

Fig.4: Labeling Methodology

# >>consider the delay for marking

# Comparison Study:

The classification of features across all four sets using various machine learning models—SVM, Random Forest, XGBoost, Gradient Boosting, KNN, and Logistic Regression—was conducted for 5 classes, each with a 1-second window.

Time features: Captured statistical properties and dynamics of EEG signals, including mean, variance, and signal changes. The best results obtained are detailed below:

| Set Number | Best Classifier | Best Accuracy | Best Precision | Best Recall | Best F1 Score |
| --- | --- | --- | --- | --- | --- |
| 1 | Random Forest | 0.553633 | 0.527038 | 0.553633 | 0.537613 |
| 2 | Gradient Boosting | 0.560554 | 0.543848 | 0.560554 | 0.550425 |
| 3 | XBoost | 0.482759 | 0.453314 | 0.482759 | 0.466427 |
| 4 | Random Forest | 0.544828 | 0.533142 | 0.544828 | 0.537486 |

Frequency Features: Derived spectral characteristics like mean power, peak frequency, and frequency bandwidth from EEG signals. The best results obtained are detailed below:

| Set Number | Best Classifier | Best Accuracy | Best Precision | Best Recall | Best F1 Score |
| --- | --- | --- | --- | --- | --- |
| 1 | XBoost | 0.572917 | 0.571039 | 0.572917 | 0.568508 |
| 2 | XBoost | 0.574394 | 0.559712 | 0.574394 | 0.565684 |
| 3 | XBoost | 0.551724 | 0.540166 | 0.551724 | 0.544609 |
| 4 | Gradient Boosting | 0.493103 | 0.464346 | 0.493103 | 0.475771 |

After fusing:  
1. Concatenation of Time and Frequency Features: Combined time and frequency-domain features into a comprehensive representation to capture diverse signal aspects.

| Set Number | Best Classifier | Best Accuracy | Best Precision | Best Recall | Best F1 Score |
| --- | --- | --- | --- | --- | --- |
| 1 | Random Forest | 0.570934 | 0.553388 | 0.570934 | 0.558619 |
| 2 | XGBoost | 0.565972 | 0.561388 | 0.565972 | 0.562153 |
| 3 | XGBoost | 0.5 | 0.475484 | 0.5 | 0.486508 |
| 4 | Gradient Boosting | 0.531034 | 0.514507 | 0.531034 | 0.520132 |

2. EEGNet: Extracted complex spatial and temporal patterns from raw EEG data using a specialized convolutional neural network.

| Set Number | Best Classifier | Best Accuracy | Best Precision | Best Recall | Best F1 Score |
| --- | --- | --- | --- | --- | --- |
| 1 | Gradient Boosting | 0.574394 | 0.573359 | 0.574394 | 0.570075 |
| 2 | XBoost | 0.582759 | 0.577831 | 0.582759 | 0.580123 |
| 3 | Gradient Boosting | 0.570934 | 0.565367 | 0.570934 | 0.567757 |
| 4 | Gradient Boosting | 0.562069 | 0.565155 | 0.562069 | 0.559727 |

3. Based on the above results, sets 1 and 2 appear to be effective labeling methods. Therefore, the autoencoder approach was applied using the best-performing model for further analysis.

| SVM | 0.6655 | 0.675 | 0.66 | 0.6675 |
| --- | --- | --- | --- | --- |
| Random Forest | 0.7334 | 0.74 | 0.73 | 0.735 |
| XGBoost | 0.7852 | 0.79 | 0.78 | 0.785 |
| Gradient Boosting | 0.7438 | 0.75 | 0.74 | 0.745 |
| Logistic Regression | 0.741 | 0.745 | 0.735 | 0.74 |

CNN features comparison:

For 5 classes classification, E2FNet showed the best result with Set 2 as Best almost in all classifiers:

| Set Number | Best Classifier | Best Accuracy | Best Precision | Best Recall | Best F1 Score |
| --- | --- | --- | --- | --- | --- |
| 1 | Random Forest | 0.878472 | 0.87951 | 0.878472 | 0.878441 |
| 2 | Gradient Boosting | 0.920139 | 0.91907 | 0.920139 | 0.919141 |
| 3 | Gradient Boosting | 0.896552 | 0.898954 | 0.896552 | 0.896494 |
| 4 | Random Forest | 0.882759 | 0.882313 | 0.882759 | 0.882304 |

For 2 classes classification:

TSception features for binary classification of Aha vs Not Aha, along with corresponding accuracy, precision, recall, and F1 score.

| Set Number | Best Classifier | Best Accuracy | Best Precision | Best Recall | Best F1 Score |
| --- | --- | --- | --- | --- | --- |
| 1 | Gradient Boosting | 0.862069 | 0.8625 | 0.862069 | 0.862028 |
| 2 | Gradient Boosting | 0.895652 | 0.897515 | 0.895652 | 0.895494 |
| 3 | Gradient Boosting | 0.801724 | 0.802534 | 0.801724 | 0.801591 |
| 4 | Gradient Boosting | 0.862069 | 0.8625 | 0.862069 | 0.862028 |

TSception features for binary classification of Impasse vs Not Impasse, along with corresponding accuracy, precision, recall, and F1 score.

| Set Number | Best Classifier | Best Accuracy | Best Precision | Best Recall | Best F1 Score |
| --- | --- | --- | --- | --- | --- |
| 1 | Gradient Boosting | 0.93087 | 0.931356 | 0.93087 | 0.930947 |
| 2 | Random Forest | 0.937093 | 0.937064 | 0.937093 | 0.936965 |
| 3 | Gradient Boosting | 0.944741 | 0.945444 | 0.944741 | 0.944533 |
| 4 | KNN | 0.926724 | 0.927598 | 0.926724 | 0.926279 |

E2FNet with Aha vs Not Aha: the Fusion of Spatio-Temporal features is highly effective for distinguishing between "Aha" and "Not Aha" states

| Set Number | Best Classifier | Best Accuracy | Best Precision | Best Recall | Best F1 Score |
| --- | --- | --- | --- | --- | --- |
| 1 | SVM/Random  Forest/XGBoost/Lo  gistic Regression | 0.938596 | 0.945419 | 0.938596 | 0.938431 |
| 2 | SVM | 0.939655 | 0.946154 | 0.939655 | 0.939435 |
| 3 | SVM | 0.921053 | 0.931818 | 0.921053 | 0.920557 |
| 4 | Gradient Boosting | 0.920139 | 0.91907 | 0.920139 | 0.919141 |

E2FNet with Impasse vs Not Impasse: the Fusion of Spatio-Temporal features is highly effective for distinguishing between "Impasse" and "Not Impasse" states:

| Set Number | Best Classifier | Best Accuracy | Best Precision | Best Recall | Best F1 Score |
| --- | --- | --- | --- | --- | --- |
| 1 | Random  Forest/XGBoost/Gr  adient Boosting | 0.982456 | 0.983061 | 0.982456 | 0.982456 |
| 2 | All except Fully Connected Layer and Logistic  Regression | 0.982456 | 0.983051 | 0.982456 | 0.982451 |
| 3 | All classifiers | 0.974138 | 0.97541 | 0.974138 | 0.974121 |
| 4 | XBoost | 0.965517 | 0.965517 | 0.965517 | 0.965517 |

Study:  
Based on the results, sets 1 and 2 demonstrated better performance compared to sets 3 and 4 labels.

Comparing the models’ performance across 5-class classification in set 2 for further highlights their relative effectiveness.

| Number of classes | Window Size | Method | Best Classifier | Best Accuracy | Best Precision | Best Recall | Best F1 Score |
| --- | --- | --- | --- | --- | --- | --- | --- |
| 5 | 1s | Time Features | Gradient  Boosting | 0.560554 | 0.543848 | 0.560554 | 0.550425 |
| 5 | 1s | Frequency  Features | XBoost | 0.574394 | 0.559712 | 0.574394 | 0.565684 |
| 5 | 1s | Time +  Frequency  Features  Concatenation | Random Forest | 0.570934 | 0.553388 | 0.570934 | 0.558619 |
| 5 | 1s | EEGNet Features | Gradient  Boosting | 0.582759 | 0.577831 | 0.582759 | 0.580123 |
| 5 | 1s | Autoencoder  Approach | XBoost | 0.7852 | 0.79 | 0.78 | 0.785 |
| 5 | 1s | E2FNeT | Gradient Boosting | 0.920139 | 0.91907 | 0.920139 | 0.919141 |

Similarly for 2-class classification on set 2 and 1:

| Method | Classification Task | Best Classifier | Best Accuracy | Best Precision | Best Recall | Best F1 Score |
| --- | --- | --- | --- | --- | --- | --- |
| TSception | Aha vs Not Aha | Gradient Boosting | 0.895652 | 0.897515 | 0.895652 | 0.895494 |
| TSception | Impasse vs Not Impasse | Gradient  Boosting | 0.944741 | 0.945444 | 0.944741 | 0.944533 |
| E2FNet | Aha vs Not Aha | SVM | 0.939655 | 0.946154 | 0.939655 | 0.939435 |
| E2FNet | Impasse vs Not  Impasse | Random  Forest/XGBoo  st/Gradient  Boosting | 0.982456 | 0.983061 | 0.982456 | 0.982456 |

Todo:  
1. Subject-Independent Analysis

2. Cross-Validation

3. Share Git repo - code, final report

Future directions:

1. Combine physiological signals and EEG signals