

Onobrychis Bitkisine Ait Meyve Tiplerinin Makine Öğrenmesi Yaklaşımıyla Sınıflandırılması

Mehmet Selim KIZGIN¹, Zafer ÇAMBAY², Hakan SEPET³, Salih Taha Alperen ÖZÇELİK⁴, Hakan UYANIK⁵

¹Milli Eğitim Bakanlığı, Elazığ Mesleki ve Teknik Anadolu Lisesi, Elazığ, Türkiye

²Fırat Üniversitesi, Sağlık Hizmetleri Meslek Yüksek Okulu, Elazığ, Türkiye

³Kırşehir Ahi Evran Üniversitesi, Mühendislik Fakültesi, Elazığ, Türkiye

⁴Bingöl Üniversitesi, Elektrik-Elektronik Mühendisliği, Bingöl, Türkiye

⁵Munzur Üniversitesi, Elektrik-Elektronik Mühendisliği, Tunceli, Türkiye

¹ mehmetseimkizgin@hotmail.com, ² zcambay@firat.edu.tr, ³hakan.sepet@ahievran.edu.tr, ⁴sozcelik@bingol.edu.tr, ⁵hakanuyanik@munzur.edu.tr

(Geliş/Received: 10/04/2023;

Kabul/Accepted: 25/09/2023)

Öz: Bu çalışmada amaç makine öğrenmesi ve Yerel İkili Örüntü (YİÖ) hakkında genel bir bilgi verip bu bilgi ışığında Türkiye’de yetişen korunga (*Onobrychis*) bitki meyvelerini makine öğrenmesi ile sınıflandırmaktır. 4 farklı Korunga (*Onobrychis*) türüne toplam 448 adet meyve görüntüsü kullanılarak bir veri tabanı oluşturulmuştur. Bu türler sırasıyla *O. cappadocica*, *O. argyrea*, *O. hypargyrea* ve *O. tournefortii*’dir. Korunga (*Onobrychis*) meyve çeşitlerini sınıflandırmasını yapmak için makine öğrenmesi yöntemleri kullanılmıştır. Kullanılan bu yöntemler sırasıyla Destek Vektör Makinesi (DVM), Naif Bayes(NB), Karar Ağaçları (KA) ve K-En Yakın Komşu (k-EYK) olmak üzere dört farklı yöntem ile sınıflandırma işlemi yapılmıştır. Bu dört farklı yöntemin pe¹rformansları karşılaştırılıp en başarılı modelin %99,6 doğru sınıflandırma başarı oranı ile Destek Vektör Makinesi Yöntemi olduğu belirlenmiştir.

Anahtar kelimeler: Makine öğrenmesi, Yerel İkili Örüntü, Örüntü Tanıma, Korunga (*Onobrychis*)

Classification of Fruit Types of Onobrychis with Machine Learning Approach

Abstract: In this study, the aim is to give general information about machine learning and Local Binary Pattern (LBP) and to classify sainfoin (*Onobrychis*) plant fruits grown in Türkiye with machine learning in the light of this information. A database was created using a total of 448 fruit images of 4 different Sainfoin (*Onobrychis*) species. These species are *O. cappadocica*, *O. argyrea*, *O. hypargyrea* and *O. tournefortii*, respectively. Machine learning methods were used to classify sainfoin (*Onobrychis*) fruit varieties. These methods were classified by four different methods, namely Support Vector Machine Method, Naive Bayes Algorithm Method, Decision Trees Method and K-Nearest Neighbor Method. The performances of these four different methods were compared and it was determined that the most suitable model was the Support Vector Machine Method with a 99.6% correct classification success rate.

Key words: Makine öğrenmesi, Local Binary Pattern (LBP), Pattern Recognition, *Onobrychis*.

1.Giriş

Günümüzde bitkilerin sınıflandırılması daha çok geleneksel yöntemler ile yapılmaktadır. Geleneksel bitki sınıflandırma yöntemleri, botanikçilerin yıllarca biriktirdiği gözlemsel bilgilere ve karmaşık sınıflandırma kılavuzlarına dayanır. Bitkilerin morfolojik ve anatomik özelliklerine göre uzmanlar tarafından sınıflandırılır ve tanımlanır. Bu deneyim ve uzmanlık gerektiren süreç yıllar süren saha çalışmaları sonucunda kazanılan bir yetenektir. Bitkilerin sınıflandırılması için uzman botanikçi morfolojisi (şekil, boyut, renk, doku), anatomisi, üreme yöntemleri ve diğer özellikleri gibi gözlemlenebilir özellikleri bilmelidir. Pratik ve kullanışlı otomatik bir bitki tanıma sistemi oluşturulması, bitkilerin ayrıştırılması ve yönetilmesi açısından oldukça verimli bir çalışma olabilir. Makine öğrenmesiyle bitki sınıfının otomatik belirlenmesi zaten az sayıda olan botanik uzmanlığı mesleğindeki uzmanlık gereksiniminden kurtaracaktır. Ayrıca uzmanlar tarafından bitki sınıflandırma eyleminin manuel gerçekleştirilmesi uzun zaman alan bir işlem olacaktır. Makine öğrenmesi büyük veri kümelerinde hızlı bir şekilde çalışabilir. Bu sayede çok sayıda bitkiyi hızlıca sınıflandırabilir. Makine öğrenmesi yanlış sınıflandırmaları öğrenerek ve düzelterek zamanla daha iyi hale gelebilir. Bu durum da sürekli iyileştirmeyi mümkün kılar. Ayrıca manuel sınıflandırma için bazı durumlarda mikroskopik inceleme gerekmesi ekipman ve özel eğitim gerektirebilir. Başarılı bir makine öğrenmesi algoritması ise buna ihtiyaç duymaz. Aynı zamanda bazı bitki türlerinin morfolojik ve anatomik açıdan benzerliği manuel sınıflandırma için hata olasılığını arttırabilir. Bitkilerin tanınabilirliğini sağlamak için her türe ait bitki çeşitlerin veri tabanını kurmak gerekir. Bu veri tabanını kurmak zor ama önemli bir adımdır. Bitkileri dijital ortama aktararak bitki türlerinin çeşitlerini sınıflandırmada

*Sorumlu Yazar mehmetseimkizgin@hotmail.com Yazarların ORCID Numarası: ¹0000-0001-8176-6259, ²0000-0002-1170-7525, ³0000-0003-0591-3397, ⁴0000-0002-7929-7542, ⁵0000-0002-6870-7569

hızlı ve iyi bir verim elde edilebilir. Meyvelerin şekil ve özelliklerini kullanarak bitki tohumlarını tanıma yaklaşımı bu çalışmanın temelini oluşturmaktadır [1-2].

Acar ve arkadaşları [3], yapmış oldukları çalışmada; Zambak yapraklarında oluşan pas lekelerinin Gabor dalgacık yöntemi kullanılarak pas lekesinin oluşup oluşmadığını veya sağlıklı olup olmadığını tespit ettiler. Bu çalışmada 53 resim ile çalıştılar. Alınan bu resimler her bir pikseline karşılık gelecek noktaya özellik vererek matris oluşturdular ve elde edilen bu matrisler normalizasyon ile sınıflandırılarak hastalıklı mı yoksa sağlıklı mı olduğunu tespit ettiler. Çalışmada zambak yapraklarındaki pas lekelerini Gabor dalgacık yöntemi %80 sınıflandırma başarısı elde edilmiştir.

Kılıç ve arkadaşları [4], narenciye yapraklarının hastalıklarını teşhis etmek amacıyla 4 farklı veri seti üzerinde çalışma yaptılar. Bu veri setleri üzerinden VGG16 ve AlexNet ile iki ESA modeli oluşturularak, narenciye hastalıklarını tespit etmek için kullandılar. Bu yöntemde VGG16 ve AlexNet modelleri oluşturulan veri setlerini kullanarak %82,64, %93,39 ve %92,56 doğruluk oranlarıyla başarılı bir şekilde sınıflandırıldılar.

Elmas ve arkadaşları [5], mantar türlerini tespit etmek için 18 aileye ait 472 sınıfın görüntülerini içeren düzensiz veri, düzenlenmiş veri ve düzenlenerek oluşturulmuş veri olmak üzere üç farklı veri seti oluşturdular. Bu çalışmada 6 farklı evrişimli sinir ağını transfer öğrenme yöntemi kullanılarak eğittiler. Bu transfer modelde, düzenlenmiş mantar verilerini %97,62 doğruluk oranı ile en başarılı veri seti olarak sınıflandırmışlardır.

Baranval ve arkadaşları [6], Elma yaprağı hastalığı tanımak için 6 farklı hastalığa ait 2462 görüntü içeren veri seti oluşturdu. Derin öğrenme ile elma yaprağı hastalıklarını tanımak için DenseNet-12 derin evrişim ağına dayalı üç regresyon yöntemi, çok etiketli sınıflandırma ve odak kaybı fonksiyonu ile oluşan üç farklı sınıflandırma yöntemi kullandılar. Üç regresyon yöntemindeki başarı oranı %93,51, çok etiketli sınıflandırma yöntemindeki başarı oranı %93,31 ve odak kaybı fonksiyonundaki başarı oranını ise %93,71 olarak tespit ettiler. Elma yaprağı hastalığını tespiti için %93,71 başarı oranı ile odak kaybı fonksiyonunu önerdiler.

Koklu ve arkadaşları [7], 5 farklı türden oluşan 500 asma yaprağı görüntüsüyle bir veri seti oluşturdular. Ardından veri artırma kullanarak verileri 5 kat çoğaltıp 2500 görüntü elde ettiler. MobileNetv2 ile sınıflandırma gerçekleştirdiler. Ayrıca kullandıkları derin öğrenme ağından aldıkları öznelik vektörlerinden Ki-Kare yöntemiyle öznelikleri seçerek destek vektör makineleri (DVM) ile sınıflandırdılar. Araştırmacılar Kübik DVM için %97,6 sınıflandırma başarısı elde ettiler.

Ganguly ve arkadaşları [8], derin öğrenme tekniğiyle Bonferroni füzyon öğrenme yöntemini birleştirerek BLeafNet isminde yeni bir yaprak görüntü sınıflandırma modeli geliştirdiler. Malayakew, Leafsnap ve Flavia veri setlerini kullanan araştırmacılar sırasıyla %98,54, %92,22 ve %98,7 sınıflandırma başarısı elde ettiler.

Belal ve arkadaşları [9], 12 türe ait 960 benzersiz bitki ile yaklaşık 5.000 adet bitki fidesi görüntüsü içeren bir veri seti oluşturdular. Bu bitki fidelerini derin öğrenme tekniği olan, Konvolüsyon Sinir Ağı (KSA) algoritmaları kullanarak sınıflandırma işlemi yaptılar. Bu sınıflandırma algoritması ile %99,48' lik başarı elde ettiler.

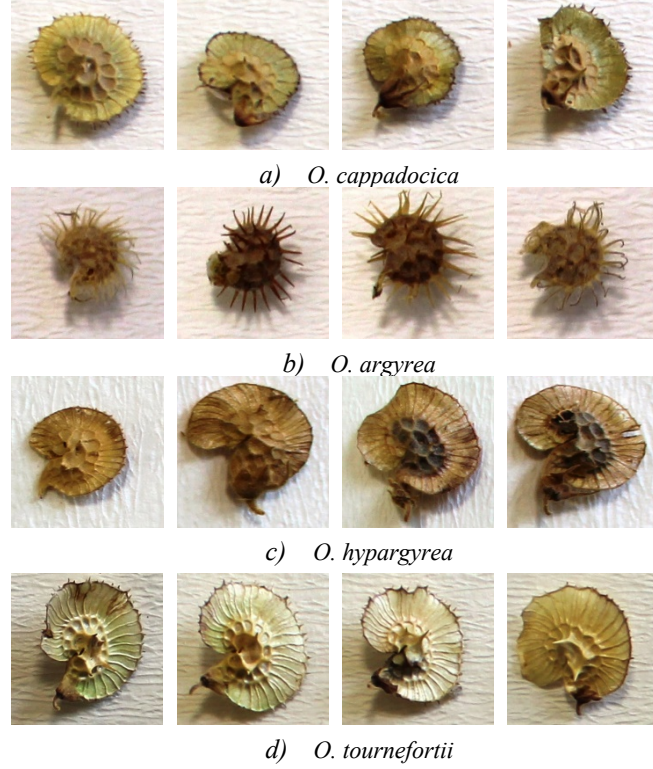
Argüeso ve arkadaşları [10], tarlada çekilen bitki görüntülerden bitki hastalığını tespit etmek için ABD'deki Penn State Üniversitesinde bulunan PlantVillage bilgi bankasından, 14 mahsul türü ve 26 hastalık içeren 38 sınıfla 54.303 yaprak görüntülerini veri seti olarak kullandılar. Bu çalışmada bitki yapraklarını sınıflandırmak için Few-Shot Learning (FSL) algoritması ile çalıştılar. Few-shot Learning algoritmasından %90 doğruluk oranı ile sınıflandırma başarısı elde ettiler.

Joggekar ve arkadaşları [11], Leafminer, Sooty Mold ve Pulvinaria üç farklı narenciye hastalıklarını tespit etmek için 1774 narenciye yaprağı görüntüsünden oluşan bir veri seti oluşturdular. Sınıflandırıcı düzeyinde çeşitlilik için AlexNet, VGG16, ResNet50 ve Inception_ResNet_v2 gibi farklı mimariler kullandılar. Evrişimli sinir ağların (ESA) yöntemi ile %99,04'lük başarı elde ettiler.

Bu çalışmadaki amaç, onobrychis (korunga) bitki meyvelerini tanımak ve tanınan meyveleri sınıflandırmaktır. Sınıflandırma için 4 farklı takson sınıfından 448 görüntüden oluşan bir veri seti oluşturulmuştur. Toplanan görüntülerden yerel ikili örüntü (YİÖ) özellikleri çıkarılmıştır. Sonra YİÖ özellikleri çeşitli makine öğrenmesi sınıflandırıcıları aracılığıyla sınıflandırılmıştır. Bu sınıflandırma işlemi sayesinde botanikçiler *Onobrychis* cinsini rahatlıkla tanıyabilecek ve sınıflandırabilecektir. Böylece iklim ve toprak yapısına uygun bitki seçimini yaparak daha verimli ürün elde etmesine olanak sağlayacaktır.

2. Materyal ve Metod

Türkiye'nin değişik bölgelerinde yetişen *Onobrychis* (korunga) cinsine ait 4 farklı taksondan 448 farklı meyve toplanarak bir veri tabanı oluşturulmuştur. *Onobrychis* cinsine ait türler sırası ile *O. cappadocica*, *O. argyrea*, *O. hypargyrea* ve *O. tournefortii* 'dir. Aşağıdaki Şekil 1'de türlere ait örnek görüntüler verilmiştir [12].



Şekil 1. Kullanılan *Onobrychis* (korunga) bitkisine ait resimler
a) *O. cappadocica*, b) *O. argyrea*, c) *O. hypargyrea* ve d) *O. tournefortii*

Onobrychis bitkisine ait taksonlar Türkiye genelinde farklı lokalitelerden toplanarak görüntüleri veri tabanına kaydedilmiştir. Her nesne ayırt edilebilecek bir fon üzerine yerleştirilerek ve meyve ile kamera arasında 40 cm yükseklik olacak şekilde görüntüleri alınmıştır. Bu alına görüntüler veri tabanına kaydedilmiştir. Bu yükseklikten, ölçümleme nesnesi olarak kullanılan uzunluğu bilinen bir nesnenin görüntüsü alınmış, paranın resimdeki boyutları ile gerçek boyutları karşılaştırılmış ve değerlerin aynı olduğu görüldüğünden kullanılan yükseklik geçerli kılınmıştır [12]. Elde edilen görüntülerin boyutu 300x300 dür.

Bu çalışmanın amacı *Onobrychis* cinsine ait meyvelerden etkili özellikler elde ederek bu özellikler üzerinden görüntülerin sınıflandırılmasıdır. Bu çalışmada Destek Vektör Makinesi, Naif Bayes Algoritması, Karar Ağaçları, K-en Yakın Komşu gibi makine öğrenmesi sınıflandırma algoritmaları kullanılarak sınıflandırma işlemi gerçekleştirilmiştir.

3. Yerel İkili Örüntü (YİÖ)

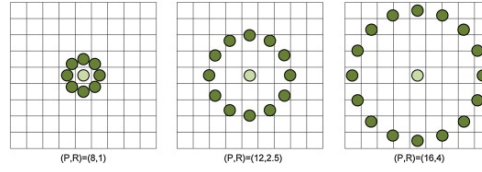
YİÖ operatörü yerel komşuluk değerlerine dayalı bir özellik çıkarma yöntemidir. Bu yöntem ilk olarak Zhou vd. tarafından ortaya çıkarılmıştır [13]. YİÖ, her bir pikselin komşu piksellerle karşılaştırılmasına dayanır, bu da yerel özelliklerin vurgulanmasına yardımcı olur. Bitki tohumları gibi nesnelerin benzersiz yerel özellikleri, tanıma ve sınıflandırmada önemlidir. YİÖ, görüntülerdeki dokusal bilgiyi yakalamak için etkili bir yöntemdir. Bitki tohumlarının yüzey dokusu ve deseni, bitki türlerini ayırt etmek için önemli bir özelliktir ve YİÖ bu özelliği vurgular. YİÖ, özellik vektörlerini düşük boyutlu tutarak, veri boyutunu azaltabilir ve hesaplama maliyetini düşürebilir. Bu durum, büyük veri setleri üzerinde etkili bir şekilde çalışmayı kolaylaştırabilir. Bitki tohumu sınıflandırmada efektif olabilecek bu güçlü yanlarından ötürü çalışma için öznitelik olarak YİÖ seçilmiştir.

YİÖ, bir görüntünün mikro örüntülerden oluştuğunu varsayar. Bu mikro örüntüler tekdüze (uniform) ve tekdüze olmayan örüntüler şeklinde ikiye ayrılmaktadır. Genel olarak YİÖ'nün uygulandığı çalışmalarda tekdüze örüntüler kullanılmaktadır. YİÖ operatörü, görüntünün her pikseli için bir etiket oluşturmaktadır [13]. Merkez piksel in komşu piksellerle karşılaştırılması sonucu elde edilen ikili tabandaki sayı merkez piksel için YİÖ değerini ifade etmektedir. YİÖ değerleri görüntüdeki her pikselin komşuları arasındaki farkın basamak fonksiyonu ile ikilileştirilmesi ile elde edilir. YİÖ etiketi aşağıdaki eşitlik ile elde edilir [14].

$$YiÖ_{P,R}(X_c) = \sum_{p=0}^{P-1} \mu(X_p - X_c) 2^p \quad (1)$$

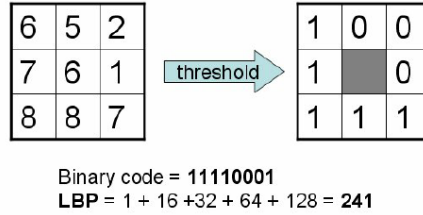
$$\mu(y) = \begin{cases} 1, & y \geq 0 \text{ ise} \\ 0, & y < 0 \text{ ise} \end{cases} \quad (2)$$

Denklem (1)'de belirtilen X_c , YİÖ etiketi üretilen merkez pikseli, X_p , merkez pikselin komşularını, R komşuların merkez piksele olan uzaklığını, P ise işleme sokulan komşu sayısını belirtir [15]. Denklem 2' de ise merkez pikselin komşuları arasındaki mesafenin sıfıra veya sıfırdan büyük olduğunda 1 sonucu ürettiğini sıfırdan küçük olduğunda ise sıfır değerini ürettiğini belirten denklemdir. Bu yapı farklı dairesel komşuluklar içinde kullanılabilirliğini göstermektedir. Farklı P ve R değerlerin kullanılması ile farklı ölçekteki dokuların analizini gerçekleştirmek mümkün olmaktadır [15]. Şekil 2'de farklı YİÖ operatörleri verilmiştir.



Şekil 2. Farklı P ve R parametrelerine göre YİÖ operatörleri [16].

Şekil 3'te de YİÖ operatörü ile bir pikselin etiketlenmesine ilişkin örnek verilmiştir.



Şekil 3. Bir piksel için YİÖ değerinin elde edilmesi [17]

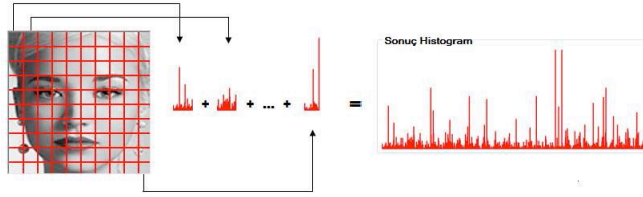
Her YİÖ sayısal değerinin tümü doku analizinde kullanılmaz [17]. Tekdüze örüntüler, YİÖ dilinde 1-0 veya 0-1 gibi geçiş sayıları 2 veya 2'den daha az olanlardır. Örnek verecek olursak 11111 ve 00000 örüntüleri herhangi bir sayı geçişi olmadığından 0 geçişe, 001100 ve 111011 örüntüleri ise 2 sayı geçişine sahip olduklarından tekdüze örüntüler olarak ifade edilir. Ancak 10111010 sayısı 5 geçişli 00100100 sayısı ise 4 geçişli bir sayı olduklarından bunlar tekdüze örüntü değildir. 3x3 alt resimler için 256 farklı YİÖ kodu oluşmaktadır. Bunların 58 tanesi tekdüze örüntüdür. Tekdüze olmayan örüntülerin tümü tek bölmede toplanarak toplam 59 bölme elde edilir. YİÖ histogramları bu 59 örüntü üzerinden çıkarılmaktadır [18].

YİÖ merkez pikselin komşular ile arasındaki sıralamasını aynen alır, fakat komşuların sıralamasını almaz. Bundan dolayı ikilik sayı arasındaki geçiş düzenine bakarak YİÖ'nün düzenli olup olmadığına karar verilir. Düzenli olabilmesi için ölçüt olarak 1-0 veya 0-1 geçişlerin en fazla iki tane olması şartı koşulur. 256 YİÖ kodu incelendiğinde ise 58 tanesi düzenli 198 tanesi ise düzensizdir. Düzenli YİÖ kodlarının bulunması için aşağıdaki denklemler kullanılır. Denklem 3' de ve Denklem 4'te belirtilen U_c merkez pikselin düzenlilik ölçüsü iken s ise özel dönüşüm fonksiyonunu belirtir [19].

$$U_c = \sum_{i=0}^{p-1} |s(g_i - g_c) - s(g_{i-1} - g_c)| \quad (3)$$

$$YiÖ_{cj} = \begin{cases} \sum_{i=0}^{P-1} |s(g_i - g_c), \text{eğer } U_c \leq j, j = 2 \\ P + 1, \text{aksi takdirde} \end{cases} \quad (4)$$

Histogram, YİÖ ile etiketlenen resim üzerindeki verilerin dağılımlarını belirlemek için kullanır. Görüntü üzerinde genel bir histogram oluşturmaktansa, lokal dağılımlara göre histogram oluşturmak görüntüyü sisteme tanıtmaya işleminde daha başarılı olacaktır. Lokal histogramı oluşturmak için resim belirli boyutlarda (20x20 gibi) bloklara ayrılır. Daha sonra elde etmiş olduğumuz her bir bloğun histogramı çıkarılır. YİÖ' llerde düzenli olanlar, dikkate alınır. Her blokta 58 adet düzenli blok olduğundan bunlar dikkate alınır. Düzensiz olanlar için ise tek bir değer tutmak yeterli olacaktır. Örnek verilecek olursa 10x10'luk bloklar için 59x100 tane özellik vektörü bulunur. Şekil 4'te görüldüğü gibi alt bloklardan oluşmuş histogramlar uç uca eklenerek YİÖ lokal histogramı oluşturulur [20].



Şekil 4. YİÖ Histogramının elde edilmesi [21]

4. Bulgular ve Tartışma

Sınıflandırma kavramı, verinin daha önceden belirlenen sonuçlara uygun olarak ayrıştırılmasını sağlayan bir yöntemdir. Sınıflandırma algoritmaları, daha önceden verilen eğitim kümesinden bu ayrışma şeklini öğrenirler ve daha sonra sınıfı belli olmayan test verileri verildiğinde doğru şekilde ayrıştırmaya çalışırlar. Sonuçlar, daha önceden bilindiği için sınıflama, veri kümesini denetimli (supervised) olarak öğrenirler. Sınıflandırma yöntemleri, verilen bir eğitim kümesinden bu dağılım şeklini diğer bir deyişle hangi parametreler ile gözlemlerin hangi kategorilere ayrışacağını öğrenirler ve daha sonra test verileri geldiğinde doğru şekilde sınıflandırmaya çalışırlar. Veri kümesi üzerinde verilen bu sınıfları belirten değerlere etiket ismi verilir ve gerek eğitim gerekse test sırasında verinin sınıfının belirlenmesi için kullanılırlar [21,22].

Bu çalışmada amacımız *Onobrychis* cinsine ait meyvelerden etkili özellikler elde ederek bu özellikler üzerinden görüntülerin sınıflandırılmasıdır. Bu çalışmada Destek Vektör Makinesi, Naive Bayes Algoritması, Karar Ağaçları, K-en Yakın Komşu ve Logistik Regrasyon yöntemleri gibi farklı sınıflandırma algoritmaları kullanılarak sınıflandırma işlemi gerçekleştirilmiştir. Bu sınıflandırmaya ait sonuçlar Tablo 1'de verilmiş ve aşağıdaki bölümlerde detaylandırılmıştır..

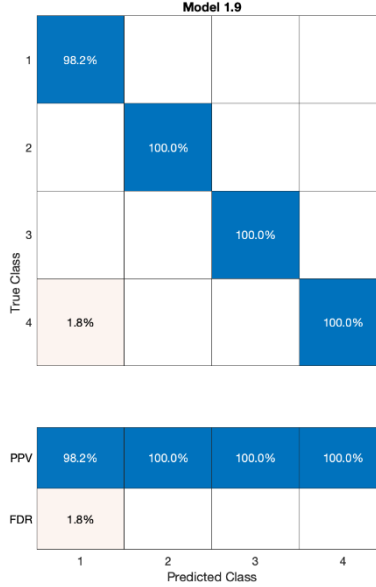
Tablo 1. Sınıflandırma Sonuçları

Sınıflandırma Metodu	Elde edilen sınıflandırma başarısı
DVM	99,6
Naive Bayes	96,2
Karar Ağaçları	92,4
k-EYK	96,2

4.1 Destek vektör makinesi yöntemi

Destek vektör makineleri (DVM), iki farklı modelde kullanılmaktadır. Birincisi regresyon analizi ikincisi ise sınıflandırma modelidir. Denetimli öğrenme modeline dayandığından sınıflandırma modeli daha çok tercih edilmektedir. Bu yöntem çalışacak verinin tipine bağlı olarak, çekirdek fonksiyonlarında da kullanılabilir. Bu özelliğinden dolayı lineer ve lineer olmayan sınıflandırma işlemlerini gerçekleştirebilmektedir. Sınıflandırma modeli uygulanırken verinin yapısına bağlı olarak iki farklı düzlem görülmektedir. Eğer tam ayrıştırılabilir veri

kullanılırsa veriler hiper düzlem ile sınıflandırılır. Tam tersi durumda veriler tam ayrıştırılmayan durumda kullanılmışsa aynı boyutta tek bir düzlemde sınıflandırma işlemi gerçekleşir. Bu özelliğinden dolayı farklı çekirdek fonksiyonları tercih edilmektedir [23].



Şekil 5. Kuadratik DVM Karmaşıklık Matrisi

Şekil 5'te gösterilen Kuadratik DVM sınıflandırma modeline göre aşağıdaki sonuçlar elde edilmiştir:

Birinci satırda gösterilen *O. cappadocica* bitki meyvesi diğer meyve çeşitlerinden %98,2 doğruluk oranı ile ayrılmıştır.

İkinci satırda gösterilen *O. argyrea* bitki meyvesi diğer meyve çeşitlerinden %100 doğruluk oranı ile ayrılmıştır.

Üçüncü satırda gösterilen *O. hypargyrea* bitki meyvesi diğer meyve çeşitlerinden %100 doğruluk oranı ile ayrılmıştır.

Dördüncü satırda gösterilen *O. tournefortii* bitki meyvesi diğer meyve çeşitlerinden %100 doğruluk oranı ile ayrılmıştır.

Yukarıdaki sonuçlara göre, Kuadratik DVM sınıflandırma modeline göre başarı oranı %99,6 çıktığı görülmektedir. Bu başarı oranına göre *O. cappadocica* bitki meyvesi %98,2 doğruluk oranı ile ortalama başarı oranının altında kaldığı, *O. argyrea*'nin meyvesi %100, *O. hypargyrea*'nin meyvesi %100 ve *O. tournefortii*'nin meyvesi %100 doğruluk oranları ile ortalama başarı oranının üstünde olduğu görülmektedir. Quadratic Svm sınıflandırma modeli *O. argyrea*, *O.hypargyrea* ve *O. tournefortii* bitki meyvelerinde daha başarılı bir sonuç verdiği söylenebilir.

4.2 Naive bayes algoritma yöntemi

Naive Bayes algoritma yöntemi, sınıflandırma işlemi için daha önce sisteme tanıtılmış olan verileri kullanarak daha sonra verilen verinin hangi sınıfa dahil olacağını tahmin eden istatistik tabanlı bir algoritma yöntemidir. Çalışma yöntemi olarak, sisteme tanıtılmış olan verilerin ve sınıfların bulunma sıklığını hesaplayıp, bu hesaba göre yeni gelen verinin hangi sınıfa dahil olacağını tahmin eder. Kolay uygulanabilirliği ve hızlı hesaplama yönteminden dolayı veri madenciliği sınıflandırma modelleri içinde en fazla tercih edilen algoritma yöntemidir [23].

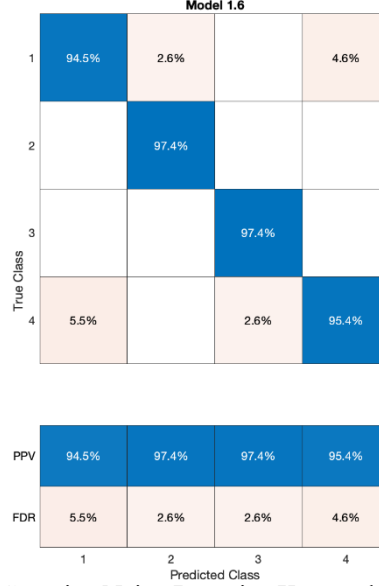
Şekil 6'da gösterilen Gaussian Naive Bayesian sınıflandırma modeline göre aşağıdaki sonuçlar elde edilmiştir:

Birinci satırda gösterilen *O. cappadocica* bitki meyvesi diğer meyve çeşitlerinden %94,5 doğruluk oranı ile ayrılmıştır.

İkinci satırda gösterilen *O. argyrea* bitki meyvesi diğer meyve çeşitlerinden %97,4 doğruluk oranı ile ayrılmıştır.

Üçüncü satırda gösterilen *O. hypargyrea* bitki meyvesi diğer meyve çeşitlerinden %97,4 doğruluk oranı ile ayrılmıştır.

Dördüncü satırda gösterilen *O. tournefortii* bitki meyvesi diğer meyve çeşitlerinden %95,4 doğruluk oranı ile ayrılmıştır.



Şekil 6. Gaussian Naive Bayesian Karmaşıklık Matrisi

Yukarıdaki sonuçlara göre, Gaussian Naive Bayesian sınıflandırma modeline göre başarı oranı %96,2 çıktığı görülmektedir. Bu başarı oranına göre *O. cappadocica* bitki meyvesi %94,5 ve *O. tournefortii* bitki meyvesi %95,4 doğruluk oranları ile ortalama başarı oranının altında kaldığı, *O. argyrea*'nin meyvesi %97,4 ve *O. hypargyrea*'nin bitki meyvesi %97,4 doğruluk oranları ile ortalama başarı oranının üstünde olduğu görülmektedir. Gaussian Naive Bayesian sınıflandırma modeli *O. argyrea* ve *O. hypargyrea*'nin meyvelerinde daha başarılı bir sonuç verdiği söylenebilir.

4.3 Karar ağaçları yöntemi

Karar ağacı modeli isminden de anlaşılacağı üzere ağaç yapısına benzetilmektedir. Ağaç yapısı gibi kök, dal ve yapraklardan oluşmaktadır. Karar ağaçları sınıflandırma işlemini ilk köklerde yapar, daha sonra sırasıyla düğümler, dallar ve yaprak (karar sınıfı) şeklinde gider. Ağaç yapısında yapraklar elde edilinceye kadar devam eder. Düğüm noktaları karar verme işleminin yapıldığı yerdir. Buradaki sonuca göre dallar oluşur ve elde edilen dallara göre yapraklar elde edilir. Eğer yaprak oluşmuyorsa düğüm oluşturulur ve bu işlem yaprak (karar sınıfı) oluşuncaya kadar devam edilir [24].

Model 1.1

True Class	1	84.6%	2.7%	2.6%	6.7%
	2	3.4%	96.4%	0.9%	
	3	0.9%	0.9%	95.7%	
	4	11.1%		0.9%	93.3%
		1	2	3	4
PPV	84.6%	96.4%	95.7%	93.3%	
FDR	15.4%	3.6%	4.3%	6.7%	

Predicted Class

Şekil 7. Fine Tree Karmaşıklık Matrisi

Şekil 7’de gösterilen Fine Tree sınıflandırma modeline göre aşağıdaki sonuçlar elde edilmiştir:

Birinci satırda gösterilen *O. cappadocica*’nın bitki meyvesi diğer meyve çeşitlerinden %84,6 doğruluk oranı ile ayrılmıştır.

İkinci satırda gösterilen *O. argyrea*’nin bitki meyvesi diğer meyve çeşitlerinden %96,4 doğruluk oranı ile ayrılmıştır.

Üçüncü satırda gösterilen *O. hypargyrea*’nin bitki meyvesi diğer meyve çeşitlerinden %95,7 doğruluk oranı ile ayrılmıştır.

Dördüncü satırda gösterilen *O. tournefortii*’nin meyvesi diğer meyve çeşitlerinden %93,3 doğruluk oranı ile ayrılmıştır.

Yukarıdaki sonuçlara göre, Fine Tree sınıflandırma modeline göre başarı oranı %92,4 çıktığı görülmektedir. Bu başarı oranına göre *O. cappadocica*’nın bitki meyvesi %84,6 doğruluk oranı ile ortalama başarı oranının altında kaldığı, *O. argyrea*’nın meyvesi %96,4, *O. hypargyrea*’nın meyvesi %95,7 ve *O. tournefortii*’nin meyvesi %93,3 doğruluk oranları ile ortalama başarı oranının üstünde olduğu görülmektedir. Fine Tree sınıflandırma modeli *O. argyrea*, *O. hypargyrea* ve *O. tournefortii* meyvelerinde daha başarılı bir sonuç verdiği söylenebilir.

4.4 K-En yakın komşu yöntemi

k-EYK sınıflandırma yöntemi en eski ve en basit yöntemlerden biridir k-EYK modelinde, yeni gelen verinin daha önceden K kategoriye ayrılmış olan verilere yakınlığını inceler. Yeni bir veri geldiğinde en yakın komşusuna bakıp, verinin hangi sınıfta olduğuna karar verir. Basit olmasına rağmen rekabetçi sonuçlar elde edilir. Bazı durumlarda, karmaşık öğrenme algoritmalarından daha iyi sonuçlar elde eder [24].

Model 1.14

True Class	1	93.9%	0.9%		3.7%
	2		98.2%		
	3			100.0%	
	4	6.1%	0.9%		96.3%
		Predicted Class			
		1	2	3	4

PPV	93.9%	98.2%	100.0%	96.3%
FDR	6.1%	1.8%		3.7%

Şekil 8. K-EYK için Karmaşıklık Matrisi

Şekil 8’de gösterilen Fine Knn sınıflandırma modeline göre aşağıdaki sonuçlar elde edilmiştir:

Birinci satırda gösterilen *O. cappadocica*’nın meyvesi diğer meyve çeşitlerinden %93,9 doğruluk oranı ile ayrılmıştır.

İkinci satırda gösterilen *O. argyrea* ’nin meyvesi diğer meyve çeşitlerinden %98,2 doğruluk oranı ile ayrılmıştır.

Üçüncü satırda gösterilen *O. hypargyrea* ’nin meyvesi diğer meyve çeşitlerinden %100 doğruluk oranı ile ayrılmıştır.

Dördüncü satırda gösterilen *O. tournefortii*’nin meyvesi diğer meyve çeşitlerinden %96,3 doğruluk oranı ile ayrılmıştır.

Yukarıdaki sonuçlara göre, Fine Knn sınıflandırma modeline göre başarı oranı %96,2 çıktığı görülmektedir. Bu başarı oranına göre *O. cappadocica*’nın meyvesi %93,9 ve *O. tournefortii*’nin meyvesi %96,3 doğruluk oranları ile ortalama başarı oranının altında kaldığı, *O. argyrea*’nin meyvesi %98,2 ve *O. hypargyrea*’nin meyvesi %100 doğruluk oranları ile ortalama başarı oranının üstünde olduğu görülmektedir. Fine Knn sınıflandırma modeli *O. argyrea*, *O. hypargyrea* ve *O. tournefortii* bitki meyvelerinde daha başarılı bir sonuç verdiği söylenebilir.

5. Sonuçlar

Bu çalışmada, Onobrychis cinsinin 4 taksonuna (*O. cappadocica*, *O. argyrea*, *O. hypargyrea* ve *O. tournefortii*) ait meyve görüntüleri kullanılarak sınıflandırma işlemi gerçekleştirildi. Bu sınıflandırma işlemi için dört farklı makine öğrenmesi algoritması kullanıldı.

İlk kullanılan sınıflandırma algoritması DVM sınıflandırma algoritmasıdır. Doğrusal DVM, ikinci dereceden Kuadratik DVM, Kübik DVM, kaba gauss DVM ve çekirdek DVM olmak üzere beş farklı tip DVM sınıflandırma yöntemi kullanıldı. Bu yöntemler arasında ikinci dereceden kuadratik DVM yöntemi %99,6 ile en yüksek sınıflandırma başarısı veren yöntem olmuştur.

Kullanılan bir diğer sınıflandırma algoritması Naive Bayes algoritmasıdır. Gauss Naive Bayes ve Kernel Naive Bayes olmak üzere iki farklı sınıflandırma yöntemi uygulandı. Bu yöntemler arasında Gauss Naive Bayes sınıflandırma yöntemi %96,2 ile en yüksek sınıflandırma başarısı veren yöntem olmuştur.

Kullanılan bir diğer sınıflandırma algoritması Karar ağaçları algoritmasıdır. Güzel ağaç, ortalama ağaç, kaba ağaç, güçlendirilmiş ağaçlar, yükseltilmiş ağaç ve torbalı ağaç olmak üzere altı farklı sınıflandırma yöntemi uygulandı. Bu yöntemler arasında güzel ağaç sınıflandırma yöntemi %92,4 ile en yüksek sınıflandırma başarısını veren yöntem olmuştur.

Kullanılan bir diğer sınıflandırma algoritması k-EYK sınıflandırma algoritmasıdır. İyi k-EYK, ortalama k-EYK, kaba k-EYK, kosinüs k-EYK, kübik k-EYK, ağırlıklı k-EYK ve artırılmış k-EYK olmak üzere yedi farklı sınıflandırma yöntemi uygulandı. İyi k-EYK sınıflandırma yöntemi %96,2 ile en yüksek sınıflandırma başarısını veren yöntem olmuştur.

Dört farklı algoritmayla toplamda 20 farklı makine öğrenmesi yöntemi ile sınıflandırma gerçekleştirilmiştir. Bu 20 farklı sınıflandırma yöntemi içerisinde en başarılı yöntem %99,6 başarı oranı ile ikinci dereceden DVM (Quadratic SVM), sınıflandırma yöntemi olduğu söylenebilir. Elde edilen sonuçlar $\pm 0,77$ hata toleransı ile doğruluk göstermektedir. Çalışmanın en büyük avantajı yüksek sınıflandırma başarı oranı ile doğru bir sınıflandırma yapabilesidir. Çalışmanın dezavantajı 448 görüntü kullanılmış olmasıdır. Bu veri sayısı yetersiz gelebilecektir. Gelecekte daha çok kategori içeren daha büyük bir veri seti kullanılması ile çalışma daha ileri bir noktaya taşınabilir.

Teşekkür

Bu makale sorumlu yazar Mehmet Selim Kızgın'ın yüksek lisans tezinden üretilmiştir.

Kaynaklar

- [1] Sağiroğlu, Ş., Beşdok, E., Erler, M. Mühendislikte Yapay Zeka Uygulamaları-1: Yapay Sinir Ağları. Ufuk Yayıncılık, Kayseri, 2003.
- [2] Schalkoff RJ, Artificial Neural Networks, McGraw-Hill Inc., Singapore, 1997.
- [3] Acar, E., Çalışkan A., Sezgin N. Gabor Dalgacık Dönüşümü Tabanlı Yapay Sinir Ağı Modeli ile Zambak Yapağı İmgelerinden Pas Hastalıklarının Tespiti, Batman Üniversitesi Yaşam Bilimleri Dergisi Cilt1, Sayı 2 Batman, 2012.
- [4] Kılıç, E., Ecemiş, İN, İlhan, H.O. Narenciye Ağaç, Yaprak Hastalıklarının Evrişimli Sinir Ağları ile Sınıflandırılması. 5. International Symposium on Multidisciplinary Studies and Innovative Technologies. 2021.
- [5] Elmas, B., Evrişimli Sinir Ağları ile Mantar Görüntülerinden Mantar Türlerinin Transfer Öğrenme Yöntemiyle Tanımlanması. Süleyman Demirel Üniversitesi Fen Bilimleri Enstitüsü Dergisi, 2021, Cilt 25, Sayı 1, 74-88.
- [6] Baranwal, S., Khandelwal, S., & Arora, A., Elma yaprakları hastalık tespiti için derin öğrenme evrişimli sinir ağı. Bilim, teknoloji ve yönetimde sürdürülebilir bilgi işlem (SUSCOM) üzerine uluslararası konferansın bildirilerinde, Amity Üniversitesi Rajasthan, Jaipur- Hindistan, Şubat 2019.
- [7] Koklu, M., Unlersen, MF, Ozkan, I. A., Aslan, M. F., & Sabanci, K., A CNN-SVM study based on selected deep features for grapevine leaves classification. Measurement, 188, 110425, 2022.
- [8] Ganguly, S., Bhowal, P., Oliva, D., & Sarkar, R., BLeafNet: a Bonferroni mean operator based fusion of CNN models for plant identification using leaf image classification. Ecological Informatics, 69, 101585, 2022.
- [9] Hoşgören, H. (2004). Güneydoğu Anadolu Bölgesinde Yetişen Onobrychis Miller Türlerinin Sistematik, Morfolojik ve Kromozom Sayısı Bakımından Araştırılması. D.Ü. Fen Bilimleri Enstitüsü. Doktora Tezi. Diyarbakır
- [10] Argüeso, D., Picon, A., Irusta, U., Medela, A., San-Emeterio, MG, Bereciartua, A., & Alvarez-Gila, A., Tarlada çekilen görüntüleri kullanarak bitki hastalıklarının sınıflandırılması için Few-Shot Learning yaklaşımı. Tarımda Bilgisayar ve Elektronik, 175 , 105542, 2020.
- [11] Jogetkar, RN ve Tiwari, N., Bitki yaprağı hastalığının tanımlanması ve teşhisi için derin öğrenme tekniklerinin gözden geçirilmesi. Bilgi İşlem ve İletişimde Akıllı Trendler: SmartCom 2020 Bildiri Kitabı, 435-441, 2021.
- [12] Elçi, Ş., Anadolu'nun Önemli Yem Bitkilerinden Birkaç Korunga (Onobrychis) Türü Üzerinde Bazı Morfolojik ve Biyolojik Araştırmalar. Ziraat Fak. Yem Bit. Yem Bit. Yem Bit. Mera Kürsüsü, Ankara, 1994.
- [13] Hedge, IC, Onobrychis Adans. in Davis, P.H., Flora of Turkey and the East Aegean Island, Vol. 3, s. 560-589, Edinburg Univ. Pres, UK, 1970.
- [14] Zhou, S. R., Yin, J. P., Zhang, JM, Local binary pattern (LBP) and local phase quantization (LBQ) based on Gabor filter for face representation, Neurocomputing, Vol. 116, 260-264, 2013.
- [15] Zhao, G., Pietikainen, M., Dynamic texture recognition using local binary patterns with an application to facial expressions, IEEE transactions on pattern analysis and machine intelligence, 29 (6), 915-928. Ph.D. Thesis, North Dakota State University, North Dakota, USA, 2008.
- [16] Kaya, Y., Uyar, M., Tekin, R., Yildirim, S., 1D-Local Binary Pattern Based Feature Extraction for Classification of Epileptic EEG Signals. Applied Mathematics and Computation. 243. 10.1016/j.amc.2014.05.128, 2014.
- [17] Maenpaa, T., Pietikainen, M., Texture analysis with local binary patterns. Handbook of Pattern Recognition and Computer Vision. 10.1142/9789812775320_0011, 2005.
- [18] Goering, R., Matlab edges closer to electronic design automation world, EE Times, 2004.
- [19] Nabiyevev, VV, Yapay Zeka: İnsan-Bilgisayar Etkileşimi, Seçkin Yayınevi .Ankara, 2010.
- [20] Anjanappa, D., Raja, D., Venugopal. K., Face Recognition Based on Windowing Technique Using DCT, Average Covariance and Artificial Neural Network. 335-342. 10.1109/ICIIBMS.2018.8549981, 2018.
- [21] Bouchrika, I., Harrati, N., Ladjailia, A., Khedairia, S., Age Estimation from Facial Images based on Hierarchical Feature Selection. 10.1109/STA.2015.7505156, 2015.
- [22] Hedge, I.C., Onobrychis Adans. in Davis, P.H., Flora of Turkey and the East Aegean Island, Vol. 3, s. 560-589, Edinburg Univ. Pres, UK., 1970.
- [23] Yılmaz, A., Yapay Zeka(5.Baskı). Ankara: Kodlab Yayıncılık, 2017.
- [24] Maenpaa,T., Pietikaainen M., Texture Analysis with Local Binary Patterns, University of Oulu, 2004. Jasitha, P., Dileep, MR ve Divya, M., Venasyon bazlı bitki, GoogLeNet ve VGG kullanarak sınıflandırmadan çıkar. 2019'da 4. Uluslararası Elektronik, Bilgi, İletişim ve Teknoloji Son Trendler Konferansı, mayıs 2019, (RTEICT) (s. 715-719). IEEE