iJRASET
International Journal For Research in Applied Science and Engineering Technology



INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 12 **Issue:** III **Month of publication:** March 2024

DOI: https://doi.org/10.22214/ijraset.2024.59189

www.ijraset.com

Call: 08813907089 | E-mail ID: ijraset@gmail.com

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 12 Issue III Mar 2024- Available at www.ijraset.com

Facial Sentimental Analysis in Hospitals

Ms. M. Kamala[1], A.S. Srinath[2], Pranav Kumar[3], Arun Teja[4]
[1]*Associate Prof., Computer Science & Engineering CMR College of Engineering & Technology, Hyderabad, India*
[2, 3, 4]*Computer Science & Engineering CMR College of Engineering & Technology,* Hyderabad, India

Abstract: ***In today's rapidly evolving healthcare landscape, the focus on patient satisfaction and well-being has become crucial. A positive experience plays an important role in aiding the recovery of the patient. A significant factor contributing to a positive patient experience is the quality care and the emotional support they receive during their hospital stay. Traditional methods of measuring patient satisfaction often rely on surveys and feedback forms, which can be subjective, time-consuming, and potentially biased. To address these limitations, this abstract proposes a novel approach: the integration of facial sentiment analysis as a real-time tool for assessing patient emotions and satisfaction levels. This abstract introduces an approach to enhance patient satisfaction and well-being in healthcare settings. By integrating facial sentiment analysis, we propose a real-time method to objectively evaluate patient emotions and treatment satisfaction. Through advanced computer vision, this approach detects and interprets facial expressions of patients in hospitals, offering insights into care effectiveness, dedication of hospital staff, and efficiency of the treatment. The system generates a "Happiness and Satisfaction Index," aiding healthcare stakeholders in making informed improvements as well and it works as a review metric for the customers who would like to visit their hospital.***

Keywords: Facial Sentiment Recognition, CNN, Deep Learning, Satisfaction Index, Hospital welfare.

I. INTRODUCTION

Facial expressions are indicators of emotions, providing insightinto a person's feelings. They serve as a tangible means of discerning honesty, as they convey emotions without verbal communication. Emotions are conveyed through various channels such as voice, hand movements, body language, and prominently, facial expressions. These expression plays an important role in interpersonal communication, transmitting nonverbal cues that influence relationships. Automatic recognition of our facial expressions is crucial for natural human- machine interactions and behavioral studies. While humans effortlessly recognize facial expressions, machines struggle with expression recognition. Despite advancements in face detection, feature extraction mechanisms, and classification techniques, developing automated systems for this task remains challenging. Mostly two approaches are commonly used in facial expression detection: explicit classification, which separates expressions, and the use of extracted facial features to aid in recognition. The Facial ActionCoding System (FACS) utilizes action units as markers for facial expression analysis.

Our facial expressions can express emotions more effectivelyin situations where words may fall short, such as moments ofshock or surprise. Furthermore, detecting deception through spoken language is generally more challenging than recognizing fake expressions. Non-verbal forms of communication, including facial expressions, account for approximately 66% of all communication. This highlights theimportance of studying, utilizing, and analyzing facial expressions. Analyzing facial expressions can be applied to various scenarios, such as a system that detects a driver's drowsiness in a car, triggering an alarm and stopping the vehicle if drowsiness is detected.

II. DATASETS

We explored five datasets suitable for training our models,each offering unique characteristics:

1) *FER 2013:* This dataset, originating from the ICML 2013 Workshop on Challenges in Representation Learning, comprises 35,887 48x48 images with six basic emotion categories: Anger, Disgust, Fear,Happiness, Neutral, Sadnes. It relies on images sourced through the Google image search API.
2) *AFEW:* Unlike many facial expression datasetsgenerated in controlled lab settings, AFEW features short video clips capturing spontaneous facial expressions from movies. It includes the same seven basic expressions as FER-2013 but also reflects natural head poses, a diverse range of races and genders, and often includes multiple individuals in a single clip. The dataset encompasses both indoor and outdoor scenes, filmed during various lighting conditions, offering a broad spectrum of expressions for algorithm training and generalization.

©IJRASET: All Rights are Reserved

ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 12 Issue III Mar 2024- Available at www.ijraset.com

III. RELATED WORK

1) AlexNet, developed by Alex et al., revolutionized the field of Image Classification by winning the ImageNet ILSVRCS-2012 competition with a top-5 error rate of 15.3%. This Deep CNN (DCNN) was trained on the Imagenet ILSVRCS training Dataset and boasts 60 million parameters. The architecture of AlexNet includes 5 Convo Layers along with 3 fully connected Dense layers. Rectified Linear Units follow the Convolutional Layers to introduce non-linearity, while Max Pooling Layers helps with reducing the number of parameters and combat overfitting. The fully connected Dense layers are supplemented by dropout layers, where neurons are randomly set to zero with a dropout coefficient of 0.5. This technique reduces overfit and the adaptation of neurons. Additionally, AlexNet employs data augmentation to artificially expand the dataset, further enhancing its performance in image classification tasks.

2) Xception - Xception is a CNN architecture that stands out for its depth and efficiency. Introduced by François Chollet in 2017, Xception takes the concept of depth-wise separable convolutions to the next level, offering a highly efficient alternative to traditional convolutional layers. Unlike traditional convolutional layers, which apply a single filter to the entire input volume, Xception's depth-wise separable convolutions split the process into two stages: a depth-wise convolution, which applies a single filter to each input channel separately, followed by a point-wise convolution, which combines the output of the depth-wise convolution using 1x1 filters. This separation of spatial and channel-wise operations allows Xception to reduce the no. of parameters and computational costs significantly compared to other architectures, such as ResNet or VGGNet, while maintaining competitive performance. This efficiency makes Xception particularly well-suited for applications with limited computational resources, such as mobile and embedded devices.

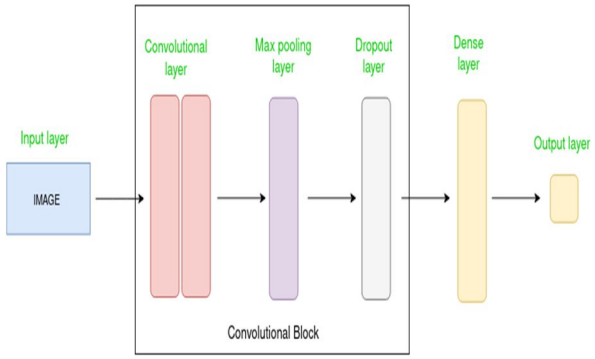


Fig. 1. Convolutional Block

3) ResNet, with its 152 layers, is significantly deeper than VGGNet, being 8 times deeper. It achieved a remarkable top-5 error rate of 3.57% on the ImageNet Dataset. The researchers observed that as they increased the number of layers by adding Convolutional Layers, the accuracy did not decrease but instead plateaued. This phenomenon occurred because the network became too complex, making it challenging to optimize using backpropagation. To address this issue, the authors have introduced ResNet Blocks, as illustrated in the figure below. These blocks utilize a shortcut connection, represented by H(x) = F(x) + x, where an identity function is added to the input. This shortcut aids in optimizing the network by providing a direct path for the gradient to flow through, allowing the addition of numerous layers without hindering performance. This capability enables the network to learn intricate features, making it highly effective in tasks such as Object Detection and Image Segmentation, where complex features are crucial.

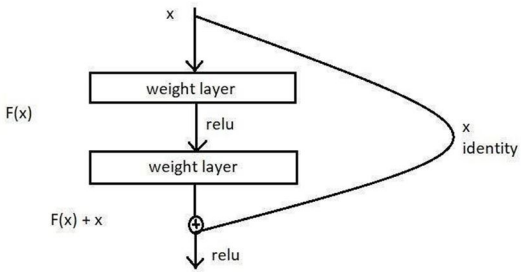


Fig. 2. ResNet layer

©IJRASET: All Rights are Reserved | SJ Impact Factor 7.538 | ISRA Journal Impact Factor 7.894 |

ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 12 Issue III Mar 2024- Available at www.ijraset.com

IV. DEEP LEARNING

In the Zhiding[16] study, the authors began by identifying faces in the images, which were sourced from labeledmovie images. Due to significant background noise, detecting theface was prioritized before employing a Deep Convolutional Neural Network (CNN) to determine the emotion. To achieve this, an ensemble approach was employed, combining multiple face recognition algorithms. This ensemble included a joint cascade detect and alignment (JDA) detector, a Deep-CNN-based detector, and Mixture of Trees (MOT). The network was initiallypretrained on the FER dataset and then fine-tuned using the SFEW dataset.

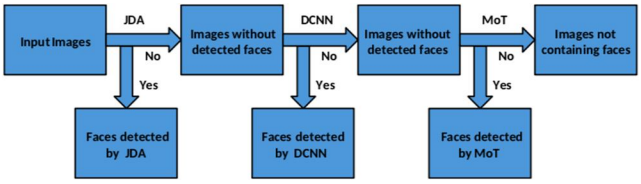


Fig. 3. Detecting faces in an image [20]

The CNN architecture utilized for emotion detection, depicted in Figure 8, operates on grayscale images with dimensions of48x48 pixels. It comprises five of the convolutional layers alongwith the three stochastic pooling layers, and three fully connected Dense layers.

In contrast to traditional max-pooling layers, stochastic poolingis employed due to its superior accuracy in scenarios with limited training data. While max pooling selects the maximum value within a region, stochastic pooling introduces randomness by probabilistically selecting values based on a normalized probability distribution. This randomization enhances the network's performance by introducing variability. Furthermore, the fully connected layers incorporate dropout, a technique that introduces statistical randomness by randomlydropping a fraction of connections during training. This randomization aids in the network's generalization capabilities and mitigate overfitting the training data. Random data augmentation to input faces is used to generates additional training samples and improves model performance.

In Octavio et al's[16] work, a convolutional neuralnetwork(CNN) was implemented in real-time on a robot platform with limited hardware capabilities to classify emotionsand gender. Two CNN models were proposed, focusing onachieving a balance between accuracy and model size.

Our first model is a fully convolutional neural network without any fully connected layers. It comprises 9 convolutional layers, utilizes ReLU activation functions, employs Batch Normalization to expedite learning, and incorporates Global Max Pooling before the application of softmax layer. This modelcontains an estimated 600,000 parameters. Omitting fully connected Dense layers helps avoid the parameter-heavy natureof such layers, thereby keeping the model size in check. The final convolutional layer replaces the fully connected Dense layers, ensuring the same number of feature maps as classes. An average pooling layer further reduces parameters, followed by asoftmax layer for output prediction.

Our second model is based on the Xception architecture, whichuses residual modules along with depth-wise separable convolutions. Depth-wise separable convolutions consist of depth-wise convolutions and point-wise convolutions. Depth- wise convolutions apply a D × D filter to each input channel individually, while point-wise convolutions use a 1 × 1 convolution to reduce the no. of channels from M to N (M > N),reducing the overall no. of parameters. This model also utilizes four residual depth-wise separable convolutions, along withbatch normalization & ReLU activation functions. The models achieved an accuracy of 96% on the IMDB gender dataset and 66% on the FER-2013 emotion dataset. The authors utilized guided-gradient backpropagation, as proposed by Springenberg,to visualize the features learned by CNN.

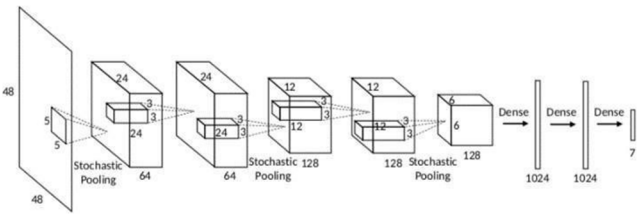


Fig. 4. Deep CNN architecture

©IJRASET: All Rights are Reserved

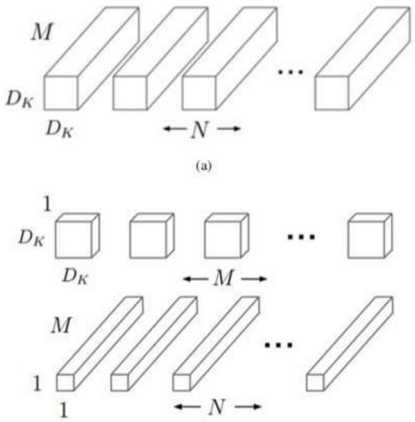


Fig. 5. Depth-wise convolution

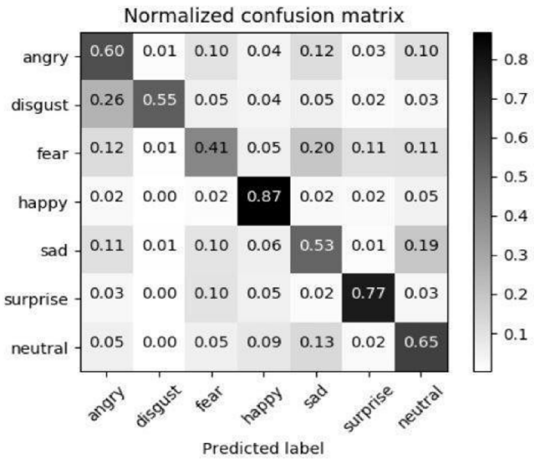


Fig. 6. Results & Analysis.

V. CONCLUSION

We have discussed the data collection and computational methods employed in Facial Emotion Recognition. It's advisable to source images from real-life settings to ensure that they adequately represent the environment, facilitating effective algorithm deployment. Acquiring additional training instances, particularly for less common emotions such as Disgust, is essential. Convolutional Neural Networks (CNNs) demonstrate superior performance compared to traditional Machine Learning techniques like Support Vector Machines (SVM) and other algorithms. A recommended approach involves initially detecting faces and subsequently training a CNN on the resulting output, yielding satisfactory results.

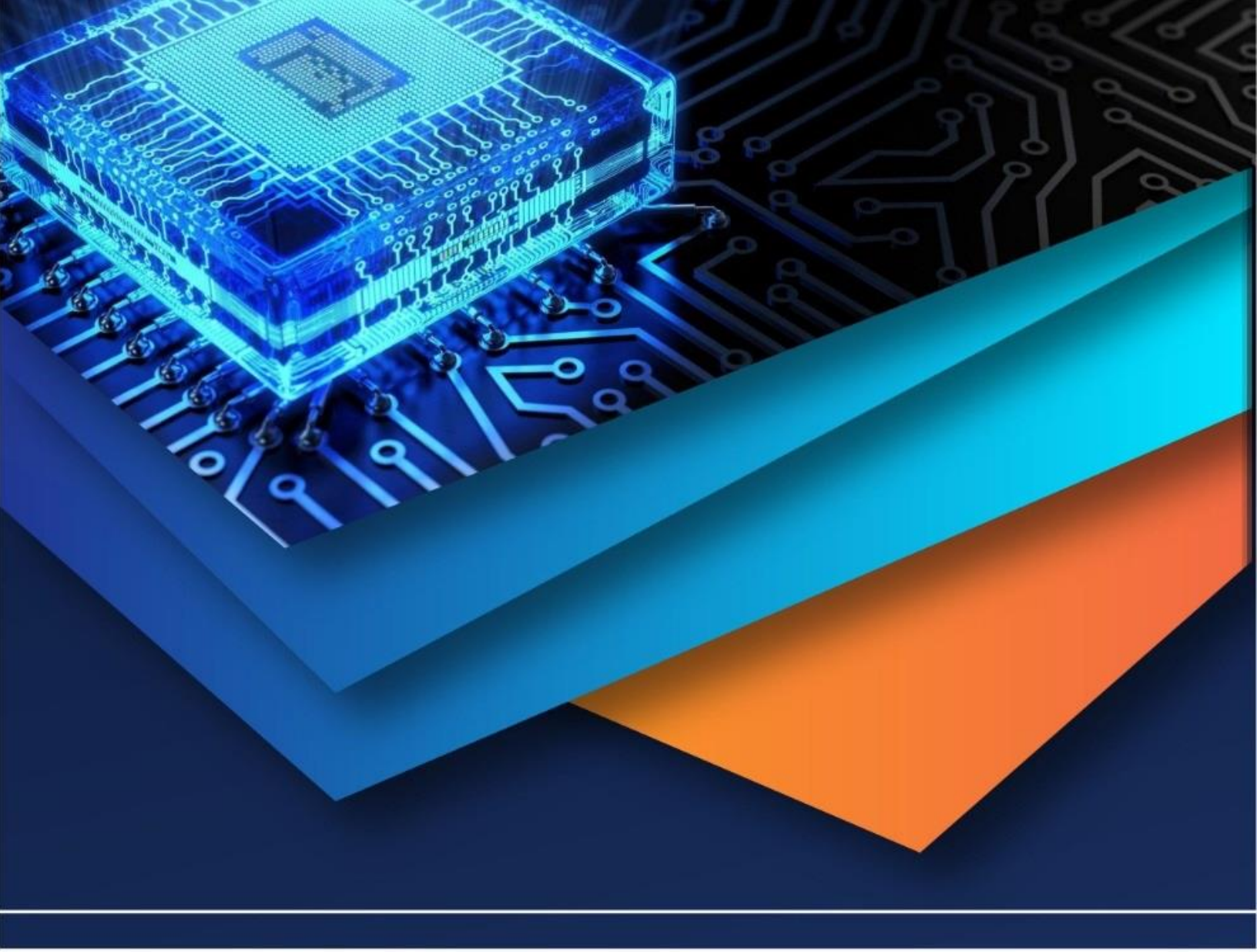
VI. ACKNOWLEDGMENT

We would like to thank all our critics for their valuable feedback on our paper, and our project guide, along with theproject faculty of CMR College of Engineering & Technology for their unconditional support.

REFERENCES

[1] Ahmed, A. R., Khaled, M., and Yusof, R. (2002). Machine Learning Using SVM of the 2002 Malaysian Science & Technology Congress. 19-21 September 2002. Johor Bahru, Malaysia. 1-8.
[2] P. Michal and R. E Kaliouby, "Facial Expression Recognition With Support Vector Machines," 2000.
[3] M.S. Bartlett, Gwen Littlewort, M. Frank, C. Lainscsek, I. Fasel, and J.Movellan, "Recognizing Facial Expression: Machine Learning and Application to Spontaneous Behavior," Proc. IEEE Int'l Conf. CV and Pattern Recognition (CVPR '05), pp. 568-573, 2005.
[4] P. S. Aleksic and A. K. Katsaggelos, "Automatic facial expression recognition using facial animation parameters and multistream hmms," TIFS, vol. 1, no. 1, pp. 3–11, 2006.
[5] P. Lucey, J. F. Cohn, T. Kanade, J. Saragih, Z. Ambadar, and I. Matthews, "The extended cohn-Kanade dataset (ck+): A complete dataset for action unit and emotion-specified expression," in Computer Vision and Pattern Recognition Workshops (CVPRW), 2010 IEEE Computer Society Conference on. IEEE, 2010,

pp. 94– 101.
[6] M. Lyons, S. Akamatsu, M. Kamachi, and J. Gyoba, "Coding facial expressions with Gabor wavelets," in Automatic Face and Gesture Recognition, 1998. Proceedings. Third IEEE International Conference
[7] I. J. Goodfellow, D. Erhan, Carrier, A. Courville, M. Mirza, B. Hamner, W. Cukierski, Y. Tang, D. Thaler, D.-H. Lee et al., "Challenges in representation learning: A report on three machine learning contests," in International Conference on Neural Information Processing. Springer, 2013, pp. 117–124.
[8] A. Dhall, R. Geecke, S. Lucey, T. Gedeon, et al., "Collecting large, richly annotated facial-expressions databases from movies," IEEE Multimedia, vol. 19, no. 3, pp. 34–41, 2012.
[9] C. F. Benitez-Quiroz, R. Srinivas, and A. M. Martine, "Emotionet: An accurate, method for the automatic annotation of a million facial expressions in the wild," in Proceedings of IEEE International Conference on Computer Vision & Pattern Recognition (CVPR), Las Vegas, NV, USA, 2016.
[10] S. Du, Y. Tao, and A. M. Martinez. Compound facial expressions of emotion. Proceedings of the National Academy of Sciences, 111(15): E1454–E1462, 2014.
[11] Alex Krizhevsky, Sutskever I, and Hinton G.E, Imagenet classificationwith deep convolutional neural networks. In NIPS, 2012
[12] J. Dang, W. Dong, R. Socher, Li, K. Li, and L. Fei-Fei. ImageNet: A Large-Scale Image Database
[13] V. Nar and G. E. Henton. Rectified linear units improve Boltzmann machines. In Proc. 27th International Conference on Machine Learning, 2010.
[14] C. Segedy, W. Liuh, Y. Jiah, P. Sermanet, S. Reed, Going Deeper witht h e convolutions, CoRR, 2014.
[15] K. He, X. Zhag, S. Ran, J. Sun, Deep learning for image recognition, Computer Vision and the Pattern Recognition (CVPR), 2016, pp. 770–778.
[16] Zhiding Yu, Cha Zhang Image-based Facial Expression Classification with Multiple Network Learning, 2015



10.22214/IJRASET

45.98

IMPACT FACTOR: 7.129

IMPACT FACTOR: 7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Call : 08813907089 (24*7 Support on Whatsapp)