

Dalgacık Dönüşümü ve RUS Geliştirilmiş Ağaç Kullanarak Otomatik Aritmi Tespiti

Özgür TOMAK^{1*}, Birkut GÜLER², Alparslan TÜFEKÇİ¹, Kenan YANMAZ³

¹Giresun Üniversitesi, Elektrik ve Enerji Bölümü, , Giresun, Türkiye

²Giresun Üniversitesi, Makine ve Metal Teknolojileri Bölümü, Giresun, Türkiye

³Giresun Üniversitesi, Elektronik ve Otomasyon Bölümü, Giresun, Türkiye

Geliş Tarihi: 30.03.2018

Kabul Tarihi: 30.05.2018

***Sorumlu Yazar:** ozgur.tomak@giresun.edu.tr

Öz

Kalp hastalıklarının çok yaygın ve ölümcül hastalıklar olduğu söylenebilir. Bu nedenle EKG incelenerek kalp rahatsızlıklarını doğru tespit etmek gerekmektedir. Bu çalışmada kişinin elektrokardiyografi (EKG) kayıtlarının otomatik incelenmesi yoluyla aritmi tespiti hedeflenmektedir. Bu süreçte dalgacık yönteminden elde edilen özellikler RUS Geliştirilmiş Ağaç yöntemiyle sınıflandırılmıştır. Kalp atışları 7 farklı sınıfa ayrılmıştır. Bu aşamada öznitelik sayısını azaltmak ve sinyal işleme sürecini hızlandırmak için RUSBoost yöntemi kullanılmıştır. Bu yöntem çok miktarda zayıf öğreniciyi bir araya getirip, bu süreçten güçlü öğrenici oluşturmak olarak bilinir. ST-Petersburg Enstitüsü Kardiyolojik Veritabanı analiz için tercih edilmiştir. Test ve eğitim doğruluğu 12 kanal EKG verilerinde bulunmuştur. Yöntem gerçek zamanlı aritmi tespiti için yeterince hızlıdır. Bütün analizler için MATLAB kullanılmıştır.

Anahtar Kelimeler: Elektrokardiyografi (EKG), Ayrık Dalgacık Dönüşümü, RUSBoost, Sınıflandırma.

Automatic Arrhythmia Detection Using Wavelet Transform and RUSBoosted Trees Classification

Abstract

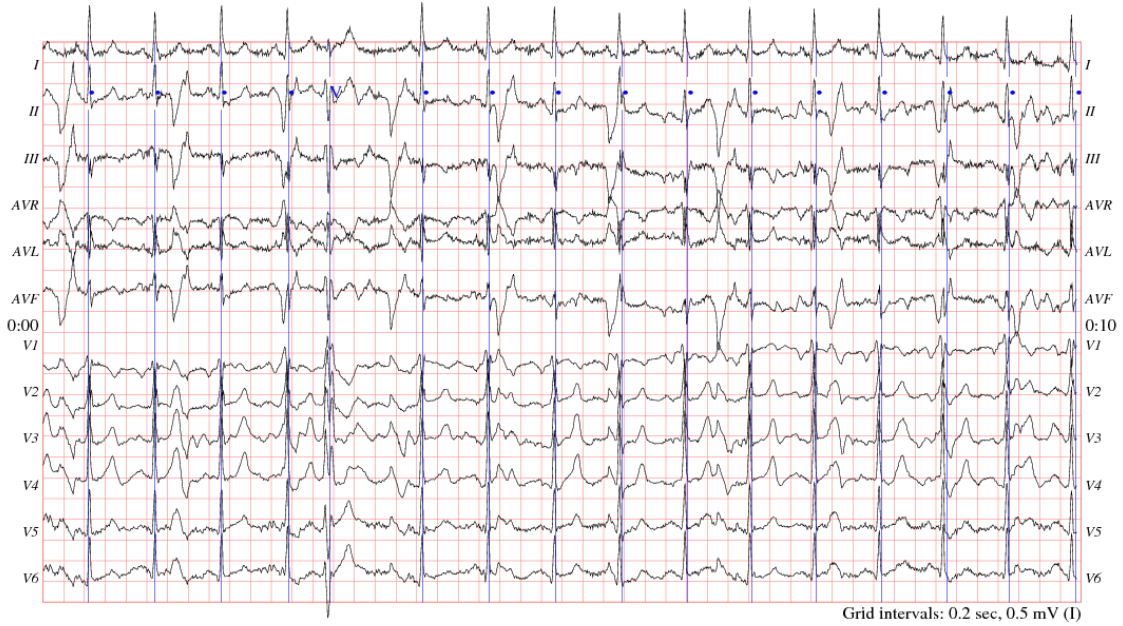
It can be said that heart diseases are very common and fatal diseases. Therefore, it is necessary to determine the heart diseases correctly by examining the ECG. In this study, it was aimed for detection of the arrhythmia by automatic examination of the person's electrocardiography (ECG) records. In this process, features obtained from the wavelet method was classified by RUSBoosted Trees method. The heartbeats were divided into seven different classes. RUSBoost method was used to reduce the number of features, and it speeds up the signal processing process. This method is known as bringing together a lot of weak learners and creating powerful learners from this process. ST-Petersburg Institute of Cardiological Database has been preferred for analysis. Test and training accuracy was found in 12 channel ECG data. The method was fast enough to detect real-time arrhythmia. MATLAB was used for all analyzes.

Keywords: Electrocardiogram (ECG), Discrete Wavelet Transform, RUSBoost, Classification.

1. Giriş

Elektrokardiyografi (EKG) kalp hastalıklarının tespit edilmesinde yaygın olarak kullanılan biyomedikal bir sinyaldir. Aritmilerin otomatik teşhisi erken tedavi imkânı sağlamaktadır. Bu çalışmada aritmi tespiti için ayrık dalgacık dönüşümünden gelen katsayılar RUSBoost yöntemiyle azaltılmış ancak daha güçlü özellikler elde edilmiştir. Bu özellikler RUS geliştirilmiş ağaç (RUSBoosted Trees) sınıflandırma metodu kullanılarak ayrıştırılmıştır.

Bu çalışmada ST-Petersburg Enstitüsü Kardiyolojik Veritabanı (ST-Petersburg Institute of Cardiological Database) kullanılmıştır. Bu kayıtlar yaşları 18 ile 80 arasında değişen 15 kadın, 17 erkekten oluşan hasta gurubuna aittir. 12 kanallı 30 dakika uzunluğunda 75 sinyal üzerinde analizler gerçekleştirilmiştir. 176177 kalp atışı harmanlanmış ve sonrasında kullanılmıştır. Her kayıt için örnekleme frekansı 257 Hz'dir. ST-Petersburg Enstitüsü Kardiyolojik Veritabanına ait örnek bir veri Şekil 1'de verilmiştir (Goldberger ve ark. 2000).



Şekil 1. ST-Petersburg Enstitüsü Kardiyolojik Veritabanına ait örnek bir veri.

EKG sinyalinde aritmi tespiti giderek önem kazanan bir çalışma konusudur. Pek çok şekilde bu konuda çalışmalar yapılmıştır. Bu çalışmaların bazılarında aşağıda bahsedilmiştir.

Dokur ve arkadaşları, çalışmalarında EKG sinyalinden Dalgacık Dönüşümü ve Fourier analizi kullanarak özellikleri belirlemişlerdir. Sınıflandırma aşamasında yapay sinir ağları ve genetik algoritmalar ve kullanılmıştır. 10 farklı aritmi tipinde sınıflandırma yapılmış, dalgacık dönüşümünde %99.4 ve Fourier Dönüşümünde %92.2 başarı sağlanmıştır (Dokur ve ark. 1999). Zhao ve arkadaşları, çalışmalarında her EKG sinyalinden dalgacık dönüşü ile elde ettikleri kat sayıları özellik olarak kullanmışlardır. Sonra özbağlımlı modelleme ile geri beslemeli ağlar oluşturup, Destek

Vektör Makinesi kullanarak 6 farklı aritmi çeşidini %99.68 başarıyla sınıflandırmışlardır (Zhao ve Zhang 2005). Jiang ve arkadaşları, dalgacık dönüşümünden elde ettikleri katsayılarla bağımsız bileşen analizini kullanarak elde ettikleri özellikler ile Destek Vektör Makinesi metodunu kullanarak sınıflandırma yapmışlar ve %98.65'lik başarı elde etmişlerdir (Jiang ve ark. 2006). Prasad ve Sahambi dalgacık ailesi olarak sym6 kullanmış ve 4.seviyede ayırtmış ve çıkan sonuçları yapay sinir ağları ile sınıflandırılmıştır (Prasad ve Sahambi, 200). Erdoğan ve Pekçakar yaptıkları çalışmada dalgacık ailesi olarak db2 ve db10 kullanmış ve 4.seviyede ayırtmadan elde edilen katsayıları kullanarak yapay sinir ağlarında sınıflandırılmıştır (Erdoğan ve Pekçakar, 2009). Saini ve arkadaşları, dalgacık dönüşümü ve k En Yakın Komşular(KNN) sınıflandırıcısını kullandılar ve % 87.5 doğruluk elde ettiler(Saini ve ark. 2015). Dewangan ve Shukla, ayırık dalgacık dönüşümü ile EKG sinyallerini incelediler(Dewangan ve ark. 2016). Yapay sinir ağları ile % 87 doğruluk elde ettiler. Sayilgan ve arkadaşları, aritmiyi tespit için kümeleme yöntemini kullanmışlardır(Sayilgan ve ark. 2017). MIT-BIH Veri Tabanı'ndan alınan veriler ile analiz yapıldı ve analizlerde % 92 doğruluğa ulaşıldı. Rad ve arkadaşları, EKG sinyallerini sınıflandırılmak için bir sistem geliştirdi ve yapay sinir ağları ve Bayesian düzenlemesi kullanarak % 78.5 doğruluğa ulaştı (Rad ve ark. 2017).

Çalışmanın takip eden kısımlarında önce özellik çıkarımı için kullanılan ayırık dalgacık dönüşümü açıklanmıştır. Sonra sınıflandırma için kullanılan RUS Geliştirilmiş Ağaç Sınıflandırması üzerinde durulmuştur. En son bölümde ise yaptığımız analiz sonuçları incelenerek yorumlanmış ve diğer çalışmalarla karşılaştırılmıştır.

2. Kullanılan Yöntem

Özellik olarak dalgacık yönteminden elde edilen katsayılar kullanılmıştır. EKG sinyallerinden çıkarılan bu özellikler sınıflandırma işlemi için RUS Geliştirilmiş Ağaç algoritmasına giriş olarak verilmiştir.

2.1. Ayırık Dalgacık Dönüşümü

Dikey dalgacık taban fonksiyonumuz denklem 1'de verilmiştir.

$$\Phi_{(s,l)}(x) = 2^{-\frac{s}{2}} \Phi(2^{-s}x - l) \quad (1)$$

$\Phi_{(s,l)}(x)$ = Dikey dalgacık taban fonksiyonu, l = Konum indeksi,
 s = ölçek indeksi

Ayrık dalgacıklar ailesini oluşturmak için ana fonksiyonu genişleten ve ölçekleyen değişkenler l ve s 'dir. Konum indeksi l konumu verir ve ölçek indeksi s dalgacık genişliğini verir (Graps, 1995).

Analiz yapan dalgacık Φ , veri alanının farklı çözünürlüklerde yayılması için bir denklem 2'deki gibi bir ölçeklendirme denkleminde kullanılır (Tomak, 2018).

$$W(x) = \sum_{k=-1}^{N-2} (-1)^k c_{k+1} \Phi(2x + k) \quad (2)$$

Φ = Analiz yapan dalgacık, $W(x)$ = Ölçekleme fonksiyonu,
 c_k = dalgacık katsayıları

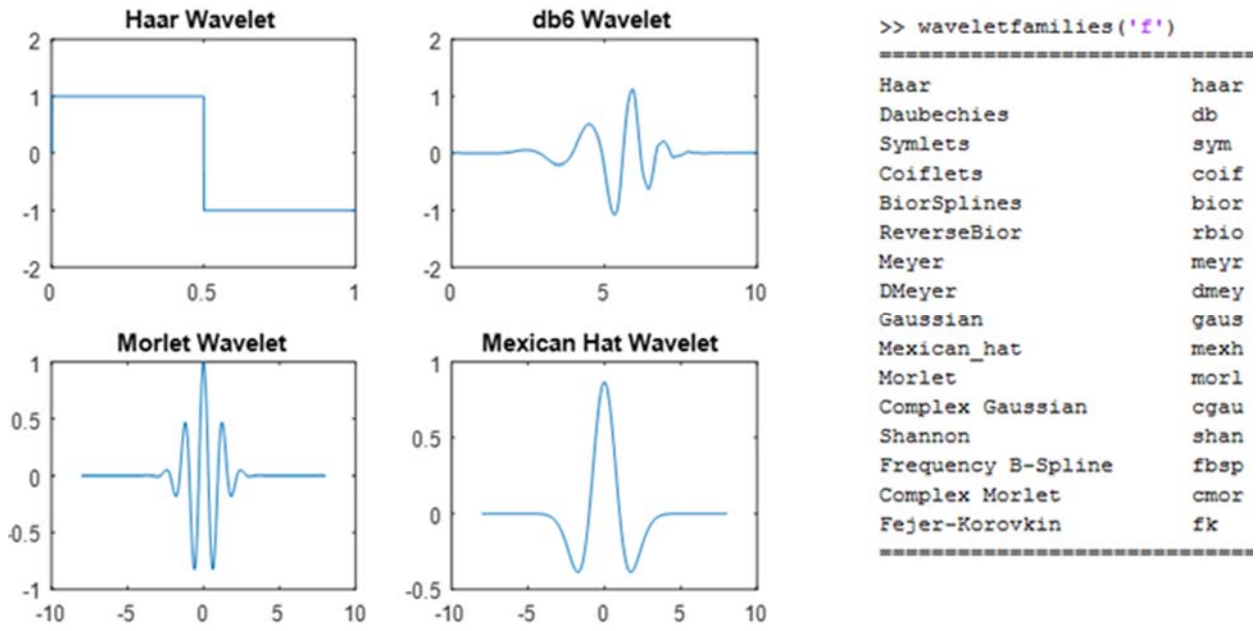
Dalgacık katsayıları denklem 3 ve 4'deki koşulları sağlamalıdır.

$$\sum_{k=0}^{N-1} c_k = 2 \quad (3)$$

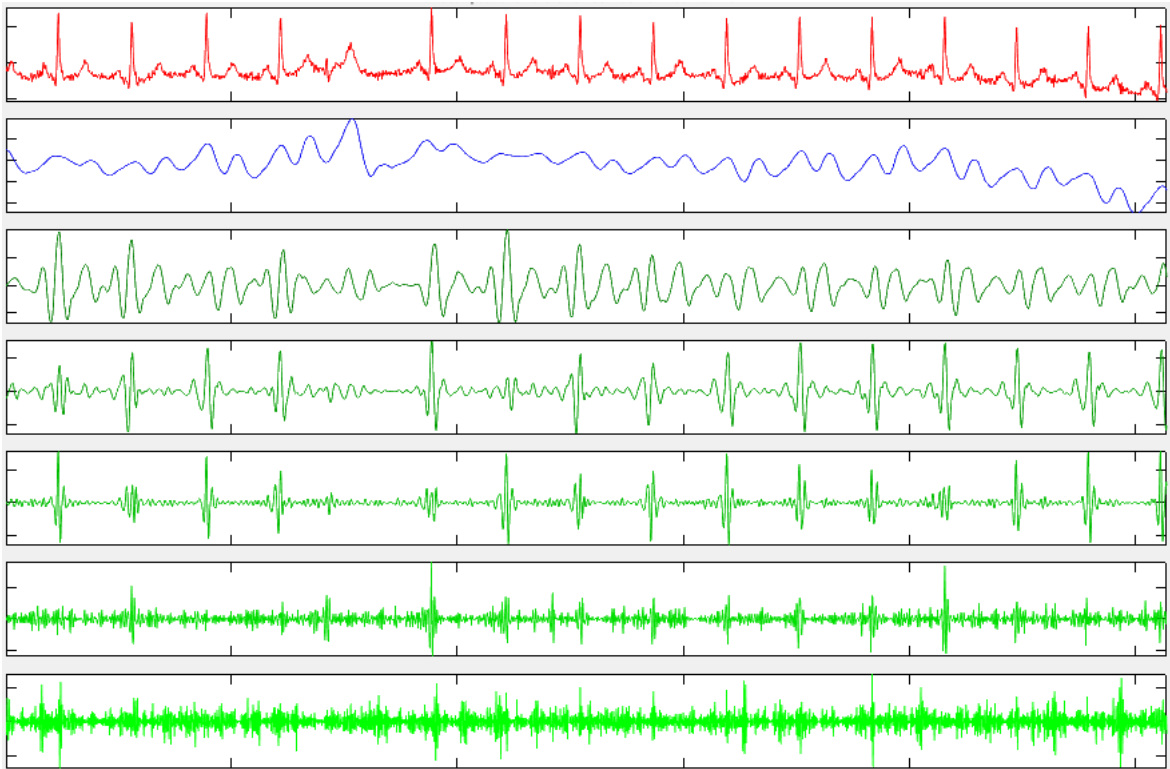
$$\sum_{k=0}^{N-1} c_k c_l = 2\delta_{l,0} \quad (4)$$

c_k ve c_l = dalgacık katsayıları, l = konum indeksi, δ = delta fonksiyonu

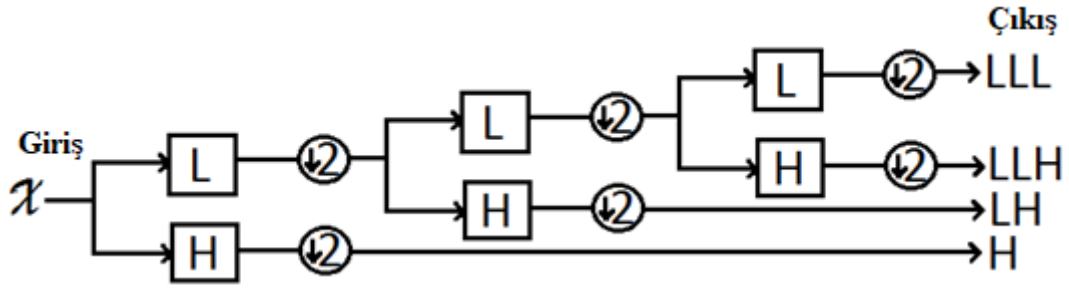
Dalgacık türlerinin isimleri ve bazı dalgacık fonksiyonlarının çizimleri şekil 2'de verilmiştir. Sinyalin 5. Seviye Db 6 ayrık dalgacık analizi şekil 3'de verilmiştir. Çok çözünürlüklü çift ağacı (Multiresolution dyadic tree) şekil 4'de verilmiştir. Bu çalışmada db6'nın seçilme nedeni yapısının EKG sinyaline benzemesidir.



Şekil 2. Dalgacık türlerinin isimleri ve bazı dalgacık fonksiyonlarının çizimleri



Şekil 3. 5. Seviye Db 6 ayrık dalgacık analizi.



Şekil 4. Çok çözünürlüklü çift ağacı.

2.2. RUS Geliştirilmiş Ağaç Sınıflandırması

Sınıflandırma yöntemi olarak RUS Geliştirilmiş Ağaç (RUSBoosted Trees) tercih edilmiştir. RUS Geliştirilmiş Ağaç için grup yöntemi RUSBoost seçilmiştir Robustboost adını verdiğimiz önerilen algoritmamız, Freund tarafından önerilen Brownboost algoritmasında bir varyasyondur. Bu ikisi arasındaki fark, eğitim hatasını en aza indirmek yerine, denklem 5’de tanımlanan tolerans bazlı maliyet fonksiyonunun asgariye indirilmesidir (Freund, 2009).

$$MF = \frac{1}{N} \sum_{i=1}^N 1[\bar{m}(x_i, y_i) \leq \theta] \quad (5)$$

Zaman değişkeni aralığı $0 \leq t \leq 1$ 'dir. Toleransı m ile gösterilir. Hedef toleransı $\theta \geq 0$ parametresidir. Çapraz-doğrulama kullanıldı ve ayarlandı. Artan θ değeri, validasyon seti ve de eğitim setine ait olan performans değerleri arasındaki farkı azaltacaktır. $\sigma_f > 0$ değeri, son potansiyel işlevindeki adımın eğimi ile tanımlanır. Robustboost kendini sonlandırma özelliği olan bir algoritmadır. Algoritma $t \geq 1$ olduğu zaman sona erer. Hata hedefi ϵ eğer küçük bir değer olarak ayarlanırsa Robustboost algoritması sonlandırılmaz. Burada ϵ değerinin doğru ayarlanması, algoritmanın uygun bir sayıda tekrarlanmasıyla sona ermesi, ϵ 'nin minimum değerinin bulunmasıyla yapılır. ϵ ayarı, aynı zamanda ρ parametresi için değeri belirler.

RUSBoost verilerin dengesiz olduğu durumda daha faydalıdır. RUS, Random-Under-Sampling kelimelerinin ilk harflerinden oluşan bir kısaltmadır. Algoritmamız, N örneklemede eğitim için kullanılan verilerin içinde en az üye barındıran sınıfa ait üye sayısını temel olarak alır. Daha da fazla üye sahibi olan sınıflar ise, her sınıfa ait N gözlemleri alır ve örneklendirilir. Farklı bir şekilde anlatsak, eğer K adet sınıflar var ise, o zaman, zayıf öğrenciler için, toplulukta, RUSBoost, K sınıf var ise N gözlemleriyle verilere ait bir alt küme alınır. AdaBoostM2'deki yolu topluluk oluşturmada ve değerlendirmede kullanabilir. AdaBoost.M2 zayıf baz sınıflandırıcıları ile çok-sınıf problemleri

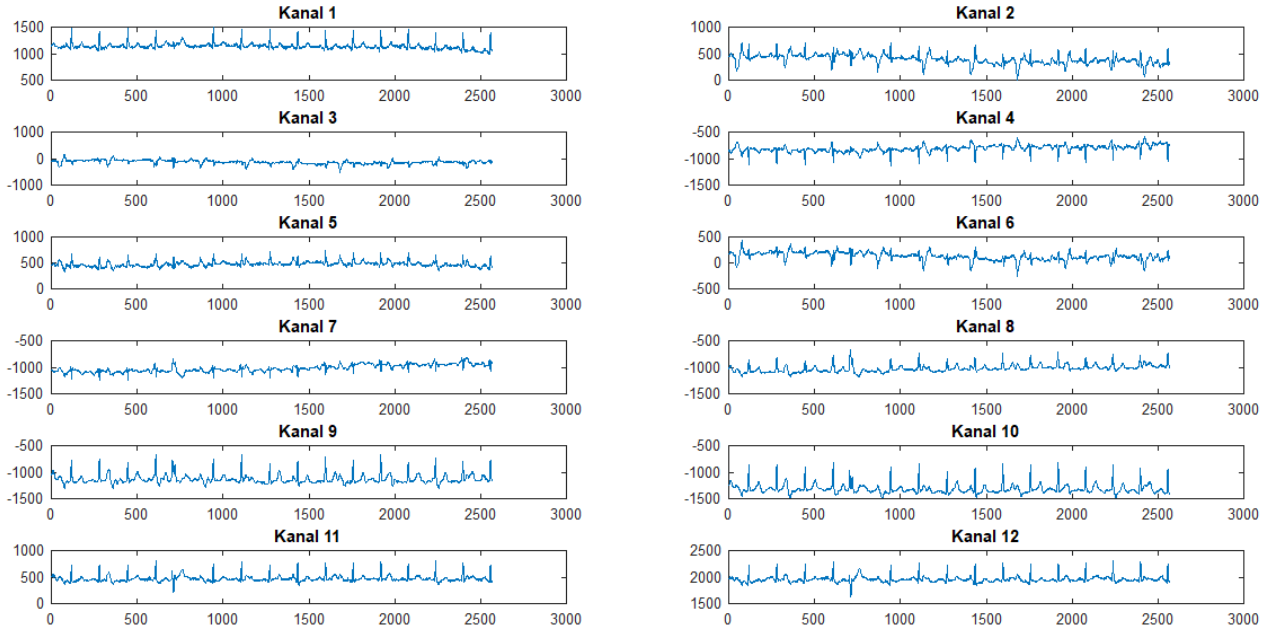
için tasarlanmış bir destek algoritmasıdır. Algoritma, eğitim hatası üzerinde gevşek olan sınırı en aza indirecek şekilde tasarlandı.

Öğrenme tipi karar ağacıdır. Karar ağaçları, hızlı sınıflandırma için en iyi seçeneklerden biridir ve ayrıca yorumlanması kolaydır. Öğrenen sayısı 30, öğrenme oranı 0.1, maksimum bölünme sayısı 20 seçildi.

Kalp atışları 7 farklı sınıfa ayrılmıştır. Bunlar normal atım, erken ventriküler kasılma, atriyal erken atım, düğüm kaçış vuruşu, supraventriküler erken atım, sağ dal bloğu, tanımlanamayan atım şeklinde sınıflandırılabilir. ST-Petersburg Enstitüsü Kardiyolojik Veritabanından aldığımız 176177 kalp atışını harmanlanarak, yarısı eğitim ve diğer yarısı test aşamasında kullanıldı. Eğitimde 10 kat katlama (10 fold) yaklaşımı kullanılmıştır. Sinyalin 12 kanal için db6 ile 5. seviye ayrıştırmasından elde edilen katsayılar RUSBoost yöntemiyle daha güçlü özniteliklere çevrildi ve RUS Geliştirilmiş Ağaç metoduyla MATLAB kullanılarak sınıflandırıldı. Bu sınıflandırma süreci için MATLAB'da bulunan fonksiyonlar kullanıldı. Elde edilen sonuçlar Tablo 1'de verilmiştir. Kullandığımız veriye ait 12 kanala ait ilk 10 saniyelik veri Şekil 5'de verilmiştir.

	Eğitim Doğruluk	Test Doğruluk
Kanal 1	94.1	93.9
Kanal 2	94.5	93.6
Kanal 3	93.7	92.5
Kanal 4	95.1	94.2
Kanal 5	93.9	93.1
Kanal 6	93.5	92.6
Kanal 7	92.3	91.6
Kanal 8	94.1	93.4
Kanal 9	94.7	93.9
Kanal 10	93.1	93.0
Kanal 11	94.4	93.6
Kanal 12	93.4	93.1

Tablo 1. Eğitim ve test doğruluk sonuçları



Şekil 5. 12 kanala ait ilk 10 saniyelik veri.

3. Sonuçlar ve Öneriler

Sınıflandırmada aşamasında kullandığımız veri ST-Petersburg Enstitüsü Kardiyolojik Veritabanından geldi ve veri miktarı 176177 kalp atışından oluşmaktadır Ayırık Dalgacık Yöntemiyle elde edilen katsayılardan RUSBoost yöntemiyle daha güçlü öznelikler elde edilmiş ve RUS Geliştirilmiş Ağaç (RUSBoosted Trees) metoduyla sınıflandırılmıştır. Benzer çalışmalarda tek kanal ile çalışılırken yaptığımız çalışmada 12 kanalda analiz yaptık. Benzer çalışmaların en yüksek başarı seviyeleri ile kendi en yüksek başarı seviyemizin karşılaştırılması Tablo 2’de verilmiştir. Elde edilen sonuçlar her kanal için başarılıdır ve yöntem gerçek zamanlı aritmi tespiti için yeterince hızlıdır yani her kalp atışına ait hesaplamalar yine bir kalp atışı süresi içinde yapılabilmektedir. Veri miktarında azalma sinyalin gönderilmesini ve işlenmesini kolaylaştırmaktadır.

	Doğruluk
Dokur ve arkadaşları	99.4
Zhao ve arkadaşları	99.68
Jiang ve arkadaşları	98.65
Saini ve arkadaşları	87.5
Dewangan ve Shukla	87
Sayilgan ve arkadaşları	92
Rad ve arkadaşları	78.5
Yöntemimiz	94.2

Tablo 2. Benzer çalışmaların doğruluk sonuçları

Kaynaklar

- Dewangan NK, Shukla SP. ECG arrhythmia classification using discrete wavelet transform and artificial neural network. In: Recent Trends in Electronics, Information & Communication Technology (RTEICT), IEEE International Conference on; 20-21 May 2016; Bangalore, India: IEEE. pp. 1892-1896.
- Dokur, Z., Olmez, T., and Yazgan, E. (1999). ECG waveform classification using the neural network and wavelet transform. In [Engineering in Medicine and Biology, 1999. 21st Annual Conference and the 1999 Annual Fall Meeting of the Biomedical Engineering Society] BMES/EMBS Conference, 1999. Proceedings of the First Joint (Vol. 1, pp. 273-vol). IEEE.
- Erdođmuş, P., and Pekçakar, A. (2009). Dalgacık dönüşümü ile ekg sinyallerinin özellik çıkarımı ve yapay sinir ađları ile sınıflandırılması. Uluslararası İleri Teknolojiler Sempozyumu, 13-15.
- Freund, Y. (2009). A more robust boosting algorithm. arXiv preprint arXiv:0905.2138.
- Goldberger, A. L., Amaral, L. A., Glass, L., Hausdorff, J. M., Ivanov, P. C., Mark, R. G., ... and Stanley, H. E. (2000). Physiobank, physiotookit, and physionet. *Circulation*, 101(23), e215-e220.
- Graps, A., An Introduction to Wavelets, IEEE Computational Science and Engineering, vol. 2, num. 2. IEEE Computer Society, Loas Alamitos-CA, USA, 1995.
- Jiang, X., Zhang, L., Zhao, Q., and Albayrak, S. (2006, November). ECG arrhythmias recognition system based on independent component analysis feature extraction. In TENCON 2006. 2006 IEEE Region 10 Conference (pp. 1-4). IEEE.
- Prasad, G. K., and Sahambi, J. S. (2003, October). Classification of ECG arrhythmias using multi-resolution analysis and neural networks. In TENCON 2003. Conference on Convergent Technologies for the Asia-Pacific Region (Vol. 1, pp. 227-231). IEEE.
- Rad AB, Eftestol T, Engan K, Irusta U, Kvaloy JT, Kramer-Johansen J, Wik L, Katsaggelos AK. ECG-based Classification of Resuscitation Cardiac Rhythms for Retrospective Data Analysis. *IEEE T Bio-Med Eng* 2017; 64.10: 2411-2418.
- Saini R, Bindal N, Bansal P. Classification of heart diseases from ECG signals using wavelet transform and kNN classifier. In: Computing, Communication & Automation (ICCCA), 2015 International Conference on; 15-16 May 2015; Noida, India: IEEE. pp. 1208-1215.
- Sayilgan E, Cura ÖK, İşler Y. Use of clustering algorithms and extreme learning machine in determining arrhythmia types. In Signal Processing and Communications Applications Conference (SIU) 2017 25th; 15-18 May 2017; Antalya, Turkey: IEEE. pp. 1-4.
- Tomak Özgür, (2018). Dalgacık Teorisi Kullanarak Güç Kalitesi Ve Geçici Durumların Analizi. Yüksek Lisans Tezi, Giresun Üniversitesi, Fen Bilimleri Enstitüsü, Giresun.
- Zhao, Q., & Zhang, L. (2005, October). ECG feature extraction and classification using wavelet transform and support vector machines. In Neural Networks and Brain, 2005. ICNN&B'05. International Conference on (Vol. 2, pp. 1089-1092). IEEE.