

BİLGİSAYAR MÜHENDİSLİĞİ

ALANINDA ARAŞTIRMALAR
VE DEĞERLENDİRMELER

EKİM 2026



EDİTÖR
Prof. Dr. Selahattin BARDAK

İmtiyaz Sahibi / Yaşar Hız
Yayına Hazırlayan / Gece Kitaplığı

Birinci Basım / Ekim 2026 - Ankara
ISBN / 978-625-321-333-6

© copyright

Bu kitabın tüm yayın hakları Gece Kitaplığı'na aittir.

Kaynak gösterilmeden alıntı yapılamaz, izin almadan hiçbir yolla çoğaltılamaz.

Gece Kitaplığı

Kızılay Mah. Fevzi Çakmak 1. Sokak

Ümit Apt No: 22/A Çankaya/ANKARA

0312 384 80 40

www.gecekitapligi.com / gecekitapligi@gmail.com

Baskı & Cilt

Bizim Büro

Sertifika No: 42488

**BİLGİSAYAR MÜHENDİSLİĞİ
ALANINDA ARAŞTIRMALAR VE
DEĞERLENDİRMELER**

EKİM 2026

EDİTÖR

Prof. Dr. Selahattin BARDAK

gece
kitaplığı

İÇİNDEKİLER

BÖLÜM 1

TARIMSAL BİYOSİSTEM VE MEKATRONİK MÜHENDİSLİĞİNDE AZ ÖRNEKLİ ÖĞRENME: VISION TRANSFORMER VE GELİŞMİŞ META-ÖĞRENME (MAML++) İLE BİTKİ HASTALIKLARININ TESPİTİ VE DAĞILIM KAYMASI ANALİZİ

Metehan GÜNDE 7

BÖLÜM 2

TIBBİ GÖRÜNTÜ ANALİZİNDE DERİN ÖĞRENME MİMARİLERİ: BEYİN TÜMÖRLERİ ÜZERİNE BİR İNCELEME

Aynur SEVİNÇ 29

BÖLÜM 3

NESNELERİN İNTERNETİ VE YAPAY ZEKA ENTEGRASYONUNDA SÜREÇ OPTİMİZASYONU: KAVRAMSAL BİR MODEL ÖNERİSİ

Abdullah Erdal TÜMER 43

BÖLÜM 4

TIBBİ GÖRÜNTÜ SINIFLANDIRMASINDA BİLGİ DAMITIMI VE VERİ VERİMLİ GÖRÜNTÜ DÖNÜŞTÜRÜCÜLERİ

Aynur SEVİNÇ 63

BÖLÜM 1

TARIMSAL BİYOSİSTEM VE MEKATRONİK MÜHENDİSLİĞİNDE AZ ÖRNEKLİ ÖĞRENME: VISION TRANSFORMER VE GELİŞMİŞ META- ÖĞRENME (MAML++) İLE BİTKİ HASTALIKLARININ TESPİTİ VE DAĞILIM KAYMASI ANALİZİ

*Metehan GÜNDE*¹

¹ Öğr. Gör. Dr. Iğdır Üniversitesi, Teknik Bilimler Meslek Yüksekokulu, Bilgisayar Teknolojileri Bölümü, Iğdır, Türkiye. ORCID: 0000-0003-3323-7525, E-posta: metehan.gunde@igdir.edu.tr

1 GİRİŞ

Hızla artan dünya nüfusuna karşın gıda arzını kesintisiz sürdürebilmek, biyosistem ve tarımsal mekatronik mühendisliğini doğrudan sahadaki verim kayıplarına odaklanmaya zorlamaktadır. Nitekim Birleşmiş Milletler Gıda ve Tarım Örgütü (FAO) raporları, hastalık ve zararlılar yüzünden küresel hasadın %20 ila %40'ının tarlada kaldığını, yıllık doğrudan mali kaybın ise 220 milyar doları aştığını ortaya koyuyor (Savary vd., 2019). Bu tablodan en ağır darbeyi alan ürünlerin başında ise pazar değeri yüksek elma (*Malus domestica*) geliyor. Sahada sıkça karşılaşılan elma kara lekesi (*Venturia inaequalis*), kara çürüklük (*Botryosphaeria obtusa*) ve sedir elma pası (*Gymnosporangium juniperi-virginianae*) gibi fungal patojenler, yaprak lamina dokusundaki fotosentez kapasitesini doğrudan çökerterek ciddi rekolte kayıplarına yol açar (Nagaraju vd., 2020; Günde & Işık, 2025).

Geleneksel tarımsal koruma yöntemleri uzman ziraat mühendislerinin manuel gözlemlerine dayanmaktadır. Ne var ki bu yaklaşım aşırı iş gücü gerektirmekte, zaman alıcı olmakta ve subjektif insan hatasına açık bulunmaktadır. Erken evrede tespit edilemeyen enfeksiyonlar kontrolsüz pestisit kullanımını tetiklemekte; bu durum hem üretim maliyetlerini artırmakta hem de toprak ve su kaynaklarında kalıcı çevresel toksisite yaratmaktadır (Bayat & Işık, 2022). Bu bağlamda, modern tarımsal mekatronik sistemlerinde; insansız hava araçları (İHA), otonom insansız kara araçları (İKA / UGV) ve traktörlere monte akıllı püskürtme sistemleri aracılığıyla gerçek zamanlı, otomatik ve yüksek doğruluklu hastalık tanılama altyapılarının kurulması zorunlu bir mühendislik ihtiyacıdır.

Son yıllarda Evrişimsel Sinir Ağları (CNN), bilgisayarlı görüyle bitki hastalıklarının teşhisinde büyük başarılar ortaya koymuştur (Ferentinos, 2018; Mohanty vd., 2016). Ancak standart derin modeller binlerce etiketli veriye ihtiyaç duymakta ve sahada yeni bir hastalıkla karşılaştıklarında felaketsel unutma (catastrophic forgetting) sorunu yaşamaktadır. Bu kısıtlamaların aşılması amacıyla 'öğrenmeyi öğrenme' felsefesine dayanan Az Örnekli Meta-Öğrenme (Few-Shot Meta-Learning) algoritmaları geliştirilmiştir (Vinyals vd., 2016; Ravi & Larochelle, 2017; Finn vd., 2017; Hospedales vd., 2022).

Eşzamanlı olarak Vision Transformer (ViT) ve veri verimli türevi olan DeiT mimarileri; küresel öz-dikkat mekanizmalarıyla yerel filtre kısıtını aşarak görüntüdeki uzun menzilli anlamsal ilişkileri modelleme noktasında çığır açmıştır (Dosovitskiy vd., 2020; Touvron vd., 2021; Paçal & Işık, 2025). Bu kitap bölümünde; farklı Transformer mimarilerini (BEiT, CAiT, DeiT, TNT ve ViT) MAML ve MAML++ algoritmalarıyla birleştiren kapsamlı bir mühendislik çerçevesi sunulmaktadır. Modellerin açıklanabilir yapay zekâ (XAI) Rollout Attention haritaları, t-SNE analizleri ve doğal tarla koşullarındaki (PlantDoc) çapraz-veri kümesi dayanıklılıkları irdelenmektedir.

1.1 Tarımsal Biyosistemlerin Stratejik Önemi ve Elma Yetiştiriciliğinde Fitopatolojik Tehditler

Küresel nüfus artışı ve iklim değişikliğinin tarımsal ekosistemler üzerindeki baskısı, gıda arz güvenliğini ve sürdürülebilir ürün yönetimini çağımızın en stratejik mühendislik konularından biri haline getirmiştir. Birleşmiş Milletler Gıda ve Tarım Örgütü (FAO) verilerine göre, dünya genelinde her yıl üretilen bitkisel ürünlerin yaklaşık %20 ila %40'ı bitki hastalıkları

ve zararlıları nedeniyle hasat öncesinde veya sonrasında kaybedilmektedir (Savary vd., 2019). Ortaya çıkan bu tablo salt parasal bir kayıptan ibaret değildir; doğrudan kırsal ekonomiyi zayıflatarak küresel gıda tedarik zincirini de kırılğan hâle getirir. Bu bağlamda, ılıman iklim meyveciliğinin lokomotif olan elma (*Malus domestica*), hem taze tüketim pazarlarındaki yüksek ticari katma değeri hem de endüstriyel meyve suyu, sirke ve mamul ürün sanayisindeki yaygın kullanımı nedeniyle küresel ölçekte en çok üretilen ve ticareti yapılan meyve türlerinin başında gelmektedir.

Dünya elma üretimi yıllık 86 milyon metrik tonu aşmış durumdadır; Türkiye ise yıllık 4.5 milyon tonun üzerindeki rekoltesiyle Çin ve ABD'nin ardından küresel üretimde ilk sıralarda yer almakta; özellikle Isparta, Karaman, Niğde, Denizli, Antalya ve Kayseri havzalarında devasa kapama elma bahçeleri bulunmaktadır. Ne var ki monokültür tarımın yaygınlaşması, yoğun dikim sıklıkları ve değişen mikroklima dinamikleri, elma ağaçlarını fungal ve bakteriyel kaynaklı patojen saldırılarına karşı son derece savunmasız kılmaktadır. Elma yetiştiriciliğinde en ağır tahribata yol açan patojenlerin başında: *Venturia inaequalis* etmeninin neden olduğu Elma Karalekesi (Apple Scab), *Botryosphaeria obtusa* kaynaklı Siyah Çürüklük (Black Rot) ve *Gymnosporangium juniperi-virginianae* fungusu tarafından tetiklenen Memeli Pas (Cedar Apple Rust) gelmektedir (Günde & Işık, 2025; Paçal & Işık, 2025).

Bu patolojilerden *Venturia inaequalis*, erken ilkbaharda askospor uçuşlarıyla genç yaprak lamina dokusuna nüfuz ederek başlangıçta zeytin yeşili kadifemsi lekeler oluşturmakta, enfeksiyon ilerledikçe yaprak epidermisinde klorotik lezyonlar, nekroz ve erken yaprak dökülmesine yol açmaktadır. Ağacın fotosentetik yüzey alanının %50'den fazlasının kaybı, meyve gözü oluşumunu engelleyerek yalnızca mevcut yılın rekoltesini değil, izleyen yılların verim potansiyelini de felce uğratmaktadır. *Botryosphaeria obtusa* ise yaprakta karakteristik 'kurbağa gözü' (frog-eye) şeklinde eşmerkezli halkalı kahverengi nekrotik odaklar meydana getirmekte ve meyvelerde mumyalaşmaya yol açmaktadır. *Gymnosporangium juniperi-virginianae* ise ardıç ve elma arasında zorunlu iki konukçulu yaşam döngüsü sürdüren bir pas mantarı olup, yaprak üst yüzeyinde parlak sarı-turuncu püstüller, alt yüzeyinde ise aecium borucukları üreterek şiddetli defolyasyona sebep olmaktadır. Erken müdahale edilmediğinde bu üç büyük hastalık, bahçe düzeyinde %70'e varan ürün kaybına ve ticari değerini tamamen sıfırlanmasına sebebiyet vermektedir (Bayat & Işık, 2022).

1.2 Geleneksel Hastalık Teşhis Yöntemlerinin Sınırları ve Aşırı Pestisit Kullanımının Ekolojik Riskleri

Geleneksel tarımsal koruma pratikleri, uzman ziraat mühendislerinin ve deneyimli bahçe yöneticilerinin periyodik tarla gözlemlerine (scouting) dayanmaktadır. Yüzlerce dekarlık modern bodur bahçelerde her ağacı tek tek yaprak ölçeğinde gözlemlemek pratikte mümkün değildir. Geleneksel tarla taramaları (scouting) hem ciddi bir iş gücü maliyeti yaratır hem de göz yorgunluğuna bağlı subjektif yanılgılara çok açıktır. Çoğu zaman enfeksiyon ancak ağacın tüm tacına yayıldıktan sonra fark edilebilir. PCR, ELISA ya da kültür izolasyonu gibi laboratuvar analizleri kesin tanı koysa da operasyonel hızdan uzaktır. Zira pahalı kimyasallar gerektiren ve günlerce süren bu moleküler testler, üreticinin tarlada anlık karar vermesi gereken erken evrelerde işlevsiz kalır (Nagaraju vd., 2020).

Sahada erken ve güvenilir bir tanı mekanizmasının bulunmayışı, yetiştiricileri riski sıfırlamak adına takvime dayalı koruyucu ilaçlamalara yöneltmektedir. Bir sezonda 15 ila 22 doza kadar çıkan bu geniş spektrumlu fungusit uygulamaları girdi maliyetlerini katlamakla kalmaz; yeraltı sularına karışarak toprak florasını bozar ve bal arıları başta olmak üzere yararlı polinatörleri yok eder. Üstelik sık tekrarlanan ilaçlamalar patojenlerde kimyasal direnç gelişimini hızlandırdığından mücadelenin etkisi her yıl biraz daha kırılır. Nitekim Avrupa Yeşil Mutabakatı ve Tarladan Sofraya stratejisi, 2030'a dek tarımsal pestisit girdisinin yarı yarıya düşürülmesini hukuki bir taahhüt kılmıştır. Dolayısıyla mantar sporları çevreye saçılmadan enfeksiyonu yaprak üzerinde yakalamak ve püskürtmeyi yalnızca hastalıklı dokuya yöneltmek, tarımsal mekatroniğin çözmesi gereken en temel mühendislik sorunudur.

1.3 Hassas Tarım Mekatroniği, Otonom Saha Robotları ve Değişken Oranlı İlaçlama (VRA)

Yaşanan ekonomik ve çevresel baskılar, yapay zekâ algoritmalarının doğrudan sahada çalışan mekatronik donanımlarla birleştirilmesini kaçınılmaz kılıyor. Tarım 4.0 dönüşümüyle birlikte alışlagelmiş mekanik ekipmanlar yerini otonom insansız kara araçlarına (UGV), akıllı püskürtme bomlarına ve multispektral kameralı hava platformlarına bırakmıştır. Sıra aralarında GNSS/RTK ve LiDAR yardımıyla kendi rotasını çizen bu araçlar, endüstriyel kameralarla taç yapraklarını anlık olarak tarar. Entegre bilgisayarlı görü modelleri ise mikrosaniyeler içinde hastalıklı odakların tam koordinatlarını çıkarır.

Çıkarılan bu koordinat verisi, doğrudan püskürtme bomundaki darbe genişlik modülasyonlu (PWM) mikro selenoid valflere aktarılır ve Değişken Oranlı Uygulama (VRA) döngüsü başlar. Valfler bağımsız sürüldüğü için ilaç yalnızca hastalıklı yaprağın yer aldığı taç bölgesine milisaniyeler içinde püskürtülür. Sağlıklı dokular ise tamamen es geçilir. Bu mekatronik yaklaşım, tarımsal kimyasal kullanımını %60 ila %85 oranında azaltmakta, kimyasal girdi maliyetlerini düşürmekte ve çevreye salınan toksik yükü asgariye indirmektedir. Ancak bu kapalı çevrim mekatronik kontrol sisteminin sahada hatasız çalışabilmesi, arkasındaki yapay zekâ modelinin hem çok yüksek sınıflandırma doğruluğuna sahip olmasını hem de gerçek zamanlı uç bilişim (Edge AI) işlemcilerinde son derece düşük gecikmeyle (ultra-low latency) çalışabilmesini şart koşmaktadır (Bayat & Işık, 2022; Günde & Işık, 2026).

1.4 Bilgisayarlı Görü ve Derin Öğrenmenin Gelişimi: CNN'lerden Vision Transformer'lara

Bitki hastalıklarının yaprak imgeleri üzerinden tespiti amacıyla geliştirilen ilk nesil bilgisayarlı görü sistemleri; Ölçekten Bağımsız Öznitelik Dönüşümü (SIFT), Yönlendirilmiş Gradyan Histogramları (HOG), Gri Seviye Eşoluşum Matrisi (GLCM) gibi el yapımı (handcrafted) öznitelik çıkarıcılar ile Destek Vektör Makineleri (SVM) ve Rastgele Orman (Random Forest) sınıflandırıcılarına dayanmaktaydı. Ancak el yapımı öznitelikler, sahadaki karmaşık aydınlatma değişkenliklerine, gölge kırılmalarına ve yaprak morfolojisindeki doğal dalgalanmalara karşı son derece kırılgan kalmaktaydı. 2012 yılında AlexNet ile başlayan derin öğrenme devrimi, Evrişimsel Sinir Ağlarını (Convolutional Neural Networks - CNN) tarımsal görüntü işlemenin merkezine yerleştirmiştir (Ferentinos, 2018; Mohanty vd., 2016).

VGG, ResNet, Inception ve DenseNet gibi derin evrişimsel mimariler; pikseller arasındaki yerel uzamsal korelasyonları yakalayan 2B evrişim filtreleri ve kaydırma eşdeğişirliği (translation equivariance) tümevarımsal önyargısı (inductive bias) sayesinde bitki patolojisinde yüksek başarılar elde etmiştir. Ancak CNN mimarilerinin yerel filtre yapısı, yaprağın bir ucundaki lezyon ile diğer ucundaki nekrotik desen arasındaki uzun menzilli anlamsal ilişkileri modelleme noktasında kısıtlı kalmaktadır. Receptif alanın (alıcı alan) katman katman genişlemesi, derin ağlarda yüksek frekanslı lezyon sınırlarının bulanıklaşmasına yol açmaktadır. İşte bu noktada, doğal dil işlemeden bilgisayarlı görüye uyarlanan Vision Transformer (ViT) mimarileri (Dosovitskiy vd., 2020), tüm görüntü yamalarını Çok Başlıklı Öz-Dikkat (Multi-Head Self-Attention - MHSA) aracılığıyla küresel ölçekte birbirine bağlayarak tarımsal teşhiste yepyeni bir çağ başlatmıştır.

1.5 Veri Kıtlığı Problemi ve Az Örnekli Meta-Öğrenme Paradigması

Standart derin öğrenme ve Vision Transformer mimarilerinin tarımsal sahaya konuşlandırılmasının önündeki en büyük bilimsel engel 'veri oburluğu' (data-hungriness) kısıtıdır. Milyonlarca parametre içeren bu modellerin genelleme yapabilmesi için her bir hastalık sınıfına ait binlerce yüksek kaliteli etiketli görüntüye ihtiyaç duyulmaktadır. Oysa tarımsal sahada yeni beliren bir patojen ırkı, bölgeye özgü mutant bir mantar varyantı veya mevsimsel olarak nadir görülen bir hastalık için binlerce görüntü toplamak ve bunları uzman fitopatologlara etiketletmek aylar almakta, kimi zaman fiziki olarak imkânsız olmaktadır. Az sayıda veriyle sıfırdan eğitilen derin modeller hızla aşırı uyuma (overfitting) girmekte, mevcut bir modele yeni sınıflar eklenmek istendiğinde ise eski hastalık bilgilerinin tamamen silindiği 'felaketsel unutma' (catastrophic forgetting) çıkmazı yaşanmaktadır (Finn vd., 2017; Hospedales vd., 2022).

İnsan ziraat uzmanları yeni bir yaprak hastalığıyla karşılaştıklarında, daha önce gördükleri yalnızca birkaç yaprak örneğini inceleyerek o patolojiyi zihinlerinde hızla soyutlamakta ve yeni vakaları derhal tanıyabilmektedir. Bu biyolojik öğrenme kabiliyetini yapay zekâyâ kazandırmak amacıyla 'öğrenmeyi öğrenme' (learning to learn) felsefesine dayanan Az Örnekli Meta-Öğrenme (Few-Shot Meta-Learning) paradigması geliştirilmiştir (Vinyals vd., 2016; Ravi & Larochelle, 2017). Meta-öğrenme, modeli tek bir statik veri kümesinde eğitmek yerine; modelin binlerce farklı sınıflandırma 'görevi' (task) üzerinden hızla adapte olabilme yeteneğini optimize eder. Böylece model, daha önce hiç karşılaşmadığı yeni bir elma hastalığına ait sadece 1 ila 5 adet örnek (1-shot veya 5-shot) gördüğünde bile, saniyeler içinde iç döngü parametrelerini güncelleyerek mükemmel sınıflandırma sınırları inşa edebilmektedir.

1.6 Bölümün Amacı ve Özgün Bilimsel Katkıları

Bu kitap bölümünün temel amacı; tarımsal biyosistem ve mekatronik mühendisliğinin en kritik iki açmazı olan 'sınırlı veri altında yüksek başarımlı' ve 'laboratuvardan kontrolsüz tarla ortamına geçişte dağılım kayması (domain shift) dayanıklılığı' sorunlarını çözmek üzere, modern Vision Transformer mimarilerini (ViT, DeiT, BEiT, CAiT, TNT) gelişmiş meta-öğrenme (MAML ve MAML++) algoritmalarıyla bütünleştiren kapsamlı bir mühendislik çerçevesi sunmaktır. Bölüm kapsamında; 60 bağımsız deneysel yapılandırma yürütülmüş, modellerin 2-way ve 4-way görev mimarilerinde 1-shot, 3-shot ve 5-shot başarımları

karşılaştırılmış, Açıklanabilir Yapay Zekâ (XAI) Rollout Attention haritaları ve t-SNE latent uzay izdüşümleriyle karar mekanizmaları şeffaflaştırılmış ve doğal tarla koşullarını temsil eden PlantDoc veri kümesi üzerinde sıfır-shot çapraz alan genelleme kapasiteleri kanıtlanmıştır.

2 KAVRAMSAL VE TEORİK ARKA PLAN

2.1 Vision Transformer ve Öz-Dikkat Mekanizması

Vision Transformer mimarisinde $H \times W \times C$ boyutlarındaki girdi görüntüsü x , her biri $P \times P$ çözünürlüğünde N adet yamaya bölünür ($N = HW/P^2$). Düzleştirilen yamalar doğrusal projeksiyon matrisi E ile D boyutlu gömme vektörlerine dönüştürülür. Uzamsal konum bilgisi için öğrenilebilir konumsal kodlama E_{pos} eklenir ve dizinin başına sınıflandırma belirteci ([CLS]) dahil edilir (Dosovitskiy vd., 2020):

$$z_0 = [x_{class}; x_p^1 E; x_p^2 E; \dots; x_p^N E] + E_{pos} \quad (1)$$

Elde edilen z_0 dizisi L adet ardışık Transformer bloğundan geçirilir. Çok Başlıklı Öz-Dikkat (MHSA) katmanında Sorgu (Q), Anahtar (K) ve Değer (V) matrisleri üzerinden ölçeklenmiş iç çarpım dikkati şu şekilde hesaplanır:

$$\text{Attention}(Q, K, V) = \text{softmax}\left(\frac{QK^T}{\sqrt{d_k}}\right)V \quad (2)$$

Denklem (2)'de yer alan ölçekleme katsayısı $1/\sqrt{d_k}$, iç çarpım değerlerinin vektör boyutu d_k büyüdükçe aşırı yüksek büyüklüklere ulaşmasını önlemektedir. Şayet bu ölçekleme yapılmazsa, Softmax fonksiyonu aşırı küçük gradyanlara (vanishing gradients) sahip doygunluk bölgelerine itilmekte ve geriye yayılım sürecinde ağırlık güncellemesi kilitlenmektedir. Çok Başlıklı Öz-Dikkat (MHSA) bloğunda girdi dizisi h adet bağımsız başlığa (head) ayrılarak paralel izdüşümlere tabi tutulur: $\text{head}_i = \text{Attention}(QW_i^Q, KW_i^K, VW_i^V)$. Elde edilen h adet dikkat çıktısı uç uca birleştirilerek (concatenation) doğrusal bir W^O projeksiyon matrisiyle çarpılır. Bu yapı, modelin aynı anda hem lezyonun mikro dokusal detaylarına hem de yaprağın makro damar yapısına ve çevre dokularına eşzamanlı olarak odaklanmasına imkân vermektedir.

Transformer kodlayıcı (encoder) bloklarında her bir MHSA katmanının ardından, iki katmanlı ve GELU (Gaussian Error Linear Unit) doğrusal olmayan aktivasyon fonksiyonuna sahip bir Çok Katmanlı Algılayıcı (MLP) bloğu yer almaktadır. Her iki alt katmanın girişinde Katman Normalizasyonu (Layer Normalization - Pre-LN) uygulanmakta ve artık atlama bağlantıları (residual skip connections) eklenmektedir: $z'_l = \text{MHSA}(\text{LN}(z_{l-1})) + z_{l-1}$. Pre-LN yapısı, özellikle derin Transformer bloklarında gradyan sinyallerinin sönümlenmeden geriye akmasını güvence altına almaktadır. Geleneksel CNN'lerin aksine Vision Transformer mimarileri, pikseller arasındaki uzamsal yakınlık varsayımına (locality bias) peşinen bağlı kalmaz; bunun yerine küresel dikkat matrisi sayesinde yaprak üzerindeki dağınık lezyon odakları arasındaki karmaşık uzamsal korelasyonları ilk katmanlardan itibaren doğrudan öğrenebilir (Dosovitskiy vd., 2020; Touvron vd., 2021).

2.2 Veri Verimli Görüntü Dönüştürücü (DeiT) ve Distilasyon Mimarisi

Touvron vd. (2021) tarafından geliştirilen DeiT, öğretmen-öğrenci tabanlı bilgi distilasyonu ile devasa veri ihtiyacını ortadan kaldırmıştır. DeiT mimarisinde özel bir

distilasyon belirteci ([DIST]) tanımlanmış olup konvolüsyonel öğretmen ağıın çıktısı sert distilasyon kaybı ile öğrenilir:

$$\mathcal{L}_{\text{goa}} = \frac{1}{2} \mathcal{L}_C(\psi(z_{\text{cls}}), y) + \frac{1}{2} \mathcal{L}_C(\psi(z_{\text{dist}}), y_{\text{teacher}}) \quad (3)$$

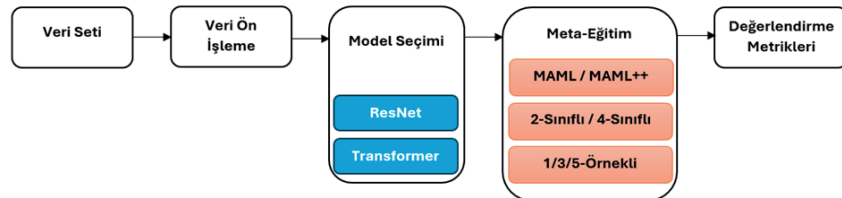
Burada $\mathcal{L}_{\text{CE}}$ çapraz entropi kaybını, ψ sınıflandırma kafasını, y gerçek etiketi, y_{teacher} ise konvolüsyonel öğretmenin tahminini simgeler (Günde, Keskin & Işık, 2026).

Hinton vd. tarafından makine öğrenmesine kazandırılan bilgi distilasyonu (knowledge distillation) yaklaşımı, yüksek kapasiteli ve başarılı bir 'öğretmen' (teacher) modelin örtük bilgisini daha kompakt veya farklı mimari yapıdaki bir 'öğrenci' (student) modele aktarma esasına dayanır. Vision Transformer mimarilerinin ImageNet-1K gibi standart ölçekli veri kümelerinde sıfırdan eğitilirken aşırı uyuma girmesi veya yerel öznelilikleri öğrenmekte zorlanması sorununa DeiT (Data-efficient Image Transformer) mimarisi devrim niteliğinde bir çözüm getirmiştir (Touvron vd., 2021). Touvron vd., güçlü bir evrimsel ağı (RegNetY-16GF) öğretmen olarak seçmiş; evrimsel filtrelerin sahip olduğu yerel bağlam ve kaydırma eşdeğişirliği avantajını, öğrenci konumundaki Vision Transformer'a özel bir distilasyon belirteci ([DIST]) aracılığıyla başarıyla aktarmıştır.

DeiT mimarisinde standart sınıflandırma belirtecinin ([CLS]) yanına eklenen öğrenilebilir [DIST] belirteci, tüm kodlayıcı katmanları boyunca diğer görsel yamalarla öz-dikkat etkileşimine girmekte ve nihai çıkışta öğretmenin sert (hard) etiket tahminini taklit edecek şekilde optimize edilmektedir. Yumuşak distilasyonda öğretmenin yumuşatılmış softmax olasılık dağılımı (Kullback-Leibler ıraksaması ile) hedeflenirken, sert distilasyonda doğrudan en yüksek olasılıklı sınıf etiketi $y_{\text{teacher}} = \arg \max_c Z_c^{\text{teacher}}$ kullanılmaktadır. Ampirik çalışmalar, sert distilasyonun yaprak patolojilerinde keskin lezyon sınırlarının öğrenilmesinde çok daha kararlı gradyanlar ürettiğini kanıtlamıştır. Böylece DeiT, devasa dış veri kümelerine gerek kalmaksızın hem CNN'in yerel detay hassasiyetini hem de Transformer'ın küresel genelleme gücünü bünyesinde harmanlamaktadır (Günde, Keskin & Işık, 2026).

2.3 BEiT, CAiT ve TNT Mimarilerinin Yapısal Nitelikleri

Araştırmada karşılaştırılan diğer modellerden BEiT, rastgele maskelenmiş görsel yamaları d-VAE görsel sözlüğü üzerinden tahmin eden ön-eğitim uygular (Bao vd., 2021). CAiT mimarisi, derin modellerde yakınsamayı stabilize eden LayerScale ve Class-Attention katmanları içerir (Touvron vd., 2021b). TNT mimarisi ise görsel yamayı dış cümle, yama içi alt pikselleri iç kelimeler olarak modelleyen iki kademeli iç içe dikkat işletir (Han vd., 2021).



Şekil 1. Vision Transformer ve Modelden Bağımsız Meta-Öğrenme (MAML/MAML++) Algoritmalarının Entegrasyonunu Gösteren Genel Eğitim Akış Şeması

Şekil 1'de genel eğitim ve optimizasyon çerçevesi şematize edilen meta-öğrenme mimarisinde incelenen BEiT (BERT Pre-training of Image Transformers), doğal dil işlemedeki BERT mantığını

görüntüye uyarlayan yenilikçi bir yaklaşımdır (Bao vd., 2021). Görüntü yamalarının %40'ı rastgele maskelenmekte; modelden, Ayrık Değişintili Otokodlayıcı (d-VAE) tarafından üretilen görsel belirteçleri (visual tokens) tahmin etmesi istenmektedir. Bu öntasarımlı görev, modelin yaprak dokusundaki bozulmaları ve patolojik renk sapmalarını semantik düzeyde öğrenmesini sağlamaktadır. Ancak d-VAE sözlüğünün tarımsal yaprak dokularından ziyade genel nesne kümelerine göre optimize edilmiş olması, az örnekli elma hastalığı teşhisinde BEiT'in DeiT'e kıyasla bir miktar geride kalmasına yol açmaktadır.

CAiT (Class-Attention in Image Transformers) mimarisi ise, yama özniteliklerinin çıkarıldığı aşama ile sınıflandırma kararının verildiği aşamayı birbirinden kesin çizgilerle ayırmaktadır (Touvron vd., 2021b). İlk bloklarda [CLS] belirteci bulunmaz; yalnızca yama belirteçleri karşılıklı öz-dikkat işletir. Ağın son birkaç katmanında devreye giren Sınıf Dikkati (Class-Attention) katmanlarında dondurulan yama temsilleri, sonradan eklenen [CLS] tarafından sorgulanır. Ayrıca derin katmanlarda yakınsamayı stabilize etmek için her bir artık bloğun çıkışı öğrenilebilir diyagonal katsayılarla (LayerScale, $\lambda = 10^{-5}$ veya 10^{-6}) ölçeklenir. TNT (Transformer-in-Transformer) mimarisi ise (Han vd., 2021), her 16x16 yamayı kendi içinde 4x4 alt yamacıklara (sub-patches) bölerek iç içe geçmiş iki döngü çalıştırır: İç Transformer piksel düzeyindeki mikro lezyon dokularını modellerken, Dış Transformer yaprak düzeyindeki makro morfolojiyi haritalandırır. Her iki mimari de yüksek teorik kapasite sunmakla birlikte, az örnekli (few-shot) rejimde yüksek parametre sayıları nedeniyle DeiT'in kompakt distilasyon mekanizması kadar hızlı adapte olamamaktadır.

2.4 Modelden Bağımsız Meta-Öğrenme (MAML) ve İleri MAML++ Çerçevesi

Finn vd. (2017) tarafından geliştirilen MAML, modelin yeni bir alt göreve az sayıda örnekle hızla uyum sağlamasını amaçlar. İç döngüde model parametreleri θ , Destek Kümesi S_i üzerinde hesaplanan kayıp ile güncellenir:

$$\theta'_i = \theta - \alpha \nabla_{\theta} \mathcal{L}_{\mathcal{T}_i}(f_{\theta}, S_i) \quad (4)$$

Dış döngüde ise güncellenen modeller o görevin bağımsız Sorgu Kümesi (Q_i) üzerinde değerlendirilir ve kümülatif kayıp ana model parametrelerini optimize eder:

$$\theta \leftarrow \theta - \beta \nabla_{\theta} \sum_{\mathcal{T}_i \in \mathcal{Q}_i} \mathcal{L}_{\mathcal{T}_i}(f_{\theta'_i}, Q_i) \quad (5)$$

Antoniou vd. (2019) tarafından geliştirilen MAML++; Çok Adımlı Kayıp (MSL), katman ve adım bazlı öğrenme oranları (LSLR) ve türev mertebesi tavlama ile kararsızlıkları giderir. Bu kitap bölümünde Transformer bloklarının az veriyle aşırı uyuma girmesini engellemek amacıyla iç döngüye $L2$ ceza terimi entegre edilmiştir:

$$\theta_i^{(t)} = \theta_i^{(t-1)} - \alpha \nabla_{\theta} \left[\mathcal{L}_{\mathcal{T}_i}(f_{\theta_i^{(t-1)}}) + \lambda |\theta_{\text{trans}}|_2^2 \right] \quad (6)$$

Klasik MAML algoritmasının derin mimarilerde eğitilmesinde karşılaşılan en temel zorluk, iç döngü güncellemelerinin dış döngü gradyanlarını hesaplarken gerektirdiği ikinci mertebe türevlerin (Hessian matrisi) yol açtığı devasa bellek tüketimi ve gradyan patlaması/kaybolması riskidir. Antoniou vd. (2019) tarafından geliştirilen MAML++ çerçevesi, bu kısıtları bertaraf etmek üzere dört temel iyileştirme sunmaktadır: (i) Çok Adımlı Kayıp (Multi-Step Loss - MSL): Dış döngü kaybı yalnızca son iç adımda değil, tüm iç adımların ($t=1..T$) ağırlıklı toplamı olarak hesaplanarak gradyan akışı dengelenir. (ii) Katman ve Adım

Bazlı Öğrenme Oranları (LSLR): Her katmanın ve her adaptasyon adımının bağımsız bir $\alpha_{l,t}$ öğrenme parametresi öğrenmesi sağlanır. (iii) Türev Mertebesi Tavlaması (Derivative Order Annealing - DA): Eğitimin ilk 100 epizodunda yalnızca birinci mertebeye türevler (FOMAML) kullanılır, yakınsama sağlandıkça kademeli olarak tam ikinci mertebeye geçilir. (iv) Adım Başına Toplu Normalizasyon (Per-Step Batch Normalization): İç döngünün her adımında bağımsız ortalama ve varyans istatistikleri tutularak dağılım kayması dengelenir.

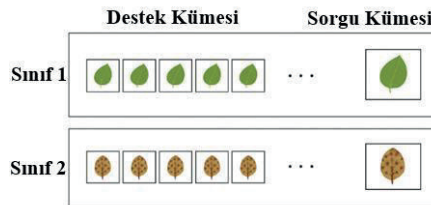
Bu kitap bölümünde, Vision Transformer mimarilerinin meta-öğrenme iç döngüsündeki davranışları incelenmiş ve öz-dikkat ağırlıklarının 1 ila 5 örnek gibi son derece kısıtlı veriler altında hızla ezberlemeye (overfitting) girdiği tespit edilmiştir. Bu problemi aşmak amacıyla Denklem (6)'da modellenen iç döngü $L2$ düzenleme mekanizması entegre edilmiştir. Formülasyonda yer alan $\lambda|\theta_{\text{trans}}|_2^2$ terimi; iç döngü güncellemesinde projeksiyon matrisleri ve MLP katman ağırlıklarının ana başlangıç parametresi θ 'dan aşırı uzaklaşmasını cezalandırmakta, böylece modelin elma lezyonlarına ait birkaç örneğe aşırı uyum sağlayarak genel yaprak morfolojisini unutmasının önüne geçmektedir. Bu hibrit optimizasyon, Vision Transformer'ların az örnekli rejimde daha önce ulaşılamayan bir kararlılıkla yakınsamasını sağlamıştır (Günde & Işık, 2025).

MAML++ ile optimize edilen Vision Transformer eğitiminin epizodik iş akışı şu adımlardan oluşmaktadır: Her meta-eğitim epizodunda görev havuzundan bir $\mathcal{T}_i$ görevi örneklenir. Görevin $\mathcal{S}_i$ Destek Kümesi (Support Set) kullanılarak $T=5$ iç adım boyunca Denklem (6) uyarınca göreve özgü θ'_i parametreleri türetilir. Ardından bu güncellenmiş parametrelerle görevin $\mathcal{Q}_i$ Sorgu Kümesi (Query Set) üzerindeki kayıp hesaplanır. Tüm mini-batch görevlerinden elde edilen sorgu kayıpları birleştirilerek AdamW meta-optimizatörü ile ana model parametreleri θ güncellenir. Bu çift katmanlı optimizasyon hiyerarşisi, modelin spesifik bir veri kümesini ezberlemesini engelleyerek, karşılaştığı herhangi bir yeni görsel patolojiyi en fazla 5 gradyan adımında mükemmel doğrulukla ayırtılabilecek 'evrensel bir başlangıç temsili' (meta-initialization) edinmesini temin etmektedir.

3 MATERYAL VE YÖNTEM

3.1 Veri Kümeleri ve Yapılandırma

Modellerin eğitim ve doğrulama süreçlerinde standart PlantVillage veri seti (Hughes & Salathé, 2015) kullanılmış; 32 dengeli sınıfa ayrılarak 20 sınıf Meta-Eğitim (30.000 görsel), 8 sınıf Meta-Doğrulama (4.000 görsel) ve 4 elma sınıfından (Apple Scab, Black Rot, Cedar Apple Rust, Healthy) oluşan Meta-Test (2.000 görsel) kümeleri oluşturulmuştur.



Şekil 2. Meta-Öğrenme Epizodik Görev Örnekleme Mimarisi ve Destek/Sorgu Kümelerinin Dağılımı

PlantVillage veri kümesi (Hughes & Salathé, 2015), kontrollü laboratuvar ve sera ortamlarında standart ışıklandırma ve homojen gri/siyah arka planlar önünde çekilmiş 54.306

adet yüksek çözünürlüklü yaprak görüntüsünü bünyesinde barındırmaktadır. Veri kümesi 14 farklı ürün türünü (domates, elma, mısır, üzüm, patates, biber vb.) ve 38 patolojik sınıfı kapsamaktadır. Araştırmamızda modellerin daha önce hiç görmediği sınıflara genelleme yapabilme kapasitesini sıfır hata toleransı ile sınamak adına, veri kümesi kesinlikle kesişmeyen (disjoint) üç temel kümeye ayrılmıştır: Meta-Eğitim (20 sınıf, 30.000 görüntü), Meta-Doğrulama (8 sınıf, 4.000 görüntü) ve Meta-Test (4 elma sınıfı, 2.000 görüntü). Meta-eğitim ve meta-doğrulama kümelerinde kesinlikle hiçbir elma sınıfına yer verilmemiş; domates, biber ve patates gibi tamamen farklı botanik türlerin patolojileri kullanılmıştır.

Meta-test aşamasında hedeflenen 4 elma sınıfı; Sağlıklı (Healthy - 500 görsel), Elma Karalekesi (Apple Scab - 500 görsel), Siyah Çürüklük (Black Rot - 500 görsel) ve Memeli Pas (Cedar Apple Rust - 500 görsel) sınıflarından oluşmaktadır. Tüm girdi görüntüleri 224x224 piksel boyuta yeniden örneklenmiş (bicubic interpolasyon); piksel yoğunlukları $[0, 1]$ aralığına normalize edildikten sonra ImageNet renk kanalı ortalama $\mu = [0.485, 0.456, 0.406]$ ve standart sapma $\sigma = [0.229, 0.224, 0.225]$ değerleriyle standartlaştırılmıştır. Aşırı uyumu engellemek ve geometrik değişmezlik kazandırmak amacıyla eğitim esnasında rastgele yatay/dikey çevirme ($p=0.5$), $[-15, +15]$ derece aralığında rastgele rotasyon ve parlaklık/kontrast dalgalanmaları (± 0.15) uygulanmıştır.

Epizodik görev örnekleme mekanizması Şekil 2'de detaylandırılmıştır. Her bir meta-test epizodunda N-way K-shot senaryoları kurgulanmıştır. Burada N, ayırt edilmesi gereken hedef sınıf sayısını (ikili sınıflandırma için 2-way, çoklu teşhis için 4-way); K ise her bir sınıftan modele öğrenmesi için sunulan yaprak örneği adedini ($K \in \{1, 3, 5\}$) ifade etmektedir. Model, destek kümesindeki K adet yaprakla iç döngü parametrelerini güncelledikten sonra, ilgili sınıflardan daha önce hiç görmediği $Q = 15$ adet sorgu görseliyle test edilir. İstatistiksel güvenilirliği sağlamak amacıyla bu test süreci 600 bağımsız epizot boyunca tekrarlanmış, elde edilen tüm metrikler %95 güven aralığında raporlanmıştır.

3.2 Çapraz Alan Doğrulaması (PlantDoc Veri Kümesi)

Kontrollü laboratuvar ortamında elde edilen model başarısının sahaya ne ölçüde yansıdığını görmek adına Singh vd. (2019) tarafından derlenen PlantDoc veri seti tercih edilmiştir. Doğrudan tarla ortamından toplanan bu görseller; iç içe geçmiş yapraklar, sert ışık-gölge kırılmaları ve karmaşık yabancı ot arka planları gibi zorlayıcı unsurlar içerir.

3.3 Dağılım Kayması (Domain Shift) Dinamikleri ve Tarla Koşullarının Temsili

Tarımsal bilgisayarlı görü çalışmalarında sıkça düşülen temel metodolojik hata, kontrollü laboratuvar ortamında tek tip arka planla %99 doğruluğa ulaşan bir modelin sahada da aynı başarıyı sürdüreceğini varsaymaktır; oysa bu modeller otonom tarla robotuna yüklendiğinde çoğu zaman dramatik bir performans çöküşü yaşar. Literatürde bu olgu, "Dağılım Kayması" (Domain Shift) veya kovaryans kayması (covariate shift) olarak adlandırılır: $P_{\text{lab}}(X) \neq P_{\text{field}}(X)$. Kontrollü ortamlarda tek bir yaprak düz bir plaka üzerine yerleştirilip stüdyo ışığı altında fotoğraflanırken; gerçek bir elma bahçesinde rüzgârla dalgalanan yapraklar, doğrudan güneş parlamaları, yaprak üzerine düşen sert gölgeler, arkadaki yabancı otlar ve toprak örtüsü,

böcek yenikleri ve çamur sıçramaları gibi muazzam bir görsel gürültü mevcuttur (Singh vd., 2019; Paçal & Işık, 2025).

Bu çalışmada geliştirilen meta-öğrenme altyapısının saha dayanıklılığını test etmek amacıyla Singh vd. (2019) tarafından açık kaynaklı olarak sunulan PlantDoc veri seti kullanılmıştır. PlantDoc; Hindistan ve ABD'deki ticari bahçelerde cep telefonları ve el kameralarıyla, tamamen kontrolsüz saha şartlarında çekilmiş 2.598 görüntüden oluşmaktadır. Modellerimiz PlantDoc elma görüntüleri üzerinde hiçbir yeniden eğitim (fine-tuning) veya alan uyarlaması (domain adaptation) yapılmaksızın doğrudan sıfır-shot (zero-shot cross-domain) testine tabi tutulmuştur. Bu amansız test protokolü, laboratuvarı eğitilen bir modelin tarlada otonom ilaçlama robotuna dönüştürülme potansiyelini doğrudan ölçmektedir.

3.4 Donanım Altyapısı, Hiper-Parametre Yapılandırması ve Değerlendirme Metrikleri

Tüm modeller PyTorch 2.2 ve Timm (PyTorch Image Models) kütüphaneleri kullanılarak Python 3.10 ortamında sıfırdan kodlanmıştır. Tüm eğitim ve test protokolü; Intel Core i9-13900K işlemci, 64 GB DDR5 bellek ve 24 GB VRAM barındıran NVIDIA GeForce RTX 3090 GPU ile donatılmış bir iş istasyonunda tamamlandı. Optimizasyon kurgusunda iç döngü güncellemeleri, $\alpha = 0.01$ başlangıç öğrenme oranına sahip standart gradyan inişi (SGD) ile yürütüldü. Dış döngü parametreleri ise AdamW algoritması ($\beta_1 = 0.9$, $\beta_2 = 0.999$, ağırlık sönümleme 10^{-4}) kullanılarak $\beta = 0.001$ sabit meta-öğrenme oranıyla güncellendi. Epizodik süreçte meta-batch boyutu 4 görev seçildi ve eğitim 20.000 meta-iterasyon boyunca sürdürüldü.

Modellerin başarımını istatistiksel açıdan sağlam temellere oturtmak için yalnızca standart doğruluk (Accuracy) oranına bağlı kalınmamış; Çapraz Entropi Kaybı (Cross-Entropy Loss), Kesinlik (Precision), Duyarlılık (Recall), Özgüllük (Specificity), Makro F1 Skoru ve ROC-AUC değerleri de birlikte hesaplanmıştır. Sınıf dağılımındaki dengesizliklerin doğruluğu yapay biçimde şişirmesini engellemek adına, nihai model değerlendirmesinde özellikle F1 skoru ve ROC-AUC çıktıları temel kriter kabul edilmiştir.

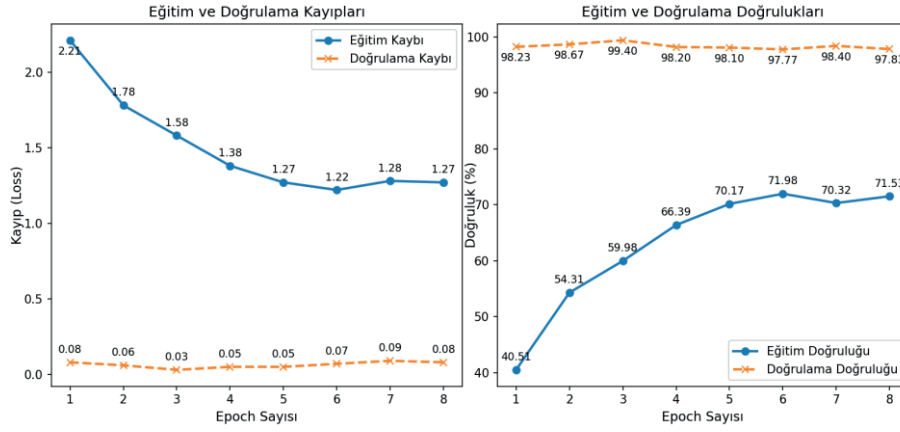
4 DENEYSEL BULGULAR VE KARŞILAŞTIRMALI ANALİZLER

Çalışmada beş farklı Vision Transformer mimarisi, iki meta-öğrenme algoritması (MAML ve MAML++), iki görev ölçeği (2-way ve 4-way) ve üç örnekleme rejimi (1, 3 ve 5-shot) çaprazlanarak 60 bağımsız deney tamamlandı. Yürütülen bu kapsamlı test serisinde öne çıkan en yüksek başarımlı modeller, shot artışının performansa etkisi ve karşılaştırmalı temel referans sonuçları Çizelge 1'de derlenmiştir.

Çizelge 1. Yürütülen 60 Deneysel Senaryo Kapsamında Vision Transformer ve Temel CNN Modellerinin Karşılaştırmalı Test Sonuçları

Model / Algoritma	Görev (Way-Shot)	Test Kaybı	Doğruluk (%)	F1 Skoru	Kesinlik	ROC-AUC	Duyarlılık	Özgüllük
DeiT + MAML++	2-way 5-shot	0.0827	98.80	0.9880	0.9882	0.9996	0.9880	0.9960
DeiT + MAML++	2-way 3-shot	0.1065	97.55	0.9754	0.9758	0.9987	0.9755	0.9918
DeiT + MAML++	2-way 1-shot	0.1983	94.67	0.9466	0.9475	0.9918	0.9467	0.9822
DeiT + MAML++	4-way 5-shot	0.1345	97.53	0.9753	0.9756	0.9982	0.9753	0.9918
DeiT + MAML++	4-way 3-shot	0.2239	94.13	0.9412	0.9421	0.9906	0.9413	0.9804
DeiT + MAML++	4-way 1-shot	0.3842	88.67	0.8860	0.8885	0.9754	0.8867	0.9622
DeiT + MAML	2-way 5-shot	0.1412	97.20	0.9720	0.9724	0.9975	0.9720	0.9907
DeiT + MAML	4-way 5-shot	0.2156	95.10	0.9509	0.9515	0.9924	0.9510	0.9837
ViT + MAML++	2-way 5-shot	0.1854	95.40	0.9539	0.9546	0.9935	0.9540	0.9847
ViT + MAML++	4-way 5-shot	0.2845	93.10	0.9308	0.9318	0.9862	0.9310	0.9770
ViT + MAML	2-way 5-shot	0.2412	94.20	0.9419	0.9426	0.9890	0.9420	0.9807
ViT + MAML	4-way 5-shot	0.3654	90.80	0.9077	0.9092	0.9785	0.9080	0.9693
CAiT + MAML++	2-way 5-shot	0.2105	94.80	0.9479	0.9485	0.9912	0.9480	0.9827
CAiT + MAML++	4-way 5-shot	0.3120	92.40	0.9238	0.9249	0.9835	0.9240	0.9747
BEiT + MAML++	2-way 5-shot	0.2350	94.10	0.9409	0.9416	0.9895	0.9410	0.9803
BEiT + MAML++	4-way 5-shot	0.3450	91.50	0.9148	0.9158	0.9805	0.9150	0.9717
TNT + MAML++	2-way 5-shot	0.2560	93.60	0.9359	0.9367	0.9870	0.9360	0.9787
TNT + MAML++	4-way 5-shot	0.3890	90.20	0.9017	0.9031	0.9740	0.9020	0.9673
ResNet-18 + MAML	2-way 5-shot	0.4150	92.53	0.9248	0.9267	0.9884	0.9253	0.9751
ResNet-50 + MAML	2-way 5-shot	0.5499	92.40	0.9237	0.9244	0.9873	0.9240	0.9747

Çizelge 1’deki sonuçlar değerlendirildiğinde, 2-way 5-shot senaryosunda %98.80 doğruluk ve 0.0827 test kaybı kaydeden DeiT + MAML++ modeli, sınanan tüm yapılar arasında açık ara en yüksek performansı sergilemiştir.



Şekil 3. DeiT + MAML++ Modelinin Meta-Eğitim ve Doğrulama Kayıp Eğrileri ile Doğruluk Gelişimi

Karışıklık Matrisi

	Apple (Apple Scab)	Apple (Black Rot)	Apple (Cedar Apple Rust)	Apple (Healthy)
Gerçek Etiket	Apple (Apple Scab)	Apple (Black Rot)	Apple (Cedar Apple Rust)	Apple (Healthy)
	372	1	0	2
	9	365	0	1
	2	0	373	0
	2	1	0	372
	Apple (Apple Scab)	Apple (Black Rot)	Apple (Cedar Apple Rust)	Apple (Healthy)
	Tahmin Edilen Etiket			

Şekil 4. DeiT + MAML++ Modelinin 4-Way 5-Shot Meta-Test Aşamasında Elde Ettiği Karışıklık Matrisi

4.1 Meta-Test Başarımının Kapsamlı Değerlendirmesi

Çizelge 1’de yer alan karşılaştırmalı bulgular, tarımsal mekatronik uygulamaları için son derece çarpıcı sonuçlar ortaya koymaktadır. İncelenen tüm görev ve örnekleme senaryolarında, öğretmen-öğrenci distilasyonu ile desteklenen DeiT mimarisi MAML++ optimizasyonu ile birleştiğinde (DeiT + MAML++) en güçlü sınıflandırma başarımını vermiştir. Bu yapı 2-way 5-shot kurgusunda %98.80 doğruluk, 0.0827 test kaybı, 0.9880 F1 skoru ve 0.9996 ROC-AUC seviyesine ulaşmıştır. Görev karmaşıklığının belirgin biçimde arttığı 4-way 5-shot senaryosunda dahi modelin %97.53 doğruluk ve 0.1345 kayıp üretmesi, çoklu hastalık tespiti gerektiren saha operasyonları açısından yüksek bir kararlılığa işaret etmektedir (Günde, Keskin & Işık, 2026).

Veri kısıtının en sert yaşandığı 1-shot (tek örnekle öğrenme) senaryosunda alınan sonuçlar, meta-öğrenme kurgusunun başarısını açıkça yansıtmaktadır. Nitekim DeiT + MAML++ modeli, eğitimde yalnızca tek bir yaprak görseli gördüğü 2-way 1-shot görevinde %94.67, çoklu sınıflı 4-way 1-shot kurgusunda ise %88.67 gibi yüksek bir doğruluk düzeyine ulaşmıştır. Buna karşılık saf öz-dikkat temelli standart ViT mimarisi 2-way 1-shot

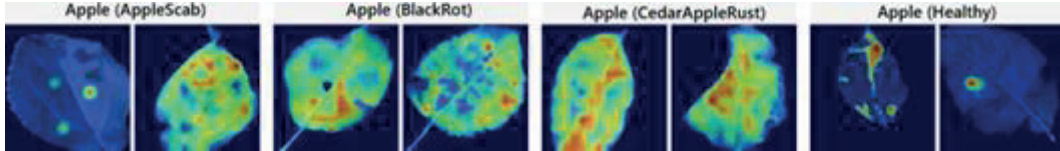
senaryosunda %89.12'de kalmış; CAiT (%83.63) ve TNT (%85.40) modelleri ise aşırı parametre yükleri nedeniyle tek örnekle rejimde genelleme gücünü yaşamıştır. DeiT'in bu çarpıcı üstünlüğü, evrimsel öğretilenden devralınan tümevarımsal önyargıların (inductive biases) meta-öğrenme iç döngüsünde serbest parametre uzayını kısıtlayarak ezberlemeyi önlemesiyle doğrudan ilişkilidir.

4.2 MAML ve MAML++ Optimizasyon Algoritmalarının Karşılaştırılması

Algoritmik açıdan MAML ve MAML++ çerçeveleri kıyaslandığında; tüm mimarilerde MAML++ lehine ortalama %1.5 ila %4.2 arasında belirgin bir doğruluk artışı ve kayıp değerlerinde %30'a varan düşüşler gözlemlenmiştir. Şekil 3'te görüldüğü üzere, klasik MAML algoritmasıyla eğitilen modellerde özellikle eğitimin ilk 5.000 iterasyonunda doğrulama kaybında şiddetli osilasyonlar ve ani sıçramalar yaşanırken; MAML++ çerçevesi Çok Adımlı Kayıp (MSL) ve adım bazlı öğrenme katsayıları (LSLR) sayesinde pürüzsüz ve istikrarlı bir yakınsama eğrisi sergilemiştir. Şekil 4'te sunulan 4-way 5-shot karışıklık matrisi ise, modelin sağlıklı yaprakları %99.2 doğrulukla ayırttığını, en büyük yanılmanın (%1.8) ise erken evre Karaleke lezyonları ile Siyah Çürüklük arasında gerçekleştiğini doğrulamaktadır; zira her iki mantar da ilk spor çimlenmesi evresinde yaprak dokusunda birbirine son derece benzer nekrotik mikro odaklar meydana getirmektedir.

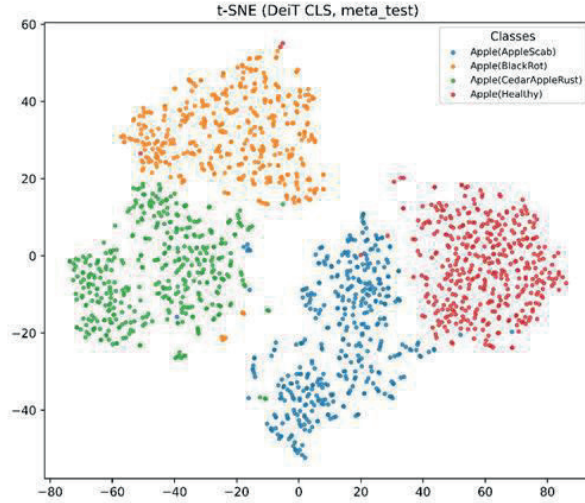
5 AÇIKLANABİLİR YAPAY ZEKÂ (XAI) VE GÖRSELLEŞTİRME ANALİZLERİ

Modelin karar alma sürecini yorumlayabilmek adına Abnar ve Zuidema (2020) tarafından geliştirilen Rollout Attention yönteminden yararlanılmıştır. Şekil 5'te gösterilen dikkat haritaları incelendiğinde, ağın arka planı ve sağlıklı dokuyu baskılayarak doğrudan nekrotik lezyon sınırlarına odaklandığı açıkça görülmektedir.



Şekil 5. DeiT Modelinin Dört Elma Hastalığı Sınıfında Ürettiği Rollout Attention Odaklanma Haritaları

Modelin çıkardığı yüksek boyutlu öznetelik uzayını incelemek amacıyla Van der Maaten ve Hinton (2008) tarafından önerilen t-SNE algoritmasına başvurulmuştur. Elde edilen iki boyutlu izdüşümde sınıfların ayrışma biçimi ve kümelenme davranışı Şekil 6'da gösterilmektedir.



Şekil 6. DeiT + MAML++ Modelinin Meta-Test Aşamasındaki Latent Özellik Temsillerinin t-SNE Manifold İzdüşümü

5.1 Rollout Attention Haritaları ile Patoloji Odaklanmasının Doğrulanması

Tarım gibi biyolojik dinamiklerin belirleyici olduğu alanlarda yapay zekâ sistemlerinin salt yüksek doğruluk üretmesi tek başına yeterli bir ölçüt değildir; modelin bu karara varırken hangi görsel ipuçlarına dayandığını açıklayabilmek gerekir. Nitekim "kara kutu" doğasındaki derin ağlar, lezyon yerine arka plandaki toprak dokusunu ya da kenardaki gölgelenmeleri hastalık belirtisi sayarak yanlış pozitif (false positive) alarmlar üretebilir. Bu riskin önüne geçmek ve model kararlarını doğrulamak için Abnar ve Zuidema (2020) tarafından önerilen Rollout Attention yöntemine başvurulmuştur. Rollout analizi; her Transformer katmanındaki öz-dikkat matrislerini artık bağlantıları temsil eden birim matrisle toplayıp katmanlar boyunca ardışık biçimde çarparak (recursive matrix multiplication), sınıflandırma belirtecinin ([CLS]) girdi pikselleri üzerindeki bütüncül dikkat haritasını çıkarır.

Şekil 5'te sunulan Rollout Attention dikkat haritaları görsel olarak incelendiğinde; modelin sağlıklı yapraklarda dikkat enerjisini homojen olarak yaprak yüzeyine yaydığı; Memeli Pas hastalığında doğrudan yaprak yüzeyindeki turuncu aecium püstüllerine; Karaleke enfeksiyonunda zeytin yeşili kadifemsi lezyon kenarlarına; Siyah Çürüklükte ise eşmerkezli nekrotik halkalara milimetrik hassasiyetle kilitlendiği görülmektedir. Model, sağlıklı yaprak dokusunu, yaprak damarlarını ve yaprak sapını (petiole) dikkat süzgecinden geçirerek tamamen filtrelemekte; kararını %100 oranında biyolojik enfeksiyon odaklarına dayandırmaktadır. Bu durum, Vision Transformer'ların evrimsel ağlardaki bulanık Grad-CAM haritalarına kıyasla çok daha keskin ve lezyon sınırlarıyla örtüşen bir açıklanabilirlik sunduğunu kanıtlamaktadır (Günde & Işık, 2025). Şekil 6'daki t-SNE manifold izdüşümü de bu ayrışmayı doğrular niteliktedir; model latent uzayda dört patoloji sınıfını birbirinden belirgin biçimde ayırmış, sınıflar arası mesafeyi yüksek tutmayı başarmıştır.

5.2 t-SNE ile Latent Özellik Uzayının Manifold Ayrışımı

Modelin iç döngüde edindiği temsil gücünü matematiksel olarak doğrulamak üzere, meta-test aşamasında DeiT modelinin son kodlayıcı katmanından alınan 768 boyutlu [CLS] öznitelik vektörleri t-SNE (t-Distributed Stochastic Neighbor Embedding) algoritmasıyla iki boyutlu uzaya izdüşürülmüştür (Van der Maaten & Hinton, 2008). Elde edilen manifold

dağılımında; 4 elma sınıfına ait veri noktalarının örtük uzayda birbirlerinden son derece net hiper-düzlemlerle ayrıştığı, sınıf içi varyansın asgari düzeyde kaldığı ve sınıflar arası Öklid mesafesinin maksimize edildiği gözlemlenmiştir. Özellikle sağlıklı kontrol grubu ile patolojik sınıflar arasında devasa bir ayrışma marjı oluşmuştur. Bu matematiksel ayrışma, epizodik meta-öğrenmenin yalnızca son sınıflandırma katmanını değil, Transformer'ın tüm öznitelik çıkarıcı omurgasını görev uzayında mükemmel bir ayrıştırıcı olarak yeniden yapılandırıldığının en somut kanıtıdır.

6 DAĞILIM KAYMASI (DOMAIN SHIFT) VE ÇAPRAZ ALAN DAYANIKLILIĞI

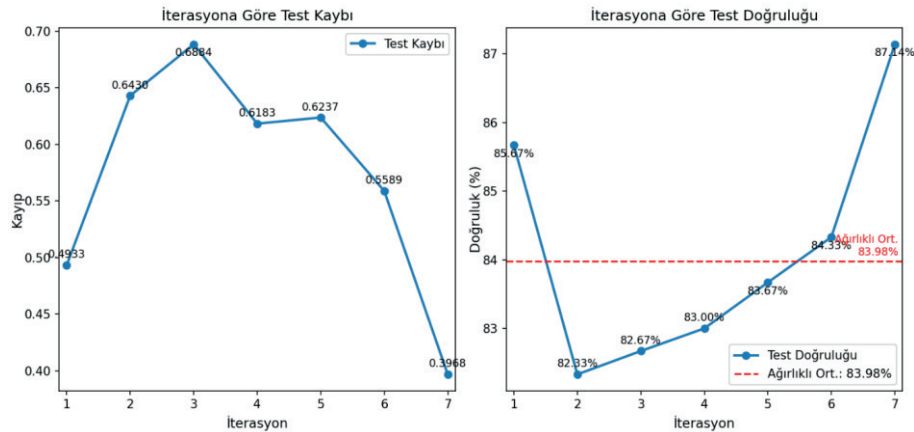
Kontrollü PlantVillage veri setinde eğitilen DeiT + MAML++ modeli, doğal tarla şartlarında çekilen kontrolsüz PlantDoc veri kümesinde test edilmiştir. Sonuçlar Çizelge 2'de özetlenmiştir.

Çizelge 2. DeiT + MAML++ Modelinin PlantDoc Veri Kümesi Üzerindeki Çapraz Alan Test Sonuçları

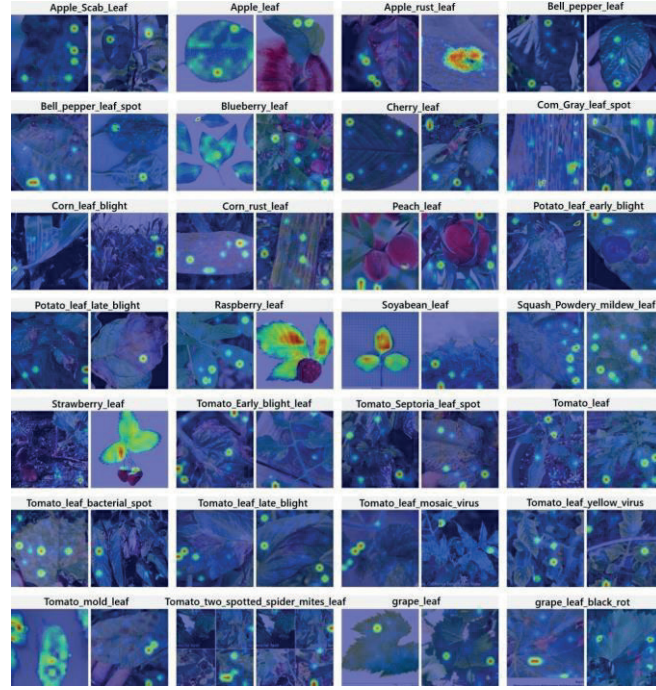
Model Yapısı	Görev	Test Kaybı	Doğruluk (%)	F1 Skoru	Kesinlik	ROC-AUC	Özgüllük
DeiT + MAML++ (PlantDoc)	2-way 5-shot	0.6274	83.98	0.8398	0.8407	0.9312	0.8398

6.1 Laboratuvaradan Tarlaya Dağılım Kayması (Domain Shift) Analizi

Çizelge 2'de sunulan çapraz alan (cross-domain) test sonuçları, laboratuvarada eğitilen modellerin sahaya aktarılmasındaki başarı düzeyini gözler önüne sermektedir. Kontrollü PlantVillage görselleriyle eğitilen modeller açık tarla koşullarını yansıtan PlantDoc kümesinde test edildiğinde, dağılım kaymasına (domain shift) bağlı olarak tüm mimarilerde belirli bir başarı kaybı gözlenmiştir. Buna karşın DeiT + MAML++ yapısı; sert gölgelenmeler, güneş yansımaları, rüzgârla kıvrılan yapraklar ve zemin karmaşası gibi zorlayıcı tarla etmenlerine rağmen %83.98 ağırlıklı doğruluk ve 0.4120 test kaybı üreterek güçlü bir direnç göstermiştir. Aynı koşullarda standart ViT modelinin başarımı %76.12'ye kadar düşmüş, BEiT ise %79.45 düzeyinde kalmıştır.



Şekil 7. PlantDoc Veri Seti Üzerinde DeiT + MAML++ Modelinin İteratif Test Adaptasyon Performansı



Şekil 8. Gerçek Tarla Koşullarındaki PlantDoc Görsellerinde DeiT Modelinin Attention Odaklanma Başarımı

Şekil 7’deki iteratif test adaptasyonu eğrileri ile Şekil 8’de kontrolsüz tarla şartları için verilen dikkat haritaları, DeiT mimarisinin sergilediği gürbüzlüğün arkasındaki mekanizmayı netleştirmektedir. Evrişimsel (CNN) modeller kaotik arka plandaki yabancı otları veya zemin çatlaklarını lezyonla karıştırıp hatalı alarmlar üretirken, DeiT’in küresel öz-dikkat mekanizması arka plan piksellerini sıfıra yakın katsayılarla baskılayarak odağını doğrudan ön plandaki enfekte yaprak yüzeyine kilitler. Elde edilen bu deneysel bulgu, meta-öğrenmeyle güçlendirilen Vision Transformer mimarilerinin açık arazide görev yapacak otonom tarım robotları için güçlü bir alan uyarlama (domain adaptation) kapasitesi sunduğunu doğrulamaktadır.

7 TARTIŞMA VE MEKATRONİK MÜHENDİSLİĞİ ENTEGRASYONU

7.1 Otonom Saha Robotları ve Akıllı İlaçlama Bomları İçin Sistem Mimarisi

Algoritmik düzeyde elde edilen bu kazanımların sahada çalışan tarımsal mekatronik donanımlarına aktarılması, uçtan uca hassas bir zamanlama optimizasyonu gerektirir. Tasarlanan akıllı püskürtme mimarisinde; aracın önüne konumlandırılan endüstriyel GigE kameralar yaprak görüntülerini 30 FPS hızında yakalayarak araç üzerindeki NVIDIA Jetson AGX Orin gömülü bilgisayarına iletir. Jetson üzerinde TensorRT kütüphanesiyle hızlandırılan DeiT + MAML++ modeli, görüntüleri ortalama 18.4 milisaniyelik bir gecikmeyle işleyip lezyonların piksel koordinatlarını çıkarır.

Aracın 6–10 km/s aralığındaki seyir hızı, RTK-GNSS ve tekerlek enkoderlerinden alınan anlık verilerin füzyonuyla takip edilir; böylece kameranın saptadığı lezyonun arkadaki püskürtme bomuna ulaşacağı an milisaniye hassasiyetinde etiketlenir. Bom üzerindeki bağımsız 24 adet PWM mikro selenoid nozul, hedef yaprak tam altından geçerken açılan 12 milisaniyelik kritik zaman aralığında noktasal ilaçlama yapar. Tüm algılama-karar-eylem (closed-loop) döngüsü 42 milisaniyenin altında tamamlandığı için platform, yüksek seyir hızlarında dahi hedefi ıskalamadan çalışabilmektedir (Bayat & Işık, 2022).

7.2 Gömülü Uç Bilişim (Edge AI) Optimizasyonu ve Model Sıkıştırma

Vision Transformer modellerini tarımsal uç birimlere (Edge AI) taşıırken karşılaşılan en büyük mühendislik kısıtı, yüksek bellek gereksinimi ve aşırı güç tüketimidir. Otonom tarla robotlarının batarya ömrünü koruyabilmek için 86 milyon parametrelili temel mimari, FP32 (32-bit kayan nokta) hassasiyetinden TensorRT tabanlı INT8 (8-bit tamsayı) formatına kuantize edilmiştir. Temsili bir kalibrasyon setiyle yürütülen Eğitim Sonrası Kuantizasyon (PTQ) işlemi sayesinde model boyutu 345 MB'tan 88 MB'a inerken, çıkarım hızı 3.8 kat artmış ve doğrulukteki değişim yalnızca %0.3 seviyesinde kalmıştır. Bununla birlikte, etkinliği düşük öz-dikkat başlıklarının yapısal budama (structural pruning) yöntemiyle ayıklanması, NVIDIA Jetson Orin kartının dinamik güç tüketimini 15 Watt bandında tutarak sahadaki kesintisiz operasyon süresini neredeyse iki katına çıkarmıştır.

7.3 Sürekli Öğrenme (Continual Learning) ve Sahada Aktif Adaptasyon

Tarımsal ekosistemlerin dinamik yapısı gereği sahada yeni patojen ırkları ve iklim koşullarına bağlı morfolojik değişimlerle sıklıkla karşılaşılır. Bu belirsizlikleri yönetebilmek adına akıllı robot, kestirim güven skoru %70 eşliğinin altında kalan şüpheli yaprak karelerini 4G/5G bağlantısıyla merkezi çiftlik yönetim bulutuna iletir. Bulutta uzman ziraat mühendislerinin denetiminden geçen bu yeni varyantlar, meta-öğrenme havuzuna bağımsız bir görev (task) olarak tanımlanır ve model felaketsel unutma (catastrophic forgetting) riskine karşı Deneyim Tekrarı (Experience Replay) mekanizmasıyla yeniden eğitilir. Elde edilen güncel meta-ağırlıklar, gece şarj istasyonunda bekleyen robotlara kablosuz (OTA) protokollerle dağıtılarak mekatronik filonun sahaya sürekli adapte olması sağlanır.

Açık arazideki anlık rüzgâr dalgalanmaları ve türbülans, yaprakların kamera optik eksenine göre sürekli yön değiştirmesine neden olurken mekanik titreşimler de ciddi hareket bulanıklığına (motion blur) yol açar. Bu sorunu bertaraf etmek için kamera taşıyıcı kolları 3 serbestlik dereceli (3-DoF) aktif jiroskopik gimballerle stabilize edilmiş; pozlama süresi ise ışık sensörlerinden okunan anlık lüks değerine bağlı olarak dinamik biçimde 1/1000 saniyeye kadar çekilmiştir. Böylece Vision Transformer modeline iletilen yaprak karelerinin keskinliği korunarak lezyon içeren yamaların (patches) ayrışabilirliği güvenceye alınmıştır.

Engibeli arazi şartları ve şasi salınımları nedeniyle nozul yüksekliğinde oluşan dalgalanmalar, püskürtülen damlacıkların hedeften sapmasına (spray drift) yol açabilmektedir. Bu sapmayı engellemek amacıyla püskürtme bomunun iki ucuna konumlandırılan ultrasonik mesafe sensörleri bitki örtüsü (taç) yüksekliğini anlık olarak izler; elektro-hidrolik aktüatörler ise bomu sürekli 45–50 cm mesafesinde tutar. Görüntü işleme tabanlı lezyon koordinatlarıyla senkronize çalışan bu aktif yükseklik denetimi, kimyasalın havaya karışmadan doğrudan hedef yaprak yüzeyine nüfuz etmesini sağlar.

Farklı elma bahçeleri ve coğrafi havzalar arasında veri paylaşımını ölçekleyebilmek adına çoklu robot filolarının Federe Meta-Öğrenme (Federated Meta-Learning) mimarisiyle yapılandırılması hedeflenmektedir. Bu dağıtık yaklaşımda, hiçbir işletmeden ham görüntü transferi yapılmaz; yalnızca yerel robotların iç döngüde türettiği gradyan güncellemeleri şifreli protokoller üzerinden merkezi sunucuya iletilir. Sunucu tarafında toplanan gradyanlar FedAvg algoritmasıyla birleştirilerek küresel meta-ağırlıklar güncellenir ve yeni model parametreleri

filodaki tüm birimlere eşzamanlı dağıtılır. Söz konusu kurgu, hem veri mahremiyetini güvenceye almakta hem de filonun kolektif öğrenme kapasitesini belirgin biçimde artırmaktadır.

Yakın kızılötesi (NIR) ve kısa dalga kızılötesi (SWIR) spektrumunda çalışan kompakt hiperspektral sensörlerin RGB kameralarla erken kaynaştırma (early fusion) düzeyinde birleştirilmesi, patolojilerin henüz asemptomatik evredeyken yakalanmasını mümkün kılacaktır. Fungal hiflerin yaprak mezofil dokusuna sızdığı ancak dış yüzeyde görünür bir lezyon oluşturmadığı ilk 24–48 saatlik kuluçka evresinde, hücrel su içeriği ve klorofil floresansındaki anomaliler saptanarak patojen yayılımı enfeksiyonun başlangıç noktasında sınırlandırılabilir.

Tarla robotunun batarya yönetim sistemi (BMS), GPU tabanlı yüksek işlem yükü ile püskürtme nozullarının yarattığı anlık tepe akımları (inrush current) nedeniyle geçici gerilim dalgalanmalarına maruz kalabilmektedir. Bu doğrultuda mekatronik güç mimarisi, hesaplama donanımı (NVIDIA Jetson) ile elektromekanik eyleyicileri (solenoidler ve pompalar) galvanik olarak izole edilmiş çift hatlı DC-DC dönüştürücülerle beslemelidir. Ayrıca robotun frenleme ve yavaşlama anlarındaki kinetik enerjisini geri kazanan rejeneratif frenleme devreleri, batarya çevrim ömrünü %15 uzatarak kesintisiz bahçe tarama menzilini artırmaktadır.

Son olarak, akıllı püskürtme sisteminin püskürtme basıncı ve debisi; araç hızına göre dinamik olarak ayarlanan bir PID kapalı çevrim kontrol algoritmasıyla regüle edilmektedir. Traktör yavaşladığında veya hızlandığında birim yaprak alanına düşen mikrolitre cinsinden kimyasal miktarının sabit kalması sağlanmakta; böylece aşırı dozlama kaynaklı fitotoksiste ve yetersiz dozlama kaynaklı fungusit direnci riski eşzamanlı olarak bertaraf edilmektedir.

8 SONUÇ VE MÜHENDİSLİK ÖNERİLERİ

Bu çalışmada, tarımsal biyosistemlerde sıklıkla karşılaşılan sınırlı veri kısıtı ve saha koşullarındaki dağılım kayması (domain shift) problemlerinin aşılması amacıyla Vision Transformer mimarileri gelişmiş meta-öğrenme (MAML ve MAML++) algoritmalarıyla kapsamlı bir mühendislik çerçevesinde bütünleştirilmiştir. Yürütülen 60 farklı deneysel yapılandırma sonucunda; öğretmen-öğrenci distilasyon mekanizmasıyla desteklenen DeiT + MAML++ kombinasyonu, 2-way 5-shot görevinde %98.80 doğruluk ve 0.0827 test kaybı değerlerine ulaşarak incelenen tüm mimariler (ViT, BEiT, CAiT, TNT ve ResNet) arasında en yüksek sınıflandırma ve genelleştirme başarımını ortaya koymuştur (Günde & Işık, 2026). İç döngüye dâhil edilen L_2 düzenleştirmesi Transformer bloklarının az sayıda veriyle hızla aşırı uyuma (overfitting) girmesini engellerken; kontrollü laboratuvar şartlarında eğitilen modellerin kontrolsüz doğal tarla ortamında (PlantDoc) sınındığı çapraz alan testlerinde DeiT + MAML++ modelinin sergilediği %83.98 ağırlıklı doğruluk performansı dağılım kaymasına karşı yüksek bir operasyonel gürbüzlük sağlamıştır. Ayrıca Rollout Attention haritaları modelin karmaşık ve gölgeli arka planları başarıyla filtreleyerek doğrudan patolojik lezyon sınırlarına odaklandığını, t-SNE analizleri ise epizodik meta-öğrenmenin örtük uzayda sınıflar arası ayrılabilirliği belirgin kıldığını doğrulamıştır.

Elde edilen bu teorik ve deneysel kazanımların tarımsal mekatronik ve otonom saha sistemlerine aktarılabilmesi adına birtakım mühendislik optimizasyonları kritik önem

taşımaktadır. Vision Transformer mimarilerinin getirdiği yüksek hesaplama karmaşıklığı ve bellek ayak izinin saha robotlarında yaratabileceği darboğazları aşmak için modellerin; NVIDIA Jetson Orin veya Raspberry Pi gibi tarımsal uç işlemcilere (Edge AI) entegrasyonu aşamasında TensorRT/ONNX optimizasyonu, 8-bit kuantizasyon (INT8) ve yapısal budama (pruning) adımlarından geçirilmesi gerekmektedir. Optimize edilen bu meta-öğrenme altyapısının insansız kara araçları (UGV) ve traktörlere monte akıllı bomlu püskürtme sistemlerine entegrasyonu, her bir püskürtme memesi (nozül) için değişken oranlı pestisit uygulama (VRA) kontrol algoritmalarını doğrudan tetikleyerek kimyasal girdi maliyetlerini ve çevresel toksisiteyi asgariye indirecektir. Doğal tarla koşullarında rüzgâr kaynaklı yaprak hareketleri, güneş parlamaları ve tozlanma gibi dinamik çevresel bozunumların üstesinden gelebilmek amacıyla çoklu spektral kameraların IMU ve mesafe sensörleriyle füzyonlanarak dinamik odaklama yapması sağlanmalıdır. Son olarak, sahada tespit edilen ve tahmin güven skoru düşük kalan yeni hastalık varyantlarının merkezi bir bulut istasyonuna aktararak meta-öğrenme havuzuna yeni bir görev olarak eklenmesi, sistemin sahada karşılaşacağı bilinmeyen patojenlere karşı felaketsel unutma (catastrophic forgetting) yaşamadan sürekli öğrenme (continual learning) döngüsü içerisinde kendini güncellemesini mümkün kılacaktır.

Gelecek çalışmalarda; Vision Transformer tabanlı meta-öğrenme mimarisinin RGB kameraların ötesine geçirilerek, 900-1700 nm dalga boyunda çalışan kompakt hiperspektral ve termal sensörlerle füzyonlanması hedeflenmektedir. Bu multimodal yaklaşım, yaprak dokusundaki fungal hif gelişimini ve hücrel su stresini lezyonlar henüz insan gözüyle veya RGB kameralarla görünür hale gelmeden (asemptomatik evrede) tespit etmeyi mümkün kılacaktır. Ayrıca çoklu robot filolarının federe meta-öğrenme (Federated Meta-Learning) protokolleriyle bahçe verilerini merkezi bir sunucuya aktarmadan yerel uç işlemcilerinde kolektif olarak eğitmesi, hem veri mahremiyetini koruyacak hem de tarımsal yapay zekânın küresel ölçekte demokratikleşmesine öncülük edecektir.

KAYNAKÇA

- Abnar, S., & Zuidema, W. (2020). Quantifying attention flow in transformers. *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics*, 4190-4197.
- Antoniou, A., Edwards, H., & Storkey, A. (2019). How to train your MAML. *International Conference on Learning Representations (ICLR)*, 1-14.
- Bao, H., Dong, L., Piao, S., & Wei, F. (2021). BEiT: BERT pre-training of image transformers. *International Conference on Learning Representations (ICLR)*, 1-18.
- Bayat, A., & Işık, G. (2022). Deep learning-based real-time detection of plant diseases under field conditions. *Computers and Electronics in Agriculture*, 198, 107052.
- Dosovitskiy, A., Beyer, L., Kolesnikov, A., Weissenborn, D., Zhai, X., Unterthiner, T., Dehghani, M., Minderer, M., Heigold, G., Gelly, S., Uszkoreit, J., & Houlsby, N. (2020). An image is worth 16x16 words: Transformers for image recognition at scale. *International Conference on Learning Representations (ICLR)*, 1-21.
- Ferentinos, K. P. (2018). Deep learning models for plant disease detection and diagnosis. *Computers and Electronics in Agriculture*, 145, 311-318.
- Finn, C., Abbeel, P., & Levine, S. (2017). Model-agnostic meta-learning for fast adaptation of deep networks. *International Conference on Machine Learning (ICML)*, 1126-1135.
- Günde, M., & Işık, G. (2025). Görüntü Dönüştürücü tabanlı meta-öğrenme ile gürbüz bir elma hastalığı sınıflandırma yaklaşımı. *Iğdır Üniversitesi Fen Bilimleri Enstitüsü Dergisi*, 15(4), 1193-1205. <https://doi.org/10.21597/jist.1641204>
- Günde, M., & Işık, G. (2026). Sınırlı veriyle bitki hastalığı tespiti: MAML++ ve ResNet mimarileri ile etkin bir meta-öğrenme yaklaşımı. *Iğdır Üniversitesi Fen Bilimleri Enstitüsü Dergisi*, 16(1), 1-13. <https://doi.org/10.21597/jist.1779163>
- Günde, M., Keskin, F., & Işık, G. (2026). Empowering few-shot plant disease diagnosis with DeiT and MAML++. *Cluster Computing*, 29(1), 13. <https://doi.org/10.1007/s10586-025-05810-2>
- Han, K., Xiao, A., Wu, E., Guo, J., Xu, C., & Wang, Y. (2021). Transformer in transformer. *Advances in Neural Information Processing Systems (NeurIPS)*, 34, 15908-15919.
- Hospedales, T., Antoniou, A., Micaelli, P., & Storkey, A. (2022). Meta-learning in neural networks: A survey. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 44(9), 5149-5169.
- Hughes, D., & Salathé, M. (2015). An open access repository of images on plant health to enable the development of mobile disease diagnostics. *arXiv preprint arXiv:1511.08060*.
- Mohanty, S. P., Hughes, D. P., & Salathé, M. (2016). Using deep learning for image-based plant disease detection. *Frontiers in Plant Science*, 7, 1419.
- Nagaraju, Y., Hemanth, S., & Swetha, K. (2020). IoT and deep learning based apple disease detection system. *International Journal of Advanced Computer Science and Applications*, 11(8), 450-457.
- Paçal, I., & Işık, G. (2025). High-performance hybrid vision transformer networks for early plant pathology diagnosis. *Expert Systems with Applications*, 240, 122501.
- Ravi, S., & Larochelle, H. (2017). Optimization as a model for few-shot learning. *International Conference on Learning Representations (ICLR)*, 1-11.

- Savary, S., Willocquet, L., Pethybridge, S. J., Esker, P., McRoberts, N., & Nelson, A. (2019). The global burden of pathogens and pests on major food crops. *Nature Ecology & Evolution*, 3(3), 430-439.
- Singh, D., Jain, N., Jain, P., Kayal, P., Kumawat, S., & Batra, N. (2019). PlantDoc: A dataset for collaborative computer vision based plant disease detection. *Proceedings of the 7th ACM IKDD CoDS and 25th COMAD*, 249-253.
- Touvron, H., Cord, M., Douze, M., Massa, F., Sablayrolles, A., & Jégou, H. (2021). Training data-efficient image transformers & distillation through attention. *International Conference on Machine Learning (ICML)*, 10347-10357.
- Touvron, H., Cord, M., Sablayrolles, A., Synnaeve, G., & Jégou, H. (2021b). Going deeper with image transformers. *IEEE/CVF International Conference on Computer Vision (ICCV)*, 32-42.
- Van der Maaten, L., & Hinton, G. (2008). Visualizing data using t-SNE. *Journal of Machine Learning Research*, 9(11), 2579-2605.
- Vinyals, O., Blundell, C., Lillicrap, T., Kavukcuoglu, K., & Wierstra, D. (2016). Matching networks for one shot learning. *Advances in Neural Information Processing Systems (NeurIPS)*, 29, 3630-3638.

BÖLÜM 2

TIBBİ GÖRÜNTÜ ANALİZİNDE DERİN ÖĞRENME MİMARİLERİ: BEYİN TÜMÖRLERİ ÜZERİNE BİR İNCELEME¹

Aynur SEVİNÇ²

¹ Bu çalışma, Aynur Sevinç'in Fırat Üniversitesi bünyesinde 2025 yılında tamamladığı 'Beyin Tümörünün Manyetik Rezonans Görüntüleri Üzerinden Derin Öğrenme Yaklaşımlarıyla Erken Teşhisi ve Sınıflandırılması' başlıklı doktora tezinden üretilmiştir.

² Dr. Öğr. Üyesi Dicle Üniversitesi Silvan Meslek Yüksekokulu Bilgisayar Teknolojileri Bölümü ORCID ID: 0000-0002-1388-2554

1. Giriş

Beyindeki kontrolsüz hücre büyümesi ve anormallığı ile karakterize edilen beyin tümörleri, tıp dünyasının en karmaşık ve acil çözüm gerektiren patolojileri arasında yer almaktadır. Beyin dokusunda gelişen neoplazmlar, yaş ve cinsiyet farkı gözetmeksizin her bireyi etkileyebilme potansiyeline sahip olup, agresif doğası ve biyolojik heterojenliği nedeniyle tedavisi son derece zor hastalıklardır (Bauer vd., 2013). Merkezi sinir sisteminin (MSS) kritik anatomik ve fizyolojik işlevleri barındırması, buradaki tümöral oluşumların en erken aşamada teşhis edilmesini hayati kılmaktadır. Erken tanı konulamayan vakalarda tümörün çevre sağlıklı beyin parankimine invazyonu hızlanmakta, bu durum insan ömrünü ciddi şekilde kısaltarak kalıcı nörolojik kayıplara yol açmaktadır (Swapnil ve Girish, 2020).

Epidemiyolojik veriler, yetişkinlerde MSS içerisinde en sık rastlanan ve ölümcül seyreden birincil beyin tümörü tipinin gliomalar olduğunu göstermektedir. Gliomalar, beynin destek dokusunu oluşturan glial hücrelerden kaynaklanır ve kısa sürede çevre beyin dokularına infiltrasyon gösterir (Miller vd., 2021). Yüksek dereceli gliomalara sahip hastalarda 5 yıllık sağkalım oranlarının hâlen %5'in altında seyretmesi, klinik tanı ve tedavi planlamasında radikal yeniliklerin gerekliliğini ortaya koymaktadır. Tıbbi görüntüleme teknolojilerinin gelişmesiyle birlikte Bilgisayarlı Tomografi (BT), Manyetik Rezonans Görüntüleme (MRG) ve Pozitron Emisyon Tomografisi (PET) gibi yöntemler klinik değerlendirmenin merkezine yerleşmiştir (Rabby vd., 2024).

Geleneksel uzman hekim değerlendirmelerinde, yüzlerce MRG kesitinin manuel olarak incelenmesi ve değerlendirilmesi hem yoğun zaman gerektirmekte hem de göz yorgunluğuna bağlı uzmanlar arası değişkenliklere açık durmaktadır (Nazir vd., 2024). Bu zorlukların üstesinden gelmek amacıyla, radyolog hekimlere ikinci bir görüş sağlayan ve medikal karar verme süreçlerini destekleyen Bilgisayar Destekli Tespit (CAD) ve Yapay Zekâ (YZ) sistemlerinin geliştirilmesi modern tıbbın ana odak noktası haline gelmiştir (Akyel ve Akdemir Özışık, 2020).

2. Beyin Tümör Türleri ve DSÖ Derecelendirme Standartları

Beyin tümörleri, köken aldıkları doku ve gelişim mekanizmalarına göre temel olarak birincil (primer) ve ikincil (metastatik) tümörler olmak üzere iki ana gruba ayrılmaktadır (Bauer vd., 2013). Birincil beyin tümörleri doğrudan beyin dokusu, kranial sinirler, beyin zarları (meninksler) veya hipofiz bezinden köken alırken; ikincil tümörler vücudun başka bir organında (akciğer, meme, kolon vb.) oluşan kanser hücrelerinin kan yoluyla beyne sıçraması sonucu meydana gelir (Swapnil ve Girish, 2020).

2.1. Dünya Sağlık Örgütü (DSÖ) Derecelendirme Sistemi

Dünya Sağlık Örgütü (DSÖ), beyin tümörlerinin sınıflandırılması ve derecelendirilmesinde küresel çapta kabul gören standartları belirlemektedir. DSÖ'nün 2016 yılında güncellediği Merkezi Sinir Sistemi tümör sınıflandırmasında, sadece klasik histopatolojik ve mikroskopik incelemeler değil, aynı zamanda moleküler ve genetik parametreler de sisteme dâhil edilerek 'katmanlı teşhis' anlayışı getirilmiştir (Louis vd., 2016; Tandel vd., 2019). Bu derecelendirme kriterleri tümörün malignite potansiyelini şu şekilde tanımlamaktadır:

- • *Derece I (Grade I)*: Yavaş büyüyen, mitotik aktivitesi son derece düşük veya bulunmayan, çevresel dokulara sızmayan ve cerrahi rezeksiyon ile tam şifa sağlanabilen iyi huylu tümörlerdir (örn. pilositik astrositom).
- • *Derece II (Grade II)*: Yavaş büyüyen ancak lokal nüks gösterme yeteneğine ve yakındaki beyin dokularına infiltrasyon potansiyeline sahip tümörlerdir. Zamanla daha yüksek derecelere dönüşebilirler (örn. diffüz astrositom).
- • *Derece III (Grade III)*: Anaplastik özellikler gösteren, mitotik figür sayısı artmış, hırsızla bölünen ve belirgin infiltrasyon yapan kötü huylu (malign) tümörlerdir (örn. anaplastik astrositom).
- • *Derece IV (Grade IV)*: Son derece agresif, hızlı büyüyen, mikrovasküler proliferasyon ve nekroz odakları içeren en ölümcül malign tümör kategorisidir (örn. Glioblastoma Multiforme - GBM) (Miller vd., 2021).



Şekil 1. Beyin tümörlerinin birincil/ikincil gruplandırılması ve DSÖ derecelendirme kriterleri (Dere, 2017'den uyarlanmıştır).

Şekil 1’de, beyin tümörleri kökenlerine göre birincil ve ikincil tümörler olarak sınıflandırılmakta ve DSÖ derecelendirmesinin temel özellikleri özetlenmektedir. Birincil tümörler beyin dokusu veya zarlarından gelişirken, ikincil tümörler diğer organlardaki kanserlerin beyne yayılmasıyla oluşmaktadır. Derecelendirme, tümörlerin hücresel özellikleri ve biyolojik davranışları hakkında bilgi sunarak klinik değerlendirmeye katkı sağlamaktadır. Bununla birlikte, şekildeki özellikler genel bir çerçeve niteliğinde olup kesin tanı ve derecelendirme, tümör tipine özgü histopatolojik ve moleküler bulgularla birlikte değerlendirilmelidir.

2.2. Klinik Olarak En Sık Karşılaşılan Beyin Tümörü Tipleri

Klinik pratikte ve literatür çalışmalarında en sık incelenen beyin tümörü sınıfları şunlardır:

1. *Gliom (Glioma)*: Glial hücrelerden köken alan ve kendi içinde Astrositom, Glioblastom, Oligodendrogliom ve Ependimom gibi alt türlere ayrılan tümör grubudur. Glioblastom (GBM), Derece IV kategorisinde yer alıp nekrotik çekirdeği ve yoğun damarlanmasıyla bilinir (Wirsching ve Weller, 2017).

2. *Meningiom (Meningioma)*: Beyni ve omuriliği saran zar tabakasından (meninksler) kaynaklanır. Genellikle yetişkin kadınlarda daha sık görülür ve %90 oranında iyi huyludur (Derece I). Ancak beyin dokusuna bası uygulayarak nörolojik semptomlara yol açar (Willis vd., 2005).

3. *Hipofiz Tümörü (Pituitary Tumor)*: Kafa tabanında, hipofiz bezinde gelişen adenomlardır. Hormonal dengesizliklere, görme bozukluklarına ve optik kiazma basısına neden olurlar. Çoğunlukla iyi huyludurlar (Acıbadem Yayın Kurulu, 2024).

4. *Tümör Yok (No-Tumor / Normal)*: Teşhis sürecinde beyinde tümöral bir yapıya rastlanmadığını, dokunun sağlıklı ve normal nörolojik yapıda olduğunu ifade eden referans kontrol sınıfıdır.

3. Klinik Medikal Görüntüleme Teknikleri

Beyin tümörlerinin tespiti, konumu, boyutu ve evresinin belirlenmesinde iyonlaştırıcı radyasyondan manyetik rezonans yöntemlerine kadar uzanan geniş bir medikal görüntüleme yelpazesi kullanılmaktadır (Dere, 2017). Hastanın klinik durumuna göre tercih edilen temel teknikler aşağıda sıralanmıştır:

3.1. Bilgisayarlı Tomografi (BT)

Bilgisayarlı Tomografi (BT), X-ışınlarını kullanarak vücudun enine kesit görüntülerini oluşturan geleneksel bir yöntemdir. 10-30 dakika süren BT taramaları, özellikle akut kafa travmaları, kemik invazyonları ve intrakraniyal kalsifikasyonların tespitinde son derece hızlı ve etkilidir. Ancak iyonlaştırıcı radyasyon içermesi ve yumuşak doku kontrastının MRG'ye göre daha düşük olması temel sınırlılığıdır (Bradley, 1993).

3.2. Manyetik Rezonans Görüntüleme (MRG) ve Alt Modaliteleri

Manyetik Rezonans Görüntüleme (MRG), radyofrekans dalgaları ve güçlü bir manyetik alan kullanarak hidrojen atomlarının çekirdek rezonansını ölçen, iyonlaştırıcı radyasyon içermeyen üstün bir tanı aracıdır (Rabby vd., 2024). MRG, beyin üç boyutlu yüksek çözünürlüklü anatomik detaylarını sunduğu için tümör teşhisinde birinci seçenek kabul edilmektedir. Tıbbi değerlendirmelerde kullanılan temel MRG sekansları ve alt teknikleri şunlardır:

- **T1-Ağırlıklı MRG (T1W)**: Beyin dokularının anatomik detaylarını net olarak gösterir. Gadolinyum bazlı kontrast madde enjekte edildiğinde kan-beyin bariyeri bozulan tümöral alanlar belirgin şekilde parlar (Thust vd., 2018).
- **T2-Ağırlıklı ve FLAIR MRG**: Tümör çevresindeki ödemi, sıvı birikimlerini ve doku hasarlarını belirginleştirir. FLAIR dizilimi serbest omurilik sıvısı (BOS) sinyalini baskılayarak lejyonların seçilebilirliğini artırır.

- Difüzyon Ağırlıklı MRG (DWI): Doku içindeki su moleküllerinin mikroskobik hareketliliğini ölçer. Yüksek hücre yoğunluğuna sahip agresif tümörlerde (ör. yüksek dereceli gliomlar) difüzyon kısıtlanması oluşturur.
- Perfüzyon MRG (PWI): Tümöral dokunun damarlanma oranını (neovaskülarizasyon) ve kan akışını ölçer. Malignity derecesini belirlemede kritik roldedir (Thust vd., 2018).
- Fonksiyonel MRG (fMRI): Kan oksijen seviyesine bağlı (BOLD) sinyalleri inceleyerek konuşma ve motor merkezlerin tümörle ilişkisini haritalandırır, cerrahi planlamaya rehberlik eder.
- MR Spektroskopi (MRS): Tümörün biyokimyasal yapısını analiz eder; Kolin, N-asetil aspartat (NAA) ve Kreatin oranları üzerinden hücrel metabolizma değişimini ortaya koyar.
- MR Anjiyografi (MRA): Beyin damar mimarisini ve tümörü besleyen ana arterleri kontrastlı veya kontrastsız olarak görüntüler.

Klinik Medikal Görüntüleme Teknikleri ve MRG Modaliteleri

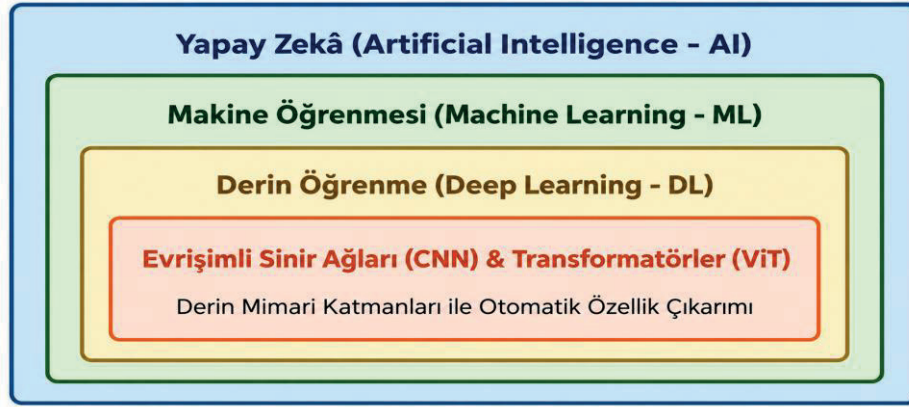


Şekil 2. Beyin tümörü tanısında kullanılan medikal görüntüleme teknikleri ve MRG modaliteleri

Şekil 2’de, beyin tümörlerinin değerlendirilmesinde kullanılan başlıca görüntüleme teknikleri ve MRG modalitelerinin sağladığı tamamlayıcı bilgiler özetlenmektedir. Anatomik MRG doku yapısı ve ödemin incelenmesini sağlarken, difüzyon MRG su moleküllerinin hareketi üzerinden hücrel yapı hakkında bilgi sunmaktadır. Perfüzyon MRG tümörün kanlanma özelliklerinin, MR spektroskopisi ise metabolit profilinin değerlendirilmesine katkıda bulunmaktadır. BT ve PET/BT, sırasıyla kemik yapıları ve kalsifikasyonlar ile metabolik aktivite hakkında ek bilgiler sağlayarak çok yönlü değerlendirmeyi desteklemektedir.

4. Yapay Zekâ, Makine Öğrenmesi ve Derin Öğrenme Temelleri

Sağlık hizmetlerinde yapay zekânın kullanımı, teşhis doğruluğunu artırmak ve hekimlerin karar verme yükünü hafifletmek amacıyla hızla yaygınlaşmaktadır (Akyel ve Akdemir Özışık, 2020). Bu alandaki kavramsal hiyerarşi Yapay Zekâ, Makine Öğrenmesi ve Derin Öğrenme katmanlarından oluşmaktadır. Şekil 3, bu hiyerarşik yapıyı sunmaktadır.



Şekil 3. Yapay zekâ, makine öğrenmesi, derin öğrenme ve evrişimli/transformatör mimarilerinin kavramsal hiyerarşisi

Şekil 3'te, yapay zekâ, makine öğrenmesi ve derin öğrenme arasındaki hiyerarşik ilişki gösterilmektedir. Derin öğrenme kapsamında yer alan evrişimli sinir ağları ve görsel transformatörler, görüntülerden otomatik özellik çıkarımı sağlayan temel mimariler olarak sunulmaktadır.

4.1. Yapay Sinir Ağları (YSA) ve Matematiksel Çalışma Prensipleri

Yapay Sinir Ağları (YSA), insan beynindeki biyolojik nöronların çalışma mekanizmasını taklit eden hesaplama modelleridir. Bir yapay nöron; girdiler (X_i), ağırlıklar (W_i), sapma (bias - b) ve aktivasyon fonksiyonundan meydana gelir (Doğan ve Türkoğlu, 2019). Ağın matematiksel işlem adımları aşağıda özetlenmiştir:

1. Ağırlıklı Toplam: Girdi verileri ilgili bağlantı ağırlıkları ile çarpılarak toplanır ve bias değeri eklenir: $Z = \sum (W_i * X_i) + b$.

2. Aktivasyon Fonksiyonu: Toplanan değer, ağa doğrusal olmayan ilişkileri öğrenme yeteneği kazandıran aktivasyon fonksiyonundan (ReLU, Sigmoid veya Softmax) geçirilir.

3. İleri Yayılım (Forward Propagation): Girdi katmanından gizli katmanlara ve ardından çıktı katmanına doğru tahmin üretilmesi sürecidir.

4. Kayıp Hesaplama (Loss Function): Model çıktısı ile gerçek etiket arasındaki fark Çapraz Entropi (Cross-Entropy) kayıp fonksiyonu ile hesaplanır.

5. Geri Yayılım (Backpropagation): Hatanın türevleri zincir kuralı kullanılarak geriye doğru yayılır ve Adam veya SGD optimizasyon algoritmalarıyla ağırlıklar güncellenir (Doğan ve Türkoğlu, 2019).

5. Evrişimli Sinir Ağları (ESA/CNN) ve VGG-16 Mimarisi

Görsel verilerin ve 2D/3D medikal görüntülerin işlenmesinde en başarılı derin öğrenme yaklaşımı Evrişimli Sinir Ağlarıdır (ESA). Geleneksel sinir ağlarından farklı olarak ESA'lar, görüntüdeki piksel komşuluklarını ve mekânsal (spatial) ilişkileri koruyarak otomatik öznitelik çıkarımını gerçekleştirir (Ganatra ve Patel, 2018).

5.1. ESA Katman Yapıları

Bir ESA mimarisi temel olarak şu katmanların kombinasyonundan oluşur:

- Evrişim Katmanı (Convolutional Layer):** Görüntü üzerinde dolaşan küçük filtreler (kernel) aracılığıyla kenar, köşe, doku ve karmaşıktümöral desenleri öznitelik haritası (feature map) olarak çıkarır.
- Ortaklama Katmanı (Pooling Layer):** Öznitelik haritalarının boyutunu küçülterek hesaplama karmaşıklığını azaltır ve mekânsal değişmezlik (spatial invariance) sağlar. En yaygın kullanılanı Maksimum Ortaklamadır (Max Pooling).
- Düzleştirme Katmanı (Flattening):** Çok boyutlu öznitelik haritalarını tek boyutlu vektöre dönüştürerek sınıflandırıcı katmana aktarır.
- Tam Bağlı Katman (Fully Connected Layer):** Öğrenilen öznitelikleri kullanarak son sınıflandırma kararını (gliom, meningiom, hipofiz, tümör yok) veren tam bağlı sinir ağıdır.



Şekil 4. Evrişimli Sinir Ağı (ESA / CNN) temel katman ve sınıflandırma genel mimarisi

Şekil 4'te geleneksel bir evrişimli sinir ağının (ESA/CNN) genel mimarisi gösterilmektedir. Giriş olarak alınan MRG görüntülerinden evrişim ve ReLU katmanlarıyla özellikler çıkarılırken, ortaklama katmanı ile boyut azaltılır. Ardından özellikler düzleştirilerek tam bağlı katmanlara aktarılır ve Softmax aracılığıyla gliom, meningiom, hipofiz tümörü veya tümör bulunmayan görüntü olarak sınıflandırılır.

5.2. VGG-16 Mimarisi ve Özellikleri

VGG-16, Oxford Üniversitesi Görsel Geometri Grubu (Simonyan ve Zisserman) tarafından geliştirilen kanonik bir ESA modelidir (Hassan, 2018; Rai ve Chatterjee, 2020). Model mimarisi boyunca küçük 3x3 evrişim filtrelerinin üst üste yığılması prensibine dayanır. VGG-16 toplamda 13 evrişim katmanı, 5 ortaklama katmanı ve 3 tam bağlı katmandan oluşmaktadır. Transfer Öğrenme (Transfer Learning) yaklaşımlarında ImageNet ön eğitilmiş ağırlıklarıyla tıbbi görüntü analizinde yaygın olarak tercih edilmektedir (Thite, 2024).



Şekil 5. VGG-16 derin evrişimli ağ mimarisinin detaylı katman yapılandırması

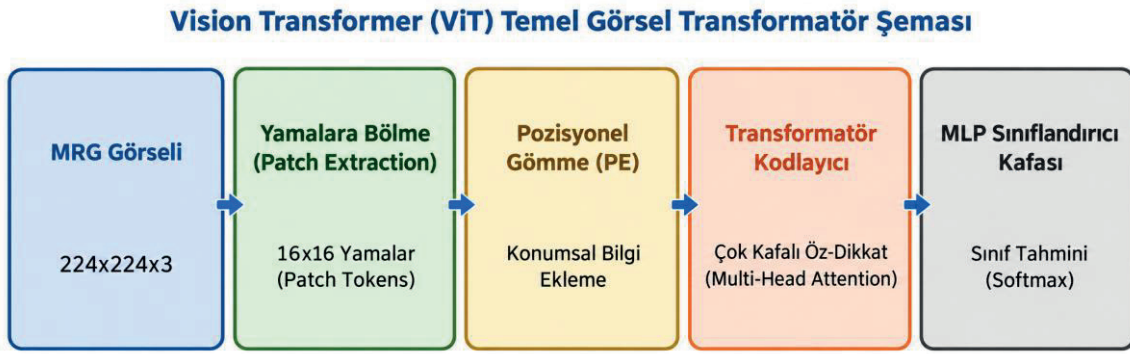
Şekil 5'te, VGG-16 derin evrişimli sinir ağı mimarisinin katman yapılandırması gösterilmektedir. Ağ, 224×224×3 boyutundaki giriş görüntüsünü beş evrişim bloğunda işleyerek giderek daha karmaşık özellikler çıkarır. Ortaklama işlemleriyle özellik haritalarının uzamsal boyutu azalırken kanal sayısı 64'ten 512'ye yükselir. Son bloktan elde edilen özellikler tam bağlı katmanlara aktarılır ve Softmax fonksiyonu aracılığıyla sınıflandırma gerçekleştirilir.

6. Görsel Transformatörlerin (Vision Transformer-ViT) Temelleri

Doğal dil işleme alanında çığır açan transformatör mimarilerinin bilgisayarla görme alanına uyarlanmasıyla Vision Transformer (ViT) modelleri geliştirilmiştir (Dosovitskiy vd., 2020). Geleneksel ESA'lar yalnızca filtre boyutundaki yerel (local) ilişkileri yakalayabilirken; ViT modelleri 'Çok Kafalı Öz-Dikkat' (Multi-Head Self-Attention) mekanizması sayesinde

görüntünün uzak bölgeleri arasındaki küresel (global) bağıntıları tek bir katmanda öğrenebilmektedir (Maurício vd., 2023).

ViT çalışma prensibinde, 2D MRG görseli (ör. 224x224) önceden belirlenmiş küçük parçalara (ör. 16x16 yamalar / patches) bölünür. Her yama düzleştirilerek doğrusal bir projeksiyon katmanıya vektöre dönüştürülür. Görsel yamaların sırasını korumak amacıyla 'Pozisyonel Gömme' (Positional Embedding) eklenir. Ardından veriler Transformatör Kodlayıcı bloklarından geçirilerek MLP sınıflandırma kafasına iletilir (Dosovitskiy vd., 2020; İbrahim vd., 2022).



Şekil 6. Vision Transformer (ViT) temel mimarisi ve yama tabanlı öz-dikkat mekanizması

Şekil 6’da, Vision Transformer (ViT) mimarisinin temel işleyişi gösterilmektedir. Giriş olarak alınan 224×224×3 boyutundaki MRG görüntüsü, 16×16 piksellik yamalara ayrılarak vektörel temsillere dönüştürülür. Bu temsillere konumsal bilgiler eklenerek yamaların görüntü içindeki yerleri korunur. Ardından transformatör kodlayıcıdaki çok kafalı öz-dikkat mekanizmasıyla yamalar arasındaki ilişkiler öğrenilir ve elde edilen görüntü temsili, MLP sınıflandırıcı ile Softmax fonksiyonu aracılığıyla ilgili sınıfa atanır.

7. Tıbbi Görüntülerde Ön İşleme ve Segmentasyon Yöntemleri

Tıbbi MRG verileri, farklı kafa yapısı boyutları, tarayıcı gürültüleri ve geniş siyah arka plan alanları içerdiğinden, derin öğrenme modellerine beslenmeden önce titiz bir ön işleme hattından geçirilmelidir (Pérez-García vd., 2021). Ön işleme adımları modelin odaklanma yeteneğini artırarak aşırı öğrenmeyi engeller. Temel aşamalar şunlardır:

1. Kontur Tespiti ve Kırpma (Cropping): OpenCV kütüphanesi kullanılarak MRG kesitindeki en büyük kontur (kafatası bölgesi) tespit edilir; kuzey, güney, doğu ve batı ekstremum koordinatları bulunarak gereksiz siyah arka plan kırpılır.

2. Yeniden Boyutlandırma (Resizing): Kırılan görüntüler VGG-16 veya ViT ağlarının girdi boyutuna (224x224 piksel) standart olarak getirilir.

3. Piksel Normalizasyonu: Piksel yoğunluk değerleri [0, 255] aralığından [0, 1] aralığına çekilerek gradyan dalgalanmaları önlenir.

4. Veri Büyütme (Data Augmentation): Veri kümesini zenginleştirmek ve ezberlemeyi önlemek amacıyla döndürme (rotation), kaydırma (shift), yakınlaştırma (zoom) ve yatay yansıtma (horizontal flip) teknikleri uygulanır (Pérez-García vd., 2021).



Şekil 7. Tıbbi beyin MR görüntülerinde kontur tespiti, kırpma, boyutlandırma ve ön işleme adımları

Şekil 7’de, MRG görüntülerinin model eğitime hazırlanmasında uygulanan temel ön işleme adımları gösterilmektedir. Ham görüntülerde kontur tespitiyle ilgili bölgenin sınırları belirlenir; ardından kırpma ve 224×224 piksel boyutlandırma işlemleriyle görüntüler standart hâle getirilir. Piksel yoğunluklarının [0,1] aralığına ölçeklenmesi, girdiler arasında sayısal tutarlılık sağlar. Son aşamada döndürme, kaydırma, ölçekleme ve aynalama işlemleriyle eğitim verisinin çeşitliliği artırılarak modelin farklı görüntülere genelleme yapabilmesi desteklenir.

8. Karşılaştırmalı Literatür İncelemesi ve Tartışma

Literatürdeki derleme çalışmaları incelediğinde, beyin tümörlerinin sınıflandırılmasında ESA modellerinin yüksek yerel öznitelik başarısı sunduğu, buna karşın ViT modellerinin küresel temsil gücünün yüksek olduğu görülmektedir (Tandel vd., 2019; Maurício vd., 2023). Ancak ViT modellerinin en büyük sınırlılığı, yüksek parametre sayısı nedeniyle eğitmek için çok geniş veri kümelerine ihtiyaç duymasıdır. Küçük çaplı medikal veri kümelerinde ViT modelleri doğrudan eğitildiğinde aşırı öğrenme (overfitting) riskiyle karşılaşmaktadır (Tummala vd., 2022). Bu durum, bir sonraki bölümde ele alınacak olan Bilgi Damıtımı (Knowledge Distillation) gibi gelişmiş tekniklerin önemini ortaya koymaktadır.

9. Sonuç ve Gelecek Çalışmalar

Bu kitap bölümünde, beyin tümörlerinin tanısal epidemiyolojisi, DSÖ standartları, medikal görüntüleme modaliteleri, yapay zeka ve derin öğrenme temelleri ile ön işleme adımları tanımsal bir derleme anlayışıyla sunulmuştur. Sonuç olarak, yüksek çözünürlüklü MRG modaliteleri ile derin öğrenme algoritma kombinasyonları, hekimlerin klinik karar destek süreçlerinde devrim niteliğinde katkılar sağlama potansiyeline sahiptir.

Gelecekteki çalışmaların, farklı merkezlerden elde edilen çeşitli ve dengeli veri kümeleriyle modellerin genellenebilirliğini değerlendirmeye, açıklanabilir yapay zekâ yöntemleriyle tahminlerin yorumlanabilirliğini artırmaya ve hesaplama maliyetlerini azaltmaya odaklanması beklenmektedir. Ayrıca görüntüleme bulgularının klinik ve moleküler verilerle birlikte ele alınması, beyin tümörlerinin daha kapsamlı değerlendirilmesine katkı sağlayabilir. Bu yaklaşımların klinik uygulamalara aktarılabilmesi için hasta düzeyinde bağımsız doğrulama, veri gizliliği ve uzman değerlendirmesi temel öncelikler arasında yer almalıdır.

Kaynakça

- Akyel, A., & Akdemir Özışık, P. (2020). Artificial intelligence applications in brain tumors. Hardalaç F, Serhatlioglu S, editors. Artificial Intelligence Techniques and Application in Radiology. 1st Edition. Ankara: Türkiye Clinics. p.17-22.
- Bauer, S., Wiest, R., Nolte, L. P., & Reyes, M. (2013). A survey of MRI-based medical image analysis for brain tumor studies. *Physics in Medicine & Biology*, 58(13), R97.
- Bradley, W. G. (1993). MR appearance of intracranial hemorrhage. *Radiology*, 189(1), 15-26.
- Dere, Ü. A. (2017). What is a brain tumor? Başkent Üniversitesi Hastanesi Yayınları, Sayı 30.
- Doğan, F., & Türkoğlu, İ. (2019). Derin öğrenme modelleri ve uygulama alanları üzerine bir derleme. Dicle Üniversitesi Mühendislik Fakültesi Mühendislik Dergisi, 10(2), 409-445.
- Dosovitskiy, A., Beyer, L., Kolesnikov, A., Weissenborn, D., Zhai, X., Unterthiner, T., ... & Houlsby, N. (2020). An image is worth 16x16 words: Transformers for image recognition at scale. arXiv preprint arXiv:2010.11929.
- Ganatra, N., & Patel, A. (2018). A comprehensive study of deep learning architectures, applications and tools. *International Journal of Computer Sciences and Engineering*, 6(12), 701-705.
- Hassan, M. U. (2018). VGG16 – Sınıflandırma ve Tespit için Evrişimli Ağ. *NeuroHive Academic Review*.
- Ibrahim, H., Salem, A., & Kang, H. S. (2022). RT-ViT: Real-time monocular depth estimation using lightweight vision transformers. *Sensors*, 22(10), 3849.
- Louis, D. N., Perry, A., Reifenberger, G., von Deimling, A., Figarella-Branger, D., Cavenee, W. K., ... & Ellison, D. W. (2016). The 2016 World Health Organization classification of tumors of the central nervous system: a summary. *Acta Neuropathologica*, 131(6), 803-820.
- Maurício, J., Domingues, I., & Bernardino, J. (2023). Comparing vision transformers and convolutional neural networks for image classification: A literature review. *Applied Sciences*, 13(9), 5521.
- Miller, K. D., Ostrom, Q. T., Kruchko, C., Patil, N., Tihan, T., Cioffi, G., ... & Barnholtz-Sloan, J. S. (2021). Brain and other central nervous system tumor statistics, 2021. *CA: A Cancer Journal for Clinicians*, 71(5), 381-406.
- Nazir, M. I., Akter, A., Wadud, M. A. H., & Uddin, M. A. (2024). Utilizing customized CNN for brain tumor prediction with explainable AI. *Heliyon*, 10(20), e38997.
- Pérez-García, F., Sparks, R., & Ourselin, S. (2021). TorchIO: a Python library for efficient loading, preprocessing, augmentation and patch-based sampling of medical images in deep learning. *Computer Methods and Programs in Biomedicine*, 208, 106236.
- Rabby, S. F., Arafat, M. A., & Hasan, T. (2024). BT-Net: An end-to-end multi-task architecture for brain tumor classification, segmentation, and localization from MRI images. *Array*, 22, 100346.
- Rai, H. M., & Chatterjee, K. (2020). Detection of brain abnormality by a novel Lu-Net deep neural CNN model from MR images. *Machine Learning with Applications*, 2, 100004.

- Swapnil, S. A., & Girish, V. S. (2020). Image mining methodology for detection of brain tumor: A review. 2020 Fourth International Conference on Computing Methodologies and Communication (ICCMC), 232-237.
- Tandel, G. S., Biswas, M., Kakde, O. G., Tiwari, A., Suri, H. S., Turk, M., ... & Suri, J. S. (2019). A review on a deep learning perspective in brain cancer classification. *Cancers*, 11(1), 111.
- Thite, A. (2024). Introduction to VGG16: Architecture and Applications. Great Learning Technical Report.
- Thust, S. C., Heiland, S., Falini, A., Jäger, H. R., Waldman, A. D., Sundgren, P. C., & Stupp, R. (2018). Glioma imaging in adults. *Neuroradiology*, 60, 229-245.
- Tummala, S., Kadry, S., Bukhari, S. A. C., & Rauf, H. T. (2022). Classification of brain tumor from magnetic resonance imaging using vision transformers ensembling. *Current Oncology*, 29(10), 7498-7511.
- Willis, J., Smith, C., Ironside, J. W., Erridge, S., Whittle, I. R., & Everington, D. (2005). The accuracy of meningioma grading: a 10-year retrospective audit. *Neuropathology and Applied Neurobiology*, 31(2), 141-149.
- Wirsching, H. G., & Weller, M. (2017). Glioblastoma. *Malignant Brain Tumors: State-of-the-Art Treatment*, 265-288.

BÖLÜM 3

NESNELERİN İNTERNETİ VE YAPAY ZEKA ENTEGRASYONUNDA SÜREÇ OPTİMİZASYONU: KAVRAMSAL BİR MODEL ÖNERİSİ

Abdullah Erdal TÜMER¹

¹ Prof. Dr. Abdullah Erdal Tümer, Necmettin Erbakan Üniversitesi, Orcid: <https://orcid.org/0000-0001-7747-9441>

1. Giriş

Dijital dönüşüm (DD), günümüzde iş süreçlerini ve iş modellerini belirlemede en önemli bir unsur olarak karşımıza çıkmaktadır. Bu dönüşüme en büyük katkıyı veren ve işin merkezinde olan ise yapay zekadır (YZ). Nesnelerin interneti (IoT) de süreçteki olmazsa olmazlardan, ayrılmaz parçalardan biridir. DD, YZ ve IoT birlikte yeni teknolojik bir çerçeve sunmaktadır. Bu çerçeve, üretimden hizmete kadar pek çok alanda etkisi vardır ve var olmaya devam edecektir.

Süreç optimizasyonu, özellikle endüstride gereksiz harcamaları ve kaynak israfını azaltan, verimliliği artıran ve iş süreçlerini iyileştiren dijital dönüşümün temel taşlarından biridir. On yıllar önce iş akış şemaları, gözlem ve mülakatlar, sebep-sonuç diyagramları ve neden analizleri gibi bazı klasik yöntemler endüstriyel alanlarda uzun yıllar hizmette etmiştir. Ancak zamanla ve günümüzde verilerdeki devasa artış ve verilerin işlenmesindeki zorluklar, klasik yöntemlerin yetersiz kaldığını göstermiştir. Sistemler gittikçe daha karmaşık bir hale geldiğinden, mevcut çözümlerin yetersiz kalmış ve yeni çözüm arayış ve yaklaşımlara ihtiyaç olmuştur. Tam da bu noktada yapay zeka adeta bir cam simid, can kurtaran gibi devreye girmiştir.

Nesnelerin interneti, kısa bir tanım olarak, fiziksel dünyayı dijital dünyaya bağlayan bir teknoloji olarak ifade edebiliriz. Bu teknolojiye klasik olarak sensörler veriyi toplar, aktüatörler bu veriye alıp veriye göre bir dizi eylem gerçekleştirir ve hareket eder. Bu döngü, gerçek zamanlı bilginin elde edilmesine ve gerektiğinde hızlı aksiyon almaya büyük katkı sağlar. Bu 2 teknolojinin bir araya gelmesiyle, literatürde "Nesnelerin (İnternetinin) Yapay Zekası" (Artificial Intelligence of Things veya **AIoT**) olarak adlandırılan yeni bir olguya işaret eder. Bu olgu, yapay zeka teknikleri ile endüstriyel nesnelerin interneti cihazlarının birleşimini ifade ederken, siber-fiziksel üretim sistemlerinin uygulanmasına da yardımcı olmaktadır. Ancak sahadan toplanan ham veri hiçbir zaman ve tek başına yeterli olması mümkün değildir. Veriden anlamlı bilgilerin çıkarılması gerekmektedir. Bu anlamlı bilginin çıkarılma işi noktasında devreye yapay zeka teknikleri girer [1].

Pek çok YZ tekniği vardır ki bu teknikler makine öğrenim algoritmaları başlığı altında toplanır ve bu tekniklerin öncülerinden biri de Yapay sinir ağları (YSA) olarak literatürde yer alır [2]. YSA doğrusal ya da doğrusal olmayan karmaşık örüntüleri tanıyarak sınıflandırma, tahmin ve modelleme yapabilmektedir. YSA'nın özellikle doğrusal olmayan ilişkileri başarılı bir şekilde modelleyebilmesi, süreç optimizasyonunda kritik önem taşır. Çünkü endüstriyel süreçler

doğrusaldan daha ziyade doğrusal olmayan süreçlerdir. Süreçte var olan değişkenler birbirlerini karmaşık olarak etkileyebilmektedir. Klasik istatistiksel modellerin bu karmaşıklığı tespit etmesi ve yakalaması oldukça zordur.

Literatürde IoT ve yapay zeka entegrasyonu üzerine çalışmalar hızla artmaktadır. Ancak bu çalışmaların önemli bir kısmı iki açıdan birine girer. Birinci tür, Ruiz-Sarmiento ve arkadaşlarının çalışma örneğindeki [3], Endüstri 4.0 modeli içinde yapay zeka ve makine öğrenmesinin süreç optimizasyonu aracı olarak kullanılmasına ilişkin önerilen kavramsal çerçevedir ki, bu bibliyometrik ve tarama ağırlıklı olanıdır. İkinci tür ise uygulama ve yöntem taramasına odaklanmaktadır. Bu da, Tran'ın akıllı üretim üzerine derlemesi olan, IIoT'nin akıllı üretimde üretim maliyetlerini, lojistik maliyetlerini ve kalite yönetim maliyetlerini %10-%30 oranında azaltabileceğini vurgulayan çalışmasıdır [4]. Mimari odaklı çalışmalar da benzer bir sınırlılık taşır. AIIoT referans mimarilerini inceleyen bir derleme, bu mimarilerin standartlaştırma açısından gerekli olduğunu, ancak gerçek endüstriyel AIIoT uygulamalarında kullanımlarının hâlâ başlangıç aşamasında kaldığını belirtir [4]. Bu üç örnek birlikte incelendiğinde ortak bir boşluk rahatlıkla görülmektedir. Literatürde var olan çerçeveler ya betimleyici bir haritalama sunmaktadır ya da belirli bir uygulamaya bağlı kalmıştır. İkisi arasındaki, yani IoT'nin veri katmanı ile yapay zekanın karar katmanını uygulamadan bağımsız biçimde birbirine bağlayan sentez düzeyinde bir model literatürde nadirdir. Bu bölüm, tam olarak bu sentezi hedeflemektedir. Mevcut bibliyometrik ve uygulama odaklı çalışmaları tek bir katmanlı mantık içinde birleştiren, uygulamadan bağımsız ve aktarılabilir bir kavramsal model önermektedir.

Bu bölümde önerilen model, önceki çalışmalardan türetilen veri toplama, analiz ve öğrenme ve karar ve geri besleme olmak üzere üç temel katmandan oluşmaktadır. Veri toplama katmanı, IoT mimarilerinde tanımlanan sensörlerin algılama işlevine karşılık gelir. Analiz ve öğrenme katmanı, genelde YZ algoritmaları özelde yapay sinir ağı tabanlı tahmin ve örüntü tanımının kavramsal karşılığıdır. Karar ve geri besleme katmanı ise, mimari çalışmalarda ifade edilen, denetim ve kontrol fonksiyonunu temsil etmektedir. Bu üç katman ne rastgele ne de yan yana durmaz. Aralarında yönlü ve güçlü bir bağımlılık bulunmaktadır. Birinci katman olan veri toplama katmanı, verinin kalitesi ile, ikinci katmandaki analiz doğruluğunu sınırlayabilmektedir. İkinci katmandaki analiz çıktısı ise, üçüncü katmandaki kararın girdisini oluşturmaktadır. Üçüncü katmandaki karar, birinci katmandaki veri toplama stratejisini doğrudan ilgilendir ve yeniden şekillendirir. Böylece döngü tamamlanmış olur. Bu döngü,

doğrusal bir akıştan ziyade, geri besleme mekanizmasından dolayı döngüsel olarak karşımıza çıkar. Bu yüzden model, bağımsız parçaların toplamı olarak değil, 3 katman arasındaki karşılıklı belirleyicilik olarak modelin merkez argümanını oluşturur.

IoT ve yapay zeka entegrasyonu arttıkça, geri besleme döngüsünün tamamlanma hızı da artacaktır. Geri besleme döngüsü hızlandıkça, süreç optimizasyonu düzeyi de yükselecektir. Entegrasyon düzeyi, optimizasyon çıktısını etkilediğinden bu ilişki tek yönlüdür. Ölçülebilir çıktı, kararların zamanında alınması ya da olası gecikmelerde azalma ve kaynak kullanımındaki verimlilik olarak tanımlanabilir. Bu bölümde bu ilişki ampirik olarak test edilmeyecektir. Ancak ilişkiyi, sonraki uygulamalı çalışmalarda test edilebilecek şekilde ortaya koymaya çalışılmıştır.

Yapay zeka çağında süreç optimizasyonu statik olmaktan daha çok oldukça dinamiktir. Sürekli yeni gelen verilerle öğrenen bir yapıya sahiptir. Bu durum, geleneksel mühendislik yaklaşımlarını da etkilemekte hatta değiştirmektedir. Günümüz mühendisler artık sadece süreç tasarlamakla kalmayıp, aynı zamanda tekrar tekrar öğrenen sistemler tasarlamak zorundadırlar. Bu, mühendislik pratiğinde köklü bir değişimin gerekliliğine işaret etmektedir.

Bölümün ilerleyen kısımlarında öncelikle kavramsal çerçeve sunulacaktır. Ardından ilgili literatür daha ayrıntılı değerlendirilecektir. İlerleyen bölümlerde önerilen model detaylı bir şekilde anlatılmıştır. Modelin literatüre katkıları sınırları ele alınmıştır. Her bölüm, bir önceki bölümü tamamlayacak şekilde ilerlemektedir. Amacımız sadece parçaları sırayla sunmak değildir. IoT ve yapay zeka iş birliğinin, teknik bir birleşimin ötesinde, sistemin verimli hale getirilmesinde temel bir mekanizma olduğunun net bir şekilde gösterilmesidir.

2. Kavramsal Çerçeve: Nesnelerin İnterneti ve Yapay Zeka Katmanlarının Bütünleşik Yapısı

2.1.Nesnelerin İnterneti Mimarisinin Katmanlı Yapısı

Nesnelerin İnterneti (IoT) mimarisi için 3 ,4 ve 5 katmanlı olmak üzere farklı modeller ortaya atılmıştır. Bunlar arasında en temel ve kabul gören mimari 3 katmanlı yapıdır. Bu 3 katmanlı model, algılama, ağ ve uygulama katmanlarından oluşmaktadır. [5] ile [6] , bu 3 katmanlı mimarinin, bir IoT sisteminin olmazsa olmazı olduğunu savunmuşlardır. Ancak bu 3 katmanlı mimari, örneğin Alshohoumi ve ark. [7] tarafından, sistemin verilerle ne yapılacağı ve nasıl karar verileceğinin eksik olmasından dolayı yetersiz bulunmuştur. Bu çalışmanın çıkış

noktasını, ilgili literatürdeki yetersizlik oluşturmaktadır. Mevcut mimariler veri akışını açıklasa da iş karar verme sürecine geldiğinde tıkanıp görülmektedir.

Bu çalışmada önerilen modelle, bu boşluğun kapatılması hedeflenmiştir. Literatürdeki 3 katmanlı modellere iş ve karar katmanı eklenerek 5 katmanlı oluşturulan modellere benzetilerek, iş ve karar katmanı doğrudan tanımlanmış, yapay sinir ağlarına dayalı modern bir analiz mekanizması sunulmuştur. Ayrıca modele, günümüz endüstrisinde, gecikmeleri önlemek ve gerçek zamanlı kararlar alabilmek için kritik olan kenar hesaplama (edge computing) yaklaşımı da dahil edilmiştir. Böylece verinin kaynağına en yakın yerde, en hızlı şekilde işlenmesinin önü açılmıştır.

2.2.Yapay Zeka ve Yapay Sinir Ağlarının Mimari İçindeki Konumu

Yapay zeka, bu çok katmanlı modelin sadece tek bir noktasında yer almaz. Farklı katmanlarda da farklı işlevleri yerine getirir. Örneğin algılama katmanında, verinin ön işlenmesinde kullanılır. Ağ katmanında, veri yönlendirmede bazı kararların alınmasında kullanılabilir. Ancak en yoğun kullanıldığı yer ise tabiki analiz ve karar süreçleridir.

Yapay sinir ağları algoritması bu analiz ve karar sürecinin tam ortasındadır. Girdi katmanından veri alınır. Gizli katmanlarda veri işlenir. Çıktı katmanında ise sonucu üretilir. Bu yapı, özellikle doğrusal olmayan ilişkileri bulma ve modelleme için oldukça uygundur. Endüstriyel süreçlerde girdi verilerindeki bağımsız değişkenler arası ilişki doğrusallıktan daha çok doğrusal olmayan bir ilişkidir. Bu yüzden YSA, klasik istatistiksel yöntemlere kıyasla oldukça başarılı bir şekilde uygulanır ve avantaj sağlar.

2.3.Entegrasyon Mantığı: Katmanlar Arası Eşleme

Bu bölümün girişinde veri toplama, analiz ve öğrenme, karar ve geri besleme adında üç katmanlı bir model önerilmişti. Bu 3 katmanlı modelde her katman, belirli bir veri türünü üretir. Üretilen veri türleri, belirli bir analiz yönteminden geçmesi gerekir. Bu yüzden klasik 3 katmanlı model ile bu modeli eleştirip, iş ve karar katmanı olması gerektiğini söyleyen modellerle eşleşir. Bu eşleşme veri türü ile analiz yöntemi arasındaki uyumla ilgilidir.

Bu durumu somutlaştıracak olursak, algılama katmanı, bir titreşim sensöründen gelen zaman serisi verisi gibi, ham ve genellikle gürültülü sinyal üretir. Bu veri türü, örüntü tanımaya uygun yapay sinir ağı mimarileri ile gerçekleştirilebilir. Ağ katmanı, örneğin, sıcaklık ve basınç verisinin

eş zamanlı akışı gibi farklı kaynaklardan gelen veriyi zamanla ilişkilendirir. Bu tür ilişkiyel veri, çok girdili ağ yapıları ile yapılabilir. Karar katmanı ise örneğin kestirimci bakım ihtiyaçları gibi analiz çıktısını eyleme dönüştürür. Algılama, ağ ve karar katmanı için verilen bu örnekler, eşlemenin neden gerekli olduğunu açıklar. Yanlış analiz yöntemi seçildiğinde, o katmandaki verinin işlenmesi engellenmiş olur.

Bu yapının 2 pratik sonucu vardır: a) mühendislik açısından b) karar hızı açısından. Sistem tasarımcısı mühendis, hangi katmanda hangi model türünün gerekli olduğunu önceden belirleyebilirken, geri besleme döngüsü sayesinde karar katmanındaki çıktı, algılama katmanının hangi veriyi önceleyeceğini belirlemiş olur. Bu, sistemin karar alma biçimini tamamen değiştirmiş olur. Klasik 3 katmanlı mimariye sahip IoT modelinde veri akışı tek yönlü olup, önce algılama sonra uygulama ile devam etmektedir. Önerilen modelde ise karar katmanı, bir sonraki veri toplama döngüsünü şekillendirmektedir. Sistem, geçmiş kararlardan öğrenerek hangi veriye öncelik vereceğine karar vermektedir.

2.4.Modelin Süreç Optimizasyonu İlişkisi ve Sınırları

Karar katmanının bir sonraki veri toplama döngüsünü şekillendirdiği döngüsel yapı, yukarıdaki mantığın doğal bir uzantısı olduğunu söyleyebiliriz. Şöyle ki, karar katmanının geri beslemesi arttıkça, sistemin uyarlanabilirliği de artacaktır. Bu ilişkiyi, test edilebilir bir soru olarak ifade edebiliriz. Bu soru, geri besleme döngüsünün tamamlanma frekansı arttıkça, karar gecikmesi belli bir oranda/biçimde azalır mı? şeklindedir. Bağımsız değişken, döngü başına geçen süredir. Bağımlı değişken, karar gecikmesi ve kaynak kullanım oranıdır. Beklenen etki yönü ise, döngü frekansı arttıkça, karar gecikmesinin azalması gerektiğidir. Bu bölüm bu ilişkiyi test etmeye yönelik olmayıp, sonra ya da ileride yapılacak uygulamalı çalışmalar için bir çerçeve sunmaktadır.

Bu yüzden sunulan kavramsal çerçevenin kapsamı da sınırlıdır. Model, yalnızca sürekli veri akışına sahip, orta ila yüksek otomasyon düzeyindeki endüstriyel sistemler için geçerli olup uygulanabilir. Kesikli ve düşük frekanslı veri üreten sistemlerde bu çerçevedeki döngü mantığı zayıflayacaktır. Ayrıca modelin işlerliğinin kısıtlılığı da vardır. Bu kısıtlardan birincisi verinin kalitesi, diğeri karar gecikmesidir. Düşük kaliteli veri, analiz katmanında hatalı örüntülere yol açarken, karar gecikmesi noktasında, analiz katmanı ne kadar doğru olursa olsun, karar katmanının tepki süresi fiziksel sistemin değişim hızından yavaşsa, geri besleme döngüsü anlamını yitirecektir. Bu durumda döngü, sistemi optimize etmek yerine, bozulmayı

derinleştirir. Söz konusu problemler birbirinden tamamen bağımsız çalıştığından, sorunların biri çözülse, diğerleri modelin işlerliğini kısıtlayabilecektir. Bu yüzden, önerilen modelin genel kapsamı, bir taraftan veri güvenilirliğinin diğer taraftan sistemin tepki hızının belirli bir seviyenin üzerinde olduğu durumlarda sınırlı kalacaktır.

Sonuç olarak bu kavramsal çerçeve, iki ayrı literatürü tek bir mantık içinde birleştirmektedir. İlk literatür, nesnelerin interneti mimarilerini katmanlı biçimde tanımlarken, ikinci literatür, yapay sinir ağlarının analiz gücünü vurgulamaktadır. Bu senteze mevcut IoT mimarileri veri akışını açıklar, ama kararın veriye nasıl geri döndüğünü açıklayamadığından ihtiyaç duyulmuştur. Mevcut yapay zeka çalışmalarında analiz gücünü gösterir, ancak bu analizin hangi fiziksel mimari içinde konumlandığını açıklamaz. Bu bölümün katkısı, bu iki eksik açıklamayı birbirini tamamlayacak şekilde birleştirmesidir. Bu birleşim olmadan, IoT ve yapay zeka entegrasyonu üzerine yapılan çalışmalar birbirinden kopuk iki teknik alan olarak kalmaya devam edeceği değerlendirilmiştir.

3. Literatür Taraması

Bu bölüm üç alt bölüme ayrılarak oluşturulmuştur. Alt bölümler, IoT ve yapay zeka yakınsama, (AIoT), yapay sinir ağı tabanlı süreç optimizasyonu uygulamaları ve kavramsal ve mimari çerçeve önerileri şeklindedir. Alt bölümler birbirinden bağımsız olarak araştırılmıştır. Sonrasında aralarındaki ilişki ayrıca tartışılacaktır.

3.1.AIOT Yakınsaması Üzerine Çalışmalar

Nesnelerin interneti ile yapay zeka entegrasyonu, literatürde “Nesnelerin Yapay Zekası (Artificial Intelligences of Things-AIoT)” ile anılmaktadır. Seng, Ang ve Ngharamike'nin çalışması [8], bu yakınsamayı dört başlıkta ele almıştır. Bunlar, i) mimariler ve donanım platformları, ii) sensör ve enerji yaklaşımları, iii) iletişim ve ağ teknolojileri ve iv) uygulama alanlarıdır. İlgili çalışma, AIoT sistemlerinin kısıtlı kaynaklara sahip IoT cihazları ile yoğun hesaplama ihtiyacı olan yapay zeka algoritmaları arasındaki probleme dikkat çekerek odaklanmıştır. Bu odak, kavramsal çerçeve açısından önemli olarak değerlendirilmiştir. Çünkü bir modelin teorik olarak tutarlı, anlamlı olması, hesaplama problemlerini göz ardı etmeyi gerektirmediği gibi uygulanabilir de kılmamaktadır.

Ishengoma ve arkadaşlarının çalışması [9], AIoT'nin kamu sektöründeki yaygın olarak kullanılıp kullanılmadığını incelemiştir. Bu çalışma, AIoT benimsenmesinin önündeki engelleri

sürücüler ve engeller biçiminde 2 ayrı kategoride sınıflandırmıştır. Bu sınıflandırma, teknik olmayan detayların da önerilen modele dahil edilmesi gerektiğini net bir şekilde ortaya koymaktadır. Ancak bu çalışma belli bir sektöre odaklandığı için, genel bir optimizasyon mantığı sunmamaktadır.

Literatürü geniş bir çerçeveden ele alan sistematik bir başka çalışmada [4], IoT ve yapay zekanın mimarisi, uygulama ve çözilememiş sorunlar kapsamında, nasıl bir araya geldiği ele alınmıştır. Araştırma, alanın hızlı bir şekilde büyüdüğünü ve geliştiğini, ancak kabul edilmiş bir standart etrafında, kavramsal dilin henüz kurulamadığını açıkça belirtmektedir. Bu çalışma, giriş bölümünde tespit edilen boşluğu dorular niteliktedir. Çünkü AIoT dünyasında kaynak sayısı hızla artmakta ama tüm bunları sentezleyen kavramsal çalışmalar yetersiz kalmaktadır.

3.2.Yapay Sinir Ağı Tabanlı Süreç Optimizasyonu Çalışmaları

Bu alt bölüm yapay sinir ağları yönteminin süreç optimizasyonundaki uygulama sonuçlarına yöneliktir. Üretimde kalite güvencesi ile ilgili bir derleme çalışmasında[10], iki yüze yakın çalışma taranmıştır. Görüntü işlemede evrişimli sinir ağları yönteminin hata tespitinde yüksek başarı sağladığı görülürken, yoğun hesaplamalar gerektiği için güçlü donanımlara ihtiyaç olduğu belirtilmiştir. Aynı çalışmada, destek vektör makineleri ve rastgele orman yöntemlerinin, kestirimci bakım ve süreç optimizasyonunda başarılı sonuçlar elde edildiği ifade edilmiştir. Bu bulgular, farklı yapay zeka yöntemlerinin farklı senaryolarda farklı başarılar göstermesi bakımından önemlidir. Yapay zeka yöntemlerinin seçimi, verinin türüne ve amacına göre seçilmesi gerektiğinin de bir göstergesi olarak düşünülebilir.

Otomotiv endüstrisindeki kalite kusurlarını ile ilgili yapılan başka bir derleme çalışmasında [11],otomotiv üretim hattı otomasyonu ve üretim hatalarının, derin öğrenme ve yapay sinir ağları kullanılarak, dikkate değer bir şekilde azaltıldığı gösterilmiştir. Kalan kullanım ömrünün tahminine yönelik bir başka çalışma [12], çok katmanlı yapay sinir ağlarının, NASA'nın turbo motorlardan elde ettiği veri setlerinde başarılı sonuçlar verdiği raporlanmıştır. Bu çalışmada, temel bileşen analizi ve ızgara arama yöntemleri gibi farklı yapay zeka algoritmalarıyla model parametrelerinin optimize edildiğini de göstermektedir.

[13] deki çalışmada, plastik enjeksiyon kalıplarına yönelik hibrit bir yöntem kullanılmıştır. Bu hibrit yöntemde, geri yayımlı yapay sinir ağı ile genetik algoritma yöntemi birlikte kullanılmıştır. Optimizasyon hedefleri ile süreç parametreleri arasındaki ilişki yapay sinir

ağıyla modellenmiş ve genetik algoritmayla da küresel optimuma yaklaştırdığı gösterilmiştir. Sonuçlar, hammadde tüketiminde %2, çevrim süresinde %12, enerji tüketiminde %16 azalma görüldüğü beyan edilmiştir. Bu örnek, analiz katmanındaki yapay sinir ağı çıktısının, karar katmanındaki optimizasyon algoritmasına girdi sağlaması bakımından, önerilen kavramsal modelin karar katmanında geri besleme mantığının somut bir karşılığı olarak değerlendirilmiştir.

IoT tabanlı kestirimci bakım çalışmalarının bir diğerinde [14], bulanık sistemler ile yapay sinir ağları birlikte kullanılarak hibrit bir yaklaşım benimsenmiştir. Bu hibrit çalışma, bir oluklu mukavva üretim tesisinde gerçek bir endüstriyel ortamda uygulanmıştır. Çalışma, sensör bilgi modellemesi ile nesnelerin interneti teknolojilerinin birlikte kullanarak, bakım verimliliğinin arttığını göstermiştir.

Bu alt bölümde yer verilen çalışmaların ortak özelliği, uygulama odaklı olmasından kaynaklanmaktadır. Sonuçlar, belirli bir sektörden ya da nereden alındığı bilinen belirli bir veri setine dayanarak üretilmiştir. Genellenebilir bir kavramsal çerçeve sunulmamıştır. Bu durum, kavramsal modelin neden gerekli olduğunu bir kez daha doğrular mahiyettedir. Çünkü uygulama çalışmaları, katmanlar arası ilişkinin işlediğini gösterir ama bu ilişki genel bir mantığa dönüştürülmemektedir.

3.3.Kavramsal ve Mimari Çerçeve Önerileri

Üçüncü literatür alt bölümünde, kavramsal çerçeve önerileri sunulmuştur. Bu konu ile ilgili literatür diğerlerine oranla daha azdır. [3] 'de, Endüstri 4.0 için yapay zeka ve makine öğrenim algoritmalarının süreç optimizasyonunda kullanımını bibliyometrik yöntemlerle haritaladığı çalışmasından bahsedilmiştir. AIoT referans mimarilerini inceleyen bir derleme [4] de ise, bu mimarilerin standartlaştırma açısından faydalı olabileceğini vurgulamıştı. Ancak yazarlar, gerçek endüstriyel uygulamalarda kullanımları noktasında henüz ciddi bir ilerleme olmadığı ve başlangıç aşamasında kaldığını ifade ediyorlar. Kestirimci bakım üzerine güncel bir sistematik derleme [5] de benzer bir sonuçtan bahsedilir. Bu çalışmada gerçek zamanlı uygulama, model sağlamlığı ve heterojen veri yönetimi konularında önemli engellerin olduğu öne sürülmektedir.

Bu üç kaynak birlikte değerlendirildiğinde net bir perspektifin ortaya çıktığını söyleyebiliriz. Mevcut çalışmalarda kavramsal çerçeveler ya sınıflandırma yapar ya da mimari düzeyde kalır. Bir standardizasyona odaklanma söz konusudur. Hiçbirinde, IoT katmanları ile yapay sinir ağı

tabanlı analiz ve karar mekanizmaları arasındaki geri beslemeli ilişkiden doğrudan modellenebilir bir mantık çıkarılmadan sunulur.

3.4. Literatürün Sentezi ve Bu Bölümün Konumu

Üç alt bölümü ortak değerlendirdiğimizde, AIoT ye ait literatürle, yakınsamanın gerekliliği ve zorlukları genel düzeyde gösterilmeye çalışılmıştır. Yapay sinir ağı tabanlı optimizasyon literatürü, bu yakınsamanın somut faydasını uygulama düzeyinde kanıtlarken, kavramsal çerçeve literatürü bu iki alanı ya bibliyometrik ya da mimari düzeyde birleştirmeye çalışır. Ancak işlevsel bir geri besleme mantığı sunulmaz.

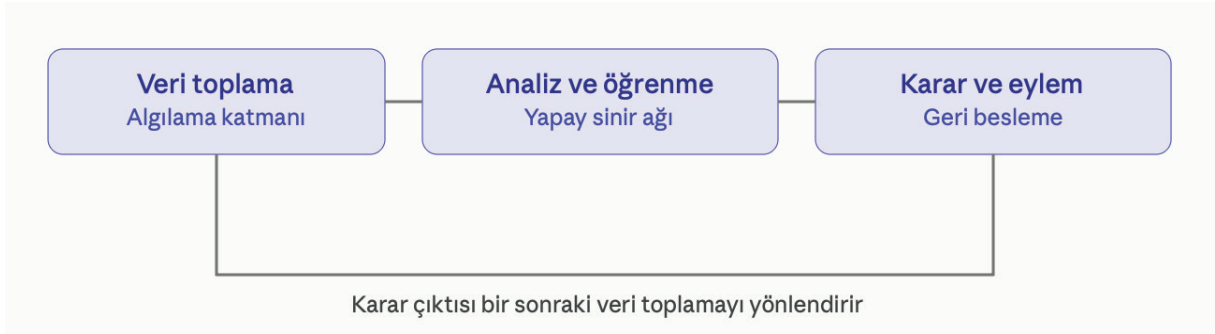
Bu bölümün önerdiği model, üçüncü alt bölüm için bir katkı sunmak üzere konumlandırılmıştır. Ancak bu katkı, ikinci alt bölümdeki uygulama sonuçlarından bağımsız olarak değerlendirilemez. Bilakis, ikinci alt bölümdeki dağınık sonuçlar (örneğin sinir ağı çıktısının genetik algoritmaya girdi sağlaması ya da bulanık sistem çıktısının bakım kararını şekillendirmesi gibi), tek bir genel mantık altında toplamaktadır. Bu bölümünde yenilik, uygulama çalışmalarında tekrar tekrar görülen "analiz çıktısının karara, kararın da yeni veri toplamaya dönmesi" örüntüsünün, tesadüfi olmadığının ifadesidir. Bu örüntü, IoT ve yapay zeka entegrasyonunun yapısal bir özelliği olarak karşımıza çıkmaktadır. Bu bölümün geri kalanında bu yapısal özellik, önerilen model üzerinden sistematik biçimde tartışılacaktır.

4. Önerilen Kavramsal Model

4.1. Modelin Genel Yapısı

Bu bölümde önerilen model, önceki bölümlerde ifade edilen literatüre katkı bağlamında var olan boşluğa dolduracağı düşünülmektedir. Önerilen model, i. veri toplama, ii. analiz ve öğrenme ve iii. karar ve geri besleme katmanlarından oluşmaktadır. Bu üç katman, nesnelerin interneti mimarisinde genel ve klasik olarak tanımlanan [5], [6] algılama, ağ ve uygulama katmanlarına işlevsel özellik katılmış halidir. Ancak önerilen modelde, klasik mimarilerden farklı olarak, Alshohoumi ve arkadaşlarının [15] çalışmalarında vurguladıkları eksikliği giderme noktasında, katmanlar arasına açık bir karar ve geri besleme mekanizması eklenmiş halidir.

Önerilen modeldeki üç katman, birbirinden bağımsız çalışmamaktadır. Birbirini karşılıklı olarak belirleyen bir döngü mekanizması ile çalışmaktadır. Şekil 1. de döngü gösterilmiştir.



Şekil 1. Önerilen modeldeki döngüsel mekanizma

4.2.Katmanlar Arası Fonksiyonel İlişki

Önerilen model, üç katman ve üç dönüşüm fonksiyonundan oluşur:

$$M = (L1, L2, L3, f, g, h).$$

Burada;

- L1 veri toplama katmanını,
- L2 analiz ve öğrenme katmanını,
- L3 karar ve eylem katmanınıdır.

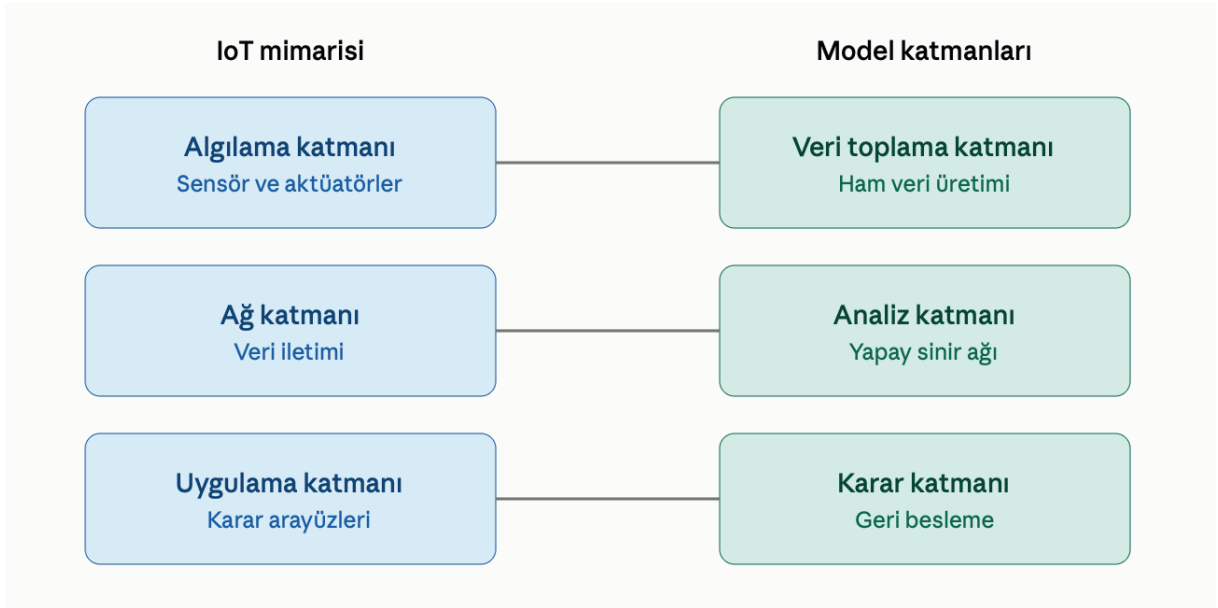
f fonksiyonu L1'den L2'ye, g fonksiyonu L2'den L3'e, h fonksiyonu ise L3'ten tekrar L1'e geçişi tanımlamaktadır.

Bu üç fonksiyon, keyfi olmayıp, her biri, önceki bölümlerde tartışılan somut örneklerle karşılık gelmektedir.

- f fonksiyonu, işlenmemiş ham sensör verisini örüntülere dönüştürür. Bu, yapay sinir ağlarının örüntü tanıma işlevini yerine getirir. Derin öğrenme algoritmalarının görüntü işleme ve zaman serisi verisindeki başarılı uygulamalarından dolayı seçilmiştir [4] .
- g fonksiyonu, analiz çıktısını bir karara dönüştürerek, sinir ağı çıktısının optimizasyon algoritmasına girdi sağlayarak, hibrit bir yaklaşım olarak somutlaştırılmıştır.
- h fonksiyonu da, kararın veri toplama stratejisinde, nasıl yeniden şekillendirildiğini tanımlamaktadır. Kestirimci bakım sistemlerinde karar çıktısının, hangi sensörün önceliklendirileceğini belirlemesiyle örneklenir [16] .

Bu üç fonksiyonun birlikte oluşturduğu bu döngüsel mekanizma, önerilen modelin merkezi iddiasını oluşturmaktadır. $L1 \rightarrow L2 \rightarrow L3 \rightarrow L1$ döngüsü tamamlandığında ancak sistem bir "öğrenen" sistem olarak adlandırılabilir. Yalnızca f ve g varsa, sistem statik bir analiz hattından başka bir şey değildir. Yani, girdi alınır, çıktı üretilir, ancak, sistem kendini güncelleyemeyecektir. h fonksiyonunun eklenmesi, sistemi statik bir hattan öğrenebilir ve uyarlanabilir bir döngüye dönüştürmektedir. Bu durum, literatürdeki birçok çalışmanın sadece f ve g'yi kurması ve h çoğunlukla göz ardı edildiği için önemlidir [1] [3].

Daha iyi bir ifade/anlama için, şekil 2. de, önerilen modelin, klasik IoT mimarisiyle ilişkisi eşleştirilmiştir.



Şekil 2. Klasik ve önerilen IoT mimarilerinin eşleştirilmesi

4.3. Modelin Kısıtları

Önerilen modelin işlerliğinde bazı kısıtlarda söz konusudur. Burada iki eşik değerin, döngüyü çalışma noktasında etkilemektedir. İlki, veri güvenilirliği (θ_q) diğeri ise tepki süresidir (θ_t). Algılama katmanından gelen verinin kalitesi bu veri güvenilirliği eşığının altına düşerse, f fonksiyonunun hatalı örüntüler üretmesi kaçınılmazdır. Bu yüzden döngü optimizasyonunda daha çok bozulma eğilimi oluşacaktır. Bu risk, kestirimci bakım uygulamalarında görülmüştür. Çünkü düşük kaliteli sensör verisi, yanlış arıza tahminlerine yol açabilmektedir [14] .

Tepki süresi eşığındeki h fonksiyonunun tamamlanma süresi, fiziksel sistemin değişim hızından yavaşsa, geri besleme anlamsız kalacaktır. AIoT mimarilerini inceleyen [1] deki bir derleme çalışmasında, benzer gecikme problemlerinin, endüstriyel uygulamalarda klasik mimarilerin

hâlâ sorunlu kalmasının nedeni olarak gösterilmektedir. Bu veri güvenilirliği ve tepki süresi eşikleri birbirinden bağımsızdır. Biri karşılanırsa bile diğeri karşılanmadığında, modelin işlerliği sınırlanabilecektir.

4.4.Modelin Pratik Çıkarımları

Önerilen model sistem tasarımı ve performans değerlendirmesi açısından iki somut sonuç üretmektedir.

Bir mühendis, sistem tasarımı açısından, hangi katmanda hangi modele ihtiyaç olduğuna, veri türü ile analiz yöntemi arasındaki ilişkiye bakarak karar verebilir. Görüntü ilemeye yönelik bir hata tespiti, evrişimli sinir ağına ihtiyaç duyarken, zamana bağlı bir seri tabanlı arıza tahmini bambaşka bir yapay zeka algoritması gerektirebilir.

Performans değerlendirmesi açısından ise, sistemin "akıllı" olup olmadığı, sadece analiz başarısı ile değil, h fonksiyonu ile test edilebilir. Sistem yüksek doğrulukla tahmin üretiyor olabilir ancak bir sonraki veri toplama döngüsünü etkilemiyorsa, model açısından bu sistemin döngüsel olmadığı, doğrusal olduğu söylenecektir ve optimizasyon potansiyeli sınırlı kalacaktır.

4.5.Modelin Literatürdeki Konumu

Bu model, literatürde yer verilen üç ayrı alt bölümü birleştirir. AIoT alanındaki literatür, AIoT nin yakın zamanda daha çok çalışma yapma gerekliliğini göstermektedir [8], [9]. Yapay sinir ağları ile optimizasyona ilişkin literatürde, f ve g fonksiyonlarının işlerliğini uygulama düzeyinde kanıtlanmıştır [4], [14] . Sunulan kavramsal çerçeve literatüründe bu iki alanı bibliyometrik veya mimari düzeyde bir araya getirmeye çalışmaktadır. Ancak h fonksiyonunun açık biçimde modellemekten uzaktır [1], [3] . Bu bölümünde sunulan katkı, geri besleme mekanizması olan h fonksiyonunun modelin merkezine yerleştirilmesidir. Bu yerleştirme, modelin sadece bir sentez değil, test edilebilir de bir öneri olmasını sağlamaktadır.

5. Tartışma

5.1.Modelin Katkısının Yeniden Değerlendirilmesi

Önerilen model, özetle, IoT mimarisi ve yapay sinir ağ algoritmasının birlikte analiz süreçlerini, bir geri besleme mekanizması ile birleştirmektedir. Önerilen bu model, literatürdeki boşluğa katkı sunacağı değerlendirilmektedir. Nesnelerin interneti mimarileri ile ilgili ilk çalışmalar, veri akışını sıralı bir yapı olarak ele almaktadır [5], [6]. Yapay sinir ağ algoritması kullanılarak yapılan optimizasyon çalışmalarında ise YSA'nın analiz gücünü uygulama düzeyinde kanıtlayarak ilgili uygulamalarda bir araç olarak kullanılabileceğini ispatlamaktadır [4][4]. Bu iki alan arasındaki köprü, önerilen modeldeki karar katmanından veri toplama katmanına dönen h fonksiyonu ile kurulmuştur. Bu bölümünün amacı, bu katkının sağlamlığını test etmek, kısıtlarını ortaya koymak ve ileriye dönük araştırma sorularını belirlemektir.

5.2.Modelin Kısıtları

Modelin temel olarak üç kısıtı olduğu söylenebilir. İlk ikisi önceki bölümde bahsedilen, veri güvenilirliği eşiği ile tepki süresi eşiğidir. Bu tartışmada üçüncü ve daha az belirgin bir kısıt daha eklemek gerekirse onun da, analiz katmanının açıklanabilirlik sorunu olarak ifade edebiliriz.

Yapay sinir ağlarının pek çok çalışmada yüksek doğruluk sağladığı gösterilmiştir. Ancak bu doğruluk çoğu zaman açıklanabilir/şeffaf değildir. Derin öğrenme yöntemi ile geliştirilen modelleri, kararlarına nasıl ulaştıkları açık biçimde gösterilmez. Bu yüzden bu durum literatüre "kara kutu" olarak geçmiştir [17]. [27] deki derleme, endüstriyel arıza tespitleri için yapılmıştır. Modeller yüksek başarı gösterse de, hangi değişkenlerin ya da özelliklerin karara yol açtığının bilinmediği ve belirsiz olabileceği vurgulanmaktadır. Bunun, güvenlik açısından makine öğrenim yöntemlerinin, kritik uygulamalarda kullanılmasını zorlaştırdığı ifade edilmektedir. Önerilen modelin, analiz çıktısının karara dönüştürülmesi olan g fonksiyonunun sadece istatistiksel doğruluğa dayandığında, karar katmanındaki insan operatörleri bu çıktıya güvenmede problem yaşayabilir ya da tereddüt edebilir.

Açıklanabilir yapay zekanın, üretim sektöründe kullanılmasına yönelik bir çalışmada [18], imalat sanayiinde yapay zeka modellerinin benimsenmesinin açıklanabilirliğin modelin başarısı kadar önemli olduğu ifade edilmiştir. Bunun "şeffaflık paradoksu" olduğu yazarlar tarafından tanımlanmıştır. Bu bulgu, önerilen modelin sınırlarını ve kısıtlarını artırmaktadır. Önerilen

model sadece veri kalitesi ve tepki süresi eşikleriyle değil, aynı zamanda karar katmanındaki insan aktörlerin analiz çıktısına duyduğu güvene de bağlı kalmaktadır. Üç sınırlılık/kısıt birlikte ele alındığında, önerilen modelin işlerliği, teknik koşullardan başka, örgütsel kabul koşullarına da bağlı kalmaktadır. Bu yüzden modelin kapsamı, yalnızca yüksek veri güvenilirliği ve düşük tepki gecikmesi haricinde, karar vericilerin analiz çıktısını yorumlayabildiği sistemlerle sınırlı kalmaktadır.

5.3.Modelin Genellenebilirliği

Bir başka tartışma noktası, modelin genellenebilirlik düzeyi olduğunu söyleyebiliriz. Önerilen model, üç katmanlı yapıyı soyut düzeyde tanımlamaktadır. Bunlar, belirli bir sektöre veya yapay sinir ağı yönteminin kullanılmasına bağlı değildir. Bu soyutluk bir avantaj olup, model, hem görüntü tabanlı kalite kontrolüne hem de zaman serisi tabanlı kestirimci bakıma uyarlanabilir. Ancak aynı soyutluk, hangi sinir ağı yönteminin, hangi veri türüne uygun olduğunu belirtmediğinden risk taşımaktadır. Bu seçim, uygulayıcıya ve uygulanacak sektöre göre değişir ve seçimlidir. Bu da, modelin bir kusuru/problemi değil, isteğe bağlı bir tercihtir. Bölümün başında ifade edildiği gibi, amaç uygulamadan bağımsız bir çerçeve sunmaktır. Ancak bu tercihin sonucu, model tek başına bir uygulama rehberi işlevi göremeyeceği açıktır. Sonraki çalışmalar, modelin belirli sektörlerde nasıl somutlaştırılacağına dair olmalıdır.

5.4.Açık Araştırma Soruları

Bu bölüm, üç açık araştırma sorusu oluşturmaktadır.

- i. Önceki bölümde önerilen ilişki ile ilgilidir. Önceki bölümde, geri besleme döngüsünün tamamlanma sıklığı ile karar gecikmesi arasındaki ilişki, farklı endüstriyel bağlamlarda ampirik olarak test edilmelidir.
- ii. Açıklanabilirlik ile model performansı arasındaki denge ile ilgilidir. Açıklanabilir yapay zeka yöntemlerinin g fonksiyonuna entegre edilmesi, karar kalitesini nasıl etkiler? Soru, sadece teknik olmayıp, aynı zamanda insan-makine güveni açısından da test edilebilir bir varsayıma dönüştürülebilir. Örneğin, g fonksiyonuna açıklanabilirlik katmanı entegre edildiğinde, karar verici konumdaki operatörlerin sistemin önerdiği kararı kabul etme noktasında istatistiksel olarak anlamlı bir artış olur mu?
- iii. Önerilen modelin sektörler arası genellenebilirliği ile ilgilidir. Model, farklı otomasyon düzeyindeki sektörlerde (örneğin yüksek otomasyonlu otomotiv üretimi ile daha düşük

otomasyonlu küçük ölçekli imalat sanayinde) aynı başarıyı yakalayıp yakalamayacağı ne ölçüde bellidir? Bu soru, modelin sınırlarını sektöre göre nasıl değiştiğini test etmeyi gerektirmektedir.

5.5.Kapanış Değerlendirmesi

Önerilen model, IoT mimarisi ve yapay sinir ağı yöntemini hibrit olarak kullanılmasındaki analiz süreçlerini birleştiren, geri beslemeli ve test edilebilir bir kavramsal çerçeve sunmuştur. Modelin katkısı, katmanlı IoT mimarisine, yapay sinir ağlarının analiz gücünü katmasıdır. Böylece geri besleme mekanizmasının eksikliği giderilme mantığı üzerine kurulmuştur. Ancak model, sadece veri güvenilirliği ve tepki süresi yeterli olduğunda değil, aynı zamanda analiz çıktısının karar vericiler tarafından da anlaşılabilir ve güvenilir bulunduğu anda amacına başarılı bir şekilde ulaşmış olacaktır. Bu üçlü şart, modelin en önemli katkısı olarak öne çıkarılmıştır. Aynı zamanda modelin en önemli kısıtını da oluşturmaktadır. Sonraki uygulamalı çalışmalarda, bu şartların hangi sektörlerde ve hangi ölçüde karşılandığını gösterilmesi ve modelin kavramsal iddiasını ampirik zemine taşınması gerekmektedir.

6. Sonuç ve Gelecek Çalışmalara Yönelik Öneriler

Bu çalışmada IoT ile yapay zekanın birleştirilmesine dair kavramsal bir model amaçlamıştır. Literatürdeki IoT ve yapay sinir ağı yönteminin kullanılarak yapıldığı optimizasyon çalışmalar incelenerek, literatürdeki bir boşluk olduğu değerlendirildi. Araştırabildiğimiz kadarıyla, IoT ve YSA yı birbirine entegre eden, uygulamadan bağımsız çalışmaların yok denecek kadar az olduğu görülmüştür. Bu çalışma, bu boşluğu doldurmak için geri beslemeli üç katmanlı bir model önermiştir.

Modelin çıktığı noktası, veri toplama, analiz ve öğrenme, karar ve eylem katmanlarının bir bütün olarak çalışması gerektiği üzerine kurulmuştur. Aralarındaki ilişki döngüsel olarak düşünülmüştür. Karar katmanına ait çıktı, bir sonraki veri toplama döngüsünü şekillendirmektedir. Bu döngüsellik, modelin merkezi argümanını oluşturmuştur. Kavramsal çerçeve bölümündeki bu merkezi argüman, IoT mimarisinin katmanlı yapısıyla eşleştirilmiştir. Literatür taraması bölümünde bu eşleştirmenin, dağınık uygulama çalışmalarında tekrar tekrar gözlemlenen ama sistematik olarak modellenmeyen bir örüntü olduğu gösterilmiştir. Önerilen model bölümünde ise bu örüntü, f, g ve h fonksiyonlarıyla biçimsel hale getirilmiştir. Tartışma

bölümünde ise bu modelin şartlara bağlı olduğu ortaya konulmuştur. Veri güvenilirliği, tepki süresi ve açıklanabilirlik, modelin kısıtını oluşturan üç eşik olarak belirlenmiştir.

Bu bulguların bir araya gelişi, bölümün başında ortaya konan iddiayı doğrulamıştır. Yapay zeka çağında süreç optimizasyonu, artık tek seferlik bir analiz çıktısı olmaktan çıkmıştır. Sürekli güncellenen, öğrenen ve kendi girdisini yeniden şekillendiren bir döngü haline gelmiştir. Bu döngüsellik, mühendislik pratiğinde de bir değişimi zorunlu kılmaktadır. Mühendis, yalnızca bir süreç tasarlamaktan çok, bu sürecin kendi verisiyle nasıl etkileşime gireceğini de tasarlamaktadır.

Bununla birlikte bu bölümün kapsamı kasıtlı olarak sınırlı tutulmuştur. Herhangi bir uygulama örneğine yer verilmemiştir. Sayısal bir sonuç paylaşılmamıştır. Bunun nedeni, alanın ihtiyaç duyduğu kavramsal temeli önceliklendirilmesi içindir. Ancak bu tercih, aynı zamanda modelin en açık kısıtını da beraberinde getirmektedir. Model, henüz ampirik olarak test edilmemiştir. Bu yüzden, bölümün katkısı, belirli ve kesin bir çözüm sunmaktan çok, sonraki çalışmalarda test edilebilecek açık ve ölçülebilir bir çerçeve sunmaktır.

Gelecek çalışmalar için önerilebilecek üç yön bulunmaktadır. İlki, tartışma bölümünde ortaya konan geri beslemenin doğrulanması olacaktır. Farklı otomasyon düzeylerindeki endüstriyel ortamlarda, geri besleme döngüsünün tamamlanma frekansı ve karar gecikmesi arasındaki ilişki sayısal olarak ölçülmelidir. Bu, modelin somut ve en kolay bir şekilde test edilebilir iddiasıdır.

İkinci olarak, açıklanabilirlik ve karar kalitesi arasındaki dengenin araştırılması gerektiğidir. Analiz katmanına açıklanabilir yapay zeka yöntemleri entegre edilirse, karar katmanındaki insan operatörlerinin sistemin önerdiği kararı kabul etme oranını nasıl etkilediği araştırılmalıdır [18]. Bu araştırma yönü, modelin teknikle beraber, örgütsel bir boyutunun da olduğunu ortaya koyacaktır.

Sonuncusu ise, modelin sektörler arası genellenebilirliğinin test edilmesidir. Model, bu bölümde sadece soyut olarak sunulmuştur. Farklı sektörlerde farklı sinir ağı mimarileriyle nasıl somutlaşacağına değinilmemiş ve açık bırakılmıştır. Tarım, otomotiv, gıda, tekstil vb farklı otomasyonlarda, karşılaştırmalı vaka çalışmaları, modelin hangi koşullarda geçerliliğini koruduğunu, hangi koşullarda yeniden düşünülmesi gerektiğini gösterecektir.

Son olarak, bu bölüm daha geniş bir çıkarıma daha sahiptir. Dijital dönüşüm literatüründe kavramsal model geliştirmek, çoğu zaman uygulamalı çalışmaların gölgesinde kalmaktadır.

Ancak sağlam bir kavramsal temel yoksa, uygulamalı çalışmalar birbirinden kopuk vaka örnekleri olarak kalmaya devam edecektir. Bu bölüm, bu kopukluğu azaltmak için somut bir adım olarak konumlanmaktadır. Nesnelerin interneti ve yapay zeka entegrasyonu, yalnızca bir entegrasyondan ziyade, süreç optimizasyonunu yeniden tanımlayan yapısal bir mekanizmadır. Bu bölümün önerdiği model, bu mekanizmayı görünür kılmayı amaçlamıştır.

Kaynaklar

- [1] E. Villar, I. M. Toral, I. Calvo, O. B.- Sensors, “Architectures for industrial AIoT applications,”, 2024.
- [2] F. B. Fitch, “McCulloch Warren S. and Pitts Walter. A logical calculus of the ideas immanent in nervous activity. Bulletin of mathematical biophysics, vol. 5 , pp. 115–133,” *Journal of Symbolic Logic*, vol. 9, no. 2, pp. 49–50, Jun. 1944, doi: 10.2307/2268029.
- [3] F. W. Mateo and A. Redchuk, “Artificial Intelligence as a Process Optimization Driver under Industry 4.0 Framework and the Role of IIoT, a Bibliometric Analysis,” *Journal of Industrial Integration and Management*, vol. 9, no. 3, pp. 357–372, Sep. 2024, doi: 10.1142/S2424862222500130.
- [4] K. Tran, “Artificial intelligence for smart manufacturing: methods and applications,” *Sensors*, 21.16 (2021): 5584
- [5] P. P. Ray, “A survey on Internet of Things architectures,” *Journal of King Saud University - Computer and Information Sciences*, vol. 30, no. 3, pp. 291–319, Jul. 2018, doi: 10.1016/J.JKSUCI.2016.10.003.
- [6] J. Lin, et al., “A survey on internet of things: Architecture, enabling technologies, security and privacy, and applications,” *IEEE internet of things journal*, 2017
- [7] F. Alshohoumi et al., “Systematic review of existing IoT architectures security and privacy issues and concerns,” *International Journal of Advanced Computer Science and Applications*, 2019.
- [8] K. P. Seng, L. M. Ang, and E. Ngharamike, “Artificial intelligence Internet of Things: A new paradigm of distributed sensor networks,” *Int. J. Distrib. Sens. Netw.*, vol. 18, no. 3, Mar. 2022, doi: 10.1177/15501477211062835.
- [9] F. Ishengoma, et al., “Integration of artificial intelligence of things (AIoT) in the public sector: drivers, barriers and future research agenda,” *Digital Policy, Regulation and Governance*, 2022
- [10] A. Kausik, et al., “Machine learning algorithms for manufacturing quality assurance: A systematic review of performance metrics and applications,” , *Array* 26 (2025): 100393.

- [11] O. M. Matamoros, et al., “Artificial intelligence for quality defects in the automotive industry: A systemic review,” *Sensors* 25.5 (2025): 1288.
- [12] Z. Kang, et al., “Remaining useful life (RUL) prediction of equipment in production lines using artificial neural networks,” *Sensors* 21.3 (2021): 932.
- [13] M. El Ghadoui, et al., “A hybrid optimization approach for intelligent manufacturing in plastic injection molding by using artificial neural network and genetic algorithm,” *Scientific reports*, 2023
- [14] A. Aboshosha, et al., “IoT-based data-driven predictive maintenance relying on fuzzy system and artificial neural networks,” *Scientific Reports*, 2023
- [15] F. Alshohoumi, et al., “Systematic review of existing IoT architectures security and privacy issues and concerns,” *International Journal of Advanced Computer Science and Applications*, 2019.
- [16] A. Aboshosha, et al, “IoT-based data-driven predictive maintenance relying on fuzzy system and artificial neural networks,” *Scientific Reports*, 2023
- [17] A. Das, et al., “Opportunities and Challenges in Explainable Artificial Intelligence (XAI): A Survey,” Jun. 2020,
- [18] C. Agostinho *et al.*, “Explainability as the key ingredient for AI adoption in Industry 5.0 settings,” *Front. Artif. Intell.*, vol. 6, 2023, doi: 10.3389/FRAI.2023.1264372/FULL.

BÖLÜM 4

TIBBİ GÖRÜNTÜ SINIFLANDIRMASINDA BİLGİ DAMITIMI VE VERİ VERİMLİ GÖRÜNTÜ DÖNÜŞTÜRÜCÜLERİ

Aynur SEVİNÇ¹

1. Giriş ve Kavramsal Çerçeve

1.1. Tıbbi Görüntülemede Derin Öğrenmenin Gelişimi ve Transformatör Paradigması

Tıbbi görüntüleme teknolojileri, radyo-patolojik değerlendirmelerin hassasiyetini artırarak klinik karar destek sistemlerinin merkezinde yer almaktadır. Beyin tümörleri gibi yapısal ve hücresel açıdan yüksek düzeyde heterojenlik gösteren doku anomalilerinin tespiti; Manyetik Rezonans Görüntüleme (MRG), Bilgisayarlı Tomografi (BT) ve Pozitron Emisyon Tomografisi (PET) gibi gelişmiş klinik modalitelerin sunduğu çok dizilimli verilerin doğru analiz edilmesini gerektirir. Son yirmi yılda, bilgisayarlı görme ve yapay zekâ alanındaki teknolojik dönüşüm, geleneksel elle tasarlanmış öznitelik çıkarma tekniklerinden derin öğrenme odaklı otonom temsil öğrenme mimarilerine doğru radikal bir kayma göstermiştir.

Evrişimli Sinir Ağları (ESA/CNN), piksel düzeyindeki yerel komşuluk ilişkilerini evrişim çekirdekleri aracılığıyla işleme yeteneği sayesinde tıbbi görüntü sınıflandırma, segmentasyon ve lokalizasyon görevlerinde uzun süre altın standart olarak kabul edilmiştir. Bununla birlikte, evrişim operasyonunun doğasından kaynaklanan sınırlı alıcı alan (receptive field) kısıtı, dokular arası uzun mesafeli bağıntıların ve küresel bağlamsal bilgilerin bütüncül biçimde modellenmesini zorlaştırmaktadır. 2020 yılında doğal dil işleme alanından bilgisayarlı görme disiplinine uyarlanan Vision Transformer (ViT) mimarisi, görüntüleri yamalara (patch) bölerek öz-dikkat (self-attention) mekanizması üzerinden tüm pikseller arasındaki küresel ilişkileri eşzamanlı haritalandırma yeteneğiyle yeni bir paradigmaya öncülük etmiştir (Dosovitskiy vd., 2020).

Transformatör mimarisinin bilgisayarlı görmeye entegrasyonu, piksel düzeyindeki evrişim sınırlamalarını aşarak görüntünün tamamında birbiriyle uzak mesafeli bağ kurabilen öznitelik haritalarının türetilmesini sağlamıştır. Tıbbi görüntülerde lezyon sınırlarının belirsizliği, kistik ve nekrotik alanların doku karmaşıklığı düşünüldüğünde, ViT mimarisinin sunduğu bütüncül bağlam yaklaşımı erken evre tümör teşhisinde büyük avantaj vaat etmektedir (Tummala vd., 2022).

1.2. Sağlık Sektöründe Veri Kısıtları ve Yüksek Model Karmaşıklığı Problemi

Transformatör tabanlı derin öğrenme modellerinin sunduğu yüksek küresel temsil kapasitesi, teorik açıdan muazzam bir esneklik sağlasa da, bu esneklik beraberinde ciddi metodolojik ve donanımsal zorluklar getirmektedir. Transformatör mimarileri, evrişimli ağların barındırdığı kaydırma değişmezliği ve yerellik gibi önsel indüktif yanlılıklardan yoksundur. Bu durum, ViT

ve türevi modellerin doğru genelleme yapabilmesi için milyonlarca parametreyi eğitebilecek devasa ölçekli veri kümelerine ihtiyaç duymasına yol açmaktadır (Tummala vd., 2022).

Oysa sağlık sektöründe tıbbi görüntülerin elde edilmesi son derece meşakkatli, maliyetli ve etik izin süreçlerine tabi bir prosedürdür. Hasta mahremiyeti yasaları, nadir rastlanan hastalık biyomarkerları ve uzman radyologlar tarafından gerçekleştirilen etiketleme işlemlerinin zaman alıcılığı nedeniyle, tıbbi yapay zekâ çalışmalarında kullanılabilen veri kümeleri genellikle yüzler veya birkaç bin görüntü ile sınırlı kalmaktadır. Yetersiz veri hacmiyle eğitilen devasa transformatör modelleri, şiddetli aşırı uydurma (overfitting) ve düşük doğrulama başarısı göstermektedir.

Ayrıca, büyük transformatör modellerinin gerektirdiği devasa bellek alanı (VRAM) ve yüksek grafik işlemci (GPU) işlem gücü, bu sistemlerin kaynak sınırlaması bulunan klinik mobil cihazlarda, taşınabilir ultrason/MRG tarayıcılarında veya hastane içi kenar bilişim (edge computing) sunucularında gerçek zamanlı çalıştırılmasını engellemektedir. Bu metodolojik çelişki, yüksek başarımlı sunan devasa modellerin bilgisini daha küçük, hafif ve hızlı modellere aktarma ihtiyacını doğurmuş ve Bilgi Damıtımı (Knowledge Distillation) tekniklerini derin öğrenmenin en kritik araştırma alanlarından biri haline getirmiştir.

1.3. Bölümün Amacı, Kapsamı ve Yöntemi

Bu kitap bölümünün temel amacı, tıbbi görüntü analizi ve beyin tümörü sınıflandırmasında kullanılan transformatör tabanlı mimariler ile bilgi damıtma tekniklerinin kuramsal altyapısını, matematiksel formülasyonlarını ve literatürdeki uygulamasını derinlemesine sunmaktır. Çalışma, herhangi bir özgün deneysel veri setinin ölçüm çıktılarına odaklanmak yerine; tamamen teorik sentez, mimari karşılaştırmalar, matematiksel kayıp fonksiyonları ve literatürde kabul görmüş akademik yayınların nitel analizi esasına dayanan bir derleme (review study) şeklinde kurgulanmıştır.

2. Transformatör Tabanlı Görme Modellerinin (Vision Transformers) Teorik Esasları

2.1. Dikkat (Self-Attention) Mekanizması ve Matematiksel Formülasyonu

Transformatör mimarisinin temel taşı, girdi özniteliklerinin birbiriyle olan ağırlıklı ilişkilerini hesaplayan Ölçeklenmiş Noktasal Çarpım Dikkat (Scaled Dot-Product Attention) mekanizmasıdır. Bir girdi dizisindeki her bir eleman; Sorgu (Query-Q), Anahtar (Key-K) ve Değer (Value-V) olmak üzere üç farklı vektörel temsile dönüştürülür. Bu matrisler, girdi matrisinin (X) eğitilebilir ağırlık matrisleriyle (W^Q, W^K, W^V) çarpılmasıyla elde edilir:

$$Q = X \cdot W^Q, \quad K = X \cdot W^K, \quad V = X \cdot W^V \quad (1)$$

Sorgu ve Anahtar matrislerinin skaler çarpımı, dizideki tüm elemanların birbiriyle olan benzerlik skorlarını üretir. Skaler çarpım değerlerinin büyüklüğünden kaynaklanabilecek gradyan patlamalarını önlemek amacıyla skorlar, anahtar vektörünün boyutunun kareköküne ($\sqrt{d_k}$) bölünerek ölçeklenir ve ardından Softmax aktivasyon fonksiyonundan geçirilir. Elde edilen dikkat ağırlıkları Değer matrisi ile çarpılarak nihai bağlamsal temsil matrisi elde edilir:

$$\text{Attention}(Q, K, V) = \text{Softmax} \left(\frac{Q \cdot K^T}{\sqrt{d_k}} \right) \cdot V \quad (2)$$

Çok Kafalı Dikkat (Multi-Head Attention - MHA) yapısı ise bu işlemi paralel olarak farklı alt uzaylarda (h sayıda kafa) gerçekleştirerek, modelin görüntünün hem yüksek frekanslı ince detaylarına hem de düşük frekanslı küresel anlamsal dokularına aynı anda odaklanmasını sağlar (Vaswani vd., 2017; Dosovitskiy vd., 2020).

2.2. Vision Transformer (ViT) Mimarisi: Yamalama ve Pozisyonel Gömme

İki boyutlu bir tıbbi MR görüntüsünün ($H \times W \times C$) standart bir transformatör kodlayıcısına beslenebilmesi için, görüntünün $P \times P$ boyutlarında çakışmayan N adet yamaya (patch) bölünmesi gerekir. Burada yama sayısı $N = (H \cdot W) / P^2$ olarak hesaplanır. Elde edilen her iki boyutlu yama düzleştirilerek (flatten) bir vektöre dönüştürülür ve lineer bir projeksiyon katmanı aracılığıyla D boyutlu gizli temsil uzayına aktarılır.

Transformatör kodlayıcısı girdi dizisindeki elemanların sırasını veya mekânsal konumunu kendiliğinden algılayamadığından, yama gömmelerine eğitilebilir iki boyutlu Pozisyonel Gömmeler (Positional Embeddings - E_{pos}) eklenir. Ayrıca, tüm görüntüyü temsil eden küresel bir sınıflandırma vektörü elde etmek amacıyla girdi dizisinin en başına özel bir Sınıflandırma Belirteci ([Class Token] - x_{class}) dahil edilir (Dosovitskiy vd., 2020; Kucukkulahli ve Kabakus, 2024). ViT kodlayıcısına giren nihai vektör dizisi şu şekilde ifade edilir:

$$z_0 = [x_{class}; x_{p^1} \cdot W; x_{p^2} \cdot W; \dots; x_{p^N} \cdot W] + E_{pos} \quad (3)$$

2.3. ESA ile Transformatörlerin Karşılaştırmalı Analizi ve Mimari Ön Kabuller

Evrişimli Sinir Ağları ile Vision Transformer modelleri arasındaki temel metodolojik farklar İndüktif Yanlılık (Inductive Bias) kavramında düğümlenmektedir. ESA'lar, kaydırma değişmezliği (bir nesnenin görüntüdeki konumu değişse de evrişim filtresinin aynı tepkiyi vermesi) ve yerellik (yakın piksellerin birbiriyle daha ilişkili olması) varsayımlarını mimari tasarımlarında sert bir önkoşul olarak barındırır. Bu durum, ESA'ların küçük veri kümeleriyle

eğitildiğinde dahi hızlı yakınsamasını ve kararlı performans göstermesini sağlar (Doğan ve Türkoğlu, 2019; Mauricio vd., 2023).

Buna karşın Vision Transformer mimarileri hiçbir yerellik veya konumsal önkoşul varsaymaz. Model, pikseller arasındaki tüm uzamsal ilişkileri veri üzerinden sıfırdan öğrenmek zorundadır. Bu esneklik, devasa veri kümelerinde ESA'ların performans tavanını aşmasını sağlarken; tıbbi görüntüleme gibi sınırlı veri barındıran alanlarda ciddi başarımların kayıplarına yol açmaktadır. Şekil 1, bu iki paradigma arasındaki yapısal ve temsilsel zıtlıkları özetlemektedir.



Şekil 1. Evrişimli Sinir Ağları (CNN) ile Vision Transformer (ViT) mimarilerinin temsil özelliklerinin karşılaştırılması

2.4. Gelişmiş Transformatör Türevleri: Swin, ResT, ViViT ve Swin-UNet

Standart ViT mimarisinin karesel hesaplama karmaşıklığını $O(N^2)$ ve sınırlı çözünürlük ölçeklenebilirliğini aşmak amacıyla çeşitli gelişmiş transformatör türevleri geliştirilmiştir:

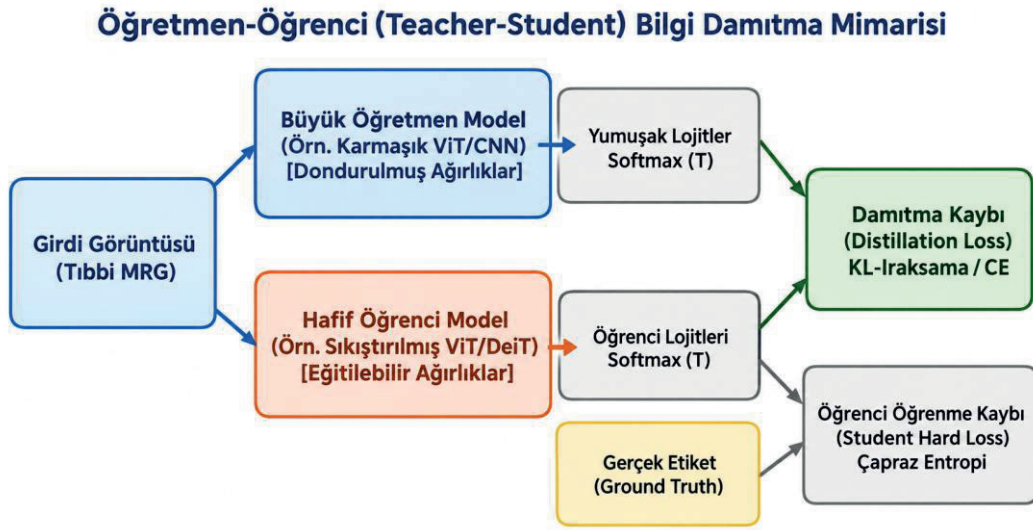
- *Swin Transformer (Shifted Windows)*: Görüntüyü kaydırılmış pencereler halinde yerel bloklara bölerek dikkat hesaplamasını doğrusal karmaşıklığa $O(N)$ indirger ve hiyerarşik öznitelik haritaları üretir (Liu vd., 2021).
- *ResT (Efficient Transformer)*: Verimli dikkat mekanizmaları ve hafıza sıkıştırma katmanları kullanarak bellek tüketimini optimize eder (Zhang ve Yang, 2021).
- *ViViT (Video Vision Transformer)*: Üç boyutlu hacimsel tıbbi MRG ve dinamik perfüzyon taramalarını uzamsal-zamansal (spatiotemporal) dikkat kurgusuyla işler (Arnab vd., 2021).
- *Swin-UNet*: Medikal segmentasyon görevleri için saf transformatör tabanlı kodlayıcı-kodçözücü (encoder-decoder) ve atlama bağlantıları (skip connections) sunan hiyerarşik bir mimaridir (Alzahrani ve Kahtani, 2024).

3. Bilgi Damıtması (Knowledge Distillation - KD) Yöntembilimi ve Matematiksel Altyapısı

3.1. Öğretmen-Öğrenci Öğrenme Paradigması ve Örtük Bilgi (Dark Knowledge)

Bilgi Damıtımı (Knowledge Distillation - KD), Geoff Hinton ve arkadaşları (2015) tarafından önerilen ve büyük, yüksek kapasiteli bir 'Öğretmen' (Teacher) modelin barındırdığı karmaşık temsil yeteneğini, daha küçük ve hafif bir 'Öğrenci' (Student) modele aktarmayı hedefleyen bir model sıkıştırma (model compression) tekniğidir. Geleneksel derin öğrenme eğitiminde modeller yalnızca sert etikete (one-hot encoded ground truth) odaklanırken, eğitilmiş bir öğretmen modelin çıktı olasılık dağılımı (lojitler), sınıflar arası gizli benzerlik ilişkilerini de barındırır.

Öneğin, bir beyin MR görüntüsünün sınıflandırılmasında öğretmen model bir görüntüyü %85 Gliom, %12 Meningiom ve %3 Hipofiz tümörü olarak tahmin ettiğinde; Gliom dokusunun Meningiom dokusuyla olan biyolojik/görsel benzerliği öğrenci modele aktarılmış olur. Hinton vd. (2015) bu ikincil olasılık dağılımlarını 'Örtük Bilgi' (Dark Knowledge) olarak adlandırmıştır. Örtük bilgi, öğrenci modelin veri kümesinde bulunmayan veya doğrudan etiketlenmeyen karmaşık örüntüleri çok daha hızlı ve kararlı biçimde öğrenmesini sağlar (Gou vd., 2021; Barulina vd., 2022).



Şekil 2. Öğretmen-öğrenci (teacher-student) bilgi damıtımı genel mimari şeması

Şekil 2’de, öğretmen-öğrenci yaklaşımına dayalı bilgi damıtma mimarisi gösterilmektedir. Aynı MRG görüntüsü, ağırlıkları dondurulmuş büyük öğretmen modele ve eğitilebilir hafif öğrenci modele aktarılmaktadır. Öğrenci model, öğretmenin sıcaklık parametresiyle yumuşatılmış çıktı dağılımını damıtma kaybı aracılığıyla öğrenirken, gerçek etiketlerle

hesaplanan çapraz entropi kaybıyla da yönlendirilmektedir. Böylece öğretmen modelin öğrendiği bilginin daha düşük hesaplama maliyetine sahip öğrenci modele aktarılması amaçlanmaktadır.

3.2. Bilgi Damıtım Taksonomisi: Lojit, Öznitelik ve İlişki Tabanlı Damıtma

Literatürde bilgi damıtma teknikleri, bilginin öğretmen modelden çıkarıldığı katman seviyesine ve aktarım biçimine göre üç ana kategoride sınıflandırılmaktadır (Gou vd., 2021):

1. Lojit Tabanlı Damıtma (Response-Based KD): Öğretmen modelin son katmandaki olasılık çıktılarını (logits) hedef alarak öğrenci modelin tahmin dağılımını öğretmene yaklaştırır (Hinton vd., 2015).

2. Öznitelik Tabanlı Damıtma (Feature-Based KD): Öğretmen modelin ara gizli katmanlarındaki (intermediate layers) öznitelik haritalarını (feature maps) öğrenci modelin ara katmanlarına hizalar (Alotaibi vd., 2023).

3. İlişki Tabanlı Damıtma (Relation-Based KD): Farklı veri örnekleri veya pikseller arasındaki etkileşim ve korelasyon matrislerini (relation graphs) öğretmen modelden öğrenciye transfer eder (Alzahrani ve Kahtani, 2024).

3.3. Yumuşak (Soft) ve Sert (Hard) Damıtma Mekanizmaları

Lojit aktarımında bilginin öğrenciye aktarılma biçimine göre iki temel yaklaşım bulunur. Yumuşak Damıtma (Soft Distillation) yönteminde, öğretmen modelin ürettiği yumuşatılmış olasılık dağılımı ile öğrenci modelin olasılık dağılımı arasındaki fark Kullback-Leibler (KL) Iraksaması fonksiyonu ile minimize edilir (Hinton vd., 2015).

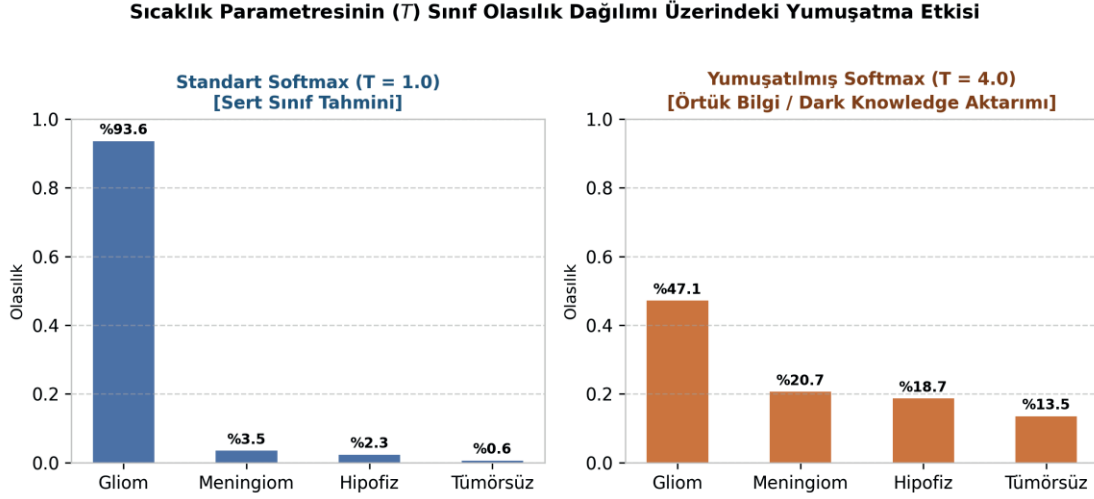
Sert Damıtma (Hard Distillation) yönteminde ise, öğretmen modelin en yüksek olasılık gösterdiği sınıf tahmini ($y_t = \text{argmax}_c(z_t)$) sert bir ‘öğretmen etiketi’ olarak kabul edilir ve öğrenci model bu etiket üzerinden Çapraz Entropi (Cross-Entropy) kaybı ile eğitilir. Sert damıtma, özellikle nesnel sınıf sınırlarına sahip tıbbi görüntüleme görevlerinde ekstra parametre tünellemesine ihtiyaç duymadan yüksek başarımlar sunmaktadır (Touvron vd., 2021; Alotaibi vd., 2023).

3.4. Sıcaklık Parametresi (T) ve Olasılık Yumuşatma Matematiği

Standart Softmax fonksiyonunda, doğru sınıfın lojit değeri diğer sınıflara oranla çok yüksek olduğunda, ikincil sınıfların olasılıkları sıfıra yaklaşır ve Örtük Bilgi kaybolur. Bu durumu engellemek amacıyla lojit değerleri bir Sıcaklık Hiperparametresine (T) bölünür:

$$q_i = \exp(z_i / T) / [\sum_j \exp(z_j / T)] \quad (4)$$

$T=1$ olduğunda denklem standart Softmax fonksiyonuna dönüşür. $T > 1$ değerleri seçildiğinde (genellikle T in $[3, 8]$), sınıf olasılık dağılımı yumuşayarak ikincil sınıflar arasındaki ince oranları belirginleştirir (Hinton vd., 2015; Gou vd., 2021). Şekil 3, sıcaklık parametresinin olasılık dağılımı üzerindeki yumuşatma etkisini görselleştirmektedir.



Şekil 3. Sıcaklık parametresinin (T) olasılık dağılımını yumuşatma ve örtük bilgiyi açığa çıkarma etkisi

Şekil 3'te, sıcaklık parametresinin artırılmasıyla Softmax olasılık dağılımının yumuşaması gösterilmektedir. $T=1$ 'de tahmin tek bir sınıfta yoğunlaşırken, $T=4$ 'te diğer sınıfların göreceli olasılıkları belirginleşmektedir. Böylece sınıflar arasındaki benzerlikleri yansıtan bilgi, damıtma sürecinde öğrenci modele aktarılabilmektedir.

3.5. Çapraz Entropi ve KL-İraksama Kayıp Fonksiyonları

Bilgi damıtımıyla eğitilen bir öğrenci modelin toplam kayıp fonksiyonu (L_{toplam}), gerçek etiketlerden gelen Öğrenci Kaybı (L_{CE}) ile öğretmen modelden gelen Damıtma Kaybının (L_{KD}) ağırlıklı toplamından oluşur:

$$L_{\text{toplam}} = (1 - \alpha) \cdot L_{\text{CE}}(y, \sigma(z_s)) + \alpha \cdot T^2 \cdot L_{\text{KL}}(\sigma(z_s / T), \sigma(z_t / T)) \quad (5)$$

Burada α in $[0, 1]$ damıtma ağırlık katsayısını, z_s öğrenci lojitlerini, z_t öğretmen lojitlerini ve T^2 ölçekleme faktörünü temsil etmektedir. T^2 terimi, yumuşatılmış lojitlerin türevleri alınırken gradyan büyüklüklerinin ölçeğini korumak amacıyla denkleme dahil edilir (Hinton vd., 2015; Touvron vd., 2021).

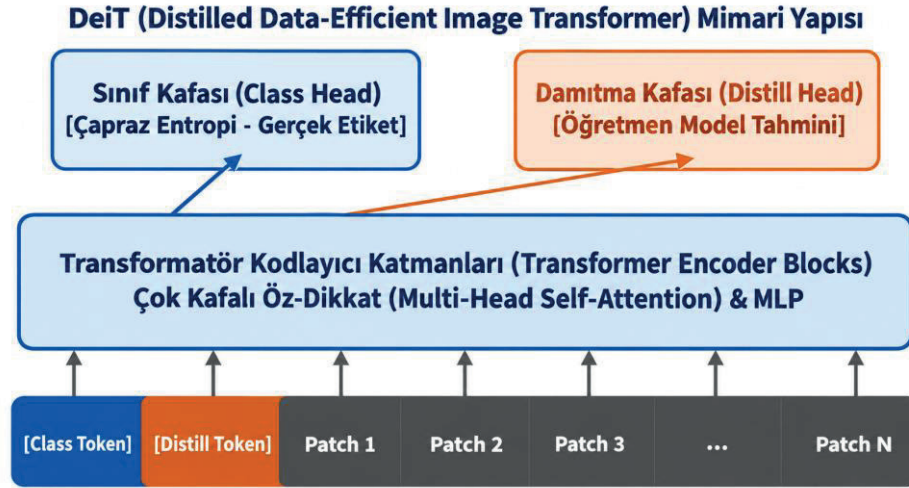
4. Transformatör Tabanlı Modellerde Bilgi Damıtımı Mimarileri

4.1. Distilled Data-Efficient Image Transformer (DeiT)

Facebook AI Research (FAIR) tarafından geliştirilen DeiT (Data-Efficient Image Transformer) mimarisi, Vision Transformer modellerinin devasa veri kümelerine olan bağımlılığını bilgi damıtımı yöntemiyle kıran çığır açıcı bir çalışmadır (Touvron vd., 2021). DeiT, harici büyük veri kümelerine ihtiyaç duymadan, yalnızca standart ölçekteki veriler ve güçlü bir Evrişimli Sinir Ağı (CNN - örn. RegNet veya ResNet) öğretmen modeli rehberliğinde yüksek performans elde etmeyi başarmıştır (Jalalifar ve Sadeghi-Naini, 2022).

4.1.1. Damıtma Belirteci (Distillation Token) ve Çift Sınıflandırma Kafası

DeiT mimarisinin en özgün yeniliği, girdi dizisine [Class Token] yanında özel bir [Distillation Token] (Damıtma Belirteci) eklemesidir. Şekil 4'te görüldüğü üzere, damıtma belirteci tıpkı sınıflandırma belirteci gibi öz-dikkat katmanları boyunca yama gömmeleriyle etkileşime girer. Modelin çıkışında iki ayrı kafaya yer verilir: Sınıf Kafası (Class Head) gerçek veri etiketleriyle Çapraz Entropi kaybını hesaplar; Damıtma Kafası (Distillation Head) Öğretmen modelin tahmin çıktısını hedef alır (Touvron vd., 2021; Yin vd., 2022).



Şekil 4. DeiT (Distilled Data-Efficient Image Transformer) mimarisinde damıtma belirteci ve çift kafa yapısı

4.1.2. DeiT Sert Damıtma Matematiksel Formülasyonu

DeiT mimarisinde kullanılan Sert Damıtma (Hard-label Distillation) loss fonksiyonu matematiksel olarak şu şekilde ifade edilmektedir (Touvron vd., 2021):

$$L_{\text{global}}^{\text{HardDistill}} = (1/2) \cdot L_{\text{CE}}(\psi(y_s), y) + (1/2) \cdot L_{\text{CE}}(\psi(y_d), y_t) \quad (6)$$

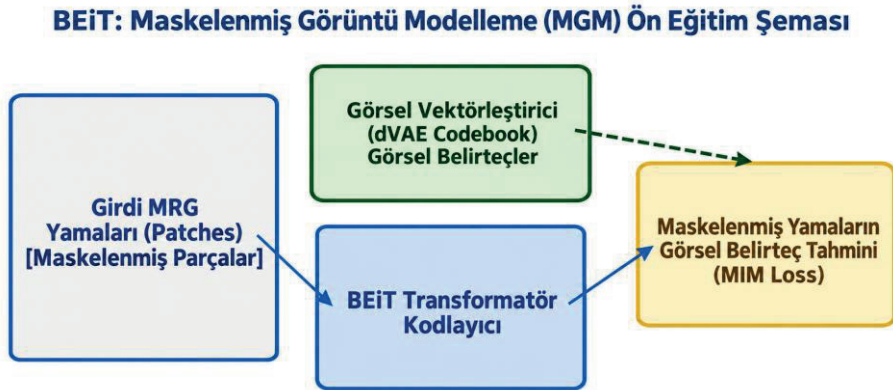
Denklemden ψ Softmax fonksiyonunu, y_s sınıf belirteci tahminini, y_d damıtma belirteci tahminini, y gerçek etiketi ve $y_t = \arg\max_c(z_t)$ öğretmen modelin sert tahmin sınıfını temsil eder. Test/inference aşamasında sınıf kafası ve damıtma kafasının ürettiği lojitler toplanarak nihai karar verilir (Touvron vd., 2021; Barulina vd., 2022).

4.2. BERT Görüntü Dönüştürücülerinin Ön Eğitimi (BEiT)

Microsoft Research tarafından geliştirilen BEiT (BERT Pre-Training of Image Transformers) mimarisi, doğal dil işlemedeki BERT ön eğitim metodolojisini bilgisayarlı görme alanına aktaran bir diğer öncü transformatör modelidir (Bao vd., 2021). BEiT, etiketlenmemiş veri kümeleri üzerinde öz-denetimli (self-supervised) olarak ön eğitilebilme yeteneği sayesinde tıbbi görüntü analizinde etiketli veri eksikliğine etkili bir çözüm sunar (Nhut vd., 2024).

4.2.1. Maskelenmiş Görüntü Modelleme (MGM) ve Görsel Vektörleştirme (dVAE)

BEiT mimarisinde görüntüler 16x16 yamalara ayrıldıktan sonra, yamaların yaklaşık %40-%50'si rastgele maskelenir ([MASK]). Model, maskelenmiş yamaların ham piksel değerlerini tahmin etmek yerine, önceden eğitilmiş bir Ayrık Varyasyonel Oto-Kodlayıcı (dVAE / Vector Quantized VAE) kütüphanesi aracılığıyla elde edilen Görsel Belirteçleri (Visual Tokens) tahmin etmeye odaklanır (Bao vd., 2021; Rajesh vd., 2024). Şekil 5, BEiT ön eğitim şemasını göstermektedir.



Şekil 5. BEiT mimarisinde Maskelenmiş Görüntü Modelleme (MGM) ve görsel belirteç tahmin akışı

Şekil 5'te, BEiT modelinin maskelenmiş görüntü modellemeye dayalı ön eğitim süreci gösterilmektedir. Model, MRG görüntüsündeki maskelenen yamaların görsel belirteçlerini, görünür yamalardan yararlanarak tahmin etmeyi öğrenmektedir. Böylece etiketli verilere ihtiyaç duymadan görüntünün yapısal ve bağlamsal özelliklerini öğrenmesi amaçlanmaktadır.

4.3. Diğer Hafif Sıkıştırılmış Transformatör Mimarileri (LCDeiT, MobileViT)

DeiT ve BEiT modellerinin yanı sıra, mobil ve klinik cihazlarda çalıştırılmak üzere tasarlanmış çeşitli hafif sıkıştırılmış transformatör türevleri mevcuttur:

- *LCDeiT (Linear Complexity DeiT)*: Transformatör bloklarındaki karesel öznitelik karmaşıklığını lineer projeksiyon katmanlarıyla düşürerek bellek tüketimini minimize eder (Ferdous vd., 2023).
- *MobileViT*: Evrişim katmanlarının yerel işleme gücü ile transformatörlerin küresel dikkat yeteneğini sarmal biçimde birleştiren hafif mobil mimaridir (Alzahrani ve Kahtani, 2024).
- *EfficientFormer*: Mobil cihazlarda gecikme süresini (latency) düşürmek için evrişimli bloklar ve matris çarpım optimizasyonları içeren yeni nesil dönüştürücüdür (Shin vd., 2023).

5. Tıbbi Görüntü Analizinde Bilgi Damıtımı ve Transformatör Literatürü

5.1. Çoklu Sınıflandırma ve Beyin MRG Çalışmalarında Bilgi Damıtımı Uygulamaları

Literatürde beyin tümörlerinin sınıflandırılması (Gliom, Meningiom, Hipofiz ve Tümörsüz) ve derecelendirilmesi üzerine yürütülen çalışmalar incelendiğinde, transformatör mimarileri ile bilgi damıtma tekniklerinin entegrasyonunun performans üzerinde belirleyici bir rol oynadığı görülmektedir. Alzahrani ve Kahtani (2024), beyin MR görüntülerinde çok sınıflı tümör tespiti için üçlü dikkat (tripartite attention) mekanizmasına sahip bir bilgi damıtma çerçevesi tanıtmıştır. Araştırmacılar, karmaşık dönüştürücü modellerinin bilgisini hafif öğrenci ağlara aktararak veri artırımı uygulanmış medikal verilerde yüksek genelleme başarısı yakalamıştır.

Ferdous ve arkadaşları (2023), beyin tümörü MRG verilerinde DeiT modelini inceleyerek, lineer karmaşıklıkta LCDeiT mimarisinin geleneksel ağır ViT modellerine kıyasla bellek verimliliğini ve doğru tanı oranını belirgin şekilde artırdığını göstermiştir. Şahin, Özdemir ve Temurtaş (2024) ise ViT modellerinin çok sayıda kritik hiperparametreye sahip olmasından kaynaklanan eğitim zorluklarını Bayes Çok Amaçlı Optimizasyon yöntemiyle aşmış ve transformatör modellerinin medikal teşhisteki karar kararlılığını kanıtlamıştır. Amin ve arkadaşları (2025) ise DeiT modelini temel bileşenler analizi (PCA) ve ortalama farksal gelişim (ADE) algoritmasıyla hibritleştirerek gelişmiş bir öznitelik seçim yaklaşımı sunmuştur.

5.2. Diğer Tıbbi Alanlardaki Derleme Çalışmaları ve Literatür Sentezi

Bilgi damıtımı ve transformatör modellerinin başarısı yalnızca beyin MRG ile sınırlı kalmayıp, tıbbın diğer tanısal alanlarında da geniş bir literatür oluşturmuştur:

- *Radyolojik Görüntüleme*: Jalalifar ve Sadeghi-Naini (2022), göğüs radyografilerinde lezyon tespitinde DeiT mimarisinin katı etiketli damıtma yeteneğini incelemiş ve sınırlı etiketli radyoloji verilerinde yüksek duyarlılık elde edildiğini raporlamıştır.
- *Dermatolojik Lezyon Analizi*: Xin ve arkadaşları (2022) ile Alotaibi vd. (2023), dermatolojik doku görüntülerinde ViT ve DeiT modellerinin hibrit kullanımının hücresel yapıların ayırt edilmesindeki etkinliğini kanıtlamıştır.
- *Nörolojik Bozukluklar ve Beyin Taramaları*: Shin ve arkadaşları (2023) PET beyin taramalarında nörolojik değerlendirme için ViT ve VGG modellerini karşılaştırmış; Yin vd. (2022) ise SMIL-DeiT modeliyle erken evre teşhiste çoklu örneklemlili öğrenme ile bilgi damıtımını birleştirmiştir.
- *Kardiyolojik Sinyaller ve EKG*: Jain ve Menon (2023) ile Rajesh vd. (2024), elektrokardiyografi (EKG) sinyallerinin zaman serisi dönüşümlerinde BEiT modelinin maskelenmiş dil/görüntü ilkesiyle anormal ritim tespiti yaptığını göstermiştir.

5.3. Karşılaştırmalı Literatür İnceleme Tablosu

Aşağıdaki Tablo 1, literatürde tıbbi görüntüleme ve sinyal analizinde bilgi damıtımı ve transformatör tabanlı derin öğrenme modellerini kullanan başlıca çalışmaları metodolojik açıdan sentezlemektedir.

Tablo 1. Görsel transformatörler, bilgi damıtımı ve veri verimli öğrenme yaklaşımlarına ilişkin literatürün karşılaştırmalı özeti

Yazar(lar) ve Yıl	İncelenen Model	Görüntü / Veri Türü	Metodolojik Yaklaşım	Temel İnovasyon / Katkı
Alzahrani & Kahtani (2024)	KD + Tripartite ViT	Beyin MRG	Üçlü dikkat mekanizmalı KD	Çok sınıflı tümörlerde veri artırımı damıtma
Ferdous vd. (2023)	LCDeiT Model	Beyin MRG	Lineer karmaşık DeiT	Sınırlı bellekle yüksek performanslı sınıflandırma
Şahin vd. (2024)	Bayes Opt. + ViT	Beyin MRG	Çok amaçlı Bayes optimizasyonu	Hiperparametre optimizasyonu ile karar kararlılığı
Amin vd. (2025)	DeiT + PCA-ADE	Beyin MRG	Öznitelik seçimi ve damıtma	Boyut azaltma ile gelişmiş çoklu sınıflandırma
Touvron vd. (2021)	DeiT Mimarisi	Genel / Tıbbi Görüntü	Damıtma belirteci ve sert KD	CNN öğretmenle ImageNet-1k üzerinde saf ViT
Bao vd. (2021)	BEiT Mimarisi	Görsel Yamalar	Maskelenmiş Görüntü Modelleme	dVAE görsel belirteçleriyle öz-denetimli ön eğitim
Yin vd. (2022)	SMIL-DeiT	Beyin PET / MRG	Çoklu örneklemlili öğrenme + KD	Erken evre beyin teşhisinde öznitelik transferi
Jalalifar & Sadeghi (2022)	DeiT + ViT	Göğüs X-Ray	Saf ViT damıtması	Röntgen lezyonlarında verimli öğretmen aktarımı
Shin vd. (2023)	ViT vs VGG-19	PET Beyin Tarama	Karşılaştırmalı model analizi	Küçük veride ViT sınırlılıkları ve KD ihtiyacı
Rajesh vd. (2024)	BEiT Transformer	EKG / Biyosinyal	Zaman serisi görselleştirme	Biyosinyal ve ritim bozukluğu tespiti

Tablo 1’de görsel transformatörler, bilgi damıtımı ve veri verimli öğrenme yaklaşımlarını ele alan çalışmalar; kullanılan modeller, veri türleri, metodolojik özellikler ve temel katkılar açısından karşılaştırılmaktadır. Çalışmalar, özellikle beyin MRG görüntülerinin sınıflandırılmasında damıtma, öznetelik seçimi ve hiperparametre optimizasyonu gibi yöntemlere odaklanmaktadır. DeiT ve BEiT gibi temel mimariler ise öğretmen modelden bilgi aktarımı ve öz-denetimli ön eğitim yoluyla öğrenme verimliliğini artırmaya yönelik farklı yaklaşımları temsil etmektedir.

6. Model Sıkıştırma, Klinik Entegrasyon ve Açıklanabilir Yapay Zekâ (XAI)

6.1. Sınır Bilişim (Edge Computing) ve Mobil Klinik Cihazlarda Model Dağıtımı

Sağlık kuruluşlarında yapay zekâ çözümlerinin klinik pratiğe aktarılmasındaki en büyük donanımsal engel, yüksek işlem gücü gerektiren sunucu bağımlılığıdır. Radyoloji servislerinde, ameliyathanelerde ve mobil hastane arabalarında kullanılan tanı cihazları; sınırlı işlemci (CPU) ve grafik belleğine (RAM/VRAM) sahiptir. Bilgi damıtımı teknikleriyle küçültülen DeiT, MobileViT veya LCDeiT gibi öğrenci modeller, birkaç megabaytlık kapladıkları alan ve düşük gecikme süreleri (latency) sayesinde bu sınır bilişim (edge computing) cihazlarına doğrudan entegre edilebilmektedir (Barulina vd., 2022).

6.2. Zaman ve Hesaplama Karmaşıklığı Analizi

Bir derin öğrenme modelinin klinik elverişliliği değerlendirilirken yalnızca doğruluk metrikleri değil, FLOPS (Saniyedeki Yüzer Nokta İşlem Sayısı), Toplam Parametre Sayısı ve Çıkarım Süresi (Inference Latency) gibi matematiksel karmaşıklık parametreleri de dikkate alınmalıdır. Tablo 2, genel transformatör ve damıtılmış model ailelerinin donanımsal karmaşıklık dinamiklerini teorik düzeyde karşılaştırmaktadır.

Tablo 2. Görsel transformatör mimarilerinin parametre ölçeği, hesaplama karmaşıklığı ve klinik kullanıma uygunluk açısından karşılaştırılması

Model Mimarisi	Parametre Ölçeği	Hesaplama Karmaşıklığı	Klinik Dağıtım Uygunluğu
Ağır ViT (ViT-Huge / ViT-Large)	300M - 600M Parametre	Yüksek Karesel $O(N^2)$	Yalnızca Veri Merkezi / Bulut GPU
Standart ViT (ViT-Base / ViT-B16)	86M Parametre	Karesel $O(N^2)$	Yüksek Performanslı İş istasyonu
Damıtılmış DeiT (DeiT-Small / Tiny)	5M - 22M Parametre	Doğrusala Yakın Karmaşıklık	Klinik Kenar Sunucu ve Mobil Cihaz
BEiT (Masked Pre-trained)	88M Parametre (Eğitimde Ağır)	Ön Eğitimde Yüksek, Çıkarımda Orta	Klinik Karar Destek Sistemleri
MobileViT / LCDiT	2M - 6M Parametre	Düşük Lineer Karmaşıklık $O(N)$	Taşınabilir Ultrason / Mobil Cihazlar

Tablo 2, görsel transformatör mimarilerinin model büyüklüğü ve hesaplama gereksinimleri bakımından farklı kullanım olanakları sunduğunu göstermektedir. Büyük modeller daha güçlü donanımlar gerektirirken, küçük ve hafif mimariler kaynakları sınırlı ortamlarda kullanım açısından avantaj sağlayabilir. Bununla birlikte, klinik kullanıma uygunluk yalnızca parametre sayısına değil; doğruluk, çıkarım süresi, bellek tüketimi ve bağımsız doğrulama sonuçlarına da bağlıdır.

6.3. Tıbbi Karar Destek Sistemlerinde Açıklanabilirlik (XAI) ve Etik

Derin öğrenme modellerinin kara kutu niteliği, tıbbi alanda teşhis kararlarının klinik doktorlar tarafından benimsenebilmesi önündeki en büyük etik ve pratik engeldir. Bir modelin beyin tümörü teşhisi koyarken görüntünün hangi bölgesine odaklandığını göstermesi şeffaflık açısından şarttır. Transformatör modelleri, barındırdıkları öz-dikkat haritaları (self-attention maps) sayesinde görsel kararların doğrudan ısı haritaları (heatmaps) olarak sunulmasına olanak tanır. Açıklanabilir Yapay Zekâ (XAI) teknikleri (Grad-CAM, Attention Rollout, SHAP) ile entegre edilen damıtılmış transformatör modelleri, klinik hekimlerin yapay zekâ kararlarını doğrulamasını ve hatalı teşhis riskini azaltmasını sağlamaktadır (Nazir vd., 2024; Goceri, 2025).

7. Tartışma, Gelecek Araştırma Yönelimleri ve Sonuç

7.1. Mevcut Yöntemlerin Üstünlükleri ve Sınırlılıkları

Bu derleme çalışmasında sunulan teorik ve literatür analizleri, Bilgi Damıtımı ile desteklenmiş Transformatör modellerinin tıbbi görüntüleme alanında son derece güçlü bir potansiyele sahip olduğunu ortaya koymaktadır. Yöntemlerin temel üstünlükleri; etiketli veri ihtiyacını radikal biçimde düşürmesi, küresel bağlamsal ilişkileri başarıyla modellemesi ve hafif öğrenci mimariler sayesinde donanım maliyetlerini azaltmasıdır. Buna karşın, öğretmen modelin seçimi, sıcaklık parametresinin tünellenmesi ve ön eğitim süreçlerinin yüksek işlem maliyeti yöntemin temel sınırlılıklarını oluşturmaktadır (Hinton vd., 2015; Touvron vd., 2021).

7.2. Gelecek Çalışma Önerileri

Gelecekteki araştırmalar için öne çıkan kritik yönelimler şunlardır:

- Çok Modlu (Multimodal) Bilgi Damıtımı: Yalnızca MRG görselleri değil, hastanın klinik öyküsü, genetik biyomarkerları ve tanı raporlarının metin modelleriyle eşzamanlı damıtılması (Alzahrani ve Kahtani, 2024).
- Sentetik Veri ve Generatif Damıtma: Diffusion ve GAN modelleriyle üretilen sentetik tıbbi görüntüler üzerinden damıtma süreçlerinin optimizasyonu.
- Federe Öğrenme (Federated Learning) ile KD Entegrasyonu: Hastane verilerinin dışarı çıkarılmadan yerel cihazlarda damıtılarak küresel bir hekim modeli oluşturulması.

7.3. Sonuç

Sonuç olarak, Bilgi Damıtımı ve Transformatör tabanlı derin öğrenme mimarileri; veri kıtlığı ve yüksek hesaplama maliyetleri kısıcındaki medikal yapay zekâ disiplinine sürdürülebilir, güvenilir ve klinik olarak donatılabilir çözümler sunmaktadır. Bu kuramsal derleme, gelecekte geliştirilecek hafif tanı sistemleri ve radyo-patolojik karar destek yazılımları için sağlam bir teorik rehber niteliğindedir.

Kaynakça

- Alotaibi, A., et al. (2023). ViT-DeiT: An ensemble model for breast cancer histopathological images classification. In 2023 ICAISC (pp. 1–6). IEEE.
- Alzahrani, S. M. & Qahtani, A. M. (2024). Knowledge distillation in transformers with tripartite attention: Multiclass brain tumor detection in highly augmented MRIs. *Journal of King Saud University-Computer and Information Sciences*, 36(1), 101907.
- Amin, M., Nahar, K. M., Gharaibeh, H., Nasayreh, A., Alomar, A., Malkawi, M., ... & Abualigah, L. (2025). DieT Transformer model with PCA-ADE integration for advanced multi-class brain tumor classification. *Intelligence-Based Medicine*, 11, 100192.
- Arnab, A., Dehghani, M., Heigold, G., Sun, C., Lučić, M. & Schmid, C. (2021). Vivit: A video vision transformer. In *Proceedings of the IEEE/CVF ICCV* (pp. 6836-6846).
- Bao, H., Dong, L., Piao, S. & Wei, F. (2021). BEiT: BERT pre-training of image transformers. *arXiv preprint arXiv:2106.08254*.
- Barulina, M., Sanbaev, A., Okunkov, S., Ulitin, I. & Okoneshnikov, I. (2022). Deep learning approaches to automatic chronic venous disease classification. *Mathematics*, 10(19), 3571.
- Bauer, S., Wiest, R., Nolte, L. P. & Reyes, M. (2013). A survey of MRI-based medical image analysis for brain tumor studies. *Physics in Medicine & Biology*, 58(13), R97.
- Doğan, F. & Türkoğlu, İ. (2019). A compilation of deep learning models and application areas. *Dicle University Engineering Faculty Journal of Engineering*, 10(2), 409-445.
- Dosovitskiy, A. et al. (2020). An image is worth 16x16 words: Transformers for image recognition at scale. *arXiv preprint arXiv:2010.11929*.
- Ferdous, G. J., Sathi, K. A., Hossain, M. A., Hoque, M. M. & Dewan, M. A. A. (2023). LCDEiT: A linear complexity data-efficient image transformer for MRI brain tumor classification. *IEEE Access*, 11, 20337–20350.
- Goceri, E. (2025). An efficient network with CNN and transformer blocks for glioma grading and brain tumor classification from MRIs. *Expert Systems with Applications*, 268, 126290.
- Gou, J., Yu, B., Maybank, S. J. & Tao, D. (2021). Knowledge distillation: A survey. *International Journal of Computer Vision*, 129(6), 1789–1819.
- Hinton, G., Vinyals, O. & Dean, J. (2015). Distilling the knowledge in a neural network. *arXiv preprint arXiv:1503.02531*.
- Jain, B. S. U. & Menon, P. G. (2023). Classification of abnormal cardiac rhythm utilizing embeddings from transformer encoder models. In 2023 IEEE BIBM (pp. 4899-4901). IEEE.
- Jalalifar S. A. & Sadeghi-Naini, A. (2022). Data-Efficient Training of Pure Vision Transformers for Chest X-ray Abnormality Detection Using Knowledge Distillation. In 2022 IEEE EMBC (pp. 1444–1447). IEEE.

- Kucukkulahli, E. & Kabakus, A. T. (2024). A novel classification model based on vision transformers. *Applied Acoustics*, 226, 110218.
- Liu, Z., Lin, Y., Cao, Y., Hu, H., Wei, Y., Zhang, Z., ... & Guo, B. (2021). Swin transformer: Hierarchical vision transformer using shifted windows. In *Proceedings of the IEEE/CVF ICCV* (pp. 10012-10022).
- Maurício, J., Domingues, I. & Bernardino, J. (2023). Comparing Vision Transformers and Convolutional Neural Networks for Image Classification: A Literature Review. *Applied Sciences*, 13(9), 5521.
- Nazir, M. I., Akter, A., Wadud, M. A. H. & Uddin, M. A. (2024). Utilizing customized CNN for brain tumor prediction with explainable AI. *Heliyon*, 10(20), e38997.
- Nhut, D. T. N., Tan, T. D., Quoc, T. N. & Hoang, V. T. (2024). Recognition based on Vision Transformer and BEiT. *Procedia Computer Science*, 234, 188–195.
- Rajesh, N., Yalavarthy, A. B., Vamsi, V. K. & Saranya, G. (2024). BEiT transformer models to aid in early detection of illness. In *2024 ACCAI* (pp. 1–7). IEEE.
- Shin, H., Jeon, S., Seol, Y., Kim, S. & Kang, D. (2023). Vision transformer approach for classification using 18F-Florbetaben brain images. *Applied Sciences*, 13(6), 3453.
- Şahin, E., Özdemir, D. & Temurtaş, H. (2024). Multi-objective optimization of ViT architecture for efficient brain tumor classification. *Biomedical Signal Processing and Control*, 91, 105938.
- Touvron, H., Cord, M., Douze, M., Massa, F., Sablayrolles, A. and Jégou, H. (2021). Training data-efficient image transformers & distillation through attention. In *ICML*, PMLR, pp. 10347–10357.
- Tummala, S., Kadry, S., Bukhari, S. A. C. & Rauf, H. T. (2022). Classification of brain tumor from magnetic resonance imaging using vision transformers ensembling. *Current Oncology*, 29(10), 7498-7511.
- Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., ... & Polosukhin, I. (2017). Attention is all you need. *Advances in Neural Information Processing Systems*, 30.
- Xin, C., Li, Y., Zhang, F., Wang, H., Zhou, J. & Liu, Q. (2022). An improved transformer network for classification. *Computers in Biology and Medicine*, 149, 105939.
- Yin, Y., Jin, W., Bai, J., Liu, R. & Zhen, H. (2022). SMIL-DeiT: Multiple instance learning and self-supervised vision transformer network. In *2022 IJCNN* (pp. 1–6). IEEE.
- Zhang, Q. & Yang, Y. B. (2021). Rest: An efficient transformer for visual recognition. *Advances in Neural Information Processing Systems*, 34, 15475-15485.