

**GÖZETİM VİDEOLARINDA ANOMALİ TESPİT
YÖNTEMLERİNİN KARŞILAŞTIRMALI BİR
DEĞERLENDİRMESİ**

**A COMPARATIVE EVALUATION OF ANOMALY
DETECTION METHODS FOR SURVEILLANCE VIDEOS**

CEM ÇETİNKAYA

PROF. DR. M. ÖZGÜR YENİAY

Tez Danışmanı

Hacettepe Üniversitesi

Lisansüstü Eğitim-Öğretim ve Sınav Yönetmeliğinin

İstatistik Anabilim Dalı için Öngördüğü

YÜKSEK LİSANS TEZİ olarak hazırlanmıştır.

*Sevgili annem Fatma Alp
ve kardeşim Özgür Çetinkaya 'ya...*

ÖZET

GÖZETİM VİDEOLARINDA ANOMALİ TESPİT YÖNTEMLERİNİN KARŞILAŞTIRMALI BİR DEĞERLENDİRMESİ

Cem ÇETİNKAYA

Yüksek Lisans, İstatistik Bölümü

Tez Danışmanı: Prof. Dr. M. Özgür YENİAY

Mayıs 2022, 60 sayfa

Kameralar ile oluşturulan veriler çeşitli sebepler ile analize ihtiyaç duymaktadır. Bunlardan en önemlisi güvenlik sebebidir. Videolarda gerçekleşen istenmeyen ve nadir durumda olan olaylar anomali olarak nitelendirilir. Videolarda anomalilerin tespiti için son yıllarda oldukça fazla yöntem ve yaklaşım önerilmiştir. Çalışmaların çoğu doğruluk bazında yapılmış olsa da, uygulanabilirlik de önem arz etmektedir. Karmaşık mimariler ile mükemmel yakın doğrulukta performans elde etmek mümkün iken, ilgili mimarilerin karmaşıklığı arttıkça bellek ve işlem gücü ihtiyacı da artmaktadır. Bu durum gerçek hayatta uygulanabilirliğini azaltmaktadır.

Bu çalışmanın amacı, görece olarak daha basit, literatürde sıkça kullanılan otokodlayıcı tabanlı mimariler ile iki yaklaşımın incelenmesidir. Bu yaklaşımlar ham veri yaklaşımı ve optik akış yaklaşımıdır. Ham veri yaklaşımında görüntü verileri direkt olarak eğitim ağına alınarak, görüntülerin uzay-zamansal öznitelikleri ile anomali tespiti yapılmaktadır. Optik akış yaklaşımında ise iki video karesi arasındaki değişimden optik akış özneliği hesaplanarak mimariye girdi olarak alınmaktadır. Böylece hareket bazlı anomali tespiti yapılmaktadır.

Çalışma üç farklı veri setinde ve üç farklı mimari ile test edilen iki yaklaşımdan hangisinin daha iyi olduğu sorusuna cevap aramaktadır.

Anahtar Kelimeler: Otokodlayıcı mimarileri, videolarda anomali tespitleri, gözetimsiz öğrenme, optik akış çıkarımı, bilgisayarlı görü

ABSTRACT

A COMPARATIVE EVALUATION OF ANOMALY DETECTION METHODS FOR SURVEILLANCE VIDEOS

Cem ÇETİNKAYA

Master of Science, Department of Statistics

Supervisor: Prof. Dr. M. Özgür YENİAY

May 2022, 60 pages

The data generated by the camera needs to be analyzed for various reasons. The most important of these is for security reasons. Undesirable and rare events that occurs in videos are considered an anomaly. In recent years, many methods and approaches have been proposed for the detection of anomalies in videos. Although most of these studies have been done on the basis of accuracy, applicability is also important. While it is possible to achieve near-perfect accuracy with complex architectures, the need for memory and processing power increases as the complexity of the related architectures increases. This situation reduces its applicability in real life.

Main goal of this study is to review the relatively simple autoencoder-based architectures and two approaches that are frequently used in the literature. These approaches are the raw data approach and the optical flow approach. In the raw data approach, video image data directly imported into the training network and anomaly detection is performed with the spatio-temporal features of the images. In the optical flow approach, the optical flow attribute is calculated from the change between two video frames and taken as an input to the architecture. Thus, motion-based anomaly detection is performed.

This study seeks to answer the question of which of the two approaches is better, tested on three different datasets and with three different architectures.

Keywords: Autoencoder architectures, video anomaly detection, unsupervised learning, optical flow extraction, computer vision

TEŐEKKÜR

Danışmanım ve deęerli hocam Prof. Dr. M. Özgür Yeniay'a, eğitim sürecim boyunca her konuda bana verdiği destekleri, bana olan inancı ve güveni için,

Hacettepe Üniversitesi İstatistik bölümünde bulunan deęerli mesai arkadaşlarım ve kıymetli hocalarıma bana verdikleri destekleri için,

Beni bu günlere getiren en deęerlim, annem Fatma Alp'e ve neşe kaynađım olan kardeşim Özgür Çetinkaya'ya tüm kalbimle teşekkürlerimi ve minnettarlıklarımı sunarım.

İÇİNDEKİLER

ÖZET	i
ABSTRACT.....	ii
TEŞEKKÜR.....	iii
İÇİNDEKİLER	iv
ŞEKİLLER DİZİNİ	vi
ÇİZELGELER DİZİNİ	ix
KISALTMALAR.....	x
1. GİRİŞ	1
1.1. Çalışmanın Amacı.....	4
1.2. Tez Düzeni	6
1.3. Literatür Özeti	6
2. DERİN ÖĞRENME	11
2.1. Yapay Sinir Ağları	11
2.1.1. Algılayıcı	11
2.1.2. Aktivasyon Fonksiyonları.....	14
2.1.2.1. Lojistik Aktivasyon Fonksiyonu.....	14
2.1.2.2. Hiperbolik Tanjant Aktivasyon Fonksiyonu.....	14
2.1.2.3. Düzeltilmiş Doğrusal Birim Aktivasyon Fonksiyonu	15
2.1.2.4. Softmax Aktivasyon Fonksiyonu.....	16
2.2. Evrimsel Sinir Ağları	16
2.2.1. Evrim.....	17
2.2.2. Filtreler.....	18
2.2.3. Biriktirme İşlemi.....	18
2.2.4. Adım Atlama İşlemi.....	19
2.2.5. Evrimsel Sinir Ağı Mimarisi	20
2.3. Uzun Kısa Vadeli Bellek	20
2.4. Otokodlayıcı.....	23
2.4.1. Evrimsel Otokodlayıcı	24
2.4.2. Değişimsel Otokodlayıcı.....	25

3. ANOMALİ TESPİTİ	26
3.1. İstatistiksel Yöntemler	26
3.2. Makine Öğrenmesi Temelli Yöntemler	28
3.2.1. Gözetimli Makine Öğrenmesi Yöntemleri	28
3.2.2. Gözetimsiz Makine Öğrenmesi Yöntemleri	29
3.3. Derin Öğrenme Temelli Yöntemler	29
3.3.1. Gözetimli Derin Öğrenme Yöntemleri	30
3.3.2. Yarı Gözetimli Derin Öğrenme Yöntemleri	30
3.3.3. Gözetimsiz Derin Öğrenme Yöntemleri	30
3.4. Anomali Çeşitleri	30
3.5. Videolarda Anomali Tespitleri	31
3.5.1. Yeniden İnşa Modelleri (Reconstructive Models).....	32
3.5.2. Tahminsel Modelleme (Predictive Modeling).....	32
3.5.3. Üretici Modeller (Generative Models)	32
3.5.4. Anomali Skorları ve Eşik Değerleri.....	32
4. UYGULAMA	34
4.1. Veri Setleri ve Problem Tanımı	34
4.2. Kullanılan Yöntemler ve Mimariler	36
4.2.1. Optik Akış Yaklaşımı	36
4.2.2. Ham Veri Yaklaşımı	39
4.2.3. Otokodlayıcı Tabanlı Mimariler	40
4.3. Eğitim ve Test Süreci.....	42
4.4. Deneysel Sonuçlar	43
4.4.1. CAE Sonuçları	43
4.4.2. ConvLSTMAE Sonuçları	47
4.4.3. VAE Sonuçları.....	50
5. SONUÇLAR VE ÖNERİLER.....	54
6. KAYNAKLAR	56

ŞEKİLLER DİZİNİ

Şekil 1.1. 2012-2021 Yılları arasında video anomali tespiti alanındaki yayın sayıları (Web of Science)	1
Şekil 1.2. Renkli görüntü ve gri tonlanmış görüntü.....	4
Şekil 1.3. AUC grafiksel gösterimi.....	5
Şekil 2.1. Temel algılayıcı modeli	11
Şekil 2.2. İleri beslemeli sinir ağı modeli	12
Şekil 2.3. Gradyan inişi ile ağırlıkların ayarlanması	13
Şekil 2.4. Lojistik aktivasyon fonksiyonu	14
Şekil 2.5. Hiperbolik tanjant aktivasyon fonksiyonu.....	15
Şekil 2.6. Düzeltilmiş doğrusal birim aktivasyon fonksiyonu ve türevi.....	15
Şekil 2.7. Örnek bir evrişim işlemi	17
Şekil 2.8. Sobel filtresi.....	18
Şekil 2.9. Gri bir resme Sobel filtresinin uygulanmasıyla köşe tespiti.....	18
Şekil 2.10. Max biriktirme işleminine bir örnek.....	19
Şekil 2.11. İki Adım atlama ile max biriktirme işlemi	19
Şekil 2.12. Evrişimsel sinir ağı mimarisi	20
Şekil 2.13. LSTM hücresi	21
Şekil 2.14. LSTM hücresinde bulunan kapı yapısı	21
Şekil 2.15. Eşitliklerin LSTM hücresinde gösterilmesi	22
Şekil 2.16. Örnek bir otokodlayıcı mimarisi.....	23
Şekil 2.17. Evrişimsel otokodlayıcı mimarisi	24
Şekil 2.18. Değişimsel otokodlayıcı mimarisi	25
Şekil 3.1. İki boyutlu düzlemde anomalilerin gösterimi.....	26
Şekil 3.2. Tek değişkenli veri için kutu grafiği örneği	28
Şekil 4.1. UCSD Ped 1 veri seti test kümesi anomali örnekleri	34
Şekil 4.2. UCSD Ped 2 veri seti test kümesi anomali örnekleri	34
Şekil 4.3. CUHK Avenue veri seti test kümesi anomali örnekleri	35
Şekil 4.4. Optik akış ile hareketin gösterimi.....	36
Şekil 4.5. Optik akış yaklaşımı	38
Şekil 4.6. Optik akış örneği	39
Şekil 4.7. Ham veri yaklaşımı.....	39
Şekil 4.8. Kullanılan ConvLSTMAE mimarisi	40

Şekil 4.9. Kullanılan CAE mimarisi	41
Şekil 4.10. UCSD Ped 1 veri seti CAE - ham veri (sol) ve optik akış (sağ) yaklaşımı düzen skorları	43
Şekil 4.11. UCSD Ped 1 veri seti 130. kare CAE – (a) gerçek kare, (b) yeniden inşa edilmiş optik akış, (c) optik akışlı kare görüntüleri	44
Şekil 4.12. UCSD Ped 1 veri seti 180. kare CAE – (a) gerçek kare, (b) yeniden inşa edilmiş optik akış, (c) optik akışlı kare görüntüleri	44
Şekil 4.13. UCSD Ped 2 veri seti CAE - ham veri (sol) ve optik akış (sağ) yaklaşımı düzen skorları.....	45
Şekil 4.14. UCSD Ped 2 veri seti 145. kare CAE – (a) gerçek kare, (b) yeniden inşa edilmiş optik akış, (c) optik akışlı kare görüntüleri	45
Şekil 4.15. CUHK Avenue veri seti CAE - ham veri (sol) ve optik akış (sağ) yaklaşımı düzen skorları.....	46
Şekil 4.16. CUHK Avenue veri seti 4. kare CAE – (a) gerçek kare, (b) yeniden inşa edilmiş optik akış, (c) optik akışlı kare görüntüleri	46
Şekil 4.17. UCSD Ped 1 veri seti ConvLSTMAE - ham veri (sol) ve optik akış (sağ) yaklaşımı düzen skorları	47
Şekil 4.18. UCSD Ped 1 veri seti 180. kare ConvLSTMAE – (a) gerçek kare, (b) yeniden inşa edilmiş optik akış, (c) optik akışlı kare görüntüleri.....	47
Şekil 4.19. UCSD Ped 2 veri seti ConvLSTMAE - ham veri (sol) ve optik akış (sağ) yaklaşımı düzen skorları	48
Şekil 4.20. UCSD Ped 2 veri seti 145. kare ConvLSTMAE – (a) gerçek kare, (b) yeniden inşa edilmiş optik akış, (c) optik akışlı kare görüntüleri.....	48
Şekil 4.21. CUHK Avenue veri seti ConvLSTMAE - ham veri (sol) ve optik akış (sağ) yaklaşımı düzen skorları	49
Şekil 4.22. CUHK Avenue veri seti 4. kare ConvLSTMAE – (a) gerçek kare, (b) yeniden inşa edilmiş optik akış, (c) optik akışlı kare görüntüleri.....	49
Şekil 4.23. UCSD Ped 1 veri seti VAE - ham veri (sol) ve optik akış (sağ) yaklaşımı düzen skorları.....	50
Şekil 4.24. UCSD Ped 1 veri seti 70. kare VAE – (a) gerçek kare, (b) yeniden inşa edilmiş optik akış, (c) optik akışlı kare görüntüleri	50
Şekil 4.25. UCSD Ped 2 veri seti VAE - ham veri (sol) ve optik akış (sağ) yaklaşımı düzen skorları.....	51

Şekil 4.26. UCSD Ped 2 veri seti 75. kare VAE – (a) gerçek kare, (b) yeniden inşa edilmiş optik akış, (c) optik akışlı kare görüntüleri	51
Şekil 4.27. CUHK Avenue veri seti VAE - ham veri (sol) ve optik akış (sağ) yaklaşımı düzen skorları.....	52
Şekil 4.28. CUHK Avenue veri seti 265. kare VAE – (a) gerçek kare, (b) yeniden inşa edilmiş optik akış, (c) optik akışlı kare görüntüleri	52

ÇİZELGELER DİZİNİ

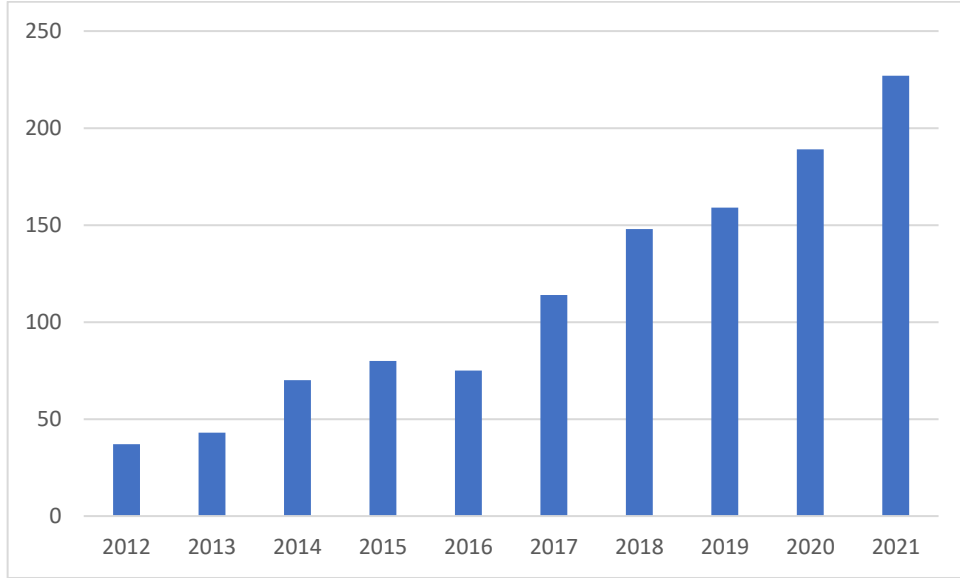
Çizelge 1.1. Ülkelere göre kapalı devre televizyon (CCTV) sayıları.....	2
Çizelge 4.1. Veri setleri için özet bilgiler	35
Çizelge 4.2. 100, 200 ve 500 faz ile eğitilmiş CAE mimarisi için AUC skorları.....	42
Çizelge 4.3. Mimari – yaklaşım / veri seti için ortalama AUC skorları	53

KISALTMALAR

AAE	Çekişmeli Otokodlayıcı
AUC	Eğri Altında Kalan Alan
CAE	Evrişimsel Otokodlayıcı
CCTV	Kapalı Devre Televizyon
CNN	Evrişimsel Sinir Ağları
ConvLSTMAE	Evrişimsel Uzun Kısa Vadeli Bellekli Otokodlayıcı
DBM	Derin Boltzman Makinesi
DBN	Derin İnanç Ağları
FCM	Bulanık C-Ortalamalar
GAN	Çekişmeli Üretici Ağ
HOF	Optik Akış Histogramı
HOG	Yönlendirilmiş Gradyanların Histogramı
KNN	K. En Yakın Komşuluk
LSTM	Uzun Kısa Vadeli Bellek
MPPCA	Olasılıksal Temel Bileşen Analizcileri Karışımı
MRF	Markov Rasgele Alan
OCSVM	Tek Sınıflı Destek Vektör Makinesi
ReLU	Düzeltilmiş Doğrusal Birim
RNN	Özyinelemeli Sinir Ağı
ROC	Alıcı İşlem Karakteristiği
SL-HOF	Uzaysal Yerelleştirilmiş Optik Akışların Histogramı
SOM	Kendini Düzenleyen Haritalar
ST-CaAE	Uzay-Zamansal Basamaklı Otokodlayıcı
VAE	Değişimsel Otokodlayıcı

1. GİRİŞ

İnsanların can ve mal güvenliğini sağlamak amacıyla kavşaklar, asansörler, hastaneler, bankalar, alışveriş merkezleri gibi çeşitli kamusal alanlara hatta evlere video gözetim kameraları konuşlandırılmaktadır. Çoğu şirket, kameraları izlemesi için kişileri işe alsada insan hatası nedeniyle kamera yayınlarındaki bazı anormal olayların kaçırılabilmesi olasıdır. Bu nedenle gözetleme kameralarını izlemek önemli zaman ve enerji kaybına neden olmaktadır. Videolarda anormal olayları otomatik tespit etmek son yıllarda gittikçe popülerlik kazanan en aktif araştırma alanlarından biridir. Son 10 yılda bu alanda Web of Science (WoS) tarafından taranan dergilerde yayınlanan yayınların sayıları Şekil 1.1’ de verilmiştir.



Şekil 1.1. 2012-2021 Yılları arasında video anomali tespiti alanındaki yayın sayıları

(Web of Science)

Çizelge 1.1’ den görüldüğü üzere birçok ülke için kapalı devre televizyon (CCTV) sayısı [1] çok fazladır ve fiyatlarının yüksek olmaması ve teknolojiye hızlı ilerlemeden dolayı her geçen yıl bu sayılar daha da artmaktadır. Kameraları izleyen sorumlu operatörün, insan başına düşen kamera sayısının fazla olması sebebiyle gün boyu videoları izlemesinin etkisiyle doğan yorgunluk, dikkat dağılması, manuel video

analizini imkânsız kılmakta, anomali tespiti başarısız olmaktadır. Operatör sayısı artırılrsa bile maliyet artacak, performans artışı sağlanamayacaktır.

Çizelge 1.1. Ülkelere göre kapalı devre televizyon (CCTV) sayıları

Ülke	Kamera Sayısı	Nüfus	100 Kişiye Düşen Kamera Sayısı
Birleşik Devletler	50.000.000	327.167.430	15.28
Çin	200.000.000	1.392.730.000	14.36
Birleşik Krallık	5.000.000	66.488.990	7.51
Almanya	5.200.000	82.927.920	6.27
Hollanda	1.000.000	17.231.020	5.80
Avustralya	1.000.000	24.992.370	4
Japonya	5.000.000	126.529.100	3.95
Vietnam	2.600.000	99.540.400	2.72
Fransa	1.650.000	66.987.240	2.46
Güney Kore	1.030.000	51.635.260	1.99

Bilgisayar donanımlarının yıllar içerisinde gelişmesiyle birlikte oldukça büyük boyutlu olan video verilerinin işlenmesi de kolaylaşmıştır. Bu durum video işleme, görüntü işleme gibi konuları içeren bilgisayarlı görü (computer vision) alanı için önemli bir gelişme sayılmaktadır.

Videolarda anomali tespitleri güvenlik alanında (kamu güvenliği, mobese kameralarında olay ve kaza tespiti), eğitim alanında (kopya tespiti), araç sistemlerinde (otonom araçlar, şerit takip sistemleri) ve tıp alanında (hastalık, kalp krizi, inme tespiti) kullanılmaktadır.

Güvenlik kameralarının kullanılmasıyla elde edilen video verileri, güvenlik sebebi başta olmak üzere anomali tespit sistemlerinin otomatize edilmesi ihtiyacını ortaya çıkarmıştır. Anomali tespiti özellikle video gibi büyük verilerle birlikte çeşitli zorlukları

bünyesinde barındırmaktadır. Bunlar, sistemlerin gerekli işlem güçleri ve bunu sağlayacak donanımlar, mimarilerin eğitimi için gerekli süreler, eğitimde ve test aşamasında kullanılacak verilerin içerik olarak değişkenlik göstermesi olarak belirtilebilir. Örneğin bir kavşağa yerleştirilen mobese kamerası verilerine göre anomaliler trafik kazaları, hatalı sollamalar, tehlikeli veya hızlı araç kullanımı olabilirken, bir parka yerleştirilmiş güvenlik kamerasına göre kavga gibi olaylar, parka giren araçlar, yaya yolu dışında yürümler anomalilere örnek olarak verilebilir.

Anomalilerin tanımsal olarak değişmesi durumuna karşın, veri seti ve anomali durumuna bağlı ayrı ağlar eğitilmelidir. Örneğin trafik kazalarının tespiti için eğitilmiş bir ağ, yayaların davranışlarında anomali tespiti yapmak için kullanılamaz. İlgili ağ mimarileri veri seti özelinde, veriden öznitelik çıkarımı yapmakta ve buna uygun olarak normal davranışları öğrenmektedir.

Anomali tespit edici sistemin öğrenmesi gereken davranışlar, örüntüler karmaşıklaştıkça mimarinin yapısının da derinleşmesi, dolayısı ile daha çok veriye, işlem gücü ve belleğe ihtiyaç duymalarına sebep olmaktadır. Öne çıkan bir husus, derin ağ yapıları sadece eğitildiği veri ile öğrenmektedir. Eğitim kümesinde olmayan davranış veya örüntüler bu ağ mimarileri için tanınmayacak, dolayısı ile anomali olarak kabul edilecektir. Sistemler aynı zamanda her davranışı öğrenemeyebilir veya öngöremeyebilirler. Örneğin arka planda hareket eden bulutların görüntülerini, parka yerleştirilmiş mobese kamerasının önünden geçen kuşları, gökyüzünden geçen uçakları, kişilerin günün saatine bağlı uzunluğu değişen gölgelerini anomali olarak tespit edebilirler.

Karşılaşılan zorluklardan bir diğeri, eğitim verilerinin dengesiz olmasıdır. Anomaliler doğaları gereği nadir olduklarından, normal veriler oldukça yüksek sayıda bulunurken, anormal olayları içeren veriler çok daha az sayıda bulunmaktadır. Veri setindeki dengesizlik, anomali tespiti işlemi için doğruluğu düşürücü bir durumdur.

Bunların yanı sıra, kullanılacak video görüntülerinin seçimi, anomalilerin başlangıçta belirlenmesi ve işaretlenmesi işlemleri insan gücüne bağlı olduğundan, ağ eğitimi hata

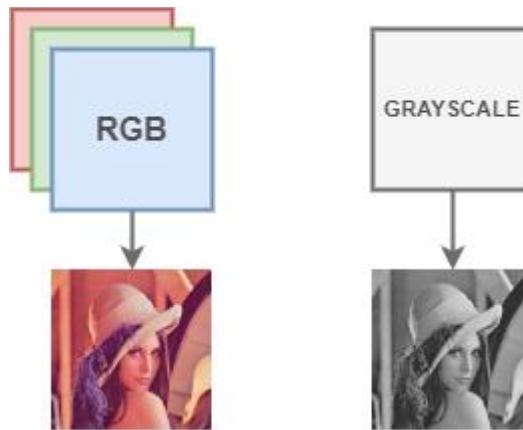
payı içermektedir. Ayrıca gizlilik ve kişisel hakların ihlali de karşılaşılan bir başka zorluktur. Video olmayan verilerde anomali tespiti yaparken, analize etkisi olmayan kişisel bilgiler sansürlenebilirken (örneğin kimlik numaraları), video verilerinde kişilerin yüzleri, hareketleri sansürlenememekte, bu durum gizliliği ihlal etmektedir.

Video anomali tespitlerinde birçok yaklaşım mevcuttur. Bunlardan ikisi, ham veri yaklaşımı ve optik akış yaklaşımıdır. Bu yaklaşımlar, anomalileri tespit ederken görüntülerin örüntüsel öznitelikleri veya hareket bazlı özniteliklerinin kullanılması durumu olarak özetlenebilmektedir.

1.1. Çalışmanın Amacı

Tez çalışmasında, videolarda anomali tespiti için iki yaklaşım karşılaştırılmıştır. Bunlar, geleneksel ham veri yaklaşımı ve video verilerinin özniteliklerinden birisi olan optik akış yaklaşımıdır. Bu iki yaklaşımın belirlenen üç veri seti için, farklı mimariler ile karşılaştırması yapılmıştır. Böylece hangi yaklaşımın daha iyi performans gösterdiğine yönelik değerlendirmelerin yapılması planlanmıştır.

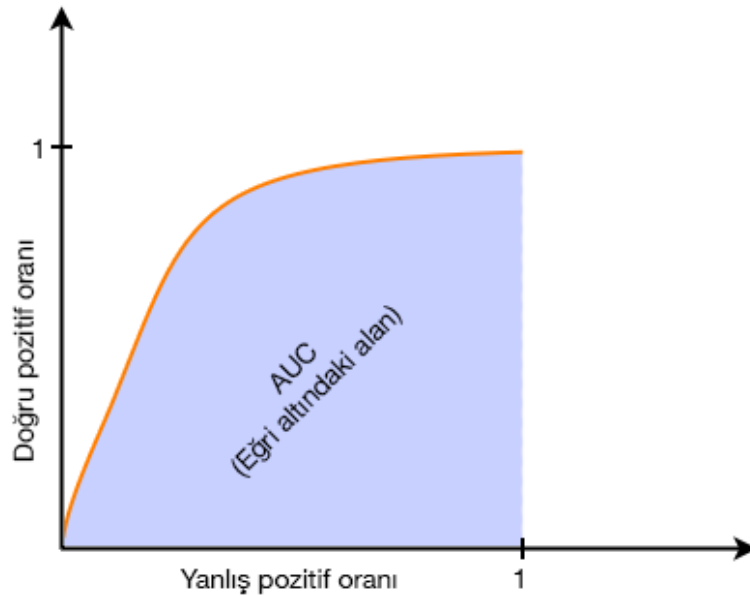
Görüntülerde ön işleme olarak ölçeklendirme ve gri tonlama (grayscale) (Şekil 1.2) yapılmıştır. Optik akış yaklaşımında kullanılan yöntem Farneback (Dense Optical Flow) yöntemidir. Hıza bağlı optik akış tespit edileceği için literatürde özellikle Farneback yönteminin kullanılması önerilmektedir [2].



Şekil 1.2. Renkli görüntü ve gri tonlanmış görüntü

Kullanılan veri setleri çeşitli anomaliler içermektedir. Bunlar yaya yoluna giren araç, bisikletli ve kaykaylılar, yanlış yönde yürüyen, kameranın önünden koşarak geçen ve kameranın önünde nesne atan yayalardan oluşmaktadır. İlgili veri setlerinde eğitim kümeleri normal yaya davranışlarını ve normal hızda yürüyüşü içermektedir. Anomali durumlarda ise her veri setinde ortak olarak, normal yaya hızından daha hızlı hareket durumu söz konusudur.

Performans ölçütü olarak Alıcı İşlem Karakteristiği (ROC) eğrisi altında kalan alan (AUC) skoru kullanılmıştır (Şekil 1.3). Bu skor gerçek pozitif sınıflandırma oranı ile yanlış pozitif sınıflandırma oranları grafiğinin altında kalan alanı içermektedir. $[0,1]$ aralığında değer almaktadır. AUC skorunun yüksek olması, daha doğru bir sınıflandırma olduğu anlamına gelmektedir. Sınıflandırma oranları için otokodlayıcı tabanlı mimari çıktısından elde edilen yeniden inşa hatası (reconstruction error) ile düzen skoru (regularity score) elde edilmiştir. Bu düzen skoru, belirlenen bir eşik (threshold) değerinden düşük ise anomali, değil ise normal görüntü karesi olarak sınıflandırılmıştır.



Şekil 1.3. AUC grafiksel gösterimi

1.2. Tez Düzeni

Tez çalışmasının ikinci bölümünde, derin öğrenme yöntemleri incelenmiştir. Yapay sinir ağları ve kullanım durumuna göre özelleştirilen, sinir ağlarından türetilmiş hibrit ağlar ele alınmıştır.

Üçüncü bölümde, anomali tespit yöntemleri incelenmiştir. Anomali ve aykırı değer kavramları, istatistiksel, makine öğrenmesi ve derin öğrenme temelli yaklaşımlar, video anomali tespiti yöntemleri ile ele alınmıştır.

Dördüncü bölümde, uygulama çalışması yapılmıştır. Problemin tanımı, veri setlerinin açıklaması, kullanılan mimariler ve sonuçlar açıklanmıştır.

Son bölüm olan beşinci bölümde, yaklaşımların performansları detaylıca yorumlanmıştır. Gelecekte yapılması planlanan çalışmalara değinilmiştir.

1.3. Literatür Özeti

Videolarda anomali tespitleri için literatürde birçok yaklaşım ve anomali tespiti çerçevesi (framework) önerilmiştir. Bunlar verinin ve anomalinin türüne göre değişebildiği gibi mimari bazlı da değişebilmektedir.

Ham veri yaklaşımında video görüntülerinden öznitelik çıkarımı (feature extraction), örüntü tanıma (pattern recognition), davranış tanıma (behaviour recognition) yapılmaktadır. Böylece eğitim setinde gerçekleştirilen tüm davranışların ve örüntülerin öznitelikleri öğrenilerek öngörüselleme (predictive modeling) yapılmaktadır.

Optik akış yaklaşımında ise, iki video karesi arasındaki farktan, parlaklık bilgisine göre hareket hesaplanmaktadır. Bu hareket ile öznitelik çıkarımı yapıldığı zaman hız, hareket yönü, hareket şiddeti, hareket eden cismin boyutu, hareket örüntüleri gibi özellikler öğrenilmektedir. Bu özellikler ile hareket bazlı tahmin yapılmaktadır.

Temelinde iki yaklaşım da aynı amaca sahiptir. Geçmiş hareketler ve davranışlar öğrenilerek gelecek hareketler ve davranışlar tahmin edilmektedir. Tahmin edilen ile gerçek olan görüntü veya öznitelik arasındaki fark, anomali durumu hakkında bilgi vermektedir. Normal örüntüleri öğrenip normal olmayan (anomali) örüntüleri ve davranışları tespit etme yöntemi uzun süredir literatürde çalışılan bir konu olmuştur [3 - 15].

Lu ve ark. [3], video görüntü verilerini seyrek (sparse) kombinasyonlarını içeren küp verilerine bölütlemişlerdir. Bu küp verilerinin üç boyutlu öznitelikleri çıkarılarak temel bileşenler analizi yapılmış, böylece boyut indirgenmiştir. Bu boyut indirgeme sonucunda ortaya çıkan veriler, anomali içeren durumlarda yüksek değerler almaktadır.

Hasan ve ark. [4], yörünge (trajectory) yaklaşımı ile yaptıkları çalışmada görüntülerin yönlendirilmiş gradyanların histogramı (HOG) ve optik akışların histogramı (HOF) özniteliklerinin (yörünge öznitelikleri de denir) çıkarılarak otokodlayıcı mimarisinde yeniden inşa edilmesini önermişlerdir. İlgili öznitelikler düzenli görüntülerde (regular frame, anomali olmayan normal kareler) hareket bilgisini içermektedir. Eğitim kümesinde öğrenilen hareket bilgileri, test kümesinde yüksek yeniden inşa skoru verecek ve anomali tespitini sağlayacaklardır.

Luo ve ark. [5], Evrimsel Uzun Kısa Vadeli Bellekli Otokodlayıcı (ConvLSTMAE) mimari bazlı çalışmada aktivite tanımayı (activity recognition) amaçlamışlardır. Video görüntülerinin geçmiş uzaysal (spatial) ve zamansal (temporal) bilgileri öğrenilerek gelecek bilgiler tahmin edilmeye çalışılmıştır. Video karelerindeki değişimler karakterize edilerek anomali tespiti işlemi hareket bazlı olarak yapılabilmektedir. Burada önerilen mimari, aynı zamanda tez çalışmasının ana unsurlarından birisidir.

Adam ve ark. [6], optik akış temelli çalışmada hareket temelli anomali tespitlerinin güçsüz yanlarını vurgulamış, buna bir çözüm önermişlerdir. Hareket temelli anomalilerin video verilerinde her zaman tespit edilemediği ve yanıltıcı olduğu, gölge,

bulut gibi arka planda gerçekleşen ve doğası gereği hareket eden cisimler sebebiyle daha güçlü bir yönteme ihtiyaç duyulduğu vurgulanmıştır. Önerilen yöntem ise, video karelerinin daha alt seviye optik akış özniteliklerinin çıkarılması olmuştur. Bu optik akış hesaplama yöntemi olasılıksal olarak yapılarak akış olasılık matrisi oluşturulması önerilmiştir.

Mehran ve ark. [7], kalabalık ortamlarda bulunan insanların davranışsal özelliklerini tespit eden Sosyal Güç (Social Force) modelini önermişlerdir. Bu model anormal kalabalık davranışını modelleyerek anomali tespiti yapmaktadır. Modelleme işlemi optik akış verilerinin her birinin ayrı olarak incelenmesi ile sağlanmaktadır.

Li ve ark. [8], Uzay-Zamansal Basamaklı Otokodlayıcı (ST-CaAE) mimarisini önermişlerdir. Mimari iki farklı otokodlayıcı içermektedir. İlk olarak ham video verilerinin optik akış küboidleri hesaplanmaktadır. Gradyan ve optik akış küboidleri ilk otokodlayıcıya girdi olarak verilmektedir. İlk aşamada anomaliden şüphelenilen hareket küboidleri kalmakta, normal veriler elenmektedir. İkinci aşamada ise kalan şüpheli hareketler ikinci otokodlayıcıya geçerek anomali tespiti yapılmaktadır.

Ravanbakhsh ve ark. [9], Çekişmeli Üretici Ağ (GAN) mimarisi içeren bir yapı önermişlerdir. GAN mimarileri üretken ağlar olduklarından teknik olarak otokodlayıcılar gibi yeniden inşa yapamazlar. Tamamen gözetimsiz ağlardır. Normal görüntü karelerinden kalabalık davranışlarını içeren görüntüleri üretirler. Anomali içeren durumu üretmeyecekleri için iki görüntü karesi arasındaki aşırı fark anomali tespiti yapmaktadır. Bu farkları hem ham veri ile hem de optik akış verileri ile test etmişlerdir.

Duman ve Erdem [10], optik akış verilerinin ConvLSTMAE mimarisi ile yeniden inşasını içeren bir çerçeve önermişlerdir. Bu önerilen yapıda otokodlayıcının yeniden inşa hatası kullanılarak düzen skoru elde edilmiş, bu düzen skorundan anomali tespiti yapılmıştır. Önerilen bu çerçeve tez çalışmasında optik akış yaklaşımı altında kullanılmıştır.

Qiao ve ark. [11], video görüntülerinin Horn-Schunck yöntemi ile elde edilen optik akış verilerinin derin otokodlayıcı mimarisine dahil edilmesi ile oluşturulan bir mimari önermişlerdir. Burada önemli nokta bu çalışmada kullanılan derin otokodlayıcı, Evrişimsel Otokodlayıcı (CAE) değildir. 240x320 çözünürlüğe sahip görüntüler için piksel sayısı kadar (76800) girdi nöronu içeren otokodlayıcı mimarisi oluşturulmuştur.

Wang ve ark. [12], Uzaysal Yerelleştirilmiş Optik Akışların Histogramı (SL-HOF) yaklaşımını önermişlerdir. Klasik HOF özneliğinde görüntüler küplere bölünür ve her küpün kendi içerisinde optik akış histogramı çıkarılır. Önerdikleri SL-HOF yönteminde ise küboidin tüm HOF öznelikleri konsantre edilmiş (birleştirilmiş) haldedir. Bölgesel olarak tüm HOF özneliklerini içeren bu veriler Tek Sınıflı Destek Vektör Makinesi (OCSVM) mimarisi ile kategorize edilerek anomali tespiti işlemi yapılır.

Kim ve Grauman [13], yeni bir optik akış çıkarımı ve bu optik akış ile yeni bir anomali tespiti yaklaşımı önermişlerdir. Optik akış çıkarımı temel bileşenler analizi temelli Olasılıksal Temel Bileşen Analizcileri Karışımı (MPPCA) yaklaşımı ile sağlanır, ardından Markov Rasgele Alan (MRF) parametreleri içeren küboid hesaplanır. Bu küboidde Bayesci yaklaşım ile normal noktalar ve anomali noktalar belirlenir. Bu yaklaşımın diğer yaklaşımlara göre güçlü olduğu nokta, zaman içerisinde değişen davranışları da hafızasına alarak kendini güncellemesidir.

Yan ve ark. [14], doğal dil işleme mimarilerinden olan Uzun Kısa Vadeli Bellekli Değişimsel Otokodlayıcı (LSTM VAE) mimarisine yenilik katarak oluşturulan Evrişimsel Değişimsel Otokodlayıcı (R-Conv VAE) mimarisini önermişlerdir. Kare bazlı anomali tespiti için geliştirilen bu yönteme ek olarak iki akışlı mimari de önerilmiştir. Akışlardan birinde ham veri ile uzay-zamansal öznelikleri öğrenilirken diğerinde optik akış ile hareket öğrenimi gerçekleştirilmektedir. Olay tespiti yapması hedeflenen mimari, uzaysal bilgileri hafızasında tutarken aynı zamanda uzun vadeli anomalileri de tespit edebilmiştir.

Luo ve ark. [15], anomali tespiti yöntemi olarak Yığılanmış Özyinelemeli Sinir Ağı (Stacked RNN) mimarisi önermişlerdir. Önerilen yöntemin dezavantajı piksel komşuluğunun zamansal değişimlerini tespit edememektir. Özyinelemeli Sinir Ağı (RNN) mimarilerinde çoklu yapıda LSTM hücreleri bulunmakta ve noktasal değerlerin zamana bağlı özelliklerini hem kısa vadeli hem de uzun vadeli olarak hafızasında tutmaktadır. Noktasal olduğu için piksel komşulukları göz ardı edilmektedir. Buna karşın RNN hafızalarının çokluğundan dolayı video verilerinde iyi performans gösterdiği görülmektedir.

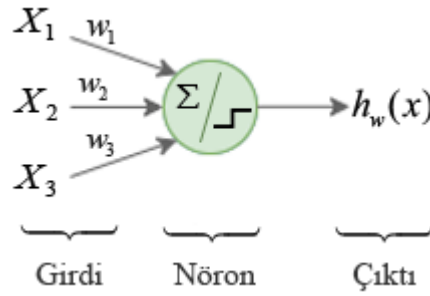
2. DERİN ÖĞRENME

2.1. Yapay Sinir Ağları

Yapay sinir ağları (ANN), beyinde bulunan sinirlerden ve özelliklerinden ilham alınarak oluşturulan bir makine öğrenmesi modelidir. En az bir nörondan oluşan bir katmanın girdi ve çıktı katmanlarına bağlanmasıyla oluşmaktadır. Katmanlarda bulunan her nöron aktivasyon fonksiyonu içermektedir ve nöronlar birbirlerine w ağırlıkları ile bağlıdır.

2.1.1. Algılayıcı

En temel yapay sinir ağlarından biridir. 1957’de Frank Rosenblatt tarafından icat edilmiştir. Algılayıcı (perceptron) modelinde girdiler w ağırlıkları ile aktivasyon fonksiyonu içeren nöronlara bağlıdır (Şekil 2.1). Bu aktivasyon fonksiyonu ise genelde adım (step) fonksiyonu veya işaret fonksiyonudur.



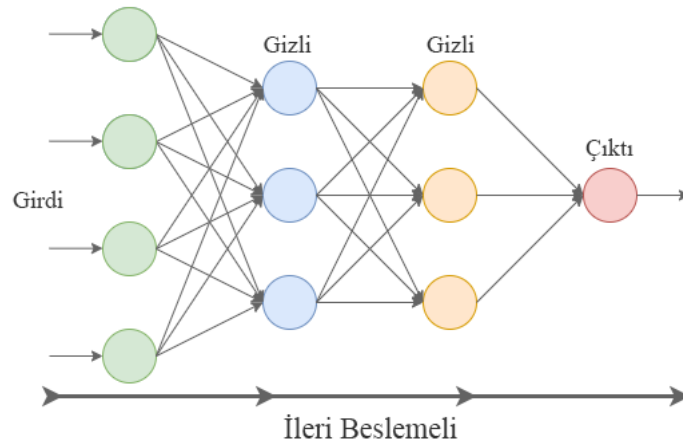
Şekil 2.1. Temel algılayıcı modeli

Burada çıktı $h_{w,x}(x) = \phi(Xw + b)$ ile hesaplanır. Adım fonksiyonu ise Eşitlik 2.1’de verilmiştir.

$$Adım(z) = \begin{cases} 0, & z < 0 \\ 1, & z \geq 0 \end{cases} \quad (2.1)$$

Aktivasyon fonksiyonu içeren nöron, ağırlıklı girdilerin toplamını hesaplar. Daha sonra aktivasyon fonksiyonunu uygulayarak elde ettiği çıktıyı sonraki nörona aktarır. Algılayıcı her bir girdinin bir nörona bağlanmasıyla oluşur. Girdiler en az bir nörona ve en az bir çıktıya bağlandığında ise oluşan yapıya yoğun katman veya tam bağlanmış katman denir. Bu şekilde oluşan modele ise çok katmanlı algılayıcı (MLP) denir. MLP girdi katmanı, gizli katman olarak da bilinen ve en az bir algılayıcı içeren katman veya katmanlardan, çıktı katmanından oluşmaktadır.

MLP modelinde nöronlar arası sinyaller tek yönlü olarak, ileri yönde (girdi katmanından çıktı katmanına) ilerlediğinde bu modele ileri beslemeli sinir ağı denir (Şekil 2.2).



Şekil 2.2. İleri beslemeli sinir ağı modeli

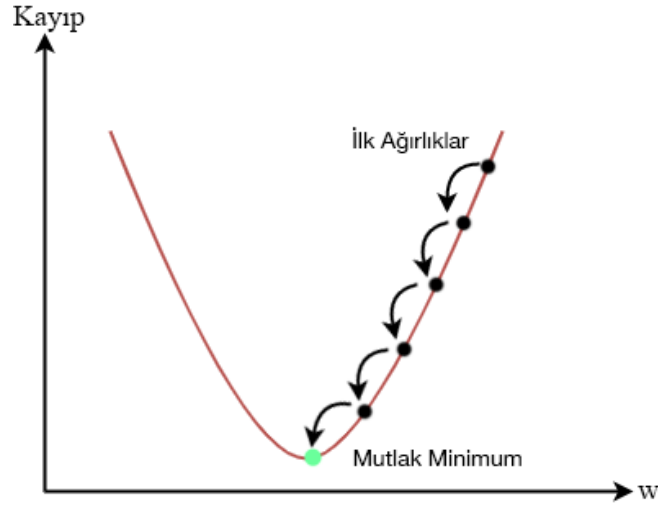
Yapay sinir ağları uzun yıllardır bilinmesine rağmen bu ağları eğitmenin yolu 1986'da D. Rumelhart ve ark. [16] tarafından keşfedilmiştir. Bu yöntem geri yayılım (backpropagation) algoritması olarak bilinmektedir. Bu algoritma temel olarak ağın her parametresine göre hataların gradyanını hesaplamaktadır. Hataların azalması için w ağırlıklarının nasıl değişmesi gerektiği hesaplanır.

Geri yayılım algoritmasının adımları şöyledir;

1. İleri besleme ile eğitim gerçekleşir ve hata hesaplanır.
2. Her w ağırlığının hataya olan katkısını hesaplamak için her katmanda geriye doğru ilerlenir.
3. Hatayı azaltmak için ağırlıklar ayarlanır. (Gradyan inişi / Gradient descent)
(Şekil 2.3)

$$w = w - \alpha \frac{\partial \text{kayıp}}{\partial w} \quad (2.2)$$

Eşitlik 2.2'de α öğrenme hızı, değeri bilinen ve 0.01 gibi küçük bir değerdir.



Şekil 2.3. Gradyan inişi ile ağırlıkların ayarlanması

Gradyan iniş düz bir yüzeyde gerçekleşemeyeceği için aktivasyon fonksiyonunun adım veya işaret fonksiyonundan farklı olması gerekmektedir [17]. Kullanılan popüler aktivasyon fonksiyonlarından bazıları;

- Lojistik (Sigmoid),
- Hiperbolik Tanjant,
- Düzeltilmiş Doğrusal Birim (ReLU),
- Softmax'dır.

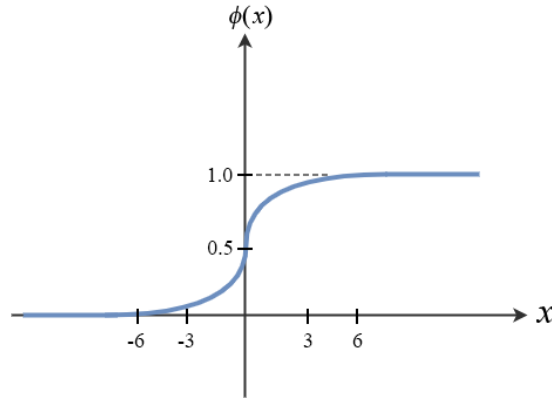
2.1.2. Aktivasyon Fonksiyonları

Girdilerden nöronlara giden bilgiler önce toplanır, daha sonra bu toplam bir aktivasyon fonksiyonuna aktarılır. İlgili aktivasyon fonksiyonunun amacı, kendi kuralları ve formülasyonu dahilinde bulunduğu nöronun aktive olmasını sağlamaktır. Böylece bilgi bir nörondan diğerine dönüştürülerek aktarılır.

2.1.2.1. Lojistik Aktivasyon Fonksiyonu

En çok kullanılan aktivasyon fonksiyonlarından birisidir. Nörona gelen bilgileri (0,1) aralığına dönüştürür. Genelde çıktı katmanında, çıktının olasılık olduğu durumlarda kullanılırlar [18]. Lojistik aktivasyon fonksiyonu Eşitlik 2.3'te, fonksiyonun grafiği Şekil 2.4'te verilmiştir.

$$\phi(x) = \frac{1}{1 + e^{-x}} \quad (2.3)$$

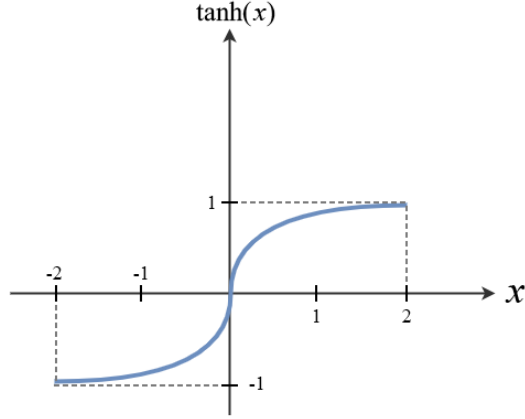


Şekil 2.4. Lojistik aktivasyon fonksiyonu

2.1.2.2. Hiperbolik Tanjant Aktivasyon Fonksiyonu

Lojistik aktivasyon fonksiyonuna benzemektedir. [-1, 1] aralığında değerler alır. Bu özelliği sebebiyle genelde ikili sınıflandırma problemlerinde çıktı katmanında kullanılmaktadır. Hiperbolik tanjant (tanh) aktivasyon fonksiyonu Eşitlik 2.4'te, fonksiyonun grafiği Şekil 2.5'te verilmiştir.

$$\tanh(x) = \frac{e^x - e^{-x}}{e^x + e^{-x}} \quad (2.4)$$

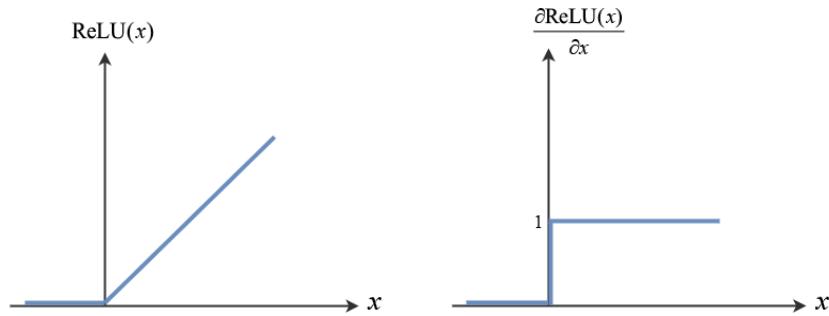


Şekil 2.5. Hiperbolik tanjant aktivasyon fonksiyonu

2.1.2.3. Düzeltilmiş Doğrusal Birim Aktivasyon Fonksiyonu

Düzeltilmiş doğrusal birim fonksiyonu (ReLU) derin öğrenmede, özellikle görüntü işlemede sıklıkla kullanılır [18]. Girdi negatif ise 0, diğer durumlarda girdinin kendisini çıktı olarak aktarır. Düzeltilmiş doğrusal birim aktivasyon fonksiyonu Eşitlik 2.5'te, fonksiyonun grafiği Şekil 2.6'da verilmiştir.

$$\text{ReLU}(x) = \begin{cases} 0, & x \leq 0 \\ x, & x > 0 \end{cases} \quad (2.5)$$



Şekil 2.6. Düzeltilmiş doğrusal birim aktivasyon fonksiyonu ve türevi

ReLU aktivasyon fonksiyonu süreklidir. Ancak negatif girdiler için 0 değerini aldığı için türevlenebilir değildir. Diğer aktivasyon fonksiyonlarına kıyasla hesaplaması çok daha hızlı ve kolaydır [18]. Parametre sayıları ve nöron sayıları evrişimsel sinir ağları (CNN) mimarilerinde diğer sinir ağlarına kıyasla çok daha fazla olduğundan, bu gibi durumlarda ReLU kullanmak işlem maliyeti açısından oldukça önemlidir.

2.1.2.4. Softmax Aktivasyon Fonksiyonu

İkiden fazla sınıf içeren çoklu sınıflandırma problemlerinde kullanılmaktadır [18]. Bu sınıflandırma çeşitlerinde ağ çıktıları her bir sınıf için negatif olabilir ve her zaman çıktılar toplamı 1'e eşit olmayabilir. Softmax kullanılarak her sınıf değeri için bir olasılık değeri elde edilebilmektedir. Standart durumda olmayan sınıf çıktıları (0,1) aralığına getirilerek olasılık dağılımı elde edilebilmektedir (Eşitlik 2.6).

$$\text{Softmax}(x_i) = \frac{\exp(x_i)}{\sum_j \exp(x_j)} \quad (2.6)$$

Örneğin 5 elemana sahip çıktı vektörü aşağıdaki gibi olsun. Softmax aktivasyon fonksiyonu, bu vektörü olasılıksal olarak dağıtmaktadır.

$$\begin{bmatrix} 1.3 \\ 5.1 \\ 2.2 \\ 0.7 \\ 1.1 \end{bmatrix} \rightarrow \text{Softmax} = \frac{\exp(x_i)}{\sum_{j=1}^5 \exp(x_j)} \rightarrow \begin{bmatrix} 0.02 \\ 0.90 \\ 0.05 \\ 0.01 \\ 0.02 \end{bmatrix}$$

2.2. Evrişimsel Sinir Ağları

Evrişimsel Sinir Ağları (CNN), matris tipinde verileri işlemek için kullanılan, yapay sinir ağlarının özel bir çeşidi olan ağlardır. Zamana bağlı tek boyutlu veriler veya iki boyutlu, piksellerden oluşan görüntü verileri matris tipi verilere örnek olarak verilebilir. Bilgisayarlı görü alanında sıklıkla kullanılmaktadır. Kişi ve nesne tanıma, görüntü sınıflandırma, görüntü işlemenin yanı sıra ses tanıma, doğal dil işleme gibi alanlarda da kullanılırlar. Evrişim terimi ise, sinir ağı içerisinde gerçekleştirilen evrişim (convolution) adı verilen matematiksel işlem den gelmektedir.

2.2.1. Evrişim

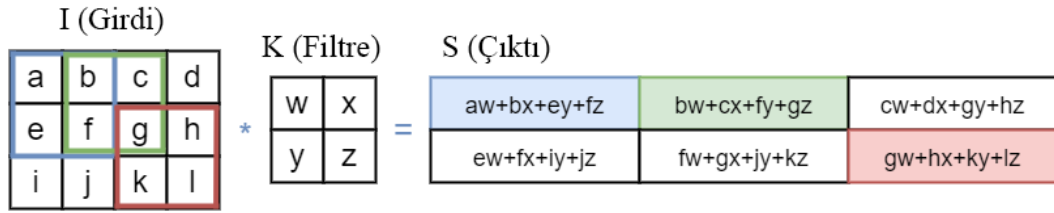
f ve g gibi iki fonksiyondan üçüncü bir s gibi fonksiyonu oluşturan matematiksel bir işlemdir [19, 20]. Evrişimsel sinir ağlarında evrişim işleminin ilk değişkeni genellikle x fonksiyonu ile çok boyutlu bir veri dizisi girdileridir. İkinci değişken w fonksiyonu ile filtreler (kernel, tensör)'dir. Bu tensörler öğrenme algoritması tarafından oluşturulan çok boyutlu parametreler dizisidir. Çıktı ise öznelik haritası (feature map) olarak adlandırılır (Eşitlik 2.7).

$$s(t) = (x * w)(t) = \sum_a x(a) w(t-a) \quad (2.7)$$

Eşitlik 2.7'de t ve a zaman indeksleridir. $*$ ise evrişim işlemini ifade etmektedir (Şekil 2.7). Genelde bir anda birden fazla eksen üzerinde evrişim işlemi kullanılır. Örneğin iki boyutlu I görüntüsü girdi olarak, iki boyutlu bir K tensörü ile kullanılırsa Eşitlik 2.7 artık Eşitlik 2.8 gibi olur. Evrişim işlemi sıra bağımsızdır. (Eşitlik 2.9)

$$S(i, j) = (I * K)_{(i,j)} = \sum_m \sum_n I(m, n) K(i-m, j-n) \quad (2.8)$$

$$\sum_m \sum_n I(m, n) K(i-m, j-n) = \sum_m \sum_n I(i-m, j-n) K(m, n) \quad (2.9)$$



Şekil 2.7. Örnek bir evrişim işlemi

2.2.2. Filtreler

Evrişim işleminde kullanılan tensörlerdir. Temelinde görüntü verileri üzerinde çeşitli işlemler yapmaya, görüntüleri işlemeye yaramaktadır. Karesel matris verileridir. Görüntü matrisi üzerinde gezerek çarpma ve toplam işlemi yaparlar. Filtreler ile işlem örneği Şekil 2.7’de verilmiştir.

1	2	1
0	0	0
-1	-2	-3

-1	0	1
-2	0	2
-1	0	1

Şekil 2.8. Sobel filtresi

Şekil 2.8’ de verilen filtre, köşe tespiti (edge detection) uygulamalarında kullanılmaktadır. Köşe tespitine bir örnek uygulama Şekil 2.9’da verilmiştir.

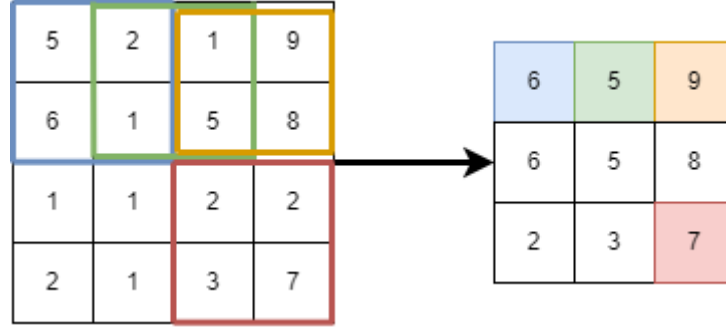


Şekil 2.9. Gri bir resme Sobel filtresinin uygulanmasıyla köşe tespiti

2.2.3. Biriktirme İşlemi

Biriktirme işleminde boyut indirgenmektedir. Evrişimsel sinir ağlarında ise bu işlemdeki amaç, minimum kayıp ile boyut indirgenerek parametre sayısının azaltılması ve aşırı öğrenmenin (overfitting) önlenmesidir [21]. Biriktirme işlemlerinden bazıları min biriktirme (min pooling), ortalama biriktirme (avg. pooling) ve en çok kullanılan

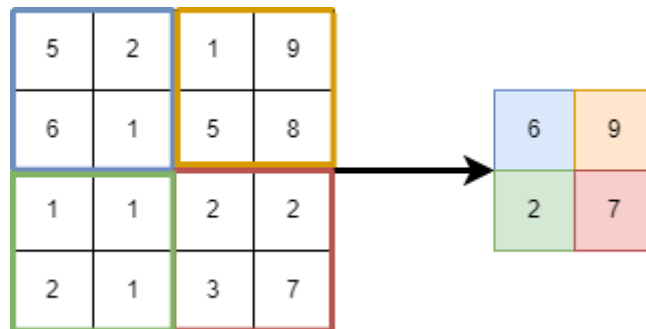
max biriktirme (max pooling)'dir. Bu işlemler evrişim işleminde olduğu gibi karesel tarama şeklinde gerçekleşir. Taranan karenin maksimum, minimum değerleri veya ortalaması hesaplanarak tek değere indirgenir. Şekil 2.10'da 4x4 boyutlu bir veri 2x2 max biriktirme işlemi sonucunda 3x3 boyutuna indirgenmiştir.



Şekil 2.10. Max biriktirme işleminine bir örnek

2.2.4. Adım Atlama İşlemi

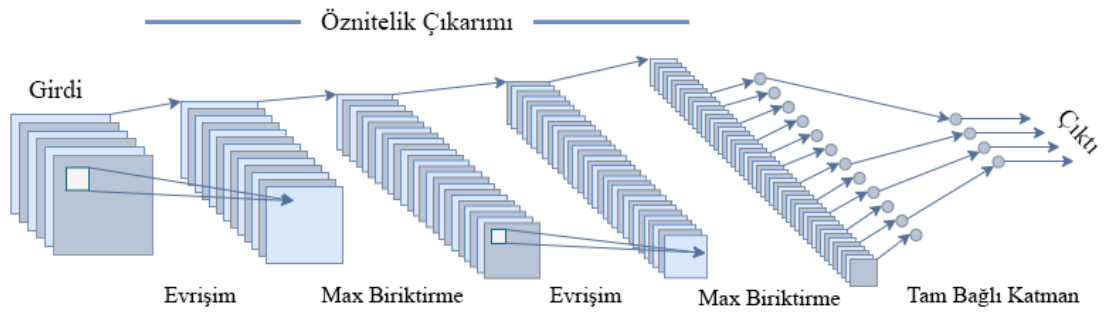
Boyut indirgemenin bir diğer yolu ise, adım atlama (stride) işlemidir. Evrişim işleminde veya biriktirme işleminde karesel tarama alanı her seferinde bir birim kayarak ilerlemektedir. Bu kayma işlemi adım atlama olarak adlandırılmaktadır. Adım atlama iki olduğunda tarama alanı her seferinde iki birim kayarak ilerleyecektir. Örneğin 4x4 boyutlu bir veri, 2x2 max biriktirme işleminde adım atlama sayısı iki olduğunda, işlem sonucunda 2x2 boyutlu bir veriye indirgenecektir (Şekil 2.11).



Şekil 2.11. İki Adım atlama ile max biriktirme işlemi

2.2.5. Evrişimsel Sinir Ağı Mimarisi

Evrişimsel sinir ağları, evrişim ve biriktirme katmanlarının art arda gelmesiyle oluşmaktadır. Girdi olarak verilen matris şekilli veri (örneğin görüntü verisi) önce evrişim katmanına girerek boyutu indirgenir ve öznitelik haritaları çıkarılır. Bu katmanda kullanılan filtreler (kernel), parametrelerin sinir ağı tarafından güncellenmesi ile oluşmaktadır. Böylece en doğru öznitelik çıkarımı hedeflenmektedir. Oluşturulan öznitelikler biriktirme katmanına giderek en anlamlı veriler seçilerek boyut indirgenir. Bu ikili katman kombinasyonu mimari seçiminde belirtildiği kadar tekrarlanır. Her evrişim katmanında daha üst seviye öznitelikler çıkarılır [22]. Her evrişim katmanı aktivasyon fonksiyonu ile biriktirme katmanına bağlanır. Görüntü verilerinde genellikle ReLU aktivasyon fonksiyonu kullanılmaktadır [18].



Şekil 2.12. Evrişimsel sinir ağı mimarisi

Son aşama olarak biriktirme katmanından sonra öznitelikler düzleştirilir, yani matris formundan dizi formuna dönüştürülür. Bu veriler yoğun katmana ağırlıklar ile bağlanır. Mimaride belirtildiği kadar yoğun katman tekrarlanarak çıktı katmanına bağlanır (Şekil 2.12). Burada çıktı katmanı probleme göre değişmektedir. Örneğin görüntü verileri ile ikili sınıflandırma işlemi yapılacaksa, çıktı katmanında tek nöron yeterlidir.

2.3. Uzun Kısa Vadeli Bellek

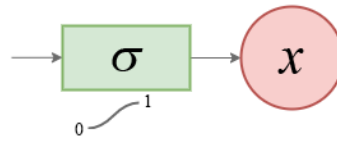
Zamana bağlı verilerin analizinde sinir ağlarının eğitim süresi oldukça uzamaktadır. Bu sorunun bir çözümü, girdi olarak verilen verinin zaman aralığını kısaltmaktır. Ancak bu yöntem uzun trendlerin tahmininde verimli olmamaktadır. Uzun sekanslar ile sinir ağı

eğitmenin yarattığı diğer sorun ise, verinin çokluğundan dolayı eğitim sürecinde ağın ilk girdileri unutulmasıdır. Ağırlıklar güncellenirken ilk verilerin ağırlığı sonraki verilerde önemini yitirmektedir. Bu duruma bir çözüm olarak, bellek sistemi olan Uzun Kısa Vadeli Bellek (LSTM) önerilmiştir [23].



Şekil 2.13. LSTM hücresi

LSTM hücre yapısında kapılar (gate) bulunmaktadır (Şekil 2.13). Bilgilerin bu kapılardan sigmoid fonksiyonu ile geçerek unutulup unutulmayacağına, güncellenme durumuna, hafızaya aktarılma durumlarına dair kararlar verilir. Şekil 2.14’te kapı örneği gösterilmiştir.



Şekil 2.14. LSTM hücresinde bulunan kapı yapısı

Bilgi ilk aşamada sigmoid fonksiyonundan geçer. Çıktı 0 değerine yakınsa bilgi unutulur, 1 değerine yakınsa sonraki hücreye aktarılır. Bu aşamada hesaplanan değer unutma kapısı kuyruğu (forget gate tail , f_t) olarak adlandırılır. (Eşitlik 2.10).

$$f_t = \sigma(w_f [h_{t-1}, x_t] + b_f) \quad (2.10)$$

Burada h_{t-1} önceki LSTM hücresini, x_t girdi bilgisini, b_f ise yan (bias) terimini ifade eder. Girdi kapısı katmanında (input gate layer, i_t) güncellenecek değerler tespit edilir. Hiperbolik tanjant katmanında (hyperbolic tangent layer, \tilde{C}_t) yeni aday değerlere karar verilerek bilgi vektörü oluşturulur (Eşitlik 2.11 ve Eşitlik 2.12).

$$i_t = \sigma(w_i [h_{t-1}, x_t] + b_i) \quad (2.11)$$

$$\tilde{C}_t = \tanh(w_c [h_{t-1}, x_t] + b_c) \quad (2.12)$$

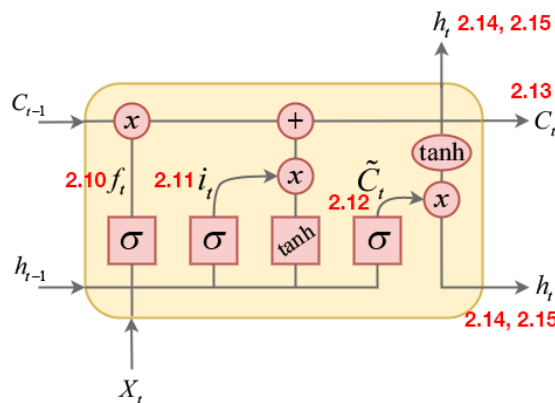
Eski hücre değerleri C_{t-1} , f_t unutmaya kapısı ile çarpılarak eski hücrede unutulacak değerler hesaplanır. \tilde{C}_t ile aday değerler girdi kapısı katmanı i_t ile çarpılır ve güncel hücreye eklenecek değerler belirlenir (Eşitlik 2.13).

$$C_t = f_t * C_{t-1} + i_t * \tilde{C}_t \quad (2.13)$$

Sigmoid fonksiyon, hücre yapısının hangi parçalarının çıktıya yönlendirileceğine karar verir. Daha sonra bu hücre yapısı hiperbolik tanjant fonksiyonuna yönlendirilerek çıktılar $[-1, 1]$ aralığına dönüştürülür. Bu değerler o_t çıktısı ile çarpılır. Böylece çıktıya yönlendirilecek değerler hesaplanmış olur (Eşitlik 2.14 ve Eşitlik 2.15). Bu eşitliklere sahip LSTM yapısı Şekil 2.15'te gösterilmiştir.

$$o_t = \sigma(w_o [h_{t-1}, x_t] + b_o) \quad (2.14)$$

$$h_t = o_t * \tanh(C_t) \quad (2.15)$$



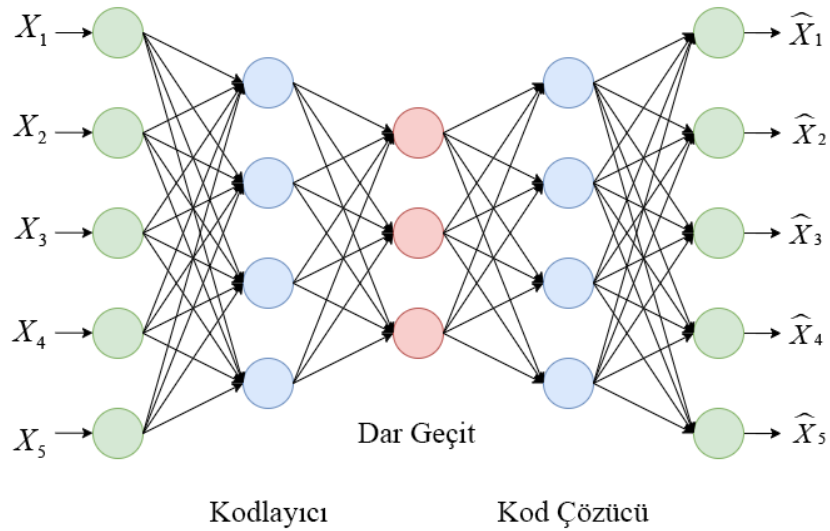
Şekil 2.15. Eşitliklerin LSTM hücresinde gösterilmesi

2.4. Otokodlayıcı

Yapay sinir ağlarının özel bir çeşidi olan otokodlayıcılar, girdilerini çıktı katmanında minimum kayıp ile yeniden inşa etmek üzere tasarlanmıştır. Girdi katmanındaki nöron sayısı ile çıktı katmanındaki nöron sayıları birbirine eşittir. Bu özelliklerinden dolayı mimarileri belirli karakteristik yapıya sahiptir. Mimari yapıları çok katmanlı algılayıcılara çok benzemektedir [19, 20].

Otokodlayıcılar temel olarak iki aşamaya sahiptir. İlk aşama kodlayıcı (encoder) olarak adlandırılır. Burada girdilerin alt öznelikleri çıkarılarak boyutu indirgenir. Bu boyut indirgeme işlemi, girdi katmanından sonraki her katmanda, bir önceki katmana göre daha az sayıda nöron içeren gizli katmanların art arda gelmesi ile sağlanır. Kodlayıcı katmanların bittiği katman, yani en küçük boyutlu katmana dar boğaz veya dar geçit katmanı (bottleneck layer) denir.

Otokodlayıcıların ikinci aşamasına kod çözücü (decoder) denir. Bu kısımda boyutu indirgenen girdi, boyut artırma işlemine girerek kendi boyutlarına gelene kadar yeniden inşa edilir. Burada kodlayıcı mimarisi ters olarak dar boğaz katmanı takip eder (Şekil 2.16).



Şekil 2.16. Örnek bir otokodlayıcı mimarisi

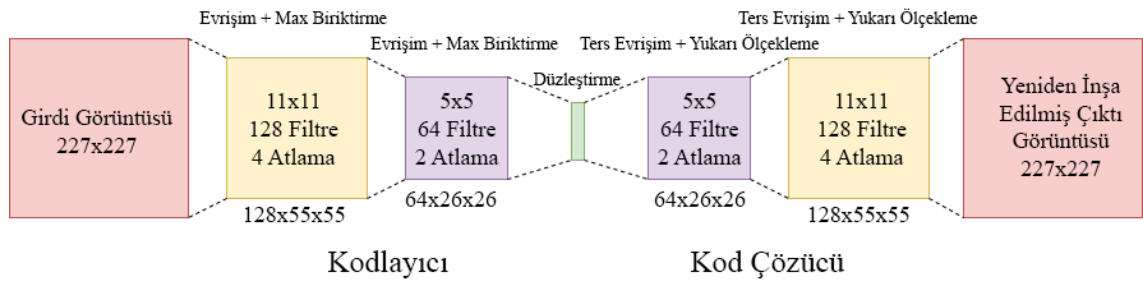
Otokodlayıcılar girdileri birebir kopyalamamayı öğrenirler, aksi halde kullanışlı olmazlar. Ağın eğitim verilerine benzeyen verileri yaklaşık olarak yeniden inşa ederler. Eğitim sürecindeki amaç, minimum kayıp ile girdiyi yeniden inşa etmektir. Böylece model girdininin hangi özneliklerini seçeceğini önceliklendirmeye zorlanır.

Otokodlayıcıların boyut indirgeme özellikleri temel bileşenler analizine benzemektedir [24]. Eğer otokodlayıcı sadece doğrusal aktivasyon fonksiyonları içeriyorsa ve hata metriği hata kareler ortalaması ise bu otokodlayıcı artık yaklaşık olarak temel bileşenler analizi yapar.

Otokodlayıcılar, sinir ağlarının özel bir çeşidi olduğu gibi, otokodlayıcıların da özel çeşitleri vardır.

2.4.1. Evrişimsel Otokodlayıcı

Görüntü verileri işlenirken, basit yapıli otokodlayıcılar çok küçük boyutlu görüntüler dışında iyi performans göstermezler. Bunun için evrişimsel katmanlara sahip otokodlayıcılar önerilmiştir [25]. Evrişimsel Otokodlayıcı (CAE) yapısının kodlayıcı aşamasında evrişim ve biriktirme katmanları yer alır. Görüntülerin öznelik haritaları çıkarılır. Dar boğaz katmanının boyutu kadar bu özneliklerin boyutu indirgenmiş olur. Kod çözücü aşamasında ise öznelik haritaları kullanılarak ters evrişim (deconvolution) ve yukarı ölçekleme (upsampling) katmanları ile işlem yapılır. Girdide kullanılan görüntü, CAE mimarisi sonunda yeniden inşa edilmiş olur (Şekil 2.17).

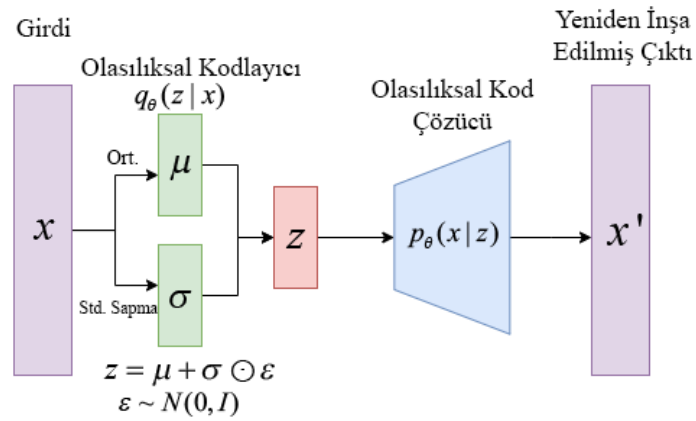


Şekil 2.17. Evrişimsel otokodlayıcı mimarisi

Kodlayıcı aşamasında bulunan evrişim katmanlarının LSTM evrişim katmanına sahip olmasıyla birlikte CAE artık Evrişimsel Uzun Kısa Vadeli Bellekli Otokodlayıcı mimarisi olur.

2.4.2. Değişimsel Otokodlayıcı

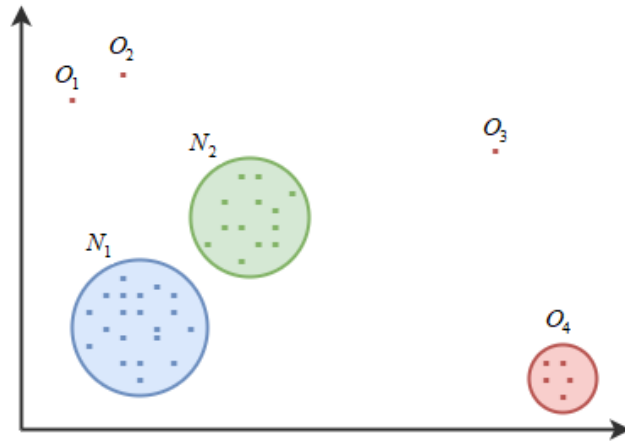
Değişimsel otokodlayıcı (VAE), olasılıksal ve üretken (generative) otokodlayıcılardır. Değişimsel Bayes çıkarımı gerçekleştirirler [26]. Girdinin olasılıksal kodlayıcı aşamasında μ ve σ parametreleri üretilerek normal dağılımdan rasgele örnekleme yapılarak boyutu indirgenir. Olasılıksal kod çözücü ise bu örnekleme çözümlerini çözerek girdileri yeniden inşa eder (Şekil 2.18).



Şekil 2.18. Değişimsel otokodlayıcı mimarisi

3. ANOMALİ TESPİTİ

Bir veri kümesinde, verinin örüntüsüne uymayan, verinin geneline göre sapması olan verilere anomali denir. Anomaliler veri madenciliği ve istatistik biliminde aykırı değer, sapma, anormallik olarak da bilinirler. Anomaliler yapıları gereği nadirlerdir. Verideki hatalar tarafından oluşabileceği gibi, bilinmeyen bir işlem tarafından da oluşturulmuş olabilirler. Anomali tespiti ise, verideki beklenmedik örüntülerin bulunması işlemidir. Kullanılacak yöntemler verinin yapısına (görüntü, video gibi), anomalinin tanımına göre değişmektedir. Şekil 3.1'de 2 boyutlu düzlem üzerinde anomaliler gösterilmiştir. Burada N_1 ve N_2 normal veri kümelerini, O_1, O_2, O_3 anomali noktaları, O_4 ise anomali kümeyi göstermektedir.



Şekil 3.1. İki boyutlu düzlemde anomalilerin gösterimi

Anomali tespitlerini 3 başlıkta özetlemek mümkündür. Bunlar istatistiksel yöntemler, makine öğrenmesi temelli yöntemler ve derin öğrenme temelli yöntemlerdir.

3.1. İstatistiksel Yöntemler

İlk geliştirilen yöntemlerdendir. Genellikle tek boyutlu, gerçel ve sayısal değerler için kullanılmıştır. Bu durum ilgili yöntemlerin problemler üzerine uygulanabilirliğini kısıtladığı gibi, karmaşık ve kompleks veri yapıları için işlem sürelerini uzatmaktadır [27, 28]. Bu yöntemler verinin θ parametrelili ve $f(x;\theta)$ olasılık yoğunluk

fonksiyonuna sahip olduğunu varsayar. Her x gözlemi için anomali skoru $f^{-1}(x; \theta)$ olasılık yoğunluk fonksiyonu ile hesaplanır. θ parametresi ise veriden tahmin edilir.

İstatistiksel yöntemlerden bir diğeri istatistiksel hipotez testleridir [29]. Yokluk hipotezi, x gözleminin θ parametrelili bir dağılımdan geldiği şeklinde kurulur. Eğer H_0 yokluk hipotezi reddedilirse x gözlemi anomali kabul edilir.

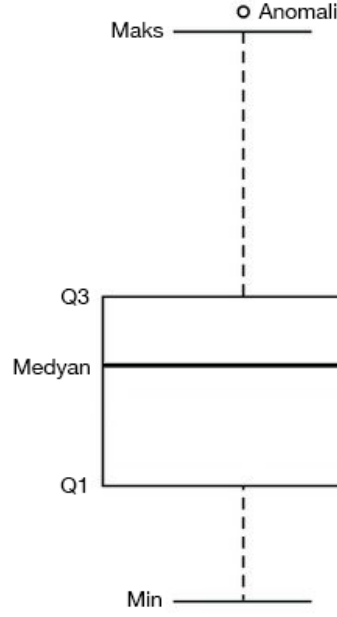
Bir diğeri yöntem ise Grubbs testidir. Tek değişkenli (univariate) veri için tek anomali (aykırı değer) tespiti yapar. Grubbs testi Eşitlik 3.1’de verilmiştir.

$$z > \frac{N-1}{N} \sqrt{\frac{t_{\alpha/2N; N-2}^2}{N-2 + t_{\alpha/2N; N-2}^2}} \quad (3.1)$$

ise x değeri anomalidir. Burada $z = \frac{|x - \bar{x}|}{S}$ ‘dir.

Zaman serisi problemlerinde anomali tespiti için regresyon temelli yöntemler de kullanılmıştır [30]. Anomali tespiti bu durumda iki aşamada gerçekleşmektedir. İlk aşamada veri modellenir. İkinci aşamada artıklar hesaplanır. Bu artıklar, ilgili noktalar için anomali skorlarıdır. Anomali olup olmama durumu aynı zamanda istatistiksel güven aralığı veya hipotez testine göre de karar verilebilir.

En çok kullanılan ve kullanması en kolay yöntemlerden biri de kutu grafiği (box plot) yöntemidir. Verilerin medyan, ortanca, dördlük gibi değerlerine göre çizilen kutu grafiğinde, alt ve üst sınırın dışında kalan noktalar kolaylıkla tespit edilebilmektedir. Kutu grafiği örneği Şekil 3.2’de verilmiştir.



Şekil 3.2. Tek değişkenli veri için kutu grafiği örneği

3.2. Makine Öğrenmesi Temelli Yöntemler

Anomali tespitinde kullanılan makine öğrenmesi temelli yöntemlerin aşamaları genellikle benzerlik göstermektedir. Parametrelendirme, eğitim ve anomali tespiti aşaması olmak üzere üç temel aşamaya sahiptirler. Parametrelendirme işleminde veri toplanır, toplanan veri modellenmelidir. Eğitim aşaması manuel veya otomatik yöntemlerin belirlenmesi ve kullanılmasını içermektedir. Anomali tespiti aşamasında ise belirli bir eşik değeri belirlenir. Bu eşik değeri, verinin anomali olup olmasında bir karar değeridir [31]. Bu yöntemler öğrenme çeşidine göre ikiye ayrılabilir.

3.2.1. Gözetimli Makine Öğrenmesi Yöntemleri

Sınıflandırma yöntemleri olarak da bilinen bu yöntemler etiketlenmiş (label) veriye ihtiyaç duymaktadır. Bu etiketler hem normal veriler hem de anomali veriler için ayrı olarak tanımlanmalıdır. Teorik olarak etiketli oldukları için yarı gözetimli veya gözetimsiz anomali tespiti yöntemlerinden daha iyi performans gösterirler [32]. Ancak etiketlenmiş veriye ihtiyaç duymamanın da dezavantajları vardır. Örneğin etiketleme

işlemi genelde hata (pürüz) içermektedir ve bu durum ağın yanlış eğitimine sebep olur. Dolayısı ile anomali tespiti işleminde doğruluk (accuracy) düşmektedir.

Kullanılan yöntemlerden bazıları K. En Yakın Komşuluk (KNN), Yapay Sinir Ağları, Destek Vektör Makineleri (SVM), Karar Ağaçları olarak verilebilir.

3.2.2. Gözetimsiz Makine Öğrenmesi Yöntemleri

Bu yöntemler genelde kümeleme yöntemlerini içermektedirler. Kümeleme işlemi veri noktaları arası uzaklıklara göre yapıldığı için benzer veriler aynı kümede, anomali veriler ise daha uzak mesafede olacakları için ayrı küme veya küme dışında bulunacaklardır. Bu yöntemler etiketlenmiş veriye veya eğitim verisine ihtiyaç duymazlar. Veride büyük çoğunluk normal, çok az bir bölümün anomali içerdiği ve istatistiksel olarak farklı olduğu varsayılır.

Anomali tespiti için kullanılan gözetimsiz makine öğrenmesi yöntemlerinden bazıları K-Ortalamlar, Kendini Düzenleyen Haritalar (SOM), Bulanık C-Ortalamlar (FCM), Tek Sınıflı Destek Vektör Makineleri olarak verilebilir.

3.3. Derin Öğrenme Temelli Yöntemler

Anomali tespiti problemlerinde derin öğrenme temelli yöntemlere duyulan ihtiyaç veriden veya verinin yapısından kaynaklanmaktadır. Bu veriler ardışık (sequential) veriler ve ardışık olmayan (non-sequential) veriler olarak ikiye ayrılır. Ardışık verilere örnek olarak ses, müzik, zaman serileri, protein sekansları verilebilir. Ardışık olmayan verilere örnek ise görüntü, video gibi çok boyutlu veriler verilebilir. Derin öğrenme temelli yöntemler çok boyutlu verilerin daha derin özneliklerini öğrenmede kullanılırlar. Mimari derinleştikçe yüksek boyutlu verilerde derin öğrenme yöntemleri daha başarılıdır [33].

3.3.1. Gözetimli Derin Öğrenme Yöntemleri

Makine öğrenmesi yöntemlerinde olduğu gibi gözetimli derin öğrenme yöntemlerinde de etiketli veriye ihtiyaç duyulmaktadır. Ancak çok boyutlu verilerde etiketleme işlemi diğer verilere göre çok daha masraflı olduğundan pratikte bu öğrenme yöntemi diğer yöntemlere göre daha az tercih edilmektedir. Bu yöntemlerde kullanılan mimarilerin performansları veri setinin anomali dengesine bağlıdır.

3.3.2. Yarı Gözetimli Derin Öğrenme Yöntemleri

Veri setlerinde normal etikete sahip verilerin elde edilmesi, anomali etiketli verilerin elde edilmesine göre, anomalilerin yapısı gereği daha kolaydır. Bunun bir sonucu olarak yarı gözetimli yöntemler daha çok tercih edilir.

Bu yöntemlerde genelde veri setinin küçük bir kısmı etiketli, çoğunluğu ise etiketsizdir. Etiketli veriler kullanılarak etiketsiz verilerde daha başarılı sonuçlar elde edilmesi amaçlanır. Örneğin sık kullanılan yöntemlerden biri, normal veriler ile otokodlayıcılar küçük yeniden inşa hatası ile eğitilir. Yüksek yeniden inşa hatası ise anomali skorunu belirler. Belirlenen eşik değeri ise anomali olup olmama durumunu tespit eder. Böylece tüm verinin etiketli olmasına ihtiyaç duyulmaz.

3.3.3. Gözetimsiz Derin Öğrenme Yöntemleri

Gözetimsiz yöntemler anomalileri sadece veri setinin yapısına göre belirler. Bu yöntemler etiketsiz verileri otomatik olarak etiketlemek için kullanılırlar [34]. Sağlık ve siber güvenlik başta olmak üzere kullanılan yöntemlerden bazıları Derin İnanç Ağları (DBN), Derin Boltzman Makineleri (DBM), Özyinelemeli Sinir Ağı (RNN), Uzun Kısa Vadeli Bellek tabanlı yöntemlerdir.

3.4. Anomali Çeşitleri

Anomaliler görülme tiplerine göre üçe ayrılmaktadır [28]. Bunlardan ilki noktasal anomalilerdir. Literatürde yapılan çalışmaların büyük çoğunluğunu oluştururlar. Veride rasgele olarak görülen düzensizliklerdir. Örneğin bir kredi kartı kullanıcısının yaptığı

harcamalarda, diğ er tutarlara göre çok yüksek kalan bir harcama tutarı noktasal anomalidir.

Bağ lamsal anomaliler, verinin davranış sal özelliklerinin (verinin örüntüsünün) dışında kalan deę erlerdir. Genelde zamana ve uzaya baę lıdır. Örneğ in kredi kartı kullanıcısının para harcama alış kanlıkları dışında yaptıę ı harcamalardır.

Grupsal anomaliler, belirli bir deę er alan birden fazla sayıda anomalinin oluşturduę u gruplar olarak tanımlanır. Verideki alış kanlığ ın dışındadır. Örneğ in kredi kartı kullanıcısının kartından üst üste yapılmış belirli miktarda harcama bu anomali çeş idine girer.

Anomaliler yapılan tanım dışında, probleme ve tanımına göre de deę iş mektedir. Örneğ in veri setinin akciğ er grafileri olduę u durumda anomaliler kitleler olabilir. Mobese kameralarının veri seti olduę u durumda insan davranış ları, olaylar, kazalar anomaliler olabilir. Her seferinde tek tip anomali tespit edilebilmektedir.

3.5. Videolarda Anomali Tespitleri

Video verileri çok boyutlu ve büyük hacimli oldukları için klasik yöntemler ile anomali tespiti yapılamamaktadır. Bunun için derin öğrenme temelli yöntemlere ihtiyaç duyulmaktadır [35]. Örneğ in 500x600 piksel boyutlarına sahip, 10 saniyelik ve 30 fps (saniye baş ına kare) hızına sahip bir videodaki veri sayısı aşağıdaki gibi olmaktadır.

$$500 * 600 = 300.000 \text{ (her bir karedeki veri sayısı)}$$

$$300.000 * 30 = 9.000.000 \text{ (her bir saniyedeki veri sayısı)}$$

$$9.000.000 * 10 = 90.000.000 \text{ (videodaki toplam veri sayısı)}$$

Oldukça yüksek boyuttaki bu veriler için mimari seç imi, eğ itimin yapılacaę ı bilgisayarın donanım seviyesi, eğ itim süresi gibi ciddi zorluklar ve masraflar vardır.

Videolardaki anomalilerin tespiti, durağan arka plana sahip görüntülerdeki optik hareketlerdeki sapmaların tespitidir. Bu sapmalar, öznelik çıkarımı ve örüntü tanıma ile sağlanır [35].

3.5.1. Yeniden İnşa Modelleri (Reconstructive Models)

Temel bileşenler analizi ve otokodlayıcılar videolardaki normal davranışların modellenmesinde kullanılırlar. Bu modellerdeki anomaliler ise kötü bir şekilde yeniden inşa edilmiş verilerdir. Girdi görüntüsü ile yeniden inşa edilmiş görüntü arasındaki fark yeniden inşa hatasını, bu hata da anomali skorunu vermektedir.

3.5.2. Tahminsel Modelleme (Predictive Modeling)

Videolardaki her bir kare zaman serisi verisi olarak kabul edilir. Bu yöntemlerdeki asıl amaç, bu karelerin koşullu olasılık dağılımının (Eşitlik 3.2) çıkarılmasıdır.

$$P(X_t | X_{t-1}, X_{t-2}, \dots, X_{t-p}) \quad (3.2)$$

Bu dağılım ile geçmiş p adet kare kullanılarak t anındaki kare yeniden tahmin edilir. Kullanılan yöntemler otoregresif modeller, Evrişimsel LSTM (ConvLSTM) modelleridir.

3.5.3. Üretici Modeller (Generative Models)

Derin öğrenme yöntemi ile normal video örnekleminin olabirliğinin modellendiği yöntemlerdir. Bu modeller ile sınıf (etiket) olasılıkları hesaplanır. Kullanılan yöntemler Değişimsel Otokodlayıcı, Çekişmeli Üretici Ağlar, Çekişmeli Otokodlayıcılar (AAE)'dir.

3.5.4. Anomali Skorları ve Eşik Değerleri

Yeniden inşa modellerinde anomali skorları yeniden inşa hatası ile tanımlanmaktadır. Bu hata, girdi olarak verilen kare ile yeniden inşa edilen kare arasındaki fark bulunarak hesaplanmaktadır (Eşitlik 3.3).

$$E_t = |X_t - \hat{X}_t| \quad (3.3)$$

Bu yeniden inşa hatası daha sonra $[0, 1]$ aralığına ölçeklenir. Elde edilen Eşitlik 3.4, anomali skoru olarak tanımlanmaktadır.

$$s(t) = \frac{\sum_{(x,y)} E_t - \min_{(x,y)}(E_t)}{\max_{(x,y)}(E_t)} \quad (3.4)$$

Bu anomali skoru 1' den çıkartıldığında artık düzen skoru elde edilmektedir (Eşitlik 3.5). Bu skor aynı zamanda zamansal düzen skoru (temporal regularity score) olarak da bilinmektedir. Burada x ve y uzaysal indislerdir (spatial indices).

$$sr(t) = 1 - s(t) \quad (3.5)$$

Düzen skoru değeri veya yeniden inşa hatası değerleri belirli bir eşik değeri ile koşullandırılır. Eğer eşik değerinden büyükse anomali, küçükse anomali değildir, şeklinde yorum yapılır. Eşik değeri objektif bir değerdir. Deneme yanılma ile, manuel kontrol ile en iyi değer saptanmaya çalışılır. Eşik değeri belirlemenin belirli bir yöntemi yoktur. Literatürde yapılan bir çalışmada [36], normal veriler ile yapılan eğitim sırasındaki en yüksek yeniden inşa hatası eşik değeri kabul edilmiştir. Bunun yanı sıra dinamik eşik değerleri üzerine de çalışmalar yapılmaktadır [37, 38].

4. UYGULAMA

4.1. Veri Setleri ve Problem Tanımı

Tez çalışmasının uygulama bölümünde ham veri yaklaşımı ve optik akış yaklaşımlarının performanslarını ölçmek amacıyla üç adet veri seti kullanılmıştır: UCSD Pedestrian 1, UCSD Pedestrian 2 [39] ve CUHK Avenue [3]. Bu veri setleri literatürde yaygın olarak kullanılan performans testi (benchmark) veri setleridir [3]. Her veri seti eğitim ve test kümeleri olarak kendi içlerinde ikiye ayrılmışlardır.

UCSD Pedestrian 1 ve UCSD Pedestrian 2 veri seti, bir parka belirli bir açıdan yerleştirilmiş güvenlik kamerasının görüntülerinden oluşmaktadır. Eğitim kümesinde sadece yayaların yürüyüşleri varken, test kümesinde ise parktan geçen bisikletliler, kaykaylılar ve araçlar görülmektedir. Sadece test kümesinde görünen bu araçlar anomali olarak kabul edilmiştir (Şekil 4.1 ve Şekil 4.2). UCSD Pedestrian 1 eğitim kümesinde 34, test kümesinde 36 video klip varken, UCSD Pedestrian 2 eğitim kümesinde 16, test kümesinde 12 video klip vardır.



Şekil 4.1. UCSD Ped 1 veri seti test kümesi anomali örnekleri



Şekil 4.2. UCSD Ped 2 veri seti test kümesi anomali örnekleri

CUHK Avenue veri seti ise, Hong Kong Üniversitesinde kampüste belirli bir yere yerleştirilmiş olan kameradan elde edilen görüntülerden oluşmaktadır. Eğitim kümesi öğrencilerin dersliklere yürüyüşlerini içerirken, test kümesi ise ters yönde yürüyen, koşan ve nesne fırlatan öğrencilerin görüntülerini içermektedir (Şekil 4.3). Bu alışılmadık örüntüler anomaliler olarak kabul edilmiştir. Eğitim kümesi 16, test kümesi 21 video klip içermektedir.



Şekil 4.3. CUHK Avenue veri seti test kümesi anomali örnekleri

İlgili üç veri setinin klip sayıları, video kliplerinin kare sayısı ve anomali içeren kare sayıları Çizelge 4.1’de özetlenmiştir.

Çizelge 4.1. Veri setleri için özet bilgiler

	Eğitim Kümesi		Test Kümesi			Kare Çözünürlüğü
	Video Sayısı	Toplam Kare Sayısı	Video Sayısı	Anomali İçeren Kare Sayısı	Toplam Kare Sayısı	
UCSD Pedestrian 1	34	6800	36	4045	7200	238x158
UCSD Pedestrian 2	16	2550	12	1648	2010	360x240
CUHK Avenue	16	15270	21	3867	15324	640x360

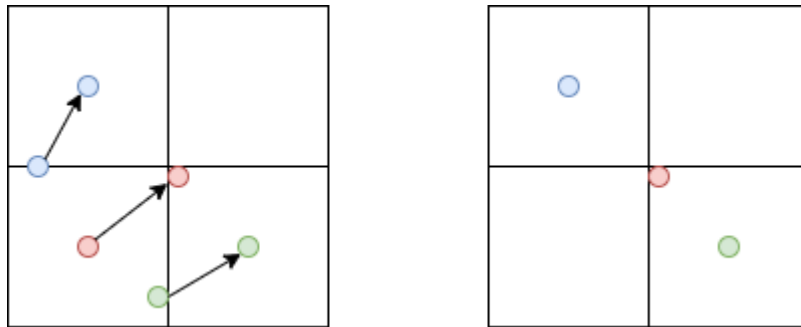
4.2. Kullanılan Yöntemler ve Mimariler

İlgili video karelerine analiz aşaması için belirli ön işlemler uygulanmıştır. İlk olarak her üç veri seti için görüntüler 256x256 piksel çözünürlüğe ölçeklenmiştir. Daha sonra her kare gri tonlama ile 3 renk kanallı görüntülerden tek kanallı görüntüye ölçeklenmiştir (Bkz. Şekil 1.2).

Son olarak verilerin zamansal yapıda hazırlanması gerekmektedir. Video verileri zamana bağlı olduğundan, ilk t tane video karesi ile $(t+1)$. video karesi tahmin edilmeye çalışılacaktır. Literatürde yapılmış bir çalışmada [40] ilk 10 kare ile 11. karenin tahmininin en iyi performansa sahip sonuçları vermiş olduğu ispatlanmıştır. Bu sebeple çalışmamızda $t=10$ alınmıştır. Dolayısı ile her veri seti için eğitim kümelerindeki kareler 10'ar parçalık bölütlere bölünmüştür.

4.2.1. Optik Akış Yaklaşımı

Optik akış, iki görüntü karesi arasındaki farktan hesaplanan hareket hızının dağılımıdır. Optik akış, parlaklık değerine sahip piksel değerlerinin t ve $t+\Delta t$ anı arasındaki hareket bilgisini içeren iki boyutlu vektörden oluşmaktadır (Şekil 4.4). Piksel değerinin, yani parlaklığın zamana bağlı olarak değişmediği varsayımına sahiptir. Video verilerinin alt özniteliklerindedir.



Şekil 4.4. Optik akış ile hareketin gösterimi

t . zamandaki x ve y koordinatlarındaki parlaklık değeri $I(x, y, t)$ olarak fonksiyonel şekilde gösterilebilir. Δt zaman sonrasındaki parlaklık değeri, ilgili piksel için değişmeyeceği için Eşitlik 4.1 elde edilir.

$$I(x, y, t) = I(x + \Delta x, y + \Delta y, t + \Delta t) \quad (4.1)$$

Hareket şiddeti hesaplanırken uzaysal ve zamansal değişimlerin tespiti için diferansiyel yöntemlerden yararlanır. Hareketin küçük olduğu varsayımı ile görüntü sinyallerinin Taylor açılımları alınırsa Eşitlik 4.2 elde edilir.

$$I(x + \Delta x, y + \Delta y, t + \Delta t) = I(x, y, t) + \frac{\partial I}{\partial x} \Delta x + \frac{\partial I}{\partial y} \Delta y + \frac{\partial I}{\partial t} \Delta t \quad (4.2)$$

x, y ve t piksel koordinatlarını, $\Delta x, \Delta y, \Delta t$ bu koordinatlar üzerindeki değişimi ifade etmek üzere Eşitlik 4.2 artık Eşitlik 4.3' deki gibi olur.

$$\frac{\partial I}{\partial x} \Delta x + \frac{\partial I}{\partial y} \Delta y + \frac{\partial I}{\partial t} \Delta t = 0 \quad (4.3)$$

Bu eşitlik kullanılarak artık hareket şiddeti hesaplanabilir. $\Delta x, \Delta y, \Delta t$ değişimlerinin t zamanına göre türevleri alınırsa Eşitlik 4.4 elde edilir.

$$\frac{\partial I}{\partial x} V_x + \frac{\partial I}{\partial y} V_y + \frac{\partial I}{\partial t} = 0 \quad (4.4)$$

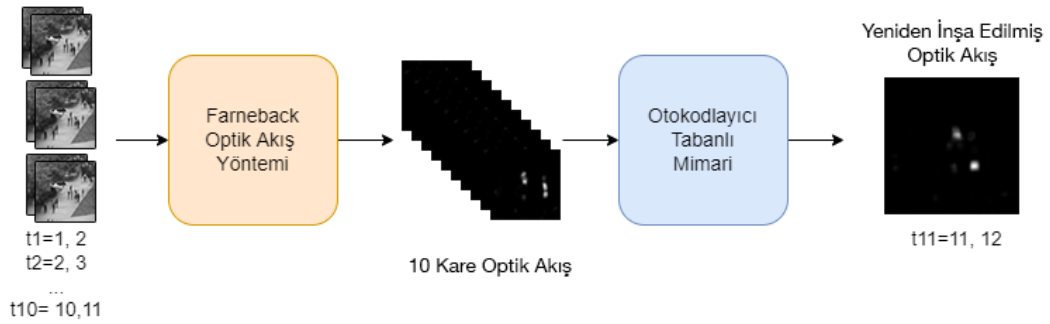
Burada $V_x = \frac{dx}{dt}$ ve $V_y = \frac{dy}{dt}$ olarak hesaplanmaktadır. Elde edilen son eşitlikte artık V_x ve V_y çözümleri optik akış bilgisini vermektedir. Tek denklem ve iki bilinmeyen içeren bu denklem sistemi normal yöntemler ile çözülememektedir. Bu bilinmeyenler için literatürde birçok yöntem geliştirilmiştir. Bunlardan bazıları Lucas-Kanade [41], Horn-Schunck [42] ve Farneback (Dense Optical Flow) [43] yöntemleridir. Lucas-Kanade yöntemi köşelerde bulunan hareketleri tespit ederken, Farneback yöntemi tüm piksellerde bulunan hareketi tespit eder. Horn-Schunck yöntemi ise optik akış hesaplamada sıklıkla karşılaşılan “Hareket Algısı (Motion Perception – Aperture Problem)” problemini çözmek için geliştirilmiştir.

Tüm pikseller üzerindeki hareketi tespit edebildiği için çalışmamızda Farneback optik akış yöntemi kullanılmıştır. Bu yöntem her piksel için, komşu piksellerin polinom açılımlarını kullanarak (Eşitlik 4.5) aradaki fark değerine göre sinyal oluşturur (Eşitlik 4.6).

$$f_1(x) = x^T A_1 x + b_1^T x + c_1 \quad (4.5)$$

$$f_2(x) = f_1(x-d) = (x-d)^T A_1 (x-d) + b_1^T (x-d) + c_1 \quad (4.6)$$

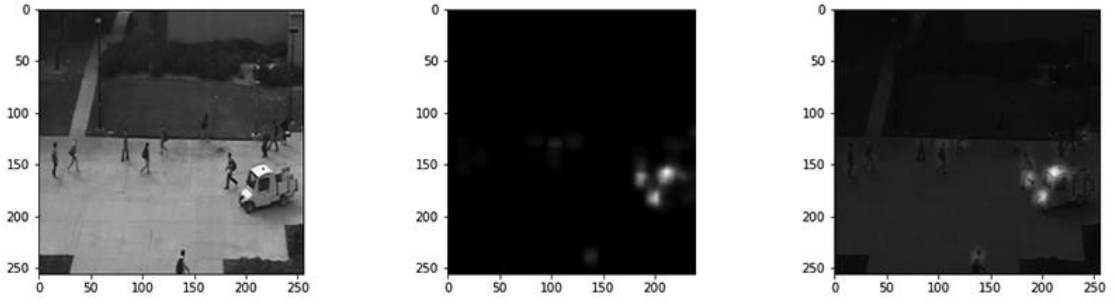
Farneback yöntemi kullanılarak oluşturulan optik akış temelli mimariler Şekil 4.5'te verilmiştir.



Şekil 4.5. Optik akış yaklaşımı

Optik akış yaklaşımı ile videolarda anomali tespitinde, harekete bağlı anomalileri bulmak amaçlanmaktadır. Eğitim ve test kümesindeki görüntülerin optik akış öznelikleri çıkarılarak, ilk 10 karede bulunan hareketlerden 11. karenin hareketi tahmin edilir. Yaya hareketlerinin olduğu veri setlerinden çıkacak optik akış sonucu yayaların hızları, hareketi, hareket yönü ve yoğunluğu tespit edilir. Yayalardan hızlı olan bir araç veya cisim geçtiğinde bu durum anomali olarak tespit edilecektir.

Optik akış yönteminin güçlü olduğu noktalardan biri, durağan arka plana sahip video görüntülerindeki arka plan hareket etmediği için, bu fazla bilgiyi veriden çıkarmasıdır (Şekil 4.6). Böylece geriye sadece hareket bilgilerini içeren veriler kalmaktadır. Optik akış sonucu ortaya çıkan tüm veri anlamlıdır.

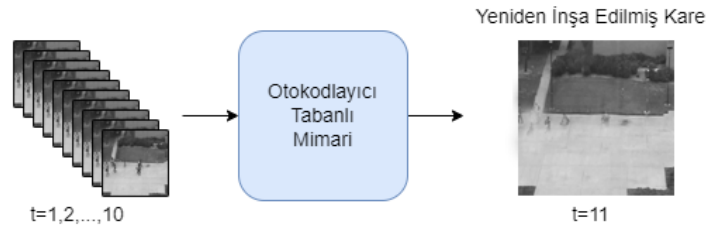


Şekil 4.6. Optik akış örneği

Bu mimarilerin sonucunda yeniden inşa edilen 11. karenin optik akışının, gerçek optik akış ile farkından yeniden inşa hatası hesaplanır. Eğitim aşamasında amaç, bu hataları minimize edecek şekilde ağı eğitmektir. Test kümesi, eğitim aşamasında olmayan hız, hareket ve hareket örüntüleri içerdiği için eğitilen ağ, test aşamasında yüksek yeniden inşa hatası oluşturacaktır. Anomali tespiti işlemi ise bu yeniden inşa hatalarından türetilen düzen skorundan yapılmaktadır.

4.2.2. Ham Veri Yaklaşımı

Literatürde en sık kullanılan yaklaşımdır. Bu yaklaşımda, video görüntüleri doğrudan otokodlayıcı tabanlı mimariye girdi olarak verilmektedir. Otokodlayıcı tarafından ilk 10 karede zamana bağlı meydana gelen öznelik değişimleri öğrenilerek, 11. kare yeniden oluşturulmaya çalışılmaktadır (Şekil 4.7).



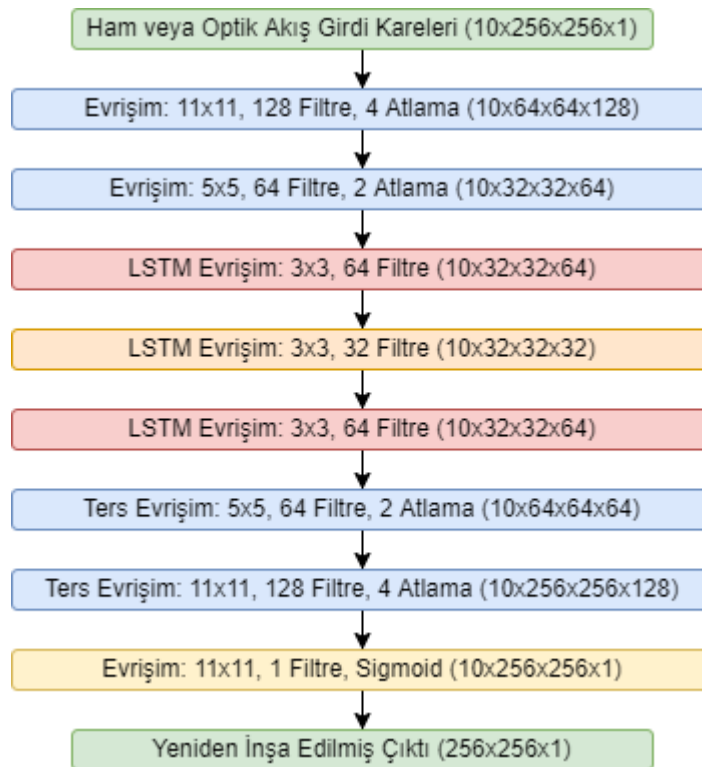
Şekil 4.7. Ham veri yaklaşımı

Optik akış yöntemine kıyasla, bu yaklaşımların güçsüz yönleri, statik arka plan bilgilerini de işlerler, bu yüzden daha çok belleğe ihtiyaç duyarlar. Gri ölçekli bile olsalar her bir piksel değerinde meydana gelen renk değişimlerine duyarlıdır.

Optik akış yaklaşımında olduğu gibi yeniden inşa edilen 11. kare ile gerçek 11. kare arasındaki fark yeniden inşa hatasını verecektir. Bu hatalardan düzen skoru elde edilerek anomali tespiti işlemi yapılacaktır.

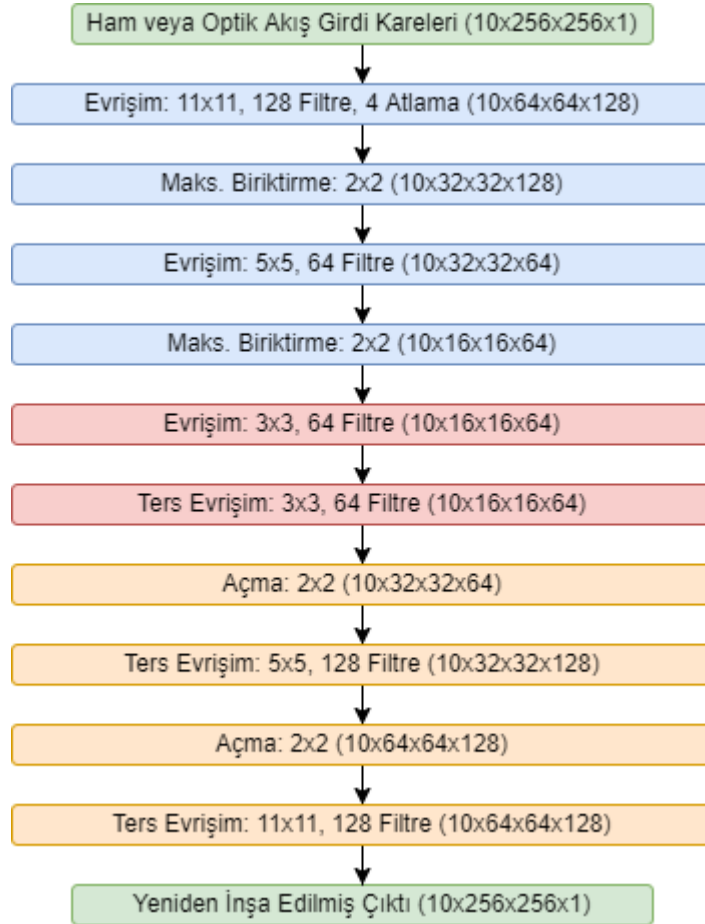
4.2.3. Otokodlayıcı Tabanlı Mimariler

Yapılan birçok çalışmada, karşılaştırma amacı ile genelde üç farklı otokodlayıcı kullanılmaktadır. Bunlar CAE, ConvLSTMAE ve VAE'dir [4, 15, 44-49]. Çalışmamızda görüntü ve optik akış verileri her veri seti için ayrı ayrı olmak üzere bu üç mimari ile eğitilmişlerdir.



Şekil 4.8. Kullanılan ConvLSTMAE mimarisi

ConvLSTMAE mimarisi Şekil 4.8’de verilmiştir. 10 karelik görüntü girdi olarak mimariye verilerek, LSTM evrişim katmanlarından geçmektedir. Zamana bağlı özellikleri ile öznitelikleri çıkarılmaktadır. Daha sonra ters evrişim işlemleri ile bu öznitelikler tekrar 10 karelik görüntü olarak inşa edilmektedir. Son aşamada da Sigmoid aktivasyon fonksiyonuna sahip evrişim katmanından 11. kare oluşturulmaktadır.



Şekil 4.9. Kullanılan CAE mimarisi

CAE mimarisinde ise 10 karelik görüntü girdi olarak mimariye girip, evrişim ve max biriktirme katmanlarına girerek boyutu küçültülmekte, öznitelikleri çıkartılmaktadır. Kod çözücü aşamasında ters evrişim ve ters biriktirme (açma / unpooling) işlemi yapılarak tekrar 10 karelik görüntü inşa edilmektedir. Son olarak yine sigmoid evrişim katmanına girerek 11. kare yeniden oluşturulmaktadır (Şekil 4.9).

VAE mimarisinde ise yeniden inşa modeli olasılıksal olarak gerçekleşmektedir. CAE mimarisine olasılıksal katmanların eklenmesi ile oluşturulmuştur. Olasılıksal üretme işlemi kodlayıcı ve kod çözücü aşamalarda standart normal dağılım hataları ile gerçekleşmektedir. VAE mimarilerinde KL kayıp (Kullback-Leibler Divergence Loss, KL Loss) fonksiyonu kullanılmaktadır [50] (Eşitlik 4.7).

$$KL(Q(Z)||P(Z|X)) = \sum_{z \in Z} Q(z|x) \log \frac{Q(z|x)}{P(z|x)} \quad (4.7)$$

Mimarilerin sonunda yeniden inşa edilmiş çıktı ile gerçek veri arasındaki fark ile yeniden inşa hatası hesaplanmış olup, bu hata ile düzen skoru hesaplanmıştır.

4.3. Eğitim ve Test Süreci

Uygulama Python programlama dilinde, Tensorflow ve Keras kütüphaneleri kullanılarak yapılmıştır. Kullanılan bilgisayarın özellikleri; Intel i7 10700K, 64 GB RAM, Nvidia RTX2070s GPU şeklindedir. Ağ eğitimi için faz (epoch) sayısına karar verirken çeşitli denemeler yapılmıştır. 100, 200, 500 faz denenmiş ve AUC skorları karşılaştırılmıştır (Çizelge 4.2).

Çizelge 4.2. 100, 200 ve 500 faz ile eğitilmiş CAE mimarisi için AUC skorları

	UCSD Ped 1	UCSD Ped 2	CUHK Avenue
100 Faz	0.67	0.65	0.80
200 Faz	0.70	0.67	0.83
500 Faz	0.71	0.67	0.83

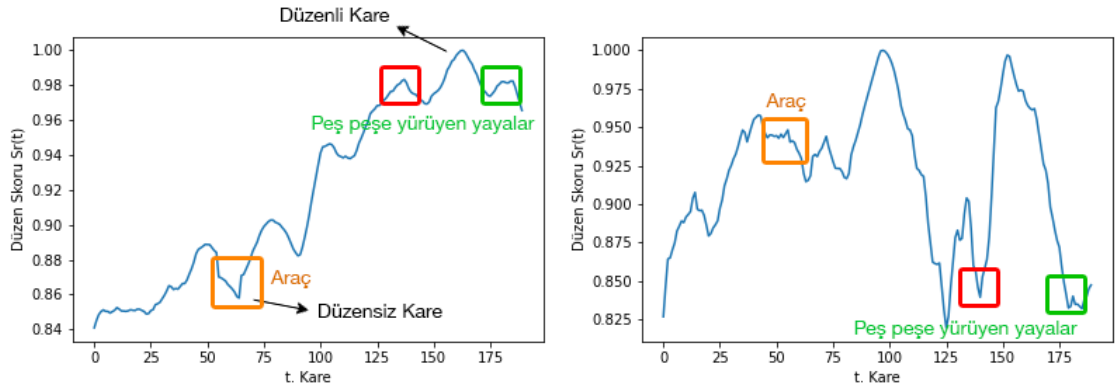
Elde edilen sonuçlar daha önce yapılan çalışmalar [51, 52] ile faz sayısı açısından benzerlik göstermektedir. Eğitim aşamasında kullanılacak faz sayısı 200 olarak belirlenmiştir. Eğitim her iki yaklaşım ile (ham veri ve optik akış), üç mimari ile (CAE, ConvLSTMAE, VAE) ve üç veri setinin eğitim kümesi ile ayrı ayrı yapılmıştır. Toplamda 18 ağ eğitilmiş olup, her ağ kendi veri setinin test kümesi kullanılarak AUC skorları elde edilmiştir.

4.4. Deneysel Sonuçlar

Yapılan analizin sonuçları incelenirken, her veri setinin test kümelerinden belirli bir video klip seçilmiştir. Bu video klipler anomalileri içermeye durumuna ve iki yaklaşım arasında en belirgin farka sahip olan kliplerden seçilmiştir. UCSD Ped 1 veri setinde 19. test klipi, UCSD Ped 2 veri setinde 4. test klipi ve CUHK Avenue veri setinde 20. test klipi sonuç yorumlamada kullanılmışlardır. Sonuçlar yorumlanırken iki yaklaşımın düzen skoru grafikleri, önemli görülen karelerin yeniden inşa edilmiş optik akış görüntüleri kullanılmıştır. Analiz sonunda ise AUC skoru tüm test kliplerinin ortalaması alınarak hesaplanmıştır.

4.4.1. CAE Sonuçları

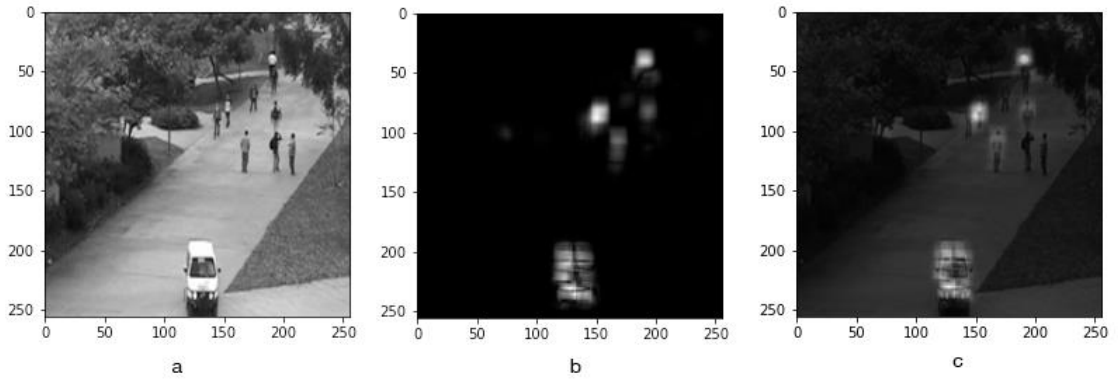
CAE mimarisi ile yapılan analizde ham veri ve optik akış yaklaşımlarının bazı önemli farklılıkları göze çarpmaktadır.



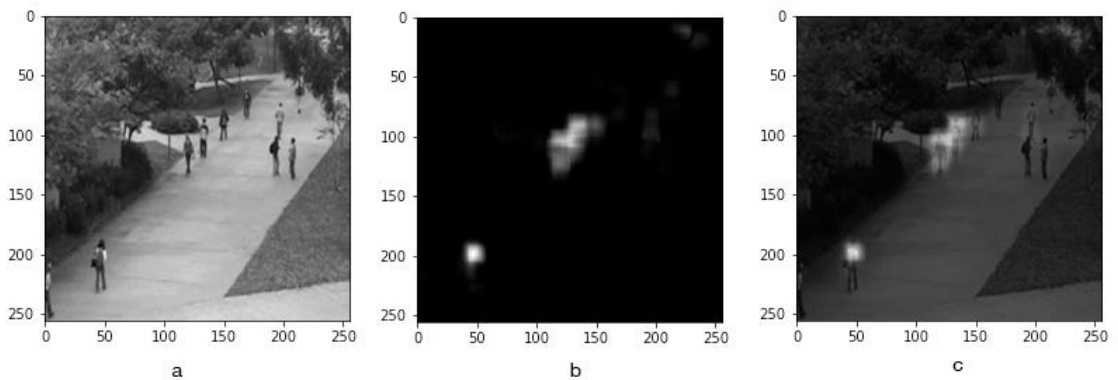
Şekil 4.10. UCSD Ped 1 veri seti CAE - ham veri (sol) ve optik akış (sağ) yaklaşımı düzen skoru grafikleri

Şekil 4.10'dan görüleceği gibi ham veri yaklaşımı ile 130. ve 180. karede düzen skoru yüksek iken, optik akış yaklaşımında düşük çıkmıştır (kırmızı ve yeşil kutu ile işaretlenmiş bölgeler). Düzensiz kare noktaları (eğrinin aşağı yönde olan kısımları) yüksek yeniden inşa hatasını, dolayısıyla ile gerçekleşen anomalileri bize göstermektedir. İki yaklaşım arasında belirgin farkın olması daha yakından incelenmelidir.

60. kare ile 130. kare arasında yaya yoluna giren aracın düzen skorunu ham veri yönteminde etkilemediği, optik akış yönteminde ise oldukça etkilediği görülmektedir. 130. kare incelenecek olursa (Şekil 4.11) yaya yoluna giren aracın optik akış bilgisi net bir şekilde çıkarılmıştır. Görüntüdeki hıza bağlı anomali tespiti burada, ham veri ile yapılan anomali tespitine göre daha doğru sonuç vermiştir. 180. kare incelenecek olursa, gerçek görüntülerde herhangi bir anomali yoktur. Ancak eğitim kümesinin aksine burada farklı bir hareket örüntüsü mevcuttur. Dört yaya arka arkaya, yakın şekilde ve aynı anda yürüdüğü için çıkarılan optik akış daha büyük yer kaplamaktadır (Şekil 4.12-b). Büyük bir nesne hareket ediyormuş gibi algılanacağı için bu durum düzen skorunu düşürmüştür.

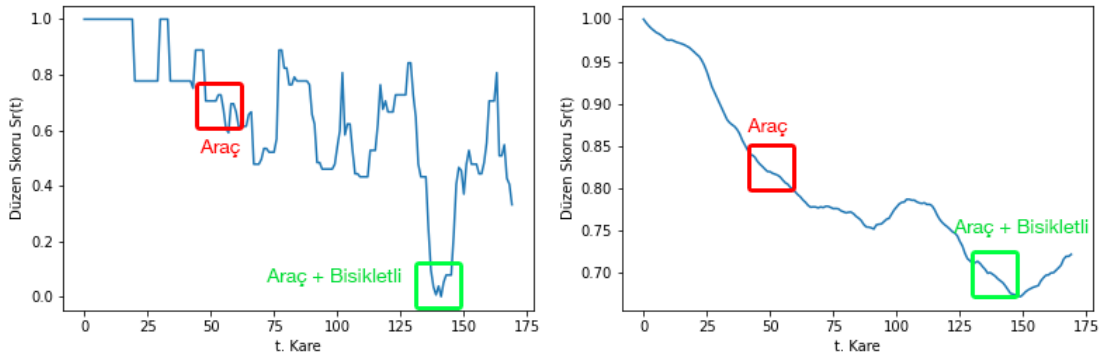


Şekil 4.11. UCSD Ped 1 veri seti 130. kare CAE – (a) gerçek kare, (b) yeniden inşa edilmiş optik akış, (c) optik akışlı kare görüntüleri

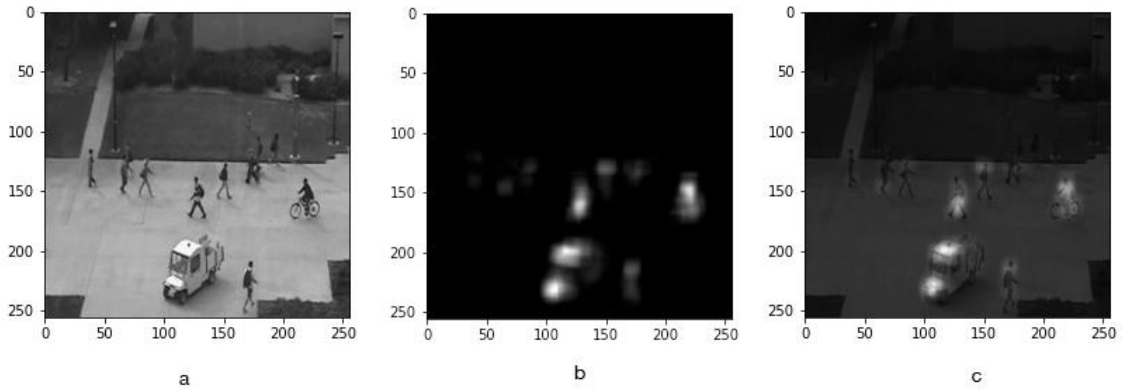


Şekil 4.12. UCSD Ped 1 veri seti 180. kare CAE – (a) gerçek kare, (b) yeniden inşa edilmiş optik akış, (c) optik akışlı kare görüntüleri

CAE mimarileri için UCSD Ped 2 veri setine baktığımızda 50. karede araç yaya yoluna giriş yapmaktadır. 130. karede ise bisikletli de araç ile görüntüye dahil olmaktadır. Düzen skoru grafiklerine baktığımızda optik akış grafiğinde düzensiz kareler zaman içerisinde artmıştır. Ham veri grafiğinde ise 130-140. kareler arasında en düşüğe, diğer durumlarda artıp azalan düzensiz grafik sergilemiştir. Test setine göre 50. kareden itibaren kareler düzensizleşmektedir. Bu durumu optik akış daha iyi açıklamıştır (Şekil 4.13 ve Şekil 4.14).



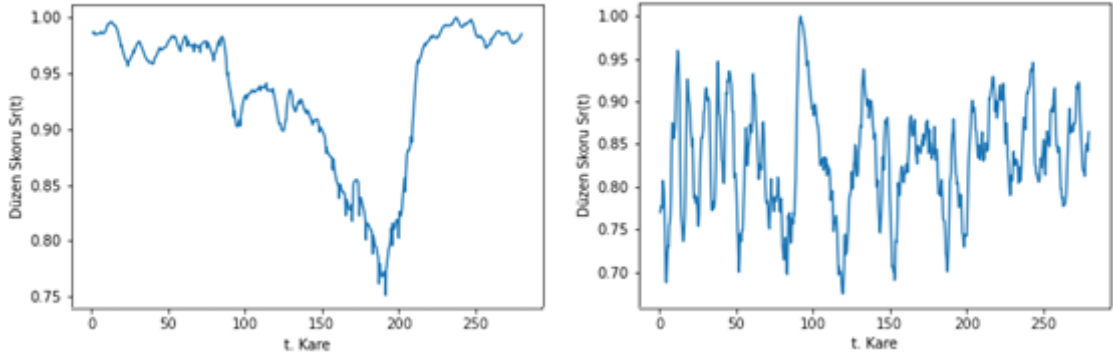
Şekil 4.13. UCSD Ped 2 veri seti CAE - ham veri (sol) ve optik akış (sağ) yaklaşımı düzen skorları



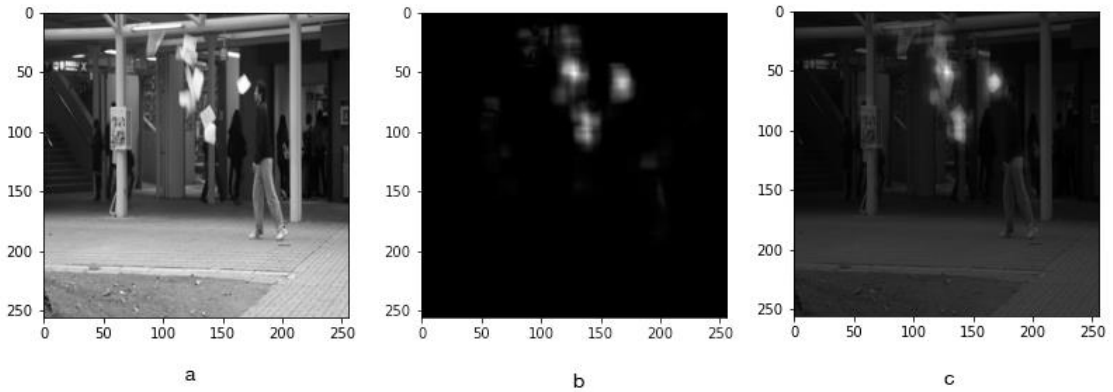
Şekil 4.14. UCSD Ped 2 veri seti 145. kare CAE – (a) gerçek kare, (b) yeniden inşa edilmiş optik akış, (c) optik akışlı kare görüntüleri

CUHK Avenue test kümesinde elindeki kağıtları havaya atan bir öğrenci görünmektedir. Düzen skoru incelendiğinde optik akış grafiği daha düzensiz gelmiştir.

Bu durum, Avenue veri setinin eğitim kümesinde görülmemiş bir davranışın test kümesinde CAE mimarisinin karşısına çıkmasından kaynaklanmaktadır. Daha önce hareket gerçekleşmeyen alanda hareket gerçekleştiği için yeniden oluşturulma sağlanamamış, dolayısı ile düzen skorları aşırı değişken gelmiştir (Şekil 4.15 ve Şekil 4.16).



Şekil 4.15. CUHK Avenue veri seti CAE - ham veri (sol) ve optik akış (sağ) yaklaşımı düzen skorları

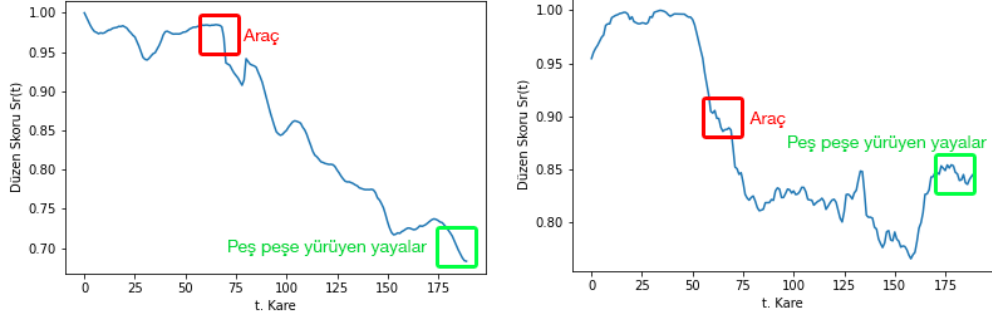


Şekil 4.16. CUHK Avenue veri seti 4. kare CAE – (a) gerçek kare, (b) yeniden inşa edilmiş optik akış, (c) optik akışlı kare görüntüleri

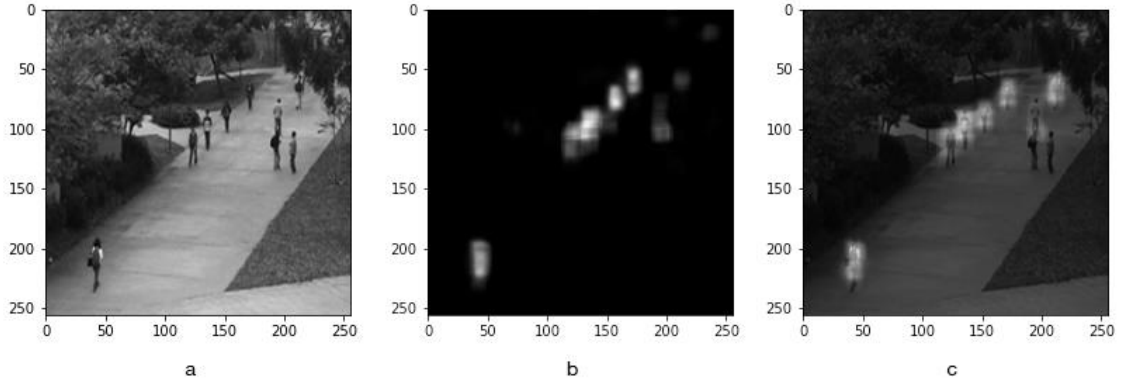
CAE ile yapılan uygulamada AUC değerlerinin optik akış yönteminde, ham veri yöntemine göre UCSD Ped 1 veri setinde 0.08 (%8), UCSD Ped 2 veri setinde 0.04 (%4), CUHK Avenue veri setinde ise 0.01 (%1) daha iyi sonuçlar verdiği görülmektedir. CAE mimarisi en iyi optik akış ile çalışmaktadır (Çizelge 4.3).

4.4.2. ConvLSTMAE Sonuçları

UCSD Ped 1 veri setinde 60. ve 130. kareler arasında yaya yoluna giren araç düzen skorunu her iki yaklaşıma göre düşürmüştür. Ancak 180. karede peş peşe yürüyen yayalar ham veri yöntemine göre düzen skoru düşük iken, optik akış yöntemine göre düzen skoru yükselmiştir. Yani ConvLSTMAE mimarisi ile optik akış yöntemi, yayaları anomali olarak tespit etmemiştir (Şekil 4.17 ve Şekil 4.18).

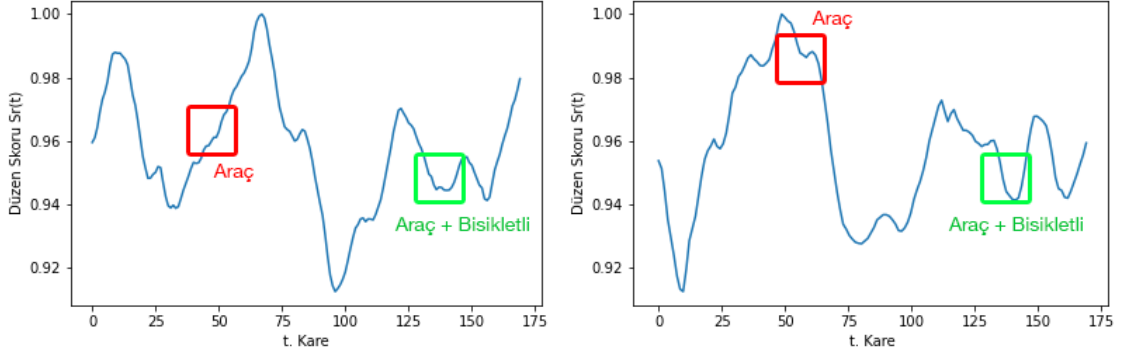


Şekil 4.17. UCSD Ped 1 veri seti ConvLSTMAE - ham veri (sol) ve optik akış (sağ) yaklaşımı düzen skorları

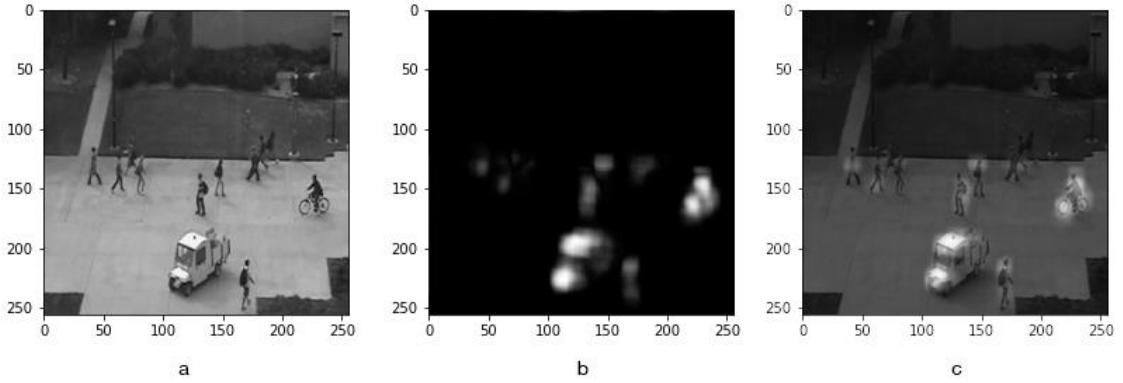


Şekil 4.18. UCSD Ped 1 veri seti 180. kare ConvLSTMAE – (a) gerçek kare, (b) yeniden inşa edilmiş optik akış, (c) optik akışlı kare görüntüleri

UCSD Ped 2 veri setinde aracın görüntüye girdiği andan itibaren düşüş görülen yaklaşım optik akıştır. Ham veri yönteminde araç görüntüye girdikten sonra düzen skorunda yükselme görülmüştür. Her iki yöntemde de araç ve bisikletlinin aynı anda görüntüde olması düzen skorunu düşürmüştür. Araç tespitini optik akış yaklaşımı anlık olarak daha iyi tespit etmiştir (Şekil 4.19 ve Şekil 4.20).

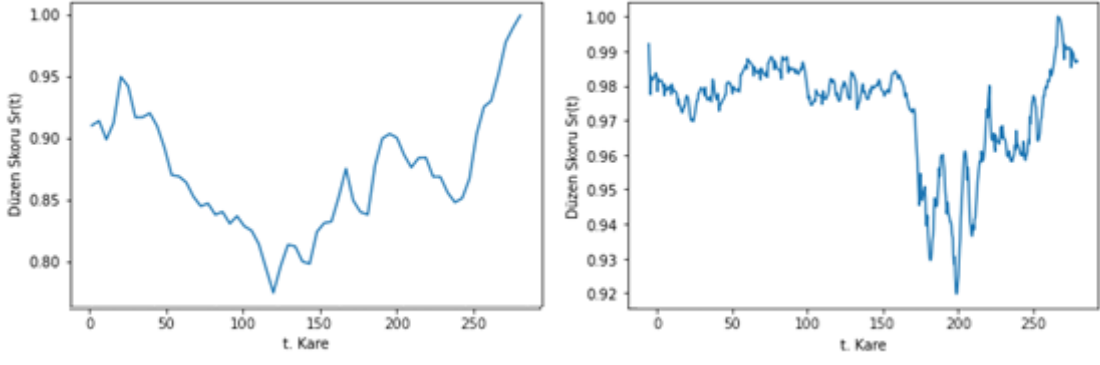


Şekil 4.19. UCSD Ped 2 veri seti ConvLSTMAE - ham veri (sol) ve optik akış (sağ) yaklaşımı düzen skorları

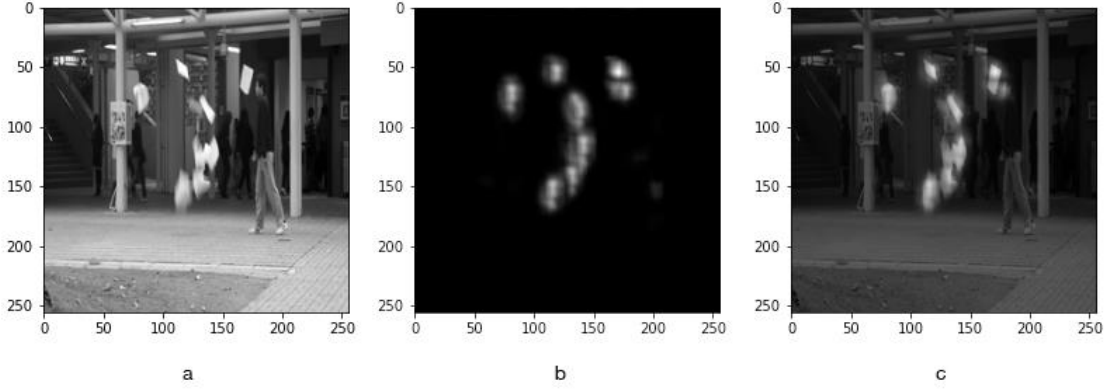


Şekil 4.20. UCSD Ped 2 veri seti 145. kare ConvLSTMAE – (a) gerçek kare, (b) yeniden inşa edilmiş optik akış, (c) optik akışlı kare görüntüleri

CUHK Avenue veri setinde test edilen video klipin tamamı anomali içerdiği halde düzen skoru optik akış yaklaşımında genel olarak yüksek çıkmıştır. Ham veri yönteminde düzen skoru azalan bir trende sahip iken 300. kare civarında minimum skora ulaşmıştır. Optik akış yaklaşımında da 500. kare civarında minimum skor vardır (Şekil 4.21 ve Şekil 4.22).



Şekil 4.21. CUHK Avenue veri seti ConvLSTMAE - ham veri (sol) ve optik akış (sağ) yaklaşımı düzen skorları

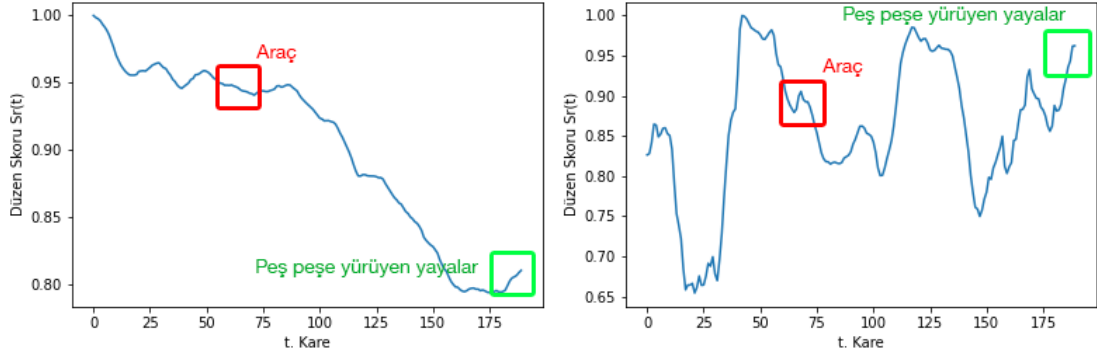


Şekil 4.22. CUHK Avenue veri seti 4. kare ConvLSTMAE – (a) gerçek kare, (b) yeniden inşa edilmiş optik akış, (c) optik akışlı kare görüntüleri

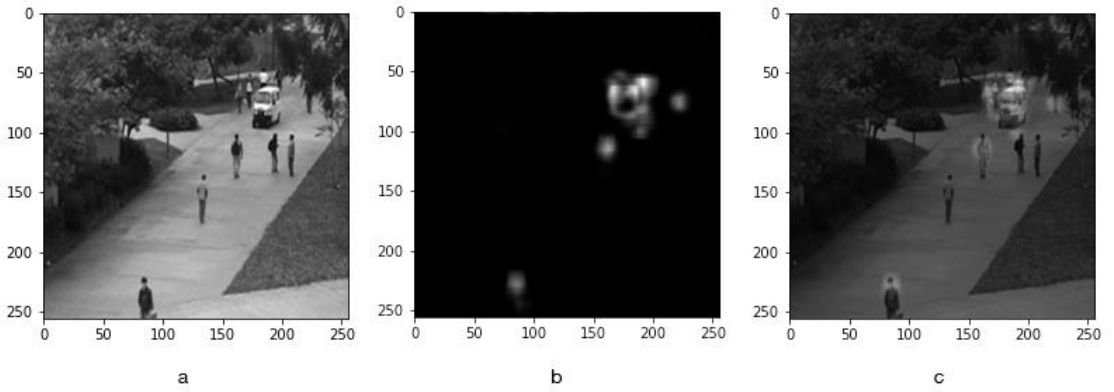
ConvLSTMAE mimarisi ile yapılan analizlerde AUC skoruna göre optik akış yaklaşımı, ham veri yaklaşımına göre, UCSD Ped 1 veri setinde 0.12 (%12), UCSD Ped 2 veri setinde 0.12 (%12) ve Avenue veri setinde 0.21 (%21) daha iyi sonuç vermiştir. Avenue veri setinin 20. test video klipine göre ham veri yaklaşımı daha iyi olmasına karşın, diğer test klipleri de analize dahil edildiğinde ortalama AUC skoru optik akışta daha yüksek çıkmıştır (Çizelge 4.3).

4.4.3. VAE Sonuçları

UCSD Ped 1 veri setinde araç görüntüye girdikten sonra ham veri yaklaşımında düzen skoru düşme eğilimi göstermiştir. Optik akış yaklaşımında ise araçtan sonra azalıp artma örüntüsü göstermiştir. Bu veri setinde ve ilgili test video klipinde VAE mimarisi ile ham veri yaklaşımı aracı tespit etmede daha başarılı iken, peş peşe yürüyen yayaları yanlış tahmin etmeme durumu optik akış yaklaşımında daha başarılı olmuştur (Şekil 4.23 ve Şekil 4.24).

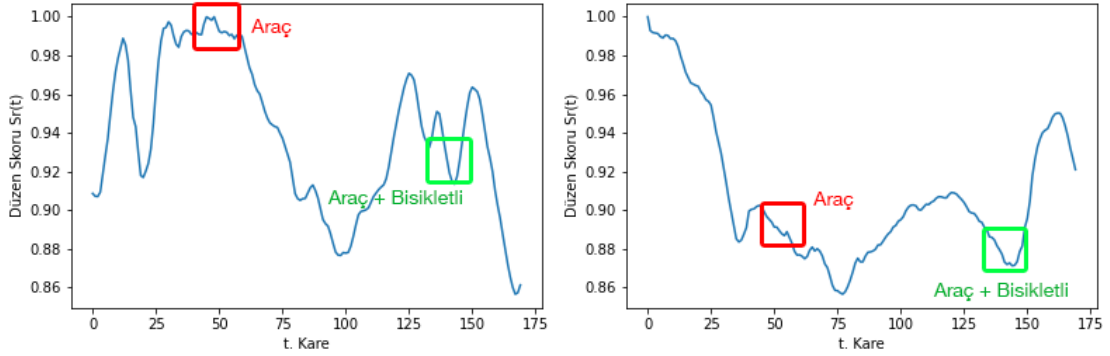


Şekil 4.23. UCSD Ped 1 veri seti VAE - ham veri (sol) ve optik akış (sağ) yaklaşımı düzen skorları

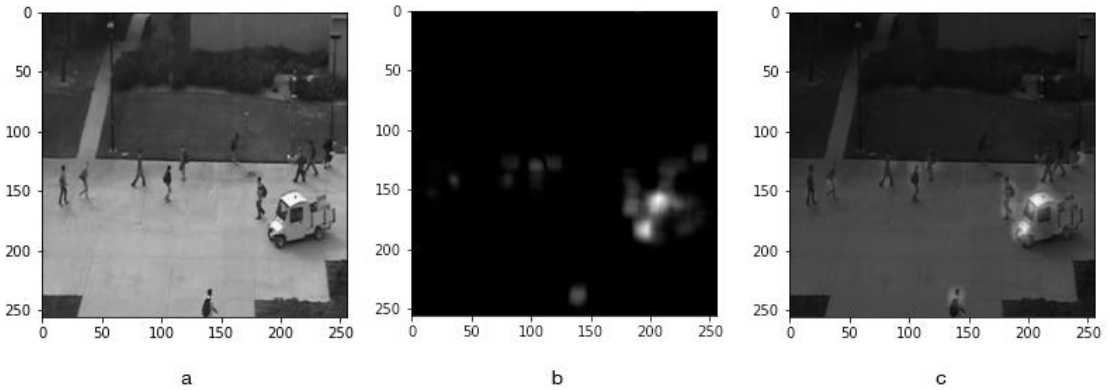


Şekil 4.24. UCSD Ped 1 veri seti 70. kare VAE – (a) gerçek kare, (b) yeniden inşa edilmiş optik akış, (c) optik akışlı kare görüntüleri

UCSD Ped 2 veri setinde ise araç ve bisikletlinin aynı anda görüntüde olduğu durumda optik akış yaklaşımı daha düşük düzen skoru vermiştir. Aracın tamamının görüntüye girdiği anda ham veri yaklaşımında düzen skoru düşme eğilimine girmiştir. Optik akış yaklaşımında ise aracın ufak bir kısmı dahi görüntüye girdiğinde düzen skoru düşmüştür. Bu durumda optik akış yaklaşımı daha iyi sonuçlar vermiştir (Şekil 4.25 ve Şekil 4.26).

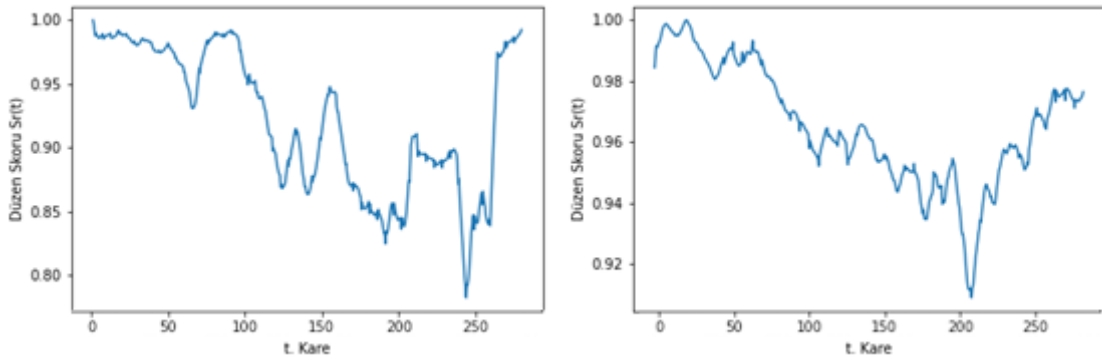


Şekil 4.25. UCSD Ped 2 veri seti VAE - ham veri (sol) ve optik akış (sağ) yaklaşımı düzen skorları

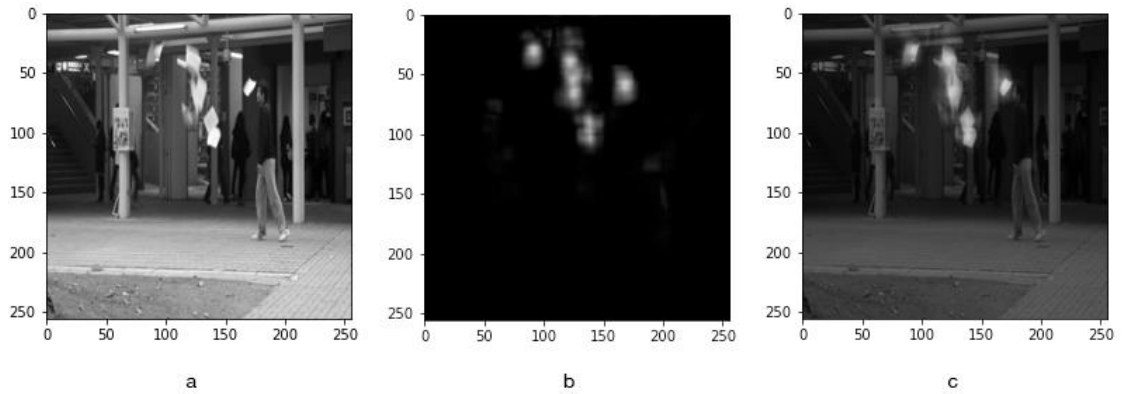


Şekil 4.26. UCSD Ped 2 veri seti 75. kare VAE – (a) gerçek kare, (b) yeniden inşa edilmiş optik akış, (c) optik akışlı kare görüntüleri

CUHK Avenue veri setinde düzen skoru grafiğinin tamamen azalma eğiliminde olması beklenmektedir. Düzen skoru ham veri yaklaşımında bazı yerlerde artarken, genel olarak azalma eğilimindedir. Optik akış yaklaşımında ise, yerel artışlar olmasına karşın sürekli azalan bir grafik karşımıza çıkmaktadır. İlgili video klipte özellikle öğrencinin kağıtları yerden topladığı zaman ham veri yaklaşımının düzen skorunda artmalar olurken, optik akış yaklaşımına ait düzen skorlarında böyle bir artış görülmemektedir (Şekil 4.27 ve Şekil 4.28).



Şekil 4.27. CUHK Avenue veri seti VAE - ham veri (sol) ve optik akış (sağ) yaklaşımı düzen skorları



Şekil 4.28. CUHK Avenue veri seti 265. kare VAE – (a) gerçek kare, (b) yeniden inşa edilmiş optik akış, (c) optik akışlı kare görüntüleri

VAE mimarisi ile yapılan analizde UCSD Ped 1 veri setinde, optik akış yaklaşımı ham veri yaklaşımına göre 0.08 (%8), UCSD Ped 2 veri setinde 0.04 (%4) daha iyi iken, CUHK Avenue veri setinde yaklaşık 0.01 (%1) daha kötü AUC skoru elde edilmiştir (Çizelge 4.3).

Çizelge 4.3'te görüleceği üzere VAE – CUHK Avenue hariç tüm durumlarda optik akış yaklaşımı, ham veri yaklaşımına göre daha yüksek AUC skorlarına sahiptir. Videolarda anomali tespitinde, veri setine bağlı olmak ile, hareket ve hız bazlı anomali tespiti, görüntü bazlı anomali tespitinden daha iyi sonuçlar vermektedir.

Çizelge 4.3. Mimari – yaklaşım / veri seti için ortalama AUC skorları

Veri Seti	CAE		ConvLSTMAE		VAE	
	Ham	Optik Akış	Ham	Optik Akış	Ham	Optik Akış
UCSD Ped 1	0.7017	0.7861	0.6555	0.7713	0.6872	0.7641
UCSD Ped 2	0.7202	0.7667	0.7612	0.8872	0.7563	0.7906
CUHK Avenue	0.8296	0.8395	0.5484	0.7531	0.8434	0.8340

5. SONUÇLAR VE ÖNERİLER

Bu çalışmada iki farklı yaklaşım, üç farklı mimaride, üç farklı veri seti kullanılarak performansları incelenmiştir. Sonuçlar incelenirken test kümesinden her veri seti için belirli bir klip, yansıttığı anomalilere göre seçilmiştir. İlgili test video kliplerinden her mimari ve yaklaşım için düzen skoru grafikleri ile yeniden inşa edilmiş optik akış görselleri kullanılmıştır. Her durum için bu skorlar ve görseller incelenip yorumlanmıştır. AUC skorları hesaplanırken her veri setinin video klipleri için ayrı ayrı AUC skorları hesaplanıp, son aşamada veri seti özelinde ortalama bir AUC skoru elde edilmiştir.

En iyi sonuçlar optik akış yaklaşımında elde edilmiştir. Hareket ve hız bazlı anomali tespiti, ham video görüntüleri ile yapılan anomali tespitinden daha doğru sonuçlar vermiştir.

Yaya yoluna giren araçların tespiti ham veri yaklaşımında mimarilerde tespit edilememiştir. Araçların hareketinin düzen skorunu azaltması beklenirken, ham veri yaklaşımında yükselttiği görülmüştür. Araç tespitinde en iyi yaklaşım optik akış yaklaşımı olmuştur. Bu durum aracın hızının yayalardan fazla olması olarak özetlenebilmektedir. Hız bazlı anomali tespiti araç içeren görüntü durumlarında önerilir.

Peş peşe ve aynı anda yürüyen yayaların hareketi ham veri yaklaşımına göre düzgün tespit edilememiştir. Bu durum düzen skorunu düşürmüştür. Burada optik akış yaklaşımının daha doğru sonuçlar verdiği görülmüştür. Yayaların yürüyüş ve yakınlığından ziyade sadece hareket hızına bağlı analiz yapıldığı için optik akış yaklaşımı daha iyi tahminde bulunmuş, düzen skorunu yükseltmiştir.

Havaya kağıtlarını atıp yerden toplayan öğrencinin bulunduğu görüntü karelerinde en iyi yaklaşım yine optik akış yaklaşımı olmuştur. Nesnelerin hareketi yeniden inşa edilememiş, düzen skoru düşmüştür. Bu durumun sebebi, kağıtların havaya atıldığı karelerde test kümesinde daha önce bir hareket olmadığından öznitelikleri çıkmamış, o bölgede yapılan hareketler öğrenilememiştir. Tahmin edilemeyecek bölgede olan, sık ve

küçük boyutlu nesnelerin hareketi yeniden inşa edilemediği için düzenli azalan bir düzen skoru grafiği elde edilmiştir.

Yapılan analiz sonucunda sadece VAE mimarisi optik akış ve ham veri yaklaşımı için CUHK Avenue veri setinde yaklaşık aynı (%0.9 fark ile) sonuçlar vermiştir. Ham veri yaklaşımının sadece bu mimaride %0.9 daha güçlü çıkması göz ardı edilebilir bir durumdur. Genel olarak çalışma sonuçlarına baktığımızda optik akış yaklaşımının daha iyi sonuçlar verdiği görülmektedir.

Gelecekte yapılması planlanan çalışma, farklı optik akış yöntemleri ile farklı mimarilerin hareket bazlı anomali tespitine etkisinin araştırılmasıdır. Optik akış verilerinin dönüşümleri, öznitelikleri kullanılarak yeni bir yapı oluşturulması, bu gibi yapılar ile anomali tespiti için yeni bir çerçeve önerilmesi planlanmaktadır. Optik akış yaklaşımının daha iyi sonuçlar vermesi, hareket bazlı anomali tespitinin güçlü yanlarını vurgulamaktadır. Optik akış çıkarımı işleminin daha doğru olması neticesinde, ince ayarla optimize edilmiş mimariler ile anomali tespitinin daha iyi performans göstereceği öngörülmektedir.

Aynı zamanda videolarda anomali tespitlerinin en büyük zorluklarından olan, gizlilik ihlali için optik akış yönteminin de ileriye dönük olarak daha uygulanabilir olduğu düşünülmektedir. Gelecekteki çalışmalarda optik akış yönteminin uygulanması, yüz verileri gibi kişisel bilgileri içermeyen analiz yaptığından ötürü daha tercih edilebilir olacaktır. Farklı bir alan için ise, trafik ve araç verilerini içeren video görüntülerinde de optik akış yaklaşımının daha iyi sonuçlar verebileceği düşünülmektedir. Geliştirilen yöntemlerin farklı alanlara ait veri setlerinde de denenmesi planlanmaktadır.

6. KAYNAKLAR

- [1] Precise Security, CCTV Kamera Sayıları En Yüksek 10 Ülke, <https://www.precisecurity.com/articles/top-10-countries-by-number-of-cctv-cameras/> (Erişim Tarihi: **25 Nisan 2022**).
- [2] N. Nemade, V. Gohokar, Comparative performance analysis of optical flow algorithms for anomaly detection, in Proceedings of International Conference on Communication and Information Processing (ICCIP), **2019**.
- [3] C. Lu, J. Shi, J. Jia, Abnormal event detection at 150 fps in matlab, in Proceedings of the IEEE international conference on computer vision, **2013**, pp. 2720-2727.
- [4] M. Hasan, J. Choi, J. Neumann, A.K. Roy-Chowdhury, L.S. Davis, Learning temporal regularity in video sequences, in Proceedings of the IEEE conference on computer vision and pattern recognition, **2016**, pp. 733-742.
- [5] W. Luo, W. Liu, S. Gao, Remembering history with convolutional lstm for anomaly detection, in 2017 IEEE International Conference on Multimedia and Expo (ICME), IEEE, **2017**, pp. 439-444.
- [6] A. Adam, E. Rivlin, I. Shimshoni, D. Reinitz, Robust real-time unusual event detection using multiple fixed-location monitors, IEEE transactions on pattern analysis and machine intelligence, 30 (**2008**) 555-560.
- [7] R. Mehran, A. Oyama, M. Shah, Abnormal crowd behavior detection using social force model, in 2009 IEEE conference on computer vision and pattern recognition, IEEE, **2009**, pp. 935-942.
- [8] N. Li, F. Chang, C. Liu, Spatial-temporal cascade autoencoder for video anomaly detection in crowded scenes, IEEE Transactions on Multimedia, 23 (**2020**) 203-215.
- [9] M. Ravanbakhsh, M. Nabi, E. Sangineto, L. Marcenaro, C. Regazzoni, N. Sebe, Abnormal event detection in videos using generative adversarial nets, in 2017 IEEE international conference on image processing (ICIP), IEEE, **2017**, pp. 1577-1581.
- [10] E. Duman, O.A. Erdem, Anomaly detection in videos using optical flow and convolutional autoencoder, IEEE Access, 7 (**2019**) 183914-183923.

- [11] M. Qiao, T. Wang, J. Li, C. Li, Z. Lin, H. Snoussi, Abnormal event detection based on deep autoencoder fusing optical flow, in 2017 36th Chinese Control Conference (CCC), IEEE, **2017**, pp. 11098-11103.
- [12] S. Wang, E. Zhu, J. Yin, F. Porikli, Anomaly detection in crowded scenes by SL-HOF descriptor and foreground classification, in 2016 23rd International Conference on Pattern Recognition (ICPR), IEEE, **2016**, pp. 3398-3403.
- [13] J. Kim, K. Grauman, Observe locally, infer globally: a space-time MRF for detecting abnormal activities with incremental updates, in 2009 IEEE conference on computer vision and pattern recognition, IEEE, **2009**, pp. 2921-2928.
- [14] S. Yan, J.S. Smith, W. Lu, B. Zhang, Abnormal event detection from videos using a two-stream recurrent variational autoencoder, *IEEE Transactions on Cognitive and Developmental Systems*, 12 (**2018**) 30-42.
- [15] W. Luo, W. Liu, S. Gao, A revisit of sparse coding based anomaly detection in stacked rnn framework, in Proceedings of the IEEE international conference on computer vision, **2017**, pp. 341-349.
- [16] D.E. Rumelhart, G.E. Hinton, R.J. Williams, Learning representations by back-propagating errors, *nature*, 323 (**1986**) 533-536.
- [17] G.-B. Huang, Q.-Y. Zhu, K. Mao, C.-K. Siew, P. Saratchandran, N. Sundararajan, Can threshold networks be trained directly?, in *IEEE Transactions on Circuits and Systems II: Express Briefs*, 53 (**2006**) 187-191.
- [18] C. Nwankpa, W. Ijomah, A. Gachagan, S. Marshall, Activation functions: Comparison of trends in practice and research for deep learning, *arXiv preprint arXiv:1811.03378*, (**2018**).
- [19] I. Goodfellow, Y. Bengio, A. Courville, *Deep learning*, MIT press, **2016**.
- [20] A. Géron, *Hands-on machine learning with Scikit-Learn, Keras, and TensorFlow: Concepts, tools, and techniques to build intelligent systems*, " O'Reilly Media, Inc.", **2019**.
- [21] A. Vouloimos, N. Doulamis, A. Doulamis, E. Protopapadakis, *Deep learning for computer vision: A brief review*, *Computational intelligence and neuroscience*, **2018**.

- [22] Y. LeCun, L. Bottou, Y. Bengio, P. Haffner, Gradient-based learning applied to document recognition, *Proceedings of the IEEE*, 86 (**1998**) 2278-2324.
- [23] S. Hochreiter, J. Schmidhuber, Long short-term memory, *Neural computation*, 9 (**1997**) 1735-1780.
- [24] D. Bank, N. Koenigstein, R. Giryes, Autoencoders, *arXiv preprint arXiv:2003.05991*, (**2020**).
- [25] J. Masci, U. Meier, D. Cireşan, J. Schmidhuber, Stacked convolutional auto-encoders for hierarchical feature extraction, in *International conference on artificial neural networks*, Springer, **2011**, pp. 52-59.
- [26] D.P. Kingma, M. Welling, Auto-encoding variational bayes, *arXiv preprint arXiv:1312.6114*, (**2013**).
- [27] V. Hodge, J. Austin, A survey of outlier detection methodologies, *Artificial intelligence review*, 22 (**2004**) 85-126.
- [28] V. Chandola, A. Banerjee, V. Kumar, Anomaly detection: A survey, *ACM computing surveys (CSUR)*, 41 (**2009**) 1-58.
- [29] R. Pincus, Barnett, V., and Lewis T.: *Outliers in Statistical Data*. J. Wiley & Sons 1994, XVII. 582 pp., Wiley Online Library, **1995**.
- [30] B. Abraham, A. Chuang, Outlier detection and time series modeling, *Technometrics*, 31 (**1989**) 241-248.
- [31] S. Omar, A. Ngadi, H.H. Jebur, Machine learning techniques for anomaly detection: an overview, *International Journal of Computer Applications*, 79 (**2013**).
- [32] N. Görnitz, M. Kloft, K. Rieck, U. Brefeld, Toward supervised anomaly detection, *Journal of Artificial Intelligence Research*, 46 (**2013**) 235-262.
- [33] R. Chalapathy, S. Chawla, Deep learning for anomaly detection: A survey, *arXiv preprint arXiv:1901.03407*, (**2019**).
- [34] J. Patterson, A. Gibson, *Deep learning: A practitioner's approach*, " O'Reilly Media, Inc.", **2017**.
- [35] B.R. Kiran, D.M. Thomas, R. Parakkal, An overview of deep learning based methods for unsupervised and semi-supervised anomaly detection in videos, *Journal of Imaging*, 4 (**2018**) 36.

- [36] R.C. Aygun, A.G. Yavuz, Network anomaly detection with stochastically improved autoencoder based models, in 2017 IEEE 4th International Conference on Cyber Security and Cloud Computing (CSCloud), IEEE, **2017**, pp. 193-198.
- [37] K. Hundman, V. Constantinou, C. Laporte, I. Colwell, T. Soderstrom, Detecting spacecraft anomalies using lstms and nonparametric dynamic thresholding, in Proceedings of the 24th ACM SIGKDD international conference on knowledge discovery & data mining, **2018**, pp. 387-395.
- [38] Y. Cong, J. Yuan, Y. Tang, Video anomaly search in crowded scenes via spatio-temporal motion context, IEEE transactions on information forensics and security, 8 (**2013**) 1590-1599.
- [39] V. Mahadevan, W. Li, V. Bhalodia, N. Vasconcelos, Anomaly detection in crowded scenes, in 2010 IEEE computer society conference on computer vision and pattern recognition, IEEE, **2010**, pp. 1975-1981.
- [40] X. Wang, Z. Che, B. Jiang, N. Xiao, K. Yang, J. Tang, J. Ye, J. Wang, Q. Qi, Robust unsupervised video anomaly detection by multipath frame prediction, IEEE Transactions on Neural Networks and Learning Systems, (**2021**).
- [41] B.D. Lucas, T. Kanade, An iterative image registration technique with an application to stereo vision, Vancouver, **1981**.
- [42] B.K. Horn, B.G. Schunck, Determining optical flow, Artificial intelligence, 17 (**1981**) 185-203.
- [43] G. Farnebäck, Two-frame motion estimation based on polynomial expansion, Scandinavian conference on Image analysis, Springer, **2003**, pp. 363-370.
- [44] R. Nayak, U.C. Pati, S.K. Das, A comprehensive review on deep learning-based methods for video anomaly detection, Image and Vision Computing, 106 (**2021**) 104078.
- [45] J.R. Medel, A. Savakis, Anomaly detection in video using predictive convolutional long short-term memory networks, arXiv preprint arXiv:1612.00390, (**2016**).
- [46] T. Li, H. Chang, M. Wang, B. Ni, R. Hong, S. Yan, Crowded scene analysis: A survey, IEEE transactions on circuits and systems for video technology, 25 (**2014**) 367-386.

- [47] D. Tran, L. Bourdev, R. Fergus, L. Torresani, M. Paluri, Learning spatiotemporal features with 3d convolutional networks, in Proceedings of the IEEE international conference on computer vision, **2015**, pp. 4489-4497.
- [48] Y. Zhao, B. Deng, C. Shen, Y. Liu, H. Lu, X.-S. Hua, Spatio-temporal autoencoder for video anomaly detection, Proceedings of the 25th ACM international conference on Multimedia, **2017**, pp. 1933-1941.
- [49] V. Patraucean, A. Handa, R. Cipolla, Spatio-temporal video autoencoder with differentiable memory, arXiv preprint arXiv:1511.06309, (**2015**).
- [50] Y. Fan, G. Wen, D. Li, S. Qiu, M.D. Levine, F. Xiao, Video anomaly detection and localization via gaussian mixture fully convolutional variational autoencoder, Computer Vision and Image Understanding, 195 (**2020**) 102920.
- [51] C. Wu, S. Shao, C. Tunc, S. Hariri, Video anomaly detection using pre-trained deep convolutional neural nets and context mining, in 2020 IEEE/ACS 17th International Conference on Computer Systems and Applications (AICCSA), IEEE, **2020**, pp. 1-8.
- [52] J. Feng, Y. Liang, L. Li, Anomaly detection in videos using two-stream autoencoder with post hoc interpretability, Computational Intelligence and Neuroscience, **2021**.