

Klasifikasi Eksekusi Gerak Tangan dan Kaki Menggunakan Arsitektur EEGNet

Panji Iman Sujatmiko (220011), Darren Christian Liharja (220043)

Pendahuluan

Kemajuan teknologi interaksi manusia dan mesin (Human-Machine Interaction) kini memungkinkan sinyal otak diterjemahkan langsung menjadi perintah digital melalui Brain-Computer Interface (BCI). Teknologi ini memberi harapan bagi pasien dengan gangguan motorik berat, seperti pasca-stroke atau ALS, untuk kembali berkomunikasi atau mengendalikan alat bantu tanpa perlu menggerakkan otot fisik (Wolpaw et al., 2002; Aggarwal & Chugh, 2019).

Alat yang paling umum digunakan untuk BCI adalah Electroencephalography (EEG). Selain murah dan portabel, EEG mampu merekam aktivitas otak dengan cepat (real-time) yang penting untuk kendali langsung (Rashid et al., 2020). Salah satu fungsi utamanya adalah mendeteksi pola sinyal saat seseorang melakukan atau membayangkan gerakan (Motor Execution/Imagery), seperti niat menggerakkan tangan atau kaki.

Namun, mengolah data EEG memiliki tantangan besar karena sifat sinyalnya yang non-stasioner, memiliki rasio signal-to-noise yang rendah, dan rentan terhadap artefak fisiologis (Altaheri et al., 2023). Metode konvensional seperti Common Spatial Pattern (CSP) yang dikombinasikan dengan Support Vector Machine (SVM) sering kali memiliki keterbatasan. Metode ini sangat bergantung pada ekstraksi fitur manual (hand-crafted features) dan pemilihan pita frekuensi yang kaku, sehingga sulit beradaptasi dengan variabilitas sinyal antar-individu yang unik (Lotte et al., 2018).

Untuk mengatasi keterbatasan ekstraksi fitur manual, pendekatan deep learning seperti Convolutional Neural Network (CNN) mulai banyak diterapkan. Meskipun demikian, penggunaan arsitektur CNN standar (yang biasa digunakan untuk pengolahan citra) pada data EEG sering kali tidak efisien. Model-model ini cenderung memiliki jumlah parameter yang sangat besar

(computationally expensive), yang menyulitkan penerapan pada perangkat BCI portabel. Selain itu, kompleksitas model yang tinggi berisiko menyebabkan overfitting, terutama mengingat ketersediaan dataset EEG yang umumnya terbatas dibandingkan dataset citra (Schirrmester et al., 2017).

Oleh karena itu, diperlukan arsitektur jaringan saraf yang dirancang khusus (domain-specific) untuk menangani karakteristik unik sinyal EEG secara efisien, yaitu EEGNet. Diperkenalkan oleh Lawhern et al. (2018), EEGNet menggunakan mekanisme Depthwise Separable Convolutions yang memungkinkannya mengekstraksi fitur temporal (frekuensi) dan spasial (lokasi elektroda) secara simultan dengan jumlah parameter yang jauh lebih sedikit dibandingkan CNN konvensional. Pendekatan ini memungkinkan model untuk mempelajari filter frekuensi layaknya band-pass filter dan pola topografi otak secara otomatis dalam satu arsitektur yang compact dan ringan.

Berdasarkan keunggulan tersebut, penelitian ini menerapkan arsitektur EEGNet untuk mengklasifikasikan gerak tangan dan kaki. Penelitian ini bertujuan untuk membuktikan bahwa arsitektur yang ringkas namun terspesialisasi ini mampu menghasilkan performa klasifikasi yang tinggi dan generalisasi yang baik, serta lebih efisien untuk diimplementasikan dalam sistem BCI nyata dibandingkan model deep learning yang kompleks.

Tinjauan Pustaka

2.1. Landasan Teori

Brain-Computer Interface (BCI)

Brain-Computer Interface (BCI) adalah sebuah sistem teknologi yang menciptakan jalur komunikasi baru antara otak manusia dengan perangkat eksternal, seperti komputer atau mesin (Wolpaw et al., 2002). Sederhananya, sistem BCI memungkinkan seseorang untuk mengendalikan perangkat hanya dengan

menggunakan sinyal dari aktivitas otaknya, tanpa melalui jalur otot dan saraf normal.

BCI umumnya dibagi menjadi dua kategori: invasif (memerlukan operasi bedah untuk menanam elektroda di dalam otak) dan non-invasif (menggunakan sensor yang ditempatkan di permukaan kulit kepala). Penelitian ini berfokus pada BCI non-invasif, yang memiliki potensi besar untuk aplikasi di dunia nyata, terutama di bidang medis untuk rehabilitasi motorik pasien stroke atau sebagai alat bantu komunikasi bagi penderita kelumpuhan (Nicolas-Alonso & Gomez-Gil, 2012).

Electroencephalography (EEG)

Untuk sistem BCI non-invasif, Electroencephalography (EEG) adalah metode yang paling umum digunakan untuk merekam aktivitas otak. EEG bekerja dengan cara mengukur sinyal listrik sangat kecil yang dihasilkan oleh aktivitas jutaan sel saraf (neuron) di otak melalui elektroda yang ditempatkan di kulit kepala (Nicolas-Alonso & Gomez-Gil, 2012).

Popularitas EEG didasari oleh beberapa keunggulan praktis. Pertama, EEG bersifat non-invasif (tidak perlu bedah). Kedua, biaya perangkatnya relatif lebih murah dibandingkan metode lain seperti fMRI atau MEG. Ketiga, dan yang paling penting untuk analisis gerak, EEG memiliki resolusi temporal yang tinggi, yang berarti alat ini mampu menangkap perubahan aktivitas otak yang terjadi dengan sangat cepat, dalam hitungan milidetik (Nicolas-Alonso & Gomez-Gil, 2012).

Paradigma Motor Execution (ME)

Untuk mengendalikan BCI, pengguna harus menghasilkan pola sinyal otak yang dapat dikenali oleh sistem. Salah satu pola yang paling sering diteliti adalah yang terkait dengan aktivitas motorik (pergerakan).

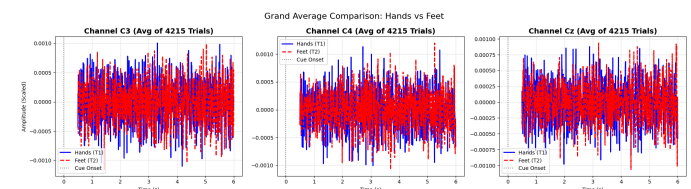
- Motor Execution (ME): Ini adalah fenomena yang terjadi saat seseorang benar-benar melakukan sebuah gerakan fisik, misalnya menggerakkan tangan kanan atau kaki kiri.
- Motor Imagery (MI): Ini adalah fenomena yang terjadi saat seseorang hanya membayangkan melakukan gerakan tersebut, tanpa benar-benar bergerak.

Saat seseorang melakukan (ME) atau membayangkan (MI) sebuah gerakan, terjadi perubahan pola sinyal EEG yang khas di area korteks motorik otak. Fenomena ini dikenal sebagai Event-Related Desynchronization (ERD) dan Event-Related Synchronization (ERS). Sinyal inilah yang menjadi "target" yang akan dipelajari dan diklasifikasikan oleh sistem BCI (Pfurtscheller & Lopes da Silva, 1999).

Distribusi spasial dari fenomena ini sangat spesifik bergantung pada anggota tubuh yang digerakkan, sesuai dengan pemetaan fungsi otak (Motor Homunculus). Berdasarkan standar penempatan elektroda internasional 10-20, aktivitas motorik tangan dan kaki terpusat pada area sensorimotor (Cortex):

1. Gerakan Tangan: Bersifat kontralateral (berlawanan arah). Gerakan tangan kanan akan memicu aktivitas ERD dominan di belahan otak kiri (elektroda C3), sedangkan tangan kiri di belahan otak kanan (elektroda C4).
2. Gerakan Kaki: Area kaki terletak di bagian medial (tengah) korteks motorik, sehingga aktivitas sinyalnya terpusat di garis tengah kepala (elektroda Cz).

Pola spasial (C3 vs Cz vs C4) menjadi fitur kunci bagi algoritma untuk membedakan antara perintah gerakan tangan dan kaki.



Gambar 2.1. Perbandingan Sinyal Tangan dan Kaki pada Channel C3, C4 dan Cz

2.2. Tantangan dan Pendekatan Klasik

Tantangan dalam Pemrosesan Sinyal EEG

Meskipun memiliki banyak kelebihan, sinyal EEG sangat sulit untuk dianalisis. Masalah utamanya adalah rasio signal-to-noise (SNR) yang rendah. Hal ini berarti sinyal otak yang ingin kita ukur sering kali "tenggelam" dalam berbagai gangguan (noise).

Gangguan ini bisa berasal dari:

- Artefak Biologis: Seperti kedipan mata, gerakan otot di dahi atau rahang, dan bahkan detak jantung.
- Noise Eksternal: Seperti interferensi listrik dari lampu atau perangkat elektronik lainnya.

Selain itu, sinyal EEG bersifat non-stasioner (karakteristik sinyalnya terus berubah-ubah) dan sangat bervariasi antar individu, bahkan pada individu yang sama di hari yang berbeda (Lotte et al., 2018). Tantangan inilah yang membuat proses klasifikasi sinyal EEG menjadi rumit.

Keterbatasan Pendekatan Konvensional

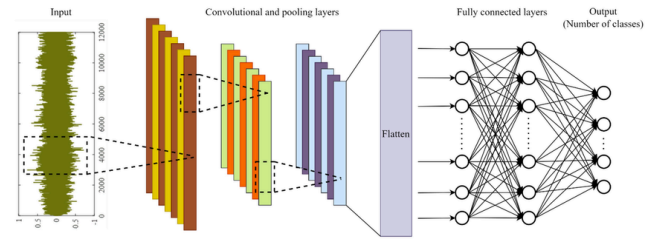
Masalah terbesar dari pendekatan klasik adalah ketergantungannya pada rekayasa fitur (feature engineering) manual. Kinerja metode CSP sangat bergantung pada keahlian peneliti dalam menentukan parameter yang tepat, seperti memilih pita frekuensi (frequency band) yang benar. Jika salah pilih, kinerjanya akan buruk.

Metode ini juga sering kesulitan menangkap pola-pola yang bersifat kompleks dan non-linear dalam data EEG. Keterbatasan inilah yang mendorong para peneliti untuk mencari pendekatan baru yang tidak lagi bergantung pada ekstraksi fitur manual (Lotte et al., 2018).

2.3. Deep Learning untuk Analisis EEG

Convolutional Neural Networks untuk Data EEG

Convolutional Neural Networks (CNN) adalah arsitektur deep learning yang awalnya sangat sukses untuk tugas pengenalan gambar. Ternyata, prinsip kerja CNN juga dapat digunakan untuk menganalisis data EEG. Data sinyal EEG dari banyak elektroda dapat direpresentasikan sebagai sebuah "gambar" 2D, di mana satu sumbu adalah waktu (temporal) dan sumbu lainnya adalah elektroda (spasial).



Gambar 2.2. Arsitektur CNN
Sumber (10.1007/s42107-023-00816-w)

CNN dapat mempelajari dua jenis fitur ini secara bersamaan:

- Fitur Temporal: Dengan lapisan konvolusi yang bergerak di sepanjang sumbu waktu, CNN bisa belajar mengenali bentuk-bentuk pola sinyal yang khas (misalnya, bentuk gelombang ERD).
- Fitur Spasial: Dengan lapisan konvolusi yang bergerak di sepanjang sumbu elektroda, CNN bisa belajar mengenali pola hubungan antar elektroda (misalnya, saat elektroda C3 aktif, elektroda C4 tidak aktif).

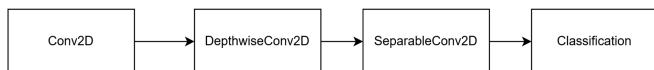
Arsitektur CNN untuk EEG

Banyak penelitian telah membuktikan keunggulan CNN untuk klasifikasi EEG. Salah satu penelitian paling penting dilakukan oleh Schirmer et al. (2017). Mereka mengusulkan dua arsitektur, DeepConvNet dan ShallowConvNet, dan membuktikan bahwa kedua arsitektur tersebut mampu mengalahkan metode klasik (FBCSP) dengan belajar langsung dari data EEG mentah.

Penelitian penting lainnya adalah EEGNet (Lawhern et al., 2018). Ini adalah arsitektur CNN yang dirancang khusus agar sangat ringkas dan efisien untuk data EEG. EEGNet menggunakan jenis konvolusi khusus (depthwise dan separable convolutions) untuk menangani data spatio-temporal EEG secara efektif dan telah menjadi baseline yang sangat populer di banyak penelitian BCI. Berbagai penelitian lain juga telah menunjukkan bahwa pendekatan Deep Learning secara konsisten memberikan kinerja yang lebih baik atau sebanding dengan metode klasik, tanpa memerlukan proses ekstraksi fitur manual yang rumit.

Arsitektur EEGNet

EEGNet adalah arsitektur Convolutional Neural Network (CNN) yang dirancang secara khusus (compact) untuk menangani sinyal EEG dalam antarmuka otak-komputer (BCI). Diperkenalkan oleh Lawhern et al. (2018), arsitektur ini bertujuan untuk mengatasi tantangan utama dalam penerapan Deep Learning pada data EEG, yaitu ketersediaan data yang terbatas dan risiko overfitting yang tinggi pada model CNN konvensional yang memiliki parameter terlalu besar.



Gambar 2.3. Arsitektur EEGNet

Keunggulan utama EEGNet terletak pada penggunaan mekanisme Depthwise Separable Convolutions, yang memungkinkan model mengekstraksi fitur penting secara efisien melalui tahapan berikut:

1. Temporal Convolution (Block 1): Lapisan pertama menggunakan filter konvolusi 1D yang bergerak sepanjang sumbu waktu. Fungsinya mirip dengan band-pass filter elektronik, yaitu mempelajari dan memisahkan frekuensi sinyal yang relevan (seperti pita Mu dan Beta pada kasus motorik) dari noise frekuensi tinggi atau rendah.
2. Depthwise Spatial Convolution (Block 1): Berbeda dengan CNN biasa yang mencampur semua channel sekaligus, lapisan ini mempelajari filter spasial untuk setiap peta fitur secara terpisah. Lapisan ini bertugas mengenali pola topografi otak, seperti perbedaan aktivitas antara elektroda C3 (kiri), Cz (tengah), dan C4 (kanan) yang krusial untuk membedakan gerak tangan dan kaki. Teknik ini secara signifikan mengurangi jumlah parameter yang harus dilatih.
3. Separable Convolution (Block 2): Lapisan ini menggabungkan fitur-fitur yang telah dipelajari untuk membentuk representasi akhir yang lebih ringkas sebelum masuk ke tahap klasifikasi. Penggunaan Separable Convolution membuat EEGNet jauh lebih ringan dibandingkan

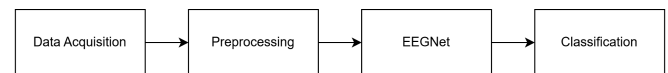
arsitektur seperti DeepConvNet, namun tetap mampu menghasilkan akurasi yang tinggi (Lawhern et al., 2018).

Penelitian menunjukkan bahwa EEGNet mampu menggeneralisasi data dengan sangat baik lintas-subjek dan lintas-paradigma (seperti Motor Imagery atau P300) dengan kebutuhan komputasi yang minimal.

Metodologi

Research Framework

Penelitian ini mengusulkan pendekatan Deep Learning berbasis arsitektur EEGNet untuk mengklasifikasikan sinyal EEG Motor Execution. Sistem dirancang secara end-to-end untuk meminimalkan rekayasa fitur manual, terdiri dari tiga tahapan utama: (1) Akuisisi & Seleksi Data, (2) Pra-pemrosesan Sinyal, dan (3) Klasifikasi menggunakan EEGNet.



Gambar 3.1. Sistem Klasifikasi Sinyal EEG menggunakan EEGNet

Tahap awal penelitian dimulai dengan proses akuisisi data (Data Acquisition) yang memuat dataset sinyal EEG mentah dalam format standar .edf (European Data Format). Sinyal ini bersumber dari dataset publik PhysioNet EEG Motor Movement/Imagery Dataset yang berisi rekaman dari subjek saat melakukan instruksi gerakan fisik (Motor Execution). Menggunakan pustaka MNE-Python, struktur data beserta penanda waktu stimulus (event markers) diekstrak. Pada tahap ini, dilakukan seleksi channel secara spesifik pada area korteks sensorimotor, yaitu C3, Cz, dan C4. Pemilihan ini bertujuan untuk memfokuskan analisis pada area otak yang merepresentasikan aktivitas motorik tangan dan kaki, serta mengurangi dimensi data yang tidak relevan.

Karena sinyal EEG mentah memiliki amplitudo rendah dan rentan terhadap noise, tahap pra-pemrosesan (Signal Preprocessing) menjadi langkah krusial. Proses ini diawali dengan penyaringan frekuensi (Band-pass Filtering) pada rentang 7–60 Hz. Rentang ini dipilih untuk mempertahankan spektrum gelombang yang luas, mencakup gelombang Mu (8–13 Hz) dan Beta (13–30 Hz) secara utuh, serta mengakomodasi komponen

frekuensi Low-Gamma (>30 Hz) yang dapat memuat informasi kognitif tambahan. Batas bawah 7 Hz efektif membuang artefak EOG berfrekuensi rendah, sementara batas atas 60 Hz menjaga integritas sinyal tanpa memotong informasi di ujung spektrum Beta secara agresif. Selanjutnya, dilakukan segmentasi data (Epoching) di mana sinyal kontinu dipotong menjadi segmen berdurasi tetap berdasarkan event markers. Setiap segmen kemudian dinormalisasi menggunakan teknik Z-score Scaling agar distribusi data seragam (rata-rata 0, standar deviasi 1), yang berfungsi mempercepat konvergensi model saat pelatihan.

Setelah data bersih dan terstandarisasi, proses dilanjutkan ke tahap pembangunan model menggunakan arsitektur EEGNet (Lawhern et al., 2018). Data pertama kali masuk ke blok Temporal Convolution, di mana filter konvolusi 1D bekerja mempelajari frekuensi sinyal secara otomatis layaknya band-pass filter digital. Output dari blok ini diteruskan ke lapisan Depthwise Spatial Convolution. Berbeda dengan CNN standar, lapisan ini tidak mencampur seluruh channel, melainkan mempelajari filter spasial untuk setiap peta fitur secara terpisah. Hal ini memungkinkan model untuk menangkap pola topografi otak yang spesifik, yaitu perbedaan distribusi energi antara elektroda C3 (kiri), Cz (tengah), dan C4 (kanan). Fitur-fitur tersebut kemudian dipadatkan melalui Separable Convolution untuk mereduksi parameter komputasi secara efisien. Pada tahap akhir (Classification), seluruh informasi spatio-temporal yang telah diekstrak dikumpulkan pada Dense Layer dan diproses menggunakan fungsi aktivasi Sigmoid. Proses ini menghasilkan skor probabilitas tunggal dalam rentang 0 hingga 1. Prediksi kelas target (gerakan kaki atau tangan) ditentukan berdasarkan nilai ambang batas (threshold) 0.5, di mana nilai di atas ambang batas diklasifikasikan sebagai kelas positif, dan sebaliknya.

Dataset dan Sumber Data

Data yang digunakan dalam penelitian ini bersumber dari repositori publik PhysioNet, secara spesifik menggunakan [EEG Motor Movement/Imagery Dataset \(eegmmidb\)](#) yang dikembangkan oleh Schalk et al. (2004). Dataset ini merupakan salah satu standar tolok ukur (benchmark) yang paling banyak digunakan dalam penelitian BCI karena volume datanya yang besar dan protokol eksperimen yang terstandarisasi.

Dataset ini merekam aktivitas gelombang otak dari 109 sukarelawan sehat. Perekaman dilakukan menggunakan sistem BCI2000 dengan spesifikasi teknis sebagai berikut:

- Jumlah Elektroda (Channel): 64 elektroda yang dipasang pada kulit kepala mengikuti standar internasional sistem 10-20.
- Frekuensi Recording(f_s): Sinyal direkam pada frekuensi asli 160 Hz. Resolusi ini dianggap cukup memadai untuk menangkap gelombang Mu (8-13 Hz) dan Beta (13-30 Hz) tanpa membebani komputasi secara berlebihan.
- Durasi Perekaman: Setiap subjek melakukan 14 sesi (runs) eksperimen, di mana setiap run berlangsung selama satu atau dua menit.

Protokol eksperimen dirancang untuk menangkap respon motorik yang berbeda. Subjek duduk di depan layar komputer dan diminta untuk melakukan gerakan fisik (Motor Execution) atau membayangkan gerakan (Motor Imagery) sesuai dengan target visual yang muncul di tepi layar (kiri, kanan, atas, atau bawah).

Karena fokus penelitian ini adalah Motor Execution (ME), maka hanya data dari sesi eksekusi gerak yang digunakan, sementara sesi Motor Imagery diabaikan. Sesi yang dipilih adalah sebagai berikut:

- Gerak Kedua Tangan vs Kedua Kaki: Subjek diminta membuka dan menutup kedua kepalan tangan atau menggerakkan kedua kaki secara nyata (Run 5, 9, 13).
- Subjek yang digunakan adalah subjek ke 1 sampai 100

Preprocessing

Tahap pra-pemrosesan bertujuan untuk meningkatkan rasio sinyal-terhadap-derau (Signal-to-Noise Ratio/SNR) dan menyusun data ke dalam format yang optimal bagi model Deep Learning. Mengingat sinyal EEG bersifat non-stasioner dan rentan terhadap artefak, diterapkan serangkaian teknik pembersihan sinyal menggunakan pustaka MNE-Python.

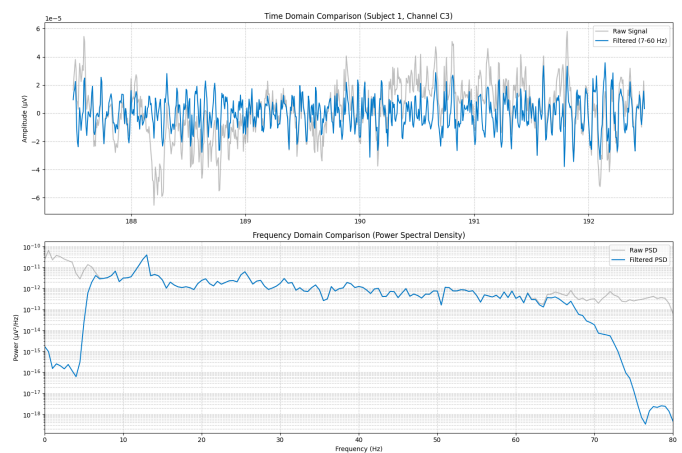
Langkah pertama adalah penerapan filter Finite Impulse Response (FIR) dengan rentang frekuensi 7.0 Hz hingga 60.0 Hz. Ambang batas bawah (lower cut-off) sebesar 7.0 Hz dipilih secara spesifik untuk membuang artefak frekuensi rendah yang dominan, seperti pergeseran lambat (slow drifts) dan sinyal Electrooculography (EOG) akibat kedipan mata. Sementara itu, ambang batas atas (upper cut-off) 60.0 Hz ditetapkan untuk mempertahankan integritas spektrum gelombang Mu (8–13 Hz) dan Beta (13–30 Hz) secara utuh, serta mengakomodasi komponen frekuensi Low-Gamma yang dapat memuat informasi kognitif tambahan, sembari membatasi noise frekuensi tinggi yang tidak relevan.

Sinyal kontinyu kemudian disegmentasi menjadi epochs berdasarkan penanda waktu stimulus (event markers). Jendela waktu pengambilan data ditetapkan mulai dari $t = 0.5$ detik hingga $t = 6.0$ detik relatif terhadap onset stimulus (munculnya instruksi di layar).

- Waktu Awal ($t_{min} = 0.5$): Penundaan pengambilan awal sebesar 0.5 detik bertujuan untuk menghindari kontaminasi sinyal oleh Visual Evoked Potential (VEP) yang muncul di korteks oksipital saat subjek melihat visual instruksi, serta mengakomodasi jeda latensi reaksi (reaction time latency) sebelum gerakan fisik dimulai.
- Waktu Akhir ($t_{max} = 6.0$): Durasi hingga detik ke-6 dipilih untuk menangkap fase eksekusi gerakan secara utuh serta fase penyelesaian (de-sinkronisasi dan sinkronisasi ulang), tanpa melebihi batas durasi antar-trial yang tersedia dalam dataset.

Selain untuk isolasi kejadian, strategi segmentasi ini juga berfungsi sebagai metode persiapan data latih. Label kategori dari dataset asli dikonversi menjadi format biner, di mana anotasi gerakan tangan dipetakan menjadi kelas 0 dan gerakan kaki menjadi kelas 1.

Terakhir, dilakukan teknik Penskalaan Amplitudo (Amplitude Scaling). Karena data mentah EEG direkam dalam satuan Volt yang memiliki nilai numerik sangat kecil (orde 10^{-6}), setiap segmen dikalikan dengan faktor 1.000.000 (atau dikonversi sesuai kebutuhan library) untuk mengubah satuan menjadi Mikrovolt (μV). Langkah ini bertujuan untuk menstabilkan rentang numerik input agar lebih mudah diproses oleh model deep learning, tanpa mengubah bentuk distribusi statistik asli dari sinyal tersebut.



Gambar 3.2. Hasil Preprocessing Sinyal EEG pada Channel C3

Perancangan Arsitektur EEGNet

Penelitian ini menerapkan arsitektur EEGNet, sebuah model Deep Learning yang dirancang kompak dan efisien untuk pemrosesan sinyal EEG. Berbeda dengan pendekatan hibrida konvensional, EEGNet memanfaatkan operasi konvolusi khusus, yaitu Depthwise dan Separable Convolutions untuk mengekstrak fitur temporal dan spasial secara terpisah namun terintegrasi. Struktur input data diatur dalam dimensi (Channels, Samples, 1), yang memungkinkan penggunaan kernel konvolusi 2D. Arsitektur model terdiri dari tiga blok utama: blok konvolusi temporal-spasial, blok konvolusi terpisah (separable), dan blok klasifikasi.

- Blok Ekstraksi Fitur Temporal dan Spasial

Blok pertama berfungsi sebagai extractor fitur dasar yang dirancang untuk meniru mekanisme pemrosesan sinyal EEG konvensional, yaitu band-pass filtering dan spatial filtering (seperti Common Spatial Pattern atau CSP). Proses ini diawali dengan tahap Temporal Convolution menggunakan lapisan Conv2D dengan kernel berukuran (1, kernLength) yang bergerak khusus di sepanjang sumbu waktu. Konfigurasi ini memungkinkan model mempelajari filter frekuensi, misalnya pada rentang 8–30 Hz secara otomatis dari data mentah tanpa mencampur informasi antar-elektroda, yang kemudian distabilkan menggunakan Batch Normalization. Output dari tahap ini selanjutnya diproses oleh Spatial Depthwise Convolution menggunakan kernel berukuran (Chans, 1) dengan parameter $\text{depth_multiplier}=D$. Lapisan ini bekerja secara unik dengan menerapkan filter spasial pada setiap peta fitur temporal secara independen, sehingga mampu menangkap pola spasial atau lokasi aktivitas otak tanpa mendistorsi informasi waktu. Rangkaian proses pada blok ini diakhiri dengan penerapan kembali Batch Normalization dan fungsi aktivasi ELU (Exponential Linear Unit) untuk mempercepat konvergensi serta mencegah masalah dying ReLU, sebelum akhirnya dimensi data direduksi dan dihaluskan melalui Average Pooling 2D berukuran (1, 4) yang dilengkapi dengan teknik Dropout untuk mencegah terjadinya overfitting.

- **Separable Convolution**

Blok kedua bertujuan untuk meringkas fitur yang telah diekstrak secara efisien. Blok ini menggunakan SeparableConv2D, sebuah teknik yang memisahkan konvolusi menjadi dua bagian: konvolusi depthwise diikuti oleh konvolusi pointwise (1x1). Penggunaan Separable Convolution secara signifikan mengurangi jumlah parameter yang harus dilatih dibandingkan konvolusi standar, menjadikan model lebih ringan dan cepat.

Sama seperti blok sebelumnya, output dari lapisan ini diproses melalui Batch

Normalization dan aktivasi ELU. Selanjutnya, dilakukan Average Pooling 2D dengan ukuran kernel (1, 8) untuk mereduksi dimensi temporal secara lebih agresif sebelum data masuk ke tahap klasifikasi. Lapisan Dropout kembali diterapkan untuk meningkatkan kemampuan generalisasi model.

- **Classification Block**

Tahap akhir bertujuan memetakan fitur abstrak menjadi keputusan probabilitas biner. Fitur 2D dari blok sebelumnya didatarkan menjadi vektor satu dimensi menggunakan lapisan Flatten. Vektor ini kemudian masuk ke lapisan Dense (Fully Connected Layer) yang dikonfigurasi dengan 1 unit neuron dan kendala bobot (max norm constraint).

Untuk menghasilkan prediksi akhir, digunakan Fungsi Aktivasi Sigmoid. Fungsi ini memetakan nilai output jaringan ke dalam rentang probabilitas 0 hingga 1. Keputusan kelas didasarkan pada nilai ambang batas (threshold) 0.5:

1. Jika nilai output >0.5 , data diklasifikasikan sebagai Gerak Kaki (Kelas 1)
2. Jika nilai output ≤ 0.5 , data diklasifikasikan sebagai Gerak Tangan (Kelas 0).

Training Scheme

Skema pelatihan dirancang untuk mengoptimalkan parameter bobot pada jaringan saraf guna meminimalkan kesalahan prediksi. Strategi ini mencakup pembagian dataset yang ketat untuk validasi dan konfigurasi hiperparameter yang disesuaikan dengan karakteristik binary classification.

Pembagian Dataset (Data Splitting) Untuk menjamin kemampuan generalisasi model, dataset yang terdiri dari total 4.215 sampel dibagi menggunakan teknik stratified random splitting. Teknik ini dipilih untuk menjaga keseimbangan proporsi kelas (gerak tangan vs. kaki) di setiap sub-himpunan data. Mekanisme pembagian dilakukan dalam dua tahap:

1. Tahap Pertama (Hold-out): Mengisolasi 20% dari total data sebagai Data Uji (Testing Set). Data ini bersifat unseen dan hanya digunakan untuk evaluasi final.
2. Tahap Kedua: Sisa data (80%) dibagi kembali dengan rasio 80:20 menjadi Data Latih (Training Set) dan Data Validasi (Validation Set).

Dengan skema tersebut, distribusi data yang digunakan adalah sebagai berikut:

- Data Latih (Training Set): 2697 sampel. Digunakan untuk pembaruan bobot via backpropagation.
- Data Validasi (Validation Set): 675 sampel. Digunakan untuk memantau performa per-epoch dan mencegah overfitting.
- Data Uji (Testing Set): 843 sampel. Digunakan untuk pengujian performa akhir.

Konfigurasi Hiperparameter Eksekusi pelatihan menggunakan algoritma pengoptimal Adam (Adaptive Moment Estimation). Algoritma ini dipilih karena efisiensinya dalam mengadaptasi learning rate secara dinamis untuk setiap parameter, memberikan konvergensi yang lebih cepat dan stabil dibandingkan Stochastic Gradient Descent (SGD) konvensional.

Mengingat arsitektur model menggunakan neuron output tunggal dengan aktivasi Sigmoid, fungsi kerugian (Loss Function) yang diterapkan adalah Binary Cross-entropy. Fungsi ini mengukur divergensi antara probabilitas prediksi model dengan label target biner (0 atau 1). Parameter ukuran batch (batch size) ditetapkan sebesar 32 untuk menyeimbangkan efisiensi memori GPU dengan stabilitas estimasi gradien. Pelatihan dijalankan selama sejumlah iterasi (epochs) tertentu dengan mekanisme Early Stopping atau pemantauan grafik loss validasi untuk memastikan model berhenti pada titik konvergensi optimal.

Hasil dan Analisis

Implementasi dan Environment

Seluruh eksperimen dalam penelitian ini diimplementasikan menggunakan bahasa pemrograman Python 3.10 di lingkungan Google Colab. Proses komputasi dipercepat menggunakan dukungan GPU

(Graphics Processing Unit) T4 yang disediakan oleh platform tersebut untuk efisiensi pelatihan model Deep Learning.

Adapun pustaka (library) utama yang digunakan meliputi:

- MNE-Python (v1.6): Untuk pemrosesan sinyal EEG (filtering, epoching).
- TensorFlow/Keras (v2.15): Untuk pembangunan dan pelatihan arsitektur EEGNet.
- Scikit-learn: Untuk evaluasi metrik kinerja (Confusion Matrix, Classification Report).
- Seaborn & Matplotlib: Untuk visualisasi data dan grafik hasil.

Parameter pelatihan final yang menghasilkan kinerja terbaik dirangkum pada Tabel 4.1 berikut:

Tabel 4.1 Konfigurasi Hiperparameter

Parameter	Nilai / Konfigurasi
Optimizer	Adam (learning rate adaptif)
Loss Function	Binary Cross-entropy
Batch Size	32
Epochs	100 (dengan EarlyStopping)
Dropout Rate	0.5 (Spatial & Temporal Dropout)
Activation (Output)	Sigmoid

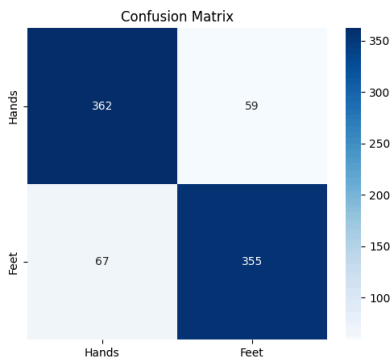
Evaluasi Kinerja Model

Bagian ini memaparkan kinerja model EEGNet setelah diuji pada Data Uji (Testing Set) yang berjumlah 843 sampel. Evaluasi ini krusial untuk mengukur kemampuan generalisasi model terhadap data yang belum pernah dilihat sebelumnya.

Analisis Metrik Utama Ringkasan hasil klasifikasi disajikan pada Tabel 4.2. Model berhasil mencapai tingkat Akurasi Total sebesar 85.05%. Meskipun sinyal EEG memiliki rasio signal-to-noise yang rendah, model mampu mempertahankan performa yang konsisten di atas 80%.

Tabel 4.2 Confusion Matrix

	Precision	Recall	F1-Support
Class 0 (Hands)	0.84	0.86	0.85
Class 1 (Feet)	0.86	0.84	0.85
Accuracy	0.85		



Gambar 4.1 Confusion Matrix

Analisis mendalam terhadap tabel klasifikasi menunjukkan karakteristik model yang sangat positif, terutama dalam aspek keseimbangan kinerja (performance balance). Hal ini dibuktikan dengan perolehan nilai F1-Score yang identik sebesar 0.85 untuk kedua kelas, yang mengindikasikan bahwa model tidak memiliki bias terhadap salah satu kategori dan memiliki kompetensi yang setara dalam mengenali pola imajinasi gerakan tangan maupun kaki. Lebih lanjut, terlihat trade-off yang sehat antara presisi dan recall; pada kelas Hands, model menunjukkan sensitivitas yang sedikit lebih tinggi (Recall 0.86), sedangkan pada kelas Feet, model unggul dalam aspek reliabilitas prediksi (Precision 0.86). Sempitnya selisih antara kedua metrik tersebut (hanya 0.02) menegaskan bahwa konfigurasi ambang batas (threshold) keputusan sebesar 0.5 sudah berada pada titik optimal.

Secara lebih spesifik, kesalahan prediksi model yang tersisa dapat diatribusikan secara signifikan pada fenomena Variabilitas Antar-Subjek (Inter-Subject Variability). Dalam konteks neurofisiologis, pola sinyal otak setiap individu memiliki karakteristik yang unik, seperti adanya pergeseran frekuensi puncak ritme Mu (biasanya 10–12 Hz) yang berbeda antar-orang, serta variasi lokasi anatomis korteks motorik akibat perbedaan bentuk kranial. Akibatnya, model Deep Learning cenderung mempelajari fitur yang representatif untuk mayoritas subjek ("pola rata-rata"), namun mengalami kesulitan dalam mengenali pola spesifik pada subjek outlier yang memiliki karakteristik sinyal menyimpang. Ketidaksesuaian antara pola umum yang dipelajari model dengan fitur unik pada subjek-subjek tersebutlah yang pada akhirnya memicu terjadinya sisa misklasifikasi pada data uji.

Standard CNN vs. EEGNet

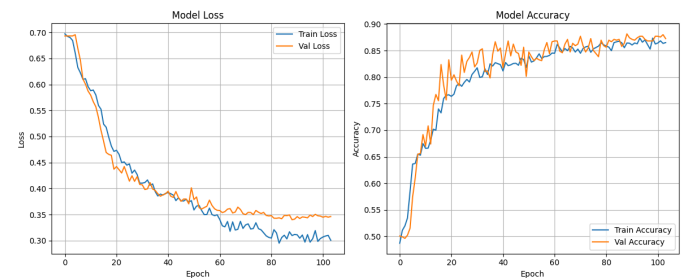
Untuk memvalidasi efektivitas arsitektur yang diusulkan, penelitian ini membandingkan kinerja EEGNet dengan model CNN Standar sebagai baseline. Perbandingan ini penting untuk melihat dampak penggunaan konvolusi khusus (Depthwise & Separable) dibandingkan konvolusi biasa.

Tabel 4.3 Confusion Matrix

Metrik Evaluasi	Standard CNN	EEGNet
Akurasi Total	75.8%	85.05%
Recall (Feet)	0.70	0.84
Recall (Tangan)	0.82	0.86
F1-Score (Average)	0.76	0.85

Training Dynamics

Kestabilan proses pembelajaran model dianalisis melalui grafik riwayat pelatihan (training history) yang membandingkan nilai Loss dan Accuracy antara data latih dan data validasi selama 100 epoch. Dinamika perubahan metrik tersebut divisualisasikan pada Gambar 4.2.



Gambar 4.2 Grafik Akurasi dan Loss saat Training

Berdasarkan visualisasi pada Gambar 4.2, dinamika proses pelatihan dapat diuraikan ke dalam dua fase utama. Pada fase awal atau pembelajaran cepat (Epoch 0–20), teramati penurunan drastis baik pada Training maupun Validation Loss dari angka 0.70 ke kisaran 0.45, yang diikuti oleh lonjakan akurasi dari 0.50 menuju 0.80; hal ini mengindikasikan kemampuan model dalam mengekstraksi fitur fundamental sinyal secara cepat di awal iterasi. Selanjutnya, memasuki fase stabilisasi (Epoch 60–100), kurva mulai mencapai titik jenuh (plateau). Meskipun terlihat divergensi minor pada grafik Loss, di mana Training Loss terus menurun mendekati 0.30 sedangkan Validation Loss tertahan stabil di kisaran 0.35, kurva validasi tidak menunjukkan

tren kenaikan kembali (U-turn), yang menandakan bahwa performa generalisasi tetap terjaga tanpa indikasi catastrophic overfitting. Fenomena menarik lainnya terlihat pada grafik Accuracy, di mana kurva validasi bergerak sangat berimpitan dan sesekali melampaui kurva pelatihan dengan nilai puncak di kisaran 87-88%. Anomali positif ini lazim terjadi akibat penerapan teknik regularisasi Dropout yang agresif; selama pelatihan, kapasitas model sengaja dibatasi dengan menonaktifkan sebagian neuron ("kesulitan belajar"), sedangkan pada fase validasi, seluruh neuron diaktifkan penuh sehingga menghasilkan kapasitas prediksi yang maksimal. Kondisi grafik yang konvergen dan stabil ini menegaskan bahwa skema pelatihan yang diterapkan telah berhasil menghasilkan model dengan status Good Fit.

Kesimpulan

Meskipun arsitektur EEGNet yang diusulkan telah menunjukkan kinerja klasifikasi yang tinggi, analisis kesalahan (error analysis) mengungkapkan bahwa model masih menghadapi tantangan dalam memprediksi sebagian kecil data secara akurat. Keterbatasan ini dapat diatribusikan secara signifikan pada fenomena Variabilitas Antar-Subjek (Inter-Subject Variability) dan BCI Illiteracy, di mana tidak semua individu memiliki kemampuan neurofisiologis untuk memodulasi pola Event-Related Desynchronization (ERD) yang distingtif dan konsisten.

Sebagai implikasi untuk pengembangan selanjutnya, penelitian mendatang disarankan untuk mengeksplorasi strategi pelatihan yang lebih adaptif, seperti penerapan teknik Transfer Learning atau Subject-Specific Fine-tuning. Pendekatan ini memungkinkan model untuk memanfaatkan pola umum yang telah dipelajari dari populasi besar (seperti model pra-latih), kemudian mengadaptasinya secara spesifik terhadap karakteristik sinyal unik setiap individu guna meminimalkan bias pada subjek outlier.

Daftar Pustaka

- Aggarwal, S., & Chugh, N. (2019). Signal processing techniques for motor imagery brain computer interface: A review. *Array*, 1–2, 100003. <https://doi.org/10.1016/j.array.2019.100003>
- Altaheri, H., Muhammad, G., & Alsulaiman, M. (2022). Physics-Informed Attention Temporal Convolutional Network for EEG-Based motor imagery classification. *IEEE Transactions on Industrial Informatics*, 19(2), 2249–2258. <https://doi.org/10.1109/tii.2022.3197419>
- Lawhern, V. J., Solon, A. J., Waytowich, N. R., Gordon, S. M., Hung, C. P., & Lance, B. J. (2016). EEGNet: a compact convolutional network for EEG-based Brain-Computer interfaces. *arXiv* (Cornell University). <https://doi.org/10.48550/arxiv.1611.08024>
- Lotte, F., Bougrain, L., Cichocki, A., Clerc, M., Congedo, M., Rakotomamonjy, A., & Yger, F. (2018). A review of classification algorithms for EEG-based brain–computer interfaces: a 10 year update. *Journal of Neural Engineering*, 15(3), 031005. <https://doi.org/10.1088/1741-2552/aab2f2>
- Nicolas-Alonso, L. F., & Gomez-Gil, J. (2012). Brain Computer Interfaces, a review. *Sensors*, 12(2), 1211–1279. <https://doi.org/10.3390/s120201211>
- Pfurtscheller, G., & Da Silva, F. L. (1999). Event-related EEG/MEG synchronization and desynchronization: basic principles. *Clinical Neurophysiology*, 110(11), 1842–1857. [https://doi.org/10.1016/s1388-2457\(99\)00141-8](https://doi.org/10.1016/s1388-2457(99)00141-8)
- Rashid, M., Sulaiman, N., Majeed, A. P. P. A., Musa, R. M., Nasir, A. F. A., Bari, B. S., & Khatun, S. (2020). Current status, challenges, and possible solutions of EEG-Based Brain-Computer Interface: A Comprehensive review. *Frontiers in Neurobotics*, 14, 25. <https://doi.org/10.3389/fnbot.2020.00025>
- Schalk, G., McFarland, D., Hinterberger, T., Birbaumer, N., & Wolpaw, J. (2004). BCI2000: A General-Purpose Brain-Computer Interface (BCI) system. *IEEE Transactions on Biomedical Engineering*, 51(6), 1034–1043. <https://doi.org/10.1109/tbme.2004.827072>
- Schirrneister, R. T., Springenberg, J. T., Fiederer, L. D. J., Glasstetter, M., Eggenberger, K., Tangermann, M., Hutter, F., Burgard, W., & Ball, T. (2017). Deep learning with convolutional neural networks for

EEG decoding and visualization. arXiv (Cornell University).
<https://doi.org/10.48550/arxiv.1703.05051>

Wolpaw, J. R., Birbaumer, N., McFarland, D. J., Pfurtscheller, G., & Vaughan, T. M. (2002). Brain–computer interfaces for communication and control. *Clinical Neurophysiology*, 113(6), 767–791.
[https://doi.org/10.1016/s1388-2457\(02\)00057-3](https://doi.org/10.1016/s1388-2457(02)00057-3)