

KLASIFIKASI EEG EPILEPSI MENGGUNAKAN SINGULAR SPECTRUM ANALYSIS, POWER SPECTRAL DENSITY DAN CONVOLUTION NEURAL NETWORK

Nurseno Bayu Aji¹⁾ dan Handayani Tjandrasa²⁾

^{1, 2)}Jurusan Teknik Informatika, Fakultas Teknologi Informasi

Jalan Raya ITS, Surabaya, Jawa Timur 60111

e-mail: bayu.nurseno@gmail.com¹⁾, handatj@its.ac.id²⁾

ABSTRAK

Epilepsi merupakan gangguan sistem syaraf otak manusia dan menyebabkan berbagai reaksi terhadap tubuh manusia. Epilepsi dapat dideteksi dengan menggunakan Electroencephalogram (EEG). Pengamatan EEG secara visual tidak mungkin dilakukan secara rutin, sehingga dibutuhkan deteksi otomatis pada EEG. Sistem deteksi EEG secara otomatis terdiri dari 2 langkah, yaitu ekstraksi fitur dan klasifikasi. Power Spectral Density (PSD) adalah metode ekstraksi fitur yang sering dipakai untuk memunculkan karakteristik EEG dengan mengelompokkan energi pada EEG. Pada proses klasifikasi metode Convolution Neural Network (CNN) dapat mereduksi fitur hasil PSD dan digunakan mengklasifikasikan multiclass dari EEG. Namun, data EEG memiliki kecenderungan bercampur noise berupa sinyal yang lain saat perekaman, oleh karena itu sebelum data EEG diklasifikasikan, perlu dilakukan pengolahan terlebih dahulu.

Pada penelitian ini diusulkan penggabungan metode Singular Spectrum Analysis (SSA) untuk penghilang noise, PSD sebagai ekstraksi fitur dan CNN sebagai klasifier. Penelitian ini dilakukan melalui beberapa fase, pertama adalah menghilangkan noise yang bercampur dengan sinyal EEG menggunakan SSA. Selanjutnya ekstraksi fitur menggunakan PSD untuk diambil energi dari sinyalnya, dan terakhir diklasifikasi dengan CNN. Pengujian klasifikasi akan dilakukan ke 500 sinyal dengan target 5 kelas dan 3 kelas. Untuk mengetahui performa terhadap metode yang diusulkan, akan dilakukan pengujian antara gabungan PSD dengan CNN yang akan dibandingkan dengan gabungan SSA, PSD dan CNN.

Berdasarkan hasil uji coba, metode yang diusulkan yaitu SSA, PSD dan CNN dapat meningkatkan rata-rata hasil akurasi klasifikasi sebesar 1,2% dari 93,2% menjadi 94,4%, untuk kasus 3 kelas dan meningkatkan 13,4% dari 78,6% menjadi 92%, untuk kasus 5 kelas dibandingkan metode PSD dengan CNN.

Kata Kunci: Epilepsi, EEG, SSA, PSD, CNN.

ABSTRACT

Epilepsy is a disorder of the human brain and causes a variety reactions to the human body. Epilepsy can be detected using Electroencephalogram (EEG). Visual analysis can't be done routinely, automatic computer system is needed. Automatic system for EEG detection consists of feature extraction and classification. Power Spectral Density (PSD) is feature extraction methods that can bring up the EEG characteristics. In the classification process, Convolution Neural Network (CNN) useful for multiclass EEG classification. However, the EEG data have a problem with noise during recording, so preprocessing needs to be done first.

In this study, we proposed the combination of Singular Spectrum Analysis (SSA) for reducing noise, PSD for feature extraction and CNN as a classifier. It is expected that the combination can improve the accuracy of EEG classification for five classes. The study was conducted through several phases, the first phase is to eliminate noise mixed with the EEG signals using SSA. Furthermore, feature were extracted using PSD and classified with CNN. The tests were for the classification of 500 signals with target of 5 classes and 3 classes. Performance of the proposed method that consists of SSA, PSD and CNN was compared with the combination method of PSD and CNN.

Based on results of the experiment, proposed method can increase the average accuracy results for epileptic disease classification by 1.2% from 93.2% to 94.4%, for 3 class case and increase 13.4% from 78.6% to 92%, for 5 class case compared to the combination method without SSA processing.

Keywords: Epilepsi, EEG, SSA, PSD, CNN.

I. PENDAHULUAN

OTAK manusia dapat melakukan berbagai aktivitas baik itu normal maupun abnormal. Kondisi normal terdiri dari kondisi fisik (seperti tidur, terjaga, dan beraktifitas) dan kondisi mental (seperti nyaman, sedih, dan marah). Kondisi abnormal dapat terjadi pada gangguan neurologis dan ketidakseimbangan akibat pengaruh obat-obatan, seperti kejang pada epilepsi dan demensia [1]. Aktivitas otak tersebut dapat diteliti dengan media penelitian berupa citra fungsional yang dihasilkan dari pengukuran sinyal otak dengan *electroencephalogram* (EEG), *Magneto-Encephalography* (MEG), dan *functional Magnetic Resonance Imaging*

(fMRI).

Pengukuran sinyal EEG merupakan suatu teknik yang paling mudah dan paling murah dibandingkan dengan pengukuran fMRI dan MEG. Resolusi temporal yang tinggi dari EEG dapat memungkinkan dideteksinya perubahan aktivitas otak lebih cepat dibandingkan dengan MEG dan fMRI, sehingga penelitian dengan menggunakan sinyal EEG lebih diminati. EEG bersifat non-invasif dan merupakan alat diagnostik yang praktis untuk penelitian berbagai kondisi otak, terutama kondisi abnormal pada kasus gangguan neurologis [2]. EEG adalah metode yang digunakan dalam mengukur aktivitas listrik spontan dari otak yang diperoleh dengan tertangkapnya sinyal listrik otak dari neuron ke neuron. Proses perekaman sinyal EEG dilakukan dalam waktu singkat, biasanya selama 20-40 menit. Rekaman diperoleh dengan menempatkan elektroda pada berbagai posisi pada kulit kepala[1]. Tingginya resolusi temporal yang EEG miliki akan membuat respon terhadap segala perubahan aktivitas otak lebih cepat.

Sinyal EEG merupakan sinyal yang sangat kompleks dan menjadi sumber informasi untuk penelitian fungsi otak dan gangguan neurologis salah satunya adalah epilepsi. Epilepsi merupakan gangguan neurologis yang mempengaruhi lebih dari 50 juta orang di seluruh dunia.

Pola dan karakteristik yang dimiliki sinyal EEG pada serangan epilepsi dapat digunakan oleh dokter atau profesional kesehatan untuk membedakan serangan epilepsi dari kondisi normal (*nonseizure*). Tetapi, analisis secara visual tidak mungkin dilakukan secara rutin, karena sinyal EEG yang dihasilkan dari sistem monitoring EEG sangat besar dan cukup memakan waktu [3]. Beberapa teknik deteksi otomatis telah diuji coba untuk mempercepat dan meningkatkan akurasi identifikasi bentuk gelombang EEG patologis yang dihubungkan dengan serangan epilepsi, dan diusulkan untuk mendeteksi lonjakan dalam EEG untuk memprediksi kejadian epilepsi.

Deteksi otomatis epilepsi berbasis EEG telah banyak dikembangkan, dalam perkembangannya terdapat beberapa fokus penelitian yang dilakukan oleh peneliti. Fokus dari peneliti dapat berupa penyempurnaan metode yang digunakan dan banyaknya kelas yang dipakai dalam proses klasifikasi. Berdasarkan banyaknya kelas yang dipakai terbagi menjadi 4 kasus penelitian yaitu untuk 2 kelas, 3 kelas, 4 kelas dan 5 kelas. Dua kelas klasifikasi merupakan penelitian yang paling sering dilakukan oleh peneliti, seperti Güler et al 2005 menggunakan Recurrent Neural Networks (RNNs) dan ekstraksi fitur Lyapunov yang di-*training* dengan algoritma Levenberg–Marquardt. Übeyli, 2006 menggunakan Multilayer Perceptron Neural Network (MLPNN), Lan Ma dkk, 2015 menggunakan Convolutional Neural Network (CNN)[4]. Untuk tiga kelas terdapat beberapa penelitian seperti yang dilakukan Tzallas et al. 2009 menggunakan *Time frequency* untuk mengekstraksi fitur yang ada pada sinyal EEG dan mengklasifikasikan EEG ke dalam 3 kelas dengan ANN. Dalam penelitian lain Faust et al. 2015 menggunakan Wavelet untuk mengklasifikasikan 3 kelas EEG epilepsi [5]. Untuk penelitian lima kelas, masih jarang dilakukan karena dianggap memerlukan analisis yang mendalam untuk masing–masing kelas sinyal. Penelitian yang telah dilakukan untuk lima kelas, seperti yang dilakukan Tjandrasa et al. 2016 dengan menggabungkan Intrinsic Mode Functions dan Power Spectrum dalam pengklasifikasian sinyal EEG untuk kasus epilepsi [6].

Dalam tahapan proses klasifikasi atau *machine learning*, seleksi fitur adalah salah satu teknik yang penting dalam *pre-processing*. Seleksi fitur adalah proses memilih *subset* dari fitur asli sehingga jumlah fitur berkurang secara optimal sesuai dengan kriteria yang ditentukan. Teknik ini terbukti efektif mengurangi fitur-fitur yang tidak relevan, berlebihan, meningkatkan efisiensi dalam proses *learning*, dan meningkatkan kinerja *learning* seperti akurasi pada proses prediksi. Oleh karena itu, seleksi fitur menjadi sangat diperlukan oleh aplikasi *machine learning* ketika menghadapi data dengan dimensi yang tinggi. [7].

SSA dalam sejarah penggunaannya digunakan untuk memisahkan sinyal *single channel* menjadi beberapa sinyal berdasarkan *trend*, komposisi dan juga *noise* [8]. SSA dapat mengurangi *noise* yang mempengaruhi pola utama atau ciri dari suatu sinyal. Lokal SSA digunakan untuk memisahkan sinyal EOG dari sinyal EEG yang akan digunakan pada tahapan selanjutnya dalam pengklasifikasian EEG [9].

Dalam ekstraksi fitur untuk data berupa sinyal metode PSD juga sering dipakai karena mampu digunakan untuk menganalisa pola dari sinyal berdasarkan kekuatan spektrum yang dimiliki sebuah sinyal. PSD digunakan untuk menganalisa aktivitas *Electrodermal* dalam penilaian fungsi simpatik [10].

Pada penelitian ini diajukan metode hybrid untuk pengolahan data awal menggunakan penggabungan Singular Spectrum Analysis (SSA) untuk menghilangkan noise dan memunculkan karakteristik sinyal, dilanjutkan Power Spectral Density (PSD) berbasis metode Welch dan Convolution Neural Network (CNN) untuk mengklasifikasikan sinyal. Penambahan metode kombinasi ekstraksi fitur tersebut diharapkan mampu meningkatkan hasil akurasi pada klasifikasi data sinyal EEG.

II. DASAR TEORI

A. Epilepsi

Epilepsi merupakan gangguan pada sistem syaraf otak manusia yang mengakibatkan terjadinya aktivitas yang berlebihan dari sel neuron pada otak. Epilepsi menyebabkan berbagai reaksi pada tubuh manusia seperti kesemutan, bengong sesaat, kejang-kejang gangguan kesadaran, dan kontraksi otot pada tubuh manusia.

Epilepsi memiliki beberapa karakteristik tertentu [11]:

- 1) Terjadi gangguan yang bervariasi yang diakibatkan adanya proses disfungsi sistem saraf pusat
- 2) Dalam sehari serangan dapat terjadi hanya 1 kali serangan sampai beberapa kali serangan tergantung beratnya serangan yang dialami.
- 3) Epilepsi dapat berakibat buruk pada kehidupan sosial penderita karena adanya stigma yang terbangun pada saat terjadinya serangan.
- 4) Disebabkan oleh efek samping pemakaian obat jangka panjang yang mungkin diremehkan.

B. Electroencephalography

Electroencephalogram (EEG) adalah metode yang digunakan dalam mengukur aktivitas listrik spontan dari otak yang diperoleh dengan menangkap sinyal listrik yang ditembakkan dari neuron ke neuron dalam otak. Proses perekaman sinyal EEG dilakukan dalam waktu singkat, biasanya selama 20-40 menit. Rekaman diperoleh dengan menempatkan elektroda di berbagai posisi pada kulit kepala. Tingginya resolusi temporal yang EEG miliki akan membuat respon terhadap segala perubahan aktivitas otak lebih cepat [2].

Terdapat dua pendekatan untuk mendapatkan sinyal EEG sesuai dengan lapisan otak mana yang akan diambil sinyalnya. Pendekatan pertama adalah pendekatan invasif, yaitu dengan menanam elektroda yang sangat kecil secara langsung di atas korteks melalui bedah saraf. Keuntungan dari pendekatan ini adalah mampu memberikan sinyal EEG dengan kualitas sangat tinggi. Pendekatan lain adalah pendekatan non invasif, yaitu dengan menempatkan elektroda pada permukaan kulit kepala. Hasil rekaman EEG non invasif memiliki kualitas sinyal yang kurang bagus karena tengkorak memperkecil sinyal, mendispersi, dan mengaburkan gelombang elektromagnetik yang dihasilkan oleh neuron. Selain itu juga memiliki resolusi spasial rendah sehingga sulit untuk menentukan wilayah otak yang menghasilkannya atau respon masing-masing neuron [2].

Sinyal EEG terbukti cukup membantu proses pemantauan dan pendiagnosis sejumlah kasus klinis, yaitu: Epilepsi, Penyakit Alzheimer, Penyakit Huntington, Gangguan tidur. EEG digunakan ahli saraf untuk mempelajari fungsi otak. Penggunaan metode dalam Computational Intelligence (CI) dan pemodelan matematika membantu proses analisis EEG. Selain itu juga telah dikembangkan beberapa aplikasi Brain Computer Interface (BCI) untuk membantu penderita cacat melakukan kegiatan sehari-hari. BCI dapat memfasilitasi komunikasi individu cacat fisik dengan bantuan komputer dengan menggunakan karakteristik sinyal EEG. Aplikasi EEG lainnya adalah pembangkitan kemampuan atau respon otak yang muncul terhadap suara, sentuhan, dan cahaya, yang berguna untuk mengevaluasi sejumlah kondisi neurologis [2].

Berdasarkan frekuensi, amplitudo, sinyal EEG dapat dibagi menjadi 4 gelombang, yaitu gelombang delta (kurang dari 4 Hz), theta (4 – 7 Hz), alpha (8 – 12 Hz), beta (13 – 20 Hz) dan gamma (21 – 40 Hz) [12].

C. Singular Spectrum Analysis

SSA adalah sebuah teknik berbasis ruang bagian yang sering digunakan untuk analisis geofisika iklim dan time series [13]. Tujuan utama dari SSA adalah untuk menguraikan sinyal yang berupa *single channel* sinyal menjadi beberapa sinyal, dengan masing-masing sinyal pecahan tersebut dapat diindikasikan sebagai *trend*, *periodic/quasi-periodic component* dan *noise* [8]. Dengan adanya pemisahan sinyal tersebut sinyal dapat dipilah mana yang paling mewakili *trend* kelas pada saat klasifikasi.

$$X = \begin{pmatrix} y_1 & y_2 & \cdots & y_K \\ y_2 & y_3 & \cdots & y_{K+1} \\ \vdots & \vdots & \ddots & \vdots \\ y_L & y_{L+1} & \cdots & y_T \end{pmatrix} \quad (1)$$

Pada dasarnya, SSA terdiri dari dua tahapan, yaitu *decomposition* dan *reconstruction*. Dalam *decomposition* dibagi lagi menjadi 2 langkah lagi, yaitu *embedding* dan Singular Value Decomposition (SVD). *Embedding* merupakan proses pemecahan *single channel* ke *multichannel* sinyal dengan memanfaatkan *lagged vector single channel*. Apabila *single channel* $Y_T = (y_1, y_2, \dots, y_T)$ akan diubah ke *multichannel* X_1, \dots, X_k dengan vektor $X_i = (y_i, y_{i+1}, \dots, y_{i+L-1})^T \in R^T$, dimana $K = T - L + 1$ seperti persamaan matriks (1) [8]. Vektor X_i adalah vektor *L-lagged* dari Y . Di fase *embedding* ini terdapat 1 parameter, yaitu *Window Length* atau L yang harus

disesuaikan dengan permasalahan yang dihadapi. Parameter L harus cukup lebar dan sesuai agar dapat memisahkan sinyal dengan baik dan berada pada range $2 \leq L \leq T$.

Setelah didapat matriks X pada langkah 1, selanjutnya adalah langkah kedua yaitu Singular Value Decomposition (SVD). Pertama definisikan *covariance* matriks $C = XX^T$ yang akan digunakan untuk menghitung *eigenvalues* dan *eigenvectors*. *Eigenvalues* akan dicari sebanyak L nilai dengan terurut dari yang paling besar ke terkecil ($\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_L \geq 0$) dan *eigenvector* (U_1, U_2, \dots, U_L). Setelah didapat *eigenvalues* dan *eigenvectors* selanjutnya SVD akan ditulis dengan matriks $X = E_1 + E_2 + \dots + E_d$, dimana $E_i = \sqrt{\lambda_i} U_i V_i^T$, d adalah banyaknya *eigenvalue* yang lebih dari 0, dan (V_1, V_2, \dots, V_d) adalah Principal Component (PC) dengan nilai $V_i = X^T U_i \sqrt{\lambda_i}$ (Enrico, 2013). Dan koleksi dari ($V_i, U_i, \sqrt{\lambda_i}$) disebut dengan *eigen triple* dari matriks X (nina, 2013).

Langkah yang kedua adalah *reconstruction* dan akan dibagi menjadi 2 step yaitu *grouping* dan *diagonal averaging*. Dalam fase *Grouping*, *eigen triple* yang didapat akan digunakan untuk memisahkan sinyal yang dianggap sebagai *trend* dan *noise* [14]. Berdasarkan urutan *eigenvalue* yang dibatasi nilai 0, akan didapatkan I dimana $I = (1, 2, \dots, m)$, dimana m banyaknya *eigenvalue* yang masuk ke dalam subset. Setelah ditentukan nilai I maka matriks X akan dapat dipisah, $X = X_I + X_{\bar{I}}$, dimana $X_I = \sum_{i \in I} X_i$ dan $X_{\bar{I}} = \sum_{i \notin I} X_i$. Apabila pemilihan I sesuai, X_I akan mencerminkan *trend* dari kelas yang sesuai dengan potongan sinyal tersebut dengan memisahkan $X_{\bar{I}}$.

Fase terakhir adalah *Diagonal Averaging*, yang ditujukan untuk mengubah matriks X_I menjadi Hankel matriks yang kemudian akan diubah menjadi form originalnya. Jika z_{ij} bagian dari matriks Z , dan k adalah optimal result dari merata rata z_{ij} dari i, j sehingga $i + j = k + 2$. Prosedur ini dinamakan *diagonal averaging* atau *Hankelization* dari matriks Z . Hankel matriks dari Z adalah $\mathcal{H}Z$, menurut riset *Hankelization procedure* matriks $\mathcal{H}Z$ merupakan matriks yang paling mirip dengan Z . Dengan menggunakan procedure tersebut digunakan untuk matriks yang ada menjadi $X = \tilde{X}_{I_1} + \dots + \tilde{X}_{I_m}$, dimana $\tilde{X}_{I_1} = \mathcal{H}Z$. Persamaan tersebut sama dengan dekomposisi dari $Y_T = (y_1, y_2, \dots, y_T)$ ke dalam penjumlahan seri d menjadi $y_t = \sum_{k=1}^m \tilde{y}_t^{(k)}$, dimana $\tilde{Y}_T^{(k)} = (\tilde{Y}_1^{(k)}, \dots, \tilde{Y}_T^{(k)})$ berhubungan ke matriks X_{I_k} .

D. Power Spectral Density

Spektrum frekuensi merupakan salah satu fitur yang dapat diambil dari EEG. Spektrum frekuensi tersebut dapat diambil dengan mentransformasikan sinyal EEG dari *time-based* ke *frequency-based* sehingga akan didapatkan fitur frekuensi yang ada sinyal tersebut. Metode yang dapat digunakan untuk perhitungan estimasi spectrum daya adalah metode Welch [15].

Pada metode ini, sinyal input dibagi menjadi beberapa segmen yang pendek dan akan dihitung periodogramnya dengan menggunakan nilai imajiner *Fast Fourier Transform*. Sebelum dilakukan perhitungan periodogram, segmen – segmen data tersebut akan dimodifikasi dengan fungsi jendela (window) yang ada. Selanjutnya periodogram yang telah didapat akan dirata sehingga menghasilkan estimasi spektrum yang lebih baik [15][16].

Algoritma Power Spektral Density menggunakan metode Welch adalah sebagai berikut [17]:

1) Sinyal input $x[N]$ yang dibagi menjadi sebanyak L *overlapping* segmen menjadi (2):

$$x_l(n) = x(n + (l - 1)M) \tag{2}$$

2) dimana, $n = 0, \dots, N - 1$ dan $l = 1, \dots, L$ dengan rekomendasi untuk metode *welch* $M = N/2$. Lalu masing-masing segmen akan dicari *window*-nya (3). Setelah didapat *windowed* segmen akan diolah dengan FFT (4).

$$A_l(k) = \sum_{n=0}^{N-1} x_l(n)w(n)e^{-j\frac{2\pi}{N}nk} \tag{3}$$

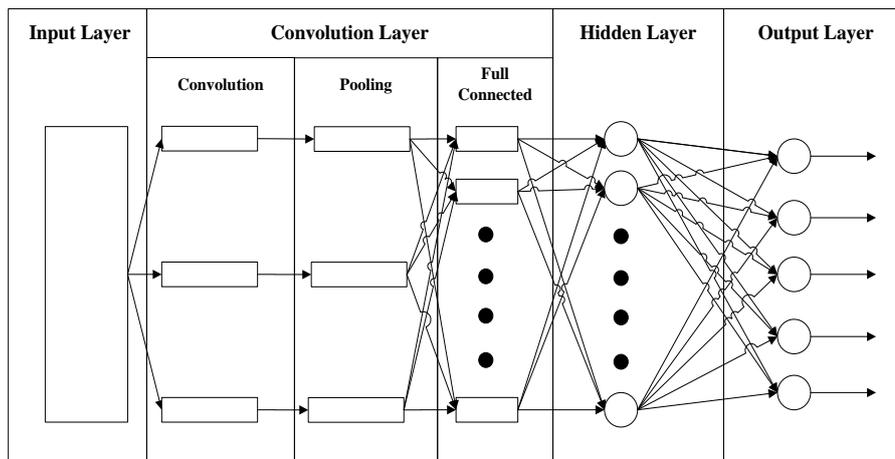
$$\Phi_l(k) = \frac{1}{NP} |A_l(k)|^2 \tag{4}$$

dimana A_l adalah FFT *windowed* segmen, Φ_l adalah periodogram, dan P menunjukkan power window ($w(n)$), dengan (5)

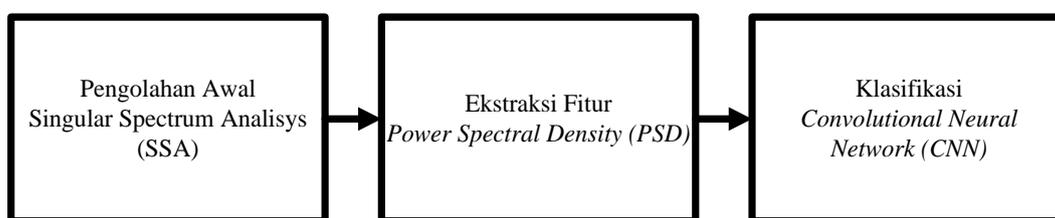
$$P = \frac{1}{N} \sum_{n=0}^{N-1} |w(n)|^2 \tag{5}$$

dan akan diestimasi dengan metode Welch dari PSD dengan melihat rata rata dari periodogram, dengan (6)

$$S(k) = \frac{1}{L} \sum_{l=1}^L \Phi_l(k) \tag{6}$$



Gambar 1. Convolution Neural Network



Gambar 2. Diagram alur sistem yang diusulkan

E. Convolutional Neural Network

Convolutional Neural Network (CNN) adalah pengembangan dari Multilayer Perceptron (MLP) yang didesain untuk mengolah data dua dimensi dan banyak diaplikasikan pada data citra dan suara. Perbedaan dari MLP adalah banyaknya layer dan fungsinya. CNN memiliki layer yang digunakan untuk mengekstraksi data awal sehingga lebih efektif digunakan pada layer klasifikasi.

Langkah dari CNN dapat dilihat pada Gambar 1 dimana data input akan dipecah menjadi beberapa bagian pada *convolution layer* yang masing-masing bagian mewakili fitur yang akan digunakan. *Convolution layer* melakukan operasi konvolusi dari layer sebelumnya yaitu input layer. Layer tersebut adalah proses utama yang mendasari sebuah CNN. Tujuan dari *convolution layer* adalah untuk memilih fitur yang akan di gunakan pada proses Multilayer Perceptron. Fitur tersebut didapatkan dengan cara memfilter data input dengan *convolution layer* yang sesuai dengan fitur yang diinginkan. Untuk mendapatkan fitur yang optimum biasanya filter yang digunakan lebih dari satu. Masing-masing filter tersebut akan diaplikasikan ke data input dan setiap filter menghasilkan fitur masing-masing untuk data input yang sama. Penggunaan filter yang banyak tersebut bertujuan untuk membatasi variabel bebas yang mengurangi bahkan menutupi karakteristik dari data input.

Setiap Convolution layer diikuti dengan subsampling layer, dimana hasil dari convolution di ambil sampling untuk mengecilkan dimensi matriks. Dalam pengolahan data citra pada khususnya, subsampling juga bertujuan untuk meningkatkan invariansi posisi dari fitur. Dalam sebagian besar CNN, metode subsampling yang digunakan adalah Max Pooling [18]. Max Pooling membagi output dari *convolution layer* menjadi beberapa grid kecil lalu mengambil nilai maksimal dari setiap grid untuk menyusun matriks citra yang telah direduksi. Langkah *convolution layer* dan *subsampling layer* akan diulang sampai didapat matriks yang dirasa cukup kecil untuk dilakukan proses klasifikasi [19].

Setelah didapatkan matriks yang cukup kecil dan tidak menghilangkan informasi, hasil dari *convolution layer* akan disatukan ke *Fully Connected Layer*. Layer tersebut adalah layer yang biasanya digunakan sebagai input layer pada *Multilayer Perceptron* dan bertujuan untuk melakukan transformasi pada dimensi data agar data dapat diklasifikasikan secara linear.

III. METODOLOGI PENELITIAN

Secara keseluruhan desain model sistem yang akan dibangun pada penelitian ini memiliki 3 tahapan seperti yang ditunjukkan pada Gambar 2. Tahapan tersebut adalah pengolahan awal sinyal, Ekstraksi fitur sinyal, dan

terakhir adalah proses klasifikasi sinyal.

A. Dataset

Data sinyal EEG yang digunakan dalam penelitian ini diperoleh dari *database* yang tersedia di Universitas Bonn. Data ini tersedia secara *online* dan dibuat oleh Dr. Ralph Andrzejak dari Pusat Epilepsi di Universitas Bonn, Jerman¹. *Dataset* secara lengkap terdiri atas lima kelas yaitu A, B, C, D, dan E. Masing-masing kelas terdiri atas 100 segmen sinyal EEG *single channel* dengan durasi selama 23,6 detik. Data disimpan dalam format teks (*.txt).

Set A dan set B adalah sinyal yang diambil dari rekaman EEG yang dilakukan pada lima sukarelawan sehat dengan skema penempatan elektroda standar (*International 10-20 system*) dalam kondisi santai, dengan set A diambil dengan kondisi mata terbuka dan set B diambil saat mata tertutup. Set C-E berasal dari arsip EEG diagnosis *presurgical* dari lima pasien dipilih, dan semua telah mencapai kontrol kejang yang lengkap, setelah reseksi dari salah satu formasi hippocampal, sehingga didiagnosis dengan benar masuk zona *epileptogenic*.

Sinyal set D direkam saat zona *epileptogenic*, dan berada pada interval tanpa kejang dan set C berasal dari pembentukan *hippocampus* pada belahan yang berlawanan dari otak. Sementara set C dan D berisi aktivitas yang hanya diukur selama interval tanpa kejang, sedangkan set E hanya berisi aktivitas kejang. Data set A - E tersebut yang digunakan dalam penelitian ini.

B. Pengolahan Awal

SSA merupakan sebuah metode yang digunakan untuk menganalisis data *time series*. Analisis yang dilakukan oleh SSA adalah dengan memisahkan sinyal yang dianggap noise dan sinyal yang dianggap sebagai sinyal yang asli. Pemisahan sinyal tersebut dilakukan dengan menerapkan *principal component* pada data *time series*. Dengan memanfaatkan *eigenvalue* pada *principal component* yang diterapkan pada *single channel time series*, data *time series* tersebut dapat direkonstruksi.

SSA terdapat 2 tahapan yang dilakukan, yaitu *decomposition* dan *reconstruction*. Pertama adalah *decomposition*, dalam tahap langkah pertama adalah *embedding* ini data sinyal EEG akan diubah dari *single channel* sinyal menjadi multichannel sinyal. Dengan memanfaatkan sinyal yang memiliki tipe *time series* dimana data berurut berdasarkan waktu sinyal dapat dipisah berdasarkan *lagged time* atau waktu delay dari sinyal tersebut. Ide dasar dari tahap ini, untuk mencari *principal component* dengan mencoba menghitung *covariance* antara nilai $X(t)$ dan $X(t+k)$, dimana k adalah jumlah delay dari sinyal tersebut. Setiap sinyal k memiliki panjang L (*window size*) dan berfungsi sebagai kolom / fitur untuk *lagged* sinyal. Panjang *window* L akan mempengaruhi jumlah *eigenvalue* pada tahapan selanjutnya. Pada penelitian ini panjang L adalah 2000 mengacu Hassani et al. pada 2009 ukuran dari L sebesar – besarnya dengan range $2 \leq L \leq T/2$ dimana T adalah panjang sinyal asli. Setelah didapatkan k sinyal sinyal tersebut disatukan sehingga seolah-olah didapat *multiclass* data yang disebut *trajectory matrix*.

Langkah kedua dari *decomposition* adalah Singular Value Decomposition (SVD), setelah didapatkan *trajectory matrix*, selanjutnya dari matrix tersebut akan dicari covariannya. Berdasarkan *covariance matrix* akan dihitung *eigenvalue* dan *eigenvector*-nya.

Tahapan kedua dari SSA adalah *reconstruction*. Langkah pertama dari rekonstruksi adalah *grouping*, pada tahapan ini *trajectory matrix* akan ditransformasikan menjadi matrix baru berdasarkan *eigenvalue*. Pemilihan *eigenvalue* dilakukan berdasarkan pengamatan hubungan *eigenvalue* dengan frekuensi sinyal dalam penelitian yang dinilai sesuai dengan karakteristik.

Langkah kedua dari tahapan *reconstruction* adalah *Diagonal Averaging*. *Diagonal Averaging* bertujuan untuk mengubah matrix hasil *grouping* menjadi *single channel* seperti pada awalnya. Diharapkan data hasil rekonstruksi terbebas dari noise dan karakteristiknya dapat diklasifikasikan oleh klasifier.

C. Ekstraksi Fitur

Setelah didapatkan hasil dari proses SSA dilanjutkan ke ekstraksi fitur. Dalam fase ini sinyal akan diambil adalah spektrum frekuensi. dengan mentransformasikan sinyal EEG dari *time-based* ke *frequency-based* sehingga akan didapatkan fitur frekuensi yang ada sinyal tersebut. Metode yang dapat digunakan untuk perhitungan estimasi spectrum daya adalah metode Welch [15].

Pada metode ini, sinyal input dibagi menjadi beberapa segmen yang pendek dan akan dihitung periodogramnya dengan menggunakan nilai imajiner Fast Fourier Transform. Sebelum dilakukan perhitungan periodogram segmen-segmen data tersebut akan dimodifikasi dengan fungsi jendela (*window*) yang ada. Selanjutnya periodogram yang telah didapat akan dirata-rata sehingga menghasilkan estimasi spektrum yang lebih baik.

¹ (www.epileptologie-bonn.de/cms/front_content.php?idcat=193&lang=3&changelang=3)

$$X' = \frac{X - \text{Min}(X)}{\text{Max}(X) - \text{Min}(X)} (0.9 - 0.1) + 0.1 \quad (7)$$

Setelah didapatkan hasil ekstraksi fitur untuk masing-masing kelompok yang akan diujikan, selanjutnya sinyal hasil PSD akan dinormalisasi untuk masing-masing sinyal. Normalisasi untuk masing-masing sinyal ini dilakukan untuk menyamakan rentang amplitudo sinyal dari semua kelas. Berdasarkan pengamatan hasil ekstraksi fitur antara sinyal set A–set D memiliki rentang yang sangat jauh dengan set E sehingga diperlukan normalisasi. Selain perbedaan rentang yang mencolok normalisasi ini juga bertujuan untuk mempermudah pemfilteran sinyal pada proses klasifikasi menggunakan CNN. Sinyal EEG hasil ekstraksi fitur akan diubah rentangnya menjadi 0,1 –0,9. Normalisasi yang akan digunakan menggunakan persamaan (7).

D. Klasifikasi

Tahap ketiga dari metode penelitian adalah melakukan klasifikasi dengan menggunakan CNN. Dalam tahapannya terdapat dua tahapan utama dalam proses uji cobanya yang pertama adalah tahapan *convolution* dan yang kedua adalah Neural Network. Pada tahapan *convolution* sinyal hasil *normalisasi* dari ekstraksi fitur akan di filter menggunakan dua filter, yang pertama adalah *low-pass filter* dan *high-pass filter*. Dari penelitian sebelumnya khususnya pada klasifikasi EEG epilepsi, fitur dengan frekuensi rendah memiliki pengaruh besar dalam pengklasifikasian EEG pada kasus epilepsi. Penggunaan *high-pass* filter ditujukan untuk menangani set E pada pengklasifikasian sinyal EEG epilepsi. Set E untuk data EEG epilepsi banyak terdapat lonjakan amplitudo sehingga fitur dengan amplitudo yang tinggi dianggap mewakili karakter dari sinyal tersebut. Data yang diolah dalam bentuk sinyal, maka untuk masing – masing filter yang digunakan akan diaplikasikan menggunakan 1-D filter untuk *convolution*. Setiap filter yang digunakan akan memiliki dimensi 1x5.

Setelah mendapatkan hasil dari masing-masing filter selanjutnya setiap data hasil filter tersebut akan di *pooling* dengan menggunakan Max Pooling. Dimana Max Pooling merupakan filter yang mereduksi dimensi fitur dengan cara mengambil salah satu nilai fitur dengan memilih nilai yang paling tinggi antar *grid* yang sudah ditentukan, dan banyaknya tetangga yang dibandingkan akan sesuai dengan *window* yang ditentukan di awal. Dalam penelitian ini *window* dari Max Pooling adalah 1x2, sehingga fitur dapat tereduksi.

Proses yang kedua adalah Neural Network, fitur hasil dari *max pooling* yang disatukan. Seperti biasa langkah dari Neural Network adalah Forward Propagation dimana data input adalah hasil filter yang di *pooling* dan dihitung dengan fungsi aktivasi yang digunakan. Setelah itu penghitungan error apabila error masih berada di atas batas atau lebih besar dari nilai yang ditentukan maka akan dilakukan proses *backpropagation* untuk memperbaiki parameter yang digunakan dengan cara Gradient Descent dengan tujuan supaya *error* berkurang, apabila *error* lebih kecil dari batas maka iterasi akan dihentikan. Model CNN dengan nilai *error* terkecil yang diperoleh pada proses *training* akan diuji dalam proses testing untuk melihat kecakapan model terhadap data lain.

IV. UJI COBA

Pada tahap ini akan dilakukan analisis dari hasil uji coba untuk metode yang akan digunakan. Untuk mengetahui performa dari metode yang digunakan pada penelitian ini akan dilakukan dua skenario percobaan, yang pertama adalah penggabungan SSA, PSD dengan CNN dan yang kedua adalah penggabungan PSD dengan CNN. Dengan kedua skenario penelitian tersebut akan dihitung akurasi, sehingga diharapkan dari analisis uji coba ini akan diperoleh hasil yang sesuai dengan tujuan penelitian.

Selain dua skenario utama yang akan diujikan, akan dibahas pengaruh parameter-parameter yang mempengaruhi uji coba yang dilakukan. Parameter tersebut adalah pengaruh pemilihan *eigenvalue* pada SSA dan penggunaan normalisasi untuk menyamakan range dari hasil dari ekstraksi fitur. Pemilihan *eigenvalue* merupakan hal yang penting karena sangat mempengaruhi hasil akurasi pada proses klasifikasi. Dalam penelitian ini dalam penentuan *eigenvalue* akan dilakukan 9 uji coba berdasarkan pengamatan dan *trial error*. Sembilan uji coba tersebut akan dilakukan pada kelompok *eigenvalue* ke - 1, 1-100, 100-200, 200-1000, 1000-2000, 1-500, 500-1000, 1000-1500, 1500-2000. Penentuan kelompok *eigenvalue* ditentukan berdasarkan besarnya *eigenvalue*, dimana untuk masing-masing sinyal input memiliki 2000 *eigenvalue* dan dikelompokkan berdasarkan besaran nilai tersebut. Pada kasus EEG untuk dataset Universitas Bonn yang digunakan pada penelitian ini kelompok *eigenvalue* yang paling bagus berdasarkan percobaan yang dilakukan adalah pada *eigenvalue* ke 200 – 1000, seperti TABEL I.

TABEL I
HASIL UJI COBA KELOMPOK *EIGENVALUE*

Kelompok <i>Eigenvalue</i>	Akurasi	
	3 kelas(%)	5 kelas(%)
1	59,2	44,2
1-100	88	71,6
1-500	92,8	75,2
100-200	91,6	84,6
200-1000	94,4	92
500-1000	93,2	89,8
1000-1500	90,6	83,5
1000-2000	90,2	85
1500-2000	90,6	87,6

TABEL II
HASIL SKEMA NORMALISASI

No	Skema	Akurasi(%)
1.	Tanpa Normalisasi	85,4
2.	Mormalisasi	92

Setelah melakukan pengamatan pada penelitian ini, hasil sinyal dari ekstraksi fitur memiliki tingkat range power berbeda-beda. Perbedaan range tersebut sangat mempengaruhi hasil dari klasifikasi yang dilakukan. Untuk mengatasi perbedaan range tersebut dilakukan normalisasi. Normalisasi yang dilakukan dengan cara mengubah range setiap sinyal yang akan diuji. Sehingga setiap sinyal yang akan diolah memiliki nilai minimal dan nilai maksimal yang sama. Dengan range yang sama tersebut karakteristik sinyal dari masing–masing kelas akan lebih terlihat dari pola masing–masing sinyal dan tidak terpengaruh oleh perbedaan nilai yang besar untuk masing–masing sinyal. Seperti pada TABEL II yang merupakan perbedaan pengolahan sinyal antara yang menggunakan normalisasi dan tidak untuk 5 kelas.

Dalam penelitian ini akan dilakukan 2 skenario percobaan untuk mengetahui hasil dari gabungan metode yang dihasilkan yaitu SSA, PSD, dan CNN dibandingkan dengan PSD dan CNN. Dari hasil percobaan kedua skenario tersebut membuktikan bahwa SSA, PSD dan CNN memiliki akurasi yang lebih baik untuk 2 kasus kelas percobaan yaitu untuk 5 kelas dan 3 kelas seperti TABEL III. Pada TABEL III terdapat nilai akurasi hasil percobaan metode yang diusulkan dengan metode sebelumnya yaitu penggabungan PSD dengan CNN. Kedua skenario yang diujikan dengan menggunakan parameter yang sama untuk parameter pada PSD dan CNN, sesuai dengan pembahasan sebelumnya. Sehingga dapat dibuktikan bahwa penambahan SSA untuk memunculkan karakteristik dari sinyal EEG epilepsi akan menghasilkan akurasi yang lebih baik.

Untuk kasus 3 kelas EEG metode yang diusulkan berhasil melakukan klasifikasi dengan akurasi 94,4 % lebih baik sebesar 1,2 % daripada metode sebelumnya 93,2 %. Sedangkan untuk kasus 5 kelas metode usulan memiliki nilai akurasi 92 % meningkat sebesar 13,4 % dari metode sebelumnya 78,6 %. Untuk 5 kelas EEG metode usulan memiliki selisih akurasi yang cukup signifikan dibandingkan dengan sebelumnya berbeda dengan hasil

TABEL III
UJI COBA SKENARIO

No.	Skenario	Akurasi	
		3 Kelas (%)	5 Kelas (%)
1.	SSA, PSD, CNN	94,4	92
2.	PSD, CNN	93,2	78,6

TABEL IV
PRECISION DAN *RECALL* KLASIFIKASI 3 KELAS

Kelas	Precision (%)		Recall (%)	
	PSD, CNN	SSA, PSD, dan CNN	PSD, CNN	SSA, PSD, dan CNN
Set AB	94	98,5	93,07	98,5
Set CD	93	93,5	92,53	92,57
Set E	92	88	94,8	89,8

TABEL V
PRECISION DAN RECALL KLASIFIKASI 5 KELAS

Kelas	Precision (%)		Recall (%)	
	PSD, CNN	SSA, PSD, dan CNN	PSD, CNN	SSA, PSD, dan CNN
Set A	86	90	85,14	91,83
Set B	93	94	88,57	91,26
Set C	68	96	64,76	96,96
Set D	52	93	62,65	87,77
Set E	94	87	88,67	92,55

pengklasifikasian 3 kelas yang memiliki perubahan yang tidak signifikan. Hal ini dapat terjadi karena terjadi perbedaan karakteristik dalam pengklasifikasian 3 kelas dan 5 kelas.

Berdasarkan TABEL IV terdapat beberapa nilai presisi dari setiap kelas untuk masing–masing metode yang diujikan. Pada set AB dan set CD metode usulan memiliki nilai presisi dan recall yang lebih baik dibandingkan metode sebelumnya. Penambahan SSA mampu meningkatkan presisi untuk kedua kelas tersebut. Pada set E metode sebelumnya memiliki nilai presisi dan recall yang lebih baik dibandingkan metode usulan. Pada metode SSA terdapat langkah yang mengharuskan untuk memilih kelompok eigenvalue yang bertujuan untuk mengelompokkan sinyal sesuai karakteristiknya. Berdasarkan pengamatan pada penelitian sebelumnya untuk memunculkan karakteristik sinyal EEG, perlu memisahkan sinyal dengan noise yang ikut terekam. Salah satu noise yang ada adalah EOG dimana noise tersebut memiliki karakteristik lonjakan amplitudo pada sinyal EEG. Dari pengamatan yang dilakukan untuk set A, set B, set C dan set D sangat terpengaruh oleh adanya lonjakan amplitudo. Dengan adanya lonjakan amplitudo karakteristik kelas tersebut menjadi kurang dominan sehingga susah diklasifikasikan. Berbeda dengan kelas yang lain untuk data set E memiliki karakteritik utama dengan banyaknya lonjakan amplitudo sehingga dengan pemilihan eigenvalue yang terfokus pada pemisahan EEG dengan sinyal dengan amplitudo yang besar maka karakteristik utama set E berkurang.

Pada kasus 5 kelas terdapat tingkat akurasi untuk metode yang diusulkan adalah 92 % meningkat sebesar 13,4 % dari metode sebelumnya 78,6 %. Penambahan SSA mampu memberi peningkatan akurasi yang signifikan pada kasus 5 kelas. Berbeda dengan kasus 3 kelas yang memiliki perbedaan yang cukup mencolok, kasus 5 kelas tidak memilikinya. Kelas dalam kasus 5 kelas merupakan pecahan dari kelas – kelas yang ada pada 3 kelas. Set A dan set B pada kasus 5 kelas merupakan pecahan set AB pada 3 kelas begitu pula dengan set C dan set D merupakan pecahan dai set CD. Kelas hasil pecahan tersebut memiliki kemiripan dengan kelas pecahan dari asal yang sama, misal set A mirip dengan set B dan set C mirip dengan set D. Kemiripan tersebut dapat dilihat dari hasil precision dan recall seperti pada TABEL V, dimana pada metode sebelumnya terdapat banyak salah klasifikasi terutama pada set C dan set D. Hasil recall dari kombinasi PSD dengan CNN, untuk semua kelas masih kurang memuaskan itu menandakan adanya banyak salah klasifikasi.

Pada metode yang diusulkan semua nilai recall untuk semua kelas lebih baik daripada kombinasi PSD dengan CNN, walaupun untuk nilai precision pada set E lebih kecil dibandingkan kombinasi PSD dengan CNN. Menurunnya nilai precision pada set E dipengaruhi oleh pemilihan eigenvalue seperti yang dijelaskan sebelumnya. Dengan demikian metode SSA mampu memunculkan karakteritik sinyal EEG dengan memisahkannya dari noise dan sinyal EOG, terutama untuk set A, set B, set C, dan set D.

V. KESIMPULAN DAN SARAN

Pada Bab ini dijelaskan kesimpulan akhir yang didapatkan dari penelitian yang telah dilakukan dan juga dipaparkan saran-saran yang bersifat membangun untuk penelitian selanjutnya di masa yang akan datang.

A. Kesimpulan

Dari hasil uji coba yang dilakukan, dapat diambil kesimpulan yaitu penambahan metode SSA dalam kombinasi metode SSA, PSD dan CNN mampu meningkatkan akurasi klasifikasi sinyal EEG untuk set A, set B, set C, set D dan set E sebesar 1,2 % untuk 3 kelas dan 13,4 % untuk 5 kelas. Tingkat akurasi gabungan SSA, PSD dan CNN lebih baik dibandingkan kombinasi PSD dan CNN untuk dua jumlah kelas yang diuji yaitu 92 % berbanding 78,6 % untuk 5 kelas dan 94,4 % berbanding 93,2 % untuk 3 kelas. Pemilihan kelompok eigenvalue yang tepat pada metode SSA dapat digunakan untuk memunculan karakteristik dari sinyal EEG epilepsi yang ditunjukkan dengan peningkatan akurasi dibandingkan dengan tidak menggunakan SSA, yaitu 78,6 % menjadi

92 % dengan kelompok eigenvalue terbaik adalah 200-1000. Penyamaan besaran *range power spectrum* dengan menggunakan normalisasi dapat meningkatkan akurasi sebesar 6,6 %.

B. Saran

Pengembangan yang bisa dilakukan untuk pada penelitian selanjutnya adalah mengembangkan atau mengoptimasi pemilihan kelompok *eigenvalue* pada SSA dan melakukan lebih banyak percobaan pada CNN sehingga filter dapat ditambah dan dapat memperbaiki hasil akurasi.

DAFTAR PUSTAKA

- [1] M. Teplan, "Fundamentals of EEG measurement," *Meas. Sci. Rev.*, vol. 2, no. 2, pp. 1–11, 2002.
- [2] M. A. Karyawan, A. Zainal, and A. Saikhu, "Klasifikasi Sinyal Eeg Menggunakan Koefisien Autoregresif," *J. TIF*, vol. 2, no. 1, pp. 1–13, 2011.
- [3] E. D. Übeyli, "Least squares support vector machine employing model-based methods coefficients for analysis of EEG signals," *Expert Syst. Appl.*, vol. 37, no. 1, pp. 233–239, 2010.
- [4] L. Ma, J. W. Minett, T. Blu, and W. S. Wang, "Resting State EEG-Based Biometrics for Individual Identification Using Convolutional Neural Networks," pp. 2848–2851, 2015.
- [5] O. Faust, U. R. Acharya, H. Adeli, and A. Adeli, "Wavelet-based EEG processing for computer-aided seizure detection and epilepsy diagnosis," *Seizure*, vol. 26, pp. 56–64, 2015.
- [6] H. Tjandrasa, S. Djanali, and F. X. Arunanto, "Feature Extraction Using Combination of Intrinsic Mode Functions and Power Spectrum for EEG Signal Classification," *Int. Congr. Image Signal Process. Biomed. Eng. Informatics*, vol. 9, pp. 1498–1502, 2016.
- [7] L. Yu and H. Liu, "Feature Selection for High-Dimensional Data: A Fast Correlation-Based Filter Solution," *Int. Conf. Mach. Learn.*, pp. 1–8, 2003.
- [8] A. Zhigljavsky, H. Hassani, and S. Heravi, "Forecasting European Industrial Production with Multivariate Singular Spectrum Analysis," *Business*, pp. 1–39, 2009.
- [9] S. Kouchaki and S. Sanei, "Supervised Single Channel Source Separation Of EEG Signals," *IEEE Int. Mach. Learn. SIGNAL Process.*, 2013.
- [10] H. F. Posada-Quintero, J. P. Florian, A. D. Orjuela-Canon, T. Aljama-Corrales, S. Charleston-Villalobos, and K. H. Chon, "Power Spectral Density Analysis of Electrodermal Activity for Sympathetic Function Assessment," *Ann. Biomed. Eng.*, vol. 44, no. 10, pp. 3124–3135, 2016.
- [11] V. Marpaung, "Depresi Pada Penderita Epilepsi Umum Dengan Kejang Tonik Klonik Dan epilepsi Parsial Sederhana," pp. 1–25, 2005.
- [12] Y. H. Mahendra and H. Tjandrasa, "Klasifikasi data eeg untuk mendeteksi keadaan tidur dan bangun menggunakan autoregressive model dan support vector machine," *JUTI*, vol. 15, pp. 35–42, 2017.
- [13] A. K. Maddirala and R. A. Shaik, "Removal of EMG Artifacts from Single Channel EEG Signal using Singular Spectrum Analysis," *IEEE Sens. J.*, vol. 16, no. 23, pp. 111–115, 2015.
- [14] V. Moskvina and K. M. Schmidt, "Approximate projectors in singular spectrum analysis," *Siam J. Matrix Anal. Appl.*, vol. 24, no. 4, pp. 932–942, 2003.
- [15] D. Titisari, I. Soesanti, and B. Winduratna, "Reduksi Suara Jantung Dari Instrumentasi Akuisisi Perekaman Suara Paru-Paru Pada Anak-Anak Menggunakan Butterworth Band Pass Filter," in *SEMINAR NASIONAL ke 8 Tahun 2013 : Rekayasa Teknologi Industri dan Informasi REDUKSI*, 2013, pp. 129–134.
- [16] P. J. R. and A. Wijayanto, "Identifikasi Sinyal Suara Paru Berdasarkan Power Spectra Density Metode Welch Untuk Deteksi Kelainan Parenkim Paru."
- [17] K. K. Parhi and M. Ayinala, "Low-complexity welch power spectral density computation," *IEEE Trans. Circuits Syst. I Regul. Pap.*, vol. 61, no. 1, pp. 172–182, 2014.
- [18] D. Palaz, M. Magimai.-Doss, and R. Collobert, "Convolutional Neural Networks-Based Continuous Speech Recognition Using Raw Speech Signal," *Icassp*, pp. 4295–4299, 2015.
- [19] A. Levinskis, "Convolutional neural network feature reduction using wavelet transform," *Elektron. ir Elektrotehnika*, vol. 19, no. 3, pp. 61–64, 2013.