






Hafif Bir Derin Öğrenme Modeli İle Bilgisayarlı Tomografi Görüntülerinden Beyin Kanaması Tespiti

Emine Betül Altun¹ , Sümeyye Engin¹ , Esmâ Başkaya¹ , Fatma Nur Şafak¹ ,
Saffet Vatansver¹ 

¹ Bursa Teknik Üniversitesi, Mühendislik ve Doğa Bilimleri Fakültesi, Mekatronik Mühendisliği Bölümü, Bursa

Geliş Tarihi / Received Date: 07.09.2024

Kabul Tarihi / Accepted Date: 26.11.2024

Öz

Beyin dokusu içine kan sızması durumu olarak ifade edilen beyin kanaması, acil tıbbi müdahale gerektiren nörolojik bir komplikasyondur. Bu sebeple, beyin kanamasında erken tanı, hastaların hayatta kalma şansını ve iyileşme sürecini önemli ölçüde etkiler. Beyin kanaması teşhisinde, radyologlarca yaygın olarak tercih edilen bilgisayarlı tomografi (BT) ve manyetik rezonans (MR) görüntüleri, derin öğrenme tabanlı yaklaşımlar ile analiz edilerek, beyin kanamasının varlığı ve kanamanın yeri hızlı ve etkili bir şekilde tespit edilebilir. Bu yöntemler, radyologların iş yükünü önemli ölçüde azaltabileceği gibi, kompleks vakalarda daha kesin teşhisler koyulmasına da yardımcı olabilir. Buna bağlı olarak, beyin kanaması kaynaklı ölümlerin veya bedensel işlev bozukluklarının önüne geçilebilir. Bu çalışmada, bilgisayarlı tomografi görüntüleri üzerinden beyin kanaması ve türünü yüksek doğrulukta tespit edebilen CNN tabanlı düşük boyutlu bir derin öğrenme modeli önerilmiştir. DenseNet121, MobileNet ve Inception V1 gibi popüler CNN modelleri ile yapılan karşılaştırmalı deneysel analizler, önerilen modelin, eğitim süresini önemli ölçüde kısalttığını ve daha başarılı bir performans sergilediğini göstermiştir.

Anahtar Kelimeler: beyin kanaması, bilgisayarlı tomografi, BT, derin öğrenme, CNN

Intracranial Hemorrhage Identification from Computed Tomography Images Using a Lightweight Deep Learning Model

Abstract

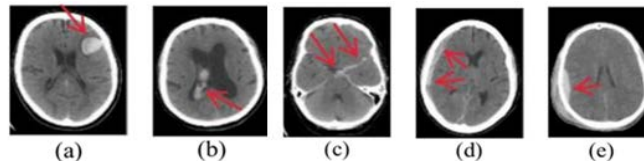
Intracranial hemorrhage, which is defined as blood leakage into the brain tissue, is a neurological complication that requires urgent medical intervention. Accordingly, early diagnosis of the hemorrhage significantly affects the chance of survival of patients and the recovery process. The presence of intracranial hemorrhage and its location can be identified quickly and effectively by deep learning-based approaches from computed tomography (CT) or magnetic resonance (MR) images, which radiologists commonly prefer to diagnose intracranial hematomas. These methods can significantly reduce the workload of radiologists and help make more accurate detections in complex cases. Accordingly, potential deaths or physical dysfunctions due to the hemorrhage can be prevented. This study proposes a CNN-based lightweight deep learning model that can accurately detect and classify intracranial hemorrhage from computed tomography images. Comparative experimental analyses with popular CNN models such as DenseNet121, MobileNet, and Inception V1 have shown that the proposed model significantly reduces training time and performs better.

Keywords: intracranial hemorrhage, computed tomography, CT, deep learning, CNN

Giriş

Kalıcı fiziksel hasarların veya bedensel işlev kayıplarının birinci nedeni, ölümlerin ise üçüncü nedeni olan beyin kanaması, acil müdahale gerektiren nörolojik bir tıbbi durumdur (Ahmed ve Prakasam, 2023; Champawat vd., 2023; Salehinejad vd., 2021). Tıp literatüründeki kullanımıyla intraserebral kanama olarak da bilinen beyin kanaması, bir beyin damarının yırtılması sonucu, kanın beyin dokusuna sızmasıyla meydana gelir. Beyin kanamasının en yaygın nedenleri kafa travması ve hipertansiyondur. Diğer nedenler arasında tümörler, venöz enfarktüs, serebral anevrizma, vasküler anormallikler ve zararlı madde kullanımı yer almaktadır (Champawat vd., 2023; Sage ve Padura, 2020). Beyin kanaması genellikle şiddetli baş ağrısı, denge kaybı, nefes almakta zorluk, görme bozuklukları ve uzuv uyuşması gibi belirtiler ile ortaya çıkmaktadır (Champawat vd., 2023). Beyin kanaması, beyinde olduğu bölgeye göre vücudun farklı kısımlarında felç oluşturabilmektedir. Bu bağlamda, beyin kanaması ve türünün hızlı bir şekilde tespit edilip inme olan bölgenin belirlenmesi, hastanın hayatta kalması ve yaşam kalitesini sürdürebilmesi bakımından kritik önem taşımaktadır (Altıntaş, 2021). Şekil 1’de belirtildiği gibi, beyin kanamasının (intraserebral kanama) epidural, subdural, subaraknoid, intraventriküler ve intraparenkimal kanama başta olmak üzere farklı türleri mevcuttur (Radiological Society of North America [RSNA], 2019).

Beyin kanamasının varlığını, konumunu ve boyutunu belirleyebilmek için kullanılan manyetik rezonans görüntüleme (MRG) ve bilgisayarlı tomografi (BT) gibi görüntüleme tekniklerinden; yüksek hızı ve duyarlılığı, nedeniyle BT, nörologlar ve radyologlar tarafından sıklıkla tercih edilmektedir (Gençtürk vd., 2023; Yeo vd., 2021). BT’nin yaygın olarak kullanılmasının bir diğer nedeni de diğer görüntüleme cihazlarına kıyasla daha düşük maliyetli ve daha kolay erişilebilir olmasındandır (Polat ve Kartal, 2023; Yeo vd., 2021).



(a) Intraparenkimal, (b) Intraventriküler, (c) Subaraknoid, (d) Subdural, (e) Epidural (RSNA, 2019).

Şekil 1. Beyin Kanaması Türleri

Bir hastaya, beyin kanaması semptomlarının görülmeye başlamasından sonraki ilk 6 saat içinde uygulanabilecek acil müdahale, kanamanın kontrol altına alınabilmesi bakımından kritiktir (Cordinnier vd., 2018; Nilsson vd., 2000; Wang vd., 2021). Bu bağlamda, beyin kanamasının erken tanısı büyük önem taşımaktadır. Ancak, radyologların iş yükündeki artış, BT görüntüleri üzerindeki görüş farklılıkları ve deneyimli uzmanların eksikliği göz önüne alındığında, BT görüntülerinin incelenmesi, doğru teşhisin koyulması ve hastaların tedaviye alınması zaman alabilmektedir (Gençtürk vd., 2023; Yeo vd., 2021).

BT görüntülerinin yapay zekâ (Artificial Intelligence (AI)) temelli uygulamalar ve sistemler ile otomatik olarak işlenip değerlendirilmesi, beyin kanaması teşhisini, dolayısıyla da tedavi sürecini hızlandırabilmektedir (Sage ve Padura, 2020; Wang vd., 2021; Yeo vd., 2021). Ayrıca karmaşık BT görüntülerinin doğru yorumlanmasına yardımcı olarak radyologların işini kolaylaştırabilmekte ve yanlış tanıların oranını azaltabilmektedir. Evrimsel sinir ağı (Convolutional Neural Network-CNN) tabanlı modeller, beyin kanamalarının tespiti ve sınıflandırılmasında en yaygın olarak kullanılan derin öğrenme mimarilerindedir (Burduja vd., 2020; Desai vd., 2017; Gautam ve Raman, 2021; Gençtürk vd., 2023; Ker vd., 2019; Mushtaq vd., 2021; Phong vd., 2017; Rane ve Warhade, 2021; Yalçın ve Vural, 2022; Zhang vd., 2022).

Bu çalışmada, bilgisayarlı tomografi görüntüleri üzerinden beyin kanaması varlığını ve türünü yüksek doğrulukta tespit edebilen CNN tabanlı düşük boyutlu (lightweight) bir derin öğrenme modeli önerilmiştir. Model eğitimi ve performans değerlendirmesi için, Kuzey Amerika Radyoloji Derneği (RSNA) tarafından 2019’da düzenlenen beyin kanaması tespiti yarışmasında katılımcılara sağlanan BT

görüntüleri (RSNA, 2019) kullanılmıştır. RSNA veri seti, pencereleme (windowing) temelli bir veri ön işleme sürecinden (Wang vd., 2021) geçirilerek model eğitimi gerçekleştirilmiştir. Yapılan deneysel çalışmalar ile elde edilen validasyon ve test doğruluk sonuçları, önerilen modelin radyolog değerlendirmelerine yakın ölçüde sonuçlar üretebildiğini göstermiştir. Önerilen modelin performans değerlendirmesi için ayrıca, DenseNet121 (Huang vd., 2017), MobileNet (Howard vd., 2017), InceptionNet V1 (Szegedy vd., 2016) gibi popüler CNN modelleri ile karşılaştırmalı bir analiz yapılmıştır. Önerilen modelin, diğer modellere kıyasla hesaplama maliyetini önemli ölçüde azalttığı, buna karşın performansı koruduğu, hatta nispeten artırdığı görülmüştür.

Genel başlıklarıyla bu makalenin ikinci bölümünde, beyin kanaması tespiti ve sınıflandırılması amacıyla bu çalışma kapsamında kullanılan veri seti, veri ön işleme adımları, önerilen CNN mimarisi, model parameter ayarları ve test süreçleri ele alınmıştır. Üçüncü bölümde, model performansı analiz edilmiştir. Kullanılan mimari çeşitlerine, veri büyüklüğüne ve hiperparametre ayarlarına göre elde edilen bulgular karşılaştırmalı olarak analiz edilmiştir. Sonuç bölümünde ise, çalışma kapsamında elde edilen önemli bulgular özetlenmiştir.

Materyal ve Yöntem

Bu bölümde, beyin kanaması tespiti ve sınıflandırılmasında, çalışma kapsamında kullanılan veri seti ve yöntem detayları sunulmuştur.

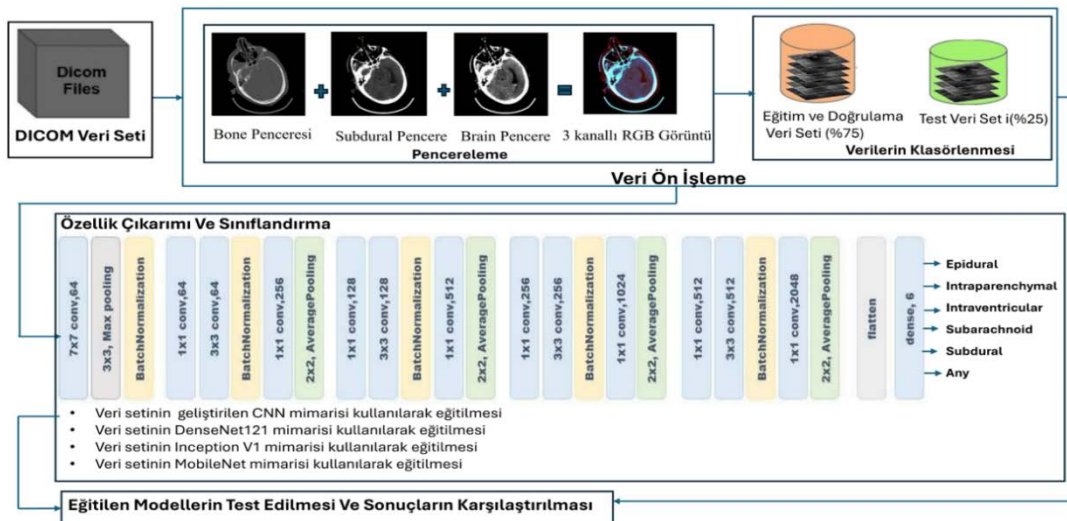
Materyal

Bu çalışmada, 2019 yılında Kuzey Amerika Radyoloji Derneği (Radiological Society of North America (RSNA)) tarafından düzenlenen intrakranyal kanama tespiti yarışmasında (Intracranial Hemorrhage Detection Challenge) sağlanan DICOM formatındaki BT görüntüleri veri seti kullanılmıştır (RSNA, 2019). Bu görüntüler, Thomas Jefferson Üniversite Hastanesi, Sao Paulo Federal Üniversitesi ve Stanford Üniversitesi olmak üzere üç kuruluştan toplanmış olup, Amerikan Nöroadyoloji Derneği (American Society of Neuroradiology (ASNR)) tarafından etiketlenmiştir.

RSNA veri seti (RSNA, 2019), tek bir kanama türü içeren ve hiç kanama içermeyen toplam 720733 BT görüntüsü içermektedir. Bu görüntülerin tamamına, Yöntem bölümünde belirtilen veri ön işleme adımları uygulanmıştır. Görüntülerin %50'si model eğitimi, %25'i validasyon ve %25'i test aşamasında kullanılmıştır. Kanama türlerinde ait detaylar Tablo 1'de verilmiştir.

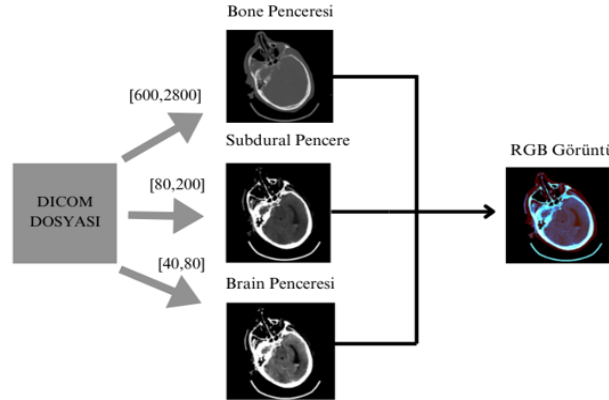
Yöntem

Bu çalışma kapsamında uygulanan yöntem, genel hatlarıyla Şekil 2'de blok diyagramı halinde gösterilmiş olup, detayları alt bölümler halinde incelenmiştir.



Şekil 2. Uygulanan Yöntem**Tablo 1.** Kanama Türlerine ait Görüntü Sayıları

Kanama Çeşidi	Görüntü Sayısı			
	Eğitim	Validasyon	Test	Toplam
Epidural	848	423	423	1694
İntraparenchymal	7832	3916	3916	15664
İntraventriküler	4940	2469	2469	9878
Subaraknoid	8213	4105	4105	16423
Subdural	16100	8050	8050	32200
Kanama olmayan	322440	161217	161217	644874

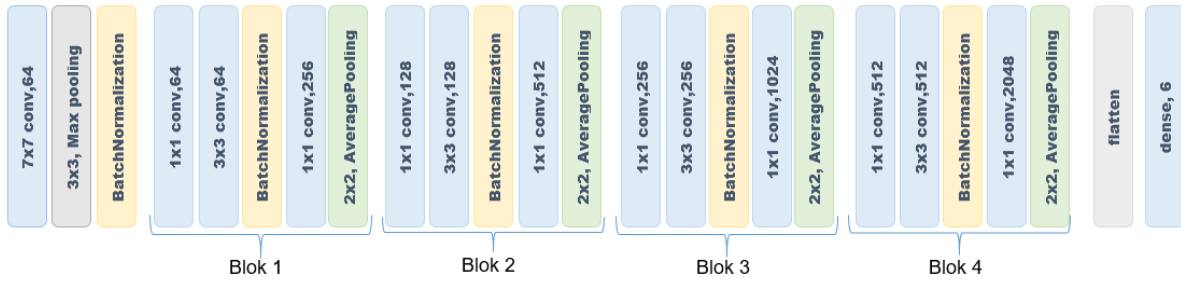
**Şekil 3.** DICOM Görüntüsünden Belirli Yoğunluk Aralıklarında RGB Görüntüsü Elde Etme**Veri Ön İşleme**

DICOM formatındaki BT görüntüleri geniş bir piksel aralığında (12 veya 16 bit) kaydedildiğinden, görüntüye ait ayrıntılar kolayca fark edilememektedir. Beyin kanaması gibi önemli detayların yakalanabilmesi için radyologlar, bu görüntüleri farklı yoğunluk aralıklarında görüntülemektedir. Pencereleme (windowing) adı verilen bu yöntemde pencere genişliği (window width (WW)) ve pencere seviyesi (window level (WL)) parametreleri kullanılmaktadır. Radyoloji uzmanları, beyin penceresi (brain window) için $WL = 40$, $WW = 80$; subdural pencere (subdural window) için $WL = 80$, $WW = 200$; kemik penceresi (bone window) için ise $WL = 600$, $WW = 2800$ olarak ayarlanmaktadır (Wang vd., 2021). Wang ve arkadaşları (2021), bu tekniğin, derin öğrenme modellerinde de etkili bir şekilde kullanılabileceğini göstermiştir.

Bu çalışmada, Wang ve arkadaşlarının (2021) yaklaşımından esinlenilerek, tüm DICOM görüntülerine pencereleme (windowing) ön işleme adımları uygulanmıştır. Bu sayede, BT görüntülerindeki kanama bölgelerinin netliği artırılarak detayların, CNN modeli tarafından daha iyi yakalanabilmesi, dolayısıyla da modelin daha iyi eğitilebilmesi hedeflenmiştir. Pencere işleminde, öncelikle, yukarıda belirtilen WW ve WL değerlerine göre minimum ve maksimum piksel değerleri hesaplanmıştır. Bu işlem, minimum piksel değeri için, pencere seviyesinden pencere genişliğinin yarısı çıkarılarak $(WL-WW/2)$; maksimum piksel değeri için, pencere seviyesine pencere genişliğinin yarısı eklenerek $(WL+WW/2)$ sağlanmıştır (Wang vd., 2021). Ardından, minimum piksel değerinin altındaki tüm piksel değerleri minimum piksel değeri ile; maksimum piksel değerinin üzerindeki tüm piksel değerleri ise maksimum piksel değeri ile değiştirilmiştir. Ardından, piksel değerleri normalize edilerek sekiz bitlik gri tonlamalı görüntü elde edilmiştir. Bu işlemler, her üç pencere tipi (beyin, subdural, kemik) için ayrı ayrı uygulanmıştır. Her bir DICOM dosyası için gri tonlamalı olarak elde edilen üç ayrı görüntü, üç kanallı bir RGB görüntüsünün birer kanalını temsil edecek şekilde birleştirilmiştir. Her bir DICOM dosyası için oluşturulan $512 \times 512 \times 3$ boyutundaki yeni görüntüler, CNN modeline giriş olarak verilebilmesi için $224 \times 224 \times 3$ boyutuna dönüştürülmüştür. Bu şekilde, CNN modelinin eğitimi için gerekli ön işlemler tamamlanmış olur. Şekil 3'te RSNA veri setine (RSNA, 2019) ait DICOM formatındaki örnek bir BT görüntüsünün, gri tonlamalı pencere görüntüleri ve birleştirilmiş RGB görüntüsü verilmiştir.

Önerilen Model

Her bir DICOM görüntüsüne veri ön işleme adımlarının uygulanması sonucu elde edilen $224 \times 224 \times 3$ boyutlarındaki yeni görüntüler, CNN modeline giriş olarak verilir. Bu çalışma kapsamında önerilen CNN modeli 4 GB hafızaya sahip GPU'larda dahi çalışabilecek düşük boyutlu (lightweight) bir mimariye sahiptir. Önerilen CNN modeline ait katman detayları Şekil 4'te görülmektedir. Önerilen model, 7×7 boyutunda 64 filtrelili evrişim katmanı, 3×3 boyutunda maksimum havuzlama ve normalizasyon katmanı ile başlamaktadır. Bu katmanların ardından birbirini izleyen dört blok gelmektedir. Her bir blok, birbirini izleyen iki evrişim katmanı, bir normalizasyon katmanı, bir evrişim katmanı ve bir ortalama havuzlama katmanından oluşur. Blok 1; 64 filtrelili 1×1 ve 3×3 evrişim katmanları, 256 filtrelili 1×1 evrişim katmanı içermektedir. Blok 2'nin evrişim katmanları 128 filtrelili 1×1 ve 3×3 , 512 filtrelili 1×1 filtre boyutludur. Blok 3; 256 filtrelili 1×1 ve 3×3 evrişim katmanları, 1024 filtrelili 1×1 evrişim katmanı içermektedir. Blok 4'ün evrişim katmanları ise 512 filtrelili 1×1 ve 3×3 , 2048 filtrelili 1×1 filtre boyutludur. Tüm evrişim katmanlarında ReLU aktivasyon fonksiyonu kullanılmaktadır. Son bloğun ardından flatten katmanı ile özellik haritası tek boyutlu bir vektöre dönüştürülmüştür. Altı sınıf çıktılı Softmax aktivasyon fonksiyonlu yoğun katman ile model sonlandırılmıştır.



Şekil 4. Önerilen CNN Mimarisi

Önerilen model ile, karmaşık özelliklerin öğrenilmesinde, işlem yükünün düşük tutulması hedeflenmiştir. Bu doğrultuda Şekil 4'te görüldüğü üzere her bloğa Batch Normalization (Küme normalizasyonu) ve Average pooling (Ortalama havuzlama) katmanları eklenmiştir. Batch Normalization ile, katman çıkışlarının kararlı bir dağılımda kalması ve hız kazanılması sağlanmıştır. Average pooling katmanı ile blok sonunda özellik haritasının boyutu küçültülmüş ve bu sayede parametre sayısı azaltılmıştır. Buna göre, önerilen modelin hafif, verimli bir mimari olduğu söylenebilir.

Hiperparametre Seçimleri ve Ayarlar

Hiperparametre seçimleri, derin öğrenme modellerinin performansı üzerinde önemli bir etkiye sahiptir. Önerilen CNN modelinin eğitiminde kullanılmak üzere seçilen hiperparametre değerleri ile CNN modelinin daha doğru öğrenmesi, etkili ve kararlı bir sınıflandırma yapabilmesi hedeflenmiştir.

Öğrenme hızı (learning rate (α)), empirik olarak ve ayrıca Wang ve arkadaşlarının (2021) çalışmasıyla da uyumlu olarak $1e-5$ seçilmiştir. Bu düşük öğrenme hızı, modelin öğrenme süresini kısmen uzatsa da, eğitim sürecinde daha küçük adımlarla ilerlenerek daha dengeli ve daha doğru öğrenme sağlanmış, dolayısıyla da kayıp fonksiyonunun global minimum değerinin kaçırılması engellenmiş olur. Beyin kanaması tespiti gibi hassas bir görevde, modelin aşırı öğrenme (overfitting) veya yetersiz öğrenme (underfitting) risklerini minimize edebilmesi için dikkatli ve yavaş bir öğrenme süreci gerekli görülmüştür. Momentum parametresi (β), empirik olarak 0.95 seçilmiştir. Literatürde, momentum değerinin 0.9 ve üzerinde seçilmesinin, model performansını genel olarak artırdığı görülmüştür (Jelassi ve Li, 2022). Yığın boyutu (batch size) olarak, öncelikle, mevcut bilgisayar donanımlarının önerilen modele izin verdiği en yüksek değer olan 32 seçilmiştir. Ardından, önerilen modeli diğer modellerle tarafsız bir şekilde karşılaştırmak üzere, her modelin ortak çalışabildiği 8 değeri seçilmiştir.

Optimizasyon yöntemi olarak α ve β değerlerini optimize ederek modeli eğittiği için Adam optimizasyon yöntemi tercih edilmiştir. Model sonunda aktivasyon fonksiyonu olarak, çok sınıflı sınıflandırma problemlerinde yaygın bir şekilde kullanılan Softmax fonksiyonu tercih edilmiştir.

Kayıp fonksiyonu olarak, modelin yüksek doğrulukla çalışabilmesinde, tahmin edilen sınıf olasılıklarının gerçek sınıf etiketleriyle uyumunun değerlendirilebilmesinde, dolayısıyla da model performansının optimize edebilmesinde oldukça etkili olan cross-entropy loss kullanılmıştır.

Model eğitiminde yetersiz veya aşırı öğrenmenin önüne geçmek için epok sayısının doğru seçilmesi kritik öneme sahiptir. Epok sayısının düşük olması, modelin yetersiz öğrenmesine yol açabilirken; fazla seçilmesi durumunda aşırı öğrenme (overfitting) sorununa neden olabilir ve modelin performansını düşürebilir. Ayrıca, aşırı öğrenme sorunu, belirlenen epok sayısına ulaşıldığında zaman kaybına yol açabilir. Bu nedenle, doğru epok sayısını belirlemek amacıyla erken durdurma (early stopping) yöntemi uygulanmıştır. Başlangıçta, epok sayısı 30 olarak set edilmiş, ancak erken durdurma yöntemiyle bu sayı optimize edilmiştir. Erken durdurma, validasyon kaybı fonksiyonunu izleyerek gerçekleştirilmiştir. Validasyon kaybının 3 ardışık epok boyunca artması durumunda, model eğitimi sonlandırılmıştır. Erken durdurma fonksiyonu, ayrıca, modelin en iyi performansını sağlamak adına en iyi parametreleri geri yükleyecek şekilde ayarlanmıştır.

Değerlendirme Metrikleri

Model başarımını ölçmek için, değerlendirmeye katılan tüm veriler arasında ne kadarının doğru tahmin edildiğini oransal olarak gösteren doğruluk (accuracy) (Grandini vd., 2020), ana performans metriği olarak kullanılmıştır. Doğruluk değeri, şu şekilde hesaplanır:

$$\text{Doğruluk} = \frac{TP + TN}{TP + TN + FN + FP} \quad (1)$$

Model performans değerlendirmesi için, ayrıca, pozitif olarak tahmin edilenlerin gerçekte ne kadarının pozitif olduğunu oransal olarak ifade eden kesinlik (precision) (Grandini vd., 2020); doğru olarak tahmin edilen pozitiflerin tüm gerçek pozitiflere oranını ifade eden duyarlılık (recall) (Grandini vd., 2020); kesinlik ile duyarlılığın dengeli bir ifadesi olan (harmonik ortalaması) F1 puanı (F1-score) (Grandini vd., 2020) metriklerinden yararlanılmıştır. Ele alınan problemin çoklu sınıf problemi olması ve her bir sınıfa ait eğitim verisi sayısı dengesizliğinin dikkate alınabilmesi amacıyla, her bir sınıf için elde edilen kesinlik, duyarlılık ve F1 puanı metrik değerlerinin ağırlıklı ortalaması tercih edilmiştir.

Belirli bir sınıfa ait kesinlik ve tüm sınıflara ait ağırlıklı kesinlik, sırasıyla Denklem (2) ve Denklem (3) yardımıyla hesaplanır:

$$\text{Kesinlik}_i = \frac{TP_i}{(TP_i + FP_i)} \quad (2)$$

$$\text{Ağırlıklı Kesinlik} = \sum_{i=1}^L \frac{N_i \times \text{Kesinlik}_i}{M} \quad (3)$$

Burada i , sınıf numarasını; L sınıf sayısını; N_i , i . sınıftaki veri sayısını; M , tüm sınıflardaki toplam veri sayısını temsil eder.

Belirli bir sınıfa ait duyarlılık ve tüm sınıflara ait ağırlıklı duyarlılık, sırasıyla Denklem (4) ve Denklem (5); belirli bir sınıfa ait F1 puanı ve tüm sınıflara ait ağırlıklı F1 puanı, sırasıyla Denklem (6) ve Denklem (7) yardımıyla hesaplanır.

$$\text{Duyarlılık}_i = \frac{TP_i}{(TP_i + FN_i)} \quad (4)$$

$$\text{Ağırlıklı Duyarluluk} = \sum_{i=1}^L \frac{N_i \times \text{Duyarluluk}_i}{M} \quad (5)$$

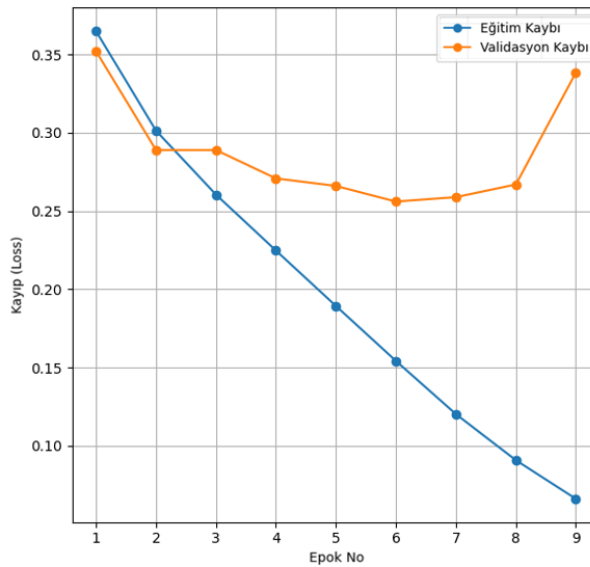
$$F1 \text{ Puan}_i = 2 \times \frac{\text{Kesinlik}_i \times \text{Duyarluluk}_i}{\text{Kesinlik}_i + \text{Duyarluluk}_i} \quad (6)$$

$$\text{Ağırlıklı F1 Puanı} = \sum_{i=1}^L \frac{N_i \times F1 \text{ Puan}_i}{M} \quad (7)$$

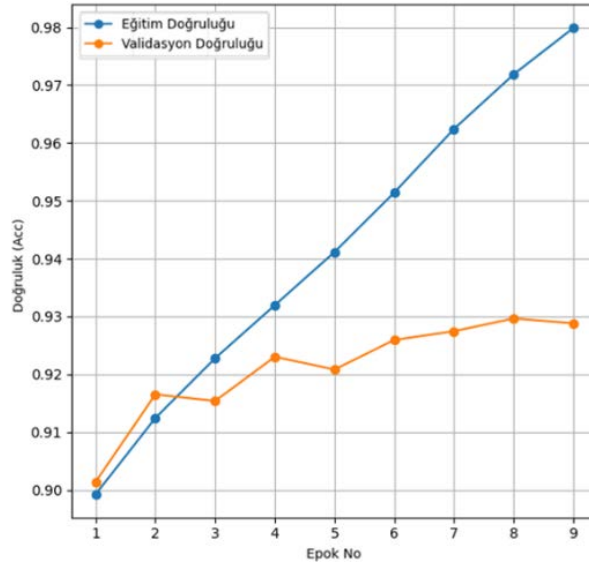
Bulgular

Önerilen CNN modeli, Yöntem bölümünde belirtilen hiperparametre ayarları ile en fazla 30 epok boyunca eğitilecek şekilde model eğitimine başlanmıştır. Ancak, Şekil 5'te görüleceği üzere, validasyon kaybı, 6. epoktan itibaren artmaya başlamış ve artışların artmaya devam etmesi sonucu eğitim, erken durdurma (early stopping) yöntemince 9. epokta otomatik olarak durdurulmuştur. Kayıp fonksiyonunun minimum değeri olan 6. epoktan sonra modelin aşırı öğrenmeye (overfitting) maruz kaldığı, bu sebeple de 6. epok en iyi eğitim sonuçlarının elde edildiği nokta olarak kabul edilmiştir. Buna göre model, o ana ait ağırlıklara göre kaydedilmiştir. Şekil 6'da epok numarasına göre elde edilen doğruluk değerleri görülmektedir. En yüksek doğruluk değeri 8. epokta görülmesine karşın, 6. epoka ait değerler dikkate alınmıştır. Tablo 2'de bu ana ait eğitim, validasyon ve test doğruluk sonuçları verilmiştir.

Veri seti ile modelin eğitilmesi ortalama 4 saat sürmüştür. 6. epokta eğitim doğruluk değeri %95,14 ve kayıp değeri 0.1542 olarak kaydedilmiştir. Bu sonuçlar, modelin eğitim verisi üzerinde oldukça yüksek bir doğruluk oranına sahip olduğunu ve kayıp değerinin düşük olduğunu göstermektedir. Eğitim aşamasındaki düşük kayıp değeri, modelin eğitim verisi üzerinde hataları minimize ettiğini ve öğrenme sürecinin verimli olduğunu göstermektedir. Modelin validasyon aşamasındaki performansı, modelin genelleme yeteneğini ve aşırı öğrenme (overfitting) sorununu değerlendirmek için kritik öneme sahiptir. Validasyon aşamasında elde edilen doğruluk değeri %92.59 ve kayıp değeri 0.2561 olarak kaydedilmiştir. Modelin gerçek dünya verilerine karşı performansını ölçmek için ayrılan test verisi üzerinde elde edilen sonuçlar, modelin genel geçerliliğini değerlendirmede önemli bir role sahiptir. Test aşamasında elde edilen doğruluk değeri %92.63 ve kayıp değeri 0.2547 olarak olarak kaydedilmiştir.



Şekil 5. Önerilen CNN Mimarisi Eğitim ve Validasyon Kayıp Fonksiyonu (Loss Function) Grafikleri



Şekil 6. Önerilen CNN Mimarisi Eğitim ve Validasyon Doğruluk Grafikleri

Tablo 2. Önerilen CNN mimarisi eğitim, validasyon ve test doğruluğu sonuçları

Eğitim	Validasyon	Test
0,9514	0,9259	0,9263

Validasyon ve test doğruluklarının birbirine yakın olması, modelin tutarlı bir performans sergilediğini ve eğitimin doğru yerde durdurulduğunu göstermektedir. Test kaybının validasyon kaybından daha düşük olması, modelin test verisi üzerinde beklenenden daha iyi performans gösterdiğine işaret edebilir. Bu durum, modelin veri kümesinin genel yapısını iyi öğrendiğini ve bilinmeyen veriler üzerinde de başarılı olduğunu göstermektedir.

Önerilen model, sırası ile 7.049 milyon, 3.224 milyon, 5.979 milyon parametreye sahip DenseNet121 (Huang vd., 2017), MobileNet (Howard vd., 2017), InceptionNet V1 (Szegedy vd., 2016) mimarileri ile doğruluk ve zaman açısından ayrıca karşılaştırılmıştır. Parametre sayısı, diğer modellere kıyasla daha yüksek olan Densenet121, tüm eğitim işlemlerinin gerçekleştirildiği 4 GB GPU'ya sahip dizüstü bilgisayarda, 8'den daha büyük bir yığın boyutu (batch size) seçildiğinde eğitilememiştir. Bu sebeple, şeffaf bir değerlendirme için, önerilen model de dâhil olmak üzere tüm modeller yığın boyutu 8 seçilerek eğitilmiştir. Diğer hiperparametre ayarları, tüm modeller için, Yöntem bölümünde verilenlerle verilenlerle aynı seçilmiştir. Modellere ait eğitim, validasyon ve test doğrulukları Tablo 3'te verilmiştir. Buna göre, önerilen model, 0.9284 validasyon doğruluğu ve 0.9281 test doğruluğu ile diğer modellerden daha başarılı bir performans göstermiştir. Önerilen modelin, diğer modellerle eğitim süreleri karşılaştırması Tablo 4'te verilmiştir. Buna göre, önerilen model, diğer modellere göre çok daha hızlı çalışmaktadır. Önerilen modelin hızlı çalışması, her işlem bloğunda BatchNormalization ve Average Pooling katmanlarına yer verilmesinden kaynaklandığı değerlendirilmektedir.

Tablo 3. Farklı CNN Modelleri ile Eğitim, Validasyon ve Test Doğruluğu Sonuçlarının Karşılaştırılması

Yöntem	Parametre Sayısı	Eğitim	Validasyon	Test
Önerilen	5.434M	0,9544	0,9284	0,9281
DenseNet 121	7.049M	0,9265	0,8828	0,8816
Mobile Net	3.224M	0,9117	0,9095	0,9087
Inception V1	5.979M	0,9359	0,9224	0,9221

Tablo 4. RSNA Veri Seti ile Eğitilen Farklı CNN Modellerinin Eğitim Süreleri Karşılaştırması (Ortalama)

Önerilen	DenseNet121	MobileNet	Inception V1
266 dakika	1004 dakika	553 dakika	422 dakika

Daha tarafsız, dolayısıyla da daha güvenilir bir performans karşılaştırması için, her bir model, test veri kümesi üzerinde, kesinlik (precision), duyarlılık (precision) ve F1 puanı (F1-score) metriklerinin ağırlıklı ortalama (weighted-average) değerleri ile ayrıca değerlendirilmiştir. Tablo 5'te görüleceği üzere, önerilen model ile, tüm metriklerde, diğer modellere kıyasla daha başarılı sonuçlar elde edilmiştir. Her bir modele ait karmaşıklık matrisi (confusion matrix), Tablo 6, 7, 8, 9'da verilmiştir. Karmaşıklık matrislerinde, köşegenler, ait oldukları sınıfa ait (ilgili satır veya sütuna karşılık gelen) TP değerini; herhangi bir köşegene ait satırda köşegen dışındaki diğer hücrelere ait değerlerin toplamı, ilgili sınıfa ait FN değerini; herhangi bir köşegene ait sütunda köşegen dışındaki diğer hücrelere ait değerlerin toplamı, ilgili sınıfa ait FP değerini; herhangi bir köşegene ait tüm satır ve sütun hücreleri dışında kalan diğer tüm hücrelere ait değerlerin toplamı ise, ilgili sınıfa ait TN değerini temsil etmektedir. Tablo 6, 7, 8 ve 9'a göre, tüm modellerin, kanama olmayan sınıfta daha başarılı performans gösterdiği söylenebilir. Bu durum, kanama olmayan sınıfa ait veri sayısının diğer beş sınıfa göre çok daha fazla olduğundan (Tablo 1'e bakılabilir), modellerin bu sınıfa ait görüntüleri çok daha iyi öğrenebilmesi ile açıklanabilir.

Tablo 5. Farklı CNN Modelleri ile Test Veri Kümesi Üzerinde Elde Edilen Kesinlik, Duyarlılık ve F1 Puanı Karşılaştırması (Her Bir Metriğe Ait Ağırlıklı Ortalama Değerler)

Yöntem	Kesinlik	Duyarlılık	F1 puanı
Önerilen	0.9268	0.9278	0.9246
DenseNet 121	0.9185	0.9132	0.9138
Mobile Net	0.8850	0.9088	0.8908
Inception V1	0.9210	0.9241	0.9190

Tablo 6. Önerilen Modele ait Karmaşıklık Matrisi

		Tahmin Edilen Sınıf (Predicted Class)					
		S1	S2	S3	S4	S5	S6
Gerçek Sınıf (Actual Class)	S1	156974	20	207	74	689	3253
	S2	170	107	5	0	15	126
	S3	1271	9	2076	60	222	278
	S4	989	1	179	1053	102	145
	S5	2041	1	82	23	1531	427
	S6	2374	15	40	5	127	5489

Tablo 7. DenseNet121 Modeline ait Karmaşıklık Matrisi

		Tahmin Edilen Sınıf (Predicted Class)					
		S1	S2	S3	S4	S5	S6
Gerçek Sınıf (Actual Class)	S1	154326	2	332	95	1539	4923
	S2	162	0	16	1	15	229
	S3	966	3	2083	115	436	313
	S4	749	0	194	1127	272	127
	S5	1760	0	105	16	1650	574
	S6	2367	12	104	4	225	5338

Tablo 8. MobileNet Modeline ait Karmaşıklık Matrisi

		Tahmin Edilen Sınıf (Predicted Class)					
		S1	S2	S3	S4	S5	S6
Gerçek Sınıf (Actual Class)	S1	158798	9	420	190	159	1641
	S2	318	8	14	2	6	75
	S3	1923	2	1492	143	23	333
	S4	1485	0	206	582	31	165
	S5	3316	1	117	76	212	383
	S6	5151	4	176	28	53	2638

Tablo 9. InceptionV1 Modeline ait Karmaşıklık Matrisi

		Tahmin Edilen Sınıf (Predicted Class)					
		S1	S2	S3	S4	S5	S6
Gerçek Sınıf (Actual Class)	S1	156468	0	193	57	1118	3381
	S2	149	0	5	0	14	255
	S3	999	0	2018	48	415	436
	S4	883	0	283	883	235	185
	S5	1979	0	56	6	1545	519
	S6	2621	1	21	1	180	5226

Sonuç

Bu çalışmada, bilgisayarlı tomografi görüntülerinden beyin kanaması ve türlerinin tespitini başarılı bir şekilde yapabilen düşük boyutlu (lightweight) bir CNN modeli önerilmiştir. Model eğitimi ve performans değerlendirmesi için BT görüntülerinden oluşan RSNA beyin kanaması veri seti (RSNA, 2019) kullanılmıştır. RSNA veri setine ait tüm görüntüler, pencereleme (windowing) tabanlı ön işlemlerden (Wang vd., 2021) geçirilerek modele verilmiştir. Önerilen model, 0,9259 validasyon doğruluğu ve 0,9263 test doğruluğu ile radyolog değerlendirmelerine yakın sonuçlar üretmiştir. Validasyon ve eğitim doğruluklarının tutarlı olması eğitimin doğru noktada sonlandırıldığını göstermektedir. Önerilen modelin performansı, ayrıca, literatürdeki düşük ağırlıklı popüler CNN modelleri olan DenseNet121, MobileNet, Inception V1 mimarileri ile eğitim süreleri, doğruluk, kesinlik, duyarlılık ve F1 puanı açısından karşılaştırılarak değerlendirilmiştir. Deneysel bulgular, önerilen modelin eğitim süresinin diğer modellere kıyasla çok daha kısa olduğunu ve nispeten daha başarılı bir performans sergilediğini göstermiştir. Buna göre, önerilen modelin kayda değer ölçüde verimli bir mimari olduğu söylenebilir.

Destek ve Teşekkür

Bu çalışma 2209-A - Üniversite Öğrencileri Araştırma Projeleri Destekleme Programı kapsamında “Bilgisayarlı Tomografi Görüntüleri Kullanılarak Derin Öğrenme ile Beyin Kanaması Tespiti” isimli proje ile TÜBİTAK tarafından desteklenmiştir.

Yazar Katkısı

Emine Betül Altun, Sümeyye Engin, Esmâ Başkaya ve Fatma Nur Şafak; deneyleri yapmış ve makaleyi yazmıştır. *Saffet Vatansever*, araştırmanın ortaya çıkmasını sağlamış, deneylerin yapılmasında, makale yazımında, düzetmelerinde ve iyileştirmelerinde yardımcı olmuştur.

Etik

Bu makalenin yayınlanmasıyla ilgili herhangi bir etik sorun bulunmamaktadır.

Çıkar Çatışması

Yazarlar herhangi bir çıkar çatışması olmadığını belirtmektedir.

ORCID

Emine Betül Altun  <https://orcid.org/0009-0006-1408-3371>

Sümeyye Engin  <https://orcid.org/0009-0009-4534-2139>

Esmâ Başkaya  <https://orcid.org/0009-0000-6317-7952>

Fatma Nur Şafak  <https://orcid.org/0009-0003-3685-311X>

Saffet Vatansever  <https://orcid.org/0000-0002-4680-1263>

Kaynaklar

- Ahmed, S. N. ve Prakasam, P. (2023). A systematic review on intracranial aneurysm and hemorrhage detection using machine learning and deep learning techniques. *Progress in Biophysics and Molecular Biology*, 183, 1-16. <https://doi.org/10.1016/j.pbiomolbio.2023.07.001>
- Altıntaş, M. (2021). *Bilgisayarlı tomografi görüntülerinde inmenin farklı derin öğrenme modelleri ile sınıflandırılması* [Yayımlanmamış yüksek lisans tezi]. Necmettin Erbakan Üniversitesi.
- Burduja M., Ionescu R. T. ve Verga N. (2020). Accurate and efficient intracranial hemorrhage detection and subtype classification in 3D CT scans with convolutional and long short-term memory neural networks, *Sensors*, 20(19), 1-21. <https://doi.org/10.3390/s20195611>
- Champawat, Y. S., Shagun ve Prakash, C. (2023). Literature review for automatic detection and classification of intracranial brain hemorrhage using computed tomography scans. In H. Muthusamy, J. Botzheim ve R. Nayak (Ed.), *Lecture notes in electrical engineering: Vol. 1009. Robotics, control and computer vision* (s. 39-65). Springer. https://doi.org/10.1007/978-981-99-0236-1_4
- Cordonnier C., Demchuk A., Ziai W. ve Anderson C. S. (2018). Intracerebral haemorrhage: current approaches to acute management. *The Lancet*, 392(10154), 1257-1268. [https://doi.org/10.1016/S0140-6736\(18\)31878-6](https://doi.org/10.1016/S0140-6736(18)31878-6)
- Desai V., Flanders A. E. ve Lakhani P. (2017). Application of deep learning in neuroradiology: automated detection of basal ganglia hemorrhage using 2D-convolutional neural networks. *arXiv preprint arXiv*, 1-7. <https://doi.org/10.48550/arXiv.1710.03823>
- Gautam A. ve Raman B. (2021). Towards effective classification of brain hemorrhagic and ischemic stroke using CNN. *Biomedical Signal Processing Control*, 63, 1-13. <https://doi.org/10.1016/j.bspc.2020.102178>
- Gençtürk T. H., Gülağız F. K. ve Kaya İ. (2023). Derin öğrenme yöntemleri kullanılarak BT taramalarında beyin kanaması teşhisinin karşılaştırmalı bir analizi. *Zeki Sistemler Teori ve Uygulamaları Dergisi*, 6(1), 75–84. <https://doi.org/10.38016/jista.1215025>
- Grandini, M., Bagli, E. ve Visani, G. (2020). Metrics for multi-class classification: An overview. *arXiv preprint arXiv*, 1-17. <https://doi.org/10.48550/arXiv.2008.05756>
- Howard A. G., Zhu M., Chen B., Kalenichenko D., Wang W., Weyand T., Andreetto M. ve Adam H. (2017). MobileNets: Efficient convolutional neural networks for mobile vision applications. *arXiv preprint arXiv*, 1-9. <https://doi.org/10.48550/arXiv.1704.04861>
- Huang G., Liu Z., Maaten L. V. D. ve Weinberger K. Q. (2017, Temmuz, 21-26). *Densely connected convolutional networks* [Sözlü sunum]. IEEE Conference on Computer Vision and Pattern Recognition, Honolulu, HI, USA. <https://doi.org/10.1109/CVPR.2017.243>
- Jelassi, S. ve Li, Y. (2022, Temmuz, 17-23). *Towards understanding how momentum improves generalization in deep learning* [Sözlü sunum]. International Conference on Machine Learning, Maryland, USA. <https://proceedings.mlr.press/v162/jelassi22a/jelassi22a.pdf>
- Ker J., Singh S. P., Bai Y., Rao J., Lim T. ve Wang L. (2019). Image thresholding improves 3-dimensional convolutional neural network diagnosis of different acute brain hemorrhages on computed tomography scans. *Sensors*, 19(9), 1-12. <https://doi.org/10.3390/s19092167>
- Mushtaq M. F., Shahroz M., Aseere A. M., Shah H., Majeed R., Shehzad D. ve Samad A. (2021). BHCNet: neural network-based brain hemorrhage classification using head CT scan. *IEEE Access*, 9, 113901-113916. <https://doi.org/10.1109/ACCESS.2021.3102740>

- Nilsson O. G., Lindgren A., Stahl N., Brandt L. ve Saveland H. (2000). Incidence of intracerebral and subarachnoid haemorrhage in southern Sweden. *Journal of Neurology, Neurosurg & Psychiatry*, 69(5), 601-607. <https://doi.org/10.1136/jnnp.69.5.601>
- Phong T. D., Duong H., Nguyen H. T., Trong N. T., Nguyen V. H., Hoa T. V. ve Snasel V. (2017, Ocak, 13-16). *Brain hemorrhage diagnosis by using deep learning*. International Conference on Machine Learning and Soft Computing, New York, NY, USA. <https://doi.org/10.1145/3036290.3036326>
- Polat Ö. ve Kartal M. S. (2023). Derin öğrenme ile pencere ayarlı görüntüler kullanılarak beyin inme segmentasyon performansının geliştirilmesi. *Gümüşhane Üniversitesi Fen Bilimleri Dergisi*, 13(4), 1094-1109. <https://doi.org/10.17714/gumusfenbil.1319024>
- Radiological Society of North America (2019). RSNA intracranial hemorrhage detection challenge. <https://www.rsna.org/rsnai/ai-image-challenge/rsna-intracranial-hemorrhage-detection-challenge-2019> adresinden 2 Ağustos 2024 tarihinde alınmıştır.
- Rane H. ve Warhade K. (2021, Mart, 5-7). *A survey on deep learning for intracranial hemorrhage detection* [Sözlü sunum]. International Conference on Emerging Smart Computing and Informatics, Pune, India. <https://doi.org/10.1109/ESCI50559.2021.9397009>
- Sage A. ve Badura P. (2020). Intracranial hemorrhage detection in head CT using double-branch convolutional neural network, support vector machine, and random forest. *Applied Sciences*, 10(21), 1-12. <https://doi.org/10.3390/app10217577>
- Salehinejad H., Kitamura J., Ditkofsky N., Lin A., Bharatha A., Suthiphosuwana S., Lin H., Wilson J. R., Mamdani M. ve Colak E. (2021). A real-world demonstration of machine learning generalizability in the detection of intracranial hemorrhage on head computerized tomography. *Scientific Reports*, 11(1), 1-11. <https://doi.org/10.1038/s41598-021-95533-2>
- Szegedy C., Vanhoucke V., Ioffe S., Shlens J. ve Wojna Z. (2016, Haziran, 27-30). *Rethinking the inception architecture for computer vision* [Poster sunumu]. IEEE Conference on Computer Vision and Pattern Recognition, Las Vegas, NV, USA. <https://doi.org/10.1109/CVPR.2016.308>
- Wang X., Shen T., Yang S. ve Lan J. (2021). A deep learning algorithm for automatic detection and classification of acute intracranial hemorrhages in head CT scans. *Neuroimage Clinical*, 32, 1-10. <https://doi.org/10.1016/j.nicl.2021.102785>
- Yalçın S. ve Vural H. (2022). Brain stroke classification and segmentation using encoder-decoder based deep convolutional neural networks. *Computers in Biology and Medicine*, 149, 1-14. <https://doi.org/10.1016/j.compbiomed.2022.105941>
- Yeo M., Tahayori B., Kok H. K., Maingard J., Kutaiba N., Russell J., Thijs V., Jhamb R. V. Chandra A., Brooks M., Barras C. D. ve Asadi H. (2021). Review of deep learning algorithms for the automatic detection of intracranial hemorrhages on computed tomography head imaging. *Journal of NeuroInterventional Surgery*, 13, 369-378. <https://doi.org/10.1136/neurintsurg-2020-017099>
- Zhang M., Gu S. ve Shi Y. (2022). The use of deep learning methods in low-dose computed tomography image reconstruction: A systematic review. *Complex and Intelligent Systems*, 8, 5545-5561. <https://doi.org/10.1007/s40747-022-00724-7>