

Uzun-Kısa Vade Hafıza Tabanlı Kalp Ritmi Analizi ve Sınıflandırması

Heart Sounds Analysis and Classification Based on Long-Short Term Memory

Emre ÇANCIOĞLU¹, Savaş Şahin¹, Yalçın İşler²

¹Elektrik & Elektronik Mühendisliği Bölümü, Kâtip Çelebi Üniversitesi, İzmir, Türkiye
emrecancioglu7@gmail.com, [sasas.sahin@ikcu.edu.tr](mailto:savas.sahin@ikcu.edu.tr)

²Biomedikal Mühendisliği Bölümü, Kâtip Çelebi Üniversitesi, İzmir, Türkiye
islerya@yahoo.com

Özetçe—Bu çalışmada, kalp sesi fonokardiyogram dalga formlarının Normal, Hırıltılı, Ekstrasistol ve Yapay gibi kategorilere sınıflandırılma çalışması yapılmıştır. Sınıflandırma için kullanılan yaklaşımı genel bir makine öğrenimi uygulama bakış açısından sunarak, özellik çıkarma, performanslarını karşılaştırarak kullanılan sınıflandırıcıların türleri detaylandırıldı. Çalışmada kullanılan Uzun-Kısa Vadeli Hafıza (LSTM) metodu, ses kayıtlarındaki her bir kardiyak döngünün sınıflandırılmasını destekler. LSTM tabanlı özelliklere ek olarak, yöntemimiz tüm ses kayıtlarının özelliklerini özetlemek için spektral özellikler içerir.

Anahtar Kelimeler—Kalp Ritmi, Sınıflandırma, LSTM, RNN

Abstract—In this study, the development of an algorithm for the classification of heart sound phonocardiogram waveforms such as Normal, Murmur, Extrasystole, Artifact. By presenting the approach used for classification from a general machine learning application point of view, the types of classifiers used were detailed by comparing their features and their performance. The Long-Short Term Memory method which supports the classification of each cardiac cycle in sound recordings. In addition to the LSTM-based features, our method incorporates spectral features to summarize the characteristics of the entire sound recording.

Keywords—Heart Sounds, Classification, LSTM, RNN

I. GİRİŞ

Kalp oskültasyonu, üstünlük ve uygun maliyetli tanı testlerinden biridir. Hemodinamik durumun primer değerlendirmesini sağlayabilir ve ventriküler septal defektler ve aortta stenoz gibi kardiyovasküler bir hastalığı tespit edebilir [1]. Kalp sesi (veya fonokardiyogram) daha ileri tıbbi değerlendirmeler için ek tanı testleri de sunabilir.

Bununla birlikte, kalp sesinin pratik uygulamaları büyük ölçüde bilişsel becerilere ve tıp uzmanının yetkinliğine bağlıdır. Kayıt bölgelerinde duyulabilir frekans aralığı, çevresel gürültü ve varyasyon sınırlaması bu testin diğer önemli eksiklikleridir. Ayaktan izlemede maliyet etkin bir tanı testlerinde bu eksikliklerin giderilmesi için kalp seslerinin otomatik analizi için çeşitli teknikler önerilmiştir [2-5].

Fonokardiyogram (PCG) sinyal analizi iki ana kategoriye ayrılabilir. İlk yaklaşımlar zamansal segmentasyona, yani kardiyak döngülerin tanımlanmasına ve birinci (S1; sistol başlangıcı) ve ikinci (S2; sistol ucunun) birincil kalp seslerinin konumunun lokalizasyonuna dayanır. S1 ve S2 süresindeki değişim ve yoğunlukları, kardiyak anomalilerin kesin belirtileri olarak kabul edilir.

PCG segmentasyonu için Shannon enerjisi [2], Shannon entropisi [3], Hilbert-Huang dönüşümü [4] ve otokorelasyon [5] gibi farklı zarf ekstraksiyon yöntemleri kullanılarak çeşitli çalışmalar yapılmıştır. Sinyal zarfı gürültüyü azaltır ve sinyalin düşük yoğunluklu bileşenlerini güçlendirir. Bazı segmentasyon yaklaşımları, S1 ve S2 sesinin frekans karakteristiklerini elde etmek için dalgacık dönüşümüne dayalı zarf çıkarımı kullanır [6]. İkinci tip yaklaşımlarda, segmentasyon olmadan anormal PCG kayıtları saptanır [7,8].

Bu çalışmada PCG sınıflandırmasına yönelik birinci tip yaklaşımı takip edildi. Bunun arkasındaki ana motivasyon, segmentasyona olan bağımlılığı ortadan kaldırmak ve hesaplama yükünü azaltmaktır. Bu çalışmanın ana katkısı, zaman-frekans özelliklerinin ayrıntılı bir şekilde araştırılması (Bölüm 2.A) ve etkili sinir ağı tasarımıdır (Bölüm 2.B). Önerilen yaklaşım, en büyük kamusal kalp sesi veri tabanından birinde değerlendirilmektedir [9]. Sonuçlar Bölüm 3'te ve Bölüm 4'te tartışılmıştır ve makaleyi sonuçlandırmakta ve gelecekteki araştırmalar için konuları önermektedir.

II. METOT

Bu çalışma için, 3454 PCG etiketli kayıtlar (eğitim ve doğrulama setleri dahil) Kalp Ritmi Sınıflandırma Yarışmasından sağlanmaktadır [9]. Önerilen özellik çıkarma ve sınıflandırma yaklaşımları bundan sonra tartışılacaktır.

A. Özellik Çıkarma

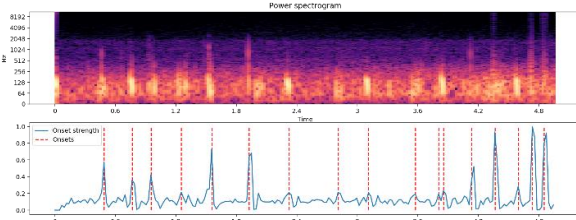
Bu çalışmanın ilk aşamasında, zaman, frekans ve zaman-frekans alanlarında 40 özellik çıkarılmıştır. Daha sonra, ardışık ileri seçim arama algoritmasının [12] kullanıldığı bir sargı temelli özellik seçim şeması [11] kullanılarak 18 özellik içeren bir alt küme seçildi.

Seçilen özellikler aşağıdaki gibi 3 kategoriye ayrılabilir:

- (1) Kök Ortalama Kare Enerjisi (RMSE): PCG sinyallerinin tam-bant enerjisini hesaplamada kullanılmıştır.

$$E_j = \frac{1}{N} \cdot \sum_{i=(j-1).N+1}^{j.N} x^2(i)$$

- (2) Başlangıç Algılama: PCG sinyallerinin başlangıcını algılamak için kullanılır.



Şekil 1. Başlangıç Algılama Gösterimi

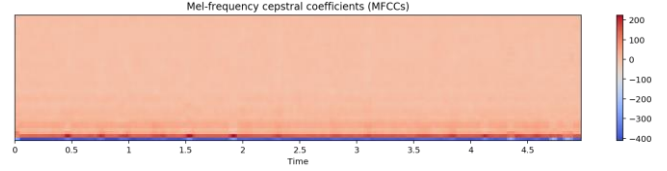
- (3) Mel Frekans Cepstral Katsayıları (MFCC) tabanlı özellikler: Her PCG sinyalinin MFCC'leri, 10 ms çakışma ile 25 ms çerçeve süresi için 13 katsayı parametresine göre hesaplanır. MFCC'lerin çıkarılması her bir çerçeve için 13 katsayı ile sonuçlanır, burada i ve j sırasıyla özellik ve çerçeve sayısıdır.

$C_{i,j}$ hesaplandıktan sonra, üç özellik aşağıdaki gibi çıkarılır:

$$AV_{MFCC} = \frac{1}{N} \sum_{j=1}^N \min_{i \in I} C_{i,j}, \quad j = 1, 2, \dots, N$$

$$\mu_r^{max,MFCC} = E \left(\max_{i \in I} C_{i,j} - \mu_j \right)^r$$

$$\mu_r^{skew,MFCC} = E \left(\text{Skew } C_{i,j} - \mu_j \right)^r$$



Şekil 2. Mel-Frekans Septral Katsayıları

Burada $\min_{i \in I} C_{i,j}$, $\max_{i \in I} C_{i,j}$ ve $\text{Skew } C_{i,j}$ her bir matris kolonunun minimum, maksimum ve eğimini temsil eder. μ ortalama anlamına gelir ve $r = 2$ 'dir.

B. Sınıflandırma

Bu çalışmada, önerilen sınıflandırma algoritması ile kaliteli sinyaller arasında Normal, Hırıltılı, Ekstrasistol ve Yapay kategorilerinde sınıflandırma yapılmaktadır. Aşağıda bu sınıflandırıcıların yapısı ayrıntılı olarak açıklanmaktadır. Topluluk tabanlı sınıflandırma sistemleri bir dizi sınıflandırıcı oluşturur ve daha iyi bir sınıflandırma performansı elde etmek için bu sınıflandırıcıların sonuçlarını birleştirerek yeni örnekleri sınıflandırır.

Tekrarlayan sinir ağları (RNN) çeşitli görevlerde umut verici sonuçlar üretmek için kullanılmıştır. Bir RNN, geçmiş bilgilerine dayanan bir bellek tutar ve bu da modelin uzun mesafe özelliklerinde koşullandırılmış akım çıkışını tahmin etmesini sağlar.

Bir giriş katmanı, t zamanındaki özellikleri temsil eder ve özellik boyutuyla aynı boyutluluğa sahiptir. Bir çıkış katmanı, t zamanında etiketler üzerindeki bir olasılık dağılımını temsil eder. Etiketlerin boyutuyla aynı boyutlara sahiptir. İleri beslemeli ağ ile karşılaştırıldığında, bir RNN önceki gizli durum ile mevcut gizli durum (ve dolayısıyla tekrarlayan katman ağırlığı parametreleri) arasındaki bağlantıyı sunar. Bu tekrarlayan katman geçmiş bilgilerini saklamak için tasarlanmıştır. Gizli ve çıktı katmanlarındaki değerler aşağıdaki gibi hesaplanır:

$$\mathbf{h}(t) = f(\mathbf{U}\mathbf{x}(t) + \mathbf{W}\mathbf{h}(t-1)),$$

$$\mathbf{y}(t) = g(\mathbf{V}\mathbf{h}(t)),$$

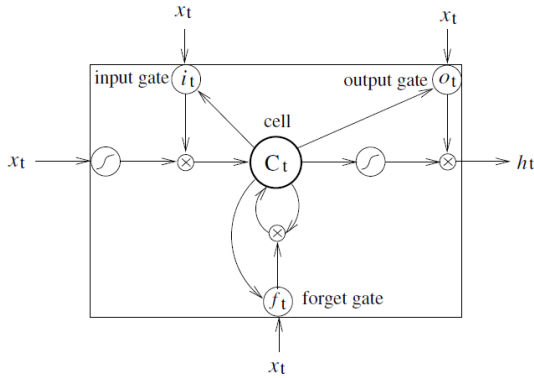
\mathbf{U} , \mathbf{W} ve \mathbf{V} 'nin bağlantı ağırlıkları eğitim zamanında hesaplanacaktır. $f(z)$ ve $g(z)$ sigmoid ve softmax aktivasyon fonksiyonları aşağıdaki gibidir.

$$f(z) = \frac{1}{1 + e^{-z}},$$

$$g(z_m) = \frac{e^{z_m}}{\sum_k e^{z_k}}.$$

Uzun-Kısa Vadeli Hafıza ağları, gizli katman güncellemelerinin, amaca yönelik oluşturulan bellek hücreleriyle değiştirilmeleri dışında, RNN'lerle aynıdır.

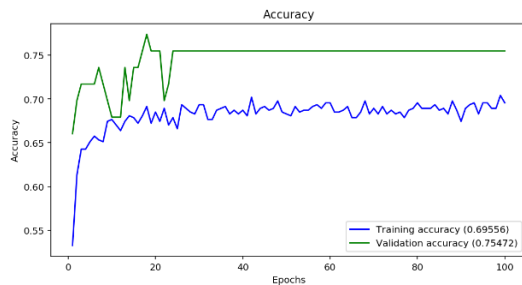
Sonuç olarak, verilerdeki uzun menzilli bağımlılıkları bulma ve kullanma konusunda daha iyi sonuç vermektedir v bu çalışmada LSTM modellemesi kullanılmıştır. Şekil 1'de tek bir LSTM hafıza hücresini göstermektedir.



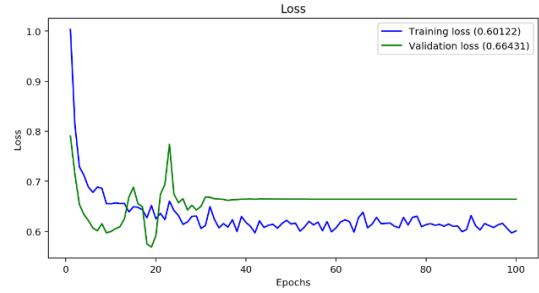
Şekil 3. Uzun-Kısa Vadeli Hafıza Hücresi [17]

III. ÇIKTILAR VE TARTIŞMA

PCG sınıflandırması için önerilen iki kombinasyon kuralını karşılaştırmak için deneyler yaptım. Her bir kombinasyon kuralının performansı, Hassasiyet (S_e), Özgüllük (S_p) ve Skor (S_c) değerlerini hesaplayarak değerlendirildi.

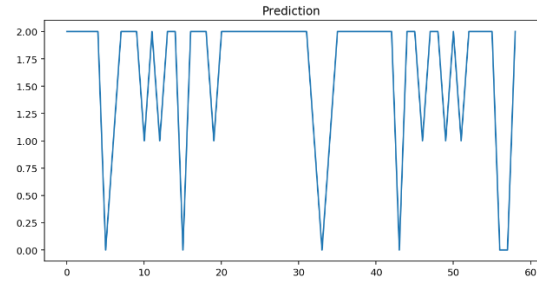


Şekil 4. Keras Doğruluk Geçmişi



Şekil 5. Keras Kayıp Geçmişi

Toplamda 29,411 parametre eğitildi ve yapılan test sonuçlarına göre %79,0 doğruluk oranına sahip bir model ortaya çıkmıştır.



Şekil 6. Tahmin Testi

Gelecekte, daha fazla performans iyileştirmesi için, buna göre seçilecek özelliklerle iki bağımsız sınıflandırma senaryosu tasarlanacaktır.

IV. SONUÇ

Bu çalışmada, doğrulama skorları test skorlarıyla tutarlıdır, bu nedenle aşırı uyum minimumdur. Genel doğruluk, anormalliklerin nedeni ve türü hakkında daha iyi bir kavrayışla geliştirilebilir ve bir anormallik gözlenirse bile, hangi PCG bölgesinin tabi tutulduğunu bilmek yararlı olacaktır.

Üstelik gürültü, ayrılmak imkânsız frekans alanı, sınıflandırmayı çok zorlaştırır, bu nedenle daha iyi ele alınmalıdır. Kalite tespiti için belirli özellikler ve ek sınıflandırıcı tasarlamak sistem performansını artırabilir.

TEŞEKKÜR

Bu çalışmayı gerçekleştirebilmek için gerekli olan teorik altyapıyı bana sağladıkları için Doç. Dr. Savaş Şahin ve Doç. Dr. Yalçın İşler'e sonsuz teşekkürlerimi sunuyorum.

KAYNAKÇA

- [1] McConnell ME, *Pediatric Heart Sounds*, Springer Science & Business Media, 2008.
- [2] Balili CC, Sobrepena M, Naval PC, *Classification of heart sounds using discrete and continuous wavelet transform and random forests*. In 3rd IAPR Asian Conference on Pattern Recognition, Kuala Lumpur, 2015.
- [3] Moukadem A, Dieterlen A, Brandt C. *Shannon Entropy based on the S-Transform Spectrogram applied on the classification of heart sounds*. In IEEE International Conference on Acoustics, Speech and Signal Processing, Vancouver, BC, 2013.
- [4] Zhang D, He J, Jiang Y, Du M. *Analysis and classification of heart sounds with mechanical prosthetic heart valves based on Hilbert-Huang transform*. International Journal of Cardiology 2011;151(1):126–127.
- [5] Kao WC, Wei CC. *Automatic phonocardiograph signal analysis for detecting heart valve disorders*. Expert Systems with Applications 2011;38(6):6458–68.
- [6] Huiying L, Sakari L, Iiro H. *A heart sound segmentation algorithm using wavelet*. In 19th international Conference IEEE/EMBS, Chicago, IL, 1997.
- [7] Yuenyong S, Nishihara A, Kongprawechnon W, Tungpimolrut K. *A framework for automatic heart sound analysis without segmentation*. In BioMedical Engineering OnLine, 2011; 10-13.
- [8] Deng SW, Han JQ. *Towards heart sound classification without segmentation via autocorrelation feature and diffusion maps*. Future Generation Computer Systems 2016; 60:13–21.
- [9] Heart Sounds Classification Challenge, <http://www.peterjbentley.com/heartchallenge/>
- [10] Liu C, Springer D, Li Q, et. al. *An open access database for the evaluation of heart sound algorithms*. Physiological Measurement 2016;37(11).
- [11] Kohavi R, John GH. *Wrappers for feature subset selection*. Artif Intell1997; 97:273–324.
- [12] Whitney AW. *A direct method of nonparametric measurement selection*. IEEETrans Comput 1971; 20:1100–3. Whitney AW. *A direct method of nonparametric measurement selection*. IEEE Trans Comput 1971; 20:1100–3.
- [13] Hagan MT, Menhaj MB. *Training feedforward networks with the Marquardt algorithm*. IEEE transactions on Neural Networks 1994;5(6):989-93.
- [14] MacKay DJC. *Bayesian interpolation*. Neural Computation 1992;4(3):415–447.
- [15] Parmanto B, Munro PW, Doyle HR. *Improving committee diagnosis with resampling techniques*. Advances in neural information processing systems 1996:882-8.
- [16] Polikar R. *Ensemble based systems in decision making*. Circuits and Systems Magazine, IEEE 2006;6(3):21–45.
- [17] Zhiheng Huang, Wei Xu, and Kai Yu. 2015. *Bidirectional LSTM-CRF models for sequence tagging*. CoRR, abs/1508.01991.