

# Peningkatan Akurasi Pengenalan Emosi pada Sinyal *Electroencephalography* Menggunakan *Multiclass Fisher Discriminant Analysis*

Evi Septiana Pane<sup>1,2</sup>, Adhi Dharma Wibawa<sup>3,4</sup>, Mauridhi Hery Purnomo<sup>3,4</sup>

**Abstract**—EEG signals have a significant correlation to emotions when compared to other external appearances such as face and voice. Due to the low accuracy of emotional recognition through EEG signals, this study proposes a dimensional reduction method for EEG data to address that problem using Multiclass Fisher Discriminant Analysis (MC-FDA). In this study, the experiment was applied on public EEG dataset with three classes of emotions, namely positive, negative, and neutral. Differential entropy features were extracted from the decomposed EEG signals in five frequency band of the delta, theta, alpha, beta, and gamma. The accuracy of emotion recognition was measured using two prevalent classifiers on EEG identification, such as LDA and SVM. To demonstrate the superiority of the MC-FDA method, the PCA dimension reduction method was applied as a comparison. Classification accuracy results from all experiment scenario showed the advantages of the MC-FDA compared to the PCA. The best emotion classification accuracy was obtained from trials on all data from twelve electrodes using the MC-FDA and LDA methods, namely 93.3%. These results show a mean increase in accuracy of 3.5 points from the original feature vector dataset.

**Intisari**—Sinyal EEG memiliki korelasi yang signifikan terhadap emosi bila dibandingkan dengan modalitas lainnya. Dikarenakan akurasi yang rendah pada pengenalan emosi melalui sinyal EEG, makalah ini mengusulkan metode reduksi dimensi pada sinyal EEG untuk menangani permasalahan tersebut menggunakan *Multiclass Fisher Discriminant Analysis* (MC-FDA). Dalam makalah ini, uji coba diterapkan pada dataset publik dengan tiga kelas emosi, yaitu positif, negatif, dan netral. Metode ekstraksi ciri *differential entropy* diterapkan pada sinyal EEG yang telah didekomposisi ke dalam frekuensi delta, theta, alpha, beta, dan gamma. Akurasi pengenalan emosi diukur menggunakan dua metode klasifikasi ternama pada identifikasi EEG, yakni LDA dan SVM. Untuk menunjukkan keunggulan metode MC-FDA, metode reduksi dimensi PCA diterapkan sebagai pembandingan. Hasil akurasi klasifikasi dari seluruh uji

coba menunjukkan keunggulan MC-FDA dibanding PCA. Akurasi klasifikasi emosi terbaik diperoleh dari uji coba pada seluruh data dari dua belas elektrode dengan metode MC-FDA dan LDA, yakni 93,3%. Hasil tersebut menunjukkan rerata peningkatan akurasi sebesar 3,5 poin dari data ciri vektor semula.

**Kata Kunci**- gelombang otak, reduksi dimensi, EEG, pengenalan emosi multikelas

## I. PENDAHULUAN

Kondisi emosi diketahui memiliki korelasi dengan tingkat kekebalan tubuh manusia. Penelitian sebelumnya menunjukkan bahwa perubahan emosi memengaruhi sistem kekebalan tubuh [1]. Sistem kekebalan berubah sesuai dengan respons terhadap perubahan emosional, seperti marah, cemas, depresi, dan emosi positif (misalnya gembira, senang, optimis). Sebuah studi melaporkan bahwa faktor emosi negatif (misalnya depresi, stres) berakibat fatal pada individu yang memiliki riwayat penyakit kronis [2]. Emosi memainkan peran penting dalam mengatur sistem dalam tubuh yang memengaruhi kesehatan. Oleh sebab itu, pengenalan emosi diperlukan untuk mencegah timbulnya dampak buruk dari emosi negatif yang berlebihan bagi kesehatan. Dikarenakan emosi muncul secara tiba-tiba dan tidak terkontrol, maka diperlukan pengenalan emosi yang cepat serta akurat.

Dalam bidang interaksi manusia-komputer, pengenalan emosi berkaitan dengan identifikasi emosi menggunakan berbagai modalitas. Sebagian besar studi pengenalan emosi menggunakan citra ekspresi wajah dan sinyal audio wicara manusia [3]. Akhir-akhir ini, banyak peneliti melakukan identifikasi emosi dari sinyal fisiologis, seperti *electrocardiography* (ECG), *galvanic skin responses* (GSR), tekanan pembuluh darah, dan *electroencephalography* (EEG). Di antara sinyal fisiologis lainnya, EEG memiliki korelasi yang signifikan terhadap emosi [4]. Hal ini terjadi karena sistem *limbic* yang terutama berpusat pada otak mengendalikan proses pengelolaan emosi [5]. Ini menjelaskan penyebab sinyal EEG pada otak dapat menjadi petunjuk dalam identifikasi pengenalan emosi.

Pada umumnya pengenalan emosi berbasis EEG mencakup empat tahapan utama yang meliputi pemrosesan awal sinyal EEG, ekstraksi ciri, pemilihan dan reduksi dimensi pada vektor ciri, serta klasifikasi emosi. Tahap pra-proses sinyal EEG bertujuan untuk menghapus artefak dan derau yang tidak diperlukan. Tahap ekstraksi ciri dilakukan untuk mengekstrak informasi dari data EEG yang berdimensi besar. Identifikasi ciri emosi dari sinyal EEG masih menjadi masalah yang belum terpecahkan [6]. Namun, penelitian terdahulu

<sup>1</sup>Mahasiswa, Departemen Teknik Elektro, Fakultas Teknologi Elektro, Institut Teknologi Sepuluh Nopember, Jl. Raya ITS, Sukolilo, Surabaya 60111, Indonesia (telp/fax.: 031-5947302; e-mail: evi.pane15@mhs.ee.its.ac.id, pyrena.eve@gmail.com)

<sup>2</sup>Widyaiswara, Balai Diklat Industri Surabaya, Kementerian Perindustrian, Jl. Gayung Kebonsari Dalam No. 12 Surabaya, 60254, Indonesia (telp/fax: 031-8292002)

<sup>3</sup>Dosen, Departemen Teknik Elektro, Fakultas Teknologi Elektro, Institut Teknologi Sepuluh Nopember, Jl. Raya ITS, Sukolilo, Surabaya 60111, Indonesia (telp/fax: 031-5947302; e-mail: adhiosa@te.its.ac.id, hery@ee.its.ac.id)

<sup>4</sup>Dosen, Departemen Teknik Komputer, Fakultas Teknologi Elektro, Institut Teknologi Sepuluh Nopember, Jl. Raya ITS, Sukolilo, Surabaya 60111, Indonesia (telp/fax: 031-5922936)

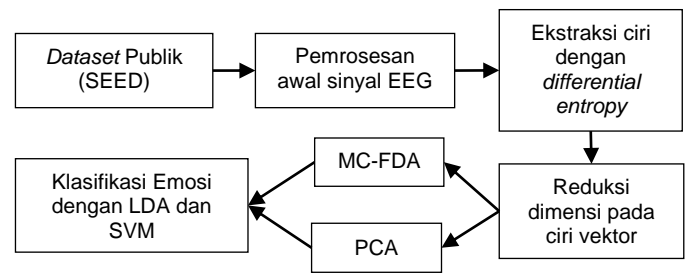
menemukan bahwa ciri *differential entropy* yang diekstrak dari domain frekuensi terbukti dapat diandalkan dalam pengenalan emosi pada sinyal EEG [7]. Oleh karena itu, pada makalah ini ciri *differential entropy* diterapkan dalam tahap ekstraksi ciri.

Hasil tahapan ekstraksi ciri adalah kumpulan vektor ciri dengan dimensi data yang besar. Hal ini berkaitan dengan jumlah elektrode EEG yang dilibatkan dalam identifikasi. Semakin banyak elektrode yang disertakan, semakin besar ukuran vektor ciri. Namun, jika elektrode yang dilibatkan dalam identifikasi sedikit akan berakibat pada kurangnya informasi EEG yang diperoleh. Untuk mengurangi ukuran dimensi vektor ciri, diusulkan metode untuk melakukan reduksi dimensi vektor ciri. Metode ini bertujuan untuk memproyeksikan data berdimensi tinggi ke dimensi yang lebih rendah, seperti 2D atau 3D. Di antara beberapa teknik dalam reduksi dimensi, metode *Principal Component Analysis* (PCA) digunakan pada sebagian besar kasus [8]–[10]. Terlepas dari popularitasnya, PCA memiliki keterbatasan dalam menangani *dataset* yang tidak bergantung pada nilai variansinya [11]. Selain itu, PCA merupakan teknik tanpa pengawasan (*unsupervised*) yang tidak memanfaatkan informasi label dari *dataset*. Untuk mengatasi kelemahan PCA, Fisher memperkenalkan *Fisher Discriminant Analysis* (FDA). Metode FDA mengeksploitasi informasi label dari *dataset* sementara melakukan proyeksi data. FDA mencari proyeksi yang berguna untuk tujuan diskriminasi [12]. Hal ini bertujuan untuk memperbesar jarak antar kelas dan mengurangi jarak di dalam kelas ketika membangun matriks bobot sebagai faktor diskriminan. Selain pengurangan dimensi, FDA dapat digunakan dalam masalah klasifikasi, yang dikenal sebagai *linear discriminant analysis* (LDA) [13]. LDA memprediksi kelas target dari sebuah contoh data sesuai dengan ciri diskriminan dari kelas dengan rata-rata vektor terdekat ke contoh data tersebut.

Sebuah penelitian pengenalan emosi berbasis EEG mengusulkan versi nonlinear FDA yang disebut dengan *kernel* FDA sebagai metode reduksi dimensi [14]. Hasil penelitian tersebut menunjukkan bahwa *kernel* FDA dapat memberikan akurasi yang baik pada klasifikasi untuk dua kelas dimensi emosi, yakni *arousal* dan *valence*. Akan tetapi, teknik *kernel* FDA hanya dapat digunakan untuk klasifikasi dua kelas emosi saja. Oleh karena itu, metode *multiclass* FDA (MC-FDA) yang diterapkan, diharapkan mampu bekerja untuk klasifikasi multikelas. MC-FDA mengembangkan metode FDA dengan melakukan pencarian sejumlah proyeksi tertentu dengan informasi jumlah kelas maksimum yang dapat dipisahkan berdasarkan nilai vektor *eigen* dan dekomposisi nilai *eigen*. Oleh sebab itu, tujuan yang hendak dicapai dari makalah ini adalah peningkatan akurasi pengenalan emosi dari sinyal EEG dengan menggunakan teknik reduksi dimensi MC-FDA dengan target kelas emosi lebih dari dua.

## II. DATA DAN METODE

Dalam penelitian ini, diusulkan metode reduksi dimensi untuk meningkatkan akurasi pengenalan emosi pada sinyal EEG. Metodologi keseluruhan yang diusulkan dalam studi ini meliputi: praproses sinyal EEG, ekstraksi ciri, reduksi dimensi, dan klasifikasi emosi, seperti pada Gbr. 1.



Gbr. 1 Metode pengenalan emosi pada sinyal EEG.

### A. Dataset Publik

*Dataset* publik EEG dari SEED terdiri atas 675 rekaman sinyal EEG [7], [15]. *Dataset* diambil dari lima belas peserta yang dilakukan dalam tiga sesi percobaan. Semua peserta adalah mahasiswa pada universitas di Cina yang dilaporkan sendiri dalam kondisi memiliki pendengaran dan penglihatan yang baik secara umum. Selang waktu antar sesi lebih atau kurang dari tujuh hari. Dalam setiap sesi, peserta direkam saat menonton lima belas klip film terstimulasi. Klip film mencakup emosi positif, negatif, dan netral dengan durasi 4 menit untuk setiap klip. Setelah menonton klip video yang distimulasi, masing-masing peserta diberi kuesioner untuk melaporkan tanggapan emosi masing-masing peserta setelah melihat setiap klip film. Sinyal EEG dalam *dataset* ini diperoleh menggunakan 62-elektrode (berdasar sistem internasional 10-20) ESI NeuroScan System. Pada sinyal awal EEG dilakukan proses *sub-sampling* dengan frekuensi 200 Hz.

### B. Praproses Sinyal EEG

Langkah-langkah praproses diperlukan untuk menghilangkan derau yang terkandung dalam sinyal EEG mentah. Ini termasuk penghapusan kedipan mata, pemilihan elektrode, dan dekomposisi kanal frekuensi. Dalam *dataset* SEED, gangguan kedipan mata dihilangkan menggunakan filter *bandpass* antara frekuensi 0,5 Hz hingga 70 Hz. Pemilihan elektrode pada studi ini menggunakan rekomendasi dua belas elektrode dari pekerjaan [7]. Pada dasarnya sinyal EEG adalah kombinasi dari lima kanal frekuensi utama, yaitu delta, gamma, beta, alfa, dan theta. Dalam studi ini, sinyal EEG diuraikan ke dalam lima frekuensi (yaitu, gamma 31-45 Hz, beta 14-30 Hz, alfa 8-13 Hz, theta 4-7 Hz, dan delta 1-3 Hz). Filter *bandpass* yang diterapkan menggunakan *Short-Time Fourier Transform* (STFT) dengan 256-titik dan jendela Hanning yang tidak tumpang tindih untuk memecah sinyal EEG kedalam lima kanal frekuensi tersebut.

### C. Ekstraksi Ciri dengan Differential Entropy

Berbagai ciri telah diusulkan sebagai ciri dari emosi pada sinyal EEG. Studi ulasan yang dilakukan tentang metode ekstraksi ciri pada pengenalan emosi melalui sinyal EEG membahas tiga domain ciri ekstraksi yang seringkali digunakan, yakni domain waktu, frekuensi, dan waktu-frekuensi [16]. Ciri yang paling sering digunakan pada domain frekuensi adalah ciri daya (*power*) yang dihitung menggunakan algoritme *Fast Fourier Transform* (FFT), algoritme STFT, dan pendekatan *Power Spectral Density* (PSD) untuk sinyal nonstasioner [17]. Sedangkan ekstraksi

pada domain ciri waktu-frekuensi umum menggunakan algoritme *Discrete Wavelet Transform* (DWT) yang seringkali diterapkan untuk keperluan analisis pada sinyal fisiologis [18].

Pada pekerjaan sebelumnya pada *dataset SEED*, ciri *differential entropy* digunakan sebagai ciri utama dalam pengenalan emosi. Ciri ini pertama kali diusulkan pada pengenalan emosi berbasis EEG [19]. Ciri *differential entropy* memperluas gagasan dari *Shannon entropy* yang digunakan untuk mengukur kompleksitas variabel acak yang kontinu. *Differential entropy* didefinisikan sebagai (1).

$$\begin{aligned} h(X) &= - \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}} \log\left(\frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}\right) dx \quad (1) \\ &= \frac{1}{2} \log(2\pi e\sigma^2). \end{aligned}$$

Rangkaian data  $X$  memenuhi distribusi Gauss  $N(\mu, \sigma^2)$ . Telah terbukti bahwa untuk segmen EEG yang tetap, nilai *differential entropy* setara dengan rerata spektrum energi pada kanal frekuensi tertentu. Selanjutnya untuk ekstraksi ciri, di hitung nilai *differential entropy* dari lima hasil dekomposisi kanal frekuensi yang telah dilakukan di tahap praproses sinyal EEG.

#### D. Principal Component Analysis (PCA)

PCA merupakan metode terkemuka dalam reduksi dimensi serta untuk masalah pemilihan ciri. PCA bekerja berdasarkan metode statistik yang menggunakan teknik proyeksi untuk mengurangi jumlah atribut dalam *dataset* [20]. Pada PCA, atribut yang sangat berkorelasi diakumulasi bersama untuk membentuk kelompok baru vektor ciri yang diperkecil. PCA memproyeksikan seluruh rangkaian data dengan mengabaikan informasi label kelas ke dalam subruang yang berbeda. PCA melakukan dekomposisi vektor tunggal pada matriks *covariance* untuk mendapatkan nilai *eigen* dan vektor *eigen*. Akibatnya, matriks linear  $w$  yang sesuai dengan matriks proyeksi dari vektor *eigen* terpilih digunakan untuk mengubah ciri asli  $f$  ke ciri yang dikurangi  $k$ , dengan jumlah  $k$  biasanya lebih sedikit dari  $f$ . Vektor ciri yang berkurang,  $y_i \in \mathcal{R}^f$ , ditentukan oleh (2).

$$y_i = w^T x_i, \quad i = 1, \dots, N \quad (2)$$

dengan  $x_i$  adalah turunan dari vektor ke- $i$  dalam *dataset* EEG, yang mewakili setiap rekaman EEG dan  $W$  adalah kolom yang berisi nilai *eigen*  $e_i$  yang diperoleh dengan menghitung nilai *eigen* dan vektor *eigen* dari matriks *covariance*. Komponen utama terbaik yang memisahkan *dataset* terbaik adalah berdasarkan  $k$  nilai *eigen* terbesar dalam urutan menurun. Vektor-vektor *eigen*  $k$  yang berhubungan dengan nilai-nilai *eigen* terbesar dipilih untuk membangun matriks bobot  $W$ . Selanjutnya, *dataset* awal  $x$  ditransformasikan ke dalam ciri-ciri dimensi  $k$  subruang  $y$  melalui  $W$ . Asumsi utama metode PCA adalah dimensi-dimensi yang menyatakan perbedaan paling signifikan dari titik data yang paling berguna [11]. Namun, asumsi tersebut tidak berlaku umum pada semua kasus.

#### E. Multiclass Fisher Discriminant Analysis (MC-FDA)

Cara kerja metode MC-FDA bermula dari metode FDA. Metode MC-FDA mencari vektor proyeksi terbaik yang memisahkan antara kelas dengan baik. Ide dasarnya adalah dengan mengeskalisasi fungsi yang merepresentasikan perbedaan antara rerata pada masing-masing kelas, dilakukan dengan mengukur sebaran data di antarkelas dan dalam kelas [12]. Untuk klasifikasi kasus multikelas, dilakukan pencarian untuk proyeksi sejumlah  $c$  kelas dikurangi satu,  $(c-1)$ , melalui proyeksi vektor matriks  $w_i$ . Oleh karena itu, proyeksi ke dimensi yang lebih rendah ditentukan oleh  $y = w_i^T x$  sebagai hasil dari satu kumpulan sampel data terdekat ke kelas yang sama berkaitan dengan matriks sebaran. MC-FDA ditentukan sebagai operasi linear dari  $w^T$  yang bertujuan memperbesar jarak antarfungsi kriteria berikut.

$$J(w) = \frac{w^T S_B w}{w^T S_W w} \quad (3)$$

Untuk mencari proyeksi optimal,  $w^T$ , ditetapkan nilai  $J(w)$  sebagai fungsi eksplisit dari matriks bobot  $w$ . Matriks sebaran digunakan untuk mengukur persebaran dalam vektor ciri  $x$ . Sebuah matriks sebaran dapat dihitung sebagai berikut.

$$S_B = \sum_{c=1}^c (\mu_c - \bar{x})(\mu_c - \bar{x})^T \quad (4)$$

$$S_W = \sum_c \sum_{i \in c} (x_i - \mu_c)(x_i - \mu_c)^T \quad (5)$$

dengan  $\bar{x}$  merupakan rerata global dari keseluruhan data, sementara  $\mu_c$  merupakan rerata untuk masing-masing kelas  $c$ ,  $S_B$  adalah matriks sebaran antarkelas, dan  $S_W$  adalah matriks sebaran dalam kelas. Pada metode MC-FDA yang digunakan, ditetapkan parameter tambahan,  $r$ , (parameter *ridge*) dengan nilai tertentu, yakni  $1e-6$ , untuk ditambahkan pada kumpulan matriks diagonal,  $S_W$ . Untuk menyelesaikan masalah optimalisasi nilai maksimum, nilai *eigen* untuk setiap kolom pada matriks bobot  $w$  digunakan sebagai pemecahan. Pada kolom dalam matriks  $w$  diwakili oleh nilai vektor *eigen* yang optimal berdasar nilai *eigen* terbesar. Matriks bobot  $w$  dihitung menggunakan kalkulasi berikut [21].

$$S_B w = \lambda S_W w \rightarrow w = S_W^{-1/2} U \quad (6)$$

dengan  $U = S_W^{-1/2} S_B S_W^{-1/2}$ , diperoleh dari vektor *eigen*. Kumpulan ciri yang telah direduksi didapatkan dari transformasi data awal ke dalam subruang baru melalui perkalian dengan matriks bobot  $w$ .

#### F. Metode Klasifikasi Emosi

Metode klasifikasi yang seringkali digunakan pada identifikasi dan pengenalan emosi melalui sinyal EEG adalah SVM [22]–[24] dan LDA [23]. Oleh karena itu, dalam makalah ini, diterapkan algoritme SVM dan LDA untuk tujuan klasifikasi kelas emosi ke dalam tiga kondisi, yakni positif, negatif, dan netral.

1) *Linear Discriminant Analysis (LDA)*: Metode LDA mengasumsikan bahwa data di setiap kelas memiliki distribusi normal yang berbeda, dengan matriks *covariance* yang sama untuk semua kelas. Analisis diskriminan menggunakan garis *hyperplane* untuk memisahkan data antarkelas. *Hyperplane*

merupakan jarak antara pusat kluster dengan data. Garis *hyperplane* tersebut digunakan untuk menentukan contoh data ke dalam target kelas yang tepat. Lebih dari satu *hyperplane* digunakan untuk memisahkan masalah klasifikasi dengan  $N$  kelas ( $N > 2$ ). LDA menggunakan jenis diskriminan linear yang berdasarkan pada perhitungan rerata sampel setiap kelas, kemudian mengurangi jumlah sampel kelas untuk menghitung matriks *covariance* dan menganggapnya sebagai hasilnya. Klasifikasi LDA mengasumsikan bahwa data distribusi untuk setiap kelas berbagi matriks *covariance* yang sama. LDA melakukan multivarians Gaussian pada setiap kelas dan menerapkan aturan Bayes untuk tujuan klasifikasi. Nilai Gaussian diperoleh dengan mengumpulkan matriks *covariance* untuk beberapa kelas yang berbeda (sehingga *classifier* yang dihasilkan akan linear). Untuk setiap kelas  $c$ , sampel data ditentukan berdasarkan distribusi multivarian Gaussian, dengan matriks *covariance*,  $\Sigma$ , dan rerata,  $\mu_c$ , tertentu. Asumsi matriks *covariance* untuk masing-masing kelas serupa sehingga distribusi dari masing-masing kelas memiliki bentuk linear. Kemudian, fungsi diskriminan linear untuk masing-masing kelas ditentukan sebagai (7) [21].

$$y_c = x^T \Sigma^{-1} \mu_c - \frac{1}{2} \mu_c^T \Sigma^{-1} \mu_c + \log\left(\frac{n_c}{n}\right) \quad (7)$$

dengan  $n$  merupakan jumlah dari keseluruhan data dan  $n_c$  merupakan jumlah data dari masing-masing kelas  $c$ . Data sampel diklasifikasi berdasarkan nilai fungsi linear  $y_c$  terbesar.

2) *Support Vector Machine (SVM)*: Dalam studi ini, SVM digunakan untuk melakukan klasifikasi kelas emosi pada sinyal EEG. Prinsip kerja utama dari SVM adalah dengan menentukan *hyperplane* yang optimal untuk memisahkan data uji coba [25]. *Hyperplane* yang dikatakan optimal adalah *hyperplane* dengan jarak  $d$  terbesar antara dua sampel data pelatihan. *Hyperplane* optimal terletak tepat di tengah antara dua kelas. Sampel pelatihan terdekat yang berada paling dekat dengan *hyperplane* disebut *support vector*. *Hyperplane* yang optimal diperoleh dengan memaksimalkan margin ( $d$ ) melalui *Lagrange multiplier*, yaitu,  $L = ((w, \bar{b}), \alpha)$ . Vektor  $w$  yang menentukan *hyperplane* optimal dapat diberikan sebagai kombinasi linear dari vektor data pelatihan.

$$w = \sum_{i=1}^l y_i \alpha_i x_i, \alpha_i \geq 0. \quad (8)$$

Secara garis besar, proses pelatihan pada metode SVM bertujuan untuk mencari *support vector* untuk menentukan *hyperplane* yang optimal. Untuk klasifikasi SVM pada makalah ini, algoritme *sequential minimum optimization* digunakan untuk membangun model klasifikasi dengan *kernel* polinomial [26].

### III. HASIL EKSPERIMEN

Pada makalah ini, evaluasi terhadap MC-FDA diterapkan pada *dataset* publik SEED. Mengingat bahwa *dataset* SEED terdiri atas lima belas subjek yang dilakukan perekaman EEG sebanyak tiga sesi dengan rentang waktu satu hingga dua minggu, maka strategi evaluasi pada studi ini terbagi atas tiga skenario uji coba, yang meliputi: (a) uji coba pada setiap subjek, (b) uji coba pada setiap sesi, dan (c) uji coba pada keseluruhan data.

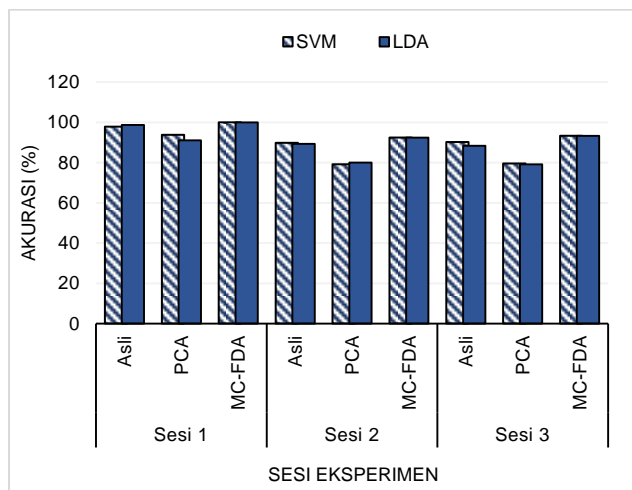
TABEL I  
HASIL AKURASI (%) PENGENALAN EMOSI PADA TIAP SUBJEK

Subjek	SVM			LDA		
	Asli	PCA	MC-FDA	Asli	PCA	MC-FDA
S01	68,9	77,8	100	57,8	84,4	100
S02	73,3	68,9	100	64,4	75,6	100
S03	75,6	75,6	100	62,2	77,8	100
S04	75,6	71,1	100	60,0	82,2	100
S05	73,3	68,8	100	57,8	73,3	100
S06	82,2	80,0	100	71,1	75,6	100
S07	93,3	93,3	100	95,6	95,6	100
S08	93,3	93,3	100	86,7	93,3	100
S09	82,2	93,3	100	80,0	82,2	100
S10	71,1	80,0	100	66,7	73,3	100
S11	82,2	86,6	100	73,3	84,4	100
S12	75,6	80,0	100	71,1	82,2	100
S13	86,7	86,7	100	71,1	86,7	100
S14	86,7	91,1	100	80,0	91,1	100
S15	91,1	95,5	100	93,3	95,5	100
<b>Rerata</b>	<b>80,7</b>	<b>82,8</b>	<b>100</b>	<b>72,7</b>	<b>83,5</b>	<b>100</b>

Pada setiap skenario, dihitung akurasi pada ciri asli sebelum dilakukan reduksi dimensi, ciri yang direduksi dengan metode PCA, dan ciri yang direduksi dengan metode MC-FDA. Untuk tujuan klasifikasi emosi, digunakan metode SVM dan LDA. Ukuran akurasi dihitung berdasarkan jumlah data yang berhasil diprediksi dengan benar dibandingkan dengan jumlah keseluruhan data. Metode uji coba yang diterapkan menggunakan validasi silang dengan *5-fold* untuk membagi data ke dalam data pelatihan dan data uji coba.

#### A. Hasil Uji Coba pada Skenario Setiap Subjek

Jumlah data pada setiap subjek terdiri atas 45 sampel data, yang terbagi sama ke dalam tiga kelas emosi (lima belas data untuk setiap kelas emosi). Pada skenario ini, dua belas elektrode dilibatkan dalam ekstraksi ciri, sehingga dimensi dari vektor ciri adalah 5 kanal frekuensi  $\times$  12 elektrode. Maka dimensi vektor ciri per subjek adalah  $45 \times 60$ . Tabel I menunjukkan hasil akurasi klasifikasi emosi di setiap subjek. Hasil pengukuran akurasi antar subjek memiliki hasil yang beragam pada klasifikasi data vektor ciri awal sebelum direduksi. Pengukuran akurasi klasifikasi dengan metode SVM dan LDA pada semua subjek tidak ditemui perbedaan yang signifikan. Pada klasifikasi data setelah direduksi dengan metode PCA dan klasifikasi SVM, beberapa subjek mengalami penurunan akurasi dibandingkan akurasi pada data asli, yaitu pada S02, S04, dan S06. Sementara pada klasifikasi LDA dengan metode PCA, seluruh subjek mengalami peningkatan akurasi. Sementara, hasil akurasi data yang direduksi dengan MC-FDA, baik menggunakan klasifikasi SVM maupun LDA, nilai akurasi meningkat hingga mencapai 100% pada seluruh subjek. Ditinjau dari hasil rerata akurasi di semua subjek dengan SVM, terlihat peningkatan akurasi dari data asli ke PCA sebesar 2,08 poin, sementara kenaikan dari data asli ke MC-FDA sebesar 19,3 poin. Pada rerata akurasi hasil klasifikasi LDA, kenaikan akurasi dari data asli ke PCA sebesar 10,8 poin, sementara dari data asli ke MC-FDA meningkat sebesar 27,3 poin.



Gbr. 2 Hasil akurasi pengenalan emosi untuk uji coba per- sesi pengambilan data.

### B. Hasil Uji Coba pada Skenario per Sesi

Perekaman data EEG pada subjek dilakukan selama tiga sesi yang berbeda. Pada setiap sesi, subjek diberi stimulasi lima belas klip film yang serupa dengan urutan yang berbeda.

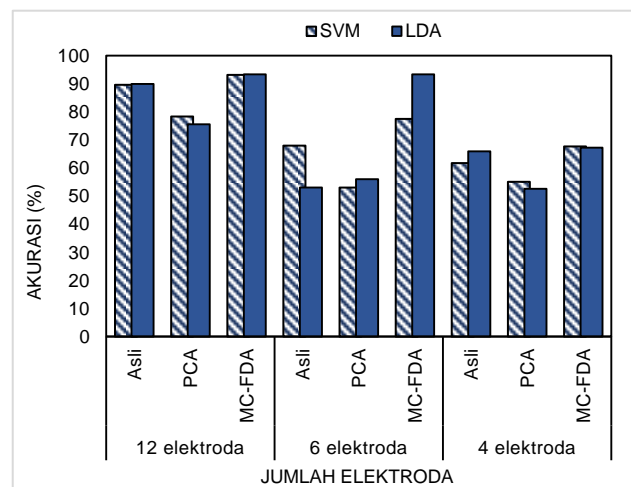
Pada skenario setiap sesi, seluruh subjek dikumpulkan berdasarkan sesi untuk melihat hasil akurasi pengenalan emosi, sehingga jumlah data pada skenario ini sebanyak 15 subjek  $\times$  15 klip film, yakni 225 data EEG. Untuk elektrode yang dilibatkan adalah dua belas elektrode yang sama dengan skenario per subjek, sehingga ukuran dimensi data adalah 5 kanal frekuensi  $\times$  12 elektrode, yakni  $225 \times 60$  vektor ciri data asli pada tiap sesi.

Gbr. 2 mengilustrasikan hasil klasifikasi emosi yang terbagi dalam tiga sesi. Secara umum, tidak terlihat perbedaan yang signifikan antara klasifikasi dengan metode SVM dan LDA untuk semua sesi dan uji coba. Dari rerata akurasi setiap sesi, sesi 1 memperoleh akurasi yang paling baik, disusul kemudian dengan sesi 3 dan sesi 2. Pada setiap sesi, akurasi metode MC-FDA tampak unggul dibanding metode PCA dan data asli.

### C. Hasil Uji Coba pada Skenario Seluruh Data

Jumlah keseluruhan data pada *dataset* SEED adalah 675 baris, yang mencakup lima belas subjek tiga sesi pengambilan data dan lima belas klip film stimulasi. Dimensi vektor ciri pada skenario seluruh data melibatkan 12, 6, dan 4 elektrode, sehingga jumlah ciri secara berurutan adalah 60, 30, dan 20. Uji coba dengan melibatkan jumlah elektrode yang berbeda bertujuan untuk memperlihatkan dampak jumlah elektrode terhadap akurasi. Gbr. 3 menunjukkan hasil akurasi pada masing-masing jumlah elektrode yang terlibat. Dari data percobaan, pada dua belas elektrode dicapai hasil akurasi terbaik, sementara pada empat elektrode mendapat akurasi rendah. Akurasi terbaik dari seluruh skenario ini pada metode MC-FDA dengan klasifikasi LDA dan jumlah elektroda 12, yakni 93,3%. Sementara dengan metode PCA akurasi terbaik dicapai pada dua belas elektrode dengan klasifikasi SVM, yakni sebesar 78,4%. Pada seluruh elektrode, metode PCA mengalami penurunan akurasi terhadap data asli di semua

jumlah elektrode, sedangkan metode MC-FDA tampak bekerja dengan baik dalam meningkatkan akurasi terhadap data asli. Secara rerata (dari SVM dan LDA) kenaikan dari data asli ke MC-FDA untuk dua belas elektrode sebesar 3,5 poin, sedangkan dari data asli ke metode PCA terjadi penurunan akurasi sebesar 12,8 poin.



Gbr. 3 Hasil akurasi pengenalan emosi dari seluruh *dataset* dengan jumlah elektrode yang berbeda.

## IV. PEMBAHASAN

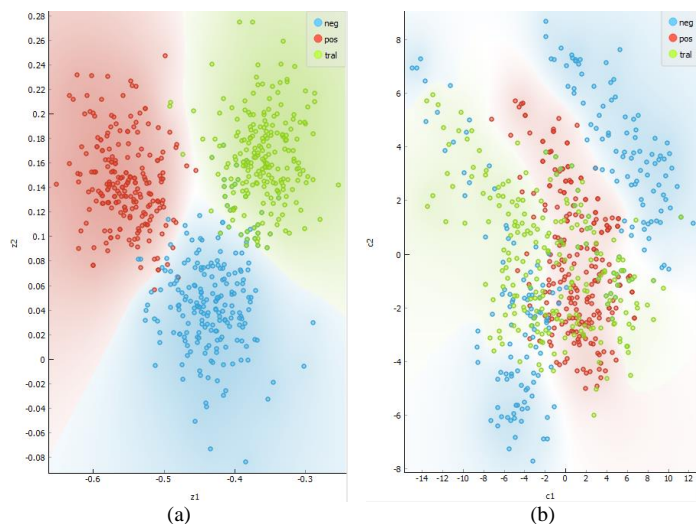
Pada bab ini, beberapa temuan penting dari hasil uji coba di bagian sebelumnya diuraikan. Berdasarkan uji coba pada seluruh skenario, penerapan metode reduksi dimensi MC-FDA terlihat mampu memberikan kontribusi peningkatan akurasi terhadap vektor ciri asli. Metode MC-FDA juga terlihat memberikan kenaikan akurasi pengenalan emosi yang lebih baik bila dibandingkan dengan PCA, baik menggunakan klasifikasi SVM maupun LDA. Pada hasil skenario uji coba dalam setiap subjek yang ditunjukkan pada Tabel I, metode MC-FDA menunjukkan hasil akurasi klasifikasi mencapai 100% akurasi. Hasil peningkatan akurasi yang serupa juga diperoleh pada skenario uji coba setiap minggu dan keseluruhan data. Baik menggunakan klasifikasi SVM maupun LDA, metode MC-FDA mengungguli hasil metode PCA dalam hal akurasi pengenalan emosi. Hal ini membuktikan bahwa metode MC-FDA mampu bekerja dengan baik untuk meningkatkan kinerja akurasi pengenalan emosi dari sinyal EEG dengan tetap mempertahankan sebanyak mungkin informasi elektrode yang dilibatkan.

Pada skenario uji coba per sesi dapat dilihat bahwa *bar plot* biru sesi minggu ke-1 dalam setiap uji coba mengungguli hasil akurasi untuk pengenalan emosi pada sesi minggu ke-2 dan ke-3. Sementara pada skenario uji coba untuk keseluruhan data, terbukti terdapat dampak terhadap akurasi klasifikasi dari jumlah elektrode yang dilibatkan dalam identifikasi. Hasil menunjukkan bahwa semakin besar jumlah elektrode yang dilibatkan, akurasi pengenalan emosi semakin meningkat. Pada dasarnya, baik metode PCA maupun MC-FDA bekerja berdasarkan nilai *eigen* dan vektor *eigen* untuk membentuk proyeksi linear untuk memperoleh dimensi ciri yang telah direduksi. Akan tetapi, perbedaan mendasar dalam hal

pemanfaatan label kelas dari data antara PCA dan MC-FDA memengaruhi hasil proyeksi dari data. Hal tersebut menjadi salah satu penyebab kegagalan PCA bekerja pada reduksi dimensi dalam kasus pengenalan emosi ini. Hal lain yang mungkin menjadi penyebab adalah kalkulasi *inverse* dari matriks *covariance* pada metode MC-FDA, seperti terlihat pada (6), yang tidak dilakukan pada metode PCA.

TABEL II  
HASIL UJI DISKRIMINAN ( $\tau^2$ ) UJI COBA SELURUH DATA DENGAN PCA DAN MC-FDA

Jumlah elektrode/jumlah ciri asli	Jumlah ciri data setelah direduksi		Koefisien tau ( $\tau^2$ )	
	PCA	MC-FDA	PCA	MC-FDA
12 / 60	16	2	0,097	0,728
6 / 30	10	2	0,036	0,452
4 / 20	8	2	0,033	0,340



Gbr. 4 Hasil *scatter plot* data pada skenario dua belas elektrode dengan metode: (a) MC-FDA dan (b) PCA.

Untuk membuktikan lebih lanjut perbedaan hasil nilai diskriminan dari PCA dan MC-FDA pada uji coba seluruh data, perhitungan nilai *Wilk's lambda* dilakukan untuk mengukur keterpisahan antar kelas dalam data. Pada *Wilk's lambda*,  $\Lambda$ , nilai indeks *tau*  $\tau^2$  digunakan sebagai ukuran diskriminan. Nilai indeks *tau*  $\tau^2$  berada antara 0 hingga 1. Semakin besar nilai indeks *tau*, semakin diskriminan variabel dalam *dataset*. Tabel II menunjukkan uji signifikansi diskriminan dari masing-masing subjek pada metode PCA dan MC-FDA. Hasil yang dapat diamati dari rerata nilai indeks *tau* MC-FDA dari semua percobaan pada Tabel II mendekati 1, sementara rerata indeks *tau* PCA mendekati 0. Hasil dari uji diskriminan ini mendukung dugaan semula bahwa MC-FDA dapat meningkatkan kinerja akurasi pengenalan emosi melalui reduksi dimensi. Jumlah ciri setelah direduksi pada Tabel II menunjukkan bahwa rerata jumlah dimensi ciri dari PCA lebih banyak dibanding jumlah ciri dari MC-FDA. Gbr. 4 menunjukkan hasil *scatter plot* dari data PCA dan MC-FDA dari skenario uji coba seluruh data dengan dua belas elektrode. Jelas terlihat dari *scatter plot* pada Gbr. 4(a), bahwa data hasil MC-FDA terbagi jelas dalam tiga kelas

dibanding dengan Gbr. 4(b) data hasil PCA. Hasil ini selaras dengan hasil uji diskriminan yang ditampilkan pada Tabel II.

TABEL III  
PERBANDINGAN KINERJA AKURASI TERBAIK DARI STUDI *DATASET SEED* SEBELUMNYA

Penulis	Metode yang diusulkan	Akurasi terbaik (%)
Zheng dan Lu, 2015 [7]	Ekstraksi ciri menggunakan <i>differential entropy</i>	86,6
Zheng, 2017 [15]	Penentuan pola pengenalan emosi dari EEG yang stabil untuk banyak dataset	91,1
W. Zheng, 2017 [27]	Metode <i>group sparse canonical correlation analysis</i> (GSCCA) untuk seleksi elektroda	80,2
Hasil studi ini	Reduksi dimensi dengan MC-FDA	93,2

Untuk menunjukkan keunggulan MC-FDA, pada Tabel III dipaparkan hasil perbandingan kinerja akurasi pengenalan emosi pada makalah-makalah sebelumnya yang memanfaatkan *dataset* publik yang serupa. Berdasarkan Tabel III, terlihat jelas bahwa akurasi pendekatan MC-FDA unggul dibanding usulan metode sebelumnya. Pada hasil studi dari pengembangan pertama kali *dataset* publik SEED berhasil diidentifikasi ciri *differential entropy* sebagai penciri utama dari pengenalan emosi dari sinyal EEG [7]. Selain itu, hasil dari investigasi pada kanal frekuensi yang penting untuk pengenalan emosi juga ikut dibahas pada penelitian tersebut. Pada studi lanjutan, hasil temuan sebelumnya dikembangkan dengan melakukan identifikasi pola pengenalan emosi dari sinyal EEG yang diterapkan pada lebih dari satu *dataset*. Hasil dari pekerjaan tersebut menunjukkan akurasi 91,1% untuk *dataset SEED*. Sementara, penelitian terdahulu pada usulan metode seleksi ciri mengajukan metode pengurangan jumlah elektrode yang dilibatkan pada proses klasifikasi dengan metode *group sparse canonical correlation analysis* (GSCCA) [27]. Hasil yang diperoleh adalah akurasi 80,2% dari delapan elektrode yang mencakup semua kanal frekuensi dalam sinyal EEG. Dari ketiga studi sebelumnya pada *dataset SEED*, akurasi yang dihasilkan menunjukkan hasil yang paling unggul. Selain itu, jumlah dimensi vektor ciri yang relatif rendah, yaitu dua ciri, menyebabkan waktu proses pembentukan model klasifikasi, baik untuk data pelatihan maupun data uji coba, menjadi lebih cepat.

## V. KESIMPULAN

Pada studi ini, telah dilakukan investigasi secara sistematis untuk menunjukkan peningkatan akurasi pengenalan emosi dari sinyal EEG dengan metode reduksi dimensi yang diusulkan, MC-FDA. Berdasarkan hasil uji coba di setiap skenario, metode MC-FDA berhasil mencapai akurasi terbaik untuk pengenalan emosi dengan kelas yang lebih banyak. Dari hasil perbandingan metode MC-FDA dan PCA, diperoleh bahwa metode MC-FDA mampu bekerja lebih baik dalam memberikan peningkatan akurasi klasifikasi emosi dari vektor ciri asli, baik dengan SVM maupun LDA. Hal ini menunjukkan bahwa metode MC-FDA berhasil menyelesaikan

permasalahan tingginya dimensi pada sinyal EEG untuk kasus pengenalan emosi. Selain itu, metode MC-FDA terlihat lebih unggul dibanding usulan metode sebelumnya terhadap data publik EEG yang sama. Jumlah ciri hasil reduksi dimensi metode MC-FDA yang kecil mengakibatkan waktu komputasi klasifikasi yang lebih singkat. Oleh sebab itu, metode ini dapat dipertimbangkan lebih lanjut untuk diterapkan pada pengenalan emosi dari EEG pada sistem yang *real-time* dan pencapaian akurasi yang tinggi.

Beberapa hal yang masih menjadi tantangan untuk diselesaikan pada pekerjaan selanjutnya berkaitan dengan mengurangi jumlah elektrode yang dilibatkan. Dari segi waktu, persiapan pemasangan perangkat EEG, tingkat kenyamanan pengguna dan jumlah ciri yang di ekstrak, jumlah elektrode yang terlibat perlu untuk dikurangi terutama bagi pengembangan sistem *real-time* pengenalan emosi dari sinyal EEG.

#### UCAPAN TERIMA KASIH

Penulis mengucapkan terima kasih kepada Lembaga Pengelola Dana Pendidikan (LPDP) Kementerian Keuangan Republik Indonesia melalui program Beasiswa Pendidikan Indonesia.

#### REFERENSI

- [1] K.B. Koh, "Emotion and Immunity," *Journal of Psychosomatic Research*, Vol. 45, No. 2, hal. 107–115, 1998.
- [2] J.F. Brosschot dan J.F. Thayer, "Heart Rate Response is Longer After Negative Emotions than After Positive Emotions," *International Journal of Psychophysiology*, Vol. 50, No. 3, hal. 181–187, Nov. 2003.
- [3] C. Busso, Z. Deng, S. Yildirim, M. Bulut, C.M. Lee, A. Kazemzadeh, S. Lee, U. Neumann, dan S. Narayanan, "Analysis of Emotion Recognition Using Facial Expressions, Speech and Multimodal Information," *Proceedings of the 6th international conference on Multimodal interfaces*, 2004, hal. 205–211.
- [4] S. Koelstra, C. Muhl, M. Soleymani, J.-S. Lee, A. Yazdani, T. Ebrahimi, T. Pun, A. Nijholt, dan I. Patras, "DEAP: A Database for Emotion Analysis; Using Physiological Signals," *IEEE Transactions on Affective Computing*, Vol. 3, No. 1, hal. 18–31, 2012.
- [5] A.M. Bhatti, M. Majid, S.M. Anwar, and B. Khan, "Human Emotion Recognition and Analysis in Response to Audio Music Using Brain Signals," *Computers in Human Behavior*, Vol. 65, hal. 267–275, Dec. 2016.
- [6] C. Mühl, B. Allison, A. Nijholt, dan G. Chanel, "A Survey of Affective Brain Computer Interfaces: Principles, State-of-the-Art, and Challenges," *Brain-Computer Interfaces*, Vol. 1, No. 2, hal. 66–84, 2014.
- [7] W.-L. Zheng dan B.-L. Lu, "Investigating Critical Frequency Bands and Channels for EEG-Based Emotion Recognition with Deep Neural Networks," *IEEE Transactions on Autonomous Mental Development*, Vol. 7, No. 3, hal. 162–175, 2015.
- [8] I. T. Jolliffe dan J. Cadima, "Principal Component Analysis: A Review And Recent Developments," *Phil. Trans. R. Soc. A*, Vol. 374, No. 2065, hal. 20150202, 2016.
- [9] A. Subasi dan M.I. Gursoy, "EEG Signal Classification Using PCA, ICA, LDA and Support Vector Machines," *Expert Systems with Applications*, Vol. 37, No. 12, hal. 8659–8666, Dec. 2010.
- [10] F. Artoni, A. Delorme, dan S. Makeig, "Applying Dimension Reduction to EEG Data by Principal Component Analysis Reduces the Quality of Its Subsequent Independent Component Decomposition," *NeuroImage*, Vol. 175, hal. 176–187, 2018.
- [11] A. M. Martínez dan A.C. Kak, "PCA Versus LDA," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 23, No. 2, hal. 228–233, 2001.
- [12] M. Welling, "Fisher Linear Discriminant Analysis," University of Toronto, Technical Note, 2005.
- [13] P. Xanthopoulos, P.M. Pardalos, dan T.B. Trafalis, "Linear Discriminant Analysis," in *Robust Data Mining*, New York, USA: Springer, 2013, hal. 27–33.
- [14] Y.-H. Liu, W.-T. Cheng, Y.-T. Hsiao, C.-T. Wu, dan M.-D. Jeng, "EEG-based Emotion Recognition Based on Kernel Fisher's Discriminant Analysis and Spectral Powers," *2014 IEEE International Conference on Systems, Man and Cybernetics (SMC)*, 2014, pp. 2221–2225.
- [15] W.-L. Zheng, J.-Y. Zhu, dan B.-L. Lu, "Identifying Stable Patterns over Time for Emotion Recognition from EEG," *IEEE Transactions on Affective Computing*, pp. 1–1, 2017.
- [16] R. Jenke, A. Peer, dan M. Buss, "Feature Extraction and Selection for Emotion Recognition from EEG," *IEEE Transactions on Affective Computing*, Vol. 5, No. 3, hal. 327–339, Jul. 2014.
- [17] R. Darmakusuma, A.S. Prihatmanto, A. Indrayanto, dan T.L. Mengko, "Deteksi Intensi Pergerakan Jari Menggunakan Metode Power Spectral Density dengan Stimulus Visual," *Jurnal Nasional Teknik Elektro dan Teknologi Informasi (JNTETI)*, Vol. 4, No. 2, hal. 125-129. Dec. 2015.
- [18] A. Surtono, T.S. Widodo, and M. Tjokronagoro, "Analisis Klasifikasi Sinyal EKG Berbasis Wavelet dan Jaringan Syaraf Tiruan," *JNTETI*, vol. 1, no. 3, hal. 60-66, 2012.
- [19] R.-N. Duan, J.-Y. Zhu, dan B.-L. Lu, "Differential Entropy Feature for EEG-based Emotion Classification," *2013 6th International IEEE/EMBS Conference on Neural Engineering (NER)*, 2013, hal. 81–84.
- [20] H. Anton, *Elementary Linear Algebra*, New York: Wiley, 1987.
- [21] I.H. Witten, E. Frank, M.A. Hall, dan C.J. Pal, *Data Mining: Practical Machine Learning Tools and Techniques*, Burlington, USA: Morgan Kaufmann, 2016.
- [22] J. Atkinson dan D. Campos, "Improving BCI-based Emotion Recognition by Combining EEG Feature Selection and Kernel Classifiers," *Expert Systems with Applications*, Vol. 47, hal. 35–41, Apr. 2016.
- [23] X.-W. Wang, D. Nie, dan B.-L. Lu, "Emotional State Classification From EEG Data Using Machine Learning Approach," *Neurocomputing*, Vol. 129, hal. 94–106, Apr. 2014.
- [24] Y.-P. Lin, C.-H. Wang, T.-P. Jung, T.-L. Wu, S.-K. Jeng, J.-R. Duann, dan J.-H. Chen, "EEG-Based Emotion Recognition in Music Listening," *IEEE Transactions on Biomedical Engineering*, Vol. 57, No. 7, hal. 1798–1806, Jul. 2010.
- [25] C. Cortes, dan V. Vapnik, "Support-Vector Networks," *Machine Learning*, Vol. 20, No. 3, hal. 273–297, 1995.
- [26] J. Platt, "Sequential Minimal Optimization: A Fast Algorithm for Training Support Vector Machines," Microsoft Research, Technical Report MSR-TR-98-14, hal. 1-21, 1998.
- [27] W. Zheng, "Multichannel EEG-Based Emotion Recognition via Group Sparse Canonical Correlation Analysis," *IEEE Transactions on Cognitive and Developmental Systems*, Vol. 9, No. 3, hal. 281–290, Sep. 2017.