Format Prepared by: Sudipto Chaki, Assistant Professor, CSE, BUBT

Department of Computer Science and Engineering Bangladesh University of Business and Technology (BUBT)

CSE 498A: Literature Review Records

Student's Id and Name	22234103382, Umma Sumaiya Riya
Capstone Project Title	High-Efficiency Micro-Expression Recognition for Automated Human Behavior Analysis Systems.
Supervisor Name & Designation	Md. Mijanur Rahman, Assistant Professor
Course Teacher's Name & Designation	Sudipto Chaki, Assistant Professor

Aspects	**Paper # 22 AFER-POADCNN: Automatic Facial Emotion Recognition Using Pelican Optimization and Deep Capsule Networks [Published Year: 2023 Publisher: MDPI Electronics]**
Problem Statement	Facial emotion recognition becomes difficult due to noise in images, variations in lighting/pose, and insufficient labeled data. Traditional CNN-based models struggle to extract robust spatial features and require manual hyperparameter tuning, which reduces accuracy. The authors aim to design a more accurate, noise-resistant, and automatically optimized FER method.
Key Contributions	The core contributions of this paper are: 1)Introduced POA-tuned CapsNet for enhanced emotion feature extraction. 2)Used Median Filtering to reduce noise before processing. 3)Proposed a full pipeline combining MF + POA + ADCNN + BiLSTM for FER.

	4)Achieved higher accuracy than SVM, AlexNet, VGG19, ResNet152, and MobileNet.
Methodology/Theory/Framework	The proposed AFER-POADCNN pipeline includes: ● Median Filter (MF) for noise reduction. ● Adaptive Deep Convolutional Neural Network (ADCNN) for initial feature extraction. ● CapsNet, which captures spatial–hierarchical relationships more effectively than CNNs. ● Pelican Optimization Algorithm (POA) to tune CapsNet hyperparameters automatically. ● BiLSTM classifier to learn sequential facial features and generate final emotion labels
SoftwareTools/Setup Details	Python (Anaconda) TensorFlow / Keras Google Colab GPU (Tesla K80) Hardware: Intel Core i5 processor, NVIDIA GTX 1050Ti (4GB)
Test/Experiment Analysis	Experiments were performed using **70:30** and **80:20** train–test splits with metrics including **Accuracy, Sensitivity, Specificity, F-score, and MCC**. Hyperparameters were tuned using POA, while CapsNet and BiLSTM were trained with optimized settings. Median Filtering helped stabilize input images. The model was compared with CNN, SVM, AlexNet, VGG19, ResNet152, and MobileNet. **Accuracy:** ● **80:20 split → 99.05%** ● **70:30 split → 97.15%**
Test Data/Dataset Source	The authors used a publicly available **Facial Emotion Recognition (FER) image dataset** containing **920 labeled images** across **8 emotion classes**: Happy, Sad, Fear, Angry, Neutral, Surprise, Disgust, Contempt. The images were collected from multiple open-access FER sources and cleaned before training.

Final Result (Assessment Criteria Wise)	The authors used a publicly available **Facial Emotion Recognition (FER) image dataset** containing **920 labeled images** across **8 emotion classes**: Happy, Sad, Fear, Angry, Neutral, Surprise, Disgust, Contempt. The images were collected from multiple open-access FER sources and cleaned before training.
Limitations (List the limitations the authors mentioned in the article)	● Dataset is small and imbalanced, which may reduce generalization. ● Performance on real-world or uncontrolled (in-the-wild) data is untested. ● POA is computationally expensive for hyperparameter tuning. ● Pose variation and occlusion robustness not fully evaluated.
Final Summary	S. Alshmrany et al. proposed the AFER-POADCNN model for facial emotion recognition by integrating Median Filtering, CapsNet, POA-based hyperparameter tuning, and a BiLSTM classifier. The method enhances feature extraction quality and achieves high accuracy through automated optimization. Experiments on a facial emotion dataset show a maximum accuracy of 99.05%, outperforming popular models like VGG19, ResNet152, and MobileNet. Training was conducted using 70:30 and 80:20 splits with multiple evaluation metrics. Although the approach performs strongly, it is limited by a small dataset and does not fully address real-world variations. The model also lacks robustness testing for pose, lighting, and occlusions. Overall, AFER-POADCNN is a promising and highly accurate FER framework.