

Klasifikasi Sinyal EEG Menggunakan Metode Fuzzy C-Means Clustering (FCM) Dan Adaptive Neighborhood Modified Backpropagation (ANMBP)

Dian Candra Rini Novitasari

Jurusan Matematika Universitas Islam Negeri Sunan Ampel, Surabaya.

Email: diancrini@uinsby.ac.id

Abstrak

Instrumen EEG (*electroencephalography*) merupakan suatu instrumen yang digunakan sebagai perekam aktivitas otak dengan memperlihatkan gelombang otak. Prinsip kerja EEG adalah dengan mendeteksi perubahan muatan secara tiba-tiba dari sel neuron yang ditandai dengan adanya interictal *spike-and-wave* pada hasil EEG (*electroencephalogram*). Terdapat suatu data set sinyal EEG, direkam pada sukarelawan normal dan epilepsi. Pada penelitian ini dengan menggunakan data tersebut akan dilakukan suatu sistem klasifikasi sinyal EEG dengan berdasar pada kondisi normal dan epilepsi.

Klasifikasi sinyal EEG menggunakan Metode *Adaptive Neighborhood Base Modified Backpropagation* (ANMBP). Hasil ekstraksi fitur dari sinyal EEG dengan menggunakan metode *Fuzzy C-Means* (FCM) *Clustering*, dimana proses awalnya melalui dekomposisi wavelet menggunakan *Discrete Wavelet Transform* (DWT) dengan level 2 didapatkan 3 koefisien wavelet kemudian pada masing masing koefisien tersebut di clustering menggunakan FCM dengan 2 cluster sehingga menghasilkan 6 fitur yang akan menjadi vektor fitur. Dari vektor fitur tersebut digunakan sebagai inputan untuk dilakukan proses klasifikasi dengan menggunakan metode ANMBP. Hasil sistem sementara didapatkan *recognition rate* sebesar 74.37%.

Kata kunci: EEG, Wavelet, FCM, Backpropagation, Modified.

Abstract

EEG (*electroencephalography*) is an instrument used to records brain activity and shows some brain waves. The working principle is to detect changes in EEG sudden charge of neuron cells characterized by interictal *spike-and-wave* on the EEG. There is a data set of EEG signals, recorded from normal volunteers and epilepsy. By using the data, we will build classification of EEG signals system and the classification system based on the EEG signals, normal and epilepsy conditions.

Classification of EEG signals is using *Adaptive Neighborhood Base Modified Backpropagation* (ANMBP) method. The result of EEG signals feature extraction, using *Fuzzy C-Means* (FCM) *Clustering*. Using wavelet decomposition level 2 and using FCM with 2 cluster each sub-band of wavelet coefficients then the probability distribution of each cluster in each sub-band is calculated. The result of probability distribution in each cluster of each sub-band is the feature vectors and will be inputed to the classification process by using ANMBP. The temporary result system gives the accuracy of 74.37%.

Keywords: EEG, Wavelet, FCM, Backpropagation, Modified.

1. Pendahuluan

Penyakit epilepsi adalah penyakit yang dapat terjadi pada siapapun walaupun dari garis keturunan tidak ada yang pernah mengalami epilepsi ini. Akan tetapi penyakit epilepsi tidak dapat menular. Epilepsi ini merupakan sebuah gangguan yang terjadi di sistem syaraf otak manusia yang disebabkan adanya aktifitas kelompok sel neuron yang terlalu berlebihan hingga akhirnya terjadi berbagai reaksi pada penderitanya. Epilepsi dapat terjadi karena lepasnya muatan listrik yang berlebihan dan mendadak pada otak sehingga penerimaan serta pengiriman impuls dari otak ke bagian-bagian lain dalam tubuh terganggu [10]. Untuk mendiagnosis penyakit epilepsi menggunakan pemeriksaan EEG. Karena instrumen electroencephalogram (EEG) dapat digunakan untuk perekaman yang menunjukkan aktivitas listrik otak, hal tersebut dapat memberikan suatu pengetahuan mengenai gangguan aktivitas otak. Dalam konteks ini, rekaman EEG diukur dalam interval bebas kejang dari pasien epilepsi dianggap sebagai komponen penting untuk proses diagnosis atau prediksi [2][6]. Meskipun terjadinya serangan epilepsi tampaknya tidak terduga [7], lebih banyak upaya yang difokuskan pada pengembangan model komputasi untuk deteksi otomatis debit epilepsi, yang kemudian dapat digunakan untuk memprediksi terjadinya kejang [2][3].

Banyak sistem diagnosis yang digunakan untuk mengklasifikasikan sinyal EEG untuk mengklasifikasikan antara epilepsi dengan kondisi normal [6][7][8].

Adaptive Neighborhood Base Modified Backpropagation (ANMBP), merupakan suatu metode klasifikasi yang sangat handal [1][9]. ANMBP ini memodifikasi metode *backpropagation*, dengan menggabungkan error linier dan non linier, adaptif *learning rate* serta *neighborhood* pada *backpropagation*.

Pada penelitian kali ini, akan dibuat suatu sistem klasifikasi, dimana menggabungkan antara *Discrete Wavelete Transform* (DWT), *Fuzzy C-Means Clustering* (FCM), dan Perhitungan Probabilitas sebagai proses ekstraksi fiturnya, dan menggunakan ANMBP sebagai metode klasifikasinya. Diharapkan dengan menggabungkan beberapa metode tersebut dapat menghasilkan sistem klasifikasi

yang handal yaitu membedakan kondisi normal atau epilepsi dengan menggunakan data EEG.

2. Data

Data sinyal EEG digital dapat diperoleh dari *database* yang tersedia di Universitas Bonn yang tersedia secara online dan dibuat oleh Dr. Ralph Andrzejak dari Pusat Epilepsi di Universitas Bonn, Jerman dan dapat di download atau di unduh dari link berikut:

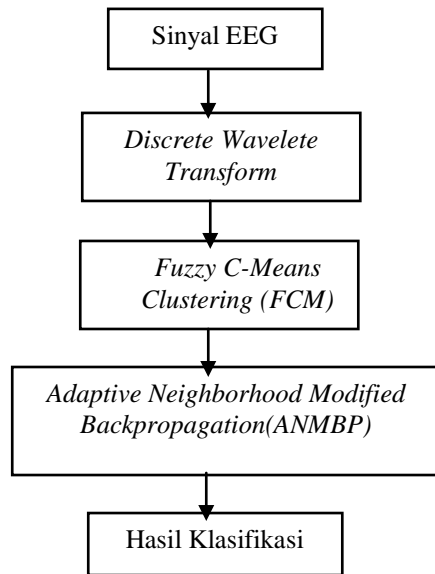
(http://epileptologie.bonn.de/cms/front_content.php?idcat=193&lang=3&changelang=3).

Selain itu, data sinyal EEG dalam bentuk digital dapat diperoleh di http://sccn.ucsd.edu/~arno/fam2data/publicly_available_EEG_data.html. Data sinyal EEG dari Universitas Bonn terdiri atas lima kelas dataset yaitu A, B, C, D, dan E. Tiap dataset berisi 100 segmen EEG saluran tunggal dengan durasi selama 23.6 detik. Setiap segmen dipilih dan dipotong dari rekaman EEG *multichannel* secara kontinyu setelah inspeksi artefak secara visual, misalnya gerakan mata atau aktivitas otot. Sebelum dilakukan proses ekstraksi fitur, sebelumnya dilakukan untuk *preprocessing* terhadap data yang akan digunakan dalam penelitian ini. Proses *preprocessing* pertama adalah memecah data menjadi segmen-segmen dengan ukuran 256 data dari 2 data set sinyal EEG, yaitu set A dan set E. Dari masing-masing kelas data, terdiri dari 100 segmen data dengan panjang masing-masing data 4097. Data tiap segmen sebanyak 4097 tersebut kemudian dilakukan *windowing* (pemotongan) dengan panjang 256. Hasil pemotongan sinyal didapatkan data yang lebih kecil sebanyak $4097/256 = 16$ segmen.

Jumlah segmen data keseluruhan yang diperoleh dari pemotongan sinyal EEG set A dan set E didapatkan data sebanyak $2 \times 100 \times 16 = 3200$ segmen data (ket: 2 merupakan kelas dari set A dan set E, 100 adalah jumlah masing-masing data awalnya, 16 merupakan segmen per 1 data awal). Jadi 3200 segmen data inilah yang nanti akan digunakan sebagai data *training* dan data *testing*.

3. Metodologi

Metode yang diterapkan dalam penelitian ini, dapat dilihat pada gambar flowchart pada gambar 1.



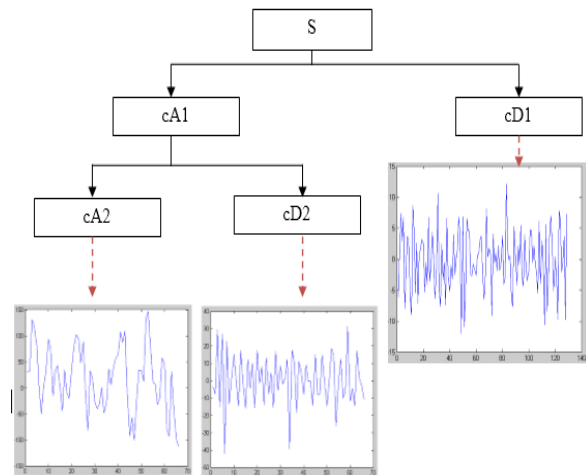
Gambar 1. Diagram alur sistem proses klasifikasi sinyal EEG

3.1 Ekstraksi Fitur

Proses ekstraksi fitur yang pertama dilakukan dengan menggunakan metode *Discrete Wavelet Transform (DWT)*, dimana proses awalnya DWT menguraikan sinyal ke dalam *sub-band* dengan menggunakan *low-pass filtering* dan *high-pass filtering* dari domain waktu sinyal. *Low-pass filtering* menghasilkan aproksimasi dan *high-pass filtering* menghasilkan koefisien detil. Dimana pada level selanjutnya dekomposisi dilakukan dengan menggunakan aproksimasi pada level sebelumnya. Pada penelitian ini tipe *wavelet* yang digunakan adalah *Daubechies Wavelet* order 2 (db2), hal tersebut didapatkan dari hasil penelitian sebelumnya.

Secara garis besar, dekomposisi membagi panjang data menjadi 2, yaitu berdasarkan *low-pass filtering* dan *high-pass filtering*. Dilakukan permisalan data mempunyai panjang 256, didekomposisikan pada level 1 berdasarkan *low-pass filtering* menghasilkan panjang data 128 yang di sebut aproksimasi 1 (A1) dan berdasarkan *high-pass filtering* menghasilkan panjang data 128 yang di sebut detil 1 (D1). Untuk melakukan dekomposisi pada level 2 maka menggunakan aproksimasi 1 dengan panjang 128, didekomposisikan *low-pass*

filtering menghasilkan panjang data 64 yang di sebut aproksimasi 2 (A2) dan berdasarkan *low-pass filtering* menghasilkan panjang data 64 yang disebut detil 2 (A2).

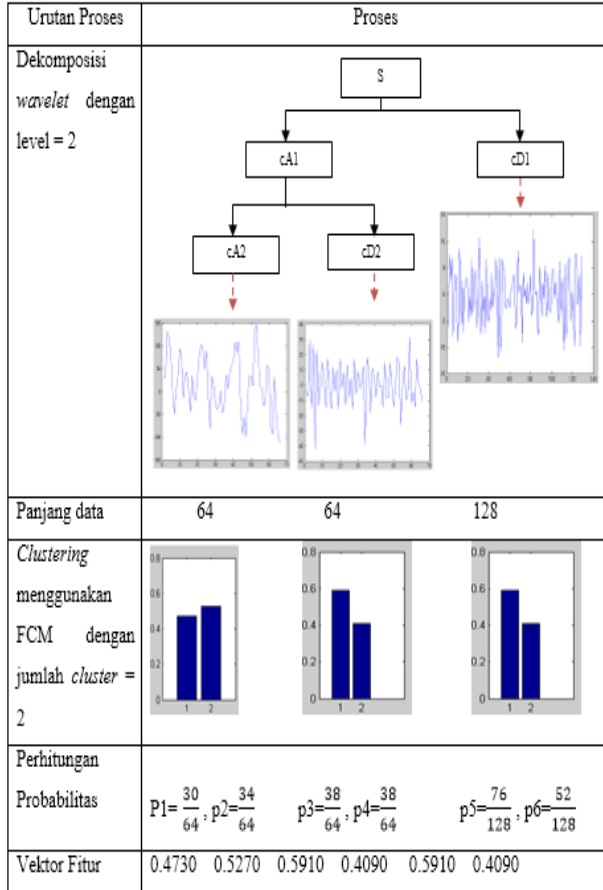


Gambar 2. Proses Dekomposisi sinyal EEG

Untuk proses selanjutnya, dilakukan *clustering* pada masing-masing koefisien *wavelet* menggunakan FCM dengan 2 *cluster*, jadi aproksimasi 2 dibagi ke dalam 2 *cluster*, detil 2 dibagi ke dalam 2 *cluster*, dan detil 1 dibagi ke dalam 2 *cluster*. FCM menggunakan fungsi keanggotaan untuk mencari kesamaan antara kumpulan data dengan *center*, menggunakan fungsi objektif dan mempartisi data masuk kedalam kluster-kluster hingga optimasi dari fungsi objektif tercapai. Dari hasil *clustering* menggunakan FCM ini didapatkan 6 buah *cluster*, selanjutnya masing-masing *cluster* dihitung nilai probabilitasnya. Proses ekstraksi fitur yang terakhir yaitu perhitungan probabilitas dari masing-masing *cluster*, karena aproksimasi 2 dari proses FCM menghasilkan 2 *cluster*, untuk perhitungan probabilitasnya dimisalkan P1 dan P2, P1 didapatkan dari jumlah koefisien *wavelet* yang masuk ke *cluster* 1 dibagi dengan panjang data aproksimasi 2 dan P2 didapatkan dari jumlah koefisien *wavelet* yang masuk ke *cluster* 2 dibagi dengan panjang data aproksimasi 2. Untuk detil 2, dari proses FCM menghasilkan 2 *cluster* untuk perhitungan probabilitasnya dimisalkan P3 dan P4, P3 didapatkan dari jumlah koefisien *wavelet* yang masuk ke *cluster* 1 dibagi dengan panjang data detil 2 dan P2 didapatkan dari jumlah koefisien *wavelet* yang masuk ke *cluster* 2 dibagi dengan panjang data detil 2. Hal tersebut juga berlaku untuk detil 1, dimana menghasilkan dua nilai probabilitas

dimisalkan P5 dan P6, P5 didapatkan dari jumlah koefisien wavelet yang masuk ke *cluster* 1 dibagi dengan panjang data detil 1 dan P6 didapatkan dari jumlah koefisien wavelet yang masuk ke *cluster* 2 dibagi dengan panjang data detil 1.

Proses ekstraksi fitur secara keseluruhan dapat dilihat pada gambar berikut ini:



Gambar 3. Contoh hasil ekstraksi fitur

3.2 Klasifikasi

Untuk klasifikasi metode yang digunakan adalah metode *Adaptive Neighborhood Modified Backpropagation* (ANMBP). Metode tersebut merupakan pengembangan dari metode backpropagation yaitu dengan menggabungkan dari penjumlahan error linier dan error non linier menggunakan *adaptive parameter learning* dan terdapat struktur *neighborhood* pada *hidden layer*.

Untuk mencari nilai net/jaringan digunakan rumus berikut:

$$u_j^s = \sum_{i=1}^n w_{ji}^s y_i^{s-1} \quad (1)$$

$$f(u_j^s) = \frac{1}{(1 + e^{-u_j^s})} = y_j^s \quad (2)$$

Dengan n menunjukkan jumlah neuron dan w_{ij}^s adalah bobot dari neuron ke i dari layer $(s-1)$ ke neuron ke j dari layer s .

Untuk menghitung error (E) yang didapatkan dari penjumlahan dari kuadrat error linier dan non linier yang dihasilkan dari output.

$$E_p = \sum_{j=1}^n \frac{1}{2} (e_{1j}^s)^2 + \sum_{j=1}^n \frac{1}{2} \lambda (e_{2j}^s)^2 \quad (3)$$

dimana λ adalah *learning rate*, error non-linier di simbolkan dengan e_1 dan error linier disimbolkan dengan e_2 yang diperoleh rumus:

$$e_{1j}^s = d_j^s - y_j^s \quad (4)$$

$$e_{2j}^s = ld_j^s - u_j^s \quad (5)$$

$$ld_j^s = f^{-1}(d_j^s) \quad (6)$$

Simbol d adalah output yang diinginkan dan y adalah output yang dihasilkan dari sistem. Sehingga perubahan bobot pada layer output menjadi seperti rumus berikut:

$$\Delta w_{ji}^s = -\mu \frac{\partial E}{\partial w_{ji}^s} \quad (7)$$

$$\Delta w_{ji}^s = \mu e_{1j}^s \frac{\partial y_j^s}{\partial w_{ji}^s} + \mu \lambda e_{2j}^s \frac{\partial u_j^s}{\partial w_{ji}^s}$$

$$\Delta w_{ji}^s = \mu e_{1j}^s \frac{\partial y_j^s}{\partial u_j^s} \frac{\partial u_j^s}{\partial w_{ji}^s} + \mu \lambda e_{2j}^s y_i^{s-1}$$

$$\Delta w_{ji}^s = \mu e_{1j}^s f'(u_j^s) y_i^{s-1} + \mu \lambda e_{2j}^s y_i^{s-1} \quad (8)$$

Error linier dan non linier pada hidden layer (L) adalah:

$$e_{1j}^L = \sum_{r=1}^{n_{L+1}} f'(u_r^{L+1}) e_{ir}^{L+1} w_{rj}^{L+1} \quad (9)$$

$$e_{2j}^L = f^{-1}(u_j^L) \sum_{r=1}^{n_{L+1}} e_{2r}^{L+1} w_{rj}^{L+1} \quad (10)$$

Sehingga perubahan bobot pada hidden layer adalah :

$$\Delta w_{ji}^L = \mu e_{1j}^L y_i^{L-1} f'(u_j^L) + \mu \lambda e_{2j}^L y_i^{L-1} \quad (11)$$

Parameter learning μ dan $\mu \lambda$ diganti dengan parameter adaptive η' dan μ'

$$\eta' = \frac{\mu \|e_{1j}\|^2}{\|f'(u_j) y_i^{s-1} e_{1j}\|^2 + \epsilon} \quad (12)$$

$$\mu' = \frac{\lambda \|e_{1j}\|^2}{\|f'(u_j) y_i^{s-1} e_{1j}\|^2 + \epsilon} \quad (13)$$

Dimana μ , λ adalah konstanta dengan nilai kecil positif dan ϵ konstanta dengan nilai kecil positif untuk menjamin adanya ketidakstabilan ketika error menuju 0.

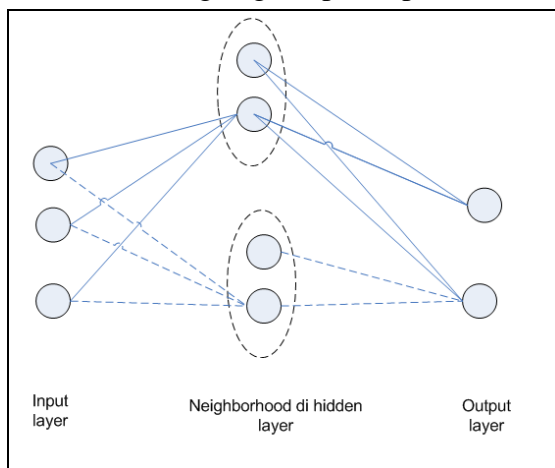
Sehingga perubahan bobot yang terjadi pada layer *output* dan *hidden layer* menjadi :

$$\Delta w_{ji}^s = \eta' e_{1j}^s f'(u_j^s) y_i^{s-1} + \mu' e_{2j}^s y_i^{s-1} \quad (14)$$

$$\Delta w_{ji}^L = \eta' e_{1j}^L f'(u_j^L) y_i^{L-1} + \mu' e_{2j}^L y_i^{L-1} \quad (15)$$

$$w(t+1) = w(t) + \Delta w(t) \quad (19)$$

Perubahan bobot hanya dilakukan pada bobot di node-node layer *hidden* dan node layer output yang merupakan *neighborhood* dari neuron-neuron yang terpilih di layer hidden. *Neighborhood* terlihat seperti gambar 4, dimana diilustrasikan dengan garis putus-putus.



Gambar 4. Struktur jaringan dengan neighborhood

4. Pembahasan Hasil

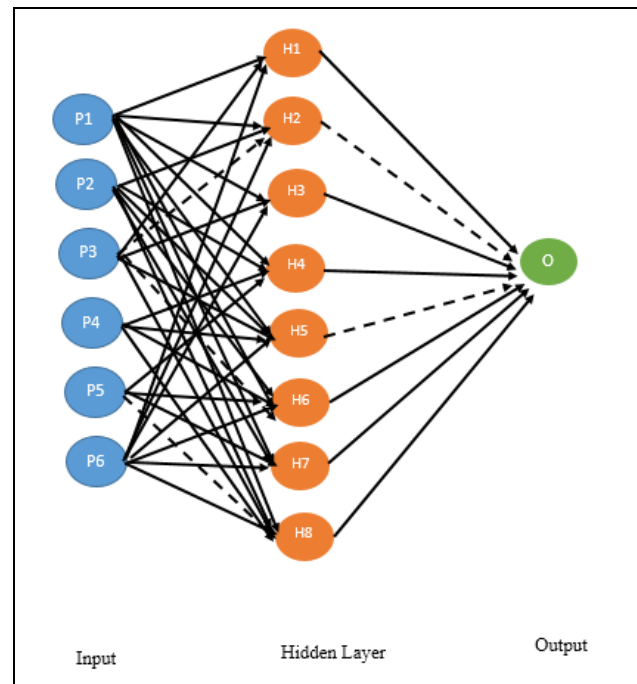
Dari proses ekstraksi fitur yang menghasilkan 6 fitur yang kemudian dijadikan satu menjadi vektor fitur digunakan sebagai inputan pada klasifikasi menggunakan ANMBP. Contoh hasil ekstraksi fitur dengan level dekomposisi waveletnya 2 dan kluster pada FCM ditentukan sebanyak 2 dapat dilihat pada gambar 5.

p1	p2	p3	p4	p5	p6	Target
0.473	0.527	0.591	0.409	0.591	0.409	Normal
0.558	0.442	0.561	0.439	0.561	0.439	Normal
0.527	0.473	0.470	0.530	0.470	0.530	Normal
0.504	0.496	0.409	0.591	0.409	0.591	Normal
0.620	0.380	0.364	0.636	0.364	0.636	Normal
0.395	0.605	0.470	0.530	0.470	0.530	Epilepsi
0.372	0.628	0.379	0.621	0.379	0.621	Epilepsi
0.341	0.659	0.348	0.652	0.348	0.652	Epilepsi
0.318	0.682	0.364	0.636	0.364	0.636	Epilepsi
0.333	0.667	0.303	0.697	0.303	0.697	Epilepsi

Gambar 5. Tabel Contoh Hasil Ekstraksi Fitur

Setelah didapatkan hasil ekstraksi fitur, vektor fitur tersebut yang digunakan sebagai masukkan / inputan pada proses klasifikasi menggunakan ANMBP. Untuk hasil klasifikasi

menggunakan ANMBP strukturnya dibuat seperti gambar 6.



Gambar 6. Jaringan ANMBP yang dibangun

Dengan menggunakan jaringan diatas, tidak mutlak menggunakan keseluruhan jaringan, karena bersifat adaptif maka tergantung *generate* dari parameter yang digunakan, sehingga garis putus-putus diatas, menunjukkan contoh bahwa ada beberapa net yang tidak dipakai.

Untuk training (pelatihan) dan testing (pengujian), maka diambil sampel secara random dari 3200 data set. Untuk kelas normal (set A) jumlah data setnya 1600, diambil 70% secara random, begitu pula dengan data set untuk kelas epilepsi (set E) yaitu sejumlah 1120 data set, sehingga total yang digunakan sebagai data training yaitu 2240. Selain data training, atau sisa 30% dari data set digunakan sebagai data testing (pengujian) yaitu sebanyak 960 data set.

Struktur jaringan adalah *input* sejumlah 6 didapatkan dari hasil ekstraksi fitur, *hidden layer* menggunakan satu layer dengan 8 node, Fungsi aktivasi yang digunakan adalah sigmoid biner. Untuk mencapai hasil yang optimal perlu diperhatikan pemilihan nilai konstanta μ , λ (*learning rate*) dan ϵ pada adaptif *learning rate*. Pemilihan nilai ini dilakukan secara heuristik. Pada penelitian ini nilai konstanta $\mu = 0.3$, $\lambda = 0.000001$, dan $\epsilon = 0.1$ serta bias = 0.5. Dengan menggunakan parameter tersebut

hasil satu kali training dan satu kali *testing* (uji data), dengan menggunakan parameter uji *recognition rate* berikut.

$$\text{Recognition rate} = \frac{\text{Jumlah terklasifikasi benar}}{\text{Jumlah Data}} \times 100,$$

didapatkan yang teruji benar adalah 714 dari 960 data, sehingga didapatkan akurasi sebesar 74.37%.

5. Kesimpulan

Dari sistem yang telah dibuat penulis masih jauh dari sempurna, dimana hasil *recognition rate* nya masih cukup rendah dibandingkan dua penelitian sebelumnya yaitu sebesar 74.37%. Pada penelitian sebelumnya bahkan ANMBP mampu sampai melebihi 90% untuk *recognition rate*-nya. Untuk penelitian selanjutnya perlu dilakukan uji coba dalam beberapa hal, antara lain:

1. Menggunakan jumlah level dekomposisi wavelet.
2. Penggunaan jumlah kluster, untuk metode FCM.
3. Parameter-parameter yang digunakan pada jaringan *adaptive neighborhood base modified backpropagation* (ANMBP) seperti jumlah *hidden layer*, *learning rate*, dan lain sebagainya.

Penelitian ini masih jauh dari kata sempurna, karena masih dilakukan satu kali *training* dan satu kali *testing* (pengujian) maka akan terus dilakukan pengujian data dengan menggunakan parameter-parameter yang ada baik pada proses ekstraksi fitur maupun klasifikasi yang ada.

Referensi

- [1] Werdiningsih, Indah. Transformasi Wavelet dan Adaptive Neighborhood Base Modified Backpropagation (ANMBP) untuk Klasifikasi Data Mamogram. Jurnal SCAN vol IX no 2. 2014.
- [2] Adeli H, Zhou Z, Dadmehr N. "Analysis of EEG records in an epileptic patient using wavelet transform". J Neurosci Methods; 123(1):69–87. Elsevier Science Ltd. 2003.
- [3] Elif Derya Übeyli. "Least squares support vector machine employing model-based methods coefficients for analysis of EEG signals", Expert Systems with Applications 37, 233–239. Elsevier Science Ltd. 2010.
- [4] Hekim, M., & Orhan, U. "Subtractive approach to fuzzy c-means clustering method". Journal of ITU-D, 10(1). 2011.
- [5] Orhan, U., Hekim, M., & Ozer, M. "Epileptic Seizure Detection Using Probability Distribution Based On Equal Frequency Discretization". Journal of Medical Systems. doi :10.1007/s10916-011-9689-y. Springer Ltd. 2011.
- [6] Subasi, A. EEG signal classification using wavelet feature extraction and a mixture of expert model. Expert Systems with Applications, 32, 1084–1093. 2007.
- [7] Subasi, A., & Ercelebi, E. Classification of EEG signals using neural network and logistic regression. Computer Methods and Programs in Biomedicine, 78, 87–99. Elsevier Science Ltd. 2005.
- [8] Novitasari, Dian Candra Rini., Klasifikasi Sinyal EEG Menggunakan Metode *Fuzzy C-Means (FCM) Clustering dan Adaptive Neuro Fuzzy Inference System (ANFIS)*. Undergraduate Thesis, Department of Information Technology, Faculty of Information Technology, Institut Teknologi Sepuluh Nopember, Indonesia, 2013
- [9] Werdiningsih, Indah. Sistem Diagnosis Kanker Payudara Menggunakan *Wavelete Transform dan Modified Backpropagation* Pada Data Mamogram. Undergraduate Thesis, Department of Information Technology, Faculty of Information Technology, Institut Teknologi Sepuluh Nopember, Indonesia, 2011
- [10] Penyakit Epilepsi. <http://penyakitpilepsi.com/>. Diakses tanggal 17 Agustus 2015.