

Farklı zaman ölçekli EEG işaretlerinden epilepsi nöbetinin otomatik tespiti

Mehmet YILDIRIM^{*1}, Abdulnasır YILDIZ²,

¹ Dicle Üniversitesi, Sivil Havacılık Yüksekokulu, Diyarbakır

² Dicle Üniversitesi, Elektrik Elektronik Mühendisliği Bölümü, Diyarbakır

Makale Gönderme Tarihi: 14.04.2017

Makale Kabul Tarihi: 25.05.2017

Öz

Epilepsi halk arasındaki adı ile sara kendini sürekli nöbetler ile tekrarlayan yaygın bir hastalıktır. Dünya nüfusunun yaklaşık olarak % 1'de görülen bu hastalık beynin bir bölümünde yahut tamamında meydana gelen ani, beklenmedik ve düzensiz elektriksel boşalma sonucu ortaya çıkan klinik bir durumdur. Beyinde bulunan sinir hücrelerinin elektriksel durum analizi anlamına gelen elektroensefalografi (EEG) epilepsinin tespiti için önemli bilgiler içermektedir. Bu sebeple EEG işaretleri birçok uzmanın ilgilendiği bir araştırma alanı haline gelmiştir.

Bu çalışmamızda Bonn Üniversitesi veri tabanından (A,B,C,D,E) alınan 23,6 saniye 4096 örnek uzunluğunda sağlıklı ve epilepsi nöbeti geçiren deneklerden alınan işaretlerden sadece A ve E işaret kümeleri kullanılarak gerçekleştirilen bir otomatik örüntü tanıma sistemi sunulmuştur.

Sunulan örüntü tanıma sistemi ön işlem, öznitelik çıkarım ve sınıflandırma olmak üzere üç aşamadan meydana gelmiştir. Birinci aşamada 23,6 saniye ve 4096 örnekten oluşan EEG işaretleri 128, 256, 512, 1024, 2048, 4096 uzunluğunda bölütlere ayrılmıştır. İkinci aşamada parametrik olmayan güç spektral yoğunluk (GSY) yöntemlerinden Peridogram ve Welch yöntemleri kullanılarak EEG işaretlerinin spektral bilgisi elde edilmiştir. Welch GSY kestirimi yapılırken her bir EEG işaret uzunluğunun dörtte biri uzunluğunda hamming penceresi kullanılmış ve parçaların örtüşme oranı %50 olarak seçilmiştir. GSY kestirimi yapıldıktan sonra veri boyutunu azaltmak için tüm EEG işaretlerine aritmetik ortalama uygulanmış ve her bir bölüt 16 örnek uzunluğunda öznitelik vektörü ile temsil edilmiştir. Üçüncü ve son aşamada her bir EEG bölütü için elde edilen ve 16 örnek uzunluğunki öznitelik vektörleri 5-katlı çapraz doğrulama yöntemi kullanılarak k en yakın komşu algoritması (k-NN), destek vektör makinesi (SVM), aşırı öğrenme makinesi (ELM) ile sınıflandırılmıştır. Tüm sınıflandırıcılar ile yapılan çalışmalarda maksimum %100 sonuç elde edilmiştir.

Anahtar Kelimeler: Epilepsi, Elektroensefalografi, Güç spektral yoğunluk, SVM, k-NN, ELM

Giriş

Epilepsi oldukça yaygın, ani ve tekrarlayan nöbetler ile karakterize nörolojik bir hastalıktır. Halk arasında sara olarak bilinen bu hastalık bio elektriksel bir ağ yapısı ile çalışan beynin bir bölümünde yahut tamında meydana gelen ani ve beklenmedik elektriksel aktivitenin oluşmasıyla ortaya çıkan klinik bir durumdur (Li vd., 2016). Bu hastalık nöbet dediğimiz kasılma, şuur kaybı, titreme, ağızda köpürme, kısa süreli göz dalması, ağız şapırdatılması, sorulara kısa süreli cevap verememe ve dalma gibi yaygın belirtiler ile kendini göstermektedir. Dünya nüfusunun yaklaşık olarak %1’de olduğu tespit edilen bu hastalık özellikle çocukluk ve yaşlılık döneminde yaygın olarak görülmektedir (Nicolaou ve Georgiou, 2012). İnsanların %9’nun ömründe en az bir kez epilepsi nöbeti geçirdiği uzmanlar tarafından ifade edilmekte fakat bu durum nöbeti geçiren insan tarafından fark edilememektedir (Toklu vd., 2012). Dünyada 40’ın üstünde epileptik nöbet çeşidi üzerinde araştırma yapılmış ve bu çalışmalar raporlanmıştır (Kaya vd., 2014).

Epilepsi şikayeti olan insanlarda nöbet geçirme sıklığı ve şiddeti insandan insana farklılık arz etmektedir. Bu sebeple hekim tarafından hastalık şiddetinin belirlenmesi, ona göre bir teşhis ve tedavi yönteminin geliştirilmesi oldukça önemlidir (Sezgin, 2016). Teşhis ve tedavi sürecinin başlatılıp devam ettirilmesinde kullanılan en önemli araçlardan biride EEG işaretleridir. Beyinde var olan sinir hücre gruplarının elektriksel aktivitesinin grafiksel gösterimi olan EEG işaretleri epilepsinin nedenlerinin araştırılması, yorumlanması, teşhis ve tedavi sürecinin geliştirilmesinde faydalı olacak değerli bilgiler taşır (Guo vd., 2010).

EEG işaretleri kullanılan diğer teknik ve yaklaşımlara göre daha hızlı, ucuz ve erişilebilir olduğundan araştırmacıların ilgi odağı haline gelmiştir (Nigam ve Graupe, 2004). EEG işaretlerinden farklı yaklaşımlarda çok sayıda otomatik tanı sistemi geliştirilmiştir. Dalgacık

dönüşümü, faz uzayı, kovaryans matrisinin hesaplanması ve M-ELM (Multiplicative extreme learning machine) kullanarak %100 doğruluk elde edilmiştir (Li vd., 2016). Farklılaşma, spektral analiz ve AR-MLPNN (Autoregressive -Multilayer perceptron neural network) kullanarak %99,97 doğruluk elde edilmiştir (Kang vd., 2015). Dalgacık dönüşümü ve LibSVM kullanarak %98,50 doğruluk elde edilmiştir (Nergiz vd., 2014). Dalgacık dönüşümü, istatistiksel yöntemler ve yapay sinir ağları kullanarak %99 doğruluk elde edilmiştir (Polat ve Özerdem, 2014). Dalgacık dönüşümü ve yapay sinir ağı kullanarak %100 doğruluk elde edilmiştir (Kumar vd., 2014).

Bu çalışmada sağlıklı kişilerden uyanık durumda gözler açıkken ve epilepsi hastalarında nöbet sırasında kaydedilmiş EEG işaretlerini sınıflandıran bir otomatik örüntü tanıma sistemi sunulmuştur. Kullanılan veri kümesi Bonn Üniversitesi veri tabanından alınmıştır (Andrzejak vd., 2001). Her bir veri kümesi 23,6 saniye uzunluğunda 4096 örnekten (sample) ibaret olup 100 adettir.

Burada sunulan örüntü tanıma sistemi ön işlem, öznitelik çıkarım ve sınıflandırma olmak üzere üç aşamadan meydana gelmiştir. Birinci aşamada 23,6 saniye ve 4096 örnekten oluşan veri kümeleri bölütlere ayrılarak farklı sayılarda, sürelerde ve uzunlukta EEG işaretleri elde edilmiştir. Bu bölütleme ile EEG işaretlerinin farklı uzunluklarda davranış biçimleri incelenmeye çalışılmıştır. İkinci aşamada elde edilen tüm EEG işaretleri için parametrik olmayan yöntemlerden Periodogram ve Welch GSY kestirim yöntemleri uygulanarak işaretlerin spektral bilgisine ulaşılmıştır. Farklı iki GSY kestirim yönteminin kullanılması ile veri bölütlerinin farklı spektrum kestirim yöntemlerindeki davranış biçimleri incelenmeye çalışılmıştır. GSY kestirimi yapıldıktan sonra veri boyutunu azaltmak amacı ile EEG işaretlerine aritmetik ortalama uygulanmış tüm işaretleri 16 örnek uzunluğundaki öznitelik vektörü ile temsil edilmiştir. Üçüncü ve son

aşamada tüm EEG işaretleri k-NN, SVM ve ELM sınıflandırıcıları ile sınıflandırılarak başarımları tablolar halinde değerlendirilmiştir. Bu sınıflandırma ile farklı veri bölütlerinin farklı GSY kestirimleri ile elde edilen öznitelik vektörlerinin farklı sınıflandırıcılardaki performansları incelenmeye çalışılmıştır. Başarımları performanslarının değerlendirmesinde 5 katlı çapraz doğrulama kullanılmıştır. Tüm sınıflandırıcılar ile yapılan çalışmalarda maksimum % 100 sonuç elde edilmiştir. Elde edilen bu sonuçların literatürdeki çalışmalar ile örtüştüğü görülmüştür.

Materyal ve Yöntem

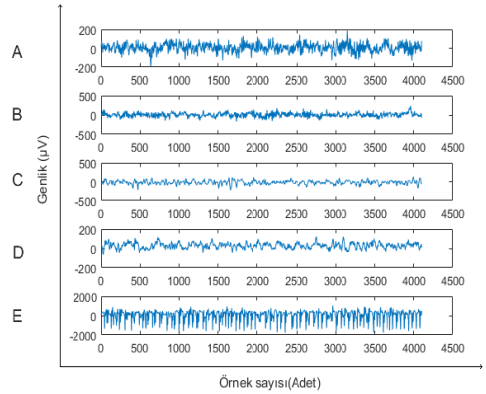
Veri Kümesi

Çalışmada kullanılan EEG işaretleri Bonn Üniversitesi veri tabanından alınmıştır. Tüm veri tabanı A, B, C, D ve E olmak üzere 5 veri kümesini içermektedir. Her bir küme 23.6 saniyelik, 100 tek-kanallı ölçümlerle kaydedilmiş EEG işaret parçalarından oluşmaktadır. A sağlıklı kişilerden uyanık durumda gözleri açıkken ve B ise uyanık durumda gözleri kapalıyken 10-20 uluslararası elektrot konumlandırma sistemine göre yerleştirilen yüzeysel elektrotlar yardımıyla ölçülen EEG işaretini kapsamaktadır. C işaret kümesi ise hasta deneklerden nöbet öncesinde hipokampal oluşumlu karşı yarım küreden alınarak kaydedilmiştir. D işaret kümesi hasta deneklerden nöbet olmayan durumda epileptojenik bölgeden alınarak kaydedilmiştir. E ise epilepsi hastalarından alınan ve nöbet sırasında kaydedilen EEG işaretlerini kapsar. Tüm EEG işaretleri, 12 bitlik analog-sayısal dönüştürücü kullanılarak, 173.61 Hz frekansında örneklenecek sayısal ortama kaydedilmiştir. Kaydedilen EEG işaretlerine 0.53-40 Hz bant geçiren filtre uygulanmıştır (Andrzejak vd., 2001). EEG işaretleri şekil 1'de gösterilmiştir.

Güç Spektral Yoğunluk Kestirimi

GSY kestirimi işaretlerin karakteristiklerinin frekans bandı üzerinde değerlendirilmesidir. Başka bir ifade ile işaretin frekans bandı

üzerindeki güç dağılımını gösteren bir ölçüdür. GSY kestirim yöntemleri parametrik olmayan yöntemler ve parametrik yöntemler olarak iki guruba ayrılır. Fakat bazı kaynaklarda alt uzay yöntemleri ayrı bir sınıf olarak değerlendirilmektedir. Parametrik olmayan yöntemler işaretin kendisini kullanarak GSY kestirimi yapan yöntemler olup az işlem yükü getirirler. Ancak parametrik yöntemlere göre spektrum bozulması sonucu zayıf işaretlerin maskelenmesi gibi dezavantajları vardır. Parametrik yöntemler sinyal modelleme sürecini kullanan ve fazla işlem yükü getiren GSY kestirim yöntemleridir. Literatürde Periodogram, Welch parametrik olmayan, Burg ve Yule Walker parametrik yöntemler arasında sıkça kullanılan yöntemlerdir (Türk vd., 2015).



Şekil 1. A,B,C,D ve E kümelerinden alınan EEG işaretlerinin örnek sayısı (adet) ve genlik (μV) gösterimi

a) Periodogram Güç Spektral Yoğunluk Kestirimi (PGSY)

1898 yılında İngiliz fizikçi Arthur Schuster tarafından ortaya atılan Periodogram parametrik olmayan GSY kestirim yöntemi olup Fourier dönüşümü temelli bir yöntemdir. Periodogram yöntemi ile GSY kestirimini elde etmek için işaret ikinin katları olacak şekilde 64, 128, 256'lı çerçevelere bölünür (Alkan, 2006).

N örnekten oluşan $x[n]$ işaretinin Periodogram GSY kestirimi işaretin Fourier dönüşümünün

karesi olup 1ve 2 numaralı denklemlerde ifade edilmiştir.

$$X(f) = \sum_{n=0}^{N-1} x[n] e^{-j2\pi fn} \quad (1)$$

$$P_{xx}(f) = \frac{1}{N} |X(f)|^2 \quad (2)$$

b) Welch Güç Spektral Yoğunluk Kestirimi (WGSY)

WGSY kestirim yöntemi PGSY yönteminin dezavantajlarını gidermek için geliştirilmiş bir yöntemdir. Bu yöntemde GSY'si kestirilen işaret birbirleri ile örtüşen bölütlere ayrılır. İşaretin kenarlarını düzleştirmek için her bölüt uygun bir pencere fonksiyonu kullanılarak pencerelenir. Pencerelenmiş her bölüt için PGSY kestirimi alınıp tahmini ortalaması hesaplanır. WGSY kestirimi iyileştirilmiş PGSY'lerin ortalaması alınarak bulunur (Koçak vd., 2016).

N örnekten oluşan işaret örtüştürülmüş K uzunluğunda L tane veri bölümüne ayrılır.

$$x_i = (x + is), n = 0, 1, \dots, K-1 \quad (3)$$

ve

$$i = 0, 1, \dots, L-1$$

3 numaralı denklemde is i. veri bölümünün başlangıç noktasıdır.

Daha sonra örtüştürülmüş her bir bölüt uygun bir w(n) pencere fonksiyonu ile pencerelenerek iyileştirilmiş PGSY'si alınır.

$$P_{xx}^{(i)}(f) = \frac{1}{K.T} \left| \sum_{n=0}^{K-1} x_i(n) w(n) e^{-j2\pi fn} \right|^2 \quad (4)$$

Denklem 4'teki T pencere fonksiyonu için güç olup 5 numaralı denklemde verilmiştir.

$$T = \frac{1}{K} \sum_{n=0}^{K-1} w^2(n) \quad (5)$$

L tane veri bölümüne ait iyileştirilmiş PGSY'lerin ortalaması alınarak N örnek uzunluğundaki

işaretin WGSY kestirimi denklem 6 ile elde edilir.

$$P^W = \frac{1}{K} \sum_{i=0}^{K-1} P_{xx}^i(f) \quad (6)$$

WGSY yöntemi işaret gürültü oranı düşük olduğunda PGSY yöntemine göre daha iyi bir çözünürlük sağlamaktadır (Proakis ve Manolakis, 1996).

Aritmetik Ortalama

Aritmetik ortalama, bir sayı dizisindeki elemanların toplamının dizinin eleman sayısına bölünmesi sonucu elde edilen değerdir.

$$\text{Aritmetik Ortalama} = \frac{\sum_{i=1}^n x_i}{n} \quad (7)$$

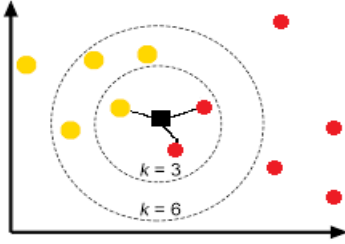
Sınıflandırma ve Başarı Ölçütü

Sınıflandırma iki veya daha fazla türü, nesneyi, olayı, örüntüyü vb. tanıma, tanımlama, kategorize etme veya birbirinden ayırma işlemidir. Bu işlem yapılırken çoğunlukla bir eğitim sonucu elde edilen bilgi ve deneyim gelecekteki problemlerin çözümü için kullanılır. Daha hızlı işlem yapmaları ve büyük miktarlardaki veriyi kolaylıkla işledikleri için sınıflandırma problemlerinde bilgisayarlar tarafından işletilen sınıflandırma algoritmaları yaygın bir şekilde kullanılmaktadır.

Bu çalışmada sunulan örüntü tanıma sisteminde k-NN, SVM ve ELM sınıflandırıcıları kullanılmıştır.

a) k En Yakın Komşu Algoritması (k-NN)

k-NN algoritması 1967 yılında T. M. Cover ve P. E. Hart tarafından ortaya atılmış denetimli bir öğrenme algoritmasıdır (Cover ve Hart, 1967). k-NN algoritması çok basit ve kolaydır.

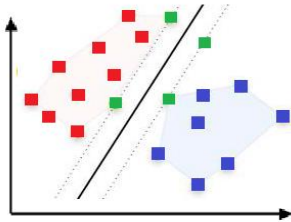


Şekil 2. k-en yakın komşuluk

Test örneğinin k adet en yakın komşuluklarını bulmak için test örneği ile eğitim örnekleri arasındaki uzaklıklar dikkate alınır. Test örneğinin en yakın komşuları bulunduktan sonra bu komşulardan en çok olan hangi sınıfa ait ise test örneği bu sınıfa dahil edilir. k-NN algoritmasında, Minkowski, Öklid, Manhattan, Chebyshev ve Dilca gibi uzaklık ölçüleri kullanılmaktadır. k-NN algoritmasında k parametresinin uygun seçilmesi çok önemlidir. k parametresinin küçük seçilmesi benzer özellikteki verileri bir sınıfta birleştirmekte faydalı bir etken iken bu parametrenin büyük seçilmesi farklı özellikteki verileri aynı sınıfta birleştirerek yanlış sınıflandırmalara sebebiyet verebilmektedir (Mitchell, 1997).

b) Destek Vektör Makinesi (SVM)

SVM sınıflandırma problemlerinde yaygın olarak kullanılan, basit, etkili ve denetimli bir öğrenme algoritmasıdır. Boser, Guyon ve Vapnik tarafından 1992’de ortaya atılmış ve pratikte çok iyi sonuçlar verdiği için 1990’ların sonlarına doğru yaygın olarak kullanılmaya başlanmıştır (Moavenian ve Khorrami, 2010).

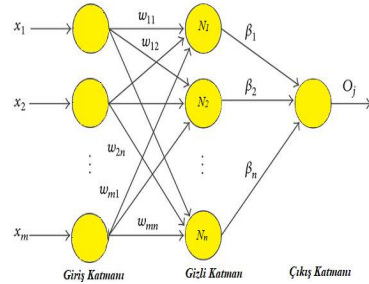


Şekil 3. Destek vektör makinesi için iki sınıf problemi

SVM’de amaç sınıfları bir birinden ayıracak bir hiperdüzlem elde etmektir. Bu ayırma esnasında sınıfları birbirinden ayıran pek çok hiperdüzlem bulunacağından asıl amaç bu hiperdüzlemin her iki sınıfa en uzak mesafede olanını bulmaktır. Bu hiperdüzleme en yakın vektörleri de destek vektörleri olarak adlandırılmaktadır. Eğitim verilerinden elde edilen destek vektörleri sınıflandırmada araç olarak kullanılmaktadır. SVM sınıflandırılacak verinin doğrusal olarak ayrılma veya ayrılmama durumuna göre genelde iki kategoride değerlendirilir. Doğrusal olarak ayrılma durumu yukarıda sunulduğu gibi gerçekleştirilmektedir. Doğrusal olarak ayrılmama durumunda ise doğrusal olmayan bir dönüşüm ile çekirdek fonksiyonları kullanılarak verilerin doğrusal olarak ayrılacakları yüksek boyutlu bir özellik uzayına taşınması ile doğrusal olarak ayrılabilir duruma gelmesi sağlanmaktadır (Demir ve Ertürk, 2010).

c) Aşırı Öğrenme Makinesi (ELM)

ELM 2006 yılında Huang ve arkadaşları tarafından geliştirilen, verimliliği yüksek, uygulaması kolay, tek gizli katmanı olan ileri beslemeli bir yapay sinir ağı modelidir. Bu modelin gizli katmanda sigmoidal, sine, Gaussian ve hard-limitin gibi aktivasyon fonksiyonları kullanılırken çıkış katmanda doğrusal fonksiyon kullanılmaktadır.



Şekil 4. Aşırı öğrenme makinesi

ELM’nin matematiksel yapısı denklem 8’de verilmiştir.

$$\sum_{i=1}^L \beta_j(\omega_i \cdot x_j + b_i) = o_j, j = 1, 2, \dots, N \quad (8)$$

Burada x_j ağız girişini, o_j ağız çıkışını, w_i ağız giriş katman düğümleri ile gizli katman düğümlerini birleştiren ağırlık vektörünü, β_i ağız gizli katman düğümleri ile çıkış katman düğümlerini birleştiren ağırlık vektörünü, $g(.)$ aktivasyon fonksiyonunu göstermektedir (Huang vd., 2006; Sezgin, 2016).

Huang ve arkadaşları ELM'nin gizli katmanındaki düğüm seçim problemlerini gidermek için açık aktivasyon fonksiyonunu örtük çekirdek haritalamasına aktaran çekirdek sürümünü 2012 yılında önerdi. Geleneksel çekirdek yöntemlerinde olduğu gibi, ELM'nin temel problemi çekirdek fonksiyonlarının seçimi olmaktadır. Sık kullanılan çekirdek fonksiyonları arasında Rbf, Gauss ve polinom fonksiyonu bulunur (Huang vd., 2012).

d) Başarı Ölçütü

Sınıflandırıcı performansı çoğunlukla test performansının doğruluk, kesinlik, duyarlılık gibi istatistiksel parametrelerinin hesaplanması ile sunulmaktadır. Bu parametreler 9, 10, 11'inci denklemlerde verilmiştir.

$$\text{Doğruluk (Accuracy)} = \frac{GP+GN}{GP+YP+DN+YN} (\%) \quad (9)$$

$$\text{Kesinlik (Specifitcity)} = \frac{GN}{YP+GN} (\%) \quad (10)$$

$$\text{Duyarlılık (Sensitivity)} = \frac{GP}{GP+YN} (\%) \quad (11)$$

GP (gerçek pozitif), GN (gerçek negatif), YP (yanlış pozitif), YN (yanlış negatif) kavramları sırasıyla sağlıklı deneğin sağlıklı olarak belirlenmesi, nöbet geçiren deneğin nöbet geçiren denek olarak belirlenmesi, nöbet geçiren deneğin sağlıklı denek olarak belirlenmesi, sağlıklı deneğin nöbet geçiren denek olarak belirlenmesi şeklinde tanımlanırlar.

Önişlem

Önceki bölümlerde belirtildiği üzere çalışmada kullanılan A ve E veri kümeleri 23,6 saniye süreden 4096 adet örnekten oluşmaktadır. Bu veri kümelerinden her biri tablo 1'de gösterildiği sürelerde ve uzunlukta bölütlere ayrılarak farklı sayılarda, sürelerde ve uzunlukta EEG işaretleri elde edilmiştir.

Tablo 1. Çalışmada kullanılan veri kümelerinin bölüt yapısı

Bölüt Örnek Sayısı	Bölüt Süresi (Sn.)	İşaret Sayısı
128	0,7375	3200
256	1,4750	1600
512	2,9500	800
1024	5,9000	400
2048	11,8000	200
4096	23,6000	100

Öznitelik Çıkartma

Sınıflandırma işleminde giriş vektörlerinin seçimi gerçekten çok hassas ve önemli bir konudur. Girişler kötü ve gereksiz bilgi içerdiğinde bilinen en iyi sınıflandırıcı kullanılsa bile başarı performansı kötü olacaktır. Sınıflandırıcıların performansını artırmak, çalışma süresini ve donanım maliyetini azaltmak için veriyi temsil eden en iyi özellikler seçilmelidir. Veri kümesinde bulunan ilgisiz veya gereksiz özellikleri uzaklaştırmak, veri dağılımının daha uygun ve ayrıştırılabilir bir hale getirmek sınıflandırıcıların daha etkin kullanılmasında faydalı bir durumdur (Güler ve Übeyli, 2005).

Burada öznitelik çıkarımı EEG işaretlerinin spektrum bilgisi üzerinde aritmetik ortalama kullanılarak yapılmıştır.

Bu çalışmada ve tablo 1'de detaylı olarak gösterilen 6 adet EEG işareti için PGSY, WGSY kestirimleri ve aritmetik ortalama kullanılarak 16 örnek uzunluğunda 12 adet öznitelik vektörü çıkarılmıştır.

PGSY kestirim yöntemi ile öznitelik vektörleri elde etmek için tablo 1'de gösterilen ve örnek sayıları 128'den başlayıp 4096'da biten 6 adet EEG işaretinin PGSY kestirimleri alınarak işaretlerin spektrum bilgisine ulaşılmıştır. EEG işaretlerinin PGSY kestirimleri şekil 5'ten başlanarak şekil 10'a kadar gösterilmiştir. PGSY kestirimi sonucunda uzunlukları 128, 256, 512, 1024, 2048, 4096 olan işaretlerin öznitelik vektörleri sırasıyla 64, 128, 256, 512, 1024, 2048 olarak elde edilmiştir. Daha sonra bu öznitelik

vektörlerinin veri boyutunun azaltılması yoluna gidilmiştir.

Veri boyutunu azaltmak için örnek uzunluğu 64 olan öznitelik vektörünün her ardışık 4, örnek uzunluğu 128 olan öznitelik vektörünün her ardışık 8, örnek uzunluğu 256 olan öznitelik vektörünün her ardışık 16, örnek uzunluğu 512 olan öznitelik vektörünün her ardışık 32, örnek uzunluğu 1024 olan öznitelik vektörünün her ardışık 64, örnek uzunluğu 2048 olan öznitelik vektörünün her ardışık 128 örneğinin aritmetik ortalaması alınarak tek örnek ile temsil edilmiştir.

Bu şekilde uzunlukları 128, 256, 512, 1024, 2048, 4096 olan ve PGSY kestirimi sonucu öznitelik vektörleri 64, 128, 256, 512, 1024, 2048 uzunluğunda elde edilen EEG işaretlerinin boyutları 16 örnek uzunluğuna indirgenmiş olup verinin boyutu azaltılmıştır.

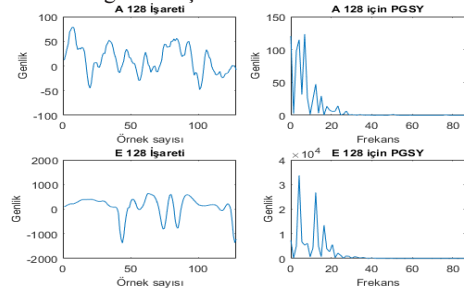
WGSY kestirim yöntemi ile öznitelik vektörleri elde etmek için tablo 1'de gösterilen ve örnek sayıları 128'den başlayıp 4096'da biten 6 adet EEG işaretinin WGSY kestirimleri alınarak işaretlerin spektrum bilgisine ulaşılmıştır. WGSY kestirimi alınırken her bir EEG işaret uzunluğunun dörtte biri uzunluğunda hamming penceresi kullanılmış ve bölütlerin örtüşme oranı %50 olarak seçilmiştir. EEG işaretleri için WGSY kestirimleri şekil 11'den başlanarak şekil 16'ya kadar gösterilmiştir. WGSY kestirimi sonucunda uzunlukları 128, 256, 512, 1024, 2048, 4096 olan işaretlerin öznitelikleri sırasıyla 16, 32, 64, 128, 256, 512 olarak elde edilmiştir. Daha sonra bu öznitelik vektörlerinin veri boyutunun azaltılması yoluna gidilmiştir.

Veri boyutunu azaltmak için örnek uzunluğu 16 olan öznitelik vektörünün her ardışık 1, örnek uzunluğu 32 olan öznitelik vektörünün her ardışık 2, örnek uzunluğu 64 olan öznitelik vektörünün her ardışık 4, örnek uzunluğu 128 olan öznitelik vektörünün her ardışık 8, örnek uzunluğu 256 olan öznitelik vektörünün her ardışık 16, örnek uzunluğu 512 olan öznitelik vektörünün her ardışık 32 örneğinin aritmetik

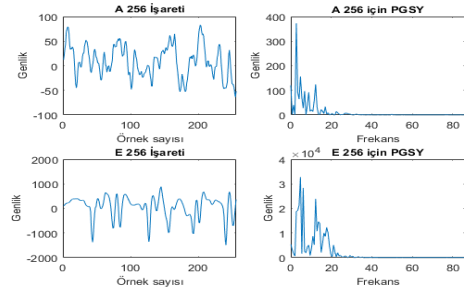
ortalaması alınarak tek örnek ile temsil edilmiştir.

Bu şekilde uzunlukları 128, 256, 512, 1024, 2048, 4096 olan ve WGSY kestirimi sonucu öznitelik vektörleri 16, 32, 64, 128, 256, 512 uzunluğunda elde edilen EEG işaretlerinin boyutları 16 örnek uzunluğuna indirgenmiş olup verinin boyutu azaltılmıştır.

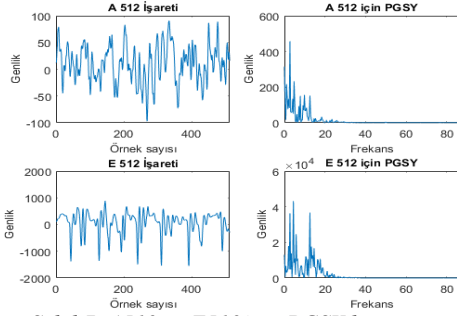
Öznitelik çıkarımından sonra 12 adet öznitelik vektörü sınıflandırıcılar ile sınıflandırılmaya hazır hale getirilmiştir.



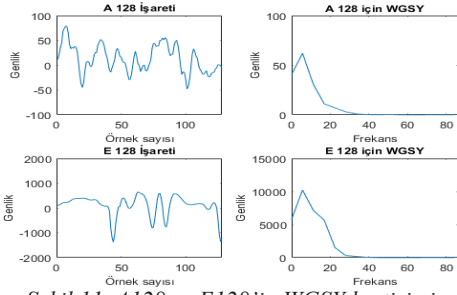
Şekil 5. A128 ve E128'in PGSY kestirimi



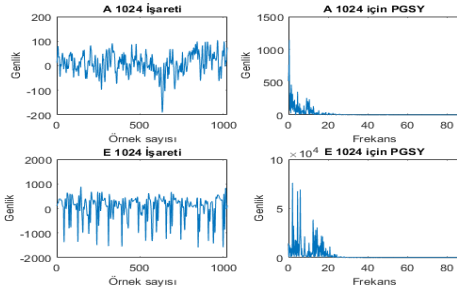
Şekil 6. A256 ve E256'nun PGSY kestirimi



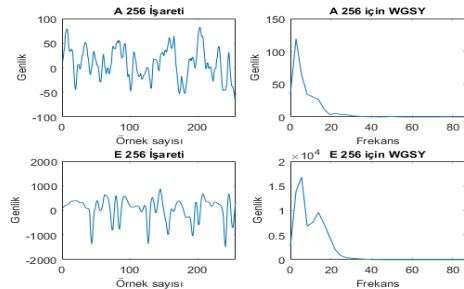
Şekil 7. A512 ve E512'nin PGSY kestirimi



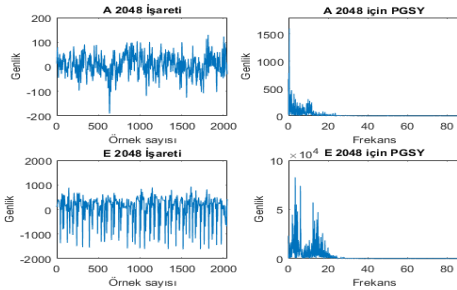
Şekil 11. A128 ve E128'in WGSY kestirimi



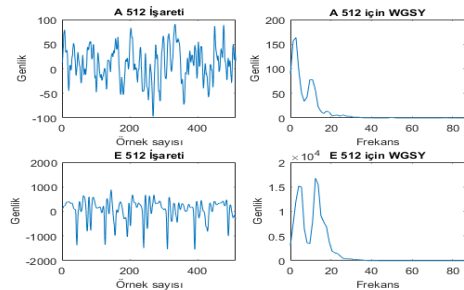
Şekil 8. A1024 ve E1024'ün PGSY kestirimi



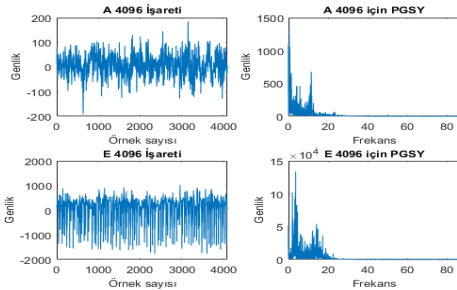
Şekil 12. A256 ve E256'nın WGSY kestirimi



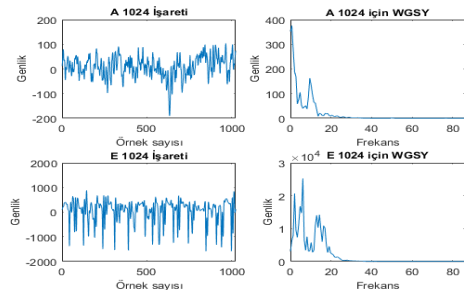
Şekil 9. A2048 ve E2048'in PGSY kestirimi



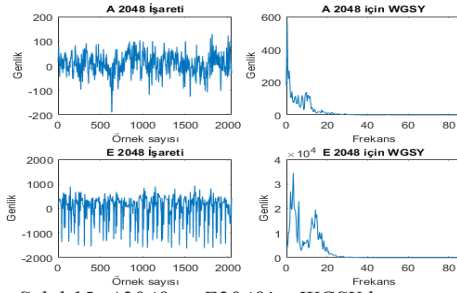
Şekil 13. A512 ve E512'nin WGSY kestirimi



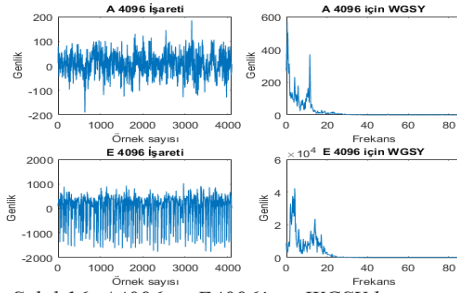
Şekil 10. A4096 ve E4096'nın PGSY kestirimi



Şekil 14. A1024 ve E1024'ün WGSY kestirimi



Şekil 15. A2048 ve E2048'in WGSY kestirimi



Şekil 16. A4096 ve E4096'nın WGSY kestirimi

Uygulama ve Başarımlar

Bu çalışmada yapılan tüm uygulamalar 5 katlı çapraz doğrulama yöntemi kullanılarak gerçekleştirilmiştir. Çalışmada kullanılan tüm veriler hem eğitim hem test için kullanılmıştır.

k-NN ile yapılan sınıflandırmada uzaklık ölçütü olarak Öklid uzaklık ölçütü kullanılmıştır. k-NN algoritmasında k parametresi 1 ile 100 aralığında seçilmiş her bir değer için sınıflandırma yapılarak maksimum başarımlar sonuçları detaylı olarak tablo 2'de gösterilmiştir.

DVM için kernel tipi radial base function (RBF) seçilmiştir. Eğitim sürecinde genişlik parametresi sigma (σ) 0,01'den başlanarak 0,01'lik artışlar ile 10 kadar, regülarizasyon parametresi (C) 1'den başlanarak 1'lik artışlar ile 100 kadar denenmiş elde edilen maksimum başarımlar sonuçları detaylı olarak tablo 3'te gösterilmiştir.

ELM için kernel tipi radial base function (RBF) seçilmiştir. Eğitim sürecinde genişlik parametresi sigma (σ) 0,001'den başlanarak 0,001'lik artışlar ile 10 kadar, regülarizasyon parametresi (C) 1'den başlanarak 1'lik artışlar ile 100 kadar denenmiş elde edilen maksimum başarımlar sonuçları detaylı olarak tablo 4'te gösterilmiştir.

Tablo 2. k-NN ile EEG işaretlerinin sınıflandırılması

GSY	Bölüt Boyutu	Komşu sayısı (k)	Doğruluk (Accuracy) (%)	Kesinlik (Specificity) (%)	Duyarlılık (Sensitivity) (%)
Periodogram	128	5	99,30	98,91	99,69
	256	3	99,66	99,38	99,94
	512	3	99,81	99,75	99,88
	1024	1	99,75	99,50	100
	2048	1	100	100	100
	1024	1	100	100	100
Welch	128	9	99,30	98,97	99,63
	256	3	99,72	99,44	100
	512	1	99,75	99,50	100
	1024	1	99,88	99,75	100
	2048	1	100	100	100
	4096	1	100	100	100

Tablo 3. SVM ile EEG işaretlerinin sınıflandırılması

GSY	Bölüt Boyutu	C katsayısı	RBF(σ) Değeri	Doğruluk (Accuracy) (%)	Kesinlik (Specificity) (%)	Duyarlılık (Sensitivity) (%)
Periodogram	128	2	0,01	99,41	98,97	99,84
	256	1	0,01	99,72	99,50	99,94
	512	1	0,01	99,88	99,75	100
	1024	1	0,01	100	100	100
	2048	1	0,01	100	100	100
	1024	1	0,01	100	100	100
Welch	128	8	0,01	99,24	99,00	99,47
	256	4	0,01	99,75	99,50	100
	512	8	0,01	99,88	99,75	100
	1024	1	0,01	100	100	100
	2048	1	0,01	100	100	100
	4096	1	0,01	100	100	100

Tablo 4. ELM ile EEG işaretlerinin sınıflandırılması

GSY	Bölüt Boyutu	C katsayısı	RBF(σ) Değeri	Doğruluk (Accuracy) (%)	Kesinlik (Specificity) (%)	Duyarlılık (Sensitivity) (%)
Periodogram	128	1	0,001	99,33	98,78	99,88
	256	4	0,001	99,72	99,50	99,94
	512	1	0,001	99,75	99,50	100
	1024	32	0,001	99,88	99,75	100
	2048	1	0,001	100	100	100
	1024	1	0,001	100	100	100
Welch	128	1	0,001	99,18	98,72	99,63
	256	4	0,001	99,72	99,44	100
	512	8	0,001	99,75	99,50	100
	1024	12	0,001	99,88	99,75	100
	2048	1	0,001	100	100	100
	4096	1	0,001	100	100	100

Sonuçlar ve Tartışma

Tablo 5'te detaylı olarak ifade edildiği üzere burada önerilen örüntü tanıma sisteminde farklı sınıflandırıcılar ile yapılan çalışma sonucunda elde edilen doğruluk performanslarının %99'un üzerinde olması tatmin edici bir durum olarak değerlendirilmiştir. Önerilen örüntü tanıma sisteminin performansı sınıflandırıcılar açısından değerlendirildiğinden en iyi performansın SVM ile elde edildiği görülmüştür. Sistemde kullanılan öznelik çıkarım yöntemleri açısından değerlendirildiğinde WGSY kestirim yönteminin PGSY kestirim yöntemine göre daha iyi sonuç verdiği görülmüştür. Sistemin başarımları

performansı kullanılan farklı EEG veri uzunlukları açısından değerlendirildiğinde ise veri uzunluğu artıka başarımların arttığı görülmüştür. Bu çalışmada kullanılan EEG işaretleri kullanılarak farklı öznelik çıkarım ve sınıflandırma teknikleri ile çok sayıda örüntü tanıma sistemi sunulduğu bilinmektedir. Literatürde sunulan örüntü tanıma sistemleri ile elde edilen sonuçlar yaptığımız çalışma ile karşılaştırıldığında sonuçların yakın olduğu gözlemlenmiştir. Bu sonuçlar yapılan çalışmanın hekimler için ön değerlendirme aşamasında faydalı olabileceğini düşündürmektedir.

Tablo 5. k-NN, SVM ve ELM doğruluk performanslarının karşılaştırılması

GSY	Bölüt Boyutu	k-NN Doğruluk (Accuracy) (%)	SVM Doğruluk (Accuracy) (%)	ELM Doğruluk (Accuracy) (%)
Periodogram	128	99,30	99,41	99,33
	256	99,66	99,72	99,72
	512	99,81	99,88	99,75
	1024	99,75	100	99,88
	2048	100	100	100
	1024	100	100	100
Welch	128	99,30	99,24	99,18
	256	99,72	99,75	99,72
	512	99,75	99,88	99,75
	1024	99,88	100	99,88
	2048	100	100	100
	4096	100	100	100

Kaynaklar

- Alkan, A., (2006). EEG İşaretlerinin Ayrıştırılmasında, Altuzay Yöntemlerinin Kullanılması, *Journal of Yasar University*, (3), 211-219
- Andrzejak, R. G., Lehnertz, K., Mormann, F., Rieke, C., David, P., Elger, C. E., (2001) Indications of nonlinear deterministic and finite-dimensional structures in time series of brain electrical activity: Dependence on recording region and brain state, *Physical Review E*, 64(6), 061907
- Cover, T.M., Hart, P.E., (1967). Nearest Neighbor Pattern Classification, *IEEE Transactions on Information Theory*, 13(1), 21–27
- Demir, B., Erturk, S., (2010). Empirical mode decomposition of hyperspectral images for support vector machine classification, *IEEE Transactions on Geoscience and Remote Sensing*, 48(11), 4071-4084
- Guo, L., Rivero, D., Pazos, A., (2010). Epileptic seizure detection using multiwavelet transform based approximate entropy and artificial neural networks, *Journal of Neuroscience Method*, (193), 156-163
- Güler, İ., Übeyli, E.D., (2005). Adaptive neuro-fuzzy inference system for classification of EEG signals using wavelet coefficients, *Journal of Neuroscience Methods*, (148), 113-121.
- Huang, G.B., Zhu, Q.Y., Siew, C.K., (2006). Extreme learning machine: theory and applications, *Neurocomputing*, 70, 1-3, 489-501
- Huang, G.B., Zhou, H., Ding, X., Zhang, R., (2012). Extreme learning machine for regression and multiclass classification, *IEEE Transactions on Systems, Man, and Cybernetics*, 42(2), 513–529
- Kanga, J., Chunga, Y.G., Kim, S., (2015). An efficient detection of epileptic seizure by differentiation and spectral analysis of electroencephalograms, *Computers in Biology and Medicine*, 66, 352–356
- Kaya, Y., Uyar, M., Ramazan Tekin, R., Yıldırım, S., (2014). 1D-Local Binary Pattern Based Feature Extraction for Classification of Epileptic EEG Signals, *Applied Mathematics and Computation*, (243), 209-219
- Koçak, O., Beytar, F., Firat, H., Telatar, Z., Eroğul, O., (2016). EEG İşaretlerinin Apne Süreci Analizinde Parametrik Olmayan GSY Tespit Yöntemlerinin Karşılaştırılması, TIPTEKNO'16. Tıp Teknolojileri Kongresi, Antalya
- Kumar, Y., Dewal, M.L., Anand, R.S., (2014). Epileptic Seizures Detection in EEG Using DWT-Based Apen and Artificial Neural Network. *Signal, Image and Video Processing*, 8(7), 1323–1334.
- Li, D., Xie, Q., Jin, Q., Hirasawa, K., (2016). A Sequential Method using Multiplicative Extreme Learning Machine for Epileptic Seizure Detection, *Neurocomputing*, (214), 692-707
- Mitchell, T., (1997). *Machine Learning*, McGraw Hill.
- Moavenian, M., Khorrani, H., (2010). A qualitative comparison of Artificial Neural Networks and Support Vector Machines in ECG arrhythmias classification, *Expert Systems with Applications*, 37, 3088–3093
- Nergiz, M., Özerdem, M.S., Akın, M., (2014). Dalgacık Dönüşümü Kullanılarak EEG İşaretlerinde Epileptik Nöbet Tespiti, TIPTEKNO'14. Tıp Teknolojileri Kongresi, Kapadokya
- Nicolaou, N., Georgiou, J., (2012). Detection of Epileptic Electroencephalogram Based on Permutation Entropy and Support Vector Machines, *Expert Systems with Applications*, 39(1), 202–209

- Nigam, V., Graupe, D., (2004). A Neural-Network-Based Detection of Epilepsy, *Neurological Research*, (26), 55-60
- Polat, H., Özerdem, M.S. (2014). Dalgacık Katsayıları ve Yapay Sinir Ağları Kullanılarak EEG İşaretlerinin Sınıflandırılması. TIPTEKNO'14. Tıp Teknolojileri Kongresi, Kapadokya
- Proakis, J.G., Manolakis, D.G., (1996). Digital Signal Processing Principles, Algorithms, and Applications. Prentice-Hall, New Jersey
- Sezgin, N., (2016). Epileptik EEG işaretlerin aşırı öğrenme makineleri ile sınıflandırılması, *Dicle Üniversitesi Mühendislik Dergisi*, 7(3), 529–551
- Toklu, Z., Kutlu, G., Demirbaş, H., Koyuncu, G., İnan, E.L., (2012). Ankara Eğitim ve Araştırma Hastanesi Epilepsi Polikliniğine Başvuran Epilepsi Hastalarının Demografik ve Klinik Bulguları. *Epilepsi Dergisi*, 13-18
- Türk, Ö., Özerdem, M.S., Akpolat, N., (2015). Gözler açık/kapalı durumunda EEG bantlarındaki frekans değişiminin Güç Spektral Yoğunluğu ile belirlenmesi, *Dicle Üniversitesi Mühendislik Dergisi*, 6(2), 131–138

Automatic detection of epileptic seizures for different time-scaled EEG signals

Extended abstract

Epilepsy is a common disease that recurs itself with constant seizures. This disease, which is seen in about 1% of the world population, is a clinical condition in which a sudden, unexpected and irregular electrical discharge occurs in a part of the brain or completely. The brain contains important information for the detection of electroencephalography (EEG) epilepsy, which is an electrical status analysis of the nerve cells. For this reason, EEG signals have become a research area that many experts are interested in.

In this study, we presented an automatic pattern recognition system using only A and E point clusters of signs from healthy subjects and epileptic seizures in a sample length of 23.6 seconds taken from Bonn University database (A, B, C, D, E). The pattern recognition system that has been presented in three stages, pre-processing, feature extraction and classification. In the first stage, the EEG signals consisting of 23.6 seconds and 4096 samples are divided into segments of 128, 256, 512, 1024, 2048, 4096 lengths. With this segmentation, we tried to study the behavior of EEG signs at different lengths. In the second step, spectral information of EEG signals was obtained by using Periodogram and Welch methods of nonparametric Power Spectral Density (PSD) methods. When Welch PSD was performed, a Hamming window was used for each of the lengths of one EEG mark length, and the overlap ratio of the parts was selected as 50%. By using two different PSD estimation methods in the study, it was tried to investigate the behavior patterns of the data segments in different spectral methods. After the PSD estimation, an arithmetic mean is applied to all EEG signals to reduce the data size, and each segment is represented by the feature vector in 16 sample lengths. In the third and last phase, the feature vectors of 16 sample lengths obtained for each EEG segment are classified by k nearest neighbors (k -NN), support vector machine (SVM) extreme learning machine (ELM) using 5-fold cross-validation method. With this classification, we tried to investigate the performances of different classifiers of feature vectors obtained by different PSD estimates of different data segments.

The accuracy performances for the 128, 256, 512, 1024, 2048 and 4096 segments resulting from the classification of the feature vectors obtained by the Periodogram PSD estimation with k -NN were 99.30%, 99.66%, 99.81%, 99.75%, 100%, 100% respectively. The accuracy performances for the 128, 256, 512, 1024, 2048 and 4096 segments resulting from the classification of the feature vectors obtained by the Welch GSY estimation with k -NN were 99.30%, 99.72%, 99.75%, 99.88%, 100%, 100% respectively.

The accuracy performances for the 128, 256, 512, 1024, 2048 and 4096 segments resulting from the classification of the feature vectors obtained by periodogram PSD estimation with SVM were 99.41%, 99.72%, 99.88%, 100%, 100%, 100% respectively. The accuracy performances for the 128, 256, 512, 1024, 2048 and 4096 segments resulting from the classification of the feature vectors obtained by the Welch PSD estimation with the SVM were 99.24%, 99.75%, 99.88%, 100%, 100%, 100% respectively.

The accuracy performance for the 128, 256, 512, 1024, 2048 and 4096 segments resulting from the classification of the feature vectors obtained by the Periodogram PSD estimation by ELM was 99.33%, 99.72%, 99.75%, 99.88%, 100%, 100% respectively. The accuracy performances for the 128, 256, 512, 1024, 2048 and 4096 segments resulting from the classification of the feature vectors obtained by Welch PSD estimation by ELM were 99.18%, 99.72%, 99.75%, 99.88%, 100%, 100% respectively.

It suggested that the pattern recognition system wherein the performance evaluated for the SVM classifier has been found that good performance is obtained. In terms of feature extraction methods used in the system Welch PSD estimation method of Periodogram PSD it was found to give better results according to the estimation. As accuracy performance is evaluated in terms of the different EEG data lengths used, in case of the data length redundancy, the accuracy performance is improved. A number of pattern recognition techniques have been proposed with different feature extraction and classification techniques than the EEG markers used in this study. The results obtained in these studies were observed to be close to the results when we performed the study.

Keywords: Epilepsy, Electroencephalography, Power spectral density, SVM, k -NN, ELM

mühendislik dergisi

