

e-ISSN: 2319-8753 | p-ISSN: 2347-6710

International Journal of Innovative Research in SCIENCE | ENGINEERING | TECHNOLOGY

INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN SCIENCE | ENGINEERING | TECHNOLOGY

Volume 10, Special Issue 2, December 2021

INTERNATIONAL STANDARD SERIAL NUMBER INDIA

Impact Factor: 7.569

9940 572 462

6381 907 438

🛿 ijirset@gmail.com





|| e-ISSN: 2319-8753, p-ISSN: 2320-6710| www.ijirset.com | Impact Factor: 7.569| || Volume 10, Special Issue 2, December 2021 ||

4th International Conference on Emerging Trends in Science, Technology and Mathematics [ICETSTM '2021]

15th & 16th September 2021

Organized by

Department of Computer Science, Parvathy's Arts and Science College, Dindigul, Tamilnadu, India

Comparison of JAFFE and CK+ Database Facial Expression Recognition using Deep Learning

Dr.N.Krishnaveni¹

HOD, Dept. of Computer Science, Annai Women's College, Karur, Tamilnadu, India¹

ABSTRACT: Face represent large range of information about Identity,age,sex,race as well as emotional and mental state. Facial expression play crucial role in social interactions and commonly used in the behavioral interpretation of emotions. The biological characteristics of the face have overall structural similarity and individual local difference. Automatic facial expression recognition is one of the interesting and challenging problem in computer vision due to its potential applications such as Human Computer Interaction (HCI), behavioral science, video games etc. In this paper, a novel method for automatically recognizing facial expressions using Deep Convolution Neural Network(DCNN) features is proposed. Here we proposed a DCNN model and investigate it for the facial expression recognition by using both JAFFE, CK⁺ dataset to evaluate the performance. The proposed model focuses on recognizing the facial expression based on Deep learning in the field of biometric combined with face recognition technology of an individual from a single image.

KEYWORDS: Deep Convolution Neural Network, JAFF, Confusion matrix, Support Vector machine-

I. INTRODUCTION

Facial expression is an major a part of gesture communication. Human expression recognition is influenced by certain context. When a subject is being investigated, the analyst could be diverted by the subject's voice tone or argument and should forget to keep track of the facial expression. Automatic facial expression recognition systems are exempt to such contextual interference. This systems are often beneficial in many fields, like gaming applications, criminal interrogations, psychiatry, animations etc. State-of-art approaches plan to recognize six basic facial expressions such as anger, disgust happiness, sadness, surprise and fear.

Facial expression recognition techniques are divided into geometry features or appearance features. Geometric features are extracted from the shape of the face and its elements like as the eyebrows, the mouth, the nose etc. Appearance features are extracted using the texture of the face caused by expression, like furrows, wrinkles etc. In 1970s Paul Ekman and Wallace V. Friesen, developed Facial Action Coding System (FACS) which is the most widely used method for describing and measuring facial behaviors. FACS is a method designed for human observers to represent changes in facial expression in terms of observable skeletal muscle actions referred to as facial Action Units or AUs. FACS is a powerful system for detecting and measuring facial expressions and currently used for feature extraction together with other techniques like Dynamic Bayesian Network (DBN) and Local Binary Pattern (LBP), Histograms of oriented gradients(HOG), Scale Invariant Feature Transform (SIFT), Local Binary Pattern (LBP) are few state-of-art techniques for extracting facial features. These techniques use handcrafted features for facial expression recognition, and they require particular efforts both in terms of computation cost and programming effort.

In recent years, deep learning using convolution neural networks(CNNs) for feature extraction of image data is becoming more popular. Their popularity stems from their ability to extract good depiction from image data. DCNN's computation intensive tasks can run on GPU, which results in high performance at very low power consumption. They have also yielded high performance for some of challenges such as the CNN based model proposed by Kim et al. CNN is extensively used for facial feature extraction for determining age, gender etc.



|| e-ISSN: 2319-8753, p-ISSN: 2320-6710| www.ijirset.com | Impact Factor: 7.569| || Volume 10, Special Issue 2, December 2021 ||

4th International Conference on Emerging Trends in Science, Technology and Mathematics [ICETSTM '2021]

15th & 16th September 2021

Organized by

Department of Computer Science, Parvathy's Arts and Science College, Dindigul, Tamilnadu, India

II. RELATED WORK

Several methods are reported within the literature to naturally recognize facial expressions. Lucey et al. ¹¹ manually labeled 68 facial points in key frames and used a gradient descent Active Appearance Model (AAM) to fit these points within the remaining frames. It is not possible to get accurate key points in many practical situations. A study on using LBP for facial expression recognition is proposed by Shan et al. ¹². Here expression recognition is accomplished using

support vector machine (SVM) classifiers with boosted-LBP features. Authors manually labeled eye positions, which isn't

feasible in many practical cases. Computer Expression Recognition Toolbox (CERT) is proposed by Littlewort et al. ¹³. CERT fold through the registered face image with Gabor filters to extract the facial features and uses SVM and multivariate logistic regression (MLR) classifiers to recognize facial expressions. Lyons and Akamatsu ¹⁴ proposed a system for coding facial expressions with 2D Gabor wavelets for feature extraction, having clustering for classification.

system for coding facial expressions with 2D Gabor wavelets for feature extraction, having clustering for classification. These methods also require particular efforts, both in terms of computation cost and programming effort.

Deep Convolution Neural Network (DCNN) framework is widely used for extraction of features from the images. DCNN uses several layers leading accurate feature learning. Here the prelearned features are used as filters and these filters convolves through the input image and produces the features which in turn are used by other layers of the network as

discussed by Krizhevsky et al. ¹⁵. Convolutional neural networks techniques have been proposed by Kahou et al. ¹⁶. But they greatly train the model with other facial dataset. Sebastien¹⁷ used Deep Convolutional Activation Feature for Generic Visual Recognition (DeCAF)¹⁸ for facial feature extraction that does not require substantial training, but DeCAF is too slow to use it for training even the small image dataset as it does not support GPU.

III. METHODOLOGY

The flowchart of the proposed method is as shown in the figure 1.



3.1 Facial Expression Database

Facial expression plays a crucial role in human communication because 55% of facial modality involve in human-to-human interaction. Two publicly database set utilized in this work JAFFE and CK⁺

3.1.1 JAFFE Database

Japanese Female Facial Expression (JAFFE) database was taken from publicly available data which consists of 213 facial expression images of 10 subjects of Japanese female. Each subject performed six basic emotions plus neutral (33 fear,31 sad,30 surprise,30 happiness,30 angry,30 neutral,29 disgust) during which expression contains 3 to 4 images per subjects. The image has the resolution in grayscale. All the facial images are been taken under strict controlled conditions of comparable lighting and no occlusion like hair or glasses. All the expression in frontal view and therefore the resolution of the first image are 256 X 256 pixels.

3.1.2 CK⁺ Database

The Expended Cohn-kanade(CK^+) dataset contains 593 video sequences from a total of 123 different subjects starting from 18 to 50 years of aged with a spread of genders and heritage. Each video shows a facial shift from the neutral expression to a targeted peak expression recorded at 30 Frames Per Second(FPS) with a resolution of either 640 X 490 or

IJIRSET © 2021



|| e-ISSN: 2319-8753, p-ISSN: 2320-6710| www.ijirset.com | Impact Factor: 7.569| || Volume 10, Special Issue 2, December 2021 ||

4th International Conference on Emerging Trends in Science, Technology and Mathematics [ICETSTM '2021]

15th & 16th September 2021

Organized by

Department of Computer Science, Parvathy's Arts and Science College, Dindigul, Tamilnadu, India

640 X480 pixels. Out of those video, 327 are labeled with one among expression classes fear, sad, surprise, happiness, angry, neutral and disgust. The CK^+ database is widely considered the foremost extensively used laboratory controlled facial expression classification database available and is employed within the majority of facial expression classification methods.

3.2 Image Preprocessing

For the aim of comparisons we crop the first image of size 256×256 into 168×120 by removing the background influences. Since the illumination condition is varied to the images in JAFFE database. We apply histogram equalization to eliminating lightning effects.

Pre-processing work must be done a little more before Identification. Two stage step in our image preprocessing is employed to eliminate disturbance within the original images namely face detection and histogram equalization.

3.3 Feature Extraction using Caffe

Feature extraction is performed using Caffe on Graphics Processing Unit (GPU). The convolution neural network architecture,

which is used for ImageNet 15 object detection is used to extract facial features. ImageNet uses eight learned layers which incorporates five convolutional and three fully connected layers for object detection. In this work, features are extracted using only first five layers. These layers are combination of convolution, Rectified Linear Units(RELU), Local Response Normalization(LRN) and pooling operations. The operations are repeated, to make different layers . The features extracted after pooling operation at fifth layer (POOL5) are used for recognizing facial expressions. This is the layer where features are most visible and also the POOL5 feature length($6 \ge 0 \le 256$) is feasible to use in classifiers to recognize facial expressions.

Algorithm 1 Algorithm for recognizing facial expressions using DCNN procedure RECOGNIZE-FACIAL-EXPRESSIONS for all image (i), depicting facial expressions face dataset do convert the image (i) to gray-scale detect frontal face in (i) and crop only the face (c) extract POOL5(256 x 6 x 6) features for cropped face (e) using DCNN copy predefined facial expression label for each image(i) as per the input dataset use the resulting *POOL5* vector of dimension §216(256 x 6 x 6) for tenfold and leave –one-out cross validation with SVM classifier to recognize facial expression

The first layer is the convolution layer. This model uses 96 filters with size $11 \times 11 \times 3$. This layer extracts the low-level edge features. Figure 2 shows the sample output after first convolution filters being applied on a face image.

The second layer is ReLU(Rectified Linear Units). This layer improves the nonlinear properties of network. For any given input value x, ReLU is defined by (f(x) = max(0, x)). Pooling is the next layer. In pooling layer, small rectangular blocks from the previous layer are taken and subsampled to produce a single output from that block. In this model, max-pooling is used and 3 x 3 rectangular blocks with an interval of 2 pixels are used for max-pooling. Figure 3a shows the facial features before performing the max-pooling



Fig. 2: Output of first convolution layer applied on face image



|| e-ISSN: 2319-8753, p-ISSN: 2320-6710| www.ijirset.com | Impact Factor: 7.569| || Volume 10, Special Issue 2, December 2021 ||

4th International Conference on Emerging Trends in Science, Technology and Mathematics [ICETSTM '2021]

15th & 16th September 2021

Organized by

Department of Computer Science, Parvathy's Arts and Science College, Dindigul, Tamilnadu, India

·····································			200	and the second	a se		4
					14 m		
						i de ste	
(all	- 	16 (b)			1 23		

(a)

(b)

Fig. 3: Extracted facial features (a) before applying max-pooling; (b) after applying max-pooling

and Figure 3b depicts the facial features after performing max-pooling. It can been seen that after pooling, facial edges are retained even though the dimension of the image is reduced.

The fourth layer is LRN (Local Response Normalization). This is the brightness normalization layer given by $(1 + (\alpha/n) i xi2\beta)$, where α and β are tuning parameters with the default values 1 and 5 respectively. This layer normalizes the brightness, so that the relevant features are more visible and irrelevant features are reduced.

Once the features are extracted, SVM is used for classification. Best kernel for SVM is estimated using grid search estimators²². Grid search exhaustively considers all parameter combinations on a dataset and the best combination is retained. Here 50% of the dataset is used for training and 50% for testing. The accuracy is evaluated and the estimator provided good accuracy for SVM linear kernel over other kernels. So SVM one-vs-one (SVC) and one-vs- all (LinerSVC) classifiers are used to recognize facial expressions. In SVM classification, parameter C controls the trade off between errors

on training data and margin maximization (C = ∞ leads to hard^omargin)²³. Leave-one-out and tenfold cross validation methods are used to estimate the performance. In case of tenfold cross validation, the dataset is divided into ten sets. One set is used for testing and remaining sets are used for training the model. Confusion matrix is used to visualize the performance of a classifier as it identifies the nature of the classification errors, as well as their quantities. The diagonal elements in the confusion matrix represent the number of correctly recognized expressions for each class. The off-diagonal elements represent the unrecognized expressions or the recognition error. Given the confusion matrix, the accuracy is

calculated using Equation 1^{24} . Accuracy is also measured using F1 score (F measure) using precision and recall as given by Equation as given by Equation 2. Python machine learning tools are used to implement the proposed model





|| e-ISSN: 2319-8753, p-ISSN: 2320-6710| <u>www.ijirset.com</u> | Impact Factor: 7.569| || Volume 10, Special Issue 2, December 2021 || 4th International Conference on Emerging Trends in Science, Technology and Mathematics [ICETSTM '2021]

15th & 16th September 2021

Organized by

Department of Computer Science, Parvathy's Arts and Science College, Dindigul, Tamilnadu, India



Fig. 4: Confusion matrix for CK+ dataset

IV. EXPERIMENTAL RESULTS

4.1 Cross Validation for CK+ Dataset

Leave-one-subject-out cross validation strategy is utilized to test the model. Figure 4a shows confusion lattice for best SVM classification with the C worth equivalent to |e|. The numbers in the confusion matrix addresses the quantity of images for a specific facial expression. For instance in Figure 4a, out of 18 contempt facial expression samples, 15 are properly identified as contempt and 3 are wrongly identified as anger ,happiness and surprise. Altogether, the facial expressions are correctly recognized for 314 pictures out of 327, with an exactness of 96.02%. The standardized confusion matrix framework is displayed in Figure 4b. It very well may be seen that the vast majority of the articulations are perceived effectively with the exception of the groupings that contain pitiful articulations. This may be on the grounds that the examples accessible for bitterness articulation are less and not many of the successions for instance S131.003, S026.002 and so forth are not clear to be perceived as sad expressions.

Since few of the state-of-art literature use only six universal expressions excluding contempt, the recognition is performed using six universal expressions (without contempt). The confusion matrix as displayed in Figure 5a portrays the best planning of valid and anticipated expressions with C=10.0 for SVM one-versus all classifier. The precision of facial expressions recognition for the proposed technique for six universal expressions is 97%. Figure 5b gives the standardized normalized confusion matrix to six universal expressions. This is 13% more precise than the gauge results $83.33\%^{11}$.







|| e-ISSN: 2319-8753, p-ISSN: 2320-6710| <u>www.ijirset.com</u> | Impact Factor: 7.569| || Volume 10, Special Issue 2, December 2021 || 4th International Conference on Emerging Trends in Science, Technology and Mathematics [ICETSTM '2021]

15th & 16th September 2021

Organized by

Department of Computer Science, Parvathy's Arts and Science College, Dindigul, Tamilnadu, India

4.2 Cross Validation for JAFFE Dataset

All 213 facial expression that are accessible with the JAFFE dataset were utilized to assess the proposed technique. The pattern similitude rank utilizing Gabor ²⁰ filter is 0.679. From the baseline literature ²⁰, it was not satisfactory in regards to the exactness of the model, so assessment of the proposed model is contrasted and the model proposed by Lyons et. al ¹⁴. They removed the components utilizing 2D Gabor wavelet portrayal and performed ten-fold cross approval which came about 92.00% acknowledgment exactness. The proposed model with DCNN elements and ten times cross approval brings about 98.12% exactness. Here the dataset is divided into ten folds utilizing Python's StratifiedKFold mama chine learning instrument by setting the shufflng parameter. Figure 6a portrays the disarray lattice for JAFFE dataset. The outcomes show that out of 213 looks pictures, the proposed model effectively perceives articulations for 209 images with an accuracy of 98.12%. Figure 6b portrays the normalized confusion matrix . It very well may be seen that the exactness for the vast majority of expressions is close to 100%. Leave-one-example out cross approval is likewise performed and the no change in accuracy is found, i.e the outcomes show a similar disarray lattice as displayed in Figure 6a.



Fig. 6: Confusion matrix for JAFFE dataset

Figure 7a portrays the Precision-Recall bend for the after effects of ten times cross approval. From the graph, it can be seen that acknowledgment precision is 100% for fear and neutral expressions. There is slight deviation for bliss articulation, as two of the articulations having a place with trouble and shock are wrongly classified as joy. Figure 7b provides the details of the facial expression, for which the prediction mismatches with the relating marks gave in the dataset. It very well may be seen that the facial expression match more with anticipated classes instead of the comparing real names gave in the dataset.

In data handling, F1 score (F Measure) is a proportion of classifier's exactness, as it thinks about both accuracy and review. Table 1 summarizes F1 score for the proposed strategy. Correlation of state-or-craftsmanship acknowledgment precision with the proposed model is displayed in Table 2. It tends to be seen that proposed model gives the best acknowledgment precision to JAFFE. For CK+, better outcomes were gotten by other condition of-craftsmanship writing. Techniques proposed by Siddiqi et al. 25 and Mlakar et al. 26 outcomes in better execution when impartial pictures are utilized alongside seven essential expressions, yet it tends to be noticed that main few subjects (10, 106) were utilized for cross validation. For confirming the proposed strategy for CK+ dataset, every one of the 118 subjects were considered with 327 arrangements without neutral expression. The proposed strategy around requires 140-145ms to extract the facial components for a picture on Tesla K20Xm GPU, which is significantly less contrasted with include extraction strategies that utilizes just CPU (700-900 ms).



|| e-ISSN: 2319-8753, p-ISSN: 2320-6710| www.ijirset.com | Impact Factor: 7.569| || Volume 10, Special Issue 2, December 2021 ||

4th International Conference on Emerging Trends in Science, Technology and Mathematics [ICETSTM '2021]

15th & 16th September 2021

Organized by

Department of Computer Science, Parvathy's Arts and Science College, Dindigul, Tamilnadu, India



Fig. 7: (a) Precision-Recall curve for JAFFE dataset; (b) Misclassified facial expressions for JAFFE dataset Table 1: Overall accuracy of the proposed method

Dataset	Cross validation method	F1 Score	
C.V.	Leave-one-subject- out(six classes)	0.9708	
CK+	Leave-one-subject- out(seven classes)	0.96024	
	Ten Fold	0.9812	
JAFFE	Leave-one-sample- out	0.9812	

Table 2: Comparison of recognition rate obtained by proposed model with state-of-art literature

Dataset	Cross validation method(Proposed Model)	Accuracy
<u>OV</u>	Six classes	97.08
CK+	Seven classes	96.02
	Ten Fold	98.12
JAFFE	Leave-one-sample- out	98.12

An application has been created on Unix stage utilizing Python AI apparatuses and Caffe. This application perceives the looks from either a picture or video succession. Alongside JAFFE, CK+ look datasets, pictures from Google search are utilized to prepare the framework utilizing the above proposed model. Around 150 pictures are utilized for every look class to prepare the model. The yield of the classifier is saved and utilized for testing the acknowledgment of looks for obscure appearances. For a video succession the look for each edge is perceived and result is shown utilizing a visual chart. Figure 8 shows yield for three facial expressions namely happy, sadness and surprise.



|| e-ISSN: 2319-8753, p-ISSN: 2320-6710| www.ijirset.com | Impact Factor: 7.569| || Volume 10, Special Issue 2, December 2021 ||

4th International Conference on Emerging Trends in Science, Technology and Mathematics [ICETSTM '2021]

15th & 16th September 2021

Organized by

Department of Computer Science, Parvathy's Arts and Science College, Dindigul, Tamilnadu, India



Fig. 8: Output of automatic facial expressions recognition module

V. CONCLUSION

Facial expression convey the emotion of a person to the spectators. A proficient and quicker technique to perceive the looks is proposed in this paper. The facial provisions are separated utilizing profound convolutional neural organization utilizing Caffee on CUDA empowered GPU framework. The proposed strategy is assessed on two freely accessible datasets and cutting edge results are accomplished. Since GPU based Caffe module is utilized to lead the analysis, the time needed to extricate an elements is significantly decreased. The proposed model can be embraced to any conventional looks acknowledgment dataset that either includes acknowledgment in static pictures or video groupings. No retraining or broad pre-handling strategies are needed to embrace the proposed strategy for facial component extraction. The future work includes investigating other DCNN pre-prepared models like GoogLeNet .

REFERENCES

[1] Tian, Y.L., Kanade, T., Cohn, J.F.. Handbook of Face Recognition; chap. Facial Expression Analysis. New York, NY: Springer New York. ISBN 978-0-387-27257-3; 2005, p. 247–275. URL: http://dx.doi.org/10.1007/0-387-27257-7_12. doi:10.1007/0-387-27257-7_12.

[2] Ekman, P., Friesen, W. Facial Action Coding System: A Technique for the Measurement of Facial Movement. Palo Alto: Consulting

Psychologists Press; 1978.

[3] Ko, K.E., Sim, K.B.. Emotion recognition in facial image sequences using a combination of aam with facs and dbn. In: *Proceedings of the Third International Conference on Intelligent Robotics and Applications - Volume Part I*; ICIRA'10. Berlin, Heidelberg: Springer-Verlag. ISBN 3-642-16583-4, 978-3-642-16583-2; 2010, p. 702–712. URL: http://dl.acm.org/citation.cfm?id=1939443.1939518.

[4] Wang, L., Li, R.F., Wang, K., Chen, J.. Feature representation for facial expression recognition based on facs and lbp. *International*

Journal of Automation and Computing 2015;**11**(5):459–468. URL: http://dx.doi.org/10.1007/s11633-014-0835-0. doi:10.1007/s11633-014-0835-0.

[5] Dalal, N., Triggs, B., Histograms of oriented gradients for human detection. In: Computer Vision and Pattern Recognition, 2005. CVPR

2005. IEEE Computer Society Conference on; vol. 1. 2005, p. 886–893 vol. 1. doi:10.1109/CVPR.2005.177.

[6] Lowe, D.G.. Object recognition from local scale-invariant features. In: *Computer Vision, 1999. The Proceedings of the Seventh IEEE International Conference on*; vol. 2. 1999, p. 1150–1157 vol.2. doi:10.1109/ICCV.1999.790410.

[7] Zhao, G., Pietikainen, M.. Dynamic texture recognition using local binary patterns with an application to facial expressions. *IEEE*

Transactions on Pattern Analysis and Machine Intelligence 2007;29(6):915–928. doi:10.1109/TPAMI.2007.1110.

[8] Kim, B.K., Roh, J., Dong, S.Y., Lee, S.Y.. Hierarchical committee of deep convolutional neural networks for robust facial expression recognition. *Journal on Multimodal User Interfaces* 2016;:1–17URL: http://dx.doi.org/10.1007/s12193-015-0209-0. doi:10.1007/s12193-015-0209-0.

[9] Wang, X., Guo, R., Kambhamettu, C.. Deeply-learned feature for age estimation. In: *Applications of Computer Vision* (WACV), 2015 IEEE

Winter Conference on. 2015, p. 534–541. doi:10.1109/WACV.2015.77.

[10] Levi, G., Hassneer, T.. Age and gender classification using convolutional neural networks. In: Computer Vision and Pattern Recognition Workshops (CVPRW), 2015 IEEE Conference on. 2015, p. 34-42.



|| e-ISSN: 2319-8753, p-ISSN: 2320-6710| www.ijirset.com | Impact Factor: 7.569| || Volume 10, Special Issue 2, December 2021 ||

4th International Conference on Emerging Trends in Science, Technology and Mathematics [ICETSTM '2021]

15th & 16th September 2021

Organized by

Department of Computer Science, Parvathy's Arts and Science College, Dindigul, Tamilnadu, India

doi:10.1109/CVPRW.2015.7301352.

[11] Lucey, P., Cohn, J.F., Kanade, T., Saragih, J., Ambadar, Z., Matthews, I.: The Extended Cohn-Kanade Dataset (CK+): A complete dataset

for action unit and emotion-specified expression. In: *Computer Vision and Pattern Recognition Workshops (CVPRW)*, 2010 IEEE Computer Society Conference on. IEEE. ISBN 978-1-4244-7029-7; 2010, p. 94–101. URL: http://dx.doi.org/10.1109/cvprw.2010.5543262. doi:10.1109/cvprw.2010.5543262.

[12] Shan, C., Gong, S., McOwan, P.W.. Facial expression recognition based on local binary patterns: A comprehensive study. *Image Vision*

Comput 2009;**27**(6):803–816. URL: http://dx.doi.org/10.1016/j.imavis.2008.08.005. doi:10.1016/j.imavis.2008.08.005.

[13] Littlewort, G., Whitehill, J., Wu, T., Fasel, I., Frank, M., Movellan, J., et al. The computer expression recognition toolbox (cert). In: *Automatic Face Gesture Recognition and Workshops (FG 2011), 2011 IEEE International Conference on.* 2011, p. 298–305. doi:10.1109/ FG.2011.5771414.

[14] Lyons, M.J., Budynek, J., Akamatsu, S.. Automatic classification of single facial images. *IEEE Trans Pattern Anal Mach Intell* 1999;

21(12):1357–1362. URL: http://dx.doi.org/10.1109/34.817413. doi:10.1109/34.817413.

[15] Krizhevsky, A., Sutskever, I., Hinton, G.E.. Imagenet classification with deep convolutional neural networks. In: Pereira, F., Burges, C.,

Bottou, L., Weinberger, K., editors. *Advances in Neural Information Processing Systems* 25. Curran Associates, Inc.; 2012, p. 1097–1105.

URL: http://papers.nips.cc/paper/4824-imagenet-classification-with-deep-convolutionalneural-networks. pdf.

[16] Kahou, S.E., Froumenty, P., Pal, C.. Facial Expression Analysis Based on High Dimensional Binary Features. Cham: Springer International

Publishing. ISBN 978-3-319-16181-5; 2015, p. 135–147. URL: http://dx.doi.org/10.1007/978-3-319-16181-5_10. doi:10.1007/978-3-319-16181-5_10.

[17] Ouellet, S.. Real-time emotion recognition for gaming using deep convolutional network features. *CoRR* 2014;**abs/1408.3750**. URL:

http://arxiv.org/abs/1408.3750.

[18] Donahue, J., Jia, Y., Vinyals, O., Hoffman, J., Zhang, N., Tzeng, E., et al. Decaf: A deep convolutional activation feature for generic visual recognition. *CoRR* 2013;**abs/1310.1531**. URL: http://arxiv.org/abs/1310.1531.

[19] Jia, Y., Shelhamer, E., Donahue, J., Karayev, S., Long, J., Girshick, R., et al. Caffe: Convolutional architecture for fast feature embedding.

arXiv preprint arXiv:14085093 2014;.

[20] Lyons, M., Akamatsu, S., Kamachi, M., Gyoba, J.. Coding facial expressions with gabor wavelets. In: *Proceedings of the 3rd. International Conference on Face & Gesture Recognition*; FG '98. Washington, DC, USA: IEEE Computer Society. ISBN 0-8186-8344-9; 1998, p. 200–. URL: http://dl.acm.org/citation.cfm?id=520809.796143. [21] Bradski, G., The OpenCV Library. *Dr Dobb's Journal of Software Tools* 2000;.





Impact Factor: 7.569





INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN SCIENCE | ENGINEERING | TECHNOLOGY





www.ijirset.com