



TÜBİTAK–2209-A ÜNİVERSİTE ÖĞRENCİLERİ ARAŞTIRMA PROJELERİ DESTEĞİ PROGRAMI

Başvuru formunun Arial 9 yazı tipinde, her bir konu başlığı altında verilen açıklamalar göz önünde bulundurularak hazırlanması ve ekler hariç toplam 20 sayfayı geçmemesi beklenir (Alt sınır bulunmamaktadır). Değerlendirme araştırma önerisinin özgün değeri, yöntemi, yönetimi ve yaygın etkisi başlıkları üzerinden yapılacaktır.

ARAŞTIRMA ÖNERİSİ FORMU

2024 Yılı

1. Dönem Başvurusu

2209/A ÜNİVERSİTE ÖĞRENCİLERİ ARAŞTIRMA PROJELERİ DESTEĞİ PROGRAMI
ARAŞTIRMA ÖNERİSİ FORMU

A. GENEL BİLGİLER

Başvuru Sahibinin Adı Soyadı: Hamza ŞENGÜL, Hüseyin KÖROĞLU
Araştırma Önerisinin Başlığı: Portakal bahçelerinde makine öğrenmesine dayalı rekolte tahmini
Danışmanın Adı Soyadı: Dr. Öğr. Üyesi Serkan NAS
Araştırmanın Yürütüleceği Kurum/Kuruluş: Tarsus Üniversitesi

ÖZET

Türkçe özetin araştırma önerisinin (a) özgün değeri, (b) yöntemi, (c) yönetimi ve (d) yaygın etkisi hakkında bilgileri kapsamı beklenir. Bu bölümün en son yazılması önerilir.

Ekonomik koşullar ve iklim değişiklikleri, tarım ürünlerinin verimliliğini etkileyerek çiftçilerin yanlış hesaplamalar yapmasına ve ekonomik zarar etmesine neden olmaktadır. Gelişmekte olan ülkelerde, ürün miktarının doğru tahmin edilememesi çiftçilerin giderlerini yanlış hesaplamalarına yol açmaktadır. Ayrıca paketlenme ve ihracat yapan fabrikalar da çiftçilerden gelen hatalı veriler nedeniyle yıl sonu hesaplamalarında yanlışlanmaktadır. Bu belirsizliklerden dolayı çiftçiler büyük şehirlere göç etmektedir. Fabrikaların ise sektörden çekilmesine veya sektör değiştirmelerine neden olmaktadır. Bu projede, portakal ağaçlarında rekolte tahmini yapabilmek için makine öğrenmesi ve görüntü işleme teknikleri kullanılacaktır. Nesne tespiti ve segmentasyon için Mask R-CNN algoritması ile portakal meyveleri yüksek doğrulukla tanınacak, böylece yaprak yoğunluğu yüksek ağaçlarda bile doğru sayım yapılacaktır. RGB görüntülerine ek olarak, doğruluğu artırmak amacıyla LiDAR sensörlerinden alınan 3D verilerle destek sağlanacaktır. LiDAR sensörleri, üç boyutlu tanıma konusunda oldukça iyi sonuçlar vermektedir. 3D veriler yaprak ve dalların yoğun olduğu bölgelerde meyve tespitini kolaylaştıracaktır. Makine öğrenmesi modelleri olarak Gradient Boosting, Random Forest ve Convolutional Neural Networks (CNN) algoritmaları tercih edilecektir. Gradient Boosting, ardışık öğrenme sayesinde doğruluğu artırırken, Random Forest aşırı öğrenmeye dayanıklı ve stabil sonuçlar vermektedir. Görsel verilerin analizinde ise CNN, portakal sayımı ve anormalliklerin tespitinde etkili bir araçtır ve LiDAR verileri ile birleştirildiğinde yüksek doğruluk sağlar. Proje kapsamında örneğin, 1000 ağaçlık bir bahçede 50 ağacın görsel ve 3D verileri analiz edilerek tüm bahçenin rekolte tahmini yapılacaktır. Bu tahminler sayesinde çiftçiler personel ve diğer maliyetlerini daha doğru hesaplayabilecek, fabrikalar ise üretim ve paketlenme planlarını daha yüksek doğruluk oranı ile yapabileceklerdir. Doğru hesaplamalar sayesinde, portakal gibi gıda ürünlerinin ekonomik etkilerinin tahminine de katkıda bulunacaktır.

Anahtar Kelimeler: Görüntü işleme, LiDAR sensörleri, rekolte tahmini, makine öğrenmesi, derin öğrenme.

1. ÖZGÜN DEĞER

1.1. Konunun Önemi, Araştırma Önerisinin Özgün Değeri ve Araştırma Sorusu/Hipotezi

Araştırma önerisinde ele alınan konunun kapsamı ve sınırları ile önemi literatürün eleştirel bir değerlendirmesinin yanı sıra nitel veya nicel verilerle açıklanır.

Özgün değer yazılırken araştırma önerisinin bilimsel değeri, farklılığı ve yeniliği, hangi eksikliği nasıl gidereceği veya hangi soruna nasıl bir çözüm geliştireceği ve/veya ilgili bilim veya teknoloji alan(lar)ına kavramsal, kuramsal ve/veya metodolojik olarak ne gibi özgün katkılarda bulunacağı literatüre atıf yapılarak açıklanır.

Önerilen çalışmanın araştırma sorusu ve varsa hipotezi veya ele aldığı problem(ler)i açık bir şekilde ortaya konulur.

Bugüne kadar insanlar tarımda verimliliği artırmak ve üretilecek ürünlerin toplam miktarının tahminlerini hızlı bir şekilde yapabilmek amacıyla birçok yöntemle başvurmuşlardır. Tarımsal üretim doğa koşullarına bağlı olduğu için risk ve belirsizlik yüksektir. Günümüzde birçok alanda teknolojik gelişmeler yaşanmıştır. Özellikle tarım alanında yapay zekânın kullanımına yönelik büyük bir gelişme gösteren bilgisayar teknolojileri sayesinde, tarımda yaşanan bu risk ve belirsizliklere daha hızlı ve tutarlı çözümler üretmek mümkün hale gelmiştir. Tarımda yaşanan risk ve belirsizliklerden birisi de rekolte tahminidir. Meyveler ağaçlardan toplanmadan önce uzman kişiler tarafından meyvenin toplam miktarı tahmin edilir ve satış işleminde ona göre fiyat belirlenir.

2209/A ÜNİVERSİTE ÖĞRENCİLERİ ARAŞTIRMA PROJELERİ DESTEĞİ PROGRAMI ARAŞTIRMA ÖNERİSİ FORMU

Bu proje, görsel ve LiDAR verilerinin tarımsal üretimde rekolte tahmini için entegre kullanımını önermekte olup, literatürde bu iki veriyi bir arada kullanan ilk çalışmalar arasında yer almaktadır. Geleneksel yöntemler genellikle gözlemci tahminlerine dayanmakta olup, bu durum insan hatası riskini artırmak ve geniş alanlarda manuel veri toplamanın zorlukları ile karşılaşmaktadır. Ayrıca, sadece görsel veriler kullanılarak yapılan çalışmalar da vardır, ancak bunlar ağaçların yalnızca 2D görsellerine dayanarak sınırlı bir tahmin doğruluğu sağlayabilmektedir. Bu noktada, önerilen proje, LiDAR teknolojisinin sunduğu 3D veri işleme kapasitesi ile bu sınırlamaları aşmayı hedeflemektedir. LiDAR, lazer tabanlı ölçüm yaparak bir ağacın tam yapısal detaylarını sağlayabilmektedir. Bu sayede ağaçların yüksekliği, genişliği, meyve yoğunluğu gibi kritik özellikler 3D olarak modellenilebilir. Bu tür veriler, görsel verilerle birleştirildiğinde, daha isabetli bir tahmin yapma olanağı sunar. Görsel veriler, renk, şekil ve yüzey pürüzsüzlüğü gibi özelliklerin analizini yaparken, LiDAR verileri sayesinde meyve yoğunluğu ve ağaç yapısal özellikleri hassas bir şekilde ölçülebilir. Dolayısıyla, proje yalnızca meyve sayımına değil, ağaçların yapısal sağlığına da odaklanarak çok daha kapsamlı bir veri seti sunmaktadır. Proje kapsamında, bir bahçede bulunan 1000 portakal ağacından yalnızca 20 ağaçtan elde edilen veriler kullanılarak tüm bahçedeki rekolte tahmini yapılacaktır. Makine öğrenmesi algoritmaları, bu verilerden anlamlı sonuçlar çıkaracak ve küçük bir örnekleme ile geniş bir alanın verimini tahmin etme kapasitesine sahip olacaktır. Bu yöntem, manuel gözlemlerden ve sadece 2D görsel veriye dayanan çalışmalardan daha hızlı, daha doğru ve maliyet etkin bir çözüm sunmaktadır. Ek olarak, mevcut literatür incelemeleri, görsel verilerin tahmin doğruluğunu artırmada önemli bir rol oynadığını, ancak LiDAR teknolojisinin sunduğu derinlemesine analiz kapasitesinin göz ardı edilmemesi gerektiğini göstermektedir. Literatürde yalnızca görsel veriler kullanılarak yapılan tahminlerde elde edilen doğruluk oranları %70-85 arasında değişmektedir. Önerilen projede ise bu oranların LiDAR ile daha da yükseleceği öngörülmektedir. LiDAR ile yapılan 3D modellemeler sayesinde ağaçların gölgedeki meyveleri ve yapraklarının altında gizlenmiş olan meyveler dahi tespit edilebilecek ve bu da tahminlerin isabetliliğini artıracaktır. Bu yenilikçi yaklaşım, tarım sektöründe rekolte tahmini için devrim niteliğinde bir adım olarak görülmektedir. Küresel iklim değişikliği ve gıda kıtlığı gibi sorunlarla mücadele etmek için, doğru ve hızlı tahmin yapabilen bu teknoloji, tarımsal süreçleri optimize ederek çiftçilerin karar alma süreçlerini geliştirecektir. Çiftçiler, bu teknoloji sayesinde ekim planlarını, hasat zamanlamalarını ve iş gücü dağılımını daha verimli bir şekilde yapabilecek, bu da genel tarımsal üretkenliğin artmasına katkı sağlayacaktır.

Bu proje, tarım sektöründeki dijital dönüşümü hızlandırarak sürdürülebilir tarımsal üretime önemli bir katkı sağlayacaktır. Çiftçiler, ürün verimini önceden tahmin edebilecek ve bu tahminler sayesinde maliyet hesaplamalarını daha doğru yaparak kar oranlarını artıracaktır. Ayrıca, proje küresel iklim değişikliklerinin getirdiği tarımsal zorluklara karşı teknoloji destekli bir çözüm sunarak, tarımda verimliliği ve sürdürülebilirliği güçlendirecektir.

İlgili literatür tarandığında bu çalışmayla aynı doğrultuda olarak ortaya konmuş çalışmalar için:

Portakal ağaçlarından alınan görsel ve LiDAR verileri ile rekolte tahmini yapılması üzerine çeşitli çalışmalar mevcuttur. Bu çalışmalarda genellikle, meyve sayımı, meyve büyüklüğü tahmini ve ağaç verimliliği üzerine yoğunlaşmaktadır. Örneğin, Şule Ataç'ın elma ağaçları üzerine yapmış olduğu tezde, yapay zekâya dayalı görüntü işleme teknikleri kullanılarak elma ağacı rekolte tahmini yapılmıştır. Ataç, bu çalışmada RGB görüntüleri kullanarak çeşitli görüntü işleme teknikleri ile elma rekoltesini %85'in üzerinde bir doğrulukla tahmin etmeyi başarmıştır. Ataç'ın çalışması, görüntü işleme ve yapay zeka tekniklerini kullanarak tarımda rekolte tahminine yönelik önemli bir katkı sunmaktadır (Ataç, Ş. (2023). *Yapay zekâya dayalı görü teknikleri kullanılarak elma ağacı rekolte tahmini* (Master's thesis, Karamanoğlu Mehmetbey Üniversitesi).

Benzer şekilde, Yalçın Işık ve arkadaşlarının limon ağaçları üzerinde yapmış oldukları çalışmada, bilgisayarlı görü teknikleri kullanılarak meyve rekoltesi tahmini yapılmıştır. Bu çalışmada, limon ağaçlarından alınan görüntüler yapay sinir ağına (ANN) aktarılmış ve bu teknikle ağaçtaki limonların toplam ağırlığı %72,8 oranında doğru tahmin edilmiştir. Limonlar toplandıktan sonra ise %99,13 oranında başarı elde edilmiştir (Işık, Y., Ünay, M., & Kayabaşı, A. (2022). Bilgisayarlı görü teknikleri kullanılarak yapay zeka temelli limon ağacı rekolte tahmini. *Mühendislik Bilimleri ve Araştırmaları Dergisi*, 4(1), 80-88.)

Varjovi ve Talu'nun kayısı çalışmasında, görüntü işleme teknikleri kullanılarak kayısı bahçelerinde yıllık rekolte tahmini yapılmıştır. Yerden 17 metre yükseklikten çekilen görüntüler, alan genişletme ve Gaussian karışım modeli gibi yöntemler ile işlenmiştir. Bu çalışmada yapay sinir ağlarına dayalı bir model ile kayısı rekoltesi %77 doğruluk oranıyla tahmin edilmiştir. Özellikle, meyve ağaçlarının otomatik olarak bölütlenmesi ve alan hesaplamaları gibi teknikler ile tahmin sonuçları elde edilmiştir. (Varjovi, M. H., & Talu, M. F. (2016). Kayısı için otomatik rekolte tahmin sistemi. In *International conference on artificial intelligence and data processing IDAP*.)

1.2. Amaç ve Hedefler

Araştırma önerisinin amacı ve hedefleri açık, ölçülebilir, gerçekçi ve araştırma süresince ulaşılabilir nitelikte olacak şekilde yazılır.

Makine öğrenmesi, son yıllarda tarım sektörü başta olmak üzere birçok alanda yaygın kullanım alanı bulan bir disiplindir. Tarımsal üretimde, özellikle verim tahmini, hastalık tespiti ve kalite kontrol süreçlerinde sağladığı yüksek doğruluk ve hız ile öne çıkmaktadır. Geleneksel yöntemlere kıyasla makine öğrenmesi algoritmaları, büyük veri setleri ile çalışabilme, karmaşık ilişkileri analiz etme ve belirli bir örüntüye dayalı olarak gelecek tahminleri yapma gibi avantajlar sunar. Görüntü işleme teknikleri ise, görsel verilerin analizini mümkün kılarak, tarımsal ürünlerin sayımı, sağlık durumu değerlendirmesi gibi süreçlerde hassasiyet sağlamaktadır. ML destekli görüntü işleme sistemleri, sensörlerden alınan çok boyutlu verilerle desteklendiğinde doğruluğu daha da artarak çiftçilere ve tarım sektörü paydaşlarına büyük kolaylık sağlamaktadır.

a) Gradient Boosting:

Gradient Boosting, birçok zayıf tahmin modelinin birleşimiyle güçlü bir model oluşturarak veri analizinde doğruluğu artıran bir ansambl algoritmasıdır. Model, her yeni tahmin modelinin, önceki modelin hatalarını öğrenerek gelişmesini sağlar. Bu ardışık öğrenme süreci, özellikle hacimsel ve sayısal verilerin analizinde doğruluk oranını artırır. Rekolte tahmininde Gradient Boosting, meyve sayısını ve toplam ürün hacmini tahmin etmek için güçlü bir araç olarak kullanılacaktır.

b) Random Forest:

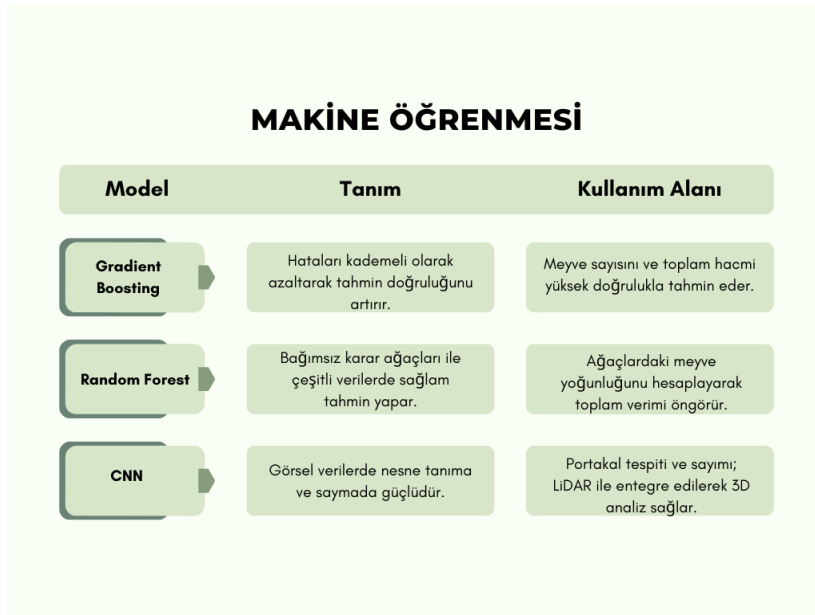
Random Forest, bağımsız karar ağaçlarından oluşan bir ansambl modelidir ve ağaç tabanlı algoritmalar arasında yüksek doğruluğu ve dayanıklılığı ile öne çıkar. Her bir karar ağacı, farklı veri alt kümeleri üzerinde eğitilerek bir tahmin sağlar; bu tahminler, tüm ağaçların çoğunluk oyuyla birleşerek sonuca ulaşır. Modelin doğası gereği aşırı öğrenmeye (overfitting) karşı daha dayanıklıdır ve veri çeşitliliği ile geniş hacim veriler üzerinde sağlam tahminler yapar. Bu model, portakal bahçelerindeki ağaçların meyve yoğunluğu ve toplam hacmini tahmin etmede verimli bir yöntemdir.

c) Convolutional Neural Networks (CNN):

Convolutional Neural Networks, özellikle görsel verilerin işlenmesi ve analizinde yaygın olarak kullanılan derin öğrenme tabanlı bir makine öğrenmesi modelidir. Görüntülerdeki öznitelikleri algılayarak portakal meyvelerinin tespit edilmesi ve sayılması için uygundur. CNN, meyve tespiti ve sayımında karmaşık sahnelerde bile yüksek doğruluk oranına ulaşmamıza olanak tanır. Portakal ağaçlarında yaprak ve dal yoğunluğu gibi karmaşık yapılar arasında meyveleri ayırt edebilmek, bu modelin öne çıkan yeteneklerindendir. CNN ayrıca LiDAR verileriyle entegre edildiğinde, ağaçların 3D yapısal analizine katkıda bulunarak daha hassas tahminler yapılmasını sağlar.

Bu modellerin bir arada kullanımı, proje kapsamında rekolte tahmini yapılırken görsel ve hacimsel verilerin bütünsel analizini sağlayarak yüksek doğruluğa ulaşmamıza yardımcı olacaktır.

Öğrenme türlerine ait açıklamalar aşağıdaki Görsel 1’de yer almaktadır.



Görsel 1) Makine Öğrenmesi Türleri (Görsel araştırmacı tarafından oluşturulmuştur)

1. **Veri Toplama:** Portakal bahçelerinden yüksek çözünürlüklü RGB görüntüleri ve LiDAR sensörlerinden 3D veri toplanacaktır. Bu veri seti, proje boyunca kullanılmak üzere hazırlanacaktır.
2. **Mask R-CNN ve LiDAR Entegrasyonu:** Portakal meyvelerini yüksek doğrulukla tespit edebilmek amacıyla Mask R-CNN algoritması LiDAR sensörlerinden alınan 3D verilerle birleştirilecektir. Bu sayede, yoğun yaprak ve dal yapısının bulunduğu sahnelerde dahi meyve tespiti sağlanacaktır.
3. **Makine Öğrenmesi Modeli Oluşturma:** Rekolte tahmini için Gradient Boosting, Random Forest ve Convolutional Neural Networks (CNN) gibi algoritmalarla bir model oluşturulacaktır. Bu model, her bir ağacın meyve sayısını tahmin ederek genel bahçe verimliliğini hesaplayacak şekilde optimize edilecektir.
4. **Veri Analizi ve İşleme:** LiDAR sensörlerinden alınan derinlik ve hacim verileri segmentasyon teknikleriyle işlenecek, her bir ağaç üzerindeki meyve miktarı ve yoğunluğu tahmin edilecektir.
5. **Rekolte Tahmin Sistemi Kurulumu:** Bahçedeki toplam ürün miktarını tahmin eden bir sistem geliştirilecek ve 1000 ağaçlık bir bahçeden alınan 50 ağacın verisi ile tüm bahçenin rekoltesi tahmin edilecektir.
6. **Sonuçların Değerlendirilmesi ve Raporlama:** Tahminlerin doğruluğunu belirlemek için model testleri ve çapraz doğrulama yapılacaktır; çiftçiler ve fabrikalar için elde edilen sonuçlar analiz edilip raporlanacaktır.

Bu hedefler doğrultusunda, tarım sektöründe doğruluk oranı yüksek tahmin sistemleri geliştirilerek, bu sistemin çiftçilerin maliyet hesaplamaları ile fabrikaların üretim süreçlerine sunduğu katkılar analiz edilecektir.

Projemiz ile aynı doğrultuda olan çalışmalar ile projemizin farklılıkları görsel 2’de yer almaktadır.

Karşılaştırma Tablosu

Diğer Projeler	Bizim Projemiz
Tüm bahçedeki ağaçların tahminini sadece birkaç ağaçtan alınan veri ile yapmaz.	Gradient Boosting ve Random Forest gibi modellerle aşırı öğrenmeye karşı dayanıklı yapı sunar.
Genellikle sadece 2D görsel veriler kullanılır; 3D analiz sınırlıdır veya yoktur.	Bahçedeki tüm portakal ağaçları için tahmin yapılacaktır; örneğin 50 ağacın verisi ile tüm bahçenin tahmini.
Portakal sayısı veya toplam hacim için spesifik doğruluk sağlamakta yetersiz kalabilmektedir.	LiDAR sensörleri kullanarak 3D verilerle desteklenmiş tahminler sağlar, doğruluğu artırır.
Farklı çevresel koşullarda (ışık, gölge vb.) performans azalabilmektedir.	Mask R-CNN ve CNN ile portakal meyvesi tespiti ve sayımını yüksek doğrulukla yapar.
Genellikle aşırı öğrenme sorununa karşı sınırlı çözüm yöntemleri içerir.	Çeşitli koşullara dayanıklı model yapısı sayesinde daha tutarlı tahmin sağlar.

Görsel 2) Proje Kıyas Tablosu (Görsel araştırmacı tarafından hazırlanmıştır.)

Özetle bu projenin amacı, portakal bahçelerinde rekolte tahminini yüksek doğrulukla yapabilen, makine öğrenmesi ve görüntü işleme tabanlı bir sistem geliştirmektir. Bu sistem, sadece birkaç ağaçtan alınan görsel ve 3D verilerle tüm bahçenin tahminini gerçekleştirmeyi mümkün kılarak, çiftçilerin maliyet ve verimlilik hesaplamalarını daha doğru hesaplamalarına yardımcı olmaktadır. Aynı zamanda, fabrikaların üretim ve paketleme süreçlerinde doğru planlama yapmalarını sağlayarak tarımsal üretimin sürdürülebilirliğine katkı sağlamayı hedeflemektedir.

2. YÖNTEM

Araştırma önerisinde uygulanacak yöntem ve araştırma teknikleri (veri toplama araçları ve analiz yöntemleri dahil) ilgili literatüre atıf yapılarak açıklanır. Yöntem ve tekniklerin çalışmada öngörülen amaç ve hedeflere ulaşmaya elverişli olduğu ortaya konulur.

Yöntem bölümünün araştırmanın tasarımını, bağımlı ve bağımsız değişkenleri ve istatistiksel yöntemleri kapsamı gerekir. Araştırma önerisinde herhangi bir ön çalışma veya fizibilite yapıldıysa bunların sunulması beklenir. Araştırma önerisinde sunulan yöntemlerin iş paketleri ile ilişkilendirilmesi gerekir.

Proje kapsamında kullanılan Gradient Boosting, Random Forest ve Convolutional Neural Networks (CNN) modelleri, tarım alanında rekolte tahmini için en uygun özellikleri sunmaktadır. Bu modellerin sağladığı avantajlar, projenin yüksek doğruluk ile sonuçlanmasına katkıda bulunmaktadır. Modellerin avantajlarını şu şekilde özetleyebiliriz:

1. **Gradient Boosting:**

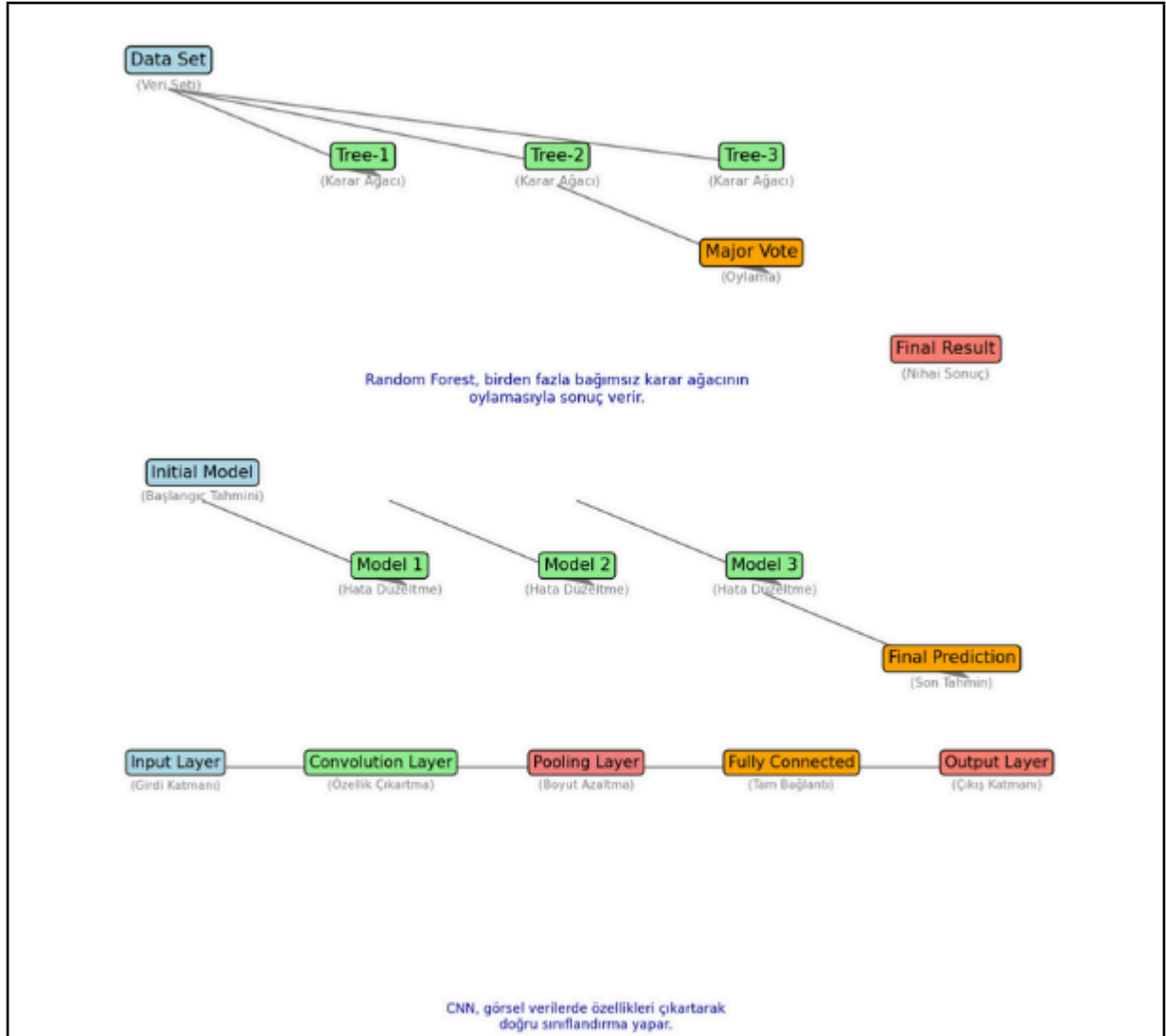
Gradient Boosting, ardışık öğrenme yaklaşımıyla her bir modelin hatalarından öğrenerek doğruluğu artırır. Bu model, verinin karmaşık yapısındaki ilişkileri çözme yeteneğine sahip olup, özellikle meyve sayısı ve hacim tahminlerinde yüksek performans gösterir. Gradient Boosting'in avantajları arasında doğruluğu artıran, hata oranını düşüren ve büyük veri setlerinde bile etkin bir şekilde çalışabilen yapısı bulunmaktadır.

2. **Random Forest:**

Random Forest, birden fazla bağımsız karar ağacından oluşan bir ansambl yöntemidir. Ağaç tabanlı bir algoritma olduğundan, veri çeşitliliğini ve ağaçlar arasındaki bağımsızlığı artırarak aşırı öğrenmeye karşı dayanıklılık gösterir. Bu model, özellikle veri çeşitliliğinin yüksek olduğu ve farklı çevresel koşulların mevcut olduğu portakal bahçelerinde sağlam tahminler yapılmasını sağlar. Hem sayısal hem de kategorik verilerle çalışabilmesi ve genelleme kapasitesinin yüksek olması, Random Forest'in tarımsal veri analizlerinde tercih edilmesini sağlar.

3. **Convolutional Neural Networks (CNN):**

CNN, görsel veri işleme için geliştirilmiş bir derin öğrenme modelidir. Portakal bahçelerinde meyve sayımı ve tespiti için yüksek doğruluk sunar ve görüntülerdeki nesneleri ayırt etme yeteneği çok güçlüdür. CNN, portakal meyvelerinin yaprak ve dalların arasında dahi tespit edilmesini sağlarken, LiDAR verileri ile birleştirildiğinde üç boyutlu analiz imkanı sunarak daha detaylı bir hacim tahmini sağlar. Modelin görsel veriler üzerinde anormallik tespiti yapabilmesi ve farklı çevresel koşullara uyum sağlayabilmesi, proje kapsamında portakal sayısı ve yoğunluğunu doğru şekilde tahmin etmemize katkı sağlar.



Görsel 3) Sırasıyla Gradient Boosting, Random Forest ve Convolutional Neural Network (CNN) modellerinin yapıları (Görsel araştırmacı tarafından oluşturulmuştur.)

A. Veri Toplama

Proje için gereken veriler, portakal bahçelerinden elde edilecek RGB görüntüleri ve LiDAR sensörlerinden toplanacak 3D nokta bulutu verileriyle sağlanacaktır. RGB görüntüleri, ağaçlardaki meyve sayısını belirlemeye yönelik görsel bilgi sağlarken, LiDAR sensör verileri ağaçların hacmini ve yaprak-dal yoğunluğunu analiz etmek için kullanılacaktır. Bu iki veri türü, portakal ağaçlarında yüksek doğrulukla rekolte tahmini yapılabilmesi için temel veri setini oluşturacaktır. LiDAR sensör verileri, nokta bulutu formatında olup, hacimsel analiz için 3D modeller oluşturmakta kullanılacaktır.

B. Veri İşleme ve Analiz

Toplanan veriler, modelin doğruluğunu artırmak için ön işlemden geçirilecektir. İlk olarak, RGB görüntüler Mask R-CNN gibi nesne tespiti ve segmentasyon teknikleriyle işlenerek portakal meyvelerini tespit edecek şekilde hazırlanacaktır. Mask R-CNN algoritması, portakal meyvelerini yüksek doğrulukla tanıyacak ve her bir meyveye maske uygulayarak sayımı kolaylaştıracaktır. Ayrıca, LiDAR verileri RANSAC algoritması gibi segmentasyon yöntemleriyle analiz edilerek yaprak, dal ve gövde gibi farklı yapılar birbirinden ayrılacak ve bu bölümlerin hacimleri hesaplanacaktır. Bu aşamada veri seti, eğitim ve test olmak üzere ikiye ayrılacaktır. Eğitim seti modelin öğrenmesi için, test seti ise modelin doğrulama sürecinde kullanılacaktır.

- **B.1) Pandas:** Veri düzenleme ve analiz işlemlerinde kullanılacak. Toplanan veriler üzerinde temizleme, filtreleme ve dönüştürme işlemleri yapılacaktır.
- **B.2) OpenCV:** Görüntü işleme ve nesne tespiti süreçlerinde etkili bir kütüphane olarak kullanılacaktır. RGB görüntüler üzerinde portakal meyvelerinin tespiti için kullanılacaktır.

2209/A ÜNİVERSİTE ÖĞRENCİLERİ ARAŞTIRMA PROJELERİ DESTEĞİ PROGRAMI
ARAŞTIRMA ÖNERİSİ FORMU

- **B.3) Scikit-Learn:** Veri ön işleme, eğitim-test ayrımı ve model değerlendirme adımlarında kullanılacaktır. Ayrıca, modelin performansını ölçmek için gerekli metrikleri sağlayacaktır.

RGB_Görsel_Sayımı	LiDAR_Hacim	Tahmini_Rekolte
120.0	15.3	500.0
115.0	15.1	520.0
130.0	15.4	530.0
140.0	15.2	510.0
135.0	15.3	525.0

görsel 4) Örnek veri seti görüntüsü (Görsel araştırmacı tarafından oluşturulmuştur.)

C. Makine Öğrenmesi Modeli Geliştirme ve Eğitim

Portakal bahçelerinde rekolte tahmini için Gradient Boosting, Random Forest ve Convolutional Neural Networks (CNN) gibi makine öğrenmesi algoritmaları kullanılacaktır. Gradient Boosting, farklı ağaç tabanlı modellerin bir araya gelmesiyle ardışık öğrenme gerçekleştirir ve doğruluğu artırır. Random Forest ise bağımsız karar ağaçları ile veri çeşitliliğini değerlendirerek sağlam tahminler yapar. CNN algoritması ise özellikle görsel verilerde portakal meyvesi tespiti ve sayımında etkili olup, LiDAR verileri ile birleştirildiğinde üç boyutlu analiz yapabilme kapasitesine sahiptir.

Eğitim sürecinde, modellerin performansını optimize etmek için hiperparametre ayarları yapılacaktır, öğrenme hızı ve ağaç sayısı gibi parametreler, model doğruluğunu maksimize edecek şekilde belirlenecektir. Modelin aşırı öğrenme (overfitting) yapmasını engellemek için çapraz doğrulama (cross-validation) ve erken durdurma (early stopping) yöntemleri uygulanacaktır. Eğitim sürecinin sonunda, modelin doğruluğu test veri seti üzerinde değerlendirilecektir. Tüm süreç boyunca, eksik veya hatalı verilerin etkisini en aza indirmek için veri temizleme teknikleri kullanılacaktır.

D. Sonuçların Değerlendirilmesi ve Raporlama

Elde edilen sonuçlar analiz edilerek projenin başarı düzeyi değerlendirilecektir. Modelin doğruluk oranı, portakal bahçesindeki toplam rekolte tahmininin doğruluğu ve farklı çevresel koşullarda elde edilen sonuçların kararlılığı gibi ölçütler incelenecektir. Proje sonucunda geliştirilen sistem, tarım sektörü paydaşlarına yönelik bir rapor halinde sunulacak ve elde edilen bulgular akademik yayınlar ile paylaşılacaktır.

3 PROJE YÖNETİMİ

3.1 İş- Zaman Çizelgesi

Araştırma önerisinde yer alacak başlıca iş paketleri ve hedefleri, her bir iş paketinin hangi sürede gerçekleştirileceği, başarı ölçütü ve araştırmanın başarısına katkısı "İş-Zaman Çizelgesi" doldurularak verilir. Literatür taraması, gelişme ve sonuç raporu hazırlama aşamaları, araştırma sonuçlarının paylaşımı, makale yazımı ve malzeme alımı ayrı birer iş paketi olarak gösterilmemelidir.

Başarı ölçütü olarak her bir iş paketinin hangi kriterleri sağladığında başarılı sayılacağı açıklanır. Başarı ölçütü, ölçülebilir ve izlenebilir nitelikte olacak şekilde nicel veya nitel ölçütlerle (ifade, sayı, yüzde, vb.) belirtilir.

İŞ-ZAMAN ÇİZELGESİ (*)

İP No	İş Paketlerinin Adı ve Hedefleri	Kim(ler) Tarafından Gerçekleştirileceği	Zaman Aralığı (... Ay)	Başarı Ölçütü ve Projenin Başarısına Katkısı
1	Literatür Tarama ve Veri Toplama	proje yürütücüsü & proje araştırmacısı	1-2 ay	Projenin teorik altyapısının oluşturulması ve veri gereksinimlerinin karşılanması.
2	Model Geliştirme ve Makine Öğrenmesi Modellerinin Seçimi	proje yürütücüsü & proje araştırmacısı	3-4 ay	Tahmin doğruluğunun sağlanması ve projede kullanılacak algoritmaların belirlenmesi.
3	Model Eğitimi ve Veri İşleme	proje yürütücüsü & proje araştırmacısı	4-6 ay	Modellerin verimli şekilde çalıştığının kanıtlanması ve test sonuçlarının gözlemlenmesi.
4	Test ve Doğrulama	proje yürütücüsü & proje araştırmacısı & proje danışmanı	6-7 ay	Modelin gerçek ortamda uygulanabilirliğinin doğrulanması.
5	Sonuçların Değerlendirilmesi ve Raporlama	proje yürütücüsü & proje araştırmacısı & proje danışmanı	7-8 ay	Projenin tüm çıktılarının özetlenmesi ve yayınlanabilir bir rapor haline getirilmesi.

(*) Çizelgedeki satırlar ve sütunlar gerektiği kadar genişletilebilir ve çoğaltılabilir.

2209/A ÜNİVERSİTE ÖĞRENCİLERİ ARAŞTIRMA PROJELERİ DESTEĞİ PROGRAMI
ARAŞTIRMA ÖNERİSİ FORMU

3.2 Risk Yönetimi

Araştırmanın başarısını olumsuz yönde etkileyebilecek riskler ve bu risklerle karşılaşıldığında araştırmanın başarıyla yürütülmesini sağlamak için alınacak tedbirler (B Planı) ilgili iş paketleri belirtilerek ana hatlarıyla aşağıdaki Risk Yönetimi Tablosu'nda ifade edilir. B planlarının uygulanması araştırmanın temel hedeflerinden sapmaya yol açmamalıdır.

RİSK YÖNETİMİ TABLOSU*

İP No	En Önemli Riskler	Risk Yönetimi (B Planı)
1	Model eğitimi için yeterli veri toplanamaması veya verinin düşük kaliteye sahip olması	Alternatif veri kaynakları araştırılacak ve uzman görüşlerinin alınması.
2	Modellerin düşük doğruluk göstermesi ve tahmin gücünün yetersiz kalması	Modellerin parametre ayarları yeniden düzenlenecek; farklı algoritmalar (ör. XGBoost gibi) denenecek ve model iyileştirme çalışmaları yapılacaktır. Gerekirse daha fazla veri toplanacaktır.

(*) Tablodaki satırlar gerektiği kadar genişletilebilir ve çoğaltılabilir.

3.3. Araştırma Olanakları

Bu bölümde projenin yürütüleceği kurum ve kuruluşlarda var olan ve projede kullanılacak olan altyapı/ekipman (laboratuvar, araç, makine-teçhizat, vb.) olanakları belirtilir.

ARAŞTIRMA OLANAKLARI TABLOSU (*)

Kuruluşta Bulunan Altyapı/Ekipman Türü, Modeli (Laboratuvar, Araç, Makine-Teçhizat, vb.)	Projede Kullanım Amacı
Tarsus üniversitesi bilgisayar laboratuvarı	Araştırma ve kodlama çalışmaları için.
Lidar sensörü	Daha doğru 3D görüntü elde etmek için.

(*) Tablodaki satırlar gerektiği kadar genişletilebilir ve çoğaltılabilir.

4. YAYGIN ETKİ

Önerilen çalışma başarıyla gerçekleştirildiği takdirde araştırmadan elde edilmesi öngörülen ve beklenen yaygın etkilerin neler olabileceği, diğer bir ifadeyle yapılan araştırmadan ne gibi çıktı, sonuç ve etkilerin elde edileceği aşağıdaki tabloda verilir.

ARAŞTIRMA ÖNERİSİNDEN BEKLENEN YAYGIN ETKİ TABLOSU

Yaygın Etki Türleri	Önerilen Araştırmadan Beklenen Çıktı, Sonuç ve Etkiler
Bilimsel/Akademik (Makale, Bildiri, Kitap Bölümü, Kitap)	Uluslararası yayın yapan bir dergide makale olarak yayınlanması, üniversiteler ve araştırma merkezleri ile işbirliği yapılarak tarımsal verim tahmini alanında yeni projelerin geliştirilmesine öncülük edebilecek potansiyeli olması.

2209/A ÜNİVERSİTE ÖĞRENCİLERİ ARAŞTIRMA PROJELERİ DESTEĞİ PROGRAMI
ARAŞTIRMA ÖNERİSİ FORMU

Ekonomik/Ticari/Sosyal (Ürün, Prototip, Patent, Faydalı Model, Üretim İzni, Çeşit Tescilli, Spin-off/Start- up Şirket, Görsel/İşitsel Arşiv, Envanter/Veri Tabanı/Belgeleme Üretimi, Telif Konu Olan Eser, Medyada Yer Alma, Fuar, Proje Pazarı, Çalıştay, Eğitim vb. Bilimsel Etkinlik, Proje Sonuçlarını Kullanacak Kurum/Kuruluş, vb. diğer yaygın etkiler)	Geliştirilecek olan yüksek doğruluklu rekolte tahmin sistemi, portakal bahçelerinde verimliliği artırmaya yönelik yenilikçi bir ürün ve prototip olarak sunulacaktır. Tarım sektöründe maliyetlerin azaltılmasına ve kaynak kullanımının daha doğru hesaplanmasına olanak tanıyarak çiftçilerin gelirini artıracak ve sürdürülebilir üretime katkı sağlayacaktır. Ayrıca, proje sonuçları tarım sektöründeki firmaları, kooperatifler ve üniversiteler gibi kurumlar tarafından kullanılabilir niteliktedir.
Araştırmacı Yetiştirilmesi ve Yeni Proje(ler) Oluşturma (Yüksek Lisans/Doktora Tezi, Ulusal/Uluslararası Yeni Proje)	Proje verileri, yüksek lisans ve doktora seviyesindeki öğrenciler için tez çalışmaları olarak değerlendirilebilirken, makine öğrenmesi, görüntü işleme ve tarımsal veri analizi konularında uzmanlaşmak isteyen araştırmacılar için de eğitim ve gelişim imkanı sunacaktır. Ayrıca, proje sonuçları temel alınarak, tarımda veri işleme ve yapay zeka kullanımını daha da ileri taşıyacak yeni araştırma projelerinin ortaya çıkması teşvik edilecektir.

5. BÜTÇE TALEP ÇİZELGESİ

Bütçe Türü	Talep Edilen Bütçe Miktarı (TL)	Talep Gerekçesi
Sarf Malzeme	-	Bulunmamaktadır.
Makina/Teçhizat (Demirbaş)	920 TL	Görsel verileri güçlendirmek için ihtiyaç olan LİDAR sensörü
Hizmet Alımı	-	Bulunmamaktadır.
Ulaşım	-	Bulunmamaktadır.
TOPLAM	920 TL	Görsel verileri güçlendirmek için ihtiyaç olan LİDAR sensörü

NOT: Bütçe talebiniz olması halinde hem bu tablonun hem de TUBİTAK Yönetim Bilgi Sistemi (TYBS) başvuru ekranında karşınıza gelecek olan bütçe alanlarının doldurulması gerekmektedir. Yukardaki tabloda girilen bütçe kalemlerindeki rakamlar ile, TYBS başvuru ekranındaki rakamlar arasında farklılık olması halinde TYBS ekranındaki veriler dikkate alınır ve başvuru sonrasında değiştirilemez.

6. BELİRTMEK İSTEDİĞİNİZ DİĞER KONULAR

Sadece araştırma önerisinin değerlendirilmesine katkı sağlayabilecek bilgi/veri (grafik, tablo, vb.) eklenebilir.

--

7. EKLER

EK-1: KAYNAKLAR