



TIP TEKNO'17

TIP TEKNOLOJİLERİ KONGRESİ

12-14 Ekim 2017 / TRABZON

Karadeniz Teknik Üniversitesi, Prof.Dr. Osman Turan Kongre Merkezi



Biyomedikal ve Klinik
Mühendisliği Derneği



Elektrik-Elektronik Mühendisliği Bölümü
Bilgisayar Mühendisliği Bölümü

C4.5 Karar Ağacı Temelli Öznitelik Seçimi ile Video-Okülografi (VOG) Sinyallerinden Diyabetik Retinopati Hastalığının Sınıflandırılması

Classification of Diabetic Retinopathy Disease from Video-Oculography (VOG) Signals with Feature Selection Based on C4.5 Decision Tree

Ceren Kaya¹, Okan Erkaymaz², Orhan Ayar³, Mahmut Özer⁴

¹Biyomedikal Mühendisliği Bölümü, Bülent Ecevit Üniversitesi, Mühendislik Fakültesi, Zonguldak, Türkiye
ceren.kaya@beun.edu.tr

²Bilgisayar Mühendisliği Bölümü, Bülent Ecevit Üniversitesi, Mühendislik Fakültesi, Zonguldak, Türkiye
okan.erkaymaz@beun.edu.tr

³Göz Hastalıkları Anabilim Dalı, Bülent Ecevit Üniversitesi, Tıp Fakültesi, Zonguldak, Türkiye
orhanayar@gmail.com

⁴Elektrik-Elektronik Mühendisliği Bölümü, Bülent Ecevit Üniversitesi, Mühendislik Fakültesi, Zonguldak, Türkiye
mahmutozer2002@yahoo.com

Özetçe— “Diyabet (Şeker Hastalığı)“, pankreas bezinden salgılanan insülin hormonu bozukluklarına dayanan bir hastalıktır. Klinik bulgular, diyabetin hayati organlarda bazı hastalıklara neden olduklarını ortaya koymaktadır. “Diyabetik Retinopati” diyabete dayalı en yaygın göz hastalıklarından birisidir ve retina tabakası damarlarının yapısal değişikliğe uğraması sonucunda ortaya çıkan görme kaybının önde gelen nedenidir. Son yıllarda yapılan araştırmalar, hayati organlardan elde edilen sinyallerin literatürdeki hastalıkların teşhisinde kullanılabileceğini göstermektedir. Bu çalışmada, sağ ve sol gözden alınan yatay ve dikey Video-Okülografi (VOG) sinyallerinin öznitelikleri, erken evre ve geç evre diyabetik retinopati hastalığını sınıflandırmak için kullanılmıştır. 25 istatistiksel öznitelik, 24 bireyden alınan VOG sinyalleri ile ayrık dalgacık dönüşümü kullanılarak elde edilmiştir. Elde edilen 25 öznitelikten C4.5 karar ağacı algoritması kullanılarak öznitelik seçimi yapılmıştır. Ayrık dalgacık dönüşümü ve C4.5 karar ağacından elde edilen istatistiksel öznitelikler yapay sinir ağlarına giriş olarak uygulanarak “Diyabetik Retinopati” hastalığının bu iki yöntemle göre sınıflandırma performansı karşılaştırılmıştır. Bulgularımız, C4.5 karar ağacı algoritması ile seçilen özniteliklerin (%96.87), ayrık dalgacık dönüşümü ile çıkarılan özniteliklere (%93.75) nazaran daha iyi sınıflandırma performansı sağladığını göstermektedir.

Anahtar Kelimeler — Video-Okülografi (VOG); Ayrık Dalgacık Dönüşümü; Öznitelik Çıkarımı; C4.5 Karar Ağacı; Öznitelik Seçimi; Yapay Sinir Ağları; Diyabetik Retinopati.

Abstract— "Diabetes Mellitus (Diabetes)" is a disease based on insulin hormone disorders secreted from the pancreas gland. Clinical findings find out that diabetes causes some diseases in vital organs. "Diabetic Retinopathy" is one of the most common eye diseases based on diabetes, and it is the leading cause of visual loss resulting from structural changes in the retinal vessels. Recent researches show that signals from vital organs can be used to diagnose diseases in the literature. In this study, the features of horizontal and vertical Video-Oculography (VOG) signals from right and left eye are used to classify non-proliferative and proliferative diabetic retinopathy disease. 25 statistical features are obtained using discrete wavelet transform with VOG signals from 24 subjects. Feature selection is performed using C4.5 decision tree algorithm from 25 features obtained. The statistical features obtained from C4.5 decision tree and discrete wavelet transform are applied as input to artificial neural networks and the classification performance of the "Diabetic Retinopathy" disease are compared according to these two methods. Our results show that feature selection by C4.5 decision tree algorithm (96.87%) provides better classification performance than feature extraction with discrete wavelet transform (93.75%).



Biyomedikal Sinyal İşleme 2

12 Ekim 2017 - 17.15-18.45 - Salon A

Keywords — Video-Oculography (VOG); Discrete Wavelet Transform; Feature Extraction; C4.5 Decision Tree; Feature Selection; Artificial Neural Networks; Diabetic Retinopathy.

I. GİRİŞ

Diyabetik retinopati ciddi görme kaybına yol açtığından erken teşhisi çok önemlidir. Oftalmoloji kliniklerinde, fundus florescein anjiyografi, optik koherens tomografi vb. yalnızca hastalık komplikasyonları oluştuğunda bilgi veren çeşitli cihazlar kullanılır. Komplikasyonlar gelişmeden önce retinal dokudaki farklılıkları saptamak için yöntemler geliştirildiğinde ise, uzman doktorlar hastalığın erken döneminde müdahale etme kabiliyetine sahip olmaktadır [1-2]. Literatür incelendiğinde; diyabetik retinopati ile ilgili yapılan çalışmaların görüntü işleme alanında yapay zeka yöntemlerini içerdiği görülmektedir. Priya ve arkadaşları yaptıkları çalışmada, destek vektör makinesi ve olasılıksal sinir ağı yöntemlerini diyabetik retinopati tanısı ve sınıflandırması sırasında fundus görüntülerini kullanarak karşılaştırmışlardır [3]. Rajput ve arkadaşları, diyabetik retinopati hastalarının retinal fundus görüntülerinden özellik çıkarımında dalgacık dönüşümü ve sınıflandırmada aşamasında ise k-ortalama kümeleme algoritmasını kullanmışlardır [4]. Noronha ve arkadaşları, fundus görüntülerinin özelliklerini ayırık dalgacık dönüşümü ile çıkararak destek vektör makinesi ile sınıflandırmışlardır [5]. Gürkan ve arkadaşları, Elektrokülografi (EOG) sinyallerinin sınıflandırılmasında destek vektör makinesi ve yapay sinir ağı yöntemlerini karşılaştırmışlardır [6]. Banerjee ve arkadaşları, göz hareketlerine göre k-ortalama kümeleme algoritmasını kullanarak yapay sinir ağları ile EOG sinyalini sınıflandırmışlardır [7]. Kim ve arkadaşları, EOG sinyalini göz hareketlerinden kaynaklanan diğer sinyallerden ayırmak için yeni bir algoritma oluşturarak bir karar destek sistemi tasarlamışlardır [8]. Kaya ve arkadaşları, VOG sinyallerinden ayırık dalgacık dönüşümü ile elde ettikleri öznelikleri kullanarak diyabetik retinopati hastalığının sınıflandırılması için yapay sinir ağları yöntemini kullanmışlardır. Önerilen yöntem %95.83 başarı ile sınıflandırma yapabilmektedir [9]. Saba ve arkadaşları, diyabet veri setlerinin sınıflandırılması için tek yöntem yerine ID3, C4.5 ve CART gibi üç tür karar ağacı algoritmasını karşılaştırmışlardır [10].

Bu çalışmada, erken evre diyabetik retinopati ve geç evre diyabetik retinopati hastalığı olan bireyleri yatay ve dikey Video-Okülografi (VOG) sinyalleri kullanarak sınıflandırabilen ve aynı zamanda diyabetik retinopatinin hastalık seviyesini tespit edebilen yeni bir karar destek sistemi ortaya konmuştur. VOG veri setinden ayırık dalgacık dönüşümü (ADD) yöntemi ile elde edilen normalize edilmiş istatistiksel özneliklerin ve C4.5 karar ağacı algoritmasıyla seçilen özneliklerin sınıflandırma performansı yapay sinir ağları kullanılarak araştırılmıştır.

II. VOG SİNYAL ÖLÇÜMÜ

Önerilen iki yöntemin performansını doğru bir şekilde ölçmek için Metrovision MonPackOne Elektrokülografi

cihazında bulunan 200 Hz'lik iç takip kamerası kullanılarak bir VOG sinyal veri seti oluşturulmuştur. "Bülent Ecevit Üniversitesi Klinik Araştırmalar Etik Kurulu"nun izniyle elde edilen veri seti 24 örnek ve üç sınıf içermektedir. (Sağlıklı: 7 birey, Erken Evre Diyabetik Retinopati: 9 birey ve Geç Evre Diyabetik Retinopati: 8 birey). Her örnek, ayırık dalgacık dönüşümü ile hesaplanan 25 istatistiksel özneliğe sahiptir.

III. AYRIK DALGACIK DÖNÜŞÜMÜ İLE ÖZNETELİK ÇIKARIMI

Dalgacık dönüşümü geleneksel Fourier dönüşümüne alternatif bir yöntem olarak türetilmiştir. Dalgacık dönüşümünün en önemli avantajı, alçak ve yüksek frekanslarda farklı pencere boyutlarının kullanılmasıdır [11]. Biyomedikal sinyal işlemede, biyolojik işaretlerin yapıları nedeniyle sürekli zaman dalgacık dönüşümü yerine ayırık zaman dalgacık dönüşümü tercih edilmektedir. Orijinal sinyal, detay (D1-D2-D3-D4) ve yaklaşım (A4) katsayılarını elde etmek için sırasıyla bir yüksek geçiren filtreden ve bir alçak geçiren filtreden geçirilir. Ardından, aynı işlem, yalnızca yaklaşım katsayıları için tanımlanmış ayrışma seviyesine kadar sürdürülür [12]. Bu çalışmada, Coiflet dalgacığı (Coif4) VOG sinyallerinin dalgacık dönüşümü katsayılarını hesaplamak için kullanılmıştır. Sınıflandırma başarımını arttırmak için ADD yöntemiyle elde edilen öznelik vektörlerinin boyutları, istatistiksel yöntemlerle katsayıların maksimumu, minimumu, ortalaması, standart sapması ve varyansı kullanılarak azaltılmıştır. Video-Okülografi (VOG) sinyal verilerinin istatistiksel özneliklerinin örneği Tablo 1'de belirtilmiştir.

Maksimum (µV)	Minimum (µV)	Ortalama (µV)	Standart Sapma (µV)	Varyans (µV)
2311,5786	2516,63723	3614,94357	3245,5144	8920,56845
2285,4076	2516,06465	2600,89071	3232,5731	4522,94737
9657,5316	9116,91229	5635,14108	3974,5500	2099,36929
1786,7207	2555,19428	3449,40526	3651,7915	7741,84778
9555,4005	8825,49503	6234,75530	6294,4114	2629,77110

Tablo 1. Video-Okülografi (VOG) sinyal verilerinin istatistiksel öznelikleri

IV. C4.5 KARAR AĞACI İLE ÖZNETELİK SEÇİMİ

Karar ağacı yapısında, örneklerin aynı sınıfa ait olması durumunda ağaç bir yaprağı temsil eder, bu nedenle yaprak aynı sınıfla etiketlenerek geri gönderilir. Olası bilgi, öznelik üzerinde yapılan bir testle verilen her öznelik için hesaplanır. Ardından, öznelik üzerindeki testten kaynaklanacak bilgi kazancı hesaplanır. Daha sonra en iyi öznelik, mevcut seçim kriterine ve dallanma için seçilen özneliğe dayanarak bulunur. Karar ağacı algoritmalarından C4.5'in uygulanması ve yorumlanması kolay olduğundan kullanımı yaygındır [13]. Bu çalışmada, ADD ile elde edilen istatistiksel özneliklerden daha başarılı

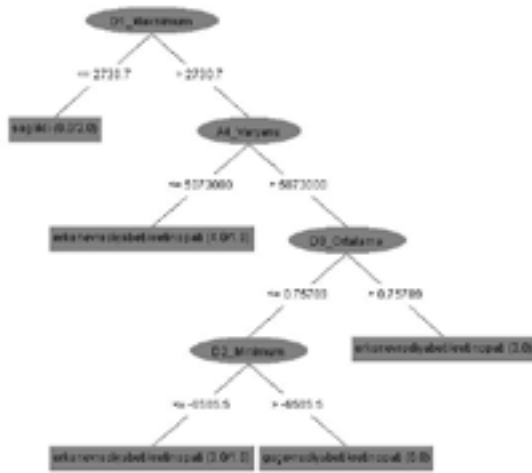
Biyomedikal Sinyal İşleme 2

12 Ekim 2017 - 17.15-18.45 - Salon A

sınıflandırma sonuçları elde edilmesini sağlayan öznelik seçimleri, Weka veri madenciliği programında C4.5 karar ağacı algoritması (J48) kullanılarak gerçekleştirilmiştir. C4.5 karar ağacı ile seçilen öznelikler Tablo 2'de belirtilmiştir. Sol göz yatay için oluşturulan karar ağacı Şekil 3'te verilmiştir.

Göz Yönü	C4.5 Karar Ağacı Öznelik Seçimi
Sağ Göz Dikey	D1-Maksimum, D1-Standart Sapma, D2-Minimum, D3-Standart Sapma, A4-Maksimum
Sol Göz Dikey	D1-Ortalama, D1-Standart Sapma, D3-Maksimum, D4-Standart Sapma
Sağ Göz Yatay	D1-Maksimum, D1-Ortalama, D1-Varyans, D2-Minimum, D3-Ortalama
Sol Göz Yatay	D1-Maksimum, D2-Minimum, D3-Ortalama, A4-Varyans

Tablo 2. C4.5 karar ağacı ile öznelik seçimi



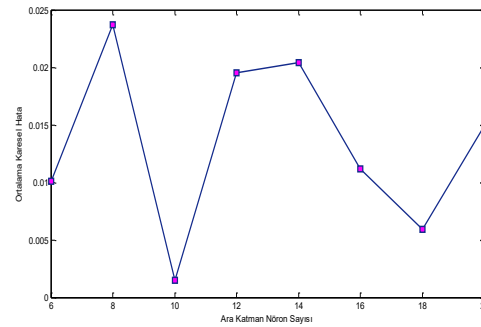
Şekil 3. C4.5 karar ağacı (Sol göz yatay)

V. YAPAY SİNİR AĞLARI İLE SINIFLANDIRMA

Yapay sinir ağları (YSA), beyin gibi biyolojik sinir sistemlerinin bilgi işleme yöntemlerinden ilham alan matematiksel modellerdir. Birbirine bağlı olarak çalışan nöronlardan oluşan YSA mimarisi veri sınıflandırması gibi özel uygulamalar için yapılandırılmışlardır [14]. Hazırlanan modelde, [0,1] aralığında normalize edilen 25 istatistiksel öznelik, veri seti YSA'nın girişi olarak kullanılmıştır. Normalize edilen veri seti ile ağ Levenberg - Marquardt geri yayılım algoritması kullanılarak 200 iterasyon boyunca eğitilmiştir. Eğitim sürecinde çapraz doğrulama yöntemi kullanılarak aşırı öğrenme engellenmiştir. Model performansının ortaya konması için regresyon katsayısı, ortalama karesel hata, ortalama mutlak hata, doğruluk, hassasiyet, duyarlılık ve özgüllük değerleri veri setinin %80 eğitim ve %20 test olarak bölünmesiyle hesaplanmıştır.

VI. SONUÇLAR

İlk olarak 25 giriş, 1 ara katman ve 1 çıkış içeren bir YSA modeli oluşturulmuştur. Bu topoloji için veri seti çapraz doğrulama yöntemi ile eğitim ve test olarak iki kısma ayrılmıştır. Ara katmandaki nöronların sayısını bulmak için ortalama karesel hata performansını üzerinde ara katman nöron sayısının değişiminin etkisi araştırılmıştır. Elde edilen sonuç Şekil 4'te gösterilmiştir.



Şekil 4. Ara katman nöron sayısının ortalama karesel hataya etkisi

Şekil 4'te açıkça görüldüğü gibi ortalama karesel hatanın minimum değeri ara katman nöron sayısı 10 olduğunda elde edilmiştir. ADD ile elde edilen 25 istatistiksel öznelik ve C4.5 karar ağacı ile seçilen (sağ göz için 5 ve sol göz için 4) istatistiksel öznelik giriş olduğu YSA modelleri, veri setinin %80'i ile eğitilip, kalan %20'si ile test edilmiştir. Tablo 3'ten görüldüğü gibi C4.5 & YSA modeli %96.87 sınıflandırma doğruluğu ile %93.75 doğruluğu olan ADD & YSA modeline göre daha yüksek başarıma sahiptir. Aynı zamanda, C4.5 & YSA modelinin, diyabetik retinopati hastalığının tespitinde %97.22 hassasiyet, %95 duyarlılık ve %96.66 özgüllük değerleri ile güçlü bir karaktere sahip olduğu görülmüştür.

Model	Göz Yönü	Doğruluk (%)	Hassasiyet (%)	Duyarlılık (%)	Özgüllük (%)
ADD & YSA	Sağ Dikey	91.67	88.89	88.89	93.33
	Sol Dikey	95.83	100	90	93.33
	Sağ Yatay	91.67	77.78	100	100
	Sol Yatay	95.83	100	90	93.33
C4.5 & YSA	Sağ Dikey	95.83	100	90	93.33
	Sol Dikey	95.83	100	90	93.33
	Sağ Yatay	95.83	88.89	100	100
	Sol Yatay	100	100	100	100

Tablo 3. Ayrık dalgacık dönüşümü ile C4.5 karar ağacının istatistiksel analizi

Buna ek olarak, test aşamasında modellerin hata analizi Tablo 4 ve Tablo 5'te sunulmuştur. C4.5 karar ağacı & YSA modelinin (0.0116), ADD & YSA modelinden

Biyomedikal Sinyal İşleme 2

12 Ekim 2017 - 17.15-18.45 - Salon A

(0.0174) daha iyi hata performansına ve yüksek regresyon katsayısına sahip olduğu gözlenmiştir.

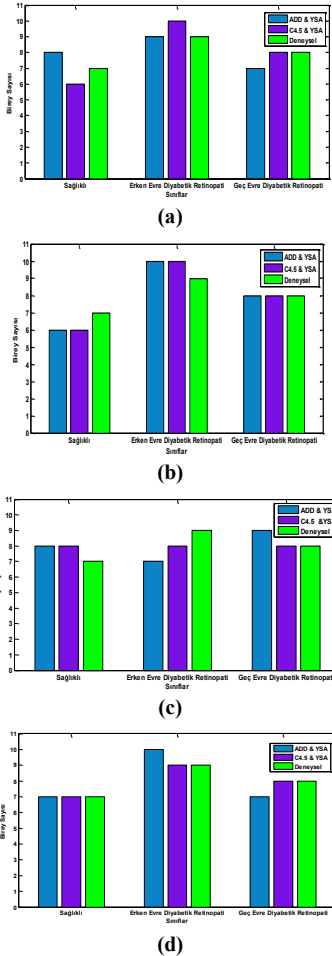
ADD & YSA	Regresyon Katsayısı	Ortalama Karesel Hata	Ortalama Mutlak Hata
Sağ Göz Dikey	0.93441	0.0222	0.0883
Sol Göz Dikey	0.94234	0.0178	0.0409
Sağ Göz Yatay	0.93640	0.0200	0.0879
Sol Göz Yatay	0.96957	0.0097	0.0491

Tablo 4. Ayrık dalgacık dönüşümü hata analizi

C4.5 & YSA	Regresyon Katsayısı	Ortalama Karesel Hata	Ortalama Mutlak Hata
Sağ Göz Dikey	0.93859	0.0210	0.0817
Sol Göz Dikey	0.97627	0.0083	0.0401
Sağ Göz Yatay	0.95151	0.0159	0.0678
Sol Göz Yatay	0.99534	0.0015	0.0156

Tablo 5. C4.5 karar ağacı hata analizi

Son olarak, diyabetik retinopati hastalık seviyesi üç sınıfa ve iki YSA modeline göre analiz edilerek, hesaplanan karışıklık matrisi sonuçları Şekil 5'te her bir göz yönü için grafiksel olarak sunulmuştur.



Şekil 5. Veri seti %80-%20 için elde edilen karışıklık matrisi sonuçları

a) Sağ göz dikey kanal b) Sol göz dikey kanal c) Sağ göz yatay kanal d) Sol göz yatay kanal

Bulgularımız, C4.5 karar ağacı algoritması ile seçilen özneliklerin, ayrık dalgacık dönüşümü ile çıkarılan özneliklere göre daha iyi sınıflandırma performansı sağladığını göstermektedir. Ön işleme aşaması için farklı öznelik çıkarma yöntemlerinin kullanılması ve VOG sinyalini kullanarak çeşitli göz hastalıklarının tespit edilmesi çalışmalarına devam edilmektedir.

TEŞEKKÜR

Bu çalışma, Bülent Ecevit Üniversitesi tarafından desteklenmiştir (BAP Proje No: 2015-38701644-01).

KAYNAKÇA

- [1] Vallabha, D., Dorairaj, R., Namuduri, K. and Thompson, H., "Automated Detection and Classification of Vascular Abnormalities in Diabetic Retinopathy", Proceedings of 13th IEEE Signals, Systems and Computers, 1625-1629, 2004.
- [2] Walter, T., Klein, J. C., Massin, P. And Erginay, A., "A Contribution of Image Processing to The Diagnosis of Diabetic Retinopathy- Detection of Exudates in Colour Fundus Images of The Human Retina", IEEE Transactions on Medical Imaging, 22(10): 1236-1243, 2002.
- [3] Priya, R. and Aruna, P., "Diagnosis of Diabetic Retinopathy Using Machine Learning Techniques", ICTACT Journal on Soft Computing, 3(4): 563-575, 2013.
- [4] Rajput, Y. M., Manza, R. R., Patwari, M. B., Rathod, D. D., Borde, P. L. and YMLPawar, P. L., "Detection of Non-Proliferative Diabetic Retinopathy Lesions Using Wavelet and Classification Using K-Means Clustering", International Conference on Communication Networks (ICCN), 381-387, 2015.
- [5] Noronha, K., Acharya, U. R., Nayak, K. P., Kamath, S. and Bhandary, S. V., "Decision Support System for Diabetic Retinopathy Using Discrete Wavelet Transform", Journal of Engineering in Medicine, 227(3): 251-261, 2012.
- [6] Gürkan, G., Gürkan, S. ve Uşaklı, A. B., "EOG Sinyalleri İçin Sınıflandırma Algoritmalarının Karşılaştırılması", 20. IEEE Sinyal İşleme ve İletişim Uygulamaları Kurultayı (SIU), 1-4, 2012.
- [7] Banerjee, A., Datta, S., Pal, M., Konar, A., Tibarewala, D. N. and Janarthanan, R., "Classifying Electrooculogram to Detect Directional Eye Movements", Int. Conf. on Comp. Intellig.: Model. Tech. and App. (CIMTA), 10: 67-75, 2013.
- [8] Kim, O., Doh, N. L., Youm, Y. and Chung, W. K., "Robust Discrimination Method of The Electrooculogram Signals for Human-Computer Interaction Controlling Mobile Robot", Intelligent Automation & Soft Computing, 13:319-336, 2013.
- [9] Kaya, C., Erkaymaz, O., Ayar, O. ve Özer, M., "Video-Okülografi (VOG) Sinyallerinden Diyabetik Retinopati Hastalığının Yapay Sinir Ağları İle Tespiti", 25. IEEE Sinyal İşleme ve İletişim Uygulamaları Kurultayı (SIU), 1-4, 2017.
- [10] Bashir, S., Qamar, U., Khan, F. H. and Javed, M. Y., "An Efficient Rule-Based Classification of Diabetes Using ID3, C4.5 & CART Ensembles", International Conference on Frontiers of Information Technology (FIT), 2014.
- [11] Goswami, J. C. and Chan, A. K., "Fundamentals of Wavelets", Second Edition, John Wiley & Sons, Inc., 2011.
- [12] Subaşı, A., "EEG Signal Classification Using Wavelet Feature Extraction and A Mixture of Expert Model", Expert Systems with Applications, 32 :1084-1093, 2007.
- [13] Huang J., Lu J. and Ling C. X., "Comparing Naive Bayes, Decision Trees, and SVM with AUC and Accuracy", 3rd IEEE International Conference on Data Mining, 553-556, 2003.
- [14] Erkaymaz, H., Özer, M. ve Orak, İ. M., "Detection of Directional Eye Movements Based on The Electrooculogram Signals Through an Artificial Neural Network", Chaos Solitons & Fractals, 56 : 202-208, 2015.