A Comprehensive Within-Subject Analysis of EEGNet for SEED-IV Emotion Recognition: Subject Variability and Per-Class Performance

Sujata Kulkarni
KLE Technological University, Karnataka, Hubballi, India
sujatakulkarni@kletech.ac.in (corresponding author)

Prakashgoud Patil
KLE Technological University, Karnataka, Hubballi, India
prpatilji@gmail.com

Received: 24 December 2025 | Revised: 4 February 2026 and 22 February 2026 | Accepted: 23 February 2026

 DOI: https://doi.org/10.48084/etasr.17161

ABSTRACT

This paper presents a comprehensive within-subject analysis of the EEGNet model's performance on the SEED-IV dataset for multi-class emotion recognition. This analysis accounts for the individual contributions of each of the 15 subjects by considering a standardized preprocessing procedure, subject-specific splitting of the data into training, validation, and test sets, and a trained model architecture and process for each subject to account for individual variability. Accuracy, precision, and recall of the model for each subject, as well as class-specific measures, were used to evaluate model performance for the differentiation of each emotion on the SEED-IV dataset, analyzing the amount of each emotion that the model can reproduce on the test dataset, presenting an accurate interpretation of the distance between classes by analyzing confusion matrices, as well as the capability of the model to differentiate learned emotions on the test dataset by means of the two-dimensional visualization of the learned features extracted by the model. The study presents three contributions: (i) subject clustering analysis, (ii) per-class recall analysis, and (iii) validation that the raw EEG waveform achieves competitive performance.

Keywords-emotion recognition; EEGNet; SEED IV; brain computer interface

I. INTRODUCTION

A. Background

Electroencephalography (EEG)-based emotion recognition is an application area associated with affective computing, Brain–Computer Interface (BCI), and Human–Computer Interface (HCI). Designing models for emotion recognition using EEG is challenging because EEG signals are typically noisy, non-stationary, and highly user-dependent. Physiological signals provide a direct, objective, and reliable representation of an individual's internal emotional state, which may not always be accurately reflected through external observations.

In [1], a framework for emotion recognition based on Principal Component Analysis and Random Forest classifier (PCA-RF) is proposed. It used the DEAP dataset. This method involved an initial processing step that included filtering based on specific frequencies, followed by PCA for noise reduction. The noise-filtered EEG features were classified using a Random Forest (RF) classifier. In [2], a hybrid fuzzy Convolutional Neural Network - Long Short-Term Memory (CNN-LSTM) framework for multimodal emotion recognition using EEG and electrocardiogram (ECG) physiological signals is proposed. Feature-level fusion of EEG and ECG was combined with fuzzy logic-based membership functions to handle uncertainty, intersubject variability, and noisy physical data. Features enhanced through fuzziness were fed into the CNN-LSTM model. In [3], hybrid deep learning approaches were applied for emotion recognition. The study specifically examines the effectiveness of the models in data from individuals who cannot express their feelings explicitly. Authors in [4] introduce a deep learning model comprising a CNN and an LSTM for emotion recognition based on EEG signals. The model employed CNN layers to represent spatial attributes in EEG channels and LSTM layers to analyze temporal dependencies in emotional responses. The results showed better accuracy in emotion recognition. In [5], transformer neural networks were used for emotion recognition. Transformers can use a self-attention mechanism to capture long-term dependencies across EEG signal channels. The proposed method relies less on feature engineering and achieves better generalization in emotion recognition. Authors in [6] proposed the use of Graph Neural Networks (GNNs) to detect emotional states from EEG signals. In this method, EEG signal channels are graphed to reflect real-world connections in the human brain. The proposed approach showed high spatial dependencies in emotional responses and higher recognition

accuracy in subjects. Authors in [7] proposed a strategy using a hybrid architecture of CNN and Bi-LSTM in the frequency domain. EEG features were extracted using the Wavelet transform. The bidirectional LSTM allowed learning from both past and future temporal dependencies, resulting in improved emotion recognition accuracy compared to unidirectional models. Authors in [8] developed a multimodal EEG-based emotion recognition system that combines EEG with other signals, such as GSR and EMG. Multimodal emotion recognition systems employing feature-level fusion with Machine Learning classifiers achieve significantly higher classification accuracy than unimodal EEG-based systems. Authors in [9] developed an emotion recognition model based on EEG signals, incorporating a genetically optimized Multi-Layer Perceptron (GA-MLP) architecture for applications in affective computing and brain-computer interfaces. Power Spectral Density (PSD) features mapped to the valence–arousal model were employed to quantify emotional states. The Genetic Algorithm (GA) was used to optimize the MLP architecture, improving classification performance. The model achieved 91.10% accuracy for valence, 91.02% for arousal, and 83.52% for four-class emotion classification. Authors in [10] provide an objective method for analyzing emotional states, however, recognizing patterns in multichannel EEG data is challenging for traditional ML methods. Therefore, the work employs deep learning models, specifically CNNs and LSTMs, which are effective at automatically learning complex features. The proposed models classify arousal, valence, and dominance emotions using EEG data from the DEAP dataset, considering subjects with and without recorded frontal videos. Authors in [11] propose a framework that integrates effective connectivity and deep learning. Fused connectivity representations derived from transfer entropy, partial directed coherence, and direct transfer function. These are then classified by pre-trained CNN and CNN-LSTM classifiers using an ensemble learning approach. Authors in [12] propose the EmotionMeter, a multimodal emotion recognition framework that integrates multiple physiological and behavioral signals to improve human emotion recognition accuracy.

B. Motivation

Most EEG emotion studies report only group-level metrics, which obscure how well models work for individual users. A within-subject analysis is needed to reveal this subject-specific variability. Also, aggregate accuracy hides which emotions are consistently easy or difficult to recognize, so detailed per-class performance is required to understand systematic confusions between emotion categories. While EEGNet is widely used, its behavior at both the individual and per-class levels is not well characterized, and clarifying this is crucial for subject-dependent applications, such as personalized mental health monitoring and adaptive human–computer interaction.

II. METHODOLOGY

The pipeline diagram in Figure 1 summarizes the complete within-subject EEG emotion recognition workflow from raw SEED-IV [12, 13] data to evaluation for each participant. The study utilizes the SEED-IV, a multimodal EEG emotion dataset comprising EEG recordings from 15 healthy subjects collected across three sessions. At the same time, participants viewed emotion-eliciting film clips corresponding to four emotions: neutral, sadness, fear, and happiness. EEG signals were recorded using a 62-channel ESI NeuroScan system at a sampling rate of 1,000 Hz and subsequently filtered and downsampled to 200 Hz. This dataset includes eye movement signals in addition to EEG. This makes the dataset suitable for either unimodal or multimodal affective computing analysis. The SEED IV dataset has been widely accepted for evaluating cross-subject and within-subject emotion recognition algorithms using EEG. The ground truths are expressed in emotion class labels of {0, 1, 2, 3} corresponding to emotions of "Neutral", "Sad", "Fear", and "Happy", respectively.

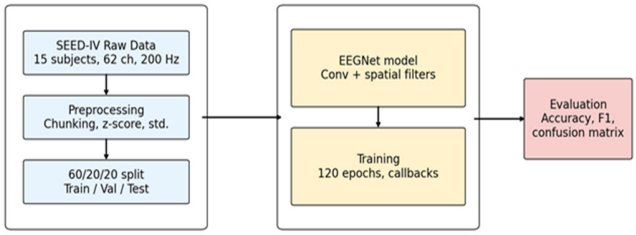

Fig. 1. Preprocessing, modeling, and evaluation workflow.

A. Data Preprocessing

The continuous EEG signal for a given subject is segmented into short-time epochs, or "chunks," which are 400 ms long and non-overlapping, derived sequentially. Having fixed, non-overlapping chunks increases the number of training examples, capturing spatial locality even when overlap between adjacent chunks makes nearby chunks highly correlated. Normalization, achieved by subtracting the mean and dividing by the standard deviation computed over the samples in the current chunk, reduces the effect of slowly drifting signals and the impact of signals being different in scale across chunks. After this chunk-based normalization, the subject's data undergo additional standardization across the entire dataset. For each subject, the chunks, along with the emotion labels for each chunk, are split into training, validation, and test sets, with 60% allocated to training, 20% to validation, and 20% to testing. This segmentation is carried out in a way that maintains the ratio of the different classes in approximately equal proportions across the test, validation, and training sets for the particular subject. This procedure ensures that: (i) Only the subject's data will be used to train and validate the model, (ii) the validation will have a fair idea of the subject's class distribution, (iii) it will give an unbiased estimate of the test accuracy in each subject and each class.

B. EEGNet Architecture

EEGNet is a compact CNN designed explicitly for EEG signals. It efficiently learns both spatial and temporal features with a minimal number of parameters, making it suitable for datasets with limited sample sizes, such as the SEED-IV. The first layer of EEGNet applies a 1D temporal convolution with a kernel length of 64 to each channel along the time axis. This acts like a learnable band-pass filter bank that captures frequency-specific patterns from the raw 400-ms segments. After temporal filtering, a depth-wise spatial convolution is applied across the

channels. Instead of mixing all channels at once, each temporal filter is spatially projected over the 62 electrodes with separate depth-wise kernels. This enables EEGNet to learn subject-specific spatial patterns while maintaining a small number of parameters and preserving the EEG's topography.

The parameters F1=8 and F2=16 specify a small number of temporal and separable-convolution filters, keeping EEGNet compact and less prone to overfitting on limited within-subject data. The dropout rate of 0.5 randomly silences half of the units during training, so the model relies on redundant features instead of memorizing subject-specific noise. The ELU activation function handles the gradient flow and convergence speed, and can be considered an improvement over ReLU. The model was trained using the Adam learning rule at a learning rate of 0.001, which enables adaptive learning for each parameter and allows for variations that may differ across individuals. ReduceLROnPlateau lowers the learning rate when validation performance saturates, enabling coarse-to-fine optimization. EarlyStopping stops training once validation performance stops improving and restores the best weights, thereby limiting overfitting to each subject's training chunks. Overall, the configuration is designed to be compact, regularized, and conservatively trained, allowing each subject-specific EEGNet to generalize well despite having relatively few labeled trials. The model was trained for 120 epochs. Training, validation, and test accuracy, along with macro precision, macro recall, and accuracy per class were used to evaluate the model.

III. EXPERIMENTS

A. Results

Each experiment was conducted in a rigorous within-subject design, with a distinct EEGNet model used for each of the 15 participants in the SEED-IV. For each subject, the continuously recorded EEG data of 62 channels were first divided into 400-ms non-overlapping chunks. Then these chunks were subjected to z-scoring. These chunks were then assigned to training, validation, and test sets for each subject. Table I reports the resulting per-subject performance, including test accuracy, macro F1-score, and per-class recall for the four emotion categories. These metrics quantify how reliably each subject's model recognizes emotions overall (accuracy), how balanced the performance is across classes, and which specific emotions are easy or difficult to detect for each participant. Table I highlights the degree of subject variability in EEGNet's behavior on SEED-IV. It serves as the basis for subsequent analyses of subject-wise differences and class-specific trends. These results demonstrate apparent subject variability, good generalization within each subject, and uneven class-wise performance, despite the use of the same EEGNet architecture and training protocol for all participants.

TABLE I. PER SUBJECT PERFORMANCE (%)

Subject No	Train Accuracy	Validation accuracy	Test Accuracy	Macro Precision	Macro Recall	Neutral Class 0	Sad Class 1	Fear Class 2	Happy Class 3
1	56.69	50.45	48.27	50.5	47.2	60.4	51.1	48.4	29.0
2	76.43	74.83	73.84	73.9	73.5	85.7	69.0	69.8	69.6
3	63.44	58.08	55.80	56.1	56.4	59.3	52.2	41.9	72.0
4	80.56	77.80	73.44	74.3	73.3	81.3	65.0	78.6	68.2
5	67.40	63.33	61.15	60.8	61.9	68.9	50.7	48.0	79.9
6	74.28	67.20	67.89	69.0	68.4	81.7	53.3	58.1	80.4
7	62.51	61.45	57.48	60.2	56.8	72.5	50.0	62.1	42.5
8	65.16	62.34	60.36	60.9	60.5	61.5	67.2	44.0	69.2
9	70.12	68.68	66.20	66.0	66.3	59.0	66.8	74.2	65.4
10	67.21	59.76	56.89	56.7	56.7	76.2	41.6	54.0	55.1
11	49.59	43.81	43.81	43.4	43.2	55.7	39.8	44.4	33.2
12	55.77	47.87	44.70	44.2	44.3	68.5	23.4	50.0	35.5
13	63.74	57.98	57.78	58.2	58.2	52.7	58.4	56.0	65.4
14	79.57	74.93	70.96	70.8	71.0	63.4	67.2	88.7	65.0
15	90.38	90.98	89.20	89.2	89.0	89.1	95.98	93.5	85.9

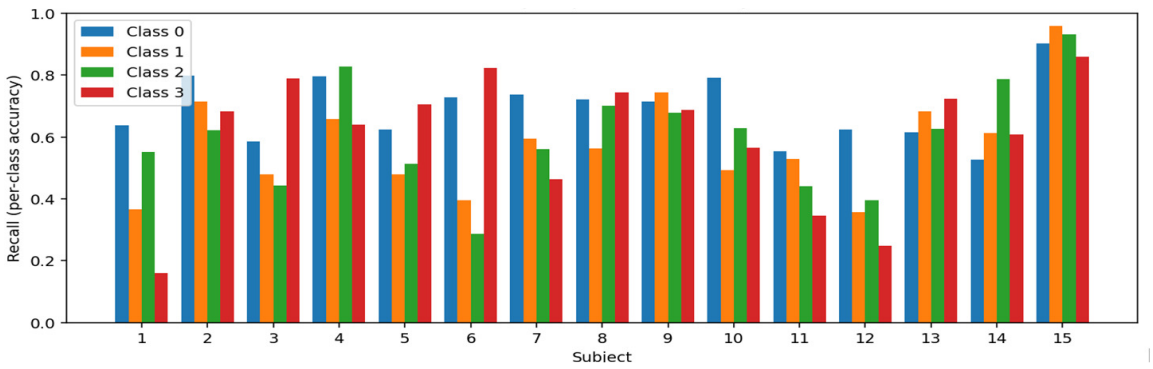

Fig. 2. Per-class accuracy for all participants.

B. Subject Variability Analysis

Subject variability refers to differences in the EEG patterns across different individuals. It is one of the most critical challenges because each person's brain signals are unique, emotions vary across individuals, and how these signals are recorded differs. Figure 2 illustrates the class-wise accuracies of all 15 participants. As shown in Table I, the range of test accuracy across the 15 subjects starts from 43.81% for Subject 11 to 89.20% for Subject 15, a difference of 45.39%. The average test accuracy across all subjects is 61.85%, with a standard deviation of 11.74%, confirming the variability among subjects despite the same architectures and training procedures. At the high-performance end, Subject 15 performs the best across all metrics with 89.20% test accuracy, 89.2% macro precision, and 89.0% macro recall. Per-class recall for Subject 15 is uniformly high: 89.1% for neutral, 95.98% for sad, 93.5% for afraid, and 85.9% for happy. The 10.08% difference between these four values suggests that the model for Subject 15 generalizes very well across all four emotion classes, with no class bias. At the low-performance end, Subject 11 records the lowest test accuracy of 43.81%, with macro precision of 43.4% and macro recall of 43.2%. Per-class recall for Subject 11 is 55.7% for neutral, 39.8% for sad, 44.4% for afraid, and 33.2% for happy. Happy is the most challenging class for this subject. Subject 6 illustrates selective class difficulty: Neutral and happy achieve high recall rates of 81.7% and 80.4%, respectively, while sad and afraid are notably lower at 53.3% and 58.1%, respectively. The 28.4% gap between the best-performing class in this subject indicates that certain emotion pairs produce overlapping EEG representations specific to this individual. Subjects 1 and 12 both exhibit strong bias toward the Neutral class. Subject 1 records neutral recall of 60.4% against happy recall of only 29.0%, a gap of 31.4%. Subject 12 exhibits a gap between neutral and sad recall of 45.1%. These findings confirm that within-subject EEG emotion recognition is highly individual-dependent.

Due to space constraints, a small representative subset of the confusion matrices and t-SNE (t-distributed Stochastic Neighbor Embedding) plots is analyzed. The three confusion matrices selected are those of Subject 1, which exhibits heavy confusion in certain emotions, Subject 6, with moderate structured confusions, and Subject 15, which demonstrates high accuracy and balanced per-class performance. In Figure 3, for Subject 1, the model recognizes well the neutral class but frequently confuses sad and happy with neutral trials, indicating a bias toward predicting neutrality. Fear achieves the highest per-class recall, yet still exhibits notable misclassification into neutral and sad categories, indicating that its representation is only partially separable from those categories. Overall, the errors are intensely focused on predictions of the neutral class, suggesting that the network tends to predict "neutral" whenever the learned features are ambiguous. For Subject 6, the Nnutral and happy classes are well recognized, but sad and fear results show many misclassifications, suggesting overlapping EEG patterns between these two negative emotions. Asymmetric misclassification toward neutral and happy is also noticed. A notable number of sad and fear instances are also misclassified as neutral or happy, indicating that the classifier sometimes collapses negative feelings into more extreme or baseline states.

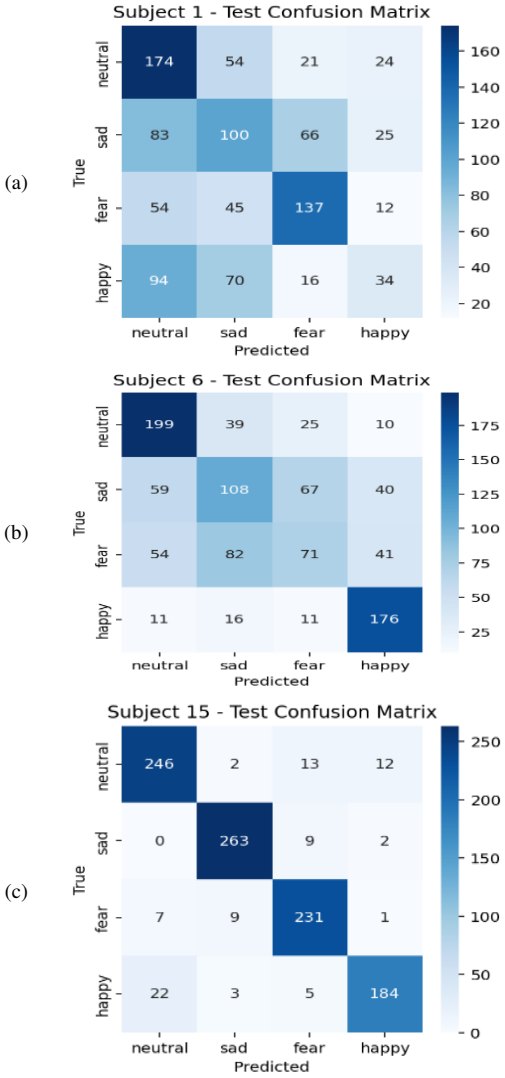

Fig. 3. Confusion matrices of Subjects (a) 1, (b) 6, and (c) 15.

For Subject 15, all four emotions are accurately recognized: The diagonal entries are high for every, indicating consistently strong classification performance. The number of off-diagonal elements is low, ranging mainly from 0 to 13, indicating that the instances of a given emotion are rarely confused with cases of other emotions. Negative emotions are well distinguished: There is minimal confusion between sad and fear trials, indicating that the model can distinguish EEG patterns for both tasks with limited bias toward any single class. Figure 4 illustrates the progressive improvement in within-subject emotion separability from heavily overlapping clusters (Subject 1) to partially structured manifolds (Subject 6) and finally to four well-defined, compact class clusters (Subject 15), reflecting an increase in inter-class discriminability and a decrease in intra-class variability across participants.

t-SNE visualizations were employed here as a supportive tool to illustrate the learned feature representations and are not considered definitive evidence of class separability. t-SNE projections are sensitive to hyperparameters such as perplexity

and learning rate, and the resulting 2D embedding may not accurately reflect the high-dimensional geometry of the feature space. The following observations from the t-SNE plots are therefore interpreted in conjunction with the quantitative metrics reported in Table I and the corresponding confusion matrices. For Subject 1, Figure 4(a) shows the t-SNE embedding with heavily intermixed points across all four classes, consistent with the low test accuracy of 48.27% and the high off-diagonal values in the confusion matrix. For Subject 6, Figure 4(b), shows a partial separation: one class forms a relatively compact cluster, while the other three classes overlap. For Subject 15, in Figure 4(c), there are four distinct clusters with little overlap. This is also reflected in the quantitative analysis, where the test accuracy is 89.20%, the per-class recall values vary from 85.9% to 95.98%, and the off-diagonal confusion matrix values are all consistently below 13. This suggests that the learned representations for Subject 15 have high inter-class separability.

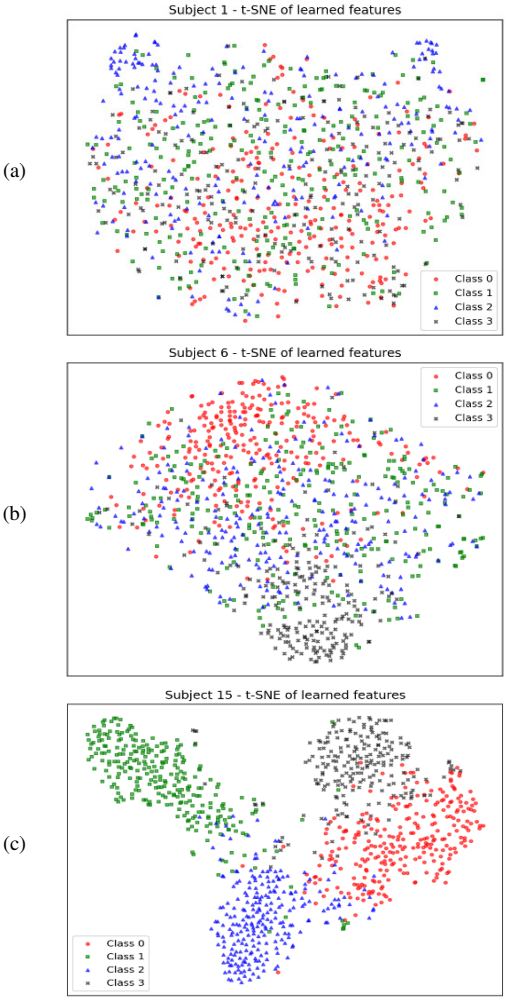

Fig. 4. t-SNE visualization of learned EEG feature representations of Subject (a) 1, (b) 6, and (c) 15.

With respect to within-subject analysis, this structure indicates that the network has learned highly discriminative subject-specific features; trials from the same emotion consistently map to the same region, while trials from different emotions occupy distinct manifolds in the feature space. In terms of variability analysis, the tight, coherent clusters imply low intra-class variability and high inter-class variability for Subject 15, which explains the very high per-class accuracies and low confusion observed in the corresponding confusion matrix. This subject, therefore, represents a case where SEED IV emotions are strongly separable in the neural feature space, contrasting with subjects such as Subject 1, where t-SNE reveals heavy mixing between classes.

Figure 5 illustrates the Subject-wise learning curves of training and validation accuracy versus epochs for Subjects 1, 6, and 15, demonstrating learning and generalization dynamics. In Figure 5(a), the model's performance varies more erratically across epochs, hence, repeated training runs or session splits would likely reveal larger performance swings for the same subject. In Figure 5(b), the learning curve is much smoother, indicating that internal representation learning is more consistent and less sensitive to mini-batch or epoch-to-epoch noise within that subject's data. In Figure 5(c), the stable high plateau suggests that the subject's data are much easier to model. It yields robust, low-variance accuracy across epochs and is likely to hold across resamples of the same subject's trials.

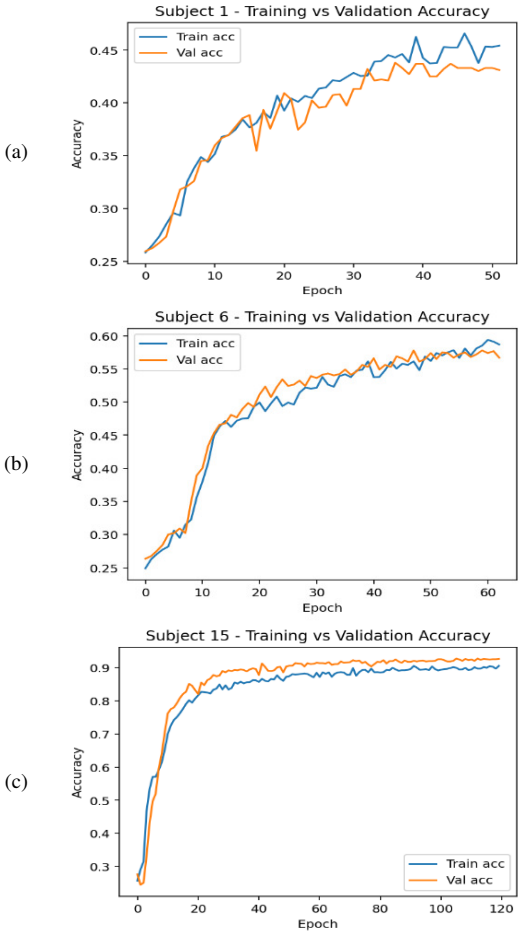

Fig. 5. Subject-wise learning curves of training and validation accuracy versus epochs for Subjects (a) 1, (b) 6, and (c) 15.

IV. CONCLUSION

This study provided a thorough within-subject analysis of EEGNet on the SEED-IV dataset for four-class emotion recognition, testing each of the 15 subjects' data individually with the same preprocessing, architecture, and training. The paper has confirmed three points described in the abstract, each supported by quantitative results from the experiments.

- The subject clustering analysis reveals two distinct performance tiers among the 15 subjects. A high-performing group, including Subjects 2, 4, 14, and 15, achieves test accuracies of 70.96%- 89.20%, with macro-precision and recall scores consistently above 70%. The low-performing group includes Subjects 1, 11, and 12 and records test accuracies between 43.81% and 48.27%. The 45.39% gap between the lowest (Subject 11, 43.81%) and highest (Subject 15, 89.20%) test accuracies, despite identical model architecture and training procedures, confirms that subject-specific neural characteristics are the dominant factor governing classification performance. This finding is further supported by the t-SNE visualizations, which qualitatively show distinct levels of class separation across subjects, and is independently validated by the quantitative per-subject metrics in Table I.
- The per-class recall analysis shows that Neutral is the most consistently recognized emotion, achieving the highest recall in 8 of the 15 subjects. Fear and sad exhibit the highest mutual confusion: in Subject 6, for instance, 67 sad instances are misclassified as fear and 82 fear instances are misclassified as sad, reflecting overlapping EEG representations for these two negative affect states. Happy is the most variable class, with per-class recall ranging from 29.0% (Subject 1) to 85.9% (Subject 15).
- THe results validate that raw EEG waveforms can achieve competitive performance, as EEGNet operates directly on 400-ms chunk-normalized EEG segments without manual feature extraction, such as differential entropy or power spectral density. Despite this, the model achieves test accuracies up to 89.20% on SEED-IV within the subject-dependent setting, which is competitive with methods that rely on hand-crafted frequency-domain features.

A few limitations affect the generalizability of these findings. First, the within-subject evaluation protocol restricts each model to training and testing on data from a single participant, and thus the reported results do not reflect cross-subject generalization, which remains a more challenging and open problem. Second, the analysis is confined to the SEED-IV dataset, comprising recordings from 15 subjects collected under controlled laboratory conditions using emotion-eliciting film clips. Third, the use of non-overlapping 400 ms segments increases the number of training samples but assumes that emotional states are adequately captured within short temporal windows, and alternative windowing strategies may influence performance. Future work will explore cross-subject transfer learning, domain adaptation, and validation on additional EEG emotion datasets.

REFERENCES

[1] R. B. Jadekar, P. Basavaraju, S. B. S. Kumar, and M. Rafi, "Emotion Recognition from EEG Signals Using Principal Component Analysis and Random Forest Classifier," *Engineering, Technology & Applied Science Research*, vol. 15, no. 5, pp. 27300–27305, Oct. 2025, https://doi.org/10.48084/etasr.12301.

[2] L. Monish and S. G. Shaila, "A Hybrid Fuzzy CNN–LSTM Approach for Emotion Recognition from EEG–ECG Physiological Signals," *Engineering, Technology & Applied Science Research*, vol. 15, no. 6, pp. 30405–30411, Dec. 2025, https://doi.org/10.48084/etasr.14864.

[3] M. Jehosheba Margaret and N. M. Masoodhu Banu, "Performance analysis of EEG based emotion recognition using deep learning models," *Brain-Computer Interfaces*, vol. 10, no. 2–4, pp. 79–98, Oct. 2023, https://doi.org/10.1080/2326263X.2023.2206292.

[4] K. Henni, N. Mezghani, A. Mitiche, L. Abou-Abbas, and A. Benazza-Ben Yahia, "An Effective Deep Neural Network Architecture for EEG-Based Recognition of Emotions," *IEEE Access*, vol. 13, pp. 4487–4498, 2025, https://doi.org/10.1109/ACCESS.2025.3525996.

[5] Z. Wang, J. Yu, J. Gao, Y. Bai, and Z. Wan, "MutaPT: A Multi-Task Pre-Trained Transformer for Classifying State of Disorders of Consciousness Using EEG Signal," *Brain Sciences*, vol. 14, no. 7, Jul. 2024, Art. no. 688, https://doi.org/10.3390/brainsci14070688.

[6] J. Pan *et al.*, "ST-SCGNN: A Spatio-Temporal Self-Constructing Graph Neural Network for Cross-Subject EEG-Based Emotion Recognition and Consciousness Detection," *IEEE Journal of Biomedical and Health Informatics*, vol. 28, no. 2, pp. 777–788, Oct. 2024, https://doi.org/10.1109/JBHI.2023.3335854.

[7] Y. Xu, Y. Gao, Z. Zhang, and S. Du, "Emotional recognition of EEG signals utilizing residual structure fusion in bi-directional LSTM," *Complex & Intelligent Systems*, vol. 11, no. 1, Dec. 2024, Art. no. 62, https://doi.org/10.1007/s40747-024-01682-y.

[8] S. Koelstra *et al.*, "DEAP: A Database for Emotion Analysis ;Using Physiological Signals," *IEEE Transactions on Affective Computing*, vol. 3, no. 1, pp. 18–31, Jan. 2012, https://doi.org/10.1109/T-AFFC.2011.15.

[9] S. Kulkarni and P. Patil, "EEG-based Emotion Recognition based on DEAP Dataset with Genetic Algorithm Augmented Multi-Layer Perceptron," in *2023 OITS International Conference on Information Technology (OCIT)*, Raipur, India, Sep. 2023, pp. 687–692, https://doi.org/10.1109/OCIT59427.2023.10431149.

[10] R. D. Gaddanakeri, M. M. Naik, S. Kulkarni, and P. Patil, "Analysis of EEG Signals in the DEAP Dataset for Emotion Recognition using Deep Learning Algortihms," in *2024 IEEE 9th International Conference for Convergence in Technology (I2CT)*, Apr. 2024, pp. 1–7, https://doi.org/10.1109/I2CT61223.2024.10543369.

[11] S. Bagherzadeh, A. Shalbaf, A. Shoeibi, M. Jafari, R.-S. Tan, and U. R. Acharya, "Developing an EEG-Based Emotion Recognition Using Ensemble Deep Learning Methods and Fusion of Brain Effective Connectivity Maps," *IEEE Access*, vol. 12, pp. 50949–50965, 2024, https://doi.org/10.1109/ACCESS.2024.3384303.

[12] W.-L. Zheng, W. Liu, Y. Lu, B.-L. Lu, and A. Cichocki, "EmotionMeter: A Multimodal Framework for Recognizing Human Emotions," *IEEE Transactions on Cybernetics*, vol. 49, no. 3, pp. 1110–1122, Mar. 2019, https://doi.org/10.1109/TCYB.2018.2797176.

[13] "SEED IV Dataset." https://bcmi.sjtu.edu.cn/home/seed/seed-iv.html.