

Sınıflandırma Algoritmalarına Dayalı VGG-11 ile Yüzde Duygu Tanıma

Facial Emotion Recognition With VGG-11 Based On Classification Algorithms

Kenan Donuk^{*1} , Davut Hanbay² 

¹Bilgisayar Programcılığı Bölümü, Şırnak Üniversitesi, Cizre Meslek Yüksekokulu, Şırnak, Türkiye

²Bilgisayar Mühendisliği Bölümü, İnönü Üniversitesi, Malatya, Türkiye
(kenandonuk@sirnak.edu.tr, davut.hanbay@inonu.edu.tr)

Received: Sep.3, 2021

Accepted: Sep.16, 2021

Published: Oct.20, 2021

Özetçe— Yüz duygu ifadeleri insanların birbirleriyle olan iletişiminde sözlü olmayan bir iletişim aracıdır. Bu ifadeler insanların düşünceleri hakkında bilgiler vermektedir. Bu bilgiler ışığında müşteri memnuniyeti, zihinsel bozuklukların tespiti, otizm, yalan ve korku tespiti gibi birçok alanda çalışmalar yapılmaktadır. Duygu tanıma görevi için yapılmış geleneksel ve derin öğrenme tabanlı birçok yöntem mevcuttur. Yapılan çalışmada duygu tanıma için FER2013 veriseti ve derin öğrenme tabanlı mimarilerden VGG-11 kullanılmıştır. VGG-11 mimarisi ile %68.32'lik test doğruluğu elde edilmiştir. Çalışmada VGG-11 mimarisinin özellik katmanına uygulanan sınıflandırma yöntemlerinin duygu tanıma doğruluğuna etkisi incelenmiştir.

Anahtar Kelimeler : Vgg-11, yüz duygu tanıma, fer2013.

Abstract— Facial emotional expressions are a non-verbal communication tool in people's communication with each other. These expressions give information about people's thoughts. In the light of this information, studies are carried out in many areas such as customer satisfaction, detection of mental disorders, autism, detection of lies and fear. There are many traditional and deep learning-based methods for emotion recognition task. In the study, FER2013 dataset and VGG-11, one of the deep learning-based architectures, were used for emotion recognition. Test accuracy of 68.32% was achieved with the VGG-11 architecture. In this article, the effect of classification methods applied to the feature layer of the VGG-11 architecture on emotion recognition accuracy is examined.

Keywords : Vgg-11, facial emotion recognition, fer2013.

1. Giriş

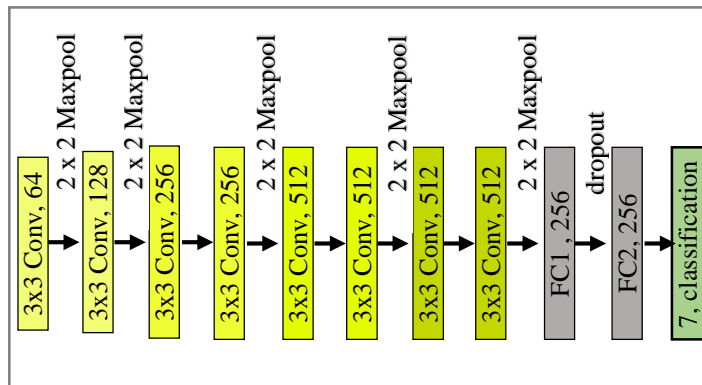
Derin öğrenme tabanlı yöntemlerin gelişmesiyle beraber birçok alanda olduğu gibi yüz ifadesi tanıma alanında da ilerlemeler kaydedilmiştir. Yüz Duygu Tanıma (YDT) sistemleri, yüzdeki ifadeden yola çıkarak insan duygularını yedi temel duygudan biri olarak sınıflandırmayı amaçlamaktadır (Mahmoudi ve diğ., 2020). YDT insanların düşünce dünyası hakkında bizlere önemli bilgiler vermektedir. Bu bilgiler ışında birçok farklı alanda çalışmalar mevcuttur (Bouzakraoui ve diğ., 2020; Mukhopadhyay ve diğ., 2020; ShanthaShalini ve diğ., 2021; Gao ve diğ., 2021; Simcock ve diğ., 2020). Yüz ifadelerinin tanınması için HOG, SIFT, LBP gibi geleneksel ve derin öğrenme tabanlı birçok farklı yöntem kullanılmıştır. Bunlardan bazıları şöyle sıralanabilir. Zahara ve ark. evrimsel sinir ağlarını (ESA) kullanarak yüz duygu ifadeleri veri seti olan FER2013 ile yüz ifadeleri tahmininde %65.97'lik bir

başarım elde etmişlerdir (Zahara ve diğ., 2020). Khairuddin ve ark. FER2013 veri setini kullandıkları çalışmalarında duygu tahmini için VGGNet mimarisini kullanmışlardır. Mimarinin hiperparametreleri üzerinde yapılan ince değişiklikler sonucunda %78.28'lik bir başarı elde etmişlerdir (Khairuddin ve Chen, 2021). Shafira ve ark. yaptıkları çalışmada duygu tanıma görevi için FER2013 ve CK+ veri setlerini kullanmışlardır. Veri setlerindeki duygu ifadelerinden öznitelik çıkarımı için Histogram of Oriented Gradient (HOG) ve Local Binary Pattern (LBP) yöntemlerini ve sınıflandırma için Extreme Learning Machine (ELM) sınıflandırıcısını kullanmışlardır. HOG yönteminin FER2013 veri setindeki doğruluğu %63.86 iken LBP yönteminin FER2013 veri seti üzerindeki doğruluğu %55.11'dir (Shafira ve diğ., 2019). Daihong ve ark. dikkat mekanizmasına dayalı bir yüz ifadesi tanıma modeli önermişlerdir. Bu modelde evrimsel sinir ağlarındaki yerel özniteliklerin kullanılmasıyla ortaya çıkan düşük verimliliğinin üstesinden gelmeye odaklanılmıştır. Önerdikleri model öz dikkat ve kanal dikkati mekanizmalarından oluşmaktadır. Model, CK+ ve FER2013 veri setleri üzerinde test edilmiştir. CK+ veri setinde %97.89 ve FER2013 veri setinde %74.15'lik bir performans elde etmişlerdir (Daihong ve diğ., 2021). Dong ve ark. yaptıkları çalışmada ESA tabanlı mimarilerin özelliklerinin birleştirilmesi ile elde edilen doğruluğun tek ESA tabanlı mimari ile elde edilecek doğruluktan daha yüksek olduğunu ortaya koymuşlardır. ESA tabanlı farklı mimarilerinin çıkış özellikleri birleştirilerek Destek Vektör Makineleri (DVM) ile sınıflandırma gerçekleştirilmiştir. Önerilen yöntemin FER2013 veri seti üzerindeki doğruluğu %73.78 olarak hesaplanmıştır (Dong, Van ve Bao 2021). Mollahosseini ve ark. Inception katmanı yapısını içeren bir ESA mimarisi önermişlerdir. Bu mimariyle yerel özellik performansında önemli bir başarı elde etmişlerdir. Mimarinin başarımı MultiPIE, MMI, CK +, DISFA, FERA, SFEW ve FER2013 veri setleri üzerinde denenmiştir. Önerilen mimarinin FER2013 veri seti üzerindeki başarısı %66.40 olarak ölçülmüştür (Mollahosseini, Chan ve Mahoor, 2016).

2. Materyal ve Yöntem

2.1. Vgg-11 Mimarisi:

Evrimsel Sinir Ağı (ESA) mimarilerinden biri olan Visual Geometry Group (VGG), 2014 yılında ImageNet Büyük Ölçekli Görsel Tanıma Yarışması (ILSVRC)'nda GoogLeNet'ten sonra en düşük hata oranıyla 2. olarak başarısını kanıtlamıştır (Simonyan ve Zisserman, 2014). VGG mimarisi 224x224x3 giriş görüntüsüne sahiptir. Mimari genel olarak evrimsel, havuzlama ve tam bağlantı katmanlarından oluşmaktadır. Evrimsel katmanlar giderek derinleşen bir yapıya sahiptir. Derinleşen bu yapıda 3x3'lük filtreler ve 2x2'lik havuzlama katmanları ile parametre miktarı azaltılır. Mimari 3 adet tam bağlantı katmanına sahiptir. Son katman 1000 sınıflı softmax katmanıdır. Mimarinin büyük başarı sağladığı VGG-16, VGG-19 modellerinin yanında VGG-11, VGG-13 gibi farklı modellerde literatürde kullanılmaktadır (Gan, Yang ve Lai, 2019; Wu, 2020). Bu çalışmada VGG mimarisinin VGG-11 modeli kullanılmıştır. Bu modelde veri setine uygun olarak mimarinin girdi verisi 48x48 olarak değiştirilmiştir. Yine tam bağlantı katmanlarındaki 4096 nöron, 256 nöron olarak değiştirilmiştir. Aynı şekilde sınıflandırma katmanı 1000 sınıf iken probleme uygun olarak 7 sınıf olarak değiştirilmiştir. Tam bağlantı katmanları arasına aşırı uyumu engellemek için bırakma (dropout) uygulanmıştır. Yapılan çalışmadaki probleme uygun olarak değiştirilmiş VGG-11 modeli Şekil 1'de gösterilmiştir.



Şekil 1. VGG-11 modeli

2.2. FER2013 Veri Seti:

VGG-11 modelinde duygu tanıma görevi için FER2013 veri seti kullanılmıştır. FER2013 duygu tanıma çalışmaları için oluşturulmuş halka açık bir veri setidir (Goodfellow ve diğ., 2013). Veri seti 28709 adet eğitim, 3589 doğrulama ve 3589 test seti olmak üzere 35887 adet duygu ifadesi görüntüsüne sahiptir. Görüntüler 48x48 uzamsal boyutlarda gri görüntülerdir. FER2013 veri setine ait bazı görüntüler Şekil 2’de gösterilmiştir.



Şekil 2. FER2013 veri seti görüntüleri

2.3. Sınıflandırma Algoritmaları:

Sınıflandırma, bir dizi etiketlenmiş verinin eğitilmesiyle farklı etiketlere sahip veriler arasındaki örüntüyü öğrenmesiyle yeni veriler üzerinde etiket tahmini yapma işlemidir. Literatürde birçok farklı sınıflandırma algoritmaları mevcuttur. Çalışmada kullanılan bazı sınıflandırma algoritmaları aşağıda verilmiştir.

2.3.1. Random Forest (RF):

RF algoritması 2001 yılında Breiman tarafından önerilmiştir (Breiman, 2001). Sınıflandırma algoritmaları arasında önemli bir yere sahip olan RF sınıflandırma ve regresyon problemlerinde kullanılabilir. Algoritmanın eğitim aşamasında bireysel tahmin edici olarak çok sayıda karar ağacı oluşturulur. Sınıflandırma tahmini karar ağaçlarının çoğunluğu tarafından belirlenir (Siji George ve Sumathi, 2021; Ooka ve diğ., 2021).

2.3.2. Support Vector Machine (SVM):

SVM, sınıflandırma problemlerinde kullanılan bir makine öğrenmesi yöntemidir. Sınıflandırılacak özellikler bir öznelik uzayına yerleştirilir. Öznelikleri sınıflara ayıracak bir karar yüzeyi oluşturulur (Cortes ve Vapnik, 1995). Sınıflar arası karar yüzeyi marjının üst düzeyde olması SVM'nin ana amaçlarından biridir. Marjın yüksek olması ile SVM yüksek genelleme yeteneği ortaya koyar (Cervantes ve diğ., 2020). SVM sınıflandırma problemlerinin yanında regresyon problemlerinde de kullanılmaktadır (Saigal ve Khanna, 2020).

2.3.3. K-Nearest Neighbors (KNN):

Bu algoritma 1967 yılında Cover ve Hart tarafından önerilmiş bir sınıflandırma algoritmasıdır (Cover ve Hart, 1967). KNN’de veriler alt gruplara ayrılır. Sınıflandırılmamış veriler önceden sınıflandırılmış verilere olan benzerliğine göre sınıflandırma gerçekleştirilir (Taunk ve diğ., 2019). Bu sınıflandırma K sayıda yakın komşuluğa bakılarak yapılır.

2.3.4. Decision Tree (DT):

DT, Quinlan tarafından 1985 yılında sunulan denetimli bir sınıflandırma tekniğidir (Quinlan, 1986). Problemin alt problemlere tekrarlayan bir şekilde bölünmesiyle oluşturulan bir karar ağacı tekniğidir

(Bastos ve diğ., 2020). Karar ağacı, veriler üzerinde sınıflandırma kurallarını belirlemek için oluşturulur (Liu ve diğ., 2020).

3. Sonuçlar ve Tartışma

3.1. VGG-11 Modeli Eğitimi ve Sınıflandırma:

VGG-11 modeli yüz ifadesi tanıma veri seti FER2013 ile eğitilmiştir. Eğitim aşaması %93.79, test aşaması ise %68.32'luk bir doğrulukla sonuçlanmıştır. Modelin eğitim verileri ve test sonuçlarının karışıklık matrisi sırasıyla Tablo 1 ve Şekil 3'te gösterilmiştir.

Tablo 1. VGG-11 Eğitim verileri

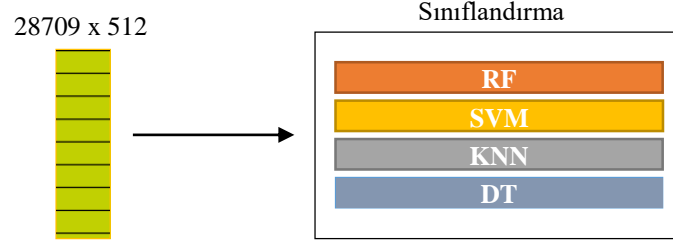
Parametreler	Değerler
Veri seti	28.709
Eğitim verisi	% 80
Doğrulama verisi	% 10
Test verisi	% 10
Giriş verileri	48 x 48 gri görüntü
Kayıp fonksiyonu	Binary crossentropy
Optimizasyon	Adam (lr=0.0001)
Parti boyutu	16
Eğitim döngü sayısı	17
Doğruluk (Eğitim)%	93.79
Doğruluk (Test)%	68.32



Şekil 3. Test verileri karışıklık matrisi

VGG-11 modelinin karmaşıklık matrisi incelendiğinde en yüksek doğruluk mutluluk (happy) verisinde ortaya çıkmıştır. Bu durumun sebebi mutluluk yüz ifadesinin belirgin hatlara ve yüksek veriye sahip olmasına bağlanabilir. Matriste en düşük sınıflandırma üzgün (sad) verisinde hesaplanmıştır. Bu durum da üzgün verisinin yüz ifadesi olarak tarafsız (neutral) verisine yakın yüz hatlarına sahip olmasına bağlanabilir. Modelde en yüksek hata (%20) iğrenme (disgust) ile kızgın (angry) verisi arasında yaptığı sınıflandırmada ölçülmüştür. Bu aşamadan sonra sınıflandırma doğruluğunu arttırmak için eğitilen

modelin son evrişim katmanındaki özelliklere RF, SVM, KNN ve DT sınıflandırma algoritmaları uygulanarak doğruluk artırılmaya çalışılmıştır. Bu durum Şekil 4'te gösterilmiştir.



Şekil 4. Öznitelikler ve sınıflandırma algoritmaları

VGG-11 modelinden elde edilen özniteliklere, python sklearn kütüphanesi fonksiyonları kullanılarak, RF, SVM, KNN ve DT sınıflandırma algoritmalarının doğrulukları ölçülmüştür (Pedregosa ve diğ., 2011). Sınıflandırma doğruluklarına ait sonuçlar ve çalışma zamanı performansları Tablo 2'de verilmiştir.

Tablo 2. Sınıflandırma algoritmalarının doğruluğa etkileri

Mimari	Doğruluk	Yöntem	Fark %	Doğruluk	Süre
VGG-11	68.32%	+ RF	1.67	69.99%	$8 \cdot 10^{-2}$ ms
		+ SVM	1.14	69.46%	3.4 ms
		+ KNN	0,33	68.65%	15.6 ms
		+ DT	-1,34	66.98%	$2 \cdot 10^{-3}$ ms

Tablo 2 incelendiğinde VGG-11+RF birleşiminin VGG-11 modelinden %1.67 'lik bir artışla sınıflandırma yaptığı görülmektedir. Yine aynı şekilde RF çalışma zamanı performansı olarak ta diğer algoritmalarla göre 2. sırada yer almaktadır. VGG-11+SVM birleşiminde ise doğruluk artışı %1.14 olarak ölçülmüş ve en iyi 2. doğruluk performansı elde edilmiştir. Fakat çalışma zamanı performansı olarak 3. Sırada yer almaktadır. Çalışma zamanı performansında en iyi ölçüm DT sınıflandırma algoritmasında tespit edilmiştir. Fakat DT sınıflandırma algoritması doğruluk olarak %1.34'lük gerileme ile VGG-11 modelinin de altında ölçülmüştür.

4. Öneriler

Yapılan bu çalışmada ESA tabanlı VGGNet mimarisinden alınan derin özellikler sınıflandırma algoritmaları ile sınıflandırılmaya çalışılmıştır. Sınıflandırma sonunda bazı algoritmaların kullanımı ile yüz ifadesi tanıma doğruluğu arttırılmıştır. Sınıflandırma doğruluğu VGG, ResNet, Xception, Inception gibi mimarilerin farklı versiyonlarının farklı katmanlardaki özellikler kullanılarak arttırılabilir.

Teşekkür

Bu çalışma İnönü Üniversitesi Bilimsel Araştırma Projeleri Koordinasyon Birimi (BAP) tarafından FDK-2020-2110 kodlu proje ile desteklenmiştir.

Kaynaklar

Mahmoudi MA, Chetouani A, Boufera F, Tabia H. (2020) Improved Bilinear Model for Facial Expression Recognition. Communications in Computer and Information Science, 1322 CCIS, pp.47-59.

Bouzakraoui MS, Sadiq A, Alaoui AY (2020) Customer satisfaction recognition based on facial expression and machine learning techniques. *Advances in Science, Technology and Engineering Systems* 5(4):594–599.

- Mukhopadhyay M, Pal S, Nayyar A, Pramanik PKD, Dasgupta N, Choudhury P. (2020) Facial Emotion Detection to Assess Learner's State of Mind in an Online Learning System. *ACM International Conference Proceeding Series*, pp.107-115.
- ShanthaShalini K, Jaichandran R, Leelavathy S, Raviraghul R, Ranjitha J, Saravanakumar N (2021) Facial Emotion Based Music Recommendation System using computer vision and machine learning techniques. *Turkish Journal of Computer and Mathematics Education (TURCOMAT)* 12(2):912–917.
- Gao Z, Zhao W, Liu S, Liu Z, Yang C, Xu Y (2021) Facial Emotion Recognition in Schizophrenia. *Frontiers in Psychiatry*.
- Simcock G, McLoughlin LT, Regt TD, Broadhouse KM, Beaudequin D, Lagopoulos J, Hermens DF (2020) Associations between Facial Emotion Recognition and Mental Health in Early Adolescence. *International Journal of Environmental Research and Public Health* 17(1).
- Zahara L, Musa P, Prasetyo Wibowo E, Karim I, Bahri Musa S. (2020) The Facial Emotion Recognition (FER-2013) Dataset for Prediction System of Micro-Expressions Face Using the Convolutional Neural Network (CNN) Algorithm based Raspberry Pi. 5th International Conference on Informatics and Computing (ICIC 2020).
- Khairuddin Y, Chen Z (2021) Facial Emotion Recognition: State of the Art Performance on FER2013.
- Shafira SS, Ulfa N, Wibawa HA, Rismiyati. (2019) Facial Expression Recognition Using Extreme Learning Machine. ICICOS 2019 - 3rd International Conference on Informatics and Computational Sciences: Accelerating Informatics and Computational Research for Smarter Society in The Era of Industry 4.0, Proceedings.
- Daihong J, Yuanzheng H, Lei D, Jin P (2021) Facial Expression Recognition Based on Attention Mechanism. *Scientific Programming*.
- Dong TN, Van L, Bao PT (2021) Facial Expression Recognition Using Multi-deep Convolutional Neural Network Encoders with Support Vector Machines. *International Journal of Machine Learning and Computing* 11(5):345–349.
- Mollahosseini A, Chan D, Mahoor MH. (2016) Going deeper in facial expression recognition using deep neural networks. 2016 IEEE Winter Conference on Applications of Computer Vision, (WACV 2016).
- Simonyan K, Zisserman A. (2014) Very Deep Convolutional Networks for Large-Scale Image Recognition. 3rd International Conference on Learning Representations, ICLR 2015 - Conference Track Proceedings.
- Gan Y, Yang J, Lai W. (2019) Video object forgery detection algorithm based on VGG-11 convolutional neural network. Proceedings - 2019 International Conference on Intelligent Computing, Automation and Systems (ICICAS 2019), pp.575–580.
- Wu M, Ma W, Li Y, Zhao X. (2020) The Optimization Method of Knowledge Distillation Based on Model Pruning. Proceedings - 2020 Chinese Automation Congress (CAC 2020), pp.1386–1390.
- Goodfellow IJ, Erhan D, Carrier PL, Courville A, Mirza M, Hamner B ve ark. (2013) Challenges in Representation Learning: A report on three machine learning contests. *Neural Networks* 64:59–63.
- Breiman L (2001) Random Forests. *Machine Learning* 2001 45(1):5-32.

Siji George CG, Sumathi B (2021) Genetic Algorithm Based Hybrid Model Of convolutional Neural Network And Random Forest Classifier For Sentiment Classification. *Turkish Journal of Computer and Mathematics Education (TURCOMAT)* 12(2):3216–3223.

Ooka T, Johno H, Nakamoto K, Yoda Y, Yokomichi H, Yamagata Z (2021) Random forest approach for determining risk prediction and predictive factors of type 2 diabetes: large-scale health check-up data in Japan. *BMJ Nutrition, Prevention & Health*.

Cortes C, Vapnik V (1995) Support-vector networks. *Machine Learning* 20(3):273–297.

Cervantes J, Garcia-Lamont F, Rodríguez-Mazahua L, Lopez A (2020) A comprehensive survey on support vector machine classification: Applications, challenges and trends. *Neurocomputing* 408:189–215.

Saigal P, Khanna V (2020) Multi-category news classification using Support Vector Machine based classifiers. *SN Applied Sciences* 2(3):1-12.

Cover TM, Hart PE (1967) Nearest Neighbor Pattern Classification. *IEEE Transactions on Information Theory* 13(1):21-27.

Taunk K, De S, Verma S, Swetapadma A. (2019) A brief review of nearest neighbor algorithm for learning and classification. 2019 International Conference on Intelligent Computing and Control Systems (ICCS 2019), pp.1255–1260.

Quinlan JR (1986) Induction of Decision Trees. *Machine Learning* 1:81-106.

Bastos NS, Marques BP, Adamatti DF, Billa CZ (2020) Analyzing EEG Signals Using Decision Trees: A Study of Modulation of Amplitude. *Computational Intelligence and Neuroscience*.

Liu S, Yang Z, Li Y, Wang S (2020) Decision tree-based sensitive information identification and encrypted transmission system. *Entropy* 22(2).

Pedregosa F, Varoquaux G, Gramfort A, Michel V, Thirion B, Grisel O ve ark. (2011) Scikit-learn: Machine Learning in Python. *Journal of Machine Learning Research* 12(85):2825-2830.