Journal of the Faculty of Engineering and Architecture of Gazi University 41:1 (2026) 383-399

Mühendislik Mimarlık Fakültesi Dergisi
Journal of The Faculty of Engineering and Architecture of Gazi University

Elektronik / Online ISSN : 1304 - 4915
Basılı / Printed ISSN : 1300 - 1884

Signal-level network traffic classification using darknet-based CNNs: A new methodological approach

Munip Geylani[1]*, Musa Çıbuk[2], Ayhan Akbal[3]

[1]Common Courses Department, Bitlis Eren University, 13100, Bitlis, Türkiye
[2]Department of Computer Engineering, Faculty of Engineering and Architecture, Bitlis Eren University, 13100, Bitlis, Türkiye
[3]Department of Electrical - Electronics Engineering, Faculty of Engineering, Fırat University, 23119, Elazığ, Türkiye

Highlights:

- Introduction of a novel approach for network traffic classification at signal level
- Detailed presentation of a newly constructed signal-based dataset
- Comparison of different visualization techniques with Darknet19 and Darknet53 architectures

Keywords:

- Network traffic classification
- Signal-level traffic classification
- Physical layer
- Signal visualization
- Darknet

Article Info:

Research Article
Received: 11.08.2025
Accepted: 12.12.2025

DOI:

10.17341/gazimmfd.1761166

Correspondence:

Author: Munip Geylani
e-mail: mgeylani@beu.edu.tr
phone: +90 506 951 0911

Graphical/Tabular Abstract

In contrast to traditional approaches, this study addresses network traffic classification at the signal level. A custom dataset was constructed by capturing raw electrical signals corresponding to network packets using an oscilloscope at the physical layer. These time-domain signals were then transformed into image-based datasets. The resulting visual representations were used to fine-tune two pre-trained convolutional neural network architectures, Darknet19 and Darknet53, through transfer learning. The proposed method demonstrates that signal-level features preserve meaningful traffic characteristics and can be effectively leveraged by deep learning models for accurate classification. The proposed methodological approach is given in Figure A.

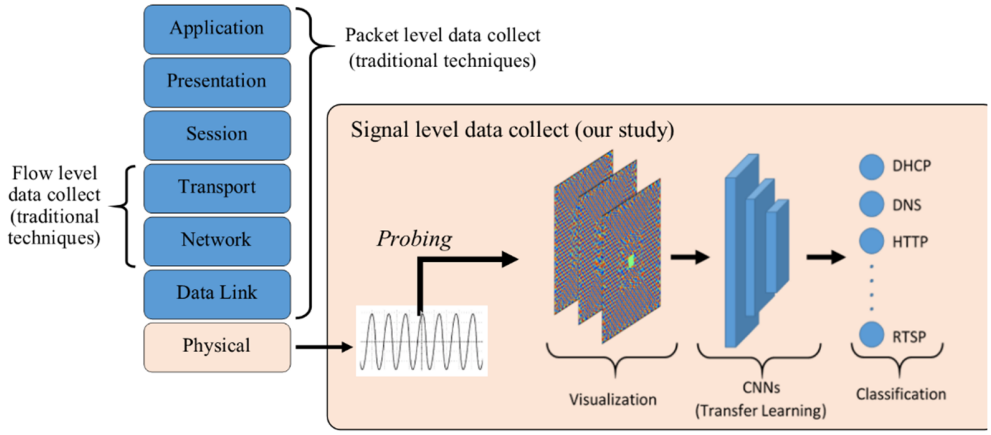


Figure A. Proposed signal level network traffic classification approach

Purpose:
This study aims to perform network traffic classification at the signal level by analyzing raw electrical signals captured from the physical layer.

Theory and Methods:
Six types of network protocols were selected, and corresponding packet data was collected. These packets were retransmitted in a controlled network, and their electrical signals were captured from the physical layer using an oscilloscope. Each signal was matched with its corresponding packet via timestamp alignment. The resulting packet-level signals were visualized using four techniques: horizontal, spiral, diagonal zigzag, and spectrogram. These images were then used to fine-tune Darknet19 and Darknet53 models via transfer learning for classification.

Results:
The proposed method achieved a highest classification accuracy of 96.24% using the Darknet53 model combined with the diagonal zigzag visualization technique, demonstrating effective learning of signal-level traffic features.

Conclusion:
This study demonstrates that network traffic classification can be effectively performed at the signal level using deep learning. The proposed approach offers a promising alternative to traditional network traffic classification techniques.

Journal of the Faculty of Engineering and Architecture of Gazi University 41:1 (2026) 383-399

Mühendislik Mimarlık Fakültesi Dergisi
Journal of The Faculty of Engineering and Architecture of Gazi University

Elektronik / Online ISSN : 1304 - 4915
Basılı / Printed ISSN : 1300 - 1884

Darknet tabanlı CNN'ler ile sinyal seviyesinde ağ trafiği sınıflandırması: Yeni bir yöntemsel yaklaşım

Munip Geylani[1]*, Musa Çıbuk[2], Ayhan Akbal[3]

[1]Bitlis Eren Üniversitesi, Ortak Dersler Bölümü, 13100, Bitlis, Türkiye
[2]Bitlis Eren Üniversitesi, Mühendislik Fakültesi, Bilgisayar Mühendisliği Bölümü, 13100, Bitlis, Türkiye
[3]Fırat Üniversitesi, Mühendislik Fakültesi, Elektrik-Elektronik Mühendisliği Bölümü, 23119, Elazığ, Türkiye

ÖNEÇIKANLAR

- Ağ trafiği sınıflandırmasının sinyal seviyesinde yapılmasıyla yeni bir yöntem sunulması
- Yeni bir sinyal tabanlı veri seti oluşturma adımlarının detaylandırılması
- Farklı görselleştirme tekniklerinin Darknet19 ve Darknet53 mimarileri ile karşılaştırılması

Makale Bilgileri
Araştırma Makalesi
Geliş: 11.08.2025
Kabul: 12.12.2025

DOI:
10.17341/gazimmfd.1761166

Anahtar Kelimeler:
Ağ trafiği sınıflandırması, sinyal seviyesinde ağ trafiği sınıflandırma, fiziksel katman, sinyal görselleştirme, Darknet

ÖZ

Ağ trafiği sınıflandırması, araştırmacılar tarafından yoğun biçimde çalışılan önemli bir alandır. Mevcut çalışmaların genellikle veri bağı katmanı ve üzerindeki katmanlarda, trafiğin çözümlenmiş (decode edilmiş) hali üzerinde yapıldığı görülmektedir. Ancak literatürde, ağ trafiğinin doğrudan fiziksel katman düzeyinde, yani sinyallerin kendisi üzerinden sınıflandırılmasına yönelik sistematik bir yaklaşım bulunmamaktadır. Bu çalışmada, ağ trafiği sınıflandırmasına fiziksel katman düzeyinde yeni bir yöntemsel yaklaşım sunulmaktadır. Önerilen yöntemde, ağ trafiği elektriksel sinyaller aracılığıyla temsil edilmekte ve sınıflandırma işlemi bu ham sinyaller üzerinden gerçekleştirilmektedir. Bu amaçla, bilgisayar ile ağ anahtarlama cihazı arasındaki fiziksel bağlantıdan osiloskop yardımıyla sinyaller toplanmış ve etiketlenerek altı farklı trafik türünü içeren yeni bir veri seti oluşturulmuştur. Oluşturulan sinyal veri seti, yatay, spiral, diyagonal zikzak ve spektrogram teknikleriyle görselleştirilerek Darknet tabanlı CNN mimarilerinin eğitilmesinde kullanılmıştır. Darknet53 ve diyagonal zikzak görselleştirme tekniğiyle yapılan sınıflandırma deneyinde %96,24 doğruluk oranı ile en yüksek performans elde edilmiştir. Elde edilen sonuçlar, paket içerikleri çözümlenmeden ve veri mahremiyeti ihlal edilmeden, ağ trafiğinin yalnızca sinyal düzeyinde dahi yüksek doğrulukla sınıflandırılabileceğini göstermektedir. Bu yönüyle çalışma, literatürdeki önemli bir boşluğu doldurarak geleneksel yöntemlere güçlü bir alternatif sunmaktadır.

Signal-level network traffic classification using darknet-based CNNs: A new methodological approach

HIGHLIGHTS

- Introduction of a novel approach for network traffic classification at the signal level
- Detailed presentation of a newly constructed signal-based dataset
- Comparison of different visualization techniques with Darknet19 and Darknet53 architectures

Article Info
Research Article
Received: 11.08.2025
Accepted: 12.12.2025

DOI:
10.17341/gazimmfd.1761166

Keywords:
Network traffic classification, signal-level traffic classification, physical layer, signal visualization, Darknet

ABSTRACT

Network traffic classification is a significant research area that has been extensively studied by researchers. Existing studies are mostly conducted at the data link layer or higher layers, typically using the decoded form of the traffic. However, there is no systematic approach in the literature that performs traffic classification directly at the physical layer using raw electrical signals. In this study, a novel classification approach is proposed at the physical layer level. In the proposed method, network traffic is represented through electrical signals, and the classification process is performed directly on these raw signals. To this end, signals were collected from the physical connection between a computer and a network switch using an oscilloscope, and a new dataset consisting of six different traffic types was constructed by labeling these signals. The created signal dataset was visualized using horizontal, spiral, diagonal zigzag, and spectrogram techniques, and was then used to train Darknet-based CNN architectures. The highest performance was achieved in the classification experiment using Darknet53 and the diagonal zigzag visualization technique, with an accuracy rate of 96.24%.The results demonstrate that network traffic can be effectively classified at the signal level, without decoding packet contents and without compromising data privacy. In this respect, the study fills a notable gap in the literature and provides a strong alternative to traditional traffic classification methods.

*Sorumlu Yazar / Yazarlar / *Corresponding Author / Authors: *mgeylani@beu.edu.tr, mcibuk@gmail.com, ayhan_akbal@firat.edu.tr / Tel: +90 506 951 0911

1. Giriş (Introduction)

Ağ trafiği sınıflandırma, güvenli ve verimli bir iletişim ortamı sunulmasında internet servis sağlayıcıları (ISS), kamu kurumları veya özel şirketler için oldukça önemli bir konudur. Kuruluşlar, trafik önceliklendirme, trafiğe izin verme ve yasaklama, trafik şekillendirme ve bant genişliği paylaşımı sağlama gibi ağ yönetim görevlerini yerine getirmek amacıyla kendi ağ trafiklerini izleyip analiz etmektedirler. Ayrıca, ağ güvenliği, dinamik erişim kontrolü ve saldırı tespiti gibi internet ağ güvenliği sağlama işlemleri için de sınıflandırmaya ihtiyaç duyulmaktadır. Yapılan sınıflandırmalar ile ağ trafiği regüle edilerek daha etkin ve verimli bir iletişim sağlanabilir. Bu nedenlerle ağ trafiği sınıflandırma, araştırmacılar için önemli bir çalışma alanı olarak karşımıza çıkmaktadır [1]. Ağ trafiğini sınıflandırma için kullanılan teknikleri üç başlık altında toplamak mümkündür [2-5];

- Port tabanlı sınıflandırma;
- Derin paket incelemesi (Deep Packet Inspection-DPI) veya yüke dayalı sınıflandırma;
- İstatistiksel ve Makine Öğrenimi Tabanlı sınıflandırma;

Port tabanlı sınıflandırma; Internet Assigned Numbers Authority (IANA), internet iletişiminin küresel çapta düzgün çalışabilmesi için belli protokol ve uygulamalar için port ataması yapmaktadır. Sınıflandırma işlemi genellikle, IANA tarafından önceden tanımlanmış portlarla eşlenen İletim Kontrol Protokolü (TCP) ve Kullanıcı Datagram Protokolü (UDP) paketlerinin port numaralarının incelenmesiyle gerçekleştirilir [5, 6].

Yüke dayalı sınıflandırma; internet paketi içerisindeki veriye (yüke) bakılarak sınıflandırma yapılır. Bu sınıflandırmayı yapabilmek için tüm protokoller ve verileri hakkında derinlemesine bilgi gerektirmektedir. Porta dayalı tekniğe göre daha karmaşık ve daha fazla sınıflandırma yapabilmektedir. Ancak buna karşın çok daha fazla işlem gücüne ve bellek kaynaklarına ihtiyaç duyulmaktadır. Bu da ilgili sistemlerin maliyetlerinin çok artmasına neden olmaktadır [5, 7].

İstatiksel ve Makine öğrenimi tabanlı sınıflandırma; Genel olarak ağ üzerindeki akış bilgisinin toplanarak daha çok istatistiki yöntem ve algoritmaların kullanılması ile sınıflandırma yapılır. Makine öğrenmesinin sınıflandırmadaki üstün başarısı network trafiği sınıflandırma çalışmalarında da yoğunlukla kullanılmasını beraberinde getirmiştir [8-10].

Ağ trafik sınıflandırma teknikleri inceleme ve veri seti oluşturma için veri toplama yönüyle ele alındığında, genel olarak paket bazlı ve akış (flow) bazlı yaklaşımların olduğu görülmektedir [6, 7, 11, 12]. Paket tabanlı yaklaşımda, veri setleri her bir paketi bağımsız bir analiz birimi olarak ele alır [5]. Paketlerin başlık ve veri yükü (payload) kısımları analiz edilir. Akış tabanlı yaklaşımda ise veri setleri, ağ trafiğini beşli tuple (demet) bilgilerine göre gruplandırarak özetler. Akış tabanlı yaklaşımlarda, bir grup paketin ortak özelliklere sahip olduğu varsayılır ve bu özelliklere dayanarak akışlar oluşturulur. Akışlar paketlerin kaynak port, hedef port, kaynak ip, hedef ip ve protokol türlerinin benzerliklerine göre gruplanır. Her bir akış, bir trafik birimi olarak değerlendirilir. Paket tabanlı yaklaşımın alternatifi olarak akış tabanlı bir veri toplama mekanizması, tüm paketler yerine akışı tarar [11]. Makine öğrenmesi tekniğinde genellikle akış tabanlı verilerle çalışma yapıldığı görülmektedir [6].

Tüm bu tekniklerin ortak noktası, ağ trafiğinin dijital olarak çözülmüş ve mantıksal olarak yapılandırılmış hali üzerinde çalışmalarıdır. Başka bir deyişle, ilgili katmanlardan başlık ve veri yükü alanlarının ayrıştırıldığı ve bu alanlardan bilgi çıkarıldığı bir paket ayrıştırma (packet parsing) sürecinin ardından analiz yapılmaktadır. Bu nedenle, sınıflandırma işlemleri OSI modelinin ikinci ve daha üst katmanlarındaki verilere dayanır. Bu da, tüm paketlerin dekapsüle edilerek ilgili katmana kadar işlenmesini gerektirir. Öte yandan, şifrelenmemiş veriler bu süreçte kolaylıkla erişilebilir ve görünür hale gelir.

Günümüz ağ altyapısında, kablolu iletişimin büyük çoğunluğu Ethernet teknolojisine dayanmaktadır. Bu teknolojide Ethernet çerçeveleri, bilgisayarlar ve ağ cihazları arasında veri taşıyan temel iletişim birimleri olarak görev yapar. Fiziksel katmanda, bu çerçeveler kodlanır ve iletim ortamı üzerinden elektrik sinyalleri olarak iletilir. İster paketin başlığına isterse yük kısmına bakılsın, sınıflandırma yapmaya yarayan bilginin bu elektrik sinyalleri üzerinde de bir örüntü veya bir desen oluşturacağı hipotezi bu çalışmanın ana motivasyonunu oluşturmuştur.

Literatürdeki mevcut çalışmalar incelendiğinde, ağ trafiği sınıflandırmasının büyük ölçüde paket veya akış seviyesindeki verilere dayandığı görülmektedir. Bununla birlikte, ağ trafiğinin doğrudan fiziksel katmanda ve ham elektriksel sinyaller üzerinden sınıflandırılmasına yönelik sistematik bir çalışma bulunmamaktadır. Bu durum, sınıflandırma işleminin sinyal düzeyinde gerçekleştirilmesi ve böylece paket içeriklerinin açılmadan (decapsulation) değerlendirilebilmesi açısından önemli bir araştırma boşluğu oluşturmaktadır. Bu çalışma, network trafiği henüz elektrik sinyali iken ele alınıp sınıflandırmanın bu aşamada yapılması yönüyle diğer çalışmalardan ayrılmaktadır. Diğer bir ifade ile fiziksel katmanda sınıflandırma yapılmakta olup literatüre farklı bir bakış açısı ve derinlik kazandırılmaya çalışılmıştır.

Bu çalışmanın literatüre katkıları aşağıda özetlenmiştir:

- Ağ trafiğinin fiziksel katmanda, ham elektriksel sinyaller üzerinden sınıflandırılmasına yönelik yeni bir yaklaşım önerilmiştir.
- Bilgisayar ile ağ anahtarlama cihazı arasındaki fiziksel bağlantıdan toplanan sinyaller kullanılarak, ilk kez sinyal tabanlı bir ağ trafiği veri seti oluşturulmuştur. Veri seti oluşturma adımları benzer çalışmalara referans olması amacıyla ayrıntılı biçimde açıklanmıştır.
- Veri setindeki sinyaller yatay, spiral, diyagonal zikzak ve spektrogram olmak üzere dört farklı teknikle görselleştirilmiş; bu temsiller Darknet19 ve Darknet53 mimarilerinin eğitiminde kullanılmıştır. Sınıflandırma deneylerinde %96'nın üzerinde doğruluk elde edilmiş ve böylece hem sinyal düzeyinde sınıflandırmanın uygulanabilirliği hem de görselleştirme yöntemlerinin başarım üzerindeki etkisi sistematik olarak ortaya konmuştur
- Elde edilen sonuçlar, paket içeriği çözülmeden ve veri mahremiyeti korunarak ağ trafiğinin yüksek doğrulukla sınıflandırılabileceğini göstermiştir.

Bu çalışmanın ikinci bölümünde, ağ trafiği sınıflandırması alanında yapılan önceki çalışmalar incelenmiş ve literatürdeki mevcut yöntemler özetlenmiştir. Üçüncü bölümde, sinyal düzeyinde veri elde etme süreci detaylandırılmış ve önerilen sınıflandırma modeli açıklanmıştır. Dördüncü bölümde, farklı görselleştirme teknikleri ve evrişimli sinir ağı (CNN) mimarileri kullanılarak gerçekleştirilen deneysel çalışmalar sunulmuş; elde edilen performans sonuçları karşılaştırmalı olarak değerlendirilmiştir. Son olarak, beşinci bölümde çalışmanın genel bir değerlendirmesi yapılmış ve gelecekteki araştırmalara yönelik önerilerde bulunulmuştur.

2. İlgili Çalışmalar (Related Works)

Ağ trafiği sınıflandırması alanında birçok çalışma bulunmasına rağmen, problemi sinyal düzeyinde ele alan bir çalışmaya literatürde rastlanmamıştır. Bu bölümde, mevcut literatürde öne çıkan çalışmalara ve yöntemlerine yer verilmiştir. Makine öğrenmesi ve derin öğrenme alanındaki gelişmeler, bu yöntemlerin ağ trafiği sınıflandırmasında yaygın biçimde kullanılmasına yol açtığı görülmektedir. Ayrıca, mevcut çalışmaların büyük çoğunluğunun paket tabanlı ve akış tabanlı yaklaşımları benimsediği gözlemlenmektedir.

Fang vd. [13], paket tabanlı bir yaklaşımla uygulama katmanı protokollerini tanımlamak üzere ResNet tabanlı bir yöntem önermişlerdir. Yöntemde, ağ paketlerinin yük verisi 784 bayta sabitlenmiş; ardından, entropiye dayalı bir ayrıştırıcı algoritma ile öznitelik bloklarına ayrılmıştır. Bu bloklardan özel bir yöntemle 224×224 boyutunda RGB görseller üretilmiş ve önceden eğitilmiş ResNet18/34/50 modelleri ile sınıflandırma yapılmıştır. Deneylerde CIC-IDS2017 ve Shodan kaynaklı iki farklı veri seti kullanılmış; HTTP, DNS, FTP, SMB, TLS, SMTP, MQTT ve SSH gibi toplam 20 uygulama protokolü sınıflandırılmıştır. Önerilen yöntemin, %98'in üzerinde doğruluk oranı ile sınıflandırma yaptığı bildirilmiştir.

Chiu vd. [5], dinamik portlar ve şifreli trafik üzerinde etkili, paket tabanlı bir sınıflandırıcı olan CAPC (Convolutional Autoencoder Packet Classifier) modelini önermişlerdir. Model, 1D evrişimli katmanlarla desteklenen bir oto-kodlayıcı yapısı ile ardından gelen bir DNN (Derin Sinir Ağı) sınıflayıcıdan oluşur. Oto-kodlayıcı, paketin ham bayt verisini özet temsillere dönüştürerek DNN ile sınıflandırmayı mümkün kılar. Tüm paketler 1500 bayt uzunluğunda normalize edilir ve bayt değerleri 0–1 aralığına ölçeklenir. Deneylerde, araştırmacılar tarafından oluşturulan 16 hizmete ait özel bir veri seti ile 24 farklı hizmet türü içeren ISCX VPN-nonVPN veri seti kullanılmıştır. Sınıflandırılan hizmetler arasında dosya aktarımı, video akışı, VoIP ve uzaktan erişim uygulamaları (ör. FTP, Skype, YouTube, RDP) yer almaktadır. CAPC modelinin; özel veri setinde %99,98, ISCX VPN-nonVPN veri setinde ise %97,42 doğrulukla en yüksek performansı gösterdiği belirtilmiştir.

Lotfollahi vd. [14], Deep Packet adlı çalışmalarında, CNN tabanlı bir model ile hem uygulama tanımlama (17 uygulama) hem de şifreli trafik sınıflandırması (6 VPN ve 6 VPN olmayan olmak üzere 12 trafik türü) gerçekleştirmişlerdir. Çalışmada ISCX VPN-nonVPN veri seti kullanılmıştır. Derin öğrenme mimarisine giriş olarak, IP başlığı dâhil 1500 bayt uzunluğunda vektörler kullanılmıştır. Ham trafik verisi (PCAP dosyası) paket seviyesinde elde edilip, ön işleme sonrasında bayt vektörlerine dönüştürülerek CNN mimarisine beslenmiştir. Önerilen model, uygulama sınıflandırmasında %98, trafik sınıflandırmasında ise %94 doğruluk elde etmiştir.

Zhou vd. [15], ağ trafiği sınıflandırması için MMN-CNN adını verdikleri geliştirilmiş bir CNN modeli önermişlerdir. Akış tabanlı bu yaklaşımda, Moore veri setindeki trafik akışlarına ait 249 istatistiksel özellik Min-Max normalizasyonu ile 16×16 gri seviye görüntülere dönüştürülerek modele giriş olarak kullanılmıştır. Çalışmada WWW, MAIL, FTP-DATA, FTP-Control, FTP-Pasv, DATABASE, P2P, ATTACK, MULTIMEDIA, INTERACTIVE ve GAME olmak üzere toplam 12 trafik türü sınıflandırılmış; %99,30 genel doğruluk oranıyla başarılı sonuçlar elde edildiği bildirilmiştir.

Yamansavascilar vd. [16], ağ trafiğinden uygulama tanımlamak amacıyla J48, Random Forest, k-NN ve Bayes Net gibi dört farklı makine öğrenmesi algoritması kullanmışlardır. UNB ISCX ve kendi oluşturdukları veri seti ile gerçekleştirdikleri çalışmada, 5'li demet (tuple) yaklaşımıyla akışlara dönüştürdükleri verilerden elde edilen 111 özellik ile sınıflandırma yapılmış; ayrıca bu özelliklerin içinden seçilen 12 öznitelikle de karşılaştırmalı analiz gerçekleştirilmiştir. ISCX veri seti için her iki durumda benzer sonuçlar elde edilirken, kendi veri setlerinde seçilmiş 12 özellik ile sınıflandırma doğruluğunda %2 oranında artış rapor edilmiştir. Tüm özelliklerle yapılan sınıflamada ISCX veri seti için en yüksek doğruluk %93,94 ile k-NN, kendi veri setleri için ise %90,87 ile Random Forest algoritmasında elde edilmiştir. Seçilen 12 özellikle bu oran %92,99'a yükselmiştir.

Salman vd. [17], ISCX VPN ve TOR veri setleri üzerinde çalışarak CNN tabanlı çok seviyeli bir trafik sınıflandırma yöntemi önermişlerdir. Çalışmada dört seviyeli bir yapı kullanılmıştır: Seviye 1'de trafik; interactive, bulk data transfer, streaming ve transaction olarak dört ana kategoriye ayrılmıştır. Seviye 2'de bu kategoriler alt türlerine ayrılarak, örneğin interactive trafiği voice call, video call, texting ve gaming gibi alt sınıflara ayrılmıştır. Seviye 3, uygulama bazlı sınıflandırma; Seviye 4 ise cihaz bazlı sınıflandırma yapılmıştır. Akışlar 5'li tuple üzerinden elde edilerek, her paketten boyut, varış süresi, protokol ve yön bilgileri çıkarılmış ve bu özelliklerden paket vektörleri oluşturulmuştur. Bu vektörlerden 4×4 ve 28×28 boyutlarında iki farklı akış matrisi türetilmiş, RGBA formatında görüntülere dönüştürülerek LeNet5, AlexNet, ConvNet, GoogleNet ve ResNet gibi CNN mimarileriyle sınıflandırılmıştır. En yüksek başarı %95,84 doğruluk oranı ile 28×28 görüntüler ve ConvNet mimarisiyle elde edilmiştir.

Ahmed vd. [8], Netscrapper adını verdikleri bir trafik sınıflandırma modeli geliştirmişlerdir. Çalışmada, Amazon, YouTube, Google ve Twitter gibi popüler çevrimiçi uygulamaları sınıflandırmak amacıyla K-En Yakın Komşu (KNN), Rastgele Orman (RF) ve Yapay Sinir Ağı (ANN) algoritmaları kullanılmıştır. 78 farklı özelliğe sahip 3.577.296 akış içeren açık bir veri setiyle eğitilen model, canlı trafik üzerinde test edildiğinde; KNN ile %88,86, RF ile %96,33 ve ANN ile %99,86 doğruluk elde edilmiştir.

Lopez-Martin vd. [18], ulaşım katmanındaki 15 yaygın internet servisini sınıflandırmak için TCP ve UDP akışlarından 20 paket ve her paketten çıkarılan 6 özellik içeren bir veri seti oluşturmuşlardır. CNN, RNN ve CNN+RNN modelleriyle yapılan sınıflandırmada, HTTPS, DNS, SSL ve Google gibi servisler dâhil olmak üzere 15 trafik türü %99'un üzerinde doğrulukla sınıflandırılmıştır.

Al-Jameel vd. [19], çok katmanlı ileri beslemeli derin öğrenme (DL) tekniğine dayalı gerçek zamanlı video akışı trafiği sınıflandırıcısı geliştirmişlerdir. Wireshark ile yakalanan trafik verilerinden JSON uygulamasıyla 38 istatistiksel özellik çıkarılarak model eğitilmiştir. Önerilen DL modeli, Amazon Prime, Netflix ve YouTube videolarını %98,4 doğrulukla sınıflandırmıştır. Ayrıca, Naive Bayes makine öğrenmesi sınıflandırıcısı ile test edilen model %96,9 doğruluk oranı sağlamıştır.

Aouini vd. [20], modern internet hizmetlerini tanımlamak için C5.0 makine öğrenimi algoritmasına dayalı bir sınıflandırma yöntemi önermişlerdir. Çalışmada, her akışın ilk 4 paketinin istatistiksel özellikleri kullanılarak Facebook, Google Services, Skype, BitTorrent, Web-Browsing ve Secure-WebBrowsing gibi uygulamalar %98,8 doğrulukla sınıflandırılmıştır. Veri seti, 34.000'den fazla konut müşterisinin gerçek trafiğini içermektedir.

Fauvel vd. [21], akış tabanlı bir yaklaşımla şifreli internet trafiği sınıflandırması için LEXNet (Lightweight, Efficient and eXplainable-by-design CNN) adını verdikleri hafif, verimli ve açıklanabilir bir CNN modeli önermiştir. Model, paket boyutu ve yön bilgisinden oluşan 20×2 boyutunda çok değişkenli zaman serisi girdilerini kullanmaktadır. LEXNet, ResNet tabanlı yeni bir hafif artık blok

(LERes) ve sınıf bazlı prototipleri otomatik öğrenebilen yeni bir prototip katmanı (LProto) içermektedir. Çalışmada Huawei tarafından toplanan 9,7 milyon akış ve 200 uygulama türü içeren ticari ölçekli AppClassNet veri seti kullanılmıştır. Değerlendirmeler sonucunda, LEXNet'in mevcut açıklanabilir CNN modellerine kıyasla daha az parametreyle, daha hızlı çalıştığı ve %89,7 doğruluk sağladığı rapor edilmiştir. Ayrıca modelin saniyede 10.000 akış sınıflandırabildiği, dolayısıyla yönlendirici gibi ağ cihazlarına uygulanabilir olduğu vurgulanmıştır.

Wang vd. [22], şifreli trafik sınıflandırması için SwinT-CNN adını verdikleri hibrit bir derin öğrenme modeli önermiştir. Model, CNN modülü ile yerel uzamsal özellikleri, Swin Transformer modülü ile de çoklu başlıklı dikkat mekanizması sayesinde küresel bağıntıları öğrenmektedir. Çalışmada ISCXVPN2016 veri seti kullanılmış; oturum seviyesinde elde edilen trafik akışları 784 bayta sabitlenmiş ve 28×28 boyutunda gri seviye görüntülere dönüştürülerek modele beslenmiştir. E-posta, dosya aktarımı, VoIP, P2P, sohbet, video akışı gibi 12 trafik türünün (VPN ve non-VPN versiyonlarıyla birlikte) sınıflandırıldığı model ile %96,7 doğruluk oranına ulaştığı rapor edilmiştir.

Zheng vd. [23], ağ trafiği sınıflandırması için çoklu görünüm (multi-view) ve çoklu etiketleme (multi-label) yaklaşımını birleştiren, MLP-Mixer tabanlı yeni bir derin öğrenme modeli önermiştir. Çalışmada, her paket bayt düzeyinde kodlanarak 1500 uzunluğunda vektörlere dönüştürülmüş, ayrıca akış düzeyinde 8 istatistiksel özellik çıkarılmıştır. Modelde paket başlıkları, yük ve akış istatistikleri olmak üzere üç farklı görünümden öznitelik çıkarımı yapılmış; bu öznitelikler MLP-Mixer katmanlarıyla hiyerarşik olarak birleştirilmiştir. Çoklu etiketleme mekanizması sayesinde aynı veri hem anomali düzeyinde hem de trafik türü düzeyinde sınıflandırılabilmiştir. Deneylerde ISCXVPN2016, ISCXTor2016 ve USTC-TFC2016 veri setleri kullanılmış; önerilen modelin 1D-CNN ve yalnızca Mixer veya MLP tabanlı yöntemlere kıyasla daha yüksek doğruluk sağladığı rapor edilmiştir. Özellikle dengesiz veri setlerinde performans kaybı minimum düzeyde kalmış ve doğruluk oranı %99'un üzerine çıkmıştır.

Yang vd. [24], hem paket hem de akış seviyesindeki özellikleri birlikte kullanarak şifreli trafik sınıflandırmasına yönelik DM-HNN (Dual-Mode Hybrid Neural Network) adını verdikleri hibrit bir model önermiştir. Çalışmada, akış seviyesinde paket uzunluğu dizisi, paket seviyesinde ise paketlerin ilk 40 baytı öznitelik olarak alınmıştır. Bu öznitelikler sırasıyla GRU (Gated Recurrent Unit) ve SAE (Stacked Autoencoder) ağlarında işlenmiş, ardından özel bir füzyon katmanında birleştirilerek sınıflandırma yapılmıştır. Deneylerde ISCXVPN2016 ve ISCXTor2016 veri setleri kullanılmış, toplamda 53 uygulama türü sınıflandırılmıştır. VPN, NonVPN, Tor ve NonTor olmak üzere dört veri grubunda yapılan deneylerde DM-HNN modeli %91–%96,42 doğruluk aralığında performans göstermiştir.

Bunlara ek olarak, makine öğrenmesi ve derin öğrenme teknikleri kullanılarak anomali tespiti [25–28], saldırı tespiti [29-32] ve şifreli trafik analizi [33-35] üzerine yapılan çalışmalar da mevcuttur.
Son olarak, Abbasi vd. [3], ağ trafiği sınıflandırma ve analiz çalışmalarında derin öğrenme tekniklerini kullanan 90'a yakın çalışmayı ele alarak kapsamlı bir karşılaştırmalı analiz sunmaktadırlar.

3. Önerilen Yöntem (Proposed Methodology)

Bu bölümde, sinyal düzeyinde veri kümesi oluşturma adımları ile önerilen metodoloji kapsamında kullanılan Darknet tabanlı evrişimli sinir ağı (CNN) mimarileri sunulmaktadır. Öncelikle sınıflandırılması hedeflenen protokol türleri belirlenmiş ve bu protokollere ait paket verileri toplanmıştır. Bir dizi ön işlemden geçirilen bu paketler yeniden ağa gönderilmiş ve her bir pakete ait elektriksel sinyaller osiloskop yardımıyla yakalanmıştır. Elde edilen sinyaller görselleştirilerek CNN mimarilerinin eğitilmesinde kullanılmıştır.

3.1. Veri Seti Oluşturulması (Dataset Construction)

Önerdiğimiz sinyal düzeyinde network trafik sınıflandırması yapan bir çalışmaya literatürde denk gelinmediği gibi bu alanda kullanabileceğimiz bir veri setine rastlanılmamıştır. Bu nedenle; Şekil 1'de verilen işlem adımları uygulanarak yeni ve özgün bir veri seti oluşturulmuştur.

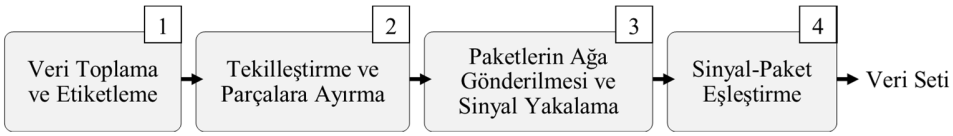
3.1.1. Veri toplama ve etiketleme (Data collection and labeling)

İlk olarak, kampüs ağlarında karşılaşılması muhtemel ve çoğunluğu uygulama katmanına ait olan altı farklı ağ protokolü belirlenmiştir: DHCP, DNS, HTTP, ICMP, RTSP ve TLS. Veri setinin gerçekçi bir yapıya sahip olması ve çok kullanıcılı trafik çeşitliliğini yansıtabilmesi amacıyla, çok kullanıcılı ortamlardan ağ trafiği örnekleri toplanmaya çalışılmıştır. Bu kapsamda, Bitlis Eren Üniversitesi (BEÜ) ağ altyapısında yer alan güvenlik duvarının web tabanlı arayüzü kullanılarak farklı ağ arabirimleri seçilmiş ve bu arabirimler üzerinden paket yakalama işlemi gerçekleştirilmiştir. Tüm protokol verileri ".pcap" formatında kaydedilmiştir. Elde edilen PCAP dosyaları, Wireshark yazılımı kullanılarak hedef protokole göre filtrelenmiş ve yalnızca ilgili protokole ait paketlerin yer alması sağlanmıştır. Etiketleme amacıyla, her PCAP dosyası içerdiği protokol türüne göre adlandırılmıştır (örneğin: {$protokol}.pcap) ve sonraki işlem adımlarına aktarılmıştır.

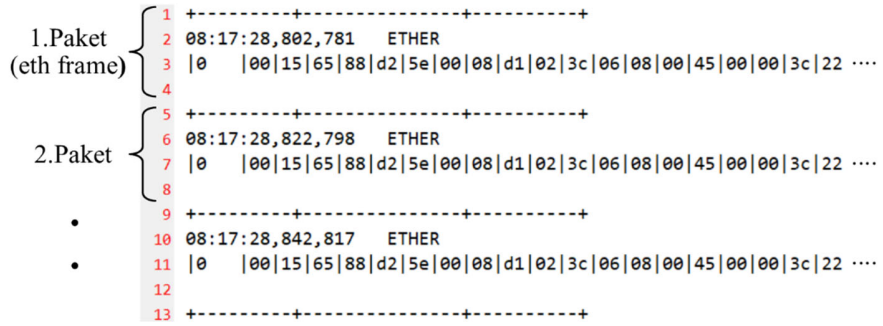
3.1.2. Tekilleştirme ve Parçalara Ayırma (Deduplication and segmentation)

Bu işlem adımında önceki adımda elde edilen PCAP dosyaları, MATLAB (2024a) kullanılarak yük (payload) tabanlı bir tekilleştirme (deduplication) işlemine tabi tutulmuştur. Aynı yüke sahip ve tekrar eden paketler tespit edilerek değerlendirme dışı bırakılmıştır. Böylece veri kümesinde yalnızca benzersiz paketlerin kalması sağlanarak CNN mimarilerinin eğitimi sırasında oluşabilecek öğrenme yanlılıkları ve ezberleme olasılıkları azaltılmıştır. Tekilleştirmiş paketleri içeren PCAP dosyaları, sinyal yakalama için kullanılan osiloskopun kayıt tutma kapasitesine göre uygun büyüklükte parçalara bölünmüştür.

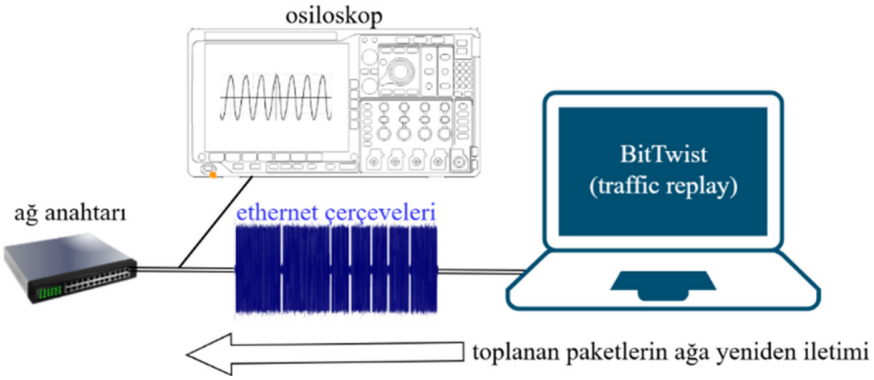
PCAP dosyalarının MATLAB ortamında işlenebilir hale getirilmesi amacıyla, Wireshark'ın komut satırı ara yüzü olan Tshark kullanılmış ve dosyalar K12 metin formatına dönüştürülmüştür. Şekil 2'de bir örneği sunulan bu K12 dosyaları, her bir paketin içeriğini hexadecimal (onaltılık) biçimde gösteren kendine özgü yapıda metin dosyalarıdır. Buna ek olarak, yine Tshark aracılığıyla yalnızca yük kısımlarını içeren ayrı metin dosyaları da üretilmiştir. Bu dosyalarda ise veriler düz hexadecimal formatta yer almaktadır.



Şekil 1. Veri seti oluşturma işlem adımları (Dataset construction process)



Şekil 2. K12 metin dosyası örneği (K12 text file example)



Şekil 3. Sinyal kayıt düzeneği (Signal capture setup)

Tablo 1. Osiloskop yapılandırma parametreleri (Oscilloscope configuration parameters)

Parametre	Değer
Örnekleme Modu (Sample Mode)	Hi-Res
Örnekleme Oranı (Sample Rate)	100MS/s (saniyede 100 milyon örnek)
Kayıt Uzunluğu	20M (20 milyon örnek)
İletişim Yolu Türü (Bus)	Ethernet
Giriş (Define Input)	10BASE-T
Eşik Değeri (Threshold)	Diff Probe ±1,2 V

Yük dosyaları MATLAB ile analiz edilerek aynı içeriğe sahip paketler tespit edilmiş ve karşılık gelen paketler K12 dosyalarından çıkarılmıştır (her paket 4 satır olarak temsil edilmektedir). Bazı ICMP paketlerinde yük bulunmadığından, bu protokole özgü olarak paketler, ICMP başlık alanından itibaren karşılaştırılarak tekilleştirme işlemi gerçekleştirilmiştir. Sonuç olarak, tekilleştirilmiş ve parçalara ayrılmış paket dosyaları {$protokol}_{partNo}_part.pcap formatında adlandırılarak sonraki işlem adımları için hazır hale getirilmiştir.

3.1.3. Paketlerin Ağa Gönderilmesi ve Sinyal Yakalama
(Replay packets to the network and signal capture)

Bu adımda, her bir *part.pcap dosyası ağa tekrar iletilmiş ve karşılık gelen elektriksel sinyaller osiloskop aracılığıyla kaydedilmiştir. Mevcut donanım kısıtlarından dolayı 10Base-T standartlarıyla çalışılmıştır. Ağ üzerinde başka trafik oluşmasını önlemek amacıyla izole bir ağ ortamı oluşturulmuştur. Bu amaçla, teknik özellikleri aşağıda verilen donanımlar kullanılarak, Şekil 3'te gösterilen deneysel düzenek kurulmuştur.

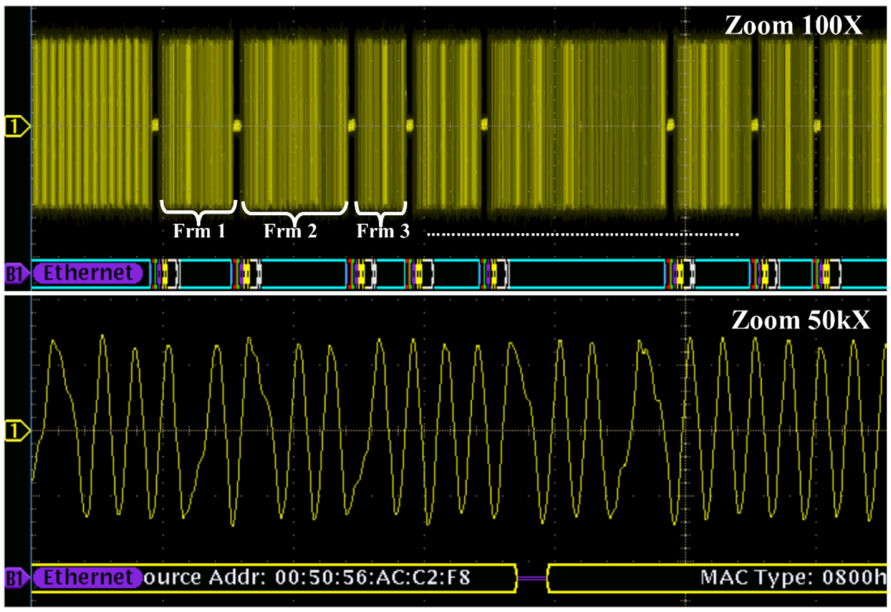
- Dizüstü Bilgisayar: HP, Intel Core i7-5600U CPU, 16 GB RAM
- Osiloskop: Tektronix MDO4104-6, DPO4ENET modülü ile birlikte
- Prob: Tektronix TDP0500 diferansiyel prob (CH1 kanalına bağlı)
- Ağ Anahtarı ve Kablo: HP 2510-24 switch ve Cat6 U-UTP kablo

Bilgisayar ile switch arasındaki bağlantı 10 Mbps, full-duplex olarak yapılandırılmıştır. CAT6 kablo üzerinde probun "+" ucu yeşil-beyaz kabloya, "–" ucu ise yeşil kabloya bağlanmıştır. Diferansiyel prob osiloskobun 1. kanalına (CH1) bağlanmıştır. Sinyal yakalama işlemleri sırasında osiloskop, Tablo 1'de belirtilen parametrelere göre yapılandırılmıştır. Bu yapılandırmada her bir bit, 10 örnekle temsil edilmektedir.

DPO4ENET modülü, osiloskobun Ethernet çerçevelerini çözümlemesini ve temel hata denetimlerini gerçekleştirmesini sağlamaktadır. Bu modül, çözümlediği çerçeveleri zaman sıralı bir olay tablosu şeklinde CSV formatında dışa aktarabilmektedir. DPO4ENET modülünün bu özelliği sayesinde, osiloskop tarafından yakalanan Ethernet sinyalleri ile PCAP dosyalarındaki paketlerin doğru ve güvenilir bir şekilde eşleştirilmesi mümkün olmaktadır.

Paketlerin ağa yeniden gönderim işlemi, açık kaynak kodlu Bit-Twist [36] aracı kullanılarak gerçekleştirilmiştir. Her part.pcap dosyası döngüsel (loop) olarak tekrar oynatılmış, böylece osiloskop kayıt ekranında tüm paketlerin eksiksiz şekilde yer alması sağlanmıştır. Şekil 4'te osiloskobun kayıt ekranı örnekleri gösterilmiştir.

Osiloskopla yapılan her kayıt öncesinde, diferansiyel prob üzerinde auto-zero kalibrasyon işlemi uygulanmıştır. Gerçekleştirilen her bir kayıt için osiloskoptan iki ayrı dosya elde edilmiştir:



Şekil 4. Osiloskop kayıt ekranı örnekleri (100X ve 50kX zoom değerleri ile)
(Oscilloscope capture screens with 100x and 50kx zoom factors)

- CH1 diye adlandırılan, 20 milyon örnek noktası (sample) içeren ham sinyal dosyası ({$protokol}{$partNo}_CH1.csv),
- ETH diye adlandırılan, yakalanan sinyalin osiloskop tarafından çözümlenmiş (dekode edilmiş) bilgilerini içeren olay tablosu (event table) dosyası ({$protokol}{$partNo}_ETH.csv),

Bu dosyalar, sonraki adımda gerçekleştirilen paket-sinyal eşleştirme sürecinde kullanılmıştır.

3.1.4. Sinyal-Paket Eşleştirme (Signal-packet matching)

Bu adımda gerçekleştirilen işlemler MATLAB ortamında gerçekleştirilmiştir. Bir önceki adımda part.pcap dosyalarının döngüsel (loop) biçimde tekrar oynatılması nedeniyle, bazı paketlerin osiloskop tarafından birden fazla yakalanmış olma ihtimali doğmaktadır. Bu nedenle bu aşamada ilk olarak, ETH (olay tablosu) dosyaları üzerinde içerik tabanlı tekilleştirme işlemi gerçekleştirilmiştir. Ardından ETH dosyalarındaki paketler, ilgili part.pcap dosyalarındaki paketlerle eşleştirilmiştir. Eşleştirmede part.pcap dosyalarının metin dosyası formatı olan K12 dosyaları kullanılmıştır. Eşleşen her paket için, sinyalin ilgili kısmı CH1 dosyasından (20 milyon örnek içeren tam kayıt) kırpılarak çıkarılmış ve ({$protokol}_{$partNo_{$SignalNo}_signal.csv formatında kaydedilmiştir.

Paketlere karşılık gelen sinyallerin doğru biçimde çıkarılabilmesi için, her sinyal parçasının başlangıç ve bitiş noktalarının belirlenmesi gerekmektedir. Bu işlem için aşağıda belirtilen işlem adımları uygulanmıştır:

- Öncelikle Eş. 1 kullanılarak CH1 sinyal dalga formunun 2 kez zarfı alınmıştır. Bu işlemde, 5 örnek uzunluğunda kayan pencere ile üst karekök ortalama (RMS – Root-Mean-Square) zarfı tekniği uygulanmıştır. Bu sayede Şekil 5 (b)'den de görüleceği üzere paketler arası boşluklar iyice belirginleşmiş ve bir eşik değeri yardımıyla kolayca tespit edilebilir hale getirilmiştir.

$$y[n] = \sqrt{\frac{1}{L}\sum_{k=0}^{L-1}\left(x\left[n-\left[\frac{L}{2}\right]+k\right]\right)^2} \quad (1)$$

Burada L: kayan pencere uzunluğudur.

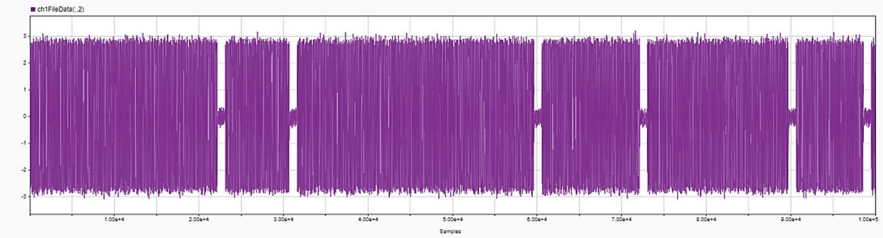
- Zarfı alınan sinyale uygulanan eşik değeri sayesinde sinyaldeki geçiş noktaları tespit edilmiş; bu noktalar sinyal parçalarının olası başlangıç ve bitiş noktalarını temsil etmektedir (Şekil 5(c)).

Bu başlangıç ve bitiş noktaları kullanılarak, paketlere ait sinyaller CH1 sinyal dosyasından kırpılarak çıkarılabilmektedir. Ayrıca, paketler arası boşlukları temsil edebilmek amacıyla, her sinyal parçasının bitiş noktasından itibaren 80 örnek (sample) sinyale dâhil edilmiştir. Şekil 6, üç farklı sinyalin (DHCP_01_0001, DHCP_01_0002 ve DHCP_01_0003_sinyal.csv) son 1000 örneğini göstermektedir. Paketler arası 80 örneklik boşluklar kırmızı renkle işaretlenmiştir.

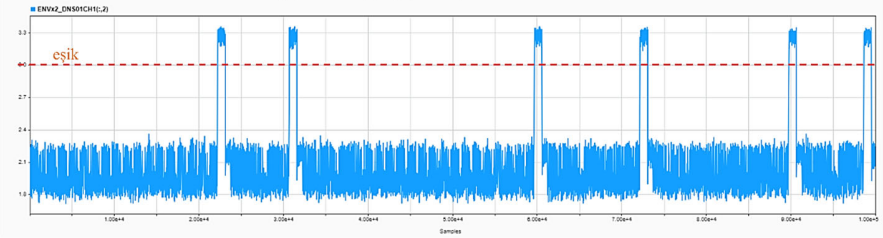
Sonuç olarak, her bir paket için benzersiz bir sinyal dosyası elde edilmiş ve toplamda 7421 sinyal dosyasından oluşan bir veri seti oluşturulmuştur. Araştırmacılar için verilerin daha anlaşılır ve kolay analiz edilebilir olmasını sağlamak amacıyla; part.pcap dosyalarındaki paketler, sinyallerin osiloskopla yakalanma sırasına (ETH dosyalarındaki paket sırasına) göre yeniden sıralandırılmıştır. Böylece, part.pcap dosyasındaki paket numaraları ile karşılık gelen sinyal dosyasının numaraları birebir eşleştirilmiştir. Ayrıca ETH dosyalarındaki her satırın sonuna ilgili sinyal parçasının adını içeren "Signal File Name" isimli bir sütun eklenmiştir. Veri setiyle birlikte güncellenmiş ETH ve part.pcap dosyaları da paylaşılmıştır. Hazırlanan veri seti tüm araştırmacıların erişimine açık olup ilgili dosyalar Mendeley Data platformunda [37] paylaşılmıştır. Tablo 2'de veri seti oluşturma sürecindeki işlem adımlarında toplanan ve işlem sonrası elde kalan paket sayıları sunulmuştur.

3.2. Önerilen Sinyal Tabanlı Sınıflandırma Modeli
(Proposed Signal-Based Classification Model)

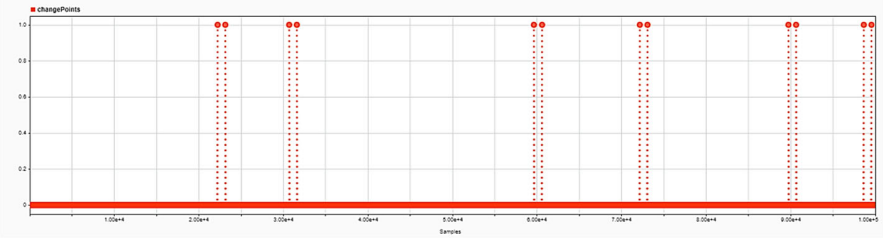
Sinyal tabanlı ağ trafiği sınıflandırması için önerilen metodolojik yaklaşım Şekil 7'de gösterilmektedir. Bu yaklaşımda, fiziksel katmandan elde edilen paket sinyalleri farklı tekniklerle görselleştirilmekte, ardından önceden eğitilmiş evrişimli sinir ağı (CNN) mimarilerine aktarılmaktadır. Çalışmada kullanılan görselleştirme teknikleri ve CNN mimarileri alt bölümlerde açıklanmıştır.



(a) DNS_01_CH1 (ilk 100.000 örnek)

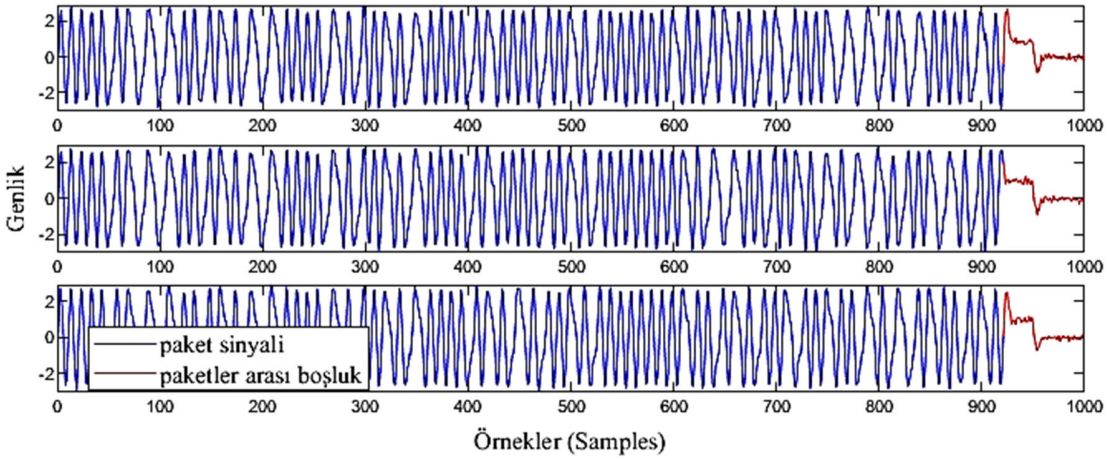


(b) 2 kez zarfı alınmış CH1



(c) Paket sinyallerinin başlangıç ve bitiş noktaları

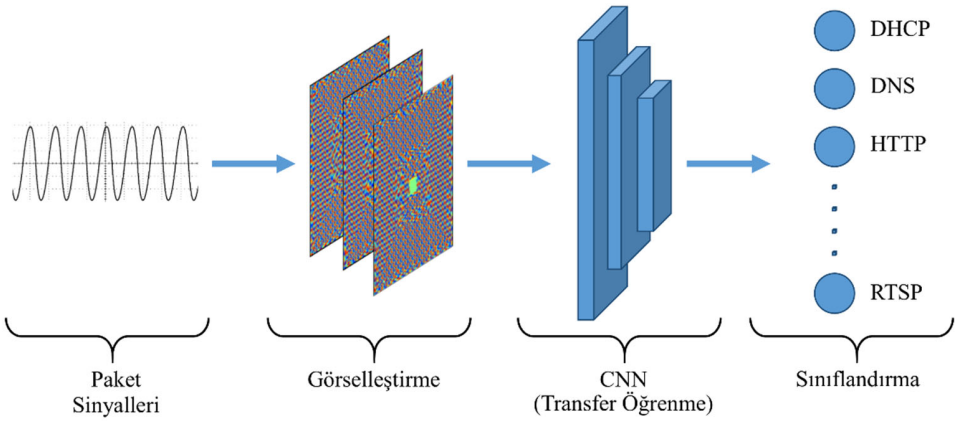
Şekil 5. Sinyal Kırpma İşlemi (Signal segmentation process)



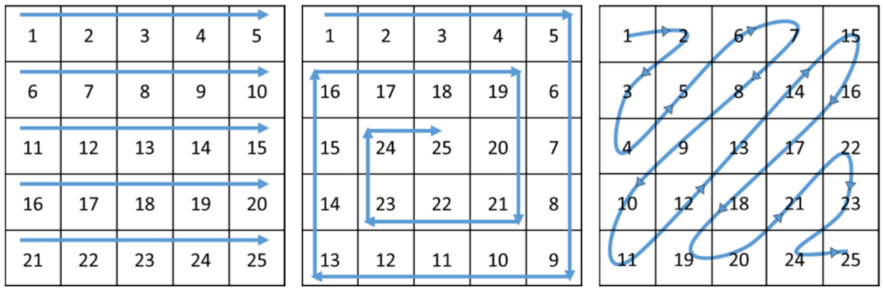
Şekil 6. Üç sinyale ait son 1000 örnek (Final 1000 samples of three signal instances)

Tablo 2. Veri seti oluşturma adımlarına göre protokol bazlı paket sayıları
(Protocol-based packet counts by dataset construction steps)

Protokoller	Paket Sayıları		
	Toplanan (1. Adım)	Tekilleştirme Sonrası (2. Adım)	Sinyali Elde Edilen (4. Adım)
DHCP	2276	1195	1191
DNS	1193	1187	1185
HTTP	1626	1385	1377
ICMP	1308	1067	1064
RTSP	1593	1587	1569
TLS	1038	1037	1035
Toplam	*9034*	*7458*	*7421*



Şekil 7. Önerilen metodolojik yaklaşım (Proposed methodological approach)



Şekil 8. Zaman serisi temelli görselleştirme matris dizilimleri (Time-series based visualization matrix arrays)

3.2.1. Kullanılan görselleştirme teknikleri
(Used visualization techniques)

Bu çalışmada, sinyallerin görsel temsillerini oluşturmak amacıyla dört farklı görselleştirme tekniği kullanılmış ve her bir teknik için ayrı bir görüntü veri seti oluşturulmuştur. Kullanılan teknikler; zaman serisi tabanlı yatay, spiral ve diyagonal zikzak görselleştirmelerin yanı sıra zaman-frekans tabanlı spektrogram yönteminden oluşmaktadır. Zaman serisi tabanlı görselleştirmelerde, sinyalin zaman ekseni boyunca değişen genlik değerleri doğrudan kullanılarak iki boyutlu görüntüler elde edilmiştir. Zaman-frekans temelli spektrogram tekniği ile ise, sinyalin zaman içerisinde değişen frekans bileşenlerini yansıtan ve hem zaman hem de frekans bilgisini eşzamanlı olarak sunan görsel temsiller üretilmiştir.

Zaman serisi tabanlı görselleştirme tekniklerinde izlenen genel yaklaşım şu şekildedir; Fiziksel katmandan elde edilen ham paket sinyalleri, öncelikle [-1,1] aralığında normalize edilerek vektör haline getirilmiştir. Ardından bu vektörler, seçilen görselleştirme tekniğine göre (yatay, spiral, diyagonal zikzak) Şekil 8'de gösterildiği gibi iki boyutlu kare matrislere aktarılmıştır. Bu matrisler, görselleştirme amacıyla önce 0-1 aralığında normalize edilip gri tonlamalı görüntülere çevrilmiştir. Gri tonlamalı görüntüler 8-bit (0–255) formatına çevrildikten sonra uygun renk haritaları (colormap) kullanılarak üç kanallı (RGB) renkli görüntülere dönüştürülmüştür. Son olarak elde edilen bu renkli görüntüler CNN mimarilerine doğrudan giriş olarak kullanılmıştır.

Yatay görselleştirme, sinyalin zamansal yapısını doğrudan koruyarak dalga formlarının yatay eksende ilerlemesini sağlar. Bu temsil, özellikle uzun süreli trendlerin, genlik değişimlerinin ve sinyalin genel formunun incelenmesi gereken durumlarda avantajlıdır. Dolayısıyla, zaman alanında sinyal bütünlüğünün korunmasının önemli olduğu uygulamalarda tercih edilmektedir [38]. Yatay görselleştirmede kullanılan matris dizilimi Eş. 2'deki gibi ifade edilebilir. N: matrisin satır/sütun boyutu, $X = [X_1, X_2, .., X_{N^2}]$ giriş sinyali olmak üzere;

$$M_{yatay_{i,j}} = X_{(i-1)\cdot N+j}, \quad 1 \le i,j \le N \tag{2}$$

Spiral görselleştirmede, zaman serisinin dairesel düzlemde yinelenen desenler halinde sunulmasıyla periyodik ve döngüsel yapılar belirgin hale getirilir. Bu yöntem özellikle periyodik bileşenlerin, ritmik tekrarların ve frekans tabanlı düzenliliklerin analizinde avantaj sağlar. Ağ trafiğinde tekrarlayan paket dizilerinin veya döngüsel protokol davranışlarının incelenmesinde kullanışlıdır [38]. Bu görselleştirme görselleştirmede kullanılan matris dizilimi Eş. 3-Eş. 5'deki gibi ifade edilebilir.

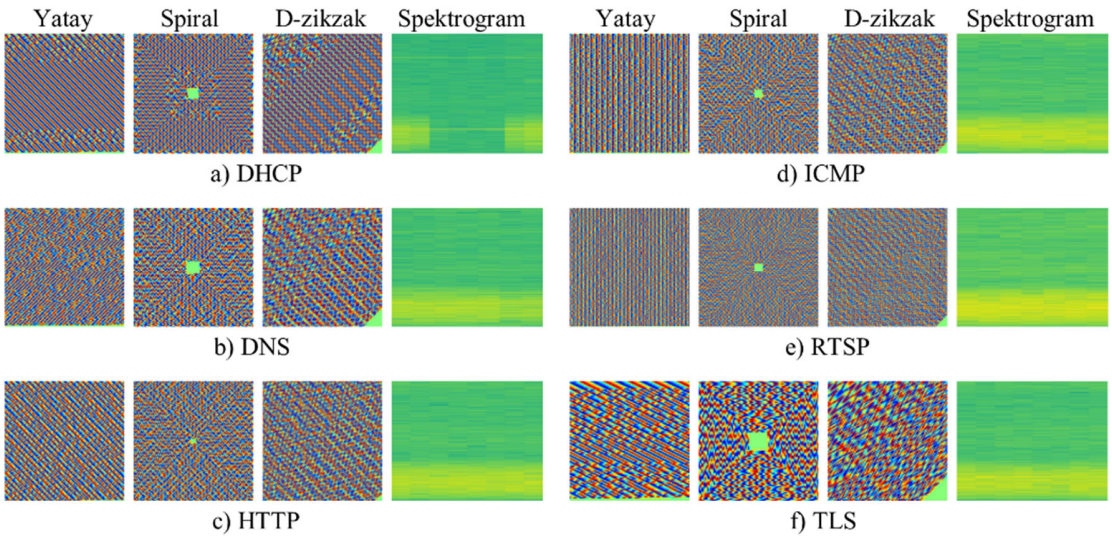
X: giriş sinyali, N: matrisin satır/sütun boyutu, b: katmanın başladığı örnek indeksi, L: katman (halka) numarası, ℓ: katman kenar uzunluğu, $\pi(i,j)$: spiral dizilimdeki indis eşleme fonksiyonu olmak üzere;

$$\begin{aligned} &X = [X_1, X_2, .., X_{N^2}]; \quad b = 1 + 4L(N - L); \\ &L = \min\{i-1, j-1, N-i, N-j\}; \quad \ell = N - 2L \end{aligned} \tag{3}$$

$$\pi(i,j) = \begin{cases} b + (j - L - 1), & i = L + 1 \\ b + (\ell - 1) + (i - L - 1), & j = L + \ell \\ b + 2(\ell - 1) + (L + \ell - j), & i = L + \ell \\ b + 3(\ell - 1) + (L + \ell - i), & j = L + 1 \\ b, & \ell = 1 \end{cases} \tag{4}$$

$$M_{spiral_{i,j}} = X_{\pi(i,j)}, \quad 1 \le i,j \le N \tag{5}$$

Diyagonal zikzak görselleştirmede, sinyalin hem yatay hem dikey eksenlerde zikzak düzeninde işlenmesiyle lokal değişimlerin ve frekans bileşenlerinin birlikte vurgulanması mümkün kılınır. Bu



Şekil 9. Örnek sinyal görüntüleri (Example signal images)

yaklaşım hem düşük hem de yüksek frekans bileşenlerinin eşzamanlı gözlenmesi gereken durumlarda avantaj sağlar. Özellikle ani dalgalanmaların, kısa süreli anomalilerin veya çok ölçekli sinyal özelliklerinin ortaya çıkarılmasında etkilidir [38]. Bu görselleştirme görselleştirmede kullanılan matris dizilimi matematiksel olarak Eş. 6-Eş. 9'daki gibi ifade edilebilir.

$X = [X_1, X_2, .., X_{N^2}]$ giriş sinyali, N: matrisin boyutu, L_d: her diyagonal üzerindeki eleman sayısı, b: her diyagonalin başlangıç örnek indeksi, $d = i + j = 2,3,\ldots,2N$, p: diyagonal yön tayin parametresi, $i_{min} = \max(1, d-N)$, $i_{max} = \min(N, d-1)$ olmak üzere;

$$L_d = \begin{cases} d-1, & d \le N+1 \\ 2N-d+1, & d > N+1 \end{cases} \quad (6)$$

$$b(d) = \begin{cases} 1 + \frac{(d-1)(d-2)}{2}, & d \le N+1 \text{ için} \\ 1 + \frac{N(N+1)}{2} + (d-N-1)N - \frac{(d-N-1)(d-N)}{2}, & \text{diğer} \end{cases} \quad (7)$$

$$p(i,j) = \begin{cases} i - i_{min}, & d = tek \\ i_{max} - i, & d = çift \end{cases} \quad (8)$$

$$M_{dzigzag_{i,j}} = X_{b(d)+p(i,j)}, \quad 1 \le i,j \le N \quad (9)$$

Spektrogram, sinyalin zaman boyunca frekans dağılımını renk yoğunluğu ile ifade ederek hem zaman hem de frekans bilgilerini aynı anda sunar. Yatay eksende zaman, dikey eksende frekans bileşenleri gösterilir. Bu yöntem, CNN'lerin farklı ölçeklerdeki öznitelikleri yakalamasına imkân tanır. Özellikle sinyalin frekans bileşenlerinin zamana göre nasıl değiştiğini görmek, kısa süreli olayları veya spektral yoğunluk farklılıklarını analiz etmek gerektiğinde avantaj sağlar. Spektrogram tekniğinde kısa süreli Fourier dönüşümü (Short-Time Fourier Transform – STFT) temel alınır. Matematiksel olarak STFT Eş. 10'daki gibi tanımlanır.

$$STFT_x(t,\omega) = \int_{-\infty}^{\infty} x(\tau) \cdot w(\tau - t) \cdot e^{-j\omega t}\, d\tau \quad (10)$$

Burada:
$x(\tau)$: Giriş sinyali,
$w(\tau - t)$: t anında merkezlenmiş pencere fonksiyonu (Hamming),
ω: Açısal frekans ifadesidir.

Elde edilen STFT çıktısının genlik karesi alınarak Eş.11'deki gibi spektrogram elde edilir.

$$Spektrogram_x(t,\omega) = |STFT_x(t,\omega)|^2 \quad (11)$$

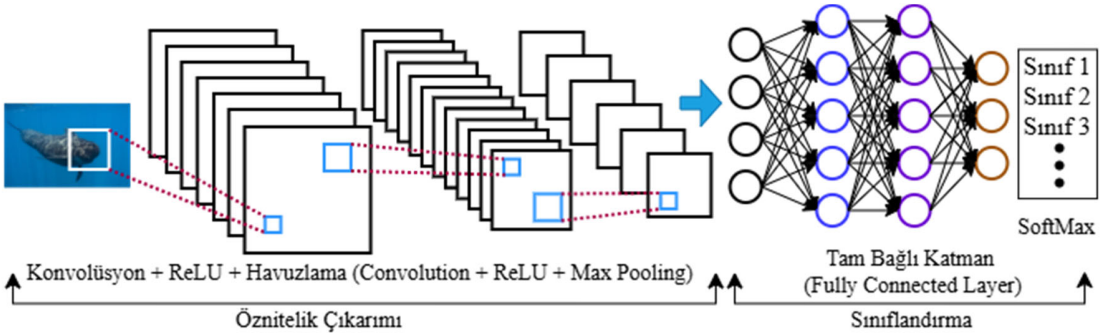
Spektrogram, bu şekilde sinyalin frekans içeriğinin zamanla nasıl değiştiğini enerji yoğunluğu (renk şiddeti) üzerinden görsel olarak sunar [28]. Bu teknikle oluşturulan renkli görseller doğrudan evrişimli sinir ağı (CNN) mimarilerine giriş olarak verilmiş ve böylece sinyalin spektral yapısına dayalı bir öğrenme süreci de gerçekleştirilmiştir.

Şekil 9'da, fiziksel katmanda yakalanan paketlere ait ham sinyallerin dört farklı görselleştirme tekniğiyle işlenmesi sonucunda elde edilen örnek görüntüler sunulmuştur. Bu görüntüler, her bir protokole ait ilk üç sinyal kullanılarak oluşturulmuştur.

3.2.2. Kullanılan CNN mimarileri (CNN architectures used)

Evrişimli sinir ağları (CNN), çok katmanlı yapay sinir ağlarının özel bir türü olup, özellikle görüntü işleme, nesne tanıma ve sınıflandırma gibi görevlerde üstün başarı göstermektedir. Şekil 10'da CNN'lerin temel mimarisi verilmiştir. İleri beslemeli bir yapıya sahip olan CNN'ler, genel olarak aşağıdaki temel katmanlardan oluşur:

- Evrişim (Convolutional) Katmanı: Giriş verisine konvolüsyon işlemi uygulanarak, yerel örüntüler çıkarılır ve çıktı olarak özellik haritaları üretilir. Bu işlemde filtreler yardımıyla giriş verisindeki kenar, doku, köşe gibi anlamlı özellikler belirlenir.
- Aktivasyon Fonksiyonu (ReLU): Evrişim işlemi sonrası elde edilen çıktılara doğrusal olmayanlık kazandırmak için uygulanır. ReLU (Rectified Linear Unit) fonksiyonu, negatif değerleri sıfıra dönüştürerek ağı daha verimli hale getirir. Çıktı olarak düzeltilmiş bir özellik haritası verir.
- Havuzlama (Pooling) Katmanı: Özellik haritalarının boyutunu azaltarak modelin hesaplama yükünü düşürür ve aşırı öğrenmeyi önlemeye yardımcı olur. Genellikle maksimum havuzlama (max pooling) işlemi tercih edilir. Havuzlanmış özellik haritası düzleştirilir ve nihai çıktıyı elde etmek için tamamen bağlı bir katmana bağlanır.
- Tam Bağlantılı (Fully Connected) Katman: CNN mimarisinin son kısmında yer alır ve sınıflandırma işlemi bu katman üzerinden gerçekleştirilir.



Şekil 10. Temel CNN mimarisi (Basic CNN architecture)

Bu katmanlar, veri içindeki örüntüleri kademeli olarak soyutlayarak sınıflandırma doğruluğunu artırmayı hedefler. CNN mimarileri, kullanılan katman sayısı, katman tipleri ve bağlantı yapıları gibi mimari özelliklere göre farklılık gösterir [39, 40].

Bu çalışmada, literatürde yaygın olarak kullanılan iki CNN mimarisi olan Darknet19 [41] ve Darknet53 [42], transfer öğrenme yöntemiyle yeniden eğitilmiştir. Darknet19; 19 evrişim ve 5 maksimum havuzlama katmanından oluşmaktadır. Daha derin bir yapı sunan Darknet53 ise toplam 53 evrişim katmanına sahiptir. Darknet53, Darknet19'dan farklı olarak maksimum havuzlama katmanlarına sahip olmayıp residual (artık) bloklar içermektedir. Tablo 3'te, Darknet mimarilerinin ayrıntılı yapısı özetlenmiştir [43].

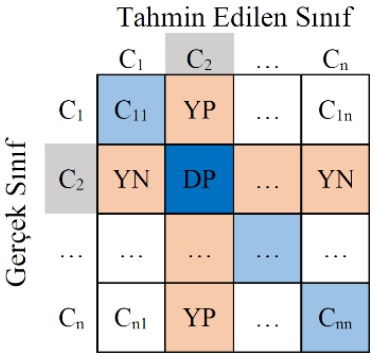
Residual blok, giriş bilgisini ara katmanlardan geçirmeden doğrudan çıkışa ekleyen skip connection yapısına sahiptir. Bu bloklar, giriş ile çıkış arasındaki farkı öğrenmeye odaklanarak çok katmanlı yapılarda bilgi kaybını azaltır ve gradyan yayılımını iyileştirir. Böylece, ağın daha derin katmanlarla daha verimli ve kararlı şekilde eğitilmesine olanak tanır.

Her iki mimaride evrişim katmanlarında 3×3 ve 1×1 boyutunda filtreler kullanılmakta, filtre sayısı ağ derinleştikçe artmaktadır. Son evrişim katmanı hariç, diğer tüm evrişim katmanlarının ardından Batch Normalization (toplu normalizasyon) ve LeakyReLU (sızdıran relu) aktivasyon fonksiyonu uygulanmaktadır. Batch Normalization, girdileri sıfır etrafında yeniden merkezleyip ölçeklendirerek eğitim sürecini hızlandırmakta ve kararlı hale getirmekte, ayrıca aşırı öğrenme eğilimini azaltmaya katkı sağlamaktadır. LeakyReLU ise negatif girdilerde küçük de olsa bir değer üreterek klasik ReLU'nun neden olduğu ölü nöron problemini büyük ölçüde ortadan kaldırmakta ve nöronların aktif kalmasını sağlamaktadır.

Orijinal mimaride, Darknet19 giriş görüntülerini 224×224 piksel boyutunda işlerken, Darknet53 256×256 veya 416×416 piksel boyutlarında giriş kabul etmektedir. Ancak, MATLAB ortamında her iki mimari için giriş görüntü boyutları 256×256 piksel olarak standartlaştırılmıştır. Transfer öğrenme sürecinde, her iki mimarinin son convolution2D katmanı ile sınıflandırma katmanları, hedef sınıf sayısına uygun şekilde yeniden yapılandırılmış ve başlangıç ağırlıkları olarak önceden eğitilmiş mimarilerin ağırlıkları kullanılmıştır. Önceki alt bölümde oluşturulan sinyal görüntüleri, 256×256 piksel boyutlarına yeniden ölçeklendirilerek, mimariler bu yeni görev için eğitilmiştir.

3.3. Performans Ölçümleri (Performance Metrics)

Çalışmada kullanılan CNN modellerinin performansını değerlendirmek için karışıklık matrisleri kullanılmıştır. Çoklu sınıflandırma problemlerinde karışıklık matrisi, Şekil 11'de gösterildiği gibi doğru pozitif (DP), yanlış pozitif (YP) ve yanlış negatif (YN) gibi temel değerlendirme bileşenlerini içermektedir.



Şekil 11. Çoklu sınıf karışıklık matrisi (Multi-class confusion matrix)

Doğru Pozitif (DP), bir örneğin hem gerçekte hem de model tarafından belirli bir sınıfa ait olarak doğru şekilde tahmin edildiği durumların sayısını ifade eder. Karışıklık matrisinde ilgili sınıfa ait köşegen hücreye karşılık gelir.

Yanlış Pozitif (YP), aslında başka bir sınıfa ait olduğu hâlde, model tarafından ilgili sınıfa ait olarak yanlış tahmin edilen örneklerin sayısını gösterir. Karışıklık matrisinde ilgili sınıfın sütunu boyunca, köşegen dışındaki hücrelerde yer alır.

Yanlış Negatif (YN), gerçekte ilgili sınıfa ait olduğu hâlde, modelin başka bir sınıfa ait olarak yanlış sınıflandırdığı örneklerin sayısını ifade eder. Karışıklık matrisinde ilgili sınıfın satırı boyunca, köşegen dışındaki hücrelerde yer alır. Son olarak, n toplam sınıf sayısını ifade eder.

Modellerin genel performansını değerlendirmek için, karışıklık matrislerinden elde edilen doğruluk (Accuracy) Eş. 12, kesinlik (Precision) Eş. 13, duyarlılık (Recall) Eş. 14 ve F1 skoru Eş. 15 gibi metrikler kullanılmıştır.

$$Doğruluk = \frac{\sum_{i=1}^{n} DP_i}{\sum_{i=1}^{n}(DP_i + YP_i + YN_i)} = \frac{\sum_{i=1}^{n} C_{ii}}{\sum_{i=1}^{n}\sum_{j=1}^{n} C_{ij}} \quad (12)$$

$$Kesinlik = \frac{1}{n}\sum_{i=1}^{n} \frac{DP_i}{DP_i + YP_i} \quad (13)$$

$$Duyarlılık = \frac{1}{n}\sum_{i=1}^{n} \frac{DP_i}{DP_i + YN_i} \quad (14)$$

$$F1\ Skoru = \frac{1}{n}\sum_{i=1}^{n} 2 \cdot \frac{Kesinlik_i \cdot Duyarlılık_i}{Kesinlik_i + Duyarlılık_i} \quad (15)$$

Tablo 3. Darknet19 ve Darknet53 yapıları (Structure of Darknet19 and Darknet53)

Darknet19				Darknet53				
Katman	Filtreler		Çıktı	Tekrar	Katman	Filtreler		Çıktı
	Sayı	Boyut /adım				Sayı	Boyut /adım	
Evrişim, BN, LReLU	32	3x3	224x224		Evrişim, BN, LReLU	32	3x3	256x256
Max. Havuzlama		2x2/2	112x112		Evrişim, BN, LReLU	64	3x3/2	128x128
Evrişim, BN, LReLU	64	3x3	112x112		Evrişim, BN, LReLU	32	1x1	
Max. Havuzlama		2x2/2	56x56	1x	Evrişim, BN, LReLU	64	3x3	
Evrişim, BN, LReLU	128	3x3	56x56		Residual (artık)			128x128
Evrişim, BN, LReLU	64	1x1	56x56		Evrişim, BN, LReLU	128	3x3/2	64x64
Evrişim, BN, LReLU	128	3x3	56x56		Evrişim, BN, LReLU	64	1x1	
Max. Havuzlama		2x2/2	28x28	2x	Evrişim, BN, LReLU	128	3x3	
Evrişim, BN, LReLU	256	3x3	28x28		Residual			64x64
Evrişim, BN, LReLU	128	1x1	28x28		Evrişim, BN, LReLU	256	3x3/2	32x32
Evrişim, BN, LReLU	256	3x3	28x28		Evrişim, BN, LReLU	128	1x1	
Max. Havuzlama		2x2/2	14x14	8x	Evrişim, BN, LReLU	256	3x3	
Evrişim, BN, LReLU	512	3x3	14x14		Residual			32x32
Evrişim, BN, LReLU	256	1x1	14x14		Evrişim, BN, LReLU	512	3x3/2	16x16
Evrişim, BN, LReLU	512	3x3	14x14		Evrişim, BN, LReLU	256	1x1	
Evrişim, BN, LReLU	256	1x1	14x14	8x	Evrişim, BN, LReLU	512	3x3	
Evrişim, BN, LReLU	512	3x3	14x14		Residual			16x16
Max. Havuzlama		2x2/2	7x7		Evrişim, BN, LReLU	1024	3x3/2	8x8
Evrişim, BN, LReLU	1024	3x3	7x7		Evrişim, BN, LReLU	512	1x1	
Evrişim, BN, LReLU	512	1x1	7x7	4x	Evrişim, BN, LReLU	1024	3x3	
Evrişim, BN, LReLU	1024	3x3	7x7		Residual			8x8
Evrişim, BN, LReLU	512	1x1	7x7		Ortalama Havuzlama		Global	
Evrişim, BN, LReLU	1024	3x3	7x7		Evrişim (Connected)		1000	
Evrişim	1000	1x1	7x7		Softmax			
Ortalama Havuzlama		Global	1000			(BN: Batch Normalization)		
Softmax						(LReLU: LeakyReLU)		

Ayrıca kullanılan modellerin sınıflandırma performansları arasındaki farkın istatistiksel anlamlılığını değerlendirmek için McNemar testi uygulanmıştır. Bu test, iki sınıflandırma modelinin aynı örnekler üzerinde farklı tahminler üretip üretmediğini inceleyerek anlamlılık düzeyini belirler. Bu testin istatistiği Eş. 16'daki gibi tanımlanır.

b: Model A'nın doğru, Model B'nin yanlış sınıflandırdığı örnek sayısı, c: Model A'nın yanlış, Model B'nin doğru sınıflandırdığı örnek sayısı, X^2: McNemar test istatistiği, $F_{X^2(1)}$: serbestlik derecesi 1 olan ki-kare dağılımının kümülatif dağılım fonksiyonu, p: elde olasılık değeri olmak üzere $p < 0{,}05$ olduğunda fark istatistiksel olarak anlamlı kabul edilir.

$$\chi^2 = \frac{(|b-c|-1)^2}{(b+c)}, \qquad p = 1 - F_{X^2(1)}(X^2) \tag{16}$$

4. Deneysel Sonuçlar (Experimental Results)

Deneyler, AMD Ryzen Threadripper PRO 3975WX işlemci, 128 GB RAM ve dört adet Nvidia RTX 3090 GPU ile donatılmış bir iş istasyonunda gerçekleştirilmiştir. CNN modelleri, yatay, spiral, diyagonal zikzak ve spektrogram tabanlı görselleştirme teknikleriyle oluşturulan veri setleri kullanılarak MATLAB R2024a ortamında 5-kat çapraz doğrulama yöntemiyle eğitilmiştir. Bu yöntemde veriler eşit boyutta rastgele seçilmiş beş parçaya bölünmüş, her yinelemede verinin %80'i eğitim, %20'si test amacıyla kullanılmış; ayrılan test kümesi aynı zamanda doğrulama amacıyla değerlendirilmiştir. Eğitim süreci bu şekilde beş kez tekrarlanmış ve her yinelemede farklı bir parça test için ayrılarak tüm verilerin hem eğitim hem test aşamasında en az bir kez kullanılması sağlanmıştır. Böylece modelin genellenebilirlik performansı daha güvenilir biçimde değerlendirilmiştir. Ayrıca, eğitim sırasında ağ ağırlıklarının büyüklüklerini sınırlandırarak aşırı öğrenme riskini azaltmak amacıyla weight decay (L2 regularization) tekniği uygulanmıştır. Eğitim sürecinde kullanılan tüm hiperparametreler Tablo 4'te sunulmuştur.

Tablo 4. CNN modelleri için kullanılan eğitim hiperparametreleri (Training hyperparameters)

Eğitim Parametreleri	Değer
Solver	sgdm
MaxEpoch	50
MiniBatchSize	128
InitialLearnRate	0.001
VerboseFrequency	20
ExecutionEnvironment	gpu
L2regularization	1×10^{-4}
LearnRateDropFactor	0.1
LearnRateDropPeriod	20
LearnRateSchedule	Piecewise

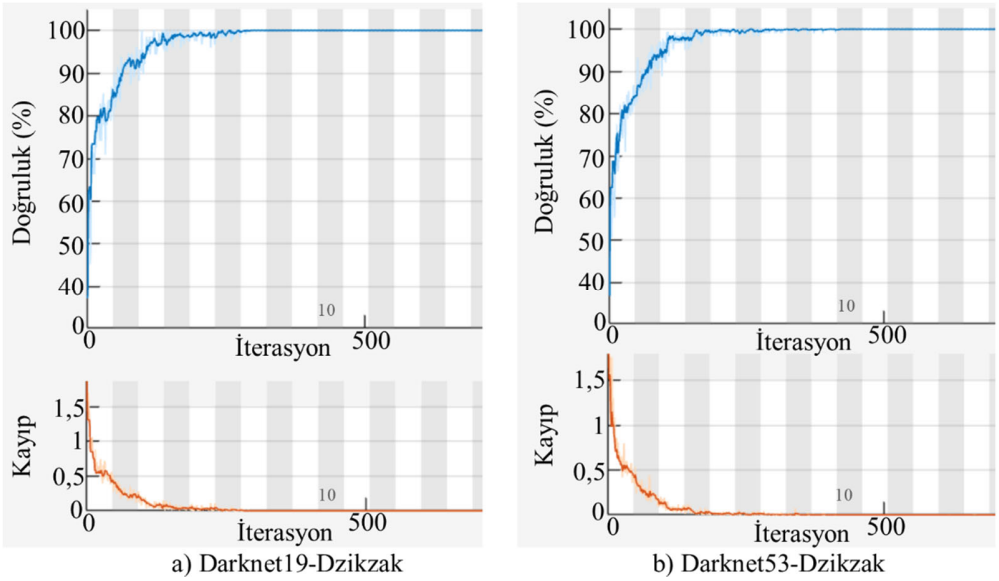
Şekil 12'de, D-zikzak görselleştirmesiyle eğitilen Darknet19 ve Darknet53 modellerinin eğitim doğruluk (accuracy) ve kayıp (loss) eğrileri sunulmuştur. Bu iki kombinasyon, yapılan deneylerde en yüksek sınıflandırma başarımını elde ettikleri için temsil edici örnekler olarak seçilmiştir. Eğriler, her iki modelin de 10 epoch sonrasında kararlı bir öğrenme süreci izlediğini ve kayıp değerlerinin istikrarlı biçimde azaldığını göstermektedir. Ayrıca, doğruluk eğrilerinin belirgin bir aşırı öğrenme (overfitting) belirtisi göstermeden doygunluğa ulaştığı gözlenmiştir.

Modellerin genel eğitim performansını gösteren karışıklık matrisleri Şekil 13'te sunulmuş, bu matrislerden elde edilen performans ölçümleri ve uygulanan McNemar test sonuçları Tablo 5'te özetlenmiştir. Tablo 5'te farklı görselleştirme teknikleri ve modellerin sınıflandırma performansları karşılaştırılmıştır. Darknet19 ve Darknet53 modelleri, tüm görselleştirme yöntemlerinde yüksek doğruluk, kesinlik, duyarlılık ve F1 skorları sergilemiştir. En yüksek performans, %96,24 doğruluk ve %96,18 F1 skoru ile Darknet53 modelinin Diyagonal Zikzak görselleştirmesi üzerinde elde edilmiştir. Diyagonal Zikzak yöntemi, sinyaldeki lokal değişiklikleri ve farklı frekans bileşenlerini eşzamanlı olarak temsil edebilmesi sayesinde CNN'lerin daha zengin öznitelikler öğrenmesine imkân tanımış ve bu nedenle en yüksek başarıyı sağlamıştır. Yatay ve Spiral görselleştirme tekniklerinde her iki modelde de benzer ve başarılı sonuçlar vermiştir. Yatay görselleştirme uzun süreli trendlerin korunması, spiral görselleştirme döngüsel desenlerin belirginleştirilmesi gibi farklı odaklara sahip olmakla birlikte, sınıflandırma için ayırt edici öznitelikleri yeterli düzeyde sağlamış ve Diyagonal Zikzak'a oldukça yakın performans sergilemiştir. Buna karşılık, Spektrogram görselleştirmesi diğer tekniklere kıyasla belirgin olarak daha düşük performans göstermiştir. Spektrogram, zaman-frekans bilgisini aynı anda sunması açısından güçlü bir yöntem olsa da elde edilen yoğunluk haritalarının karmaşık yapısı CNN'lerin öznitelik çıkarımını zorlaştırmış ve sınıflar arasındaki ayırt edici farkların zayıflamasına yol açmıştır. McNemar testi sonuçları, Darknet53 modelinin genel doğruluk bakımından Darknet19'a göre daha yüksek doğru sınıflandırma sayısına ulaşmasına rağmen, bu farkların istatistiksel olarak anlamlı olmadığını göstermektedir ($p > 0,05$). Bu sonuç, iki modelin genel performans düzeylerinin istatistiksel olarak eşdeğer kabul edilebileceğini ortaya koymaktadır. Eğitim süreleri açısından ise, Darknet53 modellerinin Darknet19'a göre iki katından daha uzun eğitim süreleri gerektirdiği görülmektedir. Bu bulgular, Diyagonal Zikzak görselleştirme yöntemi ile Darknet53 modelinin sinyal düzeyindeki ağ trafiği sınıflandırmasında en etkili kombinasyon olduğunu göstermektedir. Ayrıca, daha derin yapıya sahip olan Darknet53 modelinin, genellikle Darknet19'a göre biraz daha üstün performans sağladığı gözlemlenmiştir.

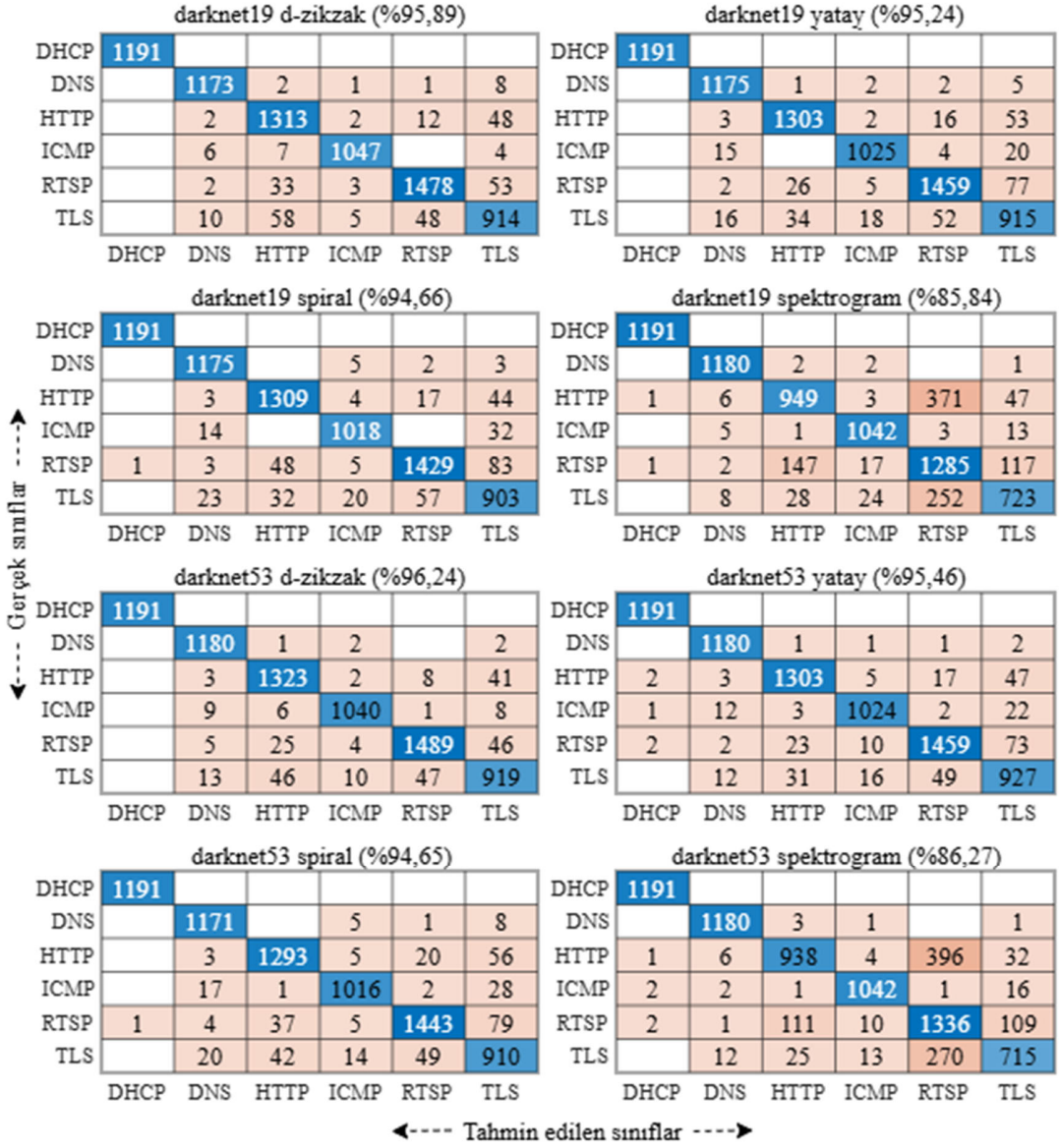
Şekil 14'te, D-zikzak görselleştirmesiyle eğitilen Darknet19 ve Darknet53 modellerinin beş kat çapraz doğrulama sonucunda elde edilen ortalama ROC eğrileri sunulmuştur. Eğriler, her iki modelin de yüksek ayırt ediciliğe sahip olduğunu ve tüm sınıflar için ortalama AUC değerlerinin 0,992 olarak hesaplandığını göstermektedir. Bu sonuçlar, modellerin farklı ağ protokollerini sinyal düzeyinde yüksek doğrulukla ayırt edebildiğini ve önerilen yöntemin kararlılığını desteklemektedir.

Şekil 15'te, D-zikzak görselleştirmesiyle eğitilen Darknet19 ve Darknet53 modellerine ait Grad-CAM aktivasyon haritaları örnek olarak sunulmuştur. Görseller, modellerin karar verme sürecinde sinyal görüntülerinin belirli bölgelerine odaklandığını göstermektedir. Aktivasyon haritalarında kırmızıya yakın alanlar modelin en yüksek dikkat yoğunluğuna sahip bölgeleri, mavi alanlar ise düşük önem derecesine sahip bölgeleri temsil etmektedir. Verilen örneklerde, modellerin genellikle sinyal görüntülerinin başlangıç ve orta kısımlarında daha yoğun aktivasyon gösterdiği görülmektedir. Bu durum, modellerin ağ trafiği sinyallerinde enerji dağılımı, genlik değişimleri ve yapısal geçişler gibi ayırt edici örüntülere karşı duyarlı olduğunu ve bu bölgelerden protokollere özgü özellikleri öğrenebildiğini göstermektedir.

Geleneksel ağ trafiği sınıflandırma çalışmalarında paketlerin decode edilmesiyle başlık bilgisi, IP adresleri, MAC adresleri ve payload içerikleri gibi hassas verilere erişim söz konusu olmaktadır. Bu durum kullanıcı gizliliği açısından önemli riskler doğurabilmektedir. Buna karşılık, bu çalışmada önerilen yöntem ağ trafiğini doğrudan fiziksel katmandan alınan ham elektriksel sinyaller üzerinden işlemektedir. Dolayısıyla paket içeriğine veya üst katman başlık bilgilerine dair herhangi bir veri elde edilmemekte, sınıflandırma yalnızca sinyal formuna dayalı olarak gerçekleştirilmektedir. Önerilen yaklaşımın, paket içeriğine erişim ihtiyacını tamamen ortadan kaldırarak yüksek doğrulukta sınıflandırma başarısı sağlayabilmesi, veri gizliliği ve



Şekil 12. D-zikzak görselleştirmesi için Darknet19 ve Darknet53 modellerine ait eğitim doğruluk ve kayıp eğrileri
(Training accuracy and loss curves of Darknet19 and Darknet53 for D-zigzag visualization)

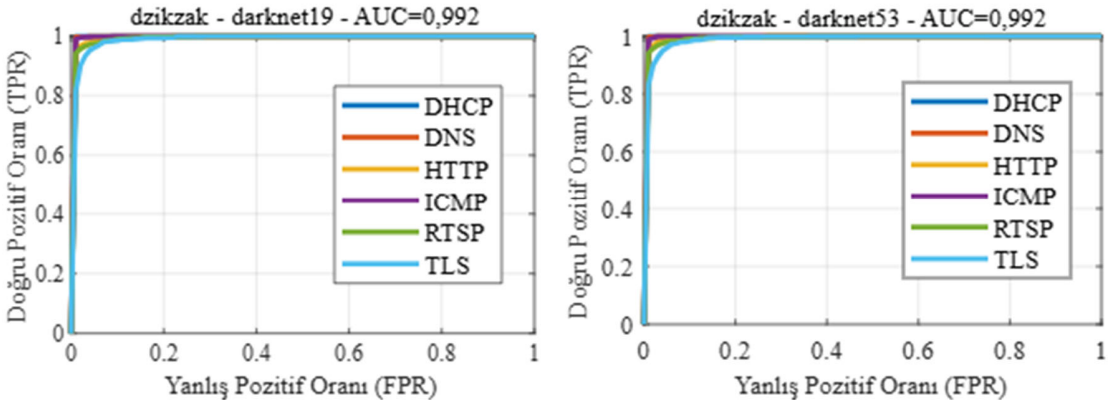


Şekil 13. Dartnet19 ve Darknet53 için karışıklık matrisleri (Confusion matrices for Dartnet19 and Darknet53)

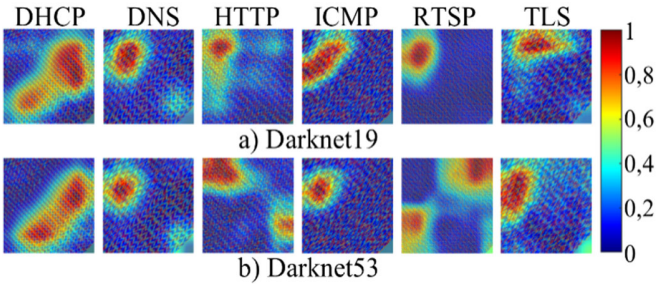
Tablo 5. Görselleştirme teknikleri ve modellerin sınıflandırma performans sonuçları
(Classification performance results of visualization techniques and models)

Görselleştirme	Model	Doğruluk (%)	Kesinlik (%)	Duyarlılık (%)	F1 Skoru (%)	Kat (Fold) Başına Ortalama Eğitim Süresi (saniye)	*p* (McNemar)
Yatay	Darknet19	95,24	95,11	95,25	95,17	1761	0,247
	Darknet53	95,46	95,32	95,50	95,40	4174	(b=76, c=92)
Spiral	Darknet19	94,66	94,52	94,70	94,60	2045	1,00
	Darknet53	94,65	94,52	94,68	94,59	4374	(b=131, c=130)
D-zikzak	Darknet19	95,89	95,87	95,88	95,87	1762	0,118
	Darknet53	*96,24*	*96,18*	*96,18*	*96,18*	*4138*	(b=115, c=141)
Spektrogram	Darknet19	85,84	87,60	86,36	86,67	2024	0,188
	Darknet53	86,27	88,47	86,64	87,09	4399	(b=261, c=293)

mahremiyeti açısından güçlü bir avantaj sunmakta; ağ trafiği sınıflandırma sürecini, içerik (paket, akış) odaklı yöntemlere kıyasla daha güvenli ve etik hale getirmektedir. Bu bulgular, derin CNN mimarilerinin ham Ethernet sinyallerinin görsel temsillerinden ayırt edici özellikleri öğrenerek, paket içeriğine erişmeden de yüksek doğrulukla trafik sınıflandırması yapabildiğini ortaya koymaktadır. Tablo 6'da önerilen çalışma ile literatürdeki ağ trafiği sınıflandırma alanında yapılmış bazı çalışmalar karşılaştırmalı olarak sunulmuştur.



Şekil 14. D-zikzak görselleştirmesi için Darknet19 ve Darknet53 modellerine ait ortalama ROC eğrileri
(Average ROC curves of Darknet19 and Darknet53 for D-zigzag visualization)



Şekil 15. D-zikzak görselleştirmesi için Darknet19 ve Darknet53 modellerine ait sınıf-bazlı Grad-CAM aktivasyon haritaları
(Class-wise Grad-CAM activation maps of Darknet19 and Darknet53 for D-zigzag visualization).

Tablo 6. Önerilen çalışma ile literatürdeki bazı çalışmaların karşılaştırılması
(Comparison of the proposed study with some works in the literature)

Çalışma	Analiz Seviyesi	Kullanılan Teknik	Veri Seti	Başarı Oranı (%)
Fang vd. [13]	Paket seviyesi	CNN (ResNet18/34/50)	CIC-IDS2017 Shodan	%98
Chiu vd. [5]	Paket seviyesi	CAPC (1DCNN+AEC+DNN)	ISCXVPN2016 Özel veri seti	%97,42 (ISCX) %99,98
Lotfollahi vd. [14]	Paket seviyesi	CNN (Deep Packet)	ISCXVPN2016	%98 (uygulama) %94 (trafik türü)
Zhou vd. [15]	Akış seviyesi	CNN (MMN-CNN)	Moore	%99,30
Yamansavascilar vd. [16]	Akış seviyesi	ML (RF, KNN, Bayes Net)	ISCXVPN2016 Özel veri seti	%93,94 (RF, ISCX) %92,99 (RF)
Salman vd. [17],	Akış seviyesi	CNN (ConvNet)	ISCXVPN2016, ISCXTor2016	%95,84
Ahmed vd. [8]	Akış seviyesi	ML	Özel veri seti	%88,86 (KNN) %96,33 (RF) %99,86 (ANN)
Lopez-Martin vd. [18]	Akış seviyesi	CNN+RNN	Özel veri seti	%99
Al-Jameel vd. [19]	Akış seviyesi	DL	Özel veri seti	%98,4
Aouini vd. [20]	Akış seviyesi	ML (C5.0)	Özel veri seti	%98,8
Fauvel vd. [21]	Akış seviyesi	CNN (LEXNet)	AppClassNet	%89,7
Wang vd. [22]	Akış seviyesi	CNN+Transformer (SwinT-CNN)	ISCXVPN2016	%96,7
Zheng vd. [23]	Paket ve akış seviyesi	MLP-Mixer	ISCXVPN2016, ISCXTor2016 USTC-TFC2016	%99
Yang vd. [24]	Paket ve akış seviyesi	GRU + SAE (DM-HNN)	ISCXVPN2016 ISCXTor2016	%91,02-%96,42
Önerilen Çalışma	*Sinyal seviyesi*	*CNN (Darknet)*	*Özgün veri seti (SigNet-6)*	*%96,24*

Tablo 6 incelendiğinde, literatürdeki çalışmaların genellikle ağ trafiğini paket veya akış seviyesinde sınıflandırmaya odaklandığı görülmektedir. Bu çalışmalarda, sinir ağı tabanlı derin öğrenme yöntemleri ile makine öğrenmesi temelli yaklaşımlar yaygın olarak kullanılmış ve elde edilen doğruluk oranlarının genel olarak %88,86 ile %99,98 arasında değiştiği rapor edilmiştir. Ancak bu yöntemler, verinin içerik veya başlık düzeyindeki bilgilerini kullanmakta ve dolayısıyla ağ trafiğinin üst katmanlarına bağımlı kalmaktadır. Buna karşın, bu çalışmada önerilen yöntem, sinyal seviyesinde gerçekleştirilen ilk sistematik sınıflandırma yaklaşımı olma özelliğini taşımaktadır. Önerilen model, ağ paketlerine ait fiziksel katman sinyallerini doğrudan kullanarak, herhangi bir paket içeriği veya başlık bilgisi gerektirmeden sınıflandırma yapmaktadır. Bu özgün yaklaşım, oluşturulan SigNet-6 veri setiyle birlikte, sinyal düzeyinde trafiğin ayırt edici özelliklerinin derin öğrenme modelleri tarafından etkili biçimde öğrenilebileceğini göstermiştir. Elde edilen %96,24 doğruluk oranı, paket veya akış tabanlı çalışmalarda raporlanan başarılarla karşılaştırılabilir düzeyde olup, yöntemin hem yüksek performans hem de veri gizliliği açısından önemli avantajlar sunduğunu ortaya koymaktadır.

5. Sonuçlar (Conclusions)

Bu çalışmada, fiziksel katman sinyallerine dayalı yeni bir ağ trafiği sınıflandırma yaklaşımı sunulmuştur. Geleneksel yöntemlerin çoğunlukla paket veya akış düzeyinde gerçekleştirdiği sınıflandırma işlemleri burada sinyal düzeyinde ele alınmış; osiloskop yardımıyla yakalanan altı farklı protokole ait paket sinyallerinden 7421 benzersiz örnek elde edilmiştir. Bu sinyaller yatay, spiral, diyagonal zikzak ve spektrogram olmak üzere dört farklı görselleştirme tekniğiyle görsel veri setlerine dönüştürülmüş ve transfer öğrenme yöntemiyle Darknet19 ile Darknet53 mimarileri eğitilmiştir. Deneyler sonucunda, %96'nın üzerinde doğrulukla Darknet53–Diyagonal Zikzak kombinasyonu en başarılı performansı göstermiştir. Bu bulgular, anlamlı trafik özelliklerinin sinyal düzeyinde korunduğunu ve bu bilgilerin derin öğrenme modelleri tarafından etkili bir biçimde öğrenilebildiğini ortaya koymaktadır. Ayrıca, önerilen yöntemin paket içeriğine veya başlık bilgilerine erişme gereksinimi olmadan yalnızca fiziksel sinyaller üzerinden yüksek doğruluk sağlaması, veri gizliliği ve mahremiyeti açısından güçlü bir avantaj sunmaktadır. Çalışma kapsamında, literatürde benzeri bulunmayan sinyal tabanlı bir ağ trafiği veri seti de oluşturulmuş ve açık erişime sunulmuştur.

Bununla birlikte, çalışmanın bazı sınırlılıkları bulunmaktadır. Öncelikle, deneysel veri kümesi yalnızca altı protokol türü ile sınırlıdır ve 10BASE-T (10 Mbps) Ethernet standardı üzerinden gerçekleştirilmiştir. Bu durum, yöntemin daha geniş protokol çeşitliliğine ve farklı ağ standartlarına doğrudan genellenebilirliğini kısıtlayabilmektedir. Ayrıca, derin CNN mimarilerinin eğitim sürelerinin uzun olması, gerçek zamanlı uygulamalarda bazı pratik kısıtlar doğurabilmektedir. Bu sınırlılıklara rağmen, elde edilen sonuçlar yöntemin sinyal düzeyinde trafik sınıflandırması için güçlü bir potansiyele sahip olduğunu göstermektedir.

Gelecek çalışmalarda, yöntemin farklı ağ standartları üzerinde ve daha büyük ölçekli, çok protokollü veri kümeleri ile test edilmesi planlanmaktadır. Ayrıca, daha verimli mimarilerin geliştirilmesiyle eğitim sürelerinin kısaltılması hedeflenmektedir. Bunun yanında, yöntemin yalnızca protokol sınıflandırmasıyla sınırlı kalmayıp, saldırı tespiti, güvenlik tehditlerinin analizi ve anomali tespiti gibi ileri kullanım senaryolarına da uyarlanabilme potansiyeli bulunmaktadır. Bu tür olası çalışmaların, sinyal düzeyinde trafik analizi alanının hem akademik hem de pratik katkılarını genişletebileceği değerlendirilmektedir.

Kaynaklar (References)

1. Wang Y., Yun X., Zhang Y., Zhao C., Liu X, A Multi-Scale Feature Attention Approach to Network Traffic Classification and Its Model Explanation, IEEE Trans Netw Serv Manag, 19 (2), 875–889, 2022.
2. Rezaei S., Liu X., Deep Learning for Encrypted Traffic Classification: An Overview, IEEE Commun Mag, 57 (5), 76–81, 2019.
3. Abbasi M., Shahraki A., Taherkordi A., Deep Learning for Network Traffic Monitoring and Analysis (NTMA): A Survey, Comput Commun, 170(February), 19–41, 2021.
4. Zheng Y., Dang Z., Lian X., Peng C., Gao X., Multi-view multi-label network traffic classification based on MLP-Mixer neural network, Comput Networks, 253(August), 110746, 2024.
5. Chiu K.C., Liu C.C., Chou L. Der., CAPC: Packet-Based Network Service Classifier with Convolutional Autoencoder, IEEE Access, 8, 218081–218094, 2020.
6. Zhao J., Jing X., Yan Z., Pedrycz W., Network traffic classification for data fusion: A survey, Inf Fusion, 72, 22–47, 2021.
7. Azab A., Khasawneh M., Alrabaee S., Choo K.K.R., Sarsour M., Network traffic classification: Techniques, datasets, and challenges, Digit Commun Networks, 10 (3), 676–692, 2024.
8. Ahmed A.A., Agunsoye G., A real-time network traffic classifier for online applications using machine learning, Algorithms, 14 (8), 250, 2021.
9. Pacheco F., Exposito E., Gineste M., Baudoin C., Aguilar J., Towards the Deployment of Machine Learning Solutions in Network Traffic Classification: A Systematic Survey, IEEE Commun Surv Tutorials, 21 (2), 1988–2014, 2019.
10. Burukanlı M., Çıbuk M., Intrusion Detection and Performance Analysis Using Copula Functions, Bitlis Eren Üniversitesi Fen Bilim Derg, 13 (4), 1335–1354, 2024.
11. Zhou D., Yan Z., Fu Y., Yao Z., A survey on network data collection, J Netw Comput Appl, 116(May), 9–23, 2018.
12. Hu Y., Zeng Z., Song J., Xu L., Zhou X., Online network traffic classification based on external attention and convolution by IP packet header, Comput Networks, 252(June), 110656, 2024.
13. Fang Z., Gao X., Zhang H., Tang J., Gao Q., Application Layer Protocol Identification Method Based on ResNet, Algorithms, 18 (1), 52, 2025.
14. Lotfollahi M., Jafari S.M., Shirali H.Z.R., Saberian M., Deep packet: a novel approach for encrypted traffic classification using deep learning, Soft Comput, 24 (3), 1999–2012, 2020.
15. Zhou H., Wang Y., Lei X., Liu Y., A Method of Improved CNN Traffic Classification, In: 2017 13th Int. Conf. Comput. Intell. Secur. IEEE, 177–181, 2017.
16. Yamansavascilar B., Guvensan M.A., Yavuz A.G., Karsligil M.E., Application identification via network traffic classification, In: 2017 Int. Conf. Comput. Netw. Commun. IEEE, 843–848, 2017.
17. Salman O., Elhajj I.H., Chehab A., Kayssi A., A Multi-level Internet Traffic Classifier Using Deep Learning, Proc 2018 9th Int Conf Netw Futur NOF 2018, 68–75, 2018.
18. Lopez-Martin M., Carro B., Sanchez-Esguevillas A., Lloret J., Network Traffic Classifier With Convolutional and Recurrent Neural Networks for Internet of Things, IEEE Access, 5, 18042–18050, 2017.
19. Al-Jameel M., Turner S., Kanakis T., Al-Sherbaz A., Bhaya W.S., Deep Learning Approach for Real-time Video Streaming Traffic Classification, In: 2022 Int. Conf. Comput. Sci. Softw. Eng. IEEE, 168–174, 2022.
20. Aouini Z., Kortebi A., Ghamri-Doudane Y., Cherif I.L., Early classification of residential networks traffic using C5.0 machine learning algorithm, In: 2018 Wirel. Days. IEEE, 46–53, 2018.
21. Fauvel K., Chen F., Rossi D., A Lightweight, Efficient and Explainable-by-Design Convolutional Neural Network for Internet Traffic Classification, In: Proc. 29th ACM SIGKDD Conf. Knowl. Discov. Data Min. ACM, New York, NY, USA, 4013–4023, 2023.
22. Wang Y., Gao Y., Li X., Yuan J., Encrypted Traffic Classification Model Based on SwinT-CNN. 2023 4th Int Conf Comput Eng Appl ICCEA 2023, 138–142, 2023.
23. Zheng Y., Dang Z., Lian X., Peng C., Gao X., Multi-view multi-label network traffic classification based on MLP-Mixer neural network. Comput Networks, 253(August), 110746, 2024.
24. Yang Y., Yan Y., Gao Z., Rui L., Lyu R., Gao B., Yu P., A Network Traffic Classification Method Based on Dual-Mode Feature Extraction

and Hybrid Neural Networks. IEEE Trans Netw Serv Manag, 20 (4), 4073–4084, 2023.
25. Kırışoğlu S., Kotan B., Kotan K., Çok Katmanlı Algılayıcı ile Ağ Trafiği Sınıflandırma Analizi, Düzce Üniversitesi Bilim ve Teknol Derg, 10 (2), 837–846, 2022.
26. Shams M., Ağ anomalisi tespitinde makine öğrenmesi algoritmalarının kullanımı ve karşılaştırmalı analizi, Yüksek Lisans Tezi, Sakarya Üniversitesi, Fen Bilimleri Enstitüsü, Sakarya, 2020.
27. Özekes S., Karakoç E.N., Makine Öğrenmesi Yöntemleriyle Anormal Ağ Trafiğinin Tespit Edilmesi, Düzce Üniversitesi Bilim ve Teknol Derg, 7 (1), 566–576, 2019.
28. Khan A.S., Ahmad Z., Abdullah J., Ahmad F., A Spectrogram Image-Based Network Anomaly Detection System Using Deep Convolutional Neural Network, IEEE Access, 9, 87079–87093, 2021.
29. Salati M., Askerzade İ., Bostancı G.E., Convolutional neural network models using metaheuristic based feature selection method for intrusion detection, Journal of the Faculty of Engineering and Architecture of Gazi University, 40 (1), 179–188, 2024.
30. Wang W., Sheng Y., Wang J., Zeng X., Ye X., Huang Y., Zhu M., HAST-IDS: Learning Hierarchical Spatial-Temporal Features Using Deep Neural Networks to Improve Intrusion Detection, IEEE Access, 6, 1792–1806, 2018.
31. Imrana Y., Xiang Y., Ali L., Abdul-Rauf Z., A bidirectional LSTM deep learning approach for intrusion detection, Expert Syst Appl, 185 (June), 115524, 2021.
32. Yu L., Dong J., Chen L., Li M., Xu B., Li Z., Qiao L., Liu L., Zhao B., Zhang C., PBCNN: Packet Bytes-based Convolutional Neural Network for Network Intrusion Detection, Comput Networks, 194 (January), 108117, 2021.
33. Hu X., Gu C., Chen Y., Wei F., tCLD-Net: A Transfer Learning Internet Encrypted Traffic Classification Scheme Based on Convolution Neural Network and Long Short-Term Memory Network, In: 2021 Int. Conf. Commun. Comput. Cybersecurity, Informatics, IEEE, 1–5, 2021.
34. Uğurlu M., Doğru I.A., Arslan R.S., Detection and classification of darknet traffic using machine learning methods, Journal of the Faculty of Engineering and Architecture of Gazi University, 38 (3), 1737–1746, 2023.
35. Wang W., Zhu M., Wang J., Zeng X., Yang Z., End-to-end encrypted traffic classification with one-dimensional convolution neural networks, In: 2017 IEEE Int. Conf. Intell. Secur. Informatics, IEEE, 43–48, 2017.
36. Bit-Twist: a flexible packet generator and editör, https://bittwist.sourceforge.io/doc.html, Erişim tarihi Ekim 11, 2024
37. Geylani M., Çıbuk M., Akbal A., Ethernet Frame Physical-Layer Signal Dataset - 10BaseT, Mendeley Data, https://data.mendeley.com/datasets/x8x39r6nmt/1, 2025.
38. Demir Ş.N., Çıbuk M., The impact of signal visualization types on the performance of image processing-based convolutional neural networks, In: Akdeniz 14th Int. Conf. Appl. Sci., Academy Global Publishing House, Kyrenia, 98–126, 2025.
39. Şeker A., Diri B., Balık H., A review about deep learning methods and applications, Gazi Journal of Engineering Sciences, 3 (3), 47-64, 2017.
40. Amri A.A., Ismail A.R., Zarir A.A., Comparative performance of deep learning and machine learning algorithms on imbalanced handwritten data, Int J Adv Comput Sci Appl, 9 (2), 258–264, 2018.
41. Redmon J., Farhadi A., YOLO9000: Better, Faster, Stronger, In: 2017 IEEE Conf. Comput. Vis. Pattern Recognit, IEEE, 6517–6525, 2017.
42. Redmon J., Farhadi A., YOLOv3: An Incremental Improvement, arXiv Prepr arXiv180402767, 1–6, 2018.
43. Wu W., Guo L., Gao H., You Z., Liu Y., Chen Z., YOLO-SLAM: A semantic SLAM system towards dynamic environment with geometric constraint, Neural Comput Appl, 34 (8), 6011–6026, 2022.