



Derin öğrenme için otomatik görüntü veri seti oluşturma düzeneği tasarımı ve ceviz cinslerine uygulanması

Design of an automatic image dataset generator for deep learning and its application to walnut species

Cihat Özil^{1,*} , Mürsel Önder² 

^{1,2} Tokat Gaziosmanpaşa Üniversitesi, Bilgisayar Mühendisliği Bölümü, 60150, Tokat Türkiye

Öz

Geleneksel ceviz sınıflandırma yöntemleri, uzman bilgisine bağımlılık, zaman alıcı süreçler ve subjektif değerlendirmeler gibi sınırlamalar içermektedir. Bu durum, tarım ve gıda endüstrisinde otomatik ve hassas sınıflandırma sistemlerine duyulan ihtiyacı artırmaktadır. Bu ihtiyacı gidermek amacıyla bu çalışmada, Türkiye’de yaygın olarak yetiştirilen beş farklı ceviz cinsinin (Bilecik, Chandler, Fernette, Yavuz ve Niksar) sınıflandırılması için özgün ve yüksek çözünürlüklü bir görüntü veri seti oluşturulmuştur. Bu amaç doğrultusunda, cevizlerin 360 derece görüntülenmesini sağlayan otomatik bir veri toplama sistemi tasarlanmış ve geliştirilmiştir. ESP32 mikrodenetleyici, step motor ve yüksek çözünürlüklü bir akıllı telefon kamerasının entegrasyonu ile gerçekleştirilmiştir. Her ceviz, yatay ve dikey eksenlerde 5 derecelik adımlarla 144 farklı açıdan görüntülenmiştir. Görüntüler, kontur kırpması ve arka plan temizleme ile ön işleme tabi tutulmuştur. Elde edilen veri seti eğitim, doğrulama ve test kümelerine ayrılarak düzenli şekilde organize edilmiştir. Bu çalışma, ceviz sınıflandırmasına yönelik gelecek araştırmalar için güçlü bir altyapı sunmaktadır.

Anahtar kelimeler: Ceviz sınıflandırması, Ceviz veri seti, Otomatik veri toplama, 360 derece görüntüleme, ESP32 mikrodenetleyici

1 Giriş

Ceviz, hem ekonomik hem de kültürel açıdan önemli bir tarım ürünüdür. Gıda endüstrisinde, yağ üretiminden tatlılara kadar geniş bir kullanım alanına sahip olan ceviz, aynı zamanda yüksek besin değerleri (protein, lif, antioksidanlar vb.) ile insan sağlığı için faydalı bir gıda olarak kabul edilmektedir [1]. Bunun yanı sıra, ceviz ağacı kerestesi mobilya ve dekorasyon sektöründe tercih edilen değerli bir materyaldir. Kültürel açıdan da ceviz, geleneksel mutfaklardan yöresel el sanatlarına kadar pek çok alanda yer bulmaktadır [2].

Ceviz cinslerinin doğru bir şekilde tanımlanması ve sınıflandırılması, tarım ve ticaret açısından büyük önem taşımaktadır. Çiftçiler için doğru cins belirleme, ürün verimliliğini, kalitesini (boyut, kabuk yapısı, iç doluluk vb.) ve hastalıklara dayanıklılığını artırırken, tüccarlar için ticari

Abstract

Traditional walnut grading methods have limitations such as dependence on expert knowledge, time-consuming processes and subjective evaluations. This situation increases the need for automatic and precise classification systems in agriculture and food industry. In order to fulfil this need, in this study, a unique and high resolution image dataset was created for the classification of five different walnut varieties (Bilecik, Chandler, Fernette, Yavuz and Niksar) widely grown in Turkey. For this purpose, an automatic data acquisition system that allows 360-degree imaging of walnuts was designed and developed. It was realised with the integration of an ESP32 microcontroller, stepper motor and a high-resolution smartphone camera. Each walnut was imaged from 144 different angles in 5-degree steps on the horizontal and vertical axes. The images were preprocessed with contour cropping and background removal. The resulting dataset was organised into training, validation and test sets. This study provides a strong foundation for future research on walnut classification.

Keywords: Walnut classification, Walnut dataset, Automatic data acquisition, 360 degree imaging, ESP32 microcontroller

değerleme süreçlerini (fiyatlandırma, kalite kontrol vb.) kolaylaştırır. Ayrıca tüketicilerin doğru bilgilendirilmesi, ürün tercihlerini bilinçli bir şekilde yapmalarına olanak tanır [3]. Geleneksel sınıflandırma yöntemleri (örneğin, morfolojik inceleme, uzman tadımı), uzman bilgisi gerektirmesi, subjektif olabilmesi ve zaman alıcı olması gibi nedenlerle sınırlılıklar taşımaktadır. Bu bağlamda, görüntü tabanlı sınıflandırma yöntemleri, hızlı, objektif, tekrarlanabilir ve otomasyona uygun olmaları nedeniyle dikkat çekmektedir [4].

Dijital teknolojilerin tarım sektöründeki uygulamaları, verimliliği arttırmak ve hata oranlarını azaltmak için yaygınlaşmaktadır. Özellikle meyve sınıflandırması gibi belirli tarım ürünlerinde, bu teknolojiler manuel süreçlere kıyasla daha hassas analiz ve hızlı sonuç sunarak büyük bir fark yaratmaktadır [5]. Özellikle meyve sınıflandırması

* Sorumlu yazar / Corresponding author, e-posta / e-mail: cihatozil@hotmail.com (C. Özil)

Geliş / Received: 26.01.2025 Kabul / Accepted: 24.04.2025 Yayınlanma / Published: 15.07.2025

doi: 10.28948/ngumuh.1627310

alanında yapılan çalışmalar, yapay zekâ ve görüntü işleme teknolojilerinin bir araya getirilmesiyle dikkat çekmektedir. Örneğin, Unay ve Gosselin [6], meyve sınıflandırmasında renk ve doku özelliklerinin kullanımını araştırmış, Taner vd. [7] ise elma sınıflandırması için makine öğrenimi yöntemlerini uygulamıştır. Benzer şekilde, Vo vd. [8], mango sınıflandırmasında görüntü işleme ve destek vektör makinelerini birleştirerek etkili bir çözüm sunmuştur. Bu çalışmada, özellikle renk ve doku özelliklerinin kombinasyonu ile sınıflandırma doğruluğunun artırıldığı görülmüştür. Meyve ve tarım ürünlerinin sınıflandırılması üzerine yapılan çalışmalar, özellikle yapay zekâ destekli sistemlerin kullanımının artmasıyla ivme kazanmıştır. Örneğin, Taner vd. [7], elma sınıflandırmasında makine öğrenimi yöntemlerini kullanmış ve büyük bir başarı elde etmiştir. Benzer şekilde, Kumar vd. [9] nar sınıflandırmasında görüntü işleme tekniklerini başarıyla uygulamıştır. Ancak, ceviz gibi özel tarım ürünleri için veri setlerinin sınırlı olması, bu alandaki çalışmalarını zorlaştırmaktadır.

Görüntü tabanlı sınıflandırma, ceviz gibi tarım ürünlerinin belirli görsel özelliklerine (boyut, şekil, renk, doku, kabuk yapısı vb.) dayanarak hızlı ve hassas analizler yapılmasını sağlar. Bu yöntem, insan hatasını en aza indirerek ve işlemleri hızlandırarak tarımsal süreçlerde önemli bir avantaj sunar. Mevcut görüntü tabanlı yöntemlerin çoğu, uzman bilgisine ihtiyaç duymakta ve sınırlı sayıda veri seti ile çalışmaktadır. Bu durum, daha genel, sağlam ve uygulanabilir otomatik sınıflandırma sistemlerinin geliştirilmesini zorlaştırmaktadır [10]. Özellikle derin öğrenme tabanlı yöntemlerin başarısı, büyük ve çeşitli veri setlerinin mevcudiyetine bağlıdır [11].

Çalışmanın amacı, ceviz sınıflandırması için kullanılacak özgün ve zengin bir görüntü veri seti oluşturmaktır. Çalışmanın kapsamı, sadece veri seti oluşturma süreciyle sınırlıdır. Veri toplama, görüntü ön işleme ve veri setinin organizasyonu süreçlerinde kullanılan yöntemler detaylı olarak açıklanmış, oluşturulan veri setinin yapay zeka, makine öğrenmesi ve görüntü işleme gibi alanlarda potansiyel kullanım alanları vurgulanmıştır. Bu veri setinin, ceviz sınıflandırması üzerine yapılacak gelecekteki araştırmalar için önemli bir temel oluşturması ve otomatik sınıflandırma sistemlerinin geliştirilmesine katkıda bulunması beklenmektedir.

Bu çalışmanın literatüre ve tarım sektörüne sağladığı temel katkılar şu şekilde özetlenebilir:

- **Özgün Bir Veri Seti:** Beş farklı ceviz cinsine ait yüksek çözünürlüklü ve çok açılı görüntülerden oluşan özgün bir veri seti oluşturulmuştur.
- **Otomatik Veri Toplama Sistemi:** Veri toplama sürecini hızlandıran, tutarlı ve tekrarlanabilir bir otomasyon sistemi geliştirilmiştir.
- **Veri Ön İşleme Yöntemleri:** Kontur kırpma ve arka plan temizleme gibi veri kalitesini artıran ön işleme yöntemleri uygulanmıştır.
- **Genelleştirilebilir Model:** Geliştirilen metodoloji, diğer tarım ürünlerinin görüntü tabanlı sınıflandırması için de uygulanabilir bir çerçeve sunmaktadır.

Bu çalışma, ceviz sınıflandırması için görüntü tabanlı bir sistem geliştirme sürecini içermektedir. Öncelikle, farklı ceviz cinslerinin belirlenmesi ve veri toplama sisteminin tasarlanması gerçekleştirilmiştir. Ardından, otomatik görüntü toplama süreci yürütülmüş ve veri seti oluşturulmuştur. Son aşamada ise, elde edilen görüntüler işlenerek sınıflandırmaya uygun hale getirilmiştir. Gelecek çalışmalarda, oluşturulan veri setinin farklı derin öğrenme modelleriyle test edilmesi ve model doğruluğunun artırılması hedeflenmektedir. Ayrıca, sistemin diğer tarım ürünlerine uygulanabilirliğini araştırmak için yeni denemeler yapılacaktır.

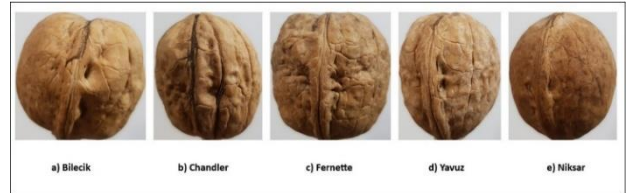
2 Materyal ve metot

Bu çalışma, Türkiye'de yaygın olarak yetiştirilen beş farklı ceviz cinsinin otomatik sınıflandırılması için bir görüntü veri seti oluşturma sürecini ve kullanılan yöntemleri detaylandırmaktadır.

2.1 Ceviz cinsleri:

Çalışmada kullanılan beş farklı ceviz cinsi, ticari önemleri, yaygınlıkları ve farklı morfolojik özelliklere sahip olmaları kriterlerine göre seçilmiştir. Numuneler, Tokat ili, Niksar ilçesi, Sorhun köyünden temin edilmiştir. Kullanılan cinsler ve açıklamaları şöyledir:

- **'Bilecik'** Cinsi: Türkiye'ye özgü, yerli bir çeşittir. Orta büyüklükte meyvelere, ince kabuklu ve yüksek iç oranına sahiptir. Yağ oranı yüksektir [12].
- **'Chandler'** Cinsi: ABD kökenli, ince kabuklu, yüksek verimli ve geç çiçeklenen bir çeşittir. Dünya genelinde en çok yetiştirilen ceviz çeşitlerinden biridir. Yan dallarda meyve verme özelliği yüksektir [13].
- **'Fernette'** Cinsi: Fransa kökenli, geç çiçeklenen ve Chandler çeşidini döleyici olarak kullanılan bir çeşittir. Soğuklara dayanıklıdır [14].
- **'Yavuz'** Cinsi: Türkiye'ye özgü, yerli bir çeşittir. Küçük-orta büyüklükte meyvelere ve yüksek iç kalitesine sahiptir. Erken meyve verme özelliği vardır [15].
- **'Niksar'** Cinsi: Türkiye'ye özgü ve coğrafi işaret almış bir çeşittir. İri, oval şekilli, dolgun içli ve yüksek yağ oranına sahiptir. Niksar ve çevresinde yetiştirilmektedir [16].

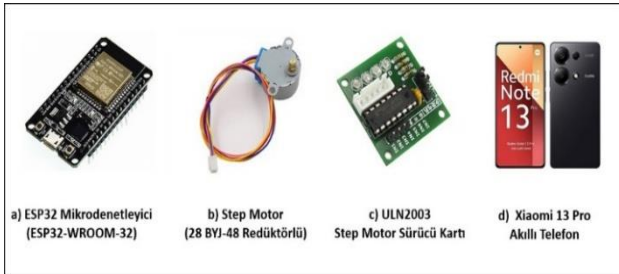


Şekil 1. Veri setinde bulunan 5 farklı ceviz cinsi

2.2 Kullanılan ekipmanlar:

- **ESP32 Mikrodenetleyici (ESP32-WROOM-32):** Fotoğraf çekim sürecini otomatikleştirmek ve step motoru kontrol etmek için kullanılmıştır. Bu mikrodenetleyici, çift çekirdekli bir işlemciye, 4 MB flaş belleğe ve Wi-Fi/Bluetooth bağlantı özelliklerine sahiptir [17]. Çalışmada Arduino IDE, C++ ve Python programlama dili kullanılmıştır.

- Step Motor (28 BYJ-48 Redüktörlü): Cevizlerin dikey ve yatay eksenlerde hassas dönüşünü sağlamıştır. Bu motor, düşük hızda yüksek hassasiyetli hareket sağlamaktadır. Bir ULN2003 motor sürücü modülü kullanılarak, hareket kontrolü gerçekleştirilmiştir [18]. Motorun küçük boyutları ve enerji verimliliği, çalışmada tercih edilme nedenleri arasında yer almaktadır.
- Xiaomi 13 Pro Akıllı Telefon: Yüksek çözünürlüklü görüntüler elde etmek için kullanılmıştır. Telefonun ana kamerası, 50 MP çözünürlüğe, 1 inç sensör boyutuna, f/1.9 diyafram açıklığına ve 23 mm odak uzaklığına sahiptir [19]. Çekimler otomatik modda, otomatik odaklama ve otomatik beyaz dengesi ayarlarıyla ve JPEG formatında gerçekleştirilmiştir.
- Aydınlatma: Fotoğraflar, ev ortamında, akşam saatlerinde ve tutarlı bir aydınlatma koşulunda çekilmiştir. Telefon, cevizlere yaklaşık 30 cm uzaklıkta konumlandırılmıştır. Bu düzenleme, gölgeleri en aza indirerek ve eşit bir aydınlatma sağlayarak görüntülerin tutarlılığını sağlamıştır.
- Arka Plan: Düz beyaz bir mat karton zemin kullanılmıştır. Bu, cevizlerin konturlarını net bir şekilde belirlemeye ve arka plan gürültüsünü en aza indirmeye yardımcı olmuştur.
- Diğer Ekipmanlar: Cevizleri sabitlemek için 3D yazıcıda PLA filament kullanılarak özel olarak tasarlanmış bir platform kullanılmıştır. Bu platform, cevizlerin çekimler sırasında aynı pozisyonda kalmasını sağlamıştır. USB kablosu, görüntülerin bilgisayara aktarılması için kullanılmıştır.



Şekil 2. Çalışmada kullanılan donanımlar.

2.3 Veri toplama prosedürü:

Çalışmada her bir ceviz cinsinden beşer adet numune kullanılmıştır. Bu numune sayısının azlığı, çalışmanın önemli bir sınırlamasıdır ve gelecekteki çalışmalarda daha fazla numune kullanılması gerekmektedir. Numuneler, ortalama boyutlarda, olgun, kabukları hasarsız, çatlak ve gözle görülür kusurları olmayan cevizlerden seçilmiştir. Cevizlerin boyutları cetvel ile ölçülerek ortalama boyutlarda olduklarından emin olunmuştur. Bu, veri setinde boyut çeşitliliğini kontrol altında tutmak ve modelin belirli boyutlara aşırı uyum sağlamasını engellemek için yapılmıştır. Nıksar, Bilecik, Chandler, Fernette ve Yavuz cinslerine ait numuneler, 2024 yılı Ekim ayında, ceviz hasat zamanı, Tokat ilinin Nıksar ilçesine bağlı Sorhun köyündeki yerel üreticilerden temin edilmiştir.

Her bir ceviz, dikey ve yatay ekseninde 5 derecelik adımlarla 144 farklı açıdan fotoğraflanmıştır. Bu yöntem, cevizlerin 360 derece tam bir görünümünü elde ederek veri setindeki çeşitliliği artırmayı amaçlamaktadır. 5 derecelik adım seçimi, yeterli detay yakalarken veri setinin boyutunu makul bir seviyede tutmak ve işlem yükünü dengelemek için yapılmıştır. Daha küçük adımlarla (örneğin 2.5 derece) daha fazla görüntü elde edilebilir ancak bu, veri setinin boyutunu ve işlem yükünü önemli ölçüde artıracaktır. 5 derecelik adımlar, cevizlerin morfolojik özelliklerini yeterince yakalamak için uygun bir denge sağlamaktadır.

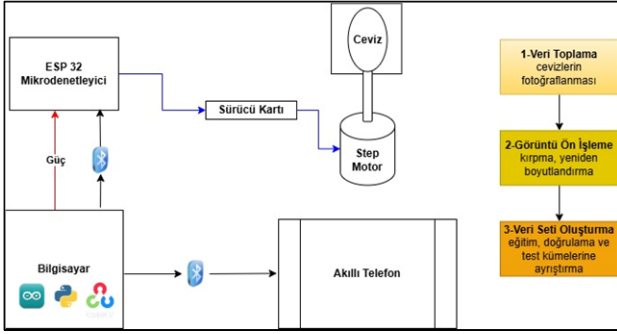
Bu işlem için step motor, ESP32 mikrodenetleyici ve telefonun kamera uygulaması senkronize çalıştırılmıştır. Step motorun kontrolü ve Bluetooth Low Energy (BLE) iletişimi Arduino IDE kullanılarak C++ ile programlanmış ESP32 mikrodenetleyici tarafından sağlanmıştır. Telefonun kamera uygulamasının açılması ve tetiklenmesi ise Python programlama dili ile geliştirilen bir "script" aracılığıyla, BLE üzerinden ESP32 ile iletişim kurularak gerçekleştirilmiştir [20].

Sistem şu şekilde çalışmaktadır: Python "script"i, BLE üzerinden ESP32'ye "start" komutu göndererek motorun dönmeye başlamasını başlatır. ESP32, motoru her 5 derecede bir döndürür ve her dönüşte Python "script"ine "Capture" mesajı gönderir. Python "script"i bu mesajı aldığında, ADB (Android Debug Bridge) komutlarını kullanarak telefonun kamerasını tetikler ve fotoğraf çeker. Motor tam bir tur (360 derece) döndükten sonra ESP32, Python "script"ine "StopCapture" mesajı göndererek fotoğraf çekme işlemini sonlandırır [21]. Bu senkronizasyonu sağlayan algoritma yapısı şu şekildedir:

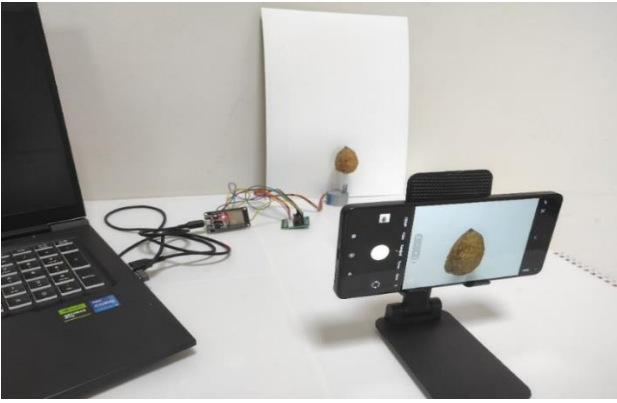
Algoritma:

- Başlangıç (Python):
 - o BLE üzerinden ESP32 ile bağlantı kurulur.
 - o ADB ile telefonun kamera uygulaması açılır.
 - o ESP32'ye "start" komutu gönderilir.
- Döngü (ESP32):
 - o Motor 5 derece döndürülür.
 - o Python "script"ine "Capture" mesajı gönderilir.
 - o Döngü 72 kez tekrarlanır (360 derece tamamlanana kadar).
- Fotoğraf Çekme (Python):
 - o ESP32'den "Capture" mesajı alındığında, Android Debug Bridge (ADB) üzerinden gönderilen input keyevent 27 komutu kullanılarak telefonun kamera uygulaması tetiklenir ve fotoğraf çekilir. Bu komut, Android cihazlarda kamera uygulamasını açmak ve fotoğraf çekmek için standart bir yöntemdir. Alternatif olarak, input tap x y komutu kullanılabilir. Bu komut, belirli ekran koordinatlarına dokunma simülasyonu yapar. x ve y koordinatları, telefonun ekranındaki kamera tetikleme butonunun bulunduğu noktanın koordinatlarıdır. Bu koordinatlar, telefonun geliştirici seçeneklerinde "Pointer location" özelliği etkinleştirilerek tespit edilebilir [22].
- Bitiş (ESP32):
 - o Tam tur tamamlandığında Python scriptine "StopCapture" mesajı gönderilir.
- Bitiş (Python):

- ESP32'den "StopCapture" mesajı alındığında fotoğraf çekme işlemi sonlandırılır.



Şekil 3. Veri seti oluşturma sisteminin blok diyagramı



Şekil 4. Veri toplama sisteminin genel görünümü

Görüntüler, telefonun hafızasına kaydedildikten sonra USB kablosu aracılığıyla bilgisayara aktarılmıştır. Aktarım sırasında veri kaybını önlemek için güvenli dosya aktarım protokolleri kullanılmıştır.

Veri toplama sürecinde bazı zorluklarla karşılaşmıştır:

- Aydınlatma koşullarındaki küçük değişiklikler: Çekimler, ev ortamında akşamları aynı aydınlık koşullarında ve aynı saatlerde gerçekleştirilmiştir. Bu sayede tutarlı bir aydınlatma sağlanmış ve çekim setinden önce beyaz dengesi kontrol edilerek olası farklar en aza indirilmiştir.

- Motorun bazen atlaması: Motor atlamaları tespit edilerek atlayan çekimler tekrarlanmıştır. Ayrıca, motor sürücüsünün akım ayarları kontrol edilerek ve motor bağlantılarının sağlamlığı kontrol edilerek atlama sorunlarının önüne geçilmeye çalışılmıştır. Mikro adım sürücüsünün doğru bir şekilde yapılandırıldığından ve motorun torkunun cevizleri döndürmek için yeterli olduğundan emin olunmuştur.

- Bluetooth bağlantı sorunları: Bazen Bluetooth bağlantısı kesilmiş veya veri iletiminde sorunlar yaşanmıştır. Bu gibi durumlarda, Bluetooth bağlantısı yeniden başlatılmış ve Python scriptinin doğru çalıştığından emin olunmuştur.

- ADB bağlantı sorunları: Bazen ADB bağlantısı kesilmiş veya komutlar doğru şekilde iletilmemiştir. Bu gibi durumlarda USB kablosu kontrol edilmiş, telefonun USB hata ayıklama modu yeniden başlatılmış ve bilgisayardaki ADB sürücülerinin güncel olduğundan emin olunmuştur.

Bu zorluklar ve çözüm yöntemleri kayıt altına alınarak veri toplama sürecinin şeffaflığı ve tekrarlanabilirliği sağlanmıştır.

2.4 Veri ön işleme:

Bu aşamada, veri setindeki görüntülerin derin öğrenme model eğitimine uygun hale getirilmesi için bir dizi ön işleme adımı, Python programlama dili ve OpenCV kütüphanesi kullanılarak uygulanmıştır. Bu adımlar boyutlandırma, kontur kırpma ve veri seti ayırıştırma.

2.4.1 Boyutlandırma

Orijinal çözünürlükteki fotoğraflar, OpenCV'nin `cv2.resize()` fonksiyonu ve bilinear interpolasyon yöntemi kullanılarak 512x512 piksel boyutuna yeniden boyutlandırılmıştır. Bu işlem, modelin giriş katmanına uygun bir boyut sağlamanın yanı sıra, gereksiz işlem yükünü azaltmayı hedeflemektedir. Bilinear interpolasyon, görüntüdeki önemli detayları korurken, daha hızlı bir yeniden boyutlandırma sağlar. Farklı interpolasyon yöntemleri (örneğin `bicubic`, `Lanczos`) de denenebilir ve sonuçlar karşılaştırılabilir [23].

2.4.2 Kontur kırpma ve arka plan temizleme

Cevizlerin görüntülerdeki arka plandan ayrılması ve sadece cevizlerin kalması için kontur kırpma yöntemi kullanılmıştır. Bu işlem için OpenCV kütüphanesindeki `findContours` ve `boundingRect` fonksiyonları kullanılmıştır [24]. İşlem adımları şöyledir:

1. `cv2.cvtColor` fonksiyonu ile görüntü gri tonlamaya dönüştürülmüştür.

2. `cv2.GaussianBlur` fonksiyonu ile görüntüdeki gürültüyü azaltmak için Gaussian bulanıklaştırma uygulanmıştır.

3. `cv2.threshold` fonksiyonu kullanılarak ikili eşikleme ile cevizler arka plandan ayrılmıştır. Kodda eşik değeri 190 olarak belirlenmiştir. Bu değer, cevizlerin ve arka planın kontrastına göre ayarlanabilir. `cv2.THRESH_BINARY_INV` parametresi kullanılarak, cevizler beyaz, arka plan siyah olarak işaretlenmiştir.

4. `cv2.findContours` fonksiyonu ile konturlar bulunmuştur. Bu fonksiyon için `RETR_EXTERNAL` (sadece dış konturları bulur) ve `CHAIN_APPROX_SIMPLE` (kontur noktalarını sıkıştırır) parametreleri kullanılmıştır.

5. En büyük kontur (ceviz konturu olduğu varsayılır) seçilmiştir.

6. `cv2.boundingRect` fonksiyonu ile bu konturu çevreleyen dikdörtgen alan bulunmuştur.

7. Görüntü bu dikdörtgen alana göre kırılmıştır.

8. Alt Kısım Temizleme: Python kodunda, kırılan görüntünün alt kısmında kalan beyaz alanları temizlemek için ek bir işlem yapılmıştır. Bu işlemde, kırılan görüntünün her satırındaki piksel değerlerinin toplamı hesaplanmış ve belirli bir eşik değerinin üzerinde beyazlık içeren satırlar tespit edilerek görüntü bu noktadan kırılmıştır. Bu, özellikle fotoğraflarda cevizlerin altında kalan tabla veya zemin gibi kısımların temizlenmesini sağlamıştır.



Şekil 5. Orijinal ve veri ön işleme yapılmış görüntü

2.4.3 Veri seti ayrıştırma

Önceden işlenmiş görüntüler, Python'da bulunan `train_test_split` fonksiyonu kullanılarak %50 eğitim, %25 doğrulama ve %25 test olmak üzere rastgele olarak ayrılmıştır. Bu oranlar, veri setinin sınırlı boyutu göz önünde bulundurularak tercih edilmiştir. Derin öğrenme uygulamalarında yaygın olarak kullanılan %70 eğitim, %15 doğrulama ve %15 test oranlarına kıyasla, bu oranlar daha dengeli bir dağılım sağlamaktadır. Daha büyük veri setlerinde farklı oranlar (örneğin %80 eğitim, %10 doğrulama, %10 test) kullanılabilir. Ayrıştırma işlemi, her bir cinsin veri setinde orantılı olarak temsil edilmesini sağlamak için katmanlı örnekleme (stratified sampling) yöntemi kullanılarak yapılmıştır. Bu, modelin tüm cinsleri eşit şekilde öğrenmesine yardımcı olur [25]. Python'da `sklearn.model_selection` modülünden `train_test_split` fonksiyonu bu işlem için kullanılmıştır ve `random_state` parametresi ile rastgelelik kontrol altına alınmıştır. Veri setinin ayrıştırılması ve kaydedilmesi için gerekli klasörler Python kodu tarafından otomatik olarak oluşturulmaktadır.

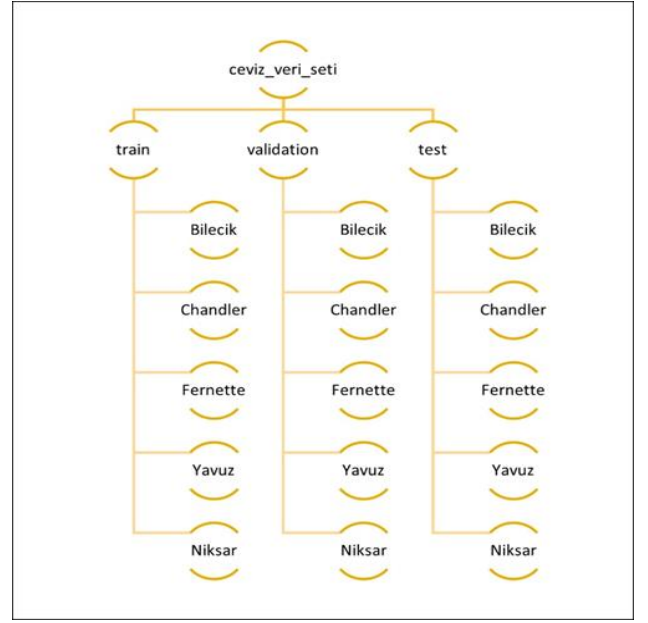
2.5 Klasör organizasyonu:

Veri setinin düzenli ve erişilebilir olması için standart bir klasör yapısı oluşturulmuştur. Bu yapı, veri setinin farklı algoritmalar ve platformlar üzerinde kullanılabilirliğini artırmaktadır. Ana klasör `ceviz_veri_seti` olarak adlandırılmıştır ve tüm verileri içermektedir. Bu klasörün altında, veri setinin eğitim, doğrulama ve test kümelerine ayrılmış halleri bulunmaktadır.

Veri seti, ön işleme adımları (2.4. Veri Ön İşleme bölümünde açıklanan boyutlandırma ve kontur kırma işlemleri) uygulandıktan sonra, `train`, `validation` ve `test` adlı klasörlere ayrılmıştır. Her bir klasörün içinde, ceviz cinsleri için alt klasörler bulunur. Her bir ceviz cinsinden 5 adet ceviz bulunmaktadır ve her ceviz için yatay ve dikey eksenlerde 144 farklı açıdan çekilmiş görüntü mevcuttur. Bu, her bir ceviz cinsi için toplamda $5 \times 144 = 720$ görüntü anlamına gelir. Bu bilgiler ışığında klasör yapısı ve görüntü sayıları dağılımı aşağıdaki gibidir:

Tablo 1. Veri setindeki ceviz cinslerinin görüntü sayıları

| Ceviz Cinsi | Eğitim Görüntü Sayısı | Doğrulama Görüntü Sayısı | Test Görüntü Sayısı |
|-------------|-----------------------|--------------------------|---------------------|
| Bilecik | 360 | 180 | 180 |
| Chandler | 360 | 180 | 180 |
| Fernette | 360 | 180 | 180 |
| Yavuz | 360 | 180 | 180 |
| Niksar | 360 | 180 | 180 |



Şekil 6. Veri seti klasör yapısı

Bu klasör yapısında:

- `ceviz_veri_seti/` ana klasördür ve tüm verileri içerir.
- `train/`, `validation/` ve `test/` klasörleri, veri setinin eğitim, doğrulama ve test kümelerine ayrılmış halini içerir. Bu klasörlerin her birinde, ceviz cinsleri için alt klasörler bulunur. Bu sayede, her kümede her cinsin örnekleri bulunur ve modelin farklı kümeler üzerinde performansını değerlendirmek mümkün olur.

Bu yapı, veri setinin yönetimi, eğitimi ve değerlendirilmesi için gerekli olan tüm organizasyonu sağlar. Veri setinin ayrıştırılması sırasında oluşturulan `train/`, `validation/` ve `test/` klasörleri, 2.4. Veri Ön İşleme bölümünde bahsedilen %50-%25-%25 oranlarına göre otomatik olarak oluşturulmuştur. Bu işlem sırasında, her bir cinsten rastgele seçilen önceden işlenmiş görüntüler ilgili klasörlere kopyalanmıştır. Katmanlı örnekleme uygulandığı için, her kümede her cinsten orantılı sayıda örnek bulunmaktadır. Örneğin, Bilecik cinsinden toplam 720 önceden işlenmiş görüntü varsa, yaklaşık 360 görüntü `train/Bilecik/`, 180 görüntü `validation/Bilecik/` ve 180 görüntü `test/Bilecik/` klasörüne kopyalanmıştır. Bu oranlar, kusurlu sayılar ortaya çıktığında en yakın tam sayıya yuvarlanmıştır.

3 Bulgular ve tartışma

Çalışma, ceviz sınıflandırması problemi için özgün ve yüksek kaliteli bir görüntü veri seti sunmaktadır. Veri toplama ve işleme süreçlerinde hem yazılım (ESP32 kontrol yazılımı, ADB komutları, Python görüntü işleme kodları) hem de donanım (ESP32 mikrodenetleyici, step motor, yüksek çözünürlüklü akıllı telefon kamerası) unsurları entegre bir şekilde kullanılmıştır. Bu entegre yaklaşım, model performansının artırılmasını hedeflemiştir.

Han vd. [26], elma sınıflandırması ve Lei vd. [27] nar sınıflandırması çalışmaları, bu çalışmanın yöntemleriyle bazı paralellikler göstermektedir. Özellikle, meyve konturlarının tespiti ve sınıflandırmada kullanılması konusunda benzer yaklaşımlar mevcuttur. Ancak, bu çalışma, cevizlerin üç boyutlu yapısını daha iyi yakalamak için 360 derece görüntüleme tekniğini kullanarak literatürdeki diğer çalışmalardan ayrılmaktadır. Bu otomasyon sistemi, cevizlerin her açısının standart ve tekrarlanabilir bir şekilde görüntülenmesini sağlayarak, veri setindeki varyasyonu ve zenginliği önemli ölçüde artırmıştır. Bu, özellikle kabuk yapısındaki ince farklılıkların tespitinde kritik bir rol oynamıştır. Literatürdeki benzer çalışmalarda genellikle manuel görüntüleme yöntemleri kullanıldığından, bu otomasyon sistemi çalışmaya özgün bir değer katmaktadır. Bu sayede, cevizlerin sadece belirli açılardan değil, tüm yüzey özelliklerinin analiz edilmesi mümkün kılınmıştır. Bu, özellikle farklı ceviz cinslerinin kabuk yapısındaki ve şeklindeki ince farklılıkların ayırt edilmesinde önemli bir avantaj sağlamıştır. Literatürde ceviz sınıflandırması üzerine yapılan çalışmalarda, Akça ve Yılmaz [28] ve Ünver vd. [29] gibi kaynaklarda ceviz çeşitlerinin morfolojik farklılıklarının tespit edilmesi hedeflenmiştir. Bu çalışmalar, ceviz morfolojisini incelemiş olsalar da, çalışmamızda kullanılan 360 derece görüntüleme tekniği ve otomasyon sistemi gibi yenilikçi yaklaşımları içermemektedir.

3.1 Çalışmanın sınırlamaları

Bu çalışmanın bazı sınırlamaları bulunmaktadır. En önemlisi, veri toplama sürecinin yalnızca beş ceviz cinsi (Bilecik, Chandler, Fernette, Yavuz ve Niksar) ve her cins için başlangıçta yalnızca beş örnek ile sınırlı tutulmasıdır. Bu durum, veri setinin genelleştirilebilirliğini sınırlayabilir ve modelin belirli örneklere aşırı uyum (overfitting) göstermesine neden olabilir. Ayrıca, numune sayısının azlığı, istatistiksel analizlerin gücünü de etkileyebilir. Gelecekteki çalışmalarda, daha fazla ceviz cinsi (Türkiye'de yetiştirilen diğer önemli ceviz çeşitleri eklenebilir) ve her bir cins için çok daha fazla örnek (en az 100-200 örnek önerilir) içeren genişletilmiş ve dengeli bir veri seti oluşturulması önerilmektedir. Bu, modelin farklı ceviz çeşitleri arasındaki varyasyonları daha iyi öğrenmesine ve gerçek dünya koşullarına daha iyi uyum sağlamasına yardımcı olacaktır.

Çevresel faktörlerin (ışık koşulları, arka plan farklılıkları, nem, sıcaklık vb.) veri toplama süreci üzerindeki etkisi detaylı olarak araştırılmamıştır. Farklı ışık koşullarında (doğal ışık, yapay ışık, farklı renk sıcaklıkları), farklı arka planlarda (farklı renkler, dokular) ve farklı çevresel koşullarda (farklı nem ve sıcaklık seviyeleri) veri toplanarak

modelin bu faktörlere karşı dayanıklılığı artırılabilir. Bu, modelin gerçek dünya uygulamalarında daha güvenilir ve tutarlı sonuçlar vermesini sağlayacaktır.

Veri seti oluşturma aşamasında karşılaşılan teknik zorluklar (aydınlatma koşullarındaki küçük değişiklikler, motorun bazen atlaması, Bluetooth ve ADB bağlantı sorunları) ve bu sorunlara getirilen çözümler (beyaz dengesi kontrolü, motor sürücüsü ayarlarının kontrolü, bağlantıların yeniden başlatılması vb.) detaylı olarak açıklanmıştır. Bu, çalışmanın tekrarlanabilirliğini artırmaktadır.

4 Sonuçlar

Bu çalışma, Türkiye'de yaygın olarak yetiştirilen beş farklı ceviz cinsinin (Bilecik, Chandler, Fernette, Yavuz ve Niksar) otomatik sınıflandırılması için özgün ve yüksek çözünürlüklü bir görüntü veri seti oluşturulmasını başarıyla gerçekleştirmiştir. Bu veri seti, ceviz sınıflandırması alanında gelecekte yapılacak çalışmalar için önemli bir temel oluşturmaktadır. Geliştirilen otomatik veri toplama sistemi, ESP32 mikrodenetleyici, step motor ve yüksek çözünürlüklü akıllı telefon kamerasının entegrasyonu ile cevizlerin 360 derece görüntülenmesini sağlayarak veri setindeki çeşitliliği önemli ölçüde artırmıştır. Uygulanan kontur kırpma gibi ön işleme adımları, görüntüdeki gereksiz arka plan gürültüsünü azaltarak veri setinin kalitesini artırmıştır.

Bu çalışmanın temel katkıları şunlardır:

- Özgün ve Yüksek Kaliteli Veri Seti: Literatürdeki çalışmalara kıyasla daha kapsamlı görsel bilgi sunan, beş farklı ceviz cinsine ait yüksek çözünürlüklü ve çok açılı görüntülerden oluşan özgün bir veri seti oluşturulmuştur.
- Otomatik Veri Toplama Sistemi: Veri toplama sürecini daha verimli, tutarlı ve tekrarlanabilir hale getiren bir otomasyon sistemi geliştirilmiştir.
- Veri Ön İşleme Yöntemleri: Veri setinin kalitesini artıran kontur kırpma gibi ön işleme yöntemleri uygulanmıştır.

Bu çalışma, ceviz yetiştiriciliği, kalite kontrolü ve otomasyonu gibi alanlarda önemli uygulamalara sahip olabilecek bir veri seti sunmaktadır. Oluşturulan bu veri seti, gelecekteki araştırmacılar tarafından cevizlerin otomatik sınıflandırılması için farklı makine öğrenmesi ve derin öğrenme modellerinin geliştirilmesinde kullanılabilir.

Gelecekteki çalışmalar, veri setinin genişletilmesi, farklı çevresel koşullarda veri toplanması ve farklı görüntüleme tekniklerinin denenmesi gibi konulara odaklanabilir. Bu çalışma, tarım sektöründe otomasyonun önemini ve potansiyelini göstermesi açısından da değerlidir. Özetle, bu çalışma, ceviz sınıflandırması alanında özgün ve yüksek kaliteli bir görüntü veri seti oluşturularak önemli bir katkı sağlamaktadır. Geliştirilen veri toplama yöntemi ve oluşturulan veri seti, gelecekteki araştırmalar için sağlam bir temel oluşturacak ve ceviz yetiştiriciliği ve işleme süreçlerinin iyileştirilmesine katkıda bulunacaktır. Bu veri setinin kamuya açık bir şekilde paylaşılması, alanın ilerlemesine önemli bir ivme kazandıracaktır.

Çıkar çatışması

Yazarlar çıkar çatışması olmadığını beyan etmektedir.

Benzerlik oranı (iThenticate): %4

Kaynaklar

- [1] M. Şimşek ve E. Gülsoy, Ceviz ve içerdiği yağ asitlerinin insan sağlığı açısından önemi üzerine yapılan bazı çalışmalar. *Journal of the Institute of Science & Technology/Fen Bilimleri Enstitüsü Dergisi*, 6(4), 2016.
- [2] A. Şimşek ve B. Temel, Gastronomi turizmi kapsamında geleneksel ürünlerin incelenmesi: Tarsus Mamülü. *Mersin Üniversitesi Sosyal Bilimler Enstitüsü Dergisi*, 5(2), 6-13, 2022. <https://doi.org/10.55044/meusbd.1000865>.
- [3] M. Kilci, Tokat ili Niksar ilçesi ceviz üretim ve pazarlama yapısı (Yüksek lisans tezi, Fen Bilimleri Enstitüsü), 2016.
- [4] A. Y. Yiğit ve M. Uysal, Nesne tabanlı sınıflandırma yaklaşımı kullanılarak yolların tespiti. *Türkiye Fotogrametri Dergisi*, 1(1), 17-24, 2019.
- [5] H. Şahin, Dijital tarım, Tarım 4.0, akıllı tarım, robotik uygulamalar ve otonom sistemler. *Tarım Makinaları Bilimi Dergisi*, 18(2), 68-83, 2022.
- [6] D. Unay and B. Gosselin, Stem and calyx recognition on 'Jonagold' apples by pattern recognition. *Journal of food Engineering*, 78(2), 597-605, 2007. <https://doi.org/10.1016/j.jfoodeng.2005.10.038>.
- [7] A. Taner, M. T. Mengstu, K. Ç. Selvi, H. Duran, İ. Gür and N. Ungureanu, Apple varieties classification using deep features and machine learning. *Agriculture*, 14(2), 252, 2024. <https://doi.org/10.3390/agriculture14020252>.
- [8] M. T. Vo, T. D. Nguyen and N. T. Dang, Embedded machine learning for mango classification using image processing and support vector machine. In 2019 6th NAFOSTED Conference on Information and Computer Science (NICS) (ss. 279-284). IEEE, 2019. <https://doi.org/10.1109/NICS48868.2019.9023803>.
- [9] P. S. Kumar, S. Sudha, P. Das, D. Pradeep and K. Vijaipriya, Pomegranate quality analysis and classification using feature extraction and machine learning. In 2022 6th International Conference on Electronics, Communication and Aerospace Technology (ss. 862-867). IEEE, 2019. <https://doi.org/10.1109/ICECA55336.2022.10009628>.
- [10] K. Kayaalp ve A. Altınalan, Ceviz sürgün görüntüleri kullanılarak ceviz çeşitlerinin sınıflandırılması. *Uluborlu Mesleki Bilimler Dergisi*, 6(2), 42-52, 2024.
- [11] Ş. E. Eze ve S. K. Berkaya, Görüntü ön işleme teknikleri ve derin öğrenme ile bitki zararlılarının sınıflandırılması. *Mühendislik Bilimleri ve Tasarım Dergisi*, 12(2), 455-465, 2024. <https://doi.org/10.21923/jesd.1490176>.
- [12] M. Erkol, E. Altuntaş ve H. Tokbaş, Bilecik ve Yalova-3 ceviz çeşitlerinin bazı fiziksel özelliklerine nem içeriğinin etkisi. *Tarım Bilimleri Araştırma Dergisi*, (2), 37-43, 2008.
- [13] M. B. Olut, Chandler x Kaplan-86 F1 ceviz populasyonunun meyvelerinde yağ asidi ve tokoferol içeriklerinin karakterizasyonu (Yüksek lisans tezi, Fen Bilimleri Enstitüsü), 2018.
- [14] F. Şen, S. Bilgin, E. Özeker ve N. A. Bilgin, Bazı ceviz çeşitlerinin Menemen ekolojisinde morfolojik ve pomolojik özelliklerinin belirlenmesi. *ÇOMÜ Ziraat Fakültesi Dergisi*, 6(1), 31-39, 2018.
- [15] S. Bayazıt, O. Caliskan and D. Kılıç, Pomological and chemical properties of some walnut genotypes in central Anatolia. *Journal of Agricultural Faculty of Gaziosmanpaşa University (JAFAG)*, 36(3), 243-249, 2019. <https://doi.org/10.13002/jafag4583>.
- [16] T. Çukur, N. Kızılaslan, F. Çukur ve H. Kızılaslan, Tüketicilerin coğrafi işaretli ürünler için ödeme istekliliğine etki eden faktörler: Niksar ceviz örneği. *Turkish Journal of Agriculture-Food Science and Technology*, 8(11), 2476-2481, 2020. <https://doi.org/10.24925/turjaf.v8i11.2476-2481.3898>.
- [17] F. Başçıftçi ve K. A. Gündüz, Nesnelerin interneti uyumlu mikrodenetleyiciler üzerine bir araştırma. *Selçuk Üniversitesi Sosyal ve Teknik Araştırmalar Dergisi*, (18), 66-76, 2019.
- [18] E. Güler, Fındık zarı uygunluğunun OpenCV ile belirlenmesi için yazılım ve donanım geliştirilmesi (Doktora tezi, Selçuk Üniversitesi), 2018.
- [19] Xiaomi. (n.d.). Xiaomi 13 Pro teknik özellikler. Xiaomi Türkiye. <https://www.mi.com/tr/product/xiaomi-13-pro/specs/>
- [20] F. J. Dian, A. Yousefi, and S. Lim. A practical study on Bluetooth Low Energy (BLE) throughput. In 2018 IEEE 9th Annual Information Technology, Electronics and Mobile Communication Conference (IEMCON) (ss. 768-771). IEEE, 2018. <https://doi.org/10.1109/IEMCON.2018.8614763>.
- [21] EazSoftware, A comprehensive guide to ADB (Android Debug Bridge): The unsung hero for Android developers. Medium, 2024. <https://medium.com/@EazSoftware/a-comprehensive-guide-to-adb-android-debug-bridge-the-unsung-hero-for-android-developers-28b349037436>
- [22] Android Developers. (t.y.). Android Debug Bridge (adb). <https://developer.android.com/tools/adb?hl=tr>
- [23] S. Balovsyak ve Y. Hnatiuk, Analysis of results of scaling digital images by interpolation algorithms. *Security of Infocommunication Systems and Internet of Things*, 2(1), 01007-01007, 2024. <https://doi.org/10.31861/sisiot2024.1.01007>.
- [24] Q. Wang, Y. Bi, B. Xue, and M. Zhang, Genetic programming with flexible region detection for fine-grained image classification. *IEEE Transactions on Evolutionary Computation*, 2024. <https://doi.org/10.1109/TEVC.2024.3379257>.
- [25] J. Sadaiyandi, P. Arumugam, A. K. Sangaiah and C. Zhang, Stratified sampling-based deep learning approach to increase prediction accuracy of unbalanced dataset. *Electronics*, 12(21), 4423, 2023. <https://doi.org/10.3390/electronics12214423>.
- [26] B. Han, J. Zhang, R. Almodfer, Y. Wang, W. Sun, T. Bai and W. Hou, Research on innovative apple grading technology driven by intelligent vision and machine

- learning. *Foods*, 14(2), 258, 2025. <https://doi.org/10.3390/foods14020258>.
- [27] X. Lei, H. Ouyang and L. Xu, Mature pomegranate recognition methods in natural environments using machine vision. *Ciencia Rural*, 49, e20190298, 2019. <https://doi.org/10.1590/0103-8478cr20190298>.
- [28] Y. Akça ve S. Yılmaz, Tüplerde yetiştirilen farklı ceviz (*Juglans regia*) çeşitlerine ait çöğürlerin bazı morfolojik özelliklerinin ve kalitelerinin belirlenmesi. *Journal of Agricultural Faculty of Gaziosmanpaşa University (JAFAG)*, 33(3), 157-166, 2016. <https://doi.org/10.13002/jafag950>.
- [29] H. Ünver, E. Sakar ve M. Sülüşoğlu Durul, Düzce ilinde doğal olarak bulunan ceviz genotipleri arasındaki morfolojik farklılıkların değerlendirilmesi. *Turkish Journal of Agricultural and Natural Science*, 10(4), 2023. <http://doi.org/10.30910/turkjans.1307868>.

