

**Atf İçin:** Akar F, Akgül İ, 2022. Derin Öğrenme Modeli ile Yüz İfadelerinden Duygu Tanıma. İğdır Üniversitesi Fen Bilimleri Enstitüsü Dergisi, 12(1): 69-79.

**To Cite:** Akar F, Akgül İ, 2022. Emotion Recognition from Facial Expressions by Deep Learning Model. Journal of the Institute of Science and Technology, 12(1): 69-79.

## **Derin Öğrenme Modeli ile Yüz İfadelerinden Duygu Tanıma**

Funda AKAR<sup>1\*</sup>, İsmail AKGÜL<sup>1</sup>

**ÖZET:** Yüz ifadesinden duygu tanıma; insan-bilgisayar etkileşimi, duygusal hesaplama vb. gibi birçok bilgisayarla görme alanında uygulanabilen güncel bir araştırma konusudur. Bu çalışmada, KDEF ve PICS veri setleri kullanılarak derin öğrenme ile duygu tanıma yönelik bir uygulama yapılmıştır. Öznitelik çıkarımı için derin öğrenme tekniklerinden olan ve yapay sinir ağları içeren bir yapay zekâ yaklaşımı olan Evrişimsel Sinir Ağı (ESA) kullanılarak yeni bir model geliştirilmiştir. Derin öğrenmenin yüksek başarımı için büyük veri setine ihtiyaç duyulmaktadır. KDEF veri setinde 4900, PICS veri setinde 322 görüntü bulunmaktadır. Bu amaçla öncelikle PICS veri setindeki görüntü sayısının az olmasından dolayı veri artırma yöntemi ile görüntü çoğaltma işlemi uygulanmıştır ve PICS veri seti 4830 görüntüye çıkarılmıştır. Daha sonra bu iki farklı veri seti üzerinde ayrı ayrı eğitim gerçekleştirilerek geliştirilen yeni model test edilmiştir. ESA modellerinden olan VGGNet temel alınarak geliştirilen yeni model ile gerçekleştirilen çalışmada, her bir veri setinde yedi farklı duygu sınıfı (korku, öfke, iğrenme, mutluluk, nötr, üzüntü, şaşırma) ele alınmıştır. Geliştirilen model ile KDEF veri setinin geçerleme kümesinde %97.44, PICS veri setinin geçerleme kümesinde %98.24 doğruluk değerleri elde edilerek yüksek bir başarı oranına ulaşılmıştır.

**Anahtar Kelimeler:** Derin öğrenme, evrişimsel sinir ağı, duygu tanıma

### **Emotion Recognition from Facial Expressions by Deep Learning Model**

**ABSTRACT:** Emotion recognition from facial expression is a current research topic that can be applied in the many fields of computer vision, such as human-computer interaction, emotional computing, etc. In this study, an application for emotion recognition through deep learning was made using KDEF and PICS datasets. A new model was established using the convolutional neural network (CNN), an artificial intelligence approach that involves artificial neural networks, which is one of the deep learning techniques for attribute inference. Large datasets are needed for the high performance of deep learning. There are 4900 images in the KDEF dataset and 322 images in the PICS dataset. For this purpose, primarily due to the small number of images in the PICS dataset, image iteration was applied with the data augmentation method, and the PICS dataset was increased to 4830 images. Then, the new model developed by conducting separate training on these two different datasets was tested. Seven different classes of emotion (afraid, angry, disgusted, happy, neutral, sad, surprised) were covered in each dataset in the study conducted with a new model developed based on VGGNet which is one of the CNN models. With the developed model, a high success rate was achieved by obtaining 97.44% accuracy values in the validation set of the KDEF and 98.24% accuracy values in the validation set of the PICS dataset.

**Keywords:** Deep learning, convolutional neural network, emotion recognition

<sup>1</sup>Funda AKAR ([Orcid ID: 0000-0001-9376-8710](https://orcid.org/0000-0001-9376-8710)), İsmail AKGÜL ([Orcid ID: 0000-0003-2689-8675](https://orcid.org/0000-0003-2689-8675)), Erzincan Binali Yıldırım Üniversitesi, Mühendislik-Mimarlık Fakültesi, Bilgisayar Mühendisliği Bölümü, Erzincan, Türkiye

\*Sorumlu Yazar/Corresponding Author: Funda AKAR, e-mail: fakar@erzincan.edu.tr

## GİRİŞ

Yüz ifadesi, insanların günlük iletişimde içsel duyguları ifade etmenin en açık göstergelerinden biridir. Bir kişinin fiziksel veya ruhsal durumu, yüz ifadeleri analiz edilerek elde edilebilir. Bu nedenle yüz ifadesi tanıma; otopilot, insan-bilgisayar etkileşimi, tıbbi tedavi ve yüz ifadesi ile alakalı diğer alanlarda büyük öneme sahiptir ve giderek daha önemli bir araştırma yönü haline gelmiştir. Yüz ifadesinden duygu tanıma, otizm ve şizofreninin belirlenmesi, uykulu bir sürücünün saptanması, alzheimer hastalığı veya şizofreninin erken aşamalarında anormalliklerin belirlenmesi ve suç tahmin sistemleri gibi çeşitli uygulamalar için kullanılmıştır (Mehta ve ark., 2019). Ayrıca yüz ifadeleri eğitim alanını uygulamalarında da kullanılmaktadır (Özdemir ve Karaman, 2017).

Otomatik duygu tanımanın kullanımı; dijital reklam, çevrimiçi oyunlar, müşterilerin geri bildirim değerlendirmesi, sağlık hizmetleri gibi çeşitli akıllı sistemlerde (e-sağlık, öğrenme, turizm için öneri, akıllı şehir, akıllı konuşma vb.) büyük bir potansiyele sahiptir. Bu sistemlerin çoğu metin veya ifade verilerine dayanmaktadır (Hossain ve Muhammad, 2019).

Makine öğrenmesinde, çeşitli yüz ifadesi tanıma algoritmaları önerilmiştir. Ancak görüntülerdeki karmaşıklık, çeşitlilik, üst üste gelme, aydınlatma problemleri ve yüz ifadesi tanımadaki diğer zorluklar nedeniyle, pratik uygulamalardaki tanıma doğruluğu hala tatmin edici değildir (J. Li ve ark., 2020).

Son yıllarda sinir ağlarında derin öğrenme, etkili modeller ortaya çıkaran çok sayıda derin öğrenme mimarisine yol açmış ve çeşitli alanlarda çok büyük başarılar elde etmiştir (Fayek ve ark., 2017). Bu alanlardan biri olan yüz ifadelerini tanıma, bilgisayarlı görme ve yapay zekâ alanında önemli bir role sahiptir (Dandıl ve Özdemir, 2019). Derin öğrenme, duygu tanıma problemleri için sınıflandırma verimliliği konusunda gerçek bir umut vadetmiştir (Ruiz-Garcia ve ark., 2016). Özellikle derin evrişimli ağlar görüntülerin, videoların, konuşmanın ve sesin işlenmesinde çığır açmıştır, yinelenen ağlar ise metin ve konuşma gibi sıralı verilere ışık tutmuştur. 2000'lerin başından beri, ConvNets, görüntülerdeki nesnelerin ve bölgelerin algılanması, bölümlendirilmesi ve tanınması için büyük bir başarıyla uygulanmaktadır. Bunlara örnek olarak trafik işareti tanıma, biyolojik görüntülerin özellikle konektomikler için bölümlere ayrılması ve doğal görüntülerde yüzlerin, metinlerin, yayaların ve insan vücutlarının tespiti gibi etiketlenmiş verilerin nispeten fazla olduğu uygulamalar, otonom robotlar, otonom araçlar verilebilir. ConvNets'in yakın zamandaki uygulamalı en büyük başarısı yüz tanımadır (Lecun ve ark., 2015).

Dandıl ve Özdemir 2019 yılında yaptıkları çalışmada klasik evrişimsel sinir ağı AlexNet ile gerçek zamanlı video karelerinden yüz ifadeleri temelinde bir duygu tanıma sistemi önermiştir (Dandıl ve Özdemir, 2019). Chen ve ark. 2015 yılında yaptıkları çalışmada resim görüntülerinden duygu sınıflandırması yapmak için kullandıkları otomatik öğrenen evrişimsel sinir ağı (ESA) ile elle hazırlanmış özelliklere dayanan geleneksel yöntemlerden daha iyi performans almışlardır (Chen ve ark., 2015). Yapılan çalışmalar ESA'nın duygu tanıma özelliği çıkarmak için kullanılabileceğini göstermektedir.

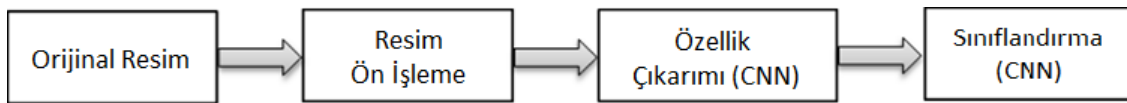
Bu çalışmada, derin öğrenme tekniklerinden olan evrişimsel sinir ağı kullanılarak yüz ifadelerini içeren görüntüler üzerinden bir duygu tanıma modeli önerilmiştir. Önerilen model klasik evrişimsel sinir ağlarından olan VGGNet temel alınarak yeniden oluşturulmuş bir modeldir. KDEF ve PICS veri setlerinden elde edilen görüntüler; korku, öfke, iğrenme, mutluluk, nötr, üzüntü, şaşırma olmak üzere yedi duygu için sınıflandırılmıştır. Kullanılan veri setleri ayrı ayrı ESA ile oluşturulan model ile eğitilmiştir.

Çalışmanın geri kalanı aşağıdaki gibi yapılandırılmıştır. Bölüm 2’de veri seti ve ön işleme, geliştirilen ESA modeli, kullanılan ortam ve kütüphaneler hakkında bilgiler verilmiştir. Bölüm 3’te deneysel sonuçlar verilmiştir. Bölüm 4’te çalışma sonuçlandırılmıştır.

## MATERYAL VE METOT

### Veri Seti ve Ön İşleme

Bu çalışmanın amacı, çeşitli görüntü türlerindeki farklı yüz ifadelerini otomatik olarak ve doğru bir şekilde tanıyan bir tanıma modeli tasarlamaktır. Genel olarak yüz ifadesi tanıma süreci; 1) resimlerin ön işlenmesi, 2) yüz ifadelerinin özellik çıkarımı ve 3) yüz ifadelerinin sınıflandırılması adımlarından oluşur. Yüz görüntüsünün ön işlemden geçirilmesi, yüz ifadesi tanımda önemli bir adım olarak kabul edilir (Eng ve ark., 2019). Önerilen yöntemin akış diyagramı Şekil 1’de gösterilmiştir.



Şekil 1. Önerilen yöntemin akış diyagramı

Resim ön işleme adımında, öncelikle orijinal boyuttaki resimlerde OpenCV CropFace görüntü işleme tekniği kullanılarak gözlerin konumu tespit edilip yüz bölgesinin, kaşların hemen üstünden ve çene hizasından olacak şekilde Python kodu ile kesilerek resimler otomatik olarak yeniden boyutlandırılmıştır. Daha sonra tüm resimler 96\*96 piksel boyutunda gri seviyeye normalize edilmiştir. Son olarak Evrişimsel Sinir Ağı kullanılarak özellik çıkarımı ve sınıflandırma yapılmıştır. Evrişimsel Sinir Ağı Modeli Bölüm 2.2 de anlatılmıştır.

Duygu analizi konusunda kullanılan birçok veri seti bulunmaktadır. Çalışmada kolay erişilebilir ve kullanıma açık olduğu için kontrollü ortamlarda hazırlanmış KDEF (Lundqvist ve ark., 1998) ve PICS-Pain Expressions (Hancock, 2008) veri setleri kullanılmıştır. KDEF veri seti fazla sayıda görüntü içermesi bakımından derin öğrenme yöntemine uygun olduğu düşünülerek seçilmiştir. Daha az sayıda görüntü içeren PICS veri seti ise görüntü artırma yönteminin de uygulanıp sonuca etkisinin görülmesi amacıyla tercih edilmiştir.

### KDEF veri seti ve ön işleme

Karolinska Directed Emotional Faces (KDEF) veri seti, insan yüz ifadelerini içeren toplam 4900 resimden oluşmaktadır. Resim kümesi; 7 temel duygusal ifade (korku, öfke, iğrenme, mutluluk, nötr, üzüntü, şaşırma) sergileyen 20-30 yaş arası sakal, bıyık, küpe, gözlük ve görünür bir makyaj olmayan özel gri tişörtler giymiş 35 kadın ve 35 erkek olmak üzere toplam 70 kişi içermektedir. Her duygu ifadesi 5 farklı açıdan (tam sol, yarım sol, düz, yarım sağ, tam sağ) iki kez (A serisi, B serisi) görüntülenmiştir. Veri seti; boyutu 562\*762 piksel, çözünürlüğü 72 \* 72 dpi olan renkli 32 bit jpeg uzantılı resimlerden oluşmaktadır (Lundqvist ve ark., 1998).

Çalışmada KDEF veri setinde yer alan toplam 4900 resmin tamamı kullanılmıştır. Ele alınan resimler, gözlerin lokasyonu dikkate alınarak python kodu ile yüz hatlarını alacak şekilde 96\*96 piksel olarak yeniden boyutlandırılmış ve gri seviyeye (8 bit) dönüştürülmüştür. Bu ön işlemden sonra KDEF veri setine ait resimler ait olduğu duygunun altında etiketlenerek gruplandırılmıştır. Böylece çalışmada kullanılacak birinci veri seti oluşturulmuştur (Şekil 2).



Şekil 2. KDEF veri setinde 7 duyguya ait farklı açılardan görüntülenmiş a) Orijinal ve b) Ön İşlenmiş örnek resimler

### PICS-Pain Expressions veri seti ve ön işleme

Psychological Image Collection at Stirling (PICS) veri seti, 2D insan yüz ifadelerini içeren toplam 599 resimden oluşmaktadır. Resim kümesi; 7 temel duygusal ifadeye (korku, öfke, iğrenme, mutluluk, nötr, üzüntü, şaşırma) ek olarak 10 ağrı ifadesi sergileyen farklı kıyafetler giymiş 13 kadın ve 10 erkek olmak üzere toplam 23 kişi içermektedir. Her duygu ifadesi tam karşıdan iki kez görüntülenmiştir. Veri seti; boyutu 720\*576 piksel, çözünürlüğü 72\*72 dpi olan renkli 24 bit jpeg uzantılı resimlerden oluşmaktadır (Hancock, 2008).

PICS veri setinde ağrı içeren resimler çıkarılarak, 7 temel duygusal ifade içeren tek açıdan iki kez görüntülenmiş resimler alınarak her duygu için 46 (23 kişi, 2 kez) resim kullanılmıştır. Bu veri setinde kullanılan resimler de KDEF veri setindeki gibi, gözlerin lokasyonu dikkate alınarak python kodu ile yüz hatlarını alacak şekilde 96\*96 piksel olarak yeniden boyutlandırılmış ve gri seviyeye (8 bit) dönüştürülmüştür. Derin öğrenme kullanılarak duygu tanımayı gerçekleştirebilmek için veri setinde yeterli sayıda görüntünün olması gerekmektedir. Ağın iyi bir genelleme becerisine ulaşmasını sağlamak için yeterli eğitim verisi gereklidir. Ancak, PICS gibi halka açık veri setlerinin çoğunda eğitim için yeterli sayıda görüntü bulunmamaktadır. Veri miktarının az olması aşırı uyum sorununa ve ağın ezberlemesine yol açabilir. Bu nedenle veri artırma işlemi, yüz ifadesi tanıma için önemlidir (Huang ve ark., 2019; Shorten ve Khoshgoftaar, 2019; J. Li ve ark., 2020; S. Li ve Deng, 2020).

PICS veri setindeki resim sayısının az olmasından dolayı, bu çalışmada yetersiz verinin neden olduğu aşırı uyum ve ezberleme sorununu gidermek amacıyla, ağı eğitmek için veri artırma yöntemi kullanılmıştır. Veri artırma işleminde; 15 derece döndürme, yüzde 5 oranında genişliği kaydırma, yüksekliği kaydırma, kesme, yakınlaştırma ve yatayda döndürme yöntemleri farklı kombinasyonlarla uygulanmıştır. Şekil 3'te görüldüğü üzere her resimden 14 tane görüntü çoğaltılarak toplam 15 resim elde edilmiştir. Dolayısıyla her duygu için 690 resim elde edilerek, 7 duygu için toplam 4830 resim ele alınmıştır. Veri artırma ön işlemleri bittikten sonra PICS veri setine ait resimler ait olduğu duygunun altında etiketlenerek gruplandırılmıştır. Böylece çalışmada kullanılacak ikinci veri seti oluşturulmuştur (Şekil 4).





Şekil 3. PICS veri setine uygulanan veri artırma yöntemi sonucunda elde edilen Artırılmış örnek resimler



Şekil 4. PICS veri setinde 7 duyguya ait Orijinal ve Ön İşlenmiş örnek resimler

### Evrişimsel Sinir Ağı (ESA) Modeli

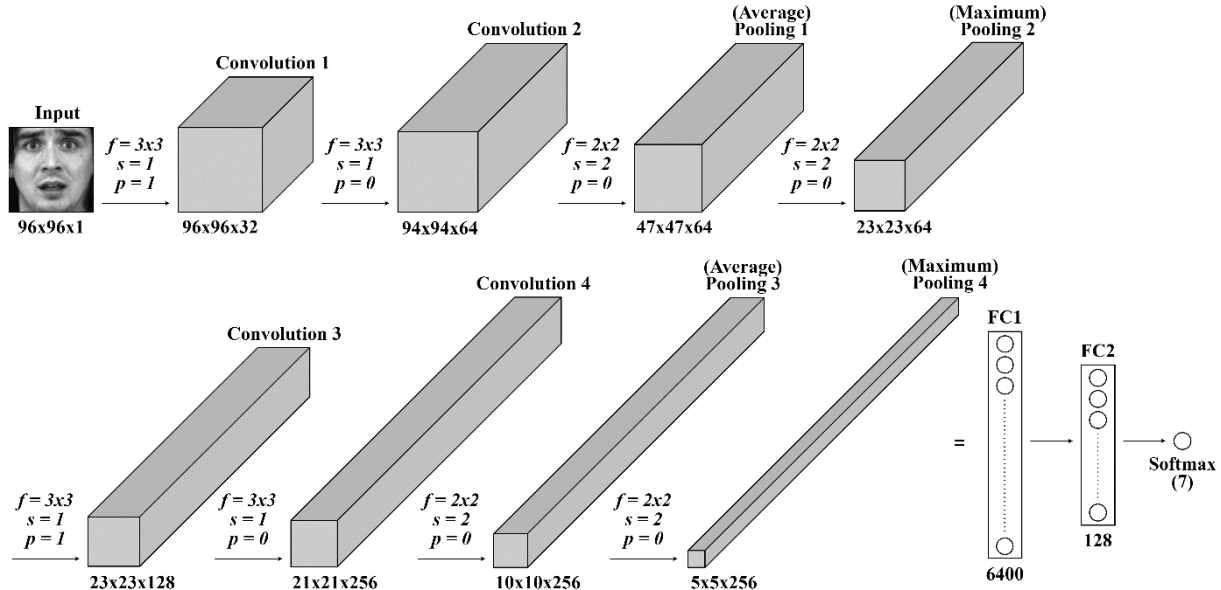
Derin modellerin sağlam ve soyut öznitelikler çıkarabildiği iyi bilinmektedir (Xu ve ark., 2015). Son yıllarda bilgisayarlı görüde en yaygın kullanılan derin öğrenme modellerinden biri ESA'dır (Altan, 2019). Derin öğrenme ve bilgisayarlı görüde temel oluşturan ESA, 1998 yılında Yann LeCun ve arkadaşları tarafından gerçekleştirilmiştir. Yann LeCun ve arkadaşları çok katmanlı ileri beslemeli yapay sinir ağlarının bir alt türü olan ESA kullanarak oluşturdukları LeNet-5 modeli ile rakamları sınıflandırmışlardır (LeCun ve ark., 1998). 1998 yılından sonra 2012 yılına kadar ESA konusunda fazla bir ilerleme gerçekleşmemiştir. ESA'nın temelini oluşturan LeNet-5 modelinden sonra bilgisayarların kapasite olarak gelişmesinden dolayı temel bileşenleri birbirine benzer olan AlexNet (Krizhevsky ve ark., 2012), VGGNet (Simonyan ve Zisserman, 2014), GoogLeNet (Szegedy ve ark., 2015) ve ResNet (He ve ark., 2016) modelleri geliştirilmiştir.

2012 yılında AlexNet'in başarısından bu yana ESA'larda kayda değer gelişmeler olmuştur (Gu ve ark., 2018). Böylece bilgisayarlı görü alanında ESA sayesinde derin öğrenmenin önemli katkı sağladığı yeni modeller geliştirilmiştir. Çalışmamızda VGGNet modelinin "2 evrişim 1 ortaklama" mantığı temel alınarak yeni bir model geliştirilmiştir. Yeni geliştirilen ESA modeli Şekil 5'te gösterilmiştir.

Şekil 5'te gösterilen modelde 4 Evrişim (Convolution), iki Ortalama (Average) ve iki Maksimum (Maximum) olmak üzere 4 Ortaklama (Pooling) katmanı bulunmaktadır. 3 Unutma (Droupout), 5 Doğrusallaştırılmış Linear Birim (ReLu-Rectified Linear Unit), 2 Tam Bağlantı (Fully Connected) ve 1 Çıkış (Softmax) katmanı ile model toplam 19 katmandan oluşmaktadır. Modeldeki toplam parametre sayısı 1.208.071'dir.

Modelin yapısı incelendiğinde; gri formatta gelen giriş görüntüsüne 2 evrişim katmanı ve 2 ortaklama (1 ortalama, 1 maksimum) katmanı uygulanır. Her maksimum ortaklama katmanından sonra %50 unutma işlemi gerçekleştirilir. Unutma katmanı sayesinde ağız ezberlemesinin önüne geçilmiş olur. 2 evrişim, 2 ortaklama ve 1 unutma işlemi farklı parametrelerle (f, s, p) ve farklı filtre kanal sayılarıyla 2 defa tekrarlanır. Evrişim katmanlarındaki filtre kanal sayıları sırasıyla (32, 64) ve (128, 256) şeklindedir. Evrişim ve ortaklama katmanları tamamlandıktan sonra düzleştirme ve tam bağlantı katmanı ile nöronlar bir dizi haline getirilir. Düzleştirme işlemi sonucunda tam bağlantı (FC1) katmanı oluşturularak 6400 nöron elde edilir. Düzleştirmenin ardından tam bağlantı (FC2) yapılarak nöron sayısı 128'e düşürülür. Tam bağlantı işlemlerinden sonra %50 unutma işlemi gerçekleştirilir. En sonunda

Softmax ile 7 sınıflı bir çıkış elde edilir. Modelin evrişim katmanlarında ReLu, çıkış katmanında Softmax aktivasyon fonksiyonu kullanılmıştır.



Şekil 5. Yeni geliştirilen ESA modeli (f: filtre boyutu, s: kaydırma, p: piksel ekleme, FC: tam bağlantı)

### Kullanılan Ortam ve Kütüphaneler

ESA modeli ile duygu analizi için yapılan deneyler; Intel Core i7-5700HQ 2.7 GHz işlemci, NVIDIA Geforce GTX 960M ekran kartı ve 16GB RAM özelliklerine sahip bir dizüstü bilgisayarda gerçekleştirilmiştir. Program kodları Windows 10 işletim sistemi üzerinde Python programlama dilinin 3.6 sürümü ile Keras, NumPy, Matplotlib, PIL, Os, OpenCV, Sklearn ve Imageio kütüphaneleri kullanılarak yazılmıştır. Python kodları Anaconda programının Spyder 4.1.4 arayüzü ile derlenmiştir.

### BULGULAR VE TARTIŞMA

Veri sayısına bağlı olarak veri seti eğitim, test ve geçerleme olarak bölünebilir. Veri sayısı çok fazla olmayan veri setlerinde test ve geçerleme için daha büyük oranda ayırma yapılabilir. KDEF ve PICS veri setlerindeki veri sayıları dikkate alınarak veri seti %60 eğitim, %20 test ve %20 geçerleme veri kümesi olmak üzere üçe bölünmüştür. Ağın eğitimi sırasında eğitim ve test veri kümeleri kullanılmış, ağın eğitimi bittikten sonra modelin eğitim sırasında hiç görmediği geçerleme veri kümesiyle başarımların sonucu elde edilmiştir. Çizelge 1’de KDEF ve PICS veri setlerinin kümelerindeki görüntülerin dağılımları gösterilmiştir.

Çizelge 1. Veri setlerinin eğitim-test-geçerleme kümelerindeki görüntü sayıları

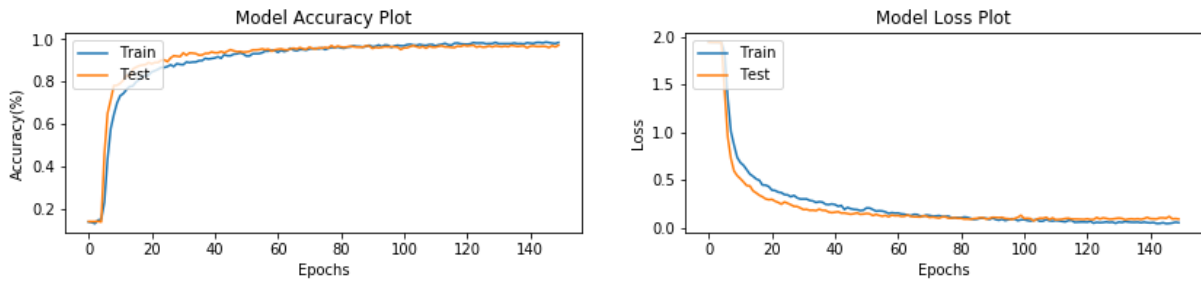
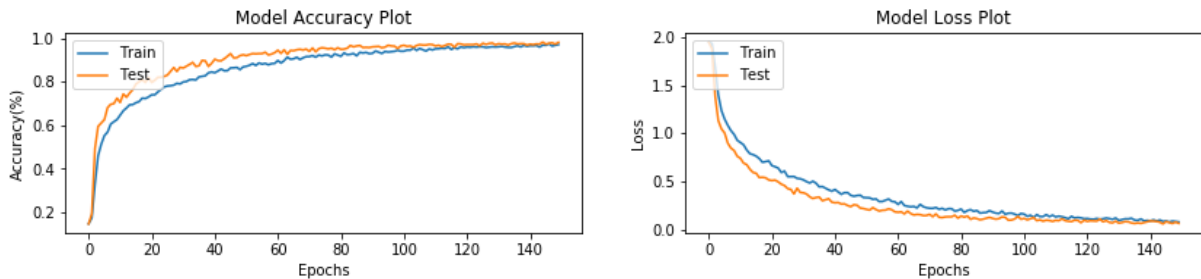
Veri Seti	Eğitim	Test	Geçerleme	Toplam
KDEF	2940	980	980	4900
PICS	2898	966	966	4830

Geliştirilen ağın eğitimi sırasında KDEF ve PICS veri setlerindeki eğitim başarımlarını karşılaştırabilmek için iki veri setinde de aynı parametreler kullanılmıştır. En yüksek doğruluk oranına ulaşmak için epoch (eğitim tur sayısı), mini batch size (her bir iterasyonda alınan küme boyutu), dropout (unutma), aktivasyon fonksiyonu ve optimizasyon algoritması parametreleri Çizelge 2’de gösterildiği gibi belirlenmiştir.

**Çizelge 2.** Eğitim parametreleri

Parametre	Değer
Epoch	150
Mini Batch Size	32
Dropout	0.5
Aktivasyon Fonksiyonu	ReLu
Optimizasyon Algoritması	Adamax

Şekil 6'da KDEF veri seti, Şekil 7'de ise PICS veri seti ile ESA modelinin eğitim (train) ve test veri kümesiyle eğitilmesi sırasında elde edilen doğruluk (accuracy) oranı ve kayıp (loss) değeri grafiksel olarak gösterilmiştir. Grafikler incelendiğinde; her iki veri setinde de ilk iterasyonlardan (epoch) itibaren ağırlık hızlı bir şekilde öğrendiği, 100 iterasyondan sonra iniş çıkışlardan da anlaşıldığı üzere ağırlık öğrenmeye devam ettiği ve 140 iterasyonun sonunda %98 civarında bir eğitimin gerçekleştiği görülmektedir.

**Şekil 6.** KDEF veri seti ile eğitilen ESA modelinin doğruluk ve kayıp grafiği**Şekil 7.** PICS veri seti ile eğitilen ESA modelinin doğruluk ve kayıp grafiği

Çizelge 3'te KDEF ve PICS veri setleri üzerinde yapılan eğitim sonucunda elde edilen başarımların oranları gösterilmiştir. Çizelge incelendiğinde; KDEF veri setinde eğitim verisinin %98.5 oranında öğrenildiği görülürken, test verisine göre öğrenme oranı %97.24 ve eğitime katılmayan geçerleme verisine göre 0.106 gibi düşük kayıpla başarımların oranı %97.44 olarak elde edilmiştir. PICS veri setinde ise eğitim verisinin %97.14 oranında öğrenildiği görülürken, test verisine göre öğrenme oranı %98.14 ve eğitime katılmayan geçerleme verisine göre 0.061 gibi daha düşük bir kayıpla başarımların oranı %98.24 olarak elde edilmiştir.

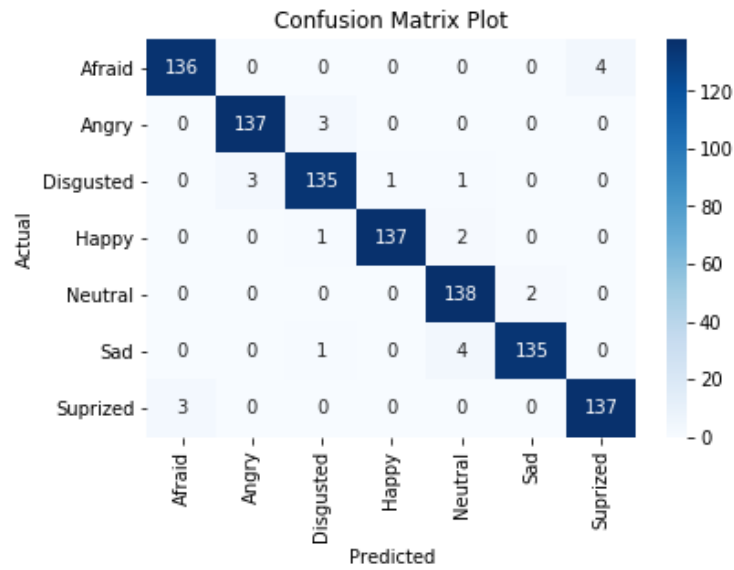
**Çizelge 3.** Veri setlerine ait başarımlar ve kayıp oranları

Veri Seti	Başarımlar/Kayıplar	Eğitim	Test	Geçerleme
KDEF	Başarımlar (%)	98.50	97.24	97.44
	Kayıplar	0.052	0.089	0.106
PICS	Başarımlar (%)	97.14	98.14	98.24
	Kayıplar	0.081	0.066	0.061

Çalışma sonucunda, hem 4900 görüntü bulunan KDEF veri setinde hem de veri artırma yöntemiyle 4830 görüntüye çıkarılan PICS veri setinde yüksek oranda öğrenme gerçekleşmiştir. İki veri seti

karşılaştırıldığında; birbirine yakın olmakla beraber, PICS veri seti KDEF veri setine göre %0.8 oranında daha yüksek sınıflandırma başarısına sahip olmuştur.

Evrişimsel Sinir Ağının hiç görmediği %20'lik geçerleme kümesinde KDEF veri seti için 980 görüntü bulunmaktadır. Dolayısıyla her bir duyguya ait 140 ar görüntü üzerinden elde edilen karışıklık matrisi (confusion matrix) Şekil 8'de gösterilmiştir. Şekilden görüldüğü üzere en başarılı sonuçlar sırasıyla nötr (%98.57), öfke, mutluluk ve şaşırma (%97.86) duygu sınıflarında, en başarısız sonuç ise iğrenme ve üzüntü (%96.43) duygu sınıflarında elde edilmiştir. En çok karıştırılan duygu sınıfları %2.86 ile korku-şaşırma ve üzüntü-nötr olmuştur. Bunu %2.14 ile öfke-iğrenme, iğrenme-öfke ve şaşırma-korku sınıfları izlemektedir. En az karıştırılan duygu sınıfları olarak; %0.71 ile iğrenme-mutluluk, iğrenme-nötr, mutluluk-iğrenme ve üzüntü-iğrenme göze çarpmaktadır.

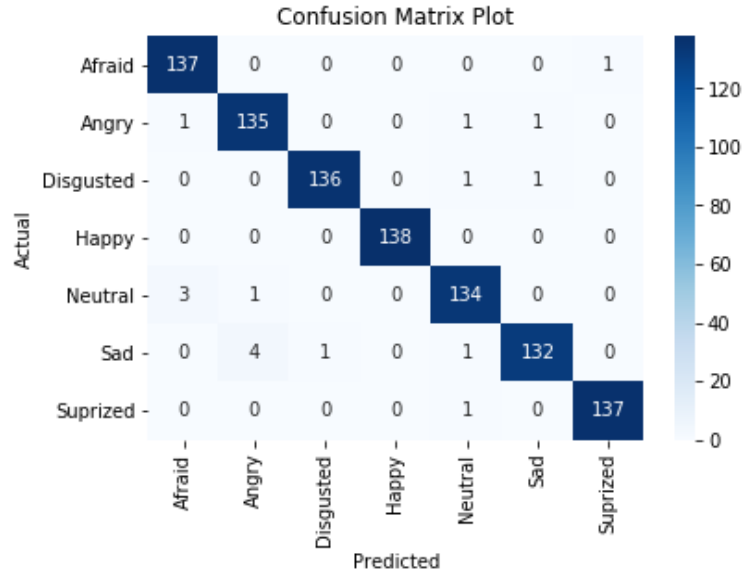


**Şekil 8.** KDEF veri setinin geçerleme kümesine ait karışıklık matrisi (afraid: korkmuş, angry: kızgın, disgusted: iğrenmiş, happy: mutlu, neutral: nötr, sad: üzgün, suprized: şaşkın)

Calvo ve Lundqvist (Calvo ve Lundqvist, 2008) tarafından KDEF veri seti üzerine yapılan çalışmada her bir duygu ifadesi için kimlik doğruluğu, hata türleri ve reaksiyon süreleri ölçümleri yapılmıştır. Elde ettikleri sonuçlara göre en çok karıştırılan duygu sınıfları olarak sırasıyla; şaşırma-korku, korku-şaşırma, iğrenme-öfke, üzüntü-iğrenme, öfke-iğrenme ve üzüntü-korku sınıflarını vermişlerdir. Ayrıca mutluluk ve nötr duygu sınıflarının en az karıştırılan ve en hızlı cevap alınan sınıf olduğunu da ifade etmişlerdir. Bu durum bu çalışmadan elde edilen sonuçları da doğrular niteliktedir.

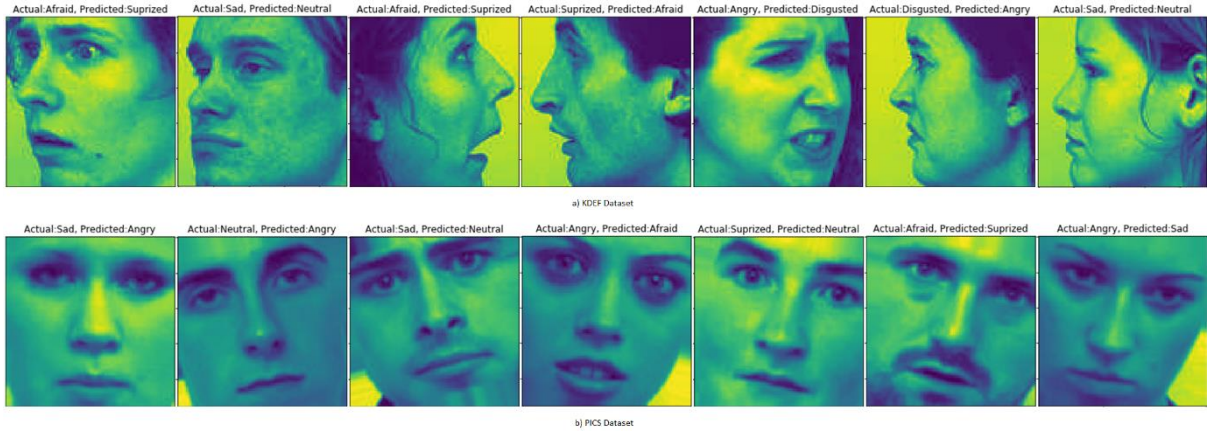
Evrişimsel Sinir Ağının hiç görmediği %20'lik geçerleme kümesinde PICS veri seti için 966 görüntü bulunmaktadır. Dolayısıyla her bir duyguya ait 138 ar görüntü üzerinden elde edilen karışıklık matrisi Şekil 9'de gösterilmiştir. Şekilden görüldüğü üzere en başarılı sonuç %100 ile mutluluk duygu sınıfında elde edilmiştir. Bunu sırasıyla %99.28 ile korku ve şaşırma duygu sınıfları izlemiştir. En başarısız sonuç ise üzüntü (%95.65) ve nötr (%97.1) duygu sınıflarında elde edilmiştir. En çok karıştırılan duygu sınıfları %2.9 ile üzüntü-öfke ve %2.17 ile nötr-korku olmuştur. Diğer korku sınıfları %0.72 gibi oldukça düşük oranda başka bir duygu sınıfı ile karıştırılmıştır.





**Şekil 9.** PICS veri setinin geçерleme kümesine ait karışıklık matrisi (afraid: korkmuş, angry: kızgın, disgusted: iğrenmiş, happy: mutlu, neutral: nötr, sad: üzgün, suprized: şaşkın)

Çalışma sonucunda hatalı duygu sınıflandırması yapılan resimlere ait örnekler Şekil 10'da verilmiştir. KDEF veri setinde 5 yönden çekilen resimler arasından önden çekilmiş olan resimlerde hiç hata yapılmadığı göze çarpmaktadır. PICS veri setinde tüm resimler tek yönden çekilmiş, orijinal resimlerde hata oranı çok düşük olup hatalı tanınan resimler çoğunlukla artırılmış resimlerdir.



**Şekil 10.** KDEF ve PICS veri setlerine ait hatalı sınıflandırma yapılan resimlere örnekler (actual: gerçek, predicted: tahmin)

Önerilen modelin aynı veri setlerini kullanarak duygu analizi yapılan diğer çalışmalarla karşılaştırılması Çizelge 4'te gösterilmiştir. Çizelge incelendiğinde, Xu ve ark. (2015) dört farklı verisetinden oluşturdukları verisetini derin evrişimli ağlardan (ConvNets) transfer özelliklerine dayanan bir model ile eğiterek %81.5, Ruiz-Garcia ve ark. (2016) ve Koç ve Özdemir (2019) ise çalışmamızda kullandığımız KDEF verisetini önerdikleri evrişimli sinir ağlarıyla eğiterek sırasıyla %96.93 ve %96.46 başarı oranı elde etmişlerdir. Çalışmamızda önerilen evrişimli sinir ağı modelinin hem evrişimli sinir ağı kullanan hem de farklı sınıflandırıcı kullanan diğer çalışmalara göre daha iyi sonuçlar verdiği görülmektedir.

**Çizelge 4.** Önerilen çalışma ile aynı veri setlerini kullanan çalışmaların karşılaştırılması.

Yazar	Veri Seti	Sınıflandırıcı	Başarım Oranı (%)
(Xu ve ark., 2015)	KDEF, PICS, CK+, JAFFE	CNN “ConvNets”	81.50
(Ruiz-Garcia ve ark., 2016)		CNN	96.93
(Eng ve ark., 2019)	KDEF	SVM	80.95
(Koç ve Özdemir, 2019)		CNN “ResNet101”	96.46
(Ahmad ve ark., 2013)		Haar	90.65
(Irtija ve ark., 2018)	PICS	SVM	84.00
(Lu ve ark., 2019)		SVM	97.33
<b>Önerilen Model</b>	<b>KDEF PICS</b>	<b>CNN</b>	<b>97.44 98.24</b>

## SONUÇ

Bu çalışmada, yeni geliştirilen bir Evrişimsel Sinir Ağı (ESA) modeli ile insan yüz ifadesinden duygu tanıma yönelik bir uygulama yapılmıştır. Duygusal ifadelerin klasik yöntemlere göre daha başarılı bir şekilde sınıflandırılması amacıyla derin öğrenme yöntemlerinden ESA kullanılmıştır. Geliştirilen modeli test etmek için iki farklı veri setindeki yedi farklı nesne sınıfına ait görüntüler kullanılmıştır. Derin öğrenme yöntemlerinin en önemli avantajı sınıflandırmada kullanılacak özniteliklerin otomatik olarak çıkarılmasıdır. Bu avantajın yanında yüksek işlem gücüne sahip donanım ihtiyacı vardır. Çalışmada kullanılan donanım alt yapısı kısıtlarından dolayı veri setlerindeki görüntü boyutları küçültülmüştür. Ayrıca veri setlerinden PICS veri setindeki görüntü yetersizliğinden dolayı veri artırma yöntemi ile görüntü sayısı artırılmıştır. Veri setlerindeki eğitim ve test veri kümelerine göre model eğitildikten sonra, ağın hiç görmediği geçerleme veri kümesine göre başarı oranları tespit edilmiştir. Başarım oranları modelin doğruluk değerlerine göre elde edilmiştir.

Test sonuçlarına göre; KDEF veri setinin geçerleme veri kümesindeki 980 görüntüde duygu tanıma başarıları %97.44, PICS veri setinin geçerleme veri kümesindeki 966 görüntüde duygu tanıma başarıları %98.24 olarak belirlenmiştir. Modelin performansını test etmek için kullanılan iki farklı veri setinde de yüksek başarı elde edilmiştir. En yüksek başarı %98.24 doğruluk oranı ile PICS veri setinden elde edilmiştir.

Sonraki çalışmalarda modelin daha fazla duygusal ifade içeren daha büyük veri setleriyle veya duygusal ifade içermeyen farklı veri setleriyle eğitilmesi faydalı olacaktır. Donanım alt yapısı kısıtlarından dolayı bu çalışmadaki sınırlılıkların giderilerek yeni denemeler yapıp modelin geliştirilmesi düşünülmektedir. Ayrıca veri setindeki görüntülerin ön işlenmesinden kaynaklı oluşabilecek kayıpların önüne geçmek amacıyla ve işlem gücünü arttırmak maksadıyla önerilen sistemin paralelleştirilmesi düşünülebilir.

## Çıkar Çatışması

Makale yazarları, aralarında herhangi bir çıkar çatışması olmadığını beyan ederler.

## Yazar Katkısı

Yazarlar makaleye eşit oranda katkı sağlamış olduklarını beyan eder.

**KAYNAKLAR**

- Ahmad F, Najam A, Ahmed Z, 2013. Image-based Face Detection and Recognition: “State of the Art.”. ArXiv Preprint ArXiv: 1302.6379.
- Altan G, 2019. DeepGraphNet: Grafiklerin Sınıflandırılmasında Derin Öğrenme Modelleri. *European Journal of Science and Technology*, October: 319–329. <https://doi.org/10.31590/ejosat.638256>
- Calvo MG, Lundqvist D, 2008. Facial expressions of emotion (KDEF): Identification under different display-duration conditions. *Behavior Research Methods*, 40(1): 109–115. <https://doi.org/10.3758/BRM.40.1.109>
- Chen M, Lu Z, Jan PA, 2015. Learning deep features for image emotion classification. 2015 IEEE International Conference on Image Processing (ICIP), 4491–4495.
- Dandil E, Özdemir R, 2019. Real-time Facial Emotion Classification Using Deep Learning. *DATA SCIENCE AND APPLICATIONS*, 2(1): 13–17.
- Eng SK, Ali H, Cheah AY, Chong YF, 2019. Facial expression recognition in JAFFE and KDEF Datasets using histogram of oriented gradients and support vector machine. *IOP Conference Series: Materials Science and Engineering*, 705(1). <https://doi.org/10.1088/1757-899X/705/1/012031>
- Fayek HM, Lech M, Cavedon L, 2017. Evaluating deep learning architectures for Speech Emotion Recognition. *Neural Networks*, 92: 60–68. <https://doi.org/10.1016/j.neunet.2017.02.013>
- Gu J, Wang Z, Kuen J, Ma L, Shahroudy A, Shuai B, Liu T, Wang X, Wang G, Cai J, Chen T, 2018. Recent advances in convolutional neural networks. *Pattern Recognition*, 77: 354–377. <https://doi.org/10.1016/j.patcog.2017.10.013>
- Hancock P, 2008. Psychological image collection at stirling (pics). Web Address: <Http://Pics.Psych.Stir.Ac.Uk>.
- He K, Zhang X, Ren S, Sun J, 2016. Deep Residual Learning for Image Recognition Kaiming. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. <https://doi.org/10.1002/chin.200650130>
- Hossain MS, Muhammad G, 2019. Emotion recognition using deep learning approach from audio–visual emotional big data. *Information Fusion*, 49: 69–78. <https://doi.org/10.1016/j.inffus.2018.09.008>
- Huang Y, Chen F, Lv S, Wang X, 2019. Facial expression recognition: A survey. *Symmetry*, 11(10). <https://doi.org/10.3390/sym11101189>
- Irtija N, Sami M, Ahad MAR, 2018. Fatigue detection using facial landmarks. *Joint Conference ISASE-MAICS 2018 - 4th International Symposium on Affective Science and Engineering 2018, and the 29th Modern Artificial Intelligence and Cognitive Science Conference*, 1–6. <https://doi.org/10.5057/isase.2018-c000041>
- Koç M, Özdemir R, 2019. Yeni Bir Veri Kümesi (RidNet) Kullanarak Kontrolsüz Ortamda Yüz İfadesi Tanımının Derin Öğrenme Yöntemleri ile İyileştirilmesi. *Bilecik Şeyh Edebali Üniversitesi Fen Bilimleri Dergisi*, 6(2): 384–396. <https://doi.org/10.35193/bseufbd.645138>
- Krizhevsky A, Sutskever I, Hinton GE, 2012. ImageNet Classification with Deep Convolutional Neural Networks. *Advances in Neural Information Processing Systems*, 1097–1105. <https://doi.org/10.1201/9781420010749>
- Lecun Y, Bengio Y, Hinton G, 2015. Deep learning. *Nature*, 521(7553): 436–444. <https://doi.org/10.1038/nature14539>
- LeCun Y, Bottou L, Bengio Y, Haffner P, 1998. Gradient-based Learning Applied to Document Recognition. *Proceedings of the IEEE*, 86(11): 2278–2323. <https://doi.org/10.1109/5.726791>
- Li J, Jin K, Zhou D, Kubota N, Ju Z, 2020. Attention mechanism-based CNN for facial expression recognition. *Neurocomputing*, 411: 340–350. <https://doi.org/10.1016/j.neucom.2020.06.014>
- Li S, Deng W, 2020. Deep Facial Expression Recognition: A Survey. *IEEE Transactions on Affective Computing*, 3045(c): 1–20. <https://doi.org/10.1109/TAFFC.2020.2981446>
- Lu Y, Wang S, Zhao W, 2019. Facial expression recognition based on discrete separable shearlet transform and feature selection. *Algorithms*, 12(1). <https://doi.org/10.3390/a12010011>
- Lundqvist D, Flykt A, Öhman A, 1998. The Karolinska directed emotional faces (KDEF). CD ROM from Department of Clinical Neuroscience, Psychology Section, Karolinska Institutet 91(630): 2–2.
- Mehta D, Siddiqui MFH, Javaid AY, 2019. Recognition of emotion intensities using machine learning algorithms: A comparative study. *Sensors (Switzerland)*, 19(8): 1–24. <https://doi.org/10.3390/s19081897>
- Özdemir, D, Karaman, S, 2017. Investigating interactions between students with mild mental retardation and humanoid robot in terms of feedback types. *Eğitim ve Bilim*, 42(191).
- Ruiz-Garcia A, Elshaw M, Altahhan A, Palade V, 2016. Deep learning for emotion recognition in faces. *International Conference on Artificial Neural Networks Springer*, 38–46. [https://doi.org/10.1007/978-3-319-44781-0\\_5](https://doi.org/10.1007/978-3-319-44781-0_5)
- Shorten C, Khoshgoftaar TM, 2019. A survey on Image Data Augmentation for Deep Learning. *Journal of Big Data*, 6(1). <https://doi.org/10.1186/s40537-019-0197-0>
- Simonyan K, Zisserman A, 2014. Very deep convolutional networks for large-scale image recognition. ArXiv Preprint ArXiv: 1409.1556.
- Szegedy C, Liu W, Jia Y, Sermanet P, Reed S, Anguelov D, Erhan D, Vanhoucke V, Rabinovich A, 2015. Going Deeper with Convolutions. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. <https://doi.org/10.1002/jctb.4820>
- Xu M, Cheng W, Zhao Q, Ma L, Xu F, 2015. Facial expression recognition based on transfer learning from deep convolutional networks. 2015 11th International Conference on Natural Computation (ICNC). *IEEE*, 702–708. <https://doi.org/10.1109/ICNC.2015.7378076>