T.C.
YILDIZ TEKNİK ÜNİVERSİTESİ
FEN BİLİMLERİ ENSTİTÜSÜ

OTONOM ARAÇLARDA KİŞİ TESPİTİ İÇİN FEDERE ÖĞRENMEYE DAİR YENİ BİR YÖNTEM

Emin AKPINAR

YÜKSEK LİSANS TEZİ
Elektronik ve Haberleşme Mühendisliği Anabilim Dalı
Haberleşme Programı

Danışman
Doç.Dr. Bülent BOLAT

Eş-Danışman
Dr.Öğr.Üyesi Murat TAŞKIRAN

Mayıs, 2024

T.C.
YILDIZ TEKNİK ÜNİVERSİTESİ
FEN BİLİMLERİ ENSTİTÜSÜ

OTONOM ARAÇLARDA KİŞİ TESPİTİ İÇİN FEDERE ÖĞRENMEYE DAİR YENİ BİR YÖNTEM

Emin AKPINAR tarafından hazırlanan tez çalışması 14.05.2024 tarihinde aşağıdaki jüri tarafından Yıldız Teknik Üniversitesi Fen Bilimleri Enstitüsü Elektronik ve Haberleşme Mühendisliği Anabilim Dalı Haberleşme Programı **YÜKSEK LİSANS TEZİ** olarak kabul edilmiştir.

Doç.Dr. Bülent BOLAT
Yıldız Teknik Üniversitesi
Danışman

Dr.Öğr.Üyesi Murat TAŞKIRAN
Yıldız Teknik Üniversitesi
Eş-Danışman

Jüri Üyeleri

Doç.Dr. Bülent BOLAT, Danışman
Yıldız Teknik Üniversitesi _____

Doç.Dr. Umut Engin AYTEN, Üye
Yıldız Teknik Üniversitesi _____

Dr.Öğr.Üyesi N. Tuğrul ARTUĞ, Üye
İstanbul Üniversitesi Cerrahpaşa _____

Danışmanım Doç.Dr. Bülent BOLAT sorumluluğunda tarafımca hazırlanan Otonom Araçlarda Kişi Tespiti için Federe Öğrenmeye Dair Yeni Bir Yöntem başlıklı çalışmada veri toplama ve veri kullanımında gerekli yasal izinleri aldığımı, diğer kaynaklardan aldığım bilgileri ana metin ve referanslarda eksiksiz gösterdiğimi, araştırma verilerine ve sonuçlarına ilişkin çarpıtma ve/veya sahtecilik yapmadığımı, çalışmam süresince bilimsel araştırma ve etik ilkelerine uygun davrandığımı beyan ederim. Beyanımın aksinin ispatı halinde her türlü yasal sonucu kabul ederim.

Emin AKPINAR

İmza

Geçtiğimiz Yıl
Depremde Hayatını
Kaybetmiş Vatandaşlarımıza
İthaf Edilmiştir

TEŞEKKÜR

Tezimi yazarken yanımda olan herkese en içten teşekkürlerimi sunarım. Öncelikle, bu yolculuğumda bana rehberlik eden, bilgi ve deneyimlerini benimle paylaşan danışman hocam Doç.Dr. Bülent BOLAT'a ve eş danışman hocam Dr.Öğr.Üyesi Murat TAŞKIRAN'a teşekkür etmek isterim. Sabır, anlayış ve bilimsel bakış açıları olmadan bu tezi gerçekleştirmem mümkün olmazdı.

Aileme; annem Hayriye AKPINAR, babam Mustafa AKPINAR ve kardeşlerim Eren AKPINAR ve Ecrin AKPINAR'a sonsuz sevgi ve destekleri için teşekkür ederim. Sizlerin varlığı, bu zorlu süreci çok daha kolay ve anlamlı kıldı.

Arkadaşlarım Volkan SUNGUN, Can AYDEMİR, Uğur CAN, Mehmet GÜLBUDAK ve Berkay YILDIRIM olmasaydı, bu yolculuk çok daha zor olurdu. Hem moral hem de pratik destekleri için onlara minnettarım.

Ayrıca sonu gelmeyen sorularıma ve eleştirel yorumlarıma göğüs geren ve akademik kariyerimin bir hayli renkli geçmesini sağlayan kahvaltı grubu fertlerine bu süreçte hep yanımda oldukları için teşekkür ederim.

Bu tez, pek çok yönden bir öğrenme süreciydi ve bu süreçte yanımda olan herkese teşekkür ederim.

Emin AKPINAR

İÇİNDEKİLER

SİMGE LİSTESİ vii

KISALTMA LİSTESİ viii

ŞEKİL LİSTESİ xi

TABLO LİSTESİ xiii

ÖZET xiv

ABSTRACT xvi

1 GİRİŞ 1

1.1 Literatür Araştırması 2

1.2 Tezin Amacı 21

1.3 Hipotez 22

2 FEDERE ÖĞRENME 23

2.1 Dikey Federe Öğrenme 24

2.2 Yatay Federe Öğrenme 25

2.3 Federe Transfer Öğrenme 26

2.4 Silolar Arası Federe Öğrenme 27

2.5 Cihazlar Arası Federe Öğrenme 29

2.6 Heterojen Federe Öğrenme 30

3 YÖNTEM 32

3.1 Veri Kümeleri 33

3.2 Yaya Tespitinde Federe Öğrenme 35

4 BULGULAR 42

5 SONUÇ 49

KAYNAKÇA 51

TEZDEN ÜRETİLMİŞ YAYINLAR **56**

SİMGE LİSTESİ

a	Yerel Model 1 için Normalize Edilmemiş Katsayı
b	Yerel Model 2 için Normalize Edilmemiş Katsayı
c	Yerel Model 3 için Normalize Edilmemiş Katsayı
A	Yerel Model 1 için Normalize Edilmiş Katsayı
B	Yerel Model 2 için Normalize Edilmiş Katsayı
C	Yerel Model 3 için Normalize Edilmiş Katsayı
S	Normalize Edilmemiş Katsayılar Toplamı
$W_{k\text{üresel}}$	Küresel Model Ağırlıkları
W_{yerel1}	Yerel Model 1 Ağırlıkları
W_{yerel2}	Yerel Model 2 Ağırlıkları
W_{yerel3}	Yerel Model 3 Ağırlıkları
p_i	Gerçek Sınıf Etiketi
$\hat{p_i}$	Tahmin Edilen Sınıf Olasılığı
y_i	Pozitif-Negatif Örnek (Pozitif 1, Negatif 0)
t_i	Temel Gerçek Kutu Koordinatları
$\hat{t_i}$	Tahmin Edilen Kutu Koordinatları
$smooth_{L1}$	Smooth L1 Kaybı
$\text{Kayıp}^{\text{RPN}}_{\text{sınıflandırma}}$	RPN Sınıflandırma Kaybı
$\text{Kayıp}^{\text{RPN}}_{\text{regresyon}}$	RPN Regresyon Kaybı
$\text{Kayıp}^{\text{R-CNN}}_{\text{sınıflandırma}}$	R-CNN Sınıflandırma Kaybı
$\text{Kayıp}^{\text{R-CNN}}_{\text{regresyon}}$	R-CNN Regresyon Kaybı

KISALTMA LİSTESİ

3B	Üç Boyutlu
AdaBoost	Adaptif Artırma
ALF	Asimptotik Yerelleştirme Uydurma
ALFNet	Asimptotik Yerelleştirme Uydurma Ağı
BIFPN	Çift Yönlü Ağırlıklı Özellik Füzyon
C2F	Çaprazdan Birleşime
CBAM	Evrişimsel Blok Dikkat Modülünü
CenterNet	Merkezi Nokta Sinir Ağı
CFNet	Basamaklı Füzyon Ağı
CNN	Evrişimli Sinir Ağları
CPB	Evrişimli Tahmin Bloğu
C-RBM	Evrişimli Kısıtlı Boltzmann Makinesi
CSP	Merkez ve Ölçek Tahmini Dedektörü
CSS	Renk Kendine Benzerliği
EEG	Elektroensefalografi
ERM	Ampirik Risk Minimizasyon
ETH	Federal Teknik Yüksekokulu
Faster R-CNN	Daha Hızlı Bölge Tabanlı Evrişimli Sinir Ağı
Fast R-CNN	Hızlı Bölge Tabanlı Evrişimli Sinir Ağı
FCCL	Federe Çapraz Korelasyon ve Sürekli Öğrenme
FedAvg	Federe Ortalama
FedDyn	Dinamik Düzenleme ile Federe Öğrenme
FedFa	Adil ve Doğru Federe Öğrenme Algoritması

FedKT	Bilgi Aktarımı Yoluyla Federe Öğrenme Algoritması
FedLDC	Kayıp Bağımlı Katsayılarla Federe Öğrenme
FedMA	Eşleştirilmiş Ortalama ile Federe Öğrenme
FPN	Özellik Piramidi Ağı
FedProx	Federe Yakınsama
HHHFL	Hiyerarşik Heterojen Yatay Federe Öğrenme
HOG	Yönlendirilmiş Gradyan Histogramları
IIoT	Endüstriyel Nesnelerin İnterneti
INRIA	Ulusal Bilgisayar Bilimi ve Otomasyon Araştırma Enstitüsü
IoT	Nesnelerin İnterneti
IoU	Kesişim Bölü Birleşim
KITTI	Karlsruhe Teknoloji Enstitüsü ve Toyota Teknoloji Enstitüsü
LBP	Yerel İkili Desenler
LSFM	Yerelleştirilmiş Anlamsal Özellik Karıştırıcıları
LT-FL	Kayıp Toleranslı Federe Öğrenme
mAP	Ortalama Hassasiyetin Ortalaması
MIME	Mimik Federe Öğrenme Algoritması
MS COCO	Microsoft Bağlamdaki Ortak Nesneler
PSI	Özel Küme Kesişimi
RBM	Kısıtlı Boltzmann Makinesi
R-CNN	Bölge Tabanlı Evrişimli Sinir Ağı
ReLU	Düzeltilmiş Doğrusal Birim
ResNet	Artık Ağ
R-FCN	Bölge Tabanlı Tam Evrişimli Ağ
ROC	Alıcı İşletim Karakteristiği
RoI	İlgi Alanı
SCAFFOLD	Federe Öğrenme İçin Stokastik Kontrollü Ortalamalama
SDD	Tek Atış Dedektörü
SpilitNN	Bölünmüş Sinir Ağı

ST	Standart Eğitim Metodolojisi
SVM	Destek Vektör Makinesi
TRB	Üç Artık Ağ Bloğu
TRBBC-Net	Merkezi Nokta Sinir Ağı Tabanlı Üç Artık Sinir Ağı Bloğu Tespit Modeli
VAFL	Dikey Eş Zamanlı Olmayan Federe Öğrenme
YOLO	Yalnızca Bir Kez Bak
YOLOv8-CB	Katmanlar Arası Birleşim ve Dikkat Mekanizması İyileştirilmiş Yalnızca Bir Kez Bak versiyon 8
ZUT	Batı Pomeranya Teknoloji Üniversitesi

ŞEKİL LİSTESİ

Şekil 1.1 2000 yılı ölüm sayısı yüz değerine normalize edilmiş, yıllara göre yaya ile sürücü ve yolcu ölüm sayılarının 2000 yılına göre oransal değişim grafiği [1] 1

Şekil 1.2 Kehtarnavaz ve Rajkotwala'nın çalışmasına ait örnek iki çıktı [7] . 4

Şekil 1.3 Papageorgiou ve diğerlerinin çalışmasından alınan, farklı görüntülere ait dalgacık katsayılarının gri seviyeli temsilleri [8]. (a) Rastgele bir görüntünün dikey katsayıları. (b-d) 32 x 32 ölçekli insan görüntülerinin dikey, yatay ve köşe katsayıları. (e-g) 16 x 16 ölçekli insan görüntülerinin dikey, yatay ve köşe katsayıları 5

Şekil 1.4 Zhao ve Thorpe'nin çalışmasından alınan görüntü ve çalışmanın aşamaları [9]. (a) stereo kameranın sol görüntüsü, (b) derinlik farkı haritası, (c) segmentasyon sonucu, (d) önerilen yöntem ile tespit edilen nesneler 6

Şekil 1.5 Gavrila ve Giebel'in çalışmasından alınan bir yaya şekilleri hiyerarşisi örneği [10] 7

Şekil 1.6 Viola ve diğerlerinin çalışmasında paylaşılan örnek filtreler [11] . 8

Şekil 1.7 Tüzel ve diğerlerinin yaptığı çalışmaya ait örnek sonuçlar [13] . . 9

Şekil 1.8 Walk ve diğerlerinin çalışmasında paylaşılan işaretli bölgelere ait CSS çıktıları [16] 10

Şekil 1.9 Vazquez ve diğerlerinin kullandığı eğitim ve test yöntemi [17] . . 11

Şekil 1.10 Ouyang ve Wang'ın ortaklaşa öğrenilmesini savunduğu dört bileşen [20] 13

Şekil 1.11 DeepCascade yönteminin hız ve doğruluk ekseninde daha önceki yöntemlerle karşılaştırılması [21] 14

Şekil 1.12 R-FCN mimarisi [23] 15

Şekil 1.13 ALFNet mimarisi ve CPB [25] 17

Şekil 1.14 CSP mimarisi [28] 18

Şekil 1.15 CenterNet mimarisi [30] 19

Şekil 1.16 TRB yapısı [30] 20

Şekil 1.17 YOLOv8-CB mimarisi [32] 21

Şekil 2.1 PyVertical düzeneği [35] 25

Şekil 2.2 FedKT yapısı [40] 28

Şekil 3.1 Çalışmada gerçekleştirilen adımların blok diyagramı 32
Şekil 3.2 Kullanılan modelin genel yapısı 35
Şekil 3.3 Çalışmada kullanılan federe öğrenme yöntemleri için genel blok diyagram 38
Şekil 4.1 Temel gerçek ve tahmin kutusu örneği 43
Şekil 4.2 FedLDC ile eğitilmiş modele ait maksimum olmayan bastırma uygulanmış örnek tahmin çıktısı 45

TABLO LİSTESİ

Tablo 1.1 YOLO-R ağının yapısı [24] 16
Tablo 3.1 Çalışmada kullanılan veri kümeleri 34
Tablo 3.2 Her bir veri kümesinin eğitim, validasyon ve test görüntüsü sayıları 34
Tablo 3.3 Faster R-CNN ResNet50 FPN toplam parametre sayısı 36
Tablo 3.4 Her bir küresel iterasyonda yerel modeller için hesaplanan kayıp değerleri ve hesaplanan katsayılar 40
Tablo 4.1 Kesişim oranı eşiği 0.5 için test sonucu elde edilen kaçırma oranları 44
Tablo 4.2 Kesişim oranı eşiği 0.75 için test sonucu elde edilen kaçırma oranları 44
Tablo 4.3 Literatürdeki çalışmaların Caltech veri kümesi için kaçırma oranları 46
Tablo 4.4 Literatürdeki çalışmaların CityPersons veri kümesi için elde ettikleri kaçırma oranları 47
Tablo 4.5 Literatürdeki çalışmaların EuroCity Persons veri kümesi için elde ettikleri kaçırma oranları 48

ÖZET

Otonom Araçlarda Kişi Tespiti için Federe Öğrenmeye Dair Yeni Bir Yöntem

Emin AKPINAR

Elektronik ve Haberleşme Mühendisliği Anabilim Dalı
Yüksek Lisans Tezi

Danışman: Doç.Dr. Bülent BOLAT
Eş-Danışman: Dr.Öğr.Üyesi Murat TAŞKIRAN

Otonom sürüş sistemlerinin en önemli avantajlarından biri olası kazaları önleme ve dolayısıyla ölüm ve yaralanma sayılarını azaltma potansiyeline sahip olmasıdır. Yapay zekanın gelişmesiyle birlikte bu alanda yapılan çalışmalarda büyük bir artış yaşanmaktadır. Bu tür uygulamalarda başarılı bir genelleme performansı elde etmek, modelin farklı arka plan ve mümkün olduğunca çok benzersiz veriyle eğitilmesiyle mümkündür. Ancak bu kadar çeşitli ve büyük veri kümelerinin tek bir topluluk tarafından oluşturulması büyük bir emek ve maliyet gerektirmektedir. Ayrıca bu tür veri kümeleri oluşturulsa bile bu veri kümelerinin doğrudan paylaşılması kişisel veri gizliliği sorunlarını da beraberinde getirecektir. Bu çalışmada, bahsedilen sorunların çözümü için federe öğrenmenin potansiyeli araştırılmakta ve performansı standart eğitim metodolojisi ile kıyaslanmaktadır. Federe öğrenme, veri paylaşımına gerek duyulmaması sayesinde veri gizliliği sorunlarını ortadan kaldırır ve farklı veri kümelerinin bir sunucuyla paylaşılmasına gerek kalmadan birçok veri kümesinin eğitim için birlikte kullanılmasına olanak tanır. Literatürde federe öğrenme için pek çok yaklaşım önerilmiştir. Ancak farklı zorluk ve kapsamlılığa sahip veri kümelerinin kullanıldığı uygulamalarda, oluşturulan küresel model ağırlıkları, uygulamada kullanılan veri kümelerinin iyi bir temsilini yakalamak için yeterli değildir. Bu nedenle yerel model olarak Özellik Piramidi Ağı Kullanan Daha Hızlı Bölgesel Evrişimli Sinir Ağı ResNet50 (Faster R-CNN ResNet50 FPN) kullanılan ve yüksek kayıp değerlerinin elde edildiği ama aynı zamanda daha çok ve benzersiz veriye sahip veri kümelerinin ait olduğu yerel

modellere, eğitim sırasındaki kayıp değerinden yararlanarak daha yüksek katsayılar atayan yeni bir federe öğrenme yaklaşımı sunan Kayıp Bağımlı Katsayılarla Federe Öğrenmeyi (FedLDC) öneriyoruz. Çalışmada FedLDC hem federe öğrenmenin ilk örneği olan Federe Ortalama (FedAvg) hem de federe olmayan standart eğitim metodolojisi ile karşılaştırılmıştır. Üç farklı veri kümesi üzerinde yapılan testler sonucunda FedLDC, Caltech Yayada (Caltech Pedestrian) 0,13 ve ŞehirKişilerinde (CityPersons) 0,07 kaçırma oranı ile en başarılı sonuca ulaşırken, AvrupaŞehir Kişileri Gün (EuroCity Persons Day) veri kümesinde 0,10 kaçırma oranı ile 0,08'lik bir kaçırma oranına ulaşan standart eğitim metodolojisi ile rekabetçi sayılabilecek bir performans elde etti.

Anahtar Kelimeler: Federe Öğrenme, Yaya Tespiti, Derin Sinir Ağları, Kayba Bağlı Katsayılar.

ABSTRACT

Federated Learning in Person Detection and a New Method of Federated Learning

Emin AKPINAR

Department of Electronics and Communications Engineering
Master of Science Thesis

Supervisor: Assoc. Prof. Bülent BOLAT
Co-supervisor: Asst. Prof. Murat TAŞKIRAN

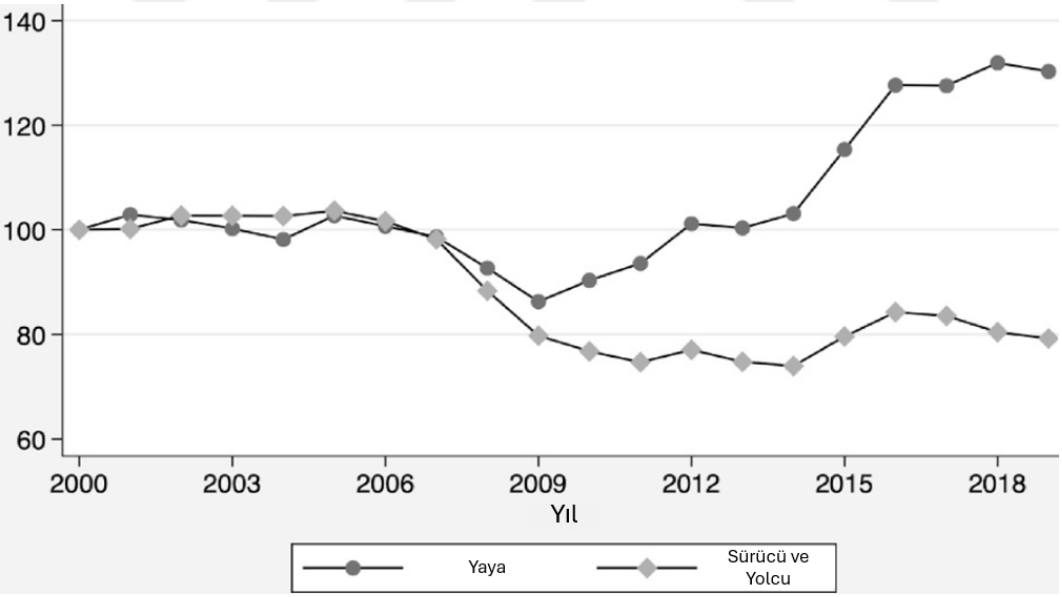
One of the most significant advantages of autonomous driving systems is their potential to prevent accidents and, consequently, reduce the number of deaths and injuries. With the advancement of artificial intelligence, there has been a significant increase in research in this field. Achieving successful generalization performance in such applications is possible through training the model with as diverse background and as much unique data as possible. However, creating such varied and large datasets by a single community requires considerable effort and cost. Moreover, even if such datasets are created, directly sharing these datasets will bring along personal data privacy issues. In our study, we explore the potential of federated learning as a solution to these problems and compare its performance with the standard training methodology. Federated learning eliminates data privacy issues by obviating the need for data sharing, allowing for the training of multiple datasets together without the need to share them with a server. Many approaches to federated learning have been proposed in the literature. Yet, in applications where data sets of different difficulty and comprehensiveness are used, the global model weights created are not sufficient to capture a good representation of the datasets used in practice. Therefore, we propose Federated Learning with Loss Dependent Coefficients (FedLDC), a new federated learning approach that uses the Faster R-CNN ResNet50 FPN as the local model, and assigns higher coefficients to local models belonging to datasets with high loss values but also with more and unique data, by taking advantage of the loss value during training. In our study,

we compared FedLDC with both FedAvg, the first example of federated learning, and the standard non-federated training methodology. Our tests on three different datasets showed that FedLDC achieved the most successful result with a miss rate of 0.13 in Caltech Pedestrian and 0.07 in CityPersons, while obtaining a competitive performance with a miss rate of 0.10 in the EuroCity Persons Day dataset, compared to the 0.08 miss rate of the standard training methodology.

Keywords: Federated Learning, Pedestrian Detection, Deep Neural Networks, Loss-Dependent Coefficients.

1
GİRİŞ

ABD'de 2021 yılında yayınlanan bir araştırmada, 2000-2019 yılları arasında ABD'de motorlu araç kazalarında 100.000'i yaya olmak üzere 741.000 kişinin öldüğü ve çalışmadan alınan Şekil 1.1'e göre sürücüler ve yolcular arasındaki ölümlerin sayısının zamanla azalırken, yayalar arasındaki sayıların zamanla arttığı raporlanmıştır [1]. Karayolu Güvenliği Küresel Durum Raporuna göre ise, motorlu araç kazaları yılda 1,2 milyondan fazla insanın hayatına mal olmakta ve 20 ila 50 milyon arasında kişi bu kazalarda yaralanmaktadır [2]. Belirtilen sayılar göz önünde bulundurulduğunda, özellikle yayalar için sürüş ve trafik güvenliğinin günümüzde halen daha önemli bir problem olduğu görülmektedir.



Şekil 1.1 2000 yılı ölüm sayısı yüz değerine normalize edilmiş, yıllara göre yaya ile sürücü ve yolcu ölüm sayılarının 2000 yılına göre oransal değişim grafiği [1]

Son yıllarda teknolojinin hızla ilerlemesi, günlük yaşamın birçok alanında köklü değişikliklere yol açmıştır. Özellikle yapay zeka ve görüntü işleme alanlarında kaydedilen önemli gelişmeler, otonom sistemlerin ve akıllı şehir uygulamalarının önünü açmıştır. Bu bağlamda, otonom araçlar ve yaya algılama sistemleri, hem

akademik araştırmalarda hem de endüstriyel uygulamalarda büyük ilgi görmektedir.

Otonom araçlar, trafikteki kazaları azaltma, enerji verimliliğini artırma ve seyahat konforunu iyileştirme potansiyeline sahiptir. Bu araçların en önemli bileşenlerinden biri, çevresel farkındalığı sağlayan yaya algılama sistemleridir. Yaya algılama, bir aracın yolda karşılaştığı yayaları doğru bir şekilde tanımlaması ve bu bilgiyi güvenli sürüş kararları almak için kullanması anlamına gelir.

Günümüzde yaya algılama sistemleri, geleneksel görüntü işleme tekniklerinden derin öğrenme tabanlı yöntemlere doğru evrilmiştir. İlk nesil sistemler, kenar ve köşe bilgilerine dayanan temel algoritmalar ve şablon eşleştirme yöntemleri kullanırken, modern sistemler büyük veri kümeleri üzerinde eğitilen derin sinir ağları ile çalışmaktadır. Ancak, bu sistemlerin başarısını artırmak için daha geniş ve çeşitli veri kümelerine ihtiyaç duyulmaktadır. Bu noktada, federe öğrenme yöntemleri, veri gizliliğini koruyarak farklı veri kaynaklarını bir araya getirmenin etkili bir yolu olarak öne çıkmaktadır.

1.1 Literatür Araştırması

Görüntü işleme ve yapay zekadaki gelişmelerle birlikte otomotiv endüstrisinde, otonom sürüş sistemleri ve yaya ve canlı tespiti alanında birçok çalışma yapılmaktadır. Bu teknolojilerin gelişmesi ve yaygınlaşmasıyla birlikte bahsedilen trafik kazalarının önemli bir kısmının önlenebileceği ve bu sayede ölüm ve yaralanmaların ciddi şekilde azalacağı düşünülmektedir.

İlk yaya tespit sistemleri, kenar, köşe gibi bilgilerin kullanıldığı elle yapılan özellik çıkarımlarını, temel algoritmaları ve şablon eşleştirmeye dayalı yöntemleri içermekteydi ancak bu ilk sistemler karmaşık senaryolarla ve yaya görünümündeki değişikliklerle baş edememeleri nedeniyle kullanım açısından sınırlı ve düşük başarıya sahiptiler.

1990'da Ali ve Dagless'ın yaptığı Araç ve Yaya Algılama ve İzleme (Vehicle and Pedestrian Detection and Tracking) adlı çalışmada gelişmiş sahne analizi algoritmalarından yararlanılarak, araçların ve yayaların hareketleri tanımlanıp takip edilmiştir ve bu veriler anlamlı trafik akışı ölçümlerine dönüştürülmüştür [3]. Araştırmada, özel olarak oluşturulmuş bir transputer-görüntü işleme sisteminden yararlanılmış ve araçların yapısal şekillerine göre sınıflandırılması için orta eksen dönüşümünün potansiyeli araştırılmıştır.

1992'de Tang ve diğerleri, farklı görüntü ve şekil yeniden yapılandırma

tekniklerini kullanarak yayaların sayısını ve yürüme yönlerini ölçen bir robot görüş sistemi önerdiler [4]. Bu sistem sekiz adımdan oluşmaktaydı, sırasıyla video görüntülerinin dönüştürülmesi, dondurulmuş karelerin sayısallaştırılması, 256 seviyeli gri görüntülerin ikili görüntülere dönüştürülmesi, fark görüntüleri kullanılarak yayanın dış hatlarının çıkarılması, çizgi gürültüsünün giderilmesi, yayanın şeklinin yeniden oluşturulması, yaya sayısının ölçülmesi ve yaya hareket yönünün belirlenmesi. Bu sistemde her adımdaki işlemler yalnızca yerel bilgilere bağlıydı. Bu sayede adımlar bağımsız ve paralel olarak gerçekleştirilebiliyorlardı.

1993'de Yamanashi Üniversitesi'nden Mori ve Charkari tarafından gerçekleştirilen çalışmada, engel tespiti için iki farklı model sunulmuştur [5]. Bu modeller ritim modeli ve gölge tespit modelidir. Ritim modeli, aydınlatma ve kıyafet gibi dış faktörlerden bağımsız olarak, hareket sıklığı ve mekansal konumlandırma gibi nitelikleri kullanarak ve yürüyüşlerinin ritmini analiz ederek yayaları tanımlar. Gölge tespit modeli ise, algılama için çeşitli yüzeylerden gelen ışık yansımasındaki farktan yararlanarak, oluşan karakteristik gölge desenlerini tanıyarak araçları tanımlamaya odaklanmıştır. Çalışmada bu yaklaşımlar, gerçek zamanlı bir görüş sistemi kullanan deneysel kurulumlarla doğrulanmış, değişen koşullarda etkinlikleri gösterilmiş ve robotik navigasyon ve güvenlik sistemlerinin geliştirilmesindeki potansiyel uygulamaları vurgulanmıştır.

1994'de yine Yamanashi Üniversitesi'nden Yasutomi ve Mori'nin gerçekleştirdiği çalışmada, yürümenin doğasında olan ritmik hareketten yararlanarak yayaların monoküler görüntü dizilerinden otomatik olarak algılanmasına yönelik bir yöntem sunulmuştur [6]. Çalışma, maksimum entropi yöntemini kullanarak yürüme hareketine karşılık gelen görüntü yoğunluğundaki periyodik değişiklikleri analiz ederek yayaları diğer hareketli nesnelerden ayırmaktadır. Çalışmada, hareketli bir nesnenin konumunu ve hızını tahmin etmek için, bir kinematik model, bir ölçüm modeli ve bir izleme filtresi kullanılarak yürüme ritmi zamansal ve uzaysal frekanslara ayrıştırılmaktadır. Bu yöntemin avantajları arasında giyim, mesafe ve hava koşullarındaki değişikliklere karşı dayanıklılığının yanı sıra basitliği ve gerçek zamanlı performansa olanak sağlaması yer almaktadır. Çalışmada, yöntemin başarısı asfalt kaplı yol ortamında gerçekleştirilen deneylerle test edilmiş ve çalışmanın ait olduğu yıla göre, önerilen yöntemin yayaları, diğer hareketli nesnelerden ayırt etme konusundaki başarısı vurgulanmıştır.

1997'de Kehtarnavaz ve Rajkotwala tarafından yapılan çalışmada, Texas Instruments TMS320C40 dijital sinyal işlemcisi kullanılarak, karşıdan karşıya geçmek için bekleyen yayalar gerçek zamanlı olarak tespit edilmiştir [7]. Çalışmada, aydınlatma koşullarındaki değişiklikleri hesaba katmak ve sabit ve

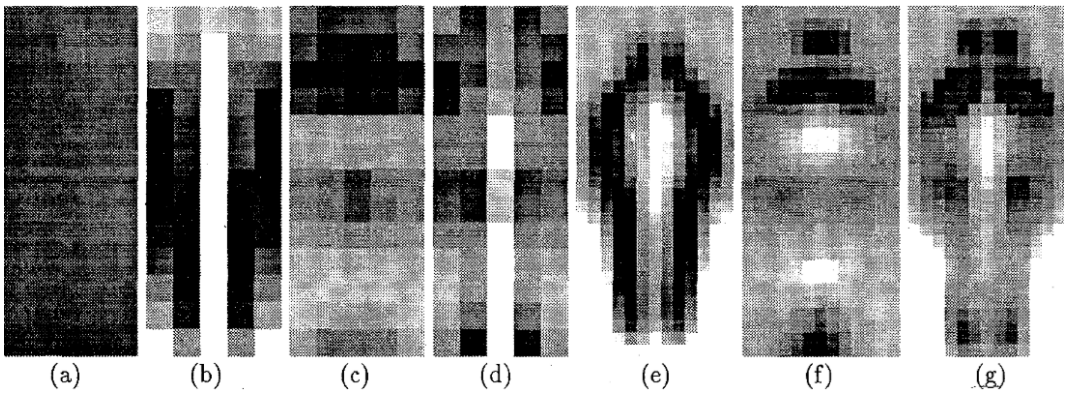


Şekil 1.2 Kehtarnavaz ve Rajkotwala'nın çalışmasına ait örnek iki çıktı [7]

hareketli yayaları ayırt etmek için arka plan güncelleme ve çıkarma tekniklerinin bir kombinasyonu kullanılmıştır. Sinyalizasyon amacıyla geliştirilen çalışma Şekil 1.2'de görüldüğü gibi ışıkta bekleyen yayaları tespit ederken, çalışmada hedeflendiği gibi karşıdan karşıya geçen hareket halindeki yayaları tespit etmiyor. Algoritmanın etkinliği, günün çeşitli zamanlarında ve hava koşullarında kapsamlı gerçek dünya testleri yoluyla doğrulanmış ve çalışmada trafik yönetimi amaçları için kabul edilebilir bir algılama oranına ulaşılmıştır.

1990'lar ve 2000'ler, daha önceleri de üzerinde çalışmalar yapılan makine öğrenmesi yöntemleri açısından devrimsel bir dönemdi. Her ne kadar tarihi 90'lardan çok öncesine uzansa da geri yayılım yönteminin keşfi ve bilgisayar teknolojilerindeki gelişmelerle makine öğrenimi için altın çağ başlamış oldu. Makine öğrenmesinin popülaritesindeki bu artış, 90'lar ortasından itibaren yaya tespiti alanında da etkisini göstermeye başladı.

1998'de Papageorgiou ve diğerleri tarafından sunulan çalışmada, istatistiksel analizle birleştirilmiş Şekil 1.3'deki gibi dalgacık tabanlı temsillerden yararlanarak, statik, karmaşık sahneler içindeki nesneleri tespit etmek için ölçeklenebilir bir sistem sunulmuştur [8]. Nesne sınıflarının kompakt ancak sağlam bir temsilini oluşturmak için bir dalgacık temel fonksiyonları sözlüğü kullanılmış ve bu daha sonra sınıf içi değişkenliği ele almayı ve çeşitli ortamlarda yanlış algılama oranlarını en aza indirmeyi amaçlayan bir destek vektör makinesi (SVM) sınıflandırıcısını eğitmek için kullanılmıştır. Çalışmada önerilen yöntem, yüz algılama ve yaya tanımlama olmak üzere iki farklı alanda test edilmiş ve sistemin nesne algılama görevlerinin doğasında olan karmaşıklığa uyum sağlama yeteneği ve çok çeşitli nesne algılama zorluklarına potansiyel olarak uygulanabilirliği ortaya koyulmuştur.

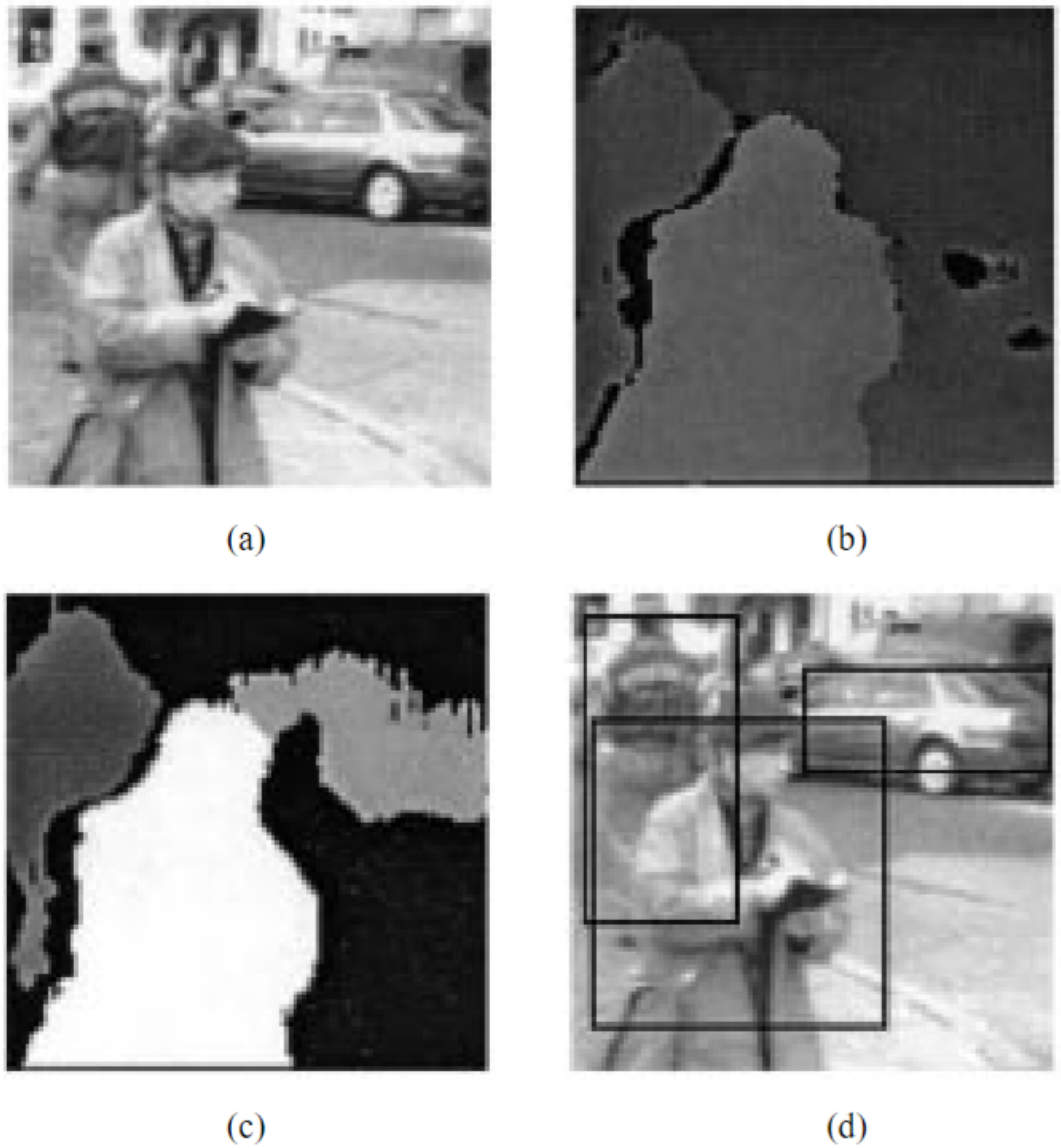


Şekil 1.3 Papageorgiou ve diğerlerinin çalışmasından alınan, farklı görüntülere ait dalgacık katsayılarının gri seviyeli temsilleri [8]. (a) Rastgele bir görüntünün dikey katsayıları. (b-d) 32 x 32 ölçekli insan görüntülerinin dikey, yatay ve köşe katsayıları. (e-g) 16 x 16 ölçekli insan görüntülerinin dikey, yatay ve köşe katsayıları

2000'de Zhao ve Thorpe, bir stereo kamera kullanarak karmaşık sahnelerde yaya tespiti için yeni, hızlı ve sağlam bir algoritma sunmuştur [9]. Bu yaklaşım, stereoskopik bölütlemeyi ve sinir ağı tabanlı tespiti üç adımlık bir süreçte birleştirir. Bu adımlar sırasıyla, derinlik farkı kullanılarak alt görüntü nesnesi adaylarına görüntü bölütlemesi gerçekleştirilmesi, bu adayların yaya boyutu ve şekli kısıtlamalarına göre birleştirilmesi ve bölünmesi ve yaya tanıma için eğitimli bir sinir ağı için girdi olarak aday alt görüntülerin yoğunluk gradyanlarının kullanılmasıdır. Şekil 1.4'de, çalışmada kullanılan stereo kameraya ait bir görüntü ve bu görüntüde derinlik farkı kullanılarak gerçekleştirilen bölütleme ve tespit sonuçları paylaşılmıştır. Gerçek zamanlı olarak çalışan bu yöntem, çeşitli pozlar, nesne boyutları, giysiler ve tıkanma (oklüzyon) durumları karşısında ve aydınlatma ve arka plandaki değişikliklere karşı dayanıklıdır. Çalışmada, geri yayılım algoritmasıyla eğitilmiş beş birimden oluşan tek gizli katmanlı bir sinir ağı, yayaları diğer nesnelerden ayırıyor, hareket ipuçlarından ziyade şekil özelliklerinden yararlanıyor ve yüksek şekil değişkenliğine rağmen hem hareketli hem de sabit yayaların tespit edilmesine olanak tanıyor.

2002 senesinde Gavrila ve Giebel'in yaptığı çalışma, şekil tabanlı nesne tanımaya odaklanarak, Chamfer sistemi kullanan yaya algılama ve izlemenin deneysel bir değerlendirmesini sunmaktadır [10]. Çalışma, sistemin şablonlarını Şekil 1.5'deki gibi hiyerarşik bir yapı halinde nasıl düzenlediğinin ana hatlarını çizerek başlamaktadır. Bu yapı, sistemin eşleştirmeyi geniş bir perspektiften gerçekleştirmesine ve giderek daha ayrıntılı analizlere doğru ilerlemesine olanak tanımaktadır. Bunu, mesafe ölçümlerine göre dönüştürülen görüntüleri ilişkilendiren ve farklı dönüşümler arasında etkili bir karşılaştırma yapılmasına

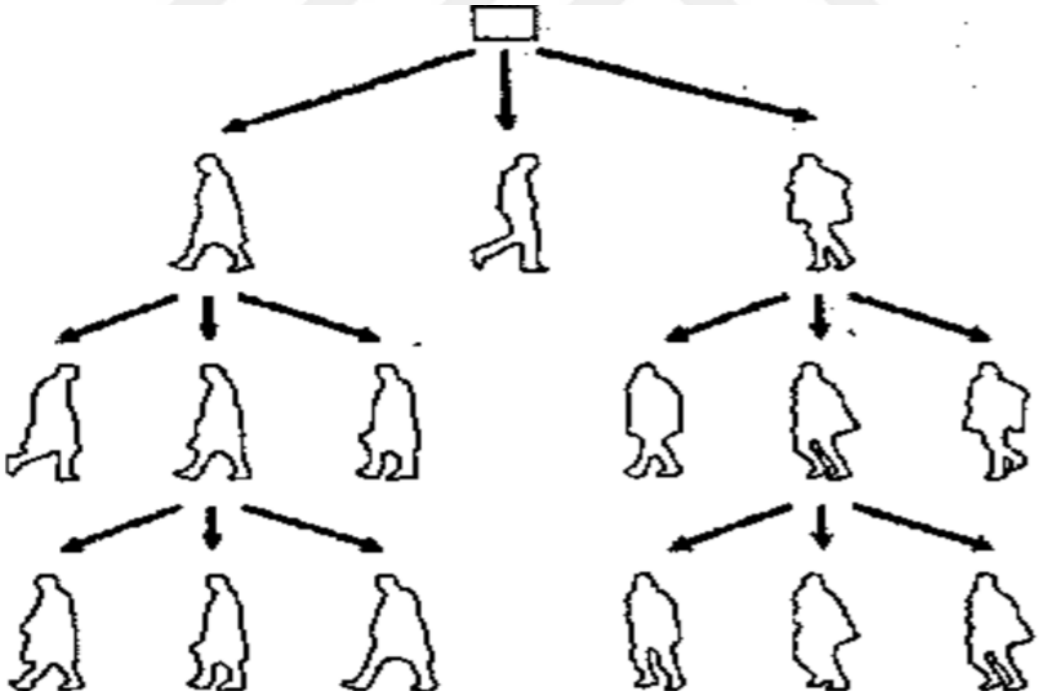
olanak tanıyan bir süreç aracılığıyla gerçekleştirmektedir. Doğruluğunu artırmak için sistem daha sonra bir sinir ağı kullanır. Bu ağ, çok çeşitli doku özelliklerini analiz etmek ve eşleştirme sürecine başka bir iyileştirme katmanı eklemek üzere tasarlanmıştır. Her adım bir öncekinin üzerine inşa edilerek nesnelerin kapsamlı ve kesin bir şekilde tanımlanmasını sağlamaktadır. Çalışmada, 4762 yaya görüntüsünden oluşan kapsamlı bir veri kümesi kullanılmıştır ve sistem performansı alıcı işletim karakteristiği (ROC) eğrileri aracılığıyla değerlendirilmiştir. Ek olarak, bir alfa-beta izleyici, zaman içinde algılamayı kullanarak, tıkanmaları ve gürültüyü gidermekte ve böylece algılama güvenilirliğini artırmaktadır.



Şekil 1.4 Zhao ve Thorpe'nin çalışmasından alınan görüntü ve çalışmanın aşamaları [9]. (a) stereo kameranın sol görüntüsü, (b) derinlik farkı haritası, (c) segmentasyon sonucu, (d) önerilen yöntem ile tespit edilen nesneler

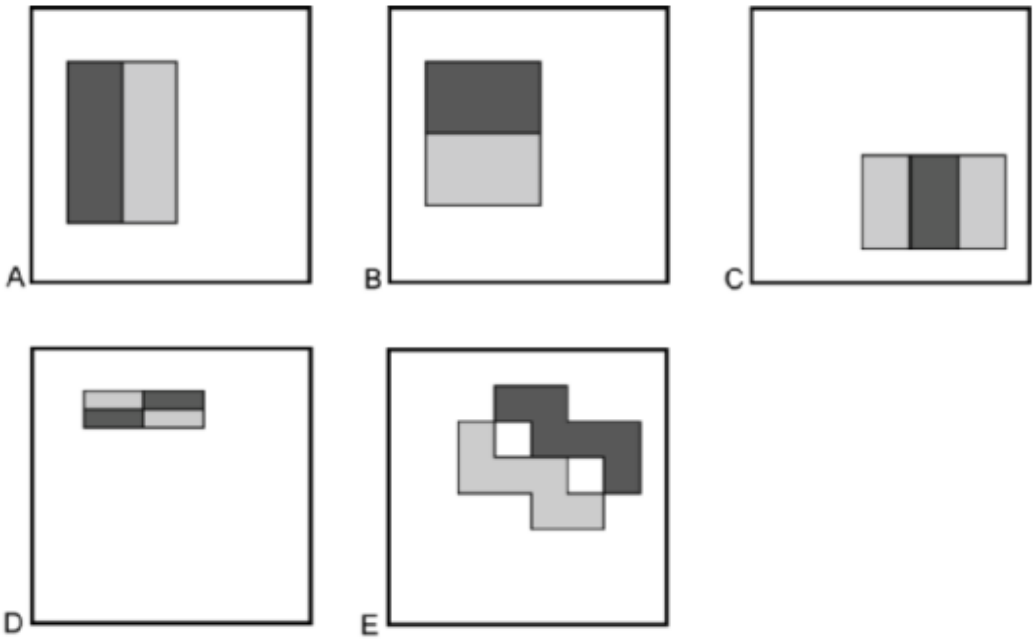
2003'de Viola ve diğerleri tarafından yapılan çalışma, yürüyen bireyleri tanımlamak için video dizilerinden elde edilen hem hareket hem de görünüm bilgilerinden

yararlanan yeni bir yaya tespit sistemi sunmaktadır [11]. Sistem, dedektörü yaya şekillerini etkili bir şekilde tanıyacak şekilde eğitmek için adaptif artırmayı (AdaBoost) kullanarak bir dedektörü ardışık çerçeveler üzerinden tarayan bir algılama algoritması kullanır. Önerilen sistem saniyede yaklaşık 4 kare hızında verimli bir şekilde çalışmakta, yağmur ve kar gibi zorlu çevre koşullarında düşük çözünürlüklü görüntülerde yayaları tespit edebilmekte ve düşük yanlış pozitif oranına ulaşmaktadır. Tespit mekanizması, hareket analizi için Şekil 1.6'daki gibi çeşitli ölçeklerde ve yönlerde hızlı hareket kapasitesine sahip dikdörtgen filtreler kullanmaktadır. Yaklaşım farklı filtre türlerini birleştirir: hareket yönü filtreleri, hareket kesme filtreleri, hareket büyüklüğü filtreleri ve görünüm filtreleri. Bunların tümü hızlı değerlendirme için bir "integral görüntü" tekniği ile işlenmektedir. Sistemin önemli bir bileşeni olan sınıflandırıcı, AdaBoost öğrenme yoluyla hareket ve görünüm özelliklerini birleştirerek geniş bir kümeden en düşük ağırlıklı hataya sahip özellikleri seçmektedir. Sistem, hesaplama verimliliğini ve doğruluğunu yönetmek ve algılama hızlarını ve hatalı pozitifleri dengelemek için algılama eşiklerini dinamik olarak ayarlayan kademeli bir mimari kullanmaktadır. Eğitim süreci, algılama doğruluğu ile hesaplama verimliliği arasında bir denge kurmayı hedefleyerek kademeyi yinelemeli olarak iyileştirmektedir.



Şekil 1.5 Gavrila ve Giebel'in çalışmasından alınan bir yaya şekilleri hiyerarşisi örneği [10]

2005'de Cheng ve diğerlerinin yaptığı çalışmada, örneklerden öğrenilen seyrek gabor filtrelerine dayalı bir özellik temsil yönteminden yararlanılarak, hareket ipuçlarını kullanmadan yaya tespitine odaklanılmış ve tespit aşaması için bir



Şekil 1.6 Viola ve diğerlerinin çalışmasında paylaşılan örnek filtreler [11]

destek vektör makinesi kullanılmıştır [12]. Makale, özellikle yaya görünümündeki farklılıkların, katı olmayan vücut hareketlerinin ve değişken arka planların algılama çabalarını zorlaştırdığı karmaşık ortamlarda, yaya algılama sırasında ortaya çıkan önemli zorlukları özetlemektedir. Önerilen tespit çerçevesi, yaya görüntülerinin benzersiz özelliklerini karakterize eden ve etkili sınıflandırma için destek vektör makinesi kullanan, seyrek bir dizi Gabor temel fonksiyonunun öğrenilmesi etrafında yapılandırılmıştır. Çalışma, yaya tespit doğruluğunda umut verici sonuçlar ortaya koymuştur, ancak önerilen yöntem belirli koşullar altında yayaları benzer arka plan unsurlarından ayırmada zorluklarla karşılaşmaktadır.

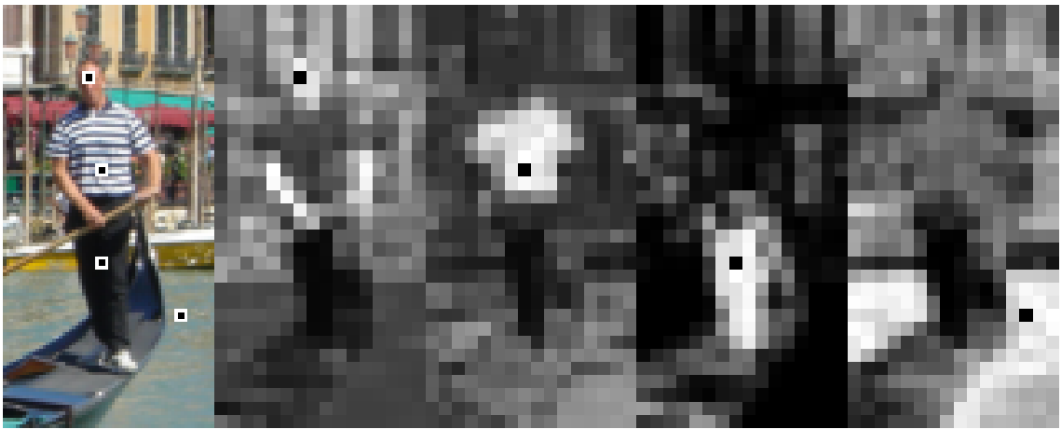
2008'de Tüzel ve diğerlerinin yaptığı çalışma, nesne tanımlayıcıları olarak kovaryans matrislerini kullanarak hareketsiz görüntülerdeki yayaları tespit etmek için yeni bir algoritma sunmaktadır [13]. Bu yöntem, nesneleri vektör olmayan bir uzayda, özellikle, bağlantılı bir Riemann manifoldu oluşturan d boyutlu tekil olmayan kovaryans matrislerinin uzayında sınıflandırma zorluğunu ele aldığı için öne çıkıyor. Bu araştırmanın temel katkısı, manifoldun geometrik özelliklerinden yararlanılarak Riemann manifoldu üzerindeki noktaların sınıflandırılmasına yönelik yeni bir yaklaşımın tanıtılmasıdır. Önerilen algoritma, Ulusal Bilgisayar Bilimi ve Otomasyon Araştırma Enstitüsü (INRIA) ve DaimlerChrysler yaya veri kümelerinde Şekil 1.7'deki örneklerde görülebileceği üzere çok başarılı tespit sonuçları elde etmiştir [14, 15]. Bu makale, kovaryans matrisleri doğal olarak geleneksel makine öğrenimi tekniklerine uygun bir vektör uzayı oluşturmadığından, verilerin geometrik yapısını dikkate almanın öneminin altını

çizmektedir. Araştırmacılar, bu matrisleri Riemann geometrisi çerçevesinde temsil ederek, verilerin altında yatan matematiksel özelliklere önem veren gelişmiş sınıflandırma stratejilerini uygulayabilmektedir. Bu yaklaşım, sınıflandırma için manifold üzerindeki noktaların Riemann ortalamasını hesaplamayı içermektedir. Ayrıca çalışma, bölge kovaryanslarının etkili bir şekilde hesaplanması için integral görüntülerin kullanımı ve özellik optimizasyonu için güçlendirme ile birleştirilmiş açgözlü bir özellik seçme algoritmasının uygulanması da dahil olmak üzere, bu sınıflandırma tekniğinin uygulanmasının pratik yönlerini araştırmaktadır.



Şekil 1.7 Tüzel ve diğerlerinin yaptığı çalışmaya ait örnek sonuçlar [13]

2010'da Walk ve diğerlerinin yaptığı çalışma, gelişmiş algılama performansı için sınıflandırıcı değerlendirmesine, yeni özelliklerin ve içgörülerin dahil edilmesine odaklanarak yaya algılamada önemli ilerlemeler sağlamıştır [16]. Araştırmacılar iki ana katkı sunmaktadır: optik akıştan türetilen hareket özelliklerinin uygulanması ve renk kanalları üzerinde kendine benzerliğe dayalı yeni bir özelliğin tanıtılması. Yönlendirilmiş gradyan histogramları (HOG) ile birleştirildiğinde bu özellikler, çeşitli veri kümelerinde mevcut en gelişmiş tekniklerden %20'ye kadar önemli ölçüde daha iyi performans göstermektedir. Çalışma, görüntü dizilerindeki

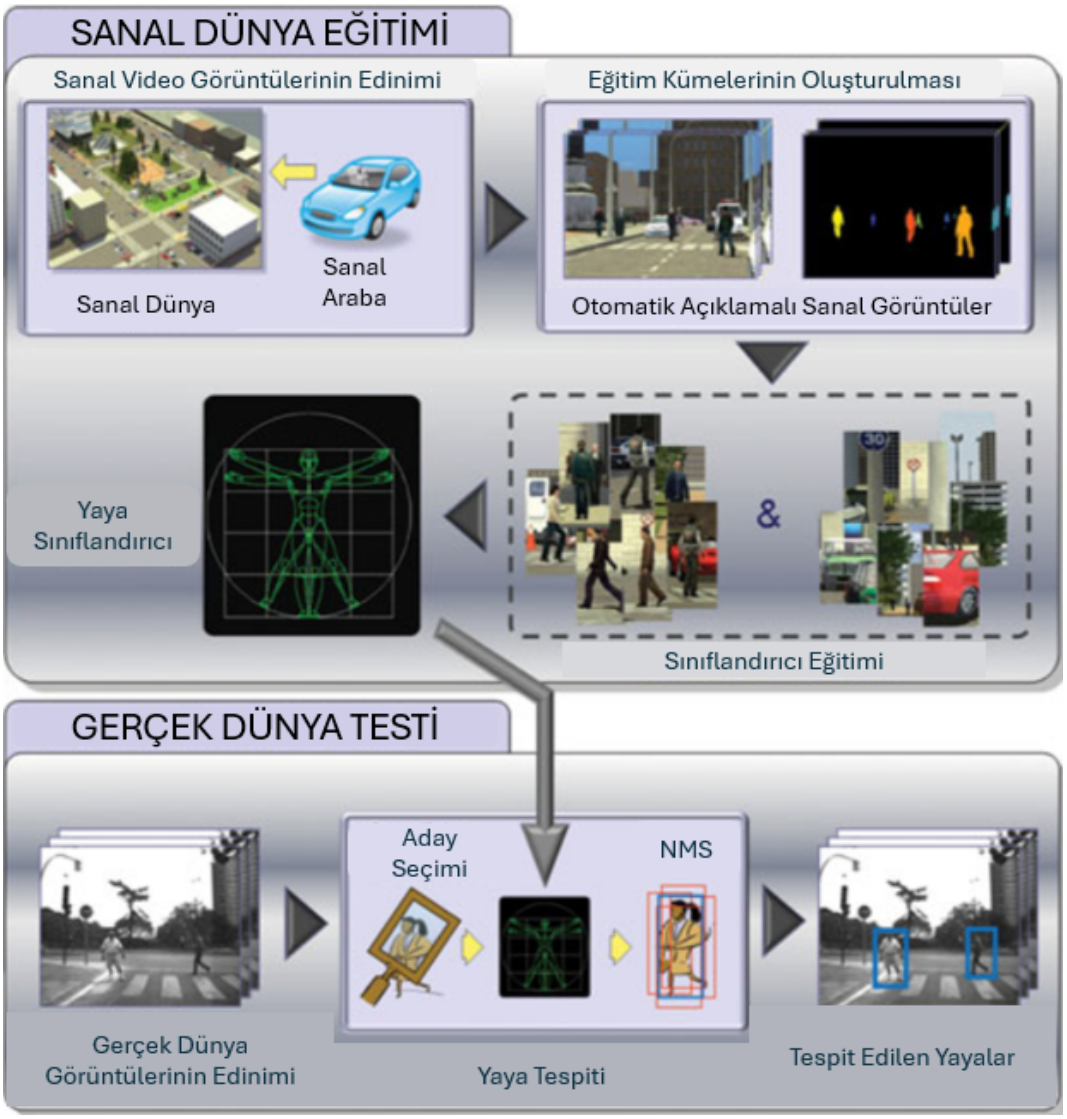


Şekil 1.8 Walk ve diğerlerinin çalışmasında paylaşılan işaretli bölgelere ait CSS çıktıları [16]

algılamayı geliştirmek için düşük kaliteli videolarda bile hareket özelliklerinin önemini vurgulamaktadır. Şekil 1.8'de örnek çıktıları verilen, uzamsal olarak yerelleştirilmiş renk dağılımlarının ikili istatistiklerini yakalayan, renk kendine benzerliği (CSS) özelliği hem statik görüntüler hem de video dizileri için algılama performansını bağımsız olarak artırır. Metodolojik olarak bu çalışma, nesne tespitinde özellik seçiminin ampirik doğasının altını çizerek, özellik değerlendirme ve sınıflandırıcı eğitimine yönelik titiz bir yaklaşım kullanmaktadır. Çalışma HOG, optik akış ve CSS özelliklerinin bir kombinasyonunu uygulayan ve doğrusal ve histogram kesişim çekirdeklerine sahip bir destek vektör makinesi yöntemini içermektedir. Çalışmada, yapılan kapsamlı değerlendirmelerle, dedektörün farklı kriterler ve koşullar karşısında üstün performansı ortaya koyulmuş ve özellikle zorlu tıkanma seviyeleri de dahil olmak üzere çeşitli senaryolarda yeni CSS özelliğinin katma değeri vurgulanmıştır.

2013'de Vazquez ve diğerlerinin yaptığı çalışma, pahalı ve öznel bir görev olan elle açıklama eklemeyle ilgili zorlukların üstesinden gelmek için hem sanal hem de gerçek dünya verilerinden yararlanarak yaya tespitine yeni bir yaklaşım getirmiştir [17]. Araştırmacılar, bir yaya sınıflandırıcısını eğitmek için sanal dünya açıklamalarını, sınırlı sayıda gerçek dünya verisiyle birleştiren, V-AYLA adı verilen bir alan uyarlama çerçevesi önermektedir. Eğitim ve test yöntemi Şekil 1.9'da paylaşılan teknik, sınıflandırıcıların sanal ortamlardan türetilen açıklamalı örnekler üzerinde eğitildiği ve daha sonra gerçek dünya örneklerinin küçük ama önemli bir alt kümesi kullanılarak uyarlandığı, ayrımcı öğrenmeye dayanmaktadır. Bu uyarlama süreci, sınıflandırıcıları farklı alanlarda eğitirken ve test ederken yaygın bir sorun olan veri kümesi kaydırma sorununu en aza indirmek için tasarlanmıştır. V-AYLA, başlangıçta eğitilmiş sanal dünya tabanlı sınıflandırıcının tespit etmesi zor olan gerçek dünyadaki yaya görüntülerini seçici olarak örneklemek

için aktif öğrenme stratejileri kullanmakta ve böylece sınıflandırıcının gerçek dünya senaryolarındaki performansını artırmaktadır. Çalışma, yönlendirilmiş gradyan histogramları (HOG) ve yerel ikili desenler (LBP) de dahil olmak üzere çeşitli tanımlayıcıların hem sanal hem de gerçek dünya bağlamlarındaki performansını kapsamlı bir şekilde değerlendirmiş ve sanal ortamlarda eğitilmiş bir sınıflandırıcının, kapsamlı gerçek dünya verileri üzerinde eğitilmiş bir sınıflandırıcıyla karşılaştırılabilir doğruluk elde edebileceğini göstermiştir.



Şekil 1.9 Vazquez ve diğerlerinin kullandığı eğitim ve test yöntemi [17]

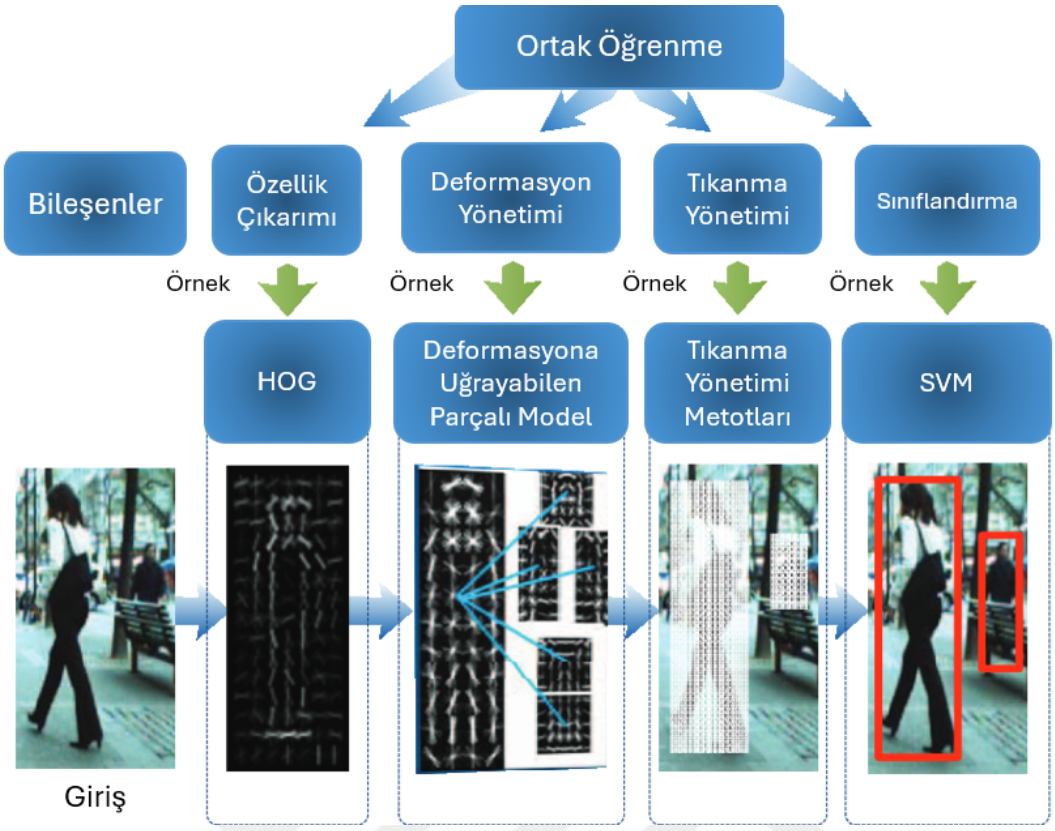
Makine öğrenimindeki popülarite 2000 ortalarında ve özellikle 2010'larda yerini derin öğrenmeye bırakmaya başladı. Önceleri destek vektör makineleri ve karar ağaçları gibi renk histogramları ve doku özellikleri tarzı elle özellik çıkarımı gerektiren yöntemler yerine, evrişimli sinir ağları (CNN) gibi özellik çıkarımını doğrudan ağın yaptığı derin öğrenme yöntemleri, elde ettikleri başarıları sebebiyle geniş bir kabul gördü. Standart makine öğrenimi tekniklerinin hem elle

özellik çıkarımı gerektirmesi hem de çıkarılan özelliklerin çoğu durumda farklı ortamlardaki insan görünümlerinin karmaşıklığını ve değişkenliğini yakalamakta derin öğrenmeye kıyasla geride kalması sebebiyle yaya tespiti alanında da araştırmacılar derin öğrenme kullanma yönüne gittiler. Araştırmacılar tarafından kullanılan büyük veri kümeleri ile eğitilen derin ağlarla birlikte yaya dedektörleri farklı koşullar altındaki sağlamlık ve başarısını önemli ölçüde artırdı.

2005'de Szarvas ve diğerlerinin yaptığı çalışma, poz, arka plan ve aydınlatma değişkenliği de dahil olmak üzere çeşitli koşullar altında yayaların tanımlanmasında daha iyi doğruluk sağlamak için evrişimli sinir ağlarının yeteneklerinden yararlanan yeni bir yaya algılama sistemi sunmaktadır [18]. Bu araştırmada kullanılan evrişimli sinir ağı modeli, özellikle yaya algılama görevi için özellik gösterimini otomatik olarak optimize ederken aynı zamanda sinir ağının genelleştirme yeteneklerini geliştirmek için düzenleme tekniklerini de uygulamaktadır. Bu çalışmanın ayırt edici bir yönü, önerilen evrişimli sinir ağı tabanlı sınıflandırıcı ile Haar-dalgacık özelliklerini kullanan geleneksel destek vektör makinesi (SVM) sınıflandırıcısı arasındaki karşılaştırmalı analizdir. Sonuçlar, evrişimli sinir ağı sınıflandırıcının yanlış pozitif oranı açısından destek vektör makinesinden önemli ölçüde daha iyi performans gösterdiğini göstermektedir. Bu araştırmada önerilen yenilikçi bir yaklaşım, destek vektör makinelerinden geniş marjlı eğitim konseptinin evrişimli sinir ağlarının eğitimine uygulanmasını içermektedir. Bu, evrişimli sinir ağının performansında önemli bir iyileşme ile sonuçlanmıştır. Çalışma, sistemin ilgi alanı tespiti, özellik çıkarma, aday sınıflandırma ve izleme modülleri bağlamında sınıflandırıcının performansı da dahil olmak üzere tüm tespit sisteminin kapsamlı değerlendirmesi, sistemin yetenekleri ve gelecekteki iyileştirmeler hakkında değerli bilgiler sağlamıştır.

2009'da Norouzi ve diğerlerinin yaptığı çalışma, standart kısıtlı boltzmann makinesi (RBM) modelini görüntülerin uzamsal yapısına uyacak şekilde değiştiren evrişimli kısıtlı boltzmann makinesinin (C-RBM) geliştirilmesi yoluyla tanıma için sınıfa özgü özelliklerin öğrenilmesine yönelik bir yöntem tanıtmaktadır [19]. Ağırlık paylaşımını kullanarak, evrişimli kısıtlı boltzmann makinesi çerçevesi, dört katmanlı bir filtreleme hiyerarşisi ve maksimum alt örnekleme yoluyla belirli bir nesne sınıfının görüntülerini oluşturabilen bir dizi özelliği öğrenebilir. Evrişimli kısıtlı boltzmann makinesi modelinin yeniliği, üretken bir çerçeve içinde uzamsal konumu ve ağırlık paylaşımını ele alma yeteneğinde yatmaktadır, böylece çok katmanlı evrişimli özellik dedektörlerinin katman bazında eğitilmesini kolaylaştırmaktadır.

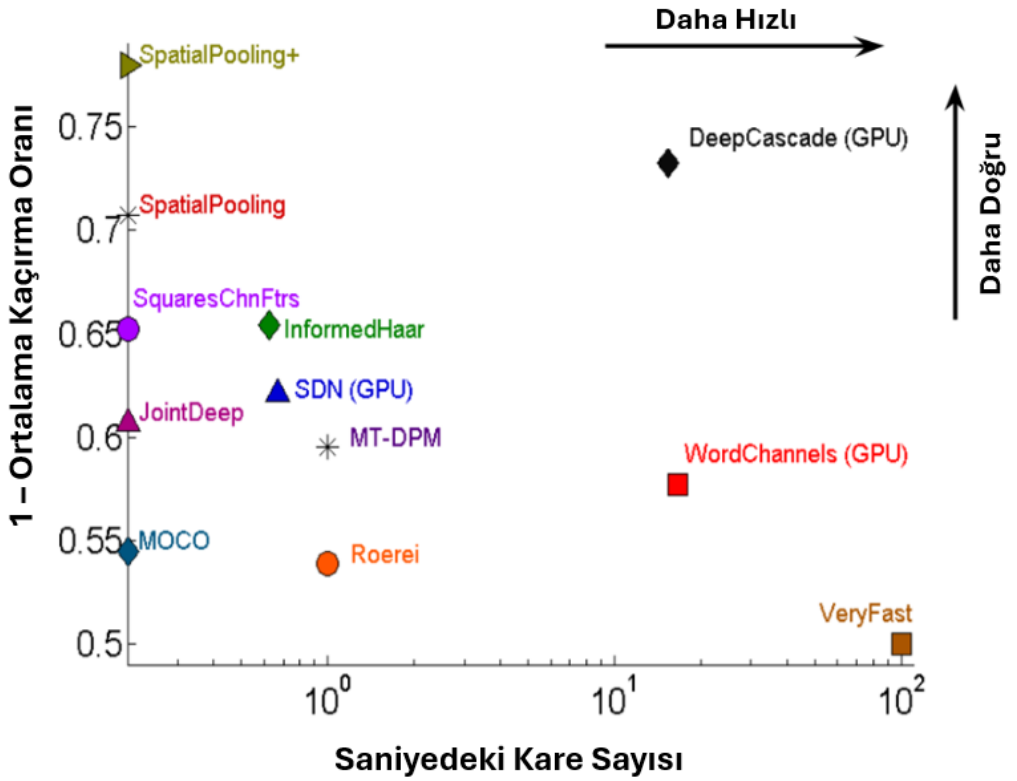
2013'de Ouyang ve Wang'ın yaptığı çalışma, yaya tespiti için ortak bir derin



Şekil 1.10 Ouyang ve Wang'ın ortaklaşa öğrenilmesini savunduğu dört bileşen [20]

öğrenme çerçevesi önermektedir [20]. Bu çalışmanın arkasındaki motivasyon, daha önceki çalışmalarda özellik çıkarma, deformasyon yönetimi, tıkanma yönetimi ve sınıflandırmanın geleneksel olarak aralarındaki potansiyel ilişkiyi keşfetmeden tek tek veya sıralı olarak öğrenilmesinin performansı olumsuz olarak etkilediği düşüncesidir. Araştırmacılar Şekil 1.10'da paylaşılan birleşik bir yaklaşımı savundular ve bu bileşenlerin kolektif güçlerini en üst düzeye çıkarmak için ortaklaşa öğrenilmesi gerektiğini öne sürdüler. Bu bileşenleri tutarlı bir çerçeveye entegre eden, otomatik karşılıklı etkileşimi ve öğrenmeyi kolaylaştıran yeni bir derin öğrenme mimarisi sundular. Önerilen yöntem, evrişimli sinir ağlarına bir deformasyon katmanı ekleyerek yaya tespitini önemli ölçüde geliştiriyor ve modelin insan duruşu ve tıkanmalarındaki değişikliklere etkili bir şekilde uyum sağlamasına olanak tanıyor. Bu, insan şekillerinin ve hareketlerinin daha esnek ve doğru modellenmesini sağlamaktadır. Çerçeve, özellikleri doğrudan piksellerden öğrenerek ve deformasyon ve tıkanma işleme modelleriyle etkileşim yoluyla özelliklerin ayırt edicilik performansını artırmaktadır. Önerilen yöntem, zamanının en iyi performans gösteren yöntemleriyle karşılaştırıldığında ortalama kaçırma oranında %9'luk bir azalma sağlamıştır.

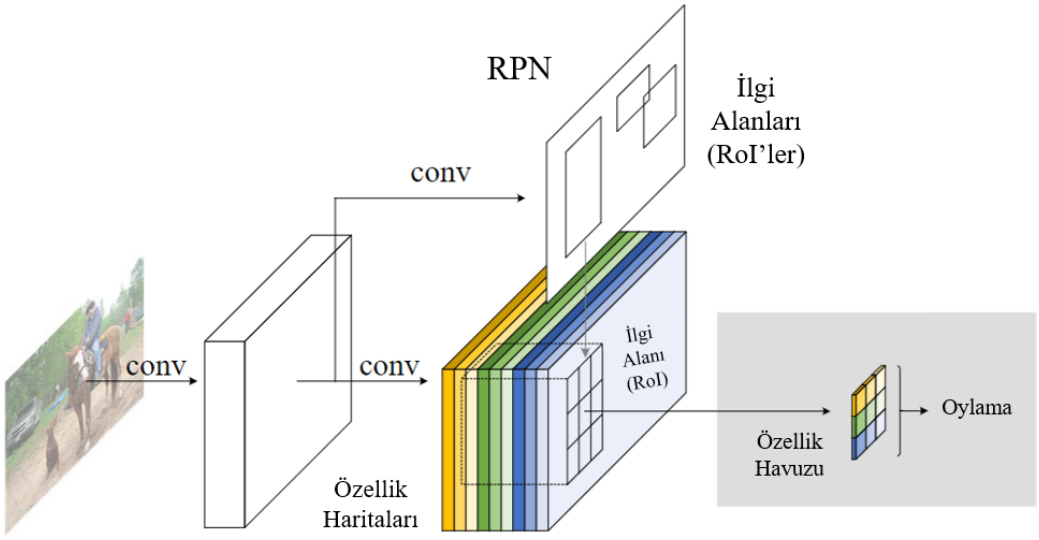
2015'de Angelova ve diğerleri tarafından yapılan çalışma, yaya tespiti için kademeli sınıflandırıcıların hızını derin sinir ağlarının hassasiyetiyle birleştiren yeni



Şekil 1.11 DeepCascade yönteminin hız ve doğruluk ekseninde daha önceki yöntemlerle karşılaştırılması [21]

bir gerçek zamanlı nesne algılama yöntemini tanıtmaktadır [21]. DerinKaskat (DeepCascade) adı verilen önerilen sistem, saniyede 15 kare işlemeyle yüksek doğruluk ve gerçek zamanlı performans elde etmek için her iki tekniğin güçlü yönlerinden yararlanıyor. Şekil 1.11'de görüleceği üzere DeepCascade, yaya tespitindeki iki temel zorluğa çözüm sunmaktadır: yüksek doğruluk elde etmek ve gerçek zamanlı işlem hızını korumak. Geleneksel yöntemler genellikle birini diğeri için feda eder, ancak DeepCascade, başlangıçtaki hızlı tespit aşamalarının kolay negatifleri hızla attığı ve daha karmaşık derin sinir ağı aşamalarının zor örneklere odaklanmasına izin verdiği katmanlı bir yaklaşım kullanmaktadır. Bu çalışmanın en önemli katkılarından biri, tüm görüntü parçalarını hızlı bir şekilde işleyen bir derin sinir ağı ile başlayıp, yüksek güvenirliğe sahip parçalar için daha karmaşık bir derin sinir ağı ile devam eden bir dizi derin sinir ağının yenilikçi kullanımıdır. Bu aşamalı işlem, algılama doğruluğundan ödün vermeden hesaplama süresini önemli ölçüde azaltmıştır. Sistemin mimarisi, evrişimli katmanların derinliğini ve boyutunu ayarlayarak hız için optimize edilmiştir. Bu, sınırlı hesaplama kaynaklarına sahip donanımlarda bile gerçek zamanlı yaya tespitini mümkün kılar. Ayrıca çalışma, eğitim veri kalitesinin ve miktarının derin sinir ağlarının performansı üzerindeki etkisini araştırmış ve ek, yüksek kaliteli eğitim verilerinin kullanılmasının algılama doğruluğunu artırabileceğini göstererek derin

öğrenmede veri hacminin ve çeşitliliğinin önemini vurgulamıştır. Araştırmacılar ayrıca, modelin genelleştirilebilirliğini ve yayaları arka plandaki dağınıklıktan ayırt etme yeteneğini geliştirmek için derin sinir ağlarının ImageNet gibi geniş bir veri kümesi üzerinde ön eğitiminin etkinliğini de göstermiştir [22].



Şekil 1.12 R-FCN mimarisi [23]

2016'da Dai ve diğerleri tarafından yapılan çalışma, hızlı/daha hızlı bölge tabanlı evrişimli sinir ağı (Fast/Faster R-CNN) gibi geleneksel bölge tabanlı dedektörlerde yapılan özelleştirmelerle, Şekil 1.12'de mimarisi paylaşılan bölge tabanlı tam evrişimli ağı (R-FCN) verimli ve doğru nesne tespiti için yeni bir yaklaşım olarak tanıtmıştır [23]. R-FCN, neredeyse tüm hesaplamanın görüntünün tamamında paylaşıldığı tamamen evrişimli bir tasarım kullanarak, bölge başına bir alt ağı defalarca uygulamanın hesaplama açısından verimsizliğini gidermektedir. R-FCN'in önemli bir yeniliği, görüntü sınıflandırması için dönüşüm değişmezliği ile nesne algılama için dönüşüm varyansının çelişen gerekliliklerini bağdaştıran, konuma duyarlı puan haritalarını tanıtmasıdır. Bu tasarım, R-FCN'nin artık ağlar (ResNet) gibi son teknoloji ürünü görüntü sınıflandırma ağlarını nesne algılama çerçevesine doğal bir şekilde entegre etmesine olanak tanımaktadır ve doğruluktan ödün vermeden verimliliği önemli ölçüde artırmaktadır. R-FCN çerçevesi, ilgi alanı (RoI) havuzlama adımından sonra öğrenilebilir katmanlara olan ihtiyacı ortadan kaldırarak mimariyi basitleştirir, böylece RoI başına hesaplama maliyetini azaltmaktadır ve sistemin uçtan uca eğitilmesine olanak tanımaktadır. Bu kolaylaştırılmış yaklaşım yalnızca yüksek doğruluk elde etmekle kalmaz, aynı zamanda hem eğitim hem de tespit hızlarını önemli ölçüde geliştirerek nesne algılama görevlerinde tamamen evrişimli tasarımların etkinliğini de gösterir.

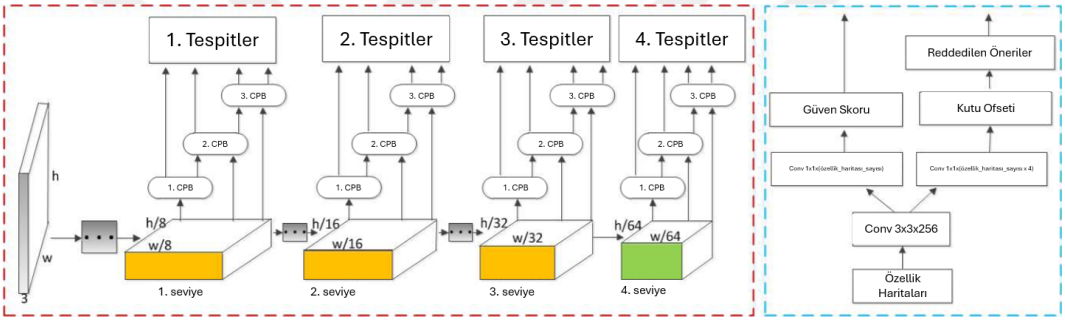
2018'de Lan ve diğerleri tarafından yapılan çalışma, yalnızca bir kez bak

Tablo 1.1 YOLO-R ağının yapısı [24]

Katman	Filtre	Boyut/Adım	Giriş	Çıkış
Konvolüsyon	32	3x3/1	416x416x3	416x416x32
Maks. Havuzlama		2x2/2	416x416x32	208x208x32
Konvolüsyon	64	3x3/1	208x208x32	208x208x64
Maks. Havuzlama		2x2/2	208x208x64	104x104x64
Konvolüsyon	128	3x3/1	104x104x64	104x104x128
Konvolüsyon	64	1x1/1	104x104x128	104x104x64
Konvolüsyon	128	3x3/1	104x104x64	104x104x128
Maks. Havuzlama		2x2/2	104x104x128	52x52x128
Konvolüsyon	256	3x3/1	52x52x128	52x52x256
Konvolüsyon	128	1x1/1	52x52x256	52x52x128
Konvolüsyon	256	3x3/1	52x52x128	52x52x256
Maks. Havuzlama		2x2/2	52x52x256	26x26x256
Konvolüsyon	512	3x3/1	26x26x256	26x26x512
Konvolüsyon	256	1x1/1	26x26x512	26x26x256
Konvolüsyon	512	3x3/1	26x26x256	26x26x512
Konvolüsyon	256	1x1/1	26x26x512	26x26x256
Konvolüsyon	512	3x3/1	26x26x256	26x26x512
Maks. Havuzlama		2x2/2	26x26x512	13x13x512
Konvolüsyon	1024	3x3/1	13x13x512	13x13x1024
Route	11			
Reorg		/2	26x26x256	13x13x1024
Konvolüsyon	512	1x1/1	13x13x1024	13x13x512
Konvolüsyon	1024	3x3/1	13x13x512	13x13x1024
Route	11			
Reorg		/2	26x26x256	13x13x1024
Konvolüsyon	512	1x1/1	13x13x1024	13x13x512
Konvolüsyon	1024	3x3/1	13x13x512	13x13x1024
Konvolüsyon	1024	3x3/1	13x13x1024	13x13x1024
Route	11			
Reorg		/2	26x26x256	13x13x1024
Konvolüsyon	1024	3x3/1	13x13x1024	13x13x1024
Route	12			
Reorg		/2	26x26x512	13x13x2048
Route	3230			
Konvolüsyon	1024	3x3/1	13x13x3072	13x13x1024
Konvolüsyon	30	1x1/1	13x13x1024	13x13x30

(YOLO) derin öğrenme çerçevesini geliştirerek, yaya tespitinde önemli bir ilerleme sunmuştur [24]. Üç geçiş katmanını orijinal YOLO ağına entegre eden bu çalışma, geleneksel olarak gradyanın kaybolmasına ve hatalı algılamaya yol açan, derin ağ katmanlarındaki ince taneli yaya özelliklerinin kaybını ele almaktadır. Yol (route) ve yeniden düzenleme (reorg) alt katmanlarından oluşan geçiş katmanı, sığ ve derin ağ özellikleri arasında etkili bir köprü kurarak algılama süreci boyunca

ayrıntılı yaya özelliklerinin korunmasını ve kullanılmasını sağlamaktadır. Yapılan değişiklik, geçiş katmanı bağlantısını orijinal 16. katmandan 12. katmana kaydırarak ağın temel yaya özelliği bilgilerini yakalama ve işleme yeteneğini optimize etmiştir. INRIA yaya veri kümesi üzerindeki deneysel değerlendirme, katmanları Tablo 1.1'de paylaşılan konfigüre edilmiş YOLO-R'ın, saniyede 25 karelik rekabetçi bir algılama hızını korurken algılama doğruluğunda dikkate değer bir artışla hassasiyet, geri çağırma ve kesişim bölü birleşim (IoU) ölçümlerinde temel YOLO modelini aştığını göstermektedir. Ayrıca, çalışmanın, stratejik katman entegrasyonu yoluyla düşük seviyeli özellik çıkarımını yüksek seviyeli anlamsal bilgilerle birleştirmeye yönelik yaklaşımı, derin öğrenme tabanlı nesne tespitindeki yaygın zorlukların üstesinden gelmek için pratik bir yöntem sergilemektedir. Bu katkı yalnızca yaya tespit sistemlerinin performansını artırmakla kalmamakta, aynı zamanda çeşitli bilgisayarlı görme uygulamaları için daha verimli ve doğru derin öğrenme modellerinin tasarımına ilişkin bilgiler de sunmaktadır.

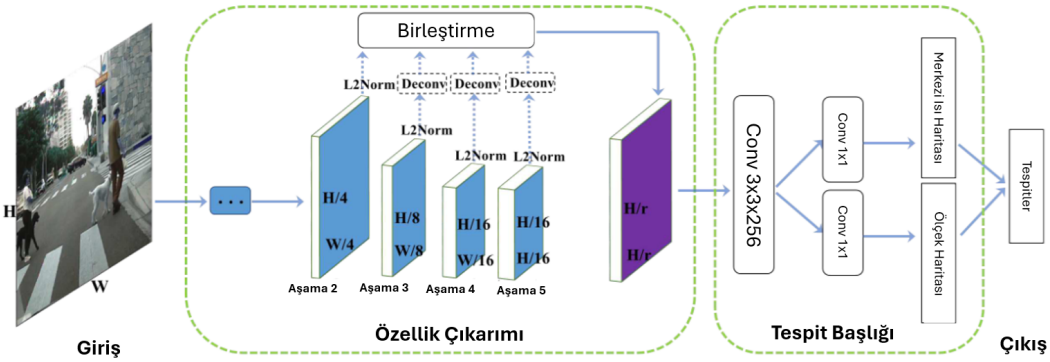


Şekil 1.13 ALFNet mimarisi ve CPB [25]

Yine 2018'de Liu ve diğerleri tarafından yapılan çalışma, Faster R-CNN gibi iki aşamalı dedektörlerin ve tek atış dedektörü (SSD) gibi tek aşamalı dedektörlerin geleneksel olarak karşılaştığı zorluklar olan hız ve doğruluk arasındaki boşluğu dolduran etkili bir tek aşamalı dedektör önererek yaya tespitine yeni bir yaklaşım sunmaktadır [25]. Araştırmacılar, varsayılan bağlantı kutularını giderek daha doğru tespit sonuçlarına yinelemeli olarak ulaştıran, yapısal olarak basit ama etkili bir bileşen olan asimptotik yerelleştirme uydurma (ALF) modülünü tanıtmıştır. Bu yinelemeli iyileştirme, artan kesişim bölü birleşim eşikleriyle daha kaliteli pozitif örneklerden ve daha zorlayıcı negatif örneklerden yararlanılmasına olanak tanıyarak tek atışlı çok kutulu algılayıcı tabanlı yaya tespitinde yaygın bir sorunu ele almaktadır. Şekil 1.13'de paylaşılan asimptotik yerelleştirme uydurma ağı (ALFNet) ve evrişimli tahmin bloğu (CPB), tek aşamalı dedektörlerin hesaplama verimliliğini korurken algılama doğruluğunu önemli ölçüde artırır ve ŞehirKişileri (CityPersons) ve Caltech gibi büyük ölçekli yaya algılama veri kümelerinde en son teknoloji performansı sergiler [26, 27]. Ayrıca, asimptotik yerelleştirme uydurma modülünün tasarımı esnek ve omurgadan bağımsız olup, nesne algılama alanında

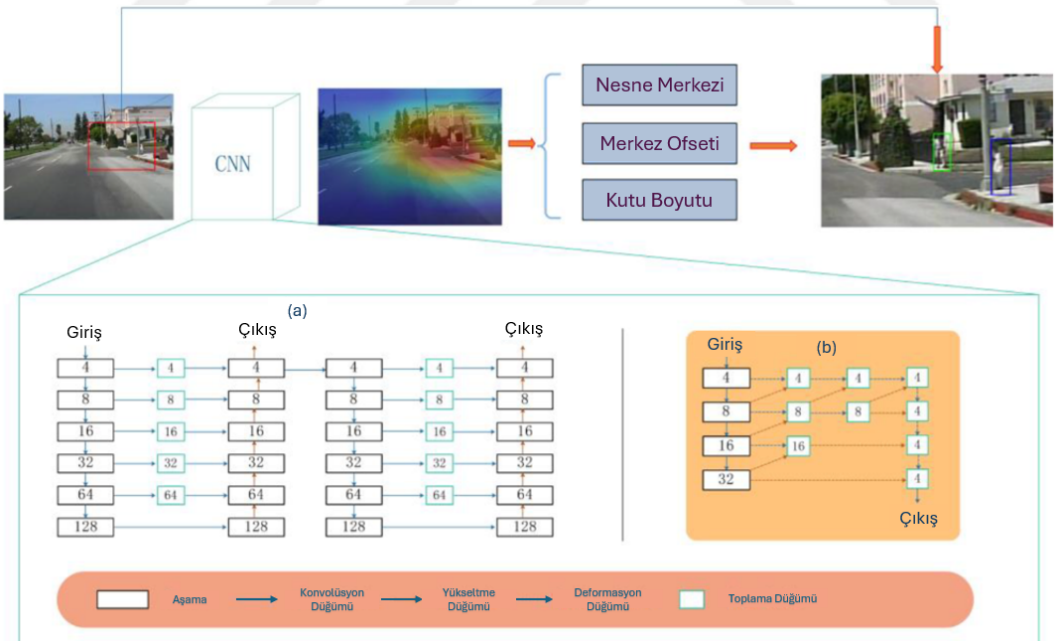
daha ileri uygulamalar ve iyileştirmeler için potansiyel olduğunu göstermektedir. Yapılan deneyler, asimptotik yerelleştirme uydurma ağının hem doğruluk hem de hız açısından üstün performansını doğrulayarak, onu gerçek zamanlı yaya tespiti gerektiren gerçek dünya uygulamaları için ilgi çekici bir çözüm olarak konumlandırıyor.

2019'da Liu ve diğerleri tarafından yapılan bir başka çalışma, yaya tespitini, evrişim işlemlerini kullanarak nesnelerin merkezlerini ve ölçeklerini tahmin etme gibi basit bir göreve indirgeyen, çapa içermeyen bir yöntem önermektedir [28]. Çalışmadaki temel fikir, nesne tespitini yüksek seviyeli anlamsal özellik tespit görevi olarak ele almaktır. Kenarlar veya köşeler gibi düşük seviyeli özelliklere odaklanan geleneksel özellik dedektörlerinden farklı olarak bu yaklaşım, yayaların merkezi noktalarını tespit etmeyi amaçlıyor ve modern derin öğrenme modellerinin yeteneklerinden yararlanarak yüksek seviyeli anlamsal bilgileri doğrudan görüntülerden soyutluyor. Yöntem aynı zamanda tespit edilen nesnelerin ölçeklerini de tahmin ederek, bir görüntü piramidini işlemeden boyutlarının tahmin edilmesini mümkün kılar ve böylece tespit görevini daha da basitleştirir. Şekil 1.14'de mimarisi paylaşılan merkez ve ölçek tahmini dedektörü (CSP) olarak adlandırılan önerilen dedektör, önceden tanımlanmış çapalara veya karmaşık işlem sonrası adımlara olan ihtiyacı ortadan kaldırarak çapasız bir şekilde çalışır. Yapısal basitliğine rağmen CSP, zorlu yaya tespit kriterlerinde rekabetçi doğruluk ve hız sergileyerek çeşitli koşullarda yayaları tespit etmek için güçlü bir yöntem olma potansiyeline sahiptir. CSP mimarisi, algılama görevini desteklemek üzere değiştirilmiş bir omurga ağı üzerine kuruludur. CSP'nin algılama kafası, evrişimli bir katmandan ve bunu takip eden iki tahmin katmanından oluşur, biri merkezi konum için, diğeri ise tespit edilen yayaların ölçeği için. Bu tasarım, doğrudan evrişimli işlemler yoluyla verimli ve doğru yaya tespitine olanak tanır ve bağlantı kutusu sınıflandırması ve regresyon gibi ek ve genellikle zahmetli işleme adımlarına olan ihtiyacı ortadan kaldırır. Çalışmada, önerilen CSP, CityPersons ve Caltech gibi standart yaya algılama veri kümeleri üzerinde test edilmiş ve başarılı değerlendirme sonuçları paylaşılmıştır.



Şekil 1.14 CSP mimarisi [28]

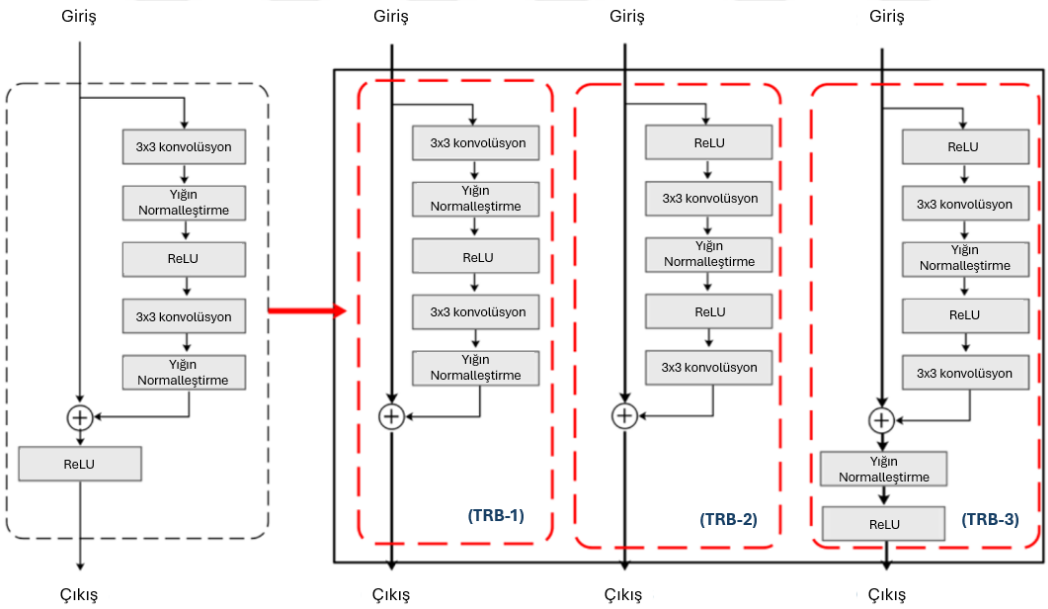
2020'de Tumas ve diğerleri tarafından yapılan çalışma, derin öğrenme tekniklerini kullanarak zorlu hava koşullarında yaya tespiti gibi zorlu bir görevi ele almıştır [29]. Çalışmada, olumsuz hava koşulları altında dört büyük Avrupa Birliği ülkesinde çekilen çok çeşitli ince taneli, açıklamalı görüntüleri kapsayan ZUT adlı yeni bir 16 bitlik termal veri kümesi tanıtılmıştır. Çalışmada, 16 bit derinlikli termal görüntüleri işleyecek şekilde değiştirilmiş, yalnızca bir kez bak versiyon 3 (YOLOv3) derin sinir ağının yaya tespiti için uyarlanması ve test edilmesi gerçekleştirilmiştir. Bu değişiklik önemlidir çünkü dedektörün daha geniş bir termal veri spektrumunu kullanmasına olanak tanımıştır ve gece veya olumsuz hava koşulları gibi zorlu görüş koşullarında yayaları ayırt etme yeteneğini artırmaktadır. Geliştirilmiş YOLOv3 modeli, %89,1'e kadar ortalama hassasiyetin ortalaması (mAP) değeri elde ederek, derin öğrenme tabanlı yaya algılama sistemleri için 16 bitlik görüntülerin kullanılmasının etkinliğini göstermektedir. Ayrıca çalışma, veri kümesinin ve derin öğrenme modelinin kapsamlı bir değerlendirmesini içeriyor ve çeşitli mesafeler ve hava koşullarında yaya algılama performansına ilişkin bilgiler sağlıyor. Bu değerlendirme, sis, yağmur veya düşük ışık koşulları gibi geleneksel görsel kameraların başarısız olabileceği senaryolarda yaya tespitini iyileştirmek için derin öğrenmeyi termal görüntülemeyle entegre etmenin başarıma etkisini açığa çıkarmaktadır.



Şekil 1.15 CenterNet mimarisi [30]

2023'de Wang ve diğerleri tarafından yapılan çalışma, özellikle Şekil 1.15'de görülen merkezi nokta sinir ağı (CenterNet) mimarisi içindeki artık blokların kullanımına odaklanan, derin öğrenme yaklaşımını kullanan, yeni bir küçük ölçekli yaya algılama yöntemini tanıtıyor [30]. Geleneksel algoritmalar küçük ölçekli

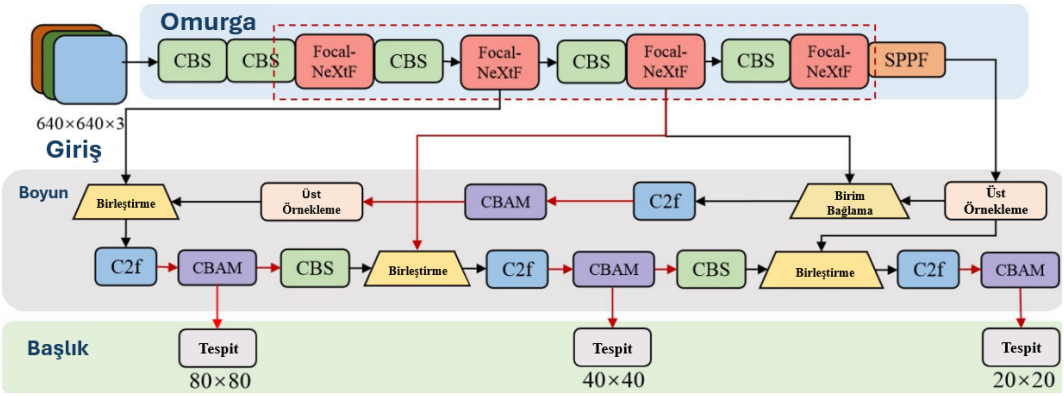
yayaları tespit etmekte zorlanır ve çoğu zaman tespit edilmesi gereken yayaların gözden kaçmasına neden olur. Bununla mücadele etmek için araştırmacılar, Şekil 1.16'da paylaşılan üç artık ağ bloğunun (TRB), CenterNet tespit modeline entegrasyonunu temel alan bir model önerdiler. Bu çok bloklu grup, ağ yapısı içindeki bilgi akışını geliştirerek daha doğru tespit sonuçlarına olanak tanımaktadır. Model, alt örneklemenin ilk aşamalarında aktivasyon fonksiyonlarına olan bağımlılığı azaltmak için TRB'yi kullanır, böylece daha fazla bilgi korunur ve tespit doğruluğu artar. Ayrıca çalışmada, modeli ondalık değerlere duyarlı hale getirerek algılama performansını daha da artıran geliştirilmiş bir aktivasyon fonksiyonu olan düzeltilmiş doğrusal birim 6 (ReLU6) tanıtılmıştır. Caltech ve Federal Teknik Yüksekokulu (ETH) yaya veri kümeleri üzerinde gerçekleştirilen kapsamlı deneyler, yöntemin özellikle küçük ölçekli yaya tespiti için iyi bir doğruluk ve hıza ulaştığını göstermektedir [31]. Önerilen merkezi nokta sinir ağı tabanlı üç artık sinir ağı bloğu tespit modeli (TRBBC-Net), çoklu TRB bloklarını ve yaya konumları için uygulanan regresyonu birleştirerek algılama hızı ve doğruluk arasında bir denge kurmaktadır. Özetle, çalışma, özellikle artık blok gruplarının CenterNet mimarisine entegrasyonu ve geliştirilmiş bir aktivasyon fonksiyonunun tanıtılması yoluyla, derin öğrenme tekniklerini etkili bir şekilde kullanan yeni bir yöntem önererek yaya algılama alanına katkıda bulunmuştur.



Şekil 1.16 TRB yapısı [30]

2024'de Liu ve diğerleri tarafından yapılan çalışma, Şekil 1.17'de paylaşılan araç içi kamera sistemleri için tasarlanmış, özellikle karmaşık cadde veya kavşak senaryolarında iyi sonuç veren, gelişmiş bir yoğun yaya algılama algoritması olan katmanlar arası birleşim ve dikkat mekanizması iyileştirilmiş yalnızca bir kez bak versiyon 8 (YOLOv8-CB)'i tanıtmaktadır [32]. Çalışmada,

yalnızca bir kez bak versiyon 8n (YOLOv8n) çerçevesine basamaklı füzyon ağı (CFNet), çift yönlü ağırlıklı özellik füzyon yöntemi (BIFPN) ve evrişimsel blok dikkat modülü (CBAM) entegre edilerek, önerilen YOLOv8-CB elde edilmiştir. Geleneksel çaprazdan birleşime (C2F) modülünün yerini alan CFNet, hafif kademeli füzyon yaklaşımından yararlanarak çok ölçekli özellik semantiğini ve konum bilgisi çıkarımını geliştirir. BIFPN, çift yönlü bir özellik birleştirme yolu oluşturarak, farklı ölçeklerdeki özelliklerin asimilasyonunu optimize ederek modelin verimliliğini daha da artırır. CBAM'ın dahil edilmesi, ağın odağını ilgili özellik kanallarına yönlendirerek küçük ve kapalı yaya hedeflerinin tespit doğruluğunu artırır. Deneysel değerlendirmeler, YOLOv8-CB'nin YOLOv8n'e kıyasla doğrulukta %2,4'lük bir artış, model parametrelerinde %6,45'lik bir azalma ve hesaplama yükünde %6,74'lük bir azalma elde ettiğini ortaya koymaktadır.



Şekil 1.17 YOLOv8-CB mimarisi [32]

1.2 Tezin Amacı

Yaya tespiti alanında literatürün dönemlere göre ilerleyişinde paylaşıldığı üzere derin öğrenmenin ortaya çıkışı ile yaya tespit sistemlerinde başarımlar artmış ve hatalar oldukça düşmüştür. Ancak özellikle otonom araçlar ve otonom sürüş sistemleri gibi uygulamalarda, idealde bu sistemlerin kusursuz olarak çalışması hedeflenmektedir. Bu ideallere mümkün olduğunca yakın sonuçlar elde etmek için kullanılan derin ağların çok kapsamlı veri kümelerinde, çok değişken arka planlarda ve mümkün olduğunca çok benzersiz insan verisiyle eğitilmesi gerekmektedir. Ancak bu tarz kapsamlı veri kümelerinin bir kişi ya da bir topluluk tarafından oluşturulması çok büyük bir emek ve iş gücü gerektirdiği için, bu şekilde kapsamlı veri kümelerinin tek bir elden oluşturulması pek mümkün değildir. Ayrıca bu tarz kapsamlı veri kümeleri oluşturulsalar bile yaya tespiti gibi alanlarda oluşturulan ve kişisel veriler içeren veri kümelerinin paylaşılması kişisel veri gizliliği problemlerini beraberinde getirecektir. Gerçekleştirilen çalışma ideal yaya

dedektörlerini oluşturmak için karşılaşılan bu problemleri odağına almıştır ve bu problemlerin çözümü için yenilikçi bir yaklaşım getirmeye çalışmaktatır.

1.3 Hipotez

Yapılan çalışma ideal yaya dedektörlerini oluşturmak için karşılaşılan zorlukların üstesinden gelmek için federe öğrenmenin potansiyelini araştırmaktadır. Federe öğrenme, farklı verilere sahip kullanıcıların tek bir küresel ağı kendi verileriyle eğitmesine imkan sağlar ve kullanıcıların verileri bir başka kullanıcı ile veya bir sunucu ile paylaşması gerekliliğini ortadan kaldırır. Bu sayede birçok kullanıcı kendi verilerini kullanarak tek bir küresel ağı eğitir. Bu, ideal yaya dedektörleri oluşturmak için gerek duyulan yüksek miktarda veri, benzersiz insan temsili ve çeşitli arka plan gerekliliği için iyi bir çözüm oluşturmaktadır. Ayrıca, kullanıcıların verilerini doğrudan paylaşmasına gerek olmadığı için olası kişisel veri gizliliği problemlerinin de önüne geçilmiş olur. Bu bağlamda federe öğrenmenin kullanımı, ideal yaya dedektörleri oluşturmak için karşılaşılan iki büyük sorun için çok önemli bir çözüm potansiyeli taşımaktadır.

Ayrıca çalışmada federe öğrenme kapsamında küresel modelin ağırlıklarını oluşturmak için yerel modellerin kayıp değerlerini dikkate alan yeni bir hesaplama metodolojisi önerilmiştir. Önerilen yöntem yüksek kayıp değerlerine sahip yerel modellere daha yüksek çarpım katsayıları atayarak, küresel modelin, yüksek kayıp değerine sahip ancak daha kapsamlı ve daha zorlu veri kümelerini daha iyi temsil etmesini ve küresel modelin bu kapsamlı veri kümelerinin kullanıldığı yerel modellere daha yakın ağırlıklar elde etmesini sağlar. Bu önerilen yönteme kayba bağlı katsayılarla federe öğrenme (FedLDC) adı verilmektedir. Çalışmada, önerilen FedLDC, FedAvg ve standart eğitim metodolojisi ile karşılaştırılmıştır. Ayrıca test sonuçlarını değerlendirmek için IoU'ya benzer yeni bir yaklaşım kullanılmıştır.

2
FEDERE ÖĞRENME

2017 yılında Google tarafından tanıtılan ve federe öğrenmeye dair ilk yaklaşım olan FedAvg, farklı veri kümelerine sahip kullanıcıların, verilerini paylaşmadan ortak bir ağı eğitmesine olanak sağladı [33]. Bu yaklaşım hem büyük veri oluşturma problemine hem de kişisel verilerin paylaşılması yoluyla oluşabilecek veri gizliliği problemlerine alternatif bir çözüm sundu. Önerilen yöntemin çalışma mantığı şu şekildedir, farklı veri kümelerine sahip farklı kullanıcılar, modellerini belirlenen yerel iterasyon sayısı kadar eğitirler. Bu belirlenen yerel iterasyon sayısının tamamlanmasıyla, her bir kullanıcı elde ettikleri model ağırlıklarını tek bir merkeze gönderirler ve bu merkezde tüm yerel model ağırlıklarının aritmetik ortalaması hesaplanarak küresel model ağırlıkları oluşturulmuş olur. Daha sonra bu elde edilen küresel modelin ağırlıkları her bir kullanıcıya iletilir ve yerel modeller, yerel eğitimlerine bu ağırlıklarla tekrar başlarlar. Elde edilen küresel model ağırlıklarının her bir kullanıcıya iletilmesi ile ilk küresel iterasyon tamamlanmış olur. Bu işlemler belirlenen küresel iterasyon sayısı kadar tekrar edilir ve elde edilen son küresel ağırlıklar eğitilmiş küresel modeli oluşturur. Küresel modelin ağırlıklarını oluşturmak için farklı veri kümeleri üzerinde eğitilmiş birçok yerel modelin tekrarlı eğitimi sonucunda elde edilen ağırlıkların yinelemeli aritmetik ortalamasının alınması ilkesine dayanan bu yöntem, büyük ölçekli verilerin tek merkezde toplanması ihtiyacına alternatif sunması ve özellikle hassas verilerin paylaşılması ihtiyacını ortadan kaldırması nedeniyle oldukça önemlidir. Bu belirtilen avantajları sebebiyle federe öğrenme araştırmacılar için popüler bir çalışma alanı haline gelmiştir ve literatürde federe öğrenme için bir araya getirme (aggregation) yöntemleri, gizlilik ve güvenlik yöntemleri, iletişim verimliliği, sağlamlık ve adalet başlıca olmak üzere birçok alanda çalışmalar yapılmıştır ve halen yapılan çalışmalarla federe öğrenme ile ilgili geliştirmeler devam etmektedir. Yapılan çalışmalar çok farklı yöntemler içerdiği ve önerilen yöntemler belirli durumlarda ve uygulamalarda kullanışlı olduğu için, federe öğrenme ile ilgili çalışmaları farklı başlıklar altında toplamak, literatürdeki çalışmaları sınıflandırmak açısından faydalı olacaktır.

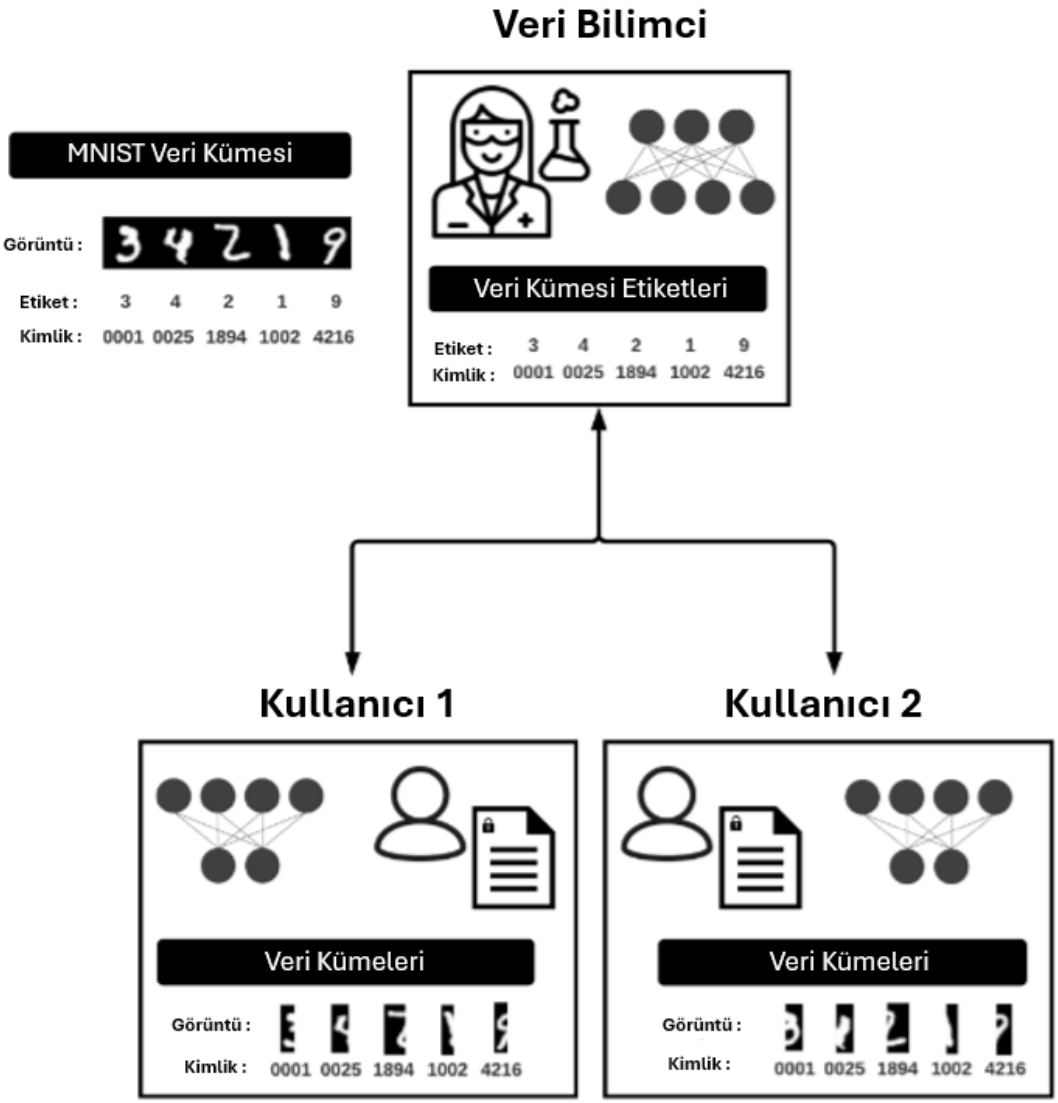
2.1 Dikey Federe Öğrenme

Dikey federe öğrenmede, eğitim sürecine katılan kullanıcılar aynı örnek kümesi için farklı özelliklere sahiptir. Bu senaryo, farklı kullanıcıların veya kuruluşların aynı kişiler veya öğeler hakkında farklı türde bilgiler topladığı ancak hassas verileri paylaşmadan model performansını artırmak için iş birliği yapmak istediği durumlarda tercih edilir.

2020'de Chen ve diğerleri tarafından yapılan çalışma, aynı konu grubu için farklı özellik alt kümelerine sahip birden fazla istemcide makine öğrenimi modellerinin eğitimini geliştirmeyi amaçlayan dikey eşzamanlı olmayan federe öğrenme (VAFL) için bir yöntem tanıtmaktadır [34]. Yöntem, her istemci için küresel bir model ve bireysel yerel yerleştirme modelleri içeren, yüksek boyutlu verilerin işlenmesini kolaylaştıran ve ham verileri kompakt verilerle eşleyerek iletişim yükünü azaltan bir optimizasyon formülasyonu sunar. Bu yöntem, istemcilerin eş zamanlı olmayan güncellemeler yapmasına olanak tanır ve aralıklı bağlantı gibi zorlukların üstesinden gelir. Bu yaklaşım özellikle gerçek zamanlı veya eşzamanlı veri paylaşımının ve işlemesinin pratik olmadığı ortamlar için kullanışlıdır. Model güncellemelerinde veri gizliliğini ve düzgünlüğü sağlamak için, modelin kullanışlılığından önemli ölçüde ödün vermeden bir diferansiyel gizli katman ekleyen bozulmuş yerel yerleştirme tekniği tanıtılmıştır. Makalede, sınırlı ve stokastik sınırsız gecikmeler de dahil olmak üzere çeşitli koşullar altında kapsamlı bir yakınsama analizi sunularak, önerilen yöntemin geleneksel merkezi ve eşzamanlı federe öğrenme yöntemleriyle karşılaştırılabilir bir performans elde ettiği paylaşılmıştır. Eşzamanlı olmayan güncellemelerden ve yerel yerleştirmelerden yararlanarak bu yöntem, bant genişliği ve veri iletim maliyetlerinin önemli olduğu federe öğrenme ortamlarında kritik bir faktör olan iletişim verimliliğini artırır. Karışık yerel yerleştirme kullanımı yoluyla bu yöntem, gizliliğin korunması ile model doğruluğu arasında bir denge sağlayarak, federe öğrenmedeki temel zorluklardan birine çözüm getirir.

2021'de Romanini ve diğerleri tarafından yapılan çalışma, bölünmüş sinir ağlarını (SplitNN) kullanan dikey federe öğrenme için yeni bir çerçeve olan PyDikey (PyVertical)'i tanıtıyor [35]. Şekil 2.1'de uygulama şekli paylaşılan PyVertical, sinir ağlarının birden fazla kullanıcı arasında dikey olarak bölünmüş veriler üzerinde eğitilmesine olanak tanır ve ham verilerin ilgili kullanıcının cihazında kalmasını sağlar. Çalışmada, farklı veri kümelerindeki varlıkları birbirine bağlamak için özel küme kesişimi (PSI) kullanılır. Çerçeve, verileri soyut bir gösterimle eşlemek için SplitNN'leri kullanmıştır ve ham verileri açığa çıkarmadan öğrenme için birden fazla kaynaktan gelen bilgilerin birleşimini sağlamıştır. Çalışmanın temel katkıları

arasında dikey federe öğrenmede SplitNN ve PSI kullanımının genişletilmesi ve PyVertical'ın açık kaynaklı bir çerçeve olarak paylaşılması yer alır.



Şekil 2.1 PyVertical düzeneği [35]

2.2 Yatay Federe Öğrenme

Yatay federe öğrenme, farklı kullanıcıların farklı örneklere ancak bu örnekler için aynı tip özelliklere sahip olduğu durumlar için tercih edilir. Bu federe öğrenme yöntemleri genellikle birden fazla farklı kullanıcının veri kümelerinin benzer türde bilgiler içerdiği durumlarda veri çeşitliliğini artırarak model performansının iyileştirilmesinin hedeflendiği uygulamalarda kullanılır.

2019'da Gao ve diğerleri tarafından yapılan çalışma, elektroensefalografi (EEG) veri analizinde veri heterojenliği ve gizlilik kaygılarının zorluklarını ele almak üzere tasarlanmış, hiyerarşik heterojen yatay federe öğrenme (HHHFL)

çerçevesinden yararlanan elektroensefalografi sınıflandırmasına yeni bir yaklaşım sunmaktadır [36]. Bu çalışmada tanıtılan HHHFL çerçevesi, heterojen EEG verilerinin ortak bir yerleştirme alanına entegrasyonunu sağlayarak bu sınırlamaların üstesinden gelmeyi ve veri gizliliğini sağlarken farklı cihazlar arasında etkili model eğitimine olanak sağlamayı amaçlamaktadır. Bu amaç, farklı veri dağılımları arasındaki tutarsızlığı en aza indirmek için sınıflandırma kaybını alana özgü kayıplarla birleştiren manifold projeksiyonu ve yeni bir kayıp fonksiyonu kullanılarak gerçekleştirilmiştir. Kullanılan metodoloji iki ana bileşenden oluşur, çeşitli kullanıcılardan gelen verileri birleşik bir yerleştirme alanına eşleyen bir manifold projeksiyon mekanizması ve ham verileri paylaşmadan işbirlikçi model eğitimine olanak tanıyan ve böylece gizliliği koruyan bir federe öğrenme süreci. HHHFL'nin hiyerarşik mimarisi, aynı türdeki cihazların bir özellik haritasını iş birliği içinde oluşturmasına olanak tanıyarak, kullanıcıların sağlam bir EEG sinyal sınıflandırıcısı geliştirmesini sağlamış olur.

2022'de Huang ve diğerleri tarafından yapılan çalışmada, yatay federe öğrenmede hem adaleti hem de doğruluğu artırmak için tasarlanmış yeni bir algoritma olan adil ve doğru federe öğrenme algoritması (FedFa) sunulmuştur [37]. Algoritma, çift momentum gradyan yöntemini ve özel bir ağırlıklandırma stratejisini birleştirerek istatistiksel heterojenlik ve iletişim ek yükünden kaynaklanan zorlukları ele almaktadır. Bu yenilikler, yakınsamayı hızlandırmayı ve çeşitli veri dağıtımlarına sahip cihazlar arasında adil model performansı sağlamayı amaçlamaktadır. Çift momentum gradyanı yöntemi hem istemci hem de sunucu düzeyinde geçmiş gradyan bilgilerini birleştirerek optimum değerlere daha hızlı ve daha kararlı yakınsamayı kolaylaştırır. Bu yaklaşım, federe öğrenme ortamlarında yaygın bir sorun olan, cihazlar arasında aynı olmayan veri dağıtımlarının olumsuz etkilerini azaltmaktadır. Ağırlıklandırma stratejisi ise ağırlıkları atamak için eğitim sırasında kullanıcıların hem doğruluğunu hem de katılım sıklığını göz önüne almaktadır. Bu yöntem, verilerin kalitesini ve her kullanıcının katkısının boyutunu dikkate alarak adaleti ön planda tutarak ağ genelinde model performansını dengelemeyi amaçlamaktadır. Ayrıca çalışma, farklı parametre ayarlarının algoritmanın performansı üzerindeki etkilerini araştırmış ve çeşitli senaryolar altında algoritmanın uyarlanabilirliği ve verimliliği hakkında fikirler vermiştir.

2.3 Federe Transfer Öğrenme

Federe transfer öğrenme, federe öğrenme çerçevesi içinde bilginin kaynak alandan hedef alana aktarılmasına izin veren hibrit bir yaklaşımdır. Bu tür yöntemler, özellikle kullanıcının sınırlı etiketli veriye sahip olduğu durumlarda kullanışlıdır.

Federe transfer öğrenmede, veriler ve kullanılan özellikler aynı olmasa bile, kullanıcı eğiteceği modelin performansını artırmak için yüksek miktarda veri içeren önceden eğitilmiş modellerden yararlanabilir.

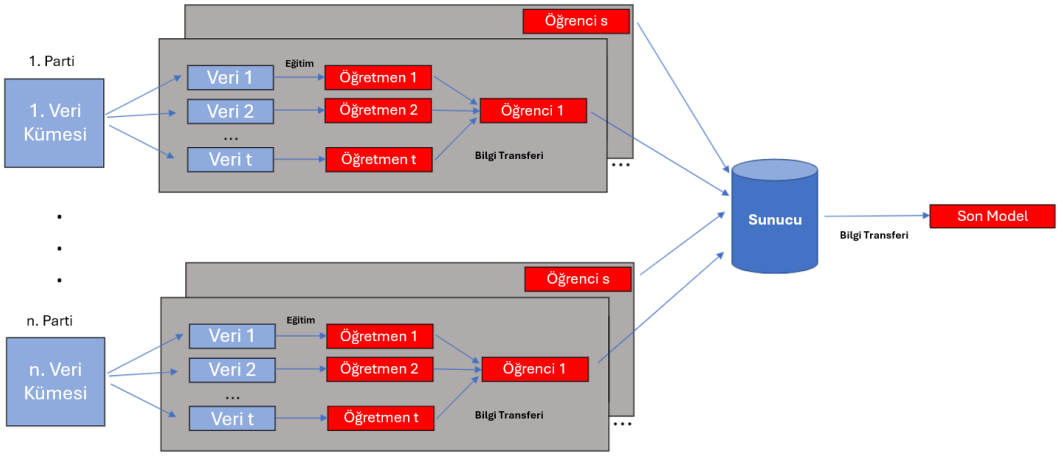
2019'da Sharma ve diğerleri, çalışmalarında, federe transfer öğrenme uygulamalarının, yüksek hesaplama yüküne sahip olmaları sebebiyle kullanışsız olmaları sorunu üzerine odaklanmıştır [38]. Araştırmacılar, gizli paylaşım ve çok taraflı hesaplama ile federe transfer öğrenmenin verimliliğini ve güvenliğini artırmayı önermişlerdir. Modeli sadece yarı dürüst değil, aynı zamanda kötü niyetli katılımcıları da dikkate alacak şekilde geliştirmiş ve böylece daha güçlü bir güvenlik modeli sağlamışlardır. Makalede, iki kullanıcının özel verilerini birbirlerine açıklamadan iş birliği içinde bir modeli eğittiği güvenli bir federe transfer öğrenme modeli ayrıntılı olarak ele alınmıştır. Sistem güvenliği, şifrelenmiş değerler üzerinde hesaplamalar yapmak için gizli paylaşım teknikleri kullanarak gerçekleştirilir ve böylece giriş verilerinin gizliliği sağlanmış olur. Çalışmada önerilen yöntemin, önceki benzer yapılı (homomorfik) şifreleme tabanlı yaklaşımlardan daha verimli olduğu ileri sürülmektedir.

2020'de Liu ve diğerleri tarafından yapılan çalışma, federe transfer öğrenme yoluyla makine öğrenimi modellerini geliştirmeye yönelik yeni bir yaklaşım sunarak, veri gizliliği ve farklı kuruluşlar arasında dağıtım zorluklarına değinmiştir [39]. Çerçeve, etiketli verilerin az olduğu veya kullanılamadığı bir hedef alanda model performansını artırmak için bir kaynak alandan gelen etiketli verilerden yararlanarak farklı alanlar arasında model eğitimine olanak tanımaktadır. Çalışmanın iki temel amacı vardır, federe transfer öğrenmenin, veri gizliliğinin korunduğu bir ortamda gerçekleştirilmesi ve ayrıca şifreli makine öğrenimiyle ilgili çalışmaların araştırılması. Araştırmacılar, kamuya açık veri kümeleri üzerinde gerçekleştirdiği deneyler ile, önerilen çerçevenin geleneksel transfer öğrenme yöntemleriyle karşılaştırılabilir model performansı elde ettiğini paylaşmaktadır.

2.4 Silolar Arası Federe Öğrenme

Silolar arası federe öğrenme genellikle potansiyel olarak büyük miktarda veriye sahip nispeten az sayıda kullanıcının bulunduğu uygulamalardır. Bu kullanıcılar genellikle hastaneler, bankalar veya devlet kurumları gibi kuruluş veya kurumlardır.

Şekil 2.2'de 2020 yılında Li ve diğerleri tarafından önerilen, birden fazla kullanıcının, verilerini paylaşmadan eğitimini mümkün kılmak için tasarlanmış yeni bir tek seferlik federe öğrenme algoritması olan bilgi aktarımı yoluyla federe öğrenmenin (FedKT) genel yapısı sunulmaktadır [40]. Federe öğrenme



Şekil 2.2 FedKT yapısı [40]

için birkaç tur haberleşme gerektiren geleneksel federe öğrenme algoritmalarının aksine, FedKT yalnızca bir tur haberleşme gerektirir. Bu özelliği onu özellikle silolar arası uygulamalar için ideal kılmaktadır. Bu, gizlilik endişeleri veya düzenleyici kısıtlamalar nedeniyle, kullanıcıların ham veriler yerine model güncellemelerini paylaşmak isteyebileceği senaryolarda faydalıdır. FedKT, herhangi bir sınıflandırma modeliyle çalışmasına olanak tanıyan ve sağlam bir gizliliğe ulaşmak için esnek bir çerçeve sağlayan bilgi aktarım tekniklerinden yararlanmaktadır. FedKT'nin omurgası iki katmanlı bir bilgi aktarım mekanizması içermektedir. Başlangıçta, kullanıcıların her biri birden fazla yerel modeli, verilerinin ayrık alt kümeleri üzerinde eğitmektedir ve bu modelleri genel, etiketlenmemiş bir veri kümesini etiketlemek için kullanmaktadır. Bu etiketli örnekler daha sonra bir öğrenci modelini yerel olarak eğitmek için kullanılır. Tüm kullanıcılardan gelen öğrenci modelleri, nihai küresel modeli eğitmek için öğretmen olarak görev yaptıkları merkezi bir sunucuya gönderilir. Bu süreç yalnızca doğrudan veri paylaşımını önleyerek mahremiyeti artırmakla kalmaz, aynı zamanda kamuya açık veri kümesini etiketlemek için kullanılan oylama sayılarına gürültü ekleyerek diferansiyel gizlilik tekniklerinin uygulanmasına olanak tanır.

2021'de Chu ve diğerleri tarafından yapılan çalışma, silolar arası federe öğrenme bağlamında adil modellerin eğitimi için yeni bir çerçeve olan FedAdil (FedFair)'i sunuyor [41]. Çalışmada, kullanıcı verilerinin mahremiyetini ihlal etmeden model adaletini doğru bir şekilde gerçekleştiren federe bir tahmin yöntemi geliştirilmiştir. Federe öğrenmedeki önemli zorluklardan biri olan, her bir veri kümesinin arasında adaletin sağlanması sorununa karşı, FedFair'de eğitim için geliştirilen yeni bir formülizasyon ile adil ve pratik bir çözüm sunulmuştur. Optimizasyon teorisine dayanan önerilen çözüm, federe öğrenme ortamı için özel olarak tasarlanmış alternatif bir gradyan projeksiyon yöntemini kullanıyor. Bu yaklaşım, model

doğruluğu ile adalet arasındaki dengeyi etkili bir şekilde sağlayarak, eğitilen modelin, belirli özelliklere karşı hassas ve dezavantajlı olmasını engellemektedir. Çalışma model adaletini ölçmek için genelleştirilmiş eşit fırsatlar farkı kısıtlamasını detaylandırarak problem formülasyonunun teknik yönlerini ele almaktadır. Bu kısıtlama, adaleti federe öğrenme paradigmasıyla uyumlu bir şekilde ölçerek FedFair çerçevesinde çok önemli bir rol oynar.

2.5 Cihazlar Arası Federe Öğrenme

Cihazlar arası federe öğrenme, paylaşılan modellerin eğitimine katkıda bulunan akıllı telefonlar, nesnelerin interneti (IoT) cihazları veya kişisel bilgisayarlar gibi çok sayıda cihazın bulunduğu uygulamalardır. Her cihaz, modelleri kendi verileri üzerinde yerel olarak eğitir ve ham veriler yerine model güncellemelerini merkezi bir sunucuya gönderir.

2021'de Karimireddy ve diğerleri tarafından yapılan çalışmada, özellikle verilerin çok sayıda cihaza dağıtıldığı uygulamalarda federe öğrenmeyi geliştirmek için tasarlanmış bir çerçeve olan merkezi olmayan stokastik algoritmaları taklit eden mimik federe öğrenme (MIME) tanıtılmaktadır [42]. Federe öğrenmede kullanıcıların sahip oldukları veriler arasındaki veri heterojenliği öğrenme açısından sorun oluşturmaktadır ve bu da istemci sürüklenmesine yol açabilir. MIME çerçevesi, herhangi bir merkezi optimizasyon algoritmasını ek hiper parametrelere ihtiyaç duymadan federe öğrenmeye uyarlayarak bu sorunu giderir ve böylece istemci sapmasını azaltır. Kullanıcılardaki yerel güncellemelerin, merkezi bir algoritmanın güncellemelerini taklit etmesini sağlamak amacıyla kontrol değişkenlerini ve sunucudaki optimizasyon durumlarını birleştirir ve federe öğrenmede iyi bir yakınsama oluşturur. Makalede MIME'nin etkinliğini gösteren ayrıntılı bir teorik analiz sunulmuş ve MIME'nin istemci sürüklenmesini azalttığı ve merkezi optimizasyon yöntemlerini yakından taklit ederek federe öğrenme algoritmalarında iyi bir yakınsama sağladığı paylaşılmıştır. MIME, özellikle sunucu ve istemciler arasında daha az haberleşme turuyla daha hızlı yakınsama ve daha yüksek model performansı elde etmektedir.

2021'de Rehman ve diğerleri tarafından yapılan çalışmada, endüstriyel nesnelerin interneti cihazları (IIoT) arası federe öğrenme sistemlerinde adaleti ve güvenilirliği sağlamayı amaçlayan blok zinciri tabanlı bir çerçeve olan adil ve güvenilir federe öğrenme (TrustFed) sunulmaktadır [43]. TrustFed, merkezi olmayan eğitim ağlarına katılan cihazların hem model sahibi hem de üretici olarak hareket etmesini sağlayarak katılımcılar arasında adaleti ve yüksek kaliteli model kullanılabilirliğini

korumaktadır. Bu, model zehirlenmesi saldırılarıyla mücadelede ve eğitim süreçlerinin bütünlüğünün sağlanmasında çok önemlidir. Çerçeve, cihazlar için güvenilirlik sistemini sürdürmek ve dürüst ve aktif model katkılarını ödüllendirmek için blok zinciri yapısını kullanmaktadır. TrustFed, istatistiksel aykırı değer tespit tekniklerinden ve Python simülasyonlu federe öğrenme çerçevelerinden yararlanarak saldırganları tespit etmekte ve eğitim dağıtımlarından bu saldırganları temizlemektedir.

2.6 Heterojen Federe Öğrenme

Heterojen federe öğrenme çalışmaları, katılımcı cihazların farklı işlem gücüne, depolama kapasitesine ve hatta yazılım ortamlarına sahip olduğu durumları ele alır. Heterojen federe öğrenme algoritmalarının bu tür değişiklikleri etkili bir şekilde ele alacak şekilde uyarlanabilir olması gerekir. Bu nedenle, algoritmaların esneklik ve uyumluluk sağlaması, federe öğrenmenin başarısı için kritik bir öneme sahiptir.

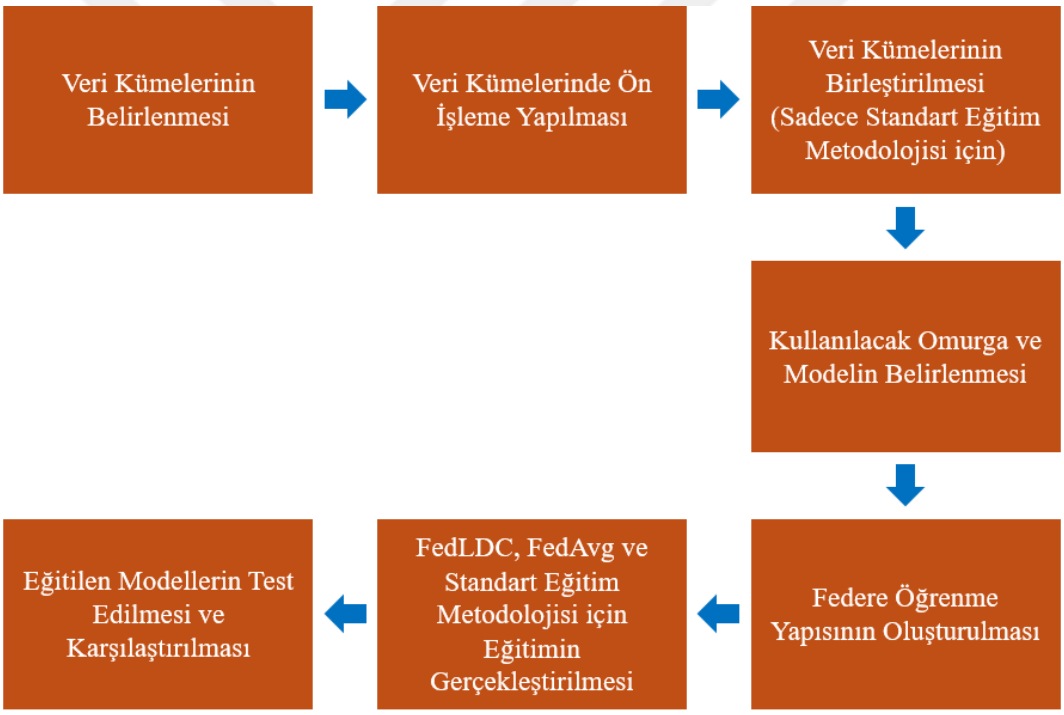
2019'da Ghosh ve diğerleri tarafından yapılan çalışma, federe öğrenmenin zorluklarını ele alıyor ve heterojen veri dağılımları ve Bizans sağlamlığı üzerine odaklanıyor [44]. Araştırmacılar bu zorlukları etkili bir şekilde ele almak için üç aşamalı bir algoritma önermektedir. Çalışmaya veri heterojenliğinden kaynaklanan hesaplama zorlukları ve keyfi veya düşmanca davranışlar sergileyen Bizans makinelerinin varlığı gibi federe öğrenme içindeki sorunları tanımlayarak başlıyorlar. Bunları ele almak için hem kullanıcı kümesi yapılarını hem de Bizans makinelerini hesaba katan ve optimal tahmin hata sınırlarını elde eden istatistiksel bir model sunuyorlar. Yaklaşımlarında ilk aşama, her çalışan düğüm tarafından yerel ampirik risk minimizasyon (ERM) araçlarının hesaplanmasını ve bunların merkezi bir makineye gönderilmesini içermektedir. İkinci aşama, temeldeki veri dağıtımlarını belirlemek ve Bizans makinelerini ayırmak için bu ERM'leri kümeler. Üçüncü aşama daha sonra model parametreleri üzerinde fikir birliğine varmak için her kümeye dağıtılmış sağlam optimizasyon uygulamaktadır.

2022'de Huang ve diğerleri tarafından yapılan çalışma, veri heterojenliği, model heterojenliği ve yıkıcı unutma dahil olmak üzere, federe öğrenmedeki zorlukları ele almak için federe çapraz korelasyon ve sürekli öğrenmeyi (FCCL) tanıtmaktadır [45]. Federe öğrenme, bağımsız ve aynı dağılıma sahip olmayan uygulamalar nedeniyle performans düşüşüne maruz kalmaktadır. FCCL, çapraz korelasyon matrisi oluşturmak için etiketlenmemiş genel verilerden yararlanır, heterojen modeller arasındaki iletişimi kolaylaştırır ve etki alanı değişikliği altında genelleştirilebilir bir temsilin öğrenilmesini sağlar. FCCL, yıkıcı unutmayla

mücadele etmek için yerel güncelleme sırasında gizlilikten ödün vermeden alanlar arası ve alan içi bilgiyi dengelemek amacıyla bilgi damıtmasını kullanır.

3
YÖNTEM

Bu bölümde, federe öğrenmenin yaya tespitindeki potansiyelini araştırmak için çalışmada benimsenen metodolojik çerçeve özetlenmektedir. Şekil 3.1'de çalışmanın gerçekleştirilmesi aşamasında uygulanan adımlar blok diyagram ile paylaşılmaktadır.



Şekil 3.1 Çalışmada gerçekleştirilen adımların blok diyagramı

Çalışmada, önerilen FedLDC yönteminin FedAvg ve standart eğitim metodolojisi ile karşılaştırılacağı daha önceki bölümlerde belirtilmişti. Adil bir karşılaştırma için her üç durum için de ortak bir omurga ve model belirlendi. Bu belirlenen omurga ve model, federe öğrenme yöntemleri için yerel modellerde, standart eğitim metodolojisi için ise ana modelde kullanılmaktadır. Ayrıca yine adil bir karşılaştırma için çalışmada kullanılan üç veri kümesi, federe öğrenme yöntemlerinde her bir veri kümesi bir kullanıcıda olacak şekilde bölünmüştür, standart eğitim metodolojisinde ise bu üç veri kümesi birleştirilerek model

eğitiminde kullanılmıştır. Bu sayede her üç yöntem için de belirlenen üç veri kümesi eğitime dahil edilmiş olur.

3.1 Veri Kümeleri

Yaya tespiti için literatürde birçok veri kümesi sunulmuştur. 2005 yılında Dalal ve Triggs, Ulusal Bilgisayar Bilimi ve Otomasyon Araştırma Enstitüsü Kişi (INRIA Person) veri kümesini tanıttılar. Bu veri kümesi 1800'den fazla açıklamalı insan görüntüsü içerir ve veri kümesindeki her görüntü standart olarak 64x128 pikselden meydana gelmektedir. 2007'de Ess ve diğerleri yoğun yaya trafiğinin olduğu bölgelerde saniyede 15 kare hızda çekilen videolardan oluşan Federal Teknik Yüksekokulu Yaya (ETH Pedestrian) veri kümesini tanıttı. Veri kümesi, 10958 yaya içeren, 640x480 çözünürlükte toplam 2293 video karesinden meydana gelmektedir. 2009 yılında Dollar ve diğerleri Caltech yaya (Caltech Pedestrian) veri kümesini tanıttılar. Bu veri kümesi, Los Angeles bölgesinde LAX, Santa Monica, Hollywood, Pasadena ve Küçük Tokyo gibi mahallelerde bir araçtan kaydedilen 640x480 çözünürlükte yaklaşık 10 saatlik 30Hz bir videonun karelerinden oluşturulmuştur ve yaklaşık 350.000 etiketli yaya ve 250.000 görüntüden oluşmaktadır. 2012 yılında Geiger ve diğerleri, otonom bir sürüş platformu kullanarak, özellikle stereoskopik, optik akış, görsel odometri ve 3B nesne algılama görevleri için tasarlanmış Karlsruhe Teknoloji Enstitüsü ve Toyota Teknoloji Enstitüsü (KITTI) veri kümesini tanıttılar [46]. Veri kümesi, 389 stereo ve optik akış görüntü çiftini, 39,2 km'lik stereoskopik görsel odometri dizilerini ve karmaşık kentsel sahnelerde 200.000'den fazla 3B nesne açıklamasını içerir. Kayıt alınan sistem yüksek çözünürlüklü video kameralar, bir Velodyne lazer tarayıcı ve gelişmiş bir yerelleştirme sistemi ile donatılmıştı. 2014'de Lin ve diğerleri Microsoft bağlamdaki ortak nesneler (MS COCO) veri kümesini tanıtmıştır [47]. Bu veri kümesi, kolayca tanınabilen 91 farklı nesnenin görüntülerinden oluşur ve oldukça kapsamlı bir veri kümesi olan MS COCO 328.000 görüntü ve toplam 2,5 milyon etiketli örnek içermektedir. 2017'de Zhang ve diğerleri yaptıkları çalışmada, Şehir manzaraları (Cityscapes) veri kümesine ek olarak oluşturulan, yaya tespitine yönelik yüksek kaliteli ek açıklamalardan oluşan kapsamlı bir veri kümesi olan ŞehirKişileri (CityPersons) veri kümesini tanıttı [48]. Bu veri kümesi, 27 farklı şehirde, üç farklı mevsimde ve çeşitli hava koşullarında çok çeşitli kentsel senaryoları yansıtan zengin bir veri kümesidir. CityPersons, özellikle tıkanma ve ölçek çeşitliliğine sahip olması sebebiyle zorlu bir veri kümesidir. CityPersons, Caltech ve KITTI gibi önceki veri kümelerine kıyasla daha yüksek görüntü başına kişi sayısına sahiptir. Veri kümesi, görüntü başına ortalama 7 kişi içermektedir ve 2048x1024 çözünürlüklü 5.000 görüntüde 35 bin kişi verisi ve 13

bin görmezden gelinen bölge verisi içerir. Bu yoğunluk ve çeşitlilik, yaya algılama modellerinin daha sağlam bir şekilde eğitilmesine ve test edilmesine olanak sağlar. 2019'da Braun ve diğerleri AvrupaŞehir Kişileri (EuroCity Persons) veri kümesini tanıttılar [49]. Bu veri kümesi, 12 Avrupa ülkesindeki 31 şehirden toplanan çeşitli kentsel ortamlardaki yayaların, bisikletlilerin ve diğer sürücülerin geniş ve çeşitli verilerinden oluşmaktadır. EuroCity Persons 1920x1024 çözünürlükte 47.300'den fazla görüntüde elle etiketlenen 238.200'den fazla kişi verisine sahip olması sebebiyle, benzer amaçlarla kullanılan önceki veri kümelerinden çok daha kapsamlıdır. EuroCity Persons veri kümesinin önemli özelliklerinden biri de toplamda 211.200'ü aşan çok sayıda kişi için yön bilgilerini içermesidir. Bu özellik, yaya davranışının ve yol tahmininin daha iyi anlaşılması için yönelim tahmini gerektiren uygulamalarda özellikle faydalıdır. Veri kümesi, günün farklı zamanları, mevsimler ve hava koşulları da dahil olmak üzere çok çeşitli çevresel koşullarda kaydedilmiştir.

Tablo 3.1 Çalışmada kullanılan veri kümeleri

Veri Kümeleri	Görüntü Sayısı	Kişi Sayısı	Çözünürlük
Caltech Pedestrian	249.884	350.000	640 x 480
CityPersons	5000	35.000	2048 x 1024
EuroCity Persons Day	40.217	183.000	1920 x 1024

Tablo 3.2 Her bir veri kümesinin eğitim, validasyon ve test görüntüsü sayıları

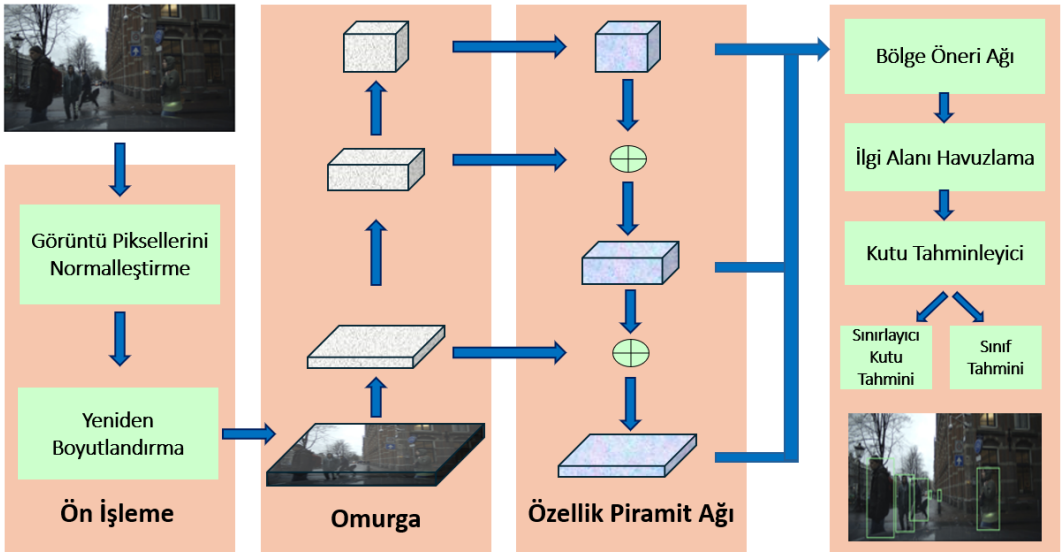
Veri Kümeleri	Eğitim	Validasyon	Test
Caltech Pedestrian	128.419	-	121.465
CityPersons	2975	500	1525
EuroCity Persons Day	23.892	4266	12.059

Çalışmada eğitim metodolojilerini karşılaştırmak için bu başlık altında tanıtılan Caltech Pedestrian, CityPersons ve EuroCity Persons Day veri kümeleri kullanılmıştır. Bu veri kümeleri ile ilgili belirli özellikler Tablo 3.1'de ve her bir veri kümesinin eğitim, validasyon ve test görüntüsü sayıları Tablo 3.2'de paylaşılmıştır. Standart eğitim metodolojisi sırasında Caltech Pedestrian, CityPersons ve EuroCity Persons Day eğitim veri kümeleri bir araya getirilerek, tek bir veri kümesi gibi standart eğitimin gerçekleştirileceği modele eğitim için girdi olarak verilmektedir. Federe öğrenme kullanılan FedAvg ve çalışmada önerilen FedLDC yöntemlerinde, yerel modelleri eğitmek için üç eğitim veri kümesinin tamamı ayrı ayrı farklı kullanıcılar için girdi olarak kullanıldı. Çalışmayı tek sınıflı bir tespit işlemi haline getirmek için, kullanılan veri kümelerindeki insan temsilleri içeren sınıflar tek bir sınıf altında birleştirildi ve bu sınıfa "kişi" (person) adı verildi. Caltech Pedestrian için kişi, kişi-fa, kişi?, kişiler, CityPersons için yaya, sürücü, oturan kişi, kişi grubu, kişi (diğer) ve EuroCity Persons için yaya, kişi grubu-uzak, binici,

binici+araç-grup-uzak sınıfları kişi sınıfı altında birleştirilmiştir. Ayrıca kullanılan veri kümelerinde bu belirtilen sınıflar dışındaki insan temsilleri içermeyen sınıflara ait veriler veri kümelerinden çıkarıldı ve ne eğitim ne de test için kullanıldı.

3.2 Yaya Tespitinde Federe Öğrenme

Çalışmada hem federe öğrenme için hem de standart eğitim metodolojisi için kullanılan modelin blok şeması Şekil 3.2'de görülmektedir. Bu model, MS COCO veri seti üzerinde önceden eğitilmiş bir özellik piramidi ağı kullanan daha hızlı bölgesel evrişimli sinir ağı resnet50 (Faster R-CNN ResNet50 FPN)'dir [50]. Bu model son zamanlarda özellikle gerçek zamanlı uygulamalarda sıklıkla tercih edilmektedir. Hafif bir model olması ve modeldeki özellik piramit ağı (FPN) yapısının küçük nesnelerin tespitindeki başarıyı artırması ve modelin genel tespit yeteneklerinin yüksek olması nedeniyle çalışma için bu model kullanılmıştır. Modelin önceden MS COCO veri kümesinde eğitilmiş ağırlıklarla başlatılması hem federe öğrenme yöntemlerinde hem de standart eğitim metodolojisinde transfer öğrenmeden yararlanarak modelin hızlı bir şekilde yakınsamasını sağlar. Transfer öğrenme, modelin önceden başka bir veri kümesinde eğitilerek edinmiş olduğu bilgi ve özellikleri kullanarak, yeni veri kümeleri üzerinde daha hızlı ve etkili bir şekilde öğrenme gerçekleştirilmesini sağlar. Bu da öğrenme süreçlerinde hız ve başarı açısından kullanıcıya önemli bir avantaj sunar.



Şekil 3.2 Kullanılan modelin genel yapısı

Model, eğitimde kullanılan görüntülerin piksel boyutlarını belirlenen oranlarda normalize eder. Bu görüntüler daha sonra minimum kenar boyutu 800 piksel olacak ve görüntünün en boy oranı korunacak şekilde yeniden boyutlandırılır. Modelin omurgası olan ResNet50, özellik haritalarının genişliğini ve yüksekliğini aşamalı

olarak azaltırken, ardışık evrişim katmanlarıyla giriş kanallarının sayısını aşamalı olarak artırarak giriş görüntülerinden çeşitli özelliklerin yakalanmasına olanak tanır. ResNet yapısı, derin sinir ağlarının eğitiminde ortaya çıkan kaybolan gradyan (vanishing gradient) sorununu çözmek için geliştirilmiş bir yapıdır. ResNet, bu sorunu residual bloklar kullanarak aşar. Residual bloklar, girişten çıkışa bir atlama bağlantısı ekleyerek gradyanların kaybolmasını önler. Bu yapı, ağın daha derin katmanlarda bile etkili bir şekilde öğrenmesini sağlar. FPN, ResNet'in aksine kanal sayısını azaltırken özellik haritalarının genişliğini ve yüksekliğini artırır ve ResNet'in çeşitli katmanlarındaki özellik haritalarını kullanarak farklı boyutlardaki nesneleri algılama başarısını geliştirir. Bölge öneri ağı (RPN), yaya tespiti için özellik piramit ağındaki özellik haritaları üzerinde kayan pencerelerle her bir bölgeyi puanlamak üzere çapa oluşturucu tarafından oluşturulan farklı boyutlardaki bağlantı kutularını kullanır ve bu puanlara dayanarak yaya içerme olasılığı yüksek olan bölgeler önerilir. Son olarak ilgi alanı başlığı bölümünde, bölge öneri ağında önerilen alanların özellik haritaları, bir ilgi alanı havuzu oluşturmak için 7x7 sabit boyuta dönüştürülür ve kutu tahminleyicisi, olası sınırlayıcı kutuları ve sınıfı tahmin etmek için bu ilgi alanı havuzundaki özellik haritalarını kullanır. Burada mevcut çalışmanın doğası gereği tek sınıf içerdiğini hatırlamakta fayda vardır. Çalışmada kullanılan Faster R-CNN ResNet50 FPN modelinin toplam parametre sayısı 41,299,161'dir ve modelin farklı bileşenlerindeki parametre sayıları Tablo 3.3'de özetlenmiştir. Daha önce belirtildiği üzere standart eğitim metodolojisinde Faster R-CNN ResNet50 FPN ana model olarak kullanılmaktadır. Federe öğrenme yöntemini kullanan FedAvg ve FedLDC için ise Faster R-CNN ResNet50 FPN yerel model olarak kullanılmıştır. Hem FedAvg hem FedLDC için tüm kullanıcılar kendilerine ait yerel modellerinde Faster R-CNN ResNet50 FPN kullanmaktadır.

Tablo 3.3 Faster R-CNN ResNet50 FPN toplam parametre sayısı

Bileşen	**Parametre Sayısı**
Omurga (ResNet50)	23,510,081
Özellik Piramit Ağı (FPN)	6,176,512
Bölge Öneri Ağı (RPN)	1,059,984
İlgi Alanı Başlığı (ROI Head)	10,552,584
Toplam	41,299,161

Çalışmada önerilen yöntem FedLDC, kayıp değerlerinden yararlandığı için kullanılan Faster R-CNN ResNet50 FPN modelinin kayıp değerlerini incelemek önemlidir. Kullanılan model Denklem 3.1 RPN sınıflandırma kaybı, Denklem 3.2 RPN regresyon kaybı, Denklem 3.3 R-CNN sınıflandırma kaybı ve Denklem 3.4 R-CNN regresyon kaybını doğrudan toplayarak Denklem 3.5'de belirtildiği gibi kayıp değerini oluşturur. Her bir kullanıcının yerel modellerinde hesaplanan

bu kayıp değerleri FedLDC yönteminde küresel model ağırlıklarını güncellerken kullanılan ana unsurdur.

$$\text{Kayıp}^{\text{RPN}}_{\text{sınıflandırma}} = -\sum_{i} \left[p_i \log(\hat{p_i}) + (1 - p_i) \log(1 - \hat{p_i}) \right] \tag{3.1}$$

$$\text{Kayıp}^{\text{RPN}}_{\text{regresyon}} = \sum_{i} \left[y_i \cdot \text{smooth}_{L1}(t_i - \hat{t_i}) \right] \tag{3.2}$$

$$\text{Kayıp}^{\text{R-CNN}}_{\text{sınıflandırma}} = -\sum_{i} \left[p_i \log(\hat{p_i}) + (1 - p_i) \log(1 - \hat{p_i}) \right] \tag{3.3}$$

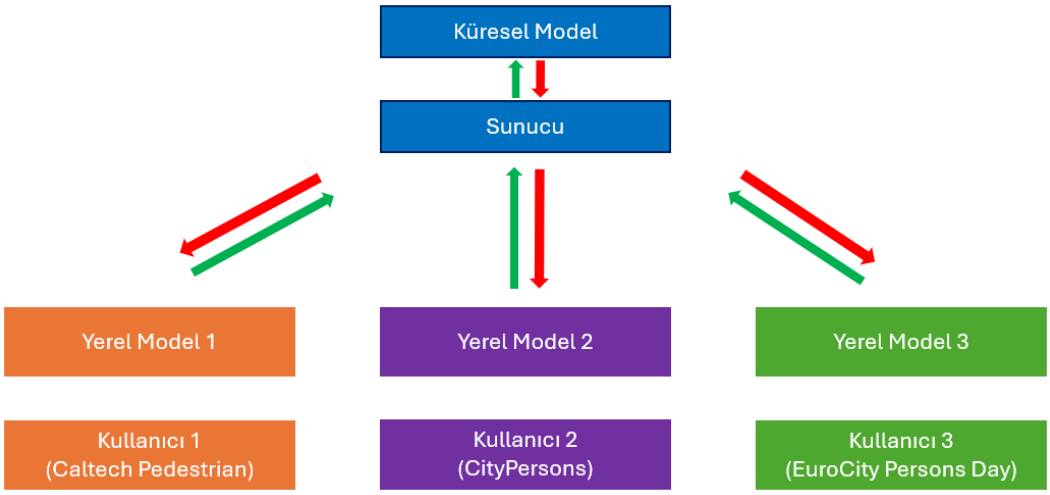
$$\text{Kayıp}^{\text{R-CNN}}_{\text{regresyon}} = \sum_{i} \left[y_i \cdot \text{smooth}_{L1}(t_i - \hat{t_i}) \right] \tag{3.4}$$

$$\text{Kayıp} = \text{Kayıp}^{\text{RPN}}_{\text{sınıflandırma}} + \text{Kayıp}^{\text{RPN}}_{\text{regresyon}} + \text{Kayıp}^{\text{R-CNN}}_{\text{sınıflandırma}} + \text{Kayıp}^{\text{R-CNN}}_{\text{regresyon}} \tag{3.5}$$

Federe öğrenmenin genel işleyişi Şekil 3.3'de paylaşılmıştır. Federe öğrenme için ilk yöntem olan ve halen daha kabul gören FedAvg, farklı kullanıcıların verilerini paylaşmasına gerek kalmadan kendi veri kümeleri üzerinde elde ettikleri yerel ağırlıkları, yinelenen aralıklarla bir sunucuyla paylaşarak küresel bir modelin ağırlıklarını güncellediği ve her bir yerel döngünün sonucunda, küresel modelin ağırlıklarının her bir yerel modele dağıtıldığı ve kullanıcıların yeni yerel döngülerine bu yeni ağırlıklarla başladığı bir eğitim yöntemi izlemektedir. Yöntemde bu eğitim aşamaları önceden belirlenen küresel döngü sayısı kadar tekrarlanır ve elde edilen en son ağırlıklarla küresel model oluşturulmuş olur.

Yapılan çalışma özelinde detaylandırmak gerekirse, standart bir eğitim yöntemi kullanılarak eğitilen bir model performansı, iki farklı küresel model ağırlık güncelleme yöntemi kullanılarak eğitilen federe öğrenme yöntemlerinin performansıyla karşılaştırılmıştır. Federe öğrenme yöntemlerinde kullanılan eğitim düzeneğinin farklı veri kümelerine ancak aynı tür özelliklere sahip büyük boyutlu veriler bulunduran az sayıda kullanıcıdan oluşması, bu uygulamayı yatay federe öğrenme ve silolar arası federe öğrenme kategorisine yerleştirir. Çalışmada kıyaslama için kullanılacak ilk federe öğrenme yöntemi olan FedAvg için, üç farklı veri kümesinin her biri için yerel modeller, yerel dönem sayısına göre eğitilir. Her veri kümesi için ilk yerel eğitim döngüleri tamamlandıktan sonra, küresel modelin ağırlıklarını oluşturmak için her yerel modelin nihai ağırlıklarının aritmetik ortalaması alınır. Bu işlemden sonra ilk küresel döngü tamamlanmış olur. Birinci küresel döngü tamamlandıktan sonra küresel modelin ağırlıkları yerel modellere dağıtılır ve her yerel model, son küresel model ağırlıkları ile yeni yerel döngüler başlatır ve ağırlıklarını kendi veri kümelerinde tekrar günceller. Bu işlem,

belirtilen miktarda küresel döngü boyunca tekrarlanır. Başka bir deyişle, her bir yerel model kendi veri kümelerinde "yerel iterasyon x küresel iterasyon = toplam iterasyon" sayısı kadar ayrı ayrı eğitilir.



Şekil 3.3 Çalışmada kullanılan federe öğrenme yöntemleri için genel blok diyagram

Çalışmada önerilen FedLDC yöntemi temel olarak FedAvg'ye benzemekle birlikte farklı bir global ağırlık güncelleme yöntemi kullanmaktadır. Yöntemde, her yerel eğitim sırasında, farklı veri kümeleri için eğitilen yerel modellerin ortalama kayıp değerleri hesaplanmaktadır. Daha sonra üç farklı veri kümesinin kullanıldığı yerel modeller için hesaplanan bu ortalama kayıp değerleri toplanır. Hesaplamanın matematiksel formülü Denklem 3.6'da verilmiştir.

$$\text{Toplam Kayıp} = \text{Kayıp 1} + \text{Kayıp 2} + \text{Kayıp 3} \tag{3.6}$$

Daha sonra her bir veri kümesinin kaybı, Denklem 3.7'de gösterildiği gibi küresel modelin ağırlıklarını güncellemek için kullanılacak katsayıları oluşturmak için toplam kayba bölünür.

$$\begin{aligned} a &= \frac{\text{Kayıp 1}}{\text{Toplam Kayıp}} \\ b &= \frac{\text{Kayıp 2}}{\text{Toplam Kayıp}} \\ c &= \frac{\text{Kayıp 3}}{\text{Toplam Kayıp}} \end{aligned} \tag{3.7}$$

Bu a, b, c katsayılarının toplamı ağırlık dengesini bozmamak adına bire eşit olacak şekilde normalize edilmiştir. Hesaplamanın matematiksel formülü Denklem 3.8'de verilmiştir. Burada S normalleştirilmemiş katsayıların toplamını tanımlar.

$$\begin{aligned} S &= a + b + c \\ A &= \frac{a}{S} \\ B &= \frac{b}{S} \\ C &= \frac{c}{S} \end{aligned} \tag{3.8}$$

Küresel modelin ağırlıkları güncellenirken her yerel modelin ağırlıkları kendi katsayısı ile çarpılarak küresel modelin ağırlıkları Denklem 3.9'daki gibi güncellenir.

$$W_{\text{küresel}} = A\,x\,W_{\text{yerel 1}} + B\,x\,W_{\text{yerel 2}} + C\,x\,W_{\text{yerel 3}} \tag{3.9}$$

Çalışmada her bir küresel iterasyonda tüm kullanıcılardan elde edilen kayıp değerleri Tablo 3.4'de paylaşılmaktadır. Görüleceği üzere her bir yerel kullanıcının kayıp değerleri her bir küresel iterasyon sonunda azalmaktadır. Bu da her bir küresel ağırlık güncellemesinin sonunda yerel model performanslarının kendi veri kümeleri üzerinde olumlu bir iyileşme sağladığını göstermektedir. Aynı tabloda bu yerel kayıp değerleri doğrultusunda, küresel model ağırlıklarını hesaplamak için yerel modeller için hesaplanan katsayılar görülmektedir. Burada kişi ve arka plan çeşitliliğinin düşük olduğu Caltech Pedestrian veri kümesinde eğitim gerçekleştiren yerel modele daha düşük katsayılar atandığı görülmektedir. Önerilen yöntem, daha yüksek kayıp değerlerine sahip yerel model ağırlıklarına daha büyük katsayılar vererek, küresel model ağırlıklarının, daha benzersiz verilere ancak daha yüksek kayıp değerlerine sahip olan EuroCity Persons Day ve CityPersons veri kümelerinde eğitilen yerel model ağırlıklarına yakınsamasına olanak tanır. Daha özgün verilere sahip bu veri kümelerinde elde edilen ağırlıklar, farklı veri kümeleri üzerinde yapılan testlerde sistemin genelleme performansına olumlu katkı sağlayacaktır.

Literatürde, önerilen FedLDC yöntemine benzer şekilde federe öğrenmede kayıp değerlerini kullanan yaklaşımlar mevcuttur. Veri heterojenliğini ele almak ortak bir zorluktur ve FedProx, SCAFFOLD ve FedMA gibi çalışmalar, müşteri kaymasını

Tablo 3.4 Her bir küresel iterasyonda yerel modeller için hesaplanan kayıp değerleri ve hesaplanan katsayılar

	Yerel Model 1 (Caltech Pedestrian)		Yerel Model 2 (CityPersons)		Yerel Model 3 (ECP Day)	
	Kayıp	Katsayı	Kayıp	Katsayı	Kayıp	Katsayı
Küresel İterasyon 1	0.156	0.137	0.511	0.448	0.474	0.416
Küresel İterasyon 2	0.138	0.135	0.435	0.426	0.447	0.439
Küresel İterasyon 3	0.126	0.129	0.416	0.426	0.434	0.444
Küresel İterasyon 4	0.118	0.125	0.403	0.427	0.423	0.449
Küresel İterasyon 5	0.111	0.121	0.392	0.427	0.415	0.452

azaltmak için çeşitli yöntemler önerir [51–53]. Örneğin, FedProx, istemci ve sunucu ağırlıkları arasındaki farkın L2 normunu düzenleyici terim olarak kullanarak eğitim stabilitesini artırır. SCAFFOLD ise kontrol değişkenleri kullanarak istemci sapmasını azaltır ve daha tutarlı model güncellemeleri sağlar. Kayıp toleranslı federe öğrenme (LT-FL) yaklaşımı, ThrowRightAway adı verilen bir yöntemle bazı paket kayıplarını kasıtlı olarak göz ardı ederek düşük bant genişliğine sahip cihazlar için veri yüklemesini hızlandırır ve bu sayede kişiselleştirme ve adalet performansını garanti eder [54]. Ayrıca, literatürde sınıf dengesizliğini ele alan çalışmalar da mevcuttur; örneğin, Oran kaybı (ratio loss) kullanarak sınıf dengesizliğini azaltmayı amaçlayan yöntemler geliştirilmiştir [55]. Oran kaybı, özellikle küresel dengesizliği azaltmayı hedefler. Çapraz entropi kaybı temelli bu yeni kayıp fonksiyonu, sınıf dengesizliğini gidermek için ağırlık vektörleri kullanır ve bu ağırlıklar, sınıfların eğitim verilerindeki oranlarına göre belirlenir. Minör sınıfların katkısını artırarak dengesizliği giderir ve bu sayede modelin genel performansını artırır. Oran kaybı, yerel eğitim sırasında bir veri örneği ağına beslendiğinde, sınıfın doğru etiketine karşılık gelen düğümdeki gradyanları dikkate alır ve diğer sınıflardan gelen örneklerin aynı düğüm üzerindeki etkilerini de değerlendirir. Bu yöntem, istemcilerin genel örnek sayılarını sunucuya yüklemelerini gerektirmez, böylece istemci gizliliğini korur. FedDyn, öneri sistemlerinde kullanılmak üzere federe distilasyon yöntemi ile veri heterojenliği sorununu ele alır ve model çıktılarını kullanarak iletişim yükünü azaltır [56]. FedDyn, yerel modellerin ürettiği kayıp değerlerinden yararlanarak dinamik bir düzenleyici kullanır. Her turda, cihazlar kendi yerel kayıp fonksiyonlarını optimize ederken, sunucu tarafından sağlanan modelle uyumlu hale gelmeleri için dinamik bir ceza terimi eklenir. Bu ceza terimi, yerel ve küresel kayıp fonksiyonlarının istasyoner noktalarının uyumlu olmasını sağlar ve bu da yerel cihaz modellerinin küresel modelle uyumlu olmasına yardımcı olur. FedLDC, özellikle kayıp-bağımlı katsayılar kullanarak diğer çalışmalardan ayrılır; bu katsayılar, her yerel modelin kayıp değerlerine dayalı olarak küresel modele katkısını belirler. Bu yöntem,

veri setlerinde farklı zorluk seviyeleri ve çeşitliliklerin olduğu durumlarda daha etkili olup, küresel modelin genelleme performansını artırır. Sonuç olarak, FedLDC, kayıp-bağımlı katsayılar kullanarak veri çeşitliliğini ve zorluğunu daha büyük ölçüde etkileyen bir model oluşturur, bu da onu diğer federe öğrenme çalışmalarından ayırır.

4
BULGULAR

Önceki bölümde hem standart eğitim metodolojisi için hem de FedAvg ve FedLDC için Faster R-CNN ResNet50 FPN kullanıldığı belirtilmişti. Çalışmada federe öğrenme yöntemleri için yerel iterasyon sayısı üç, küresel iterasyon sayısı beş olarak belirlenmiştir. Yani her bir yerel model kendi veri kümesinde 3x5=15 iterasyon eğitilir. Bu on beş iterasyon gerçekleştirilirken daha önce bahsedildiği gibi her yerel döngünün sonunda küresel modelin ağırlıkları güncellenir ve yerel modellere dağıtılır yani her üç iterasyonda bir küresel model ağırlıkları güncellenir, bu da tüm eğitim içerisinde beş kez küresel model ağırlıklarının oluşturulduğu anlamına gelir. Çalışmada standart eğitim metodolojisi ise adil bir kıyaslama için doğrudan 15 iterasyon boyunca eğitiliyor. Modellerin eğitiminde, optimizasyon için stokastik gradyan iniş kullanılmıştır ve öğrenme oranı 0.001 olarak ayarlanarak her bir yöntem için eğitim gerçekleştirilmiştir.

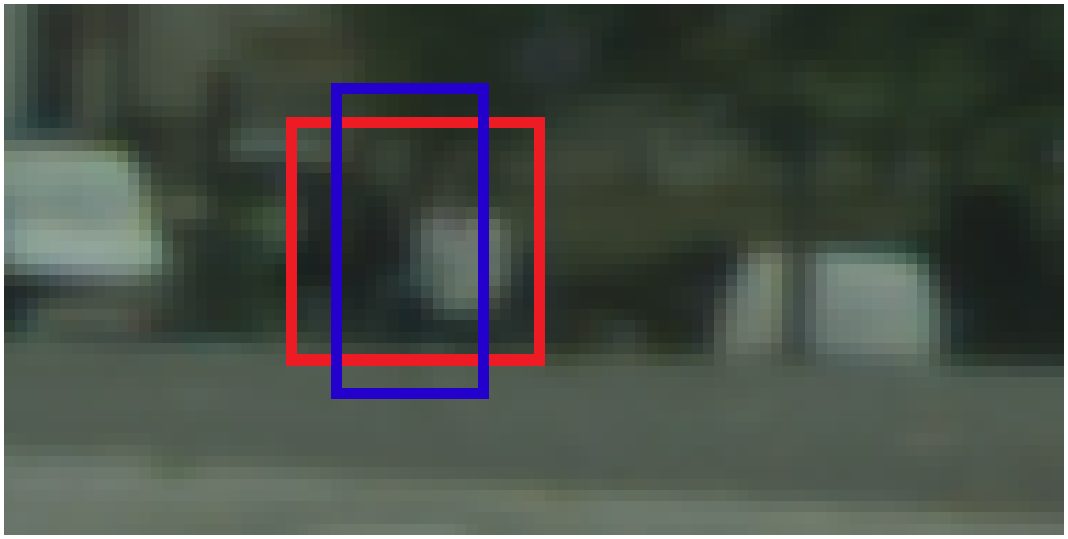
Çalışmada tahmin ile temel gerçek arasındaki kesişme değeri Denklem 4.1'de gösterildiği gibi IoU'ya benzer bir yaklaşım kullanılarak elde edilir.

$$\text{Kesişim Oranı} = \frac{\text{Temel Gerçek ve Tahmin Kutusu Kesişim Alanı}}{\text{Temel Gerçek Alanı}} \tag{4.1}$$

$$\text{IoU} = \frac{\text{Temel Gerçek ve Tahmin Kutusu Kesişim Alanı}}{\text{Temel Gerçek ve Tahmin Kutusu Birleşim Alanı}} \tag{4.2}$$

Önerilen yöntemde Denklem 4.2'de ifade edilen IoU paydasında bulunan "Temel Gerçek ve Tahmin Kutusu Birleşim Alanı" yerine "Temel Gerçek Kutu Alanı" kullanılmaktadır. Çalışmada IoU yerine farklı bir yöntemin kullanılmasının nedeni birden fazla kişi içeren sınıfların (kişiler vb.), "kişi" sınıfı adı altında tek bir sınıfta birleştirilmesidir. Önerilen değerlendirme yöntemi, temel gerçek kutusunun birden fazla kişi içerdiği durumlarda, doğru tahmin edilen temel gerçekleri değerlendirmede IoU'dan daha gerçekçidir. IoU yerine bu yöntemin

kullanılmasının sebebi Şekil 4.1 üzerinden detaylandırılacak olursa, burada içerisinde iki kişinin olduğu oldukça düşük çözünürlüklü bir görsel görülmektedir. Görselde kırmızı kutu temel gerçek kutusunu, mavi ise tahmin kutusunu temsil etmektedir. Burada IoU'nun yüzde kırk yedi civarında olduğu hesaplanıyor ancak bu değer mevcut uygulama için iyi bir değerlendirme sağlamamaktadır. Bu tür sınıfların veri kümesinden doğrudan kaldırılması ise, eğitilen modellerde performansın düşmesine neden olacaktır ve eğitilen modeller tarafından yapılacak yanlış negatiflerin sayısını artıracaktır. Dolayısıyla tahmin kutusunun temel gerçek kutusu dışında kalan kısımlarını paydaya dahil etmeyen bu yeni yaklaşım mevcut uygulama için daha tutarlı bir fikir sunmaktadır.



Şekil 4.1 Temel gerçek ve tahmin kutusu örneği

Çalışmada, modellerin başarısını karşılaştırmak için kaçırma oranı kullanılmıştır ve 0,5 ve 0,75 kesişim oranı eşik değerleri seçilerek hesaplanmıştır. Kaçırma oranı Denklem 4.3'deki gibi modellerin tespit edemediği temel gerçeklerin sayısı, yani yanlış negatifler, tüm temel gerçeklerin sayısına bölünerek hesaplanmaktadır.

$$\text{Kaçırma Oranı} = \frac{\text{Toplam Yanlış Negatif Sayısı}}{\text{Toplam Temel Gerçek Sayısı}} \tag{4.3}$$

Bu çalışmada Caltech Pedestrian, CityPersons ve EuroCity Persons Day eğitim veri kümeleri üzerinde eğitilen üç farklı eğitim yöntemi Caltech Pedestrian test, CityPersons validasyon ve EuroCity Persons Day validasyon verileri üzerinde değerlendirilmiştir. 0,5 kesişim oranı eşiğine ilişkin değerlendirme sonuçları Tablo 4.1'de, 0,75 kesişim oranı eşiğine ilişkin değerlendirme sonuçları ise Tablo 4.2'de sunulmaktadır.

Bu değerlendirmeyi yaparken veri kümeleri bölümünde anlatılan belirli sınıfları

tek bir kişi sınıfı altında birleştirme işlemi değerlendirme için kullanılan verilere de uygulanmış ve insan temsilleri içermeyen ve çalışmayla ilgili olmayan veriler yok sayılmıştır. Modellerin otonom sürüş sistemine yönelik performanslarının daha doğru değerlendirilebilmesi amacıyla veri kümeleri uzakta bulunan ve küçük piksel değerlerine sahip etiketler olmadan yeniden düzenlenmiştir.

Tablo 4.1 Kesişim oranı eşiği 0.5 için test sonucu elde edilen kaçırma oranları

	[a]ST	FedAvg	FedLDC
Caltech Pedestrian Test	0.34	0.16	**0.13**
CityPersons Validasyon	0.09	0.11	**0.07**
EuroCity Persons Day Validasyon	**0.08**	0.15	0.10

[a]Standart eğitim metodolojisi

Tablo 4.2 Kesişim oranı eşiği 0.75 için test sonucu elde edilen kaçırma oranları

	[a]ST	FedAvg	FedLDC
Caltech Pedestrian Test	0.43	0.19	**0.16**
CityPersons Validasyon	0.17	0.24	**0.16**
EuroCity Persons Day Validasyon	**0.18**	0.34	0.24

[a]Standart eğitim metodolojisi

Eğitilen modellerin, bir temel gerçek etrafında birden fazla tahmin yapmasıyla neden olduğu yanlış pozitifleri önlemek için eşik değeri 0,5 olarak ayarlanarak maksimum olmayan bastırma uygulanmıştır. FedLDC yöntemi kullanılarak eğitilen modelin yaptığı tahmin kutuları çizdirilen ve maksimum olmayan bastırma uygulanmış örnek bir görüntü Şekil 4.2'de görülmektedir. Modeller tarafından yapılan tahminlerde, tahmin kutusunun, temel gerçek kutusu ile kesişim alanının, temel gerçek kutusu alanına oranının belirlenen eşik değerinden büyük olduğu tahminler doğru pozitif olarak kabul edilmiş ve kesişim oranı bu değerden düşük olan tahminler ise yanlış pozitif olarak kabul edilmiştir. Bu varsayımlar sonucunda modeller kaçırma oranı kullanılarak birbirleriyle karşılaştırılmıştır.

Öncelikle federe öğrenme için klasik yöntem olan FedAvg, Tablo 4.1 ve Tablo 4.2'den faydalanarak standart eğitim metodolojisi ile karşılaştırılacak olursa, hem 0,5 hem de 0,75 kesişim eşik değeri için Caltech Pedestrian veri kümesinde FedAvg, standart eğitim metodolojisinden çok daha iyi bir sonuç elde etmiştir. Burada FedAvg'nin iyi performansının yanında standart eğitim metodolojisinin beklenenden çok daha kötü sonuç göstermesi de etkilidir. Tüm veri kümelerinin birleştirilerek eğitilmesi düşük çözünürlüklü bir veri kümesi olan ve çok küçük boyutlarda yaya görüntüleri içeren Caltech Pedestrian veri kümesinde kötü sonuçlar elde edilmesine neden olmuştur. CityPersons ve EuroCity Persons Day veri kümelerinde ise hem 0,5 hem de 0,75 kesişim eşik değeri için standart eğitim



Şekil 4.2 FedLDC ile eğitilmiş modele ait maksimum olmayan bastırma uygulanmış örnek tahmin çıktısı

metodolojisi, FedAvg'den daha iyi sonuçlar elde etmiştir, fakat tabloda görüleceği üzere FedAvg CityPersons veri kümesinde ne 0,5 ne de 0,75 eşik değeri için standart eğitim metodolojisinden çok farklı sonuçlar elde etmemiştir. EuroCity Persons Day veri kümesi içinse aradaki fark, Caltech Pedestrian veri kümesindeki kadar büyük değildir. Çalışmada FedAvg ile standart eğitim metodolojisi arasında adil bir kıyas oluşturmak için her iki yöntem de aynı veri kümeleri ile eğitilmiştir. Burada hem FedAvg'nin Caltech Pedestrian veri kümesinde doğrudan daha iyi sonuç aldığının hem de gerçek dünya problemlerinde federe öğrenmenin büyük veri problemlerini çözme noktasında önemli olduğunun altını çizmek lazım. Yani yapılan testlerde her ne kadar veri kümesi bakımından adil testler gerçekleştirilmiş olsa da, gerçek dünya problemlerinde standart eğitim metodolojisi kullanımı durumunda bu tarz veri kümelerini birleştirmek pek mümkün olmayacaktır. Dolayısıyla federe öğrenme yöntemleri büyük verilerin paylaşılmadan bir araya getirilmesine olanak sağlayarak genelleme performansını önemli ölçüde artıracaktır ve büyük ve çeşitli verilerin bir araya getirilemeyeceği standart eğitim metodolojisi için iyi bir çözüm sunacaktır.

Tablo 4.1 ve Tablo 4.2'de görüldüğü gibi önerilen FedLDC yöntemi hem Caltech Pedestrian hem de CityPersons veri kümeleri için 0,5 ve 0,75 eşik değerlerinde en düşük kaçırma oranı değerini elde etti. EuroCity Persons Day veri kümesinde ise önerilen FedLDC modeli, FedAvg'den daha iyi bir kaçırma oranı değeri elde eder ve bu veri kümesinde en iyi değeri elde eden, federe olmayan klasik standart eğitim metodolojisi ile rekabet edebilecek bir başarı yakalar. Bilindiği gibi standart eğitim metodolojisi, federe öğrenmenin sağladığı veri gizliliği özelliğinden yoksundur. Önerilen FedLDC yönteminin hem çoğu değerlendirmede en iyi sonucu elde etmesi

hem tüm veri kümelerinde FedAvg'den daha iyi performans göstermesi hem de EuroCity Persons Day veri kümesi için standart eğitim metodolojisiyle rekabet edebilecek bir sonuç elde etmiş olması, hem federe öğrenmenin, yaya tespiti gibi veri gizliliği ve büyük veri ihtiyacının gerektiği uygulamalarda iyi bir çözüm önerisi olduğunu hem de önerilen FedLDC yönteminin federe öğrenme için iyi bir ağırlıklandırma yöntemi olduğunu ve eğitimde kullanılan veri kümelerinin iyi bir temsilini yakaladığını göstermektedir. Burada FedLDC'nin yakaladığı başarının, FedAvg'den farklı olduğunun altını çizmek lazım, FedAvg'nin performansının gerçek dünya problemlerinde büyük verileri bir araya getirmenin mümkün olmadığı durumlarda iyi bir alternatif olduğu söylenmişti, ancak FedLDC doğrudan adil bir test ortamında standart eğitim metodolojisinden daha iyi sonuçlar elde etmiştir, yani daha fazla veriye erişme avantajı olmadan bile standart eğitim metodolojisinden daha iyi sonuçlar ortaya koymuştur. Ayrıca tekrarlamak gerekirse FedLDC federe öğrenme için klasik yöntem olan FedAvg'den ise her altı değerlendirme için daha iyi sonuçlar elde etmiştir.

Tablo 4.3 Literatürdeki çalışmaların Caltech veri kümesi için kaçırma oranları

Çalışma	**Yöntem**	**IoU Eşik Değeri**	**Kaçırma Oranı**
Learning efficient single-stage pedestrian detectors by asymptotic localization fitting [25]	ALFNet	0.5	0.061
High-level semantic feature detection: A new perspective for pedestrian detection [28]	CSP	0.5	0.05
Cascade R-CNN: High quality object detection and instance segmentation [57]	Cascade R-CNN	0.5	0.062
Localized Semantic Feature Mixers for Efficient Pedestrian Detection in Autonomous Driving [58]	LSFM	0.5	0.031
Generalizable pedestrian detection: The elephant in the room [59]	Pedestron	0.5	0.062

Tablo 4.3, Tablo 4.4 ve Tablo 4.5'de Caltech Pedestrian, CityPersons ve EuroCity Persons veri kümelerinde literatürdeki çalışmaların elde ettiği test

sonuçları paylaşılmaktadır. Yapılan uygulama gereği veri kümelerinde çeşitli konfigürasyonlar yapılması sebebiyle bu değerleri doğrudan kıyaslamak sağlıklı olmasada bu literatür çalışmalarında paylaşılan sonuçlar, mevcut uygulamada elde edilen başarımları değerlendirmek açısından yine de değerlidir. Hatırlanacağı üzere mevcut çalışmada insan temsilleri içeren sınıflar tek bir "person" adı verilen sınıfta birleştirilmiş, kişi temsili içermeyen sınıflar veri kümelerinde eğitime ve teste dahil edilmemiş ve bu konfigürasyonlar sebebiyle standart IoU yaklaşımı yerine yeni bir kesişim oranı formülü kullanılmıştı. Bu tarz konfigürasyonlar kişi tespiti odaklı olan ve federe öğrenmenin, standart eğitimle kıyaslanmasını amaçlayan bu çalışmayı, standart uygulamalardan ayrıştırmaktadır.

Tablo 4.4 Literatürdeki çalışmaların CityPersons veri kümesi için elde ettikleri kaçırma oranları

Çalışma	**Yöntem**	**IoU Eşik Değeri**	**Kaçırma Oranı**
Learning efficient single-stage pedestrian detectors by asymptotic localization fitting [25]	ALFNet	0.5	0.12
High-level semantic feature detection: A new perspective for pedestrian detection [28]	CSP	0.5	0.11
Cascade R-CNN: High quality object detection and instance segmentation [57]	Cascade R-CNN	0.5	0.112
Localized Semantic Feature Mixers for Efficient Pedestrian Detection in Autonomous Driving [58]	LSFM	0.5	0.085
Generalizable pedestrian detection: The elephant in the room [59]	Pedestron	0.5	0.112

Tablo 4.5 Literatürdeki çalışmaların EuroCity Persons veri kümesi için elde ettikleri kaçırma oranları

Çalışma	**Yöntem**	**IoU Eşik Değeri**	**Kaçırma Oranı**
Eurocity persons: A novel benchmark for person detection in traffic scenes [49]	SSD	0.5	0.105
Faster r-cnn: Towards real-time object detection with region proposal networks [50]	Faster R-CNN	0.5	0.073
Cascade R-CNN: High quality object detection and instance segmentation [57]	Cascade R-CNN	0.5	0.066
Localized Semantic Feature Mixers for Efficient Pedestrian Detection in Autonomous Driving [58]	LSFM	0.5	0.047
Generalizable pedestrian detection: The elephant in the room [59]	Pedestron	0.5	0.066

5
SONUÇ

Yapılan çalışmada federe öğrenmenin yaya tespiti gibi yüksek veri çeşitliliği ve veri gizliliği gerektiren uygulamalar için potansiyeli incelendi ve önerilen FedLDC yöntemi, Caltech Pedestrian, CityPersons ve EuroCity Persons Day veri kümeleri kullanılarak hem FedAvg hem de standart eğitim metodolojisi ile karşılaştırıldı. FedLDC yöntemi tüm testlerde FedAvg yönteminden daha iyi performans göstererek daha düşük kaçırma oranı elde etti. Önerilen yöntem, gerçekleştirilen altı testten dördünde en iyi sonuçları elde etti ve yalnızca EuroCity Persons Day veri kümesindeki testlerde standart eğitim metodolojisinin ardından ikinci sırada yer aldı. Standart eğitim metodolojisinde tüm veri kümelerinin doğrudan birleştirildiği, yani verilerin tek merkezde toplandığı dikkate alındığında, veri paylaşımı gerektirmeyen FedLDC yöntemi çoğu durumda hem en başarılı sonuçları elde etmesi hem de diğer testlerde standart eğitim metodolojisine yakın sonuçlar elde etmesi dikkat çekicidir. Doğrudan ortalama almak yerine, eğitim sırasında elde edilen kayıp değerlerine göre yerel modellere katsayılar atayan bu yöntem, daha yüksek kayıp değerlerinin elde edildiği ancak daha benzersiz verilere sahip veri kümeleri üzerinde eğitilen yerel modellere daha yüksek katsayılar atayarak, sistemin bu çeşitliliği yüksek veri kümelerine yakınsamasını sağlayarak, elde edilen küresel modelin genelleme performansının iyileştirilmesine yardımcı oldu.

Gerçekleştirilen çalışma, küresel model ağırlıklarının, kullanılan veri kümelerini iyi bir şekilde temsil etmesini sağlamak için başarılı bir yardımcı katsayı hesaplama yöntemi sunmaktadır. Ayrıca çalışmada federe öğrenmenin rekabetçi ve hatta standart eğitim metodolojisinden daha iyi sonuçlar elde edebileceğini göstermek de federe öğrenmenin geleceği açısından önemlidir. Önerilen yöntem, hem federe öğrenmeye yenilikçi bir yaklaşım sunmakta hem de federe öğrenmenin atası olan FedAvg'nin performansında dikkate değer bir iyileşme sağlamaktadır. Önerilen yöntem yaya tespiti dışındaki alanlarda gelecekteki çalışmalar için umut vericidir ve federe öğrenmenin geleceği için bir basamak görevi görmektedir. Gelecekteki

çalışmalarda, önerilen yönteme göre daha düşük kayıp oranı elde eden bir model geliştirmek için yeni bir araya getirme yöntemleri üzerine araştırmalar yapılacaktır ve elde edilen başarı geliştirilmeye çalışılacaktır.

KAYNAKÇA

[1] J. Tyndall, "Pedestrian deaths and large vehicles," *Economics of Transportation*, vol. 26, p. 100 219, 2021.

[2] C. V. Zegeer, M. Bushell, "Pedestrian crash trends and potential countermeasures from around the world," *Accident Analysis & Prevention*, vol. 44, no. 1, pp. 3–11, 2012.

[3] A. Ali, E. Dagless, "Vehicle and pedestrian detection and tracking," in *IEE Colloquium on Image Analysis for Transport Applications*, IET, 1990, pp. 5–1.

[4] Y. Y. Tang, Y. J. Lu, C. Y. Suen, "Robot vision system for pedestrian-flow detection," in *Sensor Fusion IV: Control Paradigms and Data Structures*, SPIE, vol. 1611, 1992, pp. 630–641.

[5] H. Mori, N. M. Charkari, "Shadow and rhythm as sign patterns of obstacle detection," in *ISIE'93-Budapest: IEEE International Symposium on Industrial Electronics Conference Proceedings*, IEEE, 1993, pp. 271–277.

[6] S. Yasutomi, H. Mori, "A method for discriminating of pedestrian based on rhythm," in *Proceedings of IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS'94)*, IEEE, vol. 2, 1994, pp. 988–995.

[7] N. Kehtarnavaz, F. Rajkotwala, "Real-time vision-based detection of waiting pedestrians," *Real-Time Imaging*, vol. 3, no. 6, pp. 433–440, 1997.

[8] C. P. Papageorgiou, M. Oren, T. Poggio, "A general framework for object detection," in *Sixth international conference on computer vision (IEEE Cat. No. 98CH36271)*, IEEE, 1998, pp. 555–562.

[9] L. Zhao, C. E. Thorpe, "Stereo-and neural network-based pedestrian detection," *IEEE Transactions on intelligent transportation systems*, vol. 1, no. 3, pp. 148–154, 2000.

[10] D. M. Gavrila, J. Giebel, "Shape-based pedestrian detection and tracking," in *Intelligent Vehicle Symposium, 2002. IEEE*, IEEE, vol. 1, 2002, pp. 8–14.

[11] Viola, Snow, "Detecting pedestrians using patterns of motion and appearance," in *Proceedings ninth IEEE international conference on computer vision*, IEEE, 2003, pp. 734–741.

[12] H. Cheng, N. Zheng, J. Qin, "Pedestrian detection using sparse gabor filter and support vector machine," in *IEEE Proceedings. Intelligent Vehicles Symposium, 2005.*, IEEE, 2005, pp. 583–587.

[13] O. Tuzel, F. Porikli, P. Meer, "Pedestrian detection via classification on riemannian manifolds," *IEEE transactions on pattern analysis and machine intelligence*, vol. 30, no. 10, pp. 1713–1727, 2008.

[14] N. Dalal, B. Triggs, "Histograms of oriented gradients for human detection," in *2005 IEEE computer society conference on computer vision and pattern recognition (CVPR'05)*, Ieee, vol. 1, 2005, pp. 886–893.

[15] S. Munder, D. M. Gavrila, "An experimental study on pedestrian classification," *IEEE transactions on pattern analysis and machine intelligence*, vol. 28, no. 11, pp. 1863–1868, 2006.

[16] S. Walk, N. Majer, K. Schindler, B. Schiele, "New features and insights for pedestrian detection," in *2010 IEEE Computer society conference on computer vision and pattern recognition*, IEEE, 2010, pp. 1030–1037.

[17] D. Vazquez, A. M. Lopez, J. Marin, D. Ponsa, D. Geronimo, "Virtual and real world adaptation for pedestrian detection," *IEEE transactions on pattern analysis and machine intelligence*, vol. 36, no. 4, pp. 797–809, 2013.

[18] M. Szarvas, A. Yoshizawa, M. Yamamoto, J. Ogata, "Pedestrian detection with convolutional neural networks," in *IEEE Proceedings. Intelligent Vehicles Symposium, 2005.*, IEEE, 2005, pp. 224–229.

[19] M. Norouzi, M. Ranjbar, G. Mori, "Stacks of convolutional restricted boltzmann machines for shift-invariant feature learning," in *2009 IEEE Conference on Computer Vision and Pattern Recognition*, IEEE, 2009, pp. 2735–2742.

[20] W. Ouyang, X. Wang, "Joint deep learning for pedestrian detection," in *Proceedings of the IEEE international conference on computer vision*, 2013, pp. 2056–2063.

[21] A. Angelova, A. Krizhevsky, V. Vanhoucke, A. S. Ogale, D. Ferguson, "Real-time pedestrian detection with deep network cascades.," in *Bmvc*, vol. 2, 2015, p. 4.

[22] J. Deng, W. Dong, R. Socher, L.-J. Li, K. Li, L. Fei-Fei, "Imagenet: A large-scale hierarchical image database," in *2009 IEEE conference on computer vision and pattern recognition*, Ieee, 2009, pp. 248–255.

[23] J. Dai, Y. Li, K. He, J. Sun, "R-fcn: Object detection via region-based fully convolutional networks," *Advances in neural information processing systems*, vol. 29, 2016.

[24] W. Lan, J. Dang, Y. Wang, S. Wang, "Pedestrian detection based on yolo network model," in *2018 IEEE international conference on mechatronics and automation (ICMA)*, IEEE, 2018, pp. 1547–1551.

[25] W. Liu, S. Liao, W. Hu, X. Liang, X. Chen, "Learning efficient single-stage pedestrian detectors by asymptotic localization fitting," in *Proceedings of the European Conference on Computer Vision (ECCV)*, 2018, pp. 618–634.

[26] S. Zhang, R. Benenson, B. Schiele, "Citypersons: A diverse dataset for pedestrian detection," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2017, pp. 3213–3221.

[27] P. Dollár, C. Wojek, B. Schiele, P. Perona, "Pedestrian detection: A benchmark," in *2009 IEEE conference on computer vision and pattern recognition*, IEEE, 2009, pp. 304–311.

[28] W. Liu, S. Liao, W. Ren, W. Hu, Y. Yu, "High-level semantic feature detection: A new perspective for pedestrian detection," in *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, 2019, pp. 5187–5196.

[29] P. Tumas, A. Nowosielski, A. Serackis, "Pedestrian detection in severe weather conditions," *Ieee Access*, vol. 8, pp. 62 775–62 784, 2020.

[30] M. Wang, H. Ma, S. Liu, Z. Yang, "A novel small-scale pedestrian detection method base on residual block group of centernet," *Computer Standards & Interfaces*, vol. 84, p. 103 702, 2023.

[31] A. Ess, B. Leibe, L. Van Gool, "Depth and appearance for mobile scene analysis," in *2007 IEEE 11th international conference on computer vision*, IEEE, 2007, pp. 1–8.

[32] Q. Liu, H. Ye, S. Wang, Z. Xu, "Yolov8-cb: Dense pedestrian detection algorithm based on in-vehicle camera," *Electronics*, vol. 13, no. 1, p. 236, 2024.

[33] B. McMahan, E. Moore, D. Ramage, S. Hampson, B. A. y Arcas, "Communication-efficient learning of deep networks from decentralized data," in *Artificial intelligence and statistics*, PMLR, 2017, pp. 1273–1282.

[34] T. Chen, X. Jin, Y. Sun, W. Yin, "Vafl: A method of vertical asynchronous federated learning," *arXiv preprint arXiv:2007.06081*, 2020.

[35] D. Romanini *et al.*, "Pyvertical: A vertical federated learning framework for multi-headed splitnn," *arXiv preprint arXiv:2104.00489*, 2021.

[36] D. Gao, C. Ju, X. Wei, Y. Liu, T. Chen, Q. Yang, "Hhhfl: Hierarchical heterogeneous horizontal federated learning for electroencephalography," *arXiv preprint arXiv:1909.05784*, 2019.

[37] W. Huang, T. Li, D. Wang, S. Du, J. Zhang, T. Huang, "Fairness and accuracy in horizontal federated learning," *Information Sciences*, vol. 589, pp. 170–185, 2022.

[38] S. Sharma, C. Xing, Y. Liu, Y. Kang, "Secure and efficient federated transfer learning," in *2019 IEEE international conference on big data (Big Data)*, IEEE, 2019, pp. 2569–2576.

[39] Y. Liu, Y. Kang, C. Xing, T. Chen, Q. Yang, "A secure federated transfer learning framework," *IEEE Intelligent Systems*, vol. 35, no. 4, pp. 70–82, 2020.

[40] Q. Li, B. He, D. Song, "Practical one-shot federated learning for cross-silo setting," *arXiv preprint arXiv:2010.01017*, 2020.

[41] L. Chu, L. Wang, Y. Dong, J. Pei, Z. Zhou, Y. Zhang, "Fedfair: Training fair models in cross-silo federated learning," *arXiv preprint arXiv:2109.05662*, 2021.

[42] S. P. Karimireddy *et al.*, "Breaking the centralized barrier for cross-device federated learning," *Advances in Neural Information Processing Systems*, vol. 34, pp. 28 663–28 676, 2021.

[43] M. H. ur Rehman, A. M. Dirir, K. Salah, E. Damiani, D. Svetinovic, "Trustfed: A framework for fair and trustworthy cross-device federated learning in iiot," *IEEE Transactions on Industrial Informatics*, vol. 17, no. 12, pp. 8485–8494, 2021.

[44] A. Ghosh, J. Hong, D. Yin, K. Ramchandran, "Robust federated learning in a heterogeneous environment," *arXiv preprint arXiv:1906.06629*, 2019.

[45] W. Huang, M. Ye, B. Du, "Learn from others and be yourself in heterogeneous federated learning," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2022, pp. 10 143–10 153.

[46] A. Geiger, P. Lenz, R. Urtasun, "Are we ready for autonomous driving? the kitti vision benchmark suite," in *2012 IEEE conference on computer vision and pattern recognition*, IEEE, 2012, pp. 3354–3361.

[47] T.-Y. Lin *et al.*, "Microsoft coco: Common objects in context," in *Computer Vision–ECCV 2014: 13th European Conference, Zurich, Switzerland, September 6-12, 2014, Proceedings, Part V 13*, Springer, 2014, pp. 740–755.

[48] M. Cordts *et al.*, "The cityscapes dataset for semantic urban scene understanding," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2016, pp. 3213–3223.

[49] M. Braun, S. Krebs, F. Flohr, D. M. Gavrila, "Eurocity persons: A novel benchmark for person detection in traffic scenes," *IEEE transactions on pattern analysis and machine intelligence*, vol. 41, no. 8, pp. 1844–1861, 2019.

[50] S. Ren, K. He, R. Girshick, J. Sun, "Faster r-cnn: Towards real-time object detection with region proposal networks," *Advances in neural information processing systems*, vol. 28, 2015.

[51] T. Li, A. K. Sahu, M. Zaheer, M. Sanjabi, A. Talwalkar, V. Smith, "Federated optimization in heterogeneous networks," *Proceedings of Machine learning and systems*, vol. 2, pp. 429–450, 2020.

[52] S. P. Karimireddy, S. Kale, M. Mohri, S. Reddi, S. Stich, A. T. Suresh, "Scaffold: Stochastic controlled averaging for federated learning," in *International conference on machine learning*, PMLR, 2020, pp. 5132–5143.

[53] H. Wang, M. Yurochkin, Y. Sun, D. Papailiopoulos, Y. Khazaeni, "Federated learning with matched averaging," *arXiv preprint arXiv:2002.06440*, 2020.

[54] P. Zhou, P. Fang, P. Hui, "Loss tolerant federated learning," *arXiv preprint arXiv:2105.03591*, 2021.

[55] L. Wang, S. Xu, X. Wang, Q. Zhu, "Addressing class imbalance in federated learning," in *Proceedings of the AAAI Conference on Artificial Intelligence*, vol. 35, 2021, pp. 10 165–10 173.

[56] D. A. E. Acar, Y. Zhao, R. M. Navarro, M. Mattina, P. N. Whatmough, V. Saligrama, "Federated learning based on dynamic regularization," *arXiv preprint arXiv:2111.04263*, 2021.

[57] Z. Cai, N. Vasconcelos, "Cascade r-cnn: High quality object detection and instance segmentation," *IEEE transactions on pattern analysis and machine intelligence*, vol. 43, no. 5, pp. 1483–1498, 2019.

[58] A. H. Khan, M. S. Nawaz, A. Dengel, “Localized semantic feature mixers for efficient pedestrian detection in autonomous driving,” in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2023, pp. 5476–5485.

[59] I. Hasan, S. Liao, J. Li, S. U. Akram, L. Shao, “Generalizable pedestrian detection: The elephant in the room,” in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2021, pp. 11 328–11 337.

TEZDEN ÜRETİLMİŞ YAYINLAR

Konferans Bildirisi

1. E. Akpinar, B. Bolat and M. Taskiran, "FedLDC: Federated Learning with Loss-Dependent Coefficients in Pedestrian Detection," 2024 11th International Conference on Signal Processing and Integrated Networks (SPIN), Noida, India, 2024, pp. 256-261.