

İSTANBUL TEKNİK ÜNİVERSİTESİ ★ LİSANSÜSTÜ EĞİTİM ENSTİTÜSÜ

TEK SAHNELİ SIKIŞTIRILMIŞ VİDEODA ANOMALİ TESPİTİ

YÜKSEK LİSANS TEZİ

Sümeyye ÇAVAŞ

Bilgisayar Bilimleri Anabilim Dalı

Bilgisayar Bilimleri Programı

ŞUBAT 2022

İSTANBUL TEKNİK ÜNİVERSİTESİ ★ LİSANSÜSTÜ EĞİTİM ENSTİTÜSÜ

TEK SAHNELİ SIKIŞTIRILMIŞ VİDEODA ANOMALİ TESPİTİ

YÜKSEK LİSANS TEZİ

**Sümeyye ÇAVAŞ
(704181019)**

Bilgisayar Bilimleri Anabilim Dalı

Bilgisayar Bilimleri Programı

Tez Danışmanı: Prof.Dr. Behçet Uğur TÖREYİN

ŞUBAT 2022

İTÜ, Lisansüstü Eğitim Enstitüsü'nün 704181019 numaralı Yüksek Lisans Öğrencisi Sümeyye ÇAVAŞ, ilgili yönetmeliklerin belirlediği gerekli tüm şartları yerine getirdikten sonra hazırladığı “TEK SAHNELİ SIKIŞTIRILMIŞ VİDEODA ANOMALİ TESPİTİ” başlıklı tezini aşağıda imzaları olan jüri önünde başarı ile sunmuştur.

Tez Danışmanı : **Prof. Dr. Behçet Uğur TÖREYİN**
İstanbul Teknik Üniversitesi

Jüri Üyeleri : **Dr. Öğr. Üyesi Süha TUNA**
İstanbul Teknik Üniversitesi

Dr. Öğr. Üyesi Nurullah ÇALIK
İstanbul Medeniyet Üniversitesi

Teslim Tarihi : 10 Ocak 2022
Savunma Tarihi : 14 Şubat 2022





Aileme,



ÖNSÖZ

Çalışma sürecim boyunca birlikte yol almaktan gurur duyduğum, motive edici yaklaşımı ile desteğini sürekli hissettiğim değerli danışman hocam Prof.Dr. Behçet Uğur TÖREYİN’e, yorumları ve yol göstericiliği ile çalışmama katkıda bulunan değerli hocam Arş.Gör. Muhammet Sebul BERATOĞLU’na şükranlarımı sunar ve çok teşekkür ederim.

Ayrıca, bu süreç boyunca hep yanımda olan sevgili kardeşim Ömer Faruk ÇAVAŞ’a tüm destekleri için sonsuz teşekkürlerimi sunuyorum. Son olarak, hayatı anlamlı yapan bütün değerleri kendilerinden öğrendiğim ve halen öğrenmeye devam etmekte olduğum biricik aileme tüm kalbimle sonsuz minnetlerimi sunuyorum.

Bu çalışma, TÜBİTAK ARDEB 121E378 numaralı “Videonun Sıkıştırılmış Ortamda Anlamlandırılması” başlıklı proje ve 3190538 numaralı TEYDEB projesi ile MGA-2017-40964 ve MOA-2019-42321 numaralı İTÜ BAP projeleri kapsamında desteklenmiştir.

Ocak 2022

Sümeyye ÇAVAŞ
(Bilgisayar Mühendisi)



İÇİNDEKİLER

Sayfa

ÖNSÖZ.....	vii
İÇİNDEKİLER	ix
KISALTMALAR	xi
ÇİZELGE LİSTESİ.....	xiii
ŞEKİL LİSTESİ.....	xv
ÖZET.....	xvii
SUMMARY	xxi
1. GİRİŞ.....	1
1.1 Motivasyon.....	1
1.2 Tezin Amacı	2
1.3 Tezin Organizasyonu.....	2
2. İLGİLİ ÇALIŞMALAR	3
2.1 Videoda Anomali Tespiti Hakkında Genel Bilgiler	3
2.1.1 Videoda anomali kavramaları	3
2.1.2 Videoda anomali türleri	4
2.1.2.1 Sadece görünüm anomalileri.....	4
2.1.2.2 Kısa süreli hareket anomalileri.....	4
2.1.2.3 Uzun vadeli yörünge anomalileri	4
2.1.2.4 Grup anomalileri	5
2.1.2.5 Günün saati anomalileri	5
2.1.3 Yüksek Verimli Video Kodlama Standardı (HEVC – H.265).....	5
2.2 Normal Video Verisi İle Gerçekleştirilen Anomali Tespiti Yaklaşımları.....	7
2.2.1 Mesafe tabanlı yaklaşımlar	7
2.2.2 Olasılıksal yaklaşımlar	12
2.2.3 Yeniden oluşturma tabanlı yaklaşımlar.....	15
2.3 Videonun Sıkıştırılması ile Ortaya Çıkan Verilerden Faydalanılarak Gerçekleştirilen Anomali Tespiti Yaklaşımları.....	21
3. VERİ KÜMESİNİN HAZIRLANMASI	25
3.1 Veri Kümesinin Tanıtımı	25
3.2 Veri Kümesi Ön İşleme Aşamaları	26
4. YÖNTEM.....	31
4.1 Varyasyonel Otomatik Kodlayıcı.....	31
4.2 Kayıp Fonksiyonu	33
4.3 Yeniden Parametreleştirme Hilesi.....	33
4.4 Varyasyonel Otomatik Kodlayıcı Mimarisi	34
4.4.1 Veri Kümesinin Modele Uygun Hale Getirilmesi	34
4.4.2 Varyasyonel Otomatik Kodlayıcı Mimarisi Katmanları	35
5. DENEYLER VE TARTIŞMA	39
5.1 Değerlendirme Yöntemleri.....	39
5.1.1 Düzenlilik Skoru Grafikleri	39

5.1.2 Zamansal Anomali Tespiti Performansı.....	40
5.1.3 Doğruluk (Accuracy).....	41
5.1.4 Kesinlik (Precision).....	41
5.1.5 Duyarlılık (Recall).....	41
5.1.6 F1 Skoru (F1 Score)	42
5.2 Deneylerin Gerçekleştirilmesi	42
5.2.1 Ped1 Veri Kümesi Üzerindeki Sonuçlar	42
5.2.1.1 Ped1 Düzenlilik Skoru Grafikleri.....	42
Test001 numaralı test videosunun incelenmesi.....	42
Test008 numaralı test videosunun incelenmesi.....	43
Test013 numaralı test videosunun incelenmesi.....	44
Test024 numaralı test videosunun incelenmesi.....	45
5.2.1.2 Ped1 Zamansal Anomali Tespiti Performansı.....	46
5.2.2 Ped2 Veri Kümesi Üzerindeki Sonuçlar	48
5.2.2.1 Ped2 Düzenlilik Skoru Grafikleri.....	48
Test002 numaralı test videosunun incelenmesi.....	48
Test004 numaralı test videosunun incelenmesi.....	49
5.2.2.2 Ped2 Zamansal Anomali Tespiti Performansı.....	50
5.2.3 Geçmiş Çalışmalar İle Performans Karşılaştırılması	51
5.3 Vargılar ve Tartışma.....	53
6. SONUÇ.....	55
KAYNAKLAR.....	57
ÖZGEÇMİŞ.....	63

KISALTMALAR

HEVC	: High Efficiency Video Coding
KNN	: K-Nearest Neighbors
SVM	: Support Vector Machines
DAE	: Denoising Auto Encoders
SSIM	: Structural Similarity Index Measure
CNN	: Convolutional Neural Network
GNG	: Growing Neural Gas
TCP	: Temporal CNN Pattern
PGM	: Probabilistic Graphical Models
MoPPCA	: Mixture of Probabilistic Principal Components Analyzers
MDT	: Mixture of Dynamic Textures
GPR	: Gaussian Process Regresyon
RoI	: Region of Interest
LSTM	: Long Short Term Memory
GAN	: Generative Adversarial Networks
RBM	: Restricted Boltzmann Machines
VAE	: Variational Autoencoder
AUC	: Area Under Curve
ROC	: Receiver Operating Characteristic



ÇİZELGE LİSTESİ

Sayfa

Çizelge 1: UCSD Ped1 veri kümesi zamansal anomali tespiti performansı.	47
Çizelge 2: UCSD Ped2 veri kümesi zamansal anomali tespiti performansı.	50
Çizelge 3: Modellerin UCSD veri kümesi üzerinde elde ettiği AUC değerleri.	52





ŞEKİL LİSTESİ

Sayfa

Şekil 1: UCSD veri kümesine ait solda normal çerçeve, sağda ise anomali içeren çerçeve.	26
Şekil 2: UCSD veri kümesi ön işleme adımları.	26
Şekil 3: UCSD veri kümesinin sıkıştırılması sonucu ortaya çıkan hareket vektörleri örneği.	28
Şekil 4: UCSD veri kümesinin boyutunun 480x720 şeklinde ayarlanması sonucu hareket vektörleri.	29
Şekil 5: UCSD veri kümesindeki çerçevelerin sadece hareket vektörlerini içerecek hale getirilmesi.	29
Şekil 6: UCSD veri kümesindeki çerçevelerin gri tonlamalı hale getirilmesi.	30
Şekil 7: Standart bir otomatik kodlayıcı ile varyasyonel otomatik kodlayıcının farkı.	32
Şekil 8: Yeniden parametreleştirme işlemi.	34
Şekil 9: Modelin genel mimarisi.	36
Şekil 10: Modelin kodlayıcı bölümünün mimarisi.	37
Şekil 11: Modelin kod çözücü bölümünün mimarisi.	38
Şekil 12: UCSD Ped1 veri kümesindeki Test001 numaralı test verisinin düzenlilik skoru grafiği, 120. çerçeve kesiti.	43
Şekil 13: UCSD Ped1 veri kümesindeki Test008 numaralı test verisinin düzenlilik skoru grafiği, 72. çerçeve kesiti.	44
Şekil 14: UCSD Ped1 veri kümesindeki Test013 numaralı test verisinin düzenlilik skoru grafiği, 43. çerçeve kesiti.	45
Şekil 15: UCSD Ped1 veri kümesindeki Test024 numaralı test verisinin düzenlilik skoru grafiği, 120. çerçeve kesiti.	46
Şekil 16: UCSD Ped2 veri kümesindeki Test002 numaralı test verisinin düzenlilik skoru grafiği, 140. çerçeve kesiti.	48
Şekil 17: UCSD Ped2 veri kümesindeki Test004 numaralı test verisinin düzenlilik skoru grafiği, 80. çerçeve kesiti.	49



TEK SAHNELİ SIKIŞTIRILMIŞ VİDEODA ANOMALİ TESPİTİ

ÖZET

Sürekli olarak gelişen dünyamız sağladığı birçok kolaylık ile birlikte bunların bir sonucu olarak ortaya çözülmesi gereken yeni problemler çıkarmaktadır. Bu durum, teknolojinin sürekli olarak gelişmesi ihtiyacını doğurmuştur. Bu teknolojik gelişmenin bir parçası olarak, günden güne gelişen, yeni özelliklere sahip olan, giderek daha küçük bir boyuta sığdırılan ve erişebilirliği tüm dünyayı kapsamış olan kameraları gözümüzü çevirdiğimiz her tarafta görebilmekteyiz. Bu kadar etrafımızda olan kameralardan her bir gün veri kaydının gerçekleştirildiğini düşündüğümüzde, kaydolan bu veri miktarının muazzam seviyelere ulaşabileceğini çok rahat bir şekilde söyleyebiliriz.

Ortaya çıkan bu verilerin insanlar tarafından sürekli olarak takip edilip kontrol altına alınması mümkün değildir. Bu noktada, çözüm olarak birçok sektörde insandan bağımsız akıllı sistemler geliştirilmiştir ve halen geliştirilmeye devam edilmektedir. Örnek olarak, pazarlama sektöründe bir ürüne olan talebin daha çok nasıl artırılacağı gibi yada sağlık sektöründe bir hastalık için daha doğru karar veren sistemlerin nasıl geliştirilebileceği gibi yada savunma sektöründe daha güvenli sistemlerin nasıl oluşturulabileceği ve yolunda gitmeyen birşeyler olduğunda nasıl farkedilebileceği gibi çeşitli sorulara akıllı cevaplar sunan sistemler düşünülebilir.

Bu tezde, kameralardan toplanan veriler üzerinde akıllı bir çıkarım yapma çalışması gerçekleştirilmiştir. Bu çıkarım, kameradan alınan bir video verisinin, normal seyrinde hareket edip etmediğini veya anormal herhangi bir durum içerip içermediğini tespit etmek şeklindedir. Bu çalışma, bir videonun çerçeve bazlı düzenlilik değerini grafik üzerinde gösterme ve tüm test videolarını kapsayacak şekilde modelin performansını değerlendirme şeklinde oluşturulmuştur.

Videoda anomali tespiti üzerine şuana kadar birçok çalışma gerçekleştirilmiştir. Bu çalışmayı diğerlerin ayıran en önemli özellik, çalışmanın sıkıştırılmış video alanında ortaya çıkan hareket vektörleri kullanılarak gerçekleştirilmiş olmasıdır. Video sıkıştırma üzerine birçok standart geliştirilmiştir. Bu çalışmada, günümüzde yaygın olarak kullanılan video sıkıştırma standartlarından biri olan H.265 standardına göre sıkıştırılmış olan videolar kullanılmıştır.

Çalışma boyunca veri kümesi olarak, UCSD Pedestrian veri kümesi kullanılmıştır. Bu veri kümesinde, yaya yolu üzerindeki iki ayrı kameradan elde edilen videolar çerçeveler halinde sunulmuştur. Bu çerçeveler üzerindeki hareket vektörlerini çıkarabilmek için çeşitli işlemler gerçekleştirilmiştir. İlk olarak, bu çerçeveler bir araya getirilerek videolar oluşturulmuştur. Daha sonra her bir videonun sıkıştırılması amaçlanmıştır. Bunun için International Telecommunication Union (ITU) tarafından

paylaşılan açık kaynak kod kullanılmıştır. Bu kaynak kodun derlenmesi sonucunda ortaya çıkan çalıştırılabilir dosya kullanılarak videolar üzerinde sıkıştırma işlemi gerçekleştirilmiştir. Sıkıştırma işleminden sonra, H.265 standardında, diğer bir ismi ile High Efficiency Video Coding (HEVC) standardında oluşmuş, .hevc uzantılı video verileri elde edilmiştir. Sıkıştırılan .hevc uzantılı videolar yine açık kaynak kodlu, GitlHevcAnalyzer analiz aracı kullanılarak incelenmiştir. Bu analiz aracı sıkıştırılmış videoya ait çeşitli bilgileri videonun çerçeveleri üzerinde göstermektedir. Ayrıca bu çerçeveleri üzerinde yer alan bilgiler ile indirmeye olanak sağlamaktadır. Bu araç sayesinde, sıkıştırılmış video üzerindeki hareket vektörleri videonun çerçeveleri üzerinde gösterilip, hareket vektörlerinin yer aldığı çerçeveler indirilerek kullanılmak üzere dışarıdaki bir ortama aktarılabilir. Ayrıca bu araç açık kaynak kodlu olduğu için geliştirilip özelleştirilebilir hale getirilebilmesi de mümkündür. Örnek olarak, araç tarafından her bir çerçevenin tek tek indirilmesi mümkündür. Bizim çalışmamızda kullandığımız veri kümesinde ise 9210 tane çerçeve yer almaktadır. Bunların hepsinin araç üzerinden tek tek indirilmesi oldukça zahmetli ve uzun süren bir iş olacaktır. Bu süreci kolaylaştırmak amacıyla, bir videonun tüm çerçevelerini üzerindeki bilgiler ile birlikte tek seferde indirip dış ortama aktaracak kod geliştirilmiştir. GitlHevcAnalyzer aracı bu şekilde özelleştirilerek veri çıkarma sürecinin kolaylaşması sağlanmıştır.

Hareket vektörlerini içeren çerçeveler model tarafından kullanılabilir hale gelmesi için çeşitli önileme adımlarından geçirilmiştir. İlk olarak sıkıştırma sonucunda hareket vektörlerinin asıl ortaya çıktığı 8 ve 8'in katı olan çerçeveler seçilmiştir. Diğer çerçeveler elenerek veri kümesine dahil edilmemiştir. Sonrasında, çerçeveler üzerindeki hareket vektörlerinin daha net bir şekilde görünmesi için, çerçevelerin yükseklik ve genişliği 480x720 olacak şekilde artırılmıştır. Daha sonra, hareket vektörleri dışındaki pikseller beyaz olacak şekilde tüm çerçeveler düzenlenmiştir. Böylece, çerçeveler üzerinde sadece hareket vektörleri bilgisinin yer alması sağlanmıştır. Son olarak görüntüler grayscale formatına getirilmiştir. Bu işlem ile birlikte çerçeveler üzerindeki önilemeler tamamlanarak model tarafından kullanılabilir hale getirilmiştir.

Videoda anomali tespiti gerçekleştirmek için varyasyonel otomatik kodlayıcı tabanlı model kullanılmıştır. Bu varyasyonel otomatik kodlayıcıya, hareket vektörlerini içeren çerçeveler üçerli gruplar halinde verilmiştir. Böylece modelin geçmiş çerçevelerden de öğrenebilmesi sağlanmıştır. Varyasyonel otomatik kodlayıcının ilk katmanlarını konvolüsyonel katmanlar oluşmaktadır. Bu sayede, video çerçevelerindeki uzamsal bilgilerin öğrenilmesi amaçlanmıştır. Sonraki katmanlarını ise konvolüsyonel LSTM katmanlarından oluşmaktadır. Bu sayede ise, video çerçeveleri arasındaki zamansal bilgilerin öğrenilmesi amaçlanmıştır. Böylece, videodaki anomalinin, uzam-zamansal bir öğrenme gerçekleştiren varyasyonel otomatik kodlayıcının temelinde yer alan yeniden oluşturma hatasına dayalı olarak tespit edilmesi amaçlanmıştır. İlk olarak, tamamen normal videolardan oluşan eğitim kümesi ile model eğitilmiştir. Sonrasında, gerçekleştirilen eğitim sonucunda oluşan yeniden oluşturma hatasının dağılımından yararlanılarak eşik değeri seçimi yapılmıştır. Çerçeve bazlı anomali tespiti yaparken bu eşik değeri kullanılmıştır.

Model performansını değerlendirmek için iki farklı yöntem kullanılmıştır. İlk olarak, her bir test videosu için çerçeve bazlı düzenlilik skoru grafiği çıkarılmıştır. Bu grafik sayesinde videonun hangi çerçevesinden itibaren anormal bir durum gerçekleştiği, hangi çerçevelerden itibaren normale döndüğü gözlemlenebilmektedir. Test veri kümesinde normal ve anormal olacak şekilde çerçeve bazlı bir etiketleme mevcuttur.

Bu çalışmada da, test veri kümesinde yer alan her bir videonun çerçeveleri için seçilen eşik değere göre tahminleme yapılmıştır. Böylece, ikinci olarak da, model tarafından yapılan tahminleme ile gerçek veri etiketleri kullanılarak modelin çerçeve bazlı tahminleme performansı değerlendirilmiştir. Oluşturulan düzenlilik skoru grafikleri incelendiğinde video akışına göre genel olarak tutarlı sonuçlar gösterdiği gözlemlenmiştir. Çerçeve bazlı etiketlemenin sonuçları incelendiğinde ise, UCSD Ped1 veri kümesinde anormal çerçevelerin %60'lık kısmının, UCSD Ped2 veri kümesinde ise anormal çerçevelerin %78'lik kısmının tespit edilebildiği görülmüştür. Sonuç olarak, bir videonun çerçevelerindeki tüm bilgileri kullanmak yerine, sadece bu çerçevelerdeki hareket vektörleri kullanılarak yapılmış olan bu çalışmanın etkin sonuçlar çıkardığı görülmüştür.





ANOMALY DETECTION IN SINGLE SCENE COMPRESSED VIDEO

SUMMARY

Our constantly developing world, along with the many conveniences it provides, creates new problems that need to be solved as a result of these. This situation has led to the need for continuous self-development and renewal of technology. As a part of this technological development, we can see the cameras, which are developing day by day, with new features, which are compacted in a smaller size and whose accessibility covers the whole world, wherever we turn our eyes. Considering that data is recorded every day from the cameras around us, we can easily say that this amount of recorded data can reach enormous levels.

It is not possible to continuously monitor and control these emerging data by people. At this point, intelligent systems independent of humans have been developed by many sectors as a solution, and are they are still being developed. For example, systems that provide smart answers to various problems such as how to increase the demand for a product in the marketing sector, how to develop systems that make a more accurate diagnosis for disease in the health sector, or how to create safer systems in the defense sector and how to detect when something goes wrong.

In this thesis, an intelligent inference study was carried out on the data collected from the cameras. This inference is to detect whether video data received from the camera moves in its normal course or contains any abnormal conditions. This study was created by showing the frame-based regularity score of a video on the graph and evaluating the performance of the model by covering all test videos.

Many studies have been carried out so far on anomaly detection in video. The most important feature that distinguishes this study from others is that it was carried out using motion vectors that emerged in the compressed video. Many standards have been developed on video compression. In this study, videos compressed according to the H.265 standard, which is one of the most widely used video compression standards today, were used.

The UCSD Pedestrian dataset was used as the dataset throughout the study. In this dataset, the videos obtained from two separate cameras on the pedestrian road are presented in frames. Various operations have been carried out to extract the motion vectors on these frames. First, videos were created by combining these frames. It is then intended to compress each video. For this, open-source code shared by International Telecommunication Union (ITU) was used. The videos were compressed by using the executable file resulting from the compilation of this source code. After

the compression process, video data with .hevc extension was obtained, which was formed in the H.265 standard, also called High Efficiency Video Coding (HEVC) standard. Compressed videos with .hevc extension were analyzed using GitlHevcAnalyzer analysis tool, which is also open source. This analysis tool shows various information of the compressed video on the frames of the video. It also allows downloading these frames with the information on them. Thanks to this tool, the motion vectors on the compressed video can be displayed on the frames of the video, and the frames containing the motion vectors can be downloaded and transferred to an external environment for use. In addition, since this tool is open source, it is possible to develop and customize it. For example, by default, we have to download each frame one by one. In the dataset we used in our study, there are 9210 frames. Downloading all of them one by one from the tool will be a very laborious and long-term job. To facilitate this process, a code has been developed that will download all the frames of a video together with the information on it at once and export it to the external environment. The GitlHevcAnalyzer tool has been customized in this way, facilitating the data extraction process.

Frames containing motion vectors have undergone various preprocessing steps to make them usable by the model. Firstly, frames whose index are multiples of eight in which motion vectors appear prominently are selected. Other frames were eliminated and not included in the dataset. Afterward, the height and width of the frames were increased to 480x720 so that the motion vectors on the frames could be seen more clearly. Then, all the frames are arranged so that the pixels except the motion vectors are white. Thus, only the motion vectors information is provided on the frames. Finally, the images were converted to the grayscale format. With this process, the preprocessing on the frames has been completed and they are made usable by the model.

A variational autoencoder-based model was used to perform anomaly detection in the video. Frames containing motion vectors are given in groups of three to this variational autoencoder. Thus, it is ensured that the model can also learn from previous frameworks. The first layers of the variational autoencoder are the convolutional layers. In this way, it is aimed to learn the spatial information in video frames. The next layers consist of convolutional LSTM layers. In this way, it is aimed to learn the temporal information between video frames. Thus, it is aimed to detect the anomaly in the video based on the reconstruction error, which is obtained by the variational autoencoder performing a spatio-temporal learning. First, the model is trained with a training set consisting of completely normal videos. Afterward, threshold value selection was made by making use of the distribution of the reconstruction error resulting from the training. This threshold value was used while detecting frame-based anomalies.

Two different methods were used to evaluate model performance. First, a frame-based regularity score graph is plotted for each test video. Thanks to this graphic, it can be observed from which frame of the video an abnormal situation occurs and from which frames it returns to normal. Frame-based labeling is available in the test dataset, with normal and abnormal labels. In this study, the frames of each video in the test dataset were predicted according to the selected threshold value. Thus, secondly, the frame-based prediction performance of the model was evaluated using the prediction made by the model and the actual data labels. When the created regularity score graphs are examined, it has been observed that it generally shows consistent results according to the video stream. When the results of frame-based labeling were examined, it was seen

that 60% of abnormal frames in theUCSD Ped1 dataset and 78% of abnormal frames in the UCSD Ped2 dataset could be detected. As a result, it has been seen that this study, which was made by using only the motion vectors in these frames, instead of using all the information in the frames of a video, yielded effective results.





1. GİRİŞ

Günden güne değişen, gelişen ve iyileşen bir teknoloji döneminde yaşamaktayız. Bu hızlı dönemin sonuçlarını elimizin değdiği her noktada görebilmekteyiz. Hayatımızın birçok noktasında yer alan kameralar da bu dinamik sürecin küçük bir parçasını oluşturmaktadır. Önceleri resmi kaygılar ile sadece kurumsal alanlarda yer alan kameralar artık hem boyutsal hem fiyat olarak tüm kullanıcıların rahatlıkla erişebileceği seviyelere gelerek hayatımızın her alanında var olmaktadır. Kameraların bu hızlı yaygınlaşması ile birlikte, bunlar tarafından toplanan ve depolanan veri miktarı da aynı hızla artış göstermiştir. Bu artış belli bir noktadan sonra insanlar tarafından takip edilemeyecek boyuta ulaşmıştır. Diğer taraftan elektronik ortamda depolanan veriler, insanların birbirleri ile veya tüm dünya ile paylaşması sonucu birikmiştir. Bütün bunların sonucunda, biriken ses, görüntü, yazı, video gibi farklı türdeki verilerin anlamlandırılması çalışmaları başlamıştır. Bu çalışmada, kameradan elde edilen video verileri üzerinde bir anlamlandırma metodu geliştirilecektir. Video anlamlandırma tanımını ise videodaki anomali tespiti ile sınırlı tutulmuştur. Özetle, bu çalışmada, ilk önce ilgilenilen videoda anormal bir olay olup olmadığını belirlemeye yönelik değerlendirmeler gerçekleştirilmiştir. Daha sonra eğer anormal bir durum oluşmuş ise bunu zamansal olarak ortaya çıkarmaya yönelik çalışmalar gerçekleştirilmiştir.

1.1 Motivasyon

Daha önceden videoda anomali tespitine yönelik birçok çalışma gerçekleştirilmiştir. Bu çalışmaların çoğunda girdi olarak videoyu oluşturan tüm çerçeveler ve bu çerçevelerden çıkarılan bir takım öznitelikler kullanılmıştır. Bu çalışma ise, videonun sıkıştırılması işleminde elde edilen bazı parametrelere dayanmaktadır. Herhangi bir videoyu oluşturan tüm çerçevelerdeki bütün bilgileri kullanmak yerine, çerçeveler arasındaki değişimleri işaret eden bilgilerin kullanılması tercih edilmiştir. Günümüzde videoların artan çözünürlük kabiliyetleri ile birlikte muazzam bellek alanı kaplayacağı düşünüldüğünde, bu videoların tamamını depolamak yerine sadece değişiklikleri içeren kısımlarını depolamanın, veri depolama ve depolanan verinin yönetimi açısından çok büyük miktarda kolaylık sağlayacağı açıktır. Aynı zamanda çok daha az

miktarda veri ile çok daha hızlı modeller oluşturulabileceği ön görülmüştür. Bu çalışmanın arkasındaki motivasyon, herhangi bir videoyu daha az miktardaki veri ile temsil ederek buradan anlamlı sonuçlar üretebilme arayışıdır.

1.2 Tezin Amacı

Bu çalışmadaki temel amaç, herhangi bir video üzerindeki anomalileri tespit etmektir. Bu konu daha önce birçok çalışmada incelenmiş ve başarılı sonuçlar elde edilmiştir. Ancak, daha önce gerçekleştirilen çalışmaların çoğunda, herhangi bir videoyu oluşturan tüm çerçevelerdeki bütün bilgileri kullanarak çeşitli çözüm yöntemleri sunulmuştur. Bu noktada, bu çalışmayı özel kılan şey, video sıkıştırılması esnasında elde edilen parametrelere dayalı olmasıdır. Böylece, bir videonun temel içeriğini daha az miktardaki veri ile temsil ederek, ilgili videodaki anomali durumlarını ortaya çıkarmak amaçlanmıştır.

1.3 Tezin Organizasyonu

Bu çalışmada, 2. bölümde ilk olarak tez konusu kapsamında bazı kavramlar açıklandıktan sonra videoda anomali tespitine yönelik daha önceden gerçekleştirilen çalışmalar hakkında bilgi verilmektedir. Bunu takiben video sıkıştırması sonucu elde edilen parametreleri kullanarak gerçekleştirilmiş olan anomali tespitine yönelik çalışmalar hakkında bilgi verilmektedir. 3. bölümde, bu çalışmada kullanılacak olan veri kümesinin hazırlanması aşamaları anlatılmaktadır. Sonrasında yer alan 4. bölümde, bu çalışmada önerilen yöntemin temel yapısı detaylı olarak anlatılmaktadır. En son bölüm olan 5. bölümde ise gerçekleştirilen deneyler ile bu deneylerden çıkarılan sonuçlar tartışılmaktadır.

2. İLGİLİ ÇALIŞMALAR

Bu bölümde, ilk olarak, tezin daha anlaşılır olabilmesi için bazı kavramlar açıklanmaktadır. Daha sonra videoda anomali tespiti üzerine gerçekleştiren bazı çalışmalara değinilmiştir. Son olarak ise videonun sıkıştırılması sonucu elde edilen parametrelerin kullanılması ile gerçekleştirilen anomali tespitine yönelik çalışmalar ele alınmıştır.

2.1 Videoda Anomali Tespiti Hakkında Genel Bilgiler

2.1.1 Videoda anomali kavramaları

Bu tezde ele alındığı çerçevede kapsamında videoda anomali tespitine ait bazı kavramların açıklaması aşağıdaki gibidir.

Anomali, en temel hali ile olağan dışı basit aktiviteler olarak tanımlanabilir. Anomaliler, aynı zamanda anormallik, yenilik ve aykırı değer olarak da isimlendirilir.

Videoda olağandışı görünüm veya hareket özelliklerinin ortaya çıkmasına videoda anomali denir. Bununla birlikte, olağan görünüm veya hareketlerin olağandışı lokasyon veya zamanda meydana gelmesi de videoda anomali kapsamına girmektedir [1].

Videoda yer alan anomaliyi, bulunduğu zaman veya mekan bazında tespit etme işlemine videoda anomali tespiti denir. Videoları yer aldığı arka plan veya diğer bir ismi ile sahne bazında ele alırsak, videoda anomali tespit etme işlemi tek sahneli ve çoklu sahneli olmak üzere iki türlü sınıflandırabiliriz [2]. Bu iki sınıflandırma aşağıdaki gibi tanımlanır.

Video boyunca arka planın değişmediği yani sabit bir kamera ile elde edilmiş bir video üzerinde anomali tespiti yapma işlemine tek sahneli videoda anomali tespiti denir. Dolayısı ile tanımı gereği lokasyona bağımlı anomalileri içerebilir yani bir sahnenin belli bir bölgesinde anormal olan bir aktivite veya nesne başka bir bölgesinde normal aktivite veya nesne olabilir. Bu duruma çimlerde yürümek örneğini verebiliriz. Çim üzerinde yürümek anormal bir aktivite iken aynı aktivite video sahnesinin kaldırım bölgesinde meydana gelmiş olsaydı normal olarak değerlendirilecektir.

Sabit bir arka plana sahip olmayan yani hareketli bir kamera ile elde edilen video üzerinde anomali tespiti yapma işlemine ise çok sahneli videoda anomali tespiti denir. Çok sahneli videoda anomali tespiti işlemi gerçekleştirmek için, normallik modeli oluştururken birbiri ile alakası olmayan birçok farklı sahneden normal videolar toplanır. Dolayısı ile bu tür anomali tespiti tanımı gereği lokasyon bağımlı anormallikleri içeremez. Burada amacımız herhangi bir videonun herhangi bir yerinde meydana gelen anomaliyi tespit etmektir. Videoların toplandığı sahneler arasında denklik yoktur ancak tüm sahnelerden elde edilen verilerde normal ve anormal kavramlarının tutarlı olması gerektiği dikkat edilmesi gereken bir noktadır.

Gerçek dünya uygulamalarında tek sahneli videolarda anomali tespiti işlemi çok daha yaygındır. Bu tezde de videoda anomali tespiti işlemi tek sahneleri videoda anomali tespiti kategorisinde ele alınmıştır. Dolayısı ile sabit arka plana sahip video verileri üzerinde çalışmalar gerçekleştirilmiştir ve bu şartlar altında model önerisi yapılarak testler uygulanmıştır.

2.1.2 Videoda anomali türleri

Videolarda yaygın olarak meydana gelen anomali türleri aşağıdaki gibidir [2].

2.1.2.1 Sadece görünüm anomalileri

Bir sahnede olağan dışı nesne görünümünden dolayı meydana gelen anomaliye denir. Yaya yolunda yer alan bir bisikletliyi bu anomali türüne örnek olarak verebiliriz. Bu türde bir anomali, tek bir video çerçevesinin yerel bir bölgesi incelenerek tespit edilebilir.

2.1.2.2 Kısa süreli hareket anomalileri

Bir sahnede olağan dışı nesne hareketinden dolayı meydana gelen anomaliye denir. Örnek olarak kütüphanede koşan bir kişiyi düşünebiliriz. Videonun yerel bir bölgesi kısa bir süre incelenerek bu türde bir anomali tespit edilebilir.

2.1.2.3 Uzun vadeli yörünge anomalileri

Bir sahnede olağan dışı nesne yörüngesinden dolayı meydana gelen anomaliye denir. Kaldırımda zikzak çizerek yürüyen insanı veya trafiğe girip çıkan aracı bu anomali

türüne örnek olarak verebiliriz. Bu türde bir anomali, daha uzun video segmentleri izlenerek tespit edilebilir.

2.1.2.4 Grup anomalileri

Bir sahnede olağan dışı nesne etkileşiminden dolayı meydana gelen anomaliye denir. Örnek olarak bando takımı gibi belli bir düzende yürüyen bir grup insanı düşünebiliriz. İki veya daha fazla video bölgesi arasındaki ilişkinin analiz edilmesi ile bu türde bir anomali tespit edilebilir.

2.1.2.5 Günün saati anomalileri

Burada ise aktiviteleri anormal yapan etken gerçekleştiği zamandır. Örnek olarak insanların sabahın erken saatlerinde sinema salonuna girmesini düşünebiliriz. Bu türde bir anomali, günün farklı saatleri için farklı bir normallik modeli kullanılarak tespit edilebilir.

Bir videoda anomali tespiti gerçekleştirileceği zaman yukarıda bahsedilen anomali türlerinin hepsinin bulunma zorunluluğu yoktur. Bu anomali türlerinden biri veya birkaçının yer alması anomali tespiti işlemi gerçekleştirmek için yeterlidir.

Sadece görünüm anomalileri ve kısa süreli hareket anomalileri yerel anomaliler olarak adlandırılır. Diğer anomali türleri ise global anomaliler olarak adlandırılır. Herkese açık olarak paylaşılan, videoda anomali tespiti veri kümeleri genellikle yerel anomali türlerini içeren videolardan oluşur.

2.1.3 Yüksek Verimli Video Kodlama Standardı (HEVC – H.265)

Bu çalışmada, ham bir video verisi yüksek verimli video kodlama standardı (HEVC - High Efficiency Video Coding) kullanılarak sıkıştırılmıştır. Geliştirilen modelde, bu sıkıştırma işlemi sonucunda elde edilen parametreler kullanılmıştır. Bu sebeple, bu bölümde, yöntemin daha anlaşılabilir olması açısından ilk olarak video kodlama formatlarının temeli açıklanmıştır. Sonrasında ise yüksek verimli video kodlama standardı (HEVC-H.265) açıklanmıştır.

Dijital videoyu depolamak için gereken bellek boyutu, veri miktarı nedeniyle oldukça büyüktür. Birçok farklı ortamdan elde edilen videoları işlemeyen saklamak neredeyse

mümkün değildir. Bundan dolayı, kabul edilebilir video kalitesine sahip dijital videonun veri miktarını azaltmak için farklı kodlama formatları geliştirilmiştir [3]. Bu video kodlama (sıkıştırma) formatları, dijital video içeriğinin depolanması veya iletilmesi için içerik temsil formatıdır [4]. Bunlar genellikle standart bir video sıkıştırma algoritması kullanılmaktadır. Video kodlama formatlarına örnek olarak H.262, H.264, H.265 (HEVC), Theora, RealVideo RV40, VP9, ve AV1 verilebilir. Bu sıkıştırma formatları yüksek kalitede videoları daha kolay bir şekilde iletebilmemizi ve daha düşük maliyetlerle saklayabilmemizi sağlar.

Yüksek verimli video kodlama (High Efficiency Video Coding - HEVC) standardı, ITU-T VCEG ve ISO/IEC MPEG standardizasyon kuruluşlarının ortaklaşa olarak geliştirdiği video kodlama standardıdır [5]. İlk versiyonu 2013 yılında yayınlanmıştır. Bir önceki H.264 standardına göre iki kat daha fazla sıkıştırma sağlayan, buna karşılık görüntü kalitesini düşürmeyen yeni bir sıkıştırma standardıdır [6]. H.264 standardında yer alan konseptlerin birçok yönden gelişmiş hali olarak düşünülebilir. Video çerçevesinin farklı bölümlerindeki gereksiz alanları kaldırdıkları için bu iki standart işlev bakımından benzerdirler. Yüksek verimli video kodlamanın ana odak noktası, mevcut standartlara göre sıkıştırma performansını, kodlama verimliliğini ve veri kaybı esnekliğini artırmak, ayrıca paralel işleme mimarilerinin yardımıyla veri aktarım sisteminin entegrasyonunu ve uygulamasını kolaylaştırmaktır. H.264 16×16 piksele kadar blokları tanımlarken, H.265 ise 64×64 piksele kadar daha geniş bir blok boyutu aralığını tanımlar. Hareket vektörlerini daha iyi bir hassasiyetle kodlayabilir, böylece bloklarda daha az kalıntı hatası (residual error) oluşturur. Yüksek verimli video kodlama standardı ek olarak, blok kenarlarındaki yapay dokuları azaltmaya yardımcı olan iyileştirilmiş bir blokları açma filtresi (deblocking filter) ve örnek uyarlamalı ofseti (sample adaptive offset) kullanır. Ayrıca H.265'de, değişken blok boyutlu segmentasyonda, hareket vektörü tahmininde, hareket bölgesi birleştirmede ve hareket düzeltme işlemi filtrelemesinde gelişme vardır [7].

HEVC standardı ile video sıkıştırma işleminde hareket vektörleri kullanılır. Öncelikle çerçeveler dikdörtgen veya kare dilimlere ayrılır. Daha sonra bu dilimler kodlama ağaç birimlerine (coding tree unit) ayrılır. Bu birimler de kodlama birimlerine (coding unit) ayrılır. Kodlama birimleri de tahminleme birimlerine (prediction unit) ayrılır. Sıkıştırma işlemi için ilk adım, mevcut çerçevedeki tahminleme birimini önceki

çerçevelerde veya sonraki çerçevelerde aramamaktır. En iyi eşleşmenin olduğu tahminleme biriminin bulunduğu çerçeve ve o çerçevedeki konumu kaydedilir. Bununla birlikte iki tahminleme birimi arasındaki farklılıklar da tutulur. Böylece, bir çerçevedeki tahminleme birimleri referans çerçeve olarak isimlendirilen başka bir çerçevedeki tahminleme birimlerine ve bunların konumlarına bağlı olarak hareket vektörleri çıkarılır. Bu hareket vektörleri aracılığı ile tahminleme birimleri temsil edilir.

2.2 Normal Video Verisi İle Gerçekleştirilen Anomali Tespiti Yaklaşımları

Bu bölümde videoda anomali tespitine yönelik daha önceden gerçekleştirilmiş olan bazı çalışmalar ele alınmaktadır. Bu çalışmaların birçoğunda video üzerinde zamansal olarak anomali tespiti gerçekleştirilirken, yani anomalinin meydana geldiği çerçeve tespit edilirken, bir kısmında ise mekansal olarak anomali tespit işlemi, yani anomalinin bir çerçevenin hangi bölgesinde meydana geldiğinin tespiti, gerçekleştirilmiştir. Anomali tespit biçiminin önemi kullanılacağı alana göre farklılık göstermekle birlikte, birçok çalışma için mekansal tespit oldukça önemli olmaktadır. Normal video verisi kullanarak anomali tespiti gerçekleştirmiş olan çalışmalar [2]'de olduğu gibi en genel hali ile üç şekilde kategorize edilebilir.

2.2.1 Mesafe tabanlı yaklaşımlar

Mesafeye dayalı yaklaşımlar, bir normallik modeli oluşturmak için eğitim verilerini kullanır ve anomali skorlarını belirlemek için bu modelden sapmaları ölçer. Genellikle bu modeller oldukça basittir, ancak iyi bir temsil ve formülasyon yüksek performans sağlar.

Mesafe tabanlı yaklaşımlarda kullanılan birçok farklı öznitelik türü mevcuttur. Aynı zamanda, normal özniteliklere olan mesafeyi ölçmenin de birçok yöntemi vardır. Bu yöntemlerden yaygın olanlarından biri [8-13] çalışmalarında olduğu gibi normal eğitim videosundaki özniteliklerin etrafında karar sınırı hesaplamak için tek sınıflı destek vektör makineleri (one class support vector machines) kullanmaktır. Bu tür yaklaşımların bir dezavantajı, daha önce modeli eğitmekte kullanılmamış olan yeni eğitim verileri ile mevcut modeli güncellemenin maliyetli olmasıdır. Böyle bir durumda, destek vektör makineleri öğrenme algoritması, tüm eski ve yeni eğitim verileri üzerinde yeniden çalıştırılmalıdır. Normal öznitelikleri modelleyebilmek için

destek vektör makinelerine alternatif olarak gauss karışımları (gaussian mixture) dağılımları kullanılmaktadır. Bu yöntem ile normalliğe olan mesafeyi ölçmek için mahalanobis mesafesi kullanılır. Bu fikir [14-17] çalışmalarında kullanılmıştır.

Mesafe tabanlı yaklaşımlarda videoları temsil etmek için önceleri, ön plan maskeleri (foreground masks) [18], [19], akış histogramları (histograms of flow) [20], hareket büyüklüğü (motion magnitude)[19], gradyanların histogramı (histogram of gradients) [21], hareket sınırı histogramları (motion boundary histograms) [22], yoğun gezingeler (dense trajectories) [23] ve uzam zaman ilgi noktası (space time interest point - STIP) [24] öznitelikleri dahil olmak üzere manuel bir şekilde elde edilen birçok öznitelik kullanılmıştır. [18], [8], [9], [10], [25], [15], [12], [13]'deki daha yeni çalışmalar ise, genellikle daha yüksek başarıya sahip derin ağlar tarafından öğrenilen özniteliklere odaklanmıştır. Bu derin ağ tabanlı yöntemler, derin öznitelikleri öğrenmenin çeşitli yollarını ve derin öznitelikleri farklı normallik modellerinde kullanmanın çeşitli yollarını kapsar. Bu bölümün geri kalanı, bazı mesafe temelli çalışmaları kısaca özetlemektedir [2].

[19]. çalışmada anomalilerin yerel uzam zamansal imzaya sahip olduğu önceden kabul edilir. Lokal normal verinin birleşik olasılıksal dağılımında (joint probability distribution) düşük olabilirliğe (low likelihood) sahip olması anomaliye neden olur. Örtüşen sabit boyutlu video hacimlerini çıkarırlar ve bunları düşük seviyeli hareket tanımlayıcılarla temsil ederler. Normal video ve test videosu hacimleri arasındaki toplanan K-En Yakın Komşular (KNN - K-Nearest Neighbors) mesafeleri anomali skorlarını hesaplamak için kullanılır.

[11]'de, yazarlar video kesitinden bir dizi sosyal güç (social force) [26], gradyanların histogramı (histogram of gradients) [21], optik akış histogramı (histogram of optical flow) [20], hareket sınırı histogramı (motion boundary histogram) [22] özniteliklerini ve yoğun gezingeleri (dense trajectories) [23] çıkarmışlardır. Öznitelikleri temsil etmek için Vektör Nicemleme (Vector Quantization) kodlamasını ve anomali tespiti gerçekleştirmek için lineer, RBF veya polinom çekirdekli tek sınıflı bir destek vektör makinesi [27] kullanırlar.

[12]'de yazarlar, video anomali tespiti için derin ağlarla öğrenilmiş temsilleri kullanan ilk yaklaşımlardan birini önermektedir. Düşük boyutlu temsilleri öğrenmek için çok ölçekli sabit boyutlu örtüşen video hacimlerinde yığılmış gürültü giderici otomatik

kodlayıcıların (DAE - denoising auto encoders) iki akışını (RGB ve optik akış) kullanırlar. Daha sonra, anormallik tespitinde kullanmak üzere tek sınıf sınıflandırma gerçekleştirmek için bir RBF çekirdeğine sahip tek sınıflı bir SVM'de [27] DAE'lerden gelen gizli kodları kullanırlar. Ayrıca, temsil aşamasında ve skorelama aşamasında, modeller arasında füzyon gerçekleştirmenin iki yolunu sunarlar.

[16]'da yazarlar, video anomali tespiti için hızlı bir yöntemle odaklanmaktadır. Yapısal Benzerlik İndeksi Ölçümü (SSIM - Structural Similarity Index Measure) kullanarak basit iki katmanlı otomatik kodlayıcıdan özniteliklerin dağılımını ve ayrıca her video hacminin uzam zamansal komşularına olan mesafelerinin dağılımını modellemek için Gauss'ları kullanırlar. Eğitim Gausslarına olan Mahalanobis mesafelerini hesaplayarak anormallikleri tespit ederler.

[14] numaralı çalışma yazarların daha önceki çalışması olan [16]. çalışmaya dayanmaktadır. Kademeli bir yapıda Gauss dağılımları tarafından modellenen öznitelikleri sağlamak için derin bir evrişimli sinir ağı (CNN – Convolutional Neural Network)'nın yanı sıra bir 3D otomatik kodlayıcının birçok iç katmanını kullanırlar.

Bu çalışmayı, kademeli mimariyi iki katmanlı Gauss modellerine basitleştiren ve önceden eğitilmiş tam evrişimli bir ağdan (pre-trained fully convolutional network) gelen öznitelikleri kullanan [15] numaralı çalışma izledi. Ortaya çıkan algoritma, üst düzey bir GPU'da saniyede yüzlerce video çerçevesini işleyebilir.

[9] numaralı çalışmada, video anomali tespiti için önceden eğitilmiş bir evrişimsel sinir ağından gelen özellikleri kullanan ilk yaklaşımlardan birini sunar. Bu, atomik işlem birimi olarak tek çerçeveleri kullanan tek yaklaşımlardan biridir. Ortalama çıkarılmış her çerçevede [28] bir VGG-f ağından çıkarılan derin öznitelikler üzerinde doğrusal bir çekirdeğe sahip tek sınıf bir SVM [27] eğitirler. Skor haritalarını uzam zamansal bir filtreyle yumuşatırlar ve videoyu sabit boyutlu video hacimlerine bölerek ve yama bölgeleri üzerinde anormallik skorlarını basitçe toplayarak yerleştirme gerçekleştirirler.

[13]'de yazarlar, sabit boyutlu video hacimlerinin optik akış alanlarından hareket-özellik temsillerini öğrenmek için evrişimli kazanan hepsini alır otomatik kodlayıcıları (convolutional winner-take-all auto-encoders) [29] kullanmanın bir yolunu sunarlar. Daha sonra, anormallik skorlamasını gerçekleştirmek amacıyla konuma bağlı tek sınıf SVM'ler [27] oluşturmak için öğrenilmiş hareket özelliği temsillerini kullanırlar.

[30]'de yazarlar, anomali tespiti için benzersiz bir geometrik yaklaşım sunarlar. Genişletilmiş dışbükey gövdelerden oluşan bir topluluk (ensemble of extended convex hulls) [31] oluşturmak için eğitim çerçevelerinden yoğun gezingeler (dense trajectories) kullanırlar, bir politop dahil etme testi (polytope inclusion test) kullanarak test zamanında anormallikleri tanımlar, muhtemelen dışbükey gövdeye olan mesafelerini (distance-to-convex-hull) kullanarak bireysel yörüngeleri puanlarlar. Ayrıca anormal bölgeleri tespit etmek ve yanlış pozitif tespitleri filtrelemek için potansiyel olarak anormal gezingeleri kümelerler.

[17]'da yazarlar, videolardan çıkarılan STIP öznitelikleri [24] üzerinde Büyüyen Sinir Gazı (GNG - Growing Neural Gas) [32] algoritmasını kullanarak bir normallik modeli oluştururlar. Geçmiş yöntemlerin değişen sahneler ile yeterince ilgilenmediğini iddia ediyorlar ve GNG modelini, nöron ekleme, silme, öğrenme hızı uyarlaması ve durdurma kriterleri şeklinde çevrimiçi güncellemelerle güçlendirmeyi öneriyorlar. Mesafelerin dağılımını inceleyip yeni örüntülerin GNG modelindeki en yakın komşudan önemli ölçüde farklı olup olmadığını basitçe belirleyerek anomali tespiti gerçekleştirilir.

[25]'te yazarlar, önceden eğitilmiş bir evrimsel ağ olan AlexNet'ten [33] görüntü özniteliklerini kullanmanın başka bir yolunu sunarlar. Ayrıca hem görünüm öznitelikleri hem de optik akış alanları üzerinde çalışan iki akışlı bir model öneriyorlar. Evrimsel sinir ağından çıkarılan öznitelikleri kullanarak, her bir çerçevede ikili haritalar (binary maps) oluşturmak için önceden eğitilmiş bir ikili tam evrimsel ağ (pre-trained binary fully convolutional network) aracılığıyla Yinelemeli Niceleme Hashing'i (Iterative Quantization Hashing) [34] uygularlar. Daha sonra, görünüm özniteliklerinin zaman içindeki değişim miktarının istatistiksel bir ölçüsü olan bir zamansal evrimsel sinir ağları örüntü (Temporal CNN Pattern - TCP) ölçüsü geliştirirler. İki akışın birleşimi, son anomali skor haritalarını üretir.

[35]'de yazarlar, video anomali tespiti için çekişmeli eğitim (adversarial training) kullanan ilk yaklaşımlardan birini sunarlar. Orijinal görüntü yamalarını, oluşturma rolünü oynayan gürültü giderici bir otomatik kodlayıcı ağdan (R) elde edilen gürültülü yamalarının yeniden oluşturulmasından, ayırt etmekle görevli olan bir ayırıcı ağ (D) kullanırlar. R yalnızca eğitim verilerinden gelen görüntü yamaları üzerinde eğitildiğinden, aykırı değerleri yok eder ve böylece D'nin anormal bir görüntü yamasını, onun yeniden yapılandırmasından kolayca ayırt etmesini sağlar.

[10]'da yazarlar iki aşamalı bir anomali tespit algoritması önermektedir. Eğitim videosundan sabit boyutlu video hacimlerini çıkarırlar, bunları konum, görünüm (önceden eğitilmiş bir evrişimsel sinir ağından öznetelik haritaları çıkarma) ve hareket bilgisi (3D gradyanlar şeklinde) ile zenginleştirirler. Anomali tespitinin ilk aşaması için, K-ortalama kümelemesi (K-means clustering) gerçekleştirirler ve sağlam bir temsil oluşturmak için aykırı değerlere karşılık gelen küçük kümeleri ortadan kaldırırlar. Anomali tespitinin ikinci aşaması için, bir daraltılmış normallik kümeleri modeli oluşturmak için K tane tek sınıf SVM'ler, her küme için bir tane, oluşturmayı ve test zamanında bu K tek sınıf SVM'ler altındaki bir test yaması için maksimum puanı anormallik puanı olarak ele almayı içerir.

[8]'de yazarlar, anomali tespit problemini önceki çalışmalarına [10] dayanarak k tane çok sınıflı bire karşı diğer sınıflandırma (multi-class one-versus-rest) problemine dönüştürdüler. Verileri çıkarmak için öznetelik piramit ağlarını (feature pyramid networks) [36] kullanırlar, gizli temsilleri öğrenmek için bu verilerin görünüm ve gradyan özellikleri üzerinde evrişimli otomatik kodlayıcıları eğitirler ve ardından k-ortalama kümelemesini gerçekleştirirler ve ardından ikili bire karşı diğer (binary one-versus-rest) sınıflandırma yapmak için k tane tek sınıf SVM'leri eğitirler. Test zamanında, bir anomali puanı olarak sadece maksimum k sınıflandırma puanının tersini kullanırlar. Uzamsal yerleştirme performansını rapor etmemişlerdir.

[18]'da yazarlar, yakın zamanda yayınladıkları veri kümeleri Street Scene üzerinde ilerde yapacakları karşılaştırmalar için iki temel algoritma sunar. Manuel olarak çıkarılan mesafe ölçümü, sırasıyla normalleştirilmiş L1 veya L2 voksel bazında mesafe, ile birlikte video hacimlerinin manuel olarak çıkarılmış temsillerini kullanarak, en yakın komşu konumuna bağlı basit bir anormallik algılama şeması kullanırlar. Eğitim verilerinden özlü bir örnek model temsili oluşturarak mesafe hesaplamalarının sayısını büyük ölçüde azaltırlar. Bu basit yöntemlerin diğer veri kümelerinde önceki ileri düzey yöntemlerden bazılarında daha iyi performans gösterebildiğini ifade etmişlerdir.

[37]'da yazarlar, çalışmalarını, uzaklık fonksiyonunu ve manuel olarak çıkarılmış temsilleri, Siyam sinir ağı (Siamese neural network) [38] tarafından eğitilmiş olanlarla değiştirerek basit en yakın komşu şeması üzerine inşa etmişlerdir. Siyam ağı, video yama çiftlerini benzer veya farklı olarak sınıflandırmak için eğitilmiştir ve tüm eğitim video parçacıklarından farklı olan ve bu nedenle anormal olan test video parçacıklarını

bulmak için kullanılır. Örnek bir model hedef veri kümesinin eğitim verilerinden öğrenir. Son olarak, her bir test video yamasına anormallik puanları atamak için test video kesitleri ile eğitilmiş Siyam ağını kullanan örnekler arasındaki en yakın komşu puanlaması kullanılır.

2.2.2 Olasılıksal yaklaşımlar

Olasılıksal yaklaşımlar, mesafeyi olasılıksal uzayda bir modele göre hesaplar. Bu yaklaşımlar ile oluşturulan modeller olasılıksal bir çerçevede gerçekleştirilir. Bunun için genellikle, olasılıksal grafik modelleri (probabilistic graphical models - PGM) veya olasılık dağılımlarının çok boyutlu karışımları gibi yöntemler kullanılır. Olasılığa dayalı yaklaşımların çoğu, derin öğrenme yöntemlerinin yaygınlaşmasından önce gelmektedir. Bu yaklaşımlar, uzam-zamansal gradyanlar [39], optik akış alanları[40-45] ve STIP gibi öz nitelikleri [46] markov rastgele alanı (markov random fields) [45], [47], [39] ve gauss karışımları [43], [39] gibi geleneksel modellerle birleştirmişlerdir. Bazı yeni olasılıksal yaklaşımlar ise, derin ağları [48], [49] kullanmakta ve böylece daha iyi doğruluk göstermektedirler. Bu yaklaşımlar, son derece ilkeli olmalarının yanı sıra, sürekli bir şekilde devam eden anomali yapısını iyi bir şekilde modelleme yeteneğine sahiptirler. Bununla birlikte bu yaklaşımların dezavantajı, test zamanında genellikle çok yavaş kalmalarıdır. Aşağıda çeşitli olasılıksal yaklaşımlar özetlenmiştir [2].

[44]'te yazarlar sosyal güç (social force) modelini [26] kullanırlar. Optik akış, bir pikselin optik akışı ile onun etrafındaki piksellerin ortalama optik akış arasındaki kabaca fark olan sosyal kuvvet etkileşimlerini tahmin etmek için kullanılır. Bir pikselin komşularından farklı olmasının nedeni, parçacıklar arasındaki etkileşimlerden kaynaklanmaktadır. Sosyal güç etkileşimlerini modellemek için bir kelime torbası (bag-of-words) modeli kullanılır ve modele göre anormallikler düşük olabilirlikli çerçeveler olarak tespit edilir.

[47]'da yazarlar, basit arka plan çıkarma ile her piksel için ikili hareket etiketlerini hesaplarlar. Normal verilerin hareket etiketi temsiliinde birlikte oluşum (co-occurrence) istatistiklerini hesaplamak için her pikselin etrafındaki uzam zamansal komşulukları kullanırlar ve olabilirlik oranı testi yoluyla anormallik tespiti gerçekleştirmek için oluşum matrisini (co-occurrence matrix) bir Markov Rastgele Alanındaki potansiyel fonksiyon olarak kullanırlar.

[39]'de yazarlar videoyu uzam zamansal gradyanlarla temsil ediyor. Video yamalarının dağılımlarını modellemek için çok değişkenli Gauss'ları kullanırlar. Kamera çerçevesinde belirli bir konumdaki video yamalarının dağılımını temsil etmek için Gauss'ların bir karışımını kullanırlar. Son olarak, video yamaları arasındaki uzamsal ve zamansal korelasyonların etkisini birleştirmek için birleştirilmiş bir Gizli Markov Modeli (Hidden Markov Model) kullanırlar.

[45]. çalışmada videodaki grid yapısından çıkarılan komşu eğitim video yamaları arasındaki ilişkileri modellemek için uzam zamansal Markov Rastgele Alanı kullanmanın bir yolunu sunar. Optik akış gözlemleri üzerinde bir Olasılıksal Temel Bileşen Çözümleyicilerin Birleşimini (Mixture of Probabilistic Principal Components Analyzers - MoPPCA) [50] oluşturarak, her video yamasını grafta bir düğüm olarak temsil ederler. Test zamanında, normalliğin maksimum bir sonsal tahminini hesaplayarak anormallikleri tespit ederler. Ayrıca, çevresel değişiklikleri ve konsept kaymasını hesaba katmak için modellerinin aşamalı olarak nasıl güncellenebileceğini de gösterirler.

[43]'de yazarlar [44]'e benzer şekilde videoda optik akış gözlemleri gridleri üzerindeki parçacıkları kullanan bir yaklaşım tavsiye ederler. Ancak bu parçacıkların gezingelerine odaklanırlar. Bu gezingeleri kümelerler ve iki kaotik değişmez kullanarak bu gezingelerin kaotik dinamiklerini modellerler. Anormallik tespiti, normal verilerden bu kaotik öznelilik kümesi üzerinde bir Gauss karışım modelinin parametrelerinin basitçe tahmin edilmesi ve test verilerinin olabilirliğinin değerlendirilmesi ile gerçekleştirilir.

[51]'de yazarlar, eğitim video yamalarından bir Dinamik Dokular Karışımı (Mixture of Dynamic Textures - MDT) [52], [53] öğrenmeyi ve karışımların daha büyük hücre bölgeleri arasında paylaşılmasını önermektedir. Anomalileri bir diskriminant belirginlik kriteri tarafından verilen yüksek merkez-çevre çıkıntısı olan bölgeler olarak belirlerler [54]. [55]'de, çoklu ölçeklerde çalışmak için MDT temsilini oluştururlar. Koşullu rastgele alan (conditional random field) [56] çerçevesi kullanarak çoklu ölçeklerden uzamsal ve zamansal anomali puanlarını entegre ederler.

[41]'daki yazarlar oldukça benzersiz bir önerme kullanırlar. Bu önerme, anormallik tespitinin, eğitim verilerinden öğrenilen bilgileri kullanarak test verilerindeki normalliği açıklamaya çalışarak dolaylı olarak yapılmalı görüşüdür. Bir çerçevede ön

planı ortaklaşa açıklayan ön plan nesne hipotezlerini (foreground object hypotheses) ve eşleşen normal örnek hipotezlere sahip olanları aynı anda keşfeden bir video ayırıştırma yaklaşımı ararlar. Test zamanında, ön planı açıklamak için gerekli olan ancak normal eğitim verilerinden herhangi bir örnek hipotezle eşleşmeyen nesne hipotezleri anormaldir. [42]'de, sadece görüntü yamaları yerine esnek video boruları biçimindeki nesne hipotezlerini dikkate alarak bu fikri daha da geliştirmişlerdir.

[46]'de yazarlar, anormallikleri saptamak için hiyerarşik yerel ve genel bir yöntem önermektedir. Uzam zamansal ilgi noktası (STIP - Spatio-Temporal Interest Point) öznitelikleri [24] ile videoyu modelliyorlar ve sonrasında K-ortalama kümelemesi kullanıyorlar. Yerel anomalileri k.nıncı en yakın komşuya uzak mesafeli olanlar olarak tespit ediyorlar. Genel anormallikler için, RBF çekirdeği ile Gauss Süreç Regresyonu (GPR - Gaussian Process Regresyon) modelleri [57] oluşturmak için STIP özniteliklerinin topluluklarını göz önünde bulundururlar. Daha sonra, k.nıncı en yakın komşu GPR topluluğu modeline göre düşük olabilirlikli test gruplarını anormal olarak belirlerler.

[48]'de yazarlar, anormal olayları tespit ettiklerinde yeniden saymak için benzersiz bir yöntem önermektedirler. İlk olarak büyük ölçekli COCO [58] ve Visual Genome[59] görüntü veri kümelerinden nesne, eylem ve öznitelik sınıflarını tahmin etmek için bir Fast-RCNN [60] modeli eğitirler. Daha sonra, her bir çerçeve için, ikinciden sonuncuya kadar olan tam bağlantılı katmandan (second-to-last fully connected layer) her bir ilgilenilen bölgenin (region of interest - RoI) özniteliklerini çıkarırlar. RBF çekirdeğine sahip tek sınıf bir SVM eğitim örneğine en yakın komşu mesafesi ile veya RBF çekirdekli bir çekirdek yoğunluk tahminine (kernel density estimate) göre olabilirlik fonksiyonu ile anomali tespiti gerçekleştirirler. Yeniden sayma, basitçe nesne, eylem ve nitelik sınıflarının maksimum tahminlerine bakılarak gerçekleştirilir.

[49]'de yazarlar, normal görüntü yamalarının 3D gradyanlarından öğrenilen derin temsilleri çıkarmak için PCANet [61] kullanıyor. Daha sonra, log-olabilirlik (log-likelihood) fonksiyonunda alt sınırı maksimize ederek, normal örüntülerin üretken bir sürecini modellemek için Derin GMM'leri [62] kullanırlar. Derin GMM modeli, anormallik puanları olarak kullanılan test örüntüleri için basitçe olabilirlik puanları verir.

2.2.3 Yeniden oluřturma tabanlı yaklařımlar

Yeniden oluřturma tabanlı yaklařımlar, normal videodan öğrenilen üst düzey veya kompakt bir temsil kullanarak girdiyi temsil etmeyi ve daha sonra yalnızca bu gösterimi kullanarak girdiyi yeniden oluřturmayı amaçlar. Anomaliler gibi dağılım dıřı girdilerin, normal videodan öğrenilen bir temsil kullanılarak yeniden yapılandırılmasının, dağılım içi normal verilerle karřılařtırıldığında doęası gereęi daha zor olduęundan, yeniden yapılandırma hatasının anormallik puanı için bir gösterge olarak kullanılmasına dayanırlar. Yeniden yapılandırmaya dayalı yaklařımların neredeyse tamamı, modern derin öğrenme yöntemlerini kullanır ve özellikle çoęu, evriřimli otomatik kodlayıcılara (convolutional autoencoders) [63]-[66] veya üretken çekiřmeli aęlara (generative adversarial networks - GAN) [65], [67], [68], [69] dayanır. Genel olarak, yeniden yapılandırmaya dayalı yaklařımlar, kullandıkları modellerin yeni normal eğitim videosuna uyum saęlaması için yeniden eğitilmelidirler. Bu durum da bu yaklařımlar için bir dezavantajdır. Bu yaklařımların çoęunda, yeniden yapılandırma hatası genellikle piksel düzeyinde olmasına raęmen, anomalilerin uzamsal lokalizasyonu deęerlendirilmez. Yani bu yaklařımlar, bir çerçevede meydana gelen anomalinin o çerçevenin hangi bölgesinde olduęuna dair bir bilgi içermezler. Bunun muhtemel nedeni olarak, yaklařımların uzamsal lokalizasyon doęruluklarının düşük olması düşünölmüřtür. Otomatik kodlayıcı tipi yeniden oluřturma yöntemlerinin bir bařka dezavantajı da, çerçeveler için yeniden oluřturma hatalarının çerçevedeki ön plan nesnelerinin sayısıyla orantılı olmasıdır. Bu nedenle, bu yöntemlerin çoęunda son iřleme adımı olarak video bařına normalleřtirme iřlemi gerçekteřtirilir. Bu bölümün geri kalanında çeřitli yeniden yapılanma temelli yaklařımlar özetlenmiřtir [2].

[63]'de yazarlar, eğitim video kesitlerini piksel bazında L2 kayıp fonksiyonuyla yeniden yapılandırmak için evriřimli bir otomatik kodlayıcıyı eğitiyorlar. Test video kesitlerinde, video bařına normalleřtirilmiř olan yeniden yapılandırma hatası, anormallik puanlarını hesaplamada kullanılır. Bu çalıřmada, anomali tespitinde zamansal yerelleřtirmeye odaklanılmıřtır ve uzamsal yerelleřtirme yapılmamıřtır. İlginç bir řekilde, genelleřtirilmiř bir otomatik kodlayıcıyı (generalized auto-encoder) aynı zamanda çeřitli veri kümelerinden gelen eğitim verileri üzerinde eğitilmiř ve bunun tek bir veri kümesinde eğitilmiř bir model kadar iyi performans elde ettięini göstermiřlerdir.

[64]. çalışmada, evrişimler boyunca çerçevelerin zamansal sırası korunmuştur ve darboğaz katmanındaki zamansal bilgi özel evrişimli LSTM [70] katmanlarıyla modellenmiştir. Bu çalışma, [62]. çalışmanın evrişimli otomatik kodlayıcı (convolutional auto-encoder) mimarisini temel alır.

[67]'da yazarlar videoda anomali tespiti için Üretken Çekişmeli Ağlarının (GAN - Generative Adversarial Networks) [71] ilk kullanımını denemişlerdir. Giriş olarak çerçevenin (x, z) çiftlerini ve gürültü vektörleri alırlar. Sonrasında, farklı modalitenin karşılık gelen y çerçevelerini üreten iki koşullu GAN'ı eğitirler. Ayırıcılardan çerçevelerin (x, y) temsillerini gerçek veya sahte olarak sınıflandırmaları istenir. Anormalliklerin iyi bir şekilde yeniden yapılandırılmadığını varsayarsak, her iki modaliteden gelen yeniden yapılandırma hatalarını birleştirirler. Anomali algılamada ve piksel bazında yerelleştirmede anormallik puanlaması gerçekleştirmek için, video başına normalleştirmeyi kullanırlar.

[72]'de yazarlar, Zıt Iraksama (Contrastive Divergence) [73] eğitim algoritmasını ve Kısıtlı Boltzmann Makineleri (RBM - Restricted Boltzmann Machines) [74] kullanarak sabit boyutlu ham video yamaları üzerinde öznetelik öğrenme ve yeniden yapılandırma gerçekleştirir. Bir anormallik skoru elde etmek için farklı piramit düzeylerinden ve örtüşen yamalardan test zamanında yeniden yapılandırma hatalarını birleştirirler.

[75]'de yazarlar, bir video kesitinin gelecekteki karesini tahmin etmenin anormal aktiviteler için normal olanlara kıyasla daha zor olması gerektiğini ve bu nedenle gelecekteki bir kare tahmin çerçevesi tasarlayacağını iddia ediyorlar. Girdi olarak t uzunluğunda eğitim video kesitlerini alan ve $t + 1$ zamanı için gelecekteki bir çerçeveyi tahmin eden U-net tarzı bir ağ [76] ile eğitiyorlar. Ayrıca, t 'deki çerçeve ile $t + 1$ 'deki gerçek veya yeniden oluşturulmuş çerçeveler arasındaki optik akış haritalarının çiftlerini tahmin etmek için FlowNet [77] kullanırlar. Akış haritaları arasındaki L1 kayıp fonksiyonu değerleri anomali puanlarını oluşturur. Bu kayıp, $t+1$ 'de gerçek ve yeniden oluşturulmuş çerçevelerin farkını almak için çekişmeli kayıp (adversarial loss) ile birlikte yeniden oluşturulmanın yoğunluk ve yönlü gradyanlarıdır. Bunun ardından, video başına hataların normalleştirilmesi onların anomali skorunu oluşturur. Ayrıca uzamsal yerelleştirme performansını da rapor etmemişlerdir.

[68]'de yazarlar, iki akışlı bir modelde yaygın nesne görünüşleri ve bunlarla ilişkili hareketler arasındaki bir uygunluğu öğrenerek sorunu ele alırlar. Girdi olarak tek bir çerçeve kullanarak, hem hareketi tahmin eden bir U-net kod çözücüsü, hem de lp yeniden yapılandırma hata kaybı terimleri tarafından yönetilen giriş çerçevesini yeniden yapılandıran devonkolüsyonlu bir kod çözücü ile birleştirilmiş tek bir kodlayıcı kullanırlar. Bu ağın tamamını koşullu bir GAN'da bir üretici olarak kabul ederler; burada ayırıcı, giriş çerçevesi çiftleri ile karşılık gelen gerçek/tahmini akış alanları arasında ayırım yapan ve bir ikili sınıflandırma kaybı tarafından yönetilen başka bir küçük ağıdır. Çerçeveleri test etmek için, yama seviyesinde lp puanlarını hesaplarlar. Nihai çerçeve düzeyindeki anormallik skorları için ise skorların video başına normalleştirmesini kullanırlar. Uzamsal yerelleştirme performansını rapor etmemişlerdir.

[65]'de yazarlar, geçmişteki yeniden yapılandırma temelli yöntemlerin büyük ölçüde düşük seviyeli öznitelikler üzerinde çalıştığını gözlemlemektedir. Sadece soyut özniteliklerle anomali tespiti yaparak bu sorunu çözmeye çalışırlar. İlk olarak, gürültü giderici otomatik kodlayıcıları (DAE - Denoising Auto-encoder) ham video kesitleri ve bunların akış alanı temsilleri üzerinde eğitirler. Daha sonra çoklu katmanlardaki temsilleri çıkarırlar ve her biri için [67]'ye benzer koşullu GAN'lar eğitirler. Son olarak, her bir çerçeve için bir konsensüs puan haritasına (consensus score map) ulaşmak için çoklu seviyelerden yeniden yapılandırma hata haritalarını birleştirirler.

[69]'de yazarlar, her ikisinin avantajlarından yararlanmak ve dezavantajlarını dengelemek için tahmin ve yeniden yapılandırma yaklaşımlarının birleştirilebileceğini iddia etmektedirler. Bunu, ardışık iki U-net [76] mimarisinden oluşan video kesitleri üzerinde çalışan bir üreteç oluşturarak yapmaya çalışırlar; burada birinci mimari bir ara çerçeveyi tahmin eder ve daha sonra bu ikinci mimari tarafından yakın gelecekteki çerçeveyi tahmin etmek için kullanılır. Bu yaklaşım yoğunluk ve gradyan modalitelerinde yeniden yapılandırma hatasını en aza indirerek uçtan uca eğitilir. Ayrıca, ya gelecekteki çerçevenin asıl hali ile tahmin edilen hali üzerinde ya da PatchGAN'a [78] benzer daha hassas bir seviyede çekişmeli kayıp (adversarial loss) kullanırlar.

[79]'de yazarlar, evrimsel sinir ağı tabanlı yeniden yapılandırma yaklaşımlarının yüksek temsil kapasitesi nedeniyle anormal olayları yeniden yapılandırmakta zorlandıklarını iddia ederler. Yazarlar, bir U-net stili kodlayıcı-kod çözücü ileriye

yönelik çerçeve tahmini ağını, önemli normal paternleri saklayan öğrenilmiş bir bellek modülü ile güçlendirmeyi öneriyorlar. Ayrıca, anomali puanlarını, bir çerçeve ile onun yeniden yapılandırılması arasındaki PSNR'nin ve bir kodlanan ile en yakın bellek ögesi arasındaki mesafenin bir kombinasyonunu kullanarak hesaplar. Ayrıca, anomali puanlarının video başına normalleştirilmesini de gerçekleştirirler.

[80]. çalışmada nesne düzeyinde kendi kendini denetleyen ve çok görevli öğrenme yoluyla videoda anormal olay algılamaya yaklaşımı geliştirilmiştir. Nesneleri algılamak için önce önceden eğitilmiş bir dedektör kullanılmıştır. Ardından, çoklu proxy görevlerini ortaklaşa öğrenerek ayırt edici anomaliye özgü bilgiler üretmek için bir 3B evrişimsel sinir ağı eğitilmiştir: üçü kendi kendini denetleyen ve biri bilgi damıtmasına dayalı. Kendi kendini denetleyen görevler şunlardır: (i) ileri/geri hareket eden nesneleri ayırt etme (zaman oku), (ii) ardışık/aralıklı çerçevelerdeki nesneleri ayırt etme (hareket düzensizliği) ve (iii) nesneyi yeniden yapılandırma. özel görünüm bilgileri. Bilginin damıtılması görevi, hem sınıflandırma hem de tespit bilgilerini hesaba katar ve anormallikler meydana geldiğinde öğretmen ve öğrenci modelleri arasında büyük tahmin tutarsızlıkları oluşturur. Videoda anormal olay algılamaya, birden çok kendi kendini denetleyen ve bilgi damıtma vekil görevlerini tek bir mimaride entegre ederek, çok görevli bir öğrenme sorunu olarak yaklaşan, bilinen ilk çalışma olduğu ifade edilmiştir. Hafif mimarimiz üç kriterde en gelişmiş yöntemleri geride bırakıyor: Avenue, ShanghaiTech ve UCSD Ped2. Ek olarak, çok görevli bir öğrenme ortamında kendi kendini denetleyen öğrenme ve normalliğe özgü damıtmanın entegre edilmesinin önemini gösteren bir ablasyon çalışması gerçekleştirilmiştir. Avenue, Shanghai Tech, UCSD Ped2 veri kümeleri üzerinde test yapılmıştır. Bu veri kümelerinin yıllara göre yapılan çalışmalardaki çerçeve AUC değerleri kıyaslanmış, kendilerinininkinin daha iyi sonuç çıkardığını göstermişler.

[81]. çalışmada videoda anomali tespitinin daha önceki birçok çalışmada ele alındığına ancak bu konudaki zorluğun anomali durumlarının modellenmesinde olduğu kadar veri kümelerinde bu durumların seyrekliğinden kaynaklandığı vurgulanmıştır. Bu sebeple birçok çalışmanın gerçek dünyada uygulamalarında uygulanabilir olmaktan uzak olduğu belirtilmiştir. Bu çalışmada, anomali tespiti için varyasyonel otomatik kodlayıcı yaklaşımının iki versiyonu araştırılmıştır. Varyasyonel otomatik kodlayıcıların girdinin daha zengin temsillerini öğrenerek normal ve anormal videoları ayırt etmede daha iyi olduğunu belirtilmiştir. Varyasyonel konvelasyonel LSTM

otomatik kodlayıcıların standart konvelasyonel otomatik kodlayıcılardan daha iyi performans gösterdiği, beta varyasyonel konvelasyonel LSTM otomatik kodlayıcıların ise varyasyonel konvelasyonel LSTM otomatik kodlayıcılardan daha iyi performans gösterdiği belirtilmiştir. Gauss dağılımı aracılığı ile videoların gizli alanlarının daha derinden anlaşılmasının, test verileri üzerinde performans bakımından katkıda bulunduğu ifade edilmiştir.

[82]. çalışmada yalnızca normal olayları içeren eğitim videolarından öğrenen arka plandan bağımsız (background agnostik) bir yaklaşım önerilmiştir. Bu yaklaşım, bir nesne algılayıcı, bir dizi görünüm ve hareket otomatik kodlayıcı ve bir dizi sınıflandırıcıdan oluşur. Nesne algılama için YOLO kullanılmıştır. Önerilen yaklaşım yalnızca nesne algılamalarına baktığından, normal olayların sahneler arasında aynı şekilde tanımlanması ve tek ana varyasyon faktörünün arka plan olması koşuluyla farklı sahnelere uygulanabilir. Yaklaşım, arka plana değil, anormalliklere neden olabilecek nesnelere güvendiği için, arka plandan agnostiktir. Eğitim sırasında anormal veri eksikliğinin üstesinden gelmek için, otomatik kodlayıcılar için çekişmeli bir öğrenme stratejisi önerilmiştir. Sözde anormal örnekler gradyan yükselişi uygulamadan önce otomatik kodlayıcılar tarafından doğru bir şekilde yeniden oluşturulan, sahneden bağımsız, etki alanı dışı sözde anormal örneklerden oluşan bir küme oluşturulmuştur. Normal ve anormal gizli özellikler ve yeniden yapılandırmalar arasında ayırım yapmak için görünüm tabanlı ve hareket tabanlı ikili sınıflandırıcıları eğitirken anormal örnekler olarak hizmet etmek için sözde anormal örnekler de kullanılmıştır. Ayrıca, otomatik kodlayıcıların yalnızca her sınırlayıcı kutu görüntüsü içindeki ana nesneye odaklanmasını sağlamak için ana nesneyi segmentlere ayırmayı öğrenen bir dal (yapı) tanıtılmıştır. Sunulan yaklaşım, çeşitli değerlendirme ölçütlerini kullanarak dört veri kümesinde en gelişmiş yöntemlerle karşılaştırılmıştır. Mevcut yöntemlerle karşılaştırıldığında, ampirik sonuçlar, yaklaşımın tüm veri kümelerinde olumlu performans gösterdiği ifade edilmiştir. Ek olarak, ShanghaiTech ve Subway veri kümeleri için bölge tabanlı ve iz tabanlı açıklamalar sağlanmıştır. Bu çalışmada sunulan yöntemin sahneden bağımsız olduğunu göstermek için çapraz veri kümesi deneyleri uygulanmıştır. Yani belirli bir sahneye sahip veri kümesi üzerinde model eğitildikten sonra aynı model başka bir sahneye veri kümesi üzerinde test edilmiştir. Bunun sonucunda geçerli performans değerleri elde edilmiştir.

Başka bir çalışmada ise uzun süren ve bir o kadar da yorucu olan manuel etiketlemeden kaçınmak için insanların anormallikleri nasıl algıladığından ilham alındığı vurgulanmıştır [83]. Böylece hem denetimsiz hem de uçtan uca videoda anomali tespiti sağlayan bir yöntem önerilmiştir. Bu yöntem iki temel gözleme dayanmaktadır. Birincisi; insan algısı genellikle yereldir, yani bir insan anormallikleri algılarken yerel ön plana ve bağlama odaklanır, gözlemidir. İkincisi; insanların normallik tanımını sıklıkla meydana gelen olaylar şekillendirir, gözlemidir. Bu gözlem bu çalışmayı gerçekleştiren araştırmacıları vekil bir eğitim paradigması tasarlamaya motive etmiştir. Etiketlenmemiş videolarda bir vekil görevi öğrenmek için derin sinir ağı (DNN) eğitilmiş ve böylece sıklıkla meydana gelen olayların bu derin sinir ağını şekillendirmede baskın rol oynaması sağlanmıştır. Bu sayede nadiren görülen yeni olaylar anomali olarak ortaya çıkarılmıştır. Uygulama için, hem klasik derin sinir ağı hem de yeni ortaya çıkan derin sinir ağı modellerinin yanı sıra çeşitli vekil görevleri keşfedilmiştir. Yaygın olarak kullanılan videoda anomali tespiti karşılaştırma ölçütlerine ilişkin kapsamlı değerlendirmeler yapılmıştır. Bunun sonucunda, yalnızca mevcut denetimsiz çözümlere geniş bir fark atmakla kalmayıp, aynı zamanda güncel yarı denetimli muadilleri ile karşılaştırılabilir ve hatta daha üstün bir performans elde edildiği vurgulanmıştır.

[84] numaralı çalışmada, videoda anomali tespiti için zayıf denetimli bir yaklaşım önerilmiştir. Kullanılan veri kümesinde video bazında etiket mevcut olduğu için yapılan çalışma zayıf denetimli öğrenme kategorisi altında değerlendirilmiştir. Zayıf denetim altında anomali tespitinin, eğitim sırasında anormal ve normal örneklerin belirlenmesindeki zorluk nedeniyle karmaşık olduğu vurgulanmıştır. Dolayısı ile bu modellerin optimal olmayan ayırma marjı ile sonuçlandığı belirtilmiştir. Bu bildiride, anomali örneklerinin hem öznitelik seviyesinde hem de skor seviyesinde normal olanlardan ayırt etmek için farklılık dikkat modülü (DAM) içeren bir yaklaşım önerilmiştir. DAM modülü örneklerin normal mi yoksa anormal mi olduğuna karar vermek için küresel-zamansal (global) bağlamdan çok uzamsal-zamansal (yerel) farklılıkları hesaba katar. Yani bir video içerisindeki klipleri değerlendirir. Bu yaklaşım da önerilen yöntemin ekstra zamana ihtiyaç duymadan gerçek zamanlı senaryolar için anormali tespitine olanak tanır. Ek olarak, ardışık video klipleri arasındaki farklılıkları öğrenmek için iki tane DAM modülü uygulanmıştır. DAM modülü ile birlikte önerilen yaklaşımın UCF-Crime ve ShanghaiTech gibi iki büyük

ölçekli anormallik algılama veri kümesinde doğrulandığı ve gerçek zamanlı güncel yaklaşımlardan sırasıyla %1,5 ve %3,4 daha iyi performans gösterdiği belirtilmiştir. Ayrıca gerçek zamanlı videolarda anomali tespitini sağlama gereksinimlerini karşılama açısından önceki çevrimiçi yöntemlere göre çok daha hızlı olduğu vurgulanmıştır. Bununla birlikte, önerilen yaklaşımın halen sonradan meydana gelen anomali türlerinden etkilendiği ve yanlış pozitiflere neden olabildiği eklenmiştir.

2.3 Videonun Sıkıştırılması ile Ortaya Çıkan Verilerden Faydalanılarak Gerçekleştirilen Anomali Tespiti Yaklaşımları

Ham videonun sıkıştırılması sonucu elde edilen parametreleri kullanarak yapılan çalışmalar aşağıda özetlenmiştir.

[85] makalesinde H.264 formatında sıkıştırılmış veriden elde edilen hareket vektörleri kullanılarak gerçek zamanlı anomali tespit algoritması önerilmiştir. Mevcutta yer alan çoğu çalışmanın piksel düzeyinde elde edilen bilgiler ile çalıştığı için karmaşık öznitelik çıkarma işlemleri uygulandığı ve piksel seviyesindeki verinin miktarından ötürü yavaş işleme yol açtığı vurgulanmıştır. İlgili makalede ise hareket vektörlerinden faydalanarak hesaplama maliyetinin azaltılması amaçlanmıştır. Bu çalışmada bir olay kendisini oluşturan özniteliklerin kümesi olarak tanımlanmıştır. Anomali ise gerçekleşme olasılığı belirli bir eşik değerinin altında kalan olay olarak tanımlanmıştır. Böylece bu çalışmada anomali tespit problemi, olağan davranışları olağandışı davranışlardan ayırt edebilecek ilgili özniteliklerin çıkarılmasına indirgenmiştir. Böyle bir öznitelik olarak da hareket büyüklüğü kullanılmıştır.

Video çözünürlükleri arttıkça bu videolar üzerindeki hesaplama işlemlerinin de artacağı bildiğimiz bir durumdur. Bu makalede bu sorunu çözebilmek için piramit yapısı önerilmiştir. Piramit yapısı videoda orjinal seviye yer alan bilgiler ile bu bilgilerin daha azaltılmış şeklini içeren üst seviyeler olarak tanımlanmıştır. Sunulan yaklaşım ise öncelikle işlemlerin üst seviyelerde gerçekleştirilmesi ve belirsizlik olması durumunda sorunu çözmek için işlemlerin orjinal seviyede gerçekleştirilmesi şeklindedir. Makalede bu hiyerarşik işlem sırası hareket piramitleri olarak adlandırılmıştır.

Anomali tespiti için birbirine yakın hareket vektörlerinin büyüklüklerindeki varyasyona bakılmıştır. Önerilen yöntemde öncelikle olağan davranış eğitimi yapıp

daha sonra anormal örüntü tespiti yöntemi uygulanmıştır. Olağan davranış eğitimi gerçekleştirmeden önce gürültüyü azaltmak amacı ile veri üzerinde bazı ön işlemler uygulanmıştır. Bu çalışmada hareket vektörlerine orjinal seviye ve üst seviye olmak üzerine piramitsel bir yapıda yaklaşıldığına daha önceden değinmiştik. Orjinal seviyede yer alan hareket vektörlerindeki gürültüyü azaltmak için spatio-temporal (uzlam zamansal medyan filtresi) uygulanmıştır. Üst seviyedeki hareket vektörleri, orjinal seviyedeki hareket vektörlerine bi lineer interpolasyon uygulanarak elde edilmiştir. İnterpolasyon işlemi gürültüyü önemli miktarda azaltacağı için bu seviyede ekstra filtreleme gereksiz bulunarak uygulanmamıştır.

Ön işleme adımları tamamlandıktan sonra model kurulmuştur. Model, hareket vektörlerinin büyüklüklerinin histogramı oluşturularak yapılmıştır. Histogramlar her bir konum için (i,j) hareket vektörü büyüklüklerinin zamansal gözlemlerinden oluşmuştur. Sonrasında probability density function elde etmek için histogramlar normalize edilmiştir. Bundan sonra daha az yer kaplaması için density fonksiyonlarına bir miktar gaussion model feet edilmiştir. Sonuç olarak her bir lokasyon için hareket vektörlerinin tüm çerçevelerindeki büyüklüklerinden oluşan farklı parametrelere sahip normal dağılımlar elde edilmiştir.

Test videolarında anomali tespiti tüm çerçeveler için her lokasyonda anomali oluşma olasılığına göre yapılır. Test sırasında öncelikle üst seviyedeki hareket vektörleri için işlemler gerçekleştirilir, olası bir anomali adayı gözlemlendiğinde daha fazla bilgiye sahip olan orjinal seviyedeki hareket vektörleri ile işlem gerçekleştirilir. Üst seviyede anomali olmayan durumlar için işlem yapılmaktan kaçınılır, böylece büyük miktarda gereksiz hesaplamanın yapılmaması sağlanmıştır. Bu yaklaşım yüksek çözünürlüklü videolar için muazzam bir destek sağlar. Çünkü orjinal seviyedeki hesaplamanın çoğu daha yüksek seviyelerde işlem yapılarak atlanmış olur. Düşük çözünürlüklü videolarda ise sadece orjinal seviyede işlem yapılır. Bunun sebebi, eldeki bilgi oldukça az olduğu için, bu bilgi daha da azaltılmaya çalışılırsa bu sefer model yanlış tespit yapmaya başlayacaktır.

Bu çalışmada Peds1, Peds2 ve UMN olmak üzere 3 video verisi üzerinde deneyler gerçekleştirilmiştir. Bu veri kümeleri ilk olarak H.264 standardına göre kodlanmıştır. Algoritma global anomali tespiti ve lokal anomali tespiti olmak üzere iki açıdan değerlendirilmiştir. Peds1 ve Peds2 veriseti global ve lokal anomali tespiti için kullanılmıştır. UMN veri seti ise yalnızca global anomali tespiti için kullanılmıştır.

Önerilen yöntemin yaklaşımı öncelikle anomali içeren çerçeveyi tespit etmek daha sonra ise ilgili çerçevelerdeki anomali bölgelerini tespit etmek şeklindedir. Bu çalışmada UMN veri kümesinin içeriğinden dolayı lokal anomali tespitinin yapılmadığı düşünülmektedir. Peds veri kümesi için orijinal seviye ve bir üst seviye olmak üzere iki seviyeli piramit yapısı kullanılarak işlem yapılmıştır. UMN veri kümesi düşük çözünürlüklü videolar içerdiği için sadece orijinal seviyede işlem yapılmıştır. Üst seviye anomali tespiti için karar eşiği ortalama değerinden $\pm 3,5$ standart sapma olacak şekilde ayarlanmıştır.

Önerilen algoritmanın bir miktar doğruluktan ödün vererek gerçek zamanlı tahmin yapma yeteneğine sahip olduğu söylenmiştir. Bu çalışma, Peds veri kümesinde piramidal yaklaşım kullanarak saniyede 70 çerçevede işlem gerçekleştirmiştir. Bu da mevcuttaki diğer yaklaşımlar ile karşılaştırıldığında 1750 kat ve 280 kat hızlanma sağlamıştır. UMN veri kümesinde ise saniyede 150 çerçevede işlem yapılarak mevcut yöntemlerden 120 kat hızlanma sağladığı belirtilmiştir.

Sıkıştırılmış alanda yapılan bir diğer çalışma [86] yine aynı yazarlar tarafından geliştirilmiştir. Bu çalışmada, mevcut algoritmaların çoğunun piksel düzeyinde bilgi ile çalışmasından dolayı gerçek zamanlı işleme ihtiyacını karşılayamadığı dile getirilmiştir. Piksel düzeyinde bilgi ile çalışılması ve öznitelik çıkarımı için bu büyük miktardaki bilginin kullanılması nedeni ile işlem hızının düşük olduğu vurgulanmıştır. Önerilen metod ile anomali tespitinin genel doğruluğunu etkilemeden işleme hızı artırılmaya çalışılmıştır. Önerilen metod ile anomali tespiti için sadece hareket vektörlerini kullanarak büyük ölçekteki gözetim videolarında yüksek işleme hızı elde edilmiştir. Bu çalışmada bir önceki çalışmaya [1] atıf verilerek, önceki çalışmanın sadece hareket büyüklüğündeki olağan dışı davranışların tespiti için yetenekli olduğunu ancak hareket yönündeki değişikliklerden kaynaklanan anomaliler için başarısız olduğunu söylenmiştir. Buna karşılık önerilen yaklaşımda hem hareket büyüklüğünü hem de hareket yönünü yakalayan anormalliği tespit etmek için yönlü hareket vektörü histogramı (HOMV-Histogram of Oriented Motion Vector) kullanılır. Böylece sıkıştırılmış videoda anomali tespiti için mevcut yöntemlerden daha sağlam bir yöntem geliştirdiklerini ileri sürmüşlerdir. Bu çalışmada da hareket vektörlerinde yer alan gürültüden kurtulmak için bir önceki çalışmada yer alan uzam-zamansal ortalama filtresi kullanılmıştır. Ek olarak videoda yer alan her bir çerçeve için hareket eden objelerin olduğu bölgeler tespit edilerek sadece bu bölgelerde yer alan hareket

vektörleri işleme dahil edilmiştir. Bu bölgeler ilgilenilen bölge (ROI-Region of Interest) olarak isimlendirilmiştir. Bu bölgelerdeki hareket vektörlerinin yönlerine göre histogramlar oluşturulmuştur. Aynı zamanda bu histogramlar ilgili hareket vektörlerinin büyüklüklerine göre ağırlıklandırılmıştır. Sonrasında hareket vektörlerinin sayısına bölünerek normalize edilmiştir. Bu histogramlardan ortalama ve standart sapma değerleri ile gauss (normal dağılım) fonksiyonları elde edilmiştir. Böylece video çerçevelerinde yer alan her bir konum için normal olasılık dağılımı çıkarılmıştır. Daha sonra herhangi bir konumdaki katsayının olabilirlik fonksiyonuna dayalı olarak anomali tespit edilmiştir.

Ped1 ve Ped2 veri kümesi üzerinde lokal ve genel anomali tespiti işlemi gerçekleştirilmiştir. Yapılan deneylerde piksel seviyesinde işlem yapan metodlar ile eşit düzeyde doğruluk gösterirken sıkıştırılmış alanda işlem yapan metodlardan daha iyi doğruluk gösterdiği belirtilmiştir.

3. VERİ KÜMESİNİN HAZIRLANMASI

3.1 Veri Kümesinin Tanıtımı

Bu çalışmada, videoada anomali tespiti çalışmalarında yaygın olarak kullanılan UCSD Pedestrian veri kümesi kullanılmıştır [87]. Bu veri kümesinde UCSD Ped1 ve Ped2 olmak üzere iki ayrı veri içeriği yer almaktadır. Bunların ikisinde bir yaya yolunu gözetleyen farklı statik kameralardan elde edilen videoların çerçevelerinden oluşmaktadır.

Normal şartlar altında, yaya yolunda sadece normal davranışlar sergileyen yayaların yer alması beklenmektedir. Bu sebeple, veriyi kullanıma sunanlar, yaya dışında yer alan tüm objeler ile yayaların olağan dışı davranışlarını anomali olarak tanımlanmışlardır. Örnek olarak yaya yolunda bisikletlinin, patencinin, arabanın yer alması veya yayaların karşıdan karşıya geçmesi gibi durumlar bu veri kümesinde anomaliye sebep olan durumlardır.

UCSD Ped1 ve Ped2 veri kümelerinde yer alan video çerçevelerinin train ve test isminde iki ayrı klasör altında gruplandırıldığı gözlemlenmiştir. Train klasöründe yer alan veriler, sadece normal davranışların yer aldığı videolardan oluşurken, test klasöründe yer alan veriler ise çeşitli anormal davranışları içeren videolardan oluşmaktadır. UCSD Ped1 veri kümesi 158x238 çözünürlüğünde 34 tane eğitim videosu ile 36 tane test videosu içermektedir. Video başına 200 tane çerçeve yer almaktadır. UCSD Ped2 veri kümesinde ise 240x360 piksel çözünürlüğünde 16 tane eğitim videosu ile 12 tane test videosu yer almaktadır. Bu veri kümesindeki videoların her biri 120 ile 200 arasında değişen çerçeveden oluşmaktadır. Her iki veri kümesinde de video verileri kendisini oluşturan çerçeveler şeklinde sunulmuştur. Ayrıca, veri kümesinde, hem uzamsal yani piksel seviyesinde hem de zamansal yani çerçeve seviyesinde anomaliler sağlanmıştır.

UCSD veri kümesine ait normal ve anormal durumları içeren örnek şekil aşağıdaki gibi yer almaktadır.

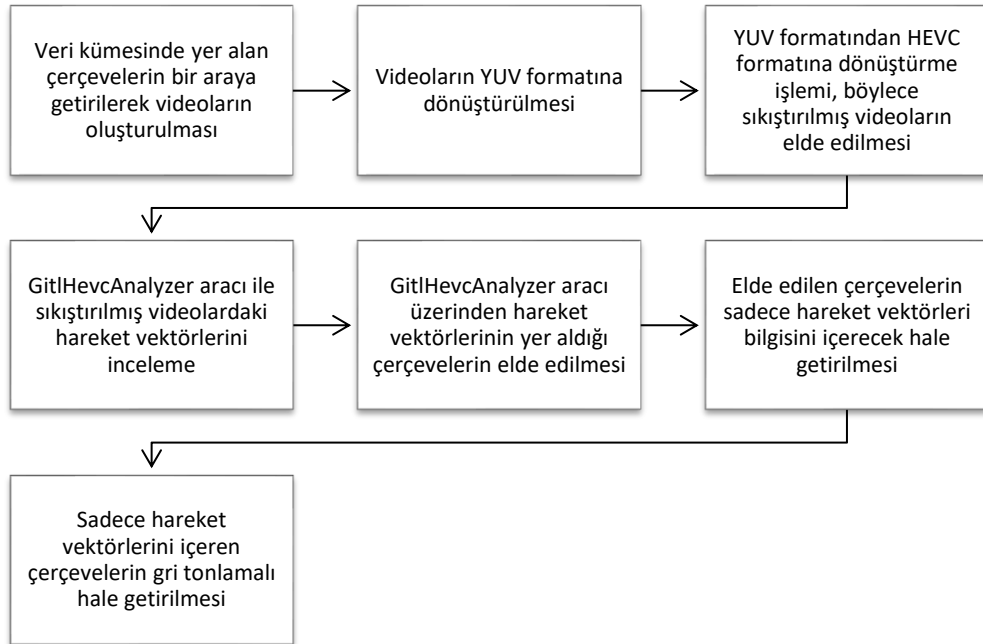


Şekil 1: UCSD veri kümesine ait solda normal çerçeve, sağda ise anomali içeren çerçeve.

Yukarıdaki şekil incelendiğinde, UCSD veri kümesinde yaya yoluna giren bir araba ile bisikletlinin anomaliye sebep olduğu görülmektedir.

3.2 Veri Kümesi Ön İşleme Aşamaları

Bu çalışmada, video verilerinin H.265 standardına göre sıkıştırılması sonucu ortaya çıkan hareket vektörlerinden faydalanılarak videoda anomali tespit edilmesi amaçlanmıştır. Bu sebeple, UCSD veri kümesi, önerilen model tarafından kullanıma hazır hale gelmesi için birtakım ön işleme aşamalarından geçirilmiştir. Aşağıda bu aşamaları özetleyen bir görsel devamında da, bu aşamaların açıklaması yer almaktadır.



Şekil 2: UCSD veri kümesi ön işleme adımları.

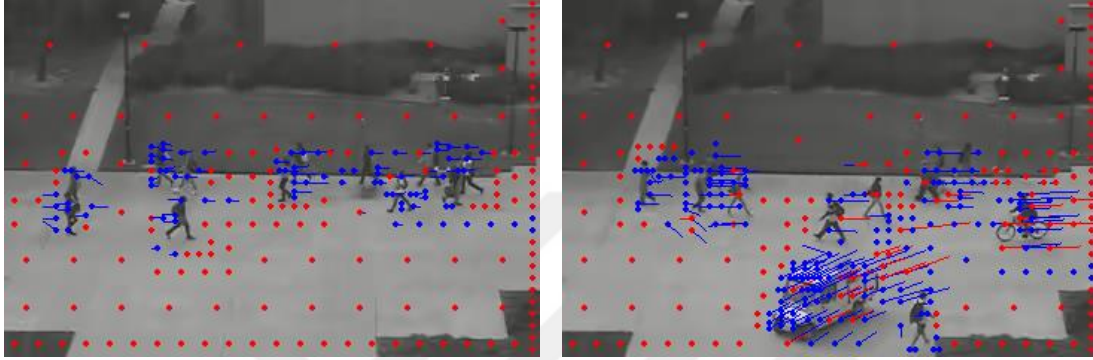
Veri kümesi incelendiğinde, videoa verilerinin çerçeveler halinde sunulduğu görülmektedir. Sıkıştırılmış formatta video elde etmek için öncelikle elimizde normal video verilerinin yer alması gerekmektedir. Bu sebeple ilk olarak bu çerçeveler bir araya getirilerek videolar elde edilmiştir. Bu videolar veri kümesinde Train001, Train002 şeklinde yer alan her bir örnek için ayrı ayrı elde edilmiştir. Yani Train001 klasörü içerisinde yer alan çerçeveler için bir video, Train002 klasörü içerisinde yer alan çerçeveler için ayrı bir video oluşturulmuştur.

HEVC standardında yer alan konfigürasyon gereği dönüştürme işlemi için giriş dosyasının YUV formatında olması gerekmektedir. Bu nedenler, videolar yukarıda anlatılan şekilde oluşturulduktan sonra, YUV formatına dönüştürülmüştür. YUV, genellikle renkli görüntü işleme sistemlerinin bir parçası olarak kullanılan bir renk modeli ve kodlama sistemidir. Renkli bir görüntü veya video için insan algısı dikkate alınarak renklerin temsili için azaltılmış bir bant genişliği sağlayarak bu renk alanında kodlanır, bu şekilde iletim hatalarına veya sıkıştırma kusurları, doğrudan RGB gösterimine kıyasla daha etkili bir şekilde insan algısından gizlenir. Y, luma bileşenini (parlaklık), U ve V, renklilik (renk) bileşenleridir [88]. Bu dönüşümden sonra UCSD Ped veri kümesi sıkıştırma işlemi için gerekli formata getirilmiştir.

Video sıkıştırma işlemi için International Telecommunication Union (ITU) tarafından paylaşılan açık kaynak kod kullanılmıştır [89]. Bu kaynak kod derlendikten sonra oluşan kodlayıcı YUV uzantısındaki videoları sıkıştırmak için kullanılmıştır. Sıkıştırma işlemi için kodlayıcı ile birlikte konfigürasyon dosyası gerekmektedir. Konfigürasyon dosyasında sıkıştırma işlemine ait birçok parametre yer almaktadır ve sıkıştırma işlemi sırasında bu parametreler kullanılmaktadır. Burada giriş videosuna ait genişlik, yükseklik, saniye başına çerçeve sayısı gibi parametreleri doğru bir şekilde vermek önemlidir.

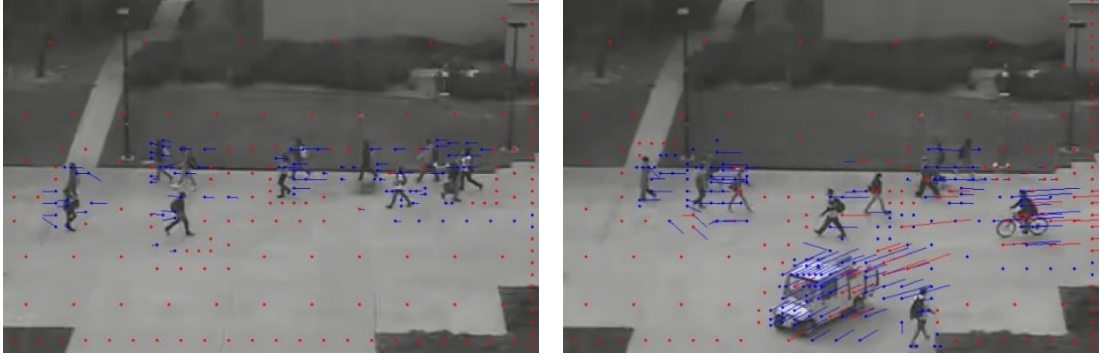
Video sıkıştırma işleminde I,P ve B olmak üzere üç tür çerçeve yer alır. I türü, kodlanmış çerçeve olarak isimlendirilir. I türündeki çerçeveler o çerçeveye ait tüm bilgileri içerirler. B türü, çift yönlü tahmin edilen çerçeve olarak isimlendirilir. Bu tür çerçeveler, önceki ve sonraki çerçevelere ait bilgiler içerir. P türü ise tahmin edilen çerçeve olarak isimlendirilir. Bu çerçeve türünde sadece önceki çerçevelerden farklı olan veri tutulur. Hareket vektörlerinin asıl ortaya çıktığı çerçeve türü P türü çerçevelerdir [90]. Sıkıştırma işlemi için kullanılan konfigürasyon dosyasına çerçeve türleri ile ilgili kaç tane olacağı, hangi aralıkta geleceği gibi düzenlemeler de

yapılabilmektedir. Konfigürasyon dosyasında P çerçevelerini 8 çerçevede 1 tane gelecek şekilde ayarladığımız için sıkıştırdığımız videoda hareket vektörleri 8 çerçevede 1 kez hesaplanmıştır. Sıkıştırma işleminden sonra HEVC uzantısında sıkıştırılmış video verileri elde edilmiştir. Aşağıda UCSD Ped veri kümesinin sıkıştırılması sonucu ortaya çıkan hareket vektörlerinin gözlemlenebildiği örnek bir çerçeve yer almaktadır.



Şekil 3: UCSD veri kümesinin sıkıştırılması sonucu ortaya çıkan hareket vektörleri örneği.

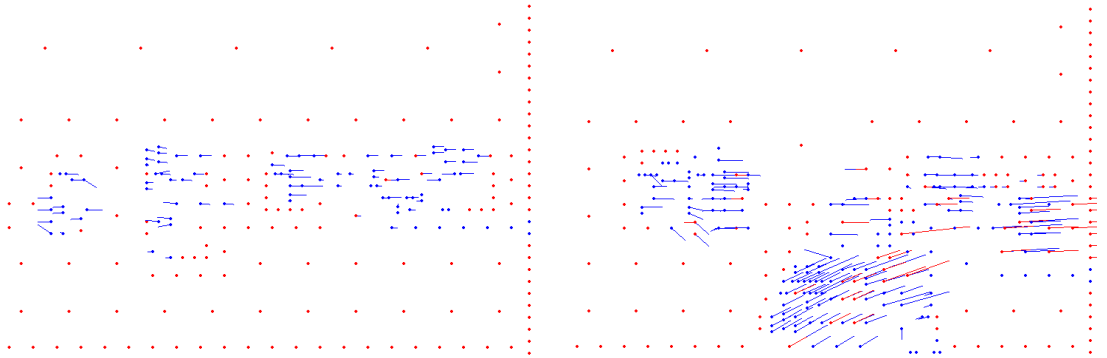
Yukarıdaki görseller HEVC formatına getirilmiş video verisinin GitlHevcAnalyzer [91] aracı üzerinde açılmasıyla elde edilmiştir. Bunlar Şekil-1’de yer alan çerçevelerin hareket vektörleri eklenmiş halidir. Önce Şekil-1’deki çerçevelerde yer alan objelerin hareketleri incelenerek sonrasında yukarıdaki şekil incelendiğinde, hareket vektörlerinin obje hareketleri ile doğru orantılı bir şekilde çıkarılmış olduğu gözlemlenebilir. Burada mavi renkli hareket vektörleri önceki çerçevelerden yararlanılarak oluşturulmuş olan hareket vektörlerini göstermektedir. Kırmızı renkli hareket vektörleri ise sonraki çerçevelerden yararlanılarak oluşturulmuş olan hareket vektörlerini göstermektedir. Ancak yukarıdaki şekil incelendiğinde hareket vektörlerinin piksel olarak aşırı fazla yer kaplayacak şekilde ve iç içe geçmiş bir şekilde olduğu gözlemlenmektedir. Bu durumun çerçeve üzerindeki herhangi bir noktayı temsil etmede yanlış sonuçlar doğuracağı öngörüsü ile hem veri kümesindeki hareket vektörlerinin birbirinden daha çok ayırt edilebilir hale getirilebilmesi için hem de çerçeve üzerindeki herhangi bir noktanın daha düzgün şekilde ifade edilebilmesi için UCSD veri kümesinin boyutu GitlHevcAnalyzer aracı üzerinde 480x720’ye çıkarılmıştır. Böylece aşağıdaki gibi daha net daha temiz çerçeveler elde edilmiştir.



Şekil 4: UCSD veri kümesinin boyutunun 480x720 şeklinde ayarlanması sonucu hareket vektörleri.

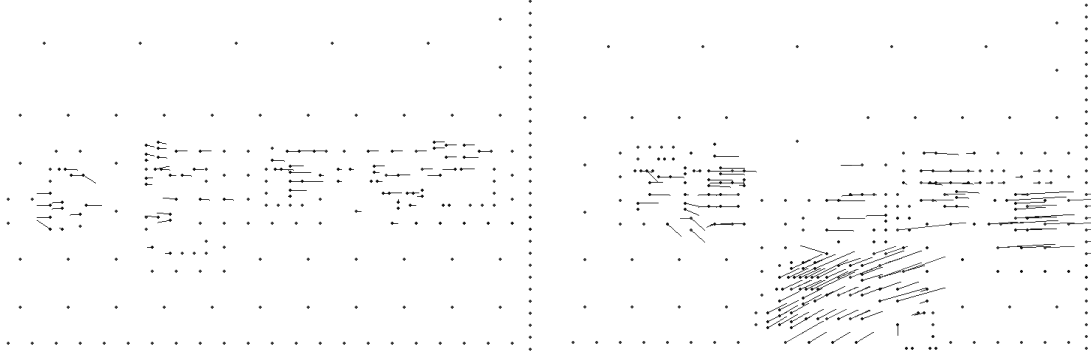
Ayrıca, GitlHevcAnalyzer aracı üzerinden her bir çerçevenin tek tek indirilmesi işlemi oldukça zahmetli olacağından, bu aracın açık kaynak kodu indirilerek bir videoya ait tüm çerçevelerin içerdiği hareket vektörleri ile birlikte otomatik olarak indirilmesini sağlayan kod geliştirilmiştir. Böylece veri kümesi oluşturma süreci hızlandırılmıştır.

UCSD Ped veri kümesi için hareket vektörleri bilgisini içeren çerçeveler yukarıdaki gibi GitlHevcAnalyzer aracı yardımıyla elde edildikten sonra, bu çalışmada önerilen modelde sadece hareket vektörü bilgisinden yararlanılacağı için çerçevelerdeki hareket vektörleri hariç diğer bilgiler temizlenmiştir. Böylece çerçevelerdeki arka plan ve tüm objeler kaldırılmıştır. Bunun sonucunda yukarıdaki şekilde yer alan örnek çerçeveler aşağıdaki gibi elde edilmiştir.



Şekil 5: UCSD veri kümesindeki çerçevelerin sadece hareket vektörlerini içerecek hale getirilmesi.

Son adım olarak, sadece hareket vektörü bilgilerini içeren yukarıdaki renkli çerçeveler gri tonlamalı hale getirilmiştir.



Şekil 6: UCSD veri kümesindeki çerçevelerin gri tonlamalı hale getirilmesi.

Elde edilen çerçevelerden sadece 8 ve 8'in katları olan çerçeveler kullanılarak eğitim ve test veri kümesi oluşturulmuştur. Bunun sebebi ise daha önceden değinildiği gibi hareket vektörlerini içeren asıl çerçevelerin 8 ve 8'in katlarında yer almasıdır. Böylece UCSD Ped veri kümesi model tarafından kullanıma hazır hale getirilmiştir.

4. YÖNTEM

Video verisini sıkıştırıp hareket vektörlerini elde ettikten sonra video üzerinde anomali tespiti yapmak amaçlanmaktadır. Anomali tespiti temel olarak ikili sınıflandırma problemi gibi ele alınabilir. Video üzerindeki anomali tespiti, önerilen yöntemin videoyu oluşturan çerçeveleri normal veya anormal olarak sınıflandırması şeklinde basite indirgenebilir. Yada çerçevelerin hangi bölgelerinin normal hangi bölgelerinin anormal olduğunu tespit etme şeklinde de düşünülebilir.

İkili sınıflandırma işleminde eğitim için kullanılan veri kümesinin etiketli olmasına ihtiyaç duyulmaktadır. Ancak etiketli veri kümesini toplamak hem anormal durumların nadir olması hem de doğası gereği çok değişken olması sebebiyle zordur. Bu zorlu görev çok fazla insan gücü gerektirmekle birlikte tüm anormal durumları içeren bir eğitim kümesi oluşturma imkansızlığı nedeniyle pek mümkün görülmemektedir. Bu sebeple, gözetimli bir öğrenme gerçekleştirilemeyeceğinden sadece normal durumları içeren veri kümesinden öğrenme gerçekleştirebilecek böylece anormal durumlar hakkında çıkarımda bulunabilecek yöntemlere yoğunlaşılmıştır. Bu kapsamda yapılan araştırmalar sonucunda anomali içermeyen video verilerini kullanarak bunlardan normalliği öğrenen sonrasında bu normallikten sapma değerine göre test videolarında anomali durumunu tespit eden bir yaklaşım türü olan otomatik kodlayıcılar uygun bulunmuştur. Bu çalışmada, otomatik kodlayıcıların bir türü olan varyasyonel otomatik kodlayıcılar kullanılmıştır. Varyasyonel otomatik kodlayıcıların katmanlarında ise görüntüden etkin bir şekilde öznetelikleri çıkarmak amacıyla konvolüsyonel katmanlar ve konvolüsyonel uzun kısa süreli bellek (Long Short Term Memory - LSTM) katmanları kullanılmıştır.

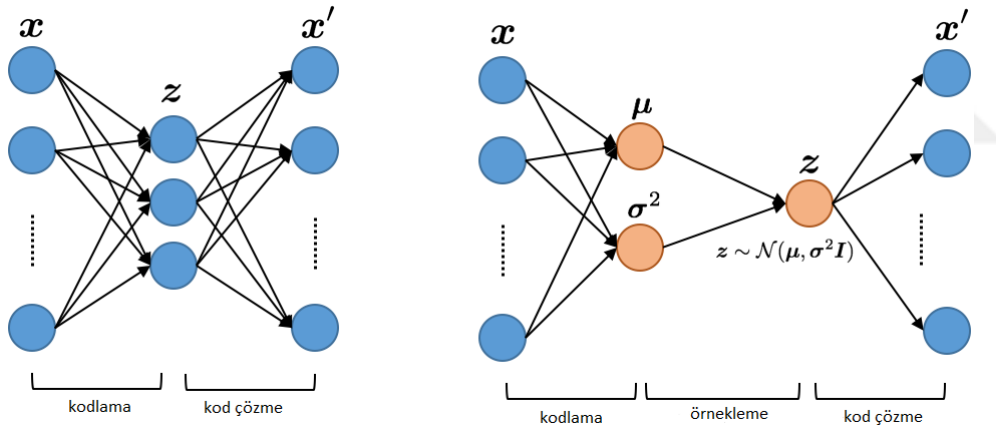
4.1 Varyasyonel Otomatik Kodlayıcı

Varyasyonel otomatik kodlayıcı, standart otomatik kodlayıcının bir türüdür. Bu sebeple ilk olarak standart otomatik kodlayıcının çalışma mantığına değinilerek sonrasında varyasyonel otomatik kodlayıcı açıklanmıştır.

Otomatik kodlayıcılar girdi verisini yeniden oluşturmaya çalışan bir yapay sinir ağıdır. Bu işlemi gerçekleştirirken kullandığı girdi verisi etiketsizdir, dolayısıyla otomatik kodlayıcılar bir gözetimsiz öğrenme çeşidi modelidir [64]. Girdi olarak aldığı veriyi öncelikle bir koda dönüştürüp daha sonra bu koddan tekrar girdi verisini oluşturmaya

amaçlarıdır. Tam bu noktada yapı olarak aslında iki bölümden oluştuğu ortaya çıkmış olur. Girdi verisini alarak bunu bir koda dönüştürdüğü bölüm kodlayıcı (encoder) bölümünü oluştururken, dönüştürülmüş kod bilgisini alarak bundan girdi verisini elde etmeye çalıştığı bölüm ise kodçözücü (decoder) bölümünü oluşturur. Bu iki bölüm otomatik kodlayıcının verilen veriyi sıkıştırıp olabildiğince az kayıpla tekrar üretmesini sağlayacak şekilde tasarlanır.

Kısaca standart bir otomatik kodlayıcılarda her girdi kodlayıcılardan geçirilerek kodlanır ve belirsiz bir vektör (latent vector) dönüştürülür. Bu vektöre kod denir. Daha sonra oluşturulan bu kod, kod çözüciden geçirilerek girdi verisi yeniden oluşturulur. Varyasyonel otomatik kodlayıcılarda ise kodlayıcıdan geçirilen girdi, bir olasılık dağılımı olarak kodlanır. Yani kodlayıcı çıktısı bir kod değil, ortalama ve varyans değerleridir. Bu ortalama ve varyans değerlerinden örnekleme yapılarak kod elde edilir. Bu kodda kod çözücü ile çözülebilir [64]. Aşağıdaki şekilde [92] standart bir otomatik kodlayıcı ile varyasyonel otomatik kodlayıcının genel şeması yer almaktadır.



Şekil 7: Standart bir otomatik kodlayıcı ile varyasyonel otomatik kodlayıcının farkı.

Yukarıdaki şekilde standart bir otomatik kodlayıcı ile varyasyonel otomatik kodlayıcının farkı görülmektedir. Otomatik kodlayıcıda girdi (x) bir koda (z) sıkıştırılır ve tekrar eski haline (x') dönüştürülür. Varyasyonel otomatik kodlayıcıda ise girdiden (x) ortalama (μ) ve varyans (σ^2) değerleri elde edilir. Bu değerler kullanılarak kod (z) örneklenir. Bu adıma örnekleme (sampling) adımı denir. Sonrasında yine bu kod çözülerek eski haline (x') dönüştürülür [64].

4.2 Kayıp Fonksiyonu

Standart otomatik kodlayıcıların eğitilmesi için kullanılan kayıp fonksiyonu yeniden oluşturma kaybı (reconstruction loss)dir. Yani modele verilen girdi ile modelin ürettiği çıktı arasındaki farktır. Bu farkı hesaplarken genellikle ortalama kare hatası (mean squared error) kullanılır.

Varyasyonel otomatik kodlayıcılarda bu kayba ek olarak düzenleme kaybı (regularization loss) kullanılır. Bu kayıp fonksiyonunun diğer bir ismi belirsizlik kaybı (latent loss)dir. Bu kayıp kodlayıcının çıktısı olan ortalama ve varyans değerlerinin oluşturduğu dağılım ile normal dağılım arasındaki farktır. Bu kaybı hesaplamak için Kullback Leibler uzaklığı (Kullback Leibler Divergence) kullanılır. Kullback Leibler uzaklığı iki olasılık dağılımının birbirinden ne kadar farklı olduğunu ölçen bir metriktir. Sonuç olarak varyasyonel otomatik kodlayıcılarda toplam kayıp aşağıdaki gibi yeniden oluşturma kaybı ile düzenleme kaybının toplamı olarak hesaplanır [64].

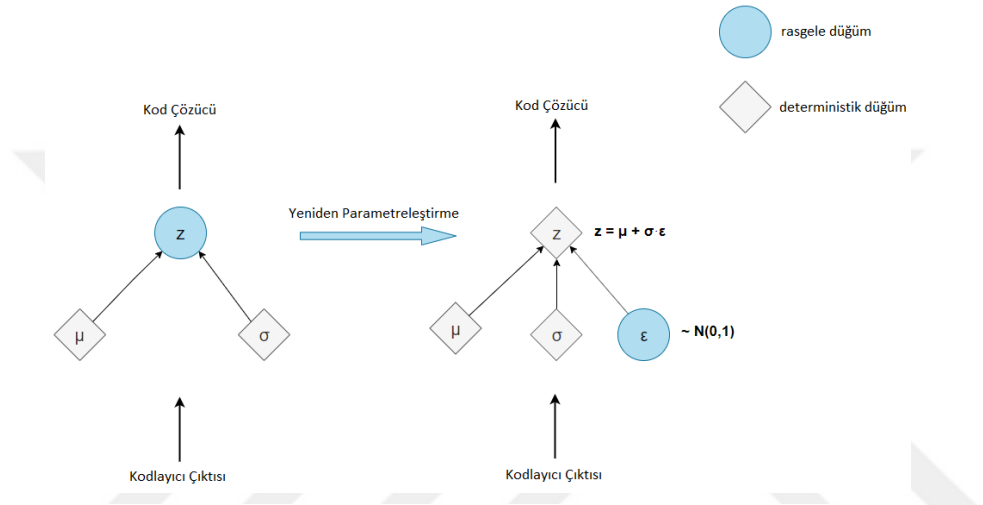
$$\text{Kayıp fonksiyonu} = \|x - x'\|_2 + KL[N(\mu_x, \sigma_x), N(0,1)] \quad (4.1)$$

Bu çalışmada model eğitilirken çerçeveler üçer üçer verilmektedir. Model sonucunda da girdi olarak verilen çerçevelerin yeniden oluşturulması işlemi gerçekleşmektedir. Yukarıdaki kayıp fonksiyonunda x modele girdi olarak verilen çerçeveleri x' ise modelin oluşturduğu çerçeveleri temsil etmektedir. Bu kayıp fonksiyonundaki ilk terim olan $\|x - x'\|_2$ kısmı yeniden oluşturma kaybını ifade etmektedir. Burada x ile x' arasındaki öklid mesafesi hesaplanmaktadır. İkinci terim olan $KL[N(\mu_x, \sigma_x), N(0,1)]$ kısmı ise düzenleme kaybını ifade etmektedir. Burada ise modelin kodlama işlemi sonucu hesapladığı μ ve σ değeri ile elde edilen normal dağılım ile standart normal dağılım arasındaki Kullback Leibler uzaklığı hesaplanmıştır.

4.3 Yeniden Parametreleştirme Hilesi

Yapay sinir ağı modelleri geri besleme (backpropagation) kullanılarak eğitilir. Ancak varyasyonel otomatik kodlayıcılarda ortalama ve varyans değerlerinden kod

örnekleme işlemi geri besleme için uygun değildir. Varyasyonel otomatik kodlayıcıları geri besleme kullanarak eğitebilmek için bir ϵ değeri kullanılır. Bu değer basitçe normal dağılımdan üretilmiş rastgele bir sayıdır. Direkt olarak μ ve σ değerinden kod (z) üretmek yerine “ $z=\mu+\sigma\epsilon$ ” formülü ile kod (z) üretilir. Geri besleme sırasında da ϵ değeri sabit tutularak sistemin eğitilmesi sağlanır. Bu işleme yeniden parametreleştirme hilesi denir. Bu işlemin tek amacı sistemin geri besleme kullanılarak eğitilebilmesini sağlamaktır [64]. Aşağıdaki şekilde [93] yeniden parametreleştirme işlemi yer almaktadır.



Şekil 8: Yeniden parametreleştirme işlemi.

4.4 Varyasyonel Otomatik Kodlayıcı Mimarisi

Veri kümesinin hazırlanması önceki bölümlerde anlatılmıştı. Bu bölümde, sadece hareket vektörlerini içerecek şekilde hazırlanan veri kümesini eğitmek ve sonrasında test verilerinde oluşan sonuçları gözlemlemek için kullanılan modelin mimarisi anlatılacaktır. Modelin mimarisi oluşturulurken aynı veri kümesi üzerinde anomali tespiti yapan [64] numaralı yaklaşımdan ilham alınmıştır.

4.4.1 Veri Kümesinin Modele Uygun Hale Getirilmesi

Veri kümesini hazırlama kısmında gri tonlamalı sadece hareket vektörlerini içeren çerçeveler elde edilmişti. Bu çerçeveler modele sunulmadan önce aşağıdaki aşamalardan geçirilmiştir.

- Gri tonlamalı çerçevelerin her bir pikseli 255'ya bölünerek 0 ile 1 arasında ölçeklendirilmiştir.
- Çerçeveler her biri ard arda gelen 3'erli pencereler halinde modele sunulacak şekile getirilmiştir.

4.4.2 Varyasyonel Otomatik Kodlayıcı Mimarisi Katmanları

Bu çalışmada videoyu oluşturan çerçeveler kendinden daha önce gelen çerçevelerden büyük ölçüde farklı ise orada anormal bir durum vardır prensibine dayanan bir çözüm modeli geliştirilmiştir. Önerilen model, giriş çerçevelerinin belirli bir kesitindeki zamansal örüntüleri zamansal kodlayıcı ve kod çözücü kullanarak öğrenir. Bunun için temelde çerçevelerdeki uzamsal öznitelikler çıkarılmıştır. Böylece uçtan uca öğrenen bir model oluşturulmuştur. Model girdi çerçeveleri ile eğitilmiş model tarafından yeniden oluşturulan çıktı çerçeveleri arasındaki yeniden oluşturma hatasını en aza indirmek amacıyla sadece normal sahneleri içeren videoların çerçeveleri ile eğitilmiştir. Model uygun şekilde eğitildikten sonra, normal video çerçevelerinin düşük yeniden yapılandırma hatasına sahip olması beklenirken, anormal sahnelerden oluşan video çerçevelerinin yüksek yeniden yapılandırma hatasına sahip olması beklenir. Sonrasında, eğitim videosu çerçevelerinden elde edilen eşik değeri kullanılarak her bir test videosu çerçeveleri için anormal olayların meydana geldiği çerçeveler tespit edilmektedir.

Bu çalışmada eğitim veri kümesindeki düzenli örüntüleri öğrenmesi için konvolüsyonel uzam zamansal varyasyonel otomatik kodlayıcı önerilmiştir. Önerilen mimari iki aşamadan oluşmaktadır. Birincisi her video çerçevesinin uzamsal yapısının öğrenilmesi için uzamsal bir varyasyonel otomatik kodlayıcıdır. İkincisi ise uzamsal yapıda kodlanan zamansal örüntüleri öğrenmesi için zamansal bir varyasyonel otomatik kodlayıcıdır. Böylece model, aynı otomatik kodlayıcı üzerinde hem uzamsal hem de zamansal olarak öğrenecek şekilde tasarlanmıştır. Uzamsal kodlayıcı ve kodçözücü sırası ile iki tane konvolüsyonel ve iki tane dekonvolüsyonel katmana sahiptir. Zamansal kodlayıcıda ise üç tane konvolüsyonel LSTM katmanına sahiptir.

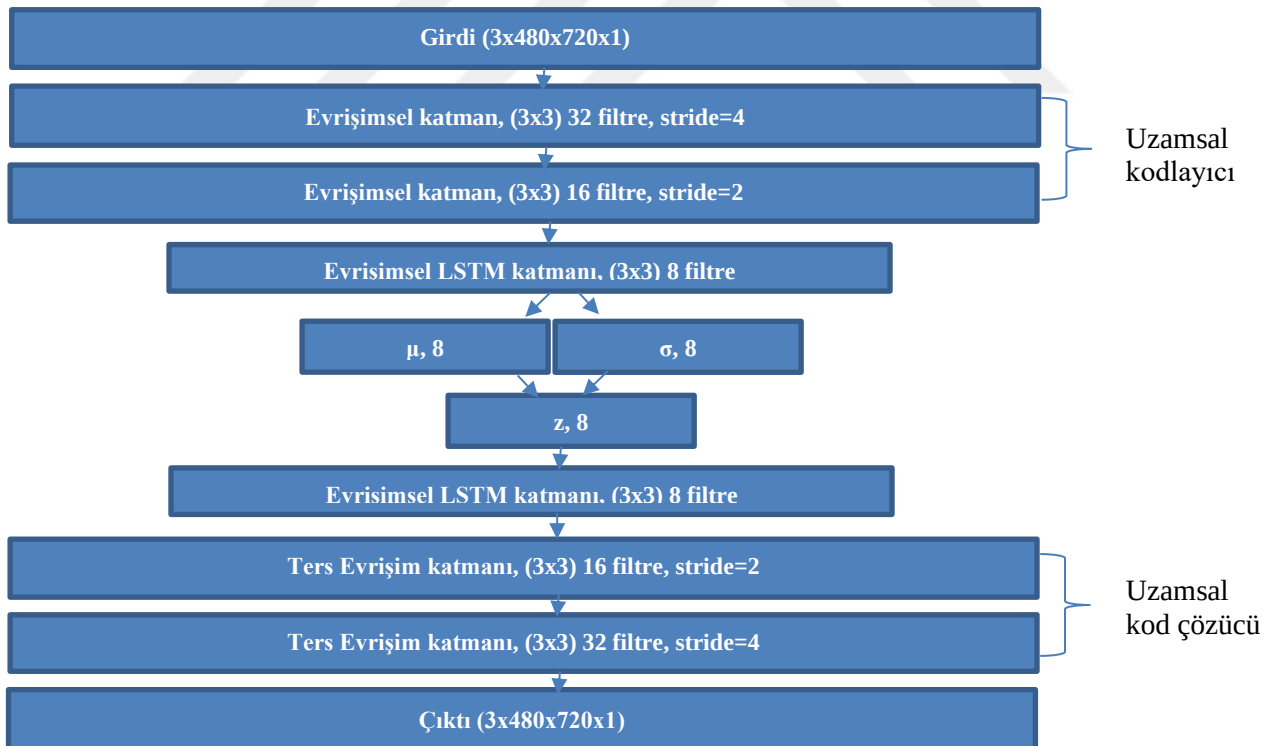
Konvolüsyonel bir ağdaki konvolüsyon işleminin temel amacı girdi görüntüsünün özniteliklerini çıkarmaktır. Konvolüsyon, girdi görüntüsünün küçük kare parçalarını kullanarak görüntü özniteliklerini öğrenip pikseller arasındaki uzamsal ilişkiyi korur.

Konvolüsyon işlemi matematiksel olarak küçük kare parçaları ile filtrenin noktasal çarpımı ile gerçekleştirilir. Filtrenin büyüklüğü arttıkça girdi görüntülerinden daha fazla öznitelik çıkarılır ve böylece ağ daha önce görülmemiş görüntülerde örüntüleri daha iyi ayırt eder. Ancak bu durum hesaplama süresinin artmasına ve belleğin daha hızlı tükenmesine neden olur. Bu nedenle bu iki durumun dengesini koruyacak bir ayarlama yapmak gerekir.

Konvolüsyonel LSTM ise geçmişteki verileride kullanarak özniteliklerin çıkarılmasını sağlar. Bu nedenle, çerçeveler üzerinde uzamsal özniteliklerin elde edilmesinde geçmişteki çerçevelerden de yararlanmak için bu yöntem tercih edilmiştir.

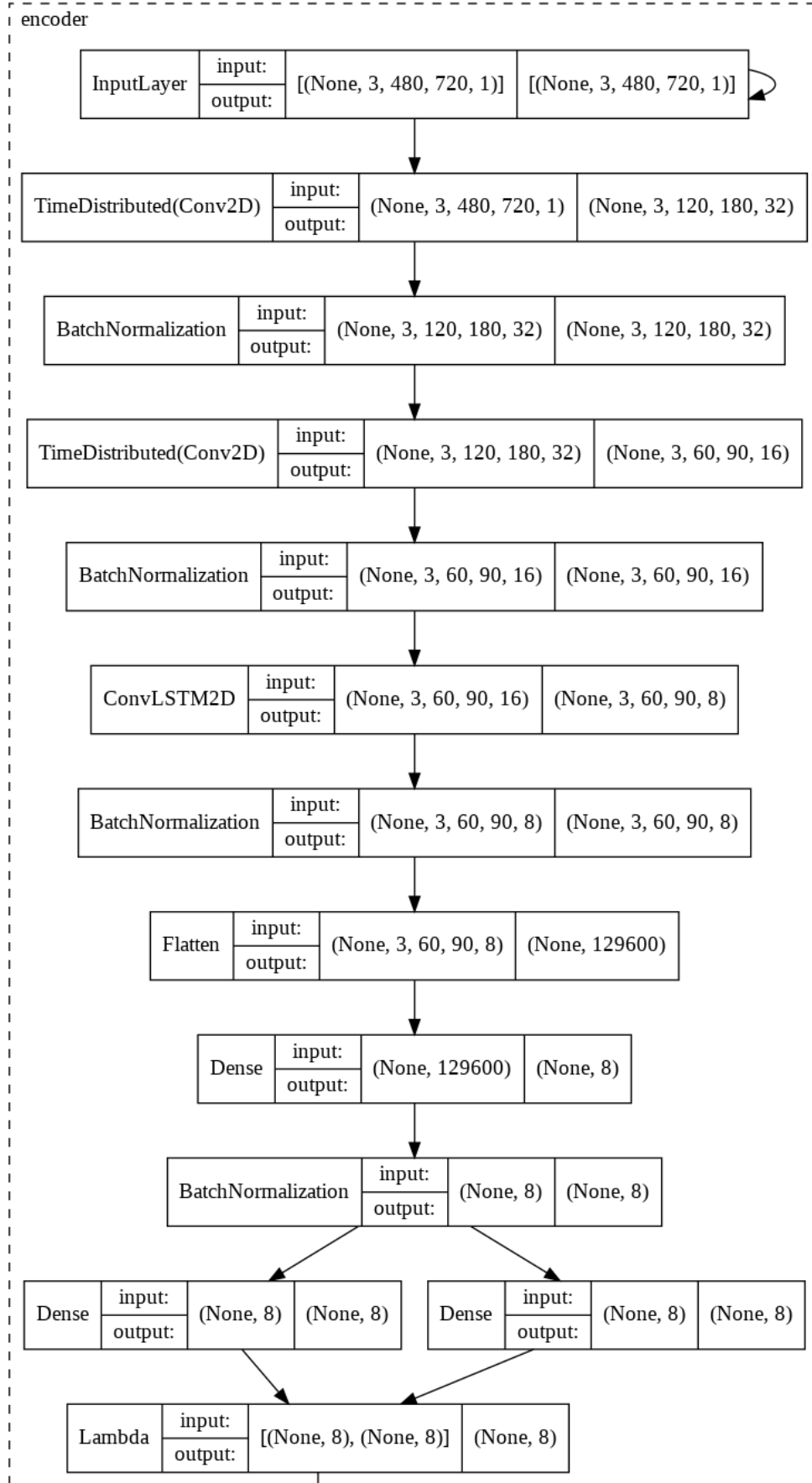
Özetle, varyasyonel otomatik kodlayıcının katmanlarına baktığımızda uzamsal öznitelikleri çıkarmak amacıyla konvolüsyonel, zamansal öznitelikleri çıkarmak amacıyla da konvolüsyonel LSTM katmanları kullanılmıştır.

Aşağıdaki şekilde bu çalışmada önerilen modelin genel mimarisi yer almaktadır. Modelin mimari yapısının kurulmasında ve geliştirilmesinde [94] numaralı çalışma referans alınmıştır.

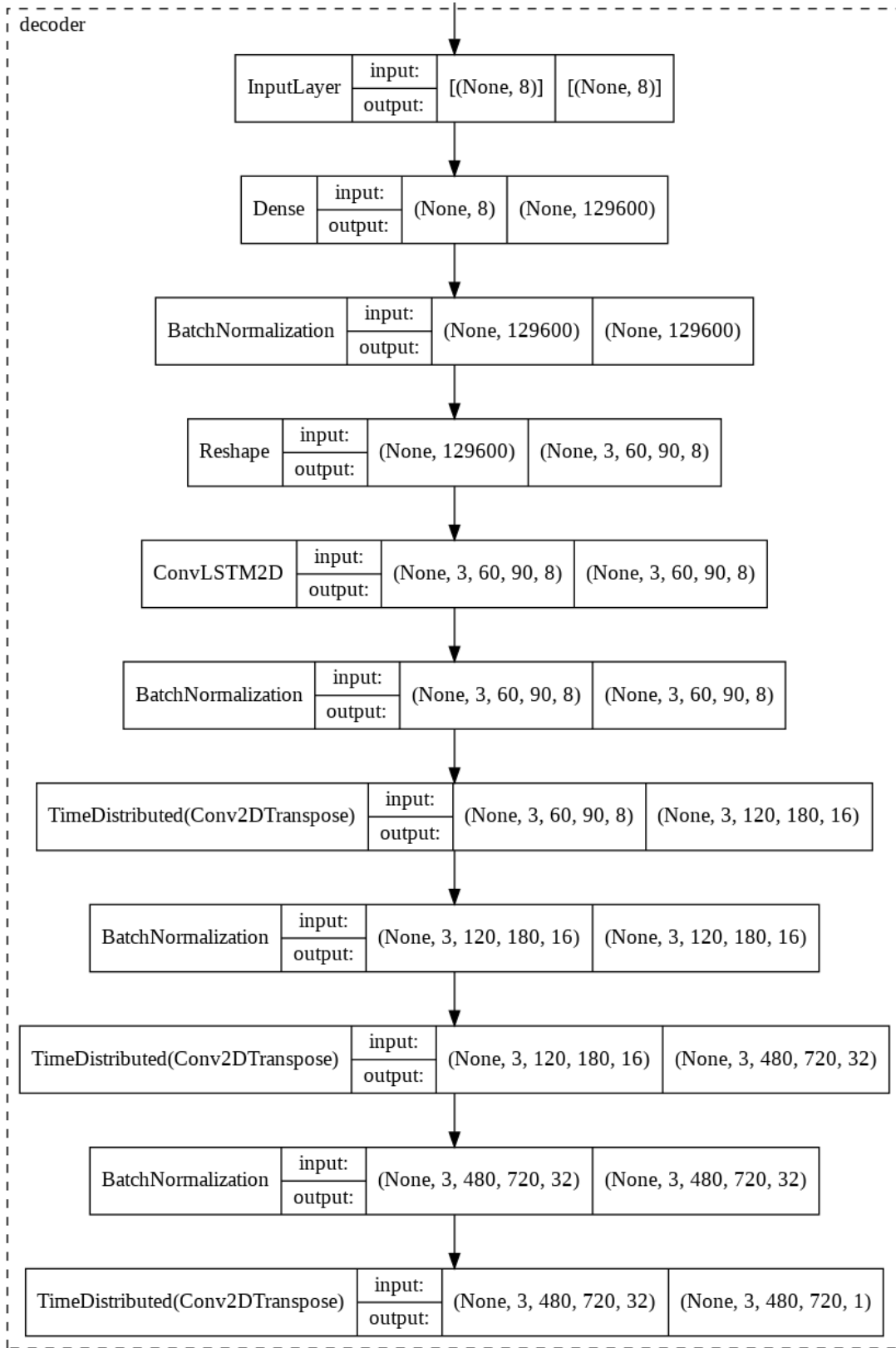


Şekil 9: Modelin genel mimarisi.

Bu modelin detaylı mimarisi ise aşağıda yer almaktadır.



Şekil 10: Modelin kodlayıcı bölümünün mimarisi.



Şekil 11: Modelin kod çözücü bölümünün mimarisi.

5. DENEYLER VE TARTIŞMA

Model eğitildikten sonra UCSD veri kümesinde yer alan test verileri kullanılarak modelin performansı değerlendirilmiştir. Bu değerlendirmeler grafikler ve çizelgeler kullanılarak sunulmuştur. Son olarak, bu çalışmadan çıkarılan vargılar eklenmiştir.

5.1 Değerlendirme Yöntemleri

Değerlendirme işlemi hem test veri kümesindeki her bir video için düzenlilik skoru grafiği çizilerek hem de tüm test veri kümesindeki videolar için zamansal anomali tespiti performansı ölçülerek yapılmıştır.

5.1.1 Düzenlilik Skoru Grafikleri

Test videoları için düzenlilik skoru grafiklerinin oluşturulmasında [64] numaralı çalışmada yer alan yöntem kullanılmıştır. Bu yöntemin detayları adım adım aşağıda anlatılmıştır.

İlk olarak model tarafından üretilen herhangi bir çerçevenin hata payı ölçümü tanımlanmıştır. Bunun için herhangi bir çerçevenin tüm piksellerinin yeniden oluşturulma hatası, girdi çerçevesi ile yeniden oluşturulan çerçeve arasındaki öklid mesafesi ile hesaplanmıştır.

Aşağıda t çerçevesi için yeniden oluşturulma hatası $e(t)$ hesaplanmıştır.

$$e(t) = \sum_{i \in S_t} \|x(i) - f_w(x(i))\|_2 \quad (5.1)$$

Burada S_t , t çerçevesindeki pikseller kümesini, f_w ise modelin tahmin ettiği piksel değerlerini ifade eder. Daha sonra her çerçeve için hesaplanan anormallik skorları $s_a(t)$ aşağıdaki formüle göre ölçeklendirilmiştir.

$$s_a(t) = \frac{e(t) - e(t)_{min}}{e(t)_{max}} \quad (5.2)$$

Ardından, düzenlilik skoru $s_r(t)$, anormallik skoru birden çıkarılarak basitçe türetilmiştir.

$$s_r(t) = 1 - s_a(t) \quad (5.3)$$

5.1.2 Zamansal Anomali Tespiti Performansı

Model tarafından oluşturulan her bir çerçevenin yeniden oluşturulma hata değeri kullanılarak çerçeveler için normal çerçeve veya anormal çerçeve şeklinde bir etiketleme yapılmıştır. Bunun için öncelikle eğitim çerçeveleri modele beslenerek modelin eğitim çerçevelerini tahminlemesi gerçekleştirilmiştir. Bu esnada eğitim çerçeveleri için oluşan yeniden oluşturma hataları histogram üzerinden gösterilmiştir ve çıkan sonuçlar gözlemlenmiştir. Sonrasında bu değerlerin ortalaması ve standart sapması kullanılarak bir çerçevenin normal çerçeve kategorisine sokulabilmesi için bir eşik değeri belirlenmiştir. Belirlenen bu eşik değerine göre test videosunda yer alan tüm çerçeveler için normal ve anormal etiketi atanmıştır. Böylece test veri kümesindeki çerçeve bazlı gerçek normal ve anormal etiketler ile belirlenen eşik değerine göre atanan değerler kullanılarak çerçeve bazlı yani zamansal anomali tespiti başarımları ölçülmüştür. Bu başarımları değerlendirmek için aşağıdaki doğruluk, kesinlik, duyarlılık, F1 skoru metriklerden yararlanılmıştır. Bu metrikleri ölçebilmek için öncelikle aşağıdaki değerlerin bulunması gerekmektedir.

DP: Doğru Pozitif – TP: True Positive

Model tarafından doğru kabul edilen değer gerçekte doğru değer olması.

YP: Yanlış Pozitif – FP: False Positive

Model tarafından doğru kabul edilen değer gerçekte yanlış değer olması.

DN: Doğru Negatif – TN: True Negative

Model tarafından yanlış kabul edilen değer gerçekte yanlış değer olması.

YN: Yanlış Negatif – FN: False Negative

Model tarafından yanlış kabul edilen değerin gerçekte doğru değeri olması.

5.1.3 Doğruluk (Accuracy)

Modelde doğru tahmin ettiğimiz alanların toplam veri kümesine oranı ile hesaplanmaktadır. Doğruluk, bir modelin başarısını ölçmek için çok kullanılan ancak tek başına yeterli olmadığı görülen bir metriktir. Bu nedenle aşağıdaki metrikler de incelenmiştir. Özellikle eşit dağılmayan veri kümelerinde sadece bu metriğe bakarak değerlendirme yapmak yanlış olacaktır. Değeri aşağıdaki formülle hesaplanır.

$$\text{Doğruluk} = (DP + DN) \div (DP + YP + DN + YN) \quad (5.4)$$

5.1.4 Kesinlik (Precision)

Kesinlik, pozitif olarak tahminlediğimiz değerlerin gerçekten kaç tanesinin pozitif olduğunu göstermektedir. Bu veri kümesinde pozitif değerler anomalinin gerçekleştiği çerçevelere karşılık gelmektedir. Değeri aşağıdaki formülle hesaplanır.

$$\text{Kesinlik} = (DP) \div (DP + YP) \quad (5.5)$$

5.1.5 Duyarlılık (Recall)

Duyarlılık, pozitif olarak tahmin etmemiz gereken işlemlerin ne kadarını pozitif olarak tahmin ettiğimizi gösteren bir metriktir. Değeri aşağıdaki formülle hesaplanır.

$$\text{Duyarlılık} = (DP) \div (DP + YN) \quad (5.6)$$

5.1.6 F1 Skoru (F1 Score)

F1 skoru değeri bize kesinlik ve duyarlılık değerlerinin harmonik ortalamasını göstermektedir. Basit bir ortalama yerine harmonik ortalama olmasının sebebi ise uç durumları da gözardı etmememiz içindir. Doğruluk yerine F1 skoru değerinin kullanılmasının en temel sebebi eşit dağılmayan veri kümelerinde doğru bir değerlendirme yapabilmek içindir. Değeri aşağıdaki formülle hesaplanır.

$$F1\ Skoru = 2 \times (Kesinlik \times Duyarlilik) \div (Kesinlik + Duyarlilik) \quad (5.7)$$

5.2 Deneylerin Gerçekleştirilmesi

Önerilen model Ped1 ve Ped2 veri kümeleri için ayrı ayrı denenmiştir. Her iki veri kümesi için elde edilen sonuçlar aşağıdaki gibidir.

5.2.1 Ped1 Veri Kümesi Üzerindeki Sonuçlar

Model Ped1 veri kümesinde yer alan eğitim videoları ile eğitilmiştir. Daha sonra, eğitilen modelin aynı veri kümesinde yer alan test videoları üzerindeki sonuçları paylaşılmıştır.

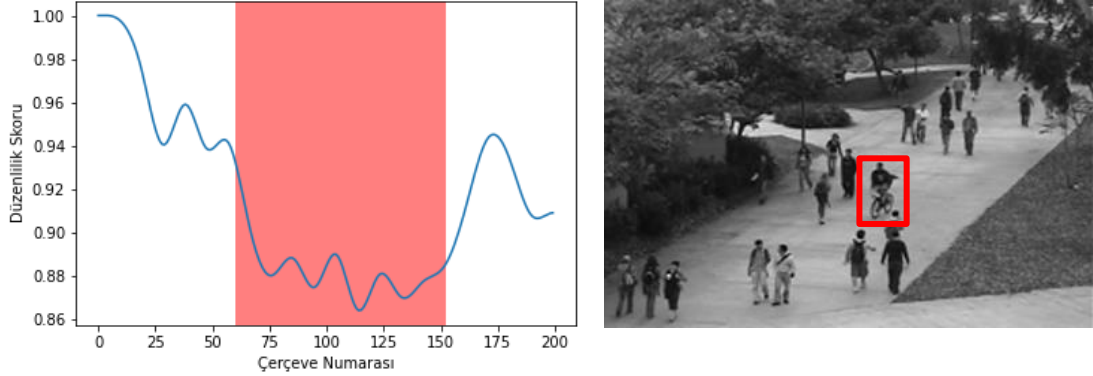
5.2.1.1 Ped1 Düzenlilik Skoru Grafikleri

36 tane test videosuna sahip olan Ped1 veri kümesi için 4 tanesinin düzenlilik skoru grafikleri ile birlikte açıklamaları aşağıdaki gibidir. Test veri kümesinde yapılan etiketleme sonucu anormal olan kısımlar, grafikler üzerinde kırmızı arka plan olacak şekilde gösterilmiştir.

Test001 numaralı test videosunun incelenmesi

Test001 numaralı test videosunu incelediğimizde 60. çerçeveden itibaren videoya bir bisikletlinin girerek anomaliye sebep olduğunu gözlemledik. Bu bisikletli yaya yolu üzerindeki hareketine devam ederek 152. çerçevede yaya yolundan ayrılmaktadır. Test videosuna ait etiketlemeye baktığımızda da 60 ile 152. çerçeveler arasının anormal olarak etiketlendiğini gördük. Test videosunda anormal olarak etiketlenen

kısımları biz de grafik üzerinde kırmızı alan olarak gösterdik. Aşağıda model sonucunda elde edilen düzenlilik skoru grafiği ile test videosuna ait anormal olayın gerçekleştiği zaman diliminden bir çerçeve yer almaktadır.

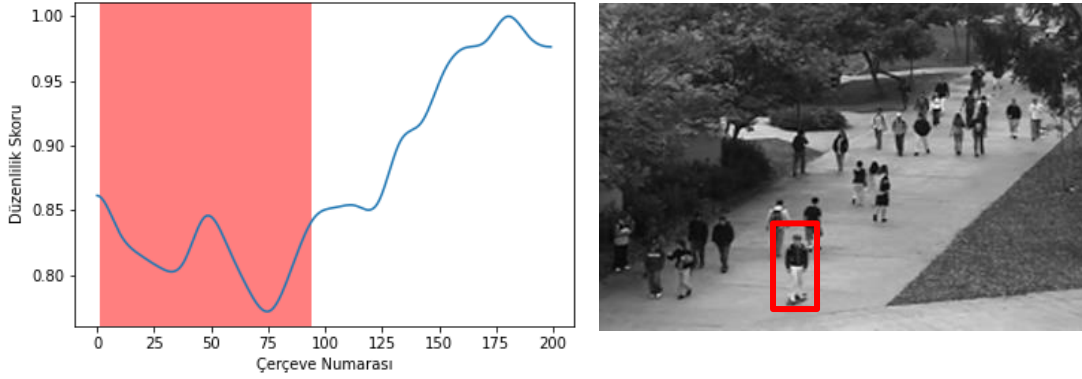


Şekil 12: UCSD Ped1 veri kümesindeki Test001 numaralı test verisinin düzenlilik skoru grafiği, 120. çerçeve kesiti.

Yukarıdaki grafiği incelediğimizde test veri kümesinde 60. çerçeveden 152. çerçeveye kadar olan kısım anormalliği ifade ettiği için kırmızı arka plan olarak gösterilmiştir. Eğittiğimiz model sonucunda oluşan düzenlilik eğrisini incelediğimizde ise test kümesinin düzenliliğin başlangıçtan itibaren yavaş yavaş azalmaya başlayıp 60 ile 152. çerçeveler arasında düzenliliğin en düşük seviyelerde olduğunu görmekteyiz. Bu aralıkta anormal bir durumun gözlemlendiği düzenlilik eğrisinden net bir şekilde gözlemlenebilmektedir. Ayrıca, oluşan düzenlilik skoru eğrisinin 152. çerçeveden itibaren artışa geçtiği yani veri kümesinin normalleşmeye başladığı da yine grafik üzerinden çok rahat bir şekilde görülmektedir.

Test008 numaralı test videosunun incelenmesi

Test008 numaralı test videosunu incelediğimizde 1. çerçeveden itibaren videoda kayak yapan birinin anormaliye sebep olduğunu gözlemledik. Bu kayaklı yaya yolu üzerindeki hareketine devam ederek 94. çerçevede yaya yolundan ayrılmaktadır. Test videosuna ait etiketlemeye baktığımızda da 1 ile 94. çerçeveler arasının anormal olarak etiketlendiğini gördük. Test videosunda anormal olarak etiketlenen kısımları biz de grafik üzerinde kırmızı alan olarak gösterdik. Aşağıda model sonucunda elde edilen düzenlilik skoru grafiği ile test videosuna ait anormal olayın gerçekleştiği zaman diliminden bir çerçeve yer almaktadır.

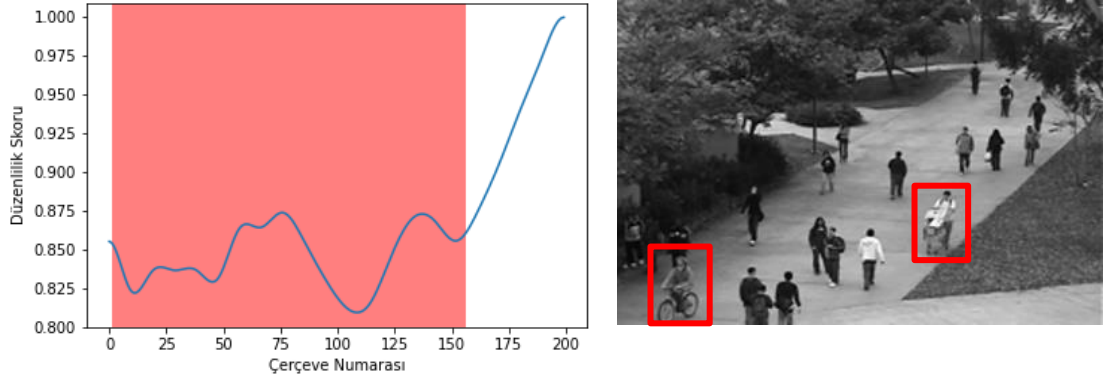


Şekil 13: UCSD Ped1 veri kümesindeki Test008 numaralı test verisinin düzenlilik skoru grafiği, 72. çerçeve kesiti.

Yukarıdaki grafiği incelediğimizde test veri kümesinde 1. çerçeveden 94. çerçeveye kadar olan kısım anormalliği ifade ettiği için kırmızı arka plan olarak gösterilmiştir. Eğittiğimiz model sonucunda oluşan düzenlilik eğrisini incelediğimizde ise test kümesinin düzenliliğin videonun geri kalanına göre düşük bir değerle başlayıp 75. çerçeve civarında en düşük düzenlilik değerine sahip olduğunu görmekteyiz. Bu aralıkta anormal bir durumun gözlemlendiği düzenlilik eğrisinden net bir şekilde gözlemlenebilmektedir. Ayrıca, kayaklının videodan ayrılmaya başladığı 94. çerçeveden itibaren düzenlilik skorunun devamlı bir yükselişte olduğu yani veri kümesinin normalleşmeye başladığı da yine grafik üzerinden çok rahat bir şekilde görülmektedir.

Test013 numaralı test videosunun incelenmesi

Test013 numaralı test videosunu incelediğimizde 1. çerçeveden itibaren videoda bir bisikletli ile bir bebek arabası sürücüsünün yer alarak anomaliye sebep olduğunu gözlemledik. Bu bisikletli yaya yolu üzerindeki hareketine devam ederek 53. çerçevede yaya yolundan ayrılırken, bebek arabası sürücüsü 156. çerçevede yaya yolundan ayrılmaktadır. Test videosuna ait etiketlemeye baktığımızda da 1 ile 156. çerçeveler arasının anormal olarak etiketlendiğini gördük. Test videosunda anormal olarak etiketlenen kısımları biz de grafik üzerinde kırmızı alan olarak gösterdik. Aşağıda model sonucunda elde edilen düzenlilik skoru grafiği ile test videosuna ait anormal olayın gerçekleştiği zaman diliminden bir çerçeve yer almaktadır.



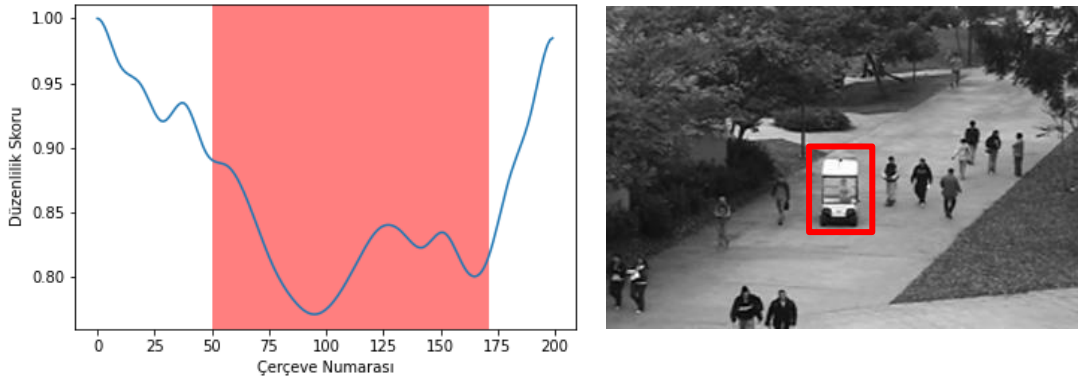
Şekil 14: UCSD Ped1 veri kümesindeki Test013 numaralı test verisinin düzenlilik skoru grafiği, 43. çerçeve kesiti.

Yukarıdaki grafiği incelediğimizde test veri kümesinde 1. çerçeveden 156. çerçeveye kadar olan kısım anormalliği ifade ettiği için kırmızı arka plan olarak gösterilmiştir. Eğittiğimiz model sonucunda oluşan düzenlilik eğrisini incelediğimizde ise test kümesinin düzenliliğin başlangıçtan itibaren videonun geri kalan kısmına göre düşük bir değerle başlayıp 100. çerçeve civarında en düşük değerine ulaştığını görmekteyiz. 100. Çerçeveden itibaren başlangıçtaki düşük değerini koruyarak artmaya başladığını söyleyebiliriz. Ancak 156. çerçeveden itibaren başlangıçtaki düşük düzenlilik skorunun değerini aşarak videonun düzenlilik skorunun giderek arttığını görebiliriz. Bu nedenle, 1. çerçeve ile 156. çerçeve arasında anormal bir durumun gözlemlendiği düzenlilik eğrisinden net bir şekilde gözlemlenebilmektedir. Ayrıca, oluşan düzenlilik skoru eğrisinin 156. çerçeveden itibaren anormallik skorunun üzerinde artışa geçtiği yani veri kümesinin normalleşmeye başladığı da yine grafik üzerinden çok rahat bir şekilde görülmektedir.

Test024 numaralı test videosunun incelenmesi

Test024 numaralı test videosunu incelediğimizde 50. çerçeveden itibaren videoya bir arabalının girerek anomaliye sebep olduğunu gözlemledik. Bu arabalı yaya yolu üzerindeki hareketine devam ederek 171. çerçevede yaya yolundan ayrılmaktadır. Test videosuna ait etiketlemeye baktığımızda da 50 ile 171. çerçeveler arasının anormal olarak etiketlendiğini gördük. Test videosunda anormal olarak etiketlenen kısımları biz de grafik üzerinde kırmızı alan olarak gösterdik. Aşağıda model

sonucunda elde edilen düzenlilik skoru grafiği ile test videosuna ait anormal olayın gerçekleştiği zaman diliminden bir çerçeve yer almaktadır.



Şekil 15: UCSD Ped1 veri kümesindeki Test024 numaralı test verisinin düzenlilik skoru grafiği, 120. çerçeve kesiti.

Yukarıdaki grafiği incelediğimizde test veri kümesinde 50. çerçeveden 171. çerçeveye kadar olan kısım anormalliği ifade ettiği için kırmızı arka plan olarak gösterilmiştir. Eğittiğimiz model sonucunda oluşan düzenlilik eğrisini incelediğimizde ise test kümesinin düzenliliğin başlangıçtan itibaren yavaş yavaş azalmaya başlayıp 50 ile 171. çerçeveler arasında düzenliliğin en düşük seviyelerde olduğunu görmekteyiz. Bu aralıkta anormal bir durumun gözlemlendiği düzenlilik eğrisinden net bir şekilde gözlemlenebilmektedir. Ayrıca, oluşan düzenlilik skoru eğrisinin 151. çerçeveden itibaren artışa geçtiği yani veri kümesinin normalleşmeye başladığı da yine grafik üzerinden çok rahat bir şekilde görülmektedir.

Ortaya çıkan düzenlilik skoru grafiklerini incelediğimizde, çerçevelerin sadece 8’de 1’inin üzerinde yer alan hareket vektörlerini kullanarak yapmış olduğumuz bu çalışmada anlamlı sonuçlar elde ettiğimiz görülmektedir.

5.2.1.2 Ped1 Zamansal Anomali Tespiti Performansı

Zamansal anomali tespiti yapmadan önce, eğitim veri kümesinden, çok fazla yayanın yer aldığı, bu sebeple yeniden oluşturma hatası oldukça yüksek çıkan veriler temizlenmiştir. Böylece anormal olan çerçevelerin daha iyi bir şekilde tespit edilmesi sağlanmıştır.

Öncelikle eğitilen modele eğitim veri kümesi girdi olarak verilip modelin eğitim veri kümesi için tahminleme yapması sağlanmıştır. Sonrasında modelin ürettiği çerçeveler ile eğitim veri kümesi çerçeveleri arasındaki yeniden oluşturma hataları bulunmuştur.

Daha sonra, bu hata değerlerinin ortalama ve standart sapma değerleri hesaplanmıştır. Bu değerler kullanılarak normal bir çerçevenin sahip olabileceği maksimum ve minimum yeniden oluşturma hata değerleri belirlenmeye çalışılmıştır. Bu yaklaşımla, bu sınırların dışında yeniden oluşturma hatasına sahip olan çerçeveler anormal olarak belirlenmiştir. Yapılan denemeler sonucunda bu sınırların eğitim verisi yeniden oluşturma hata değerlerinin ortalamasının ± 1 standart sapma aralığına denk geldiği görülmüştür. Normal şartlarda anormal bir verinin yeniden oluşturma hatasının normal bir verinin yeniden oluşturma hatasından daha küçük olmaması beklenmektedir. Bu sebeple belirlemiş olduğumuz alt sınır çok anlamlı görülme de, pratikte zaten böyle bir durumla karşılaşmamaktadır. Modelin performansını ölçmek amacıyla, test veri kümesinin etiketli çerçeve değerleri ile model tarafından tahmin edilen değerler kullanılarak aşağıdaki metrikler hesaplanmıştır.

Çizelge 1: UCSD Ped1 veri kümesi zamansal anomali tespiti performansı.

Çerçeve Türü	Kesinlik	Duyarlılık	F1 Skoru	Doğruluk
Normal	0,53	0,80	0,64	
Anormal	0,74	0,44	0,55	
				0,60

Anormal etiketli veriler için kesinlik değeri 0,74 olarak hesaplanmıştır. Bu da, anormal olarak tahmin ettiğimiz çerçevelerin %74'ünün gerçekte de anormal olduğu anlamına gelmektedir. Diğer taraftan, duyarlılık değerinin 0,44 olduğu görülmektedir. Bu da, gerçekte anormal olan çerçevelerin modelimiz tarafından anormal olarak tespit edilme performansının iyi olmadığı anlamına gelmektedir. Test veri kümesi incelendiğinde, anormal olarak etiketli olan çerçevelerdeki anormalliklerin normal yaya hareketinden ayırt edilmesinin oldukça zor olduğu görülmektedir. Örnek olarak, yaya yoluna neredeyse yaya ile aynı hızda ve aynı yönde hareket eden bisikletlinin veya kaykaylının girdiği görülmektedir. Test veri kümesinde benzer şekilde anormal olarak etiketlenen birçok veri vardır. Ayrıca, bizim modelimiz sadece hareket vektörlerinin büyüklüğüne ve yönüne dayanarak anomaliyi tahminlemeye çalıştığı için bu tür anomalileri tespit etmekte oldukça düşük performans göstermiştir.

Normal etiketli veriler için duyarlılık değeri 0,80 olarak görülmektedir. Buna göre, gerçekte normal olan çerçevelerin %80'i modelimiz tarafından normal olarak tahmin edilmektedir. Diğer taraftan, kesinlik değeri 0,53 olarak hesaplanmıştır. Bu da, normal olarak tespit edilen çerçevelerin %53'ü gerçekte normal anlamına gelmektedir. Burdaki düşük performansın da yine yukarıda ifade edilen durumdan dolayı

kaynaklandığı düşünülmektedir. Yani, normal duruma çok benzerlik gösteren anormal olaylar model tarafından çok iyi yakalanamamaktadır. Ayrıca doğruluk değeri de 0,60 olarak hesaplanmıştır. Bu durum, yapmış olduğumuz tahminlerin %60'ının gerçeğe uyduğu anlamına gelmektedir.

5.2.2 Ped2 Veri Kümesi Üzerindeki Sonuçlar

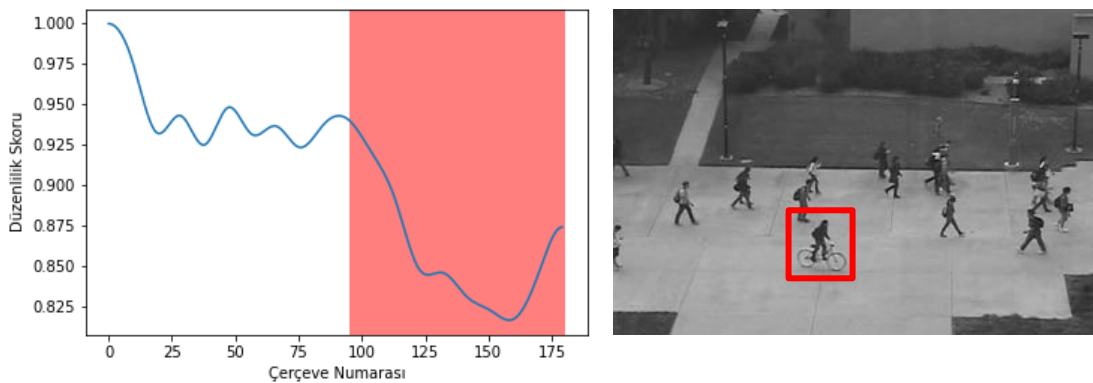
Model Ped2 veri kümesinde yer alan eğitim videoları ile eğitilmiştir. Daha sonra, eğitilen modelin aynı veri kümesinde yer alan test videoları üzerindeki sonuçları paylaşılmıştır.

5.2.2.1 Ped2 Düzenlilik Skoru Grafikleri

12 tane test videosuna sahip olan Ped1 veri kümesi için 2 tanesinin düzenlilik skoru grafikleri ile birlikte açıklamaları aşağıdaki gibidir. Test veri kümesinde yapılan etiketleme sonucu anormal olan kısımlar, grafikler üzerinde kırmızı arka plan olacak şekilde gösterilmiştir.

Test002 numaralı test videosunun incelenmesi

Test002 numaralı test videosunu incelediğimizde 95. çerçeveden itibaren videoya bir bisikletlinin girerek anomaliye sebep olduğunu gözlemledik. Bu bisikletli yaya yolu üzerindeki hareketine videonun sonuna kadar devam etmektedir. Test videosuna ait etiketlemeye baktığımızda da 95. çerçeveden 180. çerçeveye kadar yani videonun sonuna kadar anormal olarak etiketlendiğini gördük. Test videosunda anormal olarak etiketlenen kısımları biz de grafik üzerinde kırmızı alan olarak gösterdik. Aşağıda model sonucunda elde edilen düzenlilik skoru grafiği ile test videosuna ait anormal olayın gerçekleştiği zaman diliminden bir çerçeve yer almaktadır.

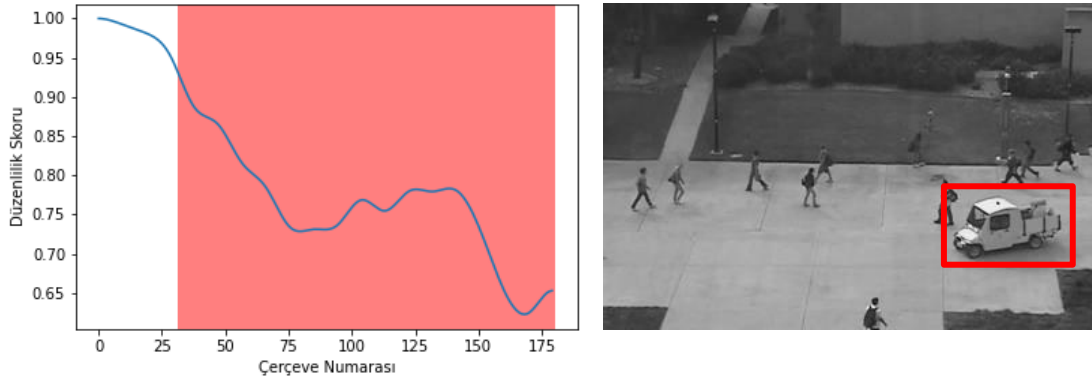


Şekil 16: UCSD Ped2 veri kümesindeki Test002 numaralı test verisinin düzenlilik skoru grafiği, 140. çerçeve kesiti.

Yukarıdaki grafiği incelediğimizde test veri kümesinde 95. çerçeveden 180. çerçeveye kadar olan kısım anormalliği ifade ettiği için kırmızı arka plan olarak gösterilmiştir. Eğittiğimiz model sonucunda oluşan düzenlilik eğrisini incelediğimizde ise test kümesinin düzenliliğin başlangıçtan itibaren yavaş yavaş azalmaya başlayıp 95 ile 180. çerçeveler arasında düzenliliğin en düşük seviyelerde olduğunu görmekteyiz. Bu aralıkta anormal bir durumun gözlemlendiği düzenlilik eğrisinden net bir şekilde gözlemlenebilmektedir. Ayrıca, başlangıçta düzenlilik skorunun 1 değerinden başlaması da videonun en düzenli halinden başlayarak zaman içerisinde anormal bir durumun meydana geldiğini göstermektedir.

Test004 numaralı test videosunun incelenmesi

Test004 numaralı test videosunu incelediğimizde 31. çerçeveden itibaren videoya bir arabalının girerek anomaliye sebep olduğunu gözlemledik. Bu arabalı yaya yolu üzerindeki hareketine videonun sonuna kadar devam etmektedir. Test videosuna ait etiketlemeye baktığımızda da 31. çerçeveden 180. çerçeveye kadar yani videonun sonuna kadar anormal olarak etiketlendiğini gördük. Test videosunda anormal olarak etiketlenen kısımları biz de grafik üzerinde kırmızı alan olarak gösterdik. Aşağıda model sonucunda elde edilen düzenlilik skoru grafiği ile test videosuna ait anormal olayın gerçekleştiği zaman diliminden bir çerçeve yer almaktadır.



Şekil 17: UCSD Ped2 veri kümesindeki Test004 numaralı test verisinin düzenlilik skoru grafiği, 80. çerçeve kesiti.

Yukarıdaki grafiği incelediğimizde test veri kümesinde 31. çerçeveden 180. çerçeveye kadar olan kısım anormalliği ifade ettiği için kırmızı arka plan olarak gösterilmiştir. Eğittiğimiz model sonucunda oluşan düzenlilik eğrisini incelediğimizde ise test kümesinin düzenliliğin başlangıçtan itibaren yavaş yavaş azalmaya başlayıp 31 ile 180. çerçeveler arasında düzenliliğin en düşük seviyelerde olduğunu görmekteyiz. Bu

aralıkta anormal bir durumun gözlemlendiği düzenlilik eğrisinden net bir şekilde gözlemlenebilmektedir. Ayrıca, başlangıçta düzenlilik skorunun 1 değerinden başlaması da videonun en düzenli halinden başlayarak zaman içerisinde anormal bir durumun meydana geldiğini göstermektedir.

Ortaya çıkan düzenlilik skoru grafiklerini incelediğimizde, çerçevelerin sadece sekizde birinin üzerinde yer alan hareket vektörlerini kullanarak yapmış olduğumuz bu çalışmada anlamlı sonuçlar elde ettiğimiz görülmektedir.

5.2.2.2 Ped2 Zamansal Anomali Tespiti Performansı

Zamansal anomali tespiti yapmadan önce, eğitim veri kümesinden, çok fazla yayanın yer aldığı, bu sebeple yeniden oluşturma hatası oldukça yüksek çıkan veriler temizlenmiştir. Böylece anormal olan çerçevelerin daha iyi bir şekilde tespit edilmesi sağlanmıştır.

Öncelikle eğitilen modele eğitim veri kümesi girdi olarak verilip modelin eğitim veri kümesi için tahminleme yapması sağlanmıştır. Sonrasında modelin ürettiği çerçeveler ile eğitim veri kümesi çerçeveleri arasındaki yeniden oluşturma hataları bulunmuştur. Daha sonra, bu hata değerlerinin ortalama ve standart sapma değerleri hesaplanmıştır. Bu değerler kullanılarak normal bir çerçevenin sahip olabileceği maksimum ve minimum yeniden oluşturma hata değerleri belirlenmeye çalışılmıştır. Bu yaklaşımla, bu sınırların dışında yeniden oluşturma hatasına sahip olan çerçeveler anormal olarak belirlenmiştir. Yapılan denemeler sonucunda bu sınırların eğitim verisi yeniden oluşturma hata değerlerinin ortalamasının ± 2 standart sapma aralığına denk geldiği görülmüştür. Normal şartlarda anormal bir verinin yeniden oluşturma hatasının normal bir verinin yeniden oluşturma hatasından daha küçük olmaması beklenmektedir. Bu sebeple belirlemiş olduğumuz alt sınır çok anlamlı görülme de, pratikte zaten böyle bir durumla karşılaşılmamaktadır. Modelin performansını ölçmek amacıyla, test veri kümesinin etiketli çerçeve değerleri ile model tarafından tahmin edilen değerler kullanılarak aşağıdaki metrikler hesaplanmıştır.

Çizelge 2: UCSD Ped2 veri kümesi zamansal anomali tespiti performansı.

Çerçeve Türü	Kesinlik	Duyarlılık	F1 Skoru	Doğruluk
Normal	0,36	0,27	0,31	
Anormal	0,85	0,90	0,87	
				0,78

Anormal etiketli veriler için kesinlik değeri 0,85 olarak hesaplanmıştır. Bu da, anormal olarak tahmin ettiğimiz çerçevelerin %85'inin gerçekte de anormal olduğu anlamına gelmektedir. Yani, modelin anormal olarak tahmin ettiği çerçevelerin gerçekte de anormal olma oranı oldukça yüksektir. Diğer taraftan, duyarlılık değerinin 0,90 olduğu görülmektedir. Bu sonuç, gerçekte anormal olan çerçevelerin modelimiz tarafından anormal olarak tespit edilme performansının oldukça iyi olduğu anlamına gelmektedir.

Normal etiketli veriler için duyarlılık değeri 0,27 olarak görülmektedir. Buna göre, gerçekte normal olan çerçevelerin %27'sinin modelimiz tarafından normal olarak tahmin edilmektedir. Diğer taraftan, kesinlik değeri 0,36 olarak hesaplanmıştır. Bu da, normal olarak tespit edilen çerçevelerin %36'sı gerçekte normal anlamına gelmektedir. Burdaki düşük performansın test veri kümesindeki dengesiz dağılımdan kaynaklandığı düşünülmektedir. Test veri kümesindeki videoların çerçevelerin %85'i anormaldir. Çok az sayıda normal çerçevenin olması ve bu çerçevelerin grup halinde modele sunulmasından dolayı normal ve anormal çerçeveler arasındaki çok ani geçişler yakalanamamıştır. Ayrıca doğruluk değeri de 0,78 olarak hesaplanmıştır. Bu durum, yapmış olduğumuz tahminlerin %78'inin gerçekte uyuştugu anlamına gelmektedir.

5.2.3 Geçmiş Çalışmalar İle Performans Karşılaştırılması

Bu bölümde, videoda anomali tespiti için, bu çalışmada önerilen yöntemin elde ettiği sonuç ile daha önceki çalışmalardan elde edilen sonuçlar kıyaslanacaktır. Bu çalışmada, önceki bölümlerde belirtildiği üzere, sadece video çerçevelerinde yer alan hareket vektörlerinden yararlanılmıştır. Kıyaslamamanın yapılacağı diğer yöntemler video çerçevelerinde yer alan tüm piksel bazlı bilgiler kullanılarak gerçekleştirilmiştir. Dolayısı ile bu çalışmada girdi olarak çok daha kısıtlı veri ile anlamlandırma yapılmaya çalışıldığı açıktır.

Karşılaştırma işlemi için AUC (Area Under Curve) metriği kullanılmıştır. AUC metriği, sınıflandırma problemlerinde iki sınıfın birbirinden ne kadar iyi ayrılabilmesini gösteren bir metriktir. Bu metriğin elde ettiği sonuç ne kadar yüksek çıkarsa modelin o kadar iyi sınıflandırma yaptığı söylenebilir.

AUC metriğinin hesaplanabilmesi için öncelikle ROC eğrisinin bulunması gerekmektedir. AUC metriği ROC (Receiver Operating Characteristic) eğrisinin

altında kalan alana eşittir. ROC eğrisi yatay ekseninde doğru pozitif oranı, düşey ekseninde ise yanlış pozitif oranı yer alacak şekilde oluşturulur. Doğru pozitif oranı ile yanlış pozitif oranının formülleri aşağıdaki gibi verilmiştir. Bu değerlerin kullanılması ile oluşan ROC eğrisi oluşturulur. Bu eğrinin altında kalan alan ile de AUC değeri hesaplanmıştır.

$$\text{Doğru Pozitif Oranı} = DP \div (DP + YN) \quad (5.8)$$

$$\text{Yanlış Pozitif Oranı} = YP \div (YP + DN) \quad (5.9)$$

ROC eğrisi için modelin ürettiği çerçeve bazlı tahminleme sonuçları kullanılmıştır. Dolayısıyla elde edilen AUC değeri de çerçeve bazlıdır. AUC değeri hem Ped1 hem de Ped2 veri kümeleri için ayrı ayrı hesaplanmıştır. Aşağıda bu veri kümelerini kullanarak videoda anomali tespiti işlemi gerçekleştirmiş olan yöntemlerin AUC metriği değerleri ile bu çalışma sonucunda elde ettiğimiz AUC değerlerini içeren çizelge yer almaktadır [2].

Çizelge 3: Modellerin UCSD veri kümesi üzerinde elde ettiği AUC değerleri.

Model	UCSD Ped1 AUC	UCSD Ped2 AUC
Adam	% 65.0	% 63.0
Social force	% 67.5	% 63.0
MPPCA	% 59.0	% 77.0
Social force + MPPCA	% 67.0	% 71.0
MDT	% 81.8	% 85.0
Video parsing	% 91.0	% 92.0
AMDN	% 92.1	% 90.8
ST video parsing	% 93.9	% 94.6
App+motion cues	% 85.0	% 90.0
Conv-AE	% 81.0	% 90.0
Compact feature sets	% 82.0	% 84.0
Conv-WTA-AE	% 91.9	% 92.8
RBM	% 70.3	% 86.4
Convex polytope ensembles	% 78.2	% 80.7
GAN	% 97.4	% 93.5
Online-GNG	% 93.8	% 94.0
Future frame prediction	% 83.1	% 95.4
Plug and play CNN	% 95.7	% 88.4
Fast SCL	% 93.8	% 95.0

MLAD0+3	% 82.3	% 99.2
Prediction+reconstruction	% 82.6	% 96.2
NN on video patch FG masks	% 77.3	% 88.3
Siamese distance learning	% 86.0	% 94.0
Önerilen model	% 63.5	% 58.7

Yukarıdaki çizelge incelendiğinde önerilen modelin AUC değeri Ped1 veri kümesinde %63.5 iken Ped2 veri kümesinde %58.7 olmuştur. Bu sonuçlar video çerçevelerindeki tüm piksel bilgilerini kullanarak yapılan çalışmaların oldukça altında kalmıştır. Bu durum, videoda anomali tespiti işleminde sadece hareket vektörlerinin kullanılması ile oluşturulan modelin piksel bazlı öğrenen modellere göre düşük performans sağladığını göstermiştir.

5.3 Vargılar ve Tartışma

Bu çalışmada, bir videodaki normallik anormallik durumunu ortaya çıkarabilmek için videoyu oluşturan çerçevelerdeki tüm bilgileri kullanmak yerine sadece bu çerçevelerdeki hareket vektörlerinin kullanılması verinin büyük çoğunluğunu kullanmak yerine sadece verinin daha önemli olan çok daha az bir kısmı ile bir sonuç çıkarılması amaçlanmıştır. Üstelik sadece asıl hareket vektörlerinin ortaya çıktığı 8 ve katlarında yer alan çerçevelerin kullanılıp diğerlerinin elenmesi durumu verinin tekrardan önemli çoğunluktaki bir kısmını daha dışarıda bırakmıştır. Tüm bunlar bu çalışmayı diğer çalışmalardan ayıran en büyük özelliğin çok kısıtlı miktardaki veri ile çalışılıp çıkarım yapılmaya çalışıldığını göstermektedir.

Tüm bu durumlar göz önüne alınarak elde edilen sonuçlar incelendiğinde test videolarındaki düzenlilik grafiklerinin ortaya koyduğu sonuçların, videonun genel izlenimi hakkında gerçekçi çıkarımları ile ciddi bir başarı yakaladığı söylenebilir. Ayrıca, çerçeve bazlı yani zamansal anomali tespiti başarımların değerleri de her ne kadar video verilerinin tamamını kullanan modeller kadar yüksek olmasada, çok az miktardaki veri ile ortaya koyduğu sonuç dikkate değer bir seviyededir.



6. SONUÇ

Bu tezde, sıkıştırılmış videoda meydana gelen hareket vektörleri kullanılarak anomali tespit çalışması gerçekleştirilmiştir. Videoda anomali tespiti üzerine birçok çalışma yapılmışken bu çalışmayı diğerlerinden ayıran özellik, çalışmanın sıkıştırılmış alanda gerçekleştirilmiş olmasıdır. Sıkıştırma işlemi için High Efficiency Video Coding (HEVC) standardı kullanılmıştır. Bu standarda uygun olarak sıkıştırılan videolar açık kaynak kodlu, GitlHevcAnalyzer analiz aracı kullanılarak incelenmiştir. Bu araç sayesinde hareket vektörlerinin yer aldığı çerçeveler elde edilmiştir. Sadece hareket vektörlerinden yararlanmak amacıyla bu çerçeveler çeşitli ön işlem aşamalarından geçirilmiştir. Böylece, veri kümesi hazır hale getirilmiştir. Daha sonra, videoda anomali tespiti gerçekleştirmek için varyasyonel otomatik kodlayıcı tabanlı model kullanılmıştır. Model performansını değerlendirmek için iki farklı yöntem kullanılmıştır. İlk olarak, her bir test videosu için çerçeve bazlı düzenlilik skoru grafiği çıkarılmıştır. Bu grafik sayesinde videonun hangi çerçevesinden itibaren anormal bir durum gerçekleştiği, hangi çerçevelerden itibaren normale döndüğü gözlemlenebilmektedir. Oluşturulan düzenlilik skoru grafikleri incelendiğinde video akışına göre genel olarak tutarlı sonuçlar gösterdiği gözlemlenmiştir. İkinci olarak da, model tarafından yapılan tahminleme ile gerçek veri etiketleri kullanılarak modelin çerçeve bazlı tahminleme performansı değerlendirilmiştir. Çerçeve bazlı etiketlemenin sonuçları incelendiğinde ise, UCSD Ped1 veri kümesinde anormal çerçevelerin %60'lık kısmının, UCSD Ped2 veri kümesinde ise anormal çerçevelerin %78'lik kısmının tespit edilebildiği görülmüştür. Sonuç olarak, bir videonun çerçevelerindeki tüm bilgileri kullanmak yerine, sadece bu çerçevelerdeki hareket vektörleri kullanılarak yapılmış olan bu çalışmanın etkin sonuçlar çıkardığı gözlemlenmiştir. Bu çalışmanın devamı olarak hareket vektörlerinin konumlarına yönelik dağılımlar kullanılarak model üzerinde geliştirmeler yapılabilir. Ayrıca, modelin elde ettiği performansı iyileştirmeye yönelik çalışmalar da gerçekleştirilebilir.



KAYNAKLAR

- [1] **Saligrama, V., Konrad, J. ve Jodoin, P.-M.** (2010). Video anomaly identification. *IEEE Signal Process. Mag.* **27** 18–33.
- [2] **Ramachandra, B., Jones, M. ve Vatsavai, R. R.** (2020). A survey of single-scene video anomaly detection. *IEEE Trans. Pattern Anal. Mach. Intell.*
- [3] **DiptiBhushan, S. ve Moharil, A. N.** (2017). Different Video Compression Standards. 83–5.
- [4] **Url-1** <https://en.wikipedia.org/wiki/Video_coding_format>, erişim tarihi 19.11.2021.
- [5] **Sullivan, G. J., Ohm, J.-R., Han, W.-J. ve Wiegand, T.** (2012). Overview of the high efficiency video coding (HEVC) standard. *IEEE Trans. circuits Syst. video Technol.* **22** 1649–68.
- [6] **Url-2** <<https://www.elektriktesisatportali.com/video-sikistirma-nedir-hayatimizi-nasil-kolaylastirir-yazi-dizisi-2.html>>, erişim tarihi 19.11.2021.
- [7] **Url-3** <<https://www.techopedia.com/definition/31897/high-efficiency-video-coding-hevc>>, erişim tarihi 19.11.2021.
- [8] **Ionescu, R. T., Khan, F. S., Georgescu, M.-I. ve Shao, L.** (2019). Object-centric auto-encoders and dummy anomalies for abnormal event detection in video. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition* pp 7842–51.
- [9] **Smeureanu, S., Ionescu, R. T., Popescu, M. ve Alexe, B.** (2017). Deep appearance features for abnormal behavior detection in video. In *International Conference on Image Analysis and Processing* pp 779–89. Springer.
- [10] **Ionescu, R. T., Smeureanu, S., Popescu, M. ve Alexe, B.** (2019). Detecting abnormal events in video using narrowed normality clusters. In *2019 IEEE Winter Conference on Applications of Computer Vision (WACV)* pp 1951–60. IEEE.
- [11] **Ma, K., Doescher, M. ve Bodden, C.** (2015). Anomaly detection in crowded scenes using dense trajectories. *Univ. Wisconsin-Madison*.
- [12] **Xu, D., Ricci, E., Yan, Y., Song, J. ve Sebe, N.** (2015). Learning deep representations of appearance and motion for anomalous event detection. *arXiv Prepr. arXiv1510.01553*.
- [13] **Tran, H. T. M. ve Hogg, D.** (2017). Anomaly detection using a convolutional winner-take-all autoencoder. In *Proceedings of the British Machine Vision Conference 2017*. British Machine Vision Association.
- [14] **Sabokrou, M., Fayyaz, M., Fathy, M. ve Klette, R.** (2017). Deep-cascade: Cascading 3d deep neural networks for fast anomaly detection and localization in crowded scenes. *IEEE Trans. Image Process.* **26** 1992–2004.
- [15] **Sabokrou, M., Fayyaz, M., Fathy, M., Moayed, Z. ve Klette, R.** (2018). Deep-anomaly: Fully convolutional neural network for fast anomaly detection in crowded scenes. *Comput. Vis. Image Underst.* **172** 88–97.
- [16] **Sabokrou, M., Fathy, M., Hoseini, M. ve Klette, R.** (2015). Real-time anomaly detection and localization in crowded scenes. In *Proceedings of the IEEE conference on computer vision and pattern recognition workshops* pp 56–62.
- [17] **Sun, Q., Liu, H. ve Harada, T.** (2017). Online growing neural gas for anomaly detection in changing surveillance scenes. *Pattern Recognit.* **64** 187–201.
- [18] **Ramachandra, B. ve Jones, M.** (2020). Street Scene: A new dataset and evaluation protocol for video anomaly detection. In *Proceedings of the*

- IEEE/CVF Winter Conference on Applications of Computer Vision* pp 2569–78.
- [19] **Saligrama, V. ve Chen, Z.** (2012). Video anomaly detection based on local statistical aggregates. In *2012 IEEE Conference on Computer Vision and Pattern Recognition* pp 2112–9. IEEE.
 - [20] **Dalal, N., Triggs, B. ve Schmid, C.** (2006). Human detection using oriented histograms of flow and appearance. In *European conference on computer vision* pp 428–41. Springer.
 - [21] **Dalal, N. ve Triggs, B.** (2005). Histograms of oriented gradients for human detection. In *2005 IEEE computer society conference on computer vision and pattern recognition (CVPR'05)* vol 1 pp 886–93. Ieee.
 - [22] **Laptev, I., Marszalek, M., Schmid, C. ve Rozenfeld, B.** (2008). Learning realistic human actions from movies. In *2008 IEEE Conference on Computer Vision and Pattern Recognition* pp 1–8. IEEE.
 - [23] **Wang, H., Klaser, A., Schmid, C. ve Liu, C.-L.** (2013). Dense trajectories and motion boundary descriptors for action recognition. *Int. J. Comput. Vis.* **103** 60–79.
 - [24] **Dollar, P., Rabaud, V., Cottrell, G. ve Belongie, S.** (2005). Behavior recognition via sparse spatio-temporal features. In *2005 IEEE International Workshop on Visual Surveillance and Performance Evaluation of Tracking and Surveillance* pp 65–72. IEEE.
 - [25] **Ravanbakhsh, M., Nabi, M., Mousavi, H., Sangineto, E. ve Sebe, N.** (2018). Plug-and-play cnn for crowd motion analysis: An application in abnormal event detection. In *2018 IEEE Winter Conference on Applications of Computer Vision (WACV)* pp 1689–98. IEEE.
 - [26] **Helbing, D. ve Molnar, P.** (1995). Social force model for pedestrian dynamics. *Phys. Rev. E* **51** 4282.
 - [27] **Schölkopf, B., Williamson, R. C., Smola, A. J., Shawe-Taylor, J. ve Platt, J. C.** (1999). Support vector method for novelty detection. In *NIPS* vol 12 pp 582–8. Citeseer.
 - [28] **Chatfield, K., Simonyan, K., Vedaldi, A. ve Zisserman, A.** (2014). Return of the devil in the details: Delving deep into convolutional nets. *arXiv Prepr. arXiv1405.3531*.
 - [29] **Makhzani, A. ve Frey, B. J.** (2015). Winner-take-all autoencoders. *Adv. Neural Inf. Process. Syst.* **28** 2791–9.
 - [30] **Turchini, F., Seidenari, L. ve Del Bimbo, A.** (2017). Convex polytope ensembles for spatio-temporal anomaly detection. In *International Conference on Image Analysis and Processing* pp 174–84. Springer.
 - [31] **Casale, P., Pujol, O. ve Radeva, P.** (2014). Approximate polytope ensemble for one-class classification. *Pattern Recognit.* **47** 854–64.
 - [32] **Fritzke, B.** (1995). A growing neural gas network learns topologies. *Adv. Neural Inf. Process. Syst.* **7** 625–32.
 - [33] **Krizhevsky, A., Sutskever, I. ve Hinton, G. E.** (2012). Imagenet classification with deep convolutional neural networks. *Adv. Neural Inf. Process. Syst.* **25** 1097–105.
 - [34] **Gong, Y., Lazebnik, S., Gordo, A. ve Perronnin, F.** (2012). Iterative quantization: A procrustean approach to learning binary codes for large-scale image retrieval. *IEEE Trans. Pattern Anal. Mach. Intell.* **35** 2916–29.
 - [35] **Sabokrou, M., Khalooei, M., Fathy, M. ve Adeli, E.** (2018). Adversarially learned one-class classifier for novelty detection. In *Proceedings of the IEEE*

- Conference on Computer Vision and Pattern Recognition* pp 3379–88.
- [36] **Lin, T.-Y., Dollar, P., Girshick, R., He, K., Hariharan, B. ve Belongie, S.** (2017). Feature pyramid networks for object detection. In *Proceedings of the IEEE conference on computer vision and pattern recognition* pp 2117–25.
 - [37] **Ramachandra, B., Jones, M. ve Vatsavai, R.** (2020). Learning a distance function with a Siamese network to localize anomalies in videos. In *Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision* pp 2598–607.
 - [38] **Bromley, J., Bentz, J. W., Bottou, L., Guyon, I., LeCun, Y., Moore, C., Sackinger, E. ve Shah, R.** (1993). Signature verification using a “siamese” time delay neural network. *Int. J. Pattern Recognit. Artif. Intell.* **7** 669–88.
 - [39] **Kratz, L. ve Nishino, K.** (2009). Anomaly detection in extremely crowded scenes using spatio-temporal motion pattern models. In *2009 IEEE conference on computer vision and pattern recognition* pp 1446–53. IEEE.
 - [40] **Adam, A., Rivlin, E., Shimshoni, I. ve Reinitz, D.** (2008). Robust real-time unusual event detection using multiple fixed-location monitors. *IEEE Trans. Pattern Anal. Mach. Intell.* **30** 555–60.
 - [41] **Anti, B. ve Ommer, B.** (2011). Video parsing for abnormality detection. In *2011 International Conference on Computer Vision* pp 2415–22. IEEE.
 - [42] **Anti, B. ve Ommer, B.** (2015). Spatio-temporal Video Parsing for Abnormality Detection. *arXiv Prepr. arXiv1502.06235*.
 - [43] **Wu, S., Moore, B. E. ve Shah, M.** (2010). Chaotic invariants of lagrangian particle trajectories for anomaly detection in crowded scenes. In *2010 IEEE computer society conference on computer vision and pattern recognition* pp 2054–60. IEEE.
 - [44] **Mehran, R., Oyama, A. ve Shah, M.** (2009). Abnormal crowd behavior detection using social force model. In *2009 IEEE Conference on Computer Vision and Pattern Recognition* pp 935–42. IEEE.
 - [45] **Kim, J. ve Grauman, K.** (2009). Observe locally, infer globally: a space-time MRF for detecting abnormal activities with incremental updates. In *2009 IEEE conference on computer vision and pattern recognition* pp 2921–8. IEEE.
 - [46] **Cheng, K.-W., Chen, Y.-T. ve Fang, W.-H.** (2015). Video anomaly detection and localization using hierarchical feature representation and Gaussian process regression. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition* pp 2909–17.
 - [47] **Benezeth, Y., Jodoin, P.-M., Saligrama, V. ve Rosenberger, C.** (2009). Abnormal events detection based on spatio-temporal co-occurrences. In *2009 IEEE conference on computer vision and pattern recognition* pp 2458–65. IEEE.
 - [48] **Hinami, R., Mei, T. ve Satoh, S.** (2017). Joint detection and recounting of abnormal events by learning deep generic knowledge. In *Proceedings of the IEEE International Conference on Computer Vision* pp 3619–27.
 - [49] **Feng, Y., Yuan, Y. ve Lu, X.** (2017). Learning deep event models for crowd anomaly detection. *Neurocomputing* **219** 548–56.
 - [50] **Tipping, M. E. ve Bishop, C. M.** (1999). Mixtures of probabilistic principal component analyzers. *Neural Comput.* **11** 443–82.
 - [51] **Mahadevan, V., Li, W., Bhalodia, V. ve Vasconcelos, N.** (2010). Anomaly detection in crowded scenes. In *2010 IEEE Computer Society Conference on Computer Vision and Pattern Recognition* pp 1975–81. IEEE.
 - [52] **Doretto, G., Chiuso, A., Wu, Y. N. ve Soatto, S.** (2003). Dynamic textures. *Int.*

- J. Comput. Vis.* **51** 91–109.
- [53] **Chan, A. B. ve Vasconcelos, N.** (2008). Modeling, clustering, and segmenting video with mixtures of dynamic textures. *IEEE Trans. Pattern Anal. Mach. Intell.* **30** 909–26.
 - [54] **Gao, D. ve Vasconcelos, N.** (2009). Decision-theoretic saliency: computational principles, biological plausibility, and implications for neurophysiology and psychophysics. *Neural Comput.* **21** 239–71.
 - [55] **Li, W., Mahadevan, V. ve Vasconcelos, N.** (2013). Anomaly detection and localization in crowded scenes. *IEEE Trans. Pattern Anal. Mach. Intell.* **36** 18–32.
 - [56] **Lafferty, J., McCallum, A. ve Pereira, F. C. N.** (2001). Conditional random fields: Probabilistic models for segmenting and labeling sequence data.
 - [57] **Rasmussen, C. E.** (2003). Gaussian processes in machine learning. In *Summer school on machine learning* pp 63–71. Springer.
 - [58] **Lin, T.-Y., Maire, M., Belongie, S., Hays, J., Perona, P., Ramanan, D., Dollar, P. ve Zitnick, C. L.** (2014). Microsoft coco: Common objects in context. In *European conference on computer vision* pp 740–55. Springer.
 - [59] **Krishna, R., Zhu, Y., Groth, O., Johnson, J., Hata, K., Kravitz, J., Chen, S., Kalantidis, Y., Li, L.-J. ve Shamma, D. A.** (2017). Visual genome: Connecting language and vision using crowdsourced dense image annotations. *Int. J. Comput. Vis.* **123** 32–73.
 - [60] **Girshick, R.** (2015). Fast r-cnn. In *Proceedings of the IEEE international conference on computer vision* pp 1440–8.
 - [61] **Chan, T.-H., Jia, K., Gao, S., Lu, J., Zeng, Z. ve Ma, Y.** (2015). PCANet: A simple deep learning baseline for image classification? *IEEE Trans. image Process.* **24** 5017–32.
 - [62] **Van Den Oord, A. and Schrauwen, B.** (2014). Factoring variations in natural images with deep gaussian mixture models. In *Neural Information Processing Systems*.
 - [63] **Hasan, M., Choi, J., Neumann, J., Roy-Chowdhury, A. K. ve Davis, L. S.** (2016). Learning temporal regularity in video sequences. In *Proceedings of the IEEE conference on computer vision and pattern recognition* pp 733–42.
 - [64] **Chong, Y. S. and Tay, Y. H.** (2017). Abnormal event detection in videos using spatiotemporal autoencoder. In *International symposium on neural networks* pp 189–96. Springer.
 - [65] **Vu, H., Nguyen, T. D., Le, T., Luo, W. ve Phung, D.** (2019). Robust anomaly detection in videos using multilevel representations. In *Proceedings of the AAAI Conference on Artificial Intelligence* vol 33 pp 5216–23.
 - [66] **Gong, D., Liu, L., Le, V., Saha, B., Mansour, M. R., Venkatesh, S. ve Hengel, A. van den.** (2019). Memorizing normality to detect anomaly: Memory-augmented deep autoencoder for unsupervised anomaly detection. In *Proceedings of the IEEE/CVF International Conference on Computer Vision* pp 1705–14.
 - [67] **Ravanbakhsh, M., Nabi, M., Sangineto, E., Marcenaro, L., Regazzoni, C. ve Sebe, N.** (2017). Abnormal event detection in videos using generative adversarial nets. In *2017 IEEE International Conference on Image Processing (ICIP)* pp 1577–81. IEEE.
 - [68] **Nguyen, T.-N. ve Meunier, J.** (2019). Anomaly detection in video sequence with appearance-motion correspondence. In *Proceedings of the IEEE/CVF International Conference on Computer Vision* pp 1273–83.

- [69] **Tang, Y., Zhao, L., Zhang, S., Gong, C., Li, G. ve Yang, J.** (2020). Integrating prediction and reconstruction for anomaly detection. *Pattern Recognit. Lett.* **129** 123–30.
- [70] **Xingjian, S. H. I., Chen, Z., Wang, H., Yeung, D.-Y., Wong, W.-K. ve Woo, W.** (2015). Convolutional LSTM network: A machine learning approach for precipitation nowcasting. In *Advances in neural information processing systems* pp 802–10.
- [71] **Goodfellow, I., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., Courville, A. ve Bengio, Y.** (2014). Generative adversarial nets. *Adv. Neural Inf. Process. Syst.* **27**.
- [72] **Vu, H., Phung, D., Nguyen, T. D., Trevors, A. ve Venkatesh, S.** (2017). Energy-based models for video anomaly detection. *arXiv Prepr. arXiv1708.05211*.
- [73] **Hinton, G. E.** (2002). Training products of experts by minimizing contrastive divergence. *Neural Comput.* **14** 1771–800.
- [74] **Freund, Y. ve Haussler, D.** (1994). Unsupervised learning of distributions of binary vectors using two layer networks.
- [75] **Liu, W., Luo, W., Lian, D. ve Gao, S.** (2018). Future frame prediction for anomaly detection-a new baseline. In *Proceedings of the IEEE conference on computer vision and pattern recognition* pp 6536–45.
- [76] **Ronneberger, O., Fischer, P. ve Brox, T.** (2015). U-net: Convolutional networks for biomedical image segmentation. In *International Conference on Medical image computing and computer-assisted intervention* pp 234–41. Springer.
- [77] **Dosovitskiy, A., Fischer, P., Ilg, E., Hausser, P., Hazirbas, C., Golkov, V., Van Der Smagt, P., Cremers, D. ve Brox, T.** (2015). FlowNet: Learning optical flow with convolutional networks. In *Proceedings of the IEEE international conference on computer vision* pp 2758–66.
- [78] **Isola, P., Zhu, J.-Y., Zhou, T. ve Efros, A. A.** (2017). Image-to-image translation with conditional adversarial networks. In *Proceedings of the IEEE conference on computer vision and pattern recognition* pp 1125–34.
- [79] **Park, H., Noh, J. ve Ham, B.** (2020). Learning memory-guided normality for anomaly detection. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition* pp 14372–81.
- [80] **Georgescu, M.-I., Barbalau, A., Ionescu, R. T., Khan, F. S., Popescu, M. ve Shah, M.** (2021). Anomaly detection in video via self-supervised and multi-task learning. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition* pp 12742–52.
- [81] **Wu, C., Waseem, F. and Martinez, R. P.** Visual Anomaly Detection in Video.
- [82] **Georgescu, M.-I., Ionescu, R. T., Khan, F. S., Popescu, M. ve Shah, M.** (2020). A Background-Agnostic Framework with Adversarial Training for Abnormal Event Detection in Video. *arXiv Prepr. arXiv2008.12328*.
- [83] **Wang, S., Yu, G., Cai, Z., Zhu, E., Liu, X., Yin, J. ve Zhu, C.** (2021). Sensing Anomalies like Humans: A Hominine Framework to Detect Abnormal Events from Unlabeled Videos. *arXiv Prepr. arXiv2108.01975*.
- [84] **Majhi, S., Das, S. ve Bremond, F.** DAM: Dissimilarity Attention Module for Weakly-supervised Video Anomaly Detection.
- [85] **Biswas, S. ve Babu, R. V.** (2013). Real time anomaly detection in H. 264 compressed videos. In *2013 Fourth National Conference on Computer Vision, Pattern Recognition, Image Processing and Graphics (NCVPRIPG)* pp 1–4. IEEE.

- [86] **Biswas, S. ve Babu, R. V.** (2014). Sparse representation based anomaly detection using HOMV in H. 264 compressed videos. In *2014 International Conference on Signal Processing and Communications (SPCOM)* pp 1–6. IEEE.
- [87] **Url-4** <<http://www.svcl.ucsd.edu/projects/anomaly/dataset.html>>, erişim tarihi 10.10.2021
- [88] **Url-5** <<https://en.wikipedia.org/wiki/YUV>>, erişim tarihi 04.12.2021
- [89] **Url-6** <https://hevc.hhi.fraunhofer.de/svn/svn_HEVCSoftware/tags/HM-16.9/>, erişim tarihi 04.12.2021
- [90] **Url-7** <https://en.wikipedia.org/wiki/Video_compression_picture_types>, erişim tarihi 04.10.2021
- [91] **Url-8** <<https://github.com/lheric/GitlHEVCAalyzer>>, erişim tarihi 04.08.2021
- [92] **Url-9** <<https://www.renom.jp/notebooks/tutorial/generative-model/VAE/notebook.html>>, erişim tarihi 04.10.2021
- [93] **Url-10** <<https://www.mathworks.com/help/deeplearning/ug/train-a-variational-autoencoder-vae-to-generate-images.html>>, erişim tarihi 04.09.2021
- [94] **Url-11** <https://github.com/rafapm/CS_230/blob/main/daelstm/notebooks/visualautoencoder-variational.ipynb>, erişim tarihi 13.11.2021

ÖZGEÇMİŞ

Ad-Soyad : Sümeyye Çavaş

ÖĞRENİM DURUMU:

- **Lisans** : 2016, İstanbul Üniversitesi, Mühendislik Fakültesi, Bilgisayar Mühendisliği
- **Yükseklisans** : 2022, İstanbul Teknik Üniversitesi, Bilgisayar Bilimleri Anabilim Dalı, Bilgisayar Bilimleri

YÜKSEK LİSANS TEZİNDEN TÜRETİLEN YAYINLAR:

Çavaş, Sümeyye; Beratoğlu, Muhammet Sebul; Töreyin, Behçet Uğur. Anomaly Detection In Compressed Video. In: 2021 29th Signal Processing and Communications Applications Conference (SIU). IEEE, 2021. p. 1-4.