P. Ekman, "Universal facial expressions of emotion," California Mental Health, vol. 8, no. 4. pp. 151–158, 1970.

[6]. J. A. Russell, "A circumplex model of affect," J. Pers. Soc. Psychol., vol. 39, no. 6, pp. 1161–1178, 1980, doi: 10.1037/h0077714.

[7]. Z. Lou, F. Alnajar, J. M. Alvarez, N. Hu, and T. Gevers, "Expression-Invariant Age Estimation Using Structured Learning," IEEE Trans. Pattern Anal. Mach. Intell., vol. 40, no. 2, pp. 365–375, Feb. 2018, doi: 10.1109/TPAMI.2017.2679739.

[8]. M. Shreve, "Automatic Macro- and Micro-Facial Expression Spotting and Applications," Grad. Theses Diss., no. January, 2013.

[9]. J. Zhou, Y. Wang Zhenan Sun, Y. Xu Linlin Shen, J. Feng Shiguang Shan, Y. Qiao Zhenhua Guo, and S. Yu, "Biometric Recognition." Online. . Available: http://www.springer.com/series/7412.

[10]. Sarma, C. A. ., S. . Inthiyaz, B. T. P. . Madhav, and P. S. . Lakshmi. "An Inter Digital- Poison Ivy Leaf Shaped Filtenna With Multiple Defects in Ground for S-Band Bandwidth Applications". International Journal on Recent and Innovation Trends in Computing and Communication, vol. 10, no. 8, Aug. 2022, pp. 55-66, doi:10.17762/ijritcc.v10i8.5668.

[11]. C. Zhang and Z. Zhang, "Improving Multiview Face Detection with Multi-Task Deep Convolutional Neural Networks."

[12]. W. M. Fang and P. Aarabi, "Robust real-time audiovisual face detection," Multisensor, Multisource Inf. Fusion Archit. Algorithms, Appl. 2004, vol. 5434, p. 411, 2004, doi: 10.1117/12.545934.

[13]. J. Nicolle, V. Rapp, K. Bailly, L. Prevost, and M. Chetouani, "Robust continuous prediction of human emotions using multiscale dynamic cues," in ICMI'12 - Proceedings of the ACM International Conference on Multimodal Interaction, 2012, pp. 501–508, doi: 10.1145/2388676.2388783.

[14]. S. K. A. Kamarol, M. H. Jaward, J. Parkkinen, and R. Parthiban, "Spatiotemporal feature extraction for facial expression recognition," IET Image Process., vol. 10, no. 7, pp. 534–541, Jul. 2016, doi: 10.1049/iet-ipr.2015.0519.

[15]. J. M. Saragih, S. Lucey, and J. F. Cohn, "Face alignment through subspace constrained mean-shifts," Proc. IEEE Int. Conf. Comput. Vis., no. Clm, pp. 1034–1041, 2009, doi: 10.1109/ICCV.2009.5459377.

[16]. X. Xiong and F. De La Torre, "Supervised descent method and its applications to face alignment," Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit., pp. 532–539, 2013, doi: 10.1109/CVPR.2013.75.

[17]. L. Zhang and D. Tjondronegoro, "Facial expression recognition using facial movement features," IEEE Trans. Affect. Comput., vol. 2, no. 4, pp. 219–229, 2011, doi: 10.1109/T-AFFC.2011.13.

[18]. X. Ding, W. S. Chu, F. D. La Torre, J. F. Cohn, and Q. Wang, "Facial action unit event detection by cascade of tasks," Proc. IEEE Int. Conf. Comput. Vis., pp. 2400–2407, 2013, doi: 10.1109/ICCV.2013.298.

[19]. A. C. Cruz, B. Bhanu, and N. S. Thakoor, "Vision and attention theory based sampling for continuous facial emotion recognition," IEEE Trans. Affect. Comput., vol. 5, no. 4, pp. 418–431, Oct. 2014, doi: 10.1109/TAFFC.2014.2316151.

[20]. O. Çeliktutan, S. Ulukaya, and B. Sankur, "A comparative study of face landmarking techniques," Eurasip J. Image Video Process., vol. 2013, no. 1, pp. 1–27, 2013, doi: 10.1186/1687-5281-2013-13.

[21]. M. Yang and L. Zhang, "Gabor feature based sparse representation for face recognition with Gabor occlusion dictionary," Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics), vol. 6316 LNCS, no. PART 6, pp. 448–461, 2010, doi: 10.1007/978-3-642-15567-3_33.

[22]. Y. li Tian, T. Kanade, and J. F. Cohn, "Recognizing upper face action units for facial expression analysis," Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit., vol. 1, no. 2, pp. 294–301, 2000, doi: 10.1109/cvpr.2000.855832.

[23]. C. Science, A. Yinka, A. Ali, D. A. R, and A. Jemai, “Information Security June 2011 International Journal of Computer Science & Information Security,” no. June, 2011.

[24]. W. S. Chu, F. De La Torre, and J. F. Cohn, “Selective transfer machine for personalized facial expression analysis,” IEEE Trans. Pattern Anal. Mach. Intell., vol. 39, no. 3, pp. 529–545, Mar. 2017, doi: 10.1109/TPAMI.2016.2547397.

[25]. M. Liu, S. Shan, R. Wang, and X. Chen, “Learning expressionlets on spatio-temporal manifold for dynamic facial expression recognition,” in Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition, Sep. 2014, pp. 1749–1756, doi: 10.1109/CVPR.2014.226.

[26]. S. Hizlisoy, S. Yildirim, and Z. Tufekci, “Music emotion recognition using convolutional long short term memory deep neural networks,” Eng. Sci. Technol. an Int. J., vol. 24, no. 3, pp. 760–767, 2021, doi: 10.1016/j.jestch.2020.10.009.

[27]. A. J. Holden et al., “Reducing the Dimensionality of,” vol. 313, no. July, pp. 504–507, 2006.

[28]. X. Han and Q. Du, “Research on Face Recognition Based on Deep Learning.”

[29]. T. F. Gonzalez, “Handbook of approximation algorithms and metaheuristics,” Handb. Approx. Algorithms Metaheuristics, pp. 1–1432, 2007, doi: 10.1201/9781420010749.

[30]. K. Simonyan and A. Zisserman, “Very deep convolutional networks for large-scale image recognition,” 3rd Int. Conf. Learn. Represent. ICLR 2015 - Conf. Track Proc., pp. 1–14, 2015.

[31]. S. Rani and P. Kumar, “Deep Learning Based Sentiment Analysis Using Convolution Neural Network,” Arab. J. Sci. Eng., vol. 44, no. 4, pp. 3305–3314, 2019, doi: 10.1007/s13369-018-3500-z.

[32]. Z. Yu and C. Zhang, “Image based static facial expression recognition with multiple deep network learning,” in ICMI 2015 - Proceedings of the 2015 ACM International Conference on Multimodal Interaction, Nov. 2015, pp. 435–442, doi: 10.1145/2818346.2830595.

[33]. A. Fernández-Caballero et al., “Smart environment architecture for emotion detection and regulation,” J. Biomed. Inform., vol. 64, pp. 55–73, Dec. 2016, doi: 10.1016/j.jbi.2016.09.015.

[34]. K. Clawson, L. S. Delicato, and C. Bowerman, “Human centric facial expression recognition,” Proc. 32nd Int. BCS Hum. Comput. Interact. Conf. HCI 2018, pp. 1–12, 2018, doi: 10.14236/ewic/HCI2018.44.

[35]. André Sanches Fonseca Sobrinho. (2020). An Embedded Systems Remote Course. Journal of Online Engineering Education, 11(2), 01–07. Retrieved from http://onlineengineeringeducation.com/index.php/joee/article/view/39

[36]. S. Nigam, R. Singh, and A. K. Misra, “Efficient facial expression recognition using histogram of oriented gradients in wavelet domain,” Multimed. Tools Appl., vol. 77, no. 21, pp. 28725–28747, 2018, doi: 10.1007/s11042-018-6040-3.

[37]. Institute of Electrical and Electronics Engineers and IEEE Signal Processing Society, 2013 IEEE International Conference on Image Processing : ICIP 2013 : proceedings : September 15-18, 2013, Melbourne, Victoria, Australia. .

[38]. 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, Jul. 2018, doi: 10.1007/s00500-017-2634-3.

[39]. Q. Wu, X. Shen, and X. Fu, “The machine knows what you are hiding: An automatic micro-expression recognition system,” in Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics), 2011, vol. 6975 LNCS, no. PART 2, pp. 152–162, doi: 10.1007/978-3-642-24571-8_16.

[40]. A. T. Lopes, E. de Aguiar, A. F. De Souza, and T. Oliveira-Santos, “Facial expression recognition with Convolutional Neural Networks: Coping with few data and the training sample order,” Pattern Recognit., vol. 61, pp. 610–628, 2017, doi: 10.1016/j.patcog.2016.07.026.

[41]. L. Bacon et al., Proceedings, 2017 15th IEEE/ACIS International Conference on Software Engineering Research, Management and Applications (SERA) : June 7-9, 2017, the University of Greenwich, London, UK. .

[42]. H. W. Ng, V. D. Nguyen, V. Vonikakis, and S. Winkler, “Deep learning for emotion recognition on small datasets using transfer learning,” in ICMI 2015 - Proceedings of the 2015 ACM International Conference on Multimodal Interaction, Nov. 2015, pp. 443–449, doi: 10.1145/2818346.2830593.

[43]. C. C. Atabansi, T. Chen, R. Cao, and X. Xu, “Transfer Learning Technique with VGG-16 for Near-Infrared Facial Expression Recognition,” J. Phys. Conf. Ser., vol. 1873, no. 1, 2021, doi: 10.1088/1742-6596/1873/1/012033.

[44]. L. D. Nguyen, D. Lin, Z. Lin, and J. Cao, “Deep CNNs for microscopic image classification by exploiting transfer learning and feature concatenation,” Proc. - IEEE Int. Symp. Circuits Syst., vol. 2018-May, no. June, 2018, doi: 10.1109/ISCAS.2018.8351550.

[45]. O. Russakovsky et al., “ImageNet Large Scale Visual Recognition Challenge,” Int. J. Comput. Vis., vol. 115, no. 3, pp. 211–252, 2015, doi: 10.1007/s11263-015-0816-y.

[46]. A. Alsayat, “Improving Sentiment Analysis for Social Media Applications Using an Ensemble Deep Learning Language Model,” Arab. J. Sci. Eng., 2021, doi: 10.1007/s13369-021-06227-w.

[47]. K. He, X. Zhang, S. Ren, and J. Sun, “Deep residual learning for image recognition,” Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit., vol. 2016-Decem, pp. 770–778, 2016, doi: 10.1109/CVPR.2016.90.

[48]. Garg, D. K. . (2022). Understanding the Purpose of Object Detection, Models to Detect Objects, Application Use and Benefits. International Journal on Future Revolution in Computer Science &Amp; Communication Engineering, 8(2), 01–04. https://doi.org/10.17762/ijfrcsce.v8i2.206646. W. Wang, Y. Li, T. Zou, X. Wang, J. You, and Y. Luo, “A novel image classification approach via dense-mobilenet models,” Mob. Inf. Syst., vol. 2020, 2020, doi: 10.1155/2020/7602384.

[49]. Dursun, M., & Goker, N. (2022). Evaluation of Project Management Methodologies Success Factors Using Fuzzy Cognitive Map Method: Waterfall, Agile, And Lean Six Sigma Cases. International Journal of Intelligent Systems and Applications in Engineering, 10(1), 35–43. https://doi.org/10.18201/ijisae.2022.265

[50]. P. Lucey et al., “IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops,” IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit. Work., vol. 4, no. July, pp. 94–101, 2003.

[51]. M. Kamachi, M. Lyons, and J. Gyoba, “The japanese female facial expression (jaffe) database,” URL http//www. kasrl. org/jaffe. html, vol. 21, no. January, p. 32, 1998.

[52]. M. Bartlett, J. Hager, P. Ekman, T. S.- Psychophysiology, and U. 1999, “Bartlett et al._1999_Psychophysiology_Measuring facial expressions by computer image analysis.pdf,” Cambridge.Org. Online. . Available: https://www.cambridge.org/core/journals/psychophysiology/article/measuring-facial-expressions-by-computer-image-analysis/B5A746609CF2198C688A0F28114CDF60.

[53]. P. S. P. Wang, “Performance comparisons of facial expression recognition in jaffe database,” vol. 22, no. 3, pp. 445–459, 2008.

[54]. M. O. Dwairi, Z. Alqadi, and R. A. Zneit, “Optimized True-Color Image Processing Optimized True-Color Image Processing,” no. August 2016, 2010.

[55]. M. S. Bartlett, G. Littlewort, M. Frank, C. Lainscsek, I. Fasel, and J. Movellan, “2005. Bartlett et al - REcognizing Facial

expression - machine learning and application."
[56]. C. Shan, S. Gong, and P. W. McOwan, "Facial expression recognition based on Local Binary Patterns: A comprehensive study," Image Vis. Comput., vol. 27, no. 6, pp. 803–816, 2009, doi: 10.1016/j.imavis.2008.08.005.
[57]. T. Jabid, M. H. Kabir, and O. Chae, "Robust facial expression recognition based on local directional pattern," ETRI J., vol. 32, no. 5, pp. 784–794, 2010, doi: 10.4218/etrij.10.1510.0132.
[58]. S. Tivatansakul, M. Ohkura, S. Puangpontip, and T. Achalakul, "Emotional healthcare system: Emotion detection by facial expressions using Japanese database," 2014 6th Comput. Sci. Electron. Eng. Conf. CEEC 2014 - Conf. Proc., pp. 41–46, 2014, doi: 10.1109/CEEC.2014.6958552.
[59]. S. TIVATANSAKUL and M. OHKURA, "Healthcare System Focusing on Emotional Aspect using Augmented Reality," Trans. Japan Soc. Kansei Eng., vol. 13, no. 1, pp. 191–201, 2014, doi: 10.5057/jjske.13.191.