



## Early diagnosis of lung cancer using deep learning and uncertainty measures

Sema Üzülmöz<sup>1\*</sup>, Mehmet Akif Çiççi<sup>2</sup>

<sup>1</sup>Department of Computer Engineering, Faculty of Natural Science, İstanbul Aydın University, İstanbul, 34150, Türkiye

<sup>2</sup>Department of Computer Engineering, Faculty of Natural Science, Bandirma Onyedil Eylül University, Balıkesir, 10200, Türkiye

### Highlights:

- A novel model develop
- Uncertainty measure
- Baresnet with UQ
- Markov approximation

### Keywords:

- CT
- Lung cancer
- UQ
- Deep learning

### Article Info:

Research Article

Received: 27.03.2022

Accepted: 26.02.2023

### DOI:

10.17341/gazimmfd.1094154

### Correspondence:

Author: Sema Üzülmöz

e-mail:

semauzulmez@aydin.edu.tr

phone: 90 506 138 8661

### Graphical/Tabular Abstract

In this study, a dataset of 663,549 lung computed tomography (CT) scans was analyzed using Deep Learning and Uncertainty Measures to determine the existence of lung cancer in patients and, if present, the benign or malignant nature of the cancer. The dataset was divided into three groups: 80% for training, 10% for testing, and 10% for validation (Figure A).

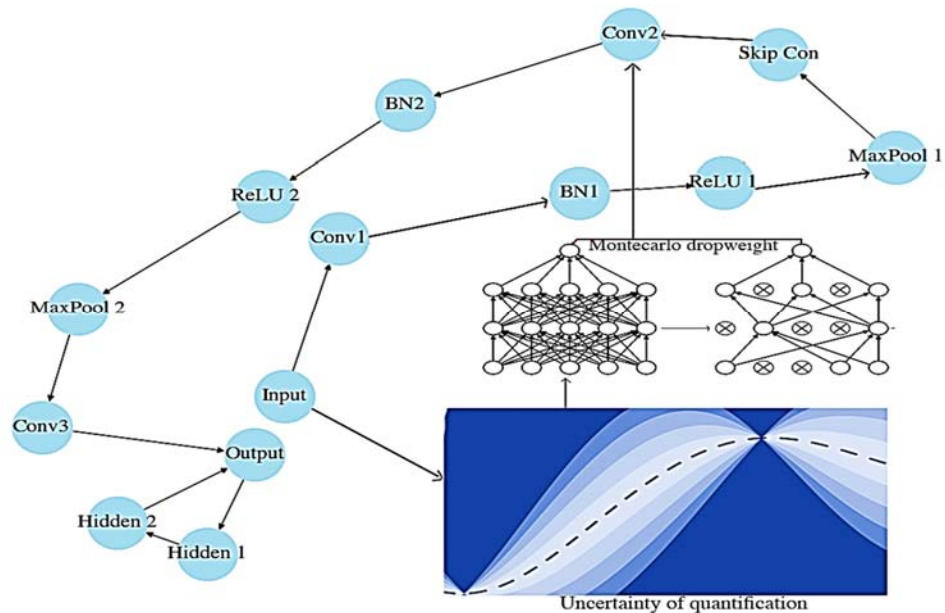


Figure A. Proposed Baresnet model architecture

**Purpose:** The study proposes two methods, with and without using Uncertainty Quantification, and proposes a new 4-layer Convolutional Neural Network model.

**Theory and Methods:** This scientific study used a dataset of 663,549 lung CT scans, split into 80% for training, 10% for testing, and 10% for validation. Two methods for detecting lung cancer in the scans were proposed. The first method employed four pre-trained neural network models (ResNet50, AlexNet, Inception v3, and VGG16) to predict cancer accuracy without utilizing Uncertainty Quantification.

**Results:** The study also proposed a new 4-layer Convolutional Neural Network model, which achieved an accuracy rate of 0.971. When Uncertainty Quantification was used, the Bayesian Neural Networks ResNet50 model achieved an accuracy rate of 0.89, the Bayesian Neural Networks AlexNet model achieved an accuracy rate of 0.86, the Inception v3 model achieved an accuracy rate of 0.91, and the VGG16 model achieved an accuracy rate of 0.77.

**Conclusion:** This study proposed two methods to improve lung cancer diagnosis accuracy using CT scans by combining Deep Learning with Uncertainty Quantification. The methods achieved high accuracy rates, demonstrating the potential of using Uncertainty Quantification to improve neural network models for lung cancer diagnosis, ultimately improving patient outcomes.



## Derin öğrenme ve belirsizlik ölççeği ile akciğer kanserinin erken teşhisi

Sema Üzülmmez<sup>1\*</sup>, Mehmet Akif Çiççi<sup>2</sup>

<sup>1</sup>İstanbul Aydın Üniversitesi, Mühendislik Fakültesi, Bilgisayar Mühendisliği Bölümü, 34250, İstanbul, Türkiye

<sup>2</sup>Bandırma Onyedli Eylül Üniversitesi, Mühendislik ve Doğa bilimleri Fakültesi, Bilgisayar Mühendisliği Bölümü, 10200 Balıkesir, Türkiye

### Ö N E Ç İ K A N L A R

- Derin Öğrenme + Belirsizlik Ölçümü, akciğer tomografisi verilerini kullanarak kanser tespiti
- Geleneksel modellerin Baresnet belirsizlik ölçümü modeli ile karşılaştırılması
- Belirsizlik ölçümü ve Bayes Ağları

#### Makale Bilgileri

Araştırma Makalesi

Geliş: 27.03.2022

Kabul: 26.02.2023

DOI:

10.17341/gazimmfd.1094154

Anahtar Kelimeler:

Derin Öğrenme,  
belirsizlik ölççeği,  
evrişimli sinir ağları,  
bilgisayarlı tomografi  
görüntüsü

#### ÖZ

Bu çalışmada Derin Öğrenme ile Belirsizlik Ölççeği tekniği birleştirilerek, 663549 akciğer tomografi veri seti kullanılarak hastaların kanser olup olmadığı; eğer kanser ise kanserin iyi huylu (benign) ya da kötü huylu (malignite) olduğuna karar verilmesi ve bu modelin ne kadar güvenilir olduğu araştırılmıştır. Bahsi geçen görüntülerin %80'i eğitim, %10'u test ve %10'u ise doğrulama için kullanılmıştır. Bu amaç doğrultusunda iki metot önerilmiştir. Birinci yöntemde Belirsizlik Ölççeği (Uncertainty Quantification) kullanılmadan yapılan ölçümlerde, ResNet50 modeline ait doğruluk oranı 0,965 olarak elde edilirken AlexNet modeli doğruluk oranı 0,837 olarak hesaplanmıştır. Ayrıca Inception v3 modelinde doğruluk oranı 0,849 ve VGG16 modelinde doğruluk oranı 0,826 olarak bulunmuştur. İlk olarak, 4 Katmanlı Evrişimsel Sinir Ağı modelinin doğruluk oranı 0,971 olarak hesaplanmıştır. İkinci yöntemde ise Belirsizlik Ölççeği (UQ) kullanılarak elde edilen ölçümlerde, Bayesian Sinir Ağları ResNet50 modelinin doğruluk oranı 0,89 olarak bulunmuştur. Ayrıca, Bayesian Sinir Ağları AlexNet modelinde doğruluk oranı 0,86, Inception v3 modelinde doğruluk oranı 0,91 ve VGG16 modelinde doğruluk oranı 0,77 olarak hesaplanmıştır. Çalışma, yeni bir model olan 4 Katmanlı Evrişimli Sinir Ağı'nın geliştirilmesi ve belirsizlik nicelleme yönteminin kullanılması gibi yenilikçi özelliklere sahiptir. Bu sonuçlar, belirli sinir ağı modellerinin performansının belirsizliğin hesaba katılmasıyla artırılabilirliğini göstermektedir.

## Early diagnosis of lung cancer using deep learning and uncertainty measures

### H I G H L I G H T S

- Deep Learning + Uncertainty Quantification improves cancer detection using lung tomography data
- Conventional models compared with Baresnet with/without uncertainty quantification
- Uncertainty quantification enhances model accuracy, especially with Bayesian Neural Networks

#### Article Info

Research Article

Received: 27.03.2022

Accepted: 26.02.2023

DOI:

10.17341/gazimmfd.1094154

Keywords:

Deep Learning,  
uncertainty quantification,  
convolutional neural  
networks,  
computed tomography image

#### ABSTRACT

In this study, the Deep Learning technique was combined with the Uncertainty Quantification scale to determine whether patients have cancer or not, and if cancer is present, to classify it as benign or malignant using a dataset of 663,549 lung tomography data. The reliability of this model was also investigated. 80% of the mentioned images were used for training, 10% for testing, and 10% for validation. Two methods were proposed for this purpose. In the first method, measurements were made without using Uncertainty Quantification, and the accuracy rate for the ResNet50 model was found to be 0.965, while the accuracy rate for the AlexNet model was calculated as 0.837. Additionally, the accuracy rate for the Inception v3 model was found to be 0.849, and for the VGG16 model, it was 0.826. Firstly, the accuracy rate of the 4-Layer Convolutional Neural Network model was calculated as 0.971. In the second method, measurements obtained using Uncertainty Quantification (UQ) resulted in an accuracy rate of 0.89 for the Bayesian Neural Networks ResNet50 model. Furthermore, the accuracy rate for the Bayesian Neural Networks AlexNet model was 0.86, for the Inception v3 model it was 0.91, and for the VGG16 model, it was 0.77. The study has innovative features such as the development of a new model called the 4-Layer Convolutional Neural Network and the use of uncertainty quantification methods. These results demonstrate that the performance of specific neural network models can be enhanced by considering uncertainty.

## 1. Giriş (Introduction)

Tıbbi görüntü analizinin sağlık sektöründe, özellikle invaziv olmayan tedavi ve klinik muayenede olağanüstü başarı sağladığı görülmüştür [1]. X-ışınları, Bilgisayarlı Tomografi (BT) ve ultrason görüntüleme gibi restoratif görüntüler, özel tanımlar için kullanılmaktadır. Tıbbi görüntüleme BT, filmlerde lezyonlu bölge görüntüsünü yakalamak için en önemli görüntüleme mekanizmalarından biridir [2]. Kanser gibi ölümcül bir hastalık söz konusu olduğunda BT görüntülerinin önemi daha da artmaktadır. Öte yandan, BT görüntüleri analizine rağmen akciğer kanseri yılda 1,61 milyon insanın ölümüne yol açmaktadır [3]. Bilgisayar Destekli Taniya (BDT) rağmen sadece 2019 yılında akciğer kanseri 1,76 milyon kişinin ölümüne sebep olmuştur [4]. Akciğer kanseri kaynaklı ölümler, tanısı yeni konulan hastaların %13'ünü, tüm kanser ölümlerinin ise %24'ünü oluşturmaktadır. Kanser kaynaklı ölümlerin yaklaşık %70'i düşük ve orta gelirli ülkelerde meydana gelmektedir [5]. BT taraması ile erken tanı, akciğer kanseri kaynaklı ölüm vakalarını azaltmasına karşın bu taramaların maliyeti yüksektir. Ayrıca, düşük ve orta gelirli ülkelerde bu durum daha büyük risk taşımakla birlikte hekimlerin görüş ayrılıklarından dolayı yanlış teşhisle sonuçlanabilmektedir.

Bu sorunlarla başa çıkabilmek için BT taramalarında otomatik lezyon saptama yöntemleri ile otomatik akciğer hastalıkları sınıflandırma yöntemleri geliştirmek, son yirmi yılda aktif bir araştırma alanı hâline gelmiştir [6]. Son yıllarda BT görüntüleri kullanılarak yapılan akciğer kanseri taramalarının sayısı artmakla beraber BDT sistemleri, yorumlamada çok önemli bir rol almıştır [7]. Birçok BDT sistemi, görüntülerden öznelik çıkarımı temel alınarak geliştirilmiştir. Geleneksel BDT sistemlerinin çoğu manuel tasarlandığı için görüntüden çıkarılan öznelikler ideal veya yeterli olabilmektedir [5]. Dolayısıyla BDT sistemlerinde istenen düzeyde başarı yakalanamamıştır. Derin Öğrenme (DÖ) son zamanlarda görüntü ve biyomedikal görüntü analizi gibi birçok alanda yoğun ilgi görmektedir [8]. DÖ, en yaygın kullanılan Makine Öğrenme (MÖ) algoritmasıdır. DÖ çeşitli alanlarda başarıyla uygulanmakla birlikte, görüntü tanımda ve sınıflandırmada en iyi performansı göstermiştir [9]. Aslında DÖ görüntü sınıflandırması, ön işlem yapılmamış görüntü piksellerine doğrudan uygulanabilmektedir. Ancak bu durumda görüntünün spektral bilgi içeriği, sınıflandırmada kullanılmamaktadır. Görüntü sınıflandırma performansını artırmak adına görüntü özneliklerinin değişmezliğini artırmak için spektral öznelik bilgisinin göz önünde bulundurulması gerekmektedir [10].

### 1.1. Literatür Taraması (Literature Review)

Literatürde yer alan pek çok araştırma ve metodoloji çerçevesinde değerlendirildiğinde, kanser hastalığına ilişkin araştırmalarda yaygın olarak karşılaşılan "yanlış sonuçlar" dikkat çekmektedir. Bu nedenle, teknolojik imkânlar sayesinde daha hızlı ve daha doğru sonuçlara ulaşmayı kolaylaştıran ve insan kaynaklı hataların en aza indirgenmesine katkı sağlayan teknolojik yöntemlerin geliştirilmesi oldukça önemlidir. Çalışmanın bu bölümünde, söz konusu önemden hareketle akciğer kanserini tanımlama yöntem ve prosedürleri, DÖ algoritmaları ile ilişkili çeşitli görüşler ve BT görüntülerinin işleme evreleri ele alınmaktadır.

Bu çalışmada, hava kirliliğinin yüksek etkisi hakkında araştırma yapmak ve farklı bireylerde farklı akciğer kanseri türlerinin oluşumunu tahmin etmek için Modifiye Evrişimli Sinir Ağı yöntemi (MESA) kullanılmaktadır. Bu araştırma çalışması, akciğer BT taramalarından elde edilen özellikleri kullanarak MESA'yı eğitmekte ve BT görüntülerinden toplanan performans ölçümlerini kullanarak performanslarını değerlendirmektedir [11]. Bu çalışmada ise akciğer dijital patolojisinde yapay zekâ (AI) araçları oluşturmak için

kullanılan metodolojiler tartışılmakta ve çeşitli el yapımı ve DÖ tabanlı denetimsiz özellik yaklaşımları açıklanmaktadır [12]. Ardından kanser, tüberküloz, idiyopatik pulmoner fibroz ve Covid-19 dâhil olmak üzere geniş bir akciğer hastalıkları yelpazesinde AI araçları gözden geçirilmektedir. PD-L1 (programmed death ligand 1), akciğer hastalığı teşhisi, risk sınıflandırması ve bağışıklık kontrol noktası inhibitörleri gibi tedavilere yanıtın tahmin edilmesi gibi biyobelirteçlerin nicelleştirilmesi dâhil olmak üzere farklı klinik problem türleri için yeni görüntüleme biyobelirteçlerinin faydası tartışılmaktadır. Ayrıca multimodal veri analizi, 3D patoloji ve transplant reddi gibi akciğer dijital patolojisinde yapay zekâ (AI) araçlarının ortaya çıkan bazı uygulamalarına da kısaca değinilmektedir. Son olarak düzenleyici onay, geri ödeme modelleri geliştirme, klinik dağıtım planlama ve AI ön yargılarını ele alma konusundaki zorluklar açıklanarak Doğru Pozitif (DP) tabanlı AI araçlarının geleceği tartışılmaktadır.

Bir başka çalışmada ise hastalarda akciğer kanserini tespit ve teşhis etmek için çeşitli doğal dil işleme teknikleri hakkında kapsamlı bir inceleme yapılmaktadır [13]. Çalışma, doğal dil işlemenin tahmin yeteneğini geliştirmek için uygulanan çeşitli teknikler içermektedir. Çalışma ayrıca akciğer kanseri hastalıklarının saptanmasında gelecekteki iyileştirmelerin sınırlarını ve kapsamını sunmaktadır.

Bir diğer çalışmada akciğer kanserinin erken teşhisi için karşılaştırmalı bir çalışma yapılmıştır. Akciğer kanseri tanısında birçok yöntem geliştirilmiştir [14]. Bazılarında röntgen görüntüleri, bazılarında ise BT tarama görüntüleri kullanılmaktadır. Ayrıca bu çalışmada görüntü veri setinden akciğer kanserini tanımlamak için çoklu sınıflandırma yöntemleri çeşitli segmentasyon algoritmaları ile birleştirilmiş ve BT tarama görüntülerinin röntgen görüntülerine göre daha doğru olduğu bulunmuştur. Bu nedenle akciğer kanseri tespiti için BT tarama görüntüleri dikkate alınmaktadır. Bunlara ek olarak diğer tekniklerle karşılaştırıldığında, işaretleyici kontrollü havza segmentasyonu daha doğru bir çıktıya sahiptir. Bu teknik, geleneksel MÖ algoritmalarına kıyasla DÖ tekniğini kullanırken daha yüksek doğruluğa sahiptir.

Bu çalışmada, akciğer kanseri riskini belirlemek için MÖ ve DÖ'yu kullanmanın farklı yöntemlerinin bir karşılaştırması tartışılmaktadır. Son on yılda, akciğer kanserini teşhis etmek için BT taramaları ve X-ray dâhil olmak üzere çeşitli yöntemler geliştirilmiştir [15]. Ek olarak, görüntüleri tanıyarak akciğer kanseri nodüllerini tespit etmek için çeşitli sınıflandırıcılar ve segmentasyon algoritmaları kullanılabilmektedir. BT tarama görüntülerinin bu çalışmada daha doğru sonuçlar verdiği bulunmuştur. Aslında akciğer kanseri ölüme neden olan yaygın bir kanser türüdür ve tanı koymanın en iyi yolu BT taramasıdır.

Burada, akciğer kanseri teşhisi için bilgisayar tabanlı tespit ve izleme sistemlerindeki son gelişmelere vurgu yapılmakta ve gelecekteki senaryoların hangi modellerden ve perspektiflerden oluşacağı tartışılmaktadır [16]. Bu stratejik derlemede, akciğer kanseri teşhisi için geleneksel yöntemler, ileriye dönük akciğer kanseri teşhis araçları için bilgisayar tabanlı tespit ve izleme sistemlerindeki son gelişmeler, farklı potansiyel bilgisayar türleri, bunların sentezleri ve biyomedikal uygulamalar için artıları ve eksileri tartışılmaktadır. Ayrıca, örneklem tekniklerini 3D baskı ve AI gibi daha ileri teknolojilerle entegre etmenin, akciğer kanserinin hızlı, verimli ve erken aşamada teşhisinin nasıl sağlanacağı tartışılmaktadır.

Bu derleme, son zamanlardaki akciğer-nodül segmentasyonu ile ilişkili temel prosedürlere genel bir bakış sağlamakta ve mevcut tekniklerin dezavantajlarını çözerek yeni güvenilir sistemlerin uygulanması için fırsatçı kapsamları tartışmaktadır [17]. %98'den

fazla doğrulukla verimli algoritma geliştirmeyi vadeden çok sayıda çalışma vardır, ancak herhangi bir sonuca varmak için çeşitli parametrelerden geçmek gerekmektedir. Bunların arasında en önemlisi, kullanılan verinin boyutu ve türüdür. Diğer bir önemli parametre, çeşitli veri işleme aşamaları sırasında algoritmaların eklenmesidir. Bunun dışında, boyutsallık ve veri toplama kaynağı da tüm prosedürü etkilemektedir. Bu makale, farklı çalışmaları, parametre değişikliğinden kaynaklanan verim sapması kolayca anlaşılabilir şekilde tanımlamaktadır.

Bu çalışmada kullanılan farklı sınıflandırma metodolojileri, akciğer kanserinin erken evrede öngörülmesi amacıyla açıklanmıştır [18]. Kanser türünün akciğer olup olmadığını belirlemek için MÖ teknikleri kullanılmaktadır. Tümörler kötü huylu veya iyi huyludur. Ayrıca çalışmada MÖ yaklaşımlarından Gömülü Çözüm Sözleşmesi (ESA), Destek Vektör Makineleri (DVM), Yapay Sinir Ağları (YSA), Çok Katmanlı Perceptron (MLP), K-En Yakın Komşu (KNN), Entropi Bozunma yöntemi ve Rastgele Orman (RF) ayrıntılı olarak tartışılmış ve performansları açısından değerlendirilmiştir. Performans ölçümleri doğruluk, duyarlılık ve özgüllük gibi metriklerden oluşmaktadır. Bu analizde, ESA küçük veri seti kullanan yaklaşım, %96 ile en iyi sonucu göstermektedir. Diğer metodolojilere kıyasla doğruluk ve EDM, %77,8 ile en kötü doğruluk oranı hesaplanmıştır.

Bu makalede, akciğer kanserinin yönetiminde yapay zekânın (AI) rolüne odaklanılmaktadır. Akciğer kanseri tespit ve teşhisinde yaygın olarak kullanılan algoritmalar, güncel uygulamalar ve yapay zekânın zorlukları özetlenmektedir [19]. Bu çalışmada üç ana derin sinir ağı kullanılmıştır ve kapsamlı bir şekilde gözden geçirilmiştir. Açık kaynak akciğer BT görüntüleri Akciğer Görüntü Veritabanı Konsorsiyumu Görüntü Koleksiyonu (LIDC-IDRI) ile tespitlerin doğruluğu değerlendirilmiştir [20]. İyi huylu ve kötü huylu akciğer nodüllerinin sınıflandırılmasında tahminler yapılmıştır. Bu makale akciğer kanserinin gözlemlenmesi hakkındadır. Burada akciğer kanserinin izlenmesi için BT kullanılmaktadır [21]. Akciğer kanseri tahminini doğru bir şekilde taramak için KNN, SVM, Karar Ağacı (DT) ve daha pek çok algoritma kullanılmıştır. Tanıtılan analizin amacı, maksimum doğrulukla bir hastada akciğer kanseri olasılığını azaltabilecek bir model tasarlamaktır. Çalışmada ilk olarak akciğer kanserini tanımlamak için normal olarak kullanılan sınıflandırma tekniklerinin kısa bir tanımını yapılarak çeşitli MÖ tekniklerine yer verilmiştir. Bu alanda MÖ yöntemleriyle akciğer kanseri

sınıflandırma analizleri üzerine odaklanılmaktadır. Ayrıca, incelenen makalenin ayrıntılı bir karşılaştırma tablosu sunulmuştur.

Bu bölümün ana amacı, akciğer nodülü sınıflandırması için iki boyutlu ve üç boyutlu dayalı farklı DÖ tekniklerinin, ortak veri tabanlarındaki etkinliklerini değerlendirerek karşılaştırmaktır. Biyomedikal mühendisliği alanındaki hızlı gelişme, tıbbi görüntü analizini ileri araştırma ve geliştirmenin önde gelen alanlarından biri haline getirmiştir [22]. Bu gelişme ve ilerlemenin ana nedenleri, tıbbi görüntü analizi için kullanılan MÖ tabanlı tekniklerdir. DÖ, otomatik özelliklerin öğrenilmesi için bir sinir ağına kullanıldığı, belirgin bir şekilde kullanılan bir MÖ tekniğidir. El ile öznelite çıkarma yöntemlerinde öznelite seçme, hesaplama gibi zorlu süreçlerden geçilir. Tıbbi görüntüleri analiz etmek için el yapımı öznelite çıkarma yöntemleri yerine aktif olarak kullanılan Derin Evrimsel Ağlar (DCN), uygulama alanları, yani Bilgisayar Destekli Tasarım (CAD), hastalık sınıflandırması, segmentasyon vb. DCN kullanan en son tıbbi görüntü analizi incelemesi ve ayrıca bu araştırma alanıyla ilgili büyük zorlukların altı çizilmiştir.

Önceki bir çalışmada, ConvPath modeline, çekirdek merkezi, nükleer sınır veya çekirdek olmayanlar dâhil olmak üzere TME'nin topolojik özellik haritalarına dayalı olarak farklı çekirdekleri otomatik olarak tanımlamayı öğretilmiştir. Bu özelliklere dayanarak hastaları yüksek ve düşük riskli alt gruplara ayırmak için görüntü özelliğine dayalı bir prognostik model kullanılmıştır [23]. Yaş, cinsiyet, sigara içme durumu ve evre gibi klinik değişkenler için düzeltme yapıldıktan sonra, yüksek riskli alt grup, her iki bağımsız kohortta da daha kötü sağkalım ile ilişkilendirilmiştir.

Başka bir çalışmada, Lopez-Ubeda destek vektör makinelerini ve rastgele orman yöntemlerini kullanarak akciğer kanseri sınıflandırma görevi için histopatolojik görüntüler üzerinde veri büyüme gerçekleştirilmiştir. Daha sonra evrimsel sinir ağı modelini kullanarak sınıflandırdıkları çalışmalarında %96 oranında doğruluk değeri elde ettiklerini ifade etmişlerdir [24]. Genel olarak, seçilen çalışmalar BT görüntülerinde pulmoner nodüllerin tespiti, segmentasyonu ve sınıflandırılmasında iyi yetenekler göstermiştir. MÖ tekniklerinin tatmin edici performans sonuçları gösterdiği, özellikle ESA olmak üzere DÖ'nün geleneksel modellerden daha iyi performans gösterdiği ve umut verici bir yaklaşım olarak ortaya çıktığı tespit edilmiştir. Tablo 1'de literatürde bulunan çalışmaların karşılaştırılması gösterilmektedir.

**Tablo 1.** Literatür Analizi (Literature Analysis)

Çalışma	Metotlar	Eğitim/Test Dağılım Oranları	Veri Seti	Başarı Oranı
Kanageswari vd. [11]	MESA	80/20	KAM	0,95
Viswanathan vd. [12]	ESA	70/30	Kanser, tüberküloz, veri merkezi	0,84
Jabir vd. [13]	Derin ESA	70/30	LUNA 16	0,90
Rastogi vd. [14]	DÖ, MÖ	80/20	LIDC-IDRI	0,82
Shivwanshi vd. [15]	MÖ	80/20	Global kanser verisi	0,98
Taher vd. [16]	ESA, DVM, YSA	70/30	LIDC-IDRI	0,96
Zhang vd. [17]	AI-Bilgisayarlı Görme	80/20	LIDC-IDRI	0,84
Kaur [18]	KNN, DVM	70/30	BT tarama görüntüleri	0,96
Yar vd. [19]	DCN, ESA	80/20	LIDC-IDRI	0,92
Li vd. [20]	RF, DT	80/20	PubMed, EMBASE	0,96
Silva vd. [21]	MÖ, DVM	70/30	Embase, PubMed	0,91
Xie vd. [22]	Naive Bayes, YSA, DVM	80/20	LUNA 16	0,97
Hsu vd. [23]	Naive Bayes, DVM	80/20	HT medica hastanesi	0,98
Lopez-Ubeda vd.[24]	DVM, RF, YSA	80/20	LIDC-IDRI	0,92
Yapılan çalışma	VGG16, ResNet50, AlexNet, ResNet, 4K-ESA Bayes Ağları, Belirsizlik modeli, DropWeights	80/20	LUNA 16	0,98

Çalışmanın bu bölümünde değinilen araştırmalar, öne çıkan bir model, genomik ve klinik gibi karışık bilgi kombinasyonlarını içermektedir.

### 1.2. Yenilik ve Katkılar (Contribution and Novelty)

Bu çalışmada, BT görüntülerinin DÖ ile Belirsizlik Ölçeği (UQ) yöntemleri birleştirilerek modelin ne kadar öğrenip öğrenmediği ve çıktının güvenilirliği ölçülmektedir. Elde edilen verilere DÖ yöntemi uygulanarak kişinin hasta olup olmadığı, eğer hasta ise hastalığın iyi veya kötü huylu sınıfı saptanabilmektedir.

Çalışmanın literatüre katkıları şu şekilde sıralanabilir:

- Önerilen algoritma ile heterojen ve dengesiz veri kümelerine daha hızlı yakınsama gösterir.
- Yenilikçi bir model geliştirilmiştir.
- UQ ile modelin daha güvenilir olması sağlanmaktadır.
- Geliştirilen yenilikçi 4 Katmanlı Evrişimsel Sinir Ağı (4K-ESA) ile neden öğrenilmediği, öğrenememe sorununun aleatorik mi yoksa epistemik mi olduğu tespit edilmektedir.

Bu makalenin ikinci bölümünde, benzer çalışmalardan örnekler verilerek bu yayın ile bulguların karşılaştırılmasına yer verilecektir. Bölüm 3'te akciğer kanserinin saptanmasında göz önüne alınan yöntemler üzerinde durulacaktır. Bölüm 4'te, önerilen model ile ilgili uygulama detaylandırılacaktır. Bu bölümü tartışma ve deneysel bulgular izleyecektir. Bölüm 5'te ise özel sonuçlar üzerinde durulacaktır.

## 2. Veri ve Yöntem (Methods and Data)

DÖ çalışmasının bu bölümünde, deneylerde kullanılan veriler ile verilerin güvenilirliği, elde edilme şekli ve kullanılabilirlik düzeyine getirilmelerine ilişkin başlıklar yer almaktadır.

### 2.1. Deneylerde Kullanılan Veri Kümeleri (Managing the Dataset)

BT görüntüleri ile akciğer kanseri tanısı koymak için kullanılan DÖ algoritmasının, ne derece başarılı olduğunun test edilebilmesi ve BDT yöntemi ile kıyasının deneysel olarak ölçülebilmesi amacıyla Kanser Arşiv Merkezi (KAM) veri tabanlarına ait çeşitli veri kümelerinden faydalanılmıştır [25]. Faydalanılan söz konusu veri kümelerine ilişkin bilgilere takip eden alt başlıklarda değinilmiştir. Çalışmada kullanılan veri seti, açık kaynak KAM'dan elde edilen akciğer BT taramalarına ait Tıpta Dijital Görüntüleme ve İletişim (DICOM) görüntüleridir. Çalışmada kullanılan veri seti Tablo 2'de görüldüğü gibi, toplamda 26.254 hastalık şüphesi taşıyan bireyden alınan 639.598 BT göğüs görüntüsünden oluşmaktadır. Örnek olarak farklı akciğer BT görüntüleri Şekil 1'de verilmiştir. Şekil 1'de görüldüğü üzere Ağustos 2002-Nisan 2004 yılları arasında yaklaşık 54.000 katılımcıdan alınan

görüntüler (Tablo 2), KAM BT veri seti testlerine ilişkin kontrollü bir klinik araştırmasına aittir. Katılımcılar eşit oranlarda (rastgele olacak şekilde) iki gruba ayrılmıştır. İlk gruptan düşük doz helisel BT alınırken ikinci gruptan tek görüntülü göğüs radyografisi alınmıştır. Ortalama takip süresi 6,5 yıl olan bu çalışmadan elde edilen BT veri seti ile ortaya çıkan kanserlerin teşhis edilme durumları ve ölüm oranları hakkında bilgiler edinilmiştir.

**Tablo 2.** Çalışmada kullanılan veri seti özellikleri (Dataset properties used in the study)

Veri seti	Özellikler
Seri sayısı	203,099
Ortalama takip süresi (yıl)	6,5
Hasta sayısı	26.254
Görüntü sayısı	639.598
Veri formatı	DICOM(BT)
Görüntü boyutu (GB)	11,3

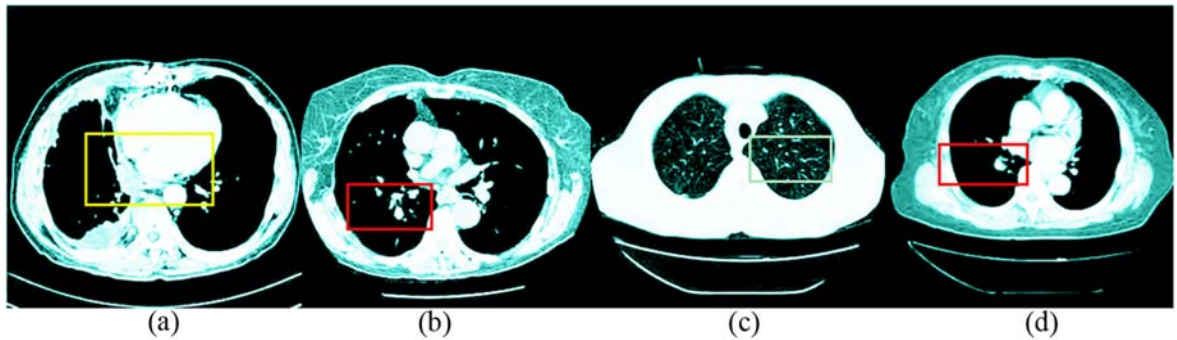
### 2.2. Geleneksel Evrişimli Sinir Ağları (Conventional Convolutional Neural Networks)

#### 2.2.1. AlexNet ağı (AlexNet architecture)

AlexNet, 8 katmandan oluşan bir derin öğrenme modelidir. Bu katmanların ilk 5'i evrişim katmanı ve son 3'ü tam bağlantılı katmandan oluşmaktadır. Son katman, softmax fonksiyonu kullanarak 1000 sınıf arasında bir çıktı üretmektedir. AlexNet'in temel özelliklerinden biri, aynı girdiye birden çok evrişim işlemi ve havuzlama işlemi uygulamasıdır. Girdi verisine, maksimum havuzlama işlemi ve 1x1, 3x3 ve 5x5 boyutlarında filtreler uygulanmaktadır. Bu işlemlerin sonuçları birleştirilerek tek bir çıktı elde edilir. Bu özellik sayesinde, AlexNet hem özel hem de genel özellikler çıkartabilir. Ancak, AlexNet'in en büyük sorunlarından biri hesaplama süresinin uzun olmasıdır. 5x5 evrişim tabakalarının çok sayıda filtre içermesi nedeniyle, boyut büyümesi kaçınılmazdır. Havuzlama katmanları da eklenince, hesaplama süresi daha da uzun hale gelir. Bu sorunu çözmek için, her işlem öncesinde boyut azaltma işlemi yapılabilir.

#### 2.2.2. VGG16 ağı (VGG16 network)

VGG16 ağı, ESA yönteminde, görüntü işlemede en yaygın olarak kullanılan DÖ mimarisidir. Verilerin büyümesi ve veri işlemedeki donanım geliştirmelerin kolaylığı, ESA görüntü işleme başarısını artırmıştır. Ağın girdisi kırmızı-yeşil-mavi (RGB) renk formatında ve 224x224 boyutlarında bir görüntüdür. Ağa verilmeden evvel görüntülerdeki her pikselden, eğitim setindeki görüntülerle hesaplanan ve ortalama resme ait olan piksel değerleri çıkartılır. VGG16 modelinde giriş kanallarının doğrusal bir dönüşümü olarak görülebilen 1 x 1 evrişim filtrelerinden yararlanılmaktadır. Evrişim adımı 1 piksel sabitlenmiş olup modelde dolgu çözünürlüğü



**Şekil 1.** a) Kötü huylu b) İyi huylu c) Nodül yok d) Sağlıklı görüntü (a) Malignant (b) Benign, (c) Not nodule (d) Healthy lung

evrişimden sonra korunacak şekilde ayarlanmıştır. Yani dolgu, 3x3 dönüşüm için 1 pikselidir. Havuzlama için beş maksimum havuzlama katmanından yararlanılmaktadır. Tablo 3'te gösterilen eğitim esnasındaki VGG16 girdisi, sabit boyutlu 224 × 224 RGB görüntüden meydana gelmektedir. Görüntü, ön işlemden sonra eğitim setinde hesaplanan ortalama RGB değerini her pikselden çıkarmaktadır.

### 2.2.3. Inception v3 ağı (Inception v3 network)

Inception v3 mimarisi modüllerden meydana gelmektedir. Her bir modül, farklı boyutlu evrişim ve maksimum havuzlamalardan oluşur. Bu nedenle her bir modüle 'inception' ismi verilmiştir. Tam olarak 9 Inception bloğundan meydana gelen mod, GoogLeNet olarak da anılmaktadır. Kısacası Inception v3, bir tür ESA modelidir. Birçok evrişim ve maksimum havuzlama adımlarından oluşmaktadır. En son katman ise tam bağlı bir sinir ağı içermektedir. Inception v3, 2012 yılında Evrişimli Sinir Ağı modellerini ve DÖ yöntemini yeniden popüler hâle getiren ilk çalışmadır. Alex Krizhevsky, Ilya Sutskever ve Geoffrey Hinton gibi araştırmacılar tarafından geliştirilmiştir [26]. Temel olarak, birbirini takip eden evrişim ve havuzlama katmanlarından oluştuğundan LeNet modeline çok benzer. ReLU aktivasyon fonksiyonu olarak kullanılır ve havuzlama katmanlarında maksimum havuzlama kullanılır. Ayrıca 5x5 ve 3x3 boyutlarındaki filtreler uygulanmadan önce 1x1 boyutlarında filtre uygulanarak boyut küçültülmektedir. 2014 yılında düzenlenen ILSVRC14 yarışmasında birinciliği kazanan bu ağ, gereken maliyeti artırmadan hesaplama için daha geniş ve daha derin bir yapı sunmaktadır. 22 katmanın meydana getirdiği ağı Inception adı verilmiştir. Ding vd. yaptıkları bir çalışma esnasında kullandıkları "Daha derine gitmeliyiz!" konulu internet söyleminden türetilmiştir [27]. Birden çok evrişim işlemi ve havuzlama işlemlerinin kombinasyonu ile özel ve genel özellikleri çıkartabilen bir yapıya sahiptir. Aynı zamanda, paket normalizasyonu kullanarak sinir ağının çıktılarını normalleştirir. Inception v3'ten farklı olarak, 3x3 boyutlu filtreler daha sık kullanılır ve ağın derinliği ve genişliği en iyileştirilir. Ağın amacı, maksimum bilgi akışı sağlamak için bu özellikleri bir araya getirmektir.

### 2.2.4. ResNet50 ağı (ResNet50 network)

ResNet50 adlı bir sinir ağı mimarisinden bahsedilmektedir. "Residual Network" ismiyle anılan bu mimari, gerçek anlamda derinleşmeye başlayan ağların bir türüdür. ResNet50, önceki modellerden farklı olarak, kalıntı değerlerin sonraki katmanlara beslendiği blokların modele eklenmesiyle oluşturulur. Microsoft tarafından geliştirilen ResNet50, kalıntı bloklarından oluşur. Kalıntı bloklar, girdi olarak  $x$  değerini alır, evrişim, aktivasyon ve bir fonksiyonla işleme tabi tutulur ve sonuç olarak  $f(x)$  fonksiyonu elde edilir. Daha sonra,  $f(x)$

fonksiyonuna orijinal girdi olan  $x$  değeri eklenerek  $h(x) = f(x) + x$  üretilir. Bu sayede, gradyanın 0'a yakınsaması probleminin ortadan kaldırılması ve öğrenmenin daha derin ağlarla gerçekleştirilmesi sağlanır. ResNet50, birkaç katman önde yer alan aktivasyon fonksiyonu ile anlık aktivasyon fonksiyonunun girdisini toplayarak işlem yapar ve bu sayede kaybolan gradyan problemi meydana gelmeden öğrenme gerçekleştirilir.

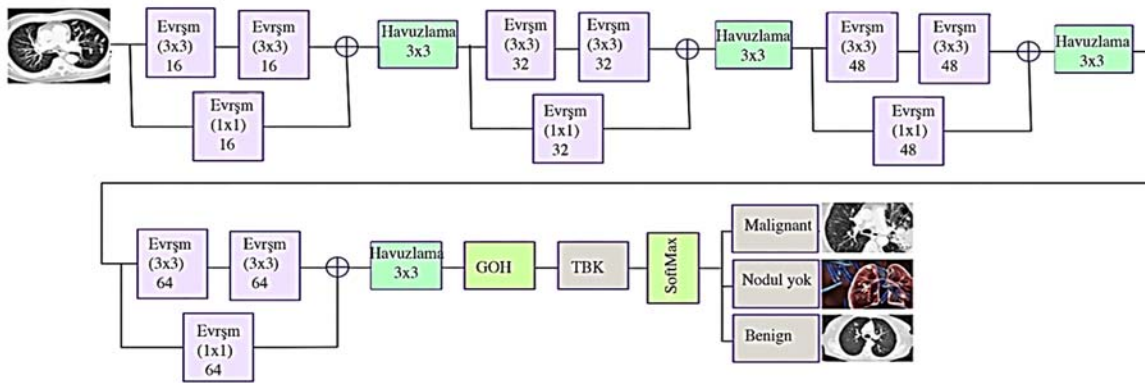
ResNet50 ağı, geleneksel ağlardan farklı bir yapıya sahiptir. Bu ağ 2015 yılında birinci olmuş, daha sonra eğitim hızını artırmak amacıyla Inception yapısıyla kalıntı ağ bağlantıları birleştirilerek Inception v3 ağı geliştirilmiştir. Teorik olarak ResNet50 modeline ne kadar katman eklenirse başarının o kadar artacağı düşünülmekte ise de pratikte bu teorinin geçerli olmadığı deneyimlenmiştir. VGG16 modelinde veri eğitimi için VGG16 kullanılmıştır. BT görüntüleri eğitim ve test için iki gruba ayrılmıştır. VGG16 bir ESA mimarisi olup 16 katmanı bulunmaktadır. Katmanlar; evrişimli katmanlar, maksimum havuzlama katmanları, aktivasyon katmanları ve TBK'lerden oluşur. 13 evrişimli, 5 maksimum havuzlama ve 3 yoğun katman olmak üzere toplam 21 katmandan oluşmasına rağmen sadece 16 ağırlık katmanı vardır. Evrişim 1'de 64 filtre bulunurken Evrişim 2'de 128, Evrişim 3'te 256, Evrişim 4 ve 5'te ise 512 filtre bulunmaktadır.

Ayrıca VGG16 basit bir ağ modelidir. Önceki modellerden en önemli farkı, evrişim katmanlarının 2'li ya da 3'lü kullanılmasıdır.

### 2.2.5. Dört katmanlı evrişimli sinir ağı (4-K CNN architecture)

Şekil 2'de görüldüğü gibi konektör evrişimli bloğun çıkışı, maksimum havuzlamadan önce ikinci ESA bloğunun çıkışı ile toplanmaktadır. Maksimum havuz oluşturma adımının 2 olması nedeniyle çözünürlük 4 faktör azalır. Her evrişimli bloktan önce boyutu korumak için sıfır dolgu uygulanır. Eğitimi hızlandırmak için her katmana mini toplu normalleştirme uygulanır. Her bir evrişimli katmana iki ardışık 3 x 3 evrişimli bloğu, ReLU'ları, ReLU ile 1 x 1 evrişimli bir bağlayıcı bloğu ve bir maksimum havuzlama bloğu vardır. İlk katmanda her bir evrişimli bloktaki filtre sayısı 16'dır. İkinci katmanda 32, üçüncü katmanda 48 ve son katmanda 64 filtre bulunmaktadır. Filtrelerdeki adım kaydırma oranı 1'dir.

4K-ESA modelindeki katman sayısı dörttür. Böylece reseptör penceresi tüm görüntüyü kaplar. 4K-ESA modeli, beş katmanla test edilmiş ancak dört katman daha iyi performans göstermiştir. Girdi görüntüsünün her pikseli, tüm veri tabanının piksellerinin ortalaması ve standart sapması ile normalleştirilir. Mini toplu normalleştirme dört kısımdır ve maliyet fonksiyonu kategorik çapraz entropidir. Her mini partiden önce, öğrenmenin tamamen rastgele hâle getirilmesini sağlamak için örnekler yönlendirilmiştir. Bu durum aynı zamanda çok



Şekil 2. 4K-ESA mimari yapısı (4 K-ESA architecture)

**Tablo 3.** Kullanılan mimariler ve özellikleri (Architectures used and their characteristics features)

Model	Katman	Filtre/nöron sayısı	Çekirdek boyutu	Kayıp fonksiyonu	Optimizasyon	Epok sayısı	Başarı oranı					
AlexNet	Conv 1	96	11 x 11	Seyrek kategorik Çapraz entropi	ReLU	40	0,86					
	Ortaklama	-	3 x 3									
	Conv 2	256	5 x 5									
	Ortaklama 2	-	3 x 3									
	Conv 3	384	3 x 3									
	Conv 4	384	3 x 3									
	Conv 5	256	3 x 3									
	Ortaklama 3	-	3 x 3									
ResNet 50	Bırakma	0,5	-	Çapraz entropi	ReLU	40	0,89					
	Conv 1	64	7 x 7									
	Ortaklama	-	3 x 3									
	Conv 2	64	3 x 3									
	Conv 3	256	1 x 1									
	Conv 4	512	3 x 3									
	Conv 5	1024	1 x 1									
	Ortaklama	-	7 x 7									
	Conv 6	2048	1 x 1									
	Ortaklama	-	5 x 5									
	Çıkış	1	-									
	VGG16	Conv2D	64					3 x 3	Kategorik çapraz entropi	ReLU	40	0,77
Ortaklama		-	2 x 2									
Conv2D		128	3 x 3									
Conv2D		192	3 x 3									
Ortaklama		-	1 x 1									
Conv2D		256	3 x 3									
Conv2D		512	3 x 3									
Ortaklama		-	1 x 1									
Conv2D		768	3 x 3									
Bırakma		0,5	-									
Conv		64	7 x 7									
Inception v3		Ortaklama	-	3 x 3	Çapraz entropi	ReLU	40	0,91				
	Inception 3a	256	-									
	Inception 3b	480	-									
	Conv	256	3 x 3									
	Ortaklama 3	-	7 x 7									
	Conv2D	32	2 x 2									
	Ortaklama2D	-	2 x 2									
	Conv2D	32	2 x 2									
	Ortaklama 2	-	2 x 2									
	4K-ESA	Düzleştirme	-	-					Çapraz entropi, ikili çapraz entropi	ReLU Softmax	40	0,98
		Dense 1	256	-								
		Dense 2	258	-								
Dense 3		4	-									
Bırakma		0,3	-									

yüksek boyutlu işlemlere başa çıkmada yardımcı olmaktadır. İlk ağırlıklar normalizasyon uygulanarak bulunmuş, ağırlıkları optimize etmek için Adam en iyilemesi kullanılmıştır. Parametreler Eş. 1’de gösterilmektedir [31].

$$\beta_1 = 0,9, \beta_2 = 0,999, \varepsilon = 1 \times 10^{-9} \quad (1)$$

Önerilen 4K-ESA mimarisi, geleneksel mimarinin bir varyasyonudur. Katman sayıları bu iki mimari arasındaki temel farktır. Önerilen mimaride modelin hafif olması daha az sayıda katman ile sağlanmıştır. Önceki geleneksel mimarilerde İlgili Alanı (İA) çıkarılmayarak kullanılmıştır. Yeni 4K-ESA mimarisi, önceden temizlenmiş ve işlenmiş olan aynı verilere uygulanmıştır. Ayrıca deneylerde örneklerin seçimi için öncelikle 3 mm'nin üzerinde kalınlığa sahip tüm taramalar kaldırılmıştır. Anlaşılması kolay, 3 mm ve üzeri boyuttaki nodüller seçilmiştir. Nodüller “malignite”,

“benign” ve “nodül değil” olarak farklı evrelerde sınıflandırılmıştır. Genel olarak veri bölümünde açıklandığı gibi aynı veriler kullanılmıştır. Dolayısıyla MCS'nin birincil hedefi, rastgele tahminler oluşturmak ve bunları olasılık dağılımından örnekler olarak yorumlamaktır.

Eğitim adımları, gradyan iniş yöntemlerinden biri kullanılarak ve her öğrenme adımından önce seyreltme ölçütlerinin yeniden değerlendirilmesiyle düzenli bir şekilde gerçekleştirilir [31].

$$\mathcal{L} = \frac{1}{N} \sum_{i=1}^N \|\hat{y}(x^{(i)}) - y^{(i)}\|^2 + \lambda \sum_{l=1}^L (p_D \|W_l\|_2^2 + \|b_l\|_2^2) \quad (2)$$

Eş. 2’de tahmin esnasında rastgele ağ çıktıları ile sonuçlanan her ileri geçişten önce rastgele seyreltme matrislerini yeniden değerlendirir. Ortalama tahmin, daha sonra T ileri geçişlerin ortalaması alınarak

yapılırken belirsizlik ve ampirik varyans açısından tahmin edilebilir. MCS uygulaması, etkinleştirme fonksiyonu  $\sigma$  ve başlangıç parametreleri  $W_1, \dots, W_L, b_1, \dots, b_L$  ile düzenli bir ÇKP derinliği olarak başlar. Yalnızca büyük modelleri dikkate alındığında bu durumda bir skaler değer olacaktır. Normal ileri besleme mimarisi yerine, daha sonra Z1... ZL seyreltme eklenir. Eş. 2’de görüldüğü üzere, birinci gizli katmandaki düğüm sayısı gösterilmektedir. ZL, tüm seyreltme katmanları için aynı olan Bernoulli olasılık pD ögesi (0,1) matrisidir. Her katman birden fazla olasılığa sahip olsa da bu çalışmada, denklem tek bir olasılık pD’si ile sınırlanır. Eğitim süreci için ağırlık azaltma faktörü tanımlanır. Eş. 2’de gösterildiği gibi tüm parametreler için aynı faktörleri kullanmak yerine, kayıp fonksiyonuyla sonuçlanan pD olasılığına göre ağırlık matrisleri için ölçeklendirilmiştir. Aleatorik Belirsizlik (Aleatoric uncertainty). İstatistiksel belirsizlik, veri oluşturma sürecindeki rastlantısallık nedeniyle ortaya çıkan bir belirsizliktir. Bazı etkiler modelde dikkate alınmayabilir veya ölçülemez olabilir, bu da kesin sonuçların bilinmemesine ve her durumda belirli bir varyansın oluşmasına neden olur. Örneğin, BT görüntüleme fotonların rastgele dağılımından kaynaklanan sensor gürültüsü nedeniyle istatistiksel belirsizlik oluşabilir. Bu tür belirsizlikler, sınıflama modelinde belirtilen hata terimine karşılık gelir. Bilinmeyen sonuçları tahmin ederken gerçek süreç bilinse bile, gerçek sonucu potansiyel gürültüden etkilenebileceğinin farkında olmak gerekir. Öte yandan, sahip olunan modelin gerçek veri üreten bir model olup olmadığı konusu da bilinmelidir. Gerçek değerli girdiler ve çıktılar dikkate alındığında ve tüm çıktılar gürültüsüz olduğu varsayıldığında, doğrusal bir model iki gözlemlerle belirlenir. Bununla birlikte, gürültü ortaya kosa dahi sonuçtan kesin olarak emin olmak mümkün değildir. Ne kadar çok veri toplarsa güvenilirlik o kadar fazla olacaktır. Eş. 3’te bu durum açıkça gösterilmektedir [31].

$$f : [0, 2\pi] \times [0, \pi] \times \mathbb{R} \rightarrow \mathbb{R}^2 \quad (3)$$

Akciğer kanseri tespitinin anlamlı olabilmesi, toleransın tipik olarak çok daha sıkı olmasına bağlıdır. BT görüntüsüne bağlı olarak öngörücü posterior dağılımla elde edilen sonuç, yüksek veya düşük bir performansla iyi bir sonuç verebilir. Bununla birlikte, DÖ’de öngörücü belirsizlik iki ayrı belirsizlik biçiminde gösterilir. Çıktıların gürültüden etkilenip etkilenmediğini ve modelin öğrenme düzeyini ölçmek için iki farklı belirsizlik türü işlenmiştir. Bu belirsizlik türleri şunlardır:

- Epistemik Belirsizlik (Epistemic Uncertainty)

“Sistemik belirsizlik” olarak da bilinen “epistemik belirsizlik”, potansiyel olarak mevcut olan veri eksikliğinden kaynaklanmaktadır. Daha fazla sayıda gözlem, bu tür bir belirsizliği ortadan kaldıracaktır. Epistemik belirsizlik, yalnızca gerçek model oluşturur. Ancak bu, veri ön işleme yapılan hataların bir sonucudur. 4K-ESA modelinde veriler üzerine İA işlemi yapıldığı için epistemik belirsizlik daha düşüktür. Ancak, farklı parametrelerle diğer geleneksel modellerde hem sistemik hem de istatistiksel belirsizlik bariz bir şekilde görülmektedir. Her ne olursa olsun, sonunda girdi katmanına yeterli miktarda BT verisi verildiğinde epistemik belirsizlik ortadan kalkacaktır. Geleneksel olarak, epistemik belirsizlik altında model doğrulamasını uygulamak daima zor olmuştur. Bu nedenle, akciğer kanserine ait bir BT görüntüsünde kullanılan DÖ yönteminin model belirsizliğini elde etmek için epistemik belirsizlik tahmin edilir. Model belirsizliğinin ölçülerinden biri, tahmin dağılımının tahmin entropisi olan H değeridir: Eş. 4’te UQ gösterilmektedir [31].

$$H(\hat{y}^*|x^*, X, Y) = - \sum_c p(\hat{y}^* = c|x^*, X, Y) \log p(\hat{y}^* = c|x^*, X, Y) \quad (4)$$

Birinci belirsizlik ölçüsü; elde edilen belirsizlik değerlerinin aralığı, veri kümeleri, ağ mimarileri, MCS örnekleme sayısı vb. değişkenlere 392

bağlıdır. Bu nedenle, sonuçları rapor etmek ve çeşitli kümeler ile konfigürasyonlar arasında karşılaştırmayı kolaylaştırmak için tahmin belirsizlik normalleştirilir. Analiz, akciğer kanserine ait BT görüntüsünün sınıflandırmasında varyasyonel ağırlık temelli iki belirsizlik ölçüsünün (Tahmin Edici Entropi (PH) ve Bayesian Aktif Öğrenme Anlaşmazlığı) karşılaştırmasını içermektedir. İkinci belirsizlik ölçüsü olan Bayesian Aktif Öğrenme Anlaşmazlığı ise model aktivasyon fonksiyonu ile tahmin dağılımının entropisi arasındaki ilişkiye dayanmaktadır [31].

$$MI [y^*i, w|x^*i, X, Y] \approx H [y^*i|x^*i, X, Y] - E [H [y^*i|x^*iw]] \quad (5)$$

Eş. 5’da test, karşılıklı bilgiyi en üst düzeye çıkararak, modelin ortalama olarak belirsiz olduğu noktadır. Ancak yüksek güvenlikle hatalı tahminler üreten model parametreleri bulunmaktadır. Bu durum, softmax katmanına (logitlere) girişte yüksek varyansa sahip noktalara eşdeğerdir. Bu nedenle modelden her bir stokastik ileri geçiş, farklı bir sınıf için en yüksek olasılığa sahip olacaktır. Uyuşmazlık Yoluyla Bayesian Sınır Ağları (UYBNN) modelinin epistemik belirsizliği ölçmesi beklenmektedir. Bu sebeple, aleatorik belirsizlik mevcutsa yüksek bir değere döndürülmeyecektir. Bayes modelindeki UQ (Belirsizlik Analizi) yöntemi, DÖ (Dijital Örnekleme) modelinin tahmin edici olasılıkları ile ilişkili epistemik belirsizliği, test süresi boyunca düşük ağırlıklarla ölçülmüştür. Kernel Yoğunluğu Tahmini kullanarak Gauss Çekirdeği yöntemi, MCS (Monte Carlo Simülasyonu) ileri besleme numunelerinde bir düşüş oranı olan p varyasyonunu temsil etmektedir. Ayrıca, doğru ve yanlış tahminlerle gruplandırılan BT (Bilgisayarlı Tomografi) test görüntülerinin çıktı arka plan dağılımlarını çizmek için kullanılır. Wasserstein Mesafesi (WM) son zamanlarda MÖ (Matematiksel Optimizasyon), stokastik programlama ve görüntü işleme alanlarında giderek daha önemli bir rol oynamaktadır. Sinkhorn mesafesi gibi, hesaplama karmaşıklığı yüksek yaklaşık veya düzenlenmiş varyasyonlara yol açan sorunları çözmek için çalışmalar devam etmektedir. Ancak, büyük düzenlilik parametreleri, birkaç önemli MÖ uygulamasında performans düşürecektir. Bu nedenle, mevcut algoritmaların sayısal kararlılık sorunları nedeniyle küçük düzenlilik parametreleri başarısız olacaktır. İki histogram arasındaki farkı belirlemek için WM değeri 1 olarak alınır. WM, bir metrik olarak, iki histogram arasındaki mesafeyi aslına sadık bir şekilde ölçülebilir ve birçok noktasal mesafeye kıyasla daha doğru sonuçlar verebilir. Örneğin WM, herhangi bir çift delta işlevi arasındaki mesafeyi ayrık desteklerle ayırt edebilir. Histogramlar düzleştirilmedikçe mesafelerin çoğu birbiriyle bağlantılı olmayan iki histogram arasındaki mesafeyi ayıramaz. Öte yandan WM, tahmin edilen belirsizliğin gerçek hatalarla ne kadar ilişkili olduğunu ölçmek için kullanılır. WM Eş. 6’da gösterilmektedir.

$$W(P_r, P_g) = \inf_{\gamma \sim \Pi(P_r, P_g)} \mathbb{E}_{(x,y) \sim \gamma} [\|x - y\|] \quad (6)$$

Burada  $P_r$ , gerçek veri dağıtımı için WM ve oluşturulan veri dağıtımı  $P_g$  ise matematiksel olarak herhangi bir taşıma planı için en büyük alt sınır (en düşük) olarak ifade edilmektedir.

- Asimetrik Maliyet Fonksiyonu

Bir hasta için akciğer kanseri olmadığı hâlde akciğer kanseri olduğuna dair bir teşhis (yanlış pozitif sonuç) koymanın maliyeti, hasta akciğer kanseri olduğu hâlde akciğer kanseri olmadığına dair bir teşhis (yanlış negatif sonuç) koymanın maliyetinden daha düşüktür. Bu nedenle amaç, bazen yanlış pozitifler ortaya çıksa bile herhangi bir yanlış negatif sonuç tespitinden kaçınmaktır. Bu asimetrik maliyet, sınıf ağırlıkları gibi farklı etkileri olan faktörleri hesaba katan bir fayda fonksiyonu ile yakalanır. Bu fayda fonksiyonu, gerçek posteriora en yakın tahminleri sağlamaktadır. Hata yapmanın bu asimetrik maliyetini ele almak için beklenen faydayı maksimize edecek bir



veya fonksiyonu ( $\alpha$ ) tanımlanır. Dolayısıyla ağırlıklı çapraz entropi kaybı işlevi şu şekilde ortaya çıkar:

$$L \approx \frac{1}{C} \sum_{c=1}^C \alpha_c * p(\hat{y} = c_i | w, \hat{x}) \quad (7)$$

Eş. 7'de  $\alpha_c$  çapraz entropi kaybındaki her bir sınıfa karşılık gelen ağırlık olup ağırlıklar üzerinden gerçek posterior'a yaklaşım için en uygun tahminleri dikte eder. En yüksek ağırlık kesin gerçekler düşük belirsizlikle doğru olduğunda, akciğer kanseri olmayan fakat enfekte olduğu için yanlış teşhis edilen bir görüntüye atanır.

Bu çalışmanın amacı, birden fazla ileri geçişin ek hesaplama maliyetini ve insan-makine kombinasyonunun tahmin doğruluğunu, belirsizliğe duyarlı yönlendirme ile iyileştirmektir. Her bir modelin doğru ve yanlış tahmin oranları verilmiş ve F-ROC ile gösterilen şekilde, önerilen sistemin bir taramasında duyarlılığa karşı YP oranı belirtilmiştir. Test aşaması, sigmoid fonksiyonunun uygulanmasından önce ve -2 değerinde eşik algılama olasılığı ile gerçekleştirilmiştir. Önerilen sistem, daha az parametre kullanarak 4K-ESA sayesinde elde edilen sonuçlara benzer sonuçlar vermiştir. Sonuçlar, MCS kullanımının daha iyi sonuçlar sağladığını göstermektedir.

### 2.3. Performans Metrikleri (Performance Metrics)

Akciğer nodüllerini tespit etmek ve sınıflandırmak için geliştirilen DÖ algoritmalarının performansını analiz etmek için farklı ölçütler kullanılmaktadır. Matthews Correlation Coefficient (MCC), özgüllük (specificity), doğruluk (accuracy), duyarlılık (sensitivity), F2-skoru, F1-skoru, İşlem Karakteristik Eğrisi (Receiver Operating Characteristic, ROC) eğrisi ve ROC eğrisinin altındaki alan olan AUC (Area Under Curve) değerlerinden yararlanılır. Hata matrisi, MÖ'de kullanılan sınıflandırma modellerinin performansını değerlendirmek için hedef niteliğe ait tahmin değerleri ve gerçek değerlerin karşılaştırıldığı özel bir tablodur. Matrisin her satırı, tahmin edilen bir sınıftaki örnekleri temsil ederken her sütun gerçek bir sınıftaki örneklere karşılık gelmektedir. Aşağıda model performansını değerlendirmek için kullanılan metrikler gösterilmektedir.

**Doğru Pozitif (DP):** BT görüntüsünde nodül bulunur ve nodül ile BT görüntüsü olarak sınıflandırılır [31].

**Yanlış Pozitif (YP):** BT görüntüsünde nodül yoktur ve nodül ile BT görüntüsü olarak sınıflandırılır [31].

**Doğru Negatif (DN):** BT görüntüsünde nodül yoktur ve BT görüntüsü olarak sınıflandırılır [31].

**Yanlış Negatif (YN):** BT görüntüsünde nodül bulunur, nodül olmayan BT görüntüsü olarak sınıflandırılır [31]. Geliştirilen modeller, yukarıda gösterilen dört değerlendirmeden en az birine sahip olmalıdır. Modelin yüksek doğruluk ve iyi performans göstermesi buna bağlıdır. Örneğin, ROC eğrileri farklı veri kümeleri kullanılarak oluşturulursa veri kümelerinin zorluk seviyeleri aynı kabul edilir. Bilgisayarda medikal görüntü analizi algoritmasıyla oluşturulan bir ROC eğrisi, genellikle 10 veya 100 değerleri arasındadır. İşlem için (DP<sub>i</sub>, YP<sub>i</sub>), i=1,...,k ve i değerinin artışıyla TP değerlerinin azalmadığı varsayılarak yamuk kuralı uygulanır. AUC değeri, Eş. 8'de görüldüğü şekilde hesaplanır [31].

$$AUC = \sum_{i=1}^{k-1} (FP_{i+1} - FP_i)(TP_i + TP_{i+1})/2 \quad (8)$$

Eş. 1'de görüldüğü gibi hata matrisi, tahminin (1 veya 0) gerçek değer (1 veya 0) ile birleşimidir. Tahmin değeri ve doğru değerler, sınıflandırılmış ya da sınıflandırılmamış olma durumlarına göre dört

bölmeye ayrılır. Bunlar, performans ölçümü bölümünde detaylı bir şekilde açıklanmaktadır.

### 3. Deneysel Sonuçlar ve Tartışma (Experimental Results and Discussion)

Bu bölümde önerilen mimariler işlenmektedir. Mimari ile ilgili genek bilgiler ve yapılan deneyler ile tartışmalar yer almaktadır. Bölümde ayrıca deneysel sonuçlar ve analizlerde ele alınmıştır.

#### 3.1. Yapılan Deneyler (Experimental Setup)

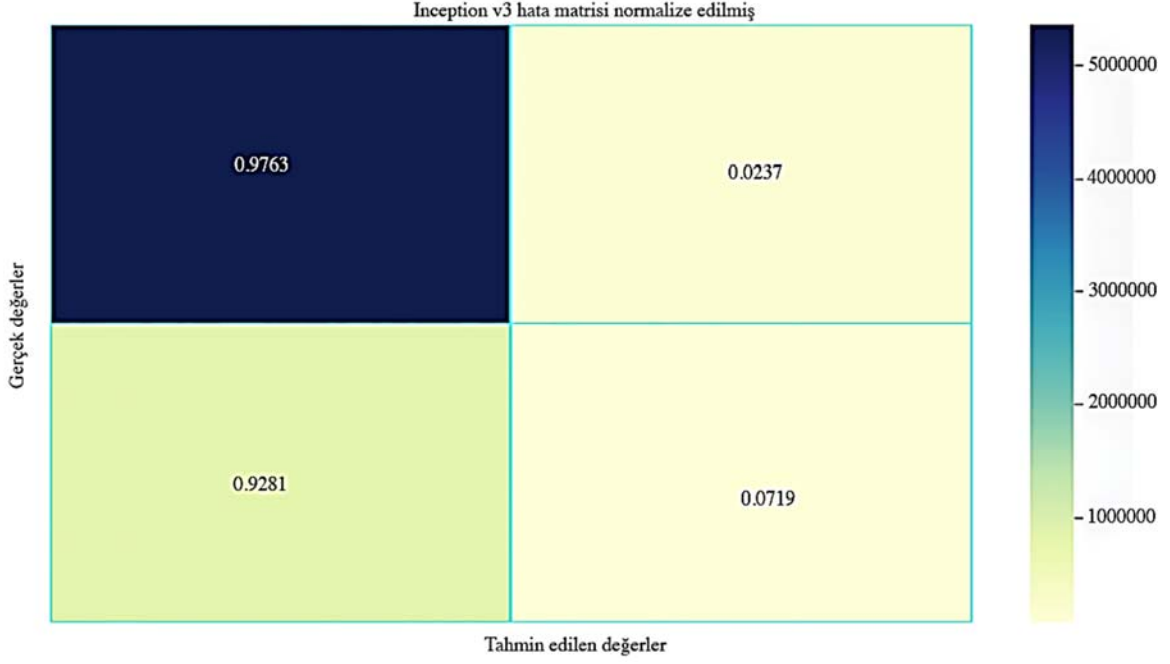
Çalışmadan elde edilen BT veri seti, BT cinsinden ve hastalık şüphesi taşıyan toplam 26.254 birey üzerinde gerçekleştirilen 73.118 çalışma neticesinde oluşturulan 13 TB veriden oluşmasına rağmen çalışmada hız ve modülerlik göz önünde bulundurularak sadece 124,8 GB veri ile çalışılmıştır. Veri setinin %80'i eğitim, %10'u test ve %10'u doğrulama için denemiştir. Bilgisayar konfigürasyonu MSI GT75 Titan 4K-071 GTX 1080 8G, 32GB 4TB SSD Windows 10 üzerinde işlem yapılarak eğitilmiştir. Programlama olarak da python programlama dili hem görüntü işleme hem de MÖ uygulamaları için kitaplıklar açısından zengindir.

Çalışmada elde edilen ROC eğrileri, Inception v3 modeli için eğitim ve test setindeki F1-skorunda doğru pozitif oranın yüksek olduğu gözlemlenmiştir. Şekil 3'te Inception v3 hata matrisi gösterilmektedir. DP 519538 iken YP 3302'de kalmıştır. Öte yandan YN 78083 değerinde kalırken DN 8975 olmuştur. Şekil 4'te VGG16 hata matrisi gösterimi ve Şekil 5 ise VGG16 ROC eğrisi gösterilmektedir. Ayrıca VGG16 ROC eğrisinde gösterilen en iyi oranın 0,87 olduğu gözlemlenmiştir. VGG16 performans sonuçlarına göre yanlış teşhis oranı düşük olmakla birlikte YPO ve YNO yüksek olduğu görülmektedir. Şekil 6'da ResNet50 hata matrisi, Şekil 7'de AlexNet hata matrisi gösterimi, son olarak Şekil 8'de AlexNet modelinin ROC performans gösterilmektedir. Şekil 9'da 4K-ESA hata matrisi sergilenmektedir. Şekilde başarı oranının yüksekliği göze çarpmaktadır. Şekil 10, 4K-ESA modelinin bir BT görüntüsünü sınıflandırma sürecini göstermektedir. Ayrıca şekilde bazı iyi huylu ve kötü huylu örneklerde kafa karıştırıcı özellikler ortaya çıkmıştır, kötü huylu örnekler en çok karışıklığa sebep olan örneklerdir. Şekil 11, 4K-ESA modelinin sonucunda ortaya çıkan sonuçları gösterilmektedir. Şekil 12, Belirsizliğe sahip örnek girdi gösterilmiştir.

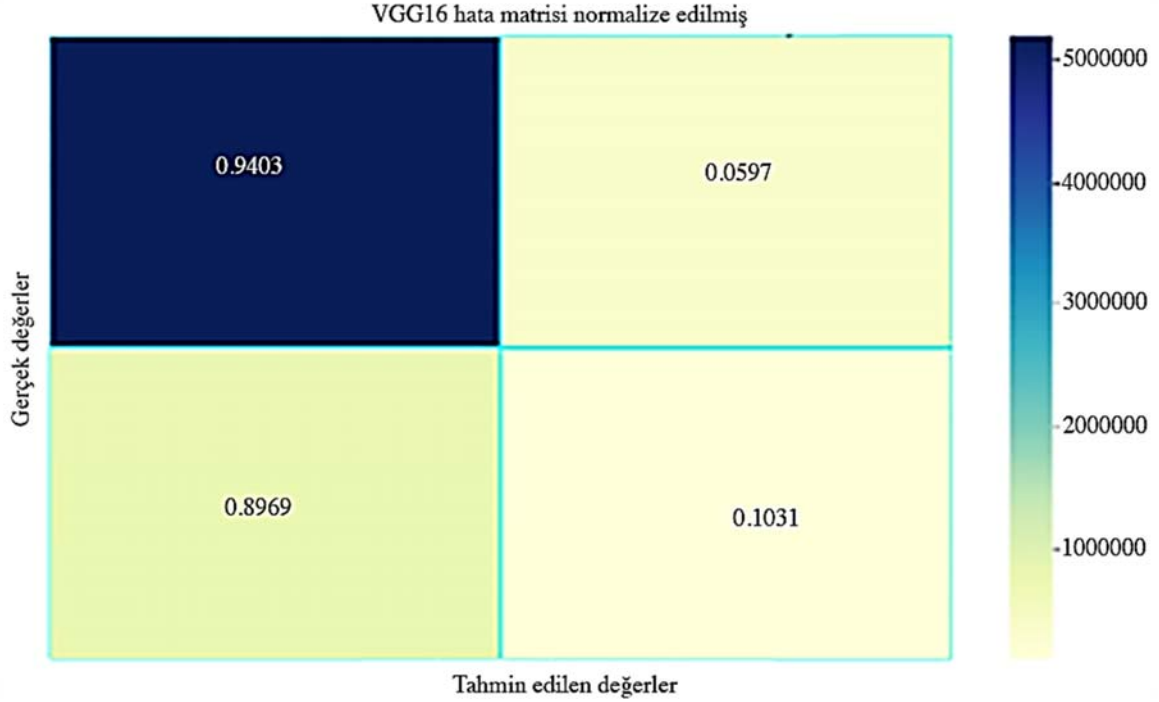
#### 3.1.1. Belirsizliğe duyarlı akciğer kanseri sınıflandırması (Uncertainty of quantification for lung cancer classification)

Tüm akciğer kanseri test görüntüleri için tahminler gerçekleştirmiş ve bu tahminler ilişkili tahmin belirsizliğine (PH) göre sıralanmıştır. Ardından daha fazla teşhis için çeşitli belirsizlik seviyelerine dayalı tahminlere başvurulmuştur. Kalan vakalar için tahminlerin doğruluğu (0,5'te eşik) ölçülmüştür. Tahmin doğruluğu, belirtilen görüntülerin fraksiyonu ile artmıştır. Tahmine dayalı doğruluğu hesaplamak için yalnızca referans verilmeyen görüntüler dikkate alınmıştır. Aynı davranış, artan model belirsizliği seviyeleri için yapılan tahmin doğruluğunda da gözlemlenmiştir. Burada dikkat edilmesi gereken noktalardan biri, yeni başlayan bir radyolog performansı için yalnızca DÖ modellerine güvenmenin genel bir tanıda daha doğru bir tahmin verdiğidir. Bununla birlikte, deneyimli bir radyolog için, en belirsiz örneklerin neredeyse %40'ını tespit edilebilirken geliştirilen model ise  $H_{norm} \gg 0,4$  olan örnekleri reddederek birleşik performans neredeyse %90'a ulaşır. Daha sonraki incelemeler için atıfta bulunulan kararların %2'sinden daha azı için örtüşmeyen iki senaryo %95 güven aralığına sahiptir.

**Belirsizliği ve Yorumlanabilirliği Görselleştirme:** DÖ modelleri genellikle "kara kutular" olmakla bilinir. Bu nedenle modellerin kesin ve yorumlanabilir olması, tahminlerdeki belirsizliğin iyi anlaşılması



Şekil 3. Inception v3 hata matrisi gösterimi (Inception v3 confusion matrix)



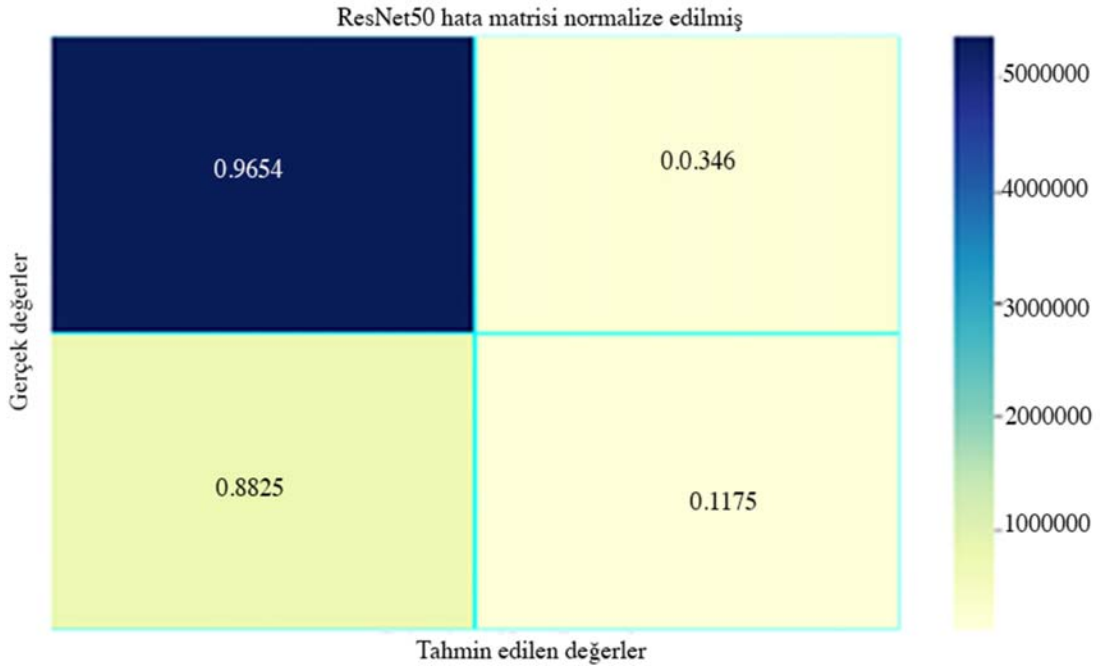
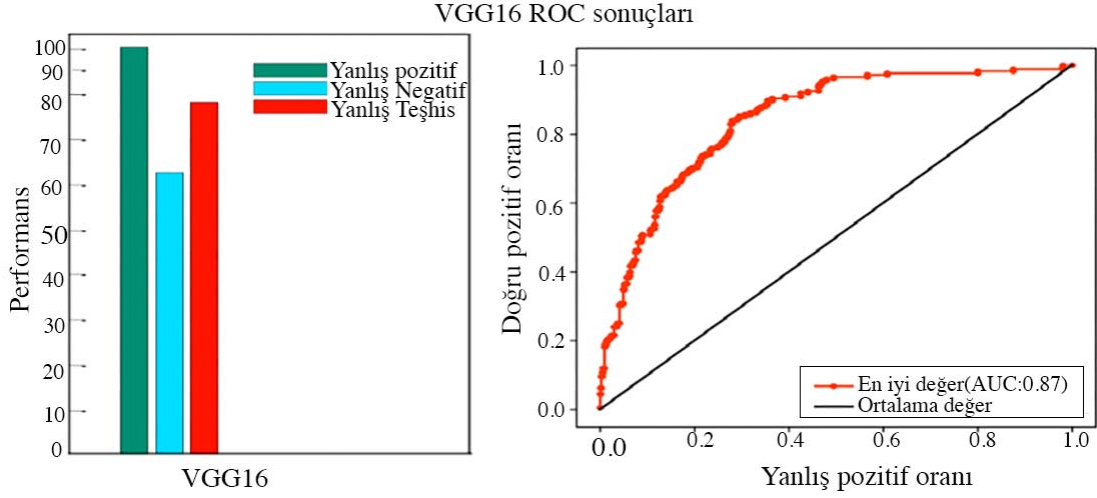
Şekil 4. VGG16 hata matrisi gösterimi (VGG16 confusion matrix)

gerekmektedir. Farklı özelliklerin görselleştirilmesinin yanı sıra güvenilir bir şekilde tahmin edilen belirsizlik, nokta tahminine ek bir içgörü olarak DÖ'de anlaşılma kolaylığını artıracak ve daha bilinçli bir karar verme süreci ile sonuçlanacaktır.

### 3.2. Deneysel Sonuçlar ve Model performans Değerlendirmesi (Results of the Models and Comparing Performance)

Mevcut verilere dayanarak Yüksek Doğruluk, DVM ve KNN sınıflandırıcı kullanılarak %100 sonuçla [32] söylenebilse de aslında

bu çalışma Yanlış Pozitif Oranından dolayı diğerlerine göre başarısız sayılabilir. Ayrıca PCA kullanılarak özellik seçimi ve görüntü olmayan veri kümesi için ve genetik algoritma için ise [32,33] KNN sınıflandırıcı kullanılarak %98,19 doğruluk elde edilmiştir. DÖ yöntemleri kullanılarak gerçekleştirilen çalışmada; DT, RF, Sınıflandırma ve Regresyon Ağaçları (CART), Radyal Temel Fonksiyonu (RBF) yöntemleri kullanılarak %97,04 doğruluk [35] elde edilmiştir. Tablo 4'te gösterildiği gibi, yüksek oranda yanlış negatif bulunmuş olan Maleki vd. [36] çalışması hariç, 4K-ESA modeli en iyi sonucu göstermiştir. 4K-ESA+BNN modelinin en iyi



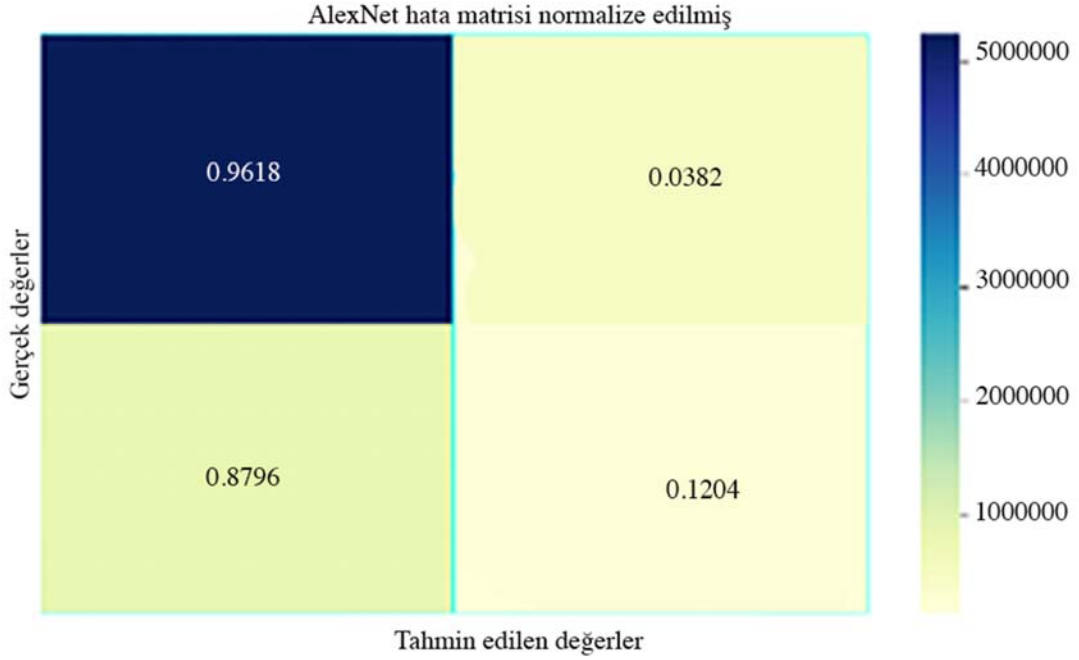
sınıflandırma performansının %98,75 doğruluk verdiği, ardından Hsu vd. [23] çalışmasının %98,19 doğruluk verdiği gözlenmektedir. Ayrıca Tablo 12'de görüldüğü üzere 4K-ESA diğer modellere nazaran daha başarılı sonuçlar vermektedir. Tablo 4'teki sonuçlar, epistemik belirsizliğin bir ölçüsü olarak olasılıkların entropisi ve tahmin hataları arasında güçlü bir korelasyon olduğunu ortaya koymaktadır ( $\rho = 0,98$ ). DropWeights [34] oranları tahmini entropi ve UYBNN'ye göre 0,1, 0,3 ve 0,5 korelasyonu ile ölçülmüştür. Buna göre tahmini entropi değerleri sırasıyla 0,984, 0,9814 ve 0,9843 iken, UYBNN'de bu değerler 0,8865, 0,8891 ve 0,9016 olarak elde edilmiştir. Yaptığımız deneyler, tahmin belirsizliğinin doğruluk ile ilişkili olduğunu ve böylece yanlış tahminlerin veya bilinmeyen durumların tanımlanmasının mümkün olduğunu göstermektedir.

Gerçek hata, tahmini değerler ile gerçek değerler arasındaki farktır. Tahmine dayalı belirsizliğin kalitesini değerlendirmek için Tahmine Dayalı Entropi (PH) ve UYBNN arasındaki Spearman [37,38]

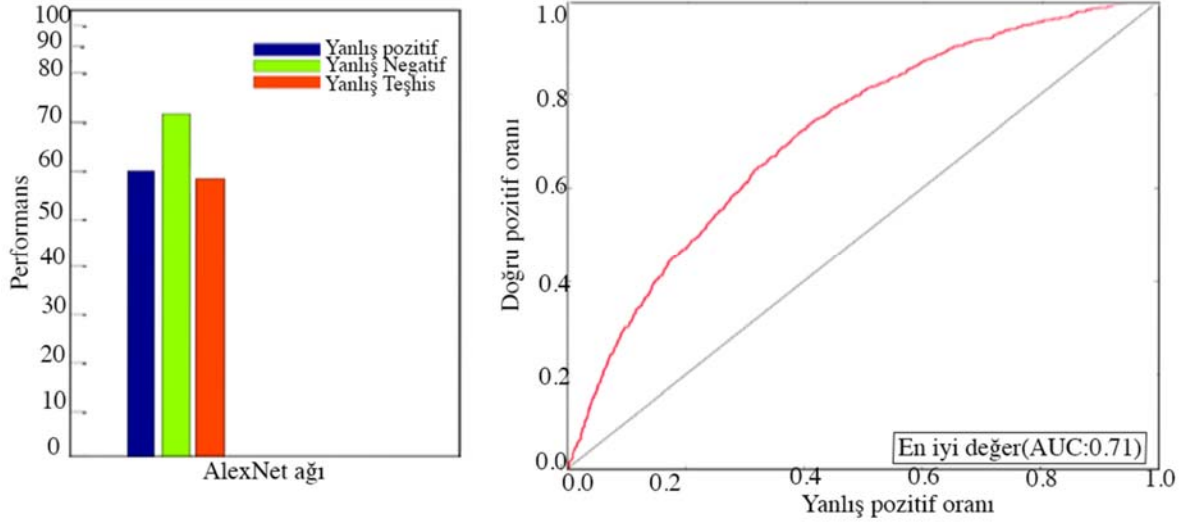
korelasyon katsayısından yararlanılmıştır. Tahmine dayalı doğruluk şu şekilde ölçülmüştür:

### 3.2.1. Deney sonucu ve modellerin karşılaştırması (Experiment results and comparison of the models)

Bu çalışmada Bayesian DÖ sınıflandırıcısı, model belirsizliğini tahmin etmek için akciğer kanseri BT görüntülerinde aktarım öğrenme yöntemi kullanılarak eğitilmiştir. Çalışma, model belirsizliği ile tahminin doğruluğu arasında güçlü bir ilişki olduğunu göstermiştir. DÖ modelindeki tahmini belirsizlik, radyologları yanlış tahminler konusunda uyarabilen daha güvenilir tahmin sağlamakta ve bu da DÖ'nün hastalık tespitinde klinik uygulamaya kabulünü artırmaktadır. BNN tabanlı sınıflandırmayla veri seti ve tedavi yanıtları ile ilişkili çalışmalar, görüntüleme belirteçleri hakkında iç görüleri ve akciğer kanseri için gelişmiş tanı ve tedaviye yönelik bulguları daha da çok ortaya çıkarmaktadır. Tablo 5'te akciğer kanseri



Şekil 7. AlexNet hata matrisi gösterimi (AlexNet confusion matrix)



Şekil 8. AlexNet modelinin ROC performans gösterimi (AlexNet ROC performance)

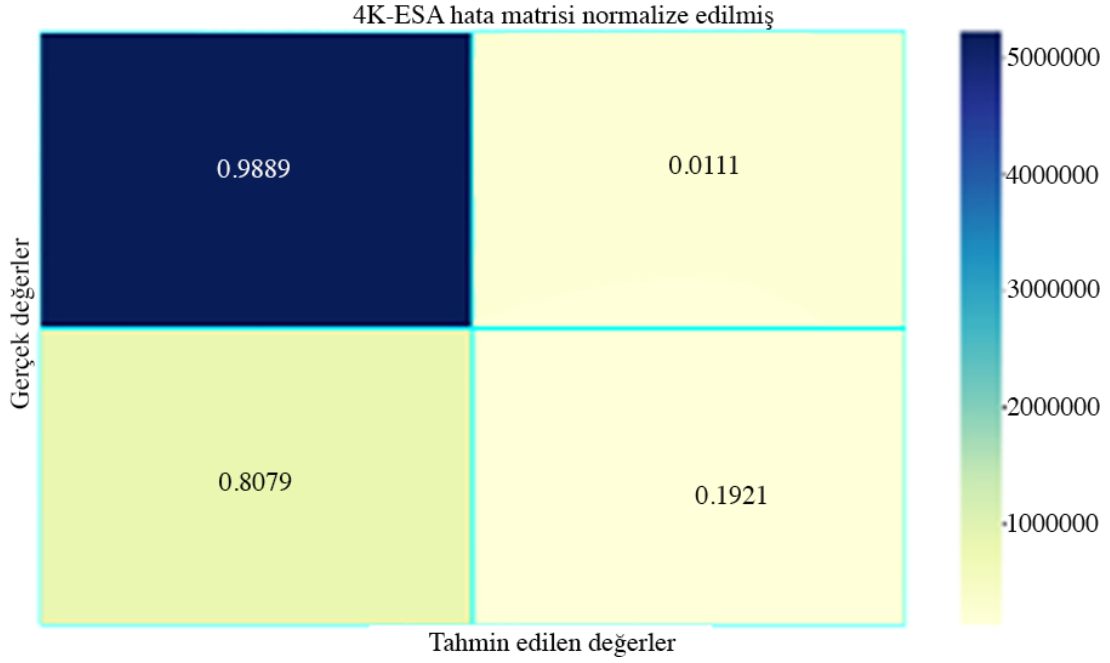
teşhisi için kullanılan metotların performansı gösterilmektedir. En iyi performans gösteren ağlar, AUC açısından aynı veri kümesi üzerinde değerlendirildiğinde diğer tüm çalışmalardan marjinal olarak daha yüksek performans elde etmekte ve bunları geride bırakmaktadır.

Tablo 5'te akciğer nodülü saptama ve sınıflandırma modellerinin değerlendirilmesi için kullanılan farklı ilişkilerden bahsedilmektedir. Tablo 5'te UQ (belirsizlik) uygulamadan elde edilen ölçümlerde ResNet50 [39-41] modelinin performansı ve yapılan deneylerden elde edilen değerler Tablo 5'te gösterilmektedir. Otomatik akciğer nodülü saptama ve sınıflandırma teknikleri, çok sayıda yanlış pozitif sonuçla yüksek algılama yeteneklerine sahiptir.

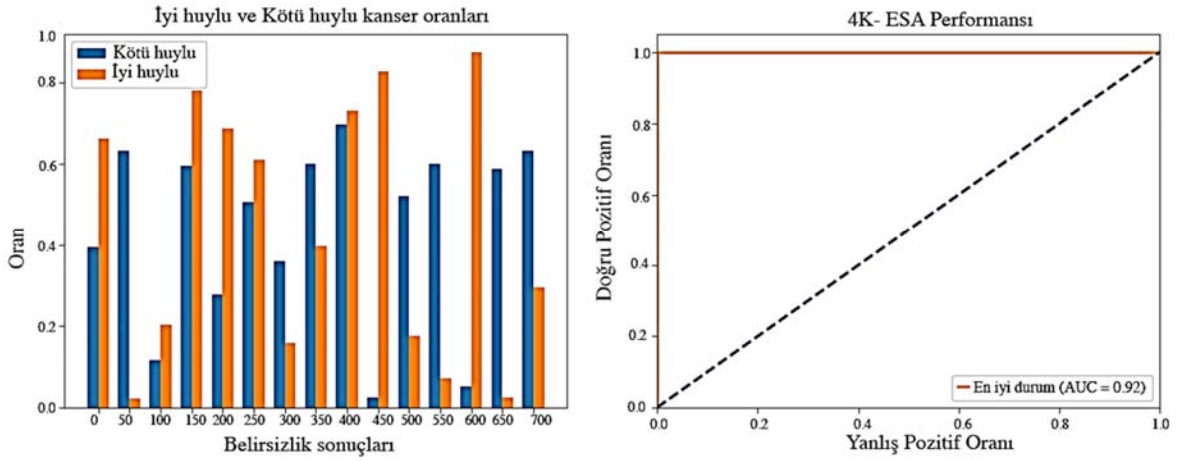
Yukarıdaki şekilde gösterilen mimaride dört evrişimli katman bulunmaktadır. Buna ek olarak global bir ortalama havuzu (GOH), iki TBK katmanı ve softmax ile çıktı katmanı bulunmaktadır. GOH,

klasik ESA'larda TBK'yi değiştirmek için tasarlanmış bir havuzlama işlemidir. Buradaki amaç, son Çok Katmanlı Perseptron (ÇKP) katmanındaki sınıflandırma görevine ait tüm ilgili kategoriler için bir özellik haritası oluşturmaktır. GOH, havuzlama amacıyla kullanılır ancak havuzlamadan daha etkilidir. Şekil 13'te 4K-ESA metodu performans sonuçları sergilenmektedir. Yukarıdaki şekilde gösterilen sonuçla 4K-ESA, akciğer kanseri teşhisi konusunda iyi bir sonuç elde etmiştir. Farklı görüntü özellikleri için farklı ağırlıklara göre farklı evrişim çekirdeklerinde, bir evrişim çekirdeği tüm evrişim sürecinde parametreleri paylaşmasına karşın, uygulanan MCS sonucunda modelin ezberden uzak başarısı ortaya konmuştur.

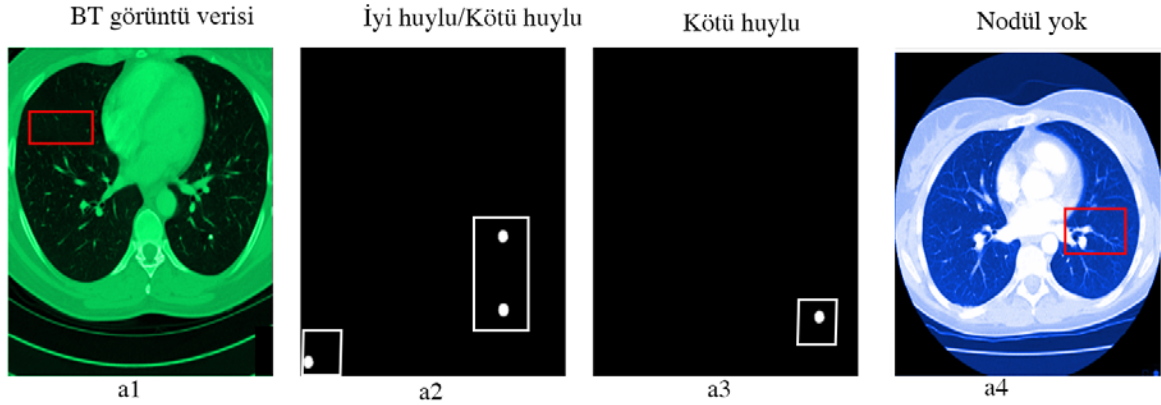
Yukarıdaki şekilde epok sayısı düşüken doğruluk az, kayıp ise fazladır. Epok sayısı yüksek olduğunda da yine aynı sonuç ortaya çıkmaktadır. Dengeli epok sayısı doğruluğu artırırken kayıp oranını en aza indirmektedir.



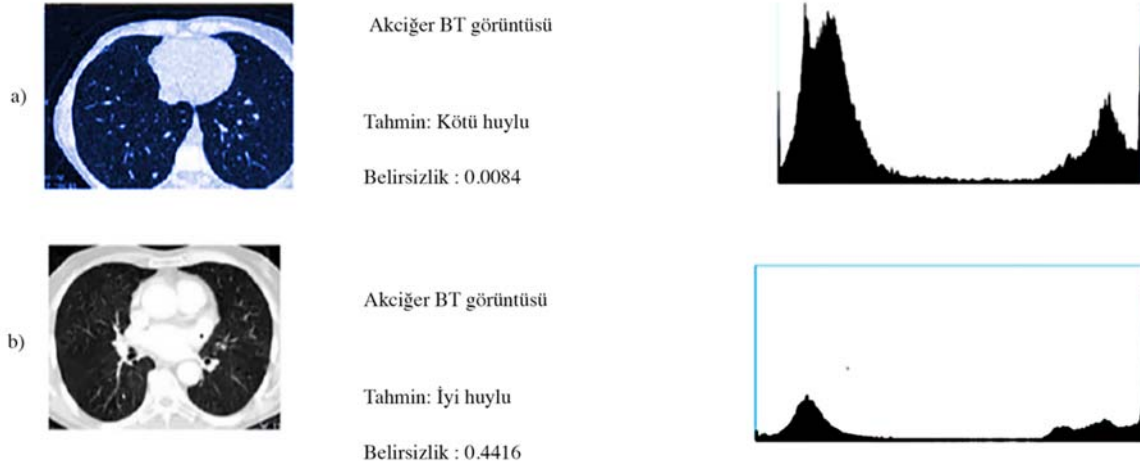
Şekil 9. 4K-ESA hata matrisi gösterimi (4K-ESA confusion matrix)



Şekil 10. 4-K ESA ROC eğrisi sonuçları (4K-ESA ROC performance)



Şekil 11. 4K-ESA modelinin sonucunda ortaya çıkan sonuçlar (Diagnosis of Lung cancer with 4K-ESA model)



Şekil 12. Belirsizliğe sahip örnek girdi (The result of and output done with uncertainty of quantification)

Tablo 4. Belirsizlik ve Tahmini Doğruluk Arasındaki İlişki (Relationship Between Uncertainty and Predicted values)

Spearmans'ın Korelasyonu	Tahmini Entropi	UYBNN
DropWeights oranı: 0,5	0,9841	0,8865
DropWeights oranı: 0,3	0,9814	0,8891
DropWeights oranı: 0,1	0,9843	0,9016

Tablo 5. Belirsizlik Ölçeği Olmadan Elde Edilen Sonuçlar (Results without uncertainty of quantification)

Modeller	MCC	F2-skoru	Doğruluk	Özgüllük	Duyarlılık	F-skoru
UQ Eklendenen Yapılan Deneysel Sonuçları						
ResNet50	0,943	0,794	0,965	0,799	0,982	0,974
AlexNet	0,962	0,854	0,837	0,781	0,854	0,911
Inception v3	0,976	0,733	0,849	0,737	0,865	0,917
VGG16	0,941	0,801	0,826	0,795	0,869	0,903
4K-ESA	0,988	0,816	0,971	0,825	0,974	0,962
UQ Eklenecek Yapılan Deneysel Sonuçları						
Modeller	MCC	F2-skoru	Doğruluk	Özgüllük	Duyarlılık	F1-skoru
BNN ResNet50	0,91	0,87	0,89	0,81	0,85	0,81
BNN AlexNet	0,89	0,81	0,86	0,78	0,83	0,85
BNN Inception v3	0,94	0,89	0,91	0,79	0,94	0,88
BNN VGG16	0,83	0,85	0,77	0,74	0,84	0,90
BNN 4K-ESA	0,98	0,88	0,98	0,94	0,97	0,96

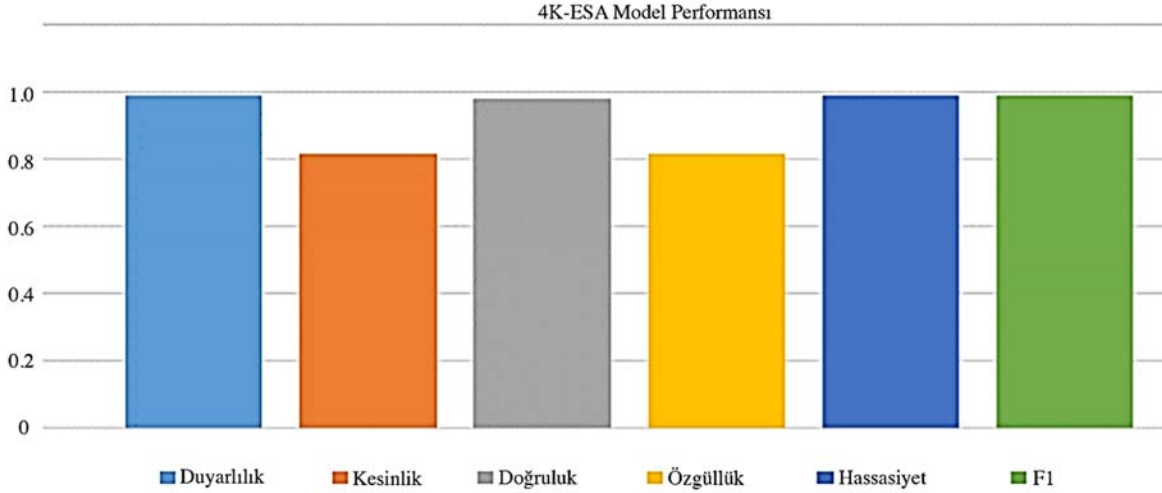
### 3.2.2. Tartışma ve karşılaştırma analizi (Discussion and comparative analysis)

Bu çalışmada, akciğer hastalıklarının tespiti ve teşhisi için dört geleneksel DÖ ve geliştirilen 4K-ESA mimarisi önerilmiştir. Bu mimariler, en yaygın göğüs hastalıkları arasında sınıflandırma yapmak için kullanılmaktadır. Önerilen farklı modeller; doğruluk, hassasiyet, kesinlik, F1 puanı ve AUC'ye göre karşılaştırılmıştır. Literatürde incelenen çalışmaların çoğunda, BT tarama ve röntgen görüntüleri kullanılmış veya görüntü olmadan çalışılmıştır. Her iki durumda da akciğer kanseri erken tanı süreci takip eden adımlardan oluşmaktadır: Ön işleme, bazı veri kümesi teknikleri hem görüntü hem de görüntü olmayan verileri içermektedir. Görüntü işleme için görüntü giderme, eşik ve normalizasyon uygulanmaktadır. Görüntü bölütleme işlemi için bölüt, BT taramalarından ve X-ray'den hem benzer hem de farklı görüntüleri içermektedir. İlgi Bölgesi (ROI), Gabor filtresi, görüntü histogramı gibi çeşitli bölütleme yöntemleri, literatürde daha yüksek doğruluk için kullanılan tekniklerdir. Öznitelik seçimi için Temel Bileşen Analizi (PCA), Genetik Algoritmalar (GA), Bilgi Kazanç Sıralamasına Dayalı Nitelik Seçimi

(ICFA), Özellik Seçimi için Hızlı Korelasyon Tabanlı Filtre (FCBF), Monte Carlo gibi yöntemlerden yararlanılmaktadır.

Deneylemlerin sonucunda Belirsizlik Tahmini Entropisi ve UYBNN oranları %90 olarak bulunmuştur. Bu oran sayesinde akciğer hastalıklarının tespiti için DÖ'nün klinik uygulamalar açısından kabulünün sağlanabileceği düşünülmektedir.

Shen, Han, Aberle, Bui ve Hsu (2019) tarafından nodülün malignite olup olmadığının teşhis edilebilmesini ve nodülde bulunan anlamsal özelliklere dair tahminlerin ortaya çıkarılabilmesini amaçlayan bir ağ (Hiyerarşik Anlamsal ESA) geliştirilmiştir. Fakat geliştirilen model eğitilirken yalnızca LIDC-IDRI veri kümesi kullanılarak geleneksel ESA'ya bağlı kalınması sebebiyle model başarısızlığa uğramıştır. Geliştirilen modelde UQ'nin kullanılmaması ve kullanılan veri kümesinin çok küçük olması ağ öğrenimini sınırlayabilecek ölçekte kalmasına neden olmuştur. Dolayısıyla bu çalışma kapsamında geliştirilen 4K-ESA ve BNN 4K-ESA modellerinin Shen vd. tarafından geliştirilen Hiyerarşik Anlamsal ESA modeline göre daha avantajlı olduğu söylenebilir.



**Şekil 13.** 4K-ESA metodu performans sonuçları (Results of 4K-ESA model)

#### 4. Sonuç ve Tartışmalar (Results and Discussions)

4K-ESA algoritması, akciğer kanseri BT veri kümesi üzerinde uygulanmadan önce, LIDC-IDRI veri kümeleri üzerinde diğer geleneksel MÖ algoritmaları ile uygulanıp karşılaştırılmıştır [31].

Sınıflandırma adımında veriler, iki aşamalı olarak gerek geleneksel modellere gerekse geliştirilen modele sırasıyla mini-paket hâlinde verilmiştir. Geleneksel modeller için başlangıç olarak 128 mini-paket boyutu ile veriler beslendikten sonra, dropout değerleri 0,2 seçilerek farklı değerler ile bu işleme devam edilmiştir. Daha sonra öğrenme oranları 0,01, 0,002 ve 0,0001 şeklinde değişik değerler ile denenerek aktivasyon fonksiyonları tanh, sigmoid ve ReLU olarak girilmiştir. Bu çalışma kapsamında geliştirilen model olan 4K-ESA modelinde ise seyreltme olarak DropWeights, aktivasyon fonksiyonu olarak da softmax denenmiştir. Böylece modellere girdi olarak verilen BT görüntülerinin kanserli olup olmadığı, kanserli ise bunun iyi mi yoksa kötü huylu mu oldukları belirlenmiştir.

UQ'nin uygulanması adımında; öncelikle geleneksel modellere UYBNN, geliştirilen modele ise BNN, UQ ve MCS eklenerek modeller çalıştırılmıştır. Geleneksel modellere eklenen UYBNN ile anlamlı, güvenli sonuçlar elde edilmiştir. Fakat makine konfigürasyonunun sınırlı olması sebebiyle parametre sayılarının yüksek oluşu 4K-ESA'ya göre geleneksel modellerin daha düşük performans sergilemelerine neden olmuştur. Dolayısıyla çalıştırılan modeller arasında en iyi sonuç veren modelin, parametre sayıları geleneksel modellere göre daha düşük olan 4K-ESA modeli olduğu görülmüştür. Bu sonucun ortaya çıkmasında, modele uygulanan MCS ve DropWeights tekniklerinin payı oldukça büyüktür.

Karar verme adımında ise; UQ eklenmeden yapılan deneylerin sonucu Tablo 5'te gösterilmektedir. Bu çalışma kapsamında geliştirilen 4K-ESA ve BNN 4K-ESA modellerinin akciğer kanserinin teşhisi aşamasında meydana gelen güven sorununu, teşhis aşamasında DÖ algoritmalarının rahatlıkla kullanabilmesiyle ortadan kaldıracığı düşünülmektedir. Ayrıca radyologlar açısından oldukça önemli olan nodül tespiti ve akciğer BT görüntülerinin sınıflandırılması hususunda meydana gelen sorunların da geliştirilen bu yöntem sayesinde ortadan kaldırılacağı düşünülmektedir.

Bu çalışma kapsamında geliştirilen modellerin, gelecek çalışmalarda kanserli dokuların bulunduğu yerin tespit edilebilmesi açısından bir ön adım niteliğinde olduğu düşünülmektedir. Geliştirilen modellerin

tek sınırlı yanı, çalışmada yerel (ulusal) veri kaynağının bulunmamasıdır. Ulusal akciğer BT taraması bulunduğu çalışma kapsamında geliştirilen modeller söz konusu ulusal veriler üzerinde de denenebilir. Geliştirilen modellerin, diğer hastalık sınıflandırmaları ve farklı görüntüleme yöntemlerinde de kullanılabileceği öngörülmektedir.

#### Kaynaklar (References)

1. Vishnubalaji R., Shaath H., Al-Alwan M., Abdelalim E.M., & Alajez N.M., Reciprocal interplays between MicroRNAs and pluripotency transcription factors in dictating stemness features in human cancers. In Seminars in Cancer Biology. Academic Press, October, 2022.
2. Xie Y., Xia, Y., Zhang, J., Song, Y., Feng, D., Fulham, M., & Cai, W., Knowledge-based collaborative deep learning for benign-malignant lung nodule classification on chest CT. IEEE transactions on medical imaging, 38 (4), 991-1004, 2018.
3. Hsu C.H., Chen, X., Lin, W., Jiang, C., Zhang, Y., Hao, Z., Chung, Y.C., Effective multiple cancer disease diagnosis frameworks for improved healthcare using machine learning. Measurement, 175, 109145, 2021.
4. Gaber, M., Bakheet, R. A. E. A., Mahmoud, F. I., Risk factors of cancer and how to prevent. Sohag Medical Journal, 21 (2), 179-183, 2017.
5. Heigener, D. F., Reck, M., Advanced non-small cell lung cancer: the role of PD-L1 inhibitors. Journal of thoracic disease, 10 (13), S1468, 2018.
6. Salaken, S. M., Khosravi, A., Khatami, A., Nahavandi, S., Hosen, M. A., Lung cancer classification using deep learned features on low population dataset. In 2017 IEEE 30th Canadian Conference on Electrical and Computer Engineering (CCECE) 1-5, IEEE, 2017.
7. Bray, F., Ferlay, J., Soerjomataram, I., Siegel, R. L., Torre, L. A., Jemal, A., Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA: a cancer journal for clinicians, 68 (6), 394-424, 2018.
8. Li, C., Zhang, L., Meng, G., Wang, Q., Lv, X., Zhang, J., & Li, J., Circular RNAs: pivotal molecular regulators and novel diagnostic and prognostic biomarkers in non-small cell lung cancer. Journal of cancer research and clinical oncology, 145, 2875-2889, 2019.
9. Chang, C. H., Lin, C. H., Lane, H.Y., Machine learning and novel biomarkers for the diagnosis of Alzheimer's disease. International journal of molecular sciences, 22 (5), 2761, 2021.
10. Rawat, A., Singh, A. P., Kumar, G. T., Emerging role of EGFR and lung cancer treatments. MOJ Anat & Physiol, 5 (5), 281-282, 2018.
11. Kanageswari, S., Gladis, D., A systematic Review and Analysis: The Impact of Air Pollution on Lung Cancer. In 2022 6th International Conference on Intelligent Computing and Control Systems (ICICCS), 1790-1797, IEEE, 2022.
12. Viswanathan, V. S., Toro, P., Corredor, G., Mukhopadhyay, S., Madabhushi, A., The state of the art for artificial intelligence in lung digital pathology, The Journal of Pathology, 257 (4), 413-429, 2022.

13. Jabir, K., Raja, A. T., A Comprehensive Survey on Various Cancer Prediction Using Natural Language Processing Techniques. In 2022 8th International Conference on Advanced Computing and Communication Systems (ICACCS), 1, 1880-1884, IEEE, 2022.
14. Rastogi, A., Yadav, K., Mishra, A., Singh, M. S., Chaudhary, S., Manohar, R., & Parmar, A. S., Early diagnosis of lung cancer using magnetic nanoparticles-integrated systems. *Nanotechnology Reviews*, 11(1), 544-574, 2022.
15. Shivwanshi, R. R., Nirala, N., Enhancement and Segmentation Methods for Lung Cancer Detection System: A Review of a Retrospective Study. In 2021 3rd International Conference on Advances in Computing, Communication Control and Networking (ICAC3N), 1742-1748, IEEE, 2021.
16. Taher, F., Prakash, N., Shaffie, A., Soliman, A., El-Baz, A., An overview of lung cancer classification algorithms and their performances. *IAENG International Journal of Computer Science*, 48 (4), 2021.
17. Chiu, H. Y., Chao, H. S., Chen, Y. M., Application of artificial intelligence in lung cancer. *Cancers*, 14 (6), 1370, 2022.
18. Rout, I., Saha, M., Nayak, S., Mohanty, N., Baral, V., Comparative Study of Machine Learning Algorithms for Predicting Lung Cancer. In *Advances in Intelligent Computing and Communication: Proceedings of ICAC*, Springer Singapore, 2020, 381-389, 2021.
19. Yar, H., Abbas, N., Sadad, T., Iqbal, S., Lung nodule detection and classification using 2D and 3D convolution neural networks (CNNs). *Artificial Intelligence and Internet of Things*, 365-386, 2021.
20. Li, J., Wu, J., Zhao, Z., Zhang, Q., Shao, J., Wang, C., Li, W., Artificial intelligence-assisted decision making for prognosis and drug efficacy prediction in lung cancer patients: A narrative review. *Journal of Thoracic Disease*, 13 (12), 7021, 2021.
21. Silva, F., Pereira, T., Neves, I., Morgado, J., Freitas, C., Malafaia, M., Oliveira, H. P., Towards machine learning-aided lung cancer clinical routines: Approaches and open challenges, *Journal of Personalized Medicine*, 12 (3), 480, 2022.
22. Painuli, D., Bhardwaj, S., Recent advancement in cancer diagnosis using machine learning and deep learning techniques: A comprehensive review. *Computers in Biology and Medicine*, 105580, 2022.
23. Hsu, C. H., Chen, X., Lin, W., Jiang, C., Zhang, Y., Hao, Z., Chung, Y. C., Effective multiple cancer disease diagnosis frameworks for improved healthcare using machine learning. *Measurement*, 175, 109145, 2021.
24. Lopez-Ubeda, P., Diaz-Galiano, M. C., Martin-Noguerol, T., Luna, A., Urena-Lopez, L. A., Martin-Valdivia, M. T., Automatic medical protocol classification using machine learning approaches. *Computer Methods and Programs in Biomedicine*, 200, 105939, 2021.
25. Clark, K., Vendt, B., Smith, K., Freymann, J., Kirby, J., Koppel, P., & Prior, F. (2013). The Cancer Imaging Archive (TCIA): Maintaining and operating a public information repository. *Journal of digital imaging*, 26, 1045-1057.
26. Huang, Y., & Khan, S. M., Advances in AI and Machine Learning for Education Research. *Computational Psychometrics: New Methodologies for a New Generation of Digital Learning and Assessment: With Examples in R and Python*, 195-208, 2021.
27. Panayides, A. S., Amini, A., Filipovic, N. D., Sharma, A., Tsaftaris, S. A., Young, A., Pattichis, C. S., AI in medical imaging informatics: current challenges and future directions. *IEEE journal of biomedical and health informatics*, 24 (7), 1837-1857, 2020.
28. Zhang, G., Zhang, H., Yao, Y., Shen, Q., Attention-Guided Feature Extraction and Multiscale Feature Fusion 3D ResNet for Automated Pulmonary Nodule Detection. *IEEE Access*, 10, 61530-61543, 2022.
29. Yılmaz A., Diagnosing COVID-19 from X-Ray images with using multi-channel CNN architecture, *Journal of the Faculty of Engineering and Architecture of Gazi University*, 36 (4), 1761-1774, 2021.
30. Yang, H., Chen, L., Cheng, Z., Yang, M., Wang, J., Lin, C., ... Li, W., Deep learning-based six-type classifier for lung cancer and mimics from histopathological whole slide images: a retrospective study. *BMC medicine*, 19, 1-14, 2021.
31. Shmatko, A., Ghaffari Laleh, N., Gerstung, M., Kather, J. N., Artificial intelligence in histopathology: enhancing cancer research and clinical oncology. *Nature cancer*, 3 (9), 1026-1038, 2022.
32. Rooiantan, A., Mir, T. A., Wani, S. I., Hussain, K. K., Ahmed, B., Abraham, S., Akhtar, M. H., Early detection of lung cancer biomarkers through biosensor technology: A review, *Journal of pharmaceutical and biomedical analysis*, 164, 93-103, 2019.
33. Lv, W., Ye, H., Yuan, Z., Liu, X., Chen, X., Yang, W., Recent advances in electrochemiluminescence-based simultaneous detection of multiple targets. *TrAC Trends in Analytical Chemistry*, 123, 115767, 2020.
34. Gülcü, A., Zeki, K.u.ş., A Survey of Hyper-parameter Optimization Methods in Convolutional Neural Networks, *Gazi University Journal of Science Part C: Design and Technology*, 7 (2), 503-522, 2019.
35. Maleki, N., Zeinali, Y., Niaki, S., A k-NN method for lung cancer prognosis with the use of a genetic algorithm for feature selection. *Expert Systems with Applications*, 164, 113981, 2021.
36. Chandrakar, P. K., Shrivastava, A. K., Sahu, N., Design of a novel ensemble model of classification technique for gene-expression data of lung cancer with modified genetic algorithm. *EAI Endorsed Transactions on Pervasive Health and Technology*, 7 (25), e2-e2, 2021.
37. Yuan, F., Lu, L., Zou, Q., Analysis of gene expression profiles of lung cancer subtypes with machine learning algorithms. *Biochimica et Biophysica Acta (BBA)-Molecular Basis of Disease*, 1866 (8), 165822, 2020.
38. Singh, G.A.P., Gupta, P., Performance Analysis of Various Machine Learning-based Approaches for Detection and Classification of Lung Cancer in Humans. *Neural Computing and Applications*, 31 (10), 6863-6877, 2019.
39. Gürkahraman, K., Karakiş, R., Brain tumors classification with deep learning using data augmentation. *Journal of the Faculty of Engineering and Architecture of Gazi University*, 36 (2), 997-1011, 2021.
40. Akalın, F., Yumuşak, N., Classification of exon and intron regions obtained using digital signal processing techniques on the DNA genome sequencing with EfficientNetB7 architecture. *Journal of the Faculty of Engineering and Architecture of Gazi University*, 37 (3), 1355-1372, 2022.
41. Han, J., Xiao, N., Yang, W., Luo, S., Zhao, J., Qiang, Y., ... Zhao, J., MS-ResNet: disease-specific survival prediction using longitudinal CT images and clinical data. *International Journal of Computer Assisted Radiology and Surgery*, 17 (6), 1049-1057, 2022.
42. Yuan, Q., Cai, T., Hong, C., Du, M., Johnson, B. E., Lanuti, M., ... Christiani, D. C., Performance of a machine learning algorithm using electronic health record data to identify and estimate survival in a longitudinal cohort of patients with lung cancer, *JAMA Network Open*, 4 (7), e2114723-e2114723, 2021.