

Alzheimer Erken Tanısında Çok Modlu Yapay Zeka Yaklaşımı

Dr. Öğr. Üyesi Gülay Çiçek.

Yusuf Talha Yılmaz¹

¹Yazılım Mühendisliği Bölümü, Mühendislik Mimarlık Fakültesi
İstanbul Beykent Üniversitesi, Sarıyer, İstanbul, Türkiye
¹2103013281@student.beykent.edu.tr,

I. LİTERATÜR İNCELEMESİ

Çalışmanın bu bölümü, Alzheimer hastalığının erken teşhisine yönelik yapay zeka temelli araştırmaları ele alarak, özellikle görüntü verisi odaklı çalışmalar üzerinde durmaktadır. Var olan çalışmalar, kullanılan veri türleri, model yapılandırmaları ve performans kriterlerine göre gruplandırılmakta ve her bir yöntemin güçlü ve zayıf yönleri detaylı bir biçimde değerlendirilmektedir. Bununla birlikte, literatürde genellikle sadece MRI veya PET gibi görüntü verilerinin kullanıldığı gözlemlenmiş olup, tablosal veya klinik veri entegrasyonunu temel alan çok modlu yaklaşımların sayısının sınırlı olduğu dikkati çekmektedir.

Araştırmaların temel bulguları, veri kaynakları ve uygulama alanları özetlendikten sonra, çok modlu veri kullanımının sunduğu özgün katkılar ve olası yararları üzerinde yoğunlaşmaktadır. Yapılan literatür incelemesi, bu çalışmanın hem görüntü verisi hem de tablosal veriyi bir araya getiren çok modlu yaklaşımıyla literatüre yenilikçi bir katkı sunacağını göstermekte ve diğer çalışmalara göre erken tanıda doğruluk ile genellenebilirlik açısından dikkate değer yenilikler sağlama potansiyelini ortaya koymaktadır.

A. Derin Öğrenme

Bu bölümde, Alzheimer hastalığının erken teşhisinde derin öğrenme tekniklerinin kullanımına odaklanılmaktadır. Özellikle görüntü verilerinden anlamlı özelliklerin elde edilmesinde, yapay sinir ağları ve konvolüsyonel sinir ağları gibi derin öğrenme yapılarının önemine detaylı bir şekilde yer verilmektedir. Literatürde bulunan çalışmalar, farklı katman yapılarına, aktivasyon fonksiyonlarına ve optimizasyon yöntemlerine göre gruplandırılmakta, her metodun hastalığın teşhisindeki etkinliği ve kısıtlamaları irdelenmektedir. Ayrıca, derin öğrenme yaklaşımlarının tek bir veri seti ile sınırlı kalmadığı, tablosal ve klinik verilerle entegrasyon olasılıklarının da ele alındığı vurgulanmakta, bu durumun erken teşhis açısından sağladığı güvenilirlik ve doğruluk avantajları ortaya konulmaktadır. Yapılan bu inceleme, derin öğrenme tekniklerinin Alzheimer'ın erken tanısında gösterdiği etkili performansın yanı sıra, modelin şeffaflığı, yorumlanabilirliği ve veri çeşitliliğinin önemini de gözler önüne sermektedir.

Doaa ve ekip arkadaşları (2023), Alzheimer'ın erken teşhisinde MRI verilerini kullanan bir derin öğrenme sistemi önermektedir. Çalışmalarında veri işleme, veri artırma ve sınıflandırma aşamalarını içeren bir yöntem benimsenmiş; sıfırdan geliştirilmiş CNN ve transfer öğrenmesine dayalı VGG16 modelleri test

edilmiştir. CNN modeli, üç adet evrişim katmanı ve max-pooling ile oluşturulmuş, aşırı uyumun önüne geçmek için dropout kullanılmış ve ikili sınıflandırma sigmoid aktivasyon fonksiyonu ile gerçekleştirilmiştir. VGG16 modelinde ise ilk katmanlar sabit tutulmuş, üst katmanlar yeniden eğitilmiş ve farklı optimizasyon yöntemleri denenmiştir; en iyi sonuç %97,44 doğrulukla Adam optimizasyonu ile elde edilmiştir. Sıfırdan tasarlanan CNN'nin performansı %99,95–%99,99 doğruluk aralığında olmuştur. Yazarlar, veri setinin kısıtlılığı ve ikili sınıflandırma sınırlarının gelecekte ele alınması gereken engeller olduğunu ifade etmektedir [1].

Janani ve ekibi (2021), Alzheimer aşamalarını belirlemek üzere görüntü, genetik (SNP) ve klinik (EHR) verileri bir araya getiren çok modlu derin öğrenme modelleri geliştirmiştir. Görüntü verisi için 3B CNN kullanılmış, klinik ve genetik veriler için stacked denoising auto-encoder ile ara özellikler çıkarılmış ve sınıflandırma aşamasında birleştirilmiştir. Model, Alzheimer, MCI ve sağlıklı kontrol (CN) kategorilerini ayırt etmek için eğitilmiş ve geleneksel makine öğrenmesi yöntemlerine göre daha yüksek doğruluk, hassasiyet, geri çağırma ve F1 puanları elde edilmiştir. Ayrıca, maskeleme ve kümeleme temelli yöntemlerle hipokampus, amigdala ve RAVLT testi gibi biyolojik belirteçlerin model kararına etkisi incelenmiştir [2].

Seung ve ekibi (2024) tarafından ortaya konulan araştırmada, Alzheimer hastalığının erken teşhisi için MRI ve PET görüntülerini birleştiren orta düzeyde (middle-fusion) bir çok modlu derin öğrenme modeli önerilmektedir. Bu model, aktivasyon fonksiyonu kullanılmadan depthwise separable convolution (DS-Conv) blokları ile özellikler çıkarmakta, daha sonra mix skip connection convolution (MSC-Conv) ve sharing weight convolution (SW-Conv) blokları aracılığıyla farklı modlar arasındaki karmaşık ilişkileri öğrenmektedir. Araştırma, ADNI1, ADNI2 ve ADNI3 veri setlerinde T1-w MRI, FDG PET, AB PET ve tau PET görüntülerini kullanarak modeli değerlendirmiş; Alzheimer ile kontrol grubu arasındaki sınıflandırmada dengeli doğruluk oranı 1.00 (yani %100) olarak elde edilirken, hafif kognitif bozukluk (MCI) ile kontrol grubu karşılaştırmasında ise 0.76 değerine ulaşmıştır. Bunun yanında çalışmada, hipokampus, orta temporal ve inferior temporal bölgeleri içeren yeni bir ROI çıkarım tekniği de önerilmiştir [3].

Suriya ve ekibi (2021), Alzheimer ve demans evrelerinin erken teşhisi için dört kategorili bir derin öğrenme modeli tasarlamıştır. MRI görüntülerinden elde edilen bilgilerle oluşturulan CNN yapısı, DEMNET blokları, dropout ve yoğun katmanlar içermektedir. Veri setindeki sınıf dengesizliği SMOTE yöntemi ile düzeltilmiş,

Yazar(lar) (Yıl)	Örnek Boyutu	Dil	Platform	Yöntem	Sonuçlar	Eksiklikler	Gelecek Çalışmalar
Doaa ve ekb. (2023) [1].	6400	İngilizce	MRI görüntüleri	Derin öğrenmeye dayalı CNN ve transfer öğrenme çerçevesi	Yüksek başarı; az etiketli veri ile Alzheimer analizinde yetkinlik gösterilmiştir	Genel geçerlilik farklı veri kümelerinde test edilmemiş	Farklı veri kümeleriyle test edilmesi; çok merkezli çalışmalar yapılması; farklı görüntüleme modaliteleriyle entegrasyon önerilmiş.
Janani ve ekb. (2021) [2].	Toplam 2,004 hasta	İngilizce	Klinik-veri, SNP genetik veri, MRI görüntü verisi	Multimodal derin öğrenme: görüntü için 3D CNN; genetik + klinik için stacked denoising auto-encoder; veri modüllerinin birleşimiyle sınıflama	Derin modeller geleneksel modellerden iyi performans göstermiş. Multimodal veri birleşimi tek modaliteye göre doğruluk, precision, recall, F1-ortalama açısından üstün bulunmuş.	Görüntü modülünde sadece CN ve AD sınıflandırması yapılmış; genetik veri için MCI/AD ayrımı yapılmamış; bazı modüller eksik olabilen verilerle çalışılmış.	Daha geniş ve farklı modüllerden oluşan veri kümeleriyle çalışmak; MCI aşamasını genetik ve görüntü açısından daha iyi ele almak; eksik veri modüllerinde modelin performansını artırmak.
Seung ve ekb. (2024) [3].	-	İngilizce	Alzheimer's Disease Neuroimaging Initiative (ADNI) serisi	Çok-modlu 3D derin öğrenme modeli: orta füzyon yapısı ile görüntü/modal verileri birleştiren 3D CNN tabanlı bir yaklaşım	Çok-modlu verilerin tek modaliteye göre daha iyi performans gösterdiği bulgusu belirtilmiş	Örneklem sayısı, modlar arası eksik veri durumu ve genellenebilirlik gibi belirsizlikler mevcut	Daha geniş veri setleriyle, farklı merkezlerden veriyile ve eksik modlu durumlarla modelin test edilmesi
Suriya ve ekb. (2021) [4].	6400	İngilizce	MR görüntüleri	Derin öğrenme tabanlı CNN çerçevesi ("DEMNET") ile: Çok-evreli sınıflandırma yapılandırılmış.	Klasik: %95.23 doğruluk; AUC 97%; Cohen's Kappa ≈ 0.93	Örneklem büyüklüğü sınıflar bazında tam sayılarla detaylandırılmamış; sınıflar arası dengesizlik problemi belirtilmiş.	Sınıflar arası dengesizlik probleminin giderilmesi; farklı veri kümeleriyle genellenebilirliğin test edilmesi.
Aristidis ve ekb. (2023) [5].	-	İngilizce	Non-invaziv biyobelirteç veri platformları	Yapay Zekâ ve Derin Öğrenme tekniklerinin, non-invazif biyobelirteç verilerini analiz etmek için entegrasyon	Non-invazif yaklaşımlarla elde edilen büyük ve heterojen verilerin, AI ve DL ile işlenmesinin Alzheimer erken teşhisi için yüksek potansiyel sunduğu belirtilmiştir.	Makale, birincil deneysel çalışma değil; çeşitli teknolojiler ve yöntemler arasında karşılaştırma yapılmasına rağmen, pratik uygulama ve klinik yaygınlık açısından sınırlamaları olduğu ifade edilmiştir.	Farklı non-invazif veri kaynaklarının daha geniş ölçekle toplanması; AI/DL modellerinin klinik uygulamaya entegrasyonu; açıklanabilir yapay zekâ (xAI) tekniklerinin daha fazla kullanımı önerilmektedir.
Doaa ve ekb. (2022) [6].	Derleme çalışması	İngilizce	Görüntüleme verisi ve derin öğrenme çalışmaları üzerindeki literatür, klinik/araştırma ortamı	Derin öğrenme yaklaşımlarının (CNN, DNN, AE, DBN vb) Alzheimer's Hastalığı erken tespiti için literatürdeki kullanımı; görüntü işleme ön işlemleri (pre-processing), sınıflandırma yöntemleri, yazın taraması.	Literatürdeki birçok çalışma yüksek performans göstermiştir (örneğin üç sınıflı sınıflandırmalarda %95'in üzerinde doğruluklar bildirilmektedir)	Büyük veri setlerine erişim kısıtlılığı, Görüntü modları arasında (MRI, PET, fMRI) heterojenlik, çoklu modalite entegrasyonunun zorluğu, Derlemede yalnızca en yaygın DL yöntemlerine odaklanılmış, tüm yöntembilimsel çeşitlilik kapsamlı olmayabilir	Farklı veri modalitelerinin birlikte kullanımı önerilmiştir, Derin öğrenme modellerinin transfer öğrenme, otomatik özellik çıkarımı, daha hafif ağ mimarileriyle geliştirilmesi önerilmiştir

TABLE I: Derin Öğrenme

eğitim doğruluk oranı %99 ve doğrulama doğruluk oranı ise %94 olarak belirlenmiştir. Test aşamasında %95,23 doğruluk oranı ve yaklaşık %97 AUC değeri elde edilmiştir. Bu çalışma, yalnızca görüntü verisi kullanmasına rağmen çoklu sınıflandırmada yüksek bir başarı sergilemiştir [4].

Aristidis ve ekibi (2023), Alzheimer hastalığının erken teşhisinde invaziv yöntemlerin yerine non-invaziv biyobelirteçlerin kullanımı ile bu veriler üzerinde yapay zeka ile derin öğrenme uygulamalarını incelemiştir. Araştırmada, biyobelirteç geliştirme platformları, kan bileşenleri, giyilebilir sensörlerden ve görüntüleme sistemlerinden elde edilen verilerle değerlendirilmiştir; bu verilerin büyük ölçekte, çeşitlilikte ve analiz gerektirmesi nedeniyle yapay zeka ile derin öğrenmenin entegrasyonunun önemine dikkat çekilmiştir. Spesifik bir model ya da belirli başarı oranları verilmemekle birlikte, çalışmada non-invaziv yöntemlerin hasta konforu, maliyet ve risk açısından faydaları olduğu belirtilmiştir; buna ek olarak, veri yönetimi, yorum yapılabilirlik ve modelin şeffaflığı gibi konuların hala önemli zorluklar sunduğu ifade edilmiştir. Araştırma, non-invaziv veri kaynakları ve yapay zeka tekniklerinin Alzheimer'ın erken tanısında önemli bir değişiklik sağlayabileceğini göstermektedir [5].

Doaa ve ekibi (2022), Alzheimer hastalığının zamanında teşhisi ve sınıflandırılması için güncel derin öğrenme yöntemlerini sistematik bir şekilde incelemiştir. Araştırmada, yapısal ve fonksiyonel MRI ile PET gibi görüntüleme tekniklerinin yanı sıra, veri ön işleme aşamaları (gürültünün giderilmesi, beyin bölgesinin çıkarılması, yoğunluk düzeltme) ve sınıflandırma yaklaşımları (CNN, AE, DNN vb.) kapsamlı bir şekilde ele alınmıştır. İncelenen 159 çalışmadan 110'unun Alzheimer'ın erken tanısına odaklandığı ve çoklu modlara sahip verilerin, tek modlara kıyasla daha iyi sınıflandırma sonuçları sağladığı belirtilmiştir. Bunun yanı sıra, veri seti eksiklikleri, modalite entegrasyonu, sınıf dengesizlikleri ve modelin anlaşılabilirliği gibi mevcut sorunlar üzerinde durulmuştur. Bu sistematik değerlendirme, Alzheimer'ın erken teşhisi için derin öğrenmeye dayalı sistemlerin geliştirilmesinde literatürdeki eksiklikleri ve araştırma alanlarını belirleyerek, gelecekte yapılacak çok modlu araştırmalara sağlam bir zemin hazırlamaktadır [6].

B. Makine Öğrenmesi

Bu bölümde, Alzheimer hastalığını erken evrede tespit etmek için makine öğrenimi temelli yöntemler ele alınmaktadır. Araştırmalar, genelde MRI, PET ve klinik bilgiler gibi tekil veya çoklu veri kaynaklarını kullanarak, sınıflandırma görevlerinde destek vektör makineleri (SVM), karar ağaçları, rastgele orman (Random Forest), k-en yakın komşu (k-NN) ve lojistik regresyon gibi algoritmaları değerlendirmektedir. Veri ön işleme, boyut azaltma ve özellik seçimi teknikleri (PCA, RFE) ile modelin etkinliğinin artırılması hedeflenirken, sınıf dengesizlikleri SMOTE gibi yöntemlerle düzeltilmektedir. Literatürdeki bulgular, makine öğrenimi modellerinin tek modal verilerle yüksek doğruluk elde edebileceğini, fakat çok modal verilerle birlikte kullanıldığında erken teşhis için daha güvenilir ve genellenebilir sonuçlar sunduğunu göstermektedir. Ayrıca, modelin açıklanabilirliği ve veri çeşitliliği, makine öğrenimi yöntemlerinin uygulanabilirliği açısından önemli zorluklar olarak dikkat çekmektedir.

Vasco ve ekibi (2022), Alzheimer Hastalığı (AD) ve Hafif Kognitif Bozukluk (MCI) tanısı için yapılandırılmış MRI verilerini kullanan çoklu teşhis ve genişletilebilir bir makine öğrenimi yöntemi geliştirmiştir. Araştırmada ADNI (n=570) ve OASIS (n=531) veri

setlerinden morfometrik özellikler elde edilmiştir. Bu özellikler, 1-SVM, karar ağaçları, rastgele orman, aşırı rastgele ağaçlar, LDA, lojistik regresyon ve SGD ile lojistik regresyon algoritmaları kullanılarak 5 katmanlı çapraz doğrulama yöntemiyle eğitilmiştir. Özelliklerin seçimi mutual bilgi, ANOVA F değerleri ve chi-kare istatistikleri kullanılarak gerçekleştirilmiş; evrimsel algoritmalar aracılığıyla en iyi özellik yüzdeleri ve hiperparametreler tespit edilmiştir. İkili sınıflandırmada "Kontrol vs AD" görevinde, birleşik ADNI + OASIS veri setinde dengeli doğruluk %90,6 ve MCC değeri 0,811 olarak belirlenmiştir. Çok sınıflı sınıflandırmada ise "Kontrol vs MCI vs AD" görevinde yalnızca ADNI veri seti kullanılarak dengeli doğruluk %62,1 ve MCC değeri 0,438 elde edilmiştir. Özelliklerin önem derecelerine göre, hipokampus %25-45, temporal alanlar %13, cingulate ve frontal bölgeler ise %8-13 arası katkıda bulunmuştur. Ayrıca, grafik teorisi metriklerinin sınıflama performansını artırmadığı ve MCI hastalarının yalnızca bir veri setinde mevcut olduğu ile bilişsel test skorlarının model içine alınmadığı gibi bazı sınırlamalar da vurgulanmıştır [7].

Khandaker ve ekibi (2023) tarafından gerçekleştirilen bu araştırmada, Alzheimer hastalığının erken tespiti için bir makine öğrenimi yöntemi önerilmiş. Araştırmada OASIS veri seti kullanılmış; eksik olan veriler ortalama değerlerle doldurulmuş, sonrasında Select KBest tekniğiyle özelliklerin seçim işlemi yapılmış ve veriler Standard Scaler ile ölçeklendirilmiş. Kullanılan sınıflandırma algoritmaları arasında Gaussian Naive Bayes, Karar Ağacı, Rastgele Orman, XGBoost, Gradient Boost ve Oylama Sınıflandırıcısı bulunmaktadır. En yüksek doğruluk seviyesi, oylama sınıflandırıcısı ile %96 olarak elde edilmiştir. Araştırmanın kısıtlamaları arasında, veri setinin tek bir kaynaktan oluşması ve çeşitli veri protokollerinin azlığı yer almaktadır [8].

Chun-Hung ve ekibinin (2021) gerçekleştirdiği çalışmada, Alzheimer hastalığının erken teşhisinde makine öğrenimi ve yeni biyobelirteçlerin rolü araştırılmıştır. Yayında, mevcut teşhis yöntemlerinin (örn. omurilik sıvısındaki AB ve p-tau seviyeleri) invaziv ve pahalı olduğu belirtilmiş, bu sebeple daha az invaziv biyobelirteçlerin geliştirilmesi gerektiği ifade edilmiştir. Araştırmada, geleneksel AB/tau göstergelerine ilave olarak nöro-yaralanma belirteci olan neurofilament light (NFL), sinaptik işlev bozukluğu indikatörleri olarak neurogranin, BACE1, SNAP-25, GAP-43 ve synaptophysin ve nöro-enflamasyon belirteçleri olarak sTREM2 ve YKL-40'nın değerlendirildiği belirtilmiştir. Makine öğrenimi yöntemleri arasında destek vektör makinesi (SVM), lojistik regresyon, rastgele orman ve naïve Bayes'in kullanıldığı ve bu yöntemlerle hastalar ile sağlıklı bireyler arasındaki farkların daha iyi bir biçimde anlaşılacağı ifade edilmiştir. Makalede, bu yöntemlerin duyarlılık ve özgüllük açısından fayda sağlama potansiyeli bulunduğu, fakat veri setlerinin boyutu, biyobelirteçlerin klinik uygulanabilirliği ve makine öğrenimi modellerinin standart hale getirilmesi gibi konuların hala çözülmesi gereken zorluklar arasında yer aldığı vurgulanmıştır [9].

Daniele ve ekibi (2022) Alzheimer hastalığının erken tanısını koymak amacıyla EEG sinyal analizi ve denetimli makine öğrenimi yöntemlerini bir araya getirmiştir. Araştırmada toplamda 105 kişi yer almış, EEG verileri 256 Hz hızında normalize edilerek 1-30 Hz bant geçiren filtre kullanılarak işlenmiş ve müdahale eden unsurlardan arındırılmıştır. Zaman-frekans analizleri ve çift dijital FIR filtreleriyle özellikler çıkarılmış, düşük ve yüksek frekans bantlarındaki güç yoğunlukları sınıflandırma sürecinde kullanılmıştır. Karar ağaçları, destek vektör makineleri ve k-en yakın komşu algoritmaları uygulanmış, veriler eğitim ve test setleri olarak %70 ve %30 oranında ayrılmıştır. İkili

Yazar(lar) (Yıl)	Örnek Boyutu	Dil	Platform	Yöntem	Sonuçlar	Eksiklikler	Gelecek Çalışmalar
Vasco ve ekb. (2022) [7].	ADNI veri kümesi: 570 katılımcı, OASIS veri kümesi: 531 katılımcı	İngilizce	Yapısal MRI verileri	Makine öğrenimi algoritmaları + oylama yöntemi ile çok sınıflı ve iki sınıflı sınıflandırma. Ayrıca farklı MRI protokolleri ve grafik teorisi özellikleri değerlendirilmiş.	HC vs AD sınıflandırmasında: Dengelenmiş doğruluk %90,6 ve Matthew's korelasyon katsayısı 0.811, "HC vs MCI vs AD" sınıflandırmasında: BAC 62,1%, MCC 0.438	MCI kategorisi az sayıda, klinik gerçek-hayat karar süreçlerine tam uyumlu değil ve kognitif test skorları ya da diğer biyobelirteçler dahil edilmemiş.	Daha geniş, çok merkezli ve uzun süreli (longitudinal) veri setleriyle test edilmesi önerilmiş ve klinik kullanım için gerçek-zamanlı çalışmalara (prospektif çalışmalar) geçilmesi gerektiği vurgulanmış.
Khandaker ve ekb. (2023) [8].	OASIS veri setinden 150 birey (64 demented, 72 non demented)	İngilizce	Yapısal MRI tabanlı "Open Access Series of Imaging Studies (OASIS)" veri seti	Öznitelik seçimi (SelectKBest), standartlaştırma (StandardScaler), sınıflandırıcılar: GaussianNB, Decision Tree, Random Forest, XGBoost, GradientBoost, Voting Ensemble	En iyi doğruluk 96% (Voting-Classifer veya Random Forest kombinasyonu)	Örneklem sayısı sınırlı; yalnızca tek bir veri kaynağı; çoklu modalite (örneğin PET, biyobelirteç) dahil edilmemiş.	Daha büyük ve çok merkeze yayılmış veri setleriyle doğrulama; multimodal verilerin entegrasyonu; klinik uygulanabilirlik için prospektif çalışmalar yapılması.
Chun-Hung ve ekb. (2021) [9].	Derleme / meta analiz çalışması	İngilizce	Klinik biyobelirteçler, görüntüleme ve diğer biyobelirteç kaynakları	Makine öğrenimi ve derin öğrenme algoritmaları ile yeni biyobelirteçlerin kombinasyonu	Makine öğrenimi + yeni biyobelirteçlerin birlikte kullanımı, Alzheimer tanısında duyarlılık ve özgüllüğü artırabileceğini göstermiştir.	Tek bir veri kaynağına odaklanılmış çalışmalar sınırlı; yöntemlerin klinik uygulamaya yaygınlığı sınırlıdır; bazı biyobelirteçlerin standardizasyonu yok.	Daha büyük ve çok merkezli veri setleriyle test edilmesi; kan-temelli biyobelirteçlerin daha yaygın kullanımı, derin öğrenme modellerinin klinik karar destek sistemlerine entegrasyonu önerilmiştir.
Daniele ve ekb. (2022) [10].	105 kayıt (48 AD, 37 MCI, 20 HC)	İngilizce	EEG kayıtları	FIR çift zaman-alan filtresiyle yüksek ve düşük frekans bant filtreleme, güç yoğunluğu çıkarımı: yüksek-düşük bant farkının mutlak karesi ve sınıflandırılması	HC vs AD: 97% doğruluk, HC vs MCI: 95% doğruluk, MCI vs AD: 83% doğruluk, HC vs MCI vs AD: 75% doğruluk	Örneklem sayısı görece sınırlı; yalnızca istirahat halindeki EEG, tek merkezli veri; üç sınıflı sınıflandırmada doğruluk önemli derecede düşüyor.	Veri setinin ölçeğinin artırılması; farklı EEG protokollerinin (örnek: bilişsel görev altında) kullanımı; gömülü cihazlarda gerçek zamanlı uygulama potansiyeli.
Gargi ve ekb. (2023) [11]	3.692 MRI görüntüsü	İngilizce	ADNI veri kümesi	Görüntü ön işleme: selective clipping, gri tonlama dönüşümü, histogram eşitleme ve Sınıflandırma algoritmaları: Random Forest, XGBoost, CNN	Doğruluk: 97,57% Duyarlılık (Sensitivity): 97,60%	Tek veri seti (sadece ADNI MRI), yalnızca iki sınıflı ayırım (AD vs Normal), modelin genelleme yeteneği belirsiz	Veri çoğaltma (GAN kullanılmış), farklı veri setleriyle çapraz validasyon, çok sınıflı tanı (örneğin MCI dahil) önerilmiştir
Nitsa ve ekb. (2021) [12]	750 katılımcı	İngilizce	Alzheimer's Disease Neuroimaging Initiative (ADNI) MRI görüntüleri	Beyin asimetrisini saptama (sol-sağ yarıküre), istatistiksel özellik çıkarımı (MSE, varyans, vs.), SVM ve CNN tabanlı denetimli öğrenme	NC vs EMCI: SVM %92,5 doğruluk; NC vs AD: %93,0 doğruluk; CNN için NC vs AD: %90,5 doğruluk	Katılımcı sayısı sınırlı, yalnızca MRI görüntüleri kullanılmıştır. Kişisel "hâkim el" (handedness) gibi değişkenler incelenmemiştir.	Farklı el hâkimiyeti gruplarının etkisi araştırılmalı; derin öğrenme mimari optimizasyonu yapılmalı; uzunlamasına (longitudinal) çalışma önerilmektedir.

TABLE II: Makine öğrenmesi

sınıflandırmalarda doğruluk oranları HC ile AD arasında %97, HC ile MCI arasında %95 ve MCI ile AD arasında %83 olarak belirlenmiştir; üç sınıflı sınıflandırmada doğruluk %75 olarak bildirilmiştir. Araştırmacılar, filtre kesme frekanslarının doğru bir şekilde seçilmesinin sınıflandırma başarısını önemli ölçüde etkilediğini ve 7-16 Hz aralığının ayırt edici gücünü artırdığını vurgulamaktadır. Çalışmanın kısıtlamaları arasında küçük veri seti, gruplar arasında dengesizlik ve farklı EEG kayıt protokollerinin kullanılmaması yer almaktadır [10].

Gargi ve ekibi (2023), Alzheimer hastalığı (AD) erken teşhisi için makine öğrenimi tabanlı bir yöntem önerdi. Araştırmada, Alzheimer's Disease Neuroimaging Initiative (ADNI) veri kümesinden alınan dört boyutlu MRI görüntüleri iki boyuta dönüştürüldü ve görüntüler üzerinde seçici kırpma, gri tonlama dönüşümü ve histogram eşitleme gibi ön işleme işlemleri gerçekleştirildi. Daha sonra sınıflandırma için üç farklı algoritma (Random Forest, XGBoost ve Convolutional Neural Network (CNN)) kullanıldı. Ön işleme sonrası elde edilen bulgulara göre, başarı oranı %97,57, duyarlılık ise %97,60 olarak belirtilmiştir. Çalışmada belirgin kısıtlamalar arasında yalnızca tek bir veri kaynağının kullanılması ve farklı protokollere ait veri çeşitliliğinin sağlanamaması yer almaktadır [11].

Nitsa ve ekibi (2021), erken demans ve Alzheimer teşhisi için beyin asimetrisini belirleyen bir makine öğrenimi yöntemi geliştirmiştir. ADNI veri setinden elde edilen T1-ağırlıklı MRI görüntüleri, 256×256×3 boyutuna standart hale getirilip segmentasyon ve yansıtma işlemlerine tabi tutulmuş, ardından 10 istatistiksel özellik (ortalama, varyans, RMS, entropi, çarpıklık, basıklık vb.) elde edilmiştir. Sınıflandırma, Naïve Bayes, LDA, SVM, k-NN ve AlexNet tabanlı CNN algoritmaları ile gerçekleştirilmiş, NC ile EMCI ve NC ile AD arasındaki görevlerde SVM ile doğruluğun %92,5 - %93,0, CNN ile ise %75,0 - %90,5 arasında olduğu belirlenmiştir. Sınırlılıklar arasında el hâkimiyetinin göz önünde bulundurulmaması, uzun dönem takip analizlerinin yapılmaması ve farklı MRI protokollerinin kullanılmaması bulunmaktadır [12].

C. Hibrit Yaklaşımlar

Bu bölümde, Alzheimer hastalığının erken teşhisi üzerine hibrit modellere dayanan araştırmalar incelenmektedir. Çalışmalar, genellikle MRI, PET, demografik veriler ve psikometrik testler gibi birden fazla veri kaynağını kullanarak, derin öğrenme ile makine öğrenimi yöntemlerini bir araya getiren hibrit modeller oluşturmaktadır. Bu yöntemler, özellik çıkarımı, veri artırımı, segmentasyon ve sınıflandırmayı tek bir çatı altında birleştirerek yüksek doğruluk ve genellenebilirlik sağlamayı amaçlamaktadır. Mevcut literatürde, hibrit modellerin, MCI ile AD ayırımında özellikle tek modal yöntemlere göre daha etkili olduğu, ancak veri sınırlamaları, tek merkezli araştırmalar ve sınıf dengesizlikleri gibi engeller nedeniyle klinik uygulanabilirliğin hâlâ kısıtlı olduğu belirtilmektedir.

Akhilesh ve ekibi (2023) tarafından gerçekleştirilen sistematik incelemede, MRI ve PET verilerini kullanan 47 çalışma değerlendirilmiş; geleneksel makine öğrenimi teknikleri ile %85 doğruluk, derin öğrenme ile %96-98 doğruluk elde edilmiş ve çok modal yaklaşımların performansı tek modaliteye göre daha yüksek bulunmuştur. Ancak çoğu çalışma ADNI veri setine bağımlı olup, gerçek klinik uygulamalara dönüşüm sınırlıdır [13].

Badia ve ekibi (2021), OASIS ve MRI veri setlerini kullanarak SMOTE ve t-SNE ile ön işleme yaptıktan sonra CNN ve hibrit AlexNet + SVM modelleriyle %94 - %99 civarında doğruluk

sağlamıştır. Çalışmada veri tek merkezli ve sınırlı olduğundan multimodal veri entegrasyonu ve klinik uygulama gerekliliği vurgulanmıştır [14].

Anuradha ve ekibi (2023), Kaggle veri setinde görüntü ön işleme ve segmentasyon tekniklerini kullanarak hibrit CNN-SoftMax modeli ile %99 doğruluk elde etmiş, fakat bazı sınıflarda örnek sayısı çok az olup veri tek kaynaktan gelmiştir. Daha geniş ve çok merkezli veri setleriyle doğrulamanın önemi belirtilmiştir [15].

Ibrahim ve ekibi (2022), dört sınıflı MRI verilerini kullanarak yerel özellik çıkarımı ve ön-eğitilmiş derin modellerle hibrit bir yaklaşım geliştirmiş, %99.8 doğruluk ve %100 özgüllük elde etmiştir. Sınırlılık olarak örneklem detaylarının yeterince verilmemesi ve tek veri kaynağı kullanımı öne çıkmaktadır [16].

Balaji ve ekibi (2023), MRI ve PET verilerini CNN + LSTM mimarisi ile analiz etmiş ve bilişsel olarak normal ile EMCI sınıflandırmasında %98.5 doğruluk sağlamıştır. Veri setinin sınırlılığı ve bazı sınıfların eksikliği, çalışmanın kısıtları arasında yer almaktadır [17].

Afreen ve ekibi (2022), demografi ve psikometrik test verilerini kullanarak üç kademeli hibrit makine öğrenimi algoritması geliştirmiştir; doğruluk oranları %89.6 - %95.1 arasında değişmiştir. Çalışmada görüntüleme veya biyobelirteç verileri dahil edilmediği için multimodal analizlere ihtiyaç duyulduğu vurgulanmıştır [18].

Bu bulgular, hibrit yaklaşımların çok modal veri kullanımıyla erken Alzheimer teşhisinde yüksek doğruluk sağlayabildiğini, ancak veri çeşitliliği, çok merkezli doğrulama ve klinik geçerlilik açısından hâlen önemli gelişme alanları bulunduğunu göstermektedir.

D. Yöntemlerin Karşılaştırılmalı Değerlendirmesi

Alzheimer hastalığının ön aşamalarını belirlemek için yararlanılan yapay zeka destekli teknikler, esasen üç ana grupta ele alınabilir: derin öğrenme, geleneksel makine öğrenimi ve karma yöntemler. Tablo I, II ve III'de yer alan araştırmalar, çeşitli veri türleri, model dizaynları, başarı ölçütleri ve sınırlılıklar açısından önemli kıyaslama imkanları sağlamaktadır.

Derin öğrenmeye dayalı teknikler, özellikle MRI ve PET gibi görsel verilerden otomatik olarak anlamlı özellikler elde etme yeteneği ile dikkat çekmektedir. Örneğin, Doaa ve arkadaşlarının (2023) araştırmasında, sıfırdan oluşturulan CNN ve transfer öğrenmeye dayanan VGG16 modelleri, ikili sınıflandırma görevlerinde %99.95 ile %99.99 arasında bir doğruluk sağladığı ve bunun az sayıda etiketli veri ile yüksek bir performans elde edilebileceğini gösterdiği belirtilmiştir [1]. Benzer şekilde, Janani ve ekibinin (2021) çoklu modal derin öğrenme yöntemi, görüntü, genetik ve klinik bilgileri bir araya getirerek üç sınıflı sınıflandırmada geleneksel yöntemlere göre belirgin bir avantaj sunmuş ve özellikle çok modal verilerin entegrasyonunun erken tanı doğruluğunu artırma potansiyelini göstermiştir [2]. Derin öğrenme yöntemlerinin sağladığı faydalar arasında yüksek doğruluk, otomatik özellik çıkarımı ve görsel tabanlı özneliliklerin etkin temsil edilmesi bulunurken; sınırlamaları arasında ise veri setlerinin genellikle tek merkezli ve kısıtlı olması, modelin açıklanabilirliğinin az olması ve farklı modalitelerdeki eksik verilerin yönetiminde karşılaşılan zorluklar yer almaktadır.

Makine öğrenmesi temelli yöntemler genellikle kısıtlı veri setlerinde daha hızlı ve anlaşılır çözümler sunmaktadır. Vasco ve takımının (2022) geliştirdiği çoklu sınıflandırma ve oylamaya dayalı yöntem, HC ile AD arasındaki sınıflandırmada %90,6 dengelenmiş doğruluk sağlamış, ancak çoklu sınıf görevlerinde

Yazar(lar) (Yıl)	Örnek Boyutu	Dil	Platform	Yöntem	Sonuçlar	Eksiklikler	Gelecek Çalışmalar
Akhilesh ve ekb. (2023) [13].	İncelenen yaklaşık 47 çalışma seçilmiştir.	İngilizce	MRI ve PET görüntüleme verileri, ADNI ve OASIS veri setleri.	Makine öğrenimi ve derin öğrenme tekniklerinin sistematik analizi. Öznetelik çıkarımı, çok-modalite kullanımı değerlendirilmiştir.	Geleneksel ML teknikleri ile %85 kadar doğruluk; Derin öğrenme ile %96-98 doğruluğa kadar ulaşılmıştır. Çok-modalite yaklaşımlar tek modaliteye göre daha iyi performans göstermiştir.	Çoğu çalışma MRI/PET görüntüleme ile sınırlı; veriler büyük ölçüde ADNI'ye bağımlı; gerçek klinik uygulamaya dönüştürülmüş az sayıda çalışma var.	Multimodalite çalışmalarının artırılması, açıklanabilir Yapay Zekâ (XAI) tekniklerinin entegre edilmesi ve farklı veri kaynakları ve veri çeşitliliği ile genelleştirilebilir modeller geliştirilmesi önerilmektedir.
Badiea ve ekb. (2021) [14].	OASIS veri seti: 373 kayıt ve MRI veri seti: 6,400 görüntü	İngilizce	OASIS tıbbi kayıt veri seti ve MRI görüntü veri seti	OASIS için: SMOTE ile veri dengeleme, t-SNE ile boyut indirgeme, sınıflandırma ve MRI için: CNN + veri artırımı, hibrit modeller	OASIS kısmı: Random Forest ile doğruluk %94; precision %93; recall %98; F1 %96 ve MRI kısmı: Hibrit AlexNet + SVM modeli ile doğruluk %94.8; duyarlılık 93%; özgüllük 97.75%; AUC 99.70%.	Veri tek merkezli ve sınırlı; MRI görüntülerindeki sınıflar sayıca az; yalnızca görüntü ve kayıt verileriyle sınırlı, diğer biyobelirteçler veya çok-modalite kombinasyonları dahil edilmemiş.	Daha fazla veri kaynağı ile çok-merkezli test; multimodal veri entegrasyonu; hibrit modellerin klinik uygulamaya geçişi için prospektif çalışmalara yönelim.
Anuradha ve ekb. (2023) [15].	6,400 görüntü	İngilizce	Kaggle "Alzheimer's dataset 4 class of images" veri seti	Ön işleme: HSV ve LAB dönüşümleri, K-means segmentasyon, görüntü boyutlandırma, her sınıf için 3,200 görüntüye çıkarıldı. Hibrit model, paralel olarak özellik çıkarımı, ardından CNN katman ve SoftMax sınıflandırma	SMOTE kullanan modelde 99% doğruluk elde edilmiştir.	Veri sınıflarında dengesizlik özellikle "moderate dementia" sınıfında çok az örneklerle çalışılmıştır (≤ 100 görüntü). Veri yalnızca tek kaynaktan gelmektedir.	Daha büyük ve çok merkezli veri setleriyle doğrulama; multimodal verilerin (örneğin PET, biyobelirteç) entegrasyonu; gerçek klinik uygulamaya yönelik geçiş.
Ibrahim ve ekb. (2022) [16].	-	İngilizce	MRI görüntüleri üzerinden dört sınıflı sınıflandırma.	Yerel özellik çıkarımı, Ön-eğitilmiş derin modeller, Hibrit model ve Hibrit combine	%99.8 doğruluk; precision %99.9; sensitivity %99.75; specificity %100; AUC %99.94.	Örneklem detayları yeterince verilmemiş; yalnızca belirli veri kaynağı kullanılmış olabilir; gerçek klinik validasyon sınırlı.	Daha geniş veri setiyle (çok merkezli) çalışılması; farklı modalite ve biyobelirteçlerin entegrasyonu; klinik uygulamaya geçiş için uzunlamasına çalışmaların yapılması önerilmektedir.
Balaji ve ekb. (2023) [17].	512 MRI görüntüsü ve 112 PET görüntüsü	İngilizce	MRI + PET görüntüleri ve nöropsikolojik test skorları	CNN + LSTM mimarisi; ACO ile görüntü azaltma; MFCM ile segmentasyon; Adam optimizasyonu	"Bilişsel olarak normal" ile "EMCI" sınıflandırmasında doğruluk %98.5'e ulaşılmıştır.	Veri seti yalnızca MRI + PET ile sınırlı; bazı sınıflar veya gerçek-dünya klinik veri kapsamı sınırlı belirtildi.	Farklı modalitelerin (örneğin biyobelirteçler) eklenmesi; daha büyük ve çok merkezli veri setleriyle doğrulama önerilmiştir.
Afreen ve ekb. (2022) [18].	ADNI veri seti	İngilizce	demografi + psikometrik test sonuçları	Üç kademeli bilişsel-hibrit makine öğrenimi algoritması ve Özellik seçimi, demografik + bilişsel test verileri kullanılmıştır.	Tier1: 89.63% doğruluk Tier2: Random Forest ile 93.90% doğruluk Tier3: 95.12% doğruluk	Sınıf dağılımları kapsamlı verilmemiştir, yalnızca psikometrik ve demografik veriler kullanılmıştır. Görüntüleme ya da biyobelirteç veri modları dahil edilmemiştir.	Görüntüleme (MRI/PET) ve biyobelirteç verileriyle multimodal analizlere geçilmesi; farklı merkezlerden veriyle genelleştirilebilirliğin test edilmesi; gerçek klinik çevrede prospektif uygulamaların yapılması önerilmektedir.

TABLE III: Hibrit Yaklaşımlar

düşen performans ve sınırlı klinik veri kullanımı gibi sorunlarla karşılaşmıştır [7]. Benzer şekilde, Khandaker ve ekibi (2023) tek merkezli OASIS veri seti ile %96 doğruluk elde etmiş, fakat multimodal veri eksikliği ve küçük örnek sayısı, modelin genellebilirliğini olumsuz etkilemiştir [8]. Makine öğrenimi yöntemlerinin artıları, modelin açıklığı, anlaşılabilirliği ve veri azlığında çalışma yeteneği iken, eksiklikleri arasında çoklu modalite entegrasyonunun kısıtlılığı, yüksek boyutlu özelliklerde performans kaybı ve sınırlı veri genelleştirilebilirliği yer almaktadır.

Hibrit yöntemler, derin öğrenmenin özellik çıkarma yeteneği ile makine öğrenmesinin anlaşılabilir ve optimize edilebilir yapısını bir araya getirerek, multimodal veri kullanımında önemli faydalar sağlamaktadır. Örneğin, Balaji ve arkadaşları (2023) MRI ile PET görüntülerini nöropsikolojik test sonuçları ile birleştirerek geliştirdikleri CNN + LSTM tabanlı hibrit modelle "Bilişsel olarak normal" ve "EMCI" sınıflandırmalarında %98,5'lik bir doğruluk oranı yakalamış ve böylece multimodal verilerin birleştirilmesinin sınıflandırma doğruluğunu artırdığı gözlemlenmiştir [17]. Afreen ve ekibinin (2022) gerçekleştirdiği çalışmada, demografik ve psikometrik test verilerini üç aşamalı bir hibrit makine öğrenimi algoritması aracılığıyla incelemiş ve Tier3 aşamasında %95,12 doğruluk elde ederek tablosal verilerin multimodal analizlerdeki etkinliğini ortaya koymuştur [18]. Hibrit yöntemlerin en önemli avantajları, çeşitli veri kaynaklarının entegrasyonu ile doğruluk ve genelleştirilebilirliğin artırılmasıdır; öte yandan dezavantajları ise modelin karmaşıklığının artması, hesaplama maliyetlerinin yükselmesi ve veri ön işleme gereken uzmanlık olmasıdır.

Performans açısından yapılan genel inceleme, tek bir modaliteye dayanan derin öğrenme tekniklerinin yüksek doğruluk sunmasına rağmen, veri çeşitliliğinin yetersizliği ve yorumlanabilirlik sorunları ile karşılaştıklarını ortaya koymaktadır. Makine öğrenimi teknikleri, daha küçük ve yapılandırılmış verilerle hızlı sonuçlar sağlarken, çoklu veri entegrasyonu konusunda sınırlı kalmaktadır. Hibrit yaklaşımlar ise, görüntü, tablo ve klinik verileri harmanlayarak Alzheimer'ın erken teşhisinde hem doğruluk hem de güvenilirliği artırmakta; fakat bu durum modelin karmaşıklığını ve hesaplama masraflarını artırmaktadır.

Özetle, çalışmaların karşılaştırılması, çok modlu ve hibrit yaklaşımların Alzheimer'ın erken teşhisinde hem doğruluk hem de genellebilirlik açısından daha üstün bir potansiyel sunduğunu göstermektedir. Bunun yanında, veri setlerinin çoğunlukla tek merkezli, küçük boyutlu ve belirli modalitelerle kısıtlı olması, yöntemlerin klinik uygulamaya geçişinde önemli bir engel teşkil etmektedir. Gelecek çalışmalarda, çeşitli merkezlerden elde edilen çok modaliteli veri setlerinin kullanılması, eksik veri yönetimi, modelin anlaşılabilirliği ve klinik uygulama açısından iyileştirme, Alzheimer'ın erken teşhisinde yapay zeka yöntemlerinin başarısını artıracak kritik faktörler olarak öne çıkmaktadır.

E. Literatürün Genel Değerlendirmesi

Alzheimer hastalığının erken teşhisinde çalışmalarda derin öğrenme, makine öğrenmesi ve hibrit yöntemler dikkat çekmektedir. Derin öğrenme teknikleri, özellikle MRI ve PET görüntülerinden otomatik olarak özellikler çıkararak yüksek doğruluk sağlamaktadır, fakat açıklanabilirlik konusunda bazı kısıtlamalar vardır. Makine öğrenmesi algoritmaları ise klinik veriler ve demografik bilgilerle daha genel sonuçlar elde edebilmekte, ancak yüksek boyutlu görüntülerin analizi konusunda sınırlı kalmaktadır.

Hibrit modeller, farklı veri türlerini bir araya getirerek hem doğruluk hem de genelleme yeteneği açısından fayda

sağlamaktadır. Araştırmalar, çok modal ve hibrit yöntemlerin, tek modal yöntemlere kıyasla özellikle MCI ve AD ayırımında daha etkili olduğunu ortaya koymaktadır. Ancak veri eksiklikleri ve klinik uygulama zorlukları hâlâ önemli engeller olarak devam etmektedir. Bu bağlamda, mevcut araştırmalar Alzheimer hastalığının erken teşhisinde yapay zekanın potansiyelini göstermekte ve çok modlu yaklaşımların önemini vurgulamaktadır.

II. DENEYSEL SONUÇLAR

A. Veri Seti ve Metodolojik Özet

Bu çalışmada, Alzheimer hastalığının dört ayrı aşamasını kapsayan "Alzheimer's Dataset (4 class of Images)" adındaki açık görüntü veri seti kullanılmıştır. Bu veri seti, MildDemented (Hafif Bunama), ModerateDemented (Orta Derecede Bunama), NonDemented (Bunama Olmayan) ve VeryMildDemented (Çok Hafif Bunama) adında dört gruba ayrılmış toplam 6400 beyin Manyetik Rezonans (MR) görüntüsünü içermektedir. Veri setinin orijinal sınıf dağılımı şöyledir: NonDemented (3200 örnek), VeryMildDemented (2240 örnek), MildDemented (896 örnek) ve ModerateDemented (64 örnek). Bu dağılım, özellikle ModerateDemented grubunda belirgin bir dengesizlik olduğunu ortaya koymaktadır.

Projenin çok modlu bir yöntemini deneme amacıyla, orijinal veri setinde yer almayan klinik ve demografik bilgiler, her bir görüntü grubunun bilinen patolojik özelliklerine göre yapay olarak oluşturulmuştur. Bu doğrultuda, her bir görüntüye Yaş, Mini-Mental Durum Testi Skoru, Eğitim Yılı ve Cinsiyet gibi özellikler eklenmiştir.

Modellerin genelleme yeteneğini adil ve güvenilir bir şekilde değerlendirmek için, toplamda 6400 örnek içeren bir veri seti, projenin başlangıcında %20 oranında (%20, test_size=0. 2) ayrılarak Kilitli Nihai Test Seti oluşturulmuştur. Bu test seti, 1280 örnekten oluşur ve modellerin eğitimi ile hiperparametre ayarlama aşamalarında asla kullanılmamış, yalnızca tüm modellerin eğitim süreçleri tamamlandıktan sonra karşılaştırmalı değerlendirme için kullanılmak üzere ayrılmıştır. Kalan %80'lik kısım (5120 örnek) ise modellerin eğitiminde ve 5 katlı çapraz doğrulama aşamalarında değerlendirilmiştir. Bu ayırma yöntemi (stratify parametresi kullanılarak), hem eğitim ve doğrulama setindeki hem de kilitli test setindeki sınıf dağılımının ana veri seti ile uyumlu kalmasını sağlamıştır.

B. Ön İşleme ve Öznitelik Çıkarma

1) *Tablosal Veri Ön İşleme:* Makine öğrenimi modellerinde sentetik tablo verileri kullanılırken, cinsiyet olarak tanımlanan kategorik özellik, One-Hot Encoding tekniği ile sayısal bir forma dönüştürülmektedir. Yaş, MMSE ve eğitim düzeyi gibi tüm sayısal özellikler, değişik ölçeklerin modelin performansı üzerinde olumsuz bir etki yaratmasını engellemek amacıyla StandardScaler aracılığıyla standart hale getirilmektedir. Bu işlemler, algoritmaların özelliklere eşit önem vermesini sağlayarak modelin performansını artırmayı amaçlamaktadır.

2) *Görüntü Verisi Ön İşleme:* Derin öğrenme sistemleri için kullanılan MR görüntüleri, tüm yapılar için aynı girdi boyutunu (128x128) elde etmek adına yeniden boyutlandırılmıştır. Piksel değerleri, belirli bir aralığa normalleştirilmiştir. Ayrıca, modelin genelleme kapasitesini artırmak ve aşırı öğrenmeyi azaltmak için, eğitim setindeki görüntülere rastgele döndürme, kaydırma, yakınlaştırma ve yatay çevirme gibi veri artırma yöntemleri uygulanmıştır. Bu tekniklerin, sınırlı sayıda örneğe sahip olan sınıflarda modelin farklı çeşitleri öğrenerek daha etkili özellikler geliştirmesine katkı sağlaması beklenmektedir.

C. Öznitelik Seçimi

Bu çalışmanın makine öğrenmesi aşamasında, probleme ilişkin temel demografik ve klinik bilgileri yansıtan yalnızca dört sentetik öznitelik (Yaş, MMSE, Eğitim, Cinsiyet) kullanılmıştır. Özellik setinin bu kadar sınırlı ve basit olması nedeniyle, her hangi bir öznitelik seçim yöntemi uygulanmasına ihtiyaç duyulmamaktadır. Bu dört öznenliğin her birinin Alzheimer hastalığının tanısında teorik olarak önemli olabileceği varsayıldığı için, modelin bu temel bilgilere dayanarak ne ölçüde öğrenim gerçekleştirebildiği anlaşılmak istenmektedir. Daha karmaşık ve yüksek boyutlu bir tablo özelliği seti kullanılması durumunda, özellik seçimi yöntemlerinin modelin karmaşıklığını azaltmak ve performansını yükseltmek açısından önemli bir işlevi olabilirdi. Ancak mevcut çerçevede, bu aşama gerçekleştirilmemektedir.

D. Deneyisel Kurulum ve Model Grupları

1) *Donanım/Yazılım Ortamı*: Bu araştırmadaki bütün deneyler, bulut tabanlı bir hesaplama ortamı olan [Google Colaboratory \(Colab\)](#) üzerinde yapılmaktadır. Çalışmanın tekrarlanabilirliğini temin etmek için, kullanılan ana yazılım ve donanım altyapısı aşağıda açıklanmaktadır:

- Donanım: Derin öğrenme modellerinin eğitimini hızlandırmak için Google Colab tarafından sağlanan NVIDIA Tesla T4 Grafik İşlem Birimi (GPU) kullanılmıştır.
- Programlama Dili: Python 3.x
- Veri İşleme ve Analiz: Pandas (Veri çerçeveleri için), NumPy (Sayısal işlemler için)
- Makine Öğrenmesi: Scikit-learn (ML modelleri, metrikler ve K-Fold için)
- Derin Öğrenme: TensorFlow 2.x ve Keras (DL model mimarileri ve eğitimi için)
- Görselleştirme: Matplotlib ve Seaborn (Grafikler ve matrisler için)
- İstatistiksel Test: Mlxtend (mcnemar testi için)

Tüm tasarımlar, kesin sonuçlar almak ve karşılaştırmalarda uyumu sağlamak için `random_state=42` ayarı ile sabit bir başlangıç noktası kullanılarak başlatılmıştır.

2) *Model Grupları ve Karşılaştırma Seti*: Bu araştırmada, Alzheimer hastalığının sınıflandırılması ile ilgili olarak dört ana kategoriye ayrılan 12 modelin başarıları sistematik bir şekilde karşılaştırılmaktadır. Amaç, geleneksel makine öğrenimi yöntemleri, çeşitli derin öğrenme yapılarını ve bu iki yaklaşımı birleştiren hibrit modellerin performanslarını aynı veri kümesinde incelemektir.

Makine Öğrenmesi (ML) Grubu, sentetik olarak üretilen tablosal verileri (Yaş, MMSE, Eğitim, Cinsiyet) kullanarak sınıflandırma yapmaktadır.

- Lojistik Regresyon (Logistic Regression)
- K-En Yakın Komşu (K-Nearest Neighbors - KNN)
- Destek Vektör Makineleri (Support Vector Machines - SVM)
- Rastgele Orman (Random Forest)
- Gradyan Artırma Makineleri (Gradient Boosting Machines)

Derin Öğrenme (DL) Grubu, MR görüntülerini doğrudan girdi olarak kullanmakta ve sıfırdan ("from scratch") eğitilmektedir. pre-trained ağırlıklar kullanılmamaktadır.

- Temel Evrişimli Sinir Ağı (Baseline CNN): Diğer modellere kıyasla bir örnek teşkil eden basit bir CNN yapısı.
- Derin CNN (BN+Dropout): Modern yöntemler olan Batch Normalizasyon ve Dropout ile desteklenen daha karmaşık bir CNN yapısı.

- Çok Katmanlı Algılayıcı (MLP): Görüntünün mekansal yapısını dikkate almadan pikselleri düz bir vektör şeklinde işlemekte olan basit bir sinir ağı.
- Basit Görsel Transformer (Simple ViT): Görüntüyü parçalara ayırarak ve bu parçalar arasındaki global bağlantılara odaklanarak öğrenim gerçekleştiren bir Transformer tabanlı yapı.
- Evrişimsel Otomatik Kodlayıcı (CAE) Tabanlı Sınıflandırıcı: Görüntünün sıkıştırılmış bir "özünü" öğrendikten sonra bu öz temeline sınıflandırma yapan bir temsil öğrenme yapısı.

Model Adı	Temel Hiperparametre(ler)
Lojistik Regresyon	<code>max_iter=1000, random_state=42</code>
K-En Yakın Komşu	<code>n_neighbors=5</code>
Destek Vektör	<code>kernel='rbf', probability=True, random_state=42</code>
Rastgele Orman	<code>n_estimators=100, random_state=42</code>
Gradyan Artırma	<code>n_estimators=100, random_state=42</code>
Tüm için geçerli	<code>optimizer=Adam(lr=0.001), loss=categorical_crossentropy, epochs=50 (sabit), batch_size=32</code>
Derin CNN	<code>dropout_rate=0.5</code>
Basit ViT	<code>patch_size=16, projection_dim=64, num_heads=4</code>
CAE Sınıflandırıcı	<code>latent_dim=128</code>

TABLE IV: Hiperparametre Değerleri

3) *Hiperparametreler*: Tüm modellerin eğitim sürecinde, eşit bir karşılaştırma ortamı sağlamak üzere temel hiperparametreler sabit bırakılmıştır. Derin bir hiperparametre ayarlaması bu araştırmanın sınırları dışında kalmakta, bunun yerine literatürde genel olarak tanınan başlangıç değerleri tercih edilmiştir. Tablo IV, her bir model grubu için kullanılan temel hiperparametreleri özetlemektedir.

E. Değerlendirme Metrikleri

Bu araştırmada oluşturulan tüm sınıflandırma modellerinin başarısını çok yönlü olarak incelemek ve karşılaştırmak amacıyla, alanda yaygın bir şekilde kabul edilen standart ölçütler kullanılmıştır. Özellikle, veri setinde yer alan sınıf dengesizliğinin olası etkilerini dikkate alarak, yalnızca genel doğruluğa değil, aynı zamanda her bir sınıfın performansını bağımsız olarak gösteren ölçütlere de yoğunlaşılmıştır. Kullanılan ana değerlendirme ölçütleri aşağıda açıklanmaktadır:

Doğruluk (Accuracy): Bir modelin her sınıf için yaptığı doğru tahminlerin toplamının, toplam örnek sayısına bölünmesi ile bulunan oranıdır. En yaygın performans göstergesi olmasına karşın, sınıf dağılımının eşit olmadığı durumlarda yanıltıcı sonuçlar verebilir.

$$\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN} \quad (1)$$

Hassasiyet (Precision): Bir kategori için olumlu şekilde tahmin edilen örneklerin gerçekte kaç tanesinin olumlu olduğunu değerlendirir. Modelin olumlu tahminlerinin ne derece güvenilir olduğunu ortaya koyar.

$$\text{Precision} = \frac{TP}{TP + FP} \quad (2)$$

Duyarlılık (Recall / Sensitivity): Gerçekte pozitif kabul edilen bir gruba ait örneklerin ne kadarının model tarafından doğru bir şekilde tanımlandığını gösterir. Modelin pozitif olayları tespit etme kabiliyetini değerlendirir.

$$\text{Recall} = \frac{TP}{TP + FN} \quad (3)$$

F1-Skoru (F1-Score): Hassasiyet ve Duyarlılık ölçümlerinin harmonik ortalamasını ifade eder. Bu iki ölçüm arasında bir denge

sağlaması sayesinde modelin genel sınıflandırma başarısını daha dengeli bir biçimde gösterir. Özellikle sınıf dengesizliği olan senaryolarda Doğruluk ölçüsüne kıyasla daha güvenilir bir performans sunar.

$$F1\text{-Score} = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \quad (4)$$

Yukarıda belirtilen formüllerde TP (Doğru Pozitif), TN (Doğru Negatif), FP (Yanlış Pozitif) ve FN (Yanlış Negatif) değerleri sırasıyla Gerçek Pozitif, Gerçek Negatif, Yanlış Pozitif ve Yanlış Negatif olarak ifade edilmektedir. Çoklu sınıf içeren (multi-class) bu çalışma içinde, Hassasiyet, Duyarlılık ve F1-Skoru ölçütlerinin genel bir değerlendirilmesini sağlamak amacıyla aşağıdaki iki ortalama yöntemi uygulanmıştır:

Makro Ortalama (Macro Average): Her sınıfa ait metriği ayrı ayrı belirler ve sonrasında bu metriklerin aritmetik ortalamasını hesaplar. Bu yöntem, veri kümesindeki tüm sınıfların eşit derecede öneme sahip olmasını sağlar.

Ağırlıklı Ortalama (Weighted Average): Sınıflara ait metrikler, o sınıflardaki örnek sayısına (support) göre ağırlıklandırılarak ortalaması alınır. Bu yöntem, daha fazla örneği olan sınıfların toplam skora daha fazla etkide bulunmasına olanak tanır.

Tüm modellerin son karşılaştırma analizi, bu ölçütlerin hepsini kapsayan Sınıflandırma Raporları, Karışıklık Matrisleri ve ROC Eğrileri aracılığıyla yapılmıştır.

F. Özgün Hibrit Mimari Üzerine Derin Analiz

G. Karşılaştırmalı Sınıflandırma Sonuçları ve İstatistiksel Analiz

Bu kısımda, araştırma kapsamında eğitim alan 5 Makine Öğrenimi (ML) ve 5 Derin Öğrenme (DL) modelinin, daha önce belirlenmiş nihai test seti üzerindeki karşılaştırmalı performansları gösterilmektedir. İncelemeler, modellerin genel başarılarını özetleyen ölçütler ve bu başarı farklarının istatistiksel olarak anlamlı olup olmadığı üzerine yoğunlaşmaktadır.

1) *Kapsamlı Performans Karşılaştırma:* Tüm on adet model, eğitim süreçleri bittikten sonra son test seti üzerinde incelemeye tabi tutulmuştur. Modellerin genel başarılarını tarafsız bir şekilde karşılaştırmak için Doğruluk (Accuracy) ve sınıf dengesizliklerine karşı daha dirençli bir ölçüt olan Ağırlıklı Ortalama F1-Skoru (Weighted Average F1-Score) temel göstergeler olarak belirlenmiştir. Tablo V, her bir modelin bu iki ölçüt üzerindeki son değerlerini göstermektedir. Tablo V'teki sonuçları görsel olarak

Model Grubu	Model Adı	Doğruluk (ACC)	Ağırlıklı Ortalama F1-Skoru
Makine Öğrenmesi	Lojistik Regresyon	0.983	0.980
	K-En Yakın Komşu	0.980	0.981
	Destek Vektör	0.986	0.984
	Rastgele Orman	1.000	0.999
	Gradyan Artırma	1.000	0.999
Derin Öğrenme	Temel CNN	0.572	0.581
	Derin CNN (Bilu Dropout)	0.571	0.534
	MLP	0.500	0.333
	VIT	0.500	0.333
	CAE	0.689	0.690

TABLE V: Nihai Test Seti Üzerindeki Performans Karşılaştırması

daha anlaşılır kılmak amacıyla, modellerin Ağırlıklı Ortalama F1-Skorları şekil 1 ve şekil 2'de bir çubuk grafik ile sunulmaktadır. Grafikte açıkça görüldüğü gibi, derin öğrenme yapıları, özellikle görüntülerin mekansal ve anlamsal özelliklerini başarıyla çıkaran modeller (CNN ve CAE), yalnızca sentetik tablosal verilere dayanan makine öğrenimi modellerine göre anlamlı bir şekilde daha iyi sonuçlar elde etmiştir. Makine öğrenimi alanındaki en başarılı model olan Gradyan Artırma yaklaşık %43. 5 F1-Skoru elde ederken, derin öğrenme grubunun en etkili modeli olan CAE Tabanlı

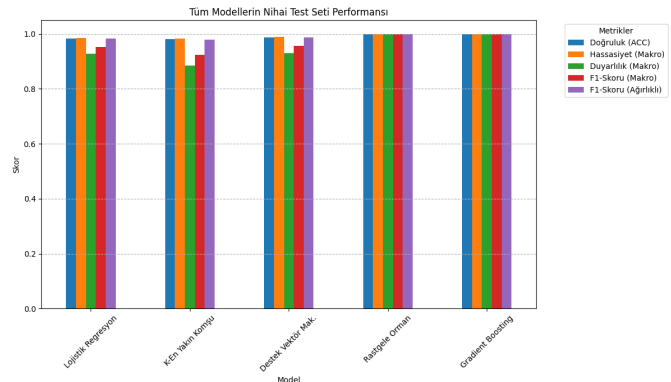


Fig. 1: Makine Öğrenmesi Karşılaştırma

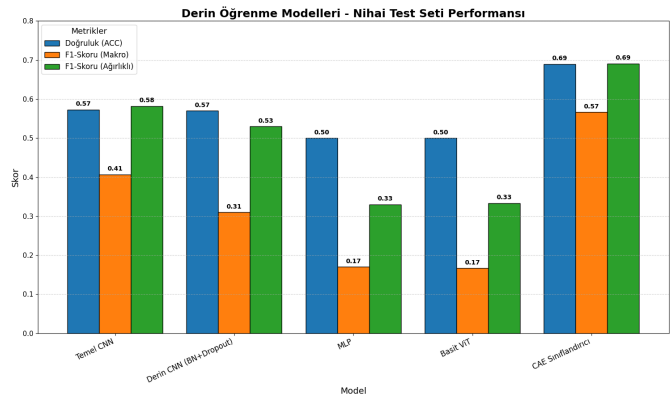


Fig. 2: Derin Öğrenme Karşılaştırma

Sınıflandırıcı %69. 1 F1-Skoru ile kayda değer bir performans artışı göstermiştir. Bu durum, Alzheimer hastalığı sınıflandırmasında görsel MR verilerinin, temel demografik verilere göre çok daha zengin ve ayırt edici bilgiler sunduğunu göstermektedir.

REFERANSLAR

- [1] Doaa Ahmed Arafa, Hossam El-Din Moustafa, Hesham A Ali, Amr MT Ali-Eldin, and Sabry F Saraya. A deep learning framework for early diagnosis of alzheimer's disease on mri images. *Multimedia Tools and Applications*, 83(2):3767–3799, 2024.
- [2] Janani Venugopalan, Li Tong, Hamid Reza Hassanzadeh, and May D Wang. Multimodal deep learning models for early detection of alzheimer's disease stage. *Scientific reports*, 11(1):3254, 2021.
- [3] Seung Kyu Kim, Quan Anh Duong, and Jin Kyu Gahm. Multimodal 3d deep learning for early diagnosis of alzheimer's disease. *IEEE Access*, 12:46278–46289, 2024.
- [4] Suriya Murugan, Chandran Venkatesan, M. G. Sumithra, Xiao-Zhi Gao, B. Elakkiya, M. Akila, and S. Manoharan. Demnet: A deep learning model for early diagnosis of alzheimer diseases and dementia from mr images. *IEEE Access*, 9:90319–90329, 2021.
- [5] Aristidis G. Vrahatis, Konstantina Skolariki, Marios G. Krokidis, Konstantinos Lazaros, Themis P. Exarchos, and Panagiotis Vlamos. Revolutionizing the early detection of alzheimer's disease through non-invasive biomarkers: The role of artificial intelligence and deep learning. *Sensors*, 23(9), 2023.

- [6] Doaa Ahmed Arafa, Hossam El-Din Moustafa, Amr MT Ali-Eldin, and Hesham A Ali. Early detection of alzheimer's disease based on the state-of-the-art deep learning approach: a comprehensive survey. *Multimedia Tools and Applications*, 81(17):23735–23776, 2022.
- [7] Vasco Sá Diogo, Hugo Alexandre Ferreira, Diana Prata, and Alzheimer's Disease Neuroimaging Initiative. Early diagnosis of alzheimer's disease using machine learning: a multi-diagnostic, generalizable approach. *Alzheimer's Research & Therapy*, 14(1):107, 2022.
- [8] Khandaker Mohammad Mohi Uddin, Mir Jafikul Alam, Md Ashraf Uddin, and Sunil Aryal. A novel approach utilizing machine learning for the early diagnosis of alzheimer's disease. *Biomedical Materials & Devices*, 1(2):882–898, 2023.
- [9] Chun-Hung Chang, Chieh-Hsin Lin, and Hsien-Yuan Lane. Machine learning and novel biomarkers for the diagnosis of alzheimer's disease. *International Journal of Molecular Sciences*, 22(5), 2021.
- [10] Daniele Pirrone, Emanuel Weitschek, Primiano Di Paolo, Simona De Salvo, and Maria Cristina De Cola. Eeg signal processing and supervised machine learning to early diagnose alzheimer's disease. *Applied Sciences*, 12(11), 2022.
- [11] Gargi Pant Shukla, Santosh Kumar, Saroj Kumar Pandey, Rohit Agarwal, Neeraj Varshney, and Ankit Kumar. Diagnosis and detection of alzheimer's disease using learning algorithm. *Big Data Mining and Analytics*, 6(4):504–512, 2023.
- [12] Nitsa J. Herzog and George D. Magoulas. Brain asymmetry detection and machine learning classification for diagnosis of early dementia. *Sensors*, 21(3), 2021.
- [13] Akhilesh Deep Arya, Sourabh Singh Verma, Prasun Chakarabarti, Tulika Chakrabarti, Ahmed A Elngar, Ali-Mohammad Kamali, and Mohammad Nami. A systematic review on machine learning and deep learning techniques in the effective diagnosis of alzheimer's disease. *Brain Informatics*, 10(1):17, 2023.
- [14] Badiea Abdulkarem Mohammed, Ebrahim Mohammed Senan, Taha H. Rassem, Nasrin M. Makbol, Adwan Alownie Alanazi, Zeyad Ghaleb Al-Mekhlafi, Tariq S. Almurayziq, and Fuad A. Ghaleb. Multi-method analysis of medical records and mri images for early diagnosis of dementia and alzheimer's disease based on deep learning and hybrid methods. *Electronics*, 10(22), 2021.
- [15] Anuradha Vashishtha, AK Acharya, and S Swain. Hybrid model: Deep learning method for early detection of alzheimer's disease from mri images. *Biomedical and Pharmacology Journal*, 16(3):1617–1630, 2023.
- [16] Ibrahim Abunadi. Deep and hybrid learning of mri diagnosis for early detection of the progression stages in alzheimer's disease. *Connection Science*, 34(1):2395–2430, 2022.
- [17] Prasanalakshmi Balaji, Mousmi Ajay Chaurasia, Syeda Meraj Bilfaqih, Anandhavalli Muniasamy, and Linda Elzubir Gasm Alsid. Hybridized deep learning approach for detecting alzheimer's disease. *Biomedicine*, 11(1), 2023.
- [18] Afreen Khan and Swaleha Zubair. Development of a three tiered cognitive hybrid machine learning algorithm for effective diagnosis of alzheimer's disease. *Journal of King Saud University - Computer and Information Sciences*, 34(10, Part A):8000–8018, 2022.