

MRI GÖRÜNTÜLERİNDEN BEYİN TÜMÖRÜ SINIFLANDIRMASI

5 FARKLI GELİŞMİŞ (ADVANCED) CNN MİMARILERİNİN KARŞILAŞTIRMALI ANALİZİ

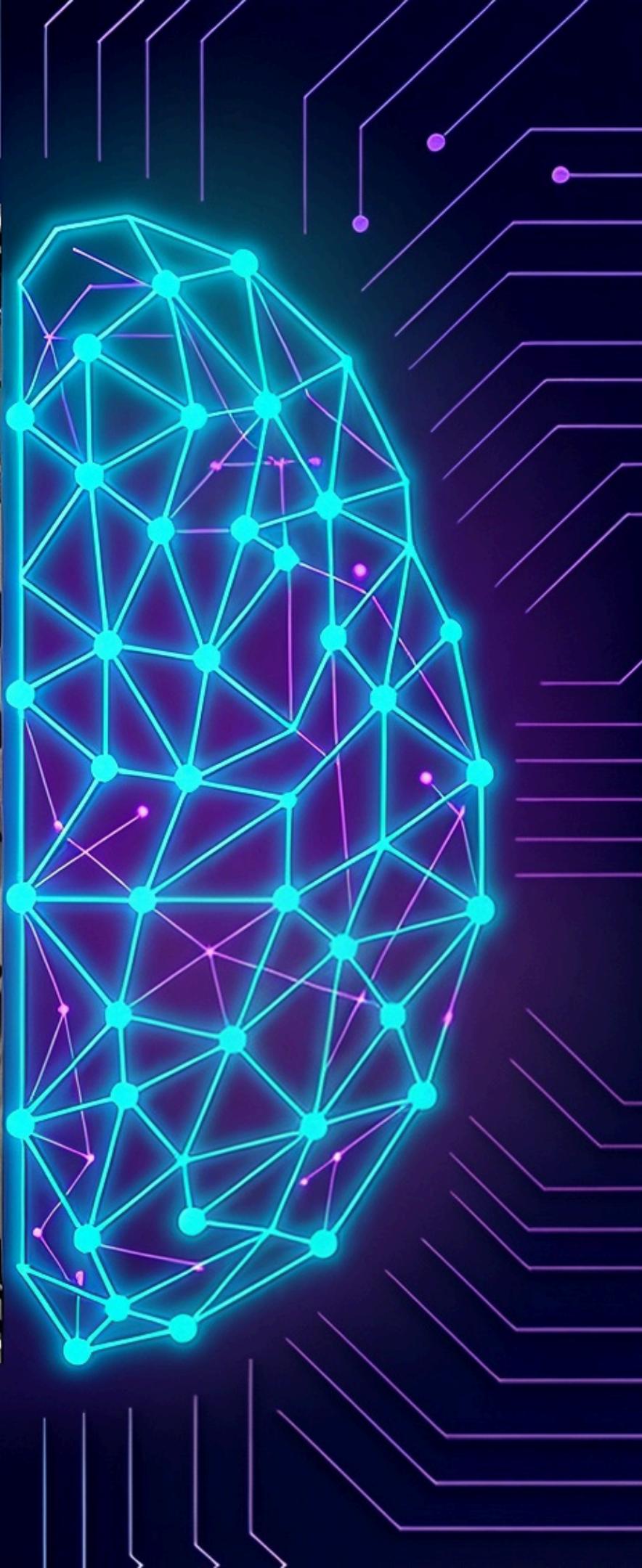
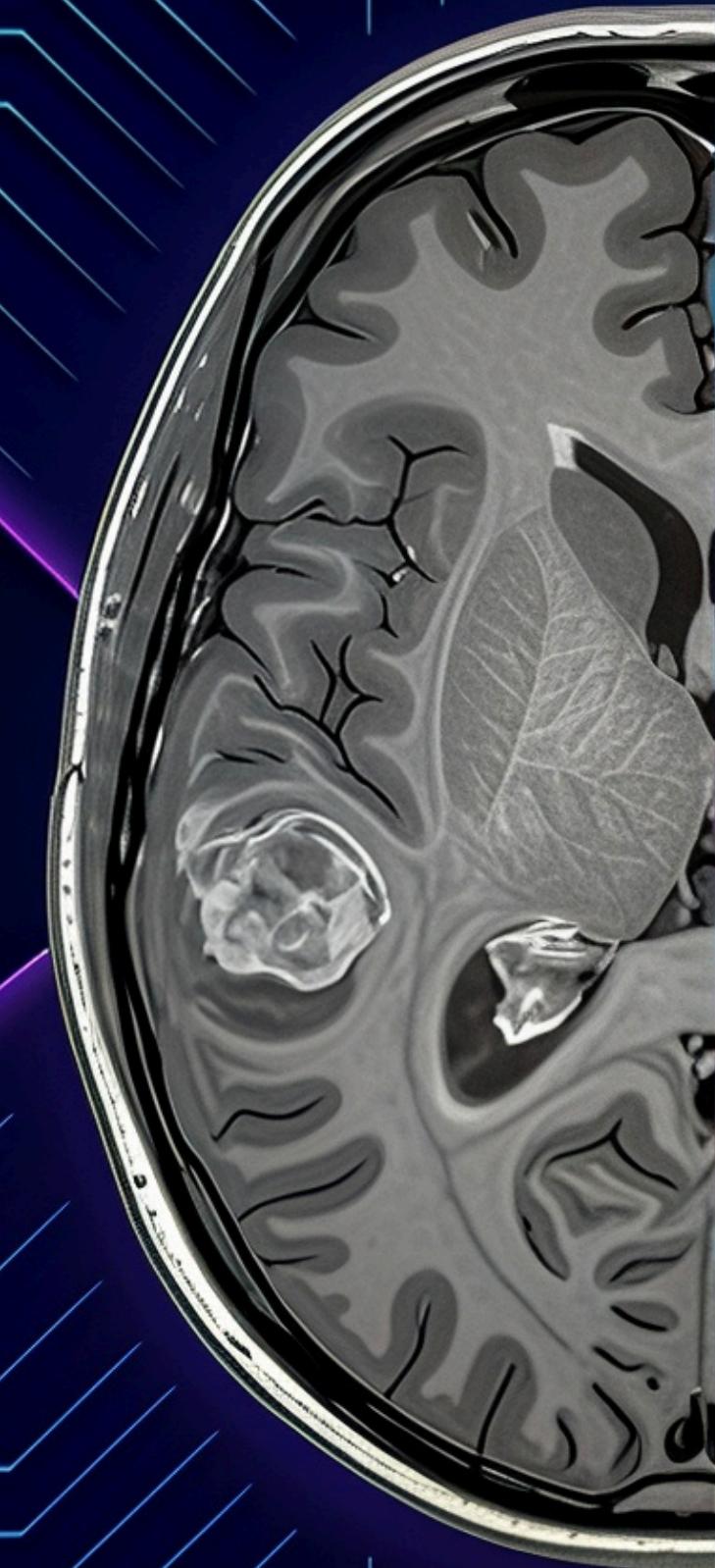


Ekip Üyeleri

Mehmet Ali Esmer	(23040101013)
Arda Ölmez	(23040101019)
Bilal Furkan Çakırgöz	(23040101021)
Burak Şahin	(23040101072)
Muhammed Emin Benzer	(23040101045)

github.com/olmezarda/FET312_BrainTumor_Project

youtube.com/watch?v=5-OFyU_zyw0&t=1s



PROBLEM TANIMI VE ÇÖZÜM YAKLAŞIMI

Manuel Teşhisin Sınırları ve Yapay Zeka Desteği

Tıbbi görüntüleme verilerinin her geçen gün artması, uzman radyologların üzerindeki iş yükünü kritik seviyelere taşıımıştir. Mevcut manuel teşhis yöntemleri, hem zaman alıcıdır hem de insan gözünün yorgunluğuna bağlı hata risklerini beraberinde getirir.

Projemiz, bu darboğazı aşmak için Derin Öğrenme teknolojilerini devreye almaktadır. Geliştirdiğimiz yapay zeka modelleri, insan gözünün yakalamakta zorlanacağı mikroskobik doku farklarını analiz ederek; radyologlara hızlı, tutarlı ve %95 üzeri doğrulukla çalışan bir 'İkinci Görüş' desteği sunmayı ve teşhis sürecindeki insan hatası faktörünü minimize etmeyi hedeflemektedir.



VERİ SETİ: HAMMADDEMİZ NEDİR?

Veri Kaynağı

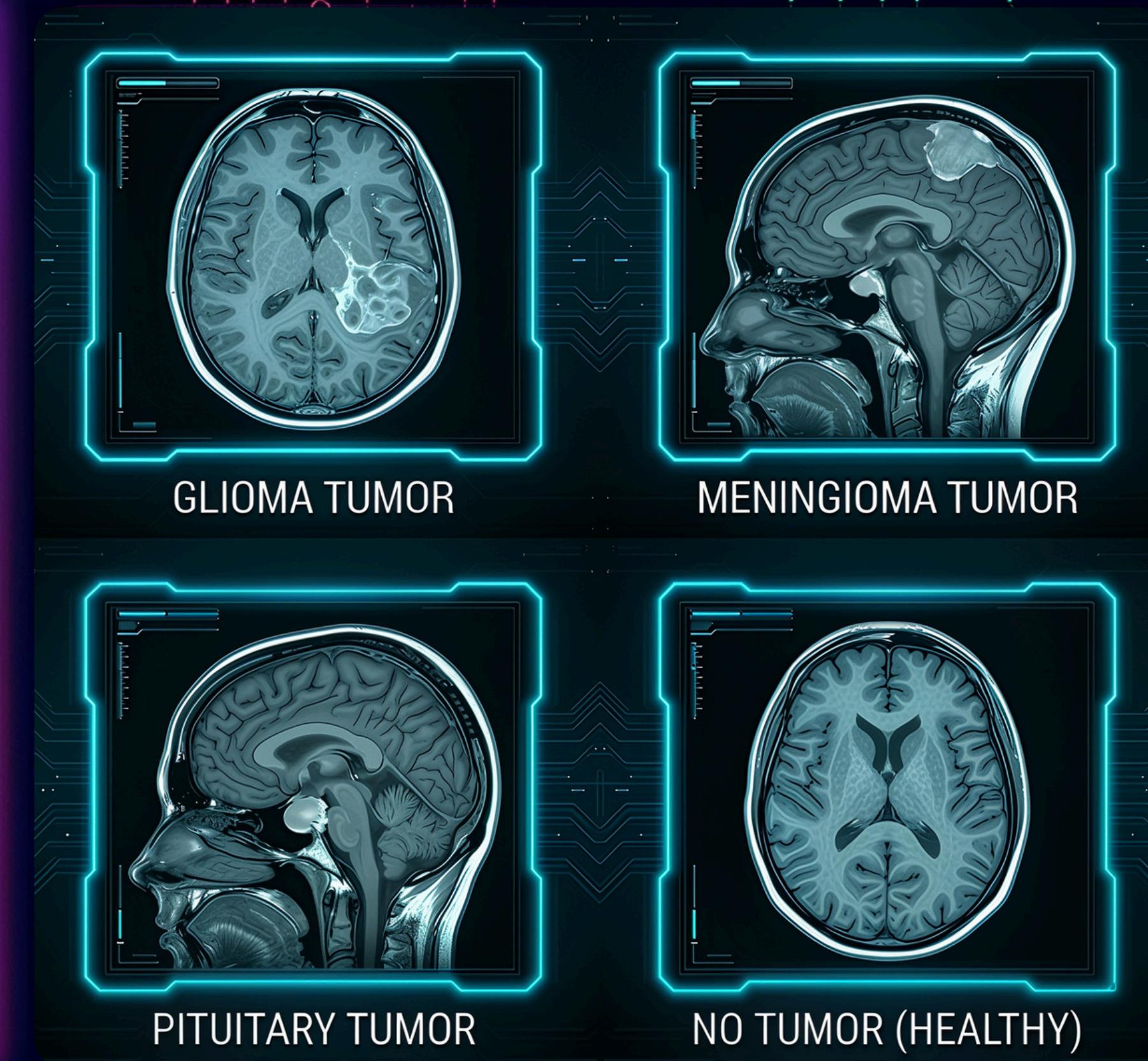
Projemizin eğitim ve test aşamalarında, tıbbi yapay zeka literatüründe güvenilirliği kanıtlanmış olan ve Kaggle platformunda araştırmacıların erişimine sunulan 'Brain Tumor MRI Dataset' kullanılmıştır.

Veri seti, beyin anatomisinin detaylı incelenmesine olanak tanıyan toplam 7022 adet T1-ağırlıklı

Manyetik Rezonans (MR) görüntüsünden oluşmaktadır. Bu geniş veri havuzu, modellerimizin farklı tümör tiplerini ve sağlıklı doku varyasyonlarını ezberlemeden (overfitting olmadan) öğrenebilmesi için gereken istatistiksel çeşitliliği ve görsel zenginliği sağlamaktadır.

Veri Seti Kaynağı:

kaggle.com/datasets/masoudnickparvar/brain-tumor-mri-dataset



VERİ SETİ: HAMMADDEMİZ NEDİR?

Veri Setindeki Teknik Zorluklar

- **Multi-Class Sınıflandırma Yapısı:** Veri setimiz; Glioma, Meningioma, Pituitary (Hipofiz) ve No Tumor olmak üzere 4 farklı etiketten oluşan bir "Multi-Class Classification" problemidir. Modelden bekłentimiz, sadece tümör var/yok demek değil, tümörün türünü de yüksek hassasiyetle ayırt etmesidir.
- **Kritik Engel:** Inter-class Similarity (Sınıflar Arası Benzerlik): Projenin en büyük zorluğu, radyolojik görünüm açısından Glioma ve Meningioma sınıflarının morfolojik olarak birbirine aşırı benzemesidir. Bu durum, modelin ayırt edici öznitelikleri (features) öğrenmesini zorlaştıran bir "Fine-Grained Classification" problemi yaratmaktadır.



GLIOMA TUMOR



MENINGIOMA TUMOR



PITUITARY TUMOR

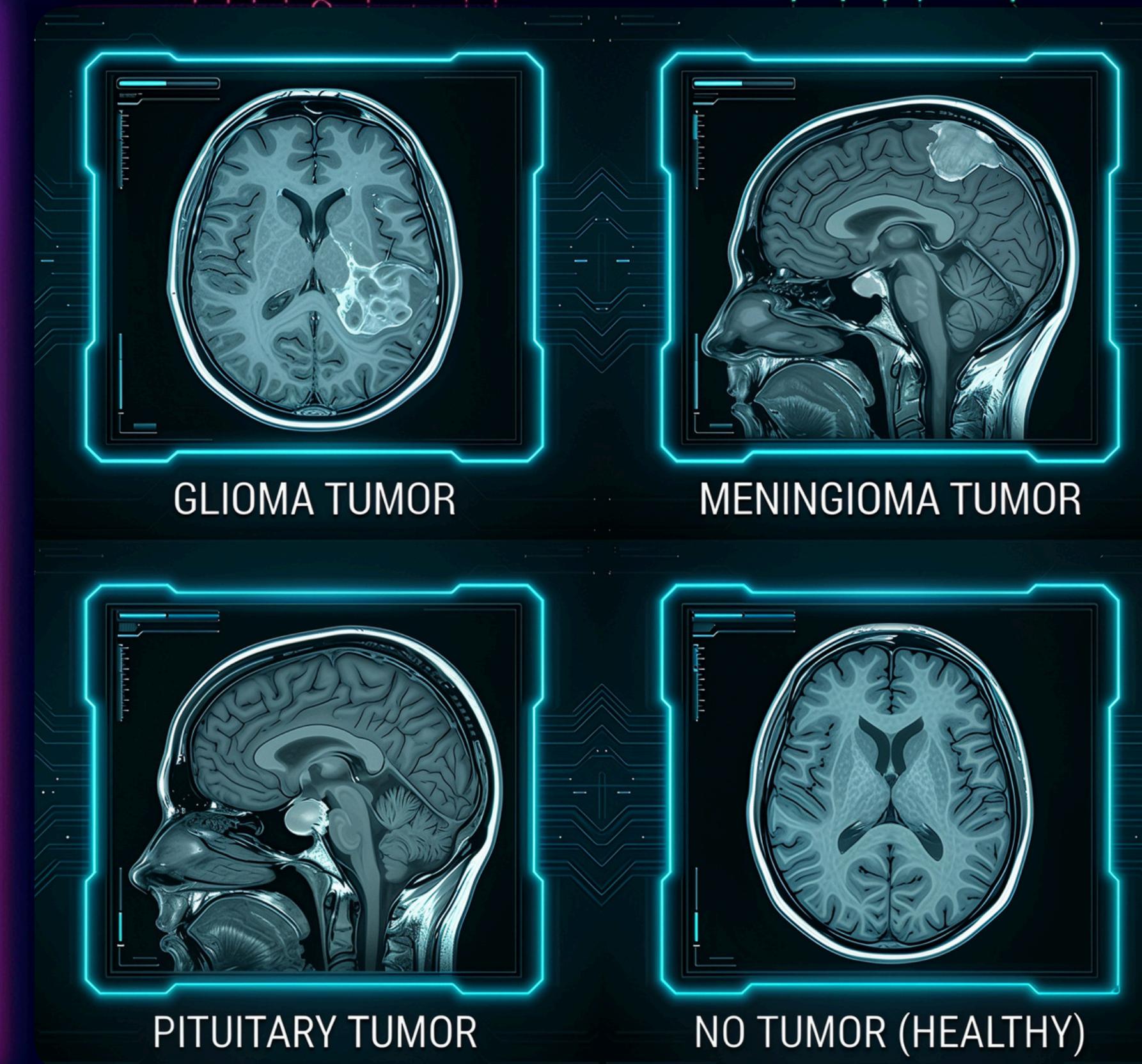


NO TUMOR (HEALTHY)

VERİ SETİ: HAMMADDEMİZ NEDİR?

Veri Setindeki Teknik Zorluklar

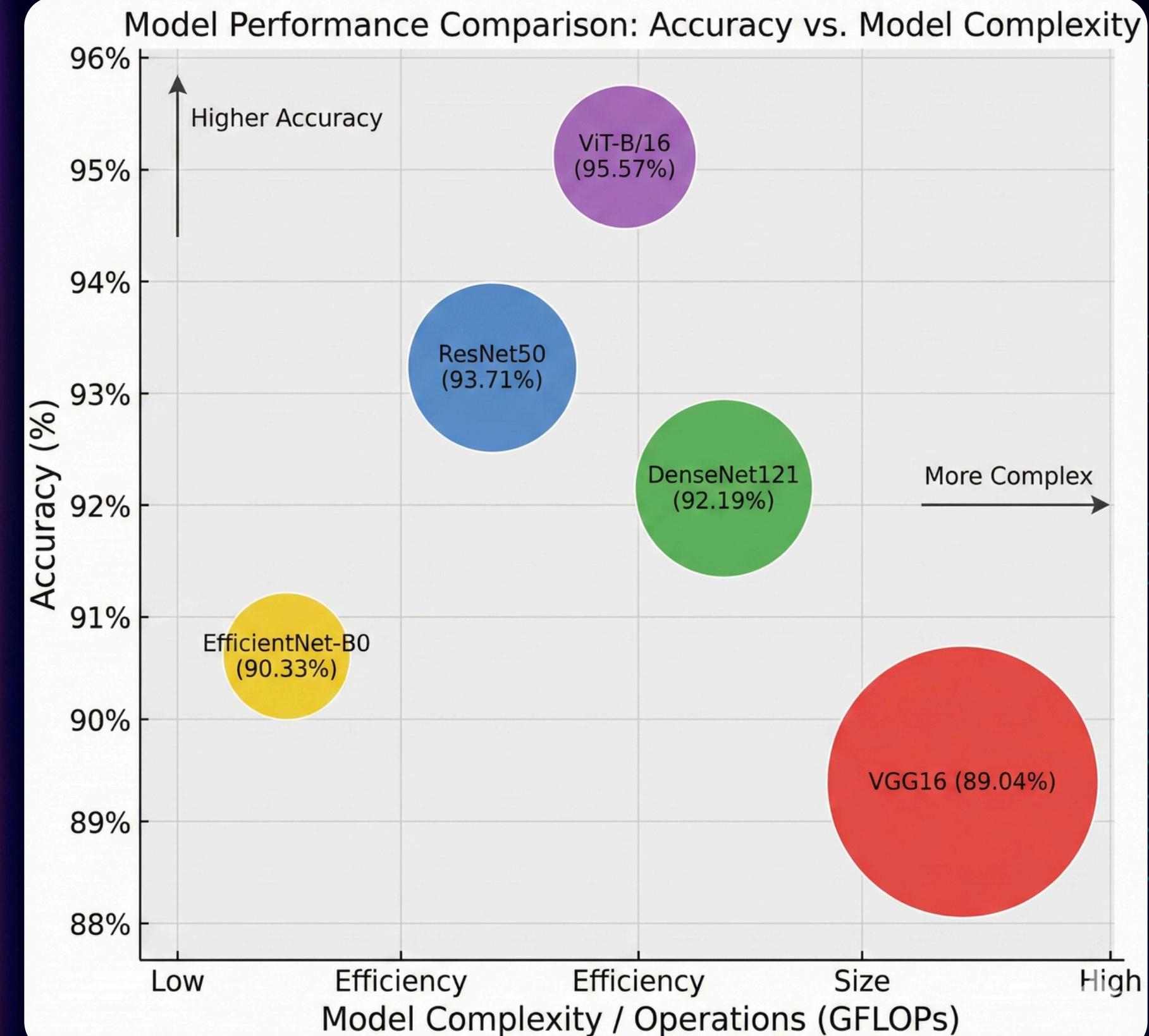
- **Dengesizlik Yönetimi (Handling Imbalance):** Veri setindeki sınıfların örnek sayıları eşit değildir. Bu durum, modelin çoğunluk sınıfına (Majority Class) eğilimli olmasına (Model Bias) yol açabilirdi. Bunu engellemek için iki kritik teknik uygulandı:
- **Class Weighting:** Kayıp fonksiyonuna (Loss Function) her sınıf için ters orantılı ağırlıklar verilerek azınlık sınıfların etkisi artırıldı.
- **Data Augmentation:** Eğitim verisi; rastgele döndürme, yansıtma ve parlaklık değişimleri ile yapay olarak çoğaltılarak modelin genelleme (Generalization) yeteneği güçlendirildi.



MİMARI YAKLAŞIMLAR VE MODEL SEÇİMİ

5 Farklı Yaklaşım, Tek Ortak Hedef

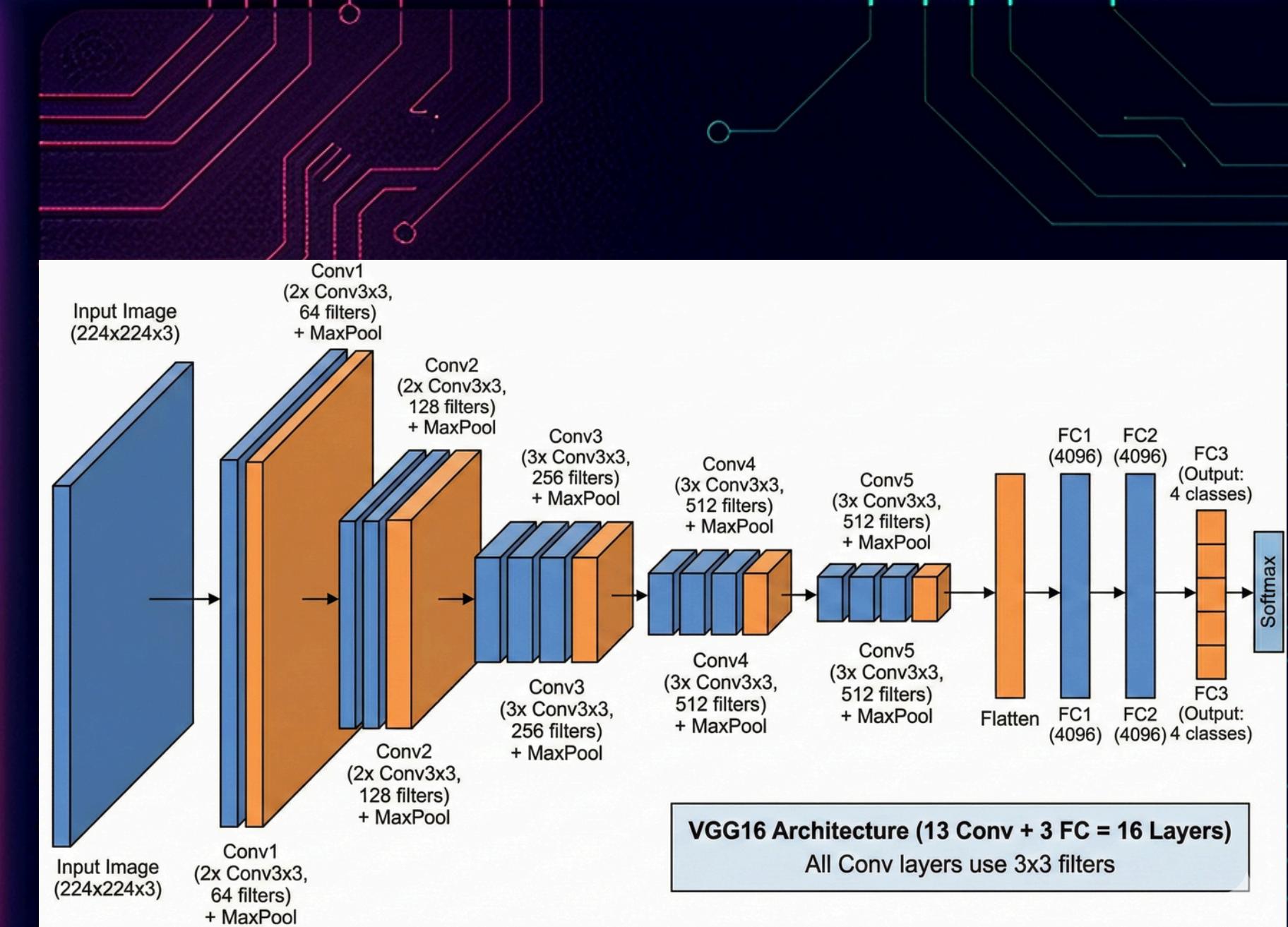
- **VGG16:** Derin öğrenmenin klasik ve güçlü standarı.
- **ResNet50:** "Residual" bloklar ile derinlik problemini çözen mimari.
- **DenseNet121:** Katmanlar arası maksimum bilgi akışı sağlayan yapı.
- **EfficientNet-B0:** Performans ve kaynak verimliliğini optimize eden model.
- **Vision Transformer (ViT):** Görüntüyü piksel piksel değil, "Attention" mekanizmasıyla bütünüyle işleyen devrimsel mimari.



VGG16: DERİN ÖĞRENMEDE KLASİK GÜC

Görüntü İşlemenin "Altın Standardı"

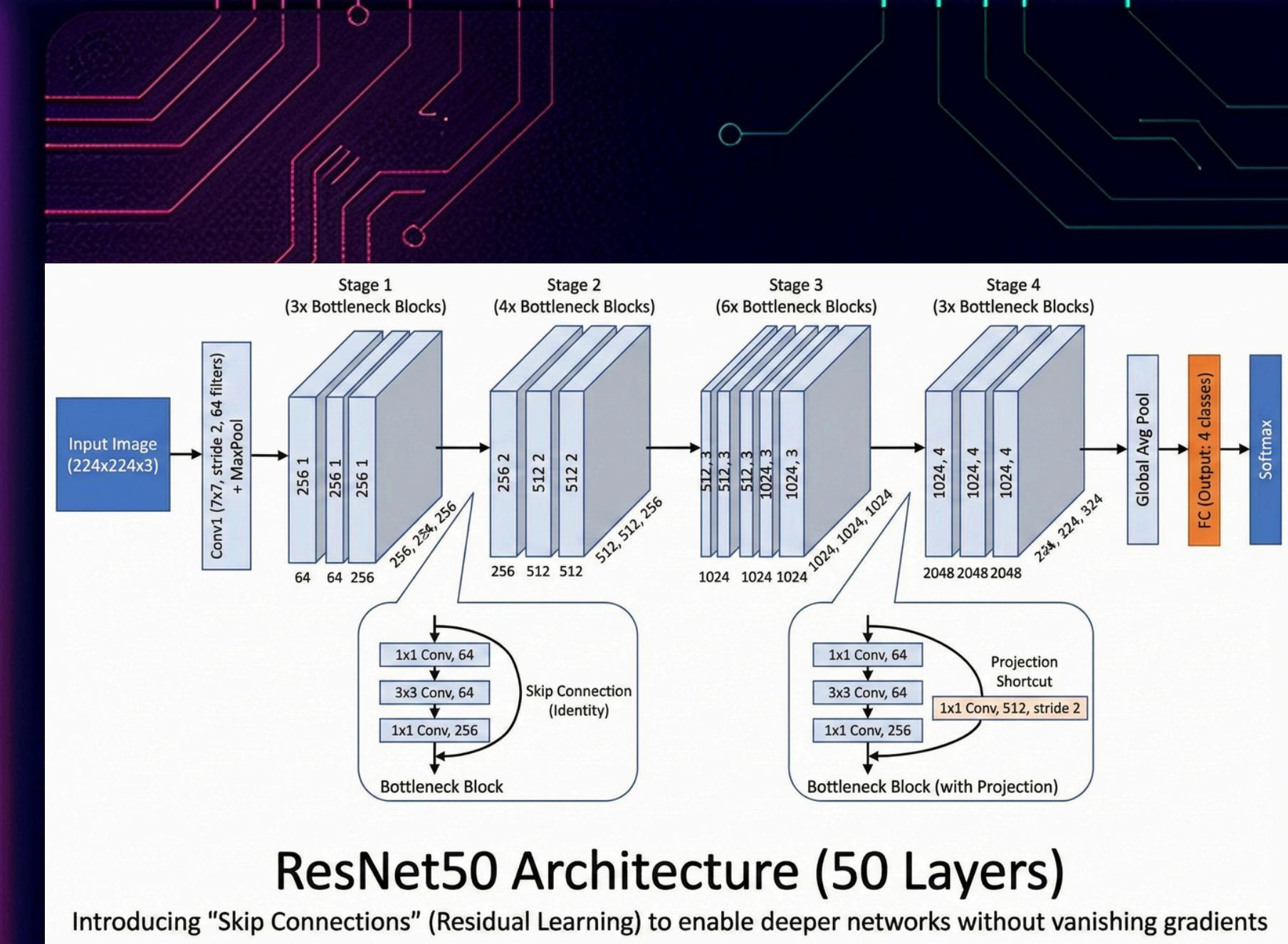
- **Mimari Yapı:** Oxford Üniversitesi (Visual Geometry Group) tarafından geliştirilen bu model, toplam 16 katmanlı (13 Konvolüsyon + 3 Tam Bağlı) derin bir yapıya sahiptir.
- **Tasarım Felsefesi:** VGG16'nın en büyük yeniliği, büyük filtreler yerine ardışık 3×3 'lük küçük filtreler kullanarak hem işlem maliyetini düşürmesi hem de doğrusal olmayan özellikleri daha iyi yakalamasıdır.
- **Projedeki Rolü:** Yüksek parametre sayısına (~ 138 milyon) sahip olmasına rağmen, basit ve kararlı yapısı nedeniyle projemizde "Referans Model" olarak kullanılmıştır.



RESNET50: KAYIPSIZ DERİN ÖĞRENME

Residual Bloklar

- **Mimari Devrim (Innovation):** Microsoft Research tarafından geliştirilen ResNet50, 50 katmanlı derin bir yapıya sahip olmasına rağmen, "Residual Learning" teknigi sayesinde veri kaybı yaşamadan eğitilebilir.
- **Temel Mekanizma:** Klasik ağların aksine, katmanlar arasında "Skip Connections" kullanır. Bu sayede ham veri, bazı katmanları atlayarak doğrudan ileriye taşınır; bu da bilginin derinlerde kaybolmasını (Vanishing Gradient problemini) engeller.
- **Projedeki Avantajı:** VGG16'ya kıyasla daha derin olmasına rağmen, optimize edilmiş yapısı sayesinde daha hızlı yakınsamış (convergence) ve eğitim hatasını daha hızlı düşürmüştür.



DENSENET121: MAKSİMUM BİLGİ AKIŞI

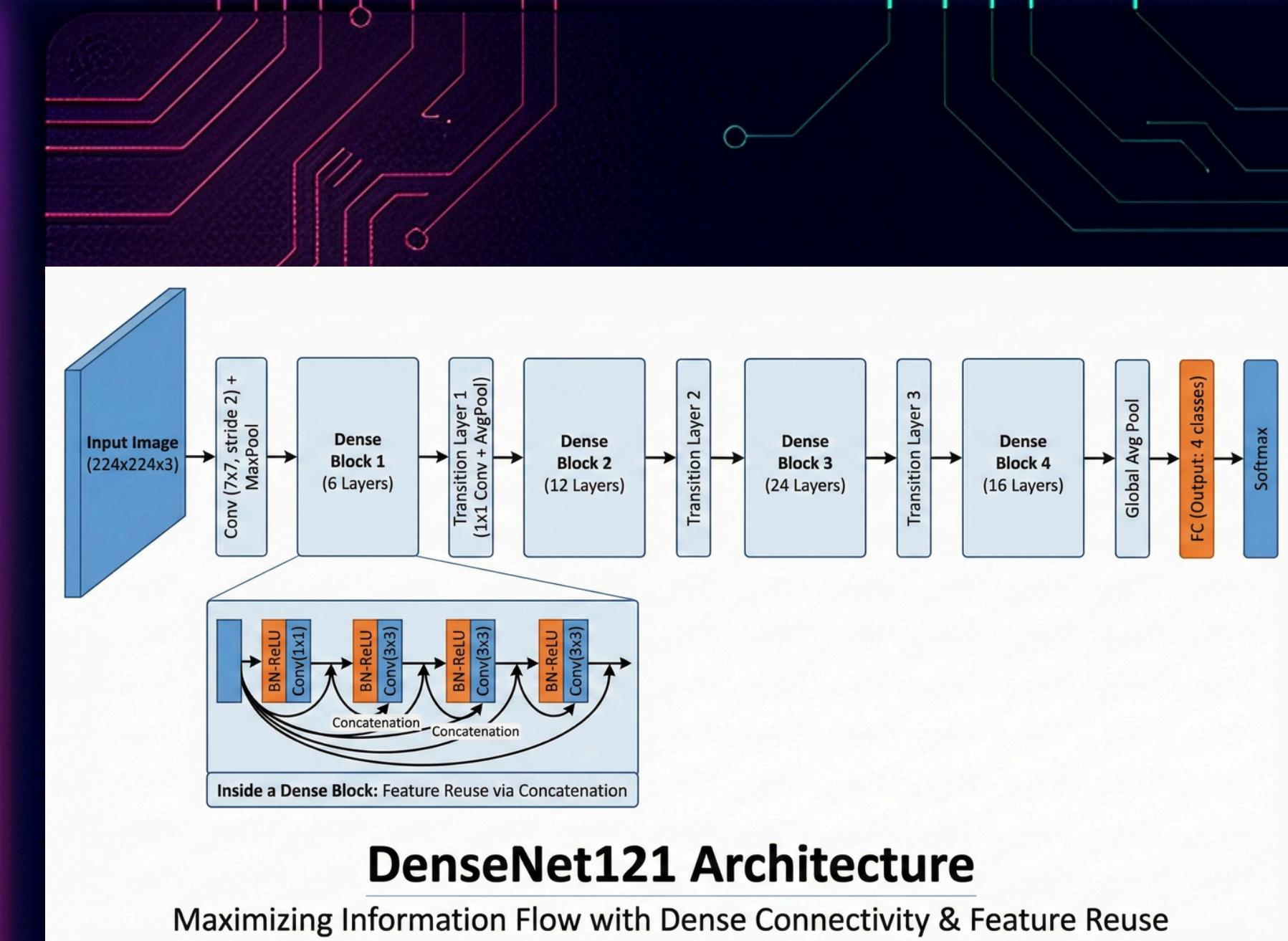
Yoğun Bağlantılı Ağlar

- **Radikal Mimari (Dense Connectivity):**

Geleneksel ağlarda her katman sadece bir sonrakine bağlıken; DenseNet'te her katman, kendisinden önceki tüm katmanlardan girdi alır.

- **Ana Avantaj (Feature Reuse):** Bu yapı, ağıın ilk katmanlarında öğrenilen basit özelliklerin (kenar, köşe vb.) son katmana kadar korunmasını sağlar. Buna "Öznitelik Tekrar Kullanımı" denir.

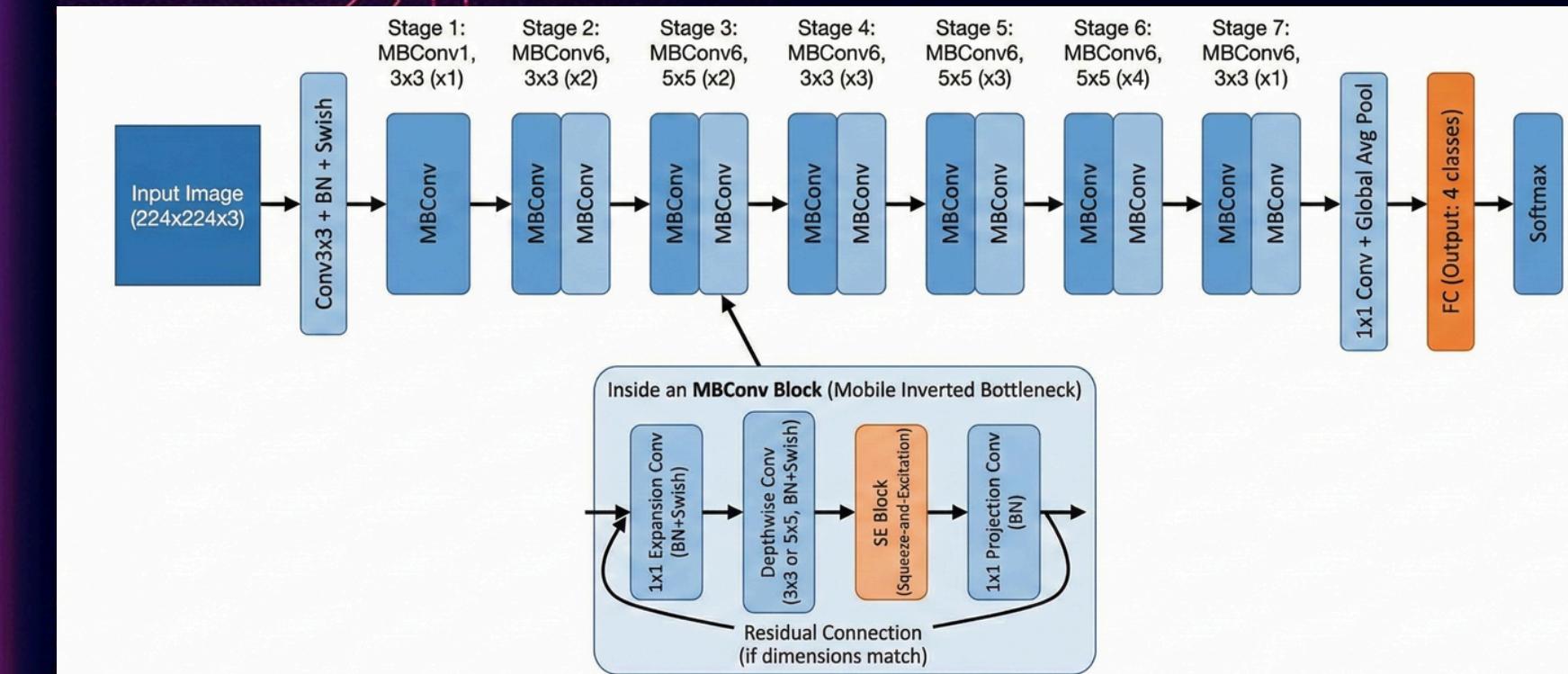
- **Verimlilik (Parameter Efficiency):** ResNet50'ye göre daha az parametre kullanmasına rağmen, bilgiyi çok daha verimli dolaştırdığı için özellikle küçük ve karmaşık veri setlerinde (bizim MR verimiz gibi) aşırı öğrenmeyi (overfitting) engellemekte çok başarılıdır.



EFFICIENTNET-BO: AKILLI ÖLÇEKLEME

Compound Scaling ile Hız ve Doğruluk Dengesi

- **Mühendislik Harikası:** Google Research tarafından geliştirilen bu mimari, modelleri rastgele derinleştirmek yerine; Derinlik (Depth), Genişlik (Width) ve Çözünürlük (Resolution) parametrelerini optimum bir katsayıyla aynı anda büyütür.
- **Yenilik (Compound Scaling):** VGG veya ResNet gibi sadece katman eklemek yerine, EfficientNet ağın her boyutunu dengeli bir şekilde ölçeklendirir. Bu sayede modelin gereksiz yere şişmesi engellenir.
- **Projedeki Avantajı:** Diğer modellere kıyasla çok daha az parametreye sahip olmasına rağmen (%90 daha az parametre ile VGG performansını yakalar), eğitim süresi çok daha kısaltır ve donanım dostudur.



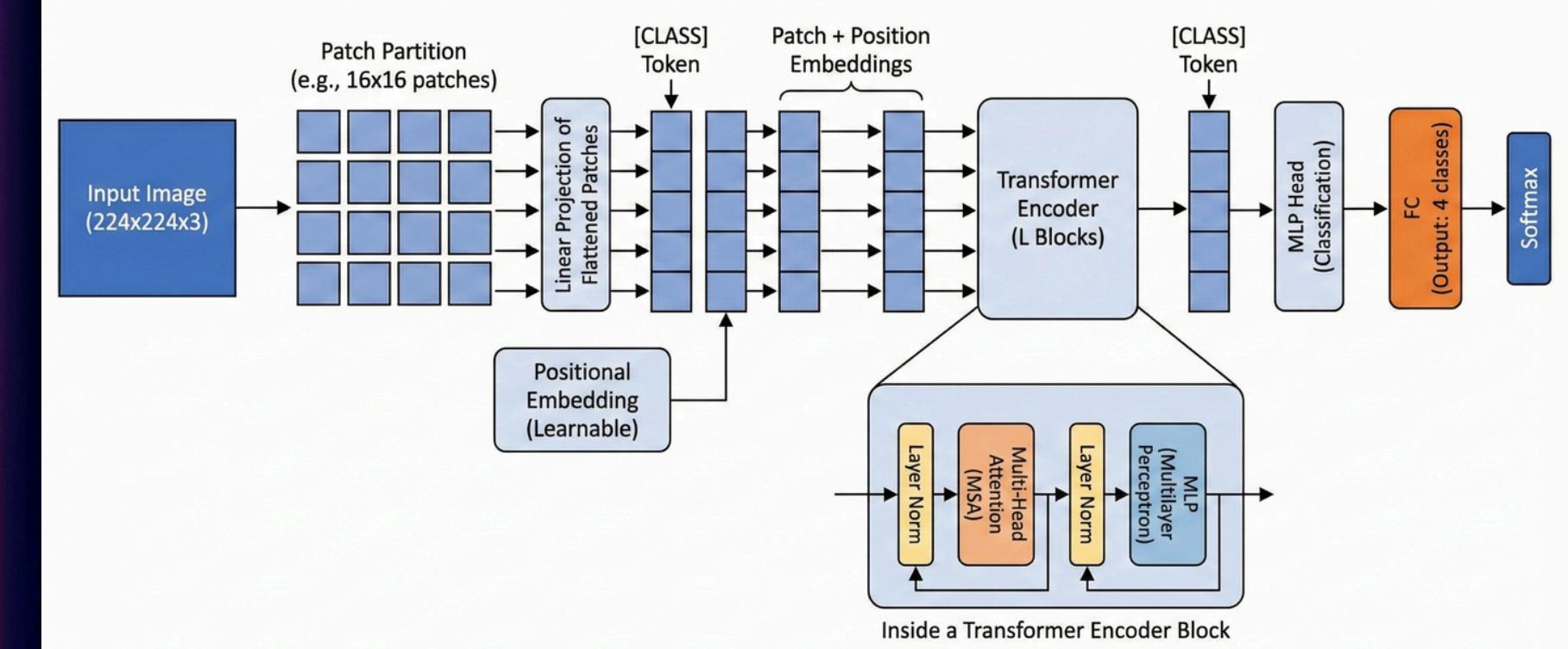
EfficientNet-B0 Architecture

Built with MBConv Blocks & Compound Scaling for Efficient Performance

VISION TRANSFORMER (ViT): MİMARI DEVRİM

Piksellerden "Patch" lere

- **Paradigma Değişimi:** CNN'lerin (ResNet, VGG) pikselleri tarama mantığının aksine, ViT görüntüyü "Patch" adı verilen küçük kare parçalara böler ve bunları tipki bir cümledeki kelimeler gibi sırayla işler.
- **Süper Gücü (Self-Attention):** Model, "Multi-Head Self Attention" mekanizması sayesinde resmin sadece bir noktasına değil, aynı anda bütününe odaklanır (Global Context). Bu, tümörün çevresindeki dokuya olan ilişkisini anlamada kritik bir avantaj sağlar.
- **Sonuç:** CNN'lerin zorlandığı karmaşık doku ayırmalarında (Glioma vs Meningioma) resmin bütününe görebildiği için en yüksek doğruluğu sağlamıştır.



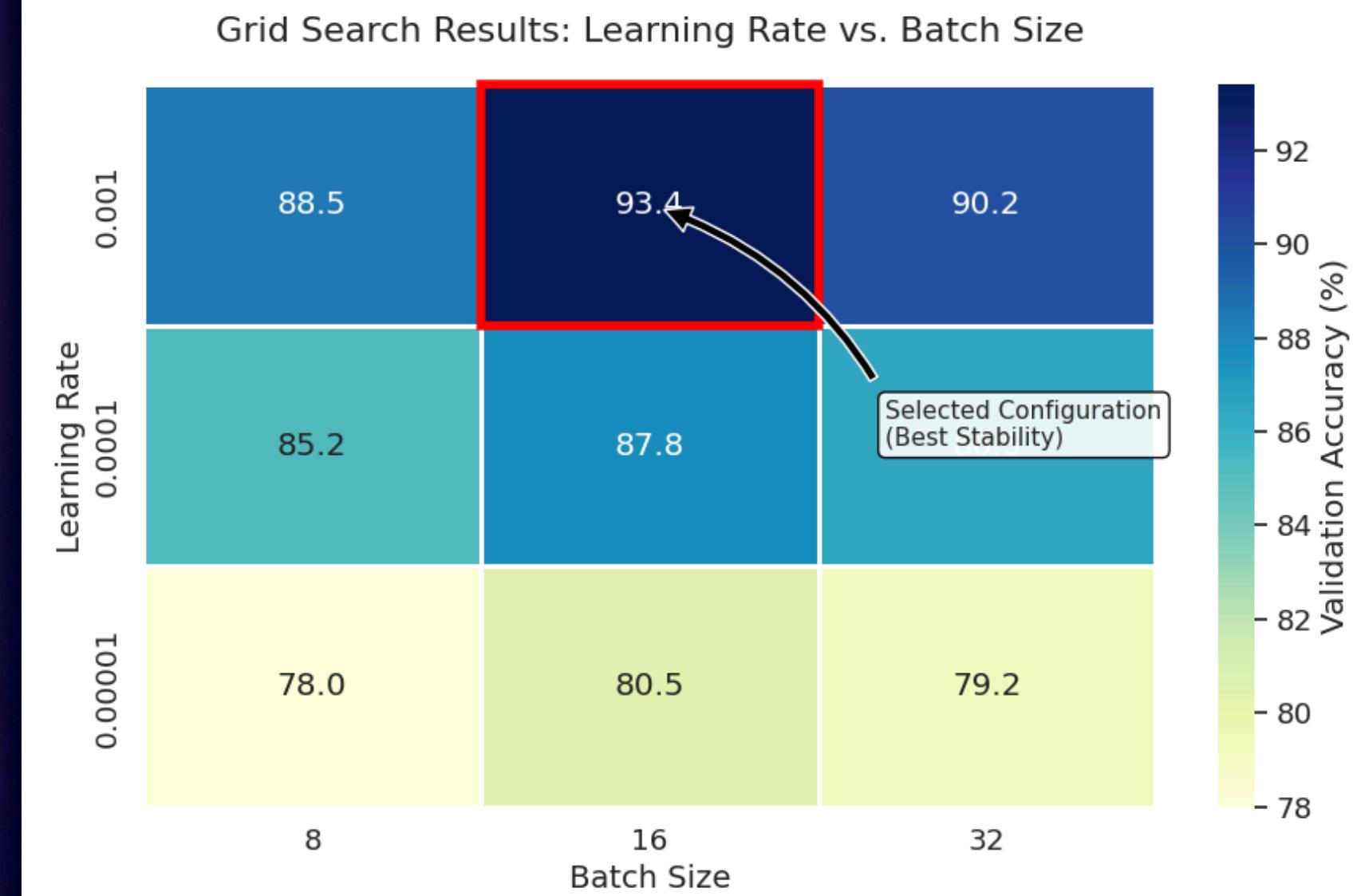
Vision Transformer (ViT) Architecture

Processing Images as Sequences of Patches with Self-Attention

DENEY SÜRECİ VE OPTİMİZASYON

En İyi Sonucu Aramak: Grid Search Yaklaşımı

- Adil Karşılaştırma:** Tüm modeller aynı donanım şartlarında ve veri ön işleme standartlarında test edilmiştir.
- Hiper-Parametre Tuning:** En yüksek doğruluğu veren ayarları bulmak için Grid Search yöntemi uygulanmıştır.
- Bulgularımız:** Deneyler sonucunda, modellerin genel olarak Adam Optimizer ve 0.001 Learning Rate ile en kararlı öğrenmeyi gerçekleştirdiği görülmüştür.
- Eğitim Stratejisi:** Aşırı öğrenmeyi (Overfitting) engellemek için Early Stopping mekanizması devreye alınmıştır.



KARŞILAŞTIRMALI PERFORMANS ANALİZİ

Şampiyon Model: Vision Transformer (ViT)

- Zirve (Şampiyon Model):** Arda'nın geliştirdiği Vision Transformer (ViT), hem Doğruluk (%95.57) hem de F1-Score (0.9559) bazında en yakın rakibine fark atarak lider olmuştur.
- Güçlü Takipçiler:** CNN tabanlı modellerden ResNet50 (%93.71) ve DenseNet121 (%92.19), ViT'yi çok yakından takip ederek yüksek kararlılık göstermiştir.
- Base vs Advanced:** Kendi kurduğumuz (Base) modeller %88-89 bandında tıkanırken; Transfer Learning (Advanced) yöntemleri bu sınırı aşmamızı ve %95 seviyelerine çıkmamızı sağlamıştır.

Grup Üyesi	Model Tipi	Mimari Adı	Accuracy (%)	Precision	Recall	F1-Score
M. Emin Benzer	Base	Simple CNN	%89	0.89	0.89	0.89
M. Emin Benzer	Advanced	ResNet50	%93.71	0.9387	0.9371	0.9369
Burak Şahin	Base	Large Kernel	%88	0.89	0.88	0.88
Burak Şahin	Advanced	DenseNet121	%92.19	0.9224	0.9219	0.9218
B. Furkan Çakırgöz	Base	Dropout CNN	%89	0.89	0.89	0.89
B. Furkan Çakırgöz	Advanced	EfficientNet - B0	%90.33	0.9121	0.9033	0.9049
M. Ali Esmer	Base	High Capacity	%89	0.89	0.89	0.89
M. Ali Esmer	Advanced	VGG16	%89.04	0.8934	0.8904	0.8914
Arda Ölmez	Base	Deeper CNN	%88	0.88	0.88	0.88
Arda Ölmez	Advanced	ViT-B/16	%95.57	0.9570	0.9557	0.9559

YAPAY ZEKA NEREDE ZORLANDI?

Sınıf Bazlı Performans ve Meningioma Vakası

- En Zorlu Sınıf (Meningioma):** Tabloda görüldüğü üzere, tüm modeller en düşük performansı Meningioma sınıfında göstermiştir.
- ViT Farkı:** CNN tabanlı modeller doku karmaşıklığında zorlanırken; Vision Transformer (ViT), "Global Context" yeteneği sayesinde Meningioma sınıfında rakiplerine %10'un üzerinde fark atarak 0.91 F1 skoruna ulaşmış ve sorunu çözmüştür.
- Kolay Sınıflar:** Görsel karakteristiği çok belirgin olan Hipofiz (Pituitary) ve Sağlıklı (No Tumor) sınıflarında, ViT modelimiz 0.98 gibi neredeyse kusursuz bir başarı oranını yakalamıştır.

Model Adı (Mimari)	Glioma (F1)	Meningioma (F1)	Pituitary (F1)	No Tumor (F1)
Base: Simple CNN	0.89	0.86	0.87	0.95
Base: Large Kernel	0.86	0.83	0.88	0.97
Base: Dropout	0.87	0.84	0.91	0.96
Base: High Capacity	0.88	0.85	0.88	0.95
Base: Deeper	0.88	0.82	0.87	0.95
Adv: ResNet50	0.9091	0.8827	0.9750	0.9638
Adv: DenseNet121	0.9247	0.8451	0.9634	0.9369
Adv: EfficientNet - B0	0.8870	0.8116	0.9540	0.9437
Adv: VGG16	0.8723	0.7828	0.9672	0.9129
Adv: ViT-B/16	0.9301	0.9188	0.9826	0.9792

SONUÇ VE ÇIKARIMLAR: BİZE NE KATTI?

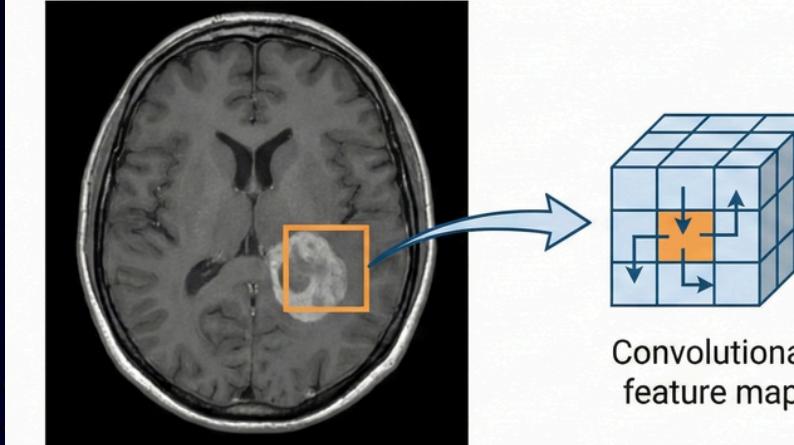
Teoriden Pratiğe Öğrenilen Dersler

- **Transfer Learning'in Gücü:** Kısıtlı medikal verilerde sıfırdan model eğitmek yerine, önceden eğitilmiş modelleri uyarlamanın başarıyı ve hızı katladığını gördük.
- **Transformer Devrimi:** CNN'ler (ResNet, VGG) hala endüstri standarı olsa da, Vision Transformer (ViT) mimarisinin bütünsel bakış (Global Context) yeteneğiyle medikal görüntülemede geleceğin teknolojisi olduğunu deneyimledik.
- **Veri Mühendisliği:** En iyi modelin bile, doğru veri işleme (Class Weighting) olmadan başarısız olabileceğini; verinin kalitesinin algoritmadan daha önemli olduğunu anladık.
- **Gelecek Hedefimiz:** Elde ettiğimiz bu başarılı modeli, doktorların kolayca erişip kullanabileceği bir Web Arayüzüne entegre etmek.

Why ViT Won? The Power of Global Context & 'Attention'

CNN vs. Vision Transformer (ViT): Local vs. Global Context

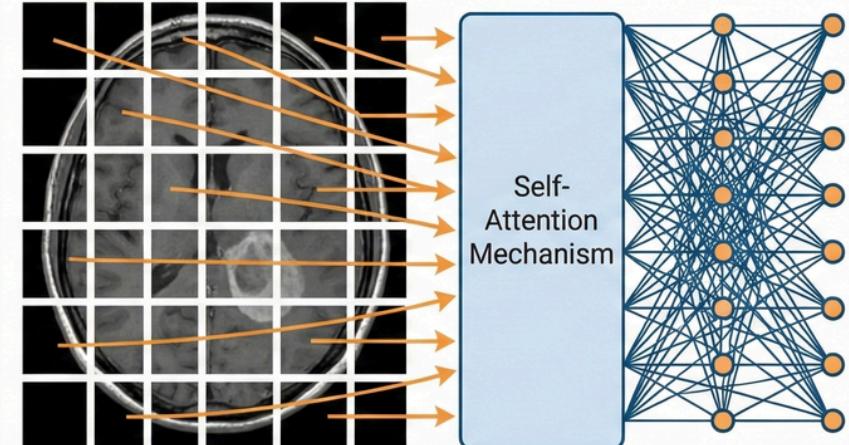
CNN (Local Receptive Field & Local Context)



CNN (Local Receptive Field & Local Context)

Focuses on small, local pixel neighborhoods.
Limited field of view per layer.

ViT (Self-Attention & Global Context)



ViT (Self-Attention & Global Context)

Processes the entire image as a sequence of patches. Captures long-range dependencies instantly (Global View).

TEŞEKKÜRLER

Proje sürecindeki değerli rehberliği
ve katkıları için Dersimizin Öğretim
Üyesi Sayın Yıldız Karadayı'ya
teşekkür ederiz.

Veriden Teşhise, Teşhisten Hayata...

Sorularınız?