Tuning In to Nerve Activity: Audio-Inspired Features for Sympathetic State Detection

Farnoush Baghestani, Yashvi Gupta, Imra Asif, Jihye Moon, Youngsun Kong, Ki H. Chon, *Fellow, IEEE*
Dept. of Biomedical Engineering, *University of Connecticut*, Storrs, USA

***Abstract*—We investigate the use of audio-inspired features to characterize ECG-derived skin sympathetic nerve activity (SKNA). Building on concepts from audio signal processing, we designed a set of SKNA-optimized spectral and energy-related features tailored to the transient, burst-like nature of sympathetic activation. High-frequency ECG signals were recorded from 20 participants during rest and during cognitive stress induced by the Stroop Color and Word Test. Applied to this two-phase protocol, the proposed features captured consistent differences between baseline and stress, achieving robust classification performance using support vector machines with leave-one-subject-out cross-validation. Results demonstrate that audio-inspired features, when physiologically constrained, provide a powerful and interpretable representation of SKNA dynamics, highlighting their promise for noninvasive assessment of autonomic function.**

***Index Terms*—SKNA, ECG, audio features, sympathetic activation, autonomic nervous system.**

I. Introduction

Imbalances in the autonomic nervous system (ANS)—particularly elevated sympathetic activity—are strongly associated with cardiovascular mortality, including sudden cardiac death and life-threatening arrhythmias [1].

ECG-derived skin sympathetic nerve activity (SKNA) is a noninvasive biosignal that reflects the activity of the sympathetic branch of the ANS [2], [3]. It is extracted from the high-frequency components of electrocardiogram (ECG) recordings using high-pass or band-pass filtering. SKNA captures transient bursts of electrical activity originating from the stellate ganglia and sympathetic pathways—an association supported by studies showing that stellate ganglion block, anesthetics, and beta-blockers such as propofol and methohexital markedly suppress SKNA [2], [3]. These signals offer a promising window into sympathetic modulation in both physiological and clinical settings, including cardiovascular regulation, stress response, and pain [4]–[7].

Despite its physiological significance, most SKNA studies still rely on relatively simple signal-level statistics such as the mean and standard deviation. Other approaches focus on burst characteristics, including burst frequency, burst duration, and total burst count, which depend on burst-detection algorithms tuned to amplitude thresholds or energy criteria [8], [9]. A notable example of an innovative feature is the SKNA Energy Ratio (SKNAER), a feature which quantifies the energy ratio of detected bursts [10].

Furthermore, spectral domain features, such as the low-frequency to high-frequency (LF/HF) power ratio derived from the power spectral density (PSD) of SKNA, have been used in the literature to characterize sympathovagal balance. These studies report a shift toward higher LF/HF values during sympathetic activation [11]–[13].

In our previous work, we proposed a novel feature to better capture the complexity of SKNA, termed TVSKNA (Time-Varying SKNA), that captures dynamic changes in SKNA by reconstructing signals from frequency bands most responsive to sympathetic stimulation [14].

Recent research demonstrates that audio-inspired features, originally developed for audio signals and speech, can effectively characterize non-speech physiological signals. For instance, Cepstral Coefficients (MFCCs) have been successfully applied in EEG and sleep stage classification [15].

Irfan et al. [16] adapted MFCCs to identify noisy channels in multi-channel surface EMG recordings. By modifying the filter banks and frequency range of MFCCs to match EMG signal characteristics, and combining them with statistical features, they trained a Support Vector Machine (SVM) classifier to distinguish between clean and contaminated channels.

Nguyen et al. [17] applied a range of speech-derived features—including MFCCs, pitch, energy, jitter, shimmer, and zero-crossing rate—to EEG signals for the purpose of biometric person identification. Based on the quasi-stationary assumption of EEG segments, they extracted these features using the openSMILE toolkit and trained SVM classifiers, achieving identification accuracies as high as 99%.

In this work, we propose applying a curated set of audio-inspired features specifically optimized for SKNA analysis. The feature set captures energy dynamics, frequency-specific intensity, spectral shape and structure, burst-related temporal variation, and SKNA-specific band ratios. Given that SKNA signals exhibit non-stationarity, transient bursts, and rapid spectral fluctuations, we hypothesize that these tailored features provide a physiologically grounded and interpretable representation of sympathetic activity.

II. Materials and Methods

A. Participants and Experimental Setup

The study included 20 healthy participants aged 18–34 years (10 females and 10 males). A well-established test was used to elicit cognitive stress and sympathetic nervous system activation: the Stroop Color and Word Test (SCWT) [18]. All participants provided written informed consent prior to the study, which was approved by the Institutional Review Board at the University of Connecticut.

Electrocardiogram (ECG) data were recorded using electrodes placed on the participants' wrists and ankles in lead III configuration, with a ground electrode positioned below the right rib cage. Signals were captured using a Bio Amp connected to a PowerLab system, and data acquisition was managed with LabChart Pro 8 software (ADInstruments, Sydney, Australia) at a sampling rate of 10 kHz. Subsequent data processing and analysis were performed using MATLAB (MathWorks, Natick, MA, USA) and Python.

The experimental protocol consisted of two phases:

Phase 1 – Baseline Recording: Participants remained at rest while ECG signals were recorded for 2 minutes.

Phase 2 – Cognitive Stress and Sympathetic Activation: Participants underwent the SCWT to induce cognitive stress to trigger a physiological stress response. ECG data were recorded during 3-minute SCWT sessions. Participants viewed a series of color-denoted words on a tablet, where the text color frequently conflicted with the color name (e.g., the word *"blue"* displayed in yellow ink). Participants were instructed to vocally identify the text color, ignoring the written word.

B. Preprocessing

For each segment, SKNA signals were extracted by high-pass filtering the ECG signals above 150 Hz. This cutoff has been shown to effectively eliminate ECG components (e.g., QRS complexes) and preserve sympathetic activity [14]. While many studies also apply a low-pass filter at 1000 Hz, given that we apply audio-inspired features designed for high-frequency signals and our sampling rate of 10 kHz, we intentionally preserved this broader frequency range in our analysis. Prior to any feature extraction, a notch filter centered at 762 Hz was applied to eliminate a persistent baseline noise observed in the recordings.

C. Audio-Inspired Feature Extraction

Signals were then segmented into short, non-overlapping windows of 0.1 s, consistent with microneurography and common SKNA practices for computing integrated SKNA (iSKNA) [3]. In each segment, we extracted audio-inspired features in 25ṁs frames with 10ṁs hop. These frame features were then aggregated over longer segments, producing averaged feature representations over time, analogous to common practice in both SKNA analysis [3] and audio signal processing [19]. Each segment served as an individual sample for subsequent analysis.

We designed a custom SKNA-optimized feature extractor, motivated by audio signal analysis but constrained to physiologically meaningful descriptors for sympathetic bursts. Speech-specific features (e.g., F0, jitter, shimmer, harmonicity) were excluded. Ten feature groups were defined:

a) Frame Energy Features (18 features): Energy was computed per frame as

$$E = \sum_{n=1}^{N} x[n]^2, \tag{1}$$

with log-energy defined as $\log(E+\epsilon)$. For both E and $\log E$, we extracted the mean, standard deviation, interquartile range (IQR), and percentiles (10th, 25th, 50th, 75th, 90th, 95th). These capture burst intensity and robustness to outliers.

b) Band-Limited RMS Features (46 features): RMS values were computed in physiologically aligned frequency bands: $(0-200)$, $(200-500)$, $(500-1000)$, $(1000-2000)$, $(2000-4000)$, and $(150-1000)$ Hz. For each band b,

$$\mathrm{RMS}_b = \sqrt{\frac{1}{K}\sum_{k\in b} |X[k]|^2}. \tag{2}$$

Features included RMS mean, standard deviation, 50th and 95th percentiles, and log-RMS statistics. Relative band contributions and inter-band ratios (e.g., vlow/low, low/mid, mid/high, high/vhigh) were also computed, quantifying spectral balance of SKNA.

c) Spectral Flux (5 features): Spectral flux quantified frame-to-frame spectral changes:

$$\mathrm{Flux}(t) = \sum_k \max(0, |X_t[k]| - |X_{t-1}[k]|). \tag{3}$$

Mean, standard deviation, median, 95th percentile, and IQR were extracted. Flux captures rapid transitions in bursts.

d) Spectral Shape Descriptors (4 features): Within 150–1000 Hz, we computed centroid, spread, skewness, and kurtosis. These describe the "brightness," concentration, asymmetry, and peakedness of SKNA bursts.

e) Spectral Structure Features (5 features): Within 150–1000 Hz, entropy was defined as

$$H = -\sum p_k \log p_k, \quad p_k = \frac{P[k]}{\sum P[k]}, \tag{4}$$

flatness as geometric/arithmetic mean ratio, and roll-off points as the frequency below which 50%, 75%, or 95% of power was contained. These features separate structured bursts from broadband noise.

f) Zero-Crossing Features (5 features): Zero-crossing rate (ZCR) was computed as

$$\mathrm{ZCR} = \frac{1}{N-1}\sum_{n=1}^{N-1} \mathbf{1}_{\{x[n]\cdot x[n-1]<0\}}, \tag{5}$$

with frame-level statistics (mean, std, median, 95th percentile). ZCR reflects high-frequency activity and burst rate.

g) SKNA Band Ratios (6 features): Energy was computed in custom SKNA bands $(150-350)$, $(350-700)$, and $(700-1000)$ Hz. Relative contributions and ratios (low/mid, mid/high, low/high) quantified energy distribution across physiologically relevant sub-bands.

h) Linear-Frequency Cepstral Coefficients (4 features): A linear filterbank spanning 150–1000 Hz (8 filters) was applied to compute log filterbank energies. A discrete cosine transform (DCT) yielded four low-order coefficients:

$$c_m = \sum_{n=1}^{N} \log(E_n)\cos\Big(\tfrac{\pi m}{N}(n-0.5)\Big), \quad m = 1,\ldots,4. \tag{6}$$

Unlike Mel-based MFCCs, these avoid psychoacoustic scaling and emphasize SKNA-specific envelopes.

i) Spectral Slope (1 feature): Log-power spectra between 200–900 Hz were regressed against frequency:

$$\log P(f) \approx \alpha f + \beta, \tag{7}$$

where α is the spectral slope, indexing overall spectral tilt.

j) Delta and Delta-Delta Features (8 features): First and second temporal derivatives were computed for centroid and log-energy:

$$\Delta f_t = f_t - f_{t-1}, \quad \Delta^2 f_t = \Delta f_t - \Delta f_{t-1}. \tag{8}$$

These capture temporal dynamics of sympathetic bursts.

D. Statistical Analysis and Classification

In the statistical analysis, significance was assessed using a mixed-effects model that accounted for inter-subject variability by including subject as a random effect.

To further evaluate the discriminative power of these features, we performed classification using a leave-one-subject-out (LOSO) cross-validation strategy. All features were standardized using z-score normalization. For dimensionality reduction, we applied feature selection based on Random Forest importance by selecting the top-10 most informative features. Classification was performed using a Support Vector Machine, with hyperparameters optionally optimized via grid search.

III. Results

The energy features displayed a consistent and characteristic pattern across phases. With the exception of variability measures, all features rose clearly in Phase 2, producing a distinct separation from Phase 1. This increase was observed across the mean and all percentile-based descriptors (p10–p95). Statistically, these features reached Fisher ratios in the range of 0.32–0.33 and AUROC values between 0.84 and 0.85, with all comparisons highly significant ($p < 0.001$). Variability-related measures behaved differently. The standard deviation decreased in Phase 2, reflecting reduced spread in the energy distribution, and this was accompanied by weaker performance (AUROC = 0.76). In contrast, the interquartile range showed no meaningful difference between phases. The log-transformed features highlighted these effects more strongly. Percentiles and the mean showed sharper increases in Phase 2, with Fisher ratios exceeding 1.20. Once again, only the variability descriptors deviated from this trend.

All absolute band-limited RMS means were significantly higher in Phase 2 compared to Phase 1, with the strongest increases observed in the mid- and high-frequency ranges (AUROC up to 0.89). However, when expressed as relative contributions, the spectral balance shifted: the low and very-low bands accounted for a larger fraction of the total power, whereas the mid-, high-, and very-high bands represented a smaller share despite their absolute increases. This redistribution resembles the interpretation of LF/HF ratios, where it is the relative balance between bands, rather than absolute power alone, that reflects underlying shifts in autonomic tone.

All spectral flux measures, including the mean, median, upper percentiles, standard deviation, and interquartile range, were significantly elevated in Phase 2 compared to Phase 1 ($p < 0.001$ for all). The increase was consistent across all metrics, with discriminative performance in the range of AUROC 0.83–0.86. Therefore, phase 2 signals exhibit more rapid and variable changes in their spectral composition, reflecting a higher degree of frequency dynamics.

Spectral shape descriptors showed an overall reduction in Phase 2. The centroid shifted slightly downward (AUROC $\approx$ 0.45), while spread, skewness, and kurtosis were all lower compared to Phase 1 (AUROCs $\approx$ 0.34–0.37, $p < 0.001$). These results suggest that the spectral distribution became narrower, more symmetric, and less heavy-tailed in Phase 2. However, the low discriminative performance indicates that these features provide only complementary information, rather than serving as strong stand-alone markers.

Spectral structure features showed small but consistent reductions in Phase 2. Both spectral entropy and spectral flatness were significantly lower (AUROCs $\approx$ 0.36–0.44, $p < 0.001$), suggesting a modest decrease in signal irregularity and noise-like character. Similarly, spectral rolloff values (50th, 75th, and 95th percentiles) were shifted downward in Phase 2, although with poor discriminative performance (AUROCs $\approx$ 0.40–0.51) despite statistical significance. Overall, these features confirm subtle changes in spectral distribution but do not serve as strong stand-alone markers.

Zero-crossing features, including the mean, median, upper percentile, and standard deviation, were consistently lower in Phase 2 compared to Phase 1 ($p < 0.001$ for all). Fisher's ratios were relatively high (0.55–1.01), indicating large relative group differences. These results suggest that while zero-crossing activity was systematically reduced in Phase 2, these features alone are not reliable markers of group separation.

Band ratios exhibited heterogeneous patterns between phases. The mid/high ratio was markedly elevated in Phase 2 (Fisher's ratio $= 0.48$, AUROC ≈ 0.74, $p < 0.001$), while the low/high ratio also increased, albeit with weaker discriminative power (Fisher's ratio $= 0.03$, AUROC ≈ 0.67, $p < 0.001$). In contrast, the low/mid ratio declined (Fisher's ratio $= 0.05$, AUROC ≈ 0.46, $p < 0.001$). Absolute band ratios showed less consistent trends: the high-band ratio decreased substantially in Phase 2 (Fisher's ratio $= 0.30$, AUROC ≈ 0.31, $p < 0.001$), while the mid-band and low-band ratios remained largely unchanged (AUROCs $\approx$ 0.51–0.52, not significant). Together, these findings indicate that relative spectral power shifted toward the mid-band at the expense of higher frequencies. This pattern is analogous to the interpretation of LF/HF ratios in heart rate variability, where changes in relative band contributions are linked to sympathovagal balance rather than absolute levels alone.

Linear-frequency cepstral coefficients (MFCCs) derived from SKNA showed divergent behaviors across components. The first coefficient (MFCC_1) was strongly elevated in Phase 2 (Fisher's ratio $= 1.49$, AUROC ≈ 0.88, $p < 0.001$), providing one of the most discriminative features overall. MFCC_2 also increased significantly (Fisher's ratio $= 0.37$, AUROC ≈ 0.73, $p < 0.001$). In contrast, MFCC_3 was lower in Phase 2 (Fisher's ratio $= 0.46$, AUROC ≈ 0.26, $p < 0.001$), while MFCC_4 showed a more modest increase (Fisher's ratio $= 0.04$, AUROC ≈ 0.64, $p < 0.001$). These results highlight MFCC_1 as a particularly powerful discriminator, with subsequent coefficients contributing complementary but less robust information.

Spectral slope was significantly steeper in Phase 2 (Fisher's

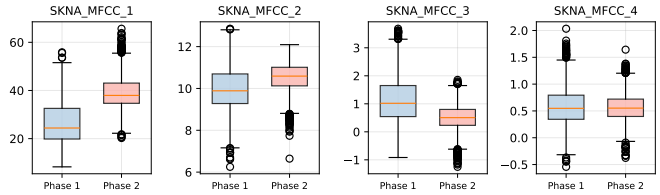

Fig. 1. MFCC features across Phase 1 (baseline recording) and Phase 2 (cognitive stress through SCWT), showing shifts in spectral energy distribution.

ratio $= 0.31$, $p < 0.001$), indicating a stronger decay of spectral energy with frequency. However, its discriminative performance was poor (AUROC ≈ 0.28), suggesting that while the slope systematically differs between groups, it does not serve as a reliable stand-alone marker of phase separation.

Delta and delta–delta features showed weak and inconsistent differences between phases. Mean values of the delta centroid and delta log-energy did not differ significantly ($p > 0.05$), while their standard deviations were modestly lower in Phase 2 (Fisher's ratios up to 0.45, $p < 0.001$). Delta–delta features contributed little additional information, with most showing no systematic separation (AUROCs ≈ 0.50).

Using top 10 selected audio-inspired features, the SVM classifier achieved an average 82% accuracy (Table I).

TABLE I
CLASSIFICATION REPORT – BASELINE VS. COGNITIVE STRESS (SCWT)

Class	Precision	Recall	F1-score	Support
Phase 1	0.86	0.69	0.77	2432
Phase 2	0.81	0.92	0.86	3379
Macro avg	0.83	0.80	0.81	5811
Weighted avg	0.83	0.82	0.82	5811

LOSO Cross-Validation Results: Overall accuracy (pooled predictions): 0.8231. Fold-wise accuracy — Mean: 0.8243 $\pm$ 0.1759 (std), Median: 0.9259.

IV. DISCUSSION AND FUTURE WORK

This study highlights the potential of audio-inspired features in capturing the fast-changing, burst-driven nature of SKNA. Unlike approaches which assumes stationarity and requires long segments for reliable power spectral density analysis, SKNA is inherently non-stationary and resembles audio signals. This limits the effectiveness of traditional LF/HF-based assessments. In contrast, features like MFCCs are well-suited for short, transient events and can be computed over brief segments without assuming stationarity. This enables fine-grained tracking of sympathetic dynamics, making them a more appropriate alternative for assessing sympathetic activity.

It is important to note that complete absence of stress in the baseline phase cannot be guaranteed. This makes the observed classification performance meaningful. To validate and generalize these findings, future work should incorporate larger datasets and clinical populations.

REFERENCES

[1] T. F. o. t. E. S. o. C. t. N. A. S. o. P. Electrophysiology, "Heart rate variability," vol. 93, no. 5, pp. 1043–1065. Publisher: American Heart Association.

[2] A. Doytchinova, J. L. Hassel, Y. Yuan, H. Lin, D. Yin, D. Adams, S. Straka, K. Wright, K. Smith, D. Wagner, C. Shen, V. Salanova, C. Meshberger, L. S. Chen, J. C. Kincaid, A. C. Coffey, G. Wu, Y. Li, R. J. Kovacs, T. H. Everett, R. Victor, Y.-M. Cha, S.-F. Lin, and P.-S. Chen, "Simultaneous noninvasive recording of skin sympathetic nerve activity and electrocardiogram," vol. 14, no. 1, pp. 25–33.

[3] T. Kusayama, J. Wong, X. Liu, W. He, A. Doytchinova, E. A. Robinson, D. E. Adams, L. S. Chen, S.-F. Lin, K. Davoren, *et al.*, "Simultaneous noninvasive recording of electrocardiogram and skin sympathetic nerve activity (neuecg)," *Nature protocols*, vol. 15, no. 5, pp. 1853–1877, 2020.

[4] R. A. Kabir, A. Doytchinova, X. Liu, D. Adams, S. Straka, L. S. Chen, C. Shen, S.-F. Lin, T. H. Everett, and P.-S. Chen, "Crescendo skin sympathetic nerve activity and ventricular arrhythmia," *Journal of the American College of Cardiology*, vol. 70, no. 25, pp. 3201–3202, 2017.

[5] T. Kusayama, J. Wan, A. Doytchinova, J. Wong, R. A. Kabir, G. Mitscher, S. Straka, C. Shen, T. H. Everett IV, and P.-S. Chen, "Skin sympathetic nerve activity and the temporal clustering of cardiac arrhythmias," *JCI insight*, vol. 4, no. 4, p. e125853, 2019.

[6] J. Mao, X. Liu, A. Kote, K. T. Andersson, X. Li, C. M. Albert, and P.-S. Chen, "Skin sympathetic nerve activity in symptomatic and asymptomatic paroxysmal atrial fibrillation," *Heart Rhythm*, vol. 21, no. 12, pp. 2437–2444, 2024.

[7] F. Baghestani, Y. Kong, W. D'Angelo, and K. H. Chon, "Analysis of sympathetic responses to cognitive stress and pain through skin sympathetic nerve activity and electrodermal activity," vol. 170, p. 108070.

[8] C. Liu, C.-H. Lee, S.-F. Lin, and W.-C. Tsai, "Temporal clustering of skin sympathetic nerve activity bursts in acute myocardial infarction patients," *Frontiers in Neuroscience*, vol. 15, p. 720827, 2021.

[9] S. Chen, G. Meng, A. Doytchinova, J. Wong, S. Straka, J. Lacy, X. Li, P.-S. Chen, and T. H. Everett IV, "Skin sympathetic nerve activity and the short-term qt interval variability in patients with electrical storm," *Frontiers in Physiology*, vol. 12, p. 742844, 2021.

[10] Y. Xing, Y. Zhang, Z. Xiao, C. Yang, J. Li, C. Cui, J. Wang, H. Chen, J. Li, and C. Liu, "An artifact-resistant feature sknaer for quantifying the burst of skin sympathetic nerve activity signal," *Biosensors*, vol. 12, no. 5, p. 355, 2022.

[11] G. Meng, W. He, J. Wong, X. Li, G. A. Mitscher, S. Straka, D. Adams, T. H. Everett, S. Manchanda, X. Liu, P.-S. Chen, and Y. Tang, "Successful continuous positive airway pressure treatment reduces skin sympathetic nerve activity in patients with obstructive sleep apnea," vol. 19, no. 1, pp. 127–136.

[12] A. K. Kirby, S. Pancholi, Z. Anderson, C. Chesler, T. H. Everett, and B. S. Duerstock, "Time and frequency domain analysis of physiological features during autonomic dysreflexia after spinal cord injury," vol. 17, p. 1210815.

[13] Y. Nakagawa, T. Kusayama, S. Tamai, Y. Nagamori, K. Takeuchi, S. Iwaisako, T. Tsutsui, T. Kamide, K. Misaki, S. Usui, K. Sakata, M. Nakada, and M. Takamura, "Association between skin sympathetic nerve activity and electrocardiogram alterations after subarachnoid hemorrhage," vol. 13, no. 1, p. e70202.

[14] Y. Kong, F. Baghestani, W. D'Angelo, I. Chen, K. H. Chon, *et al.*, "A novel approach to characterize dynamics of ecg-derived skin nerve activity via time-varying spectral analysis," *arXiv preprint arXiv:2411.08308*, 2024.

[15] W. Pei, Y. Li, P. Wen, F. Yang, and X. Ji, "An automatic method using mfcc features for sleep stage classification," *Brain informatics*, vol. 11, no. 1, p. 6, 2024.

[16] M. Irfan, K. Ullah, F. Muhammad, S. Khan, F. Althobiani, M. Usman, M. Alshareef, S. Alghaffari, and S. Rahman, "Automatic detection of outliers in multi-channel emg signals using mfcc and svm," *INTELLIGENT AUTOMATION AND SOFT COMPUTING*, vol. 36, no. 1, pp. 169–181, 2023.

[17] P. Nguyen, D. Tran, X. Huang, and D. Sharma, "A proposed feature extraction method for eeg-based person identification," in *Proceedings on the International Conference on Artificial Intelligence (ICAI)*, p. 1, The Steering Committee of The World Congress in Computer Science, Computer ..., 2012.

[18] F. Scarpina and S. Tagini, "The stroop color and word test," *Frontiers in psychology*, vol. 8, p. 557, 2017.

[19] R. Jahangir, Y. W. Teh, H. F. Nweke, G. Mujtaba, M. A. Al-Garadi, and I. Ali, "Speaker identification through artificial intelligence techniques: A comprehensive review and research challenges," *Expert Systems with Applications*, vol. 171, p. 114591, 2021.