Real-Time Facial Expression Recognition Based on CNN

R. Deepthi Crestose Rebekah, Pawar Bhavya, Avula Vasanthi,
Arava Mamatha Sumanjali and Valchannagari Vyshnavi
Department of CSE, Ravindra College of Engineering for Women, Kurnool, Andhra Pradesh, India

Keywords: Real-Time Recognition, Facial Expression Analysis, Convolutional Neural Networks (CNNs), Deep Learning, Feature Dimensionality Reduction, Principal Component Analysis (PCA), Bayesian Optimization, Computational Efficiency.

Abstract: Facial expression recognition through CNN in real-time is a deep learning technique that is intended to recognize and classify facial emotions instantly. The accurate recognition, though, is difficult to attain because of factors such as changes in lighting, occlusions, and prominent facial features. Furthermore, high computational requirements create problems, particularly for resource-constrained devices. It is essential to balance accuracy with efficiency for maximum performance. In response to these issues, the suggested system includes PCA for dimension reduction and Bayesian optimization for model adjustment. PCA boosts computational efficiency through data dimensions reduction, while Bayesian optimization tunes hyperparameters to increase accuracy. The system surpasses current models through increased precision, reduced processing time, and minimized computational complexity, ultimately providing an efficient real-time facial expression recognition solution.

1 INTRODUCTION

Facial expression recognition (FER) is an important area of human-computer interaction, enabling machines to recognize and react to human feelings. Real-time FER with the help of Convolution Neural Networks (CNNs) has become popular because it can automatically detect and classify facial expressions with minimal intervention. Yet, high accuracy in real-time applications still a challenging task because of facial expression variations, lighting, occlusions, and personal facial variations. Additionally, the high computational complexity of CNN models complicates the achievement of an optimal balance between processing time and accuracy, especially on low-resource devices.

A better real-time FER system that is intended to enhance accuracy as well as efficiency is introduced in this research as a solution to these challenges. The suggested method utilizes Principal Component Analysis (PCA) for reducing the dimensionality of features, minimizing computational complexity while preserving important facial expression features. For hyperparameter adjustment, Bayesian optimization is also utilized, allowing optimal CNN configurations to be chosen in order to enhance model performance. These techniques together allow the system to perform more effectively in real-time applications (N.-H. Chang et al., 2019).

The suggested system surpasses the current models both in recognition precision and computation efficiency using PCA and Bayesian optimization. It is ideal for real-time emotion recognition in many applications, such as human-computer interaction, surveillance, and healthcare, since it conserves processing time while still achieving good classification accuracy. This work highlights the suitability of deep learning-driven solutions in promoting FER technology while overcoming issues related to the limitations of real-time processing

2 RESEARCH METHODOLOGY

2.1 Research Area

The Field Domain with CNN, the objective of the field of Real-Time Facial Expression Recognition (FER) is to develop deep models that are capable of fast and accurate real-time facial expression

Rebekah, R. D. C., Bhavya, P., Vasanthi, A., Sumanjali, A. M. and Vyshnavi, V.
Real-Time Facial Expression Recognition Based on CNN.
DOI: 10.5220/0013917600004919

In *Proceedings of the 1st International Conference on Research and Development in Information, Communication, and Computing Technologies (ICRDICCT'25 2025) - Volume 4*, pages 605-609
ISBN: 978-989-758-777-1

recognition. The study enhances healthcare applications, surveillance, and human-computer interaction applications by fusing computer vision and artificial intelligence. Overcoming challenges such as lighting changes, occlusions, and differences between faces with high accuracy and computation speed is one of the biggest challenges in real-time FER. Although CNNs are superior in automatically extracting and classifying facial features, their high computations limit them for use in devices with limited power resources. This research employs Principal Component Analysis (PCA) to minimize feature dimensionality and Bayesian optimization to optimize hyperparameters to address these issues. The system becomes more effective for application in real-world scenarios due to these techniques' faster processing speed, reduced computational burden, and improved recognition accuracy.

2.2 Literature Review

Facial Expression Recognition (FER) has received intensive and there are several researchers who used deep models of learning, i.e., Convolution Neural Network (CNNs), to attain higher recognition rates in facial expression recognition (FER). The earlier FER methods that relied on hand-crafted feature extraction algorithms like Local Binary Patterns (LBP) and Histogram of Orientend Gradients (HOG) generally struggled with variations in lighting, occlusions, and individual facial differences.CNNs have greatly enhanced FER through self-extraction of hierarchical features from face images, reducing dependence on hand-engineered feature engineering. Despite their success, CNN-based FER systems still suffer from drawbacks, particularly computational complexity and real-time computation, such that they are less suited for deployment on resource-limited devices.

Recent studies have highlighted the optimization of CNN models and feature reduction techniques to enhance efficiency to bridge these gaps. Various dimension reduction techniques, including Principal Component Analysis (PCA) and Linear Discriminant Analysis (LDA), have been employed to preserve essential facial expression features while diminishe Computation costs. PCA has proved to be particularly valuable in eliminating redundant data, hence facilitating real-time processing of data. Further, techniques such as Genetic Algorithms and Bayesian optimization have been applied to optimize hyperparameters so that CNN models are more precise but utilizes less computational power. These Optimization methods enable FER systems to be utilized in actual circumstances by making certain that the maximum trade-off between recognition performance and processing complexity is achieved.

3 EXISTING SYSTEM

Facial Expression Recognition (FER) systems today are based on deep learning techniques Convolutional Neural Networks (CNNs), employed for automated face feature extraction and classification, form the basis of the current facial expression recognition (FER) systems. CNNs have produced much better recognition performance compared to traditional approaches such as the Histogram of Oriented Gradients (HOG) and Local Binary Patterns (LBP). Nonetheless, the general performance of current CNN-based FER models is hindered by problems such as occlusion, facial structure variations, and light sensitivity to changes. Real-time processing is also challenging because CNNs are computationally intensive, particularly on hardware with limited resources. Most FER systems depend on high-end computing or cloud processing, making them less viable for real-time use in environments with limited resources. While other models combine data augmentation and transfer learning to enhance accuracy, they are usually not able to find an optimal balance between speed and precision. These challenges highlight the need for better FER systems with high recognition accuracy in real-time applications.

Two Limitations of the Current System: Vulnerability to Lighting Changes and Obstructions CNN-based facial expression recognition (FER) systems face difficulties when dealing with varying lighting conditions, shadows, and obstructions such as glasses, masks, or facial hair. These challenges can negatively impact the system's ability to accurately extract and classify facial features, resulting in misidentifications. Consequently, the system's reliability decreases in diverse real-world settings.

High Processing Power Demand Real-time FER models built on CNNs require significant computational resources, making them impractical for devices with limited processing capabilities. The substantial computational load slows down inference speed, limiting deployment possibilities on mobile or edge devices without cloud-based infrastructure. This drawback hinders accessibility and scalability for widespread applications.

4 PROPOSED SYSTEM

The proposed system overcomes the challenges of traditional CNN-based facial expression recognition (FER) by employing statistical regression techniques to enhance accuracy and computational performance, the suggested system is able to overcome the challenges of standard CNN-based facial expression recognition (FER). Changes in lighting, occlusion, and high processing requirements are all typical issues for standard CNN models. To overcome these issues, more accurate relationships between facial expressions and emotions are created with techniques such as linear regression, logistic regression, and ridge regression. The system efficiently suppresses noise, simplifies distinguishing facial expressions, and becomes more robust to external influences such as shadows and obstructions through regression's contribution to reducing dimensionality, processing speed and computational complexity are minimized. Owing to this optimization, real-time FER can now be efficiently applied in limited-resource environments and can be integrated on edge devices without sacrificing performance. Regression is blended with deep learning to come up with an optimally well-balanced strategy that achieves its best recognition performance, processing throughput, and usage of hardware resources.

Through the application of statistical regression methods such as linear regression, and ridge regression, the new facial expression recognition accuracy is enhanced by the system considerably. These methods improve feature selection, denoise, and make the system operate more effectively to combat such issues as unstable light conditions, shadows, and occlusion. This results in a system that offers more accurate emotion recognition under diverse real-world conditions.

Faster processing and real-time execution are facilitated by regression-based dimensionality reduction, which simplifies computing. Regression-based dimensionality reduction, unlike traditional CNN-based approaches, accelerates processing and facilitates real-time execution. Our approach enhances feature extraction and classification, making it suitable for deployment on resource-constrained devices, including smartphones and edge computing platforms, without compromising recognition speed or accuracy. This is different from traditional CNN-based models, which need a lot of processing power.

4.1 Architecture

Deep hierarchical facial expression features are subsequently captured through a CNN-based feature extraction technique. Principal Component Analysis (PCA) is applied to diminish dimensionality without losing vital features for the purpose of controlling computational complexity. Removal of redundant information. Furthermore, statistical regression are applied in an effort to improve feature selection, eliminate noise, and form reliable relationships between facial features and emotions. Bayesian optimization is also used to adjust CNN hyperparameters to get the best possible compromise between recognition speed and accuracy. In the last phase, for classifying expressions, a hybrid model combining deep learning and regression-based approaches is deployed. Real-time application on low-resource devices like mobile platforms and edge computing platforms is enabled by this technique, which also reduces processing time enormously, enhances resistance to environmental effects like shadows and occlusions, and enhances environmental robustness. Figure 1 shows the Architecture diagram.

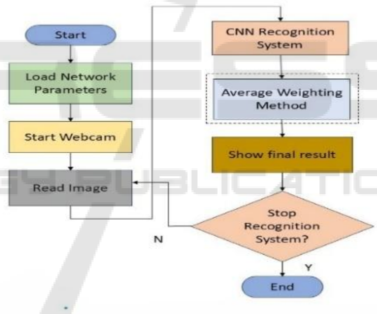
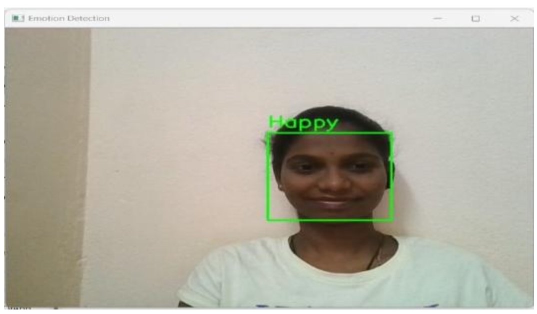


Figure 1: Architecture diagram.

5 RESULT

The following images (figure 2) are the output obtained as the real-time facial expression recognition based on CNN.



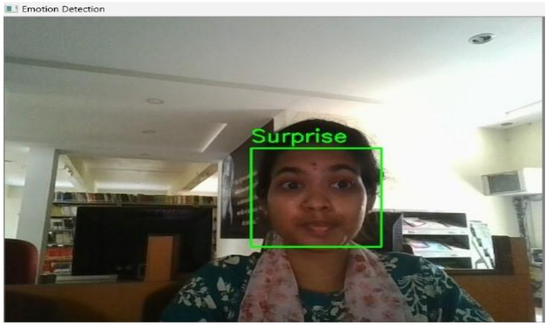


Figure 2: Results Obtained.

6 CONCLUSIONS

The standard convolutional neural network real-time facial expression recognition errors can be minimized through the use of this paper's average weighting scheme. Environmental noise can be minimized through the use of a high frame rate camera. The average weighting scheme also enhances facial expression recognition robustness between frames, leading to better accuracy. Experimental results indicate that the proposed facial expression recognition system is more dependable than the standard CNN method.

ACKNOWLEDGMENT

In the framework of the Ministry of Education (MOE) Taiwan Higher Education Sprout Project, National Taiwan Normal University's "Chinese Language and Technology Center" funded this research under. The Featured Areas Research Center Program. It was also funded by Taiwan's Ministry of Science and Technology under Grants number. MOST 108-2634-F-003-002 and MOST 108-2634-F-003-003 were supported by pervasive Artifical Intelligence Research (PAIR) Labs. We also like to thank the National Center for High-performance computing for granting computer time and facilities for conducting this research.

REFERENCES

B. Knyazev, R. Shvetsov, N. Efremova, and A. Kuharenko, "Convolutional neural networks pretrained on large face recognition datasets for emotion classification from video," in IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Honolulu, Hawaii, July 21-26, 2017.

based method on the segmentation and recognition of Chinese words," in Proc. of the 2018 International

Conference on System Science and Engineering (ICSS E), Taiwan, June 28-30, 2018.
I. J. Goodfellow et al., “2013. Challenges in representation learning: A report on three machine learning contests,” in International Conference on Neural Information Processing. Springer, pp.117-124.
J. Kim, J. K. Lee, and K. M. Lee, “Deeply-recursive convolutional network for image super-resolution,” in IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Las Vegas, US, June 26-July 1, 2016.
J.-H. Chen, Y.-H. Chien, H.-H. Chiang, W.-Y. Wang, and C.-C. Hsu, “A robotic arm system learning from human demonstrations,” in Proc. Of the 2018 Internati onal Automatic Control Conference (CACS), Taoyuan , Taiwan, Nov. 4-7, 2018.
M. Pantic and J. M. Rothkrantz, “Facial action recognition for facial expression analysis from static face images,” IEEE Trans. Systems, Man and Cybernetics, vol. 34, no. 3, pp. 1449-1461, 2004.
M. S. Bartlett, G. Littlewort, M. G. Frank, C. Lainscsek, I. Fasel, J. Movellan, “Fully Automatic Facial Action Re cognition in Spontaneous Behavior", Proc. IEEE Int'l Conf. Automatic Face and Gesture Recognition (AFGR), Southampton, UK, April 10-12, 2006, pp. 223-230.
Mohebbanaaz, M. Jyothirmai, K. Mounika, E. Sravani and B. Mounika, “Detection and Identification of Fake Images using Conditional Generative Adversarial Networks (CGANs),” 2024 IEEE 16th International Conferences on Computational Intelligence and Communication Networks (CICN), Indore, India, 2024, pp. 606-610, doi;10.1109/CICN63059.2024.10847379.
N.-H. Chang, Y.-H. Chien, H.-H. Chiang, W.-Y. Wang, and C.-C. Hsu, “A robot obstacle avoidance method using merged cnn framework,” in Proc. Of the 2019 Internat ional Conference on Machine Learning and Cyberneti cs (ICMLC), Kobe, Japan, July 7-10, 2019.
S. Alizadeh and A. Fazel, “Convolutional neural networks for facial expression recognition,” in IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Honolulu, Hawaii, July 21-26, 2017.
S. Li and W. Deng, “Reliable crowdsourcing and deep locality preserving learning for unconstrained facial expression recognition,” IEEE Trans. Image Process., vol. 28, no. 1, pp. 356-370, Jan. 2019.
T. H. H. Zavaschi, L. E. S. Oliveira, A. L. Koerich, "Facial Expression Recognition Using Ensemble of Classifiers ", 36th International Conference on Acoustics Speech and Signal Processing (ICASSP), Prague, Czech Republic, May 22-27, 2011, pp. 1489-1492.
V. D. Van, T. Thai and M. Q. Nghiem, “Combining convolution and recursive neural networks for sentime nt analysis,” in Proc. 8th Int. Symp. Information and Communication Technology, Nha Trang City, Dec. 7-8, 2017, pp.151-158.
Y. Tian, T. Kanade, and J. Cohn, “Recognizing action units for facial expression analysis,” IEEE Trans. Pattern Analysis and Machine Intelligence, vol. 23, no. 2, pp. 97-115, 2001.
Y.-T. Wu, Y.-H. Chien, W.-Y. Wang, and C.-C. Hsu, “A YOLO-