

**T.C.
YILDIZ TEKNİK ÜNİVERSİTESİ
FEN BİLİMLERİ ENSTİTÜSÜ**

İNSAN TESPİTİNDE YAPAY VERİ ARTIRIMI

BEDİR YILMAZ

**YÜKSEK LİSANS TEZİ
BİLGİSAYAR MÜHENDİSLİĞİ ANABİLİM DALI
BİLGİSAYAR MÜHENDİSLİĞİ PROGRAMI**

**DANIŞMAN
YRD. DOÇ. DR. MEHMET FATİH AMASYALI**

İSTANBUL, 2016

T.C.
YILDIZ TEKNİK ÜNİVERSİTESİ
FEN BİLİMLERİ ENSTİTÜSÜ

İNSAN TESPİTİNDE YAPAY VERİ ARTIRIMI

Bedir Yılmaz tarafından hazırlanan tez çalışması 15.04.2016 tarihinde aşağıdaki jüri tarafından Yıldız Teknik Üniversitesi Fen Bilimleri Enstitüsü Bilgisayar Mühendisliği Anabilim Dalı'nda **YÜKSEK LİSANS TEZİ** olarak kabul edilmiştir.

Tez Danışmanı

Yrd. Doç. Dr. Mehmet Fatih Amasyalı
Yıldız Teknik Üniversitesi

Jüri Üyeleri

Yrd. Doç. Dr. Mehmet Fatih Amasyalı
Yıldız Teknik Üniversitesi

Jüri Üyeleri

Doç. Dr. Sırma Yavuz
Yıldız Teknik Üniversitesi

Jüri Üyeleri

Yrd. Doç. Dr. Selçuk Sevgen
İstanbul Üniversitesi



Bu alıřma, Trkiye Bilimsel ve Teknolojik Arařtırma Kurumu TBİTAK (EEEAG-113E212) tarafından desteklenmiřtir.

ÖNSÖZ

Yüksek Lisans tezimde birlikte çalışma onuruna eriştiğim, bana çalışmalarım süresince her zaman destek olarak cesaretlendiren, bilgi ve tecrübelerini benimle paylaşarak olaylara bilimsel açıdan bakma pratiğini bana öğreten değerli hocam ve tez danışmanım Yrd. Doç. Dr. Mehmet Fatih AMASYALI'ya en içten teşekkürlerimle ve saygılarımla.

Tez çalışmalarım süresince benden desteklerini esirgemeyen aileme ve görev yaptığım Iğdır Üniversitesi'ne teşekkür ederim. Çalışmalarım süresince bana olan inançlarını yitirmeyen Y.T.Ü. Olasılıksal Robotik grubu PARS ın değerli üyelerine şükranlarımı sunarım.

Yüksek Lisans öğrenim hayatımı geçirdiğim Yıldız Teknik Üniversitesi'ne, Y.T.Ü. Bilgisayar Mühendisliği Bölümü'ne, Y.T.Ü. Bilgisayar Mühendisliği Bölümü'nün emektar ve değerli akademisyenlerine ayrıca teşekkürü bir borç bilirim.

Nisan, 2016

Bedir YILMAZ

İÇİNDEKİLER

	Sayfa
SİMGE LİSTESİ.....	viii
KISALTMA LİSTESİ.....	ix
ŞEKİL LİSTESİ.....	x
ÇİZELGE LİSTESİ	xi
ÖZET	xii
ABSTRACT	xiv
BÖLÜM 1	
GİRİŞ	1
1.1 Literatür Özeti.....	1
1.1.1 Bilgisayarla Görmenin Yakın Tarihi	2
1.1.2 Arama Kurtarma Çalışmaları ve İnsan Afettede Tespiti	5
1.1.3 Literatürde Yapay Veri Artırımı	6
1.1.3.1 Pomerlau 1989.....	6
Sürüş esnasında alınan (On-The-Fly) veri ile eğitim.....	6
Potansiyel Problemler	7
Çözüm: Kamera Resimlerini Dönüştürmek	8
Döndürülmüş ve Ötelenmiş Resimler	8
1.1.3.2 Drucker ve diğerleri 1994	9
1.1.3.3 Sung ve Poggio 1998	11
1.1.3.4 Sim 2000	11
1.1.3.5 Ke 2004	12
1.1.3.6 Tsang 2005.....	13

1.1.3.7	Laptev 2007	13
1.1.3.8	Howard 2013	14
	Kırpma işlemi	15
	Ek Renk Manipulasyonları	16
1.1.3.9	Cireşan ve diğerleri 2011	16
1.1.3.10	Simard 2003	17
	Veri Kümelerinin Elastik Bozunmalar Marifetiyle Genişletilmesi	17
1.1.3.11	Yaeger 1996	18
1.1.3.12	Khrizevsky ve Hinton 2012	19
1.1.4	Literatürden Edinilen Dersler	20
1.2	Tezin Amacı	20
1.3	Hipotez	21
BÖLÜM 2		
BAZI ÖNEMLİ ÖRÜNTÜ TANIMA SİSTEMLERİ		
2.1	Odaklı Gradyanların Histogramları	22
2.2	Artırımlı Basit Özellik Zincirleriyle Hızlı Nesne Tespiti	23
2.3	Değişebilen Parçalı Modeller	24
BÖLÜM 3		
DPM'in AFETZEDE TESPİTİNDE KULLANIMI ve ROS		
3.1	Robot İşletim Sistemi (Robot Operating System)	28
3.2	ROS un İç Yapısı ve Çalışma Dinamiği	29
	3.2.1 ROS Kurulumu, Paketler ve Derlenişleri	29
	3.2.2 ROS Düğümleri ve Çalışma Zamanı İşlemleri	29
3.3	DPM'in Gerçek Zamanlı Kullanımı Sorunu ve Tercih Edilen Çözüm	30
	3.3.1 Ana Hatlarıyla FFLD	31
	3.3.2 O penCV ve FFLD	32
3.4	FFLD nin ROS a Adaptasyonu	33
	3.4.1 cv_bridge Düğümü ve Sistemdeki İşlevi	33
	3.4.2 Tespit Edici Düğüm "watcher"	34
BÖLÜM 4		
DPM ALGORİTMASININ EĞİTİM SÜRECİNİN İYİLEŞTİRİLMESİNDE YAPAY VERİ ARTIRIMI		
..... 35		
4.1	Veri Kümelerinin Yapay Yollarla Genişletilmesi	35
4.2	Yapay Veri Artırımının Ardında Yatan Nedenler	35

4.3	Veri Artırımında Kullanılan Yöntemler	36	
4.3.1	Temel Geometrik Transformasyonlar	37	
4.3.1.1	Döndürme.....	38	
4.3.1.2	Çevirme.....	38	
4.3.1.3	Ölçekleme	39	
4.3.1.4	Öteleme	39	
4.3.2	Renk Bazlı Transformasyonlar	39	
4.3.2.1	Gauss Gürültüsü Eklenmesi	40	
4.3.2.2	Aydınlık Değerine Müdahale	40	
4.3.2.3	Karşıtlık Değerine Müdahale	41	
4.3.3	Etiket Üzerinden Uygulanan Transformasyonlar.....	40	
4.3.4	Parametrik Bileşke Transformasyonlar	41	
BÖLÜM 5			
VERİ KÜMESİ VE DENEYLER			42
5.1	INRIA İnsan Veri Kümesi.....	42	
5.1.1	DPM’de INRIA	44	
5.2	Deneyler	44	
BÖLÜM 6			
SONUÇ			50
KAYNAKLAR.....			51

SİMGE LİSTESİ

w	width (genişlik)
h	height (yükseklik)
F	filter (filtre)
p	point (piksel bazlı konum)
φ	özellik vektörü
s	skor
λ	pencere

KISALTMA LİSTESİ

AP	Average Precision
API	Application Programming Interface
CPU	Central Processing Unit
CV	Computer Vision
DPM	Değişebilen Parçalı Model
FFLD	Fast Fourier Linear Detector
HoG	Histograms of Oriented Gradients
LSVM	Latent SVM
RGB	Red Green Blue
ROS	Robot Operating System
SIFT	Scale-invariant feature transform
SVM	Support Vector Machine

ŞEKİL LİSTESİ

Sayfa

Şekil 1.1 Yapay sinir ağının sürüş esnasında kamera ve araç aksamından edinilen veri ile eğitilmesi [2]	7
Şekil 1.2 Eğitim süreci için ihtiyaç duyulan örneklerin üretilebilmesi amacıyla kaydırılmış ve döndürülmüş resimler [2]	9
Şekil 1.3 Yapay sinir ağının başarımını artırabilmek için üretilmiş yapay veriler ve orijinal örnek [25].....	11
Şekil 1.4 Pozitif etiketlere ve resim boyutlarına yapılan müdahalelerin sonuçları. Belirlenen bir bölümleri seçilerek kırılan resimler ve bu bölüm kaydırılarak üretilen ek resimler [31].....	14
Şekil 1.5 Yatay ekseninde kısaltılmış bir pozitif örnek. Hedef nesne olan kedi şekline ait kuyruk ve kulak gibi öğelerin kırılma sonucu kaybolduğu görülüyor [32].	16
Şekil 1.6 MNIST veri kümesindeki örneklerle benzer işlemlere sahip verilerin üretimi amacıyla yapılan deneylerin sonuçları [34]	18
Şekil 1.7 Orijinal el yazısı resimleri üzerinde küçük değişiklikler yapmak suretiyle stilistik birtakım varyasyonların yapay olarak taklit edilme çabalarının sonuçları [35].....	18
Şekil 3.1 cv_bridge Arayüz Şeması [38]	34
Şekil 4.1 Pozitif örneklerle uygulanan döndürmenin etkileri [44]	38
Şekil 4.2 Pozitif örneklerle uygulanan çevirmenin etkileri	39
Şekil 4.3 Pozitif örneklerle uygulanan ölçekleme işleminin sonuçları	39
Şekil 4.4 Pozitif örneklerle uygulanan öteleme işleminin sonuçları	40
Şekil 4.5 Resme Gauss gürültüsü eklenmesinin sonuçları.....	40
Şekil 4.6 Resmin aydınlık (Brightness) değerinin değiştirilmesinin resim üzerindeki sonuçları.....	41
Şekil 4.7 Resmin karşıtlık (Kontrast) değerinin değiştirilmesinin sonuçları.....	41
Şekil 4.8 Etiketlerinin konumlarının değiştirilmesinin resimlere etkisi	42
Şekil 5.1 INRIA kümesine ait bir resmin, etiketi üzerinde görselleştirilmiş hali [44]	44
Şekil 5.2 Şekil 15 teki resme ait olan ve PASCAL VOC Challenge formatına uygun olarak düzenlenmiş bir etiket.	44

Şekil 5.3 Resimlere uygulanan çeşitli transformasyon yöntemlerinin sonuçları ve tabloda karşılık geldikleri kodlar. En solda O) orjinal resim olmak üzere sağa doğru sırayla A) dikey simetriği, B) orjinalin sağa doğru 5 derece döndürülmüş hali, E) simetriğinin sola 5 derece döndürülmüş hali ve sonra simetriğin sağa 5 derece ve orjinalin sola 5 derece döndürülmüş hali [46]	48
Şekil 5.4 Her biri 16 adet DPM eğitim sürecini kapsayan deneyler sonucu ulaşılmış test sürelerinin, eğitim sürelerinin ve ortalama doğruluk değerlerinin ortalamaları (düğme simgesi) ve ağırlıklı değerleri (mavi bölgeler)	49



ÇİZELGE LİSTESİ

	Sayfa
Çizelge3.1 YTU Olasılıksal Robotik Laboratuvarı Robocup Verileri Üzerinde Yapılan Testlerin Sonuçları.....	33
Çizelge5.1 Farklı Yapay Veri ArtırımıTekniklerinin DPM Eğitime Etkileri.....	49
Çizelge5.2 INRIA Yaya Veri Kümesinin alt kümeleriyle yapılan deneylerin sonuçları...	47

İNSAN TESPİTİNDE YAPAYVERİ ARTIRIMI

Bedir YILMAZ

Bilgisayar Mühendisliği Anabilim Dalı

Yüksek Lisans Tezi

Tez Danışmanı: Yrd. Doç. Dr. M. Fatih AMASYALI

Arama kurtarma operasyonlarında robotik otonominin sağlanması sürecinin en önemli ayaklarından biri de görme yeteneğinin robotlara kazandırılmasıdır. Çünkü resimlerin anlamlandırılması, robotların çevrelerinden edindikleri görsel verileri, görevleri dahilinde ortaya çıkarmaları gereken kararların üretiminde kullanabilmeleri açısından kilit bir rol oynamaktadır. Görev konularına ve tasarım şekillerine bağlı olarak robotlara, resimleri sınıflandırma, resimleri bölümlerine ayırma ve resimlerde olay tespiti gibi maharetler, muhtelif bilgisayarla görme algoritmaları marifetiyle kazandırılmaktadır. Kurban tanıma amacıyla tasarlanmış bir robotun en temel görevi, sensörleri aracılığıyla edindiği verileri işleyerek afet alanlarındaki insanları tespit etmektir. Böyle bir görevi yerine getirmede kullanılacak görüntü işleme algoritması seçilirken, kısıtlı iletişim esnasında sorunsuz çalışabilecek, gerçek zamanda karar üretmeye elverişli olacak ve düşük yanlış tespit oranlarına sahip bir sisteme ihtiyaç duyulacağı unutulmamalıdır. Tüm bu gereksinimlerin yanında, sistemin başarımını makul seviyede tutabilmek adına, insan vücudunun yüksek eklemlili yapısının bir getirisi

olarak alabileceđi birok farklı Őeklin bulunduđu gz nne alınmalıdır. Sınıf ii varyasyonun bu denli yksek olduđu bir nesne sınıfına ait rneklerin tespit edilmesi iin seilecek algoritmalar arasında, bahsedilen diđer zellikleri de sađlayan yntemlerin en yksek baŐarımına sahip olanlarının baŐında DeđiŐebilen Para Modelleri gelmektedir [18]. Bu tez alıŐmaları dahilinde, bu yntem bir otonom arama kurtarma robotu sistemine adapte edilmiŐ ve bu modellerin daha yksek baŐarımına ulaŐmasına olanak sađlayacak veri artırım yntemleri DeđiŐebilen Paralı Modellerin eđitiminde baŐarıyla uygulanmıŐtır.

Anahtar Kelimeler: DeđiŐebilen Paralı Modeller, Bilgisayarla Grme, Afettede Tespiti, Kurban Tanıma, Arama Kurtarma, Veri Kmelerinin Yapay Yntemlerle GeniŐletilmesi



ARTIFICIAL DATASET AUGMENTATION IN HUMAN DETECTION

Bedir YILMAZ

Department of Computer Engineering

Msc. Thesis

Adviser: Assoc. Prof. Dr. M. Fatih AMASYALI

One of the most important processes in acquiring robotic autonomy in search and rescue operations is to bring the “ability of vision” to robots. Because, “image understanding” plays a critical role for using the image data in order to make decisions as a part of their ongoing missions. Depending on the tasks that have been given to the robots, skills like image classification, image segmentation and action classification from images etc. are provided to the robots via different computer vision algorithms. The main task of a robot that has been designed with the goal of victim detection is to determine the presence of a human being by using sensory information that has been gathered from its surroundings. When selecting the image processing algorithm that fits best for this mission it must be kept in mind that a system that can operate smoothly in a limited communication environment and can produce decisions on real time and has low false positive rates. Alongside all of these requirements, it should be taken in to consideration that human body can get in a high variety of shapes as a

result of its highly articulated nature. Between the algorithms that are candidates for detecting such examples that belong to a kind of object class that has this much intra-class variance and have aforementioned features, Deformable Part Models is the one with highest detection performance. In the scope of the work that has been done for this thesis, Deformable Part Models has been implemented on an autonomous robotic search and rescue system and some artificial dataset augmentation techniques that enable better performing models have been applied.

Keywords: Deformable Part Models, Artificial Dataset Augmentation, Human Victim Detection, Search and Rescue, Computer Vision



1.1 Literatür Özeti

Canlıların, duyarlarını vasıtasıyla elde ettikleri malumattan yola çıkarak buldukları fiziksel ortamı çevreleyen cisimlere ait renk ve şekil bilgilerine erişmelerini sağlayan zihinsel sürece görme adı verilir. Yapay zeka başlığı altında toplanan diğer alanlarda olduğu gibi, bilgisayarla görme alanı da, çoğunlukla insanın, görmeye olanak sağlayan zihinsel yetilerinin yapay sistemlerce taklit edilmesi çabasına dayanır.

Bilgisayarla görme ana başlığı altında toplanan sistemlerin tümü iki yahut üç boyutlu resimlerden edinilen bilgilerden yola çıkarak bir anlamlandırma süreci yürütmek üzere tasarlanırlar. Kameralarvasıtasıyla alınıp, her bir resim koordinatına istenilen sayıdaki renk kanalı adedinde (RGB, Grayscale vb.) yoğunluk değeri düşecek şekilde kodlanan veri, sayısal resim olarak adlandırılır ve bilgisayar hafızasına çeşitli formatlarda (.bmp, .png, .jpg vb.) kaydedilir.

Bu noktadan sonra, sayısallaştırılan ve genellikle çok boyutlu olan verinin analiz edilmesi süreci devreye girer. Resmin analizi, sayısal renk verisinden, sistemin karar mekanizmasının yapısına göre farklı biçimlere sahip olan özelliklerin çıkarılmasına dahil olan bütün uygulamaları içerir. Bu uygulamalar, resmi yanlış sonuçlar alma ihtimalini artıracabilecek gürültülerden arındırma, sistemi yavaşlatabilecek önemsiz ayrıntıların

budanması ve bir nesnenin yahut sahneden edinilmek istenen herhangi bir yapının bilgisine ulaşılabilme için fayda sağlayacak işlemleri içerir.

Özellik çıkarımının ardından karar mekanizması devreye girer ve görme olayının nihayet bulmasına yol açacak işlemleri başlatır [1]. Karar mekanizması, kabaca sistemin tespit etmek istediği görsel öğeye dair geliştirmiş olduğu inanca belirli bir oranı aşkın derecede uyan resim parçalarını “tespit” etmesi amacını taşır. Bahsedilen bu inanç, bazen basit bir özellik vektörü, bazen ise karmaşık bir istatistiksel model ile temsil edilir ve tespit işlemi de genellikle eldeki verinin, modelin karmaşıklığına uygun bir yöntem ile taranması sayesinde gerçekleşir.

Adı geçen inancın oluşturulması amacıyla bir veya birden fazla örnekten toplanan verilerden yola çıkılarak bir model oluşturulması işlemine, eğitim, bu modelin başarısının iyi bir model için öngörülen şartları sağlayıp sağlamadığını ortaya koymak amacıyla yapılan değerlendirme işlemine ise test adı verilir. Eğitim sonucu oluşturulan ve teste tabi tutulan ve modellerin işlevsel hale gelmesini sağlayan program yapılarına yapay öğrenici (learner), sınıflandırıcı (classifier), filtre (filter) yahut tesbit edici (detector) gibi isimler verilmektedir. Bu yapıların geliştirilmesine dair araştırmalar, makine öğrenmesi (machine learning) başlığı altında yürütülmektedir.

Günümüzde var olan, sayısal veriden örüntü tanıma amacıyla tasarlanmış sistemlerin büyük çoğunluğu makine öğrenmesi tekniklerinden faydalanarak tasarlanmıştır. Bu sistemlerin çalışmaya hazır hale gelebilmesi için bir eğitim sürecinden geçmeleri gerekmektedir. Bunun için ise özel olarak hazırlanmış bir veri kümesine ihtiyaç duyulmaktadır. Bu veri kümesi, algoritma tarafından tesbit edilmesi istenen öğelerin kategorisine giren pozitif resimlerin ve onlara ait bilgiler içeren pozitif etiketlerin bulunduğu pozitif veriler ve algoritmanın yok sayması yahut olumsuz olarak işaretlemesi tercih edilen negatif etiketli resimlerin bulunduğu negatif verilerden oluşur.

Daha önce bahsedilen eğitim ve test kümeleri de, birbirinden farklı negatif ve pozitif etiketli verilerin belirli bir oranda bir araya getirilmesiyle oluşturulur.

1.1.1 Bilgisayarla Görmenin Yakın Tarihi

Bilgisayarla görme, örüntü tanıma (pattern recognition), resmi bölümlere ayırma (image segmentation), eylem sınıflandırma (action classification) ve jest tanıma (gesture recognition) gibi bir çok çalışma alanını kapsayan ve özellikle otonom robotlar, güvenlik sistemleri, video madenciliği gibi günlük hayatımıza doğrudan etki eden uygulamaların ortaya çıkmasında hayli kritik bir rol oynamaktadır. Yine bu saha, örüntü işlemenin (image processing) ardından makine öğrenmesi (machine learning) tekniklerinde de kaydedilen ilerlemelerle birlikte büyük atılımların gerçekleştirildiği bir alan haline gelmiştir.

Makine öğrenmesi tekniklerinin bilgisayarla görme alanında kullanılmasına örnek olarak gösterilebilecek ilk çalışmardan biri, 90 lı yılların başıyla girilen dönem ile ortaya çıkan, gerçekleşmesinde yapay sinir ağlarının da kullanıldığı, otonom navigasyon için geliştirilen sayısal görüntü işleme tekniğidir. Zira kamera verilerinin aracın otonom hareket ettirilmesi amacıyla kararlar üreten bir yapay sinir ağını beslemesi fikri, günümüz bilgisayarla görme uygulamalarında yaygın bir biçimde görülen özellik çıkarımı (feature extraction) – eğitim (learning) problem çiftinin çözümü için sunulmuş ilk tekliflerdendir [2][3]. Daha sonraki süreçte bu uygulamaların performansını artırmak için yeni yöntemler keşfedilmiş, konu üzerinde hızla birikimsel bir ilerleme kaydedilmiştir [4][5].

90 lı yılların sonuna gelindiğinde, bilim dünyası artık hem özellik çıkarımı, hem de yapay öğreniciler alanında daha yüksek hız ve başarımla çalışan araçların ortaya çıktığı gelişmelere şahit olmuştur. Makine öğrenmesinde sıklıkla kullanılan SVM (Support Vector Machines) [6], AdaBoost [7], ve Markov Rastal Alanları (Markov Random Fields) [8] gibi yöntemlerin bilgisayarla görme alanında kullanılmasının yanında, özellik çıkarımında da SIFT [9], Haar-Like Features [10] ve Histograms of Oriented Gradients [11] algoritmalarının ortaya çıkışıyla bu alanda yapılan çalışmalar, 2000'li yılların ortalarına gelindiğinde özellikle örüntü tanıma altın çağının başladığının habercisi niteliğindedir.

Viola&Jones Dedektörünün (Rapid Object Detection Using a Boosted Cascade of Simple Features) ve Odaklı Histogramların Gradyanları (Histograms of Oriented Gradients)

algoritmasının örüntü tanımada nesnelerin sınıf içi varyasyonuna karşı bir dereceye kadar sağladığı gürbüzlük, önceki algoritmalara nisbeten hızlı çalışıyor olmalarıyla birlikte bu algoritmaların çok popüler olmalarını sağlamış ve günlük hayata yaptıkları etkiyle isimlerinden söz edilmesini sağlamıştır. Bunun yanında, daha yüksek başarımlı vadeden yeni algoritmaların ortaya çıkışı çok sürmemiştir. Pictorial Structures [12] adı verilen, resim üzerindeki nesnelerin alt parçalarıyla oluşturdukları yapıların matematiksel temsilini sağlayan bir yöntemi revize eden makalesinin[13] ardından, yine Felzenszwalb tarafından geliştirilen Değişebilen Parçalı Modeller (Deformable Part Models) [6], bu alanda geliştirilmiş en başarılı yöntem olma ünvanını yıllarca elinde bulundurmıştır.

İstatistiksel modellerin bu büyük başarısının ardından, özellikle Web 3.0 ile adından söz ettirmeye başlayan kitle kaynaklı efor (Crowdsourcing) ve kullanıcı tarafından oluşturulan içerik (User-generated Content) gibi kavramlarla açılan çığırın bilgisayarla görme alanına etkisi son derece çığır açıcı olmuştur. İnternette tedarik edilen ve sayısı milyonları bulan resimlerin, bu görev için özel olarak geliştirilmiş araçlar yardımıyla dünyanın dört bir yanındaki kullanıcılar tarafından etiketlenmesiyle bir veri kümesi oluşturulmasına olanak sağlayan ImageNet [14]projesi, açılan bu çığırın bir sonucudur. Derin Konvolüsyonel Yapay Sinir Ağlarının (Deep Convolutional Neural Networks) piksel değerlerinden yüze, arabadan insan gövdesine çok farklı boyut ve biçimlerde örüntülerin bilgilerini saklayarak işleyebilmesinin önünü açan keşifler ve geçmiş çalışmaların birikimiyle [15] birlikte tarihte bir ilk gerçekleşmiş; bir yapay karar mekanizması, resim sınıflandırma probleminde insan yetisine denk kabul edilenin ötesinde bir başarımlı seviyesine ulaşmıştır [16].

Ne var ki, söz konusu sistemin üzerinde koştığı donanım ve bağlı olduğu diğer operasyonel şartlar değerlendirildiğinde, algoritmanın uygulama alanlarının daha yaygın hale gelebilmesi için ortadan kalkması gereken bazı engeller olduğu görülmektedir. Örneğin kısıtlı işlem kapasitesine ve güvenilir olmayan iletişim kanallarına sahip olunan, gerçek zamanlı karar alma yetisine sahip olan robotik sistemlerin sıkça kullanıldığı arama kurtarma faaliyetlerinde ve askeri destek sistemlerinde bahsedilen şartların sağlanması çoğunlukla mümkün görülmemektedir.

Bu sebeple her ne kadar bu çalışma dâhilinde geliştirilmesi amaçlanan yöntemlerin kullanım sahası DPM ile sınırlı kalmasa da, yapılan deneylerin tamamının DPM üzerinde olduğu görülecektir. Çünkü makul ölçülerde, otonom olarak çalışan bir arama kurtarma robotunun gerçek zamanlı resim taraması yapması için bahsi geçen algoritmanın, üzerinde yapılan hız ve performansı artırıcı modifikasyonlar da göz önünde bulundurulduğunda [17][18][19][20][21] bahsi geçen tarzda bir sisteme adapte edilebilir olanlar arasında güçlüsü olduğu görülecektir.

1.1.2 Arama Kurtarma Çalışmaları ve İnsan Afetzedede Tespiti

Bir önceki bölümde bahsedilen adapte edilebilirliğin ölçütleri, arama kurtarma faaliyetlerinin doğası gereği ortaya çıkan şartlarla doğrudan bir bağlantı içerir. Arama Kurtarma çalışmaları, bir afete maruz kalmış yahut yaklaşan bir tehlikeyle karşı karşıya olan insanlara ulaşılmasını ve destek verilmesini amaçlayan çabaların genel adıdır [22][23].

Arama kurtarma çalışmaları, yüksek hayati risk ve operasyonel aciliyet taşıyan doğaları gereği, otomatize bir biçimde yürütülmesinde son derece büyük faydalar taşıyan uygulamalardan oluşur. İnsanın erişemeyeceği yahut erişmesi halinde hayati tehlikeye gireceği koşulların önemli bir kısmının robotların için aynı büyüklükte bir olumsuzluk teşkil etmeyeceği düşünüldüğünde, bu görevlerde robotların kullanımının getireceği faydalar daha belirgin bir biçimde ortaya çıkmaktadır.

Ne var ki, uzaktan kumanda edilen robotlar için göreceli olarak kolay olan bu görevlerin icrasında, yine görevlerin doğası gereği büyük bir zorlaştırıcı etmen ortaya çıkmaktadır: Erişimi zor ortamların (deprem, göçük vb.) birçoğunda elektromanyetizmaya dayalı iletişimin de sekteye uğradığı bir gerçektir. Bu tarz bir iletişimsizlik anında robotun doğru kararlarla çalışmaya devam etmesi ise yüksek önem taşımaktadır. Hal böyleyken robotların, çevrelerinden topladıkları bilgilere dayanarak yaptıkları çıkarımlarla, otonom bir faaliyet yürütebilmeleri arama kurtarma çalışmalarında yer alabilmeleri için vazgeçilmez bir ihtiyaca dönüşmektedir.

Bu otonom faaliyetlerin yürütülmesine katkı sağlayacak uygulamaların örüntü tanıma teknikleriyle gerçekleştirilebilecek türleri arasında, insan afetzedede tespiti (human victim

detection), resmi bölümlere ayırma ve engelden kaçınma gibi sorunlara getirilecek çözümler başı çekmektedir. İnsan afetzede tespiti uygulamalarının bilgisayarla görmenin çalışma sahasına giren kısımlarının temelde örüntü tanıma başlığı altında incelenebileceğini söylemek yerinde olacaktır. Sayısal resim verisi üzerinde, bir insanın gövdesine ait görsel yapıların aranması işleminde kullanılmaya aday algoritmalarından biri de DPM dir.

1.1.3 Literatürde Yapay Veri Artırımı

Tez çalışmasının bu bölümünde, geçmişten günümüze özellikle örüntü tanıma algoritmalarının eğitim kümeleri üzerinde uygulanmış yapay veri artırımı uygulamalarına dair ayrıntılı bilgi verilecektir. Söz konusu bilgiler çalışmalarda neden yapay veri artırımına gidildiği, nasıl bir yöntem izlendiği ve ne gibi sonuçlar alındığına ilişkin olacaktır.

1.1.3.1 Pomerlau 1989

Sürüş Esnasında Alınan (On-The-Fly) Veri ile Eğitim

Otonom sürüş için supervised learning konusunda yapılacak çalışmalar düşünüldüğünde, bilim dünyasının sahip olduğu en büyük avantajlardan biri insan sürücülerin her gün her an ürettiği hazır ve etiketlenmesi zor olmayan veriden yararlanma olanağıdır [2][3]. Teorik olarak bakıldığında, araca gittiği yönde çekim yapacak şekilde yerleştirilen bir kameradan okunan veri girdi ve şoförün direksiyonu kırma açısı da istenilen çıktı olarak kabul edilebilir. Bu sürüş esnasında eğitim fikri Şekil 1 de resmedilmektedir.



Şekil 1.1 Yapay sinir ağının sürüş esnasında kamera ve araç aksamından edinilen veri ile eğitilmesi[2]

Gerçek resimler üzerinden yürütülen eğitim, yeni durumlar için farklı ağlara geliştirmek için gerekli olan, elle programlanmış bir eğitim örneği oluşturucu yazılımının gerçekleşmesi ihtiyacını ortadan kaldırarak, çalışmadaki insan eforunu önemli ölçüde azaltacaktır. Sürüş esnasında eğitim ayrıca sistemin yeni ve öngörülmesi zor durumlara olan adaptasyonunu artırmalıdır.

Potansiyel Problemler

Sürüş esnasında elde edilen canlı resimlerin bir yapay sinir ağının eğitiminde kullanılmasıyla ilgili iki temel potansiyel problem bulunmaktadır. Bunlardan ilki normal bir sürüş esnasında şoförün daima yolun uygun bölümünden istenilen tarafa doğru gidiyor olduğu gerçeğidir. Başlangıçta istenilen bir durum olarak görülen bu rutin, eğitilecek olan yapay sinir ağına, aracın istenmeyen bir yön yahut konuma sahip olduğu durumda nasıl bir karar vermesi gerektiğine dair bilginin hiçbir zaman sunulmaması anlamına gelmektedir. Sistem sürüş esnasında zaman zaman yanlış bir karar verip yolun merkezinden sapabilir. Bu sebeple bu tarz durumlara hazırlık olması, yanlış bir kararın hemen ardından aracı tekrar olması gerektiği konum ve yöne getirebilmesi gerekmektedir. İkinci problem ise videolardan yola çıkarak hazırlanan veri kümelerinin sisteme daima dümdüz bir yol takip etme alışkanlığını kazandırması ihtimalidir. Verilerin elde edildiği düzenlerin genellikle dümdüz bir yol üzerinde gidip nadiren viraj alan yahut dönen bir seyrüsefere neden olduğu düşünüldüğünde, overlearning

hadisesinin gerçekleşerek, sistemin her durumda düz gitme, viraj ve dönüşlerde istenilen kararı verememe gibi sorunlarla karşılaşması muhtemeldir.

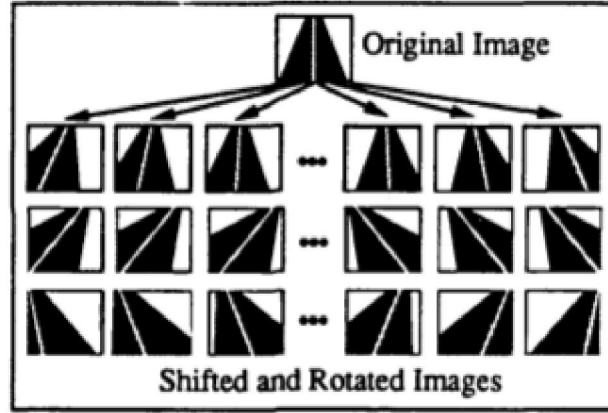
Sürüş esnasında eğitimde karşılaşılan bu iki problemin de geri yayılım (back-propagation) sürecinin, öğrenilmesi gereken görevin tamamını temsil edecek biçimde bir eğitim verisini gerekli kılmasından kaynaklandığı söylenebilir. Çalışmada bu problemi çözmek amacıyla yapılan ilk girişim, şoförden sürüş esnasında (gerekli olmadığı durumlarda da) sapmalar yaparak gitmesini rica etmek olmuştur. Bu ricanın ardındaki fikir, sistemin yoldan çıkmış bir aracı nasıl eski haline döndüreceğini öğrenmesi için gerekli olan verinin sağlanması idi. Ne var ki bu yaklaşımın iki sebepten ötürü uygunsuz olduğu anlaşılmıştır. Bunlardan birincisi, eğitim sırasında sapışlar yapmak, düz gidilmesi gereken bir yolda yanlış bir karar verilerek direksiyon kırmak anlamına gelmektedir. Bu esnada oluşturulan bir veri ile eğitilen bir sistem, eğitim halini değiştirme yeteneğine sahip olmadığı sürece istenmeyen bir şekilde insan sürücünün düz yolda sapma eylemini taklit etmeye başlayacaktır. Sürüş esnasında devreye sokulacak ek bir manuel giriş ile eğitim "kapatmak ve tekrar açmak" mümkün olmasına rağmen çalışmanın genel gidişatı boyunca önemsenen eğitim verisine müdahale edilmemesi yaklaşımı burada da korunmuştur. Sürüş esnasında sapmalara dair ikinci problem ise genel bir temsil elde edebilmek için birçok farklı durumda sıklıkla sapmalar yapılması gerekliliği olmuştur. Bu da hem zaman harcıyıcı bir işlemdir, hem de trafikte yapılan bir eğitimde tehlikeli olabilmektedir.

Çözüm: Kamera Resimlerini Dönüştürmek

Sürüş esnasında sapma yönteminde karşılaşılan problemlerle karşılaşmaksızın eğitim kümesinde ihtiyaç duyulan çeşitlilik seviyesine ulaşabilmek için, eğitim için yeni örnek verilerin üretilmesinde kameradan alınan resimlerde şeklen dönüştürmelere gidilmiştir. Yapay sinir ağının yalnızca var olan yönü ve direksiyon açısını kapsayan bir karar mekanizmasını taklit etmesinin önüne geçmek için resimler yazılım marifetiyle ötelenmiş ve döndürülmüş, bu yolla aracın Şekil 2 de görülen yeni şartlara uygun yeni tepkiler verdiği resim-karar çiftleri üretebilecek hale gelmesi sağlanmıştır. Kameradan alındıktan sonra döndürülen veriye karşılık gelen karardaki direksiyon açısının ne

olacağı bilinmektedir; zira perspektif geometri kullanılarak yeterince hassas transformasyonlar bu açının hesaplanmasında kullanılacaktır.

Döndürülmüş ve Ötelenmiş Resimler



Şekil 1.2 Eğitim süreci için ihtiyaç duyulan örneklerin üretilebilmesi amacıyla kaydırılmış ve döndürülmüş resimler[2]

Resmin transformasyonu işlemi resimde yer düzleminin ne kadarının görüldüğünün ve transformasyondan sonra ne kadarının görüneceğinin saptanmasıyla başlamaktadır. Bu iki alan birbirini üzerine gelen iki yamuk şekli alır. Bir pikselin transformasyonu için uygun değeri bulmak için, pikseller yer düzlemine yansıtılırlar, ve sonra resme geri yansıtılırlar. Orjinal resimde karşılık bulan pikselin değeri dönüştürülen resimdeki piksel değerine atanır. Burada dikkat edilmesi gereken önemli hususlardan biri bir transformasyonu gerçekleyen piksele piksel eşleştirmenin sabit oluşudur. Diğer bir ifadeyle düzlemsel kabul edilen bir evrende resime herhangi bir transformasyon uygulamak için kullanılması gereken pikseller daima aynı olacaktır. Uygulanan resim transformasyonu tekniğinde değişecek olan tek şey resim nasıl transforme edileceği olduğundan, piksellerin bir önhesaplama tabii tutulması ile işlem yüksek oranda hızlandırılmaktadır.

1.1.3.2 Drucker ve diğerleri 1994

Bir boosting algoritması sonlu bir hata payına sahip herhangi bir yapay öğreniciyi, hata payı belirgin oranda daha düşük olan bir yapay öğrenciler ensemble na dönüştürebilir [4]. Çalışmada önce, Shapire'in, Valiant'ın "Muhtemel Yaklaşık Doğru (Probably

Approximately Correct)” eğitim modelinden yola çıkarak bu durumu teorik olarak izah ettiğinden bahsedilmiştir. Sonra bu eğitim modelinin tahminlerinden yola çıkarak, Shapire’in tekniğinin, rastgele tahminden daha iyi sonuçlar veren herhangi bir öğrenici türünün performansını, hata payını çok küçük seviyelere indirebilecek bir biçimde artırabileceğinden bahsedilmiş ve tarihte ilk defa bir yapay sinir ağına boosting algoritması uygulanışının ardındaki motivasyon açıklanmıştır.

Yöntem (Boosting kullanılarak supervised learning ile yapay sinir ağlarından bir ensemble üretilmesi) üzerine verilen bilgiden sonra yapılan değerlendirmede yer verilen ifadelerden anlaşılmaktadır ki, tarif edilen işlemi gerçekleştirmek için sahip olunması gereken etiketli veri adedi, 30000 i bulmaktadır. Etiketleme maliyeti bu denli yüksek olan bir algoritmanın, bir de makalenin yazıldığı dönem göz önünde bulundurulduğunda geçerli bir çözüm sayılabilmesi için bu soruna bir çözüm bulması şarttır. Tam da bu soruna binaen geliştirilen bir yapay veri üretimi modeli, çalışma dahilinde tanıtılmıştır.

Çalışmanın ana başlıklarından biri, performansını artırmak istedikleri bir yapay sinir ağının yüksek veri adedi ihtiyacını karşılamak için Simard [24] ’ın önerisinden yola çıkarak geliştirilen bir görsel deformasyon modeli. Bu teknik sayesinde yapay veriler üretilebiliyor. Tekniği kısaca açıklamak gerekirse;

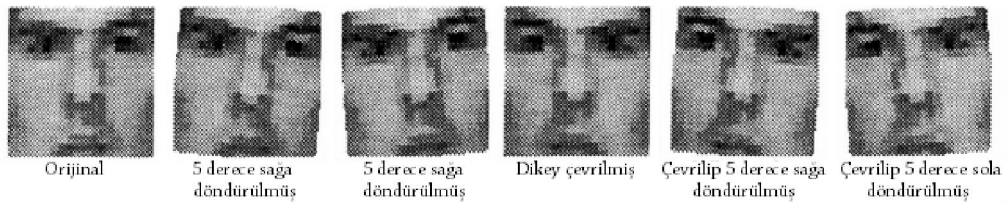
Resim kare bir pikseller dizisiyle ifade ediliyor olsun (16x16 ya da 20x20). Resmin (i, j) koordinat konumlarındaki yoğunluk değerinin (intensity) $F_{ij}(x, y)$ olduğunu, x ve y nin de F in, i ve j noktasında, x ve y nin türevlenebilir ve bu sebepten sürekli bir fonksiyonu olduğunu göstermek için kullanılan parametreler olduğunu söyleyelim. (i, j) konumlarındaki F değerinin, x ve y deki küçük ötelemelerden, dönmelerden, diagonal deformasyonlardan, eksensel deformasyonlardan, ölçeklemeden ve kalınlık deformasyonundan kaynaklı değişimlerinin değeri, aşağıdaki matrislerin birbirleriyle çarpımı sonucunda bulunabilir:

$$\Delta F_{ij}(x, y) = \left[\frac{\partial F_{ij}(x, y)}{\partial x} \quad \frac{\partial F_{ij}(x, y)}{\partial y} \right] \left\{ k_1 \begin{bmatrix} 1 \\ 0 \end{bmatrix} + k_2 \begin{bmatrix} 0 \\ 1 \end{bmatrix} + k_3 \begin{bmatrix} -y \\ x \end{bmatrix} \right. \\ \left. + k_4 \begin{bmatrix} y \\ x \end{bmatrix} + k_5 \begin{bmatrix} -x \\ y \end{bmatrix} + k_6 \begin{bmatrix} x \\ y \end{bmatrix} + k_7 \begin{bmatrix} \frac{\partial F_{ij}(x, y)}{\partial x} \\ \frac{\partial F_{ij}(x, y)}{\partial y} \end{bmatrix} \right\}$$

Tekniğin kullanımı ise, denklemin oluşturulma tarzından da anlaşılacağı gibi, birden yediye kadar sunulmuş olan k parametrelerinin farklı değerler alabilmesiyle mümkündür. Bu parametrelerin çok büyük seçilmesinin orjinal resimle sentezlenen resim arasındaki farklılığın ihtiyaç duyulan seviyenin üzerine çıkması olasılığını doğuracağı aşikardır. Tersî durumunda ise yeterli deformasyon sağlanmamış olacaktır.

1.1.3.3 Sung ve Poggio 1998

“Birçok örnek bazlı öğrenme uygulamasında bir öğrencinin görevlerindeki başarısı, eğitim sırasında kullanılmış örneklerin kalitesine ve çokluğuna yüksek oranda bağlıdır [25]. İdeal bir eğitim senaryosu, bilhassa hedef küme üzerinde kapsayıcı bir örnekleme yapabilmeyi mümkün kılmak için, yapay öğrenciye mümkün olduğunca çok veri sunulmasıyla hayata geçirilebilir.” İşte Sung ve Poggio’nun bu çalışmasında yüz tanıma görevini üstlenecek yapay öğrencilerin eğitiminde kullanılmak amacıyla azami miktarda veriye sahip olmak istenmesinin ardında yatan neden buydu. Amerikan deniz kuvvetleri için MIT tarafından yürütülen çalışmada kullanılacak resimler, ulusal vesikalık fotoğraf veri tabanından alınan verilerden oluşmaktaydılar ve uygulanan veri kümesi büyütme prosedürü ile hayli yüksek bir adede ulaştırılmıştırlar.



Şekil 1.3Yapay sinir ağının başarımını artırabilmek için üretilmiş yapay veriler ve orjinal örnek[25]

Sung ve Poggio'nun bu çalışmasında Bunun da resimlere döndürme ve aynalama (rotation and mirroring) uygulanarak uygulandığından söz ediliyor.

1.1.3.4 Sim 2000

Kişi başına düşen resim adedini en aza indirmek amacıyla uygun olarak eğitim kümesi sonradan eklenecek, sentetik olarak üretilmiş yüz resimleriyle genişletilmek istenmektedir [26]. Problemin çerçevesi yakın ve önden çekim resimler üzerinden yüz tanıma olarak belirlenmiş olduğundan, gerekli resimler basit geometrik transformasyonlar (döndürme, öteleme ve ölçekleme) ile oluşturulabilmektedirler. Daha karmaşık transformasyonlar söz konusu olduğunda kastedilen ve iki boyutlu bir düzlemin ötesine geçen perspektif müdahaleleri, özellikle üç boyutlu bir modeller oluşturma gibi yöntemlere dayanmakta, bu sebeple bütün resim türlerinde etkin olarak kullanılamamaktadır. Bunun dışında bahsedilen normalized correlation işlemi denince akla gelmesi gereken işlemde kasıt bir çeşit öteleme işlemi iken, bu işlem herhangi döndürme yahut ölçekleme içermemektedir. Bu çalışmada bahsedilen yöntem yakından incelendiğinde otonom gezinme ve otomatik rakam tanıma için oluşturulan yapay sinir ağlarının eğitiminde kullanılan ilave resimleri üreten yöntemlere benzerliğiyle dikkat çekecektir [27][15][28].

1.1.3.5 Ke 2004

Yan Ke ve Rahul Sukthankar tarafından 2004 yılında yürütülen bu çalışmanın temel amacının eş şekillilik (homografi) analizi yapmak için kullanılan Scale-invariant feature transform (SIFT) algoritmasına bir eklenti olarak PCA (Principal Component Analysis) algoritmasını devreye sokmak suretiyle bu algoritmanın performansını artırmak olduğu söylenebilir [29].

SIFT ile PCA-SIFT algoritmalarının özellik çıkarım süreçleri sonucunda üretilen verilerinin resmi temsil edişleri arasındaki farkları gözlemlemek amacıyla çalışma boyunca uygulanmış olan 3 temel deney türü vardır.

İlk deney türü özellik çıkarıcıların, gürültü eklenmesi, ışıklandırmada değişiklikler yapılması ve resim transformasyonlarının uygulanması sonucu yapay şekilde ortaya

çıkan etkilere karşı gürbzlüklerinin ölçülmesi ve karşılaştırılması amacını taşımaktadır. Çalışma dahilinde bir veri kümesi oluşturulmuş ve aşağıda zikredilen işlemler her resme uygulanmıştır:

(1) Gauss gürültüsü ($\sigma = 0.05$) resim yoğunluklarının 0 ila 1 arasında değiştiği kabul edilerek; (2) 45 derecelik bir döndürmenin ardından uygulanan % 50 lik bir ölçek büyütme; (3) %50 lik bir yoğunluk ölçeklenmesi; (4) 30 dereceye yakın bir bakış açısı değiştirmeye denk gelecek miktarda yansıtma eğme (projective warp) işlemi.

1.1.3.6 Tsang 2005

Ivor W. Tsang, James T. Kwok ve Pak-Ming Cheung, zamansal karmaşıklığı $O(n^3)$ olan Support Vector Machine (SVM) öğrencilerinin büyük veri kümeleri kullanılarak eğitilmelerinin zorluğundan yola çıkarak bu tekniği hızlandırma amacıyla Core Vector Machine (CVM) adını verdikleri bir öğrenci geliştirmiş ve çalışmalarını 2005 yılında bir yayın haline getirmişler [30]. Amaçlarının doğal bir sonucu olarak büyük bir veri kümesine ihtiyaç duymuş ve bu ihtiyaca karşılık gelecek yapay veri üretim metotları kullanmışlardır.

Eğitim kümesinde 1005 adet negatif ve 1194 pozitif resim mevcuttur. Test setinde ise 359 negatif ve 264 pozitif resim bulunmaktadır. Algoritmanın resimlerin ölçeklenmesine karşı tepkisinin nasıl olacağını ölçmek için veri kümesindeki resimler üzerinde önce çözünürlüklerini 16x16 dan 26x26 ya getirmek suretiyle değiştirmek, sonra da resimleri her ekseninde 5 er piksel ötelemek suretiyle değişiklikler yaparak veri kümesini genişletme işlemi yapılmıştır. Böylece sonuçta ortaya çıkan eğitim kümesi $(1005 + 1194) \times 112 = 266,079$ ve genişletilmiş test kümesi $(359 + 264) \times 112 = 75,383$ resimlik bir boyuta ulaşmıştır.

1.1.3.7 Laptev 2007

Ivan Laptev tarafından 2007 yılında ortaya konulan bu çalışmanın üzerine yoğunlaştığı problem, resimlerin piksel değerlerinden oluşturulan histogramlar kullanılarak nesne tespiti yapmak üzere eğitilen sınıflandırıcıların ulaşılmış olduğu performansı aşmak olarak ifade edebilir [31]. Bu problemin çözümü için geliştirilen yaklaşımı özetlemek

gerekirse, HoG ve birçok diđer özellik çıkarımı yönteminde görülen histogram oluşturma, onun kullanılmasını hızlandıran Integral Histogram yöntemi ve Viola ve Jones'un geliştirdiđine çok benzer bir zayıf öğrenici kullanılarak oluşturmuş bir sınıflandırıcının bir araya getirilmesinden söz etmek yeterli olacaktır.



Şekil 1.4 Pozitif etiketlere ve resim boyutlarına yapılan müdahalelerin sonuçları. Belirlenen bir bölümleri seçilerek kırılan resimler ve bu bölüm kaydırılarak üretilen ek resimler[31]

Sık sık karşılaşılan pozitif etiketli verilerin sınırlı sayıda olması probleminin çözümü için, veri kümesinin yapay olarak genişletilmesi faydalı görülmüştür. Eğitim kümesindeki dikdörtgen şekilli etiketlere yönelik yapılan çalışmalar şöyle tarif edilebilir: Bütün etiketlere karşılık gelen orijinal dikdörtgenlerin boyutlarına ve konumlarına eklenen gürültü sayesinde benzer dikdörtgenler elde edilmiştir. Bu gürültülü etiketler yeni pozitif resim örnekleri oluşturmak için kullanılmıştır ve bu yolla pozitif eğitim kümesi yapay bir şekilde genişletilmiştir. Prosedür Şekil 4'te görselleştirilmiştir.

1.1.3.8 Howard 2013

Andrew Howard tarafından yapılan bu çalışma, Deep Convolutional Neural Network denilen ve hali hazırda var olan en efektif nesne sınıflandırma sistemi olarak kabul edilen yapay öğrenicinin yetkinliğini artırıcı yöntemler üzerine araştırmaları ve deneyleri içermektedir [32]. Yapılan deneylerin yoğunlaştığı noktalardan biri de daha fazla eğitim verisinin elde edilmesi için ortaya konan eforlar ve sonuçlarıdır.

Deep neural network adı verilen sınıflandırıcıların eğitim verisi adedinde artırıma gidildiğinde daha iyi sonuçlar verdiği gözlemlenmiştir. Fazladan eğitim verisi var olmadığında ise, verideki varyasyonun yerini alması amacıyla var olan eğitim verisi üzerinde birtakım transformasyonlara gidilmesi eğitim kümesinin boyutunu sentetik olarak artırabilmektedir. Daha önce Imagenet sınıflandırma sistemi üzerinde 3 transformasyon yöntemi, eğitim kümesini büyütmek amacıyla uygulanmıştır [14].

Eğitim kümesini büyütmek için, pozitif resim örneklerine üç ayrı çeşit transformasyon uygulandı. Bunlardan ilki eldeki 256 x 256 lık resimlerden rasgele 224x224 lük bölgeler kırılarak bunların yeni örnekler olarak kaydedilmesi idi; bu işlem modele ötelemeden bağımsızlığı kazandırmak için yapılıyordu. İkincisi ise eldeki resimlerin yatay olarak çevrilmesine dayanıyordu ki yansıma bağımsızlığına ulaşmak için ortaya konmuş bir yöntemdi. son veri dönüştürme tekniği ise resme rastgele üretilmiş bir ışık eklemeye dayanıyordu. Bu yöntem, üretilecek modelin ışığın değişimine ve resim üzerindeki ufak renk değişimlerine karşı gürbüz olması içindi.

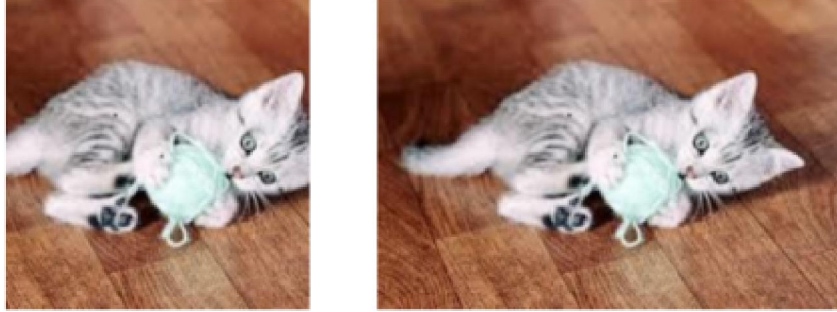
Kırpma işlemi

Önceki uygulamalarda tercih edildiği görülen 256 x 256 lık resimlerden kare biçiminde kırılan resimlerin 224x224 lük resimlerin elde edilmesi için kullanılan yöntem, orjinal veri kümesinde var olan resimlerin büyük olan kenarına karşılık gelen 256 ya eşitlenecek şekilde ölçeklenmesi, daha sonra diğer kenarın da 256 ya eşit olacak şekilde kırılmasına dayanıyordu. Kalan işlemler bu oluşan orjinal resim parçası üzerinde uygulanıyordu. Bu eldeki piksellerin %30 unun kullanılmamasından doğan bir bilgi kaybına yol açmaktaydı. Kırılmış piksellerin orta kısımda bulunan piksellerden daha bilgi verici olduğu tahmin edildiğinden bu diğer pikselleri kullanarak modelin daha da geliştirilebileceği kaydedilmiştir.

Bütün resmi kullanabilmek adına, resmin kısa tarafının 256 piksele ölçeklenmesi sonucunda 256 x N lik yahut N x 256 lık bir resim elde edilir. Sonra bu resim içerisinden 224 x 224 lük bir kesmin rastgele kırılmasıyla yeni resimler elde edilir. Bu sayede daha fazla eğitim kümesi örneği elde edilmekte ve oluşan modelin öteleme işlemine olan toleransı daha yüksek olmaktadır.

Resimde görüldüğü gibi iyi ortalanmış resimlerde dahi yanlış kırılma bilgi kaybına (kulak ve kuyruk gibi ayırt edici vücut parçaları) yol açabilmektedir. Çalışmada yeni eğitim resimleri tüm resmi kullanarak oluşturulmaktadır.

Ek Renk Manipulasyonları



Şekil 1.5 Yatay ekseninde kısaltılmış bir pozitif örnek. Hedef nesne olan kedi şekline ait kuyruk ve kulak gibi öğelerin kırılma sonucu kaybolduğu görülüyor[32].

Bundan önce kullanılmış olan rastgele yıldırım gürültüsünün dışında [14], bu çalışmada yapılan başka renk eklemeleri de kontrastı, parlaklığı ve rengi python resim kütüphanesini (python image library (PIL)) kullanarak manipule etmek suretiyle yapılmaktadır. Bu işlem yapay sinir ağına resimlerdeki renk bazlı değişikliklerden bağımsız bir şekilde çalışabilme konusunda faydalı olacak eğitim örneklerinin üretilmesini mümkün kılmaktadır. Yapılan çalışmada bu manipulasyonların sırası rastgele olarak seçilmekte ve bu işlemler, 0.5 ile 1.5 arasında bir dönüştürme katsayısına bağlı olarak uygulanmaktadır. (1 değeri resmin sabit kalması anlamına gelir.) Kontrast, parlaklık ve renk değişimlerinden sonra, önceki yöntemlerde de kullanılmış olan yıldırım gürültüsü uygulanır [14].

1.1.3.9 Cireşan ve diğerleri 2011

MNIST veri kümesi üzerinde yapılan çalışmalarda yapay sinir ağları deforme edilmiş resimlerle eğitilmiştir [33]. Bu resimler eğitim sırasında ihtiyaca binaen sürekli olarak eğitilmişlerdir. Afin (öteleme, döndürme, ölçekleme ve yatay çapraz kaydırma (horizontal shearing)) ve elastik deformasyonlar, bu ilave resimlerin üretilmesi amacıyla birleştirilmiştir [34].

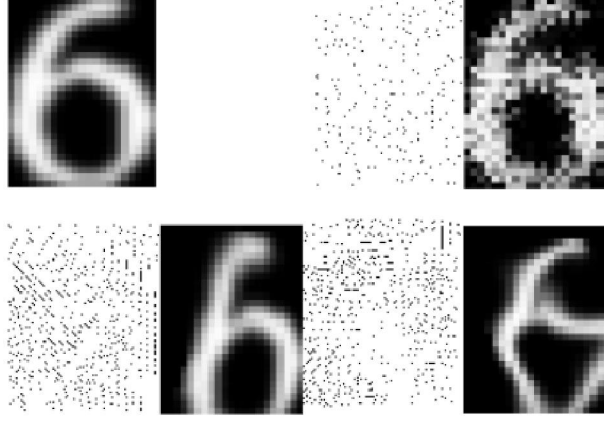
1.1.3.10 Simard 2003

Veri kümelerinin yapay yollarla genişletilmesi üzerine yapılan çalışmaların en ünlülerinden biri de Patrice Simard imzası taşımaktadır. Resimler içerisindeki altbirimler (substructures) üzerine uygulanan dönüşümlerin, yani afin deformasyonların ortaya çıktığı bu işlemler, aşağıdaki gibidir.

Veri Kümelerinin Elastik Bozunmalar Marifetiyle Genişletilmesi

Afin bozunmalar MNIST veri kümesi üzerinde alınan sonuçları büyük ölçüde iyileştirmiştir ama bu çalışma dahilinde alınan en iyi sonuçlar elastik deformasyonlar sayesinde alınmıştır [34]. Resim deformasyonları öncelikle rastgele deplasman alanlarıyla oluşturulmuştur. Bu alanların oluşturduğu deplasmanı şöyle açıklayabiliriz:

$\delta x(x, y) = 1$, ve $\delta y(x, y) = 0$ ifadesi bütün piksellerin x ekseninde 1 birim ötelenmesine karşılık gelsin. $rand(-1, +1)$ 1 ile -1 arasında uniform dağılımlı rastgele bir sayı olmak üzere, $\delta x(x, y) = rand(-1, +1)$ ve $\delta y(x, y) = rand(-1, +1)$ şeklinde ifade edilen alanlar oluşturulmuştur. δx ve δy alanları daha sonra σ piksellik bir varyansa sahip bir Gaussian ile konvolüsyona tabi tutulmuştur. σ büyük olduğunda ortaya çıkan değerler son derece küçük olacaktır çünkü rastgele değerler 0 a yakın olacaktır. Eğer deplasman alanını 1 ile normalize edersek, alan sabit ve rastgele yöne sahip olarak şekillenecektir. Eğer σ küçük ise, alan normalizasyonun ardından tamamıyla rastgele bir görünüme sahip olacaktır. (Resimde sağ üst köşede görüldüğü gibi). Bu bahsedilenlerin arasında kalan σ değerleri için ise deplasman alanları elastik deformasyona uğramış gibi görünürler, σ ise elastiklik katsayısı haline gelir. Deplasman vektörleri bu işlemde sonra bir α katsayısı ile çarpılır ki bu katsayı da deformasyonun yoğunluğunu belirlemektedir. Resimde bir saf rassal alana ($\alpha = 0.01$), bir yumuşatılmış rassal alana ($\alpha = 8$), ve çok fazla değişkenliğe sahip bir rassal alana ($\alpha = 4$) ait grafiklere yer verilmiştir. Çalışma dahilindeki deneylerde ($\alpha = 8$) değeri kullanılmıştır. $\alpha = 8$ den daha büyük olduğunda deplasmanlar afinliğe yaklaşmaya başlar, ve α çok büyük olduğunda ise deplasmanlar ötelemeye dönüşürler.



Şekil 1.6MNIST veri kümesindeki örneklere benzer işlevlere sahip verilerin üretimi amacıyla yapılan deneylerin sonuçları[34]

1.1.3.11 Yaeger 1996

Yaeger in sunmuş olduğu teknikte, eğitim sırasında (Test sırasında değil), el yazısı verisi üzerinde, çekme, dönme ve x ekseninde ve y ekseninde olmak üzere doğrusal ve karesel ölçeklemelerle rastgele varyasyonlar üretilmektedir[35]. Bu değişimler sanatçılar arasındaki stilistik varyasyonlara karşılık gelecek alternatif karakter biçimleri oluşturmakta ve standart geri yayılım (back-propagation) çerçevesinin doğrudan bir sonucu olarak algoritmaya farklı boyut oranlarından ve dönme açılarından bağımsız olma yeteneğini kazandırmaktadır. Meydana getirilen bozunmaların her birinin büyüklüğü cross-validation deneyleri sonucunda belirlenmiş, en iyi genelleştirmeyi sağlayacak değerler kullanılmıştır. Ayrıca bu tarz örneklerin doğal bir değişim hissi yarattıklarından emin olunabilmesi için, örnekler yazarlar tarafından göz ile de incelenmişlerdir. Bu tarz değişikliklere sahip küçük bir kümeyi aşağıdaki resimde görmek mümkündür.



Şekil 1.7Orjinal el yazısı resimleri üzerinde küçük değişiklikler yapmak suretiyle stilistik birtakım varyasyonların yapay olarak taklit edilme çabalarının sonuçları[35]

1.1.3.12 Khrizevsky ve Hinton 2012

Resim verisi üzerinde aşırı öğrenmeye (overfitting) karşı kullanılacak en kolay ve en yaygın yöntem etiketlerin korunması şartı ile veri kümesini yapay olarak genişletmektir

[14][15][33]. Çalışmada veri kümesi genişletiminin iki farklı biçimi uygulanmıştır, bunların ikisi de orjinal veri kullanılarak çok küçük bir hesaplama maliyetiyle üretim yapabilmeyi mümkün kılmaktadırlar. Örneğin dönüştürülmüş resimler disk üzerinde dahi tutulmak zorunda değildir. Bunun sağlanabilmesi için biliminsanlarının başvurmuş oldukları yöntem şöyle anlatılabilir: Çalışma dahilinde ortaya konan gerçekleştirilmede, dönüştürülmüş resimler CPU üzerinde koşan Python kodu ile oluşturulmuştur çünkü GPU önceki resimler grubu üzerinde yaptığı eğitime devam etmektedir. Yani bir bakıma bu veri artırımının yöntemlerinin fiilde, hesapsal olarak neredeyse bedava oldukları söylenebilir, zira eğitim için belirli bir çalışma zamanının geçecek oluşu kaçınılmazdır. Veri artırımında kullanılan ilk yöntem resim öteleme ve yatay aynalama marifetiyle veri üretimine dayanmaktadır. Bunu yapmak için 256x256 lık resimlerden 224x224 lük birçok rastgele parça kırılıp, bunların bir de yatay yansımaları alınmıştır. Sadece bu işlem ile veri kümesinin boyutu 2048 katına çıkmıştır ama eğitim örnekleri tabii ki önceki duruma kıyasla çok daha benzer hale gelmişlerdir. Bu yöntem olmaksızın yapay sinir ağı belirgin bir biçimde aşırı öğrenme göstermektedir ki bu da biliminsanlarını çalışmalarında çok daha küçük ağlar kullanmaya itmiştir. Test sırasında, yapay veri ağı tahminlerini 224x224 lük 5 adet (4 adet köşe parça ve bir merkez parça olmak üzere) parça kırarak, onların yatay yansımalarını (dolayısıyla toplamda 10 parça) alarak ve tahminlerinin en yumuşak katmanda ortalamalarını alarak yapar. Veri artırımının ikinci bir biçimi ise eğitim verilerinin RGB kanallarının yoğunluklarının değiştirilmesiyle elde edilir. Çalışmada bu değişimi sağlamak için ImageNet eğitim kümesine üzerindeki RGB piksel verilerine PCA uygulanmıştır. Bütün eğitim verilerine, büyüklükleri, bulunan principal componentlerin eigendeğerlerinin bir sıfır ortalamaya ve 0.1 standart sapmaya sahip bir Gaussian'dan çekilen bir rasgele değerle çarpımıyla elde edilmiş değerle çarpımının sonuçları eklenmiştir. Bu sebeple her RGB resim pikseli $I_{xy} = [I_{xy}^R, I_{xy}^G, I_{xy}^B]^T$ için aşağıdaki değer eklenir :

$$[p_1, p_2, p_3][\alpha_1 \lambda_1, \alpha_2 \lambda_2, \alpha_3 \lambda_3]^T$$

yukarıdaki p_i ve i sırasıyla RGB piksel verilerinin 3x3 lük bir covaryans matrisinden elde edilen eigenvektör ve eigendeğerleri iken i , bahsi geçen rassal değerdir. Her bir i , her bir resim eğitimde bir kez daha kullanılıncaya kadar her bir defa olmak üzere çekilir ve bu işlem o resmin her kullanılışında tekrarlanır. Bu işlem doğal resimlerin önemli bir

özelliğini yaklaşık olarak yakalamak ve resimlerdeki renk yoğunluğu ve aydınlık değişimlerinden bağımsız bir model oluşturarak nesne kimliğini ön plana çıkarmak üzere uygulanır. Bu yöntem sayesinde en düşük hata payına sahip düzenlemeye göre %1 lik bir azalma sağlanmıştır.

1.1.4 Literatürden Edinilen Dersler

Veri kümelerinin yapay yollarla genişletilmesinde günümüz biliminsanlarınca ulaşılan nokta incelendiğinde, kullanılan yöntemlerin değerlendirilmesinde, kullanılan yöntemlerin hesapsal ve finansal maliyetinin düşüklüğü, yeterince etkili bir bozunma meydana getirebiliyor olmaları, eğitim kümesinin sınıf içi varyasyonunu haddinden fazla yükseltmiyor olmaları gibi faktörlerin ön plana çıktığı görülmüştür.

1.2 Tezin Amacı

Örüntü tanıma algoritmaları, test süresi, eğitim süresi, etiketleme maliyeti, tesbit etme başarımı gibi birçok perspektiften değerlendirilirler [29][36][37]. Bu tez dahilinde yapılan çalışmalar, örüntü tanıma algoritmalarından Değişebilen Parçalı Modeller algoritması üzerine yoğunlaşmıştır ve bu tekniğin farklı açılardan iyileştirilmesi için gerçekleştirilen birtakım gerçeklemeleri ve deneyleri içermektedir.

Geçmişten günümüze bilim ve teknolojinin seyri, özellikle yapay zeka gibi popüler ama her zaman birikimsel olmayan çabaların hakim olduğu alanlarda, yapılan çalışmaların kalıcı etkileri olmasını sağlayacak şekilde yönlendirilmesinin ne kadar önemli olduğunu göstermektedir. Örneğin DPM gibi bir algoritmanın iyileştirilmesi için ortaya konan son derece değerli ve başarılı çabalar bulunmakla birlikte, bu çalışmaların birçoğu bir yapay sinir ağı yahut bir başka öğrenici söz konusu olduğunda etkisiz kalmakta ve bu yüzden zaman zaman geçmiş gayretlerin sağladığı avantajlardan yararlanılamamaktadır.

Bu sebeple tezin amacı, algoritma başarımını, algoritmanın içyapısından bağımsız bir şekilde yapılan müdahalelerle artırmak olarak belirlenmiştir. Bu amaç doğrultusunda algoritmaların eğitiminde ve testlerinde kullanılan veri kümelerinden azami düzeyde faydalanmalarına yönelik çalışmalar yürütülecek ve bu amaçla kullanılacak çeşitli tekniklerin başarıları modern ölçütler ışığında karşılaştırılacaktır.

Bunun yanında tez çalışmaları dâhiline alınan bir diğer iş de DPM algoritmasının daha hızlı çalışan bir gerçekleştirilmesi olan FFLD [19] nin ROS [38] a adapte edilmesi işlemidir. Böylece Yıldız Teknik Üniversitesi Olasılıksal Robotik kulübü dâhilinde yönetilen arama kurtarma robotu geliştirimi çalışmaları dâhilindeki insan kurban tespiti modülünü geliştirme amacına dair temel gereksinimleri karşılamada bir başarı kaydedilmiştir.

1.3 Hipotez

Tez çalışmaları kapsamında Değişebilen Parçalı Modeller (Deformable Part Models) algoritmasının başarımını artırmak amacıyla çeşitli deneyler gerçekleştirilmiştir. Veri kümesinin yapay yollarla genişletilmesi vasıtasıyla algoritmanın yapılan testlerde elde ettiği ortalama doğruluk değerinin yükseltilmesi hedeflenmiş, bu yükselme sırasında eğitim ve test sürelerinde gerçekleşecek değişim de mercek altına alınmıştır.

Bu tez çalışmasıyla ortaya konmuş olan temel hipotez şu şekilde ifade edilebilir :
“Değişebilen Parçalı Modeller, afet alanlarında sayısal resimler üzerinde kurban tesbitinde kullanılabilirler ve bu algoritmanın başarımını artırmada, eğitiminde kullanılan verilerin kümesinin yapay yollarla genişletilmesine dayanan yöntemler test süresine ciddi bir olumsuz etki yapmaksızın başarılı olmaktadır.”

BAZI ÖNEMLİ ÖRÜNTÜ TANIMA SİSTEMLERİ

Bu bölümde bilgisayarla görme literatüründe yer alan ve özellikle bu tez çalışmasının kapsamı ve konusu açısından önem arz eden algoritmaların tanıtımı yapılacaktır.

2.1 Odaklı Gradyanların Histogramları

Temelinde, resimlerin kıyası için onların sayısal özetlerinin çıkarılabileceği, bu özetlerinse, resimdeki nesnelerin sahip olduğu kenar/köşelerin uzanım yönlerine bunların yön değiştirme büyüklüklerini karşılayan sayılarla oluşturulabileceği varsayımını barındırır [11]. Bu varsayımdan yola çıkılarak, resim hücrelere bölünür ve her bir hücre uygulayıcının seçimine bırakılan adette histogramlar ile ifade edilir. Histogramların her biri, hücrenin içerisinde bulundurduğu piksellere ait açı (*orientation*) ve büyüklük (*magnitude*) e bağlı olarak resme ait veriler tutar.

Odaklı Gradyanların Histogramları (Histogram of Oriented Gradients) algoritması, bir resmin dikey ve yatay düzlemdeki türevlerini alır. Bu türevlerin herhangi bir pikseldeki değerlerinin oranı, yani açı (*orientation*) sı ($\frac{\partial y}{\partial x}$), o pikselin sahip olduğu yön değişiminin değişiminin açısına eşittir denir. Bu türevlerin kareleri toplamının karekökü olan büyüklük (*magnitude*) ($\sqrt{\partial y^2 + \partial x^2}$) ler ise, elde edilmiş açılara binaen 180 derecelik uzayda istenilen çözünürlükteki sepet (bin) lere dağıtılır. Her bir sepet, daha sonra

birleştirilip, resmin global birer histogramı haline getirilecek ve bütün histogramlar birleştirilip o resme ait özellik vektörünü oluşturacaklardır.

Bu vektör, resmin fonksiyonel bir özetidir ve bu resimle başka resimler arasında görsel benzerlik ilişkisi kurmak amacıyla kullanılır. Böylece aranan bir nesnenin (araştırmada bahsedilen senaryoda insan vücudu) sayısal resimlerin taranmasıyla bulunması mümkün hale gelir. Bu bir görev için Dalal ve Triggs bir ikil sınıflandırıcı (aranılan şeklin test kümesindeki resimlerde olup olmadığı kararını vermek üzere eğitilecek bir yapay öğrenici) önermiştir. HoG adı verilen bu nesne tespiti çerçevesinin uçtan uca (end to end) bir hale gelmesinin sağlanması amacıyla içerisinde bir SVM in, hedeflenen pozitif ve negatif resimlere ait HoG özellik vektörleriyle eğitilip, bir karar mekanizmasının kurulmasında rol alması öngörülmüştür.

Histograms of Oriented Gradients, başka bir yaygın adıyla HoG, nesne tespiti algoritmaları arasında en bilinenlerindedir. Nesne tespiti, hayvan tespiti, insan - yaya tespiti, insan - afetzedede tespiti, biyomedikal analiz gibi bilgisayarla görme uygulamalarında yaygın olarak kullanılmaktadır.

2.2 Artırımlı Basit Özellik Zincirleriyle Hızlı Nesne Tespiti

Viola & Jones Detector ve Boosted Cascade of Simple Features isimleriyle de bilinen bu algoritma da tıpkı HoG gibi, sayısallaştırılmış resimlerin özellik çıkarımı uygulanarak özetlenmesi ve sonrasında bu özelliklerin makine öğrenmesi tekniklerinde kullanılması temeline dayanan bir uçtan uca yazılım çerçevesidir [10]. HoG'dan farklı olarak, çok (~ 60 kat) daha fazla özellik çıkarımı yapılır. "Dikdörtgensel Özellikler" adı verilen bu bileşenler, resmin griölçek (grayscale) hali üzerinde, komşu piksellerin siyah-beyazlık değerlerindeki belirgin farklardan ortaya çıkan paternlerdir. Söz gelimi cepheden çekilmiş bir yüz resmi üzerinde dikdörtgen paternler ortaya çıkarılmak istendiğinde, yatay eksen üzerinde uzanan göz bölgesinin, bunun üzerinde yer alan alın bölgesinden ve altındaki ağız bölgesinden daha karanlık olduğu gözlemi yapıp bu özelliğin bayrak benzeri üç bölmeli bir dikdörtgene benzetilmesi Viola & Jones tarzı özellik çıkarımının tipik bir örneğidir.

Örnekten de anlaşılacağı üzere Viola & Jones algoritmasını diğerlerinden ayıran temel özelliklerin başında bu basit ve bir manada insan gözüyle rahatlıkla yorumlanabilir sayılabilecek özellik çıkarımı mantığı gelmektedir. Bu yaklaşım, algoritmanın yayımlandığı tarihte çağdaşlarının hayli ötesinde bir isabet oranına ulaşmasını sağlamıştır. Fakat algoritmanın bilgisayarla görme literatürüne yaptığı katkılar bununla sınırlı değildir.

Yukarıda anlatılmış olan dikdörtgensel özellikler belli koordinatlar arasında kalan piksel değerlerinin toplanıp, muhtelif bölgeler arasındaki değer toplamlarının farklarının hesaplanmasıyla elde edilir. Viola ve Jones makalelerinde, bu toplamları hesaplamının son derece verimli bir yöntemi olan “İntegral Resim” adını verdikleri bir ön işlem çıktısını tanıtmışlardır. Resmin sol üst köşesi başlangıç olarak kabul edilerek, her piksele kadarki satır ve sütunlardaki değerlerin kümülatif toplamlarının o piksele atanmasıyla elde edilen resme “İntegral Resim” adı verilir. Bu resmin kullanılması sonucunda bir görüntünün her hangi bir bölgesinin değerleri toplamının hesaplanmasının maliyeti, sabit sayıda indis bazlı dizi erişimi maliyetine indirgenmiş olacaktır.

Algoritmanın çalıştırılması sonucunda ortaya çıkan özellikler, AdaBoost algoritmasına benzer bir yaklaşımla, bütün özellikler kümesinin alt kümelerinin eğitim kümesi olarak ayrı ayrı belirlenmesi suretiyle öğrencilerin eğitiminde kullanılacaktır. Bu noktada yeni bir problem ortaya çıkmaktadır, o da bu nesne tesbit çerçevesinin çalışma zamanı yönünden getirmiş olduğu kazanımların, yüksek özellik adedi sebebiyle ortadan kalkması tehlikesidir. Bu sorunun çözümü olarak teklif edilmiş yaklaşım kısaca şöyle özetlenebilir: Öğrencilerin uç uca ekleneceği bir sistemde, her bir katmanın olumlu olarak işaretlediği örneğin bir alt katmana ulaşabildiği, bütün katmanların olumlu olarak işaretlediği örneğin işaretinin kesinleştiği dejenare bir karar ağacı... Yine bu yaklaşıma göre ağaç inşa edilirken, aşağı katmanlara inildikçe değerlendirmeye alınan özelliklerin adedi arttığı gibi bu özelliklerin zorluğu, yani olumsuz sonuç döndürme oranları da o nisbette artacak şekilde bir özellik seçimi yapılır. Bütün bu işlemler sonucu oluşturulmuş olan sistem, bir tesbit ediciler zinciridir (*a boosted cascade of detectors*).

Literatürde Viola & Jones adıyla da bilinen bu algoritma, yayımlandığı yıl itibarıyla hız olarak bilinen algoritmaların çoğundan daha etkili olmasının yanı sıra başarı olarak da

önde gelen bir yöntem olmuştur. Nesne tesbiti genelinde uygulamaları vardır ve özellikle yüz tesbitindeki yüksek başarısı sebebiyle yaygın olarak kullanılmıştır.

2.3 Değişebilen Parçalı Modeller

Değişebilen Parçalı Modeller algoritması, sınıf içi varyasyonu yüksek nesnelerin tespitinde rijit modellerin başarısızlığını gidermek üzere ortaya çıkmıştır [39]. Özellikle insan ve diğer canlı varlıkların fiziksel yapıları gereği yüksek bir eklemliliğe sahip oluşları, yekpare (monolithic) modellerin tespit oranının aynı hedef nesneye ait farklı pozlara sahip resimlerde bile düşük oluşu, hedef nesnelerin yapılarındaki değişimleri tolare edebilen modellerin kıymetinin artmasına sebep olmuştur. Makalede, nesne tesbiti için çalışma gruplarının asıl hedefinin gramer bazlı bir detektör tasarlamak olduğunu anlatan bilim adamları, bu zor hedefe mevcut algoritmaları adım adım ileriye götürerek ulaşmak istediklerini belirtmişler ve nihayetinde bu hedefe ulaşılmıştır[18].

Bu insan tesbit çerçevesinde, özellik çıkarımı görevi Dalal & Triggs in algoritması olan HoG tarafından üstlenilmektedir. Buradaki uygulama farkı, özellik çıkarımı yapan algoritmanın çıktılarının ele alınışında ortaya çıkmaktadır. HoG un tanıtıldığı makalede hedef objeye bir bütün halinde yaklaşıırken, DPM, resme, alt parçalarının oluşturduğu hiyerarşinin desteklenebilmesi amacıyla, bir HoG piramidi ile yaklaşır. Bu piramit farklı artan çözünürlük değerlerine sahip birden fazla HoG vektörünün (tipik olarak 40 adet) bir araya getirilmesiyle oluşturulur. Bu HoG piramidi üzerinde bir adet kök filtresi ve çok alt filtrelerden oluşan değişebilen parçalı modelle eşleştirilmeye çalışılır ve yeterince yüksek skora ulaşan kayan pencere (sliding window) konumu, bir tespit (detection) olarak adlandırılır.

DPM in skor mekanizması temel olarak HoG' a dayanır. Getirdiği temel yenilik ise Dalal ve Triggs'in teklif ettiği metodun aksine birden fazla filtre kullanılmasını öngörmesidir. DPM algoritmasının resimleri işleme prensibi, HoG dakine benzer olarak, filtreleri temel almaktadır. Buna rağmen DPM, birçok noktada HoG dan ayrılmaktadır. Herşeyden önce DPM de nesne tesbitleri tek bir HoG ile değil, HoG piramitleri sayesinde yapılmaktadır. Dolayısıyla tesbitlerin konumlarından ötürü kazandıkları

skorlar hesaplanırken, x ve y eksenine, puanı kazandıran özelliklerin dahil olduğu HoG un piramitteki konumu olan l (level) yani seviye eksenini de katılmaktadır.

Başka bir deyişle filtreler, bir HoG piramidinin alt pencerelerinin ağırlıklarını temsil eden dikdörtgensel şablonlardır. w ye h büyüklüğündeki bir F filtresinin skor u $w \times h \times 9 \times 4$ adet ağırlıkla ifade edilir ki bu değerlerden 9, açılış sepetlerinin adedine, 4 ise 2×2 lik hücre boyutuna karşılık gelmektedir.

H nin bir HoG piramidi, p nin de (x, y, l) şeklinde ifade edilen yatay-düşey eksenler ve piramid derinliğine bağlı konum bilgilerinden oluşan bir nokta olduğunu düşünelim. Bu durumda $\varphi(H, p, w, h)$ vektörü, H nin sol üst köşesi p de olan bir alt penceresine ait HoG değerlerinin birbirine ulanmasıyla (concatenation) oluşmaktadır.

Bu yüzden bu tesbit penceresi (Detection Window) üzerindeki F filtresinin skoru $F \cdot \varphi(H, p, w, h)$, veya kısaca $F \cdot \varphi(H, p)$ olacaktır. Buna karşılık DPM, şablonu sıklıkla karşılaşılan alt bölgelere ayırır ve bu alt bölgelerin her birine karşılık gelen ayrı birer model üretir. Mixture of Models adı verilen büyük bir modelin en önemli işlevi, sistemi HoG için daha önce önerilmiş olan şablonların rijitliğinden nisbeten kurtarıyor oluşudur. Zira DPM in skor ($s(\lambda)$) hesaplaması, öncelikle bir kök filtrenin (f_0) çalıştırılıp bir skor üretilmesine, daha sonra aynı hipotezine (λ) ait parça filtrelerin ($f_{1...n}$) tesbitlerinin şablondaki orjinal konumlarına $\{p_{1...n}\}$ göre yer değiştirmeleri ($d_{1...n}$) üzerinden hesaplanan bir ceza puanının bu ilk skordan çıkarılmasına dayanmaktadır. Anın da HoG özelliğini temsil ettiği aşağıdaki formüllerle bu işlem kısaca özetlenebilir :

$$s(\lambda) = f_0^T \varphi_\alpha(p_0, I) + \sum_{t=1}^n f_t^T \varphi_\alpha(p_t, I) - d_t^T \varphi_d(p_t, p_0)$$

n adet parçadan oluşan bir hipotezine atanacak nihai skoru belirlemek için hesaplanan ve denklem (1) de görülen p_t değeri, öncelikle parçaya ait filtrenin pencere içerisinde elde ettiği en büyük skorun, daha sonra bu skora ait tesbitin hipotezine olan uzaklığını kullanarak deformasyon maliyetini hesaplayan, ayrılabilir kvadratik fonksiyon (d) sayesinde bulunur :

$$p_t = \underset{p}{\operatorname{argmax}} f_t^T \varphi_\alpha(p_t, I) - d_t^T \varphi_d(p_t, p_0)$$

Nesne tesbit yaklaşımlarını geliştirip onları önceki uygulamalardan daha başarılı kılma süreci yorucudur, çünkü DPM, resme bir bütün olarak yaklaşan çoğu algoritmanın aksine hedefin küçük parçalarına kadar analiz edilmesi daha sofistike bir sistemdir fakat denemeler, bu yaklaşımın çalışma prensibi olarak kendinden çok daha naif yöntemlerle çalışan sistemler tarafından mağlup edildiğini göstermektedir. Yazarlar, bu durumu birkaç sebebe bağlamaktadır. Bunlardan en önemlisi, iyi bir sistem ortaya çıkarmanın tek başına yeterli olamayışı, bu sistemi eğitmenin başarılı ve verimli bir yönteminin bulunmamasıdır.

Eğitim sürecine ilişkin bu açığın sebebi ise yapıtaşları için gerekli ground truth verisinin olmayışı, bir örnekle açıklayacak olursak, insan tanıma görevi için üretilmiş bir eğitim setinde insanın pozisyonuyla ilgili bir çevre kutu (bounding box) mevcutken, insan silüetinin parçaları olan üyeler ve gövde gibi kısımların ayrı ayrı tanımlarına yer verilmemiştir. Bu açıdan yaklaşıldığında görülecektir ki eğitim için geliştirilecek algoritma sadece doğru veriye ait alt bölümleri “keşfetmek” durumundadır. Bu keşif, daha önce de bahsedilmiş olduğu üzere, non-convex bir problemdir ve basitleştirilmesi gerekmektedir. Bu basitleştirme işlemi, makalede “Gizli Değişkenler” (Latent Variables) denilen değişkenlerin devreye sokulmasıyla olur. Bu işlem, her bir örnek işlenmesinde, o örneğin bulundurması öngörülen yapıtaşı adedince gizli değişkenin SVM in eğitimine dahil edilmesi ve agresif bir arama ile bu verilerin bulunmasıdır. Böylece kullanılan SVM bir Latent SVM, kısa adıyla LSVM haline gelmiş olur.

Gizli değişkenlerin uygulamaya eklenmesi sayesinde işlem non-convex bir problem olmaktan çıkar ve semi-convex bir hale gelir. Böylece başarılı olmanın yanında hızlı da çalışabilecek bir algoritmaya sahip olunur. Bunlara ek olarak Felzenszwalb, bu semi-convex yapının, eğitim kümesinde nesneye ait alt parçaların konumlarının da sağlanması durumunda convex hale geleceğini belirtmiştir.

Özellikle sayısal resimlerden insan tespiti hususunda kendinden önce geliştirilmiş çalışmaların tamamından daha yüksek başarıya sahip olması DPM in kısa sürede büyük üne kavuşmasını sağlamıştır. Fakat bunun yanında HOG, Boosted Cascade of Simple Features yahut SIFT gibi yekpare modeller üretilen basit tarama işlemleri yapan

algoritmalarından çok daha yavaş çalışıyor oluşu [36] bilhassa eş zamanlı taramaya gereksinim arttığı günümüzde önemli bir eksiklik olarak öne çıkmaktadır.



DPM'in AFETZEDE TESPİTİNDE KULLANIMI ve ROS

Bu bölüm, DPM algoritmasının arama kurtarma operasyonlarında kullanılmak üzere tasarlanmış olan bir robota adapte edilmesi sürecini açıklamaktadır.

3.1 Robot İşletim Sistemi (Robot Operating System)

İngiliz Willow Garage grubu tarafından geliştirilmiş olan Robot Operating System (ROS), robotik uygulamalarında kullanılan yazılım çözümlerinin mümkün olduğunca düşük bağlaşımla bir araya getirilmelerini sağlayan bir yazılım çerçevesidir [32]. C++ ve Python dillerinde geliştirilen program parçalarının (node) özel bir derleyiciyle (catkin) sisteme dahil edilerek, ileti yayınlama (publishing), iletileri dinleme (subscribing) gibi iç protokoller (ros message) sayesinde birbirlerini etkilemeleri, bu çerçeve kapsamında mümkün kılınmış özelliklerin bazılarıdır.

ROS, robotik uygulamalarının ticari olmayan ve deneysel yapılarının, insanlığın bu alanda teknolojinin birikimselliği ilkesinden faydalanmasını engellediğinin görülmesi üzerine, dünyanın muhtelif bölgelerindeki yazılım geliştiricilerin ve bilim insanlarının açık kaynak kodlu olan bu sisteme katkıda bulunabileceği şekilde tasarlanmıştır. Böylece dünyanın neresinde olursa olsun robotik üzerine çalışma yürüten araştırmacılar, bu sistemi sahiplenip gelişmesine katkıda bulunmaya çalışabileceklerdir. ROS'un gelişiminde büyük etki sahibi olan bu vizyona ulaşılması noktasındaki büyük gayreti, kurulduğu günden bu yana artan kullanım yaygınlığıyla göstermektedir [24].

3.2 ROS un İç Yapısı ve Çalışma Dinamiği

3.2.1 ROS Kurulumu, Paketler ve Derlenişleri

ROS üzerinde geliştirilmelerin yapılacağı bir bilgisayarda var olması gereken yapılardan ilki bu sisteme ait temel bileşenlerin bulunduğu, belirli bir ROS sürümüne ait olan ana ROS yazılımını (ros-<sürüm>-ros-base) edinmektedir. Bunun üzerine ROS workspace adı verilen klasörün oluşturulmasıyla ROS un dosya yapısının en temel parçası oluşturulmuş olur. ROS yazılım parçalarının atomik bileşenlerine ROS paketi (package) adı verilir. Bir bilgisayar üzerine kayıtlı paketlerin derlenmesi işlemi ROS workspace üzerinden “catkin” sistemi ile yürütülür ve bunun için “catkin_make” komutu kullanılır.

Bir ROS paketinin çalışabilir hale getirilebilmesi için kaynak kodların dışında birtakım dosyalara ihtiyaç duyulur. Bu dosyalardan mutlaka bulunması gerekenlerden biri “package.xml” iken öteki “CmakeLists.txt” dir.

“package.xml” dosyası, ROS paketine ait bir manifesto belgesi olarak değerlendirilebilir. Bu dosyada, paketin ortaya çıkarılmasında emeği geçen kimseler ve iletişim bilgilerinden, paketin çalışabilmesi için muhtaç olduğu (depend) diğer paketlerin isimlerine kadar birçok bilgi bulunmaktadır. Diğer bir olmazsa olmaz dosya olan “CmakeLists.txt” dosyası ise adından da anlaşılacağı gibi, bir derleme protokolü olan Cmake in sayesinde çalıştığı aynı isimli dosyaya benzer bir görev yürütmektedir ve benzer bir yapıya sahiptir. Bir ROS paketinin derlenmesinde izlenecek adımların ve – varsa- derlenecek diğer paketlerin veya açıklanması gereken CPU bayraklarının tümü “CmakeList.txt” dosyasında yazılı durumda olmaktadır. Bilhassa bu dosyanın iç yapısı bir ROS paketinin istenilen şekilde derlenmesi ve çalışması açısından çok önemlidir, zira ek paketlerin derlenme sırası yahut doğru CPU bayraklarının seçimi bir ROS paketinin çalışma biçimini büyük ölçüde etkileyebilmekte, hatta derlenmemesine dahi sebep olabilmektedir.

3.2.2 ROS Dğümleri ve Çalışma Zamanı İşlemleri

Daha önce de belirtildiği gibi, ROS tabanlı sistemlerin çalışan en küçük parçalarına ROS düğümü (ros node) adı verilmektedir. Bir paket üzerinde bir yahut birden fazla ROS

düğümü tanımlı olabilmektedir. Her bir düğümün kendine ait bir çalışma ismi vardır ve bu isim, “roscore” komutunun verilmiş ve dolayısıyla ROS işletim sisteminin çekirdek yapısının (Master) aktif olduğu bir oturumda “roscun <düğüm adı>” komutu ile birlikte komut satırına girildiğinde işbu düğüm aktif hale getirilmiş olur.

ROS sisteminin atomik çalışma birimi olan düğümler, birbirleriyle iletiler sayesinde iletişim kurabilmektedirler. Böylece söz gelimi yön bulmada, haritalamada, konum saptama veya afetzede konumunun tespitinde kullanılacak GPS verileri için ayrı birer kod parçası çalıştırılmak zorunda kalmayacaktır. Üstelik GPS verisini istenilen formattaki bir ileti halinde yayınlamak için ROS düğümü daima erişilebilir kaldığı gibi, yayınladığı ileti de, paket tanımlarında ve kaynak kodlarında gerekli modifikasyonlar yapılmış olan düğümler tarafından dinlenebilecektir.

Bir “ROS ileti (Message)” si, iletilerin sistem içerisinde taşınmasını sağlayan yollar olan “ROS konu” (Topic) larına bağlı olarak yayınlanır. ROS iletilerinin yayımlandıkları format, yayımlandıkları düğümün ait olduğu paketin içerisinde bulunan .msg dosyalarında belirtilir. Bir ROS iletilerinin yayınlanabilmesi için bu mesajın yayıncı düğümünün aktif olması gerekmektedir. Bu düğümüne “yayıncı” (Publisher) iletiyi dinleyen düğümüne ise “üye” (Subscriber) adı verilmektedir. Tıpkı yayıncı düğüm iletileri yayınlayabilmesi için olduğu gibi, üye düğümünün bir iletiyi dinleyebilmesi için sağlaması gereken şartlar vardır. Bunlar, konuya üye olmak, çalışma mekanizması içerisinde asenkron olarak dinleme işlemini sürdüren ve bir ileti alındığında çalışan geri çağırım (callback) mekanizmaları bulundurmaları olarak özetlenebilir.

3.3 DPM'in Gerçek Zamanlı Kullanımı Sorunu ve Tercih Edilen Çözüm

2.9 GHz CPU ya sahip bir bilgisayar üzerinde 640 x 480 piksellik bir resmi yaklaşık 2.5 saniye sürede tarayan MATLAB kodunun yerine kullanılacak yeni programın, aynı görevi bir saniyenin altında bir süreyle yapabilmesi, gerçek zamanlı bir uygulamanın programlanabilmesi için aranılacak gerek şartlardan biriydi. Bu şartların sağlanması amacıyla yapılan araştırmaların ilk ayağında OpenCV kütüphanesi bünyesindeki “Latent SVM” [40] adındaki API nin, DPM in OpenCV üzerindeki uygulaması olduğu görülmüştür. Fakat bu gerçeğin de, benzer koşullarda tek thread üzerinde çalıştığı ve aynı büyüklüklerdeki bir resmi 2 saniyenin altında dahi işleyemiyor oluşu,

bu arayışın devam etmesine yol açmıştır. Bunun ardından ortaya çıkan seçenekler arasında libCCV [41] gibi bu süreyi 1.7 saniyeye düşürdüğü belirtilen ve FFLD [19] gibi aynı görevi 0.2 saniyede gerçekleştirebilecek kadar hızlı algoritmalar mevcuttu. DPM in doğruluk açısından sağladığı kazançtan ödün vermemeyi önceleyerek yapılan testler, FFLD (Fast Fourier Linear Detector) nin, Charles Dubout' nun ilgili makalesindeki hıza erişebildiğini, dolayısıyla gerçekleşmek istenen afettede tespit sisteminde kullanımının faydalı olacağını gösterdi ve çalışmada yeni bir döneme girilmiş oldu.

3.3.1 Ana Hatlarıyla FFLD

FFLD algoritması, yüksek derecedeki hızının yanında doğruluk açısından da DPM den geri kalmamaktadır. Bu hıza ulaşılmasında, DPM algoritmasının darboğazı olan HoG piramitleri üzerinde gezdirilen kayan pencerelerin ortaya çıkardığı maliyeti en düşük seviyeye indirecek yaklaşımın keşfedilmesi etkili olmuştur. Bu yaklaşım, kısaca şöyle özetlenebilir:

HoG algoritması ile nesne tespitinde bir filtrenin skoru, resim üzerinde eşleştiği piksellerin özellikleriyle kendi özelliklerinin uyuşmalarına bağlı olarak elde edilmektedir. K nin toplam özellik sayısını, x in $M \times N$ lik bir resmi ve y nin $P \times Q$ luk bir filtreyi temsil ettiği bir denklemde z nin de bahsi geçen filtrenin resme uygulanışı sonucu ortaya çıkan skora karşılık geldiği varsayılacak olursa; bu durumda bütün kayan pencere taramalarının adedinin denklemi aşağıdaki gibi olur.

$$z_{ij} = \sum_{k=0}^{K-1} \sum_{p=0}^{P-1} \sum_{q=0}^{Q-1} x_{i+p,j+q}^k y_{pq}^k$$

Charles Dubout ve diğerleri [19], bir lineer filtrenin resmin bütün noktalarıyla ayrı ayrı beslendiğinde ortaya çıkaracağı cevapları, resim ile lineer filtrenin tersinin konvolüsyonlarının toplamıyla da bulunabileceğini kaydetmişler ki bu da yukarıda belirtilen işlemin

$$z = \sum_{k=0}^{k-1} x^k \bar{y}^k$$

Şeklinde kısaltılabilmesi anlamına gelmektedir. Böylesi bir algoritma, zaman maliyeti açısından değerlendirildiğinde ortaya, $M \times N$ lik bir resim ve $P \times Q$ luk bir filtre için $O(MNPQ)$ luk en kötü durum maliyeti çıkar.

Bu maliyet, bir de K adet özellik ve L adet filtre için düşünüldüğünde ortaya çıkacak değer $O(KLMNPQ)$ kadar büyük olacaktır. Fourier dönüşümünün sağladığı katkıyla bir sinyal olarak algılanan resim frekans düzlemine taşınmakta ve bu sayede konvolüsyon maliyeti $O(MN \log MN)$ olmaktadır. Filtreler bu işlemle hiçbir doğruluk yahut hassasiyet kaybına uğramamaktadır.

3.3.2 OpenCV ve FFLD

Tez çalışmaları kapsamında yapılması uygun görülmüş olan testlerden biri de, FFLD nin gerçek zamanlı kullanımının OpenCV kütüphanesi yardımıyla denenmesi olmuştur. Bu deney kapsamında bilgisayarın Web Kamerasından elde edilen anlık resimlerin FFLD uygulamasına aktarılması için bir arayüz yazılmış ve bu arayüz, kamera verisi üzerinde insan tespiti yapabilecek duruma getirilmiştir.

Çizelge 1 de de görüleceği gibi FFLD, tez çalışması kapsamında hazırlanan robocup veri kümesi üzerinde yapılan testlerde başarımlı olarak da hız olarak da voc-release5 uygulamasını geride bırakmıştır.

Çizelge3.1YTU Olasılıksal Robotik Laboratuvarı Robocup Verileri Üzerinde Yapılan Testlerin Sonuçları

Uygulama Adı / Ölçüt	Doğru Olumlu Tesbitler	Doğru Olumsuz Tesbitler	Süre (Çok İşlemci)	Süre (Tek İşlemci)
voc - 5	41/134 (%30.5)	136/136	2.461 sn	8.743 sn
ffldv2	53/134 (%39.5)	136/136	0.204 sn	0.254 sn

Robocup Rescue League [42] adı verilen, her yıl uluslararası olarak düzenlenen ve robot teknolojilerini konu alan yarışmaların standartlarına uygun olarak düzenlenmiş YTU gömülü sistemler laboratuvarında elde edilen test ve eğitim verileri, voc-release5 ve

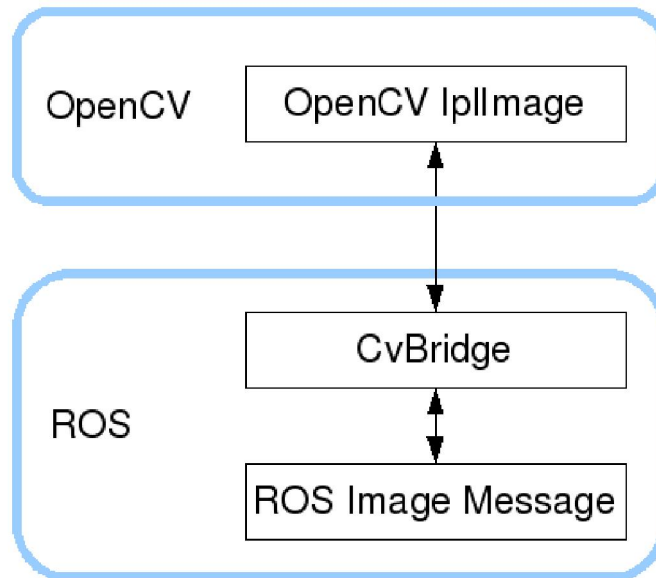
ffldv2 uygulamalarının eğitilmesinde ve testlerinde kullanılmıştır. Kullanılmış olan eğitim ve test verilerinin tamamı 640x480 piksellik üç kanallı RGB resimlerden oluşmaktadır. Eğitim verilerinin kümesi 142 resimden, test verilerinin kümesi ise 134 ü pozitif etiketli olmak üzere toplam 270 resimden oluşmaktadır.

3.4 FFLD nin ROS a Adaptasyonu

Yapılan testler sonucunda, sayısal resim verisinden afettede tespitine dayanan bir mekanizmanın YTU Olasılıksal Robotik Kulübü faaliyetleri çerçevesinde geliştirilmiş olan Arama Kurtarma Robotuna dâhil edilmesinin son basamağına gelinmiş oldu. Bu safhada, arama kurtarma robotunun operasyonunun üzerine inşa edilmiş olduğu ROS sistemi devreye girmektedir. Robotun diğer yazılımsal parçalarında olduğu gibi afettede tespiti modülünde de ROS'a ait mekanizmalardan faydalanılmıştır. Bu mekanizmaların başında "cv_bridge" düğümü gelmektedir.

3.4.1 cv_bridge Düğümü ve Sistemdeki İşlevi

ROS hiyerarşisinde "vision_opencv" paketinin altında yer alan "cv_bridge" düğümü, OpenCV kütüphanesinin kameradan aldığı anlık sayısal RGB resim verisini ROS'a ait resim iletisi formatına dönüştürüp yayınlamakla yükümlüdür.



Şekil 3.1cv_bridge Arayüz Şeması[38]

Robot üzerinde bulunan Kinect cihazının sağladığı resim verisi, OpenCV marifetiyle alınıp cv_bridge düğümü ile bir ROS iletisine dönüştürülmektedir.

3.4.2 Tespit Edici Düğüm “watcher”

FFLD nin ROS ile bütünleşmesi kapsamında tarafımızdan tasarlanan tespit edici düğüm, hem yayıncı hem de üye olarak çalışmaktadır ve C++ ile kodlanmıştır. Bu düğümün vazifesi kısaca “/camera/rgb/image_color” ve “/camera/depth/image” konularına üye olmak ve “std_msgs::Int32MultiArray” türünde bir mesaj yayınlamaktır. Üye olunan bu iletilerden ikisi de robot üzerinde bulunan Kinect cihazından yayınlanan verileri taşımaktadır. Bu iletilerden ilki cihazdan gelen renk verisini, ikincisiyse aynı sahneye ait derinlik haritası verisini taşımaktadır.

“watcher” düğümü, bu verileri FFLD uygulamasına aktarır ve bu uygulamadan dönen nesne tespiti sonuçlarını bir tamsayı dizisi olarak yayınlar. Bu tamsayı dizisinin elemanları sırayla tespit skoru, tespitin robota göre x eksenindeki konumu, tespitin robota göre y eksenindeki konumu ve tespitin robota göre z (derinlik) eksenindeki konumu olmaktadır. Üye olunan derinlik verisinin düğümün çalışmasında efektif olan kısmı budur.

Mümkün kılınan bu yayınlamayla birlikte FFLD nin sisteme bütünleşmesi tamamlanmıştır. Bunun yanında FFLD nin istenilen hızlarda çalışabilmesi için gerekli CPU bayrakları ve derleme işleminin sorunsuz gerçekleşebilmesi için “Eigen” kütüphanesinin uygun sürümünün önceden derlenmesi için gerekli komutlar da “CmakeList.txt” dosyasında belirtilmiştir. Bütün bu uygulamaların ardından “watcher” düğümü saniyede 5 defa tarama yapabilecek bir kapasiteyle işleme hazır hale gelmiştir.

DPM ALGORİTMASININ EĞİTİM SÜRECİNİN İYİLEŞTİRİLMESİNDE YAPAY VERİ ARTIRIMI

4.1 Veri Kümelerinin Yapay Yollarla Genişletilmesi

Yapay öğrenicinin eğitiminde kullanılacak örneklerin seçimi ve hazırlanması, makine öğrenmesi bazlı sınıflandırma tekniklerinin hepsi için büyük önem taşıyan bir uygulama basamağıdır. Bilgisayarla görü özelinde düşünüldüğünde bu süreç, eğitim kümesini oluşturmada kullanılacak resimlerin, elde edilmesi, özellik çıkarımına ve sınıflandırıcı eğitime en uygun biçimde etiketlenmesi, test ve eğitim kümesi olmak üzere iki ayrı kümede bir araya getirilmesi gibi alt süreçlerden oluşmaktadır.

Bu veri kümelerine yapay yollarla elde edilmiş pozitif ve negatif etiketli örneklerin ilave edilmesi ise, kimi çalışmalarda bu sürece dâhil olan bir ayrı uygulama olarak göze çarpmaktadır. Etiketleme maliyetinin azaltılması, öğrenci performansının artırımı ve hatta öğrencinin başarılı olarak eğitilebilmesinde bir gerek şart olarak, veri kümelerinin yapay yollarla genişletilmesine ilişkin bilgilere tez çalışmasının bu bölümünde yer verilecektir.

Bölümün devamında bu yöntemlerin uygulanışına ilişkin detaylar aktarılacaktır. Bu detayların ardından gelen son bölümde ise geçmişten günümüze gelen süreçte,

özellikle makine öğrenmesine dayalı şekil tanıma arařtırmalarında veri artırımının farklı bilimsel çalıřmalara has nedenleri ve uygulanıř biçimleri bir araya getirilmiřtir.

4.2 Yapay Veri Artırımının Ardında Yatan Nedenler

Makine öğrenmesi temel alınarak yürütölen bilgisayarla görü çalıřmalarının tarihsel geliřimi boyunca yapay veri artırımı, yapılan çalıřmaların konusuna ve niteliğine ve donanımsal ekipmanların sađladıđı imkânlarla bađlı olarak birçok çalıřmada farklı oranlarda kullanılmıřtır. Bunlardan bazıları řöyle sıralanabilir:

- Eđitimde ihtiyaç duyulan miktarda veri elde edinilmesinde karřılařılan güçlükler. Özellikle geçmiř dönemlerde yapılan çalıřmalar göz önünde bulundurulduđunda, bir resmin sayısallařtırılmasının maliyeti, birçok bilgisayarda yeterli adette yüksek çözünürlüklü resmin saklanmasına yetecek hafıza kapasitesinin bulunmayıřı [2] vb. durumlar yařanan bu güçlüklerle örnekle gösterilebilir.
- İstenilen sonucun üretilmesinde kritik rol oynayacak türde örneklelerin orjinal resimler kümesinde yeterli sıklıkta bulunmayıřı, veri kümesinin yapay artırım ile çeřitlendirilmesini zorunlu kılar. Aksi takdirde bu tarz bir eksiklik, eđitim kümesindeki toplam etiketli örnekle adedi ne olursa olsun eđitim sürecinin noksan kalmasına, sistemin istenilen cevapları üretecek tarzda olgunlařamamasına yol açacaktır [2][27].
- Amaçlanan standartlarda bir sınıflandırma performansına eriřmiř olsa dahi birçok sınıflandırıcının bařarımı, eđitiminde kullanılan örnekle adediyle dođru orantılı olarak önemli biçimde artmaktadır [32]. Yeni örneklelerin yapay yollarla üretimine bařvurulmasının sebeplerinden biri de bu olmaktadır.

4.3 Veri Artırımında Kullanılan Yöntemler

Veri kümesinin pozitif etiketli örnekle sayısının artırılmasına yönelik çalıřmalarda dikkat edilmesi gereken hususlar, aslında veri kümesi oluřturulurken de mevcuttur. Bunlar kısaca, örneğin hedef şekilden çok farklı bir görünüme ve oryantasyona sahip olmaması (Eđer sınıflandırıcı döndürmeden bađımsız çalıřmıyorsa) ve ortaya çıkacak örneklelerin

hali hazırda var olan örneklere yahut birbirlerine haddinden fazla benzememesi (veri kümesinde sınıf içi varyasyonun sağlanması) olarak ifade edilebilir.

Verilerin yapay olarak artırılmasına yönelik işlemler, resimler ve etiketler üzerinde yapılan değişiklikler olmak üzere birbirinin tamamlayıcısı olan iki ayrı kısımda gerçekleştirilir. Geçmişten günümüze kullanılmış yöntemler ana hatlarıyla şöyle açıklanabilir:

4.3.1 Temel Geometrik Transformasyonlar

Dikdörtgen şekle sahip bir etikete sahip olan pozitif etiketli veriler için döndürme (rotation), çevirme (flipping), öteleme (translation) ve ölçekleme (scaling) gibi seçenekler mevcuttur. Döndürme işlemi dışarıda tutulmak üzere, bu işlemlerin uygulanması, transformasyonun önce resme, sonra da etikete uygulanmasıyla mümkün olmaktadır.

4.3.1.1 Döndürme

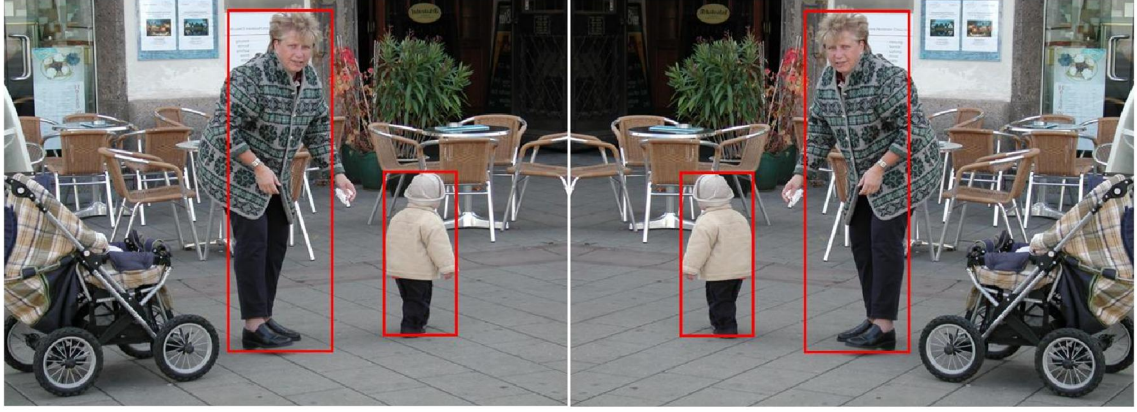
Resimlere uygulanan rotasyon işleminin ardından pozitif etiket olduğu gibi kalır [25][30].



Şekil 4.1 Pozitif örneklere uygulanan döndürmenin etkileri[44]

4.3.1.2 Çevirme

Çevirmeden (flipping) den kasıt resmin merkezinden geçen bir $x = 0$ doğrusuna göre simetriğinin alınmasıdır [25][32].



Şekil 4.2 Pozitif örneklere uygulanan çevirmenin etkileri

4.3.1.3 Ölçekleme

Bu işlem resmin boyutununun belli bir oranda küçültülmesini yahut büyütülmesini kapsar. Aşağıda 1.1 katına çıkarılmış resimle orijinalini yan yana görmek mümkündür [31].



Şekil 4.3 Pozitif örneklere uygulanan ölçekleme işleminin sonuçları

4.3.1.4 Öteleme

Bir resme ait bütün piksellerin kaydırılmasıyla gerçekleştirilen bu işlemin sonucunu aşağıda görmek mümkündür [31].



Şekil 4.4 Pozitif örneklere uygulanan öteleme işleminin sonuçları

4.3.2 Renk Bazlı Transformasyonlar

Piksellerin resim üzerindeki konumlarını değiştirmeye dayanan Geometrik transformasyonların aksine, renk bazlı transformasyonlar, onların sahip oldukları renk değerlerini (Kırmızı, Yeşil ve Mavi) modifiye etmek suretiyle uygulanır.

4.3.2.1 Gauss Gürültüsü Eklenmesi

Gauss gürültüsü eklenmesi, resimlerin sistematik deformasyonunda yaygın olarak kullanılan tekniklerdendir [29][32][14].



Şekil 4.5 Resme Gauss gürültüsü eklenmesinin sonuçları

4.3.2.2 Aydınlik Değerine Müdahale

Aydınlik değeri değiştirilerek resimlerden yeterince farklı yeni yapay resimler üretmek mümkündür [32][14].



Şekil 4.6 Resmin aydınlık (Brightness) değerinin değiştirilmesinin resim üzerindeki sonuçları

4.3.2.3 Karşıtlık Değerine Müdahale

Karşıtlık değerleri değiştirilen resimler de, orjinalleriyle kıyaslandığında makul düzeyde farklılık içerecek kadar deforme olabilirler [32][14].



Şekil 4.7 Resmin karşıtlık (Kontrast) değerinin değiştirilmesinin sonuçları

4.3.3 Etiket Üzerinden Uygulanan Transformasyonlar

Bu tür dönüşümler, adından da anlaşılacağı gibi etiketi oluşturan dikdörtgenin konumu ve boyutu üzerinde yapılan değişiklikleri içerir. Öğrenme algoritması, uygulanan bu dönüşüm sonucunda resmin pozitif olarak algılanan kısmını yeni tanım üzerinden tarayacaktır [31].



Şekil 4.8 Etiketlerinin konumlarının değiştirilmesinin resimlere etkisi

4.3.4 Parametrik Bileşke Transformasyonlar

Yapay yollarla üretilmesi gereken veri miktarı çok yüksek olduğunda başvurmanın daha uygun olduğu yöntemler birtakım parametrelere bağlı olarak gerçekleştirilirler ve çoğunlukla bir dizi transformasyonun birleşiminden oluşmaktadırlar. Veri kümesini büyütmenin kompleks fakat bir o kadar da efektif bir yolu olarak, piksellerin konumlarına ve renk değerlerine yapılacak müdahalelerin, rassallık içeren bir dizi matematiksel işlem sonucunda [43][15][35][14] tesbit edilecek olması bu tür transformasyonların birçok farklı resmi yeterince büyük bir varyans ile üretebilmesine olanak sağlamaktadır.

5.1 INRIA İnsan Veri Kümesi

Bu veri kümesi ayakta duran insanların video ve resimler üzerinden tesbiti üzerine yapılan araştırmalarda kullanılmak üzere oluşturulmuştur [44]. INRIA dataseti içerisinde iki ayrı formata sahip resimler bulunmaktadır. Bunların ilki etiket verileri özel dosyalar ile sağlanmış olan pozitif resimlerin formatı iken, ikincisi 64x128 piksellik bir standarda oturtulmuş ve adı geçen diğer formata sahip resimlerden kesilmiş pozitif verilerin formatıdır. Bu veriler, dikey simetrikleriyle birlikte veri kümesine eklenmiştir. Bu iki formatla oluşturulan resim grubuna da orjinal negatif veriler eşlik etmektedir.

INRIA insan veri kümesi, ilk olarak CVPR 2005 organizasyonunda yayınlanan makalelerinde Dalal ve Triggs tarafından kullanılmıştır [11]. Bu resimler üç ayrı kaynaktan alınmışlardır :

- GRAZ 01 veri kümesi; fakat burdan alınan örneklere ait etiketler Dalal ve Triggs tarafından oluşturulmuştur.
- Navneet Dalal'in uzunca bir sürede oluşturmuş olduğu kişisel fotoğraf koleksiyonundan; buradan alınan resimler genellikle 2592 x 1944 piksel e sahiptir, fakat kullanılabilirliklerini artırmak amacıyla bu resimler içlerindeki insanları ortaya çıkarabilmek için kırılmıştır. Veri kümesindeki insan

resimlerinin çoğu asıl fotoğraflar çekilirken kadraja plansızca girmiş art alanda bulunan kişilere aittir. Bu sayede resimlerde bir poz önyargısına yol açacak türden bir elverişsizlik bulunmamaktadır.

- Resimlerin az bir kısmı Google Images programı kullanılarak Web'den alınmıştır.

Resimlere dair bir diğer ayrıntı ise, içlerindeki insan figürlerinin dikey eksendeki boyutlarının 100 pikselden büyük olmasıdır. Ayrıca etiketlerin bir kısmı resimlerin barındırdığı insanlara ait parçaların dışına taşabilmekte yahut onları tamamen kapsamayabilmektedir.



Şekil 5.1 INRIA kümesine ait bir resmin, etiketi üzerinde görselleştirilmiş hali[44]

Bu veri kümesi ayakta duran insanların video ve resimler üzerinden tesbiti üzerine yapılan araştırmalarda kullanılmak üzere oluşturulmuştur. Bu sebeple yalnızca ayakta duran insanların etiketlendiği görülmektedir. INRIA veri kümesine ait örnek bir resim ve ona ait bir etiket Şekil 5.1 ve 5.2 de görülmektedir.

```
# PASCAL Annotation Version 1.00
# PASCAL Annotation Version 1.00
Image filename : "VOC2007/Images/person_and_bike_112.png"
Image size (X x Y x C) : 640 x 480 x 3
Database : "The INRIA Rhône-Alpes Annotated Person Database"
Objects with ground truth : 1 { "PASinriaperson" }
# Note that there might be other objects in the image
# for which ground truth data has not been provided.
# Top left pixel co-ordinates : (0, 0)
# Details for object 1 ("PASinriaperson")
# Center point -- not available in other PASCAL databases -- refers
# to person head center
Original label for object 1 "PASinriaperson" : "UprightPerson"
Center point on object 1 "PASinriaperson" (X, Y) : (339, 60)
Bounding box for object 1 "PASinriaperson" (Xmin, Ymin) - (Xmax, Ymax) : (190, 0) - (381, 370)
```

Şekil 5.2 Şekil 15 teki resme ait olan ve PASCAL VOC Challenge formatına uygun olarak düzenlenmiş bir etiket.

5.1.1 DPM’de INRIA

Veri kümesi, DPM algoritmasında kullanıma uygun hale getirilmek için fonksiyonel olarak iki temel parçaya ayrılmıştır. Bunlara eğitim ve test kümesi adı verilmiştir. 1832 resimden oluşan Eğitim kümesinde 1218 negatif resmin yanında 614 pozitif resim yer almıştır. Fakat bu pozitif resimler toplam 1237 pozitif etiket barındırmaktadır.

5.2 Deneyler

Tahmin algoritmalarının değerlendirilmesinde yaygın olarak kullanılan ortalama doğruluk (average precision) değeri, doğruluk (precision) ve dönüt (recall) değerleri temel alınarak hesaplanır [26]. DPM algoritmasının açık kaynak kodlu olarak yayınlanan programı, Linux tabanlı işletim sistemleri üzerinde koşan ve MATLAB ile C++ dillerinin beraber kullanılmasıyla yazılmış bir koda sahiptir [45].

Bu çalışma dâhilinde yapılan deneylerin tümü Dalal ve Triggs tarafından da tercih edilmiş olan [11] INRIA veri kümesi üzerinde yapılmıştır. Bu veri kümesinin PASCAL VOC Challenge için düzenlenmiş hali, Test ve Eğitim kümesi olmak üzere iki parçaya bölünmüştür [44]. 741 resimden oluşan Test kümesinde 453 negatif, 288 pozitif resim yer almaktadır ve pozitif resimlerin barındırdığı toplam insan adedi 589 dur. Eğitim kümesinde 614’ü pozitif 1832 resim bulunmaktadır. Bu pozitif etiketli veriler, 1247 adet insan figürü barındırmaktadır.

Çalışma dâhilinde INRIA veri kümesi ve Değişebilir Parça Modelleri kullanılarak yapılan deneyler, DPM in eğitiminde farklı yapay veri artırımı tekniklerinin etkilerini ortaya koyma amacı gütmektedir. Bu farklı veri artırımı teknikleri, Şekil 5.3'teki etkilere sahip beş ayrı gruba ayrılmıştır:

- A. Orijinal resimlerin dikey simetrikleri alınarak (çevrilmeleriyle) oluşturulan örnekler.
- B. Orijinal resimlerin sağa ve sola 5 er derece döndürülmesiyle oluşturulmuş örnekler.
- C. Orijinal resimlere döndürme uygulanırken, yalnızca belirli bir büyüklük oranına sahip nesnelerin etiket üzerinde korunduğu resimler.
- D. Orijinal resimlere döndürme uygulanırken, bütün nesnelerin en verimli şekilde döndürülebilmesi için ortalama bir döndürme değeri tespit edilerek döndürülen resimler.
- E. Önce A'daki gibi çevirme sonra D'deki gibi döndürme yöntemi uygulanarak orijinal resimlerden üretilmiş örnekler.

Deneylerin büyük kısmında INRIA veri kümesindeki örneklerin tamamını kullanmak yerine 1/16 lık dilimlerden, yani 38-39 resimden oluşan INRIA alt kümeleri kullanılmıştır. Bu seçimin temel nedenleri olarak, daha küçük bir veri kümesinde, başarımların artışı yahut azalışı gibi faktörlerin daha dramatik biçimde ortaya çıkacak olması, farklı veriler üzerinde aynı tekniklerin tekrar tekrar denenip sonuçların ortalamasını almanın istatistiksel olarak daha güvenilir sonuçlar verecek olması gibi etmenler sunulabilir.

Böylece deneyler sırasında her bir INRIA alt kümesine Çizelge 5.1'de belirtilmiş olan veri kümesi artırımı teknikleri gösterilen kombinasyonlar halinde uygulanmıştır. Eğitim faaliyetleri sonucunda ortaya çıkan temel üç veri Çizelge 5.1'de gösterilmektedir. Bu verilerin aktarımı sayesinde deney sonuçlarına ilişkin şu detaylara erişilebilmektedir:

Eğitim ve Test süreleri sayesinde uygulanan tekniklerin getirdiği operasyonel yük ortaya çıkmaktadır. Ayrıca yine kullanılan veri adedi ile eğitim süresi arasındaki ilişki

Çizelgedeki verilerden yola çıkarak değerlendirilebilmektedir. Başarımın test süresine olan etkisi de yine gözlemlenebilen bağıntılar arasındadır.

İlk deneyde INRIA veri kümesindeki tüm resimler kullanılmıştır. 6 ayrı model eğitilmiş

Çizelge 5.1 Farklı Yapay Veri Artırımı Tekniklerinin DPM Eğitime Etkileri

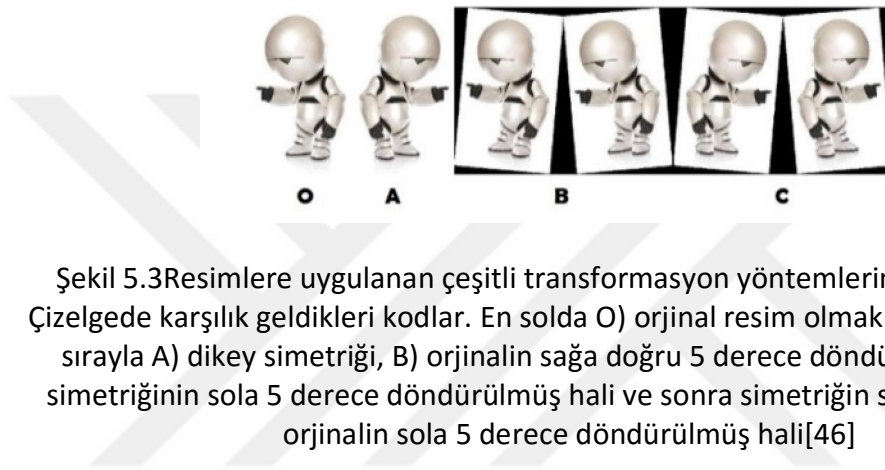
Deney No	#1	#2	#3	#4	#5	#6
Ortalama Doğruluk	0.8512	0.8568	0.8573	0.8501	0.8573	0.8636

ve bu modellerin başarımı aynı test kümesi üzerinde ölçülmüştür. Bu deneme sonuçları Çizelge5.1’de verilmiştir.

Çizelge5.1’deki 1 numaralı sonuç, orijinal INRIA veri kümesiyle yapılan deneye aittir. Orijinal makaledeki [13] sonuçlar aynen elde edilebilmiştir. 2 ve 3 numaralı sonuçlar ise orijinal verilere sırasıyla simetrik eşler ve sağa/sola beşer derece döndürülmüş resimler grubunda yer alan yapay verilerin de eğitim kümesine dâhil edilmesiyle elde edilmişlerdir. Resimlerin ayrı ayrı sağa ve sola döndürülmesiyle oluşturulan veriler, dikey simetriklerin sağladığından daha yüksek bir ortalama doğruluğa erişmektedir. 4 numaralı deneyin veri kümesi ise, orijinal veriler ile yalnız belirli bir büyüklük oranına sahip verilerin etikete dahil edildiği döndürülmüş grubun yapay verileriyle birleştirilmesiyle kurulmuştur. 5 numaralı deneydeki yaklaşım bundan küçük bir fark ile ayrılmaktadır. Bu deneyde resimlerin içerisinde bulunan ve farklı ölçeklere sahip olan nesne etiketlerinin tamamına uyarlanabilecek adaptif bir döndürme oranı tespit edilmeye çalışılmıştır. Bu yaklaşım ile yapılan döndürme işlemleri 4 numaralı deneydekinden daha başarılı sonuçlar vermiş olsa da, doğrudan yapılan döndürme işlemlerini içeren iki numaralı deneyden daha yüksek bir başarı elde etmek için yeterli olmamışlardır. Bir sonraki deney olan 6 numaralı deney ise önceki deneylerde üretilmiş yapay veri çeşitlerinin tümünü içermektedir. Bu şekilde en büyük veri kümesi de 6 numaralı deneyde kullanılmıştır. Bu veri kümesinde alınan sonuçlar deneyler içerisinde en yüksek ortalama doğruluk değerlerine sahip olanlardır. Orijinal veri

kümesiyle elde edilen 0,8512'lik ortalama doğruluk seviyesini (average precision) 0,8636'ya taşımıştır.

Yapılan bu ilk gözlemlerin ardından, gerek istatistiksel olarak daha güvenilir sonuçlar almak, gerekse başarımdaki değişim ile veri kümelerinin boyutu arasındaki bağıntıyı ortaya koymak üzere daha fazla deney yapılması uygun görülmüştür. Bu görüşe binaen, INRIA veri kümesindeki pozitif etiketli resimler grubu 1/16 lık alt kümelere ayrılmıştır ve eğitim kümesindeki pozitif resimleri sadece bu 38-39 resimlik alt kümelere ibaret kılacak yeni eğitim kümeleri oluşturulmuştur.

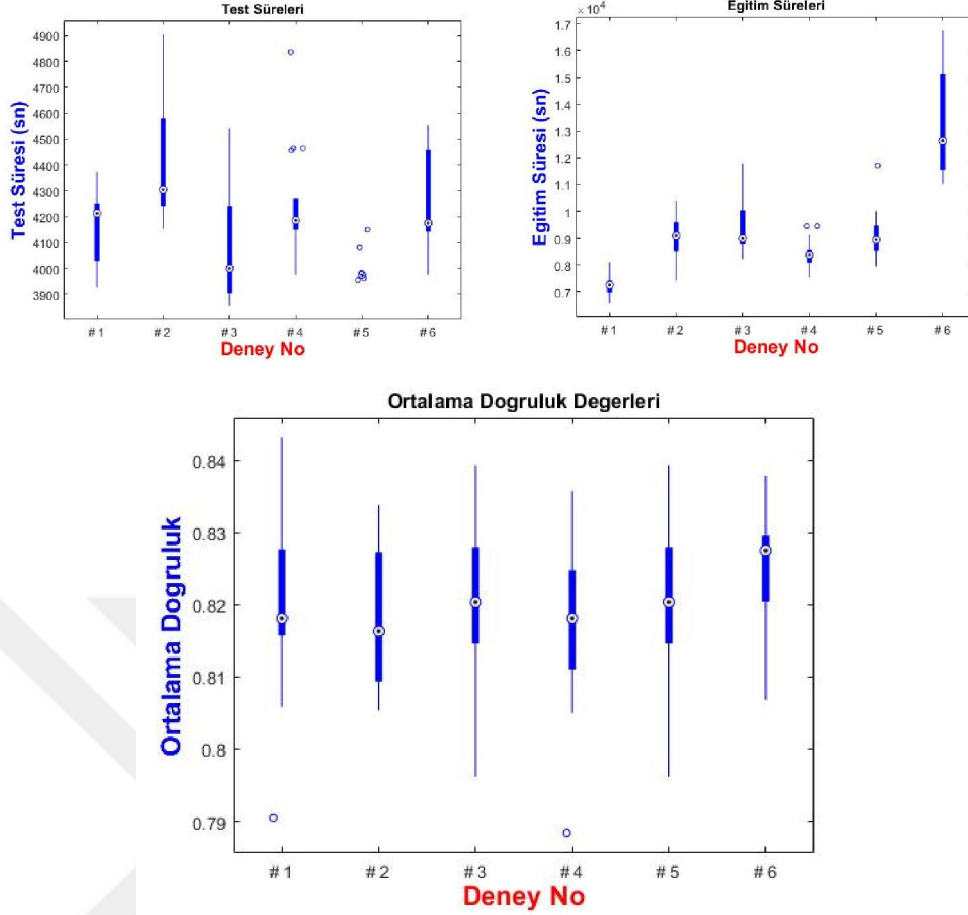


Deneyler sırasında her bir INRIA alt kümesine Çizelge5.1'de belirtilmiş olan veri kümesi artırımı teknikleri birleşimler halinde uygulanmıştır. Bu yöntemlerin resimler üzerindeki etkisi Şekil 5.3'te verilmiştir. Deneme sonuçları Çizelge5.2'de verilmiştir.

Çizelge 5.2INRIA Yaya Veri Kümesinin alt kümeleriyle yapılan deneylerin sonuçları

İsim/Kriter	Test Süresi (sn)	Eğitim Süresi (sn)	Ort. Doğruluk	Pozitif Veri Adedi
1- Orijinal resimler	4156.3 sn	7255.3	0.8171	~40
2- Orijinal resimler ile Simetrik eşler	4409.3 sn	9082.2	0.8182	~80
3- Orijinal resimler ile döndürülmüşler	4076.7 sn	9372.2 sn	0.8202	~120
4- Orijinal resimler ile Seçilerek bazıları alınarak döndürülmüşler	4250 sn	8404.9 sn	0.8178	~120
5- Orijinal resimler ile Adaptif dereceyle döndürülmüşler	4076.7 sn	9372.2	0.8202	~120
6- Orijinal resimler, Simetrik eşler, Çevrilip kaskat döndürülmüşler ve Döndürülmüşler	4240.3 sn	13330	0.8261	~240

Şekil 5.4'teki grafiklerde ise, veri kümesinin genişletilmesinde kullanılan farklı yöntemlerin, eğitimler sonucu üretilen modellerden testlerde elde edilen ortalama doğruluk değerlerine, eğitim sürelerine ve test sürelerine birbirinden farklı etkiler yaptığı görülmektedir. Bu grafikler, her biri 16 ayrı alt veri kümesi kullanılarak eğitilmiş modellerin sahip olduğu ortalama doğruluk değerlerinin, eğitim sürelerinin ve test sürelerinin ortalamalarını (düğme simgesi ile gösterilen), varyanslarını (ince çizgiler) ve testler sonucu sık olarak karşılaşılan değerlerin aralığı (koyu çizgiler) görselleştirmektedir.



Şekil 5.4 Her biri 16 adet DPM eğitim sürecini kapsayan deneyler sonucu ulaşılmış test sürelerinin, eğitim sürelerinin ve ortalama doğruluk değerlerinin ortalamaları (düğme simgesi) ve ağırlıklı değerleri (mavi bölgeler)

Test sürelerinin eğitimde kullanılan pozitif verilerin adedinin artışından etkilenmediği görülmüştür. Eğitim sürelerinin ise eğitimde kullanılan örnek adediyle arttığı ancak bu artışın lineer olması sebebiyle kabul edilebilir olduğu değerlendirilmiştir. Şekil 5.4te görüldüğü üzere, yapay veri artırımı yöntemleri DPM'in başarısını arttırmıştır. Kullanılan yöntemler arasında yine tüm yapay veri artırımı tekniklerinin birlikte kullanılmasının başarıyı en fazla arttırdığı görülmektedir.

BÖLÜM 6

SONUÇ

Deney sonuçlarının özet halinde sunulduğu Çizelge3 te görüldüğü üzere, farklı veri artırımı tekniklerinin başarımları ve süre gibi değerlere farklı etkileri olmuştur. Yine görülmektedir ki, en başarılı genişletilmiş veri kümesi, orijinal resimler , simetrik eşler, çevrilip kaskat döndürülmüş resimler ve döndürülmüş resimlerin birleşerek oluşturduğu küme olmuştur. Bu tekniğin başarısı orijinal veri kümesinden alınan örneklerle karşılaştırıldığında, 1/16 lık dilimlerde de, bütün veri kümesi üzerinde de daha yüksek başarımlar içeren sonuçlar vermektedir.

Çizelgedeki değerlerden çıkan bir diğer sonuç ise, veri sayısı arttıkça eğitim süresi mutlaka artmaktayken, aynı durum test süresi için söz konusu olmamakta, test süresi ile eğitim kümesindeki veri adedi arasında doğrudan bir korelasyon kurulamamaktadır. Ayrıca veri artırımı tekniklerinin başarısı, orijinal veri kümesinin boyutu büyüdüğünde de devam etmektedir.DPM algoritmasının başarımlarının artırılması amacıyla gerçekleştirdiğimiz bu çalışmada, veri kümesinin yapay yollarla genişletilmesinde kullanılan temel geometrik dönüşüm işlemlerinin yalnızca pozitif etiketli verilerde dahi uygulandığında başarımları artırdığı görülmüştür. Farklı dönüşüm işlemlerinin algoritma başarımına olan etkileri karşılaştırılmış ve bu dönüşüm işlemlerinin tümünün bir arada kullanıldığı bir veri kümesi genişletme operasyonunun yöntemlerin tek başına uygulanmasına kıyasla daha yüksek bir başarımlar sağladığı görülmüştür.

INRIA veri kümesi üzerinde gerçekleştirilen deneyler sonucunda, hem verinin tamamı hem de bir alt kümesi kullanıldığında yapay veri eklemenin pozitif etkisi görülmüştür. Kullanılan yöntemler, hem uygulanmasının kolaylığı, hem de resim ve verinin dışında ek bir veriye ihtiyaç doğurmaması nedeniyle başarıyı yüksek modellerin üretilmesinde önemli bir avantaj sağlamaktadır.

Gelecekte DPM üzerinde yapılabilecek çalışmalara bir örnek olarak, farklı veri kümeleri (PASCAL VOC Dataset, ImageNet vb.) ve farklı yapay veri artırımı yöntemlerinin kullanılarak yapılabilecek çalışmalar verilebilir.



KAYNAKLAR

- [1] JähneB., (2000)., Computer Vision and Applications A Guide for Students and Practitioners, 1,Academic Press, San Diego, CA, ABD
- [2] Pomerleau, D. A., (1989). "ALVINN: An Autonomous Land Vehicle In a Neural Network", Advances in neural information processing systems,1 : 305-313
- [3] Pomerleau,D. A., (1993).“Neural Network Perception for Mobile Robot Guidance”, Kluwer Academic Publishers, Norwell, MA, ABD
- [4] Drucker,H. ,Schapire R. veSimard P., (1994). "Improving Performance in Neural Networks Using a Boosting Algorithm", International Journal of Pattern Recognition and Artificial Intelligence, 4 (7) : 42-49
- [5] Cun,Y. , Jackel,L. D., Bottou,L., Cortes,C., Denker,J. S., Drucker,H., Guyon,I., Muller,U. A., Sackinger,E., Simard,P. ve VapnikV. N., (1995).“Learning Algorithms For Classification: A Comparison On Handwritten Digit Recognition”, Neural Networks: TheStatistical Mechanics Perspective, 261-276
- [6] C. Cortes ve V. Vapnik, (1995).“Support-Vector Networks”, Machine Learning 20(3) : 273-297
- [7] Freund,Y., Schapire,R. E.,(1999).“A Short Introduction to Boosting”, Proceedings of the 16th international joint conference on Artificial intelligence,2: 1401-1406
- [8] Li, S. Z.,(1995). Markov Random Field Models in Computer Vision, 1, Springer Japan, Japonya
- [9] Lowe, D. G., (1999).“Object recognition from local scale-invariant features”, The Proceedings of the Seventh IEEE International Conference on Computer Vision, 2:1150-1157

- [10] Viola P., Jones M., (2001). "Rapid Object Detection using a Boosted Cascade of Simple Features", Proceedings of the Conference on Computer Vision and Pattern Recognition (CVPR'01), 1:511-518
- [11] Dalal N., Triggs B., (2005). "Histograms of Oriented Gradients for Human Detection", IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05), 1: 886-893
- [12] Fischler M. A. and Elschlager R. A., (1973). "The Representation and Matching of Pictorial Structures", IEEE Transactions on Computers, 22(1): 67-92
- [13] Felzenszwalb P. F., Girshick R. B., McAllester D. and Ramanan D., (2010). "Object Detection with Discriminatively Trained Part Based Models", IEEE Transactions on Pattern Analysis and Machine Intelligence, 32(9): 1627-1645
- [14] Krizhevsky A., Sutskever, I., Hinton, G. E., (2012). "ImageNet Classification with Deep Convolutional Neural Networks", Advances in Neural Information Processing Systems, 25:1097-1105
- [15] Simard, P. Y., Steinkraus, D., Platt, J. C., (2003). "Best Practices for Convolutional Neural Networks Applied to Visual Document Analysis", Proceedings. Seventh International Conference on Document Analysis and Recognition, 958-963
- [16] He K., Zhang X., Ren S., Sun J., (2015). "Delving Deep into Rectifiers: Surpassing Human-Level Performance on ImageNet Classification", The IEEE International Conference on Computer Vision (ICCV'15), 1026-1034
- [17] Pedersoli, M., Vedaldi, A., González, J., Roca, X., (2015). "A Coarse-to-fine Approach for Fast Deformable Object Detection", Pattern Recognition, 48(5):1844-1853
- [18] Girschick R., (2012). "From Rigid Templates To Grammars: Object Detection With Structured Models", PhD, THE UNIVERSITY OF CHICAGO, 2012, 146
- [19] Charles Dubout and François Fleuret, (2012). "Exact Acceleration of Linear Object Detectors", 12th European Conference on Computer Vision, 301-311
- [20] Mohammad Amin Sadeghi and David Forsyth, (2014). "30Hz Object Detection with DPM V5", 13rd European Conference on Computer Vision, 65-79
- [21] Junjie Y., Zhen L., Longyin W., Stan, L., (2014). "The Fastest Deformable Part Model for Object Detection", IEEE Conference on Computer Vision and Pattern Recognition, 2497-2504
- [22] Yılmaz, B., Balcılar, M., Amasyalı, M. F., Yavuz, S., (2015). "Değişebilir Parça bazlı Modeller ile Kurban Tanıma", 23rd Signal Processing and Communications Applications Conference (SIU), 947 - 950
- [23] Kanada Silahlı Kuvvetleri'nin askeri eğitim kılavuzu, https://web.archive.org/web/20080803015913/http://www.casaraontario.ca/~webmaster1/Manuals/NationalSARmanual_full_english.pdf, 29.04.2016
- [24] Robot Operating System, <http://wiki.ros.org/Metrics>, 29.04.2016

- [25] Sung, K. ve Poggio, T., (1998). "Example-based Learning for View-based Human Face Detection", IEEE Transactions on Pattern Analysis and Machine Intelligence, 20 (1):39-51
- [26] Sim, T., Sukthankar, R., Mullin, M., Baluja, S., (2000). "Memory-based Face Recognition for Visitor Identification", Proceedings of Fourth IEEE International Conference on Automatic Face and Gesture Recognition, 214-220, 2000
- [27] Schölkopf, B., Burges, C., Vapnik, V. (1996). "Incorporating invariances in support vector learning machines", International Conference on Artificial Neural Networks, 47-52
- [28] Beymer, D. and Poggio, T., (1995). "Face Recognition From One Example View", International Conference on Computer Vision, 500-507
- [29] Ke Y., Sukthankar R., (2004). "PCA-SIFT: A More Distinctive Representation for Local Image Descriptors", IEEE Computer Society Conference on Computer Vision (CVPR), 2:506-513
- [30] Ivor, W. T., James, T. K., Cheung, P., (2005). "Core Vector Machines: Fast SVM Training on Very Large Data Sets", The Journal of Machine Learning Research, 6:363-392
- [31] Laptev, I., (2009). "Improving Object Detection with Boosted Histograms", Journal of Image and Vision Computing, 27 (5):535-544
- [32] Howard, A. G., (2013). "Some improvements on deep convolutional neural network based image classification", CoRR, abs/1312.5402
- [33] Cireşan, D. C., Meier, U., Masci, J., Gambardella, L. M. and Schmidhuber, J., (2011). "High-Performance Neural Networks for Visual Object Classification", CoRR, abs/1102.0183,
- [34] Simard, P., Victorri, B., Lecun, Y., and Denker, J., (2003) "Tangent Prop | a formalism for specifying selected invariances in an adaptive network", Advances in Neural Information Processing Systems, 4:895-903
- [35] Yaeger, L., Lyon, R., Webb, B., (1996)., "Effective Training of a Neural Network Character Classifier for Word Recognition", NIPS 9, 807-813
- [36] Andriluka, J., Schnitzspan, L., Meyer, U. et al, (2010). "Vision Based Victim Detection from Unmanned Aerial Vehicles", IROS, 461-483
- [37] Dollar, P., Wojek, C., Schiele, B. ve Perona, P., (2012). "Pedestrian detection: An evaluation of the state of the art", PAMI, 34 (4):743-761
- [38] Robot Operating System, <http://wiki.ros.org/>, 29.04.2016
- [39] Felzenszwalb, P. F., Huttenlocher, D. (2005). Pictorial Structures for Object Recognition, International Journal of Computer Vision, 61(1):55-79
- [40] OpenCV adlı açık kaynak kod kütüphanesi, http://docs.opencv.org/modules/objdetect/doc/latent_svm.html, 29.04.16

- [41] libccv adlı açık kaynak kod kütüphanesi, <http://libccv.org/doc/doc-dpm/>, 29.04.16
- [42] Robocup Rescue adlı uluslararası yarışma, <http://www.robocuprescue.org/>, 29.04.16
- [43] Drucker, H., Schapire, R., Simard, P., Improving Performance in Neural Networks Using a Boosting Algorithm, 1994
- [44] INRIA Yaya Veri Kümesi, <http://pascal.inrialpes.fr/data/human/>, 29.04.16
- [45] DPM algoritmasına ait açık kaynak kod, <https://github.com/rbgirshick/voc-dpm>, 29.04.16
- [46] Otostopçu'nun Galaksi Rehberi filmi, <http://www.imdb.com/title/tt0371724/>, 29.04.16



ÖZGEÇMİŞ

KİŞİSEL BİLGİLER

Adı Soyadı : Bedir YILMAZ
Doğum Tarihi ve Yeri : 23.03.1990 Eminönü
Yabancı Dili : İngilizce, Almanca
E-posta : bediryilmaz@igdir.edu.tr

ÖĞRENİM DURUMU

Derece	Alan	Okul/Üniversite	Mezuniyet Yılı
Y. Lisans	Bilgisayar Mühendisliği	Yıldız Teknik Üniversitesi	-
Lisans	Bilgisayar Mühendisliği	İstanbul Ticaret Üniversitesi	2012
Lise	Sayısal	Van Milli Piyango Anadolu Lisesi	2007

İŞ TECRÜBESİ

Yıl	Firma/Kurum	Görevi
2016	İğdır Üniversitesi	Araştırma Görevlisi
2013- 2014	Bilgeadam Bilişim Teknolojileri	Yazılım Danışmanı
2012	Terralabs Innovative Solutions	Yazılım Geliştirici
2011	Edevis GmbH	Yazılım Geliştirici

YAYINLARI

Bildiri

1. Virtualization of Physical Education for Engineering Laboratory Experiments, Ozpınar, Alper; Yılmaz, Bedir; Kuçukgultekin, Can, Proceedings of the 2nd International Engineering Education Confe;2012, p319
2. Değişebilir Parça bazlı Modeller ile Kurban Tanıma, Bedir Yılmaz; Muhammet Balcılar; Mehmet Fatih Amasyalı; Sırma Yavuz
Signal Processing and Communications Applications Conference (SIU), 2015 23th , vol., no., pp.947-950