

Rancang Bangun Klasifikasi Sinyal Elektroensefalografi Pada Kasus Epilepsi Dengan Random forest GSO

Siswandari Noertjahjani

Program Studi Teknik Elektro, Universitas Muhammadiyah Semarang

Email: siswandari@unimus.ac.id

INFORMASI ARTIKEL

Histori artikel:

Naskah masuk, 1 September 2023

Direvisi, 20 Desember 2023

Diiterima, 31 Desember 2023

Kata Kunci:

Epilepsi,

Normal,

Shannon,

Norm

ABSTRAK

Abstract-

There are many studies using EEG signal analysis in detecting epilepsy. This research proposes an automatic detection framework using Matlab for epilepsy based on a frequency analysis approach using DWT sym 4 EEG signals at electrodes F4 - F8 which are decomposed into 5 frequency subbands D1, D2, D3, D4, D5 with 1 second segmentation and 256 Hz sampling frequency. Each subband is then correlated to get maximum results. We extract features using norm entropy, energy log entropy and Shannon entropy of each subband. To increase maximum detection performance, the classification used is Random Forest GSO by recognizing two levels of conditions, namely epilepsy and normal. Our proposed model classifies 100 epilepsy data and 100 normal data, and multiple cross-validation is adopted to improve the modeling accuracy. Experimental results were evaluated based on accuracy, sensitivity and specificity. This chart has very significant potential in clinically diagnosing epilepsy with computer assistance so that it can help doctors in making the correct diagnosis.

Abstrak-

Banyak sekali penelitian dengan analisis sinyal EEG dalam mendeteksi adanya penyakit epilepsi. Penelitian ini mengusulkan kerangka kerja deteksi otomatis dengan matlab untuk penyakit epilepsi berdasarkan pendekatan analisis frekuensi dengan menggunakan DWT sym 4 sinyal EEG pada elektroda F4 - F8 yang didekomposisi kedalam 5 subbands frekuensi D1, D2, D3, D4, D5 dengan segmentasi 1 detik dan frekuensi sampling 256 Hz. Tiap tiap subband kemudian di korelasikan untuk mendapatkan hasil yang maksimal. Kami mengekstrak ciri dengan menggunakan entropi norm, entropi log energi dan entropi Shannon tiap subband. Untuk meningkatkan kinerja deteksi maksimal, klasifikasi yang digunakan adalah dengan Random forest GSO dengan mengenali dua tingkat kondisi yaitu epilepsi dan normal. Model yang kami usulkan mengklasifikasikan 100 data epilepsi dan 100 data normal, dan beberapa validasi silang diadopsi untuk meningkatkan akurasi pemodelan. Hasil eksperimen dievaluasi berdasarkan akurasi, sensitivitas dan spesifitas. Bagan ini mempunyai potensi yang sangat signifikan dalam mendiagnosis secara klinis penyakit epilepsi dengan bantuan computer sehingga dapat membantu dokter dalam menegakkan diagnosis dengan tepat.

Copyright © 2023 LPPM - STM IKMI Cirebon
This is an open access article under the CC-BY license

Penulis Korespondensi:

Siswandari Noertjahjani

Program studi Teknik Elektro,

Universitas Muhammadiyah Semarang

Jl. Kedungmundu Raya No. 18 Semarang, 50273

Email: siswandari@unimus.ac.id

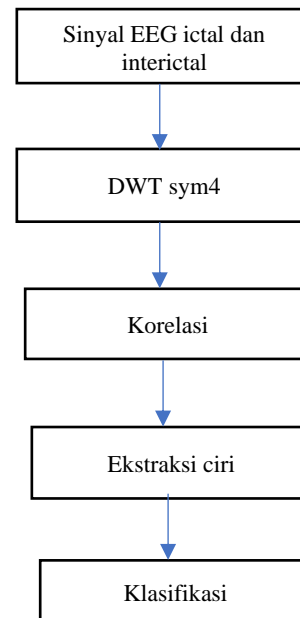
1. Pendahuluan

Penyakit epilepsi merupakan tantangan besar bagi dunia medis. Aktivitas saraf di otak manusia dimulai pada tahap awal perkembangan prenatal. Electroencephalogram (EEG) berisi informasi fisiologis penting yang dapat direfleksikan aktivitas otak manusia, membantu dalam mendeteksi epilepsi dan dapat mendiagnosis penyakit epilepsi. EEG juga merupakan alat untuk mengukur aktivitas otak dengan menempatkan sensor (elektroda) di kulit kepala[1][2]. Dengan kemajuan teknologi, banyak aplikasi telah dikembangkan yang memungkinkan klinik kesehatan menggunakan EEG untuk evaluasi diagnostik berbagai gangguan otak, termasuk penyakit alzheimer, gangguan tidur, trauma, cedera otak dll. Seorang spesialis EEG memeriksa bentuk gelombang dan menghasilkan laporan diagnostik yang menunjukkan kondisi pasien. Biasanya, langkah pertama dalam diagnosis adalah memutuskan apakah sinyal otak yang terekam menunjukkan kelainan perilaku[3]. Kelainan otak yang terkait dengan kelainan tersebut kemudian dapat dipelajari lebih lanjut dan obat-obatan yang diperlukan dapat diberikan.

Penggunaan klinis EEG meningkat pesat karena kemudahan penggunaannya sebagai prosedur non-invasif. Namun, interpretasi EEG secara manual adalah proses yang memakan waktu dan tenaga. Selain itu, untuk menjadi ahli di bidang tersebut, dokter perlu memiliki masa pelatihan yang sangat lama. Oleh karena itu, terdapat kebutuhan untuk mengembangkan prosedur interpretasi EEG secara otomatis untuk menghemat waktu dokter dan membantu mereka membuat keputusan yang akurat. Baru-baru ini, analisis EEG secara otomatis telah menarik perhatian karena ketangguhan algoritma pembelajaran mesin, kemajuan dan rendahnya biaya komputasi dengan kinerja tinggi. Selain itu, analisis sinyal EEG memiliki kesepakatan antar penilai (IRA) yang rendah, bahkan di antara para ahli yang bersertifikasi tinggi[4][5]. Penelitian mengenai deteksi epilepsi telah dilakukan dengan menggunakan model hybrid berdasarkan analisis Multi-Resolusi dan Penyaringan Adaptif [MRAF]. Tahap pertama, sinyal EEG didekomposisi menggunakan Transformasi wavelet diskrit (DWT) untuk menentukan daerah epilepsi. Tahap kedua pada frekuensi rendah dengan artefak fisiologis dihilangkan dengan pemfilteran adaptif. Hasilnya metode ini secara efektif menghilangkan efek fisiologis dan mempunyai akurasi sebesar 86,66 % [6][7]. Deteksi dini epilepsi dan diagnosis epilepsi juga dilakukan dengan model ARMA (Model autoregressive moving average) dengan kekuatan sinyal yang dihitung pada setiap titik data EEG dan klasifikasi yang digunakan RF-GSO dengan akurasi yang didapat 94% . Penelitian EEG epilepsi dengan

adaptive short-time Fourier transform-based synchrosqueezing transform akurasi mencapai 98,81% [8]. Deteksi epilepsi dengan *Gaussian Filter* dan klasifikasi *Fuzzy RBFELM* mencapai 98,48% [9]

2. Metode



Gambar 1. Blok diagram penelitian

Pada gambar 1 dijelaskan tentang data EEG, DWT sym 4, korelasi, ekstraksi ciri dan klasifikasi.

2.1. Pemilihan data

Perekaman EEG yang diambil dari kulit kepala menggunakan frekuensi sampling 256 Hz dan penempatan elektroda Internasional 10-20. Elektroda yang digunakan dalam penelitian ini ada F4,F8 data dengan segmentasi 1 detik. Data terdiri dari 100 normal dan 100 epilepsi.

2.2. DWT

Bagian Analisis transformasi wavelet diskrit (DWT) dapat mempertimbangkan fitur-fitur ini sinyal non stasioner. Rentang frekuensi efektif yang diperoleh setelah bandpass filter adalah 0 hingga 50 Hz [10][11].

DWT menguraikan sinyal EEG kulit kepala pada waktu t , yang dinyatakan sebagai:

$$s(t) = (s_1(t), s_2(t), \dots, s_n(t))^T =$$

$$\begin{pmatrix} s_{11} & \dots & s_{1p} \\ \vdots & \ddots & \vdots \\ s_{n1} & \dots & s_{np} \end{pmatrix}, (i = 1, 2, \dots, n) \quad (1)$$

dimana $s(t)$ melambangkan sinyal EEG $n \times p$, dengan $k = 1, 2, \dots, n/2$, dimana $s(t)$ melambangkan $n \times p$ sinyal EEG, $k = 1, 2, \dots, n/2$, $s_n(t)$ adalah saluran EEG ke- n , dan p adalah jumlah skala, ke dalam

wavelet, $\omega_{p,k}(t)$ dan penskala $\varphi_{p,k}(t)$ fungsi didefinisikan, masing-masing, sebagai berikut:

$$w_{p,k}(t) = 2^{-p/2} \omega(2^{-p} t - k) \quad (2)$$

$$\varphi_{p,k}(t) = 2^{-p/2} \varphi(2^{-p} t - k), p, k \in Z \quad (3)$$

DWT menghasilkan hierarki dekomposisi pada empat tingkat, seperti yang diilustrasikan pada Gambar 2. Pada masing-masing tingkat level, p , perkiraannya dapat dihitung menggunakan (4):

$$D_p(t) = \sum_{k \in Z} C(p, k) \omega_{p,k}(t) \quad (4)$$

Dimana,

$$C(p, k) = \int_{-\infty}^{\infty} s(t) \omega_{p,k}(t) dt \quad \text{menunjukkan koefisien wavelet.}$$

Sinyal EEG $s(t)$ dapat didefinisikan sebagai jumlah dari semua koefisien detail.

Pada level P , koefisien detail dengan eksponen, $p \leq P$, mewakili detail halus, sedangkan koefisien dengan $p > P$ mewakili detail kasar. Detail kasar juga disebut pendekatan (aproksimasi), $A_p(t)$ pada sinyal, $s(t)$, dan didefinisikan sebagai $A_p(t) = \sum_{p > P} D_p(t)$. Ketika level P meningkat, resolusinya 2^{-p} berkurang dan $A_p(t)$ hanya berisi fitur frekuensi rendah EEG[12]. Dalam penelitian ini, wavelet Symlet digunakan untuk mengekstraksi pita frekuensi tertentu dari sinyal EEG. Wavelet Symlet merupakan perbaikan pada wavelet Daubechies, yang mengatasi kelemahannya dari perkiraan asimetri yang ada dalam wavelet Daubechies. Nilai rentang dukungan $2N-1$ dan momen hilang wavelet Symlet N . Dasar gelombang Symlet memiliki keteraturan yang lebih baik daripada Daubechies, dan ini dapat mengurangi distorsi fase dalam analisis dan rekonstruksi sinyal non-stasioner, seperti sinyal EEG

2.3 Korelasi

Hasil dari DWT dikorelasikan

2.4 Ekstraksi ciri

Sistem biologis, seperti otak, ditentukan oleh kompleksnya dinamika. Sinyal EEG, karena sifat fisiologisnya, bersifat nonlinier dan non-stasioner. Dengan demikian, teknik ekstraksi fitur seharusnya efisien untuk memperoleh informasi yang relevan dari sinyal yang tidak jelas hanya dengan observasi visual. Dalam konteks ini, entropi telah banyak digunakan sebagai ukuran dis urutan atau ketidakpastian sinyal EEG. Mengingat aktivitas otak sebagai a sistem, jika derajat gangguan sistemnya rendah, maka ada organisasi dalam sistem itu; idealnya, ketika sistem berada dalam urutan nilai entropinya nol. Karena sifat entropi yang mengukur derajat ketidakteraturan a sistem, beberapa bentuk entropi telah dikembangkan. Dalam karya ini, empat di antaranya yang digunakan: entropi Shannon, entropi Energi Log dan entropi norm.

entropi Shannon

$$E(x) = -\sum_j x_j^2 \log(x_j^2) \quad (5)$$

Log entropi energi

$$E(x) = \sum_j \log(x_j^2) \quad (6)$$

Entropi norm

$$E(x) = \sum_j |x_j|^p \quad (7)$$

Dimana $p \geq 1$ l^p norm.

Semua variasi entropi diperoleh dengan mengambil sinyal x dan perluasan koefisiennya pada beberapa basis ortonormal x_j .

2.5 Klasifikasi Random Forest

RF-GSO untuk Model Pembelajaran Mesin Random forest (RF) adalah cara yang efektif metode pembelajaran mesin terintegrasi yang dikombinasikan dengan pohon keputusan. Metode identifikasi RF cocok untuk data dimensi tinggi dan berjalan cepat. Namun, sejumlah besar hyper-parameter dihasilkan selama operasi, dan untuk mendapatkan nilai akurasi yang lebih tinggi dalam mengenali sinyal EEG epilepsi.

Metode *random forest* biasanya didasarkan pada pengalaman memilih parameter secara manual. Khususnya, jumlah *decision tree* dalam metode *random forest* mempunyai pengaruh yang besar dalam kinerja dan untuk kategori data yang berbeda maka jumlah *decision tree* juga berbeda. Ketika kinerja model menjadi optimal. Parameter dari metode *random forest* dipilih hanya berdasarkan pengalaman, dan model identifikasi *random forest* dengan kinerja terbaik biasanya tidak diperoleh. Makalah ini menggunakan GSO yang ditingkatkan untuk mengidentifikasi RF dengan komputer. Parameternya dioptimalkan, dan metode validasi silang dalam pembelajaran mesin lebih banyak digunakan secara efektif untuk menghindari masalah over-fitting dari acak yang dilatih model. Dalam penelitian ini, dataset langsung dibagi menjadi dua kategori orang sehat, dan pasien epilepsi yang diberi label "1" dan "0".

RF adalah varian lanjutan dari Bagging. Pertama, bootstrap sampel Z^* dipilih secara acak dari data latih. Kemudian diulang Kembali, selanjutnya mengambil data yang dipilih secara acak, *decision tree* T_b . Kedua, subset fitur M dipilih secara acak dari kumpulan fitur dari setiap node *decision tree*. Dalam proses pelatihan model, menggunakan algoritma CART pohon regresi pengenalan untuk membagi node, dan nilai GINI dari indeks GINI digunakan sebagai dasar dari simpul pemisahan. Contoh set pelatihan Z^* mengandung karakteristik yang berbeda, dan indeks GINI dari set pelatihan ini adalah:

$$GINI(k) = 1 - \sum_{i=1}^k p_i \quad (8)$$

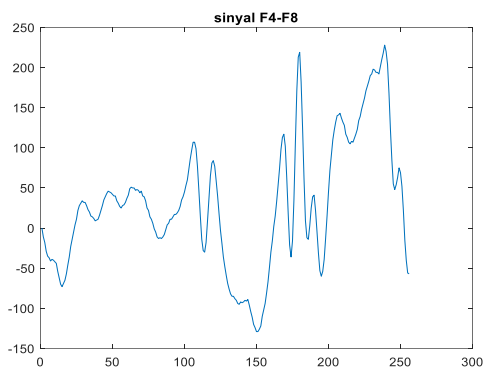
dimana, p_i merupakan probabilitas fitur kategori i . Nomor fitur yang sesuai dengan set pelatihan sampel adalah

$$(n_1, n_2, \dots, n_k), n = n_1 + n_2 + n_3$$

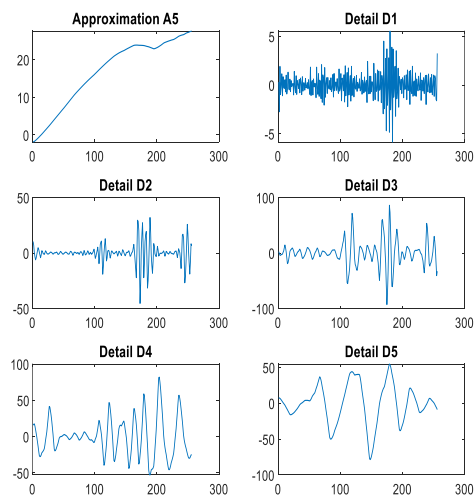
$$GINI(M^*) = \frac{n_1}{n} GINI(M_1) + \frac{n_2}{n} GINI(M_2) + \frac{n_3}{n} GINI(M_3) \quad (9)$$

3. Hasil dan Pembahasan

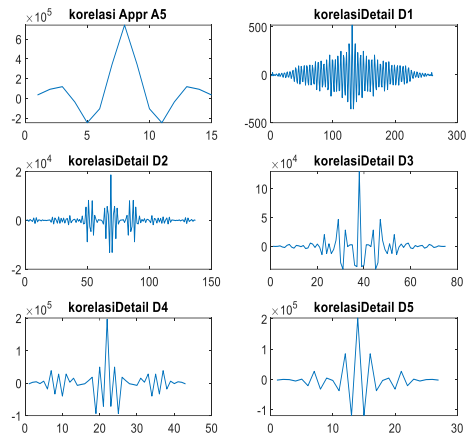
Pada gambar 2 menunjukkan sinyal EEG epilepsi superposisi F4-F8 dengan amplitudo lebih dari $200\mu V$. Pada gambar 3 menunjukkan hasil dekomposisi sinyal EEG epilepsi lima subband dengan D1, D2, D3, D4, D5 dan A5. Pada Gambar 4 menunjukkan hasil korelasi sinyal EEG epilepsi. Pada gambar 5 menunjukkan sinyal EEG normal superposisi F4-F8 dengan amplitudo maksimal $40\mu V$. Gambar 6 dan gambar 7 menunjukkan dekomposisi sinyal EEG normal dan korelasi.



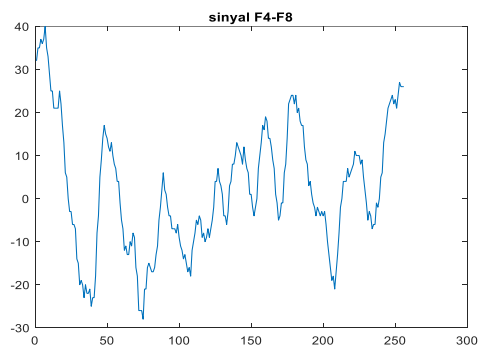
Gambar 2. sinyal EEG epilepsi F4-F8



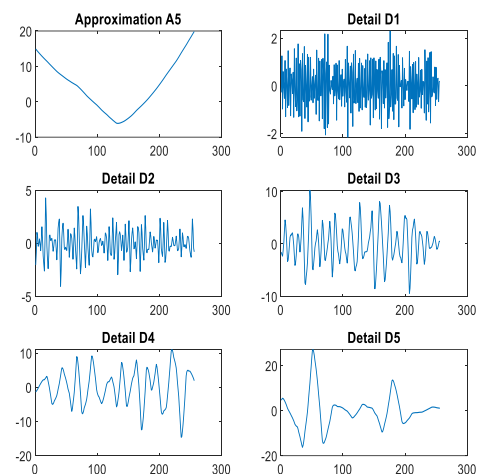
Gambar 3. Dekomposisi lima level sinyal EEG epilepsi F4-F8 menggunakan metode DWT symlet 4 (sym4).



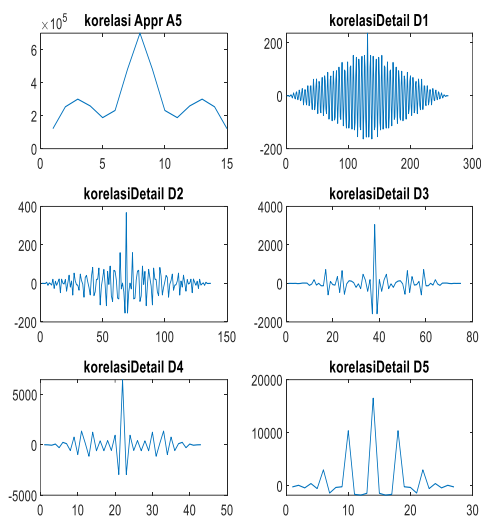
Gambar 4. Korelasi dari hasil Dekomposisi lima level sinyal EEG epilepsi F4-F8 menggunakan metode DWT symlet 4 (sym4).



Gambar 5 sinyal EEG normal F4-F8



Gambar 6. Dekomposisi lima level sinyal EEG normal F4-F8 menggunakan metode symlet 4 (sym4).



Gambar 7. Korelasi dari hasil Dekomposisi lima level sinyal EEG epilepsi F4-F8 menggunakan metode DWT symlet 4 (sym4). Kinerja klasifikasi dari algoritma RF yang diusulkan harus ditentukan berdasarkan hasil statistik semua pasien, hal ini untuk menghindari kemungkinan penyimpangan dari hasil deteksi epilepsi dari satu pasien. Masalah identifikasi dan pengambilan keputusan dalam klasifikasi data EEG epilepsi juga merupakan masalah klasifikasi yang tidak seimbang.

$$akurasi = \frac{TP+TN}{TP+TN+FP+FN} \quad (10)$$

$$spesitifitas = \frac{TN}{TN+FP} \quad (11)$$

$$sensitivitas = \frac{TP}{TP+FN} \quad (12)$$

dimana nilai true positif (TP) menandakan jumlah total yang benar kejadian normal dikenali dengan benar, tingkat negatif sebenarnya (TN) menunjukkan jumlah total kejadian sebenarnya dari epilepsi yang diidentifikasi dengan benar, dan positif palsu (FP) dan negatif palsu (FN) rate masing-masing adalah jumlah total false normal kejadian dan kejadian palsu epilepsi secara tidak benar. Keakuratan untuk klasifikasi yang benar adalah ditunjukkan pada tabel 2. Parameter ciri log entropi energi adalah parameter utama yang mempengaruhi kinerja pengklasifikasi RF baik D3,D4,D5 . Hasil uji yang ideal adalah yang mampu menghasilkan tingkat sensitivitas dan spesifitas yang tinggi. Akan tetapi kenyataannya semakin tinggi sensitivitas maka akan semakin kecil spesifitas. Begitu juga sebaliknya. Hal ini terjadi pada tabel 1 dan tabel 3. Sensitivitas berfungsi untuk mengukur proporsi positif dari subyek yang yang diidentifikasi dengan benar. Sensitivitas memperlihatkan kemampuan alat diagnostik dalam mendeteksi penyakit. Sedangkan spesifitas adalah mengukur proporsi negatif yang diidentifikasi dengan

benar. Pada tabel 1 hasil klasifikasi RF-GSO ciri entropi Shannon menunjukkan subband D2 mempunyai akurasi, sensitivitas, spesifitas dibawah 90 % . Sedangkan pada subband D3 akurasi naik menjadi 90 % . Pada ciri ini akurasi dan spesifitas maksimal terjadi pada subband D3 95% tapi sensitivitasnya turun 83%. Pada tabel 2 Hasil klasifikasi RF-GSO ciri log entropi memperlihatkan semua frekuensi band lebih tinggi daripada ciri entropi shannon dan entropi norm. Pada tabel 3 memperlihatkan nilai akurasi terbesar di D5. Deteksi epilepsi dengan DWT dan klasifikasi novel dynamic dengan akurasi 98,81% [13].

Tabel 1. Hasil klasifikasi RF-GSO ciri entropi Shannon

Band Frekuensi	Sensitivitas (%)	Spesifitas (%)	Akurasi (%)
D2	85	86	83
D3	83	95	95
D4	92	95	95
D5	90	92	90

Tabel 2. Hasil klasifikasi RF-GSO ciri Log entropi energi

Band Frekuensi	Sensitivitas (%)	Spesifitas (%)	Akurasi (%)
D2	85	87	87
D3	97	98	99
D4	95	95	96
D5	98	96	96

Tabel 3. Hasil klasifikasi RF-GSO ciri Entropi norm

Band Frekuensi	Sensitivitas (%)	Spesifitas (%)	Akurasi (%)
D2	88	86	83
D3	90	92	90
D4	92	95	92
D5	93	90	94

4. Kesimpulan

Paper ini mengusulkan metode klasifikasi epilepsi dengan Random Forest GSO menggunakan DWT symlet 4 dengan membagi sinyal menjadi 5 subband, dan dikorelasikan karena noise pada sinyal EEG terlalu tinggi maka akan mempengaruhi dalam pendeteksian epilepsi. Output dari korelasi diekstrak menggunakan entropi shannon, entropi log energi dan entropi norm. Hasil yang diperoleh klasifikasi RF-GSO dengan ciri entropi log energi mempunyai akurasi, sensitivitas, spesifitas yang tinggi pada D3. Kerangka deteksi otomatis kami menggunakan algoritma RF dan pengoptimal GSO dapat membantu diagnosis klinis untuk mendeteksi penyakit epilepsi dan membuat keputusan lebih cepat, akurat, dan secara efektif. Di masa depan, kami bermaksud untuk lebih mengoptimalkan model kami mencapai klasifikasi dan pengakuan berbagai tingkat serangan epilepsi. Hal ini juga harus diterapkan pada bidang medis lainnya

Daftar Pustaka

- [1] J. Cao *et al.*, “Biomedical Signal Processing and Control Using interictal seizure-free EEG data to recognise patients with epilepsy based on machine learning of brain functional connectivity,” *Biomed. Signal Process. Control*, vol. 67, no. August 2020, p. 102554, 2021, doi: 10.1016/j.bspc.2021.102554.
- [2] C. Li, C. Shao, R. Song, G. Xu, X. Liu, and R. Qian, “Spatio-temporal MLP network for seizure prediction using EEG signals,” *Measurement*, vol. 206, no. September 2022, p. 112278, 2023, doi: 10.1016/j.measurement.2022.112278.
- [3] B. Esmaeili *et al.*, “Clinical Neurophysiology Association between postictal EEG suppression, postictal autonomic dysfunction, and sudden unexpected death in epilepsy: Evidence from intracranial EEG,” *Clin. Neurophysiol.*, vol. 146, pp. 109–117, 2023, doi: 10.1016/j.clinph.2022.12.002.
- [4] A. Anuragi, D. Singh, and R. Bilas, “Automated FBSE-EWT based learning framework for detection of epileptic seizures using time-segmented EEG signals,” *Comput. Biol. Med.*, vol. 136, no. April, p. 104708, 2021, doi: 10.1016/j.compbiomed.2021.104708.
- [5] A. Zarei and B. M. Asl, “Automatic seizure detection using orthogonal matching pursuit, discrete wavelet transform, and entropy based features of EEG signals,” *Comput. Biol. Med.*, vol. 131, no. February, p. 104250, 2021, doi: 10.1016/j.compbiomed.2021.104250.
- [6] M. Varlı and H. Yılmaz, “Multiple classification of EEG signals and epileptic seizure diagnosis with combined deep learning,” *J. Comput. Sci.*, vol. 67, no. January, p. 101943, 2023, doi: 10.1016/j.jocs.2023.101943.
- [7] Z. Chen, G. Lu, Z. Xie, and W. E. I. Shang, “A Unified Framework and Method for EEG-Based Early Epileptic Seizure Detection and Epilepsy Diagnosis,” vol. 8, 2020.
- [8] M. Amiri, H. Aghaeinia, and H. R. Amindavar, “Biomedical Signal Processing and Control Automatic epileptic seizure detection in EEG signals using sparse common spatial pattern and adaptive short-time Fourier transform-based synchrosqueezing transform,” *Biomed. Signal Process. Control*, vol. 79, no. P1, p. 104022, 2023, doi: 10.1016/j.bspc.2022.104022.
- [9] A. Harishvijey and J. B. Raja, “Biomedical Signal Processing and Control Automated technique for EEG signal processing to detect seizure with optimized Variable Gaussian Filter and Fuzzy RBFELM classifier,” *Biomed. Signal Process. Control*, vol. 74, no. January, p. 103450, 2022, doi: 10.1016/j.bspc.2021.103450.
- [10] K. Ahmad, P. P. Shanir, Y. Uzzaman, and O. Farooq, “A hybrid Local Binary Pattern and wavelets based approach for EEG classification for diagnosing epilepsy,” vol. 140, 2020, doi: 10.1016/j.eswa.2019.112895.
- [11] D. Sunaryono, R. Sarno, J. Siswanto, D. Purwitasari, S. I. Sabilla, and R. Indarto, “Enhanced Gradient Boosting Machines Fusion based on the Pattern of Majority Voting for Automatic Epilepsy Detection,” vol. 13, no. 7, pp. 595–604, 2022.
- [12] Z. Telatar, “Biomedical Signal Processing and Control Automated temporal lobe epilepsy and psychogenic nonepileptic seizure patient discrimination from multichannel EEG recordings using DWT based analysis,” vol. 77, no. May, 2022, doi: 10.1016/j.bspc.2022.103755.
- [13] P. Datta *et al.*, “Biomedical Signal Processing and Control Automated EEG sentence classification using novel dynamic-sized binary pattern and multilevel discrete wavelet transform techniques with TSEEG database,” *Biomed. Signal Process. Control*, vol. 79, no. P1, p. 104055, 2023, doi: 10.1016/j.bspc.2022.104055.