

T.C.  
BİLECİK ŐEHY EDEBALI ÜNİVERSİTESİ  
FEN BİLİMLERİ ENSTİTÜSÜ  
ELEKTRONİK VE BİLGİSAYAR MÜHENDİSLİĐİ ANABİLİM DALI

**ENDÜSTRİYEL ORTAMLARDA GÜVENSİZ DAVRANIŐLARIN VİDEO  
ÜZERİNDEN TESPİTİ İÇİN YENİ BİR DERİN ÖĐRENME MODELİ  
GELİŐTİRİLMESİ**

DOKTORA TEZİ

OĐUZHAN ÖNAL

TEZ DANIŐMANI  
DOÇ. DR. EMRE DANDIL

BİLECİK, 2024

10669982

T.C.  
BİLECİK ŐEYH EDEBALI ÜNİVERSİTESİ  
FEN BİLİMLERİ ENSTİTÜSÜ  
ELEKTRONİK VE BİLGİSAYAR MÜHENDİSLİĐİ ANABİLİM DALI

**ENDÜSTRİYEL ORTAMLARDA GÜVENSİZ DAVRANIŐLARIN VİDEO  
ÜZERİNDEN TESPİTİ İÇİN YENİ BİR DERİN ÖĐRENME MODELİ  
GELİŐTİRİLMESİ**

DOKTORA TEZİ

OĐUZHAN ÖNAL

TEZ DANIŐMANI  
DOÇ. DR. EMRE DANDIL

BİLECİK, 2024

10669982

## BEYAN

“Endüstriyel Ortamlarda Güvensiz Davranışların Video Üzerinden Tespiti İçin Yeni Bir Derin Öğrenme Modeli Geliştirilmesi” adlı Doktora Tezinin hazırlık ve yazımı sırasında bilimsel araştırma ve etik kurallarına uyduğumu, başkalarının eserlerinden yararlandığım bölümlerde bilimsel kurallara uygun olarak atıfta bulunduğumu, kullandığım verilerde herhangi bir tahrifat yapmadığımı, tezin herhangi bir kısmının Bilecik Şeyh Edebali Üniversitesi veya başka bir üniversitede başka bir tez çalışması olarak sunulmadığını, aksinin tespit edileceği muhtemel durumlarda doğabilecek her türlü hukuki sorumluluğu kabul ettiğimi ve vermiş olduğum bilgilerin doğru olduğunu beyan ederim.

Bu çalışmanın, Bilimsel Araştırma Projeleri (BAP), TÜBİTAK veya benzeri kuruluşlarca desteklenmesi durumunda; projenin ve destekleyen kurumun adı proje numarası ile birlikte, ETİK KURUL onayı alınması durumunda ise ETİK KURUL tarih karar ve sayı bilgilerinin beyan edilmesi gerekmektedir.			
<b>DESTEK ALINMIŞTIR</b>	<b>X</b>	<b>DESTEK ALINMAMIŞTIR</b>	
<b>Destek alındı ise;</b>			
<b>Destekleyen kurum:</b> Bilecik Şeyh Edebali Üniversitesi Bilimsel Araştırma Projeleri Koordinatörlüğü			
<b>Desteğin Türü</b>		<b>Proje Numarası</b>	
Bilimsel Araştırma Projesi (BAP)		2019-02.BŞEÜ.01-03	
<b>ETİK KURUL onayı var ise;</b>			
<b>ETİK KURUL karar tarih/sayı:</b>		Bilecik Şeyh Edebali Üniversitesi Rektörlüğü Etik Kurulu Karar Tarihi: 11.06.2021 Karar No:10	

**Öğrenci Adı ve Soyadı**

**Oğuzhan Önal**

**Tarih**

.....

**İmza**

.....

## ÖNSÖZ

Lisansüstü eğitim serüvenim boyunca, kıymetli zamanlarını ve engin bilgilerini esirgemeyen, her adımda bana rehberlik eden saygıdeğer danışman hocam Doç. Dr. Emre DANDIL'a en derin şükranlarımı sunarım. Tez çalışmam esnasında, ilim yolculuğuma ışık tutan saygıdeğer hocalarım Doç. Dr. Barış BORU ve Doç. Dr. Süleyman UZUN'a, çalışmalarım da kattıkları paha biçilmez katkı ve yönlendirmeleri için minnettarım.

Doktora tez çalışmasını 2019-02.BŞEÜ.01-03 numaralı proje ile sağladığı destek ve katkılar hasebiyle Bilecik Şeyh Edebali Üniversitesi Bilimsel Araştırma Projeleri Koordinatörlüğü'ne ve çalışmanın yürütülmesinde etik açıdan bir sakınca olmadığını onaylayan Bilecik Şeyh Edebali Üniversitesi Rektörlüğü Etik Kurulu'na gönülden teşekkürlerimi arz ederim.

Tez çalışmasında, veri sağlama ve oluşturma aşamasında katkılar sunan Kafaoğlu Metal Plastik Makina Sanayi ve Ticaret A.Ş.'nin muhterem Yönetim Kurulu Başkanı Mesut YURTCAN'a, Yönetim Kurulu Başkan Yardımcısı Hasan YURTCAN'a, bu güzide şirketin değerli yöneticilerine ve emek sahibi çalışanlarına, Tek Metal Plastik Tar. Hay. Tur. San. ve Tic. Ltd. Şti.'nin kıymetli yöneticilerine ve çalışanlarına teşekkür ederim. Ayrıca, Bilecik Şeyh Edebali Üniversitesi Meslek Yüksekokulu'nun değerli idarecileri ve çalışanlarına, çalışmaya sundukları emekleri ve katkıları ve veriseti oluşturma aşamasındaki destekleri vesilesiyle sonsuz teşekkürlerimi sunarım. Ek olarak, değerli arkadaşım Metin KESLER'e manevi desteğinden dolayı teşekkürü bir borç bilirim. Bunun yanında, video veriseti oluşturma aşamasında desteğini hiçbir zaman esirgemeyen, yanımda olan ve bana güç veren muhterem kuzenlerim; Tek Metal Plastik Tar. Hay. Tur. San. ve Tic. Ltd. Şti. fabrika müdürü Vahit ÖNAL'a, Kafaoğlu Metal Plastik Makina Sanayi ve Ticaret A.Ş. Teknik Müdürü Erhan ÖNAL'a ve değerli eşi Hatice ÖNAL'a yürekten teşekkür ederim.

Ve nihayetinde, bu zorlu yolculuğun her anında yanımda olan, duaları ve desteğiyle beni ayakta tutan sevgili eşim Öznur ÖNAL'a, canım çocuklarım Irmak ÖNAL, Kerem Efe ARSLAN, Rukiye Ayza ÖNAL ve Ali Çağhan ÖNAL'a; varlığıyla bana hayat veren sevgili annem Nermin ÖNAL ve ebediyete irtihal etmiş merhum babam Ali ÖNAL'a ve kardeşlerime kalbimin en derinlerinden gelen duygularla şükranlarımı arz ederim.

**Oğuzhan ÖNAL**

**2024**

## ÖZET

# ENDÜSTRİYEL ORTAMLARDA GÜVENSİZ DAVRANIŞLARIN VIDEO ÜZERİNDEN TESPİTİ İÇİN YENİ BİR DERİN ÖĞRENME MODELİ GELİŞTİRİLMESİ

Üretim sahalarında, işletmelerde, fabrikalarda ve dolayısıyla insanın olduğu her yerde ihmal ve yeterince önemsememe nedeniyle güvenli olmayan davranışlar ortaya çıkabilmektedir. Bu tür alanlarda, güvenli olmayan tehlikeli davranışlar, birçok kaza da dahil olmak üzere ölüm veya yaralanmaların başlıca sebeplerindedir. İşyerlerinde geleneksel iş güvenliği uygulamaları ve düzenli güvenlik denetimleri yapılmasına rağmen, iş sağlığı ve güvenliği protokollerinin ihlal edilmesi sonucu birçok kaza meydana gelmektedir. Bu kazalar işyerlerinde, çalışma ortamının dinamikleri ve zorlukları ile beraber mevcut güvenlik izleme prosedürlerinin yeterince uygulanmadığını göstermektedir. Tehlikeli ortamlarda kazaların ve kayıpların önlenmesine yönelik çözümler, yıllar boyunca olay sayısını azaltmış olsa da insan davranışlarına bağlı olarak tamamen ortadan kaldırılamamaktadır. Ayrıca, bu tür üretim ortamları oldukça karmaşıktır, aydınlatma sorunları vardır ve aşırı dinamiktir. Çalışma ortamlarındaki tehlikeleri kontrol etmek için çeşitli sistemler olmasına rağmen, gerçek zamanlı yaklaşımların çok az olduğu açıktır. Özellikle, birçok bilgisayar tabanlı otomatik çözüm olmasına rağmen, bu sistemlerin eğitim ve tespit süreci düşük doğruluğa sahiptir, çok maliyetlidir ve fazla zaman gerektirir. Diğer taraftan, endüstriyel işyerlerinde, iş sağlığı ve güvenliğinin en önemli unsurlarından birisi olan Kişisel Koruyucu Donanım (KKD) kullanımı büyük önem arz etmektedir.

Bu tez çalışmasında, öncelikle, çoklu sınıflar belirleyerek, YOLO (You Only Look Once) öğrenme algoritması kullanılarak KKD'lerin uygun kullanılıp kullanılmadığı tespit edilmiştir. Bu aşamada, KKD'lere yönelik olarak yedi sınıf belirlenmiş ve bu sınıflara ilişkin özgün veriseti oluşturulmuştur. KKD'lerin tespiti için yürütülen deneysel çalışmalarda, YOLO mimarisi kullanılarak ortalama doğruluk (mean average precision, mAP) değeri %91.18 olarak başarılmıştır. Ayrıca, diğer ölçüm metrikleri Kesinlik (Precision), Duyarlılık (Recall), F1-skoru, kesiştirilmiş bölgeler (Intersection over Union, IoU) ve ortalama kayıp (Loss) için sırasıyla 0.89, 0.91, 0.90, 70.35 ve 1.1147 sonuçları elde edilmiştir. Tez çalışmasının ikinci aşamasında, işyerlerindeki güvensiz hareketlerin gerçek zamanlı sınıflandırılması için derin öğrenme modelleri destekli, hibrit bir bilgisayarla görme yaklaşımı Unsafe-Net (Önal & Dandıl, 2024a) geliştirilmiştir. Unsafe-Net altyapısı için öncelikle bir fabrikadan 39 günlük video görüntüsü toplanarak bir veriseti oluşturulmuştur. Oluşturulan bu veriseti 2024 yılında Data in Brief dergisinde yayınlanarak kullanıma sunulmuştur (Önal & Dandıl, 2024b). Çalışma için

özel olarak oluşturulan veriseti kullanılarak, hızlı ve doğru sonuçlar elde etmek için nesne algılama ve video anlamlandırılmasında YOLOv4 ve Evrişimsel Uzun Kısa Süreli Bellek (Convolutional Long Short-Term Memory, ConvLSTM) derin öğrenme mimarileri birleştirilmiştir. Tez çalışmasının bu aşamadaki deneysel çalışmalarında, önerilen Unsafe-Net mimarisi kullanılarak çalışma sahalarında güvensiz davranışların sınıflandırma doğruluğu %95.81 olarak başarılmıştır ve videolardan eylem tanıma için ortalama süre 0.14 saniye olarak hesaplanmıştır. Ayrıca Unsafe-Net mimarisinin altyapısında kullanılan YOLO algoritması sayesinde, ortalama video süresi 1.87 saniyeye düşürülerek gerçek zamanlı tespit hızı da artırılmıştır. Geliştirilen sistem Eskişehir’de faaliyet gösteren bir fabrika ortamında kurulmuş olup, güvensiz davranış gerçekleştiğinde, gerçek zamanlı olarak çalışanların sistem tarafından hem sesli hem de görsel olarak anında uyarılması sağlanmıştır. Ayrıca, sistemin fabrika ortamına kurulması ile çalışanlar belirli bir süre izlenmiş ve güvensiz davranışların tekrarlanma oranının kısa sürede yaklaşık olarak %75 oranında azaldığı tespit edilmiştir.

**Anahtar Kelimeler:** İş Sağlığı ve Güvenliği, Güvenli Olmayan Davranış Tespiti, Derin Öğrenme, Bilgisayarla Görme, CNN, YOLO, ConvLSTM.

## ABSTRACT

### DEVELOPMENT OF A NEW DEEP LEARNING MODEL FOR VIDEO DETECTION OF UNSAFE BEHAVIORS IN INDUSTRIAL ENVIRONMENTS

Unsafe behaviour can occur as a result of negligence and lack of due caution in production plants, businesses, factories and other places where people are present. In such areas, unsafe hazardous behaviour is a major cause of death or injury, including many accidents. Despite traditional occupational safety practices and regular safety inspections of workplaces, many accidents occur as a result of violations of occupational health and safety protocols. These accidents show that existing safety control procedures are not adequately implemented in workplaces, and that the dynamics and challenges of the work environment are not adequately addressed. Although solutions to prevent accidents and losses in hazardous environments have reduced the number of incidents over the years, they cannot be completely eliminated due to human behaviour. In addition, such production environments are highly complex, have lighting problems and are extremely dynamic. Although various systems exist to control hazards in work environments, it is clear that there are very few real-time approaches. In particular, although there are many computer-aided automated solutions, the training and detection process of these systems has low accuracy, is very costly and time consuming. On the other hand, the use of personal protective equipment (PPE), which is one of the most important elements of occupational health and safety, is of great importance in industrial workplaces.

In this thesis, the appropriate use of PPEs was first determined by identifying multiple classes using the YOLO (You Only Look Once) learning algorithm. At this stage, seven classes of PPEs were identified and a unique dataset was created for these classes. In the experimental studies carried out to detect PPEs, the mean average precision (mAP) value achieved using the YOLO architecture was 91.18%. Furthermore, for the other metrics, precision, recall, F1 score, intersection over union (IoU) and average loss, the results obtained were 0.89, 0.91, 0.90, 70.35 and 1.1147 respectively. In the second phase of the thesis, Unsafe-Net (Önal & Dandil, 2024a), a hybrid computer vision approach supported by deep learning models, was developed for real-time classification of unsafe movements in workplaces. For the Unsafe-Net infrastructure, a dataset was created by collecting 39 days of video footage from a factory. This dataset was published in the journal Data in Brief in 2024 and made available for available (Önal & Dandil, 2024b). Using the database created specifically for the study, YOLOv4 and Convolutional Long Short-Term Memory (ConvLSTM) deep learning architectures were combined in object recognition and video interpretation to achieve fast and accurate results. In the experimental

studies at this stage of the thesis, the classification accuracy of unsafe behaviours in workplaces using the proposed Unsafe-Net architecture was achieved as 95.81% and the average time for action recognition from videos was calculated as 0.14 seconds. In addition, thanks to the YOLO algorithm used in the infrastructure of the Unsafe-Net architecture, the average video duration was reduced to 1.87 seconds, thus increasing the real-time detection speed. The developed system has been installed in a factory environment operating in Eskisehir, and when unsafe behaviour occurs, employees are immediately alerted by the system both audibly and visually in real time. In addition, with the installation of the system in the factory environment, employees were monitored for a period of time and it was observed that the recurrence rate of unsafe behaviour decreased by approximately 75% in a short period of time.

**Keywords:** Occupational Health and Safety, Unsafe Behavior Detection, Deep Learning, Computer Vision, CNN, YOLO, ConvLSTM.

# İÇİNDEKİLER

	Sayfa
ÖNSÖZ.....	i
ÖZET .....	ii
ABSTRACT .....	iv
İÇİNDEKİLER.....	vi
TABLolar LİSTESİ.....	viii
ŞEKİLLER LİSTESİ.....	x
GRAFİKLER LİSTESİ .....	xii
KISALTMALAR VE SİMGELER LİSTESİ .....	xiii
<b>1. GİRİŞ .....</b>	<b>1</b>
1.1. İş Sağlığı ve Güvenliği.....	4
1.2. Tezin Amacı ve Hedefleri .....	8
1.3. Araştırma Sorusu ve Hipotez.....	9
1.4. Tez Çalışmasının Katkısı.....	10
1.5. Tezin Organizasyonu .....	11
<b>2. LİTERATÜR TARAMASI .....</b>	<b>13</b>
2.1. Geleneksel Yaklaşımlarla Önerilen Yöntemler .....	14
2.2. Makine Öğrenmesi Tabanlı Çalışmalar .....	17
2.3. Derin Öğrenme Tabanlı Çalışmalar .....	19
<b>3. MATERYAL VE YÖNTEM.....</b>	<b>33</b>
3.1. Verisetleri.....	33
3.1.1. KKD Tespiti İçin Hazırlanan Veriseti.....	33
3.1.2. Güvenli ve Güvensiz Davranışların Sınıflandırılması İçin Hazırlanan Video Veriseti	35
3.2. YOLO Mimarileri .....	39
3.2.1. YOLOv4 .....	40
3.2.2. YOLOv11 .....	41
3.3. Önerilen Hibrit Derin Öğrenme Mimarisi (Unsafe-Net).....	42
3.3.1. ConvLSTM .....	43

3.3.2. YOLOv4 + ConvLSTM .....	44
4. DENEYSEL ÇALIŞMALAR.....	47
4.1. KKD Tespitine Yönelik Yürütülen Deneysel Analizler .....	47
4.2. Gerçek Zamanlı Olarak Güvensiz Davranışların Tespiti İçin Geliştirilen Bilgisayarla Görme Sisteminde (Unsafe-Net) Yürütülen Deneysel Analizler .....	53
4.2.1. YOLO Mimarisi ile Uygun Çerçeve Seçimi.....	54
4.2.2. YOLOv4+ConvLSTM ile Sınıflandırma.....	60
4.2.3. Uyarı İzleme Sistemi .....	64
5. BULGULAR VE TARTIŞMA .....	69
6. SONUÇ.....	72
KAYNAKÇA .....	74

## TABLULAR LİSTESİ

	Sayfa
<b>Tablo 2.1.</b> Tez konusu ile ilişkili geleneksel yöntemlerle önerilen çalışmaların genel değerlendirmesi .....	27
<b>Tablo 2.2.</b> Tez konusu ile ilgili makine öğrenmesi kullanarak önerilen çalışmaların değerlendirmesi .....	28
<b>Tablo 2.3.</b> Tez konusu ile ilgili derin öğrenme tabanlı çalışmalar .....	29
<b>Tablo 3.1.</b> KKD tespiti için nesne tanımda kullanılan sınıflar ve etiketlenen sınıflardaki nesne sayıları .....	35
<b>Tablo 3.2.</b> Güvenli ve güvenli olmayan davranış sınıfları için sınıf kimliği, sınıf adı ve davranış türü bilgileri .....	36
<b>Tablo 3.3.</b> 05-23 Kasım 2022 tarihleri arasında toplanan video parçalarından oluşan eğitim ve test kümesi sayıları .....	37
<b>Tablo 3.4.</b> YOLOv4 kullanılarak nesne tespiti için eğitim ve test setlerinde etiketlenen görüntülerin sayısı .....	38
<b>Tablo 3.5.</b> YOLOv11 kullanılarak nesne tespitinde her bir sınıf için eğitim, doğrulama ve test kümelerinde etiketlenen görüntülerin sayısı.....	39
<b>Tablo 4.1.</b> Çalışma ortamlarından nesne tanıma için önerilen YOLOv4 ağının eğitim parametreleri ve süreleri .....	48
<b>Tablo 4.2.</b> Farklı eğitimler için YOLOv4 modeli ile elde edilen performans metrik sonuçların karşılaştırılması.....	49
<b>Tablo 4.3.</b> YOLOv4 ağı ile 5 numaralı eğitimde her bir sınıf için elde edilen AP, DP ve YP değerleri.....	51
<b>Tablo 4.4.</b> YOLOv4 ağı ile 5 numaralı eğitimde IoU eşik değeri 0.75 için her bir sınıf ait AP, DP ve YP değerleri.....	53
<b>Tablo 4.5.</b> YOLOv4 ağı ile test işleminde her bir sınıf için elde edilen AP skoru ve DP ile YP değerleri.....	56
<b>Tablo 4.6.</b> Verisetinin YOLOv11 eğitimi sonucunda elde edilen sonuçlar.....	60

<b>Tablo 4.7.</b> Video test kümesi için YOLOv4 ile video çerçeve azaltma olmadan ConvLSTM ağının her bir güvensiz hareket sınıfı için performansı .....	<b>62</b>
<b>Tablo 4.8.</b> Video test kümesi için bu çalışmada önerilen Unsafe-Net (YOLOv4+ConvLSTM) ağının her bir güvensiz hareket sınıfı için performansı .....	<b>62</b>
<b>Tablo 4.9.</b> Video test kümesi için sadece ConvLSTM ağının ve önerilen Unsafe-Net modelinin her bir güvensiz hareket sınıfında doğruluk metriği için performansı .....	<b>63</b>
<b>Tablo 5.1.</b> ConvLSTM ve önerilen Unsafe-Net için ortalama video süresi ve ortalama eylem tanımlama süresinin karşılaştırılması .....	<b>69</b>
<b>Tablo 5.2.</b> Fabrikaya gerçek zamanlı Unsafe-Net mimarisini içeren sistemin kurulmadan öncesi için belirli bir tarih arasında meydana gelen güvensiz eylemlerin sayısı.	<b>70</b>
<b>Tablo 5.3.</b> Fabrikaya gerçek zamanlı Unsafe-Net mimarisini içeren sistemin kurulmasından sonra meydana gelen güvensiz eylemlerin sayısı .....	<b>70</b>

## ŞEKİLLER LİSTESİ

Sayfa

- Şekil 1.1.** WoS üzerinde “occupational safety and health” için alınan geri dönütlerin VOSviewer ile gösterilmesi..... 8
- Şekil 2.1.** WoS üzerinde “Bilgisayarla Görme ve İş sağlığı ve güvenliği” için alınan geri dönütlerin VOSviewer ile gösterilmesi ..... 13
- Şekil 3.1.** Tez kapsamında KKD’lerin tanınması için hazırlanan verisetinden örnek görüntüler. .... 34
- Şekil 3.2.** Video veriseti hazırlanması aşamasında veri alınan fabrika ortamının şematik gösterimi ve kameraların konumu ..... 36
- Şekil 3.3.** Verisetinde güvenli ve güvensiz sınıflarına ait davranışlara örnekler. (a) Güvenli Yürüyüş Yolu, (b) Güvenli Yürüyüş Yolu İhlali, (c) Açık Pano Kapağı, (d) Kapalı Pano Kapağı, (e) Forklift ile Aşırı Yük Taşıma, (f) Güvenli Yük Taşıma, (g) Yetkili Müdahale, (h) Yetkisiz Müdahale ..... 38
- Şekil 3.4.** Çalışmada video parçalarından çerçeve seçilmesi için kullanılan YOLOv4 algoritmasının blok diyagramı..... 41
- Şekil 3.5.** Videolardan güvensiz davranışların gerçek zamanlı olarak tespiti için önerilen YOLOv4 ve ConvLSTM tabanlı Unsafe-Net hibrit derin öğrenme sisteminin blok diyagramı ve mekanizması ..... 42
- Şekil 3.6.** ConvLSTM mimarisinin blok diyagramı (Shi vd., 2015) ..... 44
- Şekil 3.7.** Güvensiz davranış tespiti için YOLOv4 ve ConvLSTM tabanlı önerilen Unsafe-Net derin öğrenme modelinin mimarisi ..... 45
- Şekil 3.8.** Güvensiz davranışların tespiti için önerilen YOLOv4+ConvLSTM derin öğrenme modelinde bulunan ConvLSTM yapısının mimarisinin blok diyagramı..... 46
- Şekil 4.1.** Önerilen YOLOv4 modeli ile video çerçeveleri üzerinde çalışma ortamında KKD tespitinin yapılması..... 52
- Şekil 4.2.** İş sahasından alınan durağan bir görüntüde önerilen YOLOv4 mimarisi ile KKD tespitinin yapılması..... 52
- Şekil 4.3.** Test kümesindeki video üzerinde güvenli ve güvensiz davranış sınıflarına ait YOLOv4 ile başarılı bir şekilde tanınmış nesne örnekleri ..... 57

<b>Şekil 4.4.</b> Başlangıçta 105 çerçeveden oluşan bir video parçasının YOLOv4 ağından geçirildikten sonra, içerisinde 3 eylem ve 40 çerçeveden oluşan bir hale gelmesi ve azaltılmış videodaki bazı örnek çerçeveler.....	<b>58</b>
<b>Şekil 4.5.</b> Önerilen Unsafe-Net yöntemi için doğru ve yanlış sınıflandırma örnekleri .....	<b>64</b>
<b>Şekil 4.6.</b> Geliştirilen uyarı ve denetim sisteminin blok diyagramı.....	<b>65</b>
<b>Şekil 4.7.</b> Uyarı izleme sisteminin kontrol kartı devre şeması.....	<b>65</b>
<b>Şekil 4.8.</b> Unsafe-Net mimarisinde sistemin izlenmesi ve kontrolü için geliştirilen programın arayüzü .....	<b>66</b>
<b>Şekil 4.9.</b> Güvensiz hareket meydana geldiğinde sistem tarafından oluşturulan metin belgesi tipinde raporlama içeriği.....	<b>67</b>
<b>Şekil 4.10.</b> Güvenli olmayan bir davranış meydana geldiğinde sesli ve görsel uyarı verilmesi .....	<b>68</b>

## GRAFİKLER LİSTESİ

Sayfa

- Grafik 4.1.** Önerilen YOLOv4 ağının 5 numaralı eğitim için ortalama kayıp ve mAP değerlerinin iterasyon sayısına göre değişimi ..... **50**
- Grafik 4.2.** Güvensiz davranışların tespiti için YOLOv4'ün eğitimi esnasında Loss ve mAP değerlerinin iterasyon sayısına göre değişimi ..... **55**
- Grafik 4.3.** Verisetinin YOLOv11 ile eğitilmesi sonucunda kayıp ve diğer metriklerin değişimi ..... **59**
- Grafik 4.4.** ConvLSTM ağının eğitimi süresinde kayıp, doğrulama kayıp, doğruluk ve doğrulama doğruluk değerlerinin değişimi ..... **61**

## KISALTMALAR VE SİMGELER LİSTESİ

<b>2B</b>	: İki Boyut
<b>3B</b>	: Üç Boyut
<b>ABD</b>	: Amerika Birleşik Devletleri
<b>B</b>	: Tahmin Edilen Değer
<b>BAP</b>	: Ortalama Kesinlik
<b>BGT</b>	: Referans Değer
<b>ACC</b>	: Doğruluk
<b>BBS</b>	: Davranış Temelli Güvenlik
<b>CBIR</b>	: İçerik Tabanlı Görüntü Alma
<b>CH</b>	: Renk Histogramı
<b>CNN</b>	: Evrişimli Sinir Ağı
<b>Conv</b>	: Evrişim
<b>CPS</b>	: Siber-Fiziksel Sistem
<b>CV</b>	: Bilgisayarla Görme
<b>DP</b>	: Doğru Pozitif
<b>DN</b>	: Doğru Negatif
<b>DVM</b>	: Destek Vektör Makinesi
<b>FFH</b>	: Yüksekten Düşme
<b>YPS</b>	: Saniyede Kare Sayısı
<b>F</b>	: F-Skor
<b>GSMH</b>	: Gayri Safi Milli Hasıla
<b>GPU</b>	: Grafik İşleme Ünitesi (Graphical Processing Unit)
<b>H</b>	: Gizli Katman
<b>HAR</b>	: İnsan Aktivitesi Tanıma (Human activity recognition)
<b>H-SVM</b>	: Hiyerarşik Destek Vektör Makinesi
<b>HMI</b>	: Hu Moment Değişmezleri
<b>I3D</b>	: Genişletilmiş 3B ConvNet (Inflated 3D ConvNet)
<b>ILO</b>	: Uluslararası Çalışma Örgütü
<b>IoU</b>	: Kesişim

<b>IP</b>	: İnternet Protokolü
<b>İSG</b>	: İş Sağlığı ve Güvenliği
<b>KKD</b>	: Kişisel Koruyucu Ekipmanı
<b>LBP</b>	: Yerel İkili Örüntüler
<b>LSTM</b>	: Uzun Kısa Süreli Hafıza (Long Short Term Memory)
<b>LRCN</b>	: Uzun Süreli Tekrarlayan Evrişimsel (Long-Term Recurrent Convolutional)
<b>mAP</b>	: Tüm Sınıfların Ortalama Kesinlik Değeri
<b>MEI</b>	: Hareket Enerji Görüntüsü (Motion Energy Image)
<b>MHI</b>	: Hareket Geçmişi Görüntüleri (Motion History Images)
<b>SSD</b>	: Çoklu Kutu Dedektörü
<b>SSP</b>	: Uzaysal Piramit Havuzlama
<b>OKB</b>	: Nesne Korelasyonu Tespiti
<b>OSB</b>	: Organize Sanayi Bölgesi
<b>P</b>	: Kesinlik (Precision)
<b>PAN</b>	: Yol Birleştirme Ağı
<b>PTSD</b>	: Travma Sonrası Stres Bozukluğu
<b>R-CNN</b>	: Bölge tabanlı Evrişimli Sinir Ağı
<b>RGB</b>	: Kırmızı Yeşil Mavi
<b>RNN</b>	: Tekrarlayan Sinir Ağı
<b>Recall</b>	: Hatırlama
<b>RFID</b>	: Radyo Frekans Tanıma Sistemi
<b>tanh</b>	: Hiperbolik Tanjant Aktivasyon Fonksiyonu
<b>T3D</b>	: Zamansal Üç Boyut (Temporal 3D)
<b>TTL</b>	: Zamansal Geçiş Katmanı, (Temporal Transition Layer)
<b>YOLO</b>	: Sadece Bir Kere Bak (You Only Look Once)
<b>YN</b>	: Yanlış Negatif
<b>YP</b>	: Yanlış Pozitif
<b>X</b>	: Ağ Girişleri
<b>C</b>	: Bellek Birimleri
<b>°</b>	: Hadamaar Çarpımı

- \* : Konvolüsyon İşlemi
- $I_t$  : Giriş Kapısı
- $f_t$  : Unutma Kapısı
- $Q_t$  : Çıkış Kapısı
- $C_t$  : ConvLSTM Ağının Bellek Birimleri
- $\sigma$  : Sigmoid Aktivasyon Fonksiyonu

## 1. GİRİŞ

Günümüzde, rekabete dayalı ekonomik kalkınma ve üretim güdüsü çoğunlukla, işyerlerinde ve farklı çalışma ortamlarında toplumsal ve ekonomik sonuçları olan bir dizi mesleki tehlikelere yol açmaktadır. Bu tehlikeler tipik olarak iş kazaları şeklinde ortaya çıkmaktadır ve çalışanların sağlık ve güvenliği açısından ciddi riskler oluşturabilmektedir. İş kazası, işyerlerinde çalışma sürecinde meydana gelen, yaralanma, ölüm veya sağlık sorunlarına neden olan olaylardır (Gulhan vd., 2012). İş kazaları genellikle beklenmeyen ve istenmeyen durumlar olarak kabul edilir. Bu olaylar, bir kişinin işyerindeki faaliyetler sırasında, işle ilgili görevleri yerine getirirken veya işyeri koşullarıyla doğrudan ilişkilendirilebilecek bir durumda meydana gelir. İş kazaları, işçilerin sağlığı ve yaşamı ile işletmenin üretim sürecini doğrudan etkilediği gibi, sonuçları bakımından toplumu da yakından ilgilendirmektedir. Bunun nedeni, iş kazalarının işçi, işveren, ülke ekonomisi ve toplum açısından önemli sosyal ve ekonomik maliyetler yaratmasıdır (Demirbilek & Pazarlıoğlu, 2007; LaDou vd., 2018; Prasad, 2017).

İnsanların güvenliğini sağlamak, çalışma ortamlarında var olan dinamik ve karmaşık çalışma koşulları nedeniyle iş yerlerinde yaygın ve oldukça zorlu bir görevdir. Düzenleyici reformlara, yasalara ve sanayi birliklerinin çabalarına ve bu sorunu gidermek için kapsamlı araştırmalara rağmen, çalışma sırasındaki kazalar ve ölümler ciddi sorun oluşturmaktadır. Heinrich'e (Heinrich, 1941) göre çalışma sırasında meydana gelen tüm kazaların yaklaşık %88'i güvensiz davranışların bir sonucu olarak ortaya çıkmaktadır.

Üretim yapılan ortamlarda ve işyerlerinde meydana gelen iş kazaları çoğunlukla işletme kusurlarından ve tehlikeli ortamlardan kaynaklandığı gibi, bunun bir yansıması olarak da çalışanlar tarafından gerçekleştirilen güvensiz davranışlar, iş kazalarının en temel sebeplerindendir (Dodoo & Al-Samarraie, 2019). Güvensiz davranış hatalara yol açabilecek standart prosedürlerin kasıtlı ihlali olarak tanımlanmıştır (Reason vd., 1990). Kazaların büyük çoğunluğunun güvensiz iş davranışlarından veya insan hatalarından kaynaklandığı düşüncesi, birçok araştırmacı tarafından desteklenmiştir (Adhikari, 2015; Alkhaldi vd., 2017; Liao vd., 2018; Omidvari & Gharmaroudi, 2015; Reyes vd., 2015; Strauch, 2017; Uğurlu vd., 2015). Heinrich (Heinrich, 1941), 75000 kaza kaydının örnek olay incelemelerinden elde ettiği verilerde, tüm endüstriyel kazaların %88'inin bireyin güvensiz davranışından, %10'unun güvensiz koşullardan ve geriye kalan %2'sinin öngörülemeyen koşullardan kaynaklandığını belirtmiştir.

İş kazalarının üretim sektöründeki ekonomik etkisi büyük, karmaşık ve yaralayıcıdır. Doğrudan maliyetler tıbbi harcamalar, tazminat ödemeleri, üretkenlik kaybı ve hukuki süreçlerden kaynaklanmaktadır. Ayrıca ekipman hasarı, üretim süreçlerindeki aksamalar ve mali yükü artıran sigorta prim artışları gibi dolaylı maliyetler de söz konusudur. Dünyada iş kazalarının etkilerinin başta raporlama kusurlarından kaynaklı olmak üzere tahmin edilenden çok daha yüksek olduğu öngörülmektedir (Altunkaynak, 2018; Tompa vd., 2021b).

Son tahminler, dünya genelinde 0.3 milyon ölümlle sonuçlanan yaklaşık 2.3 milyon iş kazası ve meslek hastalığı vakası olduğunu göstermektedir (Cavlak vd., 2022). İşle ilgili yaralanma ve hastalıkların tahmini ekonomik maliyeti, Uluslararası Çalışma Örgütü (ILO) tarafından sağlanan verilere göre ortalama %4 olmak üzere, farklı ülkelerde Gayri Safi Yurtiçi Hasılanın (GSYH) %1.8'i ile %6'sı arasında değişmektedir (Chen vd., 2019; Takala vd., 2014). İş kazaları ve meslek hastalıklarının ekonomik yükü, bazı Avrupa ülkelerinde GSYH'nin %2.9 ila %10.2'si arasında değişmektedir (Tompa vd., 2021a). Vaka başına maliyet miktarı bazı ülkelerde 75000 avro düzeylerine çıkabilmektedir.

İşyerlerinde ve çalışma ortamlarında güvenliğin sağlanması, çalışma koşullarının sürekli değişen dinamik ve karmaşık yapısı nedeniyle önemli bir zorluk teşkil etmektedir. Güvenlik kurallarına, talimatlarına ve standartlarına uyulmaması tehlikeli davranışlara yol açabilir ve çalışanları risk altına sokarak iş performansını etkileyebilir (Ding vd., 2018; Önal & Dandıl, 2021). Düzenlemelerdeki sürekli güncellemelere ve artan önleyici tedbirlere rağmen, işyeri kazaları, yaralanmalar ve ölümler önemli bir sorun olmaya devam etmektedir (Ceylan & Ceylan, 2012). Bu ortamlarda güvenlik performansını artırmak için, güvenli olmayan davranışları azaltmak veya önlemek çok önemlidir.

Son zamanlarda, mesleki riskleri azaltmak için gelişmiş araçlar uygulanmaktadır. Özellikle, sektörde çalışanların davranışlarını düzenleyen yönetmelikler değişen koşullara göre güncellenmektedir. Şirketler tarafından çeşitli eğitim programları yürütülmekte ve çalışanların güvenlik ve sağlık protokollerine sürekli olarak uymalarını sağlamak için gelişmiş önleyici adımlar atılmaktadır. Üretim ortamları çalışanların ergonomisine göre tasarlanmaktadır (Sadeghi vd., 2016). Tipik olarak, işçiler arasındaki güvensiz uygulamaları tespit etmek ve izlemek için gözleme dayalı yaklaşımlar tercih edilmektedir. Ancak bu yöntemler zaman alıcı, bireyselleştirilmiş ve yoğun emek gerektiren faaliyetler olmaları nedeniyle sınırlamalara sahiptir. Sonuç olarak, güvenli olmayan işyeri davranışlarının otomatik sistemler kullanılarak tespit edilmesi büyük önem taşımaktadır.

Her işyerinin kendine özgü tehlikeleri vardır ancak imalat endüstrisi diğer birçok modern çalışma ortamına göre daha yüksek yaralanma ve ölüm riski barındırmaktadır (Muthukumar vd., 2019). Tonlarca ağırlığında presler hareketli dönen makinalar vs. barındıran bu üretim alanlarında en ufak bir boşluk dalgalılık veya güvensiz davranış ölümcül sonuçlar doğurabilmektedir. Şirket ve üretim tesisi çalışma ortamlarında iş sağlığı ve güvenliği (İSG) yönetmeliklerine uyulması ve kişisel koruyucu ekipman kullanılması esastır. Tesislerde tehlikeli olarak kabul edilen davranışların gözlemlenmesi çok önemlidir (Barro-Torres vd., 2012).

Son zamanlarda, video gözetim teknolojisi, işyerinde iş sağlığı ve güvenliği için tehlikeli faaliyetlerin tanımlanması ve kategorize edilmesinde yaygın olarak kullanılmaktadır. Pandemi gibi dünya çapında yaşanan olaylar gözlem kameralarının önemini artırmış ve popüler bir konu haline getirmiştir. Dünyada en çok gözlem kamerasına sahip bir ülke salgın hastalıktan gözlem ağını etkin kullanımı sayesinde kısa sürede kurtulmuştur. Bu başarı diğer bazı ülkeleri de takip ağırları oluşturma konusunda harekete geçirmiştir. Bu yüzden gözlem kameraları havaalanları, terminaller, hastaneler, fabrikalar ve işletmeler gibi insanın olduğu hemen hemen her yerde yaygın olarak kullanılmaya başlanmıştır.

Video gözetim sistemleri kullanılarak insan faaliyetlerinin tanınması, içerik tabanlı video analizi, ağ tabanlı gözetim, kullanıcı arayüzü ve izleme gibi bilgisayarla görme uygulamalarını sunan çağdaş bir araştırma alanıdır (Onofri vd., 2016). Başlangıçta, eylem tanıma ve video içeriğini anlama üzerine yapılan araştırmalar, el sallama ve yürüme gibi karmaşık olmayan insan hareketlerinin tespitine odaklanmıştır (Poppe, 2010). Daha yakın zamanlarda, araştırmalar nesne etkileşimleri ve çoklu görevler içeren karmaşık faaliyetleri içeren gerçekçi ve zorlu konulara ağırlık vermiştir (Wolf vd., 2014). Derin öğrenme, son yıllarda görüntü işleme ve veri analizi alanlarında önemli ilerlemeler kaydetmiştir. Bu bağlamda, nesne tespiti ve aktivite tanıma, geniş uygulama alanları ve pratik yararları nedeniyle yoğun araştırmaların odak noktası olmuştur. Nesnelerin tespiti ve video anlamlandırma ile ilgili 20 yılı aşkın bir süredir birçok çalışma yapılmış olsa da hala aşılamayan bir takım teknik ve teknolojik zorluklar ve sınırlılıklar bulunmaktadır. Farklı algılama görevleri tamamen farklı hedeflere ve kısıtlamalara sahip olduğundan, zorlukları da birbirinden farklı olabilir. Değişik bakış açıları, aydınlatmalar ve sınıf içi varyasyonlar altındaki nesnelere gibi diğer bilgisayarla görme görevlerindeki bazı ortak zorluklara ek olarak, nesne tespitindeki zorluklar, nesne rotasyonu ve ölçek değişiklikleri, doğru nesne lokalizasyonları, yoğun ve tıkalı nesne gibi birçok zorluğu içinde barındırır (Zou vd., 2023).

Çalışma ortamlarında işçilerin/çalışanların davranışını tespit etmeye yönelik geleneksel çalışmalar ağırlıklı olarak gözlemsel yöntemlere dayanmaktadır. Bu tür yöntemler faydalı bilgiler sunsa da zaman alıcı, yoğun emek gerektiren ve doğadan özneldirler. Bu sınırlamalar nedeniyle, nesnelerin tanınması için kullanılan bilgisayarla görme teknolojisi, işçilerin güvensiz eylemlerini yerinde belirlemek için uygulanabilir (Ding vd., 2018). Mevcut çalışmalarda aktivite tanıma sistemlerinin etkinliğini belirleyen çok çeşitli etkenler vardır. Bunlardan bazıları, tanıma sürecinin çeşitli aşamalarında kullanılan yöntemlere özgüdür. Diğerleri ise edinim aygıtları, deney ortamları veya bu sistemlerin çeşitli uygulamalarıyla ilgilidir. Aydınlatma değişimi, genel olarak bilgisayarla görme tabanlı tanıma sistemlerinin karşılaştığı ana zorluktur, çünkü görüntülerin kalitesini ve dolayısıyla analiz edilen bilgileri etkilemektedir. Aynı şekilde, perspektif değişikliği, tek bir görünüm edinim aygıtları kullanılarak çalışmak üzere uygulanan mevcut sistemlerin bir başka sınırlamasıdır. Bu sorun, çıkarılan bilgi miktarını azaltır ve analiz edilen aktivitelerin sınırlı bir görselleştirilmesini sağlar. Bu, farklı türleriyle tıkanmayı içerir. Vücut parçalarının birbirini örttüğü kendi kendine tıkanma, başka bir nesnenin tıkanması ve insan vücut parçalarının kısmi tıkanması, aktivite tanıma sistemlerinin başlıca sınırlamalarıdır (Beddiar vd., 2020).

Bu tez çalışmasında, işyerlerindeki güvensiz hareketlerin ve davranışların gerçek zamanlı sınıflandırılması için derin öğrenme modelleri kullanan hibrit bir bilgisayarla görme yaklaşımı ve KKD tespiti için YOLO tabanlı bir derin öğrenme yaklaşımı önerilmektedir. Önerilen hibrit derin öğrenme modelinde hızlı ve doğru sonuçlar elde etmek için, videolardan nesne algılama aşamasında YOLOv4, verisetindeki videoların sınıflandırılarak anlamlandırılması aşamasında ise ConvLSTM mimarileri birleştirilmiştir. Bunun yanında, tez çalışması kapsamında, farklı çalışma ortamlarından toplanan videolar kullanılarak verisetleri hazırlanmıştır. Ayrıca, tez çalışması kapsamında hazırlanan gerçek zamanlı bilgisayarla görme sistemi, bir fabrika ortamına kurularak, önerilen sistemin güvenilirliği ve performansı çalışma ortamında gözlemlenmiştir. Önerilen sistemde, fabrika ortamına bir güvensiz davranış gerçekleştiğinde, çalışanları hem sesli hem de görsel olarak anında uyaracak bir gömülü sistem de tasarlanmıştır.

## **1.1. İş Sağlığı ve Güvenliği**

İş güvenliği sağlığı ve güvenliği, çeşitli işyerlerinde çalışanların sağlığını ve refahını korumaya odaklanan multidisipliner bir çalışma alanıdır. İSG'nin temel amacı, çalışanlar arasında yaralanmaları, hastalıkları ve diğer olumsuz sağlık sonuçlarını önlemek amacıyla işyerinde potansiyel olarak ortaya çıkabilecek veya işyerinden kaynaklanabilecek tehlikeleri

belirlemek, deęerlendirmek ve ynetmektir. Bu alan iř hekimlięi, endstriyel hijyen, ergonomi, psikoloji ve gvenlik mhendislięi gibi bir dizi disiplini kapsar. Bu alanlardaki profesyoneller risk tanımlama, deęerlendirme ve kontrol iin kapsamlı stratejiler geliřtirmek zere iř birlięi yapmaktadır (Jain vd., 2021; Liu vd., 2023).

İř kazalarının dnya genelinde bireyler, aileler, toplumlar ve ekonomiler zerinde derin etkileri vardır. İř kazalarının insani maliyeti byk olup iř kazaları sıklıkla yaralanmalara, hastalıklara, sakatlıklara ve lmlere yol amaktadır (Akbari vd., 2019). Bu tr olaylar, alıřanlar ve aileleri iin fiziksel acı, duygusal sıkıntı ve mali zorluklar řeklinde kendini gsteren nemli sıkıntılara neden olur. İř kazalarında sadece can kayıpları yařanmaz, aynı zamanda iřgcne katkıda bulunan deęerli kiřilerin de kaybına neden olur. İř kazalarının ekonomik etkileri de byktr ve iř kazalarının ekonomik maliyetleri olduka yksektir. İř kazalarının doęrudan maliyetleri arasında tıbbi harcamalar, tazminat demeleri ve iřyeri hasarıyla iliřkili maliyetler yer almaktadır (Jilcha & Kitaw, 2016). Ayrıca, verimlilik kaybı, iře devamsızlık, yedek iřilerin yeniden eęitilmesi ve yasal masraflar gibi dolaylı maliyetler daha da yksek olabilir. Bu tr maliyetler bireysel iřletmeler, sektrler ve tm ekonomiler zerinde ařırı baskı oluřturabilir. İř kazalarının meydana gelmesi, bir kuruluřun normal iřleyiřinin aksamasına neden olarak bir sre hareketsizlięe, gecikmelere ve iřgcnn genel retkenlięinde bir azalmaya yol aabilir. Yaralanan alıřanlar grevlerini yerine getiremeyebilir, bu da iř gn kaybına ve retim azalmasına neden olur. Ayrıca, kazalar ekipman veya tesislerde hasara yol aarak retkenlięi daha da engelleyebilir ve itibar kaybı potansiyeli de sz konusudur. Gemiřinde kt gvenlik kayıtları olan řirketler itibar kaybına uęrayabilir ve bu da alıřanlar, mřteriler ve paydařlar arasında gven kaybına yol aabilir. İřyeri kazalarıyla ilgili olumsuz tanıtım, bir řirketin imajı zerinde zararlı bir etkiye sahip olabilir, yetenekleri ekme ve elde tutma, ayrıca szleřmeleri ve yatırımları gvence altına alma becerisini engelleyebilir. İř kazalarının yasal ve dzenleyici sonuları nemlidir. Bir iř kazası durumunda, iřverenler para cezaları, cezalar ve davalar da dahil olmak zere yasal yansımalara maruz kalabilir. Ayrıca, hkmetler iřyeri gvenlik standartlarını geliřtirmek amacıyla daha sıkı dzenlemeler ve yaptırım tedbirleri uygulayabilir. Bu durum, iřletmeler iin uyum maliyetlerinin artmasına neden olabilir. İř kazalarının psikososyal etkileri nemli olabilir. İř kazalarının psikolojik etkileri, anksiyete, depresyon, travma sonrası stres bozukluęu (PTSD) ve iř memnuniyetinin azalması gibi uzun sreli olabilir (Muthukumar vd., 2019). Kazaların gzlemlenmesi veya bunlara katılım, iřyerinde bir korku ve gvensizlik kltr yaratabilir ve bu da moral ve alıřan baęlılıęı zerinde zararlı bir etkiye sahip olabilir. İř kazalarının meydana

gelmesi, sađlık hizmetleri, engelli yardım programları ve sosyal refah sistemleri de dahil olmak üzere sosyal destek sistemleri üzerinde bir baskıya neden olabilir. Sınırlı kaynaklara veya sosyal güvenlik ađlarına sahip bölgelerde, toplumlar yaralı işçilere bakmanın ve ailelerini desteklemenin yükünü taşıyabilir. Beceri ve deneyim kaybı, iş kazalarının bir diđer sonucudur. İşgücündeki deđerli beceri, bilgi ve deneyim kaybı, iş kazalarından kaynaklanan yaralanma veya ölümlerden kaynaklanabilir. Bu tür olaylar, özellikle vasıflı çalışan talebinin yüksek olduđu uzmanlaşmış sektörlerde, operasyonların verimliliğini ve etkinliğini engelleyebilir (Prasad, 2017; Tompa vd., 2021a).

Sonuç olarak, iş kazalarının sonuçları geniş kapsamlı ve çok yönlü olup, çalışanları korumak ve sürdürülebilir ekonomik ve sosyal kalkınmayı teşvik etmek için çok sağlam iş sađlığı ve güvenliği yönelik olarak çalışmalar yapılması ve gereken önlemlerin alınması gerekmektedir.

İş kazalarının insani maliyeti rakamlara şu şekilde yansımaktadır. ILO, 2018 yılında dünya genelinde yaklaşık 2,8 milyon işçinin işle ilgili kaza veya hastalıklar nedeniyle hayatını kaybettiğini bildirmiştir. Bu da günde ortalama 7600'den fazla ölüme denk gelmektedir. ILO, küresel çapta yılda yaklaşık 374 milyon ölümcül olmayan işle ilgili yaralanma meydana geldiğini ve bunların küçük olaylardan ciddi sakatlıklara kadar deđiştiğini tahmin etmektedir (ILO, 2018). Bu tür yaralanmalar, etkilenen çalışanlar için acı, ıstırap ve uzun vadeli sađlık komplikasyonlarına neden olabilmektedir. İşle ilgili yaralanma ve hastalıkların ekonomik etkisi doğrudan maliyetler açısından ölçülebilir. İşle ilgili yaralanma ve hastalıkların işletmeler ve ekonomiler üzerindeki mali etkisi kayda deđerdir. Doğrudan maliyetler, tıbbi tedavi, rehabilitasyon ve etkilenen çalışanlara yapılan tazminat ödemeleri için yapılan harcamaları içerir. Küresel olarak bu maliyetler GSYH yaklaşık %3.94'üne, yani yıllık yaklaşık 3 trilyon dolara denk gelmektedir(Lo vd., 2021).

İş kazalarının dolaylı maliyetlerinin doğrudan maliyetlerden çok daha yüksek olduđu sıklıkla görülen bir durumdur. Bunlar arasında devamsızlık, işte var olmama çalışanların yaralanma veya hastalık nedeniyle iş yerinde bulunmaları ancak tam olarak üretken olmamaları ve verimliliğin azalması nedeniyle verimlilik kaybı yer almaktadır. Ayrıca, yedek işçilerin işe alınması ve eđitilmesi, ekipman veya mülkün zarar görmesi, yasal masraflar ve şirketin itibarının zedelenmesi ile ilgili ek maliyetler de söz konusudur. Dolaylı maliyetlerin doğrudan maliyetlerden 10 kata kadar daha yüksek olabileceği ve bunun da küresel olarak yaklaşık 30 trilyon dolarlık ek bir ekonomik yüke yol açtığı tahmin edilmektedir(Tompa vd., 2021a).

ABD İşgücü İstatistikleri Bürosu'na göre, yalnızca ABD'de işyeri yaralanmaları ve hastalıkları yılda yaklaşık 1.8 milyar kayıp iş gününe neden olmaktadır (BLS, 2023). Bu, işletmeler ve endüstriler için önemli bir üretkenlik kaybını temsil eder ve bu da son teslim tarihlerini karşılama, siparişleri yerine getirme ve pazarda rekabet etme kapasitelerini bozmaktadır (BLS, 2023).

İSG uzmanları işyeri ortamındaki potansiyel tehlikeleri sistematik olarak tanımlar. Bu tehlikeler fiziksel tehlikeleri (gürültü, titreşim veya ergonomik stres gibi), kimyasal tehlikeleri (zararlı maddelere maruz kalma gibi), biyolojik tehlikeleri (bulaşıcı ajanlar gibi) ve psikososyal tehlikeleri (stres, şiddet veya zorbalık gibi) içerebilir. Tehlikeler belirlendikten sonra, İSG uygulayıcıları her bir tehlikeyle ilişkili risk düzeyini değerlendirir. Bu, maruz kalma olasılığı, potansiyel zararın ciddiyeti ve etkilenen çalışan sayısı gibi faktörlerin değerlendirilmesini içerir (Bluff, 2003).

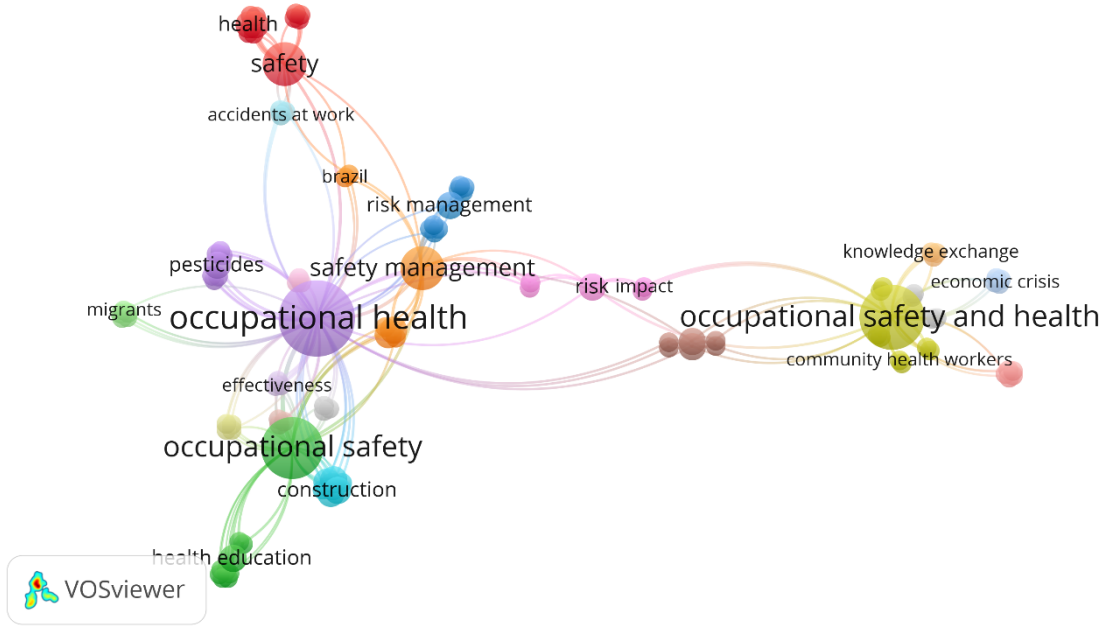
Risk değerlendirmesine dayanarak, işyeri tehlikelerini kontrol etmek veya azaltmak için önlemler uygulanır. Bu, mühendislik kontrollerini (ekipmanın yeniden tasarlanması veya havalandırma sistemlerinin uygulanması gibi), idari kontrolleri (güvenlik politikalarının geliştirilmesi veya eğitim sağlanması gibi) ve çalışanlar için KKD içerebilir.

İSG aynı zamanda çalışanların genel sağlığını ve refahını teşvik etmeyi amaçlayan girişimleri de kapsar. Bu, sağlıklı yaşam tarzlarının teşvik edilmesini, sağlık hizmetlerine erişim sağlanmasını ve kronik hastalıklara veya ruh sağlığı sorunlarına katkıda bulunan faktörlerin ele alınmasını içerebilir. İSG uzmanları, işyerlerinin devlet kurumları ve endüstri kuruluşları tarafından belirlenen ilgili sağlık ve güvenlik yönetmeliklerine ve standartlarına uymasını sağlar. Bu, uyumu izlemek ve gerektiğinde düzeltici eylemleri uygulamak için teftişler, denetimler ve soruşturmalar yürütmeyi içerir. İSG, güvenli ve sağlıklı çalışma ortamlarının yaratılmasında çok önemli bir rol oynar ve bu da nihayetinde çalışanların sağlığının ve refahının korunmasına ve işyerlerinin genel üretkenliğine ve sürdürülebilirliğine katkıda bulunur.

Güvenli olmayan davranışlar üzerine yapılan çalışmalarda iş kazasına neden olma durumunun eğitim seviyesi ile ilişkili olduğu ileri sürülmüştür (Abbasi vd., 2015; Adhikari, 2015). Bu durum gözetimin ve aktif uyarı sistemlerinin iş kazalarını sezerek önlemede çok önemli olduğunu göstermektedir.

Web of Science (WoS) tarafından indekslenmiş bilimsel çalışmalar arasında “İş sağlığı ve güvenliği” araması yapıldığında, veri tabanından alınan veriler üstünde yapılan bibliyografik incelemede bazı anahtar kelimeler için Şekil 1.1.'de görülen ilişkiler ortaya çıkmaktadır.

Burada görülen renklendirme birbiri ilişkili makalelere göre yapılmış sınıflandırmayı göstermektedir. Ayrıca hücrelerin büyüklüğü anahtar kelimenin makalenin özet bölümünde ve başlık bölümünde kullanım sıklığı ile ilişkilidir. Buradaki bibliyografik analiz için VOSviewer uygulaması kullanılmıştır (Van Eck & Waltman, 2020). Analiz sonucunda elde edilen grafikte renk farklılıklarından çalışma alanlarının başka bilim dallarına ayrıldığı da görülmektedir



**Şekil 1.1.** WoS üzerinde “occupational safety and health” için alınan geri dönütlerin VOSviewer ile gösterilmesi

## 1.2. Tezin Amacı ve Hedefleri

Bu tez çalışmasında, iş kazalarının azaltılması ve önlenmesi için, endüstriyel üretim yapılan işletmelerde ve daha genel anlamda insanın faaliyet gösterdiği alanlarda, meydana gelebilecek güvenli ve güvensiz hareketleri kapsayan bir veri seti oluşturularak, hibrit derin öğrenme ile gerçek zamanlı olarak güvensiz hareketlerin tespitini sağlayan bir bilgisayarla görme sistemi geliştirmenin yanında, önerilen sistemin sahada deneysel olarak uygulanması amaçlanmıştır. Tez çalışmasının hedefleri ise şunlardır:

- i. Tez çalışması kapsamında, Eskişehir Organize Sanayi bölgesinde faaliyet gösteren fabrikalardan elde edilen videolar ile yeni bir video veri seti oluşturulması

- ii. Çalışanların kişisel korunma ekipmanlarının (KKD) eksik olup olmadığı, tehlikeli sayılabilecek hareketleri yapıp yapmadıkları ve makine-teçhizat ile ilgili kritik çalışma kurallarına uyup uymadıklarının tespit edilmesi
- iii. Hazırlanan video veriseti üzerinde, önerilen bir derin öğrenme modeli kullanılarak davranışların içerik anlamlandırması
- iv. Video verisetinde güvenli ve güvensiz davranışların gerçek zamanlı sınıflandırılmasını sağlayan bir bilgisayarla görme sisteminin geliştirilmesi
- v. Görsel ve işitsel olarak çalışma ortamlarında çalışanların uyarılmasını sağlayan bir gömülü sistem tasarımı
- vi. Belli bir zaman diliminde çalışanları önerilen sistem tarafından izlenmesi ve güvensiz davranış eğilimlerinin değerlendirilmesi
- vii. Çalışma ortamlarından meydana gelen güvenli olmayan davranış sayılarındaki değişim eğilimlerinin sayısal olarak değerlendirilmesi
- viii. Ve nihai olarak iş kazalarının azaltılarak güvenli çalışma ortamlarının oluşturulmasına katkıda bulunulması

### **1.3. Araştırma Sorusu ve Hipotez**

İş kazalarını azaltma çabaları sanayi devriminin başlangıcından bu yana devam etmektedir. Ancak, sağlık ve çalışma örgütleri tarafından bildirildiği üzere, devam eden maddi ve manevi kayıpların da gösterdiğine göre, arzu edilen güvenlik seviyelerine ulaşılması zor olmaya devam etmektedir. Bu kayıplar ulusların gelişmişlik düzeylerine göre değişiklik gösterse de küresel olarak devam etmekte ve daha etkili önleyici tedbirlerin gerekliliğinin altını çizmektedir.

Bilgisayarla görme ve yapay zekâ alanındaki son gelişmeler, iş kazalarının azaltılması için elverişli yollar sunmaktadır. Derin öğrenme ağlarıyla güçlendirilmiş entegre uyarı ve bildirim sistemlerinin, çeşitli çalışma ortamlarında potansiyel olarak tehlikeli davranışları ve durumları ustalıkla azaltabileceği öngörülmektedir. Bu bağlamda nesne tespiti konusunda öncü olan YOLO algoritması ile aktivite tanıma çalışmalarında kullanılan ConvLSTM algoritmaları birleştirilerek oluşturulacak hibrit bir mimari ve bu çalışmaya entegre bir uyarı ve denetim sistemi, üretim ve çalışma ortamlarında tehlikeli hareketlerin azaltılması ve önlenmesi konusunda faydalı olabilir mi? sorusu bu tez çalışmasının motivasyonu olmuştur.

Çağdaş çalışma ortamlarının doğasında bulunan karmaşıklık ve dinamizm, video tabanlı güvenlik sistemleri için çeşitli çetin zorluklar ortaya çıkarmaktadır. Bu tür ortamlarda elde

edilen videoların, genellikle dağınıklık, oklüzyon ve değişken aydınlatma koşulları gibi faktörler nedeniyle anlaşılabilirliği azalmaktadır. Ayrıca, gerçek zamanlı video yorumlama zorunluluğu, maliyetli donanım ve hesaplama kaynaklarına olan talebi artırmaktadır. Bu zorlukların üstesinden gelmek için, derin öğrenme metodolojilerini geleneksel bilgisayarla görme teknikleriyle birleştiren hibrit bir sistem umut verici bir çözüm olarak ortaya çıkmaktadır. Üst düzey özellik çıkarma ve anlama için derin öğrenmeyi gerçek zamanlı işleme ve analiz için geleneksel tekniklerle entegre ederek, hibrit sistem daha verimli ve uygun maliyetli bir güvenlik çözümü sunabilir. Örneğin, derin öğrenme algoritmaları potansiyel tehlikeleri gösteren karmaşık kalıpları ve davranışları ayırt etmek için eğitilebilirken, geleneksel bilgisayarla görme teknikleri işçilere ve amirlere zamanında uyarılar ve bildirimler sağlamak için gerçek zamanlı video işlemeyi gerçekleştirebilir. Dahası, böyle bir hibrit yaklaşım farklı çalışma ortamlarına ve operasyonel zorunluluklara uyarlanabilirlik sergileyerek kaza önleme stratejilerinde daha fazla esneklik ve ölçeklenebilirlik sağlar. Bu yaklaşım, videonun anlaşılmasıyla ilgili karmaşıklıkları gidererek ve donanım gereksinimlerini azaltarak, gelişmiş güvenlik sistemlerini daha geniş bir endüstri ve kuruluş yelpazesi için daha erişilebilir hale getirmektedir.

Özetle, hibrit bir sistem çerçevesinde derin öğrenme ağlarının yeteneklerinden yararlanarak oluşturulacak bir sistem, iş kazalarını kayda değer ölçüde azaltabilir, böylece daha güvenli çalışma ortamları oluşturulabilir ve işletmeler ve genel olarak toplum için kayıplar azaltılabilir.

#### **1.4. Tez Çalışmasının Katkısı**

Tez çalışmasının birinci bölümünde üretim ortamlarında KKD kullanımının tespitine yönelik çalışmalar yapılmıştır. Bu çalışmaların katkısı şu şekilde sıralanabilir.

- i. Bu tez çalışmasında KKD'lerin kullanılıp kullanılmadığının tespiti için yüksek hızı ve doğruluğu ispatlanmış çok yeni bir model olan YOLO ile gerçek zamanlı birkaç versiyonu ile başarılı sonuçlar alınmıştır.
- ii. Bu çalışmalar için Eskişehir Organize Sanayi Bölgesinde (OSB) faaliyet gösteren iki fabrika ile görüşülmüş ve çalışmamız için gerekli izinler alınarak KKD kullanımına ilişkin, belirlenen Baret, Eldiven, Gözlük, Maske, Yelek, Baret Yok ve Maske Yok adlı yedi sınıfa ait veri seti oluşturulmuştur.
- iii. Bu çalışmalar sonucunda YOLOv4 algoritması ile gerçek zamanlı olarak yüksek doğruluk değerleri elde edilmiştir. Yapılan çalışmalar neticesinde 32 saniyede kare

sayısı (Frame Per Second, FPS) hızında %91.18'lik bir bütün sınıfların ortalama kesinliği mAP@50 değeri, 0.89'luk bir hassasiyet (Precision) değeri ve 0.91'lik bir hatırlama (Recall) değeri elde edilmiştir.

Tez çalışmasının ikinci bölümünde çalışma ortamlarında meydana gelebilecek güvensiz hareketlerin tespitine yönelik çalışmalar yapılmıştır. Bu aşamada sunulan katkıları ise aşağıdaki liste ile özetlenebilir:

- i. Tez çalışmasında işyerinde güvenli olmayan davranışların gerçek zamanlı sınıflandırılması için derin öğrenme modelleri kullanan hibrit bir bilgisayarla görme yaklaşımı olan Unsafe-Net önerilmiştir.
- ii. Unsafe-Net mimarisi için, öncelikle bir fabrikadan 39 günlük video görüntüleri alınarak özel olarak bir veri seti oluşturulmuştur.
- iii. Hazırlanan video veri seti kullanılarak, hızlı ve doğru sonuçlar elde etmek için nesne tespiti ve videodan aktivite yanıtı için YOLOv4 ve ConvLSTM yöntemleri birleştirilmiştir.
- iv. Önerilen Unsafe-Net mimarisi kullanarak güvenli olmayan davranışların sınıflandırma doğruluğu %95.81'dir ve videolardan eylem tanıma için ortalama süre 0.14 saniyedir.
- v. Unsafe-Net, ortalama video süresini 1.87 saniyeye düşürerek gerçek zamanlı tespit hızını artırmıştır.
- vi. Sistem fabrikada gerçek zamanlı bir çalışma ortamına kurulmuştur ve çalışanlar güvenli olmayan bir davranış meydana geldiğinde sistem tarafından hem sesli hem de görsel olarak hemen uyarılmaktadır.
- vii. Fabrikada yapılan pilot çalışmada, güvenli olmayan davranışların tekrarlanma oranının yaklaşık %75 oranında azaltıldığı tespit edilmiştir.

### **1.5. Tezin Organizasyonu**

Bu tez çalışması toplam 7 bölümden oluşmaktadır. Bu bölümlerden özetle bahsedecek olursak;

Bölüm 1'de iş yerlerinde, işletmelerde ve kuruluşlarda iş sağlığı ve güvenliği konusunda mevcut durumun tespiti ve geçmişe dönük olarak değerlendirilmesi, sorunların ele alınışı ve yeni gelişmeler ışığında çözüm önerilerinden ve tez çalışmasının amacından, hipotezimizden ve sunmuş olduğu katkılardan bahsedilmiştir.

Bölüm 2'de, iş kazalarının önlenmesi için literatürde yapılmış olan geleneksel görüntü işleme, makine öğrenmesi ve derin öğrenmeye dayalı çalışmalar kapsamlı olarak ele alınmıştır.

Ele alınan çalışmalar ve sonuçları incelenmiş, mevcut çalışmaların başarımları karşılaştırılmış ve sınırlılıkları ortaya koyulmuştur.

Bölüm 3'te Eskişehir OSB'sinde kurulu iki fabrikadan ve Bilecik Şeyh Edebali Üniversitesi Meslek Yüksekokulu makine atölyesinden KKD'lerin kullanımlarının tespitine yönelik olarak görüntülerin elde edilmesi, verisetinin oluşturulması, güvenli ve güvensiz davranışların sınıflandırılması için görüntü ve video verilerinin toplanması, hazırlanan video veriseti ve sınıfların belirlenmesine yönelik olarak tartışmaların ele alınması, kullanılan mimarilerin ve önerilen Unsafe-Net modelinin özellikleri ve yapısı sunulmuştur.

Endüstriyel üretim yapılan tesislerin çalışma ortamlarından toplanan video veriseti kullanılarak güvenli olmayan hareketlerin gerçek zamanlı sınıflandırılması için önerilen Unsafe-Net mimarisi açıklanmıştır. Ayrıca, nesne tespiti ve video anlamlandırma işlemleri için birleştirilen YOLOv4 ve ConvLSTM derin öğrenme modelleri üzerinde durulmuştur. Deneysel çalışmalarla ilgili olarak fabrikalarda tesis edilmiş kameralardan gelen görüntülerin önerilen hibrit yöntem ile anlık olarak denetlenmesi ve sahnenin anlamlandırılması anlatılmıştır. Devamında çalışma ortamında güvenli olmayan hareketlerin tespit edilerek, uzman müdahalesi için ön-işlemler oluşturulması, sahaya uygulanan uyarı ve denetim sisteminin yapısı çalışması ve etkileri ele alınmıştır.

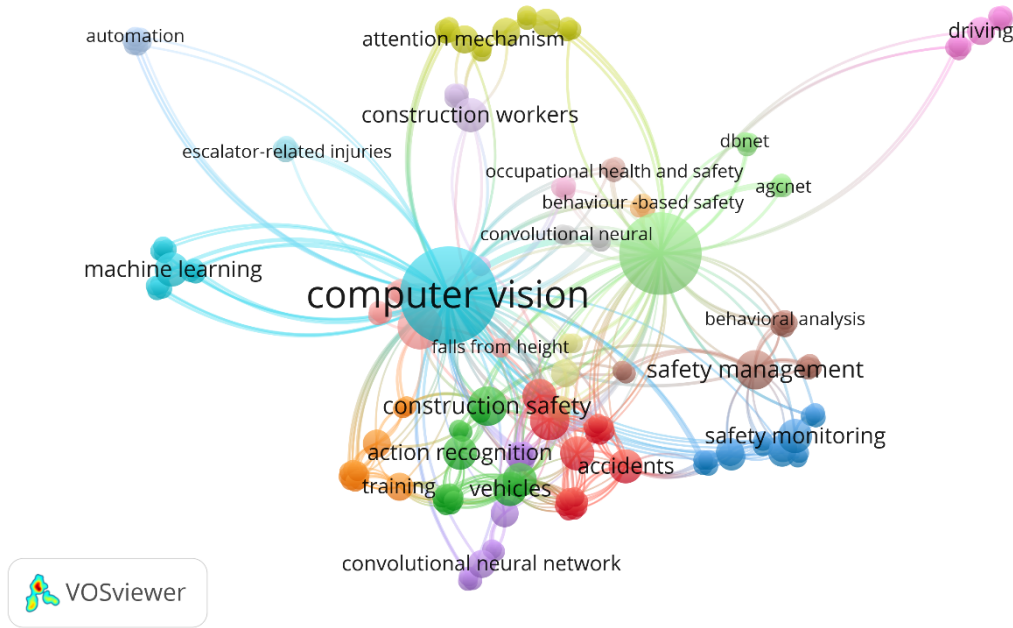
Bölüm 4'te Çalışma ortamlarında KKD tespitine yönelik yürütülen deneysel analizler, gerçek zamanlı olarak güvensiz davranışların tespiti için geliştirilen bilgisayarla görme sisteminde (Unsafe-Net) yürütülen deneysel analizler, YOLO mimarisi ile uygun çerçeve seçimi, YOLOv4+ConvLSTM ile sınıflandırma ve uyarı izleme sistemi üzerinde durulmuş deneysel sonuçlar ve çeşitli parametreler elde edilmiştir.

Bölüm 5'te Yapılan çalışmalar neticesinde elde edilen bulgular değerlendirilerek derin öğrenme ve bilgisayarla görme yöntemlerinin iş sağlığı ve güvenliği alanına olabilecek faydaları, bilime ve topluma sağlayabileceği katkıları tartışılmış ve sonuçları ele alınmıştır.

Bölüm 6'da çalışmada hedeflenen ve buna karşılık olarak elde edilen sonuç çıktıları ve yöntemin başarımları genel olarak değerlendirilmiştir. Daha başarılı sonuçlar elde edilebilmesi için öneriler sunularak tez sonlandırılmıştır.

## 2. LİTERATÜR TARAMASI

Web of Science (WoS) tarafından indekslenmiş bilimsel çalışmalar arasında “Bilgisayarla Görme ve İş sağlığı ve güvenliği” araması yapıldığında, veri tabanından alınan veriler üstünde yapılan bibliyografik incelemede bazı anahtar kelimeler için Şekil 2.1.’de görülen ilişkiler ortaya çıkmaktadır. Burada görülen renklendirme birbiri ilişkili makalelere göre yapılmış sınıflandırmayı göstermektedir. Ayrıca hücrelerin büyüklüğü anahtar kelimenin makalenin özet bölümünde ve başlık bölümünde kullanım sıklığı ile ilişkilidir. Buradaki bibliyografik analiz için VOSviewer uygulaması kullanılmıştır (Van Eck & Waltman, 2020). Analiz sonucunda elde edilen grafikte Bilgisayarla Görme anahtar kelimesi ile inşaat güvenliği kelimeleri arasındaki bağlantı şekillerdeki boyutlar büyüklüğü göze çarpmaktadır. Ayrıca görseldeki renk farklılıklarından çalışma alanlarının başka bilim dallarına ayrıldığı da görülmektedir.



Şekil 2.1. WoS üzerinde “Bilgisayarla Görme ve İş sağlığı ve güvenliği” için alınan geri dönütlerin VOSviewer ile gösterilmesi

Güvenlik kamerası görüntülerinin anlaşılmasındaki en önemli zorluklardan biri, karmaşık çalışma ortamlarındaki nesnelere tanımlama ve bu ortamdaki tehlikeli davranışları anlama becerisidir. Video anlama, yapay zekâ, gözetim, video arama ve güvenlik dahil olmak üzere bir dizi amaç için oldukça faydalıdır. Video görüntülerini girdi olarak alabilen ve videoda yer alan kişiler tarafından gerçekleştirilen herhangi bir faaliyeti tespit edebilen bir model tasarlamak çok önemlidir. Bu modeller genellikle nesne tanıma, izleme, anormallik tespiti,

aktivite anlama ve video yorumlama alanlarında yaygın uygulamaları olan bir teknik olan bilgisayarla görme kullanılarak oluşturulur. Bilgisayarla görme, derin öğrenme tekniklerini kullanarak nesnelere tespit etmek, hareketleri izlemek, anormallikleri tanımlamak, faaliyetleri tespit etmek ve videoları yorumlamak için yaygın olarak kullanılmaktadır.

Gözetim videoları, robotik, riskli davranış tespiti, video anlama, anormal durum tespiti gibi uygulamalar, genellikle, insanlar veya araçlar gibi videodaki farklı aktörlerin hareketine dayalı eylemlerin ve etkinliklerin tanınmasını gerektirir. Belirli uygulamalar, önceden tanımlanmış sınıflardan birine etkinlik atamayı gerektirebilir, diğerleri ise anormal veya seyrek görülen olağandışı faaliyetlerin saptanmasına odaklanabilir. Bu görev, birçok nedenden dolayı, statik görüntülerde nesne tanıma gibi daha geleneksel görevlerden doğası gereği daha zorludur. Aktivite tanıma, renk ve doku bilgisine ek olarak uzay-zaman segmentasyonu ve videodan hareket bilgisinin çıkarılmasını gerektirir. İkincisi, statik sahnelerdeki nesne görünüşleri, bakış açısı, oklüzyon, aydınlatma gibi görüntüleme koşulları altında da değişiklik gösterirken, kamera hareketi, eylem uzunluğu, özne görünümü ve stili de dikkate alınması gerektiğinden, insan eylemlerinin zamansal bileşenindeki değişkenlik daha da fazladır(Wolf vd., 2014).

Video içeriğinin anlaşılması ile ilgili yapılan ilk çalışmalar insanların yaptığı el sallama, yürüme gibi basit hareketlerin tespit edilebilmesi üzerine yapılmıştır (Poppe, 2010). Daha yakın zamanlarda, araştırmalar, nesnelere etkileşimler ve/veya birden çok insan ve birden çok etkinlik içeren karmaşık etkinlikler içeren daha gerçekçi ve dolayısıyla zorlayıcı problemlere odaklanmıştır (Wolf vd., 2014).

Tez çalışması konusunda literatürde bulunan mevcut çalışmalar çoğunlukla endüstriyel üretim yapılan alanlardan ziyade daha karmaşık alanlara sahip olmayan inşaat sahalarındaki güvensiz davranışları tespit etmeye, nedenlerini anlamaya ve azaltmaya odaklanmıştır. Son yıllarda yapılan çalışmalar yöntem olarak sınıflandırılacak olursa genel olarak geleneksel yaklaşımlarla önerilen yöntemler, makine öğrenmesi tabanlı yaklaşımlar ve derin öğrenme tabanlı çalışmalar olmak üzere üç alt başlık altında incelenebilir.

## **2.1. Geleneksel Yaklaşımlarla Önerilen Yöntemler**

Geleneksel güvenlik önlemlerine uyulmasına rağmen, güvensiz davranışlar nedeniyle işyerinde kazalar, yaralanmalar ve ölümler her zaman önlenememektedir. İşyerleri kapalı ortamlar olduğu için özel ekipman ve prosedürler gerektirir. Endüstriyel tesislerde, iş yerlerinde ve üretim sahalarında tehlikeli davranışların tahmin edilmesi kazaların önlenmesine ve azaltılmasına yardımcı olabilir.

Barro-Torres ve arkadaşları (Barro-Torres vd., 2012) Zigbee ve Radyo Frekans Tanıma Sistemi (RFID) teknolojilerini kullanarak KKD'ler için gerçek zamanlı bir izleme sistemi önermiştir. Yaptıkları çalışmada, inşaat sektöründe çalışanların KKD'leri doğru şekilde giymelerini sağlayan yeni bir siber-fiziksel sistem tanıtılmaktadır. Sistem işçilerin KKD kullanımını gerçek zamanlı olarak izleyerek doğru şekilde giymelerini sağlamayı hedeflemektedir. Bunun için kablosuz yerel alan ağı ve vücut alanı ağı kullanılarak bir mimari düşünülmüştür. Zigbee ve RFID teknolojileri kullanılarak bir prototip geliştirilmiş ve işçilerin bu prototipi taşıdıkları mikrodenetleyici tabanlı bir cihazla kullanmaları önerilmiştir. Bu cihaz, KKD'lerin varlığını tespit eder ve merkezi bir birime rapor gönderir. Makale, sistemi tanıtmak için izleme sistemi, donanım ve yazılım bileşenlerini açıklamakta ve işçinin cihazının kapsamı ve tüketimini analiz etmeyi amaçlamaktadır. Bu sistemin, inşaat sektöründe çalışanlar için KKD kullanımını daha etkin bir şekilde sağlayabileceği ve güvenlik standartlarını artırabileceği belirtilmiştir. Belirtilen sistemin gözleme dayalı olmadığı için, işçilerin KKD'leri doğru kullanıp kullanmadıklarının tespit edilememesi gibi bir sorunu vardır.

Ayrıca, Luo ve arkadaşları (Luo vd., 2014) çalışma ortamlarındaki güvensiz davranışları ve durumları tespit etmek için kablosuz sensör ağları ve çeşitli ekipmanlar kullanmışlardır. Çalışmada konum verilerinin kullanımıyla otonom vinç güvenliği izleme sistemlerinin geliştirilmesindeki çabaların arttığı belirtilmektedir. Ancak Küresel Konumlama Sistemi, (Global Positioning System, GPS), Radyo Frekansı ile Tanımlama (Radio Frequency Identification, RFID) ve ultra geniş bant sensörler gibi konum farkındalığı teknolojileri ile elde edilen verilerin kusurlu olduğu ve bu durumun sistem performanslarını olumsuz etkilediğini belirtmişler ve yaptıkları çalışmada, veri kusurlarını ele alarak sistem performansını iyileştirmek için farklı hesaplama yaklaşımları incelemişlerdir. Yazarlar, ölçekli bir otonom vinç güvenliği izleme test yatağı inşa etmiş ve beş temsili test durumu geliştirmiştir. 38 test noktasında 760 konum okuması gerçekleştirilmiştir. Bu konum verileri, beş farklı yaklaşım kullanılarak değerlendirilmiştir. Sonuçlar, sistem performansının daha iyi olabilmesi için sensörlerden en az on konum okumasının toplanabileceğini göstermektedir. Ayrıca, lokalizasyon hatasını azaltmak için ek verilerin (hız, ivme gibi) kullanılması da önerilmektedir. Bu yaklaşım, kusurlu verileri ek girdilerle birleştirerek yerelleştirme doğruluğunu artırmanın bir yolunu sunmaktadır. Ancak sistem maliyetleri ve mobilite sorunları vardır.

2014 yılında yapılan başka bir çalışmada Ann ve arkadaşları (Ann & Theng, 2014) aktive tanıma yönelik 32 tane makale incelenmiş; buna göre aktivite tanımanın yaygın olarak havaalanları, metro gibi halka açık alanlarda veya hasta ve yaşlı bakım yerlerin de gözetim

videoları üzerinden gözleme dayalı yapıldığını veya sensör temelli düşmelere vs. yönelik olarak kullanıldığı veya giyilebilir teknolojiler kullanılarak insan aktivitelerinin anlaşılabilceğini bildirmişlerdir.

Guo ve arkadaşlarının (Guo vd., 2015) 2015 yılında yapmış oldukları bir çalışmada Wuhan da bulunan bir metro inşaatında güvenliği artırmak için büyük veri tabanlı davranış gözlemi çerçevesini önermişlerdir. Çalışmada metro inşaatının dar alanı ve güvensiz işçi davranışları nedeniyle sık sık kazalara yol açabileceği belirtilmiştir. Davranış temelli güvenlik (BBS) yöntemi, işçi davranışlarını gözlemlemeyi, analiz etmeyi ve değiştirmeyi amaçlamıştır. Çerçeve, davranışsal risklere ilişkin bir bilgi tabanı oluşturmayı, işçilerin güvensiz davranışlarını yansıtan görüntü verilerini toplamak için akıllı bir video gözetim sistemi ve mobil uygulama kullanmayı ve verileri depolamak için bir büyük veri bulut platformu oluşturmayı içermektedir. Bu çerçevenin bir örnekle kullanılabilirliği doğrulanan ve metro inşaatında davranış gözlem verileri elde edildiği belirtilmektedir. Belirtilen çalışmada, gerçek zamanlı yaklaşımlar yerine kaza görüntülerini kullanarak işçileri eğitmek için istatistiksel temelli yaklaşımlar önerilmiştir. İstatistiksel bir çalışma ile soruna çözüm aranmıştır.

Seo ve arkadaşları (Seo vd., 2015) yaptıkları çalışmada, bilgisayarla görme teknolojilerinin alanlarda potansiyel çözümler sunmakta olduğunu, saha gözlemlerinde bilgisayarla görme yöntemlerinin uygulanmasının, mevcut manuel (el yordamı) uygulamaları tamamlayarak görüntülerden ve videolardan güvenlikle ilgili verileri çıkarmak için etkili bir otomatik araç olduğuna belirtmişlerdir. Yaptıkları çalışmada bilgisayarla görme teknikleri ile inşaat sektörü ile ilgili önceki dönemlerde yapılmış çalışmaları teknik ve pratik açılarından değerlendirerek mevcut durumu ortaya koymaktadır. Çalışmada, nesne algılama, nesne izleme ve eylem tanıma gibi alanlarda yapılan çalışmalar ele alınmaktadır. Bunun yanı sıra, sahne anlayışı, izleme doğruluğu ve gizlilik endişeleri gibi önemli zorluklar ve pratik sorunlar tartışılmaktadır.

Ayrıca, Yu ve ekibi (Yu vd., 2017) inşaat işçilerinin hareket davranışlarını belirlemek için görüntü iskeletine ve gerçek zamanlı analize dayalı bir yöntem geliştirmiştir. Bu çalışma inşaat sektöründe işçi güvenliğini sağlamak için KKD kullanımının önemini vurgulamaktadır. Ancak, güvensiz davranışlar inşaat kazalarının ana nedeni olmaktadır. Mevcut yönetim yaklaşımı olan BBS, manuel gözlem ve kayıtlara dayanmaktadır ve maliyetli ve zaman alıcı olabilmektedir. Bu nedenle, gerçek zamanlı olarak güvensiz davranışları belirlemek için görüntü-iskelet tabanlı parametrelendirilmiş bir yaklaşım önerilmektedir. Bu yaklaşımın geçerliliğini doğrulamak için deneyler yapılması gerekmektedir. Bu araştırma, yöntemin

uygulanabilirlik ve doğruluğunu incelemek amacıyla farklı güvensiz davranış türlerini içeren bir dizi deney düzenlemekte ve sonuçları göstermektedir. Ayrıca, ekipmanların doğru şekilde giyilmesi için yeni bir siber-fiziksel sistem (CPS) sunulmaktadır. Bu sistem, kablosuz ağ ve RFID teknolojileriyle KKD kullanımını gerçek zamanlı olarak izlemeyi sağlamaktadır. Yapılan çalışmada, izleme sisteminin detaylarını açıklanmakta, işçinin cihazının kapsam ve tüketimini analiz etmeyi amaçlamakta, yöntemin etkinliğini ve verimliliğini artırmayı hedeflemektedir. Yöntem, işçilerin hareketlerini kaydetmek için Kinect ve kızılötesi dedektörleri kullanmıştır. Bununla birlikte, bu tür sistemlerle ilgili sınırlamaların, ekipmanın özel üretim ve ön işleme ihtiyacını içerdiği unutulmamalıdır.

## **2.2. Makine Öğrenmesi Tabanlı Çalışmalar**

İşyerlerindeki güvensiz davranışları tespit etmek için makine öğrenimi tekniklerini kullanan çalışmalar da yapılmıştır. Örneğin, Matias ve arkadaşları (Matias vd., 2008) 2008 yılında yaptıkları bir çalışmada, işyeri kazalarının nedenleri ve türlerinin analizi için bir metodoloji önermektedir yaptıkları çalışmada özellikle kat seviyesindeki düşmelere odaklanmışlardır. Yaklaşım, Bayes ağları, sınıflandırma ağaçları, destek vektör makineleri ve aşırı öğrenme makineleri olmak üzere çeşitli makine öğrenme tekniklerine dayanmaktadır. Farklı teknikler kullanılarak elde edilen sonuçlar, her ikisi de risk önleme önlemlerinin geliştirilmesini kolaylaştıran açıklayıcı kapasite ve öngörü potansiyeli açısından karşılaştırılmıştır. Bayes ağlarının, güçlü bir yorumlayıcı kapasiteyi, mevcut en iyi tekniklerle karşılaştırılabilir bir öngörücü kapasiteyle birleştirdikleri için bu tür çalışmalar için en iyi genel teknik olduğu ortaya çıkmıştır.

Park ve Brilakis (Park & Brilakis, 2012), 2012 yılında yaptıkları çalışmada, şantiyelerdeki kamera görüntülerinde inşaat varlıklarını tanımak ve konumlarını belirlemek için görsel tabanlı nesne tespiti önermişlerdir. Görsel tabanlı nesne tespitinin, özellikle büyük ölçekli, sıkışık şantiyelerde birçok varlığın mekânsal konumlarını sağlaması açısından faydalı olduğunu ancak, yanlış nesne tespitinin yöntemi pratikte uygulanmasını engelleyebileceğini belirtmişlerdir. Bu sorunu çözmek amacıyla, çalışmalarında, tespit ve izleme algoritmalarının işlevlerini birleştiren yenilikçi bir hibrit yöntem önermişlerdir. Bu yöntemin, varlıkların hareket, şekil ve renk dağılımından yararlanarak video görüntülerinden inşaat ekipmanlarını tespit edeceğini ve bunun içinde sırasıyla Arka Plan Çıkarma (Background subtraction), Haar Benzeri Özellikler (Haar-like features) ve Öz-Görüntüler (eigen-images) gibi yöntemler kullandıklarını belirtmişlerdir. Yanlış tespitleri telafi etmek için bir izleme algoritması sürece dâhil edilmiştir. Yanlış tespitler, nesne boyutu ve görünümündeki ani değişiklikler yakalanarak

belirlenmiştir. Belirlenen yanlış tespitler izleme sonuçlarıyla değiştirilmiştir. Ön çalışmalar, izleme ile kombinasyonun tespit performansını artırma potansiyeline sahip olduğunu göstermiştir.

Alwasel ve diğerleri (Alwasel vd., 2017) güvenli ve üretken duvar ustalarının pozlarını tanımak için makine öğrenimini kullanan bir metodoloji önermişlerdir. Çalışmada, atalet ölçüm birimleri ve video kameralar kullanarak ustalardan kinematik veri toplanarak poz kümeleri belirlenmiştir. Çalışmada, beton duvar birimleri kullanılarak standart bir duvarın inşası sırasında duvar ustası grupları arasında iş pozlarının sınıflandırılması için bir çerçeve sunulmuştur. Uzman ve uzman olmayan duvar ustalarının pozları, kinematik verilerin toplanmasıyla analiz edilmiştir. Destek Vektör Makinesi (DVM) algoritması kullanılarak pozlar uzman ve uzman olmayan sınıflara ayrılmıştır. İki sınıflandırma senaryosu test edilmiştir ve her ikisi de benzer doğruluk seviyelerine ulaşmıştır. Çalışma, duvar ustalarının daha az ve daha güvenli duruşlar kullandığını ve bu duruşların daha yüksek üretkenlik ve daha düşük enerji harcaması ile sonuçlandığını göstermektedir.

Bir başka çalışmada, Wu ve Zhao (Wu & Zhao, 2018) sadece işçilerin baret takıp takmadıklarını ve baretlerin renklerini tespit etmeye odaklanmıştır. Çalışmada verileri sınıflandırmak için hiyerarşik bir DVM geliştirilmiş ve belirli bir doğruluk seviyesine ulaşılmıştır. Düzenlemelere uyulmasını sağlamak amacıyla, işçilerin baret takıp takmadığını izlemek için video gözetim sistemleri kullanılmıştır. Bununla birlikte, çok sayıda gözetim ekranıyla, herhangi bir ihlali herhangi bir zamanda tespit etmek zor olabilir ve bu da ciddi kazalara yol açabileceği ifade edilmiştir. Çalışma, görüntü tanıma teknolojilerinin hızla gelişmesi nedeniyle çok önemli bir endüstriyel uygulama alanı olan baret tanımlama için akıllı görüş tabanlı bir yaklaşım önermektedir. Yaklaşım, işçilerin baret takıp takmadığını izlemekte ve baretlerin renklerini tanımlamaktadır. Farklı renklerdeki (kırmızı, sarı ve mavi) baretlerden özellikler çıkarmak için yerel ikili örüntüler (LBP), hu moment değişmezleri (HMI) ve renk histogramlarına (CH) dayalı hibrit bir tanımlayıcı önerilmiştir. Çıkarılan özellikler daha sonra hiyerarşik bir destek vektör makinesi (H-SVM) kullanılarak dört sınıfa (kırmızı kask, sarı kask, mavi kask ve kask olmayan) ayrılmıştır. Önerilen yaklaşım bir veri kümesi üzerinde test edilmiş ve ortalama kask tanımlama doğruluğu %90.3 olarak bulunmuştur.

Mneymneh ve arkadaşları (Mneymneh vd., 2017), 2017 yılında yaptıkları bir çalışmada, kaskları tespit etmek için yararlı olduğu düşünülen bilgisayarla görme tekniklerini üzerine çalışmışlardır. Çalışmada, özellik algılama, çıkarma ve eşleştirme, şablon eşleştirme ve kademeli sınıflandırıcı modellerinin nesne algılama/tanıma konusunda umut verici olduğunu

belirtmişlerdir. Ayrıca, çalışmada, Hızlandırılmış Sağlam Özellikler (SURF) veya İkili Sağlam Değişmez Ölçeklenebilir Anahtar Noktalar (BRISK) ve Hızlandırılmış Segment Testi Özellikleri (FAST) gibi gradyan tabanlı özellikler, giriş görüntüsü (ve bir kask içeren bir referans görüntü) arasındaki nokta ilişkilerini tespit etmek için kullanılmış ve referans görüntüyle en iyi eşleşen kask tespit edilmiştir. Kask sayısı, hedef görüntüden tanımlanan kasklar gizlenerek hesaplanmış, böylece algoritmanın bir sonraki yinelemesinde bir sonraki en iyi eşleşen kask tespit edilebilmiştir. Çalışmanın devamında şablon eşleştirme ve Yönlendirilmiş Gradyanların Histogramı (HOG), Haar-like ve Yerel İkili Desen (LBP) gibi basamaklı sınıflandırıcılar kullanılarak nesne tespiti üzerinde durulmuştur.

### **2.3. Derin Öğrenme Tabanlı Çalışmalar**

Son yıllarda, görüntüler üzerinden derin öğrenme, nesne tespiti ve video anlamlandırma çalışmaları çok revaçta olan konulardır. Grafik İşleme Birimi (Graphical Processing Unit, GPU) teknolojilerinde ve veri setlerindeki meydana gelen yeni gelişmeler bu teknolojilerdeki fiyat düşüşleri, konuya olan ilgiyi çok daha fazla artırmıştır (Schmidhuber, 2015). Bu çalışmalarla ilgili her gün yeni mimariler yayınlanmakta ve birçok bilim adamı bu konularla ilgilenmektedir. Dünyanın en büyük şirketleri bu alanlara yatırım yapmakta ve çalışmalar büyük ölçekli katkılar sunmaktadırlar. Bu bağlamda geleneksel insan sağlığı iş güvenliği uygulamaları bu çalışmalardan doğal olarak etkilenmektedir.

Video anlamlandırmaya yönelik olarak Donahue ve arkadaşları (Donahue vd., 2015), Long-Term Recurrent Convolutional Networks (LRCN) kullanarak önerdikleri çalışmada, uçtan uca eğitilebilir, büyük ölçekli görsel öğrenmeye uygun yeni bir tekrarlayan evrişim mimarisi geliştirmişlerdir ve bu modellerin karşılaştırmalı video tanıma görevleri, görüntü tanımlama ve alma sorunları ve video anlatım zorlukları üzerindeki değerini belirtmişler. Eğitim sırasında videodan 16 kare klip örneklenmiştir. Mimari, Kırmızı-Yeşil-Mavi (RGB) girişi veya 16 kare klipten oluşan optik akış ile uçtan uca eğitilmiştir.

Tran ve diğerleri (Tran vd., 2015) büyük ölçekli bir denetimli video veri kümesi üzerinde eğitilmiş derin 3 boyutlu konvolüsyon ağları (3B ConvNets, C3D) kullanarak uzay-zamansal özellik öğrenimi için yeni bir yaklaşım önermişler. Bu yaklaşımın iki boyutlu (2B) CNN yapılarına göre başarımını daha yüksek olduğunu belirlemişlerdir. Eğitim sırasında, her video için, tüm videoda bildirilen eylem olarak kesin gerçeği olan beş rastgele 2 saniyelik klip çıkartılmıştır. Test aşamasında, 10 klip rastgele örneklenerek ve son tahmin için bunlar arasındaki tahminlerin ortalamasını hesaplamışlardır.

Yao ve arkadaşlarının (Yao vd., 2015) yaptığı çalışma doğrudan eylem tanıma ile ilgili olmasa da video gösterimleri açısından önemli bir çalışmadır. Bu çalışmada yazarlar, video açıklama görevi için temel mimari olarak bir 3B-CNN + LSTM kullanmışlardır. Yazarlar daha iyi sonuçlar için önceden eğitilmiş bir 3B-CNN kullanmışlardır. Bu yöntemde, iki farklı LRCN mimarisinden ayrılmaktadır. Özellikleri 3B CNN'den LSTM'e geçirmek yerine, klip için 3B CNN özellik haritaları, her kare için gösterimi zenginleştirmek için aynı kare seti için yığılmış 2B özellik haritalarıyla birleştirilmiştir. Kullanılan 2B ve 3B CNN, önceden eğitilmiş bir CNN'dir ve LRCN gibi uçtan uca eğitilmemiştir. Tüm çerçevelerde zamansal vektörlerin ortalamasını almak yerine, zamansal özellikleri birleştirmek için ağırlıklı bir ortalama kullanılmıştır.

Feichtenhofer ve arkadaşları (Feichtenhofer vd., 2016), çalışmalarında, iki yeni yaklaşımla temel iki akış mimarisini (Two Stream Fusion) kullanmışlardır ve parametrelerin boyutunda önemli bir artış olmadan performans artışı olduğunu göstermişlerdir. Yazarlar iki ana fikrin etkinliğini araştırmışlardır. Uzamsal ve zamansal akışların nasıl ve ne zaman birleştirilmesi gerektiği üzerinde durmuşlar ve erken seviyede birleşmesi gerekliliğini bildirmişlerdir. Uzun dönemli bağımlılığın da modellenmesi için zaman çerçevelerinde geçici net çıktıyı birleştirmişlerdir.

Wang ve arkadaşlarının (Wang vd., 2016) çalışmalarında, aktivite tanımda küçük boyutlar ve uzun menzilli modelleme gibi iki büyük zorluğun üstesinden gelmeye çalışmışlardır. Yapılan çalışmada, en gelişmiş sonuçları üretmek için iki akış mimarisini geliştirmişlerdir. Tüm video boyunca rastgele örnekleme yerine uzun menzilli zamansal sinyali daha iyi modellemek için video boyunca seyrek örnekleme klipleri önermişlerdir. Ayrıca optik akışa alternatif olarak iki yeni giriş yöntemini yani çarpık optik akış ve RGB farkı önermişlerdir. Çalışma, yeni Kinetics Human Action Video veri kümesinin ışığında gelişmiş mimarileri yeniden değerlendirmektedir. Kinetik, 400 insan eylemi sınıfı ve sınıf başına 400'den fazla klip ile iki kat daha fazla veriye sahiptir ve gerçekçi, zorlu YouTube videolarından toplanmıştır. Mevcut mimarilerin bu veri kümesindeki eylem sınıflandırma görevinde nasıl çalıştığına yönelik ve Kinetik üzerine ön eğitimden sonra daha küçük kıyaslama veri kümelerinde performansın ne kadar arttığına dair bir analiz sunulmuştur.

Carreira ve arkadaşları (Carreira & Zisserman, 2017), 2B ConvNet çalışmasını temel alan yeni bir iki Akışlı 3B ConvNet (I3D) önermişlerdir. Derin görüntü sınıflandırmasının filtreleri ve havuzlama çekirdekleri ConvNet'ler 3B'e genişletilerek, videodan kesintisiz uzay-zamansal özellik çıkarıcıları öğrenmeyi kullanmışlardır. ImageNet mimari tasarımlarından ve

parametrelerinden yararlanırken Kinetik üzerine ön eğitimden sonra, I3D modellerinin, eylem sınıflandırmasında teknolojiyi önemli ölçüde geliştirdiğini göstermişlerdir.

Diba ve arkadaşları (Diba vd., 2017), videolarda bulunan zamansal ipuçlarının doğru sınıflandırmaları ve özellikle insan eylemi tanınması için nasıl kullanılacağı üzerinde durmuşlardır. Çalışmada, değişken zamansal evrişim çekirdek derinliklerini modelleyen yeni bir zamansal katman sunmuşlardır. Bu yeni geçici katmanı, önerilen 3B CNN yerleştirilmiştir. Normalde 2B olan DenseNet mimarisini 3B filtreler ve havuz çekirdekleriyle genişletilmiş ve önerilen video evrişim ağına 'Temporal 3B ConvNet' (T3D) ve onun yeni zamansal katmanına 'Temporal Transition Layer' (TTL) adını vermişlerdir. T3D'nin HMDB51, UCF101 ve Kinetix üzerindeki mevcut gelişmiş yöntemlerden daha iyi performans gösterdiğini belirtmişlerdir.

Girdhar ve arkadaşlarının (Girdhar vd., 2017) çalışmalarında, videonun tüm uzamsal-zamansal kapsamı boyunca yerel konvolüsyon özellikleri toplayan eylem sınıflandırması için yeni bir video gösterimi önermişlerdir. İki akışlı ağları, öğrenilebilir uzamsal-zamansal özellik toplama algoritması ile birleştirmişler ve bu sayede algoritmayı tüm video sınıflandırması için uçtan uca eğitilebilir hale getirmişlerdir. Uzay ve zaman arasında bir havuz oluşturmak ve farklı akışlardan gelen sinyalleri birleştirmek için stratejiler önermişler.

İş ortamlarındaki güvensiz davranışları tespit etmek için daha önce çok sayıda teknik tanıtılmıştır, ancak derin öğrenmenin kapsamlı bir şekilde benimsenmesinden sonra, bu yöntemlerin tanımlama doğruluğu ve hızı önemli ölçüde artmıştır. Fang ve arkadaşları (Q. Fang, H. Li, X. Luo, L. Ding, H. Luo, vd., 2018), kuleler ve bacalar gibi yüksek yapılarda çalışanların kişisel koruyucu ekipmanlarını düşme tehlikelerine karşı otomatik olarak incelemek için derin öğrenme tabanlı bir yöntem önermiştir. Bu yöntem dış duvarların yanında havadan çalışmaya hazır olan çımacıların KKD kullanımını kontrol etmek için otomatik bir denetim yöntemi önermektedir. Derin öğrenmeye dayalı KKD kontrolü için bir oklüzyon azaltma yöntemi sunulmuştur. Yöntemin performansı çeşitli koşullar altında test edilmiş ve deneysel sonuçlar, yöntemin düşmeyi önleyici tedbirlerin denetlenmesindeki güvenilirliğini ve sağlamlığını ortaya koyduğu belirtilmiştir.

Bu arada, Ding ve arkadaşları (Ding vd., 2018) güvensiz davranışları tespit etmek için bilgisayarla görme ve örüntü tanıma tekniklerini kullanan hibrit bir derin öğrenme yaklaşımı sunmuştur. Çalışma, konvolüsyon sinir ağları CNN ve uzun kısa süreli bellek (LSTM) kullanarak bir model geliştirmiştir. İşçilerin güvenli olmayan eylemlerini otomatik olarak tanımak için bir konvolüsyon sinir ağı ve LSTM entegre edilmiştir. Önerilen model, güvenli

olmayan eylemleri tanımlamak, hareket verilerini ve saha videolarını toplamak, bir CNN modeli kullanarak videolardan görsel özellikler çıkarmak ve LSTM modellerinin kullanımıyla sağlanan öğrenme özelliklerini sıralamayı amaçlamaktadır. Modelin güvenli olmayan eylemleri tespit etme kabiliyetini test etmek için bir deney gerçekleştirilmiştir. Sonuçlar, geliştirilen hibrit modelin (CNN + LSTM) işçiler tarafından sahada gerçekleştirilen güvenli/güvensiz eylemleri doğru bir şekilde tespit ettiğini ortaya koymaktadır. Modelin doğruluğu, görüntülerdeki ilgi noktalarını tespit etmek için mevcut son teknoloji ürünü tanımlayıcı tabanlı yöntemlerin doğruluğunu aştığı belirtilmiştir.

Ayrıca, Fang ve diğerleri (W. Fang vd., 2018) sahalardaki işçileri ve ekipmanları otomatik olarak tespit etmek için daha hızlı bir R-CNN tabanlı derin öğrenme modeli önermiştir. Çalışmada kapsamlı bir nesne veri tabanı oluşturulmuş ve sonuçlar en yeni metodolojilere karşı değerlendirilmiştir. Önceki çalışmalar, şantiyelerden nesnelere ve malzemelere almak için renk tabanlı yöntemler gibi içerik tabanlı görüntü alma (CBIR) yaklaşımları geliştirmeye odaklanmıştır. CBIR yaklaşımları dijital görüntülerden düşük seviyeli özelliklerin çıkarılmasında etkili olsa da mevcut veri tabanlarından güvenli olmayan davranışları doğru bir şekilde alamamaktadır. Bu sınırlamayı ele almak için, güvensiz davranış görüntülerini daha doğru ve otomatik olarak almak için farklı modellerden çıkarılan özellikleri birleştiren gelişmiş bir CBIR yaklaşımı geliştirilmiştir. Önerilen yaklaşımın uygulanabilirliğini doğrulamak için Huazhong Bilim ve Teknoloji Üniversitesi tarafından geliştirilen dijital bir veri tabanı kullanılmıştır. Araştırma, ResNet-101 ve VGG-19 füzyonunun daha yüksek Top-K geri çağırma seviyelerine ulaştığını ve tek özellik çıkarma yönteminden daha iyi performans gösterdiğini ortaya koymaktadır.

Zhu ve arkadaşları (Zhu vd., 2019), peş peşe gelen çerçeve arasındaki hareket bilgilerini dolaylı olarak yakalayan yeni bir CNN mimarisi önermişler ve bu mimariye gizli iki akışlı CNN olarak adlandırmışlardır. Çalışmada uçtan uca yaklaşımlarının, iki aşamalı temel çalışmadan 10 kat daha hızlı olduğunu belirtmişlerdir. Dört zorlu eylem tanıma veri kümesi üzerindeki deneysel sonuçlar: UCF101, HMDB51, THUMOS14 ve ActivityNetv1.2, çalışmalarının önceki en iyi gerçek zamanlı yaklaşımlardan önemli ölçüde daha iyi performans gösterdiğini bildirmişlerdir.

Son ve diğerleri (Son vd., 2019) farklı poz ve arka planlardaki inşaat işçilerini tespit etmek için derin artık ağları ve daha hızlı R-CNN derin öğrenme mimarilerini kullanan bir model önermiştir. İnşaat işçilerinin konumunu ve davranışlarını belirlemek için şantiye görüntülerinin analizi, güvenlik yönetimi ve verimlilik analizi için yararlı bir araç olarak

tanımlanmıştır. Çalışmada inşaat işçisi tespiti için kullanılan CNN tabanlı mimari, çok derin artık ağlar (ResNet-152) kullanarak özellik haritası çıkarma ve CNN Özelliği Daha Hızlı Bölgeler (R-CNN) kullanarak orijinal görüntüden sınırlayıcı kutu regresyonu ve etiketleme içermektedir. Sonuçlar, işçilerin duruşları, görünüşleri ve arka planları hakkında sınırlı varsayımlara dayanmadan çok derin artık ağlar kullanarak şantiye görüntülerinde birden fazla işçiyi doğru ve hızlı bir şekilde tespit etme potansiyelini göstermektedir. Ancak hesaplama maliyeti yüksek olduğundan belirtilen sistem 5 FPS in üzerine çıkamamıştır.

Hu ve arkadaşları (Hu vd., 2019) 2019 yılında kask takmanın hayati öneme sahip olduğuna vurgu yaparak YOLOv3'ü algoritmasının kullanıldığı şantiyelerde kask takmayan işçilerin tespiti ile ilgili gerçek zamanlı bir çalışma yapmışlardır. Yapılan çalışmada farklı boyutlardaki görüntülerle çalışabilmek için YOLO da bazı değişiklikler yaparak bağlantı katmanını çıkarttıklarını belirtmişlerdir. Kask takmadığı tespit edilen kişilerin görüntüsünün amirlere gönderilmesi ile ilgili bir eklentinin sisteme dahil edildiği bildirilmiştir.

İşyeri güvenliğini ve üretkenliğini artırmak için, bilgisayarla görmeye dayalı denetimler çalışanlar, tesisler, ekipman ve malzemeler gibi nesnelerin varlığını belirlemede çok önemlidir. Videodan hareket algılamaya yönelik ana teknikler derin öğrenme modellerini kullanır. Bu tür yaklaşımlarda, etiketli bir veri kümesi oluşturmak dışında süreç genellikle otomatiktir. Örneğin, Wei ve diğerleri (Wei vd., 2019) video görüntülerinde bir birey tarafından gerçekleştirilen güvensiz davranışları tespit etmek için derin öğrenmeye dayalı bilgisayarla görme tabanlı bir yaklaşım kullanmıştır. Çalışmada ayrıca şantiyeler bağlamında gerçek zamanlı güvenlik yönetimi teknikleri uygulanmıştır. İnşaat alanındaki güvensiz davranışları otomatik olarak yakalamak ve tanımak için çeşitli prototip görüş tabanlı yöntemler geliştirilmiştir. Ancak bu yaklaşımlar, dijital görüntüler veya videolar kullanılarak yakalandıklarında güvenli olmayan davranışlarda bulunan bireyleri tespit edemedikleri için kullanımı zor olmuştur. Bu sorunun üstesinden gelmek için, bir videodaki gereksiz bilgileri ortadan kaldırmak için Uzamsal ve Zamansal Dikkat Birleştirme Ağı kullanan ve böylece bir kişinin kimliğinin otomatik olarak tanımlanmasını sağlayan yeni bir derin öğrenme tekniği kullandıklarını belirtmişlerdir. Benimsenilen derin öğrenme yaklaşımı uzamsal özellik haritalarının uzamsal dikkat ağı kullanılarak çıkarılması, zamansal bilginin, zamansal dikkat ağları kullanılarak çıkarılması ve özellikler arasındaki mesafenin hesaplanması ile bir kişinin kimliği tanınması üzerinde durmuştur. Bu yaklaşım, güvenli olmayan davranışlarda bulunan kişileri otomatik olarak tanımak ve eylemleri ve olası sonuçları hakkında anında geri bildirim sağlamak için şantiye yöneticileri tarafından potansiyel olarak kullanılabilir olduğu belirtilmiştir.

Wu ve arkadaşları (Wu vd., 2019) inşaat işçilerinin farklı renklerdeki baretleri kullanmasının uygunluğunu belirlemek için tek atışlı çoklu kutu dedektörü (SSD) derin öğrenme algoritmasını kullanmışlardır. Şantiyelerde güvenliği artırmak için çoğu şirket, çok aşamalı veri işleme yoluyla baretlerin varlığını ve doğru kullanımını izlemektedir. Çalışma, inşaat personelinin baret takıp takmadığını otomatik olarak izlemek ve ilgili renkleri belirlemek için konvolüsyonel sinir ağına dayalı tek aşamalı bir sistem önermiştir. Çalışmada, çeşitli saha koşullarını kapsayan 3174 görüntüden oluşan yeni ve kamuya açık bir baret takma tespiti karşılaştırma veriseti oluşturmaktadır. Önerilen sistemde, farklı ölçeklerdeki farklı katmanlardan gelen özellikleri ayırt edici bir şekilde birleştirmek ve yeni bir özellik piramidi oluşturmak için ters aşamalı dikkat kullanılmıştır. Sistem uçtan uca bir şema kullanılarak eğitilmiştir. Deneysel sonuçlar, önerilen sistemin tüm saha koşullarında etkili olduğunu göstermektedir. Çalışma tek sınıf üzerine yapılmıştır.

Chen ve Demachi (Chen & Demachi, 2021) inşaat işçileri tarafından uygun olmayan koruyucu ekipman kullanımını belirlemek için derin öğrenme tabanlı bir nesne algılama ve geometrik ilişki analizi çözümü sunmuştur. Çalışma, derin öğrenme tabanlı nesne algılama ile geometrik ilişki analizi kullanarak birey algılamayı birleştirerek KKD'nin uygunsuz kullanımını tanımlayan yeni bir çözüm önermektedir. Deneysel sonuçlar, önerilen yaklaşımın KKD'nin uygunsuz kullanımıyla ilgili tehlikeleri yüksek hassasiyet oranıyla etkili bir şekilde tanımladığını ve gerçek zamanlı performansı koruduğunu göstermişlerdir.

Kong ve diğerleri (Kong vd., 2021) şantiyelerdeki güvensiz davranışları videolardan otomatik bir şekilde tahmin etmeyi amaçlayan LSTM tabanlı bir bilgisayarla görme yaklaşımı sunmuştur. Çalışmada, videolardan güvensiz davranışları otomatik olarak tahmin etmek için bilgisayarla görmeyi LSTM ile birleştirerek alandaki boşluğu doldurmaya katkıda bulunmayı amaçlamışlardır. Geliştirilen yöntemin mevcut yöntemlere göre daha iyi sonuçlar gösterdiği çalışmada belirtilmiştir.

Khan ve diğerleri (Khan vd., 2021) mobil iskele güvenliğini izlemek ve işçilerin güvensiz davranışlarını tespit etmek için maske R-CNN ve korelasyon tabanlı bir yaklaşım geliştirmiştir. İnşaat işçilerinin konumunu ve davranışlarını belirlemek için şantiye görüntülerinin analizi, güvenlik yönetimi ve verimlilik analizi için yararlı bir araç olarak tanımlanmıştır. Çalışmada inşaat işçisi tespiti için kullanılan mimari, konvolüsyonel sinir ağlarına dayanmaktadır. ResNet-152 ile Faster R-CNN kullanılarak 3241 görüntü için doğruluk, kesinlik ve geri çağırma oranları hesaplanmıştır. Önerilen yöntemin sonuçları, işçilerin duruşları, görünüşleri ve arka planları hakkında sınırlı varsayımlara dayanmadan çok

derin artık ađlar kullanarak řantiye grntlerinde birden fazla iřiyi dođru ve hızlı bir řekilde tespit etme potansiyelini gstermektedir.

Bir diđer alıřmada, Liu ve diđerleri (Liu vd., 2022) dijital grntlerden eřitli gvensiz davranıř biimlerini kategorize etmek iin derin đrenme teknikleri uygulamıřtır. alıřmada sırayla grsel ve metinsel zellik ıkarımı, zyinelemeli alt sorgu ve sınırlayıcı kutu oluřturma iřlemleri gerekleřtirilmiřtir. alıřmada, grsel bir topraklama yaklařımı kullanarak dijital bir grntdeki eřitli gvenli olmayan davranıřları tanımlamak ve bulmak iin bir gvenlik kuralı sorgusu kullanan bir arařtırma sunmuřlardır. alıřmaları grsel ve metin zelliklerinin ıkarılması, zyinelemeli bir alt sorgu gerekleřtirilmesi ve sınırlayıcı kutunun oluřturulmasını kapsayan  adımdan oluřmaktadır. Deneysel alıřmanın sonularının, son dnemlerde yapılan alıřmalarla kıyaslandığında olduka dřk olduđu grlmektedir.

Khan ve arkadařları (Khan vd., 2022) tarafından nerile hibrit bir yaklařım, uzamsal đrenme iin CNN'i ve zamansal đrenme iin LSTM kullanılmıř ve deneysel sonular hibrit modelin diđer makine đrenmesi ve derin đrenme tabanlı sınıflandırıcılara gre stnlđđ gsterilmiřtir.

Fang ve diđerleri (Fang vd., 2022) inřaat alanlarındaki tehlikeli hareketleri belirlemek iin ierik tabanlı bir grnt alma yntemi ortaya koymuřtur. Gvensiz davranıř grntlerini daha dođru ve otomatik olarak almak iin farklı modellerden ıkarılan zellikleri birleřtiren geliřmiř bir ierik tabanlı grnt alma (CBIR) yaklařımı nermiřlerdir. nerilen yaklařım, Huazhong Bilim ve Teknoloji niversitesi tarafından geliřtirilen dijital bir veri tabanı kullanılarak dođrulanmıřtır. Arařtırma, ResNet-101 ve VGG-19 fzyonunun daha yksek Top-K geri ađırma seviyelerine ulařtıđını ve tek zellik ıkarma ynteminden daha iyi performans gsterdiđini belirtmiřlerdir.

Bir bařka alıřmada Yang ve arkadařları (Yang vd., 2023), iřyerindeki video verilerinden gvenli olmayan eylemleri tespit etmek iin dnřtrc tabanlı bir derin đrenme modeli sunmuřlardır. alıřmada, davranıřlarının meknsal ve zamansal zelliklerinin paralel video akıřlarında aynı anda ıkarıldıđı ve daha sonra zel olarak tasarlanmıř bir modl tarafından birleřtirildiđi Meknsal Zamansal İliřki Transformatr (STR-Transformator) adlı bir derin đrenme modeli aıklanmaktadır. STR-Transformatrn etkinliđini dođrulamak iin, yedi kategori inřaat iřisi davranıřı ve 1595 video klibi ieren zelleřtirilmiř bir veri kmesi geliřtirmiřlerdir. nerilen STR-Transformatrn, inřaat projelerinde daha dođru ve gvenilir otomatik gvenlik gzetimi sađlayacađı ve kaza oranlarını ve ynetim maliyetlerini azaltacađı

belirtilmiştir. Ayrıca STR-Transformatörün performansı, inşaat yönetiminde bilgisayarla görme tabanlı derin öğrenme modelleri uygulanırken daha zengin özellikleri tanımlamak, çıkarmak ve birleştirmek için gelecekteki çalışmalara ilham verebilecek verimli özellik entegrasyonuna dayandığı ifade edilmiştir. Çalışmanın yönteminin hesaplama maliyetinin ve karmaşıklığının yüksek olması ve verisetinin sınırlı olması çalışmanın bazı kısıtlarıdır.

Sing ve arkadaşları (Singh vd., 2023) farklı İnsan Aktivite Tanıma (Human activity recognition, HAR) yaklaşımları ve bunların karşılaştırmalı analizi hakkında bir çalışma sunmuşlardır. Çalışmada güncel veri setlerinin çoğu, kullanıldıkları yaklaşımlarla birlikte tartışılmış ve doğruluk açısından sonuçları gösterilmiştir. Etkinlik tanıma için kullanılan farklı yöntemlerin kategorik olarak el yapımı özellik tabanlı ve otomatik özellik çıkarımı tabanlı olmak üzere iki isim altında toplayarak kapsamlı bir analizi sunulmuştur. Her kategori, kullanıldığı makalelerle birlikte tartışılmıştır. Literatürde kullanılan popüler veri kümeleri de açıklanmıştır. Otomatik özellik çıkarma yeteneği ve daha yüksek doğruluk nedeniyle, derin öğrenme yaklaşımları günümüzde karmaşık aktivitelerin analizinde araştırmacılar arasında en çok tercih edilen yaklaşım olduğu belirtilmiş ancak derin öğrenme modellerini büyük veri kümeleri üzerinde eğitmenin çok fazla bilgi işlem gücüne ve zamana ihtiyaç duyduğunu dolayısıyla bu sorunun, verimli kaynak kullanımına sahip özel CNN mimarileri yönünde gelecekteki çalışmalarla ele alınması gerektiği bildirilmiştir. Aktivite tanıma konusunda sensör kullanımı, Hareket Enerji Görüntüsü (Motion Energy Image, MEI) ve Hareket Geçmiş Görüntüleri (Motion History Images, MHI) gibi el yapımı özelliklerin derin öğrenme yaklaşımları ile birleşimleri gibi popüler olan konuların gelecek çalışmalara yön vereceği belirtilmiştir.

2024 yılında Barlybayev ve arkadaşları (Barlybayev vd., 2024) tarafından önerilen bir diğer çalışmada inşaat sektörüne yönelik olarak YOLOv8 ile KKD'nin kullanımının tespitine yönelik sensör tabanlı modellere karşı, bilgisayarla görmenin üstünlükleri belirtilmiş YOLOv8 modelleri karşılaştırılarak elde ettikleri sonuçları sunmuşlardır. Çalışmada belirlenen 6 sınıf tüm YOLO modelleri için eğitilmiş ve bu modellerin başarımları irdelenmiştir. 2016 yılında sahneye çıkan YOLO ağının her yeni versiyonda daha iyi duruma geldiği gösterilmiş ve ayrıca diğer bazı modellerle de kıyaslama yapılmıştır.

Aşağıdaki tablolarda konuyla ilgili yapılan literatürde daha önce önerilen bazı çalışmaların özetleri değerlendirme ve kısıtları ile birlikte sunulmuştur. Buna göre Tablo 2.1.'de geleneksel yöntemlerle yapılan çalışmalar ele alınmış ve özeti sunulmuştur.

**Tablo 2.1.** Tez konusu ile ilişkili geleneksel yöntemlerle önerilen çalışmaların genel değerlendirmesi

Çalışma	Veriseti	Görüntü Sayısı	Metodoloji	Değerlendirme ve Sınırlılıklar
(Barro-Torres vd., 2012)	-	-	Ban Mess ağı	RFID ve mikroişlemci tabanlı KKD'lerin kullanımını tespiti için geliştirilen bir çalışmadan bahsedilmektedir. Ancak her KKD için kullanılan cihazlar maliyet bakımından çalışmanın en büyük çıkmazıdır.
(Luo vd., 2014)	Özgün	760 Konum	GPS, RFID	Yapılan çalışmada çalışma ortamlarındaki güvensiz davranışları ve durumları tespit etmek için kablosuz sensör ağları ve çeşitli ekipmanlar kullanılmışlardır. Sensör ve kurulum maliyetleri söz konusudur.
(Guo vd., 2015)	Özgün	150000+	VSM JHA	Çin metro inşaatı için işçi davranışlarının gözlenmesine dayalı yapılan çalışmada büyük veri kullanılmıştır. Çalışma veri seti oluşturma ve analizi üzerinde durmuştur.
(Yu vd., 2017)	Özgün	102	Kinect, Microsoft	İşçilerinin hareket davranışlarını belirlemek için görüntü iskeletine ve gerçek zamanlı analize dayalı bir yöntem geliştirmiştir. İskelet duruşuna göre tehlikeli hareket tespiti amaçlanmıştır. Bu tür sistemlerle ilgili sınırlamaların, ekipmanın özel üretim ve ön işleme ihtiyacını içerdiği unutulmamalıdır.

İşyerlerindeki güvensiz davranışları tespit etmek için makine öğrenimi tekniklerini kullanan çalışmalar da yapılmıştır. Bu çalışmalarda Bayes ağları, Destek Vektör Makinaları, Haar Benzer Özellikleri gibi yöntemler uygulanmıştır. Bu çalışmalara ilişkin daha önce sunulan bazı çalışmaların genel bir değerlendirmesi Tablo 2.2.'de özetlenmiştir.

**Tablo 2.2.** Tez konusu ile ilgili makine öğrenmesi kullanarak önerilen çalışmaların değerlendirilmesi

Çalışma ve Yılı	Veriseti	Görüntü Sayısı	Metodoloji	Değerlendirme ve Sınırlılıklar
(Matías vd., 2008)	Özgün	148	Bayes, DVM, vd.	Çalışmada, işyeri kazalarının nedenleri ve türlerinin analizi için bir metodoloji önerilmektedir yaptıkları çalışmada özellikle kat seviyesindeki düşmelere odaklanmışlardır. Yaklaşım, makine öğrenme tekniklerine dayanmaktadır.
(Park & Brilakis, 2012)	Özgün	955+697	Haar-like features, Eigen-images	Şantiyelerdeki kamera görüntülerinden inşaat varlıklarını tanımak ve konumlarını belirlemek için Görsel tabanlı nesne tespiti, yaptıklarını belirtmişlerdir.
(Abbasi vd., 2015)	American National Safety Council	3145+185	Güvenlik örnekleme tekniği, Görüşme ve doküman inceleme yöntemleri, verileri analiz etmek için Ki-kare testi ve regresyon kullanılmıştır.	Güvensiz hareketlerin istatistiksel olarak değerlendirilmesi yapılmış ve eğitim ile önleneceği belirtilmiştir. Deneysel çalışmalar ve çalışma sonrası sonuçlar ele alınmamıştır.
(Alwasel vd., 2017)	Özgün Veri Seti	70+30	Sınıflandırma için Destek Vektör makinası	Ustaların duruşları tespit edilerek uzman veya çırak olarak sınıflandırmaya çalışılmıştır.
(Mneymneh vd., 2017)	Özgün	75+164	SURF, HOG BRISK, LBP FAST,	Kaskları tespit etmek için yararlı olduğu düşünülen çeşitli bilgisayarla görme tekniklerini üzerine çalışılmıştır. Sınıf sayısı azdır.
(Wu & Zhao, 2018)	Belirsiz	26722+8292	LBP, HMI, CH, H-SVM	İşçilerin baret takıp takmadıklarını ve baretlerin renklerini tespit etmeye odaklanmıştır

Son zamanlarda, nesne algılama, sınıflandırma ve video anlama üzerine yapılan araştırmalarda derin öğrenme tekniklerinin uygulanmasıyla önemli başarılar elde edilmiştir (Wei vd., 2019). Tablo 2.3.'te tez konusu ile ilgili çalışmalara ilişkin bazı çalışmaların genel bir değerlendirmesi ele alınmıştır.

**Tablo 2.3.** Tez konusu ile ilgili derin öğrenme tabanlı çalışmalar

Çalışma ve Yılı	Veriseti	Görüntü Sayısı	Metodoloji	Değerlendirme ve Sınırlılıklar
(Yao vd., 2015)	Youtube2Text DVS	1970 49000	3D-CNN, LSTM	Çalışma videolara tanımlayıcı metin yazmak üzerine yapılmıştır. Yeni bir yaklaşım önerilmiştir.
(Tran vd., 2015)	Sports-1M	1.1 Milyon	3B ConvNets (C3D)	Videolardaki hareketli nesnelerin ve zamansal dinamiklerin tanınması için yeni bir algoritma önerilmiştir.
(Donahue vd., 2015)	UFC101	12000	LRCN (CNN+LSTM)	Bu çalışma, ardışık veri, görsel, dilsel veya başka türlü görevleri içeren zamansal dinamikleri tanımayı ve sentezlemeyi öğrenebilen bir modelle derin hiyerarşik görsel özellik çıkarıcıyı birleştiren Uzun Dönemli Tekrarlayan Evrişimsel Ağ (LRCN) modelini önermektedir.
(Feichtenhofer vd., 2016)	UFC101, HMDB51	13320 6766	Two Stream Fusion	Uzamsal ve zamansal akışların nasıl ve ne zaman birleştirilmesi gerektiği üzerinde durmuşlar ve erken seviyede birleşmesi gerekliliğini bildirmişlerdir. Yüksek hesaplama maliyeti en bariz zayıf yönüdür.
(Girdhar vd., 2017)	UFC101, HMDB51	13320 6766	ActionVLAD	Videonun tüm uzamsal-zamansal kapsamı boyunca yerel konvolüsyon özellikleri toplayan eylem sınıflandırması için yeni bir video gösterimi önermişlerdir
(Carreira & Zisserman, 2017)	Kinetics400	240000	3B ConvNet I3D	I3D, video verilerinin işlenmesi ve aksiyon tanıma gibi görevlerde güçlü bir araç olsa da, yüksek hesaplama maliyeti, büyük veri seti ihtiyacı, genelleme zorlukları ve model karmaşıklığı gibi zayıf yönleri sahiptir.

Tablo 2.3.'ün devamı

(Diba vd., 2017)	Kinetics UFC101	240000+12 000	T3D	Videodaki hareketleri ve zamansal desenleri yakalamak için geliştirilmiş bir mimari sunmuşlardır. Yüksek hesaplama maliyeti, büyük veri gereksinimi, uzun süreli bağlantıları modelleme zorluğu gibi kısıtlamalara sahiptir.
(Ding vd., 2018)	Özgün	160+40	CNN+LSTM	Yeni bir yaklaşım sunarak hibrit bir modelle güvensiz hareketlerin tespiti amaçlanmıştır. Önceki yöntemlere göre oldukça başarılıdır. Ancak, hata tespit süresini azaltmak için hareket yakalama algoritmalarında iyileştirmeler gereklidir.
(Q. Fang, H. Li, X. Luo, L. Ding, H. Luo, vd., 2018)	Özgün	18000	ASC, SSD, CNN	Yüksekten düşmeyi önleme tedbiri kapsamında KKD denetimi yapan iki algoritma kademeli olarak önerilmiştir. Oklüzyon problemlerinin önüne geçilememiştir.
(Q. Fang, H. Li, X. Luo, L. Ding, T. M. Rose, vd., 2018)	Özgün+ WIDER FACE ve CelebFaces veri kümesi	8000+5900 00 Yüz	Faster R- CNN, SORT, Hungarian Algorithm, MTCNN	İşçilerin sertifikalarına uygun işleri yapıp yapmadıklarını tespit etmek karmaşık bir sistem önerilmiştir. Sahada yüz tanıma zorlukları ve cv sorunları oldukça zorludur.
(W. Fang vd., 2018)	Özgün	10000	Faster R-CNN	Çok ölçekli nesnelere tespitinde geliştirilmiş Hızlı R-CNN yöntemi önerilmiştir. Çalışma sadece nesne tespitine dayanmaktadır.
(Zhu vd., 2019)	UFC101, HMDB51, THUMOS14, ActivityNet1.2	13320+258 4 4819+2853	Hidden Two- Stream	Videolarda insan eylemlerini tanımaya yönelik yeni bir yaklaşım önerilmiştir. Veri senkronizasyonu ve doğru birleştirme stratejileri belirlemek, performans açısından kritik
(Son vd., 2019)	ImageNet, Coco2014	14197122 320000	Faster R-CNN ResNet-152	Farklı poz ve arka planlardaki inşaat işçilerini tespit etmek için derin artık ağları ve daha hızlı R-CNN derin öğrenme mimarilerini kullanan bir model önermiştir.

Tablo 2.3.'ün devamı

(Wu vd., 2019)	Özgün	3174	SSD	Yapılan çalışmada inşaat işçilerinin uygun renklerdeki baretleri kullanıp kullanmadıkları tespit edilmiştir, sınıf tek yönlüdür.
(Wei vd., 2019)	iLIDS-VID Özgün	600 24	ASTPN	İnşaat sahalarında insanların kimliklerini belirleme üzerine Sınırlı sayıda aktivite
(Hu vd., 2019)	Özgün	20554	YOLOv3	Kask takmayan işçilerin tespiti üzerine yapılan çalışmada tek sınıf üzerine yapılmıştır.
(Chen & Demachi, 2021)	Özgün ve Web Madenciliği	6029	OpenPose YOLOv3	Kişi ve KKD ler tespit edilerek mesafe ölçülerek uygun kullanımları tespit edilmiştir. Nesnelere ilişkilendirilerek duruş sahne tahmini yapılmıştır. Karmaşık sahneler için sonuçlar irdelenebilir.
(Khan vd., 2021)	Özgün	3241	ResNet-152 Mask R-CNN	Mobil iskele güvenliğini incelemek için korelasyona dayalı bir sistem önerilmiştir. Sınırlı veri kümesi ve sınıf azlığı söz konusudur.
(Kong vd., 2021)	Davis, Özgün		Siam-Mask, Social-LST, PnPoly	İnşaatlarda kişilerin takip edilmesi ve güvensiz hareketlerin tespitine yönelik yapılan çalışmada %93 oranında başarı elde edilmiştir. Çalışmadaki tehlikeli alan ve veri seti kısıtlaması vardır.
(Fang vd., 2022)	Huazhong University of Science and Technology	40000	VGG-19, ResNet-101, CBIR Vd.	İnşaat alanlarındaki insanların güvenli olmayan davranışlarının görüntülerini elde etmek için derin öğrenme füzyon yaklaşımı önerilmiştir.
(Liu vd., 2022)	Özgün	860	YOLOv3, BERT	Dijital görüntülerden çeşitli güvensiz davranış biçimlerini kategorize etmek için derin öğrenme teknikleri uygulanmıştır. Sınıf sayısı artırılmalıdır.
(Khan vd., 2022)	Özgün	2400	CNN, LSTM	İnsan aktivitesi tanıma üzerine geliştirilen hibrit bir modelde %91 başarı elde edilmiştir.

Tablo 2.3.'ün devamı

Yang vd., 2023)	CMA	1595	STR- Transformer	Yazarlar işyerindeki video verilerinden güvenli olmayan eylemleri tespit etmek için dönüştürücü tabanlı bir derin öğrenme modeli sunmuştur. Hesaplama maliyeti yüksek bir model
(Barlybayev vd., 2024)	CHV, SHEL5K	4000+500+ 500	YOLOv8 versiyonları	CHV ve SHEL5K veri kümesini kullanarak PPE'yi tespit etmek için YOLOv8 mimarisinin çeşitli sürümlerinin kullanılmıştır. Gerçek zamanlı sonuçlar değerlendirilmemiştir

Literatür taraması neticesinde yapılan çalışmaların büyük çoğunluğunun inşaat sahalarında, şantiye alanlarında ve metro çalışmalarına yapıldığı anlaşılmıştır. Bu tez çalışmasında öncelikle özgün bir veriseti tasarlanarak endüstriyel üretim sahaları için önemli bir eksik olan veriseti, akademik çalışmalara hizmet etmek için sunulmuştur. Bu çalışma gerçek bir fabrikada, gerçek zamanlı olarak tasarlanmıştır ve aktivite tanıma ile tüm güvensiz davranışların sınıflandırılmasında %95.81 doğruluk oranı elde edilmiştir. Yapılan diğer aktivite tanıma çalışmalarına göre yaklaşık %2 oranında iyileştirme elde edilmiştir. Kullanılan donanımlar hızla gelişse de mevcut çalışmalar da saniyedeki kare sayılarının oldukça düşük olduğu görülmektedir. Bu tez çalışmasında nesne tespitine yönelik yapılan YOLOv4 algoritması ile gerçek zamanlı 32 saniyede kare sayısı (Frame Per Second, FPS) hızında %91,18'lik bir mAP@50 değeri elde edilmiştir.

### 3. MATERYAL VE YÖNTEM

Bu tez çalışmasında, endüstriyel üretim tesislerinde güvenli olmayan hareketleri tespit etmek ve gözetim kameralarından KKD'lerin kullanımını kontrol etmek için iki aşamalı bir yaklaşım sunulmuştur. Birinci aşamada, KKD'lerin kullanımını değerlendirmek için YOLOv4 derin öğrenme algoritması tabanlı bir yöntem önerilmiştir. Burada, fabrikalara kurulan kameralardan gelen videoların YOLOv4 derin öğrenme yöntemleriyle anında incelenmesi ve sahnenin anlamlı bir şekilde tanınması sağlanmıştır. Çalışmada, farklı çalışma ortamlarından toplanan videolarla bir veri seti oluşturulmuş ve çalışma kapsamında, video veri seti üzerinde, YOLOv4 derin öğrenme algoritması kullanılarak endüstriyel alanlarda faaliyet gösteren fabrikalardaki işçilerin uygun ekipmanları doğru şekilde kullanıp kullanmadıkları belirlenmiştir. İkinci aşamada ise, bir fabrika ortamında güvenli olmayan davranışların tespiti ve sınıflandırılması için YOLOv4 ve ConvLSTM derin öğrenme mimarilerine dayanan gerçek zamanlı bir bilgisayarla görme sistemi geliştirilmiştir. Her iki aşama için ayrı veri setleri oluşturularak, deneysel analizler yürütülmüştür.

#### 3.1. Verisetleri

Bu tez çalışmasında, KKD tespiti için hazırlanan bir veri seti ve çalışma ortamlarından güvensiz davranışların tespiti için hazırlanan bir başka veri seti olmak üzere iki farklı veri seti oluşturulmuştur.

##### 3.1.1. KKD Tespiti İçin Hazırlanan Veriseti

Tez çalışmasının bu aşamasında kullanılan görüntü ve video verileri Eskişehir Organize Sanayi Bölgesinde faaliyet gösteren Kafaoğlu Metal Plastik Makine San. ve Tic. A.Ş. ile Tek Metal ve Plastik Endüstriyel Mamulleri San. Tic. Ltd. Şti. adlı iki farklı işletmeden elde edilmiştir. Derin öğrenme algoritmalarında daha yüksek doğruluk başarımı elde etmek için çok fazla görüntü/video ile mimarinin eğitimi yapmak gerekmektedir. Bu nedenle, video/görüntü veri setinin genişletilmesi ve deneysel çalışmaların hızlandırılması için Bilecik Şeyh Edebali Üniversitesi Meslek Yüksekokulu Makine Atölyesindeki çalışma ortamlarından, çeşitli zamanlarda birkaç farklı kamera kullanılarak farklı çözünürlüklerde video ve görüntüler toplanmıştır. Veri alınan tüm ortamlar için ilgili kurum/kuruluşlardan gerekli izinler alınmıştır. Çalışma kapsamında hazırlanan veri seti için toplanan videolardan/görüntülerden bir kısmı için örnek çerçeveler Şekil 3.1.' de sunulmuştur.



**Şekil 3.1.** Tez kapsamında KKD'lerin tanınması için hazırlanan verisetinden örnek görüntüler.

Özellikle üretim alanlarında bilgisayarla görme açısından en büyük problem aydınlatmadır. Genellikle bu alanlar sürekli olarak yapay aydınlatma kaynakları ile ışıklandırılmaktadır. Bu yüzden, bu aşamada, görüntü verileri elde edilirken aydınlatma bakımından en doğru açılar kullanılmıştır. Farklı sahnelerden elde edilen videolardan toplam 2200 tane farklı boyutlarda görüntü elde edilerek tez çalışmasında kullanılmak üzere veriseti oluşturulmuştur. Görüntü verileri çalışanların KKD kullanımlarının tespiti sağlayabilecek şekilde tesis edilmiştir. Dolayısıyla görüntülerde işçilerin çalışırken kullanmaları gereken baret, koruyucu gözlük, eldiven, çalışma yeleği ve maske üzerine odaklanılmıştır. Veri seti oluşturulurken bazı sınıfların hem varlığı hem yokluğu sınıf etiketi olarak belirlenmiştir. Bunun sebebi, bu durumun mAP parametresini etkileyip etkilemeyeceğini öğrenmek içindir. Ayrıca, baret ve maske olma veya olmama durumu saç ve sakal ile karıştırılabileceği düşünülerek yüksek doğruluk amaçlanarak algoritma ilave iki sınıf eklenerek de eğitilmiştir. Bu yöntem daha iyi karar destek sistemleri oluşturmak için etkilidir.

Çalışma kapsamında veriseti için elde edilen videolardan saniye başına 4 çerçeve elde edilecek şekilde 4000 görüntü elde edilmiştir. Elde edilen bu görüntülerden 2200 tanesi Labelimg (Tzotalin, 2015) etiketleme programı kullanılarak ağız eğitimi yapılan YOLOv4 algoritmasının etiket yapısına formatına uygun şekilde etiketlenmiştir. Etiketleme işlemi görselde nesnenin konumunu belirleyecek şekilde kutu içine alınması şeklinde yapılmıştır.

Buna göre üretilen etiket dosyasında nesnenin x-y koordinatları, yüksekliği ve genişliği belirlenmiştir. Görsel ve görsellerde bulunan nesne sayılarının fazla olması etiketleme işleminin uzun zaman almasına neden olmuştur. Etiketleme işlemini sonrasında nesne tanıma kullanılan sınıflar ve etiketlenen sınıflardaki nesne sayıları Tablo 3.1’de görülmektedir. Buradaki nesne sayılardaki farklılık, sahnelerdeki nesne sınıflarının farklı olmasından kaynaklanmaktadır. Diğer taraftan, bu durum nesne sayılarındaki farklılıkların ağırlık eğitimi üzerindeki nasıl bir etkisi olduğunun belirlenmesine yardımcı olmuştur.

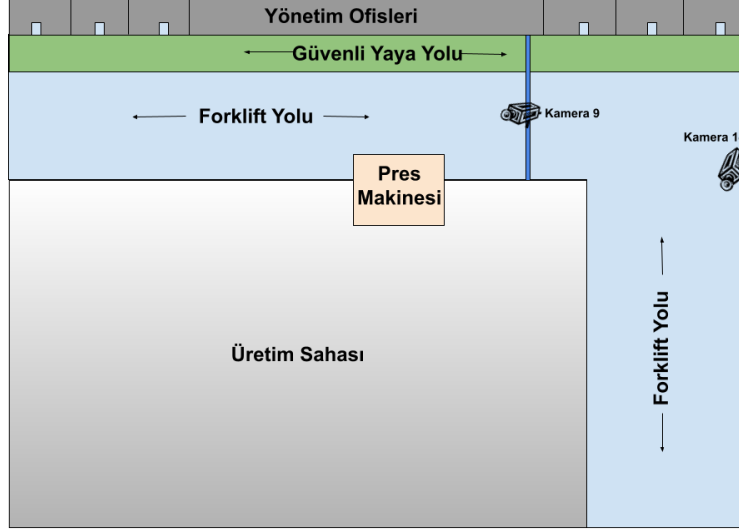
**Tablo 3.1.** KKD tespiti için nesne tanımda kullanılan sınıflar ve etiketlenen sınıflardaki nesne sayıları

Sınıf No	Sınıflar	Nesne Sayısı
1	Baret	1891
2	Eldiven	3408
3	Gözlük	2295
4	Maske	3214
5	Baret-Yok	3010
6	Maske-Yok	370
7	Yelek	2621

### 3.1.2. Güvenli ve Güvensiz Davranışların Sınıflandırılması İçin Hazırlanan Video Veriseti

Tez çalışması kapsamında, iş yerlerindeki güvensiz davranışların gerçek zamanlı otomatik tespiti için derin öğrenme modelleri ile kullanılmak üzere özgün bir video veriseti daha hazırlanmıştır. Bu verisetine ilişkin video kayıtları 5 Kasım-13 Aralık 2022 tarihleri arasında Eskişehir OSB’de faaliyet gösteren "Kafaoğlu Metal Plastik Makine San. ve Tic. A.Ş." adlı fabrikanın üretim sahasında bulunan güvenlik kameralarından alınmıştır. Video verileri iki farklı IP kameradan 39 gün boyunca, toplamda yaklaşık 4000 saat boyunca elde edilmiştir. IP kameraların markası UNV, modeli ise IPC2122CR3-PF40-A'dır. Ayrıca IP kameralar tam yüksek çözünürlükte (Full High Definition, Full HD) kayıt yapabilmektedir. Söz konusu kameraların konumları ve açıları şematik olarak Şekil 3.2.’de gösterilmiştir. Bu video veriseti hem video içeriğini anlamak hem de video içeriğindeki nesnelere tanımak için iki alt gruba ayrılmıştır. Videolarda yer alan çalışanlardan kişisel verilerin korunması ve kullanılmasına

ilişkin bir onay formu alınmıştır. Ayrıca Bilecik Şeyh Edebali Üniversitesi Rektörlüğü Etik Kurulu'nun 11.06.2021 tarih ve 10 sayılı kararı ile çalışmanın yürütülmesinde etik açıdan bir sakınca bulunmadığı onaylanmıştır.



**Şekil 3.2.** Video veriseti hazırlanması aşamasında veri alınan fabrika ortamının şematik gösterimi ve kameraların konumu

Tez çalışması için güvenli ve güvensiz davranışlar için video tabanlı veriseti oluşturulurken fabrika yöneticileri ve işyeri güvenliği uzmanlarının görüşleri alınarak çalışma ortamında en sık rastlanan 4 farklı güvensiz davranış belirlenmiştir. Bu güvensiz davranışlar şunlardır: Güvenli Yürüyüş Yolu İhlali, Yetkisiz Müdahale, Açılan Pano Kapağı ve Forklift ile Aşırı Yük Taşıma. Ayrıca, geliştirilen gerçek zamanlı tespit sisteminde sınıflandırma ve karar verme için Güvenli Yürüyüş Yolu, Yetkili Müdahale, Kapalı Pano Kapağı ve Forklift ile Güvenli Taşıma olmak üzere 4 güvenli davranış da belirlenmiştir. Çalışmada güvenli ve güvensiz davranış sınıfları için sınıf kimliği, sınıf adı ve davranış türü bilgileri Tablo 3.2.'de gösterilmiştir.

**Tablo 3.2.** Güvenli ve güvenli olmayan davranış sınıfları için sınıf kimliği, sınıf adı ve davranış türü bilgileri

Sınıf No	Sınıf İsmi (Güvenli/Güvensiz)	Davranış Tipi
0	Güvenli Yürüyüş Yolu İhlali	Güvensiz
1	Yetkisiz Müdahale	Güvensiz
2	Açık Pano Kapağı	Güvensiz
3	Forklift ile Aşırı Yük Taşıma	Güvensiz
4	Güvenli Yürüyüş Yolu	Güvenli
5	Yetkili Müdahale	Güvenli
6	Kapalı Pano Kapağı	Güvenli
7	Güvenli Taşıma	Güvenli

Veri setinin oluşturulması için öncelikle, işçilerin/çalışanların güvenli ve güvenli olmayan davranışları için işletmede bulunan 2 farklı kameradan 19 gün boyunca videolar kaydedilmiştir. Veriler toplandıktan sonra, uzmanlar eşliğinde güvenli ve güvenli olmayan davranışların olduğu kesitler belirlenerek ortalama 3-18 saniyelik video parçaları oluşturulmuştur. Şekil 3.3'te video veriseti için belirlenen sınıflara ait güvenli ve güvensiz davranışların kamera açılarının gösterildiği görsel örnekler verilmiştir. Şekil 3.3(a)'da çalışma ortamında yeşil yol boyunca Güvenli Yürüyüş Yolu sınıfı için güvenli bir davranış varken, Şekil 3.3(b)'de yeşil yol dışında Güvenli Yürüyüş Yolu İhlali sınıfı için güvensiz bir işçi davranışı görülmektedir. Şekil 3.3(c)'de bir makineye bağlı pano için Kapalı Pano Kapağı sınıfında güvenli bir davranış olmuşken, Şekil 3.3(d)'de tam tersi Açık Pano Kapağı sınıfı için güvensiz bir davranış ortaya çıkmıştır. Diğer taraftan, Şekil 3.3(e)'de forklift ile 3 blok ve daha fazlası taşıma yapılarak güvensiz bir çalışan davranışına örnek (Forklift ile Aşırı Yük Taşıma) görülmektedir. Şekil 3.3(f)'de forklift ile 2 blok ve daha aşağısı taşıma yapılarak güvenli bir davranış bulunurken (Güvenli Taşıma), Şekil 3.3(g)'de yeşil renkli yelek ile panoya yetkili bir müdahale varken (Yetkili Müdahale), son olarak Şekil 3.3(h)'de kırmızı-siyah karışımı renkli yelek ile panoya yetkisiz bir müdahale örneği (Yetkisiz Müdahale) görülmektedir.

Çalışmada, orijinal olarak oluşturulan veriseti için fabrikalardaki çalışma ortamlarında güvenli ve güvensiz davranışları içeren 8 sınıfa (Güvenli Yürüyüş Yolu İhlali, Yetkisiz Müdahale, Açık Pano Kapağı, Forklift ile Aşırı Yük Taşıma, Güvenli Yürüyüş Yolu, Yetkili Müdahale, Kapalı Pano Kapağı ve Güvenli Taşıma) ait 05 Kasım 2022 ve 23 Kasım 2022 tarihleri arasında toplanan video parçalarından oluşan eğitim ve test kümesi sayıları Tablo 3.3'te verilmiştir. Veri setinde eğitim için toplam 566 video parçası bulunurken, test kümesi için toplamda 125 video parçası hazırlanmıştır.

**Tablo 3.3.** 05-23 Kasım 2022 tarihleri arasında toplanan video parçalarından oluşan eğitim ve test kümesi sayıları

Sınıf Adı (Güvenli/Güvensiz Davranış)	Eğitim seti için Video sayısı	Test seti için Video sayısı	Ortalama video süresi (s)
Güvenli Yürüyüş Yolu İhlali	178	32	9.7
Yetkisiz Müdahale	97	11	10.2
Açık Pano Kapağı	129	13	5.8
Forklift ile Aşırı Yük Taşıma	48	8	6.9
Güvenli Yürüyüş Yolu	50	25	9.2
Yetkili Müdahale	23	15	8.6
Kapalı Pano Kapağı	19	13	4.6
Güvenli Taşıma	22	8	7.2
<b>Toplam</b>	<b>566</b>	<b>125</b>	



**Şekil 3.3.** Verisetinde güvenli ve güvensiz sınıflarına ait davranışlara örnekler. (a) Güvenli Yürüyüş Yolu, (b) Güvenli Yürüyüş Yolu İhlali, (c) Açık Pano Kapağı, (d) Kapalı Pano Kapağı, (e) Forklift ile Aşırı Yük Taşıma, (f) Güvenli Yük Taşıma, (g) Yetkili Müdahale, (h) Yetkisiz Müdahale

**Tablo 3.4.** YOLOv4 kullanılarak nesne tespiti için eğitim ve test setlerinde etiketlenen görüntülerin sayısı

Güvenli/Güvensiz Davranış	Eğitim Kümesindeki Etiketlenen Görüntü Sayısı	Test Kümesindeki Etiketlenen Görüntü Sayısı
Güvenli Yürüyüş Yolu İhlali	526	100
Yetkisiz Müdahale	180	23
Açık Pano Kapağı	590	60
Forklift ile Aşırı Yük Taşıma	200	40
Güvenli Yürüyüş Yolu	351	43
Yetkili Müdahale	82	12
Kapalı Pano Kapağı	252	40
Güvenli Taşıma	81	11
<b>Toplam</b>	<b>2262</b>	<b>329</b>

Tablo 3.4.'te görüldüğü gibi, YOLOv4'ün eğitimi için eğitim kümesindeki videolardan çerçeveler elde edilerek sınıflara ait toplam 2262 görüntüdeki nesne Labelimg (Tzutalin, 2015) yazılımı ile YOLOv4 algoritmasının etiket yapısına uygun şekilde etiketlenmiştir ve gerçek sınıf etiketleri (ground truth) oluşturulmuştur. Ayrıca YOLOv4'ün nesne tespitindeki başarısını değerlendirmek için oluşturulan test kümesindeki video çerçevelerinden ise toplamda 329 görüntüdeki güvenli/güvensiz davranış sınıflarına ait görüntünün aynı yöntemle etiketlenmesi

gerçekleştirilmiştir. Etiketleme işlemi, görselde nesnenin konumunu belirleyecek şekilde fabrikada görevli bir iş güvenliği uzmanı eşliğinde kutu içine alınarak yapılmıştır. Üretilen etiket dosyasında nesnelerin x-y koordinatları, yüksekliği ve genişliği ile ilgili değerler bulunmaktadır.

Tez çalışmasında, YOLOv4 ile kullanılarak elde edilen sonuçlar, ayrıca güncel YOLO mimarilerinin de bu veriseti üzerinde elde edilen sonuçlara etkisini değerlendirmek için YOLOv11 modeli için de deneysel çalışmalar yürütülmüştür. Tablo 3.5.'te YOLOv11 için video anlamlandırmada kullanılan video veriseti ile ilgili nesne sayıları gösterilmiştir. YOLOv11 için, YOLOv4 algoritmasının eğitiminde kullanılan veriseti etiket veya boyut değiştirilmeden %70 oranında Eğitim, %20 oranında Doğrulama ve %10 oranında Test olmak üzere yeniden düzenlenmiştir.

**Tablo 3.5.** YOLOv11 kullanılarak nesne tespitinde her bir sınıf için eğitim, doğrulama ve test kümelerinde etiketlenen görüntülerin sayısı

Güvenli/Güvensiz Davranış	Eğitim İçin Etiketlenen Görüntü Sayısı	Doğrulama İçin Etiketlenen Görüntü Sayısı	Test İçin Etiketlenen Görüntü Sayısı
Güvenli Yürüyüş Yolu İhlali	445	136	43
Yetkisiz Müdahale	146	33	19
Açık Pano Kapağı	465	132	60
Forklift ile Aşırı Yük Taşıma	160	52	25
Güvenli Yürüyüş Yolu	286	75	45
Yetkili Müdahale	67	18	8
Kapalı Pano Kapağı	208	59	30
Güvenli Taşıma	66	9	14
<b>Toplam</b>	<b>1843</b>	<b>514</b>	<b>244</b>

### 3.2. YOLO Mimarileri

Bilgisayarla görme ve görüntü tanımda yaygın olarak kullanılan derin öğrenme algoritmaları, çoklu dağıtılmış temsil seviyelerini keşfetmeyi amaçlayan makine öğrenme algoritmalarının bir alt kümesidir. Son zamanlarda, geleneksel yapay zekâ problemlerini çözmek için çok sayıda derin öğrenme algoritması önerilmiştir (Guo vd., 2016). Derin öğrenme yöntemlerinin temel yöntemi Evrimsel Sinir Ağlarıdır (CNN). CNN'in amacı, tüm özellikleri bir kaynak görüntüden çıkarmak ve ardından görüntüdeki nesneyi sınıflandırmak için tüm bu özellikleri kullanmaktır. CNN'in bu yapısı sayesinde görüntülerdeki çeşitli nesnelere otomatik olarak tanınabilmektedir (Q. Fang, H. Li, X. Luo, L. Ding, T. M. Rose, vd., 2018).

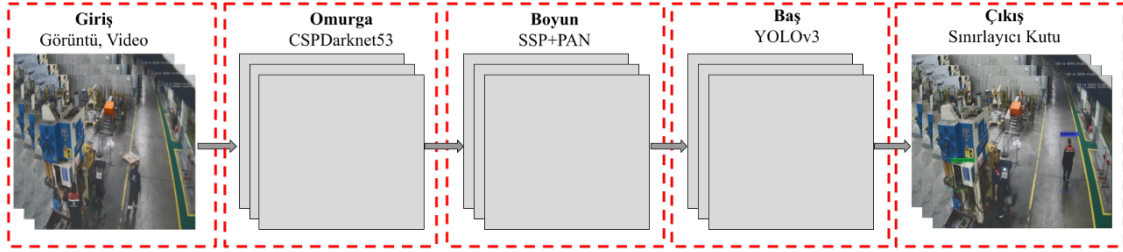
CNN'in ortaya çıkışı, nesne algılama alanının hızlı bir şekilde gelişmesine yol açmıştır (Krizhevsky vd., 2017). Bölgesel Evrişimsel Sinir Ağlarından (R-CNN, B-ESA) (Girshick vd., 2014), Uzamsal Piramit Havuzlamaya (SPP) (He vd., 2015) ve Hızlı B-CNN (Fast R-CNN)(Girshick, 2015) yöntemlerine sürekli iyileştirmeler önerilmektedir. Nesne tanıma alanındaki en son gelişmiş algoritmalar, Daha Hızlı Bölgesel-Evrişimsel Sinir Ağları (Faster R-CNN, DH B-ESA) (Ren vd., 2015), Single Shot Multibox Detector (SSD) (Liu vd., 2016) ve You Only Look Once (YOLO) (Redmon vd., 2016) olarak sıralanabilir. Bu algoritmalarından YOLO'nun diğerlerine göre aynı platformlarda yapılan çalışmalarda çok daha hızlı olduğu ve ortalama doğruluk değeri en yüksek olan Faster R-CNN'den bile daha başarılı olduğu görülmüştür (Kim vd., 2020). Çıkış yılı olan 2016'dan bu zamana her yıl geliştirilen YOLO'nun özellikleri iyileştirilmiş ve yeni versiyonları da duyurulmaktadır. Bunun yanında hızı, yüksek doğruluğu ve zamanla hesaplama maliyetinin düşmesi sebebiyle YOLO, videolarda nesne tanıma ve izleme için en uygun nesne tespit yöntemlerinden birisi haline gelmiştir (Q. Fang, H. Li, X. Luo, L. Ding, T. M. Rose, vd., 2018; Wang vd., 2024). YOLO ile son yıllarda iş güvenliğinin de kapsayacak biçimde sayısız çalışmalar yapılmıştır. Bu çalışmalarda genellikle şantiyelerde işçilerin kask takıp takmadıkları buldukları alanlarda çalışabilmek için gerekli sertifikasyona sahip olup olmadıkları denetlenmiştir. Tez çalışmasında imalat sanayinde çalışanların kullanmaları gereken tüm ekipmanları kullanıp kullanmadıkları, kendilerine ayrılan alanlarda hareket edip etmedikleri, tesis içinde hız ve taşıma kurallarına uyup uymadıkları, YOLOv4 ve YOLOv11 kullanılarak sırasıyla nesne tespiti, nesne takibi ve olay anlamlandırma çalışmaları ile tespit edilmiştir.

### 3.2.1. YOLOv4

YOLOv4 algoritması özellikle hızı ve tespit başarımını iyileştirerek gerçek zamanlı nesne tanımda yaygın olarak kullanılan bir derin öğrenme algoritmasıdır (Bochkovskiy vd., 2020; Long vd., 2020). Bochkovskiy ve vd. (Bochkovskiy vd., 2020), mevcut YOLO yapısının optimize edilmesiyle, ağın eğitiminin tek bir GPU grafik işleme kartıyla kolayca yapılabileceğini ve ayrıca Tesla V100 grafik kartıyla yüksek doğrulukta performans elde edilebileceğini belirtmiştir. Şekil 3.4.'te blok diyagramı gösterilen YOLOv4 algoritması üç alt katmandan oluşmaktadır. Bu mimaride, Omurga (backbone) katmanında özellik çıkarımı yapılırken, Boyun (neck) katmanında ise, nesnelere tahmin edilmesinde daha yüksek başarımlar sağlanması için, komşu özellik haritalarından alt-üst ve yukarıdan-aşağı akışlarla bilgi çıkarımı sağlanmaktadır. YOLOv4 mimarisinde, omurga olarak CSPDarknet53 kullanılır ve Boyun katmanında uzaysal piramit havuzlama (Spatial pyramid pooling, (SPP) ve Yol Birleştirme Ağı

(path aggregation network, PAN)'dan yararlanılmaktadır (Guo vd., 2021). Son katman olan Baş (Head) kısmında ise, sınırlayıcı kutular (bounding box) bulunur ve her kutunun sınıfı tahmin edilir. Baş katmanında, aynı zamanda YOLOv3 algoritmasının tahmin etme prosedüründen yararlanılmaktadır (Wu vd., 2020).

Tez çalışmasında, güvenli olmayan hareketlerin gerçek zamanlı olarak algılanması için nesne ile güvensiz hareket ilişkisini değerlendirmede YOLOv4 algoritması kullanılmıştır. Hazırlanan video verisetindeki güvensiz hareketlerin olduğu çerçeveler YOLOv4 derin öğrenme algoritması ile seçilmektedir. Dolayısıyla, çalışmada, belirlenmiş tehlikeli hareketler ile ilişkiz olan çerçeveler YOLOv4 ile ayrıştırılarak video anlamlandırma mimarisinin yükü hafifletilmiş ve hızı ve doğruluğu artırılmaktadır.



**Şekil 3.4.** Çalışmada video parçalarından çerçeve seçilmesi için kullanılan YOLOv4 algoritmasının blok diyagramı

### 3.2.2. YOLOv11

Ultralytics tarafından geliştirilen YOLO nesne algılama modelleri ailesinin en son versiyonu olan YOLOv11 önceki sürümlerin güçlü yönlerini temel alan, nesne algılama, örnek segmentasyonu, poz tahmini ve sınıflandırma gibi çeşitli bilgisayar görüşü görevlerini ele alma yeteneğinin yanı sıra bir dizi önemli mimari iyileştirme ve performans iyileştirmesi sunmuştur. YOLOv11'in omurga kısmı önceki sürümlerde kullanılan C2f bloğu yerine C3k2 bloğuyla değiştirilmiştir. Bu sayede önceki versiyonlarına göre karar verme süresinin kısaldığı belirtilmiştir. YOLOv11 SPPF (Spatial Pyramid Pooling Fast) katmanını korur iken modelin uzamsal yeteneğini C2PSA (Cross Stage Partial with Spatial Attention) bloğunu ekleyerek geliştirmiştir. Bu, modelin bir görüntüdeki önemli bölgelere odaklanma yeteneğini geliştirmiştir ve böylece nesne algılama doğruluğunu ve verimliliğini artırmıştır (Jocher vd., 2024).

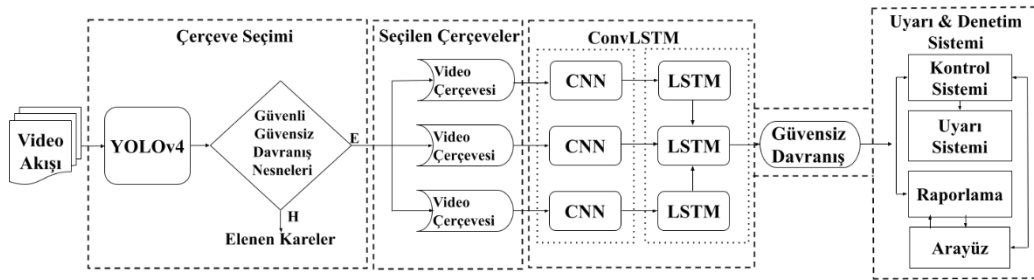
YOLO mimarisinde boyun, (neck) farklı çözünürlüklerdeki özellikleri toplamaktan ve bunları nihai tahminler için model kafasına iletmekten sorumludur. YOLOv11, eski blokları

daha hızlı ve daha verimli C3k2 bloğu ile değiştirerek boynu geliştirmiştir. Bu sayede çeşitli boyutlardaki nesnelere için daha sonuçlar elde edilebilmektedir (Jocher vd., 2024). YOLOv11 ile sadece nesne tespiti değil aynı zamanda örnek bölütleme, poz tahmini ve sınıflandırma yapılabilmektedir. Önceki versiyonlarda olduğu gibi YOLOv11 in küçük, orta, geniş ve ekstra geniş varyantları çeşitli varyantları geliştiriciler tarafından duyurulmuştur. Büyüyen varyantlarda mimari yavaşlarken doğruluk artmaktadır (Jocher vd., 2024).

Tez çalışmasında, YOLOv11'e göre daha eski bir sürüm olan YOLOv4 ile güvensiz davranışların tespiti için hazırlanan video veriseti üzerinde elde edilen sonuçların, güncel YOLO mimarileri ile değerlendirilmesi için, deneysel çalışmaların bir kısmı YOLOv11 ile de yürütülmüştür. Böylece, iki mimari arasında anlamlı bir fark olup olmadığı karşılaştırılarak, çalışmanın ilerlemesi değerlendirilmiştir.

### 3.3. Önerilen Hibrit Derin Öğrenme Mimarisi (Unsafe-Net)

Çalışma ortamlarında güvenli olmayan davranışların tespiti ve sınıflandırılması için YOLOv4 ve ConvLSTM derin öğrenme mimarilerine dayanan gerçek zamanlı bir bilgisayarla görme sisteminin (Unsafe-Net) çalışma mekanizmasını gösteren blok diyagram Şekil 3.5.'te gösterilmiştir. Çalışmaya, iş birliği yapılan fabrikada hareketli makinelerin ve iş trafiğinin olduğu farklı çalışma ortamlarından güvenli ve güvensiz davranışları içeren videolardan oluşan bir veriseti oluşturularak başlanmıştır. Çalışmada, güvensiz davranışlar içeren nesnelere tespit ederek, akan görüntüleri seçerek ve diğer tüm kareleri hariç tutarak videoyu basitleştirmek için YOLOv4 kullanılmıştır. YOLOv4, videodan güvenli olmayan davranış karelerini seçer ve ardından bu kareler içindeki güvenli olmayan eylem sınıflarına ait nesnelere için bir video segmenti oluşturur.



**Şekil 3.5.** Videolardan güvensiz davranışların gerçek zamanlı olarak tespiti için önerilen YOLOv4 ve ConvLSTM tabanlı Unsafe-Net hibrit derin öğrenme sisteminin blok diyagramı ve mekanizması

### 3.3.1. ConvLSTM

Tez çalışmasında, güvenli olmayan hareketleri sınıflandırmak için CNN'in konvolüsyonel katmanlarını LSTM ile birleştiren hibrit bir derin öğrenme mimarisi olan ConvLSTM modeli kullanılmıştır. ConvLSTM, videoyu analiz etmek ve tahmin etmek için görüntülerin şekil bilgilerini davranış ve hız gibi zamansal bilgilerle birleştirir. Model, konvolüsyon işlemleri ve daha yüksek boyutlu veri gösterimi dışında tipik LSTM yapısına benzer. Şekil 3.6.'da CNN ve LSTM ağlarını birleştiren ConvLSTM mimarisinin blok diyagramını gösterilmektedir. Ağ girişleri  $X_1, \dots, X_t$  iken, ConvLSTM ağının bellek birimleri  $C_1, \dots, C_t$  ile temsil edilmektedir. LSTM ağının gizli katman durumu  $H_1, \dots, H_t$  ile gösterilir. ConvLSTM ağı giriş kapısı, unutma kapısı ve çıkış kapısı için sırasıyla  $i_t$ ,  $f_t$  ve  $O_t$  kullanır. Ağdaki tüm değişkenler 3B tensörler olarak yapılandırılmıştır; burada '\*' konvolüsyon işlemini ve '°' Hadamard çarpımını temsil eder.  $i_t$ ,  $f_t$ ,  $O_t$ ,  $C_t$  ve  $H_t$ 'nin hesaplanması sırasıyla Eşitlik (3.1), Eşitlik (3.2), Eşitlik (3.3), Eşitlik (3.4) ve Eşitlik (3.5)'te verilmiştir. Burada  $\sigma$  sigmoid aktivasyon fonksiyonu ve  $\tan h$  hiperbolik tanjant aktivasyon fonksiyonudur. Yeni  $C_t$  belleği ve  $H_t$  çıktısı, önceki  $C_{t-1}$  belleğinin mevcut  $X_t$  girişi ve önceki  $H_{t-1}$  çıktısı ile güncellenmesiyle oluşturulur.

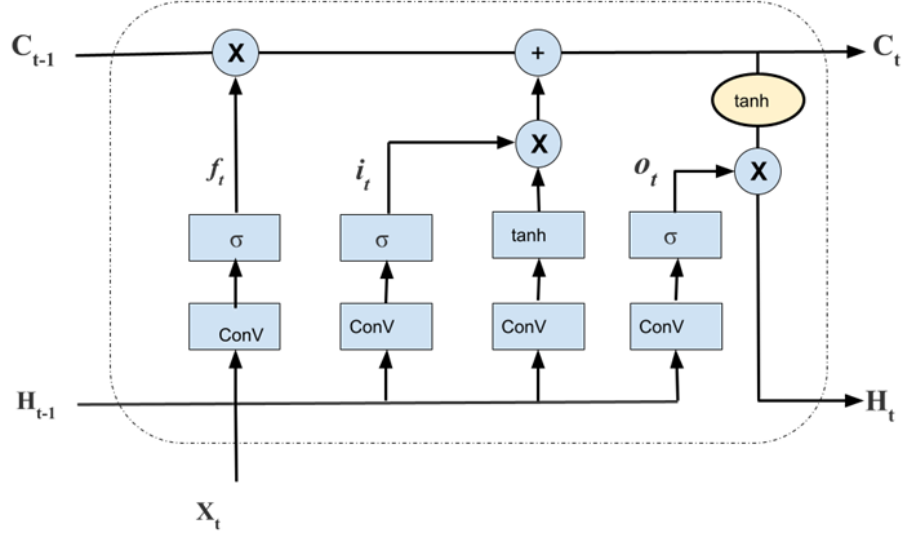
$$i_t = \sigma(W_{xi} * X_t + W_{hi} * H_{t-1} + b_i) \quad (3.1)$$

$$f_t = \sigma(W_{xf} * X_t + W_{hf} * H_{t-1} + b_f) \quad (3.2)$$

$$O_t = \sigma(W_{xo} * X_t + W_{ho} * H_{t-1} + b_o) \quad (3.3)$$

$$C_t = f_t \circ C_{t-1} + i_t \circ \tan h(W_{xc} * X_t + W_{hc} * H_{t-1} + b_c) \quad (3.4)$$

$$H_t = O_t \circ \tan h(C_t) \quad (3.5)$$



Şekil 3.6. ConvLSTM mimarisinin blok diyagramı (Shi vd., 2015)

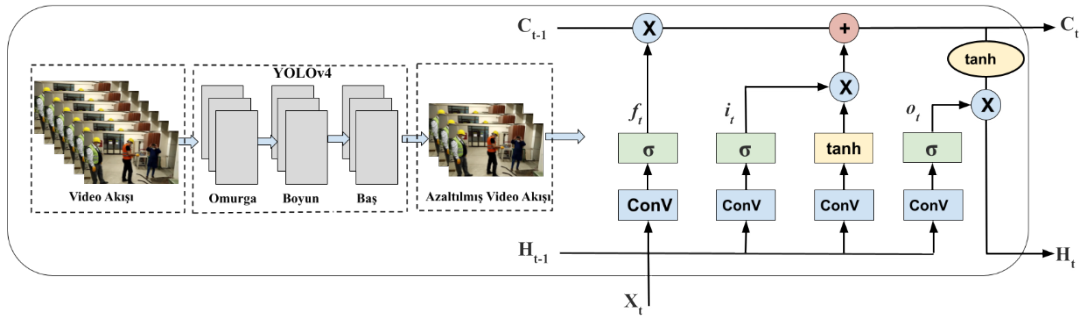
Kamerallardan veya görüntüleme cihazlarından gelen veriler genellikle bir zaman dizisi olarak temsil edilir. Bu nedenle, bu tür durumlarda LSTM tipi bir model daha uygundur. Bu mimari, modelin bir önceki gizli durumu dizinin bir sonraki adımına aktarmasına olanak tanıyarak bilgiyi tutmasını ve daha önce işlediği verilere dayanarak karar vermesini sağlar (Shi vd., 2015). ConvLSTM, görüntü modellerini hareket zamanında çalıştırmak için maksimum iş yükü ile nispeten küçük görüntü veri kümelerinden öğrenmeyi sağlayan bir mekanizmaya sahiptir. ConvLSTM birimleri genellikle geleneksel tekrarlayan sinir ağlarından (RNN) daha verimli çalışır ve giriş özelliklerini daha iyi kullanır. Bu nedenle ConvLSTM, gerçek zamanlı video görüntülerini analiz etmek ve sınıflandırmak için özellikle güçlü bir yaklaşımdır.

### 3.3.2. YOLOv4 + ConvLSTM

Aynı videoda sadece bir eylem olması sağlanarak ConvLSTM ile yapılan sınıflandırmanın doğruluğu artırılmış ve eylem tespit süresi azaltılmıştır. Unsafe-Net mimarisinde, video anlama için belirli bir zaman adımında uzamsal özellikleri çıkarmak için bir CNN oluşturulmuş, ardından çerçeveler arasındaki zamansal ilişkileri tanımlamak için bir LSTM ağı oluşturulmuştur. Ayrıca, iki farklı IP kameradan alınan canlı video karelerinin önceden eğitilmiş YOLOv4+ConvLSTM ağına uygulandığı ve dört farklı sınıfa ait güvensiz davranışlar tespit edildiğinde, çalışmada tasarlanan elektronik uyarı sisteminin devreye girdiği ve kural ihlali yapan çalışanların her seferinde sesli ve görsel olarak uyarıldığı bir pilot çalışma iş sahasında gerçekleştirilmiştir. Son olarak, geliştirilen sistemin fabrikanın çalışma ortamına

kurulmasından önce ve sonra çekilen gözetim videoları analiz edilerek çalışmanın etkinliği değerlendirilmiştir.

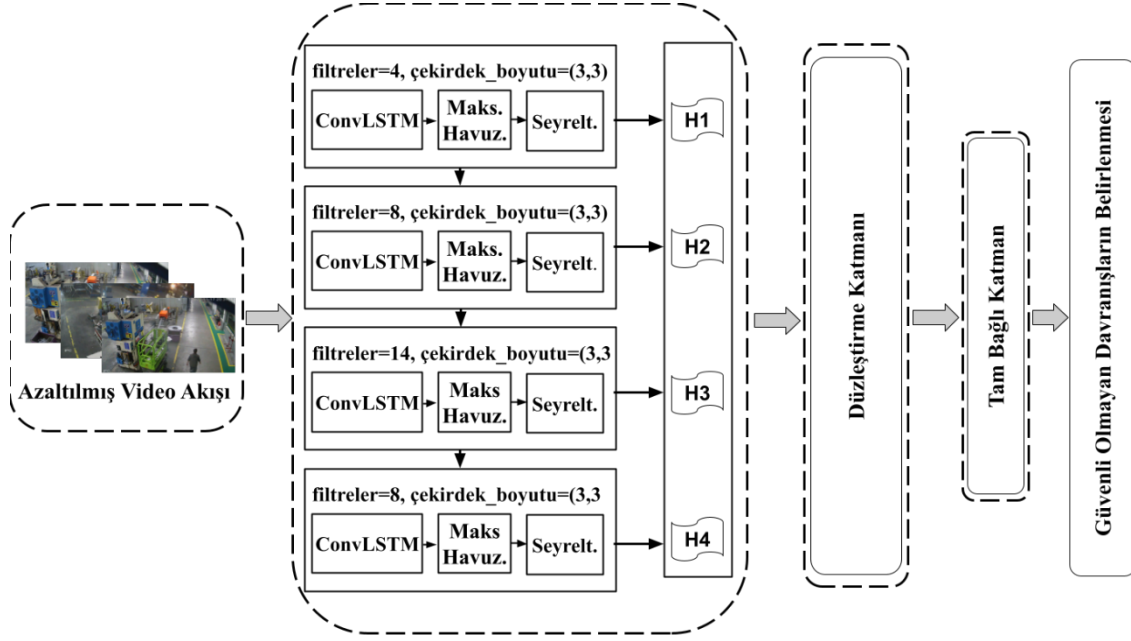
Temel bir LSTM yapısı, yalnızca tek boyutlu girdi aldığı için kendini değiştiren zamansal verileri modellemek için uygun değildir. Bu çalışmada, YOLOv4 kullanılarak videolarda güvenli olmayan davranışlar içeren kareler belirlendikten sonra ConvLSTM modeli kullanılmıştır. Model, etiketlenmiş nesnelere içeren videodan kare dizilerinin belirli bir zaman adımında uzamsal özellikleri çıkaran bir CNN'den ve ardından kareler arasındaki zamansal ilişkileri sınıflandıran bir LSTM ağından oluşur. Önerilen YOLOv4 ve ConvLSTM derin öğrenme mimarileri, endüstriyel operasyonların gerçek zamanlı olarak tespit ve analiz edilmesini sağlayarak videolardaki güvensiz davranışları tespit etme doğruluğunu artırırken modelin ağırlığını da azaltmaktadır. Şekil 3.7.'de Unsafe-Net hibrit derin öğrenme modeli yaklaşımının altyapısı gösterilmektedir.



**Şekil 3.7.** Güvensiz davranış tespiti için YOLOv4 ve ConvLSTM tabanlı önerilen Unsafe-Net derin öğrenme modelinin mimarisi

Önerilen model, YOLOv4 kullanarak video dizilerinden güvenli olmayan davranışlar içeren olası kareleri tanımlar. Bu kareler daha sonra CNN ve LSTM'in bir kombinasyonu olan ConvLSTM ağı tarafından analiz edilir. ConvLSTM hücresi, ağıdaki konvolüsyonel işlemleri içeren bir katmandır ve mimariye gömülü bir konvolüsyonel LSTM'dir. Dolayısıyla bu çalışma, verilerin mekânsal özelliklerini tespit ederken zamansal ilişkiyi de dikkate almaktadır. Bu yaklaşım, video sınıflandırması için bireysel çerçevelerdeki uzamsal ilişkiyi ve farklı çerçevelerdeki zamansal ilişkiyi etkili bir şekilde yakalar. ConvLSTM ağı, tek bir LSTM ağından farklı olarak konvolüsyonel yapısı sayesinde genişlik, yükseklik ve kanal sayısı dahil olmak üzere 3B girdileri alabilmektedir. Şekil 3.8.'de işyerindeki güvensiz davranışları tespit etmek için önerilen Unsafe-Net öğrenme modelindeki ConvLSTM mimarisinin iç yapı diyagramını göstermektedir. Mimari, YOLOv4 ağından Güvenli Yürüyüş Yolu İhlali, Yetkisiz

Müdahale, Açık Pano Kapağı ve Forkliftle Aşırı Yük Taşıma gibi güvensiz eylemleri içeren videoları girdi olarak almaktadır. Geliştirilen yaklaşım ConvLSTM, Maksimum Havuzlama ve Düşürme (Dropout) bileşenlerinden oluşan 4 katman kullanmaktadır. Konvolüsyonel sinir ağı, filtre boyutları sırasıyla 4, 8, 14 ve 16 olan dört katmandan oluşmaktadır. Çekirdek boyutu (3, 3), havuz boyutu (1, 2, 2) ve bırakma oranı 0,2 olarak ayarlanmıştır. Kullanılan aktivasyon fonksiyonu *tanh*'dir. Güvenli olmayan davranışları tanımlamak için ağıın çıkışına düz (flatten) ve yoğun katmanlar (dense) eklenmiştir.



**Şekil 3.8.** Güvensiz davranışların tespiti için önerilen YOLOV4+ConvLSTM derin öğrenme modelinde bulunan ConvLSTM yapısının mimarisinin blok diyagramı

## 4. DENEYSEL ÇALIŞMALAR

Bu tez çalışmasında, çalışma ortamlarında KKD tespiti yönelik ve gerçek zamanlı olarak güvensiz davranışların tespiti için geliştirilen bilgisayarla görme sistemi (Unsafe-Net) için deneysel analizler yürütülmüştür. Verisetlerinde elde edilen verileri eğitmek ve yürütülen deneysel çalışmalar için 2 tane NVIDIA Geforce GTX 1080Ti 11GB grafik işlem birimi (Graphics Processing Unit, GPU), Intel I9 9900 merkezi işlemci (CPU), 16 GB RAM bellek, 3 TB hard disk (HDD) ve 500 GB katı hal disk (SSD) bileşenlerine sahip bir iş istasyonu kullanılmıştır

### 4.1. KKD Tespitine Yönelik Yürütülen Deneysel Analizler

Çalışma ortamlarında KKD tespiti için yürütülen deneysel çalışmalarda öncelikle 1000 görüntü etiketlenmiş ve bu etiketlenen görüntüler düzenlenerek YOLOv4 algoritması ile eğitilmiştir. Daha sonra etiketli görüntü sayısı 2 kat daha artırılarak ağ birkaç kez daha eğitilmiş ve gerekli en uygun (optimum) ağırlıklar elde edilmiştir. YOLOv4 ağının çeşitli parametrelere göre gerçekleştirilen farklı eğitimlerde elde edilen eğitim süreleri ve diğer bilgiler Tablo 4.1.'de verilmiştir. Ağın eğitim süresi üzerinde etkili olan birçok parametre vardır. Eğitim süresi bakımından gösterilen parametrelerden en önemlisi GPU kullanımı ve sayısıdır. YOLO algoritmasının sahip olduğu katman sayısının çok olması güçlü donanımlara olan ihtiyacı artırmaktadır ve dolayısıyla GPU kullanılmadan eğitim yaptırmak çok zor hale gelmektedir. Hatta eğitim sürelerini kısaltabilmek birden fazla GPU kullanımı sayesinde mümkündür. Bu çalışmada, YOLOv4 ağının eğitimi için 2 adet NVIDIA GeForce 1080Ti grafik kartı kullanılmıştır. Tablo 4.1.'de görülebileceği gibi, eğitimler aynı parametreler kullanılarak tek GPU ve çift GPU ile yürütülerek, elde edilen eğitim süreleri karşılaştırılmıştır. Bulunan sonuçlara göre, YOLO ağının eğitiminde çift GPU kullanılırsa, eğitim süresi yaklaşık %30 civarlarında kısalmaktadır. Ayrıca, bu tabloda gösterilen eğitim öncesi konfigürasyon dosyasında değerleri ayarlanabilen görüntünün boyutu, (Subdivision), rastgele değerleri (Random) ve iterasyon sayısı hem eğitim süresini hem de mAP değerini önemli ölçüde etkilemektedir.

**Tablo 4.1.** Çalışma ortamlarından nesne tanıma için önerilen YOLOv4 ağıının eğitim parametreleri ve süreleri

Eğitim No	Eğitim Kümesindeki Görüntü Sayısı	Doğrulama İçin Ayrılan Görüntü Sayısı	Sınıf Sayısı	İterasyon Sayısı	GPU Sayısı	Görüntü Boyutu	Rastgele (Random)	Küme boyutu (Batch size)	Alt bölümler (Subdivision)	Eğitim Süreleri (dk)
1	2670	0	3	6000	1	416x416	0	64	324	403'
2	2670	0	3	6000	2	416x416	0	64	32	246'
3	888	112	7	15000	2	416x416	0	64	32	845'
4	888	112	7	15000	1	416x416	0	64	32	1210'
5	888	112	7	15000	2	416x416	0	64	16	605'
6	888	112	7	15000	2	416x416	0	64	32	700'
7	888	112	7	15000	2	416x416	1	64	32	840'
8	1374	148	7	16000	2	608x608	0	64	32	1504'
9	2006	203	7	18000	2	416x416	0	64	16	746'
10	2000	178	7	14000	1	416x416	1	64	32	1440'

Çalışma ortamlarında KKD tespitinde yürütülen deneysel analizler için önerilen YOLOv4 ağı ile çalışma ortamlarında nesne tanıma işlemlerinde elde edilen başarıyı ölçmek için Kesinlik (Precision, P), Hatırlatma (Recall, R), F1 skoru (F1-score, F1), Ortalama Kesinlik (Average Precision, AP), Ortalama kesinlik değerlerinin ortalaması (Mean Average Precision, mAP) ve IoU metriklerinin ölçümleri yapılmıştır. Bu metrikler sırasıyla Eşitlik (4.1), Eşitlik (4.2), Eşitlik (4.3), Eşitlik (4.4), Eşitlik (4.5) ve Eşitlik (4.6)'de gösterilmiştir. Kesinlik pozitif olarak tahmin edilen değerlerin, kaç tanesinin gerçekte pozitif olduğunu ifade etmektedir. Hatırlatma ise pozitif olarak tahmin edilmesi gereken verilerin, ne kadarının pozitif olarak tahmin edildiğini gösteren bir metriktir. F1-skoru ölçütü kesinlik ve hatırlatma skorlarının harmonik ortalamasını göstermektedir. mAP sınıfların ortalama kesinlik değerlerinin ortalaması alınarak hesaplanmaktadır. IoU hesabı, iki dikdörtgenin kesiştikleri alanın iki dikdörtgenin bileşiminin alanına bölümü olarak ifade edilmektedir. IoU hesabında, B tahmin edilen değeri,  $B_{gt}$  ise referans değeri temsil etmektedir.

$$\text{Kesinlik (P)} = \frac{DP}{DP + YP} \quad (4.1)$$

$$\text{Hatırlatma(R)} = \frac{DP}{DP + YN} \quad (4.2)$$

$$\text{F1 - skoru (F1)} = 2 \times \frac{P \times R}{P + R} \quad (4.3)$$

$$\text{Ortalama Kesinlik (AP)} = \int_0^1 p(r) dr \quad (4.4)$$

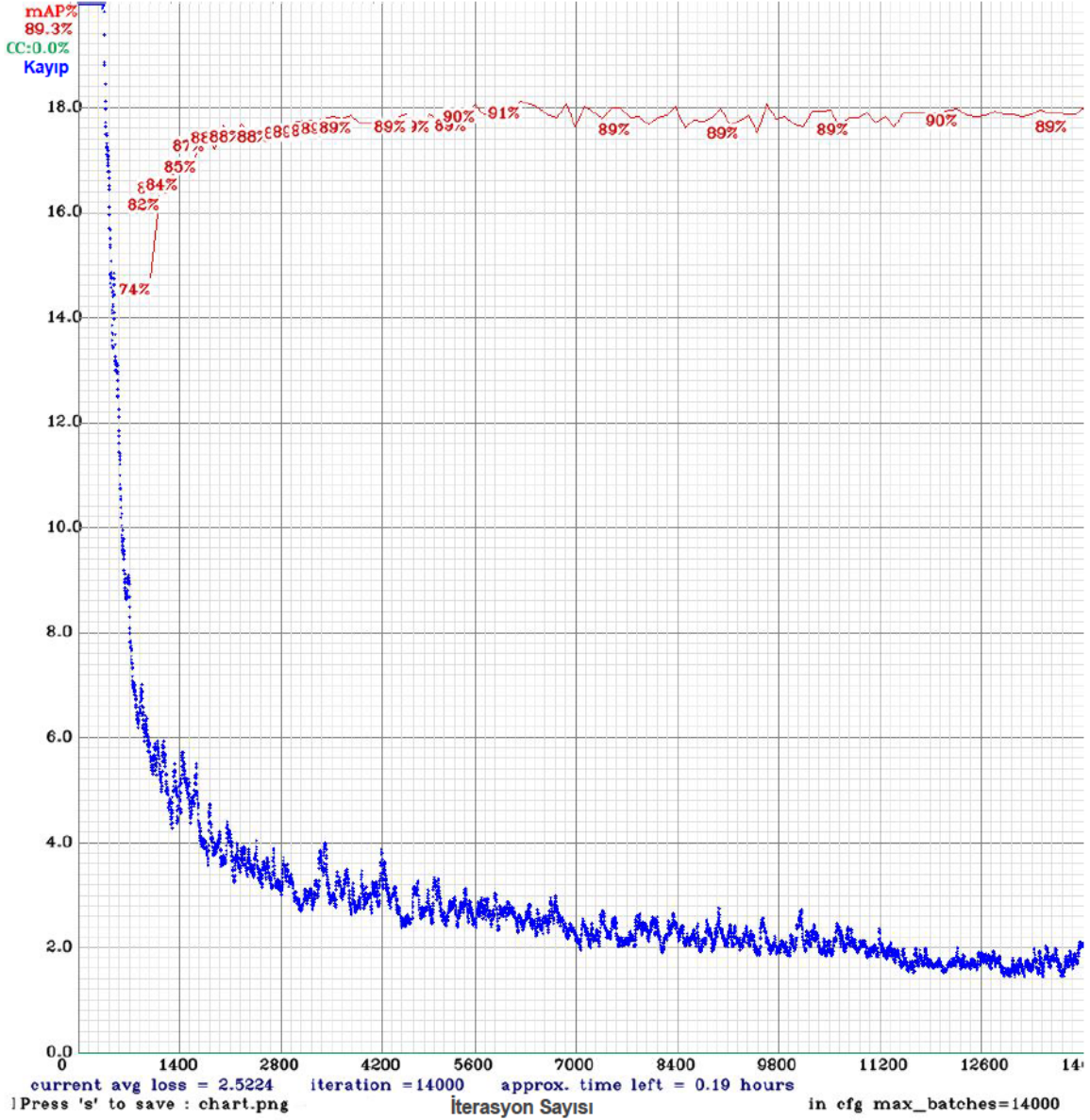
$$\text{Genel Ortalama Kesinlik (mAP)} = \frac{1}{n} \sum_{i=1}^n AP_i \quad (4.5)$$

$$\text{Birleşim Kesişimleri (IoU)} = \frac{|B \cap B_{gt}|}{|B \cup B_{gt}|} \quad (4.6)$$

KKD tespitinde nesne tanıma için kullanılan YOLOv4 algoritması farklı büyüklükteki veri kümeleri ile eğitilip, ağın başarımı test veri kümesi ile ölçülmüştür. Sonuçlar mAP, P, R, F1, doğru pozitif (DP), yanlış pozitif (YP), yanlış negatif (YN), IoU ve Ortalama Kayıp parametreleri ile kıyaslanmış ve Tablo 4.2.'de sunulan değerler elde edilmiştir. Bu tablodaki sonuçlar (3,4,5,6,7,8,9) numaralı eğitimlerde IoU değeri için 0.5 eşik değeri belirlenerek elde edilmiştir. Tablo 4.2. incelendiğinde, en yüksek mAP değeri 91.18 ile 5 numaralı eğitimde elde edilmiştir. Önerilen YOLOv4 ağı için 5 numaralı eğitimi için ortalama kayıp ve mAP değerlerinin iterasyon sayısına göre değişimi Grafik 4.1.'de gösterilmiştir. YOLOv4 ağı için yapılan 5 numaralı eğitimde diğer ölçüm metrikleri olan P, R, F1, DP, YP, YN, IoU ve Ortalama Kayıp için sırasıyla 0.89, 0.91, 0.90, 665, 79, 68, 70.35 ve 1.1147 değerleri başarılmıştır. Bu eğitimde, 888 eğitim ve 112 doğrulama görüntüsüne sahip veri kümesi ile YOLOv4 algoritmasının eğitilmesi sağlanmıştır. Bu grafik detaylıca incelendiğinde, 4500 iterasyona kadar hızlı bir şekilde öğrenme işleminin gerçekleştiği, kaybın dramatik bir şekilde düştüğü gözlemlenirken, bu noktadan sonra gözle görülür bir iyileşme olmadığı görülmektedir. Ayrıca, yaklaşık 10000 iterasyondan sonra ise mAP değerinin de daha kararlı bir hale geldiği ve son 2000 iterasyonda da bu kararlılığın iyice oturduğu görülmektedir.

**Tablo 4.2.** Farklı eğitimler için YOLOv4 modeli ile elde edilen performans metrik sonuçların karşılaştırılması

Eğitim	mAP (@0.50)	P	R	F1	DP	YP	YN	IoU	Ort. Kayıp
3	85.29	0.76	0.84	0.80	619	196	114	57.15	1.5283
4	86.50	0.80	0.84	0.82	619	155	114	61	0.9543
5	91.18	0.89	0.91	0.90	665	79	68	70.35	1.1147
6	83.98	0.79	0.83	0.81	610	161	123	59.88	1.4929
7	90.63	0.89	0.90	0.89	661	84	72	70.54	1.5082
8	83.17	0.79	0.84	0.81	618	169	115	58.10	1.5425
9	76.54	0.68	0.79	0.73	580	269	153	50.36	1.8101



**Grafik 4.1.** Önerilen YOLOv4 ağıının 5 numaralı eğitim için ortalama kayıp ve mAP değerlerinin iterasyon sayısına göre değişimi

Tablo 4.3.'te IoU için 0.5 eşik değerinde, en yüksek mAP değerinin elde edildiği 5 numaralı eğitim sonrasında, YOLOv4 ağı ile yapılan nesne tanıma işleminde, çalışma ortamlarında bulunan Baret, Eldiven, Gözlük, Maske, Baret-Yok, Maske-Yok ve Yelek sınıflarının her birisi için elde edilen AP, DP ve YP değerleri verilmiştir. Bu tablodan da görülebileceği gibi, önerilen yöntem ile nesnelerin tanınmasında yüksek başarımlar elde edilmiştir. Tablo 4.3.'te görülebileceği gibi, bazı sınıflar için (baret ve maske) hem nesnenin varlığı hem de olmaması ayrı birer sınıf olarak değerlendirilmiştir. Bunun nedeni, çalışma ortamlarında bazı KKD'lerin olması önemli olmadığı gibi, olmaması da hesaba katılması

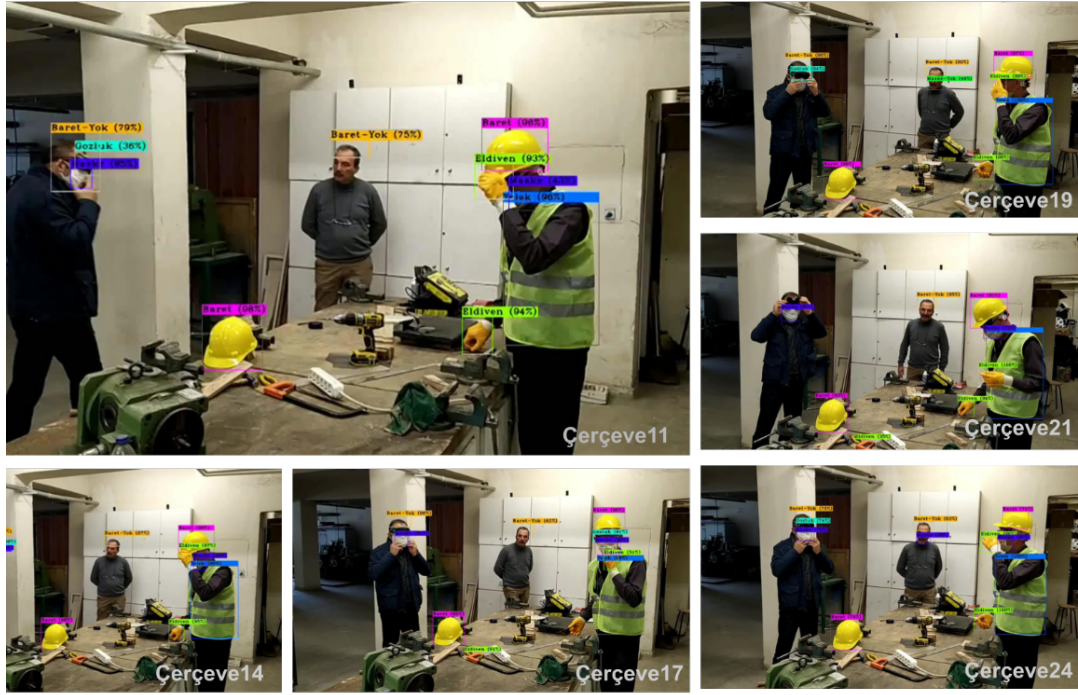
gereken bir durumdur. İşlem kolaylığı ve sonuçların daha hassas olması için bu iki sınıfta nesnenin olmaması durumu da değerlendirilmeye alınmıştır.

**Tablo 4.3.** YOLOv4 ağı ile 5 numaralı eğitimde her bir sınıf için elde edilen AP, DP ve YP değerleri

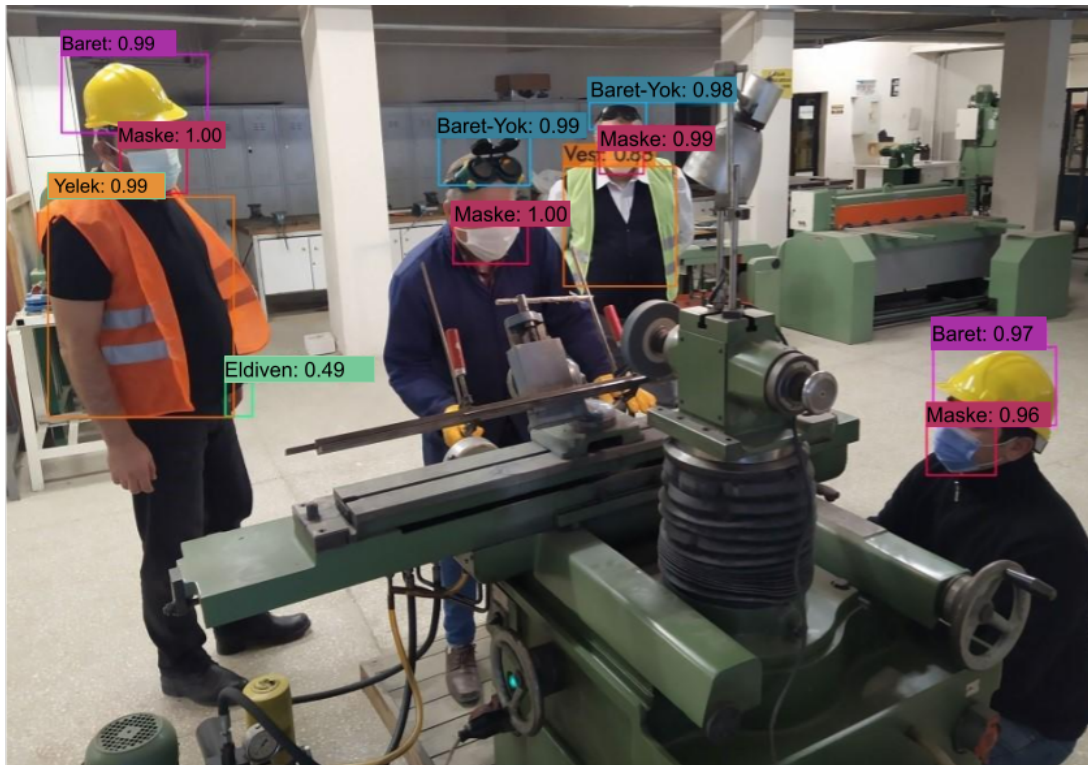
Sınıf Adı	AP (%)	DP	YP
Baret	97.62	90	2
Eldiven	96.80	122	4
Gözlük	92.97	61	10
Maske	90.35	115	14
Baret Yok	78.71	138	41
Maske Yok	82.69	18	5
Yelek	99.09	121	3

Şekil 4.1.'de, çalışma kapsamında hazırlanan ve eğitim aşamasında kullanılmayan verisetindeki bir videodan elde edilen çerçevelerde, önerilen YOLOv4 algoritması kullanılarak çalışma ortamlarındaki nesnelere (KKD) tespiti için yürütülen test çıktısı gösterilmiştir. Buradan çalışma ortamındaki nesnelere YOLOv4 algoritması ile başarılı olarak tanındığı ve doğruluk oranının da nesne etiketi olarak işaretlendiği görülmektedir. Bu videodaki test amacı ile kullanılan görüntüler etiketlenmemiş ve eğitime dahil edilmemiş, diğer görüntülerden farklı bir zamanda farklı açılar ve arka plan kullanılarak elde edilmiştir.

Şekil 4.2.'de ise YOLOv4 ağının eğitiminde kullanılmamış ve iş ortamından alınan durağan bir görüntüde, önerilen yöntem ile KKD tespitinin sonuçları görülmektedir. İş sahasında eldiven nesnesi hariç diğer tüm nesnelere tanınmasında %88'in üzerinde doğruluk sağlanmıştır.



Şekil 4.1. Önerilen YOLOV4 modeli ile video çerçeveleri üzerinde çalışma ortamında KKD tespitinin yapılması



Şekil 4.2. İş sahasından alınan durağan bir görüntüde önerilen YOLOv4 mimarisi ile KKD tespitinin yapılması

IoU için 0.75 eşik değerinde 5 numaralı eğitim sonrasında oluşan ağırlıklar kullanılarak yapılan deneysel çalışmada, YOLOv4 ağı kullanılarak çalışma ortamındaki nesnelere tanınmasında (KKD tespitinde) elde edilen AP, DP ve YP değerlerinin sonuçları Tablo 4.4.'te sunulmuştur. Bu tablodan da görülebileceği gibi, IoU için eşik değerinin yükselmesinden dolayı, ölçüm metriklerinin başarımının düşmektedir. mAP değerinin %53.88 olarak elde edildiği bu performans başarım ölçümünde, nesne sınıfları için bulunan DP değerleri düşerken, YP değerleri artmaktadır. Burada P, R ve F1 skorları ise sırasıyla 0.62, 0.63 ve 0.62 olarak hesaplanmıştır.

**Tablo 4.4.** YOLOv4 ağı ile 5 numaralı eğitimde IoU eşik değeri 0.75 için her bir sınıf ait AP, DP ve YP değerleri

Sınıf Adı	AP (%)	DP	YP
Baret	81.39	78	14
Eldiven	55.77	85	41
Gözlük	46.59	37	34
Maske	47.57	70	59
Baret Yok	31.75	66	113
Maske Yok	20.69	8	15
Yelek	93.43	115	9

#### 4.2. Gerçek Zamanlı Olarak Güvensiz Davranışların Tespiti İçin Geliştirilen Bilgisayarla Görme Sisteminde (Unsafe-Net) Yürütülen Deneysel Analizler

Tez çalışmasının bu bölümünde, toplanan videolardan fabrika çalışma ortamlarındaki güvensiz davranışları sınıflandırmak ve belirlemek için derin öğrenmeye dayalı gerçek zamanlı bir bilgisayarla görme sistemi olan Unsafe-Net için yürütülen deneysel çalışmalara değinilmiştir. Bu çalışma, çalışma ortamlarındaki güvensiz davranışları gerçek zamanlı olarak tespit edebilen bir sistemin geliştirilmesini sunmaktadır. Sistem, derin öğrenme tabanlı bir yaklaşım olan Unsafe-Net'i kullanarak çalışanlara anında görsel ve sesli uyarılar sağlamaktadır. Çalışmada, endüstriyel bir alanda bulunan bir fabrikadaki işçilerin güvenli ve güvensiz davranışlarını sınıflandırmak için YOLOv4 ve ConvLSTM algoritmaları kullanılmıştır. Çalışma sahasındaki davranışları içeren video sahne kareleri ConvLSTM ağına diziler halinde sunularak sistemin hızlı ve doğru bir şekilde çalışması sağlanmıştır.

Unsafe-Net mimarisinin performansını değerlendirmek için temel performans ölçümleri kullanılmıştır. Bu anahtar metriklerdeki DP, doğru negatifler (DN), yanlış pozitifler (YP) ve yanlış negatifler (YN) için değerler karışıklık matrisi tablosundan elde edilmiştir. Deneysel

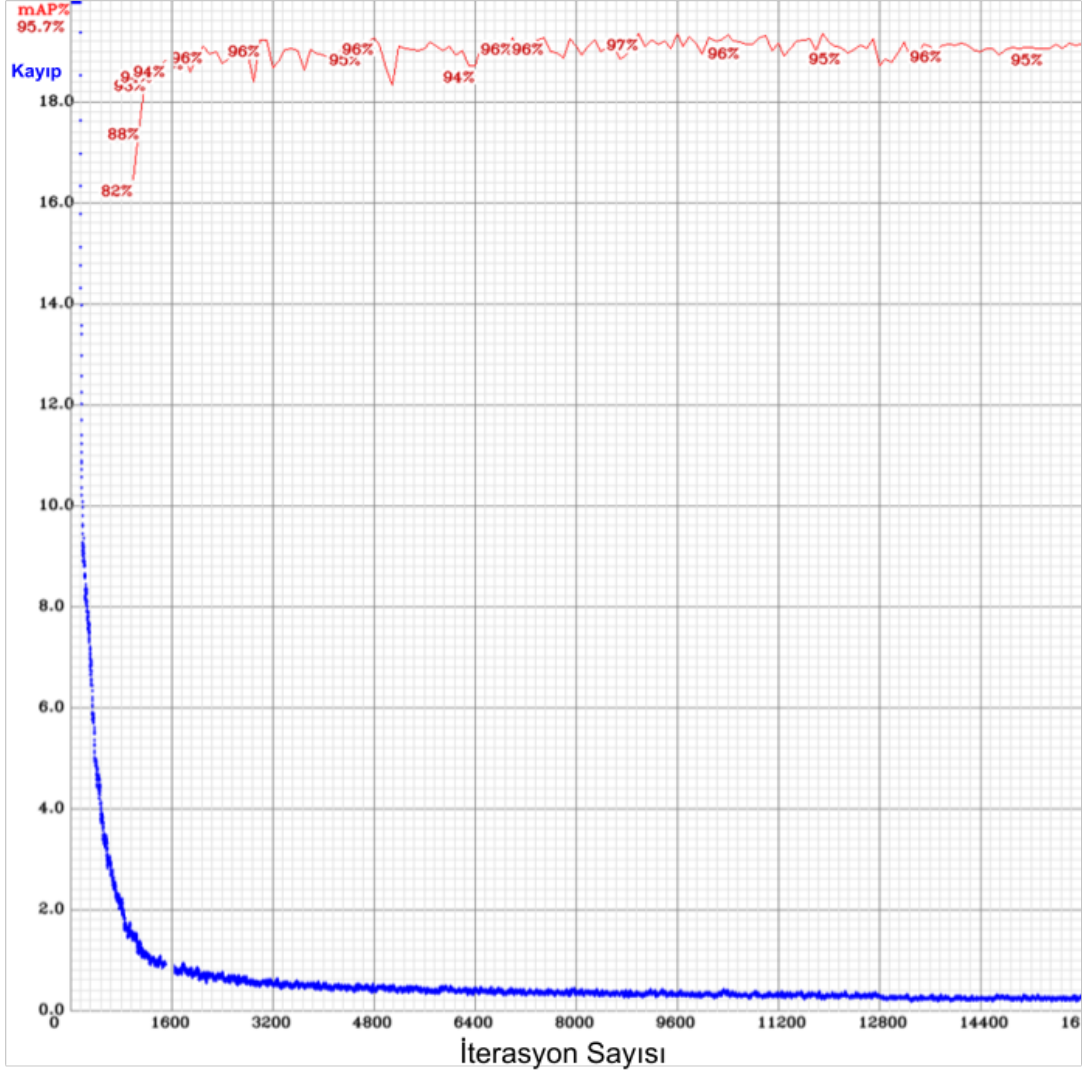
çalışmalar, önerilen YOLOv4 ağı ile çalışma ortamlarındaki güvenli ve güvensiz davranışların tespitinde elde edilen başarıyı Ortalama Kesinlik (AP), Genel Ortalama Kesinlik (mAP) ve IoU metriklerini kullanarak ölçmüştür. Kesinlik (P), DP olarak tahmin edilen değerlerin toplam pozitif değerlere oranıdır. Hatırlama (R) ise DP olarak tahmin edilen değerlerin pozitif olduğu bilinen değerlere oranıdır. AP, P ve R eğrilerinin altındaki toplam alandır. IoU, tahmin ve temel gerçeğin kesişme alanının birleşme alanına bölünmesiyle elde edilir. mAP, sınıfların AP değerlerinin ortalaması alınarak ölçülür. Unsafe-Net mimarisi, güvensiz hareketlerin sınıflandırma sonuçlarını değerlendirmek için Doğruluk (Accuracy) metriğini kullanır. mAP ve Doğruluk için metrikler sırasıyla Eşitlik (4.7) ve (4.8)'de verilmiştir.

$$\text{Genel Ortalama Kesinlik (mAP)} = \frac{1}{n} \sum_{i=1}^n AP_i \quad (4.7)$$

$$\text{Doğruluk (\%)} = \frac{TP+TN}{DP+TN+YP+YN} * 100 \quad (4.8)$$

#### 4.2.1. YOLO Mimarisi ile Uygun Çerçeve Seçimi

Video veriseti hazırladıktan sonra YOLOv4 kullanarak video parçalarından güvenli ve güvensiz davranışların nesne tanınmasını gerçekleştirilmektedir ve sadece güvensiz davranışlar ile ilgili çerçeveler seçilmektedir. YOLOv4 vasıtasıyla, videoları ConvLSTM ile sınıflandırmaya hazırlamak için Güvenli Yaya Yolu İhlali, Yetkisiz Müdahale, Açık Pano Kapağı ve Forklift ile Aşırı Yük Taşıma güvenli olmayan davranış sınıflarına ait kareleri seçilir. Güvenli davranış sınıflarına ait olan Güvenli Yürüyüş Yolu, Yetkili Müdahale, Kapalı Pano Kapağı ve Güvenli Taşıma hareketlerini içeren kareleri kaldırılır. Ayrıca, güvenli olmayan davranışlar içermeyen diğer parçaları da çıkarılır. Grafik 4.2., eğitim setindeki 2262 görüntü kullanılarak YOLOv4 ile eğitim yapılırken iterasyon süresi boyunca ortalama kayıp ve mAP değerlerindeki değişimini göstermektedir. İterasyon süresi boyunca ortalama mAP değeri yaklaşık %97'dir ve Kayıp değeri sıfıra yakındır. Dolayısıyla YOLOv4 eğitiminin başarılı olduğu sonucuna varılabilir. YOLOv4 eğitimi sırasında görüntüler 416x416 olarak yeniden boyutlandırılmış ve model 8 sınıf üzerinde eğitilmiştir. Yığın boyutu (batch size) 32, alt bölümlenme 64 olarak ayarlanmış ve model 2 GPU üzerinde eğitilmiştir.



**Grafik 4.2.** Güvensiz davranışların tespiti için YOLOv4'ün eğitimi esnasında Loss ve mAP değerlerinin iterasyon sayısına göre değişimi

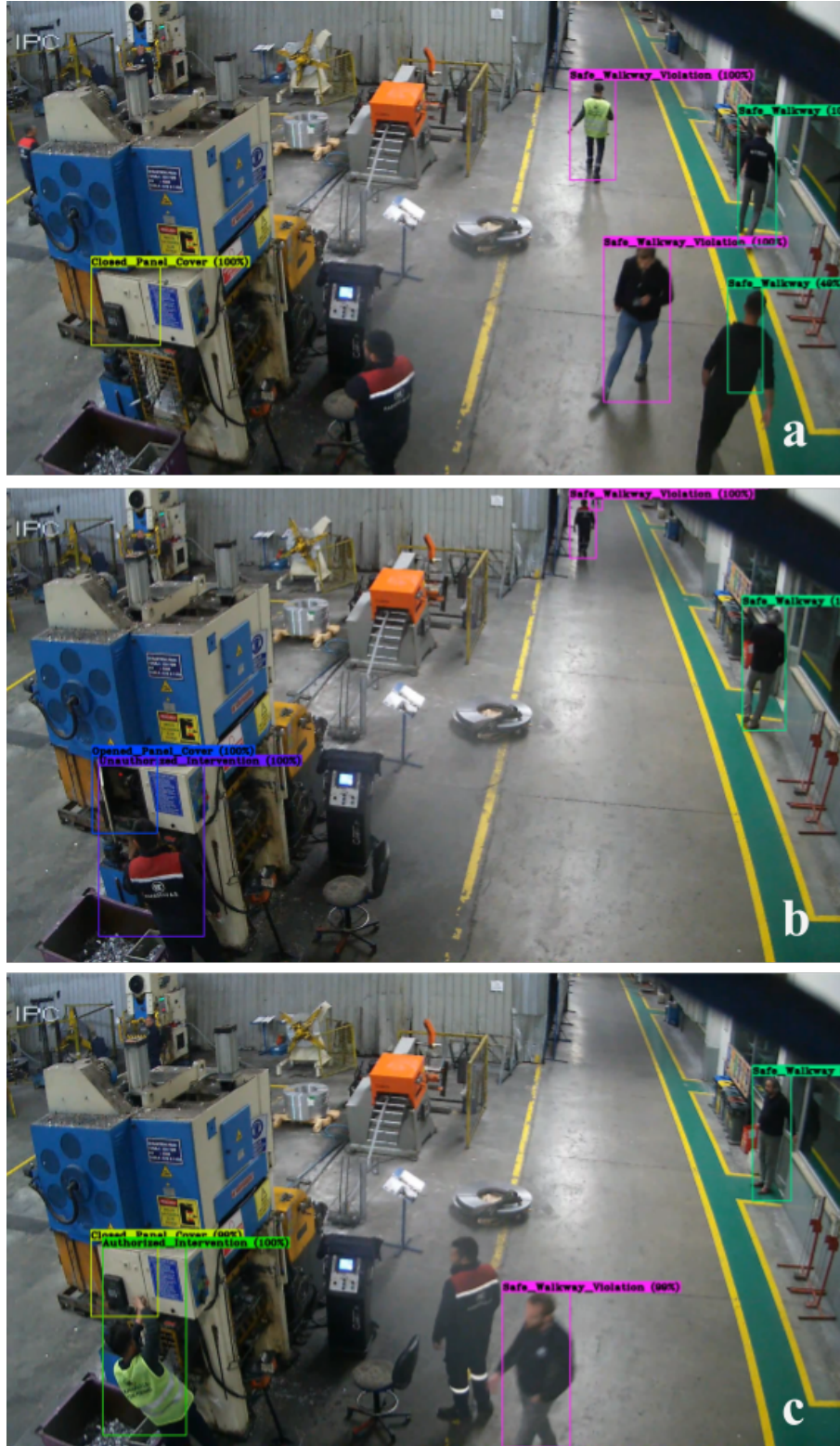
YOLOv4, 329 görüntü seti kullanılarak doğrulanmış ve test edilmiştir. Güvenli Yaya Yolu İhlali, Yetkisiz Müdahale, Açık Pano Kapağı ve Forkliftle Aşırı Yük Taşıma sınıfları için AP sonuçlarının yanı sıra DP ve YP değerleri IoU için @0.5 eşiği için elde edilmiş ve Tablo 4.5.'te gösterilmiştir. Buradan, YOLOv4'ün güvenli olmayan davranışlar sergileyen nesnelere etkili bir şekilde tespit ettiği sonucuna varılabilir. Ayrıca, YOLOv4 test süreci 329 görüntü için toplam 14 saniye sürmüştür. Ek olarak, önerilen YOLOv4 yöntemi test setinde %95.66 mAP skoru elde etmiştir. Ayrıca, test kümesindeki Güvenli Taşıma sınıfının yalnızca 11 doğru pozitif ve 8 yanlış pozitif ile %86.30'luk en düşük AP puanına sahip olduğunu göstermektedir. Bu durum, 3 veya daha fazla blokla çalışırken Forkliftle Aşırı Yük Taşıma sınıfındaki nesne tespiti

ile 2 veya daha az blokla çalışırken Güvenli Taşıma sınıfındaki nesne tespiti arasındaki benzerliğe bağlanabilir.

**Tablo 4.5.** YOLOv4 ağı ile test işleminde her bir sınıf için elde edilen AP skoru ve DP ile YP değerleri

Güvenli/Güvensiz Davranış	AP (%)	DP	YP
Güvenli Yürüyüş Yolu İhlali	97.89	97	12
Yetkisiz Müdahale	95.83	23	0
Açık Pano Kapağı	96.92	63	7
Forklift ile Aşırı Yük Taşıma	97.50	40	2
Güvenli Yürüyüş Yolu	90.82	36	5
Yetkili Müdahale	100	12	0
Kapalı Pano Kapağı	100	41	0
Güvenli Taşıma	86.30	11	8
<b>Mean (ortalama) AP (mAP)=</b>	<b>%95.66</b>		

Şekil 4.3.'te YOLOv4'ün test seti videolarındaki güvenli ve güvensiz davranış sınıfları için bazı nesne algılama sonuçlarını göstermektedir. Şekil 4.3(a), güvenli davranış sınıfları Yetkili Müdahale, Kapalı Pano Kapağı ve Güvenli Yaya Yolu ile güvenli olmayan davranış sınıfı Güvenli Yaya Yolu İhlaline ait kareleri içeren bir videoyu göstermektedir. Şekil 4.3(b), güvenli davranış sınıfı Yetkili Müdahale ve güvenli olmayan davranış sınıfı Açık Pano Kapağı olan bir videodan kareleri göstermektedir. Bu çalışma, güvenli ve güvenli olmayan davranışlarla ilişkili nesnelere tespit etmek ve Unsafe-Net mimarisinde ConvLSTM için indirgenmiş video segmentleri oluşturmak için YOLOv4'ü kullanmaktadır. Şekil 4.3(c)'de gösterilen videoya yalnızca güvensiz sınıf davranışları olan Açık Pano Kapağı ve Güvenli Yaya Yolu İhlali'ne ait kareler dahil edilirken alakasız kareler hariç tutulmuştur. Tek bir eylem için tek bir güvensiz sınıf davranışının çerçevelerini içeren video parçaları oluşturmuştur ve ConvLSTM ağına gönderilmiştir. Örneğin, Şekil 4.4'te 40 karelik bir videodan kareler gösterilmektedir. Video 3.5 saniye sürmekte ve 105 kare içermektedir. Bu video YOLOv4 ağından geçirildikten ve alakasız kareler çıkarıldıktan sonra geriye sadece 40 kare kalmaktadır ve videonun süresi 1.33 saniyeye düşmektedir. Video daha sonra ConvLSTM ağına sonraki aşamalar için Güvenli Yaya Yolu İhlali, Açık Pano Kapağı ve Yetkisiz Müdahale güvensiz davranış sınıfları dahil olmak üzere üç parça halinde iletilmektedir ve her videoda yalnızca bir eylem bulunmaktadır. Bu prosedürler test setindeki her videoya eşit olarak uygulanmaktadır.

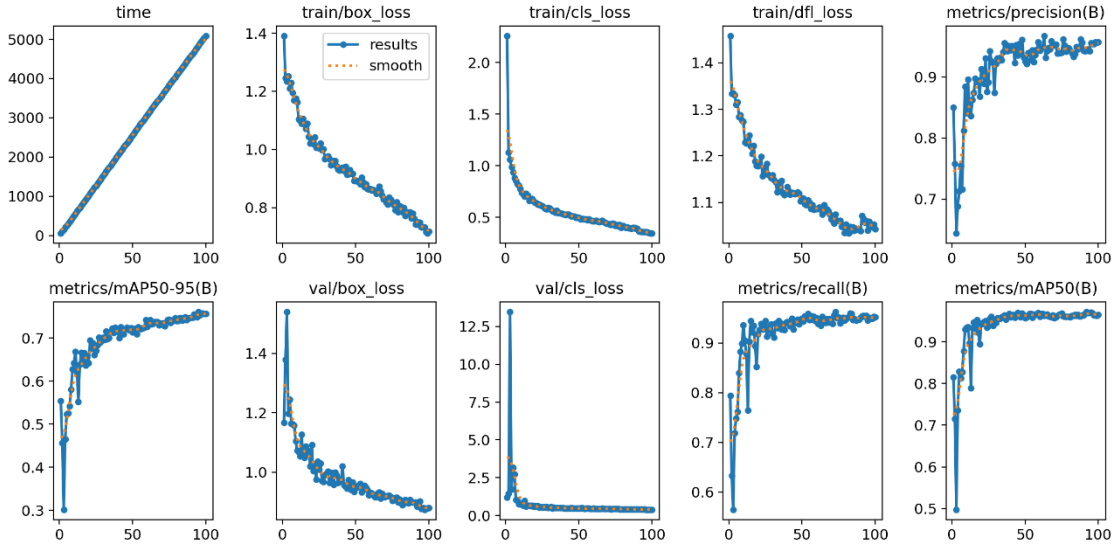


Şekil 4.3. Test kümesindeki video üzerinde güvenli ve güvensiz davranış sınıflarına ait YOLOv4 ile başarılı bir şekilde tanınmış nesne örnekleri



Şekil 4.4. Başlangıçta 105 çerçeveden oluşan bir video parçasının YOLOv4 ağından geçirildikten sonra, içerisinde 3 eylem ve 40 çerçeveden oluşan bir hale gelmesi ve azaltılmış videodaki bazı örnek çerçeveler

YOLO mimarisi ile video çerçevelerinden uygun çerçeve için YOLOv4'ün yanında daha güncel olan YOLOv11'in de nesne tanıma etkinliği de değerlendirilmiştir. Oluşturulan veriseti ve YOLO mimarisin son yayımlanan 11. versiyonu ile yapılan eğitime ilişkin sonuçlar Grafik 4.3.'te gösterilmiştir. Eğitim 100 kez tekrar (epoch) edilmiştir. Eğitim sonucunda doğrulama için ayrılan veriseti üzerinden 8 sınıfa ait performans metrikleri için elde edilen sonuçlar Tablo 4.6'da gösterilmiştir. Burada, tüm sınıflar için YOLOv11 kullanılarak ortalama mAP değerinin %97.1 olarak elde edildiği görülmektedir. YOLOv4 kullanarak aynı veriseti için ortalama mAP değerinin %95.66 (Tablo 4.5) olduğu düşünüldüğünde hem eğitim sonuçlarının hem de performans sonuçlarının her iki YOLO mimarisi için de birbirine yakın olduğu görülmüştür. Dolayısıyla, YOLOv11 kullanmanın deneysel çalışmalarda anlamlı bir değişiklik oluşturmayacağı değerlendirilmiştir.



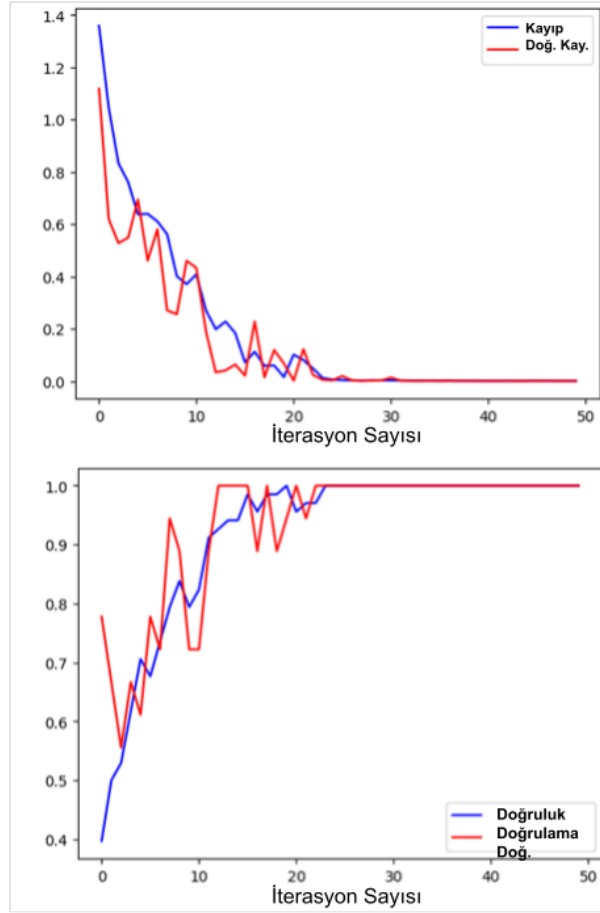
**Grafik 4.3.** Verisetinin YOLOv11 ile eğitilmesi sonucunda kayıp ve diğer metriklerin değişimi

**Tablo 4.6.** Verisetinin YOLOv11 eğitimi sonucunda elde edilen sonuçlar

Sınıflar	Görüntü	Örnekler	Kutu			
			P	R	mAP50(%)	mAP50-90
<b>Tümü (Ortalama)</b>	<b>253</b>	<b>514</b>	<b>0.956</b>	<b>0.946</b>	<b>97.1</b>	<b>0.76</b>
0	110	136	0.959	0.859	96.7	0.681
1	32	33	0.991	0.97	97.1	0.815
2	132	132	0.992	1	0.99	0.805
3	52	52	0.958	0.981	98.2	0.881
4	65	75	0.922	0.867	95.8	0.661
1	18	18	0.935	1	0.992	0.792
6	59	59	0.98	1	0.994	0.79
7	8	9	0.908	0.889	0.91	0.655

#### 4.2.2. YOLOv4+ConvLSTM ile Sınıflandırma

Unsafe-Net mimarisi, YOLOv4 modelinin önceden eğitilmiş ağırlıklarını kullanarak tehlikeli davranışlar içeren videoları işlemektedir. Videoları basitleştirmek için ilgisiz kareler atılmaktadır. Bu, sınıflandırma doğruluğunu ve gerçek zamanlı video görüntülerinden anlam çıkarma hızını artırmaktadır. Güvenli olmayan davranışlar içeren videolarda nesne tanıma için YOLOv4 kullanıldıktan ve gereksiz kısımlar çıkarıldıktan sonra, bu videoların boyutları küçültülmektedir. Daha sonra ConvLSTM kullanılarak sınıf türüne göre sınıflandırılmaktadır. ConvLSTM ile çalışma ortamlarındaki güvensiz davranışların gerçek zamanlı sınıflandırılmasını sağlamak için eğitim seti için 452 video ve test seti için 64 video seçilmiştir. Eğitim ve test kümelerindeki videoların boyutu ve algılama süresi, güvensiz davranış içermeyen kareleri kaldırmak için YOLOv4 kullanılarak azaltılmıştır. ConvLSTM ağının eğitim süresi boyunca Kayıp, Doğrulama Kaybı (val\_loss), Doğruluk ve Doğrulama Doğruluğu (val\_accuracy) değerleri Grafik 4.4.'de gösterildiği gibi iterasyon sayısının bir fonksiyonu olarak değişmiştir. Kayıp değerinin 0'a yakın olduğu ve doğruluk metriğinin 1'e yakın olduğu sonucuna varılabilir. Bu nedenle, ConvLSTM ağının eğitimi başarılı olmuştur.



**Grafik 4.4.** ConvLSTM ağının eğitimi süresinde kayıp, doğrulama kayıp, doğruluk ve doğrulama doğruluk değerlerinin değişimi

Unsafe-Net'in sınıflandırması işyerinde dört farklı güvensiz davranış/eylem sınıfı belirlemiştir: Güvenli Yürüyüş Yolu İhlali, Yetkisiz Müdahale, Açık Pano Kapağı ve Forklift ile Aşırı Yük Taşıma. Bu sınıflara sırasıyla 0, 1, 2 ve 3 sınıf kimliği değerleri atanmıştır. ConvLSTM ağını eğittikten sonra, çalışma Unsafe-Net hibrit mimarisinin performansını değerlendirmiştir. Test setinde Güvenli Geçit İhlali sınıfı için 32 video, Yetkisiz Müdahale sınıfı için 11 video, Açık Pano Kapağı sınıfı için 13 video ve Forklift ile Aşırı Yük Taşıma sınıfı için 8 video yer almıştır. Ayrıca, toplam 32 video YOLOv4 ağından geçmiştir. Tablo 4.7., video test setindeki her bir güvensiz davranış sınıfı için video karesi azaltma olmadan YOLOv4 ve ConvLSTM ağının performansını göstermektedir. Güvenli Yaya Yolu İhlali sınıfında 1 video, Yetkisiz Müdahale sınıfında 2 video, Açık Pano Kapağı sınıfında 3 video ve Forkliftle Aşırı Yük Taşıma sınıfında 1 video vardır ve bunların tümü hatalı olarak sınıflandırılmıştır. Tablo 4.8., video test setindeki her bir güvensiz davranış sınıfı için Unsafe-Net yaklaşımının performansını göstermektedir. Sonuçlar, bu çalışmada önerilen Unsafe-Net hibrit derin

öğrenme mimarisinin Güvenli Yürüyüş Yolu İhlali ve Forkliftle Aşırı Yük Taşıma davranış sınıfları için tüm videoları doğru şekilde sınıflandırdığını göstermektedir. Ancak, Unsafe-Net mimarisi kullanılarak Yetkisiz Müdahale ve Açık Pano Kapağı sınıfları için 1'er video yanlış sınıflandırılmıştır. Veri kümesinin test setindeki 4 sınıfa ait toplam 64 video verisinden 62 video Unsafe-Net mimarisi kullanılarak başarıyla sınıflandırılmıştır.

**Tablo 4.7.** Video test kümesi için YOLOv4 ile video çerçeve azaltma olmadan ConvLSTM ağının her bir güvensiz hareket sınıfı için performansı

		Tahmin			
		Sınıf No	0	1	2
Gerçek	0	31	0	0	1
	1	0	9	2	0
	2	0	3	10	0
	3	1	0	0	7

**Tablo 4.8.** Video test kümesi için bu çalışmada önerilen Unsafe-Net (YOLOv4+ConvLSTM) ağının her bir güvensiz hareket sınıfı için performansı

		Tahmin			
		Sınıf No	0	1	2
Gerçek	0	32	0	0	0
	1	0	10	1	0
	2	0	1	12	0
	3	0	0	0	8

Tablo 4.9., veri kümesinin video test setindeki her bir güvensiz davranış sınıfı için YOLOv4 ve bu çalışmada önerilen Unsafe-Net (YOLOv4+ConvLSTM) hibrit derin öğrenme modeli kullanılarak video karesi azaltma olmadan ConvLSTM ağının doğruluk metrik performansını sunmaktadır. YOLOv4 ve ConvLSTM ağı, video karelerini azaltmadan ortalama

%85.78 Doğruluk performansı elde etmiştir. Güvenli Yürüyüş Yolu İhlali ve Forklift ile Aşırı Yük Taşıma sınıfları %100 doğruluk elde etmiştir. Bununla birlikte, Yetkisiz Müdahale sınıfı %90.91 doğruluk elde ederken, Açık Pano Kapağı güvensiz davranış sınıfı %92.31 doğruluk elde etmiştir. Önerilen Unsafe-Net derin öğrenme mimarisi, tüm güvensiz davranış sınıfları için ortalama %95.81 doğruluk skoru elde etmiştir. Bu da Unsafe-Net ağının çalışma ortamlarında çekilen videolardaki güvensiz davranışları başarılı bir şekilde sınıflandırabildiğini göstermektedir.

**Tablo 4.9.** Video test kümesi için sadece ConvLSTM ağının ve önerilen Unsafe-Net modelinin her bir güvensiz hareket sınıfında doğruluk metriği için performansı

Sınıf No	Güvensiz Davranış	ConvLSTM için Doğruluk (%)	Unsafe-Net (YOLOv4+ConvLSTM) için Doğruluk (%)
0	Güvenli Yürüyüş Yolu İhlali	96.88	100
1	Yetkisiz Müdahale	81.82	90.91
2	Açık Pano Kapağı	76.92	92.31
3	Forklift ile Aşırı Yük Taşıma	87.50	100
Tüm (%)		<b>85.78</b>	<b>95.81</b>

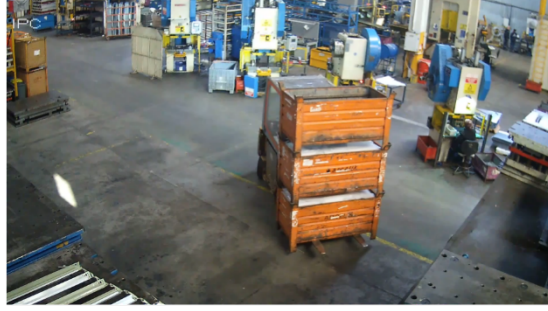
Ayrıca, Şekil 4.5.'te önerilen Unsafe-Net yöntemi için doğru ve yanlış sınıflandırma örnekleri gösterilmektedir, temel gerçek ve tahminler açıkça ayırt edilmektedir. Burada, Güvenli Yürüyüş Yolu İhlali, Yetkisiz Müdahale, Açık Pano Kapağı ve Forklift ile Aşırı Yük Taşıma olmak üzere 4 farklı güvensiz eylemin başarılı bir şekilde sınıflandırıldığını gösterilmiştir. Ancak, Şekil 4.5(e) ve 4.5(f)'de Yetkisiz Müdahale ve Açık Pano Kapağı güvensiz davranışlarının birbirleriyle karıştırıldığı görülmüştür.



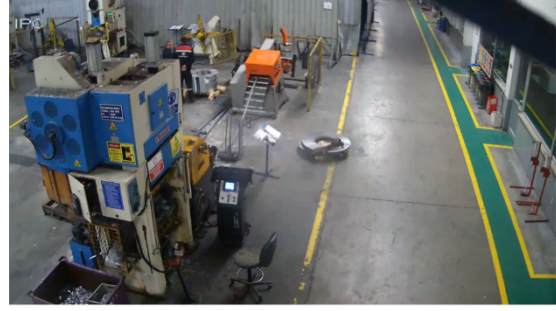
Gerçek Olan=Yaya Yolu İhlali  
Tahmin=Yaya Yolu İhlali ✓✓✓



Gerçek Olan=Yetkisiz Müdahale  
Tahmin=Yetkisiz Müdahale ✓✓✓



Gerçek Olan=Aşırı Yük Taşıma  
Tahmin=Aşırı Yük Taşıma ✓✓✓



Gerçek Olan=Panel Kapağı Açık  
Tahmin=Panel Kapağı Açık ✓✓✓



Gerçek Olan=Yetkisiz Müdahale  
Tahmin=Panel Kapağı Açık XXX

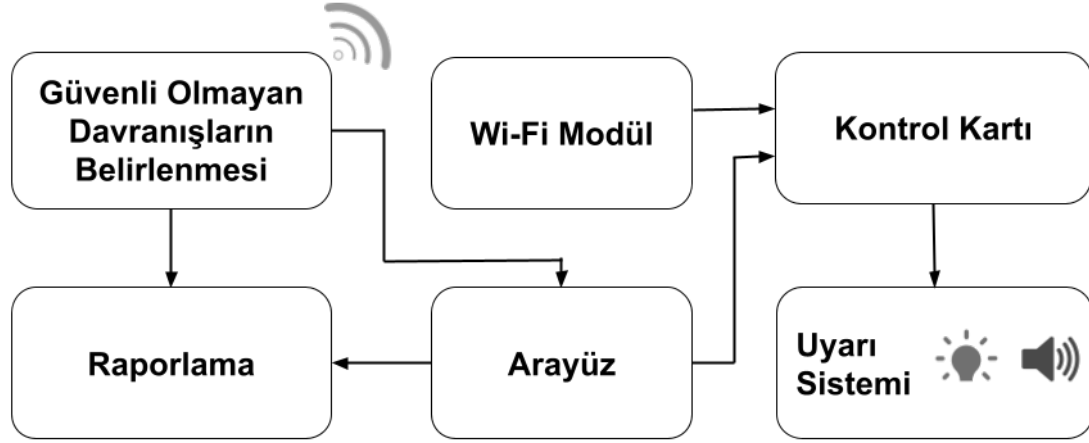


Gerçek Olan=Panel Kapağı Açık  
Tahmin=Yetkisiz Müdahale XXX

Şekil 4.5. Önerilen Unsafe-Net yöntemi için doğru ve yanlış sınıflandırma örnekleri

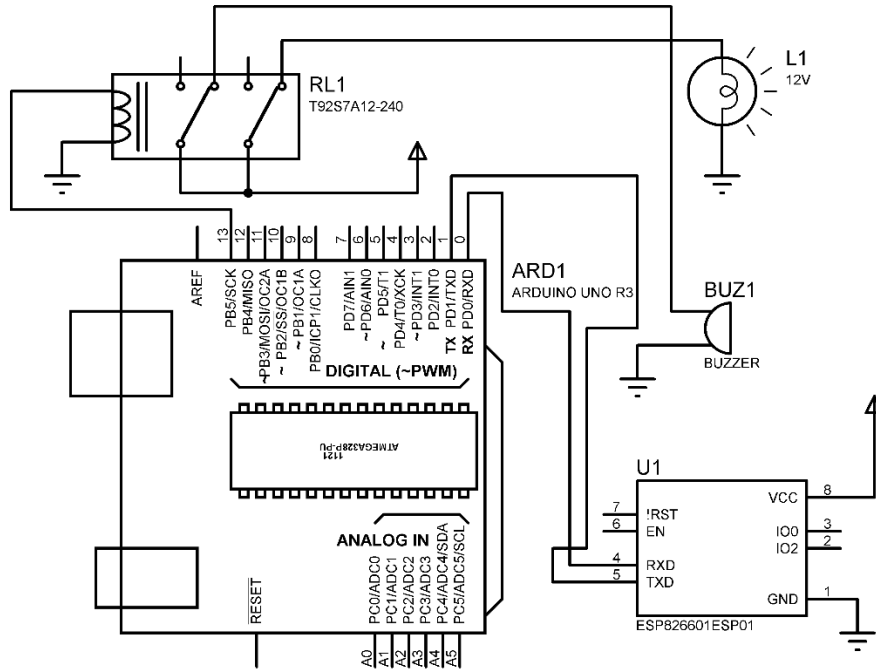
### 4.2.3. Uyarı İzleme Sistemi

Çalışma, fabrikanın çalışma ortamında güvensiz davranışların tespit edilmesi üzerine devreye giren gerçek zamanlı bir uyarı ve izleme sistemi geliştirmiştir. Sistem doğrudan güvensiz davranışın gerçekleştiği alana sesli ve görsel bir uyarı göndermektedir. Şekil 4.6'da geliştirilen uyarı ve izleme sisteminin yapısı gösterilmektedir.



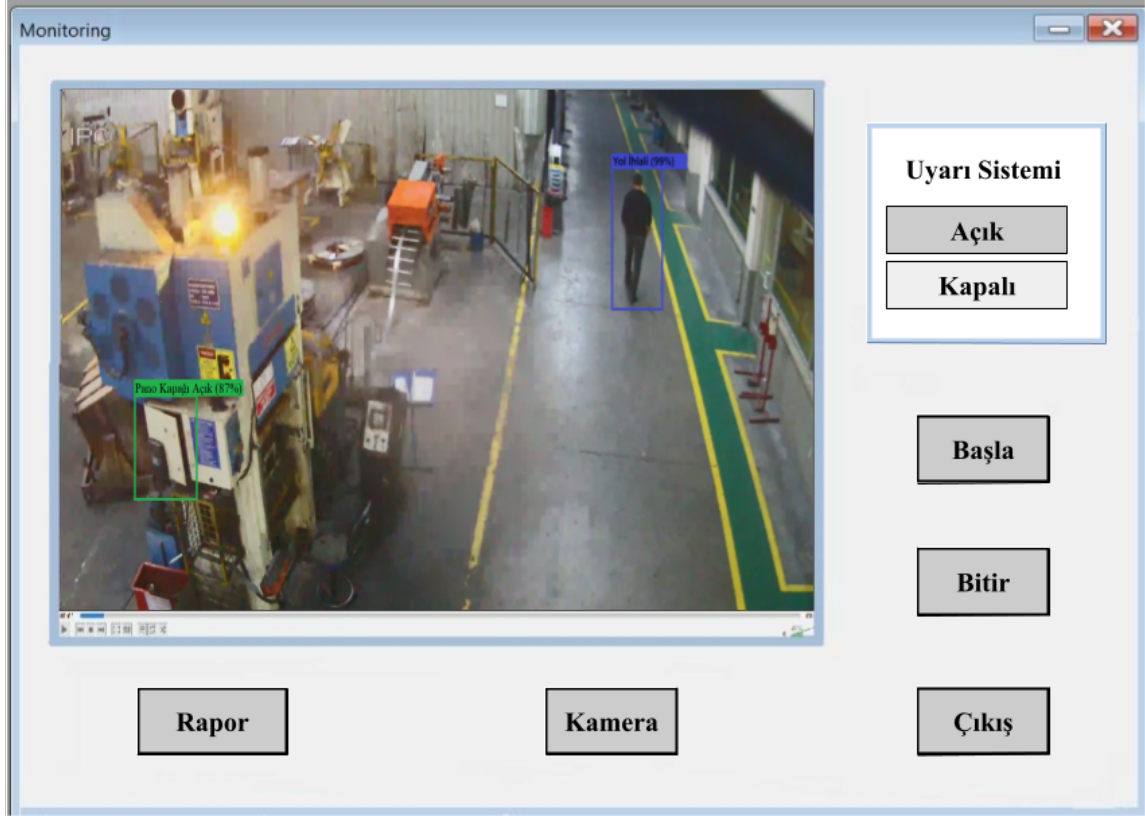
Şekil 4.6. Geliştirilen uyarı ve denetim sisteminin blok diyagramı

Şekil 4.7.'de devre şeması verilmiş olan uyarı ve izleme sistemi, Python bir betiği aracılığıyla güvenli olmayan davranışlar meydana geldiğinde devreye alınmaktadır. Güvensiz davranışla ilgili bilgiler ESP8266 Wi-Fi modülü aracılığıyla kablosuz olarak Arduino UNO tabanlı kontrol kartına iletilmektedir. Kontrol kartı, alınan bilgilere dayanarak sesli ve görsel uyarıları tetiklemektedir. Ek olarak, güvensiz davranış meydana geldiğinde otomatik raporlama ve veri kaydı gerçekleştirilmektedir. Ayrıca, kullanıcı arayüz programı hem raporlama hem de kontrol kartı altyapısına erişim sağlamaktadır.



Şekil 4.7. Uyarı izleme sisteminin kontrol kartı devre şeması

Gerçek zamanlı sistemi izlemek ve kontrol etmek için Python programlama dili ve PyQt Designer kullanılarak bir program geliştirilmiştir. Program, gerçek zamanlı sistemin çalışmasını başlatmaya ve durdurmaya, çalışma ortamındaki güvensiz faaliyetleri görüntülemeye, sistem tarafından oluşturulan metin belgesi türündeki raporlara erişmeye ve uyarı sistemini çalıştırmaya izin vermektedir. Şekil 4.8., programın grafik arayüzünü ve Şekil 4.9. ise rapor içeriklerinden bazılarını göstermektedir.



**Şekil 4.8.** Unsafe-Net mimarisinde sistemin izlenmesi ve kontrolü için geliştirilen programın arayüzü

21-Güvenli Yaya Yolu İhlali	Alan-1	09:35	25/11/2022.....
22-Yetkisiz Müdahale	Alan-1	10:45	25/11/2022.....
23-Açık Panel Kapağı	Alan-1	10:54	25/11/2022.....
24-Forklift ile Aşırı Yük Taşıma	Alan-2	12:09	25/11/2022.....
25-Yetkisiz Müdahale	Alan-1	13:05	25/11/2022.....
26-Açık Panel Kapağı	Alan-1	13:19	25/11/2022.....
27-Açık Panel Kapağı	Alan-1	13:22	25/11/2022.....
28-Forklift ile Aşırı Yük Taşıma	Alan-2	15:09	25/11/2022.....
29-Güvenli Yaya Yolu İhlali	Alan-1	15:35	25/11/2022.....
30-Yetkisiz Müdahale	Alan-1	16:45	25/11/2022.....
31-Güvenli Yaya Yolu İhlali	Alan-1	16:54	25/11/2022.....
32-Güvenli Yaya Yolu İhlali	Alan-1	16:55	25/11/2022.....
33-Güvenli Yaya Yolu İhlali	Alan-1	17:19	25/11/2022.....
34-Açık Panel Kapağı	Alan-1	10:22	26/11/2022.....
35-Yetkisiz Müdahale	Alan-1	10:45	26/11/2022.....
36-Güvenli Yaya Yolu İhlali	Alan-1	10:54	26/11/2022.....
37-Forklift ile Aşırı Yük Taşıma	Alan-2	11:09	26/11/2022.....
38-Yetkisiz Müdahale	Alan-1	13:05	26/11/2022.....
39-Güvenli Yaya Yolu İhlali	Alan-1	13:19	26/11/2022.....
40-Açık Panel Kapağı	Alan-1	13:22	26/11/2022.....
41-Forklift ile Aşırı Yük Taşıma	Alan-2	14:09	26/11/2022.....
42-Güvenli Yaya Yolu İhlali	Alan-1	14:35	26/11/2022.....
43-Güvenli Yaya Yolu İhlali	Alan-1	14:45	26/11/2022.....
44-Yetkisiz Müdahale	Alan-1	15:54	26/11/2022.....
45-Güvenli Yaya Yolu İhlali	Alan-1	16:09	26/11/2022.....
46-Yetkisiz Müdahale	Alan-1	17:05	26/11/2022.....
47-Açık Panel Kapağı	Alan-1	10:19	27/11/2022.....
48-Açık Panel Kapağı	Alan-1	10:22	27/11/2022.....
49-Forklift ile Aşırı Yük Taşıma	Alan-2	11:09	27/11/2022.....
50-Güvenli Yaya Yolu İhlali	Alan-1	14:35	27/11/2022.....
51-Yetkisiz Müdahale	Alan-1	16:08	27/11/2022.....
52-Forklift ile Aşırı Yük Taşıma	Alan-2	16:09	27/11/2022.....
53-Güvenli Yaya Yolu İhlali	Alan-1	16:35	27/11/2022.....

**Şekil 4.9.** Güvensiz hareket meydana geldiğinde sistem tarafından oluşturulan metin belgesi tipinde raporlama içeriği

Unsafe-Net sistemde, çalışma ortamlarında işçiler tarafından güvensiz davranışlardan birisi yapıldığında, sistem ile elektronik kontrol kartı Wi-Fi ağı üzerinden haberleştirilerek, çalışanların bulunduğu ortama sesli ve görsel uyarı ikazlarının gönderilmesi sağlanmaktadır. Şekil 4.10.'da görüldüğü gibi, çalışma ortamlarında Güvenli Yürüyüş Yolu İhlali sınıfına ait güvensiz davranış geliştirilen sistem tarafından başarılı bir şekilde tespit edilerek ortama sesli ve görsel uyarı verilmiştir.



Şekil 4.10. Güvenli olmayan bir davranış meydana geldiğinde sesli ve görsel uyarı verilmesi

## 5. BULGULAR VE TARTIŞMA

Bu tez çalışmasında, çalışma ortamlarındaki güvenli olmayan eylemlerin gerçek zamanlı tespiti için Unsafe-Net mimarisini önermiştir. Ayrıca, YOLO mimarisi kullanılarak durağan görüntülerde ve videolarda KKD tespiti de gerçekleştirilmiştir.

Unsafe-Net mimarisinin iki aşamalı bir uygulaması vardır. İlk aşamada videolarda güvenli olmayan eylemlerin yer aldığı karelerin seçilmesi için YOLO v4 kullanılmaktadır, diğerleri ise hariç tutulmaktadır. Önerilen sistem gerçek zamanlıdır ve eylem algılama süresini çok önemli bir faktör haline getirmektedir. Tablo 5.1.'de veri setindeki güvenli olmayan eylemlerin yer aldığı 3 ile 18 saniye arasında değişen 64 videonun ortalama video süresinin 7.72 saniye olduğu görülmektedir. YOLOv4 kullanılarak test setindeki videolar ConvLSTM ile herhangi bir azaltma yapılmadan sınıflandırıldığında dört farklı sınıf için ortalama eylem tanıma süresi 0.58 saniyedir. Çalışma, güvenli olmayan eylemlerin olduğu kareleri seçip diğerlerini hariç tutarak ortalama video süresini 1.87 saniyeye düşüren Unsafe-Net mimarisini önermektedir. Ek olarak, test setindeki 64 video için eylem algılama süresi 0.14 saniyedir ve bu da sistemin eylem tanıma için gerçek zamanlı yürütülmesini önemli ölçüde hızlandırır.

**Tablo 5.1.** ConvLSTM ve önerilen Unsafe-Net için ortalama video süresi ve ortalama eylem tanımlama süresinin karşılaştırılması

Metot	Ortalama video süresi (s)	Ortalama eylem tanımlama süresi (s)
ConvLSTM	7.72	0.58
Unsafe-Net	1.87	0.14

Çalışmada, fabrika ortamında meydana gelen güvensiz hareketlerin, bu çalışmada önerilen Unsafe-Net mimarisini içeren sistem kurulmadan önce ve sonrası değişimi de takip edilmiştir. Tablo 5.2'de görüldüğü gibi 5-23 Kasım 2022 tarihleri arasında sistem kurulmadan önce Güvenli Yürüyüş Yolu İhlali, Yetkisiz Müdahale, Açık Pano Kapağı ve Forklift ile Aşırı Yük Taşıma güvensiz sınıflara ait fabrika ortamında günlük iş güvenliği uzmanı tarafından tespit edilen davranışların sayısı verilmiştir. Burada, verilerin toplandığı ilk periyotta 4 sınıftan toplam 518 güvensiz hareketin meydana geldiği görülmektedir.

**Tablo 5.2.** Fabrikaya gerçek zamanlı Unsafe-Net mimarisini içeren sistemin kurulmadan öncesi için belirli bir tarih arasında meydana gelen güvensiz eylemlerin sayısı

Tarih (D/M/Y)	Yetkisiz Müdahale	Açık Pano Kapağı	Forklift ile Aşırı Yük Taşıma	Güvenli Yürüyüş Yolu İhlali
05/11/2022	6	7	3	10
06/11/2022	5	5	2	9
07/11/2022	7	9	2	9
08/11/2022	9	4	4	12
09/11/2022	4	5	5	9
10/11/2022	11	12	0	9
11/11/2022	3	6	1	14
12/11/2022	4	8	6	17
13/11/2022	0	11	3	8
14/11/2022	5	12	2	9
15/11/2022	6	6	2	12
16/11/2022	5	5	2	9
17/11/2022	8	3	3	11
18/11/2022	6	9	4	9
19/11/2022	7	8	1	12
20/11/2022	5	2	2	10
21/11/2022	4	11	4	9
22/11/2022	4	9	7	15
23/11/2022	9	10	5	17
<b>Total</b>	<b>108</b>	<b>142</b>	<b>58</b>	<b>210</b>

**Tablo 5.3.** Fabrikaya gerçek zamanlı Unsafe-Net mimarisini içeren sistemin kurulmasından sonra meydana gelen güvensiz eylemlerin sayısı

Tarih (G/A/Y)	Yetkisiz Müdahale	Açık Pano Kapağı	Forklift ile Aşırı Yük Taşıma	Güvenli Yürüyüş Yolu İhlali
24/11/2022	5	4	3	6
25/11/2022	4	3	2	5
26/11/2022	4	2	2	5
27/11/2022	3	2	2	5
28/11/2022	3	2	2	4
29/11/2022	2	2	0	3
30/11/2022	3	2	1	4
01/12/2022	1	1	2	3
02/12/2022	0	1	3	4
03/12/2022	1	0	2	3
04/12/2022	0	1	2	3
05/12/2022	0	1	3	1
06/12/2022	0	1	1	2
07/12/2022	1	2	1	1
08/12/2022	0	0	1	2
09/12/2022	0	0	0	2
10/12/2022	1	0	0	1
11/12/2022	0	0	0	1
12/12/2022	0	0	0	0
13/12/2022	0	0	0	0
<b>Toplam</b>	<b>28</b>	<b>24</b>	<b>27</b>	<b>55</b>

Önerilen sistem, 5-23 Kasım 2022 tarihleri arasında sonuçlar elde edildikten sonra çalışanlara tanıtılmıştır. Takip eden günlerde, güvensiz davranışlarla ilgili güvensiz eylemlerin sayısını belirlemek için video görüntüleri alınmıştır. Tablo 5.2., Unsafe-Net mimarisine sahip gerçek zamanlı sistemin fabrikaya kurulmasından sonra 24 Kasım 2022 ile 13 Aralık 2022 tarihleri arasında meydana gelen güvensiz davranışların sayısını göstermektedir. Sistemin performansı, güvensiz davranışlara ilişkin 20 günlük rapor incelenerek değerlendirilmiştir. Çalışma ayrıca sistemin iş güvenliği üzerindeki etkisini de analiz etmiştir. Tablo 5.3., önerilen sistemin uygulanmasından sonra Güvenli Yürüyüş Yolu İhlali, Yetkisiz Müdahale, Açık Pano Kapağı ve Aşırı Yük Taşıma gibi güvensiz davranışlarda önemli bir azalma olduğunu göstermektedir. Sistemin kurulmasından önceki ve sonraki dönemde güvensiz eylemlerin sayısı 518'den 134'e düşerek yaklaşık %75 oranında azalmıştır. Unsafe-Net mimarisinin dört güvensiz davranış sınıfı için verdiği kararın doğruluğu doğrulanmış ve fabrikanın iş güvenliği uzmanı sınıflandırma sonuçlarını onaylamıştır. Önerilen sistemin sınıflandırma sonuçlarını aynı dönemdeki güvenlik kamerası izleme sonuçlarıyla karşılaştıran iş güvenliği uzmanı, Açık Pano Kapağı sınıfında sadece bir güvensiz davranış örneğinin yanlışlıkla Yetkisiz Müdahale olarak sınıflandırıldığını bildirmiştir.

## 6. SONUÇ

Bu tez çalışmasında, sanayide üretim yapan tesislerde gerçek zamanlı olarak gözetim videoları vasıtasıyla çalışanların güvensiz hareketlerinin tanınması ve anında sistem tarafından uyarılarak belirtilen durumların önlenmesi başarılmıştır. Bu kapsamda, YOLOv4 derin öğrenme modeli kullanılarak üretim sanayisinde çalışanların KKD'lerinin kullanımlarının denetlenmesi ve güvenli bir çalışma ortamının tesis edilmesi amacı ile nesne tespiti gerçekleştirilmiştir. Bunun için, çalışma ortamlarında baret, yelek, maske, eldiven, gözlük gibi KKD'lerin tanınması ve çalışanların uygun donanımları doğru kullanıp kullanmadıkları YOLOv4 altyapısı kullanılarak tespit edilmiştir. Çalışma kapsamında, Kafaoğlu Metal Plastik Makine San. ve Tic. A.Ş., Tek Metal ve Plastik Endüstriyel Mamulleri San. Tic. Ltd. Şti. ve Bilecik Şeyh Edebali Üniversitesi Meslek Yüksekokulu Makine Atölyesinden görüntüler/videolar kullanılarak veri seti oluşturulmuştur. Derin öğrenme algoritmalarında daha yüksek doğruluk başarımı elde etmek için çok fazla görüntü/video ile mimarinin eğitimi yapmak gerekmektedir. Bu nedenle, çalışma ortamlarından, çeşitli zamanlarda birkaç farklı kamera kullanılarak farklı çözünürlüklerde video ve görüntüler toplanmıştır. Çalışma kapsamında yürütülen deneysel çalışmalarda, farklı ağ parametreleri ve farklı sayıda GPU kullanılarak YOLOv4 ağı eğitilmiş, eğitimler sonucunda elde edilen ağırlıklar kullanılarak, aynı şekilde benzer ortamlardan alınmış video görüntüleri üzerinde önerilen yöntemin başarımı test edilmiştir. IoU eşik değerinin 0.5 olarak belirlenmesi sonucunda, önerilen YOLOv4 ağı üzerinde yapılan eğitim sonucunda mAP değeri %91.18 olarak en yüksek değerine ulaşmıştır. Ayrıca, kesinlik, duyarlılık, F1-skoru, IoU ve ortalama kayıp için sırasıyla 0.89, 0.91, 0.90, 70.35 ve 1.1147 sonuçları başarılmıştır.

Üretim ortamlarında KKD'lerin kullanımına yönelik olarak yapılan nesne tespitinde yüksek doğrulukta ve hızda sonuçlar aldıktan sonra iş birliği yapılan firma yetkilileri ve iş güvenliği uzmanları ile görüşülerek güvensiz davranışlar belirlenmiş ve yaklaşık 40 gün boyunca güvenlik kameralarından veri alınmıştır. Toplanan bu veriler nesne tespiti ve video anlamlandırma çalışmaları doğrultusunda hazırlanıp etiketlenmiş ve düzenlenerek çalışma kapsamında orijinal veri setleri oluşturulmuştur. Güvensiz davranışlar ve hareketlerle ilişkili nesnelere içermeyen çerçeveler YOLOv4 derin öğrenme modeli ile tespit edilerek ConvLSTM modelinde işlenmeden önce videolardan ayrılmıştır. Önerilen Unsafe-Net hibrit derin öğrenme mimarisi kullanılarak, çalışma ortamlarında Güvenli Yürüyüş Yolu İhlali, Yetkisiz Müdahale, Açık Pano Kapağı ve Forklift ile Aşırı Yük Taşıma sınıflarına ait güvensiz eylemler %95.81 doğruluk oranı ile tespit edilmiştir. Bunun yanında, Unsafe-Net mimarisi ile test kümesindeki

videolar için ortalama video süresi 1.87 saniyeye düşürülerek, eylem tanıma süresi ise 0.14 saniyeye kazar azaltılmıştır ve önemli başarımlar elde edilmiştir.

Çalışma ortamlarında işçiler tarafından güvensiz davranışlardan birisi yapıldığında, Unsafe-Net tabanlı gerçek zamanlı sistem ile uyarı ve izleme sistemi tarafından çalışanların bulunduğu ortama sesli ve görsel uyarı ikazlarının gönderilmesi sağlanmıştır. Böylece, çalışanlar güvensiz davranışta bulunmamaları konusunda sistem tarafından yönlendirilmiştir. Bu amaçla, önerilen Unsafe-Net mimarisine dayalı gerçek zamanlı bir sistem pilot uygulama için bir fabrikaya kurularak, çalışanların güvensiz davranışlardaki eğilimleri bir süre izlenmiştir. Deneysel bulgular, sistem kurulmadan önce 4 farklı güvensiz davranış sınıfında 518 olan güvensiz eylem sayısı, sistem kurulduktan sonra 134'e düşerek, yani fabrikada meydana gelen güvensiz davranış sayısı yaklaşık %75 oranında azalarak önemli bir iyileşme kaydedildiğini göstermektedir. Sonuç olarak, YOLOv4 derin öğrenme modeli ile CNN ve LSTM algoritmaları birleştirilerek güvensiz davranışların tespiti için Unsafe-Net yaklaşımı getirilmiş ve güvensiz davranış videoların anlamlandırmasında doğruluk artırılmış ve eylem tanıma süresi azaltılmıştır.

Önerilen Unsafe-Net mimarisinin YOLOv4 bileşeninde 3 saniyelik bir videonun süresinin ortalama 1 saniyeye düşürülmesi ve ConvLSTM için hazır hale gelebilmesi için yaklaşık 1.2 saniye süre gerekmektedir. Unsafe-Net mimarisinde eylem tanıma süresinin ortalama 0.14 saniye olduğu hesaba katılırsa, gerçek zamanlı olarak videolardan güvensiz eylemlerin tespiti için yaklaşık 1.34 saniye gibi bir süre gerekmektedir. Bu sürenin de gerçek zamanlı bir tespit ve tanıma sistemi için yeterli olacağı değerlendirilmektedir.

Tez çalışması kapsamında çalışma ortamlarında güvensiz davranışların tespiti için gerçek zamanlı ve saha uygulaması olan bir bilgisayarla görme sistemi önerilmiş olsa da birtakım sınırlılıklar da bulunmaktadır. Yapılan çalışma belirli sayıda sınıfı içermekte ve bu alanda karşılaşılabilecek sınırsız sayıda sınıf bulunmaktadır. Bu tez çalışması derin öğrenme ve bilgisayarla görme alanları dışında iş güvenliği ve sağlığı alanına yönelik de bir çalışmadır. Dolayısıyla iş kazalarının önlenmesine yönelik yapılacak hemen hemen tüm çalışmalar o alana yönelik olarak özel yapılmalıdır. Geliştirilen yöntem ile anlık güvenli güvensiz hareketlerin tespit edilerek videoların sınıflandırılması her ne kadar yapılmış olsa da yetersiz aydınlatma, oklüzyon, perspektif gibi bilgisayarla görme sorunları vardır ve bu alanda çalışmalar hala devam etmektedir.

## KAYNAKÇA

- Abbasi, M., Gholamnia, R., Alizadeh, S. S., & Rasoulzadeh, Y.** (2015). Evaluation of workers unsafe behaviors using safety sampling method in an industrial company. *Indian Journal of Science and Technology*, 8(28), 1-6.
- Adhikari, P.** (2015). ERRORS AND ACCIDENTS IN THE WORKPLACES. *Sigurnost*, 57(2).
- Akbari, H., Kashani, M. M., Asadi, Z., Kaveh, M., & Saberi, H.** (2019). The relationship between job satisfaction and the incidence of unsafe acts in metal smelting industry workers in 2017. *International Archives of Health Sciences*, 6(3), 127-131.
- Alkhaldi, M., Pathirage, C., & Kulatunga, U.** (2017). The role of human error in accidents within oil and gas industry in Bahrain.
- Altunkaynak, B.** (2018). A statistical study of occupational accidents in the manufacturing industry in Turkey. *International journal of industrial ergonomics*, 66, 101-109.
- Alwasel, A., Sabet, A., Nahangi, M., Haas, C. T., & Abdel-Rahman, E.** (2017). Identifying poses of safe and productive masons using machine learning. *Automation in Construction*, 84, 345-355.
- Ann, O. C., & Theng, L. B.** (2014). Human activity recognition: A review. *2014 IEEE international conference on control system, computing and engineering (ICCSCCE 2014)*.
- Barlybayev, A., Amangeldy, N., Kurmetbek, B., Krak, I., Razakhova, B., Tursynova, N., & Turebayeva, R.** (2024). Personal protective equipment detection using YOLOv8 architecture on object detection benchmark datasets: a comparative study. *Cogent Engineering*, 11(1), 2333209.
- Barro-Torres, S., Fernández-Caramés, T. M., Pérez-Iglesias, H. J., & Escudero, C. J.** (2012). Real-time personal protective equipment monitoring system. *Computer Communications*, 36(1), 42-50.
- Beddiar, D. R., Nini, B., Sabokrou, M., & Hadid, A.** (2020). Vision-based human activity recognition: a survey. *Multimedia Tools and Applications*, 79(41), 30509-30555.
- BLS, U. S. B. o. L. S.** (2023). Employer-Reported Workplace Injuries and Illnesses 2021-2022. <https://www.bls.gov/news.release/pdf/osh.pdf>
- Bluff, E.** (2003). *Systematic management of occupational health and safety* (Vol. 20): Citeseer.

- Bochkovskiy, A., Wang, C.-Y., & Liao, H.-Y. M.** (2020). Yolov4: Optimal speed and accuracy of object detection. *arXiv preprint arXiv:2004.10934*.
- Carreira, J., & Zisserman, A.** (2017). Quo vadis, action recognition? a new model and the kinetics dataset. *proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*.
- Cavlak, N., Turkoglu, A., Kiliccioglu, D. B., & Tokdemir, M.** (2022). Fatal occupational injuries in eastern Turkey between 2000 and 2016. *Medicine Science, 11*(2), 766-769.
- Ceylan, H., & Ceylan, H.** (2012). Analysis of occupational accidents according to the sectors in Turkey. *Gazi university journal of science, 25*(4), 909-918.
- Chen, H., Luo, X., Zheng, Z., & Ke, J.** (2019). A proactive workers' safety risk evaluation framework based on position and posture data fusion. *Automation in Construction, 98*, 275-288.
- Chen, S., & Demachi, K.** (2021). Towards on-site hazards identification of improper use of personal protective equipment using deep learning-based geometric relationships and hierarchical scene graph. *Automation in Construction, 125*, 103619.
- Demirbilek, S., & Pazarlıoğlu, M.** (2007). Türkiye’de iş kazalarının oluşumunda etkili olan faktörler: ampirik bir uygulama. *Finans Politik & Ekonomik Yorumlar, 44*(509), 81-91.
- Diba, A., Fayyaz, M., Sharma, V., Karami, A. H., Arzani, M. M., Yousefzadeh, R., & Van Gool, L.** (2017). Temporal 3d convnets: New architecture and transfer learning for video classification. *arXiv preprint arXiv:1711.08200*.
- Ding, L., Fang, W., Luo, H., Love, P. E., Zhong, B., & Ouyang, X.** (2018). A deep hybrid learning model to detect unsafe behavior: Integrating convolution neural networks and long short-term memory. *Automation in Construction, 86*, 118-124.
- Dodoo, J. E., & Al-Samarraie, H.** (2019). Factors leading to unsafe behavior in the twenty first century workplace: a review. *Management Review Quarterly, 69*(4), 391-414.
- Donahue, J., Anne Hendricks, L., Guadarrama, S., Rohrbach, M., Venugopalan, S., Saenko, K., & Darrell, T.** (2015). Long-term recurrent convolutional networks for visual recognition and description. *Proceedings of the IEEE conference on computer vision and pattern recognition*.

- Fang, Q., Li, H., Luo, X., Ding, L., Luo, H., & Li, C.** (2018). Computer vision aided inspection on falling prevention measures for steeplejacks in an aerial environment. *Automation in Construction*, *93*, 148-164.
- Fang, Q., Li, H., Luo, X., Ding, L., Rose, T. M., An, W., & Yu, Y.** (2018). A deep learning-based method for detecting non-certified work on construction sites. *Advanced Engineering Informatics*, *35*, 56-68.
- Fang, W., Ding, L., Zhong, B., Love, P. E., & Luo, H.** (2018). Automated detection of workers and heavy equipment on construction sites: A convolutional neural network approach. *Advanced Engineering Informatics*, *37*, 139-149.
- Fang, W., Love, P. E., Luo, H., & Xu, S.** (2022). A deep learning fusion approach to retrieve images of People's unsafe behavior from construction sites. *Developments in the Built Environment*, *12*, 100085.
- Feichtenhofer, C., Pinz, A., & Zisserman, A.** (2016). Convolutional two-stream network fusion for video action recognition. *Proceedings of the IEEE conference on computer vision and pattern recognition*.
- Girdhar, R., Ramanan, D., Gupta, A., Sivic, J., & Russell, B.** (2017). Actionvlad: Learning spatio-temporal aggregation for action classification. *Proceedings of the IEEE conference on computer vision and pattern recognition*.
- Girshick, R.** (2015). Fast r-cnn. *arXiv preprint arXiv:1504.08083*.
- Girshick, R., Donahue, J., Darrell, T., & Malik, J.** (2014). Rich feature hierarchies for accurate object detection and semantic segmentation. *Proceedings of the IEEE conference on computer vision and pattern recognition*.
- Gulhan, B., Ilhan, M., & Civil, E.** (2012). Occupational accidents and affecting factors of metal industry in a factory in Ankara. *Turkish Journal of Public Health*, *10*(2), 76-85.
- Guo, F., Qian, Y., & Shi, Y.** (2021). Real-time railroad track components inspection based on the improved YOLOv4 framework. *Automation in Construction*, *125*, 103596.
- Guo, S., Luo, H., & Yong, L.** (2015). A big data-based workers behavior observation in China metro construction. *Procedia Engineering*, *123*, 190-197.
- Guo, Y., Liu, Y., Oerlemans, A., Lao, S., Wu, S., & Lew, M. S.** (2016). Deep learning for visual understanding: A review. *Neurocomputing*, *187*, 27-48.

- He, K., Zhang, X., Ren, S., & Sun, J.** (2015). Spatial pyramid pooling in deep convolutional networks for visual recognition. *IEEE transactions on pattern analysis and machine intelligence*, 37(9), 1904-1916.
- Heinrich, H. W.** (1941). Industrial Accident Prevention. A Scientific Approach. *Industrial Accident Prevention. A Scientific Approach.*(Second Edition).
- Hu, J., Gao, X., Wu, H., & Gao, S.** (2019). Detection of workers without the helmets in videos based on YOLO V3. *2019 12th International congress on image and signal processing, biomedical engineering and informatics (CISP-BMEI)*.
- ILO.** (2018). Improving the Safety and Health of Young Workers. <https://www.ilo.org/resource/news/28-april-2018-improving-safety-and-health-young-workers>
- Jain, A., Hassard, J., Leka, S., Di Tecco, C., & Iavicoli, S.** (2021). The role of occupational health services in psychosocial risk management and the promotion of mental health and well-being at work. *International journal of environmental research and public health*, 18(7), 3632.
- Jilcha, K., & Kitaw, D.** (2016). A literature review on global occupational safety and health practice & accidents severity. *International Journal for Quality Research*, 10(2), 279.
- Jocher, G., Qiu, J., & Chaurasia, A.** (2024). Ultralytics YOLO11. <https://github.com/ultralytics/ultralytics>
- Khan, I. U., Afzal, S., & Lee, J. W.** (2022). Human activity recognition via hybrid deep learning based model. *Sensors*, 22(1), 323.
- Khan, N., Saleem, M. R., Lee, D., Park, M.-W., & Park, C.** (2021). Utilizing safety rule correlation for mobile scaffolds monitoring leveraging deep convolution neural networks. *Computers in Industry*, 129, 103448.
- Kim, J.-a., Sung, J.-Y., & Park, S.-h.** (2020). Comparison of Faster-RCNN, YOLO, and SSD for real-time vehicle type recognition. *2020 IEEE international conference on consumer electronics-Asia (ICCE-Asia)*.
- Kong, T., Fang, W., Love, P. E., Luo, H., Xu, S., & Li, H.** (2021). Computer vision and long short-term memory: Learning to predict unsafe behaviour in construction. *Advanced Engineering Informatics*, 50, 101400.
- Krizhevsky, A., Sutskever, I., & Hinton, G. E.** (2017). ImageNet classification with deep convolutional neural networks. *Communications of the ACM*, 60(6), 84-90.

- LaDou, J., London, L., & Watterson, A.** (2018). Occupational health: a world of false promises. *Environmental Health*, *17*, 1-8.
- Liao, P.-C., Shi, H., Su, Y., & Luo, X.** (2018). Development of data-driven influence model to relate the workplace environment to human error. *Journal of construction engineering and management*, *144*(3), 04018003.
- Liu, J., Fang, W., Love, P. E., Hartmann, T., Luo, H., & Wang, L.** (2022). Detection and location of unsafe behaviour in digital images: A visual grounding approach. *Advanced Engineering Informatics*, *53*, 101688.
- Liu, R., Liu, H.-C., Shi, H., & Gu, X.** (2023). Occupational health and safety risk assessment: A systematic literature review of models, methods, and applications. *Safety science*, *160*, 106050.
- Liu, W., Anguelov, D., Erhan, D., Szegedy, C., Reed, S., Fu, C.-Y., & Berg, A. C.** (2016). Ssd: Single shot multibox detector. *Computer Vision—ECCV 2016: 14th European Conference, Amsterdam, The Netherlands, October 11–14, 2016, Proceedings, Part I 14*.
- Lo, J., Chan, L., & Flynn, S.** (2021). A systematic review of the incidence, prevalence, costs, and activity and work limitations of amputation, osteoarthritis, rheumatoid arthritis, back pain, multiple sclerosis, spinal cord injury, stroke, and traumatic brain injury in the United States: a 2019 update. *Archives of physical medicine and rehabilitation*, *102*(1), 115-131.
- Long, X., Deng, K., Wang, G., Zhang, Y., Dang, Q., Gao, Y., Shen, H., Ren, J., Han, S., & Ding, E.** (2020). PP-YOLO: An effective and efficient implementation of object detector. *arXiv preprint arXiv:2007.12099*.
- Luo, X., O'Brien, W. J., Leite, F., & Goulet, J. A.** (2014). Exploring approaches to improve the performance of autonomous monitoring with imperfect data in location-aware wireless sensor networks. *Advanced Engineering Informatics*, *28*(4), 287-296.
- Matías, J. M., Rivas, T., Martín, J., & Taboada, J.** (2008). A machine learning methodology for the analysis of workplace accidents. *International Journal of Computer Mathematics*, *85*(3-4), 559-578.
- Mneymneh, B. E., Abbas, M., & Khoury, H.** (2017). Automated hardhat detection for construction safety applications. *Procedia Engineering*, *196*, 895-902.
- Muthukumar, K., Janardhan, S., & Rajiev, R.** (2019). A study on work place health and safety aspects in manufacturing industry. *AIP Conference Proceedings*.

- Omidvari, M., & Gharmaroudi, M.** (2015). Analysis of human error in occupational accidents in the power plant industries using combining innovative FTA and meta-heuristic algorithms. *Journal of Health & Safety at Work*, 5(3).
- Onofri, L., Soda, P., Pechenizkiy, M., & Iannello, G.** (2016). A survey on using domain and contextual knowledge for human activity recognition in video streams. *Expert Systems with Applications*, 63, 97-111.
- Önal, O., & Dandıl, E.** (2021). Object Detection for Safe Working Environments using YOLOv4 Deep Learning Model. *Avrupa Bilim ve Teknoloji Dergisi*(26), 343-351.
- Önal, O., & Dandıl, E.** (2024a). Unsafe-Net: YOLO v4 and ConvLSTM based computer vision system for real-time detection of unsafe behaviours in workplace. *Multimedia Tools and Applications*, 1-27.
- Önal, O., & Dandıl, E.** (2024b). Video dataset for the detection of safe and unsafe behaviours in workplaces. *Data in Brief*, 110791.
- Park, M.-W., & Brilakis, I.** (2012). Enhancement of construction equipment detection in video frames by combining with tracking *Computing in Civil Engineering (2012)* (pp. 421-428).
- Poppe, R.** (2010). A survey on vision-based human action recognition. *Image and vision computing*, 28(6), 976-990.
- Prasad, R. R.** (2017). Prevention and health promotion as a management task Implementation of Marine accident analysis using PRA concepts. *Research Journal of Pharmacy and Technology*, 10(12), 4356-4358.
- Reason, J., Manstead, A., Stradling, S., Baxter, J., & Campbell, K.** (1990). Errors and violations on the roads: a real distinction? *Ergonomics*, 33(10-11), 1315-1332.
- Redmon, J., Divvala, S., Girshick, R., & Farhadi, A.** (2016). You only look once: Unified, real-time object detection. *Proceedings of the IEEE conference on computer vision and pattern recognition*.
- Ren, S., He, K., Girshick, R., & Sun, J.** (2015). Faster r-cnn: Towards real-time object detection with region proposal networks. *Advances in neural information processing systems*, 28.

- Reyes, R. M., de la Riva, J., Maldonado, A., & Woocay, A.** (2015). Association between human error and occupational accidents' contributing factors for hand injuries in the automotive manufacturing industry. *Procedia Manufacturing*, 3, 6498-6504.
- Sadeghi, L., Dantan, J.-Y., Siadat, A., & Marsot, J.** (2016). Design for human safety in manufacturing systems: applications of design theories, methodologies, tools and techniques. *Journal of engineering design*, 27(12), 844-877.
- Schmidhuber, J.** (2015). Deep learning in neural networks: An overview. *Neural networks*, 61, 85-117.
- Seo, J., Han, S., Lee, S., & Kim, H.** (2015). Computer vision techniques for construction safety and health monitoring. *Advanced Engineering Informatics*, 29(2), 239-251.
- Shi, X., Chen, Z., Wang, H., Yeung, D.-Y., Wong, W.-K., & Woo, W.-c.** (2015). Convolutional LSTM network: A machine learning approach for precipitation nowcasting. *Advances in neural information processing systems*, 28.
- Singh, R., Kushwaha, A. K. S., & Srivastava, R.** (2023). Recent trends in human activity recognition—A comparative study. *Cognitive Systems Research*, 77, 30-44.
- Son, H., Choi, H., Seong, H., & Kim, C.** (2019). Detection of construction workers under varying poses and changing background in image sequences via very deep residual networks. *Automation in Construction*, 99, 27-38.
- Strauch, B.** (2017). *Investigating human error: Incidents, accidents, and complex systems*: CRC Press.
- Takala, J., Hämäläinen, P., Saarela, K. L., Yun, L. Y., Manickam, K., Jin, T. W., Heng, P., Tjong, C., Kheng, L. G., & Lim, S.** (2014). Global estimates of the burden of injury and illness at work in 2012. *Journal of occupational and environmental hygiene*, 11(5), 326-337.
- Tompa, E., Mofidi, A., van den Heuvel, S., van Bree, T., Michaelsen, F., Jung, Y., Porsch, L., & van Emmerik, M.** (2021a). Economic burden of work injuries and diseases: a framework and application in five European Union countries. *BMC Public Health*, 21, 1-10.
- Tompa, E., Mofidi, A., van den Heuvel, S., van Bree, T., Michaelsen, F., Jung, Y., Porsch, L., & van Emmerik, M.** (2021b). Economic burden of work injuries and diseases: a framework and application in five European Union countries. *BMC Public Health*, 21(1), 1-10.

- Tran, D., Bourdev, L., Fergus, R., Torresani, L., & Paluri, M.** (2015). Learning spatiotemporal features with 3d convolutional networks. *Proceedings of the IEEE international conference on computer vision*.
- Tzutalin, D.** (2015). Labellmg. *GitHub repository*, 6, 4.
- Uğurlu, Ö., Yıldırım, U., & Başar, E.** (2015). Analysis of grounding accidents caused by human error. *Journal of Marine Science and Technology*, 23(5), 19.
- Van Eck, N., & Waltman, L.** (2020). VOSviewer Manual. Universiteit Leiden.
- Wang, A., Chen, H., Liu, L., Chen, K., Lin, Z., Han, J., & Ding, G.** (2024). YOLOv10: Real-Time End-to-End Object Detection. *arXiv preprint arXiv:2405.14458*.
- Wang, L., Xiong, Y., Wang, Z., Qiao, Y., Lin, D., Tang, X., & Van Gool, L.** (2016). Temporal segment networks: Towards good practices for deep action recognition. *European conference on computer vision*.
- Wei, R., Love, P. E., Fang, W., Luo, H., & Xu, S.** (2019). Recognizing people's identity in construction sites with computer vision: A spatial and temporal attention pooling network. *Advanced Engineering Informatics*, 42, 100981.
- Wolf, C., Lombardi, E., Mille, J., Celiktutan, O., Jiu, M., Dogan, E., Eren, G., Baccouche, M., Dellandréa, E., & Bichot, C.-E.** (2014). Evaluation of video activity localizations integrating quality and quantity measurements. *Computer Vision and Image Understanding*, 127, 14-30.
- Wu, D., Lv, S., Jiang, M., & Song, H.** (2020). Using channel pruning-based YOLO v4 deep learning algorithm for the real-time and accurate detection of apple flowers in natural environments. *Computers and Electronics in Agriculture*, 178, 105742.
- Wu, H., & Zhao, J.** (2018). An intelligent vision-based approach for helmet identification for work safety. *Computers in Industry*, 100, 267-277.
- Wu, J., Cai, N., Chen, W., Wang, H., & Wang, G.** (2019). Automatic detection of hardhats worn by construction personnel: A deep learning approach and benchmark dataset. *Automation in Construction*, 106, 102894.
- Yang, M., Wu, C., Guo, Y., Jiang, R., Zhou, F., Zhang, J., & Yang, Z.** (2023). Transformer-based deep learning model and video dataset for unsafe action identification in construction projects. *Automation in Construction*, 146, 104703.

**Yao, L., Torabi, A., Cho, K., Ballas, N., Pal, C., Larochelle, H., & Courville, A.** (2015). Describing videos by exploiting temporal structure. *Proceedings of the IEEE international conference on computer vision*.

**Yu, Y., Guo, H., Ding, Q., Li, H., & Skitmore, M.** (2017). An experimental study of real-time identification of construction workers' unsafe behaviors. *Automation in Construction*, 82, 193-206.

**Zhu, Y., Lan, Z., Newsam, S., & Hauptmann, A.** (2019). Hidden two-stream convolutional networks for action recognition. *Computer Vision–ACCV 2018: 14th Asian Conference on Computer Vision, Perth, Australia, December 2–6, 2018, Revised Selected Papers, Part III 14*.

**Zou, Z., Chen, K., Shi, Z., Guo, Y., & Ye, J.** (2023). Object detection in 20 years: A survey. *Proceedings of the IEEE*, 111(3), 257-276.