

**T.C.
BİNGÖL ÜNİVERSİTESİ
FEN BİLİMLERİ ENSTİTÜSÜ**

**ELEKTRİKLİ EV CİHAZLARININ ZAMAN SERİSİ
ÖZELLİKLERİNE GÖRE BELİRLENMESİ**

YÜKSEK LİSANS TEZİ

EBRA NUR ZUREL

YENİLENEBİLİR ENERJİ SİSTEMLERİ ANABİLİM DALI

**TEZ DANIŞMANI
Dr. Öğr. Üyesi Muzaffer ASLAN**

BİNGÖL-EKİM 2022



T.C.
BİNGÖL ÜNİVERSİTESİ
FEN BİLİMLERİ ENSTİTÜSÜ



**ELEKTRİKLİ EV CİHAZLARININ ZAMAN SERİSİ ÖZELLİKLERİNE GÖRE
BELİRLENMESİ**

Dr. Öğr. Üyesi Muzaffer ASLAN danışmanlığında, Ebra Nur ZUREL tarafından hazırlanan bu çalışma 06/10/2022 tarihinde aşağıdaki jüri tarafından Yenilenebilir Enerji Sistemleri Anabilim Dalı'nda Yüksek Lisans Tezi olarak **oybirliği** ile kabul edilmiştir.

Başkan : Doç. Dr. Abdulkerim KARABİBER *İmza* :
Üye : Doç. Dr. Ömer Faruk ALÇİN *İmza* :
Üye : Dr. Öğr. Üyesi Muzaffer ASLAN *İmza* :

Yukarıdaki sonuç;

Enstitü Yönetim Kurulunun// tarih ve/
nolu kararı ile onaylanmıştır.

Prof. Dr. Zafer ŞİAR
Enstitü Müdürü

Bu çalışma TÜBİTAK 1002 projesi kapsamında desteklenmiştir.
Proje No: 121E347

Not: Bu tezde kullanılan özgün ve başka kaynaklardan yapılan bildirişlerin, çizelge, şekil ve fotoğrafların kaynak olarak kullanımı, 5846 sayılı Fikir ve Sanat Eserleri Kanunundaki hükümlere tabidir.

ÖNSÖZ

Tez çalışmaları süresince yardımlarını ve engin bilgi birikimini esirgemeyen, çalışmaların tamamlanabilmesi için gerekli desteği veren değerli hocam Dr. Öğr. Üyesi Muzaffer ASLAN'a teşekkür ederim. Tez hazırlama sürecinde bilgi ve görüşlerine başvurduğum Doç. Dr. Ömer Faruk ALÇİN'e de teşekkür ederim. Tez çalışmasına desteklerinden dolayı TÜBİTAK'a (Proje No: 121E347) teşekkür ederim.

Son olarak dualarını esirgemeyen anne ve babama, tezin hazırlanması sırasında gösterdikleri sabır, fedakârlık ve desteklerinden dolayı teşekkürü bir borç bilirim.

Ebra Nur ZUREL

Bingöl 2022

İÇİNDEKİLER

ÖNSÖZ	i
İÇİNDEKİLER	ii
SİMGELER VE KISALTMALAR LİSTESİ	iv
ŞEKİLLER LİSTESİ	v
TABLOLAR LİSTESİ	vi
ÖZET.....	vii
ABSTRACT.....	viii
1. GİRİŞ.....	1
2. KAYNAK ÖZETLERİ.....	6
3. MATERYAL VE YÖNTEM.....	13
3.1. Veri Setlerinin Analiz Edilmesi	13
3.1.1. UK-DALE Veri Seti.....	13
3.1.2. REDD Veri Seti.....	14
3.1.3. PLAID Veri Seti.....	14
3.1.4. Ev Aletlerinin Yüksek Çözünürlüklü Profilleri.....	14
3.2. Öznitelik Çıkarım Yöntemleri.....	15
3.2.1. İstatistiksel Yöntemler.....	16
3.2.2. Entropi	17
3.2.2.1. Çok Ölçekli Dağılım Entropi.....	19
3.2.3. Kısa Süreli Fourier Dönüşümü.....	20
3.2.4. Evrişimli Sinir Ağı	21
3.2.5. Uzun Kısa Süreli Bellek Ağı.....	23
3.3. Sınıflandırma	25
3.3.1. Destek Vektör Makine Sınıflandırıcı.....	26
3.3.2. En Yakın Komşu Sınıflandırıcı	27
3.3.3. Topluluk Öğrenimi	28
3.4. Performans Değerlendirme Ölçütleri	29
4. BULGULAR VE TARTIŞMA.....	30

4.1. Konutlardaki Elektrikli Cihazların Evriřimli Sinir Ađı ile Otomatik Sınıflandırılması	30
4.2. Geleneksel Makine Öğrenim Yöntemleri Elektrikli Cihazların Tespiti.....	34
4.3. Hibrit Model ile Elektrikli Cihazların Zaman Serisi Özelliklerine Göre Sınıflandırılması	36
5. SONUÇLAR VE ÖNERİLER.....	52
KAYNAKLAR	54
ÖZGEÇMİŐ	60

SİMGELER VE KISALTMALAR LİSTESİ

MOYİ	: Müdahaleci Olmayan Yük İzleme
KLE	: Karhunen Loeve Expansion
RTFSA	: Radyal Tabanlı Fonksiyon Sinir Ağı
DVM	: Destek Vektör Makinaları
k-NN	: k En Yakın Komşu
EBT	: Topluluk öğrenimi
HFD	: Hızlı Fourier Dönüşümünü
AFD	: Ayrık Fourier Dönüşümü
TSA	: Tekrarlayan Sinir Ağı
PSO	: Parçacık Sürü Optimizasyonu
AÖM	: Aşırı Öğrenme Makinaları
ESA	: Evrişimsel sinir ağı
ÇÖDE	: Çok Ölçekli Dağılım Entropisi
LSTM (UKSB)	: Uzun kısa süreli bellek
NCDF	: Normal Kümülatif Dağılım Fonksiyonu
FGMM	: Fonksiyonel Gizli Markov Modeli
KSFD	: Kısa Süreli Fourier Dönüşümü
γ	: Gama
δ	: Delta
σ	: Sigma
α	: Alfa
β	: Beta

ŞEKİLLER LİSTESİ

Şekil 1.1.	Dünyada sektörlere göre elektrik tüketimi	1
Şekil 1.2.	Mevcut enerji üretim kaynaklarının kullanımı	2
Şekil 3.1.	Zaman serilerinden spektrogram görüntülerin elde edilmesi.....	21
Şekil 3.2.	ESA'nın genel yapısı.....	22
Şekil 3.3.	Konvolüsyon işlemi	22
Şekil 3.4.	LSTM hücre mimarisi	24
Şekil 3.5.	DVM'de en iyi hiper düzlem	26
Şekil 4.1.	Elektrikli Cihazların ESA ile Otomatik Sınıflandırılması Modeli	30
Şekil 4.2.	Zaman serilerinden spektrogram görüntülerine dönüştürülmesi	31
Şekil 4.3.	ESA modelinin doğruluk ve kayıp eğrileri	32
Şekil 4.4.	ESA modeline ait karmaşıklık matrisi	33
Şekil 4.5.	Zaman serisinden kayan pencere yaklaşımı ile istatistiksel ve ÇÖDE özniteliklerin elde edilmesi	35
Şekil 4.6.	LSTM tabanlı hibrit modelin genel çerçevesi	37
Şekil 4.7.	Zaman serisinin 4x8 boyutlu geçici öznitelikleri matrisinin elde edilmesi	38
Şekil 4.8.	LSTM ağı eğitim sürecinin doğruluk ve kayıp grafikleri.....	39
Şekil 4.9.	k-NN ve LSTM+k-NN için karmaşıklık matrisi.....	41
Şekil 4.10.	EBT ve LSTM+EBT için karmaşıklık matrisi.....	42
Şekil 4.11.	DVM karmaşıklık matrisi.....	43
Şekil 4.12.	LSTM+DVM için karmaşıklık matrisi.....	44
Şekil 4.13.	k-NN, LSTM ve LSTM+k-NN modellerine ait ROC eğrileri.....	45
Şekil 4.16.	k-NN, LSTM ve LSTM+k-NN modellerine ait AUC değerlerinin karşılaştırılması.....	48

TABLolar LİSTESİ

Tablo 3. 1. Veri setindeki elektrikli cihazlar ve özellikleri.....	15
Tablo 3.2. Karmaşıklık matrisi.....	29
Tablo 4.1. ESA modeline ait eğitim parametreleri	32
Tablo 4.2. ESA model performans sonuçları (%).....	32
Tablo 4.3. ESA modelin mevcut bazı yöntemler ile performans karşılaştırması.....	34
Tablo 4.4. Geleneksel makine öğrenim yöntemleri sınıflandırma sonuçları	35
Tablo 4.5. LSTM modeli ve LSTM modeli ile birlikte k-NN, EBT ve DVM sınıflandırma performansları	39
Tablo 4.6. Geleneksel makine öğrenim yöntemleri ve LSTM tabanlı hibrit model performans sonuçlarının karşılaştırılması	40
Tablo 4.7. LSTM ağ temelli modelin, son teknoloji bazı çalışmaların başarımlarının karşılaştırılması	49
Tablo 4. 8. Hibrit model performansının son teknoloji yöntemlerle karşılaştırılması	50

ELEKTRİKLİ EV CİHAZLARININ ZAMAN SERİSİ ÖZELLİKLERİNE GÖRE BELİRLENMESİ

ÖZET

Son dönemlerdeki teknolojik gelişmelere bağlı olarak yaşanan hızlı değişimler beraberinde yaşamış olduğumuz çevreyi de bu gelişmelere uyumlu olmaya itmektedir. Yaşanan bu değişimler enerji kaynaklarına olan talebi büyük bir oranda arttırmaktadır. Özellikle konutlar ve endüstride da kullanılan elektrikli cihaz sayılarının artması ve süreç adaptasyonu zorunlu kıldığı gibi elektrik enerji tüketimini de oldukça artmaktadır. Bu durum özellikle modern toplumlarda giderek artan bir enerji talebine neden olmaktadır. Ancak sınırlı enerji kaynakları nedeniyle enerjinin verimli kullanımını da kritik öneme sahiptir.

Arz-talep dengesini etkili bir şekilde sağlanabilmesi, konutlardaki elektriksel yüklerin tanımlanmasına ve müdahaleci olmayan yük izleme sistemlerinin tasarımına bağlıdır. Müdahalesiz yük izleme yöntemleri (MOYİ) ile konutlarda kullanılan elektrikli cihazların kullanımını kontrol edilebilir ve kullanıcıların enerji tasarrufu için davranışları ayarlanabilir. Ev aletlerinin kontrolü ancak bu aletlerin etkin tespiti ile mümkündür.

Bu çalışmada, ev aletlerini tespit etmek için kayan pencere yaklaşımı, uzun kısa süreli bellek (LSTM) ağı ve sınıflandırıcı içeren yeni bir hibrit model önerilmiştir. Geleneksel özellik çıkarımının yanı sıra önerilen bu modelde, cihaz özelliklerini belirlemek için zaman serisi özniteliklerini kullanılmıştır. Böylece her bir zaman serisi, kayan pencere tabanlı yaklaşım ile örtüşmeyen pencerelere ayrılmıştır. Her pencerenin ortalama, standart sapma, medyan ve çok ölçekli entropi değerleri hesaplanmıştır. Bu değerler birleştirilerek, derin özellik çıkarımı için LSTM ağına uygulanmıştır. LSTM ağından alınan bu derin öznitelikler ev cihazlarını tespit etmek için k yakın komşuluk (k-NN), topluluk öğrenimi (EBT) ve destek vektör makinesi (DVM) sınıflandırıcılarına uygulanmıştır.

Önerilen modelin etkinliği, ev aletleri veri setinin yüksek çözünürlüklü profilleri ile test edilmiştir. Deneysel çalışmalarda, LSTM tabanlı k-NN, EBT ve DVM sınıflandırıcılarında sırasıyla %98,25, %97,81 ve %97,38 doğrulukla ev tipi cihazlar tespit edilmiştir.

Anahtar Kelimeler: Müdahalesiz yük izleme yöntemleri, uzun kısa süreli bellek, destek vektör makinesi, sınıflandırma, zaman serileri.

DETERMINATION OF HOUSEHOLD ELECTRIC APPLIANCES ACCORDING TO THE TIME SERIES FEATURES

ABSTRACT

Rapid changes due to recent technological developments push the environment we live in to be compatible with these developments. These changes greatly increase the demand for energy resources. The increase in the number of electrical devices used in residences and industry necessitates adaptation to the process and the consumption of electrical energy. This situation causes an increasing energy demand, especially in modern societies. However, efficient energy use is also critical due to limited energy resources.

Effective supply-demand balance depends on identifying electrical loads in the residences and designing non-intrusive load monitoring systems. With non-intervention load monitoring methods (NLM), the use of electrical appliances in residences can be controlled, and the behavior of users can be adjusted for energy saving. Control of household appliances is only possible with effective detection of these appliances.

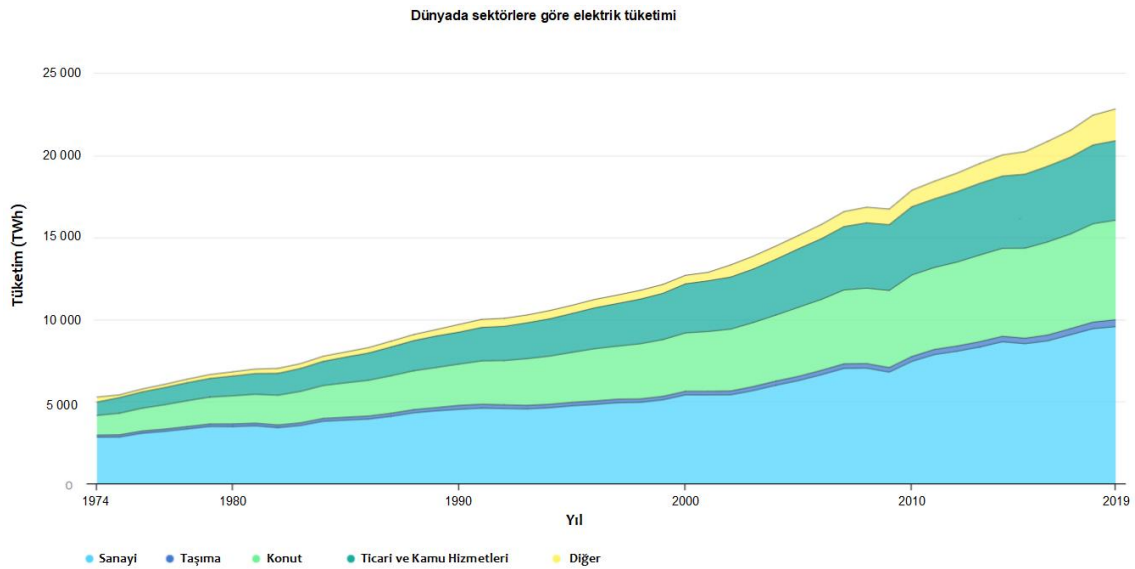
This study proposes a new hybrid model including sliding window approach, long short-term memory (LSTM) network, and a classifier to detect household appliances. In addition to traditional feature extraction, in this proposed model, time series features are used to determine device features. Thus, each time series is divided into windows that do not overlap with the sliding window-based approach. Each window's mean, standard deviation, median and multiscale entropy values were calculated. These values are combined and applied to the LSTM network for deep feature extraction. These in-depth features from the LSTM network were applied to k-neighborhood (k-NN), ensemble learning (EBT) and support vector machine (SVM) classifiers to detect home devices.

The effectiveness of the proposed model was tested with high-resolution profiles of the home appliance dataset. In the experimental studies, household appliances were detected in LSTM-based k-NN, EBT, and SVM classifiers with an accuracy of 98.25%, 97.81%, and 97.38%, respectively.

Keywords: Unattended load monitoring methods, long short-term memory, support vector machine, classification, time series.

1. GİRİŞ

Son dönemlerde, dünyamızda teknolojik anlamda yaşanan hızlı gelişmeler ve artan nüfus ile birlikte enerji talebi de önemli oranda arttırmaktadır. Özellikle 2021 yılında 7,9 milyara ulaşan dünya nüfusu ile birlikte, sanayii ve konut sektörlerinde kullanımı artan teknolojik elektrikli cihazların yaygınlaşması enerji talebini de günbegün arttırmaktadır. Uluslararası Enerji Ajansı istatistiklerine göre, 2018 yılında dünya elektrik tüketimi 22315 TWh'e ulaştı [1]. Elektrik tüketimi, sektör bazında karşılaştırıldığında konut sektörü Şekil 1.1'de gösterildiği gibi toplam kullanımın %26,92'sini oluşturmaktadır [1]. Özellikle Hindistan gibi gelişmekte olan ülkelerde elektrik tüketimi son yirmi yılda yaklaşık dört kat artmıştır [2]. Bu artışın yaklaşık %24'ü konutlarda gerçekleşmiştir [3].

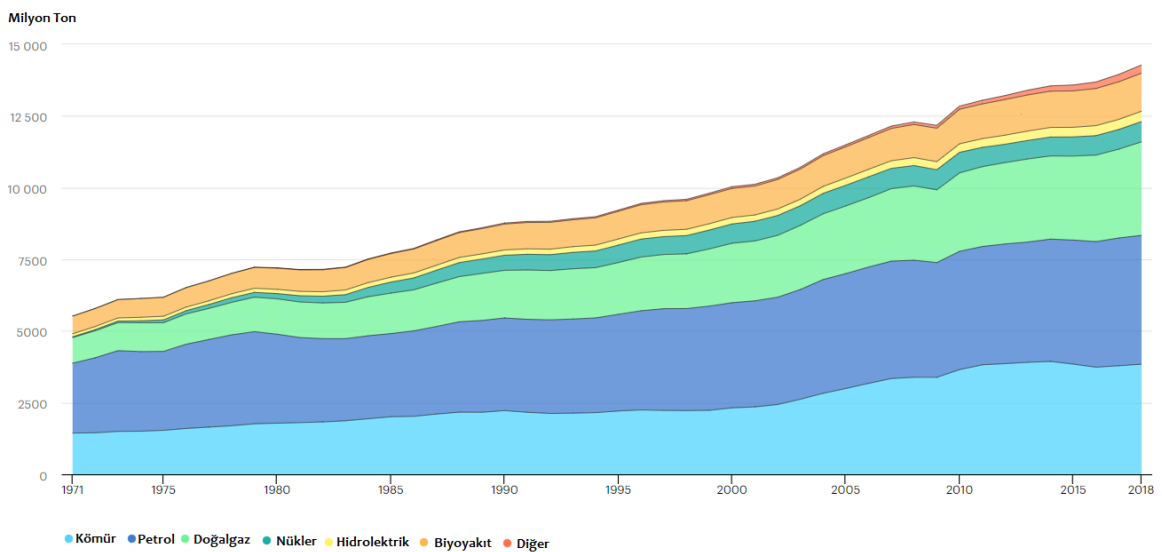


Şekil 1.1. Dünyada sektörlere göre elektrik tüketimi [4]

Daha yüksek yaşam standartları ve daha hızlı büyümeye sahip gelişmiş ülkelerde bu oran çok daha yüksektir. Örneğin Amerika'da 2017 yılında toplam elektrik tüketiminin %6,57'si konutlarda kullanılmıştır [3]. Bu artan elektrik tüketiminin arkasında, sadece hızlı nüfus artışı değil aynı zamanda dünya genelinde kırsal alanlardan şehirlere göçün yönlendirildiği, benzeri görülmemiş, onlarca yıldır süren bir kentleşme eğiliminden de

kaynaklanmaktadır. 2008'de dünya nüfusunun sadece yarısının şehirlerde yaşıyor olmasına rağmen dünyadaki enerjinin üçte ikisi bu kentsel alanlarda yaşayanlar tarafından tüketiliyordu [4]. Dünyadaki topraklarının sadece %5'ini kaplayan büyük şehirler, dünyadaki fosil yakıt kullanımının %75'ini oluşturduğundan, hızlı kentleşme ve şehirlerdeki enerji talebindeki çok büyük artışlar endişe vericidir [5]. Ayrıca, 2050 yılına kadar dünyada kentsel nüfusunun toplam nüfusunun %68,4'üne ulaşacağı beklendiğinden, gelecekte elektrik enerjisi talebinin ciddi artışlara neden olabileceğini göstermektedir.

Bu arada, yaşadığımız COVID-19 pandemi sürecinde birçok ülkede uygulanan karantina insanları evde kalmaya zorladı ve konutların rolü değişmeye başladı [6]. Bu süreçte barınma ve yaşam alanı olan konutlar esnek işyerleri, çevrimiçi öğrenme, uzaktan tıbbi hizmetler ve sosyal mesafeye yönelik diğer yaşam biçimlerinin doğal hale geldiği yeni normal döneminin ortaya çıkmasıyla birlikte eğitim, çalışma, toplantı gibi çeşitli sosyal etkinliklerin gerçekleştirilebildiği çalışma alanı olarak da kullanılmaya başlandı. Bu durum konutlarda harcanan sürenin artmasına bağlı olarak daha kaliteli ve verimli bir zamana geçirebilmek adına son teknolojik elektrikli cihazların kullanımında artışa ve bunun sonucunda konutlardaki elektrik tüketiminin de artmasına neden oldu [7]. COVID-19 karantinası sırasında Polonya, Varşova'da 7000 konut arasında yürütülen bir araştırma, Mart ve Nisan 2020 dönemindeki aktif enerji tüketiminin 2018'in aynı dönemine göre %16'lık bir artışa neden olduğu belirtilmektedir [8].



Şekil 1.2. Mevcut enerji üretim kaynaklarının kullanımı [4]

Burada belirtilen tüm bu gelişmeler elektrik enerjisi talebinin her geçen gün daha önemli oranda arttırmaktadır. Artan bu enerji talebinin karşılanabilmesi için Şekil 1.2’de görüldüğü gibi yıllara göre sınırlı olan fosil yakıt kaynaklarının daha fazla kullanımına neden olmaktadır. Fakat bu durum daha fazla sera gazı salınımı ve küresel iklim değişimi gibi ciddi problemlere sebep olmaktadır. Konutlarda tüketilen enerji, dünyadaki enerjinin %40’ını ve toplamın üçte birinden fazlasını karbondioksit salınımını oluşturduğu belirtilmektedir [9]. Bu nedenle özellikle gelişmiş ve gelişmekte olan ülkeler çeşitli yenilikçi, teknolojik ve sürdürülebilir farklı enerji kaynakları arayışına yönelmekte ve diğer bir yandan da mevcut enerji kaynaklarının nasıl daha verimli kullanılabileceğini araştırmaktadır [10]. Enerjinin verimli kullanımı ve enerji maliyetlerinin en alt seviyeye düşürmek için tüm sanayii tesislerinde önemli çalışmalar yapılmakla birlikte dünyada en büyük enerji tüketicilerinden biri olan konutlarda bu çalışmalar istenilen seviyelerde değildir.

Bu nedenle son dönemlerde araştırmacılar konutlarda artan bu enerji tüketiminin doğru ve detaylı bir şekilde tanımlayabilmek için enerji izleme çalışmalarına ağırlık vermişlerdir. Bu kapsam yapılan bazı çalışmalar konutlardaki elektrikli cihazların enerji tüketimlerinin izlenmesi ve geri bildirimlerinin konutlar ve binalarda % 12’lik elektrik gücü tasarrufuyla sonuçlanabileceği kanıtlanmıştır [11]. Bilimsel çalışmalarda konutlardaki elektrikli cihazların enerji tüketimlerinin tespitinde müdahaleci yük izleme ve müdahaleci olmayan yük izleme olmak üzere iki farklı yaklaşım benimsenmiştir. Müdahaleci yaklaşımlarda konutlarda izlenecek her bir cihaz için ayrı bir sensör kullanımı izleme maliyetini artırması nedeni ile son dönemlerde pek tercih edilmemektedir. MOYİ yaklaşımında, konutlarda tek bir merkezi noktadan toplam güç tüketimini bir sensör veya akıllı sayaç yardımıyla ölçerek ve bu güç tüketimini konutlardaki cihazlar için ayırarak her bir cihazın enerji tüketimini belirlemektedir [12,13].

Diğer bir yandan, literatürde MOYİ yaklaşımları, örüntü tanıma ve matematiksel optimizasyon veya olay tabanlı ve durum tabanlı olarak sınıflandırılmalarından bahsedilmiştir [14,15]. Her bir elektrikli ev cihazının toplam güç tüketimi bilgisine göre bir amaç fonksiyonu ve kısıtlama koşulu belirlenir. Belirlenen bu amaç ile fonksiyonunun çözümü için farklı veri setleri ve sınıflandırmalar kullanılarak karşılaştırılmalar yapılmıştır. Evlerdeki cihaz sayısının artması ve hesaplama karmaşıklığı nedeniyle cihazları kategorize

etmeyi zorlaştırmıştır. Ayrıca konutlardaki cihaz sayısındaki artış, benzer işaretlere sahip elektrikli cihazların sınıflandırılmasını daha da karmaşık hale getirmiştir [14]. Bu nedenle, araştırmacılar son zamanlarda örüntü tanıma tabanlı makine öğrenme algoritmalarını ve derin öğrenme yöntemlerini tercih etmişlerdir [16–18]. Örüntü tanıma tabanlı yaklaşımda, konutlardaki elektrikli cihazların özellikleri öğrenilir ve cihazın durumu belirlenmektedir. Fakat teknolojik gelişmelere bağlı olarak konutlarda kullanılan elektrikli cihaz sayısındaki artış ve ayrıca bu cihazların çoğunun çok fonksiyonlu olması tek bir noktada ölçülen verilerin boyutunu artırmakta ve hesaplama karmaşıklığına neden olmaktadır. Cihazların çalışma frekanslarındaki farklılıklar da sınıflandırmayı oldukça zorlaştırmaktadır.

Bu tez çalışmasında, yukarıda belirtilen problemleri çözmek ve konutlarda kullanılan elektrikli cihazları tespit etmek için bu cihazların güç tüketimine ait zaman serilerinin istatistiksel, bilgi tabanlı ve derin öğrenme yöntemlerine ait özelliklerinin kullanıldığı bir hibrit bir model önerilmiştir. Bu hibrit modelde, elektrikli ev cihazlarına ait güç tüketim zaman serilerinden kayan pencere yaklaşımı ile ortalama, standart sapma, medyan ve çok ölçekli dağılım entropi öznitelikleri birleştirilip LSTM uygulanarak derin öznitelikler elde edilmiştir. Bu derin öznitelikler konutlardaki elektrikli cihazların tespiti için k-NN, EBT ve DVM gibi popüler makine öğrenim yöntemlerine uygulanmıştır. Hibrit modelimizin katkıları şu şekilde sıralanabilir:

- Tipik özellik çıkarma perspektifinden farklı olan cihazların karakterlerini göstermek için zaman serisi özellikleri kullanılmıştır.
- Çok fazla uzman bilgisine ihtiyaç duymadan, kayan pencere yaklaşımına dayalı olarak, istatistiksel ve entropi değerleri kullanılarak ham zaman serisi verilerinden öznitelik çıkarımı gerçekleştirilerek, LSTM'nin hesaplama maliyeti azaltılmıştır.
- LSTM modelinin kritik verileri bellekte tutma özelliğinden yararlanılarak daha etkin öznitelik çıkarımı yapılmıştır.
- Önerilen hibrit modelde etkin öznitelik çıkarımı sayesinde sınıflandırma performansı artırılmıştır.

Bu tez çalışması, beş bölümden oluşmaktadır. Giriş bölümünde; son dönemlerde ekonomik ve sosyal gelişmelere paralel olarak artan enerji talebi ve bu talebin karşılanması için

yenilenebilir enerji kaynaklar ile mevcut enerji kaynaklarının verimli kullanımı hakkında özet bilgiler verilmiştir. Ayrıca bu bölümde tezin gerekçeleri ve katkılarına yer verilmiştir.

Kaynak özetleri bölümünde, literatür çalışmalarında konutların artan enerji talebinin karşılanmasına yönelik çalışmalardan bahsedilmiştir.

Materyal ve metot bölümünde, konutlardaki elektrikli cihazların tespitinde yaygın olarak kullanılan veri setleri, bu veri setlerinden öznitelik çıkarım yöntemleri ve araştırmacılar tarafından yaygın olarak tercih edilen sınıflandırıcı yöntemleri detaylandırılmıştır.

Deneysel çalışmalar ve tartışma bölümünde, konutlarda kullanılan elektrikli cihazların tespiti için önerilen yöntemler ile bu yöntemlere ait deneysel çalışmalardan elde edilen sonuçlar ve literatürde son dönemlerde yapılan bazı çalışmaların sonuçları ile karşılaştırmalar sunulmuştur.

Sonuçlar ve önerilerde, tez çalışmasında önerilen yöntemler ile gerçekleştirilen çalışmalardan elde edilen sonuçlar ve gelecekte yapılması planlanan çalışmalar hakkında bilgiler sunulmuştur.

2. KAYNAK ÖZETLERİ

Yeni nesil akıllı ev sistemlerinde yapılan çalışmalarda, enerji maliyetinin düşürülmesi ve verimli kullanılması için çeşitli yöntemler araştırılmakta ve geliştirilmektedir. Özellikle ev ve ofis gibi sıklıkla kullanılan ortamlardaki elektrikli cihazları, yapılan çalışmalar açısından büyük bir öneme sahiptir. Elektrikli cihazların harcamış oldukları güçlerin karakteristik yapısı ve enerji gibi parametreler kullanılarak yapılan sınıflandırmalar ile ilgili son dönemlerde birçok çalışmanın yapıldığını görmekteyiz. Bu alanda yapılan literatür çalışmasında çeşitli yöntemlerin performansını değerlendirmek için karşılaştırmalar yapılmıştır.

Akıllı ev sistemi uygulamaların da enerjinin tüketim, üretim ve dağıtım gibi süreçlerin bütününde kontrol edilebilir, gözlemlenebilir ve dolayısı ile yönetilebilir olması çalışma için istenilen ve hedeflenen bir durumdur [19]. Akıllı konut sistemlerinde sistem-kullanıcı arasındaki etkileşimi arttırabilmek ve doğru tahminlerde bulunabilmek için çeşitli sistemler tasarlanmaktadır. Bunlardan biri akıllı sayaçlar denilen sistemlerdir. Akıllı sayaç teknolojisindeki ilerlemeler, enerji hizmetleri ve veriye dayalı çalışmalar için yeni modeller ve sistemler geliştirilmektedir.

Bu amaçla çeşitli araştırmalar yapılmakta, yöntemler denenmektedir. Makine öğrenmesi, yapay zekâ, derin öğrenme gibi yazılım tabanlı çalışmalar yapılmakta ve bu alandaki gelişmeler gün geçtikçe fazla olmakla birlikte önemini de arttırmaktadır. Bu çalışmaların ana temasında veri izleme, analiz etme, tanımla ve gruplandırma başlıkları bulunmaktadır. Tüm bu aşamaları izleyerek sağlıklı bir veri eğitime işlemi gerçekleştirme mümkün olmaktadır. Öncelikle verilerin toplanması ve basit bir şekilde gruplandırılması gerekmektedir. Daha sonra çalışmanın en önemli noktası olan verilerin doğru izlenebilmesi ve yorumlanabilir olmasıdır. Bunun için analizi kolaylaştıran yöntemlerden biri olan MOYİ yöntemi tercih edilmektedir. Büyük ve toplu verilerin analizi yapılırken bir cihaz üzerinden elektrik tüketimini azaltmak için test ölçümleri yapılır ve tek bir noktadan izlenir. Bu çözümlerin başında derin öğrenme sistemleri gelmektedir.

MOYİ, kullanıcıların ana sayacından okunan toplam tüketim verisinin çeşitli sinyal işleme yöntemleri ile analiz edilerek, cihaz bazında tüketim verilerinin elde edilmesini sağlayan bir yaklaşımdır. Aktif olarak çalışan yüklerin anlık olarak tanınması ve tükettikleri enerjinin tahmini derin öğrenme modeli ile gerçekleştirilmiştir. Bu değerlerden aktif gücün tepkisini sisteme işleyerek ölçülmek istenen cihazın gerçek zamanlı tüketim değerlerine ulaşılabilmektedir.

Yapılan ve bahsedilen araştırmaların birçoğunda, MOYİ yöntemini kullanarak daha yüksek verimliliğe sahip sistemler tasarlanmış ve bu alanda ev aletlerinin sınıflandırılmasına büyük önem verilmiştir. Araştırmacılar, sınıflandırma yaparken öznitelik çıkarımında akım, gerilim ve güç sinyallerinin zaman ve frekans parametrelerini kullanmışlardır. Bu elektriksel parametreleri geleneksel, gelişmiş ve frekans alanı analizine dayalı sinyaller olarak üç gruba ayrılabiliriz.

Hui vd. [11], zaman serisi özelliklerine dayalı hibrit bir model önermişlerdir. Hibrit model ile geleneksel özellik çıkarma prensiplerinden farklı olarak PLAID veri seti kullanarak daha iyi performans sağladığını gözlemlemişlerdir. Frekans alanı analizine dayalı gerçek güç, reaktif güç ve akım elektriksel parametreleri iyileştirilerek faz bilgileri tahmin edilmiş ve dikkate alınmıştır. Çalışmada sinyallerin doğrusal olmayan eğrilerin 6 farklı görüntüsü elde edilmiştir. ESA ile doğrudan görüntü sınıflandırma algoritması kullanılmıştır. Sınıflandırıcı için öznitelik olarak dört test yüküne dayalı sinyallerin deneysel sonuçları test edilmiştir. DVM ve Gauss yaklaşımı ile hibrit bir sınıflandırma tasarlamışlardır. PLAID veri setine dayalı olarak k-NN yaklaşımı ile öznitelik çıkarımında bulunmuşlardır. Cihazları sınıflandırmak için ESA ve LSTM gibi derin öğrenme yöntemleri de kullanılmışlardır. Karar ağacı yöntemi baz alınarak verilerin %63'ünü eğitim ve %37'sini de her bir veri seti hatalarını tahmin etmek için ayırmışlardır. PLAID veri seti kullanarak 56 hanenin saha ölçümünü yaparak 11 farklı cihazdan 1793 tane örnek almışlardır. Her bir örnek için akım ve gerilim değerleri 30 KHz'de örneklendirmişlerdir. AÖM sınıflandırma yöntemleri kullanılarak testler gerçekleştirmişlerdir. Testlerde doğruluk ve F1 ölçüsü sırasıyla 0,8070 ve 0,7871 iken, daha önceki çalışmalarda bu oranlar sırasıyla 0,7323 ve 0,6766'dır. Sonuç olarak, doğrulukta %10,20'lik bir iyileşmeye ve F1 ölçüsünde %16,33'lük bir iyileşmeye karşılık geldiğini belirlemişlerdir. Veri setinde klima, vantilatör,

buzdolabı, saç kurutma makinesi, ısıtıcı, mikrodalga fırın, elektrik süpürgesi ve çamaşır makinesi gibi 8 farklı cihaz bulunmaktadır.

Christoforos vd. [20], konuttaki her bir cihazın güç taleplerini her birinden örnek sinyaller olarak çok etiketli bir ayrışma yöntemi ile fiili tahminlerde bulunmaktadır. İlk olarak özellik belirlemek için, yeterli boyutta veri indirilmesi yaparak daha az bilgi gücü ile algoritmaların karmaşıklığı azaltılmıştır. İkinci adım olarak basit ayrıştırma özelliği kullanmışlardır. Daha sonra çok etiketli sınıflandırma yöntemi kullanarak birden fazla cihazın kolaylıkla tanımlanmasını sağlamışlardır. Her bir model için ayrı ayrı hesaplama yapmanın önüne geçmişlerdir. Fakat MOYİ yönteminin tek etiketli sistemin sonuçlarında modellerde çakışmalar olabileceğini dezavantaj olarak nitelendirmişlerdir. Uygulamada Signal2Vec ile sistematik bir karşılaştırma yöntemleri olan, parçalı ayrıştırma yöntemi, sembolik ayrıştırma yöntemi, AFD, zaman serisi sınıflandırması, sembolik Fourier yaklaşımı gibi yöntemler ile çok etiketli sınıflandırıcılar tarafından mevcut yapılandırma ile dizi karşılaştırmaları yapılarak ileri beslemeli sinir ağı kullanılmıştır. MOYİ yönteminde zaman serisi gösterimlerinden yedisini kullanmışlardır. Bunlardan parçalı ayrıştırma yöntemi kullanılarak zaman serisini bir dizi işlemde oldukça gürültülü verilerde basit bir şekilde kolay dönüştürdüğünü gözlemlemişlerdir. Sembolik ayrıştırma yönteminin kümeleme, sınıflandırma, özetleme gibi zaman serileri ile ilgili birçok alanda uygulanan başarılı bir yöntem olduğuna vurgu yaparak mevcut verileri bu yöntem kullanılarak bir dizi işlem ve denklemler kullanılarak ortalama bir değer üzerinden yaklaşımda bulunmuşlardır. Bu yöntem ile zaman serisi değerlerini sembolize ederek her bir verinin ortalama değerini alarak lineer regresyonunu hesaplamışlardır. Verilerin ortalama değerlerinin istatistiksel özelliklerini Gauss dağılımı oluşturarak orijinalde olduğu gibi elde edebilmişlerdir. Yaptıkları diğer yöntem olan Sinyal-Vektör dönüşüm algoritmasını kullanarak verileri kümeler halinde herhangi bir zaman serisine eşitlemek için fonksiyonu eğitmişlerdir. Yalnızca bir defa eğittikleri veriyi k-NN komşuluk ile sınıflandırmışlardır. Kullanmış oldukları bu zaman serilerinde UK-DALE ve REDD veri setlerini kullanarak yukarıda bahsedilen yaklaşımlar ile 1 günden 10 güne kadar ve belirli aralıklarla testler yapmışlardır. Bu yaklaşımlarda 30 dakika içinde doğrusal artımlar gözlemleyerek hızlı ve iyi bir performans gösterdiğini gözlemlemişlerdir. Sonuç olarak, MOYİ yöntemi ile çoklu veri kümelerinde performans üstünlüğü sağladığını ve ayrıca

Sinyal-Vektör dönüşüm algoritması ile veri boyutu azaltarak daha yüksek performans elde etmişlerdir.

Jessie vd. [21], dalgacık tabanlı MOYİ yöntemi ile Procrustes analizi kullanarak dalgacık fonksiyonlarında filtrelere uygulamıştır. Verilerin karşılaştırılması sonucu MOYİ tabanlı sistemin yüksek tahmin doğruluğuna sahip olduğunu gözlemlediler. Verilerin ham bilgilerini kullanarak gerilim ve akım dalga biçimlerini KSDF ve geçici sinyal özellikleri çıkarımı yaparak kendi içlerinde sınıflandırmışlardır. Sinyallerin özniteliklerini çıkardıktan sonra makine öğrenmesi teknikleri kullanarak mevcut gürültülü sinyalleri sınıflandırmışlardır. K-NN sınıflandırıcısı yöntemi ile eğittikleri verileri bellekte saklayarak yalnızca sınıflandırma gerektiği durumda kümeler halinde böldüğü ve basit ayrıştırma yaparak öz nitelik çıkarımına yardımcı olduğunu belirtmişlerdir. Yaklaşık 441 dalgacık sinyali, filtre uzunluklarını temsil (örneğin, uzunluk-6 filtresi için 3) eden 91 tane kümeyle ayrıştırarak veri boyutunu önemli ölçüde azaltarak daha basit bir sınıflandırma yapılmasını sağlamışlardır. Önerilen algoritmanın etkinliğini test etmek için deneysel bir test sistemi tasarlamışlardır. Akkor ampul, kompakt floresan ampul, bilgisayar ve şarj cihazından oluşan dört tane yükü programlanabilir alternatif bir kaynaktan beslemişlerdir. Akım ve güç verileri toplanarak ayrık dönüşüm algoritması uygulanmıştır. Doğruluklarını eğitmek ve verileri test etmek için veri seti oluşturmuşlardır. Toplamda 32 (8×4) farklı olası kombinasyon oluşturmuşlardır. Yukarıda almış oldukları dört farklı yükten 32 kombinasyon ve 5 anahtar ile 7 farklı voltaj değerini veri setinde 1.120 farklı kayıt yapmışlardır. Veri kümesini parçalarına ayırarak çapraz doğrulama ile test ettikleri setin %90'ını yani yaklaşık 1008 veri kaydını eğitim için %10'unu da test için kullanmışlardır. 6 farklı dalga seti ve 91 farklı kümeleme ile 1120 sinyalden oluşan veri seti analizi gerçekleştirmişlerdir. Dört farklı yük için ayrı ayrı karşılaştırma gerçekleştirmişlerdir. Genel sınıflandırma doğruluklarını karşılaştırdıklarında %99,5 oranında MOYİ tabanlı sistemin yüksek tahmin doğruluğuna sahip olduğunu gözlemlemişlerdir.

Mubarok vd. [22] çalışmalarında, dört farklı doğrusal olmayan yüklere ait harmonik verilerin radyal bazlı fonksiyon özelliklerini kullandılar. Veriler 15 harmonik sıraya göre mertebelenirilmiş ve her verinin 9 farklı öznitelik içerdiği sonucuna varmışlardır. Her bir kombinasyon için deneysel çalışmalarda 1500 farklı örnek veri kullanmışlardır. Radyal bazlı sınıflandırma fonksiyonunda eğitim ve test aşaması için giriş verisi olarak dört

doğrusal olmayan yükün kombinasyonunu belirlemişlerdir. Önerdikleri algoritmanın performansını doğrulamak için dört yükün 15 farklı kombinasyonunu test etmişlerdir. Her bir kombinasyon için, yük tanımlaması 0,5 ile 0,8 hız olmak üzere iki değerle kullanılarak gerçekleştirmişlerdir. Mevcut dalga biçiminden harmonik yük verilerinin öznelikleri çıkarmak için HFD ve yük harmoniklerini belirlemek için de bir RTFSA sınıflandırıcı yönteminin birlikte kullanımını içeren bir yöntem önermiştir.

Da Yan vd. [23], Bayes sınıflandırma modeli yöntemi ile elektrikli ev cihazlarının yük eğrisi özelliklerini barındıran yeni bir sınıflandırma algoritması önerdiler. Pekin’de 15 farklı konuttaki yedi tip ev cihazın (buzdolabı, elektrikli ocak, klima, televizyon, dizüstü bilgisayar, çamaşır makinesi ve su sebili) verilerini analiz elde etmişlerdir. Bağımsız ve yüksek hassasiyetli akım örnekleme elde etmek için cihazlarda doğru akım sinyali yakalamaya çalışmışlardır. Bu algoritmanın ortalama doğruluk oranının %80 ile %90 arasında olduğunu belirtmişlerdir. Kullanmış oldukları veri setindeki her bir cihazın 1440 örneğinin her birinin bir dakikadaki değeri temsil ettiğini belirtmişlerdir. Test verilerini Bayes sınıflandırmasının etkilerini test etmek için kullandıkları verilerden oluşturdukları veri setinin ilk satırı cihazın kategorisini, ikinci satırının da özellik vektörünün uzunluğunu temsil ettiğini belirtmişlerdir. Kullandıkları ev aletlerinin özellik eşleşmesini çıkarırken belli zaman aralıkları (11:00–3:00 ve 16:30–19:00) seçmişlerdir. 14 gün boyunca ölçüm verilerinin analizi sayesinde, Bayes temelli elektrikli ev cihazların sınıflandırmasında %95,6 doğruluk oranında ulaşmışlardır.

Leen De vd. [24], cihazların gerilim ve akım değerlerin analiz etmek ve örneklem veri kümeleri oluşturarak işaretlerini tanımlamak için ağırlıklı piksel VI görüntülerini kullandılar ve akım-gerilim görüntüleri, cihaz sınıflandırması için bir ESA’ya uyguladılar. Ayrıca önermiş oldukları yöntemi WHITED ve PLAID veri setleri ile test etmişlerdir. Oluşturdukları veri setlerini oluşturup öznelik çıkardıktan sonra DVM, sinir ağları, karar ağaçları gibi farklı sınıflandırma yöntemleri kullanarak eğitim verileri oluşturmuşlardır. Denetimli ve Denetimsiz MOYİ yöntemi kullanarak evdeki cihazları test eden sınıflandırma algoritması oluşturmuşlardır. Kullanmış oldukları PLAID veri setinde akım ve voltaj değerlerini içeren 11 farklı cihaz tipinin 30 KHz’de ölçümleri yapılmış ve 1074 ölçüm gerçekleştirmişlerdir. WHITED data setinde ise 46 farklı cihaz için 44 KHz’de ölçümler gerçekleştirerek 1100 ölçüm sonucuna elde etmişlerdir. Eğitim ve test veri seti

için 22 farklı cihaz tipi tercih etmelerinden dolayı 860 farklı ölçüm sonucu elde etmişlerdir. PLAID ve WHITED üzerinden kıyaslama yapılarak, PLAID veri setindeki cihazların %39 doğru, %1 hatalı ve %60'ı tanımsız olarak sınıflandırılırken, WHITED veri setindeki cihazların %73 doğru, %6 hatalı ve %20'si tanımsız olarak sınıflandırılmıştır. Her iki veri setinde 0,5 KHz kadar düşük frekanslar için performansın değişmediğini göstermişlerdir.

Georgios [25], elektrikli ev cihazlarının günlük akım, gerilim ve güç değerlerini analiz etmek için laboratuvar ölçümleri ve gerçek veriler ile birlikte cihaz yük verilerini toplamak için yüksek çözünürlüklü bir veri seti oluşturdu. Konutlarda bulunan cihazlar arasındaki iletişimin hızlı ve doğru sağlanabilmesi için saha testleri gerçekleştirmişlerdir. İletişim kanallarında saatte tipik 20 ile 30 cihaz iletişim seviyesi yüksek bir şekilde olduğu belirtilmiştir.

Buddhahai vd. [26] çok etiketli sınıflandırma yöntemi kullanarak konutlardaki yük dağılımı makine öğrenmesi yöntemleri ile analizini önerdiler. Önerilen yöntem doğru model parametre yöntemleri ile uygun değer performansı için, farklı cihazların yük imzası kombinasyonları seçilmiş ve farklı çok etiketli sınıflandırma algoritmaları ile karşılaştırılmıştır. Veriler Bangkok, Tayland'da bulunan bir evden alınmıştır. MOYİ sistemini, veri toplama, öznelik çıkarma, veri öğrenme ve tanımlama olarak üç temel gruba ayırmışlardır. Denetimli ve denetimsiz yöntem olarak kategorize ederek hedef konumdaki cihazların verilerini denetlemişlerdir. Cihaz modelini toplu olarak doğrudan verileri belirlemek için GMM oluşturarak güç tüketimi hesaplamışlardır. Uygun performansta makine öğrenmesi aşamaları olarak, özellik seçimi, öğrenme algoritması seçimi ve öğrenme analiz eğrisi ve model seçimi olarak ayırmışlardır. 18 farklı kanaldan 10 tanesinin bulunduğu her bir elektrik kanalı için cihazların gruplandırılmış. Cihazdan akım, gerilim, aktif güç, reaktif güç ve güç faktörü gibi beş farklı elektrik parametresi toplamışlardır. Test örnekleri için uygun miktarda eğitim verisinin ne kadar olduğunu değerlendirmek için, doğrulama ve test seti için her biri iki haftalık döneme ait 20.160 veri belirlemişlerdir. Eğitim verilerini, her biri 10.000 örneğe eşit olacak şekilde bölerek veya yaklaşık bir haftalık bir veri periyodu ve değerlendirmeyi, eğitim sayısına göre oluşturulan modeli uygulamışlardır. Eğitim verisi sayısında 130.000 örnek alarak, doğrulama eğrisi puanının maksimum 0.725 seviyesine ulaştığını görmüşlerdir. Çok etiketli sınıflandırma yaklaşımının MOYİ uygulamasına uygun olduğu ve en iyi tahmin performansına sahip

olduđu kanısına varmışlardır. Yüksek güçlü cihazlarda (şofben ve klima) sırasıyla %97 ve %99 performans elde edebildiğini, daha düşük güçlü cihazlarda (akkor flemalı ampul gibi) sırasıyla %59 ve %93 performans elde etmişlerdir. Fiş çıkışı için ortalama koşul ayarlarında tahmin performansı için %75 ve %92 elde etmişlerdir. Karar ağacı tabanlı sınıflandırıcı için öğrenme yöntemi burada etiketlenmemiş veriler, mevcut etiketli verilerin sınıflandırma performansını iyileştirmek için kendi kendine eğitim için kullandığını belirterek veri açıklama süreci kısaltılabileceğini vurgulamışlardır.

Yine bu yapılan çalışmalar birçok farklı alanlarda kullanıldığı, yük izleme yöntemleri ile birlikte veri setleri sınıflandırılarak, sinir ağlarının görsele dönüştürölüp verilerin görsellerinin elde edilmesi mümkün kılınmıştır. Bu sınıflandırma ile eğitilen veriler ile doğruluklarında yüksek oranda tahmin yapılabilmektedir.

3. MATERYAL VE YÖNTEM

Bu bölümde tez çalışmalarında kullanılan veri seti detaylandırılacaktır. Ayrıca bu veri setindeki elektrik ev cihazlarının zaman serisi özellikleri kullanılarak cihazların tespitinde kullanılan öznitelik çıkarım yöntemleri ve sınıflandırma algoritmalarına yer verilecektir.

3.1. Veri Setlerinin Analiz Edilmesi

Binaların enerji tüketimi çalışmalarının çoğu için kapsamlı enerji tüketim verilerinin elde edilmesi zorlu bir iştir. Ayrıca, ortaya çıkan veri setlerinin doğruluğu ve güvenilirliği, araştırma sonuçlarının kalitesinin belirlenmesinde önemli bir rol oynamaktadır. Bu bakımdan binaların enerji tüketimindeki birçok problem makine öğrenmesi ve veri madenciliği tekniklerinden faydalanabilir.

Genel olarak, enerji tüketimi veri analizi ile ilgili araştırma faaliyetleri ya kamuya açık veri setlerine ya da özel veri setlerine dayanmaktadır. İkincisi her zaman mümkün değildir çünkü gerçek dünya verilerini toplamak zaman alıcı, maliyetli ve bazı durumlarda zordur [27]. Bu nedenle, birçok araştırmacı kamuya açık enerji tüketimi veri setlerini kullanmayı tercih etmektedir. Bu bölümde araştırmacıların bilimsel çalışmalarda yaygın olarak kullandıkları UK-DALE, REDD ve PLAID gibi bazı veri setleri ile tez çalışmasında kullanılan veri seti detaylandırılacaktır.

3.1.1. UK-DALE Veri Seti

Ayrıştırma algoritmaları hakkında araştırma yapmak için araştırmacılar, yalnızca bina başına düşen toplam talebi değil, aynı zamanda bireysel cihazların talebini açıklayan verilere ihtiyaç duyarlar [27]. Bu bağlamda, UK-DALE veri seti, İngiltere’de beş ayrı konuttan buzdolabı, bilgisayar, bulaşık makinası, çamaşır makinası ve su ısıtıcılarından alınan verileri içermektedir. Veri kaydında her altı saniyede bir konutlardaki cihazların güç kaydı ayrıca üç evin gerilim ve akım değerleri 16KHz örneklem frekansı ile kaydedilmiştir. Veriler 655 günlük bir süreci kapsamaktadır [28].

3.1.2. REDD Veri Seti

Bu veri seti, ABD metropollerinde 6 konuttan konuttaki buzdolabı, bulaşık makinası, çamaşır makinası ve mikrodalga fırınından alınan güç verilerini içermektedir. Veriler 18 aylık bir süreci kapsamaktadır. Veriler 16KHz'lik örneklem frekansına sahiptir. Her 3 saniyede bir güç ölçümü yapılmıştır. Ayrıca her bir cihazın gerilim ve akım dalga formları kaydedilmiştir [29].

3.1.3. PLAID Veri Seti

PLAID (Plug Load Appliance Identification Dataset), yük tanımlama araştırması için tasarlanmış ve iki sürümüne sahiptir. PLAID'in ilk sürümü 2013 yazında toplanan PLAID1 olarak adlandırılmıştır. PLAID'in ikinci bir sürümü 2014 kışında toplanmış ve PLAID2 adı altında yayınlanmıştır. Veri seti Pittsburgh, Pennsylvania, ABD'de 56'dan fazla evde bulunan 11 farklı cihaz tipinden 30 KHz örneklem frekansındaki akım ve gerilim ölçümlerini içermektedir [30]. Her cihaz tipi, farklı marka/modellerin düzinelerce farklı örneğiyle temsil edilmektedir. Her cihaz için, her durum geçişi için üç ila altı arasında ölçüm verisi alınmıştır.

3.1.4. Ev Aletlerinin Yüksek Çözünürlüklü Profilleri

Bu tez çalışmasında, yeni hazırlanmış ev aletlerinin yüksek çözünürlüklü profillerini içermektedir [31]. Detayları Tablo 1'de verilen veri seti, Avrupa'da konutlarda yaygın olarak kullanılan 18 farklı elektrikli cihazının 1Hz'lik örneklem frekansındaki aktif ve reaktif güç verilerini içermektedir. Çamaşır makinesi, ekmek kızartma makinesi, ütü, saç kurutma makinesi ve klima gibi elektrikli cihazların aktif güç tüketimi verileri Aralık 2018 ve Ocak 2019 dönemlerine ait iki aylık olarak ölçülmüştür. Ayrıca konutlarda ölçülen verilerin doğruluğu laboratuvar ölçümleri ile teyit edilmiştir. Isı pompasının farklı çalışma konumlarına ait aktif ve reaktif güç tüketimleri ile elektrikli bisiklet ve akülü aracın şarj tüketimleri Avrupa Komisyonu Ortak Araştırma Merkezi'nin (JRC) akıllı şebeke birlikte çalışabilirlik laboratuvarındaki (SGILab) ölçümlerinden alınmıştır [32]. Ayrıca toplanan veriler, uzaktan erişim ve işlemeyi sağlamak için bulut tabanlı bir platforma aktarılmış ve saklanmaktadır.

Tablo 3.1. Veri setindeki elektrikli cihazlar ve özellikleri

No	Elektrikli Cihazlar	Güç (W)	Güç Faktörü	Örnek Sayısı
1	Klima	1,400	0,960	120
2	Test cihazı (Test)			6
3	Akülü Araç	6,600		120
4	Bulaşık Makinası	2,100	0,635	24
5	Kurutucu	2,300	1,000	6
6	Buzdolabı	170	0,995	168
7	Saç Kurutma	2,100	1,000	44
8	Isı pompası	2,600		20
9	Ütü	2,200	1,000	12
10	Led Lamba	60	0,900	8
11	Flüorsan Lamba	200	1,000	8
12	Bilgisayar	500	1,000	78
13	Fırın	2,200	0,965	72
14	Televizyon	108	1,000	82
15	Tost Makinası	700	1,000	48
16	Su Isıtıcısı	400	1,000	12
17	Çamaşır Makinası	2,300	0,570	28
18	Elektrikli Bisiklet	200		60

3.2. Öznitelik Çıkarım Yöntemleri

Ev aletlerinin güç tüketimlerine ait zaman serisi sinyallerinden tespit edilmesindeki önemli adım, öznitelik çıkarımıdır. Öznitelik, ham verilerden tanınabilir bir ölçüyü temsil eden ayırt edici bir nitelik olarak ifade edilebilir. Öznitelik çıkarma sürecinde, zaman serisi verilerinden minimum bilgi kaybıyla bir dizi ilgili ve önemli öznitelik vektörü çıkarmak çok önemlidir. Sonuç olarak, bu süreç, daha basit ve daha hızlı hesaplama hızı sağlayarak analizi kolaylaştırır. Bu tez kapsamında yapılan çalışmalarda öznitelik çıkarımında kayan pencere yaklaşımı ile ortalama, standart sapma, medyan ve çok ölçekli dağılım entropi, kısa süreli fourier dönüşüm, evrişimsel sinir ağları ve uzun kısa süreli bellek ağ yöntemleri kullanılmıştır.

3.2.1. İstatistiksel Yöntemler

İstatistiksel analiz yöntemleri sistemli bir şekilde sayısal verilerin toplanmasıdır. Bağımlı ve bağımsız değişkenler arasındaki ilişkiyi çözümler. Sinyali en iyi temsil eden parametrelerin elde edilmesinde zaman tabanlı özellikleri kullanılarak öznelik çıkarma yöntemleri kullanılmaktadır. Sinyal işlemede, genellikle ortalama, ortanca ve standart sapma gibi temel istatistiksel yöntemlerin kullanılması neredeyse bir zorunluluktur.

Bu çalışmamızda, zaman serisi sinyallerinden öznelik elde etmek için betimsel analiz yöntemlerinden olan; ortalama, standart sapma ve medyan olmak üzere üç istatistiksel parametre dikkate alınmaktadır. Ortalama, standart sapma ve medyan, kabul edilebilir bir kesinlik derecesi olarak alabileceğimiz ve hesaplamalar doğrultusunda dağılımı doğru bir şekilde karakterize edebilen bir veri setinin en önemli üç ölçümü oldukları için seçilmiştir [33,34].

Medyan, bir sayısal veri serisi küçükten büyüğe sıralandığında ortadaki veri değerinin sayısal değeridir. Puanlar büyüklük sırasına göre dizildiğinde puan dağılımının iki eşit parçaya ayıran değer olarak tanımlanır ve N serideki veri sayısı olmak üzere matematiksel bağıntısı Denklem 2.1’de görülmektedir.

$$\begin{aligned} Med &= \frac{N+1}{2}, & N &= \text{Tek Sayı} \\ Med &= \frac{\left(\frac{N}{2}\right) + \left(\frac{N}{2} + 1\right)}{2}, & N &= \text{Çift Sayı} \end{aligned} \quad (2.1)$$

Medyan değeri, örneklem veri kümesinin nasıl dağıtıldığına ilişkin ölçülerle bir araya getirildiğinde ve veri serilerinin geneli hakkında iyi bir saptama yapmamızı sağlar.

Aritmetik ortalama, bir sayı serisindeki sayıların toplamının serinin eleman sayısına (sayı adedine) bölünmesi sonucu elde edilen değerdir [34]. Toplam puanların puan sayısına bölünmesiyle elde edilen aritmetik ortalama, puan dağılımındaki her puanın hesaplamaya dâhil edilmesi nedeniyle diğer merkezi yığılma ölçülerine göre daha yeterli bir istatistiktir [33,34]. Aritmetik ortalamaya ait matematiksel bağıntı Denklem 2.2’de görülmektedir.

$$\mu = \frac{1}{N} \sum_{i=1}^N x_i \quad (2.2)$$

Standart sapma, bir serisindeki sayıların, serinin aritmetik ortalamasından farklarının karelerinin toplamının dizinin eleman sayısının bir eksiğine bölümünün kareköküdür.

Veri setimizin standart sapma değerini bulabilmek için aşağıdaki adımları sırasıyla uygulamamız gerekmektedir [33,34].

- Veri setinin ortalama değeri Denklem 2.2 bağıntısı ile bulunur.
- Her bir veri noktası ile ortalama değer arasındaki uzaklık hesaplanır.
- Bulunan uzaklık değerlerinin kareleri alınır.
- Farkların kareleri toplanır.
- Elde edilen toplam, serinin eleman sayısının bir eksiğine bölünür.
- Bulunan sayının karekökü alınır.

Buna göre standart sapma değeri Denklem 2.3'deki matematiksel bağıntı ile tanımlanabilir.

$$\sigma = \sqrt{\frac{1}{N} \sum_{i=1}^N (x_i - \mu)^2} \quad (2.3)$$

Standart sapma, veri dağılımı hakkında, özellikle de tüm veri kümesinin dağılımın ortalama değerine ne kadar yakın olduğu hakkında bizi bilgilendirir. Büyük bir standart sapmaya sahip bir veri kümesindeki veriler, geniş bir değer aralığına genişletilir. Buna karşılık, küçük bir standart sapmaya sahip bir veri kümesi, sıkı bir şekilde gruplandırılmış kesin verilere sahiptir. Bütün değerler aynı olursa standart sapma sıfır olur. Standart sapma bize verilerin ne kadar düzgün ve dengeli dağıldığını gösterir.

3.2.2. Entropi

Entropi kavramı, ilk defa 1865 yılında Rudolf Clausius tarafından öne sürülmüş olup bir sistemin düzensizlik ve dağınıklığın bir ölçütü olarak bilinmektedir. Shannon tarafından da 1948'de farklı bir kullanıma kavuşarak enformasyon entropisi halini almıştır [35]. Shannon entropisi, bir kaynağı mümkün olan minimum değerini kullanarak bilgi kaybı olmadan sayısallaştırmak için kullanılır. Entropi yönteminde temel olarak baz aldığımız veri

kümesinin ağırlığının elde edilmesinde başlangıç matrisi yeterlidir. 5 aşamadan oluşmaktadır;

İlk aşamada, Denklem 2.4'deki gibi m satır ve n sütunlu bir karar matrisi oluşturulur.

$$X = \begin{bmatrix} X_{11} & \cdots & X_{1n} \\ \vdots & \ddots & \vdots \\ X_{m1} & \cdots & X_{mn} \end{bmatrix} \quad (2.4)$$

İkinci aşamada, verilerin ortak birimlere çevrilmesi amacıyla fayda özelliklerine göre normalizasyon yapılarak matematiksel bağıntısı Denklem 2.5'de verilen normalize edilmiş karar matrisi oluşturulur.

$$r_{ij} = \frac{X_{ij}}{\sum_{i=1}^i X_{ij}} \quad (2.5)$$

Burada;

i = alternatifler

j = kriter normalize edilmiş değerler

X_{ij} = alternatif ve kriter için faydalı değerler

Üçüncü aşamada, kriterlere ilişkin entropi değerleri Denklem 2.6 ile bulunur.

$$e_j = -k \sum_{j=1}^n r_{ij} \cdot \ln(r_{ij}), \quad i = 1, 2, \dots, m \text{ ve } j = 1, 2, \dots, n \quad (2.6)$$

Burada;

k = entropi katsayısı

r_{ij} = normalize edilmiş değerler

e_j = entropi değeri

Dördüncü aşamada, kriterlere ilişkin aralarındaki uzaklığın veya farklılaşmanın hesaplamışı Denklem 2.7 ile yapılarak bilginin farklılaşma derecesi hesaplanır.

$$d_j = 1 - e_j \quad (2.7)$$

Son olarak beşinci aşamada Denklem 2.8 yardımı ile entropi kriter ağırlığı hesaplanır.

$$w_j = \frac{1-e_j}{\sum_{i=1}^n 1-e_i} \quad (2.8)$$

3.2.2.1. Çok Ölçekli Dağılım Entropi

1948 yılı entropi alanındaki çalışmaların başlangıç yılı olarak kabul edilir. Shannon tarafından 1948'de önerilen entropi, düzensizliği veya belirsizliği ölçmek için kullanılır [36]. Bir veri kaynağının entropisinin hesaplanması, bu veriyi kayıpsız bir şekilde temsil etmek için saklanması gereken minimum bilginin bir ölçüsünü verir. Başka bir deyişle, entropi, durumlarının olabirlik dağılımını kullanarak zaman serilerindeki düzenlilik miktarının ölçülmesi olarak da ifade edilebilir [37]. Çok Ölçekli Dağılım Entropisi (ÇÖDE), iri taneli dağılım entropisine dayanır ve aşağıda açıklanan adımlardan oluşur:

İlk aşamada, $U = \{u_1, u_2, \dots, u_n\}$ şeklinde tanımlanan bir zaman serisi olduğunu varsayarsak bu seri örtüşmeyen τ uzunluğuna bölünür. Daha sonra, τ ölçek faktörüne bağlı olarak, Denklem 2.9'daki gibi iri taneli entropi prosedür gerçekleştirilir.

$$f_j^{(\tau)} = \frac{1}{\tau} \sum_{k=(j-1)\tau+1}^{j\tau} x_{i, 1 \leq j \leq \lfloor \frac{N}{\tau} \rfloor = L} \quad (2.9)$$

İkinci aşamada, $f_j^{(\tau)}$ ilgili etiketlerin 1 ile c arasında olduğu sınıflarına eşlenir. Bu işlem için, iri taneli zaman serilerine normal kümülatif dağılım fonksiyonu gerçekleştirilir ve $f_j^{(\tau)}$ Denklem 2.10 bağıntı ile 0'dan 1'e kadar olan sınıflara atanır.

$$y_j = \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{f_j} e^{-\frac{(f-\mu)^2}{2\sigma^2}} dt \quad (2.10)$$

burada, μ ve σ sırasıyla ortalama değeri ve standart sapmayı temsil eder. Bu aşamada normal kümülatif dağılım fonksiyonu ile minimum ve maksimum değerlerinin f_j zaman serisinin ortalama değeri ile karşılaştırılmasıyla sınıf sayısı düzenlenebilir. Ardından, her y_j 1 ile c arasındaki tamsayı değerlerine bölmek için Denklem 2.11 kullanılır.

$$z_j^c = \text{round}(c \cdot y_j + 0.5) \quad (2.11)$$

Zaman gecikmesi d ve gömme boyutu e göz önüne alındığında, zaman serisi Denklem 2.12'deki yazılabilir.

$$z_i^{e,c} = \{z_i^c, z_i + d^c \dots z_{i+(e-1)d^c}\} \quad (2.12)$$

Son aşamada, her değer vektörünü $z_i^{e,c}$ bir dağılım modeline atayarak;

$z_i^c = v_0, z_{i+d}^c = v_1, z_{i+(e-1)d}^c = v_{e-1}$ potansiyel modellerin orantılı olasılığı $\pi_{v_0 v_1 \dots v_{e-1}}$ Denklem 2.13 yardımı ile hesaplanır.

$$p(\pi_{v_0 \dots v_{e-1}}) = \frac{\text{Number}\{i | i \leq L - (e-1)d, z_i^{e,c} \text{ hastype} \pi_{v_0 \dots v_{e-1}}\}}{L - (e-1)d} \quad (2.13)$$

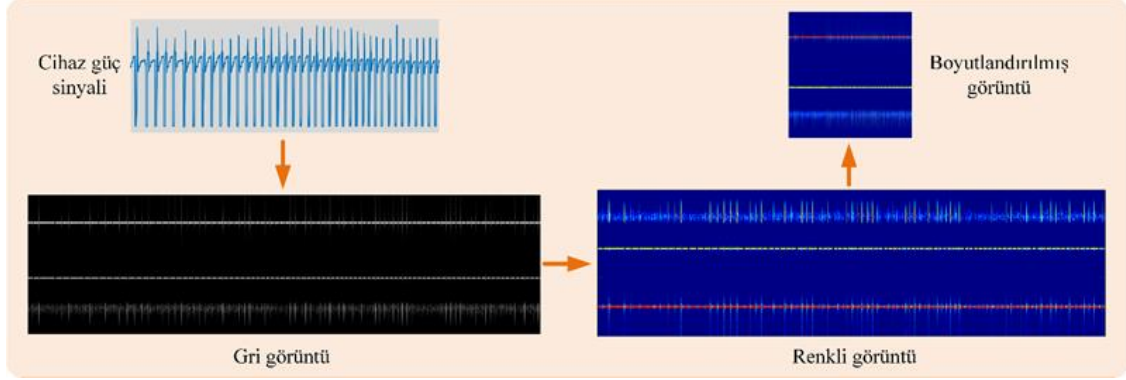
3.2.3. Kısa Süreli Fourier Dönüşümü

Bir zaman serisinin frekans dağılımını analiz etmek için Fourier dönüşümü genellikle en etkili yöntemlerden biri olarak kabul edilmektedir. Fakat bu dönüşümden sonra zaman bilgisi kaybolmaktadır. Sonuç olarak, sinyaldeki ani değişikliklerin belirlenmesinde yetersiz kalmaktadır. Literatürde, hem zaman hem de frekans bilgisinin gerekli olduğu KSFD ve Wigner-Ville dönüşümü gibi çok çeşitli uygulamalar vardır [38]. KSFD, zaman serilerindeki frekans ve faz değişimini ölçmek için Gabor tarafından önerilen bir yöntemdir [39]. Bu yöntem temel olarak tüm zaman serilerini bir kerede düşünmek yerine, kısa zaman aralıklarında (kayan bir pencerede) bir zaman serisinin frekans bileşenlerini temsilini sağlamaktadır [40]. Bu tür temsilleri zaman içinde birleştirmek, bir dalga biçiminin zaman ve frekansın 2 boyutlu bir fonksiyonu olarak görselleştirilmesine izin verir [41].

Verilen bir x zaman serisi ve ilgili STFT gösterimi f göz önüne alındığında Denklem 2.14 ile hesaplanır.

$$(\tau, \omega) = \int_{-\infty}^{+\infty} x(t) \omega(t - \tau) e^{-j\omega t} dt \quad (2.14)$$

burada ω , τ zamanda ortalanmış bir pencere fonksiyonudur. Ayrıca kayan pencereli KSFD temsilinin kare büyüklüğü, $|f(\tau, \omega)|^2$ genel olarak spektrogram olarak adlandırılır. Şekil 3.1’de deneysel çalışmalarda kullanılan bir cihazı güç sinyali zaman serisinden spektrogram görüntünün elde edilme aşamaları görülmektedir.



Şekil 3.1. Zaman serilerinden spektrogram görüntülerin elde edilmesi

3.2.4. Evrişimli Sinir Ağı

Evrişimli sinir ağları, derin öğrenmenin bir alt dalıdır ve genellikle görsel bilginin analizinde kullanılır. Evrişimli Sinir Ağı (ESA), çok boyutlu diziler biçimindeki verileri işlemek için geliştirilen ve birçok işlem katmanından oluşan bir sinir ağı biçimidir [42]. Bu açıdan sınıflandırma da, doğal dil işlemlerinde, görsel nesne analizi ve tanıma gibi birçok farklı alanlarda yoğun olarak tercih edilmekte ve kullanılmaktadır [42]. ESA, klasik sinir ağlarından farklı olarak evrişim öznitelikleri çıkarma ve sınıflama özelliklerine sahiptir [43].

Başka bir yaklaşım ile ESA, bir girdi görüntüsünü alıp görüntüdeki çeşitli görünüşleri/nesneleri birbirinden ayırabilen derin öğrenme algoritmasıdır. ESA, temel olarak görüntüleri sınıflandırmak (örneğin gördüklerini isimlendirmek), benzerlikle kümelemek (fotoğraf arama) ve sahnelerde nesne tanıma yapmak için kullanılan derin yapay sinir ağlarıdır. ESA’lar yapısı gereği giriş olarak görüntü ya da videolar alır. Elbette görüntüleri alırken sayısal formata çevrilmiş olması gerekir. Örneğin bir ESA modeline bir görüntü uygularsak bunu matris formatında vermemiz gerekiyor. Matrise uyguladığımız

filtre sayesinde görüntü üzerinden belli özellikler üzerinde karşılaştırma yapılarak görüntüden veriler elde edilmektedir [44,45].

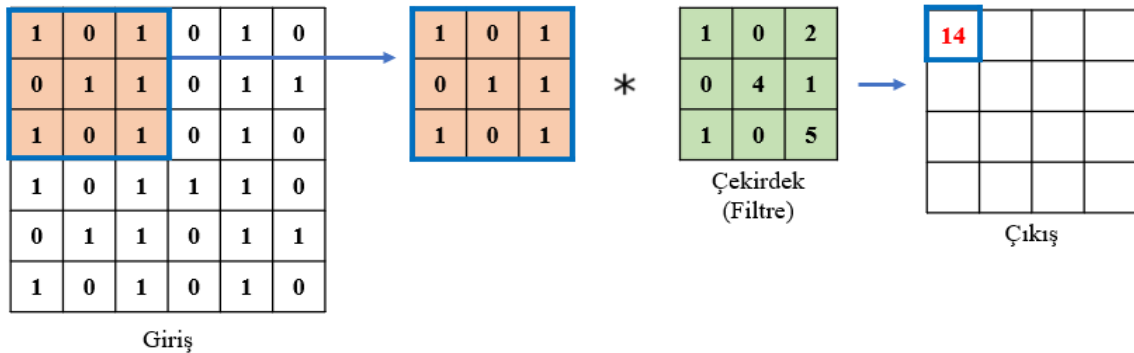
Evrışimsel sinir ağları, Şekil 3.2’de görülen genel yapısında olduğu gibi eğitilebilen birçok katmandan oluşmaktadır.



Şekil 3.2. ESA'nın genel yapısı

ESA tipik bir sinir ağı gibi katmanlar halinde organize edilen ve hiyerarşik temsilleri öğrenebilen nöronlara dayanır [46]. Genel olarak bir ESA modeli evrişim, havuzlama, tam bağlı ve softmax katmanlarından oluşur.

Evrışim Katmanı: ESA mimarisinde en önemli katmandır. Bir dizi evrişimsel filtreden (çekirdekler olarak adlandırılır) oluşur. N boyutlu olarak ifade edilen girdi görüntüsü, çıktı özellik haritasını oluşturmak için Şekil 3.3’de görüldüğü üzere (3×3, 5×5, 7×7 gibi) farklı boyutlardaki filtrelerle konvüle edilir [42].



Şekil 3.3. Konvolüsyon işlemi

Havuzlama Katmanı: Bu katmanın ana görevi, özellik haritalarının alt örneklenmesidir. Bu haritalar, evrişim işlemleri takip edilerek oluşturulur. Başka bir

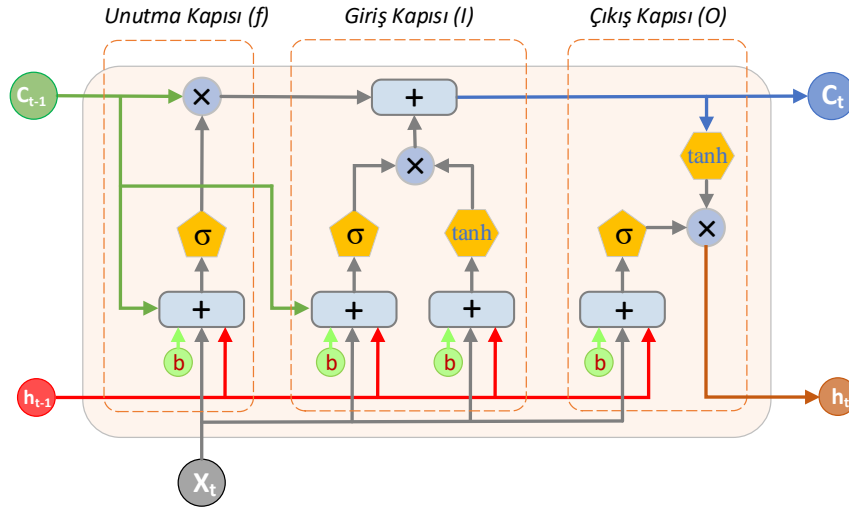
deyişle, bu yaklaşım, daha küçük özellik haritaları oluşturmak için büyük boyutlu özellik haritalarını küçültür. Eşzamanlı olarak, havuzlama aşamasının her adımında baskın bilgilerin (veya özelliklerin) çoğunluğunu korur. Ortalama, minimum ve maksimum havuzlama gibi çeşitli havuzlama yöntemleri mevcuttur [43].

Tam Bağlantılı Katman: Bu katman, genellikle ESA mimarisinin sonunda yer almaktadır. Bu katmanın içinde, her nöron, bir önceki katmanın tüm nöronlarına bağlanır. Bir tür ileri beslemeli yapay sinir ağında olduğu için geleneksel çok katmanlı algılayıcı sinir ağının temel yöntemini izler. Tam bağlantılı katmanın girişi, son havuzlama veya evrişim katmanından gelir. Bu girdi, düzleştirmeden sonra özellik haritalarından oluşturulan bir vektör biçimindedir [47,48].

Softmax: son tam bağlantılı katmandan elde edilen özelliklerin sınıflandırılmasını sağlamaktadır. Sınıflandırma işleminde olasılıksal değerler kullanılarak ilgi sınıfa atama işlemi yapılır [42].

3.2.5. Uzun Kısa Süreli Bellek Ağı

Tekrarlayan sinir ağı mimarisinin özel bir türü olarak Hochreiter ve Schmidhuber tarafından özellikle zaman serisi problemleri için geliştirilen bir derin öğrenme ağıdır [49]. Bu açıdan özellikle doğrusal olmayan zaman serilerinin sınıflandırılması ve tahmin edilmesinde önemli avantajlara sahiptir. Son yıllarda araştırmacılar tarafından tercih edilen LSTM yöntemi doğrusal olmayan karmaşık büyük verilerin tahmin tutarlılığında oldukça yüksek. LSTM ağının önemli avantajlarından bazıları, bilginin uzun süreli geri çağırılması ve geleneksel nöronların gizli katmanlarının yerini alan bellek hücresidir [50,51]. Bu şekilde zaman serisi verilerinde daha etkin öznitelik çıkarımı yapılabilir [52]. LSTM birbirini tekrar eden sıralı bloklardan oluşur. Genel olarak giriş, çıkış ve unut kapıları olmak üzere üç katmandan oluşur. Giriş ve geri besleme arasında gecikmeli bir bağlantı kurulur ve aktif nöronlarda sürekli bir akış hatasına neden olur [47,53]. Ayrıca derin öğrenme ağının bellek kapasitesiyle etkin bir şekilde iyileştirilir. LSTM hücre mimarisi Şekil 3.4 'de gösterilmektedir.



Şekil 3.4. LSTM hücre mimarisi

LSTM ağ mimarisinin ilk adımı olarak, unutma kapısından hangi bilgilerin atılacağını belirlememiz gerekmektedir. Bu süreç Denklem 2.15'deki gibi tanımlanabilir.

$$f_t = \sigma(W_f[h_{t-1}, x_t] + b_f) \quad (2.15)$$

burada, σ sigmoid fonksiyonu, W_f ağırlık matrisleri, h_t gizli durum ve b_f sapma vektörünü tanımlamaktadır.

İkinci adımda, hangi bilgilerin bellek hücresine ekleneceğine karar vermektir [54]. Benzer şekilde, giriş kapısının sigmoid işlevi hangi değerlerin güncelleneceğini belirler ve ardından bir tanh katmanı olası bir güncelleme vektörü üretir. Denklem 2.16 ve 2.17 yardım ile sırasıyla güncellenmiş giriş ve bellek hücre değerleri hesaplanır.

$$i_t = \sigma(W_i[h_{t-1}, x_t] + b_i) \quad (2.16)$$

$$C_t = \tanh(W_c[h_{t-1}, x_t] + b_c) \quad (2.17)$$

Üçüncü adımda, tutulacak ve atılacak bilgilere karar verdikten sonra hücre durumu Denklem 2.18 ile yeniden hesaplanır.

$$h_t = \sigma(W_h[h_{t-1}, x_t] + b_h) \quad (2.18)$$

Ayrıca, bir sigmoid katman, bu aşamada hücre durumundan hangi bilgilerin çıkarılacağına karar verir. Bu güncellemeler Denklem 2.19 ve 2.20 ile yapılır.

$$O_t = \sigma(W_o[h_{t-1}, x_t] + b_o) \quad (2.19)$$

$$h_t = O_t * \tanh(C_t) \quad (2.20)$$

burada σ Denklem 2.21'de verilen sigmoid fonksiyondur ve \tanh Denklem 2.22'de verilen hiperbolik tanjant fonksiyonunu ifade eder.

$$\sigma(x) = \frac{1}{1+e^{-x}} \quad (2.21)$$

$$\tanh = \frac{2}{1+e^{-2x}} \quad (2.22)$$

3.3. Sınıflandırma

Makine öğrenmesi büyük ve karmaşık veri kümelerinin analiz edilmesinde, büyük veri kümelerini temsil eden eğitim verilerini temel alan parametreleri kullanır. Farklı algoritma ve çeşitli yöntemler kullanılarak analizler yapılabilir. Makina öğrenmesi üçe ayrılır.

Denetimli öğrenme: Algoritmalar seçmiş olduğumuz örneklemi temel alarak tahmin yapar. Giriş değerleri ile istenen çıkış değerleri arasında eşleme yapan bir fonksiyon oluşturulur. Eğitim verisi hem girdilerden hem de çıktılardan oluşur. Veri setindeki çıkışlar kategorik, nümerik ise regresyon algoritmaları kullanılır.

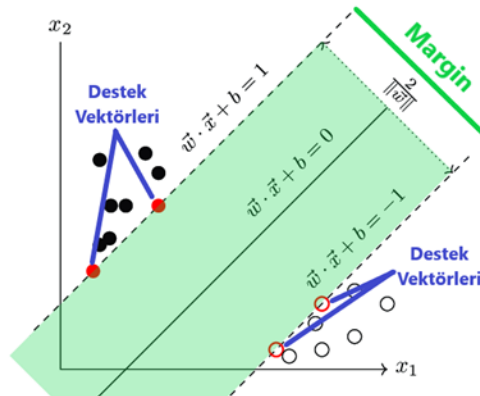
Denetimsiz öğrenme: Veri üzerinden bilinmeyen bir yapıyı tahmin etmek için bir algoritma kullanan makine öğrenmesi tekniğidir. Burada giriş verisinin hangi sınıfa ait olduğu belirsizdir. Denetimsiz öğrenmede veri noktaları etiketlenmez. Algoritma, verileri düzenleyerek veya bunların yapısını açıklayarak veri noktalarını sizin için etiketler. Daha fazla veriyi değerlendirdikçe, bu verilerle ilgili kararlar verebilme becerisi giderek artmakta ve daha rafine olmaktadır.

Pekiştirmeye dayalı öğrenme: Farklı seçenekleri ve olasılıkları kullanarak, hangisinin en uygun olduğunu belirlemek için her bir sonucu izleyerek ve değerlendirmektedir. Makine denemesini ve hatayı öğretir. Geçmiş deneyimlerden öğrenir ve mümkün olan en iyi sonucu elde etmek için duruma cevap olarak yaklaşımını adapte etmeye başlar. Asıl amaç öğrenen etken çevreyle etkileşerek çevreden geri bildirim alıp maksimuma çıkartarak hareket yönü bulmasıdır.

Öznitelik çıkarım aşamasında bahsi geçen yöntemlerden elde edilen öznitelik vektör setinden daha iyi sonuçlar üretebilecek en uygun makine öğrenmesi tekniklerini detaylandırılmıştır. Bu çalışmada, üç popüler sınıflandırma yöntemi olan Destek Vektör Makinesi (DVM), k-En yakın komşular (k-NN) ve topluluk öğrenme-yükseltme (Ensemble Boost Tree (EBT)) ele alınmıştır.

3.3.1. Destek Vektör Makine Sınıflandırıcı

Destek vektör makineleri bir düzlemi ayırarak sınıflandırma yapmamızı sağlar. Destek Vektör Makineleri genellikle sınıflandırma problemlerinde kullanılan gözetimli öğrenme yöntemlerinden biridir [55]. Bir düzlem üzerine yerleştirilmiş noktaları ayırmak için Şekil 3.5’de görüldüğü gibi en iyi hiper düzlemi bulmaya çalışır. Bu hiper düzlemin, iki sınıfının noktaları için de maksimum uzaklıkta olmasını amaçlar. Karmaşık ama küçük ve orta ölçekteki veri setleri için uygundur.



Şekil 3.5. DVM’de en iyi hiper düzlem

W 'nin normal vektörü temsil ettiği ve b 'nin yanlılık anlamına geldiği hiper düzlem, Denklem 2.23'deki gibi tanımlanabilir.

$$f(x) = w^T \cdot x + b \quad (2.23)$$

Hiper düzlem optimizasyonu Denklem 2.24 ile yapılabilir.

$$\begin{cases} \min \left\{ \frac{1}{2} \|w\|^2 \right\} \\ \left\{ \begin{array}{l} y_i - w^T x_i - b \leq \alpha \\ w^T x_i + b - y_i \leq \alpha \end{array} \right. \end{cases} \quad (2.24)$$

burada x , $f(x)$ ile hedef y_i arasındaki sapmayı tanımlar. Denklem 2.25'de, düzenleme parametresi (C) ve gevşek değişken (ξ_i) dikkate alınarak aşağıdaki gibi ifade edilebilir.

$$\begin{cases} \min \left\{ \frac{1}{2} \|w\|^2 + C \times \sum_{i=1}^n (\xi_i + \xi_i^*) \right\} \\ \left\{ \begin{array}{l} y_i - w^T x_i - b \leq \alpha + \xi_i \\ w^T x_i + b - y_i \leq \alpha + \xi_i^* \end{array} \right. \quad \xi_i, \xi_i^* \geq 0 \end{cases} \quad (2.25)$$

Denklem 7 dışbükey olmayan bir formda olduğunda, Denklem 2.26'de Lagrange çarpanı (δ_i) eklenerek optimize edilebilir.

$$f(x) = \sum_{i=1}^n (\delta_i - \delta_i^*) K(x_i, x) + b \quad (2.25)$$

burada $K(x_i, x)$ çekirdek işlevidir. Birçok çekirdek fonksiyonu mevcut olmasına rağmen, genellikle Gauss radyal tabanlı fonksiyon daha çok tercih edilmektedir.

3.3.2. En Yakın Komşu Sınıflandırıcı

K-NN, verilen bir veri noktasına göre komşuları sınıflandıran denetimli bir öğrenme yöntemidir [56]. Ayrıca, regresyon problemlerini ve uygulamaya dayalı sınıflandırma problemlerini çözmek için yaygın olarak kullanılmaktadır [57]. Sınıflandırma k değerine göre yapılır. Test edilecek veriler, eğitim kümesindeki her bir veri ögesine olan uzaklıklar hesaplandıktan sonra sınıfı bilinen en yakın komşuların k 'si olarak seçilir. Daha sonra en

son eğitim verisi olarak belirlenir [58]. Bu süreçte, belirsizliği önlemek için k'nin tek bir sayı olmasına dikkat edilmelidir. Bu mesafeler Euclid, Manhattan, Chebychev ve Minkowski yöntemleri kullanılarak hesaplanabilir. k-NN, eğitim süreci gerektirmemesi, istendiğinde yeni veriler eklenebilmesi ve uygulanmasının kolay olması nedeniyle tercih edilmiştir.

3.3.3. Topluluk Öğrenimi

Topluluk öğrenimi, denetimli öğrenme tabanlı bir makine öğrenimi tekniğidir. Zayıf öğrenen olarak bilinen bir grup sınıflandırıcı oluşturularak sınıflandırma problemlerini çözer [59]. Topluluk öğrenme tekniği tipik olarak iki alt kategoriye ayrılır. Birincisi torbalama olarak bilinir ve Breiman [60] tarafından önerilmiştir, ikincisi ise yükseltme olarak bilinir ve Freund ve Schapire [61] tarafından önerilmiştir.

Yükseltme algoritmasının birincil yaklaşımı, bir grup sınıflandırıcıyı aynı anda eğitmek ve sınıflandırıcının performansını iyileştirmek için model doğruluğunu güncellemektir [61,62]. Bu amaçla, eğitim setinin her bir parçası için ağırlığı koruyacak şekilde tasarlanmıştır. Eğitim setinin her bir bölümünden bir sınıflandırıcı öğrendikten sonra, bir sonraki sınıflandırıcı için ağırlıklar güncellenir [62]. Son olarak, Yükseltilmiş Sınıflandırıcı, sınıflandırıcının performansını iyileştirmek için her sınıflandırıcının tahminlerini birleştirir. Bu modelin dezavantajı gürültülü veriler için uygun olmamasıdır.

Torbalama algoritması, Churn Tahmin Modeli için kullanılan sınıflandırıcının varyansını azaltarak sınıflandırma performansını artırmaktadır [60,62]. Torbalama algoritmasını kullanarak Churn Tahmin Modelinin tasarımında aşağıdaki adımlar: İlk olarak, eğitim veri seti k altkümeye bölünmelidir. Daha sonra model (k-1) alt kümesi kullanılarak eğitilir [62]. Son olarak model yeni alt küme ile test edilir. Bu modelin dezavantajı, sınıflandırıcı çantasından daha doğru olan sınıflandırıcıyı kullanarak test örneklerini öngördüğü için model yapımı için daha fazla hesaplama kaynağı gerektirmesidir.

3.4. Performans Değerlendirme Ölçütleri

Tez çalışmasında gerçekleştirilen deneysel sonuçları analiz etmek doğruluk, duyarlılık, özgüllük, kesinlik ve F1-skor performans değerlendirme ölçütleri kullanılmıştır. Bu ölçütler önerilen deneysel modelin sınıflandırma sonuçlarına göre etiketlenen sınıf değeri ile gerçek sınıf değeri arasındaki ilişkiye göre belirlenir. Bunun için Tablo 2.2’de verilen karmaşıklık matrisinden faydalanılmıştır.

Tablo 3.2. Karmaşıklık matrisi

Karmaşıklık Matrisi		Tahmin Edilen Sınıf	
		Pozitif	Negatif
Gerçekteki Sınıf	Pozitif	Doğru Pozitif (DP)	Yanlış Negatif (YN)
	Negatif	Yanlış Pozitif (YP)	Doğru Negatif (DN)

Performans değerlendirme ölçütleri Denklem 2.26 ve Denklem 2.30 arasındaki bağıntılar ile hesaplanmaktadır.

$$Doğruluk = \frac{DP+DN}{DP+DN+YP+YN} \times 100 \quad (2.26)$$

$$Duyarlılık = \frac{DP}{DP+YN} \times 100 \quad (2.27)$$

$$Özgüllük = \frac{DN}{DN+YP} \times 100 \quad (2.28)$$

$$Kesinlik = \frac{DP}{DP+YP} \times 100 \quad (2.29)$$

$$F1 Skor = 2 \times \frac{Duyarlılık \times Kesinlik}{Duyarlılık + Kesinlik} \times 100 \quad (2.30)$$

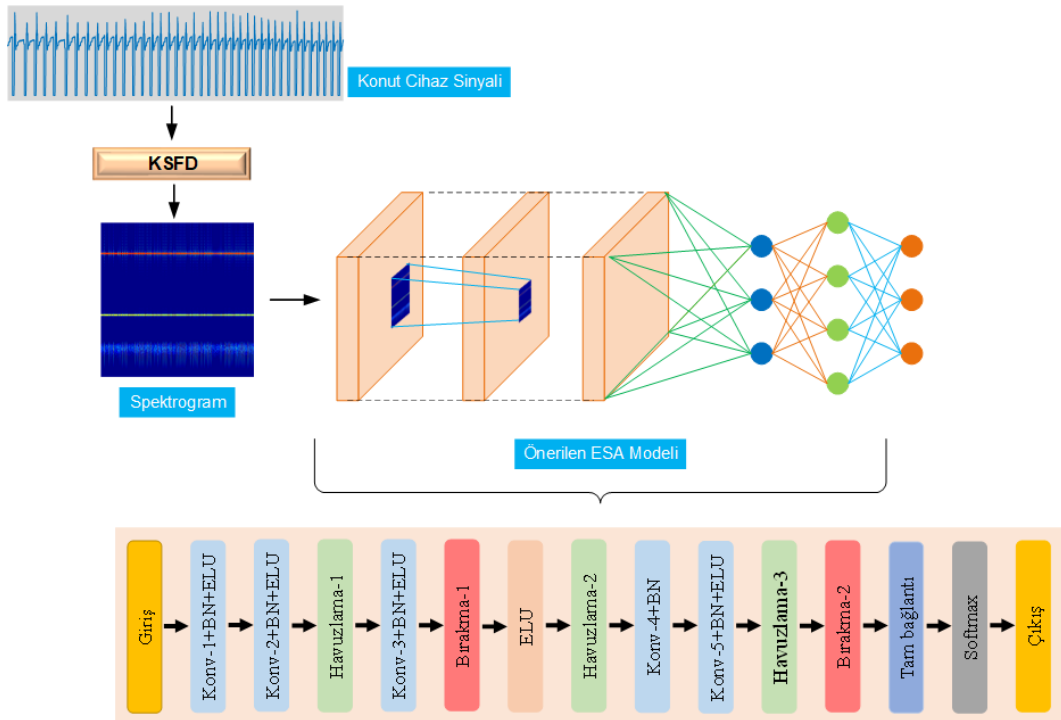
4. BULGULAR VE TARTIŞMA

Bu bölümde, bir önceki bölümde detaylandırılan veri seti, öznelik çıkarım ve sınıflandırma ışığında konutlarda kullanılan elektrikli cihazların tespiti ve cihazların sınıflandırılabilmesi için üç farklı deneysel kurulum ve deneysel sonuçları hakkında bilgiler verilmektedir.

Deneysel çalışmalar, Intel Xeon Silver 3.3 GHz işlemci, 32 GB RAM, NVIDIA P2200 Quadro GPU kartı ve MATLAB 2020a yazılımı ile donatılmış iş istasyonunda gerçekleştirilmiştir.

4.1. Konutlardaki Elektrikli Cihazların Evrişimli Sinir Ağı ile Otomatik Sınıflandırılması

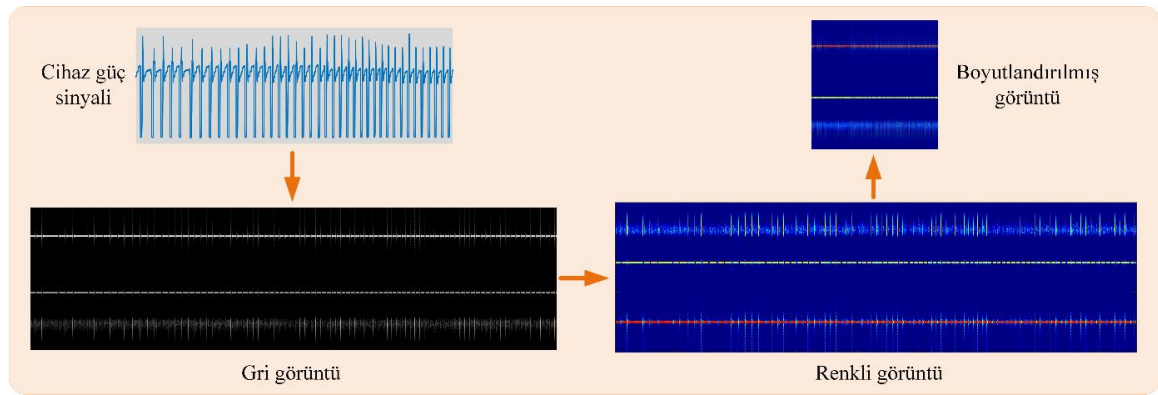
Bu deneysel çalışmada, konutlarda kullanılan elektrik cihazların otomatik tespiti için genel yapısı Şekil 4.1’de verilen yeni bir ESA modeli sunulmaktadır. ESA modelimiz,



Şekil 4.1. Elektrikli Cihazların ESA ile Otomatik Sınıflandırılması Modeli

ESA modelimiz, popüler ve sıklıkla kullanılan ön eğitilmiş ESA mimarileri yerine daha düşük kapasiteli donanımlarda kullanılabilen basit ve daha düşük hesaplama karmaşıklığına sahip etkili bir ESA mimarisi özelliklerine sahiptir. Şekil 4.1’de görüldüğü gibi bu mimaride daha az parametreye sahip konvolüsyon katmanlar ve ReLU fonksiyonu yerine ELU fonksiyonu kullanan aktivasyon katmanlarına sahiptir.

Bu deneysel çalışmanın ilk aşamasında, veri setindeki elektrik cihazlara ait zaman zaman serilerinden spektrogram görüntülerini oluşturmak için deneysel olarak belirlenen 200ms’lik Hamming penceresi kullanıldı ve FFT sayısı 1000 olarak seçilmiştir. Bu değerlere göre 1×86400 boyutlarındaki her bir zaman serisi detayları Şekil 4.2’de görüldüğü gibi önce 251×1727 boyutlarında gri görüntüye dönüştürüldü. Daha sonra gri görüntüdeki piksel parlaklığı ve kontrast değerlerindeki değişimler homojen olarak değişen bir jet128 renk haritası ile renklendirilerek $251 \times 1727 \times 3$ renkli spektrogram görüntüye dönüştürüldü ve son olarak yeni ESA model girişine uygun olarak $224 \times 224 \times 3$ olarak yeniden boyutlandırıldı.



Şekil 4.2. Zaman serilerinden spektrogram görüntülerine dönüştürülmesi

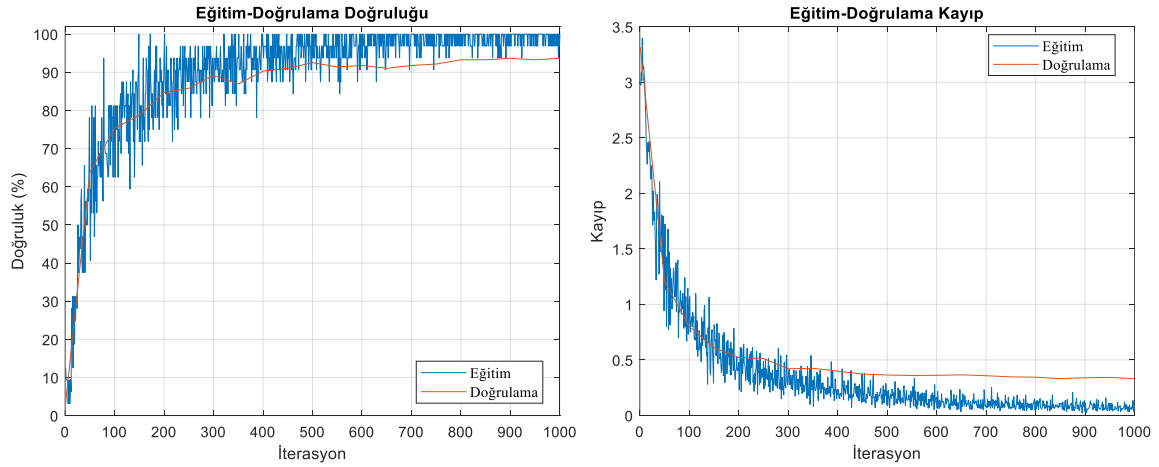
Deneyin ikinci aşamasında, yeni ESA modeli ile öznetelik çıkarımı, cihazların sınıflandırılması için elde edilen görüntülerin %70’i (648 görüntü) eğitim ve %30’u (268 görüntü) test için ayrılmıştır. Eğitim verileri önerilen ESA modelinin eğitilmesi ve test verileri önerilen ESA model parametrelerinin belirlenmesi hem de performans değerlendirmesi amacıyla kullanıldı. ESA modelinin eğitim sürecinde modele ait en iyi parametrelerinin belirlenmesi için deneyler birkaç kez tekrar edildi ve belirlenen ESA model eğitim parametreleri Tablo 4.1’de görülmektedir.

Tablo 4.1’de verilen eğitim parametrelerine göre ESA modelinin eğitim ve sınıflandırma süreçlerine ait eğitim-doğrulama ve eğitim-kayıp grafikleri Şekil 4.3’de görülmektedir.

Tablo 4.1. ESA modeline ait eğitim parametreleri

Optimizasyon Yöntemi	Öğrenme Oranı	Maksimum Tekrar Sayısı	Minimum Yığın Sayısı	Doğrulama Frekansı
Stokastik Gradyan Azalma (SGD)	10^{-3}	50	32	50

ESA model eğitimi 1000 iterasyonda 0,5’ten düşük eğitim doğrulama kaybı ile tamamlanmıştır. Tablo 4.2’de ESA modelinin performans sonuçlarına göre %94,78’lik bir doğruluk oranı ve %99,70’lik özgüllük skoru elde edildiği gözlemlenmiştir.



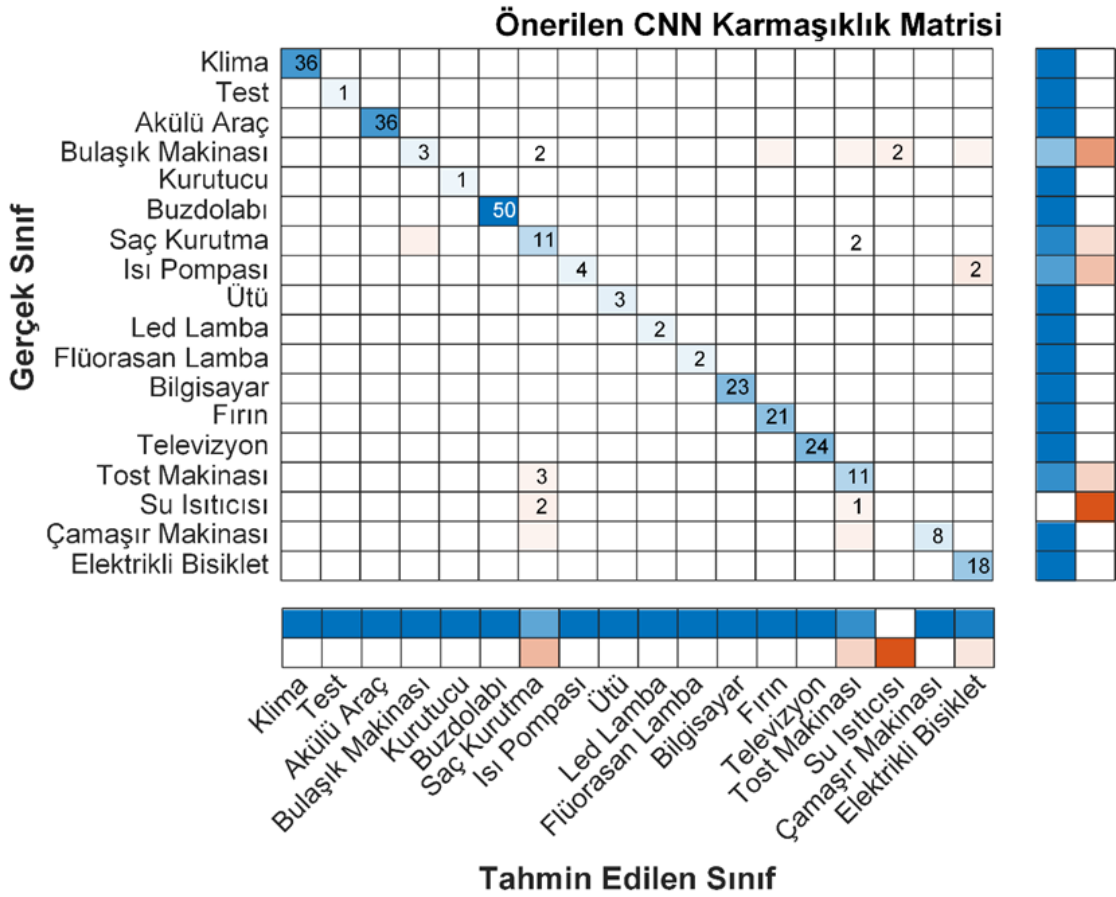
Şekil 4.3. ESA modelinin doğruluk ve kayıp eğrileri

Tablo 4.2’deki duyarlılık, keskinlik ve F1 skor diğer performans ölçütlerinde sırasıyla %87,37, %90,54 ve %88,62 oranında başarımlar sağlanmıştır.

Tablo 4.2. ESA model performans sonuçları (%)

Doğruluk	Duyarlılık	Özgüllük	Keskinlik	F1 Skor
94,78	87,37	99,70	90,54	88,62

Ayrıca deneysel çalışmalardaki test veri setindeki her bir elektrikli cihazın sınıflandırma oranlarının değerlendirilmesi için Şekil 4.4’de karmaşıklık matrisi verilmiştir. Şekil 3.4 detaylı incelendiğinde bulaşık makinası, saç kurutma, ısı pompası, tost makinası ve su ısıtıcısı dışındaki tüm elektrikli cihazlar doğru bir şekilde tahmin edilmiştir. Gerçekte bulaşık makinası olan ikişer adet test verileri sırasıyla saç kurutma ve su ısıtıcısı olarak hatalı sınıflandırılmıştır. Gerçekte saç kurutma olan iki test verisi, tost makinası olarak hatalı tahmin edilmiştir. Bununla birlikte gerçekte 3 tost makinası ve 2 su ısıtıcısı test verisi hatalı bir şekilde saç kurutma ve bir su ısıtıcısı test verisi hatalı şekilde tost makinası olarak sınıflandırılmıştır.



Şekil 4.4. ESA modeline ait karmaşıklık matrisi

Çalışmanın daha detaylı bir performans analizi için benzer veri setlerinin kullanıldığı mevcut bazı yöntemler ile performans karşılaştırmaları Tablo 3.3’de görülmektedir. Tablo 3.3’de görüldüğü üzere veri setlerindeki cihaz sayıları farklılık göstermektedir. Bununla birlikte ESA modeli %94,78 doğruluk oranı ile son teknoloji yöntemlerden daha başarılı

olduğu görülmektedir. ESA modeli [63] ve [64] numaralı çalışmalardan sırasıyla %1,78 ve %1,98 daha yüksek doğruluk oranına sahiptir. Ayrıca [65] ve [66] numaralı çalışmalardan sırasıyla %13,12 ve %11,02 daha yüksek F1 skor değerine sahiptir.

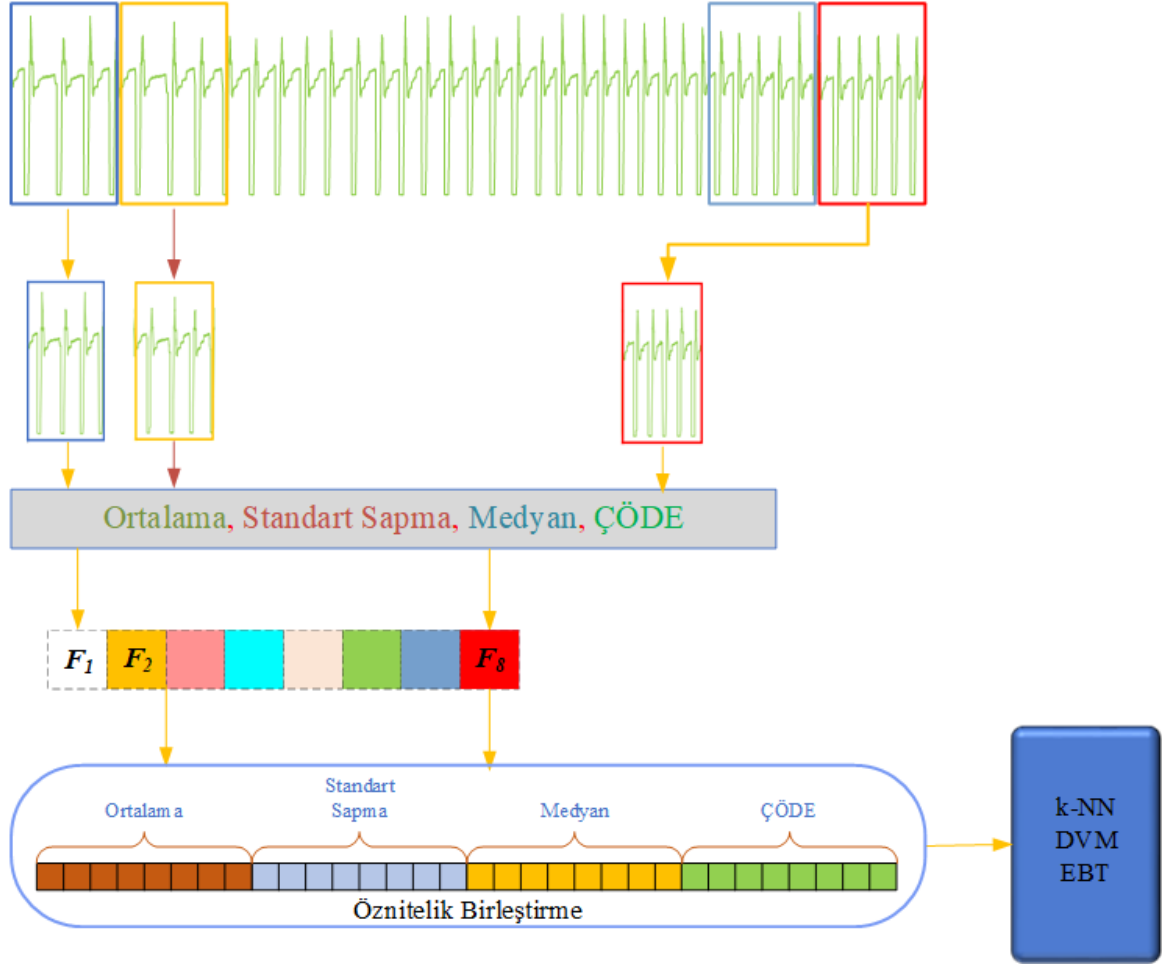
Tablo 4.3. ESA modelin mevcut bazı yöntemler ile performans karşılaştırması

Yöntem	Veri Seti	Elektrikli Cihaz Sayısı	Doğruluk	Özgüllük	F1 Skor
Athanasiadis ve diğ. [65]	WHITED	22	-	-	0,755
Devlin ve Hayes [66]	UK-DALE	7	-	-	0,776
Paiva Penha ve diğ. [63]	REDD	6	0,9300	-	0,930
Le ve diğerleri [64]	PLAID	11	0,9280	-	0,924
ESA Modeli	Yüksek Çözünürlük	18	0,9478	0,9970	0,8862

4.2. Geleneksel Makine Öğrenme Yöntemleri Elektrikli Cihazların Tespiti

İlk olarak, deneysel çalışmalarda daha iyi performans elde etmek için güç tüketim sinyalleri için kayan pencere yaklaşımı ile öznitelik çıkarımı gerçekleştirilmiştir. Bu amaçla, her bir sinyal, pencere uzunluğu ampirik olarak belirlenen sekiz örtüşmeyen pencereye bölünmüştür. Daha sonra sekiz pencerenin her birinden “ortalama, standart sapma, medyan ve çok ölçekli dağılım entropisi” değerleri Şekil 4.5’deki gibi hesaplanmıştır.

Her bir güç tüketimi zaman serisinden kayan pencere yaklaşımı ile her bir istatistiksel ve ÇÖDE yöntemleri 1×8 boyutunda öznitelik vektörü oluşturulur. Bu işlem bir sinyal için üç istatistiksel ve ÇÖDE için tekrarlandığında, 1×8 boyutunda dört farklı öznitelik vektör elde edilmiştir. Böylece, 1×86400 boyutundaki bir güç tüketim sinyali için “ortalama, standart sapma, medyan ve çok ölçekli dağılım entropisi” öznitelikleri birleştirilerek 1×32 öznitelik vektörü elde edilir. Toplamda veri setindeki 916 örnek için 916×32 öznitelik matrisi elde edilir.



Şekil 4.5. Zaman serisinden kayan pencere yaklaşımı ile istatistiksel ve ÇÖDE özniteliklerin elde edilmesi

Güç tüketim zaman serilerinden elde edilen öznitelik matrisi elektrikli cihazların sınıflandırılması için k-NN, DVM ve EBT makine öğrenim sınıflandırıcılara uygulandı. Değerlendirme ölçütlerine göre sınıflandırma sonuçları Tablo 4.4’de görülmektedir.

Tablo 4.4. Geleneksel makine öğrenim yöntemleri sınıflandırma sonuçları

Yöntem	Doğruluk	Duyarlılık	Özgüllük	Keskinlik	F1 Skor
k-NN	95,19	91,42	99,72	94,68	91,50
EBT	96,07	90,99	99,76	96,64	92,40
DVM	94,76	85,38	99,68	93,75	86,50

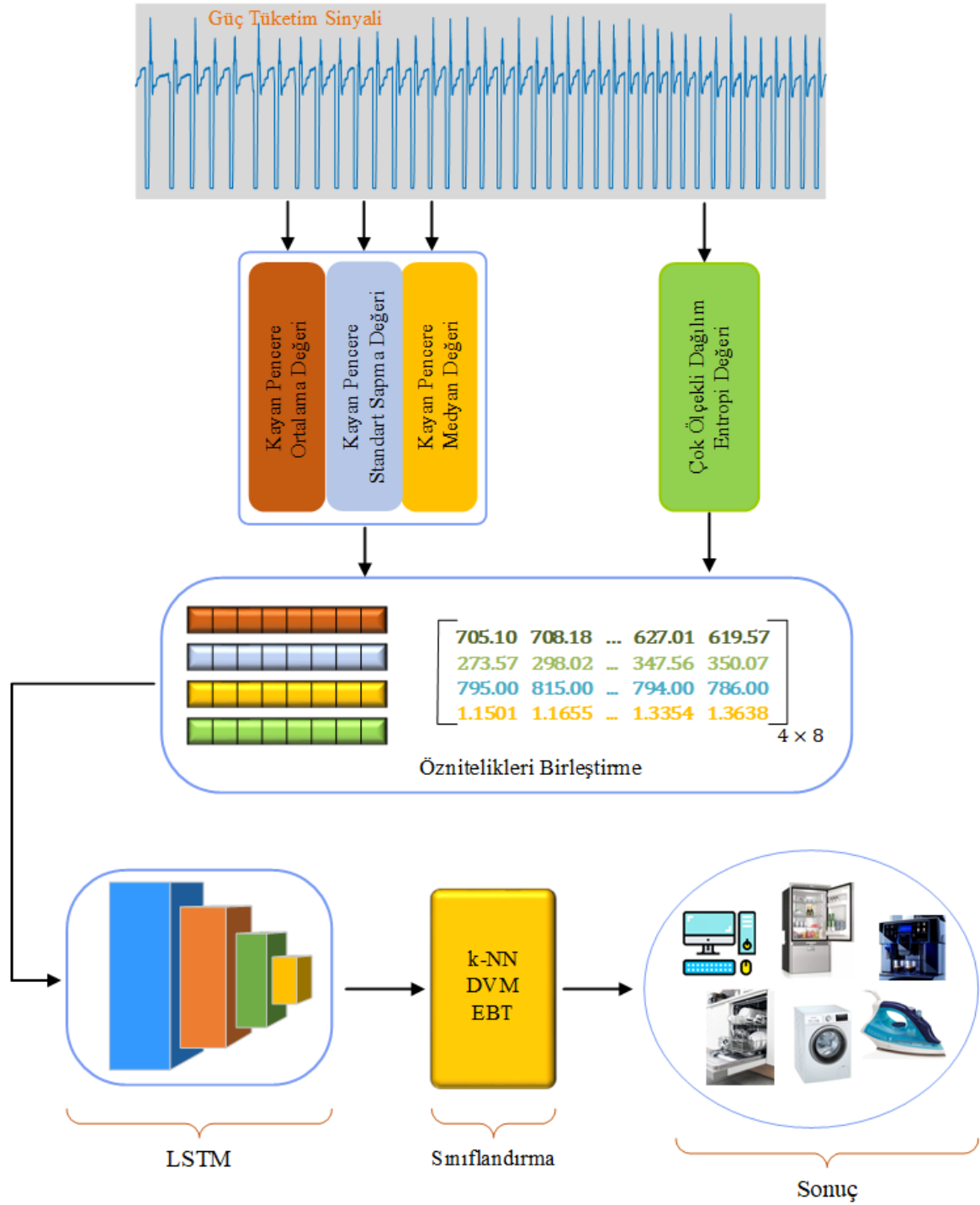
Tablo 4.4’deki en yüksek sınıflandırma performansı %96,07 ile EBT sınıflarında elde edilmiştir. k-NN ve SVM sınıflandırmalarında %95,19 ve %94,76 doğruluk oranı elde

edilmiştir. Veri setindeki örnek dağılımları dengesiz olduğu için F1 skor sonuçları performansı değerlendirmede daha etkili olduğu gözlemlenmiştir. Tablo 3.4'deki F1 sonuçlarına göre DVM, k-NN ve EBT sınıflandırıcılarında sırasıyla %86,50, %91,50 ve %92,40 puan değerlerine ulaşılmıştır. F1 puanı sonuçlarında en yüksek performans EBT sınıflandırıcısında elde edilmiştir.

4.3. Hibrit Model ile Elektrikli Cihazların Zaman Serisi Özelliklerine Göre Sınıflandırılması

Diğer bir çalışma olarak veri setindeki elektrik cihazların güç tüketim zaman serilerinden cihazların otomatik tespiti için LSTM tabanlı hibrit bir model geliştirildi. Genel çerçevesi Şekil 4.6'de verilen model, üç aşamadan oluşmaktadır. İlk aşamada, geleneksel makine öğrenim yönteminde olduğu gibi LSTM modelinin hesaplama maliyetini azaltmak için örtüşmeyen kayar pencere yaklaşımı ile zaman serileri pencerelere ayrılmakta ve her bir pencerenin ortalama, standart sapma, medyan ve entropi değerleri birleştirilerek zamansal özellikler oluşturulmaktadır. İkinci aşamada, derin öznitelikler elde etmek için bu zamansal öznitelikler LSTM'ye uygulanmaktadır. Son olarak, kontlarda kullanılan elektrik cihaz tespiti için popüler makine öğrenim sınıflandırıcılara derin özellikler uygulanarak cihazlar sınıflandırılmaktadır.

Geliştirilen hibrit model deneysel çalışmalarında, her biri 1x86400 boyutunda ve Bölüm 3.1.4'de ayrıntılı olarak açıklanan 18 farklı elektrikli ev cihazının güç tüketim sinyallerini kullanmıştır. Veri seti, her iki deneysel çalışmada da %75 (687 örnek) eğitim ve %25 (229 örnek) test için rastgele iki bölüme ayrılmıştır. Eğitim verileri önerilen modeli eğitmek için kullanılmıştır. Ayrıca modelin değerlendirilmesi için test verileri kullanılmıştır.

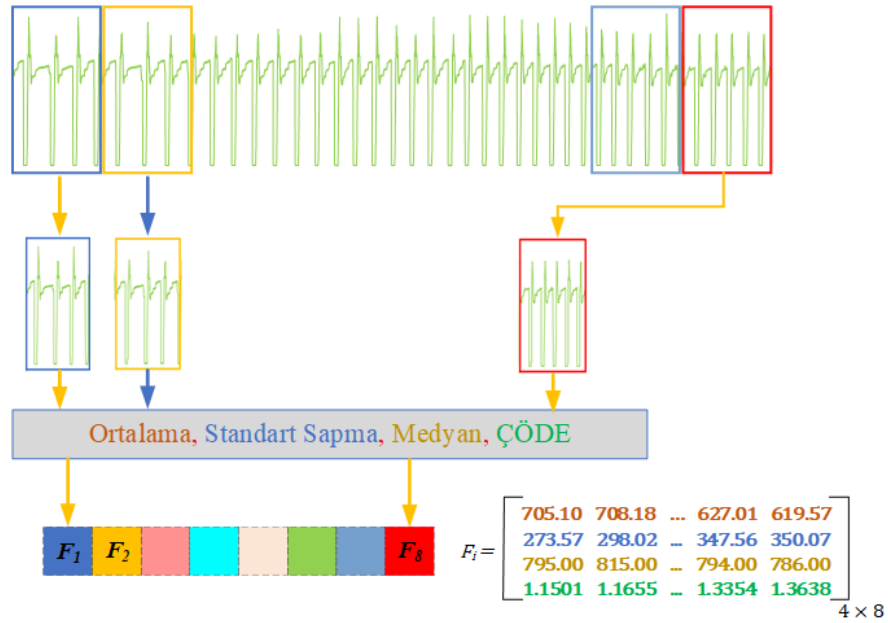


Şekil 4.6. LSTM tabanlı hibrit modelin genel çerçevesi

Burada, makine öğrenim yöntemlerinden farklı olarak her bir zaman serisi güç sinyali yine örtüşmeyen sekiz pencereye bölünmüş ve her birinden üç istatistiksel ve çok ölçekli dağılım entropi değerinden 1×8 boyutunda dört farklı öznitelik vektör elde edilmiştir. Tüm "ortalama, standart sapma (STD), ortanca ve çok ölçekli dağılımın entropisi"

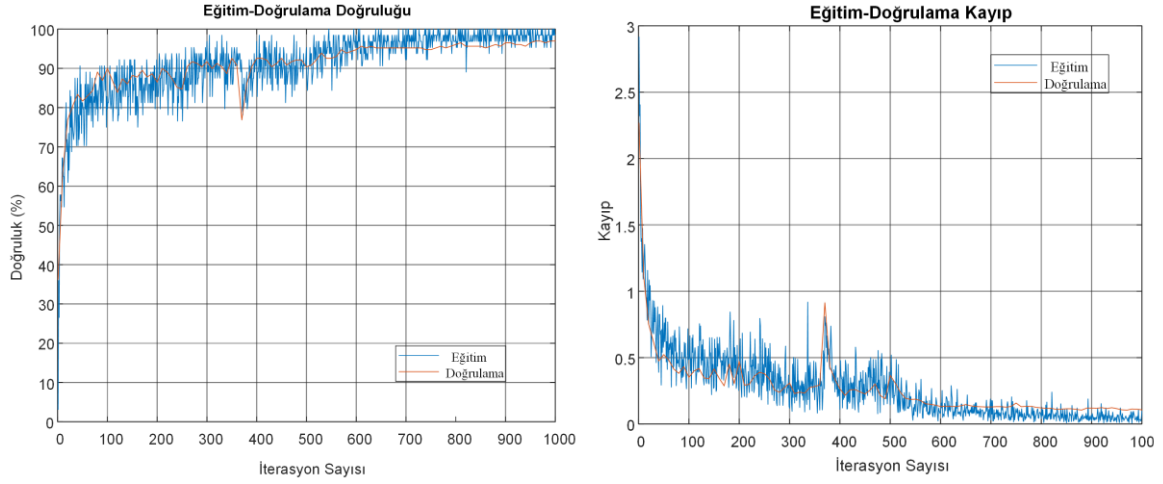
özniteliklerinin birleştirilmesiyle 4x8 geçici öznitelikler matrisi elde edilir. Şekil 4.7’de, bir zaman serisinin 4x8 boyutlu geçici öznitelikleri matrisinin nasıl elde edildiği görülmektedir. Bu işlem veri setindeki 916 zaman serisi için tekrarlanmış ve derin öznitelikler elde etmek için LSTM modeline girdi olarak $916 \times (4 \times 8)$ geçici öznitelikler kullanılmıştır.

Bu geçici özniteliklerden derin özniteliklerin elde edilmesi için geliştirilen LSTM modeli, 250 nöronlu iki LSTM katmanı, 0,5 bırakma oranına sahip iki bırakma katmanı, 100 ve 18 nöronlu iki tam bağlantılı katman, bir softmax katmanı ve bir sınıflandırma katmanından oluşmaktadır.



Şekil 4.7. Zaman serisinin 4x8 boyutlu geçici öznitelikleri matrisinin elde edilmesi

LSTM modelinin performansı; LSTM gizli nöron sayısı, öğrenme oranı, parti boyutu, eğitim periyodu sayısı ve bırakma oranları gibi birçok hiperparametreye bağlıdır. Bu çalışmada, parametrelerin uygun değerini belirlemek için çeşitli hiperparametre kombinasyonlarını göz önünde bulunduran ampirik değerlendirmeden sonra, önerilen LSTM modelimizi Adam optimizer aracılığıyla 0.01 sabit öğrenme oranı, 64 minibatch boyutu dizisi ve 100 epok iterasyon ile eğittik. Şekil 3.8’deki LSTM modelinin eğitim sürecinin doğruluk ve kayıp grafiklerinde görüldüğü gibi 1000 iterasyon sonunda eğitim doğruluğu ve kayıp puanları sırasıyla %96,94 ve 0,10 olmuştur.



Şekil 4.8. LSTM ağı eğitim sürecinin doğruluk ve kayıp grafikleri

LSTM modelinin eğitimi sırasında kayıp değeri ne kadar düşükse, o kadar iyidir. Ayrıca, daha küçük bir kayıp değeri daha az genellenebilirlik ile sonuçlanır. Şekil 4.8’de görüldüğü gibi, LSTM modelinin 0,10 kayıp değeri ile eğitim verilerine daha sağlam bir uyum performansı gösterdiği söylenebilir. Eğitilmiş LSTM modelinde, ikinci tam bağlı katmanın aktivasyon değerleri kullanılarak 916 derin öznetelik çıkarılmıştır. Bu öznetelikler, LSTM modelinin etkinliğini doğrulamak için üç popüler makine öğrenme algoritmasına uygulandı. Geliştirilen LSTM modeli ve LSTM modeli ile birlikte k-NN, EBT ve DVM sınıflandırma performansları, Tablo 4.5’te özetlenmiştir.

Tablo 4.5. LSTM modeli ve LSTM modeli ile birlikte k-NN, EBT ve DVM sınıflandırma performansları

Yöntem	Doğruluk	Duyarlılık	Özgüllük	Keskinlik	F1 Skor
LSTM	96,94	94,26	99,82	90,74	91,20
LSTM+k-NN	98,25	95,92	99,90	93,82	93,50
LSTM+EBT	97,81	95,23	99,87	92,63	92,70
LSTM+DVM	97,38	93,38	99,85	91,25	90,08

Tablo 4.5’deki sonuçlardan da görüleceği gibi, LSTM sınıflandırma doğruluğu %96,94 iken, k-NN, EBT ve DVM gibi popüler makine öğrenim sınıflandırıcılar ile birlikte kullanımından oluşan hibrit modelin sınıflandırma performansında %98,25, %97,81 ve %97,38 doğruluk oranlarına ulaşılmıştır. Ayrıca doğruluk, duyarlılık, özgüllük, keskinlik,

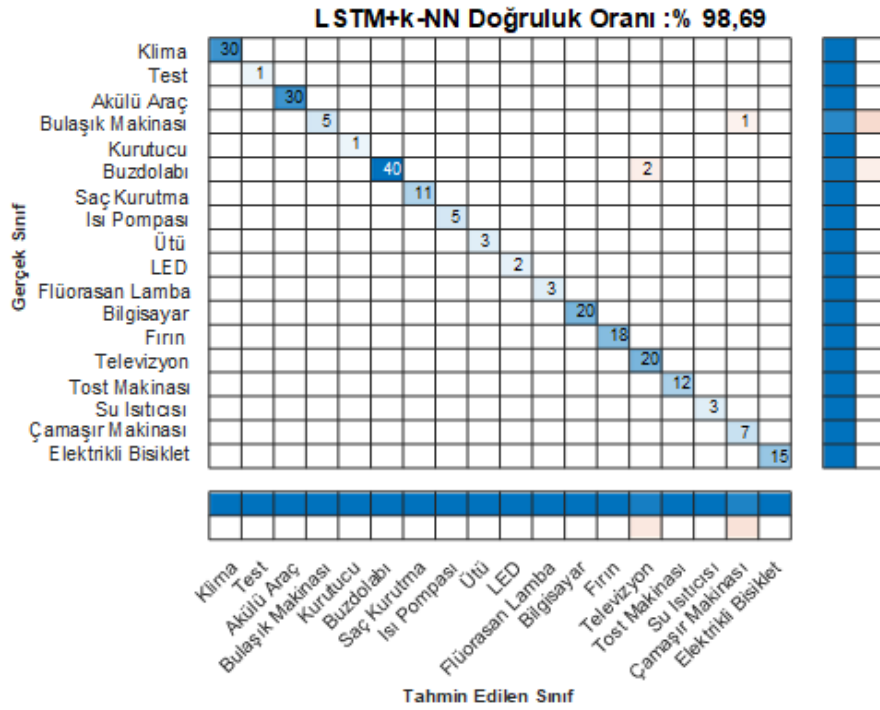
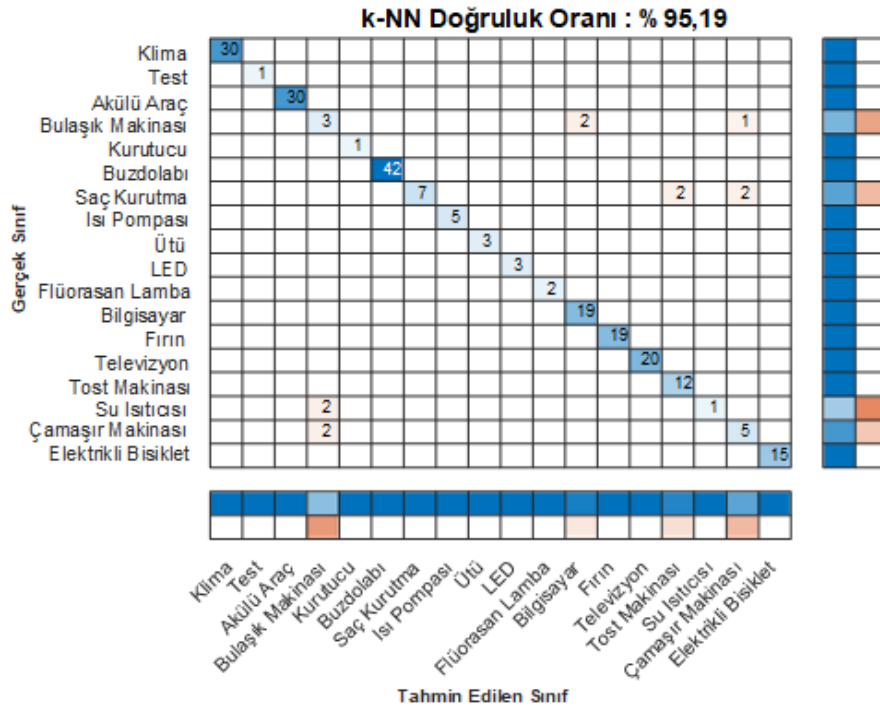
F1 skoru performans değerlendirme ölçütlerinde sırasıyla %98,25, %95,92, %99,90, %93,82 ve %93,50 ile en yüksek başarı LSTM+k-NN kombinasyonunda elde edilmiştir. Sadece LSTM ve LSTM+k-NN kombinasyonu olan hibrit modelin doğruluk sonuçları karşılaştırıldığında %1,31 daha yüksek başarı sağlanmıştır.

Ayrıca Tablo 4.6’da verilen geleneksel makine öğrenim yöntemleri ve LSTM tabanlı hibrit model performans değerlendirme sonuçları karşılaştırıldığında, sadece k-NN sınıflandırıcı doğruluk oranı %95,19 iken, LSTM ve k-NN sınıflandırıcının birlikte kullanımında elde edilen hibrit modelde doğruluk oranı %98,25’e yükseltilmiştir. Aynı şekilde sadece EBT sınıflandırıcı doğruluk oranı %96,07 iken, LSTM ve EBT sınıflandırıcının birlikte kullanımında elde edilen hibrit modelde %97,81 doğruluk oranına ve sadece DVM sınıflandırıcı doğruluk oranı %94,76 iken, LSTM ve DVM sınıflandırıcının birlikte kullanımında elde edilen hibrit modelde %97,38 doğruluk oranına ulaşılmıştır.

Tablo 4.6. Geleneksel makine öğrenim yöntemleri ve LSTM tabanlı hibrit model performans sonuçlarının karşılaştırılması

Yöntem	Doğruluk	Duyarlılık	Özgüllük	Keskinlik	F1 Skor
k-NN	95,19	91,42	99,72	94,68	91,50
EBT	96,07	90,99	99,76	96,64	92,40
DVM	94,76	85,38	99,68	93,75	86,50
LSTM	96,94	94,26	99,82	90,74	91,20
LSTM+k-NN	98,25	95,92	99,90	93,82	93,50
LSTM+EBT	97,81	95,23	99,87	92,63	92,70
LSTM+DVM	97,38	93,38	99,85	91,25	90,08

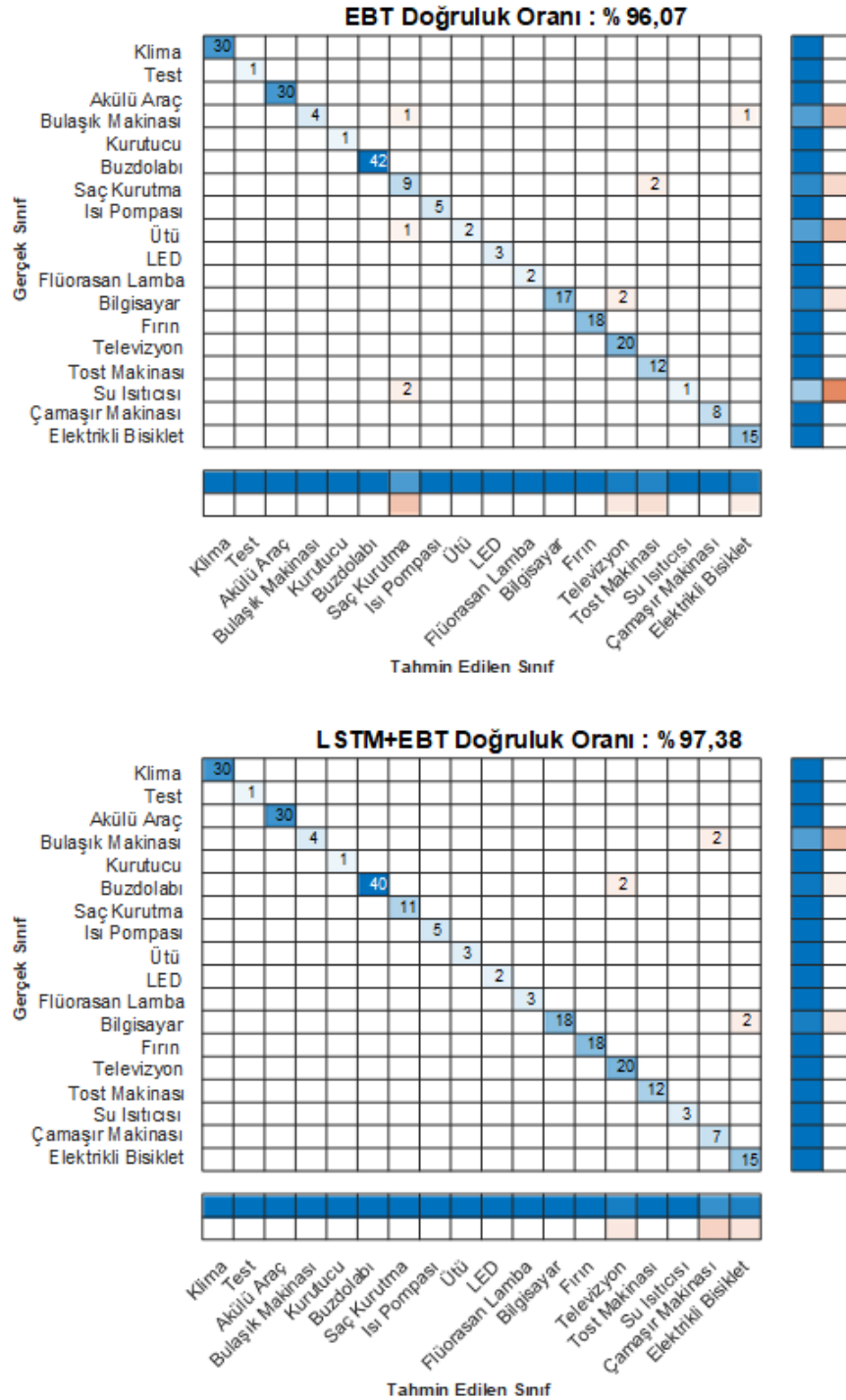
Veri setindeki her bir elektrikli cihaza özel daha detaylı performans analizi için geleneksel makine öğrenim ve LSTM tabanlı hibrit modelin performans değerlendirmeleri için karmaşıklık matrisleri kullanılabilir. Bu amaçla, Şekil 4.9’da k-NN ve LSTM+k-NN, Şekil 4.10’da EBT ve LSTM+EBT ve Şekil 3.11’de DVM ve LSTM+DVM modellerinin karmaşıklık matrisleri verilmiştir. Karmaşıklık matrislerinin amacı genellikle her bir cihazın sınıflandırma hassasiyetini, kesinliğini ve F1 puan oranlarını görme fırsatı sağlamaktır.



Şekil 4.9. k-NN ve LSTM+k-NN için karmaşıklık matrisi

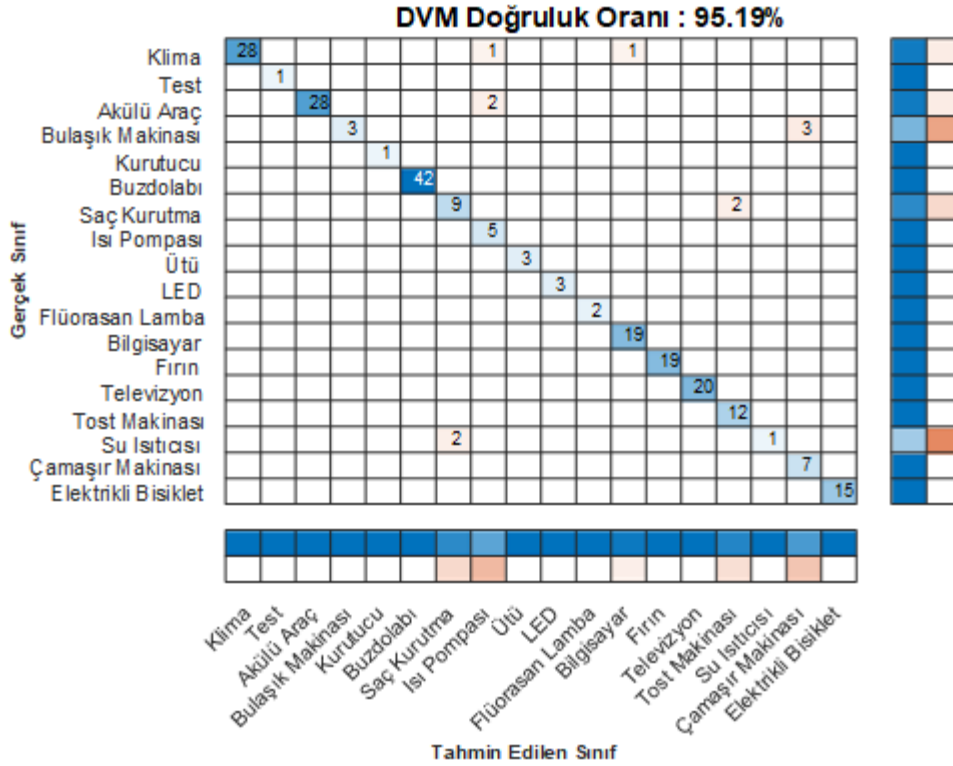
Şekil 4.9’da görüldüğü üzere sadece geleneksel makine öğrenimi k-NN sınıflandırıcı karmaşıklık matrisinde 229 test verisinden 218’i doğru tahmin edilirken LSTM+k-NN modelde 229 test verisinden 226’sı doğru tahmin edilmiştir. Gerçekte, bulaşık makinesi olarak etiketlenen iki test verisi, LSTM+k-NN karmaşıklık matrisinde hatalı bir şekilde

kurutucu ve ısı pompası olarak tahmin edildi ve ısı pompası etiketli bir test verisi, flüorosan ampul olarak hatalı bir şekilde tahmin edildi. Cihazların geri kalanı doğru şekilde sınıflandırılmıştır. Diğer bir deyişle, bulaşık makinesi ve ısı pompası hariç tüm cihazlarda %100 doğruluk, hassasiyet ve F1 skor oranları elde edildi.

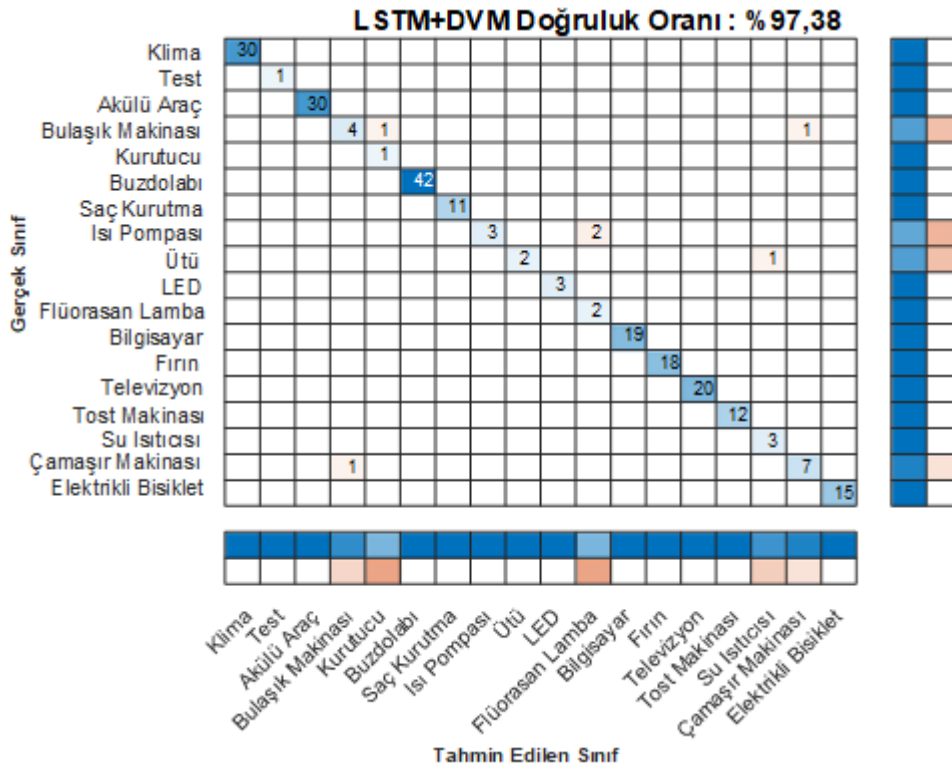


Şekil 4.10. EBT ve LSTM+EBT için karmaşıklık matrisi

Şekil 4.10’da görüldüğü üzere geleneksel makine öğrenimi EBT sınıflandırıcı karmaşıklık matrisinde 229 test verisinden 220’i doğru tahmin edilirken LSTM+EBT modelde 229 test verisinden 224’ü doğru tahmin edilmiştir. Gerçekte, ısı pompası olarak etiketlenen iki test verisi, LSTM+EBT karmaşıklık matrisinde flüorsan ampul olarak hatalı tahmin edildi ve gerçekte bulaşık makinası etiketli olan iki adet test verisi, kurutucu ve çamaşır makinası olarak hatalı bir şekilde ve gerçekte çamaşır makinası etiketli olan bir adet test verisi bulaşık makinası olarak hatalı tahmin edildi. Cihazların geri kalanı doğru şekilde sınıflandırılmıştır. Diğer bir deyişle, bulaşık makinesi, çamaşır makinası ve ısı pompası hariç tüm cihazlarda %100 doğruluk, hassasiyet ve F1 skor oranları elde edildi.



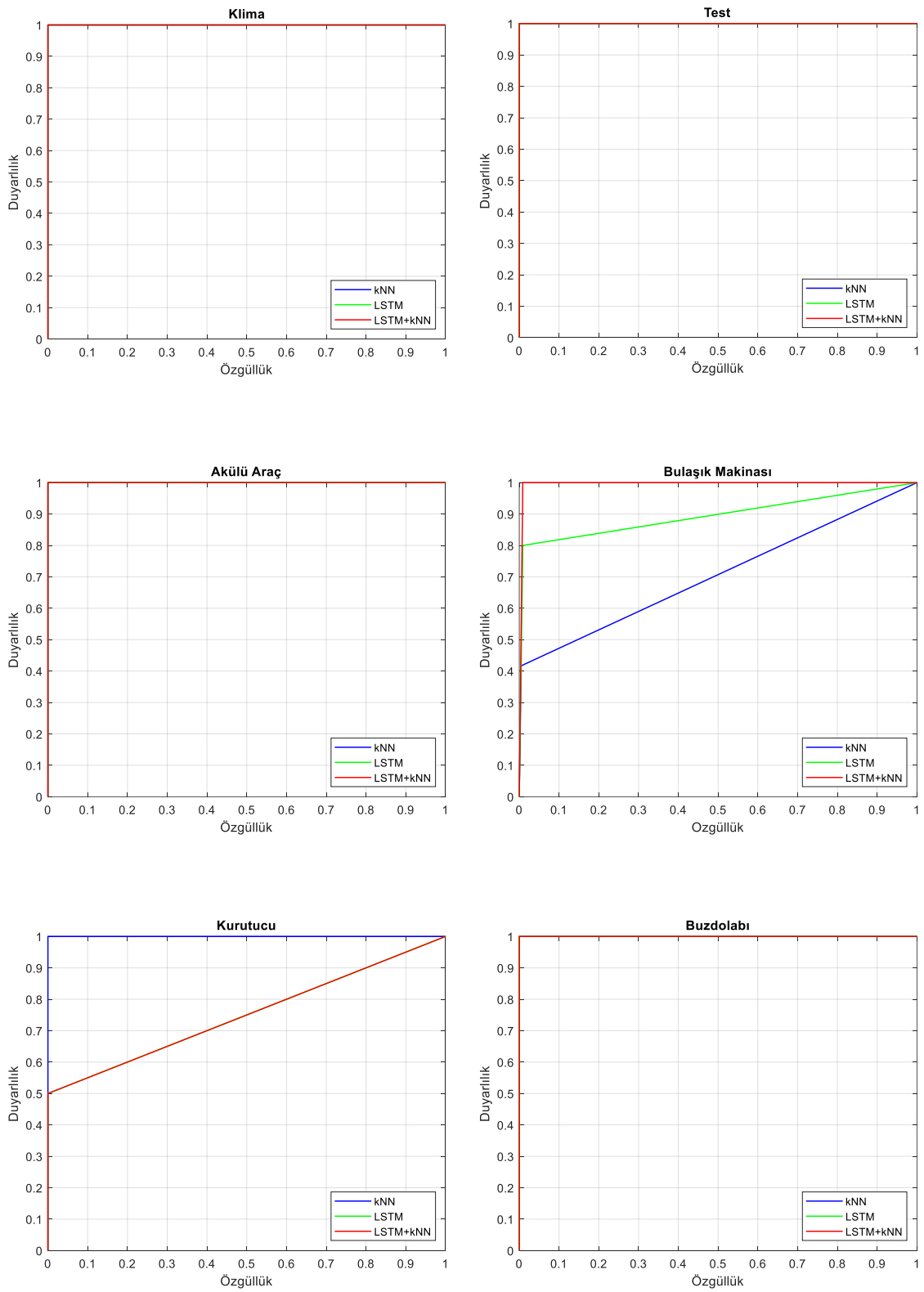
Şekil 4.11. DVM karmaşıklık matrisi



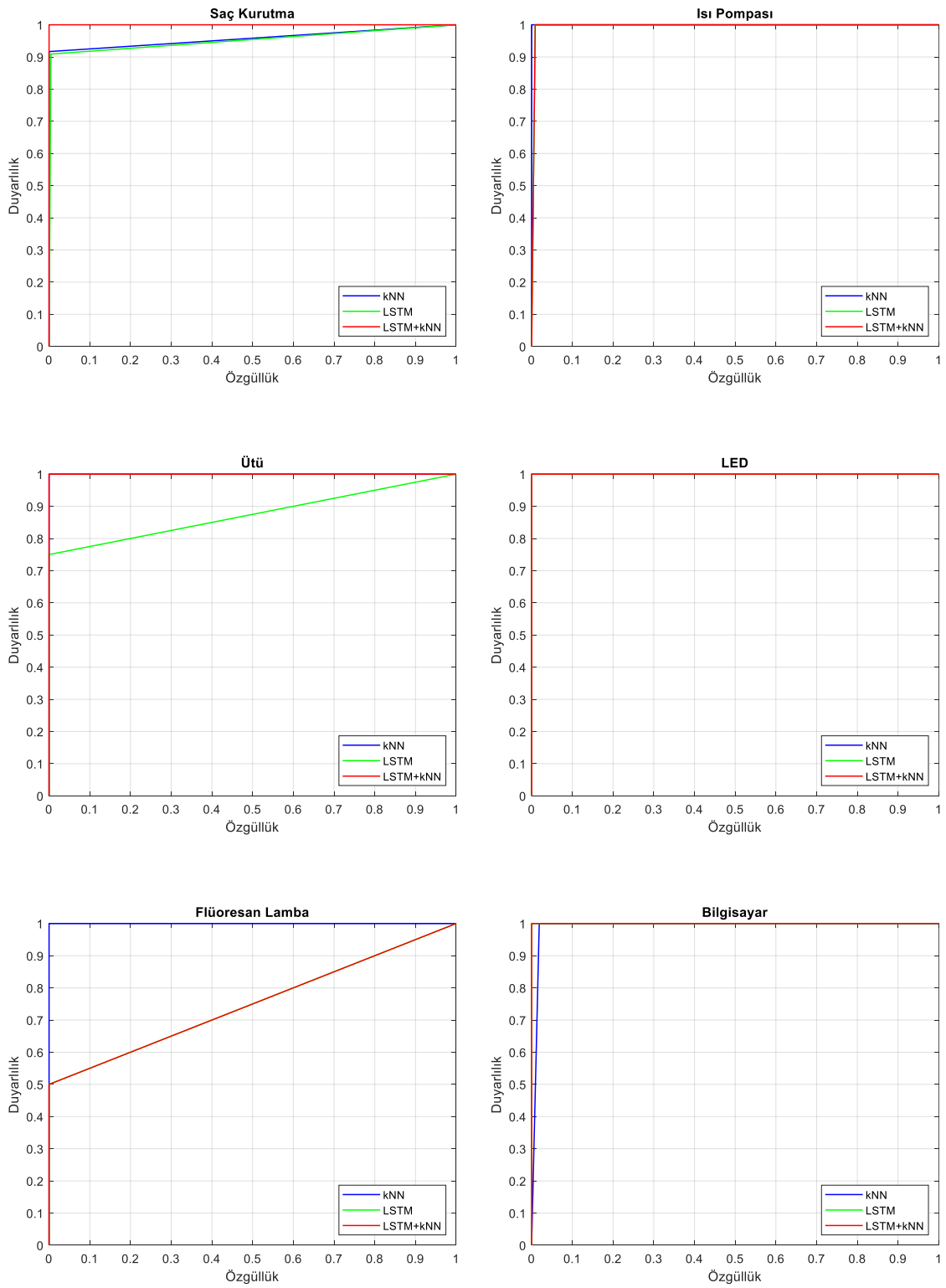
Şekil 4.12. LSTM+DVM için karmaşıklık matrisi

Şekil 4.11'deki geleneksel makine öğrenimi DVM sınıflandırıcı karmaşıklık matrisinde 229 test verisinden 218'i doğru tahmin edilirken, LSTM+DVM modelde 229 test verisinden 223'ü doğru tahmin edilmiştir. Şekil 4.12'de Gerçekte, ısı pompası olarak etiketlenen iki test verisi, LSTM+DVM karmaşıklık matrisinde flüorasan ampul ve bulaşık makinası olarak hatalı tahmin edildi, gerçekte bulaşık makinası etiketli olan iki adet test verisi, kurutucu ve çamaşır makinası olarak hatalı bir şekilde ve gerçekte ütü etiketli olan bir adet test verisi bulaşık makinası olarak hatalı tahmin edildi. Cihazların geri kalanı doğru şekilde sınıflandırılmıştır. Diğer cihazlarda %100 doğruluk, hassasiyet ve F1 skor oranları elde edildi.

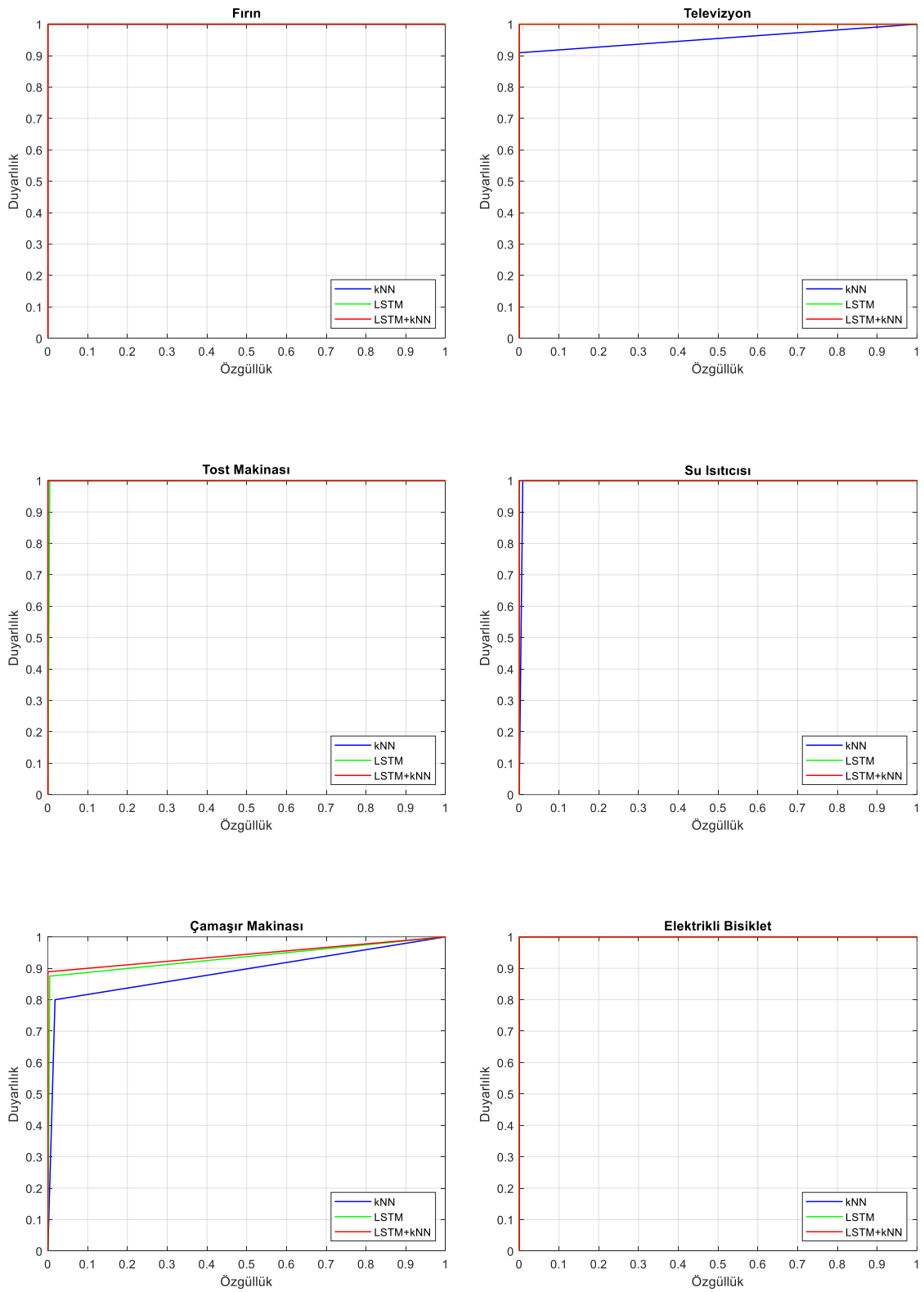
Cihaz düzeyinde daha detaylı performansını analizi için veri setindeki 18 cihazın en yüksek performanslı geleneksel makine öğrenim yöntemi olan k-NN, LSTM ve LSTM tabanlı hibrit (LSTM+k-NN) modellerinin performanslarını karşılaştıran ROC eğrileri Şekil 4.13'de verilmiştir.



Şekil 4.13. k-NN, LSTM ve LSTM+k-NN modellerine ait ROC eğrileri

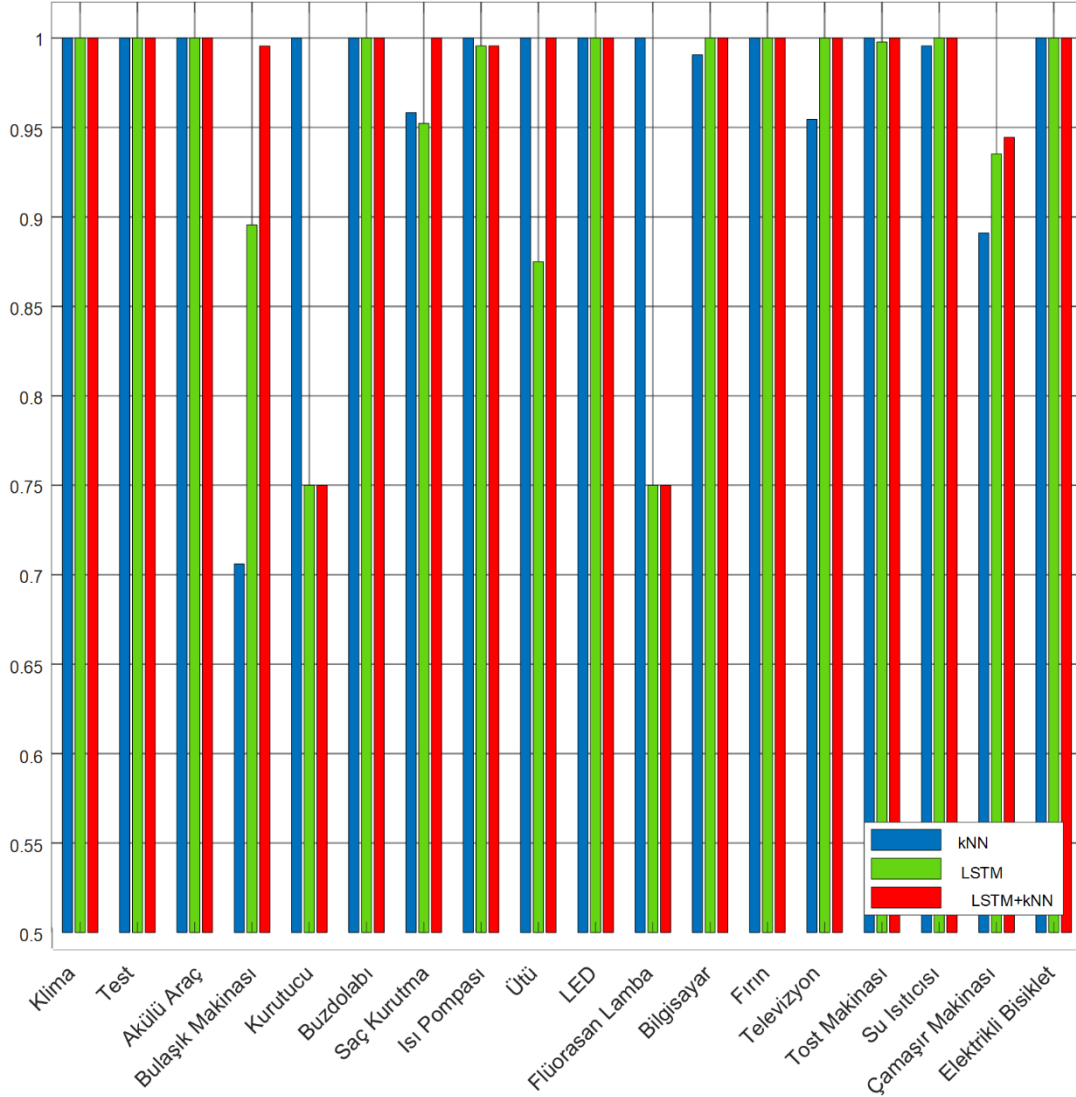


Şekil 4.13. (Devam): k-NN, LSTM ve LSTM+k-NN modellerine ait ROC eğrileri



Şekil 4.13. (Devam): k-NN, LSTM ve LSTM+k-NN modellerine ait ROC eğrileri

Ayrıca veri setindeki 18 cihazın tespit edilme performansını daha etkin bir şekilde göstermek için k-NN, LSTM ve LSTM+k-NN modellerinin AUC değerlerinin karşılaştırılması Şekil 4.14’de görülmektedir.



Şekil 4.14. k-NN, LSTM ve LSTM+k-NN modellerine ait AUC değerlerinin karşılaştırılması

Şekil 4.14’de, tüm AUC değerlerinin 0,5’ten çok daha büyük olduğunu ve ilgili tüm sınıflandırma yaklaşımlarının gayet iyi bir performansla cihazları tespit emektedir. Ayrıca LSTM+k-NN yönteminde Klima, Test, Akülü Elektrikli Araç, Buzdolabı, Saç Kurutma Makinesi, Ütü, Led, PC, Fırın, TV, Tost Makinesi, Su Isıtıcısı ve Elektrikli Bisiklet gibi 11 adet cihazda 1 AUC değerine ulaşmıştır. Bu durum, LSTM+k-NN yönteminde diğer yöntemlerden daha iyi performansa sahip olduğunu göstermektedir.

LSTM temelli hibrit model performansının değerlendirme ölçütlerine göre, bazı son teknoloji çalışmaları ile karşılaştırılması Tablo 4.7'de verilmiştir.

Tablo 4.7. LSTM ağ temelli modelin, son teknoloji bazı çalışmaların başarımlarının karşılaştırılması

Yöntem	Elektrikli Cihaz	Doğruluk	Duyarlılık	Keskinlik	F1 Skor
Son teknoloji model	Klima	1,000	1,000	1,000	1,000
	Test	1,000	1,000	1,000	1,000
	Akülü Araç	1,000	1,000	1,000	1,000
	Bulaşık Makinası	0,666	0,666	1,00	0,800
	Kurutucu	1,000	1,000	1,000	1,000
	Buzdolabı	1,000	1,000	0,989	0,977
	Saç Kurutma	1,000	1,000	0,981	0,846
	Isı pompası	0,600	0,600	1,000	0,750
	Ütü	0,333	0,333	1,000	0,800
	Led Lamba	0,666	0,666	1,000	0,800
	Flüorsan Lamba	1,000	1,000	0,911	0,667
	Bilgisayar	0,894	0,894	1,000	0,944
	Fırın	1,000	1,000	1,000	1,000
	Televizyon	1,000	1,000	0,955	0,976
	Tost Makinası	1,000	1,000	0,995	0,960
	Su Isıtıcısı	0,333	0,333	1,000	0,500
	Çamaşır Makinası	0,875	0,875	0,995	0,875
	Elektrikli Bisiklet	1,000	1,000	0,995	0,968
	Ortalama	0,947	0,853	0,996	0,865
	Önerilen LSTM+k-NN model	Klima	1,000	1,000	1,000
Test		1,000	1,000	1,000	1,000
Akülü Araç		1,000	1,000	1,000	1,000
Bulaşık Makinası		0,666	0,666	1,000	0,800
Kurutucu		1,000	1,000	0,955	0,667
Buzdolabı		1,000	1,000	1,000	1,000
Saç Kurutma		1,000	1,000	1,000	1,000
Isı pompası		0,600	0,600	1,000	0,750
Ütü		1,000	1,000	1,000	1,000
Led Lamba		1,000	1,000	1,000	1,000
Flüorsan Lamba		1,000	1,000	0,911	0,667
Bilgisayar		1,000	1,000	1,000	1,000
Fırın		1,000	1,000	1,000	1,000
Televizyon		1,000	1,000	1,000	1,000
Tost Makinası		1,000	1,000	1,000	1,000
Su Isıtıcısı		1,000	1,000	1,000	1,000
Çamaşır Makinası		1,000	1,000	0,995	0,941
Elektrikli Bisiklet		1,000	1,000	1,000	1,000
Ortalama		0,982	0,959	0,999	0,935

Tablo 4.7’de, geliştirilen hibrit modelin genel performansının, her bir cihaza ve ortalama sonuçlara göre son teknoloji modelden daha yüksek başarımlarına sahip olduğu görülmektedir. Son teknoloji modelde ütü ve su ısıtıcısı doğruluk oranı 0.333 iken, LSTM tabanlı hibrit modelde bu cihazlarda 1.000 doğruluk oranına ulaşılmıştır. Hibrit modelde bulaşık makinesi ve ısı pompasında sırasıyla 0,666 ve 0,600 doğruluk oranları elde edilirken, diğer tüm cihazlarda %100 doğruluk elde edilmiştir. Genel ortalamalarda, en son modelin doğruluğu ve F1 skoru sırasıyla 0,947 ve 0,865 iken, geliştirilen hibrit modelde sırasıyla 0,982 ve 0,935'tir. Böylece doğrulukta %13,50 ve F1 puanında %7,00 artış sağlanmıştır.

Araştırmacılar, MOYİ yaklaşımında yaptıkları çalışmalarda, WHITED, UK-DALE, REDD ve PLAID gibi genel veri kümelerini ve özel veri setlerini kullanmışlardır. Bu veri setlerinde, örnekleme frekansları ve cihaz sayısı değişmektedir. Sonuç olarak, yöntemleri karşılaştırırken bu ayrımları dikkate almak çok önemlidir. Tablo 3.8, en son yapılan bazı çalışmalar ve bu çalışmalarda kullanılan veri setleri ve ayrıca bu veri setlerinin örnekleme frekansı ve cihaz sayıları verilmiştir.

Tablo 4.8. Hibrit model performansının son teknoloji yöntemlerle karşılaştırılması

Yöntem	Veri Seti	Cihaz Sayısı	Örnekleme Frekansı	Doğruluk	Duyarlılık	F1 Skor
ESA+k-NN [67]	Özel	8	100 Hz	0,987	0,927	0,932
ESA [65]	WHITED	22	44 kHz	-	-	0,755
YSA [66]	UK-DALE	7	0,167 Hz	-	0,857	0,776
ESA [63]	REDD	6	1 Hz	0,930	0,933	0,930
PSO [68]	PLAID	11	30 kHz	0,980	0,838	0,906
HFD [68]	PLAID	11	30 kHz	0,928	-	0,924
Hibrit Modelimiz	Yüksek Çözünürlük	18	1 Hz	0,982	0,959	0,935

Tablo 4.8'de gösterildiđi gibi, [65], [66], [63] ve [64] alıřmalardaki yntemlere gre daha yksek dođruluk oranına sahipken, [67] ve [68] alıřmaları ile hemen hemen aynı dođruluk oranı elde edilmiřtir. Veri setlerindeki her bir cihaz iin rnek sayısı farklı olduđunda bazı alıřmalarda F1 puanları kullanılmıřtır. F1 puanı iin nerilen yntem diđer yntemlere gre daha yksek performansa sahiptir. Bu, nerilen LSTM+k-NN mevcut alıřmalarda daha yksek dođruluk elektrikli cihazların tespit edildiđini gstermektedir.

5. SONUÇLAR VE ÖNERİLER

Son dönemlerde konutlarda tüketilen elektrik enerji oranı gittikçe artmaktadır. Dünyada enerji kaynaklarının sınırlı olması ve bu fosil enerji kaynaklarının aşırı kullanımı çevresel problemlere neden olmaktadır. Bu nedenle konutlarda kullanılan elektrik enerjisinin verimli kullanımı, bu enerji talebinin karşılanabilmesi için yenilenebilir enerji kaynak yatırımlarının planlanması ve arz-talep dengesinin sağlanabilmesi konutlarda kullanılan elektrikli cihazların etkili bir şekilde tespitine bağlıdır.

Bu tez çalışması kapsamında konutlardaki elektrikli cihazların etkili tespiti için derin öğrenme tabanlı yöntemler geliştirildi. Bu yöntemler, oldukça popüler olan ESA ve LSTM tabanlı derin öğrenme modelleri içermektedir. Bu yöntemlerin konutlardaki elektrikli cihazların tespitinde yüksek performanslar sağladığı gözlemlenmiştir.

Tez kapsamındaki ilk çalışmada, klasik ESA modellerindeki RELU aktivasyon fonksiyonu yerine ELU aktivasyon fonksiyonu kullanılan yeni bir ESA modeli ile sınıflandırma başarımının artırılması hedeflenmiştir. Çalışmanın ön işlem aşamasında, elektrikli cihazlara ait zaman serileri önerilen ESA modeline uygun olarak renkli görüntülere dönüştürmüştür. Daha sonra bu görüntüler ESA modeline uygulanarak özellik çıkarımı yapılarak elektrikli cihazlar sınıflandırılmıştır. Deneysel çalışmalarda, cihaz sınıflandırmasında %94,78 doğruluk oranına ulaşılmıştır. Önerilen ESA yöntem veri setindeki 18 farklı cihazın 12'sini %100 doğrulukta tespit etmiştir.

Yapılan ikinci çalışmada, LSTM tabanlı bir hibrit model önerilmiştir. Hibrit modelde, geleneksel özellik çıkarma perspektifinden farklı olarak, cihazların özelliklerini tanımlamak için zaman serisi özelliklerinden faydalanılmıştır. Önerilen hibrit model, zaman serilerinde boyut patlaması problemini çözmek, MOYİ sistem maliyetini azaltmak ve düşük kapasiteli donanım kullanımına imkân sağlamaktadır. Bu amaçla, zaman serilerinden öznitelik çıkarımı için ilk olarak örtüşmeyen kayan pencere yaklaşımı ile her bir zaman serisi sekiz pencereye bölündü ve her pencerenin ortalama, standart sapma, medyan ve çok ölçekli dağılım entropi öznitelikleri çıkarılmıştır. Daha sonra bu özellikler

birleştirildi ve derin özniteliklerin elde edilmesi için LSTM'ye girdi olarak uygulanmıştır. LSTM'den elde edilen bu derin özellikler, k-NN, EBT ve SVM gibi popüler makine öğrenme yöntemleri ile sınıflandırılmış ve konutlarda kullanılan elektrikli cihazlar tespit edilmiştir. Deneysel çalışmalarda 18 farklı ev aletini içeren Ev Aletlerinin Yüksek Çözünürlüklü Profilleri adlı veri seti kullanılmıştır. Deneysel çalışmalarda k-NN, EBT ve SVM sınıflandırıcılarına uygulanan LSTM derin öznitelikleri ile sırasıyla %98,25, %97,81 ve %97,38 doğruluk oranlarına ulaşılmıştır. Ayrıca LSTM derin özniteliklerinin kullanımı ile geleneksel makine öğrenme yöntemleri olan k-NN, SVM ve EBT sınıflandırıcıların sınıflandırma başarımları ile karşılaştırıldığında sırasıyla %3,06, %2,62 ve %0,74 daha yüksek doğruluk oranı elde edilmiştir.

Gelecekteki çalışmalarda, MOYİ yaklaşımları ile konutlarda kullanılan farklı çalışma fonksiyonlarına sahip cihaz türleri ve çalışma konumlarının belirlenmesi planlanmaktadır.

KAYNAKLAR

- [1] IEA. Electricity information: Overview, IEA, Paris. 2020. [www. iea. org/reports/electricity-information-overview](http://www.iea.org/reports/electricity-information-overview). (eriřim tarihi: 20 Temmuz 2022)
- [2] Natarajan A, Krishnasamy V, Singh M. Occupancy detection and localization strategies for demand modulated appliance control in Internet of Things enabled home energy management system. *Renewable and Sustainable Energy Reviews*. 2022; 167: 112731.
- [3] Growth of Electricity Sector in India from 1947-2013. 2013. https://cea.nic.in/wp-content/uploads/pdm/2020/12/growth_2020.pdf.
- [4] Streltsov A, Malof JM, Huang B, Bradbury K. Estimating residential building energy consumption using overhead imagery. *Applied Energy*. 2020; 280.
- [5] Mohanty SP, Choppali U, Kougianos E. Everything you wanted to know about smart cities. *IEEE Consumer Electronics Magazine*. 2016; 5: 60–70.
- [6] Kim H, Choi H, Kang H, An J, Yeom S, Hong T. A systematic review of the smart energy conservation system: From smart homes to sustainable smart cities. *Renewable and Sustainable Energy Reviews*. 2021; 140: 110755.
- [7] Krarti M, Aldubyan M. Review analysis of COVID-19 impact on electricity demand for residential buildings. *Renewable and Sustainable Energy Reviews*. 2021; 143: 110888.
- [8] Bielecki S, Skoczowski T, Sobczak L, Buchoski J, Maciąg Ł, Dukat P. Impact of the Lockdown during the COVID-19 Pandemic on Electricity Use by Residential Users. *Energies*. 2021; 14: 980.
- [9] Pérez-Lombard L, Ortiz J, Pout C. A review on buildings energy consumption information. *Energy and Buildings*. 2008; 40: 394–8.
- [10] Yoon SH, Kim SY, Park GH, Kim YK, Cho CH, Park BH. Multiple power-based building energy management system for efficient management of building energy. *Sustainable Cities and Society*. 2018; 42: 462–70.
- [11] Liu H, Wu H, Yu C. A hybrid model for appliance classification based on time series features. *Energy and Buildings*. 2019; 196: 112–23.
- [12] Faustine A, Mvungi NH, Science C. A Survey on Non-Intrusive Load Monitoring

Methodies and Techniques for Energy Disaggregation Problem. arXiv.2017.

[13] Baets L De, Develder C, Dhaene T, Deschrijver D. Electrical Power and Energy Systems Detection of unidenti fi ed appliances in non-intrusive load monitoring using siamese neural networks. *Electrical Power and Energy Systems*. 2019; 104: 645–53.

[14] Liu Y, Liu C, Shen Y, Zhao X, Gao S, Huang X. Non-intrusive energy estimation using random forest based multi-label classification and integer linear programming. *Energy Reports*. 2021; 7: 283–91.

[15] Athanasiadis C, Doukas D, Papadopoulos T, Chrysopoulos A. A scalable real-time non-intrusive load monitoring system for the estimation of household appliance power consumption. *Energies*. 2021; 14: 767.

[16] Zhou G, Li Z, Fu M, Feng Y, Wang X, Huang C. Sequence-to-Sequence Load Disaggregation Using Multiscale Residual Neural Network. *IEEE Transactions on Instrumentation and Measurement*. 2021; 70.

[17] Kong W, Dong ZY, Wang B, Zhao J, Huang J. A practical solution for non-intrusive type II load monitoring based on deep learning and post-processing. *IEEE Transactions on Smart Grid*. 2020; 11: 148–60.

[18] Khan MMR, Siddique MAB, Sakib S. Non-Intrusive Electrical Appliances Monitoring and Classification using K-Nearest Neighbors. *ICIET 2019 - 2nd International Conference on Innovation in Engineering and Technology*. 2019.

[19] Akçin M, Kaygusuz A, Keleş C, Karabiber A, Alagöz BB. Akıllı Şebekelerde Kontrol ve Haberleşme : Günümüzden Geleceğe Fırsatlar Akıllı Şebekelerde Kontrol ve Haberleşme : Günümüzden Geleceğe Fırsatlar Elektrik-Elektronik Mühendisliği Bölümü. 2013.

[20] Nalmpantis C, Vrakas D. On time series representations for multi-label NILM. *Neural Computing and Applications*. 2020; 32: 17275–90.

[21] Gillis JM, Morsi WG. Non-intrusive load monitoring using wavelet design and co-testing of machine learning classifiers. 2016 IEEE Electrical Power and Energy Conference, EPEC 2016. 2016; 7: 320–8.

[22] Mubarok AF, Octavira T, Sudiharto I, Wahjono E, Anggriawan DO. Identification of harmonic loads using fast fourier transform and radial basis Function Neural Network. *Proceedings IES-ETA 2017 - International Electronics Symposium on Engineering Technology and Applications*. Vol2017-Decem. 2017: 198–202.

[23] Yan D, Jin Y, Sun H et al. Household appliance recognition through a Bayes

classification model. *Sustainable Cities and Society*. 2019; 46: 101393.

[24] De Baets L, Develder C, Dhaene T, Deschrijver D. Detection of unidentified appliances in non-intrusive load monitoring using siamese neural networks. *International Journal of Electrical Power and Energy Systems*. 2019; 104: 645–53.

[25] Hart GW. Nonintrusive Appliance Load Monitoring. *Proceedings of the IEEE*. 1992; 80: 1870–91.

[26] Buddhahai B, Wongseree W, Rakkwamsuk P. A non-intrusive load monitoring system using multi-label classification approach. *Sustainable Cities and Society*. 2018; 39: 621–30.

[27] Babaei T, Abdi H, Lim CP, Nahavandi S. A study and a directory of energy consumption data sets of buildings. *Energy and Buildings*. 2015; 94: 91–9.

[28] Kelly J, Knottenbelt W. The UK-DALE dataset, domestic appliance-level electricity demand and whole-house demand from five UK homes. *Scientific Data*. 2015; 2: 1–14.

[29] Kolter JZ, Johnson MJ. REDD: A Public Data Set for Energy Disaggregation Research. *Data Mining Applications in Sustainability*. 2011; 1–6.

[30] Gao J, Giri S, Kara EC, Bergés M. PLAID: a public dataset of high-resolution electrical appliance measurements for load identification research. *Proceedings of the 1st ACM Conference on Embedded Systems for Energy-Efficient Buildings*. New York, NY, USA: ACM, 2014: 198–9.

[31] Kryonidis GC, Barzegkar-Ntovom GA, Papadopoulos TA et al. A Bottom-up Modelling Approach for Household Power Profiles Using Time-series Measurements. *UPEC 2020 - 2020 55th International Universities Power Engineering Conference, Proceedings*. IEEE, 2020: 1–6.

[32] Joint Research Center. The smart grid interoperability laboratory. European Research Center, 2018. <https://ses.jrc.ec.europa.eu/sgil-petten>. (erişim tarihi: 20 Temmuz 2022)

[33] Siuly S, Alcin OF, Bajaj V, Sengur A, Zhang Y. Exploring Hermite transformation in brain signal analysis for the detection of epileptic seizure. *IET Science, Measurement and Technology*. 2019; 13: 35–41.

[34] Siuly S, Li Y. Improving the Separability of Motor Imagery EEG Signals Using a Cross Correlation-Based Least Square Support Vector Machine for Brain–Computer Interface. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*. 2012; 20: 526–38.

- [35] Karaatlı M. Entropi Gri İlişkisel Analiz Yöntemleri ile Bütünleşik Bir Yaklaşım:Turizm Sektöründe Uygulama. Süleyman Demirel Üniversitesi İktisadi ve İdari Bilimler Fakültesi Dergisi. 2016; 21: 63–77.
- [36] Shannon CE. A Mathematical Theory of Communication. Bell System Technical Journal. 1948; 27: 379–423.
- [37] Rostaghi M, Azami H. Dispersion Entropy: A Measure for Time-Series Analysis. IEEE Signal Processing Letters. 2016; 23: 610–4.
- [38] Samiee K, Kovács P, Gabbouj M. Epileptic seizure classification of EEG time-series using rational discrete short-time fourier transform. IEEE Transactions on Biomedical Engineering. 2015; 62: 541–52.
- [39] Gabor D. Theory of communication. part 1: The analysis of information, Journal of the Institution of Electrical Engineers-Part III: Radio and. Communication Engineering. 1946; 93: 429–441.
- [40] Ghaderpour E, Pagiatakis SD, Hassan QK. A survey on change detection and time series analysis with applications. Applied Sciences (Switzerland). 2021; 11: 6141.
- [41] Sejdić E, Djurović I, Jiang J. Time-frequency feature representation using energy concentration: An overview of recent advances. Digital Signal Processing: A Review Journal. 2009; 19: 153–83.
- [42] Aslan M. Derin Öğrenme ile Şeftali Hastalıklarının Tespiti. European Journal of Science and Technology. 2021: 540–6.
- [43] Alzubaidi L, Zhang J, Humaidi AJ et al. Review of deep learning: concepts, CNN architectures, challenges, applications, future directions. Journal of Big Data. 2021; 8: 53.
- [44] Lecun Y, Bengio Y, Hinton G. Deep learning. Nature. 2015; 521: 436–44.
- [45] Shrestha A, Mahmood A. Review of deep learning algorithms and architectures. IEEE Access. 2019; 7: 53040–65.
- [46] Kattenborn T, Leitloff J, Schiefer F, Hinz S. Review on Convolutional Neural Networks (CNN) in vegetation remote sensing. ISPRS Journal of Photogrammetry and Remote Sensing. 2021; 173: 24–49.
- [47] Turkoglu M, Alcin OF, Aslan M, Al-Zebari A, Sengur A. Deep rhythm and long short term memory-based drowsiness detection. Biomedical Signal Processing and Control. 2021; 65: 102364.

- [48] Aslan M. Derin Öğrenme Tabanlı Otomatik Beyin Tümör Tespiti. Fırat Üniversitesi Mühendislik Bilimleri Dergisi.2022 Feb 12; doi 10.35234/fumbd.1039825.
- [49] Hao W, Sun X, Wang C, Chen H, Huang L. A hybrid EMD-LSTM model for non-stationary wave prediction in offshore China. *Ocean Engineering*. 2022; 246: 110566.
- [50] Hochreiter S, Schmidhuber J. Long Short-Term Memory. *Neural Computation*. 1997; 9: 1735–80.
- [51] Demir F, Sengur A, Ari A, Siddique K, Alswaitti M. Feature Mapping and Deep Long Short Term Memory Network-Based Efficient Approach for Parkinson’s Disease Diagnosis. *IEEE Access*. 2021; 9: 149456–64.
- [52] Alhirmizy S, Qader B. Multivariate Time Series Forecasting with LSTM for Madrid, Spain pollution. 2019 International Conference on Computing and Information Science and Technology and Their Applications (ICCISTA). IEEE, 2019: 1–5.
- [53] Xu Y, Hu C, Wu Q et al. Research on Particle Swarm Optimization in LSTM Neural Networks for Rainfall-Runoff Simulation. *Journal of Hydrology*. 2022: 127553.
- [54] Karasu S, Altan A. Crude oil time series prediction model based on LSTM network with chaotic Henry gas solubility optimization. *Energy*. 2022; 242: 122964.
- [55] Vapnik VN. The Nature of Statistical Learning Theory. *The Nature of Statistical Learning Theory*.2000; doi 10.1007/978-1-4757-3264-1.
- [56] Aslan M. CNN based efficient approach for emotion recognition. *Journal of King Saud University - Computer and Information Sciences*.2021; doi 10.1016/j.jksuci.2021.08.021.
- [57] Demir F, Siddique K, Alswaitti M, Demir K, Sengur A. A Simple and Effective Approach Based on a Multi-Level Feature Selection for Automated Parkinson’s Disease Detection. *Journal of Personalized Medicine*. 2022; 12.
- [58] Akbulut Y, Sengur A, Guo Y, Smarandache F. NS-k-NN: Neutrosophic Set-Based k-Nearest Neighbors Classifier. *Symmetry*. 2017; 9: 179.
- [59] Chairatanasongporn N, Jaiyen S. A hybrid ensemble of machine and statistical learning using confidence-based boosting. 2015 7th International Conference on Information Technology and Electrical Engineering (ICITEE). IEEE, 2015: 41–5.
- [60] Breiman L. Bagging predictors. *Machine Learning*. 1996; 24: 123–40.
- [61] Freund Y, Schapire RE. A Decision-Theoretic Generalization of On-Line Learning

and an Application to Boosting. *Journal of Computer and System Sciences*. 1997; 55: 119–39.

[62] Mishra A, Reddy US. A comparative study of customer churn prediction in telecom industry using ensemble based classifiers. *Proceedings of the International Conference on Inventive Computing and Informatics, ICICI 2017*. 2018: 721–5.

[63] Paiva Penha D de, Garcez Castro AR. Home Appliance Identification for Nilm Systems Based on Deep Neural Networks. *International Journal of Artificial Intelligence & Applications*. 2018; 9: 69–80.

[64] Le TTH, Kang H, Kim H. Household appliance classification using lower odd-numbered harmonics and the bagging decision tree. *IEEE Access*. 2020; 8: 55937–52.

[65] De Baets L, Ruysinck J, Develder C, Dhaene T, Deschrijver D. Appliance classification using VI trajectories and convolutional neural networks. *Energy and Buildings*. 2018; 158: 32–6.

[66] Devlin MA, Hayes BP. Non-Intrusive Load Monitoring and Classification of Activities of Daily Living Using Residential Smart Meter Data. *IEEE Transactions on Consumer Electronics*. 2019; 65: 339–48.

[67] Athanasiadis CL, Papadopoulos TA, Doukas DI. Energy & Buildings Real-time non-intrusive load monitoring : A light-weight and scalable approach. *Energy & Buildings*. 2021; 253: 111523.

[68] Alcalá J, Urena J, Hernandez A, Gualda D. Event-Based Energy Disaggregation Algorithm for Activity Monitoring from a Single-Point Sensor. *IEEE Transactions on Instrumentation and Measurement*. 2017; 66: 2615–26.