

Perbandingan Metode Klasifikasi untuk Menentukan Tingkat Kenyamanan Suhu pada Kondisi Rileks Berbasis Sinyal EEG

Laurentius Kuncoro Probo Saputra¹, Ignatia Dhian Estu Karisma Ratri²

Fakultas Teknologi Informasi, Universitas Kristen Duta Wacana, Yogyakarta, Indonesia
kuncoro@staff.ukdw.ac.id
ignatiadhian@ti.ukdw.ac.id

Diterima 13 Desember 2018

Disetujui 31 Desember 2018

Abstract— *Temperature control of air conditioner devices is still oriented to the target environment. This control mode ignores one's physiological condition. A person's thermal comfort varies when indoors. Thermal comfort is closely related to environmental thermal satisfaction conditions. EEG signal is a signal that can reflect brain activity. This research objective is to provide classifier model for classifying person's thermal comfort based on EEG signal. This research used three conditions of a room's temperature. The features used by classifier are an average frequency band, HFD, PFD, and MSE features. Classifier performance was assessed using ROC curve evaluation. The results of the classification of thermal comfort levels with EEG signals with the KNN classifier are obtained only by using the band frequency average feature, which is equal to 0.878 with a standard deviation of 0.022. While the SVM classifier gets the highest performance by using a combination of the average band + HFD frequency feature, which is 0.877 with a standard deviation of 0.013 in the linear kernel and RBF.*

Index Terms— *EEG signal, thermal comfort, KNN, SVM, temperature.*

I. PENDAHULUAN

Semakin maraknya kegiatan penghematan energi dilakukan berbagai pihak menyebabkan banyaknya penelitian yang dilakukan untuk mendukung upaya tersebut. Terutama dengan adanya himbauan dari pemerintah Indonesia yang tertuang dalam Peraturan Menteri Energi dan Sumber Daya Mineral (ESDM) Nomor 13 tahun 2012 dengan jelas menyatakan bahwa seluruh bangunan gedung kantor pemerintah baik di pusat maupun daerah harus melaksanakan program penghematan energi listrik pada sistem Tata Udara (*Air Conditioning System*), sistem Tata Cahaya dan peralatan pendukung lainnya. Beberapa penelitian telah dilakukan untuk mengembangkan sistem kontrol suhu ruangan. Salah satunya berdasarkan sensor thermostat di ruangan [1], [2]. Adapun penelitian tersebut lebih banyak menggunakan sensor dalam melakukan kontrol suhu ruangan, antara lain sensor untuk mengukur tingkat kelembapan, suhu dan deteksi keberadaan manusia di suatu ruangan tertentu.

Kemudian terdapat penelitian lainnya yang berkaitan tentang pengembangan sistem kontrol suhu ruangan, yaitu dengan melihat faktor fisiologi dari seseorang. Adapun faktor fisiologis yang menjadi parameter dalam mengontrol suhu ruangan antara lain adalah melihat suhu yang dihasilkan dari permukaan kulit dan rekaman detak jantung seseorang. Parameter tersebut diharapkan dapat memberikan informasi yang lebih baik untuk mengetahui kenyamanan seseorang dengan suhu yang ada di sekitar mereka [3], [4].

Kenyamanan seseorang terhadap suhu di sekitar mereka dipengaruhi oleh beberapa variabel, antara lain suhu, rata-rata suhu radiasi pencahayaan, kecepatan udara, kelembapan, pakaian serta kegiatan yang dilakukan [3]. Seseorang yang tidak nyaman dengan suhu di sekitarnya biasanya berkaitan dengan tingkat stres yang dialami. Hal ini tentu saja akan memberikan dampak pada kinerja yang dilakukan serta kesehatan seseorang [3].

Tingkat stres maupun rileks dapat dilihat dari aktivitas otak seseorang. Sinyal *electrocephalogram* (EEG) dapat menampilkan aktivitas otak seseorang. Maka dapat dilakukan penelitian untuk melihat kenyamanan termal seseorang dengan melihat aktivitas sinyal EEG yang dihasilkan. Adapun sinyal EEG dari otak ditangkap oleh alat *headset MindwaveNeurosky*. Penelitian dalam mengetahui aktivitas orang dengan kondisi stres maupun rileks telah banyak dilakukan contohnya seperti yang dilakukan oleh Hilmi, dkk [5]. Akan tetapi belum ada yang menghubungkan dengan tingkat kenyamanan termal seseorang.

Proses klasifikasi menggunakan metode pembelajaran terbimbing (*supervised learning*) yang sudah banyak diterapkan dalam beberapa penelitian lainnya. Mehmood [6] melakukan perbandingan antara metode SVM dan KNN dalam mengklasifikasi emosi seseorang dengan 4 bentuk stimulus. Hasil yang didapat bahwa sinyal EEG dapat digunakan untuk mengklasifikasikan tingkat emosi seseorang. Metode

SVM dan KNN juga digunakan oleh Paul [7] untuk melakukan klasifikasi pada sinyal EMG.

Penelitian ini berdasarkan pada penelitian yang dilakukan oleh Choi [3] bahwa tingkat stres kenyamanan termal seseorang dapat dilihat dari *bio-feedback* yang merupakan respon tubuh atas kondisi yang tercapai. Penelitian ini akan menggunakan sumber *bio-feedback* yang berasal dari respon otak. Data yang digunakan ialah sinyal EEG. Sehingga penelitian ini melakukan perbandingan beberapa metode klasifikasi untuk melihat tingkat kenyamanan termal seseorang pada saat rileks menggunakan sinyal EEG. Adapun klasifikasi yang akan diteliti adalah KNN dan SVM. Data yang digunakan adalah data beberapa mahasiswa UKDW dengan kondisi kesehatan yang telah dipilih terlebih dahulu. Melalui penelitian ini diharapkan dapat memberikan perbedaan tingkat akurasi dari hasil perbandingan metode klasifikasi yang diterapkan.

II. LANDASAN TEORI

A. K Nearest Neighbors

Algoritma K- Nearest Neighbors (KNN) adalah sebuah metode yang dapat digunakan untuk melakukan klasifikasi terhadap objek dengan memperhitungkan data lain yang memiliki jarak paling dekat dengan objek tersebut [5]. Klasifikasi yang dilakukan oleh KNN ini didasarkan atas data latih yang ada dilihat dari jarak yang paling dekat dengan objek berdasarkan nilai k. Metode ini membantu untuk melakukan klasifikasi objek baru berdasarkan atribut dan data latih yang telah disimpan sebelumnya.

Penentuan data latih dan data uji perlu dilakukan pada saat menerapkan K-Nearest Neighbor (KNN). Kemudian dilakukan perhitungan untuk menghitung jarak dapat menggunakan metode Euclidean Distance atau Manhattan Distance [8]. Euclidean Distance dapat dihitung seperti pada persamaan (1):

$$d(x_i x_j) = \sqrt{\sum_r^n (a_r(x_i) - a_r(x_j))^2} \tag{1}$$

Keterangan :

$d(x_i x_j)$ = jarak Euclidean

x_i = record ke i

x_j = record ke j

a_r = data ke r

dengan $i, j = 1, 2, 3, \dots, n$

Sedangkan Manhattan Distance [9] dapat dihitung seperti pada persamaan (2):

$$\begin{aligned} dist\ Manhattan(x_i x_j) &= \left(\sum_{i=1}^m |x_i - x_j|^r \right)^{1/r} \tag{2} \end{aligned}$$

Keterangan:

x_i = data ke i

x_j = data ke j

B. Support Vector Machine

Support Vector Machine dikembangkan oleh Oser, Guyon, dan Vapnik. Pertama kali metode ini dipresentasikan pada tahun 1992. SVM pada dasarnya merupakan linear *classifier* yang kemudian dikembangkan agar dapat bekerja pada problem non-linear dengan memasukkan konsep kernel trick pada ruang kerja berdimensi tinggi [9], [10].

Secara sederhana SVM adalah metode untuk mencari *hyperplane* terbaik yang dapat digunakan untuk memisahkan dua kelas pada *input space*. *Hyperplane* pemisah terbaik didapatkan dari melakukan perhitungan margin *hyperplane* dan mencari titik maksimal yang ada. Margin adalah jarak antara *hyperplane* dengan data terdekat dari masing-masing kelas. Sedangkan data yang nantinya memiliki letak paling dekat dengan *hyperplane* tersebut nantinya yang disebut *support vector*.

Konsep kernel yang digunakan dalam SVM diterapkan untuk dapat menyesuaikan dengan pola persebaran kelas data. Jenis kernel yang dapat digunakan untuk membentuk *hyperplane* pemisah ialah kernel linear, radial basis function, dan polynomial.

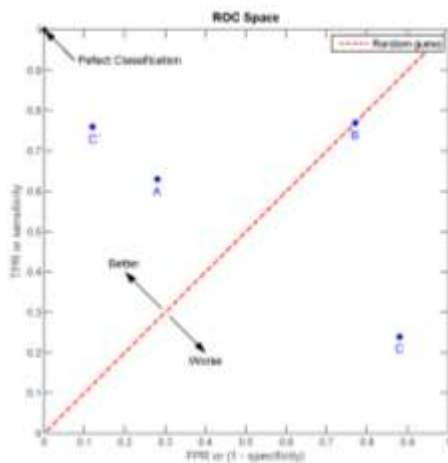
C. Receiver Operating Characteristic (ROC) Curve

Evaluasi sistem *classifier* dapat dilakukan dengan menggunakan *ROC curve*. *ROC curve* dibuat dengan cara melakukan *plotting* antara nilai *True Positive Rate* (TPR) dan *False Positive Rate* (FPR). TPR sering disebut dengan *sensitivity*, *recall*, atau probabilitas *classifier* memprediksi benar dari kelas positif. FPR sering disebut dengan *fall-out*, atau probabilitas *classifier* memprediksi benar dari kelas negatif. TPR dan FPR didapat berdasarkan hasil pembentukan *confusion matrix* pengujian, seperti yang diperlihatkan pada Gambar 1.

		Kondisi Sebenarnya	
		Positif (nyaman)	Negatif (tidak nyaman)
Prediksi	Positif (nyaman)	True Positif (TPR)	False Positif (FPR)
	Negatif (tidak nyaman)	False Negatif	True Negatif

Gambar 1. Confusion Matrix

ROC curve ditunjukkan seperti Gambar 2. Sebuah classifier dikatakan sempurna jika nilai TPR = 1 dan FPR = 0.



Gambar 2. ROC curve

III. METODOLOGI

Pengambilan data dalam penelitian ini menggunakan perangkat Neurosky Mindwave Headset. Neurosky merupakan alat penangkap sinyal EEG 1 kanal tanpa menggunakan jel untuk memperoleh data sinyal EEG, seperti pada Gambar 3.



Gambar 3. Alur penelitian

A. Data

Responden yang digunakan dalam penelitian ini sebanyak 8 responden. Responden terdiri dari 3 wanita dan 5 pria yang memiliki kondisi tubuh yang prima dan memiliki usia antara 18 – 22 tahun. Adapun untuk responden yang ikut dalam penelitian ini tidak memiliki permasalahan dalam kesehatan terutama berkaitan dengan alergi suhu dingin ataupun suhu panas, bukan perokok, memiliki tekanan darah normal dan tidak pernah melakukan operasi selama satu tahun terakhir. Hal ini diperlukan untuk menyamakan kondisi responden.

Sinyal EEG yang didapatkan dari neurosky mindwave headset disimpan dengan bantuan program OpenVibe. Program OpenVibe akan merekam data yang diterima dari Neurosky Mindwave headset selama pengambilan data. Data yang terekam disimpan dalam bentuk file dengan format *.csv.

Pada setiap pengambilan data, responden akan diperiksa kondisi tubuhnya, baik suhu tubuh dan kesehatan. Pemeriksaan ini dilakukan untuk memastikan bahwa responden dalam keadaan sehat sehingga proses pengambilan data bisa lebih valid.

Kemudian setelah itu, responden akan diberikan waktu 10 menit untuk beristirahat setelah menyelesaikan kuesioner tersebut. Langkah berikutnya adalah responden ditempatkan pada sebuah kondisi suhu ruangan tertentu dan hanya diminta untuk duduk dan memejamkan mata dalam kondisi rileks. Kemudian hal ini dilakukan untuk setiap responden dengan 3 tipe suhu yang berbeda. Adapun untuk pengambilan data pada 3 suhu yang berbeda dilakukan pada hari yang berbeda agar kondisi tubuh responden tidak mengalami perubahan suhu yang drastis dalam satu hari. Lagipula untuk penelitian ini setiap kondisi suhu, masing-masing responden diambil data EEG sebanyak lima (5) kali. Untuk kondisi suhu I adalah suhu dingin yang memiliki suhu 16 – 21° C, kemudian kondisi II adalah ruangan dengan suhu normal yaitu 21 - 26° C dan yang terakhir kondisi III adalah ruangan dengan suhu panas (lebih dari 27° C).

B. Pemilihan Fitur

Fitur yang digunakan secara umum dapat dikelompokkan menjadi 3 kategori fitur, yaitu fitur frekuensi, fitur fraktal, dan fitur statistik. Fitur Frekuensi yang digunakan ialah fitur rata-rata setiap band frekuensi sinyal EEG. Sinyal EEG memiliki 4 band frekuensi yaitu Alpha, Beta, Gamma, Delta. Masing-masing band memiliki rentang frekuensi yang berbeda.

Fitur Fraktal yang digunakan ialah Higuchi Fractal Dimension (HFD) dan Petrosian Fractal Dimension (PFD) [11].

PFD dapat dituliskan seperti Persamaan (3).

$$PFD = \frac{\log_{10} N}{\log_{10} N + \log_{10}(N/(N + 0.4N_{\delta}))} \quad (3)$$

dengan N merupakan jumlah data, dan N_{δ} ialah jumlah perubahan tanda pada sinyal. PFD merupakan fitur skalar.

Algoritma yang digunakan pada HFD ialah menyusun kembali ke dalam k kelompok deret dari data sampling sinyal $[x_1, x_2, x_3, \dots, x_N]$ dengan Persamaan (4):

$$x_m, x_{m+k}, x_{m+2k}, \dots, x_{m+|(N-m)/k|k} \quad (4)$$

dengan $m = 1, 2, \dots, k$.

Dari setiap kelompok deret yang dibentuk dari Persamaan (4), nilai panjang $L(m, k)$ dihitung dengan Persamaan (5).

$$L(m, k) = \frac{\sum_{i=2}^{|(N-m)/k|} |x_{m+ik} - x_{m+(i-1)k}| (N-1)}{|(N-m)/k|k} \quad (5)$$

Panjang rata-rata $L(k)$, dihitung dengan Persamaan (6):

$$L(k) = \left[\sum_{i=1}^k L(i, k) \right] / k \quad (6)$$

Langkah perhitungan dari Persamaan (4)-(6) diulang hingga k_{max} kali untuk setiap k dari 1 sampai k_{max} . Lalu metode *least-square* digunakan untuk menentukan slope dari garis yang terbaik dari kurva $L(k)$ vs $\ln(1/k)$. Nilai *slope* merupakan nilai HFD yang didapatkan dalam bentuk fitur skalar.

Fitur statistik yang digunakan ialah Multiscale Entropy Analysis (MSE). MSE sangat berguna dalam melakukan investigasi kompleksitas sebuah signal EEG yang saling berkorelasi di berbagai skala waktu.

C. Pengujian Classifier

Classifier yang sudah dilatihkan selanjutnya diuji dengan data uji. Data uji tidak pernah diikuti pada proses pelatihan. Data yang digunakan terbatas, sehingga pada proses pelatihan dan pengujian, data dibagi menjadi 75% sebagai data latih, dan 25% sebagai data uji. Pembagian antara data latih dan data uji dilakukan secara random. Komposisi kelas data yang digunakan pada penelitian ini tidak seimbang antara data dengan kelas nyaman dan kelas tidak nyaman. Oleh karena itu dalam pengujian *classifier* menggunakan metode *k-Fold Validation*. Pada pengujian ini menggunakan nilai $k=10$.

IV. HASIL

Tabel 1. Hasil Klasifikasi SVM dan KNN dengan menggunakan 7 Fitur

Classifier	Parameter	7 Fitur	
		Mean ROC	Standar Deviasi
SVM	1. Polynomial (n=3)	0,869	0,011
	2. Polynomial (n=5)	0,868	0,011
	3. Polynomial (n=7)	0,867	0,011
	4. RBF	0,868	0,010
	5. Linear	0,870	0,011
kNN	6. k=3, Manhattan	0,823	0,015
	7. k=3, Euclidian	0,815	0,018
	8. k=5, Manhattan	0,843	0,014
	9. k=5, Euclidian	0,841	0,013
	10. k=7, Manhattan	0,849	0,013
	11. k=7, Euclidian	0,846	0,014
	12. k=9, Manhattan	0,856	0,014
	13. k=9, Euclidian	0,851	0,014
	14. k=11, Manhattan	0,857	0,015
	15. k=11, Euclidian	0,853	0,016

Pengujian dengan menggunakan data set sinyal EEG yang diambil dalam kondisi relaks dengan mengubah suhu ruangan dalam 3 kondisi suhu telah dilakukan. Adapun hasil yang didapatkan dari hasil pengujian dengan menggunakan klasifikasi KNN dan SVM pada 7 Fitur yang telah didapatkan dapat dilihat pada Tabel 1 berikut ini.

Tabel 1 memperlihatkan hasil klasifikasi dengan metode SVC dengan fitur sebanyak 7. Klasifikasi SVC mendapatkan performa tertinggi pada klasifikasi dengan fungsi linear. Performa terlihat pada nilai mean

ROC sebesar 0,870 dengan standar deviasi sebesar 0,011. Hasil klasifikasi dengan menggunakan kNN mendapatkan performa tertinggi 0,857 dengan standar deviasi 0,15 pada nilai $k=11$ dan perhitungan jarak Manhattan.

Tabel 2. Hasil Klasifikasi SVM dan KNN pada Fitur Frekuensi Band dan Fitur Fraktal

Classifier	Parameter	Fitur Frekuensi Band		Fitur Fraktal	
		Mean ROC	sdv	Mean ROC	sdv
SVM	Polynomial (n=3)	0,575	0,208	0,824	0,020
	Polynomial (n=5)	0,650	0,228	0,824	0,020
	Polynomial (n=7)	0,396	0,229	0,824	0,020
	RBF	0,846	0,019	0,826	0,021
	Linear	0,853	0,019	0,825	0,020
kNN	k=3,Manhattan	0,866	0,019	0,725	0,024
	k=5,Manhattan	0,875	0,017	0,751	0,020
	k=7,Manhattan	0,876	0,021	0,761	0,021
	k=9,Manhattan	0,878	0,022	0,768	0,021
	k=11,Manhattan	0,877	0,022	0,772	0,022

Tabel 3. Hasil Klasifikasi SVM dan KNN pada Fitur Frekuensi Band+PFD dan Fitur Frekuensi Band+HFD

Classifier	Parameter	Fitur Frek Band + PFD		Fitur Frek Band + HFD	
		Mean ROC	sdv	Mean ROC	sdv
SVM	Polynomial (n=3)	0,848	0,014	0,875	0,013
	Polynomial (n=5)	0,781	0,217	0,873	0,014
	Polynomial (n=7)	0,151	0,017	0,493	0,369
	RBF	0,851	0,017	0,877	0,013
	Linear	0,856	0,018	0,877	0,013
	kNN	k=3,Manhattan	0,834	0,009	0,849
k=5,Manhattan		0,852	0,011	0,862	0,018
k=7,Manhattan		0,856	0,011	0,864	0,018
k=9,Manhattan		0,854	0,013	0,862	0,018
k=11,Manhattan		0,852	0,014	0,861	0,017

Perhitungan jarak Manhattan dapat menghasilkan performa klasifikasi yang lebih baik dibandingkan dengan perhitungan jarak Euclidian pada klasifikasi kNN. Sehingga pada hasil Tabel 2 dan Tabel 3 hanya akan menampilkan hasil kNN dengan perhitungan jarak Manhattan. Tabel 2 dan 3 memperlihatkan hasil pengujian *classifier* dengan menguji fitur mana yang paling berpengaruh dalam mendukung performa yang baik dari *classifier* tersebut. Dari Tabel 2 dan 3,

pengujian dilakukan dalam 4 kelompok fitur. Kelompok uji fitur pertama hanya 4 nilai fitur rata-rata frekuensi band. Kelompok uji fitur kedua hanya 2 nilai fitur fraktal. Kelompok uji fitur ketiga gabungan antara fitur rata-rata frekuensi band dengan PFD. Kelompok uji keempat gabungan antara fitur rata-rata frekuensi band dengan HFD.

Berdasarkan pengujian tersebut, dapat dilihat performa *classifier* tertinggi berdasarkan ROC curve didapatkan pada *classifier* KNN (manhattan distance) dengan menggunakan 4 nilai fitur rata-rata frekuensi band. Hasil klasifikasi sebesar 0,878 dengan standar deviasi sebesar 0,022. Sedangkan untuk *classifier* SVM, performa *classifier* tertinggi didapatkan 0,878 dengan standar deviasi 0,013 pada kelompok uji fitur gabungan antara fitur rata-rata frekuensi band+HFD. Kernel menggunakan kernel linear dan RBF.

V. KESIMPULAN

Hasil klasifikasi tingkat kenyamanan termal dengan sinyal EEG dengan *classifier* KNN didapatkan hanya dengan menggunakan fitur rata-rata frekuensi band, yaitu sebesar 0,878 dengan standar deviasi 0,022. Sedangkan *classifier* SVM mendapatkan performa tertinggi dengan menggunakan gabungan fitur rata-rata frekuensi band+HFD, yaitu 0,877 dengan standar deviasi 0,013 pada kernel linear dan RBF.

UCAPAN TERIMA KASIH

Penelitian ini didanai oleh Direktorat Riset dan Pengabdian Masyarakat Direktorat Jendral Penguatan Riset dan Pengembangan, Kementerian Riset, Teknologi dan Pendidikan Tinggi (Penelitian Dosen Pemula) No. 109/SP2H/LT/DRPM/2018. Kami juga mengucapkan terima kasih kepada Lembaga Penelitian dan Pengabdian Masyarakat (LPPM) dan Fakultas Teknologi Informasi (FTI) Universitas Kristen Duta Wacana.

DAFTAR PUSTAKA

- [1] Z. Cheng, W. W. Shein, Y. Tan, and A. O. Lim, "Energy efficient thermal comfort control for cyber-physical home system," *2013 IEEE Int. Conf. Smart Grid Commun. SmartGridComm 2013*, pp. 797–802, 2013.
- [2] S. M. Zanoli and D. Barchiesi, "Thermal and lighting control system with energy saving and users comfort features," *2012 20th Mediterr. Conf. Control Autom. MED 2012 - Conf. Proc.*, pp. 1322–1327, 2012.
- [3] J. H. Choi, "CoBi: Bio-Sensing Building Mechanical System Controls for Sustainably Enhancing Individual Thermal Comfort," *PHD Diss.*, no. May, 2010.
- [4] K. Nakayama, T. Suzuki, and K. Kameyama, "Estimation of thermal sensation using human peripheral skin temperature," *Syst. Man Cybern.*, no. October, pp. 2872–2877, 2009.
- [5] A. Hilmi, I. Wijayanto, and S. Hadiyoso, "Analisis Perbandingan Pola Sinyal Alfa Dan Beta Eeg Untuk Klasifikasi Kondisi Rileks Pada Perokok Aktif Dengan Menggunakan K-Nearest Neighbor Pattern Comparison Analysis Between Alpha and Beta Eeg Signal for Relaxed Condition Classification on Active Smok," vol. 4, no. 3, pp. 3395–3402, 2017.
- [6] R. M. Mehmood and H. J. Lee, "Emotion classification of EEG brain signal using SVM and KNN," in *2015 IEEE International Conference on Multimedia & Expo Workshops (ICMEW)*, 2015, pp. 1–5.
- [7] Y. Paul, V. Goyal, and R. A. Jaswal, "Comparative analysis between SVM & KNN classifier for EMG signal classification on elementary time domain features," in *2017 4th International Conference on Signal Processing, Computing and Control (ISPCC)*, 2017, pp. 169–175.
- [8] L.-Y. Hu, M.-W. Huang, S.-W. Ke, and C.-F. Tsai, *The distance function effect on k-nearest neighbor classification for medical datasets*, vol. 5, 2016.
- [9] Y. Helmi Mahendra, H. Tjandrasa, C. Fatchah, Y. H. Mahendra, and H. Tjandrasa, *KLASIFIKASI DATA EEG UNTUK MENDETEKSI KEADAAN TIDUR DAN BANGUN MENGGUNAKAN AUTOREGRESSIVE MODEL DAN SUPPORT VECTOR MACHINE*, vol. 15, 2017.
- [10] K. A. I. Aboalayon, H. T. Ocbagabir, and M. Faezipour, "Efficient sleep stage classification based on EEG signals," in *IEEE Long Island Systems, Applications and Technology (LISAT) Conference 2014*, 2014, pp. 1–6.
- [11] F. S. Bao, X. Liu, and C. Zhang, "PyEEG: An Open Source Python Module for EEG/MEG Feature Extraction," *Comput. Intell. Neurosci.*, vol. 2011, 2011.