

Identifikasi Emosi Melalui Sinyal EEG menggunakan 3D-Convolutional Neural Network

RINDU TEGAR SENJAWATI, *ESMERALDA CONTESSA DJAMAL,
FATAN KASYIDI

Program Studi Teknik Informatika, Universitas Jenderal Achmad Yani, Indonesia

*Email : esmeralda.contezza@lecture.unjani.ac.id

Received 17 November 2023 | *Revised* 11 Desember 2023 | *Accepted* 25 Desember 2023

ABSTRAK

Emosi memberikan peran penting dalam interaksi manusia yang didapat melalui respon yang tepat. Respon yang tak tepat menunjukkan adanya gangguan mental sehingga diperlukan identifikasi emosi. Identifikasi dapat dilakukan menggunakan aktivitas sinyal listrik di otak menggunakan Elektroensefalogram (EEG). Karena sinyal EEG pada setiap kanal merupakan urutan data maka dijadikan multi-kanal yang direpresentasikan pada matriks agar urutan-urutan data tetap terjaga. Penggunaan matriks memadukan informasi dari ketiga dimensi (kanal x frekuensi x waktu) dapat menggambarkan kompleksitas dari sinyal EEG. Sehingga dapat mengenali pola aktivitas otak pada rentang frekuensi tertentu berkembang sepanjang waktu. Untuk menangkap informasi tersebut perlu dilakukan ekstraksi fitur agar mewakili variabel-variabel emosi. Ekstraksi dilakukan pada domain frekuensi (4-45 Hz) dan waktu menggunakan Short Time Fourier Transform (STFT) kemudian idenitifikasi menggunakan 3D Convolutional Neural Network (CNN). Eksperimen menggunakan 3D CNN menghasilkan akurasi 65.45 dengan teknik koreksi bobot Adamax.

Kata kunci: *emosi, sinyal EEG, multi-kanal, STFT, 3D-CNN*

ABSTRACT

Emotions play an important role in human interaction through appropriate responses. Inappropriate responses indicate a mental disorder, so identification of emotions is required. Identification can be done using electrical signal activity in the brain with Electroencephalogram (EEG). Because the EEG signal in each channel is a data sequence, it is made into a multi-channel represented in a matrix so that the data sequence is maintained. Using a matrix combining information from all three dimensions (channel x frequency x time) can describe the complexity of the EEG signal. Allowing recognition of evolving brain activity patterns within specific frequency ranges over time. Extraction is done in the frequency domain (4-45 Hz) and time using Short Time Fourier Transform (STFT), then identification using a 3D Convolutional Neural Network (CNN). Experiments using 3D CNN resulted in an accuracy of 65.45 with the Adamax weight correction technique.

Keywords: *emotion, EEG signal, multi-channel, STFT, 3D-CNN*

1. PENDAHULUAN

Emosi adalah suatu kondisi yang mengekspresikan kesadaran manusia dan digambarkan sebagai reaksi yang didapat melalui proses penilaian terhadap sesuatu secara positif maupun negatif (**Nematollahi & Firoozabadi, 2018**). Emosi positif merupakan respon atau perasaan menyenangkan yang timbul akibat stimulus dalam konteks hal yang positif. Sedangkan emosi negatif merupakan respon atau perasaan yang tidak menyenangkan atau mengganggu dan biasanya diekspresikan sebagai bentuk ketidaksukaan yang timbul akibat stimulus dalam konteks hal yang negatif. Kedua emosi tersebut dapat berdampak buruk jika terjadi dengan intensitas berlebih sehingga mengakibatkan kesehatan mental yang terganggu seperti depresi (**Deshpande & Rao, 2018**). Seseorang yang mengalami depresi tidak mengetahui cara menunjukkan emosi yang tepat sebagai bentuk reaksinya, sehingga diperlukan identifikasi emosi.

Identifikasi emosi dapat dilakukan dengan menggunakan sinyal fisiologis untuk mengetahui aktivitas sinyal listrik di otak yang ditangkap menggunakan Elektroensefalogram (EEG) (**Yosi, dkk, 2019**). EEG digunakan karena memberikan pola yang kompleks dari aktivitas yang direkam dalam sepersekian detik setelah stimulus diberikan. Selain itu EEG bersifat *high speed* dan non-invasif (**Houssein, dkk, 2022**) yang menyebabkan EEG dapat mendeteksi emosi nyata yang muncul dari pikiran dengan menghiraukan fitur eksternal seperti gerak tubuh maupun ekspresi wajah. Penelitian terdahulu menggunakan sinyal EEG untuk melakukan identifikasi emosi saat mendengarkan musik yang terbagi ke dalam dua dimensi emosi : yaitu *arousal* dan *valence* (**Kumar, 2022**). Dan penelitian lain melakukan identifikasi emosi saat menonton film China yang terbagi ke dalam tiga kelas emosi: yaitu positif, netral dan negatif (**Zheng & Lu, 2015**).

Hasil dari perekaman sinyal EEG mengandung *noise*, sehingga diperlukan filter sebagai pembeda pola untuk mendapatkan sinyal perekaman yang sebenarnya (**Negi, dkk, 2021**). Filterisasi sinyal tersebut dapat dibagi dalam filterisasi domain waktu, domain frekuensi serta domain waktu-frekuensi. Penelitian terdahulu melakukan filterisasi sinyal dalam domain frekuensi dengan rentang frekuensi 0-45 Hz yaitu Delta (0-3 Hz), Teta (4-8 Hz), Alfa (8-12 Hz), Beta (12-30 Hz), dan Gamma (30-45 Hz) untuk identifikasi emosi: normal, fokus, sedih dan terkejut sehingga menghasilkan akurasi 83.33% (**Yudhana, dkk, 2020**). Kemudian penelitian selanjutnya menggabungkan informasi dari kedua domain yaitu waktu-frekuensi menggunakan *Short Time Fourier Transform* (STFT) melalui rentang frekuensi 4-45 Hz yaitu Teta (4-8 Hz), Alfa (8-12 Hz), Beta (12-30 Hz), dan Gamma (30-45 Hz) untuk melakukan identifikasi emosi ke dalam tiga kelas: senang, netral dan sedih sehingga menghasilkan akurasi 89,17% (**Gao, dkk, 2020**).

Setelah melalui proses filterisasi sinyal maka proses selanjutnya adalah identifikasi dan klasifikasi dengan pembelajaran mesin, seperti menggunakan Convolutional Neural Networks (CNN). CNN digunakan karena bekerja dengan melihat kondisi sesaat tanpa mengingat informasi sebelumnya. Dan representasi dari CNN dapat dibagi ke dalam beberapa jenis dimensi. Kelebihan CNN yang menggunakan lebih dari satu dimensi adalah *input* tidak akan kehilangan informasi spasialnya yang akan diekstraksi fiturnya dan diklasifikasi.

Karena sinyal EEG merupakan urutan data pada setiap kanal maka penelitian terdahulu merepresentasikannya ke dalam bentuk satu dimensi menggunakan 1D CNN untuk identifikasi emosi dalam empat kelas: yaitu *high-arousal high-valence* (HAHV), *high-arousal low valence* (HALV), *low-arousal high-valence* (LAHV) dan *low-arousal low-valence* (LALV) (**Zamani & Wulansari, 2021**). Karena urutan tersebut tidak selalu terjaga maka dilakukan penelitian selanjutnya dengan menggunakan multi-kanal yang direpresentasikan ke dalam bentuk

matriks berupa matriks dua dimensi menggunakan 2D CNN dengan representasi dimensi vertikal berupa kanal dan horizontal berupa waktu untuk identifikasi emosi ke dalam tiga kelas: yaitu positif, netral dan negatif sehingga memberikan akurasi sebesar 83,44% (**Haqqe, dkk, 2021**). Untuk mendapatkan informasi yang lebih mendalam terhadap fitur yang digunakan maka beberapa penelitian lain menggunakan 3D CNN dengan representasi fitur yang beragam (**Zheng, dkk, 2021**). Seperti penelitian lain, penggunaan 3D CNN dibangun untuk integrasi antar fitur spasial dan temporal dengan mengubah fitur spasial menjadi dua dimensi sesuai dari topologi elektroda dan menambahkan fitur temporal sehingga akurasi yang diberikan 72,1% untuk *valence* dan 73,1% untuk *arousal* (**Zheng, dkk, 2021**). Penelitian lain merepresentasikan tiga dimensi menjadi data dari setiap kanal yang disegmen menjadi *frame*. Jumlah *frame* digabungkan untuk membentuk matriks dua dimensi. Kemudian fitur temporal ditambahkan dengan memilih sejumlah *frame* berturut-turut sehingga menghasilkan akurasi 87.44% untuk *valence* dan 88.49% untuk *arousal* (**Salama, dkk, 2018**).

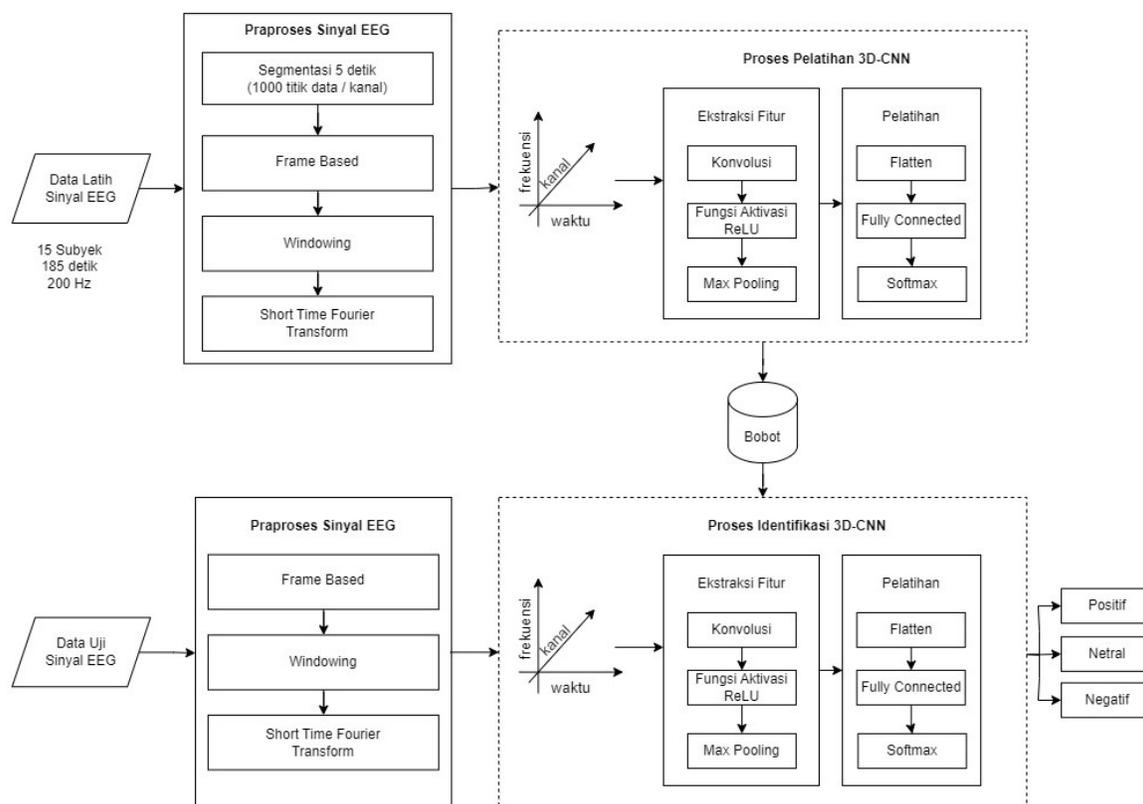
Berbeda dari penelitian terdahulu, dalam masukan tiga dimensi pada sinyal EEG hanya fokus pada dua jenis fitur (spasial versus temporal) sehingga untuk mendapatkan informasi yang lebih dalam maka menggunakan ketiga fitur (spasial versus temporal versus frekuensi) digabungkan dan digunakan secara menyeluruh. Fitur spasial pada sinyal EEG berkaitan dengan distribusi spasial potensial listrik di permukaan kepala saat merekam sinyal EEG dari beberapa elektroda. Sedangkan fitur temporal mencerminkan perubahan dalam waktu dari sinyal EEG. Ini melibatkan analisis variasi sinyal sepanjang waktu perekaman. Artinya pemrosesan sinyal EEG dalam domain frekuensi – waktu menguntungkan karena sinyal EEG dengan variabel emosi berelasi dengan frekuensi (**Zhang, dkk, 2022**).

Penggunaan 3D CNN mengintegrasikan representasi frekuensi, waktu, dan informasi kanal dalam satu model memungkinkan model untuk memahami perubahan aktivitas otak secara lebih mendalam. Model dapat secara efektif mengeksplorasi hubungan kompleks dalam data seperti mengenali bagaimana pola aktivitas otak pada rentang frekuensi tertentu berkembang sepanjang waktu di berbagai bagian otak. Hal ini memungkinkan analisis yang lebih mendalam tentang dinamika perubahan aktivitas otak dan memberikan pemahaman yang lebih holistik tentang korelasi antara waktu, frekuensi, dan kanal.

Sehingga penelitian yang akan dilakukan adalah pembelajaran mesin untuk mengidentifikasi emosi melalui sinyal EEG ke dalam tiga kelas yaitu "Negatif", "Netral", dan "Positif". Ketiga kelas tersebut dipilih karena merupakan emosi dasar dan umum dan memiliki dampak yang signifikan terhadap perilaku manusia. STFT digunakan untuk mendapatkan informasi spektrum frekuensi pada seluruh nilai amplitudo yang telah melalui proses *windowing* sekaligus menghilangkan *noise* (**Wang, dkk, 2019**). Selain itu, STFT membantu dalam menganalisis bagaimana spektrum frekuensi dari sinyal EEG berubah seiring waktu. Sinyal EEG yang telah difilter lalu diidentifikasi menggunakan CNN dengan representasi tiga dimensi. Data sinyal EEG emosi yang digunakan pada penelitian ini berasal dari *SJTU Emotion EEG Dataset (SEED)*.

2. METODE

Perancangan diawali dari perolehan data sinyal EEG dari penelitian sebelumnya (**Zheng & Lu, 2015**), kemudian dilanjutkan dengan tahap rancangan sistem identifikasi emosi yang diawali dengan proses segmentasi data kemudian ekstraksi fitur dengan STFT dan identifikasi menggunakan 3D CNN untuk menghasilkan kelas emosi berupa positif, netral dan negatif. Rancangan sistem identifikasi emosi dapat dilihat pada Gambar 1.



Gambar 1. Sistem Identifikasi Emosi dengan STFT dan CNN

2.1 Perolehan Data

Perolehan data sinyal EEG yang diperoleh dari *SJTU Emotion EEG Dataset (SEED)* (Lan, dkk, 2019). Data yang didapatkan merupakan hasil perekaman dari 62 kanal dengan frekuensi *sampling* 1000 Hz. Hasil perekaman dilakukan pada 15 subyek berusia 19-28 tahun yang merupakan mahasiswa Universitas Jiao Tong Shanghai. Setiap subyek diberikan stimulus atau rangsangan emosi dari 15 klip film China dengan durasi dari 185-265 detik yang digunakan.

Kanal yang digunakan sebanyak 12, yaitu FT7, FT8, T7, T8, C5, C6, TP7, TP8, CP5, CP6, P7 dan P8 yang dinilai lebih efektif daripada 62 kanal (Zheng & Lu, 2015). Sehingga menghasilkan 675 set data dari 15 subyek x 15 perekaman x 3 perulangan. Waktu yang diproses dalam segmentasi adalah 185 detik dari setiap percobaan. Durasi waktu tersebut disegmentasi setiap lima detik dengan frekuensi *sampling* yang telah di *down-sampling* menjadi 200 Hz, sehingga menghasilkan 37 segmen dengan panjang data pada setiap kanalnya sebanyak 1000 titik data. *Down-sampling* digunakan untuk membuang informasi yang kurang berguna dan mengurangi dimensi data tanpa mengorbankan informasi yang penting. Sehingga total set data yang didapatkan dari hasil segmentasi untuk satu kali percobaan yang menghasilkan 15 subyek x 15 perulangan x 3 kelas x 37 segmen, yaitu 24.975 set data untuk tiga kali percobaan

2.2 Short Time Fourier Transform

STFT digunakan untuk memproses sinyal EEG yang telah diproses sebelumnya, dan kemudian mengekstrak fitur spektrum daya. Proses perhitungan STFT adalah membagi sinyal waktu yang lebih panjang menjadi segmen-segmen yang lebih pendek dengan panjang yang sama dan menghitung spektrum *Fourier* pada setiap segmen yang lebih pendek (Gao, dkk, 2020). Diawali dengan pra-proses yaitu *frame based* dan *windowing*.

Frame based membagi sinyal menjadi beberapa *frame* yang jumlah titik sinyalnya telah ditentukan untuk meminimalkan pengaruh non-stasioner pada sinyal EEG dengan perhitungan *overlap* tiap *frame* sepanjang 50%, menggunakan Persamaan (1) dan (2) (Indrawan, dkk, 2017).

$$M = N * 0.5 \quad (1)$$

$$x(n) = (n + M) \quad (2)$$

Dimana $M = \text{overlapping frame}$, $N = \text{jumlah data}$, $n = \text{indeks data}$, $(n + M) = \text{nilai sinyal hasil perekaman}$, $(n) = \text{hasil frame based}$.

Sinyal yang telah dibagi menjadi beberapa *frame* dilakukan proses *windowing* untuk mengurangi *discontinue* pada ujung setiap *frame* dari proses *frame based* dan untuk mereduksi *noise*. Beberapa fungsi *window*, salah satunya adalah *Hamming* karena memiliki *side-lobe* yang tidak terlalu besar dan tidak terlalu kecil sehingga *noise* yang dihasilkan tidak terlalu besar. Perhitungan *windowing* dapat dilihat pada Persamaan (3) dan (4).

$$w(n) = 0.54 - 0.46 \cos \frac{2\pi n}{N-1} \quad (3)$$

$$z(n) = x(n)w(n) \quad (4)$$

Dimana $z(n) = \text{nilai window ke } - n$, $N = \text{jumlah sampel setiap frame}$, $n = \text{indeks sampel suatu frame}$, $(n) = \text{sinyal hasil windowing sample ke } - n$, $x(n) = \text{sample sinyal ke } - n$, $w(n) = \text{nilai jenis window ke } - n$.

Proses *windowing* menghasilkan spektrum domain waktu, agar tidak terjadi kesalahan maka spektrum domain waktu diubah menjadi sinyal frekuensi dengan menggunakan proses STFT. Fitur spektrum daya dari sinyal EEG pada rentang frekuensi yang berbeda berkorelasi dengan emosi (4-45 Hz) dengan Teta (4-7 Hz), Alfa (8-12 Hz), Beta (13-30 Hz), dan Gamma (30-47 Hz). Sehingga STFT didefinisikan melalui Persamaan (5).

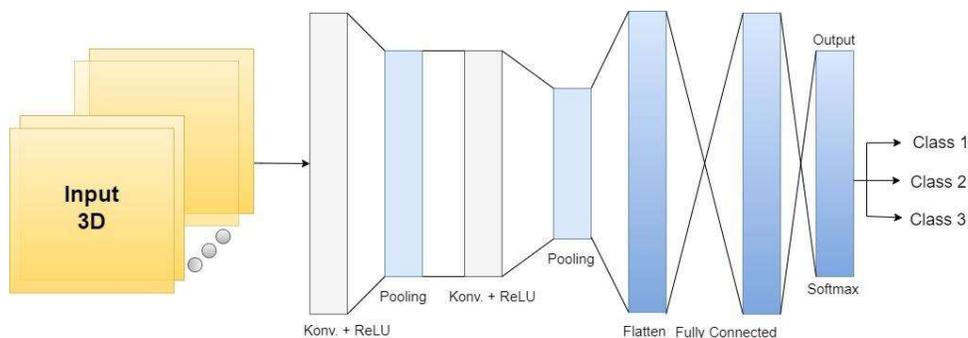
$$\sum_{n=-\infty}^{\infty} x(n)\omega(n - mR)e^{-j\omega n} \quad (5)$$

Dimana $x(n)$ adalah sinyal *input* dari waktu n , $\omega(n)$ adalah fungsi *windowing* yang panjangnya m , R adalah ukuran titik lompatan antara *Discrete Time Fourier Transform* (DTFT) dan j adalah unit imajiner.

2.3 Convolutional Neural Network

Convolutional Neural Network atau CNN/*Conv* merupakan salah satu algoritma yang ada pada *deep learning* yang dirancang untuk mengolah data dalam beberapa jenis layer yaitu satu, dua dan tiga dimensi. CNN bekerja seperti mata dengan melihat kondisi sesaat tanpa mengingat informasi sebelumnya (Jang & Tong, 2021) sehingga banyak digunakan dalam pengolahan gambar, kata, video maupun *time series* dengan menyesuaikan dimensi operasi konvolusi.

CNN terdiri dari dua buah komponen, yaitu lapisan ekstraksi fitur dan lapisan identifikasi. Pada lapisan ekstraksi fitur terdiri dari lapisan konvolusi, fungsi aktivasi *Rectified Linear Units (ReLU)* dan *Pooling*. Sedangkan pada lapisan identifikasi terdiri dari proses *Flatten*, *Fully-Connected Layer* dan *Softmax*.



Gambar 2. Arsitektur 3D CNN

2.3.1 Ekstraksi Fitur

Tahapan pertama yang dilakukan pada proses ekstraksi fitur adalah lapisan konvolusi yaitu proses mengubah data *input* dengan menggeser filter atau kernel untuk mengekstrak fitur yang ada. Sehingga fitur-fitur yang berhasil diekstrak kemudian disimpan pada *feature map*. Hasil proses konvolusi akan menghasilkan keluaran yang lebih kecil daripada data masukan, sehingga informasi yang terkandung dalam data asli tidak hilang secara drastis pada dimensi *output* maka hasil konvolusi dapat dimanipulasi menggunakan *padding* dengan menambahkan nilai nol pada piksel. Fungsi konvolusi tiga dimensi dapat dilihat pada Persamaan (6).

$$CC[m, n, o] = \sum_u \sum_v \sum_w A[m + u, n + v, o + w] \cdot B[u, v, w] \quad (6)$$

Dimana $C[m, n, o]$ diperoleh dari perkalian satu elemen A dengan satu elemen B lalu ditambahkan. Kemudian fungsi untuk menghitung lebar *feature map* dapat dihitung menggunakan Persamaan (7).

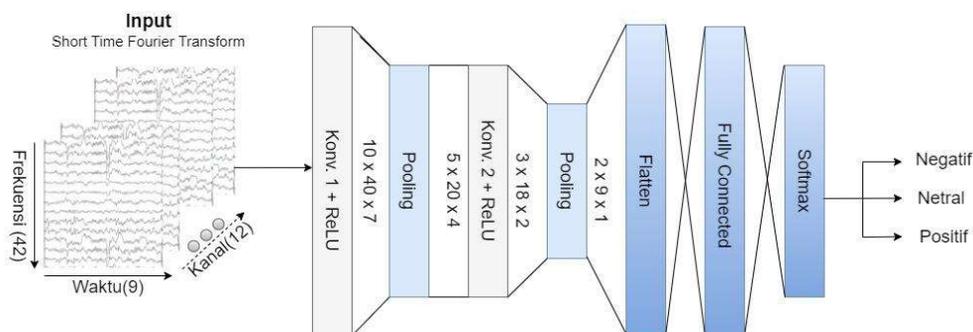
$$output = \frac{N-F+2P}{s} + 1 \quad (7)$$

Dimana N adalah lebar *input*, F adalah lebar kernel, P adalah jumlah *padding* dan S adalah *stride*.

Tahapan yang dilakukan setelah proses konvolusi selesai adalah proses aktivasi menggunakan fungsi aktivasi *Rectified Linear Units (ReLU)* untuk merubah semua nilai negatif menjadi nol. Adapun fungsi aktivasi yang digunakan pada Persamaan (8).

$$f(x) = \max(0, x) \quad (8)$$

Tahap terakhir pada proses ekstraksi fitur adalah Lapisan Pooling yang merupakan proses untuk mereduksi ukuran spasial dan jumlah parameter dalam jaringan serta mempercepat komputasi dan mengontrol terjadinya *overfitting*. Terdapat dua jenis *Pooling* yang biasa digunakan yakni *Max Pooling* dan *Average Pooling*. Dalam *Max Pooling*, *filter* hanya memilih nilai piksel maksimum di bidang reseptif sedangkan *Average Pooling* bekerja dengan menghitung nilai rata-rata nilai piksel di bidang reseptif. Hasil dari ekstraksi fitur dapat dilihat pada Gambar 3.



Gambar 3. Ekstraksi Fitur dan Identifikasi dengan 3D CNN

2.3.2 Identifikasi

Lapisan identifikasi berfungsi untuk mengidentifikasi setiap *neuron* yang telah diekstraksi pada lapisan ekstraksi fitur, karena itu neuron akan terhubung sepenuhnya ke arsitektur *Multi-Layer Perceptron (MLP)*. Tahap pertama pada lapisan ini adalah *Flatten* sehingga memuat *input* yang memiliki banyak dimensi menjadi sebuah vektor agar dapat digunakan sebagai *input* dari tahapan *Fully-Connected Layer*.

Kemudian tahapan kedua adalah lapisan *Fully-Connected*, *neuron* akan terhubung ke semua angka dalam volume sehingga skor kelas dapat dihitung. Sehingga tahap terakhir adalah melakukan klasifikasi dengan menggunakan *Softmax*, dimana *Softmax* akan menghitung probabilitas dari setiap kelas target atas semua kelas target yang memungkinkan dan akan membantu untuk menentukan kelas target untuk *input* yang diberikan.

3. HASIL DAN PEMBAHASAN

3.1 Satu Dimensi, Dua Dimensi dan Tiga Dimensi CNN

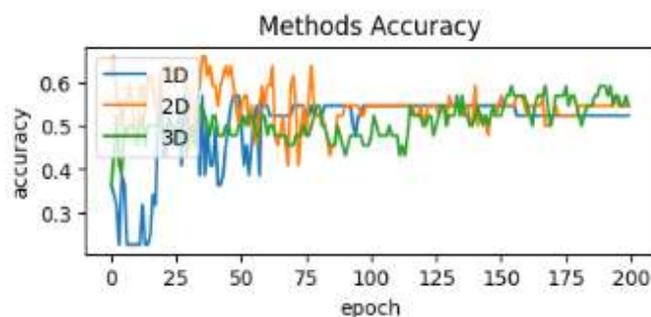
Eksperimen pertama adalah melakukan perbandingan identifikasi emosi menggunakan tipe dimensi pada CNN yaitu satu dimensi, dua dimensi dan tiga dimensi dengan hasil yang ditunjukkan pada Tabel 1.

Tabel 1. Perbandingan 1D CNN, 2D CNN dan 3D CNN

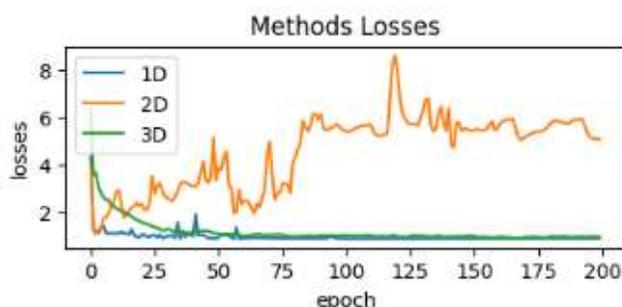
Metode	Akurasi (%)	Loss	Waktu (detik)
Satu-Dimensi CNN	60.00	1.302	7.030
Dua-Dimensi CNN	61.81	3.669	3.366
Tiga-Dimensi CNN	65.45	1.210	6.379

Berdasarkan Tabel 1, menunjukkan bahwa penggunaan tiga dimensi CNN yang merepresentasikan kanal x waktu x frekuensi menghasilkan akurasi yang lebih tinggi yaitu 65.45% dan penggunaan dua dimensi menghasilkan akurasi 61.81% dan untuk urutan data yang menggunakan satu dimensi menghasilkan akurasi 60.00%. Dengan hasil tersebut maka menggunakan multi-kanal yang merepresentasikan EEG menjadi matriks menghasilkan akurasi yang lebih tinggi dibandingkan sinyal EEG yang berbentuk dalam urutan.

Untuk itu juga penggunaan tiga dimensi memberikan akurasi lebih tinggi karena model dapat memperoleh pemahaman yang lebih mendalam tentang perubahan aktivitas otak seiring waktu dalam berbagai rentang frekuensi, dan juga bagaimana aktivitas ini berbeda di berbagai saluran otak. Namun penggunaan ketiga dimensi menghasilkan selisih akurasi yang tidak terlalu jauh terlebih penggunaan dua dimensi CNN memberikan waktu yang lebih singkat.



Gambar 4. Akurasi Perbandingan 1D CNN, 2D CNN dan 3D CNN



Gambar 5. Loss Perbandingan 1D CNN, 2D CNN dan 3D CNN

Pada Gambar 4 terlihat jika akurasi 1D CNN stabil dari *epoch* ke 160, sedangkan pada 2D CNN stabil pada *epoch* ke 177. Untuk 3D CNN belum stabil pada *epoch* ke 200 sehingga bisa ditambahkan beberapa *epoch*. Dan pada Gambar 5 dimana *loss* dari 1D CNN relatif stabil dan 3D CNN juga memberikan *loss* yang tinggi namun selama berjalannya *epoch* maka *loss* yang diberikan stabil, sedangkan untuk 2D CNN memberikan *epoch* yang belum stabil sehingga terus-menerus naik dan tinggi dibanding dengan *loss* dari kedua metode lain.

3.2 Layer Teknik Koreksi Bobot

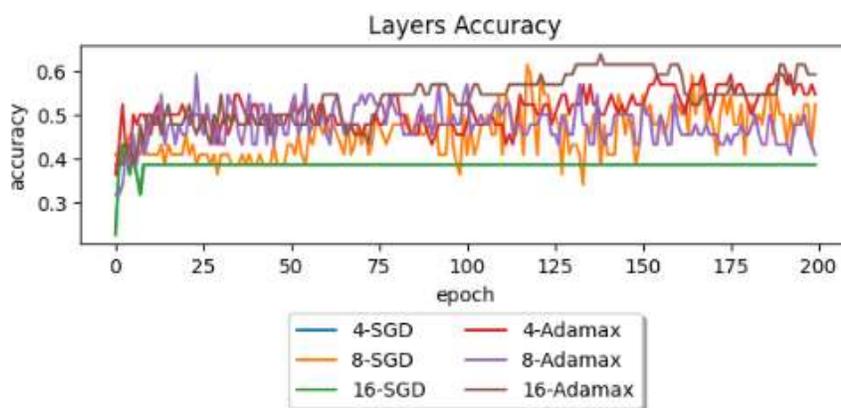
Eksperimen kedua yang dilakukan penelitian ini membandingkan konfigurasi metode yang optimal seperti menggunakan teknik koreksi bobot dan layer CNN. Ada dua teknik yang digunakan, *Stochastic Gradient Descent (SGD)*, dan *AdaMax* (perpanjangan dari *gradient descent* versi *Adam* yang menggeneralisasi pendekatan ke *infinite norm max*), serta layer yang digunakan berupa 4, 8 dan 16.

Tabel 2. Perbandingan Layer Teknik Koreksi Bobot

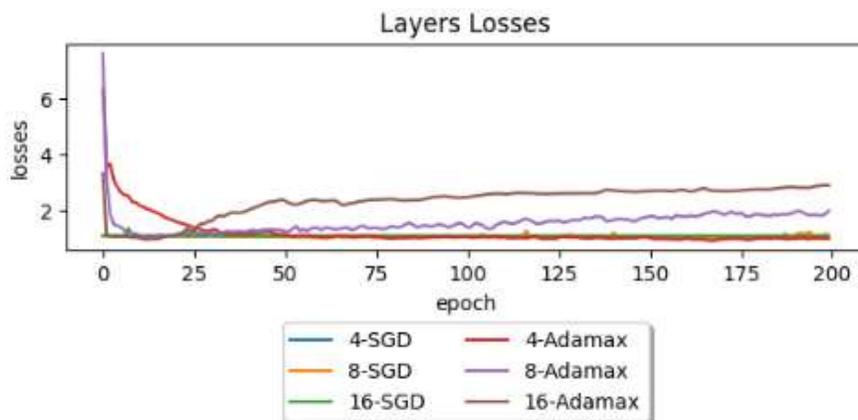
CNN Layer	Teknik Koreksi Bobot				Waktu (detik)	
	Akurasi (%)		Loss			
	SGD	Adamax	SGD	Adamax	SGD	Adamax
4	40.00	65.45	1.064	1.210	3.375	6.379
8	34.54	61.81	1,419	1.526	3.375	6.386
16	45.45	38.18	1.021	4.163	6.381	7.126

Berdasarkan Tabel 2, hasil dari penggabungan STFT dan 3D CNN memiliki akurasi mencapai 65.45% dengan teknik koreksi *AdaMax* serta menggunakan 4 layer CNN. Dimana penggunaan teknik koreksi bobot *AdaMax* memberikan akurasi yang lebih tinggi dibandingkan dengan SGD, namun SGD memberikan waktu yang lebih singkat yaitu 3-6 menit. Dalam teknik SGD maka

layer yang lebih tinggi dalam memberikan akurasi yaitu dengan 3D CNN dalam 16 layer, sedangkan untuk *AdaMax* layer yang memberikan akurasi tertinggi ada pada layer 4.



Gambar 6. Akurasi Perbandingan Layer Teknik Koreksi Bobot



Gambar 7. Loss Perbandingan Layer Teknik Koreksi Bobot

Pada Gambar 6 akurasi pada penggunaan 3D CNN 16 layer menggunakan teknik koreksi bobot SGD memberikan akurasi yang stabil pada *epoch* ke 5. Dan untuk *loss* dari SGD dan *AdaMax* baik dengan menggunakan 4, 8 dan 16 layer CNN memberikan *loss* yang cukup stabil, hanya 16 layer *AdaMax* saja yang *loss*-nya naik dan lebih tinggi daripada *loss* lainnya.

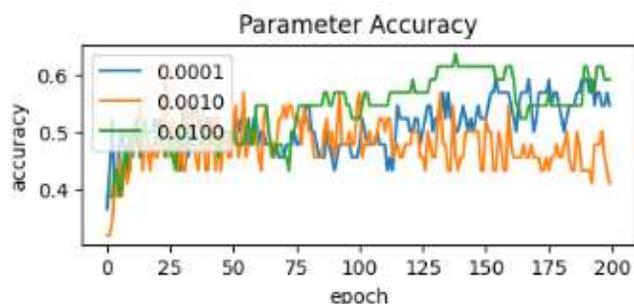
3.3 Parameter Teknik Koreksi Bobot

Eksperimen ketiga yang dilakukan penelitian ini membandingkan parameter pada teknik koreksi bobot yaitu dengan *learning rate* 0.0001, 0.0010 dan 0.0100 serta teknik koreksi bobot yang digunakan yaitu *AdaMax*. *Learning rate* digunakan sebagai optimisasi model, mencegah *divergensi* dimana model yang tidak konvergen dan menjau dari nilai minimum serta untuk menyesuaikan kecepatan pembelajaran.

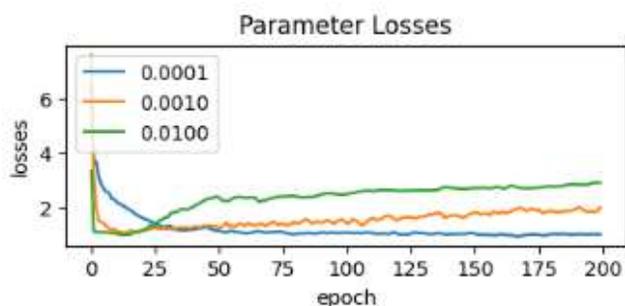
Tabel 3. Perbandingan Parameter Teknik Koreksi Bobot

Learning Rate	Akurasi (%)	Loss	Waktu (detik)
0.0001	65.45	1.210	6.379
0.0010	61.81	1.526	6.386
0.0100	38.18	4.163	7.126

Berdasarkan Tabel 3, akurasi tertinggi adalah dengan penggunaan parameter *learning rate* 0.0001. Selain akurasi yang lebih tinggi, *loss* yang dihasilkan pun lebih rendah dan waktu yang digunakan model pun lebih rendah yaitu 6.379 detik.



Gambar 8. Akurasi Perbandingan Parameter Teknik Koreksi Bobot



Gambar 9. Loss Perbandingan Parameter Teknik Koreksi Bobot

Pada Gambar 8 akurasi penggunaan *learning-rate* 0.0100 stabil pada *epoch* ke 195, dan pada Gambar 9 *loss* ketiga parameter stabil hanya saat terjadi kenaikan *loss* pada parameter 0.0100 sehingga mengakibatkan parameter tersebut memberikan *loss* yang lebih tinggi dibanding parameter lain

3.3 Perbandingan dengan Metode Lain

Performa dari model komputasi yang dibuat dibandingkan dengan beberapa penelitian dengan kasus EEG emosi.

Tabel 4. Perbandingan dengan Metode Lain

Metode	Akurasi (%)
2D CNN (Haqqe, dkk, 2021)	83,44
3D CNN (metode yang diusulkan)	65.45

Penelitian (**Haqqe, dkk, 2021**) menggunakan 2D CNN untuk melakukan identifikasi emosi sehingga menghasilkan akurasi sebesar 83,44%, menggunakan set data *SEED* dan teknik koreksi bobot *AdaMax*. Metode yang diusulkan (STFT + 3D CNN) tidak memberikan akurasi yang lebih baik daripada (Wavelet + 2D CNN) yang bisa diakibatkan karena kemungkinan adanya pengaruh non stasioner dari sinyal EEG yang berarti karakteristiknya frekuensi berubah sepanjang waktu. Padahal STFT hanya sesuai untuk sinyal stasioner. atau memiliki resolusi waktu-frekuensi yang tetap. Sementara penelitian terdahulu (**Haqqe, dkk, 2021**) menggunakan Wavelet. Apabila representasi 3D CNN tidak memungkinkan menggunakan Wavelet yang domain waktu. Selain itu, representasinya dapat lebih rentan terhadap *noise*

karena mencakup informasi dari tiga dimensi. *Noise* ikut berkontribusi dan mempengaruhi kinerja model. Namun demikian memberikan respon yang relative cepat kurang dari 0,1 detik.

4. KESIMPULAN

Identifikasi emosi menggunakan sinyal EEG dapat dilakukan dengan melakukan identifikasi terhadap variabel emosi. Penelitian ini menggunakan 3D CNN sebagai metode untuk melakukan identifikasi. Representasi tiga dimensi direpresentasikan ke dalam kanal x frekuensi x waktu karena variabel emosi berelasi dengan frekuensi dan waktu sehingga kedua domain didapatkan melalui praproses berupa STFT. Kemudian didapat hasil akurasi sebesar 65.45% dari penggunaan 3D CNN sebagai identifikasi dan STFT sebagai pra-proses. Pemilihan metode optimasi dan konfigurasi juga penting untuk mendapatkan performa terbaik. Untuk teknik koreksi bobot, AdaMax menunjukkan hasil yang lebih baik dibanding SGD. Untuk memberikan hasil akurasi yang lebih tinggi maka pemilihan ekstraksi fitur yang tepat sangat dibutuhkan mengingat bahwa sinyal EEG bersifat non-stasioner, sehingga penggunaan ekstraksi fitur yang bersifat stasioner kurang cocok.

DAFTAR RUJUKAN

- Deshpande, M., & Rao, V. (2018). Depression detection using emotion artificial intelligence. *International Conference on Intelligent Sustainable Systems, ICISS 2017*, (pp. 858–862).
- Gao, Q., Wang, C. han, Wang, Z., Song, X. lin, Dong, E. zeng, & Song, Y. (2020). EEG based emotion recognition using fusion feature extraction method. *Multimedia Tools and Applications*, 79(37–38), 27057–27074.
- Haqqe, R. H. D., Djamal, E. C., & Wulandari, A. (2021). Emotion Recognition of EEG Signals Using Wavelet Filter and Convolutional Neural Networks. *International Conference on Advanced Informatics: Concepts, Theory, and Application, ICAICTA 2021*.
- Houssein, E. H., Hammad, A., & Ali, A. A. (2022). Human emotion recognition from EEG-based brain–computer interface using machine learning: a comprehensive review. In *Neural Computing and Applications* (Vol. 34).
- Jang, H., & Tong, F. (2021). Convolutional neural networks trained with a developmental sequence of blurry to clear images reveal core differences between face and object processing. *Journal of Vision*, 21(12).
- Indrawan, R., Djamal, E. C., & Ilyas, R. (2017). Identifikasi Neuropsikologis Terhadap Video Iklan Secara Real-Time Menggunakan Fast Fourier Transform dan Support Vector Machine. *Seminar Nasional Aplikasi Teknologi Informasi (SNATI)*, (pp. 6–10).
- Kumar, D. (2022). Emotion Recognition from DEAP Dataset Using SVM Classifier. *International Journal of Scientific Research in Engineering and Management*, 6(6), 1–18.
- Lan, Z., Sourina, O., Wang, L., Scherer, R., & Muller-Putz, G. R. (2019). Domain Adaptation

- Techniques for EEG-Based Emotion Recognition: A Comparative Study on Two Public Datasets. *IEEE Transactions on Cognitive and Developmental Systems*, 11(1), 85–94.
- Negi, T., & . S. (2021). Analysis and Processing of EEG Signal: A Review. *SSRN Electronic Journal*, (Icsmdi).
- Nematollahi, J., & Firoozabadi, M. (2018). Recognition of Positive, Negative and Neutral Emotions Using Brain Connectivity Patterns. *2017 24th Iranian Conference on Biomedical Engineering and 2017 2nd International Iranian Conference on Biomedical Engineering, ICBME 2017*, (December), (pp. 330–333).
- Salama, E. S., El-Khoribi, R. A., Shoman, M. E., & Wahby Shalaby, M. A. (2018). EEG-based emotion recognition using 3D convolutional neural networks. *International Journal of Advanced Computer Science and Applications*, 9(8), 329–337.
- Wang, Z. M., Hu, S. Y., & Song, H. (2019). Channel Selection Method for EEG Emotion Recognition Using Normalized Mutual Information. *IEEE Access*, 7, 143303–143311.
- Yosi, A. N. N. M., Sidek, K. A., Yaacob, H. S., Othman, M., & Jusoh, A. Z. (2019). Emotion recognition using electroencephalogram signal. *Indonesian Journal of Electrical Engineering and Computer Science*, 15(2), 786–793.
- Yudhana, A., Muslim, A., Wati, D. E., Puspitasari, I., Azhari, A., & Mardhia, M. M. (2020). Human emotion recognition based on EEG signal using fast fourier transform and K-Nearest neighbor. *Advances in Science, Technology and Engineering Systems*, 5(6), 1082–1088.
- Zamani, F., & Wulansari, R. (2021). *Emotion Classification using 1D-CNN and RNN based On DEAP Dataset*. 363–378.
- Zhang, R., Zeng, Y., Tong, L., Shu, J., Lu, R., Li, Z., ... Yan, B. (2022). EEG Identity Authentication in Multi-Domain Features: A Multi-Scale 3D-CNN Approach. *Frontiers in Neurorobotics*, 16(June), 1–16.
- Zheng, W. L., & Lu, B. L. (2015). Investigating Critical Frequency Bands and Channels for EEG-Based Emotion Recognition with Deep Neural Networks. *IEEE Transactions on Autonomous Mental Development*, 7(3), 162–175.
- Zheng, X., Yu, X., Yin, Y., Li, T., & Yan, X. (2021). Three-dimensional feature maps and convolutional neural network-based emotion recognition. *International Journal of Intelligent Systems*, 36(11), 6312–6336.