International Journal of Software Science and Computational Intelligence
Volume 17 • Issue 1 • January-December 2025

Feature Pruned and Weighted Convolutional Neural Network for Micro-Expression Recognition:
Lightweight Deep Learning Framework

Kaja Mohideen
http://orcid.org/0000-0003-1329-3925
Vellore Institute of Technology, India

ABSTRACT

Micro-expressions are brief, involuntary facial movements that reveal concealed emotions. While Convolutional Neural Networks (CNNs) are effective for recognizing such expressions, they are often computationally intensive and memory heavy. To address this, a Feature-Pruned and Weighted Convolutional Neural Network (FPW-CNN) is proposed to compress and accelerate performance. This model prunes filters and selects only the most relevant responses, reducing memory usage and computational load. Instead of stacking outputs, convolutional layers are flattened. Experiments on five spontaneous micro-expression datasets—SMIC, CASME, CASME II, CAS(ME)^2, and SAMM—show that FPW-CNN achieves accuracies of 99.75%, 99.93%, 99.95%, 99.94%, and 99.90%, respectively. The model also reduces FLOPs by over 41% and maintains over 95% accuracy at 128×128 resolution, demonstrating robustness. These results confirm that FPW-CNN outperforms recent CNN and transformer models, making it suitable for real-time micro-expression recognition on resource-constrained devices.

KEYWORDS

Convolutional Neural Network, Edge Detection, Feature Pruning, Gabor Filters, Micro-Expression

INTRODUCTION

Facial expressions are a primary source of emotional cues, with Micro-Expressions (MEs) serving as brief, involuntary facial movements that reveal suppressed emotions, especially in high-stakes contexts. Unlike macro-expressions that last 4–5 seconds, micro-expressions endure for only 1/5 to 1/25 seconds (Yan et al. 2013), making them hard to detect with the naked eye. They are useful in areas such as security interrogations and clinical diagnostics. While polygraph tests serve a similar purpose, they are invasive and vulnerable to manipulation. In contrast, automatic ME recognition is non-invasive and harder to deceive. Verma et al. (2020) found that even professionals only achieve a 47% success rate in ME detection, highlighting the need for automated systems. They presented a gradient-based Neural Architecture Search (NAS) algorithm to identify the efficient architecture for ME recognition. Action units such as the upper face, eyes, and lip corners are key indicators in MEs. Ekman (1969) emphasized the role of MEs in deception detection and developed the Micro-Expression Training Tool (METT) to classify emotions like anger, disgust, and surprise. Given their subtlety and

DOI: 10.4018/IJSSCI.389197

fleeting nature, ME Recognition (MER) demands intelligent algorithms. Ng et al. (2015) and Takalkar & Xu (2017) pointed out that performance suffers when datasets are small or unbalanced. Traditionally, MER systems involve selecting facial regions, extracting feature descriptors, and training classifiers (Kumala et al. 2019; Huang et al. 2016; Liu et al. 2015). Deep learning, particularly Convolutional Neural Networks (CNNs), has become prominent in this space due to its ability to learn complex features. However, CNNs are computationally intensive and require careful parameter tuning (Li et al. 2019; Song et al. 2019; Wang et al. 2018; Peng et al. 2017). This paper introduces a Feature-Pruned and Weighted CNN (FPW-CNN) that reduces complexity by pruning redundant filters and applying weights to relevant ones. Key innovations include using originfeatures andence images, extracting and pruning edge/texture features, and employing weighted convolution for efficient training. The main contributions of the research work are:

- The proposed MER system considers the original and the difference image (variation between the current and the previous frame) as inputs.
- Layers of edge & textural features are extracted and pruned to choose the best feature set to reduce the complexity of the learning step.
- The pruned filters are subjected to a weighted convolution process to continue the learning procedure.

The rest of the paper is structured as follows: Section 2 provides a comprehensive review of recent Convolutional Neural Network (CNN) models applied to Micro-Expression Recognition problems. Section 3 delves into the details of the proposed FPW-CNN. Section 4 outlines the implementation setup and describes the dataset used for evaluation. The analysis of experimental results is presented in Section 5, while Section 6 offers the conclusion of the paper.

RELATED WORKS

Micro-Expression Recognition (MER) has evolved significantly from early models using posed expressions to advanced deep learning systems handling spontaneous micro-expressions. CNNs have become pivotal in this evolution due to their hierarchical feature learning capabilities, widely adopted across domains like object recognition, face verification, and image classification. Patel et al. (2016) introduced a feature selection framework aiming to enhance the performance of deep learning in the domain of Micro-Expression (ME) recognition. However, experimental results showed that their method could not rival conventional models due to overfitting the training data. On the other hand, Peng et al. (2017) proposed a spontaneous Micro-Expression Recognition (MER) system utilizing the Dual Temporal Scale Convolutional Neural Network (DTSCNN) approach. The spontaneous micro-expression videos at different frame rates are adopted with different streams of a network. And the overfitting problem is avoided by using an independent shallow network. Experiments are carried out with the CASME I/II dataset and show that the recognition rate is 10% greater than the conventional methods. Takalkar & Xu (2017) proposed a CNN-based micro-expression recognition system with synthetic images augmented from CASME I/II datasets. The experimental results indicate significant performance improvement with a higher recognition rate. The small dimension issue of micro-expression data could be handled by the transfer learning approach, where the deep learning model trained with some other data is used as a starting point rather than building a deep learning model from scratch. In their study, Ferreira et al. (2018) presented an end-to-end neural network that leverages prior facts of facial expressions to achieve improved recognition of micro-expressions. This neural network architecture learns expression-specific features. The simulation results indicate that the end-to-end neural network outperforms the conventional deep learning methods. Gan & Liong (2018) proposed a novel feature descriptor based on an apex frame for a CNN-based micro-expression recognition system. The feature descriptor is derived from optical flow fields. The investigation

Related Content