

ANKARA ÜNİVERSİTESİ
FEN BİLİMLERİ ENSTİTÜSÜ

DOKTORA TEZİ

VIDEO GÖZETİMİNDE ANORMAL HAREKET TESPİTİ

Erkan ŞENGÖNÜL

BİLGİSAYAR MÜHENDİSLİĞİ ANABİLİM DALI

ANKARA
2024

Her hakkı saklıdır

ÖZET

Doktora Tezi

VIDEO GÖZETİMİNDE ANORMAL HAREKET TESPİTİ

Erkan ŞENGÖNÜL

Ankara Üniversitesi
Fen Bilimleri Enstitüsü
Bilgisayar Mühendisliği Anabilim Dalı

Danışman: Prof. Dr. Refik SAMET

Gözetim kameralarının dünya genelinde özel ve kamusal alanlarda hızla yaygınlaşmasıyla birlikte, artan video kayıt hacminin manuel olarak incelenmesi, zaman alıcı ve zorlayıcı bir süreç haline gelmiştir. Bu durum, anormal davranışların doğru ve hızlı bir şekilde tespit edilmesi ihtiyacını artırmıştır. Dolayısıyla, video gözetiminde anormal davranışların doğru bir şekilde tanımlanması ve beklenmedik desenler ile olağandışı hareketlerin otomatik olarak belirlenmesi önemli bir araştırma alanı haline gelmiştir.

Bu doktora tezi, video gözetim sistemlerinde artan kamera kullanımının getirdiği zorlukları ele alarak, anormal olayların otomatik tespiti için yenilikçi bir metodoloji sunmaktadır. Önerilen metodoloji, veri ön işleme, özellik öğrenme ve anormallik tespiti olmak üzere üç temel aşamayı içermektedir. Ön işleme aşamasında, video verilerinden gereksiz bilgilerin temizlenmesi ve verilerin etkili bir şekilde yorumlanabilmesi için çeşitli işlemler uygulanmaktadır. Özellik öğrenme aşamasında, ConvLSTM-AE modeli kullanılarak normal video verilerinin temel özellikleri öğrenilmekte ve bu özellikler anormal durumların tespitinde kullanılmaktadır. Anormallik tespiti aşamasında ise, model tarafından üretilen düzenlilik puanlarının ürettiği yerel minima noktaları, belirli bir algorithmadan geçerek anormal hareketlerin başlangıç ve bitiş zamanlamasını belirlemede esas alınmaktadır. Bu özgün metodoloji, tezi diğer çalışmalardan ayırmaktadır. Yapılan testler, önerilen metodolojinin CUHK Avenue veri kümesinde %91,2 ve UMN veri kümesinde %94,1 AUC ile yüksek doğruluk seviyelerine ulaşarak mevcut yaklaşımlara kıyasla önemli bir başarı gösterdiğini kanıtlamaktadır. Ayrıca, tez VGAT alanındaki literatürü ve önceki çalışmaları kapsamlı bir şekilde inceleyerek, önerilen metodolojinin güçlü yönlerini ve potansiyel geliştirme alanlarını vurgulamaktadır. Bu tezin, VGAT alanında önemli bir katkı sağlayacağı ve gelecekteki araştırmalara yönelik yol gösterici bir kaynak oluşturarak, gözetim ve ilgili sektörlerdeki profesyonellere değerli bir referans sağlayacağı değerlendirilmektedir.

Ocak 2024, 149 sayfa

Anahtar Kelimeler: Video gözetimi, anomali, anormal hareket tespiti, derin öğrenme.

ABSTRACT

PhD Thesis

ABNORMAL MOTION DETECTION FOR SURVEILLANCE VIDEO

Erkan ŞENGÖNÜL

Ankara University
Graduate School of Natural and Applied Sciences
Department of Computer Engineering

Supervisor: Prof. Dr. Refik SAMET

With the rapid increase of surveillance cameras in both private and public spaces worldwide, the manual inspection of the growing volume of video recordings has become a time-consuming and challenging process. This situation has heightened the need for the accurate and swift detection of abnormal movements. Consequently, the accurate identification of abnormal movements in video surveillance and the automatic detection of unexpected patterns have become significant research areas.

This doctoral thesis addresses the challenges brought by the increasing use of cameras in video surveillance systems, presenting an innovative methodology for the automatic detection of abnormal events. The proposed methodology comprises three main stages: data preprocessing, feature learning, and anomaly detection. In the preprocessing stage, various processes are applied to clean unnecessary information from video data and to interpret the data effectively. During the feature learning stage, the ConvLSTM-AE model is used to learn the fundamental characteristics of normal video data, which are then utilized in detecting abnormal situations. In the anomaly detection stage, the local minima points produced by the regularity scores generated by the model are used, following a specific algorithm, to determine the start and end timing of abnormal movements. This unique methodology distinguishes the thesis from other studies. Tests have demonstrated that the proposed methodology achieves high accuracy levels, with 91.2% AUC in the CUHK Avenue dataset and 94.1% AUC in the UMN dataset, proving to be a significant success compared to existing approaches. Additionally, the thesis comprehensively reviews the literature in the VGAT field and previous studies, highlighting the strengths of the proposed methodology and potential areas for further development. It is evaluated that this thesis will make a significant contribution to the VGAT field and will serve as a guiding resource for future research, providing valuable reference to professionals in surveillance and related sectors.

January 2024, 149 pages

Key Words: Video surveillance, anomaly, abnormal motion detection, deep learning.

TEŞEKKÜR

Bu tezin her aşamasında, planlamasından gerçekleştirilmesine, araştırılmasından oluşturulmasına kadar uzanan süreçte, bilgi ve tecrübeleriyle bana ışık tutan, akademik ve bilimsel rehberlikleriyle çalışmamın temellerini sağlamlaştıran hem akademik ortamda hem de kişisel gelişimimde değerli fikirleriyle yol gösteren değerli hocam Prof. Dr. Refik SAMET'e derin minnettarlığımı ifade etmek isterim.

Aileme, bu yolculukta manevi destekleri ve bitmeyen sabırlarıyla yanımda oldukları ve her adımda bana güç verdikleri için kalpten teşekkür ediyorum. Onların sevgisi ve inancı, bu akademik serüvenimin en büyük motivasyon kaynağı olmuştur.

Her birinin katkısı, bu tezin oluşumundaki başarıya önemli birer katkı sağlamıştır ve bu başarı, onların destekleri olmadan mümkün olmazdı.

Erkan ŞENGÖNÜL
Ankara, Ocak 2024

İÇİNDEKİLER

TEZ ONAY SAYFASI	
ÖZET.....	i
ABSTRACT	ii
TEŞEKKÜR.....	iii
KISALTMALAR DİZİNİ	vi
ŞEKİLLER DİZİNİ	vii
ÇİZELGELER DİZİNİ	ix
1. GİRİŞ.....	1
1.1 Tezin Kapsamı	3
1.2 Tezin Özgün Değeri.....	5
1.3 Tezin Yapısı.....	6
2. KAYNAK ÖZETLERİ VE LİTERATÜR DEĞERLENDİRMESİ	8
3. VGAT ALANINDA YAYGIN KULLANILAN YAKLAŞIMLAR VE YÖNTEMLER	29
3.1 Öğrenme Türlerine Dayalı Yaklaşımlar.....	30
3.1.1 Denetimli öğrenme yaklaşımları	31
3.1.2 Denetimsiz öğrenme yaklaşımları	32
3.1.3 Yarı-denetimli öğrenme yaklaşımları.....	33
3.1.4 Öğrenme yaklaşımlarının karşılaştırılması.....	34
3.2 Algoritma Türlerine Dayalı Yaklaşımlar	37
3.2.1 İstatistik tabanlı yaklaşımlar	37
3.2.1.1 Parametrik yöntemler	38
3.2.1.1.1 Gauss karışım modeli	38
3.2.1.1.2 Bayesian yaklaşımı.....	39
3.2.1.1.3 Frekans ve spektral analiz	40
3.2.1.2 Parametrik olmayan yöntemler	41
3.2.1.2.1 Yerel aykırı değer faktörü	41
3.2.1.2.2 Büyüklük ve momentum histogramı.....	42
3.2.1.2.3 Optik akış.....	43
3.2.1.2.4 Yapısal bağlam tanımlayıcısı	44
3.2.1.2.5 Kernel yoğunluk tahmini	45
3.2.1.3 İstatistik tabanlı yaklaşımların genel değerlendirmesi	46
3.2.2 Sınıflandırma tabanlı yaklaşımlar	47
3.2.2.1 Tek sınıflı sınıflandırma.....	48
3.2.2.2 Hesaplama yoğunluğu temelli sınıflandırma	49
3.2.2.3 Destek vektör makineleri	49
3.2.2.4 Karar ağaçları ve karar kümeleri.....	50
3.2.2.5 Kümeleme.....	51
3.2.2.6 Sınıflandırma tabanlı yaklaşımların genel değerlendirmesi	52
3.2.3 Yeniden yapılandırma tabanlı yaklaşımlar.....	54
3.2.3.1 Otomatik kodlayıcı	54
3.2.3.2 Üretici çekişmeli ağlar	58
3.2.3.3 Yeniden yapılandırma tabanlı yaklaşımların genel değerlendirmesi.....	60
3.2.4 Tahmin tabanlı yaklaşımlar	61
3.2.4.1 Tekrarlayan sinir ağı.....	61

3.2.4.2 Uzun kısa süreli hafıza	64
3.2.4.3 Görüntü dönüştürücü.....	69
3.2.4.4 Tahmin tabanlı yaklaşımların genel değerlendirmesi.....	71
3.2.5 Diğer yaklaşımlar	72
3.2.5.1 Bulanık mantık	72
3.2.5.2 Kısa kodlama ve sözlük öğrenme	73
3.2.5.3 Heuristik yöntemler.....	74
3.2.6 Bölüm değerlendirmesi	75
4. ANOMALİ TESPİTİNDE KULLANILAN VERİ KÜMELERİ	77
4.1 CUHK Avenue Veri Kümesi.....	77
4.2 Subway Veri Kümesi.....	78
4.3 UCF-Crime Veri Kümesi	79
4.4 HR-Crime Veri Kümesi	81
4.5 UCSD Veri Kümesi.....	82
4.6 UMN Veri Kümesi.....	83
4.7 ShanghaiTech Veri Kümesi	84
4.8 Veri Kümelerinin Dağılımları	85
4.8.1 Kayıt ortamına göre dağılım	86
4.8.2 Renk uzayına göre dağılım	87
4.8.3 Gerçek etiket değerlerine göre dağılım	89
5. VGAT METODOLOJİSİNİN GELİŞTİRİLMESİ	92
5.1 Ön İşleme.....	93
5.1.1 Ardışık karelerin çıkarılması	93
5.1.2 Karelerin yeniden boyutlandırılması.....	94
5.1.3 Arka planın çıkarılması	95
5.1.4 Gürültülerin azaltılması.....	97
5.2 Özelliklerin Öğrenilmesi	100
5.2.1 Uzamsal kodlayıcı.....	101
5.2.2 Zamansal kodlayıcı.....	103
5.2.3 Gizli kod.....	103
5.2.4 Zamansal çözücü.....	104
5.2.5 Uzamsal çözücü.....	104
5.2.6 Kayıp fonksiyonları.....	105
5.2.7 Optimizasyon algoritmaları.....	107
5.3 Anormallik Tespiti.....	109
5.4 Değerlendirme Metrikleri.....	111
6. ÖNERİLEN METODOLOJİNİN UYGULANMASI	114
6.1 Kullanılan Teknolojiler ve Veri Kümesi	114
6.2 Metodolojinin Uygulanması.....	115
6.3 Uygulama Sonuçları	118
6.4 Değerlendirme.....	126
7. SONUÇ	128
KAYNAKLAR.....	131

KISALTMALAR DİZİNİ

AE	AutoEncoder
AI	Artificial Intelligence
CNN	Convolutional Neural Networks
ConvLSTM	Convolutional LSTM
CoP	Coherence Pursuit
CUHK	The Chinese University of Hong Kong
DCT	Discrete Cosine Transform
DL	Deep Learning
FCN	Fully Convolutional Network
FPS	Frame Per Second
GAN	Generative Adversarial Network
GMM	Gaussian Mixture Model
HoMM	Histogram of Magnitude and Momentum
HOG	Histogram of Oriented Gradients
HOFO	Histogram of the Optical Flow Orientation
HOS	Histograms of Oriented Swarms
HR	Human-Related
KDE	Kernel Density Estimation
K-NN	K-Nearest Neighbor
LOF	Local Outlier Factor
LSTM	Long Short-Term Memory
MAE	Mean Absolute Error
ML	Machine Learning
MOG	Mixture of Gaussians
MSE	Mean Squared Error
OCC	One-Class Classification
PCA	Principal Component Analysis
PED	Pedestrian
RNN	Recurrent Neural Network
SCD	Structural Context Descriptor
SGD	Stochastic Gradient Descent
SIFT	Scale Invariant Feature Transform
SVM	Support Vector Machine
UCF	University of Central Florida
UCSD	University of California, San Diego
VAE	Variational AutoEncoder
VGAT	Video Gözetiminde Anomali Tespiti
ViT	Vision Transformer
ViViT	Video Vision Transformer

ŞEKİLLER DİZİNİ

Şekil 2.1 Arka plan çıkarma işlemi uygulandıktan sonra elde edilen orijinal global görüntünün ve ön plan görüntüsünün optik akış yönü histogramı (T. Wang ve Snoussi 2014).....	13
Şekil 2.2 Küresel ve Yama-tabanlı yaklaşımların kombinasyonu (Parate vd. 2018).....	15
Şekil 2.3 Önerilen modelden tahmin edilen pozların bir örneği (Rodrigues vd. 2020) ..	16
Şekil 2.4 Önerilen anomali tespit ve konumlandırma yaklaşımı (Bansod ve Nandedkar 2020)	17
Şekil 2.5 Önerilen MIST ilk önce klip düzeyindeki sözde etiketleri (J.-C. Feng vd. 2021)	20
Şekil 2.6 Önerilen anomali tespit modeli (H. Yuan vd. 2021).....	21
Şekil 3.1 Gauss Karışım Modeli	39
Şekil 3.2 Yerel Aykırı Değer Faktörü (Alghushairy vd. 2020).....	42
Şekil 3.3 Optik Akış (Mehran vd. 2010).....	44
Şekil 3.4 Kümeleme yöntemleri: (a) Merkez tabanlı kümeleme, (b) Hiyerarşik kümeleme, (c) Yoğunluk tabanlı kümeleme	52
Şekil 3.5 Otomatik kodlayıcı yapısı	55
Şekil 3.6 Üretici Çekişmeli Ağ	59
Şekil 3.7 Tekrarlayan Sinir Ağı (Colah 2015)	62
Şekil 3.8 Kaybolan ve Patlayan Gradyan (Pandey 2022)	63
Şekil 3.9 LSTM hücresi (Colah 2015)	65
Şekil 3.10 Birden fazla LSTM hücresinin bağlantı yapısı ve veri iletimi.....	67
Şekil 3.11 ViT Tabanlı Anomali Tespiti Modeli-AnoVit (Y. Lee ve Kang 2022).....	71
Şekil 4.1 CUHK Avenue veri kümesindeki anormal kare örnekleri (Lu vd. 2013).....	78
Şekil 4.2 Subway veri kümesindeki anormal kare örnekleri (Adam vd. 2008). İlk satır Subway Entrance veri kümesinden, ikinci satır Subway Exit alt veri kümesinden alınmıştır.	79
Şekil 4.3 UCF-Crime veri kümesinden (a) normal ve (b) anormal kare örnekleri (Sultani vd. 2018).....	80
Şekil 4.4 HR-Crime veri kümesinden anormal kare örnekleri (Boekhoudt vd. 2021) ...	82
Şekil 4.5 UCSD veri kümesinden (a) Ped1 alt kümesinden ve (b) Ped2 alt kümesinden anormal kare örnekleri (Mahadevan vd. 2010).....	83
Şekil 4.6 UMN veri kümesinden (a) normal ve (b) anormal kare örnekleri (Mehran vd. 2010)	84
Şekil 4.7 ShanghaiTech veri kümesinden anormal kare örnekleri (Luo vd. 2017a).....	85
Şekil 5.1 Önerilen metodolojinin şematik gösterimi	92
Şekil 5.2 Videodan kare çıkarma işlemi	94
Şekil 5.3 Karenin yeniden boyutlandırılması işlemi	95
Şekil 5.4 Gürültü azaltma işlemi sonucunda elde edilen kare	99
Şekil 5.5 Özelliklerin öğrenilmesinde kullanılan model.....	100
Şekil 5.6 Anormallik tespit adımları	110
Şekil 6.1 Arka plan çıkarma işlemi sonucunda elde edilen kare.....	115
Şekil 6.2 Anormallik tespiti uygulanmış bir model çıktısı	117
Şekil 6.3 CUHK Avenue veri kümesindeki test videosuna uygulanan anormallik tespit işlemi.....	118
Şekil 6.4 CUHK Avenue veri kümesindeki uygulama sonuçları (klip no 1).....	119

Şekil 6.5 CUHK Avenue veri kümesindeki uygulama sonuçları (klip no 16).....	120
Şekil 6.6 CUHK Avenue veri kümesindeki uygulama sonuçları (klip no 20).....	120
Şekil 6.7 UMN veri kümesindeki uygulama sonuçları (klip no 1)	121
Şekil 6.8 UMN veri kümesindeki uygulama sonuçları (klip no 6)	121
Şekil 6.9 CUHK Avenue veri kümesindeki uygulama sonuçları.....	122
Şekil 6.10 UMN veri kümesindeki uygulama sonuçları	123



ÇİZELGELER DİZİNİ

Çizelge 2.1 Literatür özeti ve değerlendirmesi	25
Çizelge 3.1 VGAT alanındaki yaklaşımların sınıflandırılması.....	30
Çizelge 3.2 Öğrenme yöntemlerinin avantaj ve dezavantajları	36
Çizelge 3.3 İstatistik tabanlı tekniklerin tekniklerin avantaj ve dezavantajları	47
Çizelge 3.4 Sınıflandırma tabanlı yöntemlerin avantaj ve dezavantajları.....	53
Çizelge 3.5 Yeniden yapılandırma tabanlı yöntemlerin avantaj ve dezavantajları	61
Çizelge 3.6 Tahmin tabanlı tekniklerin tekniklerin avantaj ve dezavantajları	72
Çizelge 3.7 Algoritma türlerine dayalı yaklaşımların avantaj ve dezavantajları	76
Çizelge 4.1 UCF-Crime ve HR-Crime veri kümelerindeki anormal olay kategorileri ve video sayıları.....	80
Çizelge 4.2 VGAT veri kümelerinin özellikleri.....	86
Çizelge 4.3 VGAT veri kümelerinin kare özellikleri ve sayıları	86
Çizelge 4.4 Veri kümelerinin kayıt ortamlarına göre dağılımı	87
Çizelge 4.5 Veri kümelerinin renk uzaylarına göre dağılımı	88
Çizelge 4.6 Etiketleme türlerinin avantaj ve dezavantajları.....	90
Çizelge 4.7 Veri kümelerinin gerçek etiket değerlerine göre dağılımı	91
Çizelge 6.1 CUHK Avenue ve UMN veri kümeleri için kullanılan değerlendirme metrikleri.....	124
Çizelge 6.2 CUHK Avenue veri kümesinde önerdiğimiz metodolojinin kare düzeyinde performansının (AUC) diğer yöntemlerle karşılaştırılması	125
Çizelge 6.3 UMN veri kümesinde önerdiğimiz metodolojinin kare düzeyinde performansının (AUC) diğer yöntemlerle karşılaştırılması	125

1. GİRİŞ

Anomali terimi, bir teorinin veya kuramın öngördüğü davranışa uymayan veya bir veri kümesinde beklenenin dışında gerçekleşen durumu ifade eder. Gözetim videolarındaki anomali, gözetim kameraları tarafından kaydedilen görüntülerde normalden sapma gösteren veya beklenmeyen olayları belirtir. Bu olaylar, kavga, kaza, hırsızlık, soygun gibi gerçekçi anomaliler olabileceği gibi, arka planda beliren bir hayvan veya nesne gibi daha basit olguları da içerebilir. Ancak video gözetimi yalnızca güvenlikle sınırlı değildir; endüstriyel sektörde üretim süreçlerinin izlenmesi, kalite kontrolü ve ekipman durumlarının gözlemlenmesi, perakende sektöründe müşteri davranışlarının analizi ve ürün stoklarının izlenmesi, trafik yönetiminde kazaların önlenmesi ve trafik akışının düzenlenmesi, eğitimde sınıf gözetimi ve öğrenci takibi, şehir yönetiminde suçları azaltma ve acil durumlara hızlı müdahale etme gibi bir dizi farklı alanda kullanılabilir. Bu çalışma, kapsamlı bir araştırma ve geliştirme süreci sonucunda, geniş bir yelpazede kullanım alanları bulunan video gözetimindeki anomalilerin tespiti ve sınıflandırılmasına yönelik bir metodoloji önerisi sunarak, bu alana katkı sağlamayı amaçlamaktadır. Bu giriş bölümünde, çalışmanın motivasyonu, amaçları ve yapısal düzeni hakkında ayrıntılı bilgi sunulacaktır.

Yüksek çözünürlüklü kameraların yaygınlaşması ve bilgisayarlı görü teknolojilerindeki ilerlemeler, üretilen video verilerinin hızla artmasına yol açmıştır. Bu büyük miktardaki video verisini manuel olarak gözden geçirmek ve analiz etmek hem zorlayıcı hem de zaman alıcı bir görev haline gelmiştir. Bu bağlamda, video gözetiminde anomali tespiti (VGAT) yöntemleri, beklenen desenlerden sapma gösteren olağandışı olayları veya davranışları otomatik olarak tanımlamayı amaçlayan önemli bir araştırma alanı haline gelmiştir.

Anomali tespiti, gözetim videolarında anormal olayları veya davranışları otomatik olarak belirleyerek güvenlik, suç soruşturması, sağlık ve endüstriyel uygulamalar açısından kritik bir rol oynamaktadır. Bu tür tespitler, şiddet, hırsızlık, vandalizm, trafik kazaları gibi olumsuz olayları, potansiyel tehdit veya suçları belirlemeyi ve ilgilileri uyarmayı amaçlamaktadır. Bu bağlamda, anormal olayları tespit etmek, zararı en aza

indirmek veya önlemek, güvenliğini artırmak ve adli soruşturmalara kanıt sağlamak açısından son derece önemlidir. Benzer şekilde, hastane kameraları sayesinde yasak veya riskli davranışlar tespit edilerek hem hastaların hem de personelin güvenliğini sağlanabilir. Bu ve benzeri örnekler birçok alan için verilebilir.

Geleneksel anomali tespiti yöntemleri, genellikle belirli kurallar ve önceden belirlenmiş eşikler temelinde işleyiş gösterir. Ancak, gerçek dünyada karşılaşılan olaylar ve davranışlar sıklıkla bu kuralların dışında olabilir. Bu nedenle, geleneksel yöntemler özellikle karmaşık senaryolar ve değişen ortamlara uyum sağlamakta zorlanmakta ve etkinlikleri azalmaktadır. Ancak son yıllarda derin öğrenme, bilgisayarlı görü ve analitik algoritmalarında sağlanan son gelişmeler ve ilerlemeler, bu sorunları aşmak, daha doğru ve güvenilir anomali tespiti teknikleri geliştirmek için önemli bir fırsat sunmaktadır.

VGAT alanındaki çalışmalar ve literatür incelendiğinde, uygulanan çeşitli yöntemlerin kendi içinde benzersiz zorluklar ve potansiyel fırsatlar sunduğu görülmektedir. Bu durum, VGAT alanının geniş bir perspektifte ele alınması ve yeni yaklaşımların geliştirilmesi ihtiyacını belirgin hale getirmiştir.

Bu alandaki fırsatlar arasında özellikle güvenlik, sağlık veya endüstriyel uygulamalarda anormal olayların erken ve doğru bir şekilde tespit edilmesi yer almaktadır. Gelişmiş algoritmalar ve yapay zeka teknolojilerinin kullanımı, video gözetim sistemlerini sadece rutin gözetim için değil, aynı zamanda olağandışı ve şüpheli aktivitelerin otomatik tespiti için de uygun hale getirmektedir. Bu teknoloji, halka açık alanların güvenliğini artırmak, suçları önlemek ve hızlı müdahaleyi sağlamak için büyük potansiyele sahiptir. Ayrıca, veri analizi ve davranış tanıma konusundaki ilerlemeler sayesinde, bu sistemler daha etkili ve hedef odaklı hale gelebilir. Eğitim, sağlık hizmetleri ve perakende gibi çeşitli sektörlerde uygulanabilirliği, bu alandaki fırsatları genişletmektedir.

Bu alandaki zorluklar, teknolojik ve pratik boyutlarda kendini göstermektedir. Teknolojik olarak, değişken çevresel koşullar ve farklı hareket biçimlerinin doğru bir şekilde tanınması, algoritmaların karmaşıklığını artırmaktadır. Ayrıca, yüksek maliyet, kaynak gereksinimleri ve farklı sistemlerle entegrasyon zorlukları da bu teknolojinin

geniş çaplı uygulanabilirliğini sınırlamaktadır. Bir diğer zorluk da veri çeşitliliğidir. Algoritmalar genellikle normal davranışları modellemek üzere eğitilir ancak gerçek dünyada karşılaşılan olaylar ve davranışlar oldukça farklılık gösterebilmektedir. Bu durumda, eğitim veri kümesinde temsil edilmeyen veya daha önce görülmemiş anomali tiplerini tespit etmek zor olabilir. Bunun yanı sıra, farklı ışık koşulları, perspektifler, çözünürlük gibi faktörler nedeniyle gözetim kameralarının görüş alanı veya kalitesinde değişiklikler olabilir. Bu da VGAT sistemlerinin performansını olumsuz yönde etkileyebilir ve yanlış alarm veya anomalilerin gözden kaçmasına yol açabilir.

Anomali tespitinde hatalı pozitif veya hatalı negatif sonuçlar elde etme riski her zaman mevcuttur. Hatalı pozitif, normal olarak sınıflandırılması gereken bir olayın yanlışlıkla anormal olarak sınıflandırılması anlamına gelir. Bu durum gereksiz uyarılara veya alarmlara yol açabilir. Hatalı negatif ise gerçekten anormal bir olayın yanlışlıkla normal olarak sınıflandırılması anlamına gelir. Bu da gerçek tehditlerin gözden kaçırılmasına ve güvenlik açıklarına yol açabilir. Hatalı pozitif veya hatalı negatif sonuçlar, anomali tespit sistemlerinin doğruluğunu ve güvenilirliğini etkileyebilir. Bu nedenle, algoritmaların ve eğitim veri kümesinin dikkatlice ayarlanması, doğru hassaslık ve özgülük dengesinin sağlanması önemlidir.

Bu doktora tezinde, VGAT alanındaki kıyas kabul edilen veri kümeleri ve bu alanda yapılan çalışmalar değerlendirilerek, güçlü yönler ve sınırlamalar değerlendirilmektedir. Ayrıca, başarılı bir performans sunan bir metodoloji önerilmektedir. Bu tez sonuçları, video gözetimi alanında değerli bir katkı sağlamakla kalmayıp, daha güçlü ve etkili anomali tespit sistemlerinin geliştirilmesine de ışık tutacaktır.

1.1 Tezin Kapsamı

VGAT, bir video akışında meydana gelen beklenmeyen veya olağandışı olayları tespit etmeyi amaçlar. Bu tez, mevcut yaklaşımlar ve teknikler üzerine kapsamlı bir inceleme sunarak, literatür taraması ve önceki çalışmalar üzerinden bu alandaki gelişmeleri derinlemesine ele alır.

VGAT için kullanılan yöntemler, öğrenme türlerine dayalı ve algoritma türlerine dayalı olmak üzere iki ana grupta sınıflandırılmış ve detaylı bir şekilde incelenmiştir. Öğrenme türlerine dayalı yaklaşımlar denetimli, denetimsiz ve yarı-denetimli olarak alt gruplara ayrılırken, algoritma türlerine dayalı yaklaşımlar ise istatistiksel, sınıflandırma, yeniden oluşturma, tahmin ve diğer yöntemler olarak sınıflandırılmıştır. Her bir sınıflandırmanın avantajları ve dezavantajları detaylı bir şekilde açıklanmıştır.

Ayrıca, bu tez çalışmasında, VGAT alanında yaygın olarak kullanılan veri kümeleri üzerine detaylı bir inceleme yapılmıştır. Bu veri kümelerinin her biri, özellikle anormallik tespiti için kullanılan çeşitli senaryolar ve ortamlar açısından kapsamlı bir şekilde karşılaştırılmıştır. Her veri kümesinin avantajları ve dezavantajları, bu alanın gereksinimleri ve zorlukları bağlamında değerlendirilmiştir. Bu değerlendirme, VGAT alanında kullanılan veri kümelerinin çeşitliliğini ve bu kümelerin her birinin özgün niteliklerini vurgulamakta, aynı zamanda bu veri kümelerinin algoritma geliştirme ve test süreçlerindeki rolünü ve önemini ortaya koymaktadır.

Bu çalışmada ayrıca, önerilen VGAT metodolojisinin geliştirme süreçleri detaylı bir şekilde ele alınmıştır. Metodolojinin tasarım aşamaları, modelin geliştirilmesi ve kullanılan parametreler açıklanmıştır. Ayrıca, önerilen metodolojinin test sonuçları, alanındaki benzer çalışmalarının sonuçları ile karşılaştırılmıştır. Bu karşılaştırmalar, önerilen metodolojinin güçlü yönlerini ve potansiyel geliştirme alanlarını vurgulamakta, aynı zamanda VGAT alanındaki çözüm yaklaşımları ile olan ilişkisini ortaya koymaktadır.

Bu tezin kapsamında yer alan araştırma, analiz, değerlendirme ve uygulamalar ile VGAT alanındaki araştırmalara yenilikçi bir perspektif ve yol gösterici bir yaklaşım sunmayı hedeflemektedir.

1.2 Tezin Özgün Değeri

Bu tez çalışması, VGAT alanında gelecekteki araştırmalara rehberlik edecek özgün bir metodoloji sunmaktadır. Bu metodoloji, verinin ön işleme aşamalarıyla hazırlanması, ConvLSTM-AE modeli ile özelliklerin öğrenilmesi ve anormal olayların zaman dilimlerinin belirlenmesi temeline dayanmaktadır.

Önerilen metodoloji, verinin titizlikle hazırlanması için ardışık karelerin çıkarılması, yeniden boyutlandırılması, arka planın etkili bir şekilde çıkarılması ve gürültülerin azaltılmasını içeren ön işleme aşamalarıyla başlamaktadır. Özellikle, arkaplanı çıkarılmış görüntülerle eğitilen modelin hatalı pozitifleri belirgin şekilde azalttığı görülmüştür. Böylece belirgin bir başarı artışı elde edilmiştir.

Verinin uzamsal ve zamansal özelliklerini öğrenme amacıyla ConvLSTM-AE modeli kullanılmıştır. Önerilen modelde veriye farklı evrişimli işlemler ve LSTM katmanları uygulanarak, verinin özellikleri belirginleştirilmiş ve sıkıştırılarak gizli bir kod elde edilmiştir. Uygulanan modelin katmanları ve parametreleri, en iyi sonuçları elde etmek amacıyla titizlikle ayarlanmıştır. Bu sayede, veri kümesinin özgün yapısından en iyi şekilde yararlanılarak, anormal durumların etkili bir şekilde belirlenmesi hedeflenmiştir.

Bir diğer özgün değer, anomali tespiti aşamasında model tarafından üretilen düzenlilik puanlarının etkin bir şekilde değerlendirilmesinin, yerel minimum noktaların ve anormal olayların zamanlamasının kritik bir şekilde belirlenmesine katkı sağlamasıdır. Bu puanlar, her bir zaman dilimindeki özgün özellikleri ve değişiklikleri yansıtarak, beklenen “normal” olayların dışında gerçekleşen durumları tanımlamak için kullanılır. Bu süreç, tespit edilen yerel minimum noktalar ile bu noktalara en yakın yerel maksimum noktalar arasındaki bölgelerin, anormal olayların meydana geldiği bölgeleri işaret etmesi açısından önem taşır. Bu aşamalar, anormal olayların etkili bir şekilde tespit edilmesine önemli katkı sağlar.

Ayrıca önerilen metodolojinin, VGAT alanındaki anomali tespiti için sunduğu çerçeve, farklı veri kümeleri ve uygulama senaryolarında uyarlanabilir. Modelin başarı oranını artırmak için, farklı veri kümelerine uyum sağlamak üzere ön işleme adımları ve model parametreleri kolaylıkla ayarlanabilir. Bu, metodolojinin farklı çevrelerde ve uygulama alanlarında başarıyla kullanılabilmesini sağlar.

Sonuç olarak, önerilen anomali tespit metodolojisi, CUHK Avenue ve UMN veri kümeleri üzerinde gerçekleştirilen testler ile değerlendirilmiştir. Bu metodolojinin kare düzeyinde yüksek doğrulukla anormallikleri tespit edebildiği kanıtlanmıştır. Bu başarıya ek olarak, tez içeriğinin gelecekteki araştırmalara VGAT teknolojilerinin geliştirilmesi konusunda rehberlik etmesi ve ilgili sektördeki profesyoneller için değerli bir referans kaynağı olması hedeflenmektedir.

1.3 Tezin Yapısı

Bu bölümde, çalışmanın temel yapısı ve içeriği ayrıntılı bir şekilde açıklanmaktadır. Tez, “Giriş” bölümünü takip eden ikinci bölümde, VGAT alanındaki bilgi birikiminin kapsamlı bir analizi ile başlamaktadır. Bu analizde, mevcut yaklaşımların güçlü yönleri vurgulanırken, aynı zamanda karşılaşılan sınırlamalar da detaylı bir şekilde ele alınmıştır.

Üçüncü bölüm, “Yaygın Kullanılan Yaklaşımlar ve Yöntemler” başlığını taşımaktadır. Bu bölüm, VGAT alanındaki temel yaklaşımları, öğrenme türlerine ve algoritma türlerine dayalı olarak sınıflandırarak incelemektedir. Denetimli, denetimsiz ve yarı-denetimli öğrenme yöntemleri ayrıntılı bir şekilde ele alınmakta; ayrıca, istatistik tabanlı, sınıflandırma tabanlı, yeniden yapılandırma tabanlı, tahmin tabanlı ve diğer algoritma türlerine dayalı yaklaşımlar detaylı bir analize tabi tutulmaktadır.

Dördüncü bölüm, “Yaygın Kullanılan Veri Kümeleri” başlığı altında sunulmuştur. Bu bölüm, literatürde ve bu tez çalışmasında kullanılan temel veri kümelerinin ayrıntılı bir

açıklamasını içermektedir. Her veri kümesinin özellikleri, dağılımları ve potansiyel kullanım senaryoları kapsamlı bir şekilde ele alınarak sunulmaktadır.

Beşinci bölüm, “VGAT Metodolojisinin Geliştirilmesi”ne odaklanmaktadır. Bu bölümde, tez içinde önerilen metodolojinin nasıl geliştirildiği detaylı bir şekilde incelenmektedir. Bu kapsamda, veri hazırlama süreçleri, modelin tasarımı ve geliştirilmesi, parametre ayarlamaları gibi konular derinlemesine ele alınmaktadır. Ayrıca, bu metodolojinin etkinliğini ölçmek için kullanılan değerlendirme metrikleri de bu bölümde ayrıntılı olarak tartışılmaktadır.

Altıncı bölümde tez kapsamında önerilen metodolojinin gerçek dünya veri kümeleri ile nasıl test edildiği ve uygulandığı detaylıca ele alınmaktadır. Bu bölümde, metodolojinin uygulama sonuçları, etkinliği ve doğruluğunu değerlendirmek için kullanılan kriterler ve değerlendirme yöntemleri üzerine yoğunlaşmaktadır.

Sonuç bölümü, tezin temel amaçlarını ve ulaşılan sonuçları kapsamlı bir şekilde özetlemektedir. Bu bölüm, tezin VGAT alanına sağladığı katkıların yanı sıra, bu çalışmanın genel önemini ve etkisini vurgulamaktadır. Ayrıca, bu bölümde, elde edilen bulguların ışığında, gelecekteki araştırma yolları ve metodolojinin potansiyel geliştirmeleri için öneriler ve yön gösterici fikirler sunulmaktadır. Bu öneriler, alanın ilerleyişini şekillendirmede ve gelecekteki çalışmalar için bir temel oluşturmada önemli bir rol oynamaktadır.

2. KAYNAK ÖZETLERİ VE LİTERATÜR DEĞERLENDİRMESİ

VGAT alanında kullanılan çeşitli yöntemler ve teknikler, video verilerini analiz ederek anormal durumları tespit etme amacını taşır. Literatürde bu anormal durumları oluşturan “anomali” veya “aykırı” kavramına birçok farklı tanım getirilmiştir. İlk tanımlardan biri olan Frank E. Grubbs’un (1969) tarifine göre “aykırı” değer, diğer verilere göre belirgin bir şekilde sapma gösteren bir gözlem olarak nitelendirilir. Bu tür aykırı değerler, genellikle istatistiksel analizlerde özel bir ilgi gerektiren istisnai durumlar olarak öne çıkar.

Hawkins (1980) “aykırı” gözlemleri, diğer gözlemlerden belirgin şekilde farklılaşan ve farklı bir mekanizma tarafından oluşturulabileceği şüphesini uyandıracak düzeydeki gözlemler olarak tanımlamıştır. Ona göre bu aykırı değerler, analiz edilen veri kümesinin genel yapısını etkileyen ve sıradanlığın ötesine geçen özel durumları temsil etmektedir.

Barnett ve Lewis (1984) ise “anomali”yi, diğer gözlemler veya gözlem alt kümeleri ile tutarsızlık oluşturan durumlar olarak tanımlar. Bu tanıma göre, anomali diğer verilere göre aykırı bir davranış sergileyen veya beklenen desene uymayan özel bir durumu ifade eder.

Bu tanımların ortak yönü, aykırı değerlerin diğer verilere göre belirgin bir sapma gösterdiği, istatistiksel analizlerde özel bir öneme sahip olduğu ve genellikle beklenmeyen durumları temsil ettiği yönündedir. Bu değerler, veri kümesinin doğasına bağlı olarak değişken mekanizmalar ya da istisnai durumlar tarafından oluşturulabilir ve sıklıkla anormal olarak kabul edilir.

VGAT alanında, bu farklı aykırı değer tanımları temel alınarak farklı yöntemler ve teknikler kullanılarak anomali tespiti gerçekleştirilmektedir. Derin öğrenme tabanlı yaklaşımlar, geleneksel istatistiksel modeller, hareket analizi gibi çeşitli teknikler, video akışlarından anlamlı bilgileri çıkarmak ve anormal davranışları saptamak için

kullanılmaktadır. Bu yöntemler, geniş ve karmaşık veri kümelerini işleme kapasiteleri sayesinde VGAT alanında önemli bir ilerleme sağlamıştır. Bu bölümde, VGAT yöntemlerinin ve tekniklerinin, VGAT alanında nasıl ortaya çıktığına ve çeşitli senaryolarda nasıl uygulandığına odaklanılacaktır.

Video gözetimi alanındaki önemli bir hedef, anomali tespiti yöntemleri aracılığıyla normalden sapmaları tanımlamak ve anlamaktır. Bu bağlamda, farklı senaryo ve durumlarda, aynı eylem örüntülerinin önceden edinilmiş örüntülerden sapma göstermesi, potansiyel bir anormal durumun işareti olarak kabul edilmektedir. Bu nedenle, literatürde çoğu anomali tespit tekniği, önceden edinilmiş örüntülerden sapma gösteren bir eylemin anormal olarak değerlendirildiği hipotezine dayanmaktadır (Wan vd. 2021). Ayrıca, Aldayri ve Albattah (2022), Chandola ve arkadaşları (2009) ile Pannirselvam ve arkadaşları (2021) anomalilerin genellikle üç temel tipte sınıflandırılabileceğine dikkat çekmiştir:

1. Nokta anomalileri: Bu anomaliler, tek bir veri noktasının diğer verilere göre düzensiz davrandığı durumları ifade eder. Bu tür anomaliler, genellikle veri kümesindeki diğer noktalardan belirgin bir şekilde farklı davranan ve istatistiksel olarak nadir olan noktalar olarak tanımlanır. Örneğin, bir aracın normal trafik akışına uygun olmayan bir şekilde yolun ortasında durması, nokta anomalisi olarak kabul edilebilir.

Nokta anomalileri, daha karmaşık anomalilerin keşfedilmesinde de önemli bir rol oynayabilir. Bu nedenle, nokta anomalileri üzerinde çalışmak, genel anomali tespiti çalışmalarının temelini oluşturur ve daha geniş bir anomali tespiti sürecinin bir parçası olarak değerlendirilir.

2. Bağlamsal anomaliler: Bir veri değerinin belirli bir bağlamdaki diğer verilere göre düzensiz davrandığı durumları ifade eder. Bu tür anomaliler, gözlemcinin kişisel yargısı ve duruma yönelik algısıyla yakından ilişkilidir. Veri noktasının çevresindeki diğer verilere ve olaylara bağlı olarak bağlam değerlendirilir. Bağlamsal anomaliler genellikle tek bir veri noktasının beklenen bağlama göre anormal davrandığını gösteren

bir referans noktasına dayanırlar. Örneğin, binek bir aracın otobüslerin park ettiği bir alana park etmesi, bağlamsal bir anomali örneği olarak düşünülebilir.

Bağlamsal anomaliler, genellikle karmaşık ve dinamik veri kümelerinde ortaya çıkar ve genel anomali tespiti sürecinin bir parçası olarak ele alınır. Bu tür anomalilerin doğru bir şekilde tanımlanması ve tespit edilmesi, anormal durumları belirlemek ve gerektiğinde önlem almak için önemlidir.

3. Kolektif anomaliler: Bu anomaliler, bir veri kümesi örneğinin gerçek verilere kıyasla anormal olarak değerlendirildiği durumları ifade eder. Bu tür anomaliler, genellikle veri noktalarının bir araya gelerek oluşturduğu bir kümenin istatistiksel olarak belirgin bir sapma sergilediği durumlarda ortaya çıkar. Örnek olarak, belirli bir alanda beklenmedik bir şekilde bir kalabalık oluşması kolektif bir anomali olarak kabul edilebilir. Normalde, insanlar belirli bölgelerde veya etkinliklerde belirli bir dağılım gösterirken, bu grup insanlar beklenmedik bir şekilde bir araya gelerek normallikten sapmaktadır. Bu durum, genel veri kümesiyle karşılaştırıldığında anormal bir durumu temsil eder.

Kolektif anomalilerin tespiti genellikle veri kümesinin yapısı ve örüntüleri üzerinde yoğunlaşır. Bu tür anomaliler, veri noktalarının ilişkilerini ve gruplamalarını analiz ederek belirlenebilir. Kümeler arasındaki benzerlik, yoğunluk, uzaklık veya diğer özellikler kullanılarak kolektif anomalilerin algılanması sağlanır.

Bu farklı anomali türlerinin her biri, VGAT alanında özgün tespit yaklaşımları gerektirir ve genel anomali tespiti sürecinin bir parçası olarak değerlendirilir. Anomalilerin doğru bir şekilde tespiti, güvenlikten trafik yönetimine ve kalabalık kontrolüne kadar pek çok uygulama alanında büyük öneme sahiptir.

Farklı VGAT yaklaşımları, kalabalık ve kalabalık olmayan ortamlarda gerçekleşen olayların etkili bir şekilde analiz edilmesine yeni bakış açısı sunmaktadır. Bu bağlamda, özellikle Abdelghafour ve ekibinin (2021) çalışması, hareketli nesnelerin sayısı ile

videodaki anormallikler arasındaki ilişkiyi inceleyerek ortamın karmaşıklığını belirlemiş ve farklı kalabalık düzeylerini tanımlamıştır. Bu çalışma kalabalık düzeyine göre üç farklı ortamı tanımlamıştır:

1. Hafif kalabalık ortam: Bu ortamda kişi başına 0,92 metrekare alan düşer. Bu durum, biraz kalabalık olabilen ancak genel olarak rahatlıkla hareket edilebilen bir ortamı temsil eder.
2. Orta derecede kalabalık ortam: Bu ortamda kişi başına 0,41 metrekare alan düşer. Bu durumda, daha yoğun bir kalabalık söz konusudur ve kişilerin hareket etmesi biraz daha zorlaşabilir.
3. Kalabalık ortam: Bu ortamda kişi başına 0,23 metrekare alan düşer. Bu durum, yoğun bir nüfusun bulunduğu ve kişilerin sınırlı bir alanda hareket etmek zorunda kaldığı bir ortamı yansıtmaktadır.

Bu yaklaşım, videolardaki kalabalık seviyelerinin ve buna bağlı olarak anormalliklerin tespit edilmesi için kullanılabilir. Örneğin, belirli bir videodaki kişi sayısı ile toplam alan miktarı analiz edilerek, ortamdaki kalabalık seviyesi tahmin edilebilirken, beklenen seviyeden sapmalar da tespit edilebilir. Bu tür anomali tespiti, güvenlik, trafik izleme ve olay analizi gibi birçok alanda uygulanabilir potansiyele sahiptir.

VGAT, genellikle uzman sistemlerin başa çıkamadığı aykırı değerlerle mücadelede zorluklar doğurabilmektedir. Aykırı değerlerin veri kümelerinde yol açtığı sınıflandırma problemleri oldukça yaygındır (Wilmet vd. 2021). Bu tür problemler, uygulamaların etkinliğini düşürebilir ve özellikle sınırlı örnek içeren sınıfların doğru sınıflandırılmasını engelleyebilir. Aykırı değerlerin doğru bir şekilde tanınması ve sınıflandırılması, daha güvenilir ve hassas makine öğrenimi (ML) ve uzman sistemleri geliştirme potansiyelini taşır. Bilgisayar görüşü ve görüntü işleme alanındaki başarılarından ilham alan araştırmacılar, ML veya derin öğrenme (DL) yöntemlerini anormal olay tespitine uyarlamaya yönelik çalışmalara başlamışlardır (C. H. Chen 2015). Munyua ve

arkadaşları (2021) yaptıkları araştırma ile VGAT gibi karmaşık bilişsel görevin, ML ve DL yöntemleri kullanarak başarıyla gerçekleştirilebildiğini belirtmişlerdir.

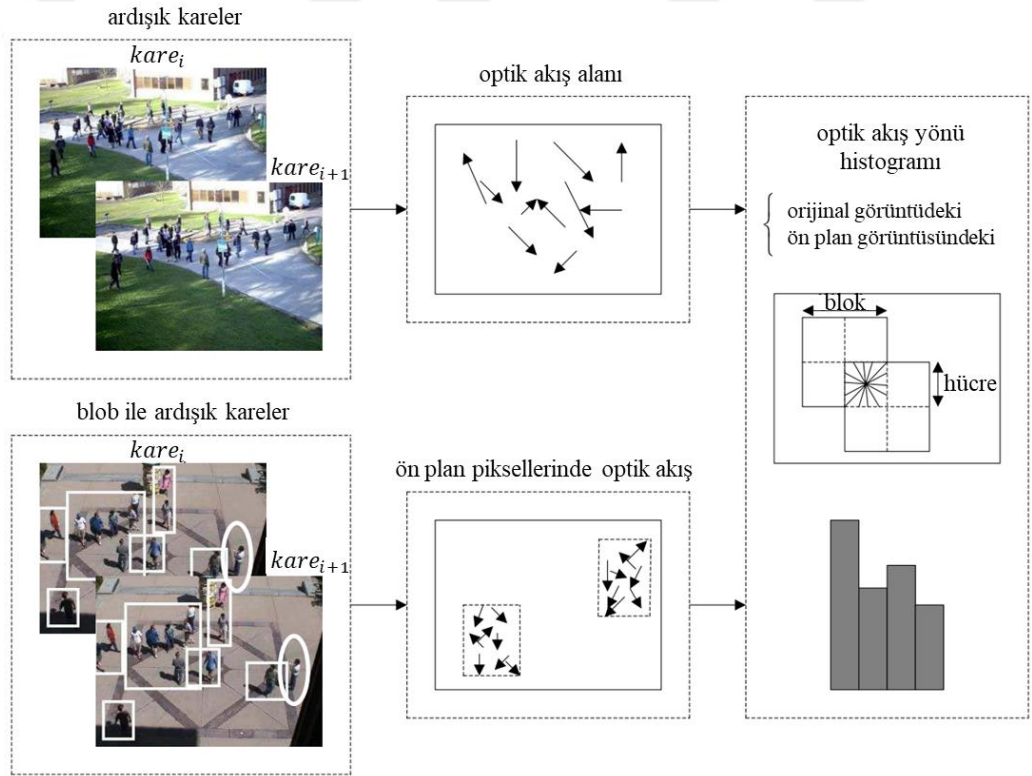
VGAT, görüntülerdeki anormal kareleri veya piksel parçalarını uzamsal ve zamansal verilere göre bulmaya çalışır (Le ve Kim 2022). Uzamsal veriler tek bir kareden, zamansal veriler ise nesne hareketlerinden ve kare sıralamasından öğrenilir. Liu ve arkadaşlarına (2022) göre videodaki anormallikler üç aşamada tespit edilmektedir: (1) Düzenli ve düzensiz olayların özellikleri ortak bir alana yansıtılır ve anormallik uzamsal dağılımın marjıyla belirlenir. (2) Olay kalıplarının anlamsal özellikleriyle bir sözlük eğitilir. Bu sözlük anormallik hesabında kullanılır. (3) Anomaliler, eğitilmiş özellik çıkarıcılarla önceki veya sonraki karelerin tahmini ve yeniden yapılandırılması sırasındaki hatalardan tespit edilir.

Mo ve arkadaşları (2013), VGAT için yeni bir sparsity modeli sunmuşlardır. Bu modelde, eğitim sözlüğü çoklu nesne olaylarının gözlemlerinden oluşur ve test olayları, eğitim sözlüğündeki olayların sparsity (seyrek) lineer bir kombinasyonu ile yeniden oluşturulur. Ortak sparsity modelinin anlamlı olması için kısıtlamalar getirilir ve denetimli durumlarda anomali önceden sınıflandırılırken, denetimsiz senaryolarda anomali önceden belirlenemeyebilir. Önerilen model, veri kümesindeki önemli özellikleri veya gözlemleri az sayıda değerle ifade ederek, veri kümesinin karmaşıklığını ve boyutunu azaltmaya çalışır. Bu durum modelin gerçek dünya veri kümelerinde etkili olduğunu ve lineer sparsity modellerine göre daha üstün bir anormallik tespiti performansı sergilediğini göstermiştir. Bunun yanı sıra, geçerli olmayan durumlarda anomali tespitini artırmak amacıyla bir çekirdek (kernelizasyon) yöntemi de önerilmiştir.

Y. Yuan ve arkadaşları (2014), insanların toplu halde sergilediği anormal davranışları tanımlamak için yeni bir yöntem önermiştir. Bu metodolojide, insanların toplu davranışlarını belirlemek için bilgilendirici yapısal bağlam tanımlayıcıları (SCD) kullanılmaktadır. SCD, parçacıklar arasındaki kuvvetlerin potansiyel enerji fonksiyonunu görsel bağlamsal ipuçlarını daha hassas bir şekilde yansıtabilecek şekilde hesaplamaktadır. Ayrıca, etkili bir biçimde SCD değişimini izlemek ve farklı

karelerdeki hedefleri güvenilir bir şekilde eşleştirmek için dayanıklı bir çoklu nesne izleyici geliştirilmiştir. Anormallikler, topluluğun SCD değişiminin çevrimiçi olarak analiz edilmesiyle belirlenmektedir. Bu çalışma, VGAT için ilişkiyi iyi temsil eden bir SCD tanımı ve 3 boyutlu ayırık kosinüs dönüşümü (3D discrete cosine transform) çoklu nesne izleyicisi ile sınırlı hedefleri dayanıklı bir şekilde eşleştiren bir yöntem sunmaktadır.

Wang ve Snoussi (2014)'nin tarafından önerilen VGAT algoritması, iki temel bileşeni içermektedir. Bu bileşenler optik akış yönü histogramının (HOFO) hesaplanması ve sınıflandırma aşamasında non-linear bir destek vektör makinesinin (SVM) uygulanmasıdır. Şekil 2.1'de temsili gösterilen HOFO yöntemi, arka plan çıkarımı sonrasında elde edilen ön plan görüntüsü ile orijinal görüntünün hızlı bir şekilde analizini yapabilme yeteneği sunmaktadır. Ayrıca, yazarlar bu yöntemin farklı ilgi alanlarına veya takip edilen nesnelere göre uyarlanabileceğini ve makine öğrenimi tekniklerinin performansını artırmak için kullanılabileceğini özellikle vurgulamışlardır.



Şekil 2.1 Arka plan çıkarma işlemi uygulandıktan sonra elde edilen orijinal global görüntünün ve ön plan görüntüsünün optik akış yönü histogramı (T. Wang ve Snoussi 2014)

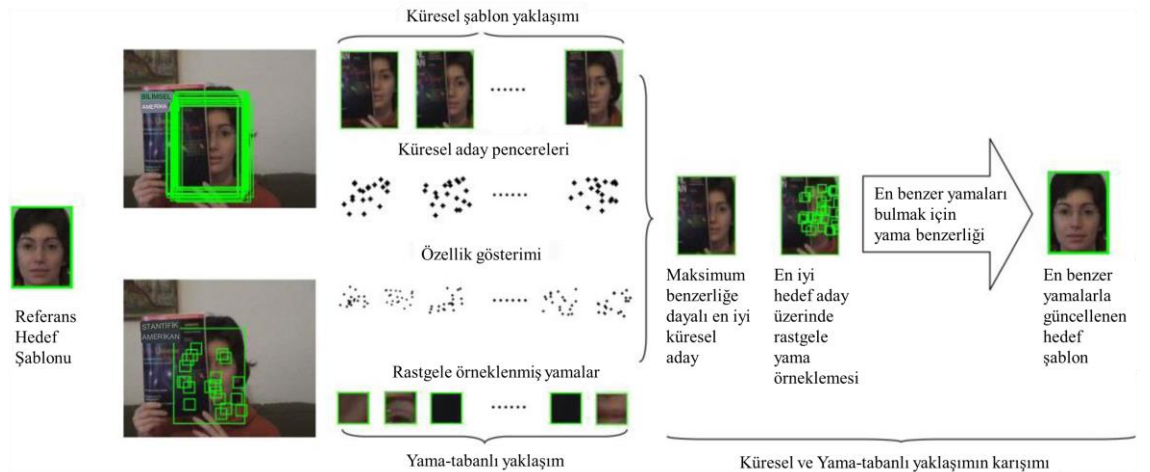
Kaltsa ve arkadaşlarının (2015) çalışması, kalabalık sahnelerdeki anormal olayları hem hareket hem de görünüm açısından analiz eden yeni bir yöntem sunmaktadır. Bu yaklaşım, sürü teorisi ile uyumlu olan “Yönlendirilmiş Sürü Histogramları (HOS)” kavramını kullanmaktadır. HOS, kalabalık ortamların dinamiklerini yakalamaya yönelik özgün bir özellik çıkarma metodudur. Yönlendirilmiş Gradyanlar Histogramı (HOG) ile birleştirilen HOS, her sahne için güçlü bir tanımlayıcı sağlamaktadır. Bu önerilen yöntem, yerel gürültülere karşı direnç gösterirken baskın olmayan anormallikleri yüksek doğrulukla tespit edebilme yeteneğine sahiptir. Ayrıca, hesaplama maliyeti düşük olup pratik bir çözüm sunmaktadır. Yöntemin genel yapısı ve gerçek hayat uygulamalarına uygunluğu, yüksek tespit başarısı, düşük hesaplama maliyeti ve farklı ortamlara uyum sağlayabilmesi, çeşitli gözetim uygulamalarına uygun olduğunu göstermektedir.

Isupova ve arkadaşları (2016) tarafından VGAT için yeni bir dinamik Bayesian parametrik olmayan konu modeli önerilmiştir. Bu model, öğrenme aşamasında batch ve online Gibbs örneklemeye yöntemlerini benimsemiş ve karar verme sürecinde yeni bir anormallik ölçüsü sunmuştur. Araştırmada, sentetik ve gerçek veri kümeleri kullanılarak önerilen yöntem değerlendirilmiş ve elde edilen sonuçlar, dinamik modelin non-dinamik modellere kıyasla anomali tespitinde sınıflandırma performansı açısından daha iyi sonuçlar ürettiği gösterilmiştir.

Colque ve arkadaşları (2018), VGAT alanında 3 farklı yöntem sunmuştur. Bunlar video verilerinden elde edilen farklı öznitelikleri (renk, hareket, yön vb.) birleştirerek ve ağırlıklandırarak anomali skorları üreten “Çoklu Öznitelikli Anomali Tespiti”, farklı öğrenme algoritmalarını (destek vektör makineleri, karar ağaçları, rasgele ormanlar vb.) bir araya getirerek ve oylama yoluyla anormallik skorları üreten “Çoklu Öğrenme Tabanlı Anomali Tespiti” ve farklı görünümlerden (ön, yan, arkadan vb.) elde edilen verileri birleştirerek ve optimizasyon yoluyla anormallik skorları üreten “Çoklu Görünüm Tabanlı Anomali Tespiti”dir. Bu yöntemler, farklı veri kaynaklarını, algoritmaları ve görünümünü birleştirerek, video verilerindeki gürültü, eksiklik, bozukluk ve değişkenliğe karşı sağlam anormallik skorları üretmektedir. Bu yöntemler hem küresel hem de yerel hem doğrusal hem de doğrusal olmayan hem tekli hem de çoklu kamera senaryolarında çalışabilen anomali tespiti algoritmalarıdır.

Sabokrou ve arkadaşlarının araştırması (2018), VGAT ve yerelleştirilme için tam evrişimli sinir ağı (FCN) ile zaman verilerini birleştiren bir yöntem önermektedir. Bu yaklaşımda, önceden eğitilmiş bir FCN, sahnelerdeki genel anormallikleri belirlemek amacıyla denetimsiz bir FCN'ye bağlanmaktadır. Bu FCN tabanlı model, özellik temsili ve ardışık aykırı değer tespiti işlevlerini etkin bir şekilde gerçekleştirerek hızlı ve doğru sonuçlar elde etmektedir. Bu önerilen yöntem, saniyede yaklaşık 370 kare işleme hızı (FPS) sunma yeteneği ile dikkat çekmektedir.

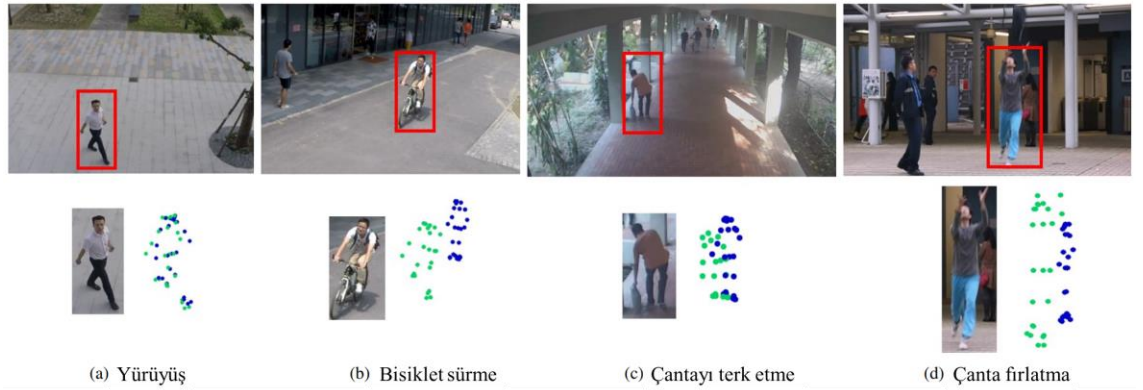
Parate ve arkadaşları (2018), GPHybrid adlı yeni bir keyfi nesne izleme yöntemi sunmuştur. Bu yöntem, Küresel Şablon Takibi (Global Template Tracking) ile Yama-Tabanlı Şablon Takibi (Patch-Based Template Tracking) yaklaşımlarını birleştirerek, zorlu senaryolarda güçlü bir izleme yeteneği sağlamaktadır. Hedef hakkında çeşitli kaynaklardan bilgi toplamak için Entegre Kanal Özellikleri (Integral Channel Features-ICF) kullanılırken, hibrit şablonun özellik yeteneklerini birleştirir. Hedef görünümünün temsili, benzer görünümlü örneklerin kümelemesi ve vektör kantilasyonu kullanılarak güncellenir. Şekil 2.2, Küresel ve Yama-tabanlı yaklaşımların birleşimini göstermektedir. Üst bölüm, küresel şablon tabanlı yaklaşımı temsil eder ve küresel hedef aday pencerelerini içerir. Alt bölüm ise, Yama-tabanlı şablon yaklaşımını yansıtır ve rastgele örneklenmiş yamaları içerir. Sağ tarafta, en iyi eşleşen hedef aday penceresini seçen ve aynı zamanda ICF hesaplaması için en iyi eşleşen hedef aday penceresinden rastgele yamaları seçen bir kombinasyon görülmektedir.



Şekil 2.2 Küresel ve Yama-tabanlı yaklaşımların kombinasyonu (Parate vd. 2018)

Güvenlik kameraları için video özetleme çalışmaları (Başkurt ve Samet 2019; Başkurt ve Samet 2018) da bulunmaktadır. Bu çalışmalar, sadece aktif etkinlikleri içeren sıkıştırılmış bir görünüm oluşturarak, video taramasını verimli hale getirmeyi hedeflemektedir. Bu yaklaşım, VGAT alanında anormal olayların özetlenmesinde kullanılmakta ve hatalı negatifleri daha etkin bir şekilde azaltmaya yardımcı olmaktadır.

Rodrigues ve arkadaşlarının çalışması (2020), anormal insan aktivitelerinin farklı zaman ölçeklerinde tespiti için bir model önermektedir. Bu model, verilen bir pozisyon takibi girdisi temel alınarak, farklı zaman ölçeklerinde gelecek ve geçmiş tahminlerini üretmek amacıyla ara katmanlar kullanmaktadır. Bu tahminler, anormal aktivitelerin tespiti için değerli bilgiler sunarak kullanılmaktadır. Şekil 2.3'teki ilk sahnede yürüme eylemi varken diğer üç sahnede anormal eylemler vardır. Yeşil ve mavi pozlar sırasıyla gerçek ve öngörülen pozları temsil etmektedir. Model, anormal pozlar için büyük tahmin hataları üretmektedir.

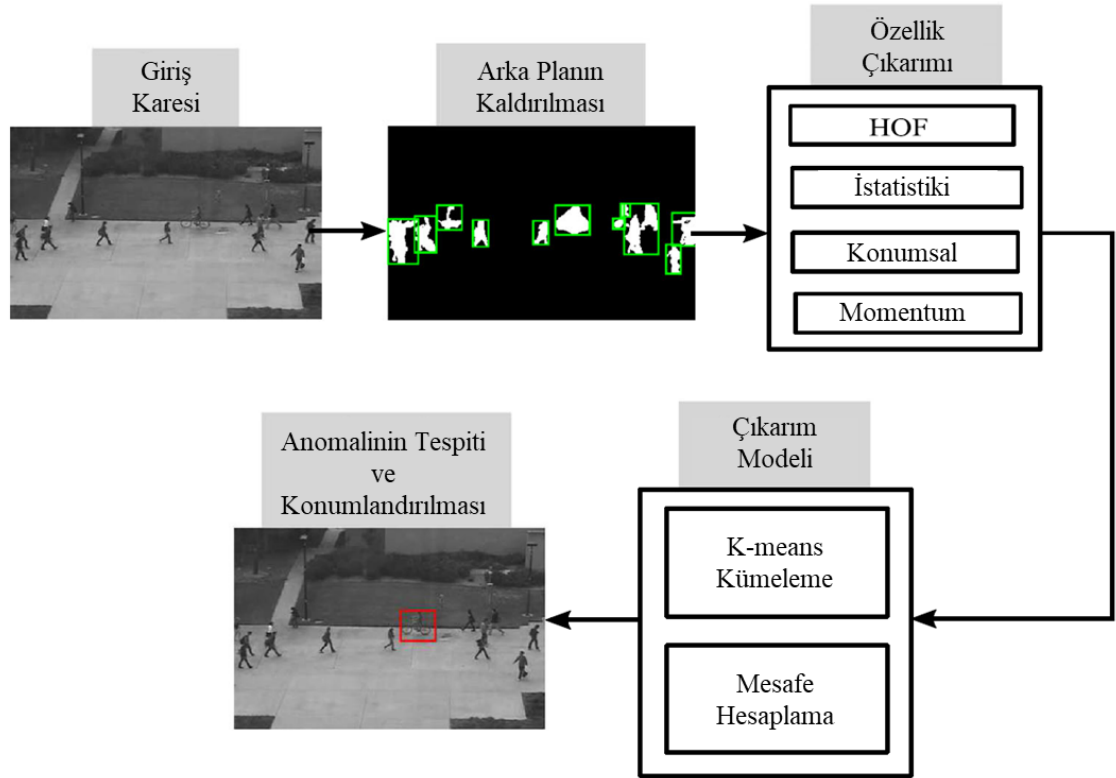


Şekil 2.3 Önerilen modelden tahmin edilen pozların bir örneği (Rodrigues vd. 2020)

Ramachandra ve meslektaşlarının (2020) VGAT alanında önerdikleri yöntem, Siamese ağı olarak adlandırılan bir yapay sinir ağı yapısına dayanmaktadır. Bu yöntem, iki veya daha fazla alt ağdan meydana gelir. Siamese ağı, girdiler arasındaki benzerlik veya farklılık düzeyini ölçmek için bir mesafe fonksiyonunu kullanarak çalışır. Bu benzersiz yaklaşım, verilerin içsel yapısını ve dağılımını yakalamayı amaçlayarak VGAT'ın sınıflandırılması için bir mesafe fonksiyonunun öğrenilmesini sağlar. Bu çerçevede, normal ve anormal video kareleri arasındaki mesafeyi hesaplamak amacıyla eğitilmiş

bir ağ kullanılır. Elde edilen bu mesafe fonksiyonu, yeni videolarda anormallikleri etkili bir şekilde belirlemek için kullanılmaktadır.

Bansod ve Nandedkar (2020), anomalinin tespiti ve konumlandırılması için Şekil 2.4'teki yaklaşımı sunmuşlardır. Bu yöntem, üç temel adımdan oluşmaktadır: (1) arka planın kaldırılması, (2) özellik çıkarımı ve davranış tanıma ve (3) anomali tespiti ve yer belirleme. Ön plandaki nesnelerin görünümü ve hareketi, her karede arka planın çıkarılmasıyla elde edilen momentum ve büyüklük histogramı gibi özelliklerle birleştirilir. Nesnelerin davranışları, denetimli olmayan bir kümeleme yöntemi kullanılarak öğrenilir. Anomali tespiti için ise pozisyonel özellikler değerlendirilir. Bu yöntem UMN veri kümesinde en yüksek AUC değeri olan %98,26'ya ulaşmıştır. Genel olarak bu yöntem, momentum özellikleri ve büyüklük histogramı gibi özellikleri kullanarak anomali tespitini etkili bir şekilde gerçekleştirmenin potansiyelini göstermektedir.



Şekil 2.4 Önerilen anomali tespit ve konumlandırma yaklaşımı (Bansod ve Nandedkar 2020)

Z.Li ve arkadaşları (2020), anormal insan davranışlarını video verilerinde tespit etmek amacıyla çok ölçekli 3D evrişimli sinir ağı kullanarak bir uzamsal-zamansal temsil öğrenme modeli geliştirmiştir. Bu model, arka plan sahnesindeki anomalileri tespit etmek üzere yüksek boyutlu karışık Gaussian modeliyle uyumlu uzamsal-zamansal özellikleri çıkarmaktadır. Model, özel eğitime ihtiyaç duymadan yüksek performans ve doğruluk sunmaktadır. Geliştirilen algoritma, farklı veri kümelerinde test edilmiş ve gerçek zamanlı uygulamalarda 18 FPS hızında çalışma başarısı göstermiştir.

Xue ve arkadaşları (2020), video verilerinden anormallikleri veya beklenmeyen davranışları belirlemek amacıyla geliştirilmiş bir zaman serisi analizi yöntemi önermiştir. Bu yöntemde, başlangıçta video verilerinden nesnelerin hareketleri, hızı ve yönü gibi uzamsal ve zamansal ilişkileri içeren önemli özellikler çıkarılması için bir özellik çıkarım algoritması kullanılmaktadır. Elde edilen bu özellikler, daha sonra bir zaman serisi analizi yöntemi ile anormallikleri veya beklenmeyen davranışları belirlemek üzere analiz edilmektedir. Ayrıca, makale yüksek kaliteli video verilerinin gerekliliği ve zaman serisi analizinin hesaplama karmaşıklığı gibi zorlukları da ele almaktadır.

Park ve arkadaşları (2020), VGAT için bellek rehberli normalite öğrenen bir yöntem sunmuştur. Bu yöntem, normal veri kalıplarının çeşitliliğini açıkça dikkate alan ve evrişimli sinir ağlarının (CNN) temsil kapasitesini azaltan bir yöntemdir. Bu amaçla, normal verilerin prototipik kalıplarını kaydeden bir bellek modülü kullanılmaktadır. Ayrıca, belleği eğitmek için yeni özellik sıklığı ve ayrılığı kayıpları sunmaktadırlar. Bu kayıplar hem bellek öğelerinin hem de normal verilerden derinlemesine öğrenilen özelliklerin ayırt edici gücünü artırmaktadır. Özellik sıklığı kaybı, bellek öğelerinin birbirine yakın olmasını sağlarken, özellik ayrılığı kaybı, bellek öğelerinin normal verilerden farklı olmasını sağlamaktadır.

Georgescu ve arkadaşlarının (2021) sundukları çerçeve, arka plan yerine anormalliklere neden olabilecek nesnelere odaklanarak, arka plana bağımlı olmayan bir yaklaşım benimsemektedir. Anormal veri eksikliğini çözmek için düşmanca öğrenme stratejisi kullanan otomatik kodlayıcılar öneriliyor. Bu yöntem hem görünüm hem de hareket

tabanlı ikili sınıflandırıcılar için sahte anormal verilerle eğitim yapmayı içerir. Çerçeve, farklı sahnelerde uygulanabilirliği ve gözetim videolarında anormal olayları etkin bir şekilde tespit etme kapasitesi ile dikkat çekmektedir.

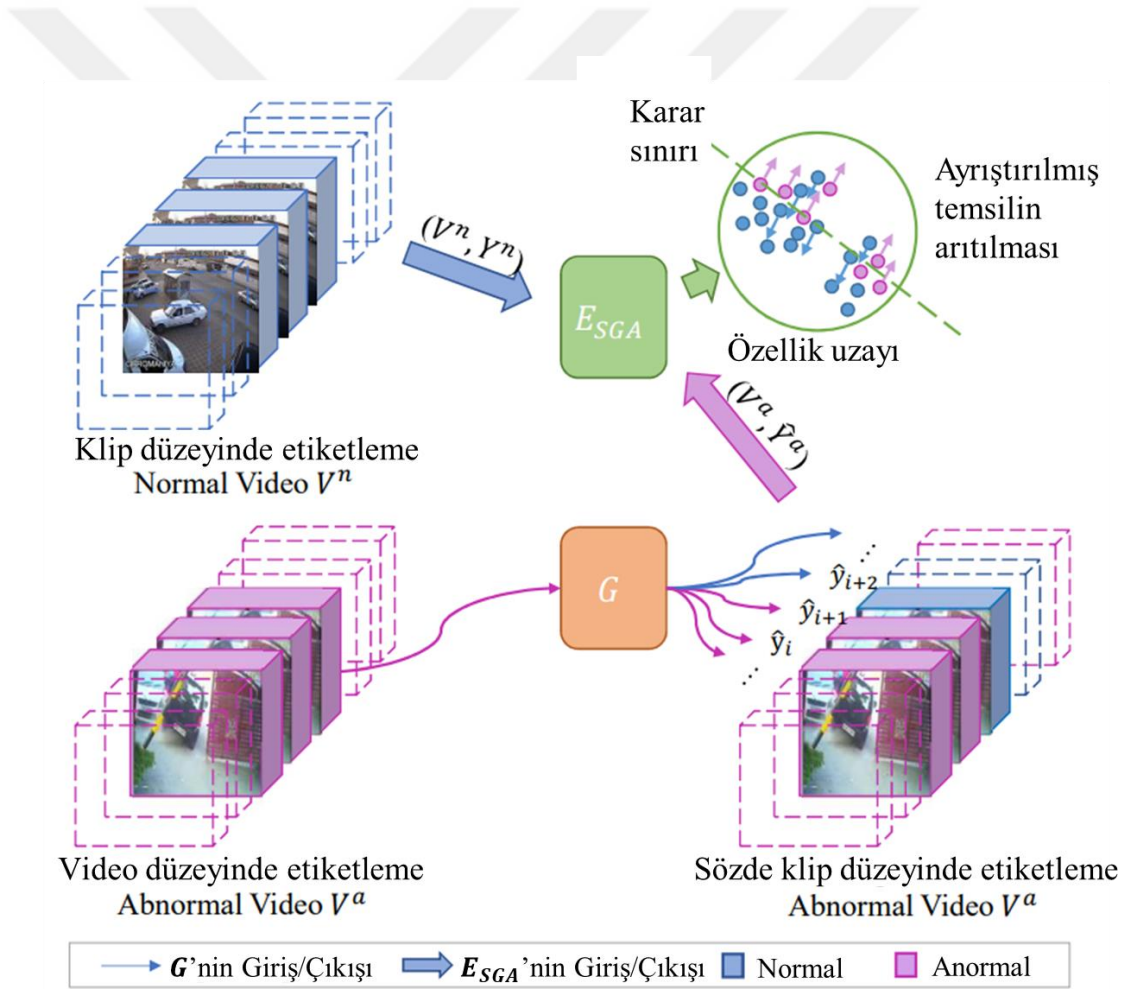
Majhi ve arkadaşları (2021), yalnızca video düzeyi etiketlerini kullanan bir zayıf denetimli öğrenme yöntemi önermiştir. Yapılan deneyler, dikkat ağırlıklarının sadece anormallik tespiti için girdi özelliklerini düzenlemekle kalmadığını, aynı zamanda görevler arasındaki karşılıklı bağımlılığı da sağlayarak anormallik sınıflandırmasına da yardımcı olduğunu göstermiştir.

Wan ve arkadaşlarının (2021) araştırması VGAT alanında yeni bir ölçüt ve model sunmuştur. Araştırmacılar, mevcut yöntemlerin gerçek dünya senaryolarında anormallikleri tespit etme konusundaki kısıtlamalarını vurgulamıştır. Bu problemi çözmek adına, gerçek dünya gözetim videolarından oluşturulan bir veri kümesi kullanılarak yeni bir ölçüt oluşturulmuştur. Model, hareket vektörleri, hareket yoğunluğu ve arka plan çıkarma gibi uzamsal ve zamansal özelliklerin birleşimini kullanarak anormallikleri tespit etmektedir.

Meta-prototip ağı kullanarak bir yöntem öneren Lv ve arkadaşları (2021), nesnelerin normal davranışlarını öğrenme ve gelecekteki hareketlerini tahmin etme amacıyla iki temel bileşen kullanmıştır. Bunlardan birincisi dinamik prototip birimi (DPU)'dur. DPU, normal dinamikleri prototipler olarak kodlama amacıyla tasarlanmıştır. Girdi kareleri ve yeniden yapılandırma hatalarına dayanarak prototipleri dinamik bir şekilde güncelleyebilme yeteneğine sahiptir. İkinci bileşen ise meta prototip birimi (MPU)'dur. MPU, meta-öğrenme modülüdür ve yeni sahnelerde hızlı bir şekilde uyum sağlama yeteneği sunar. MPU, farklı sahnelerden gelen prototipleri nasıl güncelleyeceğini az atışlı öğrenme stratejisi kullanarak öğrenebilir.

Feng ve arkadaşları (2021), zayıf denetimli video anomali tespiti (WS-VAD) için çalışma yapısı Şekil 2.5'te gösterilen Multiple Instance Self-Training Framework (MIST) adlı yöntemi önermişlerdir. WS-VAD, sadece video düzeyinde ek açıklamalarla ayırt edici temsiller öğrenerek normal olaylardan anomalileri ayırt etmeyi amaçlayan bir

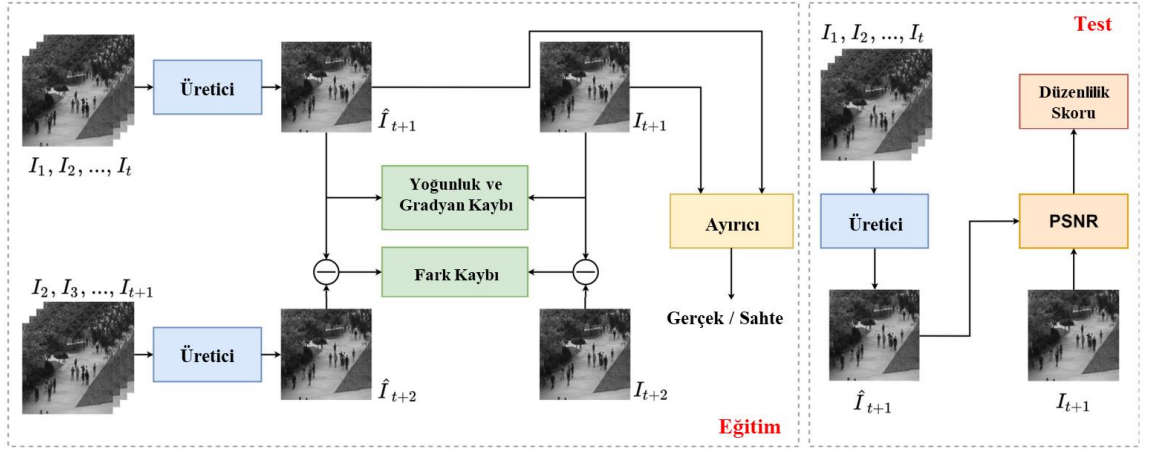
problemdir. MIST, videolardan seyrek ve sürekli klipler örnekleyerek sahte etiketler üreten bir sahte etiket üretici ve karelerdeki anomal bölgelere odaklanarak görev özel özellikler çıkaran bir dikkat artırılmış özellik kodlayıcı olmak üzere iki bileşenden oluşmaktadır. MIST, her iki bileşeni tekrarlayan bir şekilde optimize etmekte ve görev özel özellik kodlayıcısını elde etmek için kendinden eğitim prensibini benimsemektedir. Burada üretici (G)'nin yardımıyla anormal videolara $\hat{Y}^a = \{\hat{y}_i^a\}$. Daha sonra MIST, kendi kendine yönlendirilen, dikkati artıran özellik kodlayıcı E_{SGA} 'yı geliştirmek için tüm videolardan elde edilen bilgilerden yararlanır. MIST, ShanghaiTech veri kümesinde kare düzeyinde %94,83 AUC (Alan Altında Eğri) başarısı elde ederek etkileyici sonuçlar sunduğunu göstermiştir.



Şekil 2.5 Önerilen MIST ilk önce klip düzeyindeki sözde etiketleri (J.-C. Feng vd. 2021)

H. Yuan ve arkadaşları (2021), “Transanomaly” adında tahmine dayalı VGAT yöntemi önermiştir. Yöntem, önceki karelerden yola çıkarak bir videonun sonraki karesini

tahmin etmek için video işleme için bir dönüştürücü model olan Video Vision Transformer (ViViT)'in modifiye edilmiş bir sürümünü kullanmaktadır. Yöntem ayrıca giriş karelerini yeniden oluşturmak ve tahmin hatasını hesaplamak için bir U-Net mimarisi kullanmıştır. Tahmin hatası, her karenin ve kare içindeki her parçanın düzenlilik skorunu ölçmek için kullanılmıştır. Düzenlilik skoru ne kadar düşükse, karenin veya parçanın anormal olma olasılığı o kadar yüksektir. Yöntem, düzenlilik skorlarını bir eşikle karşılaştırarak hem anomali tespiti hem de anomali yerelleştirme yapabilmektedir. Şekil 2.6'da girdi olarak verilen bir video karesinin önce U-Net ile ön işlemden geçirildiği, sonra ViViT ile gelecek kareyi tahmin etmeye çalıştığı ve son olarak tahmin edilen kare ile gerçek kare arasındaki farkın anomalileri belirlemek için kullanıldığı görülmektedir.



Şekil 2.6 Önerilen anomali tespit modeli (H. Yuan vd. 2021)

Tian ve arkadaşları (2021), zayıf denetimli VGAT için sağlam zamansal özellik büyüklüğü öğrenme yöntemi sunmuşlardır. Bu yöntem, problemi birçok örnekleme öğrenme (MIL) problemi olarak formüle eder ve sınıf etiketlerinden bağımsız olarak pozitif örneklerin büyüklüklerini negatif örneklerden daha yüksek atayan bir özellik-büyüklük öğrenme fonksiyonu kullanır. Ayrıca, parçacıklar arasındaki uzun ve kısa menzilli zamansal bağımlılıkları yakalamak amacıyla genişletilmiş evrişimler ve kendinden dikkat mekanizmaları kullanılmıştır. Her bir parçacığın ve parçacık içindeki her özelliğin düzenliliği, özellik büyüklüklerini bir eşik değeri ile karşılaştırarak değerlendirilir. Bu şekilde, anormal parçacıklar ve özellikler vurgulanarak hem anomali tespiti hem de anomali yerelleştirmesi gerçekleştirilebilir.

Hirschorn ve Avidan (2022), insan duruşlarında anormallikleri tespit etmek için yeni bir yöntem geliştirmiştir. Bu yöntemin, iskelet verilerine dayanarak görünüm ve arka plan gibi etkenlerden etkilenmediği belirtilmiştir. Yöntem, duruş grafikleri arasındaki zamansal ve uzamsal ilişkileri yakalamak için zamansal-uzamsal grafik evrişim blokları kullanır. Normal veya anormal örnekler içeren veri kümelerinde denetimsiz ve denetimli durumlarda çalışabilmektedir.

Zaheer ve arkadaşları (2022), eğitim verilerinde herhangi bir etiketleme gerektirmeyen yeni bir denetimsiz Üretici İşbirlikçi Öğrenme (GCL) yaklaşımı önermiştir. Yaklaşım, bir üretici ve bir ayırt edici arasında çapraz denetim oluşturmak için anomalilerin düşük sıklığını kullanır. Temelde, her iki ağ da işbirlikçi bir şekilde eğitilir. Üretici, bozulmuş girdilerden normal kareler yeniden oluşturmaya çalışırken, ayırt edici normal ve anormal kareler arasında ayırım yapmaya çalışır. Yaklaşımın mevcut en iyi denetimsiz ve tek sınıflı sınıflandırma yöntemlerine göre tutarlı bir iyileşme gösterdiği belirtilmiştir.

Cao ve arkadaşları (2022), görüntülerdeki farklı bölgeler arasındaki uzamsal ve zamansal ilişkileri yakalamak için bir grafik temsilini kullanmıştır. Bu grafiksel temsil, videonun normal davranışını öğrenmek için kullanılan grafiksel evrişimli ağlarıyla birleştirilmiştir. Önerilen yaklaşım, videodaki içeriğe dayalı olarak grafikleri dinamik bir şekilde güncelleyerek ağın sahnelerdeki değişikliklere uyum sağlama yeteneğini artırmaktadır. Önerilen zayıf denetimli öğrenme ağı, büyük sayıda normal videonun yanı sıra etiketlenmemiş anormallikler üzerinde eğitilmiştir. Bu şekilde ağ, videonun normal davranışını öğrenir ve anormallikleri tespit etmek için sapmaları kullanabilir. Bu yöntemin anormallik tespiti performansını artırdığı belirtilmiştir.

Saleem ve arkadaşları (2022) tarafından sunulan çalışma, VGAT için maliyet-etkin ve kaynak-verimli bir çerçeve sağlamaktadır. Bu çerçeve, kısmi kaydırma stratejisi ile 2D evrişimli mimarisi tabanlı MobileNetV2 modelini birleştiren ve Temporal based Anomaly Recognizer (TAR) adını taşıyan bir yapıyı içermektedir. Gerçek zamanlı VGAT uygulamaları için tasarlanan TAR çerçevesi, altı kamera akışını aynı anda ele alma yeteneğiyle birlikte düşük gecikme oranı sunarak etkin bir performans

sağlamaktadır. Bu çalışma, gözetim sistemlerinin güvenliğini artırmayı ve daha etkin bir şekilde işlem yapmayı hedefleyen önemli bir gelişme olarak değerlendirilebilir.

VGAT için öz denetimli öğrenme ve seyrek temsil teknikleri kullanan yeni bir yöntem Wu ve arkadaşları (2022) tarafından önerilmiştir. Bu yöntem, ilk olarak öz denetimli öğrenme yaklaşımını kullanarak normal video özelliklerini öğrenmektedir. Daha sonra, seyrek temsil tekniklerini kullanarak yeni videodaki anormal kareleri tespit etmektedir. Bu yaklaşım, etiketli verilere ihtiyaç duymadan anormallikleri tespit etme konusunda etkili bir strateji olduğu vurgulanmaktadır.

Denetimli VGAT algoritmalarının değerlendirilmesi için UBnormal adında yeni bir ölçüt sunan çalışma (Acsintoae vd. 2022) bu alandaki araştırmacılar için faydalı bir araç sağlamıştır. Gerçek dünya gözetim görüntülerine dayanan UBnormal, çeşitli anormallik sınıflarını içermesi sayesinde daha etkili ve güçlü anormallik algılama algoritmalarının geliştirilmesini desteklemeyi amaçlamıştır. Bununla birlikte, yazarlar UBnormal veri kümesinde açık set tanıma için Ortalama Hassasiyet (APOSR) değerlendirme metriği önermişlerdir. Bu önerilen ölçüt ve değerlendirme metrikleri, anormallik algılama algoritmalarının performansını ölçmek ve geliştirmek amacıyla kullanılabileceği belirtilmiştir.

Madan ve arkadaşları (2022), VGAT için öz-denetimli maskeli evrişimli transformer blok yöntemi sunmuşlardır. Bu yenilikçi yaklaşım, yalnızca girdi verilerini kullanarak ve anormal etiketli örnekler olmadan model eğitimini gerçekleştirmektedir. Transformer bloğu, girdi verilerinin maskelemeli kısımlarını tahmin etmek üzere eğitilir ve bu tahmin yeteneği, verilerdeki anormallikleri tespit etmek için kullanılır. Yapılan deneyler, önerilen yaklaşımın birkaç veri kümesinde anormallikleri yüksek doğrulukla tespit etme yeteneğine sahip olduğunu göstermektedir.

Sapkota ve Yu (2022) tarafından yapılan çalışmada, VGAT için Bayesian parametrik olmayan bir yaklaşım önerilmiştir. Bu çalışma, video verilerinin karmaşık dağılımlarını esnek bir şekilde modelleyen bir yöntem sunmaktadır. Yöntem, videoları bölümlere ayırarak alt modüler optimizasyon kullanmakta ve potansiyel anormallik içeren çeşitli

ve bilgilendirici bölüm alt kümelerini seçmektedir. Geleneksel parametrik yöntemlerin karmaşık dağılımları doğru bir şekilde modelleyememe sorununa karşılık, Bayesian parametrik olmayan yaklaşım daha geniş bir dağılım yelpazesine uyum sağlayabilmektedir. Bu sayede, nadir ve zor bulunan anormallikleri tespit etme olasılığı en üst düzeye çıkarılmaktadır. Bu çalışma, VGAT alanına yeni bir bakış açısı sunmuş ve esnek bir modelleme imkânı sağlamıştır.

VGAT için öznitelik tabanlı temsiller kullanan bir yöntem Reiss ve Hoshen (2022) tarafından önerilmiştir. Yöntem, videolardan nesneler, eylemler ve sahneler gibi özniteliklerin çıkarılmasını içermekte ve bu öznitelikler kullanılarak video için bir temsil oluşturulmaktadır. Ardından, sınıflandırma ve kümeleme tekniklerinin bir kombinasyonu kullanılarak bu temsilde anormallik tespiti gerçekleştirilmektedir. Söz konusu çalışma ile öznitelik tabanlı temsillerin VGAT alanında kullanımını ve anormallik tespiti için etkili bir yöntem sunmak amaçlanmıştır.

Y. Chen ve arkadaşları (2022), zayıf denetimli öğrenme kullanarak VGAT için Büyüklük-Karşılaştırmalı Bakış ve Odaklanma Ağı (Magnitude Contrastive Gance and Focus Network-MGFN) önermişlerdir. MGFN, insan görsel sisteminin küresel-yerel bilgi bütünleştirme mekanizmasını taklit ederek önce videonun genel yapısına bakar, ardından anormallik tespiti için yerel bölgelere odaklanır. Yazarlar, mevcut yöntemlerin özellik büyüklüklerini anormallik derecesi olarak kullanırken sahne değişikliklerinin etkilerini göz ardı ettiklerini ve bu nedenle özellik büyüklüklerinin sahneler arasında tutarsızlık göstererek alt-optimal performans elde edebileceğini deneysel olarak belirlemişlerdir. Bu sorunu çözmek amacıyla, Özellik Yükseltme Mekanizması ve Büyüklük Karşıtı Kaybı önerilerek anormallikleri tespit etmek için özellik büyüklüklerinin ayırt ediciliğini artırmak hedeflemiştir. Bu çalışma, VGAT alanında yeni bir perspektif sunarak anormallik tespiti performansını artırmayı amaçlamaktadır.

Thakare ve arkadaşları (2022), VGAT için kendi kendini eğiten bir ağ olan DyAnNet (A Scene Dynamicity Guided Self-Trained Video Anomaly Detection Network)'i önermişlerdir. DyAnNet, anomali olaylarını etkili bir şekilde tespit etmek için zamansal ve uzamsal bilginin kombinasyonunu kullanır. Araştırma, önceden eğitilmiş bir model

aracılığıyla sahne dinamikliğini ölçer ve bu bilgiyi anomali tespiti sürecinde kullanarak anomali olaylarını etkili bir şekilde tespit etmeyi amaçlamaktadır. Çizelge 2.1’de kaynak özetlerinin ve literatür değerlendirmesinin genel bir özeti sunulmuştur.

Çizelge 2.1 Literatür özeti ve değerlendirmesi

Makale	Önerilen Yöntem / Çalışma	Açıklama	Yıl
Abdelghafour vd.	Hareketli nesnelerin sayısı ve videodaki anormallikler arasındaki ilişki analizi yapılarak farklı kalabalık düzeylerinin tanımlaması yapılmıştır.	Çalışma, hafif, orta ve yoğun kalabalık ortamlarını tanımlamış ve kalabalık seviyelerine göre anormallikleri tespit etmiştir. Ayrıca, güvenlik ve trafik izleme gibi alanlarda uygulanabilir hale getirilmiştir.	2021
Liu vd.	Uzamsal ve zamansal verilerin analizi için üç farklı yöntem önerilmiştir.	1) Olayların özellikleri ortak alana yansıtılmış ve anormalliklerin uzamsal dağılım marjlarıyla belirlenmiş, 2) Olay kalıplarının anlamsal özellikleriyle sözlük eğitimi gerçekleştirilmiş, 3) Özellik çıkarıcılarla tahmin yapılmış ve yeniden yapılandırma hataları üzerinden anormallik tespiti yapılmıştır.	2022
Mo vd.	Sparsity modeli, eğitim sözlüğü, lineer kombinasyonlar kullanılan yöntem önerilmiştir.	Model, çoklu nesne olaylarını içeren bir eğitim sözlüğünü kullanmış ve test olaylarını bu sözlüğün seyrek lineer kombinasyonlarıyla yeniden oluşturmuştur.	2013
Y. Yuan vd.	SCD, 3D discrete cosine transforms ile çoklu nesne izleyicisi önerilmiştir.	Çalışma, insanların toplu davranışlarını belirlemek için SCD'yi kullanmıştır. Anormallikler, topluluğun SCD değişiminin çevrimiçi analizi ile belirlenmiştir.	2014
Wang ve Snoussi	HOFO, Non-linear SVM içeren yöntem önerilmiştir.	Yöntem, arka plan çıkarımı sonrasında elde edilen ön plan görüntüsü üzerinden HOFO hesaplamış ve sınıflandırma için non-linear SVM kullanmıştır.	2014
Kaltsa vd.	HOS ve HOG içeren yöntem önerilmiştir.	HOS, kalabalık dinamiklerini yakalayan özgün bir özellik çıkarma metodu sağlamış ve HOG ile birleştirilerek güçlü tanımlayıcılar oluşturmuştur.	2015
Isupova vd.	Parametrik olmayan dinamik Bayesian konu modeli, batch ve online Gibbs örnekleme yöntemleri önerilmiştir.	Önerilen model, batch ve online Gibbs örnekleme yöntemleriyle öğrenme aşamasını gerçekleştirmiş ve yeni bir anormallik ölçüsü sunmuştur.	2016
Colque vd.	Çoklu öznitelikli, çoklu öğrenme tabanlı ve çoklu görünüm tabanlı yöntem önerilmiştir.	Üç yöntem sunulmuştur: 1) Çoklu Öznitelikli Anomali Tespiti, farklı özellikleri birleştirip ağırlıklandırmış, 2) Çoklu Öğrenme Tabanlı Anomali Tespiti, çeşitli öğrenme algoritmalarını birleştirip oylama yapmış, 3) Çoklu Görünüm Tabanlı Anomali Tespiti, farklı açılardan elde edilen verileri optimizasyonla birleştirmiştir.	2018

Çizelge 2.1 Literatür özeti ve değerlendirmesi (devam)

Makale	Önerilen Yöntem / Çalışma	Açıklama	Yıl
Sabokrou vd.	VGAT ve yerelleştirme için FCN ile zaman verilerinin birleştirildiği yeni bir yöntem	Araştırma, VGAT için önceden eğitilmiş ve denetimsiz bir FCN tabanlı modeli kullanmıştır. Bu model, sahnelerdeki anormallikleri belirlemek ve aykırı değerleri tespit etmek için etkin bir özellik temsili sağlamıştır.	2018
Parate vd.	GPHybrid isimli yeni bir nesne takip yöntemi önerilmiştir.	GPHybrid, Küresel Şablon Takibi ve Yama-Tabanlı Şablon Takibi yöntemlerini birleştirerek zorlu senaryolarda güçlü bir takip kapasitesi sunmuştur.	2018
Baskurt ve Samet	Video verilerindeki aktif etkinlikleri özetleyen yöntem önerilmiştir.	Elde edilen videolar özetlenmiş ve sadece aktif etkinlikleri içeren sıkıştırılmış bir görünüm sunulmuştur.	2019
Rodrigues vd.	Farklı zaman ölçeklerinde gelecek ve geçmiş tahminleri üreten model önerilmiştir.	Model, verilen pozisyon takibi girdilerine dayanarak gelecek ve geçmiş tahminler üretmiş ve bu tahminler, anormal aktivitelerin tespiti için kullanılmıştır.	2020
Ramachandra vd.	Siamese ağı isimli yeni bir sınıflandırma yöntemi önerilmiştir.	Önerilen Siamese ağı, girdiler arasındaki benzerlik veya farklılıkları ölçen bir mesafe fonksiyonu kullanmış ve videolarda anormallikleri belirlemede kullanılmıştır.	2020
Bansod ve Nandedkar	Kalabalıkta anomali tespiti ve yer belirlemeyi sağlayan üç adımlı yöntem önerilmiştir.	1) Arka planın kaldırılması, 2) özellik çıkarmı ve davranış tanıma, 3) anomali tespiti ve yer belirleme işlemleri gerçekleştirilmiştir.	2020
Z. Li vd.	Anormal insan davranışlarının tespiti maksatlı çok ölçekli 3D CNN.	3D evrişimli sinir ağı kullanan model, yüksek boyutlu karışık Gaussian modeliyle uyumlu uzamsal-zamansal özellikler çıkarmış ve özel eğitim gerektirmeden yüksek performans sağlamıştır.	2020
Xue vd.	Zaman serisi analizi yöntemi önerilmiştir.	Yöntem, video verilerinden nesnelerin hareketleri, hızı ve yönü gibi önemli uzamsal ve zamansal özelliklerin çıkarılmasını sağlamıştır.	2020
Park vd.	VGAT için bellek rehberli normalite öğrenme yöntemi önerilmiştir.	Normal verilerin prototipik kalıplarını kaydeden bir bellek modülü kullanılmış ve yeni özellik sıklığı ve ayrılığı kayıpları sayesinde, hem bellek öğelerinin hem de normal verilerden öğrenilen özelliklerin ayırt ediciliği artırılmıştır.	2020
Georgescu vd.	İkili sınıflandırıcı ve AE kullanan bir yöntem önerilmiştir.	Sahte anormal örnekler, sınıflandırıcıların normal ve anormal özellikler arasındaki farkı daha iyi ayırt etme yeteneklerini artırmak için kullanılmıştır.	2021
Majhi vd.	Zayıf denetimli öğrenme yöntemi önerilmiştir.	Önerilen yöntem, dikkat ağırlıklarını kullanarak anormallik tespiti için girdi özelliklerini düzenlemiş ve aynı zamanda anormallik sınıflandırmasına da yardımcı olmuştur.	2021
Wan vd.	Gerçek dünya senaryolarında anormallik tespiti için yeni ölçüt ve model önerilmiştir.	Anormallikleri tespit etme kısıtlamalarını ele alarak yeni bir ölçüt sunan yöntem, hareket vektörleri, yoğunluk ve arka plan çıkarma gibi uzamsal-zamansal özelliklerin birleşimini içeren bir model kullanmıştır.	2021
Lv vd.	Meta-prototip ağı önerilmiştir.	DPU, prototipleri dinamik olarak güncellemekte, MPU ise farklı sahneler için hızlı bir şekilde uyum sağlamaktadır.	2021

Çizelge 2.1 Literatür özeti ve değerlendirmesi (devam)

Makale	Önerilen Yöntem / Çalışma	Açıklama	Yıl
Feng vd.	Multiple Instance Self-Training (MIST) Framework isimli yöntem önerilmiştir.	Önerilen MIST yöntemi, sahte etiket üretici ve dikkat artırılmış özellik kodlayıcı bileşenlerinden oluşmuştur.	2021
H. Yuan vd.	Transanomaly yöntemi önerilmiştir.	ViViT'in modifiye edilmiş bir sürümünü kullanan yöntem, önceki karelerden yola çıkarak videonun sonraki karesini tahmin etmeye ve anormalliği yakalamaya çalışmıştır.	2021
Wang ve Snoussi	HOFO ve SVM içeren sınıflandırma yöntemi önerilmiştir.	HOFO, arka plan çıkarımı sonrası elde edilen ön plan görüntüsü ile orijinal görüntünün hızlı analizini sağlamış ve SVM sınıflandırması yapmıştır.	2014
Acintoae vd.	Ubnormal isimli yeni bir ölçüt ve model önerilmiştir.	Anormallik sınıflarını içeren bir veri kümesi sunulmuş ve yeni bir değerlendirme metriği geliştirilmiştir.	2022
Adam vd.	Çoklu sabit konumlu monitörler kullanılarak robust bir tespit yöntemi sunulmuş.	Gerçek zamanlı uygulamalar için uygun hale getirilmiştir.	2008
Aldayri ve Albattah	Farklı anomali tespit teknikleri üzerinde taksonomi geliştirilmiştir.	Çeşitli anomali tiplerini içeren kapsamlı bir sınıflandırma sunulmuştur.	2022
Bansod ve Nandedkar	Momentum ve büyüklük histogramı kullan	Kalabalıkta anomali tespiti ve yer belirleme amacıyla geliştirilen yöntem, UMN veri kümesinde yüksek başarı elde etmiştir.	2020
Boekhoudt vd.	İnsanla ilgili anormallikleri tespit etmeye yönelik bir yöntem sunulmuş.	Gözetim videolarında insanla ilgili anormalliklerin tespitine odaklanmıştır.	2021
Cao vd.	Yayıf denetimli öğrenme ağı önerilmiştir.	Grafiksel temsil kullanarak dinamik bir şekilde güncellenen ağ yöntemi kullanılmıştır.	2022
Chandola vd.	Anomali tespiti konusunda kapsamlı bir literatür incelemesi sunulmuştur.	Anomali tespiti tekniklerinin geniş bir özeti sunulmuştur.	2009
Y. Chen vd.	Büyüklük-Karşılaştırmalı Bakış ve Odaklanma Ağı (MGFN) yöntemi önerilmiştir.	Özellik Yükseltme Mekanizması ve Büyüklük Karşıtı Kayıp önerilmiş ve bu sayede özellik büyüklüklerinin ayırt ediciliği artırılarak daha etkili anormallik tespiti hedeflenmiştir.	2022

Son yıllarda VGAT alanında çeşitli yöntemler ve metodolojiler geliştirilmiştir. Bu yenilikçi yaklaşımlar arasında derin öğrenme ve uzamsal-zamansal modelleme, LSTM ağları, otomatik kodlayıcılar, kendinden denetimli çoklu görev öğrenimi, evrişimli dönüştürücü bloklar, üretici ağlar ve grafiksel sinir ağları gibi çeşitli stratejiler yer almaktadır. Derin öğrenme, videoların içerdiği karmaşık ve yüksek boyutlu verileri işlemek için güçlü bir araçtır. Uzamsal-zamansal modelleme, videolardaki nesnelerin hem 2 boyutlu uzayda hem de zaman içindeki hareketlerini ve değişimlerini daha ayrıntılı bir şekilde anlamak için kullanılır. Bu metodolojiler farklı veri kümelerinde

geniř kapsamlı deneylerle test edilmiř ve doęruluk ile hız gibi önemli kriterlerde başarılı olduklarını göstermiştir. Ayrıca, bazı yöntemler, anormallik tespiti algoritmalarının performansını değerlendirmek için yeni kıyaslama ölçütleri geliştirilmesinde de kullanılmıştır. Örneęin, grafiksel sinir aęları anormallik skorunu hesaplamak için graf teorisi temelli bir yaklaşım sunmuştur. Bu alandaki hızlı ilerlemeler, gözetim ve güvenlik gibi çok çeřitli alanlarda kullanılma potansiyeli taşımaktadır. Bu yöntemlerin gerçek dünya senaryolarında anormallikleri tespit etmede daha etkili ve güvenilir sonuçlar sağlayabilme yetenekleri, gelecekteki uygulamalar için büyük bir potansiyel olarak görölmektedir.



3. VGAT ALANINDA YAYGIN KULLANILAN YAKLAŞIMLAR VE YÖNTEMLER

VGAT alanında kullanılan yöntemlerin çeşitli gruplarda sınıflandırılması, araştırmacıların ve literatürün farklı perspektiflerine bağlı olarak değişmektedir. Hangi sınıflandırmanın daha doğru olduğu, bu perspektiflere göre değişiklik göstermektedir. Her sınıflandırma yaklaşımı, farklı açılardan anomali tespiti problemini ele alır ve çeşitli avantajlar sunar. Bu nedenle, kullanılacak sınıflandırma yaklaşımı, çalışmanın amacına, kullanılan veriye ve araştırmacının tercihinine bağlı olarak değişebilir. Bu çeşitlilik, VGAT alanındaki yöntemlerin esnekliğini ve uygulanabilirliğini vurgular, bu da her bir araştırmanın benzersiz özelliklere odaklanabilmesine olanak tanır.

Bu çalışmada VGAT için kullanılan yöntemler genel olarak öğrenme türlerine dayalı ve algoritma türlerine dayalı yaklaşımlar olmak üzere iki ana grupta sınıflandırılmıştır (Şengönül vd. 2023). Öğrenme türlerine dayalı yaklaşımlar denetimli, denetimsiz ve yarı-denetimli alt gruplara ayrılır. Algoritma türlerine dayalı yaklaşımlar ise istatistik tabanlı, sınıflandırma tabanlı, yeniden oluşturma tabanlı, tahmin tabanlı ve diğer alt gruplara ayrılır. Söz konusu taksonomi Çizelge 3.1'de gösterilmiştir.

Bazı yaklaşımlar, aynı anda farklı sınıflandırma grupları içinde yer alabilir. Örneğin, LSTM hem denetimsiz öğrenme yaklaşımları hem de tahmin tabanlı yaklaşımlar kapsamında değerlendirilebilir. Bu durum, LSTM'nin çok yönlü bir model olduğunu ve farklı amaçlar doğrultusunda kullanılabildiğini gösterir. Bu nedenle, tez kapsamında yer alan yöntemler ağırlıklı olarak tek bir grup altında detaylandırılmıştır. Ancak, bu sınıflandırma, her bir yöntemin çok yönlü özelliklerini ve çeşitli kullanım alanlarını vurgulamak adına esnek bir bakış açısı sunmaktadır.

Çizelge 3.1 VGAT alanındaki yaklaşımların sınıflandırılması

Öğrenme Türlerine Dayalı	Algoritma Türlerine Dayalı
Denetimli öğrenme	İstatistik tabanlı
Denetimsiz öğrenme	Sınıflandırma tabanlı
Yarı-Denetimli öğrenme	Yeniden yapılandırma tabanlı
	Tahmin tabanlı
	Diğer

3.1 Öğrenme Türlerine Dayalı Yaklaşımlar

Makine öğrenmesi, günümüzde birçok alanda karşılaşılan problemlere çözümler sunan yapay zekâ teknolojilerinin temelini oluşturmaktadır. Bu algoritmalar, veriler üzerinden desenleri tanımlayarak ve tahminlerde bulunarak karmaşık sorunlara çözüm aramakta kullanılan kod parçalarıdır. Bu algoritmaların işlevselliği, kullanılan veri kümesine ve seçilen öğrenme yöntemine bağlı olarak farklılık gösterir. Özellikle, derin öğrenme, büyük ve karmaşık veri kümeleri üzerinde çalışan makine öğrenmesi algoritmalarının özelleşmiş bir alt dalıdır.

Makine öğrenmesi algoritmalarının etkinliği, seçilen veri özellikleri ve bu özelliklere atanan ağırlıklarla doğrudan ilişkilidir. Özellikler, veri noktalarını tanımlayan değişkenlerdir, ağırlıklar ise model içerisinde bu özelliklerin önemini belirleyen katsayılardır. Model eğitim sürecinde, bu ağırlıklar veriye göre otomatik olarak ayarlanır. Fakat, farklı özellik ve ağırlık seçimleri modelin sonuçlarını değiştirebilir. Modelin sınırlı sayıda yüksek ağırlıklı özellikle eğitilmesi durumunda, aşırı uyum (overfitting) sorunu ortaya çıkabilir. Aşırı uyum, modelin eğitim verilerine aşırı derecede uyum sağlamasına ancak yeni verilere uyum sağlayamamasına yol açan bir durumdur (Medel 2016). Bu nedenle, makine öğrenmesi tekniklerinin ve verilerinin derinlemesine analizi ve mevcut çözümlerin kıyaslanması, etkili bir model geliştirilmesi için kritik önem taşımaktadır.

Makine öğrenmesi, genellikle denetimli, denetimsiz ve yarı denetimli olmak üzere üç ana kategoride incelenir. Denetimli öğrenme, etiketlenmiş verilerle model eğitimi yapılmasını içerirken, denetimsiz öğrenme etiketlenmemiş veriler üzerinden desen keşfini hedefler. Yarı denetimli öğrenme ise hem etiketli hem de etiketlenmemiş verilerin birlikte kullanıldığı bir yaklaşımdır.

Bu bağlamda, etkili bir makine öğrenimi modeli geliştirmek için veri kümesinin özelliklerinin dikkatlice seçilmesi, ağırlıkların doğru bir şekilde ayarlanması ve uygun bir öğrenme yönteminin belirlenmesi gereklidir. Ayrıca, benzer problemlere yönelik mevcut çözümlerin değerlendirilmesi, modelin başarısını artırmada önemli bir rol oynayabilir.

3.1.1 Denetimli öğrenme yaklaşımları

Denetimli öğrenme, önceden etiketlenmiş veriler kullanarak öğrenme modellerinin geliştirilmesine dayanır. Temelde model, belirli hedefleri gerçekleştirmek veya çıktıları tahmin etmek için eğitim verilerini analiz eder. Eğitim verileri hem girdi özelliklerini hem de bu özelliklerle ilişkili hedef çıktıları içerir. Denetimli öğrenme süreci boyunca, modelin tahminleri gerçek çıktılarla karşılaştırılarak hata oranının minimize edilmesi hedeflenir. Bu süreç, modelin girdi ve çıktı arasındaki ilişkileri anlamasını ve yeni veriler için tahminler yapmasını sağlar (Cunningham vd. 2008).

VGAT gibi alanlarda, bu yaklaşım temelli kullanılan yöntemler, önceden etiketlenmiş verilerden yararlanarak anormal olayları tespit etme ve normal davranışları öngörme kapasitesine sahiptir. Model, bu sayede anormal ve normal davranışlar arasındaki farkları öğrenir ve yeni verilerde olası anormallikleri tespit edebilir. Bu yöntemler, farklı video gözetim senaryolarına ve ortamlarına uygun, özelleştirilmiş çözümler sunabilir. Örneğin, piksel düzeyinden klip düzeyine kadar farklı seviyelerde anormalliklerin tespiti için kullanılabilirler.

Ancak, VGAT alanında denetimli öğrenme yöntemlerinin kullanımı bazı zorluklar ve sınırlamalarla karşılaşabilir. En büyük zorluklardan biri, yüksek kaliteli ve yeterli miktarda etiketlenmiş veri gereksinimidir. Etiketlenmiş veri toplama süreci zaman alıcı, maliyetli ve uzmanlık gerektiren bir işlem olabilir. Ayrıca, VGAT alanındaki anormal davranışların çeşitliliği ve karmaşıklığı, modelin genelleme yeteneklerini sınırlayabilir. Model, eğitim sürecinde karşılaşmadığı anormallikleri tanımakta zorlanabilir. Bu durum, sistemin gerçek dünya uygulamalarındaki etkinliğini ve güvenilirliğini etkileyebilir. Bu nedenle, denetimli öğrenme yöntemleri kullanılırken, etiketlenmiş veriye erişim ve anormal davranışların çeşitliliği gibi faktörler dikkate alınmalıdır.

Etkili bir denetimli öğrenme modeli geliştirmek için yüksek kaliteli eğitim verisinin seçilmesi ve hazırlanması önemlidir. Eğitim veri kümesinin boyutu ve kalitesi, modelin genel performansını belirleyici etmenlerdendir. Ayrıca, eğitim prosedürü, hiperparametre ayarları ve modelin düzenlenmesi gibi faktörler de modelin başarısında büyük rol oynar. Bu nedenle, VGAT alanında denetimli öğrenme tekniklerinin etkili bir şekilde uygulanması için bu faktörlerin dikkatle ele alınması gereklidir.

Sonuç olarak, denetimli öğrenme teknikleri VGAT alanında anormal davranış tespitinde güçlü bir araç olabilir. Ancak, doğru veri seçimi, eğitim prosedürü ve model parametrelerinin titizlikle ele alınması gerekmektedir. Bu şekilde, VGAT alanında etkili ve hassas anormallik tespiti sağlanabilir.

3.1.2 Denetimsiz öğrenme yaklaşımları

Denetimsiz öğrenme, etiketlenmemiş veriler üzerinde kalıpları, ilişkileri ve yapıları keşfetmek için kullanılan bir makine öğrenmesi yaklaşımıdır. Bu yaklaşım, çıktı değişkenlerini önceden belirlemeden, yalnızca girdi değişkenlerini kullanarak modelin kendi içindeki öğrenme mekanizmaları ile veriyi anlamlandırmasını ve sonuçlar üretmesini sağlar. Denetimsiz öğrenme yöntemlerinin ürettiği sonuçlar genellikle kesin olmayıp, farklı yöntemler farklı sonuçlar verebilir (Barlow 1989).

Anormal durumların belirtileri veya yapıları önceden bilinmediğinde bu yaklaşım temelli yöntemler, verilerdeki benzerlik ve farklılıkları keşfetmek için kullanılabilir. Ayrıca, veri kümelerinde normal grupları veya desenleri tanımlayabilir ve bu gruplara dayanarak anormallikleri tespit edebilir. Örneğin, kümeleme algoritmaları, benzer özelliklere sahip verileri gruplayarak potansiyel anormal grupları belirleyebilir. Bu sayede, anormal durumların belirgin özelliklerini belirleme ve genel veri yapısını daha iyi anlama konusunda önemli bir rol oynayabilirler.

Ancak, denetimsiz öğrenme yöntemlerinin bazı sınırlamaları ve dezavantajları da bulunmaktadır. Özellikle, öğrenme sonrası elde edilen sonuçların yorumlanması ve değerlendirilmesi daha fazla uzmanlık ve insan müdahalesi gerektirebilir. Ayrıca, bu yöntemler, verilerin yapısal özelliklerini veya anormallikleri, denetimli öğrenme kadar kesin bir şekilde tanımlamayabilir. Bu nedenle, denetimsiz öğrenme sonuçlarının doğruluğunu ve güvenilirliğini değerlendirmek genellikle daha karmaşıktır. Bu sınırlamalar, özellikle spesifik bir görev veya uygulama alanında kullanılmadan önce dikkate alınmalıdır.

Sonuç olarak, denetimsiz öğrenme yaklaşımları, önceden bilinmeyen anormalliklerin tespit edilmesi gereken senaryolarda veya geleneksel denetimli öğrenme yaklaşımlarının yetersiz kaldığı durumlarda özellikle değerlidir. Bu nedenle, VGAT alanında denetimsiz öğrenme yaklaşımlarının potansiyelini değerlendirmek ve en iyi sonuçları elde etmek için çeşitli algoritmaların ve yöntemlerin incelenmesi büyük önem taşımaktadır. Ayrıca bu yöntemlerin etkili kullanımı, verinin doğru ölçeklendirilmesi, uygun algoritmaların seçilmesi ve sonuçların doğru yorumlanması gibi koşullara bağlıdır. VGAT alanında denetimsiz öğrenme yöntemlerinin kullanımı, modelin daha geniş bir yelpazede anormallikleri tanıma yeteneğini artırabilir.

3.1.3 Yarı-denetimli öğrenme yaklaşımları

Yarı-denetimli öğrenme hem etiketli hem de etiketsiz verileri kullanarak sınıflandırıcı modeller oluşturmayı hedefleyen bir makine öğrenme yaklaşımıdır (Zhu 2005). Bu yaklaşım, özellikle etiketli verinin sınırlı olduğu durumlarda oldukça etkilidir. Yarı-

denetimli öğrenme, az miktarda etiketli veri kullanarak etkileşimli bir eğitim prosedürü ile öngörücü modeller oluşturur. Bu modeller, bir video karesinin normal veya anormal olup olmadığını tahmin etmek için eğitilir. Ardından, etiketsiz video karelerini sınıflandırır ve her tahmin için bir güven puanı sağlar. Güvenilir tahminlere sahip kareler, modelin eğitimi için yeni veri kümesine eklenerek modelin güncellenmesini sağlar.

Genel olarak yarı-denetimli öğrenme, sınırlı etiketli verilerle çalışırken etiketsiz verilerdeki bilgiyi kullanarak sınıflandırıcının doğruluğunu artırmayı hedefler. VGAT gibi alanlarda bu yöntem, önemli avantajlar sunar. Örnek olarak, sadece normal videoların etiketlenmesi ile anormal durumların etiketlenmesi gibi zorlu süreçlerden kaçınmayı sağlar (Khair ve Kumar 2022). Ayrıca, bu yaklaşımlar farklı ortamlar veya senaryolar için kolayca uyarlanabilir. Yeni bir izleme ortamına geçildiğinde, sadece o özel ortama ait normal videoların eklenmesi yeterlidir (Sarker vd. 2021).

Yarı-denetimli öğrenme yöntemleri, VGAT alanında etiketsiz ve sınırlı etiketli verilerin potansiyelini en üst düzeye çıkararak anormallik tespiti performansını artırabilir. Modelin güncellenmesi ve sınıflandırıcının doğruluğunun artırılması, etiketli ve etiketsiz verilerin uyumlu kullanımıyla mümkündür. Bu yaklaşımın VGAT alanında uygulanması, daha kesin sonuçlar ve daha güvenilir anormallik tespiti sağlayarak güvenlik, izleme ve diğer uygulama alanlarında daha etkili çözümler sunabilir.

3.1.4 Öğrenme yaklaşımlarının karşılaştırılması

Anomali tespitinde kullanılan öğrenme tekniklerini karşılaştırdığımızda, denetimli, denetimsiz ve yarı-denetimli öğrenme tekniklerinin her birinin belirgin avantajları ve dezavantajları olduğunu görmekteyiz.

Denetimli öğrenme teknikleri, etiketli verileri kullanarak anormallikleri tanımlama, sınıflandırma ve uzamsal/zamansal olarak yerleştirme konusunda yüksek etkinlik gösterir. Bu yöntemler, önceden tanımlanmış anormallik türlerini tespit etmede

başarılıdır. Ancak, geniş çaplı etiketli veri gereksinimi nedeniyle veri toplama ve etiketleme süreci zaman alıcı ve zahmetli olabilir. Ayrıca, çevresel değişikliklere karşı hassasiyetleri doğruluklarını etkileyebilir.

Denetimsiz öğrenme teknikleri, etiketli veri gerektirmeyen ve gerçek zamanlı anomali tespiti yapabilen yöntemlerdir. Bu teknikler, veri kümesindeki normdan sapmaları tanımlayarak anormallikleri belirleyebilir. Ancak, anormallik tanımının belirsizliği nedeniyle kavramı öğrenmede zorlanabilirler ve düşük anormallik oranına sahip veri kümelerinde etkisiz kalabilirler.

Yarı-denetimli öğrenme teknikleri ise hem belirli tipteki anormallikleri tanımlama hem de gerçek zamanlı anomali tespiti yapabilme yeteneğini sunar. Bu teknikler, etiketli veri sınırlılıklarını aşmak ve denetimsiz yöntemlere kıyasla daha yüksek doğruluk elde etmek için etiketli ve etiketsiz verileri birlikte kullanabilirler.

VGAT alanında, denetimli, denetimsiz ve yarı-denetimli öğrenme teknikleri, ihtiyaca ve mevcut veriye göre farklı avantajlar ve dezavantajlar sunar. Hangi yöntemin tercih edileceği, uygulamanın gereksinimlerine ve verinin niteliğine bağlıdır. Her bir yöntemin etkinliği, çevresel değişimler ve veri kalitesi gibi faktörlere bağlı olabilir.

Son yıllarda, VGAT için derin öğrenme tabanlı yöntemler popülerlik kazanmıştır. Bu yöntemler, video verilerinin yüksek boyutluluğu ve karmaşıklığı ile başa çıkmak için güçlü temsil öğrenme kabiliyetine sahiptirler. Özellikle üretici modeller gibi derin öğrenme tabanlı yaklaşımlar, başarılı sonuçlar sunmaktadır. Derin öğrenme tabanlı yöntemlerin avantajları ve dezavantajları Çizelge 3.2'de sunulmaktadır.

Çizelge 3.2 Öğrenme yöntemlerinin avantaj ve dezavantajları

Öğrenme Tekniği	Avantajları	Dezavantajları
Denetimli öğrenme	- Önceden öğretilmiş anormallik türlerini etkili bir şekilde tespit edebilir. - Anormalliklerin uzamsal konumlandırılması konusunda yüksek başarı sağlar.	- Yüksek kalitede ve yeterli miktarda etiketlenmiş veri gereksinimi vardır. - Etiketlenmiş veri toplama süreci zaman alıcı, maliyetli ve uzmanlık gerektiren bir işlemdir. - Anomalilerin çeşitliliği ve karmaşıklığı, modelin genelleme yeteneğini kısıtlayabilir. - Eğitim sürecinde karşılaşılmamış anormallikleri tanımakta zorluk yaşayabilir.
Denetimsiz öğrenme	- Veri etiketleme gerektirmez. - Veri kümesindeki gizli desenleri keşfetme yeteneğiyle bilinmeyen yapıları ortaya çıkarabilir. - Sürekli olarak gelen veri akışına hızlı bir şekilde uyum sağlayabilir. Bu da gerçek zamanlı uygulamalar için kolaylık sağlar.	- Etiketlenmemiş veri üzerinde çalıştığı için modelin anormal durumları belirleme yeteneği sınırlı olabilir. - Etiketli veri eksikliği, doğruluk ölçütlerini belirleme ve sonuçları yorumlama süreçlerini zorlaştırabilir. - Büyük veri setleri üzerinde çalışırken büyük hesaplama gücüne ihtiyaç duyabilir, bu da işlem maliyetini artırabilir.
Yarı-denetimli öğrenme	- Hem belirli anormallik tiplerini tanımlama hem de gerçek zamanlı anomali tespiti yapabilme yeteneğine sahiptir. - Etiketli verilerin azlığından kaynaklanan sorunları azaltabilir.	- Etiketli verilerin eksikliği, modelin etiketlenmiş örneklerle doğru bir şekilde öğrenememesine neden olabilir. - Hangi verilerin etikleneceğini belirlemek zor olabilir, yanlış etiketleme stratejileri modelin performansını olumsuz etkileyebilir. - Model, etiketli verilere aşırı hassas olabilir, bu da genelleme yeteneğini azaltabilir.

3.2 Algoritma Türlerine Dayalı Yaklaşımlar

Bu bölümde, algoritma türüne dayalı VGAT yaklaşımları ve yöntemleri gruplandırılmaktadır. Farklı teknikler, veri işleme, desen tanıma ve sonuç üretme açısından çeşitli yaklaşımlara sahiptir. İstatistik tabanlı, sınıflandırma tabanlı, yeniden yapılandırma tabanlı ve tahmin tabanlı gibi teknikler, farklı ihtiyaçları karşılamak üzere kullanılmaktadır.

3.2.1 İstatistik tabanlı yaklaşımlar

İstatistik tabanlı anormallik tespiti, normal davranışları ve anormallikleri tanımlamak için kullanılan bir yaklaşımdır. Bu yaklaşım, istatistiksel modeller oluşturarak normal davranışları tanımlar ve bu modele uymayan davranışları anormal olarak sınıflandırır.

İstatistik tabanlı anormallik tespiti yöntemleri, olayların sayısı, hızı, yoğunluğu, hareket yönleri gibi istatistiksel özellikleri analiz ederek normal davranış modelleri oluşturur. Oluşturulan model, bir alandaki normal durumları temsil ederken modele uymayan durumlar, potansiyel olarak anormal hareketler veya davranışlar olarak kabul edilir.

İstatistik tabanlı yaklaşımlar, belirli bir alanda normları belirlemek için çeşitli istatistiksel ölçümleri kullanır. Bu teknikler, genel olarak parametrik ve parametrik olmayan yöntemler olmak üzere iki ana kategoriye ayrılır (Santhosh vd. 2020). Parametrik yöntemler, verilerin belirli bir olasılık dağılımını takip ettiği varsayımına dayanır ve bu varsayımına göre veri dağılımının parametrelerini tahmin ederek yeni veri noktalarının olasılığını hesaplar. Parametrik olmayan yöntemler ise, verilerin dağılımı hakkında herhangi bir özel varsayım yapmadan işlem yapar. Parametrik yöntemler, video karelerinin aralık veya oran ölçeğinde olmasını gerektirirken, parametrik olmayan yöntemler, nominal veya sıralı ölçeğinde olmasını yeterli bulur. Bu farklılıklar, her iki teknik türünün uygulama alanları ve veri gereksinimleri açısından farklı yaklaşımlar sunduğunu gösterir.

3.2.1.1 Parametrik yöntemler

Parametrik yöntemler, belirli bir veri kümesinin, varsayılan bir olasılık dağılımına uyduğunu varsayar ve bu temelde veri dağılımının karakteristik parametrelerini tahmin eder. Örneğin, normal dağılım için temel parametreler ortalama ve standart sapma'dır. Bu, veri kümesinin ortalamasının ve dağılımın yayılma derecesinin belirlenmesini içerir.

Bu parametreleri hesaplamak, parametrik yöntemlerin pratik uygulamalarından biridir. Ancak, bu yaklaşımın bir dezavantajı, gerçek veri kümesinin varsayılan dağılımdan sapması durumunda, yani gerçek dünyada dağılımın başka bir formda olduğu durumlarda, sonuçların yanıltıcı olma potansiyelidir. Dolayısıyla, parametrik yöntemlerin uygulanması öncesinde, veri kümesinin özellikleri ve dağılımın uygunluğu dikkatlice değerlendirilmelidir.

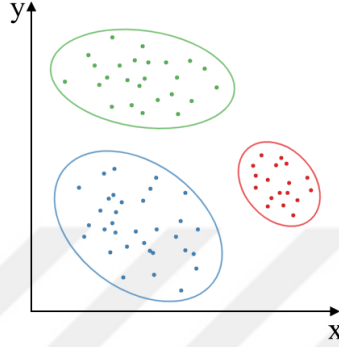
3.2.1.1.1 Gauss karışım modeli

VGAT için en popüler parametrik yöntemlerden biri olan Gauss Karışım Modeli (GMM), veri kümesini birden fazla Gauss dağılımının bir karışımı olarak modelleyerek anormal davranışları tespit etmek için istatistiksel bir yöntem sunar (Bouman vd. 1995).

GMM, veri noktalarını çeşitli Gauss dağılımlarıyla ilişkilendirir ve bu dağılımların ağırlıklarını ve parametrelerini olasılıksal bir yaklaşımla hesaplar. Bu dağılımlar, veri kümesinin farklı yapılarını temsil etmek üzere bir araya getirilir. Model, veri noktalarını kullanarak Gauss dağılımlarının parametrelerini tahmin eder, bu dağılımların ağırlıklarını belirler ve ardından yeni veri noktalarının bu dağılımlar altında olasılığını hesaplamak için bu bilgileri kullanır.

GMM'de her bileşen, veri noktalarının bir kümesini temsil eder ve bir ortalama vektörü (μ) ve bir kovaryans matrisi (Σ) ile tanımlanır. Bu parametreler, genellikle beklenti-maksimizasyon (Expectation-Maximization, EM) algoritması gibi optimizasyon

teknikleri kullanılarak tahmin edilir. Kovaryans matrisi, bileşenin şeklini belirler. Eğer kovaryans matrisi tam sıralı ise, bileşen eliptik bir şekle sahiptir. GMM, normal davranış kalıplarını yakalamak için eğitildiğinde, modelin tanımadığı veya önceden belirlenmiş kalıpların dışındaki veri noktaları anormal olarak değerlendirilir. Şekil 3.1’de GMM modeli uygulaması temsilen gösterilmektedir.



Şekil 3.1 Gauss Karışım Modeli

Bu bağlamda, GMM'nin temel yaklaşımı, veri kümesini birden çok Gauss dağılımının karışımı olarak modellemesi ve bu dağılımların parametreleri kullanılarak anormal davranışların tespiti. GMM, eğitim sürecinde öğrenilen normal davranışlara dayalı olarak, tanıdık olmayan ve önceden belirlenmiş kalıpların dışında kalan veri noktalarını anlamlandırmak için kullanılır.

3.2.1.1.2 Bayesian yaklaşımı

Bayesian yaklaşım, bir hipotezin doğruluğunu hesaplamak için bir olayın gerçekleşme olasılığına dayanır. VGAT bağlamında kullanıldığında, bu yöntem video akışındaki normal aktivitelerin olasılıklarını hesaplar ve ardından akıştaki her yeni olayın normal olup olmadığını belirlemek için bu olasılıkları kullanır (Nguyen vd. 2015).

Başlangıçta, normal durumun olasılığını belirleyen bir öncelik (prior) oluşturulur, ardından sistemin normal davranışını yansıtan bir model belirlenir. Bu model, normal durumu tanımlayan özellikleri içerir ve veri kümesi üzerinde eğitilir. Model, normal durumun belirli özelliklerini anlamak ve tanımlamak üzere tasarlanmıştır.

Modelin oluşturulmasının ardından anomali oluşturan faktörlerin olasılıkları hesaplanır. Anomali oluşturan faktörler, normal davranıştan sapma eğiliminde olan özellikler veya olaylar olabilir. Tüm durumların olasılıkları normalize edilerek genel bir olasılık dağılımı elde edilir. Bayesian teoreminden yararlanılarak olasılık dağılımı ve öncelik kullanılarak “posterior” hesaplanır. Posterior olasılığının belirli bir eşik değerini aşması durumunda, bir durum anomali olarak kabul edilir.

Bu yaklaşımda, özellikleri belirgin bir şekilde farklı olan hareketlerin modellenmesi zor olduğu için anormal hareketlerin tespiti de zor olabilir. Ayrıca, modelin doğru çalışabilmesi için yeterli sayıda normal aktivitenin eğitim veri kümesinde bulunması önemlidir. Eğitim veri kümesinde yeterli normal örneğin bulunmaması, modelin anormallikleri doğru bir şekilde tespit etme yeteneğini olumsuz etkileyebilir.

3.2.1.1.3 Frekans ve spektral analiz

Frekans ve Spektral Analiz, bir sinyalin farklı frekans bileşenlerini tanımlamak ve incelemek için kullanılan bir yöntemdir (Yokkampon vd. 2020). Bu analiz, görüntü veya video verilerini bir sinyal olarak ele alarak, bu sinyallerin frekans bileşenlerini çıkarır ve analiz eder. Bu süreç, veriler içindeki frekans bileşenlerinin incelenmesini ve anormal durumların tespit edilmesini sağlar.

Spektral analiz, bir sinyalin frekans spektrumunu detaylı bir şekilde inceleyerek sinyalin frekans bileşenlerinin özelliklerini ve karakteristiklerini belirlemeyi amaçlar. Genellikle sinyalin frekans spektrumunu grafiksel veya sayısal olarak temsil eden bu yöntem, belirli frekans aralıklarındaki değişiklikleri veya anormallikleri tespit etmede oldukça etkilidir. VGAT alanında, spektral analiz görüntü veya video verilerindeki frekans bileşenlerinin özelliklerine odaklanır ve anormal durumlar, frekans spektrumunda beklenmedik desenler veya yapısal değişiklikler olarak görülebilir.

Örneğin, bir trafik kamerası tarafından kaydedilen normal trafik akışının frekans spektrumu, genellikle düzenli ve belirli bir desene sahip olur. Ancak, bir kaza veya

yoğun trafik durumunda, bu desende belirgin farklılıklar veya anormallikler ortaya çıkabilir. Frekans ve spektral analiz yöntemleri, bu tür değişiklikleri tespit edebilir ve potansiyel anomalileri işaretleyebilir.

Frekans ve spektral analiz, VGAT alanında anormallik tespitinde güçlü bir araç olarak kabul edilir. Ancak, başarılı bir analiz için doğru analiz teknikleri ve parametrelerinin seçilmesi kritik önem taşır. Farklı veri kümeleri ve senaryolar, çeşitli frekans özellikleri gösterebilir. Bu nedenle, en uygun yöntemlerin seçimi büyük bir özen gerektirir. Diğer önemli bir nokta ise, frekans ve spektral analizin tek başına her zaman yeterli olmayabileceğidir. Karmaşık sahneler veya değişken çevresel koşullarda, hatalı pozitif veya negatif sonuçlar elde edilebilir. Bu nedenle, genellikle diğer tekniklerle birleştirilerek daha güvenilir ve kapsamlı sonuçlar elde etmek amaçlanır.

3.2.1.2 Parametrik olmayan yöntemler

Parametrik olmayan yöntemler, veri noktalarının belirli bir olasılık dağılımına uymadığı varsayımı üzerine çalışır ve veri kümesinin olasılık dağılımının yapısını kestirmeye odaklanır. Dağılım varsayımının sağlanamadığı veya verilerin dağılımının belirsiz olduğu durumlar için idealdir.

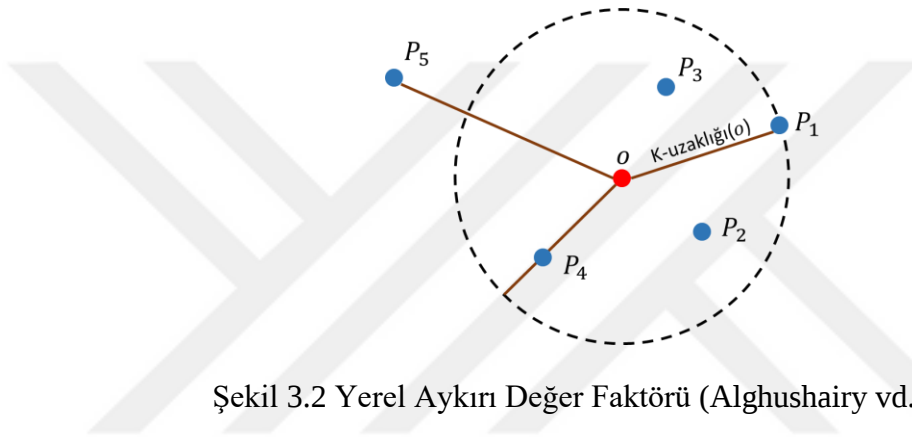
Bu yöntemlerin en büyük avantajı, gerçek veri kümesinin karmaşıklığını ve çeşitliliğini etkin bir şekilde yakalayabilme yeteneğidir. Ancak, bu yöntemlerin dezavantajları arasında hesaplama yükünün daha fazla olması ve işlem sürelerinin daha uzun sürmesi olarak ifade edilebilir.

3.2.1.2.1 Yerel aykırı değer faktörü

Yerel Aykırı Değer Faktörü (LOF), veri noktalarının yerel çevresel yoğunluklarını dikkate alarak anormallikleri tespit etmeyi amaçlar. Bu yöntem, veri kümesindeki anormallikleri belirlemek için veri noktalarının komşuluk ilişkilerini kullanarak parametrik olmayan bir yaklaşım sunar (J. Tang vd. 2002).

LOF, her veri noktası için bir LOF skoru hesaplar ve bu skor, noktanın çevresel yoğunluğunun komşularının yoğunluklarına göre ne kadar anormal olduğunu gösterir. Yüksek bir LOF skoru, noktanın anormal olduğunu işaret eder.

Şekil 3.2’de bir LOF örneği gösterilmektedir. Burada P_4 ile veri noktası o arasındaki gerçek mesafe k -uzaklığı (o)’ndan kısa olduğu için P_4 veri noktası erişilebilir olarak değerlendirilir. Ancak P_6 ile veri noktası o arasındaki gerçek mesafe k -uzaklığı (o)’ndan uzak olduğu için erişilebilir değildir ve aykırı olarak değerlendirilir.



Şekil 3.2 Yerel Aykırı Değer Faktörü (Alghushairy vd. 2020)

LOF genel bir eşik değeri yerine, her veri noktasının benzersiz çevresel özelliklerine dayalı tespitler yaparak anormallikleri belirler. Bu da karmaşık ve düzensiz veri kümelerinde etkili olmasını sağlar. Ancak, yüksek boyutlu veri kümelerinde bu yöntemin performansının düşme riski bulunmaktadır.

3.2.1.2.2 Büyüklük ve momentum histogramı

Büyüklük ve Momentum Histogramı (HoMM), video içindeki her karenin hareket büyüklüğünü ve momentumunu ölçerek olayların normallikten sapma gösterdiği bölgeleri belirlemeye yönelik bir yöntemdir.

HoMM'un çalışma prensibi, ardışık kareler arasındaki her pikselin hareket farkını hesaplamakla başlar. Bu fark, hareketin büyüklüğünü yansıtır ve genellikle piksel yoğunluk farklarının mutlak değerlerinin toplamı olarak hesaplanır. Ardından hareketli

nesnelerin momentumu, yani bir sonraki karedeki konumu ile önceki karedeki konumu arasındaki fark olarak hesaplanır. Elde edilen bu büyüklük ve momentum değerleri, histograma dönüştürülür. Histogram, hareket büyüklüğü ve momentum değerlerini belirli aralıklara bölerek her aralıkta kaç piksel olduğunu gösterir (Bansod ve Nandedkar 2020).

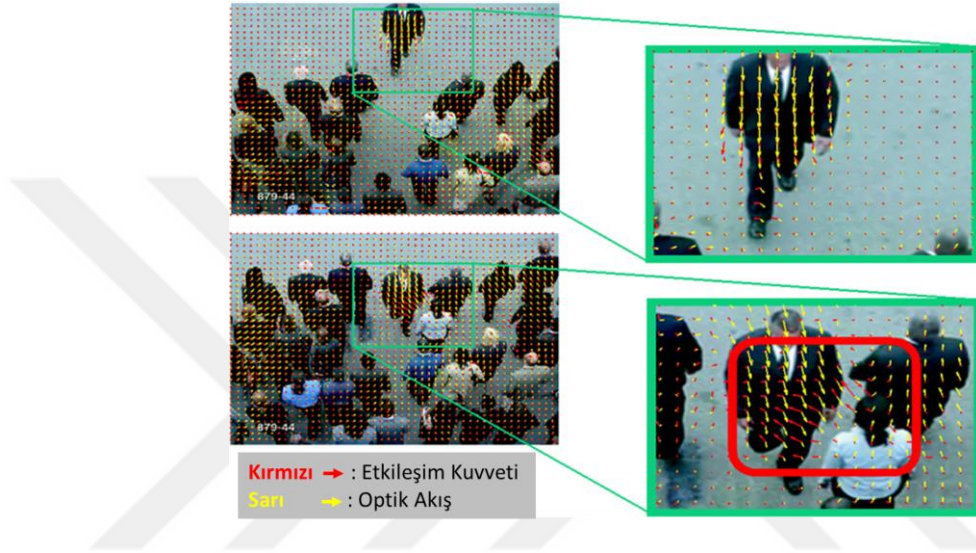
Anormallikler, bu histogramların analizi yoluyla tespit edilir. Büyük hareketler veya beklenmedik momentum dağılımları, belirlenen bir eşik değerinin üzerinde olduğunda, bu bölgeler potansiyel anormallikler olarak kabul edilebilir. Uygulamada, HoMM'un çeşitli varyasyonları ve parametreleri kullanılabilir. Örneğin, histogramın bölümlendirilme sayısı ve momentum aralıkları değişkenlik gösterebilir. Bu, analizin özelleşmiş ihtiyaçlara ve spesifik veri kümelerine uyum sağlamasını kolaylaştırır. HoMM yöntemi, hareket büyüklüğü ve momentum özelliklerini kullanarak video verilerindeki anormallikleri etkin bir şekilde saptayabilir.

3.2.1.2.3 Optik akış

Optik akış, bir video görüntüsündeki nesnelerin veya piksellerin zaman içindeki hareketini analiz etmek ve bu hareketin yönünü ve hızını hesaplamak için kullanılan bir görüntü işleme yöntemidir. Optik akış, nesnelerin hareketini izlemek, hızlarını ve yönlerini belirlemek veya görüntüler arasındaki farkları anlamak gibi birçok uygulama alanında kullanılır (Horn ve Schunck 1981).

Optik akış, birçok algoritma ve yöntemle hesaplanabilir, ancak genelde piksel parlaklık değerlerinin zaman içindeki değişimlerine dayanır. Öncelikle iki ardışık kare arasındaki piksel parlaklık farkları hesaplanır. Bu farklar, nesnelerin veya özelliklerin hareketini yansıtabilir. Piksel parlaklık farklarından yola çıkarak her piksel için bir optik akış vektörü hesaplanır. Bu vektör, pikselin hareket yönünü ve hızını temsil eder. Optik akış vektörleri, farklı hesaplama yöntemleriyle elde edilebilir. Temel yöntemler arasında Lucas-Kanade yöntemi, Horn-Schunck yöntemi ve Pyramidal yöntem bulunur.

Mehran ve arkadaşları (2010) yaptıkları çalışmada optik akış yöntemini kullanarak hareket eden piksellerin etkileşim kuvvetlerini tahmin etmiş ve bunları etkileşim kuvveti olarak adlandırarak bir haritaya dönüştürmüşlerdir. Bu harita vasıtasıyla rastgele seçilen uzamsal-zamansal hacimler, kalabalığın normal davranışını modellemek için kullanmışlardır. Şekil 3.3'te bu çalışmada kullanılan optik akış yöntemi temsili yer almaktadır.



Şekil 3.3 Optik Akış (Mehran vd. 2010)

VGAT alanında, optik akış yöntemiyle normalde statik olan bir bölgedeki ani hareketler tespit edilebilir ve bu hareketler bir anormallik işareti olabilir. Ayrıca, nesnelerin belirli bir bölgeden ayrılma veya o bölgeye doğru hareket etme desenleri de analiz edilebilir. Bu analizler, anormallik tespitine önemli katkılar sağlayabilir.

3.2.1.2.4 Yapısal bağlam tanımlayıcısı

Yapısal Bağlam Tanımlayıcısı (SCD), nesnelerin ve aralarındaki ilişkilerin grafik modelleme aracılığıyla temsil edildiği bir yöntemdir. Bu modelde, her bir nesne bir düğüm olarak temsil edilir ve nesneler arasındaki bağlantılar ise kenarlarla gösterilir. SCD, her nesneyi özelliklerini içeren bir “ağırlık vektörü” ile tanımlar. Bu özellik vektörleri, nesnenin konumu, boyutu, şekli gibi çeşitli özellikleri içerir. Ağırlık vektörleri, nesneler arasındaki ilişkileri ifade eden kenarlarla birbirine bağlanır. SCD,

normal ve anormal davranışları ayırt etmek için bu yapısal bağlantıları modelleyen özellikleri kullanarak makine öğrenme algoritmalarını entegre eder (Y. Yuan vd. 2015).

SCD, nesnelerin ve aralarındaki ilişkilerin yapısal ve bağlamsal özelliklerini kullanır. Yapısal özellikler, nesnelerin konumu, boyutu, şekli gibi fiziksel özelliklerini ifade eder. Örneğin, bir insan ve bir arabanın aynı karede bulunması, bu karenin normal bir davranışı olduğunu gösterir. Ancak, bir insan ve bir bombanın aynı karede bulunması, bu karenin anormal bir davranışı olduğunu gösterebilir.

SCD'nin doğruluğu, modelin eğitimi ve uygulanmasında nesnelerin doğru temsili ve bağlantıların dikkatlice analiz edilmesine bağlıdır. Bu bağlamda, nesnelerin doğru bir şekilde tanımlanması önemlidir; örneğin, bir insanın konumu, boyutu ve şekli gibi özellikleri dikkate alınmalıdır. Ayrıca, nesneler arasındaki bağlantılar doğru bir şekilde oluşturulmalıdır; örneğin, bir insan ve bir arabanın aynı karede bulunması, aralarında bir ilişki olduğunu gösterir. SCD, bu analizlerin yapıldığı durumlarda, karmaşık sahnelerde ve değişen koşullarda anormallikleri daha doğru ve etkin bir şekilde tespit edebilir.

3.2.1.2.5 Kernel yoğunluk tahmini

Kernel yoğunluğu tahmini, veri noktalarının yoğunluklarını tahmin etmek ve veri dağılımını anlamak için kullanılan bir istatistiksel yöntemdir. Bu yöntem, genellikle veri kümesinin yapısını anlamak, aykırı değerleri tespit etmek ve veri kümesindeki desenleri belirlemek amacıyla kullanılır (Terrell ve Scott 1992).

Kernel yoğunluğu tahmini, veri noktalarının çekirdek fonksiyonlarıyla işlenmesiyle başlar ve geçici yoğunluk dağılımlarının birleşimiyle toplam yoğunluk tahmini elde edilir. Bu tahmin, veri kümesinin yapısını grafiksel olarak görselleştirir. Çekirdek fonksiyonları, veri noktalarının etrafında simetrik olarak dağılan olasılık yoğunluklarını temsil eder. Gaussian, Epanechnikov veya Uniform gibi farklı çekirdek fonksiyonları kullanılabilir. Her bir veri noktası için bir çekirdek fonksiyonu uygulanır ve bu fonksiyonlar veri noktalarının yoğunluğuna katkıda bulunur.

VGAT için kernel yoğunluk tahmini, özellikle hareketli nesnelerin yoğunluklarının incelenmesi ve beklenmeyen veya şüpheli davranışların tespiti için kullanılır. Örneğin, trafik akışında anormal bir yoğunluğun tespiti veya kalabalık bir alanda hareket eden bir kişinin anormal davranışının belirlenmesi gibi senaryolarda kullanılabilir. Bu yöntem, verilerin parametrik bir dağılıma uymadığı durumlarda etkili bir araçtır. Ancak, çekirdek fonksiyonunun ve parametrelerinin doğru şekilde seçilmesi önemlidir.

3.2.1.3 İstatistik tabanlı yaklaşımların genel değerlendirmesi

İstatistik tabanlı yaklaşımlar, veri dağılımını ve davranış kalıplarını analiz ederek anormallikleri tespit etme yeteneğine sahiptir. Parametrik olan Gauss Karışım Modeli, Bayesian Yaklaşımı ve Frekans ve Spektral Analiz gibi yöntemler, veri dağılımının belirli bir matematiksel modele uyduğu varsayımına dayanır. Bu yöntemler, modelin veriye ne kadar uygun olduğuna bağlı olarak farklı derecelerde başarı gösterir. Parametrik yöntemlerin gücü, veri dağılımını matematiksel olarak tanımlama ve modelleme yeteneğindedir. Ancak, bu yöntemlerin kısıtlaması, varsayılan matematiksel modelin gerçek veri dağılımını tam olarak yansıtamaması durumunda yanıltıcı sonuçlar verebilir. Parametrik olmayan yöntemler ise, verinin yapısal özelliklerini herhangi bir varsayımsal modele dayandırmadan analiz eder. Bu yöntemler, veri kümesinin yapısal karmaşıklığını ve çeşitliliğini daha doğru bir şekilde yansıtabilir ve genellikle daha esnek çözümler sunar. Ancak, bu esneklik, yüksek hesaplama maliyetleri ve ayarlamaların karmaşıklığı gibi zorlukları da beraberinde getirebilir. Çizelge 3.3'te bu tekniklerin karşılaştırması sunulmuştur.

Çizelge 3.3 İstatistik tabanlı tekniklerin tekniklerin avantaj ve dezavantajları

Teknik	Avantajlar	Dezavantajlar
Parametrik Yöntemler	▪ Analizlerini belirli bir modelin varsayımlarına dayandırarak gerçekleştirir. ▪ Veri dağılımını matematiksel modellerle açıklama yeteneğine sahiptir. ▪ Teorik olarak anlaşılması ve uygulanması daha kolaydır. ▪ Bazı durumlarda hesaplama maliyeti daha düşüktür.	▪ Model, veriye uygun olmadığında yanıltıcı sonuçlar verebilir. ▪ Anormal davranışların çeşitliliği modeli olumsuz etkileyebilir. ▪ Aşırı uç değerler ve gürültülü veriler sonuçları etkileyebilir.
Parametrik Olmayan Yöntemler	▪ Model varsayımlarına bağlı kalmadan veri dağılımını analiz eder. ▪ Veri yapısının doğal özelliklerini daha etkili bir şekilde yakalar. ▪ Daha geniş veri çeşitliliği ile uyumlu olabilir. ▪ Karmaşık ve çok değişkenli verileri daha etkili bir şekilde işleyebilir.	▪ Genellikle daha yüksek hesaplama maliyeti gerektirebilir. ▪ Modelin ve sonuçların yorumlanması daha karmaşık olabilir. ▪ Eğitim ve ayarlama için daha fazla veri ve zaman gerektirebilir.

3.2.2 Sınıflandırma tabanlı yaklaşımlar

Sınıflandırma tabanlı yaklaşımlar, veri örneklerini belirli sınıflara atama prensibine dayanır. Bu yaklaşımlar, genellikle etiketlenmiş verileri içeren bir eğitim sürecini gerektirir. Eğitim aşamasında, önceden belirlenmiş normal ve bazen anormal durumları yansıtan veriler kullanılarak bir sınıflandırıcı model oluşturulur. Bu sınıflandırıcı, verilerdeki desenleri ve özellikleri öğrenir ve bu bilgileri normal durumları tanımlamak için kullanır. Daha sonra eğitilen sınıflandırıcı, gerçek zamanlı videoları analiz ederek normal ve anormal durumları ayırt etmeye çalışır.

Sınıflandırmaya dayalı yaklaşımların esnekliği, belirli bir uygulamaya özel özelliklerin ve sınıflandırıcıların seçilerek özelleştirilmesinden kaynaklanır. Bu yaklaşım, farklı gözetim senaryolarına ve veri türlerine uygun özelleştirilmiş çözümler geliştirmeyi mümkün kılar. Ancak, bu yaklaşımın en büyük zorluklarından biri, güvenilir ve doğru

sonular elde etmek iin yeterli ve temsilci eēitim verisine ihtiya duymasıdır. zellikle, anormal durumların geniř bir yelpazedeki temsili ve farklı kořullarda gsterimi, bu tekniklerin bařarısında kritik bir rol oynar. Eēer sınıflandırıcı doēru řekilde eēitilir ve ihtiya duyulan zelliklerle donatılırsa, VGAT uygulamaları iin olduka etkili bir ara haline gelebilir.

3.2.2.1 Tek sınıflı sınıflandırma

Tek Sınıflı Sınıflandırma (One-Class Classification, OCC), zellikle az sayıda anomali rneēi ieren veri kmelerinde etkili bir sınıflandırma yntemidir. Bu yntem, normal davranıřın karakteristiklerini ērenir ve bu zelliklerin dıřındaki her davranıřı anormallik olarak tanımlar. OCC, model tarafından ērenilen normal davranıř zellikleri kullanılarak normalin sınırlarını izer ve gerek zamanlı verileri bu sınırlarla karřılařtırarak anormallikleri belirler (Tax ve Duin 2001).

OCC zellikle sınıf dengesizliēi, az sayıda anomali rneēi bulunan durumlar veya genellemeye dayalı tespitlerde nemli bir rol oynamaktadır. Bu yntem, normal davranıřı ērenerek ve rneklemden anormal durumları tanımlayarak etkili zmler sunar. rneēin, havaalanı terminalinde normal insan trafiēini temsil eden veriler kullanılarak, beklenmedik bir řekilde boř bir alanda yryen bir kiři gibi anormal durumlar tespit edilebilir.

Bu yntemin bazı sınırlamaları da bulunmaktadır. Az sayıda anomali rneēi ieren veri kmelerinde yetersiz kalabilir ve anormal durumların grntlerde hassas bir řekilde yerleřtirilmesi konusunda zorluklar yařayabilir (Sabokrou vd. 2020). Ayrıca, normal ve anormal davranıřları ayırt ederken hassasiyet dzeyinin doēru belirlenmesi nemlidir. Aksi takdirde, hatalı pozitif veya hatalı sonular elde edilebilir.

3.2.2.2 Hesaplama yoğunluğu temelli sınıflandırma

Hesaplama yoğunluğu temelli sınıflandırma, olayların yoğunluğunu değerlendirerek normal ve anormal davranışları ayırt etmeyi hedefleyen bir yöntemdir. Bu yöntem, olayların anormalliğini belirlemek için zaman ve mekân bilgilerini etkin bir şekilde kullanır.

Bu yöntem, öncelikle hareket durumlarını içeren zaman ve mekân bilgisine dayanan giriş verilerini toplar. Toplanan veriler üzerinde, belirli bir zaman diliminde gerçekleşen olayların sayısını hesaplamak için hesaplama yoğunluğu yöntemleri uygulanır. Bu süreç, verilerdeki olay yoğunluğunu belirlemeyi ve analiz etmeyi içerir. Elde edilen hesaplama yoğunluğu değerleri, normal ve anormal durumları ayırmak için belirlenen bir eşik değeriyle karşılaştırılır. Bu eşik değeri altındaki yoğunluklar normal olarak kabul edilirken eşik değerini aşan yoğunluklar ise anormal olarak sınıflandırılır.

Hesaplama yoğunluğu temelli sınıflandırma, hareketin yoğunluğuna dayanarak hızlı ve etkili bir şekilde normal ve anormal durumları tespit etme potansiyeline sahiptir. Bu yöntem, özellikle yoğunluk değişimlerinin belirgin olduğu senaryolarda faydalı olabilir. Ancak, bu yöntemin etkinliği, eşik değerinin doğru ayarlanmasına ve veri kalitesine bağlıdır. Eşik değerinin yanlış belirlenmesi, hatalı pozitif veya hatalı negatif sonuçlara yol açabilir. Bu nedenle, eşik değerinin belirlenmesi ve ayarlanması, veriye özgü analizler ve testlerle dikkatlice yapılmalıdır.

3.2.2.3 Destek vektör makineleri

Destek Vektör Makineleri (SVM), sınıflandırma ve regresyon problemleri için kullanılan güçlü bir makine öğrenimi yöntemidir. SVM'nin temel amacı, verileri en iyi şekilde ayıran bir hiper düzlem belirleyerek karar sınırı oluşturmaktır. Sınıflandırma problemlerinde, SVM iki veya daha fazla sınıf arasındaki sınırları belirlemek için kullanılırken, regresyon problemlerinde tahmin edilen değerlerin belirlenmesinde etkilidir (Noble 2006).

SVM'nin avantajları arasında yüksek doğruluk oranları, küçük ve orta ölçekli veri kümeleri için uygunluk ve doğrusal olmayan verilerde çeşitli çekirdek fonksiyonlarını kullanabilme özelliği yer alır. SVM, bu çekirdek fonksiyonları sayesinde doğrusal olmayan veri kümelerinde de etkili bir şekilde çalışabilir.

Bununla birlikte, SVM'nin dezavantajları arasında çok sınıflı veriler için zorluklar, büyük veri kümeleri için yüksek hesaplama maliyeti ve hiperparametre seçiminin zorluğu bulunur. Çok sınıflı problemlerde ek yöntemler gerektiren SVM, büyük veri kümelerini işlerken hesaplama açısından yoğun olabilir ve hiperparametrelerin doğru ayarlanması deneyim gerektirir.

SVM, özellikle anormallik tespiti ve sınıflandırma problemlerinde güçlü bir araç olabilir. Ancak, veri kümesinin doğası ve problemin gereksinimlerine uygun olarak parametrelerinin doğru ayarlanması önemlidir. Doğru yapılandırıldığında, SVM karmaşık veri kümelerindeki anormallikleri etkili bir şekilde sınıflandırabilir ve doğruluk ve güvenilirliğini artırabilir.

3.2.2.4 Karar ağaçları ve karar kümeleri

Karar Ağaçları ve Karar Kümeleri, veriyi sınıflara ayırmak veya değer tahmin etmek için kullanılan bir dizi karar kuralına dayalı yöntemlerdir.

Karar ağaçları, kök düğümden başlayarak her düğümde bir özellik karşılaştırması yapar ve veri kümesini bu özelliklere göre alt kümelerde bölünür. Bu bölmeler, verinin en homojen sınıflara ayrılmasını sağlamak amacıyla seçilir. Her düğüm, dallanarak yeni alt düğümler oluşturur ve bu alt düğümler, verinin sınıflandırılmasını belirleyen kuralları içerir. Bu işlem tekrarlanarak ağaç yapısı oluşturulur ve veri kümesi bu yapıya göre sınıflara ayrılır. Yaprak düğümlerinde, verilerin en küçük alt gruplara ayrıldığı ve nihai sınıflandırma sonuçlarının elde edildiği noktalar bulunur (Quinlan 1986).

Karar kümeleri, birden fazla karar ağacının bir araya getirilmesiyle oluşturulur. Bu yöntem, her bir ağacın bağımsız olarak tahmin yapmasını sağlar. Karar kümeleri genellikle toplu öğrenme (ensemble learning) yöntemleri kullanılarak oluşturulur. Bu yöntemler arasında Rastgele Ormanlar (Random Forests-Breiman 2001) ve Gradyan Artırma (Gradient Boosting-Schapire 2003) bulunur. Her bir karar ağacı farklı alt veri kümeleriyle eğitilir ve sonuçlar bir araya getirilerek daha güvenilir ve istikrarlı bir sınıflandırma sonucu elde edilir. Bu sayede overfitting (aşırı uyum) riski azalır ve genel model performansı artar.

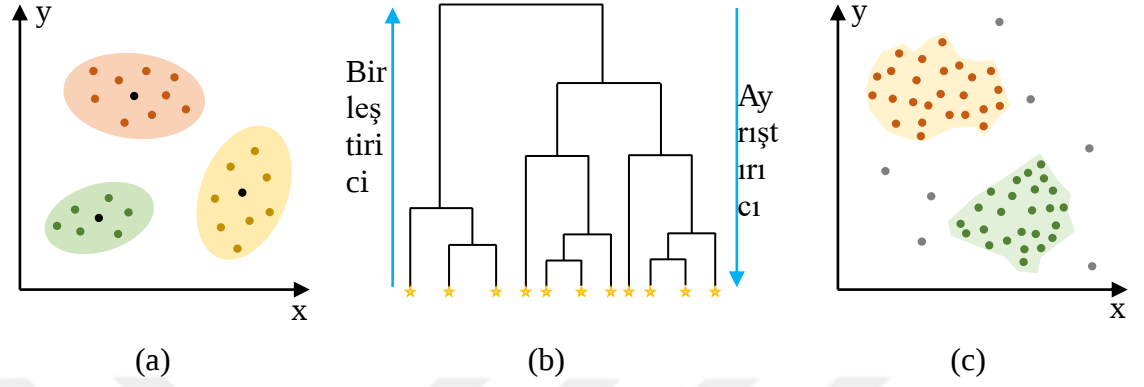
Birden fazla karar ağacını bir araya getirerek oluşturulan bir karar kümesi, her bir ağacın tek başına yapabileceğinden daha sağlam bir anomali tespiti sağlar. Her bir ağaç farklı öğrenme yolları ve hata toleranslarına sahip olabilir. VGAT için bu farklı ağaçların çoğunluk oyu veya ortalama tahminleri kullanılabilir. Bu şekilde, tek bir ağacın hatalarının etkisini azaltmak ve daha genellemeli ve güvenilir bir tespit yapmak mümkün olabilir.

Karar Ağaçları ve Karar Kümeleri'nin avantajlarından biri, verilerdeki karmaşık ilişkileri ve desenleri anlama yetenekleridir. Bu yöntemler, genellikle yüksek boyutlu veri kümeleri üzerinde de etkili sonuçlar elde ederler. Ayrıca, veri ön işleme adımlarına ihtiyaç duymazlar ve doğal olarak kategorik öznitelikleri desteklerler. Diğer bir avantajı her bir kararın nasıl alındığının ve sonuçların nasıl elde edildiğinin ağaç üzerinde açık bir şekilde görülebiliyor olmasıdır. Bu nedenle, karar ağaçları ve karar kümeleri, anlaşılabilirlik, performans ve genel uygulanabilirlik açısından önemli bir seçenek olarak değerlendirilir.

3.2.2.5 Kümeleme

Kümeleme, benzerliklere dayalı olarak verilerin gruplara ayrılmasını sağlayan bir yaklaşımdır. Bu yaklaşımda, normal veriler benzerliklerine göre gruplar oluştururken, anormal veriler herhangi bir gruba ait değildir. Alternatif bir yaklaşıma göre, anormalliklerin küçük kümelerde bulunduğu, tipik veri örneklerinin ise geniş veya

yoğun kümelerde yer aldığı düşünülmektedir (Ranjith vd. 2015). Şekil 3.4'te gösterilen ve yaygın olarak kullanılan kümeleme türleri aşağıda açıklanmıştır:



Şekil 3.4 Kümeleme yöntemleri: (a) Merkez tabanlı kümeleme, (b) Hiyerarşik kümeleme, (c) Yoğunluk tabanlı kümeleme

Merkez tabanlı kümeleme, verilerin belirlenen bir sayıda kümede toplanmasını sağlayan etkili bir kümeleme yöntemidir. Bu yöntemde her küme, bir merkez noktası etrafında oluşturulur ve veriler bu merkeze olan benzerlikleri göz önünde bulundurularak akıllıca gruplandırılır (Zhang 2003). Hiyerarşik kümeleme ise verileri iç içe geçmiş kümelere ayırmak veya birleştirmek için kullanılan bir kümeleme yöntemidir. Bu yöntemde, başlangıçta her veri noktası ayrı bir küme olarak kabul edilir ve daha sonra benzerlik ölçütlerine dayanarak kümeler birleştirilir veya bölünür. Bu işlem sonucunda bir ağaç yapısı şeklinde bir hiyerarşi oluşturulur, bu da verilerin daha ayrıntılı ve anlamlı bir şekilde gruplandırılmasını sağlar.

VGAT bağlamında veri noktaları olarak temsil edilen video kareleri, yoğunluk tabanlı kümeleme yöntemi kullanılarak yoğunluklarına göre kümelendir. Anormal durumlar genellikle yoğunluk açısından farklılık gösterir ve bu nedenle diğer normal karelerden ayrı bir küme oluşturur.

3.2.2.6 Sınıflandırma tabanlı yaklaşımların genel değerlendirmesi

Sınıflandırma tabanlı yaklaşımlar, geniş bir yelpazede, yüksek doğruluk ve hassasiyetle etkili sonuçlar sunabilirler. Eğitim verilerine iyi adapte olduklarında, farklı veri

türleriyle uyumlu şekilde çalışabilir ve çeşitli senaryolarda uygulanabilirler. Ayrıca, bu yöntemlerin esnek yapısı, farklı problemlere ve veri kümelerine göre ayarlanmalarına olanak tanır. Bu nedenle, sınıflandırma tabanlı yaklaşımlar, çeşitli uygulama alanları için çok yönlü çözümler sunabilir. Sınıflandırma tabanlı yaklaşımların etkili bir şekilde uygulanabilmesi, nitelikli ve kapsamlı eğitim verilerine büyük ölçüde bağlıdır. Model karmaşıklığı ve yorumlanabilirlik, bazı durumlarda zorluklar yaratabilir; ayrıca, aşırı öğrenme riski, modelin gerçek dünya verileri üzerindeki performansını olumsuz etkileyebilir. Çizelge 3.4'te bu yöntemlerin karşılaştırması sunulmuştur.

Çizelge 3.4 Sınıflandırma tabanlı yöntemlerin avantaj ve dezavantajları

Yöntem	Avantajlar	Dezavantajlar
Tek Sınıflı Sınıflandırma	- Normal davranışları etkin bir şekilde modelleyebilir. - Az sayıda etiketli veri ile başarılı bir şekilde çalışabilir.	- Anormalliklerin çeşitliliği, modelin performansını etkileyebilir. - Yeni türde anormalliklere karşı duyarlı olabilir.
Destek Vektör Makineleri	- Yüksek boyutlu verilerde üstün performans sergiler. - Ayrım sınırlarını optimize eder. - Farklı çekirdek fonksiyonları ile esnek bir yapıya sahiptir.	- Büyük veri kümeleri için hesaplama yoğunluğu yüksek olabilir. - Parametre seçimi, model başarısını önemli ölçüde etkileyebilir.
Hesaplama Yoğunluğu Temelli Sınıflandırma	- Yoğunluk farklarını kullanarak anormallikleri etkili bir şekilde tespit edebilir. - Farklı veri türlerine başarılı bir şekilde uygulanabilir.	- Veri dağılımının karmaşık olduğu durumlarda zorluklar yaşayabilir. - Hesaplama yoğunluğu yüksek olabilir.
Karar Ağaçları ve Karar Kümeleri	- Anlaşılır ve yorumlanabilir. - Sayısal ve kategorik verilerle uyumlu. - Hızlı eğitim süreci.	- Aşırı öğrenme riski taşır. - Karmaşık ağaçlar, anlaşılması zor sonuçlar verebilir.
Kümeleme	- Etiketsiz verilerle etkili çalışır. - Anormallikleri gruplandırarak başarılı bir şekilde tespit edebilir. - Çeşitli algoritma seçenekleri sunar.	- Optimal küme sayısının belirlenmesi zor olabilir. - Farklı kümeleme algoritmaları farklı sonuçlar verebilir.

3.2.3 Yeniden yapılandırma tabanlı yaklaşımlar

Yeniden yapılandırmaya dayalı yaklaşımlar, genellikle standart video karelerini yeniden oluşturmak için normal video verileri üzerinde eğitilen özel modeller kullanır. Bu yaklaşımda, öncelikle normal video verileri üzerinde eğitilmiş bir model, tipik görüntü kalıplarını öğrenir. Daha sonra, bu model kullanılarak yeniden oluşturulan kareler ile orijinal kareler arasındaki farklar analiz edilir. Eğer bir kare, yeniden oluşturulan karelerle belirgin bir fark taşıyorsa, bu kare anormal olarak işaretlenir. Bu yaklaşım normal verilerin, normal örneklerin ve anormalliklerin çeşitli şekillerde temsil edildiği daha düşük boyutlu bir alana entegre edilebileceği varsayımı altında çalışır (Liu vd. 2018).

Yeniden yapılandırmaya dayalı yöntemlerin çeşitli avantajları bulunmaktadır. Bunlardan ilki, eğitim için önceden etiketlenmiş anormal verilerine ihtiyaç duymadan çalışabilme yetenekleridir. Bu tür yöntemler, genellikle daha kolay elde edilebilen normal verilere odaklanarak anormal verilere kıyasla daha esnek bir yaklaşım sunarlar. Bunun yanı sıra, bir diğer avantajı da modelin normal verileri sadece ezberlemek yerine, bu verileri yeniden oluşturmayı öğrenerek çalışmasıdır. Ancak, bu yaklaşımın bazı dezavantajları da vardır. Normal verilerde belirgin farklılıklar olduğunda, model hatalı pozitif sonuçlar üretebilir. Bu durum, modelin beklenmeyen varyasyonları anormal olarak algılamasına yol açabilir.

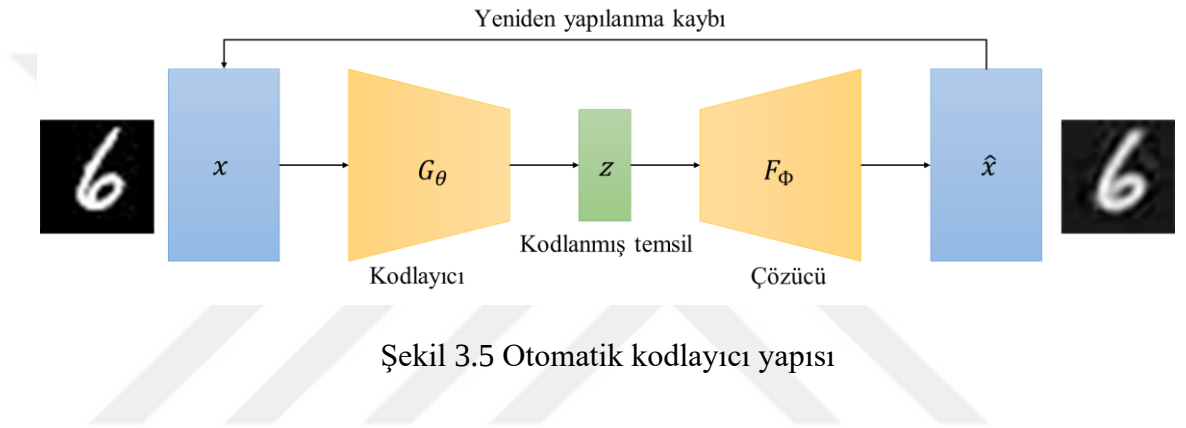
Yeniden yapılandırmada kullanılan teknikler ve yaklaşımlar aşağıda açıklanmıştır.

3.2.3.1 Otomatik kodlayıcı

Otomatik kodlayıcılar (AE), veri kümelerinin daha düşük boyutlu temsillerine dönüştürülmesinde kullanılan derin öğrenme modelleridir (Lecun 1987). Bu temsiller, genellikle verinin gürültüsüz ve önemli özelliklerini yakalamak için öğrenilir ve sıkıştırılmış bir formda saklanır. Bu tür bir temsil, veri analizi, görüntü işleme, video analizi ve birçok diğer uygulama için kullanışlı olabilir. AE'lerin temel amacı, girdi

verisini daha düşük boyutlu bir temsile dönüştürmek ve ardından bu temsili tekrar orijinal girdiye yakın bir şekilde yeniden oluşturmaktır. Bu işlem sırasında, model verinin önemli özelliklerini çıkarabilir ve gereksiz gürültüleri filtreleyebilir.

Şekil 3.5'te temsili gösterilen bir AE'nin genel yapısı en az bir kodlayıcı (encoder) ve bir çözücüden (decoder) meydana gelir. Kodlayıcı girdi verisini daha düşük boyutlu bir temsile dönüştüren kısımdır. Girdi veriyi alır ve onu daha düşük boyutlu bir uzayda temsil eden bir vektöre çevirir. Bu temsil, genellikle verinin önemli özelliklerini içerir.



Kodlayıcının matematiksel formülü Denklem 3.1'de sunulmuştur. Burada, x giriş veriyi, z düşük boyutlu kodlanmış temsili ve $G_\theta()$ kodlayıcı fonksiyonunu temsil eder.

$$z = G_\theta(x) \quad (3.1)$$

Çözücü ise kodlayıcının çıktısını alır ve orijinal girdi verisini yaklaşık olarak yeniden oluşturmaya çalışır. Bu aşamada, düşük boyutlu temsil girdi veriye geri dönüştürülür. Çözücünün matematiksel formülü Denklem 3.2'de sunulmuştur. Burada, $\hat{x}$ geri dönüştürülmüş çıktıyı, $F_\phi()$ ise çözücü fonksiyonu temsil eder.

$$\hat{x} = F_\phi(z) \quad (3.2)$$

Eğitim aşamasında, kodlayıcı ve çözücü aynı anda öğrenilir. AE, x giriş verilerini $\hat{x}$ çıktı verilerine dönüştürebilmek için en uygun $G_\theta()$ ve $F_\phi()$ fonksiyonlarını öğrenmeye çalışır. Bu, giriş ve çıktı verileri arasındaki yeniden yapılandırma kaybını

(reconstruction loss) minimize ederek yapılır. Eğitim sırasında kayıp fonksiyonu, çıktının orijinal girdiye ne kadar iyi benzemesi gerektiğini belirler. Yeniden yapılandırma kaybı, genellikle giriş ve çıktı arasındaki farkın karesinin toplamı olarak hesaplanır. Eğitim süreci, gerçek verilerin encoder ile temsil edilip çözücü ile geri çözülerek, orijinal veriyle karşılaştırılması ve hatanın geriye doğru yayılması ile gerçekleştirilir. Eğitim tamamlandığında, kodlayıcı girdi veriyi daha düşük boyutlu bir temsile dönüştürebilir hale gelir.

VGAT için, başlangıçta normal videolardan oluşturulan bir veri kümesi AE'nin eğitiminde kullanılır ve eğitim süreci boyunca normal veriyi en iyi şekilde sıkıştırmayı başarmak için optimize edilir. Anormal videolar, normal verilerle eğitilmiş AE'ye uygulandığında normal videolardan farklı bir temsil oluştururlar ve bu nedenle yeniden oluşturma işlemi sırasında daha büyük hatalar üretebilirler. Bu hatalar, anormal veriler için bir işaret olarak değerlendirilerek tespit edilmesine yardımcı olur (Jiang vd. 2020).

AE'ler, veri analizi ve işlemeye yönelik bir dizi önemli avantaja sahiptir. Birincil avantajlarından biri, boyut indirgeme ve veri sıkıştırma yetenekleridir (Michelucci 2022). AE'ler, yüksek boyutlu veriyi daha düşük boyutlu temsillerle ifade ederek, veri kümelerinin daha etkili bir şekilde yönetilmesine ve işlenmesine olanak tanır. Bu, bellek kullanımını azaltabilir, işlem sürelerini kısaltabilir ve daha hızlı analiz imkânı sağlayabilir. Ayrıca AE'ler, gürültü giderme kapasitesi sayesinde veri kümelerinin kalitesini artırabilir ve veri kümelerindeki istenmeyen veya bozulmuş bilgileri temizleyebilir. Aykırı değerlerin tespiti gibi anormallik analizlerinde de kullanılarak, veri kümelerindeki anormal örüntüleri tanımlama yeteneği sunarlar. Bunun ötesinde, varyasyonel otomatik kodlayıcılar (VAE'ler) gibi türler, veriyi özgün şekillerde üretme yeteneğiyle üretken modellemeyi destekler.

Ancak, AE'lerin bazı dezavantajları da vardır. Bunların ilki, AE'lerin eğitim süreçlerinin karmaşıklığıdır. Doğru hiperparametre ayarları, uygun kayıp fonksiyonları ve eğitim verisinin iyi seçilmesi gerekebilir. Aksi halde, kötü eğitilmiş bir AE, veri kümesini yeterince iyi temsil edemeyebilir. Ayrıca, AE'lerin daha büyük boyutlu verilerle çalışırken performansları azalabilir ve hesaplama yükü artabilir. AE'lerin doğru ve etkili

bir şekilde kullanılması, bir dizi deneme-yanılma süreci gerektirebilir ve bu da zaman alabilir. Son olarak, AE'ler, bazı veri kümelerinde aşırı öğrenmeye veya öğrenilen temsillerin anlamsız olabileceği riskine karşı duyarlı olabilirler.

Pek çok farklı AE türü mevcuttur ve bu çeşitlilik, veri analizi ve işlemeye geniş bir yelpazede uygulamalar sunmaktadır. Bu farklı türler, aşağıda daha ayrıntılı bir şekilde ele alınmıştır:

1. Evrişimli AE (ConvAE): Hem kodlayıcı hem de kod çözücü ağlarda tam bağlı katmanlar yerine evrişimli katmanlar kullanan bir AE türüdür. Özellikle görüntü verileri için tasarlanmışlardır. Evrişimli katman, girdiye bir dizi filtre uygulayan ve verilerdeki yerel örüntüleri yakalayan bir özellik haritası üreten bir katmandır. ConvAE, parametre sayısını azalttığı ve verilerin uzamsal yapısını koruduğu için görüntü verilerinden uzamsal özellikleri normal bir AE'den daha verimli ve etkili bir şekilde öğrenebilir. Uzamsal özellikler daha sonra veri dağılımının değişkenliğini dikkate alan anormallik tespiti ve sınıflandırması için kullanılır (Ehsani vd. 2022).

VGAT alanında bu yöntem, video karelerini evrişimli katmanlar ve ters evrişimli katmanlar kullanarak sıkıştırır ve genişletir. ConvAE, normal kareleri doğru bir şekilde yeniden oluştururken, anormal kareleri hatalı bir şekilde yeniden oluşturur. Bu şekilde, yeniden oluşturma hatası yüksek olan kareler anormal olarak sınıflandırılabilir.

2. Derin AE (DAE): Bu AE'ler, tamamen bağlantılı katmanlardan oluşan bir yapısı ile öne çıkar ve giriş verisini daha yüksek düzeyde soyutlamalarla temsil etmeye odaklanırlar. Genellikle aralarında girdi ve çıktı boyutlarından farklı boyutlara sahip gizli katmanlar bulunur. Bu yapı, daha fazla özellik çıkarımını mümkün kılar. DAE'ler, video karelerini çok katmanlı bir yapay sinir ağı kullanarak sıkıştırma ve genişletme işlemlerine tabi tutar (Zhou ve Paffenroth 2017).

Her bir katman, bir önceki katmandan gelen girdiyi daha düşük boyutlu bir temsile dönüştürme görevini üstlenir. Bu katmanlar, verinin özüne inerek temel özellikleri

yakalama kapasitesine sahiptir. Son katman ise, en düşük boyutlu temsili üretir. Bu temsil, verinin özünü içerir ve veri kümesinin önemli özelliklerini yansıtır. Daha sonra, bu temsil ters yönde aynı işlemlerle kullanılarak orijinal girdi verisi yeniden oluşturulur. Bu işlem, girdi verisinin en önemli özelliklerini korurken gereksiz ayrıntıları filtreleyerek veri yeniden oluşturulmasını sağlar.

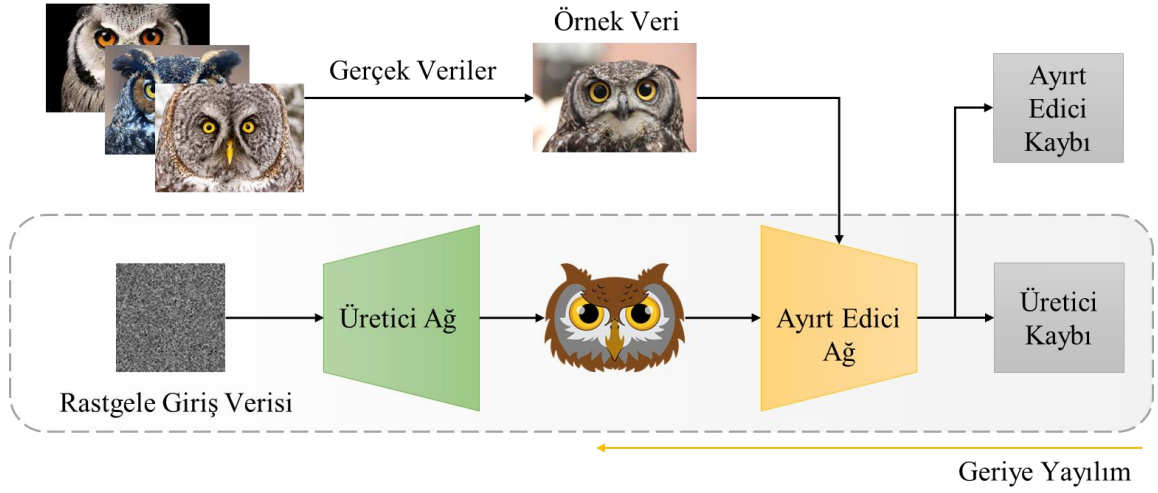
3. Varyasyonel AE (VAE): VAE, verilerin temsillerini öğrenirken, aynı zamanda gizli değişkenlerin dağılımını da modellemeyi hedefler. En belirgin özelliği, giriş verilerinin temsili tek bir nokta yerine bir olasılık dağılımı olarak yaklaşımını benimsemesidir. Bu sayede verilerin daha hassas ve gerçekçi bir temsili elde etmek mümkün olurken, veriler arasındaki benzerlikler ile farklılıklar daha iyi anlaşılır (Doersch 2016).

VAE, temelde normal bir AE'nin yapısına benzemesine rağmen, giriş verilerine ek olarak bir gizli değişken öğrenme yeteneğine sahiptir. Bu gizli değişken, verinin farklı özellik kombinasyonlarını temsil eder. VAE, bu gizli değişkenin olasılık dağılımını modeller ve veriyi temsil ederken öğrenilen bu dağılımı kullanır.

VAE'nin temsili ve veri oluşturma sürecindeki olasılık dağılımları, veriler arasındaki benzerlikleri ve farklılıkları daha hassas bir şekilde yakalamasına olanak tanır. Bu nedenle, özellikle veri varyasyonlarının ve kompleks yapının önemli olduğu durumlarda, VAE büyük bir avantaj sağlar (Cheng vd. 2020).

3.2.3.2 Üretici çekişmeli ağlar

Üretici Çekişmeli Ağlar (GAN), Ian Goodfellow ve arkadaşları (2014) tarafından geliştirilen bir derin öğrenme yöntemidir. GAN'lar, özellikle görüntü sentezi, görüntüyü yeniden oluşturma ve benzeri uygulamalarda büyük başarılar elde etmiştir. Bu yöntem, üretici (generator) ve ayırt edici (discriminator) olmak üzere iki ayrı yapay sinir ağı modelini içerir. Genel çalışma yapısı Şekil 3.6'da temsil edilmiştir.



Şekil 3.6 Üretici Çekişmeli Ağ

Üretici ağ, eğitildiği veri kümesini analiz ederek normal veri dağılımını anlamaya çalışır. Bu sayede, gerçekçi normal verilere benzeyen veriler üretir. Aynı zamanda, normal verilerden sapmalar gösteren anormal verileri de sentezleyebilir. Bu yetenek, modelin anormal durumları anlama ve tanımlama yeteneğini güçlendirebilir. Üretici ağın eğitimi sırasında kullanılan kayıba “Üretici Kaybı (Generative Loss)” denir. Üretici ağ, ürettiği verileri gerçek verilere o kadar yakın yapmaya çalışır ki ayırt edici ağ bunları ayırt edemez hale gelir. Üretici kaybı, ürettiği verilerin gerçeklik derecesini ve ayırt edici ağın bu üretilmiş verilere karşı ne kadar yanıltılabildiğini ölçer. Üretici kaybının azalması, üretici ağın daha gerçekçi veriler üretebildiğini ve ayırt edici ağı daha etkili bir şekilde yanıltabildiğini gösterir.

Ayırt edici ağ, normal ve anormal veriler arasındaki farkları öğrenerek anormal durumları tespit etmeye çabalar. Eğitilen model, bir veriyi normal olarak sınıflandıramazsa, bu veri anormal olarak kabul edilir. Ayırt edici ağın eğitimi sırasında kullanılan kayıba “Ayırt Edici Kaybı (Discriminator Loss)” adı verilir. Ağ, gerçek verileri ve üretilmiş verileri ayırt etmeyi öğrenirken eğitilir. Ayırt edici kaybı, gerçek verilerin doğru bir şekilde sınıflandırılması ile üretilmiş verilerin yanlış bir şekilde sınıflandırılması arasındaki farkı ölçer. Bu kaybın azalması, ayırt edici ağın daha iyi bir şekilde gerçek ve üretilmiş verileri ayırt etmeyi öğrendiğini gösterir.

Özellikle sınırlı anormal veri durumlarında GAN'lar, veri çeşitliliğini artırmak ve modelin genelleme yeteneğini geliştirmek için etkili bir araç olarak kullanılabilir. Üretici ağ, mevcut anormal verilere dayalı olarak yeni anormal veriler üretme yeteneğine sahiptir ve bu, veri kümesinin çeşitliliğini artırarak modelin daha genel anormal durumları tanıma yeteneğini iyileştirebilir.

VGAT için GAN yönteminin uygulanmasının adımları şu şekildedir: İlk adımda, üretici ağ normal veri kümesiyle eğitilir, böylece normal görüntüler üretebilir hale gelir. İlk adımda, üretici ağ, normal veri kümesiyle eğitilir ve böylece normal görüntüler üretebilme yeteneği kazanır. Üretici ağın mimarisi ve hiperparametreleri, veri kümesinin boyutu ve özellikleri göz önünde bulundurularak özenle ayarlanır. Ardından, test videolarındaki anormal görüntüler ayırt edici ağa girdi olarak verilir. Bu aşamada, ayırt edici ağ, üretilen görüntüler ile gerçek görüntüler arasındaki farkları değerlendirerek anormallik skorlarını üretir. Anormallik skorları, bir görüntünün ne kadar anormal olduğunu gösteren ölçülerdir. Eğer bir görüntünün anormallik skoru belirlenmiş bir eşiği aşarsa, o görüntü anormal olarak kabul edilir ve bir alarm tetiklenir. Bu, potansiyel anormalliklerin tespit edilmesi ve hızla müdahale edilmesi için hayati bir öneme sahiptir.

3.2.3.3 Yeniden yapılandırma tabanlı yaklaşımların genel değerlendirmesi

AE'ler, verinin öz niteliklerini etkili bir şekilde sıkıştırarak ve daha sonra yeniden yapılandırarak anormallikleri belirlemede kullanılır. GAN'lar ise, gerçekçi veri örnekleri üreterek ve bunları gerçek verilerden ayırt etmeye çalışarak çalışır. Bu teknikler, verinin karmaşık yapısını öğrenmede ve anormallikleri tespit etmede oldukça etkili olabilir. Çizelge 3.5'te bu yöntemlerin karşılaştırması sunulmuştur.

Çizelge 3.5 Yeniden yapılandırma tabanlı yöntemlerin avantaj ve dezavantajları

Yöntem	Avantajlar	Dezavantajlar
Otomatik Kodlayıcılar	- Verinin özniteliklerini etkili bir şekilde sıkıştırır ve yeniden yapılandırır. - Anormallikleri tespit etmede yüksek hassasiyet sunabilir. - Farklı veri türleri ve senaryolarla uyumlu olabilir. - Model eğitimi ve ayarlaması nispeten daha basittir.	- Eğitim verilerinin kalitesine ve çeşitliliğine duyarlıdır. - Yüksek hesaplama maliyeti gerektirebilir. - Sınırlı sayıda anormallik türünü tespit edebilir. - Modelin veriye uygunluğu sonuçların doğruluğunu etkiler.
Üretici Çekişmeli Ağlar	- Gerçekçi ve çeşitli veri örnekleri üretebilir. - Veri dağılımını modellemede güçlü yeteneklere sahiptir. - Verinin karmaşık yapılarını öğrenmede etkilidir. - Yaratıcı ve yenilikçi veri analizi uygulamalarına olanak tanır.	- Eğitim süreci karmaşık ve zorlayıcı olabilir. - Yüksek hesaplama gücü ve kaynak ihtiyacı. - Gerçek ve üretilmiş veriler arasındaki ayrım zorlukları. - Üretici ve ayırt edici ağlar arasındaki dengenin bozulması sonucu üretilen veri çeşidinin azalması.

3.2.4 Tahmin tabanlı yaklaşımlar

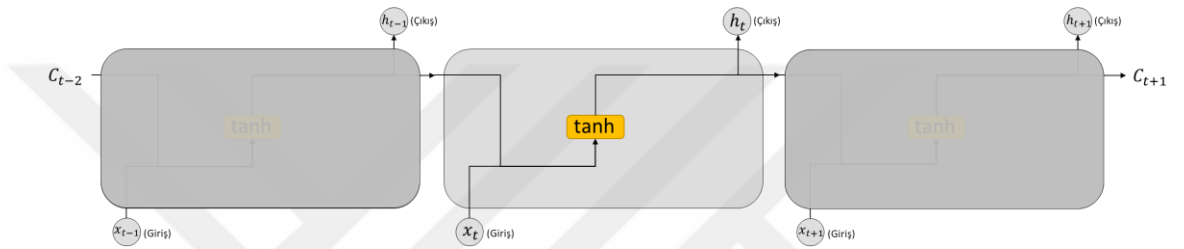
VGAT alanında tahmine dayalı yaklaşımlar, bir model kullanarak videonun gelecekteki karelerini tahmin etmek ve ardından tahmin edilen kareleri gerçek karelerle karşılaştırarak anormallikleri tespit etmek temeline dayanır (Liu vd. 2018). Bu modeller normal aktivitelerin tahmin edilebilir olduğunu varsayar ve tahminden sapma bir anormalliği gösterir. Bu alanda yaygın olarak kullanılan bazı yaklaşımlar ve yöntemler aşağıda anlatılmıştır.

3.2.4.1 Tekrarlayan sinir ağı

Tekrarlayan Sinir Ağı (RNN), zaman serileri ve sıralı verilerin işlenmesi için özel olarak tasarlanmış yapay sinir ağı türleridir. RNN'ler, önceki adımlardaki bilgileri anlamak ve mevcut adımın sonuca etkisini hesaba katmak için belleğe sahiptir. Bu

özellikleri sayesinde, zaman içinde değişen verileri işlemek ve sıralı bağlantıları anlamak için oldukça etkilidirler. Video gözetimi gibi durumlarda, ardışık kareler arasındaki ilişkileri anlamak için RNN'ler kullanılır.

RNN'lerin temel işleyi, hafıza hücreleri ya da “gizli durum” olarak adlandırılan içsel durumlar aracılığıyla gerçekleşir. Bu gizli durumlar, geçmiş zaman adımlarında taşınan bilgileri içerir (Jain ve Medsker 1999). Karmaşık problemlerde temsili Şekil 3.7’de gösterilen ve birden fazla hücre içeren RNN'ler kullanılır.



Şekil 3.7 Tekrarlayan Sinir Ağı (Colah 2015)

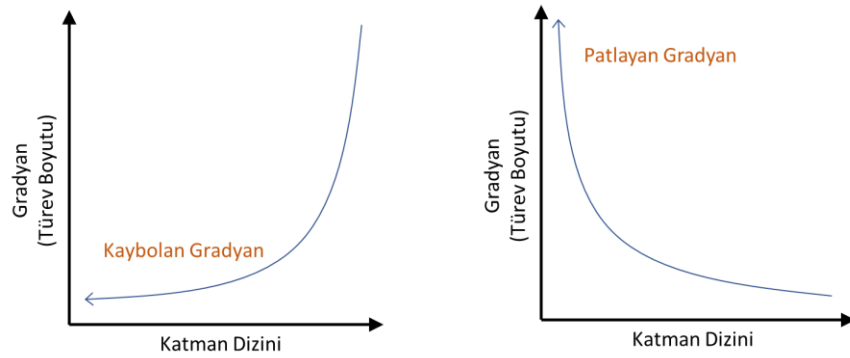
Tek bir RNN hücresinin matematiksel formülü Denklem 3.3’te sunulmuştur. Burada h_t mevcut zaman adımında RNN hücresinin çıkışını (gizli durumu), x_t mevcut zaman adımında girdi verisini, h_{t-1} gizli durumun önceki zaman adımından gelen gizli durumunu temsil eder. W_x girdi ağırlık matrisi ve W_h önceki gizli durum ağırlık matrisidir. b ise gizli durumun geri dönüşüne eklenen bias terimidir. Bu formül, mevcut girdi (x_t) ve önceki gizli durum (h_{t-1}) kullanılarak yeni gizli durumu (h_t) hesaplar. Hiperbolik tanjant aktivasyonu, gizli durumun değerlerini sınırlar ve genellikle RNN hücrelerinde yaygın olarak kullanılan bir aktivasyon fonksiyonudur.

$$h_t = \tanh(W_x x_t + W_h h_{t-1} + b) \quad (3.3)$$

RNN’in çalışma prensibi incelendiğinde, öncelikle eğitim verileri kullanılarak, normal davranış örnekleri öğrenilmektedir. Ardından, girdi verilerini kullanarak çıktı tahminleri yapar ve beklenen çıktılara göre hataları hesaplar. Bu süreç sayesinde ağ, normal davranışı anlamaya yönelik bir temel oluşturur ve beklenmeyen durumları tanımak için bir referans noktası oluşturur. Eğitilen RNN ağı video akışını analiz eder ve her zaman

adımında, mevcut girdi verisini kullanarak çıktı tahminlerinde bulunur. Anomali tespiti, tahmin edilen çıktı ile gerçek girdi arasındaki farkın belirlenen eşik değerini aşmasıyla gerçekleşir.

RNN'lerin eğitim sürecinde Şekil 3.8'de temsili gösterilen patlayan veya kaybolan gradyan sorunu ile karşılaşılabilir. Bu da gradyan değerlerinin aşırı büyümesi veya aşırı küçülmesi sonucunu doğurabilir. Patlayan gradyan sorunu, ağı eğitilirken gradyan değerlerinin aşırı büyümesi sonucunda ortaya çıkar. Gradyan değerleri çok büyük olduğunda, ağı parametreleri hızla güncellendiğinden, dalgalanmalar veya aşırı değerlerle oluşabilmektedir. Bu durum, eğitim sırasında modelin istenmeyen şekilde dalgalanmasına ve istikrarsızlaşmasına neden olabilir. Patlayan gradyan sorunu genellikle büyük öğrenme oranları veya ağı mimarileri kullanıldığında meydana gelir. Kaybolan gradyan sorunu ise, gradyan değerlerinin eğitim ilerledikçe giderek küçülmesi veya sıfıra yaklaşması sonucu ortaya çıkar. Derin yapıları ağlarda, geriye doğru yayılan gradyanlar, her katmanda çarpma işlemi ile daralır. Bu da başlangıç katmanlarına doğru gradyanların giderek zayıflamasına ve bu katmanların etkili bir şekilde eğitilememesine yol açabilir. Sonuç olarak, derin ağların alt katmanları iyi bir şekilde öğrenilmez ve modelin genel performansı düşer.



Şekil 3.8 Kaybolan ve Patlayan Gradyan (Pandey 2022)

Patlayan veya kaybolan gradyan sorunu, RNN'lerin uzun vadeli bağımlılıkları öğrenme yeteneklerini kısıtlayabilir (Van Houdt vd. 2020; Vennerød vd. 2021). Ancak, bu tür zorluklara rağmen, RNN'ler hala birçok uygulama alanında kullanılmaktadır. Zira, bu

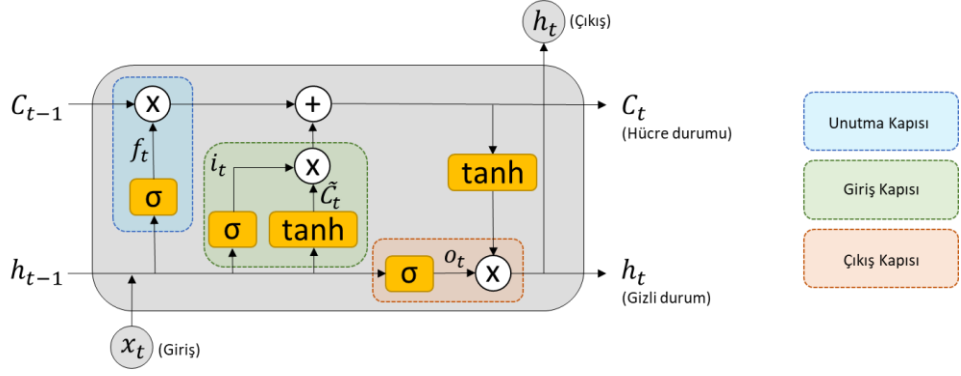
ağlar geçmiş bilgileri mevcut zamandaki tahminlere entegre ederek dinamik desenleri yakalama yetenekleri sunar.

3.2.4.2 Uzun kısa süreli hafıza

RNN'lerin, özellikle Uzun Kısa Süreli Hafıza (LSTM) olarak adlandırılan özel bir türü bulunmaktadır. LSTM'ler, geçmiş durumları etkili bir şekilde anımsayabilme yeteneğine sahip özelleştirilmiş hücreler içerir. Bu hücreler, girdileri işlerken geçmiş bilgileri muhafaza edebilir ve çıktılar üretirken de bu bilgileri dikkate alabilirler. Bu özellik sayesinde, gelecekteki tahminlerini yaparken önceki zaman dilimlerindeki kritik bilgileri kullanabilme kabiliyetine sahip olur.

LSTM'lerin öne çıkan özelliklerinden biri, RNN'lerde sıkça rastlanan patlayan gradyan ve kaybolan gradyan sorunlarının üstesinden gelmek için özel olarak tasarlanmış “bellek hücreleri” adı verilen bir yapıya sahip olmalarıdır. Bu bellek hücreleri, giriş, çıkış ve unutma kapıları gibi farklı bileşenleri içerir ve bellek hücresine gelen bilgilerin akışını düzenler. Bu yapı sayesinde, LSTM hücreleri içerdikleri bilgileri zaman içinde etkili bir şekilde koruyabilir ve gereksiz, hatırlanması gerekmeyen bilgileri unutabilirler. Bu özel tasarım, LSTM'leri RNN yapısına kıyasla daha uzun vadeli bağlantıları yakalayabilen, daha karmaşık zaman serisi verilerini işleyebilen ve daha güvenilir tahminlerde bulunabilen güçlü bir model haline getirir.

LSTM hücresinin içyapısı ve bileşenleri Şekil 3.9'da ayrıntılı bir şekilde gösterilmiştir. Bu bileşenler arasında giriş kapısı, bellek hücresine hangi yeni bilgilerin depolanacağına karar verirken, çıkış kapısı bellek hücresinden hangi bilgilerin çıkacağına karar verir. Unutma kapısı ise bellek hücresindeki belirli bilgilerin ne zaman silineceğine karar verir. Bu sayede hücre, gereksiz veya artık önem taşımayan bilgileri unutabilir. Bu mekanizma, patlayan veya kaybolan gradyan gibi sorunları aşmada da büyük bir rol oynar. Bu yönüyle, LSTM'ler, karmaşık ve uzun vadeli bağlantılara sahip zaman serilerini başarılı bir şekilde modelleme yetenekleriyle öne çıkarlar (Van Houdt vd. 2020; Vennerød vd. 2021).



Şekil 3.9 LSTM hücresi (Colah 2015)

LSTM hücresinde giriş kapısı (i_t), unutma kapısı (f_t) ve çıkış kapısı (o_t) gibi önemli bileşenler bulunmaktadır. Bu bileşenler ve işleyiş aşağıda detaylandırılmıştır:

1. Giriş Kapısı (Input Gate): Bu kapı, yeni bilgilerin hücre durumuna eklenme oranını kontrol eder. Yani, hangi bilgilerin hafızaya alınacağı bu kapı aracılığıyla belirlenir. Matematiksel olarak, giriş kapısının işleyişi Denklem 3.4'te sunulmuştur. Sigmoid aktivasyon fonksiyonu (σ), LSTM hücresindeki farklı kapıları kontrol etmek için kullanılır. Bu fonksiyon, değerleri 0 ile 1 arasında sıkıştırarak kapıların açılma veya kapanma oranını belirler.

$$i_t = \sigma(W_i * [h_{t-1}, x_t] + b_i) \quad (3.4)$$

Denklem 3.5'te gösterilen $\tilde{C}_t$ yeni hücre durumuna eklenecek olan bilgiyi temsil eder. Bu bilgi, mevcut girdi bilgilerin (x_t) ve önceki zaman adımındaki durumun (h_{t-1}) kombinasyonundan türetilir. Hiperbolik tanjant ($\tanh$) aktivasyon fonksiyonu ise, LSTM'nin içinde yer alan birçok yerde kullanılır ve değerleri $[-1, 1]$ arasında sıkıştırarak olası değerleri daraltır.

$$\tilde{C}_t = \tanh(W_c * [h_{t-1}, x_t] + b_c) \quad (3.5)$$

2. Unutma Kapısı (Forget Gate): Hücre durumundan yeni bilgilerle birlikte hangi bilgilerin unutulacağı kontrol edilir. Denklem 3.6'te eşitliği verilen W_f unutma kapısının ağırlık matrisidir. Bu matris, önceki zaman adımındaki durumun (h_{t-1})

mevcut girdi bilgilerin (h_{t-1}) arasındaki ilişkiyi öğrenir. Sigmoid (σ) aktivasyon fonksiyonu ise unutma kapısının çıkışını $[0 - 1]$ aralığına sıkıştırarak değeri ağırlıklı bir olasılık olarak yorumlamamızı sağlar. Çıkış değeri ne kadar büyükse, o kadar fazla bilgi unutulacaktır.

$$ff_t = \sigma(W_f * [h_{t-1}, x_t] + b_f) \quad (3.6)$$

3. Çıkış Kapısı (Output Gate): Hücre durumundan çıkacak olan bilginin düzenlenmesini denetler ve yeni hücre durumuna bağlı olarak çıkışın özelliklerini şekillendirir. Bu kapı, LSTM'in bilgi akışını optimize ederek, en anlamlı ve uygulanabilir çıkışları üretmesine yardımcı olur. Denklem 3.7'te eşitliği verilen " o_t " kapının çıkışıdır ve sigmoid aktivasyon fonksiyonu aracılığıyla hesaplanır. Eğer çıkış değeri yaklaşık olarak 1'e yakınsa, hücre durumunun tamamı çıkış olarak kullanılacak, 0'a yakınsa hiçbir çıkış olarak kullanılmayacaktır. W_o , giriş kapısının ağırlık matrisidir. Bu matris, önceki durum (h_{t-1}) ve mevcut girdi (x_t) arasındaki ilişkiyi öğrenir. Önceki durum (h_{t-1}), önceki zaman adımındaki durumu temsil eder. Bu, LSTM hücresinin önceki durumudur ve giriş kapısının hangi bilgileri kullanacağını belirler. " x_t " ise dış mevcut zaman adımındaki girdiyi temsil eder.

$$o_t = \sigma(W_o * [h_{t-1}, x_t] + b_o) \quad (3.7)$$

Denklem 3.8'de gösterilen mevcut zaman adımındaki çıkış veya gizli durum (h_t) LSTM'nin çıktı değeri olarak düşünülebilir. Hücre durumundan türetilir ve mevcut zaman adımında giriş verisine bağlı olarak hesaplanır.

$$h_t = o_t * \tanh(C_t) \quad (3.8)$$

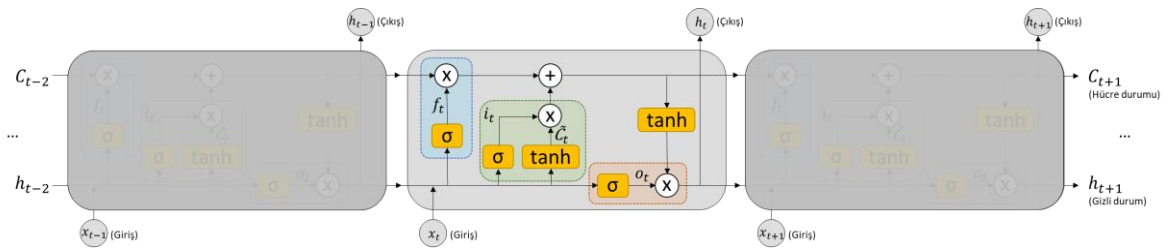
4. Hücre Durumu (Cell State): LSTM'nin ana hafıza birimidir. Bilgilerin uzun süreli olarak saklandığı bölümdür. Hücre durumu, giriş kapıları ve unutma kapıları sayesinde güncellenir ve yönetilir. Eşitliği Denklem 3.9'da verilmiştir. " f " ile önceki hücre durumundaki bilgilerin bir kısmı unutulurken " i_t " ve " $\tilde{C}_t$ " yeni bilgiler eklenir. Bu sayede, hücre durumu geçmiş bilgileri unutturken, yeni bilgileri de ekleyerek güncellenir.

$$C_t = f_t * C_{t-1} + i_t * \tilde{C}_t \quad (3.9)$$

$W_c, W_i, W_f, W_o, b_i, b_c, b_o, b_f$: Bu ağırlıklar ve sapmalar LSTM'nin öğrenilebilir parametreleridir. Eğitim sürecinde bu parametreler optimize edilerek LSTM modelinin işlevi belirlenir.

Birden fazla LSTM hücresi, birbirleriyle bir dizi katman oluşturarak bir LSTM modeli oluşturur. Bu katmanlar, verileri işlemek ve zaman içindeki bağlantıları modellemek için kullanılır. Birçok LSTM hücresi içeren bu yapı, daha karmaşık ve daha derin öğrenme yeteneklerine sahip bir model elde etmemizi sağlar.

LSTM'lerin birden fazla hücresi arasındaki ilişki, bir önceki hücrenin çıkışı (gizli durumu) bir sonraki hücrenin girişi olarak kullanılarak sağlanır. Böylece, bir LSTM katmanı içindeki hücreler, sırayla verileri işler ve bilgi akışını korur. Çok katmanlı bir LSTM modeli, her bir katmanın ardışık olarak çalıştığı bir yapıya sahiptir. Bu, bilgi işlemenin katmanlar arasında gerçekleştiği anlamına gelir. Bir zaman adımında, giriş verisi ilk katmandan başlar, ardından her bir hücreden geçerek diğer katmana ilerler. En son katmandan çıkan çıktı değeri ise modelin tahminini veya sonucunu oluşturur. Şekil 3.10'da birden fazla LSTM hücresinin bağlantı yapısı ve veri iletimi gösterilmektedir.



Şekil 3.10 Birden fazla LSTM hücresinin bağlantı yapısı ve veri iletimi

LSTM'nin performansını iyileştirmeyi veya bazı sınırlamalarını gidermeyi amaçlayan bazı varyantları veya uzantıları da vardır. Bunlardan bazıları şunlardır (Hu vd. 2020):

1. Peephole LSTM: Standart LSTM, giriş kapısı, çıkış kapısı ve unutma kapısı kullanırken, peephole LSTM'de bu kapılar aynı zamanda hücre durumunu da görebilir.

Yani, kapılar önceki gizli durumu ve mevcut girdiyi değil, aynı zamanda hücre durumunu da dikkate alır. Böylece kapıların daha fazla bilgiye erişmesi sağlanır. Bu, ağın daha hassas zamanlama ve sayma davranışları öğrenmesine yardımcı olabilir (Gers ve Schmidhuber 2000).

2. Coupled LSTM: Geleneksel LSTM, giriş kapısı, çıkış kapısı ve unutma kapısı olmak üzere üç ayrı kapı kullanır. Ancak coupled LSTM'de bu kapılar birleştirilerek tek bir kapı kullanılır. Bu kapı, neyi unutacağına ve neyi ekleyeceğine aynı anda karar verir. Bu yaklaşım, parametre sayısını azaltabilir ve eğitim sürecini hızlandırabilir.

3. Gated Recurrent Unit (GRU): Bu varyant, mimariyi daha da basitleştirerek hücre durumunu ve gizli durumu tek bir durum vektörüne birleştirir. İlave olarak sıfırlama kapısı ve güncelleme kapısı kullanır. Sıfırlama kapısı, önceki durumu ve mevcut girdiyi nasıl birleştireceğine karar verirken, güncelleme kapısı, önceki durumun ne kadarını koruyacağına veya güncelleyeceğine karar verir. GRU, LSTM'den daha hızlı ve daha kolay eğitilebilir ancak uzun süreli bağımlılıkları yakalama yeteneği biraz daha sınırlı olabilir (Chung vd. 2014).

4. Bidirectional LSTM: Bu varyant, bir girdi dizisini hem soldan sağa hem de sağdan sola işleyen iki ayrı LSTM kullanır. Her bir LSTM, girdi dizisinin farklı yöndeki bağlamını temsil eder. İki LSTM'nin çıktıları daha sonra birleştirilerek son çıktıyı oluşturur. Bu yaklaşım, ağın hem geçmiş hem de gelecek bağlamını yakalamasına olanak sağlar (Schuster ve Paliwal 1997). Bidirectional LSTM, doğal dil işleme veya konuşma tanıma gibi görevlerde sıklıkla kullanılabilir.

5. Attention LSTM: Bu varyant, ağın bağlama bağlı olarak girdi dizisinin farklı kısımlarına odaklanmasına izin veren bir dikkat mekanizması içerir. Dikkat mekanizması, LSTM'nin tüm gizli durumlarının ağırlıklı bir toplamını hesaplar. Bu ağırlıklar, her gizli durumun mevcut çıktı için ne kadar önemli olduğuna bağlı olarak belirlenir. Dikkat mekanizması ağın uzun ve karmaşık girdi dizilerini daha etkili bir şekilde işlemesine yardımcı olabilir. Ağırlıklar, ağın odaklanması gereken kısımları belirlemesine yardımcı olarak daha iyi sonuçlar elde edilmesini sağlar (Mnih vd. 2014).

VGAT alanında, LSTM normal video karelerinin zamansal dinamiklerini modellemek ve öğrenilen kalıplardan sapmaları tespit etmek için kullanılabilir. Örneğin, her bir kareden uzamsal özellikler çıkarmak ve bunları normal kareleri yeniden yapılandırmak için bir LSTM ağına beslemek için bir evrişimli otomatik kodlayıcı kullanılabilir. Yeniden yapılandırma hatası daha sonra bir anormallik göstergesi olarak kullanılabilir. Alternatif olarak hem ileri hem de geri zamansal bağımlılıkları yakalamak ve çıkarılan özelliklere dayanarak her kareyi normal veya anormal olarak sınıflandırmak için çift yönlü bir LSTM (BD-LSTM) kullanılabilir (Ullah vd. 2021). Diğer bir yaklaşım ise videolardaki hem görünüm hem de hareket bilgilerini modellemek ve anormal bölgeleri vurgulayan bir belirginlik haritası oluşturmak için çift akışlı bir evrişimli LSTM (ConvLSTM) kullanmaktır (L. Wang vd. 2022). Bunlar, LSTM'nin VGAT için kullanılabileceği yollardan bazılarıdır.

Bu yöntemlerin ortak avantajı, zayıf gözetimli veya gözetimsiz olarak çalışabilmesi ve etiketlenmiş veri gerektirmemesidir. Ancak, bu yöntemlerin de bazı dezavantajları da vardır. Örneğin, evrişimli otomatik kodlayıcılar yeniden yapılandırma hatasını tüm kareye uyguladığı için yerel anormallikleri tespit etmekte zorlanabilirler. Çift yönlü LSTM'ler ise sadece kare düzeyinde sınıflandırma yapabildikleri için anormal bölgeleri belirtmezler. Çift akışlı ConvLSTM'ler ise hem görünüm hem de hareket bilgisini modelleyebilmelerine rağmen, belirginlik haritasının kalitesi ve hassasiyeti iyileştirilebilir (Ma vd. 2021).

3.2.4.3 Görüntü dönüştürücü

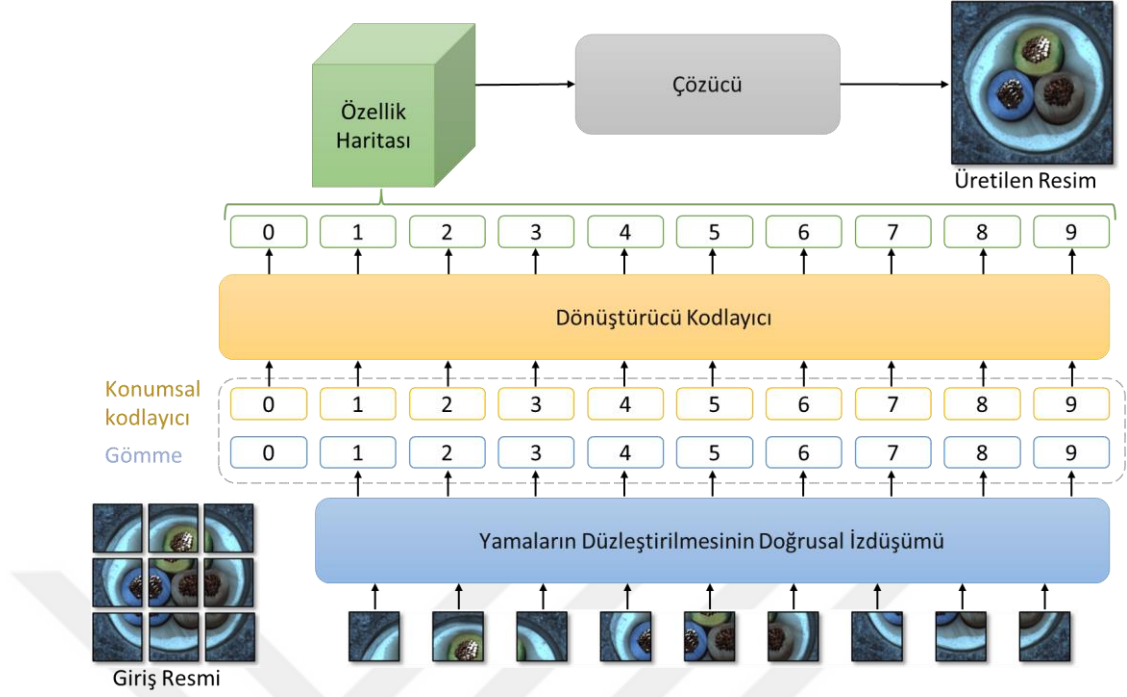
Görüntü Dönüştürücü (ViT), video verilerinden uzamsal-zamansal özellikleri çıkarmak ve bu özellikler arasındaki zaman ilişkilerini modellemek amacıyla CNN'ler ile dönüştürücüleri birleştiren bir yöntemdir (X. Feng vd. 2021; J. Lee vd. 2022; H. Yuan vd. 2021). Bu yöntem, video verilerindeki uzun dönemli bağımlılıkları etkin bir şekilde ele alabilir ve özellikle anormallikleri tespit etmek için özel bir değere sahiptir.

ViT yöntemi, video görüntülerindeki anormal davranışları veya olayları daha kolay tespit etmeyi amaçlayan bir yöntem olarak, video görüntülerini farklı görsel formlara

dönüştürme yaklaşımını benimser. Bu yöntem, video görüntülerini sabit boyutlu yamalara böler ve her bir yamayı dizi veriye dönüştürür. Daha sonra bu veriler “dönüştürücü (transformer)” adı verilen derin öğrenme modeline sunulur. Dönüştürücü modeli, yamalar arasındaki ilişkileri ve bağlamları öğrenerek video görüntülerinin anlamsal içeriğini çıkarabilir. Dönüştürücü modelinin çıktısı, video görüntülerinin gelecekteki veya geçmişteki karelerini tahmin etmek amacıyla kullanılabilir. Tahmin edilen kareler ile gerçek kareler arasındaki fark, anormallik skoru olarak hesaplanır. Anormallik skoru yüksek olan kareler, anormal olayları veya davranışları içermeye olası olduğu daha yüksek olan karelerdir.

Şekil 3.11’de VGAT için Y. Lee ve Kang (2022) tarafından önerilen görüntü yamaları arasındaki küresel ilişkiyi öğrenen bir ViT tabanlı kodlayıcı-çözücü mimarisi gösterilmiştir. Burada, öncelikle bir görüntü küçük yamalara bölünür ve her yama, bir doğrusal izdüşümü katmanı tarafından bir belirteç gömme boyutuna dönüştürülür. Bu gömülme vektörleri, dönüştürücü kodlayıcısına girdi olarak verilir. Kodlayıcı, yamalar arasındaki ilişkileri öğrenmek için çoklu kendinden dikkat katmanları kullanır. Kodlayıcıdan gelen gömülme vektörleri, bir öznitelik haritası oluşturmak için kullanılır. Bu öznitelik haritası, yamaların yerleşim bilgisini korur. Öznitelik haritası, dönüştürücü çözücüsüne girdi olarak verilir. Çözücü, anormal bölgeleri tespit etmek için çoklu kendinden dikkat katmanları ve bir softmax katmanı kullanır. Çözücünden gelen çıktı, görüntünün anormal bölgelerini işaretleyen bir maske olarak yorumlanır.

ViT yöntemi, VGAT için bazı avantajlara sahiptir. İlk olarak ViT, görüntülerin uzamsal ve zamansal özelliklerini aynı anda modelleyebildiğinden video görüntülerinin dinamik yapısını daha etkili bir şekilde yakalayabilir. İkinci olarak, video görüntülerinin genel bağlamını öğrenme yeteneğine sahip olduğundan bütünsel anlamını daha iyi anlamamıza yardımcı olabilir. Üçüncü olarak, video görüntülerinin farklı görsel biçimlerine uyum sağlayabildiğinden çok çeşitli görüntüleri kapsayabilir.



Şekil 3.11 ViT Tabanlı Anomali Tespiti Modeli-AnoViT (Y. Lee ve Kang 2022)

3.2.4.4 Tahmin tabanlı yaklaşımların genel değerlendirmesi

Tahmin tabanlı yaklaşımlar, verinin zaman içindeki evrimini anlamada ve gelecekteki durumları tahmin etmede etkili olabilir. Bu yaklaşımların kullanımı, genellikle yüksek hesaplama gücü gerektirir ve karmaşık eğitim süreçleriyle birlikte gelir. Ayrıca, modelin eğitim verilerine aşırı uyum sağlaması ve genelleştirme yeteneğinin sınırlı olması gibi zorluklar da bulunmaktadır. Veri kalitesi ve çeşitliliği, bu tekniklerin başarısını önemli ölçüde etkileyebilir. Çizelge 3.6’da bu tekniklerin karşılaştırması sunulmuştur.

Çizelge 3.6 Tahmin tabanlı tekniklerin tekniklerin avantaj ve dezavantajları

Teknik	Avantajlar	Dezavantajlar
Tekrarlayan Sinir Ağı	▪ Zaman serisi verilerindeki düzenleri etkili bir şekilde modelleyebilir. ▪ Video verilerindeki zaman bağlantılarını öğrenir.	▪ Uzun süreli bağımlılıkları modellemekte zorlanabilir. ▪ Hesaplama yoğunluğu yüksek olabilir.
Uzun Kısa Süreli Hafıza	▪ Uzun süreli bağımlılıkları başarılı bir şekilde yakalar. ▪ Zamanla değişen verilerin karmaşıklığını yönetebilir.	▪ Karmaşık mimari ve uzun eğitim süreçleri gerektirebilir. ▪ Büyük veri kümeleri ve yüksek hesaplama gücü ihtiyacı.
Görüntü Dönüştürücü	▪ Görüntü verilerindeki karmaşık desenleri öğrenmede etkilidir. ▪ Geniş çaplı ve çeşitli görsel veri kümelerinde iyi performans gösterir.	▪ Büyük miktarda eğitim verisi gerektirir. ▪ Modelin anlaşılması ve yorumlanması zor olabilir.

3.2.5 Diğer yaklaşımlar

İstatistik, sınıflandırma, yeniden yapılandırma ve tahmin tabanlı tekniklerin yanı sıra farklı yaklaşım ve yöntemler üzerinde de çalışmalar bulunmaktadır.

3.2.5.1 Bulanık mantık

Bulanık mantık, belirsizlik, kesin olmayan durumlar ve kısmi doğruluk içeren gerçek dünya problemlerini modellemek için tasarlanmış bir sistemdir. Bu sistem, belirsizliği ve karmaşıklığı ele almak için bulanık kümeler ve bulanık kurallar kullanır. Geleneksel ikili mantık sistemlerinden farklı olarak, bulanık mantık, her şeyin sadece “doğru” veya “yanlış” değil, aynı zamanda “kısmen doğru” veya “kısmen yanlış” olabileceğini kabul eder. Bu, gerçek dünya problemlerinin daha gerçekçi ve esnek bir şekilde ele alınmasını sağlar.

VGAT bağlamında, bulanık mantık, video karelerini veya bölgelerini bulanık kümeler olarak ele alır ve bulanık kurallar kullanarak anormallik skorları hesaplar. Bu yaklaşım,

görüntülerdeki belirsizlik ve karmaşıklık, özellikle karmaşık ve dinamik ortamlarda, etkin bir şekilde işler. Örneğin, karmaşık trafik akışı gibi dinamik ortamlardaki anormallikler, bulanık mantık kullanılarak başarılı bir şekilde tespit edilebilir (Y. Li vd. 2018).

Bulanık mantığın teknik avantajları arasında, gerçek zamanlı video analizi sırasında ortaya çıkan belirsizlik ve gürültüyü yönetme yeteneği yer alır. Bu yöntem, özellikle video görüntülerindeki gri alanları ve geçiş durumlarını modellemek için uygundur. Ayrıca, bulanık mantık, anormallik tespiti için çeşitli senaryolarda esnek ve dinamik kurallar sağlayarak farklı tipte anormallikleri tespit etmek ve video görüntülerindeki değişken davranışları doğru bir şekilde sınıflandırmak için özelleştirilebilir.

Bulanık mantığın uygulanmasında karşılaşılan zorluklar arasında, bulanık kuralların ve kümeleme fonksiyonlarının doğru şekilde tasarlanması ve ayarlanması yer alır. Bu süreç, özellikle karmaşık veri kümeleri ve değişken senaryolarda, dikkatli bir analiz ve deneyim gerektirir. Ayrıca, bulanık mantık modelinin doğruluğu, kullanılan kuralların ve kümeleme fonksiyonlarının kalitesine bağlıdır. Bu nedenle, modelin geliştirilmesi ve test edilmesi süreçleri, veri kümesinin doğasına ve hedeflenen uygulama senaryosuna uygun olarak dikkatle gerçekleştirilmelidir.

3.2.5.2 Kısa kodlama ve sözlük öğrenme

Kısa kodlama ve sözlük öğrenme, video verilerini işlerken yüksek boyutlu verilerin daha düşük boyutlu ve anlamlı temsillerini oluşturmak için kullanılır.

Kısa kodlama, video karelerini veya bölgelerini daha düşük boyutlu vektörler olarak kodlamak için bir kodlama fonksiyonu öğrenmeyi içerir. Bu yöntem, video verilerinin temel özelliklerini ve yapısal bileşenlerini özetleyen yoğun (dense) vektörler oluşturmayı hedefler. Bu teknik, video karelerinin veya bölgelerinin özelliklerini temsil eden kodlama vektörlerini üretir ve bu vektörler, video verilerinin özgün içeriğini daha

az veri ile temsil etmeyi sağlar. Kısa kodlama, özellikle video verilerinin yoğun olduğu durumlarda anormallik tespitinde etkili bir araç olarak ortaya çıkar.

Sözlük öğrenme, video karelerini veya bölgelerini, öğrenilmiş bir sözlük matrisinin doğrusal kombinasyonu olarak temsil etmeyi içerir. Bu teknik, video verilerinin çeşitli bileşenlerini temsil eden sözlük elemanlarını oluşturarak, video karelerinin bu elemanlar üzerinden ifade edilmesini sağlar. Anormallikler, sözlük elemanlarına göre yeniden oluşturma hatası veya kısalık ölçüsüne dayanarak tespit edilir. Bu yaklaşım, video verilerinin yeniden yapılandırılmasında ve anormal davranışların belirlenmesinde kullanılır.

Bu teknik, özellikle yüksek boyutlu video verilerinde anlamlı ve kompakt temsiller oluşturarak, anormallik tespiti sürecini kolaylaştırır. Ancak, bu tekniklerin başarısı, öğrenilen kodlama ve sözlük matrislerinin kalitesine ve veri kümesinin doğasına bağlıdır. Bu nedenle, bu tekniklerin uygulanmasında özenli bir modelleme ve parametre ayarlama süreci gereklidir.

3.2.5.3 Heuristik yöntemler

Anormallik tespitinde sezgisel kararlar almak için Heuristik yöntemler kullanılabilir. Bu yöntemler, uzman görüşüne veya deneyime dayalı olarak belirli kurallar ve yaklaşımlar geliştirmeyi içerir. Özellikle veri kümelerinin sınırlı olduğu veya kesin kuralların oluşturulmasının zor olduğu durumlarda öne çıkarlar. Heuristik yöntemler, uzman bilgisi ve sezgisel yaklaşımların entegrasyonunu sağlar. Örneğin, uzmanların deneyimlerinden elde edilen kurallar kullanılarak anomaliler tanımlanabilir veya veri değişiklikleri belirli bir eşik değerini aştığında anormal durumlar olarak değerlendirilebilir. Ancak heuristiklerin doğruluğu ve genelleme kapasitesi, her durumda dikkatlice değerlendirilmelidir.

Heuristik yöntemlerin uygulanması genellikle alan uzmanlarının detaylı bilgisi ve sezgisel anlayışına dayanır. Özellikle belirsiz ve karmaşık veri kümelerinde,

anormallikleri tanımlamak için önemli bir role sahiptir. Örneğin, trafik izleme sistemlerinde uzmanlar, belirli trafik desenlerinin beklenmeyen değişimlerini anormal olarak işaretleyebilir. Aynı şekilde, perakende mağazalarında müşteri hareketlerinin beklenmedik desenleri, potansiyel anormalliklerin bir işareti olarak kullanılabilir.

3.2.6 Bölüm değerlendirmesi

Bu bölümde, algoritmaların türlerine göre sınıflandırılan yaklaşımların detayları, avantajları, dezavantajları ve katkıları incelenmiştir.

İstatistik tabanlı yaklaşımlar, veri dağılımı ve özelliklerini analiz ederek anormallikleri tespit etme yeteneğine sahiptir. Verilerin istatistiksel özelliklerini kullanarak, anormal durumları standart davranışlardan ayırmayı hedefler ve genellikle basit ve hesaplama açısından verimli olmalarının avantajı vardır. Sınıflandırma tabanlı yaklaşımlar, öğrenilen örüntüleri ve sınıfları tanıyarak yüksek doğruluk oranları sunar. Etiketli veri kullanarak modeli eğitir ve daha sonra bu modeli yeni verilerin sınıflandırılmasında kullanır. Bu yaklaşımı kullanan yöntemlerin gücü, genellikle belirli tipteki anormallikleri etkili bir şekilde tespit edebilmeleridir. Yeniden yapılandırma tabanlı yaklaşımlar, derin öğrenme tabanlı yaklaşımlarla birleştiğinde verinin karmaşık yapısını öğrenir ve anormallikleri bu yapılar üzerinden belirler. Video verilerinin yeniden yapılandırılması sırasında oluşan hataları kullanarak anormallikleri tespit eder. Tahmin tabanlı yaklaşımlar, özellikle zaman serisi verileri ve görüntü verileri üzerinde çalışarak, gelecekteki durumları tahmin edebilir ve bu tahminlere dayanarak anormallikleri önceden belirleyebilir. Bu yaklaşımı kullanan yöntemler, özellikle video akışlarındaki zaman içindeki değişimleri takip ederek anormallikleri tespit eder. Diğer yaklaşımlar ise, özel ihtiyaçlara yönelik özelleştirilmiş ve yenilikçi çözümler sunar ve bu alandaki yöntemlerin sürekli gelişimini temsil eder.

Her yaklaşım, kendi içinde avantajlar ve dezavantajlar barındırır. Bazıları karmaşık veri yapılarını daha iyi işlerken, diğerleri belirli veri türleri veya senaryolar için daha uygundur. Çizelge 3.7’de bu yaklaşımların genel bir karşılaştırması sunulmuştur.

Çizelge 3.7 Algoritma türlerine dayalı yaklaşımların avantaj ve dezavantajları

Yaklaşımlar	Avantajları	Dezavantajları
İstatistik Tabanlı Yaklaşımlar	- Veri kümesinin genel dağılımını analiz ederek anormal davranışları tespit edebilir. - Temel istatistiksel özellikleri kullanarak büyük miktarda verinin analizinde etkilidir.	- Aşırı değerlerin varlığı, istatistiksel analizin sonuçlarını etkileyebilir. - Çok değişkenli ve karmaşık senaryolarda istatistiksel yöntemler yetersiz kalabilir.
Sınıflandırma Tabanlı Yaklaşımlar	- Eğitim verilerine iyi adapte olduklarında, bu teknikler genellikle yüksek doğruluk ve hassasiyet sunar. - Farklı senaryolara göre özelleştirilebilirler.	- Kaliteli sonuçlar için büyük ve temsil edici eğitim verilerine ihtiyaç duyar. - Modeller, eğitim verilerine aşırı uyum sağlayabilir ve genel veri kümelerinde kötü performans gösterebilir.
Yeniden Yapılandırma Tabanlı Yaklaşımlar	- Verinin karmaşık yapılarını modelleyebilir ve anormal durumları bu yapılar üzerinden tespit edebilir. - Veriden öznitelik çıkarma yeteneği sayesinde etkili analizler yapabilir.	- Model eğitimi ve veri yeniden yapılandırması genellikle yüksek hesaplama kaynakları gerektirir. - Modelin etkili bir şekilde eğitilmesi zor olabilir.
Tahmin Tabanlı Yaklaşımlar	- Zamanla değişen verilerde etkili bir şekilde çalışabilir. - Anormallikleri erken evrede tespit edebilir.	- Özellikle veride gürültü varsa tahminler bazen yanıltıcı olabilir. - Model parametrelerinin doğru ayarlanması ve eğitimi zorlayıcı olabilir.
Diğer Yaklaşımlar	- Özellikle bulanık mantık, gerçek dünya verilerinin belirsizliği ve karmaşıklığı ile başa çıkma konusunda yüksek uyarlanabilirlik sunar. - Kısa kodlama ve sözlük öğrenme gibi teknikler, verinin karmaşık yapılarını modelleyebilir ve bu yapılardan anlamlı bilgiler çıkarabilir.	- Etkili sonuçlar elde etmek için hassas ayarlamalar ve optimizasyon gerektirebilir. - Bazı durumlarda, bu tekniklerin sonuçlarının tutarlılığı ve güvenilirliği sorun oluşturabilir.

4. ANOMALİ TESPİTİNDE KULLANILAN VERİ KÜMELERİ

Bu bölümde, VGAT alanında yaygın olarak kullanılan ve bu alandaki çalışmalar için kritik öneme sahip olan başlıca veri kümeleri detaylı bir şekilde incelenmiştir. Her bir veri kümesinin özellikleri, içerdiği anormallik türleri, kullanıldığı ortamlar bu bölümde detaylandırılmıştır. Bu veri kümeleri, çeşitli ortamlarda, farklı koşullar altında kaydedilmiş videoları içermekte ve geniş bir yelpazede anormal hareket türlerini kapsamaktadır. Bu veri kümelerinin ayrıntılı analizi, VGAT alanında mevcut durumu anlamak ve bu alandaki araştırmaları ilerletmek için gelecekteki potansiyel yönleri belirlemekte yardımcı olacaktır.

4.1 CUHK Avenue Veri Kümesi

Hong Kong Çin Üniversitesi'nde (CUHK) yapılan bir çalışmanın ürünü olan CUHK Avenue veri kümesi, 2013 yılında yayımlanmıştır. Videoların CUHK kampüsünde çekilmiş olması, gerçek dünya senaryolarını yansıtır ve bu da güvenlik sistemlerinin gerçek yaşam koşullarında ne kadar etkili olduğunu değerlendirmek için önemlidir. Bu veri kümesi, dış ortamda çekilen toplamda 30.652 kareden (15.328 eğitim, 15.324 test) oluşan 47 video klipten meydana gelir (Ren vd. 2021). İçerisinde normal olayları içeren 16 video ve ters yönde yürüme, koşma, dans etme gibi anormal olayları barındıran 21 video bulunur (B. Wang ve Yang 2022).

Dış ortamda çekilen videoların çözünürlüğü 640x360 piksel olup, saniyede 25 kare (FPS) hızındadır (Nayak vd. 2021). Her video kaydı, kare ve piksel düzeyinde gerçek etiket değerlerine sahiptir. Veri kümesinde anormal hareketleri içeren örnek kareler Şekil 4.1'de gösterilmiştir.



Şekil 4.1 CUHK Avenue veri kümesindeki anormal kare örnekleri (Lu vd. 2013)

4.2 Subway Veri Kümesi

2008 yılında yayınlanan Subway veri kümesi (Adam vd. 2008), Subway Entrance ve Subway Exit olmak üzere iki alt veri kümesinden oluşmaktadır. Bu veri kümesi, metro istasyonundaki normal ve anormal davranışları kapsamlı bir şekilde inceler. Özellikle, yanlış yönde hareket etme, turnikelerden atlamak gibi belirgin anormal eylemleri içerir, bu da VGAT araştırmaları için zengin ve gerçekçi bir veri kaynağı sunar.

Veri kümesi, her bir alt kümede çeşitli anormal olayları içeren video kliplerle zenginleştirilmiştir. Şekil 4.2'de gösterilen örnekler, anormal hareketlerin ve davranışların çeşitliliğini yansıtır. Videolar 25 FPS hızında ve 512×384 piksel çözünürlüğünde kaydedilmiştir. Entrance veri kümesi, toplamda 144.250 kare ve 66 farklı anormal olayı içerirken, Exit veri kümesi ise 64.901 kare ve 19 anormal olayı barındırır. Her iki alt kümeye ait videolar, kare bazında detaylı etiketlemeler içerir. Bu etiketlemeler, anormal davranışların doğru ve hassas bir şekilde tespit edilmesi için temel teşkil eder.

Sabokrou ve arkadaşlarının çalışmasına (2018) göre, Subway veri kümesi, anormal olayların sayısı ve bu olayların öngörülebilir uzamsal konumları bakımından belirli

kısıtlamalara sahiptir. Bu kısıtlamalar, anormal hareketlerin ve davranışların çeşitliliği ve karmaşıklığı açısından veri kümesinin potansiyel kullanımını sınırlar. Bu durum, özellikle daha geniş ve karmaşık senaryoları modellemeye çalışan VGAT algoritmaları için önemli bir zorluk oluşturur. Bu bağlamda, anormal olayların sayısı ve yerleşimleri, gerçek dünya senaryolarındaki değişken ve beklenmedik durumları yansıtmakta yeterli olmayabilir.



Şekil 4.2 Subway veri kümesindeki anormal kare örnekleri (Adam vd. 2008). İlk satır Subway Entrance veri kümesinden, ikinci satır Subway Exit alt veri kümesinden alınmıştır.

4.3 UCF-Crime Veri Kümesi

UCF-Crime veri kümesi (Sultani vd. 2018), soygun, kavga, ateş etme, işyeri hırsızlığı, kötüye kullanım, patlama, kaza, tutuklama, hırsızlık, kundaklama, vandalizm, çalma ve saldırı gibi çeşitli suç ve güvenlik ihlallerini kapsayan toplam 1.900 video içerir. Her kategoriden 950 normal ve 950 anormal video olmak üzere dengeli bir dağılıma sahiptir. Çizelge 4.1, UCF-Crime veri kümesindeki kategoriye göre video sayısını göstermektedir.

Çizelge 4.1 UCF-Crime ve HR-Crime veri kümelerindeki anormal olay kategorileri ve video sayıları

Kategori	UCF-Crime	HR-Crime
Taciz	50	38
Tutuklama	50	42
Kundakçılık	50	48
Saldırı	50	47
Hırsızlık	100	96
Patlama	50	26
Kavga	50	39
Trafik kazası	150	68
Soygun	150	145
Ateş etme	50	46
İş yeri hırsızlığı	50	50
Çalmak	100	98
Vandalizm	50	46
Normal	950	782
Toplam	1.900	1.571

Videolar, YouTube ve LiveLeak gibi çevrimiçi platformlardan toplanmış gerçek dünya görüntülerinden oluşmaktadır. Normal ve anormal karelerin bazı örnekleri Şekil 4.3'te gösterilmektedir.



Şekil 4.3 UCF-Crime veri kümesinden (a) normal ve (b) anormal kare örnekleri (Sultani vd. 2018)

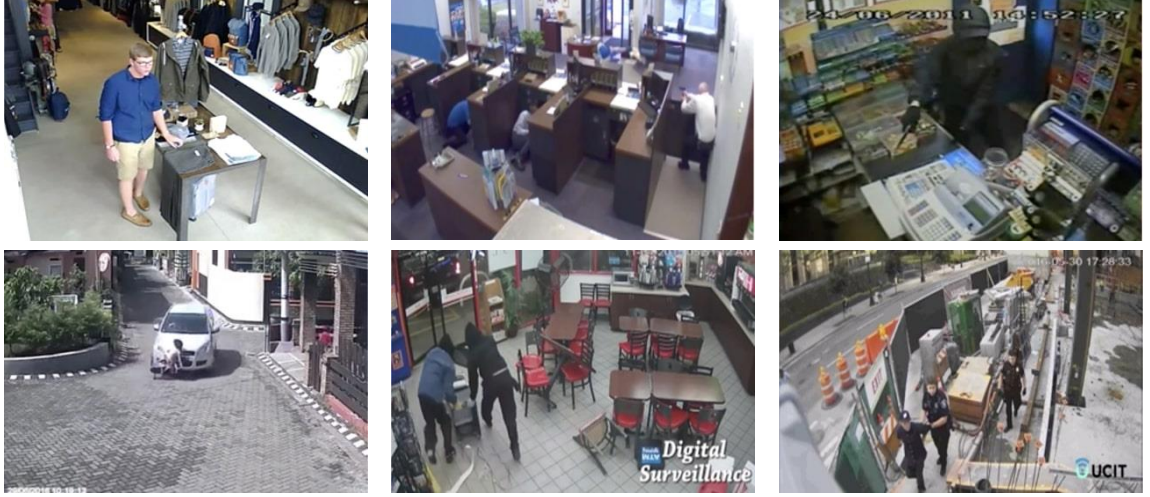
UCF-Crime veri kümesinin önemi, gerçek dünya koşullarını yansıtan ve kamu güvenliği ile ilgili çeşitli senaryoları içeren zengin içeriğinden kaynaklanmaktadır. Bununla birlikte, gerçek dünya görüntülerinin kullanılması, bazı videoların kalitesiz olmasına ve görsel gürültü içermesine neden olabilmektedir. Bu durum, VGAT algoritmalarının geliştirilmesi ve değerlendirilmesi sırasında önemli bir engel olarak karşımıza çıkmaktadır. Ayrıca, bazı videoların birden fazla kategoriye ait anormal olayları içermesi, bu olayların doğru sınıflandırılması ve analiz edilmesi açısından ek zorluklar sunmaktadır. Bu veri kümesi, algoritmaların gerçek dünya senaryolarında nasıl performans gösterdiğini test etmek için iyi bir kaynaktır.

4.4 HR-Crime Veri Kümesi

HR-Crime veri kümesi (Boekhoudt vd. 2021), daha büyük bir veri kümesi olan UCF-Crime veri kümesinin özel bir alt kümesi olarak geliştirilmiş olup, insanla ilgili anormallikler içeren videoları kapsar. HR-Crime, 13 farklı suç kategorisini içeren videolardan oluşur ve her kategori 100 ile 200 arasında değişen video sayısına sahiptir.

HR-Crime veri kümesi toplamda 1.571 video içerir; bunların 782 tanesi normal ve 789 tanesi anormal olayları temsil eder. Videoların uzunluğu, 2 ile 15 dakika arasında değişir, bu da algoritmaların uzun süreli ve karmaşık senaryolarda nasıl performans gösterdiğini değerlendirmek için ideal bir ortam sunar.

Çizelge 4.1’de veri kümesindeki videoların kategori bazında dağılımı gösterilirken, Şekil 4.4’te ise veri kümesinden alınan normal ve anormal kare örnekleri sunulmaktadır. HR-Crime veri kümesinin, VGAT alanında önemli bir katkı sağladığı ve algoritmaların çeşitli anormallik türlerini tespit etme yeteneklerini geliştirmelerine yardımcı olduğu söylenebilir.



Şekil 4.4 HR-Crime veri kümesinden anormal kare örnekleri (Boekhoudt vd. 2021)

4.5 UCSD Veri Kümesi

UCSD veri kümesi (Mahadevan vd. 2010), VGAT çalışmaları için zengin ve çeşitli anormal olayları içeren önemli bir kaynaktır. 2010 yılında yayınlanan bu veri kümesi, özellikle yaya yollarında gerçekleşen ve yayaların tipik davranış kalıplarından sapmaları içeren anormal olaylara odaklanır. Bu anormallikler arasında yaya yollarını kullanan bisikletliler, patenciler, küçük arabalar, tekerlekli sandalyeler ve diğer araçlar bulunmaktadır.

Bu veri kümesinin önemi, gerçek dünya senaryolarından alınan ve sahnelenmeyen doğal olayları içermesinden kaynaklanır. Anormal olayların doğal olarak gerçekleşmesi, VGAT algoritmalarının gerçekçi ve zorlu senaryolarda nasıl performans gösterdiğini değerlendirmek için elverişli bir ortam sunar. UCSD veri kümesi PED1 ve PED2 adı altında iki ana alt kümeye ayrılmıştır. PED1 alt veri kümesinde 34 eğitim videosu ve 36 test videosu bulunurken, PED 2 alt veri kümesinde 16 eğitim videosu ve 12 test videosu bulunmaktadır. Anormal karelerin örnekleri Şekil 4.5'te gösterilmiştir.



Şekil 4.5 UCSD veri kümesinden (a) Ped1 alt kümesinden ve (b) Ped2 alt kümesinden anormal kare örnekleri (Mahadevan vd. 2010)

Gerçek etiket değerlerinin kare seviyesinde verilmesinin yanı sıra PED1 için 10 klip ve PED2 için 12 klip için piksel düzeyinde gerçek değer etiket sağlanmıştır. Bu maskeler anomalileri içeren bölgeleri tanımlamaktadır.

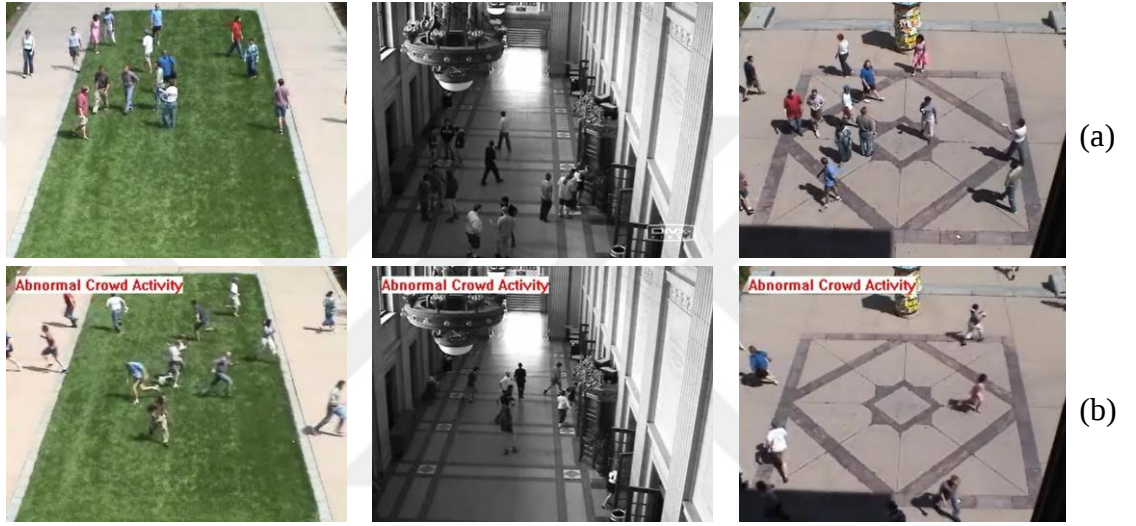
4.6 UMN Veri Kümesi

UMN veri kümesi (Mehran vd. 2009), Minnesota Üniversitesinin iç ve dış ortam sahnelerinde normal ve anormal kalabalık davranışlarını içeren videolardan oluşmaktadır (Direkoglu vd. 2017; Gu vd. 2014). Bu veri kümesi, iç ve dış ortam sahnelerinde kaydedilmiş, normal ve anormal kalabalık davranışlarını içeren videoları kapsar. Özellikle, panik halindeki kalabalıkların anormal hareketlerini gösteren 11 kısa videodan oluşan bir koleksiyon sunmaktadır. Bu videolar, iki açık hava ve bir kapalı ortamda kaydedilmiş olup, üç farklı sahneyi temsil eder.

Her video, normal yürüyüşü takiben ani bir kaçışla sonlanır ve bu durum, panik halindeki kalabalıkların dinamiklerini gözlemlemek için benzersiz bir fırsat sunar. UMN veri kümesi, özellikle kalabalık alanlardaki potansiyel güvenlik risklerini tanımlamak ve

bu tür durumlara yanıt vermek için tasarlanmış gözetim sistemlerinin geliştirilmesinde önemli bir rol oynar.

Bu veri kümesindeki anormal davranış örnekleri, kalabalık ortamlarda anormal davranış tespiti algoritmalarının performansını değerlendirmek ve geliştirmek için kullanılabilir. Anormal davranışların yalnızca kare düzeyinde işaretlendiği kare örnekleri Şekil 4.6'da gösterilmiştir.



Şekil 4.6 UMN veri kümesinden (a) normal ve (b) anormal kare örnekleri (Mehran vd. 2010)

4.7 ShanghaiTech Veri Kümesi

ShanghaiTech veri kümesi (Luo vd. 2017a), 330 eğitim ve 107 test videosu olmak üzere toplam 437 video içerir. Videolar, 13 farklı sahnede çekilmiş olup, bu sahneler karmaşık aydınlatma koşulları ve çeşitli kamera açılarına sahip ortamları içerir. Bu özellikler, VGAT algoritmalarının farklı koşullar altında nasıl performans gösterdiğini anlamak açısından önemlidir.

ShanghaiTech veri kümesi, 480×856 piksel çözünürlükte sunan toplamda 317.398 karenin, 274.515 tanesi eğitim ve 42.883 tanesi test karesinden oluşmaktadır. Veri kümesindeki anormal olaylar, kare ve piksel düzeyinde detaylı olarak işaretlenmiş, bu

da algoritmaların hem genel anormallik tespiti hem de anormalliklerin yerelleştirilmesi açısından değerlendirilmesine olanak tanır. Anormal kare örnekleri Şekil 4.7'de gösterilmektedir.



Şekil 4.7 ShanghaiTech veri kümesinden anormal kare örnekleri (Luo vd. 2017a)

4.8 Veri Kümelerinin Dağılımları

VGAT çalışmalarında kullanılan veri kümelerinin dağılımları ve özellikleri, bu alandaki araştırmaların doğruluğu ve etkililiği açısından önem taşımaktadır. Çizelge 4.1'de listelenen anormallik türleri, VGAT çalışmalarının çoğunlukla insan kaynaklı olaylara odaklandığını göstermektedir. Buna karşın, cansız nesnelerden kaynaklanan anormalliklerin oranı görece daha düşük kalmaktadır. Bu durum, VGAT algoritma ve yaklaşımlarının insan etkileşimli senaryolara daha fazla yöneltilmiş olduğunu ve cansız nesne kaynaklı anormallikleri tespit etmek için daha fazla gelişmeye ihtiyaç duyulduğunu işaret etmektedir.

Çizelge 4.2 ve 4.3'te yer alan veri kümelerinin genel özellikleri incelendiğinde, çoğunun güncel olmadığı ve düşük çözünürlüğe sahip olduğu görülmektedir. Bu durum, modern VGAT sistemlerinin yüksek çözünürlüklü ve karmaşık sahnelerde nasıl performans gösterdiğini değerlendirmek için zorluklar oluşturabilir. Ayrıca, bu veri kümelerinin

çoğu, anormallikleri yalnızca kare düzeyinde temsil etmektedir. Bu, piksel düzeyinde anormallik tespitine ve daha hassas yerleştirme tekniklerine olan ihtiyacı vurgulamaktadır.

Çizelge 4.2 VGAT veri kümelerinin özellikleri

Veri Kümeleri	Yıl	Etiketleme	Çözünürlük	FPS	Kayıt Ortamı
Avenue	2013	kare/ piksel	640 × 360	25	kapalı
UCF-Crime	2018	kare / klip	320 × 240	30	açık/ kapalı
HR Crime	2021	kare / klip	320 × 240	30	açık / kapalı
UMN	2009	kare	320 × 240	30	açık / kapalı
ShangaiTech	2016	kare / piksel	856 × 480	-	kapalı
Subway Entrance	2008	kare	512 × 384	25	açık
Subway Exit					
UCSD Ped1	2010	kare / piksel	238 × 158	-	kapalı
UCSD Ped2			360 × 240		

Çizelge 4.3 VGAT veri kümelerinin kare özellikleri ve sayıları

Veri Kümeleri	Normal Kare Sayısı	Anormal Kare Sayısı	Eğitim Kare Sayısı	Test Kare Sayısı	Toplam Kare Sayısı
Avenue	26.832	3.820	15.328	15.324	30.652
UCF-Crime	-	-	12.631.211	1.110.182	13.741.393
HR-Crime	485.227	335.378	-	-	-
UMN	6.165	1.576	-	-	7.740
ShangaiTech	300.308	17.090	274.515	42.883	317.398
Subway Entrance	132.138	12.112	76.453	67.797	144.250
Subway Exit	60.410	4.491	22.500	42.401	64.901
UCSD Ped1	9.995	4.005	6.800	7.200	14.000
UCSD Ped2	2.924	1.636	2.550	2.010	4.560

Veri kümelerinin yapısı ve içeriği incelendiğinde, kayıt ortamı ve gerçek etiket değerleri temel alınarak veri kümeleri kategorize edilebilir.

4.8.1 Kayıt ortamına göre dağılım

Veri kümelerinin yapısı, kayıt ortamına göre önemli farklılıklar göstermektedir. Bu durum, VGAT algoritmalarının ve sistemlerinin farklı ortamlarda nasıl performans

gösterdiğinin anlaşılması için önemlidir. Çizelge 4.4'te yer alan dağılım incelendiğinde, veri kümelerinin açık ve/veya kapalı ortamlarda kayıt altına alındığı görülmektedir.

Çizelge 4.4 Veri kümelerinin kayıt ortamlarına göre dağılımı

Açık Ortam	Kapalı Ortam
UCF-Crime	Avenue
UMN	UCF-Crime
Subway Entrance	UMN
Subway Exit	ShanghaiTech
	UCSD Ped1
	UCSD Ped2

Tren istasyonları, parklar, yaya geçitleri ve sokaklar gibi alanlar açık ortamlara örnek olarak verilebilir. Bu ortamlar, güneş ışığı, bulutluluk, yağmur ve kar gibi değişken hava koşullarından etkilenme eğilimindedir. Bu durum, ışık koşullarındaki değişimlerin ve hava durumunun anormallik tespit algoritmalarının performansını nasıl etkilediğini değerlendirmek için önemlidir.

Diğer yandan, alışveriş merkezleri, okullar, fabrikalar ve metro istasyonları gibi alanlar kapalı ortamlara örnek olarak verilebilir. Bu ortamlar, açık hava ortamlarına göre ışık değişimlerinden ve lens bozulmalarından daha az etkilenme eğilimindedir. Kapalı ortam veri kümelerinin analizi, VGAT algoritmalarının sabit aydınlatma koşulları ve daha kontrollü ortamlarda nasıl performans gösterdiğinin anlaşılmasında yardımcı olabilir.

4.8.2 Renk uzayına göre dağılım

Renk uzayı, görüntülerdeki piksellerin renklerinin nasıl temsil edildiğiyle ilgilidir ve VGAT algoritmalarının veri kümelerini işleme şeklini doğrudan etkiler. Bu nedenle, renk uzayının seçimi, VGAT algoritmalarının doğruluğu ve performansı üzerinde önemli bir etki yapar.

Renk uzaylarının çeşitleri arasında en yaygın olanları gri tonlamalı ve RGB renk uzaylarıdır. Gri tonlamalı görüntüler, her pikselin sadece parlaklık bilgisini içerir ve renk bilgisi taşımaz. Bu basitlik, daha az hesaplama kaynağı gerektirir ve bu yüzden bazı VGAT uygulamalarında tercih edilir. Örneğin, UCSD ve Subway veri kümeleri gri tonlamalı olarak temsil edilmektedir. Bu tür veri kümeleri, daha az boyuta sahip olduğu için hızlı işleme avantajına sahiptir. Ancak, renk bilgisinin eksikliği, bazı durumlarda anormallik tespitini zorlaştırabilir.

Diğer taraftan, RGB renk uzayı, her piksel için kırmızı, yeşil ve mavi kanalların yoğunluklarını içerir. RGB renk uzayında temsil edilen görüntüler, renk bilgisi taşıdığı için daha zengin bir veri kümesi sunar. Avenue, UCF-Crime, ShanghaiTech ve UMN veri kümeleri bu kategoride yer alır. RGB renk uzayı, daha fazla bilgi içerdiği için anormallikleri tespit etmede gri tonlamalı görüntülere göre daha etkili olabilir. Özellikle renkli nesnelerin veya arka planların önemli olduğu durumlarda, RGB renk uzayı tercih edilir.

Renk uzayının seçimi, VGAT modelinin eğitim ve test süreçlerinde uyumluluk gerektirir. Eğitim aşamasında kullanılan renk uzayı ile test aşamasında kullanılan renk uzayı arasında uyumsuzluk, modelin performansını olumsuz etkileyebilir. Örneğin, RGB renk uzayında eğitilen bir modelin gri tonlamalı bir veri kümesi üzerinde test edilmesi durumunda, model renk bilgisini kullanamayacağı için anormallik tespiti yaparken zorlanabilir. Çizelge 4.5, incelenen veri kümelerinin renk uzaylarına göre dağılımını sunmakta olup, renk uzaylarının VGAT modellerinin performansı üzerindeki etkilerini göstermektedir.

Çizelge 4.5 Veri kümelerinin renk uzaylarına göre dağılımı

Gri Seviye	RGB Seviye
UCSD Ped1	Avenue
UCSD Ped2	ShanghaiTech
Subway Entrance	UMN
Subway Exit	UCF-Crime

4.8.3 Gerçek etiket değerlerine göre dağılım

VGAT için kullanılan veri kümelerinin etiketleme yöntemleri, modelin doğruluğunu ve uygulama alanını doğrudan etkiler. Her etiketleme yöntemi, farklı seviyelerde ayrıntı ve hassasiyet sunarak, VGAT algoritmalarının eğitim ve test süreçlerini etkiler. Bu bölümde, VGAT için kullanılan üç temel etiketleme türü olan klip düzeyinde, kare düzeyinde ve piksel düzeyinde etiketlemenin avantajları ve sınırlamaları ele alınmaktadır.

1. Klip Düzeyinde Etiketleme: En genel etiketleme yöntemidir ve her video klibe bir etiket atanmasını içerir. Bu yöntemin avantajı, etiketleme sürecinin basit ve hızlı olmasıdır. Ancak bu yöntem, video içindeki anormal aktivitelerin yerini ve zamanını belirlemede yetersiz kalabilir. VGAT modelleri için klip düzeyinde etiketleme, genel bir anormallik tespiti yapabilir ancak anormalliklerin hassas yerleştirilmesi için uygun değildir.
2. Kare Düzeyinde Etiketleme: Daha ayrıntılı bir etiketleme yöntemi olup, her video karesine ayrı etiketler atanır. Bu yöntem, anormal olayların zamanlamasını daha iyi belirlemeye yardımcı olur ve modelin anormallik tespitinde daha hassas olmasını sağlar. Ancak kare düzeyinde etiketleme, anormalliklerin yerleştirilmesinde bazı sınırlamalar taşır, özellikle anormal aktivitelerin yalnızca videonun küçük bir bölümünde yer alması durumundadır.
3. Piksel Düzeyinde Etiketleme: En detaylı etiketleme yöntemidir ve her piksele ayrı etiketler atanır. Bu yöntem, anormallik tespitinde en yüksek hassasiyeti sağlar ve modelin, video içindeki anormal aktiviteleri hassas bir şekilde yerleştirmesine olanak tanır. Piksel düzeyinde etiketleme, özellikle nesne tespiti ve segmentasyon gibi görevlerde kullanışlıdır. Ancak, bu yöntem çok fazla hesaplama kaynağı gerektirir ve etiketleme süreci oldukça zaman alıcı olabilir.

Bu tez kapsamında, VGAT çalışmaları için farklı etiketleme teknikleri kullanılarak hazırlanan veri kümeleri incelenmiştir. Etiketleme yöntemlerinin seçimi, modelin genel performansı ve anormallik tespiti başarısını önemli ölçüde etkiler. Yaygın olarak kullanılan etiketleme türleri, avantajları ve dezavantajları Çizelge 4.6’te gösterilmektedir. Özellikle hassas ve kritik uygulamalarda, piksel düzeyinde detaylı etiketleme tercih edilirken, genel tespitler için kare veya klip düzeyinde etiketleme yeterli olabilmektedir. Bu yaklaşımların dikkatli seçimi ve modelin uygulama senaryosuna uygunluğu, anormallik tespitindeki başarıyı belirleyen kritik faktörlerdir. Veri kümelerinin gerçek etiket değerlerine göre dağılımı Çizelge 4.7’de gösterilmektedir.

Çizelge 4.6 Etiketleme türlerinin avantaj ve dezavantajları

Etiketleme Türü	Açıklama	Avantajları	Dezavantajları
Klip düzeyinde etiketleme	Tek bir etiket ile tüm video klipi sınıflandırmadır.	Basit ve hızlı bir süreçtir; veri etiketlemek için az zaman ve çaba gerektirir.	Video içindeki birden fazla olay veya aktivite yetersiz temsil edilebilir; genellemeye dayalı yanlışlara yol açabilir.
Kare düzeyinde etiketleme	Her bir video karesine ayrı etiket atamadır.	Daha kesin ve ayrıntılı sınıflandırma sağlar; video içinde değişen aktiviteleri ve olayları doğru bir şekilde tanımlayabilir.	Çok sayıda kare üzerinde detaylı çalışma gerektirdiğinden zaman alıcıdır; bazı ince anormallikler yine gözden kaçabilir.
Piksel düzeyinde etiketleme	Her piksele ayrı etiket atamadır.	En yüksek seviyede doğruluk ve ayrıntı sağlar; çok spesifik anomalilerin ve değişikliklerin tespitine olanak tanır.	Zaman alıcıdır; büyük veri kümeleri için uygulaması zordur; yüksek hesaplama gücü gerektirir.

Çizelge 4.7 Veri kümelerinin gerçek etiket değerlerine göre dağılımı

Klip Düzeyi	Kare Düzeyi	Piksel Düzeyi
UCF-Crime	Avenue	Avenue
	UCF-Crime	ShanghaiTech
	ShanghaiTech	UCSD Ped1/Ped2
	UMN	
	UCSD Ped1/Ped2	
	Subway Entrance/Exit	

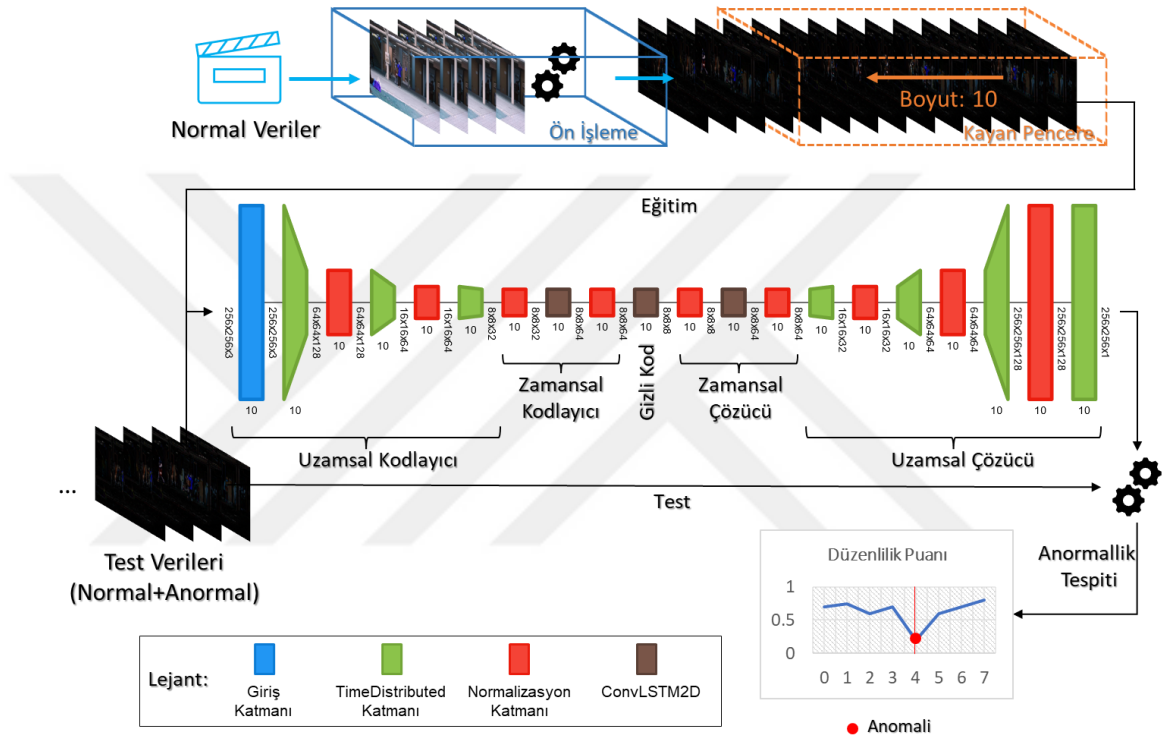
Bu tezde, önerilen metodolojinin etkinliğini test etmek için hem CUHK Avenue hem de UMN veri kümeleri kullanılmıştır. CUHK Avenue veri kümesi, yüksek çözünürlüğe sahip renkli görüntüleri ve gerçek dünya gözetim senaryolarını temsil eden zengin içeriği ile literatürde sıkça referans alınmaktadır. Bu veri kümesinin kullanımı, önerilen metodolojinin gerçek dünya koşullarında nasıl performans gösterdiğini anlamak ve araştırmanın pratik uygulanabilirliğini güçlendirmek açısından önemlidir.

UMN veri kümesi ise, iç ve dış ortam sahnelerinde normal ve anormal kalabalık davranışlarını içeren videolar sunarak, bu alanda yeni bir boyut eklemektedir. UMN veri kümesi, özellikle kalabalık alanlardaki potansiyel güvenlik risklerini tanımlamak ve bu tür durumlara yanıt vermek için tasarlanmış gözetim sistemlerinin geliştirilmesinde önemli bir rol oynar.

Bu iki veri kümesinin detaylı analizi, VGAT alanındaki araştırma trendlerini ve gelecekteki potansiyel yönleri belirlemede yardımcı olurken, bu veri kümeleri üzerinde gerçekleştirilen deneyler, VGAT algoritma ve modellerinin gerçek dünya senaryolarındaki etkinliğini ve güvenilirliğini doğrulamak için kritik bir rol oynamaktadır. Sonuç olarak, bu veri kümeleri, VGAT alanında ilerleme kaydetmek ve etkili çözümler geliştirmek için temel bir kaynak olarak kabul edilmelidir.

5. VGAT METODOLOJİSİNİN GELİŞTİRİLMESİ

Bu bölüm, önerilen VGAT metodolojisinin geliştirme sürecine derinlemesine bir bakış sunmaktadır. Metodoloji, video verilerini analiz eden ve işleyen ön işleme, özellik çıkarımı ve anormallik tespiti aşamalarını kapsar. Şekil 5.1, bu metodolojinin görsel bir özetini göstermektedir.



Şekil 5.1 Önerilen metodolojinin şematik gösterimi

Ön işleme aşaması, video verilerini analiz için hazırlama sürecini içerir ve ardışık kare çıkarma, kare boyutlandırma, arka plan çıkarımı ve gürültü azaltma gibi dört temel adımdan oluşur. Bu adımlar, verinin doğruluğunu artırarak analizin etkinliğini maksimize etmeyi hedefler. Özelliklerin öğrenilmesi aşamasında, evrişimli ağlar, LSTM ve AE gibi yapay sinir ağı yapıları, veriden anlamlı özelliklerin çıkarılması için kullanılır. Bu süreç, verinin temel özelliklerini etkili bir şekilde modellemek ve temsil etmek için kritik öneme sahiptir. Anormallik tespiti aşamasında ise, eğitim sürecinin ardından, modelin test verilerini analiz ederek anormallikleri başarılı bir şekilde tespit etmesine odaklanılır. Bu aşamada anormalliklerin doğru bir şekilde tanımlanmasında,

düzenlilik puanının değerlendirilmesinde Persistence1D ve özgün anormallik tespit yöntemi kullanılır.

Metodolojinin performansını değerlendirmek için, karelerin normal ve anormal olarak sınıflandırılması sonucunda elde edilen doğru pozitif, doğru negatif, hatalı pozitif ve hatalı negatif sonuçların yanı sıra önemli metriklerden biri olan AUC değeri de kullanılarak ölçüm yapılır. Bu değerlendirmeler, metodolojinin etkinliğini ve anormallik tespiti konusunda video gözetim sistemlerinin potansiyelini objektif bir şekilde ortaya koymaktadır.

5.1 Ön İşleme

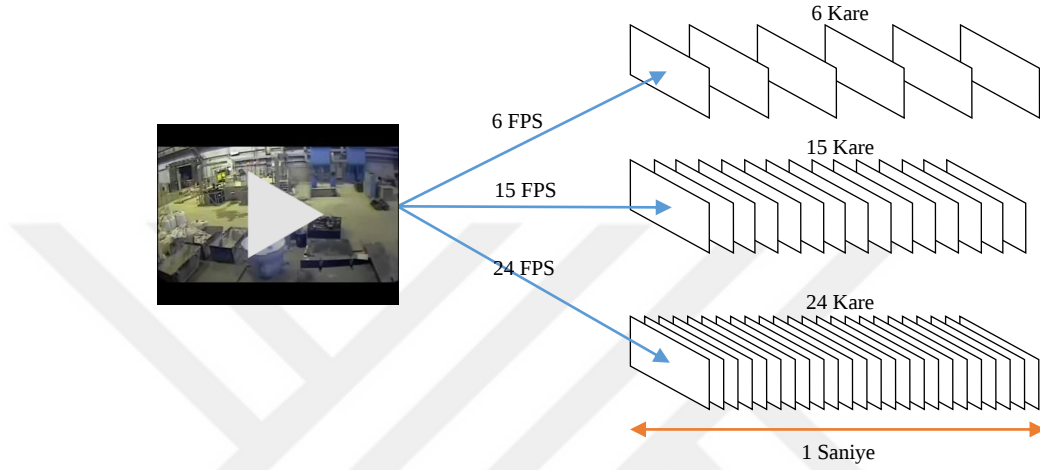
Ön işleme aşaması, verilerin analiz için uygun hale getirilmesi ve istenen sonuçların elde edilmesi için yapılan bir dizi işlemi içerir. Bu aşama, verilerin doğru bir şekilde yorumlanabilmesi, gereksiz bilgilerin temizlenmesi ve analiz sürecinin verimli bir şekilde ilerlemesi için büyük öneme sahiptir.

Arka plan çıkarma ve gürültü azaltma gibi yaygın olarak kullanılan teknikler, bu aşamada özellikle öne çıkmaktadır. Arka plan çıkarma yöntemleri, anomali tespit sürecinin doğruluğunu ve etkinliğini artırmak için gereksiz bilgileri temizler ve ilgili özellikleri belirginleştirir. Gürültü azaltma teknikleri ise istenmeyen ya da rastlantısal dalgalanmaları azaltarak analiz sonuçlarının güvenilirliğini artırır. Ön işleme aşaması dört adımdan oluşur:

5.1.1 Ardışık karelerin çıkarılması

Gözetim videosundaki kareler, nesnelerin konumu, hareketi, rengi gibi önemli bilgileri içeren 2 boyutlu parçalardır. Bu adım, her bir video akışının zaman içindeki durumlarını temsil eden karelerin ayrıştırılmasını sağlar. Temel amacı, video akışının zaman boyutundan bağımsız olarak uzamsal boyutunu ele almak ve bu kareler aracılığıyla nesnelerin değişimini anlamak ve analiz etmektir.

Bu işlemde, karelerin çıkarılma sıklığını belirleyen FPS değeri kullanılır. Bu değerin kullanılması analizin güvenilirliği ve hassasiyeti üzerinde belirleyici bir etkindir. Bu nedenle, belirlenen veri kümesi ile analiz hedeflerine göre en uygun FPS değerinin seçimi, çalışmanın başarısını artırmak için önemlidir. Şekil 5.2’de gözetim videosundan farklı FPS değerlerine uygun olarak karelerin çıkarılma işlemi gösterilmektedir.

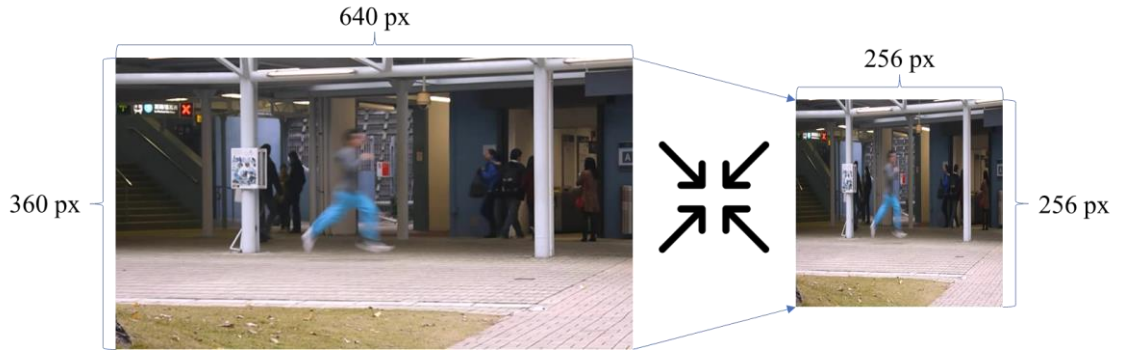


Şekil 5.2 Videodan kare çıkarma işlemi

5.1.2 Karelerin yeniden boyutlandırılması

Gözetim videosundan çıkarılmış her kare, sabit bir boyuta getirilmesi amacıyla yeniden boyutlandırılmaktadır. Bu işlem, tüm karelerin aynı boyuta sahip olmasını sağlayarak algoritmalar veya modeller için tutarlı bir giriş boyutu sunar. Sabit boyutlu karelerin tercih edilmesi, sistem kurulumunu ve yapılandırmasını değişken boyutlu karelere göre daha kolay hale getirir. Bu sayede, sahne boyutlarındaki değişikliklere göre sürekli ayarlama veya yeniden kalibrasyon gerekmeksizin işlem gerçekleştirilebilir.

Yeniden boyutlandırma adımı, her bir karenin orijinal boyutunu dikkate alarak genellikle önceden belirlenmiş bir boyut skalası kullanılarak gerçekleştirilir. Bu işlem sırasında görüntü içeriği mümkün olduğunca bozulmadan yeniden boyutlandırılır. Şekil 5.3’te orijinal boyutu 640x360 piksel olan bir karenin 256x256 piksel boyutuna ölçeklenmesi gösterilmektedir.



Şekil 5.3 Karenin yeniden boyutlandırılması işlemi

Değişken boyutlu karelere sahip sistemlerde, sahne boyutlarındaki değişikliklerle başa çıkmak için sürekli olarak ölçeklendirme, yeniden boyutlandırma veya buna benzer işlemler yapmak gerekebilir. Bu durum, sistemin karmaşıklığının artmasına ve performansının etkilenmesine neden olabilir. Ancak sabit boyutlu kareler kullanıldığında, bu tür sürekli ayarlama işlemlerine ihtiyaç duyulmaz ve sistem daha basit bir şekilde yapılandırılabilir.

5.1.3 Arka planın çıkarılması

Bu adımda arka plan çıkarma yöntemi kullanılarak yeniden boyutlandırılmış her karenin arka planı çıkarılır. Bu işlem, modelin sahnede yer alan hareketli nesnelere odaklanabilmesi için odak dışındaki verilerin sahneden çıkarılmasını sağlar. Bu, modelin eğitiminde hareketli nesneleri daha iyi tespit etmek ve izlemek için büyük bir avantaj sağlayabilir.

Video gözetim sistemlerinde, aydınlatma değişiklikleri, gölgeler ve yansımalar gibi çeşitli faktörler nedeniyle hatalı pozitif sonuçlar sıkça ortaya çıkar. Bu tür faktörlerin arka planla birlikte sahneden temizlenmesi, modelin gerçek anormal olayları ve hareketli nesneleri daha hassas bir şekilde tespit etme yeteneğini artırır. Bu süreç, hatalı pozitiflerin azaltılmasına katkı sağlar, modelin gerçek anormal durumları daha hızlı ve doğru bir şekilde tespit etmesine imkân verir ve sistemlerin genel güvenilirliğini artırır.

Hatalı pozitiflerin azaltılması, modele gerçek anormal durumları daha hızlı ve doğru bir şekilde tespit etme imkânı sunar. Ayrıca, sistemlerin güvenilirliğini artırarak, sahnedeki hareketli nesnelerin takibini daha doğru ve tutarlı hale getirir (Şengönül ve Samet 2021).

Arka plan çıkarma yöntemleri geleneksel ve güncel yaklaşımlar altında gruplandırılabilir. Geleneksel yöntemler, genellikle daha basit ve hesaplama açısından daha az yoğun olan teknikleri içerir. Örneğini “Mean ve median filtreleme” yöntemi, videodaki her piksel için ortalama veya medyan değerleri hesaplayarak arka plan modelini oluşturur. Bu yöntemler, statik arka planlar için etkili olabilir ancak dinamik sahnelerde sınırlı performans gösterir. Bir diğer yöntem olan “GMM” ise, her piksel için birden fazla Gaussian dağılımı kullanarak arka plan ve ön plan nesnelerini ayırt eder. Bu yöntem, arka planın zaman içindeki değişikliklerini daha iyi modelleyebilir.

Diğer yandan güncel yöntemler, daha karmaşık ve genellikle derin öğrenme teknikleri gibi gelişmiş algoritmaları içerir. “Derin öğrenme tabanlı” yöntemler, sinir ağları kullanarak arka planı modelleyen ve karmaşık sahnelerde daha etkili performans gösteren yöntemlerdir. Örneğin “Robust subspace modeli” yöntemi, arka planı düşük seviyeli ve seyrek matrisler olarak modelleyerek ayırt eden yöntemlerdir. Bu teknikler, arka plan ve ön plan nesneleri arasındaki farklılıkları daha etkili bir şekilde ayırt edebilir.

Çalışmamızda bir GMM yöntemi olan MOG2 (Mixture of Gaussians 2) yöntemi tercih edilmiştir. Bu yöntemin seçilmesindeki başlıca nedenler arasında, dinamik arka planlarda etkili performansı, arka plan değişikliklerine karşı esnekliği ve çeşitli aydınlatma koşullarında sağlam sonuçlar üretebilmesi sayılabilir. MOG2'nin her piksel için birden fazla Gaussian dağılımı kullanarak arka plan ve ön plan nesnelerini ayırt edebilme yeteneği, anormal olayların daha doğru ve etkili bir şekilde tespit edilmesine önemli ölçüde katkıda bulunur. MOG2 yöntemi kullanıldığında elde edilen sonuç, hareket eden nesnelerin ve arka planın ayırt edilmesini sağlayan siyah-beyaz bir maske formundadır. Ancak, bu işlem sırasında orijinal görüntünün renkli kısımlarındaki veriyi

de değerlendirmek önemlidir. Bu amaçla, Denklem 5.1’de yer alan aşağıdaki adımlar uygulanmaktadır:

$$\begin{aligned} mask_rgb &= cv2.cvtColor(fgMask, cv2.COLOR_GRAY2RGB) \\ &= cv2.resize(original_frame, (mask_rgb.shape[1], mask_rgb.shape[0])) \\ out_frame &= cv2.bitwise_and(frame_resized, mask_rgb) \end{aligned} \quad (5.1)$$

İlk olarak, MOG2 tarafından oluşturulan siyah-beyaz maske, “*cvtColor*” fonksiyonu kullanılarak RGB renk uzayına dönüştürülür. Bu işlem, maskeyi üç kanallı bir formata getirerek, orijinal renkli görüntü ile uyumlu hale getirir. Daha sonra, “*resize*” fonksiyonu kullanılarak, orijinal görüntü, maskenin boyutlarına uygun olarak yeniden boyutlandırılır. Bu adım, maske ile orijinal görüntünün boyutlarını eşleştirerek, her iki görüntünün üzerinde işlem yapmayı mümkün kılar. Son olarak, “*bitwise_and*” fonksiyonu uygulanarak, yeniden boyutlandırılmış orijinal görüntü ile RGB’ye dönüştürülmüş maske arasında bir bitwise AND işlemi gerçekleştirilir. Bu işlem, maske üzerindeki beyaz (hareket eden) bölgelerdeki orijinal görüntü verilerini korurken, siyah (statik arka plan) bölgelerdeki verileri sıfıra indirger. Böylece, hareket eden nesneler renkli bir şekilde öne çıkarılırken, statik arka plan efektif bir şekilde bastırılır. Bu yöntem, anormal hareketlerin tespitindeki doğruluğu artırarak, gözetim sistemlerinin performansını iyileştirir ve daha zengin bilgi sağlar.

5.1.4 Gürültülerin azaltılması

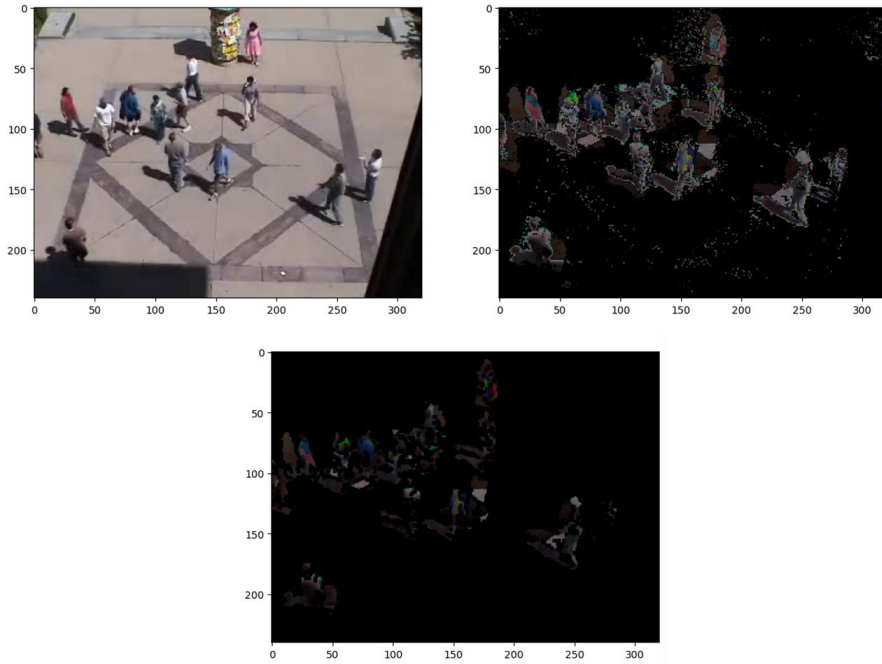
Video kaynağından gelen görüntülerdeki gürültü ve parazitler, genellikle önemli detayları örtebilmekte ve olağandışı olayların doğru bir şekilde tanımlamasını güçleştirmektedir. Bu gürültü ve parazitlerin azaltılması için kullanılan yöntemler, çeşitli gürültü türlerine ve görüntü ihtiyaçlarına göre seçilir. Yaygın olarak kullanılan yöntemler aşağıda sunulmuştur;

1. Mean Filtering: Bu yöntem, her pikselin değerini, o pikselin kendisi dahil olmak üzere, komşu piksellerin aritmetik ortalaması ile değiştirerek düzgünleştirme sağlar. Bu işlem, görüntüdeki gürültüyü azaltır ve yumuşatma etkisi yaratır. Bu işlemin sonucunda görüntüdeki detaylarda bulanıklaşma sorunu ortaya çıkmaktadır.
2. Gaussian Blurring: Piksellerin değerlerini, Gaussian fonksiyonunun belirlediği ağırlıklarla hesaplanan bir ortalama kullanarak düzgünleştirir. Bu süreç, piksel değerlerinin, çevrelerindeki diğer piksellerin değerleriyle uyumlu hale gelmesini sağlayarak görüntüdeki detayları azaltır ve yumuşatma etkisi oluşturur.
3. Median Filtering: Her pikselin değerini, o pikselin ve çevresindeki piksellerin medyan değeriyle değiştirerek gürültüyü azaltır. Bu yöntem, özellikle tuz ve biber gürültüsü gibi aşırı değerlerin etkisini azaltmak için etkilidir ve kenarları koruma özelliğine sahiptir.
4. Bilateral Filtering: Hem uzamsal yakınlık hem de piksel yoğunluğu farklılıklarını göz önünde bulundurarak gelişmiş bir yumuşatma işlemi gerçekleştirir. Bu yöntem, kenar korumalı yumuşatma yaparak görüntünün önemli yapısal özelliklerini korurken gürültüyü azaltır.
5. Wavelet Transforms: Görüntüyü düşük ve yüksek frekans bileşenlerine ayırır. Düşük frekans bileşenleri, görüntünün genel eğilimlerini veya yavaş değişen kısımlarını temsil ederken, yüksek frekans bileşenleri daha hızlı değişen detayları ve gürültüyü temsil eder. Bu yüksek frekanslı bileşenlerde, gürültüyü azaltmak için belirli bir eşik değeri uygulanır. Eşikleme işlemi sonrasında, gürültü içeren kısımlar azaltılır veya ortadan kaldırılır. Son olarak, gürültüsüz hale getirilen veri, ters Wavelet Dönüşümü uygulanarak orijinal haline yakın bir formuna dönüştürülür.
6. Morphological Filtering: Yapısal elementler kullanarak görüntüdeki nesnelerin şeklini ve yapısını iyileştirir. Nesnelerin sınırlarını küçültme, genişletme ve kenar temizleme gibi işlemlerle gürültüyü ve istenmeyen detayları azaltır.

Bu yöntemlerin her biri, belirli türdeki gürültülere ve görüntü işleme gereksinimlerine göre seçilir ve uygulanır. Morfolojik yöntemlerden olan “Erode” işlemi, basit ve etkili bir gürültü azaltma çözümü sunduğundan çalışmamızda tercih edilmiştir. Bu yöntem, özellikle küçük gürültü parçacıklarının ve kenar gürültülerinin temizlenmesinde etkili olup, modelin netlik ve doğruluk açısından daha iyi sonuçlar vermesine yardımcı olur. Yöntemin kullanımı Denklem 5.2’deki kod ile gösterilmiştir. Buradaki “*kernel*” parametresi, aşındırma işleminin şeklini belirlemektedir. Örneğin, bir dikdörtgen yapısını temsil eden bir kernel kullanılabilir. “*iterasyon*” parametresi ise aşındırma işleminin kaç kez tekrarlanacağını belirtir.

$$sonuc_kare = cv2.erode(giris_karesi, kernel, iterasyon) \quad (5.2)$$

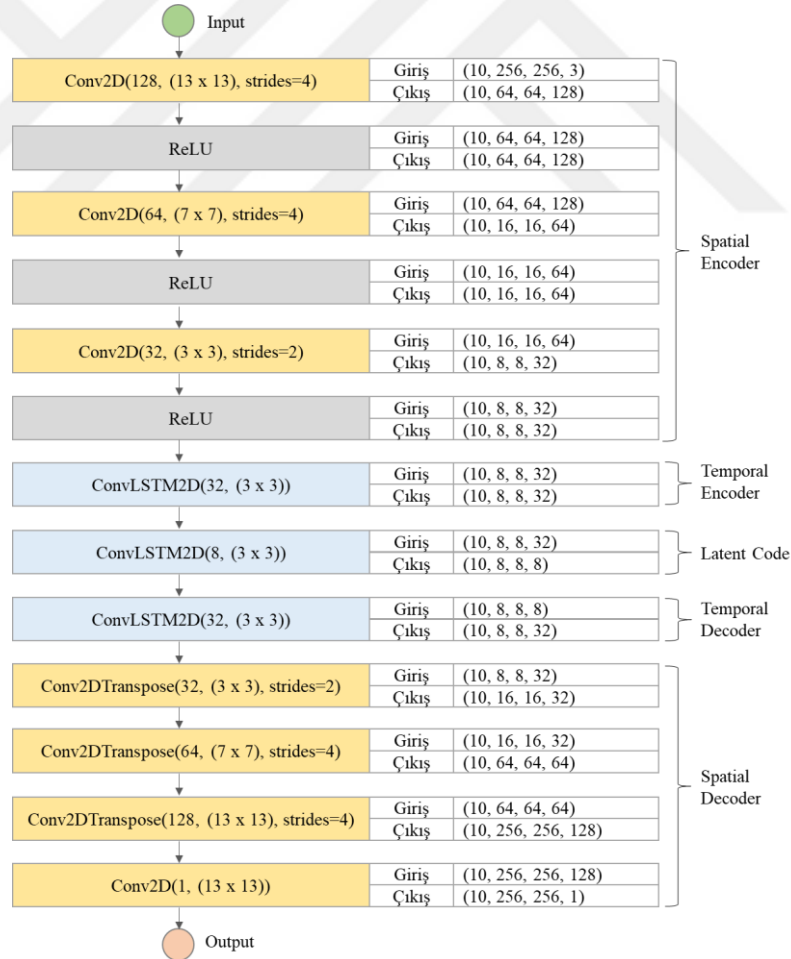
Bu yöntem özellikle düşük ışık koşullarında veya düşük kaliteli video kaynaklarında, görüntünün kalitesini iyileştirir ve modelin daha etkili bir şekilde çalışmasına yardımcı olur. Böylece modelin daha net ve temiz bir görüntü üzerinde çalışmasını sağlayarak anormal olayların daha doğru bir şekilde tanımlayabilir ve hatalı pozitiflerin sayısını azaltır. Şekil 5.4’te gürültü azaltma işlemi uygulanmış örnek bir kare gösterilmiştir.



Şekil 5.4 Gürültü azaltma işlemi sonucunda elde edilen kare

5.2 Özelliklerin Öğrenilmesi

Evrişimli sinir ağları, görsel veriler üzerinde özellik çıkarımı ve görüntü işleme alanlarında yaygın olarak kullanılan güçlü bir yapay sinir ağı türüdür. Bu ağlar, görsel verilere uygulanan evrişim işlemleri aracılığıyla özellikleri belirginleştirir ve kritik bilgileri çıkarmada etkilidir. Öte yandan, LSTM ağları, özellikle ardışık veri dizilerinin işlenmesinde kullanılır. Zaman serisi gibi veri tiplerinde, LSTM uzun vadeli bağlantıları ve bağımlılıkları etkin bir şekilde modelleyebilir. AE ise veri sıkıştırma ve özellik çıkarımında kritik rol oynayan bir diğer yapay sinir ağı türüdür. ConvLSTM-AE modeli, CNN, LSTM ve AE'nin özelliklerini birleştiren, görüntü işleme, dil işleme ve zaman serisi analizi gibi çeşitli alanlarda uygulanan kapsamlı bir derin öğrenme modelidir. Şekil 5.5, bu modelin beş aşamadan oluşan yapısını göstermektedir.



Şekil 5.5 Özelliklerin öğrenilmesinde kullanılan model

Bu model, CNN ve LSTM'in güçlü özelliklerini entegre ederek veriler üzerinde hem özellik çıkarımı hem de zaman serisi tahmini yapabilir. Çalışmamızda, ConvLSTM-AE modeli normal video verilerinin temel özelliklerini öğrenmek ve bu verileri etkili bir şekilde temsil etmek amacıyla kullanılmıştır.

5.2.1 Uzamsal kodlayıcı

Uzamsal Kodlayıcı, verinin özelliklerini belirginleştiren ve sıkıştıran aşamadır. Bu aşama için öncelikle ön işleme adımlarıyla hazır hale getirilmiş verilerin modele girdi sağlayacak bir katman haline getirilmesi gerekmektedir. Denklem 5.3'teki kod ile oluşturulan giriş katmanı, zaman boyutu dâhil olmak üzere (10, 256, 256, 3) şeklinde bir tensörü giriş olarak alır. Tensör, temelde çok boyutlu bir matris veya dizi olarak düşünülebilecek matematiksel bir nesnedir. Bu katman, 10 zaman adımında her biri 256x256 piksel çözünürlüğünde ve 3 kanala sahip bir dizi görüntüyü temsil eder.

$$input_layer = Input(shape = (10, 256, 256, 3), name = "INPUT") \quad (5.3)$$

Desen algılama amacıyla giriş katmanında Conv2D isimli evrişim katmanları ile filtreler uygulanır. Bu süreç, giriş verisinin içerisindeki özgün desenleri tanımlayarak kritik özellikleri belirginleştirir. Süreçte elde edilen önemli özellikler, daha yüksek seviyeli temsil oluşturmak üzere kullanılır. Bu aşama, veri içerisindeki kritik desenleri ayıklayarak ve özellikleri belirginleştirerek, modelin veriyi daha etkin bir şekilde işlemesine olanak tanır.

Denklem 5.4'te gösterildiği şekilde öncelikle giriş katmanına bir Conv2D işlemi uygulanır. Bu işlem, 13x13 boyutunda 128 adet filtre ile giriş verisine evrişim yapar. Her adımda 4 piksel kaydırma yapılacak şekilde “strides” parametresi 4 olarak belirlenirken, çıktının boyutunun girişle aynı tutulması için “padding” aynı bırakılır. Ayrıca, aşınmayı sağlayan “dilation_rate” genişletme oranı (1,1) olarak belirlenir ve veriler arasında hiç boşluk bırakılmadan aşındırma yapılır. Sonuç olarak, Conv2D işlemi ile 128 adet özellik haritası oluşturularak nispeten yüksek seviyedeki desenler

çıkarılmış olur. Elde edilen bu önemli özellikler daha sonra daha yüksek seviyeli temsil oluşturmak için kullanılır.

$$x = \text{TimeDistributed}(\text{Conv2D}(128, (13, 13), \text{strides} = 4, \text{padding} = "same", \text{dilation_rate} = (1, 1)))(\text{input_layer}) \quad (5.4)$$

Evrişim işlemi uygulanan katmanlara, normalizasyon işlemi ve aktivasyon fonksiyonu aşağıdaki Denklem 5.5'teki kod satırı ile entegre edilir:

$$\begin{aligned} x &= \text{LayerNormalization}()(x) \\ x &= \text{ReLU}()(x) \end{aligned} \quad (5.5)$$

Katman normalizasyonunun temel amacı, bir katmandan gelen çıktı değerlerini normalize ederek, ortalamasını "0" ve standart sapmasını "1" birime dönüştürmektir. Bu işlem, verileri stabilize ederek derin ağlarda gradyanların dengeli bir şekilde ilerlemesini ve geriye doğru yayılan gradyanların küçülmesi veya patlaması gibi problemlerin önüne geçmeyi hedefler. Bu sağlanan denge, ağın daha hızlı ve istikrarlı bir şekilde eğitilmesine olanak tanır.

ReLU (Rectified Linear Unit) katmanı ise, yapay sinir ağlarındaki yaygın bir aktivasyon fonksiyonudur ve modelde eklenmesinin ana sebebi, negatif girdileri sıfıra sınırlandırarak modelin öğrenme sürecini iyileştirmesidir. ReLU, kaybolan gradyan riskini azaltır, modelin doğrusal olmayan özellikler kazandırır ve daha hızlı eğitim süreçlerine katkıda bulunur. Bu özellikler, modelinizin daha stabil ve etkili bir şekilde eğitilmesini sağlar ve genel model performansını iyileştirir.

Aynı yaklaşım, ikinci ve üçüncü evrişim katmanlarının oluşturulmasında da uygulanır. Denklem 5.6'da gösterildiği şekilde ikinci evrişim katmanında, 64 farklı 7x7 boyutunda filtre kullanılırken, üçüncü evrişim katmanında 3x3 boyutunda 32 filtre kullanılır. Her bir evrişimli katmanın ardından normalizasyon işlemi uygulanır.

$$\begin{aligned} x &= \text{TimeDistributed}(\text{Conv2D}(64, (7, 7), \text{strides} = 4, \text{padding} = "same"))(x) \\ x &= \text{LayerNormalization}()(x) \\ x &= \text{ReLU}()(x) \end{aligned} \quad (5.6)$$

$$\begin{aligned}
x &= \text{TimeDistributed}(\text{Conv2D}(32, (3, 3), \text{strides} = 2, \text{padding} = "same"))(x) \\
x &= \text{LayerNormalization}()(x) \\
x &= \text{ReLU}()(x)
\end{aligned}$$

5.2.2 Zamansal kodlayıcı

Zamansal Kodlayıcı, ConvLSTM2D gibi katmanlar kullanılarak verinin zaman içindeki değişikliklerini modelleyen aşamadır. Bu katmanlar, 3 boyutlu verileri işlemek için tasarlanmıştır. Bu aşamada, ConvLSTM2D katmanları ile verinin zamansal özellikleri belirginleştirilir. Denklem 5.7, zamansal kodlayıcı katmanlarını içerir.

$$\begin{aligned}
x &= \text{ConvLSTM2D}(32, (3, 3), \text{padding} = "same", \text{return_sequences} = \text{True})(x) \\
x &= \text{LayerNormalization}()(x)
\end{aligned} \tag{5.7}$$

Bu aşamada ConvLSTM2D katmanı, verinin zamansal özelliklerini öne çıkarmak için kullanılır. Bu katman, 3x3 boyutunda 32 farklı filtre kullanarak veriyi işler ve verinin farklı zaman adımlarındaki özelliklerini çıkarır. Çıktının boyutunu girişle aynı tutulması amacıyla “padding” değeri aynı bırakılır. Ardından, katmanın çıktısı normalize edilir. Bu, modelin daha dengeli ve istikrarlı bir şekilde eğitilmesine olanak sağlar.

5.2.3 Gizli kod

Gizli Kod, bir modelin öğrendiği temsilin, daha küçültülmüş ve öz bir formudur. Bu kod, verinin daha düşük boyutlu bir ifadesini içerir ve genellikle giriş verisinin önemli özelliklerini koruyarak daha kompakt bir temsil elde etmek amacıyla kullanılır. Gizli Kod, daha az bellek ve işlem gücü gerektirdiğinden etkili bir şekilde saklanabilir ve kolay işlenebilir. Denklem 5.8’de ConvLSTM2D katmanı, 8 filtre ile (3,3) boyutunda bir çekirdek kullanarak giriş verisini işler. Ayrıca, bu katman zaman serisinin her bir zaman adımında bir çıktı üretir. Daha sonra, bu katmanın çıktısı “CODE” adıyla adlandırılan ayrı bir değişkene atanır. Bu, verinin sıkıştırılmış bir temsiliyi yani Gizli Kodu ifade eder.

$$\begin{aligned} \text{code_layer} &= \text{ConvLSTM2D}(8, (3, 3), \text{padding} = "same", \text{return_sequences} = \text{True})(x) \\ \text{code_layer} &= \text{LayerNormalization}(\text{name} = "CODE")(\text{code_layer}) \end{aligned} \quad (5.8)$$

5.2.4 Zamansal çözücü

Zamansal Çözücü aşamasında, Conv2DTranspose gibi katmanlar kullanılarak önceki aşamalarda sıkıştırılan verinin tekrar genişletilmesi amaçlanır. Bu katmanlar, veriyi genişleterek daha yüksek seviyeli özelliklerin belirginleşmesini sağlar. Bu aşama, modelin veriyi orijinal boyutlarına yakın bir şekilde yeniden oluşturmasına olanak sağlar. Denklem 5.9'da 3x3 boyutunda 32 filtre kullanılarak veri işlenir ve zaman serisinin her bir zaman adımında bir çıktı üretilir.

$$\begin{aligned} x &= \text{ConvLSTM2D}(32, (3, 3), \text{padding} = "same", \text{return_sequences} = \text{True})(\text{code_layer}) \\ x &= \text{LayerNormalization}()(x) \end{aligned} \quad (5.9)$$

5.2.5 Uzamsal çözücü

Uzamsal Çözücü, verinin uzamsal boyutunu geri döndürmek amacıyla kullanılan bir bileşendir. Bu aşamada, özellik piramidi konsepti kullanılarak öğrenilen özellikler farklı seviyelerde sıkıştırılır ve ardından Gizli Kodun orijinal uzamsal boyutlarına genişletilmesi hedeflenir. Bu süreç, verinin özelliklerini daha yüksek çözünürlükte ve orijinal formunda elde etmeyi amaçlar. Bu amacı gerçekleştirmek için, veriye ters evrişim (Conv2DTranspose) işlemleri uygulanır.

Denklem 5.10'da ilk olarak giriş verisi Conv2DTranspose işlemi uygulanarak daha yüksek çözünürlükte genişletilir. Bu işlem, 3x3 boyutunda 32 filtre kullanılarak gerçekleştirilir. Her adımda 2 piksel kaydırma yapılır ve çıkış boyutları girişle aynı tutulur. İkinci aşamada, 7x7 boyutunda 64 filtre kullanılır. Her adımda 4 piksel kaydırma yapılır ve çıkış boyutları girişle aynı tutulur. Son olarak, 128 farklı 13x13 boyutunda filtre kullanılır. Her adımda 4 piksel kaydırma yapılır ve çıkış boyutları girişle aynı tutulur.

Son olarak, çıkış katmanı oluşturulur. Bu katman, veriyi en geniş haliyle yeniden oluşturmak için 1 filtre kullanır. Sigmoid aktivasyonu kullanılarak çıkışın 0 ile 1 arasında olması sağlanır. Bu, ağın nihai çıkışını temsil eder. Model, Adam optimizasyon algoritması ve ortalama karesel hata kaybı fonksiyonu kullanılarak derlenir. Daha sonra, belirli bir sayıda epok için belirtilen parti boyutuyla eğitim veri kümesi üzerinde eğitilir. Bu süreçte, model normal olayları tanımak ve anomalileri tespit etmek için öğrenir.

```

x = TimeDistributed(Conv2DTranspose(32, (3, 3), strides = 2, padding = "same"))(x)
x = LayerNormalization()(x)
x = TimeDistributed(Conv2DTranspose(64, (7, 7), strides = 4, padding = "same"))(x)
x = LayerNormalization()(x)
x = TimeDistributed(Conv2DTranspose(128, (13, 13), strides = 4, padding
= "same"))(x)
x = LayerNormalization()(x)
output_layer = TimeDistributed(Conv2D(1, (13, 13), activation = "sigmoid", padding
= "same", name = "OUTPUT"))(x)
My_model = Model(input_layer, output_layer)
My_model.compile(optimizer = 'adam', loss = 'mse')

```

(5.10)

5.2.6 Kayıp fonksiyonları

Kayıp fonksiyonları (loss functions), makine öğreniminde bir modelin tahminlerinin gerçek etiket değerlerinden ne kadar sapma gösterdiğini ölçen fonksiyonlardır. Modelin performansını iyileştirmek için kullanılan bu fonksiyonlar, model eğitimi sırasında minimize edilmeye çalışılır.

1. Ortalama Kareler Hatası (Mean Squared Error- MSE): MSE, gerçek değerler ile tahmin edilen değerler arasındaki farkın karesinin ortalamasıdır. Bu, modelin tahminlerinin ne kadar doğru olduğunu ölçmek için kullanılır. Farkın karesi alınarak, büyük hatalara daha fazla ağırlık verilir. Formülü Denklem 5.11’de gösterilmiştir. “ N ” örnek sayısını, “ y_i ” gerçek değerleri ve “ $\hat{y}_i$ ” tahmin edilen değerleri göstermektedir.

$$MSE = \frac{1}{N} \sum_{i=1}^N (y_i - \hat{y}_i)^2 \quad (5.11)$$

2. Ortalama Mutlak Hata (Mean Absolute Error- MAE): MAE, gerçek değerler ile tahmin edilen değerler arasındaki mutlak farkın ortalamasıdır. Bu, modelin tahminlerinin ne kadar doğru olduğunu ölçmek için kullanılır ve büyük hataların etkisini azaltarak daha dengeli bir ölçüm sağlar. Formülü Denklem 5.12’de gösterilmiştir.

$$MAE = \frac{1}{N} \sum_{i=1}^N (y_i - \hat{y}_i) \quad (5.12)$$

3. Çapraz Entropi Kaybı (Cross-Entropy Loss): İkili sınıflandırma problemlerinde kullanılır. Gerçek değerlerle tahmin edilen olasılıklar arasındaki farkı ölçer. Bu fonksiyon, modelin doğru sınıflandırmalar yapmasını teşvik eder ve olasılık tahminlerinin doğruluğunu optimize etmeye çalışır. Formülü Denklem 5.13’te gösterilmiştir.

$$BCE = -\frac{1}{N} \sum_{i=1}^N [y_i \log(\hat{y}_i) + (1 - y_i) \log(1 - \hat{y}_i)] \quad (5.13)$$

MSE, anormallik tespiti gibi alanlarda sıklıkla tercih edilen bir kayıp fonksiyonudur ve bu tercihin altında yatan birkaç önemli sebep bulunmaktadır. İlk olarak, MSE, hataları kare olarak hesaplayarak büyük hatalara daha fazla ağırlık verir. Bu durum, özellikle gözetim ve anormallik tespiti gibi hassas uygulamalarda, küçük hata miktarlarının bile önemli olabileceği durumlarda büyük önem taşır. MSE'nin bu büyük hata duyarlılığı, modelin anormal durumları daha doğru tespit etmesine yardımcı olur. Ayrıca, MSE'nin sürekli ve türevlenebilir yapısı, gradyan tabanlı optimizasyon yöntemlerinde, örneğin geri yayılım algoritmasında, etkili kullanımı sağlar. Bu matematiksel nitelik, modelin eğitim sürecini hızlandırır ve kolaylaştırır.

MSE'nin basitliği ve uygulanabilirliği de model geliştirme sürecini hızlandırır, kayıp fonksiyonunun uygulanmasını ve ayarlanmasını kolaylaştırır. Dolayısıyla, MSE'nin modelinizde kullanımı, bu yöntemin matematiksel ve uygulamaya yönelik avantajlarını göz önünde bulundurarak, anormallik tespiti ve gözetim gibi hassas alanlarda modelin etkinliğini artırmaya yönelik bir tercihtir.

5.2.7 Optimizasyon algoritmaları

Optimizasyon algoritmaları, modelin kayıp fonksiyonunu minimize ederek, en iyi parametre değerlerini bulmasına yardımcı olur. Aşağıda, derin öğrenme modellerinde yaygın olarak kullanılan üç ana optimizasyon algoritmasının açıklaması ve formülleri bulunmaktadır:

1. Stokastik Gradyan İnişi (Stochastic Gradient Descent- SGD): SGD, gradyan tabanlı bir optimizasyon yöntemidir ve her güncelleme adımında rastgele seçilen tek bir veri noktası üzerinden hesaplama yapar. Bu yöntem, her adımda tüm veri kümesini kullanmak yerine rastgele seçilmiş alt küme üzerinden hesap yaparak hesaplama yükünü azaltır. Denklem 5.14'te bulunan formülde “ θ ” model parametrelerini, “ η ” öğrenme oranını, “ J ” kayıp fonksiyonunu, “ $x^{(i)}$ ” ve “ $y^{(i)}$ ” ise rastgele seçilen eğitim örneğini ve etiketini temsil eder.

$$\theta = \theta - \eta \cdot \nabla_{\theta} J(\theta; x^{(i)}; y^{(i)}) \quad (5.14)$$

2. RMSprop: Bu algoritma, özellikle tekrar eden sinir ağlarında kullanılan bir optimizasyon algoritmasıdır. Gradyanların karesel ortalamasını kullanarak, her parametre için özelleştirilmiş bir öğrenme oranı sağlar. Bu yaklaşım, ağırlıkların daha dengeli bir şekilde güncellenmesine ve algoritmanın daha hızlı yakınsamasına olanak tanır. Denklem 5.15'teki sunulan “ v_t ” ağırlık güncellemelerinin karesel ortalamasını, “ g_t ” gradyanı, “ β ” ağırlıklandırma katsayısını, “ η ” öğrenme oranını ve “ ϵ ” kararlılık sabitini temsil eder.

$$\begin{aligned} v_t &= \beta v_{t-1} + (1 - \beta) g_t^2 \\ \theta &= \theta - \frac{\eta}{\sqrt{v_t} + \epsilon} g_t \end{aligned} \quad (5.15)$$

3. Adam Optimizasyon Algoritması: Adam, SGD'nin bir varyasyonudur ve her bir parametre için özelleştirilmiş adaptif öğrenme oranları kullanır. Bu algoritma hem momentum hem de RMSprop'un avantajlarını birleştirerek, özellikle gürültülü veri

kümelerinde ve karmaşık modellerde daha hızlı ve etkili bir şekilde yakınsama sağlar. Momentum, özellikle derin öğrenme modellerinde sıkça karşılaşılan düzlükler ve yerel minimumlar gibi zorlu alanlarda hız kazandırmak için kullanılır. Denklem 5.16’da “ m_t ” ve “ v_t ” sırasıyla birinci ve ikinci moment tahminlerini, “ g_t ” gradyanı, “ β_1 ” ve “ β_2 ” moment tahminlerinin ağırlıklandırılma oranlarını, “ η ” öğrenme oranını ve “ ϵ ” kararlılık sabitini temsil eder.

$$\begin{aligned}
 m_t &= \beta_1 m_{t-1} + (1 - \beta_1) g_t \\
 v_t &= \beta_2 v_{t-1} + (1 - \beta_2) g_t^2 \\
 \widehat{m}_t &= \frac{m_t}{1 - \beta_1^t} \\
 \widehat{v}_t &= \frac{v_t}{1 - \beta_2^t} \\
 \theta &= \theta - \frac{\eta \cdot \widehat{m}_t}{\sqrt{\widehat{v}_t} + \epsilon}
 \end{aligned} \tag{5.16}$$

Bu algoritmalar, modelin eğitim sürecinde kayıp fonksiyonunu etkili bir şekilde minimize ederek modelin performansını artırır. Seçilen optimizasyon algoritması, modelin yapısı, veri kümesinin özellikleri ve hedeflenen performans hedeflerine bağlı olarak değişebilir. Her algoritmanın kendine özgü avantajları ve uygulama koşulları bulunur, bu nedenle modelin gereksinimlerine en uygun optimizasyon algoritmasının seçilmesi önemlidir.

Tezde önerilen modelde Adam optimizasyon algoritmasının tercih edilmesi, çeşitli avantajları beraberinde getirir. Adam, adaptif öğrenme oranları sağlayarak her bir parametre için öğrenme sürecini özelleştirir. Bu özellik, özellikle parametreler arasındaki gradyan varyansının yüksek olduğu durumlarda önemlidir. Ayrıca, geçmiş gradyan bilgilerini dikkate alarak yönlendirme yapar ve aynı zamanda parametre bazında öğrenme oranlarını ayarlar. Bu özellik, modelin yerel minimumlara sıkışma gibi sorunlara karşı daha dayanıklı olmasını sağlar. Genel olarak, Adam’ın kullanımı, hızlı konverjans, geniş uygulama alanı, dayanıklılık ve stabilite gibi özellikleriyle derin öğrenme modellerinde başarılı sonuçlar elde etmede önemli bir role sahiptir.

5.3 Anormallik Tespiti

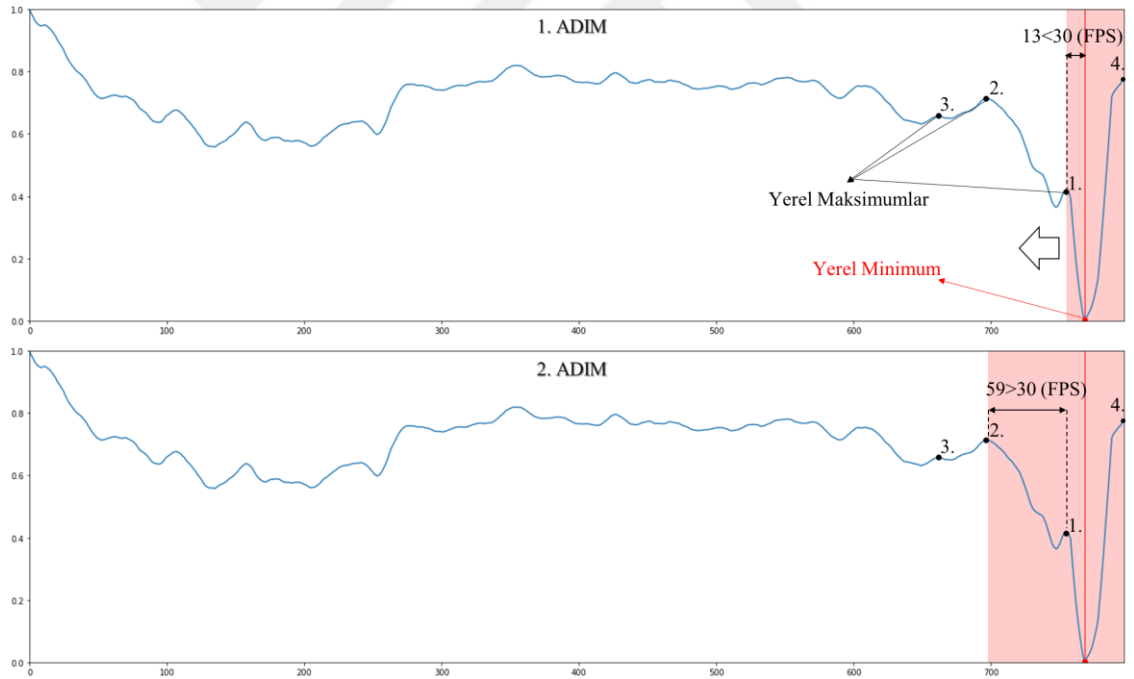
Modelin eğitimi sonrasında, normal videoların karakteristik özelliklerini öğrenir ve bu bilgiyi test veri kümesindeki videoları yeniden oluşturmak için kullanır. Normal videolar model tarafından etkin bir şekilde yeniden oluşturulur ve düşük seviyeli, düzenli çıktılar üretir. Buna karşın, anormal videolar model tarafından yeniden oluşturulduğunda daha yüksek seviyeli ve düzensiz çıktılar meydana gelir. Modelin ürettiği bu sayısal değerler, 0 ile 1 arasında değişen yeniden oluşturma hataları olarak değerlendirilir ve bu hataların “1”den çıkarılmasıyla elde edilen “düzenlilik puanı”, her bir test videosu için bir dizi olarak sunulur. Bu dizideki verilerin değişim eğilimleri gözlemlenerek yerel minimum ve maksimumlar belirlenir. Bu noktaların, bir anormal olaya mı yoksa hatalı pozitif bir duruma mı işaret ettiğini belirlemek kritiktir.

Bu nedenle, düzenlilik puanı dizisindeki yerel minima noktalarının belirli bir eşikle belirlenmesi gereklidir. Bu aşamada, metodolojimizde Persistence1d (Tino 2023) yöntemi kullanılmıştır. Persistence1D, bir boyutlu verilerde yerel ekstremumları ve bunların kalıcılık (persistence) değerlerini bulmak için kullanılan bir yöntemdir. Bu yöntem, bir dizideki yerel minimum ve maksimum noktalarını tanımlar ve bunları kalıcılık değerlerine göre sıralar. Kalıcılık değeri, bir noktanın komşularına göre ne kadar yüksek veya alçak olduğunu gösterir. Bu metodoloji ile düzenlilik skorundaki gürültülü ve anlamsız yerel minimumlar azaltılarak, anlamlı minimumlar vurgulanır.

Yerel minima noktaları, potansiyel anormal hareketlerin odak noktalarını temsil eder. Bu noktalar, genellikle hareketin azaldığı, durduğu veya ani değiştiği anları işaret eder. Bu nedenle anormallik tespitinde kritik bir rol oynar. Anormallik tespitinde, yerel minimum noktasının komşu maksimum noktalara olan uzaklığı, anormalliğin devamı ya da sonlanması hakkında kritik bilgiler sunar. Eğer bu uzaklık, videonun kare hızı (FPS) ile ölçülen süreden (örneğin 24 veya 30 kare/saniye) büyük veya eşitse, ilk komşu maksimum noktalar, anormal hareketin sınırları olarak kabul edilir. Ancak, bu mesafe videonun kare hızından daha küçükse, bu durum anormal hareketin muhtemelen devam ettiği şeklinde yorumlanır. Bu sebeple anormal hareketin sınırları, bir sonraki yerel

maksimumlarda aranmaya devam edilir. Bu döngüsel süreç, anormal hareketin başlangıç ve bitiş zamanlarını daha doğru bir şekilde tespit etmemize olanak tanır.

Veri kümelerinin incelenmesi sonucunda, çoğu anormal hareketin 2 saniyeden uzun sürdüğü belirlenmiştir. Bu durum, çerçeve sayısı açısından 2xFPS oranına karşılık gelmektedir. Şekil 5.6'da üst grafikte kırmızı nokta ile işaretlenmiş yerel minimum, siyah nokta ile gösterilen 1 numaralı yerel maksimuma 13 kare uzaklıkta yer alıyor. Bu mesafe, video hızı olan 30 FPS'ten düşük olduğundan anormal hareketin devam ettiği tahmin edilir ve bir sonraki yerel maksimum noktasında başlangıç noktası aranır. Bu adımda 1 ve 2 numaralı yerel maksimumlar arasındaki uzaklık 59 kare hesaplanır. Bu mesafe 30 FPS'ten yüksek olduğu için, süreç 3. yerel maksimum noktası uzaklığı hesaplanmadan sonlandırılır. Eş zamanlı olarak benzer süreç aynı şekilde yerel minimum noktasının diğer tarafı için de gerçekleştirilir.



Şekil 5.6 Anormallik tespit adımları

Sonuç olarak, anormal hareketin başlangıç noktası 2 numaralı yerel maksimum ve bitiş noktası 4 numaralı yerel maksimum olarak belirlenmiş olur. Hareketin başlangıç ve bitiş noktalarının tespit edilmesi sonrasında, bu iki nokta arasındaki kırmızı renkli bölge,

anormal hareketin devam ettiği alanı belirtir. Bu bölge, modelin belirlediği anormal özelliklere sahip olan karelerin yoğunlaştığı alandır. Hareketin başladığı ve bittiği zaman arasında meydana gelen olaylar, genellikle anormallik tespiti açısından önemli bilgiler içerir.

Bu anormallik bölgesi içinde kalan kareler, modelin eğitildiği normal davranıştan sapma gösteren kareler olarak işaretlenir. Bu karelerdeki anormallik, genellikle yüksek seviyeli ve düzensiz çıktılar şeklinde tezahür eder. Öte yandan, bu bölge dışında kalan kareler, model tarafından belirlenen normal davranışı sergileyen kareler olarak kabul edilir. Bu özgün analiz yöntemi, VGAT yeteneğini önemli ölçüde geliştirir.

5.4 Değerlendirme Metrikleri

Anormallik tespiti aşamasında video karelerinin normal veya anormal olarak tespit edilmesi ile bu tespitlerin gerçek etiket değerleriyle karşılaştırılması ve performansının ölçülmesi gereklidir. Bu ölçümler, aşağıdaki parametrelerin hesaplanmasıyla gerçekleştirilir. Aşağıda, bu parametrelerin açıklamaları yer almaktadır:

1. Doğru Pozitifler (True Positives-TP): Modelin gerçek anormal durumları doğru bir şekilde tespit ettiği durumları temsil eder. Yani, gerçekte var olan bir anormalliği modelin anomali olarak tespit ettiği durumları ifade eder.
2. Doğru Negatifler (True Negatives-TN): Modelin anormallik bulunmayan normal durumları doğru bir şekilde normal olarak tanımladığı durumları temsil eder.
3. Hatalı Pozitifler (False Positives-FP): Modelin normal durumları yanlışlıkla anormal olarak tanımladığı durumları temsil eder. Yani, gerçekte var olmayan bir anormalliği modelin anormal olarak sınıflandırdığı durumları ifade eder.
4. Hatalı Negatifler (False Negatives-FN): Modelin gerçek anomali durumlarını yanlışlıkla normal olarak tanımladığı durumları temsil eder.

Bir video anomalisi tespit modelinin etkinliğini deęerlendirmek için yaygın olarak Eğri altındaki alan (The Area Under the Curve-AUC) metrięi kullanılır. Bu metrik, yukarıda açıklanan parametreler kullanılarak aşağıdaki alt metriklerden oluşur:

1. Doğru Pozitif Oranı (True Positive Rate-TPR): Modelin gerçek anomalileri doğru bir şekilde tespit etme yeteneęini ölçer. Yüksek bir TPR, modelin anomalileri doğru bir şekilde tespit edebildięini gösterir. Matematiksel formülü Denklem 5.17’de sunulmuştur.

$$TPR = \frac{TP}{TP + FN} \quad (5.17)$$

2. Hatalı Pozitif Oranı (False Positive Rate-FPR): Modelin anormal olmayan durumları yanlışlıkla anormal olarak tanımlama oranını ölçer. Düşük bir FPR, modelin yanlış alarm verme eğiliminde olmadığını gösterir. Matematiksel formülü Denklem 5.18’de sunulmuştur.

$$FPR = \frac{FP}{FP + TN} \quad (5.18)$$

3. Doğruluk (Accuracy): Modelin toplam doğru tahmin sayısını ölçer. Yüksek bir doğruluk, modelin genel olarak doğru sonuçlar verdięini gösterir. Matematiksel formülü Denklem 5.19’de sunulmuştur.

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \quad (5.19)$$

4. Kesinlik (Precision): Modelin tahmin ettięi anomalilerin gerçekten anormal olanların oranını ölçer. Yüksek bir kesinlik, modelin hatalı pozitiflere eğilimli olmadığını gösterir. Matematiksel formülü Denklem 5.20’de sunulmuştur.

$$Precision = \frac{TP}{TP + FP} \quad (5.20)$$

5. F1 Skoru: Kesinlik ve duyarlılığın ağırlıklı ortalamasını hesaplar ve bu şekilde bir denge sağlar. Daha yüksek bir F1 skoru, kesinlik ve duyarlılık arasında daha iyi bir denge olduğunu gösterir. Matematiksel formülü Denklem 5.21’de sunulmuştur.

$$F1 \text{ Score} = \frac{2 \times Precision \times Recall}{Precision + Recall} \quad (5.21)$$

Bu metrikler, VGAT modellerinin performansını değerlendirmek için birbirini tamamlayan bilgiler sağlar ve modelin doğru bir şekilde anomalileri tespit etme ve yanlış alarm verme oranını anlamamıza yardımcı olur.

AUC değeri, TPR ve FPR’nin bir grafik üzerinde çizilerek hesaplanır ve eğrinin altındaki alanı temsil eder. Yüksek bir AUC değeri, modelin TPR ve FPR arasında iyi bir denge sağladığını ve videodaki anormallikleri doğru bir şekilde tanımlayabildiğini gösterir. AUC değeri ne kadar yüksekse, modelin anormallikleri tespit etmede o kadar iyi performans gösterdiği anlamına gelir ve 1.0 değeri mükemmel bir performansı temsil eder. VGAT alanında, AUC değerlendirmesi, genellikle modelin performansını kapsamlı bir şekilde değerlendirmek için hassasiyet, geri çağırma ve doğruluk gibi diğer performans ölçütleriyle birlikte kullanılır.

6. ÖNERİLEN METODOLOJİNİN UYGULANMASI

Önerilen metodolojinin performans ve doğruluğunu değerlendirmek için CUHK Avenue ve UMN halka açık veri kümeleri kullanılmıştır. Bu değerlendirme, metodolojinin gerçek dünya senaryolarında nasıl bir performans sergilediğini anlamak için önemlidir. Yapılan uygulamaların detayları aşağıdaki bölümlerde verilmiştir.

6.1 Kullanılan Teknolojiler ve Veri Kümesi

Önerilen metodolojinin uygulanmasında, Python programlama dili, Keras derin öğrenme kütüphanesi, OpenCV görüntü işleme aracı ve NVIDIA Tesla P100 grafik işlem birimi ile donatılmış bir sistem kullanılmıştır. Bu teknolojik altyapı, metodolojinin etkili ve verimli bir şekilde uygulanmasını sağlamakta ve karmaşık veri işleme süreçlerini desteklemektedir.

Python, geniş kütüphane desteği ve esnekliği ile veri analizi ve makine öğrenimi projelerinde tercih edilen bir dildir. Bu çalışmada Python, derin öğrenme model tasarımı ve geliştirilmesinde kullanılmıştır. Keras, TensorFlow üzerine inşa edilmiş, modüler ve kullanıcı dostu bir derin öğrenme kütüphanesidir. Bu kütüphane, yapay sinir ağları modellerinin kolaylıkla oluşturulması ve eğitilmesi için tercih edilmiştir. OpenCV, görüntü işleme ve bilgisayarla görü işlemleri için yaygın olarak kullanılan bir kütüphanedir. Videolardan kare çıkarma, yeniden boyutlandırma, arka plan çıkarma ve gürültü azaltma gibi ön işleme işlemlerinde önemli bir role sahiptir. OpenCV'nin sağladığı özellikler, video verilerinin analize uygun hale getirilmesinde ve modelin doğruluğunu artırmada kritik önem taşır.

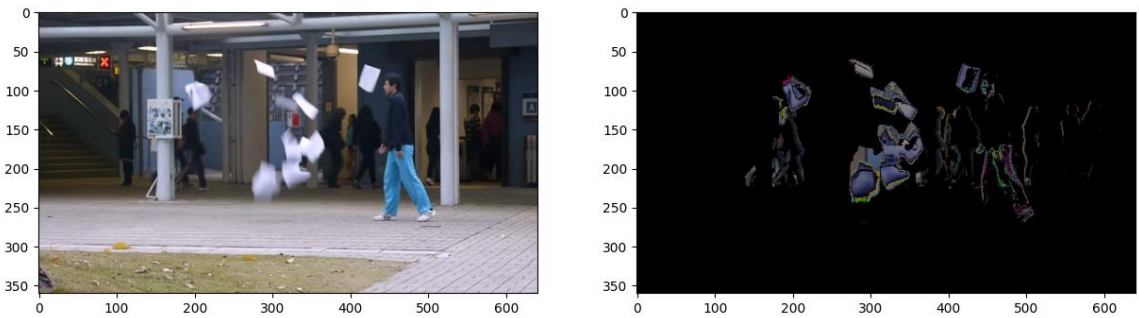
Dördüncü Bölümde detayları anlatılan CUHK Avenue ve UMN veri kümeleri, gerçek dünya senaryolarını yansıtan, zengin ve çeşitli video koleksiyonlarına sahiptirler. Her iki veri kümesi de algoritmaların gerçekçi ortamlarda nasıl bir performans sergilediğini değerlendirmek için ideal veri kümelerinden bazılarıdır. Bu çalışmada, her iki veri kümesi kullanılarak önerilen metodolojinin etkinliği ve doğruluğu ayrıntılı bir şekilde

incelenmiştir. Bu analizler, VGAT algoritmalarının gerçek dünya koşullarında nasıl performans gösterdiğine dair değerli bilgiler sunmaktadır ve bu alanın gelecekteki araştırmalarına yol gösterici olmaktadır.

6.2 Metodolojinin Uygulanması

CUHK Avenue veri kümesi 25 FPS hızında, UMN veri kümesi ise 30 FPS hızında kaydedilmiş görüntüler içermektedir. Öncelikle her veri kümesinden saniyede 25 ve 30 kare olacak şekilde ardışık kareler çıkarılır. Ardından, görüntü işleme ve model eğitimi süreçlerine uygun hale getirmek için CUHK Avenue veri kümesinin orijinal 640x360 piksel ve UMN veri kümesinin orijinal 320x240 piksel çözünürlüğündeki kareleri, 256x256 piksel boyutuna ölçeklendirilir. Bu ölçeklendirme işleminde, görüntü kalitesinin ve detaylarının mümkün olduğunca korunması sağlanır.

Modelimizde, hareketli nesnelerin belirginleştirilmesi için arka planın çıkarılması esas alınmaktadır. Bu süreç, OpenCV kütüphanesinde bulunan MOG2 yöntemi kullanılarak gerçekleştirilir. MOG2, her piksel için birden fazla Gaussian dağılımını modelleyerek, dinamik arka planları ve değişken aydınlatma koşullarını etkili bir şekilde ele alır. Bu yöntem, hareketli nesnelerin ve arka planın ayırt edilmesini sağlayan siyah-beyaz bir maske üretir. Daha sonra, bu maske RGB renk uzayına dönüştürülür ve orijinal renkli görüntüyle birleştirilir. Bu yöntemle, hareketli nesneler orijinal renk bilgisiyle öne çıkarılırken, arka plan efektif bir şekilde bastırılır. Şekil 6.1’de arka plan çıkarma örneği gösterilmektedir.



Şekil 6.1 Arka plan çıkarma işlemi sonucunda elde edilen kare

Son olarak karelerdeki gürültüyü azaltmak için “erode” yöntemi kullanılır. Bu süreçte, 3x3’lük kernel ile üç aşındırma iterasyonu gerçekleştirilir. Bu adım, gürültüyü azaltarak görüntü kalitesini artırır ve modelin daha doğru analiz yapmasını sağlar.

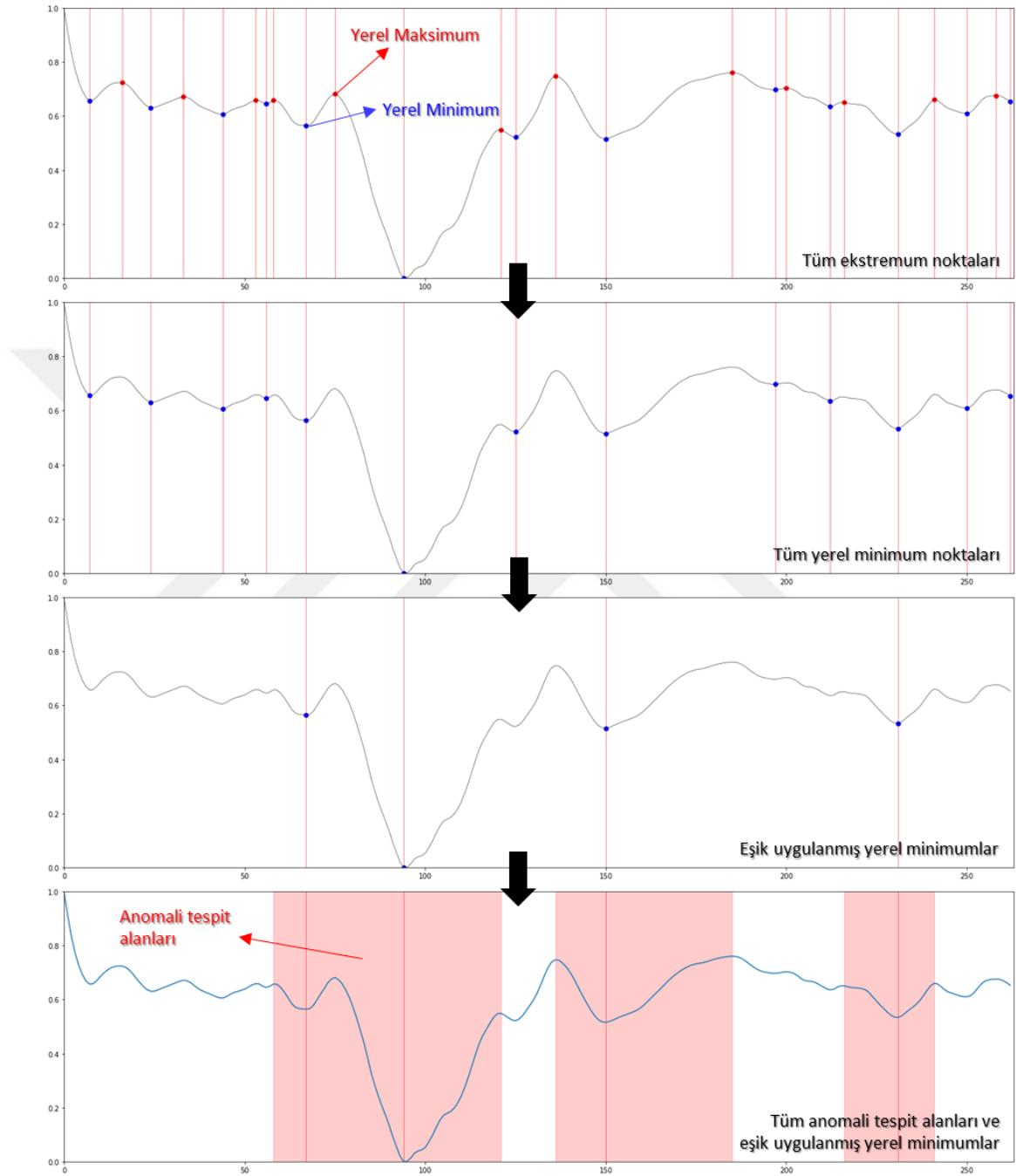
Ön işlemde geçen kareler, “sliding window” yöntemiyle 10 karelik gruplar halinde model girişine verilir. Modelin girdisi olarak alınan 10’ar karelik pencere dizileri, modelin uzamsal ve zamansal özellikleri öğrenmesi için önemlidir. Her bir kare, model tarafından işlenir ve bu süreçte Conv2D katmanları karelerin uzamsal özelliklerini, ConvLSTM2D katmanları ise zamansal dinamikleri öğrenir. Bu iki özellik türünün bütünleşmesi, modelin veri kümesinin temel karakteristiklerini öğrenmesini sağlar ve sıkıştırılmış veri temsili olan gizli kodu oluşturur. Modelin eğitimi süresince normal davranış kalıplarını temel alarak anormallikleri belirleme yeteneği kazanır.

Model, test veri kümesindeki kareleri işleyerek yeniden oluşturma süreci üzerinden anormallik tespiti yapar. Normal davranışları içeren kareler başarılı bir şekilde yeniden oluşturulurken, anormal hareketler içeren karelerde yüksek yeniden oluşturma hatası veya düzenlilik puanında düşüş meydana gelir. Düzenlilik puanı dizisi, modelin test veri kümesindeki kareleri yeniden oluşturma başarısını ifade eder. Bu dizideki düşük puanlar, potansiyel anormalliklerin işaretçisidir. Bu puanlar, 0 (düşük benzerlik, yüksek anormallik potansiyeli) ile 1 (yüksek benzerlik, düşük anormallik potansiyeli) arasında değişir. Ancak, tüm düşük puanlar gerçek anormalliklere işaret etmeyebilir. Bu nedenle, dizideki yerel minimumlar (düşük puanlar) ve bitişik maksimumlar (yüksek puanlar) arasındaki farklar dikkatlice incelenmelidir.

Persistence1D yöntemi, bu analiz sürecinde kullanılan bir tekniktir. Bu yöntem, yerel minima ve maksima noktaları sıralayarak ve kalıcılık değerlerine göre önceliklendirerek kullanılır. Böylece, gürültülü ve önemsiz düşük puanlar azaltılır ve gerçekten anlamlı olan düşük puanlar vurgulanır.

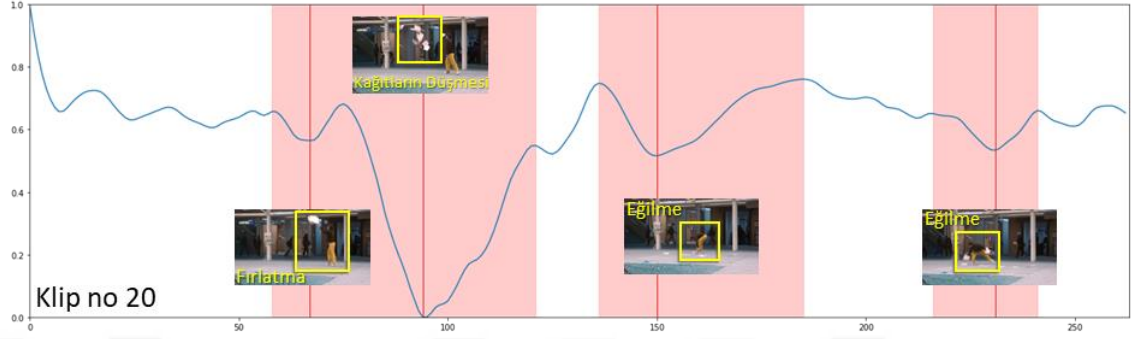
Detayları Bölüm 5.3’te açıklanan anormallik tespiti, belirlenen eşik değerlerine göre anlamlı yerel minimumlar ile bitişik maksimum noktalar arasındaki alanlarda yapılır. Bu şekilde model, veri kümesindeki anormal hareketleri başarılı bir şekilde tespit

edebilir ve görselleştirebilir. Şekil 6.2 CUHK Avenue veri kümesindeki 20 numaralı test videosunun “Anormallik Tespiti” sürecini göstermektedir.



Şekil 6.2 Anormallik tespiti uygulanmış bir model çıktısı

Şekil 6.3'te ise anormallik tespiti işlemine ait sonuçlar gösterilmektedir. Bu videoda, bir kişinin sahnenin ortasında kağıtları havaya fırlatması ve ardından bu kağıtları toplaması, modelimiz tarafından anormal olarak sınıflandırılmıştır.



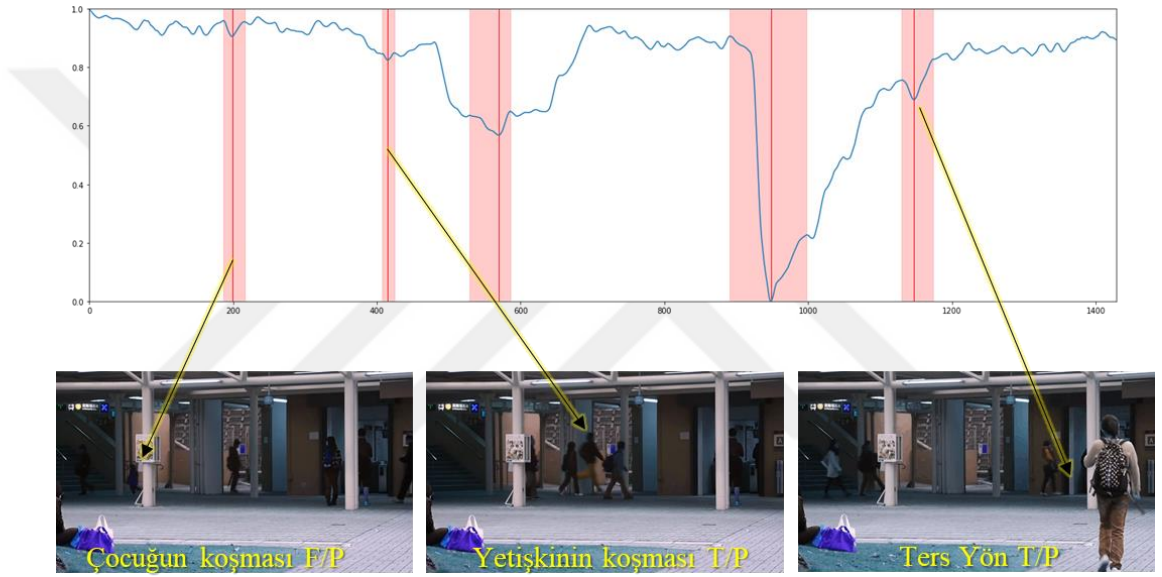
Şekil 6.3 CUHK Avenue veri kümesindeki test videosuna uygulanan anormallik tespit işlemi

Bu tip hareketler, günlük yaşamda nadiren karşılaşılan ve standart davranış kalıplarından sapmalar olarak değerlendirilmektedir. Yapılan tespitler, videodaki belirli hareket ve aktivitelerin anormal olarak algılanmasını sağlayan gerçek etiket değerlerine dayanır. Elde edilen bu tespitler, modelin performansını değerlendirmek için önemli metriklerin hesaplanmasına zemin hazırlar. Bu metrikler, modelin anormal hareketleri ne kadar doğru ve etkin bir şekilde tespit edebildiğini gösterir ve araştırmanın güvenilirliğini artırır.

6.3 Uygulama Sonuçları

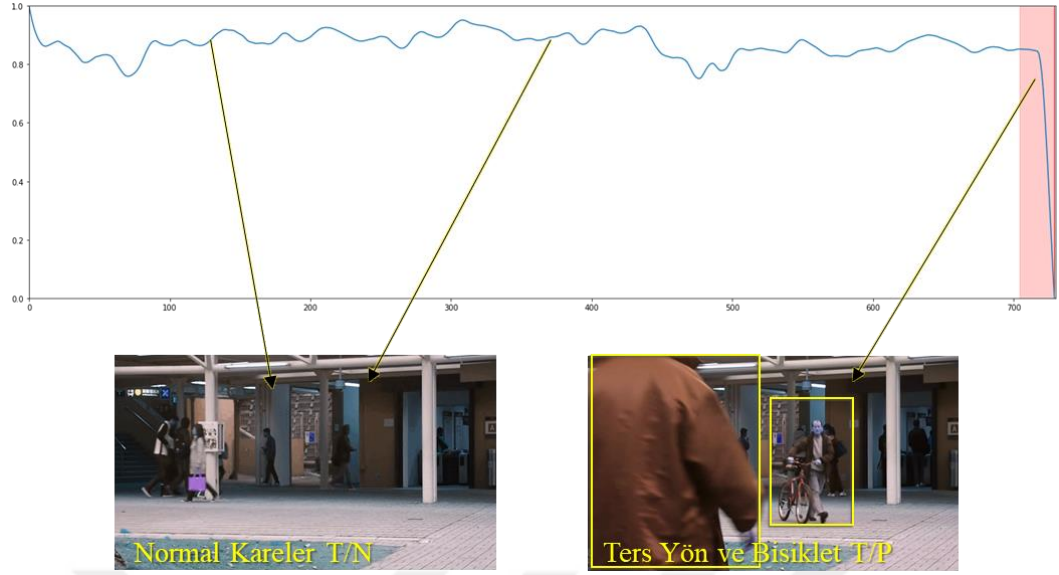
Bu bölümde, metodolojimizin uygulama sonuçları detaylı bir şekilde incelenmektedir. Oluşturulan grafiklerde yatay eksen test videosunun kare indeksleri, dikey eksen ise düzenlilik puanları yer almaktadır. Dikey kırmızı çizgiler, metodolojimizin belirlediği anormal olayların odak noktalarını gösterirken, kırmızı ile işaretlenen alanlar ise bu anormal hareketlerin başlangıç ve bitiş noktalarını belirtir. Kırmızı alanların kapsadığı bölgelerdeki kareler, anormal olarak sınıflandırılmıştır.

Şekil 6.4'te sahnede koşan bir çocuğun yaptığı hareketi metodolojimiz anormal olarak tespit etmiştir. Ancak veri kümesinin gerçek etiketleme değerine göre burada herhangi bir anormallik söz konusu değildir. Bir diğer sahnede ise bir yetişkinin koşması da aynı şekilde metodoloji tarafından anormal bir hareket olarak değerlendirilmiştir. Bu hareket ise veri kümesinin gerçek etiketleme değerine göre anormallik içermektedir. Burada veri kümesine göre yetişkinin koşması anormallik olarak değerlendirilirken çocuğun koşması anormallik olarak değerlendirilmemektedir. Bu sebeple uygulama sonucunda çocuğun koşması hatalı pozitif olarak değerlendirilmiştir.

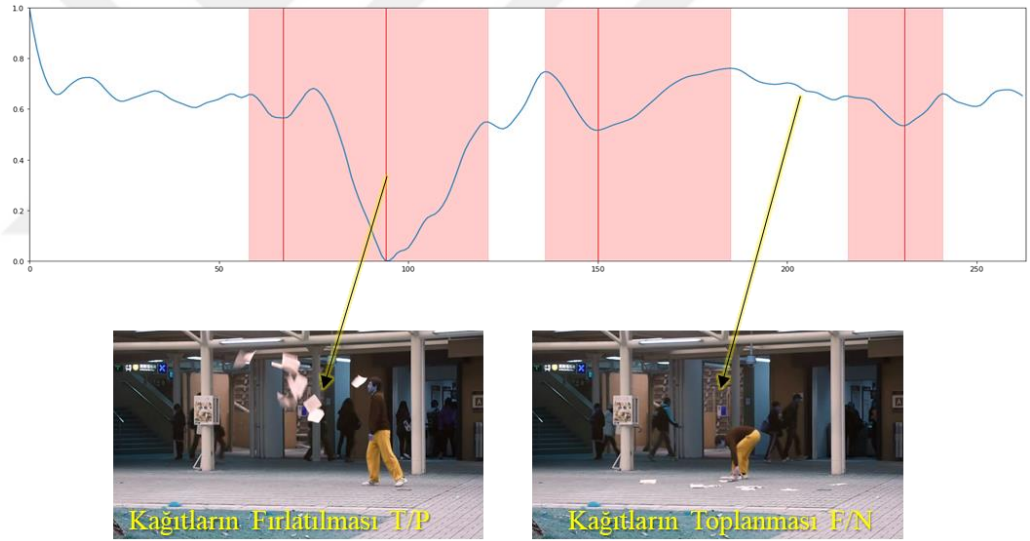


Şekil 6.4 CUHK Avenue veri kümesindeki uygulama sonuçları (klip no 1)

Şekil 6.5'teki başka bir test videosunda ise ters yönde ilerleyen ve bisikletle hareket eden kişilerin oluşturduğu anormallik metodoloji tarafından doğru olarak tespit edilmiştir. Şekil 6.6'da ise sahne ortasına gelerek elindeki kağıtları yukarıya atan bir kişinin oluşturduğu anormallik gösterilmektedir. Burada kağıtların fırlatılış hareketinden kaynaklanan anormallik metodolojimiz tarafından doğru bir şekilde tespit edilmiştir. Ancak, kağıtların yerden toplanması sırasında veri kümesi tarafından anormal olarak nitelendirilen bazı kareler metodoloji tarafından normal olarak değerlendirilmiştir. Burada hatalı negatif tespit söz konusudur.



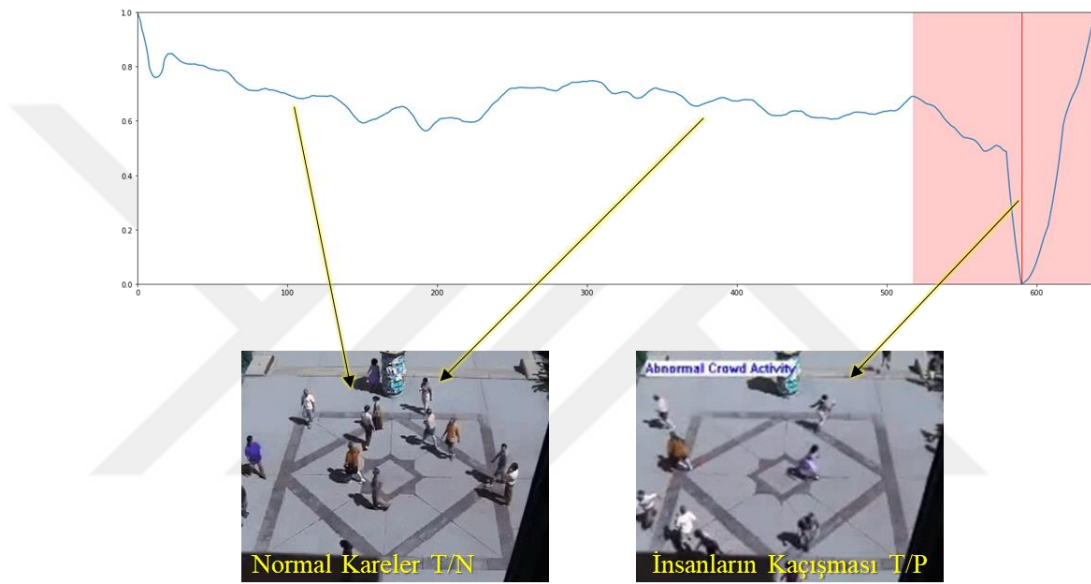
Şekil 6.5 CUHK Avenue veri kümesindeki uygulama sonuçları (klip no 16)



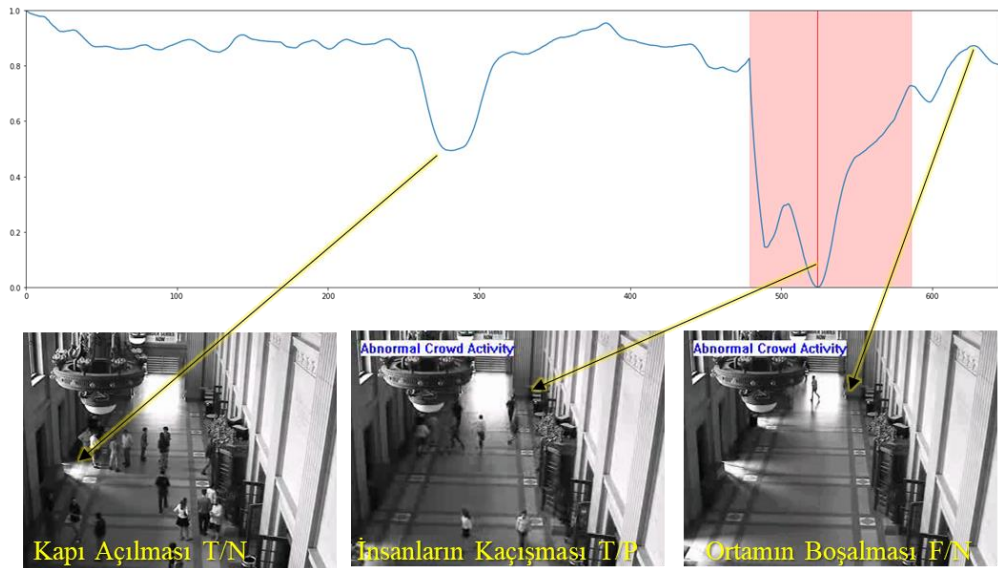
Şekil 6.6 CUHK Avenue veri kümesindeki uygulama sonuçları (klip no 20)

Metodolojimizi değerlendirdiğimiz UMN veri kümesinin açık ortamda çekilmiş ilk test videosu grafiği Şekil 6.7’de gösterilmektedir. Buradaki ilk bölümde insanlar normal bir şekilde yürüyerek gezinmektedir. Burada herhangi bir anormallik tespiti bulunmamaktadır. Ancak insanların bir anda kaçışmaya başlaması ile düzenlilik puanı hızlı bir şekilde bozulmaktadır. İnsanların kaçışması anormalliği doğru bir şekilde tespit edilmiştir. Grafiği Şekil 6.8’de verilen diğer test videosu ise nispeten karanlık ve kapalı ortamda sahnelenmiştir. Bu test esnasında ilk olarak kapının dışarıya açılması ile içeriye

yoğun ışık girmektedir. Ortamdaki ışık miktarının artması neticesinde düzenlilik puanı düşmektedir. Ancak bu düşüşün kalıcık değeri yüksek olmadığından modelimiz dayanmaktadır. Videonun devamında, insanların panikle kaçışması anormalliği metodolojimiz tarafından doğru bir şekilde tespit edilmektedir. Ancak videonun sonlarına doğru sayıca azalan kişilerin normal olarak yürümesi metodolojimiz tarafından anormal olarak nitelendirilmemiştir. Burada veri kümesinin gerçek etiket değeri anormal olarak verildiğinden yapılan tespit hatalı negatif olarak değerlendirilmiştir.

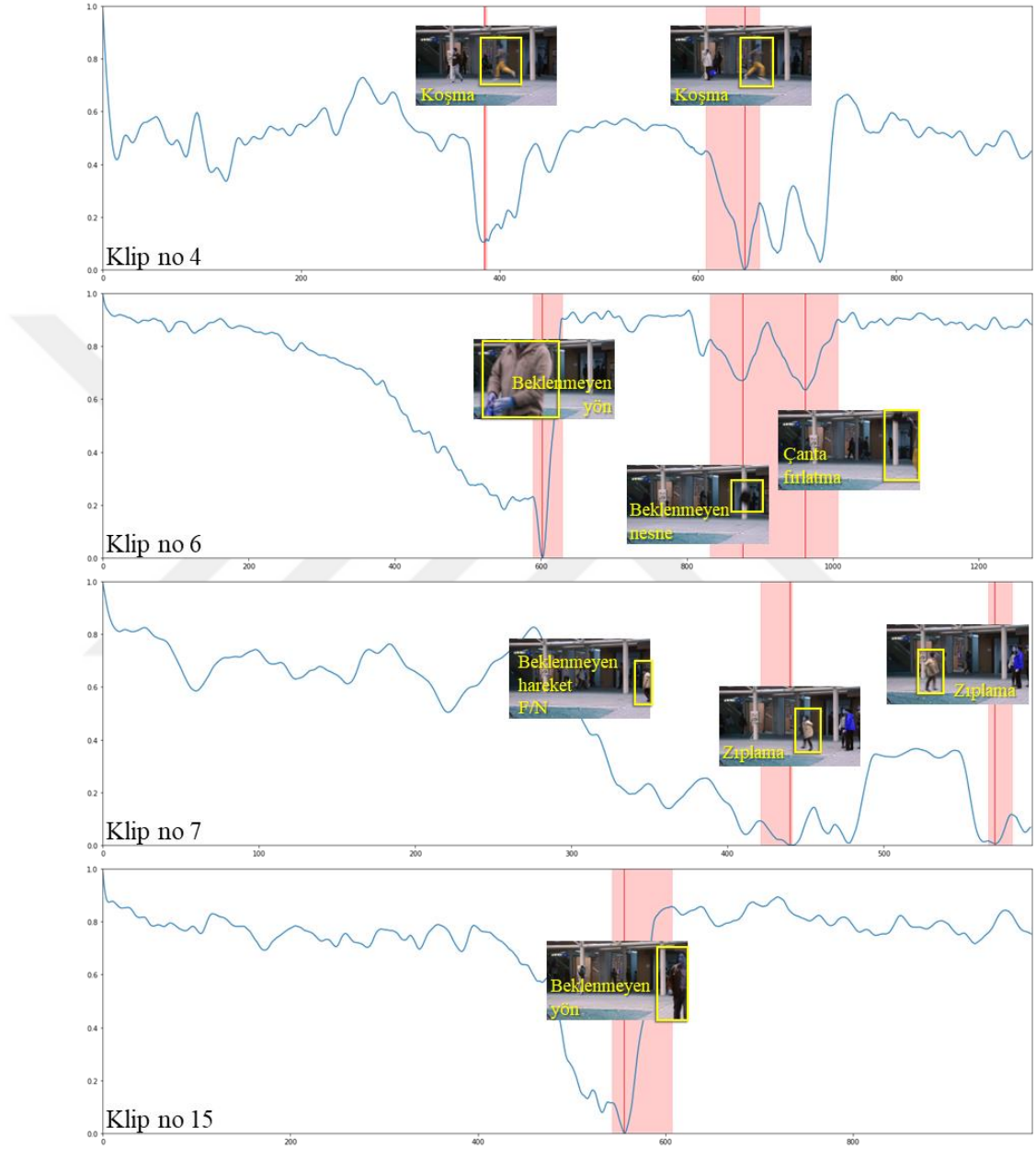


Şekil 6.7 UMN veri kümesindeki uygulama sonuçları (klip no 1)

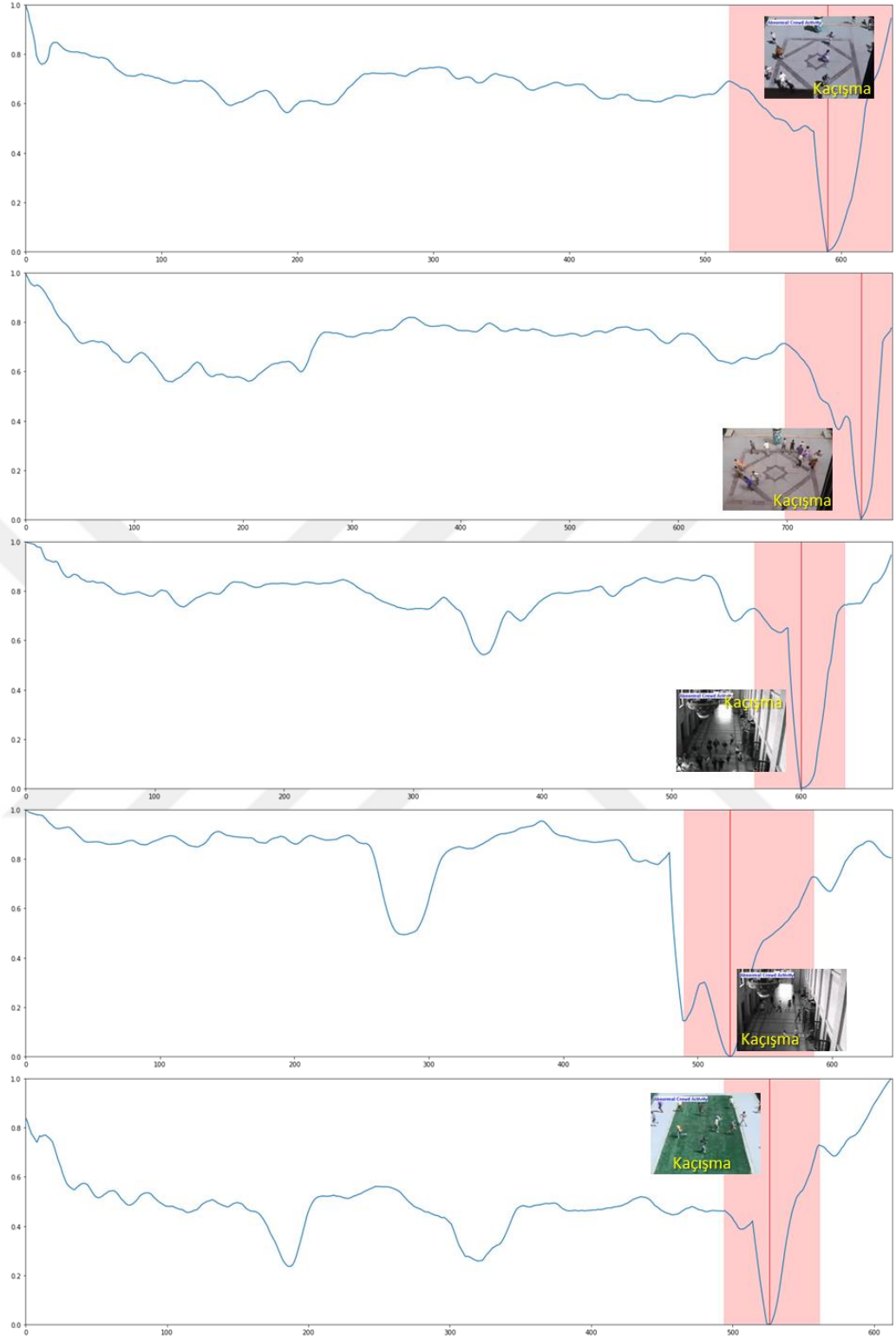


Şekil 6.8 UMN veri kümesindeki uygulama sonuçları (klip no 6)

Diğer CUHK Avenue veri kümesi örnek uygulamaları Şekil 6.9’da ve UMN veri kümesi örnek uygulamaları Şekil 6.10’da sunulmuş ve metodoloji tarafından yapılan tespitler grafiklerle görselleştirilmiştir.



Şekil 6.9 CUHK Avenue veri kümesindeki uygulama sonuçları



Şekil 6.10 UMN veri kümesindeki uygulama sonuçları

Bu grafikler, modelin veri kümesindeki anormallikleri nasıl algıladığını ve belirlediğini göstermektedir. Ayrıca, düzenlilik puanlarındaki değişimler, modelin normal ve

anormal davranışları nasıl ayırt ettiğini ve hangi noktalarda anormalliklerin yoğunlaştığını ortaya koymaktadır. Elde edilen sonuçlar, önerilen metodolojinin veri kümesindeki anormal davranışları tespit etme kabiliyetinin yanı sıra, bu tespitlerin zamanlamasının başarılı bir şekilde yapabildiğini göstermektedir.

Metodolojimizin uygulanması sırasında elde edilen değerlendirme metrikleri, Çizelge 6.1'de ayrıntılı bir şekilde sunulmuştur. Bu metrikler, modelin performansını kapsamlı bir şekilde değerlendirmek için kullanılan önemli ölçütlerdir ve modelin anormallik tespit kabiliyetini somut bir şekilde göstermektedir.

Çizelge 6.1 CUHK Avenue ve UMN veri kümeleri için kullanılan değerlendirme metrikleri

Değerlendirme Metrikleri	CUHK Avenue	UMN
Doğruluk	0,9024	0,9624
F1 Skoru	0,8253	0,8827
Kesinlik	0,7399	0.8576
Duyarlılık	0.9331	0.9094

Modelimiz, CUHK Avenue veri kümesinde gerçekleştirilen testlerde %91,2 ve UMN veri kümesinde %94,1 gibi dikkate değer AUC metrikleriyle başarılı bir performans sergilemiştir. Çizelge 6.2, CUHK Avenue veri kümesindeki test sonuçlarını, Çizelge 6.3 ise UMN veri kümesindeki test sonuçlarını sunmakta ve metodolojimizin kare düzeyindeki başarısını ortaya koymaktadır.

Bu yüksek AUC değerleri, literatürdeki benzer çalışmalarla kıyaslandığında, önerilen metodolojinin mevcut yaklaşımlara göre önemli bir iyileşme sağladığını göstermektedir. Elde edilen bu başarı özellik çıkarımı yetenekleri ve etkin anormallik tespit yöntemi sayesinde mümkün olmuştur. Ayrıca, modelin karmaşık sahneler ve değişken koşullar altında dâhil güvenilir sonuçlar sunabilmesi, pratik uygulamalar için büyük önem taşımaktadır.

Çizelge 6.2 CUHK Avenue veri kümesinde önerdiğimiz metodolojinin kare düzeyinde performansının (AUC) diğer yöntemlerle karşılaştırılması

Metot	AUC (%)
ConvLSTM-AE (Luo vd. 2017b)	77,0
Conv-AE (Hasan vd. 2016)	80,0
3DConv (Zhao vd. 2017)	80,9
StackedRNN (Luo vd. 2017a)	81,7
MemAE (Gong vd. 2019)	83,3
Frame prediction (Liu vd. 2018)	85,1
Pred+Recon (Y. Tang vd. 2020)	85,1
CT-D2GAN (X. Feng vd. 2021)	85,9
ClusterAE (Chang vd. 2020)	86,0
DESD-SVDD (Zhong vd. 2022)	87,7
MemNormality (Park vd. 2020)	88,5
VEC-AM (Yu vd. 2020)	89,6
WiderResNet (Le ve Kim 2022)	86,7
Multi-Task (M.-I. Georgescu vd. 2021)	86,9
Object-centric auto-encoders (Ionescu vd. 2019)	90,4
Önerilen	91,2

Çizelge 6.3 UMN veri kümesinde önerdiğimiz metodolojinin kare düzeyinde performansının (AUC) diğer yöntemlerle karşılaştırılması

Metot	AUC (%)
GLCM (Lloyd vd. 2017)	82,1
Optical Flow (Mehran vd. 2010)	84,0
Single Scene (Marsden vd. 2016)	92,9
Social Force (Mehran vd. 2010)	96,0
MHOF (Cong vd. 2013)	97,8
Dense Trajectories (K. Ma vd. 2015)	98,0
GAN (Alafif vd. 2022)	98,1
Chaotic (S. Wu vd. 2010)	99,0
HOOG (X. Hu vd. 2018)	99,3
Anomaly Detector (W. Li vd. 2014)	99,5
PSO-SF (Raghavendra vd. 2011)	99,6
AED-Net (T. Wang vd. 2019)	99,7
Hybrid Classifiers (Alafif vd. 2023)	99,7
Önerilen	94,1

6.4 Değerlendirme

Önerilen metodolojinin uygulamasının sonuçları incelendiğinde, ön işleme aşamasında kullanılan arka plan çıkarımının hatalı pozitif sayısını azaltmada önemli bir katkı sağladığı görülmüştür. Bu iyileştirme, genel sistemin hassasiyetini artırmakta ve anormal olayların algılanmasında daha güvenilir sonuçlar elde edilmesine olanak tanımaktadır.

Metodolojinin, diğer çalışmalara istinaden, anormal hareketin başlangıç ve bitiş zamanlarını yaklaşık olarak belirleme yeteneği önemli bir yere sahiptir. Bu yaklaşım, anormal hareketlerin sadece varlığını değil, aynı zamanda zamanlamasını da dikkate alarak, daha detaylı bir analiz imkânı sunmaktadır. Örneğin, bir kişinin elindeki kağıtları fırlatması eyleminde, kişinin elini yukarı kaldırmasının mı yoksa kağıtların uçuşmasının mı anormal hareketin başlangıcı olarak belirleneceği hareketin doğasına ve bağlamına bağlı olarak değişkenlik gösterebilir. Bu, anormal hareket tespitinde göreceli değerlendirmelerin önemini vurgulamaktadır.

CUHK Avenue veri kümesinde çeşitli anormal hareket örnekleri bulunmaktadır. Bu örnekler arasında koşmak, dans etmek, cisimler fırlatmak yer almaktadır. Bunun yanı sıra, kameranın önünden geçen ve sahneyi tamamen kapatan engeller de mevcuttur. Bu çeşitlilik, Avenue veri kümesini VGAT sistemleri için zorlu bir test ortamı haline getirmektedir. Bu durumlar, genel sahne algısını ve anormal hareket tespitini etkileyebilecek faktörler olarak ön plana çıkmaktadır. UMN veri kümesi değerlendirildiğinde ise, daha homojen bir yapı gözlemlenmekte ve anormal hareketlerin temelinde genellikle insanların panik halinde kaçması bulunmaktadır. Bu homojenlik, metodolojinin UMN veri kümesinde daha yüksek bir başarı oranı elde etmesini sağlamaktadır. Ancak, her iki veri kümesi karşılaştırıldığında, Avenue'daki anomali çeşitliliğinin ve engelleyici faktörlerin fazlalığı nedeniyle, UMN veri kümesinde elde edilen başarı seviyesine ulaşmanın zor olduğu görülmektedir. Bu durum, Avenue'daki anormal hareketlerin tespitinde daha kapsamlı ve çeşitli yaklaşımların gerekliliğini vurgular.

UMN veri kümesindeki iç ortam sahnelerinde, kapılar zaman zaman açılmakta ve gün ışığı karelerin farklı bölümlerini aydınlatmaktadır. Karelerdeki bu tür ani ışık değişiklikleri, sahnenin yapısını ve algılanan hareket dinamiklerini etkileyerek, anormal hareket tespitinde hatalara yol açmaktadır. İç ortamlarda yaşanan bu gibi ışık değişimleri, metodolojinin hassasiyetini sınamaktadır. Bu durum, ışık değişimlerine karşı daha dayanıklı bir sistem geliştirmenin ve performansı optimize etmenin önemini göstermektedir. Ani ışık değişikliklerinin etkilerini azaltmak ve güvenilirliği artırmak için ek önlemler almak, önemli bir araştırma ve geliştirme alanıdır.

Yapılan testlerde, anormal hareketliliklerin başlangıç ve bitiş noktalarının belirlenmesi konusunda farklılıklar görülmektedir. Gerçek etiket değerleri verilse bile, bu noktaların tanımlanması modelin performansını önemli ölçüde etkileyebilir. Anormal davranışların doğru sınırlarının belirlenmesi, modelin doğruluğunu ve güvenilirliğini artırırken, yanlış sınırların belirlenmesi modelin yanlış oranın artırabilmektedir. Bu durum, VGAT için kullanılan modellerin eğitimi ve değerlendirilmesinde dikkate alınması gereken kritik bir husustur.

7. SONUÇ

Bu tez, VGAT alanındaki literatürü ve önceki çalışmaları titizlikle inceleyerek, bu alandaki gelişmeleri derinlemesine ele almaktadır. Gerçekleştirilen bu analiz, VGAT alanında karşılaşılan temel zorlukları, mevcut uygulama alanlarını ve gelecekteki geliştirme potansiyellerini detaylı bir şekilde ortaya koymaktadır. Araştırma ve bulguların ışığında, video verilerinde ön işleme, özellik çıkarımı ve anormallik tespiti süreçlerini içeren özgün bir metodoloji önerilmektedir. Bu metodolojinin tasarımı ve model geliştirme süreçleri, etkinliği ve uygulanabilirliğini artırmayı hedeflemektedir.

Önerilen metodolojinin ön işleme aşaması, video verilerinin standardize edilmesi ve gereksiz bilgilerin filtrelmesi, modelin daha hızlı ve etkili bir şekilde eğitilmesine ve test edilmesine imkân tanır. Gürültü azaltma ve arka plan çıkarmanın etkili uygulanması, verideki istenmeyen ya da rastlantısal dalgalanmaları azaltarak hatalı pozitif oranını düşürmekte ve gerçek anormal durumların daha net bir şekilde belirlenmesine yardımcı olarak, VGAT sistemlerinin güvenilirliğini önemli ölçüde artırmaktadır.

Geleneksel VGAT sistemlerinin çoğu, yüzeysel veya sınırlı özellik çıkarımı tekniklerine dayanırken, bu tezde önerilen özellik öğrenme yaklaşımı, karmaşık ve dinamik sahnelerde, anormal hareketlerin daha doğru bir şekilde tespit edilmesine olanak tanıyan derin öğrenme modeline dayanmaktadır. Özellikle, ConvLSTM-AE gibi gelişmiş model yapıları kullanılarak, video verilerinin zaman ve mekân boyutlarında derinlemesine analiz edilmesi sağlanmaktadır. Bu, VGAT sistemlerinin daha önce göz ardı edilen ince detayları ve düşük seviye özellikleri yakalamasına imkân tanır. Bu yaklaşım, aynı zamanda VGAT sistemlerinin esnekliğini ve uyarlanabilirliğini artırmaktadır. Geniş ve çeşitlendirilmiş veri kümelerinde, özellikle değişken aydınlatma, hava koşulları veya arka plan hareketliliği olan sahnelerde, sistemlerin doğruluğunu ve güvenilirliğini artırır.

Önerilen metodolojinin son aşamasında, model tarafından üretilen düzenlilik puanlarının kullanılarak 3 adımda anormal karelerin tespit edilmesi sağlanmaktadır. Bu

süreç ilk adım olarak yerel minima ve maksima noktalarının tespit edilmesi ile başlamaktadır. Sonraki adımda kalıcılık değerlerine göre anlamlı yerel minima noktalarının tespit edilerek anormal hareketlerin odak noktalarının tespit edilmesi ve üçüncü adımda ise yerel maksima noktaları ile olan uzaklıklar kullanılarak hareketin başlangıç ve bitiş noktalarının bulunması işlemleri gerçekleştirilmektedir. Söz konusu anormal hareketlerin başlangıç ve bitiş zamanlarının etkin bir şekilde tahmin edilebilmesi, mevcut çalışmaların çoğunda yer alan ve yalnızca anormal kareleri tanımlama yeteneğine kıyasla önemli bir katkı sağlamaktadır.

Gerçekleştirilen testler, sunulan metodolojinin CUHK Avenue veri kümesinde %91,2 ve UMN veri kümesinde %94,1 AUC ile yüksek doğruluk seviyelerine ulaşarak bu alanda etkili sonuçlar sunduğunu göstermektedir.

Bu tez, VGAT teknolojilerinin geliştirilmesine yönelik gelecekteki araştırmacılara ve geliştiricilere yol gösterici bir kaynak oluştururken, güvenlik gözetim ve ilgili sektörlerdeki profesyonellere de değerli bir referans sağlamaktadır. Önerilen metodolojinin öğrenme kapasitesi, farklı gözetim senaryolarına kolaylıkla uyarlanabilmesini ve geniş bir yelpazede adaptasyonunu mümkün kılmaktadır. Bu esneklik, VGAT teknolojilerinin çeşitli durumlarda daha başarılı sonuçlar elde etmesini sağlayarak, uygulama alanlarının genişlemesine katkıda bulunmaktadır.

Önerilen metodolojinin ışık değişimlerine karşı hassasiyetini artırmak ve bu tür durumlarda performansını optimize etmek gerekliliğini gösteren bulguları, gelecekteki araştırmalar için önemli bir gelişme alanı olarak belirlenmiştir. Özellikle UMN veri kümesinin kapalı alanlarındaki uygulamalarda gözlemlenen, kapıların açılmasıyla birlikte içeriye giren ışığın görsel sahneyi büyük oranda değiştirmesi ve bunun metodoloji başarısını düşürmesi, metodolojinin bu tür ışık değişimlerine karşı daha dirençli hale getirilmesi gerektiğini vurgulamaktadır.

Geliştirilen metodolojilerin ve yöntemlerin büyük veri setleri üzerindeki performansının optimize edilmesi, hareketli ve çoklu kamera senaryolarında anormal hareketlerin daha etkin tespiti ve karmaşık gözetim ortamlarında modelin kullanılabilirliğinin artırılması,

gelecek arařtırmaların odaklanması gereken bařlıca konular arasında yer almaktadır. Ayrıca, modelin yorumlanabilirliđinin artırılması ve anormallikleri tespit ederken öđrendiđi özelliklerin arkasındaki nedenlerin daha iyi anlařılması bu alandaki arařtırma konularından bir diđeridir.

Teknolojinin de ilerlemesiyle birlikte, daha etkili tespit yöntemleri, gelişmiş donanım platformları ve yazılım araçları aracılığıyla algoritmaların performansını maksimize edecek gelişmeler de bu alandaki önemli yenilikler arasında yer alacaktır. Bu gelişmeler, VGAT alanındaki ilerlemeye önemli katkılar sağlayacak ve modelin karar verme süreçlerinin daha etkin bir şekilde yorumlanması için yenilikçi yaklaşımların ortaya çıkmasına zemin hazırlayacaktır.

KAYNAKLAR

- Abdelghafour, M., ElBery, M., & Taha, Z. (2021). Comparative Study for Anomaly Detection in Crowded Scenes. *International Journal of Intelligent Computing and Information Sciences*, 21(3), 84-94.
- Acsintoae, A., Florescu, A., Georgescu, M.-I., Mare, T., Sumedrea, P., Ionescu, R. T., Khan, F. S., & Shah, M. (2022). UBnormal: New Benchmark for Supervised Open-Set Video Anomaly Detection. *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 20143-20153.
- Adam, A., Rivlin, E., Shimshoni, I., & Reinitz, D. (2008). Robust real-time unusual event detection using multiple fixed-location monitors. *IEEE transactions on pattern analysis and machine intelligence*, 30(3), 555-560.
- Alafif, T., Alzahrani, B., Cao, Y., Alotaibi, R., Barnawi, A., & Chen, M. (2022). Generative adversarial network based abnormal behavior detection in massive crowd videos: a Hajj case study. *Journal of Ambient Intelligence and Humanized Computing*, 13(8), 4077-4088. <https://doi.org/10.1007/S12652-021-03323-5>
- Alafif, T., Hadi, A., Allahyani, M., Alzahrani, B., Alhothali, A., Alotaibi, R., & Barnawi, A. (2023). Hybrid Classifiers for Spatio-Temporal Abnormal Behavior Detection, Tracking, and Recognition in Massive Hajj Crowds. *Electronics* 2023, Vol. 12, Page 1165, 12(5), 1165. <https://doi.org/10.3390/ELECTRONICS12051165>
- Aldayri, A., & Albattah, W. (2022). Taxonomy of Anomaly Detection Techniques in Crowd Scenes. *Sensors*, 22(16), 6080.

- Alghushairy, O., Alsini, R., Soule, T., & Ma, X. (2020). A Review of Local Outlier Factor Algorithms for Outlier Detection in Big Data Streams. *Big Data Cogn. Comput.*, 5(1), 1-24. <https://doi.org/10.3390/BDCC5010001>
- Bansod, S. D., & Nandedkar, A. V. (2020). Crowd anomaly detection and localization using histogram of magnitude and momentum. *The Visual Computer*, 36(3), 609-620.
- Barlow, H. B. (1989). Unsupervised Learning. *Neural Computation*, 1(3), 295-311. <https://doi.org/10.1162/NECO.1989.1.3.295>
- Barnett, V., & Lewis, T. (1984). Outliers in statistical data. *Wiley Series in Probability and Mathematical Statistics. Applied Probability and Statistics*.
- Baskurt, K. B., & Samet, R. (2019). Video synopsis: A survey. *Computer Vision and Image Understanding*, 181, 26-38.
- BAŞKURT, K. B., & Samet, R. (2018). Long-term multiobject tracking using alternative correlation filters. *Turkish Journal of Electrical Engineering and Computer Sciences*, 26(5), 2246-2259.
- Boekhoudt, K., Matei, A., Aghaei, M., & Talavera, E. (2021). HR-Crime: Human-Related Anomaly Detection in Surveillance Videos. *International Conference on Computer Analysis of Images and Patterns*, 164-174.
- Bouman, C. A., Shapiro, M., Ncsa, ;, Cook, G. W., Cheng, H., & Borman, S. (1995). *CLUSTER: An Unsupervised Algorithm for Modeling Gaussian Mixtures*. <http://www.ece.purdue.edu/~bouman>
- Breiman, L. (2001). Random forests. *Machine Learning*, 45(1), 5-32. <https://doi.org/10.1023/A:1010933404324/METRICS>

- Cao, C., Zhang, X., Zhang, S., Wang, P., & Zhang, Y. (2022). Adaptive graph convolutional networks for weakly supervised anomaly detection in videos. *arXiv preprint arXiv:2202.06503*.
- Chandola, V., Banerjee, A., & Kumar, V. (2009). Anomaly detection: A survey. *ACM computing surveys (CSUR)*, 41(3), 1-58.
- Chang, Y., Tu, Z., Xie, W., & Yuan, J. (2020). Clustering driven deep autoencoder for video anomaly detection. *European Conference on Computer Vision*, 329-345.
- Chen, C. H. (2015). *Handbook of pattern recognition and computer vision*. World Scientific.
- Chen, Y., Liu, Z., Zhang, B., Fok, W., Qi, X., & Wu, Y.-C. (2022). MGFN: Magnitude-Contrastive Glance-and-Focus Network for Weakly-Supervised Video Anomaly Detection. *arXiv preprint arXiv:2211.15098*.
- Cheng, H., Liu, X., Wang, H., Fang, Y., Wang, M., & Zhao, X. (2020). SecureAD: A secure video anomaly detection framework on convolutional neural network in edge computing environment. *IEEE Transactions on Cloud Computing*.
- Chung, J., Gulcehre, C., Cho, K., & Bengio, Y. (2014). *Empirical Evaluation of Gated Recurrent Neural Networks on Sequence Modeling*. <https://arxiv.org/abs/1412.3555v1>
- Colque, R. V. H. M., & others. (2018). *Robust approaches for anomaly detection applied to video surveillance*.
- Cong, Y., Yuan, J., & Liu, J. (2013). Abnormal event detection in crowded scenes using sparse representation. *Pattern Recognition*, 46(7), 1851-1864. <https://doi.org/10.1016/J.PATCOG.2012.11.021>

- Cunningham, P., Cord, M., & Delany, S. J. (2008). Supervised learning. *Cognitive Technologies*, 21-49. https://doi.org/10.1007/978-3-540-75171-7_2/COVER
- Direkoglu, C., Sah, M., & O'Connor, N. E. (2017). Abnormal crowd behavior detection using novel optical flow-based features. *2017 14th IEEE International Conference on Advanced Video and Signal Based Surveillance (AVSS)*, 1-6.
- Doersch, C. (2016). *Tutorial on Variational Autoencoders*. <https://arxiv.org/abs/1606.05908v3>
- Ehsani, N., Aminifar, F., & Mohsenian-Rad, H. (2022). Convolutional autoencoder anomaly detection and classification based on distribution PMU measurements. *IET Generation, Transmission & Distribution*, 16(14), 2816-2828.
- Extracting and Filtering Minima and Maxima of 1D Functions*. (t.y.). Geliş tarihi 17 Ekim 2023, gönderen <https://www.csc.kth.se/~weinkauf/notes/persistence1d.html>
- Feng, J.-C., Hong, F.-T., & Zheng, W.-S. (2021). Mist: Multiple instance self-training framework for video anomaly detection. *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, 14009-14018.
- Feng, X., Song, D., Chen, Y., Chen, Z., Ni, J., & Chen, H. (2021). Convolutional transformer based dual discriminator generative adversarial networks for video anomaly detection. *Proceedings of the 29th ACM International Conference on Multimedia*, 5546-5554.
- Georgescu, M. I., Ionescu, R. T., Khan, F. S., Popescu, M., & Shah, M. (2021). A background-agnostic framework with adversarial training for abnormal event detection in video. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 44(9), 4505-4523.

- Georgescu, M.-I., Barbalau, A., Ionescu, R. T., Khan, F. S., Popescu, M., & Shah, M. (2021). Anomaly detection in video via self-supervised and multi-task learning. *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, 12742-12752.
- Gers, F. A., & Schmidhuber, J. (2000). Recurrent nets that time and count. *Proceedings of the International Joint Conference on Neural Networks*, 3, 189-194. <https://doi.org/10.1109/IJCNN.2000.861302>
- Gong, D., Liu, L., Le, V., Saha, B., Mansour, M. R., Venkatesh, S., & Hengel, A. van den. (2019). Memorizing normality to detect anomaly: Memory-augmented deep autoencoder for unsupervised anomaly detection. *Proceedings of the IEEE/CVF International Conference on Computer Vision*, 1705-1714.
- Goodfellow, I., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., Courville, A., & Bengio, Y. (2014). Generative Adversarial Networks. *Communications of the ACM*, 63(11), 139-144. <https://doi.org/10.1145/3422622>
- Grubbs, F. E. (1969). Procedures for detecting outlying observations in samples. *Technometrics*, 11(1), 1-21.
- Gu, X., Cui, J., & Zhu, Q. (2014). Abnormal crowd behavior detection by using the particle entropy. *Optik*, 125(14), 3428-3433.
- Hasan, M., Choi, J., Neumann, J., Roy-Chowdhury, A. K., & Davis, L. S. (2016). Learning temporal regularity in video sequences. *Proceedings of the IEEE conference on computer vision and pattern recognition*, 733-742.
- Hawkins, D. M. (1980). *Identification of outliers* (C. 11). Springer.

Hirschorn, O., & Avidan, S. (2022). Normalizing Flows for Human Pose Anomaly Detection. *arXiv preprint arXiv:2211.10946*.

Horn, B. K. P., & Schunck, B. G. (1981). Determining optical flow. *Artificial Intelligence*, 17(1-3), 185-203. [https://doi.org/10.1016/0004-3702\(81\)90024-2](https://doi.org/10.1016/0004-3702(81)90024-2)

How to deal with vanishing and exploding gradients | by Rahul Pandey | Geek Culture | Medium. (t.y.). Geliş tarihi 07 Ağustos 2023, gönderen <https://medium.com/geekculture/how-to-deal-with-vanishing-and-exploding-gradients-in-neural-networks-24eb00c80e84>

Hu, J., Wang, X., Zhang, Y., Zhang, D., Zhang, M., & Xue, J. (2020). Time series prediction method based on variant LSTM recurrent neural network. *Neural Processing Letters*, 52, 1485-1500.

Hu, X., Huang, Y., Duan, Q., Ci, W., Dai, J., & Yang, H. (2018). Abnormal event detection in crowded scenes using histogram of oriented contextual gradient descriptor. *Eurasip Journal on Advances in Signal Processing*, 2018(1), 1-15. <https://doi.org/10.1186/S13634-018-0574-4/FIGURES/12>

Ionescu, R. T., Khan, F. S., Georgescu, M.-I., & Shao, L. (2019). Object-centric auto-encoders and dummy anomalies for abnormal event detection in video. *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 7842-7851.

Isupova, O., Kuzin, D., & Mihaylova, L. (2016). Anomaly detection in video with Bayesian nonparametrics. *arXiv preprint arXiv:1606.08455*.

Jain, L. C., & Medsker, L. (1999). *Recurrent Neural Networks: Design and Applications*. <https://doi.org/10.1201/9781420049176>

- Jiang, T., Li, Y., Xie, W., & Du, Q. (2020). Discriminative reconstruction constrained generative adversarial network for hyperspectral anomaly detection. *IEEE Transactions on Geoscience and Remote Sensing*, 58(7), 4666-4679.
- Kaltsa, V., Briassouli, A., Kompatsiaris, I., Hadjileontiadis, L. J., & Strintzis, M. G. (2015). Swarm intelligence for detecting interesting events in crowded environments. *IEEE transactions on image processing*, 24(7), 2153-2166.
- Khaire, P., & Kumar, P. (2022). A semi-supervised deep learning based video anomaly detection framework using RGB-D for surveillance of real-world critical environments. *Forensic Science International: Digital Investigation*, 40, 301346.
- Le, V.-T., & Kim, Y.-G. (2022). Attention-based residual autoencoder for video anomaly detection. *Applied Intelligence*, 1-15.
- Lecun, Y. (1987). *PhD thesis: Modeles connexionnistes de l'apprentissage (connectionist learning models)*. Universite P. et M. Curie (Paris 6). <https://nyuscholars.nyu.edu/en/publications/phd-thesis-modeles-connexionnistes-de-lapprentissage-connectionis>
- Lee, J., Nam, W.-J., & Lee, S.-W. (2022). Multi-Contextual Predictions with Vision Transformer for Video Anomaly Detection. *2022 26th International Conference on Pattern Recognition (ICPR)*, 1012-1018.
- Lee, Y., & Kang, P. (2022). AnoViT: Unsupervised Anomaly Detection and Localization with Vision Transformer-based Encoder-Decoder. *IEEE Access*, 10, 46717-46724. <https://doi.org/10.1109/ACCESS.2022.3171559>
- Li, W., Mahadevan, V., & Vasconcelos, N. (2014). Anomaly detection and localization in crowded scenes. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 36(1), 18-32. <https://doi.org/10.1109/TPAMI.2013.111>

- Li, Y., Guo, T., Xia, R., & Xie, W. (2018). Road traffic anomaly detection based on fuzzy theory. *IEEE Access*, 6, 40281-40288.
- Li, Z., Li, Y., & Gao, Z. (2020). Spatiotemporal representation learning for video anomaly detection. *IEEE Access*, 8, 25531-25542.
- Liu, W., Cao, J., Zhu, Y., Liu, B., & Zhu, X. (2022). Real-time anomaly detection on surveillance video with two-stream spatio-temporal generative model. *Multimedia Systems*, 1-13.
- Liu, W., Luo, W., Lian, D., & Gao, S. (2018). Future frame prediction for anomaly detection—a new baseline. *Proceedings of the IEEE conference on computer vision and pattern recognition*, 6536-6545.
- Lloyd, K., Rosin, P. L., Marshall, D., & Moore, S. C. (2017). Detecting violent and abnormal crowd activity using temporal analysis of grey level co-occurrence matrix (GLCM)-based texture measures. *Machine Vision and Applications*, 28(3-4), 361-371. <https://doi.org/10.1007/S00138-017-0830-X/TABLES/6>
- Lu, C., Shi, J., & Jia, J. (2013). Abnormal event detection at 150 fps in matlab. *Proceedings of the IEEE international conference on computer vision*, 2720-2727.
- Luo, W., Liu, W., & Gao, S. (2017a). A revisit of sparse coding based anomaly detection in stacked rnn framework. *Proceedings of the IEEE international conference on computer vision*, 341-349.
- Luo, W., Liu, W., & Gao, S. (2017b). Remembering history with convolutional lstm for anomaly detection. *2017 IEEE International Conference on Multimedia and Expo (ICME)*, 439-444.

- Lv, H., Chen, C., Cui, Z., Xu, C., Li, Y., & Yang, J. (2021). Learning normal dynamics in videos with meta prototype network. *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 15425-15434.
- Ma, K., Doescher, M., & Bodden, C. (2015). *Anomaly Detection In Crowded Scenes Using Dense Trajectories*.
- Ma, Z., Machado, J. J. M., & Tavares, J. M. R. S. (2021). Weakly supervised video anomaly detection based on 3D convolution and LSTM. *Sensors*, 21(22), 7508.
- Madan, N., Ristea, N.-C., Ionescu, R. T., Nasrollahi, K., Khan, F. S., Moeslund, T. B., & Shah, M. (2022). Self-Supervised Masked Convolutional Transformer Block for Anomaly Detection. *arXiv preprint arXiv:2209.12148*.
- Mahadevan, V., Li, W., Bhalodia, V., & Vasconcelos, N. (2010). Anomaly detection in crowded scenes. *2010 IEEE computer society conference on computer vision and pattern recognition*, 1975-1981.
- Majhi, S., Das, S., Brémond, F., Dash, R., & Sa, P. K. (2021). Weakly-supervised joint anomaly detection and classification. *2021 16th IEEE International Conference on Automatic Face and Gesture Recognition (FG 2021)*, 1-7.
- Marsden, M., McGuinness, K., Little, S., & O'Connor, N. E. (2016). Holistic features for real-time crowd behaviour anomaly detection. *Proceedings - International Conference on Image Processing, ICIP, 2016-August*, 918-922. <https://doi.org/10.1109/ICIP.2016.7532491>
- Medel, J. R. (2016). *Anomaly detection using predictive convolutional long short-term memory units*. Rochester Institute of Technology.

- Mehran, R., Oyama, A., & Shah, M. (2010a). *Abnormal crowd behavior detection using social force model*. 935-942. <https://doi.org/10.1109/CVPR.2009.5206641>
- Mehran, R., Oyama, A., & Shah, M. (2010b). *Abnormal crowd behavior detection using social force model*. 935-942. <https://doi.org/10.1109/CVPR.2009.5206641>
- Michelucci, U. (2022). An introduction to autoencoders. *arXiv preprint arXiv:2201.03898*.
- Mnih, V., Heess, N., Graves, A., & Kavukcuoglu, K. (2014). Recurrent Models of Visual Attention. *Advances in Neural Information Processing Systems*, 3(January), 2204-2212. <https://arxiv.org/abs/1406.6247v1>
- Mo, X., Monga, V., Bala, R., & Fan, Z. (2013). Adaptive sparse representations for video anomaly detection. *IEEE Transactions on Circuits and Systems for Video Technology*, 24(4), 631-645.
- Munyua, J. G., Wambugu, G. M., & Njenga, S. T. (2021). A Survey of Deep Learning Solutions for Anomaly Detection in Surveillance Videos. *International Journal of Computer and Information Technology (2279-0764)*, 10(5).
- Nayak, R., Pati, U. C., & Das, S. K. (2021). A comprehensive review on deep learning-based methods for video anomaly detection. *Image and Vision Computing*, 106, 104078.
- Nguyen, V., Phung, D., Pham, D. S., & Venkatesh, S. (2015). Bayesian Nonparametric Approaches to Abnormality Detection in Video Surveillance. *Annals of Data Science*, 2(1), 21-41. <https://doi.org/10.1007/S40745-015-0030-3/FIGURES/8>
- Noble, W. S. (2006). What is a support vector machine? *Nature Biotechnology* 2006 24:12, 24(12), 1565-1567. <https://doi.org/10.1038/nbt1206-1565>

- Pannirselvam, P. M., Geetha, M. K., & Kumaravelan, G. (2021). A Comprehensive Study on Automated Anomaly Detection Techniques in Video Surveillance. *Annals of the Romanian Society for Cell Biology*, 4027-4037.
- Parate, M. R., Satpute, V. R., & Bhurchandi, K. M. (2018). Global-patch-hybrid template-based arbitrary object tracking with integral channel features. *Applied Intelligence*, 48(2), 300-314.
- Park, H., Noh, J., & Ham, B. (2020). Learning memory-guided normality for anomaly detection. *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 14372-14381.
- Quinlan, J. R. (1986). Induction of decision trees. *Machine Learning 1986 1:1*, 1(1), 81-106. <https://doi.org/10.1007/BF00116251>
- Raghavendra, R., Del Bue, A., Cristani, M., & Murino, V. (2011). Optimizing interaction force for global anomaly detection in crowded scenes. *Proceedings of the IEEE International Conference on Computer Vision*, 136-143. <https://doi.org/10.1109/ICCVW.2011.6130235>
- Ramachandra, B., Jones, M., & Vatsavai, R. (2020). Learning a distance function with a Siamese network to localize anomalies in videos. *Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision*, 2598-2607.
- Reiss, T., & Hoshen, Y. (2022). *Attribute-based Representations for Accurate and Interpretable Video Anomaly Detection*. arXiv. <https://doi.org/10.48550/ARXIV.2212.00789>
- Ren, J., Xia, F., Liu, Y., & Lee, I. (2021). Deep video anomaly detection: Opportunities and challenges. *2021 international conference on data mining workshops (ICDMW)*, 959-966.

- Rodrigues, R., Bhargava, N., Velmurugan, R., & Chaudhuri, S. (2020). Multi-timescale trajectory prediction for abnormal human activity detection. *Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision*, 2626-2634.
- Sabokrou, M., Fathy, M., Zhao, G., & Adeli, E. (2020). Deep end-to-end one-class classifier. *IEEE transactions on neural networks and learning systems*, 32(2), 675-684.
- Sabokrou, M., Fayyaz, M., Fathy, M., Moayed, Z., & Klette, R. (2018). Deep-anomaly: Fully convolutional neural network for fast anomaly detection in crowded scenes. *Computer Vision and Image Understanding*, 172, 88-97.
- Saleem, G., Bajwa, U. I., Raza, R. H., Alqahtani, F. H., Tolba, A., & Xia, F. (2022). Efficient anomaly recognition using surveillance videos. *PeerJ Computer Science*, 8, e1117.
- Santhosh, K. K., Dogra, D. P., & Roy, P. P. (2020). Anomaly detection in road traffic using visual surveillance: A survey. *ACM Computing Surveys (CSUR)*, 53(6), 1-26.
- Sapkota, H., & Yu, Q. (2022). Bayesian Nonparametric Submodular Video Partition for Robust Anomaly Detection. *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 3212-3221.
- Sarker, M. I., Losada-Gutiérrez, C., Marrón-Romera, M., Fuentes-Jiménez, D., & Luengo-Sánchez, S. (2021). Semi-supervised anomaly detection in video-surveillance scenes in the wild. *Sensors*, 21(12), 3993.
- Schapire, R. E. (2003). *The Boosting Approach to Machine Learning: An Overview*. 149-171. https://doi.org/10.1007/978-0-387-21579-2_9

- Schuster, M., & Paliwal, K. K. (1997). Bidirectional recurrent neural networks. *IEEE Transactions on Signal Processing*, 45(11), 2673-2681. <https://doi.org/10.1109/78.650093>
- Sultani, W., Chen, C., & Shah, M. (2018). Real-world anomaly detection in surveillance videos. *Proceedings of the IEEE conference on computer vision and pattern recognition*, 6479-6488.
- Şengönül, E., & Samet, R. (2021). The Effect of Background Subtraction on Abnormal Event Detection in Video Surveillance. *2021 29th Signal Processing and Communications Applications Conference (SIU)*, 1-4.
- Şengönül, E., Samet, R., Abu Al-Haija, Q., Alqahtani, A., Alturki, B., & Alsulami, A. A. (2023). An Analysis of Artificial Intelligence Techniques in Surveillance Video Anomaly Detection: A Comprehensive Survey. *Applied Sciences* 2023, Vol. 13, Page 4956, 13(8), 4956. <https://doi.org/10.3390/APP13084956>
- Tang, J., Chen, Z., Fu, A. W. C., & Cheung, D. W. (2002). Enhancing effectiveness of Outlier detections for low Density Patterns. *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 2336, 535-548. https://doi.org/10.1007/3-540-47887-6_53/COVER
- Tang, Y., Zhao, L., Zhang, S., Gong, C., Li, G., & Yang, J. (2020). Integrating prediction and reconstruction for anomaly detection. *Pattern Recognition Letters*, 129, 123-130.
- Tax, D. M. J., & Duin, R. P. W. (2001). Combining one-class classifiers. *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 2096, 299-308. https://doi.org/10.1007/3-540-48219-9_30/COVER

- Terrell, G. R., & Scott, D. W. (1992). Variable Kernel Density Estimation. <https://doi.org/10.1214/aos/1176348768>, 20(3), 1236-1265. <https://doi.org/10.1214/AOS/1176348768>
- Thakare, K., Raghuwanshi, Y., Dogra, D. P., Choi, H., & Kim, I.-J. (2022). DyAnNet: A Scene Dynamicity Guided Self-Trained Video Anomaly Detection Network. *arXiv preprint arXiv:2211.00882*.
- Tian, Y., Pang, G., Chen, Y., Singh, R., Verjans, J. W., & Carneiro, G. (2021). Weakly-supervised video anomaly detection with robust temporal feature magnitude learning. *Proceedings of the IEEE/CVF International Conference on Computer Vision*, 4975-4986.
- Ullah, W., Ullah, A., Haq, I. U., Muhammad, K., Sajjad, M., & Baik, S. W. (2021). CNN features with bi-directional LSTM for real-time anomaly detection in surveillance networks. *Multimedia Tools and Applications*, 80(11), 16979-16995.
- Understanding LSTM Networks -- colah's blog*. (t.y.). Geliş tarihi 07 Ağustos 2023, gönderen <http://colah.github.io/posts/2015-08-Understanding-LSTMs/>
- Van Houdt, G., Mosquera, C., & Nápoles, G. (2020). A review on the long short-term memory model. *Artificial Intelligence Review*, 53, 5929-5955.
- Vennerød, C. B., Kjærran, A., & Bugge, E. S. (2021). Long short-term memory RNN. *arXiv preprint arXiv:2105.06756*.
- Wan, B., Jiang, W., Fang, Y., Luo, Z., & Ding, G. (2021). Anomaly detection in video sequences: A benchmark and computational model. *IET Image Processing*, 15(14), 3454-3465.

- Wang, B., & Yang, C. (2022). Video Anomaly Detection Based on Convolutional Recurrent AutoEncoder. *Sensors*, 22(12), 4647.
- Wang, L., Tan, H., Zhou, F., Zuo, W., & Sun, P. (2022). Unsupervised anomaly video detection via a double-flow convlstm variational autoencoder. *IEEE Access*, 10, 44278-44289.
- Wang, T., Miao, Z., Chen, Y., Zhou, Y., Shan, G., & Snoussi, H. (2019). AED-Net: An Abnormal Event Detection Network. *Engineering*, 5(5), 930-939. <https://doi.org/10.1016/J.ENG.2019.02.008>
- Wang, T., & Snoussi, H. (2014). Detection of abnormal visual events via global optical flow orientation histogram. *IEEE Transactions on Information Forensics and Security*, 9(6), 988-998.
- Wilmet, V., Verma, S., Redl, T., Sandaker, H., & Li, Z. (2021). A Comparison of Supervised and Unsupervised Deep Learning Methods for Anomaly Detection in Images. *arXiv preprint arXiv:2107.09204*.
- Wu, J.-C., Hsieh, H.-Y., Chen, D.-J., Fuh, C.-S., & Liu, T.-L. (2022). Self-supervised Sparse Representation for Video Anomaly Detection. *European Conference on Computer Vision*, 729-745.
- Wu, S., Moore, B. E., & Shah, M. (2010). Chaotic invariants of lagrangian particle trajectories for anomaly detection in crowded scenes. *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 2054-2060. <https://doi.org/10.1109/CVPR.2010.5539882>
- Xue, R., Chen, J., & Fang, Y. (2020). Real-time anomaly detection and feature analysis based on time series for surveillance video. *2020 5th International Conference on Universal Village (UV)*, 1-7.

- Yokkampon, U., Chumkamon, S., Mowshowitz, A., & Hayashi, E. (2020). Anomaly Detection using Variational Autoencoder with Spectrum Analysis for Time Series Data. *2020 Joint 9th International Conference on Informatics, Electronics and Vision and 2020 4th International Conference on Imaging, Vision and Pattern Recognition, ICIEV and icIVPR 2020*. <https://doi.org/10.1109/ICIEVICIVPR48672.2020.9306570>
- Yu, G., Wang, S., Cai, Z., Zhu, E., Xu, C., Yin, J., & Kloft, M. (2020). Cloze test helps: Effective video anomaly detection via learning to complete video events. *Proceedings of the 28th ACM International Conference on Multimedia*, 583-591.
- Yuan, H., Cai, Z., Zhou, H., Wang, Y., & Chen, X. (2021). TransAnomaly: Video Anomaly Detection Using Video Vision Transformer. *IEEE Access*, 9, 123977-123986.
- Yuan, Y., Fang, J., & Wang, Q. (2014). Online anomaly detection in crowd scenes via structure analysis. *IEEE transactions on cybernetics*, 45(3), 548-561.
- Yuan, Y., Fang, J., & Wang, Q. (2015). Online anomaly detection in crowd scenes via structure analysis. *IEEE Transactions on Cybernetics*, 45(3), 548-561. <https://doi.org/10.1109/TCYB.2014.2330853>
- Zaheer, M. Z., Mahmood, A., Khan, M. H., Segu, M., Yu, F., & Lee, S.-I. (2022). Generative Cooperative Learning for Unsupervised Video Anomaly Detection. *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 14744-14754.
- Zhang, B. (2003). Comparison of the performance of center-based clustering algorithms. *Lecture Notes in Artificial Intelligence (Subseries of Lecture Notes in Computer Science)*, 2637, 63-74. https://doi.org/10.1007/3-540-36175-8_7/COVER

Zhao, Y., Deng, B., Shen, C., Liu, Y., Lu, H., & Hua, X.-S. (2017). Spatio-temporal autoencoder for video anomaly detection. *Proceedings of the 25th ACM international conference on Multimedia*, 1933-1941.

Zhong, Y., Chen, X., Jiang, J., & Ren, F. (2022). Reverse erasure guided spatio-temporal autoencoder with compact feature representation for video anomaly detection. *Science China Information Sciences*, 65(9), 1-3.

Zhou, C., & Paffenroth, R. C. (2017). Anomaly detection with robust deep autoencoders. *Proceedings of the ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, Part F129685*, 665-674.
<https://doi.org/10.1145/3097983.3098052>

Zhu, X. (Jerry). (2005). *Semi-Supervised Learning Literature Survey*.
<https://minds.wisconsin.edu/handle/1793/60444>