

Otonom Araçlarda Fren Kararı Tahmini için Dikkat Mekanizması Geliştirilmesi

Ekrem AKSOY

DOKTORA TEZİ

Elektrik Elektronik Mühendisliği Anabilim Dalı

Aralık 2020

An Attention Mechanism for Brake Decision Prediction in Autonomous Vehicles

Ekrem AKSOY

DOCTORAL DISSERTATION

Department of Electrical Electronics Engineering

December 2020

Otonom Araçlarda Fren Kararı Tahmini için Dikkat Mekanizması Geliştirilmesi

Ekrem AKSOY

Eskişehir Osmangazi Üniversitesi
Fen Bilimleri Enstitüsü
Lisansüstü Yönetmeliği Uyarınca
Elektrik Elektronik Mühendisliği Anabilim Dalı
Telekomünikasyon-Sinyal İşleme Bilim Dalında
DOKTORA TEZİ
Olarak Hazırlanmıştır

Danışman: Doç.Dr. Ahmet YAZICI

Bu Tez, Türkiye Bilimsel ve Teknolojik Araştırma Kurumu'nun (TUBİTAK) 116E731 nolu projesi tarafından desteklenmiştir.

Aralık 2020

ETİK BEYAN

Eskişehir Osmangazi Üniversitesi Fen Bilimleri Enstitüsü tez yazım kılavuzuna göre, Doç.Dr. Ahmet YAZICI danışmanlığında hazırlamış olduğum “**Otonom Araçlarda Fren Kararı Tahmini için Dikkat Mekanizması Geliştirilmesi**” başlıklı tezimin özgün bir çalışma olduğunu; tez çalışmamın tüm aşamalarında bilimsel etik ilke ve kurallara uygun davrandığımı; tezimde verdiğim bilgileri, verileri akademik ve bilimsel etik ilke ve kurallara uygun olarak elde ettiğimi; tez çalışmamda yararlandığım eserlerin tümüne atıf yaptığımı ve kaynak gösterdiğimi ve bilgi, belge ve sonuçları bilimsel etik ilke ve kurallara göre sunduğumu beyan ederim. .../.../20...

Ekrem AKSOY

ÖZET

Otonom sürüş veya İleri Sürüş Destek Sistemleri (İSDS) alanlarında insan sürüşüne yakın yaklaşımlar geliştirilmesi aktif araştırma konularındandır. Bu alandaki mevcut çalışmalar daha çok sürüşü etkileyen yaya, trafik işareti v.b. unsurları birbirinden bağımsız tespit etmeye çalışan sistemler olarak öne çıkmaktadır. Öte yandan, insan sürüşünde görsel algı en önemli karar girdisi olup ortamdaki tüm nesnelere için bütüncül olarak işlenir. İnsan görüşünde dikkat ve görsel algıda belirginlik sinirbilim alanında uzun süredir çalışılan araştırma konularıdır. Sinirbilim alanındaki bu çalışmaların ışığında makine öğrenmesi alanında da belirginlik tahmin edilmesi ve dikkat mekanizması literatüre girmiştir.

Bu tez çalışmasında, otonom sürüş veya İSDS için bilgisayar görüşüyle sürüş görüntüsündeki belirginliklere dikkat ederek fren karar tahmini üreten iki derin yapay sinir ağı modeli önerilmektedir. Önerilen her iki modelde de sürüş görüntüsündeki özellikler VGG-16 evrişimli yapay sinir ağı ile çıkarılmaktadır. Çıkarılan özellikler dikkat mekanizması ile ağırlıklandırıldıktan sonra Evrişimli Uzun Kısa Dönem Hafıza (EUKDH) ağına girdi olarak verilerek belirginlik haritası tahmini üretilmektedir. Son olarak tahmin edilen belirginlikler ikinci bir yapay sinir ağı ile ikili sınıflandırılarak fren yapma kararı tahmin edilmektedir. İlk modelde belirginlik tahmini ile fren kararı tahmin edilmesi eş zamanlı olarak üretilmektedir. Eğitim sürecinde belirginlik haritası tahmini ile fren karar tahminini birlikte kullanan bir kayıp fonksiyonu ile geri yayılım algoritması kullanılarak eğitilmektedir. Bu yaklaşımda model eşzamanlı olarak belirginlik tahminini ve fren karar tahminini en iyilemektedir. İkinci model ise önce sadece belirginlik harita tahmini üretmek için eğitilmektedir. Ardından, sürüş girdi görüntüsü, tahmin edilen belirginlik haritası ile maskelenerek fren karar tahmini üreten yapay sinir ağına girdi olarak verilmektedir. İkinci modelin nihai çıktısı fren karar tahminidir. Önerilen modeller, literatürdeki BDD-A sürüş dikkat veri kümesi ve Otonom Taşıyıcı Araç (OTA) benzetim verisiyle denenmiştir. Deney sonuçlarında belirginlik tahmin edilmesi açısından mevcut en iyi durum yakalanmıştır. Bilgisayar görüşünü bütüncül olarak ele alıp belirginliklere dikkat etmeyi öneren modeller fren tahmini açısından motive edici sonuçlar elde etmiştir.

Anahtar Kelimeler: Bilgisayarla görü, Dikkat mekanizması, Belirginlik tahmini, Otonom araçlar, Fren tahmini.

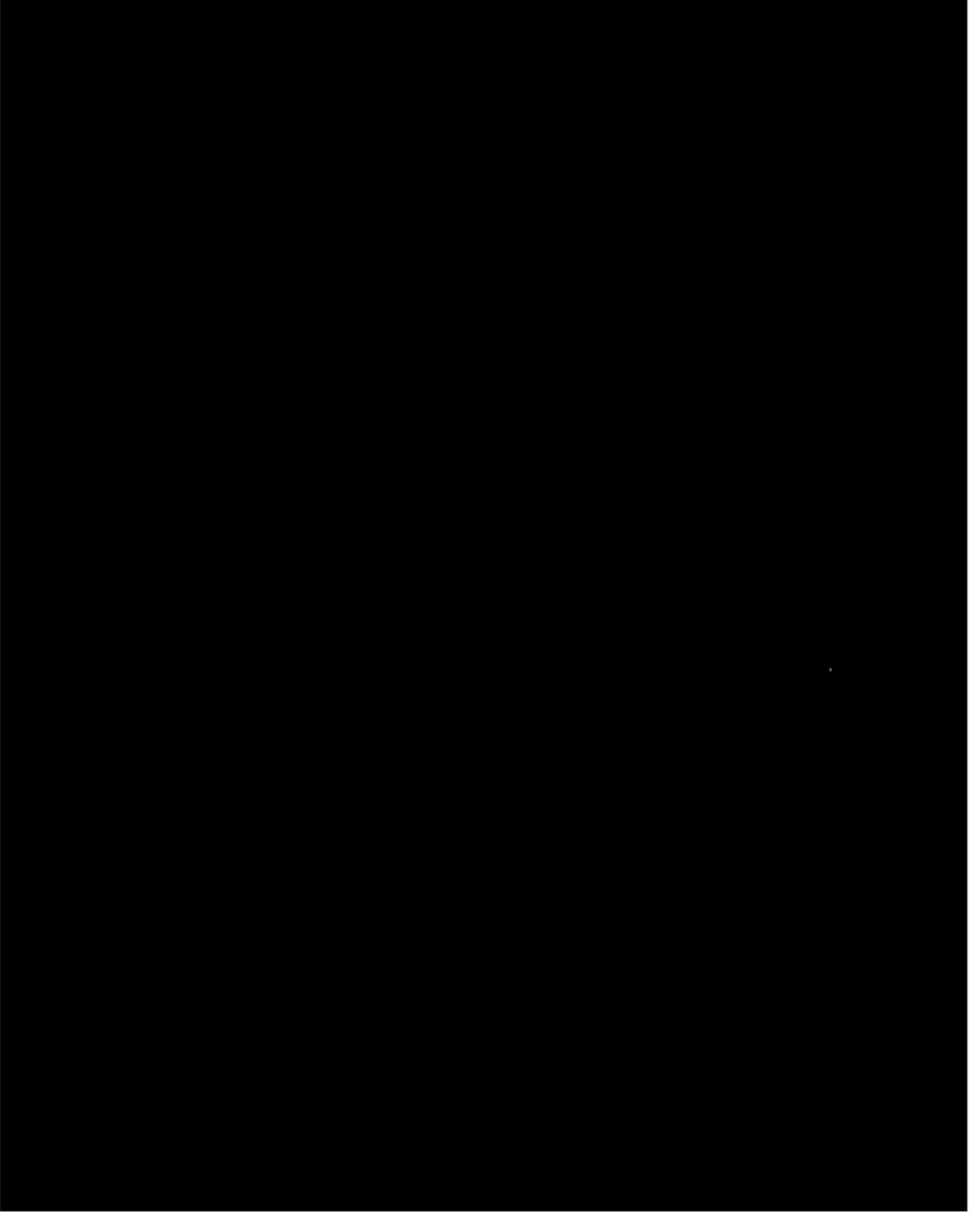
SUMMARY

Autonomous driving or Advanced Driving Assistance Systems (ADAS) with human level driving capability is one of the most active research topics. However, state-of-the-art methods are detecting contributing factors (e.g. lane detection, pedestrian detection, etc.) independently from each other. On the other hand, visual perception is the most important input for human driving decisions, and it requires evaluation of all objects in the environment. State of the art methods are capable of human-like perceiving in the area of visual perception. Therefore, developing holistic approach for detecting driving affecting factors in whole visual area is important in developing human like autonomous driving or ADAS. Visual attention and saliency are popular research topics in neuroscience. In the light of results of studies in these areas, we know see attention and saliency prediction in the machine learning literature.

In this thesis, two novel deep neural network models that predicts braking decision by attending salient features in driving scene for autonomous driving and ADAS are proposed. Both the proposed models extract features from driving input image using a modified VGG-16 convolutional neural network. Extracted features are weighted with an attention mechanism before passing into a convolutional LSTM neural network. The convolutional LSTM network generates saliency map prediction. The predicted saliency map is fed into second step neural network model to predict a binary braking decision. The first model generates braking prediction concurrently with saliency map prediction. This model optimizes both the saliency map prediction and braking prediction. The second model first optimizes saliency map prediction then the original input image is masked with optimized saliency prediction map and passed into another neural network to predict braking decision. The final output of the model is braking decision. The proposed models are trained and evaluated on the state of the art driving attention dataset BDD-A and Autonomous Transportation Vehicle (ATV) simulation dataset. Saliency prediction results for both models are at the state-of-the-art level. Furthermore, both models that are attending salient features detected with a holistic approach on visual perception are obtained motivating results for brake prediction.

Keywords: Computer vision, Attention mechanism, Saliency prediction, Autonomous Vehicles, Brake prediction.

TEŞEKKÜR



İÇİNDEKİLER

	<u>Sayfa</u>
ÖZET	vi
SUMMARY	vii
TEŞEKKÜR	viii
İÇİNDEKİLER	ix
ŞEKİLLER DİZİNİ	xi
ÇİZELGELER DİZİNİ	xiii
SİMGELER VE KISALTMALAR DİZİNİ	xiv
1. GİRİŞ VE AMAÇ	1
2. LİTERATÜR ARAŞTIRMASI	4
2.1. Otonom Sürüş ve Sürüş Destek Sistemleri	4
2.2. Görsel Algı, Dikkat ve Belirginlik	6
2.3. Makine Öğrenmesi Alanında Dikkat ve Belirginlik	9
2.4. Literatürdeki Veri Kümeleri	19
3. MATERYAL VE YÖNTEM	21
3.1. Materyal	21
3.1.1. Veri Kümeleri	21
3.2. Dikkat Mekanizmasıyla Hesaplanan Belirginlik Tahmin Etmesi ile Fren Karar Modeli	25
3.2.1. Görsel Özellik Çıkarımı	27
3.2.2. Belirginlik Tahmin Etme	27
3.2.3. Belirginlik Tahmin Etmesiyle Fren Yapma Kararı Verme	32
3.2.3.1. Eşzamanlı Model	32
3.2.3.2. Ardışık Model	34
4. BULGULAR VE TARTIŞMA	37
4.1. Dikkat Mekanizmasıyla Hesaplanan Belirginlik ile Eşzamanlı Fren Karar Model Bulguları	37

İÇİNDEKİLER (devam)

	<u>Sayfa</u>
4.2. Dikkat Mekanizmasıyla Hesaplanan Belirginlik ile Ardışık Fren Karar Model Bulguları	43
5. SONUÇ VE ÖNERİLER	49
KAYNAKLAR DİZİNİ	51
ÖZGEÇMİŞ	59

ŞEKİLLER DİZİNİ

<u>Şekil</u>	<u>Sayfa</u>
2.1 SAE sınıflandırmasına göre otonom sürüşün beş seviyesi (SAE International, 2018).	4
2.2 Otonom Araçlardaki Algılayıcılar.	5
2.3 NVidia firması tarafından geliştirilen DAVE-2 otonom sürüş model mimarisi. (Bojarski vd., 2016)	5
2.4 Yarbus'un yaptığı gözün hareketini tanımladığı deneyi.(Yarbus, 2013)	7
2.5 Dikkatin aşağıdan yukarı oluşturulması (Itti vd., 1998)	8
2.6 Dikkat ile ilgili kortikal yollar. Dikkat kazanç kontrolü yönetimiyle bir sonraki dikkat edilecek alan belirlenir (Parr ve Friston, 2017).	9
2.7 Bir DYSA modeli.	10
2.8 Görüntü üzerinde evrişim işlemi uygulaması.	10
2.9 Bir GEA içerisinde görüntü üzerinde evrişim işlemi uygulaması (Yu ve Koltun, 2015).	11
2.10 DYSA eğitiminde kullanılan geri yayılım algoritması.	11
2.11 DYSA geri yayılımında en iyileme için kullanılan SGİ algoritması.	12
2.12 UKDH mimarisi (Olah, 2020).	13
2.13 UKDH modelinde C_t durum vektörünün her zaman adımında korunması sağlanmaktadır.	13
2.14 UKDH modelinde kullanılan <i>sigmoid</i> kapı fonksiyonu.	13
2.15 EUKDH mimarisi (Xingjian vd., 2015).	14
2.16 DYSA kapsamında kodlayıcı-çözücü mimarisi.	15
2.17 Girdiden elde edilen özellikleri ağırlıklandırarak çözücüye veren dikkat mekanizması (Bahdanau vd., 2014).	16
2.18 Görüntünün tamamı yerine belirli bir alt parçasına dikkat eden bir model tanımlamışlardır (Mnih vd., 2014).	17
2.19 Belirginlik haritaları, görsel algı alanında dikkati çeken belirgin özelliklerin tahmin edilmesinde kullanımı (Simonyan vd., 2013).	17
2.20 SAM-VGG modeli durağan görüntüde belirginlik tahmin etmek için dikkat mekanizmasından faydalanması (Cornia vd., 2018).	18
2.21 ACLNet modeli akış tipi görüntüde belirginlik tahmin etmek için dikkat mekanizmasından faydalanması (Wang vd., 2019a).	19
2.22 Göz bebeği takip sistemi ile belirginlik haritası oluşturma.	20
3.1 CAT2000 veri kümesine ait görsellerden örnekler.	22

ŞEKİLLER DİZİNİ (devam)

<u>Şekil</u>	<u>Sayfa</u>
3.2 (Çizelge-3.2)'de belirtilen bir örnek veri sırasını oluşturan görseller.	23
3.3 OTA IFARLAB Gerçek ve benzetim ortamı.	24
3.4 (Çizelge-3.4)'de belirtilen bir örnek veri sırasını oluşturan görseller.	26
3.5 Tez çalışmasında önerilen dikkat mekanizmasıyla hesaplanan belirginlik haritası ile fren karar model yaklaşımlarını gösterir genel şema.	26
3.6 GEA kullanılarak değiştirilmiş VGG-16 temelli görüntü özellik çıkarıcı.	27
3.7 Girdi görüntüsü ve modele verilen ortalama piksel değeri çıkarılmış girdi görüntüsü.	28
3.8 Evrişimli UKDH ve dikkat mekanizması kullanarak belirginlik tahmin etme.	28
3.9 Sol: Girdi görüntüsü, Orta: Gerçek belirginlik, Sağ: tahmin edilen belirginlik	29
3.10 Belirginlik haritası tahmin etme ile fren yapma kararı üreten model.	32
3.11 Orijinal ResNet-18 model mimarisi.	35
3.12 Belirginlik haritası tahmin etme ile fren yapma kararı üreten model.	36
3.13 Girdi görüntüsünün belirginlik tahmin etmesi ile maskelenmesi. (a) Girdi görüntüsü, (b) tahmin edilen belirginlik haritası, (c) Girdi görüntüsü üzerinde belirginlik, (d) Belirginlik haritasıyla maskelenen girdi görüntüsü.	36
4.1 Belirginlik tahmin etme örnek bulguları.	38
4.2 CAT2000 veri kümesinden örnek sonuçlar: (a) Girdi görüntüsü, (b) Gerçek belirginlik, (c) CC-KLD-NSS-G16 modeli ile elde edilen belirginlik tahmini, (d) NSS-NCB modeli ile elde edilen belirginlik tahmini.	39
4.3 CAT2000 veri kümesinden başarıyı iyi ve kötü sonuçlar.	40
4.4 CAT2000 veri kümesinde eğitilen modellerin BDD-A veri kümesinden elde edilen test sonuçlarına örnekler.	41
4.5 Farklı eşik ve tolerans değerlerinde elde edilen sürüş karar bulgu çizelgeleri.	42
4.6 OTA benzetim veri kümesinden elde edilen belirginlik tahmin etme örnek sonuçları ve fren kararı için girdi görüntüsünün maskelenmiş biçimleri.	44
4.7 IFARLAB test parkurundaki <i>DUR</i> işareti ve yaya geçidi.	45
4.8 OTA benzetim veri kümesinde fren kararı üreten görüntüler.	46
4.9 OTA benzetim veri kümesinden elde edilen doğruluk çizelgeleri	47
4.10 OTA benzetim veri kümesinden elde edilen kayıp çizelgeleri	47
4.11 25 Epoch ile yapılan eğitim çizelgesi.	48

ÇİZELGELER DİZİNİ

<u>Çizelge</u>	<u>Sayfa</u>
3.1 Veri örneklerini etiketlemede kullanılan eşik değerleri.	23
3.2 Etiketlenen zaman adımından önceki üç zaman adımı için örnek veri sırası. . .	23
3.3 OTA için IFARLAB test parkuru Gazebo benzetim ortamı senaryo, hız ve fren mesafeleri.	25
3.4 OTA veri kümesinden örnek bir veri sırası.	25
4.1 BDD-A veri kümesinde eğitilen önerilen model ile ön eğitilmiş referans modellerin elde ettikleri belirginlik tahmin etme sonuçları.	38
4.2 CAT2000 veri kümesinde belirginlik tahmin etme sonuçları.	39
4.3 BDD-A veri kümesinde belirginlik tahmin etme sonuçları.	41
4.4 Zenginleştirilmiş BDD-A veri kümesinde elde edilen fren sürüş karar bulguları.	43
4.5 OTA veri kümesinde elde edilen fren sürüş karar bulguları.	46

SİMGELER VE KISALTMALAR DİZİNİ

Kısaltmalar	Açıklama
CC	Korelasyon katsayısı
DYSA	Derin Yapay Sinir Ağı (ing. Deep Neural Network, DNN)
EUKDH	Evrişimli Uzun Kısa Dönem Hafıza (ing. Convolutional Long Short Term Memory, C-LSTM)
EYSA	Evrişimli Yapay Sinir Ağı (ing. Convolutional Neural Network, CNN)
GEA	Genişletilmiş Evrişimli Ağ (ing. Dilated Convolutional Network, DCN)
KLD	Kullback-Leibler uzaklığı
NSS	Normalleştirilmiş Belirginlik Taramayolu (ing. Normalized Saliency Scanpath, NSS)
OTA	Otonom Taşıyıcı Araç (ing. Autonomous Transport Vehicle, ATV)
RBF	Radyal Bazlı Fonksiyon (ing. Radial Basis Function, RBF)
SGİ	Stokastik Gradyan İnişi (ing. Stochastic Gradient Descent, SGİ)
UKDH	Uzun Kısa Dönem Hafıza (ing. Long Short Term Memory, LSTM)
UMEİ	Uyarlanabilir Moment En İyileme (ing. Adaptive Moment Optimization, Adam)
YSA	Yapay Sinir Ağı (ing. Neural Network, NN)
YYSA	Yinelemeli Yapay Sinir Ağı (ing. Recurrent Neural Network, RNN)
İSDS	İleri Sürüş Destek Sistemleri (ing. Advanced Driver Assistance Systems, ADAS)

SİMGELER VE KISALTMALAR DİZİNİ (devam)

Simgeler Açıklama

L_{BCE}	İkili çapraz entropi temelli kayıp fonksiyonu
L_{BRAKE}	Fren kararı hata fonksiyonu
L_{CC}	Korelasyon katsayısı temelli kayıp fonksiyonu
L_{CE}	Çapraz entropi temelli kayıp fonksiyonu
L_{DAM}	Belirginlik karar modeli toplam kayıp fonksiyonu
L_{KL}	Kullback-Leibler uzaklığı temelli kayıp fonksiyonu
L_{NSS}	Normalize edilmiş belirginlik tarama yolu temelli kayıp fonksiyonu
L_{TOTAL}	Toplam kayıp fonksiyonu
ϵ	Düzenleyici parametre
μ	Ortalama
σ	Kovaryans
$\varphi(\cdot)$	Radyal temelli fonksiyon tanımı
$exp(\cdot)$	Üstel fonksiyon
$log(\cdot)$	Logaritma fonksiyonu
conv	Evrişim işlemi

1. GİRİŞ VE AMAÇ

Otonom sürüş veya İleri Sürüş Destek Sistemleri (İSDS) alanlarında araştırma ve teknoloji geliştirme çalışmaları artmaktadır. Görsel algılama, insanlı sürüş için en önemli karar girdisi olması nedeniyle, bu araştırma konuları arasında öne çıkmaktadır. ALVINN'den (Pomerleau, 1989) bu yana bilgisayarla görü sistemleri otonom araç sürüş karar sistemlerinde oldukça başarılı sonuçlar elde etmiştir. Bu başarımın arkasında özellikle Derin Öğrenme ve Derin Yapay Sinir Ağı (DYSA) alanındaki gelişmeler bulunmaktadır. Elde edilen başarılı sonuçlar, bu yaklaşımların sürüş görevinde insan seviyesine yakın algılama amacıyla kullanılmasını motive etmiştir. Bu sistemlere örnek olarak NVidia firmasının DAVE-2 (Bojarski vd., 2016) modeli, Xu v.d.'nin FCN-LSTM modeli (Xu vd., 2017) ve Baidu'nun (Yu vd., 2017) çalışmaları gösterilebilir.

Otonom sürüş sistemleri bilgisayarla görünün yanısıra LIDAR, RADAR v.b. algılayıcılar da kullanmakta olup bu alanda durumsal farkındalık için farklı algoritmalar kullanılmaktadır. Bütün bunların bir araya getirdiği karmaşık sistem, işletim zorluğu, maliyet gibi pek çok açıdan otonom araçların yaygınlaşmasının önündeki en büyük engeldir. Kameranın insan görüşüne eşdeğer veri üretmesi nedeniyle kamera görüşünden algılayarak sürüş yapabilmek önemlidir. Bilgisayarlı görü alanında trafikteki farklı nesnelere algılanmasına yönelik çalışmalar olgunlaşmış olmakla beraber sürüş için önemli olan görsel algı ile dikkat mekanizmasına yönelik çalışmaya rastlanmamıştır. İnsan dikkat mekanizmasının tam işleyişi henüz bilinmese de, mevcut sinir bilim alanındaki araştırmaların sonuçları makine öğrenmesi alanında kullanılmaya başlanmıştır.

Sinir bilim alanında dikkat mekanizması ele alındığında insan görüşü ile ilgili öne çıkan iki süreç dikkat ve belirginliktir. Görsel algı bağlamında dikkat ve belirginlik, birbiriyle ilişkili ama farklı sinirsel mekanizmalarla eş zamanlı çalışan mekanizmalardır (Parr ve Friston, 2017). Dikkat, zihnin sınırlı hesaplama kaynaklarının değişken biçimde kontrol edilebilmesi işlevidir (Lindsay, 2020). Öte yandan belirginlik ise, algılanan görüntü içerisinde belirli desenleri arama işlevidir (Itti ve Koch, 2000). Koch ve Ullman ile başka bir çalışmada Itti v.d. tarafından açıklandığı gibi (Koch ve Ullman, 1987; Itti vd., 1998), insan görüşü, görüntü alanını sürüş v.b. için gerekli belirgin özellikler ile arka planın bileşimi olarak algılar. Literatürde tanımlandığı gibi görsel dikkat, renk, yönelim, kenarlar v.b. aşağıdan yukarı veya kapsama bağlı olarak yukarıdan aşağı oluşturulur (Borji ve Itti, 2012). Belirginlik, veya daha açık olarak görüntüdeki belirgin özellikleri bulmak bilgisayarla görü alanındaki aktif araştırma alanlarından birisidir. Her ne kadar kapsam

bilgisi gerektirmeyen aşağıdan yukarı yaklaşımı yoğun olarak araştırılmışsa da, sürüş için görsel dikkat tüm görüntü alanı yerine belirgin özelliklere odaklanmayı gerektirdiği için, yukarıdan aşağı olarak sınıflandırılır (Tawari vd., 2018), (Ramanishka vd., 2018).

Diğer taraftan makine öğrenmesinde girdinin belirli bir alt kümesine daha fazla ağırlık verilmesi olarak tanımlanabilececek dikkat kavramı yakın zamanda literatüre girmiştir. Bu yaklaşım makine çevirisi ve doğal dil işleme/anlama alanlarında, yüksek başarımlar elde etmektedir (Bahdanau vd., 2014). Bilgisayarla görü alanında da (Mnih vd., 2014) bu yaklaşımı görsel algılama görevlerinde ilk olarak kullanmış ve başarımlar elde etmiştir. Yakın zamandaki bir diğer çalışma olan SAM (Cornia vd., 2018), dikkat ile belirgin özelliklerin tahmin edilmesi alanında SalNet (Han vd., 2018), ML-Net (Cornia vd., 2016) ve DeepGaze II (Kümmerer vd., 2016) gibi çalışmaları geride bırakarak MIT1003 (Bylinskii vd., 2015), CAT2000 (Borji ve Itti, 2015) ve SALICON (Jiang vd., 2015) veri kümelerinde en yüksek başarımları elde etmiştir. Benzer şekilde sabit imge yerine akış tipi (video) içerikte belirgin özelliklerin belirlenmesinde de ACLNet (Wang vd., 2019a) yüksek başarımlar elde etmiştir. Bu çalışmalar belirginlik tahmin etmede dikkat mekanizmasını kullanmaları açısından insan görsel algısındaki dikkat mekanizmasına en yakın makine öğrenmesi yaklaşımları olarak dikkat çekmektedir.

Sürücünün görsel dikkati konusu çalışılmış ve bu alanda özel BDD-A (Xia vd., 2019), DR(eye)VE (Palazzi vd., 2019) gibi veri kümeleri üretilmiş olsa da, literatürde dikkat yaklaşımının sürüş için belirgin özelliklerin tahmin edilmesinde ve bunların karar vermede kullanımına rastlanmamıştır. Sürüş için belirgin özelliklerin tahmin edilmesi, statik resimlerde veya herhangi bir kontrol görev önceliğinin olmadığı serbest akış tipi görüntülerde belirgin özelliklerin tahmin edilmesi probleminden farklıdır.

Bu tez çalışmasında, diğer araçlar, yayalar, trafik ışıkları gibi sürüş kararını etkileyecek sürüşe özgü belirgin özellikleri tahmin etmekte dikkat mekanizması kullanan ve bu belirgin özelliklerle fren karar tahmini üreten yeni bir model önerilmektedir. Önerilen model, yaya tespiti, trafik lambası tespiti v.b. sürüş için sadece tek bir göreve odaklanmak yerine bütüncül bir yaklaşım sunmaktadır. Bu sayede, önerilen model sürüş sistemi (otonom sürüş veya sürücü destek sistemi) için gerekli görsel algı bileşeni sayısını azaltmaktadır. Öte yandan, hesaplama performansını artıracak şekilde hiperparametre sayısını da önemli ölçüde azaltmaktadır.

Önerilen modelin fren kararı tahmini üretmek için iki farklı yaklaşımı mevcuttur. Birincisi, girdi görüntüsünde belirginlikler ile fren kararı tahminini eş zamanlı üreten bir yaklaşımdır. İkinci yaklaşımda ise model belirginlik tahmin etme için eğitildikten sonra girdi görüntüsü, tahmin edilen belirginlik haritası ile ağırlıklandırılarak fren tahmini

üretmekte kullanılmaktadır. Her iki yaklaşımda da görsel özellik çıkarımı ve belirginlik tahmin etme kısımları ortaktır.

Önerilen modelde belirginlik tahmin etme bileşeni, ACLnet (Wang vd., 2019b) ve SAM-VGG ve SAM-RESNET model mimarileri (Cornia vd., 2018) temel alınarak oluşturulmuş bir modeldir. Önerilen modelde belirginlik tahmin etme bileşeni özellik çıkarıcı olarak değiştirilmiş VGG-16 (Simonyan ve Zisserman, 2014) (veya ResNet (He vd., 2016)) Genişletilmiş Evrişim Ağı (GEA) (Yu ve Koltun, 2015) kullanılmaktadır. Daha sonra çıkarılan özellikler normal Uzun Kısa Dönem Hafıza (UKDH) (Hochreiter ve Schmidhuber, 1997) yerine Evrişimli Uzun Kısa Dönem Hafıza (EUKDH) (Xingjian vd., 2015) ağından geçirilmektedir. Çıkarılan özelliklerin EUKDH ağından geçirilme esnasında dikkat mekanizması kullanılmaktadır. Dikkat mekanizmasıyla ağırlıklandırılan özellikler evrişim ağı ile belirgin özellik haritası tahmin edilmektedir.

Belirginlik tahmin edildikten sonra fren kararı tahmin edilmektedir. Bu esnada belirginlik tahmini ile eşzamanlı olarak fren kararı üreten model yaklaşımında, tahmin edilen belirginlik haritası üç yoğun katmanlı yapay sinir ağından geçirilerek fren kararı tahmini elde edilmektedir. Belirginlik tahminiyle ardışık olarak frena karar tahmini üreten model yaklaşımında, girdi görüntüsü tahmin edilen belirginlik haritası ile maskelenerek fren karar tahmini için değiştirilmiş ResNet-18 ağına verilmektedir. Önerilen modeller literatürde fren bilgisiyle zenginleştirilmiş BDD-A veri kümesinde ve Otonom Taşıyıcı Araç (OTA) (IFARLAB, 2020) benzetim verilerinde sınanmıştır. Bu çalışmanın literatüre en önemli katkısı görsel algıdaki belirginin özelliklere bütüncül olarak dikkat edip fren tahmini üreten modellerin geliştirilmesidir.

Takip eden bölümde bu tez çalışması kapsamındaki konulara ilişkin literatür taraması verilmektedir. Üçüncü bölümde tez çalışmasında önerilen modellerin detaylı tanımı verilmektedir. Dördüncü ve beşinci bölümlerde önerilen modelin veri kümelerinde elde ettiği sonuçlar ve bu sonuçlara ilişkin yorum ve öneriler verilmektedir.

2. LİTERATÜR ARAŞTIRMASI

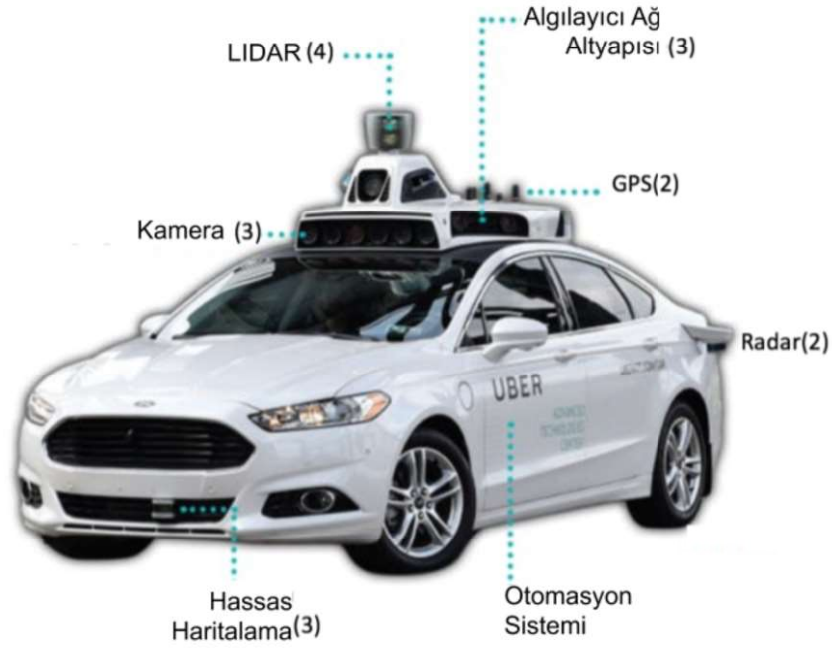
2.1 Otonom Sürüş ve Sürüş Destek Sistemleri

Otonom Sürüş ve İleri Sürüş Destek Sistemleri (İSDS) alanında son yıllarda yoğun araştırma ve geliştirme çalışmaları yapılmaktadır. Bu çalışmalarda otonom sürüş seviyesini belirtmek için Society of Automotive Engineers, SAE tarafından geliştirilen sınıflandırmada belirtilen otonom sürüş için beş seviye önerilmektedir (SAE International, 2018). İlk üç seviyedeki özellikler İSDS kapsamında yapılan çalışmaları tanımlamakta olup, seviye 3 ve üzeri ise tam veya kısmi otonom sürüş olarak nitelendirilmektedir (Şekil-2.1). Seviye 5 ise tam otonom sürüş olarak nitelendirilmekte olup bu seviyedeki bir aracın her koşul altında otonom olarak sürüş yapması beklenmektedir.

Diğer taraftan bugün mevcut otonom araçlar üçüncü ve dördüncü seviye otonom araç olarak nitelendirilmektedir. Ayrıca, otonom sürüş yapabilmeleri için Şekil-2.2’de gösterildiği gibi çok sayıda algılayıcı ile donatılmış olmak zorundadırlar. Ancak, insana benzer sürüş yetenekleri için otonom araçların bilgisayarla görü ile sürüş kararı üretebilmesi önemlidir.

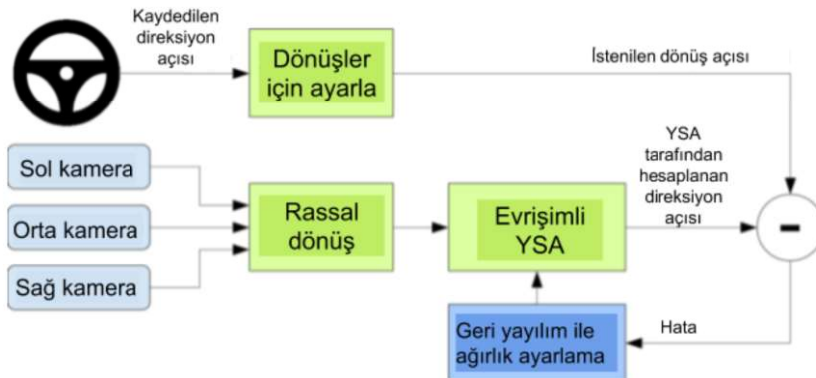


Şekil 2.1: SAE sınıflandırmasına göre otonom sürüşün beş seviyesi (SAE International, 2018).



Şekil 2.2: Otonom Araçlardaki Algılayıcılar.

DYSA son zamanlarda birçok alanda olduğu gibi otonom sürüş alanında da etkin olarak kullanılmaktadır. Yapay sinir ağları ile otonom sürüşe öncül çalışma örneği olarak ALVINN (Pomerleau, 1989) verilebilir. Yakın zamanlarda (Xu vd., 2017) tam evrişimli yapay sinir ağı ve UKDH mimarisini içeren çalışmalarını sunmuşlardır. Bu çalışmada, geniş ölçekli topluluk tarafından oluşturulmuş araç eylem veri kümesini kullanarak uçtan uca bir sürüş sistemi geliştirmişler ve veri kümesini paylaşmışlardır. Ayrıca NVidia firmasının DAVE-2 modeli (Bojarski vd., 2016) ve Baidu'nun geliştirdiği mimari (Yu vd., 2017) de bu yaklaşıma diğer örnekler olarak verilebilir (Şekil-2.3).



Şekil 2.3: NVidia firması tarafından geliştirilen DAVE-2 otonom sürüş model mimarisi. (Bojarski vd., 2016)

Öte yandan otonom sürüş sistemi içerisinde DYSA kullanımı alt bileşen seviyesinde oldukça yaygındır. Yaya tespiti, sürülebilir yol tespiti, yol çizgi tespiti, trafik işareti ve levhası tespiti v.b. amaçlarla DYSA temelli sınıflandırıcılar yaygın olarak kullanılmaktadır. Örneğin Zhang v.d. (Zhang vd., 2016) yaya tespitinde DYSA kullanan başarılı bir model sunmuştur. Otonom sürüş için sürülebilir yol tespiti alanında da Chen v.d. (Chen vd., 2019) algılayıcı verisini DYSA ile sınıflandırarak yol tespiti yapmaktadır. Farklı bir yaklaşımla Zou v.d. (Zou vd., 2020) yol çizgi tespiti yapmaktadır. Wu v.d. (Wu vd., 2013) trafik işareti tespitinde DYSA kullanarak başarılı sonuçlar elde etmiştir. Ancak bu modeller sürüş kararı üretmemektedir. Ayrıca, sürüş görüntüsünü bütüncül olarak ele almadıkları için doğru sürüş kararı üretmek için beraber kullanılmalıdır.

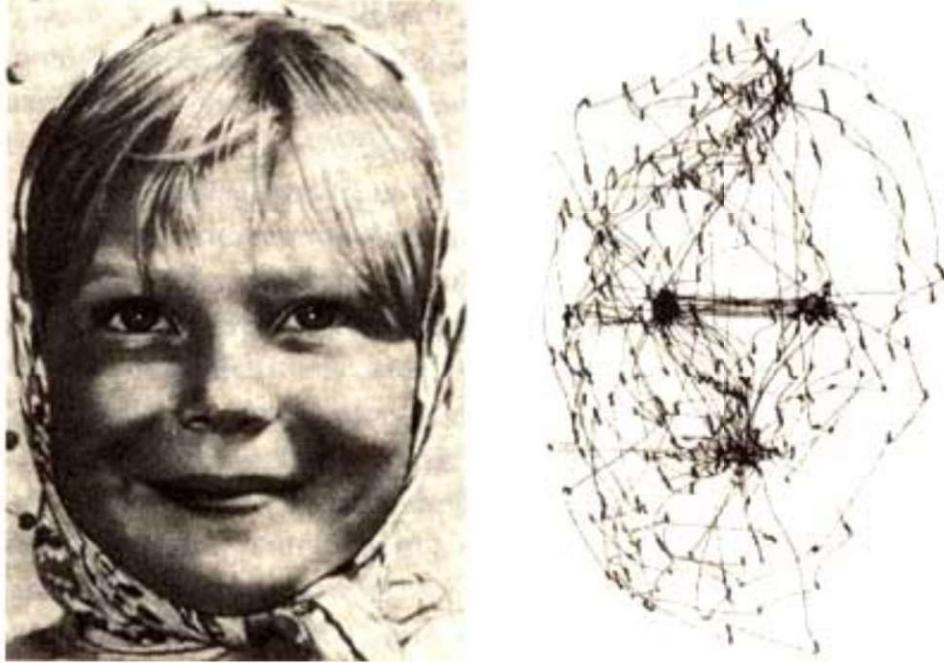
2.2 Görsel Algı, Dikkat ve Belirginlik

Görsel Algı ve Dikkat, sinirbilim alanında yoğun olarak çalışılan bir konudur. Bu alandaki temel çalışmalardan birisi Koch ve Ullman'ın (Koch ve Ullman, 1987) seçici görsel dikkat modelini tanımladıkları çalışmadır. Bu çalışmada, görsel algının öncelikli olarak renk, yönelim, hareket yönü v.b. temel özelliklerden bir erken topografik gösterim oluşturduğu, daha sonra bu gösterimden görüntü alanının belirli bir konumuna ait özellikleri içeren bir eşleştirme yaptığı ve bu eşleştirme ile merkezi bir gösterim oluşturduğu ortaya konmuştur. Fizyolojik olarak görme mekanizmasının işleyişi bilinmesine rağmen görsel olarak algının tam olarak nasıl yönetildiği henüz net olarak belirlenmemiştir. Öncelikli olarak psikolojide davranışsal deneylerle incelenen dikkat kavramı sinirbilim çalışmaları kapsamında zihinsel süreçler ile modellenmeye ve tanımlanmaya çalışılmıştır (Borji ve Itti, 2012). Bu çalışmalarda en genel haliyle dikkat, bireyin bulunduğu ortamla etkileşime girmeye hazır olması olarak tanımlanmıştır.

Lindsay'e göre dikkat, kısıtlı hesaplama kaynaklarının dinamik olarak yönetilmesidir (Lindsay, 2020). Her ne kadar sinirbilim çalışmaları kapsamında dikkatin ortak kabul edilebilir bir tanımı yapılmamış olsa da, Lindsay'in tanımı genel olarak yapılmış tüm tanımları kapsar.

Gözün fizyolojik yapısında retinanın sarı cisim (fovea) adı verilen orta kısmı en yüksek görsel algılama kabiliyetine sahip bölgesidir. Bu açıdan ele alındığında, saccade adı verilen, gözün sürekli ve aktif olarak hızlı küçük hareketler ile bulunduğu ortamı algılaması, Lindsay'ın tanımına uygundur.

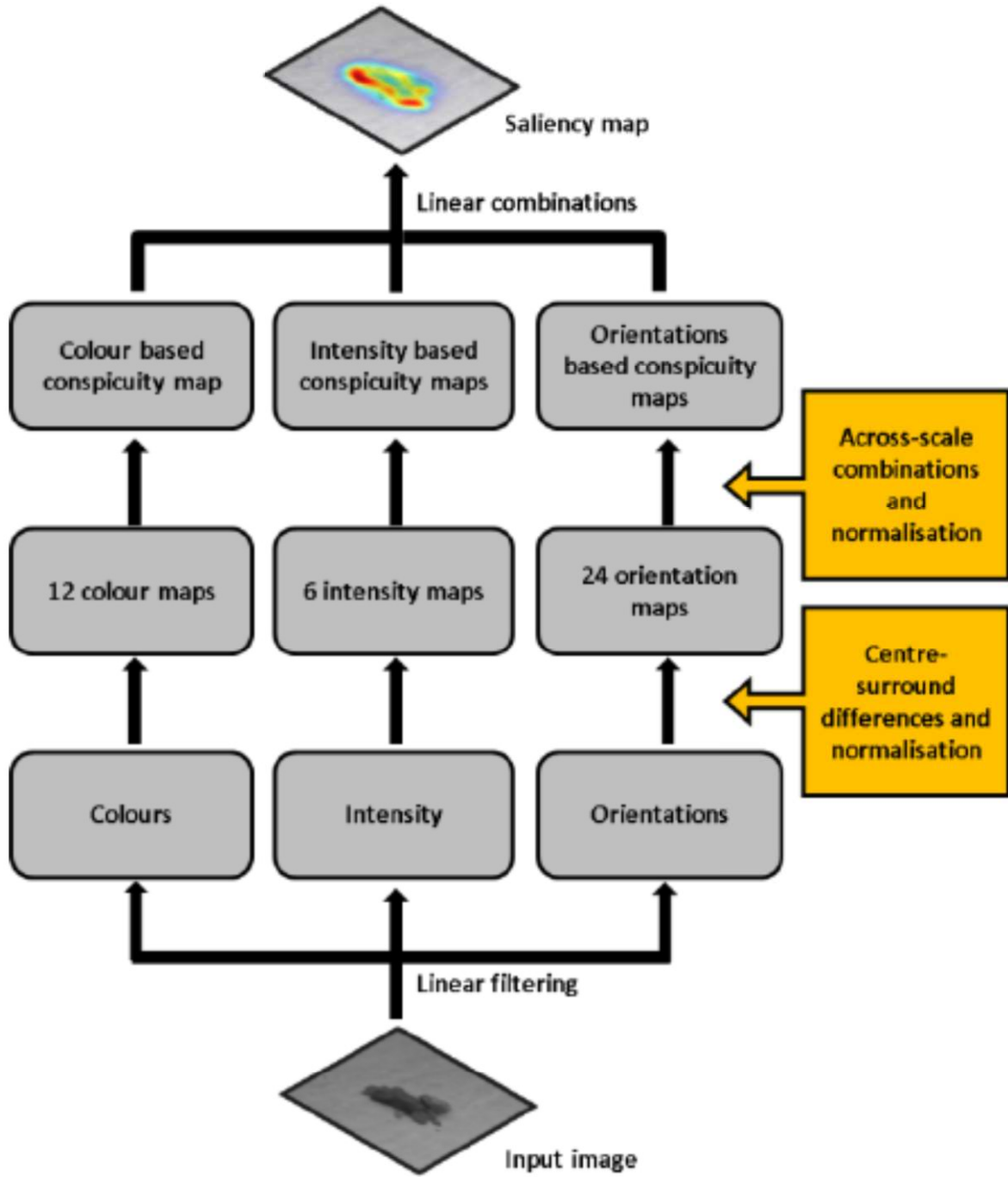
Gözün bu hareketi ile ilgili yapılan deneyler, dikkat ve belirginlik açısından görsel algıdaki süreçleri anlamak konusunda önemli ipuçları vermiştir (Şekil-2.4). Bu konuda Yarbus'un 1967'deki çalışması temel kabul edilir (Yarbus, 2013).



Şekil 2.4: Yarbus'un yaptığı gözün hareketini tanımladığı deneyi.(Yarbus, 2013)

Koch ve Ullman'ın çalışmasını esas alan Itti vd. çalışmalarında (Itti vd., 1998), görsel algının işlenmeden önce belirli bir alt kümesinin belirginliğe bağlı olarak seçildiğini öne sürmüştür. Bu seçimin muhtemelen karmaşıklığı azaltmak için olduğunu ve görsel alanın belirli bir bölgesine dikkati odaklayarak yapıldığını öne sürmüşlerdir. Burada öne çıkan husus, görsel algının bu seçimi belirgin özelliklere göre ve görevden bağımsız olarak aşağıdan yukarı oluşturduğuna yapılan vurgudur (Şekil-2.5). Öte yandan dikkat göreve bağlı olarak yukarıdan aşağı da dikkat oluşturulabilir. Bu temel çalışmada dikkati aşağıdan yukarı oluşturan belirginlik haritaları kavramı ortaya atılmıştır. Bu çalışmanın devamı olarak sayılabilecek bir diğer çalışmada, Itti ve Koch (Itti ve Koch, 2000) belirginlik temelli görsel algı modelini önermişlerdir. Önerilen bu modelde, çok boyutlu görsel karakteristik özellikler tek bir topolojik belirginlik haritasında birleştirilmiştir. Böylece görüntü alanındaki belirgin özellikler anlaşılır ve hesaplanabilir olarak tanımlanabilmiştir.

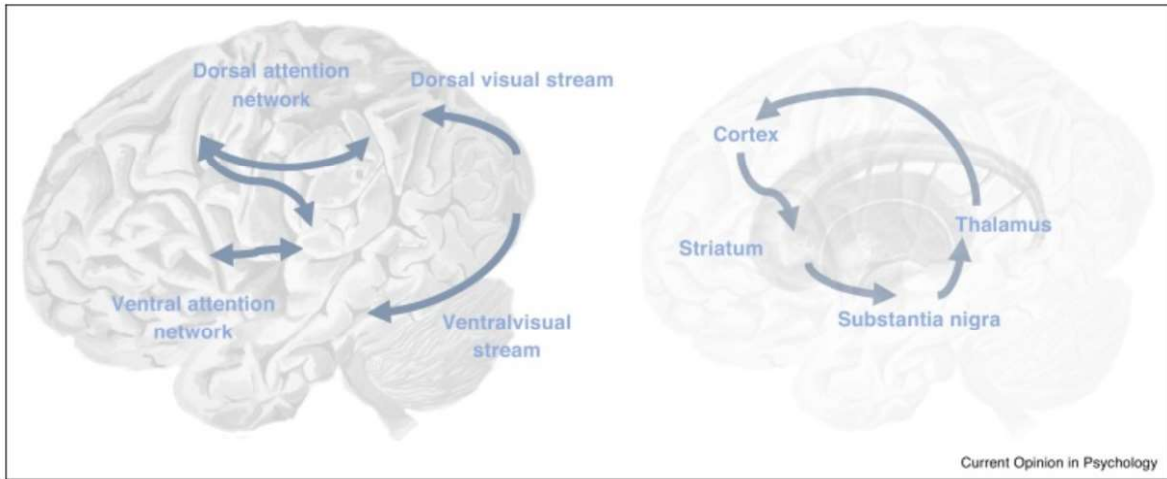
Görsel algı ve dikkat ile ilgili araştırma konularından bir diğeri de dikkatin açık ve kapalı biçimleridir. Açık dikkat, gözün hareketi ile gözlemlenebilen dikkatini vurgular. Kapalı dikkat ise, gözün sarı cisminin (foveanın) özel olarak belirli bir bölgeye yönlendirilmeden sinirsel süreçle dikkatin görüntünün belirli bir bölgesine verilmesidir. Kapalı dikkatin önemli



Şekil 2.5: Dikkatin aşağıdan yukarı oluşturulması (Itti vd., 1998)

bir özelliği, görsel girdi değişmeden dikkatin girdinin farklı bölgelerine yoğunlaşabilmesidir. Ancak belirgin özellikler kapalı dikkati değiştirebilir. Dahası, kapalı dikkatin açık dikkati yönetmek için olduğunu öne süren yaklaşımlar da vardır (Rizzolatti vd., 1987). Beyindeki FEF alanı gözün hareketlerini kontrol eden bölümdür ve bu alanın elektrik sinyalleriyle uyarılarak göz hareketliliğinin arttığı gözlenmiştir (Moore vd., 2003).

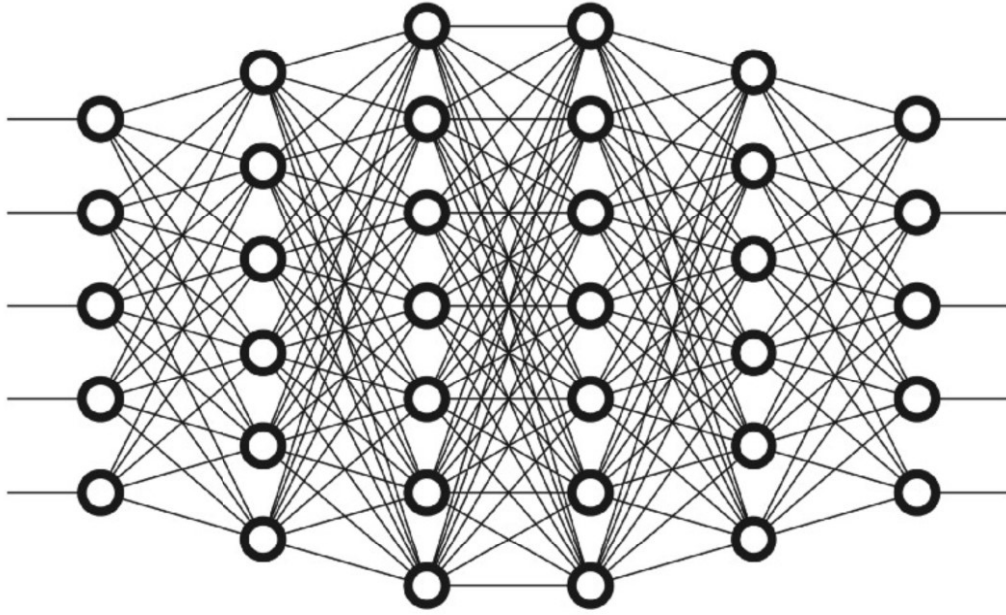
Öte yandan Parr ve Friston (Parr ve Friston, 2019) dikkat ve belirginliğin birbirlerine paralel olarak işleyen iki ayrı süreç olduğunu öne sürmektedir. Bu yaklaşıma göre dikkat, farklı girdilere farklı ağırlıklar veren bir mekanizmayken (bir kazanç kontrol mekanizması), belirginlik, aktif olarak algı arama mekanizması olarak belirtilir. Bu açıdan Moore vd.nin (Moore vd., 2003) çalışmasına benzerlik taşır. Bu yaklaşımın temelinde, Friston v.d.'nin (Friston vd., 2012) görsel algının, içerisinde bulunan çevreyle ilgili varsayımları desteklemek için Bayesçi bir yaklaşımla sürekli olarak hipotezler oluşturup bunu doğrulayacak delil araması süreci olduğunu öne sürmektedir. Buradan hareketle görsel algı aktif şekilde görsel algı içerisindeki belirgin özellikleri çıkararak bunu işler ve bununla bir sonraki dikkat edilecek alanı belirler ve bu süreç paralel şekilde birlikte işler (Parr ve Friston, 2017) (Şekil-2.6).



Şekil 2.6: Dikkat ile ilgili kortikal yollar. Dikkat kazanç kontrolü yönetimiyle bir sonraki dikkat edilecek alan belirlenir (Parr ve Friston, 2017).

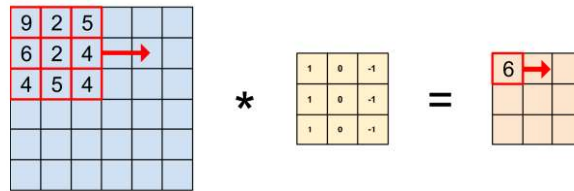
2.3 Makine Öğrenmesi Alanında Dikkat ve Belirginlik

DYSA bilgisayarla görü alanında yoğun olarak çalışılmaktadır. DYSA, Yapay Sinir Ağı'nın (YSA) birden fazla ara katmana sahip olduğu özel bir biçimidir (Şekil-2.7). Bilgisayarla görü için kullanılan en yaygın biçimi Evrişimli Yapay Sinir Ağı (EYSA) modelidir. EYSA modeli evrişim işlemine dayanmaktadır. Evrişim işlemi, bir fonksiyonun biçiminin başka bir fonksiyon ile nasıl şekillendirildiğini ifade eder. Bilgisayarlı görü açısından evrişim, girdi görüntüsündeki parçaları oluşturan kenar gibi özellikleri ortaya çıkarır (Şekil-2.8). EYSA mimarisi, DYSA çok katmanlı yapısında her katmanda görüntüyü oluşturan farklı temel öznitelikleri öğrenen yapıya sahiptir.

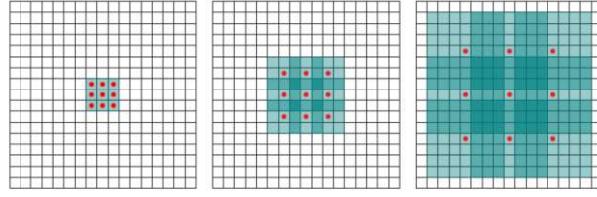


Şekil 2.7: Bir DYSA modeli.

Öte yandan, günümüz EYSA derin öğrenme mimarisinde tüm görüntü pikselleri üzerinde gerçekleştirilen yoğun evrişim işleminin yol açtığı hızlı aşağı ölçekleme, hesaplama başarımı v.b. problemleri gidermek için GEA mimarisinde kullanılan genişletilmiş evrişim işlemi geliştirilmiştir (Yu ve Koltun, 2015). Genişletilmiş evrişim işlemi, ardışık pikseller yerine parametreyle belirlenmiş aralıklı pikseller üzerinde gerçekleştirilir (Şekil-2.9).

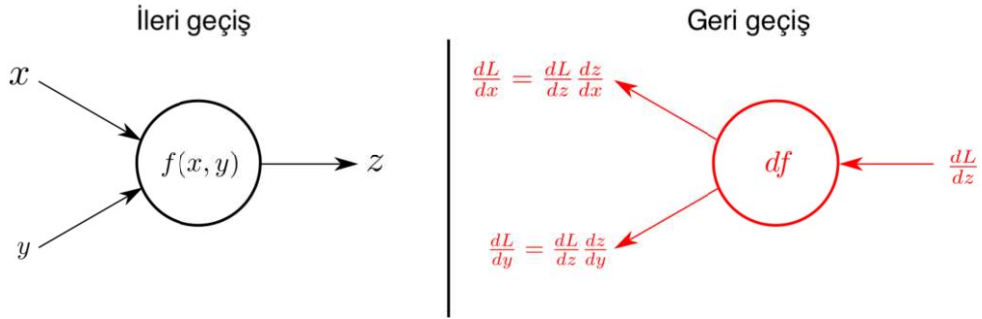


Şekil 2.8: Görüntü üzerinde evrişim işlemi uygulaması.

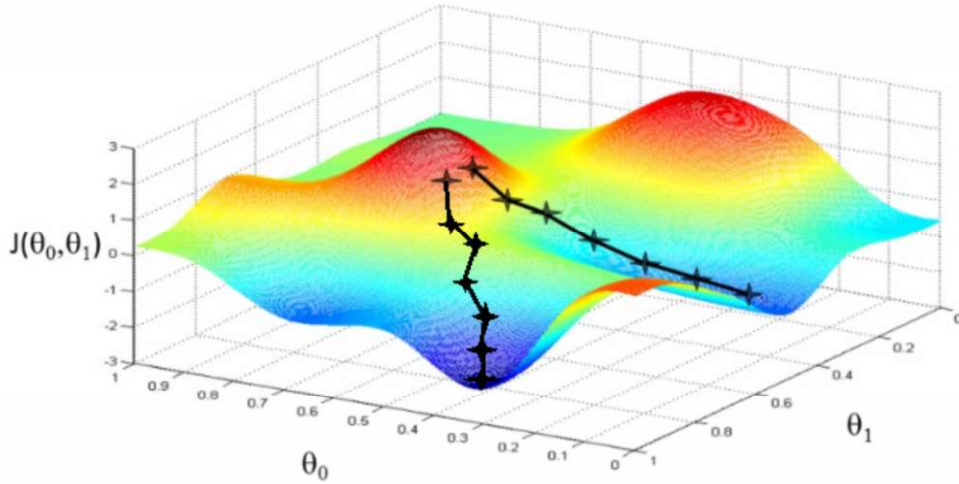


Şekil 2.9: Bir GEA içerisinde görüntü üzerinde evrişim işlemi uygulaması (Yu ve Koltun, 2015).

Genel olarak bir DYSA, geri yayılım algoritması ile eğitilir (Rumelhart vd., 1986). Geri yayılım algoritmasında önce DYSA girdi ile beslenerek çıktı üretmesi sağlanır. Daha sonra üretilen çıktı ile beklenen çıktı değeri arasındaki hata seçilen hata fonksiyonuna göre hesaplanır. Hesaplanan hata DYSA'nın son katmanından başlanarak geriye doğru her katmandaki birimin ağırlık katsayısında yapılması gereken değişiklik miktarı, çıktının ilgili katsayıya göre kısmi türevi kullanılarak hesaplanır (Şekil-2.10). Ortaya çıkan en iyileme probleminin çözümü için Stokastik Gradyan İnişi (SGİ) ve Uyarlanabilir Moment En İyileme (UMEİ) gibi en iyileme algoritmaları kullanılır (Şekil-2.11).



Şekil 2.10: DYSA eğitiminde kullanılan geri yayılım algoritması.

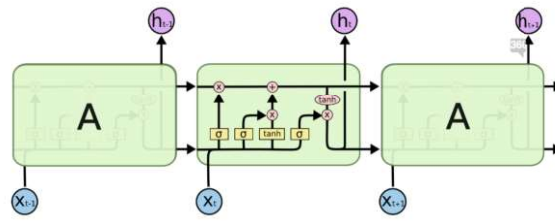


Şekil 2.11: DYSA geri yayılımında en iyileme için kullanılan SGI algoritması.

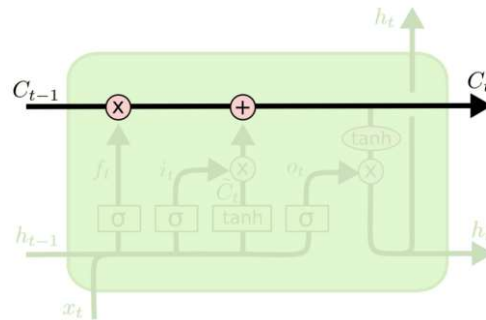
DYSA'nın eğitim sürecinde karşılaşılan en önemli sorun, kaybolan gradyan sorunudur. Geri yayılım algoritmasında çıktı katmanından girdi katmanına doğru ilerledikçe, kısmi türev zincir kuralına göre (2.1) geriye doğru gradyan hesaplanırken, gradyan değerleri hızla 0'a doğru yaklaşmaktadır. Bu durum girdi katmanına yakın ilk katmanlardaki ağırlıkların, hataya olan katkısının hesaplanmasını imkansızlaştırmasa da oldukça zor hale getirmektedir. Bu sorun, uzun yıllar boyunca YSA'nın kullanımının önünde en büyük engel olarak görülmüştür.

$$\frac{d}{dx} [f(u)] = \frac{d}{du} [f(u)] \frac{du}{dx} \quad (2.1)$$

Kaybolan gradyan problemi, özellikle zaman serisi biçimindeki verinin analizinde YSA'nın kullanımını engellemektedir. Bu sorunu giderecek şekilde Uzun Kısa Dönem Hafıza (UKDH) tipi ağlar yaygın olarak kullanılmaktadır (Hochreiter ve Schmidhuber, 1997) (Şekil-2.12). UKDH tipi ağ, her katmandaki gradyanı korumak için özel bir yapıya sahiptir (Olah, 2020) (Şekil-2.13). Bir UKDH birimi, durum bilgisine ekleme veya çıkarma yapabilir. Bu ekleme ve çıkarma işlemlerini *sigmoid* kapıları ile gerçekleştirmektedir. Özel olarak seçilen *sigmoid* kapısı, $[0,1]$ arasında değer alarak hangi bilginin ne kadar geçeceğine karar vermektedir (Şekil-2.14).

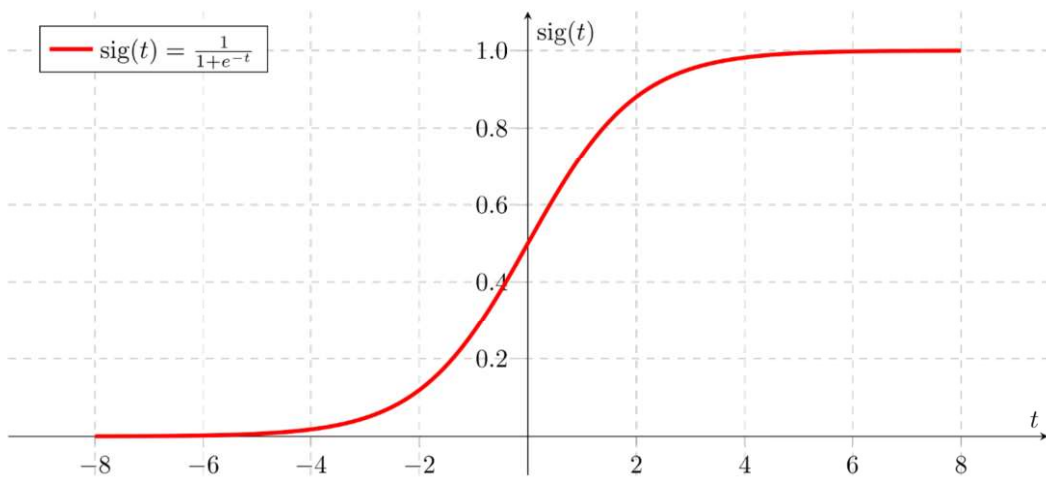


Şekil 2.12: UKDH mimarisi (Olah, 2020).



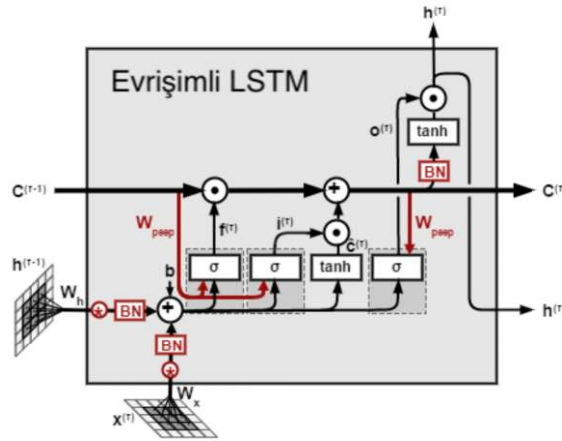
Şekil 2.13: UKDH modelinde C_t durum vektörünün her zaman adımında korunması sağlanmaktadır.

UKDH birim durum vektörü, hangi bilginin unutulacağına karar veren unutma kapısı, ne kadar bilginin ekleneceğine karar veren girdi kapısı ve birimin çıktısına karar veren çıktı kapısı bölümlerinden oluşur.



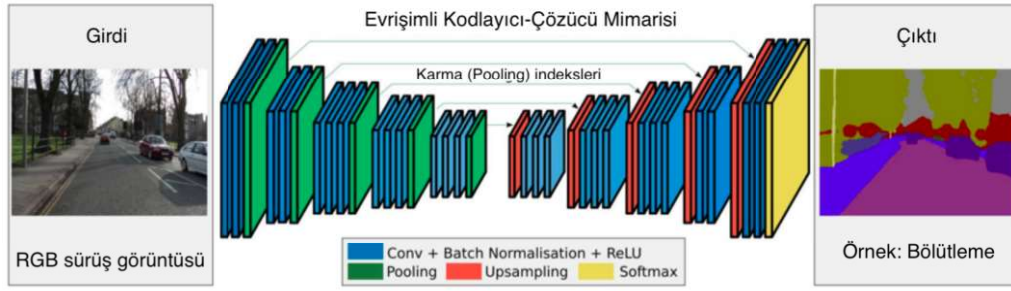
Şekil 2.14: UKDH modelinde kullanılan *sigmoid* kapı fonksiyonu.

Ancak UKDH ağırları tek boyutlu veri analiz yapabildiği için akış tipi görüntülerde kullanılmak istendiğinde, en az 2 boyutlu (veya RGB olarak 3 boyutlu) görüntünün düzlenerek tek boyuta indirilmesi gerekmektedir. Bu tek boyuta indirme, parametre sayısını önemli ölçüde arttırdığı için, akış tipi görüntülerde kullanılmak üzere iki boyutlu Evrişimli Uzun Kısa Dönem Hafıza (EUKDH) kullanılır (Xingjian vd., 2015). EUKDH, orijinal UKDH denklemlerini evrişim işlemiyle değiştirerek kullanır. Böylece, parametre sayısı dramatik olarak artmaz (Şekil-2.15).



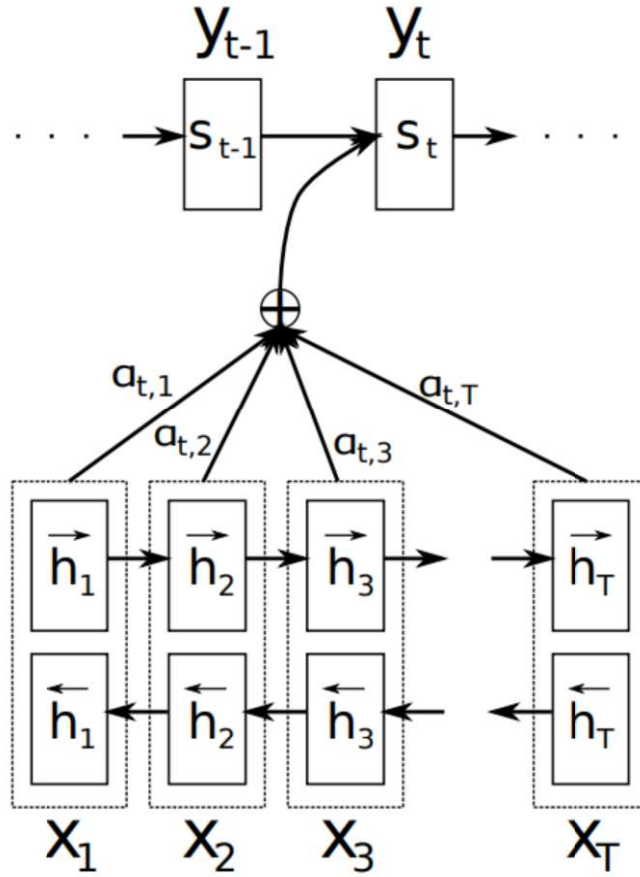
Şekil 2.15: EUKDH mimarisi (Xingjian vd., 2015).

Bütün bu temel özelliklerin yanısıra, girdi verisinin nasıl işlendiğini ve hedeflenen çıktının nasıl üretildiğinin daha anlaşılır olması için, bilgisayarlı görü alanında günümüz yapay sinir ağlarında sıkça kullanılan kodlayıcı-çözücü mimarisini ele almak gerekmektedir. Bu mimaride kodlayıcı, girdiyi temsil eden özelliklerin çıkarılmasında kullanılmaktadır. Çözücü adı verilen ikinci bölüm girdiden elde edilen özellikleri kullanarak amaca yönelik sonucun elde edilmesinde kullanılmaktadır (Şekil-2.16). Bu mimari bugün pek çok yapay sinir ağında kullanılan mimarilerin genel biçimidir.



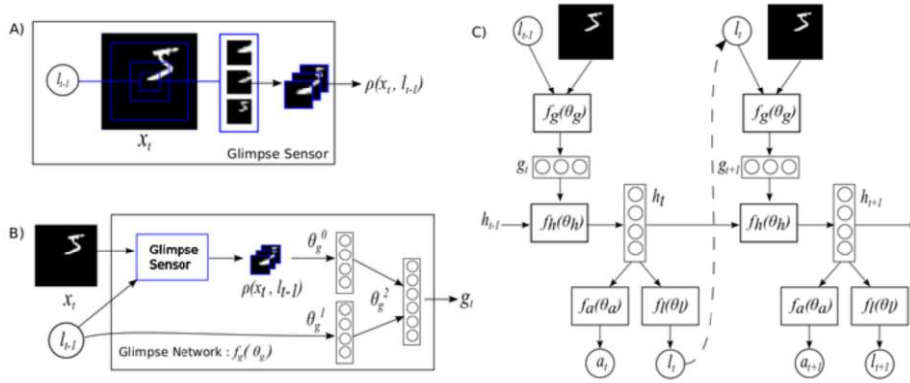
Şekil 2.16: DYSA kapsamında kodlayıcı-çözücü mimarisi.

Kodlayıcı-çözücü mimarisinde girdiyi temsil eden özelliklerin doğrudan çözücüye verilmesi dikkat çekmektedir. Çözücü, kendisine verilen özelliklerin her birini eşit önemde ele alır. Gerçek dünya problemlerinde ve daha da önemlisi insan beynindeki süreçlerde, her girdi aynı önem ve ağırlıkla ele alınmaz. Bu yaklaşımı esas alarak özellikle zamanı da içeren seri şekilde veride her girdi için ayrı önem (veya ağırlık) verecek şekilde bir ağırlıklandırma yaklaşımı makine öğrenmesi alanında yakın zamanlarda yaygınlaşan bir araştırma konusu olmuştur. Girdiden çıkarılan özellikleri ağırlıklandırarak kullanan makine öğrenmesi dikkat mekanizmasının ilk örneği, Bahdanau v.d. (Bahdanau vd., 2014) makine çevirisi, görüntüyü tanımlayan metin üretme ve ses tanıma alanlarında kullanılabilecek modelidir. Bahdanau v.d. (Bahdanau vd., 2014) doğal makine çevirisindeki sabit uzunluklu vektörlerin temel çözücü-kodlayıcı mimarisinin başarımındaki zayıf noktası olduğunu varsaymışlardır. Doğal dil işlemede hedef kelimeyi tahmin etmek için kaynak cümledeki parçaları otomatik olarak arayacak şekilde bu modeli geliştirmeyi önermişlerdir (Şekil-2.17).



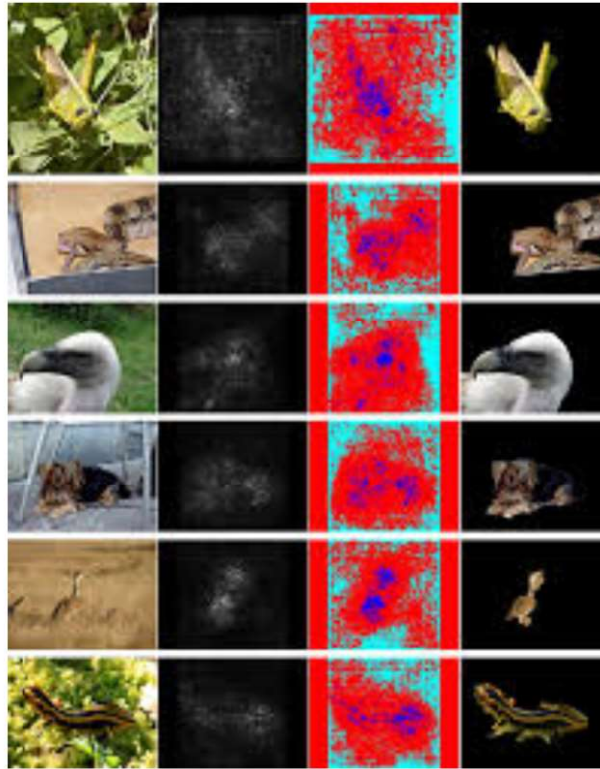
Şekil 2.17: Girdiden elde edilen özellikleri ağırlıklandırarak çözücüye veren dikkat mekanizması (Bahdanau vd., 2014).

Bilgisayarla görü alanında (Mnih vd., 2014), yüksek çözünürlüklü görüntüdeki piksel sayısını azaltarak karmaşıklığını basitleştiren farklı bir tekrarlayan yapay sinir ağı önermiştir. Önerilen bu yapay sinir ağı modeli pekiştirmeli öğrenme ile eğitilerek görsel dikkati tahmin etmektedir (Şekil-2.18). Önerdikleri modelde, görüntünün alt bölgelerine sırayla dikkat ederek görüntüyü tanımlamaktadır. Benzer şekilde Xu v.d. (Xu vd., 2015) görüntüyü otomatik olarak tanımlayan ve nasıl eğitildiğini gösteren dikkat mekanizması temelli bir model sunmuşlardır .



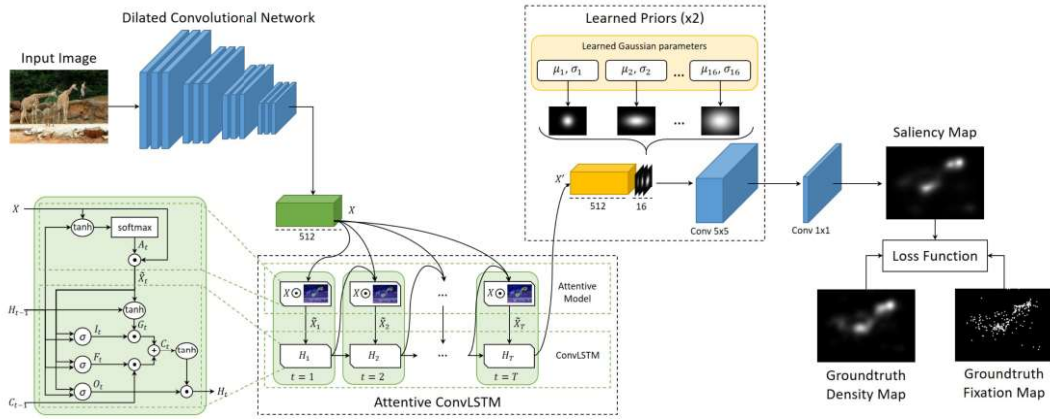
Şekil 2.18: Görüntünün tamamı yerine belirli bir alt parçasına dikkat eden bir model tanımlamışlardır (Mnih vd., 2014).

İnsan fizyolojisini esas alarak görüntüdeki belirgin özelliklerin tahmin edilmesi, bilgisayarla görü alanındaki aktif araştırma konularından birisidir. Bu alanda özellikle DYSA görüntüden belirgin özellik tahmin etmesinde durağan ve akış tipi görüntülerde yüksek başarımlar elde etmektedir.



Şekil 2.19: Belirginlik haritaları, görsel algı alanında dikkati çeken belirgin özelliklerin tahmin edilmesinde kullanımı (Simonyan vd., 2013).

Simonyan v.d. (Simonyan vd., 2013) evrişimli yapay sinir ağları ile görüntü sınıflandırmasını görselleştirmeyi öğrenen bir model geliştirmiştir. Bu model ile görüntüdeki belirginlik haritasını oluşturmaktadır (Şekil-2.19). Bahdenau v.d. tarafından ortaya konan dikkat mekanizmasını kullanarak Marcella v.d. (Cornia vd., 2018) evrişimli UKDH kullanan yapay sinir ağı tasarlamışlardır. Tasarlanan bu ağ girdi görüntüsündeki belirgin özellikleri tahmin etmekte kullanılmıştır. Önerdikleri model merkezi yanlılık (ing. central bias) Gauss parametrelerini öğrenen yapıyla dikkat çekmektedir (Şekil-2.20).

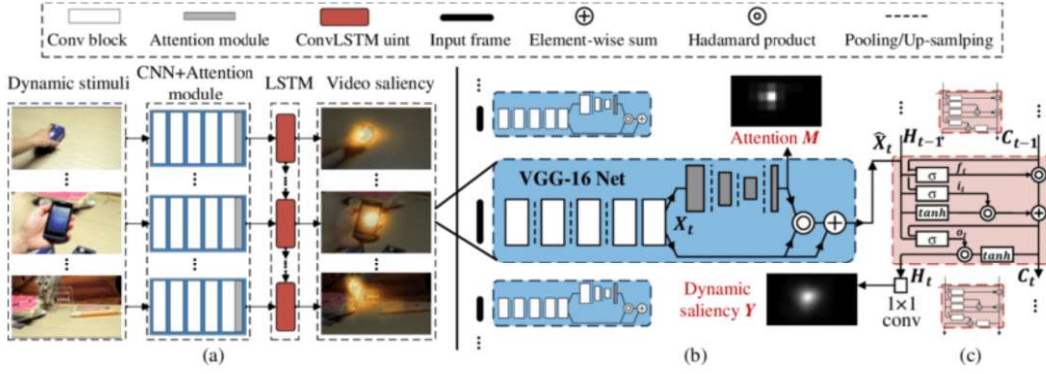


Şekil 2.20: SAM-VGG modeli durağan görüntüde belirginlik tahmin etmek için dikkat mekanizmasından faydalanması (Cornia vd., 2018).

Görüntüde belirginlik tahmin etmek için dikkat mekanizmasından faydalanan bir diğer çalışma da ACLNet çalışmasıdır (Wang vd., 2019b). Bu çalışmanın SAM çalışmasından temel farkı akış tipi görüntülere odaklanmasıdır. Ayrıca SAM modeli merkezi yanlılığı da modele dahil ederken ACLNet merkezi yanlılık hesaplamamaktadır (Şekil-2.21).

Sürüş esnasında sürücünün dikkat ettiği belirgin özellikleri belirleme alanında yapılan çalışmalar da literatürde bulunmaktadır. Örneğin, Palazzi v.d (Palazzi vd., 2017) DR(eye)VE veri kümesinden çıkardıkları kısa görüntü akışlarını kullanarak bir evrişimli yapay sinir ağı eğitmişlerdir. Eğitilen bu yapay sinir ağı sürücünün nereye ve neye baktığını tahmin etmekte kullanılmıştır.

Tawari ve Kang görsel dikkati anlamak için sürücünün bakış davranışlarını inceledikleri çalışmalarında (Tawari ve Kang, 2017) görsel dikkati modellemek için Bayesçi bir çerçeve sunmuşlardır. İlave olarak, bu çerçeve temelinde, sürüş görüntü alanındaki belirgin özellikleri tahmin etmek için tam evrişimli yapay sinir ağı geliştirmişlerdir.



Şekil 2.21: ACLNet modeli akış tipi görüntüde belirginlik tahmin etmek için dikkat mekanizmasından faydalanması (Wang vd., 2019a).

Benzer şekilde Tawari vd. (Tawari vd., 2018) belirginlik haritasını tahmin etmek için zaman dizilimli görüntü dasetini kullanan tam evreşimli bir RNN (özyinelemeli yapay sinir ağı) mimarisi önermişlerdir.

Ancak literatüre bakıldığında otonom sürüş veya sürücü destek sistemi için dikkat mekanizması öneren bir çalışmaya rastlanmamıştır.

2.4 Literatürdeki Veri Kümeleri

Görüntüdeki belirgin özelliklerin tahmin edilmesi için geliştirilen modellerin sınanabileceği farklı veri kümeleri mevcuttur. Bu veri kümeleri iki temel kategoriye ayrılır: durağan görüntü veri kümeleri ile akış tipi veri kümeleri. Belirginlik veri kümeleri, göz bebeği takip kamerası ile kaydedilen gerçek belirginlik verisi ile girdi görüntüsünden oluşmaktadır (Şekil-2.22).

Durağan görüntü için CAT2000 ve MIT1003 en önde gelen veri kümeleridir. CAT2000 veri kümesi (Borji ve Itti, 2015), 20 kategoride 2000 görüntüden oluşan ve her görüntüdeki belirgin özellikleri içeren bir veri kümesidir. Bir diğer veri kümesi de 1003 görüntüden oluşan MIT1003 (Bylinskii vd., 2015) veri kümesidir. Her iki veri kümesi de özel bir görev ve kapsam tanımlanmayan genel görüntülerden oluşmaktadır.

Akış tipi görüntü için en yaygın belirginlik veri kümeleri DHF1K, Hollywood2 ve UCF Sports Action veri kümeleridir. DHF1K veri kümesi 600 eğitim, 100 doğrulama ve 300 test akış tipi görüntüden oluşmaktadır (Wang vd., 2019b). Hollywood2 veri kümesi (Marszałek vd., 2009) 3669 akış tipi görüntüdeki insan hareketlerini içeren bir veri kümesidir. UCF Sports Action veri kümesi (Soomro ve Zamir, 2014) 10 farklı hareketi içeren 150 akış tipi görüntü içermektedir.



Şekil 2.22: Göz bebeği takip sistemi ile belirginlik haritası oluşturma.

Sürüş ile ilgili yaygın veri kümeleri var olmakla birlikte bu tez çalışması kapsamında kullanılacak sürüş esnasında sürücünün dikkat ettiği belirgin özellikleri içeren iki temel veri kümesi vardır.

(Palazzi vd., 2019) sürücülerin dikkatini tahmin etmek amacıyla aracın öz hareketini, sürücünün bakışlarını ve aracın öz bakışını içeren ham veriyi birleştirerek DR(eye)VE veri kümesini sunmuşlardır. Her biri 5 dakika uzunluğunda 74 adet akış tipi görüntüden oluşan veri kümesinde, çeşitli hava ve yol koşullarında elde edilen sürüş ve sürücü dikkat görüntüsü bulunmaktadır.

(Yu vd., 2018) BDD100K isimli geniş ölçekli sürüş görüntüleri içeren kapsamlı bir veri kümesi oluşturmuştur. Ölçeklenebilir bir etiketleme sistemi de içeren bu çalışmayla 100.000 üzerinde sürüş videosunu görüntü seviyesi etiketleme, nesne sınırları sürülebilir alan, yol çizgileri ve tam çerçeve ayrıştırma içerecek şekilde sunmuşlardır. BDD100K veri kümesi içerisindeki sürüş görüntülerinden sürücülerin dikkatini ölçtükleri yeni bir veri toplama yöntemiyle oluşturulan BDD-A ismini verdikleri bir veri kümesi oluşturulmuştur (Xia vd., 2019).

BDD-A veri kümesi farklı uzunluklardaki videolardan 3 Hz ile örneklenmiş 30161 eğitim görüntüsü, 6709 doğrulama görüntüsü ve 9834 sına görüntüsünden oluşmaktadır. DR(eye)VE veri kümesi ise 407000 eğitim görüntüsü, 37000 doğrulama görüntüsü ve 74000 sına görüntüsünden oluşmaktadır.

3. MATERYAL VE YÖNTEM

Bu bölümde, tez çalışması kapsamında önerilen modeller ve yöntemler açıklanmaktadır. Ayrıca modelleri geliştirmek ve sınamakta kullanılan teknolojiler ve veri kümeleri tanıtılmaktadır. Daha sonra gelen alt bölümlerde modellerin detaylı açıklaması verilmektedir.

3.1 Materyal

Bu tez çalışması kapsamında önerilen modeli gerçeklerken kullanılan materyal aşağıda listelenmiştir.

1. Veri kümeleri
 - a) OTA için Gazebo ortamı (Koenig ve Howard, 2004).
 - b) BDD-A veri kümesi (Xia vd., 2019).
 - c) CAT2000 veri kümesi (Borji ve Itti, 2015).
2. Yazılım Geliştirme Bileşenleri:
 - a) Python 3
 - b) Tensorflow 1.14
 - c) PyTorch 1.5
3. DYSA için Bilgisayar Kaynakları:
 - a) 8 x Nvidia Quadro P5000 GPU grafik işlemci
 - b) Intel(R) Xeon(R) CPU E5-2637 v4 @ 3.50GHz
 - c) 128GB RAM

3.1.1 Veri Kümeleri

Bu tez çalışmasında geliştirilen modellerin fren kararı vermesinde iki temel veri kümesi kullanılmıştır, ayrıca model geliştirme esnasında, belirginlik tahmin etme kısmında kavramsal doğrulama için CAT2000 veri kümesi de kullanılmıştır. CAT2000 veri kümesinde herhangi bir zenginleştirme yapılmamıştır.

1. BDD-A (Xia vd., 2019) sürüş dikkat veri kümesi.
2. OTA Gazebo benzetim ortamı veri kümesi.
3. CAT2000 (Borji ve Itti, 2015) veri kümesi.

CAT2000 veri kümesi, belirginlik tahmin etme çalışmalarında yaygın olarak kullanılan temel karşılaştırma veri kümelerindedir (Borji ve Itti, 2015). CAT2000 veri kümesi 2000 eğitim ve 2000 test verisinden oluşmaktadır. Toplamda 20 farklı kategoriye ait görüntüler, toplamda 120 gözleyici kullanılarak oluşturulmuştur (Şekil-3.1).



Şekil 3.1: CAT2000 veri kümesine ait görsellerden örnekler.

Bu veri kümelerinden BDD-A ve OTA benzetim veri kümesi, sadece görüntü haricinde aracın telemetri verisini de içermektedir. Bu tez çalışması kapsamında fren yapma kararını belirlemek için her iki veri kümesi de telemetri verisindeki anlık değişimlere göre fren yapma kararına göre etiketlenmiştir.

BDD-A verisi gerçek dış ortam araç verisinden elde edildiği için fren yapma durumuna göre etiketlemek için Federal Motor Vehicle Safety Standards (FMVSS) 49 CFR § 571.135 - Standard No. 135; Light vehicle brake systems standardındaki fren limitleri esas alınmıştır. Anılan standarda göre $-0.15g$ fren oranı, veya bir başka deyişle $-1.4715 m/sn^2$ ivmelenme durumu acil durum fren yapma olarak tanımlanmaktadır. Ne var ki, BDD-A veri kümesi genelde $5m/sn$ 'den küçük hızlar içermektedir. Bu nedenle $-1.0 m/sn^2$ ve $-0.5 m/sn^2$ olmak üzere iki eşik değeri daha belirlenmiştir. Bunun yanı sıra, yanlış etiketlemeyi engellemek için $0.5 m/sn$, $0.25 m/sn$, ve $0.1 m/sn$ olmak üzere üç tolerans değeri de belirlenmiştir. Böylelikle birer saniye aralıkla alınan telemetri örnekleri arasındaki hız değişimi seçilen eşik değerden büyük olup tolerans değerini aşmadığı süre $fren = 1$ olarak etiketleme yapılmıştır. Belirlenen eşik ve tolerans değerleri Çizelge-3.1'de verilmiştir. Ayrıca, veri etiketlerini $fren = 0$ olarak sıfırlamak için kullanılan tolerans değerleri de gösterilmiştir. Böylece, 8329 eğitim, 1079 doğrulama ve 2717 test örneği elde edilmiştir. Örnek bir veri elemanının (Çizelge-3.2) her satırı dört ardışık çerçeveye



(a) 10_06000.jpg

(b) 10_06333.jpg

(c) 10_06666.jpg

(d) 10_07000.jpg

Şekil 3.2: (Çizelge-3.2)'de belirtilen bir örnek veri sırasını oluşturan görseller.

görüntüsünden oluşmaktadır. Dördüncü çerçeveye ait fren etiketi de bu dosya yollarına eklenmiştir. İlgili örnek veri elemanlarının görselleri Şekil-3.2'de verilmiştir.

Çizelge 3.1: Veri örneklerini etiketlemede kullanılan eşik değerleri.

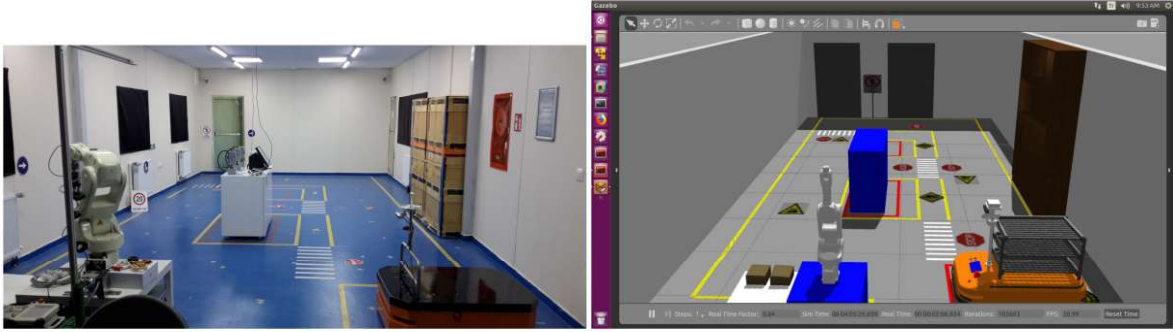
Eşik m/sn^2	Tolerans m/sn
-1.4715	0.5
-1.4715	0.25
-1.0	0.5
-0.5	0.25
-0.5	0.1

Çizelge 3.2: Etiketlenen zaman adımından önceki üç zaman adımı için örnek veri sırası.

Örnek Veri
10_06000.jpg;10_06333.jpg;10_06666.jpg; 10_07000.jpg ;brake=0
10_07000.jpg;10_08333.jpg;10_07666.jpg; 10_08000.jpg ;brake=1
10_08000.jpg;10_08333.jpg;10_08666.jpg; 10_09000.jpg ;brake=1

OTA IFARLAB test parkurunda kendisine verilen göreve bağlı olarak tanımlı rotada sürüş gerçekleştirmektedir. OTA veri kümesi üretirken gerçek ortam ve gerçekleştirme hatalarını ortadan kaldırmak için benzetim ortamı kullanılmıştır (Şekil-3.3). Benzetim ortamında insan operatör ile istenildiği zaman ve araç üzerindeki yerleşik algılayıcılar ile ön tanımlı mesafeden fren yapılabilir. Bu veriler toplanıp işlenilerek bu tez çalışmasında kullanılan veri kümesi üretilmiştir.

OTA benzetim veri kümesi Gazebo ortamında iki farklı hız ve üç farklı senaryo altında üretilen verinin işlenmesiyle üretilmiştir. BDD-A verisinden farklı olarak burada fren bilgisi araç üzerindeki mesafe ölçer algılayıcılar ile üretilmiştir. Dolayısıyla, BDD-A



(a) Gerçek ortam IFARLAB test parkurunda OTA (b) Benzetim ortamı IFARLAB test parkurunda OTA

Şekil 3.3: OTA IFARLAB Gerçek ve benzetim ortamı.

verisindeki gibi eşik ve tolerans değerleri belirlemeye gerek kalmadan fren etiket verisine sahiptir. OTA senaryolarında, $1.0 m/sn$ ve $0.5 m/sn$ hızlarla IFARLAB test parkurunda gezen araçlar için tanımlı engelleri $0.5 m$, $1.0 m$, $2.0 m$ ve $3.0 m$ uzaklıktan algılama durumuna göre fren yaptırılmıştır. Aracın ön tanımlı konumlarda engel algılamadan fren yapması, hareketsiz sabit engeli algılayarak fren yapması ve olağan ilerleme esnasında hareketli engelin araç önüne çıkması senaryolar oluşturulmuştur. Benzetim ortamında tanımlanan tüm senaryolar ve senaryo parametreleri Çizelge-3.3'de gösterilmiştir.

Gazebo ortamı verisi temizlenerek hareket başlangıcından fren kararı ile durmaya kadar olan her görüntü serisi çıkarılmıştır. Böylece toplam 24 seri elde edilmiştir. Bu serilerden 8 tanesi aracın başladığı noktadan ilk durağa kadar sürüşü içerdiği için ortaktır ve bu nedenle sadece bir defa dahil edilmiştir. Her iki hız seviyesi için telemetri verisine göre aracın fren yaptığı sürüş kareleri ile frensiz sürüş kareleri ayrılmıştır. Üretilen verinin istatistiksel olarak anlamlı bir dağılıma sahip olması için frensiz sürüş verisi, frenli sürüş oranı yaklaşık 2 : 1 olacak şekilde rassal olarak azaltılmıştır. Böylece, $0.5 m/sn$ hıza ait 261 fren yapma görüntü sayısı ve 561 olağan sürüş görüntü elde edilmiştir. Benzer şekilde, $1.0 m/sn$ hıza ait 142 fren yapma ve 317 olağan sürüş görüntü elde edilmiştir. Toplamda 403 adet fren yapma görüntüsü ile 903 adet olağan sürüş görüntü elde edilmiştir. Belirginlik tahmin etme aşaması her iki modelde de ortak olduğu için, eşzamanlı modele benzer şekilde görüntü girdisi verilmiştir. BDD-A veri kümesinden farklı olarak verilen modele ait fren yapma etiketi verilmemiştir (Çizelge-3.4). Bu şekilde verilen girdiye ait örnek bir sıra görsel Şekil-3.4'de görülebilir.

Çizelge 3.3: OTA için IFARLAB test parkuru Gazebo benzetim ortamı senaryo, hız ve fren mesafeleri.

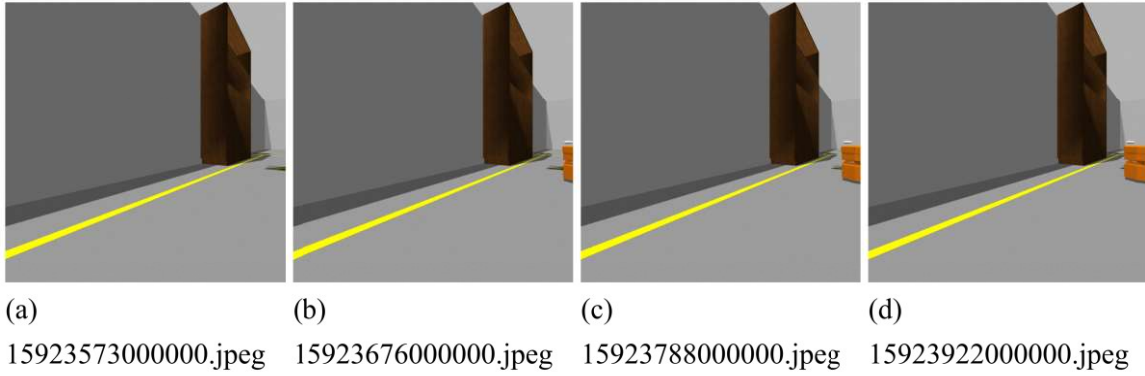
Senaryo	OTA Hızı m/sn	Fren Mesafesi m
Engelsiz Parkur	0.5	-
Engelsiz Parkur	1.0	-
Sabit Engelli Parkur	0.5	0.5
Sabit Engelli Parkur	0.5	1.0
Sabit Engelli Parkur	0.5	2.0
Sabit Engelli Parkur	0.5	3.0
Sabit Engelli Parkur	1.0	0.5
Sabit Engelli Parkur	1.0	1.0
Sabit Engelli Parkur	1.0	2.0
Sabit Engelli Parkur	1.0	3.0
Hareketli Engelli Parkur	0.5	0.5
Hareketli Engelli Parkur	0.5	1.0
Hareketli Engelli Parkur	0.5	2.0
Hareketli Engelli Parkur	0.5	3.0
Hareketli Engelli Parkur	1.0	0.5
Hareketli Engelli Parkur	1.0	1.0
Hareketli Engelli Parkur	1.0	2.0
Hareketli Engelli Parkur	1.0	3.0

Çizelge 3.4: OTA veri kümesinden örnek bir veri sırası.

Örnek Veri
15923573000000.jpeg;15923676000000.jpeg;15923788000000.jpeg;15923922000000.jpeg
16099620000000.jpeg;16099731000000.jpeg;16099963000000.jpeg;16100104000000.jpeg
15905440000000.jpeg;15905625000000.jpeg;15905802000000.jpeg;15905929000000.jpeg

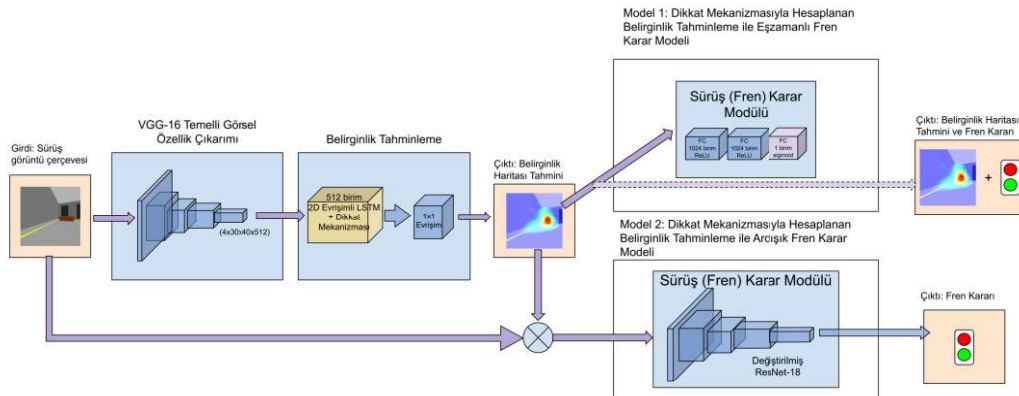
3.2 Dikkat Mekanizmasıyla Hesaplanan Belirginlik Tahmin Etmesi ile Fren Karar Modeli

Bu tez çalışmasında sürüş görüşündeki belirgin özelliklere dikkat ederek fren kararı üreten bir model önerilmiştir. Tez çalışmasında önerilen modelin genel şeması Şekil-3.5’de gösterilmektedir.



Şekil 3.4: (Çizelge-3.4)'de belirtilen bir örnek veri sırasını oluşturan görseller.

Önerilen modelin iki farklı yaklaşımı mevcuttur. İlk olarak, belirginlik haritası tahmin etme ile fren kararının eşzamanlı üretilip aynı eğitim süreciyle eğitilebildiği yaklaşım açıklanmıştır. Bu yaklaşım geliştirilirken, belirginlik tahmin etme çalışmalarında sıkça kullanılan merkezi yanlılık ta incelenmiştir. Deneysel çalışmalar esnasında (bkz. 4.1) bu yaklaşımın belirginlik tahmin etmede başarımı yüksek ancak fren kararı açısından başarımı düşük olması nedeniyle yeni bir yaklaşım geliştirilmiştir. Bu yeni yaklaşım, önce belirginlik tahmin etmekte, ardından girdi görüntüsünün tahmin edilen belirginlik haritasıyla maskelenmesi ile elde edilen girdiden yeniden özellik çıkarılarak ikili fren karar üretmesini sağlamaktadır.



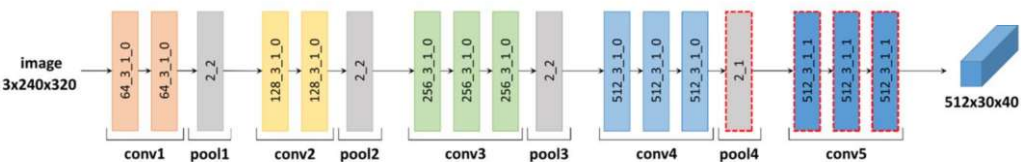
Şekil 3.5: Tez çalışmasında önerilen dikkat mekanizmasıyla hesaplanan belirginlik haritası ile fren karar model yaklaşımlarını gösterir genel şema.

Takip eden alt bölümlerde öncelikle her iki yaklaşım tarafından ortak kullanılan, girdi görüntüsünden özellik çıkarımı ve dikkat mekanizması kullanarak belirgin özellik tahmin etme alt bileşenleri açıklanmıştır.

3.2.1 Görsel Özellik Çıkarımı

Son yıllarda DYSA alanında yaşanan gelişmelere paralel olarak bilgisayarla görü alanında özelliklerin çıkarımı için DYSA'lar yaygın olarak kullanılmaya başlamıştır. Görsel özellik çıkarımı için kullanılan modeller genellikle ImageNet (Russakovsky vd., 2015) bilgisayarla görü kıyaslamasında en iyi sonuç alınan modellerin değiştirilmesiyle elde edilir. Literatürde en yaygın kullanılan iki farklı model VGG-16 (Simonyan ve Zisserman, 2014) ve ResNet (He vd., 2016) modelleridir. Bu tez çalışması kapsamında görüntüden özellik çıkarmak için geliştirilen model de VGG-16 ve ResNet-18 modellerinin değiştirilmiş hallerinden faydalanmıştır.

ImageNet yarışmasında 1000 farklı sınıfa ait görselin doğru sınıflandırılması yapılmaktadır. Bu nedenle anılan modeller, ilgili yarışmadaki halleriyle özellik çıkarmak için kullanılamaz. Özellik çıkarmak için yapılan en önemli değişiklik, ilgili modellerin sınıflandırmada kullanılan son tam bağlı katmanlarını atmaktır. Bu sayede ilgili modellerin girdi görüntüsünden elde ettikleri tüm özellik tensörlerine erişilebilir. Bu tensörler genelde boyut olarak çok küçük olmaktadır ve bu nedenle belirginlik tahmin etme amacıyla kullanılamaz. Bu sorunu düzeltmek için GEA (Yu ve Koltun, 2015) kullanılmıştır. GEA evrişim işlemini aralıklı olarak gerçekleştirmesi nedeniyle hıza aşağı ölçeklemeyi engellemektedir. Böylece, örneğin, 320x240x3 boyutlarındaki RGB formatında girdi görüntüsüne karşılık 40x30 boyutlarında 512 adet tensör elde edilmektedir (Şekil-3.6).

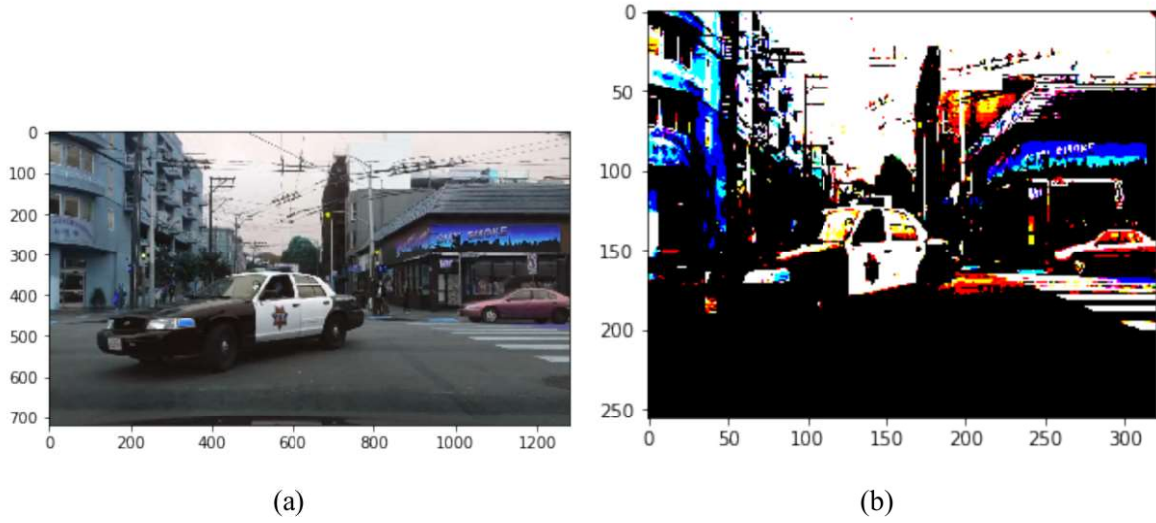


Şekil 3.6: GEA kullanılarak değiştirilmiş VGG-16 temelli görüntü özellik çıkarıcı.

Görsel özellik çıkarımı için girdi görüntüsünü modele verirken, ImageNet veri kümesinde eğitilen modellerin uyguladığı gibi, ortalama piksel değeri girdi görüntüsünden çıkarılmaktadır. Bu işlem eğitim verisini normalleştirmek ve durağanlaştırmak için yapılmaktadır (Şekil-3.7).

3.2.2 Belirginlik Tahmin Etme

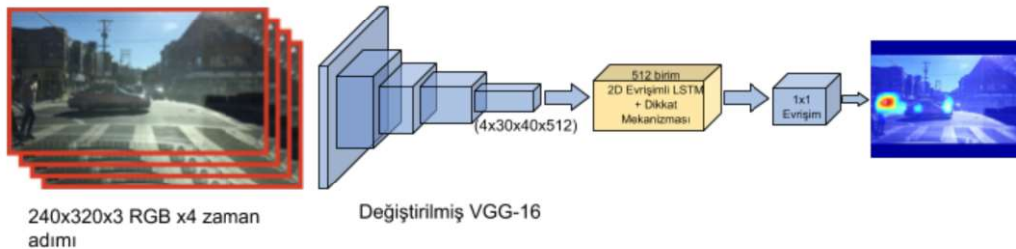
Belirginlik tahmin etme için bu alandaki mevcut durumda en iyi sonuçları elde etmiş modellerden olan SAM-VGG (Cornia vd., 2018) ve ACLNet (Wang vd., 2019b) kullanılmıştır. İki model arasındaki temel fark SAM-VGG durağan görüntü veri



Şekil 3.7: Girdi görüntüsü ve modele verilen ortalama piksel değeri çıkarılmış girdi görüntüsü.

kümelerinde sınılanmışken ACLNet akış tipi görüntü veri kümelerinde sınılanmıştır. Öte yandan DYSA mimarisi olarak iki model arasındaki en önemli farklılık SAM-VGG modelinde merkezi yanlılık kullanılmış olmasıdır. Merkezi yanlılık insan görme fizyolojisindeki görüntünün ortasına odaklanma işlevini modeller. Bunun haricinde iki model arasında yaklaşım olarak çok temel bir farklılık bulunmamaktadır.

Belirginlik tahmin etme için kullanılan bu iki model, öncelikle alt bölüm 3.2.1’de bahsedilen görüntü özellik çıkarıcılarından VGG-16 temelli özellik çıkarıcı kullanmaktadır. Çıkarılan özellik matrisleri daha sonra dikkat mekanizması uygulanmış 512 birimlik evrişimli UKDH (Xingjian vd., 2015)’den geçirilerek belirginlik haritaları tahmin etmede kullanılacak özellik matrisleri elde edilmektedir. Belirginlik matrisleri 1x1’lik evrişim işlemi ile bütünleştirilerek nihai belirginlik harita tahmini elde edilmektedir.



Şekil 3.8: Evrişimli UKDH ve dikkat mekanizması kullanarak belirginlik tahmin etme.

Bu çalışmada önerilen modelde, x , $4 \times 248 \times 320 \times 3$ RGB görüntü ve y $4 \times 32 \times 40 \times 1$ boyutunda tahmin edilen belirginlik haritasıdır, c_t kapsam vektörü, kodlanan bilginin (yani kodlayıcının h_i çıktı vektörlerinin) ağırlıklandırılmış toplamıdır. (denklem 3.1). Burada ağırlıklar, $\{\alpha_{t,i}\}$, $t = 1, \dots, n$ zaman adımlarında hesaplanan dikkat puanlarıdır. Önerilen modelde 1 sn zaman aralığına karşılık gelecek şekilde $n = 4$ seçilmiştir (denklem 3.2). Daha büyük bir n daha geniş zaman aralığına ve daha çok parametreye karşılık gelir. Puanlama fonksiyonu $driverAttn(y_t, x_i)$, girdi i ile çıktı t arasındaki ilişkiyi ölçmektedir (denklem 3.3). Puanlama fonksiyonu denklem 3.4'te gösterilen Bahdenau puanını kullanmaktadır.

$$c_t = \sum_{i=1}^n \alpha_{t,i} h_i \quad (3.1)$$

$$\alpha_{t,i} = driverAttn(y_t, x_i) \quad (3.2)$$

$$= \frac{\exp(\text{score}(s_{t-1}, h_i))}{\sum_{i'=1}^n \exp(\text{score}(s_{t-1}, h_{i'}))} \quad (3.3)$$

$$\text{score}(s_t, h_i) = v_a^\top \tanh(W_a[s_t; h_i]) \quad (3.4)$$

$s_t = f(s_{t-1}, y_{t-1}, c_t)$, t zaman adımındaki gizli durum vektörü, W_a ve v_a UKDH ağının ağırlık matrisleridir.

Model uçtan uca geri yayılım algoritması ile eğitilmektedir. Geri yayılım algoritması için kullanılan kayıp fonksiyonları, Kullback-Leibler uzaklığı (KLD) ve korelasyon katsayısı (CC) temelli kayıp fonksiyonudur. Bu kayıp fonksiyonları (Bylinskii vd., 2018)'de incelendiği şekliyle belirginlik açısından tercih edilen kayıp fonksiyonlarıdır (Şekil-3.9).



Şekil 3.9: Sol: Girdi görüntüsü, Orta: Gerçek belirginlik, Sağ: tahmin edilen belirginlik

Burada L_{KL} Kullback-Leibler uzaklığı (KLD) temelli kayıp fonksiyonu olup denklem 3.5'te verilmektedir. Bu, tahmin edilen belirginlik haritasının gerçek belirginlik haritasıyla olan uzaklığını tespit eder.

$$L_{KL}(y_{true}, y_{pred}) = \sum_i y_{i,true} \log \left(\frac{y_{i,true}}{y_{i,pred} + \epsilon} + \epsilon \right) \quad (3.5)$$

Denklem 3.5'te i , piksel indeksi ve μ düzenleyici parametredir.

Belirginlik tahmin etmede yaygın olarak kullanılan bir diğer ölçüt Korelasyon Katsayısı'dır (CC) (denklem 3.6). Korelasyon Katsayısı r , $[-1, 1]$ arasında tanımlı olduğu için kayıp fonksiyonunda $-r$ biçiminde kullanılmıştır. Kovaryans σ olmak üzere:

$$L_{CC} = -\frac{\sigma(y_{true}, y_{pred})}{\sigma(y_{true})\sigma(y_{pred})} \quad (3.6)$$

Böylece, sürüş dikkat modeli için kayıp fonksiyonu denklem 3.7 ile tanımlanmıştır.

$$L_{DAM} = L_{KL}(y_{true}, y_{pred}) + L_{CC}(y_{true}, y_{pred}) \quad (3.7)$$

Bu kayıp fonksiyonlarına ilave olarak, belirginlik tespiti modelini geliştirirken yapılan doğrulamalarda kullanılacak şekilde Normalleştirilmiş Belirginlik Taramayolu (NSS) kayıp fonksiyonu geliştirilmiştir. Geliştirilen kayıp fonksiyonu, ikili gerçek bakış sabitleme haritası gerektirdiği için sadece CAT2000 belirginlik veri kümesinde kullanılmıştır (denklem 3.8). CAT2000 veri kümesinde NSS kayıp fonksiyonunda kullanılmak üzere ikili gerçek bakış sabitleme verisi de sunulmuştur. Bu tez çalışmasında kullanılan diğer veri kümelerinde bu veri olmadığı için NSS temelli kayıp fonksiyonu kullanılamamıştır.

$$L_{NSS} = \frac{1}{N} \sum_i \frac{y_{i,pred} - \mu(y_{pred})}{\sigma(y_{pred})} \cdot y_{i,fix} \quad (3.8)$$

Burada i , i^{nci} pikselin endeksi, ve N toplam sabitlenmiş bakış alan pikselini temsil etmektedir.

Herhangi bir görev tanımı olmadan serbest görüde belirginlik çalışmalarının çoğu görüntünün ortasında belirginlik aramayı sağlayan merkezi yanlılık kullanmaktadır. Ancak Kuen v.d. (Kuen vd., 2016) tarafından açıklandığı üzere merkezi yanlılık, belirli bir kapsamda belirginlik tahmininde (yani yukarıdan aşağı belirginlik) çok az katkı sağlarlar. Merkezi yanlılık literatürdeki çalışmalarda ön tanımlı Gauss parametreleriyle kullanılmaktadır.

Belirginlik tahmin etmede dikkat mekanizması içeren çalışmasında Cornia v.d. (Cornia vd., 2018) ön tanımlı Gauss merkezi yanlılık parametreleri yerine öğrenen Gauss merkezi yanlılık parametreleri önermiştir. Bu doğrultuda, sürüşü görevi özelinde bu parametrelerin katkısı olup olmadığını incelemek açısından Gauss parametreleriyle birlikte Gauss çekirdekli Radyal Bazlı Fonksiyon (RBF) geliştirilmiş ve denenmiştir.

Gauss parametreleri veriden ortalama değer ve kovaryans öğrenir. İki boyutlu bir Gauss fonksiyonu,

$$f(x, y) = \frac{1}{2\pi\sigma_x\sigma_y} \exp \left(- \left(\frac{(x - \mu_x)^2}{2\sigma_x^2} + \frac{(y - \mu_y)^2}{2\sigma_y^2} \right) \right) \quad (3.9)$$

olarak tanımlanır.

Öte yandan, merkezi yanlılık parametresi olarak Gauss çekirdekli RBF kullanılmadaki motivasyon, RBF'nin gerçek değerli bir fonksiyon olup merkeze olan uzaklığa bağlı olmasıdır (denklem 3.10). Denklem 3.11 ile gösterilen Gauss çekirdeği ile kullanıldığında, Broomhead ve Lowe tarafından (Broomhead ve Lowe, 1988) tanımlandığı üzere genel bir fonksiyon yaklaşıklarıdır (3.12).

$$\varphi(x) = \varphi(\|x - c\|) \quad (3.10)$$

$$\varphi(r) = e^{-(\epsilon r^2)}, r = \|x - x_i\| \quad (3.11)$$

$$y(x) = \sum_{i=1}^N \omega_i \varphi(\|x - x_i\|) \quad (3.12)$$

Burada c bir merkez, ω_i , i^{nci} girdinin ağırlığı, ve ϵ biçim parametresidir. Biçim parametresi ϵ sifıra yaklaştıkça daha geniş ve düz bir temel fonksiyonuna dönüşür. Gerçeklemede $\epsilon = 0.5$ olarak seçilmiştir.

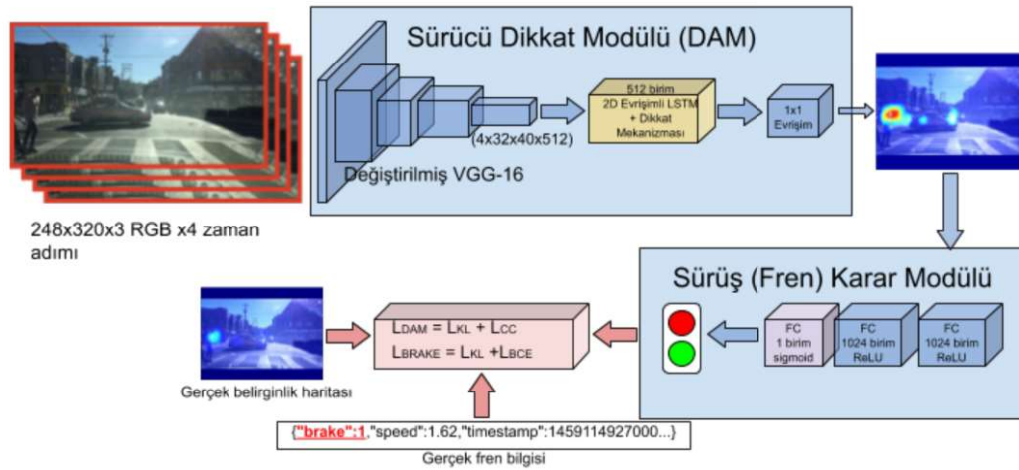
Öğrenilen merkezi yanlılık, dikkat mekanizması ile hesaplanan belirginlik tensörlerine eklenerek kullanılır. Bu tez çalışmasında Cornia v.d.'nin çalışmasıyla karşılaştırma yapabilmek için olarak 16 veya 32 adet Gauss parametresi ile 16 veya 32 RBF parametresi seçilmiştir.

3.2.3 Belirginlik Tahmin Etmesiyle Fren Yapma Kararı Verme

Belirginlik tahmin etmesiyle fren yapma kararı vermekte, Şekil-3.5’de gösterildiği gibi iki farklı model önerilmektedir. İlk model, belirginlik tahmin etmesi ile fren kararı vermenin eş zamanlı eğitildiği modeldir. Önerilen eş zamanlı model, belirginlik haritası ile fren kararını birlikte üretir. Bu model eğitilirken gerçek belirginlik haritası ile fren etiketi birlikte kullanılır. Daha sonra takip eden alt bölümde, belirginlik haritası tahmin edildikten sonra fren yapma kararının üretildiği ardışık model açıklanmaktadır. Bu ardışık model, önce belirginlik haritası tahmin etmek, daha sonra girdi görüntüsü tahmin edilen belirginlik haritası ile maskelenerek fren kararı vermek için kullanılmaktadır.

3.2.3.1 Eş zamanlı Model

Fren yapma kararı vermek için belirginlik haritası tahmin edildikten sonra fren kararı ile tahmin edilen belirginlik haritası birlikte karar modeline verilir (Şekil-3.10). Karar modeli tam bağlı bir DYSA’dır.



Şekil 3.10: Belirginlik haritası tahmin etme ile fren yapma kararı üreten model.

Bu modelde alt bölüm 3.2.1’deki VGG-16 temelli görsel özellik çıkarıcı ile alt bölüme 3.2.2’teki belirginlik tahmin etme kullanılmıştır. tahmin edilen belirginlik haritası üç aşamalı tam bağlı DYSA’dan geçirilerek fren yapma kararı üretilmektedir. Modelin eğitimi esnasında gerçek belirginlik haritası kaybı ile gerçek fren yapma kararı birlikte geri beslenerek kayıp

fonksiyonu en küçüklenir. Alt bölüm 3.2.2'teki kayıp fonksiyonuna ilave olarak sürüş yapma kararı kayıp fonksiyonu olarak denklem 3.13'te verilen İkili Çapraz Entropi (ing. Binary Crossentropy) kullanılmıştır.

$$\begin{aligned} L_{BCE} &= -\frac{1}{n} \sum_{i=1}^n [y_i \log(p_i) + (1 - y_i) \log(1 - p_i)] \\ &= -\frac{1}{n} \sum_{i=1}^n \sum_{j=1}^m y_{ij} \log(p_{ij}) \end{aligned} \quad (3.13)$$

Bu denklemde fren yapma ve yapmama kararı için ikili sınıfı belirtmek üzere $m = 2$ seçilmiştir. i örnek endeksi ve j karar sınıf endeksidir. Fren yapma kararı kayıp fonksiyonu denklem 3.14'te tanımlanmıştır:

$$\begin{aligned} L_{BRAKE} &= L_{KL}(y_{true}, y_{pred}) \\ &\quad + L_{BCE}(y_{true}, y_{pred}) \end{aligned} \quad (3.14)$$

Sonuç olarak, model uçtan uca denklem 3.15'de gösterilen kayıp fonksiyonu ile eğitilmiştir:

$$\begin{aligned} L_{TOTAL}(y_{true}, y_{pred}) &= L_{DAM}(y_{true}, y_{pred}) \\ &\quad + L_{BRAKE}(y_{true}, y_{pred}) \end{aligned} \quad (3.15)$$

Önerilen model eğitilirken geri yayılım yöntemiyle en iyilenebilir. En iyileme algoritması olarak UMEİ veya SGI kullanılabilir. Eğitim süreci *batchsize* ve *learning_rate* parametreleri ile en iyilenebilir. *batchsize* parametresi ile daha fazla sayıda girdi verisi aynı epoch içinde işlenebilir ve daha hızlı sonuç alınabilir, ancak daha fazla kaynak kullanımına yol açar. *learning_rate* parametresi de öğrenme hızını belirler. Daha küçük *learning_rate* ile daha uzun sürede hata en küçüklenir, daha büyük *learning_rate* ise lokal minimuma takılma riski taşır.

Önerilen modelin bir faydası, sürüş görev kısıtı altında tahmin edilen belirginliği fren kararı ile birlikte, uçtan uca tek bir eğitim süreciyle eğitilebilmesidir. Ayrıca sadece

fren karar bilgisi kapsamı haricinde sürüşe dair görsel algıdaki farklı belirginlikleri de genelleştirmeyi aynı eğitim sürecinde öğrenerek fren kararına aşırı uyarılığı engellemesidir. Modelin dezavantajı ise, eğitim sürecinde belirginlik veya fren karar uzayı içerisinde lokal en iyi sonuçta takılma riskidir.

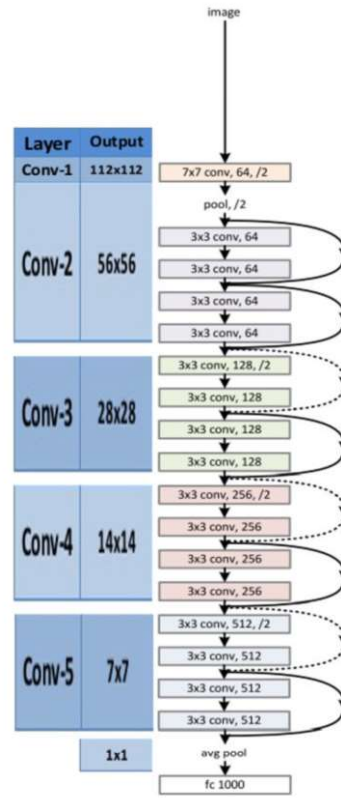
Bir sonraki alt bölümde, eğitim sürecinde lokal en iyi sonuçta takılmayı önlemek için belirginlik tahmin etme ile fren kararını ardışık olarak iki ayrı eğitim sürecinde öğrenen bir model önerilmektedir.

3.2.3.2 Ardışık Model

Önceki alt bölümde önerilen eşzamanlı model iyileştirilerek girdi görüntüsünün tahmin edilen belirginlik ile maskelenmesi ile de fren yapma kararı verilmesi mümkündür. Belirginlik haritası tahmin etme ile fren yapma kararı verme işlemlerinin ardışık olarak kullanılmasının nedeni, modelin eğitim esnasında belirginlik ile fren yapma kararını birlikte en iyileme sırasında ortam şartlarına göre lokal en iyi sonuçta takılma riskini ortadan kaldırmaktır. Modelin dezavantajı ise daha yüksek hiperparametre sayısı ve iki farklı eğitim sürecinin kullanılmasıdır.

Önerilen yeni modelde belirginlik haritası tahmini ile maskelenen girdi görüntüsü, fren yapma kararı üretebilmesi için değiştirilmiş ResNet-18 (He vd., 2016) modeline verilir. Böylece girdi görüntüsünden çıkarılan özellikler, girdi görüntüsündeki belirginlikler ile ağırlıklandırılarak fren yapma kararı üretilir. Fren yapma kararı verme sürecinde ResNet-18 modeli seçilmesinin temel nedeni ResNet mimarisinin temeli olan kalıntı fonksiyonu ile özdeş bağıntıyı birlikte kullanarak kaybolan gradyan problemini çözmesidir. ResNet-18 modelinin ImageNet yarışmasına uygun olarak 1000 farklı sınıflandırma yapan son katmanı ikili sınıflandırma yapacak şekilde değiştirilmiştir (Şekil-3.11).

Yeni model mimarisinde belirginlik haritası fren yapma kararıyla eş zamanlı olarak eğitilmez. Bu modelde önce girdi görüntüsünden belirginlik haritası tahmin edilir, daha sonra girdi görüntüsünün tahmin edilen belirginlik haritasıyla maskelenmiş hali son katmanı ikili sınıflandırmaya göre değiştirilmiş ve ImageNet ile önceden eğitilmiş ResNet-18 modeline girdi olarak verilerek fren kararı tahmini üretilir (Şekil-3.13). Üretilen fren karar tahmini denklem 3.16 ile verilen kayıp fonksiyonu ile geri yayılım yöntemiyle eğitilir.

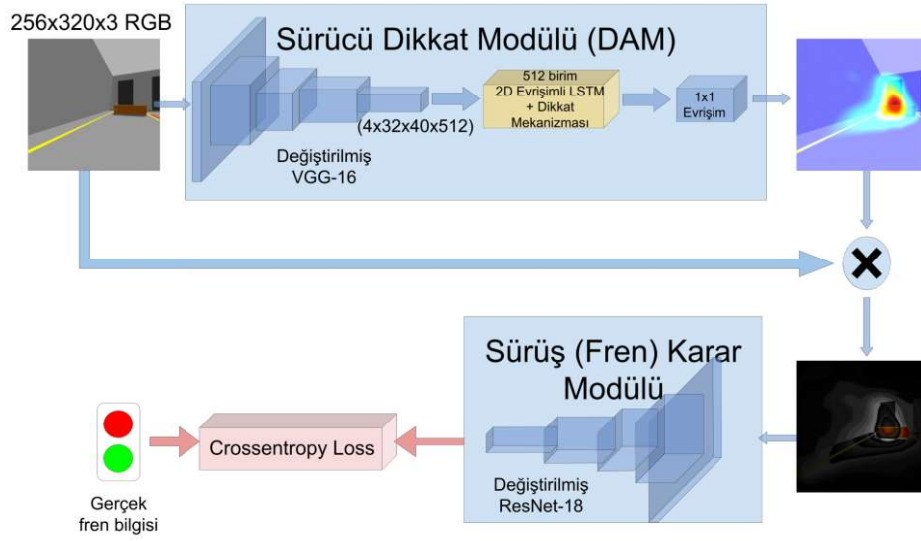


Şekil 3.11: Orijinal ResNet-18 model mimarisi.

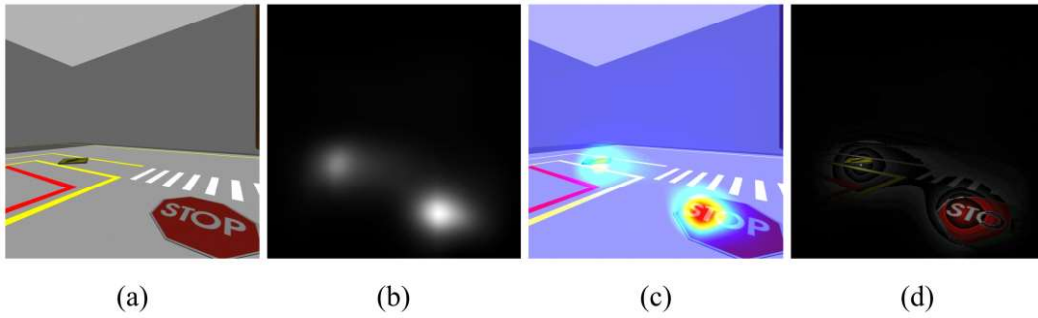
$$\begin{aligned}
 L_{CE}(x, class) &= -\log \left(\frac{\exp(x[class])}{\sum_j \exp(x[j])} \right) \\
 &= -x[class] + \log \left(\sum_j \exp(x[j]) \right)
 \end{aligned}
 \tag{3.16}$$

Nihai modelin mimari şeması Şekil-3.12’de gösterilmektedir. Bu yeni model fren yapma kararı vermek üzere maskelenen girdi görüntüsünden tekrar özellik çıkarmaktadır.

Önerilen model eğitilirken üretilen veri kümesinden eğitim-doğrulama verisi oluşturulmuştur. Ayrıca *batchsize* ve *learning_rate* öğrenim oran parametresi ile eğitim süreci en iyilenebilir. Modelin geri yayılım algoritması ile en iyilenmesinde SGI algoritması kullanılmıştır. *batchsize* parametresi ile daha fazla sayıda girdi verisi aynı epoch içinde işlenebilir ve daha hızlı sonuç alınabilir, ancak daha fazla kaynak kullanımına yol açar. *learning_rate* parametresi de öğrenme hızını belirler. Daha küçük *learning_rate* ile daha uzun sürede hata en küçüklenir, daha büyük *learning_rate* ise lokal minimuma takılma riski taşır.



Şekil 3.12: Belirginlik haritası tahmin etme ile fren yapma kararı üreten model.



Şekil 3.13: Girdi görüntüsünün belirginlik tahmin etmesi ile maskelenmesi. (a) Girdi görüntüsü, (b) tahmin edilen belirginlik haritası, (c) Girdi görüntüsü üzerinde belirginlik, (d) Belirginlik haritasıyla maskelenen girdi görüntüsü.

Önerilen ardışık modelin avantajı girdi görüntüsünü belirginlikle maskeleyerek fren kararında etkili olan belirgin özelliklere öne çıkarmasıdır. Böylece belirginlik tahmin etme ile fren kararı en iyileme sürecinin lokal en iyi sonuçta takılması önlenmektedir. Modelin fren karar sürecinde VGG-16 temelli özellik çıkarıcı yerine ResNet-18 temelli özellik çıkarıcı kullanması tamamen ImageNet sınıflandırma başarımlarındaki farklılıkları nedeniyledir. Daha çok hiperparametreye sahip ResNet-18 modeli, belirginlik tahmin etme açısından bariz bir iyileşme sunmamasına rağmen, ImageNet sınıflandırma başarımları VGG-16'dan daha iyidir.

4. BULGULAR VE TARTIŞMA

Çalışmada önerilen belirginlik tabanlı fren karar modellerine ait deneysel sonuçlar aşağıda verilmektedir. Bu kısımda önerilen modeller literatürdeki veri tabanlarına ve OTA için hazırlanan veri kümesine uygulanmıştır.

4.1 Dikkat Mekanizmasıyla Hesaplanan Belirginlik ile Eşzamanlı Fren Karar Model Bulguları

Önerilen eşzamanlı model eğitilirken geri yayılım yöntemiyle en iyilenmiştir. En iyileme algoritması olarak UMEİ kullanılmıştır. Öğrenme oranı *learning_rate* ise yaklaşım hızı incelenerek 0.001 olarak seçilmiştir. UMEİ algoritması geniş aralıklı gradyan ve gürültü durumlarında SGI gibi diğer geri yayılım en iyileme algoritmalarından daha iyi sonuç vermektedir.

Önerilen eşzamanlı modelde DHF1K (Wang vd., 2019b) veri kümesinde ön eğitilmiş ACLNet (Wang vd., 2019a) temelli belirginlik tahmin etme de referans olarak kullanılmıştır. Ayrıca ACLNet temelli belirginlik tahmin etme BDD-A veri kümesinde de eğitilerek referans modelle karşılaştırılmıştır. Merkezi yanlılık içeren SAM modelleri de (Cornia vd., 2018) karşılaştırma amacıyla BDD-A test veri kümesinde belirginlik tahmin etme için kullanılmıştır. Ön eğitimsiz olarak BDD-A veri kümesinde kullanılırken, modellerin referans eğitim süreçlerinde kullanılan parametreler korunmuştur.

Bölüm-3.2.3.1’de önerilen eşzamanlı modelin sadece belirginlik tahmin etme için elde ettiği sonuçlar Çizelge-4.1’de gösterilmiştir. (Xia vd., 2019) referans sonuçları en üst satırda verilmiştir. BDD-A veri kümesinde eğitilen ACLNet temelli model Kullback-Leibler uzaklığı metriğinde en yüksek sonucu elde etmiştir. Korelasyon katsayısı metriğinde BDD-A referans modeli daha iyi sonuç vermiştir. Ancak elde edilen korelasyon katsayısı, $CC = 0.5484$ sonucu ön eğitilmiş modellere göre daha iyi sonuç vermiştir. Ön eğitilmiş modeller, BDD-A veri kümesinde eğitilen modele göre daha kötü sonuç üretmiş olsa da referans modelden çok uzak olmadığı dikkat çekmektedir. Önerilen modelden belirginlik tahmin etme için bazı girdi görselleri için elde edilen örnek sonuçlar Şekil-4.1’de gösterilmiştir.

Çizelge 4.1: BDD-A veri kümesinde eğitilen önerilen model ile ön eğitilmiş referans modellerin elde ettikleri belirginlik tahmin etme sonuçları.

Model	KLD	CC
BDD-A referans (Xia vd., 2019)	1.24	0.59
Önerilen model	0.1092	0.5484
Ön eğitilmiş ACLNet (Wang vd., 2019b)	1.2579	0.4734
Ön eğitilmiş SAM-VGG (Cornia vd., 2018)	1.3853	0.4439
Ön eğitilmiş SAM-RESNET (Cornia vd., 2018)	1.3856	0.4403



(a) Girdi



(b) Gerçek görüntü



(c) ACLNet tahmini



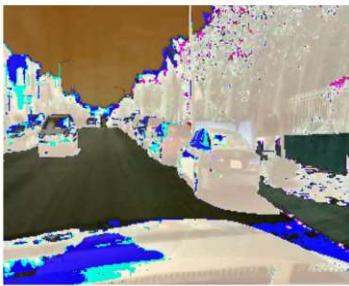
(d) Girdi



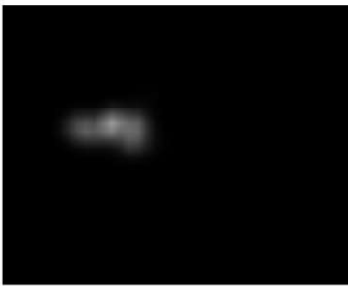
(e) Gerçek görüntü



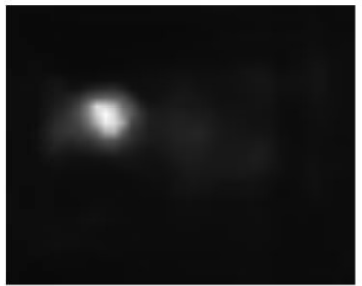
(f) SAM-VGG hatalı tahmini



(g) Girdi



(h) Gerçek görüntü



(i) SAM-VGG tahmini

Şekil 4.1: Belirginlik tahmin etme örnek bulguları.

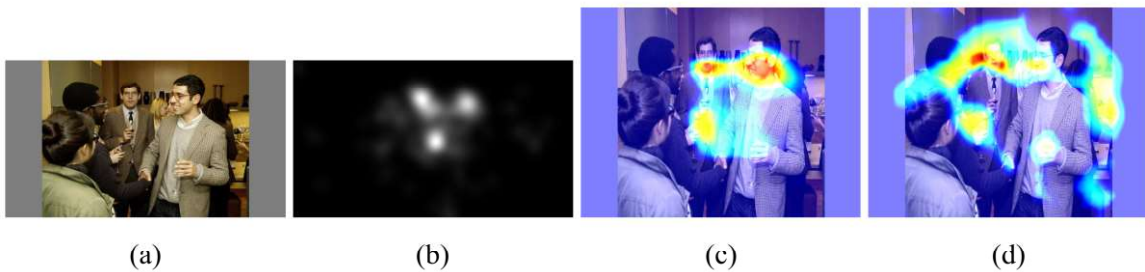
Bununla birlikte önerilen eşzamanlı model kavramsal doğrulama için CAT2000 belirginlik veri kümesinde de denenmiştir. Bu doğrulama çalışması esnasında merkezi yanlılık ta incelenmiştir. Bu çalışmanın sonuçları Çizelge-4.2’te ve Çizelge-4.3’te

Çizelge 4.2: CAT2000 veri kümesinde belirginlik tahmin etme sonuçları.

Model	CC	KLD	NSS
SAM-VGG (Cornia vd., 2018)	0.89	0.54	2.38
CC-KLD-NSS-G16	0.6604	0.6513	0.9140
CC-KLD-NSS-G32	0.6446	0.6508	0.8928
CC-KLD-NSS-RBF32	0.1685	0.6764	0.8784
CC-KLD-NSS-RBF16	0.1344	0.6696	0.8777
KLD-NCB	0.5079	0.7033	0.7842
CC-KLD-NSS-NCB	0.5380	0.7421	0.7792
CC-NCB	0.5426	0.8981	0.7763
NSS-NCB	0.1522	9.2005	0.1877

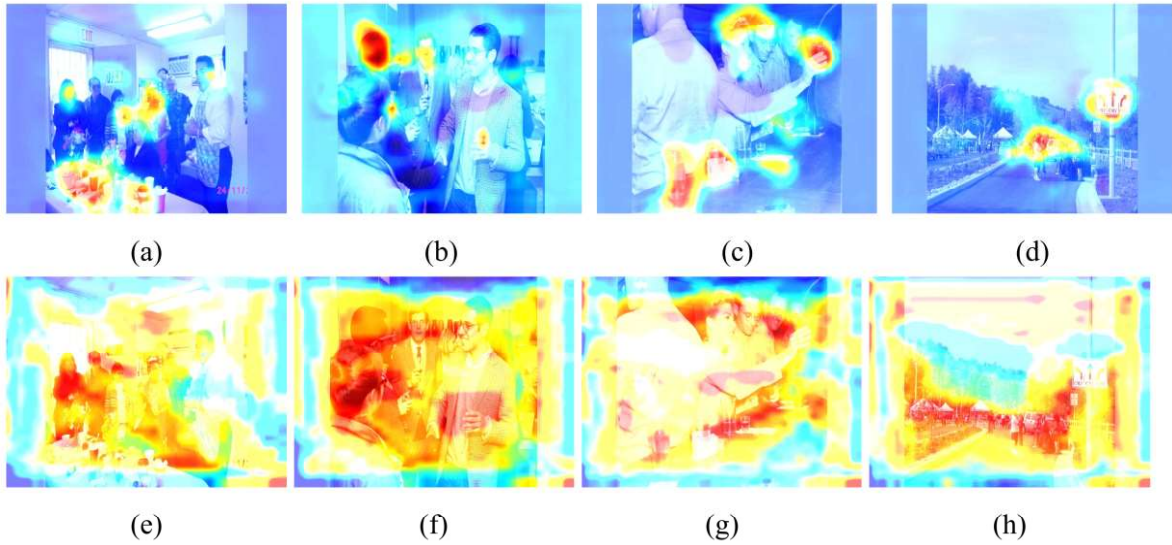
verilmiştir. Sonuçlar, (Bylinskii vd., 2018)'e göre NSS ölçütüne göre sıralanmıştır. Çizelgelerde CC, KLD ve NSS seçilen kayıp fonksiyon türünü, G ve RBF uygulanan merkezi yanlılık parametre türünü ve takip eden sayı merkezi parametre sayısını göstermektedir. NCB merkezi yanlılık uygulanmadığını göstermektedir. En üst satırlarda referans sonuçlar verilmiştir.

CAT2000 veri kümesi karşılaştırmasına katılan modellerde olduğu gibi Gauss merkezi yanlılık belirginlik tahmin etme başarımını iyileştirmektedir. RBF merkezi yanlılık parametreleri korelasyon katsayısı CC ölçütüne göre kötü sonuç verse de diğer ölçütler açısından iyi sonuç vermiştir (Şekil-4.2). Öte yandan sadece NSS ölçütü kullanıp merkezi yanlılık kullanmayan model belirginlik harita dağılımını genelleştirmeyi başaramamıştır. Bununla birlikte bu sonuçlar CAT2000 deki mevcut en iyi modellerden daha kötü sonuç vermesine rağmen ortalamanın üzerinde başarımlar elde etmiştir (Borji ve Itti, 2015).



Şekil 4.2: CAT2000 veri kümesinden örnek sonuçlar: (a) Girdi görüntüsü, (b) Gerçek belirginlik, (c) CC-KLD-NSS-G16 modeli ile elde edilen belirginlik tahmini, (d) NSS-NCB modeli ile elde edilen belirginlik tahmini.

CAT2000 veri kümesinde merkezi yanlılık haricinde (Cornia vd., 2018) ve (Wang vd., 2019b) çalışmaları olduğu gibi kayıp fonksiyonlarını ağırlıklandırarak yapılan deneylerde, tek bir kayıp fonksiyonu ve merkezi yanlılık kullanmayan modellerin başarımları daha kötü olmaktadır. Deneylerden elde edilen bazı örnek sonuçlar Şekil-4.3'de gösterilmiştir. Şekilde üst sırada merkezi yanlılık ve tüm kayıp fonksiyonlarını kullanan bir model sonucu, alt sırada merkezi yanlılık kullanmayan ve tek bir kayıp fonksiyonu kullanan bir model sonucu verilmiştir.

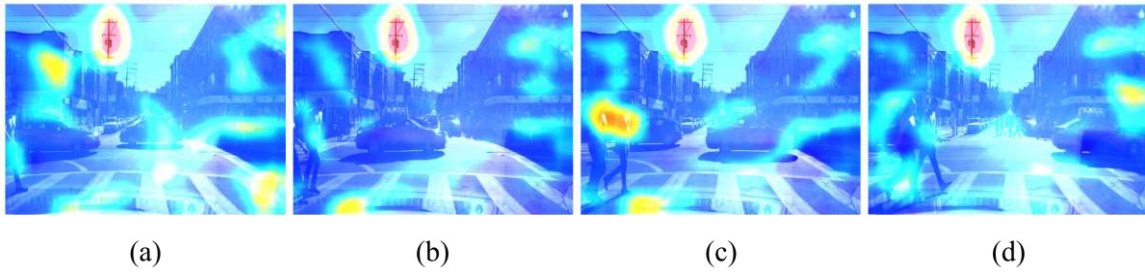


Şekil 4.3: CAT2000 veri kümesinden başarımları iyi ve kötü sonuçlar.

Yapılan deneysel çalışmaların sonuçları incelendiğinde merkezi yanlılık içerse de CAT2000 veri kümesinde eğitilen modellerin BDD-A sürüş dikkat veri kümesinde fren kararı üretmede anlamlı sonuçlar elde edemediği gözlenmiştir. Bunun nedeni, sürüş dikkatinin göreve bağlı kapsamı nedeniyle fren kararını etkileyen unsurların genelde görsel algı alanının ortasında olmamasıdır. Belirginlik tahmin etme karşılaştırmalarında ise, herhangi bir görev kısıtı olmadan serbest görüntülerin ortalamasında merkeze doğru bir yanlılık oluşmaktadır. Bu açıdan ele alındığında görev kısıtı olmadan insan görüşü kapalı dikkat ile görsel algı alanının merkezine odaklanmaya çalışırken, görev kısıtı altında, görevin gerektirdiği şekilde görsel algı alanında açık dikkat kullanarak belirginlik araması yapıyor olabileceği düşünülmektedir (Şekil-4.4).

Çizelge 4.3: BDD-A veri kümesinde belirginlik tahmin etme sonuçları.

Model	CC	KLD
BDD-A Baseline (Xia vd., 2019)	0.59	1.24
CC-KLD-RBF32	0.5685	0.4607
CC-KLD-NCB	0.3491	0.6294
KLD-NCB	0.3312	0.5392
CC-NCB	0.1921	0.8413
CC-KLD-G16	0.0000	5.1727
CC-KLD-G32	0.0000	5.1727



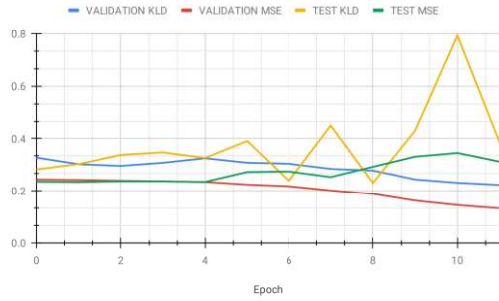
Şekil 4.4: CAT2000 veri kümesinde eğitilen modellerin BDD-A veri kümesinden elde edilen test sonuçlarına örnekler.

BDD-A sürüş dikkat veri kümesinde belirginlik tahmin etme başarımları referans model olan (Xia vd., 2019)'den daha iyidir. Ancak, merkezi yanlılık içeren model BDD-A veri kümesinde kötü başarımlar elde etmiştir. Bu bulgu Kuen v.d.'nin (Kuen vd., 2016) bulgularıyla örtüşmektedir. Kuen v.d. çalışmalarında merkezi yanlılığın, herhangi bir görev kısıtı olmayan belirginlik tahmin etmede sınırlı katkısı olduğunu ifade etmişlerdir. Sürüş gibi görsel algı olarak yoğun görev kısıtı olan problemlerde belirginlik tamamen açık dikkat ile görsel algı alanında aranmaktadır. Parr ve Friston da (Parr ve Friston, 2017) çalışmalarında görev kısıtına uygun olarak görsel algının aktif olarak belirginlik aradığını ifade etmişlerdir. BDD-A veri kümesinde belirginlik tahmin etmeye ilişkin sonuçlar Çizelge-4.3'te verilmiştir. BDD-A veri kümesi ikili bakış sabitleme nokta haritasını içermediği için *NSS* ölçütü kullanılamamıştır. Burada, G ve RBF kullanılan merkezi yanlılık tipini, takip eden sayı merkezi yanlılık parametre sayısını göstermektedir. NCB, merkezi yanlılık uygulanmadığı anlamına gelir.

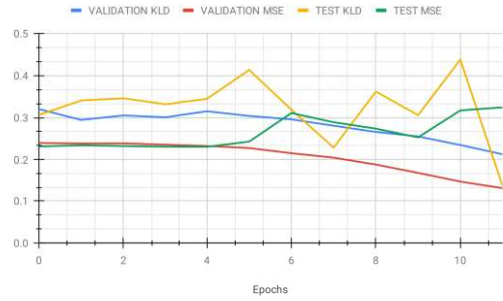
Önerilen eşzamanlı model fren kararı üretirken girdi olarak tahmin edilen belirginlik haritasını kullanmaktadır. Girdi olarak verilen belirginlik haritası, tam bağlı bir YSA'dan

geçirilerek ikili sınıflandırma yapılmaktadır. Model fren kararı ile eşzamanlı olarak tahmin edilen belirginlik haritasını da çıktı olarak vermektedir. Eğitim sürecinde gerçek fren karar etiket bilgisi ve gerçek belirginlik haritası birlikte kullanılarak en iyileme yapılmaktadır.

Önerilen eşzamanlı modelin fren yapma kararı verme başarımı sonuçları Şekil-4.5’de verilmiştir. Eşik değeri $0.5 m/sn^2$, ve tolerans değeri $0.25 m/sn$ olarak seçildiğinde en iyi sonuç elde edilmiştir. BDD-A telemetri veri kümesi ile bu sonuç uyumludur. Bununla birlikte, eşik değeri $1.0 m/sn^2$, ve $0.5 m/sn$ en kötü başarımı sergilemiştir. Bunun nedeni, seçilen bu eşik ve tolerans değerlerinin, BDD-A telemetri veri kümesinde yanlış etiketlemelere neden olmasıdır. Ölçütler açısından sonuçlar Çizelge-4.4’te gösterilmektedir.



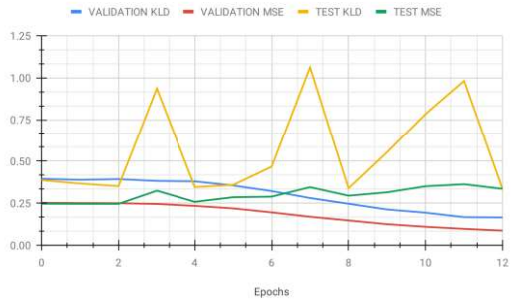
(a) Eşik=0.5, Tolerans=0.1



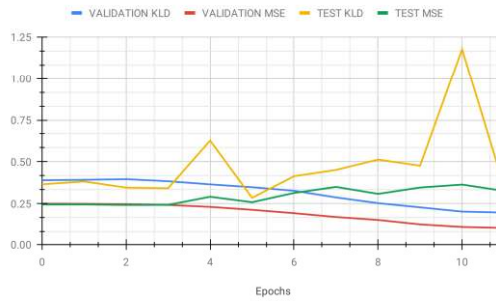
(b) Eşik=0.5, Tolerans=0.25



(c) Eşik=1.0, Tolerans=0.5



(d) Eşik=1.4715, Tolerans=0.5



(e) Eşik=1.4715, Tolerans=0.25

Şekil 4.5: Farklı eşik ve tolerans değerlerinde elde edilen sürüş karar bulgu çizelgeleri.

Çizelge 4.4: Zenginleştirilmiş BDD-A veri kümesinde elde edilen fren sürüş karar bulguları.

Model	KLD	MSE
Eşik=1.4715, Tolerans=0.5	0.3334	0.3351
Eşik=1.4715, Tolerans=0.25	0.3511	0.3258
Eşik=1.0, Tolerans=0.5	0.9866	0.4074
Eşik=0.5, Tolerans=0.25	0.1378	0.3245
Eşik=0.5, Tolerans=0.1	0.3857	0.3123

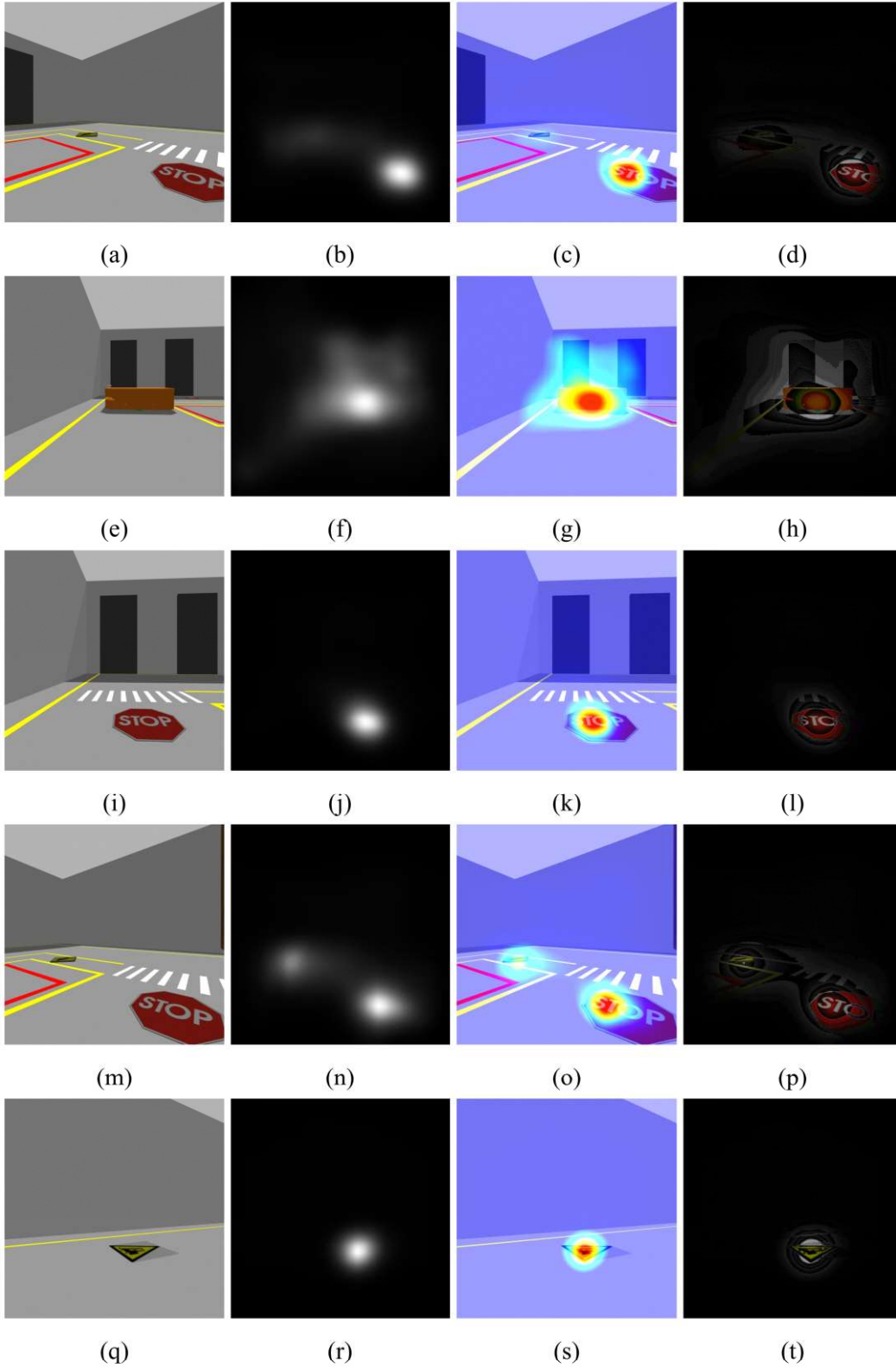
4.2 Dikkat Mekanizmasıyla Hesaplanan Belirginlik ile Ardışık Fren Karar Model Bulguları

Önerilen ardışık model eğitilirken üretilen veri kümesinden %30 oranında eğitim-doğrulama veri ayırımı kullanılmıştır. Ayrıca 1 sn zaman aralığına karşılık gelecek şekilde $batchsize = 4$ olarak seçilmiştir. Modelin geri yayılım algoritması ile en iyilenmesinde SGI algoritması kullanılmış ve öğrenim oranı olarak 0.001 seçilmiştir. Ayrıca, aşırı modellemeyi azaltmak için her 7 epoch'ta bir öğrenme hızı 0.1 çarpanıyla düşürülmüştür.

Ardışık model IFARLAB test parkurunda üç temel senaryo ile sınanmıştır. İlk senaryo, engelsiz parkurda tanımlı trafik işaretinde fren yapma senaryosudur. Bu senaryoda OTA herhangi bir engel olmadan *DUR* trafik işaretinde fren yapmaktadır (Şekil-4.7). Önerilen model, fren karar girdisi olarak tahmin edilen belirginlik haritası ile maskelenen girdi görüntüsü kullanmaktadır. Şekil-4.6'te maskeleyiş işlemine ilişkin örnekler verilmiştir. Şekilde soldan sağa sırayla, girdi görüntüsü, tahmin edilen belirginlik, girdi görüntüsü üzerinde belirginlik ve belirginlik haritası ile maskelenmiş girdi görüntüsü gösterilmektedir.

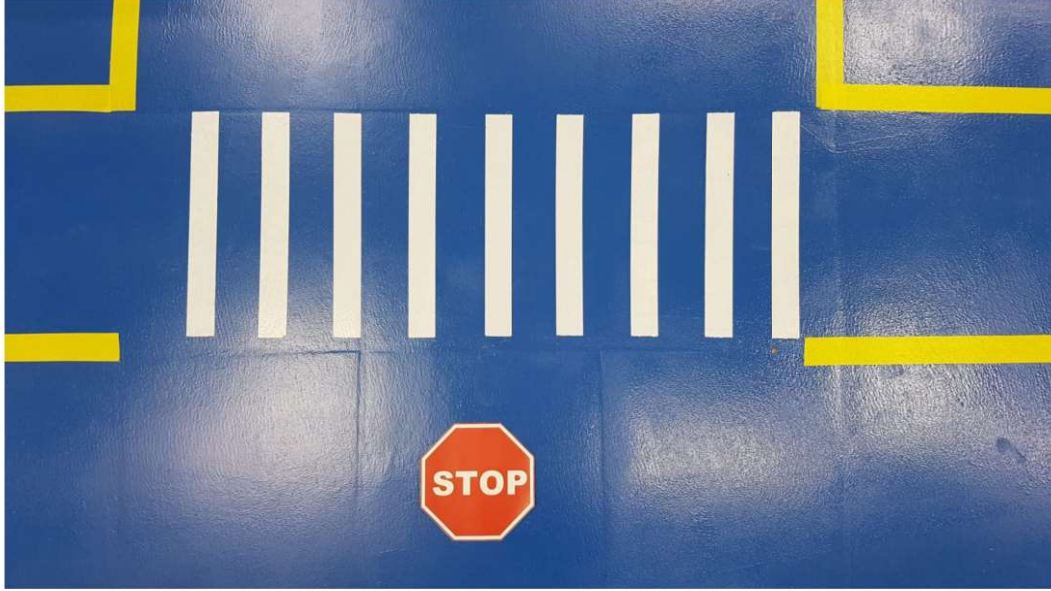
İkinci senaryo, parkur üzerinde sabit duran engel senaryosudur. Bu senaryoda parkur üzerinde önceden tanımlı konumda bir diğer OTA park edilmiştir. Senaryo gereği araç farklı mesafelerden sabit engeli algılayıp fren yapmaktadır.

Üçüncü senaryo OTA parkurda sürüş yaparken önüne çıkan hareketli engel senaryosudur. Bu senaryoda, parkur üzerinde tanımlı bir kavşaktan başka bir OTA çıkmaktadır. Senaryo gereği diğer aracın algılanıp fren yapılması sağlanmıştır.



Şekil 4.6: OTA benzetim veri kümesinden elde edilen belirginlik tahmin etme örnek sonuçları ve fren kararı için girdi görüntüsünün maskelenmiş biçimleri.

Tüm senaryolarda algılama ve fren yapma mesafeleri 0.5 m , 1.0 m , 2.0 m ve 3.0 m olarak seçilmiştir. Araç parkurda 1.0 m/sn ve 0.5 m/sn hızlarla ilerlemektedir.



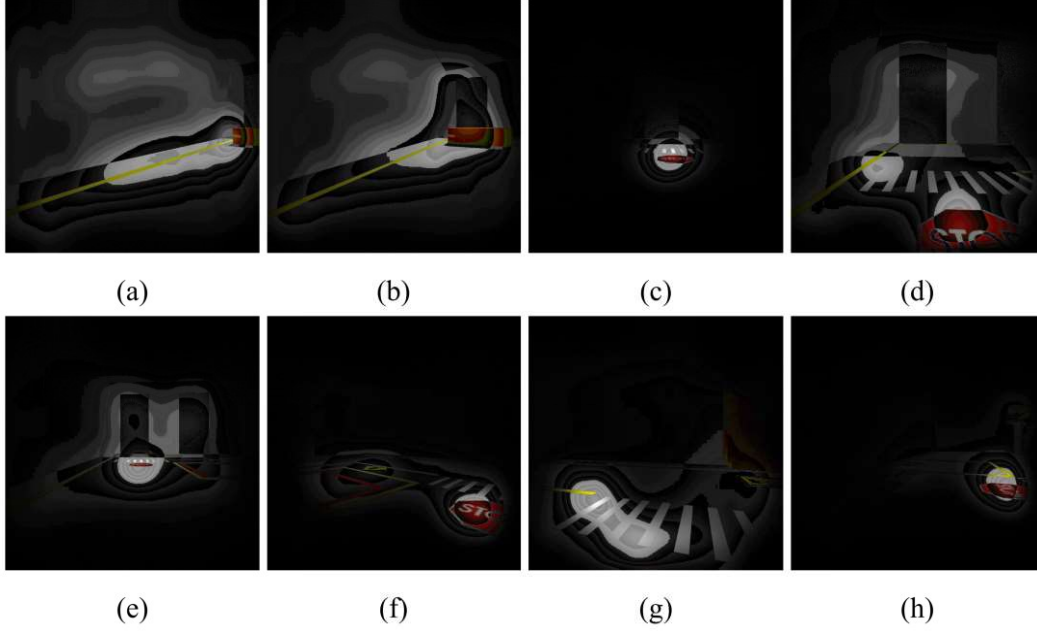
Şekil 4.7: IFARLAB test parkurundaki *DUR* işareti ve yaya geçidi.

Belirginlik tahminiyle ardışık olarak fren kararı üreten model ile OTA benzetim veri kümesinde elde edilen deneysel çalışmaların sonuçları Çizelge-4.5'te verilmiştir.

Yapılan deneylerde elde edilen bulgu ve sonuçlar değerlendirildiğinde engelli parkurda aracın çok yüksek başarımla elde ettiği göze çarpmaktadır. Engelsiz parkurda algılama ve fren yapma arasındaki anlık farklılıklardan dolayı bazı girdi görüntülerinde hatalı karar verebildiği görülmektedir. Yapılan değerlendirmede, engelsiz parkurdaki *DUR* işaretlerine ait özelliklerin hareketten bağımsız olarak aynı uzamsal belirginlik alanında yer alabilmesi olduğu düşünülmektedir. Yine de ardışık çerçevelerdeki uzamsal konumlar nedeniyle fren kararı girdi serisi bütününde yüksek başarımla doğru olarak tespit edilebilmiştir. Şekil-4.8'de üst sırada OTA benzetim veri kümesinde doğru fren kararı üreten görüntüler, alt sırada hatalı fren karar sonucu üreten *DUR* işaret örnekleri verilmektedir..

Çizelge 4.5: OTA veri kümesinde elde edilen fren sürüş karar bulguları.

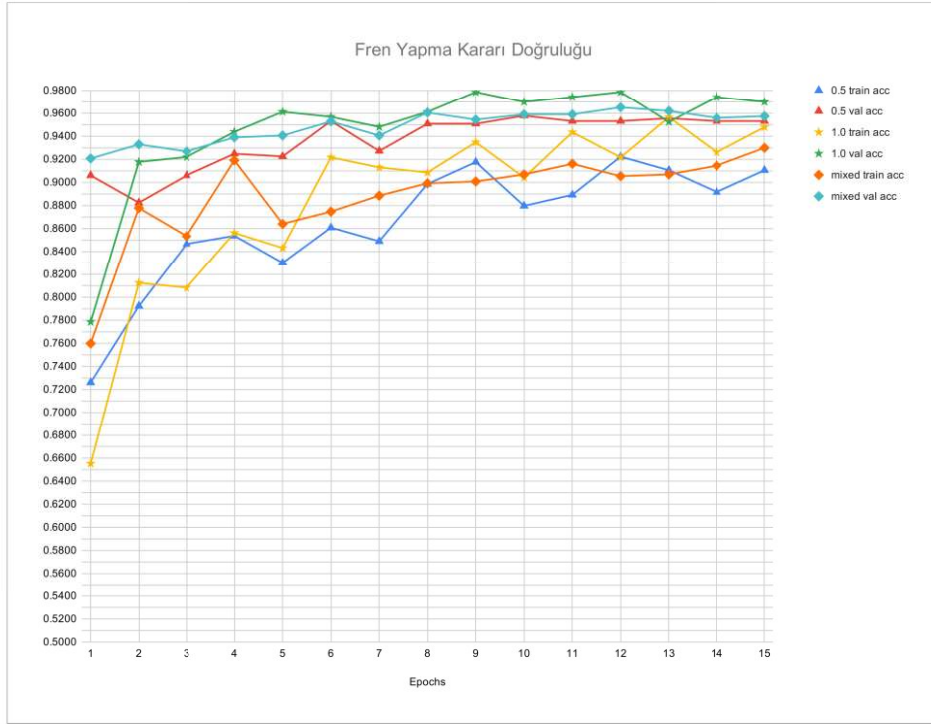
Hız	Doğruluk	F1-ölçütü	Kayıp
0.5	0.9575	0.9493	0.1233
1.0	0.9783	0.9742	0.1042
Karma	0.9648	0.9657	0.1056



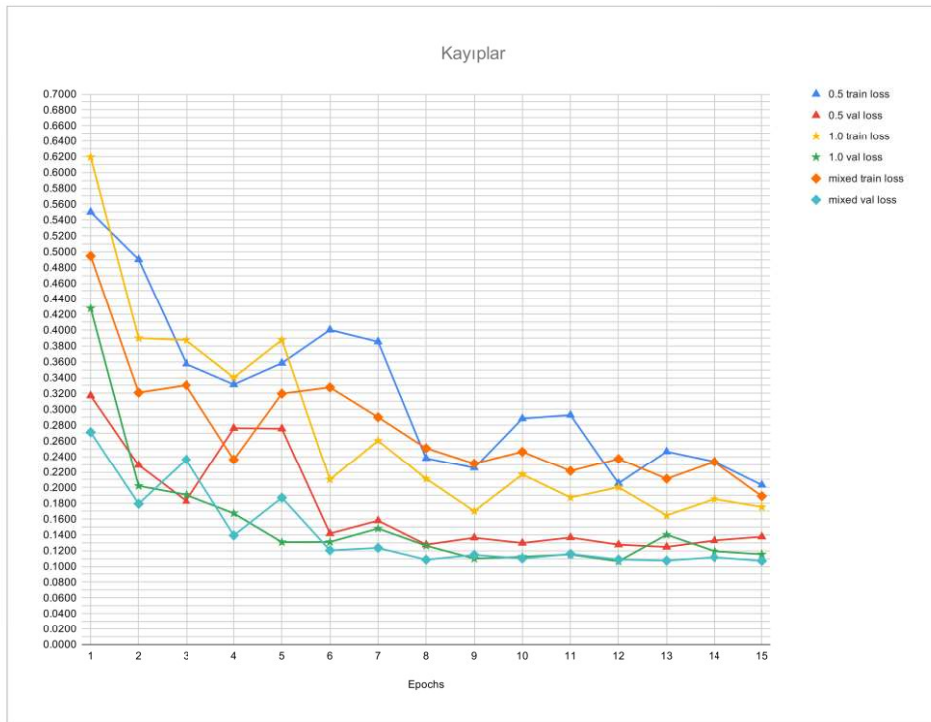
Şekil 4.8: OTA benzetim veri kümesinde fren kararı üreten görüntüler.

Engelli parkur içerisinde, engeli oluşturan diğer OTA aracı görsel algı alanına girdiğinde örnekleme aralığı içerisinde hemen belirgin olarak sürüş açısından dikkati çekmektedir. Belirginlik tahmin etme açısından araç fren kararından bağımsız olarak yüksek başarımla sergilemektedir. Eğitim sürecinde doğruluk ve kayıp fonksiyonlarının yakınsaması bu başarımla göstermektedir. Yapılan deneysel çalışmalarda eğitim ve sınamaya ait doğruluk ve kayıp çizelgeleri Şekil-4.9 ve Şekil-4.10'te verilmiştir.

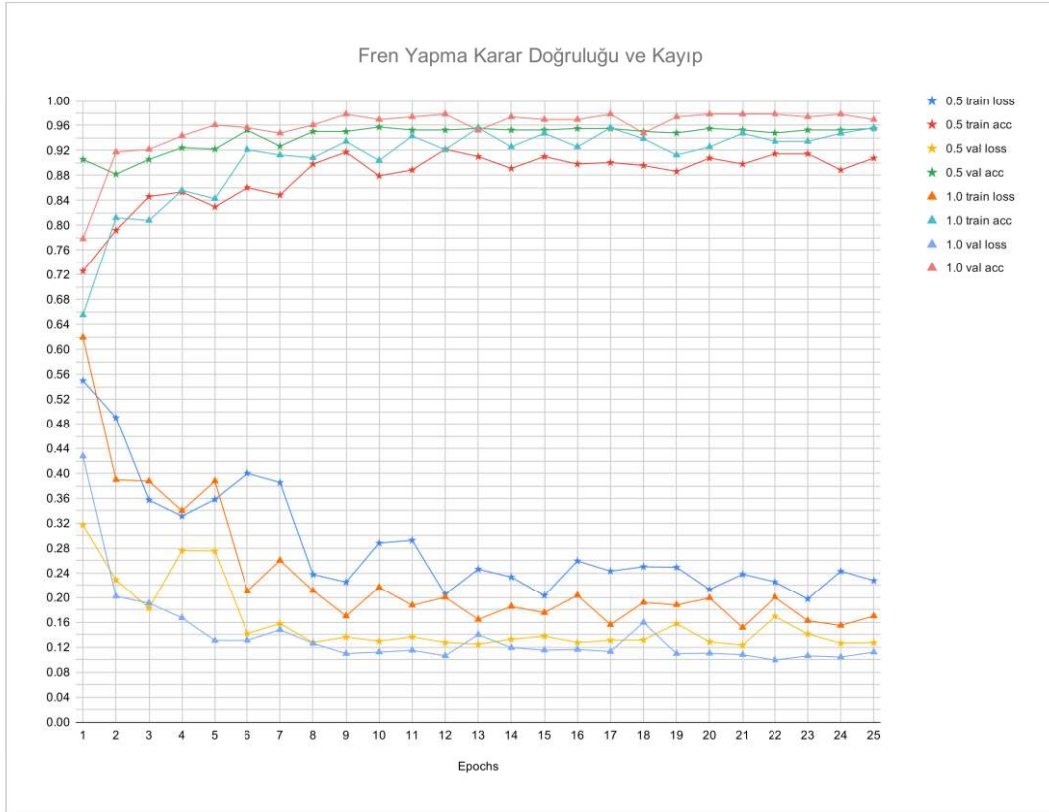
Bulgulara göre en iyi başarımla hızın $1.0 m/sn^2$ olduğu durumda ele edilmiştir. Önerilen modelin 15 epoch eğitim süresinde yakınsadığı gözlenmiş, 25 epoch a kadar eğitim süresi uzatıldığında kayda değer bir iyileşme olmadığı tespit edilmiştir (Şekil-4.11).



Şekil 4.9: OTA benzetim veri kümesinden elde edilen doğruluk çizelgeleri



Şekil 4.10: OTA benzetim veri kümesinden elde edilen kayıp çizelgeleri



Şekil 4.11: 25 Epoch ile yapılan eğitim çizelgesi.

5. SONUÇ VE ÖNERİLER

Bu tez çalışmasında, insan sürüşünün görsel algıya dayanması olgusundan hareketle bilgisayarla görüde dikkat mekanizmasının otonom sürüş veya İSDS için fren tahmininde veya durumsal farkındalığı artırmada kullanımı ele alınmıştır. Bu yaklaşımda özellikle görsel algıdaki dikkat mekanizmasına benzer şekilde, görüntüdeki belirginliklerin sürüş görevindeki fren kararına katkısı incelenmiştir.

Bu kapsamda, öncelikle belirginliğin tahmin edilmesi incelenmiştir. Belirginlik tahmin etme modeli, değiştirilmiş VGG-16 temelli EYSA'den elde edilen özellikleri EUKDH'den geçirmektedir. Elde edilen belirginlik haritası tahminleri daha sonra iki farklı yaklaşımla fren kararı vermekte kullanılmıştır. İlk yaklaşım, belirginlik haritası tahminlerini tam bağlı bir YSA'dan geçirerek fren kararı veren bir modeldir. Bu yaklaşımda belirginlik haritası tahmini ile fren kararının eş zamanlı eğitilmektedir. İkinci yaklaşım, girdi görüntüsünün belirginlik haritası tahmini ile maskelendikten sonra ResNet-18 temelli bir EYSA'den geçirilmesiyle fren kararı elde etmektedir. Bu yaklaşımda önce belirginlik tahmin etmesi modeli eğitilmekte, daha sonra fren yapma karar modeli eğitilmektedir. Bu doğrultuda, dış ortamda gerçek sürüş ve benzetim ortamında OTA sürüşünde bilgisayarlı görüde belirginliklere dikkat ederek fren yapma kararı veren bir model iki farklı biçimde önerilmiştir. Önerilen modeller, zenginleştirilmiş BDD-A veri kümesi ile OTA benzetim veri kümesinde sınanmıştır.

Önerilen modelin iki temel bileşeni bulunmaktadır: a) Sürüş Dikkat Modülü ve b) Fren Karar Modülü. Sürüş Dikkat Modülü, sürüş görüntüsündeki belirginlikleri tahmin etmektedir. Daha sonra Fren Karar Modülü'nde, tahmin edilen belirginlikler kullanılarak fren yapma kararı alınmaktadır. Her iki modül de birlikte uçtan uca eş zamanlı veya ardışık olarak geri yayılım algoritmasıyla eğitilebilmektedir. Eş zamanlı eğitimde, uçtan uca model sürüş görüntüsünden hem belirginlik hem de fren yapma kararı üretmektedir. Zenginleştirilmiş BDD-A veri kümesinde yapılan deneylerde, veri kümesinin karakteristik hız eşik ve toleranslarıyla uyumlu olacak şekilde referans modellerde elde edilen başarımlar veya biraz daha iyileştirilmiş başarımlar elde edilmiştir. Karşılaştırma açısından belirginlik tahmin etme modelleri, yaygın kullanılan belirginlik veri kümelerinde ön eğitilmiş halleriyle modüller olarak kullanılmıştır. Bu durumda, genel amaçlı görüntülerden oluşan bahsi geçen veri kümelerinde eğitilen modellerin sürüşe özgü bir belirginlik veri kümesi olan BDD-A veri kümesinde daha kötü başarımlar elde etmeleri dikkat çekmektedir. Öte yandan OTA benzetim veri kümesinde gerçekleştirilen deneylerde, model öncelikle belirginlikleri tahmin

etmekte, daha sonra girdi görüntüsünün tahmin edilen belirginlik haritasıyla maskelenmiş hali fren yapma kararına verilmektedir.

Sonuç olarak, önerilen modeller her iki veri kümesinde de motive edici sonuçlar almıştır. Bu sonuçlar, bilgisayarla görüde dikkat mekanizmasının otonom sürüş veya İSDS için önemli katkılar sunabileceğini göstermektedir. Mevcut otonom sürüş veya İSDS eş zamanlı çalışan bileşen ve farklı algılayıcıların oluşturduğu karmaşık bir sistemler bütünüdür. Sistemdeki her bir bileşen özel bir göreve tahsis edilmiştir (örn: Yaya tespiti, trafik işaret tespiti, v.b.). Önerilen model bilgisayarla görü açısından sürüşe özgü bütünlük bir yaklaşım ortaya koymaktadır. Bu çalışmanın literatüre en önemli katkısı görsel algıdaki belirginin özelliklere bütüncül olarak dikkat edip fren tahmini üreten modellerin geliştirilmesidir.

Bu konudaki ileriki araştırmalarda, bilgisayarla görüde dikkat mekanizmasının farklı bir biçimi olan gözetimsiz yaklaşıma sahip modellerin otonom sürüşteki başarımı konu olarak önerilmektedir. Genelde Pekiştirmeli Öğrenme ile gerçekleştirilen bu yaklaşım ile eğitim amacıyla veri kümesinde etiketleme yapılmasına gerek kalmayabilecektir. Benzer şekilde Bayesçi makine öğrenmesi yaklaşımı ile de gözetimsiz olarak dikkat mekanizması geliştirilebileceği düşünülmektedir. Ayrıca, insan görüşündeki dikkat mekanizması alanındaki sinir bilim çalışmalarındaki gelişmelere paralel olarak makina öğrenmesinde dikkat mekanizması konusundaki araştırmaların hız kazanacağı düşünülmektedir.

KAYNAKLAR DİZİNİ

- Bahdanau, D., Cho, K. H. ve Bengio, Y. (2014), Neural Machine Translation by Jointly Learning to Align and Translate, *arXiv e-prints*, arXiv:1409.0473. arXiv: 1409.0473.
- Bojarski, M., Del Testa, D., Dworakowski, D., Firner, B., Flepp, B., Goyal, P., Jackel, L. D., Monfort, M., Muller, U., Zhang, J. ve Others (2016), End to end learning for self-driving cars, *arXiv preprint arXiv:1604.07316*.
- Borji, A. ve Itti, L. (2012), State-of-the-art in visual attention modeling, *IEEE transactions on pattern analysis and machine intelligence* 35.1, 185–207. ISSN: 01628828. DOI: 10.1109/TPAMI.2012.89.
- (2015), CAT2000: A Large Scale Fixation Dataset for Boosting Saliency Research, *arXiv preprint arXiv:1505.03581*. arXiv: 1505.03581. URL: <http://arxiv.org/abs/1505.03581>.
- Broomhead, D. S. ve Lowe, D. (1988), *Radial basis functions, multi-variable functional interpolation and adaptive networks*. techreport. Royal Signals and Radar Establishment Malvern (United Kingdom).
- Bylinskii, Z., Judd, T., Borji, A., Itti, L., Durand, F., Oliva, A. ve Torralba, A. (2015), *Mit saliency benchmark*.
- Bylinskii, Z., Judd, T., Oliva, A., Torralba, A. ve Durand, F. (2018), What do different evaluation metrics tell us about saliency models?, *IEEE transactions on pattern analysis and machine intelligence* 41.3, 740–757.

KAYNAKLAR DİZİNİ (devam)

- Chen, Z., Zhang, J. ve Tao, D. (2019), Progressive LiDAR adaptation for road detection, *IEEE/CAA Journal of Automatica Sinica* 6.3, 693–702. DOI: 10 . 1109 / JAS . 2019 . 1911459.
- Cornia, M., Baraldi, L., Serra, G. ve Cucchiara, R. (2016), Multi-level net: A visual saliency prediction model, *European Conference on Computer Vision*. Springer, 302–315.
- (2018), Predicting human eye fixations via an LSTM-Based saliency attentive model, *IEEE Transactions on Image Processing* 27.10, 5142–5154. ISSN: 10577149. DOI: 10 . 1109/TIP.2018.2851672. arXiv: 1611.09571.
- Friston, K., Adams, R. A., Perrinet, L. ve Breakspear, M. (2012), Perceptions as Hypotheses: Saccades as Experiments, *Frontiers in Psychology* 3.MAY, 151. ISSN: 1664-1078. DOI: 10 . 3389 / fpsyg . 2012 . 00151. URL: [http : //journal.frontiersin.org/article/10.3389/fpsyg.2012.00151/abstract](http://journal.frontiersin.org/article/10.3389/fpsyg.2012.00151/abstract).
- Han, L., Li, X. ve Dong, Y. (2018), SalNet: Edge constraint based end-to-end model for salient object detection, *Chinese Conference on Pattern Recognition and Computer Vision (PRCV)*. Springer, 186–198.
- He, K., Zhang, X., Ren, S. ve Sun, J. (2016), Deep residual learning for image recognition, *Proceedings of the IEEE conference on computer vision and pattern recognition*, 770–778.
- Hochreiter, S. ve Schmidhuber, J. (1997), Long Short-Term Memory, *Neural Computation* 9.8, 1735–1780.

KAYNAKLAR DİZİNİ (devam)

IFARLAB (2020), *Eskisehir Osmangazi University Intelligent Factory and Robotics Laboratory (IFARLAB)*. URL: <https://ifarlab.ogu.edu.tr/Icerik/Detay/4>.

Itti, L. ve Koch, C. (2000), A saliency-based search mechanism for overt and covert shifts of visual attention, *Vision Research*. Vol. 40. 10-12, 1489–1506. DOI: 10.1016/S0042-6989(99)00163-7.

Itti, L., Koch, C. ve Niebur, E. (1998), A model of saliency-based visual attention for rapid scene analysis, *IEEE Transactions on Pattern Analysis and Machine Intelligence* 20.11, 1254–1259. ISSN: 01628828. DOI: 10.1109/34.730558.

Jiang, M., Huang, S., Duan, J. ve Zhao, Q. (2015), Salicon: Saliency in context, *Proceedings of the IEEE conference on computer vision and pattern recognition*, 1072–1080.

Koch, C. ve Ullman, S. (1987), Shifts in selective visual attention: towards the underlying neural circuitry, *Matters of intelligence*. Springer, 115–141.

Koenig, N. ve Howard, A. (2004), Design and Use Paradigms for Gazebo, An Open-Source Multi-Robot Simulator, *IEEE/RSJ International Conference on Intelligent Robots and Systems*. Sendai, Japan, 2149–2154.

Kuen, J., Wang, Z. ve Wang, G. (2016), Recurrent attentional networks for saliency detection, *Proceedings of the IEEE Conference on computer Vision and Pattern Recognition*, 3668–3677.

KAYNAKLAR DİZİNİ (devam)

Kümmerer, M., Wallis, T. S. A. ve Bethge, M. (2016), DeepGaze II: Reading fixations from deep features trained on object recognition, *arXiv preprint arXiv:1610.01563*.

Lindsay, G. W. (2020), Attention in Psychology, Neuroscience, and Machine Learning, *Frontiers in Computational Neuroscience* 14, 29. ISSN: 16625188. DOI: 10.3389/fncom.2020.00029.

Marszałek, M., Laptev, I. ve Schmid, C. (2009), Actions in Context, *IEEE Conference on Computer Vision & Pattern Recognition*.

Mnih, V., Heess, N., Graves, A. ve Kavukcuoglu, K. (2014), Recurrent models of visual attention, *Advances in Neural Information Processing Systems* 3.January. Ed. Z Ghahramani, M Welling, C Cortes, N. D. Lawrence ve K. Q. Weinberger, 2204–2212. ISSN: 10495258. arXiv: 1406.6247.

Moore, T., Armstrong, K. M. ve Fallah, M. (2003), Visuomotor Origins of Covert Spatial Attention, *Neuron* 40.4, 671–683. ISSN: 0896-6273. DOI: [https://doi.org/10.1016/S0896-6273\(03\)00716-5](https://doi.org/10.1016/S0896-6273(03)00716-5). URL: <http://www.sciencedirect.com/science/article/pii/S0896627303007165>.

Olah, C. (2020), *Understanding LSTM Networks*. URL: <http://colah.github.io/posts/2015-08-Understanding-LSTMs>.

Palazzi, A, Abati, D, Calderara, s., Solera, F ve Cucchiara, R (2019), Predicting the Driver's Focus of Attention: The DR(eye)VE Project, *IEEE Transactions on Pattern Analysis and Machine Intelligence* 41.7, 1720–1733. ISSN: 0162-8828. DOI: 10.1109/TPAMI.2018.2845370.

KAYNAKLAR DİZİNİ (devam)

- Palazzi, A, Solera, F, Calderara, S, Alletto, S ve Cucchiara, R (2017), Learning where to attend like a human driver, *2017 IEEE Intelligent Vehicles Symposium (IV)*, 920–925. DOI: 10.1109/IVS.2017.7995833.
- Parr, T. ve Friston, K. J. (2017), Working memory, attention, and salience in active inference, *Scientific Reports* 7.1, 1–21. ISSN: 20452322. DOI: 10.1038/s41598-017-15249-0. URL: <http://dx.doi.org/10.1038/s41598-017-15249-0>.
- (2019), Attention or salience?, *Current Opinion in Psychology* 29, 1–5. ISSN: 2352250X. DOI: 10.1016/j.copsy.2018.10.006. URL: <https://linkinghub.elsevier.com/retrieve/pii/S2352250X18301593>.
- Pomerleau, D. A. (1989), Alvin: An autonomous land vehicle in a neural network, *Advances in Neural Information Processing Systems 1*, 305–313.
- Ramanishka, V., Chen, Y.-T., Misu, T. ve Saenko, K. (2018), Toward driving scene understanding: A dataset for learning driver behavior and causal reasoning, *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 7699–7707.
- Rizzolatti, G., Riggio, L., Dascola, I. ve Umiltá, C. (1987), Reorienting attention across the horizontal and vertical meridians: Evidence in favor of a premotor theory of attention, *Neuropsychologia* 25.1, Part 1, 31–40. ISSN: 0028-3932. DOI: [https://doi.org/10.1016/0028-3932\(87\)90041-8](https://doi.org/10.1016/0028-3932(87)90041-8). URL: <http://www.sciencedirect.com/science/article/pii/0028393287900418>.

KAYNAKLAR DİZİNİ (devam)

Rumelhart, D. E., Hinton, G. E. ve Williams, R. J. (1986), Learning representations by back-propagating errors, *Nature* 323.6088, 533–536. ISSN: 1476-4687. DOI: 10.1038/323533a0. URL: <https://doi.org/10.1038/323533a0>.

Russakovsky, O., Deng, J., Su, H., Krause, J., Satheesh, S., Ma, S., Huang, Z., Karpathy, A., Khosla, A., Bernstein, M., Berg, A. C. ve Fei-Fei, L. (2015), ImageNet Large Scale Visual Recognition Challenge, *International Journal of Computer Vision* 115.3, 211–252. ISSN: 15731405. DOI: 10.1007/s11263-015-0816-y. arXiv: 1409.0575. URL: [http://image-net.org/challenges/LSVRC/..](http://image-net.org/challenges/LSVRC/)

SAE International (2018), *Taxonomy and Definitions for Terms Related to Driving Automation Systems for On-Road Motor Vehicles*. DOI: https://doi.org/10.4271/J3016_201806. URL: https://doi.org/10.4271/J3016{_}201806.

SAE Mobilus (2020), *Fusing Sensors for the Automated Driving Future*. URL: <https://saemobilus.sae.org/binaries/content/gallery/mobilus/mobilus-hub/article/feature/fusing-sensors-for-the-automateddriving-future.png/fusing-sensors-for-the-automateddriving-future.png/mobilus{\%}3Afull>.

Simonyan, K., Vedaldi, A. ve Zisserman, A. (2013), Deep inside convolutional networks: Visualising image classification models and saliency maps, *arXiv preprint arXiv:1312.6034*.

Simonyan, K. ve Zisserman, A. (2014), Very deep convolutional networks for large-scale image recognition, *arXiv preprint arXiv:1409.1556*.

KAYNAKLAR DİZİNİ (devam)

- Soomro, K. ve Zamir, A. R. (2014), Action Recognition in Realistic Sports Videos, *Computer Vision in Sports*. Ed. T. B. Moeslund, G. Thomas ve A. Hilton. Cham: Springer International Publishing, 181–208. ISBN: 978-3-319-09396-3. DOI: 10.1007/978-3-319-09396-3_9. URL: https://doi.org/10.1007/978-3-319-09396-3{_}9.
- Tawari, A ve Kang, B (2017), A computational framework for driver’s visual attention using a fully convolutional architecture, *2017 IEEE Intelligent Vehicles Symposium (IV)*, 887–894. DOI: 10.1109/IVS.2017.7995828.
- Tawari, A, Mallela, P ve Martin, S (2018), Learning to Attend to Salient Targets in Driving Videos Using Fully Convolutional RNN, *2018 21st International Conference on Intelligent Transportation Systems (ITSC)*, 3225–3232. DOI: 10.1109/ITSC.2018.8569438.
- Wang, W., Lai, Q., Fu, H., Shen, J. ve Ling, H. (2019a), Salient Object Detection in the Deep Learning Era: An In-Depth Survey, *ArXiv abs/1904.0*.
- Wang, W., Shen, J., Xie, J., Cheng, M.-M., Ling, H. ve Borji, A. (2019b), Revisiting Video Saliency Prediction in the Deep Learning Era, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 1–1. ISSN: 0162-8828. DOI: 10.1109/tpami.2019.2924417.
- Wu, Y., Liu, Y., Li, J., Liu, H. ve Hu, X. (2013), Traffic sign detection based on convolutional neural networks, *The 2013 International Joint Conference on Neural Networks (IJCNN)*, 1–7. DOI: 10.1109/IJCNN.2013.6706811.

KAYNAKLAR DİZİNİ (devam)

- Xia, Y., Zhang, D., Kim, J., Nakayama, K., Zipser, K. ve Whitney, D. (2019), Predicting Driver Attention in Critical Situations, *Computer Vision – ACCV 2018* 11365 LNCS. Ed. C. V. Jawahar, H. Li, G. Mori ve K. Schindler, 658–674. ISSN: 16113349. DOI: 10.1007/978-3-030-20873-8_42. arXiv: 1711.06406. URL: http://link.springer.com/10.1007/978-3-030-20873-8_{_}42<http://arxiv.org/abs/1711.06406>.
- Xingjian, S. H. I., Chen, Z., Wang, H., Yeung, D.-Y., Wong, W.-K. ve Woo, W.-c. (2015), Convolutional LSTM network: A machine learning approach for precipitation nowcasting, *Advances in neural information processing systems*, 802–810.
- Xu, H, Gao, Y, Yu, F ve Darrell, T (2017), End-to-End Learning of Driving Models from Large-Scale Video Datasets, *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 3530–3538. DOI: 10.1109/CVPR.2017.376.
- Xu, K., Ba, J. L., Kiros, R., Cho, K., Courville, A. C., Salakhutdinov, R., Zemel, R. S. ve Bengio, Y. (2015), Show, Attend and Tell: Neural Image Caption Generation with Visual Attention, *CoRR* abs/1502.0, 2048–2057. arXiv: 1502.03044. URL: <http://arxiv.org/abs/1502.03044>.
- Yarbus, A. L. (2013), *Eye movements and vision*. Springer.
- Yu, F. ve Koltun, V. (2015), Multi-scale context aggregation by dilated convolutions, *arXiv preprint arXiv:1511.07122*.
- Yu, F., Xian, W., Chen, Y., Liu, F., Liao, M., Madhavan, V. ve Darrell, T. (2018), BDD100K: A Diverse Driving Video Database with Scalable Annotation Tooling, *arXiv e-prints*, arXiv:1805.04687. arXiv: 1805.04687 [cs.CV].

KAYNAKLAR DİZİNİ (devam)

Yu, H., Yang, S., Gu, W. ve Zhang, S. (2017), Baidu driving dataset and end-to-end reactive control model, *2017 IEEE Intelligent Vehicles Symposium (IV)*. IEEE, 341–346.

Zhang, L., Lin, L., Liang, X. ve He, K. (2016), Is faster R-CNN doing well for pedestrian detection?, *European conference on computer vision*. Springer, 443–457.

Zou, Q., Jiang, H., Dai, Q., Yue, Y., Chen, L. ve Wang, Q. (2020), Robust lane detection from continuous driving scenes using deep neural networks, *IEEE Transactions on Vehicular Technology* 69.1, 41–54. ISSN: 19399359. DOI: 10.1109/TVT.2019.2949603.

Özgeçmiş

