

## Alzheimer Hastalığının Teşhisinde Görüntü Dönüştürücü (Vision Transformer) Yaklaşımı: Yenilikçi Bir İnceleme

Mehmet Emre SERTKAYA<sup>1\*</sup>, Burhan ERGEN<sup>2</sup>

<sup>1</sup> Uzaktan Eğitim, Uygulama ve Araştırma Merkezi, Rektörlük, Samsun Üniversitesi, Samsun, Türkiye

<sup>2</sup> Bilgisayar Mühendisliği, Mühendislik Fakültesi, Fırat Üniversitesi, Elazığ, Türkiye

\*<sup>1</sup> emre.sertkaya@samsun.edu.tr, <sup>2</sup> [bergen@firat.edu.tr](mailto:bergen@firat.edu.tr)

(Geliş/Received: 28/12/2023;

Kabul/Accepted: 27/03/2024)

**Öz:** Nörodejeneratif bir hastalık olan Alzheimer hastalığı, Demans'ın en yaygın türüdür. Şu an için kesin bir tedavisi bulunmayan bu hastalığın ilerlemesini yavaşlatıcı tedavi yöntemleri uygulanmaktadır. Bu nedenle, hastalığın erken teşhisi ve diğer hastalıklarla karıştırılmaması kritik öneme sahiptir. Bu çalışmada, Alzheimer's Disease Neuroimaging Initiative (ADNI) tarafından toplanan üç boyutlu MR görüntüleri kullanılarak Görüntü Dönüştürücü yöntemleriyle Alzheimer hastalığının tespit edilmesi amaçlanmaktadır. Alzheimer hastaları (AD), Hafif Bilişsel Bozukluk (Mild Cognitive Impairment - MCI) ve sağlıklı bireylerden (Cognitive Normal - CN) oluşan bu veri seti, %70'i eğitim, %10'u doğrulama ve %20'si test veri setleri olarak ayrılmıştır. Literatürdeki çeşitli derin öğrenme yöntemlerinin yanı sıra yeni bir yaklaşım olan Görüntü Dönüştürücü (Vision Transformer) kullanılarak sınıflandırma yapılmıştır. Çalışma sonuçları, test görüntülerinde Görüntü Dönüştürücü'nün AD/MCI ikili sınıflandırmasında %79,8 başarı, MCI/CN ikili sınıflandırmasında %80,3 başarı ve AD/CN ikili sınıflandırmada %89,3 başarı elde ettiğini göstermektedir.

**Anahtar kelimeler:** Alzheimer Hastalığı, Görüntü Dönüştürücü, Evrimsel Ağlar, ADNI veri seti.

### Vision Transformer Approach in the Diagnosis of Alzheimer's Disease: An Innovative Review

**Abstract:** Alzheimer's disease, a neurodegenerative disease, is the most common type of Dementia. Currently, there is no definitive cure for this disease and treatment methods are applied to slow down the progression of this disease. Therefore, early diagnosis of the disease and prevention of confusion with other diseases is of critical importance. In this study, it is aimed to detect Alzheimer's disease with Image Transform methods using three-dimensional MR images collected by the Alzheimer's Disease Neuroimaging Initiative (ADNI). This dataset, which consists of Alzheimer's patients (AD), Mild Cognitive Impairment (MCI) and healthy individuals (Cognitive Normal (CN)), is divided into 70% training, 10% validation and 20% test datasets. In addition to various deep learning methods in the literature, classification was performed using a new approach, Vision Transformer. The results of the study show that the Vision Transformer achieved 79.8% success in AD/MCI binary classification, 80.3% success in MCI/CN binary classification and 89.3% success in AD/CN binary classification.

**Key words:** Alzheimer's Disease, Vision Transformer, Convolutional Networks, ADNI dataset.

### 1. Giriş

Alzheimer Hastalığı (AH), Demansın en yaygın türüdür ve beyindeki nörodejeneratif süreçlerden kaynaklanan çeşitli sorunlara sebep olur [1]. Bu hastalık, günümüzdeki en tehlikeli ve ölümcül hastalıklardan biri olarak kabul edilmektedir [2]. AH'nin erken evrelerinde beyinde biriken amiloid lezyonları, nörodejenerasyona yol açan tau patolojisi ve hafif bilişsel değişiklikler gibi belirtiler görülür. Ancak, hastalarda amiloid lezyon ve nörodejenerasyon bulguları olmasına rağmen bilişsel değişiklikler her zaman gözlenmez, bu da erken teşhisi güçleştirir ve tedavi sürecini karmaşık hale getirir [3].

Dünya genelinde yaklaşık 50 milyon Alzheimer hastası bulunmaktadır ve bu rakamın her 5 yılda bir iki katına çıkması beklenmektedir [4]. AH belirtileri genellikle 60 yaşından sonra ortaya çıksa da genetik yatkınlığa bağlı olarak 30-50 yaşları arasında da görülebilir. Bu hastalık, beyinde yapısal ve işlevsel değişikliklere neden olur ve genellikle ileri aşamalarda teşhis edilir. Hasta önce Hafif Bilişsel Bozukluk (HBB) olarak tanı alır ve zamanla Alzheimer'a dönüşebilir, ancak her HBB Alzheimer'a ilerlemeyebilir. Bu değişiklikler, tıbbi görüntüleme teknikleri gibi yöntemlerle ölçülebilir [5, 6]. Bununla birlikte, AH'ye dönüşecek hafif bilişsel bozukluk, herhangi bir AH belirtisi ortaya çıkmadan 20 yıl veya daha uzun bir süre önce başlayabilir. Erken Hafif Bilişsel Bozukluk (EMCI), MCI ve Geç Hafif Bilişsel Bozukluk (LMCI) evreleri, durumun günlük aktiviteleri bozacak kadar ilerlediği zamanlardır. AH tedavi edilemezse de, erken evrelerinde yavaşlatılabilir.

Sonuç olarak, AD'nin erken teşhisi, hastaların yaşam kalitesini ve tedavi denemelerinin sonuçlarını iyileştirmek için son derece arzu edilen bir durumdur. Nörogörüntüleme teknolojilerindeki hızlı ilerleme nedeniyle

\* Sorumlu yazar: [emre.sertkaya@samsun.edu.tr](mailto:emre.sertkaya@samsun.edu.tr). Yazarların ORCID Numarası: <sup>1</sup> 0000-0001-5060-1857, <sup>2</sup> 0000-0003-3244-2615

nörogörüntüleme tanısı, AH tanısı koymanın en sezgisel ve güvenilir yolu haline gelmiştir [7]. AH erken teşhisi, beyin anormal dejenerasyonunu yavaşlatmak ve daha iyi tedavi sağlamak için araştırılmaktadır. AH üzerine yapılmış çalışmalar, erken müdahale ve teşhisin tedavinin etkinliği için çok etkili olduğu gözlenmiştir. AH teşhisinde bilgisayar destekli tanı yöntemlerinin çok yararlı olduğu görülmüştür [8].

Nörogörüntüleme alanında yapay zeka, nörobilimcilerin tıbbi görüntüleme verilerini analiz etmek için geliştirilen ve hala gelişmekte olan bir teknoloji olarak kabul edilir. Son on yılda, derin öğrenme teknikleri nörobilimcilerin Alzheimer hastalığının çeşitli aşamalarını tahmin etmelerine yardımcı oldu. Bu teknoloji, bilim insanlarının güçlü hesaplama kaynaklarıyla medikal görüntüleme verilerini önceden işleyerek, karmaşık derin öğrenme modelleri oluşturmasına ve tıp doktorlarına erken evre Hafif Bilişsel Bozukluklu (MCI) beyinlerini, normal yaşlanma sürecindeki beyinlerden yüksek derecede ayırt etmelerine olanak tanıyacak sonuçları işlemelerine imkan sağlamaktadır. Bu ilerlemeler, Alzheimer hastalığının erken teşhisi ve tedavisi için umut vadeden yeni yaklaşımlar sunmaktadır [9-11]. Literatürde çalışmalar incelendiğinde;

Ilias, L. ve arkadaşlarının yaptığı çalışmada bu veri setini kullanarak tek görevli ve çok görevli derin öğrenme modelleri kullanarak Alzheimer hastalığının teşhisi için model önerdiler. Yaptıkları çalışmada transformator tabanlı BERT modelini kullanarak %87,5 doğruluğa ulaştılar. Çok görevli tanımladıkları yapıda ise ana görev Alzheimer hastalığının tespiti iken yan görev olarak bu hastalığın safhasını tespit etmede %86,25 doğruluk sergilediği görülmüştür [12].

Jain, V. ve arkadaşlarının yaptığı çalışmada ise Generative Adversarial Networks (GAN) mimarileri kullanarak veri artırımı yaparak dengesiz veri setini düzeltmeye amaçlamıştır. Sınıflardaki eşit miktardaki görüntüleri derin öğrenme modellerine verildiğinde VGG16 ve VGG19 derin öğrenme modellerinin %84 - %87 arasında başarımlar gösterdiği gözlenmiştir [13]. Başka bir çalışmada ise bu çalışma da kullanılan ADNI veri setini kullanarak verileri kesit kesit almak yerine 3 boyutlu alan Zhang, X. ve arkadaşları Attention tabanlı derin öğrenme modeli önermişlerdir. Başarımları artırmak ve Attention mekanizmasının odak noktasını belirlemek adına veri seti ilk önce Grad-CAM(Gradyan Tabanlı Yerelleştirme Sınıfı Aktivasyon Haritalaması) ile ön işleme sokmuşlardır. Çalışma sonuçları ResAttNET modelinin AD-CN ikili sınıfını %91 ve AD-MCI %82 doğruluk sergilediği gözlenmiştir [14].

Derin öğrenme modellerindeki aktivasyon fonksiyonlarını geliştirerek başarımları artırmaya çalışan Oktavian, M. ve arkadaşları AH tespitinde önerdikleri yöntemde mesh aktivasyon fonksiyonu kullanarak AH tespitinde %88,3 doğruluk sergileyecek modeller önermişlerdir [15]. Ma, H. ve arkadaşları yapmış olduğu çalışmada AH erken teşhisinde Deep Q-Network (DQN) modeli önerdiler. Pre-trained (başlangıç ağırlıklı) bu model kullanarak yapmış oldukları çalışmada DQN sınıflandırıcı AD ve CN sınıflandırmak için kullanılmıştır. Bir permütasyon testi gerçekleştirilmiştir. DQN transfer öğrenimi %86,66 sınıflandırma doğruluğu göstermiştir [16].

Mujahid, M. ve arkadaşlarının yaptığı çalışmada EfficientNet-B2 ve VGG-16'dan oluşan bir topluluk yapılı birleşik bir model önerdiler. Bu iki modelden ayrı ayrı son katmanlarından alınan özellikler birleştirilerek çeşitli katmanlar arayıcılığı ile özellikleri çıkarılarak AH safhaları bulmayı amaçlamışlardır. Çalışma %95,89 doğruluk başarımlarını göstermiştir [17].

Bu çalışmada 3 boyutlu MR görüntülerin elde edilen Alzheimer ve MCI görüntüleri kullanılarak Bu çalışmanın amaçları şunları içerir:

- 3 boyutlu medikal görüntülerin kullanımıyla daha kesin sonuçlar elde etmek,
- Sınırlı veri miktarına sahip görüntülerde performansı artırmak,
- Vision Transformer gibi yeni bir yöntemin Alzheimer hastalığının tespitindeki başarımlarını araştırmak,
- Görüntü dönüştürücü tekniklerinin performansını değerlendirmek.

Bu hedefler, derin öğrenme ve özellikle Vision Transformer gibi yenilikçi yöntemlerin Alzheimer hastalığının teşhisi üzerindeki etkisini ve potansiyelini anlamayı amaçlamaktadır. Bu yöntemlerin, az veriyle çalışma yeteneği ve 3 boyutlu medikal görüntülerde daha hassas sonuçlar elde etme olasılığı, Alzheimer teşhisi ve erken belirtilerinin tespiti konusunda yeni bir pencere açabilir.

Bu makalenin devamında, ikinci bölüm olan “Materyal ve Yöntem” kısmında kullanılan veri seti ve modeller detaylı bir şekilde açıklanmıştır. Veri setinin yapısı, toplama süreci ve kullanılan analitik teknikler bu bölümde ayrıntılı olarak ele alınmıştır. Önerilen metodoloji ve kullanılan algoritmaların yanı sıra görüntü dönüştürücü (Vision Transformer) yönteminin nasıl entegre edildiği de bu kısımda vurgulanmıştır. Üçüncü bölüm olan “Bulgular ve Tartışma” kısmı, elde edilen sonuçların derinlemesine incelendiği ve analiz edildiği bir kısımdır. Bu bölümde, görüntü dönüştürücü yönteminin performansı, sınıflandırma sonuçları ve veri setindeki farklı sınıflar arasındaki ilişkiler detaylı olarak sunulmuştur. Bulguların kapsamlı bir şekilde tartışılmasıyla, kullanılan metodolojinin avantajları ve sınırlamaları üzerinde odaklanılmıştır. Son bölüm olan “Sonuç” kısmında, elde edilen bulguların özeti sunulmuş ve bu çalışmanın önemi vurgulanmıştır. Çalışmanın katkıları, bulguların klinik ve araştırma alanlarındaki potansiyel etkileri ve önerilen metodolojinin gelecekteki kullanımı bu bölümde öne çıkarılmıştır.

## 2. Materyal ve Yöntem

### 2.1 Veri Seti

Bu çalışmada kullanılan veri seti Alzheimer's Disease Neuroimaging Initiative (ADNI) tarafından toplanan 3 boyutlu MR görüntülerinden oluşmaktadır [18]. ADNI-1, 2004 yılında başlamış ve orijinal olarak 2009 yılında planlanan beş yıllık bir süre boyunca devam etmiştir. Ancak daha sonra 2010 yılında beş yıl uzatılmış ve toplamda on yıllık bir süre boyunca devam etmiştir. Dolayısıyla, ADNI-1 veri toplama süreci, 2004 ile 2014 yılları arasında toplam on yıl sürmüştür. Bu süre boyunca, katılımcıların farklı aşamalarda verileri düzenli aralıklarla toplanmış ve kaydedilmiştir. ADNI-1 veri seti kullanıldığında, normal yaşlanma sürecinden başlayarak hafif kognitif bozukluk ve Alzheimer hastalığına kadar geniş bir yelpazede katılımcıları içermektedir. Bu, hastalığın farklı aşamalarındaki değişikliklerin izlenmesine ve anlaşılmasına olanak tanır. ADNI-1, geniş bir katılımcı grubunu içerir ve bu da verilerin çeşitliliğini artırır. Bu, Alzheimer hastalığının farklı demografik gruplardaki etkilerini anlamak için önemlidir. Bu veri setine ait demografik bilgiler Tablo 1'de gösterilmektedir.

**Tablo 1.** Veri setine ait demografik bilgiler.

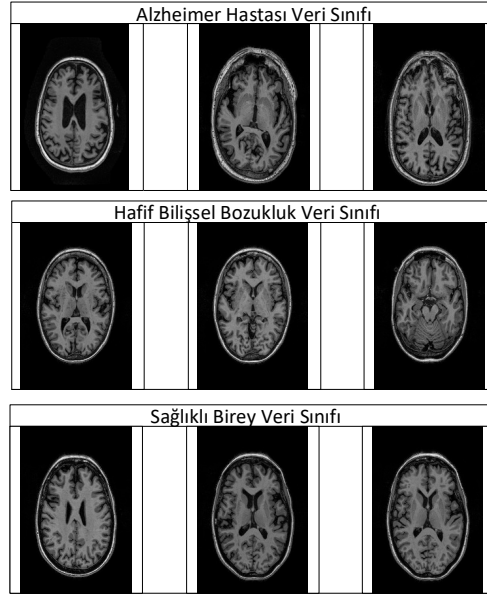
Sınıf Adı	Örnek Sayısı	Cinsiyet		Yaş		Mini Mental Sağlık Değeri-MMSE	
		Erkek	Kadın	Ortalama	Std.	Ortalama	Std.
Alzheimer	192	101	91	75,3	7,5	23,3	2,1
Hafif Bilişsel Bozukluk(MCI)	396	255	141	74,7	7,4	27,0	1,8
Normal	229	120	109	75,8	5,0	29,1	1,0

Tablo 1 incelendiğinde, 817 hastadan oluşan ve çeşitli aralıklarla kontrollere gelmeleri sonucu toplanan medikal verilerden oluştuğu görülmektedir. Bu veri seti, özellikle Alzheimer hastalığının erken dönemlerinde tanı ve ilerlemenin izlenmesi için uygun bir veri kaynağı olarak düşünülmektedir. Veri seti 3 adet sınıftan oluşmakta olup bunlar; Alzheimer hastalığı(Alzheimer Diagnosis - AD) sınıfı, Hafif Bilişsel Bozukluk(Mild Cognitive Impairment - MCI) sınıfı ve Sağlıklı(Cognitive Normal - CN) bireylerden oluşmaktadır. Bunlara ait görüntüler Şekil 1'de gösterilmektedir [18-20]. Bu görüntüler nifti uzantılı 3 boyutlu görüntülerden oluşmaktadır. Literatürde yapılan çalışmalar incelendiğinde, veri seti aksenel, sagittal ve koronal bölümlerin kesitleri alınarak ya da 3 boyutlu olarak kullanıldığı gözlemlenmiştir [21, 22]. Genel olarak 256x256x166 boyutlarında olan bu MR görüntüleri, aksenel boyutta rasgele 70 ile 130 kesitleri arasından seçilmiştir. Çalışmanın odak noktası açısından biz en çok kullanılan aksenel görüntüler ait birkaç kesiti alınarak çalışma gerçekleştirilmiştir.

**Tablo 2.** Veri seti görüntü miktarları ve dağılımları.

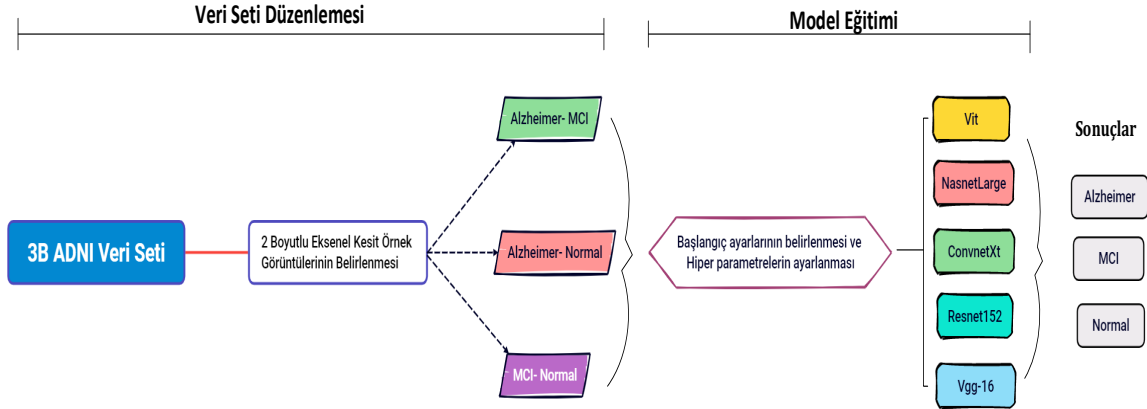
Sınıf Adı	Eğitim Görüntü Sayısı (%70)	Doğrulama Görüntü Sayısı (%10)	Test Görüntü Sayısı (%20)	Görüntü Sayısı
Alzheimer	786	112	226	1124
Hafif Bilişsel Bozukluk	1812	259	519	2590
Normal	1007	144	289	1440

Görüntüler modele verilmeden önce 3 gruba ayrıldı: eğitim, doğrulama ve test. Modeller eğitilirken her iterasyonun sonunda doğrulama görüntüleri kullanıldı. Son iterasyonda ise test görüntüleri verilerek başarımlar ölçüldü. Veri dağılım oranları %70 eğitimi, %10 doğrulama ve %20 test olarak ayarlandı. Bu dağılımlar ve görüntü miktarları ve dağılımı Tablo 2'de gösterilmektedir.



Şekil 1. Veri seti örnek görüntüleri.

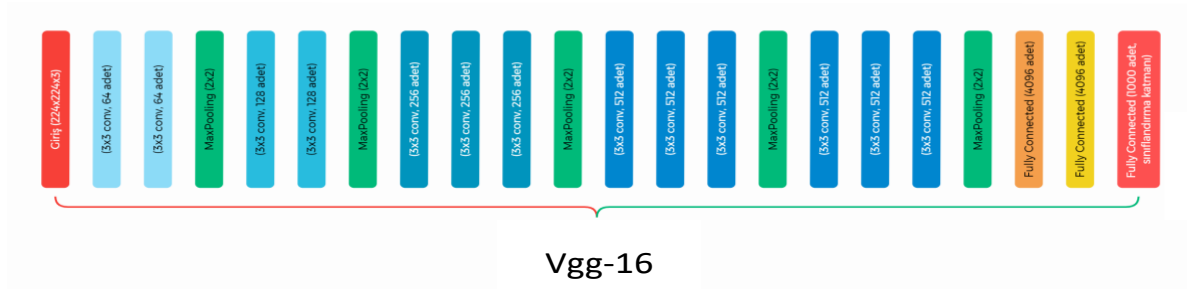
3 boyutlu MR görüntülerinden elde ettiğimiz bu görüntüleri 224x224 boyutlarında olmak üzere VGG-16, NasnetLarge, Resnet-152, ConvnetXt ve Vision Transformer modellerin hastalıkları öğrenmesi için verdik. Çalışmada diğer derin öğrenme yöntemleri de kullanıldı Aşırı öğrenme durumunda geçmeleri ve test değerlerini çok düşük olması nedeniyle bu yöntemler çalışmaya eklenmemiştir. Şekil 2’de çalışmanın iş akışı gösterilmektedir. Sınıf bazlı başarımları kıyaslamak adına ikili sınıflandırma şeklinde veri setini modellere verdik.



Şekil 2. Çalışmanın iş akış süreci.

## 2.2 Vgg-16

Evrişimsel Sinir Ağı (ESA) olan VGG16, Visual Geometry Group (VGG) tarafından Oxford Üniversitesi’nde geliştirilmiştir. Adından da anlaşılacağı üzere 16 sıralı katmandan oluşan bir yapıya sahiptir. ImageNet veri tabanı üzerinde ön eğitilmiş bu model, veri setindeki örnek sayısı az olmasına rağmen iyi sonuçlar verdiği gözlenmiştir. Modelin iç mimarisi Şekil 3’te gösterilmekte olup 5 adet havuzlama katmanı, 3 adet tam bağlı katman olmak üzere 16 evrişimsel katmandan oluşmaktadır [23, 24].



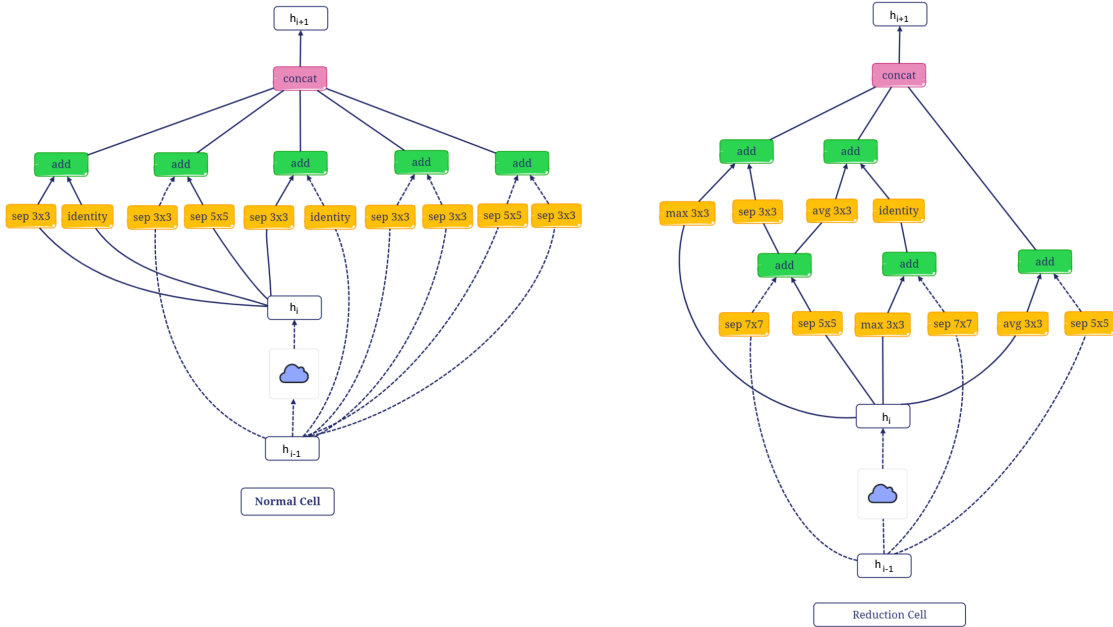
Şekil 3. Vgg-16 Mimarisi.

### 2.3 NASNetlarge

Nöral Mimari Arama (Neural Architecture Search - NAS) çerçevesinden esinlenen NASNetLarge takviyeli öğrenme arama yöntemlerinden biridir. Model Zoph, B ve arkadaşları tarafından önerilmiştir [25]. Bu model, ImageNet veri görüntü kümesinde 1 milyondan fazla görüntü ile eğitilmiştir. NASNetLarge, insan eliyle tasarlanmayan tamamıyla otomatik üretilen bir mimari arama sürecinde oluşturulmuştur. NAS, bir arama uzayında dolaşarak ve optimize edici algoritmalar kullanarak en iyi mimari yapıyı bulmaya çalışır. Evrimsel katmandaki performansı artırmak için hücre yapısı oluşturmuştur. Bu hücre yapısının ilk katmanı tipik Evrimsel katman iken sonraki katmanı azaltma hücresidir [26]. Giriş görüntülerinin özellik vektörleri 4032 eleman olarak tanımlanmıştır ve NASNetLarge için varsayılan giriş görüntü boyutu 331x331'dir [27].

Şekil 4'te NASNetLarge hücre yapıları gösterilmektedir. Normal bir hücre birden fazla evrişim ve aktivasyon katmanlarının birbirleriyle bağlantıları ve geçişleri bulunur. Bu hücre, modelin genel yapısını oluşturan önemli bir bileşendir.

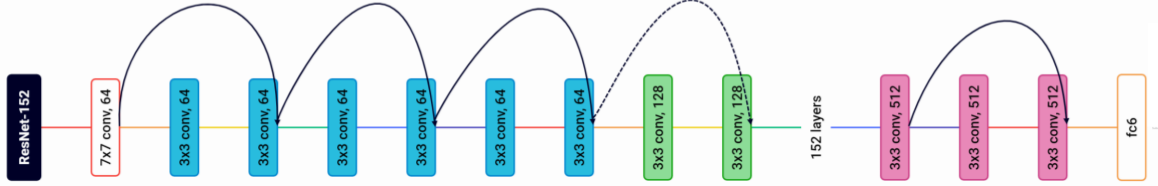
İndirgenmiş hücre ise, normal hücreye benzer bir yapıya sahiptir ancak genellikle daha küçültülmüş bir yapıdadır. Bu hücre, genellikle daha az sayıda parametreyle ve hesaplama maliyeti daha düşük bir yapıyla tasarlanır. İndirgenmiş hücre, modelin boyutunu azaltmak ve daha hafif bir model elde etmek için kullanılır.



Şekil 4. NASNetLarge hücre yapısı.

## 2.4 ResNet-152

ResNet, Kaiming He ve arkadaşları tarafından 2016 yılında artık birimlerin kullanılmaya başlamasıyla ortaya çıkan Artık Sinir Ağı (Residual Neural Network) evrimsel sinir ağı modelidir. Düz bir sinir ağı derinliğinin artması sonucu katmanlarda kaybolan gradyan kaybına ve ağ bozulması sorunlarını çözmek için oluşturulmuştur. ResNet, özellikle görüntü tanımda yepyeni bir fikir ortaya atmıştır [28, 29]. ResNet, girdiyi değiştirmeden bir katmandan diğerine sığdırmak için atlama bağlantılarını (kısayol bağlantıları olarak da bilinir) kullanır. Bu sayede hem modelin ezber yapmasını engeller hem de ara katmanlardaki kaybolan özellikleri daha sonraki katmanlara taşır [30]. Bu yapı 152 katmandan oluşmaktadır. Şekil 5'te ResNet-152 mimari yapısı gösterilmektedir.

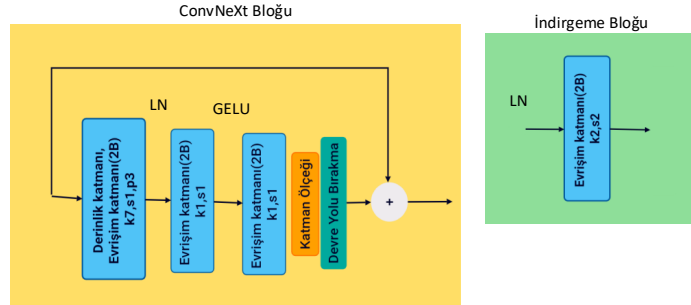


Şekil 5. ResNet Mimarisi.

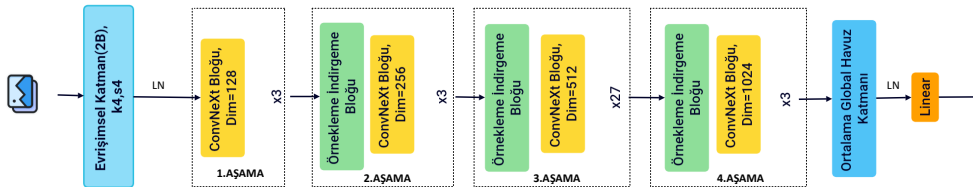
## 2.5 ConvNeXt

Bu çalışmada kullanılan diğer bir yöntem olan ConvNeXt, ilk olarak 1980 yılında tanıtılan ConvNet, geri yayılım olarak önerilmiştir. Zamanla optimizasyon, doğruluk ve verimlilik açısından sayısız gelişme kaydetmiştir. Daha sonra, özellikle ImageNet görüntü setinde gösterdiği başarıyla tam anlamıyla keşfedilmiştir [31].

ConvNeXt ağı, geleneksel ağ modellerinden farklı olarak genel yapı, derin evrişim, ters şişirme, büyük evrişim çekirdeği, GELU aktivasyon fonksiyonu ve LN katmanı gibi çeşitli alanlarda geliştirmeler yapmıştır. Genel yapı açısından, ConvNext ağının Kök katmanı, Swin Transformer'a benzer şekilde aynı boyuttaki bir evrişim çekirdeği ve dört adımlı bir evrişim işlemi kullanır. Evrişim konusunda, ConvNeXt ağı giriş ve çıkış kanal miktarlarını ayırarak derin evrişim tasarımının parametre boyutunu azaltır, bu da geleneksel evrişimden çok daha küçüktür [32, 33]. Şekil 6'da ConvNeXt blok yapıları gösterilmiştir. Bu blok dizinleri sıralı bir şekilde modelde bulunarak özellikleri çıkarmaktadır. Şekil 6'da gösterilen blokları kullanan ConvNeXt mimarisi Şekil 7'de gösterilmektedir.



Şekil 6. ConvNeXt Blok Yapıları.

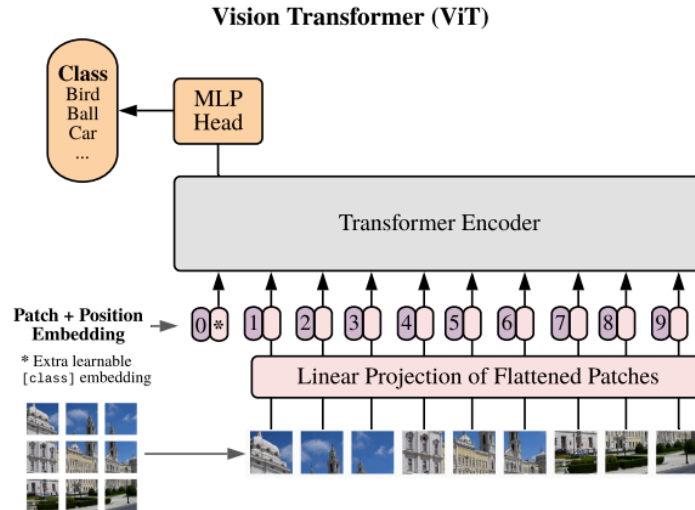


Şekil 7. ConvNeXt Mimarisi.

## 2.6 Vision Transformer

Dönüştürücülerin makine çevirilerindeki başarılarından esinlenerek, sadece dönüştürücü katmanlar kullanan evrişim katmanı olmadan özellikler çıkaran bir model olarak ortaya çıkmıştır. Özellikle Vision Transformer'ler (ViT) görüntü sınıflandırması için CNN'lerle eşleşen veya hatta onları aşan dönüştürücü tabanlı yöntemin ilk örneğidir [34, 35]. ViT birçok çeşidi yakın zamanda önerilmiştir; bunlar arasında veri verimli eğitim için damıtma [36], yapay sinir ağlarındaki gibi piramit yapı ya da tümünden kendine dikkat yerine soyut bir temsil öğrenerek verimliliği artırmak için öz-dikkat(self-attention) bulunmaktadır [37]. ViT model yapısı Şekil 8'de gösterilmektedir [35]. ViT'nin ESA modellerine göre başlıca avantajları şu şekilde ifade edilebilir: Dikkat Mekanizması, Ölçeklenebilirlik, Transfer Öğrenme Yeteneği ve Daha Az Veri Gereksinimidir. Dikkat Mekanizması, ViT'in dikkat mekanizması kullanarak görüntü işleme sürecinde pikseller arasındaki ilişkileri modelleyebilir. Bu, piksel bazında değil, pikseller arası ilişkilere dayalı olarak özellik çıkarmaya olanak tanır ve daha uzak ilişkileri yakalayabilir. Ölçeklenebilirlik, geleneksel ESA yapılarına göre daha ölçeklenebilir bir yapı sunar. Daha büyük görüntü boyutlarına ve daha fazla parametreye sahip olabilecek şekilde eğitilebilir. Bu, daha karmaşık görevlerde daha iyi performans elde etme potansiyelini artırır. Transfer Öğrenme Yeteneği, büyük ölçekli dil modellerinde olduğu gibi transfer öğrenme yeteneğine sahiptir. Önceden eğitilmiş bir ViT modeli, görsel görevler için başlangıç noktası olarak kullanılabilir ve daha az etiketlenmiş veriyle eğitim için kullanılabilir. Daha Az Veri Gereksinimi, ViT'nin daha az veriyle eğitilmesine olanak tanır. Geleneksel ESA yapılarına göre daha az veriyle iyi sonuçlar elde etme eğilimindedir. ViT, geleneksel evrişimsel sinir ağlarındaki(ESA) yapı yerine tamamen dikkat mekanizmasına dayalı(attention) bir mimari kullanır. ViT adımları şöyle açıklanabilir;

- Resim bloğu, giriş görüntülerini küçük yama dediğimiz parçalara böler. Bu parçalar gömme(Embedding) işlemiyle vektörel bir yapıya dönüştürür. ESA'lardaki gibi görüntüyü tüm olarak modele vermez.
- Patch adımı, görüntü yamaları oluşturulur ve bu yamalar gömme işlemine tabi tutulur. Her bir yamanın gömülü gösterimi, daha sonra ViT'nin girişine verilecek olan yığını oluşturur.
- Pozisyon Gömme, her yamaya bu adımda konumu veya pozisyon bilgileri verilir. Bu sayede ViT'nin dikkat mekanizması her yamanın çıkaracağı özellikler ile birlikte konum bilgisine de dikkate alır.
- Transformer Encoder, görüntüyü işlemek için birden çok odaklı transformer encoder katmanı kullanılır. Bu katmanlar, dikkat mekanizmalarını ve çoklu-kafa (multi-head) dikkat başlıklarını içerir. Giriş yığını, bu Transformer encoder katmanlarından geçer.
- Son olarak ESA'larda kullanılan sınıflandırıcı katmanına iletilir. Bu katman sınıflandırma tahmini yapar [38].



Şekil 8. ViT model yapısı [35].

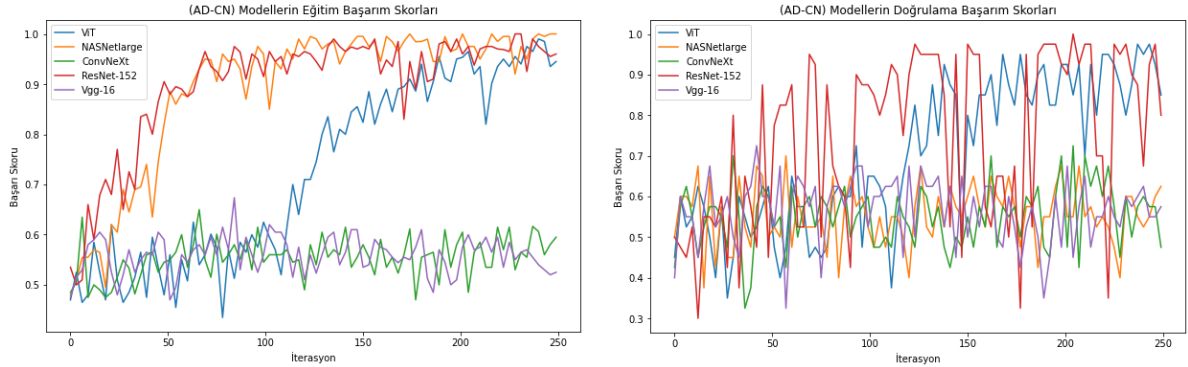
### 3. Bulgular ve Tartışma

Bu çalışmada, 3 boyutlu ADNI veri seti görüntülerine ait ekstenel kesitler alınarak yapay zeka modellerine verdik. 250 iterasyon, 8 mini-batch ve 220x220 boyutunda hiperparametre başlangıç ayarı olarak tanımlanmıştır. Veri dağılım oranları %70 eğitimi, %10 doğrulama ve %20 test olarak ayarlandı. Çalışma bulguları Alzheimer-Normal, Alzheimer-MCI ve MCI-Normal olmak üzere 3 adet ikili sınıfa ait sonuçlar alınmıştır. Az görüntüde ve 3 boyutlu MR kesitlerinden alınan ham görüntülerin başarımını artırmaktır. Tablo 3'te bu sınıflara ait Eğitim ve Doğrulama başarım sonuçları gösterilmektedir. Tablo 3 incelendiğinde NasNetLarge modeli eğitim en iyi sonuçları vermesine rağmen modellere verilen doğrulama görüntüleri verildiğinde, AD-CN ve AD-MCI ikili sınıfları incelendiğinde ViT modeli iyi sonuç verirken MCI-CN sınıfında ResNet-152 en iyi sonucu vermiştir.

**Tablo 3.** Modellere ait Eğitim ve Doğrulama Başarımları.

Model Adı	Eğitim			Doğrulama		
	AD-CN	AD-MCI	MCI-CN	AD-CN	AD-MCI	MCI-CN
<b>ViT</b>	%94,5	%85	%72	<b>%85</b>	<b>%87,5</b>	%82,5
<b>NasNetLarge</b>	<b>%100</b>	<b>%98</b>	<b>%98</b>	%62,5	%72,5	%70
<b>ConvNeXt</b>	%59,5	%69	%63,5	%47,5	%75	%60
<b>ResNet-152</b>	%95,9	%96,49	%93,99	%80	%82,49	<b>%92,5</b>
<b>Vgg-16</b>	%52,49	%72,5	%67,5	%57,49	%87,5	%62,5

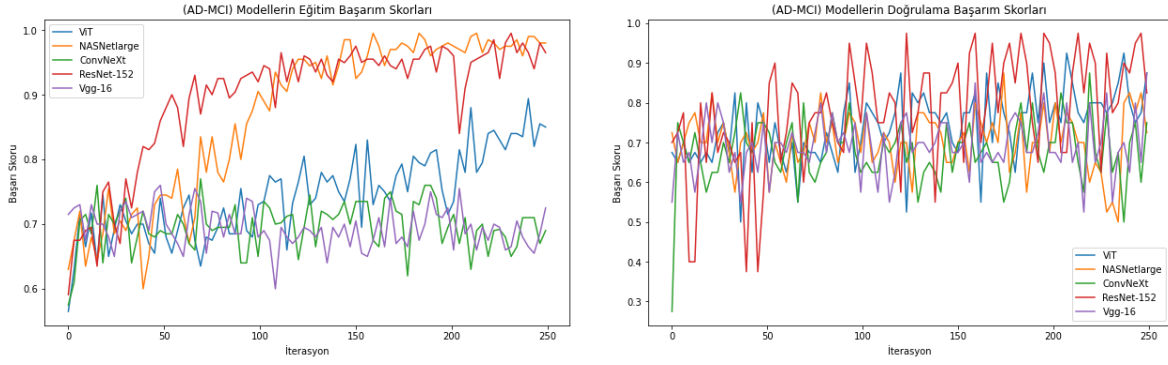
Şekil 9'da AD-CN sınıflarına ait modellerin eğitim ve doğrulama grafikleri verilmiştir. İki grafik incelendiğinde, NasNetLarge eğitimde %100 sonuç vermiş olsa da doğrulama da %62,5 başarımı düşük gözlemlenmiştir. Bu modelin ezber yaptığı ya da aşırı öğrenme yaptığı göstergesidir. Doğrulama grafiği incelendiğinde, ViT ve Resnet-152 modelleri daha iyi sonuçlar eğitime bağlı olarak iyi sonuçlar sergilemiştir. Doğrulama en iyi sonucu ViT modeli %85 vermiştir.



**Şekil 9.** (AD-CN) Eğitim ve Doğrulama Başarım Grafikleri.

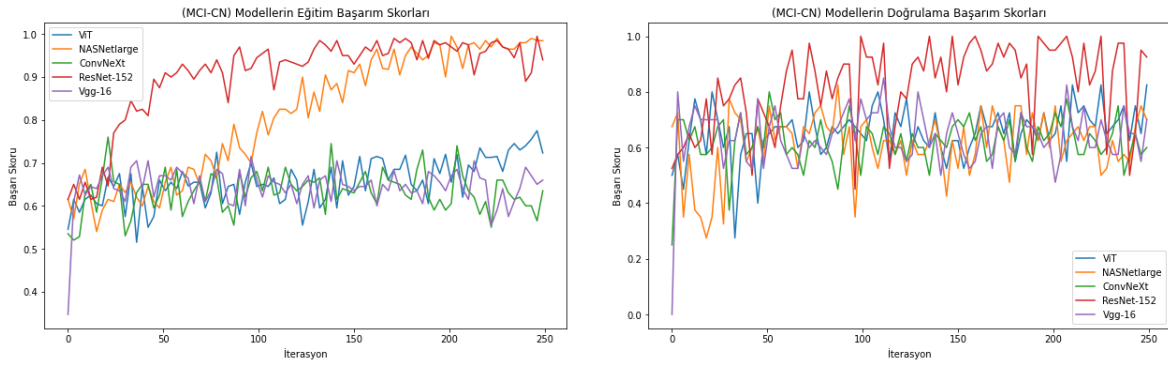
Şekil 10'da AD-MCI sınıflarına ait modellerin eğitim ve doğrulama grafikleri verilmiştir. İki grafik incelendiğinde, Şekil 9'daki gibi NasNetLarge %98 sonuç sergilemiş olsa da eğitimde, %72,5 doğrulama ile aşırı dalgalanma yaptığı gözlenmiştir. Bu dalgalanmalar, öğrenememe durumundan göstergesidir. Doğrulama en iyi sonucu ViT modeli %87,5 vermiştir.





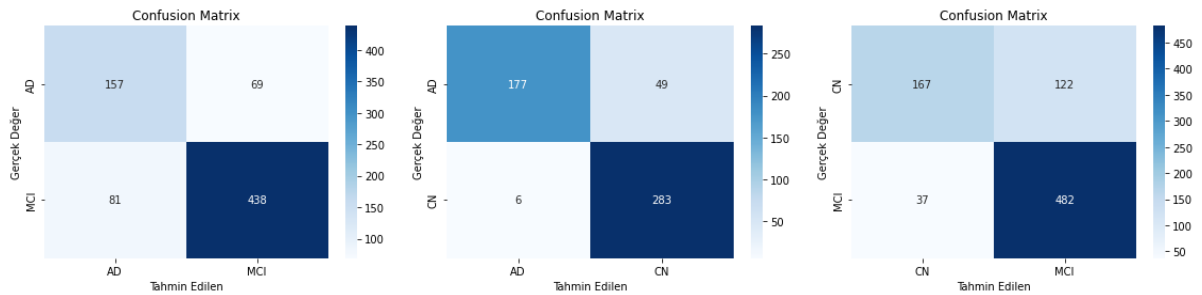
Şekil 10. (AD-MCI) Eğitim ve Doğrulama Başarım Grafikleri.

Şekil 11’de MCI-CN sınıflarına ait modellerin Eğitim ve Doğrulama grafikleri verilmiştir. İki grafik incelendiğinde, NasNetLarge %98 sonuç sergilemiş olsa da eğitimde verirken ResNet-152 modeli doğrulamada %92,5 başarımlar göstermiştir. Resnet-152’den sonra en iyi başarımları doğrulamada ViT modeli %82,5 vermiştir.



Şekil 11. (MCI-CN) Eğitim ve Doğrulama Başarım Grafikleri.

Aşağıda Şekil 12’de ViT modeline ait test sonuçlarına ait karmaşık matrisi gösterilmiştir. Bu karmaşık matrisleri sırasıyla; AD-MCI, AD-CN ve CN-MCI sınıflarına ait sonuçlardır. Diğer modellere ait karmaşık matrisi ölçütleri Tablo 3’te gösterilmektedir. MCI sınıfına ait sınıflandırmalarda modellerin tespitinin zor olduğu gözlemlenmiştir.



Şekil 12. ViT modeline ait karmaşık matrisi sonuçları (AD-MCI, AD-CN ve CN-MCI).

Tablo 4 incelendiğinde her sınıfa ait test başarımlarında ViT modeli diğer modellere göre daha iyi sonuç sergilediği gözlemlenmiştir. Özellikle Alzheimer sınıfına ait bazı tespitlerde daha iyi sonuçlar sergilemiştir. MCI sınıfın ait başarımlar incelendiğinde AD-MCI genel başarımlarında ViT iyi olmasına rağmen MCI başarımlar metrik değerleri Resnet-152 daha iyi değerler üretmiştir. Bu çalışmada, önerilen modelin başarısının değerlendirilmesi için Hassasiyet (Precision) ve F1 Skor performans kriterleri kullanılmıştır.

**Tablo 4.** Modellere ait Test Başarım Ölçütleri.

Model	Sınıf	AD-CN			AD-MCI			MCI-CN		
		Doğruluk	Hassasiyet	F1 Skor	Doğruluk	Hassasiyet	F1 Skor	Doğruluk	Hassasiyet	F1 Skor
ViT	Sınıf 1	<b>89,32</b>	78	87	<b>79,87</b>	69	68	<b>80,32</b>	58	68
	Sınıf 2	<b>89,32</b>	98	91	<b>79,87</b>	84	85	<b>80,32</b>	93	86
NasNetLarge	Sınıf 1	56,11	0	0	69,66	0	0	64,23	0	0
	Sınıf 2	56,11	100	72	69,66	100	72	64,23	100	78
ConvNeXt	Sınıf 1	56,11	0	0	69,66	0	0	64,23	0	0
	Sınıf 2	56,11	100	72	69,66	100	72	64,23	100	78
ResNet-152	Sınıf 1	87,21	99	87	74,22	24	30	78,46	58	66
	Sınıf 2	87,21	78	87	74,22	89	84	78,46	90	84
Vgg-16	Sınıf 1	56,11	0	0	69,66	0	0	64,23	0	0
	Sınıf 2	56,11	100	72	69,66	100	72	64,23	100	78

Bu çalışmada kullanılan derin öğrenme modellerinden elde edilen sonuçlar incelendiğinde, özellikle yüksek başarı sergileyen ViT modelinde herhangi bir aşırı uyum veya yetersiz uyum problemi ile karşılaşmadığı gözlemlenmiştir. ViT'ten sonra, özellikle eğitimde en iyi sonuçları NasNetLarge vermiş olsa da aşırı uyumda kaynaklı doğrulama ve teste iyi sonuçlar sergilememiştir. Bununla birlikte, özellikle MCI barındıran verilerde ResNet-152 mimarisi ayırt etmede daha iyi sonuçlar sergilemiştir. Ayrıca elde edilen eğitim doğruluklarının kanıtlanabilirliği adına, en başarılı test sonuçları sağlayan modellerin karmaşıklık matrislerine çalışmada yer verilmiştir.

#### 4. Sonuç

Alzheimer gibi nörolojik hastalıklar, insan yaşamını derinden etkileyen ve tedavi edilmesi henüz mümkün olmayan durumlardır. Bu hastalıkların etkisiyle beyinde meydana gelen değişiklikler, günlük yaşamı oldukça zorlaştırırken, bu durum hastaların yaşam kalitesini önemli ölçüde azaltabilmektedir. Dolayısıyla, bu tür hastalıkların erken teşhis edilmesi ve farkındalığın artırılması, hastalığın ilerlemesini kontrol etmek ve daha iyi bir yaşam kalitesi sağlamak için son derece önemlidir. Yapmış olduğumuz çalışmada az veri ve 3 boyutlu ham görüntü kesitleri kullanılarak özellikle Alzheimer hastalığının teşhisi için bilgisayar destekli modeller önerdik. Çalışma sonuçları, özellikle yeni bir yaklaşım olan evrişim katmanları kullanmayan ViT modelinde daha iyi sonuçlar verdiği gözlemlenmiştir.

Bu çalışmada ADNI veri seti kullanılmıştır. ADNI veri setinin kullanımı ve doğru kesit seçimi gibi parametreler, gelecek çalışmaların odak noktası olabilir ve bu faktörlerin etkisi üzerine yapılan araştırmaların, daha kapsamlı ve verimli sonuçlara ulaşmamıza katkı sağlayacağı düşünülmektedir.

#### Kaynaklar

- [1] Lopez JAS, Gonzalez HM, Leger GC. Alzheimer's disease. *Handb Clin Neurol* 2019; 167: 231-255.
- [2] Scheltens P, De Strooper B, Kivipelto M, Holstege H, Chetelat G, Teunissen CE, Cummings J, van der Flier WM. Alzheimer's disease. *The Lancet* 2021; 397(10284): 1577-1590.
- [3] Karikari TK, Pascoal TA, Ashton NJ, Janelidze S, Benedet AL, Rodriguez JL, Chamoun M, Savard M, Kang MS, Theriault J, Schöll M, Masson C, Soucy JP, Höglund K, Brinkmalm G, Mattsson-Carlgen N, Palmqvist S, Gauthier S, Stomrud E, Rosa-Neto P, Hansson O, Blennow K, Zetterberg H. Blood phospho-tau in Alzheimer disease: analysis, interpretation, and clinical utility. *Nat Rev Neurol* 2022; 18(7): 400-418.
- [4] Kivimäki M, Singh-Manoux A, Pentti J, Sabia S, Nyberg ST, Alfredsson L, Björner JB, Brunner EJ, Fransson EI, Goldberg M, Knutsson A, Koskenvuo M, Koskinen A, Kouvonen A, Kuula L, Oksanen T, Salo P, Shipley MJ, Stenholm S, Suominen S, Vahtera J, Väänänen A, Westerholm P, Zins M, Hamer M, Batty GD, Ferrie JE. Estimating Dementia Risk Using Multifactorial Prediction Models. *JAMA Network Open* 2023; 6(6): e2318132.
- [5] Bilal M, Iqbal HMN, Barceló D. Nanomaterials for the treatment and diagnosis of Alzheimer's disease: An overview. *NanoImpact* 2020; 20: 100251.
- [6] Scharre DW. Preclinical, prodromal, and dementia stages of Alzheimer's disease. *Pract Neurol* 2019; 15: 36-47.
- [7] Odusami M, Maskeliūnas R, Damaševičius R. Pixel-Level Fusion Approach with Vision Transformer for Early Detection of Alzheimer's Disease. *Electronics* 2023; 12(5): 1218.
- [8] Fathi S, Ahmadi M, Dehnad A. Early diagnosis of Alzheimer's disease based on deep learning: A systematic review. *Comput Biol Med* 2022; 146: 105634.
- [9] Sarraf S, Tofighi G, D'Souza AM, Phillips JM, Javanmardi M. OViTAD: Optimized vision transformer to predict various stages of Alzheimer's disease using resting-state fMRI and structural MRI data. *Brain Sci* 2023; 13(2): 260.
- [10] Henschel L, Conjeti S, Estrada S, Diers K, Fischl B, Reuter M. FastSurfer-a fast and accurate deep learning based neuroimaging pipeline. *NeuroImage* 2020; 219: 117012.

- [11] Puranik M, Khadidos A, Talukder A, Mohammed MA, Abbas Z, Alsadoon A, Ali S, Kannan A. Intelligent Alzheimer's detector using deep learning. In: 2018 Second International Conference on Intelligent Computing and Control Systems (ICICCS); 2018; IEEE.
- [12] Ilias L, Askounis D. Explainable identification of dementia from transcripts using transformer networks. *IEEE J Biomed Health Inf* 2022; 26(8): 4153-4164.
- [13] Jain V, Ha S, Jin K, Lee J, Jeong D. A novel AI-based system for detection and severity prediction of dementia using MRI. *IEEE Access* 2021; 9: 154324-154346.
- [14] Zhang X, Wang F, Chen X, Jiang T. An explainable 3D residual self-attention deep neural network for joint atrophy localization and Alzheimer's disease diagnosis using structural MRI. *IEEE J Biomed Health Inf* 2021; 26(11): 5289-5297.
- [15] Oktavian MW, Yudistira N, Ridok A. Classification of Alzheimer's Disease Using the Convolutional Neural Network (CNN) with Transfer Learning and Weighted Loss. *arXiv preprint arXiv:2207.01584*, 2022.
- [16] Ma H, Xu G, Wang J, Zhao Z, Liu Y, Liu X. Classification of Alzheimer's disease: application of a transfer learning deep Q-network method. *Eur J Neurosci* 2024.
- [17] Mujahid M, Khan MA, Hussain T, Ullah A, Shah JH, Naqvi SR, Balakrishnan V, Gwak J. An efficient ensemble approach for Alzheimer's disease detection using an adaptive synthetic technique and deep learning. *Diagnostics* 2023; 13(15): 2489.
- [18] Jack Jr CR, Bernstein MA, Fox NC, Thompson P, Alexander G, Harvey D, Borowski B, Britson PJ, Lainhart JE, Ward C, Dale AM. The Alzheimer's disease neuroimaging initiative (ADNI): MRI methods. *J Magn Reson Imaging* 2008; 27(4): 685-691.
- [19] Mueller SG, Weiner MW, Thal LJ, Petersen RC, Jack CR, Jagust W, Trojanowski JQ, Toga AW, Beckett L. The Alzheimer's disease neuroimaging initiative. *Neuroimaging Clin* 2005; 15(4): 869-877.
- [20] Petersen RC, Aisen PS, Beckett LA, Donohue M, Gamst AC, Harvey DJ, Jack CR, Jagust WJ, Shaw LM, Toga AW, Trojanowski JQ. Alzheimer's disease neuroimaging initiative (ADNI): clinical characterization. *Neurology* 2010; 74(3): 201-209.
- [21] Jiang J, Liu T, Xu X, Gao P, Fang Y. Deep learning based mild cognitive impairment diagnosis using structure MR images. *Neurosci Lett* 2020; 730: 134971.
- [22] Lian C, Liu M, Zhang J, Shen D. Attention-guided hybrid network for dementia diagnosis with structural MR images. *IEEE Trans Cybern* 2020; 52(4): 1992-2003.
- [23] Theekedath D, Sedamkar R. Detecting affect states using VGG16, ResNet50 and SE-ResNet50 networks. *SN Comput Sci* 2020; 1: 1-7.
- [24] Yüzgeç E, Muhammed T. Alzheimer ve Parkinson Hastalıklarının Derin Öğrenme Teknikleri Kullanılarak Sınıflandırılması. *Fırat Üniversitesi Mühendislik Bilimleri Dergisi* 2023; 35(2): 473-482.
- [25] Zoph B, Vasudevan V, Shlens J, Le QV. Learning transferable architectures for scalable image recognition. In: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*; 2018.
- [26] İlhan İ, Balı E, Karaköse M. An Improved DeepFake Detection Approach with NASNetLarge CNN. In: *2022 International Conference on Data Analytics for Business and Industry (ICDABI)*; 2022. IEEE.
- [27] Pa WP, Nwe TL. Improving Myanmar Image Caption Generation Using NASNetLarge and Bi-directional LSTM. In: *2023 IEEE Conference on Computer Applications (ICCA)*; 2023. IEEE.
- [28] Xu X, Li W, Duan Q. Transfer learning and SE-ResNet152 networks-based for small-scale unbalanced fish species identification. *Comput Electron Agric* 2021; 180: 105878.
- [29] He K, Zhang X, Ren S, Sun J. Deep residual learning for image recognition. In: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*; 2016.
- [30] Deepika D, Lakshmi AV. Identification of Breast Cancer Using RESNET152. In: *2022 Fourth International Conference on Cognitive Computing and Information Processing (CCIP)*; 2022. IEEE.
- [31] Woo S, Kim DK, Lim J, Ghaffar A, Jang B, Han S, Choi J, Park J. Convnext v2: Co-designing and scaling convnets with masked autoencoders. In: *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*; 2023.
- [32] Zhou J, Zhou W, Wei W, Zhu Y. YOLO-CIR: The network based on YOLO and ConvNeXt for infrared object detection. *Infrared Phys Technol* 2023; 131: 104703.
- [33] Zhang C, Feng C, Li Y, Lu C, Shi L, Wang H. Research on Rolling Bearing Fault Diagnosis Based on Digital Twin Data and Improved ConvNext. *Sensors* 2023; 23(11): 5334.
- [34] Vaswani A, Shazeer N, Parmar N, Uszkoreit J, Jones L, Gomez AN, Kaiser Ł, Polosukhin I. Attention is all you need. *Adv Neural Inf Process Syst* 2017; 30.
- [35] Dosovitskiy A, Beyer L, Kolesnikov A, Weissenborn D, Zhai X, Unterthiner T, Dehghani M, Minderer M, Heigold G, Gelly S, Uszkoreit J, Houlsby N. An image is worth 16x16 words: Transformers for image recognition at scale. *arXiv preprint arXiv:2010.11929*, 2020.
- [36] Touvron H, Cord M, Douze M, Massa F, Sablayrolles A, Jégou H. Training data-efficient image transformers & distillation through attention. In: *International Conference on Machine Learning*; 2021. PMLR.
- [37] Wang W, Xie E, Li X, Fan D, Song K, Liang D, Lu T, Luo P, Shao L. Pyramid vision transformer: A versatile backbone for dense prediction without convolutions. In: *Proceedings of the IEEE/CVF International Conference on Computer Vision*; 2021.
- [38] Chen CFR, Fan Q, Panda R. Crossvit: Cross-attention multi-scale vision transformer for image classification. In: *Proceedings of the IEEE/CVF International Conference on Computer Vision*; 2021.