

## Sinaptik Etkinlik Fonksiyon Tabanlı Sızdıran Entegre ve Ateşleme Nöron Modelini Kullanarak İnsan Ses Sinyallerinde Cinsiyet Tespiti

Vedat TÜMEN<sup>1\*</sup>

<sup>1</sup>Bitlis Eren Üniversitesi, Mühendislik-Mimarlık Fakültesi, Bilgisayar Mühendisliği Bölümü  
(ORCID: [0000-0003-0271-216X](https://orcid.org/0000-0003-0271-216X))



**Anahtar Kelimeler:** Ani sivri uçlu ağlar, Cinsiyet tespiti, Derin öğrenme, Sinyal işleme.

### Abstract

Günümüzdeki teknolojik gelişmeler, insanların bir ses sinyalinden konuşmacının cinsiyetini belirlemesi mümkün kılmıştır. Frekans türleri, spektral ve entropi gibi sayısal nitelikli veriler ses sinyallerinin akustik bilgilerini oluşturmaktadır. Son zamanlarda, yüksek başarı oranlarına sahip yapay zekâ tabanlı öğrenme modelleri çeşitli alanlarda ilgi görmeye başladı. Ses sinyalleri üzerinde derin öğrenme modelleri ile ilgili birçok çalışma bulunmaktadır. Bu çalışmada, derin öğrenme modellerinden esinlenerek tasarlanmış ve farklı bir mimari yapısı olan ani sivri uçlu sinir ağları kullanılmıştır. Çalışmada kullanılan veri kümesi, insan konuşmalarını ve seslerini içeren akustik bilgiye dayalı parametrelerden oluşmaktadır. Belirlenen bir veri seti kullanılarak ani sivri uçlu sinir ağı modeli eğitilmiş ve cinsiyet tespitinin gerçekleştirilmesi sağlanmıştır. Önermiş olduğumuz bu çalışmada sonuç olarak, sınıflandırma sürecinde %98,84 genel doğruluk başarısı elde edilmiştir. Bu çalışmada gerçekleştirilen deneysel analizler ile ani sivri uçlu sinir ağı modelinin başarılı bir şekilde çalıştırıldığı, yüksek başarımlar elde edildiği gözlemlenmiştir.

## Gender Determination in Human Voice Signals using Synaptic Efficacy Function-based Leaky Integrate and Fire Neuron Model

**Keywords:** Spiking network, Gender detection, Deep learning, Signal processing

### Abstract

Today's technological advances have made it possible for people to determine the gender of the speaker from an audio signal. Numerical data such as frequency types, spectral and entropy constitute acoustic information of audio signals. Recently, artificial intelligence-based learning models with high success rates have started to attract attention in various fields. There are many studies on deep learning models on audio signals. In this study, spiked neural networks with a different architectural structure, inspired by deep learning models, were used. The dataset used in the study consists of parameters based on acoustic information including human speech and voices. By using the determined data set, the spiked neural network model was trained and gender determination was achieved. As a result, 98.84% overall accuracy success was achieved in the classification process in this proposed study. With the experimental analyses carried out in this study, it was observed that the spiked neural network model was successfully run and high performances were obtained.

### 1. Giriş

Erkek ve kadın bireylerin cinsiyetlerini fark etmede kullanılan parametrelerden biri de ses sinyalleridir. Sesin kaynağından çıkış anına kadar geçen süreçte, cinsiyet farkındalığını ölçen akustik özellikler vardır.

Bu özellikler, genelde sesin entropisi veya temel frekans ile baskın frekans değerleri arasında farklılıkların oluşmasına sebep olur. Temel frekans değerleri kaynağın boyutu ile ters orantılıdır ve kolaylıkla ölçülebilmektedir [1]. Erkek yetişkin bireylerde temel frekans değeri genelde düşüktür ve

\*Sorumlu yazar: [vtumen@beu.edu.tr](mailto:vtumen@beu.edu.tr)

Geliş Tarihi: 01.12.2021, Kabul Tarihi: 05.04.2022

bayan yetişkin bireylerde ise bu frekans değeri genelde yüksektir. Erkek yetişkinlerde frekans değeri 120 hertz kabul edilirken, bayan yetişkinlerde ise 200 hertz olarak kabul edilir [2]. Fakat temel frekans değerleri her zaman için bu sonucu doğrulamayabilir. Bu sebeple cinsiyet belirlemede diğer akustik özelliklerine de gereksinim duyulmaktadır. Ses sinyalleri üzerinde cinsiyet sınıflandırılması üst düzey bir bilişsel çalışma gerektiren bir durumdur [3,4]. İnsan-makine etkileşimi gün geçtikçe daha da artmaktadır ve bilişim tabanlı cinsiyet tanıma sistemleri karmaşık bir yapıya sahiptir. Bu tür sistemlerin kendi içerisinde problemleri ve sınırlamaları vardır. Sesin karmaşıklığını artıran nedenlerden bazıları sıralanırsa; gürültülü ortamda konuşmak, sıkıştırılmış ses verileri, telefon görüşmeleri, farklı konuşma dillerinin kullanılması şeklinde ifade edilebilir. [5,6]. Neticede, karmaşıklık akustik özelliklerin çıkartılmasını olumsuz etkileyebilmektedir. Bu tür problemleri minimize edebilmek için yapay zekâ tabanlı sistemlere ihtiyaç duyulmaktadır [7].

Literatürde ses sinyallerini işleyerek konuşmacının cinsiyetini belirleyen birçok çalışma yer almaktadır. Bunlardan bazıları incelenirse; G. S. Archana ve ark. çalışmasında ses sinyallerini kullanarak cinsiyet tespitini gerçekleştirmişler. Onlar, veri eğitiminde yapay sinir ağlarını (YSA) ve sınıflandırma sürecinde destek vektör makinelerini (DVM) kullanmışlar. Onların sınıflandırma sürecinde elde ettikleri genel doğruluk oranı %80 olmuştur [8]. S.C. İleri ve ark. deneysel analizlerde konuşma sinyalleri üzerinde cinsiyet tespitini gerçekleştirmişlerdir. Onlar, veri kümesinde mel - frekans spektral katsayıları (MFSK) tekniğini kullanarak özellikler çıkarılmış ve bu özellikleri temel bileşenler analizi yöntemi (TBA) ile daha verimli çıkarımlar elde edilmiştir. Daha sonra bu çıkarımlar her bir özellik değerini normalleştirme adımından geçirerek, DVM yöntemi ile sınıflandırmışlar. Sınıflandırma sürecinde elde ettikleri genel doğruluk başarıları %98.18 olmuştur [9]. David Doukhan ve ark. Fransızca konuşma sinyallerinden oluşan veri kümesini kullanarak cinsiyet tespitini gerçekleştirmişler. Onlar, veri kümesi için ön işlem adımı olarak gauss karışım modelleri (GKM), i-vektörler yöntemini kullanmışlar ve ardından tasarladıkları evrişimsel sinir ağları (ESA) ile veri kümesini eğitip, sınıflandırmışlar. Onların, sınıflandırma sürecinde elde ettikleri en iyi genel doğruluk başarıları %96.52 olmuştur [10]. Ping Wang ve ark. Elektro Ensefalo Grafiği (EEG) sinyallerini kullanarak cinsiyet tespitini gerçekleştirmişlerdir. Onlar, rastgele orman (RO) ve lojistik regresyon (LR) makine öğrenme yöntemlerini

melezleştiren bir model önermişlerdir. Sınıflandırma sürecinde çapraz doğrulama yöntemi ile elde ettikleri en iyi genel doğruluk başarıları %99.82 olmuştur [11].

Çalışmalar incelendiğinde farklı veri türlerinden yapay zeka teknikleri ve algoritmaları kullanarak cinsiyet analizinin yapıldığı tespit edilmiştir. Bu çalışmada, insanların konuşma ve ses sinyallerinden cinsiyet tespitinin gerçekleştirilmesi amaçlanmıştır.

Ani sivri uçlu sinir ağları (ASUSA) tasarlanan bir sinir ağı modelinde nöronların sivri uçları aracılığı ile biyolojik bilgi iletimi sağlamaktadır. Sinir hücrelerinde meydana gelen ani artış stereotipik olaylar olduğu varsayıldığından burada oluşan veri iki temel adımda işlenmektedir. Ani artışın zamansal değeri ve frekans değerlerinin uyarıcı veya engelleyici olup olmadığının tespiti oldukça yüksek performans göstermiştir [12]. ASUSA'lar ayrıca, üçüncü nesil sinir ağları olarak kabul edilir ve kendilerini, büyük ağların çok hızlı gerçek zamanlı simülasyonuna ve düşük bir hesaplama maliyetine izin vererek, beynin davranış ve öğrenme potansiyelini modellemek için en başarılı yaklaşımlardan biri olarak ortaya koymuştur [13]. Daha kapsamlı ASUSA ağları geliştirilerek daha karmaşık görevleri çözebiliriz. Bu nedenle, geleneksel derin öğrenme modellerine göre kıyaslandığında daha karmaşık eğitim yöntemlerinin bilgi işlem platformlarındaki verimsiz artış simülasyonları nedeniyle eğitim süresinde oluşabilecek bir darboğazı önlemektedir. Bu önemli sorunların giderilmesi için farklı bir model yapısı olan ASUSA modeli ile eğitimin başarılı bir şekilde gerçekleştirilmesi hedeflenmektedir.

Bu makalede materyal ve metod bölümünde veriseti ve önerilen yöntem ve parametreler yer almaktadır. Bölüm 3'te bulgular bölümünde çalışmanın deneysel analizlerine ve elde edilen sonuçlara yer verilmiştir. Bu çalışmadaki performans değerleri ile literatürde bulunan yüksek başarımlı sonuçlar Bölüm 4'te karşılaştırılmıştır. Son olarak bölüm 5'te sonuçlar tekrar değerlendirilmiş ve gelecek çalışmalar hakkında öneriler verilmiştir.

## 2. Materyal ve Metod

### 2.1. Veriseti

Veri kümesi, konuşma ve ses sinyallerinin akustik özelliklerine dayalı olarak bir sesi erkek veya kadın olarak tanımlamak için oluşturulmuştur. Veri kümesi, herkesin erişebileceği “.csv” uzantılı dosyadan oluşmaktadır ve 22 akustik özellik içermektedir [14]. Bu akustik özellikler Tablo 1'de verilmiştir.

Veri kümesi, erkek ve kadın konuşmacılardan toplanan 3168 kayıtlı ses örneğinden oluşmaktadır. Ses örneklerinin eşit olarak 1584 âdeti erkek konuşmacılardan ve 1584 âdeti kadın konuşmacılarından oluşmaktadır. Ses kayıtlarında gerçekleşen konuşma dili İngilizcedir. Ses

örneklerinin frekans aralığı (0-280) Hz olarak işlenmiştir ve önceden işlenmiş WAV dosyaları bir CSV dosyasına dönüştürülmüştür [14]. Ek olarak, bu çalışmanın analizinde veri kümesinin %30'u test verisi ve %70'i eğitim verisi olarak ayrıldı.

**Tablo 1.** Veri kümesini oluşturan akustik özellikler

No	Özellik	Açıklama
1	Meanfreq	Ortalama frekans (kHz cinsinden)
2	Sd	Frekansın standart sapması
3	Median	Medyan frekans (kHz cinsinden)
4	Q25	İlk nicelik (kHz cinsinden)
5	Q75	Üçüncü nicelik (kHz cinsinden)
6	IQR	Çeyrekler arası aralık (kHz cinsinden)
7	Skew	Çarpıklık
8	Kurt	Basıklık
9	Sp.ent	Spektral Entropi
10	Sfm	Spektral düzlük
11	Mode	Mod frekansı
12	Centroid	Frekans sentroid
13	Peakf	En yüksek enerjili frekans (Tepe frekansı)
14	Meanfun	Akustik sinyalde ölçülen ortalama temel frekans
15	Minfun	Akustik sinyalde ölçülen minimum temel frekans
16	Maxfun	Akustik sinyalde ölçülen maksimum temel frekans
17	Meandom	Akustik sinyalde ölçülen ortalama baskın frekans
18	Mindom	Akustik sinyalde ölçülen minimum baskın frekans
19	Maxdom	Akustik sinyalde ölçülen maksimum baskın frekans
20	Dfrange	Akustik sinyalde ölçülen baskın frekans aralığı
21	Modindx	Modülasyon indeksi (Temel frekansların bitişik ölçümleri arasındaki birikmiş mutlak farkın frekans aralığına bölünmesiyle hesaplanır)
22	Label	Etiket (Erkek veya Kadın)

## 2.1. Ani Sivri Uçlu Sinir Ağları ile Önerilen Yaklaşım

Ani sivri uçlu sinir ağları (ASUSA), nöron ve sinaptik durumları zaman kavramı ile birleştiren ve doğal sinir yapılarını tek tek inceleyen bir ağ modelidir. ASUSA yapılarındaki nöronların yayılma döngülerinde tetiklenmemesi, bu nöronların öz kalitelerini belirli bir seviyeye yükseltmelerine olanak sağlar. Bir nöron tetiklendiğinde, etrafındaki diğer nöronların sinyal potansiyelini artıran veya azaltan bir durum yaratır. Nöronların aktivasyonunu gerçekleştirmek için belirli bir eşik değeri kullanır. Nöronlardan gelen tetikleme değeri bu eşik değerini aşarsa, tetikleyici denilen olay meydana gelir; aksi durumda nöron için bozulma denilen olay meydana gelir [15]. Belirli periyotlarla çıkan veya azalan tetiklenen çizgiler ve bunların aralıkları, gerçek değer sayısını temsil eder. Bu sayıların yorumlanması, ASUSA modelinin yapısında kullanılan kodlama ve kod çözme

yöntemleri ile yapılır. Ani sivri uçlu (ani darbeli) nöronlar, model içinde meydana gelen her geçici sinyali hesaba katar ancak hemen işlemez. Sinyalin işlenmesi için belirli bir eşik aşılması, sivri uçlu nöronların onu beklemesi ve eşik aşılar aşılmaz sinyali işlenmesi gerekir. Sonuç olarak, artan veya azalan bir sinyal çıkışı üretir. Böylece, artan veya azalan sinyaller kodlama şemaları oluşturulur [15,16]. Sinyallerin hem darbe (sivri uçlu) frekansı hem de darbe aralığı dikkate alınır ve kodlar bir sayı dizisine karşılık gelecek şekilde yorumlanarak kod çözme olayı gerçekleşir. Darbelerin hesaplanmasında ASUSA modeli, sızdıran entegre ve ateşleme (SEA / LIF) yöntemlerini kullanılır. SEA yöntemi için kullanılan integral formül Denklem 1'de verilmiştir. Denklem 1'deki değişkenler; I: akım, V: volt, C: kondansatör, R: direnç ve t: zaman değerini temsil eder [17].

$$(t) - \frac{V_m(t)}{R_m} = C_m \frac{\partial V_m(t)}{\partial t} \quad (1)$$

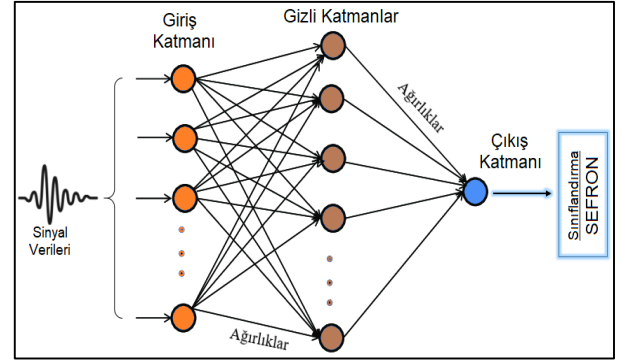
ASUSA, YSA modeline benzer derin katmanlar ve gizli ağ yapılarından oluşmaktadır; fakat YSA modeli gibi sürekli bir çıktı üretmeyen karmaşık yapılardır. Bu model YSA modellerine göre daha masraflıdır. Çünkü ASUSA modeli sinirsel durumları simülasyon olarak yorumlayabilmektedir (Ör; NEST, BindsNet, Brian ve GENESIS) [18]. ASUSA, genel olarak; giriş katmanı, gizli katmanlar ve çıkış katmanından oluşmaktadır ve bu durum Şekil 1'de gösterilmiştir. Burada, giriş katmanı girdi olarak aldığı görüntüyü önceden işler ve özellikler çıkarır. Özellik kümesi daha sonra gizli katmanlara aktarılır ve özellikler gizli katmanlara yayıldıkça, daha karmaşık bir eğitim sürecine girerler. ASUSA karmaşık yapıyı çözümlenebilmek için her bir döngüde ağırlık güncellemeleri yapar ve geri yayılım yaklaşımını kullanır. Sınıflandırma sürecinde, kolayca yorumlanamayan bir ikili çıktı üretir ve her bir nöronun yalnızca yakındaki nöronlarla ilgilendiğini gösteren, sivri uçlu nöronların uzaysal öğrenimini artırır. Bu çalışmada kullanılan sınıflandırma yöntemi, Sinaptik etkinlik fonksiyon tabanlı sızdıran entegre ve ateşleme nöronu (SEFRON)'dur [19].

SEFRON, popülasyon yaklaşımı ile tasarlanmış ve geçici kodlamaların yer aldığı bir sınıflandırıcı modeldir. SEFRON, ani artış sayacı kullanmaz bunun yerine en erken atış süresini kullanır. Ayrıca, giriş ve çıkış nöronları arasında zamanla değişen ağırlıkları (tek bir ağırlık değil) eğitim sürecinde kullanır. SEFRON, öğrenmeyi kolaylaştırmak ve sınıflandırma sürecine katkı sağlaması için spike-timed plasticity (STDP) öğrenme kurallını yapısında kullanır. STDP, ASUSA modelinin giriş ve çıkış nöronlarının zamana dayalı bağlantı ayarlarını gerçekleştirir. En erken atış süresini gerçekleştiren girdi verisi çıkış sınıfı ile etiketlenir [19]. ASUSA, ani bir tepki modeli olarak görünür ve her nöron, periyodu boyunca yalnızca bir kez ateşlenir. Ateşleme oranı dağılımı

Denklem 2'ye göre hesaplanır. Bu denklemde,  $x$  girdi katmanı,  $y = f(x)$  çıktı katmanı,  $N$  nöron sayısıdır.  $R_{max}$ , ateşleme hızını;  $XX$  koordinatındaki nöronlara bitişik nöronların sayısını ve  $\delta$  sabit bir değerini temsil eder. ASUSA' da her gecikme süresi, ön-sinaptik ateşlendiği zaman ile son-sinaptik yükselmeye başladıktan sonraki zaman arasındaki fark olarak tanımlanır. Ek olarak, bu modelde öğrenme hızı, ön-sinaptik ve son-sinaptik ateşleme zamanı arasındaki fark gecikmeyi değiştirme işlemidir [19,20].

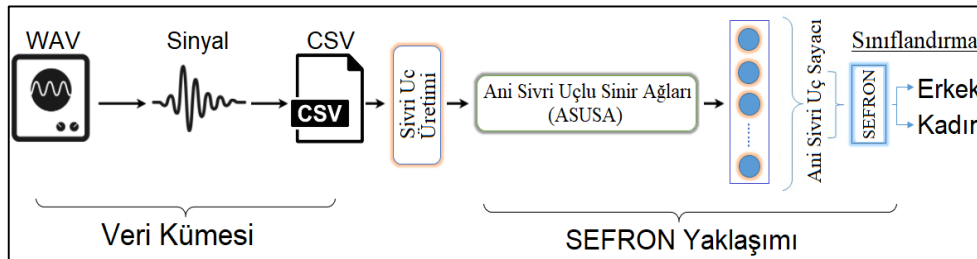
$$f_x(x') = R_{max} e^{-\frac{\cos(\frac{2\pi}{N}(XX))}{\delta^2}} \quad (2)$$

ASUSA modeli Jeyasothy ve ark.'nın SEFRON adı verdikleri açık kaynak kodları kullanılarak tasarlanan bir yaklaşımdır. Bu çalışmada tüm analizler SEFRON modelinin açık kaynak kodları güncellenerek gerçekleştirilmiştir [19].



Şekil 1. ASUSA modelinin genel adımları; giriş katmanı, gizli katmanlar ve çıktı katmanı

Şekil 1 incelendiğinde sayısal değerler içeren sinyal verileri hazırlanan modelin giriş katmanına verilir ve buradaki nöronlar, ağırlık parametreleri içeren gizli katmanlar ile eşleştirilerek hesaplanır ve aktivasyon fonksiyonlarından geçirilerek çıkış katmanına ulaştırılır. Çıkış katmanındaki özellikler SEFRON modeli ile taşıdığı değere göre sınıflandırılır. Önerilen yaklaşımın genel tasarımı Şekil 2'de verilmiştir.



Şekil 2. Önerilen yaklaşımın genel tasarımı

Şekil 2 incelendiğinde ses aygıtlarından alınan ses verileri csv formatına dönüştürülür. Bu verilerden daha sonra sivri uç üretimi yapılır ve ASUSA ağına verilir. Bu ağda elde edilen özellik

parametreleri SEFRON yaklaşımı kullanılarak sınıflandırılır. Tablo 2’de kullanılan SEFRON modelinin parametre değerleri görülmektedir.

**Tablo 2.** Veri kümesini oluşturan akustik özellikler

Özellik	Açıklama
Popülasyon kodlama şemasındaki alıcı alan nöron sayısı	18
Ön-sinaptik artış aralığı (ms)	3
Son-sinaptik artış aralığı (ms)	4
Son-sinaptik ateşleme süresi (ms)	2
Zaman adım hassasiyeti	0,01
Ağırlık güncellemesinin öğrenme oranı	0,56
Zamanla değişen ağırlıktaki çekirdeğin sigması (ms)	0,5
Ani tepki fonksiyonunun zaman sabiti (ms)	3
Maksimum devir (dönem) sayısı	100
STDP öğrenme penceresinin zaman sabiti Label	1,8

Tablo 2’ incelendiğinde SEFRON modelinin en önemli parametre değerlerinden olan alıcı nöron sayısı sınıf sayısının 2 olması ve veri sayısının da yüksek olmaması nedeniyle 18 olarak ayarlanmıştır. Farklı nöron sayıları ile denemeler yapılmış en uygun sayının 18 olduğu tespit edilmiştir. Kadın ve erkek seslerinin tespiti için önerilen bu modelde sinaptik artışı 2 ila 4 ms aralığında denemeler yapılarak seçilmiştir. Maksimum 100 dönem sayısı ile önerilen modelin tepkileri incelenmiştir.

### 3. Bulgular

Çalışmanın analiz süreci MATLAB 2020 yazılımı kullanılarak gerçekleştirildi. Donanımsal gereksinimler; işlemci Intel® Core i5 4200M 2.5 GHz, grafik kartı belleği 2 GB, geçici bellek 6 GB ve 64 bitlik Windows 10 işletim sistemi kullanılmıştır. Deneysel analizlerin sonuçlarını elde edebilmek için karışıklık matrisi kullanılmıştır ve bu matrisin metrik parametreleri şunlardır; duyarlılık, özgülük, hassasiyet, f-skor ve doğruluktur. Metrik değerlerinin hesaplanmasında Denklem 3 ile Denklem 7 arasındaki formüller kullanıldı. Bu denklemlerde kullanılan değişkenler; doğru pozitif (DP), doğru

negatif (DN), yanlış pozitif (YP), yanlış negatif (YN) anlamına gelmektedir [21–24].

$$\text{Duyarlılık (Duy.)} = \frac{DP}{DP + YN} \quad (3)$$

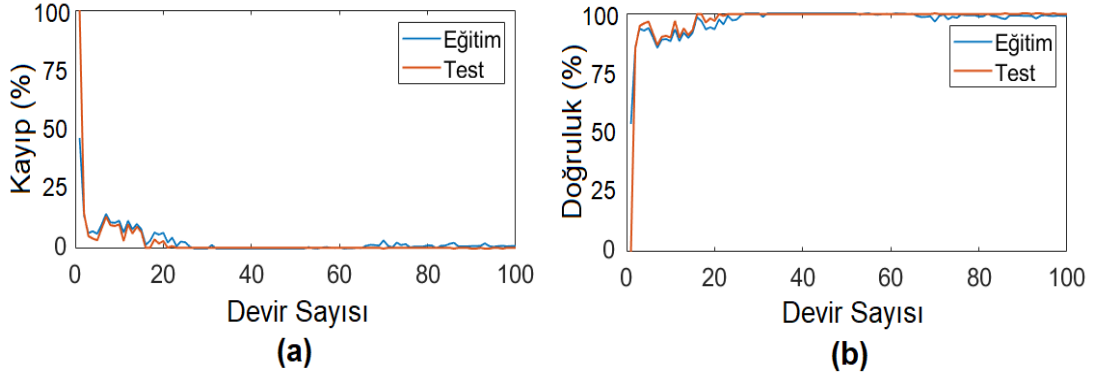
$$\text{Özgüllük (Özg.)} = \frac{DN}{DN + YP} \quad (4)$$

$$\text{Hassasiyet (Has.)} = \frac{DN + YP}{2xDP} \quad (5)$$

$$\text{F – skor (f – skr)} = \frac{2xDP + YP + YN}{DP + DN} \quad (6)$$

$$\text{Doğruluk (Dğr.)} = \frac{DP + YN + YP + DN}{DP + YN + YP + DN} \quad (7)$$

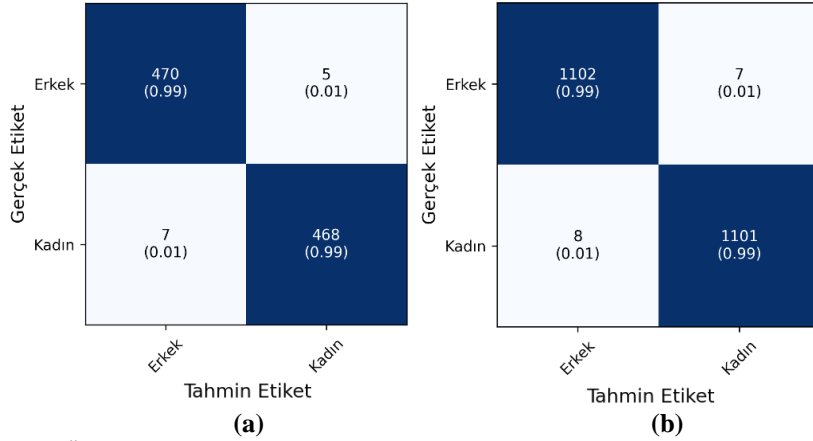
Çalışmanın deneysel analizinde ses sinyallerinden oluşan veri kümesi, “csv” uzantılı dosya formatına dönüştürüldü. Ardından veri kümesi %30’u test verisi olarak ve %70’i eğitim verisi olarak iki gruba ayrıldı. Eğitim ve test verileri önerilen yaklaşım modeli ile eğitildi. Eğitim için devir (dönem) sayısı 100 seçildi. Model eğitimi, her bir devir sayısı için yaklaşık 121 saniye sürdü ve toplam da önerilen modelin eğitimi 12100 saniyede tamamlandı. Eğitim ve test verilerinin eğitim grafikleri Şekil 3’te gösterildi. Önerilen modelin kayıp grafikleri Şekil 3(a)’da ve genel doğruluk grafikleri ise Şekil 3(b)’de gösterildi.



Şekil 3. Önerilen yaklaşımın eğitim ve test grafikleri; a) kayıp grafikleri, b) genel doğruluk grafikleri

Deneysel analizde eğitim ve test verilerinden elde edilen karışıklık matrisler Şekil 4'te gösterilmiştir. Test verilerinin karışıklık matris

değerleri Şekil 4(a)'da ve eğitim verilerinin karışıklık matris değerleri Şekil 4(b)'de verilmiştir.



Şekil 4. Önerilen yaklaşımın karışıklık matrisleri; a) test verileri, b) eğitim verileri

Test verilerinin karışıklık matris değerleri kullanılarak elde edilen metrik tabanlı analiz sonuçları Tablo 3'te verilmiştir. Eğitim verilerinin analiz sonucunda elde edilen genel doğruluk başarısı

%98,92 ve test verilerinin analiz sonucunda elde edilen genel doğruluk başarısı %98,84 olarak tespit edilmiştir.

Tablo 3. Önerilen yaklaşımda kullanılan SEFRON modelinin parametreleri ve değerleri

Parametre	Sınıf	Duy.	Özg.	Has.	f-skr	Genel Dğr.
Test	Erkek	98,73	98,74	98,84	98,84	98,84
	Kadın	98,73	98,94	98,95	98,84	
Eğitim	Erkek	99,01	98,83	98,83	98,92	98,92
	Kadın	98,83	99,01	99,01	98,92	

Tablo 3 incelendiğinde önerilen modelin ses tanımada yüksek başarı elde edildiği görülmektedir. Cinsiyete göre yaklaşık hemen hemen aynı oranlarda performans gösterdiği görülmektedir. Eğitim aşamasının performans ortalaması %98,92 olurken test aşaması %98,84 olarak tespit edilmiştir.

#### 4. Tartışma

Çeşitli biyometrik verilerden çıkartılan özellikleri kullanarak bireylerin cinsiyet tespiti, teknolojik gelişmelerin paralelinde mümkün hale gelmiştir. Biz bu çalışmada, İngilizce dili konuşanlar arasından elde edilen ses sinyallerini işleyerek ani sivri uçlu ağlar ile eğitimini gerçekleştirilmiştir. ASUSA modeli derin öğrenme modellerine göre daha zahmetli bir mimari

yapısına sahiptir ve çeşitli platformlarda yazılımcılar tarafından geliştirilmektedir. ASUSA modeli ile standart veri kümeleri (MNİST, NİST, CIFAR, vb.) kullanılarak analizler gerçekleştirilmişlerdir. Fakat diğer veri kümeleri (imgeler, sinyaller vb.) üzerinde literatürde çalışmalar sınırlıdır. ASUSA modeli ile

cinsiyet tanıma sürecinde veri kümelerinden elde edilen akustik özellikler kullanarak başarılı bir sonuç elde edilmiştir. Aynı veri kümesini kullanarak analizler gerçekleştiren diğer çalışmalar hakkında bilgiler Tablo 4'te verilmiştir.

**Tablo 4.** Önerilen yaklaşımda kullanılan SEFRON modelinin parametreleri ve değerleri

Çalışma	Yıl	Model/Yöntem	Genel Dğr. (%)
Buyukyılmaz ve Cibikdiken [6]	2016	ESA / Çok Katmanlı Algılayıcı (ÇKA)	96.74
Livieris, Pintelas ve Pintelas [25]	2019	Yarı denetimli öğrenme	98.42
Kacamarga ve ark. [26]	2019	ÇKA / Sigmoid	96.80
Önerilen bu çalışma	2022	SEFRON	98.84

Büyükılmaz ve Cibikdiken [6], ESA modeli ile ÇKA sınıflandırıcısını birlikte kullandıkları melez bir yaklaşım kullanarak %96.74 oranında başarı elde etmişlerdir. Kacamarga ve ark. [26], ÇKA modelini tam bağlantılı katmanlar ile birlikte kullanarak %96.80 oranında başarı elde etmişlerdir. İki çalışmada benzeri bir model önerdikleri için elde ettikleri başarılar yaklaşık çıkmıştır. Livieris, Pintelas ve Pintelas [25], yarı denetimli kendinden etiketli algoritma yöntemini kullanarak cinsiyet tespit başarısını %98.42'ye yükseltmiştir. Önerdiğimiz model ile ESA modellerinden farklı bir yaklaşım kullanarak cinsiyet tespit başarısı %98,84'e çıkartılmıştır. Aynı veri kümesini kullanan çalışmalarda önerilen yaklaşımlara kıyasla, ASUSA modelinin başarısı daha iyi performans sergilediği görülmüştür.

## 5. Sonuç ve Öneri

Ses sinyallerini kullanarak insanların özelliklerini ayırt etmek mümkündür. Bu özelliklerden biri de cinsiyet ayrımının ses sinyalleri üzerinde gerçekleştirilmesidir. Günümüzde, derin öğrenme modelleri ile insan sesinden elde edilen akustik özellikleri kullanarak istenilen bilgiyi elde etmek mümkündür. Akustik özellikler ile bireylerin cinsiyet

türünün doğru bir şekilde tespit edilebilmesi ve %90 üzerinde başarılı sonuçlar alınabilmektedir. Bu çalışmada akustik özellikler kullanılarak, sivri uçlu model ile cinsiyet tespitini %98,84 başarı ile ayırt edilmesi gerçekleştirildi. Önerilen yaklaşım, performans anlamında başarılı sonuçlar üretilmiştir. Önerilen yaklaşımın dezavantajlı yönü ise, diğer ESA modellerine göre daha fazla zaman kaybı olmasıdır. Bu durum, SEFRON modelinin darbelerin (sivri uçlu) oluşmasına kadar beklenen sürelerden kaynaklanmaktadır. Ayrıca SEFRON modelinin bir diğer dezavantajlı yönü olarak Tablo 2'de belirtilen parametre değerleri farklı veri kümelerinde değişkenlik göstermesidir. Bu durum, modelin başarılı sonuç alabilmesini geciktirebilmektedir. Neticede, ASUSA modeli gelişime açık bir modeldir ve bu problemlerin minimize edilmesi yakın zamanda mümkündür.

Gelecek çalışmalarda, farklı veri kümeleri üzerinde derin öğrenme modelleri ile sivri uçlu yaklaşımlar birlikte tasarlanarak analizlerin gerçekleştirilmesi planlanmaktadır.

## Araştırma ve Yayın Etiği Beyanı

Yapılan çalışmada araştırma ve yayın etiğine uyulmuştur.

## Kaynaklar

- [1] C. Pernet and P. Belin, "The Role of Pitch and Timbre in Voice Gender Categorization", *Frontiers in Psychology*, vol. 3, 2012. Doi: 10.3389/fpsyg.2012.00023.
- [2] V. Cartei and D. Reby, "Effect of Formant Frequency Spacing on Perceived Gender in Pre-Pubertal Children's Voices", *PLoS ONE*, vol. 8, no. 12, p. e81022, 2013. Doi: 10.1371/journal.pone.0081022.
- [3] M. Poon and M. Ng, "The role of fundamental frequency and formants in voice gender identification", *Speech, Language and Hearing*, vol. 18, no. 3, pp. 161-165, 2014. Doi: 10.1179/2050572814y.0000000058.



- [4] V. Skuk, L. Dammann and S. Schweinberger, "Role of timbre and fundamental frequency in voice gender adaptation", *The Journal of the Acoustical Society of America*, vol. 138, no. 2, pp. 1180-1193, 2015. Doi: 10.1121/1.4927696.
- [5] R. Alkhaldeh, "DGR: Gender Recognition of Human Speech Using One-Dimensional Conventional Neural Network", *Scientific Programming*, vol. 2019, pp. 1-12, 2019. Doi: 10.1155/2019/7213717.
- [6] M. Buyukyilmaz and A. Cibikdiken, "Voice Gender Recognition Using Deep Learning", *Proceedings of 2016 International Conference on Modeling, Simulation and Optimization Technologies and Applications (MSOTA2016)*, 2016. Doi: 10.2991/msota-16.2016.90.
- [7] E. Salomons and P. Havinga, "A Survey on the Feasibility of Sound Classification on Wireless Sensor Nodes", *Sensors*, vol. 15, no. 4, pp. 7462-7498, 2015. Doi: 10.3390/s150407462.
- [8] G. Archana and M. Malleswari, "Gender identification and performance analysis of speech signals", *2015 Global Conference on Communication Technologies (GCCT)*, 2015. Doi: 10.1109/gcct.2015.7342709.
- [9] S. İleri, a. Karabina and E. KILIÇ, "Konuşmacı Cinsiyetinin Tespitinde Değişik Normalizasyon Tekniklerinin Kıyaslanması", *Mehmet Akif Ersoy Üniversitesi Uygulamalı Bilimler Dergisi*, vol. 2, no. 2, pp. 1-12, 2018. Doi: 10.31200/makuubd.410625.
- [10] D. Doukhan, J. Carrive, F. Vallet, A. Larcher and S. Meignier, "An Open-Source Speaker Gender Detection Framework for Monitoring Gender Equality", *2018 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 2018. Doi: 10.1109/icassp.2018.8461471.
- [11] P. Wang and J. Hu, "A hybrid model for EEG-based gender recognition", *Cognitive Neurodynamics*, vol. 13, no. 6, pp. 541-554, 2019. Doi: 10.1007/s11571-019-09543-y.
- [12] M. Pfeiffer and T. Pfeil, "Deep Learning With Spiking Neurons: Opportunities and Challenges", *Frontiers in Neuroscience*, vol. 12, 2018. Doi: 10.3389/fnins.2018.00774.
- [13] J. Lobo, J. Del Ser, A. Bifet and N. Kasabov, "Spiking Neural Networks and online learning: An overview and perspectives", *Neural Networks*, vol. 121, pp. 88-100, 2020. Doi: 10.1016/j.neunet.2019.09.004.
- [14] K. Becker, "Gender Recognition by Voice", *Kaggle.com*, 2022. [Online]. Doi: <https://www.kaggle.com/primaryobjects/voicegender>. [Accessed: 01- Apr- 2022].
- [15] M. Toğaçar, B. Ergen and Z. Cömert, "Detection of weather images by using spiking neural networks of deep learning models", *Neural Computing and Applications*, vol. 33, no. 11, pp. 6147-6159, 2020. Doi: 10.1007/s00521-020-05388-3.
- [16] A. Sboev, A. Serenko, R. Rybka and D. Vlasov, "Solving a classification task by spiking neural network with STDP based on rate and temporal input encoding", *Mathematical Methods in the Applied Sciences*, vol. 43, no. 13, pp. 7802-7814, 2020. Doi: 10.1002/mma.6241.
- [17] A. Tavanaei, M. Ghodrati, S. Kheradpisheh, T. Masquelier and A. Maida, "Deep learning in spiking neural networks", *Neural Networks*, vol. 111, pp. 47-63, 2019. Doi: 10.1016/j.neunet.2018.12.002.
- [18] M. Stimberg, R. Brette and D. Goodman, "Brian 2, an intuitive and efficient neural simulator", *eLife*, vol. 8, 2019. Doi: 10.7554/elife.47314.
- [19] A. Jeyasothy, S. Sundaram and N. Sundararajan, "SEFRON: A New Spiking Neuron Model With Time-Varying Synaptic Efficacy Function for Pattern Classification", *IEEE Transactions on Neural Networks and Learning Systems*, vol. 30, no. 4, pp. 1231-1240, 2019. Doi: 10.1109/tnnls.2018.2868874.
- [20] X. Wang, X. Lin and X. Dang, "A Delay Learning Algorithm Based on Spike Train Kernels for Spiking Neurons", *Frontiers in Neuroscience*, vol. 13, 2019. Doi: 10.3389/fnins.2019.00252.
- [21] S. Sayyad, M. Shaikh, A. Pandit, D. Sonawane and S. Anpat, "Confusion Matrix-Based Supervised Classification Using Microwave SIR-C SAR Satellite Dataset", *Communications in Computer and Information Science*, pp. 176-187, 2021. Doi: 10.1007/978-981-16-0493-5\_16.
- [22] M. Toğaçar, Ergen B. Biyomedikal Görüntülerde Derin Öğrenme ile Mevcut Yöntemlerin Kıyaslanması. *Fırat Üniversitesi Mühendislik Bilim Derg* 2019;31:109–21.
- [23] E. Basaran, Z. Comert, A. Sengur, U. Budak, Y. Celik and M. Togacar, "Chronic Tympanic Membrane Diagnosis based on Deep Convolutional Neural Network", *2019 4th International Conference on Computer Science and Engineering (UBMK)*, 2019. Doi: 10.1109/ubmk.2019.8907070.



- [24] W. Alsaggaf, Z. Cömert, M. Nour, K. Polat, H. Brdese and M. Toğaçar, "Predicting fetal hypoxia using common spatial pattern and machine learning from cardiotocography signals", *Applied Acoustics*, vol. 167, p. 107429, 2020. Doi: 10.1016/j.apacoust.2020.107429.
- [25] I. Livieris, E. Pintelas and P. Pintelas, "Gender Recognition by Voice using an Improved Self-Labeled Algorithm", *Machine Learning and Knowledge Extraction*, vol. 1, no. 1, pp. 492-503, 2019. Doi: 10.3390/make1010030.
- [26] M. Kacamarga, T. Cenggoro, A. Budiarto, R. Rahutomo and B. Pardamean, "Analysis of Acoustic Features in Gender Identification Model for English and Bahasa Indonesia Telephone Speeches", *Procedia Computer Science*, vol. 157, pp. 199-204, 2019. Doi: 10.1016/j.procs.2019.08.158.