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 # 25 Geometrical Facial Expression Recognition Approach Based on Fusion CNN-SVM [Published Year: 2024, Publisher:International Journal of Intelligent Engineering and Systems]**
Problem Statement	Facial Expression Recognition (FER) remains challenging due to noise, illumination changes, pose variations, and the inability of single-feature methods to accurately capture subtle emotional changes. Appearance-based models struggle with texture variations, while geometric-only methods fail to generalize across subjects. The authors aim to develop a FER system that robustly combines shape (geometric) and texture (appearance) information to improve accuracy, especially with small datasets like CK+ and JAF
Key Contributions	(i) Introduced a hybrid Fusion CNN-SVM architecture combining appearance features from 2D-CNN and geometric features from β-skeleton-based landmark modeling. (ii) Proposed a geometry extraction method using 68 facial landmarks grouped into seven facial regions and transformed into discriminative geometric vectors.

	(iii) Utilized a 1D-CNN + 2D-CNN dual-stream feature learning mechanism. (iv) Replaced Softmax with L2-SVM for improved classification on small datasets. (v) Achieved competitive results on CK+ and JAFFE benchmarks.
Methodology/Theory/Framework	○ The authors built a two-branch CNN model. ○ **Geometric branch:** detects 68 facial landmarks, generates β-skeleton geometric features, and processes them with a 1D-CNN. ○ **Appearance branch:** preprocesses face images (crop, resize, equalize) and extracts texture features using a 2D-CNN. ○ **Fusion & classification:** both feature streams are merged, and an L2-SVM classifier is used instead of Softmax for better performance on small datasets.
SoftwareTools/Setup Details	☐ Programming Language: **Python 3.6** ☐ Libraries: **TensorFlow/Keras**, Dlib for landmarks ☐ Hardware: **Intel Core i5 CPU**, 4GB RAM ☐ Image size: **50×50 pixels** ☐ Training parameters: **Adam optimizer**, LR = 0.0001, batch size = 32, 50 epochs ☐ Data Augmentation: Rotation (±15°), flipping, blur, sharpening, mirroring
Test/Experiment Analysis	The authors performed experiments on CK+ and JAFFE datasets with an **80/20 train–test split**. They evaluated models on accuracy, precision, recall, and F1-score. Comparisons were made between: ● baseline CNN ● geometric-only features ● appearance-only features ● **proposed Fusion CNN-SVM model** The fusion model outperformed all other approaches, showing improved recognition of happy and surprise emotions, while challenging emotions like fear and disgust showed moderate accuracy.**Accuracy:** ● **CK+ dataset:** 96.19% ● **JAFFE dataset:** 89.23%

Test Data/Dataset Source	**CK+ Dataset:** 593 facial expression sequences, 7 emotion classes, frontal images; a standard FER benchmark. **JAFFE Dataset:** 213 images of 10 subjects, 7 expressions; collected in a controlled lab environment.
Final Result (Assessment Criteria Wise)	The proposed method demonstrated strong performance in combining appearance and geometric features. The model achieved high accuracy, better generalization on small datasets, improved robustness against pose/illumination variations, and superior results compared to traditional CNN and handcrafted feature methods. The SVM classifier contributed to stable performance and reduced overfitting..
Limitations (List the limitations the authors mentioned in the article)	1)Performance decreases for highly similar expressions such as fear and disgust. 20Dependent on accurate landmark detection; β-skeleton fails if landmarks are wrong. 3)Tested on only two small datasets; not evaluated in real-world, unconstrained environments. 4)No temporal information (video sequences) included.
Final Summary	The paper presents a hybrid Fusion CNN-SVM framework that integrates geometric features from β-skeleton landmark modeling with appearance features from 2D-CNN. A dual-branch CNN architecture is used to extract both shape and texture-related patterns, which are later fused and classified using an L2-SVM classifier. Experiments conducted on CK+ and JAFFE datasets show that the proposed method achieves accuracy values of 96.19% and 89.23%, outperforming baseline CNN and traditional handcrafted methods. The model demonstrates strong generalization ability despite small training datasets due to the fusion strategy and SVM-based decision-making. However, the approach faces difficulties distinguishing subtle expressions and relies heavily on accurate landmark detection. It also lacks evaluation in real-world conditions. Overall, the fusion architecture provides an effective and efficient FER solution that leverages both geometric and appearance representations..