



**T.C.
ONDOKUZ MAYIS ÜNİVERSİTESİ
LİSANSÜSTÜ EĞİTİM ENSTİTÜSÜ
TARIM MAKİNELERİ VE TEKNOLOJİLERİ MÜHENDİSLİĞİ
ANA BİLİM DALI**

**BUĞDAY ÇEŞİTLERİNİN YAPAY ZEKA YÖNTEMLERİ
KULLANILARAK SINIFLANDIRILMASI**

Doktora Tezi

Mahtem Teweldemedhin MENGSTU

Danışman
Prof. Dr. Alper TANER

SAMSUN
2024

T.C.
ONDOKUZ MAYIS ÜNİVERSİTESİ
LİSANSÜSTÜ EĞİTİM ENSTİTÜSÜ
TARIM MAKİNELERİ VE TEKNOLOJİLERİ MÜHENDİSLİĞİ
ANA BİLİM DALI



**BUĞDAY ÇEŞİTLERİNİN YAPAY ZEKA YÖNTEMLERİ
KULLANILARAK SINIFLANDIRILMASI**

Doktora Tezi

Mahtem Teweldemedhin MENGSTU

Danışman

Prof. Dr. Alper TANER

Bu çalışma Ondokuz Mayıs Üniversitesi tarafından PYO.ZRT.1904.22.032 proje numarası ile desteklenmiştir.

SAMSUN
2024

TEZ KABUL VE ONAYI

Mahtem Teweldemedhin MENGSTU tarafından, **Prof. Dr. Alper TANER** danışmanlığında hazırlanan “**BUĞDAY ÇEŞİTLERİNİN YAPAY ZEKA YÖNTEMLERİ KULLANILARAK SINIFLANDIRILMASI**” başlıklı bu çalışma, jürimiz tarafından 5.9.2024 tarihinde yapılan sınav sonucunda oy birliği ile başarılı bulunarak Doktora Tezi olarak kabul edilmiştir.

	Unvanı Adı Soyadı Üniversitesi Ana Bilim/Ana Sanat Dalı	Sonuç
Başkan	Prof. Dr. Sefa TARHAN Tokat Gaziosmanpaşa Üniversitesi Makine Mühendisliği Ana Bilim Dalı	☒ Kabul ☐ Ret
Üye	Prof. Dr. Alper TANER Ondokuz Mayıs Üniversitesi Tarım Makineleri ve Teknolojileri Mühendisliği Ana Bilim Dalı	☒ Kabul ☐ Ret
Üye	Prof. Dr. Tanzer ERYILMAZ Yozgat Bozok Üniversitesi Tarım Makineleri ve Teknolojileri Mühendisliği Ana Bilim Dalı	☒ Kabul ☐ Ret
Üye	Dr. Öğr. Üyesi İdris SANCAKTAR Ondokuz Mayıs Üniversitesi Elektrik-Elektronik Mühendisliği Ana Bilim Dalı	☒ Kabul ☐ Ret
Üye	Dr. Öğr. Üyesi Hüseyin DURAN Ondokuz Mayıs Üniversitesi Tarım Makineleri ve Teknolojileri Mühendisliği Ana Bilim Dalı	☒ Kabul ☐ Ret

Bu tez, Enstitü Yönetim Kurulunca belirlenen ve yukarıda adları yazılı jüri üyeleri tarafından uygun görülmüştür.

Prof. Dr. Ahmet TABAK
Enstitü Müdürü

BİLİMSEL ETİĞE UYGUNLUK BEYANI

Hazırladığım Yüksek Lisans tezinin bütün aşamalarında bilimsel etiğe ve akademik kurallara riayet ettiğimi, çalışmada doğrudan veya dolaylı olarak kullandığım her alıntıya kaynak gösterdiğimi ve yararlandığım eserlerin Kaynaklar'da gösterilenlerden oluştuğunu, her unsurun enstitü yazım kılavuzuna uygun yazıldığını ve TÜBİTAK Araştırma ve Yayın Etiği Kurulu Yönetmeliği'nin 3. bölüm 9. maddesinde belirtilen durumlara aykırı davranılmadığını taahhüt ve beyan ederim.

Etik Kurul Gerekli mi ?

Evet ☐ (Gerekli ise ekler kısmına ekleyiniz)

Hayır ☒

18 /06/ 2024

Mahtem Teweldemedhin MENGSTU

TEZ ÇALIŞMASI ÖZGÜNLÜK RAPORU BEYANI

Tez Başlığı : BUĞDAY ÇEŞİTLERİNİN YAPAY ZEKA YÖNTEMLERİ KULLANILARAK SINIFLANDIRILMASI

Yukarıda başlığı belirtilen tez çalışması için şahsım tarafından 19/06/2024 tarihinde intihal tespit programından alınmış olan özgünlük raporu sonucunda;

Benzerlik oranı : % 6

Tek kaynak oranı : % 1 çıkmıştır.

19 /06 / 2024

Prof. Dr. Alper TANER

ÖZET

BUĞDAY ÇEŞİTLERİNİN YAPAY ZEKA YÖNTEMLERİ KULLANILARAK SINIFLANDIRILMASI

Mahtem Teweldemedhin MENGSTU

Ondokuz Mayıs Üniversitesi

Lisansüstü Eğitim Enstitüsü

Tarım Makineleri ve Teknolojileri Mühendisliği Ana Bilim Dalı

Doktora, Eylül/2024

Danışman: Prof. Dr. Alper TANER

Üretim ve işleme aşamasında çiftçilerin, sanayicilerin ve tüketicilerin standart oluşturabilmek için ürün çeşitlerini özel olarak bilmeleri gerekmektedir. Çeşitlerin manuel olarak tanımlanması zor ve tutarsız olup tamamen insan becerilerine bağlıdır. Bu sorunları çözmek için süreci otomasyona dönüştürecek yenilikçi teknolojiler ortaya çıkmıştır. Bu çalışmada, 124 ekmeklik buğday çeşidi kullanılmıştır. Bu çeşitlerin çoklu görüntüleri alınarak 465.000 görüntüden oluşan bir veritabanı oluşturulmuştur. Çalışmada, makine öğrenmesi (Machine Learning, ML), derin öğrenme (Deep Learning, DL) ve Fourier Dönüşümlü Kızılötesi (Fourier Transform Infrared, FTIR) spektroskopisi olmak üzere üç uygulama gerçekleştirilmiştir. Her üç uygulamada da, sınıflandırma modellerinin geliştirilmesi, sınıflandırılacak çeşit sayısının artırılmasını, uygun özelliklerin çıkarılmasını, optimum özellik kombinasyonunun bulunmasını, modellerin eğitilmesini, ince ayar yapılmasını ve son olarak en iyi modellerin değerlendirilmesini içermektedir.

Birinci uygulamada, ML yöntemleri kullanılarak yönlendirilmiş gradyan histogramı (Histogram of Oriented Gradient, HOG) ve renk özelliklerinin çıkarılması, makine öğrenme modellerinin eğitimi, geliştirilmesi ve test edilmesi ele alınmıştır. Bu çalışmada, buğday çeşitlerini sınıflandırmak için HOG ve renk özelliklerine dayalı Yapay Sinir Ağları (Artificial Neural Network, ANN), Destek Vektör Makineleri (Support Vector Machine, SVM), Random Forest Sınıflandırıcısı (Random Forest Classifier, RFC) ve K-En Yakın Komşu Algoritması (K-Nearest Neighbor, KNN) modelleri geliştirilmiştir. Temel Bileşen Analizi (Principal Component Analysis, PCA) ile eğitilen en iyi modellerin sınıflandırma performansları kabul edilebilir seviyede ANN ve SVM modelleri için sırasıyla %96.8 ve %98.1 elde edilmiştir. Sonuçlar, oldukça fazla sayıda çeşit olmasına rağmen, HOG ve renk özelliklerinin başarılı ML modelleri geliştirmek için kullanılabileceğini göstermektedir.

İkinci uygulamada, DL yöntemleri kullanılarak yeni bir Evrimsel Sinir Ağı (Convolutional Neural Network, CNN) modeli, DenseNet201, MobileNet ve InceptionV3 popüler mimarileri, öğrenme aktarımı (Transfer Learning) yoluyla eğitilmiş ve sırasıyla %95.40, %92.41, %90.54 ve %83.47 doğruluk elde edilmiştir. Önerilen CNN modeli, DenseNet201 ve MobileNet modelleri başarılı bulunmuştur. Çalışmada modellerin optimum parametrelerini elde etmek için daha fazla ince ayar yapılması ve diğer popüler CNN mimarilerinin araştırılması önerilebilir.

Üçüncü uygulamada, FTIR ve ML modellerinin farklı spektral bölgelere göre uygulanması araştırılmıştır. FTIR spektrumları ve ML modellerinin entegrasyonu ile seçilen SVM modelinin elde ettiği %98.8'lik doğruluk, aynı zamanda ML modellerinin spektral temsillere dayalı dört buğday çeşidini tanıma yeteneğini ortaya koymuştur.

Anahtar Sözcükler: Ekmeklik Buğday Çeşitleri, Yapay Zeka, Makine Öğrenmesi, Derin Öğrenme, Görüntü İşleme

ABSTRACT

CLASSIFICATION OF WHEAT VARIETIES USING ARTIFICIAL INTELLIGENCE METHODS

Mahtem Teweldemedhin Mengstu

Ondokuz Mayıs University

Institute of Graduate Studies

Department of Agricultural Machinery and Technologies Engineering

Ph.D., September/2024

Supervisor: Prof. Dr. Alper TANER

At the production and processing stage, farmers, industrialists and consumers need to have specific knowledge of crop varieties in order to establish standards. Manual identification of varieties is difficult and inconsistent and entirely dependent on human skills. To solve these problems, innovative technologies have emerged to automate the process. In this research, 124 bread wheat varieties were used. Multiple-view images of these varieties were taken and an image database of 465,000 images was created. In this study, Machine Learning (ML), Deep Learning (DL) and Fourier Transform Infrared (FTIR) spectroscopy based three methods were used to develop classification models. In all three applications, the development of classification models involved increasing the number of varieties to be classified, extracting appropriate features, finding the optimal combination of features, training and fine-tuning the models, and finally evaluating the best models.

First, extraction of histogram of oriented gradient (HOG) and color features, training, development and testing of ML models were performed. Artificial Neural Network (ANN), Support Vector Machine (SVM), Random Forest Classifier (RF) and K-Nearest Neighbor Algorithm (KNN) models based on HOG and color features were developed to classify wheat varieties. The classification performances of the best models trained with Principal Component Analysis (PCA) were effective at 96.8% and 98.1% rates for ANN and SVM models, respectively. The results show that HOG and color features can be used to develop successful ML models, although there are quite a number of variants.

Second, using DL methods, a new Convolutional Neural Network (CNN) model from scratch, and popular architectures DenseNet201, MobileNet and InceptionV3 were trained by transfer learning and achieved 95.40%, 92.41%, 90.54% and 83.47% classification accuracies, respectively. The proposed CNN model, along with the DenseNet201 and MobileNet models, has been found successful. Further fine-tuning and exploring other popular CNN architectures can be recommended to obtain the optimal parameters of the models in the study.

Third, the application of FTIR and ML models with different spectral regions was investigated. The 98.8% accuracy achieved by the SVM model demonstrated the ability of ML models to recognize four wheat varieties based on spectral representations.

Keywords: Bread Wheat Varieties, Artificial Intelligence, Machine Learning, Deep Learning, Image Processing

TEŞEKKÜR

Öncelikle, hayatımdaki harika armağanları ve lütfu için Tanrı'ya şükrediyorum. Doktora çalışmam boyunca beni kabul eden ve bana yol gösteren değerli danışmanım Prof. Dr. Alper TANER'e teşekkür ederim. Onun mentörlüğü, sabrı, cömertliği ve nezaketi için her zaman minnettar olacağım.

Tez izleme komitesi üyelerim Dr. Öğr. Üyesi Hüseyin DURAN ve Dr. Öğr. Üyesi İdris SANCAKTAR'a değerli zamanları ve rehberlikleri için özellikle teşekkür ederim. Türkiye Cumhuriyeti tarafından sunulan YTB bursu yükseköğrenim destek programı için de minnettarım. Aynı zamanda TÜBİTAK kurumuna desteği için teşekkür ederim. Ayrıca Tarım Makineleri ve Teknolojileri Mühendisliği Bölümündeki tüm hocalarıma yardımları ve nezaketleri için teşekkür ederim. Bölümümdeki Doktora ve Yüksek Lisans öğrencisi arkadaşlarıma da çok teşekkür ederim.

Bunlarla birlikte, çalışmalarım boyunca duaları, sevgileri, ilhamları ve destekleri için tüm aileme teşekkür ederim.

Mahtem Teweldemedhin MENGSTU

İÇİNDEKİLER

TEZ KABUL VE ONAYI	i
BİLİMSEL ETİĞE UYGUNLUK BEYANI	ii
TEZ ÇALIŞMASI ÖZGÜNLÜK RAPORU BEYANI	ii
ÖZET	iii
ABSTRACT	iv
TEŞEKKÜR	v
İÇİNDEKİLER	vi
SİMGELER VE KISALTMALAR	ix
ŞEKİLLER DİZİNİ	x
TABLolar DİZİNİ	xii
1. GİRİŞ	1
1.1. Buğday Yetiştiriciliğinin Önemi	1
1.2. Tarım ve Makine Öğrenmesi	4
1.3. Bilgisayarla Görme	5
1.4. Özellik Çıkarma	7
1.4.1. Renk özellikleri	7
1.4.1.1. Renk uzayı	7
1.4.1.1.1. RGB	8
1.4.1.1.2. HSV	8
1.4.1.1.3. Karşıt renk uzayı	8
1.4.1.1.4. CIE L*a*b* ve CIE L*u*v*	8
1.4.1.2. Renk özelliklerini temsil etme yöntemleri	9
1.4.1.2.1. Renk histogramı	9
1.4.1.2.2. Renk momentleri	9
1.4.1.2.3. Renk korelogramı	9
1.4.1.2.4. Renk uyumluluk vektörü	10
1.4.2. Doku özellikleri	10
1.4.2.1. Dokusal özellikleri temsil etme yöntemleri	10
1.4.2.2. Tamura özellikleri	11
1.4.2.3. Eş oluşum matrisleri	11
1.4.2.4. Gabor filtreleri	11
1.4.2.5. Dalgacık dönüşümü	11
1.4.2.6. Yerel ikili örüntü	12
1.4.2.7. Yerel ikili örüntünün diğer formları	12
1.4.3. Şekil özellikleri	12
1.4.4. Şekil özelliklerini temsil etme yöntemleri	13

1.5. Araştırmanın Hedefleri	14
1.6. Çalışmanın Kapsamı ve Sınırlılıkları	15
2. KAYNAK ÖZETLERİ	16
3. MATERYAL VE YÖNTEM.....	26
3.1. Buğday Çeşitleri	26
3.2. Görüntü Alma	29
3.2.1. Kamera	29
3.2.2. Görüntü alma yazılımı.....	30
3.2.3. Görüntü elde etme süreci	30
3.3. Görüntü Ön İşleme	35
3.4. Özellik Çıkarma.....	35
3.4.1. Yönlendirilmiş gradyan histogramı	36
3.4.2. Renk özellikleri	36
3.5. Sınıflandırma Yaklaşımları.....	38
3.5.1. Makine öğrenme modelleri	38
3.5.1.1. Yapay sinir ağları.....	39
3.5.1.2. Destek vektör makineleri.....	40
3.5.1.3. Random forest sınıflandırıcısı	41
3.5.1.4. K-en yakın komşu algoritması.....	42
3.5.1.5. Veri seti.....	42
3.5.1.6. Renk ve dokuya dayalı model eğitimi	43
3.5.2. Derin öğrenme modelleri	44
3.5.2.1. Önerilen CNN modeli.....	44
3.5.2.2. Öğrenme aktarımı	45
3.5.2.3. MobileNet.....	45
3.5.2.4. DenseNet201	46
3.5.2.5. InceptionV3	47
3.5.2.6. Veri seti.....	47
3.5.3. Fourier dönüşümlü kızılötesi spektroskopisi	47
3.5.3.1. FTIR ve ML modeli.....	49
3.5.3.2. Girdi verileri	50
3.5.3.3. Temel düzeltme	52
3.5.3.4. Aşırı örnekleme	52
3.5.3.5. Temel bileşen analizi	53
3.5.3.6. Spektral verilere dayalı model eğitim yaklaşımı	53

3.5.3.7. Hiperparametre ayarlama	54
3.6. Performans Ölçütleri.....	55
3.7. Makine Öğrenme Sınıflandırıcılarının İstatistiksel Analizi.....	55
3.8. Paketler ve İşlemciler	56
4. BULGULAR	58
4.1. Görüntü Elde Etme	58
4.2. Özellik Çıkarma.....	59
4.2.1. HOG özellikleri.....	59
4.2.2. Renk momentleri.....	59
4.3. Makine Öğrenme Modelleri	60
4.3.1. Yapay sinir ağları	60
4.3.2. Destek vektör makineleri	65
4.3.3. Random forest sınıflandırıcısı	66
4.3.4. K-en yakın komşu algoritması	67
4.4. Derin Öğrenme Modelleri	69
4.4.1. CNN modeli	69
4.4.2. MobileNet	70
4.4.3. DenseNet201	71
4.4.4. InceptionV3.....	72
4.5. FTIR Tabanlı Makine Öğrenme Modelleri	73
4.5.1. Veri hazırlama ve temizleme	73
4.5.2. Baz çizgisi düzeltilmesi.....	75
4.5.3. FTIR analizi verilerine göre makine öğrenme modellerinin eğitimi	76
4.5.3.1. FTIR tabanlı yapay sinir ağları modeli.....	76
4.5.3.2. FTIR tabanlı destek vektör makineleri	79
4.5.3.3. FTIR tabanlı random forest sınıflandırıcısı	81
4.5.3.4. FTIR tabanlı K-en yakın komşu algoritması	84
4.5.4. Genel istatistiksel değerlendirme ve seçilen modeller	86
5. TARTIŞMA	88
6. SONUÇLAR VE ÖNERİLER	96
KAYNAKLAR	99
ÖZ GEÇMİŞ.....	113

SİMGELER VE KISALTMALAR

ANN	: Yapay Sinir Ağları
CNN	: Evrişimsel Sinir Ağları
DL	: Derin Öğrenme
FTIR	: Fourier Dönüşümlü Kızılötesi
GLCM	: Gri Seviye Eş-Oluşum Matrisi
HOG	: Yönlendirilmiş Gradyan Histogramı
KNN	: K-En Yakın Komşu
LBP	: Yerel İkili Desenler
LDA	: Doğrusal Diskriminant Analizi
MAE	: Ortalama Mutlak Hata
ML	: Makine Öğrenmesi
MLP	: Çok Katmanlı Algılayıcı
NIR	: Yakın Kızılötesi
PCA	: Temel Bileşen Analizi
RFC	: Random Forest Sınıflandırıcısı
RGB	: Kırmızı Yeşil Mavi
SVM	: Destek Vektör Makineleri
SWIR	: Kısa Dalga Kızılötesi
VNIR	: Görünür Yakın Kızılötesi

ŞEKİLLER DİZİNİ

Şekil 3.1. En büyük 14 buğday üreticisi ülke (2022).....	2
Şekil 3.2. Türkiye'de buğday üretimi (2016-2023).....	2
Şekil 3.3. Görüntü özellikleri ve tanımlayıcıları.....	13
Şekil 3.1 Foldio-360 görüntü alma cihazı.....	29
Şekil 3.2. Canon EOS 2000D kamera.....	29
Şekil 3.3. (a) Foldio-360 platformu ile kamera, (b) Buğday tanesinin yerleştirildiği döner tabla, (c) Görüntü alma sisteminin genel görünümü.....	30
Şekil 3.4. Farklı pozisyonlarda çekilen buğday tane örnekleri.....	31
Şekil 3.5. Görüntü alma sürecinin şematik diyagramı. (a) Uzaktan kumanda, (b) Kamera, (c) Foldio-360 cihazı, (d) Foldio-360 mobil uygulaması, (e) EOS yardımcı yazılımı, (f) Görüntülerin kaydedildiği klasör.....	31
Şekil 3.6. Çalışmada kullanılan ekmeklik buğday çeşitlerine ait örnek görüntüler .	34
Şekil 3.7. (a) Orijinal görüntü (4000x4000), (b) Kırpılmış ve yeniden boyutlandırılmış görüntü (250x250).....	35
Şekil 3.8. Buğdayın HOG özelliklerine ait bir örnek.....	36
Şekil 3.9 Buğday görüntüsünün RGB renk uzayında renk momentleri dağılımı	38
Şekil 3.10. Yapay sinir ağları yapısı.....	40
Şekil 3.11 . Destek vektör makineleri yapısı.....	41
Şekil 3.12. Random forest sınıflandırıcısının bir diyagramı.....	41
Şekil 3.13. KNN algoritması.....	42
Şekil 3.14. Makine öğrenme modellerinin diyagram gösterimi.....	43
Şekil 3.15. CNN modelinin yapısı.....	45
Şekil 3.16. MobileNet Yapısı.....	46
Şekil 3.17. DenseNet Yapısı.....	46
Şekil 3.18. InceptionV3 Yapısı.....	47
Şekil 3.19. Fourier Dönüşümlü Kızılötesi Spektroskopisi.....	49
Şekil 3.20. Dört buğday çeşidinin spektroskopik verileri, (a) Ön işlemeden sonra modeli eğitmek için kullanılan standart okuma, (b) Hatalı okumalara sahip çıkarılmış örnekler, (c) Dört buğday çeşidinin örnek FTIR spektral okuması.....	50
Şekil 3.21. Dört buğday çeşidinin tam FTIR spektral verileri.....	51
Şekil 3.22. FTIR spektral verilerinin dört bölümünün on beş kombinasyonu.....	51
Şekil 3.23. Orijinal ve baz çizgisi düzeltilmiş spektrumlardan bir örnek.....	52
Şekil 3.24. Makine öğrenme modellerinin eğitim süreci.....	54
Şekil 3.25. Grid arama hiperparametre ayarlaması ve K-katlı çapraz doğrulama	54
Şekil 4.1. Bir buğday tanesinin bir dakikada çekilen çoklu görünüşleri (75 adet)...	58

Şekil 4.2. Buğday tanelerinin HOG özelliklerinin çıkarılmasının görselleştirilmesi.	59
Şekil 4.3. Buğday çeşitlerinin renk momentlerindeki varyasyonun görselleştirilmesi	60
Şekil 4.4. ANN modelinin test doğruluğu.....	62
Şekil 4.5. (a) PCA'sız eğitilen gizli katmanlarında farklı nöron boyutlarına sahip beş modelin eğitim ve doğrulama doğruluk eğrileri, (b) Eğitim ve doğrulama kayıp eğrileri	63
Şekil 4.6. (a) PCA' ile eğitilen farklı nöron boyutlarına sahip beş modelin eğitim ve doğrulama doğruluk eğriler, (b) Eğitim ve doğrulama kayıp eğrileri.....	64
Şekil 4.7. Önerilen CNN modelinin eğitim ve doğrulama setleri için doğruluk ve kayıp eğrileri.	70
Şekil 4.8. MobileNet modelinin eğitim ve doğrulama setleri için doğruluk ve kayıp eğrileri.	71
Şekil 4.9. DenseNet201 modelinin eğitim ve doğrulama setleri için doğruluk ve kayıp eğrileri.	72
Şekil 4.10. InceptionV3 modelinin eğitim ve doğrulama setleri için doğruluk ve kayıp eğrileri.	73
Şekil 4.11. Buğday çeşitlerinin orijinal FTIR spektral okuması.	74
Şekil 4.12. Tam spektrumların dört bölümü	74
Şekil 4.13. Tüm örneklerin temel çizgi düzeltilmiş spektrumları.....	75
Şekil 4.14. Seçilen ANN modelinin hata matrisi	79
Şekil 4.15. Seçilen SVM modelinin hata matrisi	81
Şekil 4.16. Seçilen RFC modelinin hata matrisi	84
Şekil 4.17. Seçilen KNN modelinin hata matrisi	86

TABLÖLAR DİZİNİ

Tablo 3.1. Çalışmada kullanılan ekmeçlik buğday çeşitleri	26
Tablo 3.2. Makine öğrenme modellerini eğitmek için kullanılan veri setinin özeti ..	43
Tablo 3.3. Çoklu veri kümeleri ve model performansları örneği.....	56
Tablo 3.4. TRUBA'da kullanılan hesaplama kaynaklarının özellikleri	57
Tablo 4.1. ANN modellerinin hiperparametreleri.....	61
Tablo 4.2. On ANN modelinin performans metrikleri.....	61
Tablo 4.3. SVM modellerinin performans metrikleri	66
Tablo 4.4. RFC modellerinin hiperparametreleri	66
Tablo 4.5. RFC modellerinin performans metrikleri	67
Tablo 4.6. KNN modellerinin hiperparametreleri.....	68
Tablo 4.7. KNN modellerinin performans metrikleri	69
Tablo 4.8. Önceden eğitilmiş ve önerilen CNN modellerinin performansları	73
Tablo 4.9. Dört buğday çeşidinin FTIR spektral veri setleri.....	74
Tablo 4.10. Farklı spektral bölgeler ve bunların kombinasyonları kullanılarak ANN modellerinin eğitim ve test doğruluk oranları (%).....	77
Tablo 4.11. ANN modellerinin hiperparametreleri.....	77
Tablo 4.12. FTIR-tabanlı ANN model performans metrikleri	78
Tablo 4.13. Farklı spektral bölgeler ve bunların kombinasyonları kullanılarak SVM modellerinin eğitim ve test doğruluk oranları (%).....	79
Tablo 4.14. SVM modellerinin hiperparametreleri.....	80
Tablo 4.15. FTIR-tabanlı SVM model performans metrikleri	80
Tablo 4.16. Farklı spektral bölgeler ve bunların kombinasyonları kullanılarak RFC modellerinin eğitim ve test doğruluk oranları (%).....	82
Tablo 4.17. RFC modellerinin hiperparametreleri	82
Tablo 4.18. FTIR-tabanlı RFC model performans metrikleri	83
Tablo 4.19. Farklı spektral bölgeler ve bunların kombinasyonları kullanılarak KNN modellerinin eğitim ve test doğruluk oranları (%).....	84
Tablo 4.20. KNN modellerinin hiperparametreleri.....	85
Tablo 4.21. FTIR-tabanlı KNN model performans metrikleri	85
Tablo 4.22. Seçilen modellerin performans metrikleri	87
Tablo 5.1. İlgili çalışmaların ve önerilen modellerin özeti	93

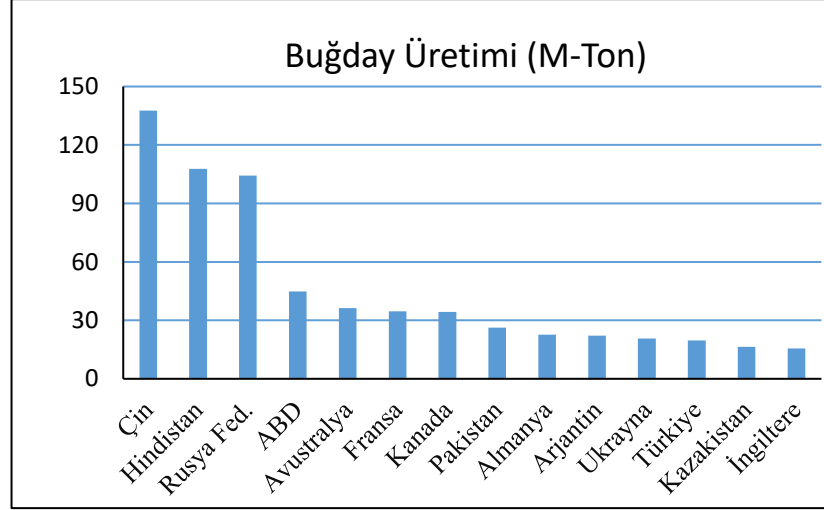
1. GİRİŞ

1.1. Buğday Yetiştiriciliğinin Önemi

Buğday (*Triticum*), Poaceae familyasına ait bir tahıl türüdür. Buğday, dünya nüfusunun yaklaşık %40'ı için temel bir besin kaynağıdır. Buğday yetiştiriciliği tarihin çok eski dönemlerine kadar uzanmaktadır (Giraldo et al., 2019). Buğday, 8.000 yıl boyunca Avrupa, Batı Asya ve Kuzey Afrika'daki büyük medeniyetlerin ana gıdası olmuştur. Bunun nedeni muhtemelen buğdayın tarımsal adaptasyon kabiliyeti, depolama kolaylığı ve birçok gıda yapımı için una dönüştürülebilme kolaylığıdır (Dixon, 2009).

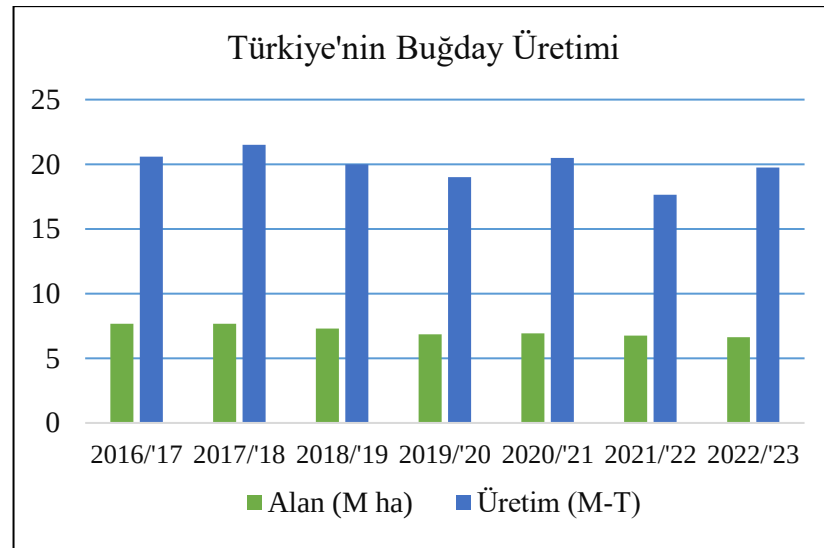
Buğday öncelikle bir enerji (karbonhidrat) kaynağı olarak kabul edilir ve bu açıdan kesinlikle önemlidir. Günlük protein ve gıda kalorisinin %20'sini sağlayarak insan beslenmesinde merkezi bir yere sahiptir. Bununla birlikte, sağlıklı bir diyet katkıda bulunabilecek lif ve lipidler, vitaminler, mineraller ve fitokimyasallar gibi küçük bileşenler de dahil olmak üzere önemli miktarda diğer önemli besin maddelerini de içerir (Shewry ve Hey, 2015). Mısır ve pirinçten sonra insan besin zincirindeki en önemli üçüncü tahıldır. Büyük üretim hacmi, tanelerin kalitesini değerlendirmek ve endüstride verimliliği artırmak için etkili yöntemlere ihtiyaç duymaktadır (Velesaca et al., 2021).

2023 yılında Türkiye'de 22 milyon ton buğday üretilmiştir (TÜİK, 2023). Türkiye, Şekil 1.1'de gösterildiği gibi dünyada buğday üreten ilk 15 ülke arasında yer almaktadır (FAO, 2022). Ekim alanı bakımından, Türkiye'de ilk sırada yer alan buğday üretimi için 6.8 milyon hektar arazi kullanılmaktadır (TÜİK, 2023). Buğday, beslenmedeki önemi ve ekonomik değeri nedeniyle Türkiye'de her zaman önemli ve stratejik bir ürün olmuştur (Cançelik et al., 2020; Keser, 2022).



Şekil 3.1. En büyük 14 buğday üreticisi ülke (2022)

2050 yılına kadar dünya buğday arzını yaklaşık %70 oranında artırma ihtiyacı öngörülürken, küresel buğday üretimi, ürün yetiştirme alanlarının genişletilmesi için kısıtlamalar ve mevcut doğal kaynakların sınırlı olması nedeniyle önemli zorluklarla karşı karşıyadır. Ayrıca küresel ısınma, artan sıcaklık ve kuraklık vakaları ile ilişkilendirilmekte ve gıda güvenliğini riske atmaktadır (Mylonas et al., 2020). Nüfus baskısından kaynaklanan artan buğday talebini karşılamak ve çiftlik gelirlerini artırmak için aşağıdakiler kullanılmalıdır: Geliştirilmiş yüksek verimli çeşitler, buğday üretiminin bir Tahıl Zincirine dönüşmesi, bir denetim sistemi ve ürün sertifikasyon programlarını içeren bir kalite değerlendirme süreci ve teknolojisi gerektirmiştir. Nihai amaç, piyasada daha fazla miktarda yüksek kaliteli buğday ürünü bulundurmaktır (Laabassi, 2021). Türkiye'nin son yedi yıllık buğday üretimi Şekil 1.2'de verilmiştir.



Şekil 3.2. Türkiye'de buğday üretimi (2016-2023)

Son birkaç on yılda tarımsal üretimde çok sayıda teknolojik dönüşüm yaşanmış ve sektör daha gelişmiş ve teknoloji tabanlı hale gelmiştir. Modern teknolojiler, insan toplumuna küresel gıda talebini karşılamak için yeterli gıda üretme olanağı ve kapasitesi sağlamaktadır.

Dünyanın her yerindeki gıda ihtiyacını karşılamak için insanlar, dünya buğday üretiminin yaklaşık %95'ini oluşturan hekzaploid ekmeklik buğday (*Triticum aestivum*) başta olmak üzere yaklaşık 30.000 buğday çeşidi yetiştirmiştir (Peng et al., 2011). Modern buğday ıslah programlarının uygulanması, yeni buğday çeşitlerinin seçilmesini ve ıslah edilmesini hızlandırmıştır. Sadece Türkiye'de, 1928 yılından 2020 yılına kadar 617 buğday çeşidi piyasaya sürülmüştür. 2016-2020 döneminde en fazla çeşit piyasaya süren ülke Türkiye olurken, Türkiye'yi İngiltere, Macaristan ve İran takip etmiştir (Keser, 2022). Bu programlar buğday endüstrisinin verimliliğini artırırken, buğday çeşitlerinin sık sık karışmasına da neden olmuştur. Günümüzde, buğday çeşitlerinin verimli, basit ve doğru bir şekilde tanımlanması önemli bir konudur. Bu konudaki derinlemesine araştırmalar, buğday çeşitlerinin mülkiyet haklarının korunmasını teşvik etmekte ve yüksek kaliteli buğday çeşitlerinin seçimi ve tanıtımı için büyük önem taşımaktadır (Gao et al., 2012).

Hangi çeşidin yetiştirileceğinin seçilmesi, tarım endüstrisi alanındaki en önemli faktörlerden biridir. Maksimum verimi elde etmek için tahıl çeşitlerinin etkin bir şekilde tanımlanması gerekir. Bu nedenle, her bir tahılda, en uygun çeşidi seçmek için verim, hastalıklara karşı direnç, mantarlara karşı direnç ve ürün kalitesi gibi önemli özelliklerin her birine göre puanlanmalıdır (Velesaca et al., 2021).

Bunu akılda tutarak, hasat sonrası faaliyetleri ele aldığımızda, ele alınması ve çözülmesi gereken ciddi sorunlar vardır. Hasattan sonra, gıda ürünleri farklı parametrelere göre ayrılır ve farklı sınıflarda derecelendirilir. Tahılların çeşitleri birbirlerine o kadar benzemektedir ki tanımlanmaları oldukça zordur. Sadece endüstri düzeyinde değil, laboratuvar düzeyinde bile çeşitlerin tanımlanması ve derecelendirilmesi kolay bir iş değildir. Modern tohum endüstrisinde, büyük ölçekli tohumların sürekli dolaşımı nedeniyle, taşıma, depolama ve işleme sırasında farklı tahıl tohum çeşitleri arasında kazara karışım meydana gelme olasılığı yüksektir ve bu da tahıl kalitesini ve verimini kesinlikle düşürmektedir (Grantina-Ievina, 2019). Kapsamlı uzman bilgisi gerektiren manuel denetim, öznel ve zaman alıcı bir prosedürdür (Khoshroo, 2014). Öte yandan, bilgisayarla görme teknolojisi son

zamanlarda tahıl çeşitlerini tanımak için otomatik ve öznel bir yol olarak geliştirilmiştir.

Tarımsal ürünlerin pazarlanmasında üretici, sanayici ve tüketiciler için söz konusu ürünlerin çeşitlerini bilmek oldukça önemlidir. Üreticiler, doğru tarım yapabilmek için ettikleri ürünün çeşidini bilmek isterler, benzer şekilde pazarlamacılar da hedef pazarlara yönelik standartları oluşturabilmek için sattıkları ürün çeşidinden emin olmak isterler. Bu nedenlerle çeşitlerin belirlenmesi için güvenilir yöntemlere ihtiyaç duyulmaktadır (Taner et al., 2018).

Çeşitlerin tanımlanması, tanelerin görünümü hakkında bazı bilgiler gerektirir ve tane görünümüne ilişkin bilgilerle desteklenir. Tane görünümündeki karışıklık, özellikle buğday genotiplerinin sayısının çok fazla olduğu Türkiye'de sıklıkla karşılaşılan bir sorundur (Olgun et al., 2016).

Yukarıdaki konuyla ilgili olarak, insan operasyonları tutarsız ve daha az etkili olduğundan, manuel operasyonlara yardımcı olmak veya manuel operasyonları otomatik sistemlerle değiştirmek için endüstrileri ve laboratuvarları donatmaya yönelik yenilikler artmaktadır. Bu nedenle, tahılları tanımlamak ve farklı çeşitler halinde sınıflandırmak için hızlı bir yaklaşım ve sistem geliştirmek önemlidir.

1.2. Tarım ve Makine Öğrenmesi

Tarım, tarımsal ürünlerin ve çiftlik hayvanlarının yetiştirilmesi süreci olup, faaliyetleri arasında tarımsal ürünlerin ve çiftlik hayvanlarının üretimi, işlenmesi ve pazarlanması yer alır. Tarım, küresel ekonomiye katkıda bulunan başlıca unsurlardan biridir. İnsanların temel ihtiyacı olan gıdayı sağlaması ve insanlığın hayatta kalmasında vazgeçilmez bir rol oynaması nedeniyle diğer ekonomik faaliyetlerden farklı olarak değerlendirilmektedir. Hemen hemen her yerde ve ağırlıklı olarak gelişmekte olan ülkelerde istihdam yaratmada büyük rol oynamaktadır. Üretim aşamasından başlayarak tarımın farklı aşamalarında toprak, hava durumu, özellikle yağmur, ürün çeşitleri, ürünlerin verim potansiyeli, pestisit uygulamaları hakkında bilgi çok önemlidir. Ayrıca, işleme aşamasında, uygun hasat sonrası teknolojilerin eksikliği ve pazar eğilimleri tarımın genel verimliliğini ve potansiyelini olumsuz yönde etkilemekte ve küresel nüfus artışına bağlı olarak sürekli artan gıda talebinin yanı sıra önemli kayıplara ve acil öngörülemeyen sorunlara yol açmaktadır (Godfray et al., 2010).

Üreticiler, her zaman daha verimli hale gelerek kârlılıklarını korumak için verimliliklerini artırmaya çalışırlar. Tarım, neyin yetiştirileceği, hangi çeşidin ekileceği, ne zaman ekim yapılacağı, hangi gübrelerin uygulanacağı, nasıl ve ne kadar sulama yapılacağı, ürünlerin ne zaman hasat edileceği ve pazarlanacağı konularında doğru kararlar vermeyi gerektirir. Tüm bu kararları verirken hava durumu, hastalık, değişen tüketici zevk ve tercihleri gibi çok çeşitli karmaşık dış etkilerin etkisi altındadır.

Tarımla ilgili her şey hakkında veri ve bilgiye sahip olmak, tarımla uğraşanların verimliliği artırmak için veriye dayalı kararlar almasına yardımcı olur. ML, DL, bulut bilişim ile desteklenen nesnelerin interneti (IoT) gibi veri ile ilgili yenilikler, kaliteyi artırmak ve nihayetinde tarımın kârlılığını artırmak için bilgi edinme ve bunları gereksinimlerine göre işleme potansiyeline sahiptir (Meshram et al., 2021). ML teknikleri temelde, belirli bir görevi yerine getirmek için deneyimlerden veya eğitim verilerinden öğrenmeye yönelik bir öğrenme prosedürünü içerir. Modelleri eğitmek için kullanılan veriler bir dizi örnek içerir. Normalde bir örnek, değişken olarak adlandırılan ve nominal, sıralı veya sayısal olabilen bir dizi parametre ile karakterize edilir. Eğitilmiş bir modelin performansı veya ne kadar iyi çalıştığı çeşitli istatistiksel ve matematiksel ölçütlerle ölçülür. Öğrenme sürecinden ve belirli bir tatmin edici performans seviyesine ulaşıldıktan sonra, eğitilen model, eğitim sürecinde kazanılan beceriyi kullanarak yeni görünmeyen verileri tahmin etmek, sınıflandırmak veya kümelemek için kullanılabilir (Liakos et al., 2018).

1.3. Bilgisayarla Görme

Bilgisayar veya makine görmesi, bilgisayarları insan benzeri görsel yeteneklerle donatan, koşulları algılayıp yorumlayan, nesneleri tanıyan ve çeşitli algoritmaların yardımıyla karar veren yapay zeka tabanlı bir mekanizmadır (Owino, 2023). Görüntülerden ve videolardan anlamlı bilgiler çıkarma ve gizli kalıpları tanıma yeteneği, makinelerin temelde insan zekası ve manuel çalışma gerektiren görevleri yerine getirmesini sağlayan yaklaşımdır.

Makine görme, hesaplama, matematik, mühendislik ve yaşam bilimleri de dahil olmak üzere birden fazla disiplinin entegrasyonu ile görsel bir durumun göreceli olarak algılanmasını kapsar (Jurj et al., 2020). DL algoritmaları ve bilgisayarla görme mimarilerindeki sürekli artan yenilikler, bilgisayarla görmenin geniş çapta yayılmasını

büyük ölçüde desteklemiştir. Görsel veri sınıflandırma, nesne algılama ve segmentasyon, makine görmenin ana uygulamalarıdır. Verilerdeki artış, daha hızlı öğrenmeyi kolaylaştırmak için yeni ve sağlam algoritmaların geliştirilmesi, hesaplama gücüne kolay erişim, özel tasarlanmış işlemciler, bilgisayarla görmenin çeşitli alanlarda araştırılmasına ve uyarlanabilirliğine büyük katkı sağlamıştır (Vazquez-Arellano et al., 2016; Mochida et al., 2019).

Görüntü verisi toplama için sensörlerle birleştirilmiş bilgisayarla görme yöntemleri, otomotiv, sağlık ve tarım gibi birçok endüstride kullanılan esnek ve uyarlanabilir mekanizmalardır. Tarım endüstrisinde, bilgisayarlı görme, hayvancılıkta hayvanların refahını izlemek için kullanılmakta, bitkisel üretimde fenotipleme, tür tanıma, çeşit sınıflandırması, verim tahmini, büyüme modelleri ve stres izleme, erken hastalık tespiti, yabancı ot tanımlama, gübre ve pestisitlerin robotik uygulaması, ürün derecelendirme ve sınıflandırma, ayrıca gıda endüstrisinde de geniş bir uygulama yelpazesine sahiptir (Gomes, 2012).

Düşük ekipman ve hesaplama maliyetlerinin avantajları ve tahribatsız gıda değerlendirme tekniklerine yönelik artan talep nedeniyle, makine görmesi ve görüntü tabanlı çözümlerin kullanımı hızlanmaktadır (Mahajan et al., 2015). Makine görmesi birçok zorlukla karşı karşıya olsa da, avantajları geleneksel yaklaşımlardan daha ağır basmaktadır (Barbedo, 2016). Makine görme algoritmalarının ve sensörler, görüntüler ve spektroskopi gibi çeşitli parametre üreten yöntemlerin entegre uygulaması, gıda değerlendirmesiyle ilgili tarımsal zorlukların üstesinden gelmek için yaygın olarak kullanılmaktadır (Vithu ve Moses, 2006). Yukarıda bahsedilen yaklaşımların uygulanması, uygulayıcıların tanımlayıcı veriler üretmesini, modeller geliştirmesini ve nihayetinde gereksiz ve sıkıcı faaliyetleri otomatikleştirerek bir sistemin genel teknik ve ekonomik verimliliğini artırmasını sağlar (Patricio ve Rieder, 2018).

Bir bütün olarak tarım ve onun gölgesi altındaki tüm alanlar, organizma, toprak ve atmosferik koşullar gibi karmaşık varlıklarla ilgilendiği için son derece karmaşık sistemlerle karakterize edilir. Yine de, yapay zeka teknolojisi tam olarak yaygınlaştırılmamış ve tarımın her alanında uygulanmamıştır ve hala önemli perspektiflerin ve zorlukların ele alınması gerekmektedir (Shamshiri et al., 2018). Kamuya açık farklı türde büyük veri kümelerinin eksikliği nedeniyle, şimdiye kadar yapılan araştırmalar, araştırmacılar tarafından sınırlı çeşit ve alanlarda hazırlanan verilere dayanmaktadır. Bu nedenle, yapay zeka tekniklerinin geliştirilmesi ve

benimsenmesine paralel olarak, büyük ve çeşitli veri tabanları oluşturmak ve makine görmesinin uygulanabileceği potansiyel alanları keşfetmek, yaygın ve gelecekteki araştırma çalışmaları olmaya devam etmektedir (Sharif et al., 2018; Tian et al., 2020).

1.4. Özellik Çıkarma

Daha önce tartışılan bilgisayar görmesi konusuyla bağlantılı olarak, özellik çıkarma, küresel bilgi teknolojisinin hızla genişlemesi ve yapay zekanın artan popülaritesi ile son yıllarda birçok sektördeki uygulamalar için çok önemli olmuştur. Özellik çıkarma yöntemi, görüntüden mümkün olduğunca fazla bilgi elde etmek için kullanılır. ML disipliniinde özellik çıkarma, büyük miktarda ham veriyi daha küçük ve yönetilebilir temsillere indirgeyen önemli bir prosedürdür. Bu büyük veri kümelerindeki çok sayıda değişken, işlemek için çok sayıda hesaplama kaynağı gerektirir. Özellik çıkarma, orijinal veri setini tam ve doğru bir şekilde tanımlarken analiz edilmesi gereken veri miktarını en aza indiren değişkenleri seçme ve özellikler halinde birleştirme tekniklerini ifade eder. Özellik çıkarmanın genel amacı, bu sayılardan anlamlı örüntüler çıkarmaktır. Bu örüntüler, renklerde, dokularda ve şekillerde bulunabilir. Bu örüntülerden elde edilen ölçümler düşük seviyeli özellikler olarak adlandırılır (Mehmood et al., 2014).

1.4.1. Renk özellikleri

Renk, bir duyu sisteminin (insan gözü veya yapay sensör) ışığa verdiği tepkiye ve ışığın nesnelerle etkileşimine bağlı bir gözlemdir. Renk, ışığın yüzey spektral yansıması ve sensör duyarlılığının bir sonucudur. Görüntü boyutu ve yönünden bağımsız olduğu için renk, arka plan karmaşıklığına karşı nispeten sağlam kabul edilir. Renk özelliği çıkarma sürecindeki önemli hususlar arasında renk uzayı, renk kuantizasyonu ve benzerlik fonksiyonunun seçimi yer alır. Bir görüntünün renk özellikleri kullanılarak tanımlanması, uygulanacak renk uzayının belirlenmesini gerektirir. Bir 3D renk uzayında, verilen görüntünün her pikseli tek bir nokta olarak temsil edilir.

1.4.1.1. Renk uzayı

Görüntülerin dijital olarak işlenmesinde çeşitli renk uzayları kullanılır. Uygun renk uzaylarını seçmek sadece doğru özellikler elde etmek için değil aynı zamanda rengi doğal görsel sistemin anlayacağı şekilde temsil etmek için de gereklidir. Çokça kullanılanlar şunlardır: RGB, HSV, karşıt renk uzayı, CIE $L^*a^*b^*$ ve CIE $L^*u^*v^*$.

1.4.1.1.1. RGB

RGB (Kırmızı-Yeşil-Mavi, Red, Green, Blue) renk uzayları en yaygın olanıdır. Bu özel renk uzayı ışığın kırmızı, yeşil ve mavi ek ana renklerinden oluşur. Görünür renk spektrumunda göreceli renk elde etmek için, üç rengin değişen seviyeleri eklenir. RGB renk uzayı bir cihaza bağlı değildir ve algısal tekdüzeliğe sahip değildir. Bu, RGB uzayında birbirine nispeten yakın olduğu varsayılan bir rengin aslında insan gözü tarafından yakın olarak değerlendirilmeyebileceği anlamına gelir.

1.4.1.1.2. HSV

HSV, Ton, Doygunluk ve Değer olmak üzere üç ögesiyle tanımlanan bir renk uzayıdır. Ton, Doygunluk ve Değer sırasıyla renk türünü, canlılığı ve parlaklığı temsil eder. Basit bir denklem uygulanarak RGB renk uzayı kolayca HSV renk uzayına dönüştürülebilir. RGB'den farklı olarak HSV algısal olarak tekdüzedir. HSV algısal olarak tekdüze olduğundan, benzer değerlere sahip renkler insan gözü tarafından benzer şekilde görülür. HSV uzayı bilgisayar grafiklerinde, özellikle de kullanıcıların herhangi bir renk örneğini seçmek için renk uzayında gezindiği kullanıcı arayüzlerinde sıklıkla kullanılır.

1.4.1.1.3. Karşıt renk uzayı

İnsanların renkleri görme biçiminin, bazı tonların asla bir arada bulunmadığının düşünüldüğü karşıt bir renk modeli içerdiğine dair kanıtlar vardır. Kırmızı ve sarı, kırmızı ve mavi, yeşil ve sarı, yeşil ve mavi kombinasyonları kolaylıkla gözlemlenirken, kırmızımsı-yeşil ve mavimsi-sarı renk algıları asla gözlemlenmemektedir. Bu keşif, karşıt renk uzayının, rengi karşıt sinyale kodlamak için kullanılabileceği kavramına yol açar. Bu, RGB değerlerinden ($R-G$, $2B-R-G$, $R+B+G$) olarak tanımlanan üç bileşene ayrılmıştır. Bu tür bir gösterimin önemi, üçüncü kanaldaki parlaklık verilerini ayırmasıdır.

1.4.1.1.4. CIE $L^*a^*b^*$ ve CIE $L^*u^*v^*$

Diğer renk uzayları CIE $L^*a^*b^*$ ve CIE $L^*u^*v^*$ 'dir. Bu uzaylar, uzamsal tekdüzelik gerekliliğini yerine getirdikleri için görüntü alma görevi için uygundur. Bu renk temsilleri cihaza bağlı değildir ve algısal olarak tekdüzedir. Modelin öğeleri açıklık (L), kırmızı ve yeşil arasındaki mesafe (a) ve sarı ve mavi arasındaki mesafedir (b). $L^*u^*v^*$ 'de, renk farklılıklarının algılanabilirliği doğrusallaştırılmaya çalışılmıştır.

1.4.1.2. Renk özelliklerini temsil etme yöntemleri

1.4.1.2.1. Renk histogramı

Renk histogramları, görüntülerin renk bilgilerini tasvir etmek için özellik çıkarımında kullanılan birincil araçtır. Renk histogramı, her bir çubuğun belirli bir renk uzayında belirli bir renge karşılık geldiği bir çubuk grafik biçimidir. Renk histogramı, üç farklı renk kanalındaki değerlerin ortak olasılığını istatistiksel olarak yaklaştırmak için uygulanan bir yöntemdir. Veri aralığı eşit büyüklükteki kutulara bölünerek tipik histogram biçimi oluşturulabilir. Bir pikselin belirli bir kutuya düşme olasılığı, daha sonra her kutuya düşen bir görüntüdeki piksellerin renklerini sayarak ve sonuçları toplam puana normalleştirerek her kutu için belirlenir. Renk histogramının piksellerin uzamsal bilgilerini hesaba katmaması temel dezavantajlarından biridir. Sonuç olarak, renk özelliği metrikleri benzer olduğu sürece çok farklı görüntüler bile benzer olarak kabul edilebilir. (Stricker ve Orengo, 1995), tarafından önerilen kümelenmiş renk histogramı, renk histogramı yaklaşımının bir modifikasyonudur. Elde ettikleri bulgular, önerilen yöntemin geleneksel renk histogramı yöntemine göre üstünlüklerini göstermiştir. Bununla birlikte, yöntemin dezavantajı, birden fazla boyutlu histogram olduğunda kutuları sıralamanın açık bir yolunun olmamasıdır.

1.4.1.2.2. Renk momentleri

Niceleme sorunlarının üstesinden gelmek için renk momentleri tekniğinin uygulanması önerilmektedir (Stricker ve Orengo, 1995). Renk olasılık dağılımlarının istatistiksel momentleri olan renk momentleri, özellikle bir resim yalnızca tek bir nesne içerdiğinde, erişim sistemlerinde sıklıkla kullanılmaktadır. Ortalama, varyans ve çarpıklığın renk momentlerinin görüntülerin renk dağılımlarını tanımlamada etkili olduğu gösterilmiştir. Buna ek olarak, renk momentleri, görüntü içindeki rengi çevreleyen uzamsal bilgilerin hiçbirini kodlamadıkları için renk histogramı tekniğiyle aynı sorunlara sahiptir.

1.4.1.2.3. Renk korelogramı

Renk korelogramı, bir görüntünün renk bilgisini yakalamak için kullanılan bir tekniktir (Huang et al., 1997). Renk histogramları ve momentlerinin aksine, bu yaklaşım kodlanmış renk bilgisine uzamsal veri ekler ve bu da önceki gösterimlerle ilgili bazı sorunları çözer. Renk korelogramı, hesaplanmasının kolay olması, renk uzamsal korelasyonunu içermesi ve yerel uzamsal korelasyonun küresel dağılımını

temsil etmesi gibi avantajlara sahiptir.

1.4.1.2.4. Renk uyumluluk vektörü

Pikselleri uzamsal uyumluluklarına göre gruplara ayıran bölünmüş histograma renk uyumluluk vektörü denir (Pass ve Zabith., 1996). Görüntünün pikselleri, daha geniş bir tek tip renk alanının parçası olup olmadıklarına bağlı olarak uyumlu ve uyumsuz olmak üzere iki kategoriye ayrılır. Özellik vektörü daha sonra uyumlu ve uyumsuz pikseller için farklı histogramlar üreterek bazı uzamsal bilgiler içerebilir.

CCV'nin, özellikle ağırlıklı olarak tek tip renk veya çoğunlukla dokusal bölgeler içeren görüntüler için uzamsal bilgisi nedeniyle renk histogramından daha iyi geri getirme sonuçları verdiği kanıtlanmıştır. Ayrıca, HSV renk uzayı hem renk histogramı hem de renk uyumluluk vektörü gösteriminde CIE $L^*u^*v^*$ ve CIE $L^*a^*b^*$ uzaylarından daha iyi performans göstermektedir.

1.4.2. Doku özellikleri

Doku, tek bir renk veya yoğunluğun varlığından bağımsız bir şekilde tekdüzelik sergileyen görsel desenleri tanımlamak için kullanılan bir terimdir. Neredeyse tüm yüzeyler doğal olarak buna sahiptir. Yüzeylerin yapısal olarak nasıl düzenlendiğine ve çevreleriyle nasıl etkileşime girdiklerine dair önemli ayrıntılar sağlar.

1.4.2.1. Dokusal özellikleri temsil etme yöntemleri

İstatistiksel yaklaşımlar, dokuyu karakterize etmek için bir görüntüyü oluşturan piksellerin gri seviyelerinin istatistiksel özelliklerini kullanır. Resimlerde, çeşitli gri seviyeleri, genellikle periyodik olarak görünür. Uzaydaki gri tonlama dağılımı hesaplanır. Doku, yapısal yaklaşımlar tarafından doku elemanlarından oluşacak şekilde tanımlanır. Bu doku elemanları, bazı yerleşik düzenleme standartlarına uygun olarak bir yüzey üzerine düzenli bir şekilde yerleştirilir. Spektral yöntemler, Fourier spektrumunun özelliklerine bağlı olarak Fourier spektrumundaki yüksek enerjili tepe noktaları bularak bir yüzeyin gri seviyelerinin küresel periyodikliğini karakterize eder. İstatistiksel yaklaşımların kullanımı dokuları sınıflandırmak için çok önemlidir. Çünkü doku nitelikleri bu yöntemler aracılığıyla hesaplanır. Tamura özellikleri, eş-oluşum matrisleri ve çok çözünürlüklü filtreleme yöntemleri (Gabor ve dalgacık dönüşümü gibi) dokunun istatistiksel temsillerine birkaç örnektir.

1.4.2.2. Tamura özellikleri

İnsanların dokuyu nasıl algıladığına dair psikolojik çalışmalardan esinlenilerek doku temsili farklı bir bakış açısıyla incelenmiştir (Tamura et al., 1978). Görsel doku özellikleri için matematiksel yaklaşımlar oluşturulmuştur. Kabalık, kontrast, yönlülük, çizgisellik, düzenlilik ve pürüzlülük görsel dokunun altı özelliğidir.

1.4.2.3. Eş oluşum matrisleri

Bu yaklaşıma göre, piksel uzaklığı ve yönelimine dayalı bir eş-oluşum matrisi oluşturulur. Doku temsili daha sonra matristen faydalı bilgilerin çıkarılmasıyla oluşturulur (Haralick et al., 1973). Çok sayıda başka araştırmacı da aynı yaklaşımı benimsemiş ve daha gelişmiş varyantlar önermiştir. (Gotlieb ve Kreyszig, 1990), ilk olarak (Haralick et al., 1973) tarafından belirlenen istatistikleri araştırmış ve kontrast, ters fark momenti ve entropinin en büyük ayırt edici gücü gösterdiğini deneysel olarak keşfetmişlerdir.

1.4.2.4. Gabor filtreleri

Gabor filtreleri olarak bilinen bir dalgacık koleksiyonu, çeşitli çözünürlüklerde ve yönlerde enerji yakalamak için kullanılır. Sonuç olarak, Gabor filtreleri tüm görüntünün yerel enerjisini tespit edebilir. Özellikle, doku özellikleri genellikle Gabor filtresi kullanılarak görüntülerden çıkarılır (Daugman, 1988).

Gabor filtreleri, görsel uzayda ve uzaysal frekansta optimum Heisenberg birleşik çözünürlüğü vermektedir. Bu nedenle, Gabor filtreleri görüntü kodlama, doku segmentasyonu, retina tanıma, belgelerin analizi, hedef tespiti, fraktal boyut ölçümü, çizgi karakterizasyonu, kenar tespiti, görüntü temsili gibi çok çeşitli uygulamalarda etkili bir şekilde kullanılmaktadır.

1.4.2.5. Dalgacık dönüşümü

Dalgacık dönüşümleri, görüntü işlemenin çeşitli alanlarında çok çözünürlüklü yöntem olarak sıklıkla kullanılmaktadır. Görüntü gürültüsünün ortadan kaldırılması, görüntü sıkıştırma, görüntü yeniden yapılandırma ve görüntü geri getirme gibi farklı görevler için bir dizi dalgacık tabanlı yöntem ve kavram ortaya konmuş ve araştırılmıştır. (Suematsu et al., 2002) görüntüleri çıkarmak için çok çözünürlüklü dalgacık dönüşümünü kullanmıştır. Dalgacık özelliklerinin geri getirme doğruluğu çok iyi değildir. Sonuç olarak, daha yüksek bir erişim doğruluğu seviyesine ulaşmak için

dalgacık dönüşümünü kullanmak üzere bir dizi teknik oluşturulmuştur. İçerik tabanlı görüntü erişiminin verimliliğini artırmak için, ayrıık dalgacık katsayılarından hesaplanan ağırlıklar dalgacık özelliklerine eklenir (Mehmood et al., 2014).

1.4.2.6. Yerel ikili örüntü

Yerel bir doku özellikleri tanımlayıcısı olan yerel ikili örüntü (YİÖ), küresel doku tanımlamaları sağlamak için kullanılır. YİÖ ve diğer türevleri dokuları etkili bir şekilde sınıflandırabilmektedir. Tanıtılmasından bu yana, YİÖ önemli ilerlemeler kaydetmiş olup görüntünün aydınlatma değişimlerine karşı duyarsızdır. O zamandan beri başka varyasyonlar da geliştirilmiştir. YİÖ ile, sağlanan görüntünün gri tonlamasındaki yerel değişikliklere karşı değişmeyen dokusal bilgiler elde edilir.

1.4.2.7. Yerel ikili örüntünün diğer formları

Fu ve Wei, (2008) çalışmasında sunulan YİÖ'nün bir türevi olan Merkezi İkili Örüntü, genişletilmiş histogramı en aza indirmek için yerellikteki piksel eşleştirmelerini karşılaştırır ve en büyük ağırlığı merkezi piksele tahsis ederek yerel etkiyi dikkate alır. Ayrıca, tamamlanmış YİÖ olarak bilinen farklı bir tür, yalnızca büyüklük bilgisini kullanan standart YİÖ'nün aksine, çevredeki pikseller ile merkez pikseller arasındaki farkın hem büyüklüğünü hem de işaretini kullanır (Guo ve Zhang, 2010). YİÖ yönsüz birinci dereceden bir türev iken, yerel türev örüntüsü daha yüksek dereceden türevleri kullanır (Zhang et al., 2010). Yüksek dereceli türevler bitişik piksellerin yönünü okuyabilir. Yerel üçlü örüntü, ikili değerler yerine -1, 0 ve 1 değerlerini kullanarak YİÖ'yü artırır (Liao, 2010). Ayrıca, -2, -1, 0, 1 ve 2 olmak üzere beş tanımlayıcı, merkez ve çevre pikseller arasındaki farkın derecesine bağlı olarak yerel dördütlü örüntüde kullanılır (Vipparthi ve Nagar, 2014). YİÖ ve türevlerinin yerel tanımlayıcılar olarak etkili olduğu bulunmuştur. Bu tekniklerin bir dizi doku analizi için geniş uyarlanabilirliği, etkinliklerinin bir kanıtıdır.

1.4.3. Şekil özellikleri

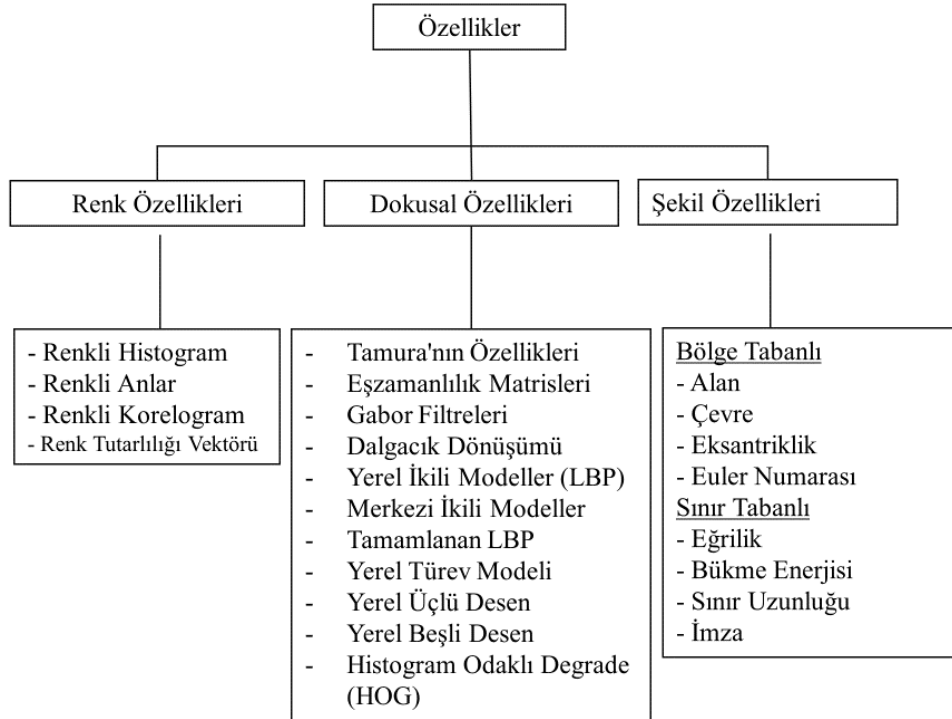
Bir nesnenin şeklini belirlemek çok zordur. İnsanlar şekli genellikle "uzun", "yuvarlak" gibi sıfatlar kullanarak sözlü veya görsel olarak tanımlar. Bilgisayar tabanlı şekil işlemenin son derece karmaşık şekilleri bile doğru bir şekilde tanımlamayı gerektirmesine rağmen, şekilleri tanımlamak için genel kabul görmüş bir metodoloji yoktur. Şekil çok önemli bir görsel unsurdur ve görüntülerin içeriğini tanımlamak için kullanılan temel unsurlardan biridir. Bir nesnenin yönü veya konumu değiştirildiğinde

bile, görüntüde yer alan geometrik bilgi genellikle değişmeden kalır. Özellik çıkarma koşullarında, çevre, alan, eksantriklik, simetri bazı temel şekil özellikleridir.

1.4.4. Şekil özelliklerini temsil etme yöntemleri

Şekil özelliği çıkarımında nesne segmentasyonu ve şekil temsili olmak üzere iki ana aşama söz konusudur. Nesneler, segmente edildiğinde şekil özellikleri temsil edilebilir ve indekslenebilir. Şekil gösterimleri tipik olarak bölge tabanlı ve sınır tabanlı olmak üzere ikiye ayrılır. Bölge tabanlı teknik, alanı, Euler sayısını, eksantrikliği, uzatılmışlığı ve kompaktlığı tanımlamak için tam şekil bölgesini kullanırken, sınır tabanlı yaklaşım, zincir kodlarını, Fourier tanımlayıcılarını ve eğrilik, bükülme enerjisi, sınır uzunluğu, imza gibi basit geometrik sınır temsillerini hesaplamak için şeklin dış çevresini kullanır.

Fourier tanımlayıcıları ve moment değişiklikleri, bu iki kategori için en iyi temsillerdir. Moment değişiklikleri, şekil özelliği olarak dönüşümden bağımsız olan bölge tabanlı momentleri kullanırken, Fourier tanımlayıcısının birincil nosyonu şekil özelliği olarak Fourier dönüşümlü sınırı kullanmaktır (Rui et al., 1996; Persoon ve Fu, 1997). Şekil 1.3 tüm özelliklerin özetini ve temsillerini gösterir.



Şekil 3.3. Görüntü özellikleri ve tanımlayıcıları

1.5. Araştırmanın Hedefleri

Buğday çeşitlerinin kalite değerlendirmesine yardımcı olmak için çeşitli görüntüleme araçları ve teknikleri kullanılmaktadır. Amaç, hem buğday çeşidinin etiketinin belirlenmesinde hem de kalitesinin değerlendirilmesinde hatalara yol açabileceğinden uzman kararlarına olan olası güvensizliği ortadan kaldırmaktır. Bu nedenle, buğday çeşitlerini daha etkili bir şekilde tanımlayabilecek ve uzman kararlarıyla ilişkili riskleri en aza indirebilecek umut verici bir çerçeveye ihtiyaç vardır (Işık et al., 2022). Ayrıca, Türkiye'de mevcut çeşit sayısı düşünüldüğünde, yapılan araştırmalar hem sınırlı olup hem de buğday çeşit sayısı çok azdır. Bu nedenle, bu araştırmanın amaçlarından biri de buğday çeşitlerinin sayısını önemli ölçüde artırmaktır.

Bunun dışında, görüntüler kullanılarak sınıflandırma modelleri geliştirilirken karşılaşılan yaygın zorluklardan biri çekişmeli saldırılardır (adversarial attacks). DL modelleri, giriş görüntülerindeki küçük değişikliklere karşı son derece savunmasızdır ve bu durum nihayetinde yanlış sınıflandırmaya yol açar. Genellikle, giriş görüntüleri belirli bir sınıfın görüntülerini mümkün olduğunca tutarlı hale getirmeye çalışarak tek bir pozisyonda çekilir. Bu nedenle, bu çalışmada bu sorunu ele almak, bu zorluğa dayanacak sağlam bir model geliştirmek ve birden fazla görünümünden çekilen görüntüleri kullanmanın etkisini değerlendirmek için 360 derecelik görüntüler kullanmayı amaçlamaktadır.

Bu nedenlerle, araştırmanın özel hedefleri şunlardır:

1. Buğday çeşitlerinin 360 derecelik görüntü veri setini oluşturmak.
2. Uygun görüntü özelliklerini çıkarmak, geleneksel ML modellerini eğitmek ve performanslarını değerlendirmek.
3. Öğrenme aktarımı (Transfer learning) yoluyla yeni ve popüler önceden eğitilmiş CNN modellerini eğitmek ve değerlendirmek.
4. ML modellerini kullanarak FTIR spektroskopisi tabanlı buğday çeşitleri sınıflandırma modeli geliştirmek.

1.6. Çalışmanın Kapsamı ve Sınırlılıkları

Bu çalışmanın kapsamı coğrafi ve yöntemsel olarak tanımlanmıştır. Sınıflandırma modelini geliştirmek için kullanılan buğday çeşitleri sadece Türkiye'den olup, yabancı çeşitler kullanılmamıştır. Çeşit sayısı ile ilgili olarak, 1928-2020 döneminde piyasaya sürülen 617 çeşitten araştırma enstitüleri tarafından ıslah edilmiş olan 124 çeşit bu çalışmaya dahil edilmiştir ve kullanılan çeşitler sadece ekmeklik buğday çeşitleridir. Ayrıca, FTIR spektroskopisine dayalı sınıflandırma modelinin geliştirilmesi için sadece dört çeşit kullanılmıştır. Orijinalliği sağlamak için yerel pazarlardan hiçbir çeşit kullanılmamıştır. Görüntü tabanlı ve spektroskopi tabanlı sınıflandırma modelleri için kullanılan tüm çeşitler, çeşitleri ıslah eden araştırma enstitülerinden temin edilmiştir.

Seçme ve özellik çıkarma tekniklerinde sadece doku ve renk özelliklerinden yararlanılmıştır. Literatür, kaynak mevcudiyeti, özellikle yüz binlerce görüntüyü çıkarmak için gereken hesaplama gücü ve zaman sınırlamaları göz önüne alındığında yalnızca HOG ve renk momenti özellikleri çıkarılmış ve ML modellerinin eğitimi için girdi olarak kullanılmıştır. Çalışmanın modellerle ilgili kapsamı da ANN, SVM, RFC ve KNN modelleriyle sınırlandırılmıştır. DL modelleri olarak ise önerilen CNN modeli, DenseNet201, MobileNet ve InceptionV3 popüler mimarileri kullanılmıştır.

Çalışma, mevcut literatürlere göre tarımsal uygulamalarda yaygın olarak kullanılan sinir ağı, çekirdek tabanlı, topluluk ve örnek tabanlı modellere ait modelleri içermeye çalışmıştır. Veri setleri ile ilgili olarak, FTIR spektroskopisi tabanlı bölümde, modellerin tüm veri setlerindeki performansını değerlendirmek için spektrumun farklı bölümlerinin olası kombinasyonlarından elde edilen on beş veri seti kullanılmış ve on beş veri setinin tamamı için PCA dahil olmak üzere, tüm modeller otuz veri setinin tamamı için değerlendirilmiştir. Görüntü tabanlı model eğitimi bölümünde ise sadece çıkarılan özellikler ve PCA kullanılmıştır. Genel olarak, bu çalışmanın yukarıda belirtilen kapsamı, buğday çeşitleri bağlamında yapay zeka tabanlı sınıflandırma modellerinin odaklanmasını ve araştırılmasını sağlayarak hem bilimsel olarak sağlam hem de ilgili tarım sektörlerinde uyarlanabilirlik potansiyeline sahip anlayışlar ortaya koymaktadır.

2. KAYNAK ÖZETLERİ

Utku ve Kijksel (1998), buğday çeşitlerini sınıflandırmak için bir görüntü işleme yazılımı oluşturmuşlardır. Yazılım, tane sınırlarını tespit etmek için medyan ve aralık dereceli filtreleme kullanmaktadır. Uydurma ve doğrudan özellik çıkarma prosedüründen elde edilen özellikler kullanılmış ve 31 ekmeklik buğday çeşidi için %69 sınıflandırma başarısı elde edilmiştir.

Schibeci et al. (2001), MALDI-TOF kütle spektrometresi verilerini kullanarak buğday çeşitlerini sınıflandırmak için ANN kullanmışlardır. Çalışma, protein profillerine dayalı olarak farklı buğday çeşitlerini sınıflandırmak için ileri beslemeli bir yapay sinir ağının başarılı bir şekilde uygulandığını göstermektedir. Elde edilen doğruluk oranı %87.5 ile %100 arasında değişmektedir ve bu durum bu alanda daha fazla araştırma ve zorluk için umut verici sonuçlara işaret etmektedir.

Zapotoczny (2011), görüntü işleme ile dokusal analiz kullanarak dokuz İran buğday çeşidini incelemiştir. Çeşitlerin görüntülerini 24-bit derinliğe ve 400 dpi yüksek çözünürlüğe sahip bir düz yataklı tarayıcı kullanarak elde etmişlerdir. Tek tane izdüşümlerinin dokuları yedi renk kanalı için hesaplanmıştır. Araştırmacılar sonuçları analiz etmek için diskriminant analizi ve sinir ağları kullandılar. Dokuz buğday çeşidinin doku tabanlı sınıflandırması %100 gibi etkileyici bir doğruluk oranına ulaştı. En iyi performans gösteren deneysel tasarım, 20 tanenin birleşik alanından doku ölçümlerini ve sinir ağına girilen yedi renk kanalından değişkenleri içermektedir.

Pazoki ve Pazoki (2011), yapay sinir ağı kullanarak altı buğday çeşidini sınıflandırmak için bir çalışma yürütmüşlerdir. Altı renk, on bir morfolojik özellik ve dört şekil faktörü hakkında veri elde etmişler ve bunları Çok Katmanlı Algılayıcı sinir ağı için girdi olarak kullanmışlardır. Bu modelinin giriş katmanında 21 nöron, çıkış katmanında altı nöron ve farklı nöronlara sahip iki gizli katman yer almıştır. Buğday çeşitlerinin sınıflandırılması için ortalama doğruluk %86.48 olarak hesaplanmış ve özellik seçimi uygulandıktan sonra doğruluk %87.22'ye yükselmiştir.

Dazhou et al. (2012), bireysel tohum örneklerine dayalı olarak çeşitli buğday çeşitlerini ayırt etmek için hiperspektral görüntülemeyi kullanmışlardır. İki farklı tohum pozisyonundan hiperspektral görüntüler yakalamışlar ve ardından her pozisyon için ortalama spektrumları hesaplamışlardır. Görüntü mozaik tekniklerini kullanarak, daha ileri analiz için tek tane bölgesinden ortalama spektrumları çıkardılar. Gizli bir

uzayda buğday çeşitlerinin ayırt edici özelliklerini keşfetmek için PCA kullanılırken, sınıflandırma modelini oluşturmak için Yumuşak Bağımsız Sınıf Analojisi Modellemesi kullanılmıştır. Sonuçlar, hiperspektral görüntülerin tek tanenin ayrıntılı özelliklerini etkili bir şekilde yakaladığını ve veri toplama konumunun hiperspektral görüntü üzerinde bir etkisi olduğunu ortaya koymuştur. Sonuçlara bağlı olarak, hiperspektral görüntüleme, gluten gücüne dayalı olarak üç tür buğday çeşidini başarılı bir şekilde ayırt ederek %96'lık bir doğruluk elde etti. Ayrıca, altı buğday çeşidi için sınıflandırma performansı %93 ve tüm ikili sınıflandırma doğrulukları %90'a ulaşmıştır.

Khoshroo et al. (2014), görüntü işleme ve ANN kullanarak buğday çeşitlerinin sınıflandırılması üzerine bir çalışma yürütmüştür. Dört adet İran buğday çeşidinin sınıflandırılması, kontrollü bir aydınlatma ortamında elde edilen, çeşit başına 164 tane görüntüden on morfolojik özellik çıkarılarak gerçekleştirilmiştir. Çeşitli ANN topolojileri geliştirilmiş ve dokuz temel morfolojik özellik ANN için girdi olarak kullanılmıştır. En iyi ANN yapısı 9-26-4 olarak bulunmuştur. Buğday çeşidi sınıflandırması için genel doğruluk %85.72 olarak rapor edilmiştir. Bu araştırma, buğdayın piyasa değerinin belirlenmesinde ve bitki ıslahçılarına verim ve kalite tahmininde yardımcı olması açısından önem taşımaktadır.

Yasar et al. (2014), çalışmalarında, X-ray tekniği ile ölçülen geometrik özelliklere dayalı olarak farklı buğday tohum türlerini sınıflandırmak için bir ANN kullanmışlardır. Sonuçlar, önerilen yöntemin buğday çeşitlerini tanımda etkili olduğunu göstermiştir. ANN kullanılarak yapılan sınıflandırma işlemi %99.78'lik test sonuçlarına ulaşmıştır.

Sabancı ve Akkaya (2016), UCI makine öğrenmesi veritabanından bir tohum veri setine dayalı olarak farklı buğday çeşitlerini sınıflandırmak için veri madenciliği algoritmalarını kullanmıştır. Çalışmada, 7 nöronlu çok katmanlı bir algılayıcı kullanılarak %97.17 sınıflandırma başarısı ve 4 komşu kullanılarak KNN ile %95.71 başarı oranı elde edilmiştir. Çalışma, buğday çeşitlerinin doğru bir şekilde kategorize edilmesinde veri madenciliğinin etkinliğini vurgulamaktadır.

Sabancı et al. (2017), buğday tanelerini ekmeklik veya makarnalık olarak sınıflandırmak için ANN kullanan bilgisayarla görmeye dayalı bir yöntem geliştirmiştir. Ana görsel özellikler görüntü işleme teknikleri kullanılarak çıkarılmış ve ANN dört farklı girdi veri alt kümesi ile eğitilmiştir. Basitleştirilmiş ANN modeli,

buğday tanelerini başarılı bir şekilde sınıflandırmıştır.

Charytanowicz et al. (2018), X-ray görüntülerini kullanarak buğday tanelerini sınıflandırmak için geometrik özelliklerin kullanımını değerlendirmişlerdir. Tanelerin iç yapısını incelemek ve aralarındaki farkları ölçmek için gelişmiş dijital görüntü işleme araçlarını kullandılar. PCA ve çok değişkenli faktör analizi yoluyla, ilk üç faktör tarafından açıklanan %89.97'lik yüksek bir varyasyon elde ederek, şekil tanımlayıcılarına dayalı olarak üç buğday çeşidinin güvenilir bir şekilde ayırt edilmesini sağladılar. Araştırma, görüntü analizi ve çok değişkenli istatistiksel yöntemleri içeren çeşitli pratik görevler için yöntemlerinin pratik kullanılabilirliğini ve etkinliğini göstermiştir.

Taner et al. (2018), yirmi sekiz ekmeklik buğday çeşidi ve on bir makarnalık buğday çeşidi, arpa, yulaf ve tritikale dahil olmak üzere çeşitli tahıl türlerini sınıflandırmak için ANN modelini eğitmiştir. Bu çeşitler arasında ayırım yapmak için bin tane ağırlığı, geometrik ortalama çap, küresellik, tane hacmi, yüzey alanı, hacim ağırlığı, özgül ağırlık, porozite ve renk parametreleri gibi tanelerin on bir fiziksel özelliği kullanılmıştır. Sınıflandırmayı gerçekleştirmek için 11-2-2 ANN modeli tasarlanmıştır. Bulgular, ANN modelinin 0.000184 RMSE ve %0.001 ortalama hata gibi oldukça düşük değerlerle yüksek doğruluk gösterdiğini ve modelin tahminleri ile gerçek veriler arasında güçlü bir uyum olduğunu ortaya koymuştur.

Sabancı et al. (2020), araştırmalarında, ürün kalitesini sağlamak için dalgacık tabanlı görüntü füzyonu kullanarak ekmeklik ve makarnalık buğday arasında ayırım yapmaya odaklanmıştır. Ultraviyole (ultraviolet, UV) ve beyaz ışık (white light, WL) görüntülerini bu teknikle birleştirerek, sırasıyla SVM ve çok katmanlı algılayıcı (Multilayer Perceptron, MLP) algoritmaları kullanarak %94.82 doğruluk oranına sahip UV ve WL görüntülerine kıyasla %98.27 daha yüksek bir sınıflandırma doğruluğu elde etmişlerdir. Çalışma, dalgacık tabanlı füzyon yaklaşımı ile her üç öğrenme algoritmasında da sınıflandırma doğruluğundaki önemli iyileşmeyi vurgulamıştır.

Zhou et al. (2020), 30 çeşitten 140.000'den fazla buğday tanesini içeren büyük bir yakın kızılötesi (Near Infrared, NIR) spektral veri kümesi kullanarak buğday tanesi çeşidi tanımlama üzerine bir çalışma yürütmüştür. Hedefle ilgili spektral kanalları tanımlamak için yeni bir CNN tabanlı özellik seçici önermişlerdir. Tek boyutlu veri sınıflandırması için dikkat içeren bir evrişimli sinir ağı çerçevesi tasarlanmıştır. Bu ağ, tam spektrumları kullanarak %93.01 ve CNN tabanlı özellik seçici kullanılmasıyla

elde edilen 60 kanallı özellikler üzerinde eğitildiğinde %90.20 doğruluk elde etmiştir.

Lingwal et al. (2021), araştırmalarında, on beş farklı buğday çeşidini görsel özelliklerinin görüntülerine göre sınıflandırmak için CNN kullanarak yeni bir model geliştirmiştir. Model, 15.000 buğday tanesi görüntüsü içeren bir veri kümesi üzerinde eğitilmiş ve sınıflandırma için en doğru ve verimli modeli oluşturmak için çeşitli hiperparametreler optimize edilmiştir. Farklı modellerin performansı, sınıflandırma doğruluğu ve kategorik çapraz entropi kaybı kullanılarak karşılaştırılmıştır. En iyi performans gösteren model %94.88 eğitim doğruluğu ve %97.53 test doğruluğu elde ederek CNN kullanarak görsel özelliklere dayalı başarılı ve hassas buğday çeşidi sınıflandırması ortaya koymuştur.

Jena et al. (2021), geometrik özellikler ve ML algoritmaları kullanarak üç buğday çeşidini tanımlamayı amaçlamıştır. Seçilen girdi özellikleri, Kama, Rosa ve Kanada çeşitlerinin X-ışını görüntülerinden çıkarılan alan, çevre, uzunluk, genişlik, kompaktlık ve asimetri katsayısını içeriyordu. SVM, KNN ve Naïve Bayes olmak üzere üç sınıflandırıcı değerlendirilmiş ve en yüksek performans %94.3 ile KNN kullanılarak elde edilmiştir.

Görünür Yakın Kızılötesi görüntülerle (Visible Near Infrared, VNIR) elli buğday türünden oluşan yeni bir görüntü veri seti oluşturan Özkan et al., (2021) buğday taneleri sınıflandırması için AlexNet ve VGG16 modellerinin performansını değerlendirmiştir. Buğday etiketlerini tahmin etmek için SVM ve SoftMax sınıflandırıcıları kullanılmıştır. Veri kümesi üzerinde yapılan deneyler, AlexNet ve VGG16 için sırasıyla FC6 ve FC7 özellikleri kullanılarak doğrusal SVM sınıflandırıcısı için %96.00 ve %99.00 genel yaklaşık doğruluk oranlarıyla sonuçlanmıştır. Softmax tahminlerine göre, VGG16 modeli örneklerin %92'sini doğru bir şekilde ayırt ederken, AlexNet %70'e ulaşmıştır.

Laabassi et al. (2021), DL tabanlı bir yaklaşım kullanarak dört buğday çeşidinin tanımlanmasını karmaşık ve zorlu olarak bildirmişlerdir. Beş standart CNN mimarisi, sınıflandırma performansını artırmak için Öğrenme Aktarımı (Transfer Learning) kullanılarak eğitilmiştir. Değerlendirme, Cezayir'deki farklı bölgelerden elde edilen tek taneli görüntülerden oluşan bir veri kümesi üzerinde gerçekleştirilmiştir. Derin öğrenme tabanlı yaklaşım, DensNet201 CNN modelinin %95.68 ile en yüksek test doğruluğunu elde ettiği doğruluk ve güvenilirlik göstermiştir.

CNN ve derin özelliklerle ilgili bir çalışmada (Çelik et al., 2022), 41 durum buğdayı tanesini tanımlamak için hibrit bir CNN modeli sunmuştur. MobileNetV2 ve SqueezeNet CNN modellerinden Logits ve Pool10 özellik katmanlarını birleştirerek, SVM'ni eğitmek için yeni derin özellikler kullanılmıştır. Hibrit model deneyleri, test doğruluğunda %91.89'luk bir performans elde etmiştir.

Khojastehnazhand ve Roostaei (2022), bir yapay görme sistemi kullanarak yedi İran buğday çeşidinin sınıflandırılmasını araştırmıştır. Gri Seviye Birlikte Oluşum Matrisi, Gri Seviye Çalışma Uzunluğu Matrisi ve Yerel İkili Örüntü çıkarılmış ve PCA ile ANN ve SVM algoritmalarının boyutsallık azaltma tekniği uygulanmıştır. Seçilen 125 ve 20 özellikten oluşan tüm özelliklerin bir kombinasyonu kullanılarak sınıflandırma doğruluğu sırasıyla %95.04 ve %98.10 olmuştur. Çalışma, görüntü işleme ve doku özelliklerinin çıkarılmasının farklı buğday çeşitlerinin daha yüksek doğrulukla sınıflandırılmasını sağladığını göstermiştir.

Işık et al. (2022), kırktan fazla buğday çeşidini sınıflandırmak için Görünür Işık (Visible Light, VL), Görünür Yakın Kızılötesi (Visible Near-Infrared, VNIR) ve Kısa Dalga Kızılötesi (Short-Wave Infrared, SWIR) görüntüleme kullanarak buğday çeşitlerini sınıflandırmak için iki yöntem geliştirmiştir. Bu tekniklerde hem geleneksel hem de modern özellik çıkarma algoritmaları, yani kelime torbası (bag of word, BoW) çerçevesi ve CNN çerçevesi kullanılmıştır. Buğday çeşitlerini tanımlamak için CNN ve BoW çerçevelerinden gelen tahminler birleştirilerek bir uzlaşma kuralı oluşturulmuştur. Sonuçlar, CNN çerçevesi için %99.94 ve BoW çerçevesi için %68.94 ile yüksek doğruluk göstermiştir. Çalışmanın deneyleri, CNN özelliklerinin, özellikle görüntülerde tekrarlanan buğday taneleri durumunda, doku desenlerini doğru bir şekilde temsil etmede ve eşleştirmede el yapımı özelliklerden daha iyi performans gösterdiğini ve CNN'i buğday çeşidi sınıflandırması için daha uygun hale getirdiğini ortaya koymuştur.

Madhavan et al. (2023), buğday tohumu sınıflandırması üzerine yaptıkları çalışmada, geometrik parametreleri ölçmek için çeşitli teknikler kullanmış ve alan, çevre, kompaktlık, tane uzunluğu, tane genişliği, asimetri katsayısı ve tane oluğu uzunluğu gibi parametreleri uygulamışlardır. Sınıflandırma bir sinir ağı örüntü tanıyıcısı kullanılarak gerçekleştirilmiş ve buğday tohumu veri tabanı UCI makine öğrenmesi havuzundan elde edilmiştir. Karışıklık matrisi ve ROC kullanılarak yapılan performans değerlendirmesi %96.7'lik bir doğruluk oranı göstermiş ve sinir ağı örüntü

tanıyıcısının buğday tohumu sınıflandırması için uygun bir araç olarak etkinliğini vurgulamıştır.

Jin et al. (2023), hiperspektral görüntülerden on buğday çeşidinin tanımlanması için maliyete duyarlı bir KNN yöntemi sunmuş ve son yöntemlerden daha iyi performans gösteren %98.09 ve %97.45 sınıflandırma doğruluğu elde etmişlerdir.

Que et al. (2023), sekiz çeşit için çeşit tanımlaması oluşturmak amacıyla gruplandırılmış bir CNN ile birleştirilmiş seçilmiş spektral özellik aralıklarını kullanmışlardır. Aralık seçim yöntemleri arasında aralık sürekliliği kaldırmanın CNN ile birleştirildiğinde sınıflandırmada %4.4'lük bir artışa yol açtığını bildirdiler.

Yaşar et al. (2024), beş buğday çeşidini sınıflandırmak için öğrenme aktarımı stratejisini benimsemişlerdir. Xception algoritmasını kullanarak derin özellikleri çıkardıktan sonra BiLSTM sinir ağlarını eğittiler ve %97.73 doğruluk elde ettiler.

Yaşar (2024), beş ekmeklik buğday çeşidini sınıflandırmak için geleneksel öğrenme modellerinden önceden eğitilmiş ResNet CNN ve SVM'nin entegrasyonunu araştırmıştır. Çalışmada en etkili bin özellik kullanılarak %94.51'lik bir sınıflandırma doğruluğu rapor edildi.

Ceyhan et al. (2024), yirmi sekiz buğday çeşidini sınıflandırmak için derin bir öğrenme modeli eğitmiş ve sınıflandırmanın performansını %99.52 olarak bulmuşlardır.

Han et al. (2024), çift kanallı derin öğrenme özelliği füzyonu ile birleştirilen hiperspektral görüntülemenin, üç çeşit, dört çeşit ve beş çeşit buğday kombinasyonları için %99.18, %97.30 ve %93.18 doğruluk oranı elde etmişlerdir.

Bu alanda yapılan diğer çalışmalarda;

Buğdayla ilgili ve ML uygulamalarında, hastalık tespiti ortak alanlardan biridir. Haider et al. (2020), hastalıklar ve pestisit kullanımı da dahil olmak üzere buğday çiftçilerinin karşılaştığı ortak sorunları ele almıştır. Manuel hastalık tespit yöntemleri hatalar, zaman tüketimi ve geniş mahsul alanlarıyla uğraşırken potansiyel yanlışlıklar ile sınırlıdır. Bu zorlukların üstesinden gelmek için araştırmacılar, çiftçilere ürün hastalıklarını etkili bir şekilde yönetmede yardımcı olmak için kendi kendine öğrenme yeteneklerine sahip kullanıcı dostu bir uygulama geliştirdiler. Verileri eğitmek için CNN kullandılar ve hastalık tespitinde %98'den daha yüksek bir doğruluk elde ettiler.

Dong et al. (2020), buğday yaprak hastalıklarını görüntülerden doğru bir şekilde tanımlamak için yeni bir yaklaşım olan Diferansiyel Yükseltme Evrişimli Sinir Ağı'nı (Differential Amplification Convolutional Neural Network, DACNN) tanıttı. DACNN, gerçek ve beklenen çıktılar arasındaki küçük farklılıklara duyarlılığı artırmak için derin evrişimsel katmanlar arasında ek dallar içerir. Çalışma, veri kümesini genişletmek için beş veri artırma yöntemi içeriyordu. Daha sonra, DACNN, SVM, KNN ve RF sınıflandırıcıları ile birlikte kullanılmıştır. Buğday yaprak hastalığı veri kümesini değerlendirmek için. LeNet-5, AlexNet, ZFNet ve Inception V3 gibi modellere karşı yapılan karşılaştırmalı analiz, buğday yaprak hastalığı veri kümesinde ortalama %95.18 tanıma doğruluğu elde eden DACNN'nin üstünlüğünü göstermiştir.

Su et al. (2021), buğdayda Fusarium başak yanıklığını (Fusarium head blight, FHB) tespit etmek için hiperspektral görüntüleme uygulamasını değerlendirmiştir. Fusarium ile enfekte olmuş 1.680 buğday başağı örneğinden 400-1000 dalga boyu aralığında hiperspektral görüntüler elde edilen görüntüler, boyutsallığı azaltmak için PCA kullanılarak işlenmiştir. Görüntülerden çeşitli teknikler kullanılarak doku ve renk özellikleri çıkarılmıştır. En alakalı özellikleri seçmek için gradyan artırma karar ağacı (Gradient Boosting Decision Tree) ve sıralı geriye doğru eleme (Sequential Backward Elimination) kullanılmış ve sonuçta 50 spektral özellik ve 40 görüntü özelliği elde edilmiştir. Random Forest Sınıflandırıcısı bu özelliklere dayalı olarak eğitilmiş ve optimum özellikler veri kümesini oluşturmak için derin bir CNN kullanılmıştır. Derin CNN, R^2 0.97 ve RMSE 3.78 ile doğru hastalık şiddeti tahminine ulaşmıştır.

Bao et al. (2021), tarla görüntülerinde buğday başak hastalıklarının otomatik olarak tanımlanması için SimpleNet adlı bir CNN modeli geliştirmiştir. Modelin hastalık özellikleri için temsil yeteneğini geliştirmek için Evrişimsel Blok Dikkat Modülünü (Convolutional Block Attention Module, CBAM) tanıttılar. Önerilen SimpleNet, test veri kümesinde %94.1 tanımlama doğruluğu elde ederek VGG16, ResNet50 ve AlexNet gibi diğer CNN modellerinden daha iyi performans gösterdi. Sadece 2.13 milyon parametre ile SimpleNet, mobil tabanlı hastalık tanımlama için uygun bulunmuştur.

Zhang et al. (2021), buğday uyuz hastalığının kontrolünde fungusitlerin etkinliğini otomatik ve etkili bir şekilde değerlendirmek için yeni bir yöntem sunmuştur. Buğday başaklarını ve hastalıklı alanları segmentlere ayırmak için sırasıyla Unet++ ağını ve Bulanık C-Means algoritmasını R-G yöntemiyle birlikte

kullanmışlardır. CNN ve bağlantılı alan yöntemini (connected domain method) kullanarak, tüm buğday başaklarını ve hastalıklı buğday başaklarını doğru bir şekilde saymışlardır. Yöntem, hastalık insidansını başarılı bir şekilde derecelendirmiş ve hastalıklı buğday başaklarının tahmin edilen ve ölçülen seviyeleri arasında %5'ten daha az hata ile beş fungusitin etkinliğini değerlendirmiştir. Bu yaklaşım, tarlada fungusit etkinliğinin manuel olarak değerlendirilmesine zaman kazandıran ve emek açısından verimli bir alternatif sunmuştur.

Chen et al. (2021), dijital görüntülerden buğday akarı tanımlama ve sayma işlemlerini otomatikleştirmek için üç aşamalı bir derin öğrenme yöntemi sunmuştur. Araştırma, görüntü segmentasyonu, CNN'ler kullanılarak özellik çıkarımı ve bölge öneri ağı (Region Proposal Network, RPN) aracılığıyla potansiyel alanların tanımlanmasını içeriyordu. ZFnet ve VGG16 gibi farklı derin öğrenme ağlarını kullanan yöntem, sırasıyla %94.6 ve %96.4 doğruluk oranlarına ulaşarak doğru akar tespiti ve sayımındaki etkinliğini göstermiştir.

Fernandez-Campos et al. (2021), Derin CNN kullanarak buğday başak yanıklığını ele almışlardır. CNN modelleri, buğday başak yanıklığı şiddetini sınıflandırmak için eğitilmiştir. Çalışmada, kontrollü koşullar altında toplanan ve sınıflandırılan verilerin güvenilirliğini değerlendirmek için değerlendiriciler arası uyum analizi kullanılmıştır. Model eğitiminden önce, değerlendiriciler arası uyum analizinde yüksek doğruluk ve düşük yanlışlık gözlemlenmiştir. Sonuçlar, eğitilen CNN modellerinin üç buğday başak yanıklığı şiddeti kategorisine sınıflandırmak için umut verici bir yaklaşım sunduğunu göstermiştir. Hem olgunlaşmamış hem de olgunlaşmış başakların görüntüleri üzerinde eğitilen modeller, sınıflandırma sırasında en yüksek hassasiyet, geri çağırma ve F1 puanı göstermiştir. Bu yüksek sınıflandırma doğruluğu, bu modellerin gelecekte buğday başak yanıklığı fenotiplemesini kolaylaştırma potansiyeline işaret etmektedir.

Goyal et al. (2021), hem ürün miktarını hem de kalitesini artırmak için umut verici bir çözüm sunan, ürün kalitesi değerlendirmesine ve fiyatlandırmasına yardımcı olan buğday hastalığı sınıflandırma modeli geliştirmişlerdir. Hastalık teşhisi ve sınıflandırmasında etkili olduğu kanıtlanmış olan derin öğrenme tabanlı görüntü analizinden yararlanılmıştır. Bu çalışmada, buğday hastalıklarını on sınıfta doğru bir şekilde sınıflandırmak için eğitilmiş yeni bir derin öğrenme modeli incelenmiştir. Önerilen yöntem %97.88 gibi yüksek bir test doğruluğuna ulaşmıştır. Özellikle, yaygın

olarak kullanılan iki derin öğrenme modeli olan VGG16 ve RESNET50'den daha iyi performans göstererek sırasıyla %7.01 ve %15.92 oranında doğruluk elde etmiştir. Deneysel sonuçlar ayrıca hassasiyet, geri çağırma ve f-skor metriklerinde üstün performans göstererek önerilen yaklaşımın etkinliğini vurgulamaktadır.

Jiang et al. (2021), pirinç ve buğday yapraklarını etkileyen hastalıkların tanımlanmasına odaklanan bir çalışma yürütmüştür. Araştırmalarında üç tür pirinç yaprağı hastalığı ve iki tür buğday yaprağı hastalığı hedeflenmiş, her bir hastalık için 40 görüntü toplanmış ve daha sonra bu görüntüler iyileştirilmiştir. Birincil amaç, çoklu görev öğrenme yaklaşımı kullanarak VGG16 modelini geliştirmektir. Bu, öğrenme aktarımı ve alternatif öğrenme stratejileri için ImageNet üzerinde ön eğitim yoluyla modelin rafine edilmesini içeriyordu. Ortaya çıkan model, pirinç yaprağı hastalıkları için %97.22 ve buğday yaprağı hastalıkları için %98.75 gibi etkileyici bir doğruluk oranına ulaşmıştır. Karşılaştırmalı deneyler, bu yöntemin tek görevli modellerden, öğrenme aktarımında yeniden kullanım modeli yaklaşımından ve ResNet50 ve DenseNet121 modellerinden daha iyi performans gösterdiğini ortaya koymuştur. Deneysel sonuçlar, geliştirilmiş VGG16 modelinin ve çalışmada önerilen çok görevli öğrenme aktarımı tekniğinin altını çizmiştir.

Kukreja ve Kumar (2021), buğday pas hastalıklarının insan müdahalesi olmadan otomatik olarak sınıflandırılmasını sağlayan bir derin öğrenme yöntemi olan DCNN'yi araştırmışlardır. Bu DCNN'nin eğitim ve test aşamaları, buğday pas hastalıklarının kesin olarak belirlenmesini sağlamış ve güçlü sınıflandırma performansı göstermiştir. CGIAR veri kümesinden 1486 buğday bitkisi görüntüsü ve ikincil kaynaklardan elde edilen 514 buğday sap pası görüntüsünden oluşan bir veri kümesi kullanılarak, DCNN'nin eğitimi ve testi için toplam 2000 buğday bitkisi görüntüsü kullanılmıştır. Sonuç olarak, önerilen yaklaşım buğday pas hastalıklarının belirlenmesinde %97.16'lık bir sınıflandırma doğruluğu elde etmiştir.

Kumar ve Kukreja (2021), çalışmalarında, buğday bitkilerindeki mantar semptomlarını doğru bir şekilde bulmak için tasarlanmış, Mask-RCNN olarak bilinen maske-bölge tabanlı bir CNN tanıttılar. Mask-RCNN bunu nesne algılama ve semantik segmentasyon süreçlerinin bir kombinasyonunu kullanarak başarmıştır. Başlangıçta, birincil ve ikincil kanallardan 400 buğday görüntüsü içeren bir veri kümesi oluşturulmuştur. Mask-RCNN modeli, özellik çıkarma ve model oluşturma için VGG-16 ve VGG-19 ön eğitilmiş modelleri kullanılarak oluşturulmuştur. Tam boyutlu

görüntülerden buğday çizgi pasını temsil eden bir maske oluşturulmasının ardından, Mask-RCNN buğday bitkisi üzerindeki çizgi pası semptomlarının yerini tahmin etmiştir. Mask-RCNN'nin eğitimi sırasında eğitim oranı ve yığın oranı dahil olmak üzere çeşitli hiperparametreler ayarlanmıştır. Mask-RCNN, F1-skoru ile ölçülen, çizgi pasından etkilenen hastalıklı alanları tanımlamak için %90.63 ve hastalık belirtilerine sahip buğday bitkilerini tanımlamak için %87.25'lik bir tahmin oranı elde ettiler.

Maqsood et al. (2021), buğday sarı pas hastalığının tespiti için derin öğrenme modellerini eğitmek üzere kullanmadan önce görüntü kalitesini artırmak için süper çözünürlüklü üretken karşıt ağlardan (super-resolution generative adversarial networks, SRGAN) yararlanan bir yöntem üzerinde çalışmıştır. Veri ön işlemeden sonra, SRGAN'lar görüntüleri yukarı örneklemek, çözünürlüklerini artırmak ve eğitim sürecinde CNN tarafından yüksek kaliteli özelliklerin öğrenilmesini kolaylaştırmak için kullanılmıştır. Çalışma, düşük çözünürlüklü görüntüler üzerinde eğitilen modellere kıyasla önemli ölçüde iyileştirilmiş sonuçlar ortaya koymuştur. Bu gelişme, düşük çözünürlüklü görüntülerle elde edilen %75'lik doğruluk oranını aşan, yukarı örneklenmiş görüntüler kullanılarak elde edilen %83'lük genel test doğruluğu ile sonuçlanmıştır. Önerilen yaklaşım, yüksek çözünürlüklü kameraların bulunmaması nedeniyle görüntü çözünürlüğünün sınırlı olduğu çeşitli gerçek dünya senaryolarında uygulanabilir.

Kendler et al. (2022), görüntüler üzerinde eğitilen CNN kullanarak ürün hastalıklarını tanımlamak için bir yöntem sunmuşlardır. Çalışma öncelikle buğday sarı pas ve diğer ürün hastalıklarının erken tespitine odaklanmıştır. Bunu başarmak için, görüntüler sahne, arka plan değiştirilmeden veya aydınlatma koşulları kontrol edilmeden doğal haliyle çekilmiştir. Daha sonra her görüntü, veri değişkenliği korunurken orijinal uzamsal çözünürlük korunarak birden fazla yamaya bölünmüştür. Bu yaklaşım, farklı yamalar arasında hastalık belirtisi, görüntüleme açıları ve aydınlatma koşullarındaki varyasyonları kapsayan oldukça çeşitli bir veri kümesi ile sonuçlandı. En etkili modeli belirlemek için bu çeşitli veri kümesi kullanılarak çeşitli CNN mimarileri eğitilmiştir. Elde edilen sınıflandırma doğruluğu 95.4 ± 0.4 olup, bitki hastalıklarının otonom erken tespiti için umut verici bir potansiyele işaret etmektedir.

3. MATERYAL VE YÖNTEM

3.1. Buğday Çeşitleri

Tarım ve Orman Bakanlığına bağlı araştırma enstitülerinden ekmeklik buğday çeşitleri toplandı. Hepsi kendi ıslahçı kuruluşundan toplanan 124 ekmeklik buğday çeşidi çalışmada kullanılmıştır (Tablo 3.1).

Tablo 3.1. Çalışmada kullanılan ekmeklik buğday çeşitleri

No	Ekmeklik Buğday	Araştırma Enstitüsü	İl
1	Ahmetağa	Bahri Dağdaş Uluslararası Tarımsal Araştırma Enstitüsü	Konya
2	Bayındır		
3	Bozkır		
4	Buhara		
5	Dağdaş		
6	Ekiz		
7	İkonya		
8	Kilistra		
9	Konya-2002		
10	Meke		
11	Selçuklu		
12	Şehzade		
13	Taner		
14	Tuğra		
15	Yavuz		
16	Adana-99	Doğu Akdeniz Tarımsal Araştırma Enstitüsü Müdürlüğü	Adana
17	Altınbaşak		
18	Altınöz		
19	Candaş		
20	Ceyhan-99		
21	Ekinoks		
22	Gökkan		
23	Karatopak		
24	Karmen		
25	Osmaniyem		
26	Pandas		
27	Seri-2013		
28	Simge		
29	Şahika		
30	Yakamoz		

31	Alparslan		
32	Ayyıldız		
33	Daphan		
34	Doğu-88		
35	Karasu-90	Doğu Anadolu Tarımsal Araştırma Enstitüsü Müdürlüğü	Erzurum
36	Kırık		
37	Lancer		
38	Nenehatun		
39	Palandöken-97		
40	Yıldırım		
41	Cumhuriyet-75		
42	Efe		
43	Egem		
44	Gönen-98		
45	Kayra	Ege Tarımsal Araştırma Enstitüsü Müdürlüğü	İzmir
46	Meltem		
47	Ünsal-35		
48	Zeybek-35		
49	Ziyabey-98		
50	Cemre		
51	Dinç		
52	Hilar		
53	Kale	GAP Uluslararası Tar. Arş. ve Eğitim Merkezi Müdürlüğü	Diyarbakır
54	Karacadağ-98		
55	Nurkent		
56	Tekin		
57	Aksungur		
58	Alpu-2001		
59	Altay-2000		
60	Beyhan		
61	Bezostaja-1		
62	Çetinel-2000		
63	Es-26		
64	Gerek-79		
65	Harmankaya-99		
66	İzgi-2001	Geçit Kuşağı Tarımsal Araştırma Enstitüsü	Eskişehir
67	Mesut		
68	Müfitbey		
69	Nacibey		
70	Reis		
71	Soyer-02		
72	Sönmez-2001		
73	Sultan-95		
74	Yıldız-98		
75	Yunus		

76	Altındane		
77	Kirve		
78	Nevzatbey	Karadeniz Tarımsal Araştırma Enstitüsü	Samsun
79	Özcan		
80	Sakin		
81	Acar		
82	Adalı		
83	Alada		
84	Altınsa		
85	Altuğ		
86	Arifbey	Mısır Araştırma Enstitüsü	Sakarya
87	Beşköprü		
88	Halis		
89	Hanlı		
90	Kaynarca		
91	Mirsa		
92	Nusrat		
93	Flamura-85	Tareks Tarım Ürünleri Araç Gereç İth. İh. ve Tic. A.Ş.	
94	Ayaz		
95	Ayten Abla		
96	Bayraktar-2000		
97	Bayram		
98	Çavuş		
99	Demir-2000		
100	Demirhan		
101	Eser		
102	Fazılbey		
103	Gün-91	Tarla Bitkileri Merkez Araştırma Enstitüsü	Ankara
104	Hamitbey		
105	İkizce-96		
106	Kenanbey		
107	Kışla		
108	Kürşad		
109	Lütfübey		
110	Selamibey		
111	Şanlı		
112	Tosunbey		
113	Zencirci-2002		
114	Abide		
115	Albaşak		
116	Anafarta		
117	Damla		
118	Değirmen		
119	Eylül	Trakya Tarımsal Araştırma Enstitüsü	Edirne
120	Harman-22		
121	Köprü		
122	Saban		
123	Toprak		
124	Yüksel		

3.2. Görüntü Alma

Görüntü alma sürecinde Foldio-360 cihazı kullanılmıştır (Şekil 3.1). Bu cihaz, nesnenin (bizim durumumuzda buğday) merkezinde yer aldığı ve görüntü alınırken 360 derece dönmeye izin veren kubbe şeklinde bir cihazdır. Ayrıca mobil uygulama kullanılarak kontrol edilebilen ayarlanabilir bir ışık sistemine sahiptir. Bunun faydası, çok boyutlu görüntü alımını kolaylaştırmasıdır. Cihaz, mobil ve kamera gibi harici cihazlara bağlanmak için Bluetooth ve IR sensörü kullanır. Cihaz, daha önce uygulanan ve bir nesnenin konumunun manuel olarak değiştirildiği fotoğraf yakalama mekanizmasına bir değer kattığı için avantajlıdır. Döner tabla uygun bir hıza ayarlanabilmektedir. 360 derecede 48 adet görüntü alınabilmektedir. Farklı yönlerden görüntü yakalayabilme özelliği, nesnenin özelliklerinin temsiliyi artırmakta ve sınıflandırma modelinin başarısına olumlu katkı sağlamaktadır.



Şekil 3.1 Foldio-360 görüntü alma cihazı

3.2.1. Kamera

Makro lensli 24MP çözünürlüğe sahip DS2000 model bir Canon kamera kullanılmıştır (Şekil 3.2). Kamera, görüntü alma süreci boyunca eşit bir mesafede konumlandırılmıştır.



Şekil 3.2. Canon EOS 2000D kamera

3.2.2. Görüntü alma yazılımı

Görüntüleri uzaktan yakalamak için EOS 2000D ile entegre olan EOS utility 3 for windows yazılımı kullanılmıştır.. Bu yazılım, görüntülerin uzaktan çekilmesine, tercih edilen aydınlatmanın seçilmesine ve görüntülerin eş zamanlı olarak kaydedilmesine ve sürecin canlı olarak izlenmesine olanak tanımaktadır.

3.2.3. Görüntü elde etme süreci

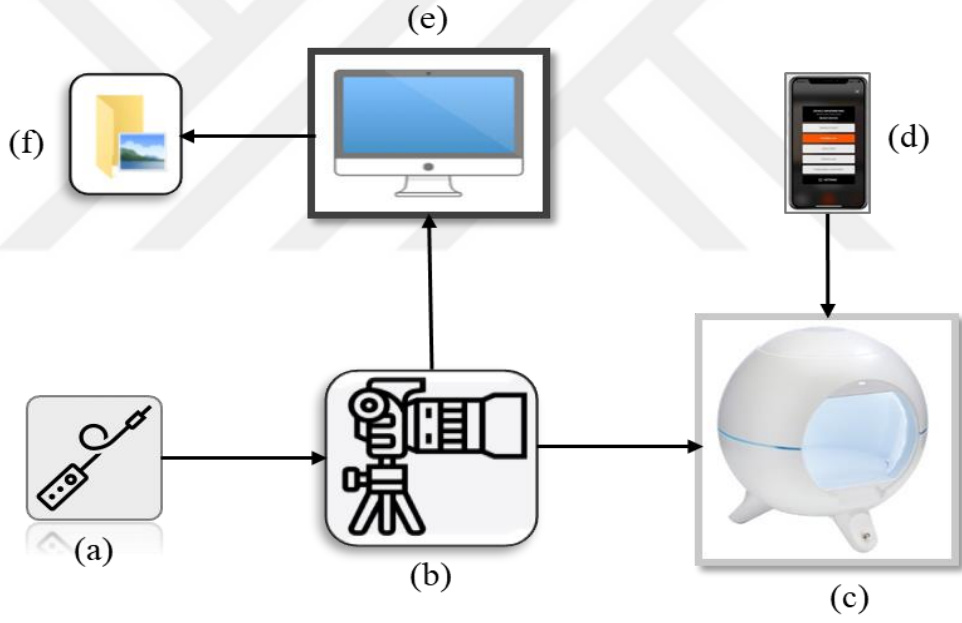
İşlem, Foldio-360 çekim cihazı, kamera, uzaktan kumanda, cep telefonu ve bilgisayar olmak üzere tüm cihazların bağlanmasıyla gerçekleştirilmiştir. Foldio-360'ın döner tablası, bir dakika boyunca 360^0 dönmesi için mobil uygulama tarafından kontrol edildi. Döner tabla üzerine yerleştirilen buğday tanesi döndükçe, kamera bir dakika boyunca her dört saniyede bir beş görüntü çekecek şekilde ayarlandı. Başka bir ifadeyle buğday tanelerinin görüntüleri her 24^0 'de bir çekildi ve her dakikada toplam 75 görüntü elde edildi (Şekil 3.3, 3.4, 3.5). Görüntülerin elde edilmesi ile ilgili örnek bir süreç Şekil 3.5'de verilmiştir. Çalışmada 124 çeşit, her çeşitten 50 örnek ve 4 saniyede bir 5 görüntü (her çeşit için 3.750 görüntü) olmak üzere toplam 465.000 görüntü alınmıştır (Şekil 3.6).



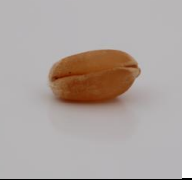
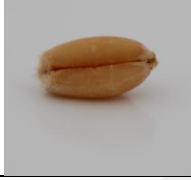
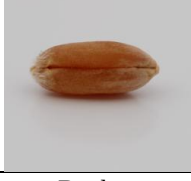
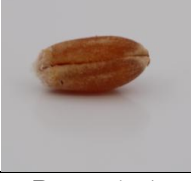
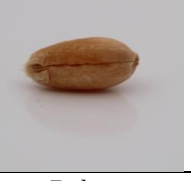
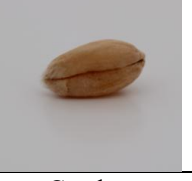
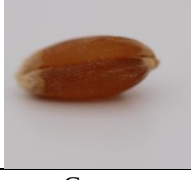
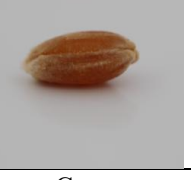
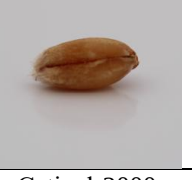
Şekil 3.3. (a) Foldio-360 platformu ile kamera, (b) Buğday tanesinin yerleştirildiği döner tabla, (c) Görüntü alma sisteminin genel görünümü

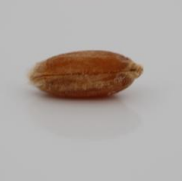
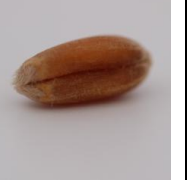
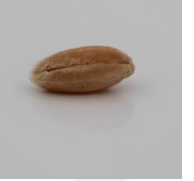
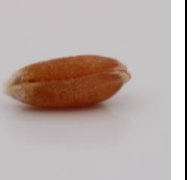
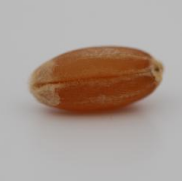
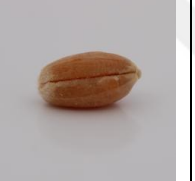
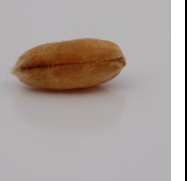
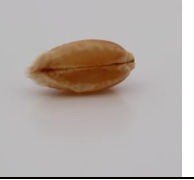
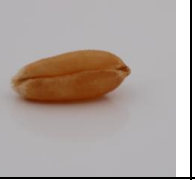
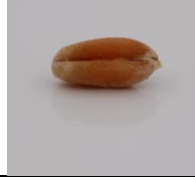
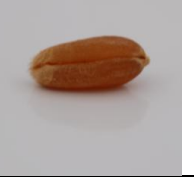
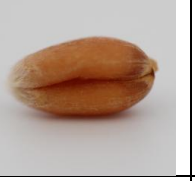


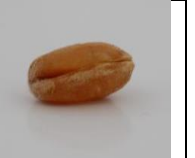
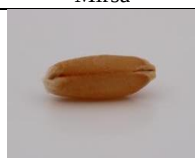
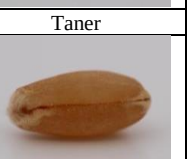
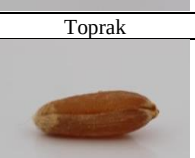
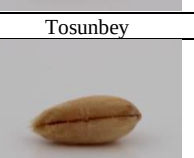
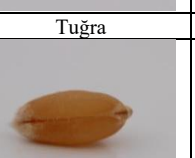
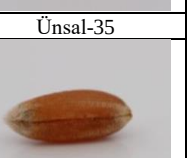
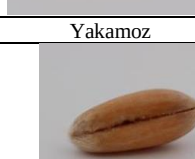
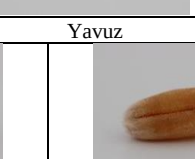
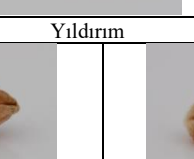
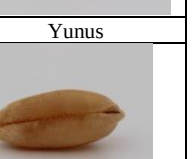
Şekil 3.4. Farklı pozisyonlarda çekilen buğday tane örnekleri



Şekil 3.5. Görüntü alma sürecinin şematik diyagramı. (a) Uzaktan kumanda, (b) Kamera, (c) Foldio-360 cihazı, (d) Foldio-360 mobil uygulaması, (e) EOS yardımcı yazılımı, (f) Görüntülerin kaydedildiği klasör

				
Abide	Acar	Adalı	Adana-99	Ahmetağa
				
Aksungur	Alada	Albaşak	Alparslan	Alpu-2001
				
Altay-2000	Altınbasak	Altındane	Altınöz	Altınsa
				
Altuğ	Anafarta	Arıfbey	Ayaz	Ayten Abla
				
Ayyıldız	Bayındır	Bayraktar-2000	Bayram	Beşköprü
				
Beyhan	Bezostaja-1	Bozkır	Buhara	Candaş
				
Cemre	Ceyhan-99	Cumhuriyet-75	Çavuş	Çetinel-2000
				
Dağdaş	Damla	Daphan	Değirmen	Demir-2000

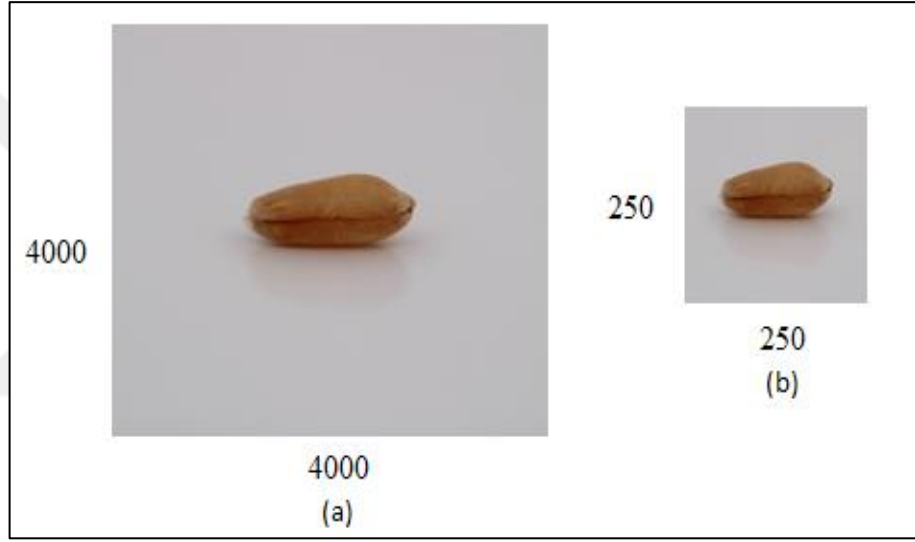
				
Demirhan	Dinc	Doğu-88	Efe	Egem
				
Ekinoks	Ekiz	Es-26	Eser	Eylül
				
Fazılbey	Flamura-85	Gerek-79	Gökkan	Gönen-98
				
Gün-91	Halis	Hamitbey	Hanlı	Harman-22
				
Harmankaya-99	Hilar	İkizce-96	İkonya	İzgi-2001
				
Kale	Karacadağ-98	Karasu-90	Karatopak	Karmen
				
Kaynarca	Kayra	Kenanbey	Kırık	Kışla
				
Kilistra	Kirve	Konya-2002	Köprü	Kürşad

				
Lancer	Lütfübey	Meke	Meltem	Mesut
				
Mirsa	Müfitbey	Nacibey	Nenehatun	Nevzatbey
				
Nurkent	Nusrat	Osmaniyem	Özcan	Palandöken-97
				
Pandas	Reis	Saban	Sakin	Selamibey
				
Selçuklu	Seri-2013	Simge	Soyer-02	Sönmez-2001
				
Sultan-95	Şahika	Şanlı	Şehzade	Taner
				
Tekin	Toprak	Tosunbey	Tuğra	Ünsal-35
				
Yakamoz	Yavuz	Yıldırım	Yıldız-98	Yunus
				
Yüksel	Zencirci-2002	Zeybek-35	Ziyabey-98	

Şekil 3.6. Çalışmada kullanılan ekmeklik buğday çeşitlerine ait örnek görüntüler

3.3. Görüntü Ön İşleme

Görüntü ön işleme, özellik çıkarma ve model eğitimi gibi diğer işlemlerin etkin bir şekilde gerçekleştirilmesi için önemlidir. Bu çalışmada, toplam görüntü verisi çok büyük olduğundan, verilerin toplam hacmini azaltmak öncelikli olarak ele alınmış, böylece görüntülerin boyutunu önemli ölçüde azaltmak için ilk adım atılmıştır. Orijinal görüntülerin boyutu 4000x4000 pikseldir. Görüntülerin dış kenarlarını kaldırmak ve ayarlamak, görüntülerin uygun bir boyutunu ve çerçevesini elde etmek için, tüm görüntüler MATLAB programı (Image Batch Processor) kullanılarak kırpılmış ve yeniden boyutlandırılmıştır. İşlenen görüntülerin boyutu 250x250 pikseldir (Şekil 3.7).



Şekil 3.7. (a) Orijinal görüntü (4000x4000), (b) Kırpılmış ve yeniden boyutlandırılmış görüntü (250x250)

3.4. Özellik Çıkarma

Görüntü tabanlı problem çözmede, görüntüleri görmek, nesneleri tanımak ve tanımlamak için özellik çıkarma esastır (Mutlag et al., 2020). Özellik çıkarma, sınıf içi örüntü değişkenliğini azaltırken ve aynı zamanda sınıflar arası örüntü değişkenliğini artırmakta, böylece ham verilerden amaçlanan amaca uygun bilgi sağlamaktadır (Kumar ve Bhatia, 2014). Özellikler çıkarılır, vektörler olarak düzenlenir ve giriş örüntülerini takip eden hedef çıktıyı tanımak için sınıflandırma algoritmaları tarafından kullanılır.

3.4.1. Yönlendirilmiş gradyan histogramı

Bu çalışmada dokusal özellik olarak yönlendirilmiş gradyan histogramı (Histogram of Oriented Gradient, HOG) yöntemi kullanılmıştır. HOG özellikleri, MATLAB HOG çıkarıcı ve görüntü toplu işleme araçları (Image Batch Processing Tools) kullanılarak çıkarılmıştır. HOG özellik tanımlayıcısı, boyut azaltma olarak kabul edilen ilgisiz bilgileri atarken görüntülerin ana noktalarını çıkarır (Ochango et al., 2022). Çıkarılan özellikler, her görüntü için benzersiz olduklarından ve bir görüntüyü diğer görüntülerden açıkça ayırt ettiklerinden, bir görüntüyü diğer görüntülerden ayıran özelliklerdir. HOG özelliklerinin çıkarılması sürecinde, hücreler ve bloklar halinde düzenlenen yerel bölgelerin piksel değerlerinin yönlendirilmiş gradyanları veya toplam gradyan büyüklüğü ve yönü hesaplanır, histogram kutularına nicelenir ve özelliklerin sağlamlığını artırmak için normalleştirilir. Son olarak, özellikler bir sınıflandırma ve tespit modellerini eğitmek için girilebilecek bir vektöre dönüştürülür (Dalal ve Triggs, 2005; Chen et al., 2013) (Şekil 3.8). Satırlar (G_x), sütunlar (G_y) ve yön (θ) boyunca toplam gradyan büyüklüğünün (Total Gradient Magnitude, TGM) matematiksel ifadesi denklem (3.1) ve (3.2)'de verilmiştir.

$$TGM = \sqrt{G_x^2 + G_y^2} \quad (3.1)$$

$$\theta = \tan^{-1} \frac{G_x}{G_y} \quad (3.2)$$



Şekil 3.8. Buğdayın HOG özelliklerine ait bir örnek

3.4.2. Renk özellikleri

Renk özellikleri grubundan renk momentleri de dikkate alınmıştır. Renk momentleri, görüntüleri renk özelliklerine göre ayırt etmek için kullanılan ölçütlerdir. Hesalandıktan sonra, bu momentler görüntüler arasındaki renk benzerliğini ölçer ve

sınıflandırma modelleri geliştirmek için girdi olarak kullanılır (Kodituwakku et al., 2011; Wang et al., 2021). Buğday görüntüsüne ait RGB renk uzayında renk momentleri dağılımı Şekil 3.9’da verilmiştir. Bu çalışmada, renk dağılımlarını hesaplamak için RGB renk uzayı dikkate alınmış, böylece her bir R, G ve B kanalının ortalaması (denklem 3.3), standart sapması (denklem 3.4), çarpıklığı (denklem 3.5), basıklığı (denklem 3.6) ve entropisi (denklem 3.7), MATLAB görüntü toplu işlemcisi (Image Batch Processor) kullanılarak hesaplanmış ve her görüntü için 15 renk dağılımı ölçüsü hesaplanmıştır. Aşağıdaki matematiksel denklemler her bir anın ifadesidir; burada X_i , N ve P_i sırasıyla giriş verilerini, veri noktası sayısını ve sayımın normalleştirilmiş histogramını göstermektedir.

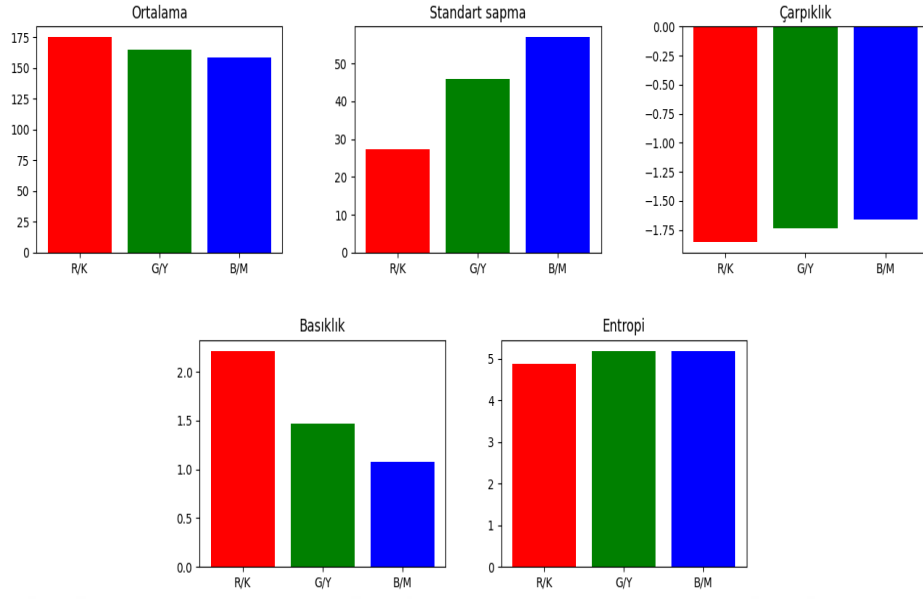
$$M = \frac{1}{N} \sum_{i=1}^N X_i \quad (3.3)$$

$$SD = \frac{1}{N-1} \sum_{i=1}^N (X_i - M)^2 \quad (3.4)$$

$$Sk = \frac{\frac{1}{N-1} \sum_{i=1}^N (X_i - M)^3}{SD^3} \quad (3.5)$$

$$k = \frac{\frac{1}{N-1} \sum_{i=1}^N (X_i - M)^4}{SD^4} \quad (3.6)$$

$$E = \sum_{i=1}^m P_i \log_2 P_i \quad (3.7)$$



Şekil 3.9 Buğday görüntüsünün RGB renk uzayında renk momentleri dağılımı

3.5. Sınıflandırma Yaklaşımları

Çalışma üç uygulamadan meydana gelmektedir. Bu uygulamalar, makine öğrenmesi (Machine Learning, ML), derin öğrenme (Deep Learning, DL) ve Fourier Dönüşümlü Kızılötesi (Fourier Transform Infrared, FTIR) spektroskopisinden oluşmaktadır.

Birinci uygulamada, HOG ve renk özelliklerinin çıkarılması, makine öğrenme modellerinin eğitimi, geliştirilmesi ve test edilmesi ele alınmıştır. Bu amaçla, buğday çeşitlerini sınıflandırmak için HOG ve renk özelliklerine dayalı ANN, SVM, RFC ve KNN modelleri geliştirilmiştir.

İkinci uygulamada, yeni önerilen CNN modeli ve öğrenme aktarımı (Transfer Learning) yöntemlerinden olan DenseNet201, MobileNet ve InceptionV3 popüler mimarileri kullanılmıştır.

Üçüncü uygulamada, FTIR ve ML modellerinin farklı spektral bölgelere göre uygulanması araştırılmıştır. ML modellerinden ANN, SVM, RFC ve KNN modelleri değerlendirilmiştir.

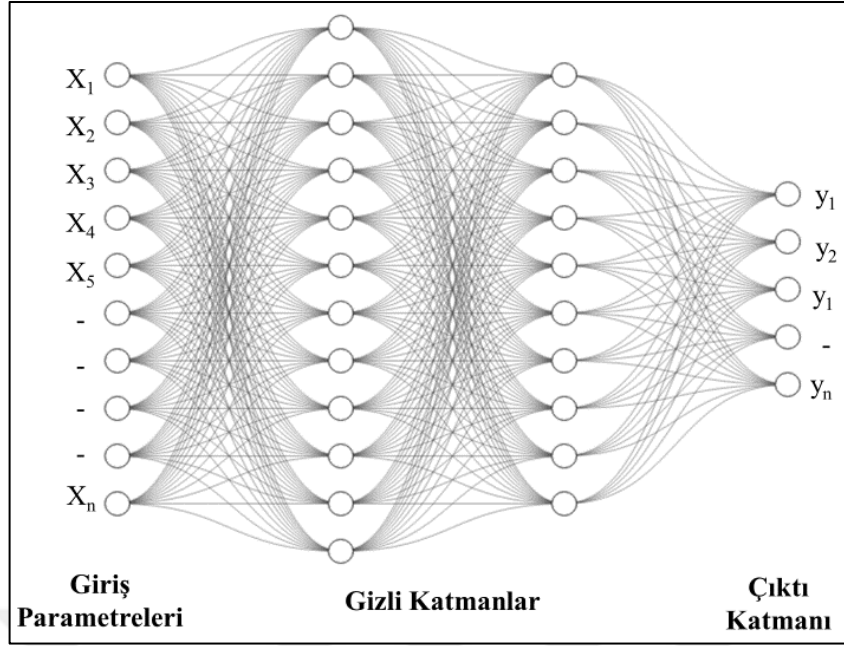
3.5.1. Makine öğrenme modelleri

Genel olarak, makine öğrenmesinin amacı, örüntüleri ve önceki bilgileri kullanarak görev akışını optimize etmektir (Benos et al., 2021). ML, belirli bir problem, veri ve bağlama bağlı olarak geniş bir uygulama yelpazesinde farklı biçimlerde ortaya çıkabilir (Smola ve Vishwanathan, 2008). Belirli bir girdi verisi için,

ML gizli ilişkileri keşfeder ve düzenlenmiş ve yeniden yapılandırılmış bir bilgi ortaya çıkarır. İnsanlara benzer şekilde, veri odaklı bir yaklaşım olduğundan, ML, veri veya deneyim miktarı arttıkça daha iyi performans gösterir (Vieira et al., 2020). Eğitilmiş bir ML modelinin, önceki deneyimlerinden öğrendiği örüntülere dayalı olarak öngörülemez verilere doğru ilişkiler ve tahminler sağlaması, öğrenme sürecinin temel sonucu olan genelleme yeteneği olarak bilinir (Domingos, 2012). Genel anlamda, ML yaklaşımları eğitim ve test olmak üzere iki aşamadan oluşur. Bu aşamalar, farklı parametrelerden oluşan örneklerin bir koleksiyonu olarak tanımlanabilecek verileri kullanarak gerçekleştirilir. Eğitim sürecinde, ikili, sayısal, sıralı ve nominal gibi her türlü özellik, öğrenme mekanizmasını kolaylaştırmak için genellikle diziler veya vektörler halinde düzenlenir ve girilir (Lopez-Arevalo et al., 2020). Öğrenme aşamasında, verilere ve her algoritmaya özgü diğer eğitim parametrelerine bağlı olarak, model kabul edilebilir bir performans düzeyine ulaştığında eğitim sona erer. Bir sonraki bölümde, sınıflandırma görevi için kullanılan dört ML algoritması (ANN, SVM, RFC ve KNN) tartışılmaktadır.

3.5.1.1. Yapay sinir ağları

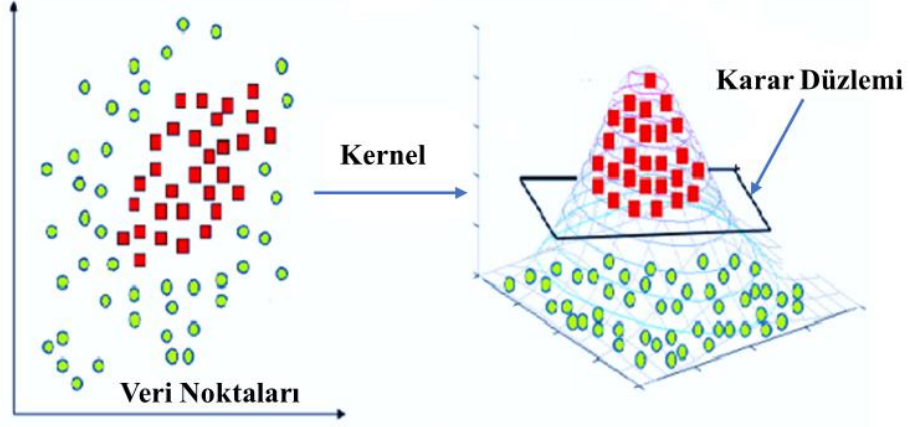
Yapay Sinir Ağları (Artificial Neural Network, ANN), regresyon ve sınıflandırma problemlerini çözmek için uygulanan çok popüler bir ML algoritmasıdır. Belirli bir yapıda bağlanmış bir dizi düğümden oluşan ve çeşitli katmanlarda düzenlenmiş olan bu algoritma, bağlantılarla donatılmış çok sayıda paralel işlemciden oluşur (Guresen ve Kayakutlub, 2011) (Şekil 3.10). Temel olarak, ANN'ler bağımsız özelliklerin girdi olarak girildiği giriş katmanına, perceptron ve backpropagation gibi öğrenme algoritmaları aracılığıyla öğrenme sürecinin gerçekleştiği gizli katmanlara ve sonuçları sağlamak için bir çıkış katmanına sahiptir. Tarım sektöründe, toprak sıcaklığı ve nem tahmini gibi su ve toprak yönetiminde (Patil et al., 2016; Nahvi et al., 2016; Johann et al., 2016), verim tahmini, çeşitlerin sınıflandırılması ve hastalık tespiti gibi ürün yönetimi alanlarında (Moshou et al., 2006; Pantazi, et al., 2016) ve hayvancılık üretiminde (Craninx et al., 2008) kullanılmaktadır.



Şekil 3.10. Yapay sinir ağları yapısı

3.5.1.2. Destek vektör makineleri

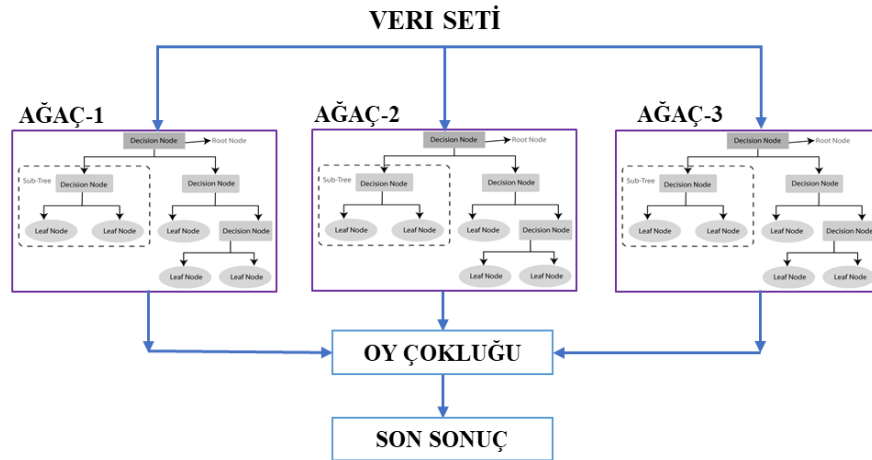
Destek Vektör Makineleri (Support Vector Machines, SVM), destek vektörleri olarak bilinen sınıf tarafının en temsili eğitim örneklerinin tanımlanmasıyla sınıflar arasındaki en uygun ayırma yüzeyinin bulunmasını içeren ML teknikleridir (Vapnik, 1995). SVM, temelde veri örneklerini sınıflandırmak için doğrusal bir ayırma hiper düzlemi oluşturan ikili bir sınıflandırıcıdır. Geleneksel SVM'lerinin sınıflandırma yetenekleri, eğitim veri seti doğrusal olarak ayrılabilir değilse, sınıfların doğrusal olarak ayrılabilmediği daha yüksek boyutlu bir uzayda verilen verilerin doğrusal olmayan bir projeksiyonunu oluşturarak bir çekirdek yöntemi ile önemli ölçüde geliştirilebilir (Mercier ve Lennon, 2003). SVM'ler, sınıflandırma, tespit ve regresyon problemlerinin yaygın olduğu tarımın hemen her alanında mahsul, toprak ve su yönetiminde yaygın olarak uygulanmaktadır (Zhang et al., 2017; Ebrahimi et al., 2017; Su et al., 2017) (Şekil 3.11).



Şekil 3.11 . Destek vektör makineleri yapısı

3.5.1.3. Random forest sınıflandırıcısı

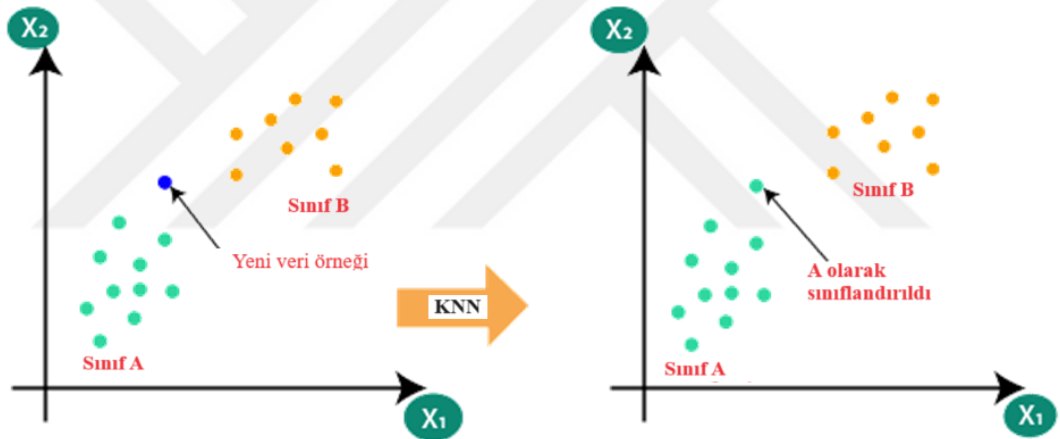
Random Forest Sınıflandırıcısı (Random Forest Classifier, RFC), bir grup karar ağacı üzerinde çalışan Ensemble öğrenme algoritmalarına aittir (Breiman, 2001). Eğitim sırasında, sınıflandırma ağaçlarının çeşitliliğini artırmak için bootstrapping uygulanır. Her bir Karar Ağacı (Decision Tree, DT), girdiler için bir sınıflandırma sağlar ve yeni bir sınıflandırma yapmak için kullanılabilir. Bundan sonra, RFC, sınıflandırmaları değerlendirir ve en çok oy alan tahmini seçer. RFC, çok sayıda bağımsız ağaçtan oluştuğu için nihai modelin ve sonraki sonuçların yorumlanması zor olsa da büyük miktarda veriyi işleme yeteneğine sahiptir (Oshiro et al., 2012). RFC, ürün ve meyve çeşitlerinin sınıflandırılması (Zawbaa et al., 2014; Fletcher ve Reddy, 2016; Fabiyi et al., 2020; Wójtowicz et al., 2021; Cinar ve Koklu 2022) ve hastalık tespiti (Adam et al., 2017; Mekha ve Teeyasuksaet 2021) gibi tarımsal alanlarda yaygın olarak kullanılmaktadır (Şekil 3.12).



Şekil 3.12. Random forest sınıflandırıcısının bir diyagramı

3.5.1.4. K-en yakın komşu algoritması

K-En Yakın Komşu (K-Nearest Neighbor, KNN) algoritması, eğitim veri setindeki örneklerle yeni örnekleri karşılaştırarak öğrenen örnek tabanlı modellerin bir türüdür. KNN algoritması, sınıflandırma problemlerini çözmek için kullanılan algoritmalarından biridir. En yakın komşu sınıflandırıcıları, verilen bir test öbeğini kendisine benzeyen eğitim öbekleri ile karşılaştırarak benzetme yoluyla öğrenmeye dayanır (Han et al., 2012). Hesaplama yapılırken, sınıflandırılacak veri ile öğrenme kümesindeki normal davranış verileri arasındaki benzerlik karşılaştırılarak önceden belirlenmiş eşik değerine bağlı olarak en yakın görünen k verinin ortalaması ile eşleştirme yapılır. KNN, çeşit sınıflandırmasında (Dheer ve Singh, 2019; Xu et al., 2021; Patil ve Lad, 2021; Aslam et al., 2021) ve hastalık tespitinde yaygın olarak kullanılmaktadır (Hossain et al., 2019; Hayit et al., 2014; Daniya ve Vigneshwari, 2024) (Şekil 3.13).



Şekil 3.13. KNN algoritması

3.5.1.5. Veri seti

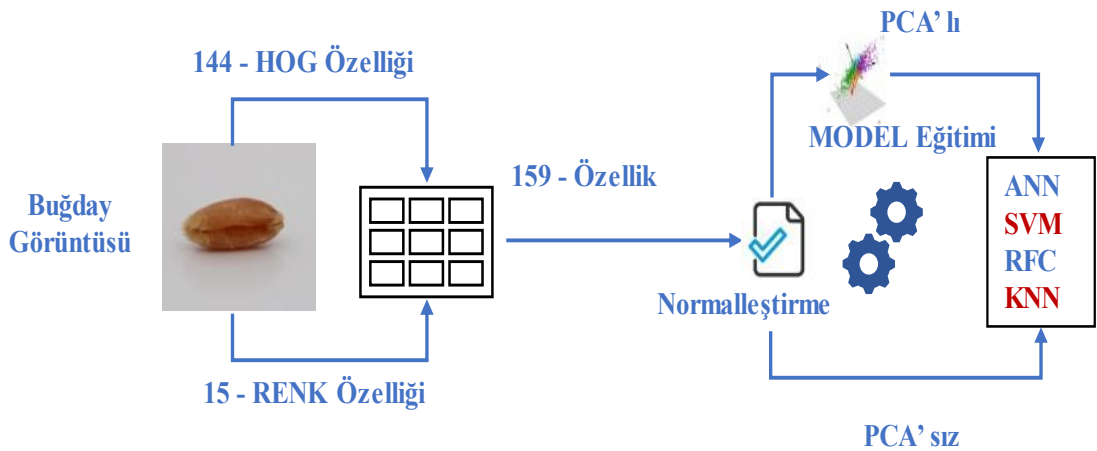
Bu çalışmada, 124 ekmeklik buğday çeşidi yer almıştır. Bu çeşitler için 3.750 satıra sahip, 144 HOG özelliği ve 15 renk momentini içeren nispeten büyük bir veri seti kullanılmıştır. Veri seti, modelleri eğitmek için kullanılan %76'lık eğitim, eğitim sırasında modelleri doğrulamak için kullanılan %12'lik doğrulama ve eğitimden sonra model genellemesini değerlendirmek için kullanılan %12'lik test setlerine bölünmüştür. Tüm modeller için veri seti, üç alt kümeye bölünmeden önce normalizasyon uygulaması yapılarak standartlaştırılmıştır. Veri setinin özeti Tablo 3.2'de sunulmuştur.

Tablo 3.2. Makine öğrenme modellerini eğitmek için kullanılan veri setinin özeti

Veri Seti Özellikleri	
Eğitim Seti	%76
Doğrulama Seti	%12
Test Seti	%12
Özellikler	159 (144 HOG Ve15 Renk)
Sınıf Sayısı	124
Her Sınıftaki Satır Sayısı	3.750
Toplam Satır Sayısı	465.000
Toplam Örnek Sayısı	~74 milyon

3.5.1.6. Renk ve dokuya dayalı model eğitimi

Birleştirilmiş 159 renk ve HOG özelliği kullanılarak dört ML modeli eğitilmiştir (Şekil 3.14). Her buğday çeşidi için 3.750 görüntü, renk ve doku özelliklerini çıkarmak için kullanılmıştır. Tüm 124 buğday çeşidi için, özellikleri çıkarmak üzere toplam 465.000 görüntü kullanılmıştır. Görüntüler, eğitim verilerinin kalitesini zenginleştiren, daha iyi temsil sağlayan ve birleşik karakteristik özelliklerden tek bir tanenin tanınmasına yardımcı olan on beş açıdan çekilmiş tek bir buğday tanesinin görüntülerini içermektedir. Veriler normalize edildikten sonra eğitim, doğrulama ve test kümelerine ayrılmış ve önceki bölümlerde açıklanan dört ML modeli (ANN, SVM, RFC, KNN) ile eğitilmiştir. Eğitimde, PCA'lı ve PCA'sız eğitim olmak üzere iki strateji benimsenmiş ve tüm eğitim süreci her model için aynı tutulmuştur.



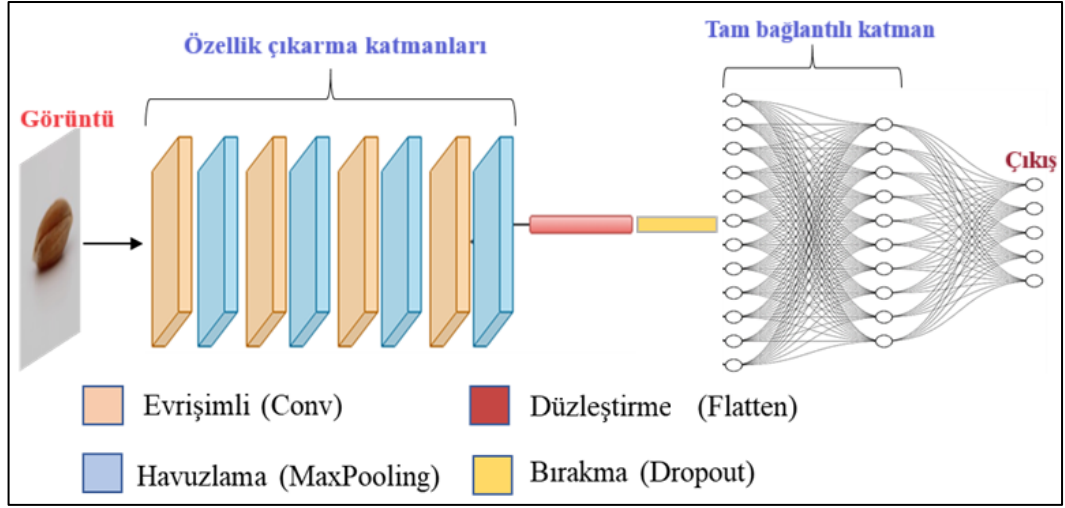
Şekil 3.14. Makine öğrenme modellerinin diyagram gösterimi

3.5.2. Derin öğrenme modelleri

Evrişimsel Sinir Ağları (CNN), girdi verisi olarak görüntüleri kullanan bir derin öğrenme algoritmasıdır. Derin öğrenmenin geniş sınıfında yer alan evrişimsel sinir ağı, yapay sinir ağının bir parçasıdır. CNN, görüntülerin özelliklerini kullanarak bu görüntülerin sınıflandırma işlemini gerçekleştirmektedir. CNN'in diğer sınıflandırma algoritmalarına kıyasla gerekli olan ön işlem miktarının düşük olmasının sebebi, filtreleri öğrenme yeteneğine sahip olmasıdır. Evrişimsel sinir ağları, evrişimli katmanlar (convolutional layers), doğrusal olmayan birimler (nonlinear units), havuzlama katmanları (pooling layers) ve tam bağlantılı katmanlardan (fully connected layers) oluşur. Bu katmanlardan geçen görüntüler, çeşitli işlemlerden geçirilerek derin öğrenme modeline uygun hale getirilir. Sınıflandırma probleminde, evrişimli katman, doğrusal olmayan birim ve havuzlama katmanı, özellikleri çıkarmak için özellik çıkarma katmanını oluşturur; tam bağlantılı katman ise sınıflandırma için sınıflandırma katmanını oluşturur. Havuzlama katmanı, çıkarılan özelliklerin boyutunu azaltırken, önemli bilgileri korumak için bir seçme işlemidir (Jiang et al., 2021). Bu algoritma serisi, görüntü sınıflandırma, nesne algılama, görüntü segmentasyonu ve daha birçok gelişmiş görev için kullanılabilen zengin özellik çıkarıcılar olarak hizmet eder (Patrick et al., 2021).

3.5.2.1. Önerilen CNN modeli

Bu çalışmada, 5.9 milyon eğitilebilir parametreye sahip nispeten sığ bir CNN modeli uygulanmıştır. Model, dört evrişimli katman, dört maksimum havuzlama katmanı (maxpooling layers), bir düzleştirme (flatten) ve bırakma (dropout) katmanı, her biri 256 nörondan oluşan iki yoğun katman (dense layers) ve 124 sınıflı bir çıkış katmanı (output layer) içermektedir (Şekil 3.15). Model, erken durdurma (early stopping) ve doğrulama kaybını (validation loss) izleyen model kontrol noktası geri aramaları ile eğitilmiştir.



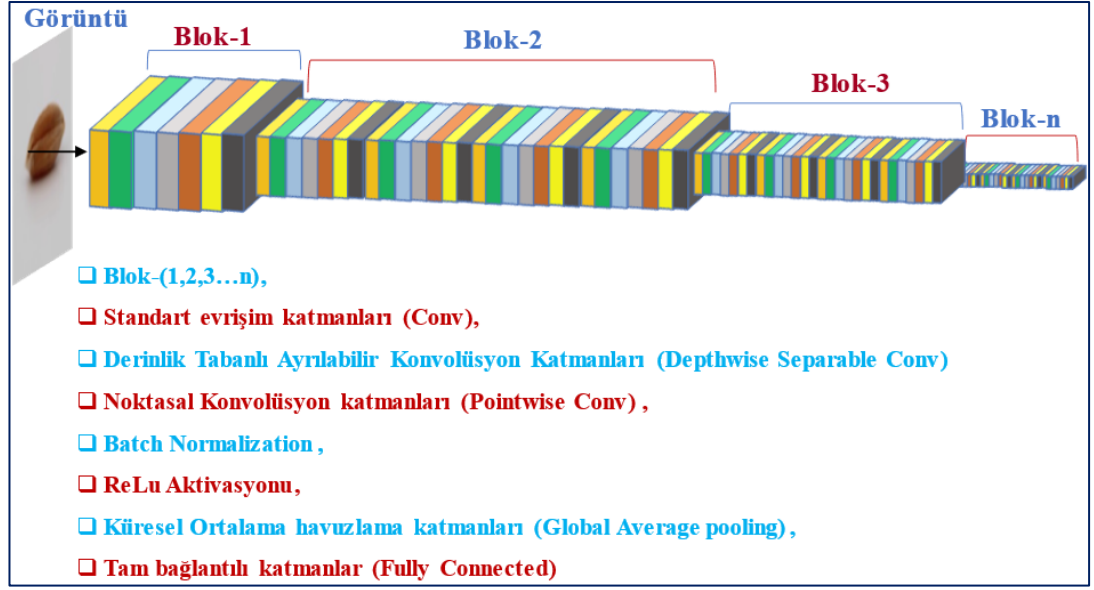
Şekil 3.15. CNN modelinin yapısı

3.5.2.2. Öğrenme aktarımı

Öğrenme Aktarımı (Transfer Learning), bir görev için eğitilen CNN'lerin ikinci bir görevdeki bir model için başlangıç noktası olarak yeniden kullanıldığı bir makine öğrenimi yaklaşımıdır (Lumini ve Nanni, 2019). Ağırlıkları rastgele başlatarak eğitimi sıfırdan başlatmak yerine, ağırlıkları büyük etiketli veri setleri, örneğin halka açık görüntü veri setleri gibi, üzerinde önceden eğitilmiş bir ağ kullanarak başlatılır. Bu çalışmada, geniş çaplı tipik bir veri seti olan ImageNet'te öğrenilen önceden eğitilmiş modelleri kullanmayı ve ardından bu modelleri, hedef veri seti üzerinde eğitilen belirli bir göreve aktarmayı ele alıyoruz. Çalışmada kullanılan popüler Öğrenme Aktarımı modelleri, bir sonraki bölümde sunulmaktadır.

3.5.2.3. MobileNet

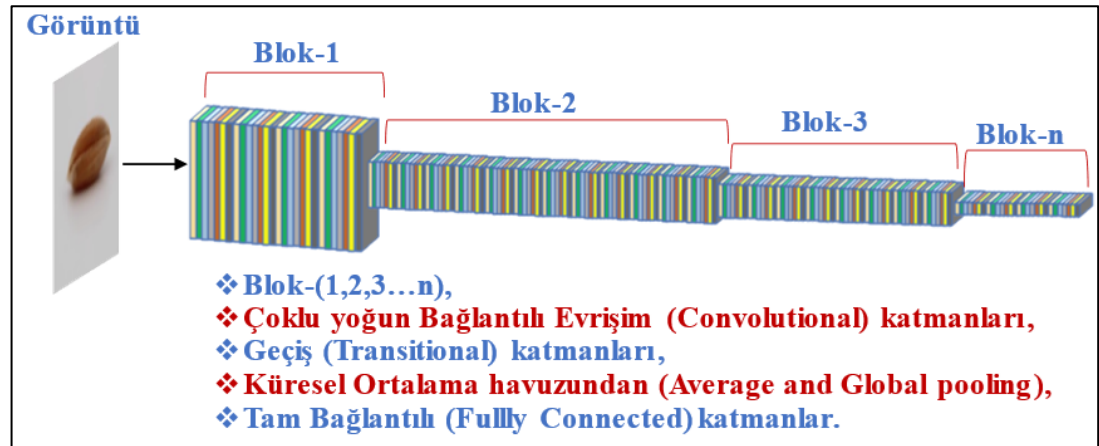
MobileNet, mobil ve gömülü uygulamalar için (Howard et al., 2017) tarafından geliştirilen bir evrışimsel sinir ağı mimarisidir. MobileNet, standart evrışimli katmanlardan farklı, ana yapı taşı olarak derinlik bazlı ayrılabilir evrışimle bilinir. Hafif bir model olması nedeniyle, tarımsal uygulamalarda MobileNet, çeşitlerin sınıflandırılması (Venkatesh et al., 2021; Miah et al., 2021; Alhawas ve Tüfekci, 2022) ve hastalık tespiti (Bir et al., 2020; Elfatimi et al., 2022) için yaygın olarak kullanılmaktadır. Şekil 3.16, MobileNet'in bir temsili göstermekte olup, 12.9 milyon eğitilebilir parametreye sahip bir model uygulanmıştır.



Şekil 3.16. MobileNet Yapısı

3.5.2.4. DenseNet201

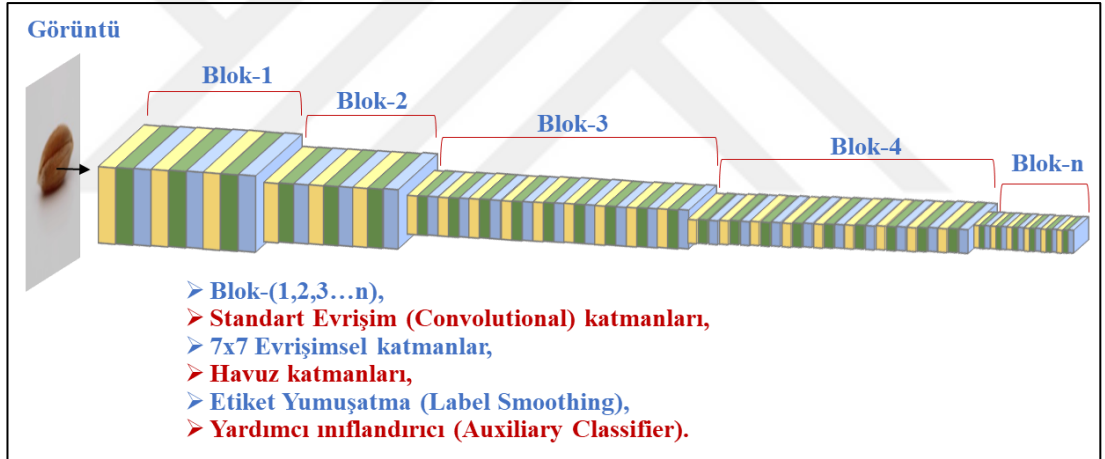
DenseNet, ağdaki katmanlar arasında maksimum bilgi akışını sağlamak amacıyla her katmanı diğer tüm katmanlara ileri beslemeli bir şekilde bağlayan bir ağıdır (Huang et al., 2017). ImageNet gibi halka açık veri setlerinde umut verici doğruluk elde eden bir CNN mimarisidir. Öğrenme aktarımı kullanılarak çeşitli DenseNet mimarileri, tarım alanındaki sınıflandırma ve ürün tanıma (Lu et al., 2021; Verma et al., 2022; Salim et al., 2023) ve bitki hastalığı tespiti (Tiwari et al., 2021; Bakr et al., 2022; Nandhini et al., 2022) gibi farklı görevler için uygulanmıştır. Şekil 3.17, DenseNet modelinin genel mimarisini göstermektedir. Bu çalışmada, 24.1 milyon eğitilebilir parametreye sahip DenseNet201 modeli benimsenmiştir.



Şekil 3.17. DenseNet Yapısı

3.5.2.5. InceptionV3

Inception, (Szegedy et al., 2015) tarafından geliştirilen bir evrişimsel sinir ağı mimarisidir. Nesne tespiti ve görüntü analizi için kullanılan bu mimari, başlangıçta GoogLeNet mimarisi için bir modül olarak geliştirilmiştir. Inception, paralel olarak gerçekleştirilen birden fazla evrişim işlemi ile karakterize edilen inception modülleriyle bilinir. Bu çalışma için benimsenen InceptionV3 modeli, 7 x 7 evrişimlerin ayrıştırılması, etiket yumuşatma ve ağın alt katmanlarına etiket bilgisini yaymak için kullanılan bir yardımcı sınıflandırıcı ile karakterize edilir. InceptionV3, tarım alanında hastalık tespiti (Teja et al., 2021; Upadhyay, 2022; Shah et al., 2023), tür ve çeşit tanıma (Al Sahili ve Awad, 2022; Malvade et al., 2023), zararlı ve yabancı ot sınıflandırması (Tetila et al., 2020; Subeesh et al., 2022) gibi çeşitli görevlerde yaygın ve başarılı bir şekilde kullanılmaktadır. Şekil 3.18, InceptionV3'ün temel temsilini göstermektedir. Bu çalışmada kullanılan InceptionV3 modeli, 18.9 milyon eğitilebilir parametreye sahiptir.



Şekil 3.18. InceptionV3 Yapısı

3.5.2.6. Veri seti

Bu çalışmada 124 ekmeklik buğday çeşidi sınıflandırılmıştır. Her bir çeşide ait 3.750 görüntü yer almıştır. Toplam 465.000 görüntü DL modellerinde kullanılmıştır. Bu görüntüler, modelleri eğitmek için kullanılan %80'lik eğitim, eğitim sırasında modelleri doğrulamak için kullanılan %10'luk doğrulama ve eğitimden sonra model genellemesini değerlendirmek için kullanılan %10'luk test setlerine bölünmüştür.

3.5.3. Fourier dönüşümlü kızılötesi spektroskopisi

Gıda kalitesi söz konusu olduğunda sağlık ve güvenlik öncelikli unsurlardır.

Tüketicinin gelirinin ve bilincinin artmasıyla birlikte kaliteli ve güvenli gıdaya olan talep de artmaktadır. Küresel içecek ve gıda üretim ve işleme endüstrileri, ürünlerini güvenli tutmayı amaçlamaktadır. Hedefleri, gıda ürünlerinin sağlıklı ve herhangi bir kötü etkiden arınmış olduğundan emin olmaktır (Zareef et al., 2020). Gıda güvenliği güvencesi süreci, üretimden teste kadar uzanır ve niteliksel ve niceliksel olarak hızlı, basit ve tahribatsız olması gerekir. Ayrıca, yazılım tabanlı değerlendirmeye olanak tanıyan veri toplamayı içerir (Hashimoto et al., 2008).

Spektroskopi, herhangi bir gıda örneğinin içeriğiyle ilgili yararlı bilgiler elde etmeye yardımcı olduğu için, endüstriyel gıda kalite kontrolü ve izlemesinde yaygın olarak uygulanmaktadır. Temel olarak spektroskopi, ışık ve fiziksel nesnelerin etkileşimi ile ilgili bir çalışma alanıdır. Spektroskopik testlerde, ışık ışını çeşitli yöntemler yardımıyla bir nesneye ışınlanır (Ricci et al., 2013). Işık ve numunenin etkileşimi sonucu ortaya çıkan spektrum kaydedilir ve spesifik numuneye ait yorumlanabilir moleküler bilgiler sağlar. FTIR spektroskopisi titreşimsel spektroskopik uygulamaların bir türüdür. Matematiksel Fourier dönüşümü yöntemini kullanarak ışığın kızılötesi yoğunluğuna karşı dalga sayısını ölçer. Elektromanyetik ışık dizisinin kızılötesi bölgesi 10cm^{-1} ile 14000cm^{-1} arasındadır ve yakın dalga boyu kızılötesi (near wavelength infrared, NIR: $4000\sim14000\text{cm}^{-1}$), orta dalga boyu kızılötesi (medium wavelength infrared, MIR: $400\sim4000\text{cm}^{-1}$) ve uzak dalga boyu kızılötesi (far wavelength infrared, FIR: $4\sim4000\text{cm}^{-1}$)'den oluşur.

FTIR, bir dizi kimyasal içeriği algılayabilen ve moleküler yapıdaki değişikliklere duyarlı olan hızlı, tahribatsız ve zaman kazandıran bir tekniktir. FTIR, tüm numunenin kimyasal bileşimine ve fiziksel durumuna dayalı bilgi sağlar (Cocchi et al., 2004). Yüksek sinyal-gürültü oranına sahip olması FTIR spektroskopisinin ana avantajıdır (Zhang et al., 2022). FTIR spektroskopisi buğday, pirinç, çay, üzüm gibi tarımsal ürünlerin farklı uygulamalarına başarıyla uygulanmaktadır (Amir, et al., 2013; Ribero, et al., 2020; Zhou et al., 2020; Murru, et al., 2019). FTIR, çeşitli buğday kalite parametrelerini ölçmek için hızlı ve hassas bir teknik olarak aktif bir şekilde kullanılmaktadır. Ayrıca, FTIR dedektörlerinin hassasiyeti ve doğruluğu, birçok algoritma ile birleştğinde, kantitatif analiz için kızılötesinin pratik kullanımını önemli ölçüde artırmıştır (Dowell et al., 2006).

Ayrıca, gıda ve tarım ürünlerinin yeni ve kayda değer tahribatsız ölçümlerinin çoğu, optik ve spektroskopik prosedürlere dayanmaktadır. Spektroskopi

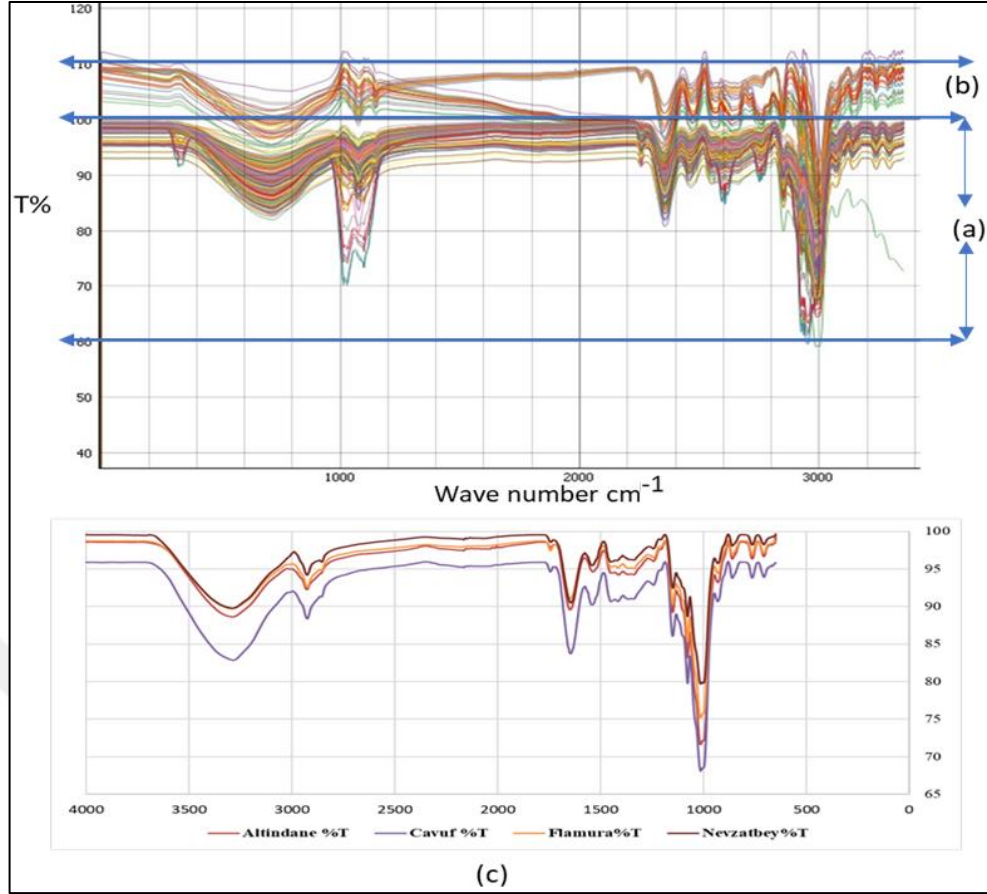
standardizasyonları, spektral verileri, incelenen numunenin kimyasal veya fiziksel bilgileriyle ilişkilendirir (Aleixandre-Tudo et al., 2020). Bununla birlikte, dolaylı ölçümlere dayandığından, oldukça karmaşık ve geniş spektrumların çıplak gözle yorumlanması neredeyse imkansızdır ve ilgili spektrumlardan analitik bilgi çıkarmak için matematiksel ve istatistiksel araçlarla kalibrasyon gerektirir (Porep et al., 2015). Daha sonra toplanan spektral veriler, analiz edilir. Bu yöntemle, tahmin veya sınıflandırma modelleri kullanılarak içerik miktarının belirlenmesi (Qi et al., 2023; Wang et al., 2023), belirli yabancı maddelerin tanımlanması (Gok vd, 2015; Gouvinhas et al., 2015; Reis et al., 2017; Miaw et al., 2018; Lim et al., 2023), menşe belirleme, çeşit sınıflandırma (Ozturk et al., 2023; Yu et al., 2023; Reale et al., 2023) gibi sorunlar etkili bir şekilde çözülebilir.

3.5.3.1. FTIR ve ML modeli

Bu çalışmada, dört ekmeklik buğday çeşidini sınıflandırmak için FTIR spektroskopisi ile birlikte bir ML modeli önerilmiştir. ML modellerinden ANN, SVM, RFC ve KNN kullanılmıştır. FTIR spektral verileri bir FTIR spektrometresi ile toplanmıştır (Şekil 3.19). Her biri 80 örnek içeren Altındane, Çavuş, Flamura-85 ve Nevzatbey olmak üzere 4 buğday çeşidine ait 320 spektral örnek toplanmıştır. Tüm örneklerin toplam spektrumları, geçirgenlik ve dalga sayılarının yüzdesi olarak Şekil 3.20'de verilmiştir. Sunulan şekilde, spektral ön işleme adımlarını takip eden veri temizleme sürecinin bir parçası olarak 100'ün üzerinde olan hatalı okumalar çıkarıldı. Bu adımdan sonra baz çizgisi düzeltilmesi gibi spektral ön işleme adımları uygulandı.



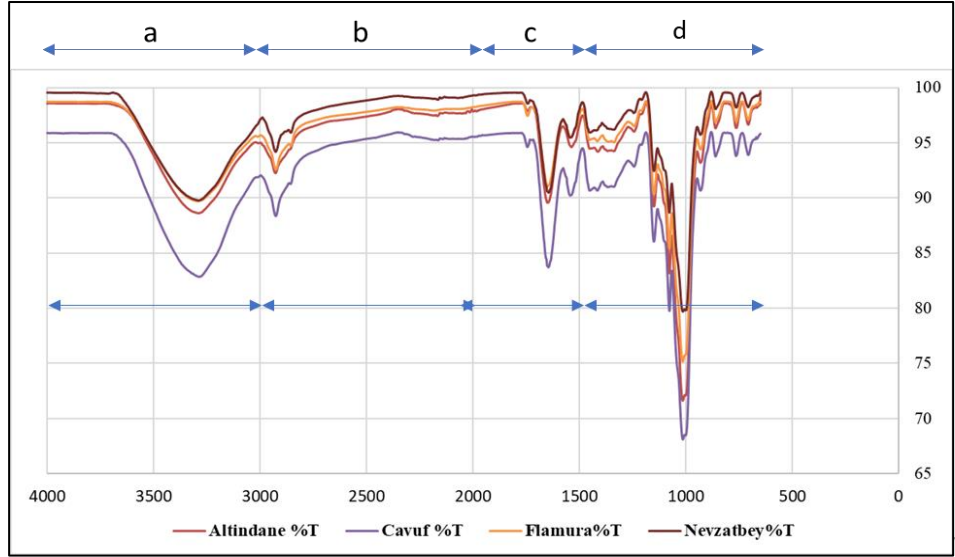
Şekil 3.19. Fourier Dönüşümlü Kızılötesi Spektroskopisi



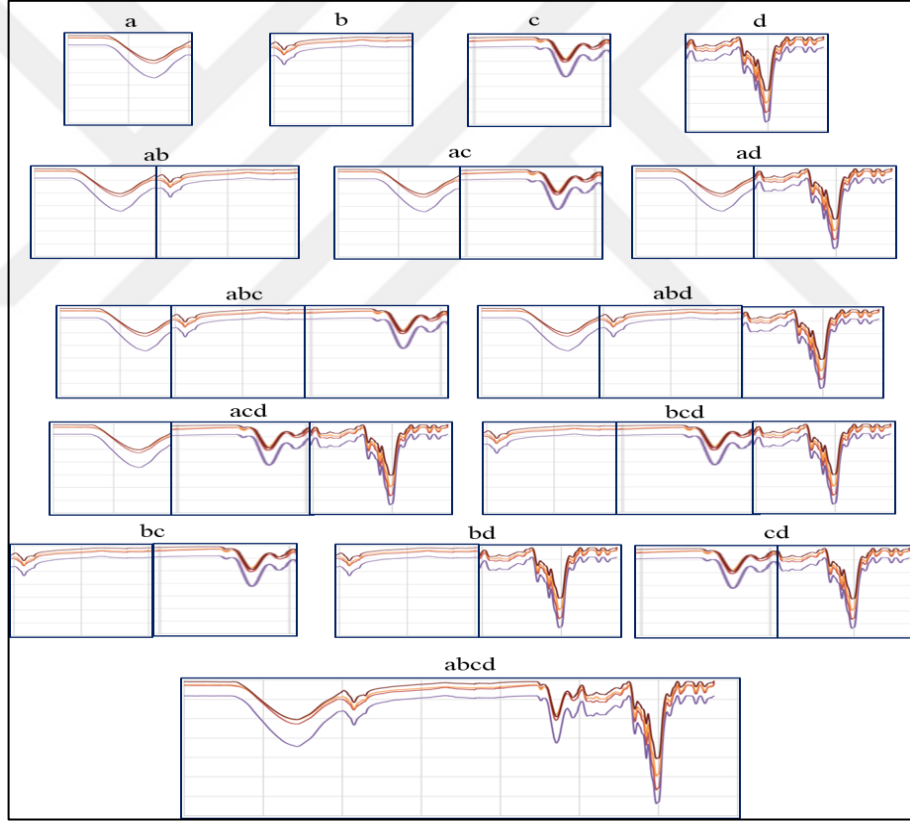
Şekil 3.20. Dört buğday çeşidinin spektroskopik verileri, (a) Ön işlemeden sonra modeli eğitmek için kullanılan standart okuma, (b) Hatalı okumalara sahip çıkarılmış örnekler, (c) Dört buğday çeşidinin örnek FTIR spektral okuması.

3.5.3.2. Girdi verileri

Spektral veriler, bölüm-a ($4000-3000\text{cm}^{-1}$), bölüm-b ($3000-2000\text{cm}^{-1}$), bölüm-c ($2000-1500\text{cm}^{-1}$) ve bölüm-d ($1500-650\text{cm}^{-1}$) olmak üzere dört bölüme ayrılmıştır (Şekil 3.21). Tam spektral veri, ayrı veya tam bir veri seti olarak kullanılmış ve dört bölümün on dört kombinasyonu hazırlanmıştır. Bu nedenle, model eğitimi sürecinde veri girişi olarak toplam on beş veri seti kullanılmıştır (Şekil 3.22).



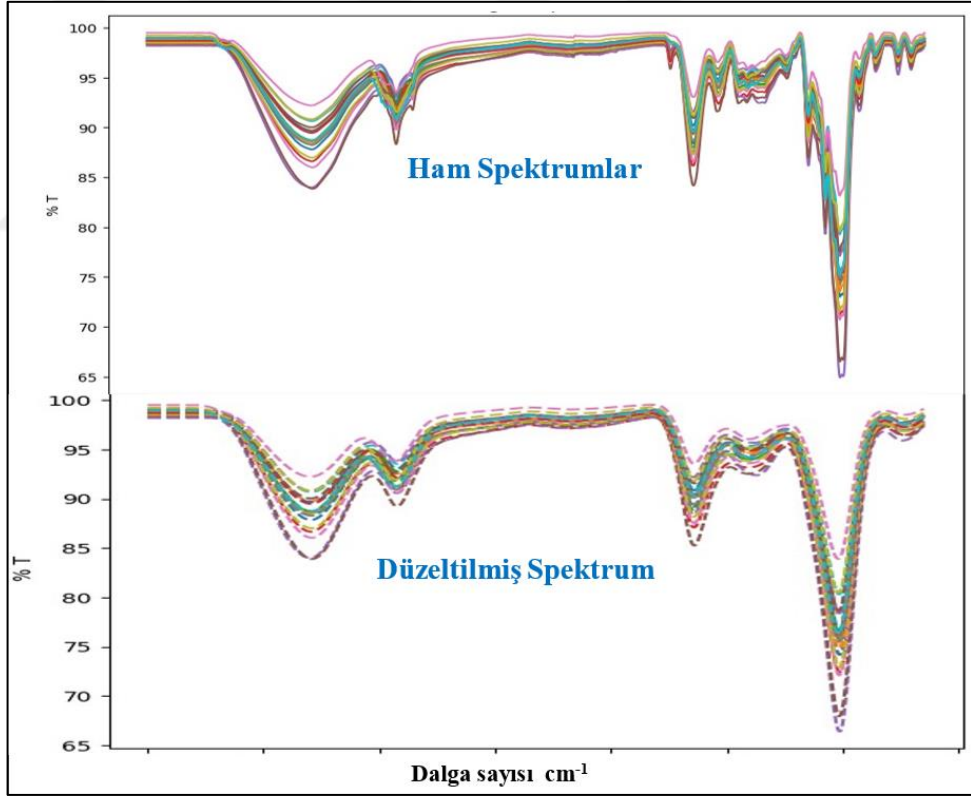
Şekil 3.21. Dört buğday çeşidinin tam FTIR spektral verileri



Şekil 3.22. FTIR spektral verilerinin dört bölümünün on beş kombinasyonu

3.5.3.3. Temel düzeltme

Bir FTIR spektrometresinde, baz çizgisi sıfır gözlemin gözlemlendiği noktayı ifade eder. İdeal bir durumda, gözlem olmadığı için baz çizgisi doğrusal bir çizgi olmalıdır. Bununla birlikte, çeşitli çevresel ve enstrümantal faktörler, baz çizgisinde kayma oluşturabilir. Sonuç olarak, bu kayma, her bir örnek okumasını etkiler ve yanlış okumaya neden olur (Zhang et al., 2022). Bu nedenle, daha iyi bir nicel ve nitel analiz için spektral doğruluğu iyileştirmek ve bizim durumumuzda olduğu gibi sınıflandırma modeli geliştirmek için baz çizgisi düzeltmesi olarak bilinen bir ön işleme adımı gereklidir. Bu nedenle, bir baz çizgisi düzeltmesi gerçekleştirmek için, eğrileri düzeltmek amacıyla asimetrik en küçük kareler bazlı düzeltme yöntemi benimsenmiştir. Bu yöntem, spektral pik noktalarının sol ve sağ yanlarında asimetrik işlemler uygulayarak spektral veri noktaları aracılığıyla en küçük kare yöntemini kullanarak en uygun eğriyi bulmayı içerir (Şekil 3.23).



Şekil 3.23. Orijinal ve baz çizgisi düzeltilmiş spektrumlardan bir örnek

3.5.3.4. Aşırı örnekleme

Aşırı örnekleme (Oversampling), dengesiz veri setlerinin etkisini telafi etmek için kullanılan bir tekniktir. Spektral veri setinden, toplam veri boyutuna kıyasla, önemli veri noktaları hatalı okumalar nedeniyle elenmiştir. Bu nedenle, iki küçük sınıfı

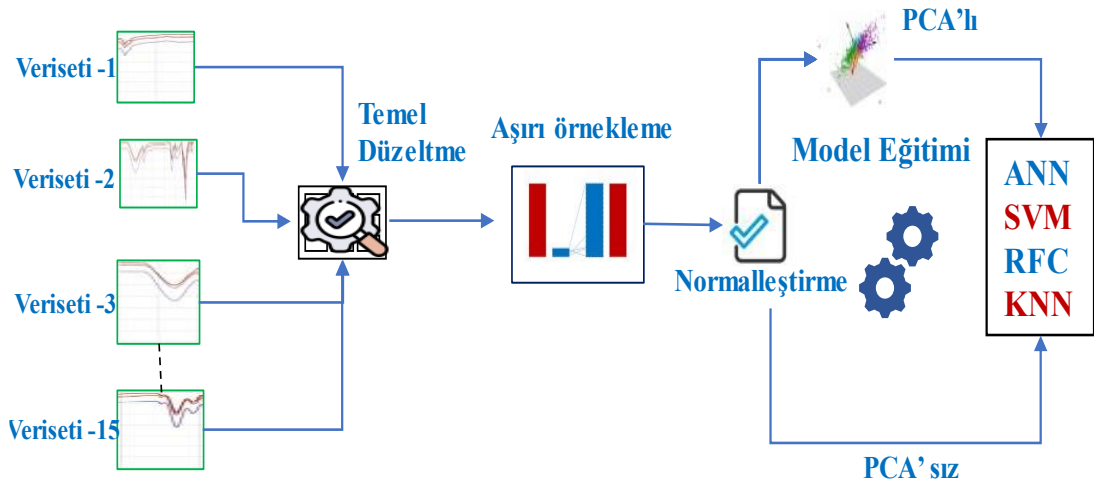
telafl etmek ve etkili model eğitimi sağlamak için veri kümesini dengelemek amacıyla aşırı örnekleme tekniğı uygulanmıştır. Özellikle, uygulanan aşırı örnekleme tekniğı, Sentetik Azınlık Aşırı Örnekleme Tekniğı (Synthetic Minority Oversampling Technique, SMOTE) olmuştur. SMOTE, basit ve yaygın olarak kullanılan bir azınlık örnekleme aracıdır. Rastgele seçilen bir veri noktası ile K-en yakın komşularından birine bağlayan doğru parçası boyunca örnekler üretir (Chawla et al., 2002). Basitliğı ve uyarlanabilirliğı, SMOTE'nin spektroskopi gibi dengesiz veri kümesi durumlarında yaygın olarak benimsenmesini sağlamıştır (Lemaître et al., 2017; De medeiros et al., 2022).

3.5.3.5. Temel bileşen analizi

Boyut azaltma sağlamak ve dolayısıyla modelin aşırı uyumunu önlemek için PCA kullanılmıştır. Toplanan FTIR spektral verileri 3.501 sütun ve 320 satırdan oluşmaktadır. Yüksek boyutluluk ve nispeten az veri miktarı nedeniyle aşırı öğrenme potansiyel bir risktir. Bu nedenle, verileri, açıklanan varyans yüzdesine dayalı olarak yeni bir dönüşüme dönüştüren bir teknik olan PCA uygulanmıştır (Khalid et al., 2014). PCA, yalnızca bağımsız değişkenleri ele alması nedeniyle denetimsiz ML tekniğı olarak kabul edilir (Jolliffe, 2002). Bu nedenle, bu çalışmada PCA ön işleme aşamasına dahil edilmiştir. Her veri kümesi analiz edilmiş ve en fazla varyansı koruyan bileşenler, modelleri eğitmek için kullanılmıştır.

3.5.3.6. Spektral verilere dayalı model eğitim yaklaşımı

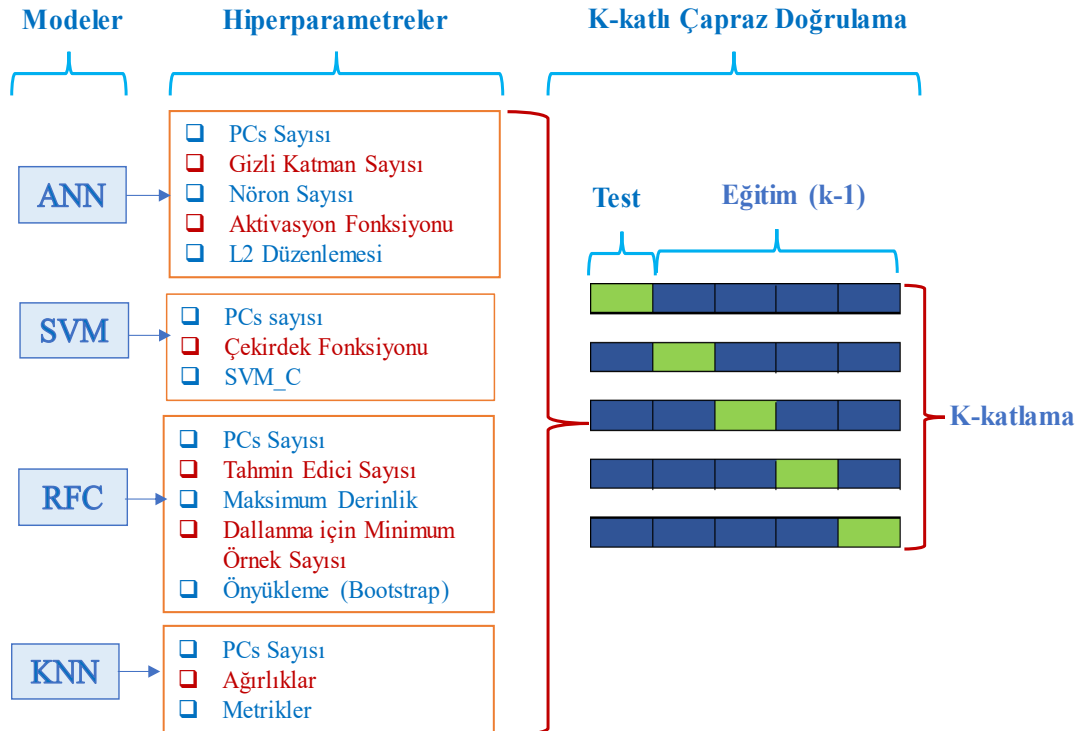
ML modellerinin eğitimi, veri setlerinin kombinasyonu, baz çizgisi düzeltmesi, aşırı örnekleme ve dört modelin K-katlı çapraz doğrulama kullanılarak eğitilmesini içermektedir (Şekil 3.24). Süreç asimetrik en küçük kare baz çizgisi düzeltmesi (asymmetric least square baseline correction), dengesiz sınıfların aşırı örneklenmesi (oversampling the imbalanced classes), normalleştirme (normalizing) ve PCA'lı ve PCA'sız olmak üzere iki hattı (pipeline) gerçekleştirirken on beş veri kümesi üzerinde yinelenmektedir. 15 veri kümesi, dört ayrı model ve iki eğitim seçeneğı (PCA'lı ve PCA'sız) kullanılarak toplamda 120 model eğitilmiştir.



Şekil 3.24. Makine öğrenme modellerinin eğitim süreci

3.5.3.7. Hiperparametre ayarlama

Her ML modelinin, aşırı öğrenme riskini azaltırken en iyi doğruluğu elde etmek için ince ayar yapılması gereken hiperparametreleri vardır. Bu çalışmada, ANN, SVM, RFC ve KNN olmak üzere dört model, grid arama stratejisi kullanılarak ince ayar yapılmış, mümkün olan en iyi modeli elde etmek için her modelle ilgili hiperparametrelerin bir listesi kullanılmıştır (Şekil 3.25). Bu durumda veriler, toplam veri nispeten az olduğu için eğitim, doğrulama ve test ayrımı yerine eğitim ve test kümelerine ayrılmıştır.



Şekil 3.25. Grid arama hiperparametre ayarlaması ve K-katlı çapraz doğrulama

3.6. Performans Ölçütleri

Modellerin performansını değerlendirmek için doğruluk (Accuracy), kesinlik (Precision), duyarlılık (Recall) ve F1-skoru gibi performans metrikleri kullanılmıştır. Bu metrikler, Doğru Pozitifler (TP), Doğru Negatifler (TN), Yanlış Pozitifler (FP) ve Yanlış Negatifler (FN) üzerinden hesaplanır. Tüm metriklerin ifadeleri (3.8-3.11) denklemlerinde verilmiştir.

$$\text{Doğruluk} = \frac{TP + TN}{TP + TN + FP + FN} \quad (3.8)$$

$$\text{Kesinlik} = \frac{TP}{TP + FP} \quad (3.9)$$

$$\text{Duyarlılık} = \frac{TP}{TP + FN} \quad (3.10)$$

$$F1 - \text{SKoru} = 2 \cdot \frac{\text{Precision} \cdot \text{Recall}}{\text{Precision} + \text{Recall}} \quad (3.11)$$

3.7. Makine Öğrenme Sınıflandırıcılarının İstatistiksel Analizi

Bu çalışmada, çoklu algoritmaları karşılaştırmak için uygun olan parametrik olmayan bir istatistiksel analiz olan Friedman testi uygulanmıştır. Friedman testi, farklı veri kümelerindeki algoritmalar arasında anlamlı bir fark olup olmadığını belirlemeyi amaçlamaktadır (Trawiński et al., 2012). Testler, her bir veri seti için algoritmaların sıralamalarını belirler. Friedman'a göre null hipotez, anlamlı bir fark olmadığı göz önünde bulundurulduğunda tüm algoritmaların sıralamasının eşit olduğunu ifade eder.

Ancak, null hipotezi reddedilirse, algoritmaların önemli ölçüde farklı olduğu boyutu belirlemek için bir sonraki adımda bir post-hoc prosedürü gerçekleştirilir. Post-hoc prosedürlerden biri olan Nemenyi prosedürü, grupların birbirinden önemli ölçüde farklı olup olmadığını belirlemek için çiftli karşılaştırma yapmak üzere benimsenmiştir (García ve Herrera 2008). FTIR spektroskopisi verilerine dayalı buğday çeşitleri sınıflandırıcılarının geliştirilmesiyle ilgili bölümde, tüm veriler dört bölüme ayrılarak on beş veri kümesi kombinasyonu kullanılmış, ayrıca PCA'da yer aldığı için toplam otuz veri kümesi kullanılmıştır. Dolayısıyla, bu otuz veri seti ve dört model kullanılarak istatistiksel bir analiz gerçekleştirilmiştir (Tablo 3.3). İlk aşamada,

sınıflandırıcıların (ANN, SVM, RFC ve KNN) performansları arasında anlamlı bir fark olup olmadığını test etmek için Friedman testi yapılmıştır. Ayrıca, Nemenyi testi hangi model çiftlerinin önemli ölçüde farklı olduğunu göstermektedir (Demšar, 2006).

Tablo 3.3. Çoklu veri kümeleri ve model performansları örneği

	Model-1	Model-2	Model-3	Model-4
Data-1	Acc-1	Acc-1	Acc-1	Acc-1
Data-2	Acc-2	Acc-2	Acc-2	Acc-2
Data-3	Acc-3	Acc-3	Acc-3	Acc-3
-	-	-	-	-
-	-	-	-	-
-	-	-	-	-
Data-30	Data-30	Data-30	Data-30	Data-30

3.8. Paketler ve İşlemciler

Buğday sınıflandırma modellerinin eğitimi, Akya-Cuda kümesindeki Türk Ulusal Bilim e-Altyapısı (TRUBA) sunucularında gerçekleştirilmiştir. Küme 24 adet Supermicro 1029GQ-TRT sunucudan oluşmaktadır. Her sunucu iki adet Intel Xeon Scalable Gold 6148 işlemci ile donatılmış olup, sunucu başına 2.4 GHz işlem hızına sahip toplam 40 işlemci çekirdeği sağlamaktadır. Her sunucu, her biri 16GB HBM belleğe sahip dört Nvidia Tesla V100 GPU kartıyla donatılmıştır. Hesaplama altyapısının özelliklerinin bir özeti Tablo 3.4'te verilmiştir. Ayrıca, NumPy, Pandas, Scikit-learn, Keras ve TensorFlow isimlerindeki python veri işleme, ön işleme ve ML kütüphaneleri kullanılmıştır.

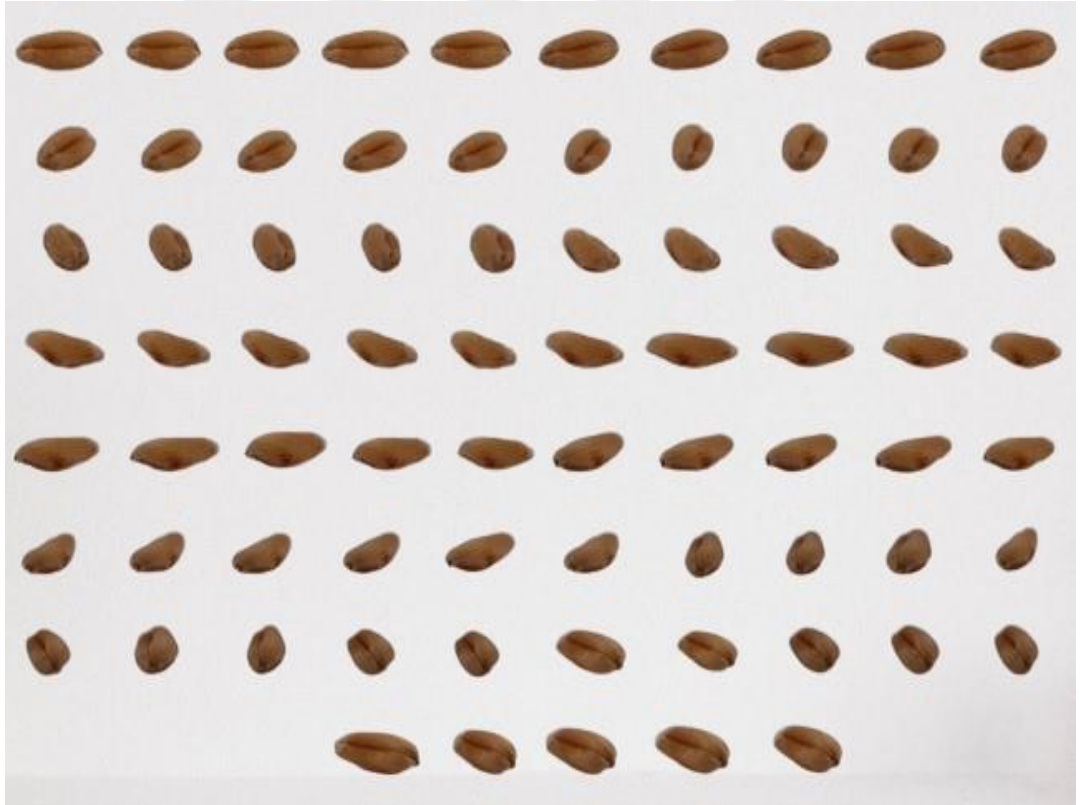
Tablo 3.4. TRUBA'da kullanılan hesaplama kaynaklarının özellikleri

Bileşenler	Özellikler
Küme	Akya-Cuda
Sunucu Sayısı	24
Sunucu Modeli	Supermicro 1029GQ-TRT
İşlemci Modeli	Intel Xeon Scalable Gold 6148
İşlemci Hızı	2.40 GHz
Toplam İşlemci Çekirdeği	Sunucu başına 40 çekirdek (2 CPUs x 20 cores)
GPU Modeli	Nvidia Tesla V100
GPU Sayısı	Sunucu başına 4
GPU Belleği	16GB HBM per GPU
Toplam Teorik GPU Performansı	31200 Gflops per server (4 GPUs x 7800 flops)
Toplam Bellek	384GB per server

4. BULGULAR

4.1. Görüntü Elde Etme

Buğday çeşitlerinin görüntüleri, Canon EOS 2000D kullanılarak Foldio360 olarak bilinen ve döner tablaya sahip özel bir görüntü alma düzeneğinde çekilmiştir. Döner tabla, manuel olarak yapılması son derece sıkıcı ve zaman alıcı olabilecek buğday tanelerinin çoklu görünümlerinden görüntü alma işlemini kolaylaştırmıştır. Ayrıca, birden fazla kameraya duyulan gereksinim, döner tabla ve sürekli çekim uygulaması ile karşılanmıştır. Birden fazla görünümünden görüntü almak, nesnenin farklı perspektiflerinden bilgi toplamaya yardımcı olur ve sınıflandırmayı ve nihayetinde modellerin geliştirilmesini geliştirir. Ayrıca, kamera her dört saniyede bir beş görüntü yakalayacak şekilde ayarlandığından, çıplak gözle benzer ve gereksiz görünen görüntülerden bile küçük farklılıklar yakalanmıştır (Şekil 4.1). Her 124 buğday çeşidi için, örneklerin belirli bir çeşidi temsil ettiğinden emin olmak amacıyla rastgele elli tane seçilmiştir. Toplamda 465.000 görüntü elde edilmiştir.

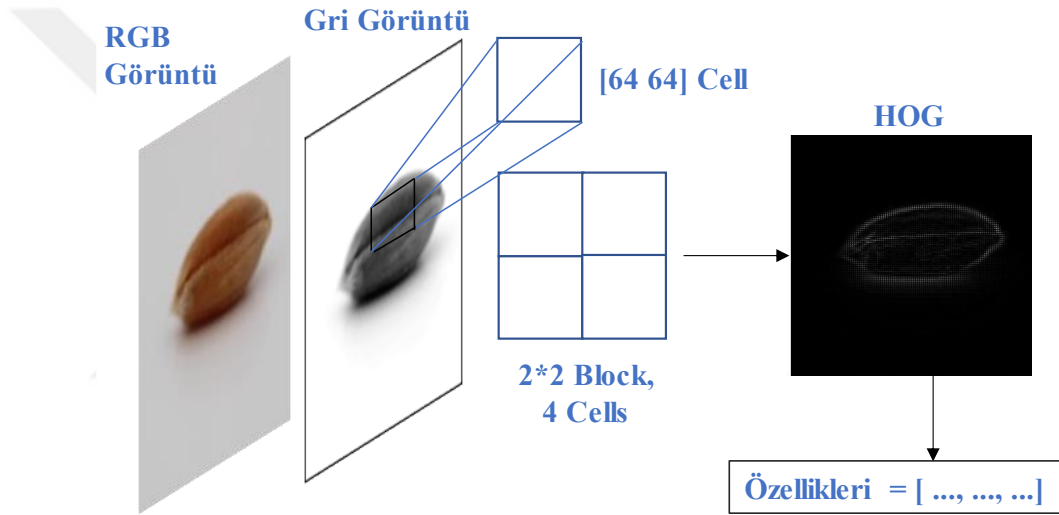


Şekil 4.1. Bir buğday tanesinin bir dakikada çekilen çoklu görünümüleri (75 adet)

4.2. Özellik Çıkarma

4.2.1. HOG özellikleri

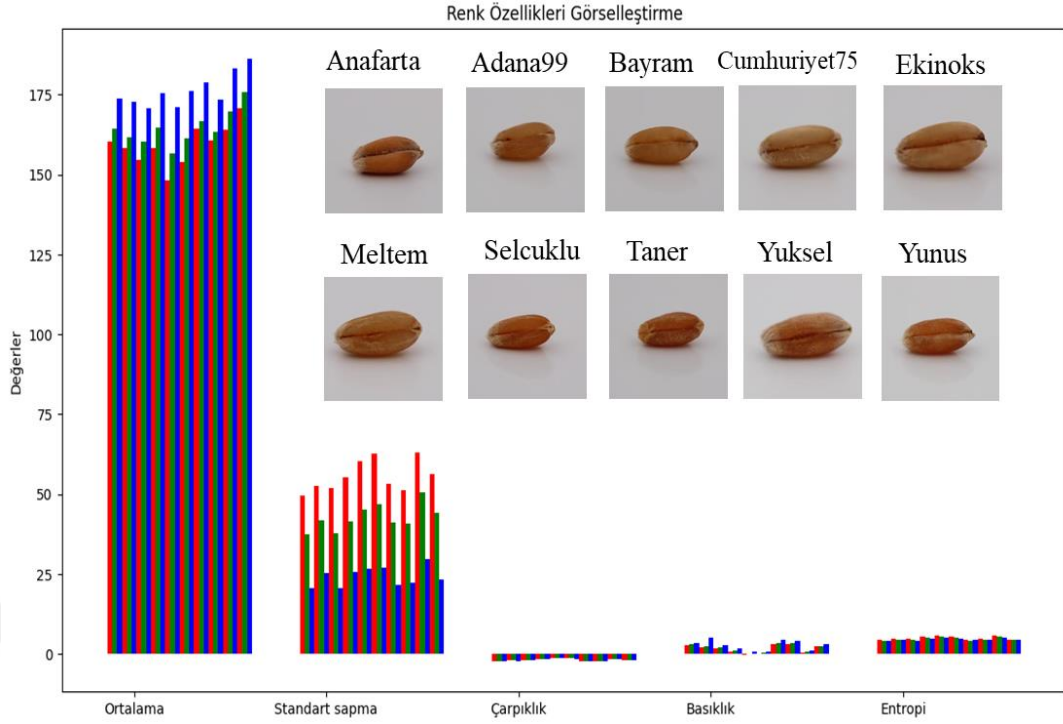
HOG Özellikleri, MATLAB HOG çıkarıcı kullanılarak çıkarılmıştır. İşlem, hücre boyutunun atanmasını içemektedir. Bu tezde, [64x64] boyutundaki hücre, yönlendirilmiş gradyanları hesaplamak için bir birim olarak atanmıştır. Böylece, blok örtüşmesi hariç, 250x250 boyutundaki bir görüntü için 4 hücre içeren bir blok kullanılmıştır. Gradyanlar, yönelimlerine göre bir histograma bölünür. Bu nedenle, girilen RGB görüntüsü önce gri tonlamaya dönüştürülmüş ve sonuçta 144 HOG özelliği çıkarılmış ve tüm buğday çeşitleri için özellik vektörleri $[H_1, H_2, H_3, \dots, H_{144}]$ olarak kaydedilmiştir (Şekil 4.2).



Şekil 4.2. Buğday tanelerinin HOG özelliklerinin çıkarılmasının görselleştirilmesi

4.2.2. Renk momentleri

RGB renk uzayı kullanılarak, ortalama-ortalama piksel yoğunluğu, standart sapma-piksel yoğunluklarının yayılımının ölçüsü, çarpıklık-piksel yoğunluğu dağılımının asimetrisinin ölçüsü, basıklık-piksel yoğunluğu dağılımının "kuyrukluluğunun" ölçüsü ve entropi-her RGB kanalı için piksel yoğunluklarındaki rastgeleliğin ölçüsü olmak üzere on beş renk momenti özelliği hesaplanmıştır. 124 buğday çeşidinin her bir görüntüsü için bu özellikler hesaplandı. Renk momentlerinin buğday çeşitleri arasında nasıl farklılık gösterdiğini görselleştirmek amacıyla, rastgele seçilen on çeşit için beş renk momentinin tümü Şekil 4.3'te sunulmuştur. Diyagramda da görüldüğü üzere, farklı sınıflandırma modellerini eğitmek için girdi parametreleri olarak kullanılacak RGB kanalları arasında renk özellikleri açısından kayda değer ve önemli bir fark bulunmaktadır.



Şekil 4.3. Buğday çeşitlerinin renk momentlerindeki varyasyonun görselleştirilmesi

4.3. Makine Öğrenme Modelleri

4.3.1. Yapay sinir ağları

Süreç, yapılandırılmış bir süreci benimseyen on ANN modelinin eğitilmesini içermektedir. Veri kümesi bir excel dosyasından yüklenmiş ve özellikler ile hedef etiketlere ayrılarak ön işleme tabi tutulmuştur. Hedef değişken sklearn'den LabelEncoder kullanılarak kodlanmıştır. Özellikler, tüm girdi özellikleri arasında tutarlı ölçeklendirme sağlamak için StandardScaler kullanılarak standartlaştırılmıştır. Eğitimde, PCA olmadan ve PCA ile eğitim olmak üzere iki strateji benimsenmiş ve diğer tüm eğitim süreci tüm modeller için aynı tutulmuştur. İkinci yaklaşımda PCA uygulanmış ve verilerin ilişkisiz bileşenlere dönüştürülmesinin avantajını elde etmek için tüm ana bileşenler giriş parametresi olarak kullanılmıştır. Her iki yaklaşımda da veriler, eğitim, doğrulama ve test setlerine bölünmüş, %76'sı eğitime ayrılmış ve kalan %24'ü doğrulama ve test setleri arasında eşit olarak paylaştırılmıştır.

Model, Keras'ın Sıralı API'si (Keras' Sequential API) kullanılarak, relu aktivasyonu ile başlayarak, iki gizli katman ve hedef sınıfların sayısına karşılık gelen softmax aktivasyonlu bir çıkış katmanı ile tasarlanmıştır. Model, Adam optimizier, 0.001 öğrenme oranı ve sparse_categorical_crossentropy kayıp fonksiyonu ile derlenmiş ve performans ölçütü olarak doğruluk kullanılmıştır. Aşırı uyum sorununu

ele almak için iki geri çağırma kullanılmıştır. Bunlardan birincisi, 50 epok boyunca doğrulama kaybında iyileşme olmadığında eğitimi durdurmak için earlyStopping, ikincisi ise doğrulama kaybına dayalı en iyi modeli kaydetmek için ModelCheckpoint'tir. Eğitim, modelin performansını izlemek ve aşırı uyumu önlemek için doğrulama seti kullanılarak 32'lik bir batch boyutuyla 200 epok üzerinden gerçekleştirilmiştir. Eğitimin hiperparametreleri Tablo 4.1'de verilmiştir.

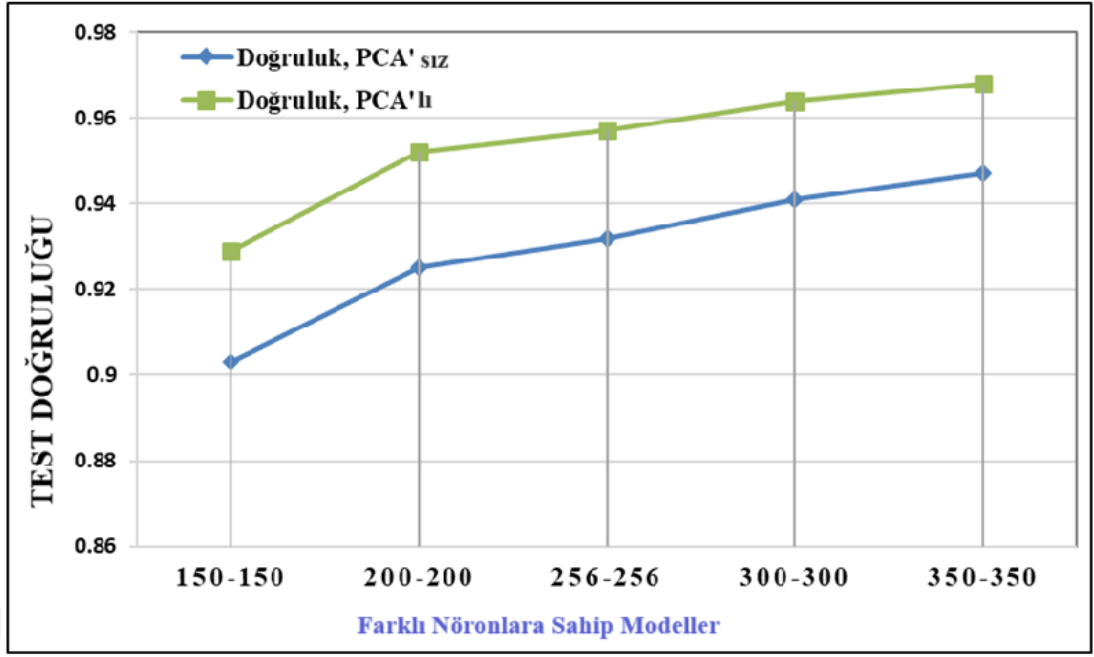
Tablo 4.1. ANN modellerinin hiperparametreleri

Parametre	Model-1	Model-2	Model-3	Model-4	Model-5
Nöron	150-150	200-200	256-256	300-300	350-350
Parti (batch) büyüklüğü	32	32	32	32	32
Epoch	200	200	200	200	200
Patience	50	50	50	50	50
Öğrenme Oranı	0.001	0.001	0.001	0.001	0.00

Sonuç olarak, iki yaklaşım ve gizli katmanlarında farklı sayıda nöron bulunan beş model benimsenerek on ANN modeli eğitilmiştir. Modellerin performans metriklerinin özeti Tablo 4.2'de sunulmuş olup, iki gizli katmanda farklı nöron boyutlarına sahip ANN modellerinin test doğruluğunu ve yalnızca özellikler ile HOG ve renk momenti özelliklerinin PCA'sı olmak üzere iki girdi durumunu dikkate alarak gösterilmiştir.

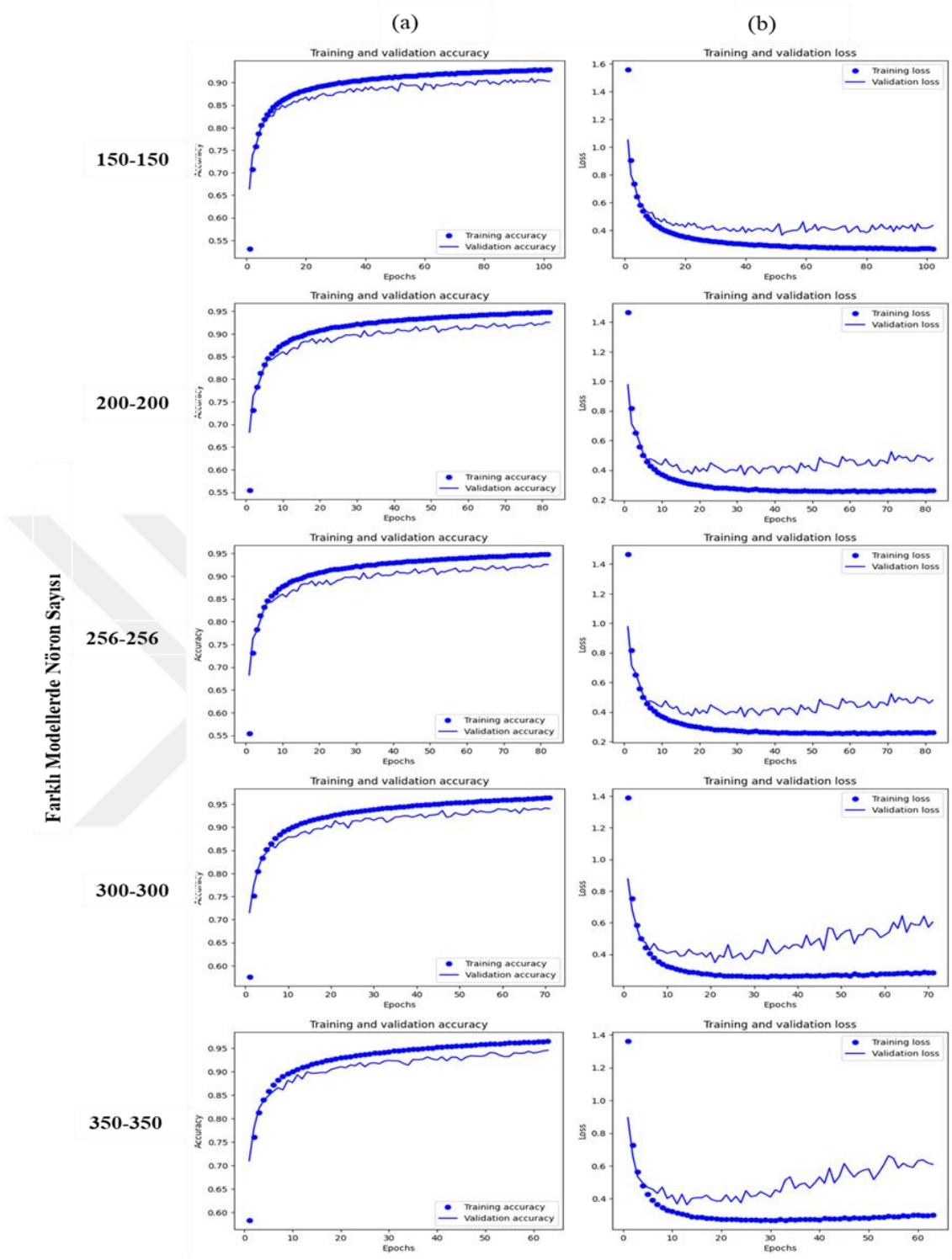
Tablo 4.2. On ANN modelinin performans metrikleri

Veri girişi	İki Gizli katmandaki nöronlar	Doğruluk	Kesinlik	Duyarlılık	F1-Skor
Sadece Özellikler	150-150	90.3	91.0	90.0	90.0
PCA'lı	150-150	92.9	93.0	93.0	93.0
Sadece Özellikler	200-200	92.5	93.0	93.0	93.0
PCA'lı	200-200	95.2	95.0	95.0	95.0
Sadece Özellikler	256-256	93.2	93.0	93.0	93.0
PCA'lı	256-256	95.7	96.0	96.0	96.0
Sadece Özellikler	300-300	94.1	94.0	94.0	94.0
PCA'lı	300-300	96.4	97.0	96.0	96.0
Sadece Özellikler	350-350	94.7	95.0	95.0	95.0
PCA'lı	350-350	96.8	97.0	97.0	97.0

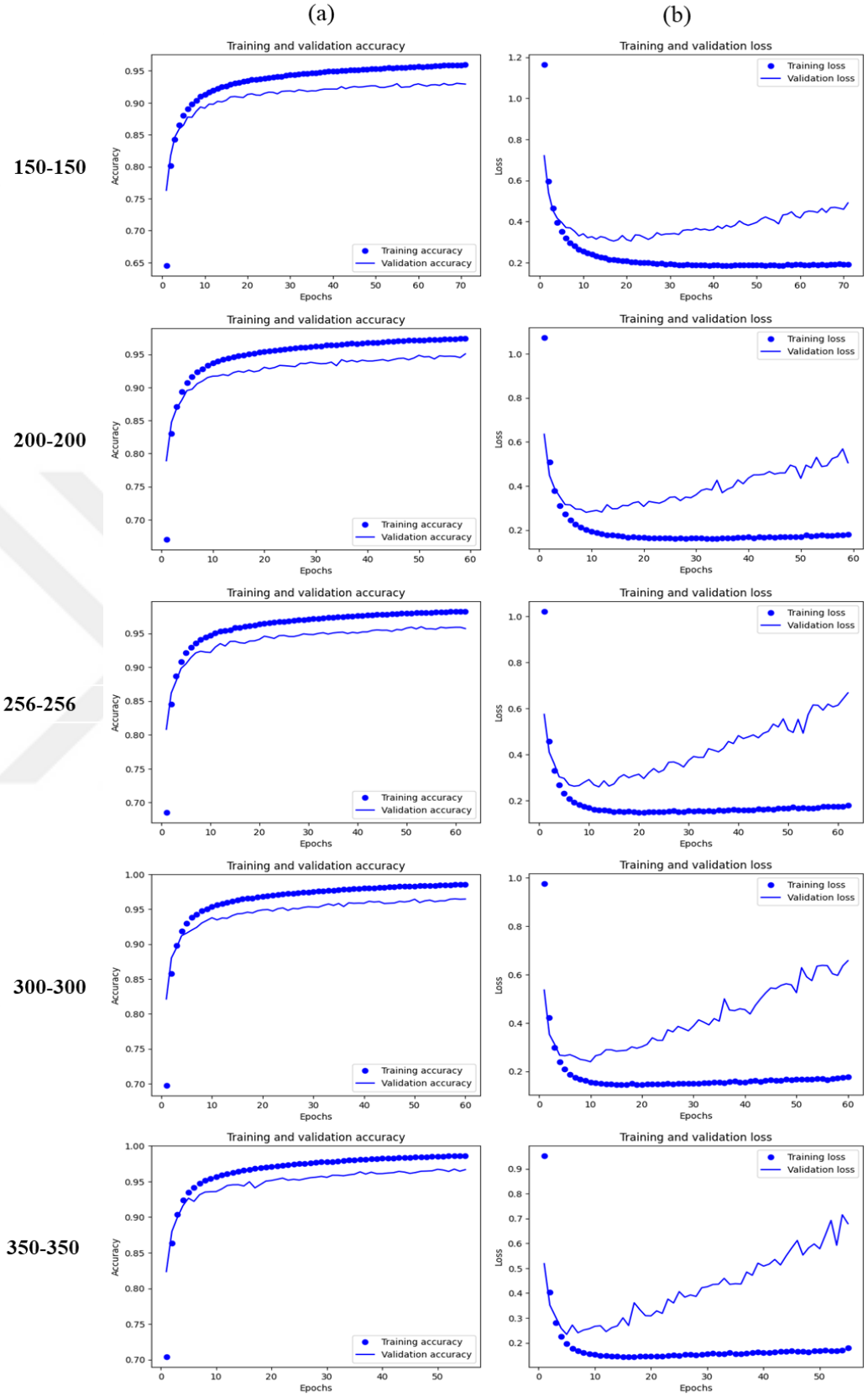


Şekil 4.4. ANN modelinin test doğruluğu

Şekil 4.4'te belirtildiği gibi, gizli katmanlardaki nöron sayısı arttıkça 10 modelin test doğrulukları da artmıştır. PCA olmadan sadece özelliklerin takip edildiği yaklaşımda, en düşük ve en yüksek doğruluk oranları sırasıyla 150-150 ve 350-350 olan modeller için %90.3 ve %94.7 olarak elde edilmiştir. Nöron sayısındaki artış, doğruluk oranında %4.87'lik bir yüzdelik artış göstermiştir. Benzer şekilde, PCA ile eğitilen modellerin test doğruluklarını incelediğimizde, en düşük ve en yüksek doğruluk oranları, 150-150 ve 350-350 nörona sahip olan modellerde sırasıyla %92.9 ve %96.8 olup, doğruluk oranındaki yüzdelik artış %4.2 elde edilmiştir. Nöron sayısındaki artış, test doğruluklarının istikrarlı bir şekilde artmasıyla sonuçlanmıştır.



Şekil 4.5. (a) PCA'sız eğitilen gizli katmanlarında farklı nöron boyutlarına sahip beş modelin eğitim ve doğrulama doğruluk eğrileri, (b) Eğitim ve doğrulama kayıp eğrileri



Şekil 4.6. (a) PCA' ile eğitilen farklı nöron boyutlarına sahip beş modelin eğitim ve doğrulama doğruluk eğrileri, (b) Eğitim ve doğrulama kayıp eğrileri.

Epok sayısına karşı verilen eğitim ve doğrulama doğruluk eğrileri ile eğitim ve kayıp eğrilerini incelemek, epok sayısının artışıyla modellerin aşırı uyum özelliklerini değerlendirmek için önemlidir (Şekil 4.5-4.6). Erken durdurma ve model kontrol noktası stratejilerine sahip olan eğitim, doğrulama kaybı iyileşmeyi durdurduğunda yani doğrulama kaybı iyileşmediğinde durdurulacak şekilde ayarlanmıştır. Böylece, PCA'sız eğitimde, eğitimin durduğu noktadaki en yüksek ve en düşük epok sayısı 150-150 ve 350-350 yapılandırılmalı modeller için sırasıyla 100 ve 65 olarak elde edilmiştir. PCA tabanlı aşamada ise en yüksek ve en düşük eğitim epok sayısı, aynı modeller için 75 ve 55 olmuştur. Bu sonuçlar, PCA uygulanmasının avantajlı olduğunu, epok sayısındaki azalmayla kanıtlandığı üzere modellerin daha hızlı yakınsamasına yardımcı olduğunu göstermektedir. Ayrıca, çoklu bağlantıyı ve gürültüyü azaltma potansiyeli, nispeten daha yüksek hesaplama verimliliği ile test doğruluklarının artmasına neden olmuştur.

Ancak, kayıp eğrilerini yakından incelediğimizde, eğitim ve doğrulama eğrileri arasındaki farkın nöron sayılarının artışıyla önemli ölçüde arttığı ve nöron sayılarının daha da artmasının aşırı uyum riskine yol açabileceği açıkça görülmektedir. Bu nedenle, 350-350 nöron sayısının ötesine geçmenin bu noktada uygun olmayacağı anlaşılmıştır.

4.3.2. Destek vektör makineleri

SVM algoritmasının eğitim süreci, hem PCA ile hem de PCA olmadan model seçimi ve hiperparametre ayarına yönelik kapsamlı bir yaklaşım içermektedir. İş akışı, özellik ölçeklendirme ve grid arama entegrasyonu ile eğitim, en iyi performans gösteren modeli elde etmek için en uygun yapılandırmayı belirlemeyi amaçlamıştır. ANN'ların eğitiminde olduğu gibi, tutarlı ölçeklendirme sağlamak için özellikler, StandardScaler kullanılarak standartlaştırılmıştır. Veriler daha sonra eğitim, doğrulama ve test setlerine ayrılmıştır. SVM modeli için boyutsallık azaltma amacıyla, biri PCA içeren, diğeri PCA içermeyen iki işlem hattı oluşturulmuştur. Her bir işlem hattı, özelliklerin ölçeklendirilmesi ve bir SVM sınıflandırıcısının uygulanmasından oluşmuştur. En iyi hiperparametreleri belirlemek için her iki iş akışı için 5 katlı çapraz doğrulama ile grid aramalar gerçekleştirilmiştir. Hiperparametrelerin listesi, çekirdek tipi (doğrusal ve rbf), düzenlilik (c: 0.1, 1.0, 10) ve temel bileşenlerin sayısını (10, 20, 100) içeriyordu. PCA içeren iş akışı için temel bileşen sayısı, düzenleme parametresi ve çekirdek türü ayarlanırken, PCA içermeyen iş akışı için yalnızca düzenleme

parametresi ve çekirdek türü ayarlanmıştır. Her iki grid araması da model performansının sağlam bir şekilde doğrulanmasını sağlamak için 5 katlı çapraz doğrulama kullanılmıştır. Eğitimin sonuçları Tablo 4.3'te verilmiştir. PCA içeren SVM, tüm metriklerde PCA içermeyen SVM'den biraz daha yüksek bir performans elde ederek sırasıyla %98.1 ve %97.02 sınıflandırma doğrulukları sağlamıştır.

Tablo 4.3. SVM modellerinin performans metrikleri

PCA'lı SVM			
Doğruluk	Kesinlik	Duyarlılık	F1-skor
98.1	98.0	98.0	98.0
En İyi Parametreler: PCs: 100, SVM_C: 10, çekirdek: rbf			
PCA'sız SVM			
Doğruluk	Kesinlik	Duyarlılık	F1-skor
97.02	97.0	97.0	97.0
En İyi Parametreler: SVM_C: 10, çekirdek: rbf			

4.3.3. Random forest sınıflandırıcısı

HOG ve renk momenti özellikleri kullanılarak bir RFC modeli de değerlendirilmiştir. Boyut azaltmanın modelin performansı üzerindeki etkisini değerlendirmek için model hem PCA ile hem de PCA olmadan eğitilmiştir. Aynı veri bölme işlemi takip edilmiş ve modelin optimum hiperparametrelerini belirlemek için 5 katlı çapraz doğrulama ile Grid arama yöntemi kullanılarak ince ayar yapılmıştır. İlgili hiperparametreler Tablo 4.4'te gösterildiği gibi araştırılmıştır.

Tablo 4.4. RFC modellerinin hiperparametreleri

RFC Modelinin Parametreleri	
PCs	{10,30, 50, 70, 124}
Estimators	{50, 100, 150, 200}
Max Depth	{10,30}
Min Samples to split	{20,25}
Min Samples leaf	{5,10}
Bootstrap	{True, False}

PCA ile en iyi performans gösteren model, şu hiperparametreler kullanılarak elde edilmiştir; 50 temel bileşen (principal component), 30 maksimum ağaç derinliği

(max tree depth) , her yaprak için minimum 10 örnek (Min samples leaf), 200 estimators ve Bootstrap olmaksızın. Bu model, eğitim verileri üzerinde %99.68 doğruluk elde ederek, modelin eğitim verilerine oldukça iyi uyum sağladığını göstermektedir. Ancak, doğrulama setinde değerlendirildiğinde doğruluk %82.58'e düştü ve test verileri üzerinde benzer bir şekilde %82.77 doğrulukla performans gösterdi. Eğitim doğruluğu ile doğrulama ve test doğruluğu arasındaki nispeten yüksek fark, modelin eğitim verilerine aşırı uyum sağlamış (overfitting) olabileceğini düşündürmektedir.

Aynı şekilde, model PCA kullanılmadan, orijinal 159 özellik kullanılarak eğitilmiştir. Bulunan en uygun hiperparametreler, PCA tabanlı modelinkilerle aynı olup, 30 maksimum ağaç derinliği (max tree depth), her yaprak için minimum 10 örnek (Min samples leaf), 200 estimators ve Bootstrap olmaksızındır. Bu konfigürasyonda, model %99.78 ile PCA tabanlı modelden biraz daha yüksek bir eğitim doğruluğu sergilemiştir. Ayrıca, doğrulama ve test doğrulukları sırasıyla %85.40 ve %85.30'a yükselerek, PCA'sız modelin görülmemiş verilere daha iyi genelleme yapabildiğini, ancak aşırı uyum (overfitting) potansiyelinin hâlâ çözülemediğini göstermektedir. Genel olarak, bu durum, bu veri seti için boyut azaltma yapılmadan tam özellik setinin korunmasının daha iyi tahmin performansı sağladığını göstermektedir. Tablo 4.5, RFC modelinin performans metriklerini göstermektedir; PCA olmadan kullanılan model, tüm metriklerde PCA kullanılan modelden daha iyi performans göstermiştir.

Tablo 4.5. RFC modellerinin performans metrikleri

PCA'lı RFC			
Doğruluk	Kesinlik	Duyarlılık	F1-skor
82.7	84.0	83.0	83.0
PCA'sız RFC			
Doğruluk	Kesinlik	Duyarlılık	F1-skor
85.3	86.0	85.0	85.0

4.3.4. K-en yakın komşu algoritması

Benzer şekilde, bir KNN modeli de her iki eğitim hattı (pipeline) üzerinden eğitilmiştir. Hem PCA tabanlı hem de PCA kullanılmayan modeller için en uygun hiperparametreleri belirlemek amacıyla 5 katlı çapraz doğrulama kullanılarak Grid

arama gerçekleştirildi. İnce ayar yapmak için Tablo 4.6'da verilen hiperparametreler test edildi.

Tablo 4.6. KNN modellerinin hiperparametreleri

KNN Modelinin Parametreleri	
PCs	{50, 100}
KNN metrics	{Euclidean, manhattan}
KNN weights	{uniform, distance}

PCA içeren en iyi KNN modeli, 100 temel bileşen (principal component), Manhattan mesafe metriği (distance metric) ve mesafeye dayalı ağırlıklandırma (distance-based weighting) ile 124 komşu büyüklüğünde oluşturulmuştur. Model, eğitim verileri üzerinde %100 doğruluk sağlayarak veriye tam uyum göstermiştir. Ancak, bu yüksek eğitim doğruluğu, doğrulama ve test setlerinde iyi bir performans anlamına gelmemiştir. Model, sırasıyla doğrulama ve test setlerinde %65.74 ve %65.81 doğruluk elde etmiştir. Performanslardaki bu tutarsızlık, modelin eğitim verileri için iyi olduğunu, ancak görülmemiş verilere genelleme yapmadığını ve potansiyel olarak aşırı uyum gösterdiğini (overfitting) ortaya koymaktadır.

PCA kullanılmayan durumda, aynı hiperparametreler en iyi parametreler olarak bulunmuştur. Model yine %100 eğitim doğruluğuna ulaşmıştır. Ancak, doğrulama ve test doğrulukları sırasıyla %63.33 ve %63.21'e hafifçe düşmüştür. Aynı hiperparametre kümesine rağmen, PCA kullanılmayan model, PCA tabanlı modele kıyasla daha düşük bir genelleme yeteneği sergilemiştir ve her iki model de zayıf genelleme göstermiştir. Tablo 4.7, KNN'nin performans metriklerini göstermektedir. PCA kullanılan model, PCA kullanılmayan modele göre daha iyi performans göstermiştir. Bu sonuçlar, PCA uygulamanın modelin genelleme yeteneğini artırmaya yardımcı olduğunu, ancak KNN'nin genel performansının düşük olduğunu ve veriyi başarıyla sınıflandırmada potansiyel zorluklar yaşadığını göstermektedir.

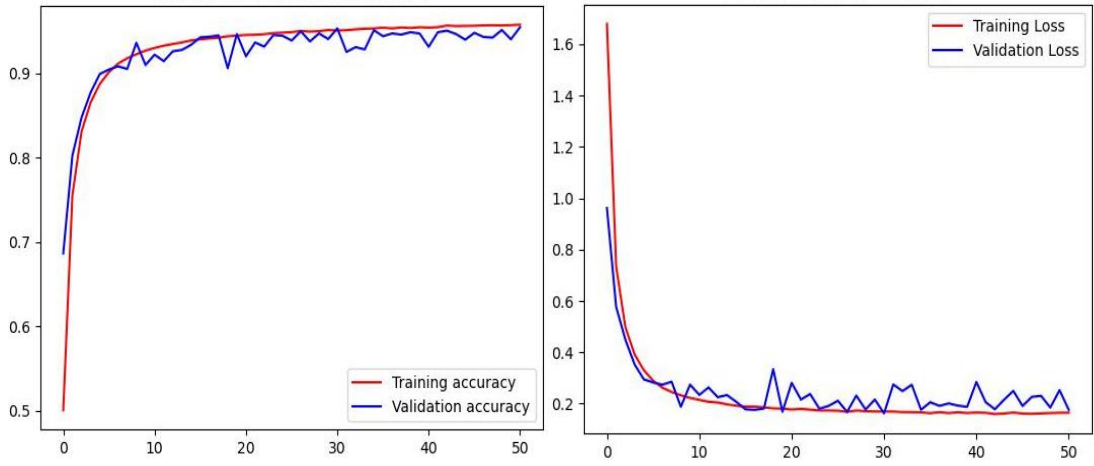
Tablo 4.7. KNN modellerinin performans metrikleri

PCA'lı KNN			
Doğruluk	Kesinlik	Duyarlılık	F1-skor
65.8	69.0	66.0	66.0
PCA'sız KNN			
Doğruluk	Kesinlik	Duyarlılık	F1-skor
63.2	66.0	63.0	63.0

4.4. Derin Öğrenme Modelleri

4.4.1. CNN modeli

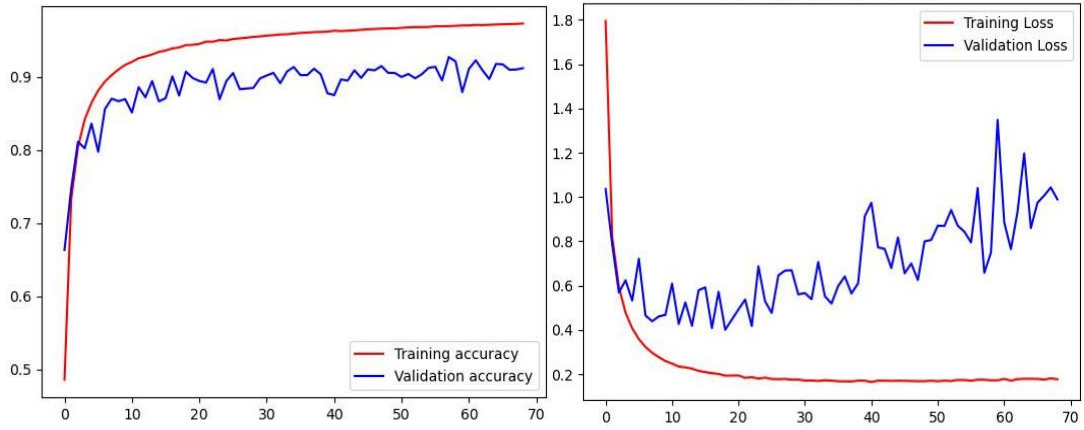
Dört evrişimli ve dört maksimum havuzlama katmanına sahip önerilen sığ (shallow) CNN modeli ve %5 dropout ile eğitim, epoch sayısını 100, batch boyutunu 32, model kontrol noktası olarak 20 patience değeri ile 'adam' optimizasyonunu kullanarak gerçekleştirilmiştir. Toplam 465 bin görüntü, %80 eğitim, %10 doğrulama ve %10 test seti olarak bölünmüştür. Eğitim ve doğrulama doğruluk eğrilerinde görüldüğü gibi, her iki doğruluk da ilk 10 epoch değerinde belirgin bir şekilde arttı ve eğitim devam ettikçe stabilize olmaya başladı (Şekil 4.7). Eğitim ve doğrulama doğruluklarının çok yakın olması, aşırı uyum riskinin çok düşük olduğunu gösterirken, doğrulama eğrisinin dalgalanmasına rağmen, eğitim eğrisine yakınlığı modelin görülmemiş doğrulama seti ile iyi performans gösterdiğini işaret etmektedir. Eğitim, epoch 51'de durdu ve en iyi modelin epoch 31'de %94.03 doğrulukla kaydedildiği, doğrulama doğruluğunun ardışık 20 epoch boyunca iyileşmediği gösterilmektedir. Bu nedenle, Tablo 4.8'de verilen performans metrikleri açısından doğruluk, kesinlik, duyarlılık ve F1-skoru sırasıyla %95.40, %95.44, %95.39 ve %95.36 olarak belirlenmiştir. Metrik değerleri, modelin sınıflandırma performansının oldukça yüksek olduğunu göstermektedir. Doğruluk oranının %95.40 gibi yüksek bir seviyede olması, modelin genel olarak doğru sınıflandırmalar yapma kapasitesini ortaya koymaktadır. Bununla birlikte, kesinlik ve duyarlılık değerlerinin birbirine yakın olması, modelin dengeli bir performans sergilediğini ve hem yanlış pozitif hem de yanlış negatif oranlarının düşük olduğunu ortaya koymaktadır. F1-skorunun da bu değerlere paralel olarak yüksek olması, modelin genel başarımının güçlü olduğunu ve sınıflandırma görevlerinde güvenilir sonuçlar ürettiğini kanıtlamaktadır.



Şekil 4.7. Önerilen CNN modelinin eğitim ve doğrulama setleri için doğruluk ve kayıp eğrileri.

4.4.2. MobileNet

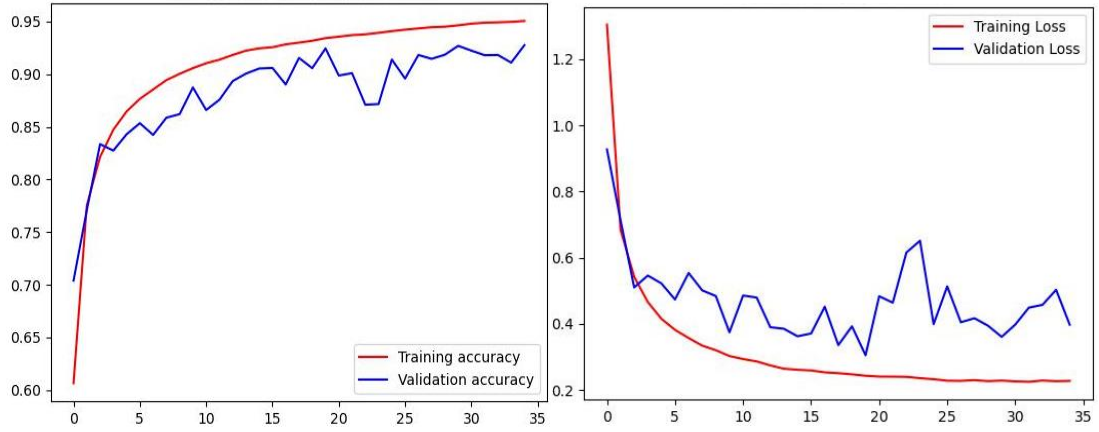
MobileNet CNN mimarisi kullanılarak öğrenme aktarımı yoluyla yalnızca tam bağlantılı katmanlar değiştirilmiş ve her biri 256 ve 128 nöron içeren iki gizli katman eklenmiştir. Önerilen CNN modelin de olduğu gibi, tüm eğitim parametreleri aynı tutulmuştur. Şekil 4.8'de görülen öğrenme eğrisi, eğitim doğruluğunun özellikle 10. epoch'tan sonra istikrarlı bir şekilde arttığını göstermektedir. Küçük dalgalanmalarla birlikte benzer bir desen doğrulama eğrisinde de gözlemlenmiştir. Eğitim, epoch 69'da durmuş ve epoch 39'da en iyi model olarak kabul edilmiştir. Bu modelin eğitim ve doğrulama doğrulukları sırasıyla %96.62 ve %91.49'dur. Genel test doğruluğu ise %90.54 olarak belirlenmiştir. Eğitim ve doğrulama doğrulukları arasındaki hafif yüksek fark, modelde az bir aşırı uyum (overfitting) olduğunu göstermektedir. Ayrıca, performans metrikleri doğruluk, kesinlik, duyarlılık ve F1-skoru sırasıyla %90.54, %91.02, %90.39 ve %90.44 olarak bulunmuştur (Tablo 4.8). Bu değerler, modelin genel performansının oldukça iyi olduğunu göstermektedir. Doğruluk oranının %90.54 seviyesinde olması, modelin doğru sınıflandırmalar yapma konusunda güçlü bir performans sergilediğini ortaya koymaktadır. Kesinlik ve duyarlılık değerlerinin birbirine yakın olması, modelin yanlış pozitif ve yanlış negatif oranlarını dengelediğini ve böylece her iki hata türünü de minimize edebildiğini işaret etmektedir. F1-skorunun %90.44 ile bu değerlerle uyumlu olması, modelin genel başarımının tutarlı ve güvenilir olduğunu kanıtlamaktadır. Bu sonuçlar, modelin sınıflandırma görevlerinde etkin bir şekilde kullanılabileceğini göstermektedir.



Şekil 4.8. MobileNet modelinin eğitim ve doğrulama setleri için doğruluk ve kayıp eğrileri.

4.4.3. DenseNet201

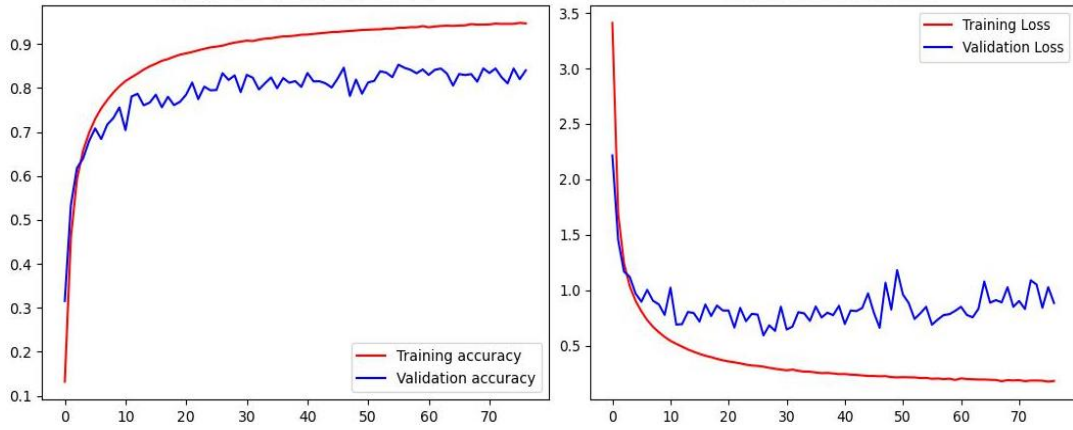
Benzer bir stratejide, popüler CNN mimarilerinden DenseNet201'de de benimsenmiştir. Model, 35 epoch boyunca eğitilmiş ve en iyi model epoch 15'te kaydedilmiştir. Öğrenme eğrisi, eğitim ve doğrulama doğruluklarının yaklaşık 20. epoch'a kadar istikrarlı bir şekilde arttığını göstermektedir (Şekil 4.9). Sonrasında ise eğitim ve doğrulama kaybında bir artış veya doğrulama doğruluğunda bir düşüş yaşamıştır ki bu durum aşırı uyum (overfitting) riskinden kaynaklanabilir. Ancak, model geri çağırma işlevi, modelin bu noktadan önce kaydedilmesini sağlamış ve doğrulama doğruluğu eğrisi istikrarlı kalmıştır. Eğitim ve doğrulama doğrulukları arasındaki fark, MobileNet'tekinden nispeten daha küçük olmuştur. Performans metrikleri doğruluk, kesinlik, duyarlılık ve F1-skoru sırasıyla %92.41, %92.88, %92.26 ve %92.32 olarak bulunmuştur (Tablo 4.8). Bu sonuçlar, modelin sınıflandırma performansının yüksek ve dengeli olduğunu göstermektedir. Doğruluk oranının %92.41 olması, modelin genel olarak doğru sınıflandırmalar yapma kapasitesini vurgulamaktadır. Kesinlik ve duyarlılık metriklerinin birbirine oldukça yakın olması, modelin yanlış pozitif ve yanlış negatif oranlarını etkili bir şekilde dengelediğini ortaya koymaktadır. Ayrıca, F1-skorunun %92.32 olarak belirlenmesi, modelin genel başarımının güçlü olduğunu ve sınıflandırma görevlerinde tutarlı sonuçlar üretebildiğini kanıtlamaktadır.



Şekil 4.9. DenseNet201 modelinin eğitim ve doğrulama setleri için doğruluk ve kayıp eğrileri.

4.4.4. InceptionV3

InceptionV3 modeli de DenseNet201 ve MobileNet'i eğitmek için kullanılan aynı yöntemle değerlendirilmiştir. Model, 20 epoch'luk patience kriterine dayalı erken durdurma ile epoch 77'ye kadar eğitilmiştir. Şekil 4.10, epoch'lar boyunca eğitim ve doğrulama doğruluğunun eğitim sürecini göstermektedir. Model, başlangıç epoch'larında hızlı bir öğrenme sergilemiş, eğitim doğruluğu %90 civarında plato yapmadan önce istikrarlı bir şekilde artmıştır. Doğrulama doğruluğu ise biraz dalgalı olmakla birlikte kademeli bir artış göstermiş ve yaklaşık 40 epoch'tan sonra %83 civarında stabilize olmuştur. Model, %83.37'lik bir son doğrulama doğruluğu ve buna karşılık gelen %83.47'lik bir test doğruluğu elde etmiştir. Test ve doğrulama doğruluklarının birbirine yakın değerlerde olması, modelin genelleme yeteneğini göstermektedir. Performans metrikleri sırasıyla %83.47, %84.11, %83.24 ve %83.25 olarak doğruluk, kesinlik, duyarlılık ve F1-skoru için belirlenmiştir (Tablo 4.8). Bu sonuçlar, modelin sınıflandırma performansının makul düzeyde olduğunu göstermektedir. Doğruluk oranının %83.47 olması, modelin genel olarak doğru sınıflandırmalar yapma kapasitesinin iyi olduğunu ancak daha yüksek performans seviyelerine ulaşabileceğini işaret etmektedir. Kesinlik ve duyarlılık değerlerinin birbirine yakın olması, modelin yanlış pozitif ve yanlış negatif oranlarını dengelemekte başarılı olduğunu ortaya koymaktadır. F1-skorunun %83.25 ile bu değerlere uyumlu olması, modelin genel performansının tutarlı ve dengeli olduğunu göstermektedir. Bununla birlikte, bu metriklerin daha yüksek değerlere ulaşması için modelin iyileştirilmesi gerektiği düşünülebilir.



Şekil 4.10. InceptionV3 modelinin eğitim ve doğrulama setleri için doğruluk ve kayıp eğrileri.

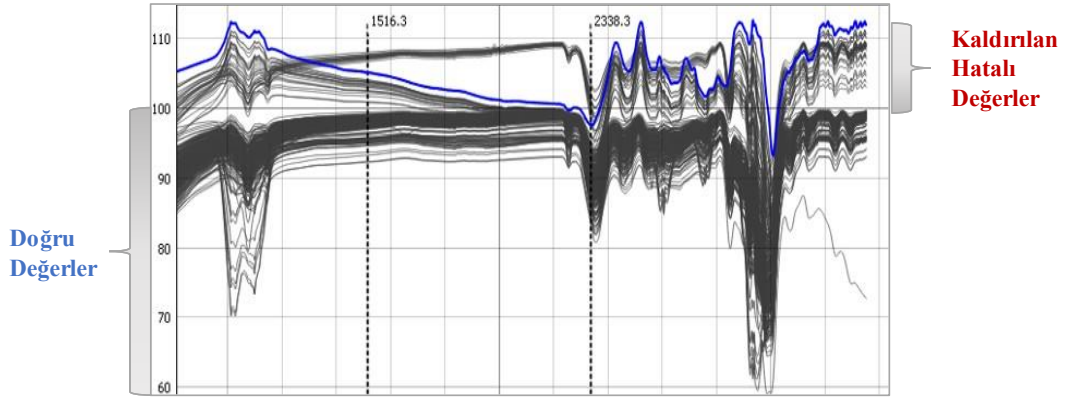
Tablo 4.8. Önceden eğitilmiş ve önerilen CNN modellerinin performansları

Modeller	Eğitilebilir Parametreler	Metrikler (%)			
		Doğruluk	Kesinlik	Duyarlılık	F1-skoru
Önerilen CNN	5.876.540	95.40	95.44	95.39	95.36
MobileNet	12.894.204	90.54	91.02	90.39	90.44
DenseNet201	24.133.628	92.41	92.88	92.26	92.32
InceptionV3	18.923.516	83.47	84.11	83.24	83.25

4.5. FTIR Tabanlı Makine Öğrenme Modelleri

4.5.1. Veri hazırlama ve temizleme

Çıkarılan özelliklerin girdi parametreleri olarak kullanıldığı görüntü tabanlı çeşit sınıflandırmasının yanı sıra, bu çalışma, girdi parametreleri olarak buğday spektral verilerini kullanmayı, dört buğday çeşidi için sınıflandırma modelleri geliştirmeyi ve değerlendirmeyi amaçlamıştır. FTIR spektrometresi kullanılarak ölçülen dört buğday çeşidinin spektral veri seti orijinal olarak 320 spektral veriden oluşmaktadır. Her spektrum, 650 - 4000 dalga sayısı cm^{-1} arasında uzanan 3.351 sütun içeriyordu. FTIR spektrumları yüzde geçirgenlik cinsinden ölçülmüş, Şekil 4.11'de gösterildiği gibi 100'den büyük değerlere sahip veri noktaları hatalı okuma olarak değerlendirilerek çıkarılmıştır. Verilerin ayrıntıları Tablo 4.9'da sunulmuştur.

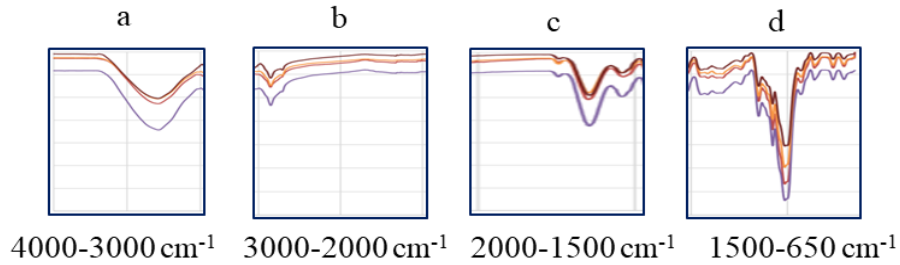


Şekil 4.11. Buğday çeşitlerinin orijinal FTIR spektral okuması.

Tablo 4.9. Dört buğday çeşidinin FTIR spektral veri setleri

Çeşitler	Toplanan Örnekler	Temizlik Sonrası Veriler
Altındane	80	60
Çavuş	80	80
Flamura-85	80	65
Nevzatbey	80	80
Toplam	320	285

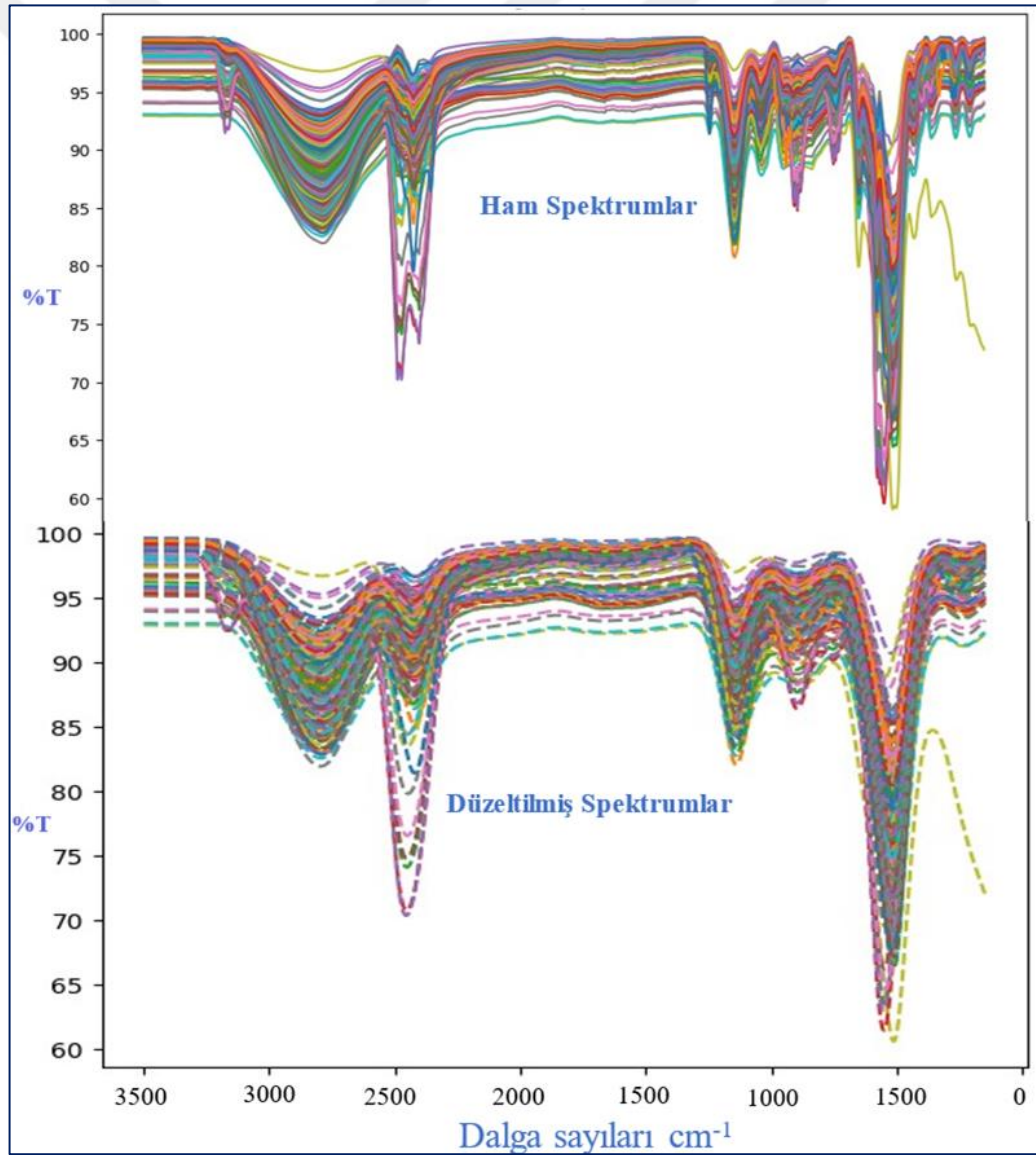
Girdi özellikleri açısından toplam 285 veri kümesine sahip olan, 3.351 nokta içeren tam spektrumların kullanılması yaygın olarak uygulanmaktadır. Ancak bu durumda, tam spektrum Şekil 4.12'de görüldüğü gibi dört bölüme (a, b, c ve d) ayrılmış ve tüm bölümlerin olası tüm kombinasyonları yapılarak yöntemler bölümünde gösterildiği gibi 15 veri seti hazırlanmış ve hem tek bölüm hem de bunların kombinasyonları kullanılarak etkiyi değerlendirmek için ML model eğitiminde kullanılmıştır.



Şekil 4.12. Tam spektrumların dört bölümü

4.5.2. Baz çizgisi düzeltmesi

Temel bir ön işleme adımı olarak, veri seti üzerinde asimetrik en küçük kareler yumuşatma uygulaması kullanılarak baz çizgisi düzeltmesi yapılmıştır. Baz çizgisi düzeltmesi, verileri temsil eden düzgün bir eğri bulmak ve düzgünleştirilmiş bir eğri üretmek için bu eğriyi orijinal eğriden çıkarmaktır. Düzeltme işlemi, düzgün bir taban çizgisi elde etmek için ağırlıkların iteratif olarak ayarlanmasını ve bir doğrusal denklem sisteminin çözülmesini içermektedir. Süreç, asimetrik gürültüleri ele alıp düzeltilmiş ve daha ileri analizler için uygun temizlenmiş bir sinyal üretmiştir. Orijinal, düzeltilmiş ve pürüzsüzleştirilmiş spektral eğriler, Şekil 4.13'de gösterilmektedir. Düzeltme, grafiğin alt kısmında görüldüğü gibi daha düzgün ve gürültüsüz bir spektral gösterimle sonuçlanmıştır.



Şekil 4.13. Tüm örneklerin temel çizgi düzeltilmiş spektrumları

4.5.3. FTIR analizi verilerine göre makine öğrenme modellerinin eğitimi

Buğday spektral veri tabanı, ML modelleri eğitimi için ANN, SVM, RFC ve KNN olmak üzere dört algoritma ile eğitilmiştir. Eğitim süreci, baz çizgisi düzeltme, aşırı örnekleme ve normalizasyon ön işleme adımlarını içermiştir. PCA'lı ve PCA'sız iki iş akışı, grid arama yaklaşımı kullanılarak hiperparametrelerin ince ayarını yapmak amacıyla ilgili modeller için hiperparametrelerin bir listesiyle birlikte kullanılmıştır. Veri seti nispeten küçük olduğundan, veri seti 5 katlı çapraz doğrulama kullanılarak eğitim ve test setlerine ayrılmıştır. FTIR spektral verilerine dayalı tüm modellerin eğitim ve test doğrulukları belirlenmiştir.

Birçok modelin doğruluğu benzer ve çok yüksek olmasına rağmen, en iyi ve uygun modeli seçmek söz konusu olduğunda bazı parametreler dikkate alınması uygun olmaktadır. Benzer durumlarda PCA kullanılmayan uygulamanın tercih edilmesi, kullanılan spectral bölgeden dolayı daha az girdi özelliği kullanmanın hem eğitim hem de çıkarım sırasında hesaplama avantajı sağlaması, eğitim doğruluk oranının test sonuçlarından daha yüksek olması, aynı zamanda eğitim ve test doğruluk oranları arasındaki farkın düşük olması bu parametreler arasındadır. Ayrıca parmak izi bölgesi de (bu çalışmada "d" spektral bölgesi) model seçiminde dikkate alınması gereken parametrelerdendir.

4.5.3.1. FTIR tabanlı yapay sinir ağları modeli

ANN modelinin test doğrulukları genel olarak yüksek sonuçlar sergilemiştir. ANN modeli kullanılarak, PCA olmadan eğitilen "a" spectral bölgesinde (4000-3000 cm^{-1}) en düşük sınıflandırma doğruluğu (%93.8) elde edilmiştir. Hem tam spektrumda hem de çeşitli veri seti kombinasyonlarında PCA ile ve PCA olmadan en yüksek sınıflandırma doğruluğu %100 olmuştur (Tablo 4.10).

PCA ile yapılan eğitimde, "ab", "ac", "ad", "abd", "bd" ve "abcd-full spectra" bölümlerinde %100 doğruluk elde edilmiştir. PCA'sız eğitimde ise "abd", "bc" ve "abcd-full spectra" bölümlerindeki modeller %100 doğruluk sağlanmıştır.

Tablo 4.10. Farklı spektral bölgeler ve bunların kombinasyonları kullanılarak ANN modellerinin eğitim ve test doğruluk oranları (%)

Spektral Bölge	ANN			
	PCA'lı		PCA'sız	
	Eğitim	Test	Eğitim	Test
a	99.6	98.8	97.9	93.8
b	99.6	98.8	99.1	97.5
c	99.6	97.5	98.75	97.5
d	100	98.8	100	98.8
ab	100	100	99.2	97.5
ac	99.2	100	98.8	97.5
ad	99.6	100	99.2	98.8
abc	99.6	97.5	99.2	97.5
abd	99.6	100	98.3	100
acd	100	97.5	98.8	98.8
bc	100	98.8	97.5	100
bd	99.6	100	99.2	98.8
bcd	99.6	98.8	99.6	97.5
cd	100	98.8	100	97.5
abcd	100	100	98.3	100

ANN modeli, en iyi hiperparametreleri bulmak için ince ayar yapılmıştır. Bunlar arasında gizli katmanlar, aktivasyon fonksiyonları ve aşırı uyumu önleyen düzenleme parametreleri yer almıştır (Tablo 4.11).

Tablo 4.11. ANN modellerinin hiperparametreleri

ANN Modelinin Parametreleri	
PCs	{20, 50, 100}
Gizli Katman	{30, 50, 100}
Alpha	{0.0001, 0.001, 0.01}
Aktivasyonlar	{relu, tanh}

Birçok modelin doğruluğu yüksek ve benzer olmasına rağmen, “bd” spektral bölgesi, PCA olmadan, relu aktivasyonlu, 50 nöronlu bir gizli katmana sahip, basit ve en iyi olan bir model seçilmiştir (Tablo 4.12). Bunun seçimin nedeni, PCA olmadan, spektral bölgeden dolayı daha az girdi özelliği kullanan, hem eğitim hem de çıkarım sırasında hesaplama avantajı sağlamasıdır. İlave olarak, parmak izi bölgesi olarak kabul edilen "d" bölgesinin de "bd" kombinasyonuna dahil olmasıdır. Seçilen modelin test doğruluk oranı %98.8 olarak elde edilmiştir. Ayrıca seçilen modelin eğitim (%99.2) ve test (%98.8) doğruluk oranları arasında makul bir farka sahip olduğunu da belirtmek gerekir.

Seçilen modelin performans metrikleri, en yüksek ve en düşük modellerle birlikte ilgili spektral bölgeler için Tablo 4.12'de verilmiştir. Modelin performansını değerlendirmek için kullanılan doğruluk (accuracy), kesinlik (precision), duyarlılık (recall) ve F1 skoru gibi parametreler, modelin genel başarısını ve güvenilirliğini ortaya koymaktadır. %98.8 doğruluk oranı, modelin veri setindeki örneklerin %98.8'ini doğru sınıflandırdığını göstermektedir. Bu yüksek doğruluk oranı, modelin genel olarak başarılı bir performans sergilediğini ve yanlış sınıflandırmaların çok az olduğunu ifade etmektedir. Kesinlik değeri %98.7 olan model, tahmin edilen sınıflarda pozitif olarak sınıflandırdığı çeşitlerin %98.7'sinin gerçekten pozitif olduğunu belirterek, yanlış pozitif oranının çok düşük olduğunu göstermektedir. Benzer şekilde, %98.7 duyarlılık oranı, modelin pozitif olarak tahmin etmesi gereken çeşitlerin %98.7'sini pozitif olarak tahmin ettiğini ve pozitif çeşitleri kaçırma oranının çok düşük olduğunu ifade etmektedir. Son olarak, %98.7'lik F1 skoru, modelin hem yüksek kesinlik hem de yüksek duyarlılık arasında mükemmel bir denge kurduğunu ve dengeli bir performans sergilediğini göstermektedir. Bu sonuçlar, modelin yüksek performansını ve güvenilirliğini vurgulamakta olup, sınıflandırma görevinde modelin üstün başarısını ortaya koymaktadır. Özellikle bu parametreler, seçilen modelin doğru sınıflandırma yapma ve pozitif örnekleri doğru bir şekilde tanımlama yeteneğinin güçlü olduğunu kanıtlamaktadır.

Tablo 4.12. FTIR-tabanlı ANN model performans metrikleri

Spektral Bölge	Doğruluk Oranı	Metrikler (%)			
		Doğruluk	Kesinlik	Duyarlılık	F1-skoru
abd	En Yüksek Model	100	100	100	100
a	En Düşük Model	93.75	94.18	94.34	93.92
bd	Seçilen Model	98.8	98.7	98.7	98.7

Seçilen modelin hata matrisi, modelin dört farklı buğday çeşidini (Altındane, Çavuş, Flamura-85, Nevzatbey) yüksek doğrulukla sınıflandırabildiğini göstermektedir (Şekil 4.14). Altındane, Flamura-85 ve Nevzatbey çeşitlerinin tümü doğru şekilde sınıflandırılmıştır. Bu da seçilen modelin bu üç çeşidi mükemmel doğrulukla ayırt edebildiğini göstermektedir. Çavuş çeşidinin 19'u doğru sınıflandırılmış ve sadece 1'i Altındane çeşidi olarak yanlış sınıflandırılmıştır. Bu sonuçlar, seçilen modelin genel performansının oldukça iyi olduğunu ve buğdayların spektral özelliklerini etkili bir şekilde ayırt edebildiğini göstermektedir.

Gerçek	ANN	Tahmin			
		Altındane	Çavuş	Flamura-85	Nevzatbey
	Altındane	18	0	0	0
	Çavuş	1	19	0	0
	Flamura-85	0	0	22	0
	Nevzatbey	0	0	0	20

Şekil 4.14. Seçilen ANN modelinin hata matrisi

4.5.3.2. FTIR tabanlı destek vektör makineleri

Genel olarak SVM modeli, hem tek spektral bölgelerde hem eşleştirilmiş bölgelerde, hem de üç bölgenin kombinasyonlarında ve hatta tam spektrumlarda tüm veri kümelerinde iyi performans göstermiştir. PCA ile ve PCA olmadan sınıflandırma doğruluğu %98,88 ile %100 arasında değişmiştir (Tablo 4.13).

Tablo 4.13. Farklı spektral bölgeler ve bunların kombinasyonları kullanılarak SVM modellerinin eğitim ve test doğruluk oranları (%)

Spektral Bölge	SVM			
	PCA'lı		PCA'sız	
	Eğitim	Test	Eğitim	Test
a	99.2	98.8	99.2	98.8
b	99.6	100	99.6	100
c	100	98.8	100	98.8
d	99.6	100	99.6	100
ab	99.6	98.8	99.6	98.8
ac	99.6	100	99.6	100
ad	100	100	100	100
abc	100	100	100	100
abd	100	100	100	100
acd	99.6	98.8	99.6	98.8
bc	99.6	100	99.6	100
bd	99.6	100	99.6	100
bcd	99.6	98.8	99.6	98.8
cd	99.6	98.8	99.6	98.8
abcd	99.2	98.8	99.2	98.8

SVM modeli, en iyi hiperparametreleri bulmak için bu süreçte temel bileşen sayısı, düzenlileştirme parametresi ve kernel türleri dahil edilmiştir (Tablo 4.14).

Tablo 4.14. SVM modellerinin hiperparametreleri

SVM Modelinin Parametreleri	
Number of PCs	{20, 50, 100}
Kernel	{Linear, rbf}
SVM_C	{0.1, 1, 10}

Tüm modeller arasından, PC olmadan, "cd" parmak izi bölgesi kullanılarak eğitilen, lineer bir kernel ve svm_c değeri 1 olan model, en uygun model olarak seçilmiştir. Bu model, daha az girdi parametresi ve basit bir doğrusal çekirdek kullanarak %98.8 doğruluk oranı sağlamıştır. Ayrıca seçilen modelin eğitim (%99.6) ve test (%98.8) doğruluk oranları arasında makul bir fark korunmuştur.

Seçilen SVM modelinin performans metrikleri, en yüksek ve en düşük modellerle birlikte ilgili spektral bölgeler için Tablo 4.15'de verilmiştir. Bu modelin spektral bölge "cd" için performans metrikleri değerlendirildiğinde, %98.8'lik doğruluk oranı dikkat çekmektedir. Bu yüksek doğruluk, modelin gerçek sınıfları doğru tahmin etme yeteneğinin oldukça güçlü olduğunu göstermektedir. Kesinlik değeri %98.7 olan model, tahmin edilen sınıflarda pozitif olarak sınıflandırdığı çeşitlerin %98.7'sinin gerçekten pozitif olduğunu belirterek, yanlış pozitif oranının çok düşük olduğunu göstermektedir. Benzer şekilde, %98.7 duyarlılık oranı, modelin pozitif olarak tahmin etmesi gereken çeşitlerin %98.7'sini pozitif olarak tahmin ettiğini ve pozitif çeşitleri kaçırma oranının oldukça düşük olduğunu ifade etmektedir. %98.7'lik F1 skoru, modelin hem kesinlik hem de duyarlılık açısından dengeli ve yüksek performans sergilediğini ortaya koymaktadır. Bu metrikler, modelin güvenilir ve tutarlı bir sınıflandırma performansı gösterdiğini ve spektral bölge "cd" üzerinde oldukça etkili olduğunu göstermektedir.

Tablo 4.15. FTIR-tabanlı SVM model performans metrikleri

Spektral Bölge	Doğruluk Oranı	Metrikler (%)			
		Doğruluk	Kesinlik	Duyarlılık	F1-skoru
a	En Yüksek Model	100	100	100	100
cd	En Düşük Model	98.8	98.18	98.75	98.46
cd	Seçilen Model	98.8	98.7	98.7	98.7

Hata matrisi incelendiğinde, seçilen modelin Altındane, Çavuş, Flamura-85 ve Nevzatbey buğday çeşitlerini sınıflandırma performansı oldukça yüksek görünmektedir (Şekil 4.15). Altındane çeşidinin tamamı doğru bir şekilde Altındane olarak sınıflandırılmıştır. Çavuş buğday çeşidi için ise 20 örnekten 19'u doğru bir

şekilde sınıflandırılmış, ancak bir örnek yanlışlıkla Altındane olarak tahmin edilmiştir. Flamura-85 ve Nevzatbey buğday çeşitlerinin tamamı ise doğru bir şekilde sınıflandırılmıştır. Bu sonuçlar, modelin genel olarak yüksek bir doğruluk ve güvenilirlik ile çalıştığını ve sadece Çavuş buğdayında tek bir yanlış sınıflandırma yaparak çok düşük bir hata oranı sergilediğini göstermektedir. Dolayısıyla, modelin spektral özellikler temelinde buğday çeşitlerini ayırt etme yeteneğinin oldukça başarılı olduğu söylenebilir.

Gerçek	SVM	Tahmin			
		Altındane	Çavuş	Flamura-85	Nevzatbey
	Altındane	18	0	0	0
	Çavuş	1	19	0	0
	Flamura-85	0	0	22	0
	Nevzatbey	0	0	0	20

Şekil 4.15. Seçilen SVM modelinin hata matrisi

4.5.3.3. FTIR tabanlı random forest sınıflandırıcısı

RFC modelleri, PCA ile ve PCA olmadan tüm veri seti kombinasyonları için benzer prosedür izlenerek eğitilmiştir. Veri setleri eğitim ve test setlerine bölünmüştür. PCA ile ve PCA olmadan sınıflandırma doğruluk oranları %97.5 ile %100 arasında değişmiştir (Tablo 4.16).

Tablo 4.16. Farklı spektral bölgeler ve bunların kombinasyonları kullanılarak RFC modellerinin eğitim ve test doğruluk oranları (%)

Spektral Bölge	RFC			
	PCA'lı		PCA'sız	
	Eğitim	Test	Eğitim	Test
a	100	99.8	100	100
b	100	100	100	100
c	100	97.5	100	100
d	100	98.8	100	100
ab	100	98.8	100	100
ac	99.6	98.8	100	98.8
ad	100	98.8	100	100
abc	100	98.8	100	100
abd	100	98.8	100	100
acd	100	98.8	100	98.8
bc	99.6	97.5	100	100
bd	99.6	97.5	100	100
bcd	100	98.8	100	100
cd	100	98.8	100	100
abcd	100	98.8	100	100

İlgili hiperparametrelerin ince ayarını yapmak için 5 katlı çapraz doğrulama yapılırken bir grid araması uygulanmıştır. Hiperparametrelerin listesi Tablo 4.17'de verilmiştir.

Tablo 4.17. RFC modellerinin hiperparametreleri

RFC Modelinin Parametreleri	
PCs	{20,50,100}
Estimators	{50,100}
Max Depth	{10,20}
Min Samples to split	{5,10}
Min Samples leaf	{5,10}
Bootstrap	{True, False}

Tablo 4.16'da gösterildiği gibi, RFC, hem PCA ile hem de PCA olmadan yapılan yaklaşımlarda istisnai derecede yüksek eğitim ve test doğrulukları sergilemiştir. En düşük ve en yüksek doğruluk oranları sırasıyla %97.5 ve %100'dür.

RFC modelleri arasında, aşırı öğrenme (overfitting) sorununu ele alan, optimum hesaplama verimliliğini ve genel doğruluğu koruyan en basit ve en iyi model, 10 maksimum ağaç derinliği, bir yaprağı ayırmak için minimum 5 örnek, bir iç

düğümü ayırmak için minimum 5 örnek ve 50 tahminci veya ağaç sayısı ile eğitilmiş 100 ana bileşenli "bd" bölümüne ait bootstrap modelidir. ANN ve SVM durumlarından farklı olarak, RFC modelleri listesinde, %97.5 doğruluğa sahip en düşük performans gösteren model, eğitim ve test doğrulukları arasındaki kabul edilebilir fark nedeniyle aşırı öğrenme riskinin minimum olması temelinde en uygun model olarak seçilmiştir. Ayrıca bu model, parmak izi bölgesi "d"den girdi parametrelerini de içermektedir.

Seçilen RFC modelinin performans metrikleri, en yüksek ve en düşük modellerle birlikte ilgili spektral bölgeler için Tablo 4.18'de verilmiştir. Seçilen modelin performans metrikleri değerlendirildiğinde, modelin doğruluk oranının %97.5 olduğu görülmektedir. Bu yüksek doğruluk oranı, modelin genel sınıflandırma performansının oldukça tatmin edici olduğunu göstermektedir. Kesinlik değeri %97.75 olan model, pozitif sınıfların %97.75'ini doğru bir şekilde tahmin ederek yanlış pozitif oranının düşük olduğunu ortaya koymaktadır. %97.5'lik duyarlılık oranı, modelin doğru olarak pozitif sınıflandırması gereken örneklerin %97.5'ini doğru bir şekilde pozitif olarak tahmin ettiğini ve pozitif örnekleri kaçırma oranının düşük olduğunu ifade etmektedir. Aynı şekilde, %97.5'lik F1 skoru, modelin hem kesinlik hem de duyarlılık açısından dengeli bir performans sergilediğini ve yüksek bir doğruluk oranına sahip olduğunu göstermektedir. Bu metrikler, seçilen modelin spektral bölge "bd" üzerinde güvenilir ve tutarlı bir sınıflandırma performansı sergilediğini ifade etmektedir.

Tablo 4.18. FTIR-tabanlı RFC model performans metrikleri

Spektral Bölge	Doğruluk Oranı	Metrikler (%)			
		Doğruluk	Kesinlik	Duyarlılık	F1-skoru
abcd	En Yüksek Model	100	100	100	100
bd	En Düşük Model	97.5	97.75	97.5	97.5
bd	Seçilen Model	97.5	97.75	97.5	97.5

Hata matrisi incelendiğinde, RFC modelinin Altındane, Çavuş, Flamura-85 ve Nevzatbey buğday çeşitlerini sınıflandırma performansının oldukça yüksek olduğu görülmektedir (Şekil 4.16). Altındane örneklerinin tamamı doğru bir şekilde sınıflandırılmıştır. Çavuş buğday çeşidi için ise 20 örnekten 18'i doğru olarak sınıflandırılmış; ancak bir örnek yanlışlıkla Altındane, bir diğer örnek ise Flamura-85 olarak tahmin edilmiştir. Flamura-85 ve Nevzatbey buğday çeşitlerinin tamamı doğru bir şekilde sınıflandırılmıştır. Bu sonuçlar, modelin genel olarak yüksek bir doğruluk

ve güvenilirlik ile çalıştığını, ancak Çavuş buğdayı için bazı sınıflandırma hataları yaptığını göstermektedir. Yanlış sınıflandırmalar, özellikle Çavuş buğday çeşidi için modelin daha fazla iyileştirilmesi gerektiğini işaret etmektedir. Bununla birlikte, genel olarak modelin diğer buğday çeşitlerini doğru bir şekilde ayırt etme yeteneği oldukça başarılıdır ve spektral özellikler temelinde yüksek performans sergilediğini göstermektedir.

Gerçek	RFC	Tahmin			
		Altındane	Çavuş	Flamura-85	Nevzatbey
	Altındane	18	0	0	0
	Çavuş	1	18	1	0
	Flamura-85	0	0	22	0
	Nevzatbey	0	0	0	20

Şekil 4.16. Seçilen RFC modelinin hata matrisi

4.5.3.4. FTIR tabanlı K-en yakın komşu algoritması

KNN modelleri, PCA ile ve PCA olmadan tüm veri seti kombinasyonları için benzer prosedür ile eğitilmiştir. Eğitim ve test doğruluk oranları Tablo 4.19’de verilmiştir.

Tablo 4.19. Farklı spektral bölgeler ve bunların kombinasyonları kullanılarak KNN modellerinin eğitim ve test doğruluk oranları (%)

Spektral Bölge	KNN			
	PCA’lı		PCA’sız	
	Eğitim	Test	Eğitim	Test
a	100	62.5	100	63.8
b	100	71.3	100	68.8
c	100	72.5	100	62.5
d	100	76.3	100	71.3
ab	100	71.3	100	63.8
ac	100	77.5	100	71.3
ad	100	72.5	100	73.8
abc	100	76.3	100	71.3
abd	100	76.3	100	70
acd	100	77.5	100	75
bc	100	76.3	100	71.3
bd	100	77.5	100	73.8
bcd	100	78.8	100	73.8
cd	100	78.8	100	76.3
abcd	100	80	100	70

İlgili hiperparametrelerin ince ayarını yapmak için kullanılan hiperparametrelerin listesi Tablo 4.20'de verilmiştir.

Tablo 4.20. KNN modellerinin hiperparametreleri

KNN Modelinin Parametreleri	
PCs	{20, 50, 100}
KNN metrics	{Euclidean, manhattan}
KNN weights	{uniform, distance}

En iyi ve uygun KNN modeli, tam spektral bölgede 100 temel bileşen (PCs) ile eğitilmiş, Manhattan uzaklık (distance) metriği ve uzaklık tabanlı KNN ağırlığı kullanılan bir model olarak seçilmiştir. Ancak, KNN modeli, diğer tüm modellere kıyasla en kötü performansı sergiledi. En düşük ve en yüksek test doğruluk oranları %62.5 ile %80 arasında değişmiştir. (Tablo 4.21). Ayrıca, tüm durumlarda eğitim ve test doğrulukları arasındaki belirgin farklar, aşırı öğrenme riskinin yanı sıra modelin test setlerinde ciddi genelleme sorunları yaşadığını göstermektedir.

Seçilen KNN modelinin performans metrikleri, en yüksek ve en düşük modellerle birlikte ilgili spektral bölgeler için Tablo 4.21'da verilmiştir. Verilen paragraf, "abcd" spektral bölgesi için bir KNN modelinin performansını açıklamaktadır. Modelin %80.0 doğruluk oranı, genel olarak kabul edilebilir bir sınıflandırma performansı sergilediğini göstermektedir. Kesinlik ve duyarlılık değerleri sırasıyla %80.78 ve %80.57 olarak ölçülmüştür, Bu durum modelin pozitif tahminlerinin çoğunun doğru olduğunu ve pozitif sınıfları kaçırma oranının düşük olduğunu göstermektedir. Ayrıca, %80.58'lik F1 skoru, modelin hem kesinlik hem de duyarlılık arasında iyi bir denge sağladığını ve genel olarak dengeli bir performans sergilediğini belirtmektedir. Ancak, KNN modelinin bu bölgede gösterdiği performans, diğer spektral bölgelerdeki modellerle karşılaştırıldığında ortalamadan düşük kalmaktadır.

Tablo 4.21. FTIR-tabanlı KNN model performans metrikleri

Spektral Bölge	Doğruluk Oranı	Metrikler (%)			
		Doğruluk	Kesinlik	Duyarlılık	F1-skoru
abcd	En Yüksek Model	80.0	80.78	80.57	80.58
c	En Düşük Model	62.5	64.62	62.97	63.43
abcd	Seçilen Model	80.0	80.78	80.57	80.58

Verilen hata matrisi, KNN modelinin performansını değerlendirmek için kullanılmaktadır (Şekil 4.17). Altındane, Çavuş, Flamura-85 ve Nevzatbey buğday çeşitlerini sınıflandırmak için kullanılan modelin doğruluğunu analiz etmek mümkündür. Altındane çeşidi için tüm örneklerin doğru bir şekilde sınıflandırıldığı görülmektedir. Çavuş çeşidi için ise 20 örneğin 14'ü doğru bir şekilde sınıflandırılmış, ancak 2 örnek yanlışlıkla Flamura-85 olarak ve 4 örnek de Nevzatbey olarak tahmin edilmiştir. Flamura-85 çeşidi için ise 22 örneğin 17'si doğru bir şekilde sınıflandırılmış, ancak 3 örnek yanlışlıkla Çavuş olarak ve 2 örnek de Nevzatbey olarak tahmin edilmiştir. Nevzatbey çeşidi için ise 20 örneğin 15'i doğru bir şekilde sınıflandırılmış, ancak 4 örnek yanlışlıkla Çavuş olarak ve 1 örnek de Flamura-85 olarak tahmin edilmiştir. Bu matris, modelin bazı sınıflar için doğru tahminler yapabildiğini ancak özellikle Çavuş ve Flamura-85 çeşitleri için karışıklıklar yaşandığını göstermektedir. Modelin buğday çeşitlerini sınıflandırma becerisi genel olarak orta düzeydedir ve ancak belirli sınıflar için iyileştirmeler gerekebilir.

Gerçek	KNN	Tahmin			
		Altındane	Çavuş	Flamura-85	Nevzatbey
	Altındane	18	0	0	0
	Çavuş	0	14	2	4
	Flamura-85	0	3	17	2
	Nevzatbey	0	4	1	15

Şekil 4.17. Seçilen KNN modelinin hata matrisi

4.5.4. Genel istatistiksel değerlendirme ve seçilen modeller

FTIR spektroskopisi tabanlı ML sınıflandırma modeli, bu çalışmada dört buğday çeşidinin sınıflandırılmasında spektroskopi ve ML entegrasyonunun potansiyelini ortaya koymuştur. Ayrıca, tüm otuz veri kümesi ve bunların karşılık gelen test doğrulukları (Tablolar 4.10, 4.13, 4.16, ve 4.19) Friedman testi ile dört model arasında önemli farklılıkların olduğunu göstermiştir ($P < 0.01$). Nemenyi testi kullanılarak yapılan ardışık ikili karşılaştırmalar, modeller arasında istatistiksel olarak anlamlı farklılıklar sergileyen model çiftlerini belirlemiştir. Bu karşılaştırmalara göre, ANN ve SVM modelleri ile SVM ve RFC modelleri, aynı zamanda ANN ve RFC modelleri performans açısından anlamlı farklılık göstermemiştir. Ancak, ANN ve KNN modelleri ile SVM ve KNN modelleri ile RFC ve KNN modelleri arasında test

doğrulukları açısından istatistiksel olarak anlamlı farklar bulunmuştur.

Tablo 4.22, seçilen dört modelin performans ölçütlerini ve bunların ilgili spektral bölgelerini göstermektedir. Toplam 120 model içinden performans ve optimum girdi ve hiperparametrelerine göre, bu dört model seçilmiştir. ANN ve SVM modelleri, KNN'den önemli ölçüde daha yüksek performans gösterirken RFC'den ise nispeten daha iyi sonuçlar elde edilmiştir. ANN ve SVM karşılaştırıldığında, her iki model de benzer doğruluk elde etmesine ve 1500-650 cm⁻¹ parmak izi bölgesini kullanmasına rağmen, SVM'nin eğitiminde, "cd" bölgesindeki kullanılan spektral süreklilik göz önüne alındığında SVM daha tercih edilebilirdir. Bunun nedeni, ANN'yi eğitmek için kullanılan spektral bölgelerin komşu bölgeler olmamasıdır. Bununla birlikte, SVM, "cd" bölgesinin spektral sürekliliğinin, herhangi bir müdahale olmaksızın kapsamlı bir örüntü ve özellik kümesini temsil eden ardışık dalga sayılarını içerdiği için daha faydalı olduğu sonucuna varılabilir. Dolayısıyla, yukarıdaki nedenler ve SVM'nin daha küçük veri kümelerindeki sağlamlığı dikkate alındığında, SVM'nin FTIR-spektroskopik verileri kullanarak buğday çeşitlerini sınıflandırmak için en uygun model olduğu görülmektedir.

Tablo 4.22. Seçilen modellerin performans metrikleri

Seçilen Modeller	Spektral Bölge	Metrikler (%)			
		Doğruluk	Kesinlik	Duyarlılık	F1-skoru
ANN	bd	98.8	98.7	98.7	98.7
SVM	cd	98.8	98.7	98.7	98.7
RFC	bd	97.5	97.75	97.5	97.5
KNN	abcd	80.0	80.78	80.57	80.58

5. TARTIŞMA

Buğday çeşitlerinin sınıflandırılmasıyla ilgili yapılan çalışmalara göre, literatürde sınıflandırılan çeşitlerin sayısı, kullanılan özellikler-veriler, sınıflandırma modellerini geliştirmek için kullanılan ilgili algoritmalar ve nihayetinde elde edilen sınıflandırma doğruluğu araştırmanın kapsamını belirlemektedir. Bu nedenle, bulgularımızı daha önce yapılan çalışmalarla ilişkili olarak tartışmak için yukarıdaki hususlar önemlidir. Buğday genellikle makarnalık ve ekmeklik olarak sınıflandırıldığından, ikili sınıflandırmalar yaygın olarak yapılmaktadır (Sabancı et al., 2017; Sabancı et al., 2020). Bununla birlikte, agronomik ve yönetsel önlemler amacıyla ve sürekli artan buğday çeşit sayısı ve bunlara bağlı olarak belirli çeşitlerin bilinmesi ihtiyacı ile araştırmacılar, birden fazla kategoriye tanıma yeteneğine sahip yapay zeka tabanlı sınıflandırma modelleri geliştirmeye yönelmişlerdir.

Türkiye dışındaki araştırmalarda, (Zapotoczny, 2011), ANN modelleri geliştirerek on bir buğday çeşidini %100 doğrulukla sınıflandırmıştır. (Pourreza et al., 2012), dokusal özellikleri kullanarak dokuz İran buğday çeşidini sınıflandırmak için bir model geliştirmiştir. Gri seviye eş-oluşum matrisi (Gray Level Cooccurrence Matrix, GLCM), gri seviye çalışma uzunluğu matrisi (Gray Level Run-Length Matrix, GLRM), yerel ikili desenler (Local Binary Patterns, LBP), yerel benzerlik desenleri (Local Similarity Patterns, LSP) ve yerel benzerlik sayıları (Local Similarity Numbers, LSN) özelliklerinden 131 özelliği birleştiren doğrusal bir diskriminant analizinin (Linear Discriminant Analysis, LDA) %98.15'lik bir doğruluk elde ettiğini bildirmişlerdir. (Khoshroo et al., 2014) de dört buğday çeşidinin sınıflandırılması için ANN modelini eğitmek üzere dokuz morfolojik özellik uygulamış ve %85.72'lik bir doğruluk oranı elde etmişlerdir. Üç buğday çeşidini sınıflandırmak amacıyla yapılan bir çalışmada, (Sabancı ve Akkaya, 2016), yedi geometrik özellik kullanarak ANN modeli eğitmiş ve %97.17 sınıflandırma başarısına ulaşmışlardır. (Zhou et al., 2020), NIR spektral verilerini kullanarak 30 buğday çeşidini sınıflandırmak için bir CNN modeli geliştirmiş ve benzer gelişmeler (Khojastehnazhand ve Roostaei, 2022)'nin çalışmalarında da gerçekleştirilmiştir. İlgili çalışmaların ve bu çalışmada önerilen modelin özeti Tablo 5.1'de verilmiştir.

Türkiye ile ilgili çalışmalarda, çeşit sayısı ondan az olan çok sınıflı buğday çeşit sınıflandırma çalışmalarının yanı sıra, çok daha büyük sınıflar için sınıflandırma çalışmaları da yapılmıştır. (Taner et al., 2018), 28'i ekmeklik ve 11'i makarnalık

buğday olmak üzere toplam 39 çeşidi, geometrik özellikleri kullanılarak %99.99 gibi yüksek bir doğrulukla sınıflandırmak için ANN modeli geliştirilmişlerdir. (Özkan et al., 2021), VGG16'yı kullanarak elde ettikleri derin özellikleri, SVM ile entegre ederek 50 buğday çeşidini %99 doğrulukla başarıyla sınıflandırmışlardır. Benzer şekilde, (Çelik et al., 2022), MobileNetV2'den elde edilen derin özelliklerin SVM ile birleştirilmesiyle 41 makarnalık çeşidi %91.89 doğrulukla tanımlamıştır. Son zamanlarda, (Ceyhan et al., 2024), spektrometre yansımasından elde edilen görüntüleri kullanarak öğrenme aktarımı yaklaşımını araştırmış ve sıfırdan bir CNN modeli eğiterek 24 buğday çeşidini %99.52 doğrulukla sınıflandırmayı başarmışlardır. Bu sonuçlar ışığında, sınıf sayısındaki önemli artış, çok sınıflı buğday çeşitleri sınıflandırmasını çözmek için uygun özelliklerin ve modellerin uygunluğunu göstermektedir. Bu nedenle, bu çalışmada Türkiye'ye ait buğday çeşitlerinin sayısı önceki çalışmaların yaklaşık iki buçuk katına çıkarılarak 124'e yükseltilmiş ve sonuçlar bölümünde gösterildiği gibi, ANN ve SVM modelleri kullanılarak %98.8'in üzerine çıkan umut verici sınıflandırma doğrulukları elde edilmiştir.

Kullanılan özelliklerle ilgili olarak, bu çalışma, renk ve HOG tanımlayıcılarından piksel tabanlı ve yapısal özelliklerin avantajlarından yararlanmak için iki farklı özellik tanımlayıcının entegrasyonuna odaklanmıştır. RGB renklerinin istatistiksel dağılımı, buğday çeşitlerinin temel renk özelliklerine ulaşarak kompakt bir temsili yakalamak için kullanılırken, HOG özellikleri kenarları ve bunların değişim büyüklüğünü yakalayarak yapısal ve dokusal özellikleri temsil etmiştir. Bu iki tanımlayıcının entegre edilmesi, sınıflandırıcıların geliştirilmesi üzerinde tamamlayıcı bir etkiye sahiptir. Birleştirilmiş temsil, renk dağılımı benzer olan ancak dokusal özellikleri farklı olan nesneleri ayırt etmeyi sağlar. Özellikle, HOG gibi özellikler, modellerin renk dağılımlarından daha ayırt edici ve belirleyici olan sınır bilgilerini kullanmasına olanak tanır. (Hoang et al., 2020), derin öğrenme ve elle oluşturulan özelliklerin karşılaştırmalı çalışmasında HOG özelliklerinin avantajını göstermiştir. Tüm elle oluşturulan özellikler arasında HOG, en iyi doğruluğu elde etmiştir. (Taner et al., 2023) tarafından yapılan çalışma, HOG ve renk momentlerinin birlikte kullanılmasının avantajını göstermiştir. Bu çalışmada SVM, RFC, ANN ve KNN modelleri sırasıyla %98.17, %96.67, %98.62 ve %91.28 doğruluk oranlarına ulaşmıştır. (Arif et al., 2023) de çalışmalarında, HOG ve diğer özelliklerin birleşik faydasını göstermiş ve XGBoost algoritmasını kullanarak %94.1 doğruluk elde

etmiştir. Tarım sektöründe çeşit sınıflandırmanın yanı sıra, (Chu et al., (2018), zararlı izleme için HOG tanımlayıcılarının ve SVM modelinin uygulanabilirliği ile etkinliğini göstermiş ve diğerleri arasında en iyi doğrulukları sergilemiştir. Benzer şekilde, Kasinathan ve Uyyala (2021) zararlıları tanımlamak ve sınıflandırmak için yaptıkları çalışmada, doku, renk, şekil, HOG ve genel görüntü tanımlayıcılarının kombinasyonunda topluluk sınıflandırıcıları ile daha iyi doğruluk elde etmişlerdir. Dolayısıyla, mevcut literatürün bir uzantısı olarak, bu çalışma, birleştirilmiş özelliklerin tutarlı etkinliğini pekiştirmekte ve bu özelliklerin sınıflandırma görevlerindeki dayanıklılığını doğrulamaktadır. Ayrıca, bu çalışma, renk ve dokusal özelliklerin yüzlerce kategorideki çok sınıflı sınıflandırma problemlerinde güvenilirliğini göstermiştir.

Eğitim veri setinin büyüklüğü, sınıf sayısına orantılı olarak eğitim başarısına büyük katkı sağlar. (Ceyhan et al., 2024) tarafından yapılan çalışmada, 28 sınıfın sınıflandırma modelini geliştirmek için 57.600 görüntü kullanılmıştır. Ferentinos (2018), 87.848 görüntü kullanarak ikili bir hastalık sınıflandırma modeli geliştirmiş ve %99.53 doğruluk elde etmiştir. Bu çalışmada, sınıf (124 çeşit) sayısının nispeten daha fazla olması nedeniyle, sınırlı eğitim verilerinden kaynaklanabilecek sorunları önlemek için 465.000 görüntü kullanılmıştır. Nispeten büyük eğitim verilerine ek olarak, buğday tanelerinin çoklu görünümü, makine ve derin öğrenme modellerinin giriş görüntülerinde düşük performansa neden olabilecek küçük değişikliklere karşı savunmasızlığını hafifletmeyi sağlamıştır. Veri artırımı (data augmentation), orijinal görüntülerdeki küçük değişikliklerle birden çok görüntü üretmeyi içeren yaygın olarak uygulanan bir stratejidir (Çelik et al., 2022). Ayrıca (Taner et al., 2023) ve (Taner et al., 2024) çalışmalarında, elma görüntülerinin üç farklı açıdan çekilmesinin, on elma çeşidini sınıflandırmada yüksek doğruluk oranlarına ulaşılmasını sağlamış ve bu stratejinin uygulanabilirliğini ve uygunluğunu kanıtlamıştır.

Veri artırmanın yanı sıra, çoklu görünüme sahip veri setleri, çeşitli açılardan elde edilen özelliklerin zenginleştirilmesini sağlar. Bu durum, modelin belirli özellikleri ezberlemesini engelleyerek aşırı öğrenme riskini azaltır ve modelin yeni daha önce görülmemiş veriler üzerinde genelleme yeteneğini artırır. Ayrıca, modelin gerçek varyasyonlara odaklanmasını sağlar ve çok küçük varyasyonlardan kaynaklanan yanlış sınıflandırma olasılığını azaltır. Eğitim sırasında model, bir nesnenin temel özelliklerini derinlemesine anlamak için geniş bir dönüşüm yelpazesine maruz kalır ve

rastgele ve kasıtlı gürültülerden ayırt edici gerçek özellikleri ayırt etme konusundaki sağlamlığını artırır. Dolayısıyla, bu tez kapsamında yürütülen çalışmada, eğitim verisi setinin hem büyüklüğü hem de çok yönlü görünümü, sınıflandırma modellerinin genel başarısına önemli ölçüde katkıda bulunmuştur

Ayrıca, HOG ve renk momentlerini birleştirme stratejisi, özellikle ANN modellerinin eğitiminde önemli bir rol oynamıştır. Bu strateji, epok açısından eğitim süresini azaltırken genel test doğruluğunu ise artırmıştır. Özellikle 350-350 gizli katmanlı model için yapılandırılmış PCA, doğruluğun %94.7'ten %96.8'e yükselmesine katkıda bulunmuştur (Tablo 4.2). Başka bir deyişle, PCA'nın entegrasyonu ile modelin daha önce yanlış sınıflandırılan 1.180 görüntüü daha doğru sınıflandırabildiği söylenebilir. Benzer şekilde, SVM modelinde PCA'nın eğitime entegre edilmesi doğruluğun %97'den %98.1'e çıkmasını sağlamıştır (Tablo 4.3).

Bu nedenle, Şekil 4.4'de verilen bağımsız deneyde PCA ile eğitilen ANN modellerinin tutarlı doğruluk artışı, doğruluğun rastgele bir şans eseri olmadığını kanıtlamaktadır. Ayrıca, doğası gereği farklı olan iki modelde de aynı etkinin görülmesi, PCA'nın orijinal HOG ve renk özelliklerinin ayırt edici karakterlerine daha fazla ağırlık kattığını göstermektedir. Sayısal olarak %1.1'lik bir artış küçük gibi görünebilir, ancak büyük ölçekli bir sınıflandırmada binlerce daha fazla örneğin doğru bir şekilde sınıflandırılması anlamına gelir. Bu sonuçlar, PCA'nın model eğitimi sürecinde daha sağlam ve güvenilir sınıflandırma sağlayan önemli bir bileşen olduğunu doğrulamaktadır.

Laabassi et al. (2021), dört buğday çeşidini sınıflandırmak için öğrenme aktarımını kullanmış ve sırasıyla MobileNet ve DenseNet201 modelleri ile %98.1 ve %98.3 başarı elde etmiştir. Taner et al. (2021) ise öğrenme aktarımının yanı sıra, 17 fındık çeşidini sınıflandırmak için dört evrişimli katmandan oluşan yeni ve sıkı bir CNN modeli tasarlamış ve önceden eğitilmiş modelleri geride bırakarak %98.63 genel doğruluk oranı yakalamıştır. Bu çalışmada da sıkı CNN ile MobileNet, DenseNet201 ve InceptionV3 modellerine dayalı öğrenme aktarımı sonuçları incelenmiş ve sırasıyla %95.40, %90.54, %92.41 ve %83.47 test doğrulukları ile umut verici performanslar elde edilmiştir. Önerilen CNN modeli, %95.40 genel test doğruluğu ile popüler mimarileri geride bırakmıştır. DenseNet201, her katmanı ileri besleme yoluyla diğer tüm katmanlara bağlayarak özellik kullanımını maksimize etmek için tasarlanmıştır. Bu yoğun bağlantı, özellikle 124 buğday çeşidi gibi birçok sınıf arasında küçük

farkların olduğu durumlarda, verideki karmaşık desenleri yakalamak açısından kritik öneme sahiptir. DenseNet201'in derin özellik temsil yetenekleri, %92.41 test doğruluğuna ulaşmasında önemli bir rol oynamıştır. MobileNet, genellikle daha hafif bir model olarak kabul edilir ve eğitilebilir parametre sayısı DenseNet201 modeline oldukça yakındır. Verimliliği optimize edilmiş olup, çok fazla doğruluk kaybı yaşanmadan kullanılır. Derinlemesine ayrılabilir evrişimler kullanarak ilgili özellikleri verimli bir şekilde yakalama yeteneği, MobileNet'i etkili kılmış ve %90.54 doğruluk oranına, DenseNet201 ve InceptionV3'ten daha küçük bir parametre sayısı ile ulaşmasını sağlamıştır. InceptionV3 mimarisi, eğitilebilir parametre sayısı açısından MobileNet'ten büyük, ancak DenseNet201'den küçüktür ve paralel olarak farklı boyutlardaki birçok evrişim filtresi kullanarak görüntülerdeki farklı soyutlama seviyelerini yakalayabilmektedir. InceptionV3, üç önceden eğitilmiş model arasında en düşük doğruluk olan %83.47 doğruluk oranını elde etmiş olmasına rağmen, tasarımı, çok sayıda sınıfı ayırt etmede gerekli olan çeşitli özellikleri yakalama ve görünmeyen verilere genelleme yapmada hala avantajlıdır.

Önerilen CNN modeli, daha hafif ve en iyi performansı sergileyen model olması açısından avantajlıdır. Eğitilebilir parametrelerin boyutunu karşılaştırdığımızda, 5.9 milyon eğitilebilir parametreye sahip önerilen CNN modelinin, öğrenme aktarımı (transfer learning) modellerine göre 2 ila 3 kat daha küçük olduğu görülmektedir. Bu durum, öğrenme aktarımının genellikle iyi bir başlangıç noktası sağladığını, ancak hedef problemin orijinal modellerin eğitildiği görevlerden önemli ölçüde farklı olması durumunda her zaman en iyi sonuçları vermeyebileceğini göstermektedir. Bu özel problem için, daha büyük ve daha karmaşık ağlar yerine, daha basit bir mimari daha etkili olabilir. Ayrıca, bu çalışmada kullanılan çoklu görünüm görüntü veri kümesi, 124 buğday çeşidinin daha zengin ve çeşitli bir temsili sağladığı için modelin başarısına önemli ölçüde katkıda bulunmuştur. Bu da modelin sağlamlığını artırmıştır. İlgili literatürle desteklenen bu bulgular, özellikle karmaşık sınıflandırma görevleriyle başa çıkarken, sadece önceden eğitilmiş modellere güvenmek yerine, özel CNN mimarilerinin göz önünde bulundurulmasının önemini vurgulamaktadır. Yukarıda belirtilen ve ilgili literatür kaynaklarıyla desteklenen bu çıkarımla, yüksek hesaplama kaynaklarına ilişkin zorluklar bulunmasına rağmen, özel CNN modelinin avantajlarının dezavantajlarını aştığı söylenebilir. Çoklu görünüm görüntü veri kümesine sahip önerilen CNN modeli, 124 buğday çeşidini sınıflandırmak gibi zorlu

bir görevde modelin genel başarısına katkıda bulunmuştur. Çünkü sınıfların daha zengin ve daha çeşitli bir temsilini sağlar ve nihayetinde modelin sağlamlığını artırır.

Tablo 5.1. İlgili çalışmaların ve önerilen modellerin özeti

Çeşit Sayısı	Özellikler	Modeller ve Doğruluk	Referanslar
11	Dokusal özellikler	ANN - %100	(Zapotoczny, 2011)
6	Renk ve şekil özellikleri	MLP - %87.2	(Pazoki ve Pazoki, 2011)
9	Dokusal özellikler	LDA - %98.15	(Pourreza et al., 2012)
4	Morfolojik özellikler	ANN - %85.7	(Khoshroo et al., 2014)
3	Geometrik özellikler	KNN - %95.71	(Sabancı ve Akkaya, 2016)
2	Renk, şekil ve doku özellikleri	ANN- (MAE) 9.8×10^{-6}	(Sabancı et al., 2017)
39	Geometrik özellikler	ANN - %99.9	(Taner et al., 2018)
2	Dalgacık tabanlı	ANN - %98.28	(Sabancı et al., 2020)
30	NIR spektral verileri	CNN - %93.01	(Zhou et al., 2020)
50	Derin özellikler	SVM - %99.0	(Özkan et al., 2021)
41	Derin özellikler	SVM - %91.89	(Çelik et al., 2022)
15	RGB Görüntüler	CNN - %97.52	(Lingwal et al., 2021)
7	GLCM, LBP	SVM - %98.1	(Khojastehnazh ve Roostaei, 2022)
40	VL, VNIR ve SWIR Görüntüleri	CNN - %99.94	(Işık et al., 2022)
4	Derin özellikler	SVM - %99.91	(Sonmez et al., 2024)
4	FT-MIR spektroskopisi	ANN - %100	(Fattahi et al., 2024)
5	Derin özellikler	BiLSTM - %97.73	(Yasar et al., 2024)
24	Spektrometre yansımaları	CNN - %99.52	(Ceyhan et al., 2024)
12	Tetrahetrz Spektroskopisi	CNN - %97.8	(Chen et al., 2024)
124	HOG ve Renk Özellikleri	ANN - %96.8, SVM - %98.1	Önerilen Modeller
124	RGB Görüntüler	CNN - %95.40	Önerilen Model

FTIR spektral verilerine dayalı yaklaşımla ilgili olarak, literatürler spektroskopinin taşıyıcı tespiti ve menşee tanımlamasında yaygın olarak uygulandığını ortaya koymaktadır. (Murru et al., 2019) FTIR spektroskopisinin üzümlerin olgunluğu ve çeşit sınıflandırması ile ilgili kimyasal içerikleri belirleme yeteneğini göstermiş,

FTIR spektrumunun istatistiksel olarak benzer olmasına rağmen sınıflandırmada iyi sonuçlar elde ettiği sonucuna varmışlardır. (Zhou et al., 2020), üç çay çeşidini sınıflandırmak için yaptıkları araştırmada, FTIR yansıma spektral verilerini ve kümeleme algoritmalarını kullanarak %98.5'lik en yüksek doğruluğu elde ettiklerini bildirmişlerdir.

(Giang et al., 2020), FTIR spektroskopisi, PCA ve Parçalı En Küçük Kareler (Partial Least Squares, PLS) ile Ortogonal Parçalı En Küçük Kareler- Diskriminant Analiz (Orthogonal Partial Least Square-Discriminant Analysis, OPLS-DA) yöntemlerini kullanarak yedi pirinç çeşidini %100 doğrulukla sınıflandırdıklarını rapor etmiştir. (Srinuttrakul et al., 2021) ise Tayland'ın iki bölgesinden getirilen pirincin coğrafi konumunu, FTIR spektroskopisi ve Doğrusal diskriminant analizi (Linear discriminant analysis, LDA) kullanarak iki üretim yılından toplanan bir veri seti için %96.97 ve %100 doğruluk elde etmiştir. (Unuvar et al., 2021), farklı spektroskopik yöntemlerle buğdayın ekmeklik ve makarnalık buğday olarak ayrımını incelemişlerdir. PCA ve parçalı en küçük kareler-diskriminant analizi kullanarak, senkron floresans spektroskopisi (Synchronous Fluorescence Spectroscopy, SFS) ile %100 ayırım oranı elde edilmiştir. Son olarak, (Fattahi et al., 2024), FT-MIR spektroskopisi ve ML modelleri kullanarak dört buğday unu çeşidini %100 mükemmel doğrulukla sınıflandırmayı başarmışlardır.

Bu çalışmada elde edilen sonuçlar, ML modellerinin FTIR geçirgenlik verilerinden buğdayın özgün özelliklerini ayırt etme yeteneğini ortaya koymuştur, Önceki araştırmalarla uyumlu bir şekilde, ANN, SVM ve RFC modellerinin yüksek doğrulukları, spektral verilerin yüksek boyutluluğuna rağmen incelenen dört buğday çeşidinin benzersiz ve tanımlanabilir spektral özelliklere sahip olduğunu göstermektedir. Ayrıca, giriş veri kümesini spektral bölgelere göre bölmek stratejisinin etkili olduğu gözlemlenmiştir. Yani belirli bölgelerin kullanımı tüm spektrumun kullanımı kadar etkili olabilmektedir. Spektrumları bölmek, giriş parametrelerini azaltma avantajı sağlamaktadır. ANN, SVM ve RFC modelleri, her biri ayrı ayrı kullanıldığında (a, b, c ve d bölgeleri), %97.5 ile %100 arasında değişen en yüksek doğruluklara ulaşmıştır. Ayrıca, spectral bölgelerin tüm potansiyel kombinasyonları (ikili ve üçlü gruplar halinde) çok yüksek doğruluklarla sonuçlanmış, bu da 3.000'den fazla özelliğe sahip tam spektrumların kullanılmasının özellikle çeşit sınıflandırma görevinde gerekli olmadığını kanıtlamıştır. Bununla birlikte, kimyasal analiz söz

konusu olduğunda, spektrumlardaki her bilgi eşit derecede önemlidir. Özellikle RFC ve SVM modellerinde, parmak izi bölgesi olarak bilinen “d” bölgesi (1500-650 dalga sayısı cm^{-1}), PCA uygulamasına gerek olmaksızın %100 doğruluk elde etmiştir. Bu durum, bu bölgenin spektrumun tamamını bağımsız bir şekilde temsil edebildiğini ve ilave boyutsal azaltma gerektirmeksizin hesaplama verimliliği sağladığını ortaya koymaktadır. Ancak model güvenilirliği söz konusu olduğunda %98,8 sınıflandırma doğruluğuna sahip SVM modeli hem verimlilik hem de spektral süreklilik açısından tercih edilmektedir.



6. SONUÇLAR VE ÖNERİLER

Buğday, en çok tüketilen tahıllar arasında yer almaktadır. Artan talebi karşılamak için tarım sektöründe çeşitli teknolojik gelişmeler yaşanmış ve bu durum endüstriyi birçok yönden dönüştürmüştür. Modern buğday ıslahı yaklaşımlarının yaygın olarak kullanılması, yeni buğday çeşitlerinin seçimi ve ıslahını geliştirmiştir. Buğday gibi tarımsal ürünlerin işlenmesi aşamasında, üreticiler, sanayiciler ve tüketiciler, ürünün çeşitlerini tam olarak bilmek isterler. Daha iyi tarım uygulamaları için, çiftçilerin de ektikleri ürünün çeşidini bilmeleri gerekmektedir. Aynı şekilde, pazarlamacılar da hedef pazarlar için standartlar oluşturmak amacıyla sattıkları ürünün çeşidinden emin olmak istemektedirler. Ayrıca, gıda ürünlerine olan artan talep ve karmaşık küresel tedarik zinciri, gıda sahtekarlığı ve taşımasına yol açmıştır. Bu nedenlerle, çeşitlerin tanımlanması için güvenilir yöntemlere ihtiyaç vardır.

Benzer çeşitlerin manuel olarak tanımlanması zordur, tutarsızdır ve insan becerilerine dayandığı için etkisizdir. Bu sorunları çözmek için, sanayi ve laboratuvarların manuel operasyonları otomatik sistemlerle değiştirerek süreci en verimli şekilde kontrol etmelerine yardımcı olacak yeni ve yenilikçi teknolojiler ortaya çıkmıştır. Bu nedenle, tahılların tanımlanması ve sınıflandırılması için verimli bir yöntem geliştirilmesi önemlidir.

Bu tez, 124 buğday çeşidinin görüntülerini toplayarak ve farklı açılardan 465.000 görüntü kullanılarak yapay zeka yöntemleriyle buğday çeşitlerini sınıflandırmayı sunmaktadır.

Tezde 465.000 görüntüye sahip büyük bir veri tabanı oluşturulmuştur. Bu alanda yer alan araştırmacıların ImageNet'te olduğu gibi yeni çalışmalar yapmasına imkan sağlanacaktır.

Özellikle, bu çalışma büyük görüntü veri setinin oluşturulmasına, HOG ve renk özelliklerinin çıkarılmasına, ML ve DL modellerinin eğitilmesine, geliştirilmesine ve test edilmesine odaklanmıştır. Ayrıca, FTIR spektroskopisi ve ML modellerinin buğday çeşitlerinin sınıflandırılmasındaki uygulamasını araştırmıştır. Sınıflandırma modellerinin geliştirilmesinin genel kapsamı, sınıflandırılacak çeşitlerin sayısını artırmayı, uygun özellikleri çıkarmayı, en uygun özellik kombinasyonunu bulmayı, modelleri eğitmeyi, ince ayar yapmayı ve nihayetinde en iyi modelleri değerlendirmeyi içermektedir.

Bu bağlamda, bu çalışma 124 buğday çeşidini sınıflandırmak için özellik tabanlı ANN ve SVM modellerini geliştirmiştir. Bu çeşit (sınıf) sayısı, buğday çeşitleri ile ilgili yayınlanmış araştırma çalışmalarında en yüksek sayıdır. Bu durum, bu çalışmanın literatürdeki en kapsamlı sınıflandırma modeli olduğunu göstermektedir ve buğday çeşitlerinin tanımlanmasında önemli bir adım olarak değerlendirilebilir.

Görüntü elde edildikten sonra, gerekli ön işleme adımları gerçekleştirilmiştir. Büyük veri seti göz önüne alındığında, kırpma ve yeniden boyutlandırmanın etkisi son derece önemlidir. Çünkü bu işlemler veriyi yüzlerce kat azaltarak yönetilebilir hale getirmektedir. ML modellerini eğitmek için girdi çok önemlidir. Bu nedenle, toplamda 159 HOG ve renk özelliği çıkarılmıştır. Bu özellikler kullanılarak ML modelleri eğitilmiştir. En başarılı model olarak, ANN (%96.8) ve SVM (%98.1) modelleri bulunmuştur. Sonuçlar, fazla sayıda çeşit olmasına rağmen, HOG ve renk özelliklerinin entegrasyonunun başarılı ML modelleri geliştirmek için kullanılabileceğini göstermektedir. Her iki modelde de temel başarının, HOG ve renk özelliklerinin güçlü temsilinden kaynaklandığı ve sınıflandırmadaki ek artışın, PCA'nın orijinal özelliklerden daha iyi dönüşüm ve daha ayırt edici bir temsil yaratmasından kaynaklandığı açıktır. Bu durum, PCA'nın model performansına olan olumlu etkisini vurgulamaktadır. RFC ve KNN modellerinden ise beklenen başarı elde edilememiştir.

Bu çalışmada ayrıca, 124 buğday çeşidini sınıflandırmak için CNN modeli ile MobileNet, DenseNet201 ve InceptionV3 gibi popüler önceden eğitilmiş mimariler kullanılmıştır. Yaklaşık 465.000 görüntüden oluşan büyük bir veri kümesi kullanılarak, modeller 100 epoch boyunca benzer hiperparametrelerle eğitilmiştir. Modellerin verimli bir şekilde eğitilmesi ve aşırı öğrenme riskinin önlenmesi amacıyla model kontrol noktası ve erken durdurma geri çağırımları kullanılmıştır. Modeller arasında, önerilen CNN modeli (%95.40), DenseNet201 (%92.41) ve MobileNet (%90.54) en başarılı model olarak belirlenmiştir.

Sınırlı veri seti ve yalnızca dört çeşit olmasına rağmen, FTIR spektrumları ve ML modellerinin entegrasyonu ile sınıflandırma modelleri geliştirilmiştir. Eğitim ve ince ayar süreçlerinin ardından, yüksek doğruluk sonuçları, ML modellerinin (ANN, SVM ve RFC) buğday çeşitlerini tanımlamada başarıyla uygulanabileceğini göstermiştir.

Büyük veri setlerini işlemek için zaman ve kaynak sınırlamaları nedeniyle,

çalışmanın kapsamı ANN, SVM, RFC ve KNN modellerinin eğitimi ile sınırlı kalmıştır. Sınıflandırma performansını artırmak amacıyla, kullanılan özelliklerin doğasını ele alabilecek potansiyele sahip modeller, (Uzun Kısa Süreli Bellek (Long Short-Term Memory, LSTM) sinir ağları gibi) hesaplama verimliliği ve ölçeklenebilirlik göz önünde bulundurularak farklı eğitim, doğrulama ve test veri seti oranları kullanılarak denenmelidir. Ayrıca, CNN ve öğrenme aktarımının (transfer learning) uygulandığı bölümde daha ileri araştırmalar yapılmalıdır. Önerilen CNN modelinin doğruluğunu daha da artırmak için katman derinliğini artırmak, nöron boyutu, öğrenme oranı, batch boyutu ve dropout oranları gibi hiperparametrelerin ince ayarının yapılarak en uygun yapılandırmanın bulunması gerekmektedir. Öğrenme aktarımı (transfer learning) yaklaşımında ise ek olarak popüler modellerin test edilmesi gerçekleştirilebilir. Bunun yanı sıra, belirli erken katmanların dondurulması ve daha sonraki katmanların özel veri setine uyum sağlamasına izin verilerek sınıflandırma doğruluğunu en üst düzeye çıkarmak için ince ayar yapılabilir.

Ayrıca, FTIR spektroskopi bölümünde, eğitim veri setinin boyutunun artırılması ve buğday çeşitlerinin sayısının artırılması, uygunluğunun daha fazla keşfedilmesi için dikkate alınmalıdır. Baz çizgisi düzeltmesi için bu çalışmada kullanılan yöntemin dışında farklı ön işleme yöntemleri keşfedilmeli ve test edilmelidir. Yine, agronomi ve spektroskopi alanında uzmanların multidisipliner araştırmalar yaparak, araştırmanın gerçek dünya durumlarındaki önemini ve uygulanabilirliğini derinlemesine incelemeleri yapılabilir.

Elde edilen sonuçların uygulama alanında kullanılabilmesi için kullanıcı dostu arayüze sahip bir program yapılmalıdır.

KAYNAKLAR

- Adam, E., Deng, H., Odindi, J., Abdel-Rahman, E. M., and Mutanga, O. (2017). Detecting the early stage of phaeosphaeria leaf spot infestations in maize crop using in situ hyperspectral data and guided regularized random forest algorithm. *Journal of Spectroscopy*. <https://doi.org/10.1155/2017/6961387>.
- Al Sahili, Z., & Awad, M. (2022). The power of transfer learning in agricultural applications: AgriNet. *Frontiers in Plant Science*, 13, 992700.
- Aleixandre-Tudo, J. L., Castelló-Cogollos, L., Aleixandre, J. L., and Aleixandre-Benavent, R. (2020). Bibliometric insights into the spectroscopy research field: A food science and technology case study. *Applied Spectroscopy Reviews*, 55(9-10), 873-906. DOI: 10.1080/05704928.2019.1694936
- Alhawas, N., and Tüfekci, Z. (2022). The effectiveness of transfer learning and fine-tuning approach for automated mango variety classification. *European Journal of Science and Technology*, Special Issue 34, 344-353. <https://doi.org/10.31590/ejosat.1082217>
- Amir, R. M., Anjum, F. M., Khan, M. I., Khan, M. R., Pasha, I., and Nadeem, M. (2013). Application of Fourier transform infrared (FTIR) spectroscopy for the identification of wheat varieties. *Journal of Food Science and Technology*, 50(5), 1018–1023. <https://doi.org/10.1007/s13197-011-0424-y>
- Arif, S., Hussain, R., Ansari, N. M., and Rauf, W. (2023). A novel hybrid feature method for weeds identification in the agri-culture sector. *Res. Agr. Eng*, 69, 132-142.
- Aslam, T., Qadri, S., Saher, N., Nawaz, S. A., Shehzad, M., Qadri, S. F., ... and Khan, D. M. (2021). A spatial model of K-nearest neighbors for classification of cotton (gossypium) varieties based on image segmentation. *International Research Journal of Computer Science*, 5(1), 24-3
- Bakr, M., Abdel-Gaber, S., Nasr, M., and Hazman, M. (2022). DenseNet Based Model for Plant Diseases Diagnosis. *European Journal of Electrical Engineering and Computer Science*, 6(5), 1–9. <https://doi.org/10.24018/ejece.2022.6.5.458>
- Bao, W., Yang, X., Liang, D., Hu, G., and Yang, X. (2021). Lightweight convolutional neural network model for field wheat ear disease identification. *Computers and Electronics in Agriculture*, 189. <https://doi.org/10.1016/j.compag.2021.106367>.
- Barbedo, J. G. A. (2016). A review on the main challenges in automatic plant disease identification based on visible range images. *Biosystems Engineering*, 144, 52-60.
- Benos, L., Tagarakis, A. C., Dolias, G., Berruto, R., Kateris, D., and Bochtis, D. (2021). Machine learning in agriculture: A comprehensive updated review. *Sensors*, 21(11), 3758.
- Breiman, L. Random Forests. *Mach. Learn.* 2001, 45, 5–32
- Cançelik, M., Şahin, Z., Sevinç, M., Küçük, N., and Aydoğdu, M. (2020). The analysis of the recent periods of wheat market in Turkey. *Journal of Engineering and Technology for Industrial Applications*. <https://doi.org/10.5935/jetia.v7i27.734>.
- Ceyhan, M., Kartal, Y., Özkan, K., and Seke, E. (2024). Classification of wheat varieties with image-based deep learning. *Multimedia Tools and Applications*, 83(4), 9597-9619.
- Charytanowicz, M., Kulczycki, P., Kowalski, P. A., Łukasik, S., and Czabak-Garbacz, R. (2018). An evaluation of utilizing geometric features for wheat grain classification using X-ray images. *Computers and Electronics in Agriculture*, 144, 260–268. <https://doi.org/10.1016/j.compag.2017.12.004>.

- Chawla, N. V., Bowyer, K. W., Hall, L. O., and Kegelmeyer, W. P. (2002). SMOTE: synthetic minority over-sampling technique. *Journal of artificial intelligence research*, 16, 321-357.
- Chen, P. Y., Huang, C. C., Lien, C. Y., and Tsai, Y. H. (2013). An efficient hardware implementation of HOG feature extraction for human detection. *IEEE Transactions on Intelligent Transportation Systems*, 15(2), 656-662.
- Chen, P., Li, W. L., Yao, S. J., Ma, C., Zhang, J., Wang, B., Zheng, C. H., Xie, C. J., and Liang, D. (2021). Recognition and counting of wheat mites in wheat fields by a three-step deep learning method. *Neurocomputing*, 437, 21–30. <https://doi.org/10.1016/j.neucom.2020.07.140>
- Chu, H., Zhang, D., Shao, Y., Chang, Z., Guo, Y., and Zhang, N. (2018). Using HOG descriptors and UAV for crop pest monitoring. In 2018 Chinese Automation Congress (CAC) (pp. 1516-1519). IEEE.
- Cinar, I., and Koklu, M. (2022). Identification of rice varieties using machine learning algorithms. *Journal of Agricultural Sciences*, 9-9.
- Craninx, M., Fievez, V., Vlaeminck, B., and De Baets, B. (2008). Artificial neural network models of the rumen fermentation pattern in dairy cattle. *Computers and Electronics in Agriculture*, 60(2), 226–238.
- Çelik, Y., Başaran, E., and Dilay, Y. (2022). Identification of durum wheat grains by using hybrid convolution neural network and deep features. *Signal, Image and Video Processing*, 16(4), 1135–1142. <https://doi.org/10.1007/s11760-021-02094-y>.
- Dalal, N., and Triggs, B. (2005). Histograms of oriented gradients for human detection. In 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05) (Vol. 1, pp. 886-893). IEEE. <https://doi.org/10.1109/CVPR.2005.177>
- Daniya, T., and Vigneshwari, S. (2022). Amalgamation of Rice Plant Disease Classification and Detection with KNN and Butterfly Algorithm. In *Innovations in Computer Science and Engineering: Proceedings of the Ninth ICICSE, 2021* (pp. 127-135). Singapore: Springer Singapore.
- Daugman, J. G. (1988). Complete discrete 2-D Gabor transforms by neural networks for image analysis and compression. *IEEE Transactions on Acoustics, Speech, and Signal Processing*, 36(7), 1169-1179.
- Dazhou, Z., Cheng, W., Binshuang, P., Fuhua, S., Qiong, W., and Chunjiang, Z. (2012). 1313 total citations on Dimensions. *Journal of Nanoelectronics and Optoelectronics*, 7(2), 167-172. American Scientific Publishers. DOI: <https://doi.org/10.1166/jno.2012.1243>.
- De Medeiros Back, H., Junior, E. C. V., Alarcon, O. E., and Pottmaier, D. (2022). Training and evaluating machine learning algorithms for ocean microplastics classification through vibrational spectroscopy. *Chemosphere*, 287, 131903.
- Demšar, J. (2006). Statistical comparisons of classifiers over multiple data sets. *The Journal of Machine learning research*, 7, 1-30.
- Dheer, P., and Singh, V. (2019). Classifying wheat varieties using machine learning model. *Journal of Pharmacognosy and Phytochemistry*, 8(3), 47-49.
- Dixon, J., Braun, H. J., Kosina, P., and Crouch, J. H. (Eds.). (2009). *Wheat facts and futures 2009*. Cimmyt.
- Domingos, P. (2012). A few useful things to know about machine learning. *Communications of the ACM*, 55(10), 78-87. <https://doi.org/10.1145/2347736.2347755>.
- Dong, M., Mu, S., Shi, A., Mu, W., and Sun, W. (2020). Novel method for identifying wheat leaf disease images based on differential amplification convolutional neural network.

- International Journal of Agricultural and Biological Engineering*, 13(4), 205–210.
<https://doi.org/10.25165/j.ijabe.20201304.4826>.
- Dowell E., Maghirang B., Xie F., Lookhart L., Pierce O., Seabourn W., Bean R., Wilson D., Chung K. (2006). Predicting wheat quality characteristics and functionality using near-infrared spectroscopy. *Cereal Chem.* 83:529–536
- Ebrahimi, M. A., Khoshtaghaza, M. H., Minaei, S., and Jamshidi, B. (2017). Vision-based pest detection based on SVM classification method. *Computers and Electronics in Agriculture*, 137, 52–58.
- Elfatimi, E., Eryigit, R., and Elfatimi, L. (2022). Beans leaf diseases classification using MobileNet models. *IEEE Access*, 10, 9471-9482.
<https://doi.org/10.1109/ACCESS.2022.3142817>
- Fabiyi, S. D., Vu, H., Tachtatzis, C., Murray, P., Harle, D., Dao, T. K., ... and Marshall, S. (2020). Varietal classification of rice seeds using RGB and hyperspectral images. *IEEE Access*, 8, 22493-22505.
- Fattahi, S. H., Kazemi, A., Khojastehnazhand, M., Roostaei, M., and Mahmoudi, A. (2024). The classification of Iranian wheat flour varieties using FT-MIR spectroscopy and chemometrics methods. *Expert Systems with Applications*, 239.
<https://doi.org/10.1016/j.eswa.2023.122175>.
- Ferentinos, K. (2018). Deep learning models for plant disease detection and diagnosis. *Computers and Electronics in Agriculture*, 145, 311–318.
- Fernández-Campos, M., Huang, Y. T., Jahanshahi, M. R., Wang, T., Jin, J., Telenko, D. E. P., Góngora-Canul, C., and Cruz, C. D. (2021). Wheat Spike Blast Image Classification Using Deep Convolutional Neural Networks. *Frontiers in Plant Science*, 12.
<https://doi.org/10.3389/fpls.2021.673505>.
- Fletcher, R. S., and Reddy, K. N. (2016). Random forest and leaf multispectral reflectance data to differentiate three soybean varieties from two pigweeds. *Computers and Electronics in Agriculture*, 128, 199-206.
- Fu, X., and Wei, W. (2008). Centralized binary patterns embedded with image Euclidean distance for facial expression recognition. In *Proceedings of the Fourth International Conference on Natural Computation (ICNC)* (pp. 115–119).
- Gao, J., Liu, C., Han, J., Lu, Q., Wang, H., Zhang, J., Bai, X., and Luo, J. (2021). Identification method of wheat cultivars by using a convolutional neural network combined with images of multiple growth periods of wheat. *Symmetry*, 13(11), 2012.
<https://doi.org/10.3390/sym13112012>.
- García, S., and Herrera, F. (2008). An Extension on "Statistical Comparisons of Classifiers over Multiple Data Sets" for all Pairwise Comparisons. *Journal of Machine Learning Research*, 9, 2677-2694.
- Giang, L. T., Thien, T. L. T., and Yen, D. H. (2020). Rapid classification of rice in Northern Vietnam by using FTIR spectroscopy combined with chemometrics methods. *Vietnam Journal of Chemistry*, 58(3), 372–379. <https://doi.org/10.1002/vjch.202000001>.
- Giraldo, P., Benavente, E., Manzano-Agugliaro, F., and Gimenez, E. (2019). Worldwide research trends on wheat and barley: A bibliometric comparative analysis. *Agronomy*, 9(7), 352. <https://doi.org/10.3390/agronomy9070352>.
- Godfray, H. C., Beddington, J. R., Crute, I. R., Haddad, L., Lawrence, D., Muir, J. F., Pretty, J., Robinson, S., Thomas, S. M., and Toulmin, C. (2010). Food security: the challenge of feeding 9 billion people. *Science* (New York, N.Y.), 327(5967), 812–818.
<https://doi.org/10.1126/science.1185383>.

- Gok, S., Severcan M., Goormaghtigh E., Kandemir I., Severcan F. (2015). Differentiation of Anatolian honey samples from different botanical origins by ATR-FTIR spectroscopy using multivariate analysis. *Food Chem.* 170:234–240. doi:10.1016/j.foodchem.2014.08.040.
- Golcuk, A., and Yasar, A. (2023). Classification of bread wheat genotypes by machine learning algorithms. *Journal of Food Composition and Analysis*, 119. <https://doi.org/10.1016/j.jfca.2023.105253>.
- Gomes, J. F. S., and Leta, F. R. (2012). Applications of computer vision techniques in the agriculture and food industry: A review. *European Food Research and Technology*, 235, 989-1000. <https://api.semanticscholar.org/CorpusID:95716096>.
- Gotlieb, C. C., and Kreyszig, H. E. (1990). Texture descriptors based on co-occurrence matrices. *Computer Vision, Graphics, and Image Processing*, 51(1), 70-86. [https://doi.org/10.1016/S0734-189X\(05\)80063-5](https://doi.org/10.1016/S0734-189X(05)80063-5).
- Gouvinhas I., Jose M., De Almeida T., Carvalho M., Barros A. (2015). Discrimination and characterization of extra virgin olive oils from three cultivars in different maturation stages using Fourier transform infrared spectroscopy in tandem with chemometrics. *Food Chem.* 174:226–232. doi:10.1016/j.foodchem.2014.11.037.
- Goyal, L., Sharma, C. M., Singh, A., and Singh, P. K. (2021). Leaf and spike wheat disease detection and classification using an improved deep convolutional architecture. *Informatics in Medicine Unlocked*, 25. <https://doi.org/10.1016/j.imu.2021.100642>.
- Grantina-Ievina, L., Ievina, B., Evelone, V., Berga, S., Kovalcuka, L., Bergspica, I., Jakovele, A., Malisevs, A., Valcina, O., Rodze, I., and Rostoks, N. (2019). Potential risk evaluation for unintended entry of genetically modified plant Propagating material in Europe through import of seeds and animal feed - the experience of Latvia. *GM crops and food*, 10(3), 159–169. <https://doi.org/10.1080/21645698.2019.1638721>.
- Guo, Z., Zhang, L., and Zhang, D. (2010). A completed modeling of local binary pattern operator for texture classification. *IEEE Transactions on Image Processing*, 19(6), 1657-1663. <https://doi.org/10.1109/TIP.2010.2044957>.
- Guresen, E., Kayakutlu, G., and Daim, T. U. (2011). Using artificial neural network models in stock market index prediction. *Expert systems with Applications*, 38(8), 10389-10397. <https://doi.org/10.1016/j.eswa.2011.02.068>.
- Haider, W., Ur Rehman, A., Maqsood, A., and Javed, S. Z. (2020). Crop disease diagnosis using deep learning models. In 2020 *Global Conference on Wireless and Optical Technologies* (GCWOT) (pp. 1-6). <https://doi.org/10.1109/GCWOT49901.2020.9391605>.
- Han J., Kamber M., Pei J. (2012). *Data Mining Concepts and Techniques*. Third Edition. Morgan Kaufmann Publishers. Waltham, MA 02451, USA. ISBN 978-0-12-381479-1.
- Han, L., Tian, J., Huang, Y., He, K., Liang, Y., Hu, X., ... and Huang, D. (2024). Hyperspectral imaging combined with dual-channel deep learning feature fusion model for fast and non-destructive recognition of brew wheat varieties. *Journal of Food Composition and Analysis*, 125, 105785. DOI: 10.1016/j.jfca.2023.105785
- Haralick, R. M., Shanmugam, K., and Dinstein, I. (1973). Texture features for image classification. *IEEE Transactions on Systems, Man, and Cybernetics*, 3(6), 610-621. <https://doi.org/10.1109/TSMC.1973.4309314>.
- Hasan, M. M., Chopin, J. P., Laga, H., and Miklavcic, S. J. (2018). Detection and analysis of wheat spikes using Convolutional Neural Networks. *Plant Methods*, 14(1). <https://doi.org/10.1186/s13007-018-0366-8>.

- Hashimoto, A., and Kameoka, T. (2008). Applications of Infrared Spectroscopy to Biochemical, Food, and Agricultural Processes. *Applied Spectroscopy Reviews*, 43(5), 416-451. <https://doi.org/10.1080/05704920802108131>
- Hayit, T., Endes, A., and Hayit, F. (2024). KNN-based approach for the classification of fusarium wilt disease in chickpea based on color and texture features. *European Journal of Plant Pathology*, 168(4), 665-681. <https://doi.org/10.1007/s10658-023-02791-z>
- Hoang, V. T., Van Hoai, D. P., Surinwarangkoon, T., Duong, H. T., and Meethongjan, K. (2020). A comparative study of rice variety classification based on deep learning and hand-crafted features. *ECTI Transactions on Computer and Information Technology (ECTI-CIT)*, 14(1), 1-10. <https://doi.org/10.37936/ecti-cit.2020141.204170>
- Hossain, E., Hossain, M. F., and Rahaman, M. A. (2019). A color and texture based approach for the detection and classification of plant leaf disease using KNN classifier. In 2019 *International Conference on Electrical, Computer and Communication Engineering (ECCE)* (pp. 1-6). <https://api.semanticscholar.org/CorpusID:96432008>.
- Howard, A. G., Zhu, M., Chen, B., Kalenichenko, D., Wang, W., Weyand, T., Andreetto, M., and Adam, H. (2017). MobileNets: Efficient convolutional neural networks for mobile vision applications. *arXiv*, arXiv:1704.04861.
- Huang, G., Liu, Z., Van Der Maaten, L., and Weinberger, K. Q. (2017). Densely connected convolutional networks. In 2017 *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)* (pp. 2261-2269). IEEE. <https://doi.org/10.1109/CVPR.2017.243>
- Huang, J., Kumar, S. R., Mitra, M., Zhu, W.-J., and Zabih, R. (1997). Image indexing using color correlograms. In *Proceedings of IEEE Computer Society Conference on Computer Vision and Pattern Recognition* (pp. 762-768). <https://doi.org/10.1109/CVPR.1997.609412>.
- Işık, Ş., Özkan, K., Demirez, D. Z., and Seke, E. (2022). Consensus rule for wheat cultivar classification on VL, VNIR and SWIR imaging. *IET Image Processing*, 16(11), 2834–2844. <https://doi.org/10.1049/ipr2.12206>.
- Jena, L., Behera, S. K., and Sethy, P. K. (2021). Identification of wheat grain using geometrical feature and machine learning. 2021 2nd *International Conference for Emerging Technology*, INCET 2021. <https://doi.org/10.1109/INCET51464.2021.9456281>.
- Jiang, Z., Dong, Z., Jiang, W., and Yang, Y. (2021). Recognition of rice leaf diseases and wheat leaf diseases based on multi-task deep transfer learning. *Computers and Electronics in Agriculture*, 186. <https://doi.org/10.1016/j.compag.2021.106184>.
- Jin, S., Zhang, F., Zheng, Y., Zhou, L., Zuo, X., Zhang, Z., ... and Pan, X. (2023). CSKNN: Cost-sensitive K-Nearest Neighbor using hyperspectral imaging for identification of wheat varieties. *Computers and Electrical Engineering*, 111, 108896. <https://doi.org/10.1016/j.compeleceng.2023.108896>
- Johann, A. L., de Araújo, A. G., Delalibera, H. C., and Hirakawa, A. R. (2016). Soil moisture modeling based on stochastic behavior of forces on a no-till chisel opener. *Computers and Electronics in Agriculture*, 121, 420-428. <https://doi.org/10.1016/j.compag.2015.12.020>
- Jolliffe, I.T. (2002). *Principal Component Analysis*. Springer Series in Statistics. Verlag New York, Inc.: Springer. ISBN 0-387-95442-2.
- Jurj, S., Opritoiu, F., and Vladutiu, M. (2020). Deep Learning-Based Computer Vision Application with Multiple Built-In Data Science-Oriented Capabilities. *Proceedings of the 21st EANN (Engineering Applications of Neural Networks) 2020 Conference*, 2, 47 - 69. https://doi.org/10.1007/978-3-030-48791-1_4.

- Kasinathan, T., and Uyyala, S. R. (2021). Machine learning ensemble with image processing for pest identification and classification in field crops. *Neural Computing and Applications*, 33(13), 7491-7504. <https://doi.org/10.1007/s00521-020-05497-z>
- Kendler, S., Aharoni, R., Young, S., Sela, H., Kis-Papo, T., Fahima, T., and Fishbain, B. (2022). Detection of crop diseases using enhanced variability imagery data and convolutional neural networks. *Computers and Electronics in Agriculture*, 193. <https://doi.org/10.1016/j.compag.2022.106732>.
- Keser, M. (2022). Variety Development of Wheat in Turkey and the Impact of Regulations. *Ekin Journal of Crop Breeding and Genetics*. 8(1):1-16.
- Khalid S., Khalil T., and Nasreen S. (2014). A survey of feature selection and feature extraction techniques in machine learning. *Science and Information Conference*, pp. 372-378.
- Khojastehnazhand, M., and Roostaei, M. (2022). Classification of seven Iranian wheat varieties using texture features. *Expert Systems with Applications*, 199. <https://doi.org/10.1016/j.eswa.2022.117014>.
- Khoshroo, A., Arefi, A., Masoumiasl, A., and Jowkar, G. H. (2014). Classification of wheat cultivars using image processing and artificial neural networks. *Agricultural Communications*, 2(1), 17-22. <https://api.semanticscholar.org/CorpusID:110279993>
- Koçer, H., Mehmet, and Konaşkan, Ö. (2020). Structure of wheat and corn farming: A survey on Amik Plain farmers. *Turkish Journal of Agriculture - Food Science and Technology*, 8(8), 1612-1617. <https://doi.org/10.24925/turjaf.v8i8.1612-1617.2935>
- Kukreja, V., and Kumar, D. (2021). Automatic Classification of Wheat Rust Diseases Using Deep Convolutional Neural Networks. 2021 9th International Conference on Reliability, Infocom Technologies and Optimization (Trends and Future Directions), ICRITO 2021. <https://doi.org/10.1109/ICRITO51393.2021.9596133>.
- Kumar, D., and Kukreja, V. (2021). An Instance Segmentation Approach for Wheat Yellow Rust Disease Recognition. 2021 International Conference on Decision Aid Sciences and Application, DASA 2021, 926–931. <https://doi.org/10.1109/DASA53625.2021.9682257>.
- Kumar, G., and Bhatia, P. K. (2014). A detailed review of feature extraction in image processing systems. In 2014 Fourth International Conference on Advanced Computing and Communication Technologies (pp. 5-12). <https://doi.org/10.1109/ACCT.2014.74>.
- Laabassi, K., Belarbi, M. A., Mahmoudi, S., Mahmoudi, S. A., and Ferhat, K. (2021). Wheat varieties identification based on a deep learning approach. *Journal of the Saudi Society of Agricultural Sciences*, 20(5), 281–289. <https://doi.org/10.1016/j.jssas.2021.02.008>.
- Lemaître, G., Nogueira, F., Aridas, C.K. (2017). Imbalanced-learn: a python toolbox to tackle the curse of imbalanced datasets in machine learning. *Journal of Machine Learning Research*. 18,1-5.
- Li, K., Lin, J., Liu, J., and Zhao, Y. (2020). Using deep learning for image-based different degrees of Ginkgo leaf disease classification. *Information*, 11(2), 95. <https://doi.org/10.3390/info11020095>.
- Liakos, K. G., Busato, P., Moshou, D., Pearson, S., and Bochtis, D. (2018). Machine learning in agriculture: A review. *Sensors*, 18(8), 2674. <https://doi.org/10.3390/s18082674>.
- Liao, W.-H. (2010). Region description using extended local ternary patterns. In 2010 20th International Conference on Pattern Recognition (pp. 1003-1006). <https://doi.org/10.1109/ICPR.2010.251>.
- Limm, W., Karunathilaka, S. R., and Mossoba, M. M. (2023). Fourier Transform Infrared Spectroscopy and Chemometrics for the Rapid Screening of Economically Motivated

- Adulteration of Honey Spiked With Corn or Rice Syrup. *Journal of Food Protection*, 86(4), 100054. <https://doi.org/10.1016/j.jfp.2023.100054>
- Lingwal, S., Bhatia, K. K., and Tomer, M. S. (2021). Image-based wheat grain classification using convolutional neural network. *Multimedia Tools and Applications*, 80(28–29), 35441–35465. <https://doi.org/10.1007/s11042-020-10174-3>.
- Lopez-Arevalo, I., Aldana-Bobadilla, E., Molina-Villegas, A., Galeana-Zapién, H., Muñiz-Sanchez, V., and Gausin-Valle, S. (2020). A memory-efficient encoding method for processing mixed-type data on machine learning. *Entropy*, 22(12), 1391. <https://doi.org/10.3390/e22121391>.
- Lu, T., Han, B., Chen, L., Yu, F., and Xue, C. (2021). A generic intelligent tomato classification system for practical applications using DenseNet-201 with transfer learning. *Scientific Reports*, 11, 15824. <https://doi.org/10.1038/s41598-021-95218-w>
- Madhavan, J., Salim, M., Durairaj, U., and Kotteeswaran, R. (2023). Wheat seed classification using neural network pattern recognizer. *Materials Today: Proceedings*, 81, 341–345. <https://doi.org/10.1016/j.matpr.2021.03.226>.
- Mahajan, S., Das, A., and Sardana, H. K. (2015). Image acquisition techniques for assessment of legume quality. *Trends in Food Science and Technology*, 42(2), 116–133. <https://doi.org/10.1016/j.tifs.2015.01.001>
- Malvade, N. N., Yakkundimath, R., Saunshi, G. B., & Elemmi, M. C. (2023). Paddy variety identification from field crop images using deep learning techniques. *International Journal of Computational Vision and Robotics*, 13(4), 405–419.
- Maqsood, M. H., Mumtaz, R., Haq, I. U., Shafi, U., Zaidi, S. M. H., and Hafeez, M. (2021). Super resolution generative adversarial network (Srgans) for wheat stripe rust classification. *Sensors*, 21(23). <https://doi.org/10.3390/s21237903>.
- Mehmood, I., Sajjad, M., Ejaz, W., and Baik, S. W. (2015). Saliency-directed prioritization of visual data in wireless surveillance networks. *Information Fusion*, 24, 16–30. <https://doi.org/10.1016/j.inffus.2014.07.002>
- Mekha, P., and Teeyasuksaet, N. (2021). Image classification of rice leaf diseases using random forest algorithm. In *2021 Joint International Conference on Digital Arts, Media and Technology with ECTI Northern Section Conference on Electrical, Electronics, Computer and Telecommunication Engineering* (pp. 165–169). Cha-am, Thailand. <https://doi.org/10.1109/ECTIDAMTNCN51128.2021.9425696>.
- Mercier, G., and Lennon, M. (2003). Support vector machines for hyperspectral image classification with spectral-based kernels. In *IGARSS 2003. 2003 IEEE International Geoscience and Remote Sensing Symposium. Proceedings* (pp. 288–290, vol. 1). Toulouse, France. <https://doi.org/10.1109/IGARSS.2003.1293752>.
- Meshram, V., Patil, K., Meshram, V., Hanchate, D., and Ramkteke, S. D. (2021). Machine learning in agriculture domain: A state-of-art survey. *Artificial Intelligence in the Life Sciences*, 1, 100010. <https://doi.org/10.1016/j.ailsci.2021.100010>.
- Miah, M. S., Tasnuva, T., Islam, M., Keya, M., Rahman, M. R., and Hossain, S. A. (2021). An advanced method of identification fresh and rotten fruits using different convolutional neural networks. In *2021 12th International Conference on Computing Communication and Networking Technologies (ICCCNT)* (pp. 1–7). IEEE. <https://doi.org/10.1109/ICCCNT51525.2021.9580117>
- Miaw, C. S. W., Assis, C., Silva, A. R. C. S., Cunha, M. L., Sena, M. M., and de Souza, S. V. C. (2018). Determination of main fruits in adulterated nectars by ATR-FTIR spectroscopy combined with multivariate calibration and variable selection methods. *Food chemistry*, 254, 272–280. <https://doi.org/10.1016/j.foodchem.2018.02.015>.

- Mochida, K., Koda, S., Inoue, K., Hirayama, T., Tanaka, S., Nishii, R., and Melgani, F. (2019). Computer vision-based phenotyping for improvement of plant productivity: A machine learning perspective. *GigaScience*, 8(1), giy153. <https://doi.org/10.1093/gigascience/giy153>.
- Murru, C., Chimeno-Trinchet, C., Díaz-García, M. E., Badía-Laiño, R., and Fernández-González, A. (2019). Artificial Neural Network and Attenuated Total Reflectance-Fourier Transform Infrared Spectroscopy to identify the chemical variables related to ripeness and variety classification of grapes for Protected. Designation of Origin wine production. *Computers and Electronics in Agriculture*, 164. <https://doi.org/10.1016/j.compag.2019.104922>.
- Mutlag, W. K., Ali, S. K., Aydam, Z. M., and Taher, B. H. (2020). Feature extraction methods: A review. *Journal of Physics: Conference Series*, 1591(1), 012028. <https://doi.org/10.1088/1742-6596/1591/1/012028>.
- Mylonas, I., Stavrakoudis, D., Katsantonis, D., and Korpetis, E. (2020). Chapter 1 - Better farming practices to combat climate change. In M. Ozturk and A. Gul (Eds.), *Climate Change and Food Security with Emphasis on Wheat* (pp. 1-29). Academic Press. ISBN 9780128195277.
- Nahvi, B., Habibi, J., Mohammadi, K., Shamshirband, S., and Al Razgan, O. S. (2016). Using self-adaptive evolutionary algorithm to improve the performance of an extreme learning machine for estimating soil temperature. *Computers and Electronics in Agriculture*, 124, 150-160. <https://doi.org/10.1016/j.compag.2016.03.025>
- Nandhini, S., & Ashokkumar, K. (2022). An automatic plant leaf disease identification using DenseNet-121 architecture with a mutation-based henry gas solubility optimization algorithm. *Neural Computing & Applications*, 34(11), 5513–5534. <https://doi.org/10.1007/s00521-021-06714-z>
- Ochango, V. M., Wambugu, G. M., and Ndia, J. G. (2022). Feature extraction using histogram of oriented gradients for image classification in maize leaf diseases. *International Journal of Computer and Information Technology*, 11(4). <https://doi.org/10.24203/ijcit.v11i4.244>
- Olgun, M., Onarcan, A. O., Özkan, K., Işık, Ş., Sezer, O., Özgişi, K., Ayter, N. G., Başçiftçi, Z. B., Ardiç, M., and Koyuncu, O. (2016). Wheat grain classification by using dense SIFT features with SVM classifier. *Computers and Electronics in Agriculture*, 122(C), 185–190. <https://doi.org/10.1016/j.compag.2016.01.033>.
- Oshiro, T. M., Perez, P. S., and Baranauskas, J. A. (2012). How many trees in a random forest? In *IAPR International Conference on Machine Learning and Data Mining in Pattern Recognition*. Retrieved from <https://api.semanticscholar.org/CorpusID:12211379>.
- Owino, A. (2023). Challenges of computer vision adoption in the Kenyan agricultural sector and how to solve them: A general perspective. *Advances in Agriculture*. <https://doi.org/10.1155/2023/1530629>.
- Ozturk, S., Bowler, A., Rady, A., and Watson, N. J. (2023). Near-infrared spectroscopy and machine learning for classification of food powders during a continuous process. *Journal of Food Engineering*, 341, 111339. <https://doi.org/10.1016/j.jfoodeng.2022.111339>
- Özkan, K., Seke, E., and Işık, Ş. (2021). Wheat kernels classification using visible-near infrared camera based on deep learning. *Pamukkale University Journal of Engineering Sciences*, 27(5), 618–626. <https://doi.org/10.5505/pajes.2020.80774>
- Pantazi, X.-E., Moshou, D., Alexandridis, T. K., Whetton, R. L., and Mouazen, A. M. (2016). Wheat yield prediction using machine learning and advanced sensing techniques.

- Computers and Electronics in Agriculture*, 121, 57–65. <https://doi.org/10.1016/j.compag.2015.11.018>
- Pass, G., and Zabih, R. (1996). Histogram refinement for content-based image retrieval. In *Proceedings Third IEEE Workshop on Applications of Computer Vision (WACV'96)* (pp. 96-102). <https://doi.org/10.1109/ACV.1996.572008>.
- Patil, A. P., and Deka, P. C. (2016). An extreme learning machine approach for modeling evapotranspiration using extrinsic inputs. *Computers and Electronics in Agriculture*, 121, 385-392. <https://doi.org/10.1016/j.compag.2016.01.016>
- Patil, A. P., and Lad, K. (2020). Chili plant leaf disease detection using SVM and KNN classification. In *Proceedings*. Retrieved from <https://api.semanticscholar.org/CorpusID:224925723>.
- Patrício, D.I., and Rieder, R. (2018). Computer vision and artificial intelligence in precision agriculture for grain crops: A systematic review. *Computers and Electronics in Agriculture*, 153, 69-81. <https://doi.org/10.1016/j.compag.2018.08.001>
- Pazoki, A., and Pazoki, Z. (2011). Classification system for rain-fed wheat grain cultivars using artificial neural network. *African Journal of Biotechnology*, 10(41), 8031-8038. DOI: 10.5897/AJB11.488.
- Peng, J., Sun, D., and Nevo, E. (2011). Wild emmer wheat, 'Triticum dicoccoides', occupies a pivotal position in wheat domestication process. *Australian Journal of Crop Science*, 5, 1127-1143. <https://api.semanticscholar.org/CorpusID:9938610>.
- Persoon, E., and Fu, K.-S. (1977). Shape discrimination using Fourier descriptors. *IEEE Transactions on Systems, Man, and Cybernetics*, 7(3), 170-179. <https://doi.org/10.1109/TSMC.1977.4309681>.
- Porep, J. U., Kammerer, D. R., and Carle, R. (2015). On-line application of near infrared (NIR) spectroscopy in food production. *Trends in Food Science and Technology*, 46(2), 211-230. <https://doi.org/10.1016/j.tifs.2015.10.002>.
- Pourreza, A., Pourreza, H., Abbaspour-Fard, M.-H., and Sadrnia, H. (2012). Identification of nine Iranian wheat seed varieties by textural analysis with image processing. *Computers and Electronics in Agriculture*, 83, 102-108. ISSN 0168-1699. <https://doi.org/10.1016/j.compag.2012.02.005>.
- Qi, H., Shen, C., Chen, G., Zhang, J., Li, H., and Zhang, C. (2023). Rapid and non-destructive determination of soluble solid content of crown pear by visible/near-infrared spectroscopy with deep learning regression. *Journal of Food Composition and Analysis*, 123, 105585. <https://doi.org/10.1016/j.jfca.2023.105585>
- Que, H., Zhao, X., Sun, X., Zhu, Q., and Huang, M. (2023). Identification of wheat kernel varieties based on hyperspectral imaging technology and grouped convolutional neural network with feature intervals. *Infrared Physics and Technology*, 131, 104653. <https://doi.org/10.21203/rs.3.rs-1951327/v1>
- Reale, S., Biancolillo, A., Foschi, M., Di Donato, F., Di Censo, E., and D'Archivio, A. A. (2023). Geographical discrimination of Italian carrot (*Daucus carota* L.) varieties: A comparison between ATR FT-IR fingerprinting and HS-SPME/GC-MS volatile profiling. *Food Control*, 146, 109508. <https://doi.org/10.1016/j.foodcont.2022.109508>
- Reis, N., Botelho, B. G., Franca, A. S., and Oliveira, L. S. (2017). Simultaneous detection of multiple adulterants in ground roasted coffee by ATR-FTIR spectroscopy and data fusion. *Food Analytical Methods*, 10(8), 2700-2709. <https://doi.org/10.1007/s12161-017-0832-3>.
- Ribeiro, M. C. S., Senesi, G. S., Cabral, J. S., Cena, C., Marangoni, B. S., Kiefer, C., and Nicolodelli, G. (2020). Evaluation of rice varieties using LIBS and FTIR techniques

- associated with PCA and machine learning algorithms. *Applied optics*, 59(32), 10043–10048. <https://doi.org/10.1364/AO.409029>.
- Ricci, A., Parpiniello, G., Laghi, L., Lambri, M., and Versari, A. (2013). Application of infrared spectroscopy to grape and wine analysis. In *Proceedings*. Retrieved from <https://api.semanticscholar.org/CorpusID:92817051>.
- Rui, Y., She, A. C., and Huang, T. S. (1996). Modified Fourier descriptors for shape representation - A practical approach. In *Proceedings*. Retrieved from <https://api.semanticscholar.org/CorpusID:2553533>.
- Sabancı, K., and Akkaya, M. (2016). Classification of Different Wheat Varieties by Using Data Mining Algorithms. *International Journal of Intelligent Systems and Applications in Engineering*, 4(2), 40. <https://doi.org/10.18201/ijisae.62843>.
- Sabancı, K., Aslan, M. F., and Durdu, A. (2020). Bread and durum wheat classification using wavelet-based image fusion. *Journal of the Science of Food and Agriculture*, 100(15), 5577–5585. <https://doi.org/10.1002/jsfa.10610>.
- Sabancı, K., Kayabasi, A., and Toktas, A. (2017). Computer vision-based method for classification of wheat grains using artificial neural network. *Journal of the Science of Food and Agriculture*, 97(8), 2588–2593. <https://doi.org/10.1002/jsfa.8080>.
- Salim, F., Saeed, F., Basurra, S., Qasem, S. N., and Al-Hadhrami, T. (2023). DenseNet-201 and Xception pre-trained deep learning models for fruit recognition. *Electronics*, 12, 3132. <https://doi.org/10.3390/electronics12143132>
- Sandhu, K. S., Lozada, D. N., Zhang, Z., Pumphrey, M. O., and Carter, A. H. (2021). Deep Learning for Predicting Complex Traits in Spring Wheat Breeding Program. *Frontiers in Plant Science*, 11. <https://doi.org/10.3389/fpls.2020.613325>.
- Schibeci, D. A., Potter, R., Wathen-Dunn, K., Jones, M., and Bellgard, W. (2001). Applying artificial neural networks to the classification of wheat varieties processed via MALDI-TOF mass spectrometry. In *Advances in Neural Networks and Applications* (pp. 262-267).
- Shah, S. R., Qadri, S., Bibi, H., Shah, S. M. W., Sharif, M. I., & Marinello, F. (2023). Comparing inception V3, VGG 16, VGG 19, CNN, and ResNet 50: a case study on early detection of a rice disease. *Agronomy*, 13(6), 1633.
- Shamshiri, R. R., Weltzien, C., Hameed, I. A., Yule, I. J., Grift, T. E., Balasundram, S. K., Pitonakova, L., Ahmad, D., and Chowdhary, G. (2018). Research and development in agricultural robotics: A perspective of digital farming. *International Journal of Agricultural and Biological Engineering*, 11(4), 1-14. DOI: 10.25165/ijabe.20181104.4278
- Sharif, M.U., Khan, M.A., Iqbal, Z., Azam, M., Lali, M.I., and Javed, M.Y. (2018). Detection and classification of citrus diseases in agriculture based on optimized weighted segmentation and feature selection. *Computers and Electronics in Agriculture*, 150, 220-234. <https://doi.org/10.1016/j.compag.2018.04.023>
- Shewry, P. R., and Hey, S. J. (2015). The contribution of wheat to human diet and health. *Food and energy security*, 4(3), 178–202. <https://doi.org/10.1002/fes3.64>.
- Smola, A., and Vishwanathan, S. V. N. (2008). Introduction to machine learning. Cambridge University, UK, 32(34).
- Sonmez, M. E., Sabancı, K., and Aydin, N. (2024). Convolutional neural network-support vector machine-based approach for identification of wheat hybrids. *European Food Research and Technology*, 250(7), 1353-1362. <https://doi.org/10.1007/s00217-024-04473-4>.

- Srinuttrakul, W., Mihailova, A., Islam, M. D., Liebisch, B., Maxwell, F., Kelly, S. D., and Cannavan, A. (2021). Geographical differentiation of hom mali rice cultivated in different regions of thailand using ftir-atr and nir spectroscopy. *Foods*, 10(8). <https://doi.org/10.3390/foods10081951>.
- Stricker, M. A., and Orengo, M. (1995). Similarity of color images. In *Proc. SPIE 2420, Storage and Retrieval for Image and Video Databases III*. Retrieved from <https://doi.org/10.1117/12.205308>.
- Su, T., Min, S., Shi, A., Cao, Z., and Dong, M. (2019). A CNN-LSVM model for imbalanced images identification of wheat leaf. *Neural Network World*, 29(5), 345-361. <https://doi.org/10.14311/NNW.2019.29.021>.
- Su, W. H., Zhang, J., Yang, C., Page, R., Szinyei, T., Hirsch, C. D., and Steffenson, B. J. (2021). Automatic evaluation of wheat resistance to fusarium head blight using dual mask-rcnn deep learning frameworks in computer vision. *Remote Sensing*, 13(1), 1–20. <https://doi.org/10.3390/rs13010026>.
- Su, Y., Xu, H., and Yan, L. (2017). Support vector machine-based open crop model (SBOCM): Case of rice production in China. *Saudi Journal of Biological Sciences*, 24(3), 537–547. <https://doi.org/10.1016/j.sjbs.2017.01.024>
- Su, Y., Xu, H., and Yan, L. (2017). Support vector machine-based open crop model (SBOCM): Case of rice production in China. *Saudi Journal of Biological Sciences*, 24(3), 537–547. <https://api.semanticscholar.org/CorpusID:24069582>
- Subeesh, A., Bhole, S., Singh, K., Chandel, N. S., Rajwade, Y. A., Rao, K. V. R., ... & Jat, D. (2022). Deep convolutional neural network models for weed detection in polyhouse grown bell peppers. *Artificial Intelligence in Agriculture*, 6, 47-54.
- Suematsu, N., Ishida, Y., Hayashi, A., and Kanbara, T. (2002). Region-based image retrieval using wavelet transform. In *Proceedings*. Retrieved from <https://api.semanticscholar.org/CorpusID:15037638>.
- Szegedy, C., Liu, W., Jia, Y., Sermanet, P., Reed, S., Anguelov, D., Erhan, D., Vanhoucke, V., & Rabinovich, A. (2015). Going deeper with convolutions. In *2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)* (pp. 1-9). IEEE. <https://doi.org/10.1109/CVPR.2015.7298594>
- Tamura, H., Mori, S., and Yamawaki, T. (1978). Textural features corresponding to visual perception. *IEEE Transactions on Systems, Man, and Cybernetics*, 8(6), 460-473. <https://doi.org/10.1109/TSMC.1978.4309999>
- Taner, A., Mengstu, M. T., Selvi, K. Ç., Duran, H., Gür, İ., and Ungureanu, N. (2024). Apple varieties classification using deep features and machine learning. *Agriculture*, 14(252). <https://doi.org/10.3390/agriculture14020252>.
- Taner, A., Mengstu, M. T., Selvi, K. Ç., Duran, H., Kabaş, Ö., Gür, İ., Karaköse, T., and Gheorghiiță, N.-E. (2023). Multiclass apple varieties classification using machine learning with histogram of oriented gradient and color moments. *Applied Sciences*, 13(13), 7682. <https://doi.org/10.3390/app13137682>.
- Taner, A., Öztekin, Y. B., and Duran, H. (2021). Performance analysis of deep learning CNN models for variety classification in hazelnut. *Sustainability*, 13(12), 6527. <https://doi.org/10.3390/su13126527>
- Taner, A., Öztekin, Y., Tekgüler, A., Sauk, H., and Duran, H. (2018). Classification of Varieties of Grain Species by Artificial Neural Networks. *Agronomy*, 8(7), 123. MDPI AG. Retrieved from <http://dx.doi.org/10.3390/agronomy8070123>.
- Teja, K. U. V. R., Reddy, B. P. V., Kesara, L. R., Kowshik, K. D. P., & Panchaparvala, L. A. (2021, October). Transfer Learning based Rice Leaf Disease Classification with

- Inception-V3. In 2021 *International Conference on Smart Generation Computing, Communication and Networking (SMART GENCON)* (pp. 1-6). IEEE.
- Tetila, E. C., Machado, B. B., Astolfi, G., de Souza Belete, N. A., Amorim, W. P., Roel, A. R., & Pistori, H. (2020). Detection and classification of soybean pests using deep learning with UAV images. *Computers and Electronics in Agriculture*, 179, 105836.
- Tian, H., Wang, T., Liu, Y., Qiao, X., and Li, Y. (2020). Computer vision technology in agricultural automation—A review. *Information Processing in Agriculture*, 7(1), 1-19. <https://api.semanticscholar.org/CorpusID:204167355>
- Tiwari, V., Joshi, R. C., & Dutta, M. K. (2021). Dense convolutional neural networks based multiclass plant disease detection and classification using leaf images. *Ecological Informatics*, 63, 101289. <https://doi.org/10.1016/j.ecoinf.2021.101289>
- Trawiński, B., Smętek, M., Telec, Z. and Lasota, T. (2012). Nonparametric statistical analysis for multiple comparison of machine learning regression algorithms. *International Journal of Applied Mathematics and Computer Science*, 22(4) 867-881. <https://doi.org/10.2478/v10006-012-0064-z>
- TÜİK (2023). Türkiye İstatistik Kurumu. <https://data.tuik.gov.tr/Bulten/Index?p=Bitkisel-Uretim-Istatistikleri-2021-37249>.
- Unuvar, A., Boyaci, I. H., and Koksel, H. (2021). A novel approach for rapid discrimination of common and durum wheat flours using spectroscopic analyses combined with chemometrics. *Journal of Cereal Science*, 100, 103269. <https://doi.org/10.1016/j.jcs.2021.103269>
- Upadhyay, S. K. (2022, May). Deep transfer learning-based rice leaves disease diagnosis and classification model using inceptionv3. In 2022 *international conference on computational intelligence and sustainable engineering solutions (CISES)* (pp. 493-499). IEEE.
- Utku, H., and Kijksel, H. (1998). Use of Statistical Filters in the Classification of Wheats by Image Analysis. *Journal of Food Engineering* (Vol. 36). 10.1016/s0260-8774(98)00072-7
- Vapnik, V. (1995). Support vector machine. *Mach. Learn.* 1995, 20, 273–297.
- Vázquez-Arellano, M., Griepentrog, H. W., Reiser, D., and Paraforos, D. S. (2016). 3-D imaging systems for agricultural applications—a review. *Sensors*, 16(5), 618. <https://doi.org/10.3390/s16050618>
- Velesaca, H. O., Suárez, P. L., Mira, R., and Sappa, A. D. (2021). Computer vision-based food grain classification: A comprehensive survey. *Computers and Electronics in Agriculture*, 187. <https://doi.org/10.1016/j.compag.2021.106287>.
- Venkatesh, Y., Nagaraju, Hegde, S. U., and Stalin, S. (2021). Fine-tuned MobileNet classifier for classification of strawberry and cherry fruit types. In 2021 *International Conference on Computer Communication and Informatics (ICCCI)* (pp. 1-8). IEEE. <https://doi.org/10.1109/ICCCI50826.2021.9402444>
- Verma, R., & Verma, A.K. (2022). Fruit classification using deep convolutional neural network and transfer learning. In V.E. Balas, G.R. Sinha, B. Agarwal, T.K. Sharma, P. Dadheech, & M. Mahrishi (Eds.), *Emerging Technologies in Computer Engineering: Cognitive Computing and Intelligent IoT. ICETCE 2022. Communications in Computer and Information Science*, Vol. 1591. Springer, Cham. https://doi.org/10.1007/978-3-031-07012-9_26
- Vieira, S., Pinaya, W. H. L., and Mechelli, A. (2020). Main concepts in machine learning. In *Machine learning* (pp. 21-44). Academic Press. <https://doi.org/10.1016/B978-0-12-815739-8.00002-X>

- Vithu, P., and Moses, J. A. (2016). Machine vision system for food grain quality evaluation: A review. *Trends in food science and technology*, 56, 13-20. <https://doi.org/10.1016/j.tifs.2016.07.011>
- Vlasov, A. v., and Fadeev, A. S. (2017). A machine learning approach for grain crop's seed classification in purifying separation. *Journal of Physics: Conference Series*, 803(1). <https://doi.org/10.1088/1742-6596/803/1/012177>.
- Wang, Z., Huang, W., Li, J., Liu, S., and Fan, S. (2023). Assessment of protein content and insect infestation of maize seeds based on on-line near-infrared spectroscopy and machine learning. *Computers and Electronics in Agriculture*, 211, 107969. <https://doi.org/10.1016/j.compag.2023.107969>.
- Wójtowicz, A., Piekarczyk, J., Czernecki, B., and Ratajkiewicz, H. (2021). A random forest model for the classification of wheat and rye leaf rust symptoms based on pure spectra at leaf scale. *Journal of Photochemistry and Photobiology B: Biology*, 223, 112278. <https://doi.org/10.1016/j.jphotobiol.2021.112278>
- Xu, P., Yang, R., Zeng, T., Zhang, J., Zhang, Y., and Tan, Q. (2021). Varietal classification of maize seeds using computer vision and machine learning techniques. *Journal of Food Process Engineering*, 44(11), e13846. <https://doi.org/10.1111/jfpe.13846>
- Yasar, A. (2024). Analysis of selected deep features with CNN-SVM-based for bread wheat seed classification. *European Food Research and Technology*, 250, 1551–1561. <https://doi.org/10.1007/s00217-024-04488-x>
- Yasar, A., Golcuk, A., and Sari, O. F. (2024). Classification of bread wheat varieties with a combination of deep learning approach. *European Food Research and Technology*, 250(1), 181-189. <https://doi.org/10.1007/s00217-023-04375-x>
- Yasar, A., Kaya, E., and Saritas, I. (2016). Classification of Wheat Types by Artificial Neural Network. *International Journal of Intelligent Systems and Applications in Engineering*, 4(1), 12–15. <https://doi.org/10.18201/ijisae.64198DOI>.
- Yu, H., Erasmus, S. W., Wang, Q., Liu, H., and van Ruth, S. M. (2023). Rapid classification of peanut varieties for their processing into peanut butters based on near-infrared spectroscopy combined with machine learning. *Journal of Food Composition and Analysis*, 120, 105348. <https://doi.org/10.1016/j.jfca.2023.105348>
- Zapotoczny, P. (2011). Discrimination of wheat grain varieties using image analysis and neural networks. Part I. Single kernel texture. *Journal of Cereal Science*, 54(1), 60–68. <https://doi.org/10.1016/j.jcs.2011.02.012>.
- Zareef, M., Chen, Q., Hassan, M.M., Ali, M., and Yu, Y. (2020). An Overview on the Applications of Typical Non-linear Algorithms Coupled with NIR Spectroscopy in Food Analysis. *Food Engineering Reviews*, 12, 173–190. <https://doi.org/10.1007/s12393-020-09210-7>
- Zawbaa, H. M., Hazman, M., Abbass, M., and Hassanien, A. E. (2014, December). Automatic fruit classification using random forest algorithm. In *2014 14th international conference on hybrid intelligent systems* (pp. 164-168). <https://api.semanticscholar.org/CorpusID:16373395>
- Zhang, B., Gao, Y., Zhao, S., and Liu, J. (2010). Local derivative pattern versus local binary pattern: Face recognition with high-order local pattern descriptor. *IEEE Transactions on Image Processing*, 19, 533–544. <https://doi.org/10.1109/TIP.2009.2035882>
- Zhang, D., Gu, C., Wang, Z., Zhou, X., and Li, W. (2021). Evaluating the efficacy of fungicides for wheat scab control by combined image processing technologies. *Biosystems Engineering*, 211, 230–246. <https://doi.org/10.1016/j.biosystemseng.2021.09.008>.

- Zhang, F., Tang, X., and Li, L. (2022). Origins of Baseline Drift and Distortion in Fourier Transform Spectra. *Molecules*, 27(13), 4287. <https://doi.org/10.3390/molecules27134287>.
- Zhang, M., Li, C., and Yang, F. (2017). Classification of foreign matter embedded inside cotton lint using short wave infrared (SWIR) hyperspectral transmittance imaging. *Computers and Electronics in Agriculture*, 139, 75-90. <https://doi.org/10.1016/j.compag.2017.05.005>
- Zhou, H., Fu, H., Wu, X., Wu, B., and Dai, C. (2020). Discrimination of tea varieties based on FTIR spectroscopy and an adaptive improved possibilistic c-means clustering. *Journal of Food Processing and Preservation*, 44(10). <https://doi.org/10.1111/jfpp.14795>.
- Zhou, H., Fu, H., Wu, X., Wu, B., and Dai, C. (2020). Discrimination of tea varieties based on FTIR spectroscopy and an adaptive improved possibilistic c-means clustering. *Journal of Food Processing and Preservation*, 44(10). <https://doi.org/10.1111/jfpp.14795>
- Zhou, L., Zhang, C., Taha, M. F., Wei, X., He, Y., Qiu, Z., and Liu, Y. (2020). Wheat Kernel Variety Identification Based on a Large Near-Infrared Spectral Dataset and a Novel Deep Learning-Based Feature Selection Method. *Frontiers in Plant Science*, 11. <https://doi.org/10.3389/fpls.2020.575810>

ÖZ GEÇMİŞ

Mahtem Teweldemedhin MENGSTU, 2013 yılında Keren, Eritre'deki Hamelmalo Ziraat Koleji'nden Ziraat Mühendisliği alanında lisans derecesi aldı. Ayrıca 2019 yılında Moskova Rusya Federasyonu'ndaki Rusya Devlet Tarım Üniversitesi K.A. Timiryazev'den Ziraat Mühendisliği alanında yüksek lisans derecesine sahiptir.

İletişim Bilgileri

ORCID ID : 0000-0001-5768-9150

Yayınlar:

1. Gürdil, G. A. K., **Mengstu**, M., and Medhn, T. A. (2021, June). Biomass energy potential from agricultural residues in Eritrea. Black Sea Journal of Agriculture, 4(3). <https://doi.org/10.47115/bsagriculture.908502>.
2. Gürdil, G. A. K., **Mengstu**, M., and Omer, A. (2021, September). Utilization of agricultural wastes for sustainable development. Black Sea Journal of Agriculture, 4(4). <https://doi.org/10.47115/bsagriculture.953415>.
3. **Mengstu**, M., Taner, A., and Duran, H. (2022, June 4-6). Convolutional neural networks in the production and processing of wheat: A survey. In Proceedings of the Agro International Conference on Agriculture (pp. 299-307). Azerbaijan State Agrar University. Retrieved from <https://www.scienceazerbaijan.org/agro>.
4. Taner, A., **Mengstu**, M. T., Selvi, K. Ç., Duran, H., Kabaş, Ö., Gür, İ., Karaköse, T., and Gheorghită, N.-E. (2023). Multiclass apple varieties classification using machine learning with histogram of oriented gradient and color moments. Applied Sciences, 13(13), 7682. <https://doi.org/10.3390/app13137682>.
5. Taner, A., **Mengstu**, M. T., Selvi, K. Ç., Duran, H., Gür, İ., and Ungureanu, N. (2024). Apple varieties classification using deep features and machine learning. Agriculture, 14(252). <https://doi.org/10.3390/agriculture14020252>.
6. **Mengstu**, M., Taner, A., and Duran, H. (2024, January). Spectroscopy and machine learning in food processing survey. BIO Web of Conferences, 85(5). <https://doi.org/10.1051/bioconf/20248501022>.
7. Alan, M., Taner, A., Duran, H., and **Mengstu**, M. (2024, January). Evaluation of agricultural machinery presence and usage activities in Tokat. BIO Web of Conferences, 85(4). <https://doi.org/10.1051/bioconf/20248501023>.