

**ANKARA ÜNİVERSİTESİ
FEN BİLİMLERİ ENSTİTÜSÜ**

YÜKSEK LİSANS TEZİ

İKİLİ KAMERA İLE VİDEO ÖZETİ YAKLAŞIMI

Talha DİLBER

BİLGİSAYAR MÜHENDİSLİĞİ ANABİLİM DALI

**ANKARA
2021**

Her hakkı saklıdır

ÖZET

Yüksek Lisans Tezi

İKİLİ KAMERA İLE VIDEO ÖZETİ YAKLAŞIMI

Talha DİLBER

Ankara Üniversitesi
Fen Bilimleri Enstitüsü
Bilgisayar Mühendisliği Anabilim Dalı

Danışman: Doç. Dr. Mehmet Serdar GÜZEL

Günümüz dünyasında her alanda üretilen veri miktarı beklenmeyecek seviyede artış göstermiştir. Artan veri karşısında veri işlemenin önemi dikkate değer bir şekilde artmıştır. Tez konumuz, artan veri içerisinde önemli yeri olan video verilerinin işlenmesi ve özet videoların üretilmesi üzerinedir. Bu tez kapsamında, video özeti oluşturulurken, nesne tabanlı denetimsiz öğrenme ile aykırılık tespiti yapan ve tespit edilen aykırılıkları video kesitleri belirlemede kullanan yeni bir yöntem geliştirilmiştir. Bu yöntem kullanılarak video verisi piksel olarak işlenmekte ve video kesiti olarak sonuç üretilmektedir. İşlem akışı kısaca şu şekilde özetlenebilir. Video üzerindeki nesneler tiplerine göre tespit edilir ardından takip edilirler. Sonrasında nesnelerin takip geçmişi verileri işlenerek nesne tipi ile sınıflandırıcı eğitilir. Bu sınıflandırıcı sayesinde nesnelerin aykırılık (anomali) davranışları tespit edilir. Aykırılık davranışları içeren video anları işlenerek video kesitleri saptanır. Saptanan video kesitleri asıl videodan çıkartılıp birleştirilerek video özeti oluşturulur. Geliştirdiğimiz model tek kamera ve ikili kamera sistemleri için ayrı ayrı test edilip doğrulanmıştır.

Temmuz 2021, 57 sayfa

Anahtar Sözcükler: Sınıflandırma, anomali tespiti, video işleme, video özetleme

ABSTRACT

Master Thesis

VIDEO SYNOPSIS BASED APPROACH USING STEREO CAMERA

Talha DİLBER

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

Supervisor: Assoc. Prof. Dr. Mehmet Serdar GÜZEL

In today's world, the amount of data produced in every field has increased at an unexpected level. In the face of increasing data, the importance of data processing has increased remarkably. Our thesis topic is on the processing of video data, which has an important place in increasing data, and the production of summary videos. Within the scope of this thesis, a new method has been developed that detects anomaly with object-based unsupervised learning and uses the detected anomalies to determine video segments while creating a video summary. By using this method, the video data is processed as pixels and the result is produced as a video segment. The process flow can be briefly summarized as follows. Objects on the video are detected according to their type, and then they are tracked. Then, the tracking history data of the objects are processed, and the classifier is trained with the object type. Thanks to this classifier, anomaly behavior of objects is detected. Video segments are determined by processing video moments containing anomaly behaviors. The video summary is created by extracting the detected video segments from the original video and combining them. The model we developed has been tested and verified separately for single camera and dual camera systems.

July 2021, 57 pages

Key Words: Anomaly detection, video processing, video summarizing, video synopsis

ÖNSÖZ ve TEŞEKKÜR

Yüksek lisans tez yazım sürecinde her türlü maddi manevi desteği veren tez danışmanım Sayın Doç. Dr. Mehmet Serdar GÜZEL hocama gönülden teşekkürlerimi sunmaktayım. Tez yazım sırasında beni her zaman destekleyen, motive eden sevgili hayat arkadaşım Zeynep'e tüm kalbimle minnet duyuyorum.

Yüksek lisans eğitimimin her safhasında verdikleri katkılardan dolayı başta Babam olmak üzere tüm aileme teşekkürlerimi sunarım.

Talha DİLBER

Ankara, Temmuz 2021

İÇİNDEKİLER

TEZ ONAY SAYFASI	
ETİK.....	i
ÖZET.....	ii
ABSTRACT	iii
ÖNSÖZ ve TEŞEKKÜR.....	iv
ÇİZELGELER DİZİNİ	vii
ŞEKİLLER DİZİNİ	viii
KISALTMALAR DİZİNİ	ix
1. GİRİŞ.....	1
2. YÖNTEM.....	7
2.1 Yolo.....	7
2.2 Sınıflandırma	9
2.2.1 İstekli öğrenen sınıflandırma algoritmaları.....	10
2.2.1.1 Karar ağacı	10
2.2.1.2 Naif Bayes	11
2.2.1.3 Yapay sinir ağları.....	12
2.2.2 Tembel öğrenen sınıflandırma algoritmaları	14
2.2.2.1 K-En yakın komşu (KNN).....	14
2.3 Anormallik Tespiti	15
2.4 Video Özeti Çıkartma	16
3. LİTERATÜR	17
4. METODOLOJİ	22
4.1 Hedef ve Motivasyon.....	23
4.2 Videonun Hazırlanması.....	23
4.2.1 Videoya zamanlayıcı eklenmesi	24
4.2.2 Video üzerine anomali tespitlerinin çizimleri.....	25
4.3 Video Özet Parçalarının Belirlenmesi.....	26
4.3.1 YOLO kullanımı	26
4.3.2 Sınıflandırıcı kullanımı	27
4.3.3 Anomalilerin işlenmesi	31
4.4 Video Özeti Oluşturulması	33
4.5 İkili Kamera ile Video Özeti Oluşturulması	34
5. DENEYLER.....	36
5.1 1. Kısa Video Deneyi	36
5.1.1 Senaryo 1.....	37
5.1.2 Senaryo 2	37
5.1.3 Senaryo 3	38
5.1.4 Senaryo 4	38
5.2 2. Kısa Video Deneyi	38
5.2.1 Senaryo 1.....	39
5.2.2 Senaryo 2	39
5.2.3 Senaryo 3	40
5.2.4 Senaryo 4	40

5.3 3. Kısa Video Deneyi	40
5.3.1 Senaryo 1	41
5.3.2 Senaryo 2	41
5.3.3 Senaryo 3	42
5.3.4 Senaryo 4	42
5.4 1. Uzun Video Deneyi.....	42
5.4.1 Senaryo 1.....	43
5.4.2 Senaryo 2	43
5.4.3 Senaryo 3	44
5.4.4 Senaryo 4	44
5.5 2. Uzun Video Deneyi.....	44
5.5.1 Senaryo 1	45
5.5.2 Senaryo 2	45
5.5.3 Senaryo 3	46
5.5.4 Senaryo 4	46
5.6 İkili Kamera Kısa Video Deneyi.....	46
5.7 İkili Kamera Uzun Video Deneyi.....	47
5.8 Deney Karşılaştırmaları	48
6. SONUÇ VE GELECEK ÇALIŞMALAR	51
KAYNAKLAR.....	53
ÖZGEÇMİŞ.....	57

ÇİZELGELER DİZİNİ

Çizelge 5.1 Tek kamera deney sonuçları karşılaştırılması.....	48
Çizelge 5.2 Tek-çift kamera deney sonuçları karşılaştırılması	48



ŞEKİLLER DİZİNİ

Şekil 1.1 Video özetleme çeşitlerinin tablosu (Madhushree 2019).....	2
Şekil 2.1 Cascade Classifier ile yüz tespitin çerçeve seçim örneği	8
Şekil 2.2 YOLO ile tespit edilmiş objelerin görüntüsü.....	9
Şekil 2.3 Karar ağacı eğitim sonrası sonuç örneği.....	10
Şekil 2.4 Naif Bayes eğitim sonrası sonuçların görselleştirilmiş örneği	11
Şekil 2.5 Sınıflandırmada kullanılabilecek bir yapay sinir ağları örneği.....	13
Şekil 2.6 K-En Yakın Komşu öğrenimi sonucunun görselleştirilmesi örneği	14
Şekil 2.7 Anormallik tespiti görselleştirilmiş örneği	15
Şekil 4.1 Video özeti çıkartma akış şeması.....	22
Şekil 4.2 Örnek video üzerine zaman bilgisi eklenmesinin örnek görseli	24
Şekil 4.3 Video özetinde işaretlenmiş olan anomali gösteren objenin görseli.....	25
Şekil 4.4 YOLO ile işlenmiş bir çerçevenin görseli	27
Şekil 4.5 YOLO ve sınıflandırıcının arasındaki ilişkiyi gösteren şekil	31
Şekil 4.6 Temsili anomali görseli (Xiufeng Liu vd. 2016)	33
Şekil 4.7 Özetleme işlemi ile ilgili temsili bir görsel (Rav-Acha vd. 2006).....	34
Şekil 5.1 1. Kısa Video. Deneylerimizde kullandığımız videodan bir görsel.....	37
Şekil 5.2 2. Kısa Video. Deneylerimizde kullandığımız videodan bir görsel.....	39
Şekil 5.3 3. Kısa Video. Deneylerimizde kullandığımız videodan bir görsel.....	41
Şekil 5.4 1. Uzun Video. Deneylerimizde kullandığımız videodan bir görsel	43
Şekil 5.5 2. Uzun Video. Deneylerimizde kullandığımız videodan bir görsel	45

KISALTMALAR DİZİNİ

CPU	Merkezi İşlem Birimi (Central Process Unit)
GPU	Grafik İşlem Birimi (Graphic Process Unit)
SGD	Stokastik Gardiyan İniş (Stochastic Gradient Descent)
YOLO	Yalnızca Birkez Bakarsınız (You Only Look Once)

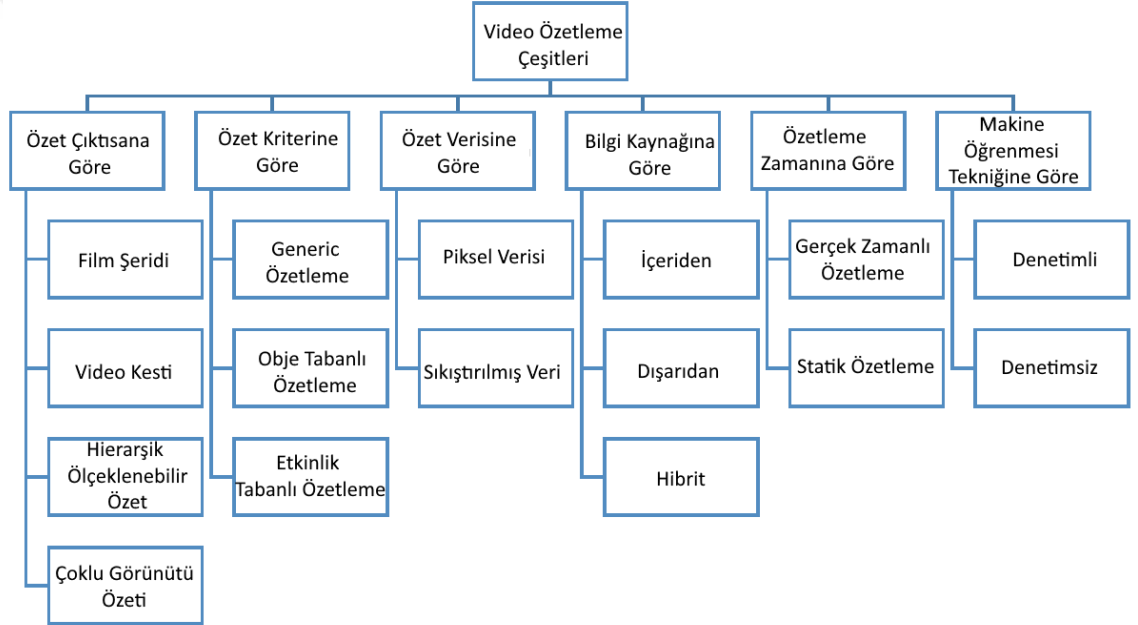


1. GİRİŞ

Günümüzde hızla gelişen teknoloji hayatımızın her alanına dahil olmuş bununla beraber işitsel, görsel ve daha pek çok çeşitli veri birikmeye başlamıştır. Büyük veri yakın geçmiş zamanlarda yazılım alanının öncüleri tarafından farklı tanımlamalar ile ifade edilmiştir. Büyük verinin etimolojisi, 1990'ların ortalarına kadar dayanır. İlk olarak eski Baş Bilim Adamı John Mashey tarafından, büyük veri kümelerinin işlenmesi ve analizine atıfta bulunmak için kullanılmıştır (Diebold, 2012). Büyük verinin içerisindeki veriler çok boyutlu olabilir veya işlenmemiş veriler içerebilir. Bu sebepten dolayı eski usul yöntemler büyük verinin işlenmesinde yetersiz kalmaktadır. Büyük veriyi işlemek hem teknoloji hem de analiz gerektirir. İşlemler sonucunda geleceği şekillendirebilecek, hayatı kolaylaştırabilecek veya karlılığı arttıracak veriler elde etmek mümkündür.

Büyük veriyi oluşturan ve aynı zamanda günlük yaşantımıza dokunan veri tipleri içerisinde kamera, video görüntüleri dikkat çeker bir boyuta ulaşmıştır (Tiejun Huang, 2014). Bankadan müzeye, ATM'den resmi devlet dairelerine kadar onlarca hatta yüzlerce kamera farklı amaçlarla kullanılmaktadır. Örneğin bazı kameralar yerleşim yerleri giriş çıkışına konumlandırılmış olup otoyollardan geçen araçları kaydetmektedir. Bazı kameralar da sokaklarda daha çok kaldırımdaki yayaları görecektir şekilde konumlandırılmış olabilmektedir. Bu kameralar çoğunlukla yüksek çözünürlükte olup, 7/24 hizmet sağlamakta ve aynı zamanda tüm ürettiği verileri belirli yerlerde depolamaktadır. Bu sebeple üretilen veri miktarı çok fazla olmakta ve veri depolama alanları aslında üretilmiş veriler ile dolmaktadır. Bu verilerin çok büyük bir kısmının gereksiz olduğu bilinen bir gerçektir. Bir güvenlik kamerası görüntüsünün incelemesi gerektiğinde bakılması gerekli olan görüntünün uzunluğu ve muhtemel oluşabilecek farklı kamera sayısı da dikkate alındığında, verilerin içerisinde anlamlı olanları tespit etmek gerçekten insan potansiyelini zorlayacak sınırlara ulaşabilmektedir. Verilerin özetlenmesi, kısaltılması, gereksiz kısımlarının elenmesi pek çok yönüyle büyük faydalar getirecektir. Öncelikle karşımıza birkaç soru çıkmaktadır. Gereksiz kısımlar hangileridir? Önemli ve saklamamız icap eden görüntüler nasıl tespit edilir? Düzenli hareketin olduğu bir kamerada gereksiz hareketler ile saklanması icap eden hareket nasıl ayırt edilir? Bu sorular tezimizin başlıca problemleridir.

Verilerin azaltılması için çeşitli çözümler mevcuttur. Fakat her bir çözümün farklı avantajları ve dezavantajları vardır. Çözümlerden biri videoların özetlerinin çıkarılmasıdır. Bu yöntem iki şekilde olabilir. Birinci yöntem, video üretimi sırasında videonun gereksiz olduğu tespit edilen yerlerinin daha ana kaydediciye kaydedilmeden elenmesi ve sadece gerekli görülen yerlerin ana kaydediciye kaydedilmesi esasına dayanır. İkinci yöntem ise video kaydı tamamlandıktan sonra günün, haftanın belirli periyotlarında üretilmiş verilerin analiz edilmesi ve gereksiz kısımların atılması esasına dayanır. Bu iki yöntemi de kapsayan başlıca bir problem vardır. Bu başlıca problem, özetin çıkarılması sırasında verinin hangi bölümünün anlamlı hangi bölümünün anlamsız olduğunun tespitinin nasıl yapılması gerektiğidir.



Şekil 1.1 Video özetleme çeşitlerinin tablosu (Madhushree 2019).

Video verilerinin özetlenmesinin çözümlerinden bir diğeri ise üretilmiş videoların belirlenen zaman periyotları içerisinde belirlenmiş zaman aşımı kuralları ile bu kurala uymayanların silinmesidir. Örneğin kural olarak kaydedilmiş video verisi 1 haftadan eski ise silinecektir. Bu kural ile her gece 8 gün (1 hafta + 1 gün) önceki veriler silinir. Video verilerinin kural bazlı periyodik silinme çözümünün başlıca problemi, eskiye ait verilerin bir daha erişilemez olmasıdır. Örneğin bir dava sırasında 3 ay önceki bir olayın kamera kaydı mahkemeden istendiğinde bu verinin temini yapılamayacaktır. Video özeti çıkarma çözümü ile video verilerinin periyodik silinmesi çözümünü melez olarak birleştirerek

zaman aşımına uğramış video verilerinin özeti çıkarılarak saklanması da bir seçenek olarak karşımıza çıkmaktadır. Verilerin azaltılmasına yönelik üçüncü çözüm olarak sıkıştırma yöntemi de vardır. Sıkıştırma, veri boyutlarını düşürmeyi sağlar. Bu çözümün zayıf yanı da sıkıştırılmış veri içerisinde arama yapmanın kolay olmamasıdır. Sıkıştırma çözümüne benzer olarak videolardaki gereksiz parçalar atılabilir. Mesela kamera kayıtlarında genellikle ses verisi önemli değildir. Ses verisi içermeyen video formatları kullanılması büyük veriyi verimli saklama çözümüne bir miktar katkı sağlayacaktır.

Çözümler arasında akılcı ve yenilikçi olanlardan biri de nesne tabanlı video özeti çıkarmadır (Madhushree 2019). Nesne tabanlı özet çıkarma işleminin temel mantığı, video üzerindeki nesnelerin arkada tekrar eden video görüntüsünden ayrıştırılıp sadece nesnelerin video üstüne koyulmasıdır. Akıllıca ve yenilikçi bu çözümün dezavantajı ise şudur. Video özetini inceleyen insanlar gördükleri görüntülerdeki doğallığın kaybolduğunu hissederek olan olayları ayırt edemediklerini söylemektedirler. Verinin özeti teorik olarak akıllıca çıkarılmış olsa da aslında işlevini yitirmeye başlamıştır.

Video özeti çıkarırken kullanabileceğimiz bir konu da yapay zekadır (Madhushree 2019). Yapay zekâ (AI); tekrarlı işleri öğrenmek, bir örüntüyü öğrenmek konusunda oldukça iyidir. Yapay zekâ günümüzde dünyanın en iyi satranç oyuncusunu, satrançtan daha kompleks olan Çin kökenli olan go oyununun dünyanın en iyi oyuncusunu yenebilecek seviyelere gelmiştir. Sadece oyun alanında değil dil işleme, otonom giden arabalar gibi farklı alanlarda yapay zekânın dikkat çeken örneklerini görebilmekteyiz. Yapay zekâ kullanarak programlar, video verilerinin özetinin çıkartılması işlemlerinde kullanılabilir ve örnekleri mevcuttur.

Sınıflandırma, bir tanımlama işlemidir. Belirli veya belirli olmayan sınıf ayrımları üzerinde bir verinin hangi sınıfa ait olduğunu tespiti etmeyi amaçlamaktadır. Sınıflandırıcının öğrenme safhası genel olarak şu şekildedir. Sınıflandırıcıya öncelikle öğrenim verisi ve test verisi gerekmektedir. Hem öğrenim verisinde hem de test verisinde katalog şeklinde veriler ve bu verilerin sınıf bilgisi bulunmaktadır. Öğrenim verisi ile sınıflandırıcı eğitilir. Sonrasında test verisi ile sınıflandırıcının başarı oranı çıkartılır. Sınıflandırıcının sınıfını belirlemek üzere işlediği verilerin sonucu kesinlik içermeyebilir. Örneğin bir sınıflandırıcı bir yazı verisinin hangi dilde olduğunu algılamaya çalışıyor.

Sınıflandırıcı bu yazının uzunluk kısalık durumlarından dolayı %40 Almanca, %60 İngilizce gibi sonuçlar üretebilir. Sınıflandırıcının eğitim ve test veri kümesinin kalitesine göre başarı oranı ve kesinlik oranları daha iyi olabilmektedir.

Sınıflandırma algoritmalarında tembel öğrenen sınıflandırma algoritmaları ve hevesli öğrenen sınıflandırma algoritmaları bulunur (Marco Grimaldi, 2006).

Tembel öğrenenler: Tembel öğrenen sınıflandırma algoritmaları öğrenim safhasında verileri az işlem ile depolar fakat bu sınıflandırıcıların sınıflandırma işlemleri sırasında daha fazla işlem yaparlar. Çünkü kaydettikleri öğrenimlerinin üzerinden tahminde bulunmaya çalışırlar. Bu sebepten dolayı tembel sıfatını almışlardır. En bilinen örneği K-en yakın komşu sınıflandırma algoritmasıdır.

Hevesli öğrenenler: Hevesli öğrenen sınıflandırma algoritmaları test süreçlerinde veriler üzerinde çok detaylı işlemde bulunurlar. Fakat sınıflandırma işlemleri sırasında daha az işlem ile sonuç üretmektedirler. Genelde öğrenim sonrasında oluşturdukları veriler kolayca anlaşılabilen veriler değildirler. En bilindik örnekleri Yapay sinir ağları, Naif Bayes ve karar ağacı algoritmalarıdır.

Derin öğrenme işlemi makine öğrenme işlemleri arasında en karmaşık işlem ağına sahip öğrenme işlemidir. Birçok farklı tekniği vardır ama çoğunlukla bir yapay sinir ağının çalışma prensiplerini taklit etmek üzerine çalışmalar yürütülmüştür. Örneğin geri beslemeli yapay sinir ağları sinir hücreleri benzeri katmanlar ve bu katmanlar arasında bağlantılar içermektedir. 1944 yılında Warren McCullough, Walter Pitts adlı bilim insanları tarafından sinir ağları ilk kez önerilmiştir.

Sinir ağları aynı zamanda MIT matematikçileri Marvin Minsky ve Seymour Papert tarafından araştırılan hem sinir bilimi hem de bilgisayar bilimlerinde ait önemli bir araştırma alanıydı (Marvin Minsky, 1972).

Derin öğrenmede, bir evrişimli sinir ağı (CNN), görsel görüntülerin analizinde en yaygın olarak uygulanan derin sinir ağları sınıfıdır (Saad Albawi, 2017). Görüntü ve video tanıma, tavsiye sistemleri, görüntü sınıflandırması, tıbbi görüntü analizi, doğal dil işleme,

beyin-bilgisayar ara yüzleri ve finansal zaman serilerinde evrişimli sinir ağı kullanımının örnekleri vardır.

CNN'ler çok katmanlı algılayıcının düzenli versiyonlarıdır. Çok katmanlı algılar genellikle tamamen bağlı ağlar anlamına gelir, yani bir katmandaki her nöron bir sonraki katmandaki tüm nöronlarla bağlı vardır. CNN'ler verilerdeki hiyerarşik düzenden yararlanırlar, küçük ve basit desenler kullanarak daha karmaşık desenleri tespit ederler.

Evrişimli sinir ağı biyolojik süreçlerden esinlenilerek üretilmiştir. Nöronlar arasındaki bağlantı kalıbı hayvan görsel korteks organizasyonunu andırır. Bireysel nöronların kısıtlı bir görme alanı vardır. Bu sebeple bireysel nöronlar sadece belirli bölgeye göre uyarılara veri gönderir. Farklı nöronların alıcı alanları, tüm görsel alanı kaplayacak şekilde örtüşür.

Görüntü sınıflandırma algoritmaları, CNN'lere kıyasla nispeten daha fazla ön işlem kullanır. Bu, ağı geleneksel algoritmalarda el mühendisliği ile yapılan filtreleri öğrendiği anlamına gelir. Bu bağımsızlık ön bilgi ve özellik tasarımı insan çabası açısından büyük bir avantajdır.

Anormallik algılama, aslında basit bir sınıflandırma işlemidir. Bu sınıflandırma işleminin içerisinde sadece iki adet sınıf bulunur. Bu sınıflar normal ve normal olmayan şeklindedir. Anormallik algılama işleminin temelinde yatan sorunlardan bir tanesi normal ve normal olmayan verilerin ayırma noktasıdır. Bu problem adına birçok yöntem geliştirilmiştir.

Güncel gelişmeler ekseninde Evrişimli Sinir Ağı (CNN) kullanarak geliştirilen Çoklu nesne algılayıcı ve izleyicileri oldukça yaygındır. Aynı zamanda bu yazılımlar birçok alanda da kullanılmaktadır.

Tezimde video özeti çıkarılması gereken alanların belirlenmesinde kullanılacak algoritma tasarımı olarak Evrişimli Sinir Ağı (CNN) kullanarak geliştirilen Çoklu nesne algılayıcı ve izleyicisini, sınıflandırma algoritmaları ile denemek istedim. Bir videonun veya bir canlı izlenen kameranın görüntüsünün üzerinde bir adet Çoklu nesne izleyicisi çalışacaktır. Çoklu nesne izleyicisi tüm tespit ettiği objelerin tiplerini ve hareketlerini bir tane sınıflandırıcıya gönderecek. Aynı zamanda Çoklu nesne izleyicisinin tespit ettiği

objenin hareketlerini sınıflandırıcıya göndererek sınıflandırıcıdan yapılan hareketin hangi obje tiplerine ait olabileceğinin sonuçlarını alıp, çoklu nesne izleyicisinin tespit ettiği objenin tipi ile uyumlu hareket edip etmediği kontrol edilecektir. Uyumlu ve uyumsuz hareket ayrımı, sınıflandırıcının obje hareketi ve obje tipi özelinde üretmiş olduğu sonucun belirli bir eşik değerden yüksek olup olmadığı ile belirlenmektedir. Yapılmak istenen bu yöntem aslında bir çeşit anormallik tespit işlemidir. Hareket uyumsuzluğu tespit edilmiş videoya ait parçalar kesilerek özet kısmına eklenecektir. Bu sebeple her bir nesne tipi özelinde rutin yapılan dikkate değmeyen işler elenmiş olup, dikkat çeken hareketler video özet parçası olarak belirlenip diğer kısımlardan ayrılıp özet kısmında birleştirilecektir.

İkili kamera ile video özetlenmesi yaklaşımı için düşündüklerimiz şu şekildedir. Öncelikle 2 kamera da fark açılara sahip olsalar bile aynı gözlem alanına sahip olmalılardır. 2 kamera aynı yerde aynı anda çoklu nesne izleyici ve sınıflandırıcı kullanarak tespit ettikleri anomali durumunda video özeti çıkarma işlemi uygulanacaktır. Böylece çoklu nesne izleyici ve sınıflandırıcı kullanılarak tespit edilen anormallikler iki farklı görüntü açısıyla teyit edilmiş olacak ve anormallik ihtimalini kuvvetlendirecektir.

2. YÖNTEM

2.1 Yolo

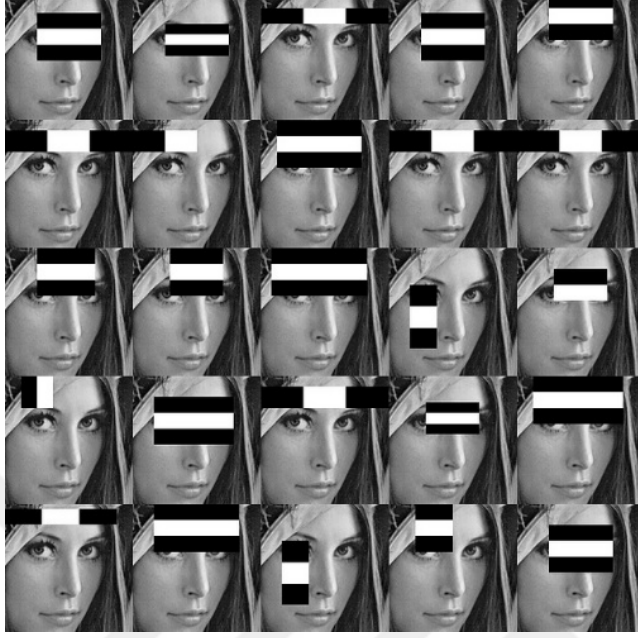
Video özeti için geliştirmiş olduğumuz çalışmamızda Evrişimli Sinir Ağı kullanarak geliştirilen son zamanlarda güzel referanslara sahip olan YOLO görüntü işleme aracı kullanıldı.

İnsanlar çevrelerindeki gerçek dünyanın üç boyutlu yapısını belirgin bir şekilde algılayabilir. Yanınızdaki masada duran bir kitaba baktığınızda üç boyutlu öngörünün ne kadar canlı olduğunu düşünün. Kitabın şeklini, rengini söyleyebilirsiniz. Benzer şekilde, herhangi bir resimdeki insan topluluğuna bakarak, insan sayısını kolayca sayabilirsiniz. Derin öğrenme kavramlarından sonra, bilgisayar görüşünün bu insan gözünü aratmadığı kanıtlanmıştır.

Nesne algılama, bir görüntüdeki farklı sınıflarda olabilecek nesneleri tanımlama ve konumlandırma ile ilgilenen bir Makine Görüş sorunudur. YOLO (Yalnızca Bir Kez Bakarsınız) nesneleri algılamak için kullanılabilecek bir yöntem veya yoldur. Görüntüde bulunan nesneleri algılamak için kullanılan bir algoritmadır. Adından da anlaşılacağı gibi (YOLO – You Only Look Once), tüm görüntüye sadece bir kez bakar, sinir ağından bir kez geçer ve nesneleri algılar (Sujata Chaudhari, 2020).

YOLO mimarisi F-CNN (tam evrişimsel sinir ağı) kullanılarak ve esinlenilerek geliştirilmiştir. Bir görüntüyü işleme sırasında işlenen görüntü yalnızca bir kez F-CNN üzerinden geçirilir ve çıktı tahmin edilir. Nesne algılandıktan sonra nesnenin boyutunun ve cinsinin kesinleştirilmesi için tespit edilmiş tüm olasılıklardaki cisim tipleri ve cisim boyutlarının olasılıklarını bir regresyon problemi çözer gibi işlem yapıp en uygun cisim tipini ve boyutunu tespit eder. Tespit edilmiş cisim boyutunun karşılığı çerçeve olarak ekranda gösterilebilmektedir. Bahsi geçen cisim boyutu sadece ekrandaki kapladığı alandır, cismin gerçek boyutu hakkında ölçüm yapılamamaktadır. Tek bir değerlendirmede, tek bir sinir ağı cisim boyutlarını ve sınıf olasılıklarını doğrudan tam boyuttaki görüntülerden tahmin eder. Çoğu görüntü işleme tekniklerinde görüntünün

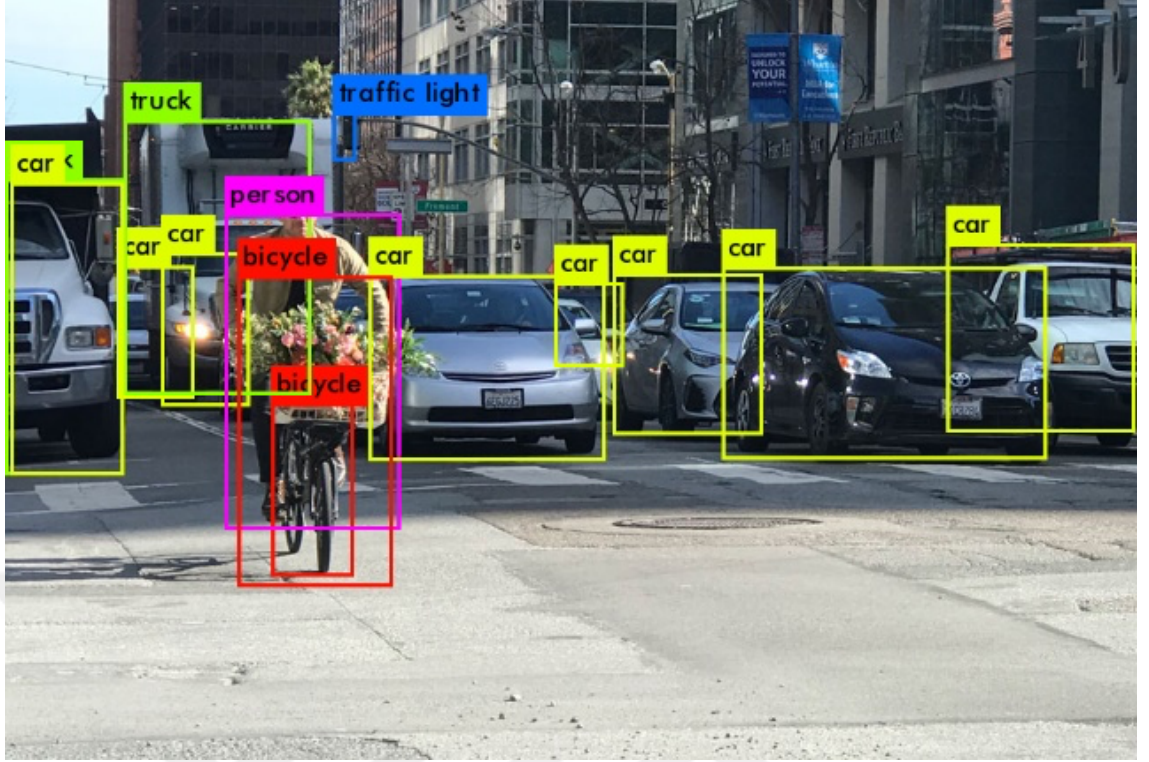
içerisindeki farklı boyuttaki tüm olasılıklar için işlem yapılarak cisim tespiti yapılmaktadır. Bu da cisim tespiti yapmanın yükünü kat be kat arttırmaktadır.



Şekil 2.1 Cascade Classifier ile yüz tespitinin çerçeve seçim örneği (Greg Borenstein tarafından, CC BY-NC-SA 2.0 Lisansı ile paylaşılmıştır)

YOLO için geliştirilmiş CNN modeli, eğitim için ayrılmış veri kümeleri ile eğitildikçe bir dizi işlenmiş veri çıkarılır. CNN modelinin son katmanı olan tamamen bağlı katmanda işlenmiş veri ile cisimler eşleştirilir. Modelin eğitimi bu şekilde gerçekleşir ve doğru sonuçlar elde etmek için Tensor Flow ortamında birden fazla test verisi ile başarı oranı ölçülür. Bu şekilde hem eğitim hem test işlemleri ile bir CNN eğitilmiş olur.

Bilgisayarların çevrelerini anlayabilmesini veya etkileşime geçmesini sağlamak için önemli ihtiyaçlarından biri resimlerdeki ve videolardaki nesneleri otomatik olarak tanımak ve bulmaktır. Çok sayıda akademisyen ve endüstri araştırmacısı otonom sürüş, otomatik güvenlik kamerası gözetlemesi, yardım ve kurtarma operasyonları yardımcı araçları, fabrikalarda çalışan robotlar gibi uygulamalara odaklanarak çeşitli başarılarla imza atmışlardır. YOLO temel sayılabilecek ve yapı taşı oluşturacak nesne algılama ve takip etme konusunda yenilikçi bir çözüm ortaya sunarak dikkatleri üzerine çekmiştir.



Şekil 2.2 YOLO ile tespit edilmiş objelerin görüntüsü

2.2 Sınıflandırma

Günümüze kadar olan süreçte geliştirilmiş birçok sınıflandırma algoritması vardır. Sınıflandırma algoritmalarının birbirleri arasında üstünlük durumları koşullar doğrultusunda elde edilebilir. Sınıflandırma algoritmalarının kendilerine ait artıları eksikleri vardır ya da sınıflandırma algoritmalarının karakterleri vardır diyebiliriz. Bu sebeple sınıflandırma algoritmasını seçerken veri kümesinin içeriğine bakmak gerekmektedir. Örneğin, büyük veri kümelerinde tembel sınıflandırıcıları kullanmak performans sıkıntılarına sebep olabilir veya kompleks veri kümelerinde yapay sinir ağlarını kullanmak sınıflandırıcının veri içerisinde örüntüyü daha iyi algılayıp başarılı bir performans çıktısı üretmesini sağlayabilir.

İlk olarak eğitim verilerini genelleştirmek ve daha sonra tahmin için kullanmak üzere modelleme yapan ve önceden eğittikleri için istekli öğrenen sınıflandırıcılar şunlardır:

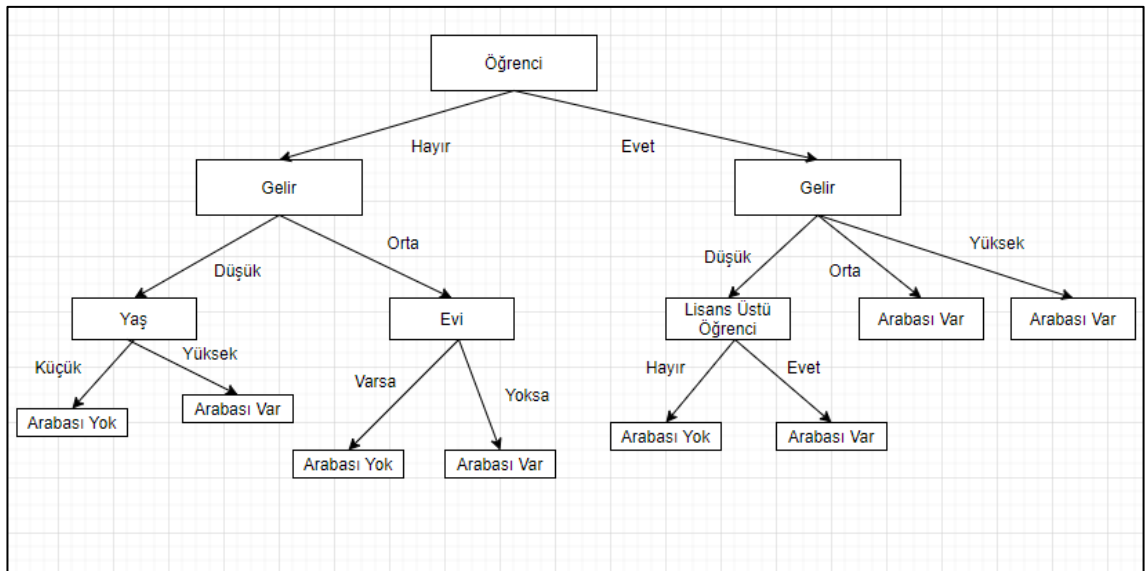
2.2.1 İstekli öğrenen sınıflandırma algoritmaları

2.2.1.1 Karar ağacı

Karar ağacı, bir ağaç yapısı şeklinde bölünmeler oluşturarak sınıflandırma oluşturur. Sınıflandırma için birbirini dışlayan ve kapsamlı bir eğer-sonra (if-then) kural kümesini kullanır. Kural ağacı oluşturulurken eğitim verileri tek tek işlenir. Her bir verinin kuralına göre bölünme gerçekleşir. Eğitim setinde bu bölünme bir sonlandırma koşuluna ulaşana kadar devam eder.

Karar ağacı yukarıdan aşağıya özyinelemeli böl ve fethet tarzında inşa edilmiştir. Karar ağacı elemanlarının nitelikleri kategori edilebilir olmalıdır. Bu koşul sağlanması için ön işlemler yapılabilir. Örneğin bir insanın yaşı kategorize edilemez eğer 40 yaşından büyükler ve küçükler şeklinde kural oluşturulursa kategorize edilebilir duruma gelir.

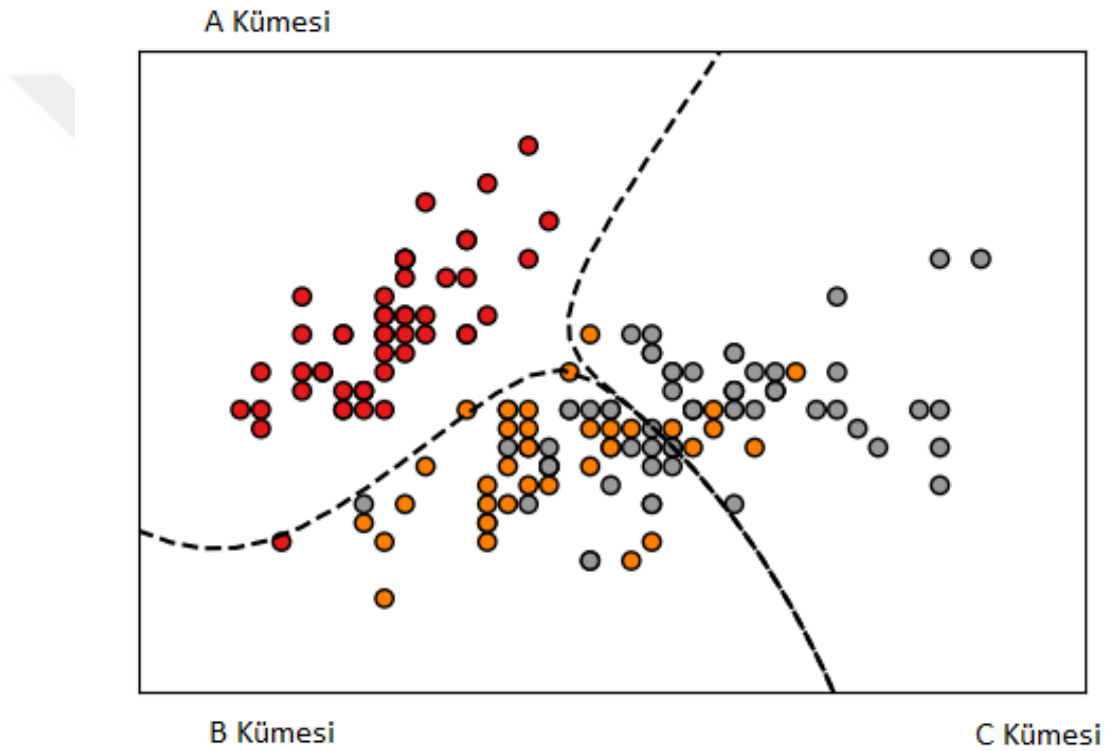
Bir karar ağacı kolayca karmaşıklaşabilir ve çok fazla dal üretebilir, gürültü veya aykırı örnekler nedeniyle anomalileri yansıtabilir. Karmaşıklaşan bir model, eğitim verilerinde etkileyici bir performans sağlamasına rağmen, görünmeyen veriler üzerinde çok düşük bir performansa sahiptir. Karar ağaçlarındaki karmaşıklık budama teknikleri ile önlenilmektedir.



Şekil 2.3 Karar ağacı eğitim sonrası sonuç örneği

2.2.1.2 Naif Bayes

Bayes teoreminden esinlenerek oluşturulmuş olasılıksal bir sınıflandırıcıdır. Bu sınıflandırıcının çalışma şekli bir elemanın her durumun olasılığını bayes teoremine göre hesaplar ve hesaplamada olasılığı en yüksek olan sınıfa dahil edilir. Bu varsayım, yalnızca sınıf dağılımını sayarak hesaplama maliyetini büyük ölçüde azaltır. Varsayım çoğu durumda geçerli olmasa da öznitelikleri bağımlı olduğundan, şaşırtıcı bir şekilde Naif Bayes etkileyici bir performans sergileyebilmiştir.



Şekil 2.4 Naif Bayes eğitim sonrası sonuçların görselleştirilmiş örneği

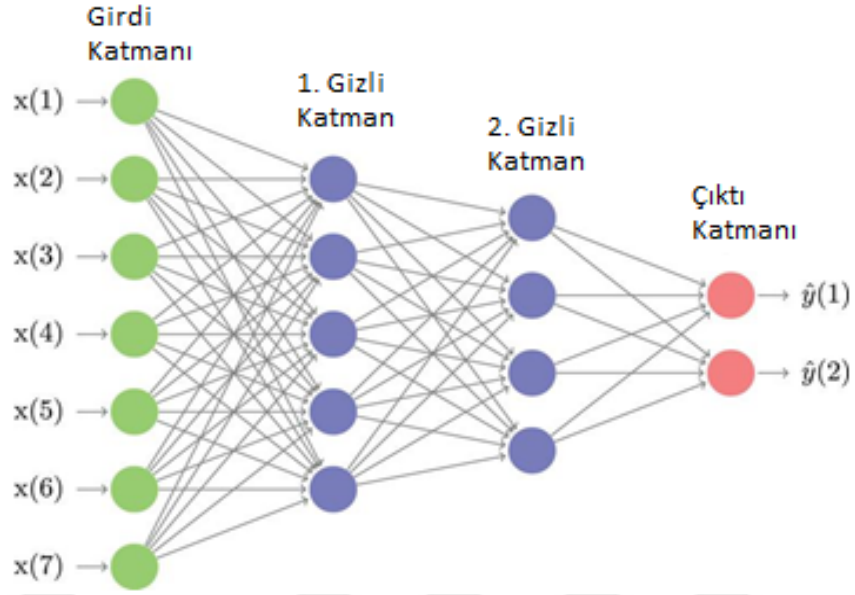
Naif Bayes uygulaması çok basit bir algoritmadır ve çoğu durumda iyi sonuçlar elde edebilmektedir. Hesaplanma karmaşıklığı doğrusal zaman aldığından ($O(n)$) daha büyük veri kümelerinde çalışmak kolay olmaktadır, diğer sınıflandırıcılar genellikle doğrusal zamanda performans göstermediği için büyük verilerde Naif Bayes algoritmasına göre yavaş kalabilmektedirler.

Naif Bayes sınıflandırmasında sıfır olasılık sorunu denilen bir sorun oluşabilmektedir. Bir öznitelik sıfır olduğunda, geçerli bir tahmin veremeyebilir. Bu tarz durumlarla karşılaşmamak için tavsiye edilen çözüm yöntemi Laplacian tahmin edicisidir.

2.2.1.3 Yapay sinir ağları

Bilim insanları dünyanın en iyi bilgisayarı olan beynimizi modellemeye çok fazla kafa yormuşlardır. Beynimizi modelleme çalışmaları sırasında tabi ki öncelikli olarak beynimizin çalışma birimleri üzerinde incelemelerde bulunmuşlardır. Bu incelemelerde sinir hücrelerini ve bu sinir hücrelerinin bağlı olduğu iletim yönü belirli olan sinir hücreleri bağlantılarını görmüşlerdir. Bu modeli bilgisayarın anlayacağı şekle getirmeye çalışmışlardır. Bu çalışmalar neticesinde nörona benzer şekilde bir düğüm oluşturulmuş ve birden fazla düğüm birbirine bağlantılar sayesinde bağlanmıştır. Her bir nöronun ve bağlantının kendi sayısal ağırlığı tanımlanmıştır. Bu ağırlık üzerinden gönderilen sinyal değerleri yeniden hesaplanır. Nöronların ürettiği sinyaller belirli eşiği geçerse gönderilecek şekilde bir eşiğe sahip olabilir. Tipik olarak, nöronlar katmanlar halinde toplanır. Farklı katmanlar, girişlerinde farklı dönüşümler gerçekleştirebilir. Girdi ve çıktı sinyalleri içeren giriş ve çıkış katmanı mevcuttur.

İleri beslemeli, geri beslemeli, evrimsel vb. gibi mevcut birçok ağ mimarileri vardır. Yapay sinir ağlarındaki verim uygun mimari modelin uygulamasına bağlıdır. Çoğu durumda geri besleme modelleri makul derecede doğru sonuçlar verir ve özellikle görüntü işleme uygulamaları için evrimsel ağlar daha iyi çalışır.



Şekil 2.5 Sınıflandırmada kullanılabilecek bir yapay sinir ağı örneği (Matthew Dixona vd. 2017)

Model tarafından eşlenecek olan fonksiyonun karmaşıklığına bağlı olarak modelde birden fazla gizli katman olabilir. Derin sinir ağı gibi karmaşık ilişkilerin modellenmesini sağlayabilmek için, sayısal olarak daha fazla gizli katman kullanılabilir. Ancak kullanılan verinin karmaşıklığı ile yapay sinir ağı modelinin uyumlu olmasına dikkat edilmelidir. Eğer fazla gizli katman koyarsanız eğitim modelini öğrenmekten daha çok ezberlemesine sebebiyet verebilir. Bununla birlikte, birçok gizli katman olduğunda, eğitim çok zaman alır. Diğer dezavantajı, öğrenimi gerçekleştiren yapılar karmaşık matematiksel işlemler içerdiği için Karar Ağaçları gibi diğer modellere kıyasla modelin zayıf yorumlanabilir olmasıdır. Örneğin bir Karar Ağacını sınıflandırma algoritmasının sonucunun haritasına baktığınızda hangi verinin ne içerdiğini gözlemleyebilirsiniz fakat Yapay Sinir Ağlarında sadece uzun virgüllü sayılar göreceksiniz.

Ancak Yapay Sinir Ağları gerçek dünyadaki uygulamaların çoğunda etkileyici bir performans sergilemiştir. Gürültülü verilere karşı yüksek toleranslıdır ve eğitimsiz kalıpları (Unsupervised Learning) sınıflandırabilir olmasından ötürü etkileyici performans göstermiştir. Genellikle, Yapay Sinir Ağları sürekli giriş ve çıkış değerleri içeren modeller ile daha iyi performans gösterir.

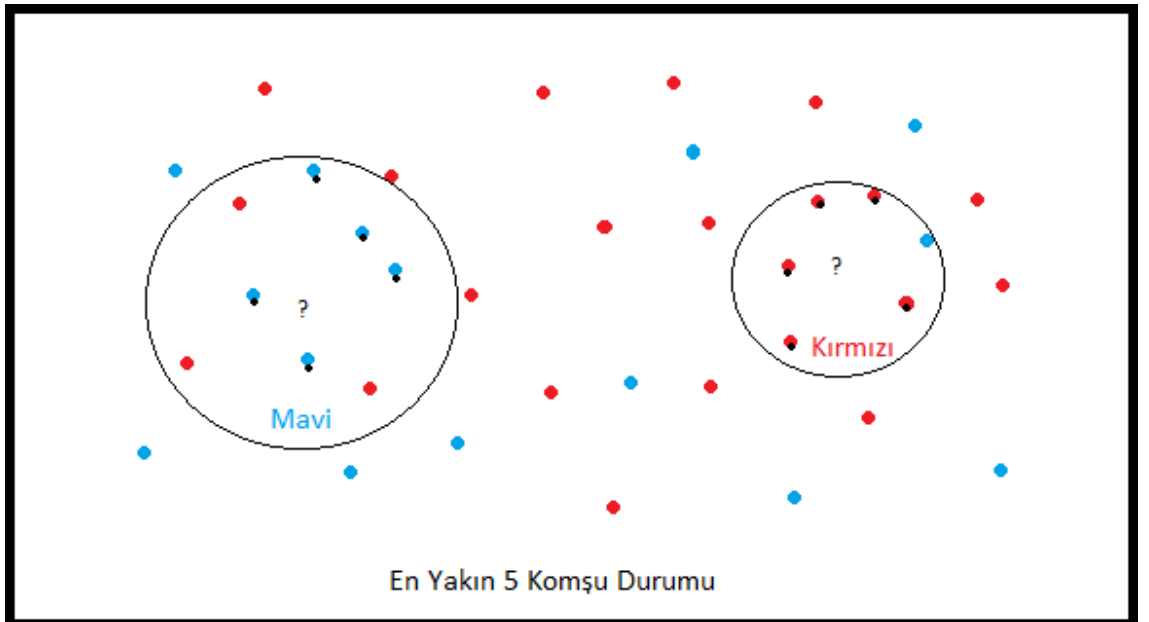
2.2.2 Tembel öğrenen sınıflandırma algoritmaları

Pek çok tembel öğrenen sınıflandırma algoritması bulunmaktadır. Bunlar arasında en dikkat çeken tembel öğrenen sınıflandırma algoritması K-En yakın komşu olduğu için sadece bu algoritma anlatılmaktadır.

2.2.2.1 K-En yakın komşu (KNN)

K-En yakın komşu sınıflandırma algoritması öğrenmiş olduğu eğitim verilerini N boyutlu bir uzaya yerleştiren bir tembel öğrenme algoritmasıdır. Bir sınıflandırma işlemini yaparken yerleştirmiş olduğu eğitim verilerini içeren uzay içerisinde K tane en yakın komşuya bakarak, en çok hangi sınıftan veri varsa o sınıf bilgisini cevap olarak döndürür.

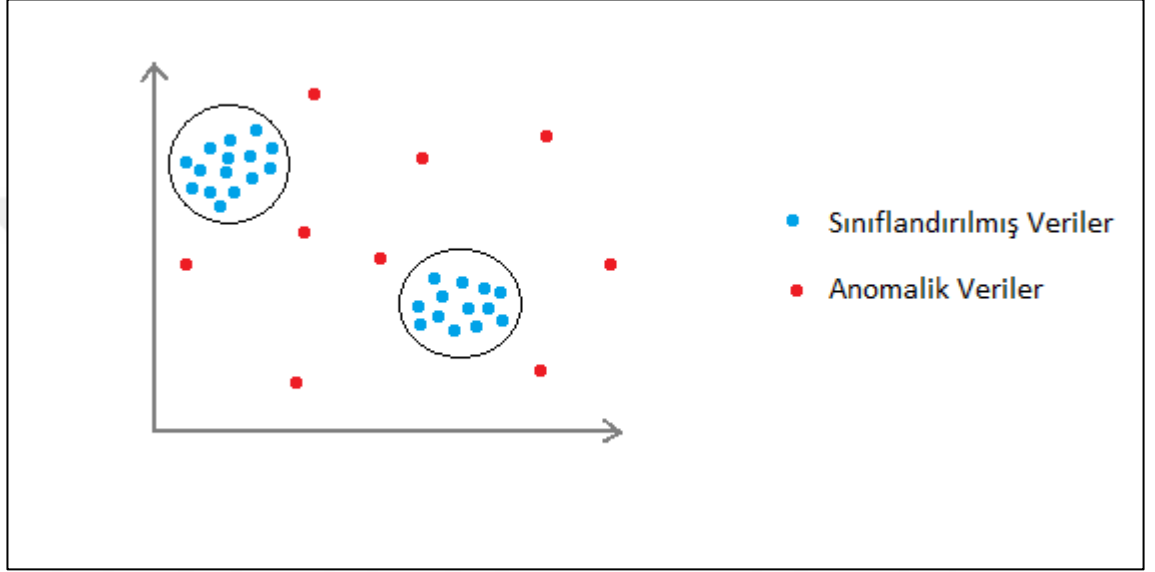
Mesafe ağırlıklı en yakın komşu algoritmasında, en yakın komşulara daha fazla ağırlık veren aşağıdaki sorguyu kullanarak k komşularının her birinin katkısını kendi mesafelerine göre ağırlaştırır. Genellikle KNN, en yakın komşulara bakarak ortalama cevap ürettiği için gürültülü verilerden daha az etkilenir.



Şekil 2.6 K-En Yakın Komşu öğrenimi sonucunun görselleştirilmesi örneği

2.3 Anormallik Tespiti

Anomali tespiti, bir veri gözlemi sırasında daha önce görülenlerden büyük ölçüde farklılık gösteren verileri bulma işine denir. Yapay zekanın temel sorunlarından biridir. Tez çalışmamızda bulunan işleyişte anomali tespiti yapmak kilit görevlerimizden birisidir. Çünkü Video Özeti çıkartılacak yerler için anomali içeren durumlar seçilecektir.



Şekil 2.7 Anormallik tespiti görselleştirilmiş örneği

Algılanan verilerdeki anormallikleri tespit etmek insanlar ve yapay zeka için önemli bir yetenektir. İnsanlar genellikle tehlikenin erken belirtilerini saptamak ya da fırsatları önceden keşfetmek için anomalileri tespit etmek ister. Anomali algılama sistemleri yapay zekâ tarafından kredi kartı dolandırıcılığını saptamak için kullanılmaktadır. Siber güvenlikte izinsiz girişi tespit etmek, endüstriyel ekipmanların bozulabileceği öngörüsünde bulunarak bakımının yapılması konusunda uyarmak ve borsada cazip fırsatları keşfetmek için anomali tespiti kullanılır. Tipik anomali algılama yöntemi, hedefin verilerini normal veya anormal olarak sınıflandırmanın olduğu tek sınıflı bir sınıflandırma görevidir.

Anomali algılama yöntemleri için farklı olası senaryolar vardır. Denetimli anomali tespitinde, normal ve anormal kalıplara ilişkin eğitim örnekleri anomali tespiti yapılma işleminden önce eğitim amacıyla verilmektedir. Bu şekilde ön eğitim verilerek geliştirilen

anomali tespit araçları eğitim verilmeyenlere göre daha iyi çalışabilir. Fakat her zaman eğitim verebiliyor olmak mümkün değildir. Eğitimin verilebiliyor olması için karşılaşılabileceğiniz anormalliklerin ne olduğunu biliyor olmak zorundasınız. Bazı alanlarda, hiç karşılaşılmamış yeni anomaliler tespit etmek gerekebilmektedir. Örneğin siber güvenlikte kullanılan güvenlik duvarında, bilinmeyen bilgisayar virüslerinin tespit edilmesi gerekir. Karşılaşılabilecek tüm anormalliklerin örnekleri olmamasına rağmen denetimsiz öğrenme ile geliştirilmiş algoritmalar bildikleri ve bilmedikleri anormalliklerin tespitini yapabilmektedir.

Son zamanlarda, sınıflandırma tabanlı yöntemlerin anormallik tespitleri üzerinde oldukça iyi sonuçlar ürettiğini gözlemliyoruz. Sınıflandırma tabanlı yöntemlerde aynı zamanda yarı denetimli senaryolar da çalıştırılabilmektedir.

2.4 Video Özeti Çıkartma

Video özetinin oluşturulması, videonun içerdiği kilit karelerin veya klip şeklindeki parçaların videodan çıkarılmasına dayanan bir yöntemdir. Video başlangıçta televizyon sistemleri için geliştirilmiş olsa da birçok kullanıcı grubu tarafından kullanılan, çeşitli uygulamalarda günümüzde yaygın olarak tercih edilen bir araçtır. Video içerikler, filmler, gözetim amaçlı kullanımlar, spor müsabaka kayıtları, ders kayıtları, tıbbi kayıtlar ve haberler gibi birçok alanda kendine yer bulmuştur. Bu içeriklerin artması sonucu video işleme teknolojileri ihtiyaç haline gelmiştir. Video işleme teknolojisinin ilerleme kaydetmesi ile içerik miktarında da belirgin bir artış gözlenmiştir.

Video özetinin adından da anlaşılacağı gibi, uzun videoların kısa sürede yorumlanabileceği bir videonun özetini elde etmek için video özeti kullanılır. Kullanıcının önemli olarak tanımlayabileceği parçalar özet videonun içeriğini oluşturmaktadır.

3. LİTERATÜR

Video özetleme konusundaki literatür taramasının kısa özeti şunlardır:

Video özetleme uygulamaları ve video özetlemesinin ilk kullanımı medya, TV yayıncılığındaydı. Film konusunda videoyu, birleştirilmiş resim özetine çıkartan bir çalışma yapıldı (Tao Chen, 2012). Bir filmin hikayesini otomatik olarak oluşturarak özet çıkartma işlemi için uygulama yapılmıştır.

Filmlerin yanı sıra spor müsabakaları da öne çıkan kısımlarının video özeti haline getirilmesi bilim insanlarının hem dikkatini çekmiş konulardır hem de ticari kaygı güdülebilen bir alandır. 2011 yılında bir çalışmada futbol maçlarının maç içerisindeki durum özelinde özetinin çıkartılması hedeflenmiştir (Hossam M. Zawbaa, 2011). Maç içinde gol olması, penaltı kullanılması gibi durumlar özet içine dahil edilerek ticari bir yanı olan araştırmaya imza atılmıştır.

Video kayıt cihazlarının günümüzde ortalama yaşantıdaki birisinin alabiliyor olması kolaylaştıkça, herhangi bir zamanda çeşitli yerlerde insanlar tarafından kaydedilen video miktarını arttırmıştır. 2011 yılında yapılan bir araştırmada günlük çekilen videoların analizlerinin sonucunda şahıs davranış analizinin yanı sıra videolarının verimli depolanmasını sağlayan bir video özetleme araştırması yapılmıştır (Wei-Ting Peng, 2011). Kural tabanlı geliştirilmiş olan algoritmalarla videolar analiz edilmektedir.

2015 yılında yayımlanan bir rapora göre, 2015 yılında profesyonel olarak kurulmuş gözetim sistemlerinde 250 milyondan fazla kamera yüksek çözünürlüklü ve internet ortamına kapalı ağlarda yer almaktadırlar. Çok ciddi miktarda güvenlik personeli sadece bu gözetim videolarını izlemek için görevlendirilmiştir. Bunun yerine, bu zahmetli çalışma, uzun gözetim videolarının bir vurgulanmasını sağlayan akıllı bir video özetleme sistemi tarafından gerçekleştirilebilirdi. Kritik durumların video özeti çıkartıldığında ve özellikle bir yerin gözetlemek için birden fazla kamera kullanılmış olduğunda, güvenlik personeli tarafından yorumlamak çok kolaylaşacaktır. Çoklu ekran özetler ile, videoların önemli kliplerini hızlı bir şekilde taranmasını kolaylaştırmaktadır. Bu sorun aslında bizim tez çalışmamızın ilgi alınana girmektedir. Bu alanla ilgili bir araştırma

2016'da yayımlanmıştır (Po Kong Lai, 2016). Obje tespiti üzerinden bir algoritma kullanmaktadırlar.

Kapalı devre işi olup aynı zamanda kendine özgü içeriği olan bir alan da endüstri binalarının içerisindeki kamera gözlemleridir. Endüstri binalarının içerisine kameraların yerleştirilmesi yaygınlaşmıştır. Bu kameralardan bilgi alabilmek, oluşmuş olan hataların Makine Öğrenmesi teknikleri ile tespit edilmesini sağlamak için araştırma yapan bir grup 2011 yılında bir yayın yapmıştır (Athanasios S. Voulodimos, 2011). Bu yayında Sınıflandırma tabanlı bir Video Özetleme algoritmasını anlatmaktadırlar.

Tıp dünyasında da Video Özetleme algoritmaları geliştirilmiştir. 2011 yılında bu konuda bir araştırma yayımlanmıştır (Sourya Bhattacharyya, 2011). Bu çalışmada özel olarak yeni doğan bebekleri gözlemleyen kameralara bebeklerin nöbet geçirme durumlarının saptanması sağlanmıştır. Yoğun bakımlardaki her bebek için bir kişinin 24 saat gözetim yapması gerekirken, yapılan çalışma sayesinde bebekleri izleyen kameralar yardımı ile güvenli bir şekilde 24 saat kesintisiz bebeklerin nöbet geçirme durumları gözlemlenmektedir.

Giyilebilir kameraların gelişi, şahsın gözünden video kaydedilmesini sağladı. Bu videolar benmerkezci video olarak adlandırıldı. Benmerkezci videoları kaydetme yeteneği, sosyal davranışsal modellemeyen bireyin günlük kayıtlarındaki tekrarlanan kalıpları incelemeye kadar değişen sayısız bilgisayar görme zorlukları için bir kapsam yarattı. Doğası gereği uzun videolar olduklarından, kamera taşıyıcısının önünde görünen nesneleri ve olayları verimli bir şekilde analiz etmek için benmerkezci videolar özetlenmelidir. 2015'te yapılmış olan bir çalışmada, benmerkezci videolardan gün içerisinde yapılanların video özeti oluşturma üzerinde çalışma yayımlandı (Jia Xu, 2015). Karşılaştıkları problemler içerisinde kalitesiz net olmayan görüntülerin tespiti, benmerkezci video çeken kişinin niyetinin tespit edilmesi bulunmaktadır. Çalışmanın kapsamı bu problemler doğrultusunda genişlemektedir.

Benmerkezci videolar, kısa süreli hafıza kaybı yaşayan insanlar için bir nimettir, çünkü tüm gün boyunca faaliyetleri kaydedilir. Günlük kayıtlar içerisinde aranan bir eylemi bulmak oldukça zordur. Bu tür sağlık koşullarına sahip insanlara, bu kayıtların kısa bir

özetini oluşturmak çok yardımcı olacaktır. 2017’de hafıza kaybı ve benzeri şekilde hastalığı olan kişiler için bir çalışma yayımlandı (Rameswar Panda, 2017). Bu çalışmada günlük çekilen benmerkezci videolar özetlenerek depolanması sağlanıyor. Hasta birisi bir konu, yer, cisim vs. hakkında bilgi almak istediğinde arama motoru desteği ile aranan konuya ait özet videolar karşısına çıkması sağlanmaktadır. Aynı zamanda özet videonun kesildiği tam videoya da erişimi sağlamaktadır.

2011 yılında yayımlanmış olan yazıda, esas olarak çevrimiçi videolara uygulanmak üzere özelleştirilmiş video özetleri oluşturmak için yeni bir yöntem önerilmektedir (Dim P. Papadopoulos, 2011). Bu yaklaşımın yeniliği, video özetleme probleminin her sorguda sadece tek bir görüntü alabilme olarak kabul edilmesidir. Bahsedilen yöntem şudur, her çerçeve ayrı bir görüntü olarak kabul edilir ve Kompakt Tanımlayıcıları (CCD) ve görsel bir sözcük histogramı tarafından açıklanması sağlanacaktır. Yöntem, çerçeveleri kümeler halinde sınıflandırmak için güçlü bir Kendi Kendini Yetiştiren ve Kendi Kendini Organize Etmiş Nöral Gaz (SGONG) ağı kullanılmıştır. Başlıca avantajı, oluşturulan nöronların sayısını ve topolojisini otomatik bir şekilde ayarlamasıdır. Böylece, eğitimden sonra SGONG bize uygun sayıda çıktı sınıfları ve merkezleri vermesi sağlanmıştır. Her kümeden temsili bir anahtar karenin çıkarılması, video özetinin oluşturulmasını sağlamıştır.

2014 yılında çıkartılmış olan bir yazıda, Gauss karışım modeli kullanılarak çevrimiçi özetleme yöntemi denenmiştir (Shun-Hsing Ou, 2014). Deneyler neticesinde kullanılan yöntem hem özetleme kalitesinde hem de hesaplama verimliliğinde diğer çevrimiçi yöntemlerden daha iyi performans gösterdiği söylenmektedir. Daha kısa bir gecikme süresi ve çok daha düşük hesaplama kaynağı gereksinimleri ile özetleme işlemi oluşabileceği savunulmuştur.

2017 yılında yapılmış olan bir yayımda içerik tabanlı video alma, indeksleme ve öneri sistemleri çalışması anlatılmaktadır (Yashar Deldjoo, 2017). Hedeflerinin kullanıcılara, videoların görsel içeriğine dayalı öğeler önermek olduğu belirtilmiştir.

Video özetlemesinin özel bir sınıfı olan film özeti hakkında 2013 yılında çalışma yapılmıştır (Chia Ming Tsai, 2013). Yayımlanan yazıda, bir filmin içerisinde önceki

sahnelerdeki rol toplulukları genellikle sonraki sahnelerde rol ilişkisini geliştirmek için kullanıldığından; rol toplulukları arasındaki ilişkiyi tespit edip, ilişkiler ağını çıkartan sahne tabanlı bir film özetleme yöntemi geliştirilmiştir. Analiz aşamasında, rol toplulukları arasındaki etkileşimleri karakterize etmek için bir sosyal ağ oluşturulmuş. Özetleme aşamasında, sahnelerin uygulanabilir özet kombinasyonları belirlenip ve sosyal güç korumasına dayalı olarak bu adaylardan bilgi açısından zengin bir özet seçilmesi önerilmiştir.

2012 yılındaki bir araştırmada büyük videolarda kullanılan özetleme yöntemlerini kıyaslanmıştır (Muhammad Ajmal, 2012). Her yöntemin en uygun olduğu durum tartışılmıştır. Gereksiz video karelerini işleme yükünü azaltacak ve çok büyük videolarda hızla gezinmeye yardımcı olacak öneriler verilmiştir.

Video özetlemesinin bir diğer önemli uygulaması, endoskopik videolar, tıbbi prosedürlerin kayıtları, teşhis histeroskopi videoları ve benzeri tıbbi videoların analizinde kullanılmasıdır. Tıbbi videolar, tıbbi prosedürlerin süresine bağlı olarak büyük ölçüde değişir. Uzun tıbbi videoların özetlenmesi, prosedürlerin verimli analizini sağlar ve tıp öğrencilerine prosedürlerin öğretilmesine yardımcı olur. 2011 yılında yayınlamış olan bir yazıda Bronkoskopi muayeneleri üzerine bir araştırma yapılmıştır (Mikołaj I. Leszczuk, 2011). Tüm prosedür bir video haline getirildiğinde önemli kısımlar özetlenerek benzer vakalar kıyaslanmaktadır. Bu şekilde 20 dakikaları bulan tanı, teşhis aşaması hızlandırılması sağlanacaktır.

Uzmanların genellikle benzer vakaları araması, birçok tanısal histeroskopi videosu içeren kütüphanelere göz atmaları veya belirli bir vakanın videosunu gözden geçirmeleri gerekir. Video arama ve tarama, daha önce teşhis edilen vakaların videolarının karşılaştırıldığı vaka tabanlı tanıda, vaka yönlendirmelerinde, hasta kayıtlarının gözden geçirilmesinde ve tıbbi araştırmaların desteklenmesi (ör. insan üremesinde) gibi birçok durumda kullanılabilir. 2011 yılında yayınlanmış olan bir yazıda, tarama görevini kolaylaştırmak, tanısal histeroskopilerdeki video parçalarının klinik alaka düzeyini tahmin etmek için bir teknik önerilmiştir (Wilson Gavião, 2011). Temel olarak, önerilen teknik, klinik ilgiyi, video edinimi sırasında bir tanısal histeroskopi video parçası tarafından çekilen dikkat ile ilişkilendirilir. Mesela bir görüntüde bir miktar bekleme

süresi varsa orada doktorlar inceleme yapıp dikkatleri çektiği anlamına geliyor. Yapılmış bu çalışmada, genel videoları özetleme yöntemleri ve bunların tanısal histeroskopi videolarına nasıl uygulandığını gözden geçirilmiş. Önerilen yöntemin video için özel bir özetleme ve gösterim yöntemi ile histeroskopi alanına katkıda bulunduğu sonucuna vardıklarını söylemektedirler.

Haber videolarını özetlemek, haber videolarında görünen önemli kalıpları hızlı bir şekilde geçmemizi sağlar. Bu alanda 2011 yılında bir yazı kaleme alınmıştır (Jie Ling Lai, 2011). Hareket tespiti, özellik çıkartma gibi bilinen yöntemleri kullanarak sonuca ulaşmaya çalışmışlardır.

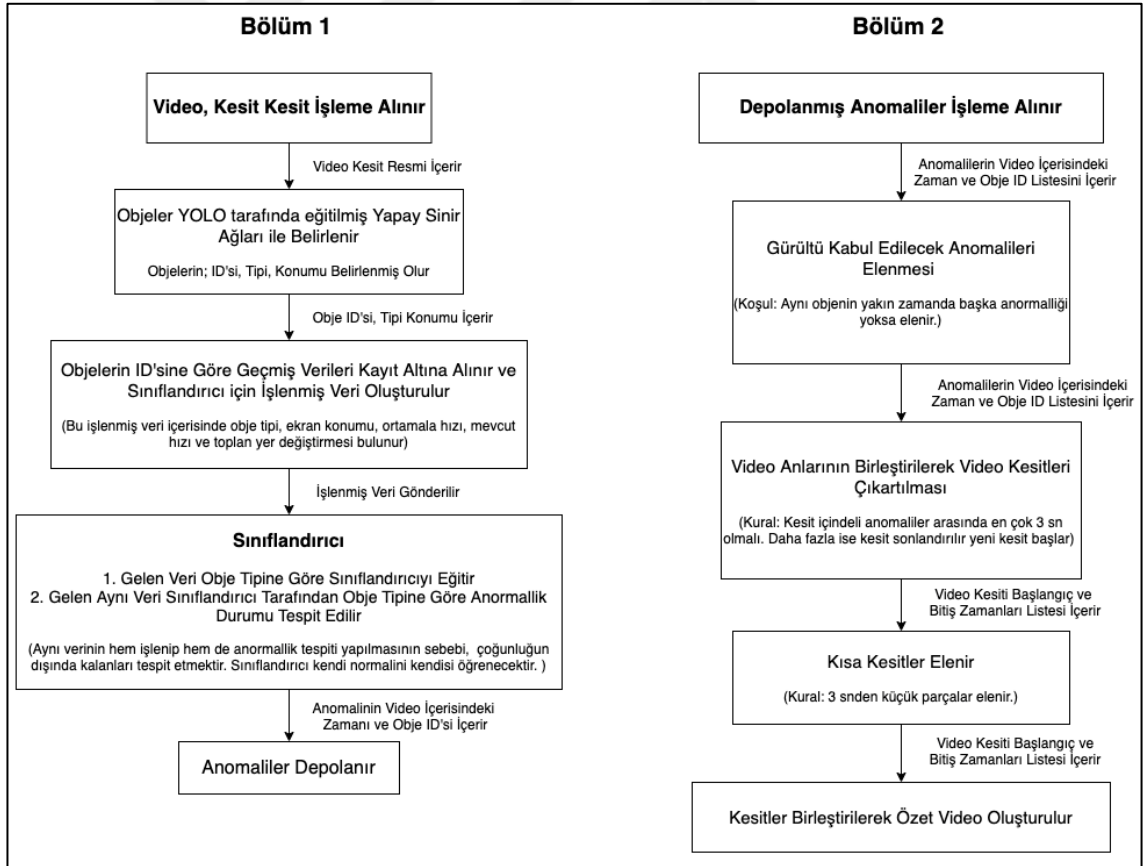
İnsansız hava aracı (İHA) ve uzaktan kumandalı uçağa dayalı gözetim her geçen gün yaygınlaşıyor. Kameralar, insansız hava araçları ve robotlar gibi yeni cihazlara entegre ediliyor. Bu sayede insanların ulaşamayacağı şartlarda kayıt imkanları sağlanıyor. Bu videoları özetlemek, bu yeni video kategorisinin yorumlanmasını kolaylaştıracaktır. Bu amaç ile alakalı 2012 de yayımlanmış bir yazı kaleme alınmıştır (Athanasios S. Voulodimos, 2012).

Literatür taramasında birçok farklı türde amaçta video özeti uygulaması, araştırması incelendi. Genel olarak aynı tarz değil de farklı amaçlar içeren video özeti çıkartma amaçlayan yazıları incelenmeye çalışıldı. Daha birçok alanda video özeti kullanıldığını, yukarıda yer alan çalışmaların video özetleme alanındaki tüm literatürün çok küçük bir kısmı olduğu görülmüştür.

4. METODOLOJİ

Tez çalışmamızda Python programlama dili tercih edilmiştir. Bunun başlıca sebebi, Python programlama dili ile günümüzde kullanılan birçok yapay zekâ kütüphanesi geliştirilmiştir. Geniş kütüphane imkanları sağlamaktadır.

Python programlama dili ile OpenCV kütüphanesini kullanarak önceden eğitilmiş olan YOLO aracının obje tespiti yapan yapay sinir ağları ile bilgisayar görme işlemi yapılmıştır. Tespit edilen objelerin bilgileri normalleştirilerek sınıflandırıcıya gönderilir. Bu objenin bilgileri ile objenin önceden bilinen tipine göre sınıflandırıcıdan anormallik içerip içermediği tespit edilir. Aynı objenin bilgileri, anormallik durumuna bakılmaksızın sınıflandırıcıya eğitim verisi olarak gönderilir. Anormallik tespiti yapılmış olan video kesitleri çeşitli filtre ve işleminden geçerek özet işlemi oluşturulur.



Şekil 4.1 Video özeti çıkartma akış şeması

Bu akış üzerinde kullanılan algoritmalar şunlardır: görüntü tespiti için evrişimli yapay sinir ağları algoritması ve anomali tespiti için de sınıflandırma algoritması kullanılmıştır. Tez çalışmamızın video özeti çıkartma akış şeması Şekil 4.1’de anlatılmaktadır. Akış şeması içerisinde kurallar örnek olarak verilmiş olup değişkenlik gösterebilmektedir.

4.1 Hedef ve Motivasyon

Tez çalışmamızın hedefi sabit kameralarla çekilmiş olan video görüntüleri üzerinde anomali tespiti yaparak, video özeti çıkartmaktır. Anomali tespiti için bir adet obje tespiti ve obje takibi yapan YOLO isminde araç vardır. Tespit edilmiş objeler kameraya girdikleri andan itibaren tüm hareketleri kaydedilir ve sınıflandırıcıya gönderilmeye uygun veri tipine dönüştürülür. Sınıflandırıcı dönüştürülmüş veri ile eğitilir. Sonrasında sınıflandırıcıya objenin kendi sınıfına ne kadar ait olduğu sorulur. Mesela bir araba objesi tanındı ve takip ediliyor. Sınıflandırıcıya bu araba ne kadar araba gibi davranıyor diye sorulur. Eğer sınıflandırıcı araba gibi davranmadığını söylüyorsa demek ki bir anomali mevcut olduğunu varsayılır.

Sınıflandırıcı hem eğitilir hem de sınıflandırma sorgusunda kullanılır. Bu bir tezat gibi gözükebilir. Fakat sınıflandırıcı uzun vadede sürekli eğitime tabi tutulduğu için farklı davranış gösteren ile göstermeyenleri ayırt edebilir duruma gelmektedir. Bunu “çan eğrisinin altında kalan öğrenciler sınıfta kalır” kuralı ile benzer olduğu düşünülebilir.

4.2 Videonun Hazırlanması

Videonun hazırlanma kısmının olmasının sebebi, özeti çıkartılmış bir videonun parçaları sadece birleştirilmiş şekilde sergilenemez. Videonun özeti oluşturulmuş orijinal video hakkında bazı bilgiler içermelidir. Özet videoda orijinal videonun hangi saniyesinde olduğunu bilmek oldukça önemlidir. Bizim tez çalışmamızın temelinde yatan anomali tespitini de özet videoda sergilemek önemli olduğunu düşünülmüş olup bu konular hakkında çalışma yapıldı.

4.2.1 Videoya zamanlayıcı eklenmesi

Video içerisinde istenilen parçalarının sadece kesilip çıkarılması ve birleştirilmesi yetmemektedir. Çünkü özet video izlenirken orijinal videonun hangi saniyede olduğu anlaşılması mümkün kılınmamaktadır. Bu sebepten dolayı özet haline getirilmiş videoya orijinal videonun sol üst kısmına video içerisinde zaman göstergesi yerleştirildi.

Başka yöntemlerle de özet videoya zaman bilgisinin eklenmesi mümkündür. Örneğin video özetleyen büyük bir proje içerisinde olursa kendi video formatını yaratmak oldukça olasıdır. Tabi ki böyle bir video formatı için kod çözücü (codec), video oynatıcı vs. ek geliştirmeler yapılmak zorundadır. Ek geliştirmelerden kaçınmak istenmesi oldukça doğaldır. Fakat ek geliştirmelerin ek kabiliyetler getireceğini unutmamak gerekir.

Başka bir yöntem olarak özetlenmiş videonun çözünürlüğü aşağı kenarından veya sağ kenarından büyütülüp video detaylarını oraya eklemek de mümkündür.



Şekil 4.2 Örnek video üzerine zaman bilgisi eklenmesinin örnek görseli

4.2.2 Video üzerine anomali tespitlerinin çizimleri

Tez çalışmamız test edilirken ihtiyaç hissedilen bir başka özellik daha eklendi. Video özeti izlenirken hangi özet parçasının hangi anomaliye ait olduğu belirtilmek istendi. Anlatıldığı üzere tez çalışmamız objeleri takip ederek objelerin anomalilerinin tespit edilmesi üzerine çalışmaktadır. Anomali tespiti yapılmış objeler bilinmektedir aynı zamanda özet videonun içerisinde de bu objelerin anomali anlarını işaretli görmek video özeti izlenme kalitesini yükseltecektir.

Bu fikir üzerine anomali tespiti yapılmış olan objelerin anomali anları kare içine alındı. Video özeti izleyen kişilerin takip kalitesi artırıldı.

Video özeti sırasında tespit edilmiş anomalilerin gösterimi için başka yöntemler de kullanılabilir. Örneğin video özet formatı yapıp bu format içerisinde ses ve görüntü verilerinin paraleline anomali durumların bilgileri de eklenebilir. Böyle bir formatın oynatılması için bir video özeti oynatıcı da bu ekseninde geliştirilmelidir.



Şekil 4.3 Video özetinde işaretlenmiş olan anomali gösteren objenin görseli

4.3 Video Özet Parçalarının Belirlenmesi

Video özet parçalarının belirlenmesi, video özet çıkartma işleminin en önemli safhasıdır. Neredeyse tüm algoritma tabanlı işlemler bu safhada yapılmaktadır. Bu safhada 2 teknoloji kullanıldı ve devamında çeşitli algoritmalar geliştirilerek bu safha tamamlandı.

4.3.1 YOLO kullanımı

YOLO (Yalnızca Bir Kez Bakarsınız), evrişimli sinir ağları ile geliştirilmiş olan obje tanıma ve takip etme kabiliyetleri olan bir makina görme aracıdır. Geliştirilmesindeki temel farklılık bir görüntünün işlenmesi sırasında resim içindeki tüm alt resim parçaları işlenmemekte bir resim parçası sadece bir kere tüm olarak işlenmektedir. Birçok makine görme algoritmalarında bir resim alt parçalarına bakılarak işlenmektedir.

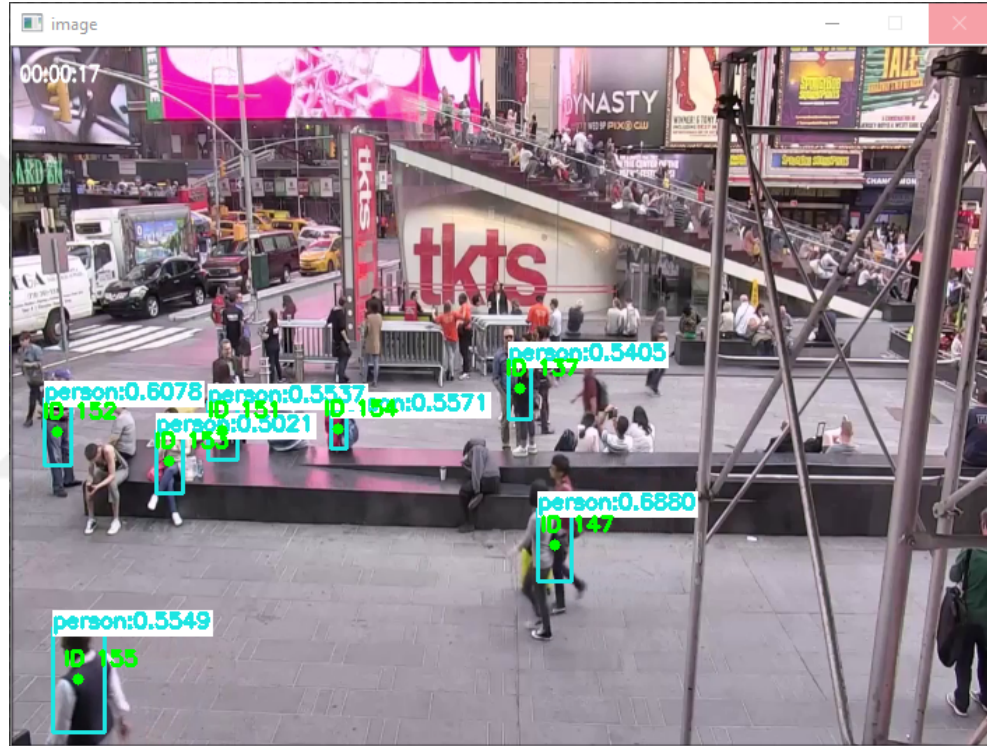
Bilindiği üzere evrişimli sinir ağları çok fazla işlem gerektiren CPU ile çalıştırıldığında fazlaca yavaş olabilen bu sebepten dolayı GPU ile işlem yapmak için uygun olan makina görme algoritmalarıdır. Tez çalışmasında kullanılmış olan YOLO aracının yavaş olmasından kaçınmak için GPU ile çalıştırılması gerekmektedir. Bu yüzden çalışmamızda kullanılan bilgisayarlarda en az bir adet GPU ve CUDA araç takımı yüklü olması gerekmektedir. YOLO aracında python programlama dilinin içerisinde CUDA destekli Open-CV kütüphanesi kullanılmaktadır. CUDA araç takımının spesifik olarak CuDNN araçları kullanılmaktadır.

Tezimiz için yapmış olduğumuz çalışmada YOLO aracı ilk işlemi yapan referans noktamızdır. Bir videonun çerçevesi ilk olarak YOLO programına gönderilmektedir. YOLO aracı çıktı olarak bir liste dönüş yapmaktadır. Bu listenin içindeki her bir eleman aşağıdaki verileri bulunmaktadır:

1. Objenin YOLO tarafından atanmış olduğu kimlik numarası (Eşsiz olacak şekilde sayı içerikli oluşturulur)
2. Objenin tipi (araba, insan, kamyon, köpek vs. toplam 86 adet obje tipi içermektedir)

3. Objenin ekran üzerindeki konumu (piksel olarak yatay ve dikey konumu)
4. Objenin ekran üzerindeki genişliği (piksel olarak eni ve boyu)
5. Objenin tipi ile uyum olasılığı (0 ile 1 arasında çıktı üretilir. Uyum olasılığına bir sınır konularak objeler tanımlanır. Örneğin 0.5'in altı sonuçlar görmezden gelinir.)

YOLO'dan alınmış olan bu liste işlenmek üzere sınıflandırıcı adımına gönderilmektedir.



Şekil 4.4 YOLO ile işlenmiş bir çerçevenin görseli

4.3.2 Sınıflandırıcı kullanımı

Tezimiz için yaptığımız çalışmada sınıflandırıcı kullanmak istenmesinin sebebi, bir video işlerken her durum için ayrı ayrı olasılıkları düşünmek yerine kendi doğrusunu öğrenen bir yapı olmasıdır. Örneğin bir yolu gözleyen kameranın görüntüleri incelenirken ters yönde giden araçları tespit etmek istiyoruz. Normalde ilk aklımıza gelen çözüm yolun yönünü çalışmamızda tanımlarız ve kim ters yönde gidiyorsa tespit edebiliriz. Bu çözüm sadece bir yol ve bir zaman aralığı için geçerli olacaktır. Yani aynı programa sahip

kamerayı görüş açısı başka bir tarafa bakan başka bir yola koyduğumuzda program düzgün çalışmayacaktır. Ya da yol üzerinde bir tadilat olduğunda yol çift şeritli olduğunda program yine düzgün çalışmayacaktır. Tezimiz için yapmış olduğumuz çalışmada bu tarz ezbere çözümlerin denemesi yerine tüm ihtimalleri kapsayacak bir çözüm üretilmeye çalışıldı.

Tüm ihtimalleri kapsayacak olan bu çözümün içerisinde sınıflandırıcının görevi kilit noktadadır. Sınıflandırıcı sabit bir kamera görüşündeki YOLO aracı ile tespit edilmiş ve takip edilmekte olan objeleri sınıflandırmaya çalışmaktadır. Yani objenin zaten cinsi belli iken sınıflandırıcıya bu objenin hareketi öğretilir. Sonrasında aynı sınıflandırıcıya bu cisim diğer türdeşlerine göre doğru hareket ediyor mu diye sorulur. Eğer sınıflandırıcı objenin hareketlerini sakıncalı buluyorsa bu durum anomali kategorisine alınıp özetleme için depolanır. Örnek verecek olursak, bir yol üzerinde tek yönde giden arabalar olsun. Tek yönde giden arabaların bilgileri sürekli sınıflandırıcıya öğretilmiş oluyor. Eğer bir araç ters yönden girdiyse sınıflandırıcı şu zamana kadar öğrenmiş olduğu araba davranışlarında çok farklı bir hareket olduğunu tespit edip anomali tespitinde bulunmuş olur.

Tez çalışmamızda sınıflandırıcı seçimi yapılırken en önemli olmazsa olmaz kriter parçalı öğrenme olmuştur. Parçalı öğrenme, mevcut çalışan bir sınıflandırma sistemine yeni öğretim verisi geldiğinde bu veriyi kabul etmesi ve sonraki işlemlerinde bu verileri dikkate alarak sınıflandırma yapmasıdır. Her sınıflandırma algoritmasına uygun olmayan bir özellik olduğu için sınırlı sınıflandırma algoritması kümemiz kalmıştır.

Sınıflandırma için YOLO aracından gelen verilerin dönüştürülmesi gerekir. Bu dönüşüm için bir yapı düşünüldü. Öncelikle tespit edilen cismin konumu önemli olduğu kadar geçmiş verileri de önemliydi. Bu sebepten YOLO aracından alınan verileri obje kimliği ile depolayarak geçmişe dayalı bir eğitim verisi oluşturulmaya çalışıldı. Örneğin bir insanın veya bir arabanın yolun ortasında uzun bir süre bekliyor olması normal bir durum değildir. Bu sebepten dolayı sınıflandırma birimine eğitim için gönderilecek veriler aşağıda belirtildiği gibi hazırlandı:

1. Ekranın yatay eksenine göre cismin merkezi (X). Piksel olarak sonuç dönülmektedir.

2. Ekranın dikey eksenine göre cismin merkezi (Y). Piksel olarak sonuç dönülmektedir.
3. Cismin mevcut hızı (Bir önceki video çerçevesi merkezi ile şimdiki merkezi arasındaki uzaklık)

$$\sqrt{(X[n] - X[n-1])^2 + (Y[n] - Y[n-1])^2} \quad (4.1)$$

4. Cismin baştan sona toplam yer değiştirmesi

$$\sqrt{(X[n] - X[0])^2 + (Y[n] - Y[0])^2} \quad (4.2)$$

5. Cismin en baştan beri ortalama hızı

$$\frac{(\sqrt{(X[1] - X[0])^2 + (Y[1] - Y[0])^2} + \dots + \sqrt{(X[n] - X[n-1])^2 + (Y[n] - Y[n-1])^2})}{n} \quad (4.3)$$

4.1, 4.2 ve 4.3 formüllerinde n değişkeni tespit edilen cismin takip edildiği video kesit miktarını belirtir. X[0] takip edilen cismin ilk anındaki yatay konumu, X[n] cismin son anındaki yatay konumu, Y[0] cismin ilk anındaki dikey konumu Y[n] cismin son anındaki dikey konumunu belirtir.

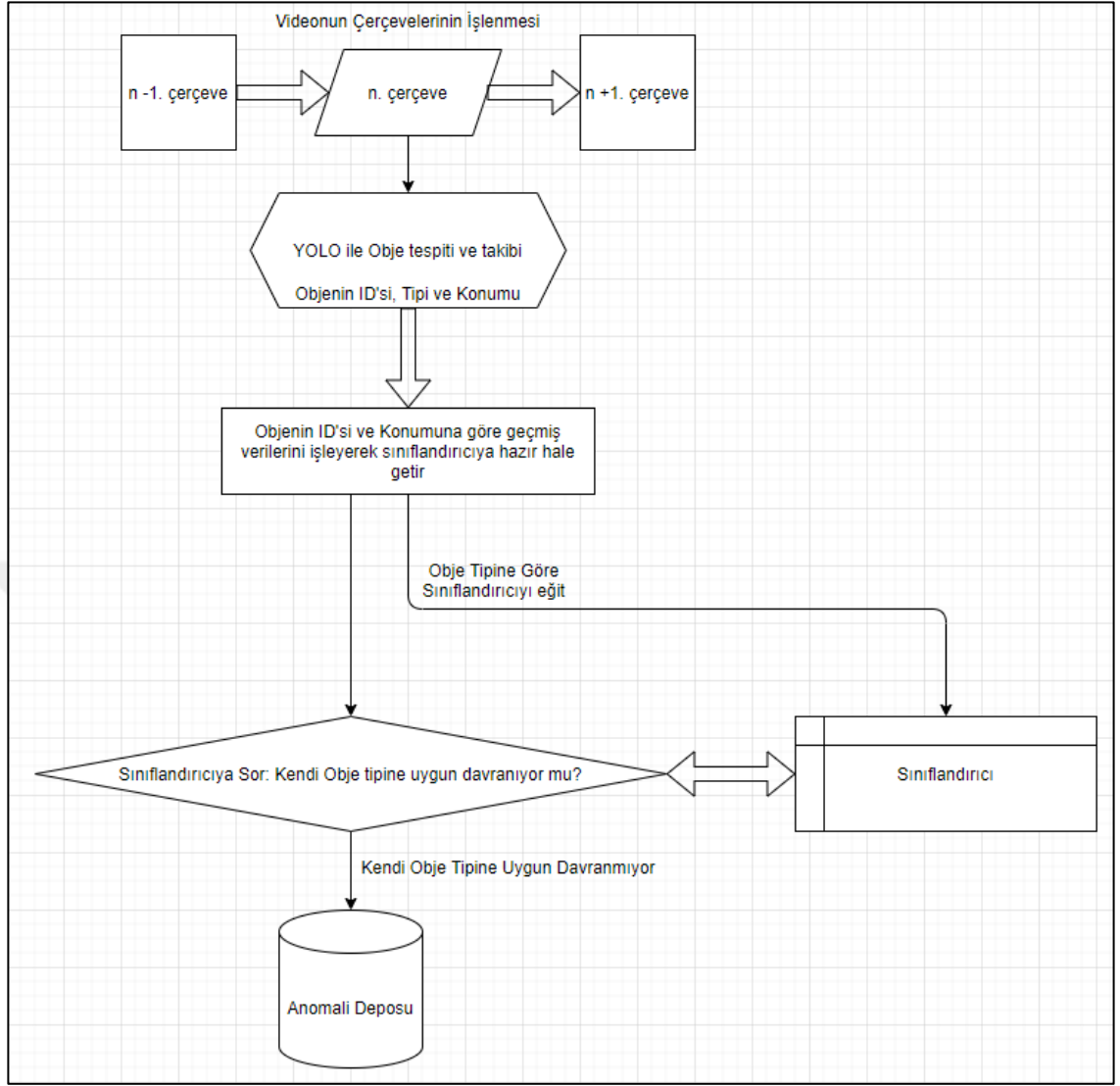
Yukarıda gördüğünüz üzere sınıflandırıcıya göndermekte olan verilerin geçmiş hareketleri ile bir bağ kurmasını sağlanarak uzun süreli dikkat çekici durumların tespiti sağlanmaya çalışılmıştır.

Eğitim verileri her obje tespitinde tek tek sınıflandırıcıya gönderilmemektedir. Çünkü bu durum sınıflandırıcının performansını kötü etkileyebilir. Bu sebepten 100 civarı obje tanımı olmadan sınıflandırıcının öğrenme metotları çağrılmamaktadır.

Tez çalışmamızda anlatmış olduğumuz üzere sistem kendi kendine öğrenme işlemi yapmaktadır. Her obje tipine göre genel hareketleri öğrenilecek genel olmayan hareketler

tespit edilecektir. Fakat burada gözden kaçırılmayacak bir husus vardır. Sınıflandırıcı öğrenim olarak belirli bir olgunluğa ulaşmadan yaptığı tahminler doğru olamayacaktır. Bu sebepten dolayı sınıflandırıcıya sorulma sınırları oluşturuldu. Deneylerimizde genel olarak bir obje tipinde 200 adet öğrenme yapılmadan sınıflandırıcının sonuçlarını dikkate alınmadı.

Günümüzde yapay zekâ ile uğraşılan birçok projenin temel sorunlarından biri sınırların en optimum düzeyde belirlenmesidir. Bizim çalışmamızda da optimumu araştırılması gereken sınırlar mevcuttur. En önemlisi sınıflandırıcının anomali olup olmadığına dair vermiş olduğu değerin sınırıdır. Bir başka sınır YOLO aracının tespit etmiş olduğu nesnelerin kendi belirtmiş doğruluk oranlarının kabul sınırıdır. Yani YOLO aracı bir resimde bir çöp bidonunu 0.1 ihtimalle (0 ile 1 arasında bir değer) insan sanabilir. Tez çalışmamızda bu ihtimaller için bir sınır değeri belirlenmekte. Bir diğer sınır, sınıflandırıcıya kaç tane eğitim verisi verdikten sonra cevapların dikkate alındığı ile ilgilidir. Deneylerimizde bu sınır 400 olarak kullanıldı. Yani 400 eğitim verisinden sonra sınıflandırma algoritması ile anomali tespiti yapılmaya başlandı. Son sınır değeri sınıflandırma algoritmasına eğitim için gönderilmesi gereken verilerin sayısını ne kadar olması gerektiğidir. Deneylerimizde çoğunlukla 100 olarak bu sınır değeri kullanıldı.



Şekil 4.5 YOLO ve sınıflandırıcının arasındaki ilişkiyi gösteren şekil

4.3.3 Anomalilerin işlenmesi

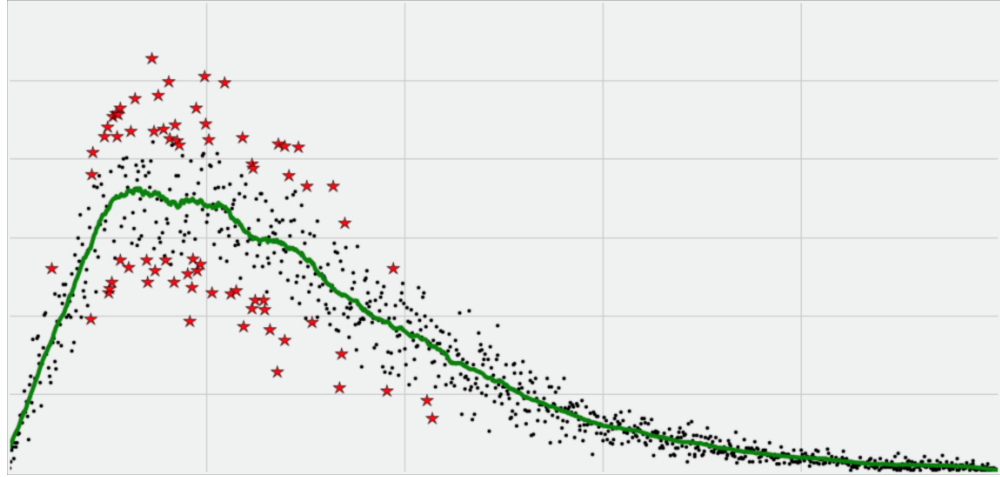
Anomalilerin işlenmesi, deneysel çalışmalar sonrasında yapmış olduğumuz gözlemler neticesinde olgunlaştı. Çünkü anomalilerin içerisinde görüntü işleme tabiri olan “gürültü” denebilecek anomaliler olduğu gözlemlendi. Mesela bir araba 10 saniye boyunca ekranda, gayet normal seyrinde bulunuyor. Bu seyir sırasında 1 tane veya 2 tane anomali tespiti yapılıyor fakat bu anomalileri destekleyen veya takip eden aynı obje ile ilgili başka anomali yok. Bu durumlarda bu anomalilerin özeti oluşturulmamalıdır. Çünkü yarım saniyelik video özet parçaları oluşabilmekte. Bu istenmeyen durumdan kaçınmak için video anomali incelemesi tamamlandıktan sonra tüm tespit edilen anomaliler obje kimliği

bazlı gruplandırıldı. Bu gruplandırma içerisinde eğer belirli bir miktarın altında anomali tespiti yapılmış obje varsa bu anomaliler listeden kaldırıldı. Örneğin 5 tane anomaliden daha az anomali olan objeler listeden kaldırıldı. Kalan objeler arasında her bir objenin anomalilerini tek tek gezerek birbirine uzak olan anomali varsa bu anomaliler silindi. Mesela bir objenin 15 tane anomali var ve 13 tanesi 1 saniyeden daha kısa aralıklarla peş peşe tespit edilmiş durumda. Fakat 2 tanesi diğerlerinden uzak zamanlarda tespit edilmiş. Bu durumlarda bahsettiğimiz 2 tanesi anomali listesinden silinir. Tüm anomalilerin obje bazlı birbirine yakın ve tekrarlı olması sağlanmış olunur.

Tespit edilmiş olan anomalilerin gürültülü kısımları temizlendikten sonra bu anomalilerin video parçaları haline getirilmesi gerekmektedir. Çünkü her bir anomali için sadece bir video çerçevesi çıkarıp bu çerçeveler birleştirilirse, videonun temel yapısını bozmuş olur ve anlaşılamaz hale gelebilir. Tespit edilmiş ve ayıklanmış anomalilerin video parçalarının belirlenmesi için basit bir algoritma geliştirildi.

1. Objeye tipi bazlı gruplanmış anomaliler tek bir listede toplanacak
2. Oluşturulmuş olan listenin tespit saniyesine göre yeniden sıralanacak
3. Bu listedeki elemanlar tek tek bakılacak
4. Bir önceki eleman ile mevcut eleman arasında 1 saniyeden fazla zaman varsa başlangıç saniyesi ile önceki elemanın saniyesi arasında bir video kesiti belirlenmiş olacak. Mevcut eleman bir sonraki kesitin başlangıç saniyesi olacak.

Algoritmaya karışık olmamasın adına yazılmayan bir durum da şudur. Eğer bir video kesiti içerisinde 5 ten daha az anomali varsa bu kesitin iptal edilmesidir. Çünkü bu kesitte gürültü sayabileceğimiz kadar anomali tespiti yapılmış demektir.



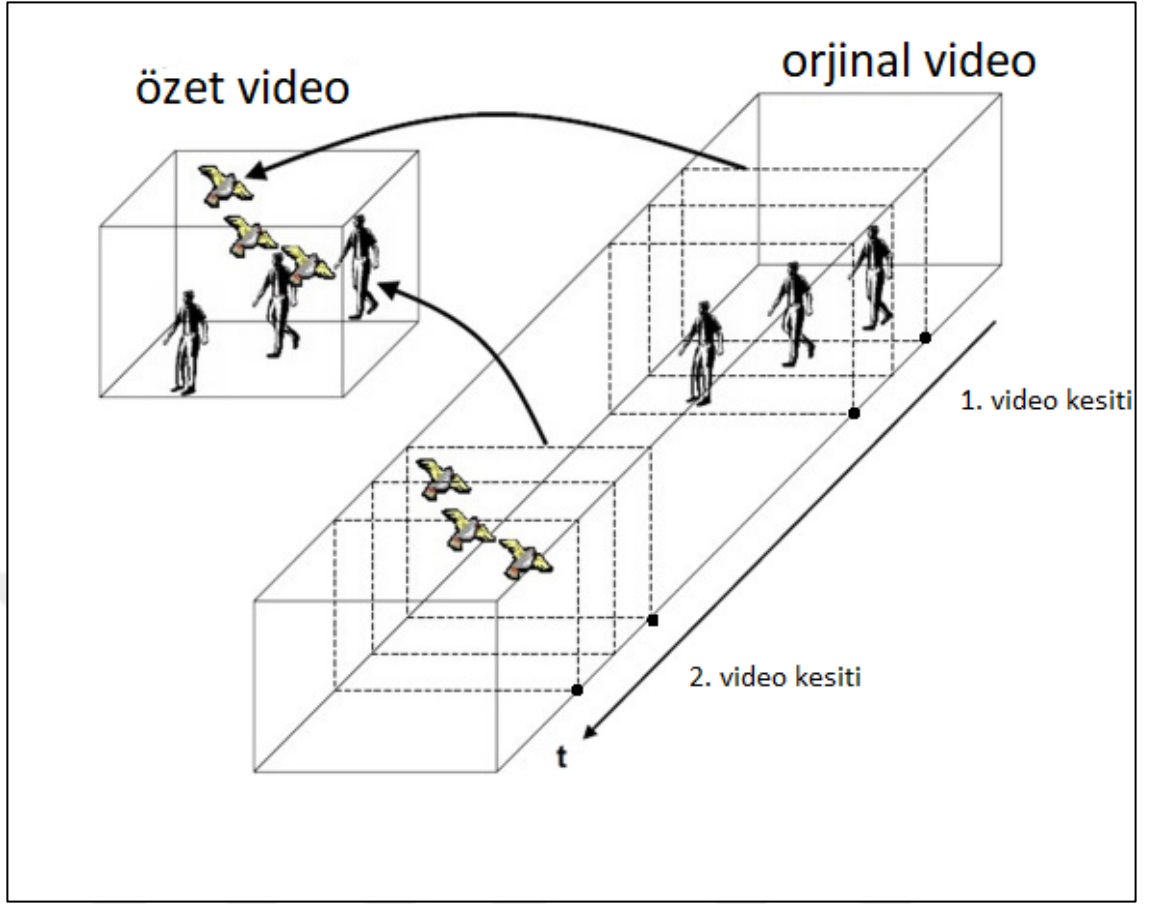
Şekil 4.6 Temsili anomali görseli (Xiufeng Liu vd. 2016)

4.4 Video Özeti Oluşturulması

Video özeti oluştururken öncelikle anomali çerçeveleri ve video saniyelerini göstermek için bir orijinal videonun aynısı üretilir. Özet çıkartma işlemi bu tekrar işlenerek üretilmiş video üzerinden yapılmaktadır.

Video özeti çıkartma sırasında yapılan işlem oldukça basittir. Bir listenin içerisindeki her bir eleman videonun bir parçasını göstermektedir ve bu elemanlar işleme sırasına göre sıralı halde olmalıdır. Listenin içindeki elemanların içerisinde video parçasının başlangıç ve bitiş saniyeleri bulunmaktadır. Video çıkartma özelliğimiz bahsettiğimiz listeyi alıp tek tek elemanları okuyup her bir eleman için belirtilmiş olduğu saniyeler arasındaki video parçasını kesip sonuç videosu olarak istediğimiz dosya adresine kaydedilmektedir. Kaydedilmesi gereken dosya dizini ve adı yapılandırma dosyalarında özel olarak belirtilmiştir.

Video özeti çıkartılması sırasında karşılaşılabilecek bir sorun da şudur. Video kesitleri arasındaki farklar kısaldığı zaman örneğin 2 saniye aralıklarda 3-4 video kesiti birleştirildiği zaman rahatsız edici bir görüntü oluşmaktadır. Bu tip rahatsızlıkların yaşanmaması için bir çözüm geliştirildi. Örneğin 3 saniyeden daha kısa yakınlıkta olan video kesitlerinin birleştirilmesi sağlandı. Bu geliştirme ile daha akıcı bir video özet deneyimi sunmaya başlanmış oldu.



Şekil 4.7 Özetleme işlemi ile ilgili temsili bir görsel (Rav-Acha vd. 2006).

4.5 İkili Kamera ile Video Özeti Oluşturulması

Tezimiz için yapmış olduğumuz çalışmamızda ikili kamera ile video özeti oluşturma yaparken karşılaştığımız en önemli sorun, iki kameranın ayrı ayrı tespit etmiş olduğu video kesitleri nasıl olup da tek bir video haline getirilmelidir. Bu sorun üzerinde şu çözüm kabul edildi. İki kameranın da aynı anda tespit ettiği durumlar her iki özete de konulacak, eğer sadece bir kamera bir anomali tespit ediyorsa bu geçersiz sayılacaktır.

Bahsettiğimiz çözümün koda dökülmesi ise şu şekilde oldu. İki kameranın iki videosu seçilerek uygulamamıza verildi. İki videonun anomalileri, video kesitleri çıkartıldı. İki videoya ait kesitlerin kesişim kümeleri alınarak video özetleme işleme yapıldı.

İkili kamerada video özeti oluşturma bizim kullandığımız yöntem için bir avantajı bulunmaktadır. Eğer gürültü kapsamına giren alakasız anomaliler ile iki kamerada da

karşılaşılma ihtimali oldukça düşecek ve daha stabil sonuçlar doğması sağlanmış olacaktır.

İkili kamera ile video özetlenmesinin bir problemi daha vardır. Bu problem; eğer bir video diğer video ile aynı anda başlamamış ise nasıl bu iki video senkronize edilir. Bu sorunun çözümü olarak çalışmamızda girdi olarak verilen videoların yanına video zaman farklılığı girdi bölümü konuldu. Videolar arası zaman farklılıkları belirtilerek diğer işlemlerin etkilenmemesi sağlandı.



5. DENEYLER

Tez çalışmamızın deneylerinde tek kamera ile 20 adet çift kamera 2 adet deney yapıldı. Tek kamera deneylerinde 5 farklı video ile çalışmamızda kullanılan değişkenler denenmiş oldu. Çift kamera deneylerinde değişkenler sabit tutularak sonuçlar üretildi.

Çift kamera deneylerimizde, aynı anlarda çekim yapmış aynı görüş alanına sahip sabit kameralar olması gerekmektedir. Bu koşullara sahip video verileri bulunamadığı için benzer koşullar üretilmeye çalışıldı. Tek bir videonun farklı fakat kesişen kesitleri çıkartıldı. Bu şekilde aynı yöne bakan yan yana duran kamera görüntüsü elde edilmek istendi. Aynı video hem tekli kamera deneylerinde hem de çiftli kamera deneylerinde kullanılarak aslında kontrollü deney imkânı sundu.

Tüm deneylerimizde sınıflandırma algoritması olarak SGD (Stokastik Gradyan İnişi) ve YOLO aracında “yolov3-tiny” çalışması kullanıldı.

5.1 1. Kısa Video Deneyi

Bu deneyimizde kullandığımız video Şekil 5-1’de gördüğünüz gibi bir şehrin 4 yol kavşağında insanların ve arabaların kameranin görüş alanında olduğu 20 dakikalık bir videodur.



Şekil 5.1 1. Kısa Video. Deneylerimizde kullandığımız videodan bir görsel

5.1.1 Senaryo 1

Video özeti sonucunda 3 dakika 29 saniyelik bir video özeti oluşmuştur. Toplamda 34 adet kesit bulunmaktadır. Ortalama kesit aralıkları 6,1 saniye olarak görülmektedir. Oluşturulmuş video özeti kısaltma oranı yaklaşık %17,4 olmuştur.

Video YOLO obje tespit sınırı olarak 0.5 kullanıldı.

Video Sınıflandırma algoritması anomali tespit sınırı olarak 0.5 kullanıldı.

Videoda bir objenin en az anomali sayısı sınırı 5 adet.

Videoda bir objenin en az sıralı anomali sayısı sınırı 3 adet.

5.1.2 Senaryo 2

Video özeti sonucunda 4 dakika 54 saniyelik bir video özeti oluşmuştur. Toplamda 46 adet kesit bulunmaktadır. Ortalama kesit aralıkları 6,3 saniye olarak görülmektedir. Oluşturulmuş video özeti kısaltma oranı yaklaşık %24,5 olmuştur.

Video YOLO obje tespit sınırı olarak 0.3 kullanıldı.

Video Sınıflandırma algoritması anomali tespit sınırı olarak 0.5 kullanıldı.

Videoda bir objenin en az anomali sayısı sınırı 5 adet.

Videoda bir objenin en az sıralı anomali sayısı sınırı 3 adet.

5.1.3 Senaryo 3

Video özeti sonucunda 3 dakika 29 saniyelik bir video özeti oluşmuştur. Toplamda 34 adet kesit bulunmaktadır. Ortalama kesit aralıkları 6,1 saniye olarak görülmektedir. Oluşturulmuş video özeti kısaltma oranı yaklaşık %17,4 olmuştur.

Video YOLO obje tespit sınırı olarak 0.5 kullanıldı.

Video Sınıflandırma algoritması anomali tespit sınırı olarak 0.8 kullanıldı.

Videoda bir objenin en az anomali sayısı sınırı 5 adet.

Videoda bir objenin en az sıralı anomali sayısı sınırı 3 adet.

5.1.4 Senaryo 4

Video özeti sonucunda 2 dakika 53 saniyelik bir video özeti oluşmuştur. Toplamda 21 adet kesit bulunmaktadır. Ortalama kesit aralıkları 5 saniye olarak görülmektedir. Oluşturulmuş video özeti kısaltma oranı yaklaşık %14,4 olmuştur.

Video YOLO obje tespit sınırı olarak 0.5 kullanıldı.

Video Sınıflandırma algoritması anomali tespit sınırı olarak 0.5 kullanıldı.

Videoda bir objenin en az anomali sayısı sınırı 10 adet.

Videoda bir objenin en az sıralı anomali sayısı sınırı 6 adet.

5.2 2. Kısa Video Deneyi

Bu deneyimizde kullandığımız video Şekil 5-2’de gördüğünüz gibi bir şehir meydanının videosudur. Daha çok insanların kamera görüş açısında olduğu ve arabaların uzaktan kameranın görüş alanına girdiği 22 dakikalık bir videodur.



Şekil 5.2 2. Kısa Video. Deneylerimizde kullandığımız videodan bir görsel

5.2.1 Senaryo 1

Video özeti sonucunda 3 dakika 13 saniyelik bir video özeti oluşmuştur. Toplamda 61 adet kesit bulunmaktadır. Ortalama kesit aralıkları 3,1 saniye olarak görülmektedir. Oluşturulmuş video özeti kısaltma oranı yaklaşık %14,8 olmuştur.

Video YOLO obje tespit sınırı olarak 0.5 kullanıldı.

Video Sınıflandırma algoritması anomali tespit sınırı olarak 0.5 kullanıldı.

Videoda bir objenin en az anomali sayısı sınırı 5 adet.

Videoda bir objenin en az sıralı anomali sayısı sınırı 3 adet.

5.2.2 Senaryo 2

Video özeti sonucunda 8 dakika 25 saniyelik bir video özeti oluşmuştur. Toplamda 75 adet kesit bulunmaktadır. Ortalama kesit aralıkları 6,7 saniye olarak görülmektedir. Oluşturulmuş video özeti kısaltma oranı yaklaşık %38,8 olmuştur.

Video YOLO obje tespit sınırı olarak 0.3 kullanıldı.

Video Sınıflandırma algoritması anomali tespit sınırı olarak 0.5 kullanıldı.

Videoda bir objenin en az anomali sayısı sınırı 5 adet.

Videoda bir objenin en az sıralı anomali sayısı sınırı 3 adet.

5.2.3 Senaryo 3

Video özeti sonucunda 3 dakika 29 saniyelik bir video özeti oluşmuştur. Toplamda 59 adet kesit bulunmaktadır. Ortalama kesit aralıkları 3,3 saniye olarak görülmektedir. Oluşturulmuş video özeti kısaltma oranı yaklaşık %15,0 olmuştur.

Video YOLO obje tespit sınırı olarak 0.5 kullanıldı.

Video Sınıflandırma algoritması anomali tespit sınırı olarak 0.8 kullanıldı.

Videoda bir objenin en az anomali sayısı sınırı 5 adet.

Videoda bir objenin en az sıralı anomali sayısı sınırı 3 adet.

5.2.4 Senaryo 4

Video özeti sonucunda 2 dakika 12 saniyelik bir video özeti oluşmuştur. Toplamda 46 adet kesit bulunmaktadır. Ortalama kesit aralıkları 2,8 saniye olarak görülmektedir. Oluşturulmuş video özeti kısaltma oranı yaklaşık %10,1 olmuştur.

Video YOLO obje tespit sınırı olarak 0.5 kullanıldı.

Video Sınıflandırma algoritması anomali tespit sınırı olarak 0.5 kullanıldı.

Videoda bir objenin en az anomali sayısı sınırı 10 adet.

Videoda bir objenin en az sıralı anomali sayısı sınırı 6 adet.

5.3 3. Kısa Video Deneyi

Bu deneyimizde kullandığımız video Şekil 5-3'te gördüğünüz gibi bir şehir meydanının videosudur. Daha çok insanların kamera görüş açısında olduğu ve arabaların uzaktan kameranın görüş alanına girdiği 23 dakikalık bir videodur.



Şekil 5.3 3. Kısa Video. Deneylerimizde kullandığımız videodan bir görsel

5.3.1 Senaryo 1

Video özeti sonucunda 1 dakika 5 saniyelik bir video özeti oluşmuştur. Toplamda 32 adet kesit bulunmaktadır. Ortalama kesit aralıkları 2 saniye olarak görülmektedir. Oluşturulmuş video özeti kısaltma oranı yaklaşık %4,8 olmuştur.

Video YOLO obje tespit sınırı olarak 0.5 kullanıldı.

Video Sınıflandırma algoritması anomali tespit sınırı olarak 0.5 kullanıldı.

Videoda bir objenin en az anomali sayısı sınırı 5 adet.

Videoda bir objenin en az sıralı anomali sayısı sınırı 3 adet.

5.3.2 Senaryo 2

Video özeti sonucunda 2 dakika 4 saniyelik bir video özeti oluşmuştur. Toplamda 49 adet kesit bulunmaktadır. Ortalama kesit aralıkları 2,5 saniye olarak görülmektedir. Oluşturulmuş video özeti kısaltma oranı yaklaşık %9,1 olmuştur.

Video YOLO obje tespit sınırı olarak 0.3 kullanıldı.

Video Sınıflandırma algoritması anomali tespit sınırı olarak 0.5 kullanıldı.

Videoda bir objenin en az anomali sayısı sınırı 5 adet.

Videoda bir objenin en az sıralı anomali sayısı sınırı 3 adet.

5.3.3 Senaryo 3

Video özeti sonucunda 25 saniyelik bir video özeti oluşmuştur. Toplamda 12 adet kesit bulunmaktadır. Ortalama kesit aralıkları 2 saniye olarak görülmektedir. Oluşturulmuş video özeti kısaltma oranı yaklaşık %1,8 olmuştur.

Video YOLO obje tespit sınırı olarak 0.5 kullanıldı.

Video Sınıflandırma algoritması anomali tespit sınırı olarak 0.8 kullanıldı.

Videoda bir objenin en az anomali sayısı sınırı 5 adet.

Videoda bir objenin en az sıralı anomali sayısı sınırı 3 adet.

5.3.4 Senaryo 4

Video özeti sonucunda 31 saniyelik bir video özeti oluşmuştur. Toplamda 16 adet kesit bulunmaktadır. Ortalama kesit aralıkları 2 saniye olarak görülmektedir. Oluşturulmuş video özeti kısaltma oranı yaklaşık %2,2 olmuştur.

Video YOLO obje tespit sınırı olarak 0.5 kullanıldı.

Video Sınıflandırma algoritması anomali tespit sınırı olarak 0.5 kullanıldı.

Videoda bir objenin en az anomali sayısı sınırı 10 adet.

Videoda bir objenin en az sıralı anomali sayısı sınırı 6 adet.

5.4 1. Uzun Video Deneyi

Bu deneyimizde kullandığımız video Şekil 5-4'te gördüğünüz gibi bir şehrin 4 yol kavşağında insanların ve arabaların kameranın görüş alanında olduğu 60 dakikalık bir videodur.



Şekil 5.4 1. Uzun Video. Deneylerimizde kullandığımız videodan bir görsel

5.4.1 Senaryo 1

Video özeti sonucunda 32 dakika 31 saniyelik bir video özeti oluşmuştur. Toplamda 107 adet kesit bulunmaktadır. Ortalama kesit aralıkları 18,2 saniye olarak görülmektedir. Oluşturulmuş video özeti kısaltma oranı yaklaşık %51,4 olmuştur.

Video YOLO obje tespit sınırı olarak 0.5 kullanıldı.

Video Sınıflandırma algoritması anomali tespit sınırı olarak 0.5 kullanıldı.

Videoda bir objenin en az anomali sayısı sınırı 5 adet.

Videoda bir objenin en az sıralı anomali sayısı sınırı 3 adet.

5.4.2 Senaryo 2

Video özeti sonucunda 45 dakika 13 saniyelik bir video özeti oluşmuştur. Toplamda 46 adet kesit bulunmaktadır. Ortalama kesit aralıkları 34,6 saniye olarak görülmektedir. Oluşturulmuş video özeti kısaltma oranı yaklaşık %75,5 olmuştur.

Video YOLO obje tespit sınırı olarak 0.3 kullanıldı.

Video Sınıflandırma algoritması anomali tespit sınırı olarak 0.5 kullanıldı.

Videoda bir objenin en az anomali sayısı sınırı 5 adet.

Videoda bir objenin en az sıralı anomali sayısı sınırı 3 adet.

5.4.3 Senaryo 3

Video özeti sonucunda 29 dakika 9 saniyelik bir video özeti oluşmuştur. Toplamda 34 adet kesit bulunmaktadır. Ortalama kesit aralıkları 15,9 saniye olarak görülmektedir. Oluşturulmuş video özeti kısaltma oranı yaklaşık %48,3 olmuştur.

Video YOLO obje tespit sınırı olarak 0.5 kullanıldı.

Video Sınıflandırma algoritması anomali tespit sınırı olarak 0.8 kullanıldı.

Videoda bir objenin en az anomali sayısı sınırı 5 adet.

Videoda bir objenin en az sıralı anomali sayısı sınırı 3 adet.

5.4.4 Senaryo 4

Video özeti sonucunda 26 dakika 23 saniyelik bir video özeti oluşmuştur. Toplamda 21 adet kesit bulunmaktadır. Ortalama kesit aralıkları 14,5 saniye olarak görülmektedir. Oluşturulmuş video özeti kısaltma oranı yaklaşık %43,4 olmuştur.

Video YOLO obje tespit sınırı olarak 0.5 kullanıldı.

Video Sınıflandırma algoritması anomali tespit sınırı olarak 0.5 kullanıldı.

Videoda bir objenin en az anomali sayısı sınırı 10 adet.

Videoda bir objenin en az sıralı anomali sayısı sınırı 6 adet.

5.5 2. Uzun Video Deneyi

Bu deneyimizde kullandığımız video Şekil 5-5'te gördüğünüz gibi bir şehrin 4 yol kavşağında çoğunlukla arabaların kameranin görüş alanında olduğu 65 dakikalık bir videodur.



Şekil 5.5 2. Uzun Video. Deneylerimizde kullandığımız videodan bir görsel

5.5.1 Senaryo 1

Video özeti sonucunda 11 dakika 52 saniyelik bir video özeti oluşmuştur. Toplamda 105 adet kesit bulunmaktadır. Ortalama kesit aralıkları 6,7 saniye olarak görülmektedir. Oluşturulmuş video özeti kısaltma oranı yaklaşık %18,4 olmuştur.

Video YOLO obje tespit sınırı olarak 0.5 kullanıldı.

Video Sınıflandırma algoritması anomali tespit sınırı olarak 0.5 kullanıldı.

Videoda bir objenin en az anomali sayısı sınırı 5 adet.

Videoda bir objenin en az sıralı anomali sayısı sınırı 3 adet.

5.5.2 Senaryo 2

Video özeti sonucunda 23 dakika 14 saniyelik bir video özeti oluşmuştur. Toplamda 149 adet kesit bulunmaktadır. Ortalama kesit aralıkları 9,3 saniye olarak görülmektedir. Oluşturulmuş video özeti kısaltma oranı yaklaşık %35,3 olmuştur.

Video YOLO obje tespit sınırı olarak 0.3 kullanıldı.

Video Sınıflandırma algoritması anomali tespit sınırı olarak 0.5 kullanıldı.

Videoda bir objenin en az anomali sayısı sınırı 5 adet.

Videoda bir objenin en az sıralı anomali sayısı sınırı 3 adet.

5.5.3 Senaryo 3

Video özeti sonucunda 11 dakika 40 saniyelik bir video özeti oluşmuştur. Toplamda 101 adet kesit bulunmaktadır. Ortalama kesit aralıkları 6,9 saniye olarak görülmektedir. Oluşturulmuş video özeti kısaltma oranı yaklaşık %16,9 olmuştur.

Video YOLO obje tespit sınırı olarak 0.5 kullanıldı.

Video Sınıflandırma algoritması anomali tespit sınırı olarak 0.8 kullanıldı.

Videoda bir objenin en az anomali sayısı sınırı 5 adet.

Videoda bir objenin en az sıralı anomali sayısı sınırı 3 adet.

5.5.4 Senaryo 4

Video özeti sonucunda 11 dakika 1 saniyelik bir video özeti oluşmuştur. Toplamda 99 adet kesit bulunmaktadır. Ortalama kesit aralıkları 6,6 saniye olarak görülmektedir. Oluşturulmuş video özeti kısaltma oranı yaklaşık %16,9 olmuştur.

Video YOLO obje tespit sınırı olarak 0.5 kullanıldı.

Video Sınıflandırma algoritması anomali tespit sınırı olarak 0.5 kullanıldı.

Videoda bir objenin en az anomali sayısı sınırı 10 adet.

Videoda bir objenin en az sıralı anomali sayısı sınırı 6 adet.

5.6 İkili Kamera Kısa Video Deneyi

5.1 1. Kısa Video Deneyi bölümünde yapmış olduğumuz deneylerimizin videosu işlenerek 2 video haline getirildi. Bu videolar yan yana duran kamera olduğu varsayılarak deney gerçekleştirildi.

Video özeti sonucunda 1 dakika 11 saniyelik bir video özeti oluşmuştur. Toplamda 27 adet kesit bulunmaktadır. Ortalama kesit aralıkları 2,6 saniye olarak görülmektedir. Oluşturulmuş video özeti kısaltma oranı yaklaşık %5,5 olmuştur.

Video YOLO obje tespit sınırı olarak 0.5 kullanıldı.

Video Sınıflandırma algoritması anomali tespit sınırı olarak 0.5 kullanıldı.

Videoda bir objenin en az anomali sayısı sınırı 5 adet.

Videoda bir objenin en az sıralı anomali sayısı sınırı 3 adet.

5.7 İkili Kamera Uzun Video Deneyi

5.5 2. Uzun Video Deneyi bölümünde yapmış olduğumuz deneylerimizin videosu işlenerek 2 video haline getirildi. Bu videolar yan yana duran kamera olduğu varsayılarak deney gerçekleştirildi.

Video özeti sonucunda 10 dakika 32 saniyelik bir video özeti oluşmuştur. Toplamda 149 adet kesit bulunmaktadır. Ortalama kesit aralıkları 4,2 saniye olarak görülmektedir. Oluşturulmuş video özeti kısaltma oranı yaklaşık %16,1 olmuştur.

Video YOLO obje tespit sınırı olarak 0.5 kullanıldı.

Video Sınıflandırma algoritması anomali tespit sınırı olarak 0.5 kullanıldı.

Videoda bir objenin en az anomali sayısı sınırı 5 adet.

Videoda bir objenin en az sıralı anomali sayısı sınırı 3 adet.

5.8 Deney Karşılaştırmaları

Çizelge 5.1 Tek kamera deney sonuçları karşılaştırılması

Senaryo No	Deney Değerleri				Çıktı Değerleri			
	YOLO Sınırı	Sınıf. Sınırı	O. A. Sınırı	S. A. Sınırı	Özet Uzunluğu	Parça Adeti	Ort. Parça Uzunluğu	Özet Oranı
	0-1	0-1	Adet	Adet	dk.	Adet	sn.	%
5.1.1	0.5	0.5	5	3	3:29	34	6.1	17.4
5.1.2	0.3	0.5	5	3	4:54	46	6.3	24.5
5.1.3	0.5	0.8	5	3	3:29	34	6.1	17.4
5.1.4	0.5	0.5	10	6	2:53	21	5	14.4
5.2.1	0.5	0.5	5	3	3:13	61	3.1	14.8
5.2.2	0.3	0.5	5	3	8:25	75	6.7	35.8
5.2.3	0.5	0.8	5	3	3:29	59	3.3	15
5.2.4	0.5	0.5	10	6	2:12	46	2.8	10.1
5.3.1	0.5	0.5	5	3	1:05	32	2	4.8
5.3.2	0.3	0.5	5	3	2:04	49	2.5	9.1
5.3.3	0.5	0.8	5	3	0:25	12	2	1.8
5.3.4	0.5	0.5	10	6	0:31	16	2	2.2
5.4.1	0.5	0.5	5	3	32:31	107	15.2	51.4
5.4.2	0.3	0.5	5	3	45:13	46	34.6	75.5
5.4.3	0.5	0.8	5	3	29:09	34	15.9	48.3
5.4.4	0.5	0.5	10	6	26:23	21	14.5	43.4
5.5.1	0.5	0.5	5	3	11:52	105	6.7	18.4
5.5.2	0.3	0.5	5	3	23:14	149	9.3	35.3
5.5.3	0.5	0.8	5	3	11:40	101	6.9	16.9
5.5.4	0.5	0.5	10	6	11:01	99	6.6	16.9

Çizelge 5.2 Tek-çift kamera deney sonuçları karşılaştırılması

Senaryo No	Deney Değerleri				Çıktı Değerleri			
	YOLO Sınırı	Sınıf. Sınırı	O. A. Sınırı	S. A. Sınırı	Özet Uzunluğu	Parça Adeti	Ort. Parça Uzunluğu	Özet Oranı
	0-1	0-1	Adet	Adet	dk.	Adet	sn.	%
5.1.1	0.5	0.5	5	3	3:29	34	6.1	17.4
5.6	0.5	0.5	5	3	1:11	27	2.6	5.5
5.5.1	0.5	0.5	5	3	11:40	101	6.9	16.9
5.7	0.5	0.5	5	3	10:32	149	4.2	16.1

Çizelge 5.1 ve Çizelge 5.2’de Deney Değerleri kısmına ait başlık açıklamaları aşağıdadır:

- YOLO Sınırı: Video YOLO obje tespit sınırı.
- Sınıf. Sınırı: Video Sınıflandırma algoritması anomali tespit sınırı.
- O. A. Sınırı: Videoda bir objenin olabilecek en az anomali sayısı sınırı.
- S. A. Sınırı: Videoda bir objenin olabilecek en az sıralı anomali sayısı sınırı.

Çizelge 5.1 üzerinde gösterdiğimiz tek kamera deneylerinde parametre karşılaştırmaları yapıldı. Her bir video (5.1, 5.2 vs.) için 4 farklı parametre değeri ile deneyler yapıldı. 1. ve 2. deneyler (Örneğin 5.1.1 – 5.1.2) arasında YOLO obje tespit sınırı karşılaştırılması yapıldı. 1. ve 3. deneyler (Örneğin 5.2.1 – 5.2.3) arasında sınıflandırıcının anomali tespit sınırı karşılaştırılması yapıldı. 1. ve 4. deneyler (Örneğin 5.4.1 – 5.4.4) arasında da anomali sayı sınırları değiştirilerek karşılaştırma yapıldı. YOLO obje tespit sınırı 0.5’ten 0.3 değerine düşürüldüğünde video özet oranları %50 ile %120 aralığında arttığı gözlemlendi. Sınıflandırıcı anomali tespit sınırı 0.5’ten 0.8 değerine yükseltildiğinde 5.1, 5.2, 5.4 ve 5.5 deneylerinde video özet oranı değişimi %0 ile %15 arasında iken 5.3 deneyinde video özeti oranı yarıdan daha aşağı düştüğü gözlemlendi. Anomali sayı sınırları 5 ve 3 değerinden 10 ve 6 değerine yükseltildiğinde video özeti oranları %10 ile %50 arasında kısaldığı gözlemlendi.

Deneylerimiz sonucunda karşılaştığımız çizelgelerde her bir video (5.1, 5.2, 5.3, 5.4, 5.5) için farklı video özet oranları görmektedir. Bu gösteriyor ki video içeriği, özetleme oranlarını fazlaca etkilemektedir. Tezimizin amacı denetimsiz öğrenme ile anomali tespiti yaparak video özetleme işlemidir. Denetimsiz öğrenme üzerine bir sonuç üretildiği için çıkan sonuçların video özet oranları, yorumlanabilecek ve kıyaslanabilecek en temel unsur olmaktadır. Üretilmiş deney sonuçlarında genellikle %5- %15 aralığında video özet oranı çıktısı olsa da 5.4 deney verisinde %40 üstünde deney oranı bulunmaktadır. Deneylerimiz sonucunda kullandığımız parametrelerin deney oranlarını hem tutarlı bir şekilde hem de fazlaca etkilediği (Örneğin 5.4.2 – 5.4.4 deneyleri arasında %32,1’lik video özet oranı farkı bulunmaktadır) gözlemlenmektedir. Bu sebepten her bir video özelinde parametre optimizasyonlarının gerekli olduğu anlaşılmaktadır.

İkili kamera deney sonuçlarında, aynı videonun tek kamera ile üretilmiş özet uzunluklarından daha kısa özet miktarları alınmıştır. Bu beklediğimiz bir çıktıdır. Çünkü iki kameranın tek bir alana bakıp aynı zamanda aynı anomalileri görmesi gürültü verilerini azaltacaktır.



6. SONUÇ VE GELECEK ÇALIŞMALAR

Günümüzde verinin ve veri işleminin önemi hız kesmeden artmaktadır. Büyük verilerin işlenmesi bu alanda yapılan çalışmalar günden güne artmakta ve önem kazanmaktadır. Videoların işlenmesi konusu, büyük veri işlenmesi alanında önemli bir yere sahiptir.

Tez kapsamında yapmış olduğumuz çalışma video özetleme alanına farklı bir bakış açısı getirmeyi hedeflemektedir. Bu kapsamda, Video işlenmesi sırasında denetimsiz olarak kendi kendine eğitebilen, nesne tabanlı öğrenme içeren, denetimsiz eğitim sonrasında sınıflandırma algoritması ile aykırılık saptaması kullanan bir yöntem geliştirilmiştir. Bu yöntem kullanılarak bir pikseller yığını olan videoyu işleyip video kesitleri tespit edilmekte, sonrasında video kesitleri orijinal videodan çıkartılarak video kesitler birleşimi olarak video özeti oluşturulmaktadır.

Tez kapsamında önerdiğimiz sistem sabit kamera ile çekilmiş görüntüler ile çalışılabilmektedir. Tezimizin ana amacı olan ikili kamera sisteminde video özetlenmesi gerçekte kısıtlı alanlara hitap etmektedir. Örnek olarak, ikili kamera sistemi aynı yola veya aynı alana bakan kameralar için kullanılabilir durumdadır.

Tekli kameralı deneylerimizde çalışmalarda kullandığımız sınırlamalar test edilip gözlemlendi. İkili kamera deneylerinde ise iki kameranın tekli kamera performans kıyası sağlandı. Bilindiği üzere denetimsiz öğrenim yapan bir sistemin sonuçlarının belirsizlik içermesi çok olasıdır. Bu çalışmanın ilgili deney sonuçları incelendiğinde de bu durum hâkim olduğu gözlemlenmiştir. Deneylerimizde dikkatimizi çeken en önemli nokta sonuçların tutarlı olmasıdır. Yapmış olduğumuz deneylerde verdiğimiz değişkenlerin etkisi bir tutarlılık göstermektedir. Benzer videolar üzerinde yapılan deneylerde, çıkarılan video özet oranı benzerlik göstermektedir. Ek olarak, benzer videolar üzerinde yapılan deneylerdeki farklı değişkenlerin video özet oranı sonucuna etkisi oldukça yakın olmaktadır. Bu tutarlılıklara ve benzerliklere bakarak yapmış olduğumuz yorum; eğer belirli bir oranda video özeti beklentiniz varsa o orana uygun şekilde çalışmamızın değişkenlerini değiştirip sonucu istediğiniz oranda yakınlaşabileceğidir.

Deneyler kapsamında bazı veriler üretildi, fakat tez çalışması kapsamında gerçek manada testlerini yapmak mümkün olmadı. Çünkü doğru bir deneysel kurgu için bir kameranın uzun süre kayıt yaptıktan sonra, kaza benzeri durumun oluşması gereklidir. Uzun süreli deneyler çalışma olgunlaşana kadar tekrarlanmalıdır. Bu seviyede deneysel çalışma bu tez çalışması için konulan kapsamın ötesinde olduğu için erişebildiğimiz girdiler üzerinden deneyler yapılmıştır.

Gelecek çalışması olarak yapılabilecek geliştirmeler ve iyileştirmeler mevcuttur. Önerilen sistemin en çok karar verme inisiyatifi olan birimi sınıflandırıcıdır. Sınıflandırıcı biriminde geliştirme çalışmaları yapılması sistemin verimliliğini arttırabilecektir. Sınıflandırmanın önerilen çalışmada iki tane işlevi vardır. Bu işlevler nesnelerin hareketlerini öğrenmek ve anomali içeren nesneleri tespit etmektir. Nesnelerin hareketlerini öğrenme işlevi üzerine iyileştirmeler yapılabilir. Örneğin, nesne hareket formüllerinin geliştirilmesi sağlanabilir veya öğrenme verimi optimize edilebilir. Sınıflandırıcının aykırılık tespit verimi, nesne hareketlerinin sorgulama formülleri değiştirilerek arttırılabilir. Sınıflandırma üzerinde köklü değişiklikler denenebilir. Mesela farklı sınıflandırma algoritmalarının verileri kıyaslanabilir. Sınıflama algoritması yerine eğitilmiş öğrenme yöntemleri kullanılabilir (Örneğin yapay sinir ağı vb.). Bu çalışma kapsamında önerilen sisteme verimlilik kaybettiren bir diğer unsur, YOLO aracının aynı nesneyi takip ederken nesneyi bir anlık takibini sürdüremeyerek nesnenin geçmiş hareket verisinin kaybolmasına sebebiyet vermesidir. Bu kapsamda YOLO'nun takip kalitesi alanında çalışmalar ve iyileştirmeler yapılabilir.

Tez çalışması kapsamında geliştirilen yöntemler, video özeti dışında başka alanlar için de kullanılabilir. Örneğin alarm sistemlerinin kurgulanması içinde oldukça uygun sistemlerdir.

Video özetleme, bilim insanları tarafında üzerine son yıllarda yoğun çalışmalar yapılan bir alanıdır. Bu tez çalışması kapsamında nesne tabanlı denetimsiz eğitim modeli üzerinden ikili kamera sistemi ile yeni bir video özet işleme yaklaşımı geliştirilerek, ilgili literatüre katkı sağlanmıştır.

KAYNAKLAR

- A. A. Yilmaz, M. S. Guzel, E. Bostanci, and I. Askerzade, "A novel action recognition framework based on deep-learning and genetic algorithms," *IEEE Access*, vol. 8, pp. 100631–100644, 2020.
- A. Rav-Acha, Y. Pritch, and S. Peleg, "Making a long video short: Dynamic video synopsis," in *Proc. IEEE Conf. Comp. Vis. Patt. Recogn.*, New York, NY, June 2006, pp. 435–441.
- Athanasios S. Voulodimos, Anastasios D. Doulamis, Dimitrios I. Kosmopoulos, and Theodora A. Varvarigou. 2011. Video summarization guiding evaluative rectification for industrial activity recognition. In *Proceedings of the IEEE International Conference on Computer Vision (ICCV'11)*. 950–957.
- Athanasios S. Voulodimos, Anastasios D. Doulamis, Dimitrios I. Kosmopoulos, and Theodora A. Varvarigou. 2011. Video summarization guiding evaluative rectification for industrial activity recognition. In *Proceedings of the IEEE International Conference on Computer Vision (ICCV'11)*. 950–957.
- B. Mutlu, M. Hacıömeroğlu, G. M. Serdar, M. Dikmen, and H. Sever, "Silhouette extraction from street view images," *Int. J. Adv. Robot. Syst.*, vol. 11, no. 7, p. 114, 2014.
- C. R. Huang, H. C. Chen, and P. C. Chung, "Online surveillance video synopsis," in *Proc. IEEE Int. Symp. on Circuits and Systems*, Seoul, South Korea, May 2012, pp. 1843–1846.
- C. R. Huang, P. C. J. Chung, D. K. Yang, H. C. Chen, and G. J. Huang, "Maximum a posteriori probability estimation for online surveillance video synopsis," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 24, no. 8, pp. 1417–1429, Aug 2014.
- Chia Ming Tsai, Li Wei Kang, Chia Wen Lin, and Weisi Lin. 2013. Scene-based movie summarization via rolecommunity networks. *IEEE Trans. Circuits Syst. Video Technol.* 23, 11 (2013), 1927–1940.
- Diebold F (2012) A personal perspective on the origin(s) and development of 'big data': The phenomenon, the term, and the discipline. Available at: http://www.ssc.upenn.edu/fdiebold/papers/paper112/Diebold_Big_Data.pdf (accessed 5 February 2013). F (2012) A personal perspective on the origin(s) and development of 'big data': The phenomenon, the term, and the discipline. Available at: http://www.ssc.upenn.edu/fdiebold/papers/paper112/Diebold_Big_Data.pdf (accessed 5 February 2013).

- Dim P. Papadopoulos, Savvas A. Chatzichristofis, and Nikos Papamarkos. 2011. Video summarization using a selfgrowing and self-organized neural gas network. In Proceedings of International Conference on Computer Vision/Computer Graphics Collaboration Techniques and Applications (MIRAGE'11). Springer, 216–217.
- Hossam M. Zawbaa, Nashwa El-Bendary, Aboul Ella Hassanien, and Ajith Abraham. 2011. SVM-based soccer video summarization system. In Proceedings of the 3rd World Congress on Nature and Biologically Inspired Computing (NaBIC'11). 7–11.
- J. Jin, F. Liu, Z. Gan, and Z. Cui, “Online video synopsis method through simple tube projection strategy,” in Int. Conf. on Wireless Commun. Signal Process., Yangzhou, China, Oct 2016, pp. 1–5.
- Jia Xu, Lopamudra Mukherjee, Yin Li, Jamieson Warner, James M. Rehg, and Vikas Singh. 2015. Gaze-enabled egocentric video summarization via constrained submodular maximization. In Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'15). 2235–2244.
- Jie Ling Lai and Yang Yi. 2012. Key frame extraction based on visual attention model. J. Vis. Commun. Image Rep. 23, 1 (2012), 114–125.
- M. Guzel , "Çok Yönlü Araç Takibi ve Sayımı Uygulaması", Karaelmas Fen ve Mühendislik Dergisi, c. 7, sayı. 2, ss. 622-626, Haz. 2017
- M. S. Güzel, “Performance evaluation for feature extractors on street view images,” imaging sci. j., vol. 64, no. 1, pp. 26–33, 2016.
- Madhushree Basavarajaiah and Priyanka Sharma. 2019. Survey of Compressed Domain Video Summarization Techniques. ACM Computing Surveys, Vol. 52, No. 6, Article 116. Publication date: October 2019.
- Marco Grimaldi, Pádraig Cunningham and Anil Kokaram. 2006. Discrete wavelet packet transform and ensembles of lazy and eager learners for music genre classification. Multimedia Systems (2006) 11(5): 422–437 DOI 10.1007/s00530-006-0027-z
- Marvin Minsky and Seymour Papert. 1972. Artificial Intelligence Process Report. MIT Artificial Intelligence Memo No: 252. 129 – 224 of the 1971 Project MAC Process Report VIII.
- Matthew Dixona, Diego Klabjanb, and Jin Hoon Bangc. 2017. Classification-based financial markets prediction using deep neural networks. Algorithmic Finance 6 (2017) 67–77

- Mikołaj I. Leszczuk and Mariusz Duplaga. 2011. Algorithm for video summarization of bronchoscopy procedures. *BioMed. Eng. Online* 10, 1 (2011).
- Muhammad Ajmal, Muhammad Husnain Ashraf, Muhammad Shakir, Yasir Abbas, and Faiz Ali Shah. 2012. Video summarization: Techniques and classification. In *Proceedings of International Conference on Computer Vision and Graphics (ICCVG'12)*. Springer, 1–13.
- Po Kong Lai, Marc Decombas, Kelvin Moutet, and Robert Laganier. 2016. Video summarization of surveillance cameras. In *Proceedings of IEEE International Conference on Advanced Video and Signal Based Surveillance (AVSS'16)*. 286–294.
- R. Panda, N. C. Mithun, and A. Roy-Chowdhury, “Diversity-aware multi-video summarization,” *IEEE Transactions on Image Processing*, vol. 26, no. 10, pp. 4712–4724, 2017.
- Rameswar Panda and Amit K. Roy-Chowdhury. 2017. Sparse modeling for topic-oriented video summarization. In *Proceedings of IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP'17)*. 1388–1392.
- S. K. Kuanar, K. B. Ranga, and A. S. Chowdhury, “Multi-view video using bipartite matching constrained optimum-path forest ,” *IEEE Transactions on Multimedia*, vol. 17, no. 8, pp. 1166–1173, 2015.
- Saad ALBAWI, Tareq Abed MOHAMMED. 2017. Understanding of a Convolutional Neural Network. 978-1-5386-1949-0/17 IEEE. ICET2017
- Shun-Hsing Ou, Chia-Han Lee, V. Srinivasa Somayazulu, Yen-Kuang Chen, and Shao-Yi Chien. 2014. Lowcomplexity on-line video summarization with Gaussian mixture model based clustering. In *Proceedings of IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP'14)*. 1260–1264.
- Sourya Bhattacharyya, Aditi Roy, Debi Prosad Dogra, Arunava Biswas, Jayanta Mukherjee, Arun Kumar Majumdar, Bandana Majumdar, Suchandra Mukherjee, and Arun Kumar Singh. 2011. Summarization of neonatal video EEG for seizure and artifact detection. In *3rd National Conference on Computer Vision, Pattern Recognition, Image Processing and Graphics (NCVPRIPG'11)*. 134–137.
- Sujata Chaudhari, Nisha Malkan, Ayesha Momin and Mohan Bonde. 2020. Yolo Real Time Object Detection. *International Journal of Computer Trends and Technology (IJCTT)* – Volume 68 Issue 6 – June 2020. ISSN: 2231-2803
- Tao Chen, Aidong Lu, and Shi Min Hu. 2012. Visual storylines: Semantic visualization of movie sequence. *Comput. Graph.* 36, 4 (2012), 241–249.

- Tiejun Huang. 2014. February 2014 Theme - Surveillance Video: The Biggest Big Data. <https://www.computer.org/publications/tech-news/computing-now/surveillance-video-the-biggest-big-data>
- W. Wolf, “Key frame selection by motion analysis,” in *IEEE Trans. Acoust., Speech, Signal Process.*, vol. 2, Atlanta, GA, May 1996, pp. 1228–1231.
- W.-S. Chu, Y. Song, and A. Jaimes, “Video co-summarization: Video summarization by visual co-occurrence,” in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2015, pp. 3584–3592.
- Wei Ting Peng, Wei Ta Chu, Chia Han Chang, Chien Nan Chou, Wei Jia Huang, Wen Yan Chang, and Yi Ping Hung. 2011. Editing by viewing: Automatic home video summarization by viewing behavior analysis. *IEEE Trans. Multimedia* 13, 3 (2011), 539–550.
- Wilson Gavião, Jacob Scharcanski, Jan Michael Frahm, and Marc Pollefeys. 2012. Hysteroscopy video summarization and browsing by estimating the physician’s attention on video segments. *Med. Image Anal.* 16, 1 (2012), 160–176.
- Xiufeng Liu, Per Sieverts Nielsen. 2016. Regression-based Online Anomaly Detection for Smart Grid Data. *arXiv preprint arXiv:1606.05781*, 2016.
- Y. Fu, Y. Guo, Y. Zhu, F. Liu, C. Song, and Z.-H. Zhou, “Multi-view video summarization,” *IEEE Transactions on Multimedia*, vol. 12, no. 7, pp. 717–729, 2010.
- Y. He, Z. Qu, C. Gao, and N. Sang, “Fast online video synopsis based on potential collision graph,” *IEEE Signal Process. Lett.*, vol. 24, no. 1, pp. 22–26, Jan 2017.
- Yashar Deldjoo, Paolo Cremonesi, Markus Schedl, and Massimo Quadrana. 2017. The effect of different video summarization models on the quality of video recommendation based on low-level visual features. In *Proceedings of the 15th International Workshop on Content-Based Multimedia Indexing (CBMI’17)*. 1–6.
- Z. Zivkovic and F. van der Heijden, “Efficient adaptive density estimation per image pixel for the task of background subtraction,” *Patt. Recogn. Lett.*, vol. 27, no. 7, pp. 773–780, May 2006.