

Evrişimsel Sinir Ağı Modellerinde Özellik Seçim Yöntemlerini Kullanarak Çiçek Görüntülerinin Sınıflandırılması

Mesut TOĞAÇAR^{1*}, Burhan ERGEN², Fatih ÖZYURT³

¹ Bilgisayar Teknolojileri Programı, Teknik Bilimler Meslek Yüksekokulu, Fırat Üniversitesi, Elazığ, Türkiye

² Bilgisayar Mühendisliği, Mühendislik Fakültesi, Fırat Üniversitesi, Elazığ, Türkiye

³ Yazılım Mühendisliği, Teknoloji Fakültesi, Fırat Üniversitesi, Elazığ, Türkiye

*¹ mtogacar@firat.edu.tr, ² bergen@firat.edu.tr, ³ fatihozyurt@firat.edu.tr

(Geliş/Received: 03/06/2019;

Kabul/Accepted: 28/10/2019)

Öz: Görüntü işleme yöntem ve tekniklerinin gün geçtikçe daha iyi sonuç vermesi, ekolojik dengenin duyarlılığı açısından önem arz etmektedir. Bu makale ekolojik dengenin temel unsuru olan çiçek görüntülerinin sınıflandırılmasını ele almaktadır. Son zamanlarda çiçek görüntülerini üzerinde derin öğrenme yöntemlerinin kullanımı artmıştır. Bu çalışmada, çiçek görüntülerinin sınıflandırılması için internette erişime açık olan veri seti kullanılmıştır. Veri seti 4326 görüntüden oluşmaktadır. Elde edilen veri kümesinde özellik çıkarımı için derin öğrenme modellerinden evrişimsel sinir ağı (ESA) kullanılmıştır. ESA mimarilerinden AlexNet, VGG-16 ve VGG-19 modelleri kullanılmıştır. Üç modelinde ortak özelliği 1000 özellik veren tam bağlantılı katmana sahip olmalarıdır. Çiçek görüntülerinden elde edilen özellikler destek vektör makineleri (DVM) ile sınıflandırılmış ve elde edilen sonuçlar karşılaştırılmıştır. Karşılaştırma sonucunda en iyi sınıflandırma performansını VGG-16 mimarisi ile sağlanmıştır. Elde edilen sınıflandırma doğruluk oranı %86,56'dır. Sonraki aşamada ESA mimarilerinin son tam bağlantılı katmanından elde edilen 1000 özellik birleştirilerek 3000 özellik seti oluşturuldu. Ardından, özellik seçim yöntemlerinden; Maksimum Bilgi Katsayı (MBK), Ridge regresyonu ve Özinelemeli Özellik Eleme (ÖÖE) yöntemleri kullanılarak özellik sayısı 300'e düşürülmüştür. Özellik seçim yöntemleri ile çıkartılan en verimli özellikler DVM yöntemi ile yeniden sınıflandırılmıştır. Sınıflandırma başarı oranı yaklaşık %4,54 artarak %91,10 olmuştur. Bu çalışma ile çiçek görüntülerinin sınıflandırılmasında ESA mimarileri ile birlikte özellik seçim yöntemlerinin kullanımının etkili olduğu gözlemlenmiştir.

Anahtar kelimeler: ESA, derin öğrenme, görüntü işleme, özellik seçimi, çiçek görüntülerı.

Classification of Flower Images by Using Feature Selection Methods in Convolutional Neural Network Models

Abstract: It is important for the sensitivity of ecological balance that image processing methods and techniques give better results day by day. This article deals with the classification of flower images which are the basic elements of ecological balance. Recently the use of deep learning methods on flower images has increased. In this study, the dataset available on the internet was used for the classification of flower images. The dataset consists of 4326 images. In the obtained dataset, convolutional neural network (CNN), one of the deep learning models, was used for feature extraction. The CNN architectures AlexNet, VGG-16 and VGG-19 are used in this study. The common feature of the three models is that they have a fully connected layer that give 1000 features. The features obtained from flower images were classified with support vector machines (SVM) and the results were compared. As a result of the comparison, the best classification performance was achieved with VGG-16 architecture. Obtained classification accuracy rate is 86.56%. In the next phase, 1000 features obtained from the last fully connected layer of CNN architectures were combined and 3000 feature set were created. Then, the most efficient 300 features were selected from the 3000 set of features using the Maximum Information Coefficient (MIC), Ridge regression, and Recursive Feature Elimination (RFE) feature selection methods. The most efficient features extracted by feature selection methods were reclassified by SVM method. The classification success rate increased by approximately 4.54% to 91.10%. In this study, it has been observed that using CNN architectures and feature selection methods are effective in the classification of flower images.

Key words: CNN, deep learning, image processing, feature selection, flower images.

1. Giriş

Bitki organlarından biri olan çiçekler, ekolojik düzenin vazgeçilmez bir unsurudur. Çiçekler, insanlara faydalı olacak alanlarda da kullanılmıştır [1]. İlaç sanayisi, gıda ve kozmetik sektörü, peyzaj alanları, özel ve anlamlı günlerde vs. sürekli arz edip, talep görmektedir. Çiçeklerin şekil ve renk benzerliğinden dolayı ayırt etmek zor

* Sorumlu yazar: mtogacar@firat.edu.tr. Yazarların ORCID Numarası: ¹ 0000-0002-8264-3899, ² 0000-0003-3244-2615, ³ 0000-0002-8154-6691

bir iştir [1,2]. Günümüzde yaklaşık 390 bin çiçek çeşidi bulunmaktadır [3]. İnsan gözüyle belli bir mesafeden sonra birbirine renk ve şekil olarak benzeyen çiçekleri ayırt etmek daha da zorlaşmaktadır. Derin öğrenme ise insan gözünün yaptığı işlevleri daha hızlı ve doğru sonuçlar vermeye yarayan bir yazılım teknolojisidir [4,5]. Bu teknolojide çeşitli modeller geliştirilmiştir. Geliştirilen bu modeller içerisinde son zamanlarda popülerliğini artırmış olan ESA modelleridir. ESA modelleri, uzman bilgisine ihtiyaç duymaksızın özellikleri otomatik olarak çıkarır. Ayrıca ESA modelleri girdi olarak vektör (özellik vektörü) değil, matris (ham görüntü) kullanır [6,7]. Sonuç olarak, çiçeklerin sınıflandırılması ve segmentasyonu gibi işlemlerde derin öğrenme mimarileri gün geçtikçe daha da ilgi görmeye başlamıştır [8].

Ciçek görüntülerinin sınıflandırılması işleminde birçok teknik ve yöntem önerilmiştir. Hazem Hiary ve ark. [8], çalışmalarında çiçek görüntülerinin sınıflandırılması için iki aşamalı derin öğrenme modeli önermişlerdir. Birinci aşamada çiçek bölgesini otomatik olarak yerelleştirmek için segmentasyonu işlemini gerçekleştirmiştir. İkinci aşamada ise çoklu sınıflandırıcı Softmax'ı kullanarak %97,1 oranında bir başarı elde etmişlerdir. Musa Çubuk ve ark.[9], çalışmalarında minimum artıklık maksimum alaka düzeyi (MAMAD) yöntemini kullanarak özellik seçimi yapmışlardır. Çıkarlıklarını verimli özellikleri DVM sınıflandırıcısını kullanarak %96,39 oranında bir başarı elde etmişlerdir. Philipe A.Dias ve ark. [10], yeni ESA tabanlı bir algılama algoritması geliştirmiştir. %90 oranında bir sınıflandırma başarısı elde etmişlerdir. MostafaMehdipour Ghazi ve ark. [11], çalışmalarında çiçek görüntülerinin sınıflandırılması için önceden eğitilmiş (transfer öğrenme) ESA modellerini kullandılar. Veri setine görüntü çoğaltma tekniği kullanıp yeni görüntü kümesi elde ederek sınıflandırma işlemini gerçekleştirdiler. Çalışmalarında elde ettikleri başarı oranı %80'dir. Marco Seeland ve ark. [12], çiçeklerin sınıflandırılması işleminde geleneksel yöntemleri kullandılar. Yerel şekil ve renk tanımlayıcıları yöntemleri ile özellik seçim detektörlerini birlikte kullandılar. Sınıflandırıcı olarak kullandıkları DVM yöntemi ile %94 oranında bir başarı elde ettiler. Son olarak D.S.Guru ve ark. [13], geleneksel yöntem olan renk dokusu momentlerini, gri-seviye eş-olşum matrisi yöntemini ve Gabor teknğini birleştirerek bir sınıflandırma işlemi gerçekleştirdiler. Üç yöntemin birleşiminde yaklaşık %79 oranında bir başarı elde ettiler.

Bu çalışmada yazarlar, çiçek görüntülerinin ayırt edilebilmesi için ESA modelleri ile özellik çıkarma yöntemlerinin birlikte kullanılmasını önermektedirler. Çiçek görüntülerinin sınıflandırılmasında yaygın olarak kullanılan sınıflandırıcı DVM [14,15], bu çalışmada da kullanılmıştır.

Bu makale kısaca şöyle düzenlenmiştir: 2. bölümde, veri kümesi ve ESA modelleri hakkında kısa bilgiler verilmektedir. Bölüm 3'te özellik seçimi yöntemleri, sınıflandırıcı ve önerilen yöntem hakkında bilgi sunulmuştur. Bu çalışmada elde edilen analiz sonuçları Bölüm 4'te verilmiştir. Tartışma bölümü bölüm 5'de verilmiştir. Tartışma bölümünde aynı veri setini kullanan diğer çalışmalara degeinilmiştir. Son bölüm ise sonuç bölümünden oluşmuştur.

2. Veri Kümesi ve ESA Modelleri

2.1 Veri kümesi

Bu çalışmada, beş bitki türüne ait 4326 çiçek görüntüsünden oluşan erişime açık veri kümesi [16] kullanılmıştır. Veri kümesini oluşturan papatya, karahindiba, gül, ayçiçeği ve lale bitki türlerine ait örnek görüntüler Şekil 1'de gösterilmiştir. Veri görüntülerini Google, Flicker ve Yandex tarayıcıları üzerinden toplanmış ve erişime sunulmuştur. Çiçek sınıflarına ait görüntü sayısı Tablo 1'de gösterilmiştir. Veri kümesindeki görüntüler yüksek çözünürlükte değildir. Görüntülerin çözünürlüğü 134×240 piksel ile 1024×198 piksel arasında değişmektedir.



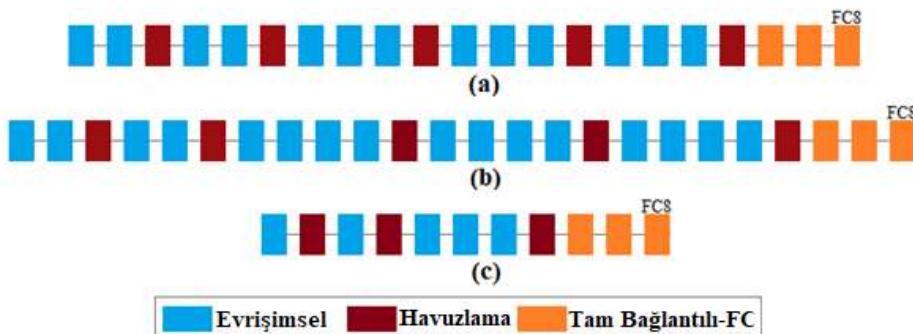
Şekil 1. Deneyde kullanılan çiçek görüntüler: (a) papatya, (b) karahindiba, (c) gül (d) ayçiçeği ve (e) lale.

Tablo 1. Veri setinde bulunan görüntülerin sınıf bazında sayısal dağılımı.

Çicek Adı	Papatya	Karahindiba	Gül	Ayçiçeği	Lale	Toplam
Görüntü Sayısı	769	1055	784	734	984	4326

2.2 ESA modelleri

Bu bölümde, AlexNet, VGG-16 ve VGG-19 modelleri hakkında bilgi verilmiştir. Üç modelin mimari yapısı birbirine benzer özellikler gösterdiği için seçilmiştir. ESA mimarilerinin şematik gösterimi Şekil 2'de gösterilmiştir. AlexNet mimarisi, 2012 yılında ImageNet yarışmasında adını duyuran, yaklaşık bir milyon görüntü ile eğitilme özelliğine sahip, görüntülerini sınıflandırmada son derece başarılı bir ESA modelidir. AlexNet mimarisinin ana katmanları evrişim, havuzlama ve tam bağlı katmanlardan oluşur. Evrişim katmanı, seçilmiş filtre boyutu ile giriş yapan görüntü üzerinde dolaştırılması işlemeye dayanır. Filtrenin boyutu 3×3 veya 5×5 piksel olabilir. Filtre ile elde edilen görüntü bir sonraki katmanın giriş görüntüsü olur. Kısaca evrişimsel katmanlardan geçen görüntünün aktivasyon haritaları oluşmaktadır. Aktivasyon haritalarının en önemli özelliği kendine has özelliklerden oluşmasıdır [17]. Son olarak havuzlama katmanı, görüntü özelliklerini korur, görüntü boyutunu ve maliyetlerini azaltır ve ayrıca görüntü bilgilerini bozulmadan tutar [18].

**Şekil 2.** ESA modellerinin şematik gösterimi: (a) VGG-16, (b) VGG-19 ve (c) AlexNet.

VGG-16 mimarisi, evrişimsel, havuz ve tamamen bağlı katmanlardan oluşur. Toplam 21 ana katmandan oluşur [19]. Bu mimari artan bir ağ yapısına sahiptir. Görüntü giriş çözünürlüğü 224×224 pikseldir. Evrişimsel katmanındaki filtre boyutu 3×3 pikseldir. Bu mimaride son katmanlar özellik çıkarımı için kullanılan tam bağlı katmanlardan oluşur.

VGG-19 mimarisinin ana katmanları 16 evrişimsel, beş adet havuzlama ve üç tam bağlı katmandan oluşur. Bu mimari toplam 24 ana katmana sahiptir. VGG-19 derinlemesine bir ağa sahip olduğundan evrişimsel katmanda kullanılan filtreler parametre sayısını azaltmak için kullanılır. Bu mimaride seçilen filtrenin boyutu 3×3 pikseldir. VGG-19 mimarisi yaklaşık 138 milyon parametre içermektedir [20].

Bu makalede kullanılan ESA mimarileri özellik çıkarımı için kullanılmıştır. Özellikler, üç modelde de FC-8 katmanından elde edilmiştir. FC-8 katmanı, 1000 tane özellik çıkarır [21]. Üç modelde filtre 3×3 , adım sayısı iki ve havuzlama türü maksimum seçilmiştir. Ayrıca bu çalışmada kullanılan veri seti tüm deneylerde %30'u test verisi, %70'i ise eğitim verisi olarak kullanıldı. Yani, çapraz doğrulama yerine geniş veri kümesi olduğu için veri kümesi test ve eğitim olarak ikiye ayrılmış sınıflandırma işlemi gerçekleştirildi. ESA mimarilerinde eğitim süreci transfer öğrenme ile gerçekleştirildi. Ek olarak, veri seti görüntülerinin ESA modellerinin giriş boyutlarına dönüştürülmesi işleminde "ImageDatastore" ve "augmentedImageDatastore" aktivasyon kodları kullanıldı.

3. Yöntemler

3.1 Özellik seçim yöntemleri

Özellik seçimi, görüntü işlemede sınıflandırma doğruluğunu artırmak ve daha hızlı sonuç vermek için kullanılır. Diğer bir ifadeyle, mevcut veri kümesinden bir veri alt kümesi oluşturma yöntemidir [22]. Bu çalışmada üç tip özellik seçim yöntemi kullanıldı.

3.1.1 MBK yöntemi

Kullanılan yöntemlerden ilki MBK yöntemidir. Bu yöntem, büyük veri kümelerindeki çift değişkenler arasında ilişki bağıntısını inceler ve MBK'nın doğruluğu, her çift için bağımlılık ölçüsü üzerinde etkilidir. Doğrusal veya doğrusal olmayan değişkenlerin eşleştirilmesi ile birlikte olabilme durumunu ölçen istatistik tabanlı bir yöntemdir [23]. Değişkenler arasındaki karşılıklı bilginin maksimum olacağı şekilde seçip yapar. Bu durum her zaman için mümkündür. Denklem 1'de X ve Y değişkenleri temsil etmektedir. Burada X ve Y değişkenlerinin boyutları kabaca aynı olmak zorundadır. $X=2$ ve $Y=2$ değerleri genel olarak kabul edilir ancak bu değerler $H(X)$ tarafından değişimlembektedir. Eğer veri kümesinde eşleştirilmiş değişkenlerin boyutları farklı ise, o kümeye elde edilmiş en büyük değere bölünerek normalizasyonu gerçekleştirilir. Böylece değişkenler arasında oluşması muhtemel düzensizlik (entropi) en aza indirgenmiş olur [24]. Bu durum Denklem 2'de gösterilmiştir. Denklem 2'nin normalizasyonu ise Denklem 3'te gerçekleşmiştir.

$$H(X_b) = H(Y_b) = H(X_b, Y_b) \quad (1)$$

$$I(X; Y) = H(X) + H(Y) - H(X, Y) \quad (2)$$

$$\log(\min(n_x, n_y)) \quad (3)$$

3.1.2 Ridge yöntemi

Ridge regresyonu, çok değişkenli neden-sonuç ilişkisine dayanan verilerin analizini gerçekleştiren istatistik tabanlı bir yöntemdir. Tercih edilmesinin sebebi en küçük kareler (EKK) yönteminin varyans değerlerinden daha düşük varyans değerleri verdiği için kullanılmaktadır. Ridge regresyon formülü Denklem 4'te gösterilmiştir. Bu denklemde k sabit bir değişken olup, $0 < k < 1$ arası değer alır. X ve Y veri kümesinde yer alan değişkenleri ifade eder. I ise birim matrisini temsil eder [25,26].

$$(XX' + kI)^{-1}X'Y \quad (4)$$

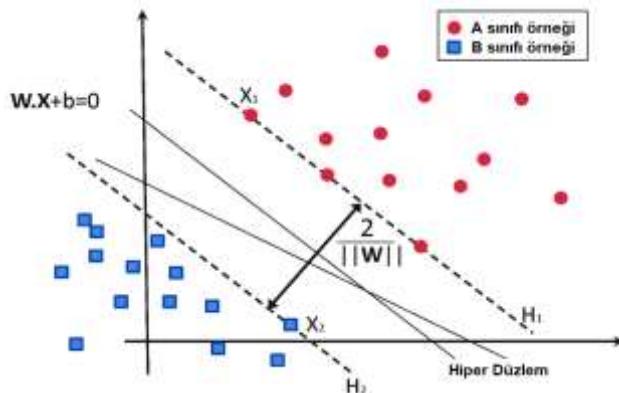
3.1.3 ÖOE yöntemi

ÖOE, özyinelemeli özellik azaltma yöntemi olarak bilinir. Bu yöntem, veri setinde elde edilen tüm özellikler ele alır ve önem sırasına göre sıralama yapar. S özellikler verisini içeren bir dizi olarak kabul edilirse, birden fazla özellik dizisini $S_1 > S_2 > S_3$ olacak şekilde önem derecesine göre sıralar. Bu yöntemde her yineleme işlemi gerçekleştirildiğinde, S_i dizisine ait üst sıralı özellikler (en verimli özellikler dizisi) korunur, iyileştirilme yapılır ve doğruluk değerlendirilir. Sonraki aşamada, en iyi doğruluk değerleri ile S_i dizisi kıyaslatılır ve en verimli S_i özelliklerinden daha iyi sonuç vermiş ise değerler modelde önem sırasına göre değiştirilir [27].

Bu makalenin çalışmasında, MBK, Ridge ve ÖOE özellik seçim algoritmalarından elde edilen özelliklerin ağırlık katsayıları çıkartılmıştır. Sonraki adımda üç özellik seçiminden elde edilen ağırlık katsayıları toplanıp aritmetik ortalaması alınmıştır. Aritmetik ortalaması yüksek olan özellikler sırası ile sınıflandırma işleminde kullanılmıştır.

3.2 DVM sınıflandırıcısı

DVM yöntemi, regresyon ve sınıflandırma problemleri için kullanılabilcek kontrollü bir makine öğrenme yöntemidir. Bu yöntem, verilerden elde edilen özelliklerini yeni bir koordinat düzlemine yerleştirir. Ardından, sınıflandırma işlemi sınıfları ayıran hiper düzlem bulmak suretiyle gerçekleştirilir. Başka bir ifadeyle DVM, pozitif veri alanını, negatif alan kümesinden maksimum alanı olan ayıran bir sınıflandırıcıdır [28]. Şekil 3'te DVM sınıflandırıcısının çalışma prensibi gösterilmiştir. Denklem 6'nın amacı oluşan sınıflandırma sürecinde oluşan problemlerin en aza indirgemesini sağlar. Denklem 7'de X_i ve Y_i özelliklerini temsil eder. Bu denklem özelliklerin hangi sınıfı ait olduğunu tahmin eder [28].



Şekil 3. DVM sınıflandırma süreci işlem diyagramı.

$$u = \vec{w} \cdot \vec{x} - b \quad (5)$$

$$\frac{1}{2} \|\vec{w}\|^2 \quad (6)$$

$$y_i(\vec{w} \cdot \vec{x}_i - b) \geq 1, \forall i \quad (7)$$

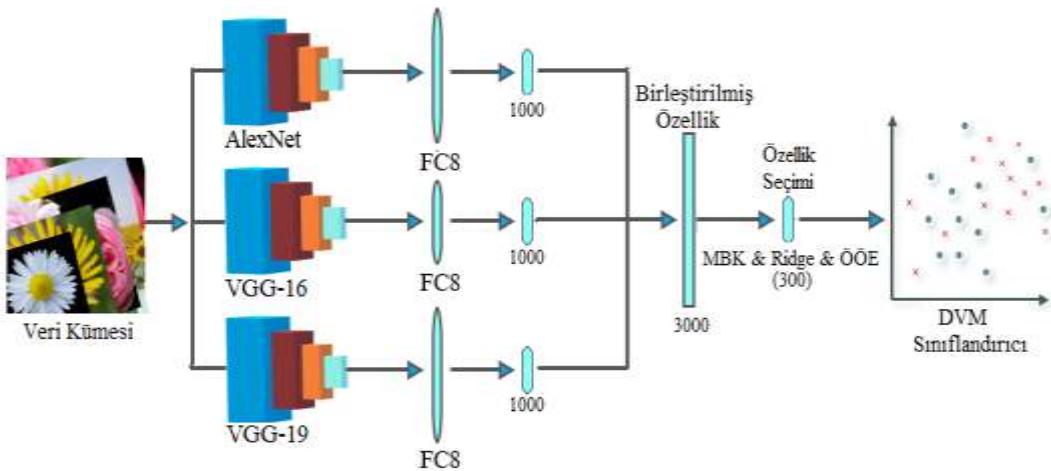
Bu çalışmada DVM sınıflandırıcı kodları ESA modellerinin mimari yapısında kullanılmıştır. Ayrıca, özellik seçim yöntemleri ile oluşturulan özellik setleri içinde Matlab yazılımının “Sınıflandırma Öğrenmesi” araç kutusu kullanılarak DVM yöntemi ile sınıflandırma gerçekleştirilmiştir. DVM parametre değerleri, makul değerler kabul edilmiş ve tercih edilen parametre bilgileri şunlardır; Kutu kısıtlama düzeyi bir tercih edilmiştir. Kernel fonksiyonu doğrusal ve kuadratik olarak seçilmiştir. Kernel skaliası otomatik kabul edilmiştir. Çok sınıflı yöntem parametresi ise bire-bir seçilmiştir.

3.3 Önerilen yöntem

Önerilen yöntemde AlexNet, VGG-16 ve VGG-19 mimarileri özellik çıkarımı için kullanılmıştır. Üç modelinde FC8 katmanına sahip olması bu modellerin birlikte kullanımında etkili olmuştur. Her modelin FC8 katmanı 1000 adet özelliğe sahiptir [20]. Önerilen yöntemde özellik çıkarımları eğitim ve test aşamalarında gerçekleştirilmiştir. Toplamda ($1000 + 1000 + 1000$) 3000 adet özellik birleştirilmiş özellik seçim yöntemleriyle (MBK, Ridge, ÖOE) 300 adet özelliğe düşürülmüştür. Bu yaklaşımda kullanılan özellik seçim yöntemleri için her bir özelliğe ait ağırlıklar hesaplanmıştır. Sonraki adımda tüm ağırlıkların aritmetik ortalaması alınarak yeni bir ağırlık özelliğe çıkartılmıştır. Bir başka deyişle, üç özellik seçiminden çıkan ağırlık değerlerinin toplamının aritmetik ortalamaları alınarak özellik seçimi gerçekleştirilmiştir. En son adımda sınıflandırma işlemi DVM ile gerçekleştirilmiştir. Önerilen yaklaşımın tasarımı Şekil 4'te gösterilmiştir. Üç modelde Bölüm 2'de belirtilen parametre değerleri dışında Tablo 2'de belirtilen parametre değerleri kullanılmıştır. Mini-batch parametresi birden fazla girdinin model üzerinde aynı anda işlemesi durumudur [29]. Bu durum ise kullanılan bilgisayarın donanım özellikleri ile alakalı olduğu için maliyetli bir iştir. Bu çalışmada mini-batch değeri 32 seçildi.

Tablo 2. Önerilen yaklaşımda kullanılan ESA mimarilerinin parametre değerleri.

Kullanılan Yazılım	ESA Mimarisi	Resim Boyutu	Optimizasyon	Momentum	Decay	Beta	Mini Batch	Öğrenme Oranı
Matlab	AlexNet	227×227	Sigmoid Gradient Descent	0.9	1e-6	-	32	0.0001
	VGG-16	224×224						
	VGG-19	224×224						



Şekil 4. Önerilen yaklaşımın tasarımlı.

4. Bulgular

Bu makaledeki deneyler, 64 bitlik Windows 10 işletimi üzerinde kurulan Matlab R2018b görüntü işleme yazılımı kullanılarak derlendi. Kullanılan bilgisayarın özellikleri; NVIDIA GeForce 2 GB grafik kartı, Intel® i5 - Core 2.5 GHz işlemci ve 8 GB RAM'e sahiptir.

Bu çalışmada kullanılan veri kümesi beş bitki türüne ait 4326 çiçek görüntüsünden oluşmaktadır. Veri kümelerini oluşturan örneklerin 3225'i sınıflandırıcının eğitimi için, geriye kalan 1101 örnek test işlemleri için kullanılmıştır. Modellerin performanslarının değerlendirilmesi için doğruluk, duyarlılık, özgünlük ve f-skor metrikleri kullanılmıştır. Çalışmanın analizinin iyi performans göstermesi için, olumlu (pozitif) ve olumsuz (negatif) kararların doğrulukları önemlidir. Bu kararlar, doğru pozitif (DP), doğru negatif (DN), yanlış pozitif (YP), yanlış negatif (YN) olarak ifade edilmektedir [30-32]. Aşağıda belirtilen denklemlerde metrik değerleri bu kararlara göre değerlendirilmektedir.

$$\text{Duyarlılık} = \frac{DP}{DP+YN} \quad (8)$$

$$\text{Özgünlük} = \frac{DN}{DN+YP} \quad (9)$$

$$\text{F-skor} = \frac{2xDP}{2xDP+YP+YN} \quad (10)$$

$$\text{Doğruluk} = \frac{DP+DN}{DP+DN+YP+YN} \quad (11)$$

Her sınıf için hesaplanan doğruluk, duyarlılık, özgünlük ve f-skor performans metriklerine ek olarak, modelin genel başarısını ölçmek için genel doğruluk metriği hesaplanmıştır. Genel doğruluk, modelin doğru tahmin sayısının, toplam tahmin sayısına oranı olarak tanımlanır.

Bu çalışma iki deney analizinden oluşmuştur. Birinci deneyde ESA mimarileri özellik seçim yöntemleri kullanılmadan DVM ile sınıflandırılmıştır. Analiz sonuçları Tablo 3'te gösterilmiştir. Analiz sonuçlarında en iyi sınıflandırma doğruluğu VGG-16 mimarisile sağlanmıştır. Elde edilen sınıflandırma başarısı %86,56'dır. ESA mimarilerinin öğrenim grafikleri ise Şekil 5'te gösterilmiştir. İkinci deneyde ise ESA mimarilerine özellik seçim yöntemleri uygulanarak analiz gerçekleştirilmiştir. Üç özellik seçim yöntemi birleştirilip aritmetik ortalamaları alınarak en iyi özellikler seçilmiştir. Özellik seçim oranı orijinal özelliklerin %10'unu alacak şekilde belirlendi. Örneğin; üç modelin bireşiminden oluşan 3000 özelliğin %10 oranını alarak özellik seçimi gerçekleştirildi ve 300 özellik seçildi. Veya 1000 özellikli AlexNet mimarisinin %10 oranını alarak özellik seçimi ile 100 özellik seçildi. İkinci deneye ilişkin analiz sonuçları Tablo 4'te gösterilmiştir. Tablo 4 incelendiğinde birleştirilmiş özellik seçim yöntemleri ESA mimarilerine ayrı ayrı ve üç mimarinin özelliklerinin bireşimine uygulanmıştır. Önerilen yöntem

ile üç modelin özellikleri birleştirilmiştir ve 3000 özellik elde edilmiştir. Özellik seçim yöntemlerinden elde edilen aritmetik ağırlık ortalamaları göz önünde alınarak bu özellikler 300'e düşürülmüştür. Neticede elde edilen sınıflandırma başarısında %4,54 oranında bir artış sağlanmıştır.

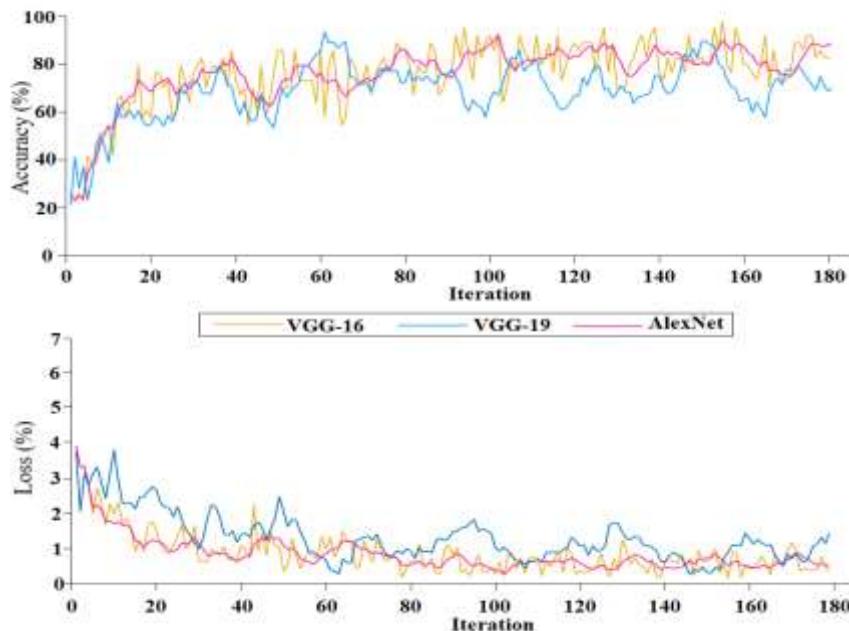
Sonuç olarak beş çiçek görüntüyü kümelerinin sınıflandırılmasında önerilen yaklaşım ile %91,10 oranında bir genel başarı sağlanmıştır. Bu başarının karmaşıklik matrisi ve ROC eğrisi grafiği Şekil 6' ta gösterilmiştir. Ayrıca özellik seçim yönteminin CNN mimarilerinde zaman olarak kazanç sağladığı Tablo 5'te gösterilmiştir.

Tablo 3. Özellik seçimi yöntemleri kullanılmadan ESA modellerinin DVM ile sınıflandırma sonuçları.

ESA Mimarisi	Sınıf Kümesi	# Özelliğ (FC8)	Doğruluk (%)	Duyarlılık	Özgünlük	F-skor	Genel Doğruluk (%)
AlexNet	Papatya	1000	93,87	0,87	0,96	0,86	86,29
	Karahindiba		95,19	0,85	0,98	0,89	
	Gül		93,13	0,86	0,95	0,84	
	Ayçiçeği		94,81	0,90	0,96	0,88	
	Lale		93,13	0,84	0,96	0,84	
VGG-16	Papatya	1000	94,83	0,85	0,98	0,88	86,56
	Karahindiba		94,92	0,90	0,96	0,89	
	Gül		93,24	0,88	0,95	0,85	
	Ayçiçeği		94,82	0,89	0,97	0,88	
	Lale		92,97	0,81	0,96	0,83	
VGG-19	Papatya	1000	92,58	0,79	0,97	0,82	83,92
	Karahindiba		93,05	0,84	0,96	0,84	
	Gül		91,94	0,85	0,94	0,82	
	Ayçiçeği		94,76	0,90	0,96	0,89	
	Lale		92,12	0,82	0,95	0,82	

Tablo 4. Özellik seçimi yöntemleri kullanılarak ESA modellerinin DVM ile sınıflandırma sonuçları.

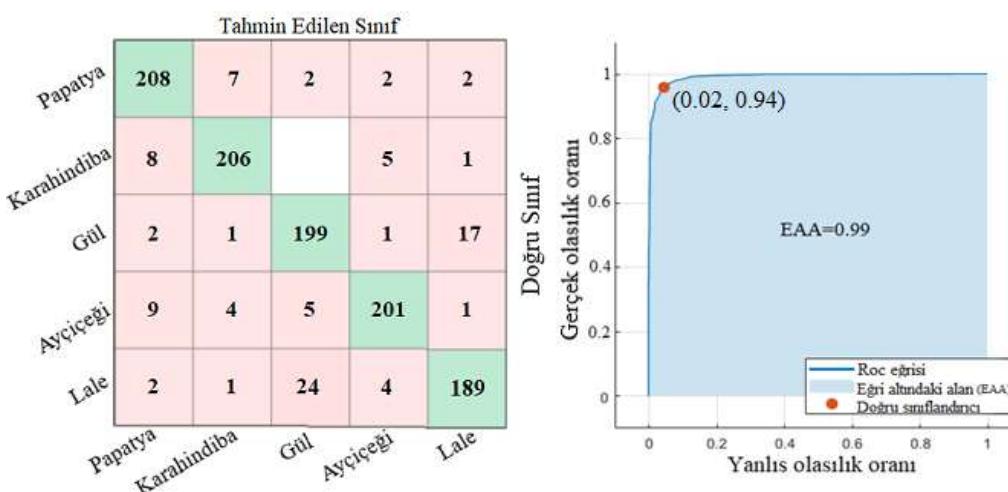
ESA Mimarisi	Sınıf Kümesi	# Özelliğ (FC8)	Doğruluk (%)	Duyarlılık	Özgünlük	F-skor	Genel Doğruluk (%)
AlexNet & (MBK + Ridge + ÖOE)	Papatya	100	92,78	0,85	0,95	0,84	84,11
	Karahindiba		94,01	0,81	0,98	0,86	
	Gül		91,68	0,83	0,94	0,81	
	Ayçiçeği		93,72	0,89	0,95	0,86	
	Lale		92,69	0,84	0,95	0,83	
VGG-16 & (MBK + Ridge + ÖOE)	Papatya	100	93,44	0,83	0,97	0,85	82,92
	Karahindiba		93,83	0,86	0,96	0,86	
	Gül		90,93	0,78	0,95	0,79	
	Ayçiçeği		92,59	0,85	0,95	0,84	
	Lale		91,20	0,82	0,94	0,81	
VGG-19 & (MBK + Ridge + ÖOE)	Papatya	100	92,84	0,82	0,96	0,84	83,74
	Karahindiba		94,08	0,92	0,95	0,87	
	Gül		92,20	0,81	0,95	0,82	
	Ayçiçeği		93,41	0,85	0,96	0,85	
	Lale		91,46	0,80	0,95	0,80	
Önerilen Yaklaşım (AlexNet + VGG16 + VGG19) & (MBK + Ridge + ÖOE)	Papatya	300	96,91	0,94	0,98	0,92	91,10
	Karahindiba		97,55	0,94	0,99	0,94	
	Gül		95,28	0,90	0,96	0,88	
	Ayçiçeği		97,18	0,91	0,99	0,93	
	Lale		95,28	0,86	0,98	0,88	



Şekil 5. CNN mimarilerinin öğrenme grafikleri.

Tablo 5. ESA modellerinin DVM sınıflandırıcısını kullanarak performans karşılaştırılması (saniye olarak).

ESA Mimarisi	Özellik Seçim Yöntemlerinin Kullanımı	Özellik Sayısı	Eğitim Süresi (sn)	Test Süresi (sn)
AlexNet	Hayır	1000	75,443	11,758
VGG-16	Hayır	1000	509,62	24,436
VGG-19	Hayır	1000	604,19	24,538
AlexNet	Evet	100	-	5,147
VGG-16	Evet	100	-	5,407
VGG-19	Evet	100	-	5,972
AlexNet + VGG16 + VGG19	Evet	300	-	7,458



Şekil 6. Önerilen yaklaşımın analiz sonuçlarının karmaşıklık matrisi ve ROC eğrisi.

5. Tartışma

Çiçek görüntü kümelerinin sınıflandırılması işlemi oldukça zor bir süreçtir. Özellikle çiçek görüntülerinin sahip olduğu yaprak, dal gibi unsurlar sınıflandırma sürecini daha da zorlaştırmaktadır. Öte yandan çiçeklerin renklerinin benzer veya aynı olması görüntü sınıflama başarısını düşürmüştür. Bu çalışmada sınıflandırma başarısını düşüren en büyük faktörün bu olduğu değerlendirilmektedir. Ancak bu alanda teknolojinin her geçen gün ilerlemesi bu zorlukları kolaylaştırmıştır. ESA mimarilerinin yetersiz kaldığı durumlarda geleneksel yöntemlerle birlikte kullanılması sınıflandırma performansına katkıda bulunabileceği nitekim bu makalenin çalışmasında gözlemlenmiştir.

Tablo 6. Aynı veri kümelerini kullanan çalışmaların sonuçları.

Yazar, Yıl	Yöntem	Başarı (%)
Baosu Guo ve ark. [33], 2019	ESA + Genetik algoritma	74,34
Francois Luus ve ark. [34], 2019	ESA + Boyut azaltma yöntemleri + Yarı denetimli öğrenme	79,33
Önerilen Yöntem	ESA + Özellik seçim yöntemi	91,10

Aynı veri setini kullanan diğer çalışmalar ile karşılaştırılması Tablo 6'da gösterilmiştir. Bu makalenin çalışmasında önerilen yöntemin başarı oranı Baosu Guo ve arkadaşlarının çalışmalarına göre daha iyi sonuç verdiği görülmüştür. Baosu Guo ve arkadaşları, genetik algoritma yöntemini kullanarak çiçek görüntülerini sınıflandırmışlardır. Elde ettikleri başarı oranı %74,34 idi. Bu çalışmalarında sınıflandırma başarısını artırıbilmeleri için tek bir ESA mimarisine (LeNet) bağlı kalmamalıdır. Farklı mimarileri kullanarak sınıflama doğruluğunu artırabilirler. Ayrıca ESA mimarisinde kullandıkları parametre değerleri ve optimizasyon yöntemlerinin de sınıflandırma sonuçlarını olumsuz etkilediği düşünülmektedir. Francois Luus ve arkadaşları yarı denetimli öğrenme yöntemi ile görüntü boyutlarını düşüren yöntemleri birlikte kullanarak sınıflandırma işlemini gerçekleştirmiştirlerdir. Çiçeklerin sınıflandırılmasında çeşitli ESA mimarilerini kullanmışlardır ve en iyi başarıyı Inception ve ResNet mimarilerinin birleştirilmesi ile yakalamışlardır. Sınıflandırma başarı oranı yaklaşık %79,33 elde etmişlerdir Francois Luus ve arkadaşlarının çalışmasında boyut düşürme yöntemleri zaman kullanımını olumlu etkilese de sınıflandırma işlemine katkı sağladığı söylenemez.

6. Sonuç

Son zamanlarda derin öğrenme modelleri ile bitki görüntülerinin sınıflandırılmasına ilgi artmıştır. Bu makalede, beş çeşitli oluşan çiçek görüntülerinin ESA mimarileri ile sınıflandırılması gerçekleştirildi. Önerilen yaklaşımda ESA mimarileri ile özellik çıkarım yöntemlerinin iç içe kullanabileceğinin görülmüştür. Ayrıca özellik seçim yöntemlerinin ayrı ayrı performansını ele almaktansa kendi aralarında ağırlık derecelerinin birleştirilebileceği ve verimli olabileceği yapılan bu çalışmada doğrulanmıştır.

Sonuç olarak beş çeşit çiçek sınıfının, sınıflandırma doğruluğu özellik seçim yöntemleri ile %91,10 oranında elde edildi. Bu çalışmanın diğer kazanımı daha az ama verimli özelliklerle daha iyi performans sağlanması sınıflandırma işleminde göstermiştir. Gelecek çalışmalarda özellik seçim yöntemlerinin ayrı bir şekilde kullanılarak sınıflandırma başarısına olan performansları kıyaslatılacaktır.

Kaynaklar

- [1] Roddy AB, Jiang G-F, Cao K, Simonin KA, Brodersen CR. Hydraulic traits are more diverse in flowers than in leaves. URL-1, <https://doi.org/10.1111/nph.15749>
- [2] Madoui S, Charef N, Arrar L, Baghianni A, Khennouf S. In vitro Antioxidant Activities of Various Extracts from Flowers-Leaves Mixture of Algerian *Cytisus triflorus*, *Annu Res Rev Biol.* 2018;26(3):1–13.
- [3] MJM C, W. BYNG J. The Number of known Plants Species in the Word and its Annual Increase. *Phytotaxa.* 2016;261(May):201–17.
- [4] Solak S, Altınışık U. Görüntü İşleme Teknikleri ve Kümeleme Yöntemleri Kullanılarak Fındık Meyvesinin Tespit ve Sınıflandırılması, SAÜ Fen Bilim Enstitüsü Derg. 2018;(February):1–1.
- [5] Toğaçar M, Ergen B. Biyomedikal Görüntülerde Derin Öğrenme ile Mevcut Yöntemlerin Kıyaslaması, Fırat Üniversitesi Mühendislik Bilim Derg. 2019;31(1):109–21.
- [6] Reichstein M, Camps-Valls G, Stevens B, Jung M, Denzler J, Carvalhais N, et al. Deep learning and process understanding for data-driven Earth system science. *2019;566(7743):195–204.* URL-2 <https://doi.org/10.1038/s41586-019-0912-1>

- [7] Altuntas Y, Cömert Z, Kocamaz AF. Identification of haploid and diploid maize seeds using convolutional neural networks and a transfer learning approach. *Comput Electron Agric.* 2019;163:104874.
- [8] Hiary H, Saadeh H, Saadeh M, Yaqub M. Flower classification using deep convolutional neural networks. *IET Comput Vis.* 2018;12(6):855–62.
- [9] Cibuk M, Budak U, Guo Y, Ince MC, Sengur A. Efficient deep features selections and classification for flower species recognition. *Measurement.* 2019;137:7–13.
- [10] Dias PA, Tabb A, Medeiros H. Apple flower detection using deep convolutional networks. 2018;99:17–28. URL-3, <http://www.sciencedirect.com/science/article/pii/S01636151730502X>
- [11] Mehdi Pour Ghazi M, Yanikoglu B, Aptoula E. Plant identification using deep neural networks via optimization of transfer learning parameters. 2017;235:228–35. URL-4, <http://www.sciencedirect.com/science/article/pii/S0925231217300498>
- [12] Seeland M, Rzannay M, Alaaqraa N, Wäldchen J, Mäder P. Plant species classification using flower images—A comparative study of local feature representations. 2017 Feb 24;12(2):e0170629. URL-5, <https://doi.org/10.1371/journal.pone.0170629>
- [13] Guru DS, Sharath Kumar YH, Manjunath S. Textural features in flower classification. *Math Comput Model.* 2011;54(3):1030–6. URL-6, <http://www.sciencedirect.com/science/article/pii/S0895717710005236>
- [14] Chai Y, Rahtu E, Lempitsky V, Van Gool L, Zisserman A. TriCos: A Tri-level Class-Discriminative Co-segmentation Method for Image Classification BT - Computer Vision – ECCV 2012. In: Fitzgibbon A, Lazebnik S, Perona P, Sato Y, Schmid C, editors. Berlin, Heidelberg: Springer Berlin Heidelberg; 2012. p. 794–807.
- [15] Chen Q, Song Z, Hua Y, Huang Z, Yan S. Hierarchical matching with side information for image classification. In: 2012 IEEE Conference on Computer Vision and Pattern Recognition. 2012. p. 3426–33.
- [16] Alexander Mamaev. Flowers Recognition | Kaggle, 2 Haziran 2019. URL-7, <https://www.kaggle.com/alxmamaev/flowers-recognition>
- [17] Togacar M, Ergen B, Sertkaya ME., Subclass Separation of White Blood Cell Images Using Convolutional Neural Network Models. *Elektron Ir Elektrotehnika*, 2019. p. 63–68. URL-8, <https://doi.org/10.5755/j01.eie.25.5.24358>.
- [18] O’Shea K, Nash R. An Introduction to Convolutional Neural Networks. 2015; URL-9, <http://arxiv.org/abs/1511.08458>
- [19] Simonyan K, Zisserman A. Very deep convolutional networks for large-scale image recognition. *arXiv Prepr arXiv14091556*. 2014;
- [20] Huang Z, Nasrullah, Wen J, Song S, Mateen M. Fundus Image Classification Using VGG-19 Architecture with PCA and SVD. *Symmetry (Basel)*. 2018;11(1):1.
- [21] Zhong G, Yan S, Huang K, Cai Y, Dong J. Reducing and Stretching Deep Convolutional Activation Features for Accurate Image Classification, *Cognit Comput.* 2018;10(1):179–86. URL-10, <https://doi.org/10.1007/s12559-017-9515-z>
- [22] Toğaçar M, Ergen B, Cömert Z., A Deep Feature Learning Model for Pneumonia Detection Applying a Combination of mRMR Feature Selection and Machine Learning Models. *Irbm* 2019. URL-11, <https://doi.org/10.1016/j.irbm.2019.10.006>.
- [23] Zhang Y, Jia S, Huang H, Qiu J, Zhou C. A novel algorithm for the precise calculation of the maximal information coefficient. *Sci Rep.* 2014;4(April).
- [24] Chen Z, Kiat Yeo C, Sung Lee Francis B, Tong Lau C. Combining MIC feature selection and feature-based MSPCA for network traffic anomaly detection, 2016. 176–181 p.
- [25] Topal M. Çoklu Doğrusal Bağlantı Durumunda Ridge ve Temel Bileşenler Regresyon Analiz Yöntemlerinin Kullanımı. 2010;41(1):53–7.
- [26] Kirpich A, Ainsworth EA, Wedow JM, Newman JRB, Michailidis G, McIntyre LM. Variable selection in omics data: A practical evaluation of small sample sizes, *PLoS One.* 2018 Jun 21;13(6):e0197910–e0197910. URL-12, <https://www.ncbi.nlm.nih.gov/pubmed/29927942>
- [27] Chen X, Jeong JC. Enhanced recursive feature elimination. In: Sixth International Conference on Machine Learning and Applications (ICMLA 2007). 2007. p. 429–35.
- [28] Huang S, Cai N, Pacheco PP, Narrandes S, Wang Y, Xu W. Applications of Support Vector Machine (SVM) Learning in Cancer Genomics, Cancer Genomics Proteomics. 2017 Dec 26;15(1):41–51. URL-13, <https://www.ncbi.nlm.nih.gov/pubmed/29275361>
- [29] Reeskamp P. Is comparative advertising a trade mark issue ? *Eur Intellect Prop Rev.* 2008;30(4):130–7.
- [30] Tharwat A, Gaber T, Ibrahim A, Hassanien AE. Linear discriminant analysis: A detailed tutorial. *AI Commun.* 2017;30(2):169–90.
- [31] Powers DMW, Ailab. Evaluation: From Precision, Recall and F-Measure To Roc, Informedness, Markedness & Correlation. 2011;2(1):37–63. URL-14, <http://www.bioinfo.in/contents.php?id=51>
- [32] Togacar M, Ergen B, Sertkaya ME. Subclass Separation of White Blood Cell Images Using Convolutional Neural Network Models. *Elektron Ir Elektrotehnika* 2019;25:63–8. URL-15, <https://doi.org/10.5755/j01.eie.25.5.24358>.
- [33] Guo, Hu, Wu, Peng, Wu. The Tabu_Genetic Algorithm: A Novel Method for Hyper-Parameter Optimization of Learning Algorithms, *Electronics.* 2019;8(5):579. URL-16, <https://www.mdpi.com/2079-9292/8/5/579>
- [34] Luus F, Khan N, Akhalwaya I. Active Learning with TensorBoard Projector. 2019;1–7. URL-17, <https://arxiv.org/abs/1901.00675>