.

‘‘Deep Learning Algorithms for Estimation of Demographic and Anthropometric Features from Electrocardiograms’’ Çalışma Değerlendirmesi

Dr. Mehmet Hakan UZUN

‘‘Deep Learning Algorithms for Estimation of Demographic and Anthropometric Features from Electrocardiograms’’ Çalışma Değerlendirmesi

Hazırlayan: Dr. Mehmet Hakan UZUN
Afyonkarahisar Sağlık Bilimleri Üniversitesi Tıp Fakültesi,

1. **Çalışmanın Adı:** Deep Learning Algorithms for Estimation of Demographic and Anthropometric Features from Electrocardiograms1

2. **Çalışmanın Yayınlandığı Dergi:** Journal of Clinical Medicine

3. **Çalışmanın Yayınlandığı Tarih:** 12 Nisan 2023

4. **Çalışmanın Sponsoru:** Kore Hükümeti Ministry of Science and ICT (Bilim Bakanlığı Bilgi ve İletişim Teknolojileri) tarafından çalışma tasarlanmış ve sponsorluk verilmiştir.

5. **Çalışmanın Amacı:** Hastalardan çekilen elektrokardiyografilerinden (EKG) elde edilebilecek demografik ve antropometrik veriler sayesinde yeni geliştirilebilecek fizyolojik belirteçler vasıtasıyla kronolojik yaşa göre daha hassas bir şekilde bireysel sağlık durumunun gösterilebilmesi ve EKG’den elde edilen bilgilerin sisteme girilmesi esnasında insan kaynaklı hataların en aza indirilebileceği öngörülmektedir. Bu amaçla EKG üzerinden yaş, cinsiyet, ABO kan grubu ve vücut-kitle endeksi (VKİ) gibi demografik ve antropometrik özellikleri tahmin edebilecek bir derin-öğrenme modelinin (DLA) geliştirilerek bu model ile elde edilen sonuçların sensitivitesi ve spesitivitesinin değerlendirilmesi amaçlanmıştır.

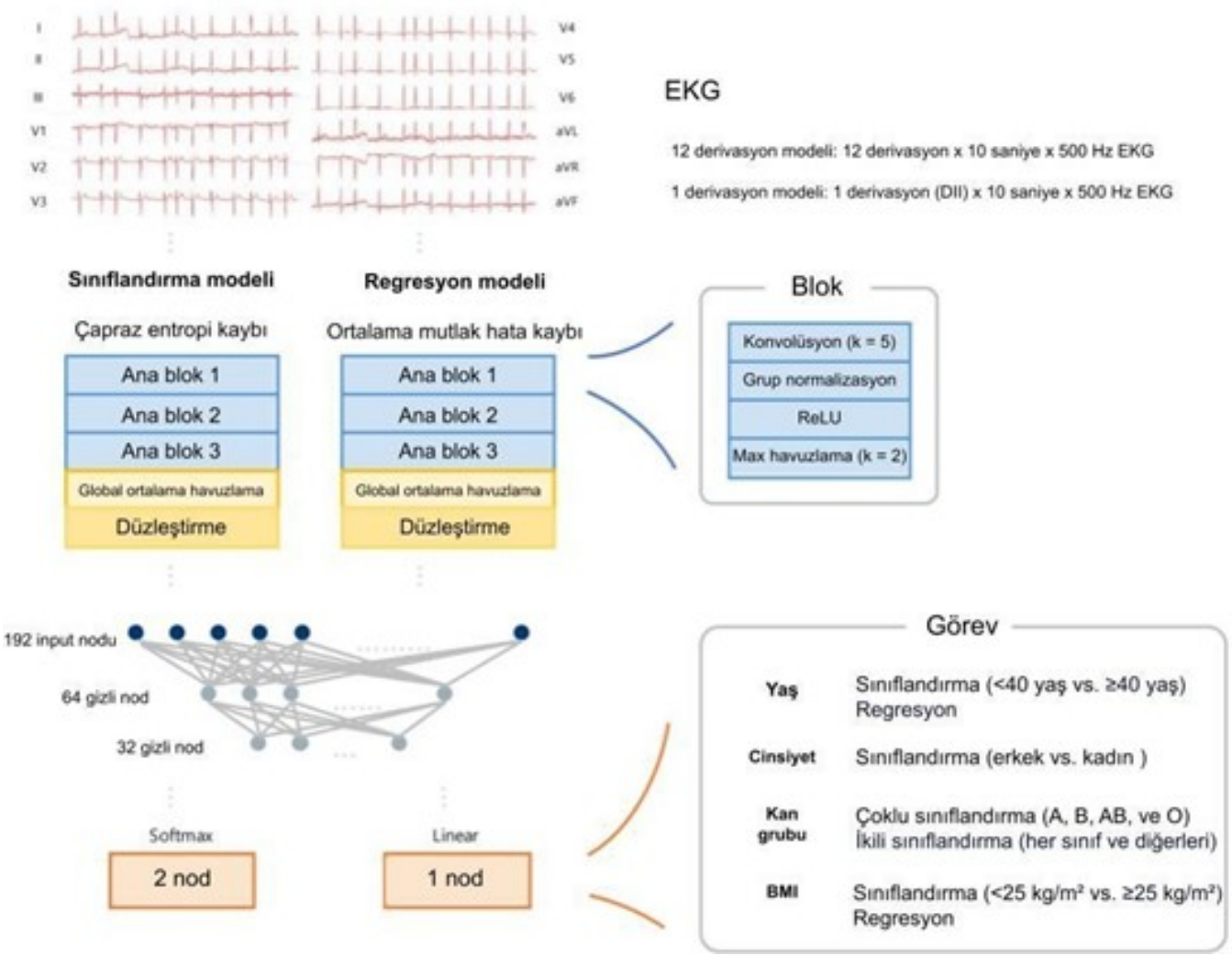
6. **Çalışmanın Dizaynı:** Ekim 2010 ile Şubat 2020 tarihleri arasında Yonsei Üniversitesi Wonju Severance Christian Hastanesi’ne başvuran, 18 yaş ve üzerinde olup bilinen kardiyak hastalığı olan veya kardiyolojik hastalığı olmayıp genel muayene için başvuran hastalar çalışmaya dahil edilmiştir. Hastaların EKG’leri MUSE Kardiyoloji Sistemi’nden retrospektif olarak taranmarak veri setine aktarılmıştır. DLA modeli ile tahmin edilen yaş, cinsiyet, ABO kan grubu ve VKİ değerlerinin korelasyonunun yapılması için hastaların bu demografik ve antropometrik değerleri, hastane veri sisteminden retrospektif olarak taranmıştır. Hastanın birden fazla EKG’si varsa ilk çekilen EKG kullanılmış olup çoklu EKG çekimi mevcut ise antropometrik ölçümlerin alındığı tarihe en yakın olarak çekilmiş EKG kullanılmıştır.

   DLA model eğitimi için veri seti, veriler arası örtüşme olmayacak şekilde eğitim, doğrulama ve test setleri 3’e ayrılmıştır. Eğitim setine hastaların %40’ı, doğrulama ve test setlerine ise %30’ar olarak dağıtılmıştır. Ayrılan setlerin demografik ve antropometrik dağılımları birbirine eşit veya yakın tutulmuştur. Oluşturulacak DLA modelinin parametrelerini belirlemek ve ayarlamak için eğitim ve doğrulama setleri kullanılmıştır. Doğrulama setinde en yüksek performansı gösteren modeller, test seti kullanılarak nihai performans için değerlendirilmiştir.

   Yaş, cinsiyet, ABO kan grubu ve VKİ olarak 4 değişken, DLA modelleri tarafından tahmin edilmiştir. Tahmin aşamasında sınıflama ve regresyon olarak iki bağımsız tahmin modeli kullanılmıştır. Sınıflandırma modellerinin varyasyonlara uygun olarak sınıflandırabildiğini doğrulamak için ikili (binary) sınıflandırma yapılmıştır. Tahmin için ise regresyon modeli kullanılmıştır. Yaş için ikili sınıflandırmada 40 değeri, VKİ’de ise 25 değeri cut-off değeri olarak kullanılmıştır. Sınıflandırma sonrası regresyon modeli ile doğrudan bir değer tahmini yapılmıştır. Cinsiyet tahmininde ise erkek veya kadın olarak sınıflandırma modeli kullanılmıştır. ABO kan grubu tahmininde ise çok sınıflı sınıflandırma modeli kullanılarak A, B, ABO veya O kan grubu olarak hastalar sınıflandırılmıştır.

   DLA modeli; konvolüsyonel nöral ağ (convolutional neural network [CNN]) modeli temel alınarak oluşturulmuştur. Bu model sıklıkla görüntüleme alanında kullanılmakta olup modeli oluşturan katmanlar, EKG’yi tarayacak şekilde programlanmıştır. EKG verisinin tek boyutlu (1D) bir veri olması sebebiyle değerlendirilen örüntünün (pattern) diğer görüntüleme modellerine kıyasla daha küçük bir ebatta olmasına bağlı olarak elde edilen verilerin büyüklüğü göreceli olarak daha azdır. Modelin konvolüsyonel ve havuzlama katmanları ile EKG’lerden verilerin özellikleri elde edilmiştir. Model, üç primer blok ve tam bağlantılı katmandan oluşturuldu. Birincil blokların her biri 4 ayrı bloktan meydana gelmekte olup bu bloklar 1D konvolüsyon katmanı (Kernel ebatı 5), grup normalizasyon katmanı, ReLU (rektifiye öğrenme ünitesi) aktivasyon fonksiyonu, ve 1D max havuzlama katmanından meydana gelmektedir. 12 derivasyonlu EKG’yi analiz eden modellerde 12 derivasyon x 10 saniye x 500 hz’lik EKG verisi ana bloğa input olarak konulmuş ve

sonucunda EKG verisi elde edilmiştir. Buna karşılık 1 derivasyonlu EKG'yi analiz eden modellerde 1 derivasyon x 10 saniye x 500 hz'lik EKG verisi analiz edilmiştir. Elde edilen analiz verileri havuzlama ve düzleştirme katmanlarında işlenerek 192 x 1 formatında bir özellik haritası haline getirilmiştir. Bu harita, 3 katmanlı ve birbirleriyle bağlantılı bir ana katmandan geçirilmiş olup her ara katmanda 192, 64 ve 32 nod bulunmaktaydı. Sınıflandırma modeli için outputta Softmax olarak 2 nod, regresyon modelinde ise Linear olarak 1 nod bulunmaktaydı (Şekil 1). Nihai input olarak ise sınıflandırma modelinde veri girdisinin hangi sınıfa ait olduğuna dair olasılığı sağlamakta, regresyon modeli ise doğrudan tahmin edilen değeri sağlamaktaydı. Çapraz entropi ve ortalama mutlak hata kaybı (mean absolute error [MAE]), kayıp fonksiyonları olarak kullanılmıştır.



Şekil 1. Modellerin şematik görünümü. Konvolüsyonel nöral ağın yapısal şeması. Şekil, orjinal metinden uyarlanmıştır.

Hiperparametreler tekrarlanan deneyler ve sistem araştırmasıyla belirlenmiştir. Grup ebatı, öğrenme hızı ve max havuzlama katmanının kernel boyutu potansiyel olarak aday parametreler olarak kabul edilmiştir. Grup ebatlarının 64 ile 512 arasında, öğrenme hızının ise 0.0001 ile 0.1 arasında olması değerlendirilmiş olup DLA modelinde konvolüsyon katmanının kernel boyutunun değeri, grup araması ile belirlenen parametrelerle sabitlenmiştir. Eş zamanlı olarak max havuzlama katmanının kernel boyutunun 2 ile 5 arasında olması gerektiği düşünülerek en yüksek AUROC (area under receiver operating characteristics) değerinin, doğrulama seti için en uygun olduğu düşünülen değere tanımlanması yapılmıştır.

Sınıflandırma modellerinin performansı, AUROC ile değerlendirilmiştir. DLA modelinin tahmini ile verilerin gerçek sonuçları karşılaştırılarak model sonuçları gerçek pozitif, yanlış negatif, gerçek negatif ve yanlış pozitif olarak sınıflandırılmıştır. Hassasiyet, sensitivite, spesifite, pozitif prediktif değer ve negatif prediktif değerler; cut-off değerlerinin ROC eğrisinin (receiver operating characteristics) maksimal Youden indeksi (J) ile hesaplanarak elde edilmiştir.

Gözlemlenen ve tahmin edilen değerler arasındaki farkı belirlemek amacıyla regresyon modellerinde kullanılmak üzere MAE ile birlikte Pearson korelasyon katsayısı (Pearson R) ve sınıf içi korelasyon katsayısı ölçülmüştür.

Çalışmanın veri setinin standart 12 derivasyonlu EKG'ler üzerinden oluşturulması sebebiyle 12 derivasyonlu EKG tabanlı modellerin performansı, çalışmanın ana sonucu olarak belirlenmiştir. Tüm istatistiksel veriler nokta tahimini ve %95 güven aralığı (CI) ile belirtilmiş olup verilerin analizi ve görselleştirilmesi Python 3.8.5 (Python Software Foundation) ile yapılmıştır.
Eğitilen CNN modelinin aldığı kararların yorumlamasını yapmak için gradiyent ağırlıklı sınıf aktivasyon haritalaması (Grad-CAM) kullanılmıştır. 12 derivasyonlu EKG'nin her kanalı, 2 boyutlu bir düzleme yerleştirilerek görüntüye dönüştürülmüş olup daha sonra tüm kanallar bu görüntülere eklenerek Grad-CAM ile kullanılacak nihai görüntüler elde edilmiştir. Modelin 1D CNN modeli olması sebebiyle ortaya çıkan ısı haritası, tek satırlı dizi olarak oluşturulmuştur. İnterpolasyon ile görüntüler yeniden ebatlandırıldıktan sonra, daha önceden oluşturulan görüntüler ile birleştirilerek Grad-CAM oluşturulmuştur. Grad-CAM ısı haritasındaki kırmızı-sarı alanların CNN'in kararını önemli ölçüde etkilediği görülmüş olup değerlendirilen bölge maviye yakınlaşmasının sonuca olan etkisi azalmaktaydı. Bu çalışmada bulunan 3 sınıflandırma grubu için Grad-CAM kullanılarak açıklayıcı model analiz edilmiştir.

7. **Çalışmaya Dahil Edilen Hastaların Demografik Özellikleri:**
    - Toplam hasta grubu:
        - Hasta sayısı 124,415
        - Yaş aralığı 18-108 arası
        - Median yaş 55
        - Kadın hasta sayısı 60,835 (%48,90), erkek hasta sayısı 63,580 (%51,1)

        - Eğitim seti için:
            - Hasta sayısı **49,762**
            - Ortalama yaş 55,25
                - 40 yaş altı %19,87
                - 40 yaş ve üzeri %80,12
            - Cinsiyet %51.10 erkek, %48,89 kadın
            - Kan grubu:
                - A grubu %31,97
                - B grubu %28,86
                - AB grubu %28,30
                - 0 grubu %10,85
            - VKİ:
                - ≥25 kg/m2 %39,84
                - <25 kg/m2 %60,15
        - Doğrulama seti için:
            - Hasta sayısı **37,324**
            - Ortalama yaş 55,25
                - 40 yaş altı %19,87
                - 40 yaş ve üzeri %80,12
            - Cinsiyet %51,10 erkek, %48,89 kadın
            - Kan grubu:
                - A grubu %31,97
                - B grubu %28,86
                - AB grubu %28,30

                1. 0 grubu %10,85
            1. VKİ:
                1. ≥25 kg/m2 %39,95
                2. <25 kg/m2 %60,04

        1. Test seti için:
            1. Hasta sayısı **37,329**
            2. Ortalama yaş 55,22
                1. 40 yaş altı %17,26
                2. 40 yaş ve üzeri %80,12
            3. Cinsiyet %51,10 erkek, %48,89 kadın
            4. Kan grubu:
                1. A grubu %31,97
                2. B grubu %28,86
                3. AB grubu %28,30
                4. 0 grubu %10,85
            5. VKİ:
                1. ≥25 kg/m2 %39.78

2. <25 kg/m2 %60.21

1. **Çalışmaya Dahil Edilme Kriterleri:**
    1. MUSE Kardiyoloji Sistemi'ne EKG'leri yüklenmiş olması
    2. Yaş, cinsiyet, ABO kan grubu ve VKİ verilerinin sisteme kaydedilmiş olması

1. **Çalışmadan Dışlama Kriterleri:**
    1. 500 Hz dışında örnekleme hızıyla çekilmiş EKG'ye sahip olma
    2. Eksik derivasyonlu kayıt gibi eksik veriyle sisteme yüklenen EKG'ye sahip olma
    3. 18 yaşın altında olma
    4. Yaş, cinsiyet, ABO kan grubu veya VKİ verilerinin sisteme kaydedilmemiş olması
1. **Çalışmanın Sonuçları:**

Normalizasyon veya gürültü azaltıcı gibi sinyal işlemeli EKG kullanımında sonuçların anlamlı bir şekilde değişmediğinin görülmesi üzerine çalışmanın tüm aşamalarında sinyal işlemesiz EKG'ler kullanılmıştır. Optimal grup sayısı 512 ve başlangıçtaki öğrenme hızı 0.001'di. DLA'nın max havuzlama kerneli 2'ye ayarlanmıştır. Ancak hiperparametre ayarlarında performans farkı minimum düzeyde izlenmiştir. Adam optimizer kullanılarak 50 döngü (epoch) boyunca model eğitilmiştir. Doğrulama setinde en yüksek AUROC değerine ulaşarak en iyi performansı gösteren sınıflandırma modeli, öğrenme süreci boyunca oluşan en iyi model olarak seçilmiştir. Regresyon modeli için ise en düşük MAE değerine sahip olan seçilmiştir. Elde edilen en iyi model, çalışmanın sonunda test seti üzerinden son doğrulama aşamasından geçmiştir.

Test seti üzerinden yaş tahiminini en iyi yapan modelin test setinde AUROC değeri 0,923 (%95 CI, 0,922 – 0,923) olarak izlenmiş olup maksimum Youden indeksi sensitivitesi %81,96 (%95 CI, %81,55–82,37) ve spesifitesi %86,67 (%95 CI, %86,27–87,08) olarak gözlendi. Bir derivasyonlu EKG ile çalışan model, 12 derivasyonlu EKG ile çalışan modele kıyasla önemli düzeyde daha düşük performans göstermiştir (1D modeline ait AUROC 0,816, CI %95, 0,814 – 0,818).

Cinsiyet tahminini en iyi yapan modelin test setinde AUROC değeri 0,947 (%95 CI, 0,945 – 0,948) olarak sonuçlanmıştır. Maksimum Youden indeksi sensitivitesi %87,42 (%95
CI, %85,92–88,91), spesifitesi %86,25 (%95 CI, %85,07–87,42) olarak sonuçlanmıştır. Bir derivasyonlu EKG ile çalışan model, 12 derivasyonlu EKG ile çalışan modele kıyasla önemli düzeyde daha düşük performans göstermiştir [AUROC, 0,947 (%95 CI, 0,945–0,948)].
ABO kan grubu tahmininde kullanılan modelin çok sınıflı sınıflama hassasiyeti
%31,98 (%95 CI, %31,98–31,98) olarak izlenmiştir. İkili sınıflandırma modelinin ortalama AUROC eğrisi 0,501 (%95 CI, 0,496–0,501) olup ortalama sensitivitesi %56,12 (%95 CI,
%4,96–107,27) ve ortalama spesifitesi %44,03 (%95 CI, – %7,31–95,37) olarak izlenmiş olup tahmin gücü genel olarak zayıftır. 12 derivasyonlu EKG ile çalışan ana modelin tahmin gücünün düşük olması sebebiyle test, 1 derivasyonlu EKG ile tekrarlanmamıştır.
En iyi performanslı VKİ sınıflama modelinin AUROC eğrisi 0,765 (%95 CI, 0,763–0,766) olup maksimum Youden indeks sensitivitesi %70,00 (%95 CI, 67,80–72,20%),
spesifitesi %69,82 (%95 CI, 67,30–72,35%) olarak izlendi. 1 derivasyonlu EKG ile çalışan
modelin performansı [AUROC, 0,765 (%95 CI, 0,763–0,766)], 12 derivasyonlu EKG ile çalışan modelin performansına kıyasla (AUROC, 0,633 [%95 CI, 0,629–0,639) belirgin düzeyde daha düşük izlenmiştir.

11) Çalışma Hakkında Yorumlar:

Bu çalışmada DLA kullanılarak hastaların yaş, cinsiyet ve VKİ'leri gibi antropometrik bilgilerinin EKG üzerinden tahmin edilebileceği öne sürülmüştür. ABO kan grubu tahmininde ise DLA kullanımının uygun sonuçlar vermediği görülmüştür. Aynı zamanda bu çalışmada 1 derivasyonlu EKG ile antropometrik bilgilerin tahmin hassasiyetinin, 12 derivasyonlu EKG'nin hassasiyetinden daha düşük olmasına rağmen DLA kullanılarak güvenilir bir şekilde tahmin edilebildiği belirtilmiştir. Daha önceki çalışmalarda yaşın ilerlemesi ile birlikte EKG'de PR ve QT interval uzaması, QRS süresinin kısalması ve farklı morfolojilerde T dalgalarının görülebilmesi sayesinde EKG'nin yaş tahminine yardımcı olabileceği öne sürülmüştür (2). Bununla birlikte diğer bir çalışmada ise cinsiyet farklılığının EKG üzerinde etkisi olduğu gösterilmiştir (3, 4). VKİ'nin de EKG üzerinde değişikliğe sebep olabileceğini gösteren çalışmalar mevcuttur (5). Ek olarak bu çalışmada 1 derivasyonlu EKG üzerinden de DLA kullanarak antropometrik çalışma yapılabildiği gösterilmiştir. ABO kan grubunun ise DLA kullanılarak EKG üzerinden tahmin edilemeyeceği öne sürülmektedir.
Bu çalışmanın yayınlanmasından önce yayınlann ve sadece 12 derivasyonlu EKG'yi analiz eden bir DLA kullanılarak yapılan çalışmalarda yaş sınıflandırması AUROC değeri 0,923 ve cinsiyet sınıflandırması AUROC değeri 0,947 olarak tespit edilmiş olup bu çalışmada kullanılan modelin AUROC değerlerinden daha iyi performans sergilediği belirtilmektedir (6,7,8). Önceki çalışmalara kıyasla bu çalışmanın performans farkının; veri seti farklılığı ve çalışmadaki hastaların büyük çoğunluğunun Asya kıtası kökenli olmasından dolayı kaynaklanabileceği belirtilmektedir.
Güncel teknolojik gelişmeler ışığında birçok farklı marka, 1 veya 2 derivasyonlu EKG kaydı alarak EKG analizi yapabilen akıllı saatleri piyasaya sürmüş olduğundan ve elde edilen bu EKG'lerin klinik pratikte tanısal bir değeri olabileceğine dair güncel literatürde gün geçtikçe artan veriler olmasından dolayı çalışmada 1 derivasyonlu EKG üzerinden antropometrik bilgilerin DLA kullanılarak doğru bir şekilde elde edilip edilemeyeceği çalışılmıştır. 12 derivasyonlu EKG kullanılarak yapılan tahminlere kıyasla tahmin hassasiyetinin daha düşük olmasına rağmen, kabul edilebilir sınırlar içerisinde tahmin hassasiyetine sahip olduğu görülmüştür.

Hastaların EKG'lerinde tespit edilen antropometrik ölçümlerin kardiyovasküler sağlık durumunu belirlemede kronolojik yaştan daha doğru bir yönlendirme sunabilecek kardiyovasküler biyolojik yaş tayin etmeye doğru bir adım olabileceği, yazarlar tarafından düşünülmüştür. Çalışmanın kısıtlılıkları arasında ise tek bir veri setinin kullanımı sebebiyle harici doğrulama analizleri yapılamaması mevcuttur. Antropometrik olarak farklı özellikteki hasta popülasyonlarından elde edilen veri setlerine aynı DLA'nın kullanılarak analizi yapıldığında ise sonuçlarda tutarsızlıkların oluştuğu görülmekte olup farklı antropometrik özelliklerin beraber toplandığı daha geniş veri setleri ile ileri araştırmaların yapılması gerekmektedir. Son olarak DLA'nın EKG'ye ait hangi parametrelerin hangi detaylarıyla tahminde bulunduğuna dair belirsizlikler mevcut olup yapay zeka algoritmalarının parametre seçim tercihlerinin daha iyi açıklanabilir olması için yapay zeka algoritmaları üzerinde daha fazla çalışmaya ihtiyaç olduğu unutulmamalıdır.

Kaynaklar

1. Ryu, J.S.; Lee, S.; Chu, Y.; Koh, S.B.; Park, Y.J.; Lee, J.Y.; Yang, S. Deep Learning Algorithms for Estimation of Demographic and Anthropometric Features from Electrocardiograms. J. Clin. Med. 2023, 12, 2828. https://doi.org/ 10.3390/jcm12082828
2. Bachman, S.; Sparrow, D.; Smith, L.K. Effect of aging on the electrocardiogram. Am. J. Cardiol. 1981, 48, 513–516.
3. Surawicz, B.; Parikh, S.R. Prevalence of male and female patterns of early ventricular repolarization in the normal ECG of males and females from childhood to old age. J. Am. Coll. Cardiol. 2002, 40, 1870–1876.
4. Mieszczanska, H.; Pietrasik, G.; Piotrowicz, K.; McNitt, S.; Moss, A.J.; Zareba, W. Gender-related differences in electrocardiographic parameters and their association with cardiac events in patients after myocardial infarction. Am. J. Cardiol. 2008, 101, 20–24.
5. Hassing, G.J.; van der Wall, H.E.C.; van Westen, G.J.P.; Kemme, M.J.B.; Adiyaman, A.; Elvan, A.; Burggraaf, J.; Gal, P. Body mass index related electrocardiographic findings in healthy young individuals with a normal body mass index. Neth. Heart J. 2019, 27, 506–512.
6. Attia, Z.I.; Friedman, P.A.; Noseworthy, P.A.; Lopez-Jimenez, F.; Ladewig, D.J.; Satam, G.; Pellikka, P.A.; Munger, T.M.; Asirvatham, S.J.; Scott, C.G.; et al. Age and Sex Estimation Using Artificial Intelligence From Standard 12-Lead ECGs. Circ. Arrhythm Electrophysiol. 2019, 12, e007284.
7. Lima, E.M.; Ribeiro, A.H.; Paixão, G.M.M.; Ribeiro, M.H.; Pinto-Filho, M.M.; Gomes, P.R.; Oliveira, D.M.; Sabino, E.C.; Duncan, B.B.; Giatti, L.; et al. Deep neural network-estimated electrocardiographic age as a mortality predictor. Nat. Commun. 2021, 12, 5117.
8. van der Wall, H.E.; Hassing, G.-J.; Doll, R.-J.; van Westen, G.J.; Cohen, A.F.; Selder, J.L.; Kemme, M.; Burggraaf, J.; Gal, P. Cardiac age detected by machine learning applied to the surface ECG of healthy subjects: Creation of a benchmark. J. Electrocardiol. 2022, 72, 49–55.