Chapter 4

Results

In this chapter, the classification results of four-class emotional states in each stage mentioned in the process pipeline, such as optimization of EDA components and EDA decomposition methods, optimization of segments (First-half, Second-half, and Whole phasic signals), and finally, optimization of the windowing approach on optimized segments, are presented. The results are organized as follows: The characteristics of representative EDA signals in all the considered emotional groups are discussed in section 4.1. The results from decomposing EDA signals into tonic and phasic signals using cvxEDA and BayesianEDA methods, extracting temporal features, assessing statistical significance, and subsequently classifying emotions using SVM, RF, and XGB models are outlined in section 4.2. Section 4.3 presents the results of segment optimization using time-frequency representation spectrograms (STFT and MFC), feature extraction through GLCM and GLRLM, statistical significance assessment, and emotional classification using SVM, RF, and XGB. Finally, Section 4.4 details the results of window optimization using time-encoded image representations such as GASF, GADF, MTF, and RP, feature extraction through GLCM, GLRLM, FDTA, ZM, HM, and FOS, statistical significance assessment, and emotional classification using SVM, RF, and XGB.

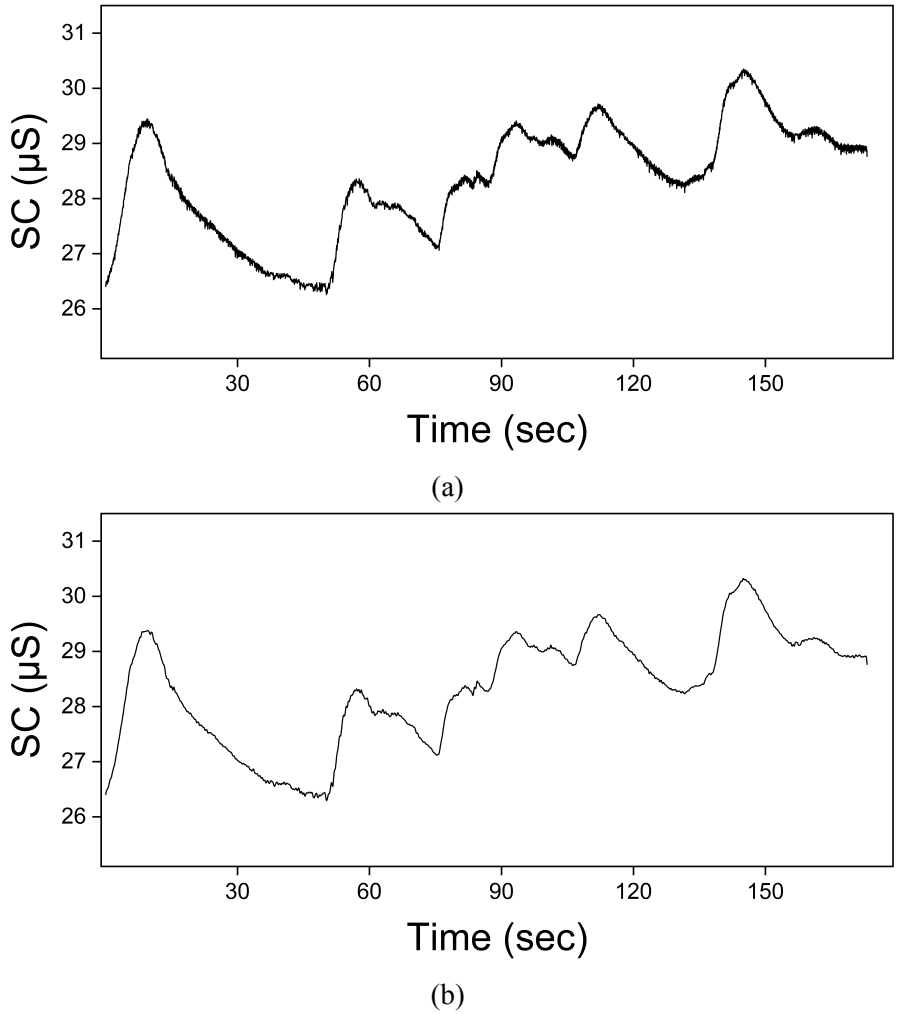

Figure 4.1: Representative EDA signals of a random subject's emotion: (a) raw signal and (b) low-pass filtered with a 2Hz cut-off frequency.

4.1 Reprentative of EDA Signals and their Analysis

This study uses a Butterworth low-pass filter with a 2Hz cutoff frequency to pre-process all the raw EDA signals (obtained from the CASE dataset) representing various emotions. Figure 4.1 displays the raw and filtered EDA signals for a randomly selected subject experiencing a specific emotional state. The raw EDA signal is affected by artifacts, as seen in Figure 4.1a, resulting in undesired high fluctuations. When selecting the frequency band of a low-pass filter, it's crucial to consider the frequency of EDA components, such as tonic and phasic, as it greatly influences emotional classification. Notably, the tonic signal ranges from 0.05 to 0.5 Hz, while the phasic component spans from 0.51 to 2 Hz (Ganapathy, Veeranki, and Swaminathan 2020). These signals undergo filtration to en-

sure accurate representation and facilitate further analysis in emotional classification. It's evident that the filtering technique employed in the study produces a cleaned and smooth filtered signal, as observed in Figure 4.1b. Additionally, it's worth noting that the cleaned signal maintains the original signal structure, potentially indicating preserved emotional information.

Figure 4.2a, 4.2c, 4.2e, 4.2g depicts the representative EDA signals from four distinct subjects experiencing the same emotional stimuli. Each subject exhibits unique amplitudes and fluctuations in their EDA signals. This variability can be attributed to subject-specific factors and may differ significantly from one individual to another. Key physiological parameters are pivotal in shaping these EDA patterns and their fluctuations across various emotional states. Factors such as sweat gland density, sweat production per gland, the density of nerve fibers around sweat glands, individual gland output, and conduction velocity all contribute to the intricate variations observed in EDA signals among different individuals. Normalization techniques were employed to ensure unbiased analysis for emotion classification despite these variations, converting the EDA signals to a standardized range of 0 to 1. This normalization process allows for fair comparison and interpretation of the normalized signals, as illustrated in Figure 4.2b, 4.2d, 4.2f, 4.2h.

Figure 4.3a, 4.3c, 4.3e, 4.3g displays the EDA signals of a random subject experiencing four distinct emotional states: amusing, boring, relaxing, and scary. These signals offer insights into the response of EDA to varying emotional stimuli. In emotions perceived as relaxing or boring, the EDA signals exhibit low variability and a declining trend, suggesting reduced physiological arousal. Conversely, in situations eliciting amusing, the EDA signals show minimal variability initially, followed by a gradual increase, indicating a rising arousal level corresponding to the heightened emotional response. However, when confronted with fear-inducing stimuli, the EDA signals display heightened variability and an upward trend, reflecting a significant increase in physiological arousal associated with the perception of danger. Skin conductance tends to be notably higher in response to fear than other emotions. This elevation in skin conductance might stem from the intensified

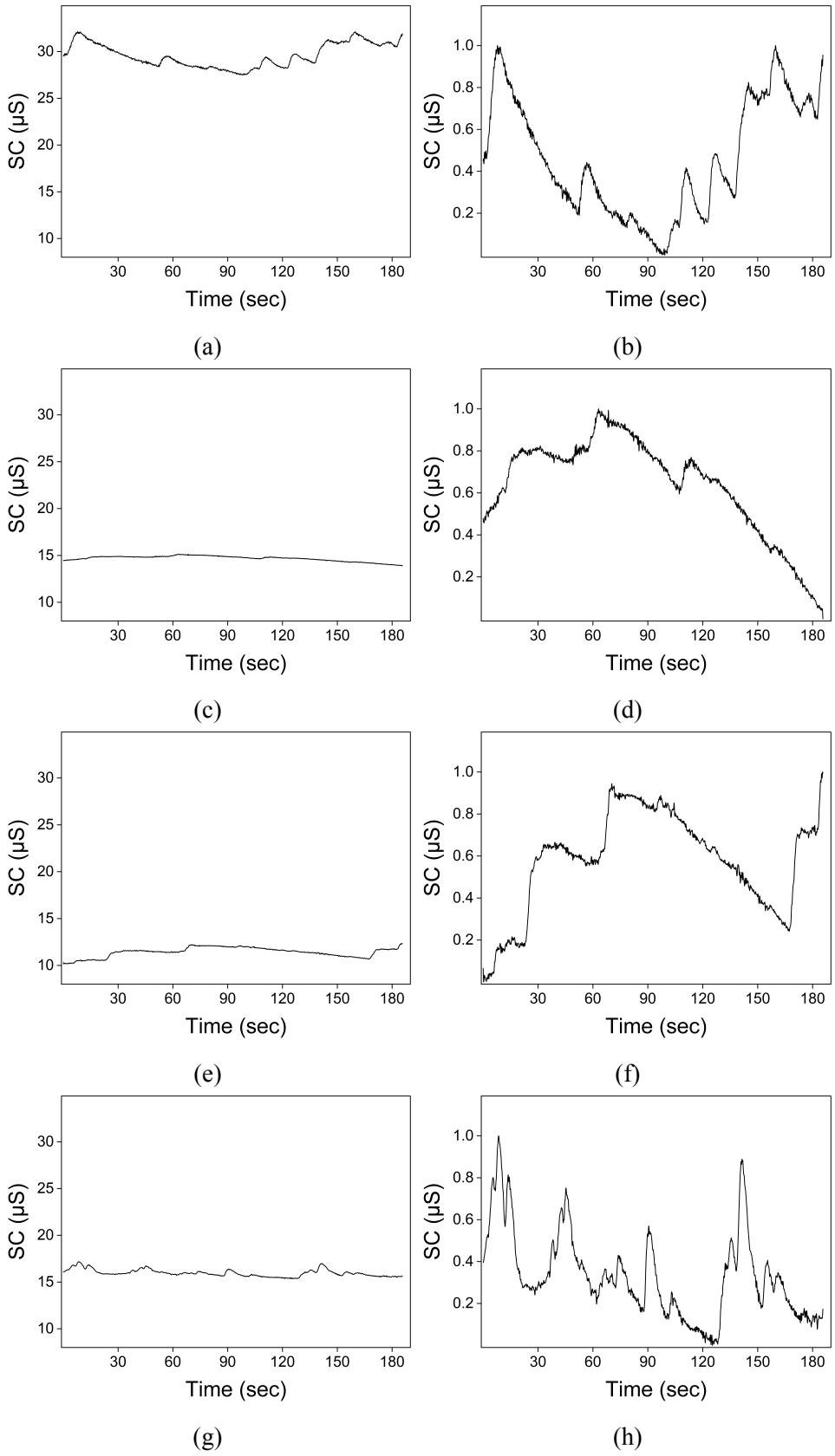

Figure 4.2: Representative EDA signals of four random subjects for the same emotional state (a, c, e, g) before normalization and (b, d, f, h) after normalization.

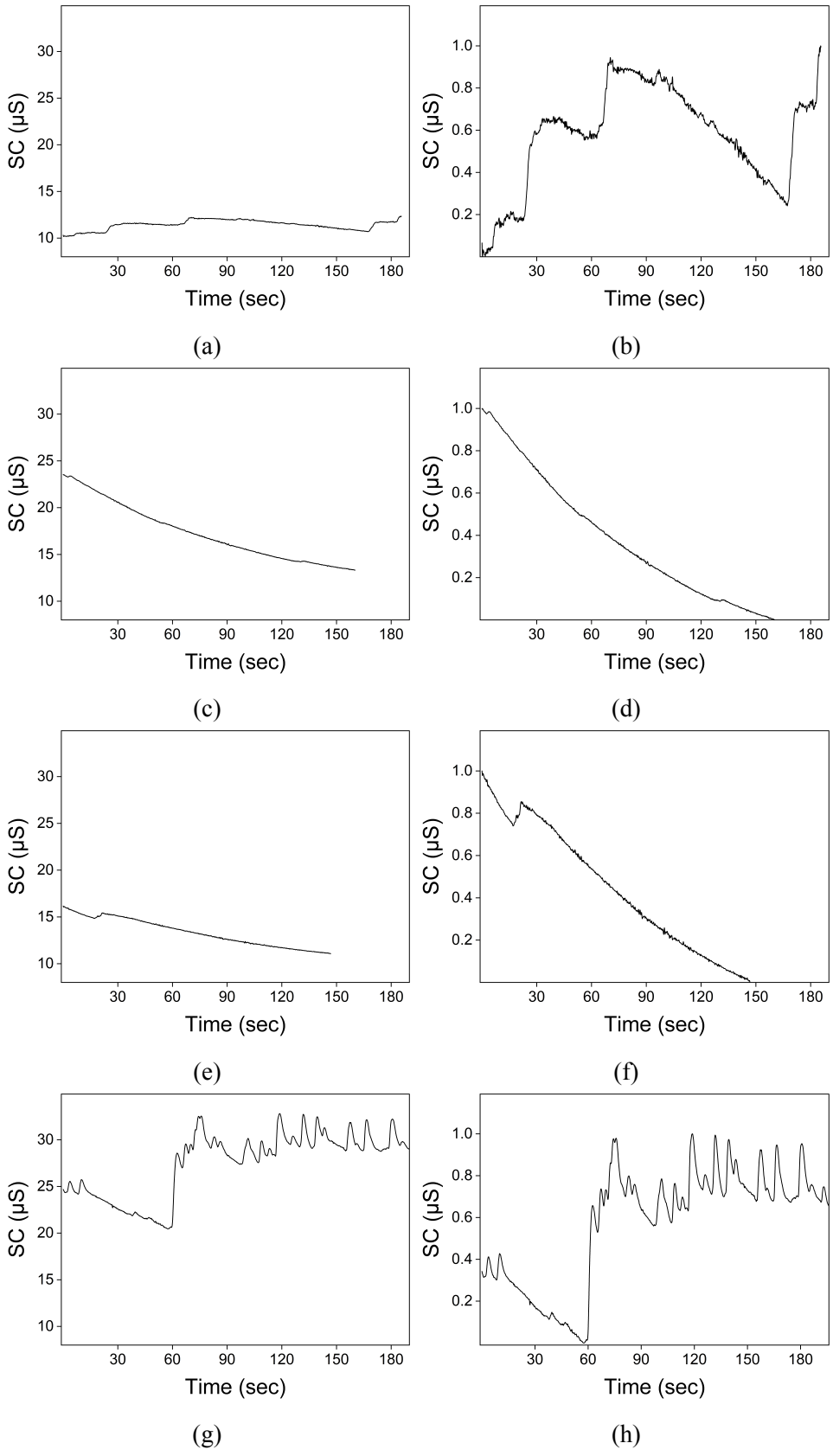

Figure 4.3: Representative EDA signals of various emotional states amusing, boring, relaxing, and scary: (a, c, e, g) before normalization and (b, d, f, h) after normalization.

activation of eccrine sweat glands triggered by the heightened emotional state induced by fear-inducing stimuli. The corresponding normalized EDA signals are shown in Figure 4.3b, 4.3d, 4.3f, 4.3h. Notably, the duration of the emotion-induced stimuli influences the length of the EDA signals across different emotions.

4.2 Optimization of EDA and Decomposition Methods

The cvxEDA decomposition method was utilized to extract the tonic and phasic components from the EDA signals, providing insight into the complex physiology underlying various emotions. Figure4.4 depicts the representative EDA signals and their corresponding tonic and phasic signals obtained through the cvxEDA decomposition method for four emotions such as amusing, boring, relaxing, and scary. Essentially, the tonic component mirrors the underlying EDA signal, preserving its low-pass frequency elements. This component essentially encapsulates the baseline arousal level associated with each emotion. Conversely, the phasic components capture the rapid fluctuations superimposed upon the tonic baseline, reflecting the dynamic response to emotional stimuli. The tonic signal has a higher magnitude than the phasic signal as it maintains the original signal magnitude levels of the EDA signals as shown in Figures 4.4b, 4.4e, 4.4h,4.4k. In contrast, the phasic signal has a lower magnitude, which indicates the fluctuations superimposed on the tonic signals. In the figures, notable distinctions can be observed among the emotional states. For instance, in Figure 4.4l the scary phasic signals have more pronounced fluctuations with higher magnitudes in skin conductance, suggesting a significant influence of fear as evidenced by increased sweat gland activity. Conversely, in Figure 4.4c, the amusing phasic signal displays relatively minor fluctuations and moderate magnitude levels of skin conductance. This reflects a more subdued physiological response. Interestingly, in the figures 4.4f, 4.4i, the boring and relaxing emotional states demonstrate no significant fluctuations in corresponding phasic signals. This suggests a very low magnitude level of physiological arousal, which might be because these emotions typically evoke less intense or stimulating responses than their more arousing counterparts.

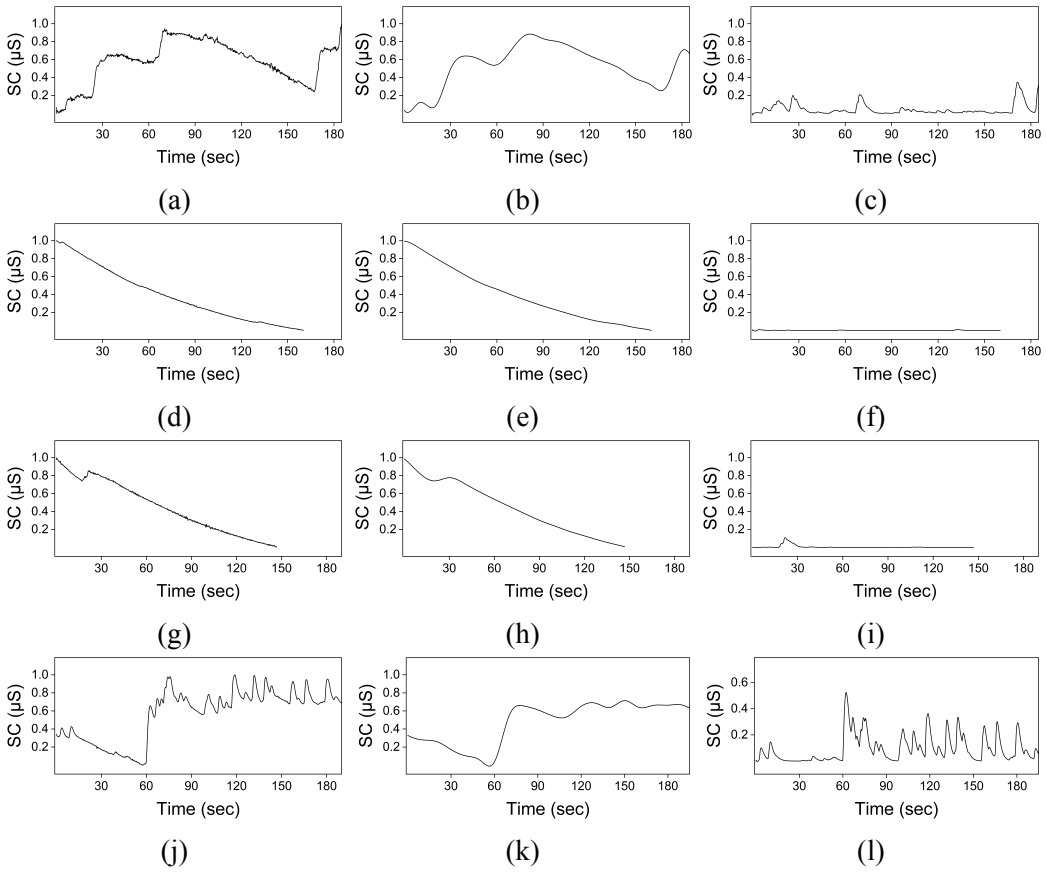

Figure 4.4: The representative signals of the amusing, boring, relaxing, and scary emotional states of a random subject (a, d, g, j) EDA, along with their corresponding decomposed components by the cvxEDA decomposition method (b, e, h, k) tonic and (c, f, i, l) phasic signals.

The BayesianEDA decomposition method was employed to extract the tonic and phasic components from the EDA signals corresponding to four emotions such as amusing, boring, relaxing, and scary. Figure 4.5 showcases the representative EDA signals alongside their respective tonic and phasic signals obtained through the BayesianEDA decomposition method. Notably, the tonic and phasic signals have been downsampled to 4 Hz, as this algorithm defaults to reducing the signal to this frequency while decomposing the tonic and phasic components. Like cvxEDA, the tonic signals have a higher magnitude than the phasic signal. Moreover, the tonic signals maintain the shape of the EDA signals as shown in Figure 4.5b, 4.5e, 4.5h, 4.5k. On the other hand, the phasic reflects the fluctuations superimposed on the tonic as shown in Figure 4.5c, 4.5f, 4.5i, 4.5l. These characteristics show that the BayesianEDA is able to decompose the EDA signals into tonic and phasic components. However, the tonic signal structure slightly varies from the original EDA

signal structure, which can be observed in boring and relaxing emotions. Moreover, the phasic signals do not capture all the fluctuations observed in the EDA signals, particularly evident in the scary and amusing. In summary, the phasic signals generated by the BayesianEDA approach exhibit fewer fluctuations and lower magnitude levels than the cvxEDA method.

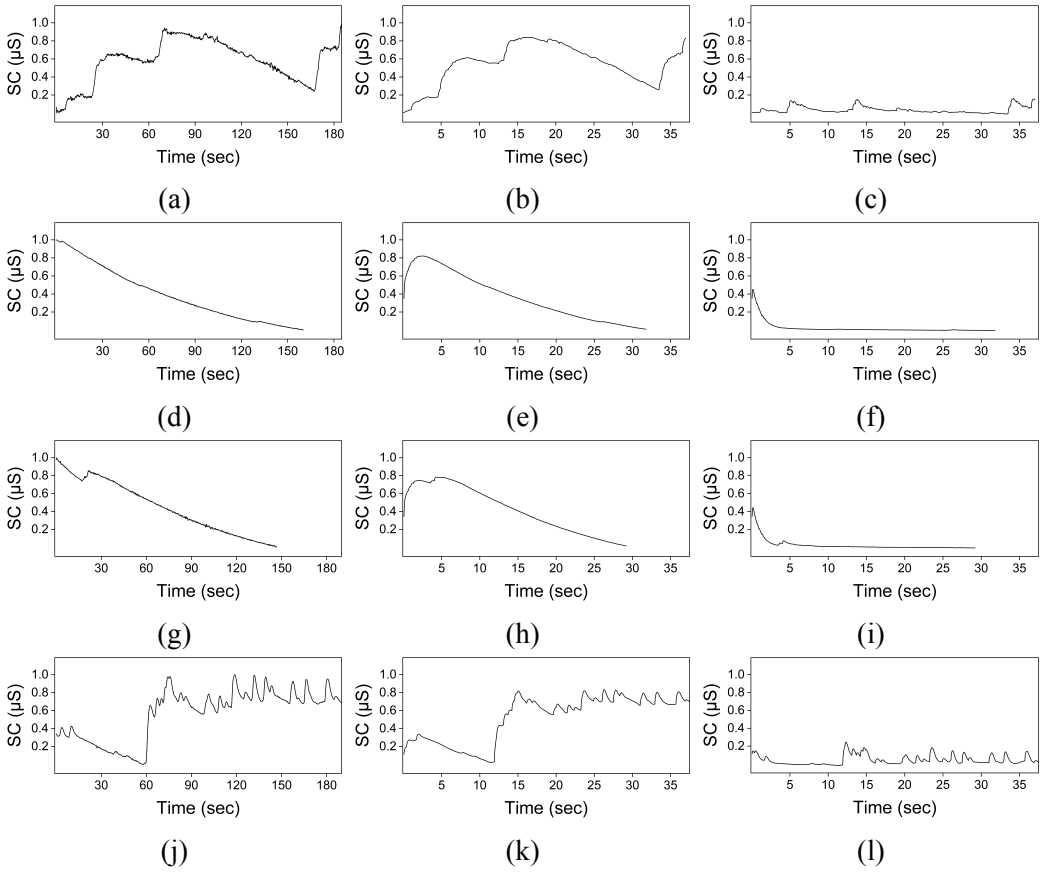

Figure 4.5: The representative signals of the amusing, boring, relaxing, and scary emotional states of a random subject (a, d, g, j) EDA, along with their corresponding decomposed components by the BayesianEDA decomposition method (b, e, h, k) tonic and (c, f, i, l) phasic signals.

4.2.1 Statistical Significance of temporal Features

We tested the normality of ten temporal features for four emotional states extracted from various signals, including EDA, cvxEDA tonic, cvxEDA phasic, BayesianEDA tonic, and BayesianEDA phasic, using the Kolmogorov–Smirnov test, which revealed that these features failed to follow a normal distribution. Hence, we used the KW test to assess the statistical significance of these features in discriminating between the four emotions, re-

jecting the null hypothesis at a significance level of 0.05. The results of the statistical significance of the temporal features calculated from the KW test are presented in Table 4.1.

Regarding features extracted from EDA signals, only four features, such as MN, MDN, MFD, and SSD, exhibited statistical significance, indicating their ability to differentiate between more than two classes, as indicated by $p < 0.05$. Conversely, the remaining features were insignificant. In contrast, both cvxEDA tonic and cvxEDA phasic signals displayed statistical significance for nine features (excluding SSD). For BayesianEDA signals, tonic and phasic components showed significance for eight features (excluding MIP and SSD in tonic; and STD and DR in phasic).

In summary, features extracted from cvxEDA tonic and phasic signals consistently demonstrated higher statistical significance compared to those derived from BayesianEDA signals, suggesting that the cvxEDA decomposition method effectively captures relevant physiological information related to emotional states. Furthermore, while both cvxEDA tonic and phasic signals exhibit significant features, the phasic signals tend to show greater significance overall, implying that the dynamic fluctuations captured by the phasic component play a crucial role in differentiating between emotional states, highlighting the importance of considering phasic components in EDA analysis.

Figure 4.6 depicts the box plot representation of the top eight significant temporal features extracted from cvxEDA phasic signals. Notably, amusing and scary emotions are discriminated from other emotional states by MFD, MSD, MAP, STD, and DR. Specifically, there is a distinct visual contrast between the scary and the remaining emotions and between the amusing and the others. However, MN, MIP, and MDN are able to discriminate only amusing from the other emotions. It can be noted that boring, relaxing, and scary are overlapping in these features. Additionally, the box plots corresponding to the MFD and MSD features show notably higher mean values for the scary emotion than the other emotions. Conversely, the box plots corresponding to the MAP, STD, DR, MN, MIP, and MDN features exhibit higher mean values for amusing than the other emotions. Even

Table 4.1: Statistical significance of features extracted from EDA, cvxEDA tonic, cvxEDA phasic, BayesianEDA tonic, and BayesianEDA phasic signals.

	EDA	cvxTonic	cvxPhasic	BysnTonic	BysnPhasic
MN	*	***	***	***	**
MDN	*	**	***	***	*
STD	NS	***	***	***	NS
MAP	NS	***	***	***	**
MIP	NS	***	***	NS	**
DR	NS	***	***	***	NS
MFD	***	***	***	***	***
MSD	NS	***	***	***	***
SFD	NS	**	***	***	***
SSD	***	NS	NS	NS	***

Notes:- cvxTonic and cvxPhasic represent the tonic and phasic signals derived from the cvxEDA decomposition method, while BysnTonic and BysnPhasic denote the tonic and phasic signals obtained through the BayesianEDA decomposition method; *$p <$ 0.05, *$p < 0.005$, *$p < 0.0005$, NS: Not significant.

though the considered features have the ability to discriminate scary and amusing emotions from other emotions, an overlap is observed between boring and relaxing emotions in the box plots.

Similarly, Figure 4.7 illustrates the box plot representation of the top eight significant temporal features extracted from BayesianEDA phasic signals, arranged from highest to lowest significance. The visualization clearly shows the differentiation among emotions across these features. Notably, the mean values of the scary emotion are higher than those of other emotions in features such as MFD, SFD, SSD, and MSD. Unlike cvxEDA, MFD and MDN features computed from BayesianEDA can be used to discriminate between boring and relaxing. Moreover, scary has higher mean values than amusing. Further,

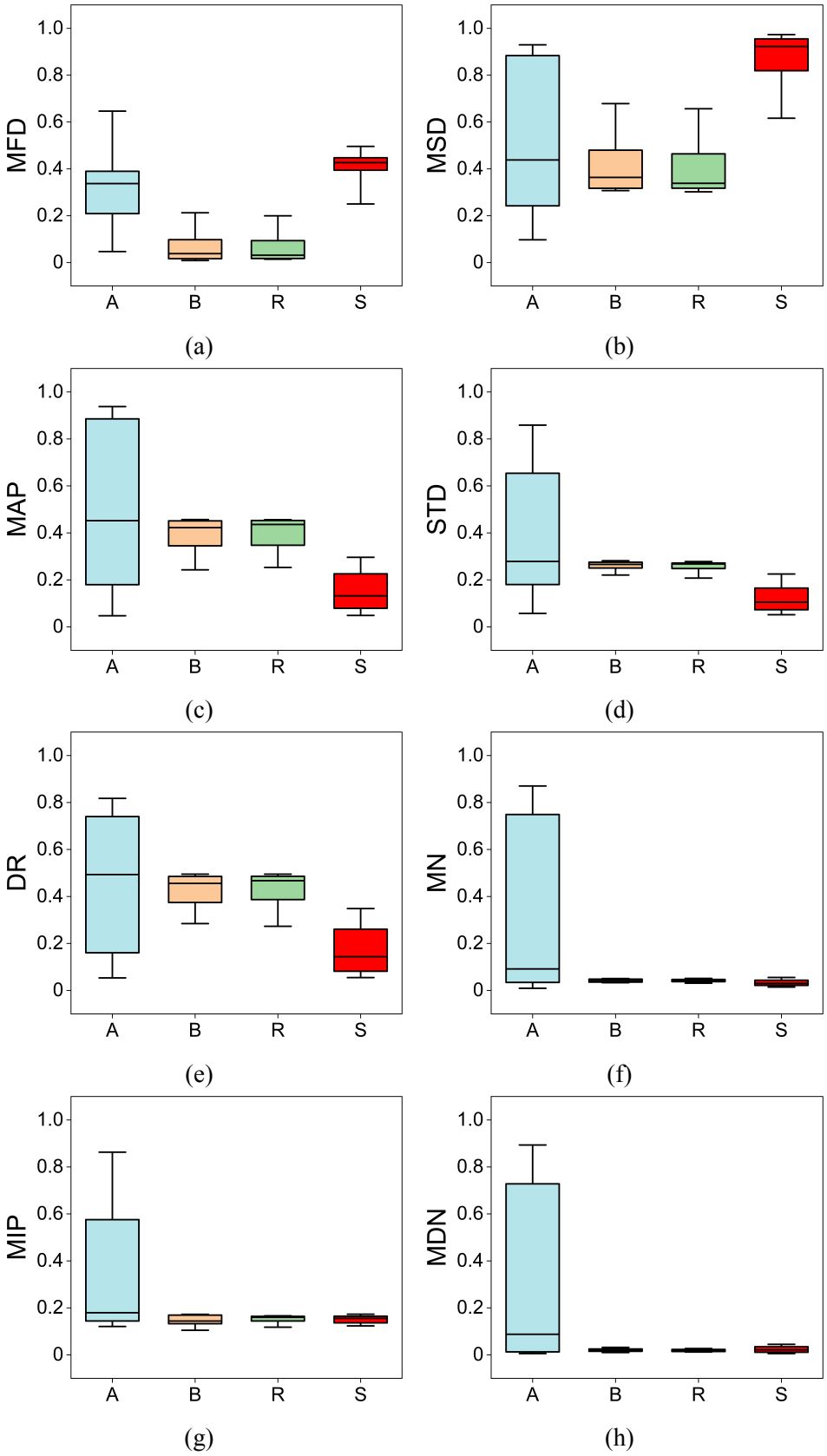

Figure 4.6: Representative box plots illustrating features extracted from cvxEDA phasic signal: (a) MFD, (b) MSD, (c) MAP, (d) STD, (e) DR, (f) MN, (g) MIP, and (h) MDN. Note: A stands for Amusing, B for Boring, R for Relaxing, and S for Scary. This convention applies to all subsequent box plots.

amusing emotions exhibit a wider distribution across values, particularly in the MFD feature. A noteworthy observation is the discrimination of amusing and scary emotions from other emotional states by features such as MFD, SFD, SSD, MIP, and MDN.

4.2.2 Classification Results for Optimization of EDA and Decomposition Methods

Classification Results for EDA Features

Table 4.2 presents the classification results for four emotional states using raw EDA signals with XGB, SVM, and RF classifiers. XGB achieved the highest performance across multiple metrics, including classification accuracy of 78.96%, sensitivity of 57.92%, specificity of 85.97%, precision of 58.93%, F1-score of 58.17%, and AUC of 81.82%. Followed by, RF and SVM classifiers yielded classification accuracies of 78.75% and 74.17%, respectively.

Table 4.2: Classification results of SVM, RF, and XGB using temporal features extracted from EDA signals for four-class emotions.

Metrics (%)	**SVM**	**RF**	**XGB**
Accuracy	74.17	78.75	78.96
Sensitivity	48.33	57.5	57.92
Specificity	82.78	85.83	85.97
Precision	48.32	57.56	58.93
F1-score	48.16	57.48	58.17
AUC	77.28	80.87	81.82

Classification Results for cvxEDA Tonic and Phasic Features

Table 4.3 presents the classification results for four emotional states utilizing tonic and phasic signals obtained through the cvxEDA decomposition method. The analysis is con-

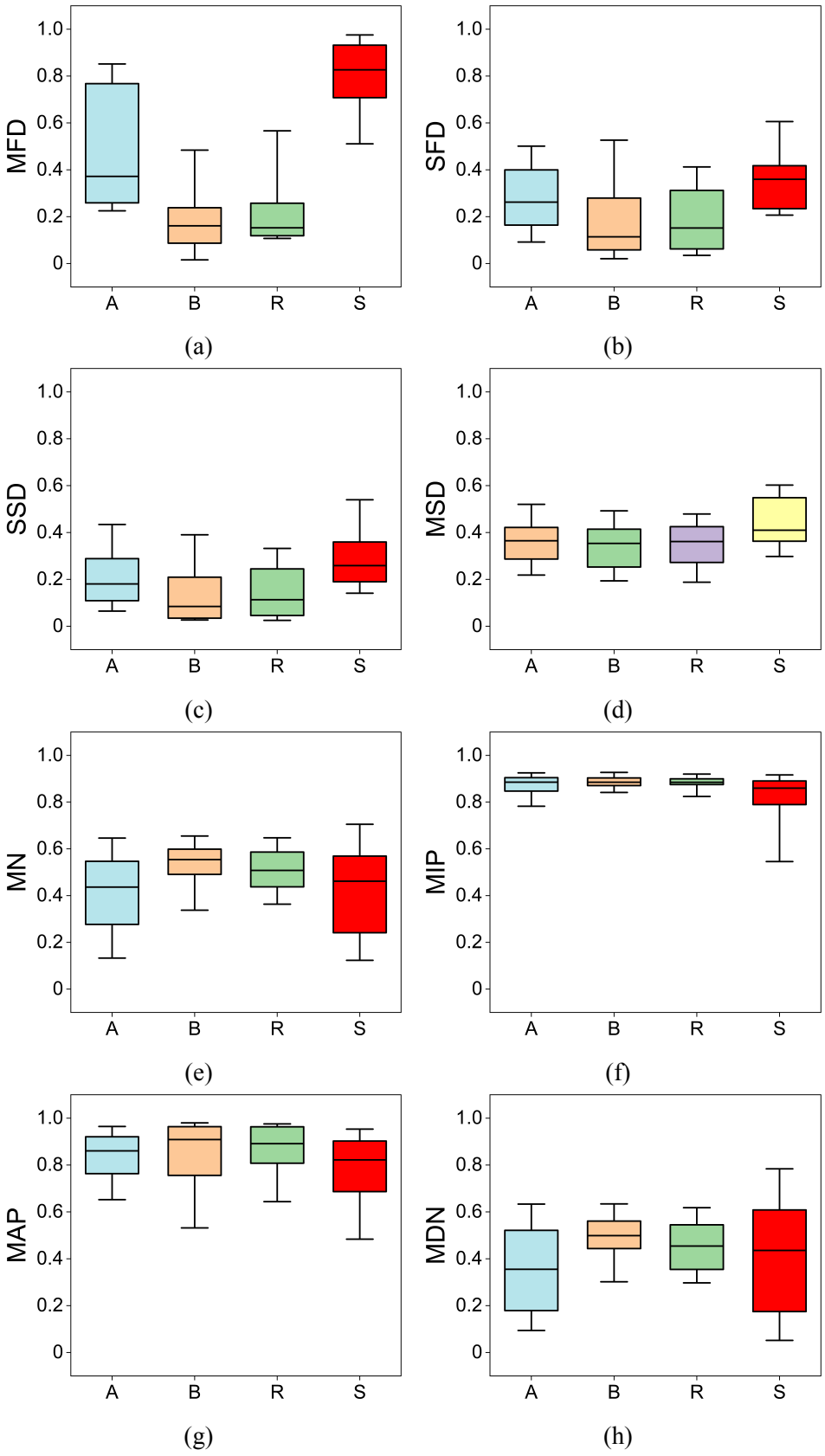

Figure 4.7: Representative box plots illustrating top eight significant features extracted from BayesianEDA phasic signal (a) MFD, (b) SFD, (c) SSD, (d) MSD, (e) MN, (f) MIP, (g) MAP, and (h) MDN.

ducted using SVM, RF, and XGB classifiers. For the tonic signals, XGB demonstrated the highest performance across multiple metrics, achieving an accuracy of 85.41%, sensitivity of 70.83%, specificity of 90.28%, precision of 71.39%, F1-score of 70.77%, and AUC of 86.53%. Following XGB, RF exhibited a classification accuracy of 85%, while SVM lagged with a lower accuracy of 79.38%. In contrast, phasic signals showcased superior performance compared to tonic signals across all classifiers: XGB continued to lead with an accuracy of 86.87%, sensitivity of 73.75%, specificity of 91.25%, precision of 74.22%, F1-score of 73.58%, and AUC of 89.92%. SVM and RF trailed behind but still performed well, with 86.25% and 85.21% classification accuracies, respectively.

Table 4.3: Classification results of SVM, RF, and XGB using temporal features extracted from tonic and phasic signals obtained by cvxEDA decomposition method for four-class emotions.

	cvxEDA Tonic			**cvxEDA Phasic**		
Metrics (%)	**SVM**	**RF**	**XGB**	**SVM**	**RF**	**XGB**
Accuracy	79.38	85	85.41	86.25	85.21	86.87
Sensitivity	58.75	70	70.83	72.5	70.42	73.75
Specificity	86.25	90	90.28	90.83	90.14	91.25
Precision	62.16	70.85	71.39	72.64	70.81	74.22
F1-score	53	70	70.77	72.56	70.31	73.58
AUC	80.82	86.51	86.53	86.92	83.18	89.92

Classification Results for BayesianEDA Tonic and Phasic Features

Table 4.4 displays the classification results for four emotional states using tonic and phasic signals derived from the BayesianEDA decomposition method, analyzed with SVM, RF, and XGB classifiers. For tonic signals: XGB achieved the highest classification metrics, including accuracy of 71.67%, sensitivity of 43.33%, specificity of 81.11%, precision of 41.69%, F1-score of 41.94%, and AUC of 67.13%. Following XGB, both RF and SVM

classifiers yielded the same classification accuracy of 69.79%. Conversely, phasic signals outperformed tonic signals across all classifiers: XGB led with an accuracy of 81.87%, sensitivity of 63.75%, specificity of 83.72%, precision of 63.5%, F1-score of 63.53%, and AUC of 85.73%. SVM and RF trailed with classification accuracies of 78.33% each.

Table 4.4: Classification results of SVM, RF, and XGB using temporal features extracted from tonic and phasic signals obtained by BayesianEDA decomposition method for four-class emotions.

	BayesianEDA Tonic			**BayesianEDA Phasic**		
Metrics (%)	**SVM**	**RF**	**XGB**	**SVM**	**RF**	**XGB**
Accuracy	69.79	69.79	71.67	78.33	78.33	81.87
Sensitivity	39.58	39.58	43.33	56.67	56.67	63.75
Specificity	79.86	79.86	81.11	85.56	85.56	83.72
Precision	34.37	38.26	41.69	56.40	56.08	63.5
F1-score	35.87	38.44	41.94	56.38	56.15	63.53
AUC	67.78	37.62	67.13	83.32	82.42	85.73

In summary, our analysis revealed that the phasic signals displayed notably higher efficacy when compared to the tonic signals. Moreover, our investigation found that the cvxEDA decomposition method consistently outshone the BayesianEDA decomposition method regarding performance metrics. Furthermore, the XGB model consistently demonstrated superior performance compared to the RF and SVM models. Notably, our findings observed a particularly pronounced significance in the temporal features, such as MFD, MN, and MDN, in all the considered cases, indicating their potential as key indicators in our analysis. We used the phasic EDA signals obtained through the cvxEDA for further analysis.

4.3 Optimization of Phasic Segments

Figure 4.8 illustrates the STFT and MFC spectrogram depictions of scary emotional cvxEDA phasic signals, encompassing the First-half, Second-half, and Whole phasic Signals. In the First-half phasic signal, a distinctive flattened shape persists until the midpoint and further emerges occasional fluctuations. This characteristic is discernible in the corresponding STFT and MFC spectrograms. Conversely, the Second-half phasic signal exhibits heightened fluctuations throughout, suggesting the subject has increased emotional experience during this interval compared to the initial phase. This heightened activity may be attributed to more activated eccrine sweat glands, and such fluctuations manifest in the STFT and MFC spectrograms, characterized by elevated energy bands during these periods. The Whole phasic signal is the concatenation of the First-half and Second-half phasic signals, and corresponding characteristics can be observed in the respective spectrograms.

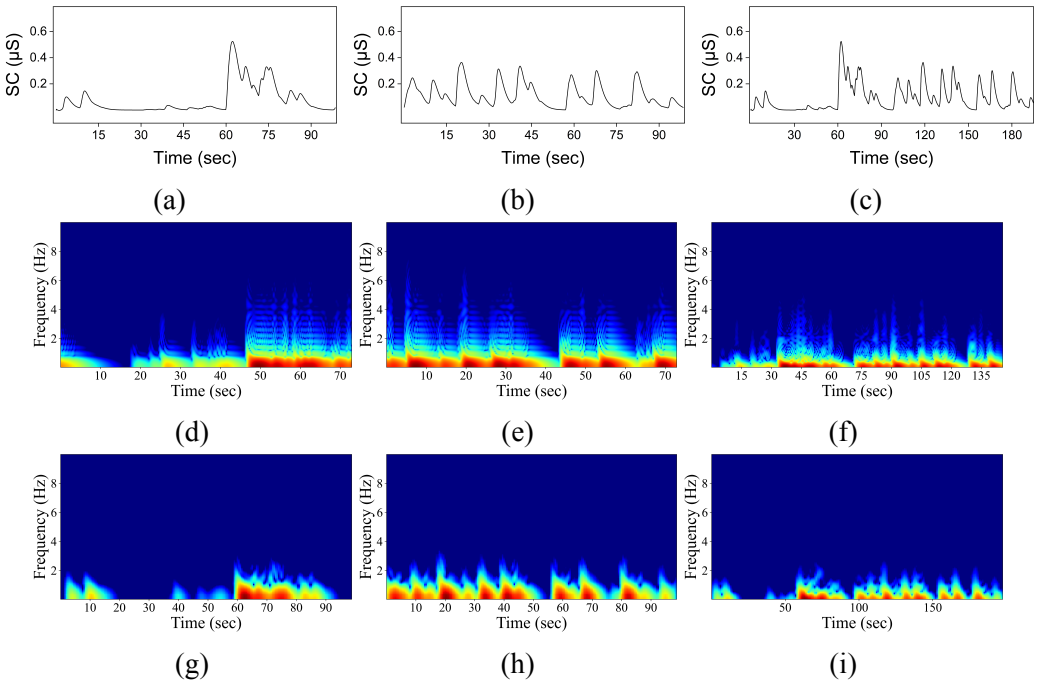

Figure 4.8: Spectrogram representation of Scary emotion: comparison of STFT and MFC for First-half, Second-half, and Whole phasic signals. (a) First-half phasic signal with corresponding (d) STFT and (g) MFC spectrograms.(b) Second-half phasic signal with corresponding (e) STFT and (h) MFC spectrograms. (c) Whole Phasic Signal with corresponding (f) STFT and (i) MFC spectrograms.

Figure 4.9 presents the STFT and MFC spectrogram representations of boring emotional cvxEDA phasic signals comprising the First-half, Second-half, and Whole phasic Signals.

Both the First-half and Second-half phasic signals display a consistent flattened pattern throughout, punctuated by minor fluctuations. These characteristics translate to reduced energy bands, evident in the STFT and MFC spectrograms. Similarly, the Whole phasic signal exhibits analogous traits, maintaining a steady profile with minimal variation, as reflected in the spectrograms.

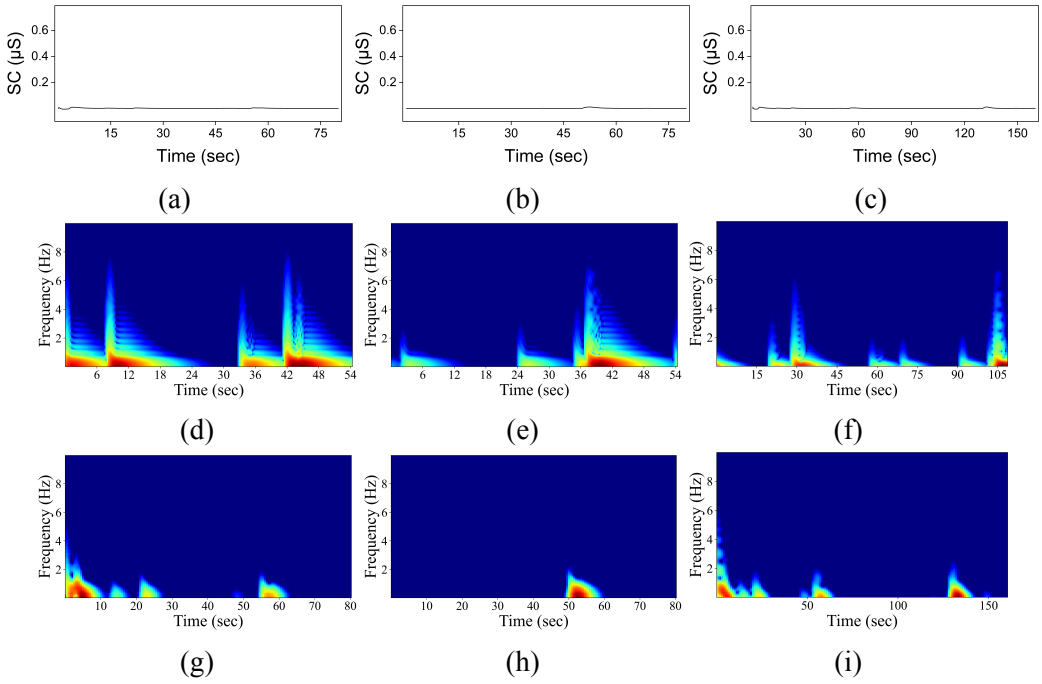

Figure 4.9: Spectrogram representation of Boring emotion: comparison of STFT and MFC for First-half, Second-half, and Whole phasic signals. (a) First-half phasic signal with corresponding (d) STFT and (g) MFC spectrograms.(b) Second-half phasic signal with corresponding (e) STFT and (h) MFC spectrograms. (c) Whole Phasic Signal with corresponding (f) STFT and (i) MFC spectrograms.

4.3.1 Statistical Significance of GLCM and GLRLM Features

Table 4.5 illustrates the KW test results to evaluate the statistical significance of 43 GLCM and GLRLM features for four emotional states derived from STFT spectrograms of three distinct phasic segments: First-half, Second-half, and Whole phasic. These features did not follow a normal distribution, validating using the KW test for the analysis. In the First-half of the phasic signal, 39 out of 43 features exhibited significance (p<0.05), with the top 5 significant features being SRE, MIMC1, RASM, MMCC, and RLU. Similarly, in the Second-half of the phasic signal, 42 features were significant, excluding the MCN feature. The top 5 significant features in this case were SRE, MMCC, RASM, RMCC, and

RLU. In the case of the Whole phasic signal, 42 features showed significance (p<0.05), with the MCN being an exception. The top 5 significant features were SRE, MIMC1, MMCC, RMCC, and RLU.

Similarly, the KW test assessed the statistical significance of 43 GLCM and GLRLM features for four emotional states extracted from MFC spectrograms of the same three phasic signals. We used the KW test as these features also did not follow a normal distribution. For the First-half phasic signal, only 10 features were significant out of 43, including MMCC, MIMC1, RIMC1, SRE, RMCC, RLU, contrast, RASM, homogeneity, and dissimilarity. In the case of the Second-half phasic signal, 32 features exhibited significance, while 11 failed to produce significance. The top 5 significant features were the MMCC, RMCC, MIMC1, and RLU. Finally, for the Whole phasic signal, 10 features were significant out of 43, including SRE, MIMC1, MMCC, RIMC1, RLU, contrast, RMCC, dissimilarity, homogeneity, and RASM.

In summary, the statistical test results suggest that most GLCM and GLRLM features hold statistical significance in emotional classification across various phasic segments, with minimal variations in feature importance between STFT and MFC spectrograms. Features like dissimilarity, homogeneity, MIMC1, MMCC, RASM, RIMC1, RMCC, SRE, and RLU consistently emerge as significant features across different signal segments and spectrogram types. However, notable differences in feature significance emphasize the importance of considering both signal segmentation and spectrogram representation methods in emotion classification tasks. The high number of significant features from the Second-half phasic segment of both STFT and MFC spectrograms suggests its suitability for emotion classification.

Figure 4.10 depicts the box plots of the top eight statistically significant GLCM and GLRLM features extracted from STFT spectrograms of Second-half phasic signals. Notably, amusing shows high variation in most features compared to other emotions, while scary appears squeezed, unlike earlier temporal feature analyses. Features like homogeneity, dissimilarity, and correlation discriminate between amusing and scary. Additionally,

Table 4.5: KW test for GLCM and GLRLM features extracted from STFT and MFC spectrograms for First-half, Second-half, and Whole phasic segments.

	STFT			MFC		
Feature	**FP**	**SP**	**WP**	**FP**	**SP**	**WP**
Contrast	***	***	***	***	NS	***
Dissimilarity	***	***	***	**	***	**
Homogeneity	***	***	***	**	***	**
Energy	*	**	**	NS	**	NS
Correlation	***	***	***	NS	*	NS
ASM	*	**	**	NS	NS	NS
Entropy	*	**	**	NS	NS	NS
Mean	*	***	***	NS	*	NS
Variance	**	***	***	NS	*	NS
SS	**	***	***	NS	*	NS
MASM	*	**	**	NS	NS	NS
MCN	***	NS	NS	NS	NS	NS
MCR	***	***	***	NS	*	NS
MSSV	**	***	***	NS	*	NS
MIDM	***	*	*	NS	**	NS
MSA	*	***	***	NS	*	NS
MSV	**	***	***	NS	*	NS
MSE	*	**	**	NS	NS	NS
ME	*	**	**	NS	NS	NS
MDV	***	**	*	NS	**	NS
MDE	***	**	*	NS	**	NS
MIMC1	***	***	***	***	***	***
MIMC2	*	***	**	NS	NS	NS
MMCC	***	***	***	***	***	***
RASM	***	***	***	***	***	*
RCN	***	***	***	NS	**	NS
RCR	***	***	***	NS	*	NS
RSSV	*	***	***	NS	NS	NS
RIDM	NS	*	*	NS	**	NS
RSA	*	***	***	NS	*	NS
RSV	***	***	***	NS	*	NS
RSE	*	***	***	NS	**	NS
RE	NS	*	**	NS	**	NS
RDV	NS	*	*	NS	*	NS
RDE	NS	*	*	NS	**	NS
RIMC1	**	***	***	***	***	***
RIMC2	*	***	***	NS	NS	NS
RMCC	***	***	***	***	***	***
SRE	***	***	***	***	***	***
LRE	**	*	*	NS	**	NS
GLU	***	**	**	NS	**	NS
RLU	***	***	***	***	***	***
RPC	***	**	**	NS	**	NS

Notes:- FP: First-half Phasic, SP: Second-half Phasic, and WP: Whole Phasic.

clear discrimination between boring and relaxing emotions is seen in features like SRE, RASM, RCN, homogeneity, and dissimilarity. Unlike the previous temporal feature analysis, Homogeneity and dissimilarity stand out for discriminating all four emotions, playing a crucial role in emotion classification. This highlights the effectiveness of the STFT spectrograms of Second-half phasic segments and GLCM and GLRLM features in identifying significant patterns corresponding to different emotions.

Figure 4.11 presents the box plots depicting the top eight statistically significant GLCM and GLRLM features extracted from MFC spectrograms of Second-half phasic signals. Features such as RMCC, MIMC1, RIMC1, RLU, homogeneity, and RASM contribute to distinguishing between amusing and scary emotions. Moreover, clear discrimination between boring and relaxing emotions is evident in features like RMCC, MIMC1, RLU, and SRE. Notably, RMCC, MIMC1, and RLU stand out for their ability to discriminate all four emotions, playing a crucial role in emotion classification. This underscores the effectiveness of GLCM and GLRLM features extracted from MFC spectrograms of Second-half phasic segments in identifying significant patterns corresponding to different emotions.

4.3.2 Classification Results for the Optimization of Phasic Segments

Classification Results for STFT-based Features

The comprehensive set of 43 GLCM and GLRLM features extracted from STFT spectrograms of First-half, Second-half, and Whole phasic signals were subsequently fed into SVM, RF, and XGB classifiers for four class emotional classification. The results are shown in Table 4.6. In the analysis of First-half phasic signals, SVM demonstrated the highest accuracy at 95.62%, along with sensitivity of 91.25%, specificity of 97.08%, precision of 91.72%, F1-score of 91.33%, and AUC of 97.44%. Following closely, XGB achieved an accuracy of 95.21%, while RF achieved 89.38%. Moving to the Second-half phasic signals, SVM again led with an accuracy of 97.08%, sensitivity of 94.17%, specificity of 98.06%, precision of 94.46%, F1-score of 94.22%, and AUC of 98.94%. XGB followed with an accuracy of 95.62%, and RF with an accuracy of 89.58%. In the case

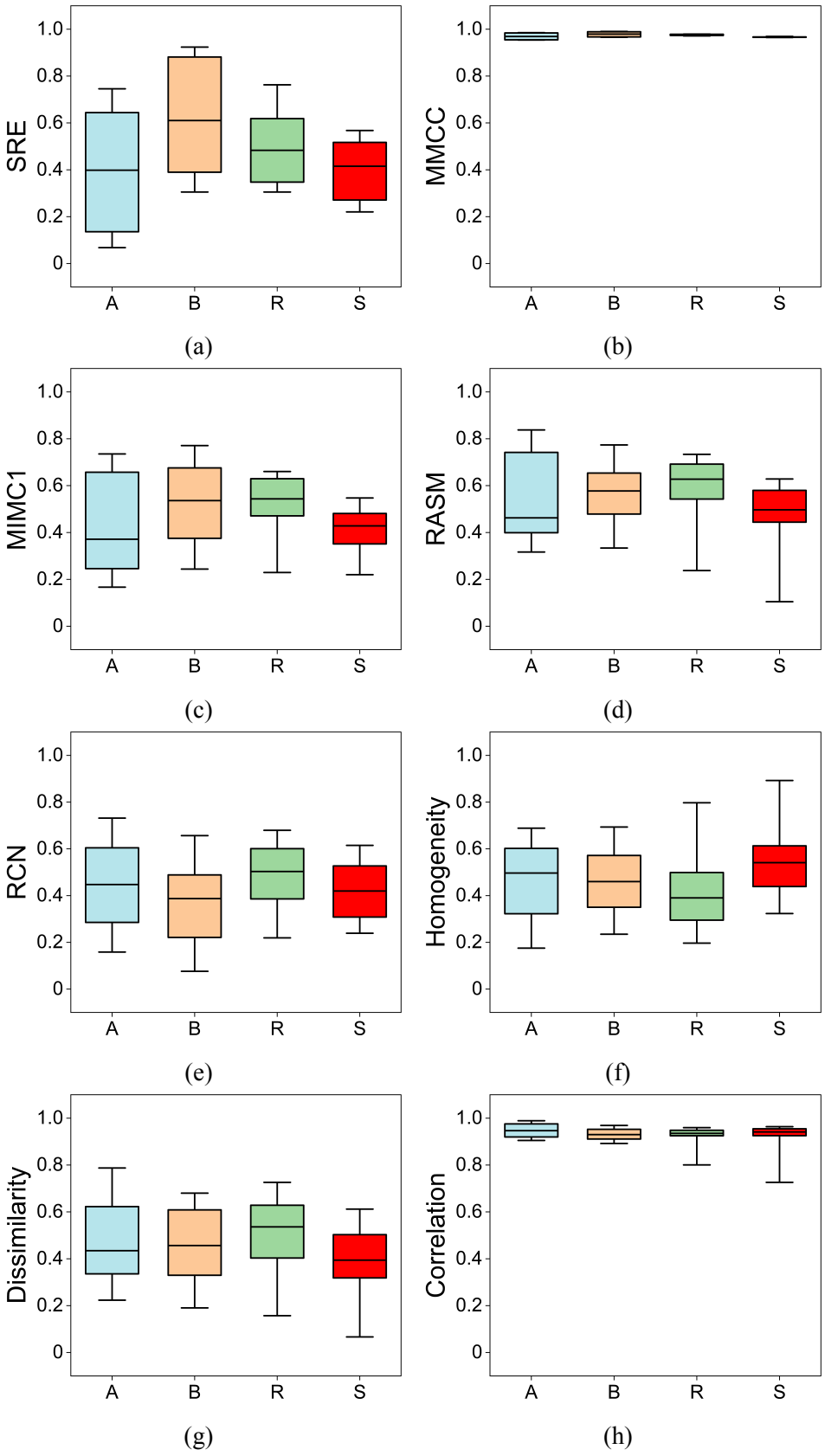

Figure 4.10: Representative box plots of top eight significant features extracted from STFT: (a) SRE, (b) MMCC, (c) MIMC1, (d) RASM, (e) RCN, (f) Homogeneity, (g) Dissimilarity, and (h) Correlation.

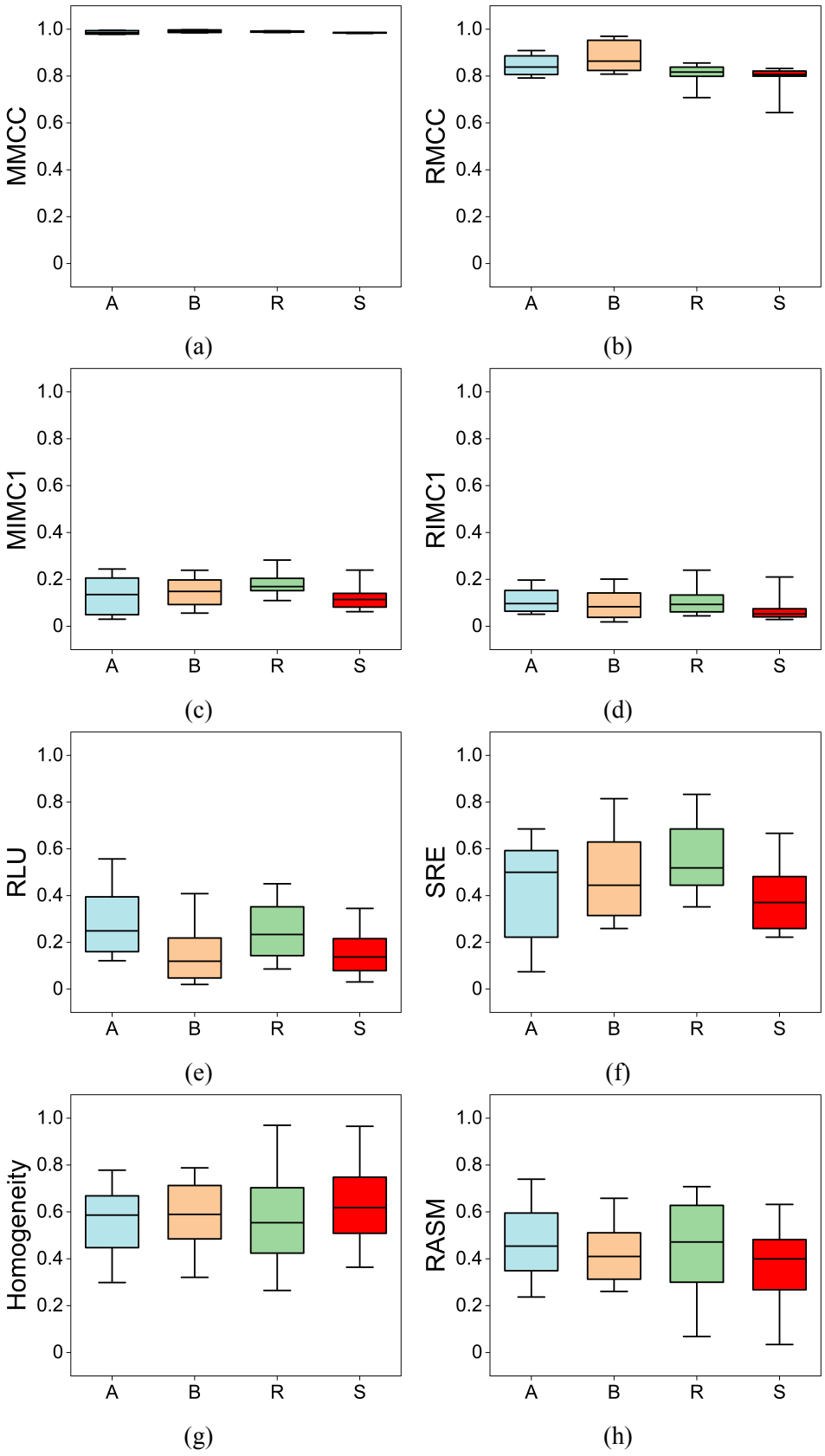

Figure 4.11: Representative box plots of top eight significant features extracted from MFC: (a) MMCC, (b) RMCC, (c) MIMC1, (d) RIMC1, (e) RLU, (f) SRE (g) Homogeneity, and (h) RASM.

of the Whole phasic signal, SVM attained an accuracy of 96.25%, sensitivity of 92.5%, specificity of 97.5%, precision of 92.63%, F1-score of 92.53% and AUC of 97.76%. However, the classification results were comparatively lower across XGB and RF classifiers, with 90.42% and 86.04%, respectively. It's evident that Second-half phasic signals consistently yield superior classification results across all classifiers compared to First-half and Whole phasic signals. This suggests that Second-half phasic signals possess enhanced discriminative features.

Table 4.6: Classification results of SVM, RF, and XGB using First-half, Second-half, and Whole phasic STFT.

	First-half Phasic			**Second-half Phasic**			**Whole Phasic**		
Metrics (%)	**SVM**	**RF**	**XGB**	**SVM**	**RF**	**XGB**	**SVM**	**RF**	**XGB**
Accuracy	95.62	89.38	95.21	97.08	89.58	95.62	96.25	86.04	90.42
Sensitivity	91.25	78.75	90.42	94.17	79.17	91.25	92.5	72.08	80.83
Specificity	97.08	92.92	96.81	98.06	93.06	97.08	97.5	90.69	93.61
Precision	91.72	79.27	90.76	94.46	79.59	91.41	92.63	73.17	80.93
F1-score	91.33	78.85	90.48	94.22	79.26	91.26	92.53	72.24	80.62
AUC	97.44	94.82	97.64	98.94	94.43	98.11	97.76	90.51	92.9

Classification Results for MFC-based Features

Table 4.7 shows the classification results of SVM, RF, and XGB models using the GLCM and GLRLM features extracted from MFC spectrograms. In the case of First-half phasic signals, SVM demonstrated the highest accuracy at 90.21%, along with sensitivity of 80.42%, specificity of 93.47%, precision of 80.89%, F1-score of 79.85%, and AUC of 93.92%. Following closely, XGB achieved an accuracy of 87.92%, while RF achieved an accuracy of 85%. In the case of the Second-half phasic signals, SVM again led with an accuracy of 91.88%, sensitivity of 83.75%, specificity of 94.58%, precision of 84.34%, F1-score of 83.93%, and AUC of 95.83%. XGB followed with an accuracy of 90.21%,

and RF with 87.08%. However, the classification results for Whole phasic signals were comparatively lower across SVM, RF, and XGB classifiers. XGB attained an accuracy of 89.79%, while SVM and RF achieved 87.71% and 86.46%, respectively. Notably, Second-half phasic signals consistently outperformed First-half and Whole phasic signals across different classifiers, highlighting their superior discriminative power and the importance of considering signal segmentation in emotion classification tasks.

Table 4.7: Classification results of SVM, RF, and XGB using First-half phasic, Second-half, and Whole phasic for MFC.

	First-half Phasic			**Second-half Phasic**			**Whole Phasic**		
Metrics (%)	**SVM**	**RF**	**XGB**	**SVM**	**RF**	**XGB**	**SVM**	**RF**	**XGB**
Accuracy	90.21	85	87.92	91.88	87.08	90.21	87.71	86.46	89.79
Sensitivity	80.42	70	75.83	83.75	74.17	80.42	75.42	72.92	79.58
Specificity	93.47	90	91.94	94.58	91.39	93.47	91.81	90.97	93.19
Precision	80.89	71.05	76.09	84.34	75.99	80.75	75.92	74	80.21
F1-score	79.85	70.2	75.67	83.93	74.29	80.43	75.49	72.94	79.69
AUC	93.92	88.29	90.8	95.83	90.85	93.53	91.56	92.22	93.35

In summary, the Second-half phasic signals consistently outperformed the First-half and Whole phasic signals in the analysis. STFT spectrograms showed impressive classification accuracies with the chosen classifiers, while MFC spectrograms, although slightly less accurate, still produced respectable results. Among the classifiers, SVM consistently outperformed RF and XGB. Unlike the analysis of temporal features, the GLCM and GLRLM features extracted from these spectrograms effectively identified individual emotions, particularly distinguishing between boring and relaxing states. Features such as homogeneity and dissimilarity were key in discriminating all four emotions, highlighting their importance in emotion classification.

4.4 Optimization of windowing approach

Figure 3.3 shows the 5 and 9 windowing approaches followed in the study. In the five-window approach, we partitioned the Second-half phasic signal into windows, each covering one-third (33.33%) of the signal with a 50% overlap, resulting in a total of five windows across the entire signal. Consequently, we obtained an array of (240 x 5) 1200 samples distributed across the five windows for each emotional state, containing 240 samples for the four emotions. For the nine-window approach, we partitioned the signal into windows, each covering one-fifth (20%) of the signal with a 50% overlap, resulting in a total of five windows across the entire signal. Consequently, for each emotional state, we obtained an array of 2160 samples distributed across the nine windows, with each window containing 240 samples for the four emotions.

Figure 4.12 showcases four distinct time-encoded representations applied to phasic signals with different fluctuation levels: low, medium, and high. GADF and GASF present patterns resembling an ogee or checkered-shaped pattern, with the complexity of the pattern increasing with higher signal fluctuation levels. These representations encode temporal dynamics into images by transforming time series data into matrices based on cosine and sine transformations. Meanwhile, MTF displays a clover-like pattern, with lower fluctuations represented by simpler grid structures and higher fluctuations by more intricate grids. MTF captures the transition probabilities between signal values, providing insights into the temporal dynamics of the signal.

Similarly, RP exhibits a plaid pattern, with lower fluctuations characterized by sparse square boxes and higher fluctuations by denser distributions of square boxes. RP visualizes recurrent patterns within the signal, highlighting repetitive structures and time dependencies. These visual representations serve as effective tools for capturing the temporal characteristics of phasic signals. However, it's crucial to note that the interpretation of these patterns may vary depending on the specific context or application. Additionally, while these representations offer valuable insights into signal dynamics, they may not directly convey specific emotional attributes. Thus, further analysis, such as feature

extraction and ML classification, may be necessary to discern emotional states accurately from these time-encoded images.

4.4.1 Statistical Significance of GLCM, GLRLM, FDTA, FOS, HM, and ZM Features

We conducted KW tests to evaluate the statistical significance of 85 GLCM, GLRLM, FDTA, FOS, HM, and ZM features for emotional classification across four time-encoded images: GADF, GASF, MTF, and RP, using both five-window and nine-window approaches on the Second-half phasic signals. The KW test was necessary because these features did not follow a normal distribution.

In the five-window approach, GADF revealed 37 significant features (p<0.05), GASF had 50, MTF identified 59, and RP found 53. With the nine-window approach, the number of significant features increased across all methods: GADF had 35 significant features, GASF identified 54, MTF yielded 73, and RP found 56.

These results highlight that features computed from the MTF method significantly discriminate between emotions. Notably, features such as dissimilarity, homogeneity, entropy, MIDM, ME, MDV, MDE, MIMC1, MMCC, RASM, RIDM, RSE, RE, RDV, RDE, RIMC1, RIMC2, RMCC, SRE, LRE, GLU, RLU, RPC, HC2, HC3, FOSERG, and FOSEN were significant in both five-window and nine-window approaches across all four time-encoded image methods. Additionally, features like FOSKRT, MIMC2, and HC4 showed significance, specifically in the nine-window approach.

Furthermore, GLRLM features played a crucial role in discriminating emotions, while HM and ZM features showed minimal significance. These findings underscore the impact of windowing on feature extraction, demonstrating that finer segmentation leads to a higher number of extracted features and provides more comprehensive insights into the analyzed data.

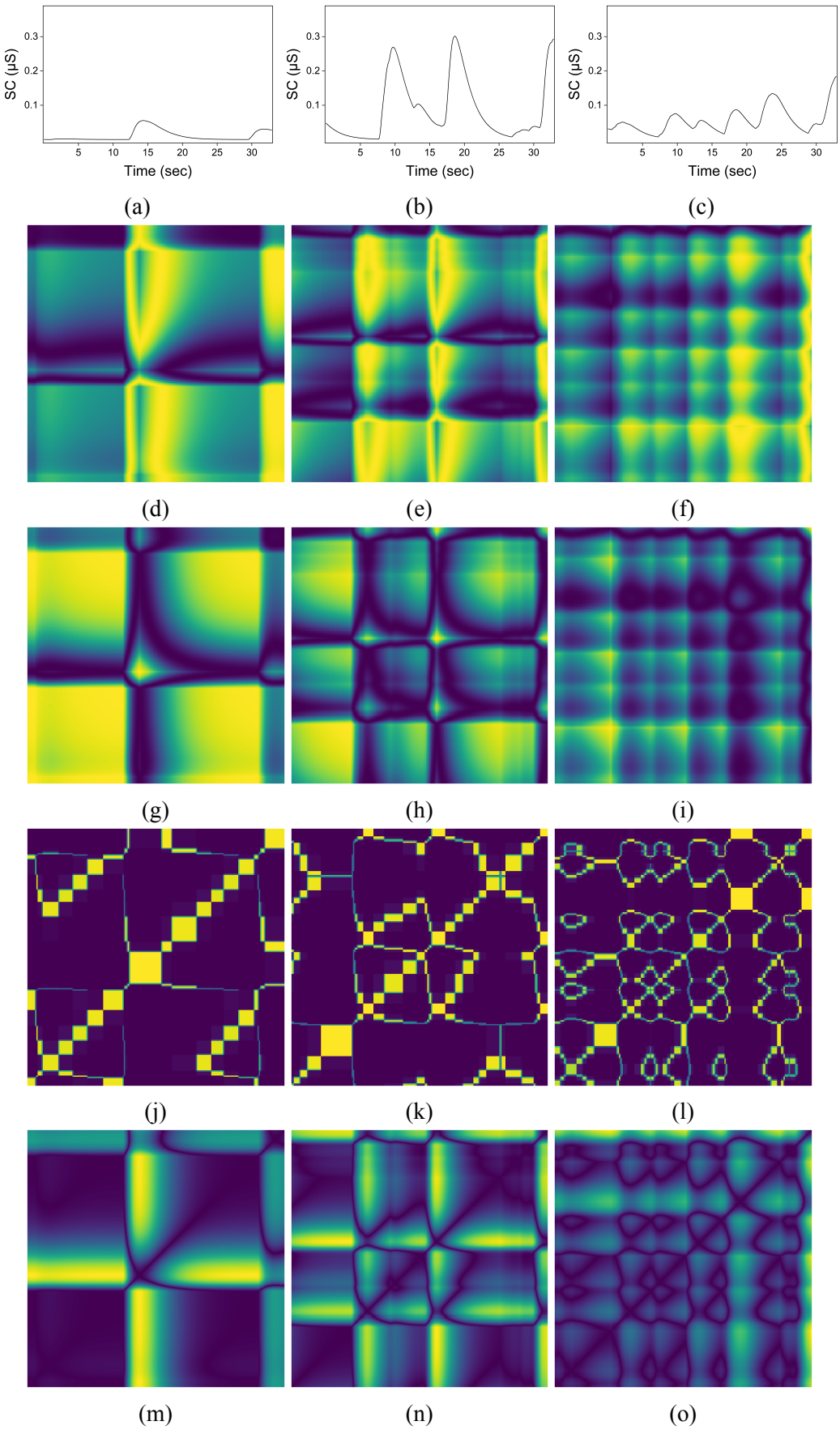

Figure 4.12: Representative (a) Low, (b) Medium, and (c) High fluctuations window signals and corresponding time-encoded images representation (d, e, f) GADF, (g, h, i) GASF, (j, k, l) MTF, and (m, n, o) RP.

Table 4.8: Significance of features using KW test: GADF, GASF, MTF, and RP across 5 and 9 window approaches

	5 Windows				9 Windows			
Feature	**GADF**	**GASF**	**MTF**	**RP**	**GADF**	**GASF**	**MTF**	**RP**
Contrast	NS	*	***	NS	NS	**	***	NS
Dissimilarity	***	***	***	***	***	***	***	***
Homogeneity	***	***	***	***	***	***	***	***
Energy	NS	*	***	***	NS	NS	***	***
Correlation	NS	*	***	***	NS	**	***	**
ASM	NS	*	***	***	NS	**	***	**
Entropy	***	***	***	***	***	***	***	***
Mean	*	NS	***	***	NS	NS	***	***
Variance	NS	*	***	*	NS	*	***	NS
SS	NS	*	***	*	NS	*	***	NS
MASM	NS	*	***	***	NS	*	***	***
MCN	***	*	***	NS	***	**	***	NS
MCR	*	*	***	***	***	**	***	**
MSSV	NS	*	***	*	NS	*	***	**
MIDM	***	***	***	***	***	***	***	***
MSA	*	NS	***	***	NS	*	***	***
MSV	NS	*	***	***	NS	*	***	**
MSE	NS	NS	***	***	***	*	***	***
ME	***	***	***	***	***	***	***	***
MDV	***	***	***	***	***	***	***	***
MDE	***	***	***	***	***	***	***	***
MIMC1	***	***	***	***	***	***	***	***
MIMC2	***	NS	***	***	***	**	***	***
MMCC	***	***	***	***	***	***	***	***

Continued on next page

Table 4.8 – continued from previous page

	5 Windows				9 Windows			
Feature	**GADF**	**GASF**	**MTF**	**RP**	**GADF**	**GASF**	**MTF**	**RP**
RASM	***	***	***	***	***	**	***	***
RCN	NS	*	***	NS	NS	**	***	NS
RCR	NS	*	***	***	NS	**	***	**
RSSV	NS	NS	***	***	NS	**	***	***
RIDM	***	***	***	***	***	***	***	***
RSA	*	NS	***	***	**	NS	***	***
RSV	NS	*	***	***	NS	*	***	***
RSE	**	**	***	***	***	***	***	***
RE	***	***	***	***	***	***	***	***
RDV	***	***	***	***	***	***	***	***
RDE	***	***	***	***	***	***	***	***
RIMC1	***	***	***	***	***	***	***	***
RIMC2	***	**	***	***	***	NS	***	***
RMCC	***	***	***	***	***	***	***	***
SRE	***	***	***	***	***	***	***	***
LRE	***	***	***	***	***	***	***	***
GLU	***	***	***	***	***	***	***	***
RLU	***	*	***	***	***	NS	***	***
RPC	***	***	***	***	***	***	***	***
HM1	NS	**	***	**	NS	NS	***	***
HM2	*	NS	***	NS	NS	NS	***	***
HM3	NS	NS	NS	NS	NS	NS	***	***
HM4	*	NS	NS	*	NS	NS	***	***
HM5	NS	NS	NS	NS	NS	NS	***	***
HM6	NS	NS	NS	NS	NS	NS	***	***

Continued on next page

Table 4.8 – continued from previous page

	5 Windows				9 Windows			
Feature	**GADF**	**GASF**	**MTF**	**RP**	**GADF**	**GASF**	**MTF**	**RP**
HM7	NS	NS	NS	NS	NS	NS	***	***
HC1	NS	NS	NS	NS	NS	NS	***	***
HC2	***	***	***	***	***	***	***	***
HC3	*	***	***	**	***	***	***	***
HC4	NS	*	***	***	**	**	***	***
FOSMN	*	***	***	***	**	***	***	***
FOSVRN	NS	**	***	**	NS	**	***	**
FOSSKW	NS	***	***	***	NS	***	***	***
FOSKRT	NS	***	***	***	NS	***	***	***
FOSERG	*	***	***	***	**	***	***	***
FOSEN	*	***	***	***	**	***	***	***
FOSCV	*	***	***	*	**	***	***	***
ZM1	NS	*	**	NS	NS	NS	NS	NS
ZM2	NS	NS	NS	NS	NS	**	***	NS
ZM3	NS	**	NS	*	NS	NS	***	NS
ZM4	NS	NS	*	NS	NS	*	**	NS
ZM5	NS	NS	*	NS	NS	NS	NS	NS
ZM6	NS	NS	NS	NS	NS	**	***	NS
ZM7	NS	NS	*	NS	NS	NS	***	NS
ZM8	NS	NS	NS	NS	NS	*	**	NS
ZM9	NS	NS	NS	NS	NS	**	***	NS
ZM10	NS	NS	NS	NS	NS	NS	***	NS
ZM11	NS	NS	NS	NS	NS	*	**	NS
ZM12	NS	NS	NS	NS	NS	NS	***	NS
ZM13	NS	NS	NS	NS	NS	NS	***	NS

Continued on next page

Table 4.8 – continued from previous page

	5 Windows				9 Windows			
Feature	**GADF**	**GASF**	**MTF**	**RP**	**GADF**	**GASF**	**MTF**	**RP**
ZM14	NS	NS	NS	NS	NS	NS	**	NS
ZM15	NS	NS	NS	NS	NS	NS	NS	NS
ZM16	NS	NS	NS	NS	NS	NS	NS	NS
ZM17	NS	NS	NS	NS	NS	NS	NS	NS
ZM18	NS	NS	NS	NS	NS	NS	NS	NS
ZM19	NS	NS	NS	NS	NS	NS	NS	NS
ZM20	NS	NS	NS	NS	NS	NS	NS	NS
ZM21	NS	NS	NS	NS	NS	NS	NS	NS
ZM22	NS	NS	NS	NS	NS	NS	NS	NS
ZM23	NS	NS	NS	NS	NS	NS	NS	NS
ZM24	NS	NS	NS	NS	NS	NS	NS	NS

Figure 4.13 illustrates the box plot representation of the top eight statistically significant features extracted from GADF images using the nine-window approach: RASM, MIMC2, RMIC2, MIMC1, MMCC, RIMC1, RMCC, and MIDM. Notably, the RASM feature distinctly discriminates the four emotions: amusing, boring, relaxing, and scary. This suggests that RASM could be a robust discriminator of emotional states. Visual inspection reveals negligible differences in the distribution of MIMC2, RIMC2, and MMCC values, as they predominantly cluster around values of either 0 or 1. Interestingly, features other than RASM struggle to differentiate between boring and relaxing emotions. This indicates their limited effectiveness in capturing emotional nuances.

Figure 4.14 shows the box plot representation of the top eight statistically significant features, such as MIDM, homogeneity, RIDM, MDV, MDE, RE, RDV, and RDE, extracted from GASF images using the nine-window approach. A similar pattern between MIDM and homogeneity can be observed, with minor variations in the values. This consistency

suggests these features might exhibit similar trends in representing emotional states. It can be observed that amusing and scary fall into different ranges, from boring and relaxing, in all the features. Moreover, the boring and relaxing emotions mostly overlap in each considered feature. This implies the challenges in accurately distinguishing between these two emotional states based solely on these features.

Figure 4.15 displays the box plot representation of the top eight statistically significant features, FOSSKW, MDV, RDV, MIMC1, MDE, RSA, mean, and MSA extracted from MTF images using the nine-window approach. A clear discrimination between the emotion in MIMC1 and MDE features can be observed. Further, it's evident that amusing and scary emotions share a common range, whereas boring and relaxing emotions are delineated into separate ranges in both MDV and RDV. This suggests that different emotions might manifest distinct patterns in specific feature spaces. It becomes evident that there are minimal disparities in the distribution patterns of FOSSKW, RSA, mean, and MSA values. These values predominantly cluster around either 0 or 1.

Figure 4.16 depicts the box plot representation of the top eight statistically significant features, RMCC, MMCC, RSE, SRE, MDV, RLU, RDV, and homogeneity extracted from MTF images using the nine-window approach. It can be observed that all the top eight features can discriminate between the four emotions. This suggests a promising potential for using RP image features in emotion recognition tasks.

4.4.2 Classification Results for the Optimization of Windowing Approach

Classification Results for GADF-based Features

Table 4.9 presents the classification results using GADF images for both five-window and nine-window approaches, employing SVM, RF, and XGB algorithms. For the five-window approach, XGB achieved the highest performance metrics, with an accuracy of 66.64%, along with sensitivity, specificity, precision, F1-score, and AUC values of

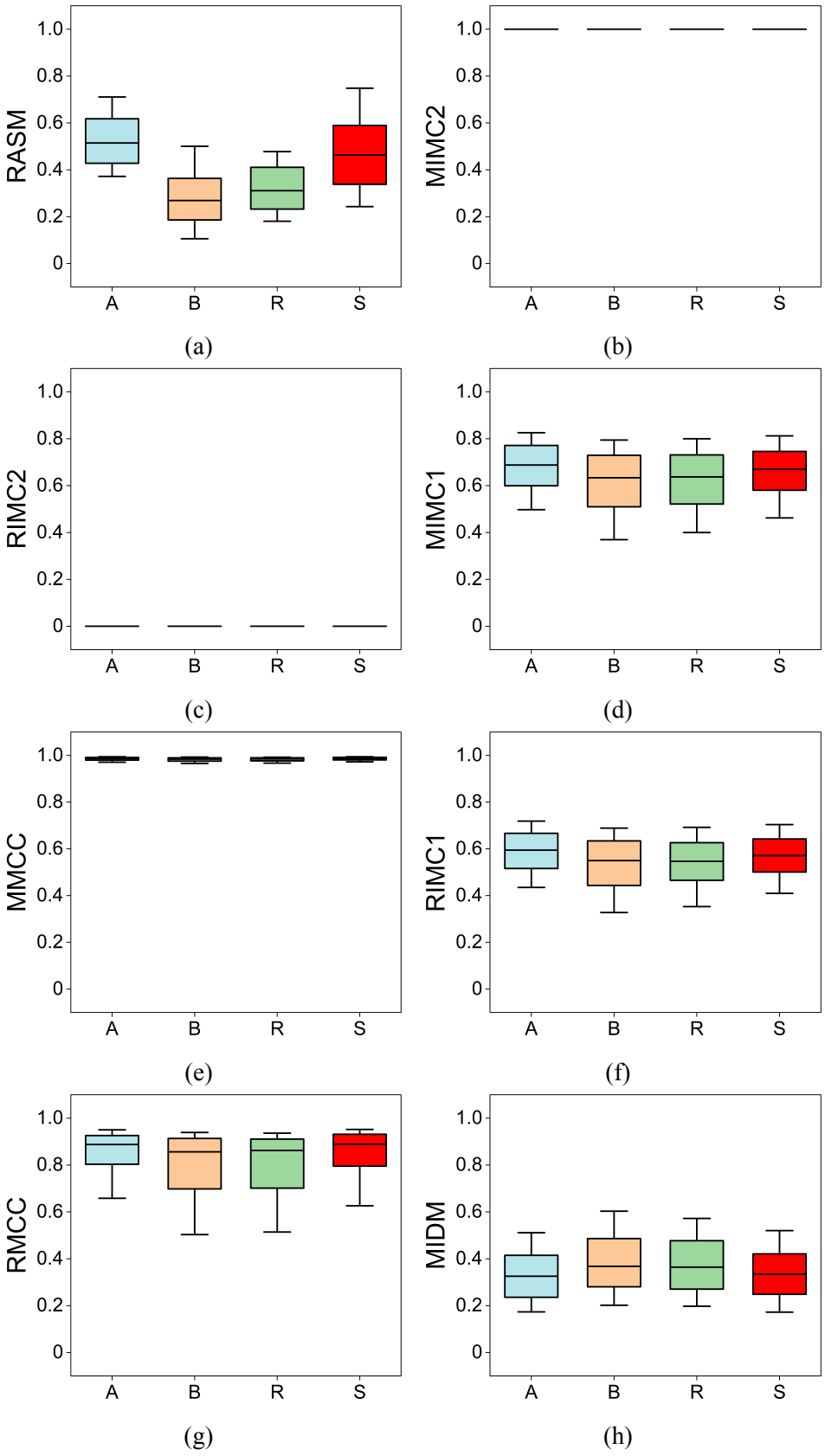

Figure 4.13: Representative box plots of top eight significant features extracted from GADF: (a) RASM, (b) MIMC2, (c) RMIC2, (d) MIMC1, (e) MMCC, (f) RIMC1, (g) RMCC, and (h) MIDM.

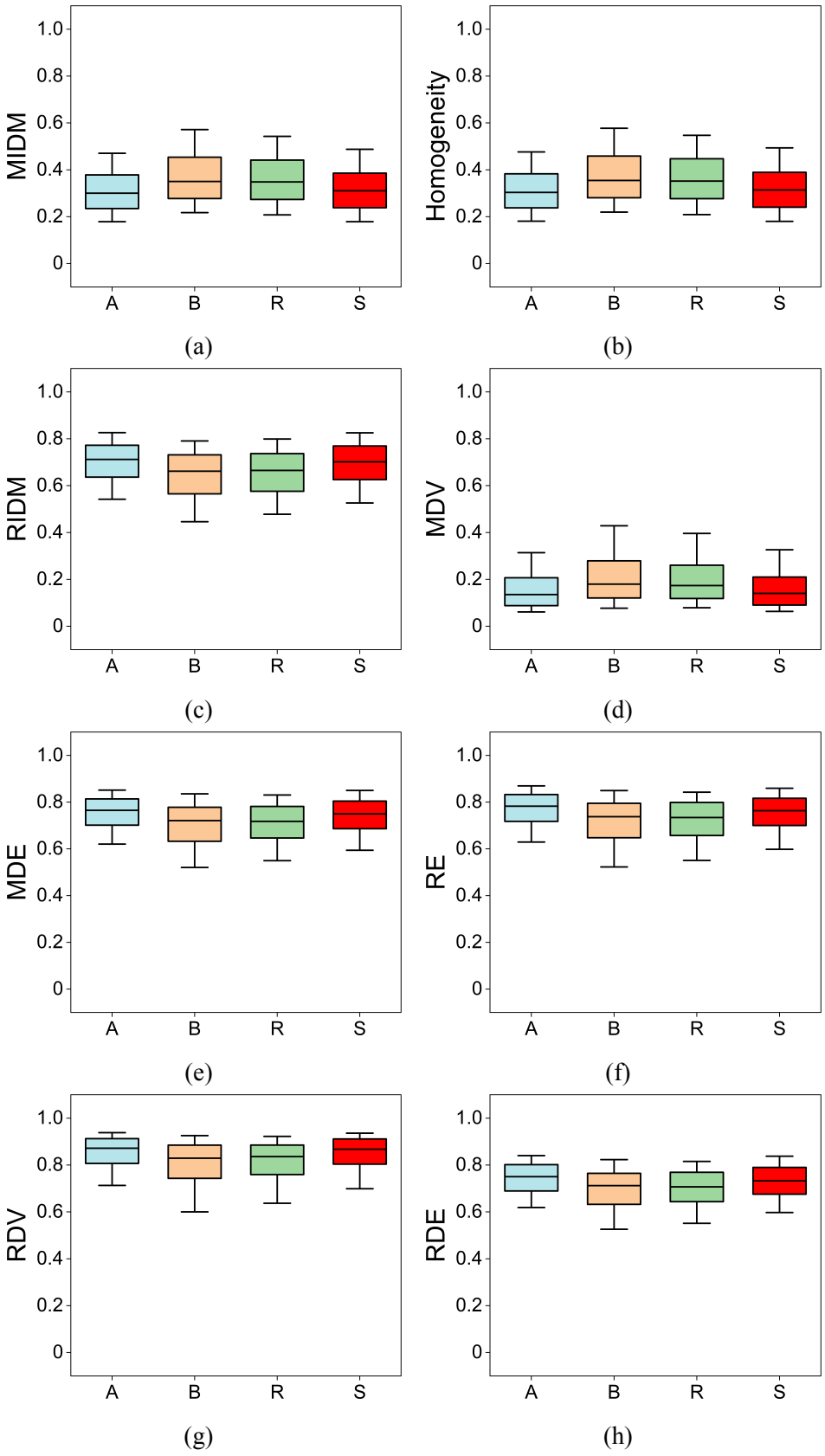

Figure 4.14: Representative box plots of top eight significant features extracted from GASF: (a) MIDM, (b) Homogeneity, (c) RIDM, (d) MDV, (e) MDE, (f) RE, (g) RDV, and (h) RDE.

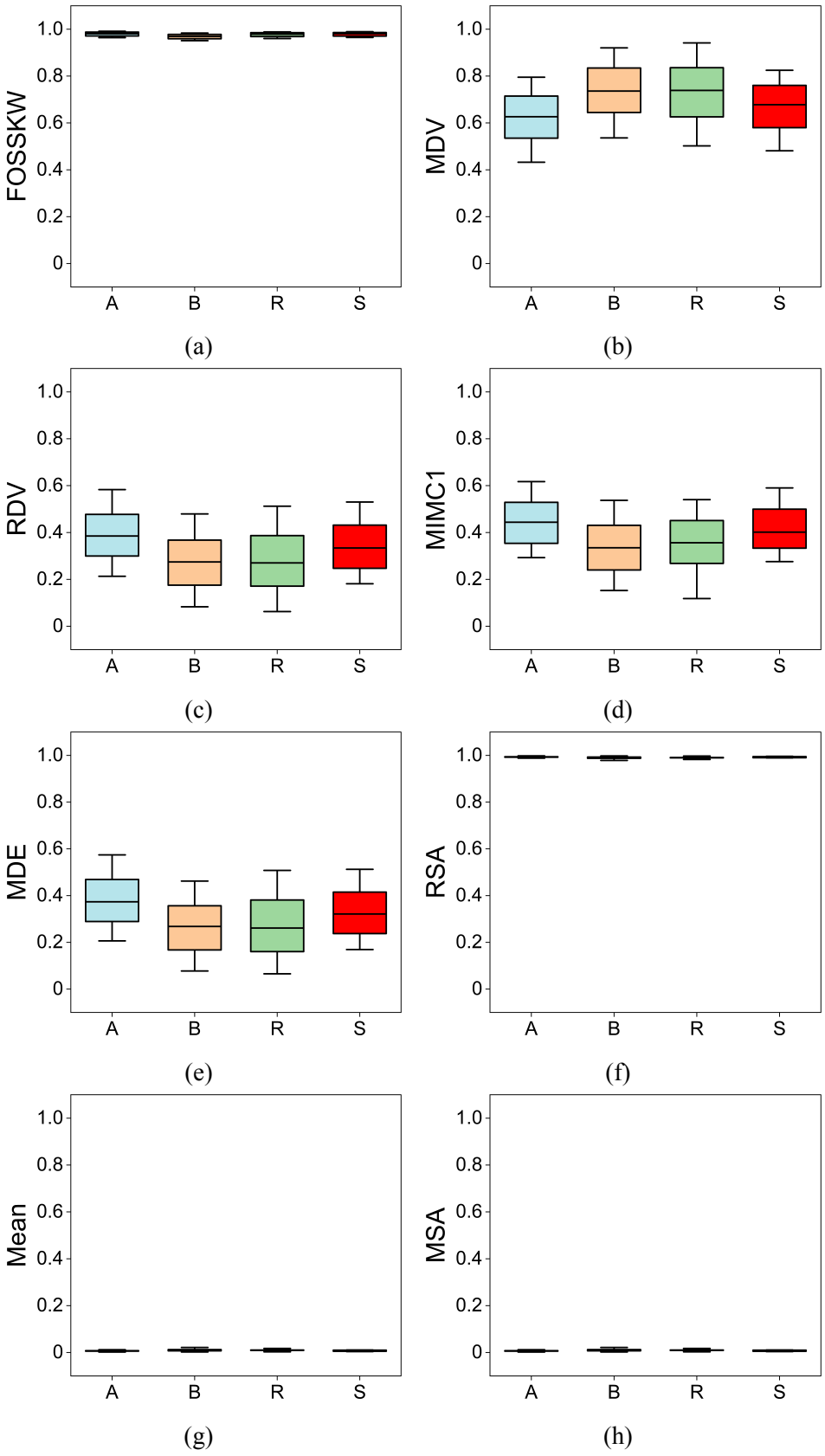

Figure 4.15: Representative box plots of top eight significant features extracted from MTF: (a) FOSSKW, (b) MDV, (c) RDV, (d) MIMC1, (e) MDE, (f) RSA, (g) Mean, and (h) MSA.

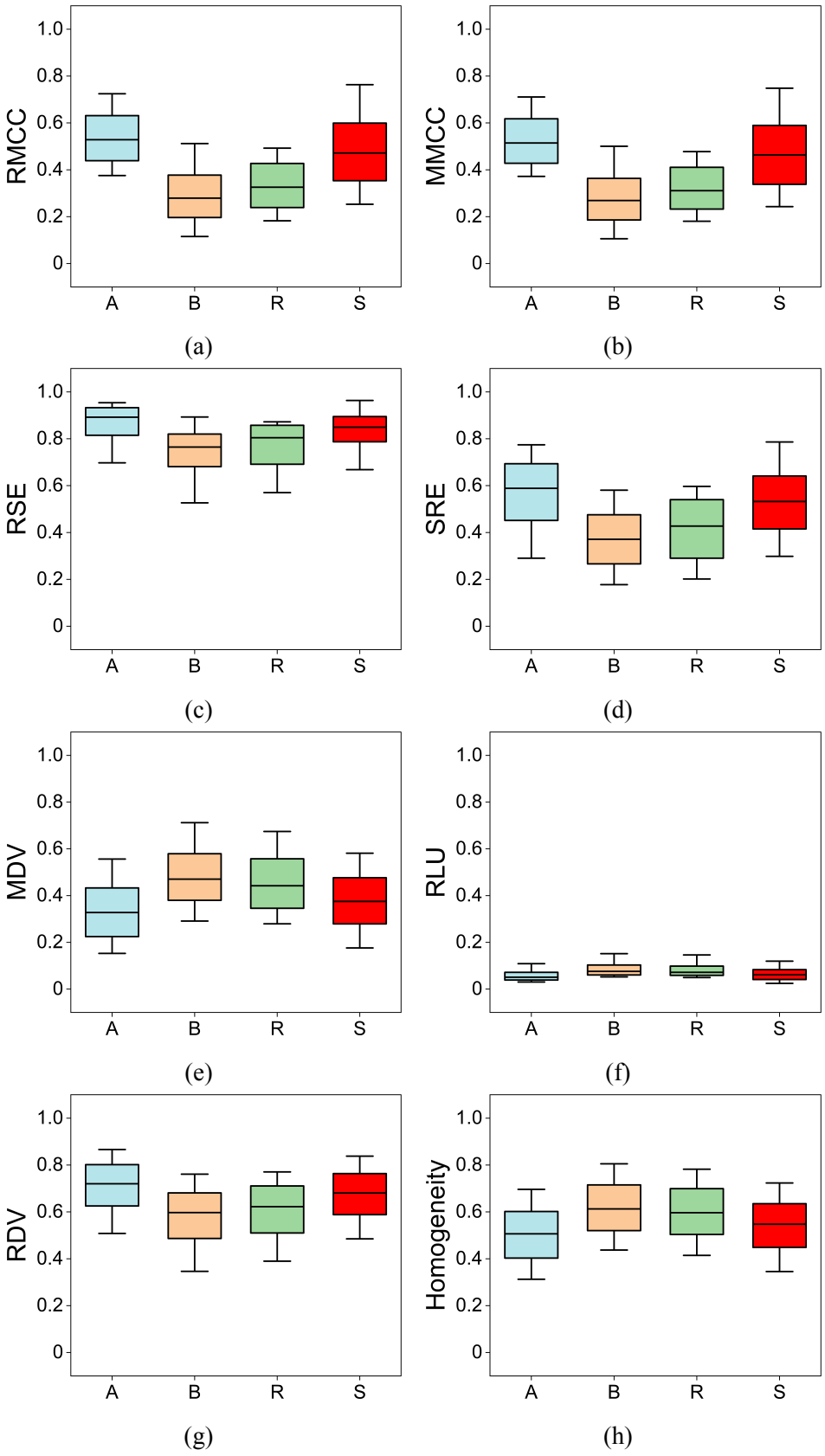

Figure 4.16: Representative box plots of top eight significant features extracted from RP: (a) RMCC, (b) MMCC, (c) RSE, (d) SRE, (e) MDV, (f) RLU, (g) RDV, and (h) Homogeneity.

36.85%, 76.89%, 35.48%, 35.04%, and 63.64%, respectively. Following closely were SVM and RF, with accuracies of 65.35% and 63.48%, respectively. Transitioning to the nine-window approach, XGB maintained its lead with an accuracy of 67.04% and demonstrated improvements across other metrics, including sensitivity, specificity, precision, F1-score, and AUC, reaching 37.83%, 78.03%, 37.48%, 37.48%, and 62.77%, respectively. Meanwhile, SVM and RF exhibited accuracies of 66.09% and 65.33%, respectively. Across both windowing approaches, XGB consistently outperformed RF and SVM, demonstrating its robustness in capturing the nuances of emotional states. While RF and SVM showed comparable performance metrics, XGB's superiority suggests its ability to leverage the dataset's features for classification tasks effectively. Notably, using a finer windowing approach improved accuracy and AUC values, highlighting the importance of detailed data representation.

Table 4.9: Classification results of SVM, RF, and XGB using GADF images for five-window and nine-window approaches.

	Five Windows			**Nine Windows**		
Metrics (%)	**SVM**	**RF**	**XGB**	**SVM**	**RF**	**XGB**
Accuracy	65.35	63.48	66.64	66.09	65.33	67.04
Sensitivity	36.5	33.43	36.85	38.98	35	37.83
Specificity	76.23	73.56	76.89	77.39	75.26	78.03
Precision	35.72	32.91	35.48	38.08	34.55	37.48
F1-score	35.71	33.04	35.04	38.17	34.61	37.48
AUC	66.06	62.38	63.64	67.29	61.95	62.77

Classification Results for GASF-based Features

Table 4.10 offers a comprehensive overview of the classification results derived from SVM, RF, and XGB models utilizing GASF images under both five-window and nine-window approaches. In the five-window approach, SVM stood out with a high accuracy of

67.04%, accompanied by sensitivity, specificity, precision, F1-score, and AUC values of 35.5%, 77.86%, 34.3%, 34.2%, and 62.27%, respectively. RF and XGB followed with accuracies of 64.79% and 66.79%, respectively. Transitioning to the nine-window approach, SVM continued to excel, showcasing the highest performance metrics with an accuracy of 70.53% and commendable sensitivity, specificity, precision, F1-score, and AUC values of 41.06%, 80.35%, 40.99%, 41%, and 68.43%, respectively. RF and XGB displayed slightly lower results than SVM, with accuracies of 67.78% and 68.58%, respectively. The results suggest that SVM excelled in the five-window approach, and its performance remained dominant in the nine-window approach, with superior metrics across the board. While RF and XGB also showed strong performance, they were slightly outperformed by SVM. Across both windowing scenarios, SVM demonstrated superior capability in handling complex data structures and effectively capturing underlying patterns. Adopting a finer windowing approach consistently improved classification outcomes, highlighting the importance of granularity in data representation.

Table 4.10: SVM, RF, and XGB classification results using GASF images for five-window and nine-window approaches.

	Five Windows			**Nine Windows**		
Metrics (%)	**SVM**	**RF**	**XGB**	**SVM**	**RF**	**XGB**
Accuracy	67.04	64.79	66.79	70.53	67.78	68.58
Sensitivity	35.5	34.07	37.75	41.06	35.56	37.18
Specificity	77.86	76.53	78.03	80.35	78.52	79.06
Precision	34.3	33.48	34.46	40.99	35.59	37.29
F1-score	34.2	33.61	33.83	41	35.49	37.05
AUC	62.27	61.59	61.02	68.43	62.84	62.90

Classification Results for MTF-based Features

Table 4.11 shows the classification results of SVM, RF, and XGB models using MTF images for both the five-segment and nine-segment approaches. In the five-segment approach, SVM achieved the highest performance metrics with an accuracy of 81.33%, sensitivity of 62.67%, specificity of 87.56%, precision of 62.82%, F1-score of 62.68%, and AUC of 83.94%. Conversely, XGB and RF yielded lower accuracies of 78.46% and 76.25%, respectively. Transitioning to a finer segmentation with nine segments, XGB demonstrated superior performance, achieving the highest accuracy of 88.24%, sensitivity of 76.48%, specificity of 92.16%, precision of 76.42%, F1-score of 76.37%, and AUC of 93.72%. In comparison, RF and SVM achieved accuracies of 87.27% and 85.67%, respectively. Adopting a finer segmentation consistently improved performance metrics such as accuracy and AUC, indicating enhanced model classification capabilities. Across both segmentation scenarios, XGB consistently outperformed SVM and RF, suggesting its superior ability to handle data complexity and nuances, resulting in enhanced predictive performance. SVM's performance slightly decreased when transitioning from five to nine segments, potentially indicating challenges in capturing nuances in more finely segmented data. RF performed comparatively lower among the models in both segmentation scenarios, suggesting its inferior effectiveness in capturing underlying data patterns compared to SVM and XGB.

Classification Results for RP-based Features

Table 4.12 illustrates the classification outcomes achieved by SVM, RF, and XGB models using RP images with both five and nine-window approaches. In the five-window scenario, XGB demonstrated the highest performance metrics, boasting accuracy, sensitivity, specificity, precision, F1-score, and AUC of 81.58%, 63.17%, 87.72%, 62.34%, 62.44%, and 85.45%, respectively. Conversely, SVM and RF achieved accuracies of 75.71% and 79.33%, respectively. Transitioning to nine windows approach, SVM emerged as the top performer, showcasing accuracy, sensitivity, specificity, precision, F1-score, and AUC

Table 4.11: SVM, RF, and XGB classification results using MTF images for five-window and nine-window approaches.

	Five Windows			**Nine Windows**		
Metrics (%)	**SVM**	**RF**	**XGB**	**SVM**	**RF**	**XGB**
Accuracy	81.33	76.25	78.46	85.67	87.27	88.24
Sensitivity	62.67	52.50	56.92	71.34	74.54	76.48
Specificity	87.56	84.17	85.64	90.45	91.51	92.16
Precision	62.82	52.61	57.17	70.99	74.38	76.42
F1-score	62.68	54.43	56.99	71.14	74.21	76.37
AUC	83.94	79.57	80.82	90.14	93.19	93.72

values of 84.75%, 69.49%, 89.83%, 70.1%, 68.45%, and 87.86%, respectively. Meanwhile, RF and XGB achieved accuracies of 82.41% and 84.14%, respectively. These findings suggest that a more detailed windowing approach enhances the models' classification capabilities, improving performance metrics across the board. While XGB demonstrates impressive performance in the five-window scenario, its performance slightly declines in the nine-window scenario compared to SVM. On the other hand, SVM exhibits robust performance in the nine-window scenario, surpassing both XGB and RF. This underscores SVM's adaptability to a finer window approach and its ability to capture underlying patterns accurately. RF consistently delivers moderate performance across both windowing approaches, indicating its competence in handling the data but lagging behind XGB and SVM in terms of predictive accuracy and robustness.

In summary, it is evident that the nine-window approach outperformed the five-window approach in terms of classification results. This suggests that employing a finer windowing approach leads to improved performance metrics across the board, indicating enhanced classification capabilities of the models. Notably, features such as GLRLM and GLCM emerged as significant influencers in classifying emotional states, while FOS and FDTA

Table 4.12: Classification results of SVM, RF, and XGB using RP images for five-window and nine-window approaches.

	Five Windows			**Nine Windows**		
Metrics (%)	**SVM**	**RF**	**XGB**	**SVM**	**RF**	**XGB**
Accuracy	75.71	79.33	81.58	84.75	82.41	84.14
Sensitivity	51.42	58.67	63.17	69.49	88.27	68.29
Specificity	83.81	86.22	87.72	89.83	64.39	89.43
Precision	50.83	58.13	62.34	70.1	63.91	68.12
F1-score	50.79	57.30	62.44	68.45	87.82	67.63
AUC	78.18	82.53	85.45	87.86	87.82	88.47

exhibited moderate influence. Conversely, HM features have a minor influence, and ZM features have a negligible impact on emotion classification. XGB consistently showcased superior performance, particularly excelling in finer windowing scenarios, highlighting its ability to handle intricate data structures and capture subtle emotional nuances. SVM also demonstrated robust performance across both the windowing approaches, showcasing its adaptability and reliability in discerning underlying data patterns. However, RF lags behind XGB and SVM in predictive accuracy and robustness, indicating potential limitations in capturing nuanced features. These findings underscore the critical role of method selection and windowing approach in achieving accurate and reliable classification outcomes in emotion recognition tasks.