T.C.
DİCLE ÜNİVERSİTESİ
FEN BİLİMLERİ ENSTİTÜSÜ

BOYUT ARTTIRMA YÖNTEMLERİ KULLANILARAK EEG SİNYALLERİNDEN DERİN ÖĞRENME TABANLI ŞİZOFREN DURUM TESPİTİ

Zülfikar ASLAN

DOKTORA TEZİ

ELEKTRİK-ELEKTRONİK MÜHENDİSLİĞİ ANABİLİM DALI

DİYARBAKIR
ŞUBAT - 2021

T.C.
DİCLE ÜNİVERSİTESİ
FEN BİLİMLERİ ENSTİTÜSÜ

BOYUT ARTTIRMA YÖNTEMLERİ KULLANILARAK EEG SİNYALLERİNDEN DERİN ÖĞRENME TABANLI ŞİZOFREN DURUM TESPİTİ

Zülfikar ASLAN

DOKTORA TEZİ

ELEKTRİK-ELEKTRONİK MÜHENDİSLİĞİ ANABİLİM DALI

DİYARBAKIR
ŞUBAT- 2021

TEŞEKKÜR

Doktora programım süresince paha biçilmez rehberliğini ve desteğini esirgemeyen, değerli vaktini bana harcayıp bilgi ve tecrübesi ile her türlü imkânı sunan, beraber çalışmaktan mutluluk duyduğum, değerli danışmanım sayın Prof. Dr. Mehmet AKIN 'a çok teşekkür ederim.

Ayrıca doktora eğitimim boyunca raporlarımı dikkatle değerlendirerek, tez çalışmamda önerileri ve yol göstericiliğiyle desteğini her zaman yanımda hissettiğim tez izleme jüri üyelerim sayın Prof. Dr. İbrahim TÜRKOĞLU ve Prof. Dr. Mehmet Siraç ÖZERDEM 'e teşekkürlerimi sunuyorum.

Sürekli olarak yanımda olan ve fikirleri ile bana destek olan değerli arkadaşlarıma ve sevgili eşime teşekkür etmek isterim.

İÇİNDEKİLER

TEŞEKKÜR 1
İÇİNDEKİLER II
ÖZET IV
ABSTRACT VI
ÇİZELGE LİSTESİ VIII
ŞEKİL LİSTESİ IX
KISALTMA VE SİMGELER XI
1. GİRİŞ 1
1.1. Tezin Amacı 2
1.2. Tezin Önemi 3
1.3. Tezin İçeriği 3
2. KAYNAK ÖZETLERİ 5
3. MATERYAL VE METOT 11
3.1. Veri Setleri 11
3.1.1. Veri Seti A 11
3.1.2. Veri Seti B 12
3.2. METOT 13
3.2.1. Elektroensefalogram (EEG) 13
3.2.2. Geleneksel Makine Öğrenmesi Algoritmaları 14
3.2.2.1. K-En Yakın Komşu Sınıflandırıcı 14
3.2.2.2. Naive Bayes 15
3.2.2.3. Çok Katmanlı Algılayıcı (ÇKA) 16
3.2.2.4. Destek Vektör Makinaları (DVM) 19
3.2.2.5. Rastgele Orman 20
3.2.3. Değerlendirme Metrikleri 22
3.2.4. Özellik Seçim Algoritmaları 24
3.2.4.1. Genetik Arama Algoritması 25
3.2.4.2. Parçacık Sürüsü Optimizasyon Arama Algoritması 25
3.2.4.3. Evrimsel Arama Algoritması 26
3.2.4.4. En İyi İlk Algoritması 26
3.2.4.5. Karınca Kolonosi Optimizasyonu Arama Algoritması 27
3.2.5. Kısa Zamanlı Fourier Dönüşümü 27

3.2.6. Dalgacık Dönüşümü 28
3.2.7. Göreceli Dalgacık Enerjisi (GDE) 30
3.2.8. Derin Öğrenme 32
3.2.9. Keras Derin Öğrenme Çatısı 32
3.2.10. Üretken ve Ayrımcı Modeller 33
3.2.11. Kısıtlanmış Boltzmann Makinesi 34
3.2.12. Derin İnanç Ağları 35
3.2.13. Otomatik Kodlayıcı 36
3.2.14. Evrişimli Sinir Ağları (ESA) 37
3.2.15. Vgg-16 ESA Mimarisi 39
3.2.16. Spektrogram Görüntülerinin Oluşturulması 39
3.2.17. Skalogram Görüntülerinin Oluşturulması 40
3.2.18. Tez Çalışmasında Geleneksel Makine Öğrenme Metotları Kullanılarak Önerilen BDT Yönteminin Genel Mimarisi 41
3.2.19. Tez Çalışmasında Derin Öğrenme Kullanılarak Önerilen BDT Yönteminin Genel Mimarisi 43
4. BULGULAR VE TARTIŞMA 45
5. SONUÇ VE ÖNERİLER 71
5.1. Tez Kapsamında Yapılan Yayınlar 71
-Science Citation Index (SCI) veya Science Citation Index Expanded (SCI-E) Yayınlar . 71
-Konferans / Sempozyum Bildirileri 72
6. KAYNAKLAR 73
ÖZGEÇMİŞ 82

ÖZET

BOYUT ARTTIRMA YÖNTEMLERİ KULLANILARAK EEG SİNYALLERİNDEN DERİN ÖĞRENME TABANLI ŞİZOFREN DURUM TESPİTİ

DOKTORA TEZİ
Zülfikar ASLAN

DİCLE ÜNİVERSİTESİ
FEN BİLİMLERİ ENSTİTÜSÜ

ELEKTRİK – ELEKTRONİK MÜHENDİSLİĞİ ANABİLİM DALI

2021

Günümüzde pek çok çalışma, elektroensefalogram (EEG) verisi inceleyerek beyin aktivitelerini tespit etme, anlama veya yorumlama amacıyla yapılmış ve çeşitli sonuçlar ortaya konulmuştur. Beyin aktivitelerinin anlaşılması, şizofreni ve epilepsi gibi pek çok hastalığın teşhisi, uyku evreleri gibi farklı durumların veya hastalıkların çözümlenmesi için önemli bir adımdır. Dolayısıyla bu yöndeki çalışmaların sonuçlarının iyileştirilmesi veya bir adım öteye götürülmesi için yeni modeller ortaya konulması oldukça kritik bir öneme sahiptir. Bu tez çalışmasında, ilk olarak literatürde önceden yapılmış, klasik makine öğrenme algoritmalarıyla sınıflandırılmış EEG verilerinden elde edilen analizlere göre başarımı daha yüksek metotlar ortaya konulması hedeflenmektedir. İkinci aşamada ise EEG verileri derin öğrenme metotlarıyla incelenerek özellik çıkarımı yapılmadan daha yüksek başarım ile sınıflandırma yapılması ve şizofreni (SZ) hastaları ile sağlıklı bireyler arasındaki farklılığın yorumlanabilir görüntüler olarak tasvir edilmesi amaçlanmıştır.

Bu tez çalışmasında, EEG sinyalleri kullanılaılarak veri analiz yöntemleriyle beyin aktivitelerinin incelenmesi amaçlanmıştır. Bu amaç doğrultusunda özellikle SZ hastalarına ait EEG kayıtları hem klasik makine öğrenmesi algoritmalarıyla hemde derin öğreme algoritması kullanılarak sınıflandırılmıştır. Bu çalışmada SZ hastalarına ait farklı iki veri seti kullanılarak uygulanan metotların doğruluğu teyit edilmiştir. İlk veri kümesinde(Veri seti A) 39 sağlıklı çocuk ve 45 SZ hastalığı olan 16 kanallı EEG kaydı bulunmaktadır. İkinci veri seti (Veri seti B) ise 28 yetişkinin (14 kontrol ve 14 SZ hastası) 19 kanallı EEG kayıtları içermektedir. Literatürde özellikle klasik makine öğrenme algoritmaları SZ hastalarının sınıflandırılmasında yaygın olarak kullanılmıştır. Ancak bu tez çalışmasının ilk aşamasında farklı özellik çıkarım ve özellik seçim metotları kullanılarak daha iyi sınıflandırma başarım sonuçları elde edilmiştir. Bunun yanısıra

farklı sınıflandırma algoritmalarıyla elde edilen başarım sonuçları karşılaştırılmıştır.

Tez çalışmasının ikinci aşamasında derin öğrenme algoritmalarıyla, SZ hastalarını EEG kayıtlarını kullanarak otomatik olarak teşhis etmeyi amaçlayan bir yöntem sunulan bir çalışma yapılmıştır. Birçok literatür çalışmasından farklı olarak, önerilen yöntem EEG kayıtlarındaki özellikleri manuel olarak çıkarmak yerine frekans-zaman özelliklerinin yararlı bir temsiline sahip olmak için Kısa Zamanlı Dalgacık Dönüşümü (KZDD) ve Sürekli Dalgacık Dönüşümü(SDD) kullanarak ham EEG'yi iki boyuta dönüştürür. Bu çalışma, ilgili literatürde SZ hastalarının otomatik teşhisi için 2D zaman frekansı özelliklerinin kullanılmasında ilktir. 2D alanda mevcut olan tüm özelliklerden en kullanışlı özellikleri çıkarmak ve örnekleri yüksek doğrulukla sınıflandırmak için, son teknoloji bir Evrişimsel Sinir Ağı (ESA) mimarisini kullanan VGG-16 algoritması ile eğitilmiştir. Farklı ESA algoritmaları (CIFAR, VGG-19, ResNet, DenseNet, MobileNet, Xception … gibi) ile de sonuçlar denenmiş ancak en az iterasyon ve sürede, en iyi başarım değeri VGG-16 ile elde edilmiştir. Deneysel sonuçlar, sunulan yöntemin farklı yaş gruplarının iki veri kümesinde SZ hastalarını ve sağlıklı kontrolleri, spektrogram görüntülerini % 95 ve % 97, skalogram görüntülerini % 98 ve % 99.5 sınıflandırma doğruluğu ile sınıflandırma görevinde başarılı olduğunu göstermektedir. Bu performansla, önerilen yöntem literatür yöntemlerinin çoğundan daha iyi performans gösterir. Çalışmanın deneyleri ayrıca bir EEG kaydının frekans bileşenleri ile SZ hastalığı arasında bir ilişki olduğunu ortaya koymaktadır. Ayrıca, tez çalışmamızda sunulan Aktivasyon Maksimizasyonu, Grad-CAM ve Belirginlik Haritası yöntemleriyle bir SZ hastasını sağlıklı bir kontrolden ayırırken yorumlanabilir görüntüler elde edilmesi sağlanmıştır. Bu görüntüler ile uzman görüşüne destek niteliğinde SZ hastası ve sağlıklı kontroller arasındaki fark kolayca anlaşılabilir.

Anahtar Kelimeler: Şizofreni, EEG, Sınıflandırma, Spektogram, Derin Öğrenme, Skalogram

ABSTRACT

DEEP LEARNING BASED SCHIZOPHRENIA STATUS DETERMINATION FROM EEG SIGNALS USING DIMENSION AUGMENTATİON METHODS

PhD THESIS

Zülfikar ASLAN

DEPARTMENT OF ELECTRICAL ve ELECTRONICS ENGINEERING
INSTITUTE OF NATURAL ve APPLIED SCIENCES
UNIVERSITY OF DICLE

2021

Today, many studies have been conducted to detect, understand or interpret brain activities by examining electroencephalogram (EEG) data, ve various results have been presented. Understanding brain activity is an important step in diagnosing many diseases such as schizophrenia ve epilepsy, ve resolving different conditions or diseases such as sleep stages. Therefore, it is of critical importance to develop new models in order to improve the results of such studies or to take them one step further. In this thesis, firstly, it is aimed to reveal methods with higher performance compared to the analyzes obtained from EEG data classified with classical machine learning algorithms, which were previously made in the literature. In the second stage, it was aimed to make classification with higher performance without extracting features by examining the EEG data with deep learning methods ve to depict the difference between schizophrenia (SZ) patients ve healthy individuals visually.

In this thesis, it was aimed to examine brain activities by using data analysis methods using EEG signals. For this purpose, EEG records of especially SZ patients were classified using both classical machine learning algorithms ve deep learning algorithms. In this study, the accuracy of the methods applied using two different datasets of SZ patients was confirmed. In the first dataset (Data set A), there are 39 healthy children ve a 16-channel EEG record with 45 SZ diseases. The second dataset (Data set B) contains 19-channel EEG recordings of 28 adults (14 control ve 14 SZ patients). In the literature, especially classical machine learning algorithms have been widely used in the classification of SZ patients. However, at the first stage of this thesis, better classification performance results were obtained by using different feature extraction ve feature selection methods. In addition, the performance results obtained with different classification algorithms were compared.

In the second stage of the thesis, a study was conducted with deep learning algorithms, in which a method aiming to automatically diagnose SZ patients using EEG records was conducted. Unlike many literature studies, the proposed method converts raw EEG to 2D using Short Time Wavelet Transform (STFT) ve Continuous Wavelet Transform (CWT) to have a useful representation of frequency-time properties rather than manually extracting features in EEG recordings. This study is the first to use 2D time frequency properties for automatic diagnosis of SZ patients in the relevant literature. It has been trained with the VGG-16 algorithm, which uses a state-of-the-art Convolution Neural Network (CNN) architecture to extract the most useful features from all the features available in the 2D space ve to classify samples with high accuracy. The results were also tested with different CNN algorithms (such as CIFAR, VGG-19, ResNet, DenseNet, MobileNet, Xception…), but the best performance value was obtained with VGG-16 with the least iteration ve time. The experimental results show that the presented method is successful in classifying SZ patients ve healthy controls in two datasets of different age groups, spectrogram images with 95% ve 97%, ve scalogram images with 98% ve 99.5% classification accuracy. With this performance, the proposed method outperforms most literature methods. The experiments of the study also reveal a relationship between the frequency components of an EEG recording ve SZ disease. In addition, with the AktivizationMaximization, Grad-CAM ve Saliency Map methods presented in our thesis, interpretable images were obtained while separating an SZ patient from a healthy control. With these images, the difference between SZ patients and healthy controls, which support the expert opinion, can be easily understood.

Keywords: Schizophrenia, EEG, Clasification, Spectrogram, Deep Learning, Scalogram

ÇİZELGE LİSTESİ

Çizelge No		Sayfa
Çizelge 2.1.	Makine öğrenimini kullanarak şizofreniyi tespit eden ilgili literatür çalışmaları	7
Çizelge 2.2.	Derin öğrenme kullanarak şizofreniyi tespit eden ilgili literatür çalışmaları	10
Çizelge 2.3.	Şizofreni tespitinde derin öğrenme metodu ile segmantasyon yapan ve kullandığımız veri setlerini kulllanan çalışmalar	10
Çizelge 4.1.	Sınıflandırma işlemi için sezgisel arama algoritmalarının kullanılarak gerçekleştirilen özellik seçimiyle elde edilen özellik skoru ve doğruluk performans değerleri.	45
Çizelge 4.2.	Her bir elektrottan ve bütün elektrotlardan alınan veriler için farklı sınıflandırma yöntemleriyle elde edilen doğruluk değerleri.	47
Çizelge 4.3	Yeniden örnekleme işlemi sonucu elde edilen başarım sonuçları.	48
Çizelge 4.4.	Spektrogram görüntüleri ile eğitilmiş ağın, Veri Kümesi A'ya göre performans sonuçları.	53
Çizelge 4.5.	Spektrogram görüntüleri ile eğitilmiş ağın, Veri Kümesi B'ye göre performans sonuçları.	53
Çizelge 4.6	Skalogram görüntüleri kullanılarak, Veri Kümesi A ile elde edilen doğruluk değerlendirme ölçütleri.	63
Çizelge 4.7.	Skalogram görüntüleri kullanılarak, Veri Kümesi B ile elde edilen doğruluk değerlendirme ölçütleri.	63

ŞEKİL LİSTESİ

Şekil No		Sayfa
Şekil 1.1.	Tez çalışmasının aşamalarının gösterildiği akış şeması	4
Şekil 3.1.	Veri Kümesi A için 16 kanallı elektrot yerleşimi	12
Şekil 3.2.	Veri Kümesi B için 19 kanallı elektrot yerleşimi	12
Şekil 3.3.	EEG elektrotları ve sinyalinin bir örneği	14
Şekil 3.4.	(KNN) k-en yakın komşu sınıflandırma algoritması.	15
Şekil 3.5.	Naive Bayes sınıflandırma algoritması.	16
Şekil 3.6.	Tek gizli katmanlı ÇKA	17
Şekil 3.7.	Lineer olarak ayrılabilen DVM	20
Şekil 3.8.	Rastgele Orman sınıflandırma algoritması	22
Şekil 3.9.	Dalgacık dönüşümü kullanılarak yaklaşımlı ve detaylı alt kanallara ayrıştırma.	23
Şekil 3.10.	Karışıklık Matrisi	30
Şekil 3.11.	Üretken ve Ayrımcı model sınıflandırması	34
Şekil 3.12.	Üç görünür birim ve dört gizli birim ile sınırlı bir Boltzmann makinesinin şeması	35
Şekil 3.13.	Derin inanç ağına şematik genel bakış	36
Şekil 3.14.	Otomatik kodlayıcı sinir ağı	37
Şekil 3.15.	ESA Mimarisi blok yapısı	38
Şekil 3.16.	VGG-16 derin öğrenme ağ mimarisi	39
Şekil 3.17.	Matlab yazılımı tarafından oluşturulan örnek bir spektrogram	40
Şekil 3.18.	Matlab yazılımı yardımıyla Morlet dalgacığı kullanılarak oluşturulan örnek bir skalogram görüntüsü	41
Şekil 3.19.	Farklı sezgisel arama algoritmaları kullanılarak SZ tespiti	42
Şekil 3.20.	GDE ve istatistiksel özelliklerim farklı makine öğrenme algoritmaları ile performanlarının karşılaştırılması.	42
Şekil 3.21.	Spectrogram görüntülerinin derin öğrenmede kullanılarak SZ tespiti	43
Şekil 3.22.	Skalogram görüntülerinin derin öğrenmede kullanılarak SZ tespiti	44

Şekil 4.1	Sınıflandırma işlemi için sezgisel arama algoritmalarının kullanılarak gerçekleştirilen özellik seçimiyle elde edilen özellik skoru ve doğruluk performans değerleri.	46
Şekil 4.2.	Yeniden örnekleme gerçekleştirilmeden önce çeşitli sınıflandırma algoritmaları ile elde edilen doğruluk performası grafiği	48
Şekil 4.3	Yeniden örnekleme gerçekleştirildikten sonra çeşitli sınıflandırma algoritmaları ile elde edilen doğruluk performası grafiği.	49
Şekil 4.4	Veri Kümesi A'dan örnek spektrogram görüntüleri (a) sağlıklı kontrol ve (b) SZ hastası	50
Şekil 4.5	Veri Kümesi B'den örnek spektrogram görüntüleri (a) sağlıklı kontrol ve (b) SZ hastası	51
Şekil 4.6	(a) Veri Kümesi A ve (b) Veri Kümesi B için iterasyon ve eğitim süresine karşı doğruluk değerleri	52
Şekil 4.7.	(a) Veri Kümesi A ve (b) Veri Kümesi B için önerilen yöntemin karışıklık Matrisleri	54
Şekil 4.8.	(a) Veri Kümesi A ve (b) Veri Kümesi B için ROC eğrisi	56
Şekil 4.9.	(a) sağlıklı kontrol ve (b) SZ hastası için aktivasyon maksimizasyon görüntüleri	57
Şekil 4.10.	(a) sağlıklı kontroller ve (b) SZ hastalarından oluşan bir grup bireyden elde edilen Grad-CAM görüntüleri	58
Şekil 4.11.	Veri Kümesi A'dan farklı kişilere ait farklı segmentlerden alınan örnek skalogram görüntüleri	60
Şekil 4.12.	Veri seti B'dan farklı kişilere ait farklı segmentlerden alınan örnek skalogram görüntüleri	61
Şekil 4.13.	Ağın eğitimi esnasında, iterasyon değerlerine karşı doğruluk ve hata değerinin değişimi: (I) Veri Kümesi A için (II) Veri Kümesi B için	62
Şekil 4.14.	Karışıklık matrisleri: (I) Veri Kümesi A için (II) Veri Kümesi B için	64
Şekil 4.15.	ROC eğrileri: (I) Veri Kümesi A için (II) Veri Kümesi B için	66
Şekil 4.16.	ESA ağından elde edilen Sağlıklı kontrol (soldaki) ve SZ hastasına (sağdaki) ait aktivasyon maksimizasyon görüntüleri	68
Şekil 4.17.	Farklı bireylerden ve farklı segmentlerden alınan Grad-CAM görüntüleri : (I) sağlıklı kontrollerden alınan (II) SZ hastalarından alınan	68
Şekil 4.18.	Farklı bireylerden ve farklı segmentlerden alınan Belirginlik Haritası görüntüleri : (I) sağlıklı kontrollerden alınan (II) SZ hastalarından alınan	69

KISALTMA VE SİMGELER

EEG	: Elektroensefalogram
SZ	: Şizofreni
KZDD	: Kısa Zamanlı Dalgacık Dönüşümü
SDD	: Sürekli Dalgacık Dönüşümü
ESA	: Evrişimsel Sinir Ağı
BDT	: Bilgisayar Destekli Tanı
DL	: Derin Öğrenme
HFD	: Hızlı Fourier Dönüşümü
DD	: Dalgacık Dönüşümü
AMD	: Ampirik Mod Ayrıştırması
DVM	: Destek Vektör Makineleri
ÇKA	: Çok Katmanlı Algılayıcı
GDE	: Göreceli Dalgacık Enerjisi
KNN	: k- En Yakın Komşu
DDA	: Doğrusal Diskriminant Analizi
RO	: Rastgele Orman
TBSA	: Tam Bağlı Sinir Ağı
PSO	: Parçacık Sürüsü Optimizasyonu
KBM	: Kısıtlanmış Boltzmann Makinesi
OK	: Otomatik Kodlayıcı
DİA	: Derin İnanç Ağları
ADD	: Ayrık Dalgacık Dönüşümü
GA	: Genetik Algoritma
EA	: Evrimsel Algoritma
ONEIROS	: Açık Uçlu Nöro-Elektronik Akıllı Robot İşletim Sistemi
GPU	: Grafik İşlem Birimi
TPU	: Tensör İşleme Birimi
Grad-CAM	: Gradyan Ağırlıklı Sınıf Aktivasyon Haritalama
VGG	: Görsel Geometri Grubu
LSTM	: Uzun Kısa Süreli Bellek

N	: Negatif
P	: Pozitif
ACC	: Doğruluk
ERR	: Hata
FP	: Yanlış Pozitifler
TP	: Doğru Pozitifler
FN	: Yanlış Negatifler
TN	: Doğru Negatifler
FPR	: Yanlış Pozitif Oranı
TPR	: Doğru Pozitif Oranı
TNR	: Doğru Negatif Oranı
PRE	: Kesinlik
SEN	: Hassasiyet
REC	: Duyarlılık
SPC	: Özgünlük

1. GİRİŞ

Şizofreni (SZ), kronik ve şiddetli etkileri bulunan ve dünya genelinde yaklaşık 20 milyon insanı etkileyen ruhsal bozukluktur (Cooper 2018; WHO 2020). Doğru olmayan garip sabit inançlar (sanrılar) , var olmayan şeyleri duyma veya görme (halüsinasyonlar), düzensiz duygular, algılama veya konuşma gibi belirtilerle kendini gösteren, uzun süreli bir beyin bozukluğudur(Oh ve ark. 2019). Normal sağlıklı insanlara göre, SZ hastalarının ölümü önlenebilir fiziksel hastalıklar nedeniyle daha yüksektir (Laursen, Nordentoft, ve Mortensen 2014).

SZ hastalığının tespiti, genellikle uzman bir psikiyatrist tarafından hasta ile görüşmelerine ve hasta davranışlarının izlenmesine dayalı olarak gerçekleştirilir (Devia ve ark. 2019). SZ farklı psikolojik rahatsızlıkları ile karıştırılabilir semptomlar gösterdiği için, bazı durumlarda uzman kişinin hata yapması mümkün olduğundan tespit aşaması tam olarak güvenilir değildir. SZ genellikle uzman tarafından tespit edilmesine rağmen, son zamanlarda EEG ve nörogörüntüleme teknikleri (MRI, CT, FMRI, PET gibi) gibi araçlar ile otomatik olarak tespit etmek için kullanılmaya başlamıştır. Nörogörüntüleme tekniklerinin çok maliyetli olması ve yüksek hesaplama süresinden dolayı EEG kayıtlarına göre fazladan ek kayıt gerektirmesi açısından dezavantajlıdır (Acharya ve ark. 2012; Gudigar ve ark. 2019; Oh ve ark. 2019; Siuly ve ark. 2018; Talo ve ark. 2019). EEG daha az maliyetli ve daha pratik kullanıma sahip olduğu için SZ tespitinde EEG kullanımı daha yaygın olarak tercih edilmektedir. SZ tespitinde EEG verilerinin kullanımı ise daha az maliyetli ve kullanımı daha pratik bir yöntemdir.

EEG, beyindeki elektrik sinyallerinin aktivitesinin kaydedilmesine olanak sağlayan cihazlardır. EEG kayıtları, kafa derisine yerleştirilen elektrotlardan alınan sinyallere ait periyodik veya durağan olmayan zamana bağlı olarak değişen verileri içermektedir (Supriya ve ark. 2018; Yin ve ark. 2019). EEG verileri, beyin aktivitelerinin incelenmesine ve analiz edilmesine olanak sağlayan önemli bir materyal olarak kullanılmaktadır. SZ hastalığında erken tespit, beyindeki bozulmaların azalmasını yardımcı olabilir buna rağmen hastalığın tespiti bir uzmana göre bile zordur (Zhang 2019).

Bu sebeplerden dolayı, uzmana karar verme aşamasında yardımcı olması için bilgisayar destekli yazılımların gelişmesi önemli bir araştırma konusu olmuştur.

Literatürde, geleneksel makine öğrenme algoritmaları kullanılarak yapılmış birçok Bilgisayar Destekli Tanı (BDT) çalışma bulunmaktadır (Cao ve ark. 2018; Shim ve ark. 2016). Geleneksel makine öğrenme algoritmalarına göre, özellikle iki boyutlu verilerde oldukça tatmin edici sonuçlar veren Derin Öğrenmedeki (DL) son gelişmeler araştırmacıların bu alana yönelmesinde önemli bir neden olmuştur.

EEG verisine ait sinyaller, genellikle dijital hale getirilerek bilgisayar yazılımları ile analiz edilir. Böylece doğrudan herhangi bir uzman tarafından yorumlanabilmesi zor olan bu bilgi daha anlaşılır duruma getirilmiş olur. Bu amaçla, SZ gibi beyin hasarlarına sebep olan rahatsızlıkların veya herhangi bir hastalıkla ilgili anormalliği göstermediği durumlarda uzmanın değerlendirmesine yardımcı olması için geliştirilmiş BDT sistemleri bulunmaktadır. EEG verileri, SZ hastalığı gibi birçok hastalığı tespit etmek için geliştirilmiş BDT sistemine girdi verisi olarak kullanılmaktadır (Oh ve ark. 2019).

Bundan dolayı, ilgili literatürde SZ'yi tespit etmek için geliştirilmiş, EEG sinyallerini kullanan BDT sistemleri yaygın olarak kullanılmaktadır. Bu alanda yapılmış literatür çalışmalarının çoğunda, EEG verisinden çıkarılan özellikler ile geleneksel makine öğrenme algoritmalarının beslendiği çalışmalar gerçekleştirilerek hastalık tespit edilmeye çalışılmıştır. Hastalığın yüksek doğrulukla tespiti, sinyalden elde edilen özelliklere bağlıdır. Farklı özellik çıkarım yöntemleri kullanılarak, elde edilen özelliklerin farklı makine öğrenme algoritmalarıyla SZ hastalığının tespiti için sunulmuş birçok çalışma bulunmaktadır. Bunun yanında, derin öğrenmede sunulan özelik çıkarımında yüksek doğruluk elde edilen yeni ESA mimarilerinin gelişimi ile literatüde DL kullanımıda yaygınlaşmaktadır

Bu bölümde tezin amacı, hedefi ve kapsamı genel hatlarıyla anlatılmıştır.

1.1. Tezin Amacı

Bu tez çalışması, EEG kayıtlarından Şizofreni hastalığını otomatik olarak tespit etmeyi amaçlayan yüksek doğruluk performansına sahip bir yöntem sunmaktadır. Tez çalışması, iki aşmadan oluşmaktadır. Birinci aşamada, geleneksel makine öğrenme metotları kullanılarak SZ hastası ve sağlıklı bireylerin yüksek doğrulukla sınıflandırılması amaçlanmaktadır. İkinci aşamada ise, geleneksel makine öğrenme algoritmaları kullanan ilgili literatür çalışmalarının aksine, EEG kayıtları boyut arttırma yöntemleri kullanılarak elde edilen görüntüler üzerinden ağın eğitimi için gerekli olan

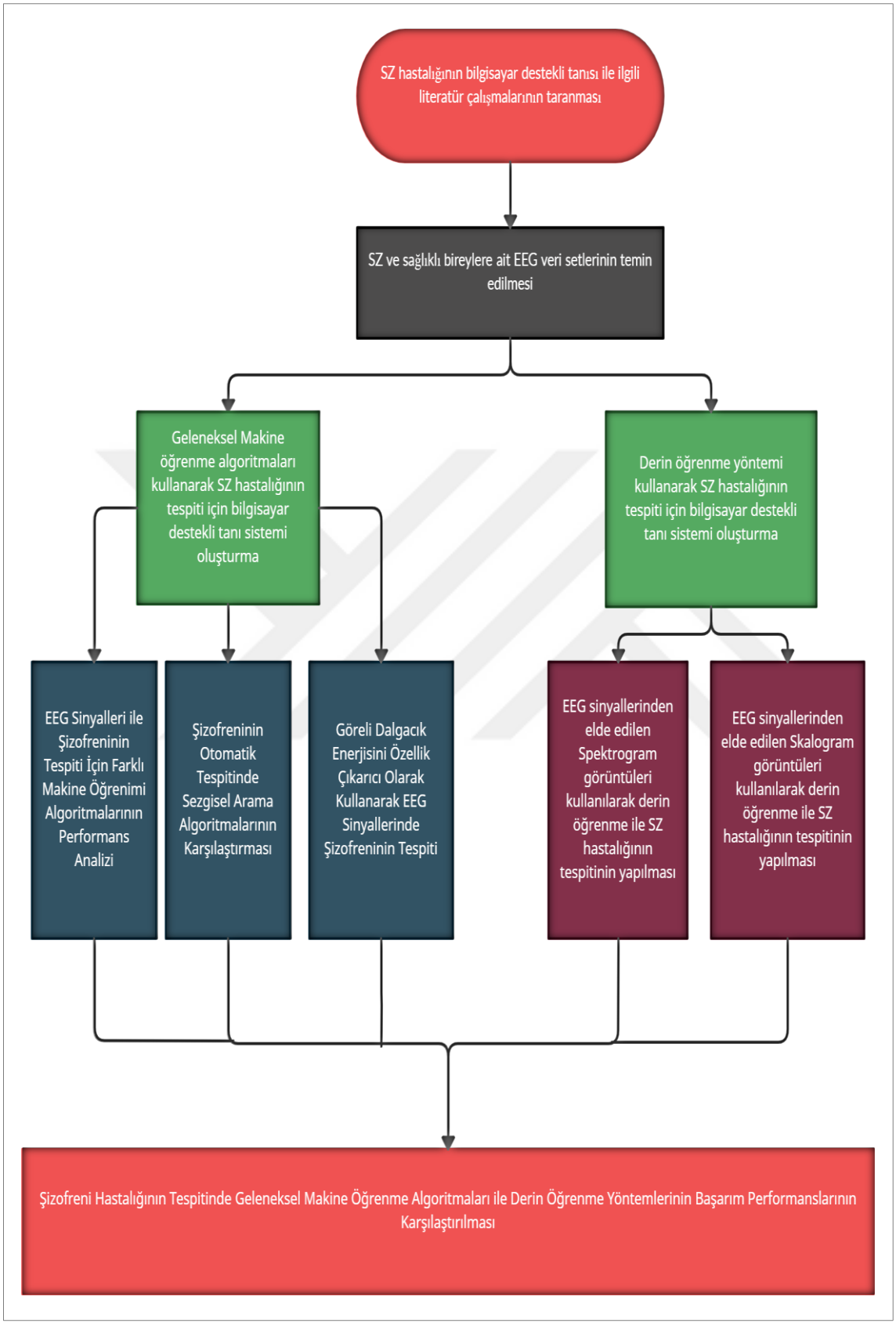
özelliklerin otomatik olarak çıkarıldığı DL kullanılarak sınıflandırma işlemi gerçekleştirilmiştir. Geleneksel makine öğrenme metotlarından farklı olarak DL ağının çıktısı farklı teknikler ile yorumlanabilir görüntüler elde edilmiştir. Dahası, bu yorumlanabilir çıktılar ile uzman görüşüne destek olarak SZ hastası ve sağlıklı bireyler arasındaki fark çok kolay bir şekilde ayırt edilebilmektedir.

1.2. Tezin Önemi

Bu tez çalışmasında, özellikle ham EEG verilerinden otomatik olarak özellik çıkarımı gerçekleştiren DL kullanılarak SZ hastası ve sağlıklı kontroller yüksek doğrulukla ayırt edilmiştir. Ham EEG sinyalleri üzerinde boyut arttırma işlemleri gerçekleştirilerek spektrogram ve skalogram görüntüleri elde edilmiştir. Farklı yaş gruplarından oluşan iki veri seti için spektrogram görüntüleriyle her veri seti için %95 ve %97 sınıflandırma başarımı elde edilirken, skalogram görüntüleriyle %98 ve %99.5 sınıflandırma başarımı elde edilmiştir. Bu yönüyle SZ tespitinde daha önceki literatür çalışmalarıyla karşılaştırıldığında oldukça tatmin edici bir sınıflandırma gerçekleştirmiştir. Ayrıca DL ağının görselleştirilerek SZ ve sağlıklı kontrolleri ayırt etmek için yorumlanabilir görüntüler elde edilmiştir. DL sonucu elde edilen yorumlanabilir görüntüler SZ tespiti ile yapılmış önceki literatür çalışmalarından farklı olarak ilk defa kullanılmış ve yorumlanabilir görüntüler arasında oldukça bariz farklar tespit edilmiştir.

1.3. Tezin İçeriği

Bu tez çalışmasının ikinci bölümünde, ilgili literatürde EEG sinyalleri kullanılarak yapılan çalışmalar değerlendirilmiştir. Literatürde yapılan çalışmalar; a) Geleneksel Makine Öğrenme Yöntemlerine Dayalı Çalışmalar, b) Derin Öğrenme Yöntemlerine Dayalı Çalışmalar olmak üzere iki ayrı sınıfta değerlendirilmiştir. Materyal ve metot kısmını oluşturan üçüncü bölümde, tezde kulanılan veri setleri, tez kapsamında kullanılan geleneksel makine öğrenme yöntemleri ve önerilen yaklaşım yer almaktadır. Dördüncü bölümde, çalışma sonucu elde edilen sınıflandırma performansları ve literatürdeki benzer çalışmalar ile karşılaştırılmaları değerlendirilmiştir. Tezin sonuçlar bölümü olan son bölümde ise, önerilen yöntemin katkıları ve yapılabilecek yeni çalışmaların detayları verilmiştir. Tez çalışmasının aşamaları Şekil 1.1'de gösterilmektedir.



Şekil 1.1. Tez çalışmasının aşamalarının gösterildiği akış şeması.

2. KAYNAK ÖZETLERİ

Bu tez çalışmasında, farklı yaş gruplarından SZ hastası ve sağlıklı kontrollerden oluşan iki farklı veri setinin hem geleneksel makine öğrenme hem de DL metoduyla sınıflandırma başarımları analiz edilmiştir. Bundan dolayı, EEG verileri kullanılarak yapılan ilgili literatür çalışmaları iki ayrı sınıfta değerlendirilmiştir. Bu çalışmalar; a) Geleneksel Öğrenme Yöntemlere Dayalı Çalışmalar, b) Derin Öğrenme Yöntemlerine Dayalı Çalışmalar.

a) Geleneksel Yöntemlere Dayalı Çalışmalar: EEG sinyallerinin geleneksel makine öğrenme algoritmaları ile sınıflandırılabilmesi için öncelikle ham EEG sinyali üzerinden önemli özelliklerin çıkarılması gerekmektedir. Bu özellik çıkarım işlemi için genellikle, sinyali zaman frekans domenine taşıyan Hızlı Fourier Dönüşümü (HFD), Kısa Zamanlı Fourier Dönüşümü (KZFD), Ampirik Mod Ayrıştırması, Dalgacık Dönüşümü (DD) gibi yöntemler kullanılmaktadır. Geleneksel makine öğrenme algoritmaları kullanarak SZ tespiti yapan ilgili literatür çalışmalarından aşağıda bahsedilmiştir.

(Kim ve ark. 2015), 21 elektrotlu bir EEG cihazından aldıkları kayıtlara ait 5 frekans bandını seçmiştir. Hızlı Fourier Dönüşümünü (HFD) seçilen bu 5 frekans bandına uygulayarak bu bantların spektral gücünü EEGLAB yazılımı ile hesaplamışlardır (Delorme ve Makeig 2004). Delta frekansını kullanarak SZ ve sağlıklı kontrolleri 62.2% doğruluk ile sınıflandırmışlardır.

Ayrıca, (Dvey-Aharon ve ark. 2015) EEG sinyallerini ön işleme tabi tutmuşlardır ve Stockwell dönüşümünü ile özellik çıkarımı yapmışlardır (Stockwell, Mansinha, ve Lowe 1996). TFFO (Zaman-Frekans dönüşümü ve ardından Özellik Optimizasyonu) yöntemini kullanarak 92% ile 93.9% arasında bir sınıflandırma başarımı elde etmişlerdir.

(Johannesen ve ark. 2016), SZ tespitinde, bellek çalışma performansını tahmin için elde edilen EEG kayıtlarını kullanarak özellikleri (Guyon ve Elisseeff 2003) çıkarmışlardır. Elde ettikleri özellikleri kullanarak Destek Vektör Makineleri (DVM) ile sınıflandırmış ve 87% doğruluk değeri elde etmişlerdir.

Bunun yanında, (Santos-Mayo, San-José-Revuelta, ve Arribas 2016), farklı

makine öğrenme tekniklerini (elektrot gruplama işlemi, özellik seçim algoritmaları ve filtreleme) çalışmışlardır. Elde ettikleri özellikleri DVM ve Çok Katmanlı Algılayıcı (MLP) algoritmalarıyla sırasıyla 92.23% ve 93.42% doğrulukla sınıflandırma işlemini gerçekleştirmişlerdir. Son olarak elde ettikleri özelliklere, J5 özellik seçme algoritması (Devijver ve Kittler 1982) uygulayarak sınıflandırmış ve daha yüksek başarım gösterdiğini gözlemlemişlerdir.

(Aslan ve Akın 2019b) çalışmasında, EEG sinyallerinden Göreceli dalgacık enerjisi (GDE) metodu uygulayarak özellik çıkarım işlemi gerçekleştirmişlerdir. Bu şilem sonucunda edinilen bu özellikleri KNN (K-En Yakın Komşular) algoritması kullanarak eğitmişler ve bu yöntemle, yaklaşık 90% doğruluk değeri elde ettiklerini bildirmişlerdir.

(Thilakvathi ve ark. 2017), DVM algoritması kullanarak SZ hastalarını ayırt etmeye çalışmışlar. Özellikleri elde etmek için farklı yöntemler (Higuchi'nin Fraktal Boyutu, Spektral Entropi, Hannon Entropy, Bilgi Entropisi ve Kolmogorov Karmaşıklık değerleri gibi) kullanmışlar. Bu özellikleri DVM ile sınıflandırarak 88.5% doğruluk elde etmişler.

(Devia ve ark. 2019), sağlıklı ve hasta bireyler arasındaki ERP (olayla ilgili potansiyeller) karşılaştırması yapmıştır. Sonrasında sınıflandırıcı tasarımını ve performansını incelemişler. Çalışmalarında sırasıyla Doğrusal Diskriminant Analizi (DDA) sınıflandırıcılar, sonra kural tabanlı sınıflandırıcılar ve son olarak birleşik yaklaşım metotlarımı sunmuşlardır. En iyi sınıflandırıcı, genel doğruluk oranı 71%, hastaların tespiti için 81% duyarlılığa ve kontrollerin tespiti için 59% özgüllüğe sahiptir.

(Piryatinska, Darkhovsky, ve Kaplan 2017), 84 ergen bireye (39 sağlıklı- 45 SZ) ait 16 kanallı EEG kayıtlarının sınıflandırılmasını gerçekleştirmiştir. Rastgele orman (RO) sınıflandırıcısı ve DVM sınıflandırıcılarını kullanılmıştır. RF ile en iyi performans, ortalama 83,6% doğrulukla yedi boyutlu özellik alanı kullanılarak elde etmişlerdir.

(Shim ve ark. 2016), 34 sağlıklı ve 34 SZ hastasına ait EEG kanalları üzerinde özellik çıkarımı işlemini sensör düzeyinde ve kaynak düzeyinde yaparak, elde ettikleri özellikleri DVM ile sınıflandırmış ve 88.24% doğrulukla SZ hastalarını tespit

etmişlerdir.

(Sui ve ark. 2014), 48 SZ hastası ve 53 sağlıklı bireye ait EEG kayıtları üzerinde özellik çıkarımı işlemini Çoklu set kanonik korelasyon analizi yöntemi ile gerçekleştirmiş ve elde edilen özellikleri DVM ile sınıflandırarak 74% doğrulukla SZ hastaları ile sağlıklı bireyleri ayırt etmiştir.

(Boostani, Sadatnezhad, ve Sabeti 2009), EEG kayıtlarına Otoregresif (AR) model parametreler, bant gücü ve fraktal boyut yöntemleri uygulayarak elde ettikleri özellik değerlerini DDA ile sınıflandırmış ve 87.5% başarım elde etmişlerdir.

(Siuly ve ark. 2020), 49 SZ hastası ve 32 sağlıklı bireye ait EEG kayıtlarından ampirik mod ayrıştırma (AMD) tabanlı özellikler elde ederek Karar Ağacı (KA), KNN, DVM ve toplu torbalı ağaç (TTA) sınıflandırıcıları ile sınıflandırılmıştır. TTA, bildirilen diğer sınıflandırıcılara kıyasla daha iyi performans elde ederek 89.59% başarım elde etmiştir. Çizelge 2.1.'de, SZ tespitinde geleneksel makine öğrenme algoritmaları kullanılanan ilgili literatür çalışmaları yer almaktadır.

Çizelge 2.1. Makine öğrenimini kullanarak şizofreniyi tespit eden ilgili literatür çalışmaları

Yazar	Sınıflandırma	EEG veri seti	Özellik Çıkarımı	Başarım Oranı
(Johannesen ve ark. 2016)	DVM	40 SZ- 12 healthy	Alpha, beta, theta1, theta2, gamma frekans bileşenleri.	87.0 %
(Devia ve ark. 2019)	DDA	11 SZ- 9 healthy	ERPs(Event-related potentials) özellikleri	71.0%
(Kim ve ark. 2015)	ROC analizi	90 SZ- 90 healthy	Frekans bantlarını kullanarak spektral analiz	62.20%.
(Piryatinska, Darkhovsky, ve Kaplan 2017)	RF	45 SZ- 39 healthy	ϵ-complexity katsayıları	83.60%
(Shim ve ark. 2016)	DVM	34 SZ- 34 healthy	kaynak düzeyinde özellikler ve sensör düzeyinde özellikler	88.24%

Çizelge 2.1. Makine öğrenimini kullanarak şizofreniyi tespit eden ilgili literatür çalışmaları (devamı)

Yazar	Sınıflandırma	EEG veri seti	Özellik Çıkarımı	Başarım Oranı
(Sui ve ark. 2014)	DVM	48 SZ- 53 sağlıklı	Çoklu set kanonik korelasyon analizi	74.0%
(Boostani, Sadatnezhad, ve Sabeti 2009)	DDA	13 SZ- 18 sağlıklı	Bant gücü, fraktal boyut ve otoregresif (AR) model parametreleri	87.50%
(Thilakvathi ve ark. 2017)	DVM	55 SZ- 23 sağlıklı	Boyut, Yaklaşık Entropiler, Shannon entropisi, Bilgi entropisi, Kolmogorov karmaşıklığı, Spektral entropi, Higuchi'nin Fraktal	88.50%
(Siuly ve ark. 2020)	EBT	49 SZ- 32 sağlıklı	AMD (Ampirik Mod Ayrıştırma)	89.59%
(Aslan ve Akın 2019b)	KNN	45 SZ- 39 sağlıklı	RWT(Göreceli Dalgacık Enerjisi)	90%
(Santos-Mayo, San-José-Revuelta, ve Arribas 2016)	DVM ÇKA	16 SZ- 31 sağlıklı	J5 özellik çıkarma ve EEGLAB özellik çıkarma,	ÇKA: 93.42% DVM: 92.23%
(Dvey-Aharon ve ark. 2015)	Zaman - Frekans dönüşümü ve ardından Özellik Optimizasyonu	50 SZ- 25 sağlıklı	Stockwell dönüşümü	92% -93.9%

b) Derin Öğrenme Yöntemlerine Dayalı Çalışmalar:

Yukarıda bahsedilen ilgili alana ait literatür çalışmalarında, EEG kayıtları üzerinden özellik çıkarımı yapılarak elde edilen özellikler geleneksel makine öğrenme algoritmaları ile sınıflandırılmıştır. Bu özellik çıkarım yöntemi tekniğinin, başarılı tahmin gerçekleştirmesi gibi avantajları bulunmasına rağmen bu alanda kapsamlı bilgiye sahip uzmanlar gerektirmesi önemli bir dezavantajdır. Yeni gelişmelerle birlikte DL, geleneksel makine öğrenme algoritmalarının yerine yeni bir alternatif olarak araştırılmaktadır. DL'nin herhangi bir ön işlem ve özellik çıkarımı yapılması gerektirmemesinden dolayı son zamanlarda bu alan oldukça yoğun olarak çalışılmaktadır. ESA, bir çeşit DL ağıdır ve bir girdi verisine ait özellikler ESA'nın ilgili katmanlarından otomatik olarak elde edilir. SZ hastalarına ait EEG kayıtlarını kullanarak ESA ile sınıflandırma yapan yalnızca birkaç literatür çalışması bulunmaktadır.

Bir çalışmada (Phang, Noman, ve ark. 2019), kullandıkları yöntemde beyin fonksiyonel bağlantı bilgilerini özellik olarak kullanmış ve bu özellikleri birkaç yöntem ile test etmişlerdir. Özellikleri, kısmi yönlendirilmiş tutarlılık, vektör otoregresif modeli ve ağ topolojisinin karmaşık ağ ölçümleri gibi yöntemler kullanarak elde etmişler. Sonrasında, bu özelikleri kullanarak iki ESA modeline girdi olarak beslemişlerdir. En son aşama da ise, sağlıklı kontrolleri ve SZ hastalarını sınıflandırabilen Tam Bağlı Sinir Ağı (TBSA) bulunmaktadır. Sınıflandırma sonucunda 93.06% başarım elde etmişlerdir. Yöntemlerinin tatmin edici bir doğruluğa ulaştığı bildirmektedirler. Ancak yöntemleri, beyin bağlantı özellikleri gibi ek verilere ihtiyaç duymaktadır.

(Oh ve ark. 2019), Bir ESA modeli kullanarak, 28 bireye ait (14 sağlıklı - 14 SZ) 19 kanallı EEG kayıtlarını sınıflandırmışlardır. Çalışmalarında kullandıkları ESA modeli toplam 11 katmandan oluşmaktadır. Ön işleme yapmamışlar ve ESA modelini doğrudan ham EEG kanallarından alınan veriler ile beslemişlerdir. Yöntemlerinin, konu tabanlı olmayan testler için 98.07% ve konu tabanlı testler için 81.26% 'lik bir başarım elde ettiklerini bildirilmiştir. SZ'yi tespit etmek için DL kullanan her iki yöntem de, yorumlanabilir sonuçlar elde etmekten yoksundur.

Başka bir çalışmada, (Aslan ve Akın 2020), EEG kanallarından elde ettikleri 5sn'lik segmetleri spectrogram görüntülerine dönüştürmüşlerdir. Elde ettikleri

görüntüleri, ESA ile eğittiklerinde farklı iki veri seti için 95% ve 97% başarım ile oldukça başarılı sonuçlar elde etmişlerdir. Çizelge 2.2.'de, SZ tespitinde DL algoritmaları kullanılanan ilgili literatür çalışmaları yer almaktadır.

Çizelge 2.2. Derin öğrenme kullanarak şizofreniyi tespit eden ilgili literatür çalışmaları.

Yazar	Derin öğrenme metodu	Başarım Oranı	EEG veri seti
(Phang, Ting, ve ark. 2019)	ESA	93.06%	45 SZ - 39 sağlıklı
(Oh ve ark. 2019)	ESA	98.07%	14 SZ -14 sağlıklı
(Aslan ve Akın 2020)	ESA	95%. 97%	45 SZ - 39 sağlıklı 14 SZ -14 sağlıklı

Çizelge 2.3. Şizofreni tespitinde derin öğrenme metodu ile segmantasyon yapan ve kullandığımız veri setlerini kulllanan çalışmalar

Yazar	EEG veri seti	Derin öğrenme metodu	Segment boyutu
(Oh ve ark. 2019)	14 SZ -14 sağlıklı	ESA	25sn
(Naira ve Del Alamo 2019)	45 SZ-39 sağlıklı	ESA	10sn
(Shalbaf ve ark. 2020)	14 SZ -14 sağlıklı	ESA	5sn
(Chandran ve ark. 2020)	14 SZ -14 sağlıklı	LSTM	5sn

Çizelge 2.3' de kullandığımız veri setlerini kullanan çalışmaları ve bu çalışmaların segmantasyon işleminin hangi aralıklarla yapıldığı gösterilmektedir. Segmantasyon boyutu daha çok deneysel çalışmalarla belirlenmektedir. Ancak kullandığımız veri seti ile yapılmış çalışmalarda genellikle 5sn kullanılmıştır. Bu nedenle tez çalışmasının ikinci aşamasında derin öğrenme çalışması kısmında segmantasyon işlemi 5sn olarak belirlenmiştir.

3. MATERYAL VE METOT

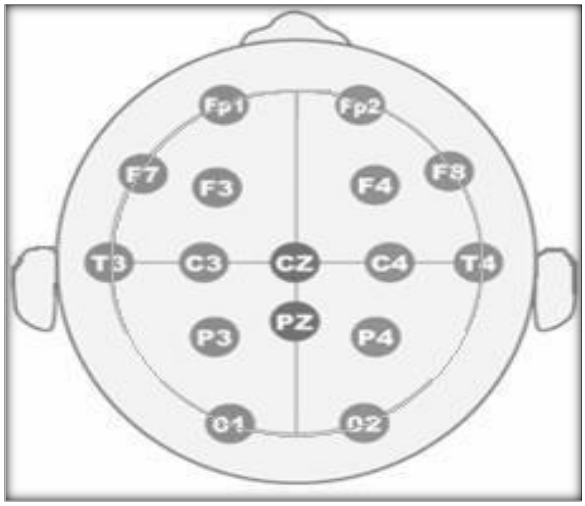
Bu tez çalışmasında, farklı yaş gruplarından oluşan SZ hastası ve sağlıklı bireylere ait iki EEG veri seti kullanılmıştır. Bu veri setleri öncelikle, farklı geleneksel makine öğrenme algoritmaları ve farklı özellik çıkarım yöntemleri kullanılarak SZ tespitindeki başarımları analiz edilmiştir. Sonrasında ise EEG sinyallerine ait önemli özelliklerin alınabilmesi için zaman frekans dönüşümüyle elde edilen spektrogram ve skalogram görüntüleri DL ile sınıflandırılarak yöntemin başarımları incelenmiştir.

3.1. Veri Setleri

SZ hastalığının tespitnde yöntemin doğruluğunun açık bir şekilde ortaya konabilmesi için farklı yaş gruplarına ait iki EEG veri seti kullanarak deneysel çalışmalar gerçekleştirilmiştir.

3.1.1. Veri Seti A

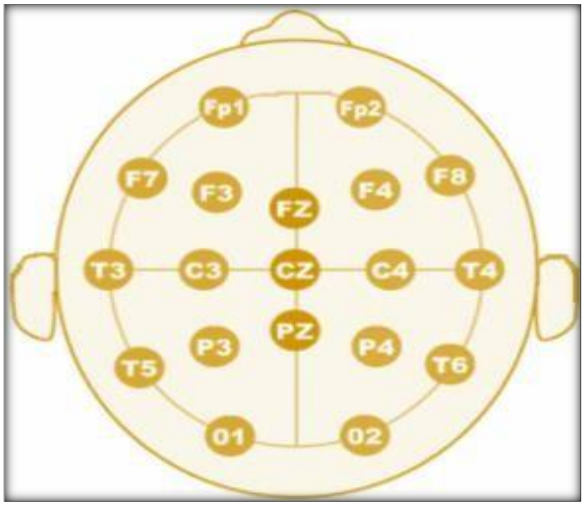
Bu çalışmada kullanılan ilk veri seti, Ruh Sağlığı Araştırma Merkezi (MHRC) tarafında SZ hastalarının en büyüğü 14, en küçüğü 10 yaşında 8 aylıkken, şizofreni hastası olmayan kontrollerin en büyüğü 13 yaşında 9 aylık, yaşı en küçük olan ise 11 yaşındadır. SZ hastası olan 45 çocuğa ve 39 sağlıklı çocuğa ait 60 saniyelik EEG kayıtlarından oluşmaktadır. Veri setindeki tüm SZ hastaları, Ruh Sağlığı Araştırma Merkezi (MHRC) uzmanları tarafından gözden geçirilmiş ve veri setinde kontrollerin hiçbiri kimyasal tedavi almamıştır. SZ hastası ve sağlıklı kontrollerin yaş ortalaması 12 yıl 3 aydır. EEG verileri, 16 kanaldan alınmış ve örnekleme frekansı 128Hz olarak kaydedilmiştir. Kaydetme esnasında denekler uyanık ve gözleri kapalı durumdadır. EEG elektrot dizilimi, uluslararası 10-20 standardına uygun olarak T3,T4, T5, T6, O1, O2, F3, F4, F7, F8, P3, P4, Pz, C3, C4 ve Cz şeklinde konumlandırılmıştır (Borisov ve ark. 2005). Şekil 3.1’ de 16 elektrotlu EEG’ye ait elektrot dizilimi gösterilmektedir.



Şekil 3.1. Veri Kümesi A için 16 kanallı elektrot yerleşimi

3.1.2. Veri Seti B

Bu çalışmada kullanılan diğer veri seti, Polonya, Varşova'daki Psikiyatri ve Nöroloji Enstitüsü tarafından 14 SZ hastası ve 14 SZ hastası olmayan kontrolden oluşan toplam kişiye ait 19 kanallı EEG kayıtlarından oluşmaktadır. EEG kayıtları 27.3 ± 3.3 ve 28.3 ± 4.1 ortalama yaşları olan 14 erkek ve 14 kadına ait 12 ile 15 dakika arasında değişen sürelerdedir. EEG verileri, örnekleme frekansı 250Hz ile deneklerin gözleri kapalı olarak kaydedilmiştir. Kayıtları 19 elektrottan alınmış ve O1, O2, C3, Cz, C4, Fp1, Fp2, F3, F7, Fz, F4, F8, T3, T4, T5, P3, Pz, P4 ve T6 şeklinde dizilmiştir (Olejarczyk ve Jernajczyk 2017). Şekil 3.2' de 19 elektrotlu EEG'ye ait elektrot dizilimi gösterilmektedir.



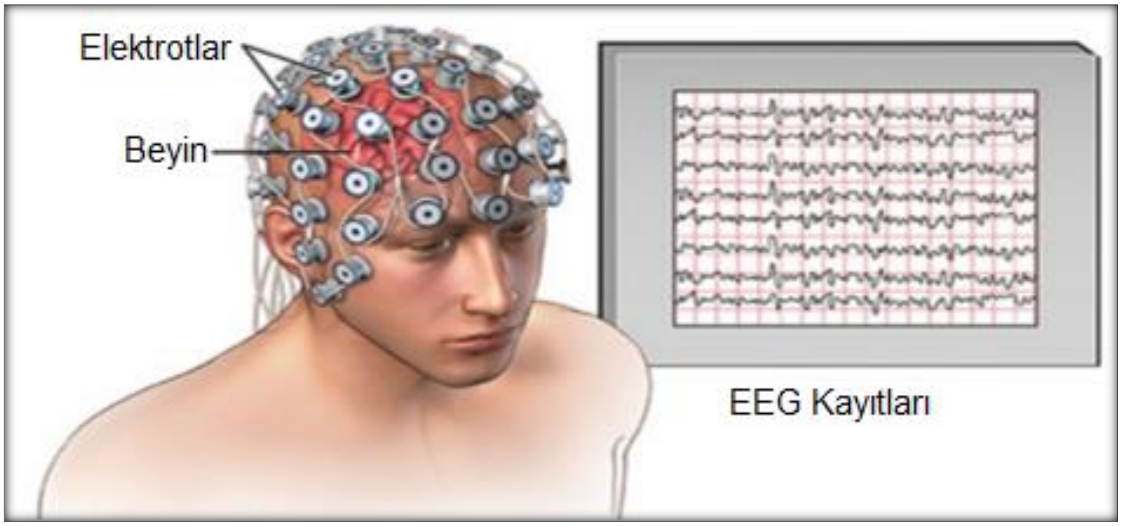
Şekil 3.2. Veri Kümesi B için 19 kanallı elektrot yerleşimi

3.2. METOT

3.2.1. Elektroensefalogram (EEG)

Elektroensefalogram (EEG) beyinde fonksiyonlarının ve beyindeki elektriksel değişimlerin incelenmesi için kullanılan önemli bir araçtır. Beyindeki nöronlar elektriksel bağlantılarla iletişim sağlarlar. Nöronlar arasındaki bu iletişim esnasında var olan sorunları bulmak için EEG kullanılabilir. EEG kaydındaki elektriksel değişimler, zirveleri ve vadileri olan dalgalı çizgiler gibi görünür. Bu çizgilerdeki değişimler, uzmanların sıradışı bir yapıda olup olmadığını hızlı bir şekilde değerlendirmesini olanak tanır. Bu çizgilerdeki herhangi bir sıradışılık, nöbet veya diğer beyin deformasyonlarını belirlemek için çnemli bir gösterge olabilir. Beyindeki nöronlar arasındaki tüm iletişim temelde bir elektrik ağıdır - tüm zihinsel aktiviteler (bilinçli ya da başka türlü), elektrik akımları kullanarak birbirlerine sinyaller gönderen bir nöron ağından oluşur. Nöronlar arasındaki iletişim de meydana gelen daha fazla elektrik sinyali, beyinde meydana gelen daha fazla aktivitenin bir işaretidir.

EEG beyindeki az miktardaki nöron araasındaki ilişkiyi tespit edemez, ancak aynı anda sinyal gönderecen binlerde nöronun elektriksel varyasyonlarını tespit edebilir. Elektrotlardan alınan sinyal bilgisi sonradında bir amplifikatöre gönderilerek, sinyalin güçlendirilmesi sağlanır. Bilgisayar yardımıyla alınan sinyal bilgisi, beyin aktivitelerinin haritası oluşturulmak üzere hızlı bir zamansal çözünürlükle beyindeki çeşitli etkinliklerin haritaları elde edilebilir. Elektrotlar beynin yüzeyindeki elektriksel değişimi ölçtüğünden, sinyalin tam olarak nerede meydana geldiği (kortekste) mi yoksa daha derin bir alanda mı oluşturulduğunu bilmek zordur bu nedenle uzaysal çözünürlük EEG'nin en önemli dezavantajıdır. Bu dezavantajı önlemek için geliştirilmiş hesaplamalar bulunmaktadır, ancak uzaysal çözünürlüğün tam olarak belirlenememesi EEG için önemli bir dezavantaj olma durumunu sürdürmektedir.(Aslan ve Akın 2018; Blocka, Karla 2020). Şekil 3.3.'de EEG elektrotları ve sinyalinin bir örneği gösterilmektedir.



Şekil 3.3. EEG elektrotları ve sinyalinin bir örneği (Cheema, Zhang, ve Chang 2016)

3.2.2. Geleneksel Makine Öğrenmesi Algoritmaları

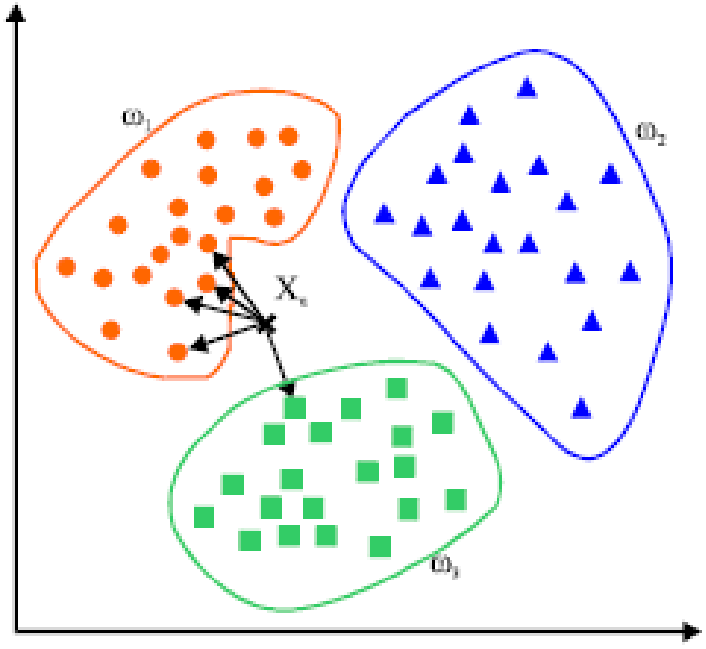
3.2.2.1. K-En Yakın Komşu Sınıflandırıcı

K-en yakın komşu (KNN) sınıflandırmarıcı geleneksel makine öğrenme algoritmaları arasında en basit bir sınıflandırıcıdır. Girdi verisi olarak kullanılan bir özellik vektörünün sınıflandırılması (X), en yakın eğitim doğrularını (k) münasip bir uzaklık metrik değerine göre tespit ederek yapılır. X özellik vektörü daha sonra, en yakın komşuların çoğunluğunun bağlı olduğu sınıfa atanır (Latifoğlu ve ark. 2008). KNN sınıflandırıcısı, metrik öklit uzaklığını kullanır ve en yakın komşuluklarda bir mesafe işlevi ve bir oylama işlevine dayanmaktadır. KNN sınıflandırıcı, en uygun değerler k için tatmin edici bir başarım sonuçları gösterdiği belirtilen klasik parametrik olmayan denetimli sınıflandırmacıdır (Duda, Hart, ve Stork 1998).

Öğreticili öğrenme algoritmalarının geneli bir eğitim ve bir test evrelerini içermektedir. KNN öğrenme algoritmasıda bu evreleri içermektedir. Eğitim evresinde, n-boyutlu bir bölgede veri noktaları belirtilir. Eğitim verisine ait bu noktalarda, hangi sınıfa ait olduklarını belirten etiketler ile ilişkilendirilmişlerdir. Etiketlenmemiş veriler test evresinde ağa verilir ve sınıflandırıcı etiketlenmemiş noktaya en yakın k (en önceden sınıflandırılmış) veri noktalarının dökümünü meydana getirir. Sınıflandırıcı daha sonra o veri noktalarının dökümünde hangisinin çok olduğuna bakarak etiketlenmemiş veriyi çok olan sınıfa dahil eder. (Hall 1998). KNN sınıflandırıcısı aşağıdaki aşamalardan oluşmaktadır (El-Dahshan, Hosny, ve Salem 2010):

1. Verileri eğitim ve test verilerine bölün.
2. Bir değer seçin K
3. Hangi mesafe fonksiyonunun kullanılacağını belirleyin.
4. Sınıflandırılması gereken test verilerinden bir örnek seçin ve n eğitim numunesine olan mesafeyi hesaplayın.
5. Elde edilen mesafeleri sıralayın ve k-en yakın veri örneklerini alın.
6. Test sınıfını, k komşularının oy çoğunluğuna göre sınıfa atayın.

(Aslan ve Akın 2019b,Chakure 2021). Şekil 3.4'de k-en yakın komşu sınıflandırma algoritmasının çalışma yapısı gösterilmektedir.



Şekil 3.4. (KNN) k-en yakın komşu sınıflandırma algoritması.

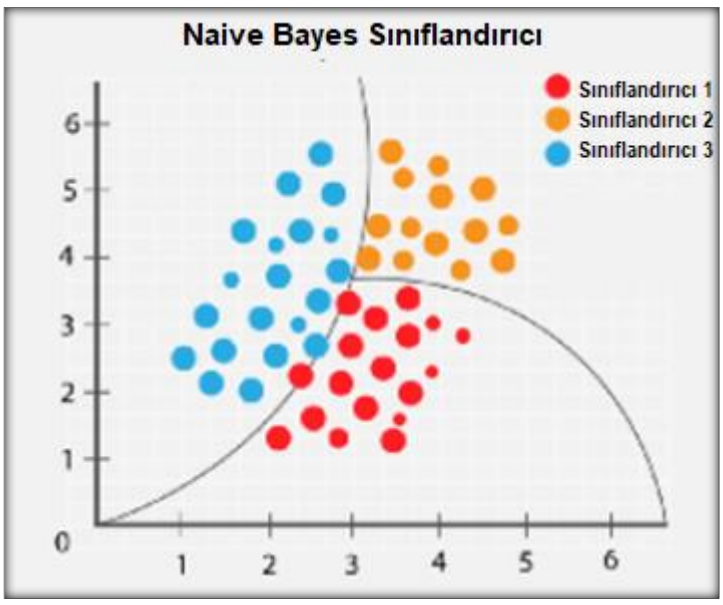
3.2.2.2. Naive Bayes

Naive Bayes, bir olasılık modelini bir karar kuralı ile birleştiren basit bir sınıflandırıcıdır. Sınıflandırıcı, özellik vektörünü oluşturan özelliklerin istatistiksel olarak bağımsız olduğu temel varsayımına göre çalışır. Naive Bayes, Bayes karar teorisine dayanmaktadır ve amaç, sınıflandırma hatası olasılığının en aza indirilmesi ve arka olasılığın maksimizasyonudur.

Sınıflandırıcı, Gauss varsayımını benimser, yani her bir marjinal iki parametre

(ortalama ve varyans) tarafından tanımlanır, bu da hesaplama karmaşıklığını ve veri ihtiyacını azaltır. Sınıflandırıcının gerekli sınıflandırma parametrelerini tahmin etmesi, aşırı uyumdan kaçınması ve Naif Bayes'i karmaşık sınıflandırma problemleri için uygun hale getirmesi için az sayıda eğitim verisine ihtiyaç vardır.

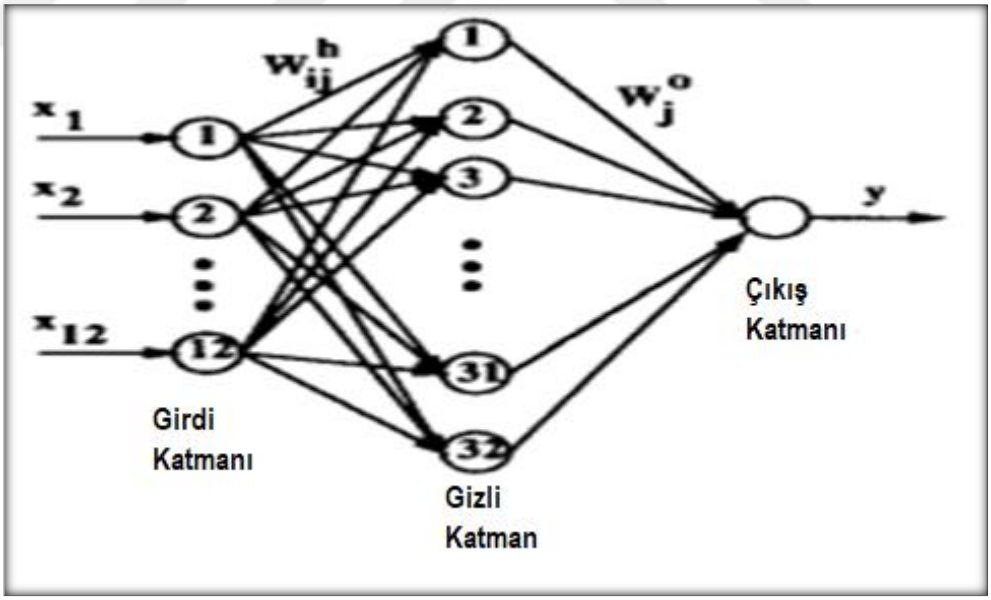
Naive Bayes sınıflandırıcı, önemli veri sınıflarını tanımlayan modelleri çıkaran veri analizi biçimidir. Naive bayes sınıflandırıcı, dolandırıcılık tespiti, performans, hedef pazarlama ve tıbbi analiz gibi çeşitli uygulamalara sahiptir. Naive Bayes sınıflandırıcı, EEG sinyaline uygulanan normal veya anormal olarak sınıflandırmak için kullanılır ve büyük veri tabanına uygulanan yüksek doğruluk sergiler. Bu yöntemde Naive bayes sınıflandırıcı, EEG sinyalini normal veya anormal sinyale göre sınıflandırmaktır. Sınıf üyeliğini tahmin eden ve başka bir sınıflandırıcıya göre yüksek doğruluğa sahip istatistiksel sınıflandırıcıdır. Belirli bir sınıftaki öznitelik değerinin diğer özniteliklerin değerlerinden bağımsız olduğunu ve buna sınıfın koşullu bağımsızlığı denildiğini varsayalım; hesaplamayı basitleştirir ve bu duruma Naive denir (Dimitoglou, Adams, ve Jim 2012, Tzimourta ve ark.2018,Celin ve Vasanth 2018, Aslan ve Akın 2018; Pawlak 2003). Şekil 3.5’ de Naive Bayes sınıflandırma algoritmasının çalışma yapısı gösterilmektedir.



Şekil 3.5. Naive Bayes sınıflandırma algoritması.

3.2.2.3. Çok Katmanlı Algılayıcı (ÇKA)

Yapay sinirsel disiplinin doğuşu, bu tür ağlarla prensipte herhangi bir aritmetik veya mantıksal işlevi hesaplayabileceğimizi gösteren W. McCulloch ve W. Pitts ile 1940'lara dayanmaktadır. Eğitim, çevresinden alınan uyaranlara tepki olarak ağın parametrelerini değiştirmek için kullanılan dinamik ve yinelemeli bir süreçtir. Denetimli eğitim, gerçek ağ çıkışı ile istenen çıktı arasında doğrudan bir karşılaştırma yaparak ağ parametrelerini ayarlar. Denetimsiz eğitim hiçbir hedef değer içermez. Girdilerden gelen bilgileri, veri boyutluluğundaki içsel bir azalma veya toplam girdi verisi miktarı ile ilişkilendirmeye çalışır. Eğitim türü, parametre değişikliklerinin nasıl gerçekleştiğine göre belirlenir. ÇKA (Çok katmanlı algılayıcı) (Rosenblatt 1957), nöronları ağırlıklı bağlantılarla birbirine bağlayan ardışık ileri beslemeli katmanlardan oluşan bir sinir ağıdır. Giriş katmanı, giriş sinyallerini toplamak için kullanılır ve çıkış katmanı yanıtlar sağlar. Aktarım için bir veya daha fazla gizli katman eklenir. ÇKA eğitimi, hata gradyan yayılımı ile gerçekleştirilir. (Guezzaz 2019, Freund ve Mason 1999; Latifo\uglu ve ark. 2008, Farrell, Mammone, ve Assaleh 1994; Morgan ve Scofield 1991) . Şekil 3.6’ de tek gizli katmanlı ÇKA yapısı gösterilmektedir.



Şekil 3.6. Tek gizli katmanlı ÇKA

Çevresel modelleme, muhtemelen kombinasyon halinde çeşitli yaklaşımların kullanılmasını içerir. En uygun yaklaşımı seçmek, ele alınan sorunun karmaşıklığına ve sorunun anlaşılma derecesine bağlıdır. Yeterli veri ve bilgi işlem kaynaklarını

varsaymak ve problemin güçlü bir teorik anlayışı varsa, o zaman tam bir sayısal model belki de en çok arzu edilen çözümdür. Bununla birlikte, genel olarak, bir problemin karmaşıklığı arttıkça teorik anlayış azalır (sistemler arasındaki kötü tanımlanmış etkileşimler nedeniyle) ve istatistiksel yaklaşımlar gereklidir. Son zamanlarda, sinir ağlarının ve özellikle ÇKA'nın kullanımının, daha geleneksel istatistiksel tekniklere etkili alternatifler olduğu gösterilmiştir. ÇKA hemen hemen her türlü pürüzsüz, ölçülebilir işleve yaklaşacak şekilde eğitilebilir. Diğer istatistiksel tekniklerin aksine, ÇKA, veri dağıtımı ile ilgili önceden hiçbir varsayımda bulunmaz. Son derece doğrusal olmayan işlevleri modelleyebilir ve yeni, görünmeyen verilerle sunulduğunda doğru bir şekilde genelleme yapmak için eğitilebilir. ÇKA'nın bu özellikleri, onu sayısal modeller geliştirmeye ve ayrıca istatistiksel yaklaşımlar arasında seçim yaparken çekici bir alternatif haline getirir.

ÇKA, Şekil 3.6.'da gösterildiği gibi, bir giriş vektörü ile bir çıkış vektörü arasındaki doğrusal olmayan bir eşleştirmeyi temsil eden bir model olan, birbirine bağlı basit nöronlardan veya düğümlerden oluşan bir sistemden oluşur. Düğümler, basit bir doğrusal olmayan transfer veya aktivasyon fonksiyonu ile modifiye edilen düğüme girişlerin toplamının bir fonksiyonu olan ağırlıklar ve çıkış sinyalleri ile bağlanır. Çok katmanlı algılayıcının son derece doğrusal olmayan fonksiyonları yaklaşık olarak tahmin etmesini sağlayan, birçok basit doğrusal olmayan transfer fonksiyonunun üst üste binmesidir. Transfer fonksiyonu doğrusal olsaydı, çok katmanlı algılayıcı yalnızca doğrusal fonksiyonları modelleyebilirdi. Kolayca hesaplanan türevi nedeniyle, yaygın olarak kullanılan bir transfer fonksiyonu, lojistik fonksiyondur. Bir düğümün çıkışı, bağlantı ağırlığı ile ölçeklenir ve ileriye doğru beslenerek, sonraki katmanındaki düğümlere bir girdi olarak ağa beslenir. Bu, bir bilgi işlem yönü anlamına gelir, bu nedenle çok katmanlı algılayıcı, ileri beslemeli bir sinir ağı olarak bilinir. Çok katmanlı bir algılayıcının mimarisi değişkendir, ancak genel olarak birkaç nöron katmanından oluşacaktır. Giriş katmanı hiçbir hesaplama rolü oynamaz, yalnızca giriş vektörünü ağa geçirmeye hizmet eder. Giriş ve çıkış vektörleri terimleri, çok katmanlı algılayıcının girdi ve çıktılarını ifade eder ve Şekil 3.6.'da gösterildiği gibi tek vektörler olarak temsil edilebilir. Çok katmanlı bir algılayıcı, bir veya daha fazla gizli katmana ve son olarak bir çıktı katmanına sahip olabilir. Çok katmanlı algılayıcılar, her bir düğümün sonraki ve önceki katmandaki her düğüme bağlı olduğu, tamamen bağlı olarak tanımlanır.

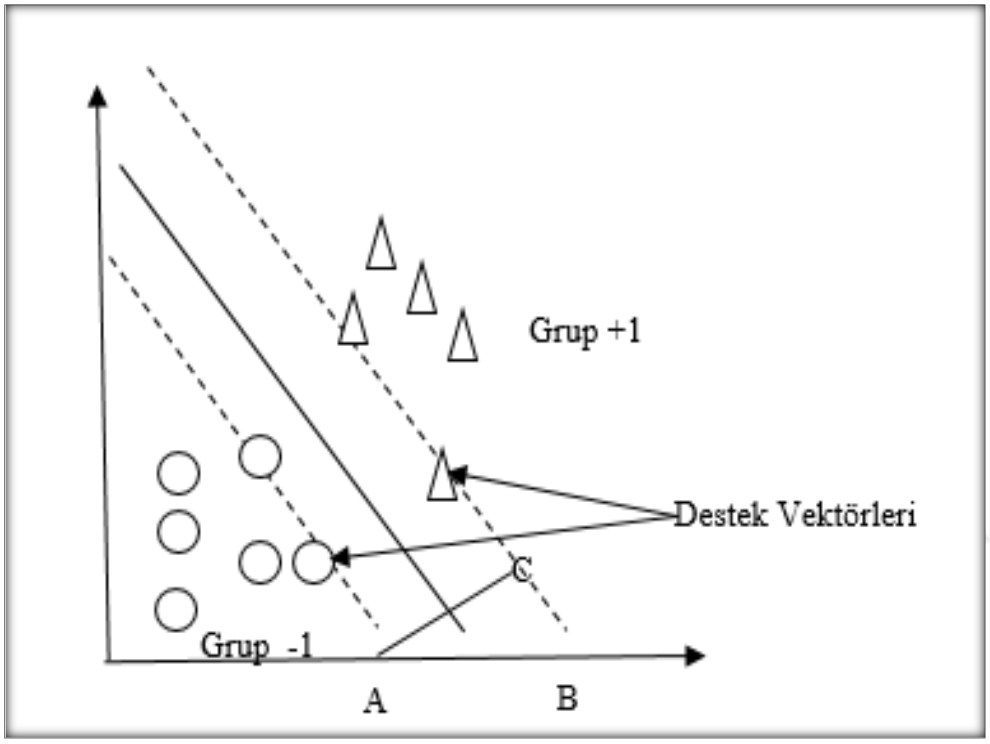
Uygun bir bağlantı ağırlıkları ve transfer fonksiyonları seti seçerek, çok katmanlı bir algılayıcının giriş ve çıkış vektörleri arasındaki herhangi bir pürüzsüz, ölçülebilir fonksiyonu yaklaşık olarak tahmin edebildiği gösterilmiştir. ÇKA'lar, eğitim yoluyla öğrenme yeteneğine sahiptir. Eğitim, bir dizi girdi ve ilişkili çıktı vektörlerinden oluşan bir dizi eğitim verisi gerektirir. Eğitim sırasında çok katmanlı algılayıcı eğitim verileriyle tekrar tekrar sunulur ve ağdaki ağırlıklar istenen girdi-çıktı eşlemesi gerçekleşene kadar ayarlanır. ÇKA'lar denetimli bir şekilde öğrenirler. Eğitim sırasında, belirli bir giriş vektörü için çok katmanlı algılayıcıdan gelen çıktı istenen çıktıya eşit olmayabilir. Bir hata sinyali, istenen ve gerçek çıkış arasındaki fark olarak tanımlanır. Eğitim, ağdaki ağırlıkların ne derece ayarlanması gerektiğini belirlemek için bu hata sinyalinin büyüklüğünü kullanır, böylece ÇKA'nın genel hatası azaltılır. ÇAKA'yı eğitmek için kullanılabilecek birçok algoritma vardır. Uygun şekilde temsili eğitim verileriyle eğitildikten sonra, ÇKA yeni, görünmeyen girdi verilerini genelleştirebilir.(Gardner ve ark. 1998, Aslan ve Akın 2019c; Haykin 1994).

3.2.2.4. Destek Vektör Makinaları (DVM)

Destek vektör makinesi (DVM), güçlü teorik temeli nedeniyle hızlı bir şekilde aktif bir araştırma alanı olarak gelişen çağdaş bir makine öğrenme aracıdır. DVM'ler, Vapnik'in istatistiksel öğrenme teorisinden icat edilmiştir. DVM, ikili bir sınıflandırma problemini denetimli bir şekilde çözmek için özel olarak formüle edilmiştir ve öğrenme problemi, hata yüzeyinin herhangi bir yerel minimumdan bağımsız olduğu ve global bir optimuma sahip olduğu ikinci dereceden bir optimizasyon problemi olarak formüle edilmiştir. İkili sınıflandırma problemleri durumunda, DVM'nin temel amacı, iki sınıf arasındaki ayrım marjını maksimize edecek şekilde optimal bir ayırıcı hiper düzlem oluşturmaktır. Makine, bu istenen özelliği, yapısal risk minimizasyonu ilkesi temelinde elde eder. SRM ilkesi, eğitim hatası ile bir güven aralığının toplamından oluşan genelleme hatasının üst sınırını en aza indirmeyi amaçlamaktadır.

Doğrusal olarak ayrılabilir örüntüler için DVM tabanlı sınıflandırıcılar geliştirmek için, x_i'nin n boyutlu girdi özelliği vektörü ve d_i'nin hedef çıktıyı temsil ettiği bir eğitim kümesini ele alalım. Hedef çıktı $d_i = 1$ ile temsil edilen girdi örüntüleri pozitif grubu, hedef çıktı $d_i = -1$ ise negatif grubu oluşturur. Doğrusal olarak ayrılabilir durum için, giriş vektörlerinin iki sınıfını ayırmak için bir hiper düzlem bulmak gerekir.

Giriş modellerini ayırmak için çok fazla hiper düzlem olabilir, ancak tek ve yalnızca bir tane optimal ayırıcı hiper düzlem olacaktır. En uygun ayırma alt düzlemi, en yakın eğitim noktaları ile kendisi arasındaki mesafenin mümkün olduğunca maksimum olacağı şekilde iki sınıf arasında geçen bir hiper düzlemdir.(Chandaka ve ark. 2009, Cristianini, Shawe-Taylor,ve ark. 2000, Collobert ve Bengio 2001, Aslan ve Akın 2019c; Campbell ve ark. 2006). Şekil 3.7' de lineer olarak ayrılabilen DVM yapısı gösterilmektedir.



Şekil 3.7. Lineer olarak ayrılabilen DVM

3.2.2.5. Rastgele Orman

Breiman tarafından önerilen rastgele orman sınıflandırıcısı, birçok bireysel sınıflandırma ağacından oluşur, burada her ağaca kendi sınıflandırma çıktısı için belirli bir ağırlık verilen bir sınıflandırıcıdır (Weka'da tüm ağaçlara aynı ağırlık verilir). Tüm ağaçların sınıflandırma çıktıları, tüm ağaçların sınıflandırma çıktısının modunun (en çok oyu alan çıktı) seçilmesiyle yapılan genel sınıflandırma çıktısını belirlemek için kullanılır. Rastgele orman sınıflandırıcısının oluşturulması için adımlar şu şekilde özetlenebilir:

- Rastgele orman (RO) sınıflandırıcısının öğrenme prosedürü, her bir ağacın

oluşturulmasıyla başlar. Literatürde rastgele ağaçlar, CART, J48, REPTree, C4.5, CHAID, vb. Gibi birçok ağaç oluşturma yöntemi kullanılmıştır. Weka yazılımı, sınıflandırıcıyı budama yapmadan oluştururken rastgele ağaçları kullanır.

- Bir sonraki adım, önyükleme tekniği kullanılarak eğitim veri setinin örneklenmesiyle oluşturulan torba içi veri seti adı verilen her ağaç için eğitim veri setini hazırlamaktır (değiştirme ile örnekleme). Torba içi veri setindeki epoch sayısı, eğitim veri setindeki epoch sayısına eşittir. Torba içi setteki verilerin yaklaşık üçte biri kopyalanır (değiştirme ile örnekleme), bu nedenle, ağaç sınıflandırma performansını test etmek için torba dışı olarak adlandırılan kalan eğitim veri seti kullanılır. Her ağaç farklı bir önyükleme örneği kullanılarak oluşturulur.

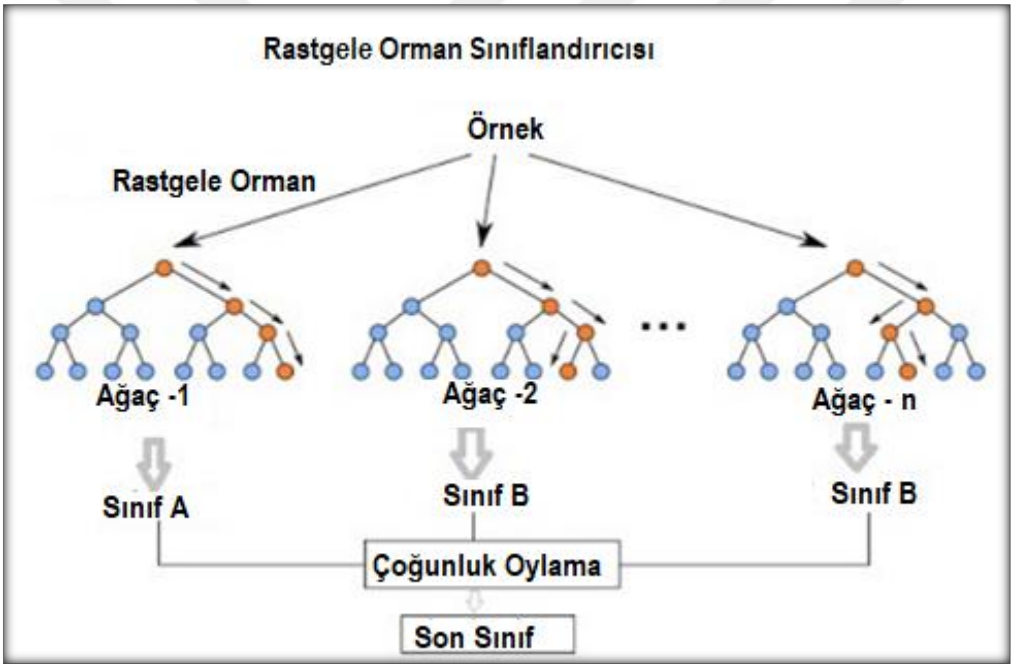
- Her ağaç için, rastgele sayıda özellik seçililir ve rastgele bir ağaç sınıflandırıcısının düğümlerini ve yapraklarını oluşturmak için kullanılır. Breiman'a göre, özelliklerin rastgele seçilmesi, sınıflandırmanın doğruluğunu artırmakta, verilerdeki gürültüye duyarlılığı azaltmakta ve özellikler arasındaki korelasyonu en aza indirmektedir.

- Rastgele bir ağaç oluşturmak, tüm eğitim veri setiyle (çanta içi veri seti) ağacın tepesinden başlar. İlk adım, kök düğümde bir öznitelik (özellik) seçmeyi ve ardından eğitim verilerini özniteliğin her olası değeri için alt kümelere ayırmayı içerir. Bu, özniteliğin her olası değeri için bir dal oluşturur. Kök düğümün bölünmesi ve seçimi, özniteliğin bölünmesinin bilgi kazancına dayalı olarak yapılır (en yüksek bilgi kazanımına sahip öznitelik, kök düğümde seçilir).

- Bilgi kazancı pozitifse düğüm bölünür; aksi takdirde düğüm bölünmez ve eğitim alt kümesindeki en yaygın uyku aşamasının bir çıktısı verilen bir yaprak düğüm haline gelir.

- İşlem, dala ve kalan özniteliklere ulaşan alt küme kullanılarak her dal düğümünde özyinelemeli olarak tekrarlanır (özyinelemeli bölümleme). Bir sonraki özniteliğin seçimi de kalan özniteliklerin en yüksek bilgi kazanımına göre yapılır. Bölme işlemi (ağaç büyütme), tüm öznitelikler seçilene kadar devam eder. Bu düğüme ulaşan eğitim alt kümesinde en çok meydana gelen uyku aşaması, sınıflandırma çıktısıdır.

- İşlem, tüm ağaçları oluşturmak için tekrarlanır.
- Rastgele orman sınıflandırıcısının sınıflandırma çıktısı, en çok meydana gelen uyku aşamasıdır.
- Tüm ağaçları oluşturduktan sonra, torba dışı veriler tek tek ağaçları ve tüm ormanı test etmek için kullanılır. Ortalama yanlış sınıflandırma hatası, torba dışı hata, her bir ağacın oyunun ağırlıklarını ayarlamak için kullanılabilir. Rastgele orman sınıflandırıcısının Weka uygulaması, her ağaca aynı ağırlığı verir (Fraiwav ve ark. 2012,Aslan ve Akın 2019c; Emanet 2009). Şekil 3.8' de Rastgele Orman sınıflandırma algoritmasının yapısı gösterilmektedir.

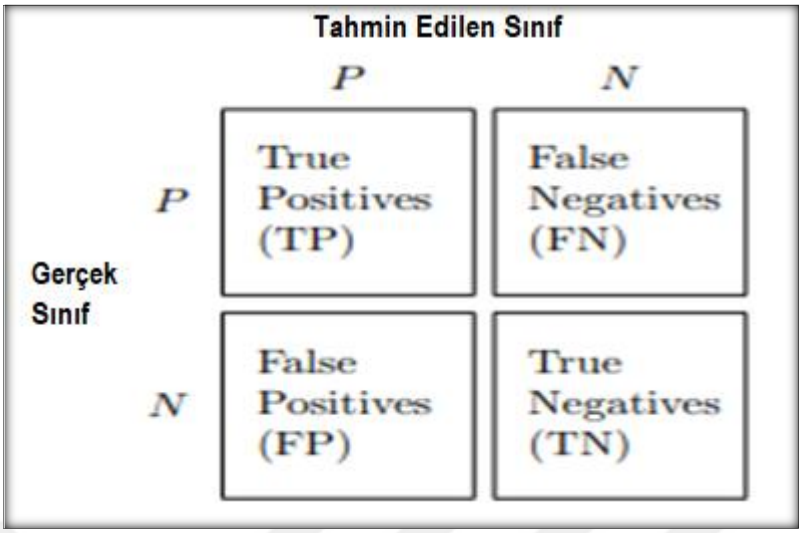


Şekil 3.8. Rastgele Orman sınıflandırma algoritması.

3.2.3. Değerlendirme Metrikleri

Çalışmanın deneylerinden elde edilen sonuçlar, iyi bilinen ve yaygın olarak kullanılan bir dizi değerlendirilmiş metrikle doğrulanır. Bu ölçütlerin ayrıntıları ve yorumları diğer belgelerde (Goutte ve Gaussier 2005; Raschka 2014) açıklanmıştır ve bu nedenle bu ölçütlerin yalnızca temel hesaplamalarını dâhil ediyoruz. Şekil 3.9'da karışıklık matrisinin yapısı gösterilmektedir.

3.2.3.1. Karışıklık Matrisi



Şekil 3.9. Karışıklık Matrisi

3.2.3.1.1. Tahmin hatası ve doğruluğu

$$ERR = \frac{FP+FN}{FP+FN+TP+TN} = 1 - ACC \quad (1)$$

$$ACC = \frac{TP+TN}{FP+FN+TP+TN} = 1 - ERR \quad (2)$$

3.2.3.1.2. Yanlış ve doğru pozitif oranlar

$$\text{FPR} = \frac{FN}{N} = \frac{FP}{FP+TN} \quad (3)$$

$$\text{TPR} = \frac{TP}{P} = \frac{TP}{FN+TP} \quad (4)$$

3.2.3.1.3. Kesinlik, duyarlılık, F1 skoru

F1 Skoru, bir sınıflandırıcı tüm yönleriyle tek bir FPR veya TPR ölçüsünden daha iyi bir şekilde değerlendirilebilir daha yetenekli bir metriktir. Hesaplamak için, TPR veya Hassasiyet (SEN) olarak bilinen F1 Skoru, Kesinlik (PRE) ve Duyarlılık metrikleri önceden hesaplanmalıdır (Goutte ve Gaussier 2005; Raschka 2014).

Denklemler. (5) - (7) F1 skorunun hesaplanmasını gösterir.

$$PRE = \frac{TP}{TP+FP} \quad (5)$$

$$REC = TPR = \frac{TP}{P} = \frac{TP}{FN+TP} \quad (6)$$

$$F_1 = 2.\frac{PRE,REC}{PRE+REC} \quad (7)$$

Yaygın olarak kullanılan ek bir ölçü olan Özgüllük (SPC), sınıflandırıcının yanlış sınıflandırmalardan kaçınmada ne kadar iyi olduğunu ölçer ve ideal olarak 1'e eşittir.

$$SPC = TNR = \frac{TN}{N} = \frac{TN}{FP+TN} \quad (8)$$

3.2.4. Özellik Seçim Algoritmaları

Özellik seçimi, özellik seçimi kriterine göre orijinal özellik kümesinden bir özellik alt kümesi bulma işlemi olarak tanımlanır. Ana hedefi, tüm özellik setiyle karşılaştırılabilir tahmini performansa sahip bir özellik alt kümesi bulmaktır. Özellik seçimi, istenmeyen yapılar oluşturmaz, yani orijinal özellik alanında gerçekleştirilir. Bu, orijinal veri kavramını kaybetmeden gereksiz ve / veya ilgisiz özelliklerin kaldırılmasıyla elde edilir. Özellik seçimi, verileri basitleştirmek için ana tema olan boyut azaltmanın bir yönüdür. Buradaki asıl soru, sonuçların kalitesinden ödün vermeden hazırlanan verilerin bir kısmı olabilir. Boyut indirgeme, verileri yüksek boyutlu bir uzaydan alıp boyutsallığı çok daha küçük olan yeni bir alana eşleştirme işlemidir. Verilerin boyutluluğunu azaltmak için birkaç neden vardır. Örneğin, yüksek boyutlu veriler, hesaplama açısından verimlilik zorluğunu getirir. Ayrıca, yüksek boyutlu veriler, öğrenme algoritmasının zayıf genelleme yeteneklerine yol açabilir (Abd-Alsabour 2014, Lotte ve ark. 2007, Elder ve Pregibon 1996).

3.2.4.1. Genetik Arama Algoritması

Genetik Algoritmalar (GA), 1975'te Holland tarafından geliştirilmiştir. Holland, bunları karmaşık özelliklerin ve geniş arama alanlarının optimizasyon problemlerine çözüm aramak için bir araç olarak görmüştür. GA'lar esasen Charles Darwin'in 1830'larda evrim teorisine getirdiği "en uygun için hayatta kalma" ilkesini taklit etmek için geliştirilmiştir. Bu perspektiften ve optimizasyon, uygunluk veya hayatta kalma becerisine benzer olduğundan, iyileştirme ve optimizasyona GA yaklaşımını uygulamak mantıklıdır. Standart GA metodolojisinde arama, rastgele oluşturulmuş bir ilk çözümle başlar. İlk çözüm, bir amaç fonksiyonu kullanılarak değerlendirilir. Üremenin doğasına (gelecek neslin üyeleri) karar vermek için ilk çözüme bir seçim şeması uygulanır. Çoğaltılan her nesil, rekombinasyon operatörlerine (mutasyon ve çapraz geçiş) tabi tutulur. Her nesil çözümün sonunda bir yakınsama testi yapılır. Yakınsama testine dayalı olarak, başka bir nesil çözümü durdurma veya devam etme kararı verilir. Bu arama süreci, bazı yakınsama koşulları karşılanana kadar devam eder. (Aslan ve Akın 2019a; Huang ve Wang 2006, Al-Aomar R. 2000).

3.2.4.2. Parçacık Sürüsü Optimizasyon Arama Algoritması

Parçacık Sürüsü Optimizasyon Arama Algoritması, verileri bir sınıflandırıcıya atamadan önce veri setinin dikkatli bir şekilde değerlendirilmesini içerir. Pek çok alakasız özellik eklemek yerine, sınıflandırma sürecini çok daha kolay hale getirmek için sadece gerekli özellikleri dikkate almaktır. Bu nedenle, ilgili ve önemli özellikleri seçebilen yeterli tekniklere sahip olmak faydalıdır. Ayrıca, sınıflamada özellik seçimi benimsenirse, sınıflandırıcının önemli özelliğinin bulunmasına ve iş yükünün azaltılmasına yardımcı olur ve bu da sınıflandırma doğruluğunu artırır. Mevcut literatürün gözden geçirilmesine dayanarak, parçacık sürüsü optimizasyonu, diğer mevcut özellik seçimine kıyasla, sınıflandırma doğruluğu açısından daha iyi seçimden yararlanır. Özellik seçimi için PSO'nun potansiyel avantajları aşağıdaki gibidir:

- PSO, optimum çözüm bulunana kadar güçlü bir keşif yeteneğine sahiptir, çünkü farklı parçacıklar çözüm uzayının farklı kısımlarını keşfedebilir.

- Parçacık sürüsü hafızaya sahip olduğundan ve çözüme ilişkin bilgi, sorun alanı içinde uçarken tüm parçacıklar tarafından korunduğundan, özellik seçimi için özellikle caziptir.
- PSO'nun çekiciliği, aynı zamanda, hala iyi performans sağlayan hesaplama açısından ucuz uygulamasından kaynaklanmaktadır.
- PSO, tek bir çözüm yerine potansiyel çözüm popülasyonuyla çalışır.
- PSO, ikili ve ayrık verileri ele alabilir.
- PSO, bellek ve çalışma süresi açısından diğer özellik seçme tekniklerine kıyasla daha iyi performansa sahiptir ve karmaşık matematiksel operatörlere ihtiyaç duymaz.
- PSO'nun birkaç parametre ile uygulanması kolaydır, gerçekleştirmesi kolaydır ve umut verici sonuçlar verir.
- PSO'nun performansı, sorunun boyutundan neredeyse hiç etkilenmez.

PSO, sosyal davranış simülasyonundan esinlenen bir hesaplama tekniğidir. PSO'nun orijinal konsepti, uçan kuşların davranışlarını ve problemleri çözmek için bilgi alışverişi araçlarını simüle etmektir. (Sakri 2018, Aslan ve Akın 2019a; Shirvany ve ark. 2013).

3.2.4.3. Evrimsel Arama Algoritması

Temelinde, Evrimsel arama (EA) algoritmaları tasarım ve uygulamalarında anahtar unsurlar olarak evrimsel süreçlerin hesaplamalı modellerini kullanan bilgisayar tabanlı problem çözme algoritmalarını tanımlamak için kullanılır. Geçtiğimiz birkaç yıl içinde çeşitli evrimsel algoritmalar geliştirildi. Bunların her biri farklı bir yaklaşım oluşturur; ancak, aynı doğal evrim ilkesinden ilham alırlar. (Dehuri ve ark. 2013, Goldberg 1989, Aslan ve Akın 2019a; Ng ve ark. 1994).

3.2.4.4. En İyi İlk Algoritması

En genel haliyle en iyi ilk arama, basit bir sezgisel arama algoritmasıdır. Burada "Sezgisel", en iyi çözümü veya hatta herhangi bir çözümü garanti etmeyen, ancak problem çözme için yararlı bir kılavuz görevi gören genel bir problem çözme kuralı veya bir dizi kural anlamına gelir. En iyi ilk arama, grafik tabanlı bir arama algoritmasıdır

(Dechter ve Pearl 1985), yani arama alanı, yollarla birbirine bağlanan bir dizi düğüm olarak gösterilebilir.

"En iyisi" adı, önce en iyi "puana" sahip düğümü keşfetme yöntemini ifade eder. Her aday düğüme bir puan atamak için bir değerlendirme işlevi kullanılır. Algoritma, biri henüz keşfedilmemiş adayların bir listesini içeren (AÇIK) ve diğeri ziyaret edilen düğümlerin bir listesini içeren (KAPALI) iki liste tutar. Her ziyaret edilen düğümün ziyaret edilmemiş tüm ardıl düğümleri AÇIK listesine dahil edildiğinden, algoritma yalnızca en son ziyaret edilen düğümün ardıl düğümlerini keşfetmekle sınırlı değildir. Diğer bir deyişle, algoritma, yakın komşular gibi yalnızca küçük bir alt kümeyle sınırlı kalmak yerine, her zaman grafiği çizilen tüm ziyaret edilmemiş düğümlerin en iyisini seçer. Önce derinlik ve en başta gibi diğer arama stratejileri bu kısıtlamaya sahiptir. Bu stratejinin avantajı, algoritma çıkmaz bir düğüme ulaşırsa, diğer düğümleri denemeye devam etmesidir (Pearl 1984).

3.2.4.5. Karınca Kolonosi Optimizasyonu Arama Algoritması

Gerçek hayatta karınca topluluklarında, feromon maddesi karıncalar arasında dolaylı bir iletişim aracı olarak kullanılmaktadır. Karıncalar besin kaynağını tespit ettiklerinde, diğer karıncalarında erişimini sağlamak amacıyla yolu berlirlemek için bir miktar feromon bırakırlar. Bırakılan feromonun ölçüsü, besin kaynağının uzaklığına, miktarına ve kalitesine bağlı olarak değişir. Besin kaynağına giden yol üzerinde bulunmayan bir karınca feromonu tespit ettiğinde yolu izlemeye karar vermesi olasıdır. Karınca besin kaynağına giderken bir miktar feromon bırakacağından yol üzerindeki feromon izini daha belirgin yapacaktır. Bundan dolayı kullanılan yolun daha fazla karınca tarafından izlenmesi daha ilgi çekici hale gelecektir. Dorigo ve ark. (Dorigo ve Di Caro 1999) karınca kolonisinin bu şekildeki davranışlarını kabulleniş ve karmaşık optimizasyon problemlerini çözümlemek için karıncaların koloni davranışına benzer bir yapay algoritma sunmuştur (Aslan ve Akın 2019a; Khushaba ve ark. 2008).

3.2.5. Kısa Zamanlı Fourier Dönüşümü

Ayrık Fourier dönüşümü, sinüzoidlerin toplamı olarak bir sinyali ifade eder. Sinüzoidlerin zaman süresi sonsuz olduğundan, sinyalin ayrık Fourier dönüşümü, zaman içinde tüm bir sinyalin spektral içeriğini yansıtır, ancak spektral içeriğin ne zaman oluştuğunu göstermez. Bununla birlikte, bazı durumlarda, bir sinyalin spektral içeriğini

kısa bir zaman ölçeğinde değerlendirmek faydalı olabilir. Kısa zaman ölçeklerinde spektral içeriği değerlendirmek için KZFD kullanılabilir.

KZFD, aynı zamanda pencereli Fourier dönüşümü veya kayan Fourier dönüşümü olarak da adlandırılır, sinyali bir pencere işlevi ile çarparak zaman alanı giriş sinyalini birkaç ayrık veya üst üste binmiş bloğa böler ve ardından ayrı Fourier dönüşümünü her bloğa uygular. Kayan pencereler olarak da adlandırılan pencere işlevleri, genliğin kenarlarda kademeli ve yumuşak bir şekilde sıfıra doğru daraldığı işlevlerdir. Her blok farklı zaman periyotlarını işgal ettiğinden, ortaya çıkan KZFD, karşılık gelen her zaman periyodunda sinyalin spektral içeriğini gösterir. Kayan pencere hareket ettirildiğinde, sinyalin spektral içeriğini farklı zaman aralıklarında elde edilir. Bu nedenle, KZFD, bir sinyalin spektral içeriğinin zaman içinde nasıl geliştiğini gösteren bir zaman ve frekans fonksiyonudur. KZFD katsayıları adı verilen karmaşık değerli, 2 boyutlu bir dizi, pencereli Fourier dönüşümlerinin sonuçlarını depolar. KZFD katsayılarının büyüklükleri bir zaman-frekans spektrumu oluşturur ve KZFD katsayılarının fazları bir faz zaman-frekans spektrumu oluşturur.

KZFD, zaman-frekans analizi yapmak için en basit yaklaşımlardan biridir ve zaman-frekans analizi kavramını kolayca anlamanıza yardımcı olabilir. KZFD, hızlı Fourier dönüşümü (HFD) kullandığı için hesaplama açısından verimlidir.

Doğrusal KZFD yöntemini faz spektrumuna ihtiyaç duyduğunuzda veya sinyal yeniden oluşturulmasına ihtiyaç olmadığında kullanılabilir. Örneğin, faz spektrumu otomatik konuşma tanıma uygulamalarında yardımcı olabilir. Bir uygulamada yalnızca büyüklük spektrumuna ihtiyaç varsa, doğrusal KZFD karesi olan ikinci dereceden KZFD spektrogram yöntemi kullanılmalıdır. KZFD spektrogramı, basitliği nedeniyle en popüler ikinci dereceden zaman-frekans analiz yöntemlerinden biridir (LabVIEW 2010 Advanced Signal processing 2020).

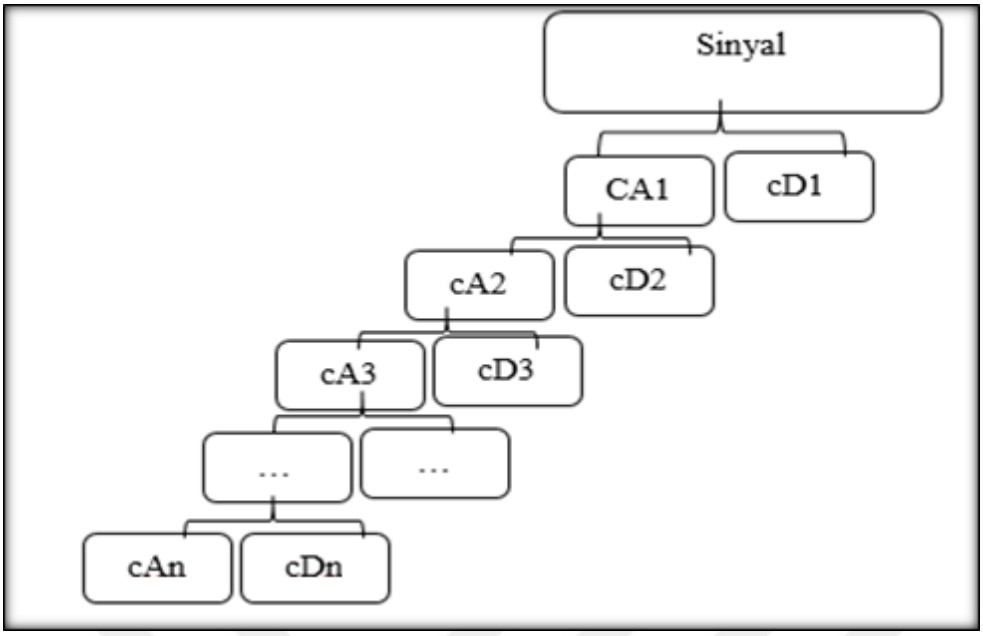
3.2.6. Dalgacık Dönüşümü

Dalgacık geliştirme ve kullanımının arkasındaki temel motivasyon, durağan olmayan sinyaller için verimli bir sinyal analiz aracına ihtiyaç duyulmasıdır. Sinyal analizi için yaygın olarak kullanılan Fourier dönüşümü, yalnızca sinyalin sinüs dalgalarının toplamları olarak temsil edildiği durağan sinyallerin analizi için uygundur. Bununla birlikte, dar bir geçici sinyal gibi durağan olmayan bir sinyal için, frekans içeriği,

sinyali tüm zaman için değerlendirerek belirlenir. Bu nedenle, geçici olayın kesin zaman konumu bilinmemektedir.

"Yerel" bir Fourier dönüşümü elde etmek için, pencereli bir Fourier dönüşümü tanımlanmıştır. Kısa süreli Fourier dönüşümü (KZFD) veya Gabor dönüşümü olarak adlandırılan Fourier dönüşümünün değiştirilmiş versiyonudur, sonlu uzunluktaki bir pencereden bakıldığında durağan olmayan sinyalin durağan olduğunu varsayar. Ancak en önemli dezavantaj, pencere bir kez seçildiğinde, zaman-frekans çözünürlüğünün tüm zaman frekansı düzleminde sabitlenmesidir. Bunun nedeni, tüm frekanslarda aynı pencerenin kullanılmasıdır.

Sonuç olarak, zaman çözünürlüğü bazı frekanslarda tatmin edici olabilirken, frekans çözünürlüğü diğer frekanslarda tatmin edici olabilir, ancak her iki çözünürlük aynı anda tatmin edici olmayabilir. Bununla birlikte, dalgacık dönüşümünün yüksek frekanslarda keyfi olarak iyi zaman çözünürlüğü sağlayabildiği gösterilmiştir. Aynı zamanda, dalgacık dönüşümü frekans çözünürlüğünü düşük frekanslarda keyfi olarak iyi hale getirebilir. Dalgacık dönüşümlerinin bu özelliği, görüntülerdeki kenarlar gibi birçok önemli özelliğin tespit edilmesinde büyük bir sonuca sahiptir. Dalgacıklar, küçük görüntü yapılarının dış hatlarını ve büyük nesnelerin sınırlarını tespit ederek, belirli görüntü modellerinin tanınmasına yardımcı olabilir. Bu nedenle dalgacık analizi, sinyal analizi için yeni ve etkili bir araç olarak ortaya çıkıyor ve bilgisayarla görme, konuşma ve görüntü sıkıştırma, kenar algılama, doku analizi ve örüntü tanıma dâhil olmak üzere çeşitli alanlarda uygulamalar bulmaktadır. (Aslan ve Akın 2019b, Shanmugam 2000). Şekil 3.10'da dalgacık dönüşümü kullanılarak yaklaşımlı ve detaylı alt kanallara ayrıştırma işlemi gösterilmektedir. Şekil 3.10' da dalgacık dönüşümü kullanılarak yaklaşımlı ve detaylı alt kanallara ayrıştırma işleminin yapısı gösterilmektedir.



Şekil 3.10. Dalgacık dönüşümü kullanılarak yaklaşımlı ve detaylı alt kanallara ayrıştırma.(Garg ve ark. 2011).

3.2.7. Göreceli Dalgacık Enerjisi (GDE)

Biyomedikal sinyallerin araştırma alanlarında, frekans alanı, zaman alanına kıyasla daha fazla bilgi sağlar. Biyomedikal sinyaller genellikle zaman alanı biçiminde çıkarıldığından, bu sinyallerin frekans alanına dönüştürülmesi gerekir. Bununla birlikte, beyin sinyallerinden duygu durumunun analizi için, hem frekans hem de sinyalin zaman bölümüyle ilişkisi daha önemli bilgiler sağlar, böylece bir Fourier Dönüşümü yöntemi daha esnek bir zaman penceresi yaklaşımı sağlayamaz. Fourier Transform yönteminden farklı bir yaklaşımda dalgacık analizi, belirli bir dalga formu şekli olan, sınırlı süreli ve ortalama değeri sıfır olan bir dalgacık kullanmıştır. Dalgacık analizi, sinyali analiz etmek için temel olarak bir "ana" dalgacık kullanırken, Fourier analizinde bir sinüs dalgası kullanılmıştır. Sinyali ayrıştırmak için, temel olarak ana dalgacık kaydırılır ve ölçeklenir, böylece orijinal sinyale birkaç kaydırılmış ve ölçeklenmiş dalgacıkların toplamı ile yaklaştırılabilir.

Birçok biyomedikal sinyal, yüksek frekansındakine kıyasla düşük frekansta daha önemli bilgiye sahiptir. Bu nedenle dalgacık analizinde düşük frekansa yaklaşım kısmı denir, detay kısmı ise yüksek frekans ile ilgilidir. Ayrık dalgacık dönüşümünde (ADD), yaklaştırma katsayıları için ölçekleme ve kaydırma faktörü ve ayrıştırılmış sinyallerin detay katsayıları, ikisinin gücüne göre seçilir. ADD, istenen temsili frekans bandını elde

edene kadar bir dizi yüksek geçiren ve alçak geçiren filtre kullanır. Bir dalgacık yapısı genel olarak şu şekilde tanımlanır:

$$\Psi_{a,b}(t) = |a|^{-\frac{1}{2}}\Psi(\frac{t-b}{a}) \quad (9)$$

a ölçekleme parametresidir, b kayan parametresidir, a, b∈R, a ≠ 0'dır. a ve b değerleri ayrıktır, dalgacık dönüşümünü bir ADD haline getirir. j, k∈Z ile $a_j = 2^{-j}$ ve $b_{j.k} = 2^{-jk}$ olduğunda, sonlu bir enerjiden oluşan bir ortonormal Hilbert uzayı L^2 (R) aşağıdaki dalgacık ailesinden oluşturulabilir:

$$\psi_{j,k}(t) = 2^{1/2}\,\psi(2^j t - k) \quad (10)$$

M uzunluğunda ayrık bir zaman sinyali S verildiğinde, birleştirici ADD katsayıları aşağıda verilen formülden elde edilir:

$$C_j(k) =< S, 2^{\frac{j}{2}}\Psi(2^j t - k) > \quad (11)$$

j = −m,…, −1 ve m = $\log_2$M ile. Bu denklem, 2^{-j} ölçeğiyle 2^{-jk} sinyal hakkında bilgi verir. Her ayrışma seviyesindeki dalgacık katsayısı sayısı 2^jM'ye eşit olacaktır.

Belirli bir biyomedikal sinyalden belirli bilgileri çıkarmak ve ayırt etmek için, seçilebilecek bir yöntem sinyal enerjisine bakmaktır. Ayrık bir sinyalin (E) enerjisi şu şekilde tanımlanır:

$$E = \sum_{i=1}^{N} {X_i}^2 \quad (12)$$

i = 1,… ile N, sinyal X_i'deki örnek sayısıdır.

Bu çalışmada, sınıflandırıcıya girdi olarak toplam enerji hesaplaması yerine her ayrıştırılmış sinyalin enerji yüzdesini veya görecei dalgacık enerjisini (GDE) kullanıyoruz. Bu araştırmada kullanılan EEG sinyali ADD'den sonra sinyalin toplam enerjisi daha sonra şu şekilde hesaplanır:

$$E_{total} = \sum_{j=1}^{M+1} E_j \quad (13)$$

E_j ile ADD dönüşümünden gelen yaklaşıklık ve detay sinyalinin enerjisi ve M ayrışma seviyesi hesaplanır. Böylece, RWE, ρi, şu şekilde tanımlanır:

$$\rho_j = \frac{E_j}{E_{total}} \; j = 1, \ldots\ldots, M+1 \quad (14)$$

$\Sigma\rho = 1$ ile. gde, farklı frekans bantları arasındaki enerji dağılımına ilişkin sinyal karakteristikleri hakkında bilgi sağlar. Bu, biyomedikal sinyallerin analizi için varsayımsal olarak uygundur (Aslan ve Akın 2019b; Kumar, Dewal, ve Anand 2012, Krisnandhika 2017).

3.2.8. Derin Öğrenme

Derin öğrenme (DL), bir verideki özelliklerin hiyerarşik olarak öğrenilmesini sağlayan ve daha derin sinir ağları içeren yeni bir makine öğrenme yöntemidir. DL ağlarının geleneksel makine öğrenme algoritmalarında üstün kılan en önemli yanı özellik çıkarım işleminin önceden yapılmasına gerek olmamasıdır. Geleneksel makine öğrenmesinde özellik değerlerinin önceden çıkartılması gerekmekteyken DL ağlarında önemli görülen özellikler otomatik olarak öğrenilir.

DL algoritmaları büyük ölçekli verilere ve verilerin işlenmesi için yüksek maliyetli donanıma ihtiyaç duymaktadır. Bu tür ihtiyaçların daha erişilebilir hale gelmesiyle geleneksel makine öğrenme algoritmalarının yerine DL kullanımı gittikçe yaygınlaşmaktadır. DL için gerekli olan ihtiyaçlara erişim olanakları son zamanlarda arttığında, tıbbi görüntülerin genellikle ESA gibi derin öğrenme ağlarıyla işlenmesi, analizi ve değerlendirmesi için alternatif bir yöntem olarak yaygın şekilde kullanılmaya başlamıştır. Bu tez çalışmasının ikinci aşamasında, ESA derin ağları yapısına sahip VGG-16 mimarisi kullanılarak spektrogram ve skalogram görüntüleri kullanılarak SZ tespiti yapılmaktadır (Aslan 2019).

3.2.9. Keras Derin Öğrenme Çatısı

Keras, yapay sinir ağları için bir Python arabirimi sağlayan açık kaynaklı bir kitaplıktır. Keras, TensorFlow kitaplığı için bir arabirim görevi görmektedir. 2.3 sürümüne kadar TensorFlow, Theano, Keras, Microsoft Cognitive Toolkit, R, ve PlaidML dahil olmak üzere çoklu arkaplan platformlarını desteklemektedir. 2.4 sürümünden itibaren yalnızca TensorFlow desteklenmektedir. Derin sinir ağları ile hızlı deneyler sağlamak için tasarlanmış olup, kullanım kolaylığı, modüler ve büyütülebilir olmaya odaklanmıştır. ONEIROS (Açık Uçlu Nöro-Elektronik Akıllı Robot İşletim Sistemi) projesinin araştırma çabasının bir parçası olarak geliştirilmiştir ve ana yazarı ve bakımcısı bir Google firmasında mühendis olarak çalışan François Chollet'dir ve XCeption derin sinir ağı modelinin yazarıdır(Chollet 2017; Keras 2020).

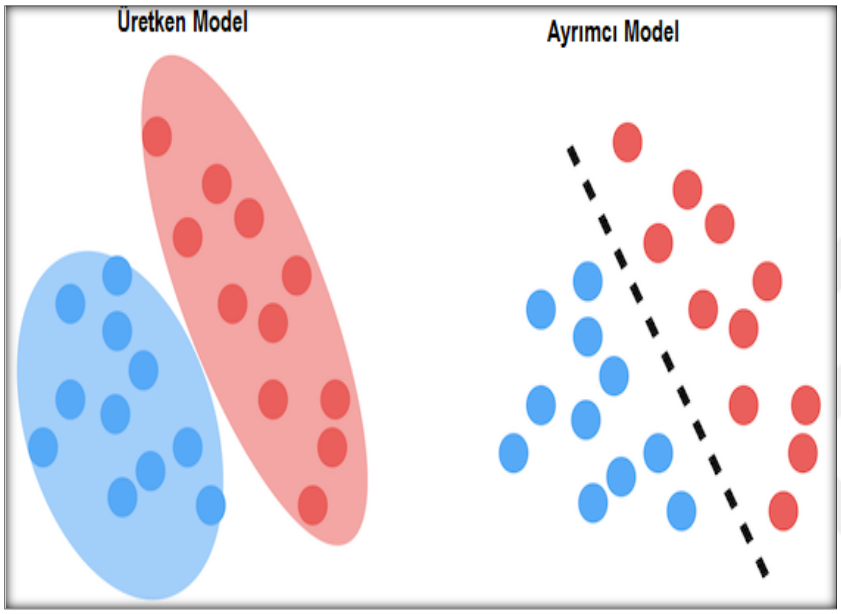
Keras, derin sinir ağı kodu yazmak için gerekli olan kodlamayı basitleştirmek ve görüntü ve metin verileriyle çalışmayı kolaylaştırmak için katmanlar, hedefler, etkinleştirme işlevleri, optimize ediciler ve bir dizi araç gibi yaygın olarak kullanılan sinir ağı yapı taşlarının sayısız uygulamasını içerir. Kod GitHub'da barındırılır ve topluluk destek forumları GitHub sorunları sayfasını ve bir Slack kanalını içerir. Standart sinir ağlarına ek olarak Keras, evrişimli ve tekrarlayan sinir ağları için desteğe sahiptir. Bırakma, toplu normalleştirme ve havuzlama gibi diğer yaygın yardımcı program katmanlarını destekler. Keras, kullanıcıların akıllı telefonlarda (iOS ve Android), web'de veya Java Virtual Machine'de derin modeller üretmesine olanak tanır. Ayrıca, Grafik işleme birimleri (GPU) ve tensör işleme birimleri (TPU) kümeleri üzerinde derin öğrenme modellerinin dağıtılmış eğitiminin kullanımına da izin verir (Keras 2020).

3.2.10. Üretken ve Ayrımcı Modeller

Üretken modeller, odak noktasının bir veri kümesindeki bireysel sınıfların dağılımı olduğu ve öğrenme algoritmalarının veri noktalarının temel modellerini / dağılımını modelleme eğiliminde olduğu modellerdir. Bu modeller teorik olarak ortak olasılık sezgisini kullanır ve belirli bir özellik (x) / girdi ve istenen çıktı / etiketin (y) aynı anda var olduğu durumlar oluşturur. Üretken modeller olasılık tahminlerini ve veri noktalarını modelleme ve bir veri kümesindeki farklı sınıf etiketleri arasında ayrım yapma olasılığını kullanır. Bu modeller yeni veri örnekleri oluşturabilir. Bununla birlikte, büyük bir dezavantajları da vardır. Aykırı değerlerin varlığı bu modelleri önemli ölçüde etkilemektedir. Naif Bayes, Gizli Markov Modeli ve DDA makine öğrenimi üretken modellerine örnek olarak verilebilir (Rufai 2020).

Koşullu modeller olarak da adlandırılan ayrımcı modeller, bir veri kümesindeki sınıflar / etiketler arasındaki sınırı öğrenme eğilimindedir. Üretken modellerden farklı olarak, buradaki amaç, bir sınıfı diğerinden ayıran karar sınırını bulmaktır. Dolayısıyla, üretken bir model, veri noktalarının ortak olasılığını modelleme eğilimindeyken ve olasılık tahminlerini ve maksimum olasılığı kullanarak yeni örnekler oluşturma yeteneğine sahipken, ayrımcı modeller (tam anlamıyla olduğu gibi) koşullu olasılığı modelleyerek sınıfları ayırır ve veri noktası hakkında herhangi bir varsayımda bulunmaz. Ayrıca yeni veri örnekleri üretemezler. Ayrımcı modeller, üretici modellerin aksine, aykırı değerlere karşı daha sağlam olma avantajına sahiptir. Bununla birlikte, büyük bir

dezavantaj, bir yanlış sınıflandırma problemidir, yani bir veri noktasını yanlış sınıflandırmaktır. Bu iki model türü arasındaki bir diğer önemli fark, üretken bir model verilerin nasıl üretildiğini açıklamaya odaklanırken, ayrımcı bir modelin verilerin etiketlerini tahmin etmeye odaklanmasıdır. Lojistik regresyon, DVM, Karar Ağacı ve RO makine öğreniminde ayrımcı modellere örnek olarak verilebilir (Rufai 2020). Şekil 3.11' de üretken ve ayrımcı model sınıflandırmasının çalışma yapısı gösterilmektedir.

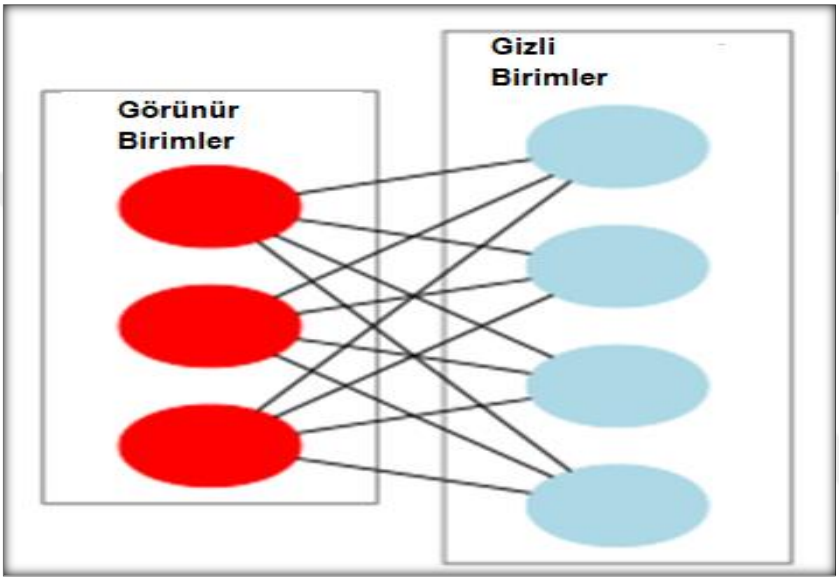


Şekil 3.11. Üretken ve Ayrımcı Model sınıflandırması

3.2.11. Kısıtlanmış Boltzmann Makinesi

Kısıtlı Boltzmann makinesi (KBM), kendi girdileri üzerinden bir olasılık dağılımını öğrenebilen üretken bir stokastik yapay sinir ağıdır. KBM'ler ilk olarak Harmonium adı altında 1986'da Paul Smolensky tarafından icat edildi ve Geoffrey Hinton ve işbirlikçilerinin 2000 yılının ortalarında onlar için hızlı öğrenme algoritmaları icat etmesinden sonra öne çıkmıştır. KBM'ler boyut azaltma, sınıflandırma, işbirliğine dayalı filtreleme, özellik öğrenimi, konu modelleme gibi alanlarda uygulamalar bulmuşlardır. Göreve bağlı olarak denetimli veya denetimsiz yollarla eğitilebilirler. Adından da anlaşılacağı gibi, KBM'ler, nöronlarının iki taraflı bir grafik oluşturması gerektiği kısıtlamasıyla Boltzmann makinelerinin bir çeşididir: iki birim grubunun her birinden bir çift düğüm (genellikle "görünür" ve "gizli" olarak adlandırılır) aralarında simetrik bir bağlantıya sahip olabilir ve bir grup içindeki düğümler arasında bağlantı yoktur. Aksine, "sınırsız" Boltzmann makinelerinin gizli birimler arasında bağlantıları olabilir. Bu

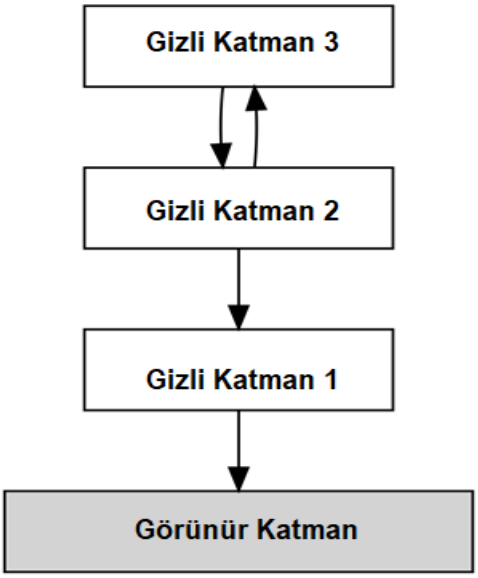
kısıtlama, genel Boltzmann makineleri sınıfı için mevcut olandan daha verimli eğitim algoritmalarına, özellikle de gradyan tabanlı kontrast sapma algoritmasına izin verir. Kısıtlanmış Boltzmann makineleri, derin öğrenme ağlarında da kullanılabilir. Özellikle derin inanç ağları, KBM'leri "istifleyerek" ve isteğe bağlı olarak ortaya çıkan derin ağın gradyan inişi ve geri yayılımı ile ince ayarını yaparak oluşturulabilir (Wikipedia 2020b). Şekil 3.12' de üç görünür birim ve dört gizli birim ile sınırlı bir Boltzmann makinesinin şeması gösterilmektedir.



Şekil 3.12. Üç görünür birim ve dört gizli birim ile sınırlı bir Boltzmann makinesinin şeması.

3.2.12. Derin İnanç Ağları

Makine öğreniminde, derin bir inanç ağı (DİA), katmanlar arasındaki bağlantılara sahip, ancak her katmandaki birimler arasında olmayan, çok sayıda gizli değişken katmanından ("gizli birimler") oluşan, üretken bir grafik model veya alternatif olarak bir derin sinir ağı sınıfıdır. Denetimsiz bir dizi örnek üzerinde eğitildiğinde, bir DİA girdilerini olasılığa dayalı olarak yeniden yapılandırmayı öğrenebilir. Katmanlar daha sonra özellik algılayıcıları olarak işlev görür. Bu öğrenme adımından sonra, bir DİA, sınıflandırma yapmak için gözetimle daha fazla eğitilebilir (Wikipedia 2020a). Şekil 3.13' de derin inanç ağını şematik olarak gösterilmektedir.



Şekil 3.13. Derin inanç ağına şematik genel bakış

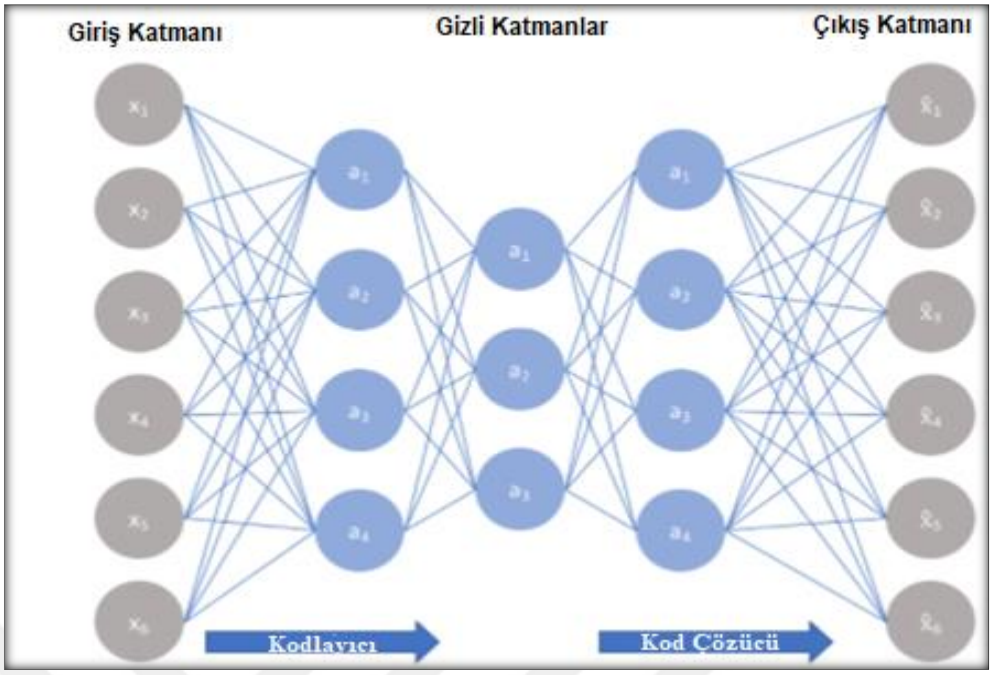
3.2.13. Otomatik Kodlayıcı

Otomatik kodlayıcı (OK), girişlerini çıkışlarıyla eşleştirmeyi amaçlayan bir sinir ağıdır. Verilerin nasıl verimli bir şekilde sıkıştırılıp kodlanacağını öğrenen ve ardından veriyi indirgenmiş kodlanmış gösterimden orijinal girdiye mümkün olduğunca yakın bir temsile yeniden yapılandıran, denetimsiz bir yapay sinir ağıdır. Bu nedenle, otomatik kodlayıcılar eğitim verilerinin en göze çarpan özelliklerini öğrenir ve bu da verilerin dağıtımı hakkında yararlı bilgiler elde etmelerini sağlar (Johannesen ve ark. 2016)(Team 2020).

Bileşenler:

Kodlayıcı: Modelin giriş boyutlarını nasıl azaltacağını ve giriş verilerini kodlanmış bir gösterime nasıl sıkıştıracağını öğrendiği yerdir.

Kod çözücü: Modelin, kodlanmış gösterimden verileri orijinal girdiye mümkün olduğunca yakın olacak şekilde nasıl yeniden yapılandıracağını öğrendiği yerdir. Şekil 3.14’ de otomatik kodlayıcı sinir ağının yapısı gösterilmektedir.

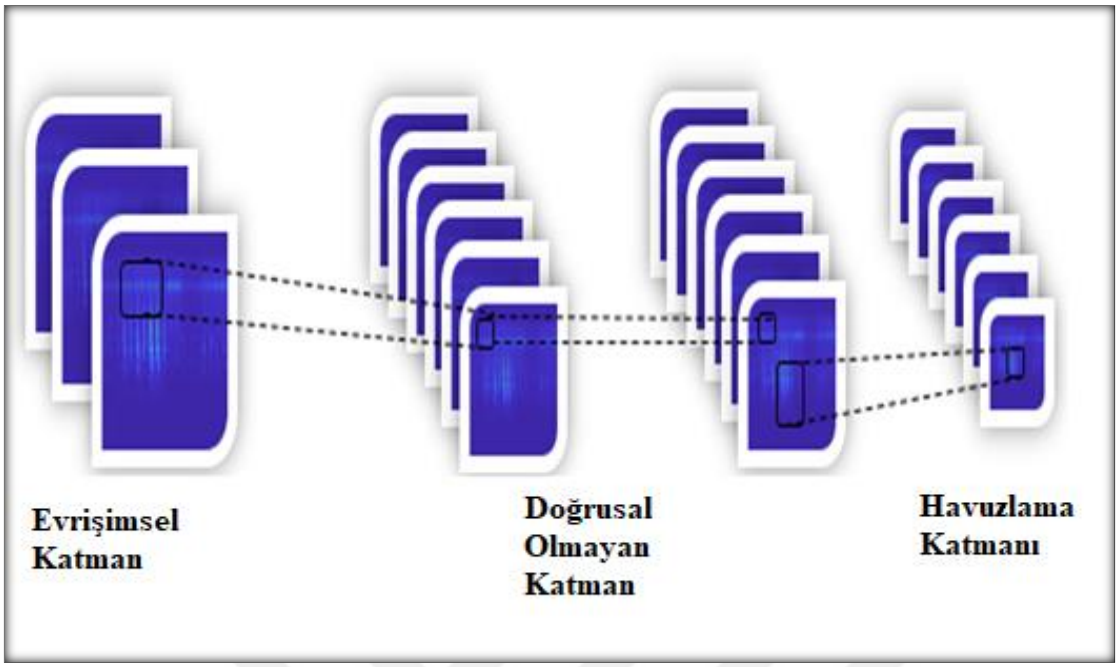


Şekil 3.14. Otomatik Kodlayıcı Sinir Ağı

3.2.14. Evrişimli Sinir Ağları (ESA)

Evrişimli Sinir Ağları, özelliklerin hiyerarşik olarak çıkarılabileceği ağ katmalarından meydana gelen bir çeşit derin öğrenme ağıdır. Özellikle görüntü gibi mutimedya veri türlerinde yaygın olarak kullanılmaktadır (Hubel ve Wiesel 1968). Bir ESA modeli, multimedya veri türlerini (ör. Görüntüler), ağ katmanları aracılığıyla özelliklerin hiyerarşik şekilde otomatik olarak çıkarılabileceği yapıda tasarlanmış DL ağlarıdır (Hubel ve Wiesel 1968). Bir ESA modelinin yapısı, genellikle iki önemli bölümden meydana gelmektedir. İlk bölüm özellik çıkarma işlemi için evrişimli ve bir araya getirme katmanlarını içerir. İkinci bölüm ise, sınıflandırma aşaması için Geleneksel Çok Katmanlı Algılayıcı (ÇKA) gibi çalışan Tam Bağlantılı Sinir Ağı katmanlarını içerir. Genel bir ESA modeli, özellik çıkarım işleminin her biri önceki katmanın çıktısından yapıldığı birkaç ardışık evrişimli katman ve havuzlama katmanından oluşmaktadır.

Basit özellikler (örneğin, görüntü üzerindeki basit çizgiler gibi), ağın ilk katmalarından elde edilir. Sonraki katmanlarda ise elde edilen bu basit özellikler kullanılarak daha karmaşık özellikler (örneğin, bir görüntüdeki nesne gibi) çıkartılır. Hücrelerin görsel öğelere benzer hiyerarşik bir şekilde yanıt verdiği insan korteksinden esinlenilerek özelliklerin bu tür hiyerarşik öğrenimi gerçekleştirilmiştir (Min, Lee, ve Yoon 2017). Şekil 3.15' de ESA mimarisi blok yapısı gösterilmektedir.



Şekil 3.15. ESA Mimarisi blok yapısı

Evrişimli bir katmanda, birçok filtre bulunmaktadır. Evrişimli katmanda, her biri görüntünün bir özellik haritasını Xk hesaplamak için kullanılan giriş görüntüsü, bu filtreler kullanılarak döndürülür, W = W1, W2, Wk. Buna bağlı olarak, evrişimli katmandaki özellik haritası sayısı ile filtre sayısı aynıdır. Her özellik haritası Denklem 15'deki formül kullanılarak hesaplanır. (1) formülde σ (·) doğrusal olmayan bir transfer fonksiyonudur ve b önyargı (Litjens ve ark. 2017):

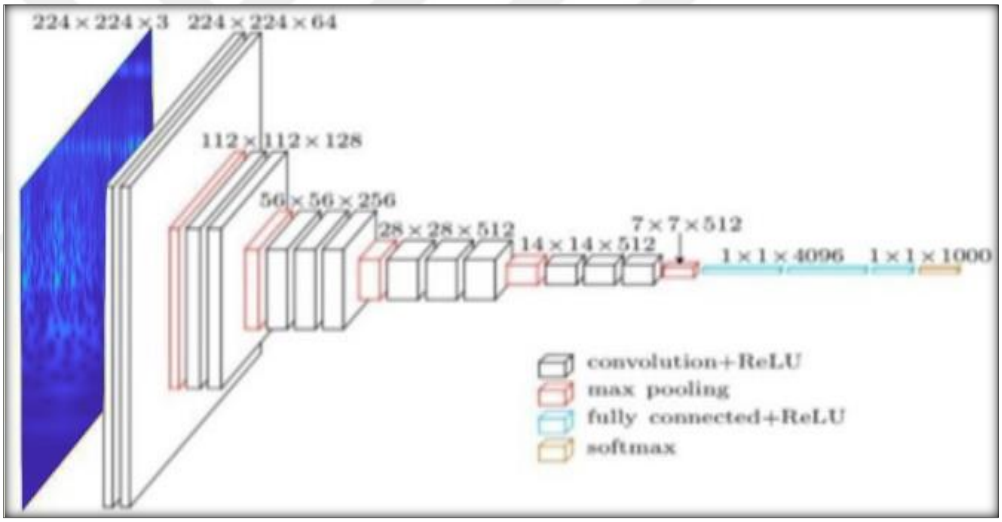
$$X_k^l = \sigma\left(W_k^{l-1} * X^{l-1} + b_k^{l-1}\right) \quad (15)$$

Özellik haritalarının seçilen havuzlama işlevi, maksimum, minimum veya ortalamadır. Özellik haritasındaki her piksel grubuna Seçim fonksiyonu uygulanır ve fonksiyonun sonucu (örneğin, o gruptaki maksimum değer), yeni alt örneklenmiş özellik haritasındaki grubu betimleyecek şekilde seçilir. Genel olarak, normal bir ESA derin öğrenme ağı, evrişimli katmanlar, havuz katmanları ve bir TBSA'dan oluşur (Aslan ve Akın 2020; Goodfellow, Bengio, ve Courville 2016).

3.2.15. Vgg-16 ESA Mimarisi

VGG-16 ESA mimarisi, 16 katmandan oluşan ve Oxford Üniversitesi Görsel Geometri Grubu tarafından ILSVRC-2014 müsabakası için tasarlanmış bir ESA derin ağ modelidir. Daha derin bir öğrenme mimarisine sahip olması, önceki derin öğrenme mimarilerinden en büyük farkıdır. VGG-16 ağı, 224x224x3'e (RGB) boyunda görüntüler ile beslenir. Her görüntü 3x3 filtre boyutuna sahip 5 adet evrişimli katmandan geçirilir. Girdi değerlerinin 2 kat alt örneklendiği maksimum havuzlama katmanıyla her blok sonlanır.

En son aşamada elde edilen özellik değerleri, sınıflandırma işlemi için bir TBSA'ya aktarılır. Şekil 3.16.'de, VGG-16 derin öğrenme ağ mimarisinin (Tindall, Luong, ve Saad 2015) genel yapısı gösterilmektedir. Bu çalışmada VGG-16 derin öğrenme mimarisi kullanılarak skalogram görüntüleri sınıflandırılmıştır.



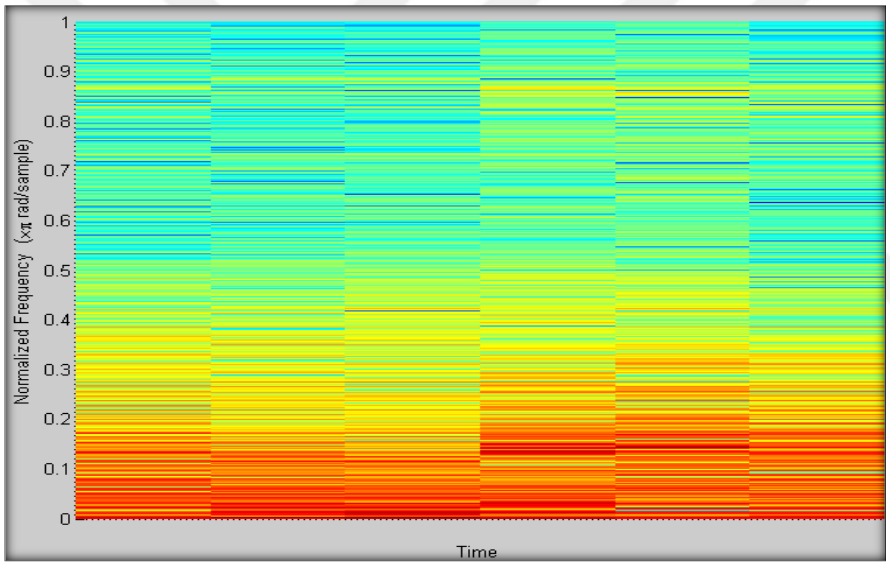
Şekil 3.16. VGG-16 derin öğrenme ağ mimarisi

3.2.16. Spektrogram Görüntülerinin Oluşturulması

Kısa Zamanlı Forier Dönüşümü (KZFD), zaman alanındaki bir sinyali frekans alanına dönüştüren genel amaçlı bir araçtır. KZFD dönüşümü, transfer fonksiyonunun bir pencere fonksiyonu ile çarpılmasıyla hesaplanır. Spektrogram, sinyalin bir zaman aralığı içinde frekans alanındaki görsel bir tasviridir (Yuan ve Cao 2017). Bu nedenle, sinyalin frekans bileşenlerinin zamanla nasıl değiştiğini gösterir. Bu çalışmada, tek bir görüntüde farklı zaman noktalarının frekans bileşenlerine sahip olabilmemiz için EEG sinyalinin kısa bölümleri (örneğin, 5 saniye uzunluğunda) bir spektrograma dönüştürülmüştür. Bu kısa EEG segmentlerinin spektrogram görüntülerini elde etmek için MATLAB yazılımını

kullandık. MATLAB'ın spektrogram fonksiyonunun varsayılan konfigürasyonunda, sinüzoidlerin toplamından oluşan bir sinyalin Nx = 1024 örneğini üretir. Sinüzoidlerin normalleştirilmiş frekansları 2π / 5 rad / örnek ve 4π / 5 rad / örnektir.

Şekil 3.17.'deki örnek spektrogramda, y ekseni 0 ile 1 arasında normalize edilen frekansı temsil ederken x ekseni zaman anlamına gelir. Kırmızıya yaklaşan renkler bu frekansta yüksek değerler gösterirken, maviye yakın renkler düşük yoğunluk değerlerini göstermek için kullanılır. Bu nedenle, örnek spektrogramda, sinyaldeki düşük frekans bileşenlerinin çoğu zaman diliminde yüksek frekans bileşenlerinden daha yoğun olduğu görülmektedir. Yüksek frekanslı bileşenler yalnızca birkaç zaman değerinde yüksek değerlerde ortaya çıkar ve bu nedenle örnek spektrogramda çoğunlukla farklı mavi renk tonları ile gösterilir (Aslan ve Akın 2020).



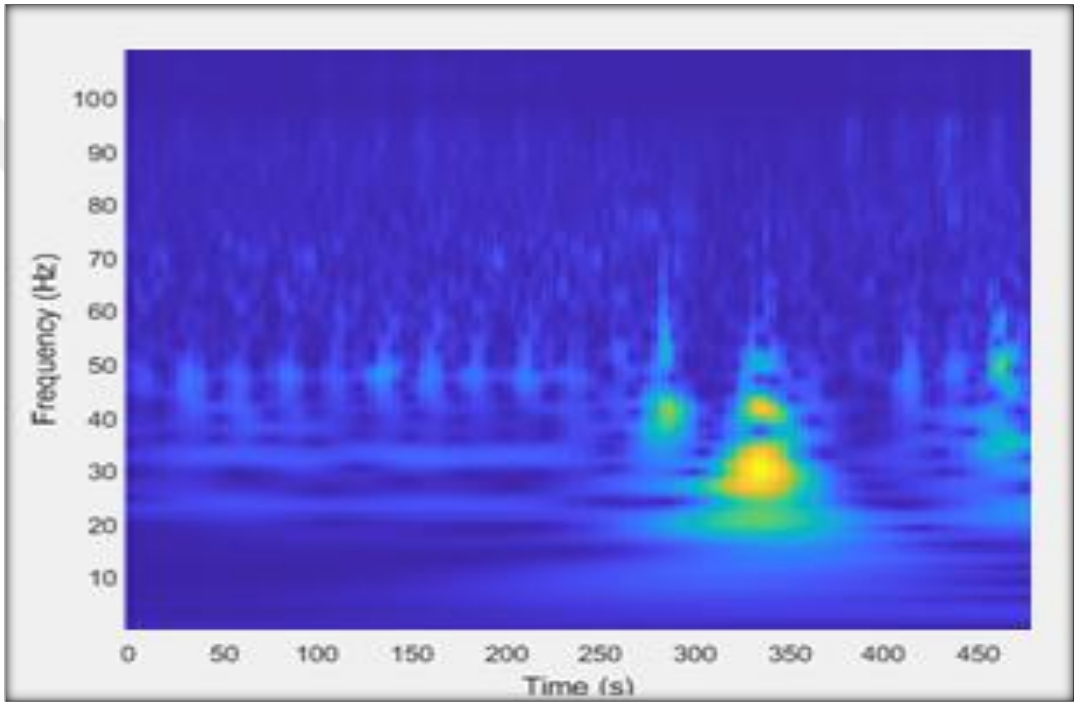
Şekil 3.17. Matlab yazılımı tarafından oluşturulan örnek bir spektrogram

3.2.17. Skalogram Görüntülerinin Oluşturulması

Sürekli Dalgacık Dönüşümü (SDD), bir dalgacık fonksiyonu yardımıyla, ölçeklerin sürekli olarak değişmesine izin vererek sinyali 2D dönüştüren bir zaman frekans analizi için kullanılan bir metottur. SDD, bir sinyalin zaman frekans bilgisinin 2D görüntüsünü oluşturmak için çok iyi frekans ve zaman lokalizasyonu sunmaktadır. SDD, değişken sinyallerin, değişimlerine ait özellikleri haritalamak için kullanılan oldukça yararlı bir yöntemdir. SDD bir sinyalin sabit veya değişken olduğunu belirlemek için kullanılan önemli bir zaman frekans analizi yöntemidir. Bir sinyalin değişken olması durumunda, sinyalin sabit bölümlerini belirlemek için kullanılabilir (Torrence ve Compo

1998).

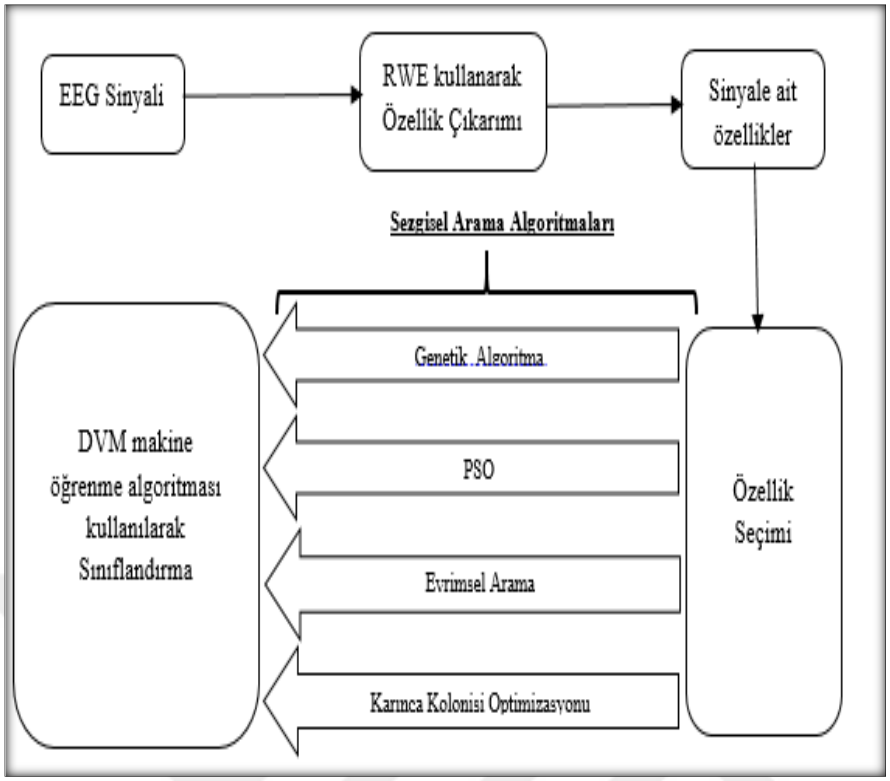
Bu çalışmada EEG sinyali verimizin farklı zaman noktalarının frekans bilgilerine erişebilmek için EEG sinyali 5sn'lik segmentlere ayrılmış ve tüm kanallardan alınan her segment birleştirilerek elde edilen vektör SDD yöntemi ile skalograma dönüştürülmüştür. Bu vektörler matlab yazılımı kullanılarak oluşturulmuştur. Skalogram görüntüleri oluşturulurken ana dalgacık olarak morlet dalgacık kullanılmıştır. Şekil 3.18.' de SDD yöntemiyle morlet dalgacığı kullanılarak elde ettiğimiz görüntülerden alınan örnek bir skalogram görüntüsü gösterilmektedir.



Şekil 3.18. Matlab yazılımı yardımıyla Morlet dalgacığı kullanılarak oluşturulan örnek bir skalogram görüntüsü

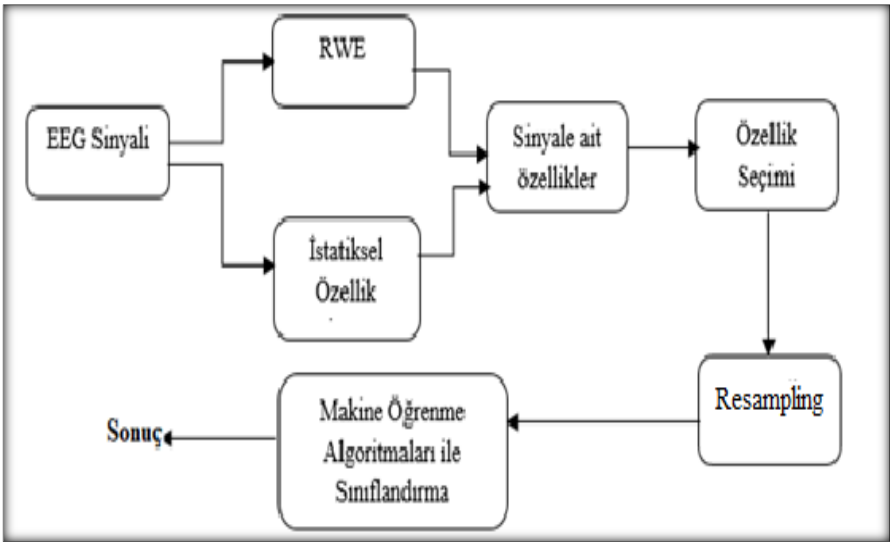
3.2.18. Tez Çalışmasında Geleneksel Makine Öğrenme Metotları Kullanılarak Önerilen BDT Yönteminin Genel Mimarisi

Tez çalışmasının ilk aşamasında, Veri Seti A' da bulunan SZ hastası ve SZ hastası olmayan kontrollere ait EEG verisi üzerinde GDE yöntemi ve istatistiksel özellik çıkarımı yapılmış ve elde edilen özellik değerleri farklı sezgisel arama algoritmalarıyla özellik seçimi yapılarak başarım performanları test edilmiştir. Şekil 3.19.'da bu çalışmaya ait akış diyagramı görülmektedir (Aslan ve Akın 2019a).



Şekil 3.19. Farklı sezgisel arama algoritmaları kullanılarak SZ tespiti

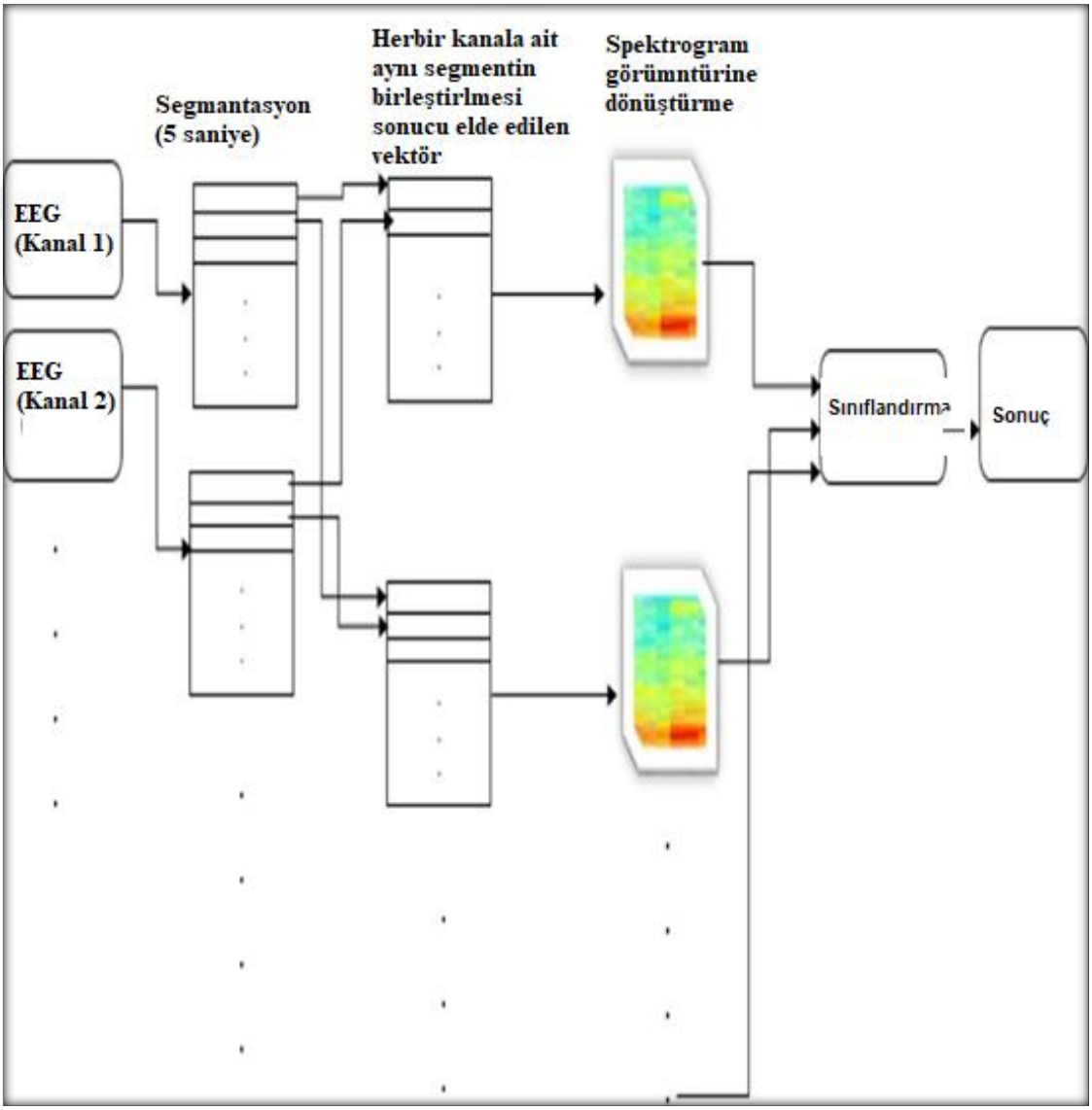
Bu çalışmanın sonucunda, sezgisel arama algoritmaları içerisnde en yüksek başarım PSO algoritması ile elde edilmiştir. Sonraki çalışmada ise SZ tespitinde farklı makine öğrenme algortimalarının performansı değerlendirilmiştir. Şekil 3.20.'de bu çalışmaya ait akış diyagramı görülmektedir (Aslan ve Akın 2019c).



Şekil 3.20. GDE ve istatistiksel özelliklerim farklı makine öğrenme algoritmaları ile performanlarının karşılaştırılması.

3.2.19. Tez Çalışmasında Derin Öğrenme Kullanılarak Önerilen BDT Yönteminin Genel Mimarisi

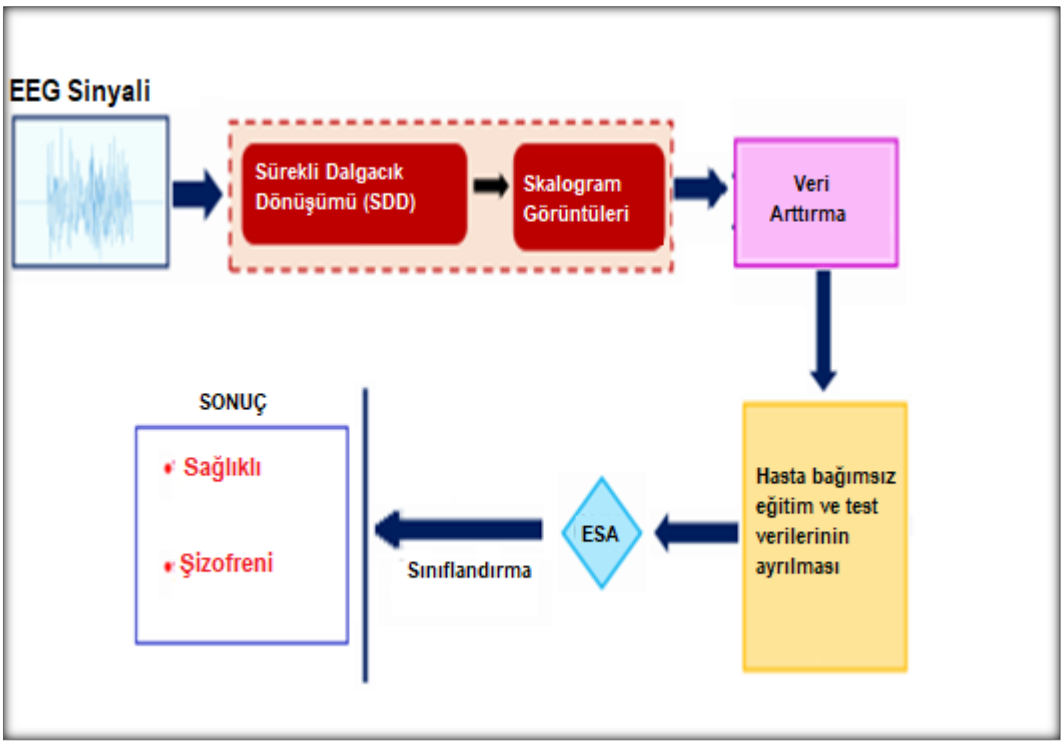
Tez çalışmasının ikinci aşasında ise, yöntemimiz herhangi bir manuel işlemeyi veya girişin geçici olarak değiştirilmesini içermez. Her girdi, yani bir dizi EEG kanalı verisi, bir dizi dönüşümden (segmentasyon, spektrogram üretimi ve ESA değerlendirmesi) geçirilir ve son olarak, SZ hastası veya sağlıklı bir kontrol olmak üzere bir sınıf değeri ile sonuçlanır. Şekil 3.21' de sürecin genel görünümünü göstermektedir.



Şekil 3.21. Spektrogram görüntülerinin derin öğrenmede kullanılarak SZ tespiti.

Diğer bir çalışmamızda, skalogram görüntüleri kullanarak derin öğrenme ile SZ tespitidir. Bu çalışmada, önerdiğimiz BDT yöntemi, her hangi bir önişlem ve özellik çıkarımı yapılması gibi aşamaları gerektirmemektedir. Şekil 3.22' de skalogram

görüntülerinin derin öğrenmede kullanılarak SZ tespiti çalışmamızda kullandığımız yöntemin akış şemasını içermektedir. ESA ağı tarafından eğitilecek olan her girdi verisi öncesinde sadece segmantasyon ve sonrasında skalogram görüntüleri aşamalarını içerir. ESA ağı tarafından verilen girdi verisi, sınıflandırılarak hasta veya sağlıklı kontrol olmak üzere bir sınıfa dâhil edilerek sonuçlandırılır.



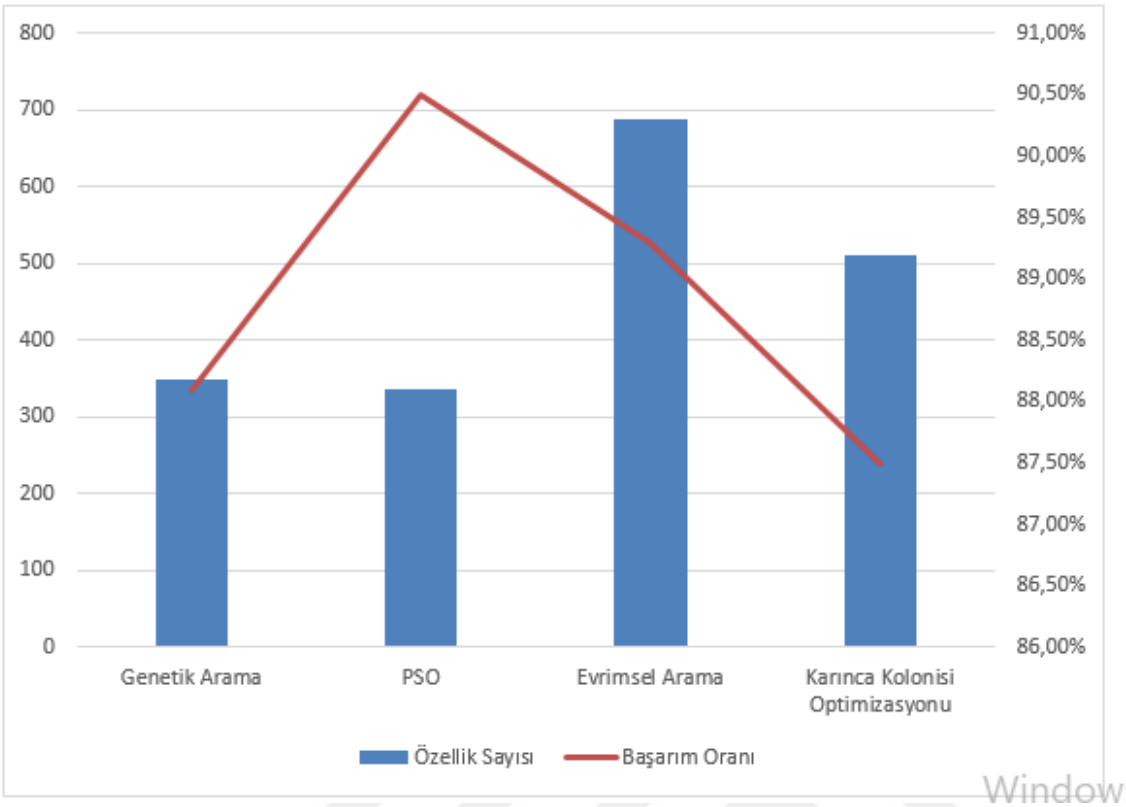
Şekil 3.22. Skalogram görüntülerinin derin öğrenmede kullanılarak SZ tespiti.

4. BULGULAR VE TARTIŞMA

Tez çalışmasının ilk aşamasında, toplam 84 bireye (45 SZ hastası - 39 sağlıklı kontrol) ait 16 elektrotlu EEG verisinden oluşan Veri seti A kullanılmıştır. Kontrollerden herbirinin EEG verisi 122280 uzunluklu bir vektör biçimindedir ve elektrotların herbirinden elde edilen veri ise 7680 bir vektördür. Veri boyutunun çok büyük olması sınıflandırma başarım performasını olumsuz olarak etkileyeceğinden, sınıflandırma başarımı arttırmak adına ham veri üzerinde özellik çıkarım işlemi yapılmıştır. Ham EEG siyaline ait veri üzerinde GDE yöntemi uygulanarak özellik çıkarım işlemi gerçekleştirilmiştir. Bu işlem sonucunda sağlanan veri setine, sınıflandırma performansını arttırmak adına özellik seçim işlemi uygulanarak sınıflandırmada etkili olacak daha önemli özellik değerlerinin ön plana çıkarılması sağlanmıştır. Bu maksatla, özellik seçim işlemi için literatürde yaygın olarak kullanılan sezgisel arama algoritmaları kullanılmıştır. Son aşamada oluşturulan özellik setleri DVM makine öğrenme sınıflandırıcısı ile sınıflandırılmış ve literatürde yaygın olarak kullanıılan sezgisel arama algoritmalarının sınıflandırma doğruluk performansları karşılaştırılmıştır. Kullanılan sezgisel arama algoritmalarına ait özellik verileri ve doğruluk performansları Çizelge 4.1'de ve Şekil 4.1.'de gösterilmektedir. Sınıflandırma işlemi sonucunda en iyi doğruluk değeri "PSO" özellik seçim algoritması ile 90,5% olarak bulunmuştur (Bakınız Çizelge 4.1.). Sınıflandırma işlemi Weka yazılımı ile gerçekleştirmiş ve default parametreler kullanılarak gerçekleştirilmiştir.

Çizelge 4.1. Sınıflandırma işlemi için sezgisel arama algoritmalarının kullanılarak gerçekleştirilen özellik seçimiyle elde edilen özellik skoru ve doğruluk performans değerleri.

Sezgisel Arama Algoritması	Özellik Sayısı	Başarım Oranı
Genetik Arama	350	88,1%
PSO	336	90,5%
Evrimsel Arama	689	89,3%
Karınca Kolonisi Optimizasyonu	510	85,7%



Şekil 4.1. Sınıflandırma işlemi için sezgisel arama algoritmalarının kullanılarak gerçekleştirilen özellik seçimiyle elde edilen özellik skoru ve doğruluk performans değerleri.

Daha sonraki çalışmada, farklı geleneksel makine öğrenme algoritmalarının performansları değerlendirimiştir. Yine veri seti A için GDE metodu ve standart istatiktiksel hesaplamalar ile ham EEG sinyalinden özellik çıkarım işlemi yapılmıştır. Bu işlemler sonucunda edinilen özellik boyutunun çok yüksek olmasına ve aksine veri boyutunun (kontrol sayısı) düşük olması sonucunda sınıflandırma işleminin başarım performasının doğru bir şekilde değerlendirilemeyeceği farkedilmiştir. Bu sebepten dolayı, var olan özellik değerlerinin daha uygun bir sayıya indirgenebilmesi amacıyla "En iyi ilk arama" sezgisel arama algoritması uygulanarak bu özellik seti üzerinden özellik seçimi gerçekleştirilmiştir. Özellik seçim işleminden sonra özellik boyutunun daha uygun seviyelerde olduğu farkedilmiş, elde edilen bu özellik değerleri makine öğrenme algoritmalarıyla sınıflandırıldığında tatmin edici sınıflandırma başarım oranları elde edildiği gözlemlenmiştir. En son aşamada edinilen özellik değerleri, Naive Bayes, DVM, KNN, ÇKA ve RO sınıflandırma algoritmaları kullanılarak sınıflandırılmıştır. Her bir kanala ait ve tüm kanallara ait başarım oranları Çizelge 4.2.'de gösterilmektedir.

Çizelge 4.2. Her bir elektrottan ve bütün elektrotlardan alınan veriler için farklı sınıflandırma yöntemleriyle elde edilen doğruluk değerleri.

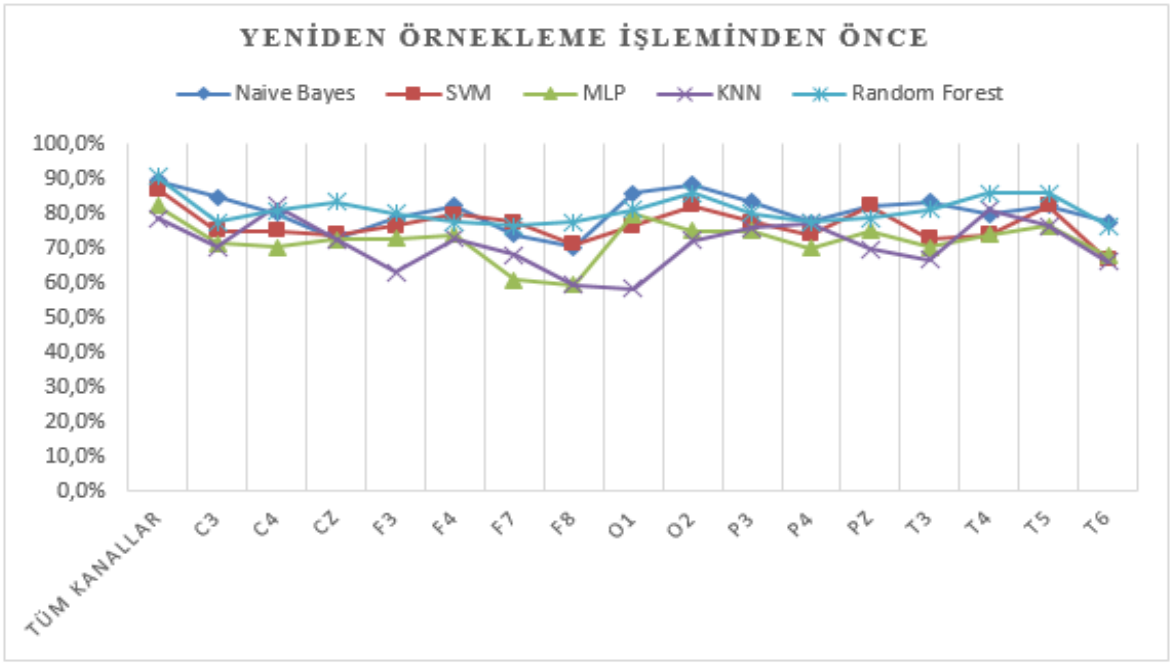
	Tüm kanallar	C3	C4	CZ	F3	F4	F7	F8	O1	O2	P3	P4	PZ	T3	T4	T5	T6
Naive Bayes	89,3 %	84,5 %	79,8 %	72,6 %	78,6 %	82,2 %	73,8 %	70,2 %	85,6 %	88,1 %	83,3 %	77,4 %	82,2 %	83,3 %	79,7 %	82,1 %	77,4 %
DVM	86,9 %	75,0 %	75,0 %	73,8 %	76,2 %	79,7 %	77,4 %	71,1 %	76,2 %	82,1 %	77,4 %	73,8 %	82,2 %	72,7 %	73,8 %	82,2 %	66,6 %
ÇKA	82,1 %	71,4 %	70,2 %	72,5 %	72,7 %	73,8 %	60,7 %	59,5 %	79,7 %	75,0 %	75,0 %	70,0 %	75,0 %	70,3 %	73,8 %	76,2 %	67,7 %
KNN	78,4 %	70,3 %	82,0 %	72,5 %	63,1 %	72,6 %	67,9 %	59,5 %	58,2 %	72,1 %	75,6 %	77,1 %	69,7 %	66,7 %	81,0 %	76,2 %	66,1 %
Rastgele Orman	90,5 %	77,4 %	80,8 %	83,3 %	79,8 %	77,4 %	76,2 %	77,4 %	81,0 %	85,7 %	79,8 %	77,4 %	78,6 %	81,0 %	85,7 %	85,7 %	76,2 %
Özellik Sayısı	75	21	15	23	20	17	11	13	49	49	31	30	29	24	25	19	17

Çizelge 4.2. incelendiğinde Rastgele Orman sınıflandırma algoritması ile en yüksek başarım oranı yaklaşık 90% olarak elde edilmiştir. Burada görüldüğü üzere tüm kanallardan edinilen özellikler kullanılarak gerçekleştirilen sınıflandırmanın başarım performansının daha yüksek olduğu görülmektedir. Veri sayısının (birey sayısı) az olmasından dolayı var olan özelliklere yeniden örnekleme (Resampling) işlemi uygulanarak sınıflandırma başarım performansının arttırılması amaçlanmıştır. Yeniden örnekleme işlemi sonucu elde edilen başarım oranları Çizelge 4.2.'de gösterilmektedir. Çizelge 4.2. incelendiğinde edinilen sınıflandırma başarım performansında belirgin bir fark olduğu gözlemlenmiştir. Farklı sınıflandırıcılar kullanılarak yapılan sınıflandırma işleminde, en iyi sınıflandırma doğruluk performası DVM ile 95,2% olarak elde edildiği gözlemlenmiştir. Sınıflandırma işlemleri Weka yazılımı ile gerçekleştirmiş ve default parametreler kullanılarak gerçekleştirilmiştir.

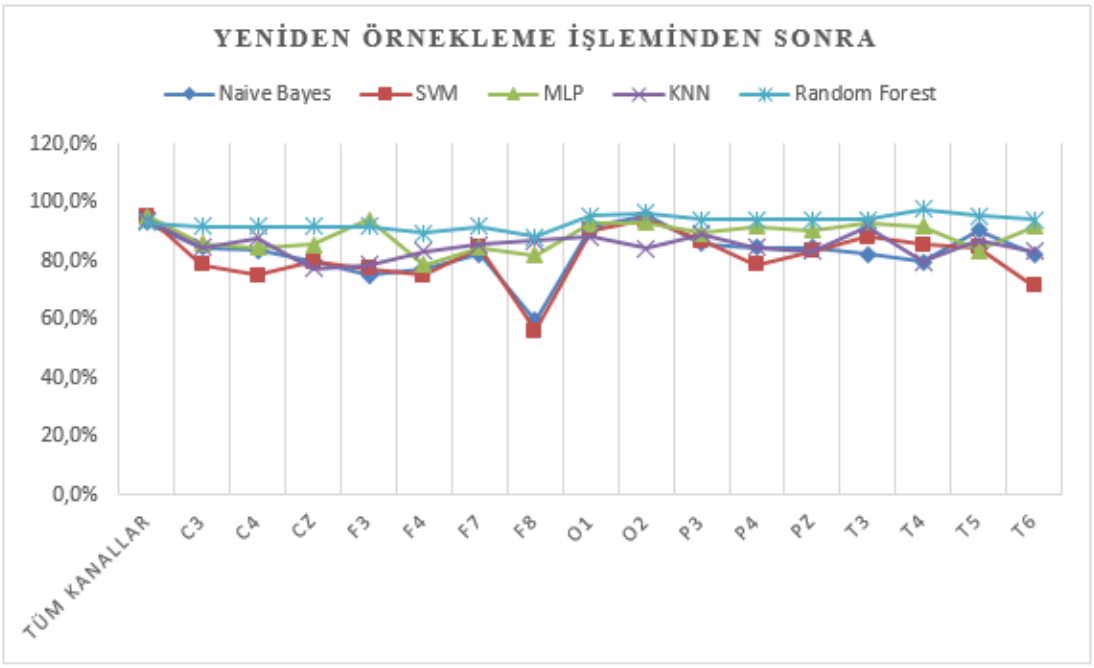
Çizelge 4.3. Yeniden örnekleme işlemi sonucu elde edilen başarım sonuçları.

	Tüm kanal	C3	C4	CZ	F3	F4	F7	F8	O1	O2	P3	P4	PZ	T3	T4	T5	T6
Naive Bayes	92,9 %	84,5 %	83,4 %	79,7 %	75,0 %	77,3 %	82,2 %	59,6 %	91,7 %	95,2 %	85,7 %	84,5 %	84,5 %	82,2 %	79,7 %	90,5 %	82,2 %
DVM	95,2 %	78,6 %	75,0 %	79,7 %	77,2 %	75,0 %	84,5 %	55,9 %	90,4 %	94,1 %	86,9 %	78,6 %	83,2 %	88,1 %	85,6 %	84,5 %	71,4 %
ÇKA	95,2 %	85,7 %	84,5 %	85,6 %	94,1 %	78,6 %	84,5 %	82,0 %	92,8 %	92,9 %	89,3 %	91,7 %	90,5 %	92,9 %	91,7 %	83,4 %	91,7 %
KNN	94,1 %	84,5 %	87,8 %	77,3 %	78,5 %	83,3 %	85,6 %	86,9 %	88,0 %	84,3 %	89,2 %	84,4 %	83,2 %	91,7 %	79,7 %	86,8 %	83,2 %
Rastgele Orman	92,9 %	91,7 %	91,7 %	91,7 %	91,7 %	89,3 %	91,6 %	88,1 %	95,2 %	96,4 %	94,1 %	94,0 %	94,0 %	94,0 %	97,6 %	95,2 %	94,0 %
Özellik sayısı	75	21	15	23	20	17	11	13	49	49	31	30	29	24	25	19	17

Literatürde yaygın olarak kullanılan beş farklı makine öğrenmesine ait başarım performansları yeniden örnekleme işleminden önce Şekil 4.2.'de, yeniden örnekleme işlemi sonrası ise Şekil4.3.'de gösterilmektedir.



Şekil 4.2. Yeniden örnekleme gerçekleştirilmeden önce çeşitli sınıflandırma algoritmaları ie elde edilen doğruluk performası grafiği.



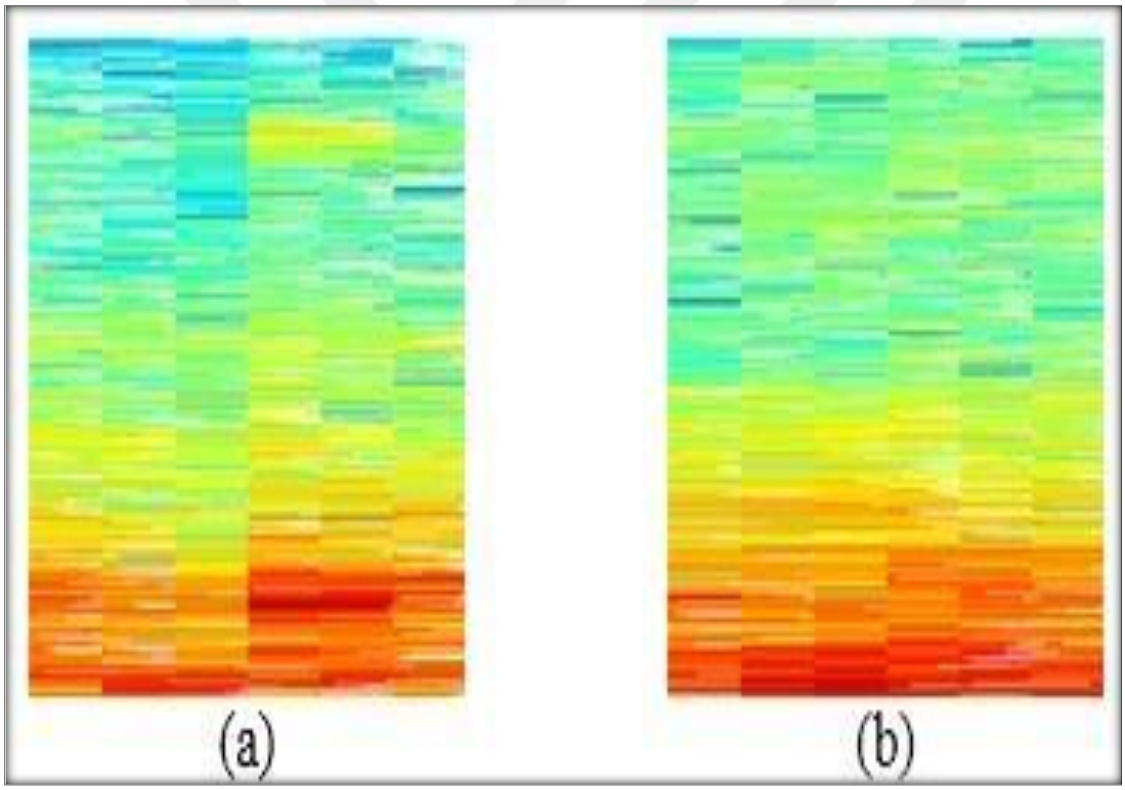
Şekil 4.3. Yeniden örnekleme gerçekleştirildikten sonra çeşitli sınıflandırma algoritmaları ile elde edilen doğruluk performası grafiği.

Geleneksel makine öğrenme algoritmalarının ön işlem, özellik çıkarımı gibi işlemler gerektirmesinden dolayı tezin ikinci aşamasında EEG sinyallerinin zaman frekans domenine dönüştürülerek herhangi bir ön işleme tabi tutulmadan doğrudan derin öğrenme algoritmalarıyla SZ tespitindeki başarımları değerlendirilmiştir.

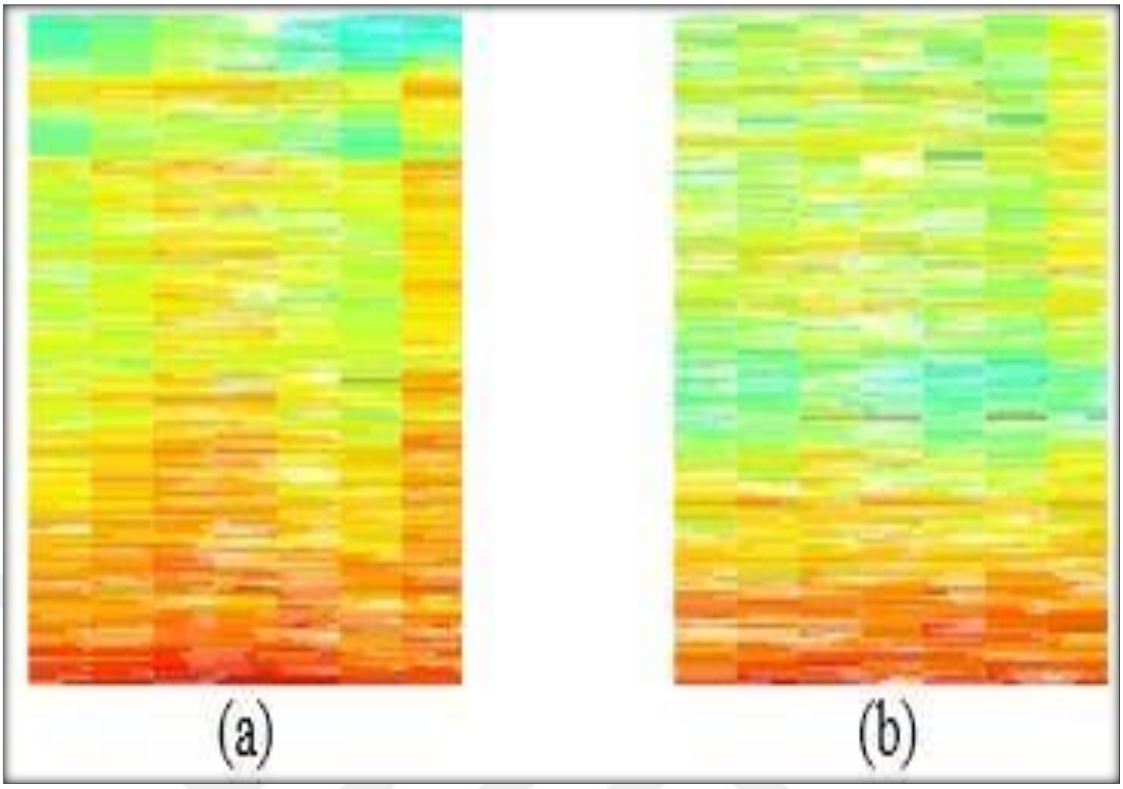
Bu aşamada, SZ hastalarını EEG kayıtları kullanarak otomatik olarak teşhis etmeyi amaçlayan bir BDT yöntemı sunulmaktadır. Birçok literatür çalışmasından farklı olarak, önerilen yöntem EEG kayıtlarından özellikleri manuel olarak ayıklamaz, bunun yerine frekans-zaman özelliklerinin yararlı bir temsiline sahip olmak için Kısa Süreli Fourier Dönüşümü (STFT) kullanarak ham EEG'yi 2D'ye dönüştürür. Bu çalışma, SZ hastalarının otomatik teşhisi için 2D zaman-frekans özelliklerinin kullanıldığı ilgili literatürdeki ilk çalışmadır. 2D alanda mevcut görüntülerden en kullanışlı özelliklerini çıkarmak ve örnekleri yüksek doğrulukla sınıflandırmak için son teknoloji ürünü bir Evrişimli Sinir Ağı mimarisi olan VGG-16 eğitilmiştir. Deneysel sonuçlar, sunulan yöntemin, farklı yaş gruplarından oluşan iki veri kümesinde % 95 ve % 97 sınıflandırma doğruluğu ile SZ hastalarını ve sağlıklı kontrolleri sınıflandırma görevinde başarılı olduğunu göstermektedir. Bu performansla, önerilen yöntem literatür yöntemlerinin çoğundan daha iyi performans gösterir. Çalışmanın deneyleri ayrıca bir EEG kaydının

frekans bileşenleri ile SZ hastalığı arasında bir ilişki olduğunu ortaya koymaktadır. Ayrıca, makalede sunulan Grad-CAM görüntüleri, orta düzey frekans bileşenleri, bir SZ hastasını sağlıklı bir kontrolden ayırırken daha önemlidir.

Bu aşamada, önerilen yöntem örneklerin iki veri kümesine karşı değerlendirilmiştir. İlk veri setinde (A), 39 sağlıklı çocuk ve 45 SZ hastalığı olan çocuğun 16 kanallı EEG kayıtları vardır. Her örnek, her biri 10240 uzunluğunda bir vektör ile temsil edilen 5 saniyelik uzunluklu bölümlere bölünmüştür (saniyede 128 değer x 16 kanal x 5 saniye). Bu vektörler daha sonra 224x224 boyutunda spektrogramlara dönüştürülür. Bu nedenle, Veri Kümesi A'da 84 kişi için 1008 görüntümüz bulunmaktadır. Şekil 4.4. ve Şekil 4.5. sırasıyla Veri Kümesi A ve B için sağlıklı bir kontrol ve bir SZ hastası için örnek spektrogram görüntülerini göstermektedir.



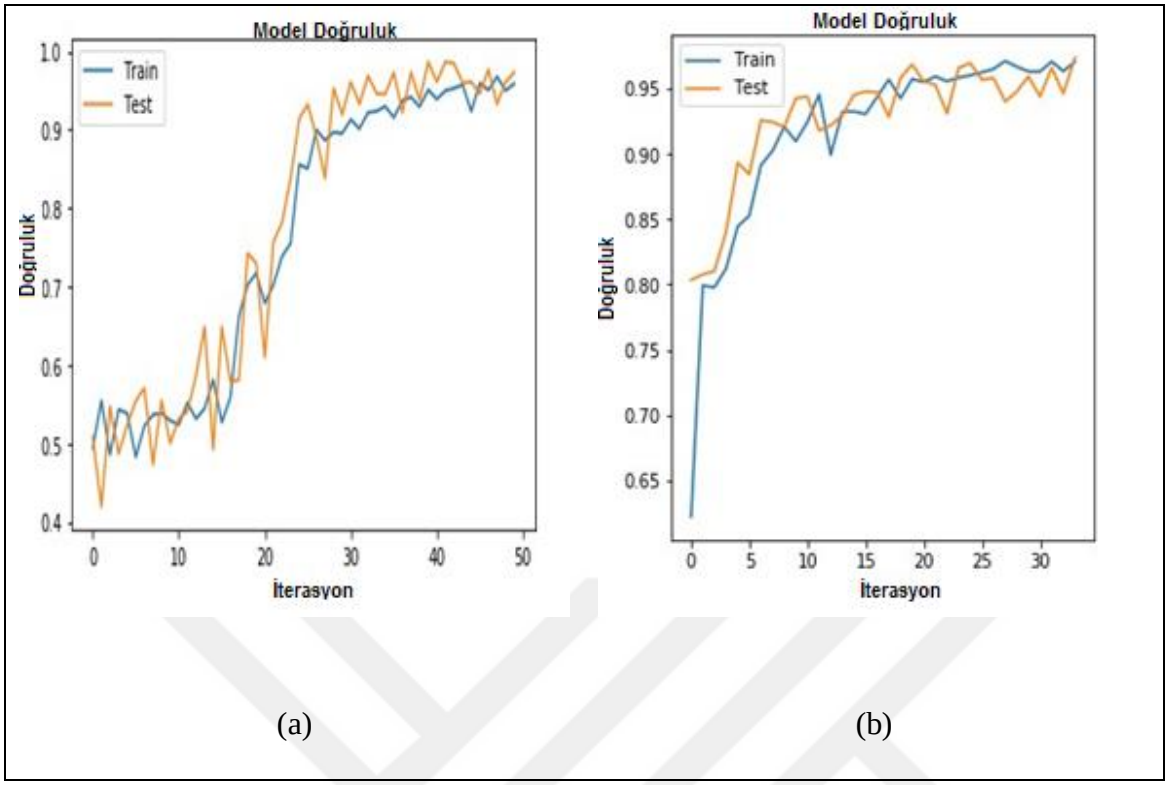
Şekil 4.4. Veri Kümesi A'dan örnek spektrogram görüntüleri (a) sağlıklı kontrol ve (b) SZ hastası



Şekil 4.5. Veri Kümesi B 'den örnek spektrogram görüntüleri (a) sağlıklı kontrol ve (b) SZ hastası

Spektrogram görüntülerinin veri seti sırasıyla % 80 ve % 20 oranında eğitim ve test setlerine ayrılmıştır. Bu görüntüler, her birini sağlıklı veya SZ hasta olarak sınıflandırmak için VGG-16 ESA modeline beslenir. Ağ için hiper parametreler şu şekilde alınır: girdi görüntü boyutu 112x112, batch size 128, 1.0e-4 öğrenme oranı ve optimize edici Adam (Kingma ve Ba 2014). Deneyler sayesinde, ağın yakınsamasına ulaşmak için 50 iterasyon eğitimin yeterli olduğunu gözlemledik. Ortalama olarak, ağ test seti için% 95 doğruluk değeri elde edilmiştir.

İkinci veri setinde 28 yetişkinin (14 kontrol ve 14 SZ hastası) 19 kanallı EEG kaydı vardır. Kayıtların uzunluğu 12 ila 15 dakika arasında değişmektedir. Bu nedenle, sağlıklı ve SZ gruplarında, kaydın uzunluğu gruptaki en kısa kaydın uzunluğuna göre ayarlanmıştır. Daha sonra her sağlıklı kontrol için 173 segment ve her SZ hastası için 148 segment elde edildi ve her segmentin uzunluğu yine 5 saniye alındı. Bu noktadan sonra, Veri Kümesi A'dakilere benzer şekilde aynı ayarlar ve hiper parametrelere sahip toplam 4494 spektrogram görüntüsü elde edilmiştir. Sonuç olarak, ağ 30 iterasyon değerinde % 97.4 doğruluğa ulaştı. Şekil 4.6. iterasyon eğitim süresine göre doğruluktaki değişikliği göstermektedir.



Şekil 4.6. (a) Veri Kümesi A ve (b) Veri Kümesi B için iterasyon ve eğitim süresine karşı doğruluk değerleri

Şekil 4.6.'da görülebileceği gibi, yeterli miktarda eğitim önemlidir ve modelin doğruluğunu önemli ölçüde etkilemektedir. Maalesef, literatürde gerekli eğitim miktarı için önceden tanımlanmış bir sayı yoktur ve genel olarak deneyler yoluyla her veri seti için deneysel olarak belirlenir.

İdeal bir sınıflandırıcı, hastalıklı kontrolleri yüksek oranda tespit ederken, tüm sağlıklı kontrolleri hastalıksız olarak dışlamalıdır. Bu nedenle, sınıflandırıcının hem Hassasiyet hem de Geri Çağırma ölçümleri aynı anda yüksek olmalıdır ve bu da sonuçta yüksek F1 skoru ile sonuçlanır. Çizelge 4.4. ve Çizelge 4.5.'de gösterilen deneylerin sonuçları, önerilen yöntemimizin, Veri Kümesi A ve B için sırasıyla% 95 ve% 97 F1-skor değerleri ile sınıflandırma görevinin bu yönlerinde iyi performans gösterdiğini göstermektedir. Çizelge 4.4. ve Çizelge 4.5. 'deki destek değerinin, her satır için gerçek örnek sayısı anlamına geldiğine dikkat edilmelidir.

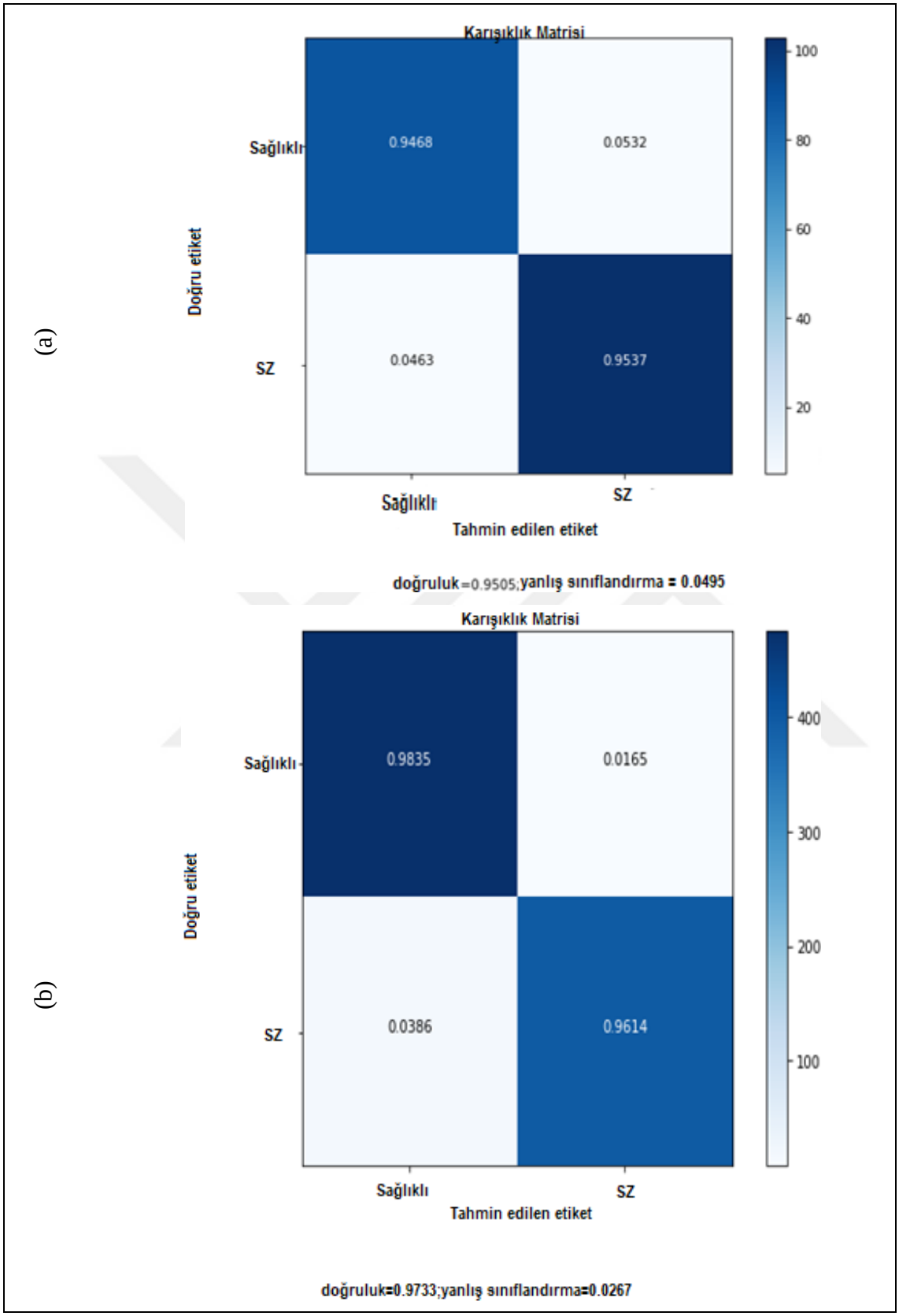
Çizelge 4.4. Spektrogram görüntüleri ile eğitilmiş ağın, Veri Kümesi A'ya göre performans sonuçları

Veri Seti A	Precision	Recall	F1-score	Support
Sağlıklı Birey	0.95	0.95	0.95	94
SZ Hastası	0.95	0.95	0.95	108
Doğruluk			0.95	202
Makro Ortalama	0.95	0.95	0.95	202
Ağırlıklı Ortalama	0.95	0.95	0.95	202

Çizelge 4.5. Spektrogram görüntüleri ile eğitilmiş ağın, Veri Kümesi B'ye göre performans sonuçları

Veri Seti B	Precision	Recall	F1-score	Support
Sağlıklı Birey	0.97	0.98	0.98	484
SZ Hastası	0.98	0.96	0.97	414
Doğruluk			0.97	898
Makro Ortalama	0.97	0.97	0.97	898
Ağırlıklı Ortalama	0.97	0.97	0.97	898

Şekil 4.7.'de verilen karışıklık matrisleri, hastalıklı ve hastalıksız numuneler için doğru sınıflandırma oranının yüksek (> = 0.94) olduğunu ve her iki veri kümesi için 0'a yakın değerlerle (<= 0.05) yanlış sınıflandırmanın çok düşük olduğunu açıkça göstermektedir.

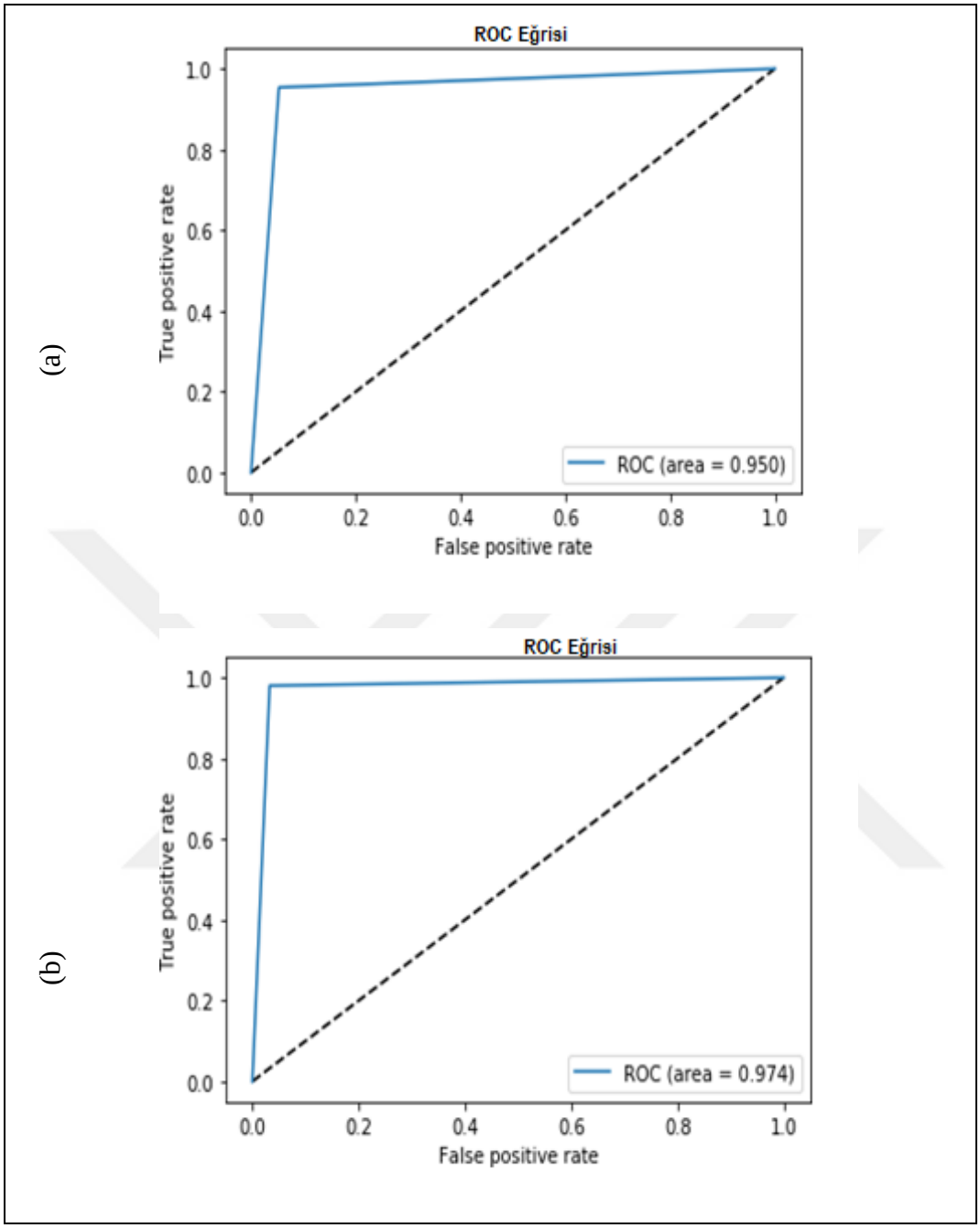


Şekil 4.7. (a) Veri Kümesi A ve (b) Veri Kümesi B için önerilen yöntemin karışıklık matrisleri

Belirli bir konfigürasyon için elde edilen yüksek TPR ve düşük FPR, ikili sınıflandırıcının tüm eşik değerlerinde iyi olduğu anlamına gelmez ve bu nedenle "0" lar ve "1" ler için üretilen yakın değerler nedeniyle düşük performanstan zarar görebilir. Ancak, AUC değerleri Veri Kümesi A ve B için sırasıyla 0.95 ve 0.974 olarak elde edildiğinden önerilen yöntem için durum böyle değildir (bkz. Şekil 4.8). Yani, "0" ve "1" leri temsil etmek için üretilen değerler "0" ve "1" e çok yakındır ve bu nedenle eşik değerlerinin çoğu ikisini ayırt etmek için yeterlidir. Bu, önerilen yöntemi çoğu durumda sağlam bir sınıflandırıcı yapar.

Veri Kümesi A daha önce Phang ve diğerleri tarafından değerlendirilmiştir. (Phang, Ting, ve ark. 2019), % 93.06 sınıflandırma doğruluğuna ulaştığını bildirmiştir. Yöntemleri, girdi olarak ham EEG sinyallerini kullanan bir ESA modeliydi. Yöntem, her biri verilen sinyalleri beyin bağlantı özellikleri gibi farklı özellik kümelerini kullanarak sınıflandıran bir 1D ve 2D ESA'lar topluluğu kullandı. İkinci veri seti de Oh ve diğerleri tarafından önerilen bir ESA modeline karşı değerlendirildi. (Oh ve ark. 2019), yöntemlerinde, ham EEG sinyalleri bir ESA'ya ön işlem yapılmadan beslenir, sınıf değerini % 98.07 doğrulukla doğrudan tahmin eder.

Bu aşamada, ilgili bu önceki literatür yöntemleriyle karşılaştırıldığında, yöntemimiz bunların çoğundan daha iyi performans gösterir ve bir çalışma ile karşılaştırılabilir şekilde performans gösterir (Oh ve ark. 2019). Halka açık SZ hasta verilerinin bolluğu gibi nedenlerden dolayı, çoğu yöntem performansı tek bir veri setiyle ölçerken, yöntemimiz iki ayrı veri setiyle (çocuklar ve yetişkin) değerlendirilmiştir. Bu, önerilen yöntemin farklı durumlar için sağlamlığını açıkça kanıtlamaktadır. İkinci olarak, ham EEG sinyallerini kullanan yöntemlerle ilgili olarak iyileştirilmiş performansımız, EEG'lerin, sinyalin frekans bilgisini açıkça yorumlanabilen spektrogram görüntülerine dönüştürülmesinin kullanımıyla bağlantılı olabilir.



Şekil 4.8. (a) Veri Kümesi A ve (b) Veri Kümesi B için ROC eğrisi

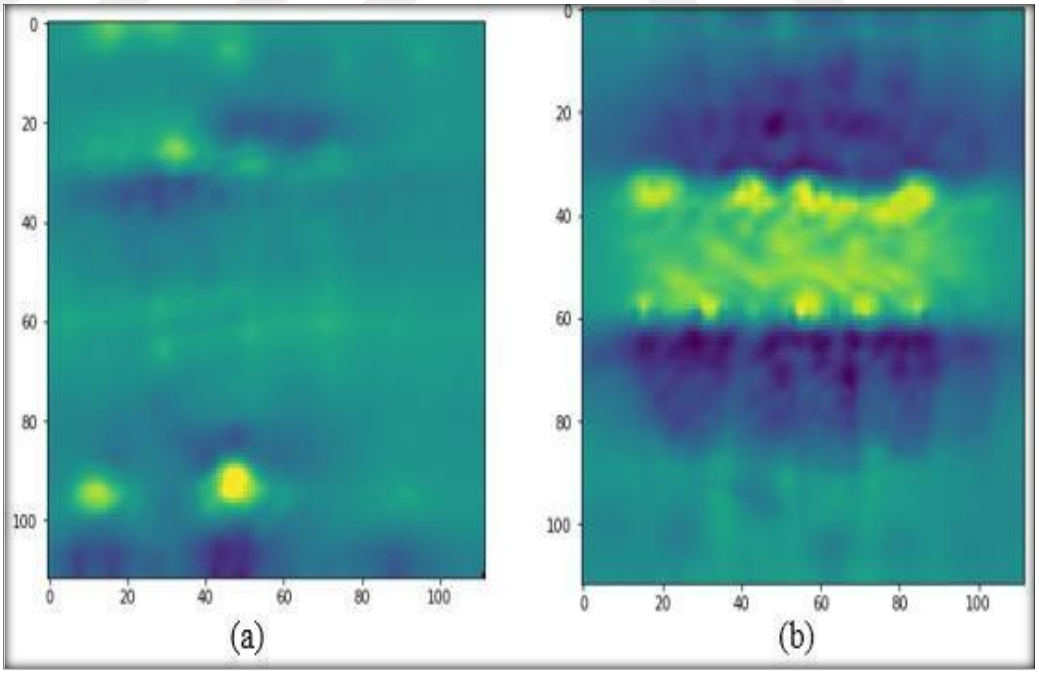
Ayrıca, ESA'lar, pikseller arasında uzamsal ilişkilere sahip olan ve bir nesneyi sınıflandırırken çoğunlukla bu uzamsal özelliklere dayanan görüntüleri sınıflandırmak için tasarlanmıştır. Ham EEG sinyalinde, yalnızca zamansal ilişkiler açıkça mevcuttur. Spectrogram, frekans ve zaman bilgisini uzamsal ilişkilerle tutan bir görüntü olan bir ESA modeli için daha uygun bir girdidir.

Dahası, bu çalışmada önerilen yöntem daha fazla insan tarafından yorumlanabilir

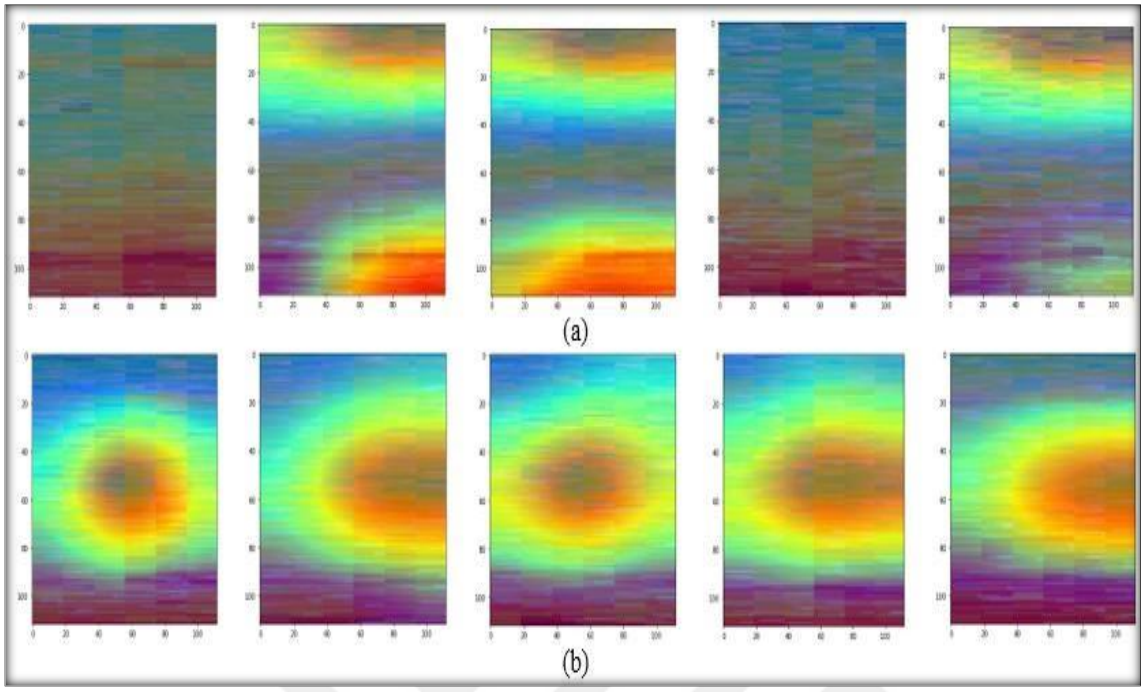
çıktı üretme yeteneğine sahiptir. Bu bağlamda Şekil 4.4. ve Şekil 4.5.'de gösterilen sağlıklı bir kontrol ve SZ hastası için spektrogram görüntüleri incelendiğinde, bu spektrogramların farklı zamanlarda farklı frekans değerleri gibi farklılıklar içerdiği görülebilir.

Bununla birlikte, bir insan için bu farklılıkları bir veri kümesindeki tüm örnekler için genellemek ve nicel olarak formüle etmek zordur. Bu bağlamda, Aktivasyon Maksimizasyonu (AC) (Kotikalapudi ve ark. 2017) adlı bir teknik kullanarak bu farklılıkların neler olduğunu ortaya çıkarmaya çalışıyoruz.

Bir AC görüntüsü, tek bir sınıf için ağın çıktısını en üst düzeye çıkaran değerleri yinelemeli olarak bularak sentezlenen yapay bir görüntüdür. Bu nedenle, bir sınıfı temsil eden ideal bir girdi olarak düşünülebilir. Sağlıklı bir kontrol ve bir SZ hastası için Şekil 4.9' da sunulan AC görüntülerinde açık farklılıklar olsa da, bu görüntüler yalnızca ağın farklı sınıflardan spektrogramlara çok çeşitli bir şekilde yanıt verdiğini göstermektedir. Bunun nedeni, AC görüntülerinde görülen özelliklerin, filtre aktivasyonunu en üst düzeye çıkarmak için yapay olarak üretilmesidir ve bu nedenle, gerçek spektrogram görüntülerini temsil etmemektedir.



Şekil 4.9. (a) sağlıklı kontrol ve (b) SZ hastası için aktivasyon maksimizasyon görüntüleri



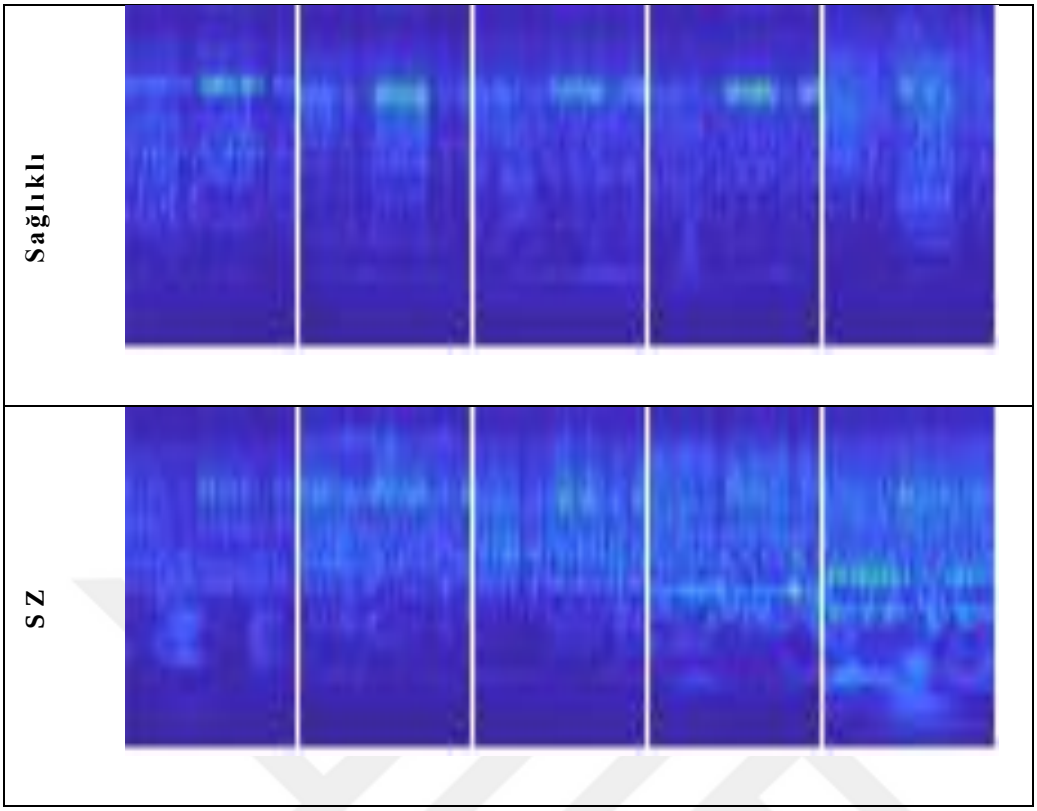
Şekil 4.10. (a) sağlıklı kontroller ve (b) SZ hastalarından oluşan bir grup bireyden elde edilen Grad-CAM görüntüleri

Bu nedenle, sağlıklı bir kontrol ve bir SZ hastası için gerçekten neyin farklı olduğunu anlamak için ikinci bir girişim olarak Gradyan Ağırlıklı Aktivasyon Haritaları (Grad-CAM) (Selvaraju ve ark. 2017) adlı başka bir teknik kullandık. Bir Grad-CAM görüntüsünde, renklerin tahmin edilen sınıfa göre görüntüdeki bölgenin alaka düzeyini gösterdiği bir giriş görüntüsünde bir renk haritası gösterilir. Bu nedenle, görüntüdeki kırmızıya yakın renkler, o bölgenin tahmin edilen sınıf için önemli uzamsal özelliklere sahip olduğu anlamına gelir. Şekil 4.10' da sınıfa göre gruplandırılmış farklı bireylerin Grad-CAM görüntülerini göstermektedir. Hastalıklı gruba ait olan Grad-CAM görüntüleri, genellikle renklerin görüntünün ortasında ortalanmış olduğu bir deseni açıkça ortaya koyarken, sağlıklı kontrol görüntülerinde yüksek yoğunluklu renkler ya hiçbir zaman görünmez ya da görüntünün sadece üstünde ve altında görünür.

Şekil 4.10.'da gösterilen modelin şekilde gösterilmeyen diğer kişiler için ortak olduğu unutulmamalıdır. Bu nedenle, bu görüntülerden birkaç sonuç çıkarılabilir. Her şeyden önce, frekans bileşenleri SZ hastaları ve sağlıklı kontrollerin ayırt edilmesinde önemlidir. Ayrıca, orta düzey frekans bileşenleri, SZ hastalarını ayırt etmek için en önemli olanlardır çünkü bu gruptaki hemen hemen tüm örnekler, orta düzey frekans bileşenlerini temsil eden spektrogram bölgesinde benzer bir modele sahiptir.

Wavelet yönteminin, sinyalin zaman frekans bileşenlerine ait önemli özellikleri daha detaylı bir şekilde çıkarttığı bilinmektedir. Bundan dolayı, geleneksel makine öğrenme algoritmalarıyla birlikte bu yöntemin kullanıldığı birçok çalışma bulunmaktadır. Literatürde birçok farklı biyomedikal sinyalin analizinde oldukça yoğun olarak kullanılmakta ve tatmin edici sonuçlar üretmektedir. Biz bu doğrultuda tezin son aşamasında, SZ'yi çok daha yüksek sınıflandırma başarımı, basit işlem süreci ve aynı zamanda uzman görüşüne destek olarak yorumlanabilir sonuçlarla tespit etmeye çalışan bir yöntem öneriyoruz. Elde edilen çıktılar, hastalığın tespitinde uzmana yardımcı olabilecek yorumlanabilir görüntülerden oluşmaktadır. Önerilen çalışma, bir EEG'nin zaman frekans özelliklerini değerlendirilmesi açısından da önem arz etmektedir. SZ'yi tespit etmek için EEG kayıtları dışında herhangi bir işleme veya materyale ihtiyaç duyulmamaktadır. Önceki çalışmaya benzer şekilde spectrogram görüntüleri yerine, EEG kayıları, SDD yöntemi ile morlet dalgacığı kullanılarak 2D zaman frekans domenine dönüştürülerek elde edilen skalogram görüntüleri kullanılmıştır.

Veri seti A'da bulunan EEG kayıtlarından, her kanal 5sn'lik uzunluklu segmentlere bölünmüştür. Her kanala denk gelen segmentler birleştirilerek tek boyutlu ve her biri 10240 uzunluğunda bir vektör elde edilmiştir. Her kanaldan alınan veri 5 saniyelik uzunluklu bölümlere bölündüğünde örnekleme frekansı 128Hz olduğundan elde edilecek vektör (5 saniye x 128 örnekleme frekansı x 16 kanal) şeklinde hesaplanacaktır. Elde edilen vektörler, 224x224 boyutunda SDD yöntemi ile morlet dalgacığı kullanılarak skalogram görüntülerine dönüştürülür. Bu veri kümesinde, yukarıda bahsedilen segmantasyon işlemi sonucunda 1008 skalogram görüntüsü elde edilmiştir. Şekil 4.11.'de veri seti A'dan elde edilen farklı kişilere ait farklı segmentlerden alınan örnek skalogram görüntüleri gösterilmektedir.

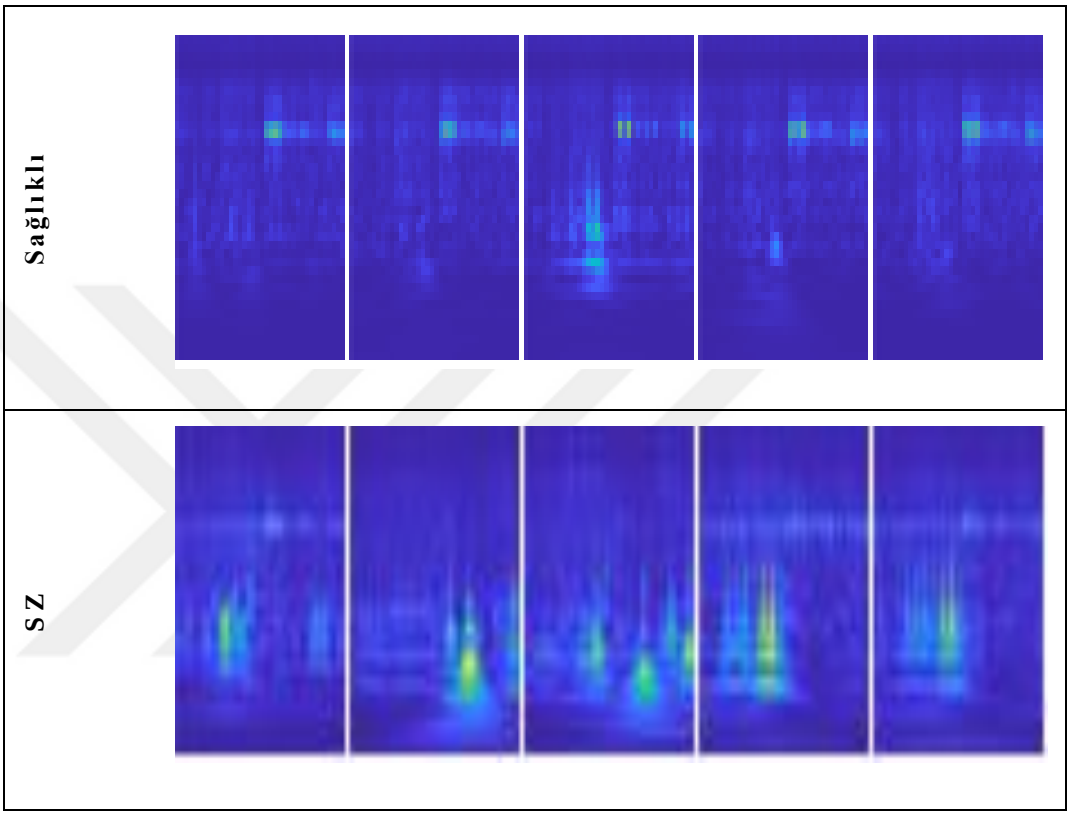


Şekil 4.11. Veri Kümesi A'dan farklı kişilere ait farklı segmentlerden alınan örnek skalogram görüntüleri

Elde edilen skalogram görüntüleri homojen olarak %80 oranında eğitim ve %20 test veri setlerine ayrılmıştır. VGG-16 derin öğrenme ağ mimarisi kullanılarak SZ hastalarını ve sağlıklı kontrolleri ayırt etmek için elde edilen skalogram görüntüleri ağa verilerek sınıflandırma işlemi gerçekleştirilir. VGG-16 derin öğrenme ağı için, girdi verisi olarak verilen görüntü boyutu 112x112, 1.0e-4 öğrenme oranı, batch size 128 ve optimize edici Adam (Kingma ve Ba 2014) olacak şekilde hiper parametreler belirlenmiştir. Bu çalışmada yapılan deneylerde ağın eğitilebilmesi için 100 iterasyon değerinin yeterli olduğu belirlenmiştir. Ağın eğitimi sonunda ortalama olarak 78. iterasyonda 98% gibi yüksek bir doğruluk değeri elde edilmiştir.

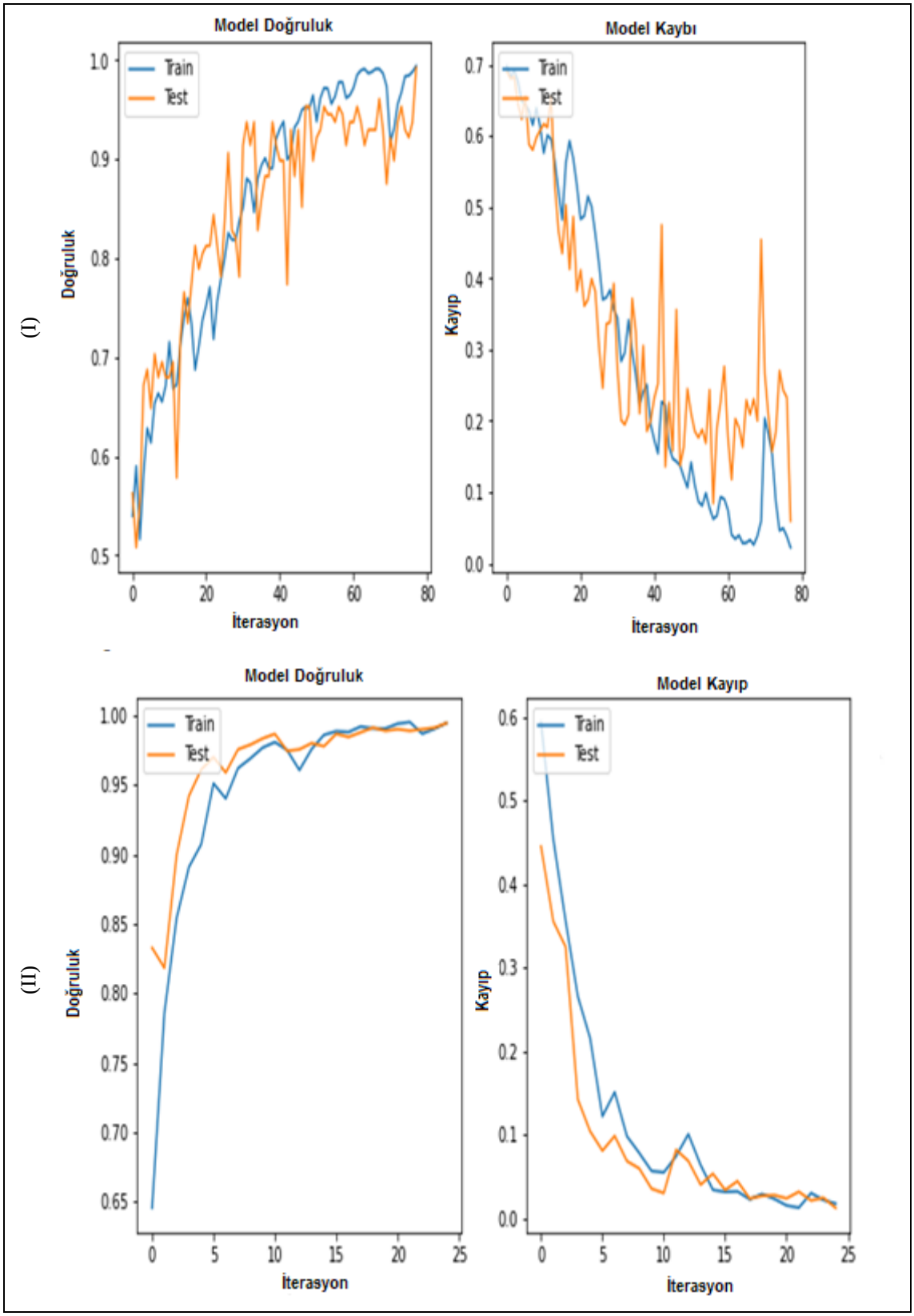
Veri seti B'ye ait EEG verisi, 12 ile 15 dakika arasında değişen sürelerden yapılan kayıtlardan oluşmaktadır. Bundan dolayı segmantasyon işlemi gerçekleştirilirken en kısa kaydın uzunluğu dikkate alınarak gerçekleştirilmiştir. Her kanaldan alınan EEG kayıtları 5sn'lik segmentlere ayrılmıştır. Bu segmantasyon sonucunda EEG kayıtlarının farklı uzunlukta olması göz önüne alınarak her SZ hastası için 148 ve her sağlıklı kontrol için ise 173 segment elde edilmiştir. Şekil 4.12.'de Veri seti B'den elde edilen farklı kişilere

ait farklı segmentlerden alınan örnek skalogram görüntüleri gösterilmektedir. Veri seti A'da VGG-16 derin öğrenme ağını eğitmek için kullanılan hiper parametreler aynı şekilde Veri seti B'de de kullanılmıştır. Veri seti A'da kullanılan yöntemin aynısı veri seti B'ye uygulanarak toplam 4494 skalogram görüntüsü elde edilmiştir. Ağın eğitimin sonrasında, 25. iterasyon değerinde 99.5% doğruluğa ulaşılmıştır.



Şekil 4.12. Veri seti B'dan farklı kişilere ait farklı segmentlerden alınan örnek skalogram görüntüleri

Bu iki veri seti ile de elde edilen doğruluk değerleri literatürde bu veri setleri ile yapılmış çalışmalardan daha yüksek başarım elde etmiştir. Şekil 4.13' de iterasyon ve eğitim süresine göre Veri seti A ve B için doğruluk ve hata değerinin değişimini göstermektedir. Ağın yeteri kadar eğitimi modelin sınıflandırma başarımını ciddi oranda etkileyen önemli bir faktördür. Literatür çalışmalarında ağın ne kadar eğitilmesi ve bunun için gerekli olan belirli değerler bulunmamaktadır. En iyi başarım değerlerinin elde edilmesi deneysel çalışmalara bağlı olarak belirlenebilmektedir.



Şekil 4.13. Ağın eğitimi esnasında, iterasyon değerlerine karşı doğruluk ve hata değerinin değişimi: (I) Veri Kümesi A için (II) Veri Kümesi B için

Veri seti A için 98% ve Veri seti B için 99.5% F1-skor değerleri sınıflandırma görevinin başarılı bir şekilde gerçekleştiğini göstermektedir. Çizelge 4.6. ve Çizelge 4.7.'de gösterilen deney sonuçları incelendiğinde bu durum açıkça görülmektedir. Çizelge 4.6. ve Çizelge 4.7.' de gösterilen destek değeri ise her satırda bulunan toplam örnek sayısını belirtmektedir.

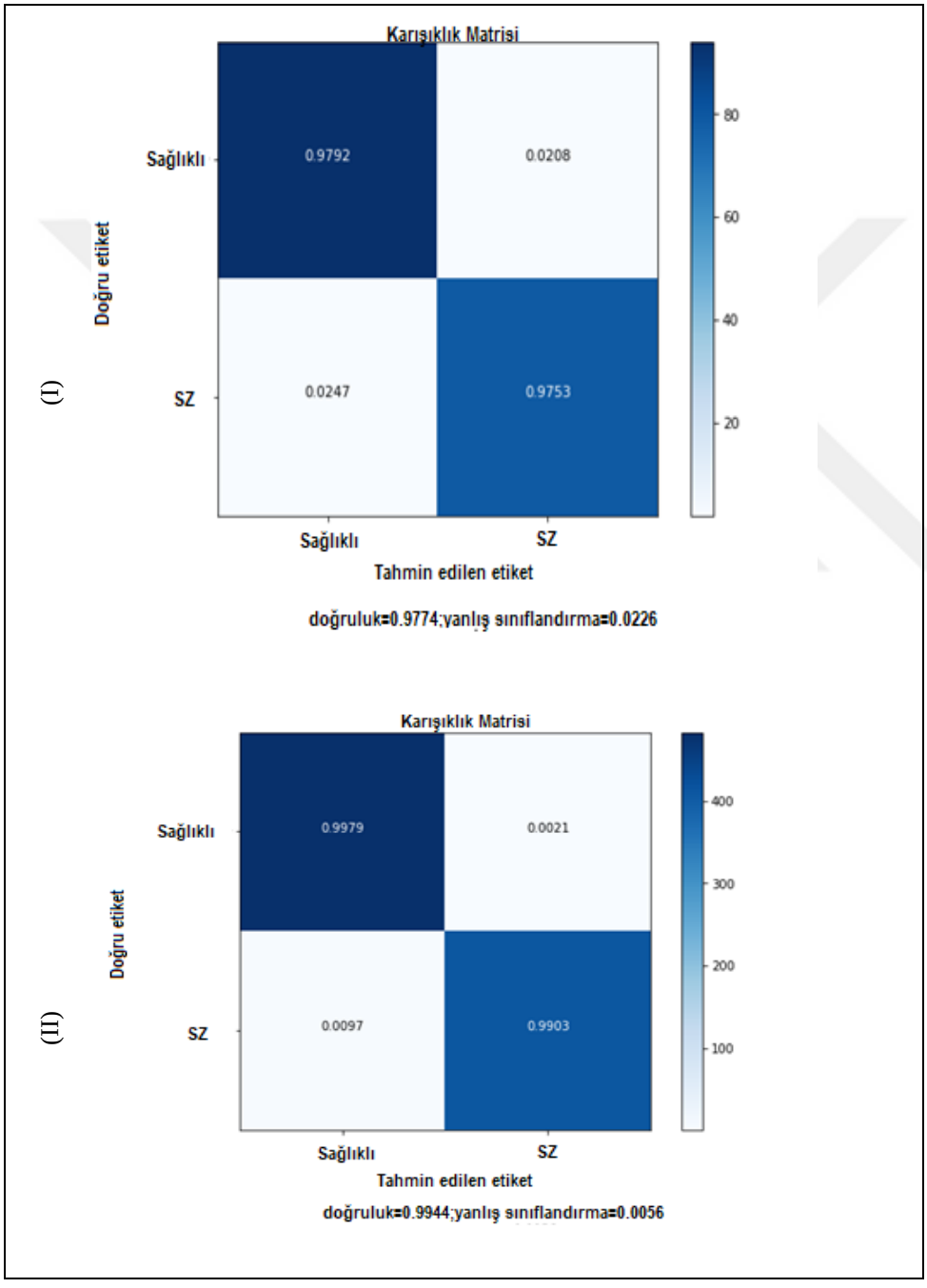
Çizelge 4.6. Skalogram görüntüleri kullanılarak, Veri Kümesi A ile elde edilen doğruluk değerlendirme ölçütleri

Veri Seti A	Precision	Recall	F1-score	Support
Sağlıklı Birey	0.98	0.98	0.98	94
SZ Hastası	0.98	0.98	0.98	108
Doğruluk			0.98	202
Makro ortalama	0.98	0.98	0.98	202
Ağırlıklı ortalama	0.98	0.98	0.98	202

Çizelge 4.7. Skalogram görüntüleri kullanılarak, Veri Kümesi B ile elde edilen doğruluk değerlendirme ölçütleri

Veri Seti B	Precision	Recall	F1-score	Support
Sağlıklı Birey	0.99	1.00	0.99	484
SZ	1.00	0.99	0.99	414
Doğruluk			0.99	898
Makro ortalama	0.99	0.99	0.99	898
Ağırlıklı ortalama	0.99	0.99	0.99	898

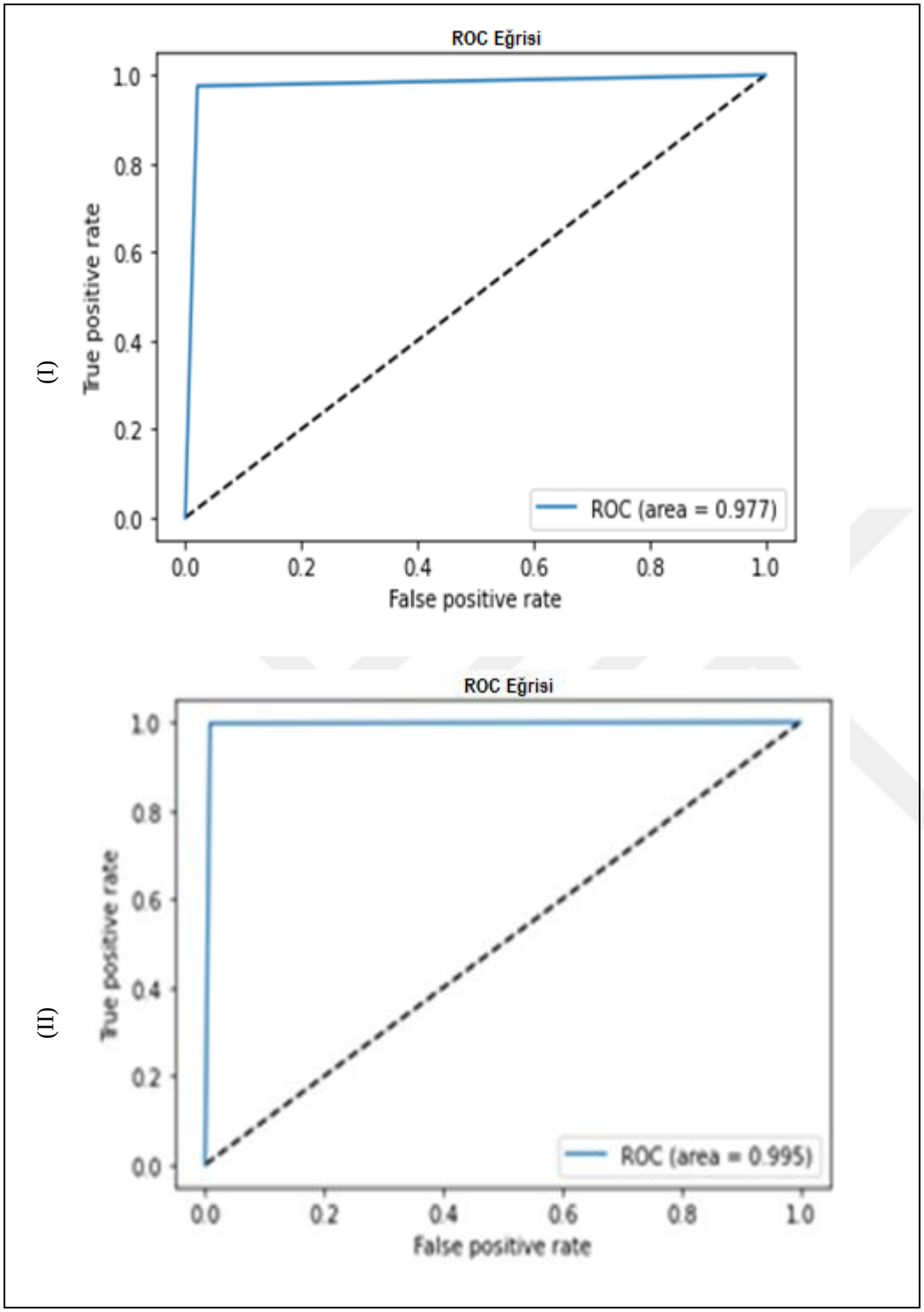
Hastalıklı ve sağlıklı kontroller için sınıflandırmanın doğruluk oranının 0.97'den yüksek olduğu ve her iki veri kümesi için sınıflandırmanın yanlış değerlendirme oranının 0.03'den düşük olduğu karışıklık matrislerinde gösterilmektedir (Bakınız Şekil 4.14). Şekil 4.15' de gösterildiği ROC eğrilerinde AUC değerinin veri seti A için 0.98 ve Veri seti B için 0.995 gibi oldukça yüksek olması ve grafikte gösterildiği üzere oldukça başarılı bir sınıflandırma gerçekleştirilmiştir.



Şekil 4.14. Karışıklık matrisleri: (I) Veri Kümesi A için (II) Veri Kümesi B için

Phang ve diğerleri tarafından veri seti A'da bulunan veriler çalışılmıştır (Phang, Ting, ve ark. 2019). Çalışmalarında 93.06% sınıflandırma başarımı elde ettiklerini belirtmişlerdir. Kullandıkları yöntemde ham EEG sinyallerinin doğrudan verildiği bir ESA derin öğrenme ağı kullanmışlardır. Veri seti B ise Oh ve diğerleri tarafından EEG sinyallerini ön işlem yapmadan doğrudan EEG sinyallerini ile beslenen bir ESA modeli kullanarak yöntemlerinin başarımlarını test etmişlerdir (Oh ve ark. 2019). Önerdikleri yöntem sonucunda çalışmalarında 98.07% başarım elde etmişlerdir. Bir önceki çalışmamızda, her iki veri setinden elde ettiğimiz spektrogram görüntülerini ESA derin öğrenme ağı ile eğiterek Veri seti A için 95% ve Veri seti B için ise %97 doğruluk ile sınıflandırma işlemini gerçekleştirdik.

Bu çalışma, bu alanda yapılmış literatür çalışmaları ile karşılaştırıldığında oldukça yüksek performans ve doğruluk değerleri elde edilmiştir. Kullandığımız yöntemin doğruluğunu kanıtlamak için farklı yaş gruplarından oluşan iki farklı veri seti ile çalışılarak değerlendirilmiştir. Çalışmamızda gerçekleştirdiğimiz yöntem her iki veri seti içinde oldukça yüksek doğruluk değerleri elde etmesi, yöntemin güvenilirliğini ispat etmektedir. Diğer literatür çalışmalarına göre, çalışmamızın öne çıkan en önemli unsurlarından bir diğeri de uzman görüşüne değerlendirmede yardımcı olabilecek daha kolay yorumlanabilen görüntüler oluşturmasıdır.



Şekil 4.15. ROC eğrileri: (I) Veri Kümesi A için (II) Veri Kümesi B için

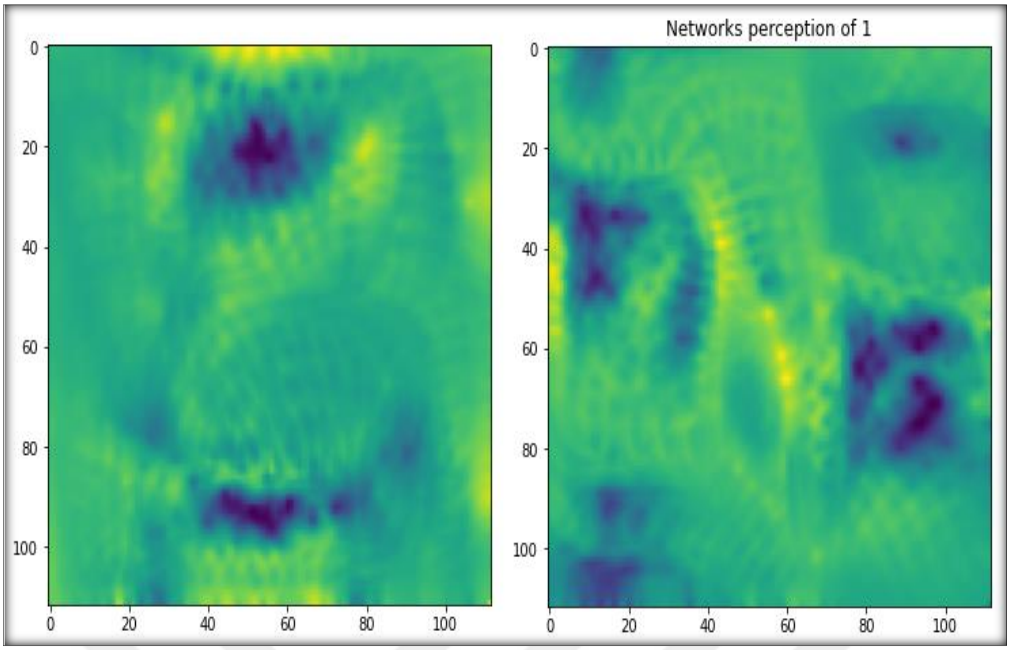
Ham EEG sinyalinde sadece zamansal ilişkiler bulunurken, elde ettiğimiz skalogram görüntülerinde daha fazla özellik bilgisi içermektedir. ESA derin öğrenme ağları pikseller arasındaki uzamsal ilişkilere sahip olan bir nesneyi sınıflandırmak için

uzamsal özellikleri kullandığından ham EEG sinyallerine göre Skalogram görüntüleri ESA ağı ile daha iyi bir sınıflandırma yapar. Bunun nedeni, ESA derin öğrenme ağı, skalogram görüntüleri üzerinden çok daha fazla özellik çıkarımı yaparak ağın daha yüksek başarım ile öğrenmesini sağlar.

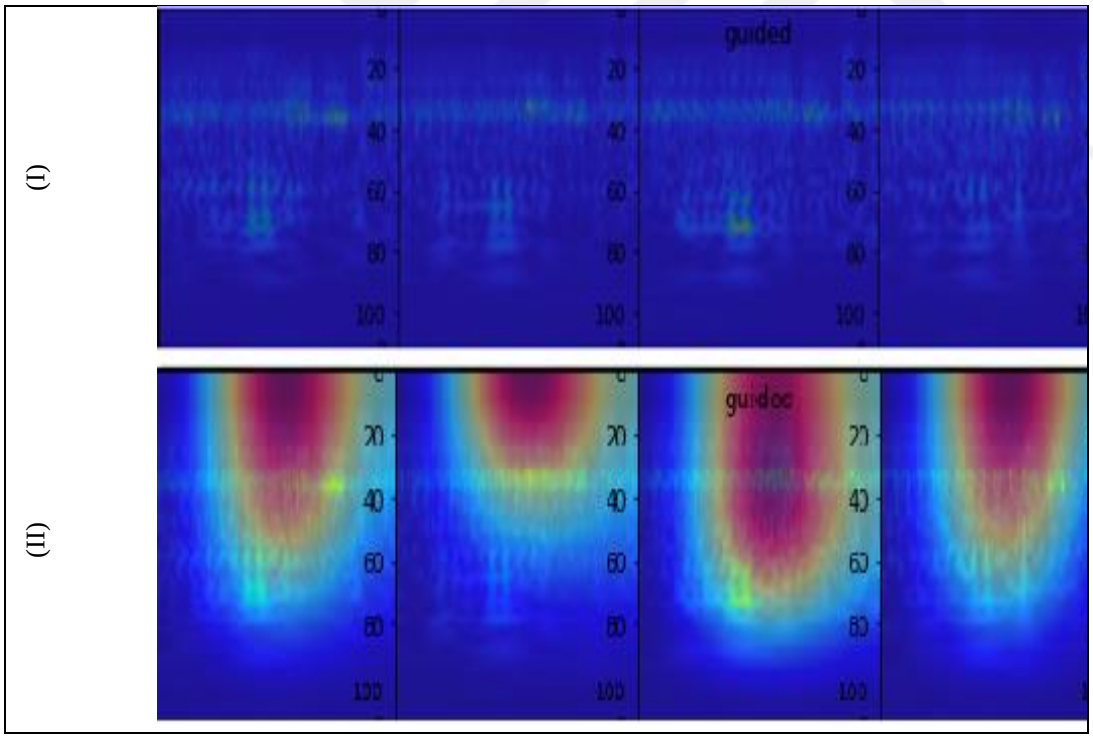
SDD, bir ana dalgacık ile dönüştürme işlemi esnasında sinyal boyunca farklı ölçekler kullanarak ve kaydırma yaparak zaman frekans analizini gerçekleştirir. Bu durum düşük frekanslarda uzun zaman aralıklı pencereleme ve yüksek frekanslarda kısa zaman aralıklı pencereleme sağlar. SDD ile farklı boyutlarda pencereler kullanmasıyla zaman serilerindeki yüksek ve düşük frekans bilgileri en iyi şekilde analiz edilir. Skalogram ise, SDD sonucu bir sinyalin 2D olarak tasvir edilmiş halidir. Dolayısıyla, Skalogram görüntülerinde sinyalin yüksek ve düşük frekans özellikleri hakkında daha fazla bilgi bulunmaktadır. Skalogram, düşük frekanslı, uzun süreli olaylar için daha iyi frekans lokalizasyonu ve kısa süreli, yüksek frekanslı olaylar için daha iyi zaman lokalizasyonu istendiği durumlardan yaygın olarak kullanılır.

Bir skalogram görüntüsünden, görüntünün hastaya veya sağlıklı bireye ait olup olmadığı doğrudan anlaşılmamaktadır. Şekil 4.11. ve Şekil 4.12.'de gösterilen sağlıklı ve SZ hastası bireylere ait skalogramlar incelendiğinde farklı zaman ve frekanslarda değişim olduğu görülebilir ancak açık bir şekilde yorumlanabilir olmadığı görülmektedir. Bir veri setindeki tüm veriler için genellemek ve formüle etmek mümkün değildir. ESA derin öğrenme ağının çıkışına ait görüntüleri Aktivasyon Maksimizasyonu (AC) (Kotikalapudi ve ark. 2017) adlı bir yöntem ile SZ hastası ve sağlıklı kontroller arasındaki farklılıkların daha net olarak yorumlanabilirliğini göstermeye çalışıyoruz.

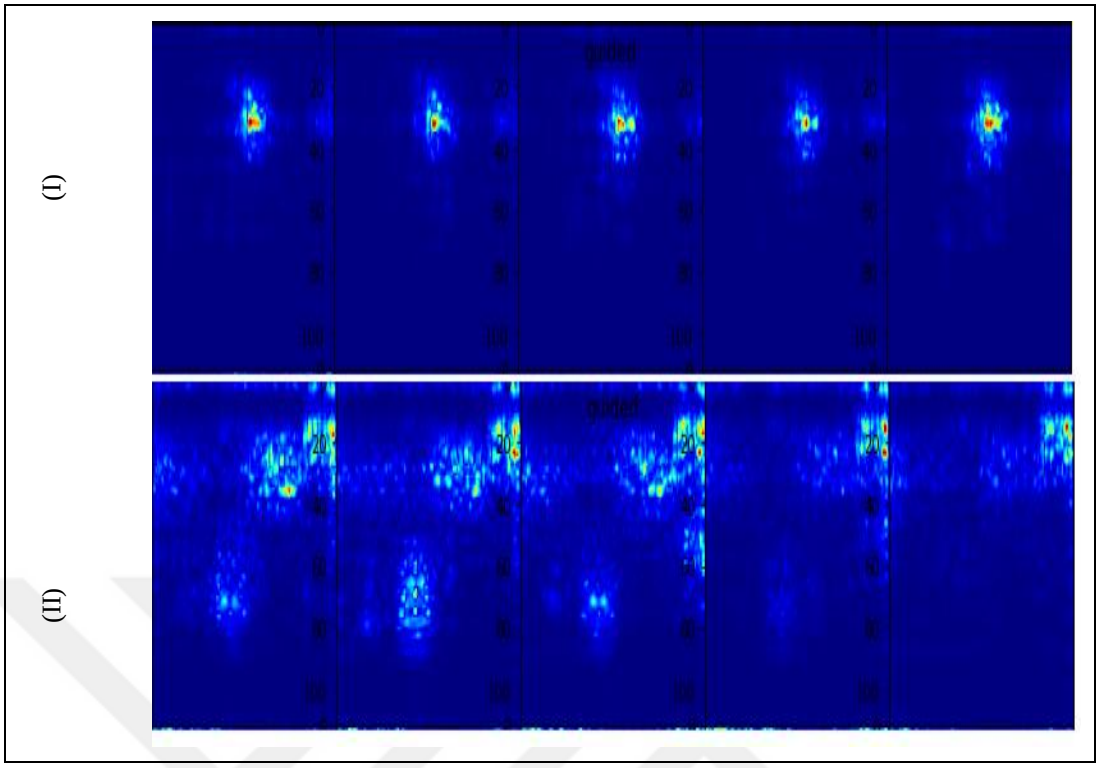
Bir sınıfa ait ağın çıktısını en üst düzeye çıkararak yinelemeli olarak bu değerleri bularak sentezlenen yapay bir görüntü oluşturduğundan, AC görüntüsü bir sınıfı betimleyen bir girdi olarak dikkate alınabilir. Şekil 4.16' da SZ hastası ve sağlıklı kontrole ait AC görüntüsünde bariz farklılıklar görülmektedir. Ancak AC görüntüleri, farklı skalogram görüntülerine ağın çok farklı şekilde yanıt verdiğini göstermektedir. AC görüntülerinde belirtilen özelliklerin, filtre aktivasyonunu en üst düzeye çıkarmak için suni olarak üretilmesinden dolayı, gerçek skalogram görüntülerini belirtmemektedir.



Şekil 4.16. ESA ağından elde edilen Sağlıklı kontrol (soldaki) ve SZ hastasına (sağdaki) ait aktivasyon maksimizasyon görüntüleri



Şekil 4.17. Farklı bireylerden ve farklı segmentlerden alınan Grad-CAM görüntüleri : (I) sağlıklı kontrollerden alınan (II) SZ hastalarından alınan



Şekil 4.18. Farklı bireylerden ve farklı segmentlerden alınan Belirginlik Haritası görüntüleri : (I) sağlıklı kontrollerden alınan (II) SZ hastalarından alınan

AC görüntülerinin bahsedilen dezavanjından dolayı, SZ hastasına ait farklılıkların daha net olarak ortaya konabilmesi için Grad-CAM (Gradyan Ağırlıklı Aktivasyon Haritaları) (Selvaraju ve ark. 2017) ve Belirginlik Haritası (Simonyan, Vedaldi, ve Zisserman 2013) tekniklerini kullandık. Grad-CAM, görüntü üzerindeki belirli bir sınıfa ait önemli ve ilgili özelliklerin renk haritası olarak gösterildiği görüntüler elde etmek için kullanılan bir tekniktir. Tahmin edilen sınıfa ait önemli uzamsal özellikler, grad-CAM görüntüsünde kırmızıya yakın renklerle gösterilmektedir. Şekil 4.17.'de, farklı bireylerden alınan farklı segmentlere ait skalogram görüntülerinde elde edilen Grad-CAM görüntüleri gösterilmektedir. SZ hastası bireylere ait Grad-CAM görüntüler incelendiğinde, üst kısımdan ortaya doğru uzanan belirgin bir yüksek yoğunluk görülürken, sağlıklı kotrollere ait görüntülerde ise hiçbir zaman bu yoğunluklu renkler görülmemektedir.

Skalogram görüntülerinin yorumlabilir olmasını sağlamak için kullandığımız diğer bir teknik ise Belirginlik Haritası tekniğidir. Belirginlik Haritası, bir pikselin parlaklığının pikselin ne kadar belirgin olduğunu temsil ettiği, yani bir pikselin parlaklığının belirginliğiyle doğru orantılı olduğu bir görüntüdür. Belirginlik haritaları,

sıcaklığın görüntünün nesnenin ait olduğu sınıfı tahmin etmede büyük etkisi olan bölgelerine atıfta bulunduğu bir ısı haritası olarak da adlandırılır. Belirginlik haritasının amacı, görsel alandaki her konumda öne çıkan veya göze çarpan bölgeleri bulmak ve çıkıntının mekansal dağılımına göre gidilen yerlerin seçimine rehberlik etmektir (D_believer 2020). Şekil 4.18.' de gösterilen ve farklı bireylerden alınan sağlıklı ve SZ hastası kişilere ait Belirginlik Haritası görüntüleri gösterilmektedir. Şekil 4.18.'de görüldüğü üzere sağlıklı ve SZ hastalarına ait Belirginlik Haritaları bariz farklarla ayırt edilebilmektedir. SZ hastalarına ait Belirginlik Haritaları incelendiğinde daha yoğun bir ısı haritası oluşturduğu gözlemlenmektedir.

Şekil 4.17. ve Şekil 4.18 'de gösterilen Grad-CAM ve Belirginlik Haritası yöntemlerinin şekilde gösterilmeyen diğer kişiler için ortak olduğu dikkate alınmalıdır. Bu iki yöntemle de açıkca görüldüğü üzere, Skalogram görüntüleri üzerinde frekans bileşenleri SZ hastaları ve sağlıklı kontrollerin ayırt edilmesinde önemli uzamsal özellikler içermektedir. İlgili literatür çalışmaları incelendiğinde önerilen yöntemin, tüm çalışmalardan daha yüksek doğruluk elde ettiği ve daha iyi performans sergilediği ve önerilen yöntemle karşılaştırılabilir bir performansa sahip olduğu görülmektedir.

Bu çalışma basit işlem süreci ve uzman görüşüne destek niteliğinde yorumlanabilir çıktılar üretmesiyle ilgili literatür çalışmalardan daha üstün durumdadır. SZ tespitinde bildiğimiz kadarıyla şu ana kadar elde edilmiş en yüksek doğruluk değeri elde edilmiştir. Önerilen çalışmamızın, yüksek performansının ötesinde daha önceki literatür yöntemlerine göre önemli avantajları bulunmaktadır. Öncelikle geleneksel makine öğrenme algoritmalarına kıyasla özellik çıkarımı ve ön işlem gerektirmemektedir. Ayrıca oldukça basit ama bir o kadar da yüksek performansa sahip işlem süreci içermektedir. Bir diğer avantaj durumu ise, SZ hastarının açık bir şekilde yorumlanabilir görüntülerini oluşturmak için 3 farklı teknik sunmaktadır. AC, Grad-CAM ve Belirginlik Haritası görüntülerinden, uzman görüşüne destek niteliğinde SZ hastasını ve sağlıklı kontrolü çok kolay bir şekilde ayırt etmeyi mümkün kılmaktadır.

5. SONUÇ VE ÖNERİLER

Bu tez çalışmasında, SZ hastası ve SZ hastası olmayan bir kontrolü otomatik olarak ayırt etmek için bir BDT yöntemi önerilmektedir. Önerilen yöntemi, farklı yaş gruplarını içeren iki farklı veri setiyle değerlendirilmiş olması yöntemin sınıflandırma başarımını net olarak ortaya koymaktadır. Tez çalışmasının ilk aşamasında SZ tespiti için bir takım ön işlem aşamalarından sonra geleneksel makine öğrenme algoritmaları kullanılmıştır. Tezin ikinci aşamasında ise önişlem aşamasına gerek kalmadan spektrogram ve skalogram görüntüleri oluşturularak derin öğrenme ağı ile eğitilmiştir. Skalogram görüntüleri kullanılarak en yüksek Veri seti B için 99.5% doğruluk ve veri seti A için ise 98% doğruluk ile literatürde bu alanda elde edilmiş en yüksek doğruluk değerlerine ulaşılmıştır.

Bu çalışma, EEG verilerindeki frekans bileşenlerinin analizinin, bir çeşit beyin deformansyonu olan SZ hastalığını tespit etmede güçlü bir yöntem olduğunu göstermektedir. Önerilen yöntemde, 3 farklı ESA ağı görselleştirme yöntemi (AC, Grad-CAM ve Belirginlik Haritası) kullanılarak SZ hastası ve sağlıklı kontrollerin açık bir şekilde ayırt edilebileceği görüntüler oluşturulmuştur. Bu tekniklerle elde edilen görüntülerin bariz farklar göstermesinin temelinde, girdi olarak verilen spektrogram ve skalogram görüntülerinin, zaman frekans düzeyinde önemli uzamsal özellikleri taşıması yer almaktadır.

Bu tez çalışmasında, EEG kayıtlarından farklı hastalıkların tespitinde de kullanılabilecek BDT çalışmaları için örnek teşkil edebilir. Yeni ESA modellerinin kullanımının, sınıflandırma performansına önemli ölçüde yarar sağlayacağı düşünülmektedir. Bunun yanında, ağın eğitimi için karmaşık modellerin yerine daha az katmandan oluşan daha basit yapıya sahip ve daha az hesaplama süresi gerektiren modeller tercih edilebilir.

5.1. Tez Kapsamında Yapılan Yayınlar

Science Citation Index (SCI) veya Science Citation Index Expanded (SCI-E) Yayınlar

- Automatic Detection of Schizophrenia by Applying Deep Learning over Spectrogram Images of EEG Signals. ASLAN,

Z. & AKIN, M. (2020). Traitement du Signal, 37(2), 235-244.

- A Deep Learning Approach in Automated Detection of Schizophrenia Using Scalogram Images of Eeg Signals. ASLAN, Z. & AKIN, M. (2020). (Değerlendirme Aşamasında).

Konferans / Sempozyum Bildirileri

- Performance Analysis Of Different Machine Learning Algorithms For The Detection of Schizophrenia with EEG Signals Aslan Zülfikar, Akın Mehmet Yayın Yeri: 4th International Energy Engineering Congress 2019

- A Comparison of Heuristic Search Algorithms in Automatic Detection of Schizophrenia. Aslan Zülfikar, Akın Mehmet Yayın Yeri: Uemk 2019

- Detection of Schizophrenia on EEG Signals By Using Relative Wavelet Energy As A Feature Extractor Aslan Zülfikar, Akın Mehmet Yayın Yeri: Uemk 2019

6. KAYNAKLAR

Abd-Alsabour, N. (2014, October). A review on evolutionary feature selection. In 2014 European Modelling Symposium (pp. 20-26). IEEE.

Acharya, U Rajendra ve ark. 2012. "Application of Non-Linear AND Wavelet Based Features for the Automated Identification of Epileptic EEG Signals." ***International journal of neural systems*** 22(02): 1250002.

Al-Aomar, R. (2000). A process-oriented system design methodology with genetic algorithm search and simulation-based evaluation (Order No. 9966114). Available from ProQuest Dissertations & Theses Global. (304634608).

Aslan, Zülfikar. 2019. "On The Use Of Deep Learnıng Methods On Medıcal Images." **The International Journal of Energy & Engineering Sciences** 3(2): 1–15.

Aslan, Zülfikar, ve Mehmet Akın. 2018. "Investıgatıon of Sımultaneous Eeg-Fmrı Analysıs Methods." Uemk 2018 proceedıngs book 18-19 ekim: 445.

Aslan, Zülfikar, ve Mehmet Akın. 2019a. "A Comparison of heuristic search algorithms in automatic detection of schizophrenia." In 4th Internatıonal energy & engıneerıng congress**,** 1248–58.

Aslan, Zülfikar, ve Mehmet Akın. 2019b. "Detection Of Schizophrenia On Eeg Signals By Using Relative Wavelet Energy As A Feature Extractor." In 4th ınternatıonal energy & engıneerıng congress, 301–310.

Aslan, Zülfikar, ve Mehmet Akın. 2019c. "Performance Analysis Of Different Machine Learning Algorithms For The Detection Of Schizophrenia With Eeg Signals." In 4th ınternatıonal energy & engıneerıng congress, 1066–81.

ASLAN, Zülfikar, ve Mehmet AKIN. 2020. "Automatic Detection of Schizophrenia by Applying Deep Learning over Spectrogram Images of EEG Signals." ***Traitement du Signa****l*.

Blocka, Karla, Erişim: [https://www.healthline.com/health/eeg], Erişim Tarihi :11.11.2020. 2020. "EEG (Electroencephalogram)."

Boostani, Reza, Khadijeh Sadatnezhad, ve Malihe Sabeti. 2009. "An Efficient Classifier to Diagnose of Schizophrenia Based on the EEG Signals." ***Expert Systems with Applications*** 36(3): 6492–99.

Borisov, S. V., A. Ya Kaplan, N. L. Gorbachevskaya, ve I. A. Kozlova. 2005. "Analysis of EEG

Structural Synchrony in Adolescents with Schizophrenic Disorders." **Human Physiology.**

Campbell, William M, Douglas E Sturim, Douglas A Reynolds, ve Alex Solomonoff. 2006. "SVM Based Speaker Verification Using a GMM Supervector Kernel AND NAP Variability Compensation." In 2006 IEEE International Conference on Acoustics Speech ve Signal Processing Proceedings, , I--I.

Cao, Bo ve ark. 2018. "Treatment Response Prediction AND Individualized Identification of First-Episode Drug-Na{\"\i}ve Schizophrenia Using Brain Functional Connectivity." ***Molecular psychiatry***: 1.

Celin, S., & Vasanth, K. (2018). ECG signal classification using various machine learning techniques. ***Journal of medical systems***, 42(12), 241.

Chakure, Afroz, Erişim: [https://aaaanchakure.medium.com/k-nearest-neighbors-knn-algorithm-bd375d14eec7], Erişim Tarihi :10.01.2021. 2021. "K-Nearest Neighbors (KNN) Algorithm"

Chandaka, S., Chatterjee, A., & Munshi, S. (2009). Cross-correlation aided support vector machine classifier for classification of EEG signals. ***Expert Systems with Applications***, 36(2), 1329-1336.

Chandran, A. N., Sreekumar, K., & Subha, D. P. (2021). EEG-Based Automated Detection of Schizophrenia Using Long Short-Term Memory (LSTM) Network. In Advances in Machine Learning and Computational Intelligence (pp. 229-236). **Springer**, Singapore.

Cheema, Muhammad Aamir, Wenjie Zhang, ve Lijun Chang. 2016. 9877 Databases Theory AND Applications: 27th Australasian Database Conference, ADC 2016, Sydney, NSW, September 28-29, 2016, Proceedings. Springer.

Chollet, François. 2017. "Xception: Deep Learning with Depthwise Separable Convolutions." In Proceedings - 30th IEEE Conference on Computer Vision AND Pattern Recognition, CVPR 2017,.

Collobert, Ronan, ve Samy Bengio. 2001. "SVMTorch: Support Vector Machines for Large-Scale Regression Problems." ***Journal of machine learning research*** 1(Feb): 143–60.

Cooper, Cyrus. 2018. "Global, Regional, AND National Incidence, Prevalence, AND Years Lived with Disability for 354 Diseases AND Injuries for 195 Countries AND Territories, 1990-2017: A Systematic Analysis for the Global Burden of Disease Study 2017." ***The Lancet*** 392(10159): 1789–1858.

Cristianini, Nello, John Shawe-Taylor,ve ark.. 2000. An Introduction to Support Vector Machines AND Other Kernel-Based Learning Methods. Cambridge university press.

D_believer. 2020. "What Is Saliency Map?" Erişim: [https://www.geeksforgeeks.org/what-is-saliency-map/], Erişim Tarihi : 21.11.2020.

Dechter, Rina, ve Judea Pearl. 1985. "Generalized Best-First Search Strategies AND the Optimality of A." ***Journal of the ACM (JACM)*** 32(3): 505–36.

Dehuri, S., & Ghosh, A. (2013). Revisiting evolutionary algorithms in feature selection and nonfuzzy/fuzzy rule based classification. ***Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery***, 3(2), 83-108.

Delorme, Arnaud, ve Scott Makeig. 2004. "EEGLAB: An Open Source Toolbox for Analysis of Single-Trial EEG Dynamics Including Independent Component Analysis." ***Journal of neuroscience methods*** 134(1): 9–21.

Devia, Christ ve ark. 2019. "Eeg Classification during Scene Free-Viewing for Schizophrenia Detection." ***IEEE Transactions on Neural Systems AND Rehabilitation Engineering*** 27(6): 1193–99.

Devijver, Pierre A, ve Josef Kittler. 1982. Pattern Recognition: A Statistical Approach. Prentice hall.

Dimitoglou, George, James A Adams, ve Carol M Jim. 2012. "Comparison of the C4. 5 AND a Na{\"\i}ve Bayes Classifier for the Prediction of Lung Cancer Survivability." ***arXiv preprint*** arXiv:1206.1121.

Dorigo, Marco, ve Gianni Di Caro. 1999. "Ant Colony Optimization: A New Meta-Heuristic." In Proceedings of the 1999 Congress on Evolutionary Computation-CEC99 (Cat. No. 99TH8406), , 1470–77.

Duda, Richard O, Peter E Hart, ve David G Stork. 1998. "Pattern Classification (2nd Ed .)." ***Computational Complexity***.

Dvey-Aharon, Zack, Noa Fogelson, Avi Peled, ve Nathan Intrator. 2015. "Schizophrenia Detection AND Classification by Advanced Analysis of EEG Recordings Using a Single Electrode Approach." ***PloS one*** 10(4): e0123033.

Eberhart, Russell, ve James Kennedy. 1995a. "New Optimizer Using Particle Swarm Theory." In Proceedings of the International Symposium on Micro Machine AND Human Science,.

Eberhart, Russell, ve James Kennedy. 1995b. "Particle Swarm Optimization." In Proceedings of

the IEEE International Conference on Neural Networks, , 1942–48.

El-Dahshan, El Sayed Ahmed, Tamer Hosny, ve Abdel Badeeh M. Salem. 2010. "Hybrid Intelligent Techniques for MRI Brain Images Classification." ***Digital Signal Processing: A Review Journal.***

Elder, J, ve D Pregibon. 1996. "A Statistical Perspective on KDD." Advances in knowledge discovery AND data mining.

Emanet, Nahit. 2009. "ECG Beat Classification by Using Discrete Wavelet Transform AND Random Forest Algorithm." In 2009 Fifth International Conference on Soft Computing, Computing with Words AND Perceptions in System Analysis, Decision AND Control, , 1–4.

Farrell, Kevin R, Richard J Mammone, ve Khaled T Assaleh. 1994. "Speaker Recognition Using Neural Networks AND Conventional Classifiers." ***IEEE Transactions on speech AND audio processing*** 2(1): 194–205.

Fraiwan, L., Lweesy, K., Khasawneh, N., Wenz, H., & Dickhaus, H. (2012). Automated sleep stage identification system based on time–frequency analysis of a single EEG channel and random forest classifier. ***Computer methods and programs in biomedicine***, 108(1), 10-19.

Freund, Yoav, ve Llew Mason. 1999. "The Alternating Decision Tree Learning Algorithm." In Icml, , 124–33.

Gardner, M. W., & Dorling, S. R. (1998). Artificial neural networks (the multilayer perceptron)—a review of applications in the atmospheric sciences. ***Atmospheric environment***, 32(14-15), 2627-2636.

Garg, Girisha, Vijander Singh, J R P Gupta, ve A P Mittal. 2011. Relatıve Wavelet Energy As A New Feature Extractor For Sleep Classıfıcatıon Usıng Eeg Sıgnals. ***International Journal of Biomedical Signal Processing***

Goldberg, David E. 1989. "Genetic Algorithms in Search, Optimization, AND Machine Learning, Addison Wesley, Reading, MA." Summary The Applıcatıons Of Ga-Genetıc Algorıthm For Dealıng Wıth Some Optımal Calculatıons In Economıcs.

Goodfellow, Ian, Yoshua Bengio, ve Aaron Courville. 2016. *Deep Learning*. MIT press.

Goutte, Cyril, ve Eric Gaussier. 2005. "A Probabilistic Interpretation of Precision, Recall AND F-Score, with Implication for Evaluation." In European Conference on Information Retrieval, , 345–59.

Guezzaz, A., Asimi, A., Asimi, Y., Tbatou, Z., & Sadqi, Y. (2019). A Global Intrusion Detection System using PcapSockS Sniffer and Multilayer Perceptron Classifier. ***IJ Network Security***, 21(3), 438-450.

Gudigar, Anjan ve ark. 2019. "Application of Multiresolution Analysis for Automated Detection of Brain Abnormality Using MR Images: A Comparative Study." ***Future Generation Computer Systems*** 90: 359–67.

Guyon, Isabelle, ve André Elisseeff. 2003. "An Introduction to Variable AND Feature Selection." ***Journal of machine learning research*** 3(Mar): 1157–82.

Hall, M. 1998. "Correlation-Based Feature Selection for Machine Learning." Waikato Uni.

Haykin, Simon. 1994. "A Comprehensive Foundation." *Neural Networks*.

Huang, Cheng-Lung, ve Chieh-Jen Wang. 2006. "A GA-Based Feature Selection AND Parameters Optimizationfor Support Vector Machines." ***Expert Systems with applications*** 31(2): 231–40.

Hubel, David H, ve Torsten N Wiesel. 1968. "Receptive Fields AND Functional Architecture of Monkey Striate Cortex." ***The Journal of physiology*** 195(1): 215–43.

Johannesen, Jason K. ve ark. 2016. "Machine Learning Identification of EEG Features Predicting Working Memory Performance in Schizophrenia AND Healthy Adults." ***Neuropsychiatric Electrophysiology***.

Keras. 2020. "Keras." Erişim : [https://en.wikipedia.org/wiki/Keras], Erişim Tarihi : 07.12.2020.

Khushaba, Rami N, Ahmed Al-Ani, Akram AlSukker, ve Adel Al-Jumaily. 2008. "A Combined Ant Colony AND Differential Evolution Feature Selection Algorithm." In International Conference on Ant Colony Optimization AND Swarm Intelligence, , 1–12.

Kim, Jun Won ve ark. 2015. "Diagnostic Utility of Quantitative EEG in Un-Medicated Schizophrenia." ***Neuroscience letters*** 589: 126–31.

Kingma, Diederik P, ve Jimmy Ba. 2014. "Adam: A Method for Stochastic Optimization." ***arXiv preprint*** *arXiv:1412.6980*.

Krisnandhika, B., Faqih, A., Pumamasari, P. D., & Kusumoputro, B. (2017, July). Emotion recognition system based on EEG signals using relative wavelet energy features and a modified radial basis function neural networks. In 2017 International Conference on Consumer Electronics and Devices (ICCED) (pp. 50-54). IEEE.

Kotikalapudi, Raghavendra, ve ark.. 2017. "Keras-Vis."

Kumar, Yatindra, Mohan Lal Dewal, ve Radhey Shyam Anand. 2012. "Relative Wavelet Energy AND Wavelet Entropy Based Epileptic Brain Signals Classification." ***Biomedical Engineering Letters***.

LabVIEW 2010 Advanced Signal processing. 2020. "Short-Time Fourier Transform." Erişim: [https://zone.ni.com/reference/en-XX/help/371419D-01/lvasptconcepts/aspt_stft/], Erişim Tarihi:12.11.2020.

Latifo\uglu, Fatma, Kemal Polat, Sad\ik Kara, ve Salih Güne\cs. 2008. "Medical Diagnosis of Atherosclerosis from Carotid Artery Doppler Signals Using Principal Component Analysis (PCA), k-NN Based Weighting Pre-Processing AND Artificial Immune Recognition System (AIRS)." ***Journal of Biomedical Informatics*** 41(1): 15–23.

Latifoğlu, Fatma, Kemal Polat, Sadik Kara, ve Salih Güneş. 2008. "Medical Diagnosis of Atherosclerosis from Carotid Artery Doppler Signals Using Principal Component Analysis (PCA), k-NN Based Weighting Pre-Processing AND Artificial Immune Recognition System (AIRS)." ***Journal of Biomedical Informatics***.

Laursen, Thomas Munk, Merete Nordentoft, ve Preben Bo Mortensen. 2014. "Excess Early Mortality in Schizophrenia." ***Annual review of clinical psychology*** 10: 425–48.

Litjens, Geert ve ark. 2017. "A Survey on Deep Learning in Medical Image Analysis." ***Medical image analysis*** 42: 60–88.

Lotte, F. ve ark. 2007. "A Review of Classification Algorithms for EEG-Based Brain-Computer Interfaces." ***Journal of Neural Engineering***.

McClelland, James L., ve David E. Rumelhart. 1988. "A Simulation-Based Tutorial System for Exploring Parallel Distributed Processing." ***Behavior Research Methods, Instruments, & Computers.***

Min, Seonwoo, Byunghan Lee, ve Sungroh Yoon. 2017. "Deep Learning in Bioinformatics." ***Briefings in bioinformatics*** 18(5): 851–69.

Morgan, David P, ve Christopher L Scofield. 1991. "Neural Networks AND Speech Processing." In Neural Networks AND Speech Processing, ***Springer***, 329–48.

Naira, C. A. T., & Alamo, C. J. L. D. (2019). Classification of People who Suffer Schizophrenia and Healthy People by EEG Signals using Deep Learning. **International Journal of Advanced Computer Science and Applications**, 10(10), 511-516.

Ng, Sin Chun, Chi Yin Chung, S I Leung, ve Andrew Luk. 1994. "An Evolutionary Search Algorithm for Adaptive IIR Equalizer." In Proceedings of IEEE International Symposium on Circuits AND Systems-ISCAS'94, , 53–56.

Oh, Shu Lih ve ark. 2019. "Deep Convolutional Neural Network Model for Automated Diagnosis of Schizophrenia Using EEG Signals." ***Applied Sciences*** 9(14): 2870.

Olejarczyk, Elzbieta, ve Wojciech Jernajczyk. 2017. "Graph-Based Analysis of Brain Connectivity in Schizophrenia." ***PloS one*** 12(11): e0188629.

Pawlak, Zdzisław. 2003. "A Rough Set View on Bayes' Theorem." ***International Journal of Intelligent Systems*** 18(5): 487–98.

Pearl, Judea. 1984. "Intelligent Search Strategies for Computer Problem Solving." Addision Wesley.

Phang, Chun-Ren, Fuad Noman, ve ark. 2019. "A Multi-Domain Connectome Convolutional Neural Network for Identifying Schizophrenia from EEG Connectivity Patterns." ***IEEE Journal of Biomedical AND Health Informatics*** 24(5): 1333–43.

Phang, Chun-Ren, Chee-Ming Ting, Fuad Noman, ve Hernando Ombao. 2019. "Classification of EEG-Based Brain Connectivity Networks in Schizophrenia Using a Multi-Domain Connectome Convolutional Neural Network." ***arXiv preprint*** *arXiv:1903.08858*.

Piryatinska, Alexandra, Boris Darkhovsky, ve Alexander Kaplan. 2017. "Binary Classification of Multichannel-EEG Records Based on the ϵ-Complexity of Continuous Vector Functions." ***Computer methods AND programs in biomedicine*** 152: 131–39.

Raschka, S. 2014. "An Overview of General Performance Metrics of Binary Classifier Systems." ***arXiv preprint*** *arXiv:1410.5330*.

Rosso, O. A. ve ark. 2006. "EEG Analysis Using Wavelet-Based Information Tools." ***Journal of Neuroscience Methods***.

Rosso, O. A., M. T. Martin, ve A. Plastino. 2002. "Brain Electrical Activity Analysis Using Wavelet-Based Informational Tools." **Physica A: Statistical Mechanics AND its Applications.**

RP Lippmann. 1987. "An Introduction to Computing with Neural Nets." ***IEEE ASSP Magazine.***

Rufai, Aminah Mardiyyah. 2020. "Generative vs. Discriminative Models in Machine Learning." Erişim: [https://medium.com/better-programming/generative-vs-discriminative-models-d26def8fd64a], Erişim Tarihi :07.12.2020.

Sakri, S. B., Rashid, N. B. A., & Zain, Z. M. (2018). Particle swarm optimization feature selection

for breast cancer recurrence prediction. ***IEEE Access***, 6, 29637-29647.

Santos-Mayo, Lorenzo, Luis M San-José-Revuelta, ve Juan Ignacio Arribas. 2016. "A Computer-Aided Diagnosis System with EEG Based on the P3b Wave during an Auditory Odd-Ball Task in Schizophrenia." ***IEEE transactions on biomedical engineering*** 64(2): 395–407.

Selvaraju, Ramprasaath R ve ark. 2017. "Grad-Cam: Visual Explanations from Deep Networks via Gradient-Based Localization." In Proceedings of the IEEE International Conference on Computer Vision, , 618–26.

Shalbaf, A., Bagherzadeh, S., & Maghsoudi, A. (2020). Transfer learning with deep convolutional neural network for automated detection of schizophrenia from EEG signals. **Physical and Engineering Sciences in Medicine**, 43(4), 1229-1239.

Shanmugam, N. (2000). Automated mammographic mass shape classification using wavelets and neural networks (Order No. 1399930).

Shim, Miseon ve ark. 2016. "Machine-Learning-Based Diagnosis of Schizophrenia Using Combined Sensor-Level ve Source-Level EEG Features." ***Schizophrenia research*** 176(2–3): 314–19.

Shirvany, Yazdan ve ark. 2013. "Application of Particle Swarm Optimization in Epileptic Spike EEG Source Localization." ***Applied Soft Computing*** 13(5): 2515–25.

Simonyan, Karen, Andrea Vedaldi, ve Andrew Zisserman. 2013. "Deep inside Convolutional Networks: Visualising Image Classification Models AND Saliency Maps." ***arXiv preprint*** *arXiv:1312.6034*.

Siuly, Siuly ve ark. 2018. "Exploring Hermite Transformation in Brain Signal Analysis for the Detection of Epileptic Seizure." ***IET Science, Measurement & Technology*** 13(1): 35–41.

Siuly, Siuly ve ark. 2020. "A Computerized Method for Automatic Detection of Schizophrenia Using EEG Signals." ***IEEE Transactions on Neural Systems AND Rehabilitation Engineering***.

Stockwell, Robert Glenn, Lalu Mansinha, ve R P Lowe. 1996. "Localization of the Complex Spectrum: The S Transform." ***IEEE transactions on signal processing*** 44(4): 998–1001.

Sui, Jing ve ark. 2014. "Combination of FMRI-SMRI-EEG Data Improves Discrimination of Schizophrenia Patients by Ensemble Feature Selection." In 2014 36th Annual International Conference of the IEEE Engineering in Medicine AND Biology Society, , 3889–92.

Supriya, Supriya, Siuly Siuly, Hua Wang, ve Yanchun Zhang. 2018. "EEG Sleep Stages Analysis AND Classification Based on Weighed Complex Network Features." ***IEEE Transactions on***

Emerging Topics in Computational Intelligence.

Talo, Muhammed, Ulas Baran Baloglu, Özal Y\ild\ir\im, ve U Rajendra Acharya. 2019. "Application of Deep Transfer Learning for Automated Brain Abnormality Classification Using MR Images." ***Cognitive Systems Research*** 54: 176–88.

Team, Edpresso. 2020. "What Is an Autoencoder?" Erişim : [https://www.educative.io/edpresso/what-is-an-autoencoder], Erişim Tarihi:07.12.2020.

Thilakvathi, B, S Shenbaga Devi, K Bhanu, ve M Malaippan. 2017. "EEG Signal Complexity Analysis for Schizophrenia during Rest AND Mental Activity. ***Biomedical Research***"

Tindall, Lucas, Cuong Luong, ve Andrew Saad. 2015. "Plankton Classification Using Vgg16 Network."

Torrence, Christopher, ve Gilbert P. Compo. 1998. "A Practical Guide to Wavelet Analysis." ***Bulletin of the American Meteorological Society***.

Tzimourta, K. D., Tsilimbaris, A., Tzioukalia, K., Tzallas, A. T., Tsipouras, M. G., Astrakas, L. G., & Giannakeas, N. (2018). EEG-based automatic sleep stage classification. ***Biomed J***, 1, 6.

WHO. 2020. "Schizophrenia." Erişim: [https://www.who.int/ mental_health/management /schizophrenia/en/], Erişim Tarihi : 11.11.2020.

Wikipedia. 2020a. "Deep Belief Network." Erişim: [https://en.wikipedia.org/wiki/ Deep_belief_network], Erişim Tarihi: 07.12.2020.

Wikipedia. 2020b. "Restricted Boltzmann Machine." Erişim : [https://en.wikipedia.org/wiki/ Restricted_Boltzmann_machine], Erişim Tarihi : 07.12.2020.

Yin, Jiao, Jinli Cao, Siuly Siuly, ve Hua Wang. 2019. "An Integrated MCI Detection Framework Based on Spectral-Temporal Analysis." ***International Journal of Automation AND Computing*** 16(6): 786–99.

Yuan, Longhao, ve Jianting Cao. 2017. "Patients' EEG Data Analysis via Spectrogram Image with a Convolution Neural Network." In International Conference on Intelligent Decision Technologies, , 13–21.

Zhang, Lei. 2019. "EEG Signals Classification Using Machine Learning for The Identification AND Diagnosis of Schizophrenia." In 2019 41st Annual International Conference of the IEEE Engineering in Medicine AND Biology Society (EMBC), , 4521–24.