



POLİTEKNİK DERGİSİ

JOURNAL of POLYTECHNIC

ISSN: 1302-0900 (PRINT), ISSN: 2147-9429 (ONLINE)

URL: <http://dergipark.org.tr/politeknik>



Yaya özellik tanıma için LM filtre temelli derin evrimsel sinir ağı

LM filter-based deep convolutional neural network for pedestrian attribute recognition

Yazar(lar) (Author(s)): Hüseyin ÜZEN¹ , Kazım HANBAY²

ORCID¹: 0000-0002-0998-2130

ORCID²: 0000-0003-1374-1417

Bu makaleye şu şekilde atıfta bulunabilirsiniz (To cite to this article): Üzen H. ve Hanbay K., “Yaya özellik tanıma için lm filtre temelli derin evrimsel sinir ağı”, *Politeknik Dergisi*, 23(3): 605-613, (2020).

Erişim linki (To link to this article): <http://dergipark.org.tr/politeknik/archive>

DOI: 10.2339/politeknik.525600

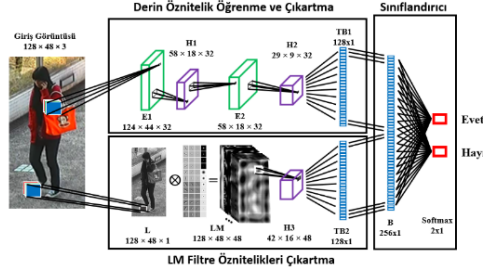
Yaya Özellik Tanıma için LM Filtre Temelli Derin Evrişimsel Sinir Ağı

LM Filter-Based Deep Convolutional Neural Network for Pedestrian Attribute Recognition

Önemli noktalar (Highlights)

- ❖ Hibrit Derin Evrişimsel Sinir Ağı Mimarisi (Deep Hybrid Convolutional Neural Network Architecture)

Grafik Özet (Graphical Abstract)



Şekil. Geliştirilen HESA mimarisi / Figure. The developed HESA architecture

Amaç (Aim)

Bu ağ mimarisinin temel amacı, LM filtrelerinden elde edilen geleneksel özellikler ile ESA mimarisinden elde edilen derin özellikleri birleştirerek güçlü bir özellik verisi oluşturmaktır. / The main purpose of this network architecture is to combine the traditional features obtained from the LM filters and the deep features obtained from the CNN architecture so thus create a strong feature data for classification.

Tasarım ve Yöntem (Design & Methodology)

Önerilen modelde yaya görüntüsünden elde edilen LM filtre özellikleri ve derin özellikler eşzamanlı olarak hesaplanmaktadır. Daha sonra bu özellikler birleştirilir ve 1×256 farklı özellikten oluşan bir özellik vektörü oluşturulur. Bu özellik vektörü tam bağlı katmanlar yardımı ile sınıflandırma işlemine alınır. / In the proposed model, the LM filter features and deep features of the pedestrian image are calculated simultaneously. Then, these features are combined and features vector consisting of 1×256 different features is built. This feature vector is taken into the classification process with the help of fully connected layer.

Özgünlük (Originality)

Bu çalışmada hem geleneksel hem de derin özellikleri hesaplamak için yeni bir Hibrit Evrişimsel Sinir Ağı (HESA) mimarisi önerilmiştir. / In this paper, a new Hybrid Convolutional Neural Network (HESA) architecture is proposed to calculate both the traditional and the deep features.

Bulgular (Findings)

Geliştirilen HESA mimarisi çok zor bir problem olan yaya özellik sınıflandırması için uygulanmıştır. Önerilen model PETA veri tabanında SVM ve MRF tabanlı yöntemlerden önemli ölçüde daha iyi performans göstermiştir. Ayrıca, ReduceLRonPlateau modelinin HESA yönteminde kullanılması yüksek başarıların elde edilmesine önemli bir katkı sağlamıştır. / The developed HESA architecture has been applied for the pedestrian attribute classification which is a very difficult problem. The proposed model significantly outperforms the SVM and MRF based methods on the PETA database. In addition, the use of the ReduceLRonPlateau model in the HESA method has made a significant contribution to achieving high successes.

Sonuç (Conclusion)

Önerilen HESA mimarisi güçlü bir özellik çıkartma ve sınıflandırma yapısı oluşturulmuştur. Oluşturulan bu hibrit ESA mimarisi yaya özellik tanıma problemine uygulanmıştır. Elde edilen sonuçlara bakıldığında, bu mimari sayesinde daha güçlü bir ESA mimarisi geliştirme olanağı sağlanabileceği görülmüştür. / The proposed HESA architecture has a strong feature extraction and classification structure. This hybrid ESA architecture has been applied to the pedestrian feature recognition problem. Considering the results obtained, it was seen that this architecture could provide the opportunity to develop a stronger ESA architecture.

Etik Standartların Beyanı (Declaration of Ethical Standards)

Bu makalenin yazar(lar)ı çalışmalarında kullandıkları materyal ve yöntemlerin etik kurul izni ve/veya yasal-özel bir izin gerektirmediğini beyan ederler. / The author(s) of this article declare that the materials and methods used in this study do not require ethical committee permission and/or legal-special permission.

Yaya Özellik Tanıma için LM Filtre Temelli Derin Evrişimsel Sinir Ağı

Araştırma Makalesi / Research Article

Hüseyin ÜZEN*, Kazım HANBAY

Bingöl Üniversitesi, Bilgisayar Mühendisliği Bölümü, Türkiye
(Geliş/Received : 11.02.2019 ; Kabul/Accepted : 17.06.2019)

ÖZ

Günümüzde Evrişimsel Sinir Ağı (ESA) mimarileri güvenlik, endüstri ve büyük veri gibi birçok farklı alanda aktif olarak kullanılmaktadır. Bu mimarilerdeki evrişim katmanları, bir sınıflandırma veya tanımlama problemi için istenilen sonuçları verebilecek en iyi öznelikleri otomatik olarak çıkartabilmektedir. Bu çalışmada, hem geleneksel hem de derin öznelikleri hesaplamak için yeni bir Hibrit Evrişimsel Sinir Ağı (HESA) mimarisi önerilmiştir. Bu ağ mimarisinin temel amacı, LM filtrelerinden elde edilen geleneksel öznelikler ile ESA mimarisinden elde edilen derin öznelikleri birleştirerek güçlü bir öznelik verisi oluşturmaktır. Önerilen modelde yaya görüntüsünden elde edilen LM filtre öznelikleri ve derin öznelikler eşzamanlı olarak hesaplanmaktadır. Daha sonra bu öznelikler birleştirilir ve 1×256 farklı öznelikten oluşan bir öznelik vektörü oluşturulur. Bu öznelik vektörü tam bağlı katmanlar yardımı ile sınıflandırma işlemine alınır. Geliştirilen HESA mimarisi çok zor bir problem olan yaya özellik sınıflandırması için uygulanmıştır. Önerilen model PETA veri tabanında SVM ve MRF tabanlı yöntemlerden önemli ölçüde daha iyi performans göstermiştir. Ayrıca, *ReduceLRonPlateau* modelinin HESA yönteminde kullanılması yüksek başarıların elde edilmesine önemli bir katkı sağlamıştır.

Anahtar Kelimeler: Yaya özellik tanıma, derin öğrenme, öznelik çıkartma.

LM Filter-Based Deep Convolutional Neural Network for Pedestrian Attribute Recognition

ABSTRACT

Today, Convolutional Neural Network (CNN) architectures have been used actively in many different areas such as security, industry and big data. Thanks to the convolution layers in these architectures, they can automatically extract the best features that can give the desired results for a classification or definition problem. In this paper, a new Hybrid Convolutional Neural Network (HESA) architecture is proposed to calculate both the traditional and the deep features. The main purpose of this network architecture is to combine the traditional features obtained from the LM filters and the deep features obtained from the CNN architecture so thus create a strong feature data for classification. In the proposed model, the LM filter features and deep features of the pedestrian image are calculated simultaneously. Then, these features are combined and features vector consisting of 1×256 different features is built. This feature vector is taken into the classification process with the help of fully connected layer. The developed HESA architecture has been applied for the pedestrian attribute classification which is a very difficult problem. The proposed model significantly outperforms the SVM and MRF based methods on the PETA database. In addition, the use of the *ReduceLRonPlateau* model in the HESA method has made a significant contribution to achieving high successes.

Keywords: Pedestrian attribute recognition, deep learning, feature extraction

1. GİRİŞ (INTRODUCTION)

Geniş bir alanı gözlemleyen kameralardan bir nesne veya yayanın görsel olarak tanımlanması sınır kontrolü, güvenlik ve reklamcılık gibi çeşitli alanlar için kritik bir öneme sahiptir. Özellikle yayaların cinsiyet, giyim tarzı gibi kişisel özelliklerin sınıflandırılması [1, 2] restoranlarda, süpermarketlerde müşteri istatistiklerinin toplanması ve kamu güvenliği gibi alanlarda oldukça kullanışlı olabilmektedir [3]. Ancak ortak yaşam alanlarında insanların özelliklerinin (giyim, cinsiyet, aksesuar vb.) otomatik bir şekilde sınıflandırılması ortam şartları ve kameradan kaynaklanan problemlerden dolayı zor bir işlemdir.

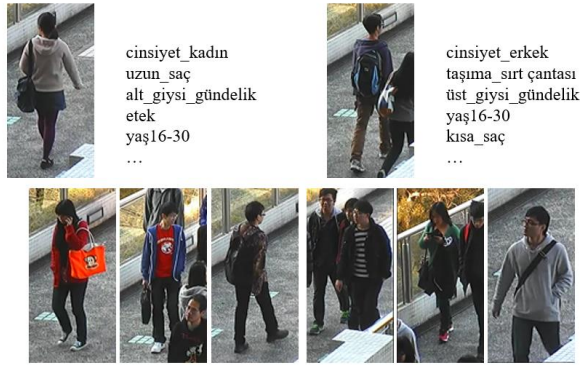
Yayaların fiziksel özelliklerinin sınıflandırılması temel olarak iki problem içermektedir. Bunlar sırasıyla insan

özelliklerinin oldukça çeşitli olması ve kameralardan alınan görüntülerde bulanıklığın ve belirsizliklerin bulunmasıdır [4, 5]. Birinci problem olan kişi özelliklerin çok çeşitli olması durumu, bir bireyin sahip olduğu giyim tarzı, saç uzunluğu, takılar gibi özelliklerin çok öznel olabilmesinden kaynaklanmaktadır. İkinci problem ise analiz edilen görüntülerdeki gürültü, kontrast eksikliği, nesnelerin ve insanların çakışık olması ve izlenen ortamın dinamikliğidir [4, 6]. Bu zorlukların temel sebebi görüntülerin alındığı ortamların genellikle açık alanlar olması ve görüntülerin uzak bir mesafeden alınması durumlarıdır.

Literatürde yaya özelliklerinin tanımlanması ve sınıflandırılması için geliştirilmiş olan yöntemler birçok farklı veri kümesi üzerinde test edilmiştir. Bu veri kümeleri genellikle farklı insan görüntüsü ve her bir görüntüye ait birçok farklı özellikleri içerisinde barındırmaktadır (Şekil 1). Literatürdeki birçok çalışma

*Sorumlu Yazar (Corresponding Author)
e-posta : huzen@bingol.edu.tr

içerdiği özellik sayısının çok olması ve 19000'den fazla görüntü içermesinden dolayı Peta veri kümesini kullanmıştır.



Şekil 1. Peta veri tabanından örnek insan görüntüleri ve etiketleri (Sample human images and labels from Peta dataset)

Literatürde bulunan yaya özellik sınıflandırma veya tanımlama yöntemleri genellikle her bir özellik için ayrı bir sınıflandırıcı eğitmişlerdir. Her bir eğitim kümesindeki görüntüler ele alınan özelliğe sahip olma durumlarına göre pozitif veya negatif olarak etiketlenmiştir. Bununla birlikte sınıflandırıcılara verilmek üzere görüntülerden genellikle düşük seviyeli özneliklerin elde edildiği görülmektedir [3, 7]. Ancak günümüzde otomatik ve görüntüyü daha yüksek seviyede analiz eden öznelik çıkartma ve sınıflandırma yöntemleri de mevcuttur [4, 8]. Literatürdeki yaya özellik sınıflandırma yöntemlerinde sınıflandırıcı olarak destek vektör makinesi de (DVM) yoğun bir şekilde kullanılmıştır [1, 9, 10]. Bunun yanında AdaBoost (Adaptive Boosting) [11], KNN [7, 12], rastgele ormanlar (Random Forest) [3] ve softmax [4, 8] gibi sınıflandırıcıları kullanan yöntemlerde mevcuttur. Bununla birlikte bazı çalışmalar fiziksel insan özelliklerinin birbirleriyle olan ilişkisinin önemli olduğunu kabul ederek ortak bir sınıflandırıcı eğitmeyi amaçlamışlardır [6, 13]. Bourdev vd. [14] yaptıkları çalışmada her bir özellik için elde edilen ayrı ayrı sınıflandırıcı sonuçlarını ortak bir DVM sınıflandırıcı ile birleştirmiştir. Zhu vd. her bir özelliğe ait sınıflandırıcı sonuçlarını ağırlıklandırarak birleştirmeye çalışmıştır [13]. Fakat bu çalışmalarda en büyük problem bazı fiziksel insan özelliklerinin birbiri ile olan ilişkisinin az ya da tanımlanmasının zor olmasıdır. Örneğin cinsiyet tespitinde kişinin erkek veya kadın olarak sınıflandırılması tek bir sınıflandırıcı ile kolay olabilmektedir. Ancak şapka, yaş ve takı gibi birbiri ile etkileşimi zor olan fiziksel özelliklerin ortak bir sınıflandırıcı ile tespit edilmesi zordur.

Cao vd. tarafından geliştirilen cinsiyet tespit yöntemi yaya özellikleri temel alınarak yapılan ilk çalışmalardan biridir [3]. Bu çalışmada öncelikle insan görüntülerinden ham piksel değerleri ve HOG özellikleri çıkartılmıştır. Zhu vd. insan görüntülerindeki kişisel fiziksel özellikleri sınıflandırmak için görüntülerin renk (HSV renk kanalları), doku (LBP) ve HOG özneliklerini

kullanmışlardır [12]. Her bir görüntüden elde edilen bu özneliklerden ikili özellik tespiti için AdaBoost sınıflandırıcı, çoklu fiziksel özellik sınıflandırması için ise KNN sınıflandırıcı kullanılmıştır. Layne vd. fiziksel insan özelliklerini sınıflandırmak için 8 farklı renk kanalı (RGB, HSV ve YCbCr) ve bu kanallar üzerinde gezdirmek üzere 21 farklı doku filtresi (8 Gabor filtresi ve 13 Schmid filtresi) kullanmıştır [7].

Son yıllarda yaya özellik tespitinde derin öğrenme tabanlı yöntemler kullanılarak başarılı çalışmalar yürütülmüştür. Bu çalışmalardan biri Su vd. tarafından Peta veri kümesini kullanarak yapılan derin öğrenme tabanlı ve yarı denetimli insan özellik tespiti çalışmasıdır [15]. Bu çalışma insan özelliklerini 3 aşamalı bir eğitim şeması kullanarak öğrenmeyi hedeflemiştir. Zhu vd. geliştirdikleri derin öğrenme tabanlı ağ mimarisi sayesinde yaya özelliklerini başarılı bir şekilde sınıflandırmıştır [16]. Yaptıkları çalışmada öncelikle görüntüler 15 parçaya bölünmüştür. Her bir görüntü parçası oluşturulan özel ağ mimarisinden geçirilerek bu görüntülerin öznelikleri ayrı ayrı hesaplanmıştır. Bir başka çalışmada ise tüm parçalardan elde edilen öznelikler tam bağlı bir katman aracılığıyla birleştirilmiştir [4]. Daha sonra bu tam bağlı katmandaki çıktılar bir softmax katmanına gönderilmiştir. Görüntülerin sadece evrimsel özelliklerinin kullanılmasının bazı fiziksel özellikleri ayırt etmede yetersiz kaldığı bazı çalışmalarda vurgulanmıştır. Liu vd. yaptıkları çalışmada ele alınan görüntüden ayrı ayrı bölgesel ve genel öznelikleri çıkartan güçlü bir ağ mimarisi geliştirilmiştir [17]. Bununla birlikte birbirine benzeyen fiziksel özelliklerin ayırt edici piksel ve bölge tabanlı özneliklerinin sadece geleneksel derin öğrenme algoritmaları kullanılarak hesaplanması istenen düzeyde başarılı sonuçlar vermemektedir.

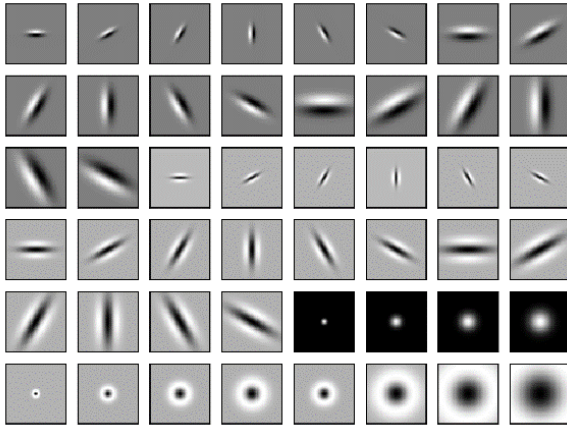
Bu çalışmada yaya fiziksel özelliklerinin tanımlanması ve sınıflandırılabilmesi için Leung-Malik (LM) filtre bankası ile derin öğrenme tabanlı bir ağ mimarisi birlikte kullanılmıştır. Böylece görüntülerin evrimsel özellikleri ile birlikte güçlü dokusal öznelikleri de LM filtreleri ile elde edilmiş ve sınıflandırma sürecine aktif olarak dâhil edilmiştir. LM filtreleri farklı ölçek ve açısız yönlerde tasarlanarak ölçekten ve açısız yönelimden bağımsız bir sınıflandırma yöntemi geliştirilmiştir. Geliştirilen evrimsel ağ mimarisinde LM filtresinden elde edilen dokusal öznelikler ile evrimsel sinir ağından elde edilen yüksek seviyeli öznelikler birleştirilmiş ve hibrit bir özellik vektörü inşa edilmiştir. İnşa edilen ESA mimarisi elde edilen her iki özellik vektöründen herhangi bir bilgi kaybına ve manipülasyonuna yol açmayacak şekilde ve çıkarılan öznelikleri tam kullanacak şekilde tasarlanmıştır. Elde edilen bu öznelikler softmax katmanı öncesinde birleştirilerek yüksek sınıflandırma başarısı elde edilmiştir. Bölüm 2'de LM filtre bankası ve ESA hakkında bilgiler verilmiştir. Bölüm 3'te insan fiziksel özelliklerinin sınıflandırılmasında geliştirilen hibrit yaklaşım detaylandırılmıştır. Bölüm 4'te önerilen yöntemin performansı güncel diğer yöntemlerle

karşılaştırmalı olarak sunulmuştur. Bölüm 5'te ise elde edilen sonuçlara yer verilmiştir.

2. ESA MİMARİLERİ VE LEUNG-MALIK (LM) FİLTRELERİ (CNN ARCHITECTS AND LEUNG-MALIK (LM) FILTERS)

2.1. Leung-Malik (LM) Filtreleri (Leung-Malik (LM) Filters)

LM filtre bankası içerdiği farklı yön ve ölçeklerdeki filtrelerden oluşan bir filtre bankasıdır. Bu filtre bankası 48 adet filtre içermektedir. Bunların 36 tanesi 6 farklı açısız yön ve 3 farklı ölçekte tasarlanan Gaussian birinci ve ikinci türev filtreleridir. 8 tanesi Gaussian filtrelerin merkez-çevre farkı filtresi ve kalan 4 tanesi ise alçak-geçiren Gaussian filtredir. Herhangi bir görüntü üzerinde bu filtrelerin uygulanması ile her bir pikselin 1×48 boyutunda filtre yanıtlarından oluşan öznelik vektörü elde edilmiş olur. LM filtreleri sonucu elde edilen özellikler dönmeden bağımsızdır. Bu filtre bankası kullanılarak görüntülerdeki özel yapılar güçlü filtre yanıtlarının içerdikleri dokusal öznelikler ile tanımlanabilmektedir. Şekil 2'de LM filtre bankası gösterilmiştir.



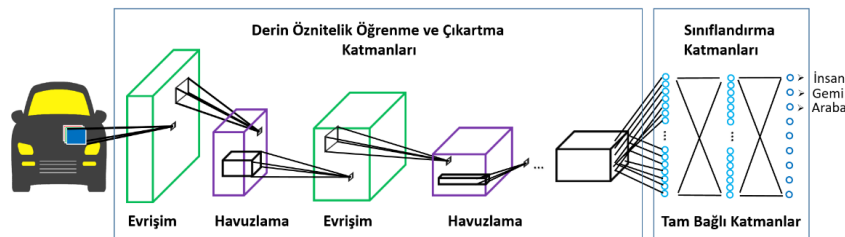
Şekil 2. Leung-Malik (LM) Filtreleri (Leung-Malik (LM) Filters)

2.2. Evrişimsel Sinir Ağı (ESA) (Convolutional Neural Network (CNN))

Günümüzde evrişimsel yapay sinir ağları (ESA, Convolutional Neural Network) görüntü tanıma ve sınıflandırmak için en yaygın kullanılan yöntemlerden biridir [18–20]. ESA mimarilerinin temel çalışma

prensibi hiyerarşik olarak giriş katmanından alınan görüntünün özneliklerinin çıkartılması ve çıkartılan özneliklerin tam bağlı katmanlarda sınıflandırılması şeklinde özetlenebilir [19–21]. Geliştirilmiş bir ESA mimarisinin geleneksel yöntemler gibi eğitim ve test aşamaları bulunmaktadır. Eğitim aşamasında; eğitim veri kümesi kullanılarak ESA mimarisindeki katmanların parametrelerinde en yüksek başarıyı elde edecek şekilde güncellemeler yapılmaktadır. Test aşamasında ise ağ katmanlarındaki parametreler sabit tutularak ESA mimarisinin sınıflandırma başarıları hesaplanmaktadır. ESA tabanlı makine öğrenmesi yöntemlerinin geleneksel makine öğrenmesi yöntemlerinden temel farkı ise, ESA'da öznelikleri otomatik bir şekilde çıkaran evrişim katmanlarının bulunmasıdır. ESA mimarileri bu evrişim katmanları sayesinde ele alınan problem için en uygun çözüme dayalı bir öznelik çıkartma modeli oluşturmaktadır [21, 22].

Bir ESA mimarisi temel olarak evrişim (convolution), havuzlama (pooling), aktivasyon ve tam bağlı katmanlardan oluşmaktadır (Şekil 3). Evrişim katmanları giriş verileri üzerinde bir filtreleme işlemi uygulayarak çıktı değerlerini oluşturmaktadır. Bu katmanın çıktısı $G \times Y \times D$ (Genişlik \times Yükseklik \times Derinlik) boyutları ile ifade edilmektedir [22]. Havuzlama katmanı boyut küçültme katmanı olarak da bilinmektedir. Bu katman genellikle evrişim katmanından sonra elde edilen yüksek boyutlu verilerin boyutunu azaltmak için kullanılmaktadır. Bu katman sayesinde önemli veriler alınarak bir boyut küçültme işlemine gidilmektedir. Aktivasyon katmanı ise gelen verilerde anlamlı bilgileri kaybetmeden bazı doğrusal değerlerin değiştirilmesine olanak sağlamaktadır. En çok bilinen aktivasyon katmanları ReLu ve Dropout katmanlarıdır. Son olarak sınıflandırma işlemleri için tam bağlı katmanlar kullanılmaktadır. Tam bağlı katmanlarda önceki katmanlardan gelen tüm veriler ağırlıklandırılarak birleştirilmektedir. Bu katmanlarda önceden belirlenmiş bir kayıp fonksiyonu kullanılarak eğitim boyunca en uygun nöron ağırlıkları bulunmaktadır. Bir ESA mimarisinde kullanılan tam bağlı katmanların sayısı ve katmanlardaki nöron sayısı (boyutu) ihtiyaca göre değişiklik gösterebilmektedir. Genellikle ağ mimarilerinin son katmanındaki tam bağlı katman için softmax gibi bir sınıflandırıcı görevi alan hata fonksiyonu kullanılmaktadır. Bu katmandaki nöron sayısı ayrıştırılmak istenilen sınıf sayısına eşittir. Bu nöronların çıktıları olasılıksal olarak sınıflandırma sonucunu vermektedir [22].



Şekil 3. Evrişimsel sinir ağı mimarisinin katmanları (Layers of convolutional neural network architecture)

3. YAYA ÖZELLİK TANIMA İÇİN HİBRİT EVRİŞİMSSEL YAPAY SİNİR AĞI MİMARİSİ (HYBRID CONVOLUTIONAL NEURAL NETWORK ARCHITECTURE FOR PEDESTRIAN ATTRIBUTE RECOGNITION)

Günümüzde birçok alanda kullanılan ESA mimarileri güçlü bir eğitim kümesine ihtiyaç duymaktadır. Fakat böyle bir eğitim kümesinin olmadığı durumlarda ESA mimarisinin katmanlarındaki parametreler doğru bir şekilde eğitilemeyecektir. Dolayısıyla yüksek seviyeli öznelikler ve başarılı bir sınıflandırıcı elde etmek imkânsız olacaktır. Bu problemlerden dolayı birçok çalışmada *Transfer Learning (TF)* metodlarına başvurulmuştur. TF metodunun temel çalışma prensibi sıfırdan bir ağ mimarisini kullanmak yerine önceden eğitilmiş bir ağ mimarisini güncelleyerek kullanmaya dayanmaktadır. Ancak ele alınan problem için önceden eğitilmiş benzer bir ağ mimarisini bulunmuyorsa, VGGnet gibi genel bir ağ mimarisini kullanılmaktadır. Fakat bu durumda kullanılan ESA mimarisini istenilen probleme çok uygun olmadığından, yüksek başarılar için ilk durumdaki gibi yine güçlü bir eğitim kümesine ihtiyaç duyulmaktadır.

Bu çalışmada yukarıdaki problemde yola çıkarak Hibrit Evrişimsel Yapay Sinir Ağı Mimarisi (HESA) önerilmiştir. Bu ağ mimarisindeki temel amaç LM filtrelerinden elde edilen öznelikler ile ESA mimarisinden elde edilen derin öznelikleri birleştirmek ve bu şekilde sınıflandırma için güçlü bir öznelik uzayı oluşturmaktır.

3.1. Geliştirilen HESA Mimarisi (The Developed HESA Architecture)

Şekil 4'te geliştirilen HESA mimarisinin temel yapısı görülmektedir. HESA mimarisini giriş olarak $128 \times 48 \times 3$

3 boyutunda alınan görüntülerin özneliklerini hesaplamaktadır. HESA mimarisinin üst kısmındaki ESA mimarisine derin öznelikler elde edilmektedir. HESA mimarisinin alt kısmında ise LM filtre bankaları kullanılarak geleneksel dönme ve açılma değişimden bağımsız öznelikler elde edilmektedir. İleriki katmanlarda bu iki öznelik grubu birleştirilerek çıkışa doğru ilerletilmektedir.

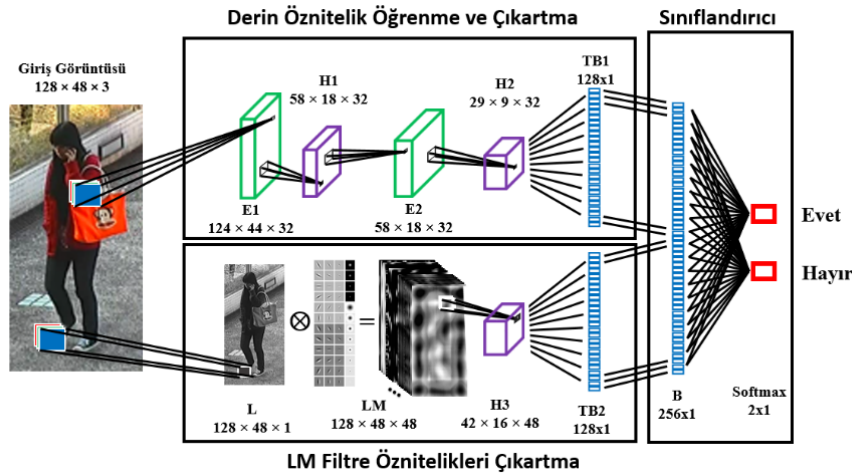
Geliştirilen HESA mimarisinin üst kısmını temsil eden ESA mimarisini evrişim (E), yığın normalizasyon (YN, Batch Normalization), aktivasyon (A) ve havuzlama (H) katmanlarından oluşmaktadır. İnşa edilen HESA mimarisinin katmanları hakkında detaylı bilgiler Çizelge 1'de verilmiştir. Evrişim katmanlarında denklem 1'de gösterildiği gibi giriş verileri üzerinde bir evrişim işlemi uygulanmaktadır.

$$Y_j = \sum_{j \in N} W_j * X + b_j \quad (1)$$

Buradaki X değeri giriş veri kümesini, Y_j değeri j 'nci evrişim filtresi için çıkış öznelik verisini, N değeri katmandaki filtre kümesini, $*$ simgesi evrişim işlevini ve son olarak b değeri bias değerini göstermektedir. Her evrişim katmanından sonra elde edilen veri kümesi yığın normalizasyon katmanından geçirilmektedir. Normalizasyon katmanlarının kullanılmasındaki temel amaç eğitim boyunca ESA mimarisinin eğitim veri kümesini ezberlemesini önlemektir. Normalizasyon katmanından sonra verilere ReLu aktivasyon işlemi uygulanmaktadır (Denklem 2).

$$f(v_i) = \max(0, v_i) \quad (2)$$

Buradaki temel işlem giriş değerindeki tüm negatif değerlerin sıfıra yükseltilmesidir.



Şekil 4. Geliştirilen HESA mimarisini (The developed HESA architecture)

Yukarıdaki işlemler sonucunda yüksek boyutlu bir veri kümesi elde edilmektedir. Bu veri kümesinin boyutunu küçültmek için en büyük değerleri alan bir havuzlama (Maxpooling) işlemi uygulanmaktadır.

Çizelge 1'den de görüleceği üzere ESA mimarisinde her bir evrişim katmanından sonra sırasıyla normalizasyon, aktivasyon ve havuzlama işlemi uygulanmıştır. Bu katmanlardan sonra elde edilen veri küpü, tam bağlı

katmanlara aktarılmak üzere boyut değiştirme (D, *Düzleştirme*) katmanından geçirilmiştir. Bu işlemler sonucunda Çizelge 1'de verildiği gibi 8352×1 boyutunda bir öznelik vektörü elde edilmiştir. Elde edilen bu öznelikler daha sonra bir tam bağlı katman yardımıyla 128 boyutuna indirgenmiştir. Sonuç olarak HESA mimarisinin üst kısmını oluşturan ESA mimarisi eğitim boyunca en yüksek başarıları elde edebilecek 128 farklı derin öznelik parametrelerini öğrenmektedir.

Geliştirilen HESA mimarisinde derin öznelikler ile birlikte yaya özellik tespitinde dönme, boyut değişikliği ve diğer çeşitli problemlerden etkilenmeyen dokusal öznelikler elde etmek için geleneksel LM filtre öznelikleri de hesaplanmıştır. LM filtre öznelikleri için öncelikle mimarinin giriş katmanından alınan görüntü *gri formata çevirme (L)* katmanından geçirilmektedir. Daha sonra 128×48 boyutundaki gri görüntüye LM filtreleriyle ayrı ayrı evrişim işlemi uygulanmaktadır (Denklem 3).

$$Y_j = \sum_{j \in N} W_j * X_{gri} \quad (3)$$

Bu denklemdeki X_{gri} değeri 128×48 boyutundaki gri giriş görüntüsünü, Y_j değeri J 'ninci LM filtresi için filtrelenmiş çıkış öznelikliğini ve * simgesi ise evrişim işlevini göstermektedir. Buradaki yapılan işlemler ESA mimarisindeki evrişim katmanlarından farklı olarak bias

değeri içermemektedir. LM filtreleme işlemi sonucunda giriş görüntüsünden $48 \times (128 \times 48)$ boyutunda bir öznelik küpü elde edilmiştir. Sayısal olarak çok fazla olan bu öznelik küpü Çizelge 1'de gösterildiği gibi havuzlama, düzleştirme ve tam bağlı (128 nöronlu) katmanlardan geçirilmektedir. Sonuç olarak HESA mimarisinin alt kısmında giriş görüntüsünden 128 farklı LM filtre öznelikleri elde edilmektedir.

HESA mimarisinin alt ve üst kollarından paralel şekilde elde edilen ayrı ayrı 1×128 boyutundaki öznelik vektörleri birleştirme katmanından geçirilerek 1×256 boyutunda hibrit bir öznelik vektörü oluşturulmuştur. Elde edilen bu öznelik vektörü görüntünün derin ve LM filtre temelli dokusal özneliklerinin birleşiminden oluşmaktadır. Son işlem olarak görüntünün aranan yaya öznelikliğini içerip içermediğini belirlemek için tüm öznelikler softmax sınıflandırma katmanına gönderilmektedir. Önceki katmanlardan gelen öznelikler işlenerek sınıflandırma sonucu elde edilmektedir (Denklem 4).

$$S = \text{Softmax}(FW + B) \quad (4)$$

Bu denklemde F öznelik kümesini ve S ise sınıflandırma tahminini göstermektedir. W ve B değerleri sırasıyla ağırlıklandırma ve bias değerlerini ifade etmektedir.

Çizelge 1. HESA mimarisinin katmanları (Layers of HESA architecture)

Katman Adı	Katman Tipi	Filtre Boyutu	Adım Genişliği	Çıkış Boyutu	Eğitilebilirliği
Giriş	Giriş			$128 \times 48 \times 3$	Hayır
E1	Evrışim	5×5	3	$124 \times 44 \times 32$	Evet
YN1	Yığın normalizasyonu			$124 \times 44 \times 32$	Evet
A1	Aktivasyon			$124 \times 44 \times 32$	Evet
H1	Havuzlama	3×3	3	$58 \times 18 \times 32$	Evet
E2	Evrışim	3×3	2	$58 \times 18 \times 32$	Evet
YN2	Yığın normalizasyonu			$58 \times 18 \times 32$	Evet
A2	Aktivasyon			$58 \times 18 \times 32$	Evet
H2	Havuzlama	3×3	2	$29 \times 9 \times 32$	Evet
D1	Düzleştirme			8352×1	Evet
TB	Tam bağlı katman			128×1	Evet
A3	Aktivasyon			128×1	Evet
L	Gri formata çevirme katmanı			$128 \times 48 \times 1$	Hayır
LM	LM filtre ile evrişim	49×49	1	$128 \times 48 \times 48$	Hayır
H3	Havuzlama	3×3	2	$42 \times 16 \times 48$	Hayır
D2	Düzleştirme			32256×1	Evet
TB2	Tam bağlı katman			128×1	Evet
A4	Aktivasyon			128×1	Evet
B	Birleştirme			256×1	Evet
Softmax	Tam bağlı katman			2×1	Evet

3.2. Uygulama Detayları (Application Details)

Geliştirilen HESA mimarisi *Tensorflow* tabanlı *Keras* kütüphanesi kullanılarak *Python* dili üzerinde tasarlanmıştır. Geliştirme ve uygulama testleri i7 6700HQ işlemci, 16 GB ram ve GeForce GTX 960M ekran kartına sahip bir bilgisayarda yürütülmüştür. Oluşturulan ağ mimarisinin eğitimi için *Adam* optimizasyonu kullanılmıştır. *Adam* optimizasyon algoritmasının öğrenme oranı başlangıç değeri olarak $\lambda = 1e - 3$ alınmıştır. Daha sonrasında eğitim oranının dinamik bir şekilde değişmesi için *Keras* kütüphanesindeki *ReduceLRonPlateau* modeli kullanılmıştır. Bu model temel olarak eğitim boyunca doğrulama veri kümesini kullanarak öğrenme oranında (λ) güncellemeler yapmaktadır. *ReduceLRonPlateau* modelinin temel çalışma prensibi şu şekildedir: doğrulama veri kümesinin başarı oranının belirli bir tur sayısı (*patience*) kadar sabit kaldığı durumlarda, öğrenme oranı sabit bir değer (*factor*) ile çarpılır ($\lambda_{yeni} = \lambda \times factor$) ve böylece öğrenme oranı düşürülür. Öğrenme oranı önceden belirlenmiş en düşük öğrenme oranına (λ_{min}) kadar azalma gösterebilmektedir. HESA mimarisi için *ReduceLRonPlateau* modelinin parametreleri sırasıyla *patience* = 5, *factor* = 5 ve $\lambda_{min} = 1e - 5$ şeklinde ayarlanmıştır.

ESA mimarilerinde tur sayısının doğru bir şekilde seçilmesi oldukça önemlidir. Tur sayısının küçük seçilmesi ağ mimarisinin eğitim verisini tam öğrenmemesine yol açarken, bu değerinin yüksek seçilmesi ise gereksiz zaman harcanmasına sebep olmaktadır. Bu problemten kaçınmak için *EarlyStopping* modeli kullanılmıştır. Bu model temel olarak doğrulama veri kümesinin başarı oranının belirli bir tur sayısı kadar sabit kaldığı durumlarda eğitim tamamlanma sinyali göndererek eğitimi durdurmaktadır. Bu çalışmada HESA mimarisi için *EarlyStopping* modeli kullanılmıştır. Yapılan uygulama sonuçlarında her bir yaya özneliğinin sınıflandırılması için *EarlyStopping* modeli sayesinde HESA mimarisinin 75 ile 150 arasında değişen tur sayıları kullandığı tespit edilmiştir.

Çizelge 2. Eğitim, doğrulama ve test veri tabanlarının boyutları ve yüzde oranları (%) (Dimensions and percentages (%) of training, validation and test datasets)

Peta	Eğitim	Doğrulama	Test
19003	13681	1521	3801
%100	%72	%8	%20

Peta veri kümesi uygulamalar için %72 eğitim, %8 doğrulama ve %20 test olacak şekilde ayrılmıştır (Çizelge 2). Veri kümeleri oluşturulduktan sonra HESA mimarisi kullanılarak Peta veri kümesindeki 35 farklı yaya özelliği ayrı ayrı sınıflandırılmıştır.

Peta veri kümesindeki bazı insan özelliklerine sahip örnek sayısı oldukça azdır. Örneğin eğitim veri kümesindeki bireylerin ceket giyenlerinin oranı % 8.6'dır. Bunun gibi bazı insan özelliklerinin sınıflandırılmasında en basit bir sınıflandırıcı tüm örnekleri negatif olarak öğrenmiş olsa bile %91.4 oranında başarılı olacaktır. Bundan dolayı etiketler arasında dengesizliğin olduğu yaya özelliklerinin olumlu ve olumsuz örnek sayılarının eşit alınması daha doğru sınıflandırma başarısı sağlayacaktır. Yapılan çalışmada yukarıda bahsedilen problemten kaçınmak için her bir yaya özellik sınıflandırılmasında etiketler arasındaki oran 0.35 ile 0.65 arasında değilse, eğitim modeli için etiketler arası örnek sayısı eşit sayıda almıştır.

4. DENEYSSEL SONUÇLAR (EXPERIMENTAL RESULTS)

Geliştirilen HESA mimarisinin performans değerlendirmeleri için PETA veri kümesi üzerinde testler gerçekleştirilmiştir. Elde edilen sınıflandırma sonuçları ikSVM ve MRFG2/MRFR2 yöntemleri ile kıyaslanmıştır [5, 6]. HESA yöntemi ile kıyaslanan yöntemler kısaca şu şekilde özetlenebilir. ikSVM yöntemi DVM tabanlı bir sınıflandırıcı kullanmış ve öznelik olarak öncelikle her bir görüntüyü 8 renk kanalına (RGB, HSV ve YCbCr) ayırtmıştır. Ayrıca Gabor ve Schmid filtreleri kullanılarak görüntülerden 21 farklı doku özneliği çıkartılmıştır. MRFG2 ve MRFR2 yöntemlerinde ise yaya görüntülerindeki insan özellik tespiti için Markov Rastgele Alan (MRF) yöntemi kullanılmıştır [6]. MRF yöntemi ile oluşturulan grafik topolojisi yardımı ile yaya görüntüleri arasındaki ikili benzerlikler ağırlıklandırılmıştır. Bu şekilde bir görüntüdeki yaya özellik tespiti için komşu görüntülerdeki bilgiler de kullanılarak performans başarısının artırılması hedeflenmiştir. Bu doğrultuda MRF yönteminin çekirdek fonksiyonu değiştirilerek 4 farklı alt yöntem geliştirilmiştir. Yaptığımız çalışmada bu yöntemlerden yüksek başarılar elde eden MRFG2 ve MRFR2 alt yöntemlerinin sonuçları kullanılmıştır. Bu iki model arasındaki temel fark ise yöntemlerde Gauss (MRFG2) ve rastgele ağaçlar (MRFR2) çekirdek fonksiyonlarının ayrı ayrı kullanılmasıdır. ikSVM, MRFG2 ve MRFR2 yöntemleri hakkında daha fazla bilgiye [5, 6] nolu çalışmalardan ulaşılabilir. Geliştirilen HESA yönteminin ikSVM, MRFG2 ve MRFR2 yöntemlerinden temel farkı, geleneksel ve derin öznelikleri bir arada kullanmasıdır. Bu sayede sınıflandırıcı için görüntülerden geleneksel özneliklerin yanında daha güçlü derin öznelikleri çıkartarak sınıflandırıcının ayırt edici gücü artırılmıştır.

Çizelge 3. Peta veri tabanındaki yaya özellik sınıflandırma sonuçları (Pedestrian attribute classification results in Peta dataset)

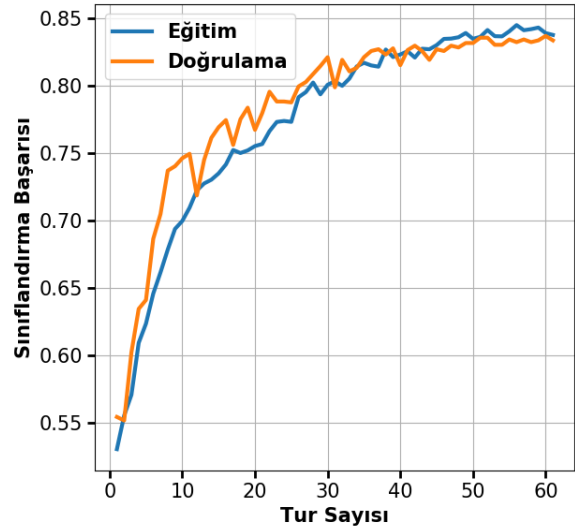
Yaya Özelliği	Yöntem			
	HESA	ikSVM	MRFg2	MRFr2
yaş16-30	81.08	80.40	81.70	83.80
yaş31-45	79.08	73.60	76.20	78.80
yaş46-60	79.35	73.10	75.20	76.40
yaş-60_ve_üstü	93.66	87.20	88.20	89.00
taşıma_sirt_çantası	73.79	66.70	67.10	67.20
taşıma_diğer	71.06	64.60	66.80	68.00
alt_giysi_gündelik	82.66	70.70	71.60	71.30
üst_giysi_gündelik	80.79	70.30	71.20	71.30
alt_giysi_resmi	85.98	71.00	71.80	71.90
üst_giysi_resmi	86.48	70.00	70.40	70.00
takı_şapka	93.19	82.30	84.30	86.70
ceket	80.21	67.70	68.40	67.90
kot	69.64	74.90	76.10	76.00
deri_ayakkabı	82.66	78.90	80.90	81.70
uzun_saç	86.92	71.50	72.60	72.80
cinsiyet_erkek	83.27	79.70	80.90	81.40
bel_çantası	78.16	71.80	74.30	75.50
takı_atkı	78.61	88.00	89.50	91.30
takı_yok	83.11	76.80	78.60	80.00
taşıma_yok	72.01	70.40	71.60	71.50
taşıma_naylon_poşet	87.32	74.90	75.50	75.50
sandalet	97.92	50.30	50.30	50.30
ayakkabı	77.58	70.60	72.50	73.60
şort	96.61	56.00	56.50	56.50
giysi_kısa_kol	72.90	71.30	71.80	71.60
spor_ayakkabı	69.35	67.50	69.00	69.30
takı_güneş_gözlüğü	96.84	52.40	52.40	51.70
pantolon	55.33	74.00	75.70	76.50
tişört	91.08	64.30	64.60	64.20
giysi_diğer	83.69	80.70	81.80	83.90
v_yaka	98.63	51.10	51.10	51.10
giysi_kareli	97.63	64.00	64.50	65.00
logo	95.76	51.10	51.10	50.70
giysi_çizgili	97.24	51.50	52.30	52.30
etek	92.79	64.00	64.00	64.30
Ortalama Doğruluk	83.78	69.52	70.59	71.11

Çizelge 3'teki sonuçlara bakıldığında, sınıflandırma için kullanılan 35 yaya özelliğinin 31 tanesinde önerilen HESA yöntemi diğer yöntemlere göre daha yüksek sonuçlar vermiştir. Çizelgenin son satırında görüleceği gibi, önerilen HESA yöntemi %83.78 oranında bir ortalama başarıya ulaşmıştır. Diğer yöntemler ise HESA yöntemine göre ortalama %13 oranında bir fark ile geride kalmıştır. Bunun temel sebebi geleneksel yöntemlerin yüksek seviyeli olmayan ve veriyi temsil etme kabiliyeti düşük olan öznelikler kullanmalarıdır. Yaya görüntülerinde bulunan ve daha önce bahsedilen zorluklar geleneksel özellik çıkartma yöntemlerini ciddi bir oranda zayıflatmıştır. Fakat önerilen yöntem bu zorluklardan kaçınmak için geleneksel özneliklerin yanında problem ve veri kümesine göre otomatik bir şekilde elde edilen derin öznelikleri de kullanmaktadır. Diğer yöntemlerin düşük başarılar gösterdiği logo, v_yaka, etek, kareli_giysi ve çizgili_giysi gibi insan özelliklerinde HESA yöntemi ortalama olarak %90 oranında başarılı sonuçlar vermiştir. Bunun temel sebebi; kıyafet gibi insan özellik tespitinde doku özneliği ifade eden aksesuarların tespit edilmesinde LM filtrelerinin elde ettiği dokusal özelliklerin sınıflandırmada HESA yöntemine büyük bir avantaj sağlamasıdır

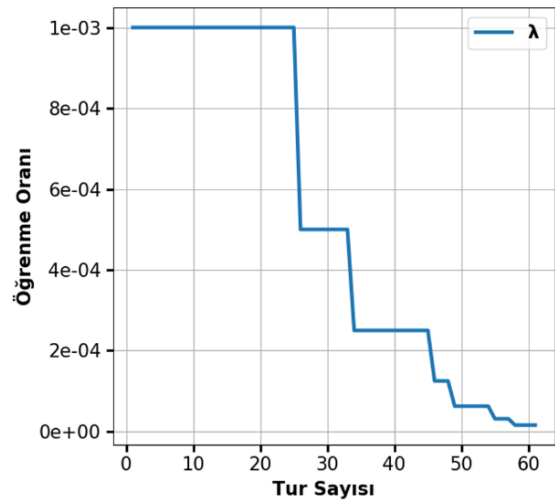
Bunun yanında logo, v_yaka, sandalet gibi LM filtreleri için oldukça zor problemler teşkil eden yaya aksesuarlarında ise HESA yöntemi uyarlamalı derin öznelikleri kullanarak diğer yöntemler karşısında avantajlı durumdadır. Bu sınıflar karşısında diğer yöntemler ise sadece renk ve dokusal öznelikleri kullandıkları için düşük sınıflandırma sonuçları üretmişlerdir. Bununla birlikte, veri kümesindeki kot, takı_atkı ve pantolan gibi yaya özelliklerinde MRF temelli yöntemler daha başarılı sonuçlar elde etmiştir. Kot pantolon gibi görüntüde belli oranda belirgin şekilde yer alan yaya özelliğinde piksel komşuluk bilgisi temelli çizge topoloji tekniği ile ayırt edici öznelikler çıkartılabilmektedir. 8 renk kanalı ve 21 doku kanalı ile üç farklı yaya özelliği daha iyi bir şekilde analiz edilebilmiştir. MRF temelli yöntemler 2784×1 boyutunda öznelik vektörü inşa etmiştir. Önerilen HESA yönteminde ise öznelik vektörü 256×1 boyutundadır. Sonuçlardan anlaşılacağı üzere bazı yaya özelliklerinde yüksek boyutlu özneliklerin MRFg2 ve MRFr2 yöntemlerine avantaj sağladığı anlaşılmıştır. Dolayısıyla daha yüksek başarılar elde etmiştir. ikSVM yönteminde ise, oldukça zor bir problem olan yaya özellik sınıflandırmada çıkartılan renk kanalı temelli öznelikler yetersiz kalmış ve dolayısıyla daha düşük sonuçlar elde edilmiştir. Genel olarak bakıldığında, özellikle benzer renk ve doku bilgisi içeren yaya aksesuarları sınıflarında MRFg2, MRFr2 ve DVM tabanlı yöntemler gürültü ve görüntülerin üst üste çakışması gibi istenmeyen durumlardan olumsuz etkilenmiştir.

Önerilen HESA yönteminin yüksek sınıflandırma başarıları elde etmesinde oldukça önemli bir katkıyı ise kullanılan *ReduceLRonPlateau* modeli sağlamıştır. Bu model sayesinde geliştirilen derin ağ mimarisi dinamik

bir öğrenme oranı kullanarak eğitim sürecinde ağ mimarisini daha güçlü bir şekilde eğitebilmiştir



(a)



(b)

Şekil 5. Cinsiyet sınıflandırma için öğrenme oranı ve sınıflandırma oranının analizi. (a) sınıflandırma başarıları, (b) öğrenme oranı. (Classification accuracy (a) and learning rate (b) Analysis of learning rate and classification accuracy for gender classification. (a) classification success, (b) learning rate).

ReduceLRonPlateau modelinde derin ağ mimarisi eğitim sürecinin hemen başında yüksek öğrenme oranı kullanmış ve böylece ağ mimarisinin belli bir noktaya kadar büyük adımlar atarak hızlı bir öğrenme stratejisine sahip olması hedeflenmiştir. Bu noktadan sonra öğrenme oranı düşürülerek daha yavaş fakat detaylı bir öğrenme modeline geçilmiştir. Örnek olarak veri tabanında bulunan yaya cinsiyet sınıflandırmasının eğitim süresince öğrenme oranındaki geçişler ve elde ettiği başarı oranı Şekil 5'te verilmiştir. Burada görüldüğü gibi öğrenme oranının sırasıyla 26, 33 ve 46'inci tur sayılarında ezberlemeden uzak olmakla birlikte kararlı ve tutarlı değişimler gösterdiği görülmektedir.

5. SONUÇ (CONCLUSION)

Günümüzde oldukça popüler olan ESA mimarileri birçok farklı problemde oldukça yüksek başarılar sunmaktadır. Fakat ESA mimarileri yüksek başarılar karşısında güçlü bir eğitim veri kümesine ihtiyaç duymaktadır. Önerilen HESA mimarisi sayesinde veri kümesinin kısıtlı olduğu problemler için doku tabanlı öznelik çıkartan LM filtreleri ile derin öznelikleri birleştiren güçlü bir özellik çıkartma ve sınıflandırma yapısı oluşturulmuştur. Oluşturulan bu hibrit ESA mimarisi oldukça zor bir problem olan yaya özellik tanıma problemine uygulanmıştır. Elde edilen sonuçlara bakıldığında, bu mimari sayesinde daha güçlü bir ESA mimarisi geliştirme olanağı sağlanabileceği görülmüştür.

İleriki dönemlerde ESA mimarisi için LM filtre özneliklerinin yanında, farklı geleneksel öznelikler eklenerek daha güçlü derin öğrenme mimarileri geliştirmek için çalışmalar yapılacaktır. Ayrıca otokodlayıcı temelli derin öğrenme mimarilerinin geleneksel özellik çıkartma yöntemleri ile birlikte kullanılmalarının etkilerini gözlemek üzere çalışmalar yapılması planlanmaktadır.

6. KAYNAKLAR (REFERENCES)

- [1] R. Layne, T. M. Hospedales, and S. Gong, "Attributes-Based Re-identification", *Person Re-Identification*, London: Springer London, 93–117, (2014).
- [2] M. Hirzer, C. Belezni, P. M. Roth, and H. Bischof, "Person Re-identification by Descriptive and Discriminative Classification", *Scandinavian conference on Image analysis*, Springer, Berlin, Heidelberg, 91–102, (2011).
- [3] L. Cao, M. Dikmen, Y. Fu, and T. S. Huang, "Gender recognition from body", *Proceeding of the 16th ACM international conference on Multimedia*, New York, New York, USA: ACM Press, 725–728, (2008).
- [4] J. Zhu, S. Liao, Z. Lei, and S. Z. Li, "Multi-label convolutional neural network based pedestrian attribute classification", *Image Vis. Comput.*, 58: 224–229, (2017).
- [5] Y. Deng, P. Luo, C. C. Loy, and X. Tang, "Pedestrian Attribute Recognition At Far Distance", *Proceedings of the ACM International Conference on Multimedia - MM '14*, New York, New York, USA: ACM Press, 789–792, (2014).
- [6] Y. Deng, P. Luo, C. C. Loy, and X. Tang, "Learning to Recognize Pedestrian Attribute", *arXiv Prepr. arXiv1501.00901*, (2015).
- [7] R. Layne, T. M. Hospedales, and S. Gong, "Towards Person Identification and Re-identification with Attributes", *European Conference on Computer Vision*, Springer, Berlin, Heidelberg, 402–412, (2012).
- [8] T. Matsukawa and E. Suzuki, "Person re-identification using CNN features learned from combination of attributes", *2016 23rd International Conference on Pattern Recognition (ICPR)*, 2428–2433, (2016).
- [9] E. S. Jaha and M. S. Nixon, "Soft biometrics for subject identification using clothing attributes", *IEEE International Joint Conference on Biometrics*, 1–6, (2014).
- [10] L. An, Xiaojing Chen, M. Kafai, S. Yang, and B. Bhanu, "Improving person re-identification by soft biometrics based reranking", *2013 Seventh International Conference on Distributed Smart Cameras (ICDSC)*, 1–6, (2013).
- [11] J. Friedman, T. Hastie, and R. Tibshirani, "Additive logistic regression: a statistical view of boosting (With discussion and a rejoinder by the authors)", *Ann. Stat.*, 28(2): 337–407, (2000).
- [12] J. Zhu, S. Liao, Z. Lei, D. Yi, and S. Li, "Pedestrian Attribute Classification in Surveillance: Database and Evaluation", *The IEEE International Conference on Computer Vision (ICCV) Workshops*, 331–338, (2013).
- [13] J. Zhu, S. Liao, Z. Lei, and S. Z. Li, "Improve Pedestrian Attribute Classification by Weighted Interactions from Other Attributes", *Computer Vision - ACCV 2014 Workshops. ACCV 2014. Lecture Notes in Computer Science*, 545–557, (2014).
- [14] L. Bourdev, S. Maji, and J. Malik, "Describing people: A poselet-based approach to attribute classification", *2011 International Conference on Computer Vision*, 1543–1550, (2011).
- [15] C. Su, S. Zhang, J. Xing, W. Gao, and Q. Tian, "Deep Attributes Driven Multi-Camera Person Re-identification", *European conference on computer vision*, 475–491, (2016).
- [16] J. Zhu, S. Liao, D. Yi, Z. Lei, and S. Z. Li, "Multi-label CNN based pedestrian attribute learning for soft biometrics", *2015 International Conference on Biometrics (ICB)*, 535–540, (2015).
- [17] P. Liu, X. Liu, J. Yan, and J. Shao, "Localization Guided Learning for Pedestrian Attribute Recognition", *Localization Guid. Learn. Pedestr. Attrib. Recognit.*, (2018).
- [18] A. J. O'Toole, C. D. Castillo, C. J. Parde, M. Q. Hill, and R. Chellappa, "Face Space Representations in Deep Convolutional Neural Networks", *Trends Cogn. Sci.*, 22(9): 794–809, (2018).
- [19] Y. Seo and K. Shin, "Hierarchical convolutional neural networks for fashion image classification", *Expert Syst. Appl.*, 116: 328–339, (2019).
- [20] L. A. Gatys, A. S. Ecker, and M. Bethge, "Texture and art with deep neural networks", *Curr. Opin. Neurobiol.*, 46, 178–186, (2017).
- [21] K. Simonyan and A. Zisserman, "Very Deep Convolutional Networks for Large-Scale Image Recognition", *arXiv preprint arXiv:1409.1556*, (2014).
- [22] "CS231n Convolutional Neural Networks for Visual Recognition." [Online]. Available: <http://cs231n.github.io/convolutional-networks/#conv>. [Accessed: 17-Dec-2018].