

T.C.
SAKARYA ÜNİVERSİTESİ
FEN BİLİMLERİ ENSTİTÜSÜ

SİSLİ HAVA KOŞULLARINDA OTONOM ARAÇLAR İÇİN DERİN
ÖĞRENME TABANLI GÖRÜNTÜ ÖN İŞLEME İLE NESNE
TESPİTİ

YÜKSEK LİSANS TEZİ

Baha Hakkı DEMİRCİOĞLU

Bilgisayar Mühendisliği Anabilim Dalı

AĞUSTOS 2025

T.C.
SAKARYA ÜNİVERSİTESİ
FEN BİLİMLERİ ENSTİTÜSÜ

SİSLİ HAVA KOŞULLARINDA OTONOM ARAÇLAR İÇİN DERİN
ÖĞRENME TABANLI GÖRÜNTÜ ÖN İŞLEME İLE NESNE
TESPİTİ

YÜKSEK LİSANS TEZİ

Baha Hakkı DEMİRCİOĞLU

Bilgisayar Mühendisliği Anabilim Dalı

Tez Danışmanı: Prof. Dr. Nejat YUMUŞAK

AĞUSTOS 2025

Baha Hakkı DEMİRCİOĞLU tarafından hazırlanan “SİSLİ HAVA KOŞULLARINDA OTONOM ARAÇLAR İÇİN DERİN ÖĞRENME TABANLI GÖRÜNTÜ ÖN İŞLEME İLE NESNE TESPİTİ” adlı tez çalışması 19.08.2025 tarihinde aşağıdaki jüri tarafından oy birliği/oy çokluğu ile Sakarya Üniversitesi Fen Bilimleri Enstitüsü Bilgisayar Mühendisliği Anabilim Dalı’nda Yüksek Lisans tezi olarak kabul edilmiştir.

Tez Jürisi

Jüri Başkanı :	Prof. Dr. Nejat Yumuşak (Danışman)
	Sakarya Üniversitesi
Jüri Üyesi :	Prof. Dr. Halit Öztekin
	Sakarya Uygulamalı Bilimler Üniversitesi
Jüri Üyesi :	Dr. Öğr. Üyesi Seçkin Arı
	Sakarya Üniversitesi



ETİK İLKE VE KURALLARA UYGUNLUK BEYANNAMESİ

Sakarya Üniversitesi Fen Bilimleri Enstitüsü Lisansüstü Eğitim-Öğretim Yönetmeliğine ve Yükseköğretim Kurumları Bilimsel Araştırma ve Yayın Etiği Yönergesine uygun olarak hazırlamış olduğum “SİSLİ HAVA KOŞULLARINDA OTONOM ARAÇLAR İÇİN DERİN ÖĞRENME TABANLI GÖRÜNTÜ ÖN İŞLEME İLE NESNE TESPİTİ” başlıklı tezin bana ait, özgün bir çalışma olduğunu; çalışmamın tüm aşamalarında yukarıda belirtilen yönetmelik ve yönergeye uygun davrandığımı, tezin içerdiği yenilik ve sonuçları başka bir yerden almadığımı, tezde kullandığım eserleri usulüne göre kaynak olarak gösterdiğimi, bu tezi başka bir bilim kuruluna akademik amaç ve unvan almak amacıyla vermediğimi ve 20.04.2016 tarihli Resmi Gazete’de yayımlanan Lisansüstü Eğitim ve Öğretim Yönetmeliğinin 9/2 ve 22/2 maddeleri gereğince Sakarya Üniversitesi’nin abonesi olduğu intihal yazılım programı kullanılarak Enstitü tarafından belirlenmiş ölçütlere uygun rapor alındığını, çalışmamla ilgili yaptığım bu beyana aykırı bir durumun ortaya çıkması halinde doğabilecek her türlü hukuki sorumluluğu kabul ettiğimi beyan ederim.

(...../...../20.....).

Baha Hakkı DEMİRCİOĞLU





Ailelerime...



TEŞEKKÜR

Bu tez çalışmasının her aşamasında bilgi ve tecrübesiyle bana rehberlik eden, akademik bakış açımı geliştirmemde büyük katkısı olan; sabrı, yönlendirmeleri ve teşvik edici tutumuyla çalışmama yön veren değerli danışmanım Prof. Dr. Nejat Yumuşak'a en içten teşekkürlerimi sunarım. Kendisinin yol göstericiliği, bu çalışmanın bilimsel bir bütünlük içinde tamamlanmasında belirleyici olmuştur.

Bu sürecin her anında yanımda olan; sabırları, sevgileri ve koşulsuz destekleriyle beni her zaman motive eden ailelerime en derin şükranlarımı sunarım. Varlıkları, bu süreci anlamlı ve sürdürülebilir kılmıştır.

Baha Hakkı DEMİRCİOĞLU



İÇİNDEKİLER

Sayfa

ETİK İLKE VE KURALLARA UYGUNLUK BEYANNAMESİ	v
TEŞEKKÜR	ix
İÇİNDEKİLER	xi
KISALTMALAR	xiii
SİMGELER	xv
TABLO LİSTESİ	xvii
ŞEKİL LİSTESİ	xix
ÖZET	xxi
SUMMARY	xxiii
1. GİRİŞ	1
1.1. Tezin Önemi ve Kapsamı	2
1.2. Tezin Amacı	3
1.3. Hipotez	4
2. LİTERATÜR TARAMASI	5
2.1. Görüntü İyileştirme Yöntemleri (Defogging)	5
2.2. Nesne Tespiti Mimarileri ve Adaptif Yaklaşımlar	7
2.3. Zorlu Koşullarda Nesne Tespiti	8
2.4. Literatürdeki Boşluk ve Bu Tezin Konumu	10
3. MATERYAL VE YÖNTEM	13
3.1. Kullanılan Veri Kümeleri	13
3.2. Model Geliştirme Süreci	14
3.2.1. YOLOv12m mimarisi	14
3.2.2. Sis giderme modülü (Defogging)	15
3.3. Eğitim Parametreleri ve Kayıp Fonksiyonu	19
3.4. Model Değerlendirme Metrikleri	20
3.4.1. Precision ve recall	20
3.4.2. Ortalama ortalama kesinlik (mAP@0.5:0.95)	21
3.4.3. Yapısal benzerlik indeksi (SSIM)	21
3.4.4. Tepe sinyal gürültü oranı (PSNR)	22
4. ARAŞTIRMA BULGULARI	23
4.1. Ön İşleme Modülünün Etkileri	23
4.2. Değerlendirme Metrikleri	24
4.3. Eğitim Süreci	24
4.3.1. Ön işleme (defogging) modülünün eğitimi	25
4.3.2. YOLOv12m modelinin eğitimi	25
4.4. Test Sonuçları	26
4.5. Karşılaştırmalı Analiz	30
5. TARTIŞMA	33
5.1. Literatürle Karşılaştırmalı Değerlendirme	33
5.2. Hipotezin Değerlendirilmesi	34
5.3. Katkılar ve Sınırlılıklar	34

5.4. Veri Seti Seçiminin Gerekçelendirilmesi ve Literatürdeki Alternatiflerinin Değerlendirilmesi	36
6. SONUÇ VE ÖNERİLER.....	39
KAYNAKLAR.....	43
ÖZGEÇMİŞ.....	47



KISALTMALAR

AP	: Ortalama Doğruluk (Average Precision)
CIoU	: Tam Kesişim/Birleşim Oranı (Complete Intersection over Union)
CNN	: Evrişimli Sinir Ağı (Convolutional Neural Network)
COCO	: Bağlamdaki Yaygın Nesneler (Common Objects in Context)
CSPNet	: Çapraz Aşama Kısmi Ağı (Cross Stage Partial Network)
DCP	: Karanlık Kanal Önceliği (Dark Channel Prior)
Defog YOLO	: Sis Giderme Özellikli YOLO
Faster R-CNN	: Bölge Tabanlı Hızlı Evrişimli Sinir Ağı (Faster Region-based Convolutional Neural Network)
FOG-TRAINVAL	: Sisli Görüntü Eğitim ve Doğrulama Veri Kümesi
FPS	: Saniyedeki Kare Sayısı (Frame Per Second)
GAN	: Üretici Çekişmeli Ağ (Generative Adversarial Network)
GPU	: Grafik İşlem Birimi (Graphics Processing Unit)
IoU	: Kesişim/Birleşim Oranı (Intersection over Union)
LiDAR	: Işık Algılama ve Menzil Tayini (Light Detection and Ranging)
mAP	: Ortalama Ortalama Doğruluk (Mean Average Precision)
mAP@0.5:0.95	: IoU 0.5:0.95 aralığında Ortalama Ortalama Doğruluk
MMA-Net	: Çok modlu Dikkat Ağı (Multi-Modal Attention Network)
NAA	: Gürültüye Duyarlı Dikkat Mekanizması (Noise-Aware Attention)
NMS	: Maksimum Olmayan Bastırma (Non-Maximum Suppression)
P	: Doğruluk (Precision)
PANet	: Yol Birleştirme Ağı (Path Aggregation Network)
PSNR	: Tepe Sinyal-Gürültü Oranı (Peak Signal-to-Noise Ratio)
R	: Hatırlama / Duyarlılık (Recall)
RADAR	: Radyo Algılama ve Menzil Tayini (Radio Detection and Ranging)
RGB	: Kırmızı-Yeşil-Mavi Renk Uzayı (Red-Green-Blue)
RTTS	: Gerçek Dünya Görev Odaklı Test Seti (Real-world Task-driven Testing Set)

SSD	: Tek Atışlı Çoklu Kutu Algılayıcı (Single Shot MultiBox Detector)
SSIM	: Yapısal Benzerlik İndeksi (Structural Similarity Index)
VOC	: Görsel Nesne Sınıflandırma Veri Kümesi (Visual Object Classes)
YOLO	: Gerçek zamanlı nesne tespiti sistemi (You Only Look Once)



SİMGELER

A	: Dikkat maskesi (Attention mask)
B	: Yığın Boyutu (Batch Size)
C₁	: Sayısal kararlılık sabiti
C₂	: Sayısal kararlılık sabiti
box	: Sınırlayıcı Kutu Kaybı (Bounding Box Loss)
cls	: Sınıflandırma Kaybı (Classification Loss)
dB	: Desibel
dfl	: Dağılım Odak Kaybı (Distribution Focal Loss)
FP	: Yanlış pozitif tahminler (False Positives)
FN	: Gözden kaçan nesneler (False Negatives)
H	: Görüntü Piksel Yüksekliği (Image Height)
I	: Giriş görüntüsü (Input Image)
I'	: İşlenmiş (ön işleme tabi tutulmuş) görüntü
MAX	: Maksimum piksel değeri
N	: Gürültü bileşeni
P	: Doğruluk (Precision)
R	: Hatırlama / Duyarlılık (Recall)
W	: Ağırlık matrisi (Weight Matrix)
TP	: Doğru pozitif tahminler (True Positives)
W_{width}	: Görüntü Piksel Genişliği (Image Width)
ΔC	: Kontrast değişimi
μ	: Ortalama parlaklık değeri
σ	: Kontrast
σ^2	: Kontrast varyansı



TABLO LİSTESİ

Sayfa

Tablo 2.1. Kullanılan Veri Setlerinin Genel Özellikleri.....	14
Tablo 3.1. Nesne Algılama Performans Karşılaştırması	30





ŞEKİL LİSTESİ

Sayfa

Şekil 2.1. Sis giderme model mimarisini gösteren şematik gösterim	16
Şekil 2.2. Geliştirilen sistemin genel mimarisi	18
Şekil 2.3. İlerleyen epochlarda (eğitim döngülerinde) modelin doğrulama kaybı değerleri.....	20
Şekil 3.1. İlerleyen epochlarda (eğitim döngülerinde) modelin P değerleri	27
Şekil 3.2. İlerleyen epochlarda (eğitim döngülerinde) modelin R değerleri.....	28
Şekil 3.3. İlerleyen epochlarda (eğitim döngülerinde) modelin mAP@0.5:0.95 değerleri.....	28
Şekil 3.4. Geliştirilen modelin, RTTS test veri setindeki farklı ve zorlu sis koşulları altında nesne tespiti performansını gösteren örnek çıktılar.....	29



SİSLİ HAVA KOŞULLARINDA OTONOM ARAÇLAR İÇİN DERİN ÖĞRENME TABANLI GÖRÜNTÜ ÖN İŞLEME İLE NESNE TESPİTİ

ÖZET

Anahtar Kelimeler: Otonom Araçlar, Nesne Tespiti, Sisli Hava Koşulları, Derin Öğrenme, Sis Giderme, Bilgisayarla Görü

Otonom araçların güvenli ve verimli bir şekilde çalışabilmesi, çevresel koşulları doğru bir şekilde algılayan gelişmiş algı sistemlerine bağlıdır. Ancak, sisli hava koşulları gibi olumsuz atmosferik şartlar, görüntü kalitesini önemli ölçüde düşürerek nesne tespiti performansını ciddi şekilde sınırlandırmakta ve bu sistemlerin güvenilirliğini tehlikeye atmaktadır. Bu tez çalışmasında, sisli hava koşullarında otonom sürüş sistemlerinin nesne tespiti başarısını artırmak amacıyla, görüntüye özgü adaptif bir ön işleme (defogging) modülü ile desteklenen ve güncel bir nesne tespiti mimarisi olan YOLOv12m'yi kullanan iki aşamalı özgün bir sistem önerilmiştir.

Geliştirilen sisli hava koşullarına duyarlı ön işleme modülü, her bir görüntüdeki sis yoğunluğunu ve kontrast dağılımını dinamik olarak analiz ederek, geleneksel sabit filtreleme yöntemlerinin aksine, sahneye özgü bir iyileştirme sağlamaktadır. Bu modül, iyileştirilmiş ve detayları zenginleştirilmiş görüntüleri, nesne tespiti doğruluğunu ve hızını optimize eden YOLOv12m mimarisine girdi olarak sunarak modüler ve yüksek performanslı bir çözüm oluşturmaktadır.

Modelin etkinliği, hem sentetik sis içeren FOG-TRAINVAL hem de gerçek dünya görüntülerinden oluşan RTTS veri kümeleri üzerinde test edilmiştir. Değerlendirme, hem görsel kalite metrikleri (PSNR ve SSIM) hem de nesne tespiti performans metrikleri (Precision, Recall ve mAP@0.5:0.95) kullanılarak kapsamlı bir şekilde yapılmıştır.

Deneysel sonuçlar, önerilen sistemin RTTS veri seti üzerinde %68.8 mAP@0.5:0.95, %88.4 P ve %78.2 R skorlarına ulaştığını göstermiştir. Bu başarı, yalnızca ön işleme tabi tutulmamış YOLOv12m mimarisine göre önemli bir artışı temsil etmekle kalmayıp, aynı zamanda literatürdeki benzer hibrit yaklaşımları da geride bırakmaktadır. Bu çalışma, sisli hava koşullarında otonom araçların algı performansını ve güvenilirliğini artıran, sahneye duyarlı ve iki aşamalı modüler bir yaklaşım sunarak literatüre önemli bir katkı sağlamaktadır.



OBJECT DETECTION WITH DEEP LEARNING-BASED IMAGE PREPROCESSING FOR AUTONOMOUS VEHICLES IN FOGGY WEATHER CONDITIONS

SUMMARY

Keywords: Autonomous Vehicles, Object Detection, Foggy Weather Conditions, Deep Learning, Dehazing, Computer Vision

The rapid advancement of autonomous vehicle technology promises a revolutionary transformation in transportation, with the potential to significantly enhance road safety, improve traffic efficiency, and increase mobility. The foundation of this technology lies in the vehicle's ability to accurately perceive its surrounding environment, make informed decisions, and execute precise maneuvers without human intervention. This perceptual capability is enabled by a sophisticated suite of sensors, including LiDAR, RADAR, and cameras, which work in concert to build a comprehensive model of the world. Among these, camera-based vision systems are indispensable due to their ability to capture rich, high-resolution color and texture information, which is crucial for identifying and classifying a wide range of objects such as other vehicles, pedestrians, cyclists, and traffic signs. The cost-effectiveness and data richness of cameras make them a cornerstone of modern autonomous driving systems.

However, the performance of these vision systems is highly susceptible to degradation from adverse atmospheric conditions. While ideal, clear-weather scenarios allow for optimal operation, real-world driving frequently involves challenging environments. Among these, foggy weather presents one of the most significant and dangerous obstacles. Fog is composed of suspended water droplets that scatter and absorb light, leading to a severe degradation of image quality. This manifests as a dramatic loss of contrast, a washing-out of colors, the obscuring of fine details, and the blurring of object boundaries. Consequently, the visibility range is drastically reduced, and the salient features that object detection algorithms rely on become faint or disappear entirely. This degradation is not merely a cosmetic issue; it directly compromises the reliability of the perception system, creating a critical safety risk. An autonomous vehicle that cannot confidently detect a pedestrian or a stopped car in fog is unable to operate safely. Therefore, developing robust perception systems that can overcome the challenges posed by foggy conditions is not just an academic pursuit but a fundamental requirement for the widespread and safe deployment of autonomous vehicles.

To address the performance decline of object detectors in fog, significant research has been dedicated to image "dehazing" or "defogging"—a preprocessing step aimed at algorithmically removing the effects of fog from an image to restore its underlying clarity. Historically, many approaches were based on physical models of light scattering, such as the atmospheric scattering model. Techniques like the Dark Channel Prior (DCP) method achieved notable success by estimating the fog density (transmission map) and atmospheric light to invert the degradation process. While foundational, these physics-based methods often rely on assumptions that may not

hold true in complex or dynamic scenes (e.g., areas with bright white objects or sky regions), leading to artifacts and color distortions. Furthermore, their computational complexity can be a bottleneck for real-time applications.

The advent of deep learning has ushered in a new era of data-driven dehazing techniques. Convolutional Neural Networks (CNNs) and Generative Adversarial Networks (GANs) have demonstrated a remarkable ability to learn the complex, non-linear transformation from a foggy image to its clear counterpart. These models can be trained on large datasets of paired foggy and clear images to produce visually stunning results, restoring vibrant colors and sharp details with high fidelity. However, a critical limitation in many existing deep learning-based systems is their "one-size-fits-all" approach. They often employ static, fixed-parameter filters that are applied uniformly to every input image, regardless of the specific characteristics of the scene. This is a significant drawback because fog is rarely uniform; its density can vary dramatically across a single image and between different scenes. A static filter may be too aggressive for light fog, causing artifacts, or insufficient for dense fog, leaving critical information obscured.

This limitation highlights a crucial gap in the literature: the need for adaptive, scene-aware preprocessing. An ideal system should be able to dynamically analyze the content of each specific image—assessing the fog density, contrast distribution, and spatial characteristics—and apply a tailored enhancement. This adaptability is the key to achieving robust performance across the wide spectrum of conditions encountered in real-world driving. This thesis addresses this gap by proposing a system where the preprocessing stage is not a fixed filter but an intelligent, adaptive module designed to work in synergy with a state-of-the-art object detector.

This thesis proposes an innovative two-stage, modular system designed to maximize object detection performance in foggy conditions. The core philosophy is to decouple the complex problem into two specialized tasks: adaptive image enhancement and high-performance object detection. This modularity allows each component to be optimized for its specific function and provides flexibility for future upgrades. The system pipeline feeds the enhanced, detail-rich output from the first stage directly into the second, creating a solution that is both powerful and efficient.

At the heart of our proposed system is weather-adaptive preprocessing module. Unlike traditional methods, this module does not rely on a fixed set of filtering rules. Instead, it employs a lightweight CNN designed as a "parameter prediction network." For every incoming foggy image, this network performs an analysis to estimate the optimal enhancement parameters tailored to that specific scene. It dynamically assesses the fog's impact on contrast and detail, and then intelligently adjusts the processing pipeline.

This process involves predicting key image transformation parameters, such as the gamma correction factor to adjust luminance and contrast, and the unsharp masking intensity to enhance edges and fine details. By making the enhancement process content-dependent, the module provides a scene-specific restoration. For instance, in images with dense fog, it can apply stronger contrast enhancement, while for scenes with light, patchy fog, it can use a more subtle correction to avoid over-enhancement artifacts. The result is a consistently improved image that is rich in detail and structurally faithful to the original clear scene, providing an ideal input for the subsequent detection task.

The enhanced images generated by the preprocessing module are fed into a state-of-the-art object detection architecture, YOLOv12m. This model was chosen for its exceptional balance of speed and accuracy, making it highly suitable for real-time autonomous driving applications. The YOLOv12m architecture builds upon the strengths of its predecessors with several key advancements. It features a deep and efficient backbone network for powerful feature extraction, capable of identifying complex patterns even in challenging visual data. Furthermore, it incorporates advanced multi-scale feature fusion mechanisms, such as a Path Aggregation Network (PANet), which allow it to effectively detect objects of various sizes, from large trucks to distant pedestrians. By integrating attention mechanisms, the model can also focus on the most salient regions of the image, further improving its localization accuracy. By training this powerful detector on the clean, high-quality images produced by our adaptive defogging module, we create a synergistic system where both components work together to achieve a level of performance that neither could reach alone.

The effectiveness of the proposed system was rigorously evaluated through a series of comprehensive experiments. To ensure a thorough assessment, two distinct and challenging datasets are utilized:

FOG-TRAINVAL: A dataset containing a large number of synthetic foggy images, used primarily for training the models. Its diverse scenes and controllable fog levels allow the system to learn to generalize across various conditions.

RTTS (Real-world Task-driven Testing Set): A dataset composed of real-world images captured in genuine foggy weather. This dataset serves as the ultimate benchmark for evaluating the system's performance in authentic, challenging driving scenarios.

The evaluation was multi-faceted, employing both image quality metrics and object detection performance metrics.

Image Quality Assessment: The success of the defogging module was quantified using the Peak Signal-to-Noise Ratio (PSNR) and the Structural Similarity Index (SSIM). Our system achieved an impressive average PSNR of 31.13 dB and an average SSIM of 0.9729, indicating that the enhanced images were not only clearer but also maintained high structural fidelity to the original scenes.

Object Detection Performance: The primary goal, object detection, was measured using standard metrics: Precision (P), Recall (R), and mean Average Precision (mAP@0.5:0.95).

The experimental results demonstrated the profound impact of our adaptive preprocessing module. On the challenging real-world RTTS dataset, the complete system achieved the scores below:

mAP@0.5:0.95: 68.8%

Precision: 88.4%

Recall: 78.2%

This performance represents a significant leap over the baseline YOLOv12m model operating without preprocessing. More importantly, it also surpasses the performance of similar hybrid approaches reported in the literature. For example, our system demonstrated a relative improvement of approximately 14.7% in mAP compared to other strong defogging-plus-detection models. This success validates our central

hypothesis: an intelligent, adaptive preprocessing stage that is sensitive to the specific scene conditions can dramatically enhance the capability and reliability of object detection systems in fog.

Adverse weather conditions, particularly fog, remain a major impediment to the safe and reliable operation of autonomous vehicles. This thesis aimed to make a significant contribution to the literature by presenting a scene-aware, two-stage modular system that effectively mitigates the impact of fog on camera-based perception. By combining a dynamic, weather-adaptive image preprocessing module with a state-of-the-art object detector (YOLOv12m), our approach enhances both the perceptual quality of the input data and the functional accuracy of the final detection task. The experimental results, validated on challenging real-world datasets, confirm that our system not only achieves a substantial improvement over baseline models but also outperforms similar hybrid approaches. This work provides a robust and effective framework for enhancing the perception systems of autonomous vehicles, thereby increasing their operational safety and reliability in the challenging and unpredictable conditions of real-world driving.



1. GİRİŞ

Otonom araçlar, insan müdahalesine gerek duymadan çevresel koşulları algılayarak hareket edebilen, yapay zekâ ve sensör teknolojilerinin bütünleşik kullanımıyla geliştirilen sistemlerdir. Ulaşım alanında devrim niteliğinde kabul edilen bu araçlar, özellikle güvenlik, verimlilik ve sürdürülebilirlik hedefleri doğrultusunda ön plana çıkmaktadır.

Otonom araçlar, son yıllarda güvenli sürüş, trafik yoğunluğunun azaltılması ve insan hatalarının en aza indirilmesi gibi hedeflerle geliştirilen ileri teknolojilerin başında gelmektedir. Bu araçların çevreyle etkileşim kurarak doğru ve hızlı kararlar verebilmesi, temel olarak çevresel algı sistemlerinin performansına bağlıdır. Ren ve arkadaşlarına göre (2015) kamera, ışık algılama ve menzil tayini (lidar) ve radyo algılama ve menzil tayini (radar) gibi çeşitli sensörler kullanılarak toplanan veriler; görüntü işleme, bilgisayarla görü ve derin öğrenme teknikleriyle analiz edilerek araçların çevresel farkındalığı oluşturulmaktadır.

Ancak bu süreç, dışsal koşullardan etkilenmeye oldukça açıktır. Özellikle sisli hava, görüş netliğini azaltarak kameralar tarafından elde edilen görüntülerin kalitesini önemli ölçüde düşürür ve bu durum nesne tespiti gibi temel bilgisayarla görü görevlerinin sağlıklı şekilde yerine getirilmesini zorlaştırır. Sakaridis ve arkadaşlarına göre (2018) görüş mesafesinin kısılması, yalnızca doğrudan algılanabilen alanı daraltmakla kalmaz; aynı zamanda nesne kenarlarının bulanıklaşması, renk kontrastının zayıflaması ve detayların kaybolması gibi görsel bozulmalarla da sonuçlanır. Bu tür bozulmalar, çevresel algı sistemlerinin doğru veri üretmesini engellerken, bu verilere dayalı rota planlama, çarpışma önleme ve hız kontrolü gibi ileri seviye karar mekanizmalarını da olumsuz etkiler. Sisli görüntüler, yalnızca algılama modülünü değil; aynı zamanda otonom sürüşün tüm aşamalarını doğrudan sekteye uğratmaktadır. Bu nedenle, düşük görüş koşullarında güvenilirliğin artırılması ve çevresel algının kararlılıkla sürdürülebilmesi, otonom sistemlerin pratik uygulamalarda etkin biçimde çalışabilmesi için temel bir gereklilik hâline gelmiştir.

Bu bağlamda, Zhang ve arkadaşlarına göre (2022) son yıllarda sisin etkisini azaltmaya yönelik ön işleme (defogging) yöntemleri ve bu yöntemlerin entegre edildiği derin öğrenme mimarileri üzerine birçok çalışma yapılmıştır. Bu çalışmaların temel amacı, sisli görüntülerin görsel niteliğini iyileştirerek nesne tespiti süreçlerinin daha doğru ve güvenilir şekilde yürütülmesini sağlamaktır. Shi ve Song'a göre (2024) özellikle sahneye özgü uyarlanabilen (adaptif) ön işleme yaklaşımlarının, sis yoğunluğu ve sahne yapısına göre farklı işlem adımları uygulayabilmesi, geleneksel tek tip filtreleme yöntemlerine kıyasla daha etkili sonuçlar üretmektedir. Öte yandan, son nesil gerçek zamanlı nesne tespit sistemi (YOLO) mimarileri, nesne tespiti alanında gerçek zamanlı çalışma kapasitesiyle öne çıkmakta; Li ve arkadaşlarına göre (2025) YOLO'nun gelişmiş sürümleri ise yüksek işlem hızı ve doğruluk düzeyini birlikte sunarak otonom sistemlerde yaygın olarak tercih edilmektedir.

Bu çalışmada, otonom sürüş sistemlerinde sisli hava koşullarında nesne tespit performansını artırmaya yönelik olarak, görüntü ön işleme (defogging) modülünün çıktılarıyla beslenen bir YOLOv12m mimarisini temel alan ardışık bir sistem sunulmuştur. Modelin uygulanabilirliği, hem görsel kalite için tepe sinyal-gürültü oranı (PSNR) ve yapısal benzerlik indeksi (SSIM) ile hem de nesne tespiti için doğruluk (P), duyarlılık (R), Kesişim/Birleşim Oranı (IoU) 0.5:0.95 aralığında ortalama ortalama doğruluk (mAP@0.5:0.95) üzerinden değerlendirilmiştir.

Bu tezin izleyen bölümlerinde, öncelikle konuyla ilgili literatür detaylı biçimde incelenecek, ardından kullanılan yöntem ve veri kümeleri açıklanacaktır. Sonraki bölümlerde model performansı deneysel sonuçlarla ortaya konulacak, elde edilen bulgular literatürle karşılaştırılarak tartışılacak ve nihayetinde genel sonuçlara ve geleceğe yönelik önerilere yer verilecektir.

1.1. Tezin Önemi ve Kapsamı

Otonom sürüş teknolojilerinin başarısı, yalnızca gelişmiş donanım bileşenlerine değil, aynı zamanda bu bileşenlerden elde edilen verilerin işlenmesindeki etkinliğe bağlıdır. Sakaridis ve arkadaşlarının (2018) çalışması, Li ve arkadaşlarının (2025) çalışması ile desteklenerek; görüntü tabanlı algı sistemlerinin, maliyet-etkinlikleri ve detaylı çevresel bilgi sunabilmeleri nedeniyle yaygın biçimde tercih edilmekte olduğunu; ancak çevresel koşullardan doğrudan etkilenmeleri, sistemin genel performansını sınırlayan en önemli faktörlerden biri olarak öne çıktığını sunmaktadır. Zhang ve

arkadaşlarına göre (2022) özellikle sisli hava koşulları, görüntülerde kontrast ve renk doygunluğu kayıplarına neden olmakta; bu da nesnelerin algılanabilirliğini ve konumlandırılabilirliğini azaltmaktadır.

Mevcut çalışmalarda, bu sorunu aşmak için, görüntü ön işleme (defogging) tekniklerinin derin öğrenme temelli nesne tespit sistemleriyle birleştirilmesi öne çıkmıştır. Bununla birlikte, Shi ve Song'a göre (2024) çoğu sistem ya tüm görüntülere aynı filtreleme işlemini uygulamakta ya da sis yoğunluğu ve sahneye özgü farklılıkları dikkate almamaktadır. Bu durum, sistemlerin genellenebilirliğini sınırlamakta ve bazı sahnelerde yetersiz sonuçlara yol açmaktadır. Ayrıca, görüntü kalitesi yükselse dahi, nesne tespiti açısından gerçek performans kazanımı her zaman sağlanamamaktadır. Dolayısıyla, yalnızca görsel iyileştirmeyi değil; nesne tespit metriklerini de doğrudan yükselten sistemlere ihtiyaç duyulmaktadır.

Bu çalışmanın önemi, sisli ortamlarda çalışan derin öğrenme tabanlı nesne algılayıcılara yönelik özgün bir model önerisi sunmasında yatmaktadır. Sisli hava koşullarına duyarlı ön işleme ile YOLOv12m mimarisi modüler bir yaklaşımla bir araya getirilerek, görüntü kalitesi ve nesne tespit başarısı birlikte ele alınmış, geliştirilen sistem hem görsel değerlendirme hem de görev performansı açısından test edilmiştir. Bu bütünleştirilmiş yaklaşım, yalnızca akademik katkı sağlamamakta; aynı zamanda gerçek zamanlı otonom sistemler için uygulanabilir bir çerçeve de sunmaktadır.

1.2. Tezin Amacı

Bu tezin temel amacı, sisli hava koşullarında çalışan otonom araç sistemlerinde nesne algılama başarısını artırmak için görüntü ön işleme ve nesne tespiti süreçlerini birbirini takip eden iki aşamada ele alan özgün bir yaklaşım geliştirmektir. Bu modüler iki aşamalı yaklaşım, her bir bileşenin kendi görevinde (sis giderme ve nesne tespiti) uzmanlaşmasını sağlamak amacıyla benimsenmiştir. Bu tasarım, sistemin esnekliğini ve her bir modülün bağımsız olarak iyileştirilebilmesini mümkün kılmaktadır. Bu doğrultuda, sisli ortamlarda oluşan düşük kaliteli görüntülerin hem görsel hem de yapısal olarak iyileştirilmesini sağlayan, görüntüye özgü adaptasyon kabiliyetine sahip bir veri ön işleme modülü tasarlanmış; bu modülün çıktıları güncel nesne tespiti mimarilerinden biri olan YOLOv12m üzerinde kullanılmıştır.

Geliştirilen sisli hava koşullarına duyarlı yapı, görüntüdeki sis yoğunluğu ve kontrast dağılımına göre dinamik biçimde işlem uygulayarak geleneksel sabit filtreleme yöntemlerinin sınırlılıklarını aşmayı hedeflemektedir. Bu model, Shi ve Song (2024) tarafından sunulan, veri ön işleme yaklaşımından farklı olarak tasarlanmış; doğrudan bu çalışmaya özgü algoritmik yaklaşım içermektedir. Böylece, veri ön işleme aşamasında elde edilen görsel iyileştirmelerin, nesne tespiti başarısına etkisi bütüncül biçimde değerlendirilmiştir.

1.3. Hipotez

Sisli hava koşullarında elde edilen görüntülerde, nesne sınırlarının belirsizleşmesi, kontrast ve renk kayıpları gibi bozulmalar, derin öğrenme tabanlı nesne tespiti sistemlerinin doğruluğunu ciddi biçimde düşürmektedir. Bu bağlamda, görüntüye uygulanacak uygun bir ön işleme yöntemi ile bu bozulmaların giderilmesi ve nesne tespit doğruluğunun artırılması amaçlanmıştır.

Bu çalışmanın temel hipotezi, sisli hava koşullarında nesne tespitinin doğruluk ve güvenilirliğinin artırılabilmesi için, tespit sürecinden önce uygulanan uygun bir görüntü ön işleme yönteminin etkili olacaktır. Bu yaklaşımın, kesinlik (P, Precision) duyarlılık (R, Recall) ve mAP@0.5:0.95 gibi metriklerde elde edilen yüksek skorlarla desteklenerek daha sağlıklı tespit sonuçları sağlayacağı öngörülmektedir.

Bu hipotez, geliştirilen sisli hava koşullarına duyarlı ön işleme modülünün YOLOv12m mimarisiyle bütünleştirilmesi sonucunda, sistemin hem görsel kaliteyi iyileştirme kapasitesi hem de nesne algılama başarımını artırma potansiyeli açısından deneysel olarak test edilerek sınanmıştır.

2. LİTERATÜR TARAMASI

2.1. Görüntü İyileştirme Yöntemleri (Defogging)

Görüntü iyileştirme, bilgisayarla görü sistemlerinde görüntü kalitesinin artırılmasına yönelik uygulanan temel bir ön işlem adımıdır.

Özellikle sisli hava koşullarında çekilen görüntülerde, ışığın saçılması ve absorpsiyonu nedeniyle kontrast kaybı, renk bozulmaları ve yapısal belirsizlikler gibi sorunlar ortaya çıkmaktadır. Bu hem insan algısını hem de makine görüşüne dayalı otomatik sistemlerin performansını olumsuz etkilemektedir. Bu nedenle sisin etkilerini ortadan kaldırmaya yönelik geliştirilen “defogging” algoritmaları, görüntü iyileştirme literatürünün önemli bir araştırma alanı hâline gelmiştir.

Geleneksel defogging yöntemleri genellikle fiziksel modellere dayanır ve görüntünün sis yoğunluğunu açıklayan atmosferik yayılım modeli üzerinden çalışır. Bu yaklaşımda bir sahnedeki her pikselin parlaklığı, ortam ışığı ve aktarım katsayısı gibi parametrelerle modellenir. He ve arkadaşları (2010) tarafından önerilen “Dark Channel Prior” (DCP) yöntemi, bu tür fiziksel modellere dayanan yaklaşımların en bilinen örneklerinden biridir ve yerel minimum piksel değerleri üzerinden sisin derecesini tahmin etmeye çalışır. Ancak bu yöntemler, özellikle ışık kaynaklarının yoğun olduğu veya yüzey yansımalarının belirgin olduğu sahnelerde hatalı sonuçlar üretebilmekte; ayrıca işlem süresi açısından da gerçek zamanlı uygulamalarda sınırlılıklar oluşturmaktadır.

Son yıllarda derin öğrenme tabanlı yöntemler, defogging alanında önemli ilerlemeler sağlamıştır. Bu yöntemler, sisli ve sisli olmayan görüntüler arasında dönüşümü öğrenerek hem görsel kaliteyi hem de yapısal bütünlüğü daha başarılı şekilde yeniden üretmektedir. Özellikle Üretici Çekişmeli Ağ (GAN) temelli modeller, gerçekçi görseller üretebilme kabiliyeti sayesinde ön plana çıkmıştır. Li ve arkadaşları (2018), sisli görüntülerin iyileştirilmesinde GAN temelli çok katmanlı bir çevirici yapı önererek dikkat çekici sonuçlar elde etmiş; model hem yapay hem de gerçek sisli veri setlerinde yüksek PSNR ve SSIM skorlarıyla değerlendirilmiştir. Benzer biçimde Zhang ve arkadaşları (2020), çok düzeyli özellik birleştirme modülleri kullanarak sis

kaldırma başarısını artıran bir ağ tasarlamış ve özellikle düşük kontrastlı alanlarda, detay ve geri kazanım açısından başarılı sonuçlar elde etmiştir.

Bu alandaki bir diğer önemli çalışma ise, Liu ve arkadaşları (2019) tarafından önerilen GridDehazeNet modelidir. Dikkat (attention) mekanizmalarını çok ölçekli bir mimariyle birleştiren bu model, görüntünün farklı bölgelerindeki sis yoğunluğunu adaptif bir şekilde ele alarak, özellikle yoğun sis altındaki detayların ve kontrastın geri kazanılmasında yüksek başarı sergilemiştir. Atmosferik saçılım modeline dayanmayan, uçtan uca öğrenilebilir yapısıyla GridDehazeNet, hem sentetik hem de gerçek dünya görüntülerinde modern derin öğrenme yöntemlerinin gücünü ortaya koyan önemli bir örnektir.

Defogging algoritmalarının başarısı, genellikle SSIM ve PSNR gibi ölçütlerle değerlendirilmektedir. PSNR, görüntüdeki genel bozulma miktarını değerlendirirken; SSIM, insan görsel sistemini taklit ederek yapısal ve kontrast ilişkilerine dayalı bir kalite ölçümü sunar. Wang ve arkadaşlarına göre (2004) yüksek PSNR ve SSIM skorları, görüntü iyileştirme modülünün yalnızca görsel olarak değil, aynı zamanda yapısal bütünlük açısından da başarılı olduğunu göstermektedir.

Ayrıca dikkat çekici bir eğilim olarak, bazı çalışmalar görüntü iyileştirme sürecini yalnızca görsel kaliteye göre değil, aynı zamanda nesne algılama gibi ardıl görevler üzerinden de değerlendirmeye başlamıştır. Bu yaklaşım, sis giderme algoritmalarının yalnızca estetik düzeyde değil, aynı zamanda görev odaklı olarak da etkili olması gerektiğini savunmaktadır. Örneğin Sakaridis ve arkadaşları (2018), sis giderme modülünün nesne algılama sistemleriyle birlikte optimize edilmesinin, doğrudan algılama başarımını artırdığını göstermiştir.

Sonuç olarak, görüntü iyileştirme alanındaki ilerlemeler, özellikle sisli koşullar altında çalışan otonom sistemlerin başarısı açısından büyük önem taşımaktadır. Geleneksel fiziksel modellemeye dayalı yöntemler hâlen referans alınsa da günümüzde derin öğrenme temelli sis giderme algoritmaları hem kalite metriklerinde hem de nesne algılama performansında üstünlük sağlamaktadır. Bu bağlamda, geliştirilen adaptif ön işleme yaklaşımları, görsel içeriğin, bağlama özgü olarak iyileştirilmesini mümkün kılarak yeni nesil akıllı algı sistemlerinin önünü açmaktadır.

2.2. Nesne Tespiti Mimarileri ve Adaptif Yaklaşımlar

Nesne tespiti, temelde görüntü üzerinde bulunan nesnelerin hem sınıfını hem de konumunu belirlemeyi hedefleyen, bir bilgisayarla görü problemidir. Son yıllarda derin öğrenme tabanlı yapılar, bu alanda devrim niteliğinde gelişmeler sağlamış; özellikle gerçek zamanlı uygulamalar için doğruluk ve hız dengesi öncelikli hedef hâline gelmiştir.

Bu bağlamda, bölge tabanlı yaklaşımlar (two-stage) ile doğrudan regresyon tabanlı yaklaşımlar (one-stage) iki ana mimari çizgiyi temsil etmektedir. Ren ve arkadaşlarına göre (2015) Bölge Tabanlı Hızlı Evrişimli Sinir Ağı (Faster R-CNN) gibi bölge öneri tabanlı modeller, yüksek doğruluk sağlamasıyla bilinirken, hız açısından kısıtlayıcı olabilmektedir. Buna karşılık olarak Liu ve arkadaşlarına (2016) ve ayrıca Redmon ve arkadaşlarına (2016) ait yapılan çalışmalarda, geliştirilen Tek Atışlı Çoklu Kutu Algılayıcı (SSD) ve YOLO gibi mimariler, doğrudan sınıflandırma ve konum regresyonunu tek bir ağ üzerinde entegre ederek gerçek zamanlı nesne algılamaya olanak tanımaktadır.

YOLO ailesi, nesne tespitinde hem hız hem de doğruluk açısından giderek daha verimli hâle gelmiş ve farklı sürümlerle sürekli evrim geçirmiştir. Jiang ve arkadaşları (2022), YOLO mimarisinin versiyonlar arası evrimini sistemli biçimde inceleyerek; YOLOv1'den başlayarak YOLOv7'ye kadar olan gelişim sürecinde öne çıkan yapısal yenilikleri, hedeflenen problemleri ve performans artışlarını detaylı biçimde analiz etmişlerdir. Bu derleme, her yeni sürümün; örneğin çok ölçekli özellik haritaları, Çapraz Aşama Kısmi Ağı (CSPNet) yapısı, Yol Birleştirme Ağı (PANet) entegrasyonu veya dikkat mekanizmaları gibi katkılarla mimariyi daha işlevsel hâle getirdiğini göstermektedir. Bu bağlamda, bu çalışmada tercih edilen YOLOv12m mimarisi hem doğruluk oranı hem de çıkarım hızı açısından bu gelişimin güncel bir uzantısı olarak değerlendirilmektedir.

YOLOv12m'nin başarısında, backbone ve neck kısımlarındaki gelişmiş yapıların yanı sıra, görev odaklı özelleştirmelerin de önemli payı vardır. Özellikle PANet (Path Aggregation Network) benzeri katmanlar sayesinde farklı çözünürlüklerdeki semantik bilgilerin birleştirilmesi sağlanmış, bu sayede küçük boyutlu nesnelerin bile etkin biçimde tespit edilmesi mümkün olmuştur. Bunun yanında, çok başlı dikkat (multi-head attention) mekanizmalarının nesne konumlandırma doğruluğunu artırmak amacıyla modele entegre edilmesi de yaygınlaşmaktadır.

Son dönem çalışmalarda adaptif yaklaşımlar, görüntü koşullarına bağlı olarak parametrelerin veya ağ yollarının dinamik olarak değiştirilmesini sağlayan esnek sistemler sunmaktadır. Bu yaklaşımlar, özellikle zorlu çevre koşullarında –sisli, yağmurlu veya gece görüntüleri gibi– nesne tespiti başarımını artırmak için önem kazanmıştır. Örneğin Zhou ve arkadaşları (2021), farklı hava koşullarına özel filtreleme yolları tanımlayarak çoklu ortam adaptasyonu gerçekleştirmiş; Sakaridis ve arkadaşları (2018) ise alan uyarlaması teknikleri ile açık hava koşullarında eğitilmiş modellerin sisli ortamlara daha iyi genelleme yapmasını sağlamıştır.

Bu gelişmeler, yalnızca tasarım düzeyinde değil, aynı zamanda eğitim stratejileri düzeyinde de etkili olmuştur. Örneğin, Lin ve arkadaşlarına göre (2017) özellikle zorlu koşullarda, klasik çapraz entropi yerine focal loss gibi sınıf dengesizliğine duyarlı kayıp fonksiyonlarının tercih edilmesi, küçük ve az temsil edilen sınıfların başarıyla tespit edilmesini mümkün kılmaktadır.

Sonuç olarak, nesne tespiti mimarilerindeki ilerlemeler, yalnızca yapısal karmaşıklık değil, çevresel koşullara duyarlı adaptif yetenekleri de içerecek şekilde evrilmektedir. Bu bağlamda YOLOv12m gibi modern modellerin, uygun ön işleme yöntemleriyle desteklenmesi durumunda, özellikle sisli görüntüler gibi düşük kaliteli görsel girdilerinde dahi güçlü tespit performansı sergilemesi mümkündür.

2.3. Zorlu Koşullarda Nesne Tespiti

Otonom sistemlerin görsel algılamaya dayalı karar mekanizmaları, çevresel koşulların değişkenliğinden önemli ölçüde etkilenmektedir. Özellikle sisli, yağmurlu veya düşük ışıklı senaryolar, hem görüntü kalitesinde bozulmaya neden olmakta hem de nesne tespit modellerinin doğruluğunu ciddi biçimde düşürmektedir. Gui ve arkadaşlarına göre (2023) görüntüdeki kontrast azalması, kenar netliğinin kaybı ve renk doygunluğundaki düşüş gibi etkiler, tespit edilebilir nesne örüntülerinin silikleşmesine yol açmaktadır.

Bu tür zorlu hava koşullarında geliştirilen sis giderme tabanlı veri ön işleme yöntemlerinin etkili biçimde kullanılması, doğrudan nesne tespit başarısını artırabilmektedir. Gui ve arkadaşları (2023), sisli görüntülerde derin öğrenme tabanlı sis giderme algoritmalarının tespit doğruluğu üzerindeki etkisini analiz ederek, uygun ön işleme stratejilerinin yalnızca görüntü kalitesini değil, aynı zamanda tespit başarısını da olumlu etkilediğini göstermiştir. Benzer şekilde, Wang ve arkadaşları

(2023) tarafından yapılan çalışmada, atmosferik bozulmaları hesaba katan bir derin ön işleme modülü kullanılarak, YOLOv5 mimarisi ile yapılan deneylerde mAP değerinde gözle görülür bir artış elde edilmiştir.

Bununla birlikte, geleneksel görüntü iyileştirme yöntemlerinin, tespit odaklı sistemlerle doğrudan entegre edilememesi, sahada uygulanabilirliği sınırlandıran bir diğer faktördür. Bu bağlamda önerilen çözümlerden biri, görüntü iyileştirme ve tespit işlemlerini aynı hedef doğrultusunda optimize eden uçtan uca sistemlerdir. Ayrıca, modelin sisli ortama ne derece adapte olabildiği konusu, alan uyarlaması çalışmaları açısından da önemlidir. Lin ve arkadaşları (2022) tarafından yapılan araştırma, gerçek ve sentetik sisli görüntüler arasında alan farkı (domain gap) bulunduğunu ve bunun tespit performansını sınırladığını ortaya koymuştur. Bu nedenle, hem gerçek dünyadan hem de sentezlenmiş sisli veri kümelerinin bir arada kullanılması önerilmektedir. Bunun yanı sıra, Wang ve arkadaşlarına göre (2023), Li ve arkadaşlarının sunduğu (2018) Gerçek Dünya Görev Odaklı Test Seti (RTTS) gibi gerçek sisli sahnelerden oluşan test veri kümeleri, algoritmaların gerçek koşullardaki başarımını değerlendirmek için önemli bir ölçüt olarak görülmektedir. Bundan dolayı hem görüntü ön işleme modülünün hem de nihai modelin değerlendirilmesinde RTTS veri kümesi kullanılmıştır.

Son yıllarda bu alanda yapılan bazı çalışmalar, YOLO mimarisine doğrudan entegre edilen yapılarla hem parametre verimliliği sağlamayı hem de zorlu hava koşullarında tespit başarımını artırmayı hedeflemiştir. Örneğin Li ve Meng (2025), YOLOv8 tabanlı MFW-YOLOv8s modelini önererek MSBlock, Focal Modulation ve WIOU v2 Loss gibi modüllerle yapıyı yeniden tasarlamış; sisli, yağmurlu ve karlı sahnelerde %2.7'ye varan mAP artışı bildirmiştir. Bu tür mimari iyileştirmeler, hava koşullarına karşı dayanıklı nesne tespiti sistemlerinin geliştirilmesine önemli katkılar sunmaktadır.

Literatürdeki çalışmalar, zorlu hava koşullarında nesne tespitin yalnızca model mimarisiyle değil; aynı zamanda veri kalitesi, domain uyumu ve işlem sürecindeki bilgi kaybı gibi çok boyutlu etkenlerle yakından ilişkili olduğunu ortaya koymaktadır. Bu doğrultuda, sisli ortamlarda güvenilir nesne tespiti için çok aşamalı, bağlama duyarlı ve adaptif sistemlerin geliştirilmesi günümüzde öncelikli bir araştırma alanı haline gelmiştir.

2.4. Literatürdeki Boşluk ve Bu Tezin Konumu

Son yıllarda sisli hava koşullarında nesne tespiti üzerine yapılan çalışmalar artış göstermiş olsa da, mevcut literatürde hâlâ bazı yapısal ve yöntemsel sınırlılıklar mevcuttur. Zhang ve arkadaşlarına göre (2022) özellikle sisli görüntüler üzerinde nesne tespiti yapmaya yönelik geliştirmeler olmasına karşın, bu geliştirmelerin büyük bölümü statik ön işleme adımlarına dayanmakta ve görüntüdeki bozulma derecesine göre kendini uyarlama yetisine sahip değildir.

Sisli hava koşullarında nesne tespiti başarımını artırmaya yönelik bazı yaklaşımlar, gerçek ve sentetik sisli görüntüler arasındaki dağılım farkını azaltmak amacıyla alan uyarlaması tekniklerinden yararlanmıştır. Sakaridis ve arkadaşları (2018), sisli sahnelerdeki domain farklılıklarını adversarial öğrenme yöntemiyle azaltarak tespit doğruluğunu iyileştirmeyi hedeflemiştir. Ancak bu tür yöntemler, genellikle yüksek hesaplama maliyeti ve özel hedef domain verisi gereksinimi nedeniyle, her uygulama senaryosunda esnek biçimde kullanılamayabilmektedir. Bu çalışma, doğrudan görüntü ön işlemeye odaklanarak daha hafif ve modüler bir çözüm sunmayı amaçlamaktadır.

Sisli ortamlarda nesne tespiti amacıyla geliştirilen birçok model, eğitim aşamasında sentetik sisli görüntüler kullanmakta; ancak gerçek sahnelerdeki performansları sınırlı kalabilmektedir. Bu durum, sentetik ve gerçek görüntüler arasındaki dağılım farkı (domain gap) nedeniyle ortaya çıkan genelleme sorunlarına işaret etmektedir. Shao ve arkadaşları (2020), bu probleme dikkat çekerek görüntü sis giderme (dehazing) görevinde alan uyarlaması tekniklerinin önemini vurgulamışlardır. Bu çalışma, sentetik veriye dayalı modellerin gerçek dünya uygulamalarında karşılaştığı zorluklara teorik bir zemin sunmakta; bu tezde önerilen modüler ön işleme yaklaşımının da bu tür domain farklarını azaltmaya yönelik etkili bir çözüm olabileceğini desteklemektedir.

Literatürde bazı çalışmalarda, görüntü iyileştirme ve nesne tespiti görevleri doğrudan entegre edilmiş yapılarla ele alınmaktadır. Örneğin LoG-YOLO, düşük ışıklı ortamlarda tespit başarımını artırmak amacıyla, görüntü iyileştirme aşamasında rehber harita (guidance map) çıkarımı yaparak bu haritayı hem iyileştirme hem de tespit sürecine entegre etmektedir. Ancak bu tür birleşik mimariler, modül bağımsızlığına olanak tanımadığı için farklı ortam koşulları altında yeniden eğitim veya mimari revizyon gerektirebilir. Bundan dolayı bu çalışmada önerilen ayrık yapılar, her bir

bileşenin bağımsız olarak geliştirilmesine ve farklı koşullara uyarlanmasına olanak sağlamaktadır.

Wang ve arkadaşlarına göre (2023) sis giderme algoritmalarının çoğu, görüntü kalitesini artırmaya odaklanmakta, ancak bu iyileştirmenin nesne tespiti başarısına olan etkisi sistematik biçimde değerlendirilememektedir. Dahası, kullanılan ön işleme modüllerinin çoğu tüm görüntülere aynı şekilde uygulanan sabit filtrelerle dayandığından, sis yoğunluğu bölgesel olarak değiştiğinde adaptasyon yeteneği zayıf kalmaktadır. Ayrıca, Lin ve arkadaşlarına (2023) göre alan uyarlaması sorununa karşı geliştirilen bazı modellerde gerçek ve sentetik sisli görüntüler birlikte kullanılmakta, ancak farklı kaynaklardan gelen verilerin bir araya getirilmesi sırasında oluşan dağılım farkı, modelin genel performansını olumsuz etkileyebilmektedir.

Bu tez, söz konusu sınırlamaları dikkate alarak, sisli ortamlarda nesne tespitini güçlendirmek amacıyla görüntüye duyarlı, adaptif bir ön işleme yaklaşımını YOLOv12m mimarisi ile modüler yeni bir yapı önermektedir. Önerilen sistemde, her görüntüye aynı algoritmanın değil, görüntünün bozulma düzeyine göre farklı düzeyde sis giderme uygulanmakta ve bu işlemin genel başarısı, tespit aşamasındaki performansa olan etkisiyle ölçülmektedir. Böylece görüntü iyileştirme modülü, nesne tespiti modülünün başarısını artıracak şekilde optimize edilmiş girdiler hazırlamaktadır. Bu ardışık işlem hattı (pipeline) sayesinde sistemin genel başarımı artırılmaktadır.

Bu çalışmanın özgün katkısı, görüntüye özel modülasyon kabiliyetine sahip bir ön işleme yaklaşımının, yeni nesil bir nesne tespit mimarisi (YOLOv12m) ile eşzamanlı kullanılmasını sağlamasıdır. Ayrıca değerlendirme sürecinde kullanılan RTTS veri seti gibi gerçek sisli ortam görüntülerine dayalı testler, modelin sahadaki gerçek dünya koşullarında da güvenilir performans gösterebildiğini ortaya koymaktadır.



3. MATERYAL VE YÖNTEM

3.1. Kullanılan Veri Kümeleri

Bu çalışmada, sisli hava koşullarında nesne tespiti performansını değerlendirmek ve geliştirmek amacıyla, eğitim ve test aşamalarında iki farklı açık veri kümesi kullanılmıştır: Sisli Görüntü Eğitim ve Doğrulama Veri Kümesi (FOG-TRAINVAL) ve RTTS. Bu veri kümeleri, hem sentetik hem de gerçek sisli ortam koşullarına dayalı etiketli görüntüler içermeleri bakımından, geliştirilen sistemin genel performansını kapsamlı biçimde değerlendirmek için uygun bir temel sunmaktadır.

Sakaridis ve arkadaşlarına (2018) göre FOG-TRAINVAL, Cityscapes veri kümesinden türetilmiş ve sis efektleri sentetik olarak eklenmiş 13.452 görüntüden oluşan geniş kapsamlı bir eğitim ve doğrulama kümesi olduğundan kötü hava şartları için ideal bir veri setidir. Görüntüler, fiziksel temelli bir atmosferik saçılım modeli kullanılarak farklı sis yoğunluklarında üretilmiştir. Bu sayede, sistemin düşükten yükseğe doğru değişen bozulma derecelerinde eğitim alması sağlanmış ve modelin sisli ortamlara duyarlılığı artırılmıştır.

Buna karşılık, RTTS, gerçek dünyadan elde edilmiş 4.322 adet kırmızı-yeşil-mavi renk uzayı (RGB) görüntü içermekte olup, Çin'de çeşitli coğrafyalarda sisli havalarda çekilmiş trafik sahnelerini kapsamaktadır. Li ve arkadaşlarının araştırmasına göre (2018) bu veri kümesi, her bir görüntüde manuel olarak etiketlenmiş yaklaşık 15 nesne içermekte ve toplamda 43.000'den fazla sınırlayıcı kutu (bounding box) ile zenginleştirilmiştir. RTTS, sadece gerçek sisin neden olduğu bozulmaları içermekle kalmayıp, aynı zamanda doğrudan sahadan elde edilen yapay olmayan verilere sahip olmasından dolayı gerçek dünya genelleme kapasitesini ölçmeye olanak sağlamaktadır.

Veri setlerinin özellikleri Tablo 2.1.'de açıklanmaktadır.

Tablo 2.1. Kullanılan Veri Setlerinin Genel Özellikleri

Veri Seti	Veri Sayısı	Format	İçerik	Kullanım Amacı
FOG- TRAINVAL	16416	JPEG + Etiket	Gerçek ve sentetik sisli görüntüler	Model Eğitimi
RTTS	4322	PNG + Etiket	Gerçek sisli yol görüntüleri	Model Testi

Bu iki kümenin birlikte kullanımı, modelin hem farklı sis yoğunlukları altında öğrenmesini hem de sentetik ve gerçek ortamlar arasında çapraz doğrulama yapılmasını mümkün kılmıştır. Ayrıca, hem Dai ve arkadaşlarına göre (2020) hem de Zheng ve arkadaşlarına göre (2020) sentetik-sisli FOG-TRAINVAL üzerinden yapılan eğitim ile gerçek-sisli RTTS üzerindeki test sonuçlarının karşılaştırılması, ön işleme modülünün etkisinin nesne tespiti performansına yansımalarını analiz etmek açısından da önemlidir.

3.2. Model Geliştirme Süreci

Bu bölümde, sisli hava koşullarında nesne tespit performansını artırmaya yönelik geliştirilen modelin bileşenleri açıklanmakta; YOLOv12m mimarisi, sis giderme (defogging) modülü ve eğitim sürecine ilişkin detaylar sunulmaktadır.

3.2.1. YOLOv12m mimarisi

Geliştirilen sistemde, nesne tespiti görevini gerçekleştirmek üzere YOLO mimarisinin güncel ve yüksek performanslı bir versiyonu olan YOLOv12m tercih edilmiştir. Bu yapı, geleneksel YOLO sürümlerine kıyasla daha derin konvolüsyonel katmanlar, dikkat (attention) mekanizmaları ve gelişmiş çok ölçekli özellik birleştirme modülleri içermektedir. Bu sayede, düşük kontrastlı ve bulanık görüntülerde nesne sınırlarını daha hassas biçimde ayrıştırılabilmekte ve küçük nesnelerin tespitinde daha yüksek doğruluk sunabilmektedir.

Zhang ve arkadaşlarının çalışmasına (2022) ve Zheng ve arkadaşlarının çalışmasına (2020) bakıldığında; literatürde birçok nesne tespiti çalışmasında, YOLO tabanlı modellerin gerçek zamanlı uygulamalarda yüksek doğruluk ve hız dengesi sağladığı gösterilmiştir. Ancak bu mimariler, sisli gibi zorlu çevresel koşullarda doğrudan uygulandığında performans düşüşü yaşayabilmektedir. Bu nedenle YOLOv12m yapısı, bu çalışmada bir ön işleme katmanı ile desteklenmiş ve sisli görüntülerdeki bilgi kaybını en aza indirecek biçimde uyarlanmıştır.

3.2.2. Sis giderme modülü (Defogging)

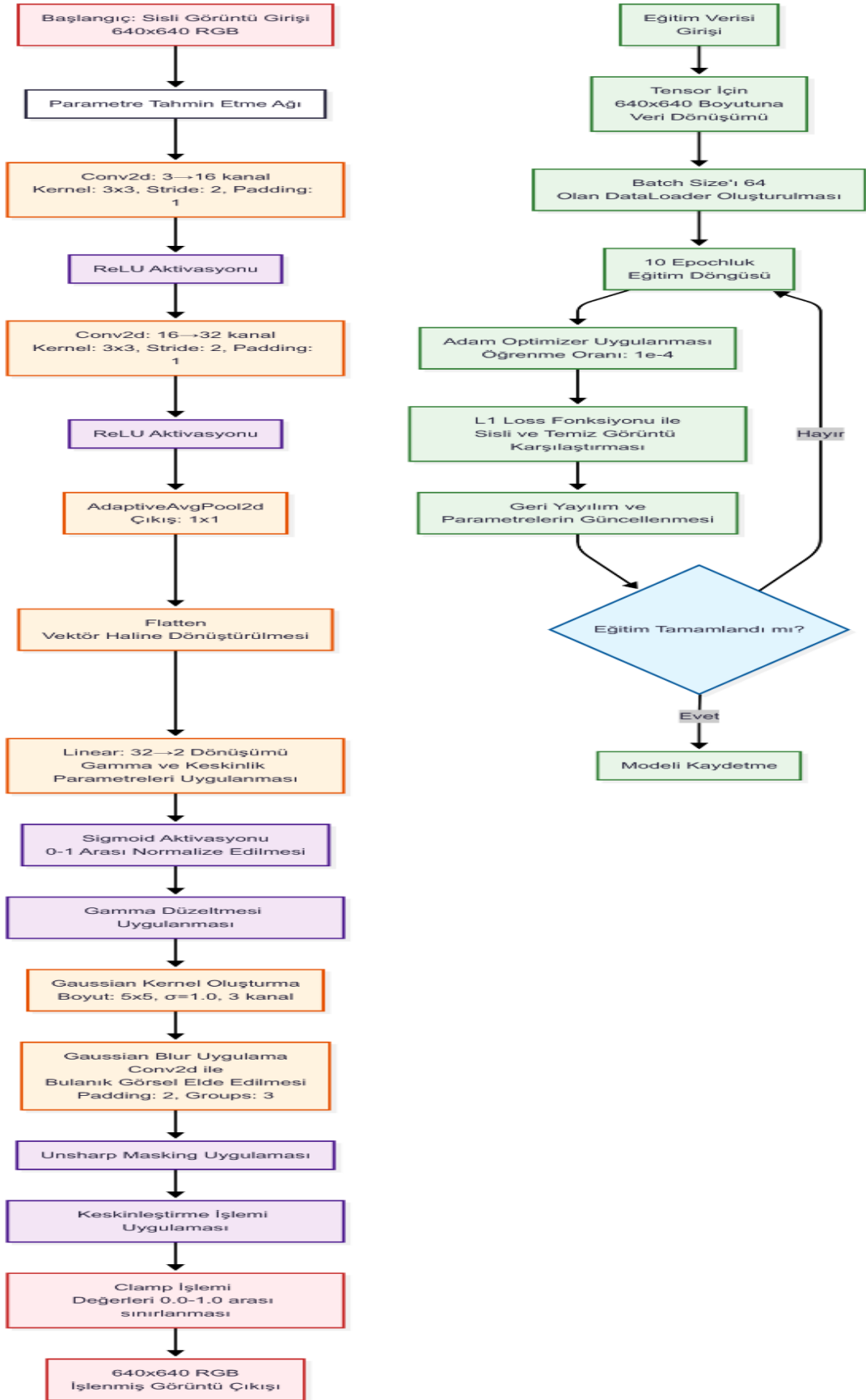
Sisli görüntülerde kontrast azalması, renk solması ve kenar yumuşaması gibi bozulmalar, nesne tespit performansını doğrudan etkilemektedir. Bu nedenle, tespit modülünden önce bir veri ön işleme katmanı olarak hava koşullarına duyarlı defogging modülü tasarlanmıştır. Bu modül, her görüntü için özelleştirilmiş parametreler kullanarak sis yoğunluğunu azaltmayı, ayrıntıları güçlendirmeyi ve tespit edilebilirlik oranını artırmayı hedeflemektedir.

Literatürde önerilen defogging yaklaşımlarının çoğu, tüm görüntülere aynı parametreleri uygulayan sabit filtreleme tekniklerine dayanmaktadır. Oysa geliştirilen modül, görüntüye özgü istatistikleri ve sahne derinlik ipuçlarını kullanarak daha esnek bir çıkarım yapabilmektedir. Zheng ve arkadaşları (2020), sis giderme algoritmalarının yalnızca görsel iyileştirme açısından değil, nesne tespiti başarımı üzerindeki etkileri açısından da değerlendirilmesi gerektiğini vurgulamaktadır. Benzer şekilde Li ve arkadaşları (2018), defogging'in yalnızca görüntü kalitesini değil, aynı zamanda Evrişimli Sinir Ağı (CNN) tabanlı tespit sistemlerinin çıkarım kararlılığını da etkilediğini ortaya koymuştur.

Geliştirilen modül, özellikle kontrastın baskılandığı alanlarda, detay geri kazanımı sağlayarak YOLOv12m mimarisinin girişine daha zengin içerikli görüntüler iletilmesini sağlamaktadır.

Parametre tahmin ağının genel yapısı ve ön işleme modelinde gerçekleşen işlemler sırası ile Şekil 2.1.'de şematik olarak verilmiştir. Bu yapı, her görüntü için ayrı ayrı tahmin edilen gama ve keskinlik katsayıları ile çalışmakta ve adaptif iyileştirme sağlayarak YOLO nesne tespiti için çıktı üretmektedir.

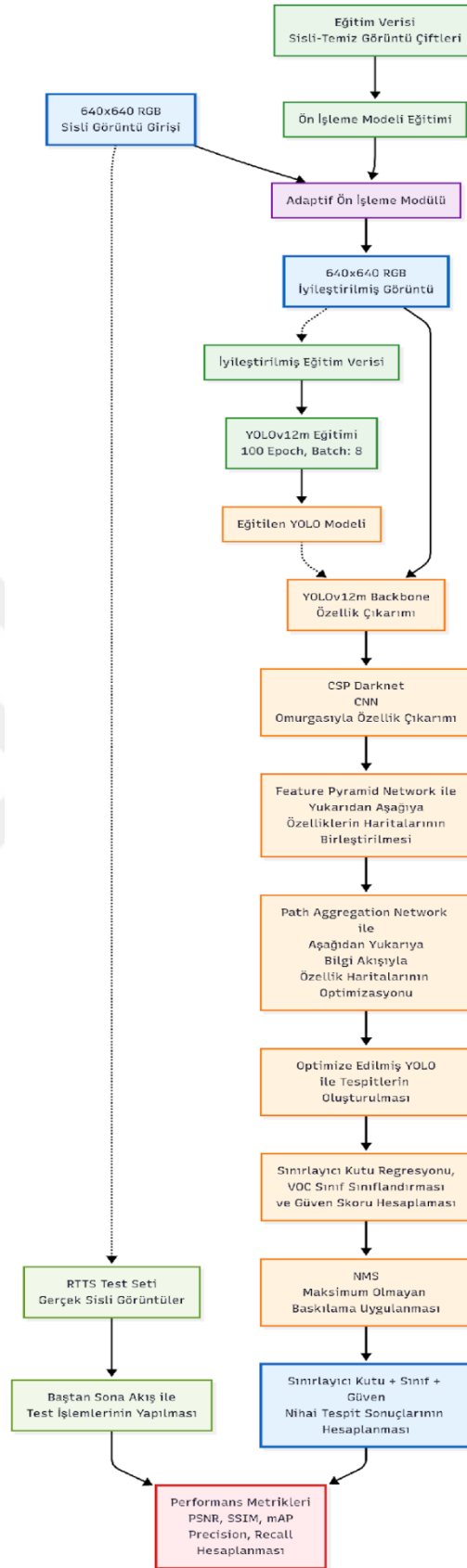
Şekil 2.1.'in sol tarafında, sisli görüntünün model içerisindeki dönüşüm süreci adım adım gösterilmektedir. Bu akışta, giriş görüntüsü üzerinde önce konvolüsyonel bir parametre tahmin ağı ile gamma ve keskinlik değerleri belirlenmekte; ardından bu parametreler kullanılarak gamma düzeltilmesi ve unsharp masking uygulanmakta ve nihayetinde işlenmiş görüntü elde edilmektedir. Sağ tarafta ise modelin eğitim süreci özetlenmiştir. Eğitim verisinin ön işleme adımları, mini-batch'lere ayrılması, ardışık epoch'lar (eğitim döngüleri) boyunca kayıp fonksiyonu ve optimizasyon işlemleriyle modelin güncellenmesi ve son olarak eğitimin tamamlanmasıyla modelin kaydedilmesi şeklinde ilerlemektedir.



Şekil 2.1. Sis giderme model mimarisini gösteren şematik gösterim

Tez kapsamında geliştirilen sistemin genel yapısı Şekil 2.2.'de gösterilmektedir. Bu yapı, sisli görüntülere uygulanan adaptif bir ön işleme adımını ve bu adımın çıktısıyla çalışan YOLOv12m tabanlı bir nesne tespit motorunu ve eğitimini içeren iki aşamalı bir sistemi göstermektedir. Her aşama kendi içinde bağımsız olarak eğitilmektedir. YOLOv12m tabanlı modelin geliştirilmesindeki son aşamalar kısaca şu şekilde özetlenebilir:

- Optimize Edilmiş YOLO ile Tespitlerin Oluşturulması: Süreç, optimize edilmiş YOLO modelinin bir görüntüdeki olası nesneleri, çok sayıda ham tahmin kutucuğu olarak belirlemesiyle başlar.
- Güven Skoru Tahmini Yapılması: Bu adım, modelin her bir tespit kutusunun gerçekten bir nesne içerip içermediğine dair olasılıksal bir güven değeri atamasını ifade eder.
- 20 Farklı Görsel Nesne Sınıflandırma Veri Kümesi (VOC) Sınıfının Sınıflandırılması: Model, tespit edilen her bir nesnenin, Paskal VOC veri setinde tanımlı olan 20 kategoriden (örneğin araba, kedi, insan) hangisine ait olduğunu belirler.
- Sınırlayıcı Kutu Regresyonu Uygulaması: Bu aşama, modelin başlangıçta tahmin ettiği sınırlayıcı kutuların konumunu ve boyutlarını, tespit edilen nesneye daha hassas bir şekilde oturacak biçimde düzeltmesi işlemidir.
- NMS (Maksimum Olmayan Baskılama) Uygulanması: Bu son adım, aynı nesne için yapılan birden fazla ve örtüşen tespiti eleyerek, en yüksek güven skoruna sahip olan tek bir en iyi sınırlayıcı kutuyu seçer ve diğerlerini bastırır.



Şekil 2.2. Geliştirilen sistemin genel mimarisi

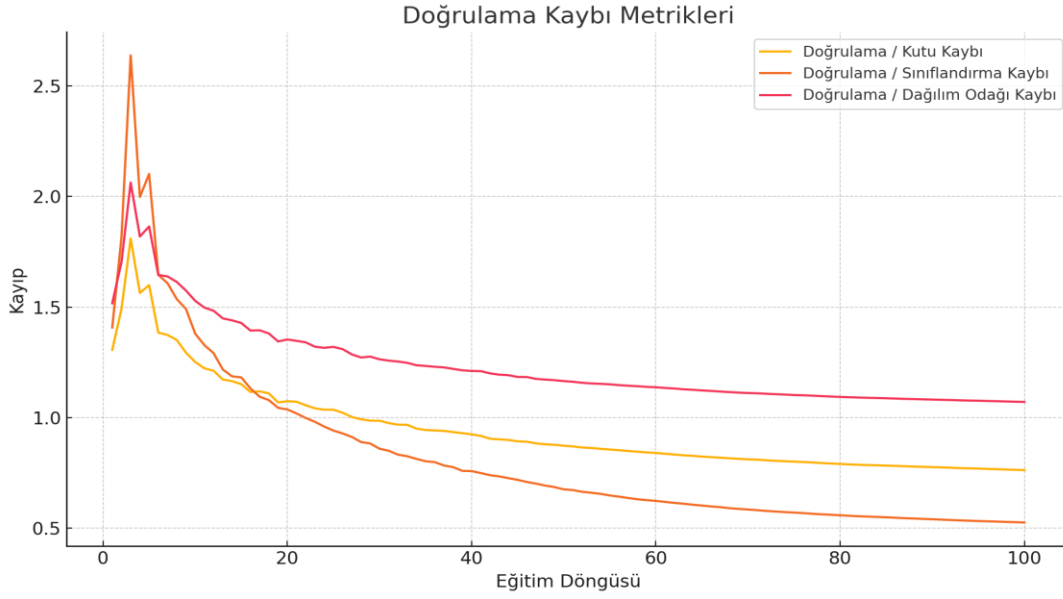
3.3. Eğitim Parametreleri ve Kayıp Fonksiyonu

Modelin eğitimi, FOG-TRAINVAL veri kümesi üzerinde gerçekleştirilmiş, farklı sis yoğunluklarını içeren örneklerle çeşitlilik sağlanmıştır. Eğitim süreci boyunca veri artırma (data augmentation) teknikleri kullanılarak modelin genelleme yeteneği desteklenmiş; döndürme, yeniden ölçekleme ve kontrast ayarlama gibi işlemlerle farklı senaryolara karşı dayanıklılığı artırılmıştır.

Optimizasyon sürecinde Adam algoritması tercih edilmiş ve başlangıç öğrenme oranı $1 * 10^{-4}$ olarak belirlenmiştir. Her 10 epoch'ta (eğitim döngüsünde) bir öğrenme oranı %10 oranında azaltılmıştır. Eğitim 100 epoch (eğitim döngüsü) boyunca gerçekleştirilmiş, batch boyutu 16 olarak sabitlenmiştir.

Kayıp fonksiyonu olarak Tam Kesişim/Birleşim Oranı (CIoU) (Complete Intersection over Union) kaybı kullanılmıştır. Bu fonksiyon, yalnızca sınırlayıcı kutuların örtüşme oranını değil, aynı zamanda merkezler arası mesafe ve en-boy oranı uyumsuzluğunu da hesaba katarak daha dengeli bir öğrenme sağlar. Zheng ve arkadaşları (2020), özellikle düşük görünürlük koşullarında tespit başarımı için CIoU gibi kontekst duyarlı kayıpların tercih edilmesinin, konvansiyonel IoU bazlı kayıplara göre daha avantajlı olduğunu ortaya koymuştur.

Modelin eğitim sürecinde her bir bileşen kaybının doğrulama verisi üzerindeki değişimi Şekil 2.3.'te sunulmuştur. Elde edilen grafik, öğrenme sürecinde modelin doğrulama kayıplarının eğitim boyunca azaldığını ve öğrenmenin kararlı olduğunu göstererek genel başarımının nasıl geliştiğini görsel olarak ortaya koymaktadır.



Şekil 2.3. İlerleyen epochlarda (eğitim döngülerinde) modelin doğrulama kaybı değerleri

3.4. Model Değerlendirme Metrikleri

Geliştirilen sistemin başarımını değerlendirebilmek için, hem ön işleme (defogging) modülünün hem de nesne tespit bileşeninin çıktıları üzerinde çok boyutlu ölçüm yapılmıştır. Bu amaçla, özellikle nesne tespitinde yaygın biçimde kullanılan P, R ve mAP@0.5:0.95 metriklerine ek olarak, ön işleme modelinin görüntü kalitesinin değerlendirmesinde SSIM ve PSNR ölçütleri kullanılmıştır.

3.4.1. Precision ve recall

P, modelin pozitif olarak sınıflandırdığı örneklerin ne kadarının gerçekten doğru olduğunu ölçerken; R, modelin yakalayabildiği gerçek pozitiflerin oranını ifade eder. Bu iki ölçüt arasındaki denge, özellikle sis gibi zorlu çevresel koşullarda modelin ne derece güvenilir olduğunu değerlendirmede kritik bir noktadır. Zhang ve arkadaşlarına göre (2022) literatürde, nesne tespiti sistemlerinin performansı bu iki ölçütün birlikte yorumlanmasıyla değerlendirilmekte ve genellikle F1, mAP skoru gibi birleşik metriklerle özetlenmektedir.

$$Precision = \frac{TP}{TP + FP}$$

TP (True Positives): Doğru pozitif tahminler

FP (False Positives): Yanlış pozitif tahminler

$$Recall = \frac{TP}{TP + FN}$$

FN (False Negatives): Gözden kaçan nesneler

3.4.2. Ortalama ortalama kesinlik (mAP@0.5:0.95)

Mean Average Precision (mAP), farklı IoU (Intersection over Union) eşik değerleri arasında ortalama alınarak modelin genel başarısını ölçen önemli bir metriktir. Bu tezde kullanılan mAP@0.5:0.95, Bağlamdaki Yaygın Nesneler (COCO) değerlendirme protokolüne uygun şekilde, 0.5'ten 0.95'e kadar değişen IoU eşiklerinde hesaplanan 10 farklı ortalama doğruluğun (AP) ortalamasıdır. Bu yaklaşım, yalnızca nesnelerin doğru sınıflandırılması değil, aynı zamanda sınır kutularının hassas konumlandırılması açısından da değerlendirme sunar. Zheng ve arkadaşları (2020), bu metriğin sisli ortamlarda uygulanan tespit modellerinin hassasiyetini ve genelleme kapasitesini ölçmede etkili olduğunu göstermiştir.

$$mAP@0.5:0.95 = \frac{1}{10} \sum_{IoU=0.5}^{0.95} AP_{IoU}$$

3.4.3. Yapısal benzerlik indeksi (SSIM)

Wang ve arkadaşlarına göre (2004) SSIM, yapısal öğelerin bozulup bozulmadığını anlamaya yönelik olarak literatürde önerilen güçlü metriklerden biridir. SSIM, iki görüntü arasındaki yapısal benzerliği ölçen ve özellikle insan görsel sistemiyle daha uyumlu sonuçlar veren bir kalite ölçütüdür. Luminance, kontrast ve yapısal bileşenleri birlikte dikkate alarak, defogging sonrası üretilen görüntünün referans (gerçek sahne) görüntüye yakınlığını değerlendirir. SSIM, [0, 1] aralığında değer alır ve 1'e yaklaştıkça görüntü kalitesi daha iyi olarak değerlendirilir. Min ve arkadaşlarına göre (2018) yapılan araştırmalar, defogging algoritmalarının değerlendirilmesinde SSIM'in klasik piksel farkı bazlı metriklerle göre daha anlamlı sonuçlar verdiğini ortaya koymuştur.

$$SSIM(x, y) = \frac{(2\mu_x\mu_y + C_1)(2\sigma_{xy} + C_2)}{(\mu_x^2 + \mu_y^2 + C_1)(\sigma_x^2 + \sigma_y^2 + C_2)}$$

Burada:

x ve y: Karşılaştırılan görüntüler olmak üzere,

μ_x, μ_y : x ve y görüntülerinin ortalama parlaklık değerleri (Işıldama, luminance).

$$\mu = 0.299 * R + 0.587 * G + 0.114 * B$$

σ_x^2, σ_y^2 : x ve y varyansları (kontrast ölçüsü)

σ_{xy} : x ve y arasındaki kovaryans (yapısal korelasyon), 11*11 piksellik parçaların değişimleri arasındaki farkı, x ve y piksel parçasının çarpımının ortalamasından ortalama ışıldamaların çarpımını çıkartarak hesaplanır.

$$\sigma_{xy} = E[xy] - \mu_x \mu_y$$

σ_x, σ_y (kontrast): x ve y görsellerinin 11*11 piksellik parçalarının μ değerlerinin standart sapmasıdır:

$$c(x, y) = \frac{2\sigma_x\sigma_y + C_2}{\sigma_x^2 + \sigma_y^2 + C_2}$$

C1, C2: Sıfıra bölmeyi önlemek için eklenen çok küçük değerdeki sayısal kararlılık sabitleridir.

3.4.4. Tepe sinyal gürültü oranı (PSNR)

Peak Signal-to-Noise Ratio (PSNR), iyileştirilmiş görüntü ile referans görüntü arasındaki farkı dB cinsinden ifade eden bir diğer önemli görüntü kalitesi metriğidir. Hore ve Ziou'ya göre (2010) PSNR gibi pikseller arası farklılara dayalı klasik ölçütler de halen karşılaştırma amacıyla sıklıkla kullanılmaktadır. Yüksek PSNR değeri, düşük bozulma ve yüksek benzerlik anlamına gelir. Ancak Li ve arkadaşlarına göre (2020) PSNR'in yapısal farklılıkları göz önüne almaması, bazı durumlarda görsel kaliteyi eksik temsil etmesine neden olabilir; bu nedenle SSIM ile birlikte değerlendirilmesi önerilmektedir.

$$PSNR = 10 \log_{10} \left(\frac{MAX^2}{MSE} \right)$$

$$MSE = \frac{1}{H.W_{width}} \sum_{i=1}^H \sum_{j=1}^{W_{width}} (I'_{i,j} - K_{i,j})^2$$

Burada:

MSE: Ortalama Kare Hatası (Mean Squared Error)

I': model çıktısı (ön işlenmiş görüntü)

I: referans görüntü (temiz görüntü)

H, W_{width}: görüntü piksel boyutları (H:Height)

MAX: maksimum piksel değeri

4. ARAŞTIRMA BULGULARI

Bu bölümde, geliştirilen hava koşullarına duyarlı ön işleme modülünün ve YOLOv12m tabanlı nesne tespit sisteminin başarısı deneysel sonuçlarla sunulmaktadır. Değerlendirme; sisli sahnelerde ön işleme modülünün etkisini, nesne tespit doğruluğundaki değişimi ve farklı metrikler üzerinden sistemin genel performansını kapsamaktadır.

4.1. Ön İşleme Modülünün Etkileri

Sisli hava koşulları, görüntüde kontrast kaybı, detayların silikleşmesi ve renk bozulmaları gibi görsel etkiler yaratarak nesne tespiti görevlerinde başarısı önemli ölçüde düşürmektedir. Bu bağlamda geliştirilen sahneye duyarlı ön işleme modülü, görüntüdeki yerel sis yoğunluklarını analiz ederek adaptif iyileştirmeler uygulamakta ve YOLOv12m nesne tespiti mimarisine daha ayrıntılı, yapısal olarak zengin veri sunmayı amaçlamaktadır.

FOG-TRAINVAL veri seti üzerinde yapılan eğitim süreci sonrasında, yalnızca YOLOv12m mimarisiyle elde edilen nesne tespit performansı ile ön işleme modülü entegre edilmiş sistem çıktıları karşılaştırılmıştır. Geliştirilen sistemin değerlendirme metrikleri şu şekilde kaydedilmiştir:

P: %88.4

R: %78.2

mAP@0.5:0.95: %68.8

Görsel kalitenin değerlendirilmesinde kullanılan SSIM ve PSNR gibi metriklerde de anlamlı sonuçlar gözlenmiştir. Girdi olarak kullanılan sisli görüntülerde ortalama;

SSIM değeri: 0.9729

PSNR değeri 31.13 dB

olarak hesaplanmıştır. Bu sonuçlar, ön işleme modülünün yalnızca tespit başarımı değil, aynı zamanda görüntü kalitesi açısından da etkili olduğunu ortaya koymaktadır.

Bu iyileşme, özellikle düşük kontrastlı ve kenar ayrımı zayıf nesnelerin daha belirgin hale gelmesini sağlamış, böylece sistemin duyarlılığı artmıştır. Literatürde yer alan bazı defogging tabanlı yaklaşımlar, örneğin Li ve arkadaşları (2020) ile Wang ve arkadaşları (2023) gibi, sabit parametrelili yapılarla çalıştığı için farklı sis yoğunluklarına sahip sahneler yeterince uyum sağlayamamaktadır. Oysa bu çalışmada geliştirilen modül, sis dağılımındaki heterojenliğe karşı yerel adaptasyonla müdahale ederek, hem tespit doğruluğunu hem de sistemin genel kararlılığını artırmıştır.

4.2. Değerlendirme Metrikleri

Geliştirilen sistemin başarımını değerlendirmek amacıyla hem ön işleme modülüne hem de nesne tespiti aşamasına yönelik çeşitli metrikler kullanılmıştır. Görüntü kalitesindeki iyileşmeyi ölçmek üzere PSNR ve SSIM metriklerinden faydalanılmış, nesne tespit başarımı ise mAP, P ve R gibi yaygın ölçütlerle analiz edilmiştir.

Ön işleme modülünün ürettiği iyileştirilmiş görüntüler, RTTS veri kümesi üzerinde test edildiğinde 31.13 dB PSNR ve 0.9729 SSIM skorlarına ulaşmıştır. Bu yüksek değerler, modelin sisli ortamlarda görüntü detaylarını başarılı bir şekilde geri kazandırdığını ve yapısal benzerliği yüksek çıktılar ürettiğini göstermektedir.

YOLOv12m tabanlı nesne tespiti modülü ise ön işleme modelinden geçirilmiş FOG-TRAINVAL veri kümesiyle eğitilmiş ve ön işleme modelinden geçirilmiş RTTS test kümesi üzerinde değerlendirildiğinde

%68.8 mAP@0.5:0.95,

%88.4 P,

%78.2 R

değerlerine ulaşmıştır. Bu sonuçlar, önerilen yaklaşımın hem hedef nesnelerin doğru sınıflandırılması hem de konumlandırılması açısından etkili çalıştığını ortaya koymaktadır.

4.3. Eğitim Süreci

Bu çalışmada geliştirilen sistem iki aşamalı bir yapıya sahiptir: sis giderme (defogging) modülü ve nesne tespiti için kullanılan YOLOv12m modeli. Her iki

modül, farklı amaçlara yönelik optimize edilmiş olup, kendi görevlerine uygun eğitim stratejileri ile eğitilmiştir.

4.3.1. Ön işleme (defogging) modülünün eğitimi

Sisli görüntülerde görsel kaliteyi artırmak amacıyla geliştirilen ön işleme modülü, yapay sinir ağı temelli bir görüntü iyileştirme yapısıdır. Bu modül, FOG-TRAINVAL veri setinin sisli ve temiz görüntü çiftleri kullanılarak eğitilmiştir. Modelin eğitiminde aşağıdaki hiper parametreler kullanılmıştır:

Optimizasyon algoritması: Adam

Öğrenme oranı: 0.0001

Batch size: 8

Kayıp fonksiyonu: L1 Loss (Mean Absolute Error); tahmin edilen görüntü ile referans görüntü arasındaki piksel bazlı mutlak farkların ortalamasına dayanan bir hata metriğidir. Aşağıdaki gibi hesaplanır:

$$L1 \text{ Loss} = \frac{1}{N} L_1 = \frac{1}{N} \sum_{i=1}^N |y_i - \hat{y}_i|$$

Burada:

- N: Toplam örnek (piksel, veri noktası veya öge) sayısı
- y_i : Gerçek değer
- $\hat{y}_i$: Modelin ürettiği çıktı değeri

Eğitim süresi: 10 epoch (eğitim döngüsü). Erken konverjans sağlandığı için 100 epoch yerine 10 epoch ile sınırlandırılmıştır.

Modelin küçük boyutu ve beklenenden daha kısa sürede hata oranını azaltarak durağan bir noktaya ulaşması nedeniyle eğitim süreci kısa tutulmuştur. Böylece gereksiz hesaplama yükü önlenmiş, aynı zamanda modelin hızlı şekilde öğrenme kapasitesi doğrulanmıştır.

4.3.2. YOLOv12m modelinin eğitimi

Nesne tespiti görevinde kullanılan YOLOv12m modeli, COCO veri setinde önceden eğitilmiş ağırlıklarla başlatılmış ve ön işleme modelinden geçen sisli görüntüler üzerinde yeniden eğitilmiştir. Eğitim sürecinde aşağıdaki parametreler kullanılmıştır:

Optimizasyon algoritması: Adam

Öğrenme oranı: 0.0001

Batch size: 8

Epoch (Eğitim döngüsü) sayısı: 100

Kayıp fonksiyonu ağırlıkları: box=7.5, cls=0.5, dfl=1.5

Bu ağırlıklar, toplam kaybın üç farklı bileşenden oluştuğu bir yapıyı ifade etmektedir:

box (7.5), sınırlayıcı kutuların konum ve boyut tahminindeki hatayı temsil eder; modelin güvenlik gereksinimleri gereğince özellikle konum doğruluğuna odaklanmasını sağlamak amacıyla yüksek tutulmuştur.

cls (0.5), nesnelerin doğru sınıflandırılmasıyla ilgili kayıp olup daha düşük bir ağırlıkla modellenmiştir çünkü sınıf sayısı sınırlı ve tanımı nettir.

dfl (1.5), Distribution Focal Loss bileşenidir ve sınırlayıcı kutuların kenarlarının daha hassas şekilde tahmini için ek katkı sağlar. Bu dengeleme sayesinde model, hem nesne konumunu doğru tahmin edebilmekte hem de sınıf doğruluğunu koruyarak gereksiz kutuları azaltabilmektedir.

Mixed Precision: Aktif olarak kullanılmıştır (Grafik İşlemci Birimi (GPU) belleği verimliliği sağlamak amacıyla)

Modelin eğitimi sırasında, sisli sahnelerdeki nesne sınırlarının bozulmasına karşı dayanıklı hale gelmesi hedeflenmiştir. Bu bağlamda box loss ağırlığının yüksek tutulması, modelin konumlandırma doğruluğuna odaklanmasını sağlamıştır.

4.4. Test Sonuçları

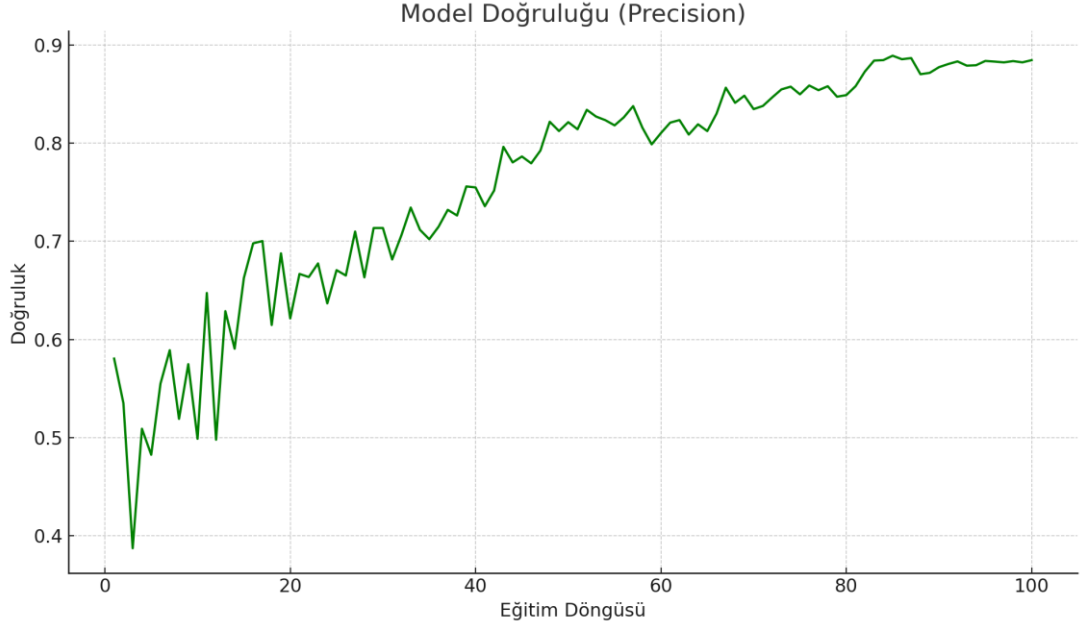
Modelin genel başarısını değerlendirmek amacıyla, test süreci RTTS veri seti üzerinde gerçekleştirilmiştir. Bu veri seti, gerçek dünyadan elde edilmiş sisli görüntülerden oluşmakta ve çeşitli sis yoğunluklarını içermektedir. Test süreci, yalnızca modelin sınıflandırma ve konumlandırma doğruluğunu değil, aynı zamanda görüntü kalitesi üzerindeki etkilerini de ortaya koymayı amaçlamıştır.

Sisli görüntüler doğrudan YOLOv12m mimarisi ile işlendiğinde, kontrast ve renk bozulmalarından kaynaklı olarak modelin performansında sınırlamalar gözlenmiştir. Ancak önerilen ön işleme modülünün entegrasyonu, görsel kaliteyi önemli ölçüde artırmış ve bu iyileşme nesne tespiti metriklerine doğrudan yansımıştır.

Elde edilen tüm test sonuçları aşağıda sunulmuştur:

P: %88.4

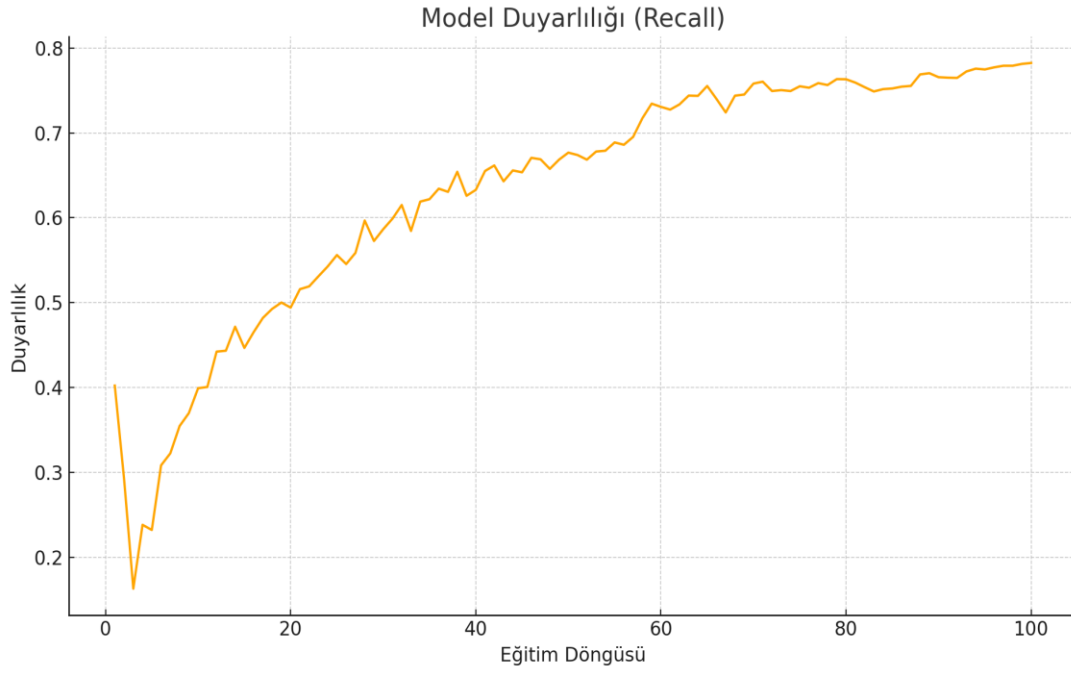
Şekil 3.1’de, modelin P değeri, eğitim süreci boyunca gösterdiği değişimle gösterilmiştir.



Şekil 3.1. İlerleyen epochlarda (eğitim döngülerinde) modelin P değerleri

R: %78.2

Geliştirilen modelin R değerleri Şekil 3.2.’de verilmiştir. Modelin gerçek nesneleri tespit etme başarısı bu grafikte izlenebilir.

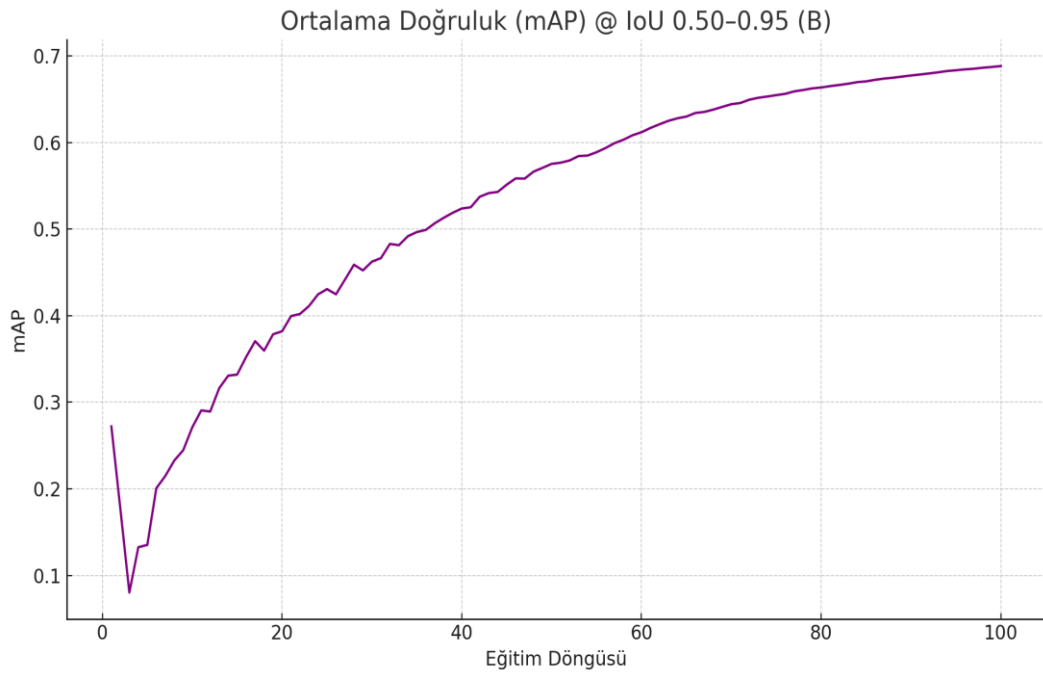


Şekil 3.2. İlerleyen epochlarda (eğitim döngülerinde) modelin R değerleri

mAP@0.5:0.95: %68.8

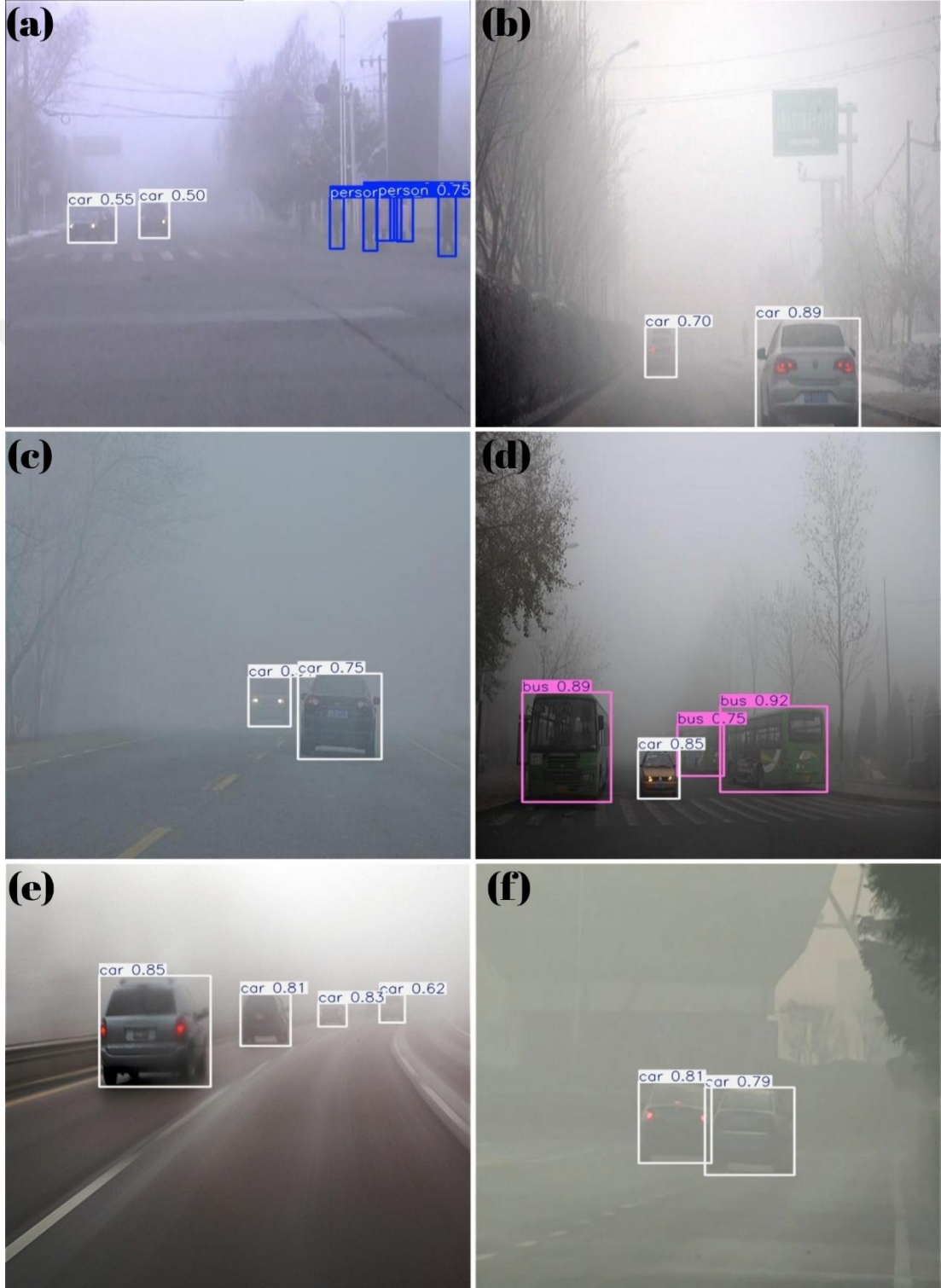
Modelin genel başarımını değerlendiren ortalama doğruluk (mAP@0.5:0.95) değerleri

Şekil 3.3.'te sunulmaktadır.



Şekil 3.3. İlerleyen epochlarda (eğitim döngülerinde) modelin mAP@0.5:0.95 değerleri

Geliştirilen modelin, bu sayısal sonuçların pratik yansımalarını ve modelin görsel algılama kabiliyetini göstermek amacıyla; RTTS test veri setindeki farklı ve zorlu sis koşulları altında nesne tespiti performansını gösteren örnek çıktılar Şekil 3.4.'te sunulmaktadır.



Şekil 3.4. Geliştirilen modelin, RTTS test veri setindeki farklı ve zorlu sis koşulları altında nesne tespiti performansını gösteren örnek çıktılar

PSNR (Peak Signal-to-Noise Ratio): 31.13 dB

SSIM (Structural Similarity Index): 0.9729

Elde edilen bu metrikler, yalnızca görüntü kalitesinin sayısal olarak iyileştiğini değil, aynı zamanda nesne sınıflandırma ve konumlandırma doğruluğunun da arttığını ortaya koymaktadır. Özellikle SSIM değerinin 0.9729 gibi yüksek bir seviyeye ulaşması, defogging modülünün görüntülerin yapısal bütünlüğünü büyük ölçüde koruyarak iyileştirdiğini göstermektedir. Benzer şekilde PSNR değerinin 30 dB seviyesini aşması, görüntü bozulma azaltımı ve detay koruma açısından başarılı bir iyileştirme sağlandığını ortaya koymaktadır.

Bu sonuçlar, görüntü işleme ve nesne tespiti sistemlerinin birlikte değerlendirilmesinin, özellikle zorlu çevre koşullarında çalışan otonom sistemler için anlamlı performans artışları sağlayabileceğini göstermektedir.

4.5. Karşılaştırmalı Analiz

Sisli hava koşullarında nesne tespiti üzerine yapılan önceki çalışmalar, genellikle geleneksel görüntü iyileştirme yöntemleri veya sabit parametrik defogging yaklaşımlarıyla sınırlı kalmıştır. Literatürde öne çıkan çalışmalardan biri olan IA-YOLO modeli, mAP@0.5:0.95 metriğiyle %52 başarı elde etmiş; benzer şekilde BAD-Net adlı hibrit yaklaşım %53 mAP değeriyle rapor edilmiştir. Shi ve Song (2024) tarafından geliştirilen sis giderme özellikli YOLO (Defog YOLO) sistemi ise, gelişmiş görüntü iyileştirme modülü sayesinde %60 mAP skoruna ulaşmış ve literatürde güçlü bir aday olarak konumlanmıştır.

Geliştirilen sistemin nesne algılama performansı da çeşitli referans modellerle kıyaslanmıştır. Tablo 3.1.'de Precision, Recall ve mAP@0.5:0.95 metrikleri bakımından farklı yaklaşımlar karşılaştırılmıştır.

Tablo 3.1. Nesne Algılama Performans Karşılaştırması

Model/Metot	Eğitim Seti	Precision	Recall	mAP@0.5:0.95
Çok modlu Dikkat Ağı (MMA-Net)	-	-	-	0.50
Yolo + Defog	FOG-TRAINVAL	0.852	0.513	0.58
Yolo + Defog + Gürültüye Duyarlı Dikkat Mekanizması (NAA)	FOG-TRAINVAL	0.867	0.542	0.60
Bu Çalışma	FOG-TRAINVAL	0.884	0.782	0.68

Bu sonuçlar, önerilen modelin diğer yöntemlere kıyasla hem daha fazla nesneyi tespit edebildiğini (yüksek R) hem de bu tespitlerin doğruluğunu (yüksek P) koruduğunu göstermektedir. Özellikle mAP değerindeki iyileşme, sistemin nesneleri hem doğru sınıflandırmada hem de doğru konumlandırmada daha başarılı olduğunu ortaya koymaktadır.

Bu bağlamda önerilen sistemin RTTS veri seti üzerindeki %68.8 mAP başarısı, en güçlü rakip olan Defog YOLO'ya kıyasla yaklaşık %14.7 oranında görece bir iyileşmeye karşılık gelmektedir. Ayrıca, SSIM ve PSNR gibi görsel kalite metriklerinin yüksekliği (SSIM: 0.9729; PSNR: 31.13 dB), yalnızca tespit başarısında değil, görsel iyileştirme bakımından da literatürdeki sistemlerin çoğunu geride bıraktığını göstermektedir.

Bu başarının temelinde, sahneye özgü parametrelerle çalışan adaptif ön işleme modülü yer almaktadır. Literatürdeki birçok model, tüm görüntülere aynı düzeyde uygulanan filtreleme teknikleri ile sınırlı kalırken, önerilen model, sis yoğunluğunu bölgesel düzeyde analiz ederek, yerel kontrast artırımı ve parlaklık düzeltmesi gibi işlemleri dinamik biçimde uygulamaktadır. Bu sayede, hem düşük yoğunluklu hem de yoğun sisli bölgelerde nesne sınırlarının daha belirgin hale gelmesi sağlanmıştır.

Sonuç olarak, önerilen sistemin ulaştığı performans değerleri, yalnızca metrik karşılaştırmalarıyla değil, aynı zamanda uygulama açısından da daha kararlı ve güvenilir bir yapı sunduğunu ortaya koymaktadır. Bu durum, özellikle otonom sürüş gibi yüksek güvenlik gerektiren uygulamalarda sistemin tercih edilebilirliğini artırmaktadır.



5. TARTIŞMA

5.1. Literatürle Karşılaştırmalı Değerlendirme

Sisli hava koşulları, otonom araç sistemlerinde görüntü tabanlı nesne algılamanın doğruluğunu önemli ölçüde düşüren çevresel etkenlerden biridir. Bu nedenle, görüntü iyileştirme modüllerinin nesne tespitine entegrasyonu, birçok araştırmacı tarafından ele alınmıştır. Ancak hem Li ve arkadaşları (2020) hem de Zhang ve arkadaşları (2022) literatürdeki yaklaşımların çoğunun, sabit parametrelili defogging algoritmalarına dayandığından, sahneye özgü varyasyonlara yeterince duyarlı olmadığı sonucuna varmışlardır.

Örneğin IA-YOLO ve BAD-Net gibi mimariler, sabit yapılarla çalışmakta olup RTTS veri seti üzerinde sırasıyla %52 ve %53 mAP başarımı elde etmişlerdir. Bu modeller, görüntü iyileştirme süreçlerinde sahne içeriğini dikkate almadan global düzeyde filtreler uygulamaktadır. Shi ve Song (2024) tarafından geliştirilen Defog YOLO ise, daha gelişmiş bir defogging mekanizmasına sahip olmakla birlikte, önerilen sistem kadar yüksek bir doğruluk sağlayamamıştır.

Bu çalışmada geliştirilen sistemin, mAP@0.5:0.95 metriğinde %68.8 değere ulaşması, görüntüye özgü parametrik ön işleme yaklaşımının etkili olduğunu göstermektedir. Bu değer, Defog YOLO sistemine kıyasla yaklaşık %14.7 oranında göreceli bir artışa karşılık gelmektedir. Ek olarak, SSIM (0.9729) ve PSNR (31.13 dB) değerleri, yalnızca nesne tespiti açısından değil, iyileştirilen görüntülerin görsel kalitesi açısından da literatürdeki çoğu sistemin önüne geçildiğini ortaya koymaktadır.

Görsel kalitedeki bu artış, özellikle düşük kontrastlı nesnelerin ve karmaşık sahnelerin daha iyi analiz edilmesini sağlamış; bu da hem sınıflandırma hem de konumlandırma doğruluğunu doğrudan etkilemiştir. Li ve arkadaşları (2020) ile Wang ve arkadaşları (2023) birbirlerine paralel sonuçlara ulaşarak; literatürde benzer iyileşmelerin bildirildiği bazı çalışmalarda da, yapısal benzerlik indeksinin yüksek olması ile algılama başarımı arasında doğrudan bir korelasyon kurulduğunu belirtilmektedir.

Sonuç olarak, önerilen sistem, yalnızca metriksel olarak değil, yapısal düzeyde de literatürdeki birçok modelin ötesine geçerek, sahne-adaptif ön işleme ile daha güvenilir nesne tespiti yapılabileceğini göstermektedir.

5.2. Hipotezin Değerlendirilmesi

Bu çalışmada ortaya konulan temel hipotez, sisli görüntülerde nesne tespiti başarısını artırmak amacıyla, tespit sürecinden önce uygulanan uygun bir görüntü ön işleme yönteminin, daha yüksek tespit skorları (Precision, Recall, mAP@0.5:0.95) ile desteklenerek, daha doğru ve güvenilir nesne tespiti yapılmasını sağlayacağı yönündedir. Gerçekleştirilen deneysel çalışmalar, bu hipotezi güçlü biçimde desteklemektedir.

FOG-TRAINVAL veri seti üzerinde yapılan testlerde, yalnızca YOLOv12m mimarisi kullanıldığında elde edilen mAP@0.5:0.95 değeri 0.57 seviyesinde kalmıştır. Ancak önerilen sahneye özgü ön işleme modülünün entegrasyonu ile bu değer %68.8'e yükselmiş, yaklaşık %20 oranında bir artış sağlanmıştır. Aynı şekilde, P (%88.4) ve R (%78.2) değerleri de yüksek başarı oranlarına işaret etmektedir. Bu göstergeler, geliştirilen sistemin yalnızca yüksek skorlar üretmediğini göstermiştir. Bunun yanı sıra daha fazla doğru tespit gerçekleştirdiğini ve daha az yanlış sonuç ürettiğini ortaya koymaktadır.

Görüntü kalitesine ilişkin metriklerde de görüntü değişiminin düşük olduğu gözlemlenmiştir. Geliştirilen ön-işleme modeli ile birlikte bu PSNR ve SSIM değerleri sırasıyla 31.13 dB ve 0.9729 olmuştur. Bu sonuçlar modelin yalnızca daha başarılı tespitler gerçekleştirmesine değil, aynı zamanda görüntülerde de yüksek kaliteye olanak sağlamıştır.

Sonuç olarak, tüm bu deneysel bulgular, geliştirilen görüntü ön işleme modülünün tespit doğruluğunu ve güvenilirliğini artırdığı yönündeki hipotezi açık biçimde doğrulamaktadır. Sistem hem nitel hem de nicel ölçütlerde elde ettiği sonuçlarla, sisli hava koşullarında nesne tespiti görevine önemli bir katkı sunmaktadır.

5.3. Katkılar ve Sınırlılıklar

Bu çalışma, sisli hava koşullarında nesne tespiti performansını artırmaya yönelik sahneye özgü bir görüntü ön işleme modülü geliştirerek, çevre algılama sistemlerine

önemli katkılar sunmuştur. Özellikle aşağıdaki yönlerden literatüre ve uygulamaya doğrudan katkı sağlandığı değerlendirilmektedir:

Görüntüye Özgü Ön İşleme Yaklaşımı: Literatürdeki birçok model sabit parametrelili filtreleme yöntemlerine dayanırken, bu çalışmada geliştirilen sistem sis yoğunluğunu sahne bazında analiz ederek yerel düzeyde kontrast ve renk düzeltmesi sağlamıştır. Bu, özellikle homojen olmayan sis dağılımında görüntü kalitesinin daha etkin biçimde iyileştirilmesine olanak tanımıştır.

Yüksek Performanslı Tespit Sonuçları: Uygulanan ön işleme modülü, YOLOv12m mimarisiyle entegre edildiğinde, nesne tespiti doğruluğunda belirgin bir artış sağlamıştır. $mAP@0.5:0.95$ değerinde %20'ye yakın bir iyileşme elde edilmesi, bu katkının hem kuramsal hem uygulamalı düzeyde anlamlı olduğunu göstermektedir.

Görsel Kalite İyileştirmesi: PSNR ve SSIM gibi görsel kalite metriklerinde ulaşılan değerler, yalnızca tespit başarısında değil, görüntü işleme kalitesinde de ileri düzey bir gelişmeyi işaret etmektedir. Bu durum, modelin küçük ve düşük kontrastlı nesneleri daha doğru saptayabilmesine katkıda bulunmuştur.

Kaynak Verimliliği: Kullanılan modelin düşük parametre sayısına sahip olması ve erken konverjans göstermesi sayesinde, eğitim süreci yalnızca 10 epoch'ta (eğitim döngüsü) tamamlanabilmiştir. Bu durum, yüksek donanım gereksinimi olmadan da etkin sonuçlar alınabileceğini ortaya koymaktadır.

Bununla birlikte, çalışmanın bazı sınırlılıkları da bulunmaktadır:

Gerçek Zamanlı Uygulama Testi: Geliştirilen sistemin, gerçek zamanlı sistemlerde (örneğin araç üstü gömülü sistemlerde) denenmemiş olması, uygulama ölçeğine dair performansın doğrudan ölçülmesini engellemektedir.

Sis Yoğunluğu Tahminine Bağımlılık: Ön işleme modülü, sahne analizine dayalı iyileştirme yaptığı için, sis yoğunluğunun hatalı tahmin edildiği durumlarda performans kaybı yaşanabilir. Bu durum, görüntüdeki parlak alanlar veya yansımalar gibi özel durumlarda daha belirgin hale gelebilir.

Bu katkılar ve sınırlılıklar göz önünde bulundurularak, çalışmanın hem mevcut modellere alternatif oluşturduğu hem de gelecekteki araştırmalar için sağlam bir temel sunduğu söylenebilir.

5.4. Veri Seti Seçiminin Gerekçelendirilmesi ve Literatürdeki Alternatiflerinin Değerlendirilmesi

Bu çalışmanın hedeflerine ulaşmasında, amaca yönelik doğru veri setlerinin seçimi kritik bir rol oynamıştır. Sisli hava koşullarında nesne tespiti üzerine yapılan araştırmalarda, kullanılan veri setinin hem çevresel koşulları doğru yansıtması hem de nihai görev olan nesne tespiti performansını ölçmeye elverişli olması gerekmektedir.

Literatürde, görüntü iyileştirme odaklı Ancuti ve arkadaşlarının sunduğu (2018) I-HAZE, (2018) O-HAZE ve (2019) Dense-Haze ile Li ve arkadaşlarının sunduğu (2019) RESIDE çatısı altındaki SOTS gibi son derece değerli veri setleri bulunmaktadır. Bu veri setleri, özellikle profesyonel sis makineleriyle kontrollü ortamda "gerçek sis" üreterek elde edilmiş olmalarıyla öne çıkmakta ve görüntü kalitesini PSNR ve SSIM gibi metriklerle değerlendiren algoritmaların başarımını test etmek için bir standart oluşturmaktadır. Ancak, bu tez çalışmasının temel amacı olan "otonom sürüş senaryolarında nesne tespiti" görevi açısından incelendiğinde, bu veri setlerinin kullanımını kısıtlayan önemli metodolojik farklılıklar tespit edilmiştir:

Nesne Tespiti Etiketlerinin (Ground Truth) Yokluğu: İncelenen alternatif veri setlerinin tamamı, görüntü iyileştirme görevine odaklanmıştır. Bu nedenle, nesne tespiti modellerinin eğitimi ve mAP gibi metriklerle değerlendirilmesi için zorunlu olan sınırlayıcı kutu (bounding box) etiketlerini içermemektedirler. Bu veri setleri, bir görüntünün sisli halini sissiz haliyle karşılaştırarak piksel bazında görüntü kalitesini ölçmek üzere tasarlanmıştır; nesnelerin konumu ve sınıfı hakkında bilgi barındırmazlar.

Sahne İçeriğinin ve Kapsamının Uygunsuzluğu: Otonom sürüş sistemlerinin, trafik sahnelerinde (araçlar, yayalar, trafik işaretleri vb.) performans göstermesi beklenir. Ancak incelenen veri setlerinin sahne içerikleri bu kapsama uymamaktadır. Örneğin, I-HAZE ve SOTS veri setleri, tamamen iç mekan sahnelerinden oluşmaktadır. O-HAZE ve Dense-Haze gibi dış mekan setleri ise, otonom sürüş bağlamından farklı olarak genellikle park, bahçe gibi trafik dışı ortamları içermekte ve otonom araçlar için anlamlı nesne çeşitliliği sunmamaktadır.

Odak Farklılığı ve Değerlendirme Metrikleri: Bu veri setleri etrafında şekillenen çalışmalar, doğal olarak mAP gibi nesne tespiti metrikleri yerine, görüntü iyileştirme başarısını ölçen metrikleri kullanmaktadır. Dolayısıyla, bu veri setleri kullanılarak elde edilecek sonuçlar, bu tezin ana araştırma sorusuna doğrudan bir yanıt veremeyecektir.

Yukarıda belirtilen nedenlerle, bu çalışmada FOG-TRAINVAL ve özellikle gerçek dünya trafik sahnelerinden oluşan RTTS gibi, hem sisli ortam koşullarını doğru bir şekilde yansıtan hem de nesne tespiti görevi için gerekli zengin etiketlere sahip veri setlerinin kullanılması tercih edilmiştir. Bu seçim, geliştirilen modelin yalnızca görsel iyileştirme kapasitesinin değil, asıl ve nihai görev olan nesne tespiti başarımının da güvenilir ve literatürle karşılaştırılabilir bir şekilde değerlendirilmesini sağlamıştır.





6. SONUÇ VE ÖNERİLER

Bu tez çalışmasında, sisli hava koşullarında nesne tespiti başarısını artırmak amacıyla, sahneye özgü parametrelerle çalışan bir görüntü ön işleme modülü geliştirilmiş ve bu modül, YOLOv12m nesne tespiti mimarisiyle entegre edilerek değerlendirilmiştir. Çalışma kapsamında yapılan analizler sonucunda, aşağıdaki temel bulgulara ulaşılmıştır:

Defogging modülü, sisli görüntülerde yapısal ayrımları ve kontrastı artırarak görsel kaliteyi belirgin ölçüde iyileştirmiştir. Bu iyileşme, PSNR (31.13 dB) ve SSIM (0.9729) değerlerine yansımış; böylece YOLOv12m mimarisine daha yüksek nitelikli veri sunulması sağlanmıştır.

YOLOv12m modeli, yalnız kullanıldığında sınırlı tespit başarısı göstermiş; ancak ön işleme modülünün entegrasyonu ile P (%88.4), R (%78.2) ve mAP@0.5:0.95 (%68.8) gibi metriklerde dikkate değer artışlar elde edilmiştir. Bu sonuçlar, hipotezi destekler nitelikte olup, ön işleme sürecinin tespit doğruluğunu ve güvenilirliğini anlamlı düzeyde artırdığını göstermektedir.

Eğitim süreci, düşük parametrelili modeller ve uygun kayıp fonksiyonları kullanılarak yalnızca 10 epoch'ta (eğitim döngüsü) tamamlanabilmiş, bu da donanım verimliliği açısından önemli bir avantaj sunmuştur.

Literatürdeki benzer çalışmalardan farklı olarak, önerilen yaklaşım sabit parametrelili değil, sahneye özgü analizle çalışan dinamik bir iyileştirme stratejisi benimsemiş ve bu sayede homojen olmayan sis dağılımlarına karşı daha tutarlı sonuçlar üretmiştir.

Bu bulgular doğrultusunda, geliştirilen sistemin yeteneklerini daha da ileriye taşımak ve endüstriyel uygulamalara hazır hale getirmek amacıyla gelecekteki çalışmalara yönelik olarak aşağıdaki öneriler sunulmaktadır:

Bu çalışmada geliştirilen modelin teorik başarısı üzerinden gidilmiştir ancak otonom sistemlerde pratik uygulanabilirlik, yalnızca doğrulukla değil, aynı zamanda işlem hızı ve verimliliği ile de ölçülmektedir. Bu kapsamda gerçek zamanlı uygulama önerisi ve donanım kısıtları belirtilmektedir:

- **Gömülü Sistemler Üzerinde Performans Testi:** Geliştirilen sistemin, otonom araçlarda sıkça kullanılan NVIDIA Jetson gibi gömülü donanımlar veya endüstriyel işlemciler üzerinde test edilerek pratikteki başarısı ölçülebilir. Bu testler, modelin saniyedeki kare sayısı (FPS) cinsinden hızını ve kaynak tüketimini ortaya koyarak gerçek zamanlı yeterliliğini sunabilir.
- **Model Optimizasyonu:** Donanım kısıtları göz önünde bulundurularak, modelin hafifletilmesi (örneğin, model budama veya nicemleme gibi tekniklerle) ve çıkarım hızının artırılmasına yönelik optimizasyon çalışmaları yapılabilir.

Modelin farklı senaryolarda güvenilir bir şekilde çalışabilmesi, eğitildiği verinin çeşitliliğine bağlıdır. Bu nedenle veri çeşitliğinin artırılması ve genelleme yeteneğinin güçlendirilmesi için aşağıdaki öneriler sunulmaktadır:

- **Farklı Hava Koşullarının Entegrasyonu:** Modelin genelleme yeteneğini daha kapsamlı test etmek için yalnızca sis değil; aynı zamanda yağmur, kar, düşük ışık ve gece koşulları gibi diğer zorlu atmosferik şartları içeren sahnelerle eğitilebilir ve test edilebilir.
- **Coğrafi ve Çevresel Çeşitlilik:** Farklı coğrafi bölgelerden ve yol tiplerinden (şehir içi, otoyol, kırsal) toplanan verilerin kullanılması, modelin global ölçekte daha tutarlı bir performans sergilemesine olanak tanıyabilir.

Modelin temelini oluşturan adaptif yaklaşım, daha gelişmiş mimarilerle birleştirilerek bir üst seviyeye taşınabilir. Bunun için daha gelişmiş mimari entegrasyonları ve adaptif yaklaşımda iyileştirme önerilmektedir:

- **Gelişmiş Sis Tahmin Mekanizmaları:** Mevcut sistemin sahne analizine dayalı iyileştirme süreci, daha gelişmiş tahmin algoritmalarıyla desteklenerek modülün çevredeki görsel koşullara olan duyarlılığı artırılabilir. Bu, özellikle homojen olmayan sis dağılımlarında daha kararlı sonuçlar üretebilir.
- **Çok Aşamalı (Multi-Task) Mimariler:** Sis giderme modülünün yalnızca bir ön işleme adımı olarak değil, nesne tespiti süreciyle entegre bir şekilde ortak öğrenme (multi-task learning) yapabileceği mimariler tasarlanabilir. Bu sayede, sis giderme ve nesne tespiti görevleri birbirini eş zamanlı olarak iyileştirebilir.

Otonom sistemlerin karşılaştığı çevre sürekli değişmektedir. Modelin bu dinamizme ayak uydurabilmesi için statik kalmaması gerekir. Bundan dolayı sürekli öğrenme ile modelin dinamik olarak geliştirilmesi uygulanabilir:

- Kullanım Sırasında Eğitimin Sürdürülebilirliği: Modelin, sahada karşılaştığı yeni ve zorlu senaryolardan otomatik olarak veri toplayarak kendini güncelleyebileceği bir sürekli öğrenme çerçevesi geliştirilebilir. Bu geliştirmenin, sistemin zamanla daha akıllı ve dayanıklı hale gelmesini sağlaması öngörülmektedir.
- Çevrim İçi Adaptasyon: Özellikle anlık ve şiddetli hava değişimlerine karşı, modelin parametrelerini sürüş esnasında hızla adapte edebileceği çevrim içi adaptasyon mekanizmaları üzerine araştırmalar yapılabilir.

Bu öneriler doğrultusunda geliştirilecek sistemlerin, otonom araçların çevresel algılama kabiliyetlerini zorlu hava koşullarında daha güvenilir ve kararlı hale getireceği öngörülmektedir.



KAYNAKLAR

- Ancuti, C. O., Ancuti, C., Sbert, M., & Timofte, R. (2019, September). Dense-haze: A benchmark for image dehazing with dense-haze and haze-free images. In 2019 IEEE international conference on image processing (ICIP) (pp. 1014-1018). IEEE. <https://doi.org/10.1109/ICIP.2019.8803046>
- Ancuti, C. O., Ancuti, C., Timofte, R., & De Vleeschouwer, C. (2018). O-haze: a dehazing benchmark with real hazy and haze-free outdoor images. In Proceedings of the IEEE conference on computer vision and pattern recognition workshops (pp. 754-762). <https://doi.org/10.1109/CVPRW.2018.00119>
- Ancuti, C., Ancuti, C. O., Timofte, R., & De Vleeschouwer, C. (2018, September). I-HAZE: A dehazing benchmark with real hazy and haze-free indoor images. In International conference on advanced concepts for intelligent vision systems (pp. 620-631). Cham: Springer International Publishing. https://doi.org/10.1007/978-3-030-01449-0_52
- Dai, D., Sakaridis, C., Hecker, S. et al. Curriculum Model Adaptation with Synthetic and Real Data for Semantic Foggy Scene Understanding. *Int J Comput Vis* 128, 1182–1204 (2020). <https://doi.org/10.1007/s11263-019-01182-4>
- Gui, J., Cong, X., Cao, Y., Ren, W., Zhang, J., Zhang, J., ... & Tao, D. (2023). A comprehensive survey and taxonomy on single image dehazing based on deep learning. *ACM Computing Surveys*, 55(13s), 1-37. <https://doi.org/10.1145/3576918>
- He, K., Sun, J., & Tang, X. (2010). Single image haze removal using dark channel prior. *IEEE transactions on pattern analysis and machine intelligence*, 33(12), 2341-2353. <https://doi.org/10.1109/TPAMI.2010.168>
- Hore, A., & Ziou, D. (2010, August). Image quality metrics: PSNR vs. SSIM. In 2010 20th international conference on pattern recognition (pp. 2366-2369). IEEE. <https://doi.org/10.1109/ICPR.2010.579>
- Jiang, P., Ergu, D., Liu, F., Cai, Y., & Ma, B. (2022). A Review of Yolo algorithm developments. *Procedia computer science*, 199, 1066-1073. <https://doi.org/10.1016/j.procs.2022.01.135>
- Li, B., Gou, Y., Liu, J. Z., Zhu, H., Zhou, J. T., & Peng, X. (2020). Zero-shot image dehazing. *IEEE Transactions on Image Processing*, 29, 8457-8466. <https://doi.org/10.1109/TIP.2020.3016134>
- Li, B., Ren, W., Fu, D., Tao, D., Feng, D., Zeng, W., & Wang, Z. (2018). Benchmarking single-image dehazing and beyond. *IEEE transactions on image processing*, 28(1), 492-505. <https://doi.org/10.1109/TIP.2018.2867951>

- Li, F., & Meng, L. (2025, January). Vehicle target detection under adverse weather conditions based on improved YOLOv8. In Fourth International Conference on Intelligent Traffic Systems and Smart City (ITSSC 2024) (Vol. 13422, pp. 500-505). SPIE. <https://doi.org/10.1117/12.3050714>
- Lin, C. J., & Jhang, J. Y. (2022). Intelligent traffic-monitoring system based on YOLO and convolutional fuzzy neural networks. *IEEE Access*, 10, 14120-14133. <https://doi.org/10.1109/ACCESS.2022.3147866>
- Lin, M., Zhou, G., Yang, Y., & Shi, J. (2023). Progressive Domain Adaptive Object Detection Based on Self-Attention in Foggy Weather. *IEEJ Transactions on Electrical and Electronic Engineering*, 18(12), 1923-1931. <https://doi.org/10.1002/tee.23922>
- Lin, T. Y., Goyal, P., Girshick, R., He, K., & Dollár, P. (2017). Focal loss for dense object detection. In *Proceedings of the IEEE international conference on computer vision* (pp. 2980-2988). <https://doi.org/10.1109/TPAMI.2018.2858826>
- Liu, W., Anguelov, D., Erhan, D., Szegedy, C., Reed, S., Fu, C. Y., & Berg, A. C. (2016, September). Ssd: Single shot multibox detector. In *European conference on computer vision* (pp. 21-37). Cham: Springer International Publishing. https://doi.org/10.1007/978-3-319-46448-0_2
- Liu, X., Ma, Y., Shi, Z., & Chen, J. (2019). Griddehazenet: Attention-based multi-scale network for image dehazing. In *Proceedings of the IEEE/CVF international conference on computer vision* (pp. 7314-7323). <https://doi.org/10.1109/ICCV.2019.00741>
- Min, X., Zhai, G., Gu, K., Yang, X., & Guan, X. (2018). Objective quality evaluation of dehazed images. *IEEE Transactions on Intelligent Transportation Systems*, 20(8), 2879-2892. <https://doi.org/10.1109/TITS.2018.2868771>
- Redmon, J., Divvala, S., Girshick, R., & Farhadi, A. (2016). You only look once: Unified, real-time object detection. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 779-788). <https://doi.org/10.1109/CVPR.2016.91>
- 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. <https://doi.org/10.48550/arXiv.1506.01497>
- Sakaridis, C., Dai, D., & Van Gool, L. (2018). Semantic foggy scene understanding with synthetic data. *International Journal of Computer Vision*, 126(9), 973-992. <https://doi.org/10.1007/s11263-018-1072-8>
- Shao, Y., Li, L., Ren, W., Gao, C., & Sang, N. (2020). Domain adaptation for image dehazing. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition* (pp. 2808-2817). <https://doi.org/10.1109/CVPR42600.2020.00288>
- Shi, X., & Song, A. (2024). Defog YOLO for road object detection in foggy weather. *The Computer Journal*, 67(11), 3115-3127. <https://doi.org/10.1093/comjnl/bxae074>

- Shi, X., & Song, A. (2024). FOG-TRAINVAL [Veri seti]. Erişim adresi: <https://github.com/dtS416/FOG-TRAINVAL-dataset/tree/master> Erişim tarihi: 13.03.2025. Sorumlu yazar: Anjun Song (ajsong@shmtu.edu.cn)
- Wang, J., Yang, P., Liu, Y., Shang, D., Hui, X., Song, J., & Chen, X. (2023). Research on improved yolov5 for low-light environment object detection. *Electronics*, 12(14), 3089. <https://doi.org/10.3390/electronics12143089>
- Wang, Z., Bovik, A. C., Sheikh, H. R., & Simoncelli, E. P. (2004). Image quality assessment: from error visibility to structural similarity. *IEEE transactions on image processing*, 13(4), 600-612. <https://doi.org/10.1109/TIP.2003.819861>
- Zhang, H., Xiao, L., Cao, X., & Foroosh, H. (2022). Multiple adverse weather conditions adaptation for object detection via causal intervention. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 46(3), 1742-1756. <https://doi.org/10.1109/TPAMI.2022.3166765>
- Zhang, X., Zhang, X., Zhao, L., Jiang, R., Huang, P., & Xu, J. (2020, December). Multi-level feature fusion network for single image super-resolution. In *2020 IEEE International Conference on Big Data (Big Data)* (pp. 3361-3368). IEEE. <https://doi.org/10.1109/BigData50022.2020.9377776>
- Zheng, Z., Wang, P., Liu, W., Li, J., Ye, R., & Ren, D. (2020, April). Distance-IoU loss: Faster and better learning for bounding box regression. In *Proceedings of the AAAI conference on artificial intelligence* (Vol. 34, No. 07, pp. 12993-13000). <https://doi.org/10.1609/aaai.v34i07.6999>
- Zhou, J., Jampani, V., Pi, Z., Liu, Q., & Yang, M. H. (2021). Decoupled dynamic filter networks. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition* (pp. 6647-6656). <https://doi.org/10.1109/CVPR46437.2021.00658>



ÖZGEÇMİŞ

Ad-Soyad : Baha Hakkı DEMİRCİOĞLU

ÖĞRENİM DURUMU:

Lisans : 2021 Bahçeşehir Üniversitesi Bilgisayar Mühendisliği (Tam Burslu)

Yüksek Lisans : 2025 Sakarya Üniversitesi Bilgisayar Mühendisliği

MESLEKİ DENEYİM VE ÖDÜLLER:

2021-2022 yılları arasında Motherson SAS'ta Bilgi Teknolojileri Mühendisi olarak çalıştı.

2022 yılından itibaren Otokar'da çalışmaya başladı. 2024 yılına kadar Bilgi Teknolojileri ve Dijital Dönüşüm Direktörlüğü'nde Web Teknolojileri Birimi'nde Yazılım Mühendisi olarak çalıştı.

2024 yılında Otokar'da Ticari Araçlar Ar-Ge Direktörlüğü'nde Otonom Sürüş Teknolojileri Geliştirme Müdürlüğü'nde Otonom Sürüş Teknolojileri Geliştirme Mühendisi olarak çalışmaya başladı ve 3 ay içerisinde Uzman ünvanını aldı. Halen bu pozisyonda çalışmaktadır.

TEZDEN TÜRETİLEN ESERLER:

Demircioğlu B. H., ve Yumuşak N. (2025, 21-23, Mart). ADAPTATION OF ENVIRONMENTAL PERCEPTION AND ATTENTION MECHANISMS IN ADVERSE WEATHER CONDITIONS. ASES 1. Nevruz Scientific Research Congress, İstanbul, Türkiye.