T.C.
ÇANAKKALE ONSEKİZ MART ÜNİVERSİTESİ
LİSANSÜSTÜ EĞİTİM ENSTİTÜSÜ

BİLGİSAYAR MÜHENDİSLİĞİ ANABİLİM DALI

MEDİKAL GÖRÜNTÜLERDE DERİN ÖĞRENME YÖNTEMLERİNİN BAŞARIM DEĞERLENDİRMESİ

YÜKSEK LİSANS TEZİ

AYŞE MELİKE İDİNAK

Tez Danışmanı
DR.ÖĞR.ÜYESİ YONCA BAYRAKDAR YILMAZ

ÇANAKKALE – 2024

T.C.
ÇANAKKALE ONSEKİZ MART ÜNİVERSİTESİ
LİSANSÜSTÜ EĞİTİM ENSTİTÜSÜ

BİLGİSAYAR MÜHENDİSLİĞİ ANABİLİM DALI

MEDİKAL GÖRÜNTÜLERDE DERİN ÖĞRENME YÖNTEMLERİNİN BAŞARIM DEĞERLENDİRMESİ

YÜKSEK LİSANS TEZİ

AYŞE MELİKE İDİNAK

Tez Danışmanı
Dr. Öğr. Üyesi YONCA BAYRAKDAR YILMAZ

ÇANAKKALE – 2024

ETİK BEYAN

Çanakkale Onsekiz Mart Üniversitesi Lisansüstü Eğitim Enstitüsü Tez Yazım Kurallarına uygun olarak hazırladığım bu tez çalışmasında; tez içinde sunduğum verileri, bilgileri ve dokümanları akademik ve etik kurallar çerçevesinde elde ettiğimi, tüm bilgi, belge, değerlendirme ve sonuçları bilimsel etik ve ahlak kurallarına uygun olarak sunduğumu, tez çalışmasında yararlandığım eserlerin tümüne uygun atıfta bulunarak kaynak gösterdiğimi, kullanılan verilerde herhangi bir değişiklik yapmadığımı, bu tezde sunduğum çalışmanın özgün olduğunu, bildirir, aksi bir durumda aleyhime doğabilecek tüm hak kayıplarını kabullendiğimi taahhüt ve beyan ederim.

(İmza)
Ayşe Melike İDİNAK
29/01/2024

TEŞEKKÜR

Bu tezin gerçekleştirilmesinde, çalışmam boyunca benden bir an olsun yardımlarını esirgemeyen saygıdeğer danışman hocam Dr. Öğretim Üyesi Yonca BAYRAKDAR YILMAZ, çalışma süresince tüm zorlukları benimle göğüsleyen canım eşime, bana destek olan değerli aileme sonsuz teşekkürlerimi sunarım.”

Ayşe Melike İDİNAK

Çanakkale, Ocak 2024

ÖZET

MEDİKAL GÖRÜNTÜLERDE DERİN ÖĞRENME YÖNTEMLERİNİN BAŞARIM DEĞERLENDİRMESİ

Ayşe Melike İDİNAK
Çanakkale Onsekiz Mart Üniversitesi
Lisansüstü Eğitim Enstitüsü
Bilgisayar Mühendisliği Anabilim Dalı Yüksek Lisans Tezi
Danışman: Dr. Öğretim Üyesi Yonca BAYRAKDAR YILMAZ
29/01/2024, 55

Beyin tümörü, kontrolsüzce büyüyen beyin hücrelerine denir. Bu kontrolsüzce büyüyen anormal dokuların tespitinde derin öğrenmenin faydalı olabileceği iddia edilmiştir. Medikal görüntüler üzerinde derin öğrenme yöntemleri kullanılarak tümörün erken teşhis edilebileceği öngörülmüştür.

Bu tezde, hastalıkların erken teşhisini kolaylaştırmak için medikal görüntüler üzerinde yaygın olarak kullanılan derin öğrenme algoritmalarının başarımları karşılaştırılarak beyin tümörleri alanı üzerindeki başarımı yüksek algoritmanın tespit edilmesi amaçlanmıştır. Çalışmada BraTS dataseti kullanılmıştır. Medikal görüntüler üzerinde yaygın olarak kullanılan Vgg19, Resnet, Unet ve InceptionV3 öğrenme algoritmalarından model oluşturularak bu algoritmalar eğitilmiştir. Daha sonra eğitilen algoritmalara BraTS dataseti verilerek başarımları ölçülmüşür. Bu eğitim ve başarım ölçümü sırasında python, tensorflow ile yazılmış algoritma modelleri kullanılarak, Google Colab üzerinde eğitimleri ve başarım ölçümleri yapılmıştır. Elde edilen sonuçlar karşılaştırılarak başarımı yüksek algoritma tespit edilmiş ve nedenleri ortaya konmuştur. Başarım değerlendirmesi yapılırken doğruluk, dice benzerliği, f1 skor ve meanIoU metrikleri kullanılmıştır. Metriklerin karşılaştırılması sonucunda en performanslı modelin InceptionV3 olduğuna karar verilmiştir.

Anahtar Kelimeler: Derin Sinir Ağları, Beyin Tümörü, CNNs, RNNs, ResNet, Unet

ABSTRACT

PERFORMANCE EVALUATION OF DEEP LEARNING METHODS IN MEDICAL IMAGES

Ayşe Melike İDİNAK

Çanakkale Onsekiz Mart University

School of Graduate Studies

Master of Science Thesis in Computer Science

Advisor: Assistant Professor Yonca BAYRAKDAR YILMAZ

01/29/2024, 55

A brain tumor is a brain cell that grows uncontrollably. It has been claimed that deep learning can be useful in detecting these uncontrollably growing abnormal tissues. It is predicted that tumors can be detected early by using deep learning methods on medical images.

In this thesis, it is aimed to determine the algorithm with high performance in the field of brain tumors by comparing the performance of deep learning algorithms commonly used on medical images to facilitate early diagnosis of diseases. BraTS dataset was used in the study. Models were created from Vgg19, Resnet, Unet and InceptionV3 learning algorithms commonly used on medical images and these algorithms were trained. The trained algorithms were then given the BraTS dataset and their performance was measured. During this training and performance measurement, algorithm models written in python and tensorflow were used to train and measure their performance on Google Colab. By comparing the results obtained, the algorithm with high performance was identified and the reasons were revealed. Accuracy, dice similarity, f1 score and meanIoU metrics were used for performance evaluation. As a result of the comparison of the metrics, it was decided that InceptionV3 was the best performing model.

Keywords: Deep Neural Networks, Glioma, CNNs, RNNs, ResNet, Unet

İÇİNDEKİLER

Sayfa No

JÜRİ ONAY SAYFASI……..…….... i
ETİK BEYAN……... ii
TEŞEKKÜR……. iii
ÖZET ……...……..…… iv
ABSTRACT ……. v
İÇİNDEKİLER ……..…….. vi
SİMGELER ve KISALTMALAR…….....……... viii
TABLOLAR DİZİNİ…….... x
ŞEKİLLER DİZİNİ……...……...……..…… xi

BİRİNCİ BÖLÜM
GİRİŞ

İKİNCİ BÖLÜM
KURAMSAL ÇERÇEVE/ÖNCEKİ ÇALIŞMALAR 4

2.1. KURUMSAL ÇERÇEVE ....…….…..... ....……..... 4
2.1.1. Beyin Tümörü ...……...…...……...……...……...……. 4
2.1.2. Veri Madenciliği ve Makine Öğrenmesi ……...…...……...……..... 5
2.1.3. Derin Öğrenme ve Medikal Alanda Kullanılan Modeller ……...…...…… 6
2.2. ÖNCEKİ ÇALIŞMALAR ……...…...……...…...……...…...……...…...……. 7

ÜÇÜNCÜ BÖLÜM
ARAŞTIRMA YÖNTEMİ/MATERYAL YÖNTEM 25

3.1. MATERYAL ....…….....……...……...……...……...……...……. 25
3.2. YÖNTEM …….. 26
3.2.1. Veri Seti ...…….…….....…….…….....…….. 27
3.2.2. Derin Öğrenme Mimarileri ...…….…….....…….…… 27
3.2.3. Veri Ön İşleme ...…….…….....…….…….....…… 32
3.2.4. Model Tanımı ve Eğitimi ...…….…….....…….... 34

3.2.5. Model Tahmin ve Test 35
3.2.6. Değerlendirme 36

DÖRDÜNCÜ BÖLÜM
ARAŞTIRMA BULGULARI 37

4.1. Model Eğitim Metrikleri 37
4.2. Değerlendirme Metrikleri 40

BEŞİNCİ BÖLÜM
SONUÇ ve ÖNERİLER 47

5.1. Sonuç 47
5.2. Limit ve Sonraki Çalışmalar 49
5.3. Araştırma Yansımaları 50
KAYNAKÇA 51
ÖZGEÇMİŞ I

SİMGELER VE KISALTMALAR

CNN	Evrişimli Sinir Ağı
T1-Ce	Post-contrast T1-weighted
T1	Native T1
T2	T2-weighted
Flair	T2 Fluid Attenuated Inversion Recovery
AI	Yapay Zeka
U-Net	U Şeklinde Evrişimli Sinir Ağı
Vgg	Görsel Geometri Grubu / Visual Geometry Group
ResNet	Residual Network
DenseNet	Dense evrişimli sinir ağı mimarisi
GoogleNet	Paralel evrişimli sinir ağı
CapsNet	Kapsül ağ / Capsule Network
3D CNN	Üç boyutlu evrişimli sinir ağıdı
FCM	Fuzzy C-Means
SWT	The Stationary Wavelet Transform
ReLU	Rectified Linear Unit
AlexNet	Derin sinir ağı
EfficientNet	Derin sinir ağı
MFO	Multi-Fidelity Optimization
NLCMFO	Nonlinear Complexity Multi-Fidelity Optimization
SegCaps	Kapsül ağına dayalı bir görüntü segmentasyon modeli
DICE	Sørensen–Dice coefficient
IRNet	Intelligent Reflective Network
ImageNet	Veri kümesi
Sota	State of the Art (Belirli bir alandaki en iyi performansı ifade eder.)
ROI	Region of Interest (Bir görüntüdeki belirli bir bölgeyi belirtir.)
SCU-Ne	Self-Calibrated Unsupervised Neural Networks (Öğrenme modelidir.)
HDC	Hybrid dilated convolution
GPU	Graphics Processing Unit (Bilgisayar donanımı)
TPU	Tensor Processing Unit (Yapay zeka için tasarlanmış Google Donanımı)

Conv.	Evrişimli
Acc.	Doğruluk
Mean iou	Ortalama Üzerinde Kesişim
TP	Gerçek Pozitif
TN	Gerçek Negatif
FN	Yanlış Negatif
FP	Yanlış Pozitif

TABLOLAR DİZİNİ

Tablo No	Tablo Adı	Sayfa No
Tablo 1	MRI sınıflandırması ve tespiti ile ilgili çalışmaların karşılaştırması	18
Tablo 2	Modellerin test verisi üzerindeki değerlendirme sonucu doğruluk ve kayıp fonksiyon metrikleri	38
Tablo 3	Modellerin test verisi üzerindeki değerlendirme sonucu metrikleri	45
Tablo 4	Karşılaştırmasını yaptığımız modellerin diğer çalışmalardaki metrik değerleridir.	45

ŞEKİLLER DİZİNİ

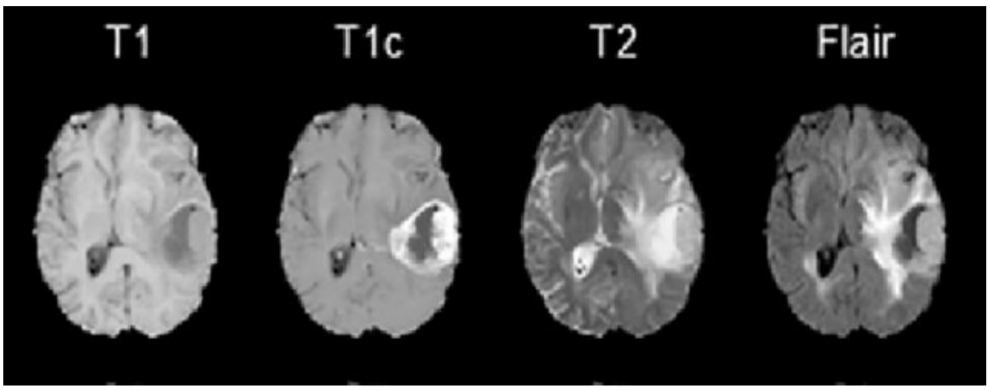
Şekil No	Şekil Adı	Sayfa No
Şekil 1	Beyin MRI'ına ait görüntü türler	2
Şekil 2	Çalışmanın takriben akış şeması	26
Şekil 3	Artık bağlantı	29
Şekil 4	U-net mimarisidir	30
Şekil 5	InceptionV3 modelinin mimarisi	32
Şekil 6	Flair, T1, T1CE, T2 ve Segmentation sınıflarına ait image görüntüleri	33
Şekil 7	MRI ın 3 farklı düzlem görüntüsü	33
Şekil 8	Karmaşık matrisinin temel yapısıdır	41

BİRİNCİ BÖLÜM
GİRİŞ

İnsan zekası ve biyomedikal ilerlemeye rağmen kanser, istikrarsız doğası nedeniyle çözülemeyen hala hastalıklardandır. En çok büyüyen ve ölümcül olan kanser türlerinden biri beyin tümörü kanseridir. Beyin ve Sinir Cerrahisi hastalıklarının önemli bir kısmını beyin tümörü oluşturmaktadır. Beyin tümörü; beyinde kontrolsüz çoğalan ve büyüyen hücrelerin oluşturduğu kitledir. Beyin tümörleri primer, metastatik (sekonder), benign ya da malign olarak görülebilir (Şimşek ve Dicle, 2013). Benign beyin tümörü iyi huylu beyin tümörüdür, kanser olmayan tümörlere denir. Vücudun diğer organlarına sıçramazlar ve yaşamı nadiren tehdit ederler. Malign beyin tümörü kötü huylu tümörlerdir ve kanser olarak adlandırılırlar.

Ulusal Beyin Tümörü Derneği'ne göre ABD'de yaklaşık bir milyon insan beyin tümörü ile yaşamaktadır. 2023'de 94.390 amerikalıya primary beyin tümörü tanısı konulmaktadır. Kötü huylu beyin tümörü olan tüm hastalar için bağıl hayatta kalma oranı %35.7'dır. 2023'de 18.990 amerikalı malignant tümöründen ölecekleri varsayılmaktadır (National Brain Tumor Society, 2023). Tümör türüne bağlı olarak hayatta kalma süresi 2 yıl denilmektedir. Bu nedenle hızlı teşhis edilmelidir. Teşhis için genelde kullanılan yöntemler; bilgisayarlı tomografi (CT) ve manyetik rezonans görüntülemedir (MRI). Her ne kadar bu iki yöntem kullanılsa da MRI taramaları kullanılarak tümör segmentasyonu ve derecelendirilmesi yaygın bir işlemdir. MRI ile tümörün konumunun belirlenmesi yanı sıra şekli ve boyutu da görülmektedir. Fakat patolojik inceleme kullanılarak tümör derecelendirilmesi için biyopsi gerektirmektedir. Bu da zaman alan, maliyetli bir işlemdir. Ayrıca gözlemci yanılgısına açıktır. Oysaki beyin tümörü tespiti için doğru segmentasyon çok önemli bir süreçtir fakat radyologlar için manuel segmentasyon zaman alıcı bir süreç olmaktadır. Bu nedenlede otomatik, yarı otomatik teknikler gerekmektedir.

Beyin analizi için T1-Ce, T1, Flair (sıvı ile zayıflatılmış inversiyon kurtarma) ve T2 en sık kullanılan MRI yöntemleridir. T1 yöntemi genelde sağlıklı dokularla çalışmak için kullanılır. T2 görüntüleri, ödem bölgelerinin sınırlarını tespit etmek için daha uygundur. T1-CE görüntüleri, tümör sınırlarını vurgulamaktadır. FLAIR görüntüleri, beyin omurilik sıvısındaki ödem bölgelerinin tespitini kolaylaştırmaktadır (Díaz-Pernas vd., 2021). Beyine ait MRI görüntülerinden örnekler Şekil 1'de verilmiştir.



Şekil 1. Beyin MRI'ına ait görüntü türleri (Daimary vd., 2020).

Daha öncede bahsettiğimiz gibi bu tümörlerin patolojik ve radyolojik kontrolleri zaman alıcı ve maliyetlidir. Ayrıca insan yanılgısına da açıktır. Hastalığa yakalanan insanların ömür sürelerinin de tümör tipine bağlı olarak kısalma durumu söz konusu olduğu için teşhis süresi çok önemli bir durum olmaktadır. Erken teşhis edildiğinde hayatta kalma olasılıklarının artacağı düşünülmektedir. Bu nedenle bilgisayar destekli tanı tespitleri için çalışmalar yapılmaktadır. Basit bir bilgisayar destekli tanı sistemi, MRI taramalarının ön işlenmesi, önceden işlenmiş MRI taramalarının özelliklerinin çıkarılması, özellik azaltımı ve son olarak denetimli bir öğrenme algoritması kullanılarak sınıflandırma gibi sıralı bir şekilde bir çok önemli adımdan oluşmaktadır (Raja vd., 2018).

Verinin ön işlem sürecinin önemi kadar seçilen algoritmanın başarım oranı da bu noktada önemli bir yer edinmektedir. Günümüze dek MRI verilerinin işlenerek en doğru şekilde beyin tümör tespit edilmesi için birçok çalışma yapılmıştır. Bir çok derin sinir ağı modeli bu alanda kullanılmıştır. Aslında bunun temelinde yatan en büyük nedenlerden biri medikal alandaki görüntüler üzerinde başarım oranı yüksek olan modelin henüz tam olarak tespit edilememiş olmasıdır. Bunun birçok nedeni olabilir; verinin iyi şekilde temizlenmemesi, yeterli olmaması vb. gibi süreçler modelin eğitimini ve buna bağlı olarak başarım oranını etkileyecektir. Bu etkenler dışında medikal alandaki görüntülerde anlamsal bölümleme çok önemli olmaktadır. Bu alanda kullanılan piksel tabanlı görüntü segmentasyonu açısında başarılı olan popüler modellerden bir tanesi, kod çözücü ve kodlayıcı özellik haritaları arasında anlamsal bir boşluğa sahiptir. Bu nedenle yüksek ve düşük düzey görüntü özelliklerinin basit bir şekilde birleştirilmesine ve piksellerin anlamsal sınıflandırılmasına elverişli değildir. Bu noktadan da yola çıkarak bu çalışmada medikal

alanda kullanılan bazı derin sinir ağlarının belirli bir veri seti ile eğitilerek test edilmesini ve başarımlarının karşılaştırılmasını hedefledik.

Bu tez çalışması; giriş, kuramsal çerçeve/önceki çalışmalar, materyal ve yöntem, araştırma bulguları, sonuç ve öneriler bölümlerinden oluşmaktadır. İlk bölüm olan giriş bölümünde; beyin tümörünün ne olduğu, hangi türlere sahip olduğu, istatiksel olarak Türkiye'de elimizde net veri olmamasına karşın Amerika'daki verilere göre 2023'de tahmini kaç tane insanın bu hastalığı yakalandığından bahsedilmiştir. Buna bağlı olarak çalışmanın amacı ve neden böyle bir çalışma yapıldığı belirtilmiştir. İkinci bölüm, kuramsal çerçeve başlığında beyin tümörü, veri madenciliği ve makine öğrenmesi, derin öğrenme ve medikal alanda kullanılan modeller hakkında ayrıntılı açıklamalar yapılmıştır. Önceki çalışmalar başlığında, beyin tümörü tespiti ve medikal alanda kullanılan algoritmalar ile ilgili yapılan çalışmalar hakkında kapsamlı bir literatür taraması yapılmıştır. Önemli çalışmalar hakkında bilgi verilmiştir. Üçüncü bölümde; çalışmada kullanılan materyal ve yöntemler ayrıntılı olarak anlatılmıştır. Kullanılan veri seti, kullanılan makine ve derin öğrenme algoritmaları, kullanılan geliştirme ortamları ve programlama dilleri hakkında bilgi verilmiştir. Dördüncü bölümde, model eğitim metrikleri ve değerlendirme metrikleri sunularak, seçilen derin öğrenme mimarilerinin karşılaştırılması sunulmuştur. Son bölümde, yapılan çalışmada sonucundan en performanslı modelin InceptionV3 olduğu belirlenmiştir. Bu alandaki limitler göz önüne serilerek ilerleyen çalışmalarda neler yapılacağından bahsedilmiştir. Son olarak bu çalışmanın yansımalarından bahsedilmiştir.

İKİNCİ BÖLÜM
KURAMSAL ÇERÇEVE/ÖNCEKİ ÇALIŞMALAR

2.1. Kuramsal Çerçeve

2.1.1. Beyin Tümörü

Tümör, anormal hücreler bir araya geldiğinde oluşan katı bir doku kütlesidir. Tümörler dokuları, cildi, kemikleri, organları ve bezleri etkileyebilir. İyi huylu (benign) ve kötü huylu (malign) tümör olarak ayrılırlar . Genelde iyi huylu olanlara kitle derken kötü huylu olanlara kanser tümör denilmektedir. Peki bu tümörlere ne sebep oluyor dediğimizde; zamanında ölmeyen hücreler ve kontrolsüz şekilde büyüyen ve bölünen hücrelerdir diyebiliriz.

Kötü huylu tümörler primary ve secondary olarak ikiye ayrılırlar. Primary malign tümörler beyinden kaynaklanan tümörlerdir. Secondary malign tümörler ise vücudun başka bir yerinden beyine yayılan tümörlerdir. Malign beyin tümörlerinin birçok farklı türü vardır. Fakat en sık görülen türleri Meningioma, Glioma ve Pituitary türleridir (Lather ve Singh, 2020).

Meningioma, beyin ve omuriliği çevreleyen zar hücrelerinde gelişir. Meningioma genellikle yavaş büyüyen iyi huylu tümörlerdir ancak yine de beyin, sinirler ve kan damarları gibi yakındaki yapılara baskı yaparak semptomlara neden olabilirler. Meningioma genellikle farklı bir tıbbi nedenden dolayı yapılan muayene veya tarama sırasında bulunur. Kadınlarda daha sık görülürler, bunun nedeni muhtemelen bu tümörlerin östrojen reseptörleri taşımasıdır. Tipik olarak ameliyat ile çıkarılabilir veya radyasyon ile tedavi edilebilir. Bazı meningiomalar için ameliyata gerek olmayabilir ve yıllarca fark edilmeden kalabilir.

Glioma, beyin tümörünün birincil türüdür. Beynin glial hücreler adı verilen destek dokusundan büyümeye başlar. Glial hücreler beyindeki nöronlara (sinir hücreleri) destek ve koruma sağlar. Gliomlar biraz iyi huylu veya kötü huylu olabilir ve astrositomlar, oligodendrogliomlar ve ependimomlar gibi alt tipleri içerir.

Pituitary, beynin tabanında yer alan ve büyüme, metabolik, üreme ve diğer vücut işlevlerinden sorumlu hormonların üretimini kontrol eden hipofiz bezinde gelişir. Hormonal dengesizliklere yol açabilir.

Tümörün erken dönemde fark edilmesi ve doğru tedavi yaklaşımları ile başarılı sonuçlar elde etmek mümkündür. Modern görüntüleme teknikleri sayesinde daha küçük boyuttayken saptanması mümkündür. Bu sayede hastalığın ilerlemesi önlenerek daha az invaziv (zarar verici) tedaviler kullanılabilir. Bir tümörün kanserli olup olmadığını belirlemek için biyopsi yapılması gerekmekte ve bir patolog tarafından incelenerek tanı konulmalıdır. Bu hem maliyetli hem de zaman alan bir süreçtir. Bunların dışında gözlemciler arası değişken sonuçlar ortaya çıkabilir. Bu nedenle otomatik tümör tespiti bu alanda tercih sebebi olmuştur.

2.1.2. Veri Madenciliği ve Makine Öğrenmesi

Veri madenciliği; varolan büyük veri kümelerinden anlam ve değerli bilgiyi çıkarma işlemidir. Veri madenciliği aşamalarını şu şekilde sıralayabiliriz; veri seti toplama, veri temizleme (smoothing), veri bütünleştirme (damy-optimization), veri indirgeme, veri dönüştürme (normalization), özellik seçimi, ilgili veri madenciliği algoritmalarını uygulamanması (kümeleme, sınıflandırma, karar destek ağaçları), değerlendirme ve yüklemedir (Bramer, 2007).

Makine öğrenmesi; bilgisayar sistemlerinin deneyimlerinden öğrenmesine ve geliştirilmesidir. Makine öğrenmesinde algoritmalar büyük veri kümelerinde ilişkileri ve anlamları bulmak ve analize dayalı en iyi kararları ve tahminleri yapmak için eğitilir. Birincil amaç, makinelerin insan müdahalesine gerek kalmadan öğrenmesini sağlamaktır. Makine öğrenmesi yapay zeka (AI) altındaki ilk alt kümedir ve diğer alt kümeleri kapsar. Sıralama şu şekildedir; AI, Machine Learning, Deep Learning, Neural network. Makine öğrenme türleri; denetimli öğrenme, denetimsiz öğrenme, takviyeli öğrenmedir (El Naqa ve Murphy, 2015).

2.1.3. Derin Öğrenme ve Medikal Alanda Kullanılan Modeller

Gözetimli öğrenme algoritmalarından olan yapay sinir ağları, bilgisayar sistemlerinin, insan beyninin özelliklerinden olan öğrenme yolu ile yeni bilgiler üretebilme, yeni bilgiler oluşturabilme keşfedebilme gibi yetenekleri, herhangi bir yardım almadan yapabilme amacı ile geliştirmişlerdir. Büyük verileri analiz etmek için etkili bir araç olarak ortaya çıkan derin öğrenme, karmaşık algoritmalar ve yapay sinir ağları kullanarak makineleri eğiterek denetimlerden öğrenmelerini, verileri tıpkı insan beyni gibi sınıflandırabilmelerini ve tanıyabilmelerini sağlıyor. Derin öğrenme içerisinde Evrişimsel Sinir Ağı (CNN), görüntü/nesne tanıma ve sınıflandırma için yaygın olarak kullanılan bir tür yapay sinir ağıdır (Lecun vd., 1998).

Medikal görüntüler üzerinde çalışma yapılan modelleri incelediğimizde; Unet, Vgg-16, Vgg-19, ResNet, Deeplab, DenseNet, GoogleNet, CapsuleNetworks (CapsNet), 3D CNN olduğunu görüyoruz.

U-Net, ilk olarak biyomedikal görüntü segmentasyonu için icat edilip kullanılmıştır. Mimarisi genel olarak bir kodlayıcı ağı ve ardından bir kod çözücü ağı olarak düşünülebilir (Ronneberger vd., 2015).

VGG-16, görüntü tanıma için kullanılan evrişimli bir sinir ağı modelidir. Çok sayıda hiper parametreye dayanmak yerine, ağırlıkları olan 16 katmana sahip olması bakımından benzersizdir. Önceki modellerden en önemli farkı evrişim katmanlarının 2'li ya da 3'lü kullanılmasıdır (Simonyan ve Zisserman, 2015).

VGG-19, 19 katman derinliğinde evrişimli bir sinir ağıdır. Görüntü tanıma ve sınıflandırma görevleri için en iyi bilinen ve yaygın olarak kullanılan sinir ağlarından biridir (Simonyan ve Zisserman, 2015).

ResNet (Residual Network), modelin bir veya daha fazla katmanı atlamasını sağlayan "kimlik kısayol bağlantısı" olarak adlandırılan bir yapay sinir ağıdır. Vgg-19 mimarisinden esas alınarak vgg ağlarından daha az filtreye ve daha düşük karmaşıklığa sahip 34 katmanlı

düz bir ağ mimarisi kullanır. Daha sonra bu düz ağa atlama bağlantıları veya artık bloklar eklenerek mimari artık ağa dönüştürülür (He vd., 2015).

DeepLab, görüntü segmentasyonu için kullanılan bir modeldir. Klinik görüntü segmentasyonu, biyoreaktör optimizasyonu, fotovoltaik tesis yönetimi ve kaldırım çatlak tespiti gibi çeşitli alanlarda uygulanır (Chen vd., 2014).

DenseNet (Dense Convolutional Network), girişe yakın katmanlar ile çıkışa yakın katmanlar arasında daha kısa bağlantılara sahip olmaları durumunda evrişimli ağların önemli ölçüde daha derin, daha doğru ve eğitilmesinin daha verimli olabileceği fikri üzerinde çalışmaktadır (Huang vd., 2016).

GoogleNet, Google'daki araştırmacılar tarafından geliştirilen bir derin evrişimli sinir ağı olan Inception Network'ün bir çeşidi olan 22 katmanlı derin evrişimli bir sinir ağıdır (Szegedy vd., 2014).

CapsuleNetwork (CN), bir özelliğin farklı durumlarını yakalamak yerine, bir özelliğin farklı durumlarının benzerliğini elde etmek için eğitilirler. Böylece kapsül yalnızca bir özellik tespit etmek yerine farklı durumları tespit edilebilmektedir. Kapsül ağı, görüntü içindeki hiyerarşik ilişkiyi daha verimli modelleyen yepyeni bir derin öğrenme yapısıdır (Sabour vd., 2017).

3D Convolutional Neural Network (3D CNN), hiyerarşik veri temsillerini öğrenebilen çok katmanlı sinir ağı mimarilerini ifade eder. Her katman, verilerin giderek daha karmaşık hale gelen mekansal özelliklerini öğrenir (Alakwaa vd., 2017).

2.2. Önceki Çalışmalar

Tıbbi görüntü işlemede beyin tümörü segmentasyonu önemli bir rol oynar. Bu tümörlerin erken teşhisi hastalara tedavi verilebilmesi açısından büyük önem taşımaktadır. Erken teşhis ile hastanın yaşam şansı artmaktadır. Beyin tümörlerinin doktorlar tarafından teşhis edilme süreci normalde manuel bir segmentasyon yöntemi kullanılarak

gerçekleştirilmektedir. Zaman alıcı ve zor bir süreç olmaktadır (Pereira vd., 2016). Bu sorunları çözmek amacıyla literatürde çok fazla çalışma olduğu görülmektedir.

Kayalibay ve arkadaşları, semantik segmentasyondaki gelişmelerin evrişimli sinir ağlarının tıbbı görüntü segmentasyonunda uygulanmaya başlamasını sağlamasıyla üç boyutlu filtreye sahip CNN tabanlı bir yöntem göstererek bu çalışmayı el ve beyin MRI'larına uygulamaktadırlar. U-net mimarisinde uygulanan iki mimari değişikliğin testlerini yapmışlardır. Bunlardan biri, farklı ölçeklerde oluşturulan birden fazla segmentasyon haritasının birleştirilmesinin performansa zarar vermeden yakınsamayı hızlandırdığını gözlemlemişlerdir. Tıbbi görüntülerin segmentasyonun zorlu yönlerinden biri olan sınıf dengesizliği sorununu çözmek için jaccard benzerlik indeksine dayalı bir kayıp fonksiyonu kullanmışlardır. Etkili tıbbi görüntülerin az olmasına rağmen, 3D CNN mimarilerinin iyi kalitede sonuçlar elde edildiğini göstermişlerdir (Kayalibay vd., 2017).

Tang ve arkadaşları, tıbbi görüntü sınıflandırması için AlexNet'i (Krizhevsky vd., 2012) incelemişlerdir. Sürekli gelişen derin öğrenme, tıbbi görüntü sınıflandırma alanında geniş bir yelpaze oluşturmuş ve sürekli artan performanslı sinir ağı modelleri, aşırı uyumu azaltmak (overfitting), genellemeyi geliştirmek (generalization), gradyan (gradient) kaybolması ver patlamasını önlemek gibi ortak noktaları paylaşmışlardır. AlexNet aşırı uyumluluk ve gradyan ile ilgili sorunlar için ReLU aktivasyon fonksiyonlarını kullanmışlardır. Bu nedenle bu çalışmada CNN'lerin gelişiminde büyük katkısı olan ALexNet üzerinde çalışılmış ve bu modellerin daha hızlı model eğitimi ve daha iyi özellik çıkarımı için çok etkili olduğunu bulmuşlardır. AlexNet'in tıbbi görüntülemedeki uygulamalarının geniş kapsamlı ve etkili olduğunu öne sürmüşlerdir (Tang vd., 2023).

Kanade ve Gumaste, MRI görüntülerini kullanarak beyin tümör tespiti yapmışlardır. Anormal ve kontrolsüzce hücre çoğalması olan beyin timörünün tespit edilmesini ve veritabanını test ederek, eğiterek tümörün aşamalarını sınıflandırmasını yapmaktadırlar. Test amaçlı segmentasyonlarını FCM ile yapmışlardır. Glial beyin tümörü olan ve herhangi bir ameliyat geçirmemiş 15 hastadan oluşan bir veri setini test görüntüsü için kullanmışlardır. Tüm yoğunluk değerlerini maksimum yoğunluk değerine bölerek görüntülerin yoğunluk aralığını normalleştirmişlerdir. Sinyal-gürültü oranını iyileştirmek için, ön işleme adımı

olarak görüntülere anizotropik difüzyon filtresi uygulamışlardır. Görüntünün gürültüsünü gidermek için medyan tabanlı filtreler ve SWT tekniği kullanılmıştır. MRI'larda kullanılacak özellikleri çıkarmak için SWT kullanmışlardır (The Stationary Wavelet Transform). Bu çalışmanın amacı segmentasyon için FCM kullanılmasıdır ki bu aşama için FCM (Fuzzy C-Means) yöntemini kullanmaktadırlar. Çalışma sonucunda önerilen algoritmanın yüksek doğruluk ve düşük hata oranlarına sahip olacağı öngörülmüştür. Deneysel sonuçların istatiksel analizi, geliştirilen algoritmanın beyin MRI'ların iyi bir doğrulukla bölünebildiğini göstermişlerdir. Yaptıkları araştırma sonucunda tıbbi görüntü bölütleme ve sınıflandırma algoritmalarında doğruluk, güvenilirlik, tekrarlanabilirlik, sağlamlık ve en az operatör bağımlılığı olması gerektiğini öne sürmüşlerdir (Kanade ve Gumaste, 2015).

Milica ve Marko, üç tümör tipinin beyin tümörü sınıflandırması için yeni bir CNN mimarisi sunmuşlardır. Geliştirilen ağı, halihazırda var olan önceden eğitilmiş ağlardan daha basit olarak tasarlamışlardır ve T1 ağırlıklı kontrastlı manyetik rezonans görüntüleri üzerinde test etmişlerdir. Ağın performansını dört yaklaşım kullanarak değerlendirmişlerdir. Bunlardan ilki, iki adet 10 kat çapraz doğrulama yöntemi ve iki veri tabanı kombinasyonudur. Önerilen ağın genelleme kabiliyeti, 10 kat çapraz doğrulama yöntemlerinden biri olan konu bazlı çapraz doğrulama kullanılarak test edilmiştir. Artırılmış bir görüntü veri seti kullanılarak iyileşme oranı test edilmiştir. 10-kat çapraz doğrulama yöntemi için en iyi sonuç, artırılmış veri seti için kayıt bazlı çapraz doğrulamada elde edilmiştir ve bu durumda doğruluğun %96,56 olduğu tespit edilmiştir. İyi genelleme kabiliyeti ve iyi yürütme hızı ile yeni geliştirilen CNN mimarisinin, tıbbi teşhis alanında radyologlar için etkili bir karar destek aracı olacağını öne sürmüşlerdir (Badža ve Barjaktarović, 2020).

Filatov ve Yar, beyin tümörü tespitinde manuel süreci teşhis sürecinden çıkarmak ve bunun yerine makine öğrenimi kullanılması üzerine çalışmışlardır. Beyin tümörünün teşhisi ve sınıflandırılması için önceden eğitilmiş konvolüsyonel sinir ağlarını (CNN) kullanmışlardır. Kullanılan ağlar ResNet50, EfficientNetB1, EfficientNetB7, EfficientNetV2B1'dir. Beyin tümörü için makine öğrenimi hattındaki ilk adım sınıflandırma, veri kümesinin yeniden şekillendirilmesi ve büyütülmesidir. Görüntü büyütme işleminden sonra sınıflandırma, önceden eğitilmiş farklı modeller kullanılarak gerçekleştirilmiştir.

EfficientNet modellerinin ResNet modellerini önemli bir farkla geride bıraktığı açıkça gösterilmektedir. ResNet50 için toplam epoch sayısı 30 olduğundan, doğrulama kaybının epoch 21'den sonra azalmayı bıraktığı da not edilmiştir. Diğer modellerde ise doğrulama kaybı daha sonraki aşamalarda azalmayı dudurmuştur. Bu aynı zamanda ResNet50 modelinin yüksek doğrulama kaybını da açıklamaktadır. EfficientNetB1 en iyi eğitim doğruluğu (training accuracy) ve doğrulama doğruluğuna (validation accuracy) sahip olarak en iyi performansı sağlamıştır. EfficientNetB1 için eğitim kaybı (training loss) da en düşük seviyededir. Ancak doğrulama kaybı (validation loss) üçüncü en düşük seviyeye sahiptir. Tüm EfficientNet modelleri eğitim doğruluğunda istikrarlı bir artış ve eğitim kaybında istikrarlı bir düşüş göstermişlerdir. ResNet50 modeli için eğitim doğruluğu (training accuracy) önce azalıp sonra arttığı için istikrarsızdır. Ancak ilk ve son doğruluk arasındaki farkın büyük olmadığı gözlemlenmiştir. Gelecekteki çalışmalarda daha iyi sınıflandırma doğruluğu için daha fazla veri toplanması gerektiğini öngörmüşlerdir (Filatov ve Yar, 2022).

Maurya ve arkadaşları, daha önceden işlenmiş MRI görüntülerini kullanarak glioblastoma beyin tümörlerinin tespiti ve segmentasyonu için Unet ve Deeplabv3 olmak üzere iki derin öğrenme modelini önermişlerdir. Bu modeller için doğruluk ve hesaplama verimliliği açısından performans değerlendirilmesi yapmışlardır. DeepLabv3'ün beyin omurilik sıvısı gibi geniş homojen doku bölgelerini içeren görüntülerle sorun yaşama eğiliminde olduğunu bulmuşlardır. Bu durumlarda, model genellikle ilgilenilen bölgeyi doğru bir şekilde segmentlere ayırma konusunda başarısız olmuş ve bunun yerine görüntünün tamamını arka plan olarak sınıflandırmayı seçmişlerdir. Ayrıca modelin bazen, özellikle de aralarındaki sınırlar iyi tanımlanmadığında, farklı doku türlerini karıştırdığını da gözlemlemişlerdir Fakat buna rağmen hem U-Net hem de Deeplabv3 modellerinin glioblastoma beyin tümörlerinin doğru tespitini ve segmentasyonunu sağladığını göstermişlerdir. Ancak Deeplabv3'ün, daha fazla hesaplama kaynağı gerektirmesine rağmen doğruluk açısından UNet'ten daha iyi performans gösterdiğini öne sürmüşlerdir (Maurya vd., 2023).

Gupta ve arkadaşları, tıbbi görüntü segmentasyonu amacıyla, beyin tümörü segmentasyonu için kullanılan veri setinde tutarlı sonuçlar gösteren yetenekli derin öğrenme modelini incelemişlerdir. Halka açık bir MRI görüntüleme veri seti kullanmışlardır. Veri

kümesi dosyaları, çeşitli omurgalara sahip U-Net ve Attention U-Net, Deep Residual U-Net, ResUnet++ ve Recurrent Residual gibi bazı iyi bilinen görüntü bölümleme derin öğrenme modellerinin uygulanması ve eğitilmesi öncesi dönüştürülmekte ve ön işleme tabi tutulmaktadır. Tüm modeller önceden herhangi bir ağırlık olmaksızın tamamen sıfırdan eğitilmiştir. Bu çalışmadan elde edilen sonuçlar, Tekrarlayan(Recurrent) ve Artık (Residual) birimlerle birleştirilmiş kodlayıcı-kod çözücü tabanlı Evrişimli Sinir Ağı mimarisinin, Adam optimizer kullanılarak beyin tümörü MRI taramalarının bir veri kümesi üzerinde eğitildiğinde 0,8495 F1 puanı ve 0,8665 IoU ürettiği gösterilmiştir. U-Net ve Attention U-Net, varyantları ile birlikte deep Residual U-net gibi diğer karşılaştırılan modellerden daha iyi performans gösterdiğini öne sürmüşlerdir (Gupta vd., 2023).

Imam ve Alam, MRI verilerini kullanarak beyin tümörü sınıflandırması için Transfer Öğrenme-CNN adı verilen yeni bir derin öğrenme tabanlı yaklaşım sunmuşlardır. Önerilen model, önceden eğitilmiş ağırlıkları kullanarak ve bu ağırlıkları CNN'e aktararak mevcut halka açık modellerin tahmin yeteneklerinden yararlanmaktadır. Halka açık bir beyin MRI veri kümesinden yararlanılarak, menenjiom, gliom ve hipofiz tümörleri dahil olmak üzere farklı tümör türlerini sınıflandırmak için çeşitli transfer öğrenme modellerini değerlendirmişlerdir. 8 transfer öğrenme + CNN modelini kullanırken 5 farklı metodolojiyi (artırma, odak kaybı, smote ve adasyn) karşılaştırmışlardır. Önerilen modelin etkinliğini değerlendirmek için doğruluk, kesinlik, hatırlama ve F1 puanı dahil olmak üzere çeşitli değerlendirme kriterleri kullanmışlardır. Bu dengesiz veri kümesi, halihazırda eğitilmiş derin öğrenme modellerinin yanı sıra önerilen transfer öğrenme + CNN modeli kullanılarak kategorize edilmiş; veri kümesinin %90'ı eğitim için ve %10'u test için kullanılmıştır. CNN sürecinde sınıflandırma görevleri sıklıkla bir karışıklık matrisi kullanılarak gerçekleştirilmiştir. Deneylerin sonucu, VGG-16'nın doğruluk değerlendirme ölçütü açısından tüm teknikler değerlendirildiğinde genel olarak en iyi performansı gösterdiğini öne sürmüşlerdir. Artırma ile birlikte 0,96, şimdiye kadar genel olarak elde edilen en yüksek doğruluk olmuştur. Bu çalışmada, MobileNet ve EfficientNets sırasıyla 0,23 ve 0,31 genel doğruluk oranına sahiptir, bu da en kötü doğruluk sonuçlarını göstermektedir. VGG16, EfficientNets, DenseNet201, MobileNet ve GoogleNet üzerindeki ilgili dengesiz veri seti için sırasıyla 0,85, 0,66, 0,42 ve 0,78 değerlerine sahip kayıp fonksiyonları karşılaştırıldığında odak kaybının çapraz entropi kaybına göre daha yüksek doğruluk elde ettiği bulunmuştur. Çapraz entropi kaybıyla birlikte geri kalan modeller daha iyi performans

göstermişlerdir. Özetlemek gerekirse, diğer modellere göre çok daha az katmana sahip olmasına rağmen, VGG-16 modelini kullanan transfer öğrenimi genel olarak en iyi performansı göstermiştir. Dengesiz veri kümesinde çeşitli transfer öğrenme + CNN modellerini kullanan performans ölçümleri karşılaştırılmıştır. Burada, çapraz entropi kaybına sahip artırılmış bir veri kümesindeki performansın, çeşitli değerlendirme parametrelerinde alternatif yöntemlerden daha iyi performans gösterdiğini öne sürmüşlerdir (Imam ve Alam, 2023).

Qian ve arkadaşları, kodlayıcı kod çözücü mimarisine sahip bir dizi UNet benzeri ağ, UNet2+ ve UNet3+'nın atlama bağlantılarını yeniden tasarladığı, sırasıyla yoğun atlama bağlantısı ve tam ölçekli atlama bağlantısı önerdiği ve tıbbi görüntü segmentasyonunda UNet ile karşılaştırıldığında önemli ölçüde iyileştiği, adına UNet-sharp (UNet#) dedikleri bir model önermişlerdir. Önerilen model, kodlayıcı ve kod çözücü alt ağındaki farklı anlamsal ölçeklerin özelliklerini, iç içe geçmiş ve yoğun atlama bağlantıları aracılığıyla bir araya getirmektedir. Ayrıca, doğru organ veya lezyon konumu algısını ve segmentasyonunu gerçekleştirmek için çok ölçekli özellik haritalarında yüksek seviyeli anlambilimin ve düşük seviyeli ayrıntıların birleşimini en üst düzeye çıkarmak için tam ölçekli bağlantıları kullanmışlardır. Aşırı segmentasyonu önlemek amacıyla model için iki tür CGM tasarlamışlardır. Farklı organ ve desen veri kümeleri üzerinde anlamsal ve örnek bölümlendirmeden elde edilen deneyler sonucunda önerilen modelin tüm son teknoloji yöntemlerden daha iyi performans gösterdiğini öne sürmüşlerdir (Qian vd., 2022).

Dehkordi ve arkadaşları, doğru beyin görüntüsü sınıflandırması için geliştirilmiş bir evrişimsel sinir ağı geliştirmişlerdir. Geliştirilmiş evrişimli sinir ağı yapısı, özellik çıkarımı ve optimal sınıflandırma için bileşenlerden oluşmaktadır. Evrişimli sinir ağının hiper parametrelerini optimize etmek için MFO algoritmasının kaotik ve doğrusal olmayan bir versiyonu önerilmiştir. Meta-sezgisel bir optimizasyon algoritması olmasına rağmen, MFO algoritmasının teorik temellerinin kapsamlı bir incelemesi, evrişimli sinir ağlarının hiper parametrelerinin optimize edilmesi gibi büyük ölçekli problemlere uygulandığında verimsizliğini ortaya çıkarmışlardır. Evrişimli sinir ağının hiper parametrelerini optimize etmek için MFO algoritmasının kaotik ve doğrusal olmayan bir versiyonu önerilmiştir. Meta-sezgisel bir optimizasyon algoritması olmasına rağmen, MFO algoritmasının teorik

temellerinin kapsamlı bir incelemesi, evrişimli sinir ağlarının hiper parametrelerinin optimize edilmesi gibi büyük ölçekli problemlere uygulandığında verimsizliğini ortaya çıkarmışlardır. Büyük ölçekli sorunları ele alırken MFO'nun düşük doğruluğunun nedenini belirlemek için çeşitli ön deneyler yapmışlardır. MFO'nun bu tür konulardaki temel zayıflığının, keşif ve işletme aşamaları olduğu belirlenmiş ve sorun tanımının ardından, arama ve kullanım süreçlerinin verimliliğini en üst düzeye çıkarmak için optimum stratejiyi belirlemek amacıyla MFO'ya Levy uçuş teorisi ve kaotik değerler gibi değişkenler eklemişlerdir. Ancak Levy uçuş teorisi değerleri ve kaotik haritaların kullanılmasının algoritmanın performansını arttırmak için yetersiz kaldığı tespit edilmiştir. Bu sorunun üstesinden gelmek için doğrusal olmayan bir ağırlık katsayısı parametresi tasarlanıp, bu algoritmaya dahil edilmiştir. Bu stratejinin bir avantajı, hem arama hem de kullanım aşamaları arasında uygun dengeyi sağlamak için doğrusal olmayan ağırlık katsayısı parametrelerinin kullanılması olduğunu öne sürmüşlerdir. Ayrıca, Levy uçuş teorisi değerlerini kullanarak ve rastgele değerleri kaotik değerlerle değiştirerek MFO'nun keşif ve kullanım aşamalarını iyileştirmişlerdir. NLCMFO, çok çeşitli değişkenler kullanılarak çok çeşitli tek modlu, çok modlu ve bileşik kıyaslama fonksiyonları üzerinde test etmişlerdir. Çoğu deneysel durumda, sonuçlar NLCMFO'nun MFO ve diğer karşılaştırmalı algoritmalardan daha iyi performans gösterdiğini göstermiştir. NLCMFO, beyin tümörü sınıflandırması için bu tür bir ağın hiper parametrelerinin optimize edilmesinde umut verici sonuçlar elde etmiştir. Sonuçlar açısından NLCMFO, gelişmiş algoritmalardan önemli ölçüde daha iyi performans göstererek güçlü yönlerini ve faydalarını göstermektedir. Optimize edilmiş CNN modeli, %97,4 doğruluk, %96,0 duyarlılık, %98,6 özgüllük, %98,4 kesinlik ve %96,6 F1 puanı sağlayarak literatürdeki diğer modellerden daha iyi performans gösterdiği öne sürülmüştür (Dehkordi vd., 2022).

Santos, bilgisayarla görme işlemlerinin performansıyla bilinen bir evrişimli sinir ağı olan VGG-16 modelini kullanarak beyin tümörü tespiti üzerine bir çalışma yapmıştır. Veri ön işleme, model oluşturma ve değerlendirmeyi içeren sistematik bir metodoloji takip ederek umut verici sonuçlar elde etmiştir. VGG-16 modeli, doğrulama setinde yaklaşık %88 ve test setinde %80 doğrulukla beyin tümörü görüntülerinin sınıflandırılmasında yüksek doğruluk göstermiştir. Bu bulgular, VGG-16 modelinin, beyin tümörlerinin erken tespiti ve teşhisinde sağlık profesyonellerini desteklemek için değerli bir araç olma potansiyelini ortaya koymuştur. Özetle, VGG-16 modelinin sağlık profesyonellerinin MRI görüntülerinden

beyin tümörlerini doğru bir şekilde tespit etmelerine yardımcı olma potansiyelini göstermiştir (Santos, 2023).

Speet, yaptığı çalışma ile capsule network'ün otomatik beyin tümörü segmentasyonunu ne ölçüde gerçekleştirildiğini araştırmıştır. Birkaç konvolüsyonel ve dekonvolüsyonel kapsülden oluşan kodlayıcı-kod çözücü tarzı bir çerçeve kullanan SegCaps algoritmasını uygulamıştır. SegCaps çerçevesinin tüm tümör yapılarının yanı sıra tümör çekirdek yapılarını da bölütlemek için kullanıldığı çeşitli deneylerin sonuçlarına dayanarak, kapsül tabanlı bir mimarinin U-Net'ten daha kötü performans gösterdiği sonucuna varmıştır. Önerilen SegCaps yaklaşımının, bu özel kurulumda, doğru beyin tümörü segmentasyonunu gerçekleştirmek için yeterince uygun olmadığını tespit ederek, bu çalışmada, beyin tümörü segmentasyonu ve kapsül ağları alanındaki mevcut literatüre katkıda bulunmuştur (Speet, 2020).

Soumik ve Hossain, CNN'lerin daha iyi sonuç vermek için büyük miktarda eğitim verisine alışkın olması ikilemine transfer öğrenmenin yararlı olduğunu ileri sürmüşlerdir. Glioma, Meningioma ve Hipofiz tümörlerini sınıflandırmak için 3 sınıflı derin öğrenme modeli önermişlerdir. Transfer öğrenme kavramını benimseyerek önerdikleri model, beyin MRI'lardan öznitelikler çıkararak bunları sınıflandırmak için softmax sınıflandırıcı yöntemini kullanan önceden eğitimli bir InceptionV3 modeli olmuştur. Önerilen sistem, figshare'den alınan CE-MRI veri kümesi üzerinde hasta düzeyinde beş kat çapraz doğrulama işlemi ile değerlendirilen önceki tüm yöntemlerden daha iyi performans göstererek ortalama %99 sınıflandırma doğruluğu elde etmişlerdir (Soumik ve Hossain, 2020).

Tamilarasi ve Gopinathan, beyin kanseri teşhisi için vücuda dışarıdan müdahale etmeden teşhis destek sistemi sağlamak amacıyla standart mimari tasarımlarından VGG deki daha derine inmenin temel dezavantajlarını da aşmak amacıyla inception ağını (GoogleNet) kullanmışlardır. Bu ağ ile daha derine gitmek yerine genişleyerek paralel evrişim katmanları kullanmışlardır. Manyetik rezonans görüntüleme (MRI) görüntüleri üzerinde beyin görüntüsü sınıflandırmasına yardımcı olmak için başlangıç mimarisinin ortalama %95,1 doğruluk, %96,2 duyarlılık ve %94 özgünlük ile kullanışlılığını göstermişlerdir. VGG-16 da

ise ortalama doğruluğun %92,8 olduğunu belirterek, inception mimarisini VGG mimarisine göre iyileşme gösterdiğini ortaya koymuşlardır (Tamilarasi ve Gopinathan, 2021).

Afshar ve arkadaşları, CNN'lerin tümör sınıflandırması için özellikle zararlı olan uzamsal ilişkileri tam olarak kullanmakta başarısız olduğunu savunmuşlardır. Bu eksikliği gidermek için CapsNet'leri kullanmışlardır. Ancak CapsNet'ler çeşitli görüntü arka planlarına karşı oldukça hassas oldukları için CapsNet'i ana hedeften uzaklaştırmadan tümörü çevreleyen dokulara erişim sağlayacak şekilde donatmak amacıyla yeni bir CapsNet mimarisi önermişlerdir. Bu mimari, CapsNet'in odağını artırmak için tümörün kaba sınırlarını kendi boru hattı içinde ekstra girdiler alarak belirlemektedir. Önerilen yaklaşımın benzerlerinden belirgin şekilde daha iyi performans gösterdiğini öne sürmüşlerdir. Önerilen CapsNet mimarisinin iki ana avantajı olduğunu savunmuşlardır. Bunlardan ilki, tümörü tam açıklama ihtiyacının ortadan kalkmış olmasıdır. İkincisi ise CapsNet'in ana alana odaklanması ve aynı zamanda çevre dokularla ilişkisini dikkate aldığını öne sürmüşlerdir (Afshar vd., 2018).

Walsh ve arkadaşları, beyin tümörü segmentasyonu için U-net'in basit bir uygulanışını önermişlerdir. Önerilen mimari ile MRI taramalarının gerçek zamanlı bölümlendirilmesini sağlamanın yanı sıra, önerilen hafif U-Net'i eğitmek için büyük miktarda veriye ihtiyaç duymamışlardır. Üstelik hiçbir ek veri artırma adımına gerek olmamıştır. BITE'ın veri setini kullanmışlardır. Bu üç boyutlu MRI beyin taramalarını iki boyutlu görüntülere dönüştürerek, bu verileri, anormallikleri iki boyutlu görüntülerden doğru bir şekilde bölümlere ayırabilen çok hafif bir U-Net uygulamasının performansını değerlendirmek için kullanmışlardır. Bu mimari, BITE veri kümesinde çok ümit verici sonuçlar vermiştir ve standart kıyaslama algoritmalarından daha iyi performans gösterirken %89'luk bir ortalama kesişme-birleşim (IoU) elde etmişlerdir. Ayrıca bu çalışma karmaşıklaştırılmış beyin tümörü segmentasyonu için orijinal üç boyutlu hacimsel görüntüler yerine üç perspektif düzleminin (Axial,Coronal, Sagittal) etkili bir şekilde kullanıldığını göstermektedir (Walsh vd., 2022).

Nasim ve arkadaşları, beyin tümörlerinde nekrotik, ödemli, büyüyen ve sağlıklı dokuların örtüşmelerini ararken görüntülerden ilgili bilgileri elde etmenin zor olabileceğini

öne sürerek 2D U-Net önermişlerdir. 2D U-Net ağı, bu dört alanı bulmak için BraTS veri kümeleri ile geliştirilip eğitilmiştir. Beyin tümörü segmentasyonunu tahmin etmek amacıyla referans beyin tümörü veri seti ile en son teknolojiye sahip modeller arasında kapsamlı bir karşılaştırmalı analiz sunmuşlardır. Veri seti olarak BraTS 2017, 2018, 2019 ve 2020 veri kümelerini kullanmışlardır. 2017-2020 veri setlerinin çok fazla farklılığa sahip olmadığı sonucuna varmışlardır. Dolayısıyla, herhangi bir model bir veri kümesinde iyi çalışıyorsa, muhtemel diğerleri için de iyi bir sonuç üreteceğini öne sürmüşlerdir. BraTS veri kümeleri üzerinde yapılan deneyler, beyin tümörlerini MRI'dan bölümlere ayırmak için önerilen modelin iyi çalıştığını göstermektedir. Segmentasyon için farklı CNN veya değiştirilmiş CNN modelleri, geleneksel makine öğrenimi modellerinden daha iyi sonuçlar vermektedir. UNet ise diğer CNN modellerinden daha iyi bir skor elde etmektedir. 2019 BraTS veri kümesini kullanarak 0,8409 DICE puanıyla bölümlere ayırmak için 2D UNet kullanmışlardır. 2D UNet, derinlemesine bilgiden yararlanılmaması nedeniyle bilgileri kaybetmektedir. Bu nedenle 3D Unet üzerinde çalışma yapmaya karar vermişlerdir (Nasim vd., 2023).

Kokkala ve arkadaşları, üç sınıflı beyin tümörü sınıflandırması için derin, deep dense inception residual network önermişlerdir. Inception, ResNetV2'nin çıktı katmanını derin yoğun bir ağ (3 katmanlı) ve softmax katmanıyla özelleştirmişlerdir. Bu yoğun katmanlar beyin tümörlerinin özelliklerini içerirken, önerilen IRNet'in parametrelerini Imagenet ağırlıklarıyla kullanmışlardır. Adam optimizasyonun kullanılması nedeniyle hızlı öğrenme sergilemiş ve modelin aşırı uyum problemlerini önlemişlerdir. Derin yoğun ağ, önerilen modelin sınıflandırma doğruluğunu arttırmaktadır. Performans değerlendirmesi için yaygın olarak kullanılan 3064 T1 ağırlıklı görüntülerden oluşan beyin tümörü veri seti kullanılmıştır. Önerilen model mevcut modellerden %99,69 doğrulukla daha iyi performans göstermektedir. Ayrıca önerilen model, gürültülü görüntü verileri üzerinde benzer performans sergilemiştir (Kokkalla vd., 2021).

Casamitjana ve arkadaşları, 3D Evrişimli Ağların uygulanmasına yönelik çalışmalarını sunmuşlardır. Genel görüntü segmentasyonu için kullanılan iyi bilinen 2D modelden ilham alan, tamamen evrişimli 3D CNN mimarisi önermişlerdir. 3DNet'in, kaba yüksek katmanlı bilgileri ince, düşük katmanlı bilgilerle birleştiren atlama bağlantılarına

sahip, VGG mimarisini temel alan 3 boyutlu, tamamen evrişimli bir ağ olduğunu öne sürmüşlerdir. Bu modelin yoğun çıkarım yapma, yani tek geçişte tüm hacimler için tahminler yapma kapasitesine sahip olduğunu belirtmişlerdir ve gelecekteki çalışmaları için bir ön çalışma sunmuşlardır (Casamitjana vd., t.y.).

Çetiner ve Metlek, multimodal görüntüleri kullanarak yüksek doğrulukla segmentasyon gerçekleştirmek için yeni bir derin öğrenme tabanlı yaklaşım olan DenseUNet+'ı önermişlerdir. DenseUNet+ modelinde yoğun blok yapılarında dört farklı modaliteden gelen veriler bir arada kullanılmıştır. Daha sonra bu verilere doğrusal işlemler uygulanmış ve ardından birleştirme işlemi gerçekleştirmişlerdir. Bu şekilde elde edilen sonuçlar kod çözücü katmanına aktarılmıştır. Önerilen yöntem aynı zamanda BraTS2021 ve FeTS2021 veri setlerinde dice ve jaccard metrikleri kullanılarak aynı veri setini kullanan son teknoloji çalışmalarla da karşılaştırılmıştır. Karşılaştırma sonucunda; BraTS2021 veri seti için dice ve jaccard değerlendirme metrikleri sırasıyla %95 ve %88 olurken, FeTS2021 için sırasıyla %86 ve %87 performans değerleri elde edilmiştir. Performans sonuçlarının birçok SOTA beyin tümörü segmentasyon yönteminden daha iyi olduğunu belirtmişlerdir. Ayrıca DenseUNet+ modeline özel bir ön işleme yöntemi de geliştirmişlerdir. Ön işleme yönteminde dört farklı modalite de ROI (ilgi alanı)bölgeleri belirlemişlerdir. ROI bölgelerini DenseUNet+ modelinde girdi verileri olarak kullanmışlardır. DenseUNet+ modelinin kodlayıcı aşamasında, bu ROI bölgeleri yoğun bloklar aracılığıyla birbirlerine bağlanarak güçlü özellik haritaları üretilmiştir. Kodlayıcı ve kod çözücü arasında kullanılan ağırlıklı atlama düğümleri, kod çözücü katmanının sonunda yeni bir segmente edilmiş görüntü oluşturmuştur. Kodlayıcı ve kod çözücü arasında kullanılan ağırlık değerlerinde, üst katmanlarda küçük değerler kullanılırken, alt katmanlara inildikçe daha yüksek ağırlıklar kullanılmıştır. Bu, kod çözücü katmanında oluşturulan görüntünün, aynı seviyede yer alan kodlayıcı bloktan gelen değerler nedeniyle orijinal görüntüden anormal bir şekilde sapmasını önlemiştir. Bu çalışma sayesinde, özellikle UNet mimarilerinde sıkça karşılaşılan gradyan kaybolma problemini önemli ölçüde ortadan kaldırmışlardır (Çetiner ve Metlek, 2023).

Zheng ve arkadaşları, beyin tümörlerini segmente etmek için SCU-Net adını verdikleri geliştirilmiş bir U-Net algoritması önermişlerdir. İki U-Net modelini VGG16'nın ana gövdesi olarak bir arada kullanmış ve her katmanda özellik birleştirme ve kod çözme

modülü birleştirme işlemleri yapmışlardır. Böylece önce ve sonra gelen iki kodlama-çözme bloğu daha sıkı bir bağlantı oluşturarak, daha fazla anlamsal bilgi elde etmiştir. Bununla özellik tekrarını azaltabilir ve modelin genelleme yeteneğini daha da artırabileceğini öne sürmüşlerdir. Beyin tümörü kategorisi sınıflandırması için konum bilgisi son derece önemli olduğundan, daha fazla konum bilgisi yakalamak için U-Net kodlama ağına başka bir HDC modülü eklemişlerdir. Önerilen yöntemin hedef segmentasyonunu ve tümör tahmin performansını önemli ölçüde iyileştirebileceğini göstermişlerdir (Zheng vd., 2022).

Tablo 1

MRI sınıflandırması ve tespiti ile ilgili çalışmaların karşılaştırması

Yazarlar ve araştırmacı	**Veri Seti**	**Model**	**Sonuç**
(Kayalibay vd., 2017)	Brats 2013 and 2015		CNN tabanlı bir yönem üç boyutlu filtrelerle desteklenen bir mimari kullanılmaktadır.
(Tang vd., 2023)		AlexNet	102 tıp metni üzerinde araştırma yapılarak Alexnet'in tıbbi görüntülerde kapsamlı olduğu belirtilmektedir
(Kanade ve Gumaste, 2015)	Veriler, Ankara, Türkiye'deki 29 Mayıs Hastanesi Radyoloji Bölümündeki tıbbi görüntü veri tabanından elde edildi	SWT (The Stationary Wavelet Transform) ve FCM (Fuzzy C-Means)	MRI Görüntülerine dayalı olarak Beyin Tümörü Tespiti için önerilen algoritmalar daha yüksek doğruluğa ve düşük hata oranlarına sahip olacaktır.

Tablo 1’in devamı

(Badža ve Barjaktarović, 2020)	2005'ten 2010'a kadar Çin'deki Tianjin Tıp Üniversitesi'ndeki Nanfang Hastanesi ve Genel Hastane'den alınan 3064 T1 ağırlıklı kontrastı artırılmış MRI görüntüsü içermektedir.	10-fold cross-validation kullanılmış yeni bir CNN mimarisidir.	Doğruluk oranı %96,56 olarak tespit edilmektedir.
(Filatov ve Yar, 2022)	Veri kümesi üç veriden oluşan bir kombinasyondur; figshare, SARTAJ ve Br35H.	ResNet50, EfficientNetB1, EfficientNetB7, EfficientNetV2B1'	EfficientNetB1, en iyi sonuçları göstermektedir. Training ve validation için doğruluk oranı sırasıyla %87,67 ve %89,55’dir.
(Maurya vd., 2023)	BraTS (253 MRI images)	UNet, Deeplabv3	Deeplabv3, daha fazla hesaplama kaynağı gerektirmesine rağmen doğruluk açısından UNet'ten daha iyi performans göstermektedir. Doğruluk oranı: 90.39%
(Gupta vd., 2023)	Beyin T1 ağırlıklı CE-MRI veri seti, 2005 ile 2010 yılları arasında Çin'in Guangzhou kentindeki Tianjing Tıp Üniversitesi ve Nanfang Hastanesi'nden elde edilmiştir.	U-Net ve Attention U-Net, Deep Residual U-Net, ResUnet++ ve Recurrent Residual U-Net	Recurrent Residual U-Net mimarinin diğer mimarilere göre daha iyi performans gösterdiği sonucuna varılmaktadır.

Tablo 1'in devamı

(Imam ve Alam, 2023)	figshare, SARTAJ ve Br35H	Transfer Learning-CNN	VGG-16 ve CNN'in birleştirildiği stratejide, diğer sunulan yaklaşımlara göre en iyi doğruluk oranına sahip olduğu öne sürülmektedir. Doğruluk oranı %96'dır.
(Qian vd., 2022)	Dsb2018 (Çekirdek segmentasyonu), BraTs19 (Beyin tümörü segmentasyonu), Lits17 (Karaciğer segmentasyonu), LIDC-IDRI (Akciğer 2D segmentasyonu), Luna16 (Akciğer 3D segmentasyonu)	Unet-Sharp	Farklı organ ve desen veri kümeleri üzerinde anlamsal ve örnek bölümlendirmeden elde edilen deneyler, önerilen modelin tüm son teknoloji yöntemlerden daha iyi performans gösterdiğini göstermektedir.
(Dehkordi vd., 2022)	BraTS 2015, Harvard Medical Üniversitesinin beyin görüntüsü veri seti	Non-linear Lévy Chaotic Moth Flame Optimizer (NLCMFO)	Optimize edilmiş CNN modeli, %97,4 doğruluk, %96,0 duyarlılık, %98,6 özgüllük, %98,4 kesinlik ve %96,6 F1 puanı sağlayarak literatürdeki diğer modellerden daha iyi performans gösterdiği öne sürülmektedir.

Tablo 1'in devamı

(Santos, 2023)	Navoneel Chakrabarty ve Eindhoven Teknoloji Üniversitesi (TU/e), kaggle'daki açık verisetleri	VGG-16	VGG-16 modeli, doğrulama setinde yaklaşık %88 ve test setinde %80 doğrulukla beyin tümörü görüntülerinin sınıflandırılmasında yüksek doğruluk göstermektedir.
(Speet, 2020)	BraTS 2019	SegCaps (CapsuleNet)	Kapsül tabanlı bir mimarinin U-Net'ten daha düşük performans gösterdiği sonucuna varılmaktadır.
(Soumik ve Hossain, 2020)	Figshare (CE-MRI)	softmax sınıflandırıcı yöntemi kullanan önceden eğitimli bir Inception V3 modeli	Beyin tümörü sınıflandırması için en iyi sınıflandırma doğruluğunu göstermesine karşın, tümörsüz beyin MRI'ları işleme de başarısız olmaktadır.
(Tamilarasi ve Gopinathan, 2021)	REMBRANDT	Inception V3	Ortalama %95,1 doğruluk, %96,2 duyarlılık ve %94 özgünlük ile kullanışlılığını göstermektedirler.
(Afshar vd., 2018)	üç beyin tümörü tipinden biri tanısı alan 233 hastadan alınan 3064 görüntüden oluşan bir beyin MRI	CapsNet	Önerilen yaklaşımın önceki CapsNet'ler ve CNN'lerle karşılaştırıldığında sınıflandırma doğruluğunu artırabileceğini göstermektedir.

Tablo 1'in devamı

(Walsh vd., 2022)	BITE ve 3d'den 2d'ye dönüştürülen image'lar	Lightweight U-Net	Çalışma sonucunda mevcut bir derin öğrenmesi olan UNet'in beyin tümörü için nasıl değiştirilebileceği gösterilmektedir.Ağ, tüm veri kümesi üzerinde eğitildiğinde ortalama %84'lük bir ortalama IoU sağlamaktadır.
(Nasim vd., 2023)	BraTS 2017, BraTS 2018, BraTS 2019, BraTS 2020	2D Unet	Segmentasyon için farklı CNN veya değiştirilmiş CNN modelleri, geleneksel makine öğrenimi modellerinden daha iyi sonuçlar vermektedir. 2D Unet'in daha iyi bir skor elde etmesine karşın derinlemesine bilgiden yararlanılmaması nedeniyle bilgileri kaybettiği ortaya konulmaktadır.
(Kokkalla vd., 2021)	3064 T1-weighted MRI	Inception ResNet v2 (deep dense network ve softmax katmanlı)	Önerilen modelin mevcut modelden %99,69 ortalama doğrulukla daha iyi performansa sahip olduğu göstermektedir.

Tablo 1'in devamı

(Casamitjana vd., t.y.)	BRATS 2015	3DNet	Modelin yoğun çıkarım yapma, yani tek geçişte tüm hacimler için tahminler yapma kapasitesine sahip olduğu öne sürülmektedir. Ek olarak, eğitilebilir yukarı örnekleme katmanlarının kullanılması, bellekte veya hesaplama maliyetinde bir artış olmadan etkin toplu iş boyutunu arttırmaktadır.
(Çetiner ve Metlek, 2023)	BraTS 2021, FeTS2021	DenseUNet+	Özellikle UNet mimarilerinde sıklıkla karşılaşılan degrade kaybolma sorunu, önemli ölçüde ortadan kaldırılmaktadır. BraTS2021 üzerinde yapılan diğer çalışmalarla kıyaslandığında daha iyi bir başarım elde etmektedirler.
(Zheng vd., 2022)	Brain Tumor MRI Image Classification (Kaggle)	SCU-Net	Deneysel sonuçlar, önerilen yöntemin hedef segmentasyonunu ve tümör tahmin performansını önemli ölçüde artırabileceğini göstermektedir.

Tablo 1'de, yapılan literatür taramasında MRI sınıflandırması ve tespiti ile ilgili uygulanan farklı yöntemler ve kullanılan teknikler özet olarak yazılmıştır. Her bir çalışmada farklı bir yaklaşım öne sürülmüştür. Bu tez kapsamında medikal alanda kullanılan mimariler arasında en popüler olanların karşılaştırılması amaçlanmıştır. Yapılan literatür taraması sonucunda beyin tümörü üzerinde en çok kullanılan algoritmalardan VGG-19, ResNet, U-Net, InceptionV3 algoritmalarının performans değerlendirmeleri karşılaştırılmıştır. Özgünlük, hassasiyet(precision), sensitivity(hassasiyet), dice benzerliği, f1 skaru, mean iou, doğruluk, kayıp gibi metriklerin değerleri tespit edilmiştir. Bu tespit edilen metriklerden, performans değerlendirmesinde derin sinir ağlarında kullanılan ve performans ölçümü açısından doğru tespit için yardımcı olabilecek; dice benzerliği, mean iou, f1 skor, doğruluk gibi metrikler kullanılarak değerlendirme yapılmıştır. Göz önüne alınan bu metrikler sonucunda, karşılaştırılan mimariler arasında verilen veri seti ve epoc sayısı kısıtları da göz önünde bulundurularak, yüksek değerlere sahip olan derin öğrenme mimarisi tespit edilmiştir.

ÜÇÜNCÜ BÖLÜM
ARAŞTIRMA YÖNTEMİ/MATERYAL VE YÖNTEM

3.1. Materyal

Bu tez kapsamında medikal alanda kullanılan algoritmaların doğruluk oranları karşılaştırılmaktadır. Karşılaştırma için Kaggle sitesinden alınmış beyin tümörü MRI görüntülerinden oluşan BraTS 2020 veri seti kullanılmaktadır (Kaggle,2020).

Bu çalışma kapsamında veri setine karar verirken öncelikli olarak medikal alanda erken teşhisin önemli olduğu hastalık türleri araştırılmaktadır. Araştırma sonucunda beyin tümöründe erken teşhisin önemli olduğu görülerek bu alanda daha önce yapılmış çalışmalar incelenmektedir. Medikal alanda en çok kullanılan mimariler tespit edilmektedir. Daha sonra bu mimarilerden beyin tümörü tespitinde en çok kullanılanlar arasında doğruluk karşılaştırılması yapılmasına karar verilerek, bu modellerin mimari yapıları araştırılmaktadır.

Oluşturulacak modellerin tanımlanması, eğitilmesi ve tahminde bulunabilmesi için kodlama işlemi python ve tensorflow kullanılarak yapılmaktadır. Python, makine öğrenmesi ve veri analizi için geniş bir kütüphane ekosistemine sahiptir. Pandas, NumPy, Scikit-learn gibi kütüphaneler veri manipülasyonu, ön işleme ve modelleme için kullanılmaktadırlar. Python'ın basit ve okunabilir sözdizimi, öğrenme sürecini kolaylaştırmaktadır. Tensorflow, makine öğrenimi ve derin öğrenme uygulamaları için güçlü bir açık kaynaklı kütüphanedir. Graf tabanlı hesaplama yapısı, hesaplamaların optimize edilmesine yardımcı olmaktadır. Python ve TensorFlow birlikte modüler bir yapı sunarak farklı işlevlerin kolayca birleştirilip kullanılabilmesini sağlamaktadır. Bu avantajlar göz önüne alınarak bu çalışmada python, tensorflow kullanılmaktadır.

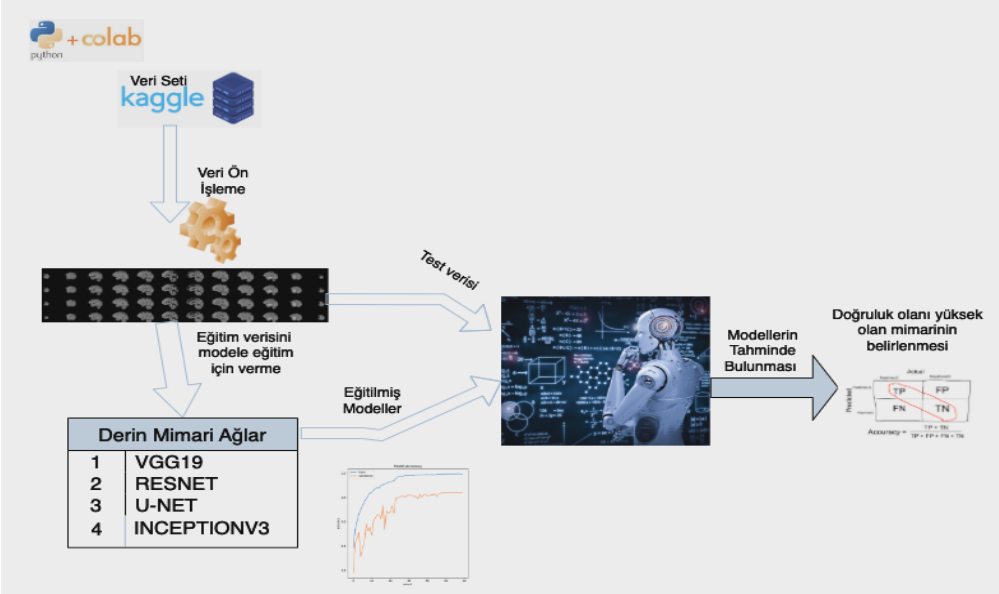
Büyük veri setlerini işlemek ve derin öğrenme modellerini eğitmek için yüksek hesaplama gücünden faydalanmak gerekmektedir. Google Colab, GPU ve TPU gibi yüksek performanslı donanım kaynakları sağlamaktadır. Fakat büyük veri setlerine sahip olan modellerin eğitimi uzun sürdüğü için, Colab yerine Colab pro plus kullanımı tercih

edilmektedir. Google Colab pro plus çevrimdışı ve 24 saat çalışabilirlik açısından Colab'a göre daha avantajlı olmaktadır. Bu nedenle Colab'ın paralı versiyonu olan Colab pro plus bu kodların çalıştırılması için kullanılmaktadır. Her ne kadar Colab pro plus kullanılsa da burada veri setinin büyüklüğü göz önünde bulundurulduğunda eğitimin uzun zaman alması nedeniyle epoch sayısında kısıtlamalar söz konusu olmaktadır.

Veri seti incelemesi yapılmış, incelenen veri setindeki veriler, modelleri eğitebilmek için ön işlemden geçirilmiştir. Veri işleme sürecinden sonra model eğitimi gerçekleştirilmiştir. Model eğitildikten sonra doğruluk oranları ile ilgili grafikler elde edilmiştir. Veri, modelin tahmin (predicted) süreci için işlenmiştir. İşlenen veri modelin tahminde bulunması için modele verilmiştir. Çıkan sonuçlar analiz edilerek değerlendirme yapılmıştır.

3.2. Yöntem

Bu çalışmada takriben izlediğimiz adımları Şekil 2.'de görebilirsiniz. Çalışmada modellerin tanımlanması ve eğitim süreçlerinin gerçekleştirilmesi, verilerin işlenmesi için python ve colab kullanılmıştır.



Şekil 2. Çalışmanın takriben akış şemasıdır.

3.2.1. Veri Seti

Bu çalışmada veri seti olarak kaggle sitesinden alınmış BraTS veri seti kullanılmaktadır (Kaggle,2020). Bu veri seti 369 hastanın MRI görüntüsünü içermektedir. BraTS veri setleri her zaman MRI görüntüleri üzerindeki çalışmalara değer katmak üzerine odaklanmaktadır. BraTS 2020, birden fazla kurumdan alınmış preoperatif MRI taramalarını kullanmaktadır. Temel olarak gliomlarda odaklanmaktadır.

Veri seti, dört farklı hacim üzerinde ayrıntılandırılmaktadır:

A. Native T1: Beyindeki farklı doku türlerinin yapısını göstermek için kullanılır, tümörleri, kistleri ve diğer farklılıkları belirlemek için kullanılmaktadır.
B. Post-contrast T1-weighted(T1Gd/T1ce): Kontrast ajanı enjeksiyonu ile anormalliklerin görünürlüğünü arttırmak için yapılmaktadır.
C. T2-weighted (T2): Beyindeki farklı dokuların sıvı içeriğini göstermek için kullanılmaktadır.
D. T2 Fluid Attenuated Inversion Recovery (T2-FLAIR): T1 ve T2 görüntülerinde net olarak görülmeyen lezyonların tanımlanmasında kullanılmaktadır.

Veri seti içindeki veriler, NIfTI (.nii.gz) dosyası şeklindedir. NIfTI görüntüleri, 3 boyutlu bir nesnenin dijital bir temsilidir. Görüntülerin shape'ini kontrol ettiğimizde 3 boyutlu oldukları görülmektedir. Bu ölçüler genişlik, yükseklik ve derinlik boyutlarına karşılık gelmektedir. İnsan vücudunda ise aksiyel/diagonal,koronal, sagittal düzlemlere denk gelmektedirler.

3.2.2 Derin Öğrenme Mimarileri

Bu çalışma medikal alanda beyin tümörü üzerinde en çok kullanılan derin öğrenme ağları olan Vgg19, Resnet, U-Net, InceptionV3 mimarilerinin karşılaştırılması yapılmaktadır. Bu bölümde kısaca bu derin ağ mimarileri açıklanmaktadır.

Vgg19

Oxford Üniversitesi'ndeki Visual Geometry Group (VGG) tarafından geliştirilen derin öğrenme modellerinden biridir. Bu model, İmageNet veri kümesi üzerinden eğitilmiş ve genellikle evrişimli sinir ağlarının bir örneği olarak kullanılmaktadır (Simonyan ve Zisserman, 2015).

Mimarinin ana özellikleri şunlardır;

Derinlik: "19" adı verilen isim, modelin 19 katmanlı olduğunu belirtmektedir. 16 katmanlı VGG'nin üzerine üç katman daha eklenerek bu derin ağ mimarisi oluşturulmaktadır.

Convolution Layers: ardışık evrişimli katmanlardan oluşmaktadır. 3x3 çekirdek boyutlarına ve 1 piksellik adımlara sahip bir dizi evrişimli katman bulunmaktadır. Bu katmanlar öğrenme kapasitesini arttırmaktadır.

Pooling Layers: Bu katmanlar özellik haritasının boyutunu küçültmek ve hesaplama yükünü azaltmak için kullanılmaktadır.

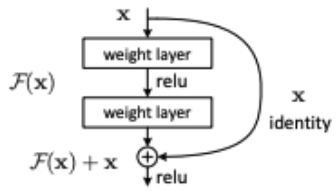
Fully Connected Layers: Özelliklerin karmaşıklığını arttırarak, sınıflandırma işlemi için kullanılmaktadır.

ReLU Activation: Ağın daha hızlı eğitilmesi ve aşırı uyumluluğa karşı daha dirençli olmasını sağlamaktadır.

ResNet

ResNet (Residual Neural Network), Microsoft tarafından 2015 yılında tanıtılan bir derin öğrenme mimarisidir. ResNet, aşırı derin sinir ağlarının eğitimini kolaylaştırmak için tasarlanmıştır ve derin ağlarda ortaya çıkan geriye doğru yayılma sırasında gradyan kaybı sorununu azaltmak için geliştirilmiştir (He vd., 2015). ResNet'in en belirgin özelliği, ağın içine "atlamalı bağlantılar" (skip connections) veya "artık bağlantılar" (residual connections) olarak adlandırılan doğrudan bağlantılar eklemesidir. Normalde bir katmanın çıktısı, bir sonraki katmana giriş olarak kullanılırken, atlamalı bağlantılarla, bir katmanın çıktısı doğrudan birkaç katman sonrasındaki çıktıya eklenir. Şekil 3'de artık bağlantı görülmektedir. Bu, ağın daha derin hale gelmesine olanak tanırken, gradyanların daha

verimli bir şekilde geriye yayılmasını sağlayarak eğitimi kolaylaştırır. ResNet ayrıca farklı blok türlerini kullanır. Özellikle, "identity block" ve "convolutional block" adı verilen iki temel blok türü vardır. Identity block, giriş ve çıkış boyutları aynı olan katmanlardan oluşurken, convolutional block, boyutları değiştiren ve çıkış katmanını içeren bir yapıya sahiptir. Bu bağlamda, ResNet, derin sinir ağlarının eğitiminde aşırı derinlik ve gradyan kaybı sorununu azaltarak daha başarılı sonuçlar elde etmek için önemli bir adım olarak kabul edilmiştir.

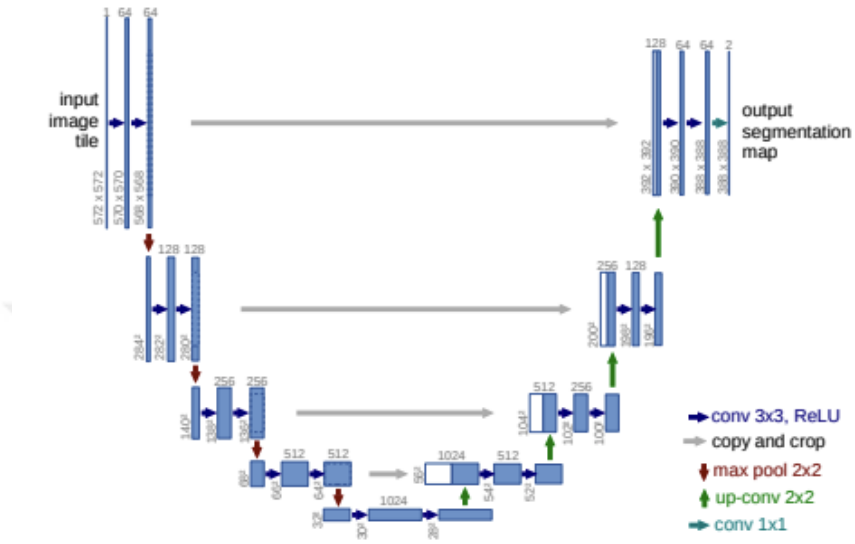


Şekil 3. Artık bağlantı.(He vd., 2015)

U-Net

U-Net mimarisi, biyomedikal görüntü segmentasyonu için geliştirilmiş bir derin öğrenme modelidir. Hem konvolüsyonel ağlar hem de “encode-decoder” yapılarını birleştirmektedir. U-Net’in ana özelliği, girdi görüntüsünün tam boyutunda bir çıktı üretirken, ağın derinlik seviyesini korumasını sağlayan simetrik bir yapıya sahip olmasıdır. Encoder kısmı, girdi görüntüyü daha küçük bir boyuta indirgerken, decoder kısmı bu küçültülmüş öznitelik haritasını tekrar girdi boyutuna getirmektedir. U-Net mimarisi, genellikle encoder kısmında konvolüsyonlarla öznitelikleri çıkaran, ardından her katmanda boyutu küçülten ve öznitelik haritasını sıkıştıran evrişim katmanlarından oluşmaktadır. Decoder kısmında ise öznitelik haritası tekrar genişletilir ve orjinal boyuta geri getirilmektedir. Encoder ve decoder arasında, öznitelik haritalarının birleştirilmesi için atlamalı bağlantı bulunmaktadır. Şekil 4’de Unet mimarisinin bağlantılarını görebilmekteyiz. Bu bağlantılar, daha düşük seviyedeki özniteliklerin decoder tarafında daha yüksek seviyedeki özniteliklerle birleştirilmesini sağlamaktadır. Bu sayede, daha detaylı lokalizasyon ve segmentasyon yapılabilmektedir (Ronneberger vd., 2015).

U-Net, özellikle medikal görüntü segmentasyonunda, organların,dokuların veya lezyonların hassas bir şekilde ayrıştırılması için yaygın olarak kullanılmaktadır.



Şekil 4. U-net mimarisi (en düşük çözünürlükteki 32x32 piksel için örnek) (Ronneberger vd., 2015).

Her mavi kutu, çoklu kanallı bir özellik haritasına karşılık gelmektedir. Kanal sayısı kutunun üstünde belirtilir. x-y boyutu kutunun sol alt köşesinde sağlanmaktadır. Beyaz kutular kopyalanmış özellik haritalarını temsil etmektedir. Oklar farklı işlemleri göstermektedir (Ronneberger vd., 2015).

InceptionV3

InceptionV3, Google tarafından geliştirilen ve 2015'te tanıtılan bir evrişimli sinir ağı mimarisidir. Bu mimari, derin ve evrişimli sinir ağlarında ağın hesaplama yoğunluğunu azaltmak ve daha verimli bir yapı elde etmek için geliştirilmiştir. İnceptionV3, önceki İnception modellerine kıyasla daha derin ve karmaşık bir yapıya sahiptir. Ağ, karmaşık yapıları daha etkili bir şekilde öğrenmek için çeşitli evrişim katmanlarını ve paralel yapıları kullanılmaktadır. Şekil 5'de InceptionV3 modelinin mimari şeması gösterilmiştir (Szegedy vd., 2015).

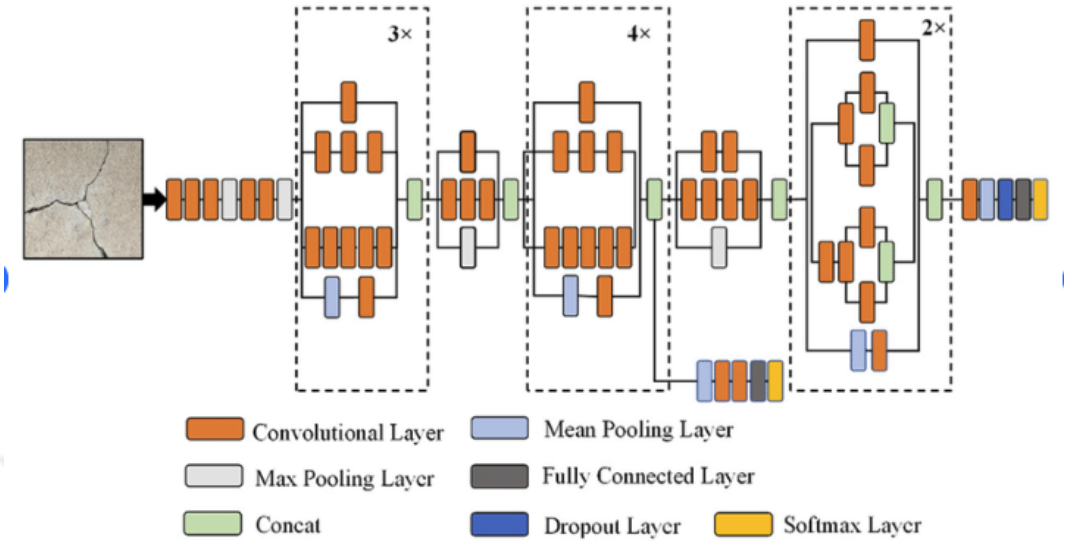
İnception Modülü: İnceptionV3, farklı evrişim çekirdek boyutlarını ve ölçekleri bir araya getiren "İnception" adı verilen modülleri kullanmaktadır. Bu modüller, paralel olarak çalışan farklı boyutlardaki filtrelerden elde edilen çıktıları birleştirmektedir. Bu, ağın aynı ölçekte hem küçük hem de büyük özellikleri yakalamasını sağlamaktadır.

Azaltma Blokları(Reduction blocks): İnceptionV3, evrişim katmanlarının arasına yerleştirilen azaltma blokları sayesinde boyut azaltma ve özellik haritası boyutunu küçültme işlemlerini gerçekleştirmektedir. Bu bloklar, ölçeklenebilir ve hesaplama yoğunluğunu azaltıcı yapılar içermektedir.

Yardımcı sınıflandırıcı(Auxiliary Classifiers): Eğitim sırasında kaybolan gradyanları geri getirmek için eğitimde ek bir yardımcı sınıflandırma yolu kullanılmaktadır. Bu, daha derin ağlarda kaybolan gradyanları izlemeye ve eğitimi daha kolay hale getirmeye yardımcı olmaktadır.

Transfer Learning Yetenekleri: Önceden eğitimli İnceptionV3 ağı, farklı görevler için transfer learning amaçlarıyla kullanılabilmektedir. Özellikle, görüntü sınıflandırma, nesne tanıma ve yerelleştirme gibi görevlerde başarılı sonuçlar vermektedir.

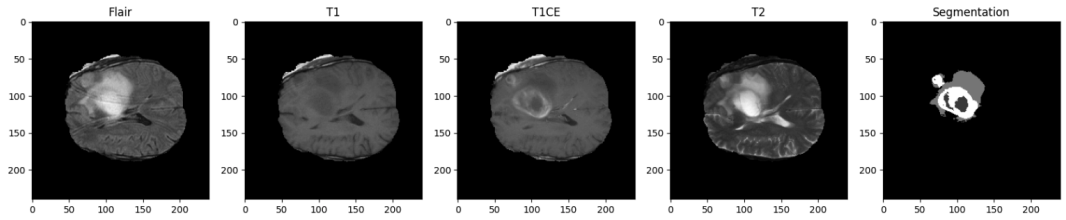
İnceptionV3, bilgisayarlı görü, nesne tanıma ve diğer görsel görevlerde etkili performans gösteren derin öğrenme modellerinden biridir. Hem hesaplama verimliliği hem de doğruluk açısından yüksek sonuçlar elde edilebilmektedir ve genellikle ölçeklenebilmektedir ve genişlemiş yapıları nedeniyle tercih edilebilmektedir.



Şekil 5. InceptionV3 modelinin mimarisi (Ali vd., 2021).

3.2.3 Veri Ön İşleme

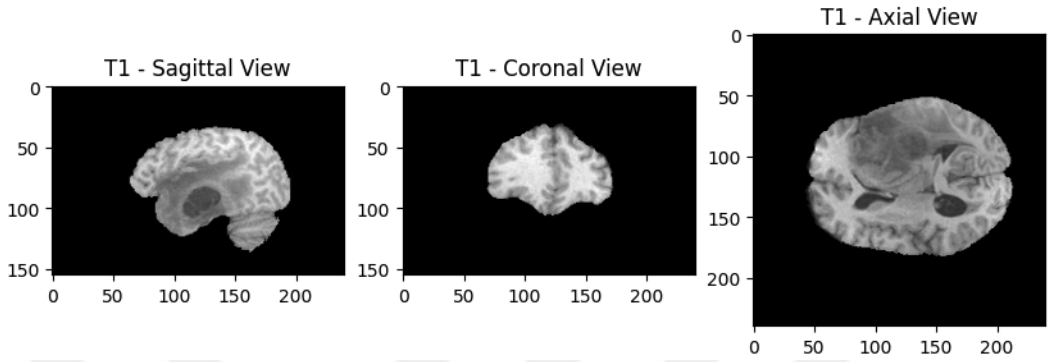
Veri yolunu belirlediğimiz verilerimizi "nibabel" kütüphanesi kullanarak veri seti içerisindeki T1, T1CE, T2, FLAIR ve segmentasyon görüntüleri yüklenmektedir. Şekil 6'da bu sınıflar arasındaki fark net bir şekilde görülmektedir. Bunun için "nib.load" fonksiyonu kullanılarak yüklenip değişkenlere atanmaktadır. "get_fdata()" fonksiyonu kullanılarak her bir MRI modalitesi ve segmentasyon görüntüsünün veri kısmı alınıp değişkenlere atanmaktadır.



Şekil 6. Flair, T1, T1CE, T2 ve Segmentation sınıflarına ait image görüntüleri.

Verilerin 3 boyutlu görüntülerini düzlem üzerinde görselleştirmek için "rotate()" fonksiyonu kullanılarak, belirlenen görüntünün sırasıyla sagittal (yan), koronal (ön) ve

aksiyal (üst) düzlem üzerindeki görüntüsü incelenmektedir. Şekil 7'da MRI'ın 3 farklı düzlemdeki görüntüsü gösterilmiştir.



Şekil 7. MRI ın 3 farklı düzlem görüntüsü.

Veri setindeki benzersiz değerleri elde edebilmek için veri seti gezilerek her bir verinin segmentasyon görüntülerinin dosya yolu oluşturulmaktadır. "os.listdir()" fonksiyonuyla veri setindeki dosya isimleri alınırken .csv dosyaları hariç tutularak sadece verinin alınması sağlanıyor. "nib.load" vd "get_fdata" fonksiyonu for döngüsün içinde kullanılarak benzersiz değerler bulunup(np.unique()), liste oluşturulmaktadır. Bu listeye göre segmentation sınıfının 4 farklı etiket değeri vardır.

Veri setini hazırlamak için data generator sınıfı tanımlanıp, Keras'ın Sequence sınıfıyla genişletilmiştir. Bu sınıf, beslenecek veri gruplarını hazırlamak için kullanılmaktadır. Veri seti "train_test_split()" fonksiyonu ile train, test ve validation verilerine ayrılmaktadır. Veri seti dağılımı %68 eğitim verisi, %12 test verisi ve %20 doğrulama verisi olarak ayrılmıştır. Son olarak, bu veri oluşturucuları model eğitimi için kullanılmıştır. Verileri belirlenen oranlar doğrultusunda bölerek ilgili alanlar için veri oluşturma işlemini yapmak için training_generator, valid_generator ve test_generator fonksiyonları kullanılmıştır. Sırasıyla eğitim, doğrulama ve test veri setleri için oluşturulmuştur.

Veri setinin parametre olarak verildiği bir görüntü veri hazırlama fonksiyonu tanımlanmıştır. Modelin tahmin etmesi gereken veri boyutlarına ve kanal sayısına sahip bir

dizi oluşturulmuştur. Veri setindeki farklı türlerin (flair,t1ce ve segmentation) veri yolu tanımlanmıştır. Bu tanımlarla beraber nifti formatındaki veriler NumPy dizilerine dönüştürülmektedir. Her bir görüntü yeniden boyutlandırılarak tanımlanan diziye yerleştirilmiştir. Böylece modelin tahmin etmesi gereken formata uygun bir hale getirilerek test ve tahmin aşamasında kullanılabilmiştir.

3.2.4 Model Tanım ve Eğitimi

Resnet, U-Net, Inceptionv3, Vgg19 derin öğrenme ağlarının karşılaştırmasını yapabilmek adına, model tanımı yaparken aktivasyon fonksiyonunun aynı olmasına önem verilmektedir. Aktivasyon için softmax fonksiyonu kullanılmaktadır.

Softmax, çok sınıflı sınıflandırma problemlerinde kullanılan bir aktivasyon fonksiyonudur. Bu fonksiyon, modelin çıktılarını olasılık dağılımına dönüştürmektedir. Softmax, çıktı katmanındaki her bir nöronun çıktısını, [0,1] aralığında ve tüm çıktıların toplamının 1 olduğu bir olasılık dağılımı olarak yorumlamaktadır. Softmax Fonksiyonu 1.1'de verilmiştir.

$$\text{Sigma}(z)_i = e\text{^}(z_i) / \Sigma(j=1)\text{^}(K)\ e\text{^}(z_j) \qquad (1.1)$$

Bu formül K adet sınıfın aktivasyon değerlerini içeren bir vektörü (z) alır ve her bir sınıfın olasılığını hesaplamak için bu değerleri ölçekler. Formülde, z_i, i. sınıfın aktivasyon değerini temsil eder ve e^(z_i) ifadesi, bu aktivasyon değerini e tabanında üs olarak hesaplar. Ardından, bu hesaplanan aktivasyon değerleri, tüm sınıfların aktivasyon değerlerinin toplamına bölünerek normalleştirilir.

Σ(j=1)^(K) e^(z_j) ifadesi, tüm sınıfların aktivasyon değerlerinin e^(z_j) toplamını temsil eder. Her bir sınıfın aktivasyon değeri, tüm sınıfların toplam aktivasyon değerine bölünerek normalleştirilir. Bu, her bir sınıfın olasılığını ifade eden bir dağılım elde etmek için kullanılır.

Sonuç olarak, sigma(z)_i ifadesi, i. sınıfın hesaplanmış olan olasılığını temsil eder ve bu şekilde tüm sınıflar için olasılık dağılımı elde edilir. Bu fonksiyon, sınıflar arasında yüksek aktivasyon değerlerine sahip olan sınıfların daha yüksek olasılıklara sahip olmasını sağlar.

Kısaca softmax fonksiyonu, her bir sınıfın çıktısını, tüm sınıfların çıktılarının toplamına oranlayarak normalize etmektedir. Bu şekilde, her bir sınıfın tahmini olasılık değerini elde etmekteyiz. Softmax genellikle çok sınıflı sınıflandırma problemlerinde kullanılır çünkü bu tür problemlerde her bir sınıfın olasılığını belirlemek istenir ve softmax çıktılarını bu olasılıklara dönüştürmektedir. Örneğin, bu çalışmada gerçekleştirdiğimiz gibi görüntü segmentasyonunda her bir pikselin hangi sınıfa ait olduğunu belirlemek için kullanılmaktadır.

Modelin çıktı sınıf sayısı, değişken olarak model oluştururken verilmektedir. Bu çalışmada n_classes değeri 4 dür, 4 farklı sınıf için bir segmentasyon problemidir. Modelin giriş şeklini (boyutlarını) belirtmekteyiz. Modelin nasıl eğitileceğini "model.compile" ile belirledikten sonra eğitim sırasında belirli metrikleri yakalayabilmek adına callback fonksiyonu tanımlıyoruz ve "model.fit" fonksiyonu ile modelleri daha önce hazırladığımız verilerle eğitiyoruz.

3.2.5 Model Tahmin ve Test

Model eğitimini tamamladıktan sonra artık bu aşamada test veri setini kullanarak segmentasyonları tahmin ettireceğiz. Bu noktada model eğitim verilerine bağlı olarak en iyi ağırlığı kullanabiliriz ya da son ağırlığı kullanarak test edebiliriz. Normal şartlarda her zaman son ağırlık iyi sonuçlara sebep olur anlamına gelmemektedir. Bunu da göz önünde bulundurarak model eğitimi sonuçlarını incelemek ve doğru ağırlığı kullanmak önemlidir. "load_weights" fonksiyonu ile bu aşamada kullanacağımız ağırlıkları modelimize yüklüyoruz.

Daha önce veri işlemede bahsettiğimiz tahmin için veri işleme adımını gerçekleştirip test verilerimizi modele tahmin etmesi için yüklüyoruz. Böylece model analizi yapabiliyoruz. Modelimizin performansını test veri seti üzerinde değerlendirebilmek için "evaluate()" fonksiyonunu kullanarak modellerin doğruluk ve kayıp oranlarını buluyoruz.

3.2.6 Değerlendirme

Değerlendirme aşamasında uygulanan derin öğrenme algoritmalarının yaptığı tahminler doğrultusunda ve evaluate sonuçlarındaki metriklerin ,doğruluk oranı (accuracy score), mean iou (Ortalama intersection over Union), f1-skor, dice benzerliği değerlendirilmesiyle performanslı olan belirlenmektedir. En iyi sonucu veren derin öğrenme algoritması diğer bir ifadeyle model, bu aşamada değerlendirilerek seçilir. Çalışmada, uygulanan derin öğrenme algoritmalarının verdiği, doğruluk oranı (accuracy score), mean iou (Ortalama intersection over Union), f1-skor, dice benzerliği metrikleri birbirleriyle karşılaştırılmıştır. Çalışmanın kısıtları ve sonraki çalışmalarda nasıl bir yol izleneceği belirlenmektedir.

DÖRDÜNCÜ BÖLÜM
ARAŞTIRMA BULGULARI

Bu çalışmada son durum olarak modellerin performansının bir değerlendirmesi sunulmaktadır. Modellerin performansını ölçmek için doğruluk, f1-score gibi metrikler kullanılmaktadır. Modellerin performansı flair ve t1 MRI'ları üzerinde incelenmiştir. Bu bölümde genel olarak modellerin performansı ve birbirleri ile karşılaştırıldığında güçlü ve zayıf yönleri sunulmaktadır.

4.1. Model Eğitim Metrikleri

BraTS 2020 veri setini ön işleme adımın veri; eğitim, doğrulama ve test aşamaları için 3'e ayrılmıştır ve bu ayrılmış veriler ilk olarak model eğitimi sonrasında da test için kullanılmıştır. Verileri ayırırken doğrulama verisini kullanmamızın nedeni overfitting dediğimiz aşırı uyumluluk kontrolünü sağlayabilmektir. Eğitim ve test aşamaları sonrasında karşılaştırılan mimarilerin doğruluk ve kayıp fonksiyon yüzdelik değerleri Tablo 2 de gösterilmektedir. Tabloda değerleri verilen metrikler, eğitim doğruluğu, eğitim kaybı, doğrulama doğruluğu, doğrulama kaybı, test doğruluğu, test kaybıdır. Bu değerlerin kısaca açıklamaları aşağıdaki gibidir:

Eğitim Doğruluğu (Training Accuracy): Bir modelin eğitim veri seti üzerinde ne kadar doğru tahmin yaptığını ölçen bir metriktir. Modelin eğitim verilerine ne kadar iyi uygulandığını gösterir. Genellikle bir oran veya yüzde olarak ifade edilir.

Eğitim Kaybı (Training Loss): Modelin eğitilirken yaptığı tahminler ile gerçek değerler arasındaki farkın bir ölçüsüdür. Eğitim boyunca modelin ürettiği tahminler ile gerçek etkiler arasındaki fark, kayıp fonksiyonu aracılığıyla hesaplanır.

Doğrulama Doğruluğu (Validation Accuracy): Modelin ayrılmış doğrulama veri seti üzerinde ne kadar doğru tahmin yaptığını ölçer. Modelin eğitim veri setine aşırı uyum sağlaması durumunda (overfitting), genelleme yeteneği düşebilir. Bu nedenle, modelin

gerçek dünyadaki performansını daha objektif bir şekilde değerlendirebilmek adına ayrı bir doğrulama seti kullanılır.

Doğrulama Kaybı (Validation Loss): Bir modelin tahminlerinin gerçek değerlerden ne kadar farklı olduğunu ölçer. Doğrulama setindeki veriler üzerinden yapılan tahminlerin ne kadar doğru ne kadar yanlış olduğunu ölçen fonksiyondur.

Test Doğruluğu (Test Accuracy): Modelin test veri setinde ne kadar doğru ölçüm yaptığını ölçer. Test seti, modelin daha önce görmediği, eğitim ve doğrulama süreçlerinde kullanılmamış verilerden oluşur. Eğitim ve doğrulama süreçlerinde iyi performans gösteren bir modelin test setinde de iyi performans göstermesi, modelin genelleme yeteneğinin güçlü olduğunu göstermektedir.

Test Kaybı (Test Loss): Modelin test veri setindeki tahminlerin ne kadar doğru veya yanlış olduğunu ölçer. Modelin gerçek değerlerle karşılaştığında ne kadar yanıldığını ölçer.

Tablo 2

Modellerin test verisi üzerindeki değerlendirme sonucu doğruluk ve kayıp fonksiyon metrikleridir.

Model	Eğitim Doğruluğu	Eğitim Kaybı	Doğrulama Doğruluğu	Doğrulama Kaybı	Test Doğruluğu	Test Kaybı
VGG19	0.9873	0.5890	0.9863	0.6198	0.9863	0.6492
RESNET	0.9900	0.4153	0.9887	0.5495	0.9884	0.5766
U-NET	0.9815	0.0625	0.9805	0.0665	0.9801	0.0618
INVEPTIONV3	0.9915	0.3388	0.9897	0.4632	0.9904	0.4698

Doğruluk doğruluğu olarak ifade edilen terim, modelin eğitim sürecinde ayrılan doğrulama seti ile genellikle modelin ne kadar iyi genelleme yaptığını değerlendirerek ve aşırı uyum (overfitting) kontrolü yapmak için kullanılmaktadır.

Overfitting, bir modelin eğitim verilerine aşırı derecede uyum sağlaması ve bu nedenle yeni, görülmemiş verilere karşı genelleme yapma yeteneğinin düşmesi durumunu ifade eder. Böyle durumlarda eğitim verileri üzerinde yüksek doğruluk gösterirken, yeni verilerde performansı düşebilir. Overfitting'i kontrol etmek için çeşitli yöntemler kullanılır (Tukey, 1962). Bu yöntemler kısaca:

Eğitim-Doğrulama Bölmesi: Veri seti eğitim ve doğrulama olarak ikiye ayrılır. Model eğitim seti üzerinde eğitilirken, doğrulama seti üzerinde doğruluk kontrol edilir. Doğrulama setindeki performans düşüşü, overfitting'in bir işareti olabilir.

Cross-Validation (Çapraz Doğrulama): Veri seti, farklı alt kümelerde eğitim ve doğrulama olarak bölünerek modelin farklı veri kümeleri üzerinde nasıl performans gösterdiği değerlendirilir. Bu, genel model performansının daha iyi değerlendirilmesini sağlar.

Düzenleme (Regularization): Modelin karmaşıklığını azaltmak için kullanılır. L1 ve L2 düzenlemesi gibi teknikler, modelin aşırı uyum yapmasını önlemek için kullanılan parametrelerin değerlerini sınırlar.

Dropout: Rastgele nöronları eğitim sırasında devre dışı bırakarak aşırı uyumu azaltır.

Overfitting'i belirlemek için, genellikle eğitim ve doğrulama (validation) setleri arasındaki doğruluk ve kayıp (loss) değerlerine odaklanılır. Eğer model eğitim setinde yüksek doğruluk gösterirken, doğrulama setindeki performansı düşükse, overfitting olma

ihtimali vardır. VGG19 modeli için değerleri kontrol ettiğimizde eğitim setinde ve doğrulama setinde yüksek doğruluk oranları ve benzer kayıp değerleri gözlemlenmektedir. Bu durum overfitting olmadığını göstermektedir. ResNet modeli için değerleri kontrol ettiğimizde farklar oldukça küçük ve genellikle bir modelin overfitting olmadığını göstermektedir. Unet ve InceptionV3 için eğitim ve doğrulama setleri arasındaki doğruluk ve kayıp değerleri birbirine oldukça yakındır. Bu durumda verilen değerlere dayanarak Unet ve InceptionV3 içinde belirgin bir overfitting gözlemlenmemiştir. Bu çalışma kapsamında kullanılan modellerin karşılaştırılmasında overfitting olmadığı görülmüştür ancak, overfitting'i tam olarak değerlendirmek için daha geniş bir veri seti kullanılması gerekebilir.

4.2. Değerlendirme Metrikleri

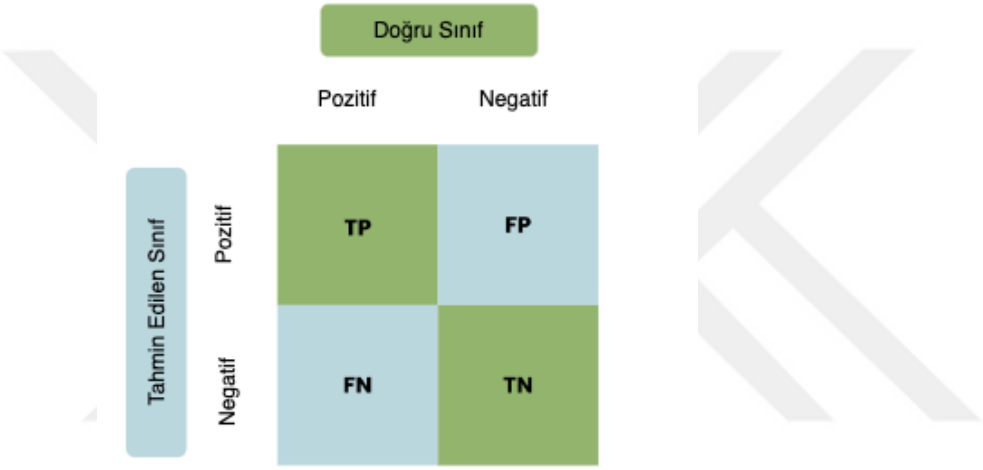
Modellerin doğruluğu, kayıpları ve performans parametreleri sonuç analizinde incelenmiştir. Performans analizine dayanarak, doğruluk, F1 skoru ve diğer ilgili parametreleri dikkate alan optimal model seçilmiştir. Beyin tümörü görüntüleri üzerinde yapılan değerlendirme sayesinde derin öğrenme modellerinin hangisinin daha performanslı olduğu bulunabilmektedir. Derin öğrenme mimarilerinin performans değerlendirmesinde yaygın olarak kullanılan bazı metrikler bulunmaktadır. Bunlar; hassasiyet (precision), F1 skoru, doğruluk (acc), duyarlılık (sensitivity) ve özgünlük (specificity), ortalama üzerinde kesişim metrikleri olmaktadır (Goutte ve Gaussier, 2005; Voulodimos vd., 2021). Bu metriklerin formülleri gerçek pozitif, gerçek negatif, yanlış negatif ve yanlış pozitif ile ilişkili olmaktadır (Google, n.d).

TP (True Positive - Gerçek Pozitif): Modelin doğru bir şekilde bir sınıfı belirlediği ve bu sınıfın gerçekten var olduğu durumları ifade etmektedir.

TN (True Negative - Gerçek Negatif): Modelin doğru bir şekilde bir sınıfı belirleyemediği ve bu sınıfın gerçekten bulunmadığı durumları ifade eder. Modelin negatif olarak tanımladığı ve gerçekte negatif olan durumları tanımlamaktadır.

FN (False Negative - Yanlış Negatif): Gerçekte var olan bir durumu modelin yanlışlıkla yok olarak sınıflandırması durumudur. Yani, modelin kaçırdığı veya tanımlayamadığı gerçek pozitif durumları tanımlamaktadır (Cardinal, 2016).

FP (False Positive - Yanlış Pozitif): Gerçekte olmayan bir durumu modelin yanlışlıkla var olarak sınıflandırması durumudur. Yani, modelin tanımladığı ancak gerçekte var olmayan pozitif durumları tanımlamaktadır (Cardinal, 2016).



Şekil 10. Karmaşık matrisin temel yapısıdır.

Karmaşık Matris (Confusion Matrix), bir sınıflandırma modelinin performansını değerlendirmek için kullanılan bir tablodur. Bu matris, gerçek ve tahmin edilen sınıfları karşılaştırarak modelin doğruluğunu ölçer. Şekil 10'da karmaşık matrisi gösterilmektedir.

Doğruluk

Doğruluk (accuracy), sınıflandırma problemlerinde modelin doğru tahmin ettiği örneklerin oranını ifade etmektedir. Doğruluk kavramı, modelin doğru tahmin ettiği veri noktalarının toplam veri noktalarına oranı olarak tanımlanır ve genellikle performans değerlendirmesinde kullanılır. Genellikle 0 ile 1 arasında bir değer olarak ifade edilir ve daha

yüksek bir doğruluk, daha iyi bir performansı göstermektedir. Bu ölçüm, genellikle sınıflandırma problemlerindeki modelin doğruluğunu belirlemek için kullanılır.

Doğruluk (Accuracy) = (TP + TN) / (TP + TN + FP + FN)

Dice Benzerliği

Dice benzerliği, özellikle görüntü segmentasyonunda kullanılan bir metriktir. İki kümenin (örneğin, bir modelin tahmini ile gerçek etiketler arasındaki) benzerliğini ölçmektedir. Değer aralığı 0 ile 1 arasındadır, 1'e yaklaşan bir değer daha iyi bir benzerlik anlamına gelmektedir. Özellikle sınıf dengesizliği olan durumlarda (örneğin, tıbbi görüntüleme gibi), bu metrik daha etkili olabilmektedir. Çünkü sınıflar arasındaki dengesizlik, sadece doğruluk gibi basit metriklerle değerlendirilmeye bilmektedir. Ancak, tek başına bir metrik olarak kullanıldığında, bazı durumlarda yetersiz olabilir ve diğer performans metrikleri ile birlikte değerlendirilmelidir.

Dice Benzerliği = (2 × TP) / (TP + FP + TP)

Hassasiyet

Hassasiyet (Precision), sınıflandırma problemlerinde kullanılan bir performans ölçüsüdür. Modelin belirlediği pozitif örnekle arasında gerçekten pozitif olanların ne kadarının doğru bir şekilde belirlendiğini gösterir. Yani modelin yanlış pozitif tahminler yapma olasılığın gösterir. Yüksek hassasiyet, yanlış pozitif tahminlerin az olduğunu ve modelin gerçek pozitifleri daha doğru bir şekilde belirlendiğini gösterir.

Hassasiyet (Precision) = TP / (TP + FP)

Duyarlılık

Duyarlılık (Sensitivity), gerçek pozitif oranı olarak da bilinir ve bir modelin gerçekten pozitif olan örnekleri ne kadar doğru bir şekilde tanımlandığını ölçer. Gerçek pozitiflerin ne kadarının doğru bir şekilde tanımlandığını ölçer (Monaghan vd., 2021).

Duyarlılık (Sensitivity) = TP / (TP + FN)

Özgünlük

Özgünlük (Specificity), bir sınıflandırma modelinin gerçek negatifleri doğru bir şekilde belirleme kabiliyetini ölçen bir performans metriğidir. Özgünlük, genellikle duyarlılık (sensitivity) ile birlikte kullanılır ve modelin negatif sınıfı doğru bir şekilde belirleme yeteneğini yansıtır (Monaghan vd., 2021).

Özgünlük, özellikle tıp alanında ve sınıflandırma problemlerinde kullanışlıdır. Örneğin, bir testin belirli bir durumu doğru bir şekilde dışlaması gereken bir durumda, özgünlük bu testin gerçek negatifleri ne kadar doğru bir şekilde belirlediğini gösterir. Yüksek özgünlük, modelin gerçek negatifleri doğru bir şekilde tanımladığını ve istenmeyen durumların doğru bir şekilde dışlandığını gösterir. Düşük özgünlük, modelin gerçek negatifleri yanlış pozitif olarak tanımlama eğiliminde olduğunu gösterebilir.

Özgünlük (Specificity) = TN / (TN + FP)

F1 Skoru

F1 skoru, sınıflandırma modellerinin performansını ölçen bir metriktir ve hem hassasiyet (precision) hem de duyarlılık (sensitivity) metriklerini birleştirerek modelin genel performansını değerlendirir. F1 skor veriler de var olabilecek dengesizlikler veya yanlış pozitif ve yanlış negatiflerin eşit öneme sahip olduğu bir durum var ise performans değerlendirmesi için önemli bir metrik olmaktadır. F1 skor 0 ile 1 arasındadır ve en iyi performans 1'e yakın olandır. F1 skoru, hem hassasiyetin hem de duyarlılığın iyi olduğu durumlarda yüksek bir değere sahip olur. Bu nedenle, dengeli bir şekilde hem yanlış pozitiflerin hem de yanlış negatiflerin az olması durumunda yüksek bir F1 skoru elde edilir. F1 skoru, dengeli bir performans gösteren modelleri değerlendirmede ve karşılaştırmada sıkça kullanılır. F1 skor 0 ile 1 arasındadır ve en iyi performans 1'e yakın olandır (Srivastava, 2019).

F1 skoru = (2 × Hassasiyet × Hatırlama) / (Hassasiyet + Hatırlama)

Ortalama Üzerinde Kesişim

Mean İntersection over Union (meanIoU), özellikle görüntü segmentasyonunda performansın ölçülmesi için kullanılan bir metriktir. Bu metrik, bir modelin tahmin ettiği ve gerçek etiketler arasındaki örtüşme alanını ölçmektedir. IoU, tahmin edilen bölgenin gerçek bölgeyle kesişimini gerçek bölgenin birleşimine oranlamaktadır. Değerler 0 ile 1 arasında olur, 1'e yaklaşan bir değer daha iyi bir örtüşme oranı göstermektedir. Özellikle çoklu sınıf segmentasyonu gibi alanlarda, her sınıf için ayrı IoU değerleri hesaplanır ve bu değerlerin ortalaması alınarak meanIoU hesaplanmaktadır. Modelin genel olarak tüm sınıflar için ne kadar iyi performans gösterdiğini ölçmek için kullanılmaktadır. Yüksek meanIoU değerleri, modelin nesnelerin sınırlarını daha iyi tanımladığı ve segmentasyonun daha doğru olduğu anlamına gelmektedir.

Ortalama Üzerinde Kesişim (MeanIoU) = TP / (TP + FP + FN)

Tablo 2'de modellerin test verisi üzerindeki değerlendirmeleri sonucunda ortaya çıkan metriklerin değerleri görülebilmektedir. Modelin sınıflandırma performansını ölçmek için özellikle çoklu sınıflı segmentasyon problemlerinde yaygın olarak MeanIoU metriği değerlendirilmektedir. Yüksek bir MeanIOU değeri, modelin nesneleri doğru bir şekilde tanımladığını ve bölgeler arasında iyi bir örtüşme sağladığını gösterirken, düşük bir değer modelin performansının zayıf olduğunu gösterebilmektedir. Tabi ki bu değer tek başına yeterli değildir. Dice benzerliği dediğimiz metrikte önemli metriklerden bir tanesi olmaktadır. Bu metrik iki bölgenin (gerçek ve tahmin edilen) ne kadar örtüştüğünü ölçmek için kullanılmaktadır. Özellikle piksel bazlı tahminlerin benzerliğini ölçmek için kullanılan bir metriktir. F1 skoru, bir sınıflandırma modelinin performansını değerlendirmek için kullanılan bir metriktir. Hassasiyet (precision) ve duyarlılık (recall/sensitivity) değerlerinin harmonik ortalamasıdır. F1 skoru, özellikle dengesiz sınıflandırma problemlerinde, yani sınıflar arasında veri sayısı açısından büyük farklar olduğunda daha güvenilir bir performans ölçüsü olarak kabul edilir. Bu durum, tıp alanında hastalık teşhisi gibi önemli uygulamalarda oldukça önemlidir. F1 skoru, bir modelin genel olarak ne kadar iyi performans gösterdiğini anlamak için önemli bir metrik olarak kabul edilmektedir. Tüm bu nedenler göz önüne alındığında performans değerlendirmesi için MeanIoU, F1 skor, doğruluk ve dice benzerliği kullanılmıştır.

Bahsettiğimiz değerlendirme metriklerinin bu çalışmada karşılaştırılan algoritmalar için değerleri Tablo 3 de görülmektedir.

Tablo 3

Modellerin test verisi üzerindeki değerlendirme sonucu metrikleridir.

Model	Doğruluk Accuracy (%)	Dice Benzerliği (%)	Özgünlük Specificity (%)	Hassasiyet Precision (%)	Hassasiyet Sensitivity (%)	F1 skoru (%)	Mean IoU (%)
VGG19	0.9863	0.3610	0.9956	0.9869	0.9858	0.9863	0.4266
RESNET	0.9884	0.4330	0.9961	0.9883	0.9883	0.9883	0.3971
U-NET	0.9801	0.3486	0.9945	0.9836	0.9794	0.9529	0.7927
INCEPTION V3	0.9904	0.5372	0.9968	0.9904	0.9903	0.9903	0.8826

Bu çalışmada kullandığımız algoritmalar ile ilgili yapılan literatür taramasında, değerlendirme için kullandığımız doğruluk, dice benzerliği, f1skor ve meaniou metriklerini bir arada kullanan bir çalışma örneğine rastlamadık. Fakat diğer yapılan çalışmalardan çıkan bazı metriklerin değerlerini Tablo 4'de görebilirsiniz.

Tablo 4

Karşılaştırmasını yaptığımız modellerin diğer çalışmalardaki metrik değerleridir.

Model	Veriseti	Doğruluk Accuracy (%)	Dice Benzerliği (%)	Hassasiyet Precision (%)	Hassasiyet Sensitivity (%)	F1 skoru (%)	Mean IoU (%)
VGG19 (Islam vd., 2023)	Brain Tumor MRI	0.96	-	1	0.88	0.94	-
VGG19 (Misu, 2023)	Kaggle(brain MRI)	0.9522	-	-	-	-	-

Tablo 4'ün devamı

ResNet50 (Onur, 2021)	253 MRI	0.8249	-	1	0.8294	0.9067	-
ResNet50 (Misu, 2023)	Kaggle(brain MRI)	0.9954	-	0.9954	0.9954	0.9954	
U-NET (Qian vd.,2022)	BraTs19	-	-	-	-	-	0.8996
U-NET (Zhou vd., 2020)	Brats2013,VisionGate, Data Science Bowl, MICCAI 2017,...	-	0.7573	-	-	0.244	0.9103
INCEPTIONV3 (Onur, 2021)	253 MRI	0.6813	-	0.4012	0.5048	0.4470	-
INCEPTIONV3 (Islam vd., 2023)	Brain Tumor MRI	0.98	-	1	0.92	0.96	-

BEŞİNCİ BÖLÜM
SONUÇ VE ÖNERİLER

Bu bölümde çalışmanın sonucunda elde edilen bulgular doğrultusunda varılan sonuçtan, bu çalışmanın limitleri ve sonraki çalışmalardan bahsedilmektedir. Ayrıca araştırma yansımaları kısmında bu çalışmanın kattığı değerden bahsedilmektedir.

5.1. Sonuç

Bu tezin amacı beyin tümörünün erken teşhisini kolaylaştırmak için medikal görüntüler üzerinde yaygın olarak kullanılan derin öğrenme mimarilerinden Vgg19, Resnet, Unet ve InceptionV3 mimarilerinin performansını karşılaştırmaktır. Bu karşılaştırma için bu mimarilerden modeller oluşturulmuş ve oluşan modeller BraTS 2020 veri seti ile eğitilmiş ve test edilmiştir. Kullanılan veri seti tüm algoritmalar için performans değerlendirmesi göz önünde bulunarak en doğru şekilde işlenmekte ve modelleri eğitmek için kullanılmaktadır. Veri setinin test için ayrılan kısmı ile test edilen eğitilmiş modellerin metrik sonuçları Tablo 3'te görülebilmektedir. Performans değerlendirmesi için doğruluk, dice benzerliği, f1 skor ve meanIoU, dikkate alınmış olan metriklerdir.

Tablo 3'te verilen değerlendirme sonuçları, Inception V3'ün en yüksek doğruluk oranına sahip olduğunu göstermektedir. Ancak, bazı durumlarda (örneğin, dengesiz sınıflandırma problemleri veya yanlış sınıflandırma maliyeti yüksek olan durumlar) doğruluk yeterli olmayabilir. Özellikle sınıf dengesizliği gibi durumlarda, doğruluk yanıltıcı olabilir ve modelin gerçek performansını tam olarak yansıtamaz. Sınıflar arasındaki dengesizlik, sadece doğruluk gibi basit metriklerle değerlendirilmeye bilmektedir. Bu nedenle, doğruluk metriği, diğer performans metrikleri ile birlikte değerlendirilmektedir.

Karşılaştırma yapılan algoritmaların dice benzerliğini karşılaştırdığımızda İnception V3 algoritmasının dice benzerliğinin daha yüksek olduğu Tablo 3'te görülmektedir. Ancak, tek başına bir metrik olarak kullanıldığında, bazı durumlarda yetersiz olabilir ve bu sebeple diğer performans metrikleri ile birlikte değerlendirilmelidir. Diğer metriklerle birlikte göz önünde bulundurulacağı için karşılaştırılma yapılan algoritmaların dice benzerliğini

incelediğimizde İnception V3 algoritmasının dice benzerliğinin daha yüksek olduğu görülmektedir. F1 skor metriği için Tablo 3'te verildiği gibi, değerlendirme sonucunda Inception V3'ün en yüksek f1 skor değerine sahip olduğu görülmektedir.

Özellikle çoklu sınıf segmentasyonu gibi alanlarda, her sınıf için ayrı IoU değerleri hesaplanır ve bu değerlerin ortalaması alınarak meanIoU hesaplanmaktadır. Modelin genel olarak tüm sınıflar için ne kadar iyi performans gösterdiğini ölçmek için kullanılmaktadır. Yüksek meanIoU değerleri, modelin nesnelerin sınırlarını daha iyi tanımladığı ve segmentasyonun daha doğru olduğu anlamına gelmektedir. Bu, doğrultuda Tablo 3 üzerinde elde edilen değerler incelendiğinde Inception V3 derin ağının meanIoU değerinin daha yüksek olduğu görülmektedir.

Tüm bu bahsettiğimiz metrikler göz önünde bulundurulduğunda yaptığımız çalışmada Inception V3 derin ağı mimarisinin diğer algoritmalara göre daha iyi performans göstermiştir. İnceptionV3'nün diğer algoritmalardan daha iyi performans gösterme sebepleri arasında şu noktalar öne çıkabilmektedir: mimari yapısı, transfer öğrenme, geniş görüntü temsili.

Mimari yapısı: Karmaşık ve derin mimari yapısıyla bilinmektedir. Bu yapısı, çoklu paralel katmanlar ve karmaşık modüller içermektedir. Bu karmaşıklık, daha derin ve karmaşık özelliklerin daha iyi temsil edilmesine olanak tanımaktadır.

Transfer Öğrenme: InceptionV3, transfer öğrenme konusunda güçlü bir yapıya sahiptir. Önceden eğitilmiş büyük bir veri seti olan ImageNet üzerinde eğitilmiş ve geniş bir veri tabanından gelen özellikleri öğrenmiştir. Bu özellikler, farklı görevler için daha iyi genelleme yapabilmesini sağlar.

Geniş Görüntü Temsili: InceptionV3, çok çeşitli ve derin özelliklerin temsil edilmesine odaklanmaktadır. Bu durum ise farklı nesneleri, özellikleri ve yapıları daha etkili bir şekilde tanımlamasına olanak tanımaktadır.

Bunlar göz önüne alındığında, InceptionV3'ün verilen metriklerde daha yüksek performans gösteren bir algoritma olması beklenmektedir. Yani bu çalışma sonucunda kullandığımız veri seti büyüklüğü ve kullanılan aktivasyon fonksiyonu da göz önüne alınarak Vgg19, ResNet, U-Net ve InceptionV3 modelleri arasında beyin tümörü tespiti açısından en başarılı model InceptionV3 olmuştur.

Ancak, yapılan literatür taramaları ve edinilen bilgiler göz önünde bulundurulduğunda metrik değerlerine dayanarak bir modelin kesin olarak daha performanslı olduğunu söylemek bazen yeterli olmayabilir. Herhangi bir modelin performansını etkileyebilecek bir dizi değişken olduğunu unutmamak önemlidir. Veri seti, eğitim süreci ve diğer hiper parametreler gibi faktörler de modelin performansını etkileyebilmektedir.

Bu bağlamda bazı kısıtlamalar mevcuttur. Bu kısıtlamalar ve sonraki çalışmalar 5.2'de anlatılmaktadır.

5.2. Limitler ve Sonraki Çalışmalar

Bu çalışmada, beyin tümör tespiti ve sınıflandırmasında derin öğrenme mimarileri kullanılarak bu mimarilerin performans değerlendirilmesi yapılmaktadır. Fakat yapılan literatür taramaları ve edinilen bilgiler doğrultusunda bir modelin performans değerlendirmesini yapabilmek için daha kapsamlı çalışmalar yapılması gerektiği ortaya konmaktadır. Bu nedenle sonraki çalışmalarımızda modelleri farklı veri setleri ile eğiterek performans değerlendirmesi üzerine çalışmalar gerçekleştirilecektir. Her geçen gün gelişen teknoloji göz önüne alınarak farklı mimarilerinde bu alanda değerlendirmeye alınıp alınmaması gerektiğine karar verilecektir. Fakat gerek bu çalışmalarda gerekse gelecekte yapılacak çalışmalarda bazı kısıtlamalar bulunmaktadır.

Herhangi bir makine öğreniminde doğru sonuçları tahmin etme doğruluğu her tahmin yapıldığında önemli ölçüde değişebilir ve eğitilmediği farklı bir türdeki MRI verildiğinde zaman zaman farklı sonuçlar verebilmektedir. Ayrıca bu mimariler veriye dayanan modeller oldukları için veri seti boyutu performans üzerinde önemli bir etkiye sahip olmaktadır. Veri

setinin büyümesi demek kaynak gerekliliği anlamına gelmektedir. Bu modeller yüksek performanslı bilgi işlem altyapısının yanı sıra yüksek GPU ihtiyacı olan mimarilerdir. Bu kısıtlama nedeniyle, modelin sınırlı kaynaklara sahip durumlarda kurulması zor olmakta, erişilebilirliği ve uygulanabilirliği de azalmaktadır.

Sonuç olarak, bir modelin performans incelmesi için büyük veri setleri ve buna bağlı olarak yüksek kaynaklara ihtiyaç vardır. Bu kısıtlarda göz önüne alınarak çıkardığımız sonuçların kısıtlı kaynak ve veri seti ile yapıldığı göz önüne alınarak değerlendirilmelidir.

5.3. Araştırma Yansımaları

Bu çalışmaya başlamadan önce yapay zeka, derin öğrenme ve medikal alanda kullanımları hakkında hiçbir bilgiye sahip değildim. Danışman hocamında yönlendirmeleriyle araştırma yaparken bir çok bilgi edindim. Veri seti hazırlanması, derin ağ mimarileri tanımlanması, eğitilmesi ve işlemesi gibi bir çok süreci detaylı olarak çalıştım. Bu çalışmamda tezimde yer vermediğim fakat gerek literatür taramaları olsun gerek model incelemeleri olsun, Deep belief, DenseNet, Vgg16, 3D CNN, CapsNet, Long Short Term Memory Network gibi bir çok farklı derin öğrenme mimarisi ile karşılaştım ve kullanım alanlarını inceleme şansı buldum. Bu çalışma bana yapay zekanın insan hayatında ne kadar büyük bir öneme sahip olabileceğini ve bu alanda öğrenme ve ilerleme isteği oluşturmuştur.

KAYNAKÇA

Afshar, P., Plataniotis, K. N., Mohammadi, A. (2018, Kasım 1). "Capsule Networks for Brain Tumor Classification based on MRI Images and Course Tumor Boundaries". *arXiv.org*. Erişim: 17 Aralık 2023, https://arxiv.org/abs/1811.00597

Alakwaa, W., Nassef, M., Badr, A. (2017). "Lung Cancer Detection and Classification with 3D Convolutional Neural Network (3D-CNN)". *International Journal of Advanced Computer Science and Applications, 8*(8). doi:10.14569/IJACSA.2017.080853

Ali, L., Alnajjar, F., Jassmi, H., Gochoo, M., Khan, W., Serhani, M. (2021). "Performance Evaluation of Deep CNN-Based Crack Detection and Localization Techniques for Concrete Structures". *Sensors*, *21*, 1688. doi:10.3390/s21051688

Badža, M. M., Barjaktarović, M. Č. (2020). "Classification of Brain Tumors from MRI Images Using a Convolutional Neural Network". *Applied Sciences*, *10*(6), Article 6. doi:10.3390/app10061999

Bramer, M. (2007). *Principles of Data Mining*. doi:10.1007/978-1-84628-766-4

Cardinal, L. J. (2016). "Diagnostic testing: A key component of high-value care." *Journal of Community Hospital Internal Medicine Perspectives*, *6*(3). doi:10.3402/jchimp.v6.31664

Casamitjana, A., Puch, S., Aduriz, A., Sayrol, E., Vilaplana, V. (t.y.). "3D Convolutional Networks for Brain Tumor Segmentation". *arXiv.org*. Erişim: 17 Aralık 2023, https://arxiv.org/abs/1705.08236v1

Chen, L.-C., Papandreou, G., Kokkinos, I., Murphy, K., Yuille, A. L. (2014, Aralık 22). "Semantic Image Segmentation with Deep Convolutional Nets and Fully Connected CRFs". *arXiv.org*. Erişim: 29 Aralık 2023, https://arxiv.org/abs/1412.7062v4

Çetiner, H., Metlek, S. (2023). "DenseUNet+: A novel hybrid segmentation approach based on multi-modality images for brain tumor segmentation". *Journal of King Saud University - Computer and Information Sciences*, *35*(8), 101663. doi:10.1016/j.jksuci.2023.101663

Daimary, D., Bora, M., Amitab, kh, Kandar, D. (2020). "Brain Tumor Segmentation from MRI Images using Hybrid Convolutional Neural Networks". *Procedia Computer Science*, *167*, 2419-2428. doi:10.1016/j.procs.2020.03.295

Dehkordi, A. A., Hashemi, M., Neshat, M., Mirjalili, S., Sadiq, A. S. (2022, Nisan 26). "Brain Tumor Detection and Classification Using a New Evolutionary Convolutional

Neural Network". *arXiv.org*. Erişim adresi: https://arxiv.org/abs/2204.12297v1

Díaz-Pernas, F. J., Martínez-Zarzuela, M., Antón-Rodríguez, M., González-Ortega, D. (2021). "A Deep Learning Approach for Brain Tumor Classification and Segmentation Using a Multiscale Convolutional Neural Network". *Healthcare*, *9*(2), Article 2. doi:10.3390/healthcare9020153

El Naqa, I., Murphy, M. (2015). *What Is Machine Learning?* (ss. 3-11). doi:10.1007/978-3-319-18305-3_1

Filatov, D., Yar, G. N. A. H. (2022, Temmuz 27). "Brain Tumor Diagnosis and Classification via Pre-Trained Convolutional Neural Networks". *arXiv.org*. Erişim: 26 Eylül 2023, https://arxiv.org/abs/2208.00768v1

Google (n.d). Machine Learning, Crash Course, *Classification: True vs. False and Positive vs. Negative.* Erişim adresi: https://developers.google.com/machine-learning/crash-course/classification/true-false-positive-negative?hl=en

Goutte, C., Gaussier, E. (2005). A Probabilistic Interpretation of Precision, Recall and F-Score, with Implication for Evaluation. Içinde D. E. Losada J. M. Fernández-Luna (Ed.), *Advances in Information Retrieval* (ss. 345-359). (Springer) doi:10.1007/978-3-540-31865-1_25

Gupta, A., Dixit, M., Mishra, V. K., Singh, A., Dayal, A. (2023, Nisan 29). "Brain Tumor Segmentation from MRI Images using Deep Learning Techniques". *arXiv.org.* Erişim: 16 Aralık 2023, https://arxiv.org/abs/2305.00257v1

He, K., Zhang, X., Ren, S., Sun, J. (2015, Aralık 10). "Deep Residual Learning for Image Recognition". *arXiv.org*. Erişim: 25 Kasım 2023, https://arxiv.org/abs/1512.03385v1

Huang, G., Liu, Z., van der Maaten, L., Weinberger, K. Q. (2016, Ağustos 25). "Densely Connected Convolutional Networks". *arXiv.org*. Erişim: 16 Aralık 2023, https://arxiv.org/abs/1608.06993v5

Imam, R., Alam, M. T. (2023). "Optimizing Brain Tumor Classification: A Comprehensive Study on Transfer Learning and Imbalance Handling in Deep Learning Models". *arXiv.org*. Erişim: 5 Aralık 2023, https://arxiv.org/abs/2308.06821

Islam, M. M., Brau, P., Rahman, M., Ahammed, T., Akter, L., Uddin, J. (2023). "Transfer Learning Architectures with fine-tunning for brain tumor classification using magnetic resonance imaging". doi:10.1016/j.health.2023100270

Kaggle (2020). *Brain Tumor Segmentation (BraTS2020*). Erişim adresi:

https://www.kaggle.com/datasets/awsaf49/brats2020-training-data/data

Kanade, P. B., Gumaste, P. P. (2015). "Brain Tumor Detection Using MRI Images". *IJIREEICE*, 146-150. doi:10.17148/IJIREEICE.2015.3231

Kayalibay, B., Jensen, G., van der Smagt, P. (2017). "CNN-based Segmentation of Medical Imaging Data". *arXiv.org*. Erişim: 9 Aralık 2023, https://arxiv.org/abs/1701.03056

Kokkalla, S., Kakarla, J., Venkateswarlu, I. B., Singh, M. (2021). "Three-class brain tumor classification using deep dense inception residual network." *Soft Computing*, *25*(13), 8721-8729. doi:10.1007/s00500-021-05748-8.

Krizhevsky, A., Sutskever, I., Hinton, G. E. (2012). "ImageNet Classification with Deep Convolutional Neural Networks". *Advances in Neural Information Processing Systems*, *25*. Erişim: 5 Ekim 2023,

https://proceedings.neurips.cc/paper_files/paper/2012/hash/c399862d3b9d6b76c8436e924a68c45b-Abstract.html

Lather, M., Singh, P. (2020). "Investigating Brain Tumor Segmentation and Detection Techniques". *Procedia Computer Science*, *167*, 121-130. doi:10.1016/j.procs.2020.03.189.

Lecun, Y., Bottou, L., Bengio, Y., Haffner, P. (1998). "Gradient-Based Learning Applied to Document Recognition". *Proceedings of the IEEE*, *86*, 2278-2324. doi:10.1109/5.726791

Maurya, U., Kalyan, A. K., Bohidar, S., Sivakumar, D. S. (2023, Nisan 18). "Detection and Classification of Glioblastoma Brain Tumor". *arXiv.org*. Erişim: 19 Aralık 2023, https://arxiv.org/abs/2304.09133

Misu, R. S. (2023). "Brain Tumor Detection Using Deep Learning Approaches". *arXiv.org*. Erişim: 21 Aralık 2023, https://arxiv.org/abs/2309.12193

Monaghan, T. F., Rahman, S. N., Agudelo, C. W., Wein, A. J., Lazar, J. M., Everaert, K., Dmochowski, R. R. (2021). "Foundational Statistical Principles in Medical Research: Sensitivity, Specificity, Positive Predictive Value, and Negative Predictive Value". *Medicina*, *57*(5), 503. doi:10.3390/medicina57050503

National Brain Tumor Society (2023). *Brain Tumor Facts.* Erişim: 10 Ekim 2023, https://braintumor.org/brain-tumors/about-brain-tumors/brain-tumor-facts/

Nasim, M. A. A., Munem, A. A., Islam, M., Palash, M. A. H., Haque, M. M. A., Shah, F. M. (2023, Ekim 24). "Brain Tumor Segmentation using Enhanced U-Net Model with Empirical Analysis". *arXiv.org*. Erişim adresi: https://arxiv.org/abs/2210.13336v2

Pereira, S., Pinto, A., Alves, V., Silva, C. (2016). "Brain Tumor Segmentation Using Convolutional Neural Networks in MRI Images". *IEEE Transactions on Medical Imaging*, *35*, 1-1. doi:10.1109/TMI.2016.2538465

Qian, L., Zhou, X., Li, Y., Hu, Z. (2022, Mayıs 24). "UNet#: A UNet-like Redesigning Skip Connections for Medical Image Segmentation". *arXiv.org*. Erişim: 2 Aralık 2023, https://arxiv.org/abs/2205.11759

Onur, S. (2021). "Performance Comparison of Different Pre-Trained Deep Learning Models in Classifying Brain MRI Images". doi:10.26650/acin.880918

Raja, N. S. M., Fernandes, S. L., Dey, N., Satapathy, S. C., Rajinikanth, V. (2018). "Contrast enhanced medical MRI evaluation using Tsallis entropy and region growing segmentation". *Journal of Ambient Intelligence and Humanized Computing*. doi:10.1007/s12652-018-0854-8

Ronneberger, O., Fischer, P., Brox, T. (2015, Mayıs 18). "U-Net: Convolutional Networks for Biomedical Image Segmentation". *arXiv.org.* Erişim: 9 Aralık 2022, https://arxiv.org/abs/1505.04597

Sabour, S., Frosst, N., Hinton, G. E. (2017, Ekim 26). "Dynamic Routing Between Capsules". *arXiv.org*. Erişim: 31 Aralık 2023, https://arxiv.org/abs/1710.09829v2

Santos, D. (2023). "Brain Tumor Detection using the VGG-16 Model: A Deep Learning Approach". doi:10.20944/preprints202308.0983.v1

Simonyan, K., Zisserman, A. (2015, Eylül 4). "Very Deep Convolutional Networks for Large-Scale Image Recognition". *arXiv.org*. Erişim: 31 Kasım 2023, https://arxiv.org/abs/1409.1556

Soumik, Mohd. F. I., Hossain, D. M. A. (2020, Temmuz 6). "Brain Tumor Classification With Inception Network Based Deep Learning Model Using Transfer Learning". doi:10.1109/TENSYMP50017.2020.9230618

Speet, S. (2020). "Brain Tumor Segmentation Using a Capsule Network". *Data Science*.

Srivastava, T. (2019, Ağustos 6). 12 Important Model Evaluation Metrics for Machine Learning Everyone Should Know (Updated 2023). *Analytics Vidhya*. Erişim adresi: https://www.analyticsvidhya.com/blog/2019/08/11-important-model-evaluation-error-metrics/

Szegedy, C., Liu, W., Jia, Y., Sermanet, P., Reed, S., Anguelov, D., Erhan, D., Vanhoucke, V., Rabinovich, A. (2014, Eylül 17). "Going Deeper with Convolutions". *arXiv.org.* Erişim adresi: https://arxiv.org/abs/1409.4842v1

Szegedy, C., Vanhoucke, V., Ioffe, S., Shlens, J., Wojna, Z. (2015, Aralık 2). "Rethinking the Inception Architecture for Computer Vision". *arXiv.org*. Erişim: 22 Kasım 2023, https://arxiv.org/abs/1512.00567

Tamilarasi, R., Gopinathan, S. (2021, Temmuz 1). "Inception Architecture for Brain Image Classification". *Journal of Physics: Conference Series*, *1964*(7), 072022. doi:10.1088/1742-6596/1964/7/072022

Tang, W., Sun, J., Wang, S., Zhang, Y. (2023, Kasım 15). "Review of AlexNet for Medical Image Classification". *arXiv.org*. Erişim: 1 Aralık 2023, https://arxiv.org/abs/2311.08655v1.

Tukey, J. W. (1962). The Future of Data Analysis. *The Annals of Mathematical Statistics*, *33*(1), 1-67.

Voulodimos, A., Protopapadakis, E., Katsamenis, I., Doulamis, A., Doulamis, N. (2021). "A few-shot U-Net deep learning model for COVID-19 infected area segmentation in CT images". *Sensors*, *21*, 2215. doi:10.3390/s21062215

Walsh, J., Othmani, A., Jain, M., Dev, S. (2022). "Using U-Net network for efficient brain tumor segmentation in MRI images". *Healthcare Analytics*, *2*, 100098. doi:10.1016/j.health.2022.100098

Zheng, P., Zhu, X., Guo, W. (2022). "Brain tumour segmentation based on an improved U-Net". *BMC Medical Imaging*, *22*(1), 199. doi:10.1186/s12880-022-00931-1

Zhou, Z., Siddiquee, M., M., R., Tajbakhsh, N., Liang, J. (2020, Aralık 11). "Unet++: Redesigning Skip Connections to Exploit Multiscale Features in Image Segmentation". *arXiv.org*. Erişim: 23 Kasım 2023, https://arxiv.org/abs/1912.05074v2