

Evrişimsel Sinir Ağları Tabanlı MR Görüntülerinden Beyin Tümörü Tespiti

Kübra BOZOĞLAN¹ , Burakhan ÇUBUKÇU^{2*} 

Öz

Bu çalışmanın amacı Manyetik Rezonans (MR) görüntülerini kullanarak derin öğrenme yöntemleri ile beyin tümörlerini tespit etmek ve derin öğrenme yöntemlerinin başarımlarını kıyaslamaktır. Beyin tümörleri, günümüzde artış gösteren ölümcül bir hastalık tehdidi haline gelmiştir. Hızlı büyüme eğilimleri göz önüne alındığında, erken teşhis ve doğru tedavi, hastaların hayatta kalma şansını artırmak için oldukça önem arz etmektedir. MR görüntülerinin incelenmesi, bu teşhis sürecinin temelini oluşturmaktadır. Bu çalışmada, beyin MR görüntülerinden tümörleri otomatik olarak tespit eden ve sınıflandıran, uzmanlara yardımcı olabilecek yeni bir bilgisayar destekli sistem sunulmaktadır. Geliştirilen sistem, Evrişimsel Sinir Ağları (CNN) adı verilen derin öğrenme mimarisine dayanmaktadır. Çalışmada, farklı öğrenme aktarım mimarilerinden VGG16, ResNet50 ve DenseNet121 kullanılmıştır. Bu modeller Figshare, SARTAJ ve Br35H veri setlerinin birleşiminden oluşturulan bir veri seti üzerinde test edilerek kıyaslanmıştır. Elde edilen bulgular, VGG16 mimarisinin %99,05'lik doğruluk oranıyla en yüksek başarıyı yakaladığını göstermiştir. ResNet50 mimarisi ise %73,40'lık başarı oranıyla modeller arasında en düşük başarı gösteren model olmuştur. Bu bulgular ışığında, evrişimsel sinir ağları tabanlı otomatik tümör tespit sisteminin, beyin tümörlerinin erken teşhisinde ve tedavisinde önemli bir rol oynayabileceği öngörülmektedir. Sistemin, uzman radyologların iş yükünü hafifletmesi ve teşhis sürecinin daha hızlı ve doğru bir şekilde gerçekleşmesine katkıda bulunması beklenmektedir.

Anahtar Kelimeler: Beyin Tümörü, Derin Öğrenme, CNN, VGG16, ResNet50, DenseNet121.

Brain Tumor Detection from MR Images Based on Convolutional Neural Networks

Abstract

The purpose of this study is to detect brain tumors using deep learning methods on Magnetic Resonance (MR) images and to compare the performances of various deep learning models. Brain tumors are increasingly becoming a serious threat as a fatal disease. Given their rapid growth, early diagnosis and accurate treatment are essential for improving patient survival rates. Examining MR images is fundamental to this diagnostic process. This study introduces a novel computer-aided system that automatically detects and classifies tumors in brain MR images, assisting medical experts. The developed system is based on a deep learning architecture known as Convolutional Neural Networks (CNN). The study employed different transfer learning architectures, including VGG16, ResNet50, and DenseNet121. These models were tested and compared using a dataset created from the combination of the Figshare, SARTAJ, and Br35H datasets. The results showed that the VGG16 architecture achieved the highest accuracy rate at 99.05%. In contrast, the ResNet50 architecture had the lowest performance among the models with an accuracy rate of 73.40%. Based on these findings, it is anticipated that the CNN-based automatic tumor detection system can play a significant role in the early diagnosis and treatment of brain tumors. The system is expected to alleviate the workload of expert radiologists and contribute to a faster and more accurate diagnostic process.

Keywords: Brain Tumor, Deep Learning, CNN, VGG16, ResNet50, DenseNet121.

¹Bilecik Şeyh Edebali Üniversitesi, Bilgisayar Mühendisliği Bölümü, Bilecik, Türkiye, kubrabooglan@gmail.com

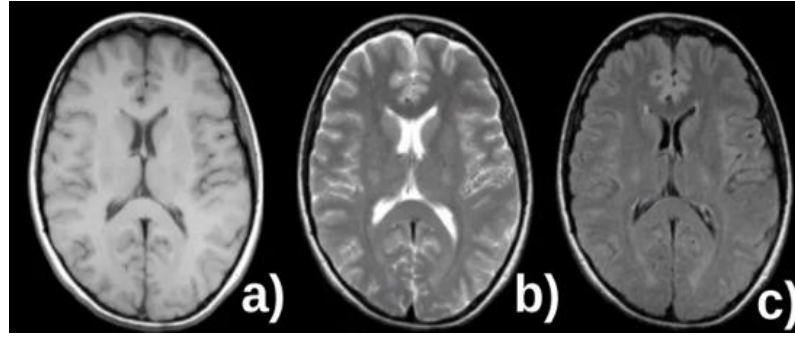
²Bilecik Şeyh Edebali Üniversitesi, Bilgisayar Mühendisliği Bölümü, Bilecik, Türkiye, burakhan.cubukcu@bilecik.edu.tr

*Sorumlu Yazar/Corresponding Author

1. Giriş

İnsan beyni, tıp dilinde “cerebrum” olarak adlandırılan ve kafatasının içerisinde bulunan, bir dizi karmaşık işlevi olan önemli bir organdır (Ersoy & Karal, 2012). Beyin tümörleri, normal dışı bir şekilde çoğalan beyin hücrelerinin oluşturduğu bir kütleyi temsil eder ve bu hücrelerin büyüklüğü, merkezi sinir sistemi üzerinde olumsuz etkilere neden olarak normal beyin fonksiyonlarının sapmasına ve bireyin sağlığının olumsuz etkilenmesine yol açabilir (Logeswari & Karnan, 2010). Beyin tümörleri birincil ve ikincil olmak üzere iki türde kayımıza çıkmaktadır. Genellikle iyi huylu olan birincil beyin tümörleri, beyinde doğrudan başlayan tümörlerdir. İkincil beyin tümörleri ise kanser hücrelerinin akciğer veya meme gibi vücudun başka bölgelerinden, beyine yayılması sonucunda oluşur (Tamilselvi vd., 2020). Bu tür tümörler genellikle zor teşhis edilir ve hayatta kalma şansı en düşük olan tümörlerden biridir (Taşci, 2022). Beyin tümörünün nedenini tam olarak söylemek oldukça zordur. Belirtiler tümörün türüne, konumuna ve büyüklüğüne bağlı olarak değişir ve baş ağrısı, bulantı, epileptik nöbetler gibi durumları içerebilir. Tedavi seçenekleri cerrahi müdahale, radyoterapi, kemoterapi ve ilaç tedavisini içerir ve tedavi planı hastanın durumuna bağlı olarak belirlenir (Hazra vd., 2017).

Tümör tespiti için ultrason, bilgisayarlı tomografi (BT) ve manyetik rezonans görüntüleme (MR) gibi görüntüleme yöntemleri kullanılmaktadır. MR görüntüsü, diğer tıbbi görüntüleme yöntemlerine kıyasla daha fazla ayrıntı sağladığı için tümör tespiti için sıkça tercih edilen bir yöntemdir. Özellikle beyin MR görüntüsü, tümör tespiti ve tümör ilerleme modelleme süreçlerinde kullanılır. Bu bilgiler, tümör tespiti ve tedavi süreçlerinde kritik öneme sahiptir. Beyin MR görüntüsü, beyin yapısındaki anormallikleri ve dokular arasındaki farklılıkları inceleyerek detaylı bilgiler sunar. Bu nedenle, beyin MR görüntüsü, tümörlerin tespiti ve tedavisi için önemli bir araçtır (Seetha & Raja, 2018a). Tümör, çeşitli biyolojik dokulardan oluştuğu için sadece tek bir görüntüleme sekansı kullanılarak yapılan beyin MR görüntüleri, tümör dokularıyla ilgili tüm bilgileri sağlayamayabilir. Bu nedenle, aynı bölgeye ait üç farklı sekansın görüntülenmesi için bir MRG kullanılır. Bu sekanslar T1 ağırlıklı, T2 ağırlıklı ve proton ağırlıklı görüntülerdir. Aksiyal düzleme ait bu üç farklı sekans görüntüsü, Şekil 1'de gösterilmiştir (Herek & Karabulut, 2010).



Şekil 1. Her bir sekansa ait MR görüntüleri a) T1 ağırlıklı b) T2 ağırlıklı c) Proton ağırlıklı

Beyin tümörleri hızlı bir şekilde büyüyebilir, normal boyutlarını iki katına çıkarabilir. Bu nedenle uzmanlar, Manyetik Rezonans (MR) görüntülerini dikkatlice ve hızlı bir şekilde inceleme sürecini gerçekleştirmelidir. Erken teşhis, kanser tanısı, tedavi planlaması ve tedavi sonuçlarının değerlendirilmesi açısından kritik öneme sahiptir. Eğer bir hasta, beyin tümörüne doğru ve erken tedavi alamazsa, hayatta kalma şansı azalabilir ve durum ölümcül olabilir. Bu sebeple beyin tümörünün hızlı ve doğru bir şekilde tespiti için bilgisayar destekli tespit yöntemleri uygulanmaya başlanmıştır. Literatürde, MR görüntülerinin kullanıldığı beyin tümörü tespiti ile ilgili birçok çalışma bulunmaktadır. Tablo 1’de MR görüntülerinden tümör tespiti yapılan çalışmalar verilmiştir.

Tablo 1. Literatürde MR görüntülerinden tümör tespiti yapılan çalışmalar.

Yazar (Yıl)	Veri Seti	Veri Ön işleme Artırma	Yöntem Model	Avantajlar	Dezavantajlar	Kullanılan Metrikler	Sonuçlar
Wu ve ark. (2014)	BRATS 2012	Süper piksel renk/doku ayrıştırması	CRF (Koşullu Rasgele Alanlar)	Hızlı, verimli, düşük bellek, gürültüye dayanıklı	Düşük dereceli tümörlerde yetersiz, karmaşık modelleme	Segmentasyon başarımları	Dice skoru düşük derecelilerde düşüktür
Karuppathal & Palanisamy (2014)	MR Görüntüleri	FIS için üyelik fonksiyonu tanımı	Fuzzy Inference System (FIS)	Hızlı, umut verici sınıflandırma aracı	Düşük doğruluk, karmaşık modelleme	Doğruluk	%88 doğruluk
Yaqub ve ark. (2014)	BRATS	-	Weighted Fast Local Random Forest	Yüksek doğruluk, farklı varyantlarla test	-	Doğruluk	%90 (WFLRF), %88 (LRF), %87 (WL-RF)
Soltaninejad ve ark. (2017)	19 MRI FLAIR, BRATS 2012	Süper piksel tekniği	ERT (Extremely Randomized Trees)	SVM’den daha iyi performans	-	Doğruluk	Dice: 0.85 (tümör), 0.79 (çekirdek), 0.75 (enhancing)
Huda ve ark. (2016)	BRATS	Hibrid özellik seçimi	GANNIG MAC, Decision Tree, Bagging C	Yüksek doğruluk, daha hızlı süre	-	Doğruluk, Çalışma süresi	%88 doğruluk

Gupta ve ark. (2020)	MR Görüntüleri	PCA ile boyut indirgeme	PCA + SVM	Basit, uygulanabilir	Düşük doğruluk	Doğruluk	%98 doğruluk
Citak-Er ve ark. (2018)	MR Görüntüleri	-	SVM, MLP, Lojistik Regresyon	Yüksek doğruluk	-	Doğruluk	%93.33 doğruluk, %93.33 duyarlılık, %96.6 özgüllük
Charfi ve ark. (2014)	MR Görüntüleri	Histogram eşitleme, PCA	Feedforward ANN	%90 doğruluk, basitlik	-	Doğruluk	%90 doğruluk
Pan ve ark. (2015)	Çok aşamalı MR Görüntüleri	-	CNN vs. geleneksel NN	CNN daha hassas	-	Hassasiyet, belirginlik	CNN, %18 daha iyi performans
Seetha & Raja (2018b)	BRATS	-	CNN + Fuzzy C-Means	Çok yüksek doğruluk	-	Doğruluk	%97.5 doğruluk
Afshar ve ark. (2019)	Figshare	-	Önerilen CNN modeli	Güvenilirlik	-	Doğruluk	%86.56 doğruluk
Shahzadi ve ark. (2019)	BRATS 2015	-	AlexNet-LSTM, VGGNet-LSTM	Zaman serisi bilgisi	AlexNet düşük doğruluk	Doğruluk	AlexNet: %71, VGG-LSTM: %84
Saxena ve ark. (2021)	BRATS	Transfer öğrenme	VGG16, InceptionV3, ResNet50	Yüksek doğruluk, güçlü modeller	-	Doğruluk	VGG16: %90, InceptionV3: %55, ResNet50: %95
Kaur & Gandhi (2019)	MR Görüntüleri	Transfer öğrenme	VGG16	%100 doğruluk, duyarlılık, özgüllük	İdealize sonuçlar	Duyarlılık, Özgüllük, Doğruluk	%94 doğruluk (optimum)
Zebari ve ark. (2023)	BRATS 2020	-	DenseNet121	Yüksek doğruluk ve hassasiyet	-	Doğruluk, Hassasiyet	%93.82 hassasiyet, %94.83 doğruluk

Beyin tümörlerini tespit etmek için Wu ve ark. (2014), MR görüntülerindeki tümörleri tespit etmek için süper piksellerin renk ve doku bilgilerini kullanan bir yöntem geliştirmiştir. Bu yöntem, CRF (koşullu rasgele alanlar) çerçevesinde çalışır ve hızlı, verimli, düşük bellek gereksinimi ve gürültüye karşı sağlamlık gibi avantajlara sahip olduğu ancak, düşük dereceli tümörlerin segmentasyonu için yeterince doğru olmadığı ve karmaşık bir modelleme sürecine sahip olduğu gözlemlenmiştir.

2014 yılında beyin tümörlerini segmentlemek ve sınıflandırmak için bulanık tabanlı kontrol teorisi kullanan bir çalışma yapılmıştır. Bu yöntemde, beyin tümörü segmentasyonu için özel bir teknik olan Fuzzy Interference System (FIS) kullanılmıştır. Denetimli sınıflandırma, FIS'in üyelik fonksiyonunu oluşturmak için kullanıldı. Yöntem yüksek performanslı ve hızlı olmasına rağmen,

doğruluk oranı düşüktü ve karmaşık bir modelleme sürecine sahipti. Bulanık tabanlı kontrol teorisinin beyin tümörlerini sınıflandırmak için umut verici bir yaklaşım sunduğu, ancak doğruluğunu iyileştirmek için daha fazla araştırmaya ihtiyaç duyulduğu sonucuna varılmıştır (Karuppathal & Palanisamy, 2014).

Yaqub ve ark. (2014), 2014 yılında yaptığı çalışmada, MR görüntülerini tümörlü ve tümörsüz olarak sınıflandırmak için farklı Rastgele Orman modelleri kullanılmıştır. En yüksek doğruluk, Ağırlıklı Hızlı Yerel Rasgele Ormanlar modeli ile %90 olarak elde edilmiştir. Ağırlıklı Lokal Rasgele Ormanlar modeli %87, Yerel Rasgele Ormanlar modeli ise %88 başarı elde etmiştir. Bu çalışma, Rastgele Orman sınıflandırıcısının MR görüntülerinde beyin tümörlerini yüksek doğrulukla sınıflandırabileceğini göstermiştir.

Soltaninejad ve ark. (2017), MRI görüntülerini otomatik olarak tanılamak ve sınıflandırmak için Süper Piksel Tekniği ve ERT sınıflandırıcısını kullanan bir yöntem geliştirmiştir. 19 MRI FLAIR ve BRATS 2012 veri seti üzerinde yapılan testler, ERT sınıflandırıcısının SVM'ye kıyasla daha yüksek doğruluk sağlayarak iyi bir performans gösterdiğini ortaya koymuştur. Bu çalışma, Süper Piksel Tekniği ve ERT sınıflandırıcısının beyin tümörlerini otomatik olarak tanımada ve sınıflandırmada kullanılabilecek potansiyele sahip olduğunu göstermektedir.

Huda ve ark. (2016), beyin tümörü teşhisinde daha yüksek doğruluk ve hızlı çalışma süresi sağlamak için hibrid özellik seçimi ve kolektif sınıflandırma içeren bir yöntem geliştirmiştir. GANNIGMAC, Decision Tree ve Bagging C algoritmalarını kullanan bu yöntem, beyin tümörlerini teşhis etmek için optimal bir özellik seti oluşturur ve farklı sınıflandırıcıların sonuçlarını birleştirerek daha yüksek doğruluk elde eder. Çalışmada elde edilen sonuçlar, hibrid yöntemin Decision Tree'den daha yüksek doğruluk ve GANNIGMAC'tan daha hızlı çalışma süresi sağlayarak beyin tümörü teşhisinde etkili bir araç olabileceğini göstermiştir. Bu çalışmalar dışında, Gupta ve ark. (2020), PCA ve SVM kullanarak tümör tespitinde %80 doğruluk oranına, Citak-Er ve ark. (2018), SVM, Çok Katmanlı Algılayıcılar ve Lojistik Regresyon kullanarak %93 doğruluk oranına ulaştığı çalışmalar bulunmaktadır. Bu çalışmalar dışında yapay sinir ağları kullanılarak yapılan çalışmalarda bulunmaktadır.

Charfi ve ark. (2014), beyin tümörlerini MR görüntülerinden teşhis etmek için histogram eşitleme ve PCA (Principal Component Analysis) ile görüntü segmentasyonu ve boyut azaltma, ardından beslemeli ileri yayılım sinir ağı ile sınıflandırma içeren bir yöntem geliştirmiştir. Yöntemin normal ve anormal beyin MR görüntülerini %90 doğrulukla ayırt edebildiği belirtilmiştir. Derin öğrenmenin görüntüler üzerinde kullanımının artmasıyla bu alanda evrimsel sinir ağları (CNN) kullanılarak yapılan çalışmalarda popüler hale gelmiştir.

Pan ve ark. (2015), beyin tümörlerini MR görüntülerinden tespit etmek amacıyla derin öğrenme ve temel sinir ağlarını karşılaştırmalı olarak incelemiştir. Çok aşamalı MR görüntüleri kullanılarak

yapılan çalışmada, CNN'lerin hassasiyet ve belirginlik açısından %18 daha iyi performans gösterdiği ortaya çıkmıştır. Bu da CNN'lerin beyin tümörlerini daha doğru ve net bir şekilde tespit edebildiğini göstermektedir. Sonuç olarak, çalışma CNN'lerin beyin tümörlerinin teşhisinde önemli bir rol oynayabileceğini ve temel sinir ağlarından daha etkili bir araç olduğunu göstermiştir. Seetha ve Raja, CNN ve Fuzzy C-Means algoritmalarını kullandıkları bir çalışmada %97,5 oranında doğruluk elde ederken (Seetha & Raja, 2018b), Afshar ve ark. (2019), önerdikleri CNN modeliyle %86,56 doğruluk oranına ulaşmıştır. Bunların dışında literatürde öğrenme aktarımı yöntemlerini kullanılarak da beyin tümörü tespiti yapan çalışmalar bulunmaktadır.

Shahzadi ve ark. (2019), AlexNet ve VGGNet CNN modellerini LSTM ile entegre ederek beyin tümörlerini tespit etmek için bir model geliştirmiştir. AlexNet-LSTM modeli %71 doğruluk, VGGNet-LSTM modeli ise %84 doğruluk elde etmiştir. Saxena ve ark. (2021), beyin tümörü verilerini sınıflandırmak için Vgg16, InceptionV3 ve Resnet50 modellerini kullandılar. Transfer öğrenme yöntemleriyle yaptıkları bu çalışmada, Resnet50 modelinde %95'lik en yüksek doğruluk oranını elde etmişlerdir. Kaur ve Gandhi. (2019), transfer öğrenmeli önceden eğitilmiş Vgg16 modeli kullanarak gerçekleştirdikleri çalışmada beyin tümörü tespiti için duyarlık, özgüllük ve doğruluk değerlerinin %100 olduğu sonucuna varmıştır. (Asaad Zebari ve ark. (2023), yaptıkları çalışmada beyin tümörü görüntüleri normal beyin görüntülerinden ayıran bir beyin tümörü sınıflandırma modeli geliştirmişlerdir. Densenet121 modelini kullanarak gerçekleştirdikleri bu çalışmada hassasiyet ve doğruluk değerlerini sırasıyla %93,82 ve %94,83 olarak ölçmüşlerdir

Bu çalışma, literatürde var olan derin öğrenme temelli beyin tümörü sınıflandırma yaklaşımlarına katkı sunmayı hedeflemektedir. Daha önce farklı araştırmalarda dört sınıflı sınıflandırma ve çeşitli performans metrikleri kullanılmış olsa da (Kaur & Gandhi, 2020; Gupta et al., 2023), bu çalışmada Figshare, SARTAJ ve Br35H gibi üç ayrı kaynaktan elde edilen görüntülerle oluşturulan Masoudnickparvar veri seti kullanılarak geniş, dengeli ve çeşitli bir veri kümesi üzerinde kapsamlı bir analiz gerçekleştirilmiştir (Masoudnickparvar , 2023). Aynı veri kümesi üzerinde hem klasik CNN hem de yaygın transfer öğrenme mimarileri (VGG16, ResNet50, DenseNet121) birlikte karşılaştırılmış ve bu modellerin doğruluk, precision, recall, F1 skoru gibi metrikler üzerinden detaylı analizleri yapılmıştır (Gupta et al., 2023). Bu bağlamda çalışmanın özgün katkısı; çeşitli kaynaklardan elde edilen verilerin bütünleştirilmesiyle oluşturulan veri kümesinin kullanımı, dört sınıflı yapıya sahip tümör tiplerinin birlikte değerlendirilmesi ve farklı mimarilerin performanslarının çok yönlü karşılaştırılmasıdır. Ayrıca, elde edilen bulgular klinik açıdan uygulanabilir düşük maliyetli modellerin de yüksek doğruluklar verebileceğini göstermesi bakımından literatürdeki çalışmalara tamamlayıcı bir bakış açısı sunmaktadır (Kaur & Gandhi, 2020).

2. Materyal ve Metot

Çalışmada Figshare, SARTAJ ve Br35H isimli üç farklı veri setinin birleştirilmesiyle elde edilen 7023 beyin MR görüntüsü üzerinde üç evrişim katmanlı CNN, VGG16, ResNet50, DenseNet121 modelleri ile glioma, meningioma, tümörsüz ve hipofiz olmak üzere 4 sınıflandırma yapılmıştır. Sınıflandırma, beyin tümörlerinin türlerini belirlemek için yapılmıştır.

Önerilen beyin tümörü tespit yöntemlerin her birinde, veri yükleme, veri arttırma, model oluşturma, model eğitimi ve model değerlendirmesi olmak üzere beş temel adımdan oluşmaktadır. Eğitim ve test veri setlerindeki görüntülerin boyutları 224x224 olarak ayarlanmıştır. Var olan görüntülere sınıf dağılımını dengelemek ve modelin genelleme kabiliyetini artırmak amacıyla ImageDataGenerator kullanılarak veri artırma yöntemleri (kaydırma, yakınlaştırma, çevirme, döndürme) uygulanmıştır. ImageDataGenerator sınıfı, doğrudan belirli sayıda yeni görüntü üretmekten ziyade, eğitim süreci boyunca gerçek zamanlı (on-the-fly) olarak veri artırımı uygular. Bu yöntemle, her epoch sırasında model, eğitim kümesindeki her görüntüyü bir kez görür; ancak bu görüntülere her seferinde rastgele dönüşümler uygulanır. Dolayısıyla, her epoch'ta model aynı sayıda görüntü ile eğitilse de, bu görüntülerin varyasyonları birbirinden farklıdır. Örneğin, 100 görüntü içeren bir veri kümesinde her epoch'ta yine 100 görüntü işlenir; ancak uygulanan rastgele dönüşümler sayesinde model her epoch'ta farklı biçimlerdeki verilerle karşılaşır. Epoch sayısı arttıkça, aynı görüntüler üzerinden sürekli yeni varyasyonlar üretilerek modelin genelleme kabiliyeti artırılır. Yani veri artırma ile veri setindeki fiziksel görüntü sayısı artırılmamış, disk üzerindeki görüntü sayısı sabit kalmıştır. Veri setinin sınıf bazlı sayıları Tablo 2'de verilmiştir.

Tablo 2. Veri setinin sınıf bazlı sayıları.

Sınıf	Eğitim Görüntüsü	Test Görüntüsü	Toplam Görüntü
Glioma	1.321	300	1.621
Meningioma	1.339	306	1.645
Hipofiz Tümörü	1.457	300	1.757
Sağlıklı	1.595	405	2.000
Toplam	5.712	1.311	7.023

Bu arttırma işlemi sadece eğitim sırasında, bellekte geçici olarak gerçekleştirilen dönüşümleri ifade etmektedir. Yani veri artırma ile veri setindeki fiziksel görüntü sayısı artırılmamış, disk üzerindeki görüntü sayısı sabit kalmıştır. Arttırılmış veriler üç evrişim katmanlı CNN, VGG16, ResNet50, DenseNet121 modelleri kullanılarak eğitilmiştir. Literatürde sınıflandırma görevleri için pek çok derin öğrenme modeli bulunmasına rağmen, VGG16, ResNet50 ve DenseNet121 modellerinin tercih edilmesinin temel nedeni, her birinin kendine özgü güçlü yönleri sahip olmasıdır. VGG16, basit ve derin yapısı sayesinde güçlü özellik çıkarımı sağlar ve transfer öğrenme için yaygın

olarak kullanılır. ResNet50, residual bağlantıları ile derin ağlarda gradyan kaybolma problemini çözerek daha verimli bir öğrenme süreci sunar, bu da modelin daha derin ağlarda bile yüksek doğrulukla çalışmasını sağlar. DenseNet121 ise, her katmanın önceki katmanlarla doğrudan bağlantı kurarak, parametrelerin daha verimli kullanılmasını sağlar ve daha iyi özellik temsilleri elde etmesine olanak tanır. Bu modellerin karşılaştırılması, farklı mimarilerin ve öğrenme stratejilerinin sınıflandırma görevindeki performans farklarını incelememize olanak tanımaktadır. Çalışma kapsamında model bazlı özgünlük, bu üç farklı mimarinin sunduğu güçlü yönlerin bir arada değerlendirilmesinde ve hangi modelin belirli veri setlerinde daha iyi sonuçlar verdiğinin analiz edilmesinde yatmaktadır.

Elde edilen sonuçlar doğruluk, hassasiyet, geri çağırma, F1 puanı AUC, Cohen's Kappa ve MCC metrikleri ile kıyaslanmıştır.

Modellerin uygulanması için 256 GB disk, macOS Monterey 12.0.1 işletim sistemi, 8 GB belleğe ve Apple M1 işlemciye sahip bir bilgisayar kullanılmıştır. Yazılım geliştirme süreci Python programlama dili ile gerçekleştirilmiştir. Derin öğrenme modelinin oluşturulması için TensorFlow framework'ü kullanılmıştır. Model mimarileri, Keras API aracılığıyla tasarlanmış ve eğitilmiştir. Çalışmada, transfer öğrenme yöntemine dayalı olarak ResNet50, DenseNet121, VGG16 gibi farklı önceden eğitilmiş modeller kullanılmıştır. Bu modellerin her biri, tensorflow.keras.applications modülünden yüklenmiş ve sınıflandırma problemine uyarlanmıştır.

Veri ön işleme ve veri artırma işlemleri için tensorflow.keras.preprocessing.image.ImageDataGenerator sınıfı kullanılmıştır. Eğitim sırasında veri artırma teknikleri, modelin genelleme yeteneğini artırmak amacıyla yatay çevirme, yakınlaştırma ve kırpma gibi yöntemleri içermektedir. Görüntülerin normalizasyonu (rescale) işlemi de ImageDataGenerator içinde gerçekleştirilmiştir.

Model değerlendirmelerinde kullanılan metrikler için scikit-learn kütüphanesi kullanılmıştır. classification_report, confusion_matrix ve ConfusionMatrixDisplay gibi fonksiyonlarla, doğruluk, precision, recall ve f1-score gibi metrikler hesaplanmıştır. Ayrıca modelin performansını görselleştirmek için Matplotlib kütüphanesinden faydalanılmıştır. NumPy ise sayısal hesaplamalar ve matris işlemleri için kullanılmıştır.

Çalışmada kullanılan modeller, görüntü sınıflandırma problemlerinde yaygın olarak kullanılan derin öğrenme mimarileridir ve her biri farklı yapısal özelliklere sahip olsa da transfer öğrenme yaklaşımıyla güçlü sonuçlar üretmektedir. Bu modeller, ImageNet veri seti üzerinde önceden eğitilmiş ağırlıklar kullanarak, belirli bir görevde sıfırdan eğitim yapmaya kıyasla daha hızlı ve daha verimli sonuçlar elde edilmesine olanak tanımaktadır.

Bu açık kaynaklı yazılım kütüphaneleri ve araçlar, derin öğrenme tabanlı görüntü sınıflandırma problemlerinde yaygın olarak kullanılan, doğruluğu ve güvenilirliği literatürle kanıtlanmış araçlardır.

Çalışmada kullanılan veri seti, orijinal haliyle %80 eğitim ve %20 test olarak önceden ayrılmış şekilde sunulmuştur. Bu ayırım, veri setinin kaynak çalışmasında tanımlanmış olup, çalışmamızda da bu yapı korunarak değerlendirme yapılmıştır. Böylece benzer çalışmalarla doğrudan karşılaştırma yapılması ve metodolojik tutarlılığın sağlanması hedeflenmiştir. Test veri seti model performansını değerlendirirken de kullanılmıştır.

Çalışmada kullanılan tüm derin öğrenme modellerinin eğitimi sırasında çeşitli hiperparametreler dikkatle seçilmiş ve deneysel olarak belirlenmiştir. Öğrenme oranı olarak 0.0001 değeri tercih edilmiştir. Bu düşük değer, transfer öğrenmede önceden eğitilmiş ağırlıkların aşırı güncellenmesini önlemek ve öğrenme sürecini daha stabil hale getirmek amacıyla seçilmiştir (Tan et al., 2018). Optimizasyon algoritması olarak ise Adam kullanılmıştır. Adaptif öğrenme oranı sağlayan Adam algoritması, literatürde derin öğrenme modellerinde başarılı sonuçlar vermesi ve hesaplama açısından verimli olması nedeniyle yaygın şekilde tercih edilmektedir (Kingma & Ba, 2015).

Modelin eğitim süresi 100 epoch olarak belirlenmiştir. Yapılan farklı denemeler sonucunda, eğitim başarımları ve süre dikkate alındığında bu değer ideal olduğu gözlemlenmiştir (Yamashita et al., 2018). Eğitim sırasında overfitting (aşırı öğrenme) riski göz önünde bulundurularak doğrulama kaybı izlenmiş, ancak erken durdurma mekanizması uygulanmamıştır. Eğitim sürecinde batch size değeri 32 olarak seçilmiştir. Bu değer, GPU bellek kapasitesi ve eğitim süresi gibi sınırlayıcı faktörler dikkate alınarak optimize edilmiştir (Li et al., 2020). Aşırı öğrenmeyi önlemek adına, tam bağlantılı katmanda %50 oranında dropout uygulanmıştır. Bu oran, küçük ve orta ölçekli veri setlerinde modelin genelleme yeteneğini artırmak amacıyla sıklıkla tercih edilmektedir (Srivastava et al., 2014).

Aktivasyon fonksiyonları olarak, ara katmanlarda doğrusal olmayanlığı sağlayan ReLU kullanılmış, çıkış katmanında ise çok sınıflı sınıflandırma problemine uygun olarak softmax tercih edilmiştir. Kayıp fonksiyonu olarak categorical crossentropy kullanılmıştır; çünkü bu fonksiyon, çok sınıflı sınıflandırma problemlerinde modelin tahmin ettiği olasılık dağılımları ile gerçek etiketler arasındaki farkı etkili biçimde yansıtmaktadır (Goodfellow et al., 2016). Hiperparametre değerlerinin tamamı, literatürdeki benzer çalışmalar temel alınarak ve küçük ölçekli ön testlerle elde edilen sonuçlara göre belirlenmiştir.

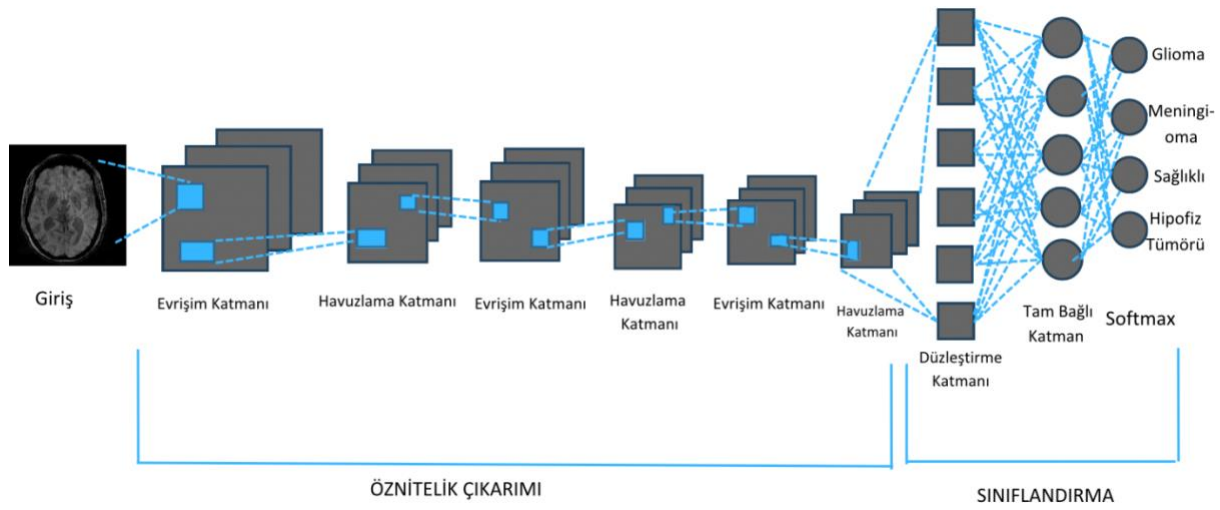
Bu bölümün devamında kullanılan derin öğrenme modellerinin detaylar, kullanılan veri seti ve modellerin kıyaslamasında kullanılan metriklere yer verilmektedir.

2.1. Üç Evrişim Katmanlı Evrişimsel Sinir Ağı (3kCNN)

Makine öğrenmesi, bir bilgisayarın belirli bir görevi yapabilmesi veya bir problemi çözebilmesi için, deneyim ve veri kullanarak kendiliğinden öğrenme yeteneğine sahip olduğu bir yapay zeka alt dalıdır. Geleneksel programlamadan farklı olarak, makine öğrenmesi algoritmaları, belirli bir görevi

doğrudan programlamak yerine veri üzerinden öğrenirler. Derin öğrenme (Deep Learning), yapay sinir ağları kullanarak karmaşık görevleri gerçekleştirmek için tasarlanmış bir makine öğrenmesi alt dalıdır. Derin öğrenme, çok katmanlı yapay sinir ağlarını kullanarak büyük miktarda veriyi öğrenmeyi ve genelleştirmeyi amaçlar. Temelde, derin öğrenme modelleri, birbirine bağlı çok sayıda sinir hücresinden oluşan katmanlardan oluşur. Bu modeller, öğrenme yeteneklerini geliştirmek ve daha karmaşık görevleri başarabilmek için genellikle birçok katmana sahiptir. Derin öğrenme, özellikle büyük veri setleri üzerinde yüksek başarı elde etme potansiyeline sahiptir (Cengil vd., 2017; Zhang vd., 1990). Aynı zamanda ses tanıma, doğal dil işleme ve diğer alanlarda da başarıyla uygulanabilir. Eğitim sürecinin uzun olması ve eğitim aşamasında yerel bir minimuma sıkışma riski taşıması bu modelin dezavantajları arasında sayılabilir (Habibi Aghdam vd., 2016).

Bu çalışmada derin öğrenme modellerinden evrimsel sinir ağları kullanılmıştır. Şekil 2’de gösterilen model 3 evrişim katmanı, 3 havuzlama katmanı, bir düzleştirme ve bir tam bağlı katmandan oluşmaktadır. Kullanılan model girdi olarak beyin MR görüntülerini kullanırken, çıktı olarak beyin tümörü türünü vermektedir.



Şekil 2. 3kCNN modelinin katman mimarisi

Çalışmada sunulan 3kCNN modelinin her bir katmanı için kullanılan parametreler Tablo 3’de verilmiştir.

Tablo 3. 3kCNN modelinin detayları.

Katman	Katman Türü	Filtre/Nöron Sayısı	Kernel Boyutu	Stride	Padding	Aktivasyon Fonksiyonu
1	Evrişim (Conv2D)	32	3x3	1	Same	ReLU
2	Havuzlama (MaxPooling2D)	-	2x2	2	Valid	-

3	Evrişim (Conv2D)	64	3x3	1	Same	ReLU
4	Havuzlama (MaxPooling2D)	-	2x2	2	Valid	-
5	Evrişim (Conv2D)	128	3x3	1	Same	ReLU
6	Havuzlama (MaxPooling2D)	-	2x2	2	Valid	-
7	Düzleştirme (Flatten)	-	-	-	-	-
8	Tam Bağlantılı (Dense)	256	-	-	-	ReLU
9	Dropout Katmanı	%50	-	-	-	-
10	Çıkış Katmanı (Dense)	4	-	-	-	Softmax

Önerilen 3kCNN modelinde, her bir evrişim katmanı için 3x3 kernel boyutu, stride=1 ve padding="same" tercih edilmiştir. 3x3 boyutundaki çekirdekler, küçük ve yerel özelliklerin (kenar, köşe, doku gibi) daha hassas şekilde yakalanmasını sağlar. Daha büyük kernel boyutlarına göre daha az parametre içerir, bu da modelin daha hızlı eğitilmesini ve aşırı öğrenmenin (overfitting) azaltılmasını destekler. Ayrıca, art arda birkaç 3x3 evrişim katmanı kullanılması, daha büyük alansal kapsamalara ulaşmayı mümkün kılar. Evrişim katmanlarında stride değerinin 1 olarak seçilmesi, giriş görüntüsünün çözünürlüğünü korurken daha detaylı özellik haritaları üretilmesini sağlar. Böylece tümör gibi küçük yapısal farklılıkların daha hassas şekilde tespit edilmesine imkan tanınır. "Same" padding kullanılması, giriş ile çıkış boyutlarının aynı kalmasını sağlar. Bu yöntem, kenar bilgilerini kaybetmeden görüntü üzerinde tam tarama yapılmasına olanak tanır. Özellikle küçük boyutlu tümör gibi önemli detayların kaybolmaması için padding uygulanması büyük önem taşımaktadır.

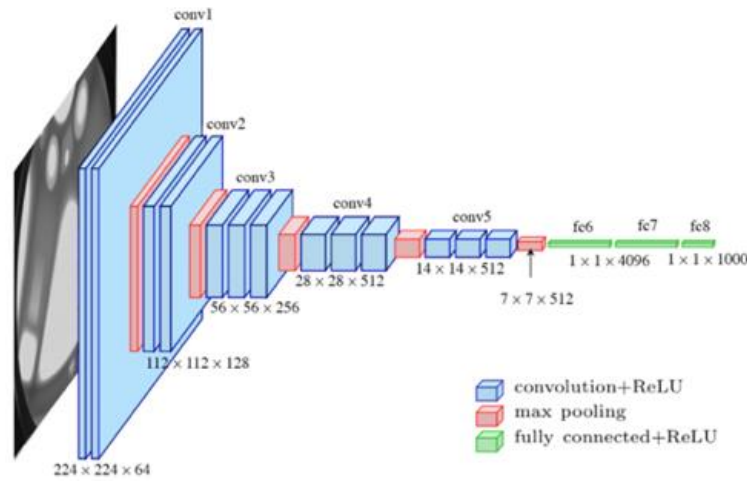
Havuzlama katmanlarında (MaxPooling2D) ise 2x2 kernel boyutu, stride=2 ve padding="valid" tercih edilmiştir. 2x2 boyutundaki havuzlama çekirdekleri, uzamsal boyutları yarıya indirerek modelin parametre sayısını ve hesaplama yükünü azaltır. Aynı zamanda öznitelik haritalarında en baskın özelliklerin korunmasına yardımcı olur. Stride değerinin 2 olması, her adımda 2 piksel ilerleyerek veri boyutunun hızlı ve kontrollü şekilde küçültülmesini sağlar. Bu da modelin daha verimli çalışmasına katkı sunar. Padding kullanılmaması ("valid"), sadece tam kapsanan bölgelerde havuzlama yapılmasını sağlar. Bu yöntem, evrişim katmanlarından sonra gereksiz kenar bilgilerini elimine ederek daha etkili bir öznitelik çıkartımı sağlar.

Bu seçimler, çalışmanın amacı doğrultusunda modelin hem eğitim süresini kısaltmak, hafif bir mimari sunmak hem de yüksek doğruluk oranı yakalamak için yapılmıştır. Sonuç olarak, önerilen 3kCNN modeli, düşük hesaplama maliyetine sahip olmasına rağmen beyin tümörü tespitinde üstün bir performans sergilemiştir.

Önerilen 3kCNN modeli, günümüzde sıklıkla kullanılan derin ve parametre sayısı yüksek transfer öğrenme tabanlı modellerin aksine, daha sade ve hafif bir yapıda tasarlanmıştır. Bu basit yapısına rağmen beyin MR görüntülerinde tümör tespiti probleminde oldukça yüksek doğruluk oranı elde etmiştir. Model, düşük hesaplama maliyeti, kısa eğitim süresi ve donanım ihtiyacı azlığı sayesinde, sınırlı kaynaklara sahip sistemlerde de etkili bir şekilde kullanılabilecek bir alternatif oluşturmaktadır. Bu yönüyle literatürdeki büyük ölçekli derin öğrenme mimarilerine kıyasla farklı ve özgün bir katkı sunmaktadır.

2.2. VGG16

VGG16, 16 katmana sahip bir evrişimli sinir ağıdır. Bu katmanlar, evrişimli (convolutional) katmanlar, tam bağlantılı (fully connected) katmanlar ve softmax sınıflandırma katmanını içerir. VGG16 modeli, özellikle 3x3 boyutundaki küçük filtrelerin ve ardışık olarak dizilmiş evrişimli katmanların kullanılmasıyla dikkat çeker. Bu tasarım, diğer modellerde kullanılan büyük filtrelerin aksine daha fazla parametre içerir, ancak genellikle daha etkili bir öğrenme sağlar. VGG16, genellikle görüntü sınıflandırma görevlerinde kullanılır ve ImageNet gibi büyük veri setlerinde eğitilmiştir (Qassim vd., 2018). VGG16'nın yapısı Şekil 3'te gösterilmiştir.

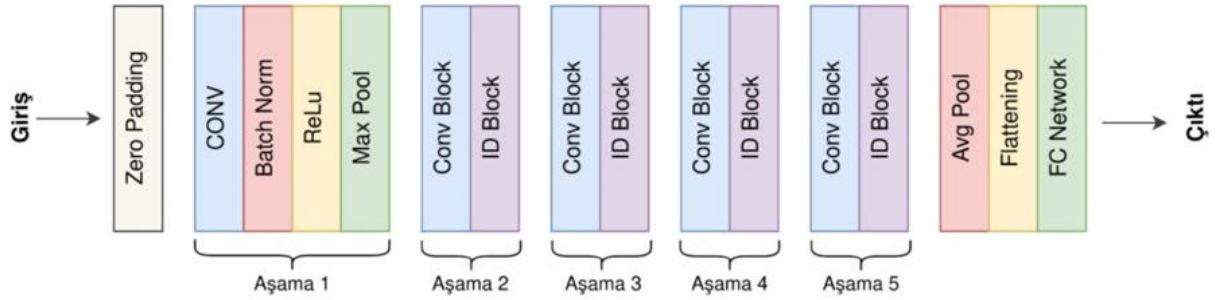


Şekil 3. VGG16 modelinin katman mimarisi

2.3. ResNet50

ResNet50, "Residual Network" kısaltmasının birleşimi olan ve evrişimli sinir ağlarının geliştirilmiş bir versiyonu olarak kabul edilen değerli bir nöral ağ mimarisidir. Geleneksel ardışık ağ mimarilerinden (örneğin, AlexNet, VggNet) farklı bir yapıya sahiptir. Temel özelliği, ağın

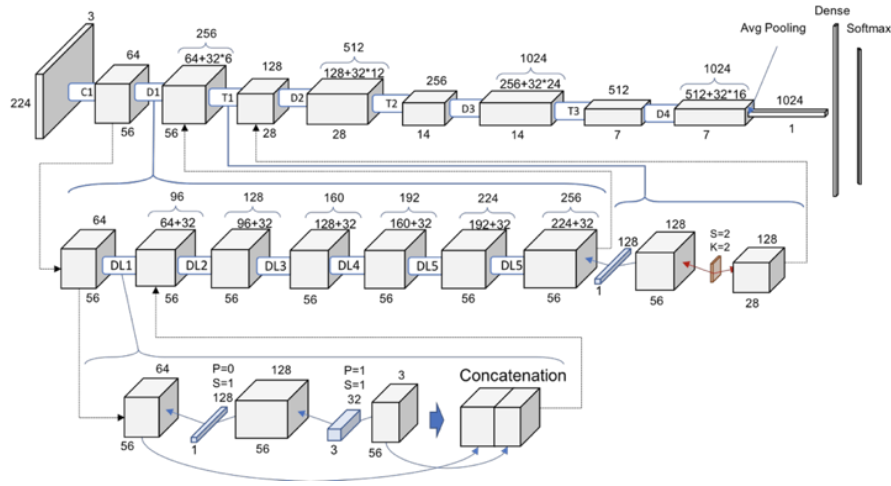
derinleştikçe ortaya çıkan performans düşüşünü ele almak için özgün bir bağlantı stratejisi kullanmasıdır. ResNet50 mimarisinde, 177 katmandan oluşan bir ağ bulunmaktadır. Bu ağ, katmanlar arasındaki geçişleri yönetmek ve daha iyi bir öğrenme sağlamak için "residual blocks" adı verilen özel yapıları içerir. Bu bloklar, bir önceki katmanın çıkışını doğrudan bir sonraki katmanın girişine ekleyerek bilgi akışını kolaylaştırır. Bu sayede, ağın daha derin olması ve daha fazla parametreye sahip olması mümkün olur. Giriş katmanı, ResNet50'de 224x224x3 boyutlarındadır, yani 224 piksel genişliğinde, 224 piksel yüksekliğinde ve 3 renk kanalına sahiptir. Bu giriş, ağın karmaşık özellikleri öğrenmeye başlamak için kullanılır. ResNet50 modeli, derin ağlar yakınsamaya başladığında ortaya çıkan performans düşüşü problemini azaltmayı hedefleyen bir yapıya sahiptir. Bu, residual bloklar aracılığıyla gerçekleştirilen katmanlar arası bağlantılar sayesinde sağlanır. Bu bağlantılar, ağın daha etkili öğrenmesine ve daha derin modellerin başarılı bir şekilde eğitilmesine olanak tanır (Gökalp & Aydın, 2021). ResNet50 yapısı Şekil 4'de gösterilmiştir.



Şekil 4. ResNet50 mimarisi

2.4. DenseNet121

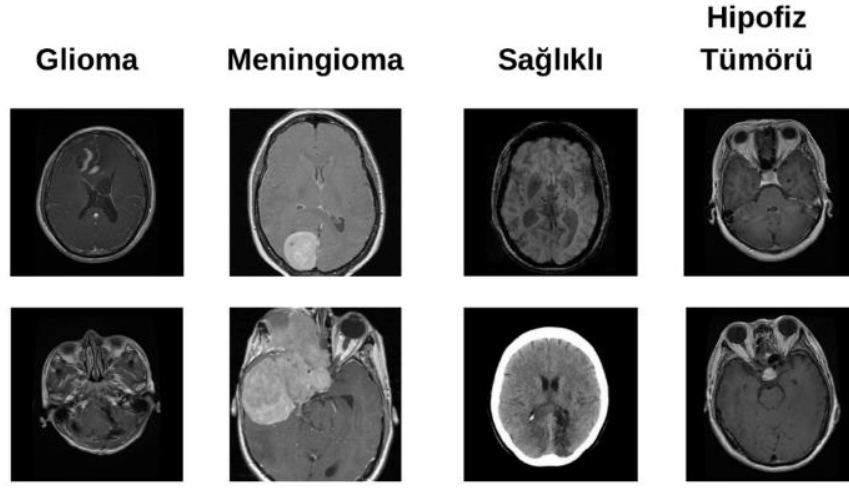
DenseNet121, bir derin öğrenme modelidir ve görüntü sınıflandırma gibi görevlerde kullanılır. Yoğun bağlantılara sahip bloklardan oluşan özel bir ağ mimarisini temsil eder. Her katman, giriş olarak kendi çıkışını ve önceki katmanlardan gelen tüm çıkışları alır, bu da bilgi akışını artırır. 121 katmanlı bir versiyonu olan DenseNet121, önceden eğitilmiş olarak kullanılabilir ve transfer öğrenme uygulamalarında başlangıç noktası olarak yaygın olarak kullanılır (Daşgün, 2023). DenseNet121 mimarisi Şekil 5'de gösterilmiştir.



Şekil 5. DenseNet121 mimarisi

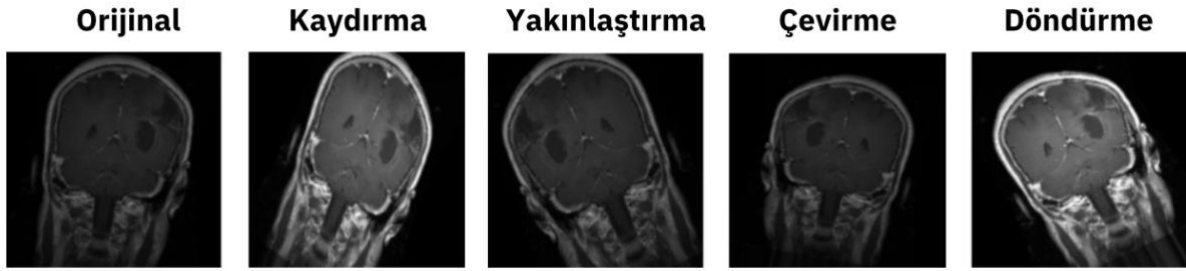
2.5. Veri Seti

Bu çalışmada Figshare, SARTAJ ve Br35H isimli üç farklı veri setinden oluşan Masoudnickparvar (Masoudnickparvar, 2023) veri seti kullanılmıştır. Bu veri setinin tercih edilmesinin temel nedeni, bu veri setinin sahip olduğu çeşitlilik, açıklık ve etiketleme kalitesidir. Literatürde birçok farklı beyin tümörü veri seti bulunmasına rağmen, bu veri setlerinin çoğu sınıf dengesizliği, düşük örnek sayısı veya yetersiz açıklayıcı etiketleme gibi sınırlılıklara sahiptir (Ali et al., 2020). Seçtiğimiz bu veri seti ise glioma, meningioma, hipofiz ve tümörsüz olmak üzere dört farklı sınıfı dengeli şekilde içermekte ve yüksek kaliteli MR görüntüler sunmaktadır. Ayrıca bu veri setinin açık erişimli olması, çalışmanın yeniden üretilebilirliğini kolaylaştırmakta ve literatürde sıklıkla referans gösterilen veri setleri olmaları nedeniyle sonuçların önceki çalışmalarla karşılaştırılmasına imkân tanımaktadır. Üç farklı veri setinin birleştirilmesiyle elde edilen 7023 görüntüden oluşan geniş veri kümesi, derin öğrenme modellerinin daha doğru genelleme yapmasına olanak sağlamıştır. Br35H veri setinin uzmanlar tarafından etiketlenmiş olması, Figshare ve SARTAJ veri setlerinin de farklı kaynaklardan alınan görüntülerle çeşitlilik sağlaması, çalışmamızda bu veri setini tercih etmemizde etkili olmuştur. Bu gerekçeler doğrultusunda, seçilen veri setinin model performansını artırmak, farklı tümör tiplerini kapsamak ve çalışmanın güvenilirliğini desteklemek açısından en uygun seçenek olduğu değerlendirilmiştir (Masoudnickparvar, 2023). Sınıflandırma, beyin tümörlerinin türlerini belirlemek için yapılmıştır. Farklı boyutlarda olan görüntülerin genişlik ve yükseklikleri 224x224 olarak ayarlanmıştır. 4 sınıfa ait görüntü örnekleri Şekil 6'da gösterilmiştir.

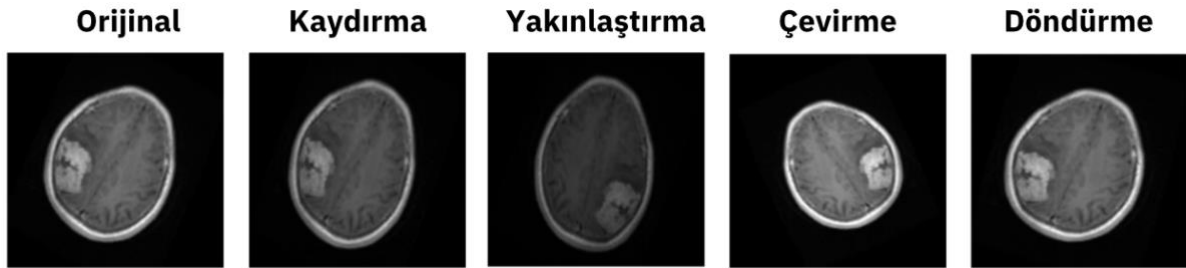


Şekil 6. Veri setindeki sınıflara ait örnek görüntüler

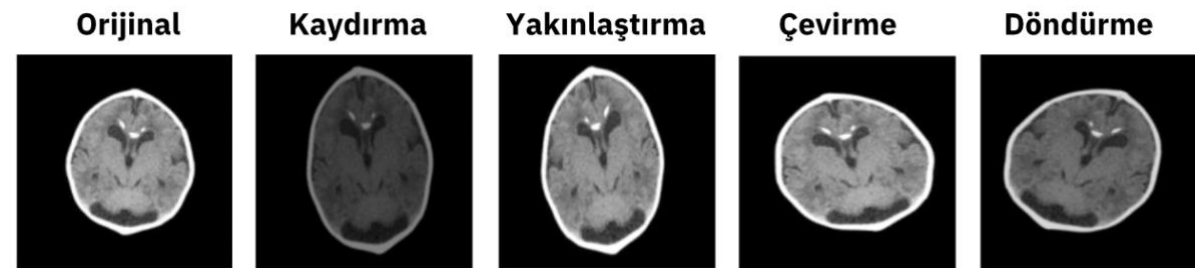
Veri setine uygulanan veri arttırma işlemi sonrasında oluşan her sınıfa ait görseller Şekil 7’de verilmiştir.



(a) Veri arttırma sonrası glioma sınıfına ait görseller



(b) Veri arttırma sonrası meningioma sınıfına ait görseller



(c) Veri arttırma sonrası sağlıklı sınıfına ait görseller



(d) Veri arttırma sonrası hipofiz tümörü sınıfına ait görseller

Şekil 7. Veri arttırma sonrası her bir sınıfa ait görseller

2.6. Performans Değerlendirme Kriterleri

Çalışmada kullanılan görüntü sınıflandırma modelinin değerlendirilmesi için doğruluk, hassasiyet, geri çağırma, F1 puanı metrikleri kullanılmıştır. Bu ölçütlerin hesaplanabilmesi için karmaşıklık matrisi $\begin{bmatrix} DN & YP \\ YN & DP \end{bmatrix}$ olacak şekilde Doğru Pozitif (DP), Doğru Negatif (DN), Yanlış Pozitif (YP), Yanlış Negatif (YN) verilerine ihtiyaç duyulmaktadır. Bu verilerden DP; gerçek pozitif örneklerin doğru bir şekilde pozitif olarak tahmin edilmesi, DN; gerçek negatif örneklerin doğru bir şekilde negatif olarak tahmin edilmesi, YP; gerçek negatif örneklerin yanlış bir şekilde pozitif olarak tahmin edilmesi, YN; gerçek pozitif örneklerin yanlış bir şekilde negatif olarak tahmin edilmesi olarak tanımlanmaktadır. Elde edilen bu değerler sayesinde doğruluk, hassasiyet, duyarlılık, F1-puanı Eşitlik 1-4’de gösterildiği şekilde hesaplanmaktadır.

$$\text{Doğruluk (Accuracy)} = \frac{DP + DN}{DP + DN + YP + YN} \quad (1)$$

$$\text{Hassasiyet (Precision)} = \frac{DP}{DP + YP} \quad (2)$$

$$\text{Geri Çağırma (Recall)} = \frac{DP}{DP + YN} \quad (3)$$

$$\text{F1 puanı (F1Score)} = 2 \times \frac{\text{hassasiyet} \times \text{geri çağırma}}{\text{hassasiyet} + \text{geri çağırma}} \quad (4)$$

Bu çalışmada ayrıca Sınıflandırma modelinin başarımını daha kapsamlı değerlendirmek amacıyla ROC eğrisi altında kalan alan (AUC), Cohen’s Kappa katsayısı ve Matthews Korelasyon Katsayısı (MCC) gibi metriklerden de yararlanılmıştır.

AUC (Area Under the ROC), modelin pozitif ve negatif sınıfları ayırt etme yeteneğini ölçer. AUC değeri 1’e yaklaştıkça sınıflandırma başarımı artar. ROC eğrisi üzerinden hesaplanır ve formülü genellikle integral üzerinden tanımlanır, ancak doğrudan formül genellikle verilmez çünkü kümülatif TPR (doğru pozitif oranı) ve FPR (yanlış pozitif oranı) değerlerine göre hesaplanır.

Cohen's Kappa (κ), rastlantısal doğruluk göz önüne alınarak sınıflandırma başarımını ölçer Eşitlik 5'de gösterildiği şekilde hesaplanmaktadır.

$$\kappa = \frac{p_o - p_e}{1 - p_e} \quad (5)$$

Eşitlik 5'de verilen p_o gözlenen doğruluk, p_e ise beklenen (rastlantısal) doğruluktur.

Matthews Korelasyon Katsayısı (MCC), ikili sınıflandırmalarda pozitif ve negatif sınıflar arasındaki korelasyonu değerlendirir ve Eşitlik 6'da gösterildiği şekilde hesaplanmaktadır.

$$MCC = \frac{((DP \times DN) - (YP \times YN))}{\sqrt{(DP + YP)(DP + YN)(DN + YP)(DN + YN)}} \quad (6)$$

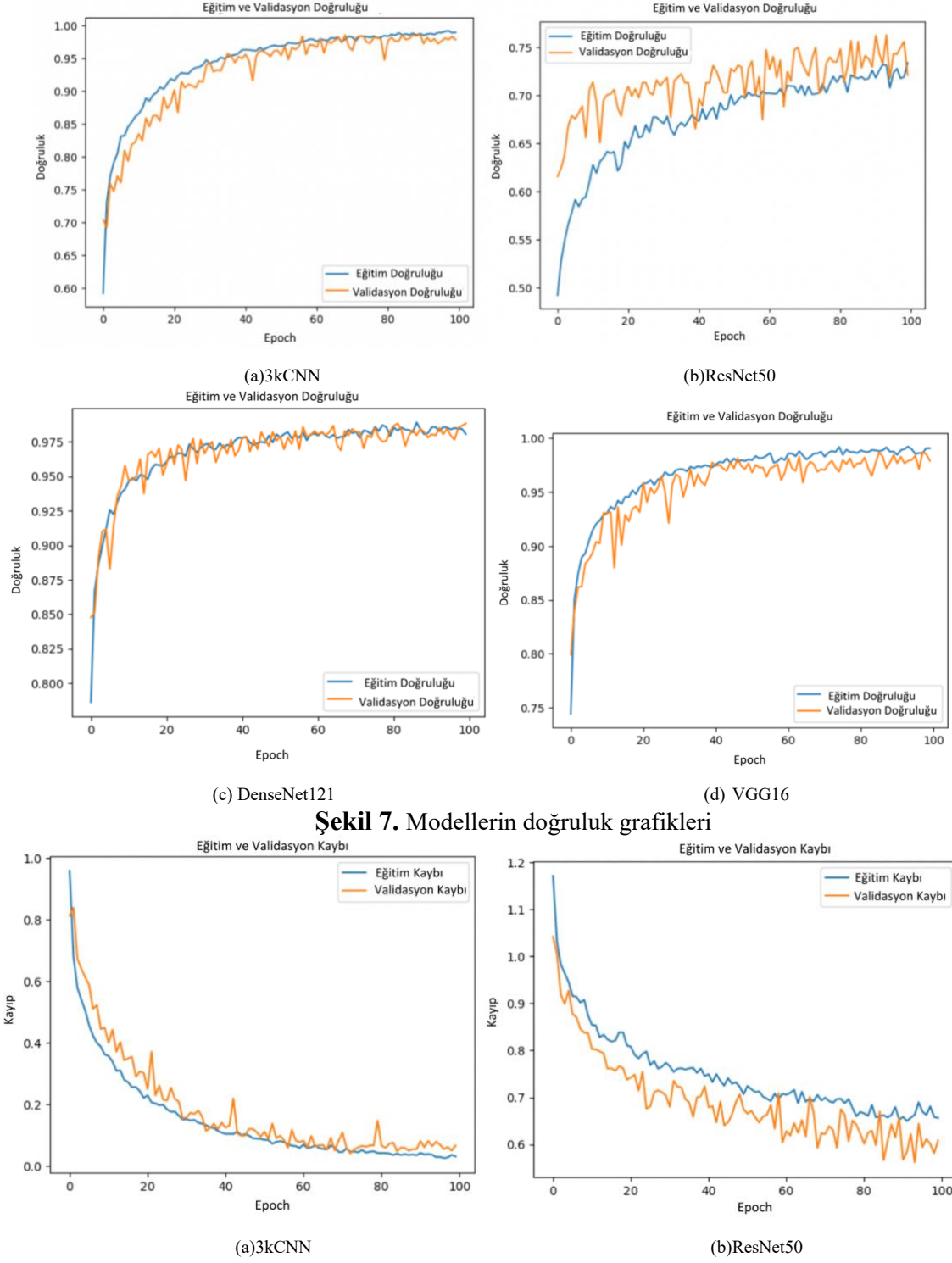
Bu metrikler, modelin genel performansını, sınıflar arası dengeyi ve hata durumlarını daha net şekilde yansıttığı için özellikle dengesiz veri setlerinde önem kazanmaktadır.

2.7. Transfer Öğrenme Stratejileri

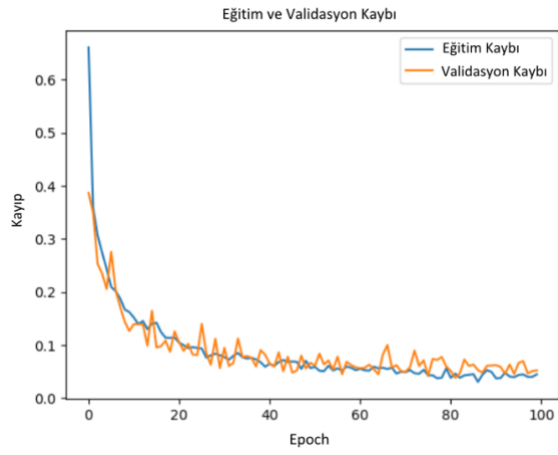
Bu çalışmada, önceden eğitilmiş derin öğrenme modelleri olan VGG16, ResNet50 ve DenseNet121 mimarileri üzerinde iki farklı transfer öğrenme stratejisi uygulanmıştır: feature extraction ve fine-tuning. Her iki strateji, modelin eğitim süreci ve çıktı üzerindeki etkilerini kıyaslayabilmek amacıyla ayrı ayrı değerlendirilmiştir. Özellikle fine-tuning stratejisi, modelin üst katmanlarının yeniden eğitilmesine olanak tanıyarak, hedef probleme daha uyumlu bir öğrenme gerçekleştirmeyi amaçlamaktadır. Literatürde de belirtildiği üzere, transfer öğrenme yaklaşımlarının uygulanması, modelin mevcut bilgilerini yeni görevler için kullanmasını sağlayarak hem eğitim süresini azaltmakta hem de başarıyı artırabilmektedir (Fırıldak & Talu, 2019). Bu kapsamda uygulanan stratejiler ile elde edilen sonuçlar, ilgili bölümlerde karşılaştırmalı olarak sunulmuştur.

3. Bulgular ve Tartışma

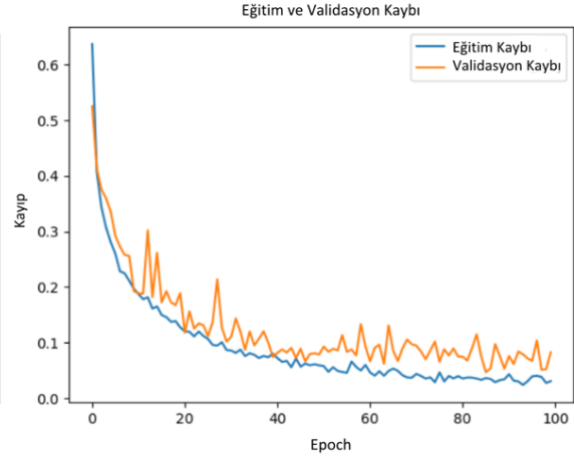
Çalışmanın bu bölümünde üç farklı veri setinden elde edilen beyin MR görüntülerden 3kCNN, VGG16, ResNet50, DenseNet121 modelleri ile glioma, meningioma, tümörsüz ve hipofiz olmak üzere yapılan dört sınıflandırma işlemi sonucunda elde edilen sonuçlar karşılaştırmalı olarak sunulmaktadır. Şekil 7'de sonuçların doğruluk-epoch grafiği, Şekil 8'de kayıp – epoch grafikleri, Şekil 9'da karmaşıklık matrisleri, Şekil 10'da 3kCNN modeline ait ROC eğrisi grafiği, Şekil 11'de transfer öğrenme modellerinin feature extraction aşamalarına ait ROC eğrisi grafikleri, Şekil 12'de ise transfer öğrenme modellerinin fine tuning aşamalarına ait ROC eğrisi grafikleri verilmiştir.



Şekil 7. Modellerin doğruluk grafikleri

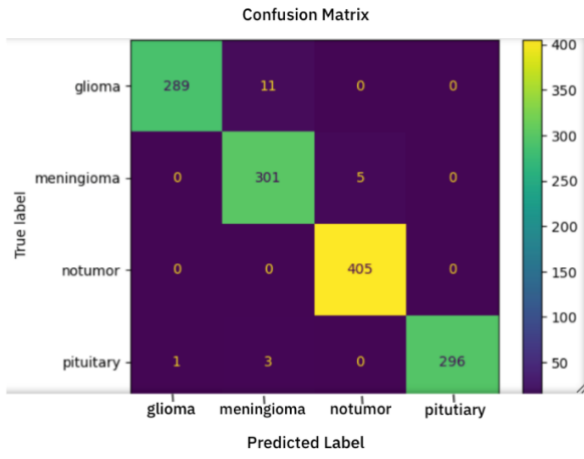


(c) DenseNet121

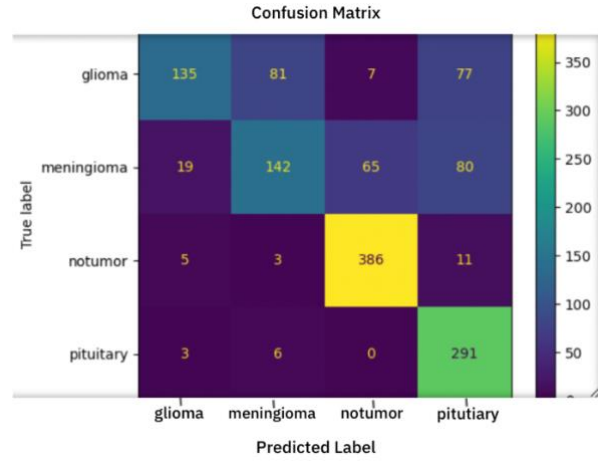


(d) VGG16

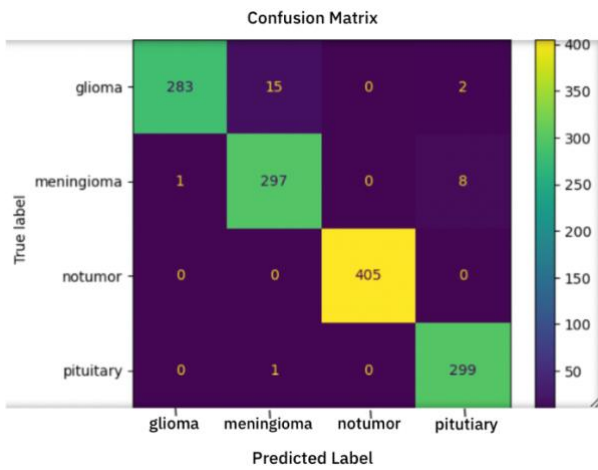
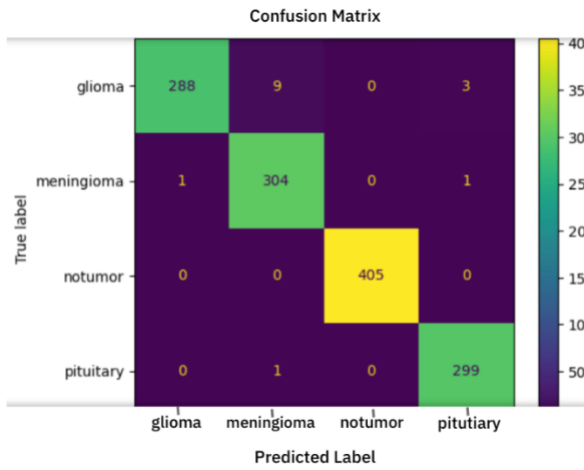
Şekil 8. Modellerin kayıp grafikleri



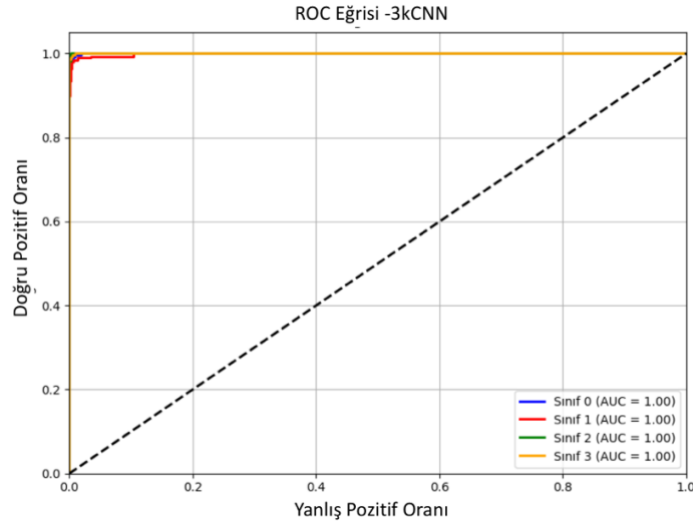
(a) 3kCNN



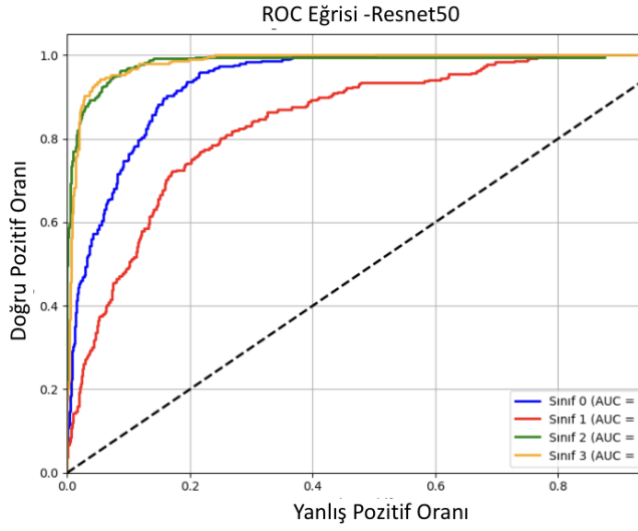
(b) ResNet50



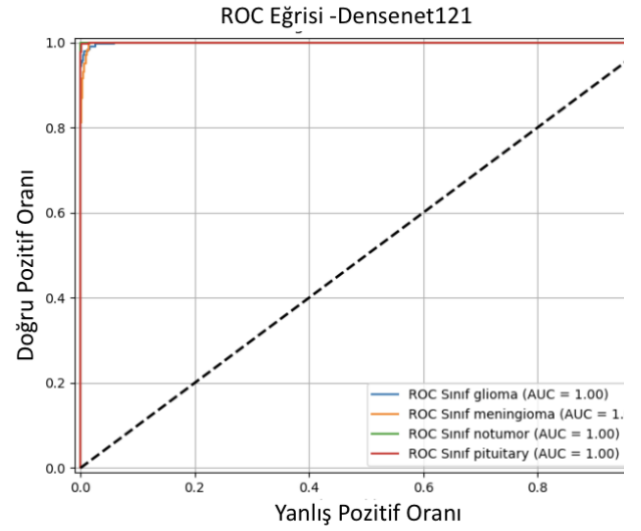
Şekil 9. Modellerin karmaşıklık matrisleri



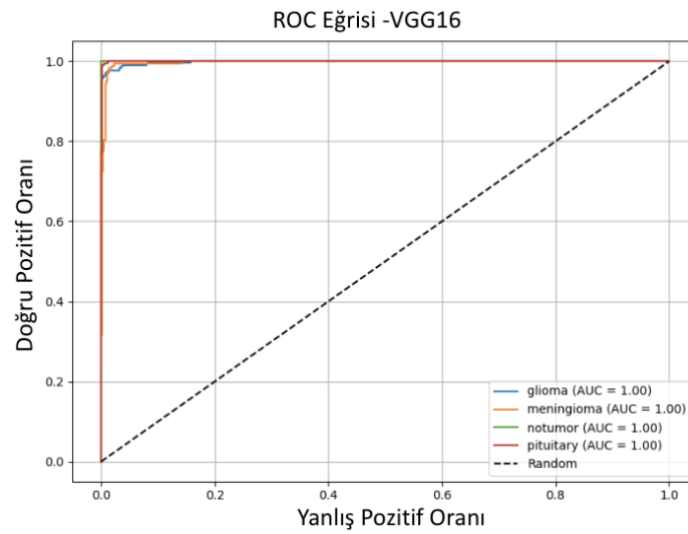
Şekil 10. 3kCNN modeline ait ROC eğrisi garfıđı.



(a)ResNet50

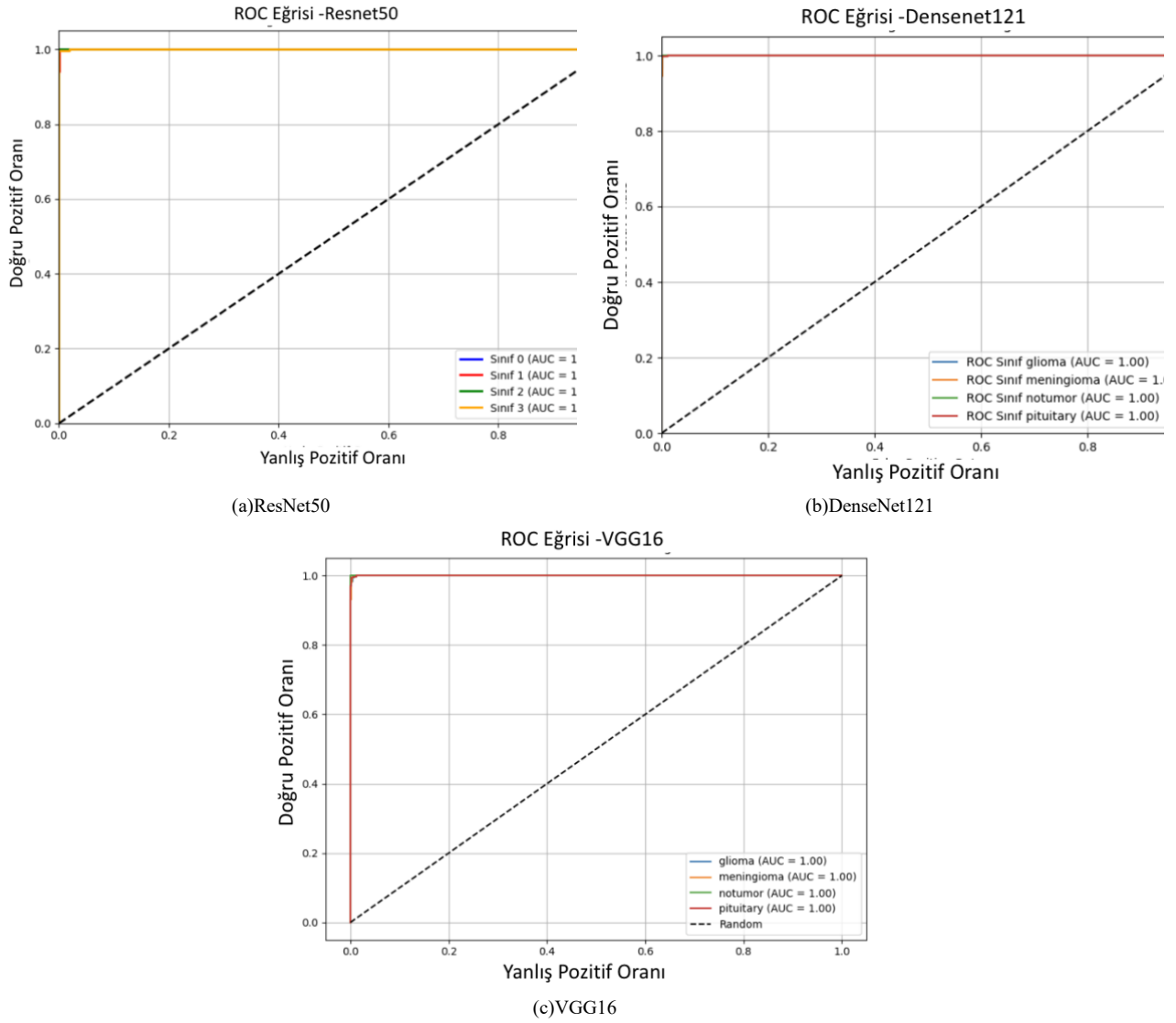


(b)DenseNet121



(c)VGG16

Şekil 11. Transfer öğrenme modellerinin feature extraction aşamalarına ait ROC eğrisi grafikleri.



Şekil 12. Transfer öğrenme modellerinin fine tuning aşamalarına ait ROC eğrisi grafikleri.

Beklendiği üzere tekrar sayısı arttıkça Şekil 7’de doğruluğun arttığı Şekil 8’de de kayıp oranının azaldığı gözlemlenmektedir. Her bir model için 100 epoch sonunda doğruluk artışının ve kaybın azalışının çok düşük seviyelere indiği görülmektedir. Bu sebepten ve eğitim süresini daha fazla uzatmamak adına epoch sayısının artırılmasına daha fazla gerek duyulmamıştır. Şekil 9’da verilen karmaşıklık matrislerinden gözlemlendiği üzere en az doğruluğu ResNet50 ederken diğer üç model daha başarılı sınıflandırmalar gerçekleştirmiştir. Daha iyi sınıflandırmalar gerçekleştiren bu üç modelde en çok glioma olması gereken sınıflar meningioma olarak tahmin edilmiştir. Bundan sonra göze çarpan en hatalı sınıflandırma 3kCNN modelinin 5 meningioma sınıfına ait görüntüyü tümör yok olarak tahmin etmesidir. Yine üç başarıyı yüksek modelin de tümör olmayan 405 görünün tamamını doğru tahmin etmesi de çalışmanın sonuçlarının umut verici olduğunu göstermektedir. Kullanılan tüm modellerinin başarımlarının daha iyi kıyaslanabilmesi için en çok kullanılan metrikler üzerinde kıyaslaması yapılmış ve Tablo 4’de sunulmaktadır.

Tablo 4. Kullanılan modellerin performans karşılaştırmaları

Model	Doğruluk	Kayıp	Hassasiyet	Geri Çağırma	F1 Skoru	AUC	Cohen's Kappa	MCC
CNN	0.9900	0.0314	0.9800	0.9794	0.9796	0.9995	0.9795	0.9796
Resnet50	0.7340	0.6563	0.7389	0.7276	0.7087	0.9370	0.6171	0.6348
Densenet121	0.9806	0.0445	0.9887	0.9885	0.9885	0.9997	0.9785	0.9786
VGG16	0.9905	0.0305	0.9798	0.9794	0.9795	0.9991	0.9724	0.9726

Tablo 4’de gösterilen elde edilen sonuçlara göre en düşük doğruluk oranını ResNet50 elde ederken en yüksek doğruluk oranını VGG16 modeli elde etmiştir. VGG16 modelini sırasıyla 3kCNN ve DenseNet121 takip etmektedir. Modellerin başarımı F1-Skorlarına göre sıralanması ise DenseNet121, 3kCNN, VGG16 ve ResNet50 olarak gerçekleşmiştir

Tablo 4’de sunulan genel model başarımlarına ek olarak, bu çalışmada transfer öğrenme kapsamında uygulanan feature extraction ve fine-tuning stratejilerinin etkileri de ayrı ayrı değerlendirilmiştir. Bu kapsamda elde edilen model sonuçları karşılaştırmalı olarak Tablo 5’de sunulmakta ve her iki stratejinin sınıflandırma başarısı üzerindeki etkisi detaylı biçimde ortaya konmaktadır.

Tablo 5. Modellerin farklı transfer öğrenme stratejilerine göre elde ettikleri başarı değerleri.

Model	Strateji	Doğruluk	Kayıp	AUC	Cohen's Kappa	MCC
Resnet50	Feature Extraction	0,7276	0,6025	0,9370	0,6171	0,6348
	Fine Tuning	0,9969	0,0107	1,0000	0,9980	0,9980
Densenet121	Feature Extraction	0,9805	0,0654	0,9997	0,9785	0,9786
	Fine Tuning	0,9935	0,0196	1,0000	0,9908	0,9908
VGG16	Feature Extraction	0,9823	0,0644	0,9991	0,9724	0,9726
	Fine Tuning	0,9896	0,0326	0,9999	0,9877	0,9877

Tablo 5’de yer alan sonuçlar incelendiğinde, fine-tuning stratejisinin üç modelde de feature extraction yöntemine kıyasla daha yüksek doğruluk ve daha düşük kayıp değerleri sağladığı görülmektedir. Özellikle ResNet50 modelinde fine-tuning sonrası doğruluk oranı %99,69’a ulaşırken, AUC değeri 1.0000 olarak hesaplanmıştır. Bu, fine-tuning stratejisinin modelin genelleme yeteneğini önemli ölçüde artırdığını göstermektedir. DenseNet121 ve VGG16 modellerinde ise her iki strateji de yüksek başarımlar göstermiş; ancak fine-tuning, Cohen’s Kappa ve MCC gibi istatistiksel uyum

ölçütlerinde küçük farklarla da olsa daha üstün performans sergilemiştir. Bu bulgular, transfer öğrenme uygulamalarında yalnızca önceden eğitilmiş ağırlıkların kullanılmasıyla yetinmenin her zaman yeterli olmayabileceğini; modelin belirli katmanlarının yeniden eğitilmesinin, hedef veri setine daha iyi adapte olunmasını sağladığını ortaya koymaktadır.

Elde edilen sonuçlara göre beyin MR görüntülerinden tümör tespiti yapmak için 3kCNN, DenseNet121 ve VGG16 modellerinin kullanılmasının pozitif sonuçlar doğuracağı değerlendirilmektedir. ResNet50 modelinin ise 0.7340 doğruluk oranıyla böylesi hassas bir konuda tahmin için kullanılmasının uygun olmadığı görüşüne varılmaktadır. Başarımı yüksek üç model arasında ise doğruluk anlamında çok ufak farklar bulunmaktadır. Bu sebepten eğitim süresi daha kısa olan basit bir yapıyla kurulan 3kCNN modelinin kullanılmasının diğer modellere göre süre açısından daha avantajlı olacağı değerlendirilmektedir.

4. Sonuçlar ve Öneriler

Beyin tümörü gibi önemli bir hastalığın erken ve doğru teşhisi hayati bir önem taşımaktadır. Bu sebeple bu çalışmada, beyin MR görüntüleri kullanılarak beyin tümörünün tespit edilmesi ve sınıflandırılması için derin öğrenme tabanlı yöntemler kullanılmıştır ve kıyaslanmıştır. Yapılan çalışmada geleneksel evrimsel sinir ağlarının yanında üç farklı öğrenme aktarım mimarisi test edilerek elde edilen sonuçlar hem görsel hem de sayısal olarak değerlendirilmiştir. Mimarilerin ortalama doğruluk başarımı üç katmanlı evrimsel sinir ağı için %99,00 olarak gözlemlenirken transfer öğrenme modellerinden olan ResNet50 için %73,40, DenseNet121 için %98,06 ve VGG16 için %99,05 olarak hesaplanmıştır. VGG16 mimarisinin ufak bir farkla kullanılan diğer mimarilere oranla daha başarılı sonuçlar ürettiğini görülmüştür. Önerilen DenseNet121, üç katmanlı evrimsel sinir ağı ve VGG16 modellerinin erken teşhis de faydası olacağı bunda hastanın tedavi ihtimalini arttıracakı öngörülmektedir. Bu çalışma sonucunda insan hayatını riske atan bir hastalığın erken teşhisinde faydalı olabilecek birden çok sistem kıyaslanmış ve sonuçlar umut verici olarak görülmüştür.

Yapılan bu çalışma, literatürde beyin MR görüntülerinden tümör tespiti üzerine yapılan güncel çalışmalarla kıyaslandığında hem yöntemsel çeşitlilik hem de elde edilen yüksek başarımlar açısından önemli katkılar sunmaktadır. Literatürde Shahzadi ve ark. (2019), VGGNet-LSTM ile %84 doğruluk elde ederken; Afshar ve ark. (2019) önerdikleri CNN modeliyle %86.56 doğruluğa ulaşmıştır. Saxena ve ark. (2021), ResNet50 modelinde %95 doğruluk oranı bildirmiştir. Bu çalışmada ise transfer öğrenme stratejileri kullanılarak DenseNet121 ile %98.06, VGG16 ile %99.05 doğruluk oranlarına ulaşılmıştır. Ayrıca, önerilen 3 katmanlı CNN modeli de %99 doğrulukla dikkat çekici bir performans göstermiştir.

Bu doğrultuda, yaygın transfer öğrenme mimarilerinin performansları doğrudan karşılaştırılarak, hangi modellerin tıbbi görüntü sınıflandırma görevlerinde daha uygun olabileceği değerlendirilmiştir. Bu durum, derin öğrenme temelli otomatik teşhis sistemlerinin etkinliğini artırma ve uzmanlara destek olabilecek karar destek sistemlerinin geliştirilmesine katkı sağlama açısından literatüre anlamlı bir yenilik sunmaktadır

Gelecek çalışmalarımızda farklı sınıflara ait beyin tümörlerini tespit edebilmek amacıyla daha fazla veri içeren veri kümeleri üzerinde yeni modeller oluşturulması hedeflenmektedir. Bunun yanı sıra kullanılan öğrenme aktarımı modelleri çeşitlendirilmesi farklı veri ön işleme yöntemlerinin denenmesi, farklı eğitim yöntemleri kullanılarak eğitim süresinin kısaltılması, hiperparametre seçimlerinin farklı optimizasyon yöntemleri ile elde edilmesi hedeflenmektedir. Özellikle çok sınıflı sınıflandırma problemlerinde modelin genellenebilirliğini daha güvenilir biçimde değerlendirmek için çapraz doğrulama (cross-validation) yöntemlerinin kullanımı oldukça önemlidir. Bu çalışmada sabit bir test seti ile değerlendirme yapılmış olsa da ilerideki çalışmalarımızda k-fold cross validation gibi yöntemlerin uygulanması planlanmaktadır. Ayrıca bundan sonraki çalışmalarda state-of-the-art derin öğrenme modelleri kullanılması da hedeflenmektedir.

Yazarların Katkısı

Tüm yazarlar çalışmaya eşit katkıda bulunmuştur.

Çıkar Çatışması Beyanı

Yazarlar arasında herhangi bir çıkar çatışması bulunmamaktadır.

Araştırma ve Yayın Etiği Beyanı

Yapılan çalışmada araştırma ve yayın etiğine uyulmuştur.

Kaynaklar

- Afshar, P., Plataniotis, K. N., & Mohammadi, A. (2019). Capsule Networks' Interpretability for Brain Tumor Classification Via Radiomics Analyses. *International Conference on Image Processing, ICIP, 2019-September*, 3816-3820. <https://doi.org/10.1109/ICIP.2019.8803615>
- Ali, H., Khan, M. A., Tariq, U., & Zhang, Y. (2020). Brain tumor classification using deep learning: A comprehensive review. *Journal of Healthcare Engineering, 2020*, Article 8897181.
- Asaad Zebari, N., A. H. Alkurdi, A., B. Marqas, R., & Shamal Salih, M. (2023). Enhancing Brain Tumor Classification with Data Augmentation and DenseNet121. *Academic Journal of Nawroz University, 12*(4), 323-334. <https://doi.org/10.25007/ajnu.v12n4a1985>

- Cengil, E., Çinar, A., & Güler, Z. (2017, Ekim 30). A GPU-based convolutional neural network approach for image classification. *IDAP 2017 - International Artificial Intelligence and Data Processing Symposium*. <https://doi.org/10.1109/IDAP.2017.8090194>
- Charfi, S., Lahmyed, R., & Rangarajan, L. (2014). A Novel Approach For Brain Tumor Detection Using Neural Network. *IMPACT: International Journal of Research in Engineering & Technology (IMPACT: IJRET)*, 2(7), 93-104. www.impactjournals.us
- Citak-Er, F., Firat, Z., Kovanlikaya, I., Ture, U., & Ozturk-Isik, E. (2018). Machine-learning in grading of gliomas based on multi-parametric magnetic resonance imaging at 3T. *Computers in Biology and Medicine*, 99, 154-160. <https://doi.org/10.1016/J.COMPBIOMED.2018.06.009>
- Daşgın, A. (2023). *Covid19 Yayılımını Azaltmak İçin Yüz Maskesinin Evrişimsel Sinir Ağı Modelleri ile Tespiti*. Aksaray Üniversitesi.
- Ersoy, E., & Karal, Ö. (2012). Yapay Sinir Ağları ve İnsan Beyni. *İnsan ve Toplum Bilimleri Araştırmaları Dergisi*, 1(2), 188-205.
- Fırıldak, K., & Talu, M. F. (2019). Evrişimsel sinir ağlarında kullanılan transfer öğrenme yaklaşımlarının incelenmesi. *Computer Science*, 4(2), 88-95.
- Goodfellow, I., Bengio, Y., & Courville, A. (2016). *Deep learning*. MIT Press.
- Gökalp, S., & Aydın, İ. (2021). Farklı Derin Sinir Ağı Modellerinin Duygu Tanımadaki Performanslarının Karşılaştırılması. *Muş Alparslan Üniversitesi Mühendislik-Mimarlık Fakültesi Dergisi*, 2(1), 35-43.
- Gupta, R., Koundal, D., & Goyal, N. (2023). Brain tumor classification using deep transfer learning with DenseNet-121. *Frontiers in Computational Neuroscience*, 17, 11719945.
- Gupta, T., Gandhi, T. K., Gupta, R. K., & Panigrahi, B. K. (2020). Classification of patients with tumor using MR FLAIR images. *Pattern Recognition Letters*, 139, 112-117. <https://doi.org/10.1016/J.PATREC.2017.10.037>
- Habibi Aghdam, H., Jahani Heravi, E., & Puig, D. (2016). A practical approach for detection and classification of traffic signs using Convolutional Neural Networks. *Robotics and Autonomous Systems*, 84, 97-112. <https://doi.org/10.1016/J.ROBOT.2016.07.003>
- Hazra, A., Dey, A., Gupta, S. K., & Ansari, A. (2017). Brain Tumor Detection Based on Segmentation using MATLAB. *International Conference on Energy, Communication, Data Analytics and Soft Computing (ICECDS-2017)*, 425-430.
- Herek, D., & Karabulut, N. (2010). Manyetik Rezonans Görüntüleme. *TTD Toraks Cerrahisi Bülteni*, 1(3), 214-222.
- Huda, S., Yearwood, J., Jelinek, H. F., Hassan, M. M., Fortino, G., & Buckland, M. (2016). A Hybrid Feature Selection with Ensemble Classification for Imbalanced Healthcare Data: A Case Study for Brain Tumor Diagnosis. *IEEE Access*, 4, 9145-9154. <https://doi.org/10.1109/ACCESS.2016.2647238>
- Karuppathal, R., & Palanisamy, V. (2014). Fuzzy Based Automatic Detection and Classification Approach For MRI-Brain Tumor. *ARPN Journal of Engineering and Applied Sciences*, 9(12), 2770-2779. www.arpnjournals.com
- Kaur, T., & Gandhi, T. K. (2020). Automated brain image classification based on VGG-16 and transfer learning. *Journal of Critical Reviews*, 7(8), 742-747.
- Kingma, D. P., & Ba, J. (2015). Adam: A method for stochastic optimization. *International Conference on Learning Representations (ICLR)*. <https://arxiv.org/abs/1412.6980>
- Li, Y., Yu, K., Zhang, Y., Li, X., & Yang, L. (2020). A review of training and evaluation strategies for deep learning in medical image analysis. *IEEE Access*, 8, 26122-26134.
- Logeswari, T., & Karnan, M. (2010). An Improved Implementation of Brain Tumor Detection Using Segmentation Based on Hierarchical Self Organizing Map. *International Journal of Computer Theory and Engineering*, 2(4), 591-595.
- Masoudnickparvar. (2023). *Brain Tumor MRI Dataset*. <https://www.kaggle.com/datasets/masoudnickparvar/brain-tumor-mri-dataset>
- Masoudnickparvar, A., & Subramaniam, M. (2022). Brain Tumor Classification Using Hybrid Deep Learning Techniques on Figshare, SARTAJ, and Br35H MRI Datasets. *Biomedical Signal Processing and Control*, 71, 103144.
- Pan, Y., Huang, W., Lin, Z., Zhu, W., Zhou, J., Wong, J., & Ding, Z. (2015). Brain tumor grading based on Neural Networks and Convolutional Neural Networks. *Proceedings of the Annual International Conference of the IEEE Engineering in Medicine and Biology Society, EMBS, 2015-November*, 699-702. <https://doi.org/10.1109/EMBC.2015.7318458>

- Qassim, H., Verma, A., & Feinzimer, D. (2018). Compressed Residual-VGG16 CNN Model for Big Data Places Image Recognition. *2018 IEEE 8th Annual Computing and Communication Workshop and Conference (CCWC)*, 169-175.
- Saxena, P., Maheshwari, A., & Maheshwari, S. (2021). Predictive Modeling of Brain Tumor: A Deep Learning Approach. İçinde *Advances in Intelligent Systems and Computing* (C. 1189, ss. 275-285). Springer, Singapore. https://doi.org/10.1007/978-981-15-6067-5_30
- Seetha, J., & Raja, S. S. (2018a). Brain tumor classification using Convolutional Neural Networks. *Biomedical and Pharmacology Journal*, 11(3), 1457-1461. <https://doi.org/10.13005/bpj/1511>
- Seetha, J., & Raja, S. S. (2018b). Brain tumor classification using Convolutional Neural Networks. *Biomedical and Pharmacology Journal*, 11(3), 1457-1461. <https://doi.org/10.13005/bpj/1511>
- Shahzadi, I., Meriadeau, F., Tang, T. B., & Quyyum, A. (2019). CNN-LSTM: Cascaded framework for brain tumour classification. *2018 IEEE EMBS Conference on Biomedical Engineering and Sciences, IECBES 2018 - Proceedings*, 633-637. <https://doi.org/10.1109/IECBES.2018.8626704>
- Soltaninejad, M., Yang, G., Lambrou, T., Allinson, N., Jones, T. L., Barrick, T. R., Howe, F. A., & Ye, X. (2017). Automated brain tumour detection and segmentation using superpixel-based extremely randomized trees in FLAIR MRI. *International Journal of Computer Assisted Radiology and Surgery*, 12(2), 183-203.
- Srivastava, N., Hinton, G., Krizhevsky, A., Sutskever, I., & Salakhutdinov, R. (2014). Dropout: A simple way to prevent neural networks from overfitting. *Journal of Machine Learning Research*, 15(1), 1929-1958.
- Tamilselvi, R., Nagaraj, A., Beham, M. P., & Sandhiya, M. B. (2020). BRAMSIT: A Database for Brain Tumor Diagnosis and Detection. *2020 Sixth International Conference on Bio Signals, Images, and Instrumentation (ICBSII)*, 1-5.
- Tan, M., Wang, R., Pang, R., Le, Q. V., & Fei-Fei, L. (2018). MnasNet: Platform-aware neural architecture search for mobile.
- Taşci, B. (2022). Beyin MR Görüntülerinden mRMR Tabanlı Beyin Tümörlerinin Sınıflandırması. *Journal of Scientific Reports-B*, 6, 1-9.
- Yamashita, R., Nishio, M., Do, R. K. G., & Togashi, K. (2018). Convolutional neural networks: An overview and application in radiology. *Insights into Imaging*, 9(4), 611-629.
- Wu, W., Chen, A. Y. C., Zhao, L., & Corso, J. J. (2014). Brain tumor detection and segmentation in a CRF (conditional random fields) framework with pixel-pairwise affinity and superpixel-level features. *International Journal of Computer Assisted Radiology and Surgery*, 9(2), 241-253. <https://doi.org/10.1007/S11548-013-0922-7/TABLES/4>
- Yaqub, M., Javaid, M. K., Cooper, C., & Noble, J. A. (2014). Investigation of the role of feature selection and weighted voting in random forests for 3-D volumetric segmentation. *IEEE Transactions on Medical Imaging*, 33(2), 258-271. <https://doi.org/10.1109/TMI.2013.2284025>
- Zhang, W., Itoh, K., Tanida, J., & Ichioka, Y. (1990). Parallel distributed processing model with local space-invariant interconnections and its optical architecture. *Appl. Opt.*, 29(32), 4790-4797.