

Atıf İçin: Aydın, T., Bozkurt, F. ve Karademir, R.M. (2025). Makine Öğrenimi ve Derin Öğrenme Metotları Kullanılarak Hurma Meyvesi Türü Tespiti. *İğdır Üniversitesi Fen Bilimleri Enstitüsü Dergisi*, 15(2), 382-395.

To Cite: Aydın, T., Bozkurt, F. & Karademir, R.M. (2025). Date Palm Fruit Type Detection Using Machine Learning and Deep Learning Methods. *Journal of the Institute of Science and Technology*, 15(2), 382-395.

Makine Öğrenimi ve Derin Öğrenme Metotları Kullanılarak Hurma Meyvesi Türü Tespiti

Tolga AYDIN¹, Ferhat BOZKURT¹, Rüstem Muhammed KARADEMİR^{1*}

Öne Çıkanlar:

- Tür sınıflandırması
- Derin öğrenme
- Model karşılaştırması

Anahtar Kelimeler:

- Meyve görüntülerinin sayısal analizi
- Makine öğrenimi
- ESA

ÖZET:

Meyve ve sebze görüntülerinin analizi endüstriyel tarımda tanımlama ve sınıflandırma süreçlerinde önemli bir rol oynamaktadır. Hızla artan insan nüfusu ve bu doğrultuda yükselen tüketim miktarı analiz ve sınıflandırma işlemlerini otomatikleştirmeyi mecbur kılmaktadır. Geleneksel olarak uzmanlar tarafından yapılan görsel inceleme ve yorumlama yöntemleriyle gerçekleştirilen bu analiz süreci zaman alıcı ve öznel olabilir. Ancak son yıllarda derin öğrenme tekniklerinin gelişimi meyve ve sebze görüntülerinin otomatik analizinde büyük bir potansiyel sunmaktadır. Bu çalışmada klasik makine öğrenimi ve derin öğrenme modellerinin hurma meyvesinin türlerine ayrılması için nasıl kullanılabileceği incelenmektedir. Lojistik Regresyon, GaussianNB, KNN, SVM ve Random Forest gibi yaygın olarak kullanılan makine öğrenimi modelleri ile Evrişimli Sinir Ağları (ESA) temeline dayanan CNN, RNN ve ANN gibi derin öğrenme modelleri kullanarak, hurma meyvesine ait görüntü tabanlı olmayıp morfolojik özelliklerden oluşan sayısal bir veri seti üzerinde eğitimler ve testler yapılmıştır. Ayrıca deneysel bir çalışma gerçekleştirilerek bu farklı makine öğrenimi ve derin öğrenme modellerinin performansı karşılaştırılmıştır. Yapılan testlere göre en yüksek %92.44 doğruluk oranı RNN modeli ile elde edilmiştir. Sonuç olarak makine öğrenimi ve derin öğrenme tabanlı modellerin meyve görüntülerinin analizi alanında önemli bir potansiyele sahip olduğu görülmektedir. Bu teknikler sınıflandırma aşamasında yüksek doğruluk ile süreçlere hız kazandırarak endüstriyel tarımın gelişimine katkı sağlayabilir.

Date Palm Fruit Type Detection Using Machine Learning and Deep Learning Methods

Highlights:

- Species classification
- Deep learning
- Model comparison

Keywords:

- Digital analysis of fruit images
- Machine learning
- CNN

ABSTRACT:

The analysis of fruit and vegetable images plays a crucial role in the detection and classification processes in industrial agriculture. The rapidly growing global population, coupled with the subsequent rise in consumption, necessitates the automation of these processes. This analysis process, traditionally carried out by experts through visual inspection and interpretation, can be subjective and time-consuming. However, the development of deep learning techniques in recent years offers great potential in the automatic analysis of fruit and vegetable images. In this study, we investigate how classical machine learning and deep learning models can be used to classify date palm fruit into species. Using widely used machine learning models such as Logistic Regression, GaussianNB, KNN, SVM and Random Forest, and deep learning models such as CNN, RNN and ANN based on Convolutional Neural Networks (CNN), training and testing were performed on a numerical dataset consisting of non-image based morphological features of date palm fruit. An experimental study was also conducted to compare the performance of these different machine learning and deep learning models. According to the tests, the highest accuracy rate of 92.44% was obtained with the RNN model. In conclusion, machine learning and deep learning-based models have significant potential in the field of fruit image analysis. These techniques can contribute to the development of industrial agriculture by accelerating the processes with high accuracy in the classification phase.

¹ Tolga AYDIN (Orcid ID: 0000-0002-8971-3255), Ferhat BOZKURT (Orcid ID: 0000-0003-0088-5825), Rüstem Muhammed KARADEMİR (Orcid ID: 0009-0003-7567-6961), Atatürk Üniversitesi, Bilgisayar Mühendisliği Bölümü, 25240, Erzurum, Türkiye

***Sorumlu Yazar/Corresponding Author:** Rüstem Muhammed KARADEMİR, e-mail: rustemkarademir@gmail.com

GİRİŞ

Tarım, dünya nüfusunun temel gıda ihtiyacını karşılayan kritik bir sektördür. (Godfray ve ark., 2010). Hızla artan dünya nüfusu ve bu artışa karşılık gıda ihtiyacının karşılanabilmesi adına endüstriyel tarımda verimli ve hızlı üretime olan ihtiyaç her geçen gün artmaktadır. Ancak iklim değişikliği, toprak verimliliğinin azalması, bitki hastalıkları ve zararlılar gibi faktörler tarımsal üretim üzerinde baskı yaratmaktadır (Selçuk ve Gülümser, 2023). Bu baskıların yönetilmesi ve üretim süreçlerinin optimize edilmesi, özellikle yüksek besin değerine sahip meyvelerin yetiştirilmesi gibi endüstriyel tarım uygulamalarında büyük önem taşımaktadır. Yetiştirilen yüksek miktardaki meyvelerin kendi içinde birçok türü vardır. Bu farklı türlerin sınıflandırılması sektör açısından önem arz etmektedir. Zira farklı türler arasındaki farklı kalitedeki ürünlerin finansal açıdan getirisi de farklı olmaktadır. Meyve sınıflandırma işlemleri önceki dönemlerde uzmanlar tarafından yapılırken, günümüzde yapılan yüksek miktardaki üretilere insan hızının yetişemeyeceği açıktır. Yapay zekâ tekniklerinin yaygın olarak benimsenmesi ve kullanılması, tarım süreçlerini daha verimli ve etkili bir şekilde yönetmeye olanak tanıyarak tarım endüstrisindeki otomasyonun gelişimini hızlandırmıştır (Jordon ve Mitchell, 2015; Paçal ve Kunduracıoğlu, 2024).

Yapay zekâ, tarım süreçlerinde verimliliği artırarak bitki hastalıklarının tespiti, hasat planlaması ve ürün sınıflandırması gibi birçok alanda yenilikçi çözümler sunmaktadır (Savary ve ark., 2019). Bu bağlamda, makine öğrenimi (ML) ve derin öğrenme (DL) teknikleri, tarımda geniş bir uygulama yelpazesi ile dikkat çekmektedir. Örneğin, tarım arazilerinde mobil cihazlarla çekilen fotoğraflar aracılığıyla bitki hastalıklarının otomatik tespiti (Picon ve ark., 2019), mango, badem ve elma meyvelerinin sınıflandırması için Faster R-CNN adlı derin öğrenme tabanlı bir nesne algılama modelinin geliştirilmesi (Bargoti ve Underwood, 2017), bitki kökleri ile yapraklarına ait görüntüleri zaman serisi ve GAN (Çekişmeli Üretici Ağlar) kullanarak bitki büyümesi ve gelişiminin tahmini (Yasrab ve ark., 2021) gibi birçok görevde bu teknolojilerin kullanımı başarılı sonuçlar elde edilmesini sağlamıştır. Hızlı ve doğruluğu yüksek bir oranda çalışabilen çeşitli makine öğrenimi ve derin öğrenme modelleri ile yapılan sınıflandırma çalışmalarında, meyvelerin görüntüleri üzerinden elde edilen veriler kullanılmaktadır. Meyve sınıflandırma çalışmaları ayrıca, biyolojik çeşitliliği anlamak ve korumak amacıyla yapılan araştırmaların önemli bir bileşenini de oluşturmaktadır (Van Zonneveld ve ark., 2023). Phoenix cinsine ait palmye familyasından olan hurma, Orta Doğu, Kuzey Afrika ve Güneydoğu Asya'da yetişir. Hurma besin değeri yüksek, lezzetli ve dayanıklı bir meyvedir, bu özellikleri ile üretimi her geçen gün artmaktadır (Elleuch ve ark., 2008). Ancak, hurma yetiştiriciliği de bitki hastalıkları (Bayoud hastalığı), zararlılar (hurma kırmızı palmye böceği), iklim değişikliğinin etkileri (kuraklık ve aşırı sıcaklar) ve yetersiz sulama uygulamaları gibi tehditlerle karşı karşıyadır (Abdullah ve ark., 2010; Allbed ve ark., 2017). Bu tehditler, hurma üretiminde önemli kayıplara neden olabilir ve sektörün ekonomik sürdürülebilirliğini tehlikeye atabilir. Hurma yetiştiriciliğinde karşılaşılan bu zorlukların üstesinden gelmek için yapay zekâ destekli çözümler geliştirilebilir. Görüntü işleme ve derin öğrenme teknikleri, hurma ağaçlarının ve meyvelerinin incelenmesinde kullanılarak hastalıkların erken tespiti, zararlıların tespiti, meyve kalitesinin değerlendirilmesi ve sınıflandırılması gibi görevlerde kullanılabilir. Meyve sınıflandırması ve kalite kontrolü için yapay zekâ tabanlı sistemler geliştirilebilir. Bu sistemler, meyvelerin boyut, renk, şekil ve diğer özelliklerini analiz ederek otomatik olarak sınıflandırabilir ve bu sayede insan hatası en aza indirilebilir. Farklı hurma türlerine ait örnekler üzerinde yapılan bu sınıflandırma çalışmasında, mevcut olan çeşitli makine öğrenimi teknikleri incelenmekte ve bu tekniklerin hurma meyvesi türlerini doğru bir şekilde tanımlama yetenekleri değerlendirilmektedir. Meyve sınıflandırma sürecinde kullanılan morfolojik özellikler, hurma türlerini ayırt etme konusunda sağlam bir temel oluşturur. Görüntü işleme teknikleri uygulanarak Köklü ve diğerleri tarafından elde

edilen morfolojik özelliklere ait sayısal veriler kullanılarak sınıflandırma çalışması gerçekleştirilmiştir (Köklü ve ark., 2021).

Son yıllardaki gelişmeler ile derin öğrenme algoritmaları büyük veri setlerinden öznetelikler öğrenerek nesneleri algılama, lokalizasyon, sınıflandırma ve segmentasyon gibi görevleri başarıyla gerçekleştirebilir hale gelmiştir. Yapay zekâ destekli sistemlerdeki bu gelişmeler ile meyve görüntülerinden çıkarılacak bilgileri, daha hızlı ve doğruluğu yüksek seviyede anlamlı hale getirmek mümkün olmuştur. Derin öğrenmeye artan ilgi derin öğrenmenin bir alt dalı olan evrişimli sinir ağlarını (CNN) segmentasyon ve sınıflandırma için kullanılan en yaygın araç haline getirmiştir (Mikołajczyk ve Grochowski, 2018). CNN algoritması görseller gibi matrislerden oluşan verileri işlemek için LeCun ve diğerleri tarafından kullanılan evrişimli bir yapay sinir ağıdır (LeCun ve ark., 1989). Yaygın kullanım alanları özellikle görüntüler üzerinde sınıflandırma ve segmentasyon, resim ve video tanıma ile doğal dil işleme olarak sıralanabilir.

İlgili Çalışmalar

Literatür incelendiğinde son yıllarda meyve sınıflandırma işlemlerinde görüntü işleme, makine öğrenimi ve yapay sinir ağı teknolojilerinin kullanımı artmıştır. Gulzar ve arkadaşları (2022), sekiz farklı hurma sınıfını içeren bir veri seti oluşturmuş ve MobileNet-V2 mimarisine dayalı bir model geliştirmişlerdir. Bu model ile %99 doğruluk oranına ulaşmış ve AlexNet, VGG16, InceptionV3, ResNet gibi diğer modellerle karşılaştırıldığında en yüksek performansı göstermiştir. Nasiri ve arkadaşları (2019), sağlıklı hurma meyvelerini kusurlu olanlardan ayırt etmek için VGG-16 mimarisine sahip bir CNN (Evrişimsel Sinir Ağı) modeli kullanmışlardır. Elde edilen deneysel sonuçlar, bu modelin hurma meyvesi görüntülerini ayırt etmede %96.98 doğruluk oranına ulaştığını göstermiştir. Joseph ve arkadaşları (2021), meyve sınıflandırması için TensorFlow tabanlı konvolüsyon, havuzlama ve tam bağlantılı katmanlardan oluşan bir CNN modeli geliştirerek, 131 farklı meyve ve sebze sınıfını içeren Fruits 360 veri kümesini kullanmışlardır. 50 tur sayısıyla eğitilen model ile test aşamasında %94.35 doğruluk oranı elde etmiştir. Kunduracıoğlu ve Paçal (2024) tarafından yapılan çalışmada, derin öğrenme tekniklerinin üzüm yapraklarının sınıflandırılması ve hastalıklarının teşhisindeki etkinliği incelenmiştir. Çalışmada, 5 farklı üzüm türüne ait 500 adet ve 4 farklı üzüm hastalığını içeren 4062 adet yaprak görüntüsünden oluşan iki ayrı veri seti kullanılarak, 14 farklı CNN ve 17 farklı Vision Transformer (ViT) tabanlı model geliştirilmiştir. CNN tabanlı VGG-16, VGG-19 ve Densenet121 modelleri ile ViT tabanlı SwinV2-Base-Win8, DeiT3-Base, MaxViT-Small ve MaxViT-Large modellerinde %100 doğruluk oranı elde edilmiştir.

Hasan ve Sattar (2021) tarafından gerçekleştirilen çalışmada, çeşitli hurma türlerini sınıflandırmak amacıyla MobilNet_V1 tabanlı bir CNN modeli geliştirilmiştir. 6 farklı hurma türü ve 2246 görüntüden oluşan veri seti üzerinde bu model ile %82.67 doğruluk oranı sağlanmıştır. Zhang ve arkadaşları (2019), 13 katmanlı bir CNN modeli tasarlayarak meyve kategorilerini belirlemeyi hedeflemişlerdir. 3600 görüntü ve 18 sınıftan oluşan bir veri tabanını kullanmışlar ve verileri %50 eğitim, %50 test olarak ayırmışlardır. 30 tur sayısıyla eğitilen modelin genel doğruluk oranı %94.94 olarak tespit edilmiştir. Büyükarıkan ve Ülker (2020) çalışmalarında, aydınlatma özelliğini kullanarak CNN modelleri ile meyve türü sınıflandırması gerçekleştirmiştir. Çalışmalarında, farklı aydınlatma koşullarının meyve tanıma performansı üzerindeki etkileri incelenmiş ve bu etkilere bağlı olarak modellerin doğruluğu değerlendirilmiştir. 29 farklı meyve türünün 12 farklı renk sıcaklığındaki ön profil görüntülerinden oluşan veri setinde, görüntü işleme ve örnek çoğaltma teknikleri uygulayarak 1479 görüntü ile yaptıkları bu çalışmada VGG19 mimarisi ile %86.49 doğruluk oranı elde edilmiştir. Alsirhani ve arkadaşları (2023), hurma meyvesi türlerini tespiti için yaptıkları çalışmada 5 aşamalı transfer öğrenme yaklaşımını kullanan DenseNet201 modeli ile %97.21'lik eğitim ve %95.21'lik test doğruluğu sonucunu elde

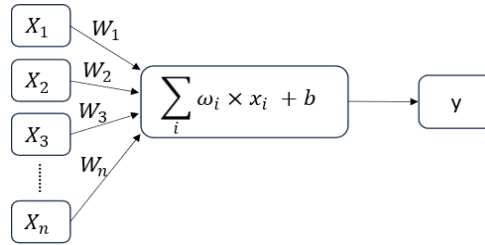
etmiştir.

Bu çalışmada hurma meyvesine ait morfolojik özellikler üzerinde makine öğrenimi ve derin öğrenme tabanlı CNN, RNN ve ANN yaklaşımlarının kullanılması ele alınmaktadır. 7 farklı hurma türünü sınıflandırabilen makine öğrenmesi ve derin öğrenme tekniklerinin aynı veri kümesi üzerinde performans değerlendirmesi ve karşılaştırması üzerinde durulacaktır. Elde edilen sonuçlar ile hurma meyvesinin sınıflandırması için en verimli yöntemin tespiti için çalışılacaktır.

Çalışmanın geri kalan bölümlerinde; ikinci kısımda çalışmada kullanılan materyal ve metotlar hakkında bilgi verilecek ve bu metotlar ve algoritmalar detaylı olarak anlatılacaktır, üçüncü kısımda çalışmada elde edilen deneysel sonuçlar aktarılacaktır, dördüncü kısımda ortaya çıkan deneysel sonuçlar tartışılacak ve sonuçlar sunulacaktır.

MATERYAL VE METOT

Yapay sinir ağları ile insan beynindeki öğrenme sürecine benzer bir yaklaşım, ilk kez 1943 yılında insan beynindeki hücrelerin yapısının matematiksel modeli ile oluşturulmuştur (McCulloch ve ark., 1943).



Şekil 1. Yapay sinir ağını oluşturan hücrelerin matematiksel modeli (Kaftan, 2010)

Şekil 1’de makine öğrenmesinde kullanılan perceptron sinir hücresine ait bir temsil gösterilmiştir. X_1, X_2, X_3 bağımsız değişkenleri (özellikleri), $\omega_1, \omega_2, \omega_3$ bağımsız değişkenler ait ağırlık değerlerini b bias değerini y ise çıktıyı temsil eder. Her epochta ağırlıklar ve bias değeri yeniden belirlenerek en iyi sonuç elde edilmeye çalışılır, bu adımlar öğrenme süreci olarak adlandırılır (Kaftan, 2010).

Makine öğrenimi, bilgisayarlı sistemlerde tanımlanan görevleri öğrenme ve öğrendiği bilgiler ile eğitimi geliştirme kapasitesi kazandıran bir yapay zekâ alt dalıdır. Bu yöntem, algoritmaların geniş veri setlerinden desenler ve ilişkiler çıkarmasına dayanır. Bilgisayarlar, belirli bir görevi gerçekleştirmek için verilerden öğrenir, deneyim kazanır ve bu deneyimi kullanarak gelecekteki kararlarını ve tahminlerini iyileştirir (Ranjana ve ark., 2023). Sınıflandırma, regresyon, kümeleme ve desen tanıma gibi çeşitli uygulama alanlarında kullanılan makine öğrenimi, bilgisayarların verilerden öğrenmesini ve kendi kendini geliştirmesini sağlar (Mitchell, 1997; Bishop, 2006; Uğuz, 2019).

Makine öğrenimi için yaygın olarak kullanılan bazı popüler sınıflandırma algoritmaları arasında Lojistik Regresyon (LR), Gaussian Naive Bayes (GNB), En Yakın Komşu (KNN), Destek Vektör Makineleri (SVM) ve Rastgele Ağaç (RF) gibi modeller bulunur. Çalışmamızda da sıklıkla kullanılan bu modellerden faydalanılmıştır.

Çok katmanlı ve geri beslemeli yapay sinir ağı modellerinin ortaya çıkması ile lineer olmayan denklemlerin çözümünün önü açılmıştır (Eberhart ve Kennedy, 1995). Çok katmanlı sinir ağı modellerinin oluşmasıyla, giriş katmanlarının miktarı artırılıp daha verimli çıktılar verebilen Convolutional Neural Network (CNN) geliştirilmiştir. Evrişimli sinir ağı modellerinde gizli ara katmanlar bulunmakta ve bu katmanlarda bağımsız değişkenlere farklı ağırlıklar verilerek başarı oranı arttırılmaktadır (Pan ve ark., 2000).

Derin öğrenme, yapay sinir ağlarının daha derin ve karmaşık yapılarını kullanarak karmaşık veri modellerini öğrenme yeteneği sunan bir makine öğrenimi yaklaşımıdır. Bu yöntem, büyük veri setlerinden özniteliklerin otomatik olarak öğrenilmesini ve karmaşık ilişkilerin modelin içinde temsil

edilmesini sağlar. Derin öğrenme, birçok alanda başarıyla uygulanmıştır (LeCun ve ark., 1998). Özellikle görüntü işleme, doğal dil işleme, konuşma tanıma, otomatik sürüş, hastalık teşhisi ve sınıflandırma gibi teknolojilerde önemli ilerlemeler kaydedilmiştir (Szegedy ve ark., 2015; Pacal ve ark., 2024; Kunduracioglu, 2024). Bunun yanı sıra, tarımda mahsul sınıflandırma ve hastalık tespiti (Kamilaris ve ark., 2018), sağlıkta hastalık teşhisi ve görüntüleme analizleri (Esteva ve ark., 2017), eğitimde kişiselleştirilmiş öğrenme sistemleri (Holmes, 2019), finans sektöründe dolandırıcılık tespiti ve risk analizi (Ryman-Tubb ve ark., 2018), enerji sektöründe akıllı şebeke yönetimi (Sağ ve Tuğcu, 2022) gibi pek çok farklı uygulama alanında etkili bir şekilde kullanılmaktadır. Bu çeşitlilik, derin öğrenme yöntemlerinin disiplinler arası uygulama potansiyelini ortaya koymaktadır. Derin öğrenmenin bu başarısı, temel özelliklerinden kaynaklanmaktadır. Derin öğrenme, veriden özellikleri otomatik olarak öğrenir. Giriş verisi ile çıktı arasındaki ilişkiyi modellemek için hiyerarşik öznitelikleri keşfeder ve bu sayede, özellik çıkarma adımını ortadan kaldırır ve öğrenme sürecini daha verimli hale getirir. Derin öğrenme modelleri birçok katmanın birleşiminden oluşur. Katmanlar sayesinde oluşan derinlik, modelin daha yüksek seviyede çıkarım yapmasını sağlar ve daha karmaşık problemlerin çözülmesini mümkün kılar (Zeiler ve Fergus, 2013). Derin öğrenme modelleri veri içerisindeki özellikleri otomatik olarak çıkarabilir. Geleneksel yöntemlerde özelliklerin el ile belirlenmesi gerekebilirken, derin öğrenme modelleri bu süreci otomatikleştirir, insan müdahalesini azaltır ve daha genel geçerli özelliklerin keşfedilmesini sağlar. Ayrıca büyük veri setlerinden yararlanarak daha iyi sonuçlar elde etme potansiyeline sahiptir.

Derin öğrenme tabanlı modellerin, klasik makine öğrenimi modellerine kıyasla temel avantajlarından biri, ham veri üzerinde çalışarak öznitelik çıkarma işlemini otomatikleştirmeleridir. Ancak bu çalışmada kullanılan veri setinde öznitelikler manuel olarak çıkarılmış olmasına rağmen, öznitelik sayısının fazla olması ve öznitelikler arasındaki karmaşık ilişkilerin varlığı, bu ilişkilerin daha yüksek boyutlu temsillerle ifade edilmesi ihtiyacını doğurmuştur. Derin öğrenme modellerinin sahip olduğu derin katman yapıları, özniteliklerin daha soyut ve güçlü bir şekilde temsil edilmesini sağlamaktadır (Goodfellow, 2016). Bu durum, çalışmada kullanılan modellerin daha yüksek sınıflandırma doğruluğuna ve genelleme kabiliyetine ulaşmasına olanak tanımaktadır. Ayrıca çalışmada deneysel bir amaçla derin öğrenme ve klasik öğrenim modellerinin başarısı kıyaslanmıştır.

Derin öğrenme için yaygın olarak kullanılan bazı popüler algoritma ve modeller arasında Evrişimli Sinir Ağları (Convolutional Neural Networks-CNN), Tekrarlayan Sinir Ağları (Recurrent Neural Networks-RNN) ve Yapay Sinir Ağları (Artificial Neural Networks-ANN-MLP) gibi modeller bulunur.

LR - Lojistik Regresyon

Makine öğrenimi alanında yaygın olarak kullanılan bir sınıflandırma yöntemi olan lojistik regresyon, bağımsız değişkenler ile bağımlı değişkenler arasındaki ilişkiyi kurulan korelasyonlar yardımıyla modellemek amacıyla kullanılan istatistiksel bir yöntemdir. Genellikle ikili sınıflandırma problemlerinde tercih edilen lojistik regresyon, doğrusal regresyonun bir genişlemesi olarak düşünülebilir. Ancak, lojistik regresyonun önemli bir farkı, çıktının bir olasılık değeri olmasıdır (Bishop, 2006). Ayrıca doğrusal regresyonda, bağımlı değişken sürekli iken, lojistik regresyonda bağımlı değişkenin kategorik olması beklenir.

Lojistik regresyon, sigmoid fonksiyonunu kullanarak çıktıyı $[0, 1]$ aralığına sıkıştırır. Bu sayede, çıktıyı olasılık olarak yorumlamak mümkün olur. Bu yöntem, özellikle hastalık teşhisi, spam filtreleme, finansal analiz ve birçok diğer alanlarda kullanılmaktadır. Bu çalışmada hurma türlerinin sınıflandırılması için lojistik regresyon algoritması kullanılmıştır.

GNB - Gaussian Naive Bayes

Naive Bayes sınıflandırıcıları, Bayes teoremine dayalı bir sınıflandırma yöntemidir ve olasılık temelli bir yaklaşım benimserler. Gaussian Naive Bayes, bu sınıflandırma yöntemlerinden biridir ve özellikle sayısal veri setlerinde kullanılır. "Naive" (saf) olasılık varsayımını benimser, yani her özellik birbirinden bağımsızdır. Bu yaklaşım, modelin her özelliğin sınıf etiketi üzerindeki etkilerini birbirinden bağımsız olarak değerlendirdiği anlamına gelir. Diğer yandan özelliklerin veri setindeki dağılımını normal (Gaussian) olarak varsayar. Her sınıfa ait özellik değerlerinin normal bir dağılımı olduğu varsayımıyla, sınıflandırma işlemini gerçekleştirir. Bu yöntem, özellikle sayısal değerlerle ifade edilen özniteliklere sahip veri setlerinde iyi performans gösterir (Bishop, 2006). Bu çalışmada hurma türlerinin sınıflandırılması için Gaussian Naive Bayes algoritması kullanılmıştır.

KNN - En Yakın Komşu

K-Nearest Neighbors (KNN), klasik makine öğrenimi modellerinde sıkça kullanılan sınıflandırma ve regresyon algoritmasıdır. Özellikle veri madenciliği alanında yaygın olarak kullanılan bir tekniktir. KNN yaklaşımı, veri setindeki benzer örneklerin birbirine yakın olduğu varsayımına dayanır ve yeni örneklerin sınıflandırılması veya tahmin edilmesi için bu benzer örneklerin kullanılmasını amaçlar.

KNN'nin temel prensibi oldukça sade ve basittir. Sınıflandırma durumunda, bir örneği sınıflandırmak için çevresindeki k en yakın komşusunu kullanır. Öklidyen mesafe veya diğer benzer mesafe metrikleri kullanılarak bu benzerlik ölçülür. En yakın komşuların etiketleri, sınıflandırılacak örnek için bir tahmin oluşturmak için kullanılır. Bu algoritmanın gücü, esnek ve adaptif olmasıdır. Ancak, büyük veri setlerinde ve yüksek boyutlu veri uzaylarında hesaplama maliyeti artabilir. Bunun önüne geçebilmek için komşuların sınıf etiketlerine göre ağırlıklandırılmasını içeren ağırlıklı yaklaşım önerilmiştir (Gazalba ve Reza, 2017). Bu çalışmada hurma türlerinin sınıflandırılması için KNN algoritması kullanılmıştır.

SVM - Destek Vektör Makineleri

Destek Vektör Makineleri (SVM), makine öğrenimi alanında sınıflandırma problemleri için kullanılan bir algoritmadır. Bu algoritmanın temeli, sınıfları birbirinden ayıran en iyi hiper düzlemi bulmaya dayanır. Sınıfları ayıran bir sınır çizgisi kullanarak, bu sınır üzerinde bulunan verileri "destek vektörleri" olarak adlandırır (Burges, 1998).

SVM, özellikle doğrusal olarak ayrılabilir olmayan veri setlerinde, yani veri noktalarının birbirine çok yakın olduğu ve basit bir çizgiyle ayrılmasının zor olduğu durumlarda etkilidir. Algoritma, veri setindeki sınıflar arasındaki en geniş marjı maksimize eden bir hiper düzlem seçer. Bu, modelin genelleme yeteneğini artırır ve yeni veri örneklerini daha doğru bir şekilde sınıflandırmasına olanak tanır. Öte yandan, doğrusal olarak ayrılabilen veri setlerinde de etkilidir. Çekirdek fonksiyonları kullanarak doğrusal olmayan sınırları da modellenebilir (Vapnik, 1999). Bu çalışmada hurma türlerinin sınıflandırılması için SVM algoritması kullanılmıştır.

RF - Rastgele Orman

Random Forest, rastgele orman algoritması birden çok karar ağacının bir araya getirilmesiyle oluşturulur ve bu sayede daha güçlü ve genelleştirilebilir bir model elde edilir. Random Forest'un temel prensibi, "karar ağacı" adı verilen basit modellerin bir araya getirilerek güçlü bir model oluşturmasıdır. Her bir karar ağacı, rastgele seçilen özellikler ve veri noktalarının alt kümeleri üzerinde eğitilir (Bozkurt, 2022). Modelin yapısı sayesinde her bir ağacın farklı bir bakış açısına ve veri alt kümesine sahip olmasını sağlar. Algoritmanın avantajlarından biri, overfitting eğilimini azaltmasıdır. Ayrıca, büyük ve karmaşık veri setlerinde yüksek performans gösterir. Random Forest, karışık veri setleri de dahil sınıflandırma ve geriye doğru regresyon için etkili bir yöntemdir (Lin ve Jeon, 2006).

CNN - Konvolüsyonel Sinir Ağı

Konvolüsyonel ya da evrişimli Sinir Ağı (Convolutional Neural Network-CNN), özellikle görüntü işleme ve desen tanıma gibi görsel veri analizi alanlarında kullanılan bir derin öğrenme modelidir.

Geleneksel yapay sinir ağlarına kıyasla, görüntü verilerinin özelliklerini daha iyi çıkarabilme özelliğiyle öne çıkar. CNN'ler, giriş verilerini evrişim (convolution) ve havuzlama (pooling) katmanlarıyla işlerler. Evrişim katmanları, giriş görüntüsü üzerinde bir veya daha fazla filtreyle (kernel) kayan pencere işlemi uygular ve bu sayede görüntünün farklı özelliklerini çıkarır. Havuzlama katmanları ise maksimum veya ortalama değeri alarak özellik haritalarını küçültür ve bu sayede hesaplama yükünü azaltır. Bu katmanların ardından, tam bağlantılı (fully connected) katmanlarla özellikleri sınıflandırma veya çıktı üretme işlemi gerçekleştirilir (Krizhevsky ve ark., 2012). Modelin en büyük artısı veri setlerindeki karmaşıklığı ve büyüklüğü işleyebilmesidir. Bu çalışmada hurma türlerinin sınıflandırılması için CNN algoritması kullanılmıştır.

RNN - Tekrarlayan Sinir Ağı

Sıralı veri yapıları üzerinde etkili olan bir derin öğrenme modelidir. Geleneksel sinir ağlarından farklı olarak, RNN'ler önceki zaman adımlarındaki bilgileri hatırlayabilir ve bu özelliği sayesinde sıralı veri setlerindeki bağıntıları öğrenme yeteneğine sahiptir. RNN'lerin temel bileşeni, döngüsel bağlantılara sahip olan bir nöron katmanıdır. Bu döngüsel bağlantılar, ağı önceki adımlarda elde ettiği bilgileri saklayarak gelecekteki adımlarda kullanabilmesine olanak tanır (Şeker ve ark., 2017). Bu çalışmada hurma türlerinin sınıflandırılması için RNN algoritması kullanılmıştır.

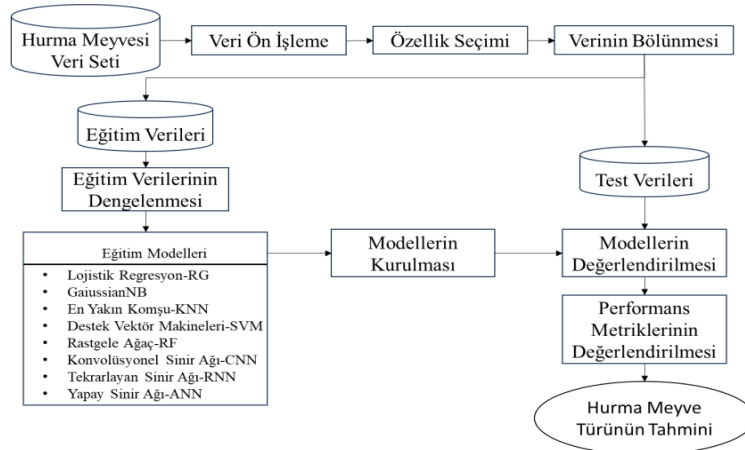
Materyal

Bu çalışmada, hurma meyvesi sınıflandırmasında farklı makine ve derin öğrenme tekniklerinin performansını karşılaştırmak amacıyla "Date Fruit Datasets" isimli Kaggle veri tabanından elde edilen veriler kullanılmıştır (Date Fruit Datasets, 2021). Bu veri kümesi, 7 farklı (Barhee, Deglet Nour, Sukkary, Rotab Mozafati, Ruthana, Safawi, Sagai) kategorik hurma türüne ait toplam 898 örnek ve bu örneklerle ait 34 farklı morfolojik özellikten oluşmaktadır. Şekil, renk ve boyut gibi dışsal özelliklerinden oluşan bu veri setindeki kategorilerin örnek görselleri Şekil 2'de gösterilmiştir.



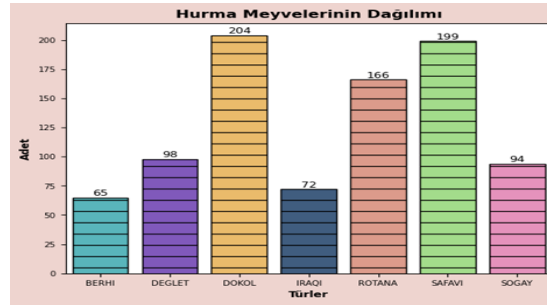
Şekil 2. Veri setindeki 7 farklı hurma türüne ait örnek görseller (Köklü ve ark., 2021)

Hurma meyvesi veri seti üzerinde 5 farklı makine öğrenimi ve 3 farklı derin öğrenme teknikleri kullanılarak hurma türü tahmini yapılmıştır. Sınıflandırma işlemi için önerilen modelin mimari yapısı Şekil 3'te verilmiştir. Şekilde gösterildiği üzere veri seti ön işleme adımlarından geçirilip özellik seçimi yapılarak eğitim ve test verilerine ayrılmıştır. Eğitim verileri ile modellerin eğitilmesinin ardından test verileri kullanılarak bu modellere ait performans metrikleri değerlendirilmiştir. Çalışmada kullanılan modellerin hurma meyve türü tahminleri üzerinden karşılaştırmalı analizleri yapılmıştır.



Şekil 3. Çalışmada kullanılan sınıflandırma mimarisi

Şekil 4'te çalışmada analiz edilen makine ve derin öğrenme modellerin eğitim ve test aşamalarında kullanılan veri setindeki toplam 898 örnekten oluşan 7 farklı hurma türünün kategorik dağılımı gösterilmiştir. Çapraz doğrulama (cross-validation) kullanılarak hurma türlerine ait verinin %75'i eğitim, %25'i test için ayrılmıştır.



Şekil 4. Çalışmada kullanılan hurma veri setinin kategorik dağılımı

Ön işleme

Veri ön işleme adımlarında, eksik veriler sorgulanmış, korelasyon matrisi üzerinden özellikler arasındaki ilişkiler incelenmiş, modellerin istikrarını ve performansını arttırmak adına ölçeklendirme işlemi uygulanmış ve kategorik veriler sayısal değerlere dönüştürülmüştür. Modelin performansını arttırmak için Boruta algoritması ile özellik seçimi (feature selection) yapılmıştır. Boruta algoritması, önemli özellikleri koruma ve modelin performansını artırma amacıyla tercih edilmiştir. Rastgele orman tabanlı yapısı sayesinde özelliklerin önem düzeyini etkin bir şekilde değerlendirir ve anlamlı olmayanları ayıklayarak aşırı uyumu (overfitting) engeller (Kursa ve Rudnicki, 2010). Bu yöntemle yapılan analiz sonucunda en yüksek önceliğe sahip 21 özellik belirlenmiştir.

Verinin eğitim ve test olarak bölünmesi aşamasında, sıralı dağılan veriler öğrenimi olumsuz etkileyebileceği için karıştırma (shuffle) işlemi uygulanmıştır. Veri setinde dengesiz bir dağılım olması modellerin performansı olumsuz etkileyebilir ve elde edilen sonuçların geçerliliğini sorgulatabilir. Bu sebeple azınlık sınıfındaki örnek sayısı artırılarak dengeli bir dağılım amaçlanmıştır. SMOTE (Synthetic Minority Over-sampling Technique) yöntemi, dengesiz veri setlerinde azınlık sınıfındaki örnek sayısını artırmak için kullanılan etkili bir yöntemdir (Chawla ve ark., 2002). Bu yöntem ile azınlık sınıfındaki her bir örnek için rastgele seçilen komşularını kullanarak sentetik örnekler oluşturulmuştur. Çapraz doğrulama (cross-validation) ile veriyi parçalara bölüp her seferinde farklı parçalar üzerinde test işlemi yapılarak modelin güvenilirliği artırılmıştır. Her denemede aynı sonuçları alarak rastgelelikten uzak tutarlı denemeler yapabilmek adına random state değeri 42 belirlenmiştir. Verileri normal dağılıma yaklaştırmak için standartlaştırma yapılmıştır.

Öğrenme sürecinde doğrulama kaybı ve test kaybı arasındaki fark arttığında, modelin aşırı öğrenme eğiliminde olduğu veya gürültüyü öğrendiği söylenebilir. Bu durumda, doğrulama hatası

artmaya başlar ise, eğitim durdurulur ve bir önceki adıma geri dönülür. Geri dönebilmek için her öğrenme adımında önceki adımın verileri saklanır. Bu tür durumları önlemek amacıyla çalışmada erken durdurma (early stopping) kullanılmıştır.

BULGULAR VE TARTIŞMA

Çalışmada makine öğrenimi ve derin öğrenme modelleri kullanılarak Date Fruit Datasets veri seti üzerinde deneysel analizler yapılmıştır. Analizlerde kullanılan geleneksel makine öğrenimi modellerine performansı arttırmak için model ayarlama (model tuning) uygulanmıştır. Model tuning ile makine öğrenimi modelinin hiperparametrelerini optimize ederek, modelin en iyi performansı göstermesini sağlamak ve genellemenin arttığı, aşırı uydurmanın azaldığı bir noktaya ulaşmak amaçlanır. Bu çalışmada hiperparametrelerin optimizasyonu için GridSearchCV algoritması kullanılmıştır. GridSearchCV algoritması ile belirlenen parametrelerin tüm kombinasyonları değerlendirilerek en iyi parametre değerlerinin tespit edilebilmesi amaçlanır (Pedregosa ve ark., 2011).

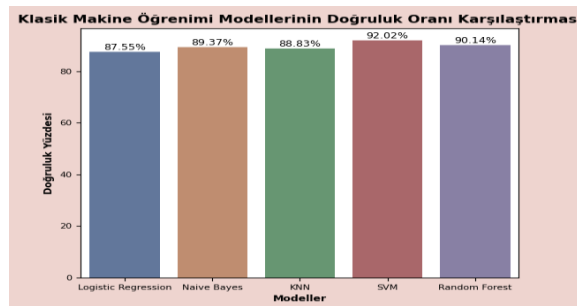
Çizelge 1. Çalışmada kullanılan klasik makine öğrenimi modellerine ait parametreler

Model	Parametreler
Lojistik Regresyon	solver=liblinear, C=100, penalty=l2
KNN	n_neighbors=2, p=1, weights=distance
SVM	C=10, gamma=scale, kernel=rbf
Random Forest	max_depth=20, max_features=10, min_samples_split=2, n_estimators=200

Çizelge 1’de gösterilen parametreler kullanılarak eğitilen modellerin performansları, çeşitli değerlendirme metrikleri aracılığıyla analiz edilmiştir. Bu değerlendirme metrikleri arasında doğruluk (Accuracy), kesinlik (Precision), duyarlılık (Recall) ve F1 score bulunmaktadır. Her bir metrik, modellerin tahmin performanslarını farklı açılardan değerlendirmekte ve modellerin güçlü ya da zayıf yönlerini daha net bir şekilde ortaya koymaktadır. Eğitilen klasik makine öğrenimi modellerine ait bu metrikler Çizelge 2’de gösterilmiştir.

Çizelge 2. Çalışmada kullanılan klasik makine öğrenimi modellerinin performans değerlendirmesi

Model	Doğruluk (%)	Kesinlik (%)	Duyarlılık (%)	F1 score (%)
Lojistik Regresyon	87.55	90.05	91.18	90.57
Gaussian Naive Bayes	89.37	82.39	84.13	82.67
KNN	88.83	85.82	86.71	85.88
SVM	92.02	88.18	89.89	88.86
Random Forest	90.14	85.12	86.12	85.52



Şekil 5. Klasik makine öğrenimi modellerine ait doğruluk grafiği

Çizelge 2’deki metrikler ve Şekil 5’teki grafikte gösterildiği üzere klasik makine öğrenimi modellerinden elde edilen sonuçlara göre SVM (Destek Vektör Makinesi) algoritması kullanılarak,

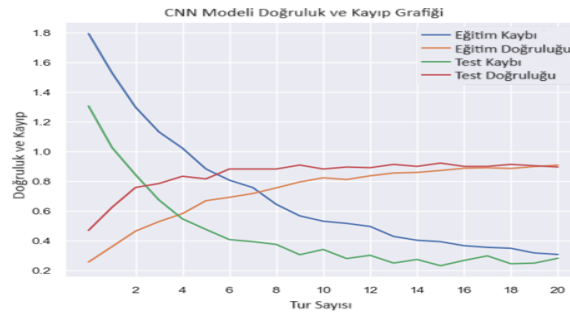
%92.02 ile en yüksek doğruluk ve %88.86 ile diğer modellere kıyasla yüksek bir F1 score elde edilmiştir. SVM, sınıflar arasındaki en geniş marjı maksimize eden hiper düzlemleri belirleyerek, özellikle dengesiz ve doğrusal olmayan veri setlerinde azınlıkta kalan sınıfları da doğru tahmin ederek, yüksek performans gösterebilme yeteneğine sahiptir. Ayrıca SVM’nin çekirdek (kernel) fonksiyonları, veri uzayını daha yüksek boyutlara taşıyarak doğrusal olmayan sınırları etkin bir şekilde modellemesini

sağlamıştır (Amaya-Tejera ve ark., 2024). Bu durum, SVM'nin diğer modellere kıyasla sınıflar arasındaki ayrımı hassasiyetle yapmasını sağlamıştır. Sonuç olarak, veri setinin az örnek içermesi, dengesiz bir dağılıma sahip olması ve özellik sayısının fazla olması SVM'nin diğer modellere kıyasla daha iyi bir performans göstermesinin başlıca sebepleri olarak yorumlanabilir.

Çizelge 3. Çalışmada kullanılan derin öğrenme modellerinin performans değerlendirilmesi

Model	Tekrar	Doğruluk (%)	Kayıp (%)	F1 score (%)
CNN	21/30	92.00	22.83	92.08
RNN	13/30	92.44	23.71	92.53
ANN	16/30	90.67	26.79	90.87

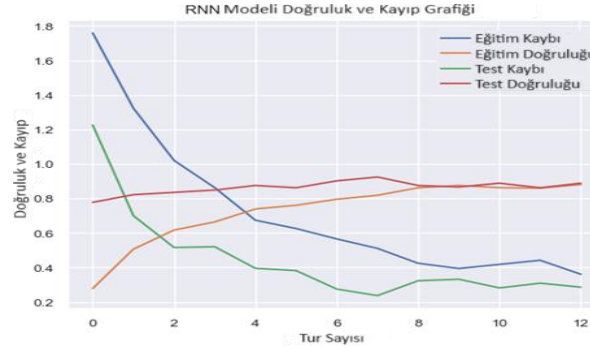
Bu çalışmada kullanılan derin öğrenme modelleri, CNN, RNN ve ANN olmak üzere üç farklı mimariden oluşmaktadır. Tüm modeller için giriş katmanında sırasıyla 256, 128, 64 ve 32 nöron bulunan dört katmanlı bir yapı tercih edilmiştir. Aşırı öğrenme (overfitting) riskini azaltmak amacıyla her katmanda sırasıyla %60, %40, %20 ve %20 oranlarında dropout uygulanmıştır. ANN modeli sadece yoğun (dense) katmanlardan oluşurken, CNN modeli evrimsel (convolutional) katmanında 3x1 boyutunda bir çekirdek (kernel) ve havuzlama (pooling) katmanında 2 boyutlu bir havuzlama yapısı kullanılarak desteklenmiştir. RNN modelinde ise SimpleRNN katmanı ile ardışık veriler üzerindeki başarının artırılması amaçlanmıştır. Modellerin giriş katmanlarında ReLU (Rectified Linear Unit) aktivasyonu kullanılırken, çıkış katmanında yedi sınıf için softmax aktivasyonu kullanılmıştır. Optimizer olarak "adam", loss fonksiyonu olarak "categorical_crossentropy" ve modelin doğruluğunu ölçmek için "accuracy" metriği seçilmiştir. Eğitim süreci, 30 tur (epoch) olarak planlanmış ancak erken durdurma (early stopping) mekanizması ile doğrulama doğruluğunda (validation accuracy) üst üste 5 tur boyunca bir düşüş gözlemlendiğinde eğitimi sonlandırarak aşırı öğrenme ya da gürültü (noisy) durumlarının önüne geçilmesini sağlamıştır. Kullanılan modellere ait doğruluk (accuracy), kayıp (loss) ve F1 score değerleri karşılaştırmaları Çizelge 3'te gösterilmiştir.



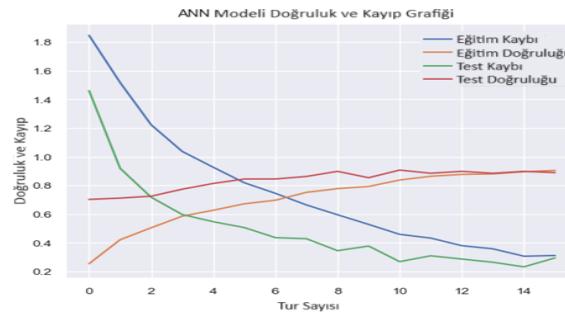
Şekil 6. CNN modelinin doğruluk ve kayıp grafiği

Derin öğrenme modellerine ait doğruluk ve kayıp eğrileri Şekil 6'da, Şekil 7'de ve Şekil 8'de gösterilmiştir. Modellere ait eğitim ve test kayıplarının düzenli olarak azalması, eğitim ve test doğruluklarının ise yine düzenli şekilde artması beklenir. Grafiklerdeki ani artış ve azalışlardan kaynaklı dalgalanmalar beklenmeyen ve istenmeyen bazı durumlar oluştuğunu gösterebilir. Bu anomaliler, modelin aşırı öğrenmeye ya da yetersiz öğrenmeye yöneldiğini veya veri setinde bir problem olduğunu gösterebilir. Eğitim kaybının ve doğruluğunun, test kaybı ve doğruluğuyla paralel ilerlemesi, modelin hem eğitiminin hem de genelleme yeteneğinin iyi olduğunu gösterir. Çalışmada kullanılan derin öğrenme modellerine ait grafikler incelendiğinde düzenli bir şekilde eğitim doğruluğunun arttığı ve eğitim kaybının azaldığı görülmektedir. Bu durum modellerin eğitim aşamasında iyi bir performans gösterdiği şeklinde yorumlanabilir. CNN modelinde 9. turdan, RNN modelinde 7. turdan ve ANN modelinde 10. turdan sonra test doğruluklarında bir durağanlık ve test kayıplarında artış sonrası dalgalanmaların görülmesi, modellerin overfitting (aşırı öğrenme) yapmaya başladığını göstermektedir;

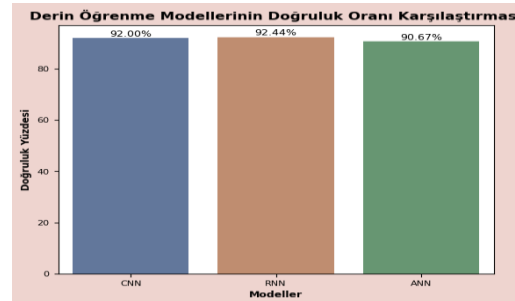
yani model eğitim setini iyi öğrenmiş olsa da test aşamasında geliştirilmeye ihtiyaç olduğu söylenebilir. Bu sorunların önüne geçmek için veri seti boyutu artırılarak, hiperparametre optimizasyonu yapılarak ve model karmaşıklıkları azaltılarak kullanılan modellerin performansları artırılabilir.



Şekil 7. RNN modelinin doğruluk ve kayıp grafiği



Şekil 8. ANN modelinin doğruluk ve kayıp grafiği



Şekil 9. Derin öğrenme modellerine ait doğruluk grafiği

Çizelge 3'teki metrikler ile doğruluk ve kayıp grafikleri incelendiğinde kullanılan derin öğrenme modelleri arasında Şekil 9'da da gösterildiği üzere RNN (Tekrarlayan Sinir Ağı), %92.44 ile en yüksek doğruluk ve %92.53 ile en yüksek F1 score elde edilmiştir. RNN'ler, zaman bağımlılıkları ve ardışık ilişkileri modellemek için geliştirilmiştir (Goodfellow, 2016). Çalışmada kullanılan veri setindeki öznitelikler (boyutlar, renk tonları veya doku gibi) arasında doğal bir sıralı ilişki olabilir. Bir özneliğin diğer özniteliklerin ardışık fonksiyonu gibi davranması, RNN modelinin bu ardışık ilişkileri CNN ve ANN modellerine kıyasla daha iyi tespit etmesini sağlamış olabilir. Kullanılan veri seti boyutunun nispeten az olması ve öznitelik sayısının fazla olması, RNN'nin önceki adımlarda elde ettiği bilgiyi bir sonraki adımda kullanan yapısı ile özniteliklerin daha anlamlı bir şekilde bir araya getirilmesini sağlamış olabilir. CNN modelleri genellikle görüntülerden oluşan veri setlerinde daha başarılı olurken, ANN modelleri bağımsız ve ilişkisel olmayan öznitelikler üzerinde yüksek performans gösterebilir (Goodfellow, 2016). RNN'nin ardışık ilişkilere duyarlılığı ve öznitelikler arasındaki ilişkileri modelleme yeteneği, diğer modellere kıyasla daha iyi sonuçlar elde edilmesini açıklayabilir.

SONUÇ

Bu çalışmada hurma meyvesinin morfolojik özellikleri üzerinden hangi tür hurma olduğu tespit edilebilmesi için çeşitli klasik makine öğrenimi ve Evrişimli Sinir Ağı (ESA) modelleri kullanılmıştır. Makine öğrenimi modellerinden Destek Vektör Makineleri (SVM) ile %92.02 doğruluk ve %88.86 F1 score elde edilmiştir. ESA modellerinden Tekrarlayan Sinir Ağı (RNN) ile %92.44 doğruluk ve %92.53 F1 score elde edilmiştir. Doğru tespit ve değerlendirme yapabilmesi veri seti büyüklüğü ile doğru orantılıdır. Kullanılan veri setinin doğruluk oranlarını arttırdığı ancak daha geniş ölçekli veri setleri ile daha doğru ve %100'e yakın sonuçlara ulaşılabilceği düşünülmektedir. Diğer sonuçlar incelendiğinde lojistik regresyon modeli %87.55 doğruluk oranı ile kıyas yapılan diğer modellerin gerisinde kalmasına karşın %90.57 ile yüksek bir F1 score sahiptir. Doğruluk oranı, doğru tahmin edilen örneklerin toplam örnek sayısına oranıdır. F1 score ise hassasiyet (precision) ve duyarlılık (recall) değerlerinin harmonik ortalamasıdır (Çelik ve ark., 2021). Bu durum, lojistik regresyon modelinin belirli sınıflarla ilgili olarak daha dengeli ve güvenilir bir performans sergilediğini gösterebilir. Daha büyük ve daha dengeli bir veri seti ile daha yüksek doğruluk ve F1 score ulaşılabilir. Çalışmada en yüksek doğruluk oranının sahip RNN modelinin doğruluk ve kayıp grafiği (Şekil 7) incelendiğinde en düşük tekrar (epoch) sayısı ile her ne kadar dalgalanmalar olsa da eğitim doğruluğu ve test doğruluğu eğrileri arasındaki farkın diğer modellere kıyasla az olması, modelin eğitim verilerinin yanı sıra doğrulama verilerine de uyum sağladığını ve yeni verilerde de benzer performansı verebileceğini gösterir. Bu durum modelin aşırı öğrenmeye (overfitting) karşı dirençli olduğu ve daha güvenilir sonuçlar üretebileceği anlamına gelir. Elde edilen sonuçlar doğrultusunda endüstriyel tarımda üretim sonrası ayrıştırma ve sınıflandırma süreçlerinin hızlandırılması için yapay zekâ destekli sistemlerin verimli şekilde kullanılabileceği söylenebilir. Gelecek çalışmalarda, daha geniş veri setleri üzerinde, farklı klasik ve derin öğrenme modelleri ile mevcut çalışmada kullanılan modellere ait parametrelerin değiştirilip geliştirilerek daha başarılı sonuçlar elde edilmesi planlanmaktadır.

Çıkar Çatışması

Makale yazarları herhangi bir çıkar çatışması olmadığını beyan ederler.

Yazar Katkısı

Yazarlar makaleye eşit oranda katkı sağlamış olduklarını beyan eder.

KAYNAKLAR

- Abdullah, S. K., Lorca, L., & Jansson, H. (2010). Diseases of date palms (Phoenix dactylifera L.). *Basrah Journal for Date Palm Researches*, 9(2), 1-44.
- Allbed, A., Kumar, L., & Shabani, F. (2017). Climate change impacts on date palm cultivation in Saudi Arabia. *The Journal of Agricultural Science*, 155(8), 1203-1218.
- Alsirhani, A., Siddiqi, M. H., Mostafa, A. M., Ezz, M., & Mahmoud, A. A. (2023). A novel classification model of date fruit dataset using deep transfer learning. *Electronics*, 12(3), 665.
- Amaya-Tejera, N., Gamarra, M., Vélez, J. I., & Zurek, E. (2024). A distance-based kernel for classification via Support Vector Machines. *Frontiers in Artificial Intelligence*, 7, 1287875.
- Bargoti, S., & Underwood, J. (2017, May). Deep fruit detection in orchards. In *2017 IEEE international conference on robotics and automation (ICRA)* (pp. 3626-3633). IEEE.
- Bishop, C. M. (2006). Pattern recognition and machine learning. *Springer google schola*, 2, 645-678.
- Bozkurt, F. (2022). A comparative study on classifying human activities using classical machine and deep learning methods. *Arabian Journal for Science and Engineering*, 47(2), 1507-1521.
- Burges, C. J. (1998). A tutorial on support vector machines for pattern recognition. *Data mining and knowledge discovery*, 2(2), 121-167.
- Büyükarıkan, B., & Ülker, E. (2020). Aydınlatma özneteliği kullanılarak evrişimsel sinir ağı modelleri ile meyve sınıflandırma. *Uludağ Üniversitesi Mühendislik Fakültesi Dergisi*, 25(1), 81-100.

- Çelik, E., Dal, D., & Aydın, T. (2021). Duygu Analizi İçin Veri Madenciliği Sınıflandırma Algoritmalarının Karşılaştırılması. *Avrupa Bilim ve Teknoloji Dergisi*, (27), 880-889.
- Date Fruit Datasets. (2021). Erişim adresi: <https://www.kaggle.com/datasets/muratkokludataset/date-fruit-datasets/data> (Erişim adresi: 3 Mayıs, 2024)
- Eberhart, R., & Kennedy, J. (1995, October). A new optimizer using particle swarm theory. In *MHS'95. Proceedings of the sixth international symposium on micro machine and human science* (pp. 39- 43). Ieee.
- Elleuch, M., Besbes, S., Roiseux, O., Blecker, C., Deroanne, C., Drira, N. E., & Attia, H. (2008). Date flesh: Chemical composition and characteristics of the dietary fibre. *Food chemistry*, 111(3), 676- 682.
- Esteva, A., Kuprel, B., Novoa, R. A., Ko, J., Swetter, S. M., Blau, H. M., & Thrun, S. (2017). Dermatologist-level classification of skin cancer with deep neural networks. *nature*, 542(7639), 115-118.
- Gazalba, I., & Reza, N. G. I. (2017, November). Comparative analysis of k-nearest neighbor and modified k-nearest neighbor algorithm for data classification. In *2017 2nd international conferences on information technology, information systems and electrical engineering (ICITISEE)* (pp. 294-298). IEEE.
- Godfray, H. C. J., Beddington, J. R., Crute, I. R., Haddad, L., Lawrence, D., Muir, J. F., ... & Toulmin, C. (2010). Food security: the challenge of feeding 9 billion people. *science*, 327(5967), 812-818.
- Gulzar, Y., Hamid, Y., Mehmood, A., & Soomro, A. B. (2022). A Deep Learning-Based Model for Date Fruit Classification: Sustainability (Switzerland). *Sustainability*, 14(10).
- Hasan, M. S., & Sattar, A. (2021, February). Arabian date classification using CNN algorithm with various pre-trained models. In *2021 Third International Conference on Intelligent Communication Technologies and Virtual Mobile Networks (ICICV)* (pp. 1431-1436). IEEE.
- Haykin, S. (2009). *Neural networks and learning machines*, 3/E. Pearson Education India.
- Holmes, W. (2019). Artificial intelligence in education: Promises and implications for teaching and learning. *Center for Curriculum Redesign*.
- Jordan, M. I., & Mitchell, T. M. (2015). Machine learning: Trends, perspectives, and prospects. *Science*, 349(6245), 255-260.
- Joseph, J. L., Kumar, V. A., & Mathew, S. P. (2021). Fruit classification using deep learning. In *Innovations in Electrical and Electronic Engineering: Proceedings of ICEEE 2021* (pp. 807- 817). Springer Singapore.
- Kaftan, İ. (2010). *Batı Türkiye gravite ve deprem katalog verilerinin yapay sinir ağları ile değerlendirilmesi* (Tez No. 276504) [Doktora tezi, Dokuz Eylül Üniversitesi-İzmir]. Yükseköğretim Kurulu Ulusal Tez Merkezi.
- Kamilaris, A., & Prenafeta-Boldú, F. X. (2018). Deep learning in agriculture: A survey. *Computers and electronics in agriculture*, 147, 70-90.
- Koklu, M., Kursun, R., Taspınar, Y. S., & Cinar, I. (2021). Classification of date fruits into genetic varieties using image analysis. *Mathematical Problems in Engineering*, 2021, 1-13.
- Krizhevsky, A., Sutskever, I., & Hinton, G. E. (2012). Imagenet classification with deep convolutional neural networks. *Advances in neural information processing systems*, 25.
- Kunduracioglu, I. (2024). CNN models approaches for robust classification of apple diseases. *Computer and Decision Making: An International Journal*, 1, 235-251.
- Kunduracioglu, I., & Pacal, I. (2024). Advancements in deep learning for accurate classification of grape leaves and diagnosis of grape diseases. *Journal of Plant Diseases and Protection*, 131(3), 1061- 1080.
- Kursa, M. B., & Rudnicki, W. R. (2010). Feature selection with the Boruta package. *Journal of statistical software*, 36, 1-13.
- LeCun, Y., Boser, B., Denker, J. S., Henderson, D., Howard, R. E., Hubbard, W., & Jackel, L. D. (1989). Backpropagation applied to handwritten zip code recognition. *Neural computation*, 1(4), 541-551.
- LeCun, Y., Bottou, L., Bengio, Y., & Haffner, P. (1998). Gradient-based learning applied to document recognition. *Proceedings of the IEEE*, 86(11), 2278-2324.
- Lin, Y., & Jeon, Y. (2006). Random forests and adaptive nearest neighbors. *Journal of the American Statistical Association*, 101(474), 578-590.
- McCulloch, W. S., & Pitts, W. (1943). A logical calculus of the ideas immanent in nervous activity. *The bulletin of mathematical biophysics*, 5, 115-133.
- Mikołajczyk, A., & Grochowski, M. (2018, May). Data augmentation for improving deep learning in image classification problem. In *2018 international interdisciplinary PhD workshop (IIPhDW)* (pp. 117-122). IEEE.
- Mitchell, T. M. (1997). *Machine learning* (Vol. 1, No. 9). New York: McGraw-hill.

- Nasiri, A., Taheri-Garavand, A., & Zhang, Y. D. (2019). Image-based deep learning automated sorting of date fruit. *Postharvest biology and technology*, 153, 133-141.
- Nishimura, J., & Shimasaki, S. (2017). Combining the complex Langevin method and the generalized Lefschetz-thimble method. *Journal of High Energy Physics*, 2017(6), 1-16.
- Pacal, I., Kunduracioglu, I., Alma, M. H., Deveci, M., Kadry, S., Nedoma, J., ... & Martinek, R. (2024). A systematic review of deep learning techniques for plant diseases. *Artificial Intelligence Review*, 57(11), 304.
- Paçal, İ., & Kunduracioğlu, İ. (2024). Data-efficient vision transformer models for robust classification of sugarcane. *Journal of Soft Computing and Decision Analytics*, 2(1), 258-271.
- Pan, Z., Rust, A. G., & Bolouri, H. (2000, July). Image redundancy reduction for neural network classification using discrete cosine transforms. In *Proceedings of the IEEE-INNS-ENNS International Joint Conference on Neural Networks. IJCNN 2000. Neural Computing: New Challenges and Perspectives for the New Millennium* (Vol. 3, pp. 149-154). IEEE.
- Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., ... & Duchesnay, É. (2011). Scikit-learn: Machine learning in Python. *the Journal of machine Learning research*, 12, 2825-2830.
- Picon, A., Alvarez-Gila, A., Seitz, M., Ortiz-Barredo, A., Echazarra, J., & Johannes, A. (2019). Deep convolutional neural networks for mobile capture device-based crop disease classification in the wild. *Computers and Electronics in Agriculture*, 161, 280-290.
- Ranjana, R., Narendra Kumar Rao, B., Raja, J., Panini Challa, N., & Madhavi, K. R. (2023). Machine learning and computer vision-beyond modeling, training, and algorithms. *Institution of Engineering and Technology* (pp. 293-307).
- Ryman-Tubb, N. F., Krause, P., & Garn, W. (2018). How Artificial Intelligence and machine learning research impacts payment card fraud detection: A survey and industry benchmark. *Engineering Applications of Artificial Intelligence*, 76, 130-157.
- Sağ, R., & Tuğcu, Z. H. (2022). Akıllı Şebeke Uygulamalarında Derin Öğrenme Tekniklerinin Kullanımına İlişkin Kısa Bir İnceleme. *EMO Bilimsel Dergi*, 13(1), 41-61.
- Savary, S., Willocquet, L., Pethybridge, S. J., Esker, P., McRoberts, N., & Nelson, A. (2019). The global burden of pathogens and pests on major food crops. *Nature ecology & evolution*, 3(3), 430-439.
- Selçuk, F., & Gülümser, A. A. (2023). İKLİM DEĞİŞİKLİĞİ ETKİSİNDE TÜRKİYE'DE TARIMSAL ÜRÜN VERİMLİLİĞİ: BÖLGESEL BİR DEĞERLENDİRME. *Bölgesel Kalkınma Dergisi*, 1(04), 425-451.
- Szegedy, C., Liu, W., Jia, Y., Sermanet, P., Reed, S., Anguelov, D., ... & Rabinovich, A. (2015). Going deeper with convolutions. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 1-9).
- Şeker, A., Diri, B., & Balık, H. H. (2017). Derin öğrenme yöntemleri ve uygulamaları hakkında bir inceleme. *Gazi Mühendislik Bilimleri Dergisi*, 3(3), 47-64.
- Uğuz, S. (2019). Makine öğrenmesi teorik yönleri ve Python uygulamaları ile bir yapay zekâ ekolü. *Nobel Yayıncılık. Ankara*.
- Van Zonneveld, M., Volk, G. M., Dulloo, M. E., Kindt, R., Mayes, S., Quintero, M., ... & Guarino, L. (2023). Safeguarding and using fruit and vegetable biodiversity. In *Science and Innovations for Food Systems Transformation* (pp. 553-567). Cham: Springer International Publishing.
- Vapnik, V. (2013). *The nature of statistical learning theory*. Springer science & business media.
- Yasrab, R., Zhang, J., Smyth, P., & Pound, M. P. (2021). Predicting plant growth from time-series data using deep learning. *Remote Sensing*, 13(3), 331.
- Zeiler, M. D., & Fergus, R. (2013). Stochastic pooling for regularization of deep convolutional neural networks. *arXiv preprint arXiv:1301.3557*.
- Zhang, Y. D., Dong, Z., Chen, X., Jia, W., Du, S., Muhammad, K., & Wang, S. H. (2019). Image based fruit category classification by 13-layer deep convolutional neural network and data augmentation. *Multimedia Tools and Applications*, 78, 3613-3632.