



T.C. İstanbul Medeniyet Üniversitesi
Mühendislik ve Doğa Bilimleri Fakültesi
Bilgisayar Mühendisliği Bölümü

Derin Öğrenme Tabanlı Beyin Tümörü Sınıflandırması

BİL485 – Derin Öğrenme
Dr. Öğr. Üyesi İshak DÖLEK

21120205306 Beyza Nur Karahasanoğlu
Aralık 2025

İÇİNDEKİLER

1. Giriş.....	1
2. Veri Seti.....	2
3. Yöntem.....	4
3.1. Problem Tanımı.....	4
3.2. Yöntem Seçim Gerekçesi.....	4
3.3. Literatürde Kullanılan Yöntemlerin Karşılaştırmalı İncelenmesi.....	4
3.4. Bu Çalışmada Kullanılan Model Mimarileri.....	4
4. Model Eğitimi ve Model Değerlendirme.....	6
4.1. Veri Hazırlama ve Ön İşleme.....	7
4.2. Temel CNN Modelinin Eğitimi.....	7
4.3. Transfer Learning ile ResNet50 Modelinin Eğitimi ve Değerlendirilmesi.....	9
4.4. Transfer Learning ile EfficientNet-B0 Modelinin Eğitimi ve Değerlendirilmesi.....	10
4.5. Transfer Learning ile MobileNetV2 Modelinin Eğitimi ve Değerlendirilmesi.....	12
4.6. Model Karşılaştırma ve Nihai Model Seçimi.....	14
5. Sonuç.....	15
6. Kaynakça.....	17

1. Giriş

Beyin tümörleri, merkezi sinir sisteminde ortaya çıkan ve yaşam kalitesini ciddi biçimde etkileyen önemli sağlık problemleri arasında yer almaktadır. Erken ve doğru teşhis, hastalığın прогнозunun belirlenmesi ve uygun tedavi stratejilerinin uygulanması açısından kritik öneme sahiptir. Bu çalışmada ele alınan beyin tümörü türleri, klinik literatürde yaygın olarak tanımlanan ve manyetik rezonans görüntülerinde ayırt edilebilir morfolojik özellikler gösteren tümör gruplarından oluşmaktadır.

Glioma, beynin destek hücreleri olan glial hücrelerden köken alan ve genellikle infiltratif yapıda büyümeye götüren tümörleri ifade etmektedir. Gliomalar, heterojen doku yapıları ve düzensiz sınırları nedeniyle MR görüntülerinde belirgin yoğunluk ve kontrast farklılıklarını sergileyebilmektedir.

Meningioma, beyn zarlarından (meninksler) köken alan ve çoğulukla iyi huylu karaktere sahip olan tümörlerdir. MR görüntülerinde genellikle daha düzgün sınırlara sahip, kompakt ve homojen yapılar olarak gözlemlenmektedir. Bu özellikleri, meningiomaların diğer tümör türlerinden görsel olarak ayırtılabilirliğini mümkün kılmaktadır.

Pituitary tümörleri, hipofiz bezinde gelişen ve hormonal dengesizliklere yol açabilen tümörlerdir. Bu tümörler, anatominik konumları nedeniyle MR görüntülerinde belirli bir bölgeye lokalize şekilde görülmekte ve çevre dokulara göre farklı yoğunluk dağılımları göstermektedir.

No tumor sınıfı ise, herhangi bir tümöral oluşum içermeyen sağlıklı beyn MR görüntülerini temsil etmektedir. Bu sınıf, modelin patolojik ve normal görüntüler arasındaki ayrimı öğrenebilmesi açısından kritik bir referans kategorisi olarak veri setine dahil edilmiştir.

Bu tümör türlerinin farklı morfolojik ve yoğunluk özellikleri, derin öğrenme tabanlı modellerin ayırt edici uzamsal örüntüleri öğrenebilmesi için uygun bir problem alanı sunmaktadır.

Etiketleme işlemi, görüntülerin sınıflara göre ayrılmış klasörlerde tutulması esasına dayanmaktadır. Bu yaklaşım, görüntü sınıflandırma çalışmalarında yaygın kullanılan bir etiketleme biçimi olup, veri okuma sürecinin standardize edilmesine ve etiketleme hatalarının azaltılmasına katkı sağlamaktadır.

Günümüzde beyn tümörlerinin teşhisinde en yaygın kullanılan görüntüleme yöntemi manyetik rezonans görüntüleme (MRG) olup, yüksek çözünürlüklü ve detaylı anatominik bilgi sunması nedeniyle klinik uygulamalarda tercih edilmektedir.

Ancak MRG görüntülerinin manuel olarak değerlendirilmesi, radyolog deneyimine bağlı olması, zaman alıcı bir süreç içermesi ve subjektif farklılıklara açık olması gibi çeşitli sınırlamalara sahiptir. Bu durum, otomatik ve bilgisayar destekli tanı sistemlerine olan ihtiyacı artırılmıştır. Son yıllarda makine öğrenmesi ve özellikle derin öğrenme yöntemleri, tıbbi görüntü analizi alanında başarılı sonuçlar ortaya koymuş ve beyn tümörü sınıflandırması gibi problemlerde etkin bir şekilde kullanılmaya başlanmıştır.

Evrişimsel Sinir Ağları (Convolutional Neural Networks – CNN), görüntülerden otomatik özellik çıkarımı yapabilme yetenekleri sayesinde klasik makine öğrenmesi yöntemlerine kıyasla önemli avantajlar sunmaktadır. Literatürde, CNN tabanlı modellerin beyn MR görüntülerinden tümör tespiti ve sınıflandırılmasında yüksek doğruluk oranlarına ulaştığı rapor edilmiştir. Bununla birlikte, büyük veri gereksinimi ve eğitim süresinin uzunluğu gibi zorluklar, transfer learning yaklaşımının önemini artırmıştır.

Transfer learning, önceden büyük veri setleri üzerinde eğitilmiş derin öğrenme modellerinin, hedef probleme uyarlanarak kullanılması esasına dayanmaktadır. Bu yaklaşım, sınırlı veri setlerine sahip tıbbi görüntüleme problemlerinde hem eğitim süresini azaltmakta hem de model performansını artırabilmektedir. Literatürde ResNet, VGG, EfficientNet gibi mimarilerin beyin tümörü sınıflandırma problemlerinde başarıyla uygulandığı görülmektedir.

Bu proje kapsamında, beyin MR görüntülerinden farklı tümör türlerinin otomatik olarak sınıflandırılması amacıyla derin öğrenme tabanlı bir yaklaşım geliştirilmesi hedeflenmektedir. Çalışmada, transfer learning temelli farklı CNN mimarileri karşılaştırmalı olarak ele alınacak, en iyi performansı gösteren model belirlenerek literatürde raporlanan sonuçlarla karşılaştırılacaktır. Böylece, tıbbi görüntü analizi alanında derin öğrenme yöntemlerinin etkinliği objektif ölçütlerle değerlendirilmiş olacaktır.

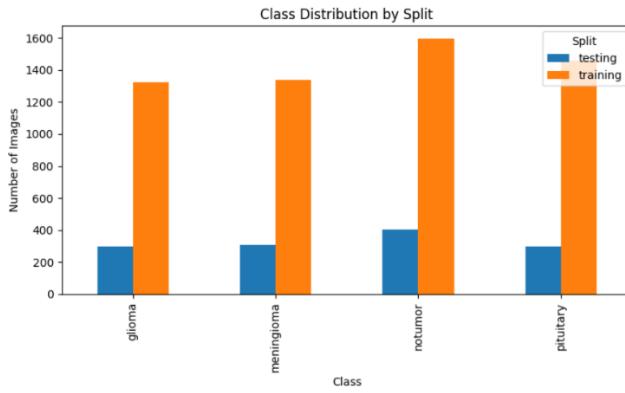
2. Veri Seti

Bu çalışmada, manyetik rezonans (MR) görüntülerinden beyin tümörü sınıflandırması amacıyla hazırlanmış, klasör tabanlı etiketleme yapısına sahip bir görüntü veri seti kullanılmıştır. Bu çalışmada kullanılan veri seti, Kaggle platformu üzerinden açık erişimle paylaşılan *Brain Tumor MRI Dataset* olup, dört sınıfından oluşan çok sınıflı bir görüntü sınıflandırma problemi olacak şekilde düzenlenmiştir: glioma, meningioma, pituitary ve no tumor. Etiketleme işlemi, görüntülerin sınıflara göre ayrılmış klasörlerde tutulması esasına dayanmaktadır. Bu yaklaşım, görüntü sınıflandırma çalışmalarında yaygın kullanılan bir etiketleme biçimi olup, veri okuma sürecinin standardize edilmesine ve etiketleme hatalarının azaltılmasına katkı sağlamaktadır.

Veri seti dizin yapısı iki ana alt küme şeklinde organize edilmiştir: training ve testing. Her bir alt küme altında dört sınıfa karşılık gelen klasörler bulunmaktadır: dataset/training/{glioma, meningioma, notumor, pituitary} ve dataset/testing/{glioma, meningioma, notumor, pituitary}. Bu yapı, modelin eğitimi sırasında kullanılan veriler ile performans değerlendirmesi için ayrılan verilerin birbirinden ayrılmasını sağlayarak, elde edilen sonuçların daha güvenilir biçimde yorumlanması olanak tanımaktadır.

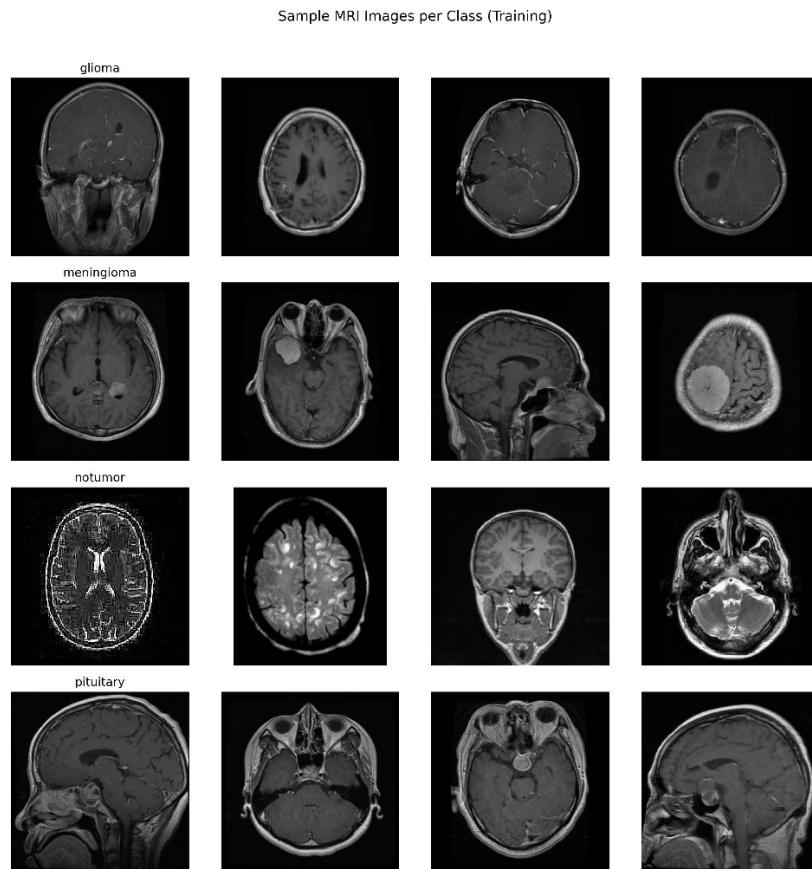
Class	Train	Test	Total
Glioma	1321	300	1621
Meningioma	1339	306	1645
Notumor	1595	405	2000
Pituitary	1457	300	1757
Total	5712	1311	7023

Tablo 2.1 Sınıf Bazında Görüntü Sayıları.



Şekil 2.1. Sınıf bazlı toplam görüntü sayılarının görsel temsili.

Veri setindeki görüntülerin sınıf bazlı dağılımı Tablo 2.1'de sunulmaktadır. Toplam örnek sayısının sınıflar arasında eşit olup olmadığı, veri setindeki olası sınıf dengesizliğinin tespiti açısından önemlidir. Sınıf bazlı toplam görüntü sayılarının görsel temsili ise Şekil 2.1'de gösterilmektedir. Bu dağılım, model eğitimi ve değerlendirme aşamalarında kullanılacak stratejilerin (ör. sınıf ağırlıkları, veri artırma) belirlenmesi bakımından raporda dikkate alınmıştır.



Şekil 2.2. Her bir sınıfa ait örnek MR görüntüleri.

Her bir sınıfa ait örnek MR görüntüleri Şekil 2.2'de sunulmuştur. Bu örnekler, sınıflar arasındaki görsel örüntü farklılıklarını ve veri setinin içерdiği örnek çeşitliliğini nitel olarak göstermeyi amaçlamaktadır. Görüntülerin çözünürlükleri ve yoğunluk dağılımları gibi temel özellikler, bir sonraki bölümde gerçekleştirilen veri ön işleme adımlarının gerekçelendirilmesinde kullanılacaktır.

3. Yöntem

Bu çalışmada, beyin MR görüntülerinden tümör türlerinin otomatik olarak sınıflandırılması amacıyla derin öğrenme tabanlı bir yaklaşım benimsenmiştir. Problem, her bir MR görüntüsünün dört sınıfından birine (glioma, meningioma, pituitary, no tumor) atanmasını hedefleyen çok sınıflı bir görüntü sınıflandırma problemi olarak ele alınmıştır. Girdi verisi olarak iki boyutlu MR görüntüleri kullanılmış, çıktı olarak ise görüntünün ait olduğu tümör sınıfı üretilmiştir.

3.1 Problem Tanımı

Ele alınan problem, etiketli MR görüntülerinden öğrenilen ayırt edici görsel örüntüler aracılığıyla, daha önce görülmemiş görüntülerin doğru sınıfa atanmasını amaçlamaktadır. Sınıflar arasında gözlemlenen morfolojik farklılıklar; tümörün konumu, sınır yapısı, doku yoğunluğu ve heterojenlik gibi görsel özellikler üzerinden ortaya çıkmaktadır. Bu özelliklerin manuel olarak tanımlanması zor ve zaman alıcı olduğundan, otomatik özellik çıkarımı yapabilen derin öğrenme modelleri tercih edilmiştir.

3.2 Yöntem Seçim Gerekçesi

Görüntü verilerinin analizinde evrişimsel sinir ağları (Convolutional Neural Networks – CNN) yaygın olarak kullanılan ve başarılı sonuçlar üreten yöntemlerdir. CNN mimarileri, görüntülerden düşük seviyeli (kenar, köşe) ve yüksek seviyeli (şekil, doku) özellikleri hiyerarşik olarak öğrenebilme yeteneğine sahiptir. Bu özellik, beyin MR görüntülerinde tümör ve normal doku arasındaki karmaşık yapısal farklılıkların modellenmesi açısından önemlidir.

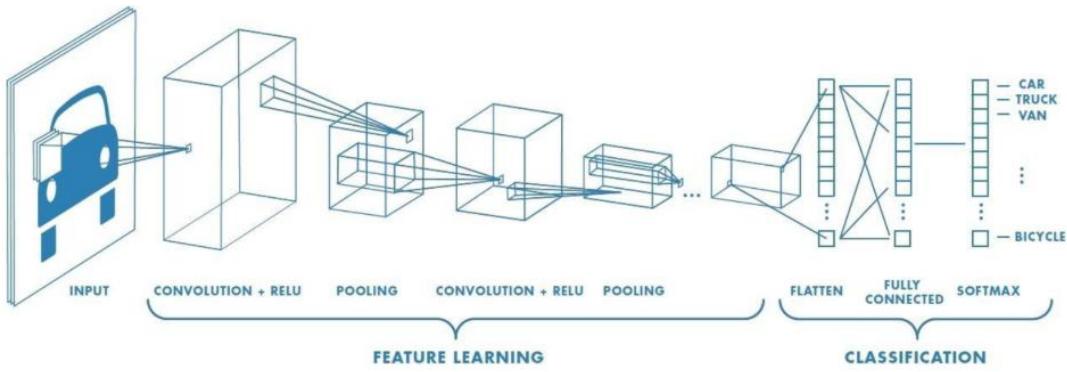
Tıbbi görüntüleme alanında veri setlerinin genellikle sınırlı sayıda örnek içermesi, derin öğrenme modellerinin sıfırdan eğitilmesini zorlaştırmaktadır. Bu nedenle bu çalışmada, büyük ölçekli görüntü veri setleri üzerinde önceden eğitilmiş CNN mimarilerinin kullanıldığı transfer learning yaklaşımı benimsenmiştir. Bu yaklaşım sayesinde, genel görsel özellikleri öğrenmiş modellerin, beyin MR görüntülerine özgü problemlere daha etkin biçimde uyarlanması hedeflenmiştir.

3.3 Literatürde Kullanılan Yöntemlerin Karşılaştırılmalı İncelenmesi

Literatürde beyin tümörü sınıflandırması problemlerinde farklı derin öğrenme mimarilerinin kullanıldığı görülmektedir. Özellikle VGG, ResNet, DenseNet ve EfficientNet gibi mimariler, MR görüntülerinden otomatik özellik çıkarımı ve sınıflandırma görevlerinde yaygın olarak tercih edilmektedir. Bu çalışmalarda genellikle transfer learning ve ince ayar (fine-tuning) stratejilerinin, model performansını artırdığı rapor edilmektedir.

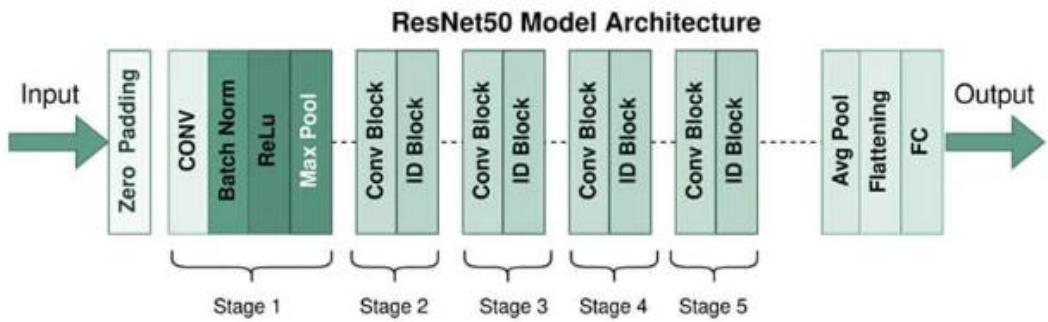
3.4 Bu Çalışmada Kullanılan Model Mimarileri

Bu çalışmada, beyin MR görüntülerinden tümör sınıflandırma performansını değerlendirmek amacıyla farklı derin öğrenme mimarileri seçilmiştir. Model seçimi yapılırken, literatürde tıbbi görüntü sınıflandırma problemlerinde yaygın olarak kullanılan, yüksek temsil gücüne sahip ve transfer learning yaklaşımına uygun mimariler tercih edilmiştir. Modeller, hem mimari karmaşıklık hem de hesaplama maliyeti açısından farklı özellikler sunacak şekilde belirlenmiştir.



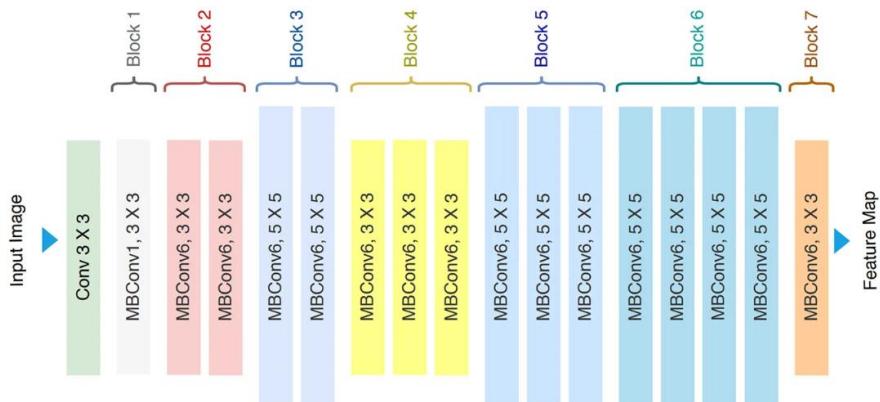
Şekil 3.4.1. CNN mimarisinin genel yapısı.

İlk olarak, karşılaştırma için bir referans noktası oluşturmak amacıyla sıfırdan eğitilen temel bir CNN modeli kullanılmıştır. Bu model, sınırlı sayıda evrişimsel ve havuzlama katmanından oluşmakta olup, görüntülerden temel uzamsal özelliklerin öğrenilmesini hedeflemektedir. Temel CNN mimarisinin genel yapısı Şekil 3.4.1’te sunulmuştur. Bu model, transfer learning tabanlı yaklaşımların sağladığı performans artışının daha net biçimde gözlemlenebilmesi açısından önem taşımaktadır.



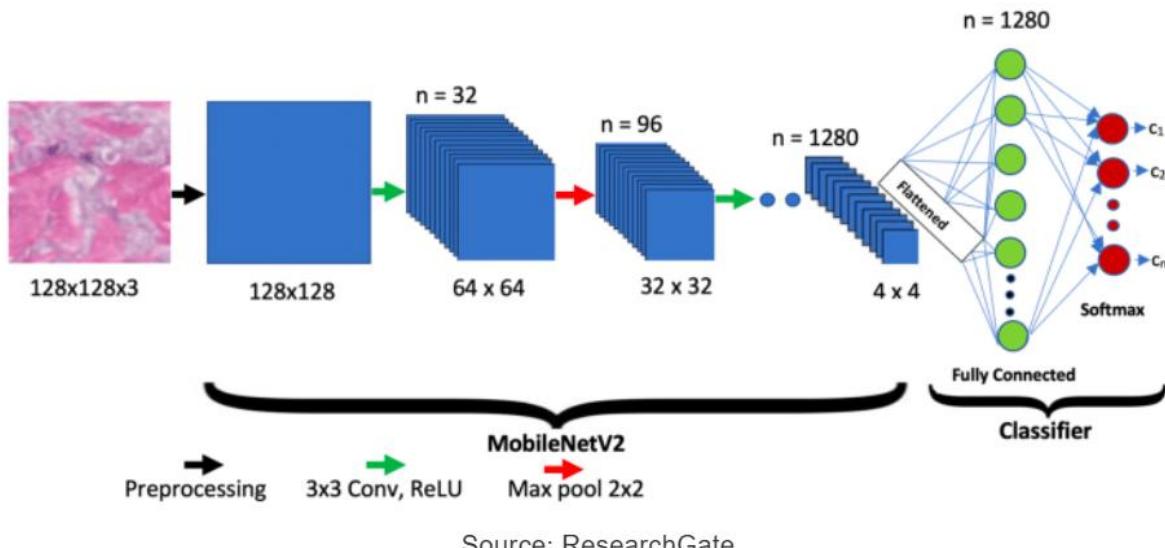
Şekil 3.4.2. ResNet50 mimarisinin genel yapısı.

Transfer learning kapsamında kullanılan ilk derin model olarak ResNet50 mimarisini tercih edilmiştir. Artık bağlantılar (residual connections) sayesinde derin ağlarda görülen gradyan kaybolması problemini azaltan ResNet mimarisini, derin yapısına rağmen etkili biçimde eğitilebilmektedir. ResNet50’nin evrişimsel blok yapısı ve artık bağlantı mekanizması, derin özellik temsillerinin öğrenilmesini mümkün kılmaktadır. Bu mimarinin genel yapısı Şekil 3.4.2’te gösterilmiştir.



Şekil 3.4.3. EfficientNet-B0 mimarisinin genel yapısı.

İkinci transfer learning modeli olarak EfficientNet-B0 mimarisi kullanılmıştır. EfficientNet ailesi, ağı derinliği, genişliği ve giriş çözünürlüğünü dengeli biçimde ölçeklendiren yapısıyla dikkat çekmektedir. EfficientNet-B0, daha az parametre ile yüksek doğruluk oranları elde edebilmesi nedeniyle özellikle sınırlı veri setlerine sahip tıbbi görüntüleme problemleri için uygun bir alternatif sunmaktadır. EfficientNet-B0 mimarisinin katman organizasyonu ve ölçeklendirme yapısı Şekil 3.4.3.'te sunulmuştur.



Şekil 3.4.4. MobileNetV2 mimarisinin genel yapısı.

Ek olarak, hesaplama maliyeti ve model boyutunun daha düşük olduğu bir yaklaşımı değerlendirmek amacıyla MobileNetV2 mimarisi de deneysel çalışmalarla dahil edilmiştir. Derinlik ayrılabılır evrişim (depthwise separable convolution) yapısı sayesinde MobileNetV2, daha hafif bir mimari sunmakta olup, performans–verimlilik dengesinin incelenmesine olanak tanımaktadır. MobileNetV2 mimarisinin genel yapısı Şekil 3.4.4'te gösterilmiştir.

Bu çalışmada, literatürde yaygın olarak kullanılan tüm mimarileri kapsamlı biçimde denemek yerine, problem bağlamına uygun, temsil gücü yüksek ve hesaplama maliyeti açısından dengeli mimariler seçilmiştir. VGG ve DenseNet gibi mimariler yüksek parametre sayısı ve bellek gereksinimleri nedeniyle tercih edilmezken; ResNet ve EfficientNet mimarileri, derinlik, verimlilik ve genelleme başarısı açısından uygun alternatifler sunduğu için deneysel çalışmalarla dahil edilmiştir.

4. Model Eğitimi ve Model Değerlendirme

Bu bölümde, bir önceki bölümlerde tanımlanan veri seti ve seçilen model mimarileri doğrultusunda gerçekleştirilen model eğitimi ve değerlendirme süreci ayrıntılı olarak ele alınmaktadır. Çalışma kapsamında, derin öğrenme tabanlı modellerin eğitimi için izlenen veri hazırlama adımları, kullanılan eğitim stratejileri ve performans değerlendirme yöntemleri sistematik bir biçimde sunulmaktadır. Ayrıca, farklı model mimarilerinin karşılaştırılmalı performans analizi yapılarak, en iyi sonuçları veren model belirlenmektedir.

Bu kapsamda, öncelikle veri seti üzerinde gerçekleştirilen keşifsel veri analizi ve ön işleme adımları açıklanmaktadır; ardından model eğitimi, transfer learning ve ince ayar (fine-tuning)

süreçleri detaylandırılmaktadır. Son olarak, eğitilen modellerin performansları çeşitli değerlendirme metrikleri üzerinden karşılaştırılarak elde edilen sonuçlar yorumlanmaktadır.

4.1 Veri Hazırlama ve Ön İşleme

Bu bölümde, model eğitim sürecinden önce veri seti üzerinde temel keşifsel analiz (EDA) gerçekleştirilmiş ve verinin derin öğrenme modellerine uygun hale getirilmesi için ön işleme adımları uygulanmıştır. Görüntülerin biçimsel özelliklerini incelemek amacıyla rastgele örnekler üzerinden görüntü boyutları ve kanal yapıları kontrol edilmiştir. Sınıflardan seçilen örnek MR görüntülerini, veri setindeki görsel çeşitliliğin nitel olarak gözlemlenebilmek için bakınız (Şekil 2.2).

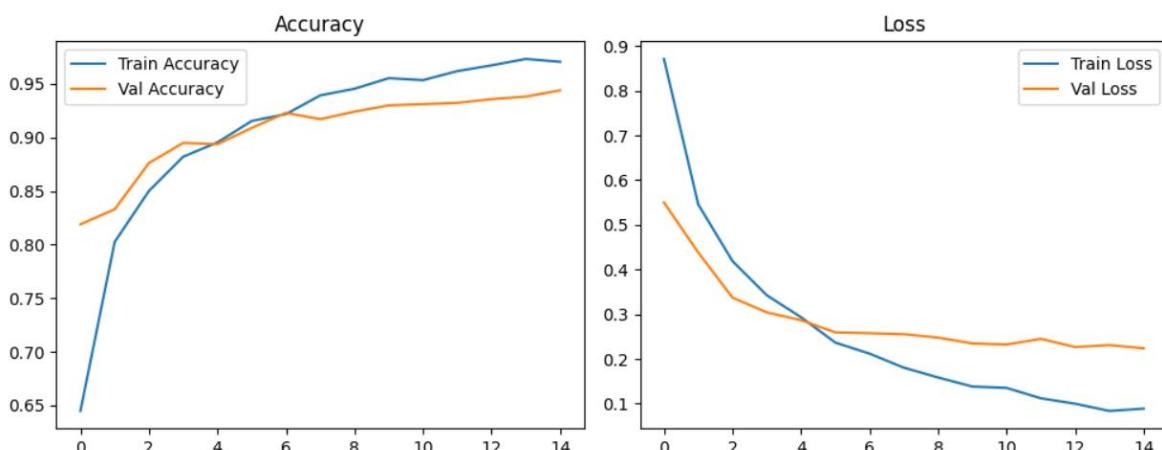
Görüntüler, eğitim sürecinde ortak bir giriş boyutuna ihtiyaç duyulması nedeniyle 224×224 çözünürlüğe yeniden boyutlandırılmıştır. Piksel değerleri, modelin daha kararlı öğrenebilmesi amacıyla $[0, 1]$ aralığına normalize edilmiştir. Ayrıca eğitim verisi üzerinde genelleme başarısını artırmak ve aşırı öğrenmeyi azaltmak amacıyla veri artırma (data augmentation) uygulanmıştır. Bu kapsamında yatay çevirme, küçük açısal döndürme ve yakınlaştırma işlemleri eğitim sırasında rastgele şekilde gerçekleştirilmiştir.

Eğitim verisi, eğitim ve doğrulama alt kümelerine ayrılmış; doğrulama kümesi, hiperparametre seçimi ve aşırı öğrenmenin izlenmesi amacıyla kullanılmıştır. Test alt kümesi ise model performansının nihai değerlendirilmesi için ayrı tutulmuştur. Sınıflar arası örnek sayılarının tam dengeli olmaması nedeniyle, eğitim sürecinde sınıf dengesizliğinin etkisini azaltmak amacıyla sınıf ağırlıkları hesaplanarak eğitim aşamasında kullanılmıştır.

4.2 Temel CNN Modelinin Eğitimi

Transfer learning tabanlı modellerin performansını daha sağlıklı biçimde değerlendirebilmek amacıyla, öncelikle sıfırdan eğitilen temel bir evrişimsel sinir ağı (CNN) modeli kullanılmıştır. Bu model, sınırlı sayıda evrişimsel ve havuzlama katmanından oluşmakta olup, görüntülerden temel uzamsal özelliklerin öğrenilmesini hedeflemektedir.

Temel CNN modeli, artan filtre sayısına sahip üç evrişimsel bloktan oluşmaktadır. Evrişimsel katmanları takiben maksimum havuzlama katmanları kullanılarak uzamsal boyutlar kademeli olarak azaltılmıştır. Elde edilen özellikler, tam bağlantılı katmanlar aracılığıyla sınıflandırma işlemine aktarılmıştır. Aşırı öğrenmeyi azaltmak amacıyla dropout katmanı kullanılmıştır.



Şekil 4.2.1. Temel CNN modeli için eğitim ve doğrulama doğruluk ve kayıp eğrileri.

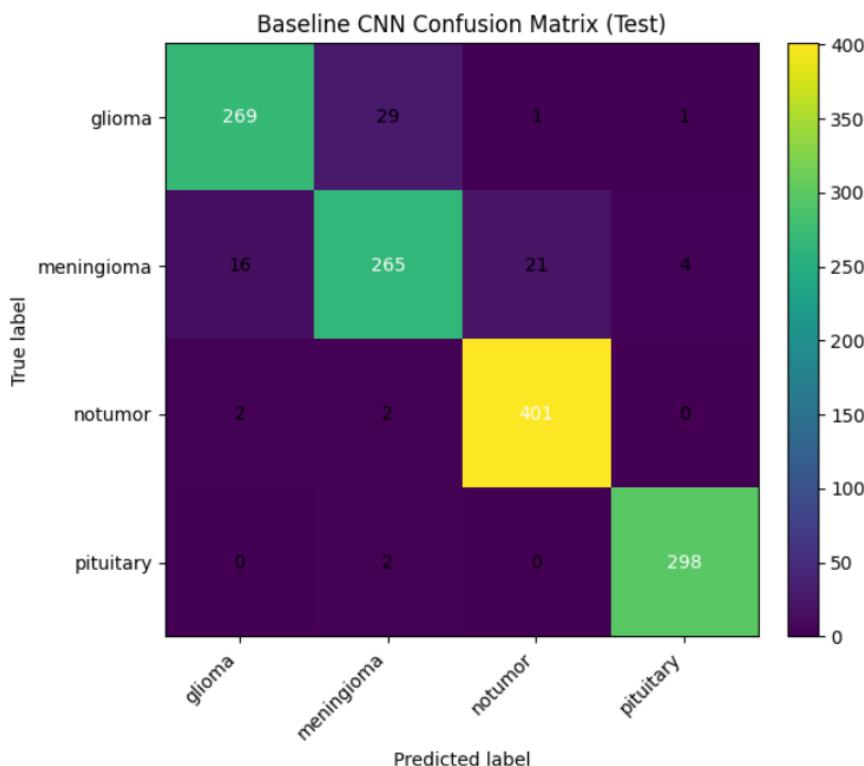
Model, Adam optimizasyon algoritması kullanılarak eğitilmiş; kayıp fonksiyonu olarak çok sınıflı sınıflandırma problemleri için uygun olan seyrek kategorik çapraz entropi tercih edilmiştir. Eğitim sürecinde sınıf dengesizliğinin etkisini azaltmak amacıyla sınıf ağırlıkları kullanılmıştır. Eğitim ve doğrulama doğrulukları ile kayıp değerlerinin epoch bazlı değişimini Şekil 4.2.1'de verilmiştir.

Eğitim süreci incelediğinde, modelin ilk epoch'lardan itibaren hızlı bir öğrenme eğrisi sergilediği ve eğitim doğruluğunun artan epoch sayısıyla birlikte istikrarlı biçimde yükseldiği görülmektedir. Doğrulama doğruluğu da eğitim doğruluğuna paralel bir seyir izleyerek yaklaşık %94 seviyesine ulaşmıştır. Eğitim ve doğrulama kayıp eğrileri arasındaki farkın sınırlı olması, modelin aşırı öğrenme göstermeden genelleme yeteneğini koruduğunu ortaya koymaktadır.

	Precision	Recall	F1-Score	Support
Glioma	0.9373	0.8967	0.9165	300
Meningioma	0.8893	0.8660	0.8775	306
Notumor	0.9480	0.9901	0.9686	405
Pituitary	0.9835	0.9933	0.9884	300
Accuracy			0.9405	1311
Macro Avg	0.9395	0.9365	0.9373	1311
Weighted Avg	0.9415	0.9405	0.9406	1311

Tablo 4.2.1 Temel CNN Modeli için Sınıf Bazlı Performans Metrikleri.

Eğitilen temel CNN modelinin performansı, daha önce ayrılmış test alt kümesi üzerinde değerlendirilmiş ve %94.05 doğruluk oranı elde edilmiştir. Genel doğruluk metriğine ek olarak, modelin sınıf bazlı performansını ayrıntılı biçimde incelemek amacıyla precision, recall ve F1-skoru değerleri hesaplanmıştır. Temel CNN modeline ait sınıf bazlı performans metrikleri Tablo 4.2.1'de verilmiştir.



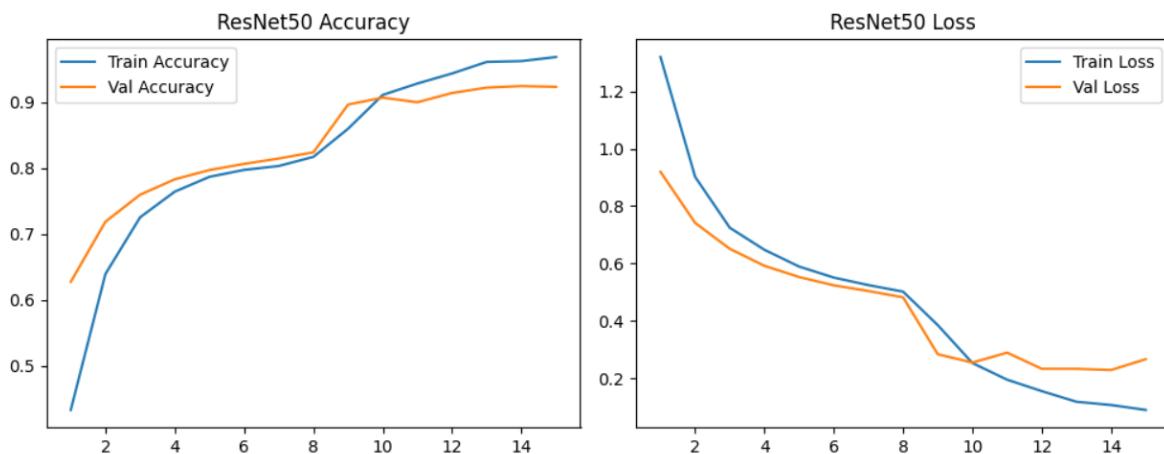
Şekil 4.2.2. Temel CNN modeli için karışıklık matrisi.

Modelin sınıflar arası hata dağılımını incelemek amacıyla karışıklık matrisi de analiz edilmiştir. Temel CNN modeline ait karışıklık matrisi Şekil 4.2.2'de verilmiştir. Karışıklık matrisi incelendiğinde, modelin özellikle *no tumor* ve *pituitary* sınıflarında yüksek doğru sınıflandırma oranlarına ulaştığı, sınıflar arası hataların ise daha çok glioma ve meningioma sınıfları arasında yoğunlaştığı görülmektedir.

Bu sonuçlar, temel CNN modelinin yalnızca genel doğruluk açısından değil, sınıf bazlı metrikler açısından da dengeli ve güçlü bir performans sergilediğini göstermektedir. Elde edilen bu performans, bir sonraki aşamada kullanılan transfer learning tabanlı modeller için sağlam bir referans noktası oluşturmaktadır.

4.3 Transfer Learning ile ResNet50 Modelinin Eğitimi ve Değerlendirilmesi

Bu çalışmada, transfer learning yaklaşımının beyin tümörü sınıflandırma performansına etkisini incelemek amacıyla ResNet50 mimarisini kullanılmıştır. Model, ImageNet veri seti üzerinde önceden eğitilmiş ağırlıklar ile başlatılmış ve iki aşamalı bir eğitim stratejisi izlenmiştir. İlk aşamada, ResNet50 taban modeli dondurularak yalnızca sınıflandırma katmanları eğitilmiş; ikinci aşamada ise ağıın son katmanları kademeli olarak açılarak ince ayar (fine-tuning) işlemi uygulanmıştır.



Şekil 4.3.1. ResNet50 modeli için eğitim ve doğrulama doğruluk ve kayıp eğrileri.

Eğitim süreci boyunca doğrulama doğruluğunun istikrarlı biçimde arttığı, fine-tuning aşaması sonrasında ise modelin beyin MR görüntülerine özgü ayrıt edici özelliklerini daha etkin biçimde öğrendiği gözlemlenmiştir. Eğitim ve doğrulama doğrulukları ile kayıp değerlerinin epoch bazlı değişimi Şekil 4.3.1'te verilmiştir.

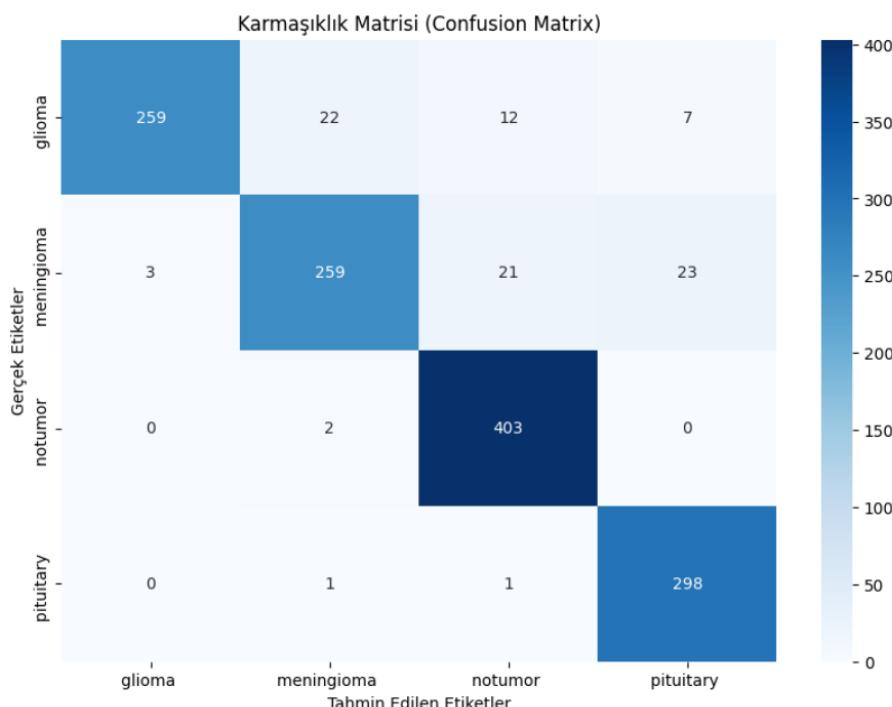
Eğitilen ResNet50 modelinin performansı, daha önce ayrılmış test alt kümesi üzerinde değerlendirilmiştir. Test sonuçlarına göre model, %92.98 doğruluk elde etmiştir. Sınıf bazlı performans metrikleri incelendiğinde; glioma sınıfı için 0.99 kesinlik (precision), 0.86 duyarlılık (recall) ve 0.92 F1-skoru; meningioma sınıfı için 0.91 precision, 0.85 recall ve 0.88 F1-skoru elde edilmiştir. No tumor sınıfında modelin performansının belirgin biçimde yüksek olduğu görülmüş; bu sınıf için 0.96 F1-skoru hesaplanmıştır. Pituitary tümörleri için ise 0.95 F1-skoru elde edilmiştir.

	Precision	Recall	F1-Score	Support
--	-----------	--------	----------	---------

Glioma	0.9885	0.8633	0.9217	300
Meningioma	0.9120	0.8464	0.8780	306
Notumor	0.9222	0.9951	0.9572	405
Pituitary	0.9085	0.9933	0.9490	300
Accuracy			0.9298	1311
Macro Avg	0.9328	0.9245	0.9265	1311
Weighted Avg	0.9319	0.9298	0.9287	1311

Tablo 4.3.1 ResNet50 modeli için sınıf bazlı precision, recall ve F1-skoru değerleri.

Modelin genel performansı değerlendirildiğinde, makro ortalama F1-skoru 0.93, ağırlıklı ortalama F1-skoru ise 0.93 olarak hesaplanmıştır. Bu değerler, modelin sınıflar arasında dengeli bir performans sergilediğini göstermektedir. Sınıf bazlı değerlendirme sonuçları Tablo 4.3.1'de verilmiştir.



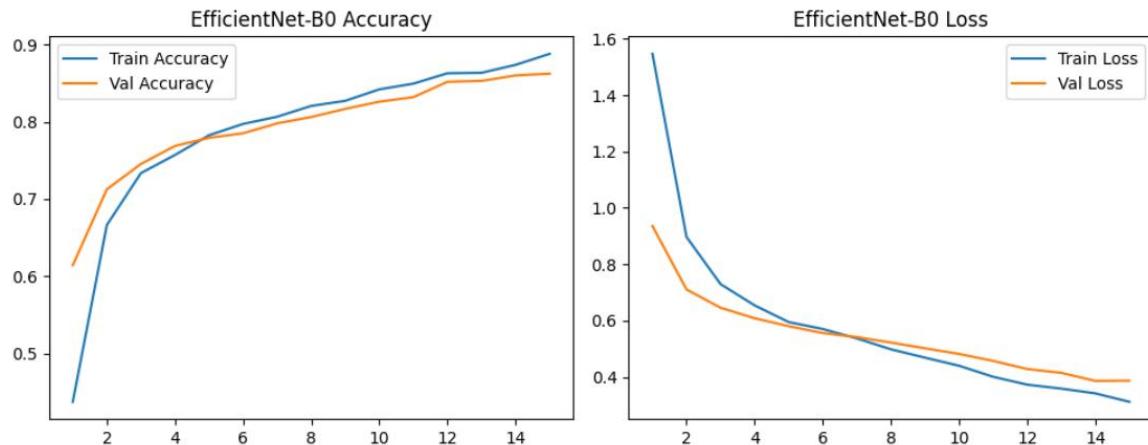
Şekil 4.3.2. ResNet50 modeli için karışıklık matrisi.

Karışıklık matrisi incelendiğinde, yanlış sınıflandırmaların büyük ölçüde glioma ve meningioma sınıfları arasında yoğunlaştığı görülmektedir. Buna karşın, no tumor ve pituitary sınıflarının büyük oranda doğru sınıflandırıldığı dikkat çekmektedir. Karışıklık matrisinin görsel temsili Şekil 4.3.2'te verilmiştir. Bu bulgular, bazı tümör türlerinin MR görüntülerinde benzer morfolojik özellikler gösterebilmesi nedeniyle sınıflar arası karışıklığın ortaya çıkabildiğini göstermektedir.

Elde edilen sonuçlar, ResNet50 mimarisinin transfer learning ve fine-tuning stratejileri ile birlikte kullanıldığından beyin tümörü sınıflandırma probleminde yüksek ve dengeli bir performans sunduğunu ortaya koymaktadır. Bu performans, temel CNN modeli ile birlikte değerlendirildiğinde, farklı model mimarilerinin sınıflandırma başarısına etkisinin objektif biçimde karşılaştırılmasına olanak tanımaktadır.

4.4 Transfer Learning ile EfficientNet-B0 Modelinin Eğitimi ve Değerlendirilmesi

Bu bölümde, transfer learning yaklaşımı kapsamında EfficientNet-B0 mimarisinin beyin tümörü sınıflandırma problemindeki performansı değerlendirilmiştir. EfficientNet-B0 modeli, ImageNet veri seti üzerinde önceden eğitilmiş ağırlıklar kullanılarak başlatılmış ve iki aşamalı bir eğitim stratejisi izlenmiştir. İlk aşamada modelin evrişimsel tabanı dondurularak yalnızca sınıflandırma katmanları eğitilmiş, ikinci aşamada ise ağın son katmanları kademeli olarak açılarak ince ayar (fine-tuning) işlemi uygulanmıştır.



Şekil 4.4.1. EfficientNet-B0 modeli için eğitim ve doğrulama doğruluk ve kayıp eğrileri.

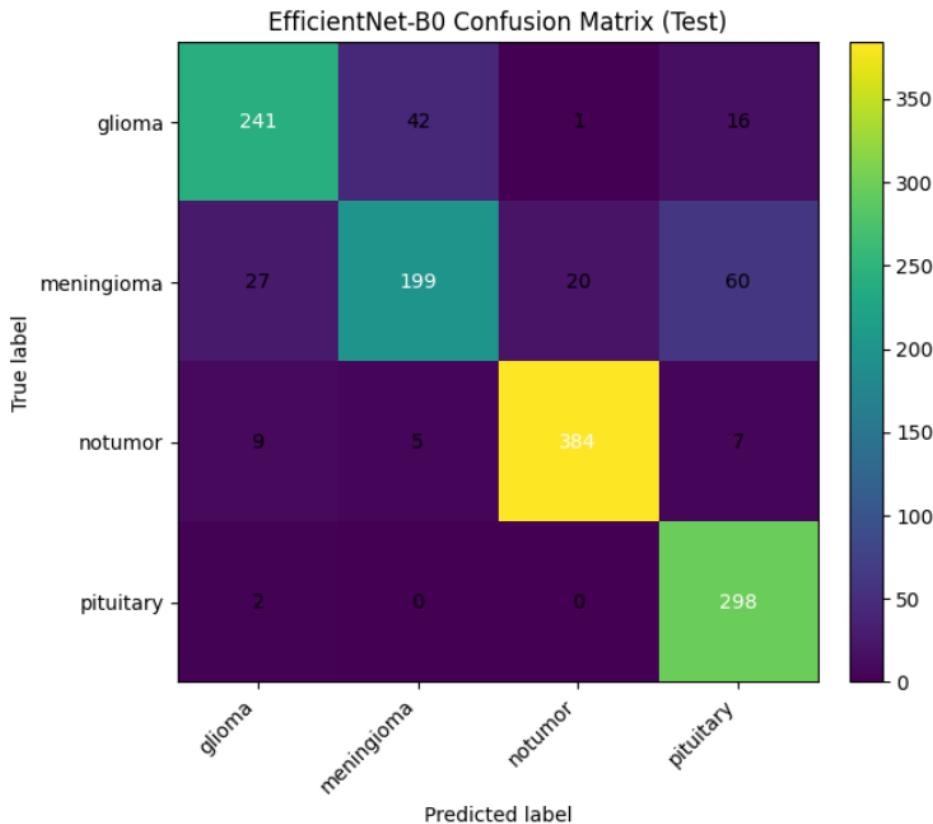
Eğitim süreci boyunca modelin eğitim ve doğrulama doğruluklarının istikrarlı bir artış gösterdiği, kayıp değerlerinin ise epoch sayısı arttıkça azaldığı gözlemlenmiştir. Eğitim ve doğrulama doğrulukları ile kayıp değerlerinin epoch bazlı değişimi Şekil 4.4.1'te verilmiştir. Bu eğriler, modelin öğrenme sürecinin kararlı ilerlediğini ve aşırı öğrenme belirtilerinin sınırlı düzeyde kaldığını göstermektedir.

	Precision	Recall	F1-score	Support
glioma	0.8638	0.8033	0.8325	300
meningioma	0.8089	0.6503	0.7210	306
notumor	0.9481	0.9481	0.9481	405
pituitary	0.7822	0.9933	0.8752	300
Accuracy			0.8558	1311
Macro avg	0.8508	0.8488	0.8442	1311
Weighted avg	0.8584	0.8558	0.8520	1311

Tablo 4.4.1 EfficientNet-B0 modeli için sınıf bazlı precision, recall ve F1-skoru değerleri.

EfficientNet-B0 modelinin performansı, daha önce ayrılmış test alt kümesi üzerinde değerlendirilmiştir. Test sonuçlarına göre model, %85.58 doğruluk elde etmiştir. Sınıf bazlı performans metrikleri incelendiğinde; glioma sınıfı için 0.83 F1-skoru, meningioma sınıfı için 0.72 F1-skoru, no tumor sınıfı için 0.95 F1-skoru ve pituitary sınıfı için 0.88 F1-skoru elde edilmiştir. Bu sonuçlar, modelin özellikle tümör bulunmayan görüntüleri yüksek doğrulukla sınıflandırabildiğini, buna karşın meningioma sınıfında görece daha düşük bir ayırt edicilik sergilediğini ortaya koymaktadır. Sınıf bazlı precision, recall ve F1-skoru değerleri Tablo 4.4.1'te verilmiştir.

Modelin genel performansı değerlendirildiğinde, makro ortalama F1-skoru 0.84, ağırlıklı ortalama F1-skoru ise 0.85 olarak hesaplanmıştır. Bu değerler, sınıflar arasında performans farklılıklarının bulunduğu ancak modelin genel olarak dengeli bir sınıflandırma yeteneği sunduğunu göstermektedir.



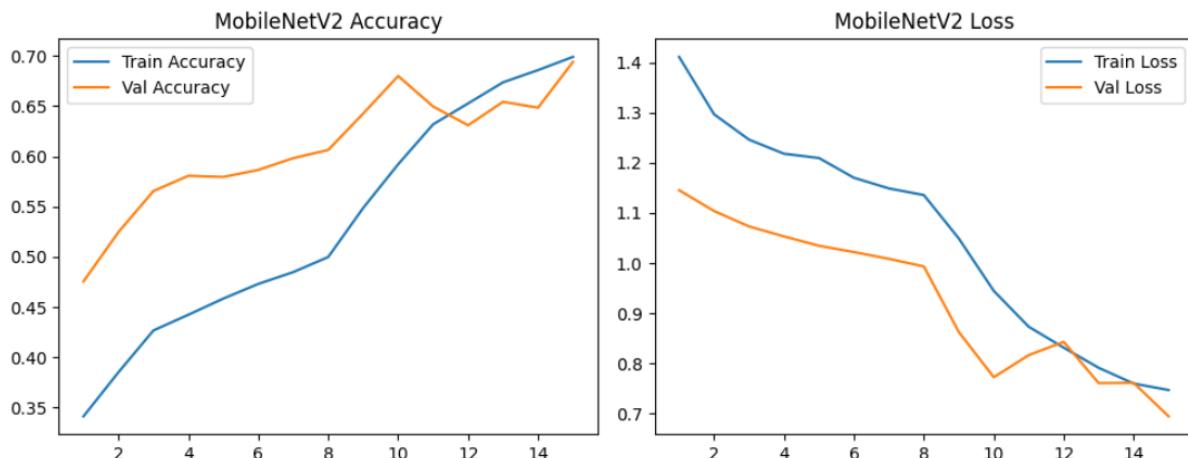
Şekil 4.4.2. EfficientNet-B0 modeli için karışıklık matrisi.

Karışıklık matrisi incelendiğinde, yanlış sınıflandırmaların özellikle glioma ve meningioma sınıfları arasında yoğunlaştığı görülmektedir. Buna karşılık, no tumor ve pituitary sınıflarının büyük oranda doğru sınıflandırıldığı dikkat çekmektedir. EfficientNet-B0 modeline ait karışıklık matrisi Şekil 4.4.2'da verilmiştir. Bu durum, bazı tümör türlerinin MR görüntülerinde benzer morfolojik özellikler gösterebilmesi nedeniyle sınıflar arası karışıklıkların ortaya çıkabildiğini düşündürmektedir.

Elde edilen sonuçlar, EfficientNet-B0 mimarisinin hesaplama maliyeti açısından verimli bir yapı sunmasına rağmen, bu çalışma kapsamında kullanılan veri seti üzerinde temel CNN ve ResNet50 modellerine kıyasla daha düşük bir sınıflandırma performansı sergilediğini göstermektedir. Bu bulgu, model mimarisinin karmaşıklığı ile veri setinin özellikleri arasındaki uyumun sınıflandırma başarısında belirleyici bir rol oynadığını ortaya koymaktadır.

4.5 Transfer Learning ile MobileNetV2 Modelinin Eğitimi ve Değerlendirilmesi

Bu bölümde, hesaplama maliyeti ve model boyutu açısından daha hafif bir mimari sunan MobileNetV2 modelinin beyin tümörü sınıflandırma problemindeki performansı değerlendirilmiştir. MobileNetV2 mimarisi, derinlik ayrılabilir evrişim (depthwise separable convolution) yapısı sayesinde daha az parametre ile çalışabilen bir yapı sunmakta olup, performans-verimlilik dengesinin incelenmesi amacıyla deneySEL çalışmalarına dahil edilmiştir.



Şekil 4.5.1. MobileNetV2 modeli için eğitim ve doğrulama doğruluk ve kayıp eğrileri.

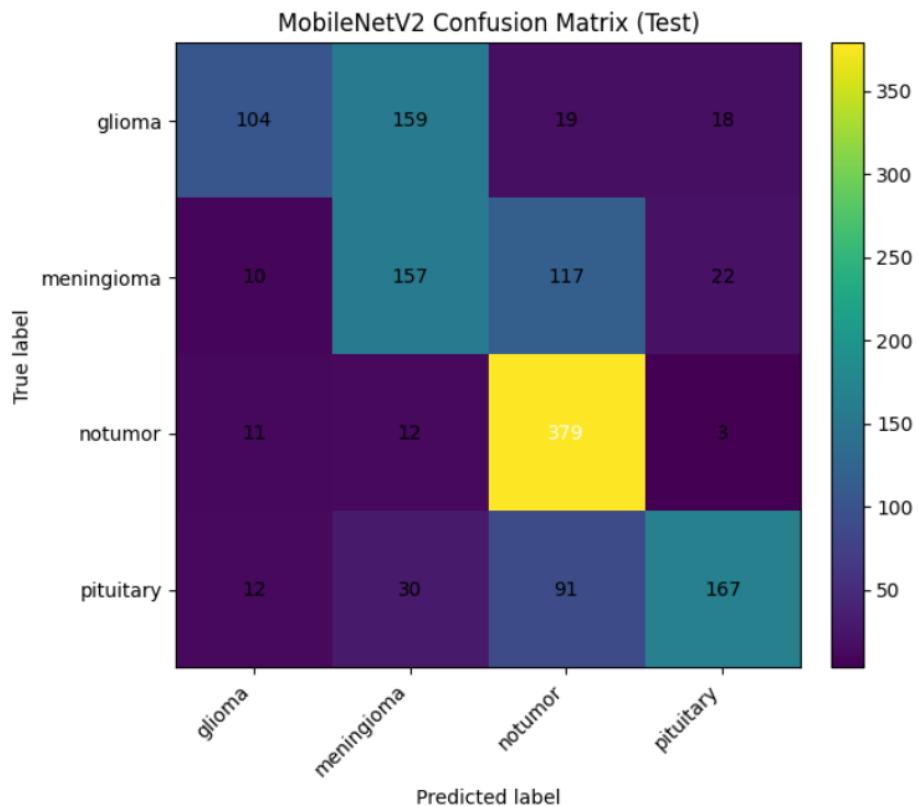
MobileNetV2 modeli, ImageNet veri seti üzerinde önceden eğitilmiş ağırlıklar kullanılarak başlatılmış ve iki aşamalı bir eğitim stratejisi izlenmiştir. İlk aşamada modelin evrişimsel tabanı dondurularak yalnızca sınıflandırma katmanları eğitilmiş; ikinci aşamada ise ağın son katmanları açılarak ince ayar (fine-tuning) işlemi uygulanmıştır. Eğitim ve doğrulama doğrulukları ile kayıp değerlerinin epoch bazlı değişimi Şekil 4.5.1.'de verilmiştir.

Eğitim süreci incelediğinde, modelin doğruluk değerlerinin epoch sayısı arttıkça kademeli olarak yükseldiği, ancak doğrulama doğruluğunun sınırlı bir seviyede kaldığı görülmektedir. Bu durum, MobileNetV2 mimarisinin daha karmaşık uzamsal örüntülerini öğrenme kapasitesinin sınırlı olabileceğine işaret etmektedir.

	Precision	Recall	F1-score	Support
Glioma	0.7591	0.3467	0.4760	300
Meningioma	0.4385	0.5131	0.4729	306
Notumor	0.6254	0.9358	0.7498	405
Pituitary	0.7952	0.5567	0.6549	300
Accuracy			0.6156	1311
Macro Avg	0.6546	0.5881	0.5884	1311
Weighted Avg	0.6513	0.6156	0.6008	1311

Tablo 4.5.1 MobileNetV2 modeli için sınıf bazlı precision, recall ve F1-skoru değerleri.

MobileNetV2 modelinin performansı, test alt kümesi üzerinde değerlendirilmiş ve %61.56 doğruluk elde edilmiştir. Sınıf bazlı performans metrikleri incelediğinde; *no tumor* sınıfında 0.75 F1-skoru ile görecek yüksek bir başarı elde edilirken, *glioma* (0.48 F1-skoru) ve *meningioma* (0.47 F1-skoru) sınıflarında performansın belirgin biçimde düşüğü görülmektedir. *Pituitary* sınıfı için ise 0.65 F1-skoru hesaplanmıştır. Sınıf bazlı precision, recall ve F1-skoru değerleri Tablo 4.5.1'te verilmiştir.



Şekil 4.5.2. MobileNetV2 modeli için karışıklık matrisi.

Modelin genel performansı değerlendirildiğinde, makro ortalama F1-skoru 0.59, ağırlıklı ortalama F1-skoru ise 0.60 olarak hesaplanmıştır. Bu değerler, sınıflar arasında performans dengesizliğinin bulunduğu göstermektedir. MobileNetV2 modeline ait karışıklık matrisi Şekil 4.5.2.'de verilmiştir. Karışıklık matrisi incelendiğinde, özellikle glioma ve meningioma sınıflarının sıkılıkla birbirleriyle karıştırıldığı, *no tumor* sınıfının ise büyük ölçüde doğru sınıflandırıldığı görülmektedir.

Elde edilen sonuçlar, MobileNetV2 mimarisinin düşük hesaplama maliyeti ve hafif yapısına karşın, bu çalışma kapsamında kullanılan veri seti üzerinde temel CNN, ResNet50 ve EfficientNet-B0 modellerine kıyasla daha düşük bir sınıflandırma performansı sergilediğini ortaya koymaktadır. Bu bulgu, modelin verimlilik avantajına rağmen, temsil gücü gerektiren tıbbi görüntü sınıflandırma problemlerinde daha derin ve karmaşık mimarilerin daha uygun olabileceğini göstermektedir.

4.6 Model Karşılaştırma ve Nihai Model Seçimi

Bu bölümde, önceki alt bölümlerde eğitilen tüm derin öğrenme modellerinin performansları karşılaştırılmış olarak değerlendirilmiş ve beyin tümörü sınıflandırma problemi için en uygun model belirlenmiştir. Karşılaştırma sürecinde, yalnızca genel doğruluk (accuracy) metriği değil; aynı zamanda sınıf bazlı performansı yansıtan makro ortalama ve ağırlıklı ortalama F1-skorları da dikkate alınmıştır. Bu yaklaşım, veri setinde yer alan sınıflar arasındaki örnek sayısı farklılıklarının değerlendirme sonuçları üzerindeki etkisini denelemeyi amaçlamaktadır.

Bu çalışma kapsamında değerlendirilen modeller; sıfırdan eğitilen temel CNN modeli, transfer learning yaklaşımı ile eğitilen ResNet50, EfficientNet-B0 ve MobileNetV2 mimarileridir. Her bir model, aynı veri seti, benzer eğitim stratejileri ve değerlendirme

metrikleri kullanılarak test edilmiştir. Elde edilen test sonuçlarına ait karşılaştırmalı performans değerleri Tablo 4.6.1’te sunulmaktadır.

Model	Accuracy	Macro F1	Weighted F1
Baseline CNN	0.9405	0.9373	0.9406
ResNet50 (Transfer Learning)	0.9298	0.9265	0.9287
EfficientNet-B0 (Transfer Learning)	0.8558	0.8442	0.8520
MobileNetV2 (Transfer Learning)	0.6159	0.5884	0.6008

Tablo 4.6.1 Kullanılan Modellerin Karşılaştırmalı Performans Analizi.

Tablo 4.6.1 incelendiğinde, sıfırdan eğitilen temel CNN modelinin tüm değerlendirme metriklerinde en yüksek performansı sergilediği görülmektedir. ResNet50 modeli, sınıf bazlı metrikler açısından dengeli bir performans sunsa da temel CNN modelinin genel doğruluk ve F1-skoru değerlerini aşamamıştır. EfficientNet-B0 modeli, daha verimli bir mimari sunmasına rağmen performans açısından ilk iki modelin gerisinde kalmıştır. MobileNetV2 modeli ise düşük hesaplama maliyetine sahip olmasına karşın, sınıf bazlı ayırt ediciliğinin sınırlı olması nedeniyle en düşük performansı göstermiştir.

Bu çalışmada eğitilen modellerin performansları karşılaştırıldığında, temel CNN modeli test doğruluğu (%94.05), makro F1-skoru (0.94) ve ağırlıklı F1-skoru (0.94) değerleri ile en yüksek performansı sergilemiştir. ResNet50 modeli, sınıf bazlı metrikler açısından dengeli bir performans sunsa da temel CNN modelinin genel doğruluk ve F1-skoru değerlerini aşamamıştır. EfficientNet-B0 ve MobileNetV2 modelleri ise daha düşük performans göstermiştir.

Bu nedenle, hem genel doğruluk hem de sınıf bazlı performans metrikleri dikkate alındığında, temel CNN modeli nihai model olarak seçilmiştir. Seçilen model, beyin tümörü sınıflandırma problemi için en yüksek ayırt ediciliği sağlamış ve genelleme yeteneğini koruduğunu göstermiştir.

5. Sonuç

Bu çalışmada, manyetik rezonans (MR) görüntülerinden beyin tümörü sınıflandırması amacıyla derin öğrenme tabanlı farklı evrişimsel sinir ağrı mimarileri karşılaştırmalı olarak incelenmiştir. Çalışma kapsamında, dört sınıfından oluşan (glioma, meningioma, pituitary ve no tumor) çok sınıflı bir görüntü sınıflandırma problemi ele alınmış; sıfırdan eğitilen temel bir CNN modeli ile birlikte transfer learning yaklaşımı kullanılarak ResNet50, EfficientNet-B0 ve MobileNetV2 mimarileri değerlendirilmiştir.

Deneysel sonuçlar, temel CNN modelinin test alt kümesi üzerinde %94.05 doğruluk, 0.94 makro F1-skoru ve 0.94 ağırlıklı F1-skoru değerleri ile en yüksek performansı sergilediğini göstermiştir. ResNet50 modeli, sınıf bazlı metrikler açısından dengeli bir performans sunsa da temel CNN modelinin genel başarısını aşamamıştır. EfficientNet-B0 modeli, daha verimli bir mimari sunmasına rağmen performans açısından ilk iki modelin gerisinde kalmıştır.

MobileNetV2 modeli ise düşük hesaplama maliyeti ve hafif yapısına karşın, özellikle tümör türleri arasındaki ayırt edici özelliklerini öğrenme konusunda sınırlı bir performans göstermiştir.

Sınıf bazlı analizler ve karışıklık matrisleri incelendiğinde, tüm modellerde *glioma* ve *meningioma* sınıfları arasında belirli düzeyde karışıklık gözlemlendiği; buna karşılık *no tumor* ve *pituitary* sınıflarının daha yüksek doğrulukla sınıflandırıldığı görülmüştür. Bu durum, söz konusu tümör türlerinin MR görüntülerinde benzer morfolojik özellikler sergileyebilmesi ile ilişkilendirilebilir.

Elde edilen bulgular, derin öğrenme tabanlı yöntemlerin beyin tümörü sınıflandırma problemlerinde yüksek doğruluk ve genelleme yeteneğini sunabildiğini ortaya koymaktadır. Ayrıca, bu çalışmada olduğu gibi veri setinin özellikleriyle uyumlu bir mimari seçiminin, model karmaşıklığından bağımsız olarak performans üzerinde belirleyici bir etkiye sahip olabileceği görülmüştür.

Bu çalışmanın bazı sınırlılıkları bulunmaktadır. Kullanılan veri seti, görüntü sayısı açısından orta ölçekli olup, daha büyük ve çok merkezli veri setleri ile yapılan çalışmalar modelin genellenebilirliğini artırabilir. Ayrıca, yalnızca 2B MR görüntülerinin kullanılması, tümörlerin üç boyutlu yapısının tam olarak modellenmesini sınırlamaktadır. Gelecek çalışmalarda 3B görüntüleme verilerinin kullanılması, çoklu MR sekanslarının (T1, T2, FLAIR) birlikte değerlendirilmesi ve dikkat mekanizmaları (attention-based models) gibi daha gelişmiş mimarilerin incelenmesi planlanabilir.

Sonuç olarak, bu çalışma kapsamında geliştirilen derin öğrenme tabanlı yaklaşımın, beyin tümörü sınıflandırma probleminde etkili ve güvenilir sonuçlar ürettiği gösterilmiştir. Nihai model olarak seçilen temel CNN mimarisi, yüksek doğruluk ve dengeli sınıf bazlı performansı ile klinik karar destek sistemlerine entegre edilebilecek potansiyele sahip bir çözüm sunmaktadır.

6. Kaynakça

- 1) Y. LeCun, Y. Bengio ve G. Hinton, “Deep learning,” *Nature*, c. 521, s. 436–444, 2015.
- 2) A. Krizhevsky, I. Sutskever ve G. E. Hinton, “ImageNet classification with deep convolutional neural networks,” *Advances in Neural Information Processing Systems (NeurIPS)*, ss. 1097–1105, 2012.
- 3) K. He, X. Zhang, S. Ren ve J. Sun, “Deep residual learning for image recognition,” *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, ss. 770–778, 2016.
- 4) M. Tan ve Q. V. Le, “EfficientNet: Rethinking model scaling for convolutional neural networks,” *Proceedings of the 36th International Conference on Machine Learning (ICML)*, ss. 6105–6114, 2019.
- 5) M. Sandler, A. Howard, M. Zhu, A. Zhmoginov ve L. C. Chen, “MobileNetV2: Inverted residuals and linear bottlenecks,” *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, ss. 4510–4520, 2018.
- 6) O. Ronneberger, P. Fischer ve T. Brox, “U-Net: Convolutional networks for biomedical image segmentation,” *Medical Image Computing and Computer-Assisted Intervention (MICCAI)*, ss. 234–241, 2015.
- 7) S. Pereira, A. Pinto, V. Alves ve C. A. Silva, “Brain tumor segmentation using convolutional neural networks in MRI images,” *IEEE Transactions on Medical Imaging*, c. 35, no. 5, ss. 1240–1251, 2016.
- 8) H. A. T. Ashour, M. M. Dessouky ve A. S. AboEleneen, “Brain tumor classification using deep learning and MRI images,” *International Journal of Advanced Computer Science and Applications*, c. 11, no. 6, ss. 417–424, 2020.
- 9) I. Goodfellow, Y. Bengio ve A. Courville, *Deep Learning*. Cambridge, MA, ABD: MIT Press, 2016.
- 10) F. Chollet, *Deep Learning with Python*. New York, NY, ABD: Manning Publications, 2017.
- 11) M. Nickparvar, “Brain Tumor MRI Dataset,” Kaggle, 2021.
<https://www.kaggle.com/datasets/masoudnickparvar/brain-tumor-mri-dataset>