



**T.C.
BİLECİK ŞEYH EDEBALI ÜNİVERSİTESİ
BİLİMSEL ARAŞTIRMA PROJESİ
SONUÇ RAPORU**

**HİSTOPATOLOJİK GÖRÜNTÜLERDE DERİN SİNİR
AĞLARI İLE GÖĞÜS KANSERİ TANIMA VE
SINIFLANDIRMA SİSTEMİ**

PROJE YÜRÜTÜCÜSÜ : DR. ÖĞR. ÜYESİ EMRE DANDIL
PROJE NUMARASI : 2019-01.BŞEÜ.25-02
ARAŞTIRMACILAR : DR. ÖĞR. ÜYESİ SÜLEYMAN UZUN
DR. ÖĞR. ÜYESİ KERİM KÜRŞAT ÇEVİK
ÖĞR. GÖR. ALİ OSMAN SELVİ
ÖĞR. GÖR. MEHMET SÜLEYMAN YILDIRIM

BAŞLAMA TARİHİ : 12.06.2019
**BİTİŞ TARİHİ : 11.12.2020 (12.06.2021 tarihine kadar olan 6 ay uzatma
kullanılmadı)**

Sonuç Dönemi Tarihi : 11.12.2020

**BİLECİK ŞEYH EDEBALI ÜNİVERSİTESİ
BİLECİK 2020**

ÖZET

Günümüzde kadınlar arasında en sık görülen kanser türü meme kanseridir ve akciğer kanserinden sonra oldukça yüksek bir ölüm oranına sahip olarak ikinci sırada yer alır. Geç tespit durumunda ise meme kanseri tedavisi oldukça zor bir duruma gelmektedir. Meme kanserinin tespiti için çeşitli yöntemler bulunmasına karşın halen yardımcı tespit ve tedavi yöntemlerine olan ihtiyaç önem arz etmektedir. Bu projede, histopatolojik görüntülerde zamanla bazal benzeri meme tümörlerinin gelişimini incelemek ve bazal benzeri meme kanseri tespiti için histopatolojik görüntüleri analiz etmek için otomatik bir sistem geliştirmiştir. Yapılan çalışmada ilk önce tümörlü ve tümörlü olmayan bölgeler arasında doğru sınıflandırmayı destekleyen kullanışlı özelliklerin çıkartılması sağlanmıştır. Sonraki çalışma adımlarında ise, gürbüz piksel tabanlı ve parçalı tabanlı segmentasyon elde etmek amacıyla da derin öğrenme temelli teknikler kullanarak sınıflandırma işlemi gerçekleştirilmiştir. Histopatolojik görüntüler üzerinde göğüs tümör alt tiplerinin sınırlı tedavi olanaklarından dolayı araştırmacılar tarafından daha sık çalışılmaktadır. Bazal benzeri tümörler ayırt edici morfolojik, genetik, immünofenotipik ve klinik özellikler ile karakterize edilmesine rağmen bu tümörün alt tipinin tanımlanması konusunda kabul görmüş bir fikir birliği olmadığı gibi sistematik olarak sınıflandırılması ile ilgili bir yol tanımlanmamıştır. Sınıflandırma için immünohistokimyasal belirteçler kullanılmakta fakat bu belirteçlerin başarımları ise %60 seviyelerinde kalmaktadır. Projende histopatolojik görüntülerden elde edilen özellikler birleştirilerek, DNN (Deep Neural Networks, CNN) ve BoW (Bag of Words) sınıflandırıcıları ile sistemin başarımlarının artırılması sağlanmıştır.

TEŐEKKÜR

Bu bilimsel arařtırma proje alıřması Bilecik Őeyh Edebalı Őniversitesi Bilimsel Arařtırmalar Projeleri Koordinatörlüğü tarafından **2019-01.BŐEÜ.25-02** nolu proje ile desteklenmiřtir. Proje ekibi katkılarından Bilecik Őeyh Edebalı Őniversitesi Bilimsel Arařtırmalar Projeleri Koordinatörlüğü' ne teőekkür etmektedir.

İÇİNDEKİLER

ÖZET.....	2
TEŞEKKÜR.....	3
İÇİNDEKİLER.....	4
ŞEKİLLER DİZİNİ.....	5
ÇİZELGELER DİZİNİ	7
1. GİRİŞ.....	8
2. LİTERATÜR ARAŞTIRMALARI VE ÖNCEKİ ÇALIŞMALARIN DEĞERLENDİRİLMESİ.....	11
3. I. ARA RAPOR DÖNEMİNDE YAPILAN ÇALIŞMALAR.....	16
3.1. Açık Verisetleri	16
3.2. Proje Kapsamında Sağlanan Materyal Teminleri.....	18
3.3. Ara Rapor Döneminde Proje kapsamında yapılan ön deneysel çalışmalar.....	19
3.3.1. ResNet18 ve ResNet50 ön-eğitilmiş ağları üzerinde yapılan testler.....	20
3.3.2. InceptionV3 ön-eğitilmiş ağı üzerinde yapılan testler	22
3.3.3. AlexNet ön-eğitilmiş ağı üzerinde yapılan testler.....	23
3.3.4. SqueezeNet ön-eğitilmiş ağı üzerinde yapılan testler	24
4. II. ARA RAPOR DÖNEMİNDE YAPILAN ÇALIŞMALAR	27
4.1. Metodoloji	27
4.2. Veri Setinin Hazırlanması	28
4.3. Veri Artırma(Data Augmentation) İşlemi	29
4.4. Evrimsel Sinir Ağı (CNN – Convolutional Neural Network) ve Ön-Eğitilmiş Ağlar. 30	30
4.5. Bırakma(Dropout) İşlemi	32
4.6. Proje kapsamında yapılan deneysel çalışmalar	32
5. III. ARA RAPOR DÖNEMİNDE YAPILAN ÇALIŞMALAR.....	40
5.2. Özellik Çıkarımı ve Özellik Birleştirme	41
5.2.1. Texton (Texture/Dokusal) özellikler	41
5.2.2. Morfometrik özellikler	42
5.2.3. Network özellikleri.....	42
5.2.4. GLCM(gray-level co-occurrence matrix) özellikleri	43
5.2.5. Özellik Birleştirme	44
5.2.6. Özellik Seçimi	44
5.3. Kullanılan Sınıflandırma Yöntemleri.....	45
5.4. Deneysel Çalışmalar.....	47
6. SONUÇLAR VE DEĞERLENDİRMELER	50
7. PROJE KAPSAMINDA YAPILAN AKADEMİK ÇALIŞMALAR.....	51

ŞEKİLLER DİZİNİ

Şekil 1. Bazal benzeri meme kanserlerinin varsayımsal spektrumu	9
Şekil 2. Breast histopathology images veri seti içerisinde bulunan IDC negatif örnek görüntüler	16
Şekil 3. Breast histopathology images veriseti içerisinde bulunan IDC pozitif örnek görüntüler	16
Şekil 4. BreCaHAD veri seti içerisinde bulunan örnek görüntü ve etiketlemeleri	17
Şekil 5. BACH veri seti içerisinde bulunan örnek görüntüler (a) normal, (b) benign, (c) in situ carcinoma, (d) invasive carcinoma.....	18
Şekil 6. Proje kapsamında temin edilen iş istasyonu ve GPU donanım yapıları	19
Şekil 7. Ön-egitimli ağlar üzerinde test işlemleri için kullanılan yazılım ortamı	19
Şekil 8. ResNet mimarisinin atlamalı eğitim örneği	20
Şekil 9. ResNet mimarileri arasındaki farklar.....	20
Şekil 10. ResNet ağının genel mimarisi	21
Şekil 11. ResNet18 ile gerçekleştirilen eğitimin detayları, başarımları ve hata grafiği	21
Şekil 12. ResNet50 ile gerçekleştirilen eğitimin detayları, başarımları ve hata grafiği	22
Şekil 13. (a) Solda klasik evrişimli ağ katman yapısı, (b) Sağda network in network yapısı..	22
Şekil 14. Naive Inception V3 modülü ve boyut küçültme inception modülü	23
Şekil 15. GoogLeNet mimarisi.....	23
Şekil 16. Inception V3 ile gerçekleştirilen eğitimin detayları, başarımları ve hata grafiği	23
Şekil 17. AlexNet'in mimarisi	24
Şekil 18. AlexNet ile gerçekleştirilen eğitimin detayları, başarımları ve hata grafiği	24
Şekil 19. SqueezeNet Fire modülü	25
Şekil 20. SqueezeNet mimarisi	25
Şekil 21. SqueezeNet ile gerçekleştirilen eğitimin detayları, başarımları ve hata grafiği.....	26
Şekil 22. Histopatolojik görüntülerde meme kanserinin tespiti ve sınıflandırılması için önerilen derin sinir ağları tabanlı modelin blok diyagramı	28
Şekil 23. BreakHis verisetinden iyi huylu ve kötü huylu örnek histopatolojik görüntüler.....	29
Şekil 24. Örnek bir histopatolojik görüntüye veri artırma (data augmentation) işlemlerinin uygulanması	30
Şekil 25. Çalışmada kullanılan evrişimsel sinir ağı mimarisinin genel yapısı.....	31
Şekil 26. DenseNet201, InceptionV3, ResNet50 ve Xception önceden eğitilmiş ağlarının 200X yakınlaştırma oranındaki başarımları grafikleri	33
Şekil 27. 200X yakınlaştırma oranındaki görüntülerde ağların TTA'lı ve TTA'sız karmaşıklık matrisleri.....	36
Şekil 28. Yakınlaştırma oranlarına göre ağların AİK eğrilerini ve EAA değerleri.....	37
Şekil 29. Xception ağının 200X yakınlaştırma oranındaki tahminleri ve gerçek sonuçlar.....	38
Şekil 30. Xception ağının 200X yakınlaştırma oranındaki tahminleri ve gerçek sonuçlar.....	38
Şekil 31. Özellik birleştirmeye dayalı önerilen sistemin çalışma şekli.....	40
(a) negatif	41
(b) pozitif.....	41
Şekil 32. Breast histopathology verisinde (a) pozitif ve (b) negatif görüntü örnekleri	41
Şekil 33. MR8 filtresi, Gaussian filtresi ve Laplacian of Gaussian filtrelerinden Texton özelliklerin çıkarılması.....	42
Şekil 34. Morfometrik özelliklerin çıkarılması	42
Şekil 35. Texton özellikler, network özellikleri, morfometrik özellikler ve GLCM özelliklerinin birleştirilmesi	44
Şekil 36. Özellik seçimi için KNIME ortamında tasarlanan blok diyagram.....	44

Şekil 37. KNIME yazılımı üzerine test ve eğitim kümeleri ile SFFS ile elde edilen sınıflandırma sonuçları	45
Şekil 38. Özelliklerin önem derecelerine göre sıralanması.....	45
Şekil 39. Histopatolojik görüntülerden çıkarılan özelliklerin sınıflandırılması için tasarlanan DNN mimarisi	46
Şekil 40. BoW modelinin çalışma yapısını gösteren blok diyagram	46
Şekil 41. BoW yöntemi ile elde edilen sınıflandırma sonuçları.....	47
Şekil 42. DNN yöntemi ile elde edilen sınıflandırma sonuçları	47
Şekil 43. BoW+CNN hibrit yöntemi ile yapılan eğitim aşamasına ait ulaşılan bazı parametre değerleri.....	48
Şekil 44. Önerilen hibrit ağın eğitimdeki doğruluk ve Loss değerlerinin değişimini gösteren board ekranı.....	48
Şekil 45. Hibrit yöntem ile sırasıyla eğitim kümesi üzerinde elde edilen sınıflandırma sonuçlarını gösteren karmaşıklık matrisi ve sınıflandırma sonuçları.....	49
Şekil 46. Hibrit yöntem ile sırasıyla eğitim kümesi üzerinde elde edilen sınıflandırma sonuçlarını gösteren karmaşıklık matrisi ve sınıflandırma sonuçları.....	49

ÇİZELGELER DİZİNİ

Tablo 1. Ön-eğitimli ağlarda elde edilen test başarımları sonuçları.....	26
Tablo 2. BreakHis veri setindeki yakınlaştırma oranları ve görüntü sayıları.....	28
Tablo 3. Bütün ağlarda elde edilen başarımların karşılaştırılması (D.S.: Doğruluk Skoru, Has. D.: Hassasiyet Değeri, Hat.D.: Hatırlama Değeri, M.O.: Makro Ortalama, A.O.: Ağırlıklandırılmış Ortalama)	35
Tablo 4. Histopatolojik görüntülerden çıkarılan network özellikleri.....	43
Tablo 5. Histopatolojik görüntülerden çıkarılan GLCM özellikleri	43

1. GİRİŞ

Kanser, nüfusu nispeten fazla bir ülke olan ABD’de kalp hastalıklarından sonra ikinci sırada ölümlerin nedeni olarak gösterilmiş ve kısa süre içerisinde de birinci sırada yer alması beklenen bir hastalıktır [1]. Dünya Sağlık Örgütü’nün (WHO) raporuna göre; dünya çapında kadınlar arasında en yaygın olarak görülen kanser %25.2’lik oran ile meme kanseridir [2]. Meme kanserinin hastalıklılık (Morbidity) oranı ise %14.7 olup son yıllarda kanser ölümcüllüğü (Mortality) akciğer kanserinden sonra ikinci sırada yer almaktadır [3]. Yine bu araştırmaya göre, yaklaşık yarım milyon meme kanseri hastası hayatını kaybetmiş ve her yıl 1.7 milyon yeni vaka ortaya çıkmaktadır. Bu vaka sayılarının giderek artması beklenmektedir. 2019 yılında yapılan bir çalışmaya göre ise toplamda 271.270 yeni meme kanseri vakası ve aynı çalışmaya göre toplamda 42.260 meme kanseri nedeniyle ölüm beklendiği öngörülmüştür [4].

Meme kanserinin teşhisinde mamografi, ultrason ve manyetik rezonans (MR) görüntüleme gibi klasik yöntemler kullanılmaktadır. Mamografi özel bir röntgen çeşidi olarak betimlenir. Ultrason ise ses dalgaları kullanılarak tümör tespiti yapılan bir yöntemdir. MR, manyetik alanın ve radyo dalgalarının kullanılarak teşhisin yapıldığı bir yöntemdir [5]. Her ne kadar meme kanserinin teşhisinde bu yardımcı görüntüleme yöntemleri kullanılsa da, kesin tanı için histopatolojik görüntülerde testler yapılması gerekmektedir. Patologlar biyopsi ile kesin hastalık tanıları koymaktadır ancak bir hastalığın mikroskop yardımıyla detaylı bir şekilde incelenmesi çok uzun süreler alabilmektedir. Ayrıca kimi durumlarda insan kaynaklı (yorgun olma veya deneyimsizlik gibi) hatalar oluşabilmekte ve her zaman doğru sonuca ulaşmak mümkün olmamaktadır. Son yıllarda bu nedenlerden ötürü patoloji alanında makine öğrenmesi, derin öğrenme ve görüntü işleme gibi teknolojilerden faydalanılmaktadır [6]. Patologlar meme kanseri teşhisi için dokusal özellikleri, meme parankiminin normal yapılarındaki farklılıkları tespit ederek Hematoksilin ve Eosin (H&E) ile boyanmış dokuları mikroskop ile inceler ve buna göre bir sonuca varmaya çalışır. Uygun tedavi yönteminin seçilmesi için dokunun iyi huylu veya kötü huylu olduğunun zamanında tespit edilmesi gereklidir. Buna karşın H&E yöntemi pataloglar için çok zahmetli ve zaman alıcıdır. Ayrıca uzmanlar kimi zaman kanser tespitinde ortak bir sonuca varamamaktadır. Yapay zeka ve makine öğrenmesi gibi teknolojiler kanser tespitinde daha hızlı ve tek bir insana göre daha başarılı sonuçlar ortaya koymasından ötürü bu alanlarda daha sık kullanılmaya başlamıştır. Ayrıca bu yöntem sınıflandırma doğruluğunu arttırmakta ve uzmanlar arasındaki görüş farklılıklarını azaltmaktadır [7].

Bu tümör alt tiplerinin sınırlı tedavi olanaklarından dolayı arařtırmacılar tarafından daha sık alıřılmaktadır. Bazal benzeri tümörler ayırt edici morfolojik, genetik, immünofenotipik ve klinik özellikler ile karakterize edilmesine rağmen bu tümörün alt tipinin tanımlanması konusunda kabul görmüş bir fikir birlięi olmadığı gibi sistematik olarak sınıflandırılması ile ilgili bir yol tanımlanmamıştır. Sınıflandırma için immünohistokimyasal belirteçler kullanılmakta fakat bu belirteçlerin başarımları ise %60 seviyelerinde kalmaktadır.

Patoloji ve onkoloji alanlarında kullanılan makine öğrenmesi modelleri histopatolojik görüntüdeki tümörün iyi huylu veya kötü huylu olduğunun sınıflandırılmasını sağlamaktadır. Tümör sadece bir kitle olarak tanımlanabilir ve bu tanıma göre her tümör insanlar için bir tehlike arz etmez ve bazı tümörler iyi huylu olabilir. Kanser dediğimiz hastalık hücrelerin kontrolsüz bir şekilde bölündüğü ve yayıldığı durumları ifade etmektedir [8]. Derin öğrenme algoritmalarında kullanılan modeller eğitim aşamasında kendisine gösterilen onlarca örneğin özellik çıkartmasını(hücre büyüklüğü, rengi ve şekli gibi) yapar ve daha önce görmediği bir takım görüntüler üzerinde test edilerek modelin başarılı olup olmadığı gösterilir. Kullanılan modeller çok fazla piksel ve matris işlemleri yapması nedeniyle güçlü donanımlara sahip bir bilgisayara ihtiyaç duyar. Yapılan eğitimler kimi durumlarda günler sürebilmektedir. Modeller bu kadar çok gereksinime ve zamana ihtiyaç duymalarına karşın kimi durumlarda yanlış tahminlerde bulunabilmekte ve hatalar yapabilmektedir. Hatalı sonuçlar elde edilmesinin en büyük sebepleri: tümörlü hücrenin arka plandan ve diğer dokulardan ayırt edilmesinde zorluk çekilmesi, görüntüye uygulanan bazı işlemlerin özellik kaybına neden olması vb.dir. Modelleri bu eksikliklerinden kurtarmak ve hatta daha iyi sonuçlar elde etmek için önceden eğitilmiş modeller(pre-trained) ortaya çıkmıştır [9].

2. LİTERATÜR ARAŞTIRMALARI VE ÖNCEKİ ÇALIŞMALARIN DEĞERLENDİRİLMESİ

Literatürde yapılan çalışmalarda, meme kanserinin erken teşhis ve tedavisinin, kadın hastaların geç tanı konan hastalardan en az beş yıl daha uzun süre hayatta kalmasına yardımcı olduğunu göstermektedir [10]. Maliyet etkin bir yaklaşım olan mikroskopik görüntü değerlendirmesi, meme kanserlerinin tanı ve tedavisinde en sık kullanılan yöntemlerden biridir [11]. Patologlar, östrojen reseptörü (ER), progesteron reseptörü (PR) ve Her2 gibi üç ana biyobelirteci inceleyerek; etiyolojiiyi belirler, davranışı tahmin eder ve hastalığın klinik sunumuyla bağlantı kurar [12]. Bu uygulamalarda meme kanserinin derecelendirilmesi ve sınıflandırılması, yüksek işgücü maliyetleri ve araştırmacılar arasında değişkenlik gösteren mesleki deneyimlere dayanmaktadır. Bu sorunların üstesinden gelmek için, klinik uygulamada meme kanserinin yönetimine yardımcı olmak üzere bir dizi otomatik görüntü analiz yöntemi önerilmiştir.

Bazal benzeri tümörler, tüm meme kanserlerinin% 16-37'sini oluşturur. Diğer meme kanseri alt sınıflarının aksine, bazal benzeri alt tip, ER, PR ve Her2/neu ekspresyonu eksikliğini gösterir. Bazal benzeri tümörlerin kötü tanımlanması klinik tedavi eksikliğine neden olabilir. Deneysel sonuçlar, toplumda yerleşmiş histolojik özelliklerin yeterince belirli olmadığını ve bazal benzeri tümörlerin çalışmalarında tutarlı bir şekilde tanımlanmadığını göstermektedir [12].

Görüntü işleme teknikleri, tıbbi görüntülerin yorumlanmasında ve otomatik tanı için önemli bir role sahiptir. Özellikle son yıllarda, bütün slaytlı görüntüleme (whole-slide imaging) tekniklerinin geliştirilmesi ve kanser vakalarındaki artış, birçok araştırmacının otomatik histopatolojik görüntü analizine dikkatini çekmiştir [13]. Bütün görüntüler çok yüksek çözünürlükte ve uzmanlar tarafından incelenmesi çok uzun sürmektedir. Bu ayrıntılı işlemi kolaylaştırmak için literatürde bilgisayar destekli otomatik teşhis yöntemleri üzerinde durulmuştur [14]. Bu yöntemler uzmana görüntü analizinde, karar vermesinde yardımcı olur ve bazı durumlarda karar vericinin rolünü üstlenir [15].

Literatür incelendiğinde araştırmacılar, bazal benzeri meme kanseri tespiti veya herhangi bir kanserin tespiti için histopatolojik görüntüleri analiz etmek için çalışmışlardır. Son yıllarda bu alanda çok miktarda çalışma bulunmaktadır.

2013 yılında yapılan bir araştırmada meme kanseri sınıflandırması için PNN (Probabilistic Neural Networks), MLP (Multi Layer Perceptron) ve RBF (Radial Basis Function) modelleri ile bir çalışma yapılmıştır. Yürütülen deneysel çalışmalarda, PNN ile daha başarılı sonuçların elde edildiği görülmüştür [16].

Bir başka çalışmada, meme kanserinin tespiti için global veriseti üzerinde geri yayılım ağı ve Liebenberg Marquardt öğrenme fonksiyonunun kullanıldığı, ayrıca ağırlıkların DBN-NN (deep belief network path) ile başlatıldığı bir model ile yüksek bir doğruluk oranı elde edilmiştir [17].

2018 yılında Öztürk ve Akdemir gerçekleştirdikleri çalışmada, iyi bilinen özellik çıkarım algoritmaları ve sınıflandırma algoritmaları histopatolojik görüntüler üzerinde karşılaştırılmıştır. Kesilmiş görüntü parçalarından GLCM (Gray level co-occurrence matrix), LBP (Local binary pattern), LBGLCM (Local Binary Gray Level Co-occurrence Matrix), GLRLM (Gray Level Run Length Matrix) ve SFTA (Segmentation-based Fractal Texture Analysis) ile çıkartılmış özellik matrisleri SVM (Support Vector Machine), KNN (K-nearest neighbors), LDA (Linear Discriminant Analysis) ve Boostted Tree modelleri ile sınıflandırmıştır. Çalışma sonucunda yazarlar SFTA algoritması tarafından elde edilen özellik matrisi sonuçları, diğer özellik çıkarma algoritmalarından daha başarılı sonuçlar verdiğini belirtmişlerdir. Ayrıca LBP algoritması, diğer özellik çıkarma algoritmalarından daha başarısız sonuçlar verdiğini bildirmişlerdir. Sınıflandırma algoritmaları arasında SVM ve Boostted Tree algoritmaları daha başarılı olduğunu ve en başarılı kombinasyon %94,3 ile SFTA ve Boostted Tree kombinasyon olduğunu belirtmişlerdir [14].

2018 yılında yapılan bir çalışma da BreakHis veri seti kullanılarak meme kanseri sınıflandırması yapılmıştır. Bu çalışma da CNN (Convolutional Neural Network) ve LSTM (Long Short Term Memory)'nin bir kombinasyonu kullanılmıştır. [18].

Bir diğer çalışmada, 72'si tümörlü ve 184'ü tümörsüz görüntüden oluşan toplam 184 meme ultrason görüntüsü içeren verisetinde göğüslerdeki kitlelerin tespiti yapılmıştır. Çalışmada çoklu fraktal boyut özellikleri kullanılarak iyi düzeyde bir hassasiyet değerine ulaşılmıştır [19].

Yine 2018 yılında Wang ve ark., yayınladıkları makalede histopatolojik görüntülerin sınıflandırılması için özellik tabanlı analiz şeması önermişlerdir. Yöntemde ilk olarak, hücre çekirdeği ortak hücre lokalizasyonu ve gelişmiş morfoloji yöntemine göre bölümlendirilir.

Daha sonra, şekil özellikleri, istatistiksel özellikler ve metinsel özellikler (Gabor ve Markov rastgele alan özellikleri) dahil olmak üzere çeşitli özellikler çıkarılır. Yazarlar çalışmanın deneysel sonuçlarına bakılarak, önerilen bölümlendirme yönteminin histopatolojik görüntüleri otomatik olarak bölümleyebileceği ve etkili bölümlendirme sonuçlarına sahip olduğunu iddia etmişlerdir [20].

Khameneh ve ark., 2019 yılında yaptıkları çalışmada IHC meme kanseri görüntülerini etkili bir şekilde bölümlere ayırmak, sınıflandırmak ve ölçmek için bir makine öğrenmesi (ML) yapısı sunmuşlardır. Önerilen teknik üç adımdan oluşmaktadır. İlk adımda, epilasyonel ve stromal bölgeleri WSI'dan sınıflandırmak için süper piksel tabanlı bir destek vektör makinesi (SVM) özelliği öğrenme sınıflandırıcısı önerilmiştir. İkinci aşamada, sınıflandırılmış epitelyal bölgelerde, segment membran (zar) bölgelerine bir evrişimsel sinir ağı (CNN) bazlı segmentasyon yöntemi uygulanmıştır. Son olarak, bölünmüş alanlar birleştirilmiş ve her slaydın toplam puanı değerlendirilmiştir. 127 slayt için deneysel sonuçlar sunulmuş ve modern yöntemler, derin öğrenme temelli yaklaşımlarla karşılaştırılmıştır. Deneysel sonuçlar, önerilen yöntemin IHC ile lekelenmiş veriler üzerinde ümit verici bir performans elde ettiğini göstermektedir. Sunulan otomatik algoritmanın, CNN kullanarak epitel bölgelerinin süper piksel bazlı sınıflandırılması ve membran (zar) boyama bölümleri bakımından diğer yaklaşımlardan daha iyi performans gösterdiği söylenmektedir [21].

Aynı yıl (2019) Budak ve ark., gerçekleştirdikleri çalışmada meme kanseri tespiti için tamamen evri ağı (FCN: Fully Convolutional Network) ve çift yönlü uzun kısa süreli belleğe (Bi-LSTM: Bidirectional Long Short Term Memory) dayanan uçtan uca bir model tanıtmışlardır. Deneysel sonuçlarını <http://web.inf.ufpr.br/vri/breast-cancer-database> adresinde halka açık olan BreakHis veritabanı üzerinde yapmışlar ve beş kat çapraz değerlendirme tekniği göz önüne alınarak doğruluk ölçümü yapıldığında, önerilen yöntemin performansı daha önce bildirilmiş sonuçlardan daha iyi olduğunu iddia etmişlerdir [22].

Hekler ve ark., Almanya'da 2019 yılında yaptıkları bir çalışmada histopatolojik melanom görüntülerinin sınıflandırılmasında derin öğrenmenin performansını literatürde bir ilk olarak doğrudan uzmanlar ile karşılaştırmayı amaçlamışlardır. Gerçekleştirilen çalışmada ResNet50 pre-trained ağı yardımıyla oluşturulan CNN modeli kullanılmıştır. 595 adet görüntü CNN ağını eğitmek için kullanılmıştır. Çalışmanın testi için 11 patalog sonucu ile CNN sonucu kıyaslanmıştır. CNN, 11 test çalışmasında % 76.6 / %60.68 ortalama hassasiyet / özgüllük / doğruluk sağlamıştır. Buna karşılık, 11 patalog, %51.8 / %66.5 / %59.2'lik bir ortalama

hassasiyet / özgüllük / doğruluk sağlamıştır. Böylece, CNN kesilen görüntülerin sınıflandırılmasında önemli ölçüde ($p=0.016$) üstün olmuştur. Sonuçta sadece iki patalogun CNN yapısına yakın sonuçlar verebildiği belirtilmiştir [23].

Öztürk ve Akdemir 2019 yılında yaptıkları bir başka çalışmada ise tüm-slayt histopatolojik görüntüler (WSI) üzerindeki kanserli alanları otomatik olarak tanımlamak için evrişimli bir sinir ağı (CNN) modeli önermişlerdir. WSI sınıflandırma ağı için HIC-net mimarisi kullanmışlardır. HIC-net algoritmasının değerlendirilmesinde 30.656 görüntü içeren geniş bir veri kümesi kullanmışlardır. Çalışma sonucunda HIC-net başarı oranını %96.71 duyarlılık, %95.7 özgüllük, %96.21 doğruluk olarak bildirmişlerdir [8].

Sudharshan ve ark., gerçekleştirdikleri uygulamada meme kanseri hastalarının bilgisayar destekli teşhisi için, Çoklu Örnek Öğreniminin (MIL: Multiple Instance Learning) histopatolojik görüntülerin analizine dayalı kısmen denetimli bir öğrenme çerçevesi (framework) önermişlerdir. Çalışmada MIL metotlarının öncülerinden olan APR, Diverse Density, MI-SVM, citation-kNN ve derin öğrenme temelli parametrik olmayan bir yöntem bir yaklaşım olan (MIL-CNN) testleri gerçekleştirilmiştir. Deneyler, 82 hastadan alınan, yaklaşık 8000 mikroskopik biyopsi görüntüsü olan benign ve malign meme tümörlerinin yer aldığı halka açık BreakHis veri setinde gerçekleştirilmiştir. Sonuç olarak MIL metotları arasında parametrik olmayan yaklaşım en iyi genel sonuçlara sahip olduğunu ve bazı durumlarda konvansiyonel (tek örnek) sınıflandırma çerçeveleriyle asla ulaşılmayan sınıflandırma oranları elde edilmesine izin verebileceğini belirtmişlerdir[24].

Şu anda basım aşamasında olan bir çalışmada Yan ve ark., meme kanseri histopatolojik görüntü sınıflaması için yeni bir hibrid evrişimli ve tekrarlayan derin sinir ağı önermişlerdir. Histopatolojik görüntü parçalarının daha zengin çok seviyeli özellik gösterimini temel alan yöntem, evrişimli ve tekrarlayan sinir ağlarının avantajlarını bütünleştirir ve parçalar arasındaki kısa ve uzun vadeli uzamsal korelasyonlar korunur. Çalışma sonuçlarında önerilen metodun, 4 sınıflı sınıflandırma görevi için modern yöntemle elde edilen ortalama doğruluğu %91.3'ten daha iyi bir performans gösterdiğini belirtmişlerdir. Şu an halka açık olan http://ear.ict.ac.cn/?page_id=1616 adresinde yayınlanabilen bilimsel topluluk için 3771 meme kanseri histopatolojik görüntüleri içeren bir veri seti de yayınlamışlardır [25].

2020 yılında yayınlanacak bir çalışmada Kumar ve ark., Köpek meme tümörleri (CMT: Canine mammary tumors) ve insan meme kanserinin histopatolojik görüntü sınıflaması için

derin öğrenme temelli bir yaklaşım sunmuşlardır. VGGNet-16'ya dayanan bir CNN mimarisi ile transfer öğrenme yöntemi kullanmışlardır. Sistemin performansını CMT veri kümesi (CMTHis) ve insan meme kanseri veri seti (BreakHis) üzerindeki farklı sınıflandırıcılarla birlikte değerlendirmişlerdir. İnsan meme kanseri ve CMT'nin ikili sınıflandırılması için önerilen Destek vektör makineleri kullanılan yöntem sistemin etkinliğini doğrulayan sırasıyla %97 ve %93'lük ortalama doğruluklarla sonuçlanmıştır. Ayrıca ilk defa CMT histopatolojik görüntülerinin (CMTHis) bir veri kümesi çalışma tarafından sunulmuştur [7].

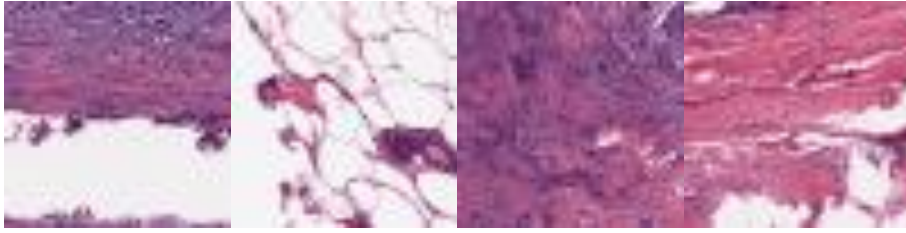
2020 yılında yapılan bir çalışma da meme tümörünün (CMT) köpekler için de oldukça ölümcül olduğu ve insan meme kanseri çalışmaları için oldukça güzel örnekler olduğu görülmüştür. Her iki tümör içinde hematoksilin ve eosin (H&E) ile boyanmış görüntüler kullanılan bu çalışma histopatolojik görüntüler üzerinde bir derin sinir ağı yöntemi önermiştir. Bu çalışma da VGGNet-16 ön-eğitilmiş ağı kullanılarak sınıflandırma yapılmıştır [7].

3. I. ARA RAPOR DÖNEMİNDE YAPILAN ÇALIŞMALAR

3.1. Açık Verisetleri

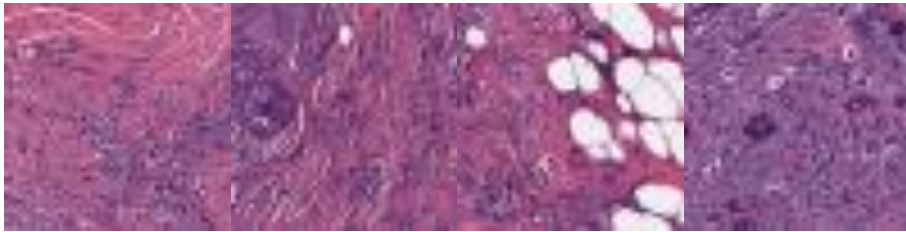
3.1.1. Breast Histopathology Images Veriseti (Kaggle)

Breast Histopathology Images veriseti [26] meme kanseri histopatoloji görüntülerinden elde edilmiştir. Görüntüler orijinal olarak 40x taranan 281 meme kanseri (BCa) örneğinin 281 slayt görüntüsünden çıkartılan 277.524 adet 50×50 parçadan oluşmaktadır. Elde edilen görüntülerin 198.738 adetine IDC negatif, 78.786 adetine ise IDC pozitif olarak tanı konulmuştur. Bölgesel evrişimsel sinir ağı ile geliştirilecek modelde elde edilen bu veri seti üzerinde çalışılacak olup çalışma esnasında 5.000 adet IDC pozitif ve 5.000 adet IDC negatif görüntüler bir alt kümede kullanılacaktır. Yine bu veri seti ön eğitilmiş ağlar ile kurulan sınıflandırma sisteminde kullanılarak tasarlanan sistemin başarımı ölçülecektir.



Şekil 2. Breast histopathology images veri seti içerisinde bulunan IDC negatif örnek görüntüler

Şekil 2’de Breast Histopathology Images veriseti içerisinde bulunan IDC negatif örnek görüntüler görülmektedir. Şekil 3’te ise aynı veri seti içerisinde bulunan IDC pozitif örnek görüntüler görülmektedir.

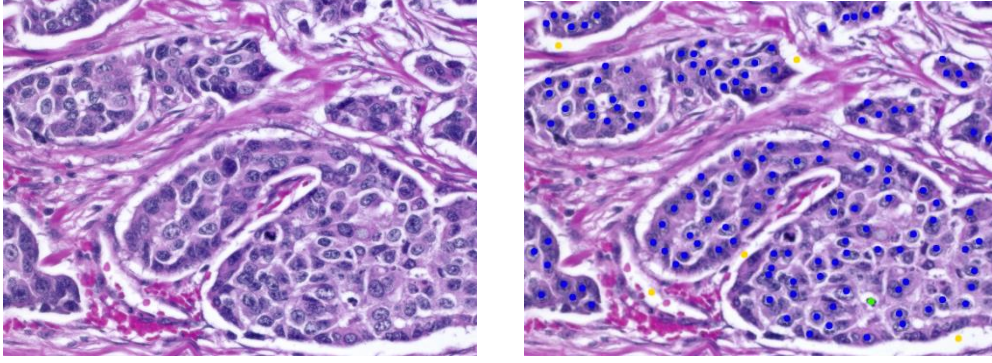


Şekil 3. Breast histopathology images veriseti içerisinde bulunan IDC pozitif örnek görüntüler

3.1.2. BreCaHAD Veriseti

BreCaHAD veriseti [27], 162 hastadan alınan meme kanseri histopatoloji görüntülerinden oluşmaktadır. Görüntüler piksel başına $0.514\mu\text{m} \times 0.527\mu\text{m}$ olup 40× taranan doku numunesi slaytlarından oluşmaktadır. Elde edilen görüntülerin her biri TIFF formatında, üç kanallı RGB ve her bir kanal 8-bit derinliğine sahip 1360×1024 piksel boyutlarında kaydedilmiş mikroskobik biyopsi görüntüleridir. Veri seti içerisinde bulunan bütün görüntülerde tümör

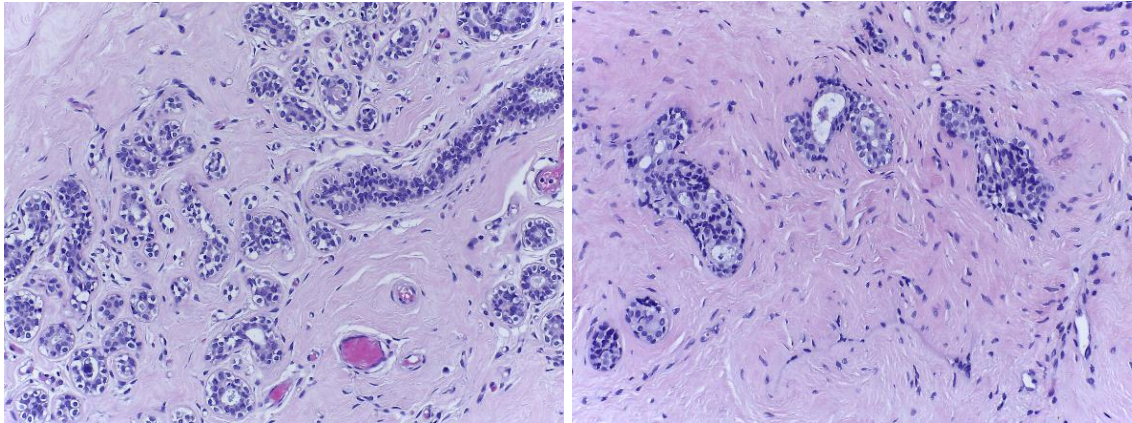
çeşitleri bulunmaktadır. Konunun uzmanlar tarafından bu tümörlü bölgeler mitoz, apoptoz, tümör çekirdekleri, tümör dışı çekirdekler, tübül ve tübül olmayan şeklinde farklı şekillerde etiketlenerek JSON formatında kaydedilmiştir. Bu veri seti, maskeli bölgesel evrimsel sinir ağı ile gerçekleştirilecek olan sistem ile bölgesel modüllerin tespitinin yapılması çalışmalarında kullanılacaktır. Şekil 5’de BreCaHAD veri seti içerisinde bulunan örnek görüntüler gösterilmektedir.

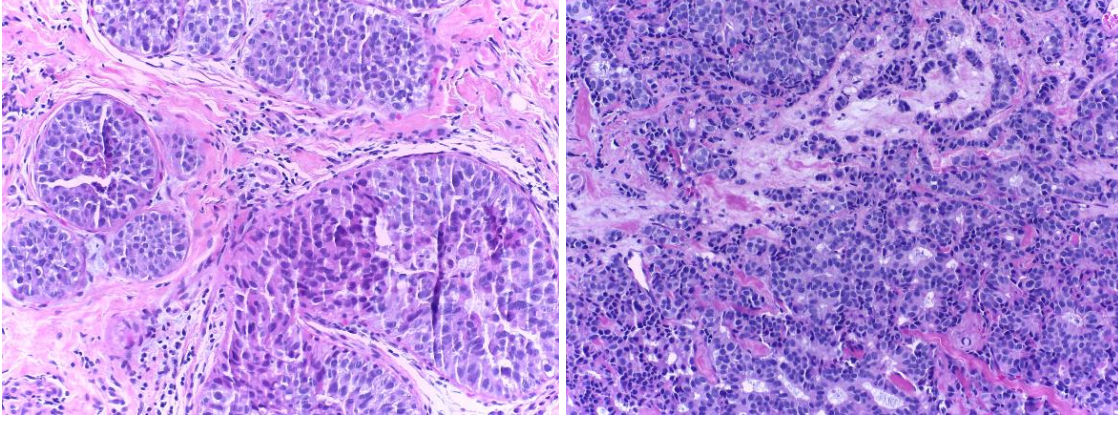


Şekil 4. BreCaHAD veri seti içerisinde bulunan örnek görüntü ve etiketlemeleri

3.1.3. BACH Veriseti

BACH veriseti [28], meme kanseri histopatoloji görüntülerinden oluşmaktadır. Görüntüler piksel başına $0.42\mu\text{m} \times 0.42\mu\text{m}$ olup 2048×1536 piksel boyutlarında görüntülerden oluşmaktadır. Veri seti içerisindeki görüntüler normal, iyi huylu (benign), in situ carcinoma, invasive carcinoma şeklinde sınıflandırılmış 100’er adet toplamda 400 görüntü bulunmaktadır. Proje kapsamında geliştirilecek sistemde sınıflandırma işlemlerinde kullanılacak olup sistemin başarımının analiz edilmesinde yardımcı olacak bir veri setidir. Şekil 5’te veri seti içerisinde bulunan her bir sınıf için örnek birer görüntü görülmektedir.





a b
c d

Şekil 5. BACH veri seti içerisinde bulunan örnek görüntüler (a) normal, (b) benign, (c) in situ carcinoma, (d) invasive carcinoma

3.2. Proje Kapsamında Sağlanan Materyal Teminleri

Projede kullanılmak üzere Şekil 6'da gösterilen ve aşağıda yazılı olan özelliklere sahip bir iş istasyonu bilgisayar temin edilmiştir;

- *Intel® Core™ i9-9900K CPU (8 Core/ 16 Thread),*
- *32GB DDR4 2666MHz Memory,*
- *Workstation Class Motherboard,*
- *3TB SATA 6Gb 3.5" Enterprise Hard Drivers (1 Adet),*
- *250GB SATA 6Gb Internal SSD (1 Adet),*
- *1200W Power Supply,*

Nvidia Geforce® RTX 2080Ti 352-bit 11GB GDDR6 Memory (1 Adet),

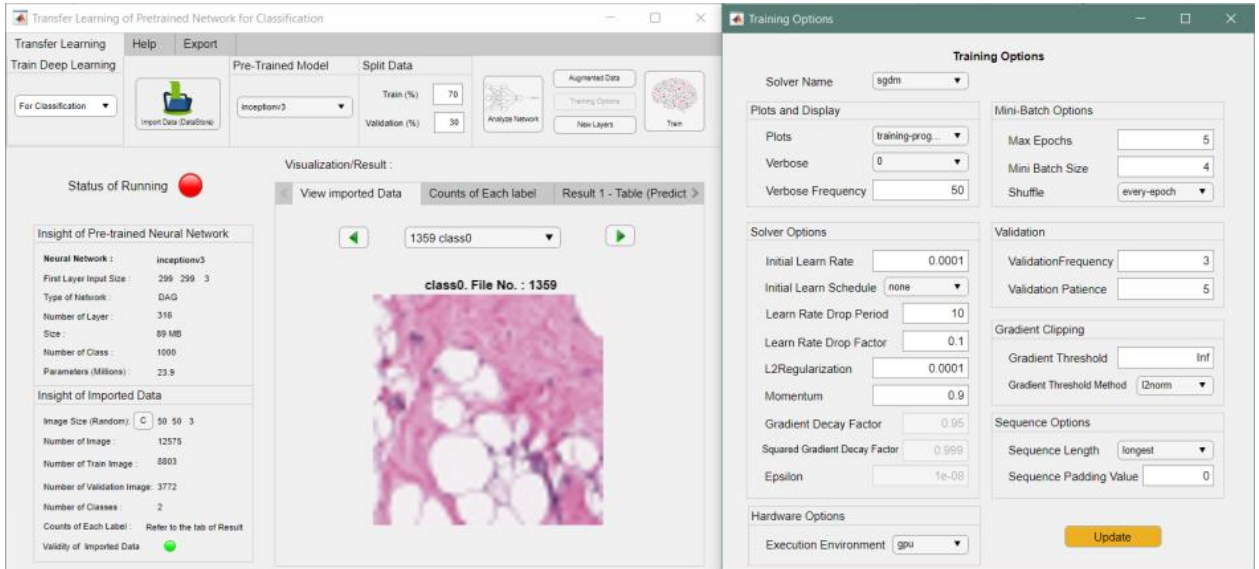
Proje kapsamında temini edilen iş istasyonu üzerinde geliştirilen derin öğrenme temelli sistemler çalıştırılmaktadır.



Şekil 6. Proje kapsamında temin edilen iş istasyonu ve GPU donanım yapıları

3.3. Ara Rapor Döneminde Proje kapsamında yapılan ön deneysel çalışmalar

Breast Histopathology Images verisetindeki görüntülerden rastgele seçilen %10 oranındaki toplam 26504 (19249 Class 0, 7255 Class1) görüntü üzerinde çeşitli ön eğitilmiş ağlar üzerinde testler gerçekleştirilmiştir. Test için Resnet18, Resnet50, AlexNet, Squaezenet ve InceptionV3 ağları kullanılmıştır. Kullanılan ağlar hakkında bilgi ve veri kümesinin sonuçları şu şekildedir. Yapılan test çalışmalarında karşılaşılan en önemli sorun eğitimin eşik değerlere hızlı ulaşması nedeniyle tüm turların tamamlanamamış olmasıdır. Şekil 7' de test işlemlerinin yapıldığı yazılım üzerinde ön-eğitilmiş ağlar için kullanılan parametrelerin ayarlanması ve eğitim işleminin başlatılması aşamaları görülmektedir.

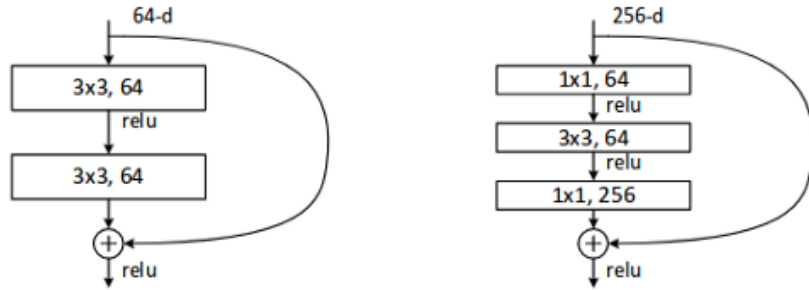


Şekil 7. Ön-eğitilmiş ağlar üzerinde test işlemleri için kullanılan yazılım ortamı

3.3.1. ResNet18 ve ResNet50 ön-eğitilmiş ağı üzerinde yapılan testler

Ağ modelinin gerçek anlamda derinleşmeye başladığı kendinden önceki modellerden farklı bir mantığı barındıran ResNet; artık değerlerin (residual value) sonraki katmanlara besleyen blokların (residual block) modele eklenmesiyle oluşmaktadır [29]. ResNet bu özelliği ile klasik bir model olmaktan çıkmaktadır.

Doğrusal ve ReLU arasında iki katmanda bir eklenen bu değer Şekil 8’de görüldü gibi sistemdeki hesabı değiştirir. Önceden Layer A ya gelen değeri Layer A+2’de veya ResNet’in bazı modellerinde Layer A+3’de tekrar hesaba katılmış olur. Şekil 9’da ResNet mimarileri arasındaki katman farklılıkları görülmektedir.

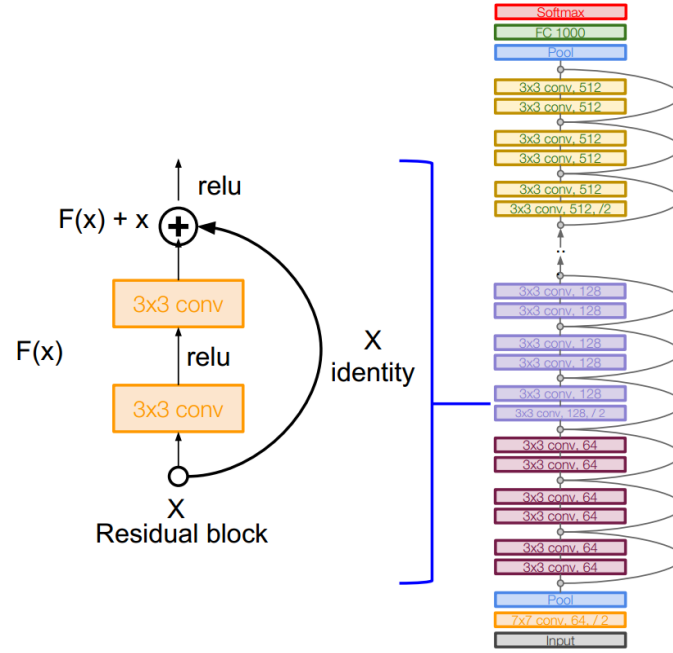


Şekil 8. ResNet mimarisinin atlamalı eğitim örneği

layer name	output size	18-layer	34-layer	50-layer	101-layer	152-layer
conv1	112×112	7×7, 64, stride 2				
		3×3 max pool, stride 2				
conv2_x	56×56	$\begin{bmatrix} 3 \times 3, 64 \\ 3 \times 3, 64 \end{bmatrix} \times 2$	$\begin{bmatrix} 3 \times 3, 64 \\ 3 \times 3, 64 \end{bmatrix} \times 3$	$\begin{bmatrix} 1 \times 1, 64 \\ 3 \times 3, 64 \\ 1 \times 1, 256 \end{bmatrix} \times 3$	$\begin{bmatrix} 1 \times 1, 64 \\ 3 \times 3, 64 \\ 1 \times 1, 256 \end{bmatrix} \times 3$	$\begin{bmatrix} 1 \times 1, 64 \\ 3 \times 3, 64 \\ 1 \times 1, 256 \end{bmatrix} \times 3$
conv3_x	28×28	$\begin{bmatrix} 3 \times 3, 128 \\ 3 \times 3, 128 \end{bmatrix} \times 2$	$\begin{bmatrix} 3 \times 3, 128 \\ 3 \times 3, 128 \end{bmatrix} \times 4$	$\begin{bmatrix} 1 \times 1, 128 \\ 3 \times 3, 128 \\ 1 \times 1, 512 \end{bmatrix} \times 4$	$\begin{bmatrix} 1 \times 1, 128 \\ 3 \times 3, 128 \\ 1 \times 1, 512 \end{bmatrix} \times 4$	$\begin{bmatrix} 1 \times 1, 128 \\ 3 \times 3, 128 \\ 1 \times 1, 512 \end{bmatrix} \times 8$
conv4_x	14×14	$\begin{bmatrix} 3 \times 3, 256 \\ 3 \times 3, 256 \end{bmatrix} \times 2$	$\begin{bmatrix} 3 \times 3, 256 \\ 3 \times 3, 256 \end{bmatrix} \times 6$	$\begin{bmatrix} 1 \times 1, 256 \\ 3 \times 3, 256 \\ 1 \times 1, 1024 \end{bmatrix} \times 6$	$\begin{bmatrix} 1 \times 1, 256 \\ 3 \times 3, 256 \\ 1 \times 1, 1024 \end{bmatrix} \times 23$	$\begin{bmatrix} 1 \times 1, 256 \\ 3 \times 3, 256 \\ 1 \times 1, 1024 \end{bmatrix} \times 36$
conv5_x	7×7	$\begin{bmatrix} 3 \times 3, 512 \\ 3 \times 3, 512 \end{bmatrix} \times 2$	$\begin{bmatrix} 3 \times 3, 512 \\ 3 \times 3, 512 \end{bmatrix} \times 3$	$\begin{bmatrix} 1 \times 1, 512 \\ 3 \times 3, 512 \\ 1 \times 1, 2048 \end{bmatrix} \times 3$	$\begin{bmatrix} 1 \times 1, 512 \\ 3 \times 3, 512 \\ 1 \times 1, 2048 \end{bmatrix} \times 3$	$\begin{bmatrix} 1 \times 1, 512 \\ 3 \times 3, 512 \\ 1 \times 1, 2048 \end{bmatrix} \times 3$
	1×1	average pool, 1000-d fc, softmax				
FLOPs		1.8×10^9	3.6×10^9	3.8×10^9	7.6×10^9	11.3×10^9

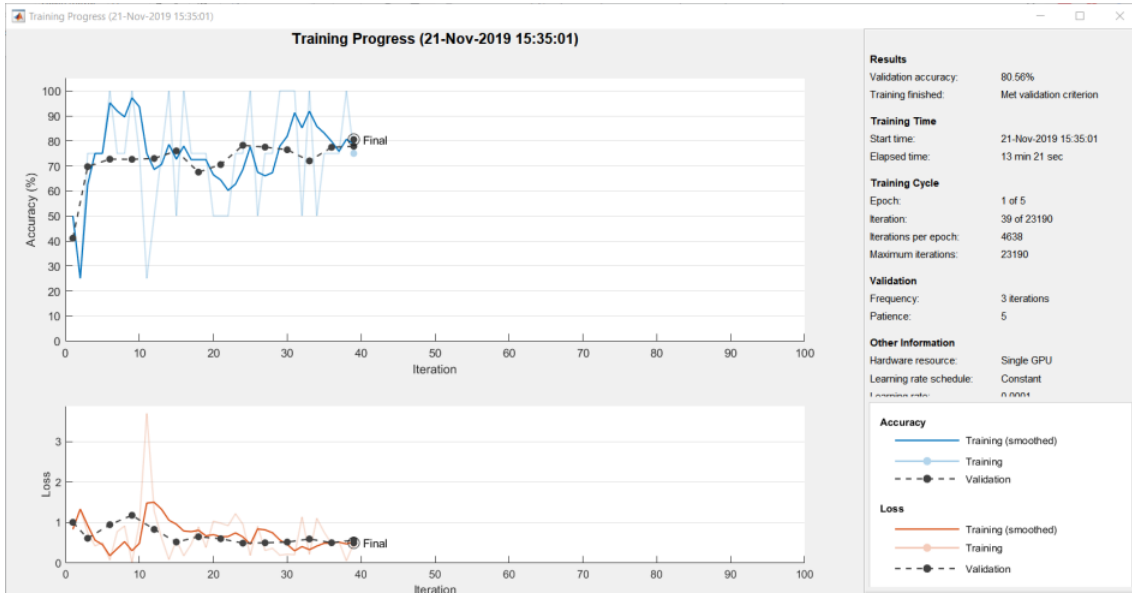
Şekil 9. ResNet mimarileri arasındaki farklar.

Şekil 10’da ResNet Mimarilerinin genel yapısı görülmektedir. Bu mimarinin 18 katmanlı versiyonu ResNet 18, 50 katmanlı versiyonu ResNet50 olarak isimlendirilir.



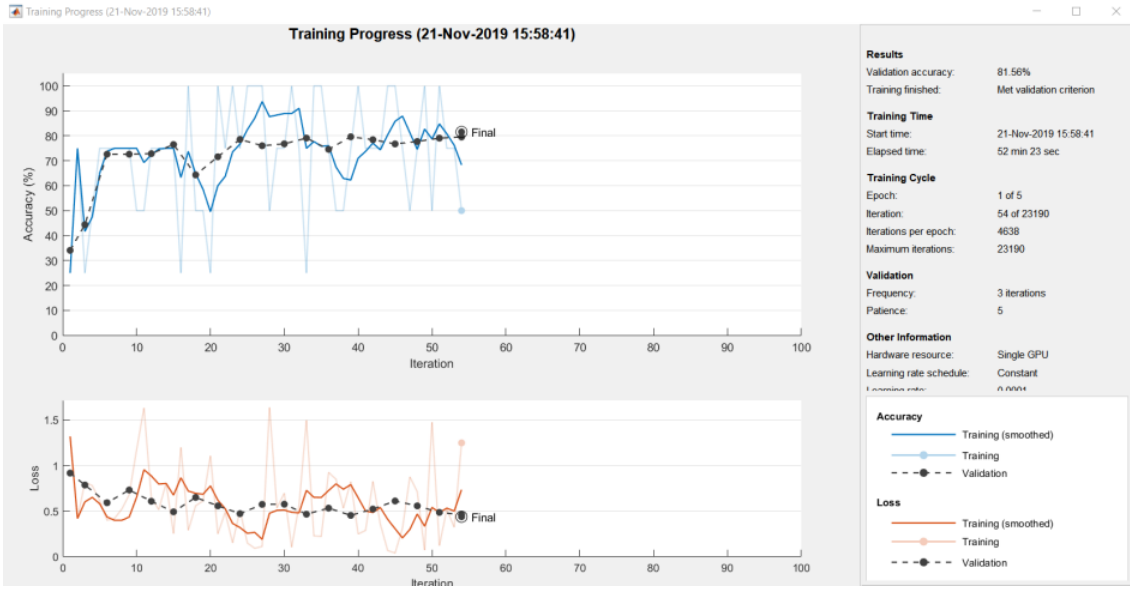
Şekil 10. ResNet ağının genel mimarisi

ResNet 18’de yapılan testlerde eğitimin başarımı **%80,56** olmuştur. Şekil 11’de ResNet18 ile gerçekleştirilen eğitimin detayları, başarım ve hata grafiği görülmektedir.



Şekil 11. ResNet18 ile gerçekleştirilen eğitimin detayları, başarım ve hata grafiği

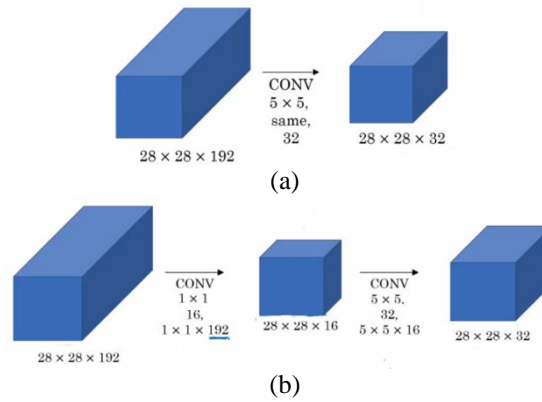
ResNet 50’de yapılan testlerde eğitimin başarımı **%81,56** olmuştur. Şekil 12’de ResNet50 ile gerçekleştirilen eğitimin detayları, başarım ve hata grafiği görülmektedir.



Şekil 12. ResNet50 ile gerçekleştirilen eğitimin detayları, başarımları ve hata grafiği

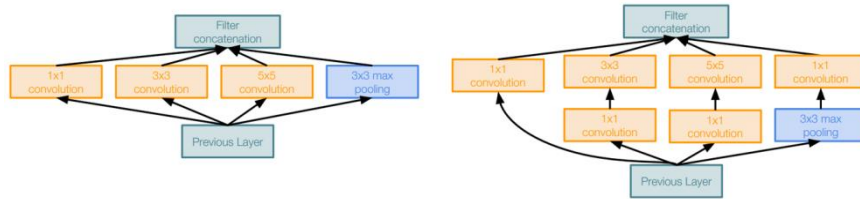
3.3.2. InceptionV3 ön-eğitilmiş ağı üzerinde yapılan testler

Bir Google ekibi tarafından geliştirilen Inception V3, Inception V2'den ayıran tek özelliği sadece evrişim katmanları değil yardımcı sınıflayıcı olarak küme normalizasyonu (batch-normalized) ve tam bağlı (FC) katmanı eklenmesidir [30]. Inception v2 bu şekilde düzenlendiğinde buna Inception v3 denmektedir. Inception ağlarının en önemli özelliği Network In Network prensibiyle ağ içindeki hesaplanması gereken parametre miktarlarını azaltmasıdır. Şekil 13(a)'da Klasik Evrişimsel sinir ağları görülürken Şekil 13(b)'de Network In Network modeline örnek görülmektedir. İlk durumda 120 Milyon parametre hesaplanması gerekirken, ikinci ağın tamamında 12,4 Milyon parametre ile aynı sonuç elde edilebilmektedir.



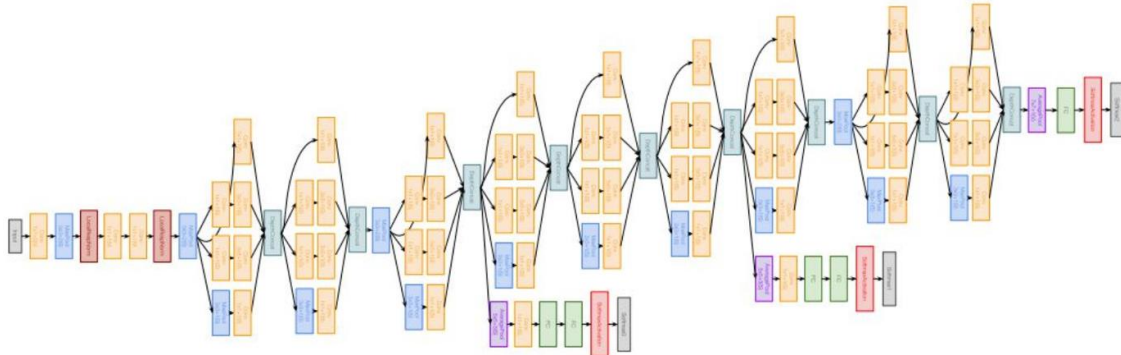
Şekil 13. (a) Solda klasik evrişimli ağ katman yapısı, (b) Sağda network in network yapısı

Inception ağları Naive Inception V3 Modülü ve Boyut Küçültme Inception Modülü gibi alt blokların birleşmesi ile oluşur. Şekil 14'te modüllerin yapısı görülmektedir.



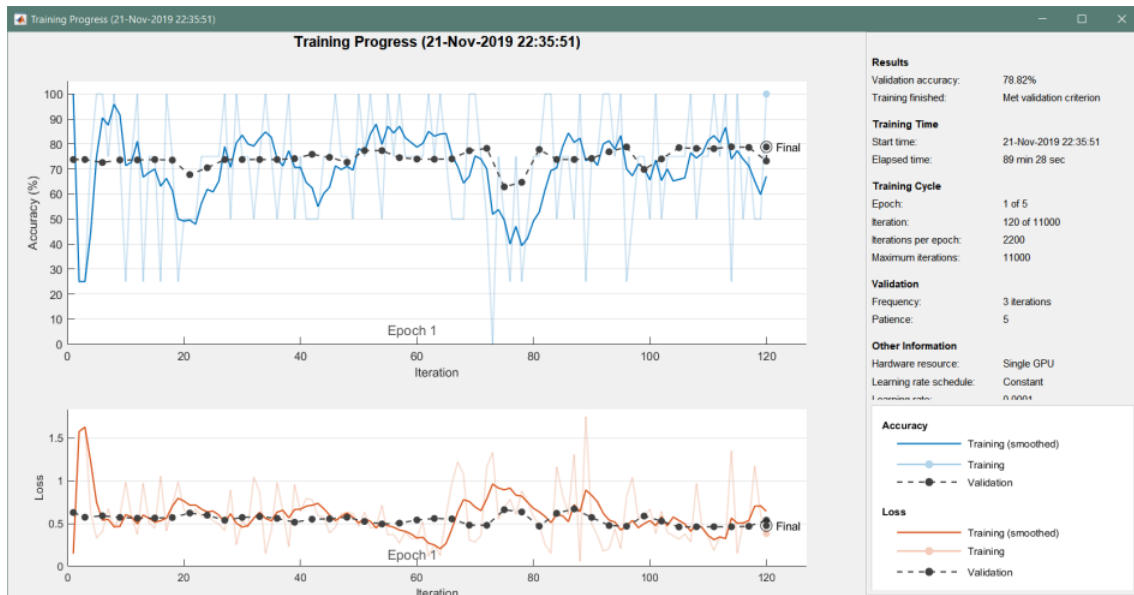
Şekil 14. Naive Inception V3 modülü ve boyut küçültme inception modülü

Şekil 15'te GoLeNet mimarisi görülmektedir. Bu ağı yapısında Inception Mimarisine dayanmaktadır.



Şekil 15. GoLeNet mimarisi

Inception V3'de yapılan testlerde eğitimin başarımı **%78,82** olmuştur. Şekil 16'da Inception V3 ile gerçekleştirilen eğitimin detayları, başarım ve hata grafiği görülmektedir.

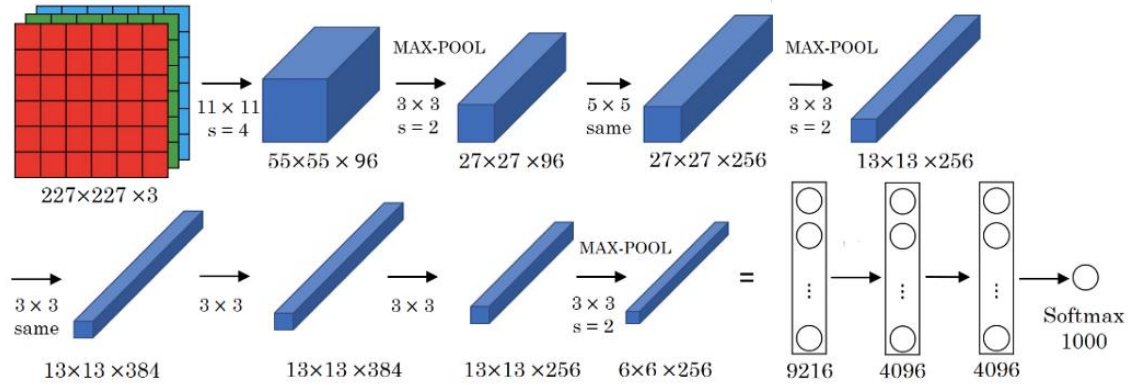


Şekil 16. Inception V3 ile gerçekleştirilen eğitimin detayları, başarım ve hata grafiği

3.3.3. AlexNet ön-eğitimli ağı üzerinde yapılan testler

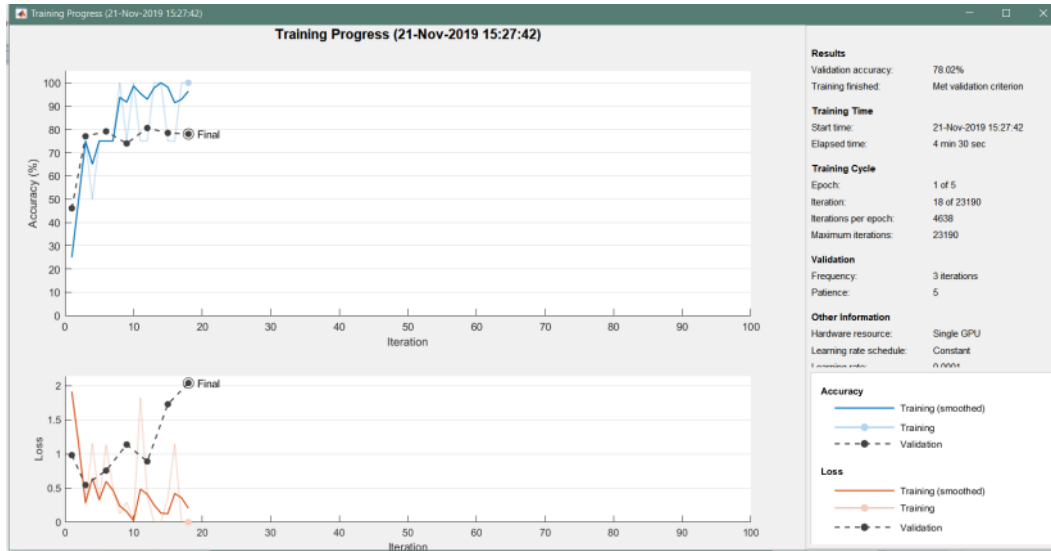
2012 yılında evrişimli sinir ağı modellerini ve derin öğrenmenin tekrar popüler hale gelmesini sağlayan ilk çalışmadır. Alex Krizhevsky, Ilya Sutskever and Geoffrey Hinton tarafından

geliştirilmiştir. Temel olarak LeNet modeline, birbirini takip eden evrişim ve pooling katmanları bulunmasından dolayı, çok benzemektedir. Aktivasyon fonksiyonu olarak ReLU (Rectified Linear Unit), pooling katmanlarında da max-pooling kullanılmaktadır [31]. Şekil 17’ de AlexNet’in mimarisi görülmektedir.



Şekil 17. AlexNet’in mimarisi

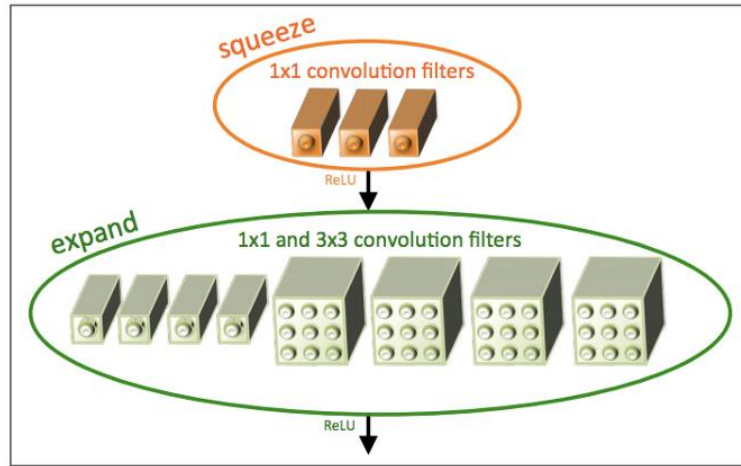
AlexNet’de yapılan testlerde eğitimin başarımı **%78,02** olmuştur. Şekil 18’de AlexNet ile gerçekleştirilen eğitimin detayları, başarımları ve hata grafikleri görülmektedir.



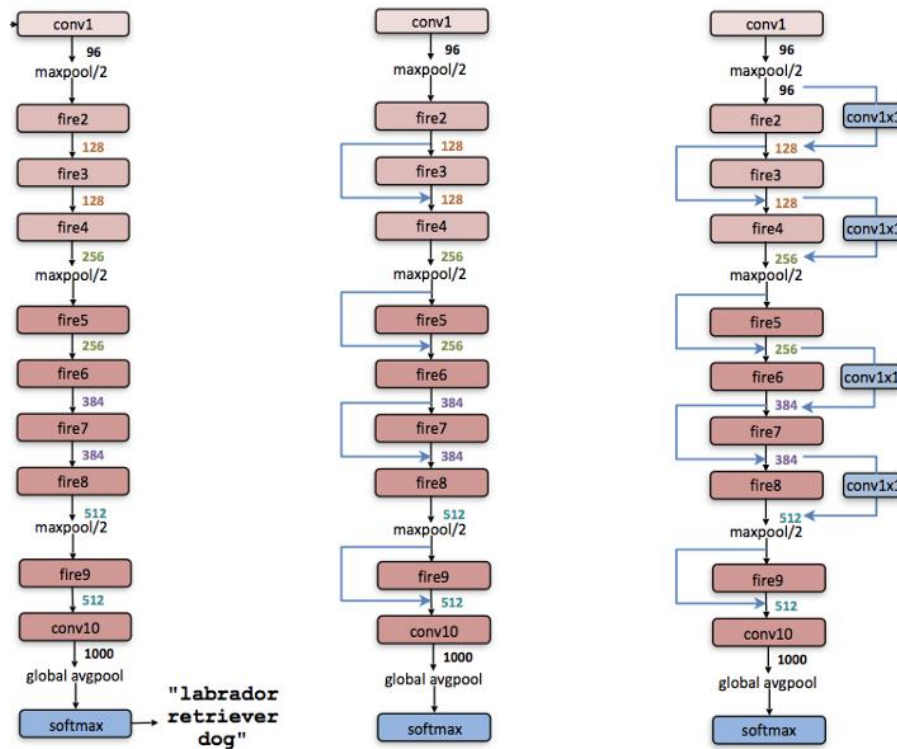
Şekil 18. AlexNet ile gerçekleştirilen eğitimin detayları, başarımları ve hata grafikleri

3.3.4. SqueezeNet ön-eğitilmiş ağı üzerinde yapılan testler

AlexNet’in geliştirilmesi ve modifiye edilmesi ile ortaya çıkan bir modeldir. AlexNet’in 1/510 ‘u kadar özellik hesaplatır. Şekil 19’da Squeeze işlemini yapan Fire Modülü isimli katman modeli görülmektedir. Filtreleri 1×1 ’e indirgeyerek hesaplanacak özellik miktarı azaltılmıştır [32]. Şekil 20’ de SqueezeNet Mimarisi görülmektedir.

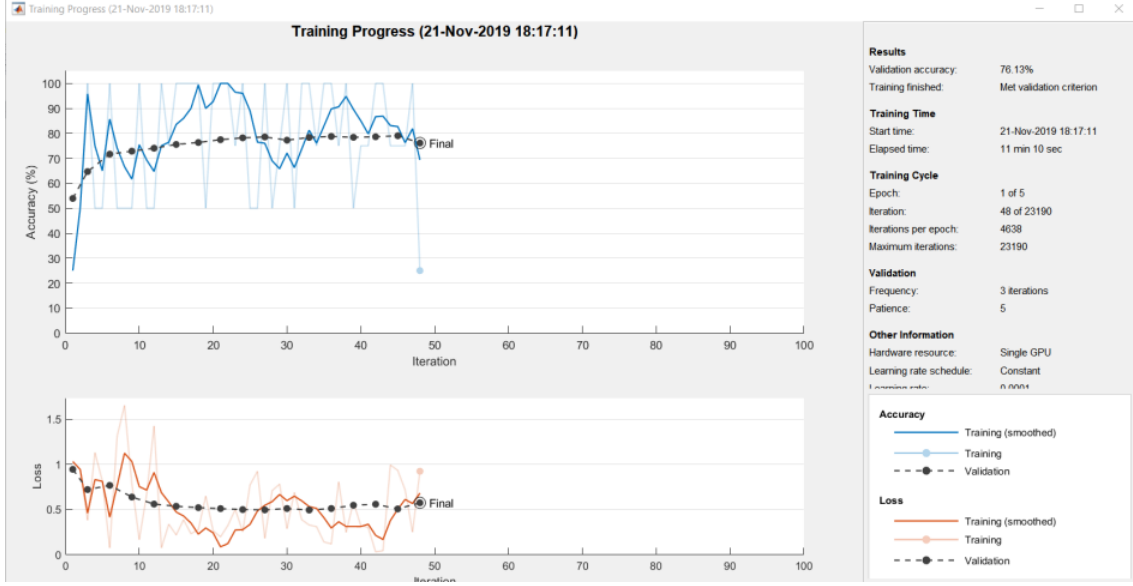


Şekil 19. SqueezeNet Fire modülü



Şekil 20. SqueezeNet mimarisi

SqueezeNet’de yapılan testlerde eğitimin başarımı **%78,02** olmuştur. Şekil 21’ de SqueezeNet ile gerçekleştirilen eğitimin detayları, başarıım ve hata grafiği görülmektedir.



Şekil 21. SqueezeNet ile gerçekleştirilen eğitimin detayları, başarımları ve hata grafiği

Tablo 1’de tüm ön-eğitilmiş ağlarda yapılan test çalışmalarında elde edilen sonuçlar sunulmuştur. Şu ana kadarki çalışmalarda, en yüksek başarımları ResNet50 ön-eğitilmiş ağı ile elde edilmiştir.

Tablo 1. Ön-eğitilmiş ağlarda elde edilen test başarımları

Ön-Eğitilmiş ağ mimarisi	Test Başarımları (%)
ResNet18	80.56
ResNet50	81.56
Inception V3	78.82
Alexnet	78.02
SqueezeNet	78.02

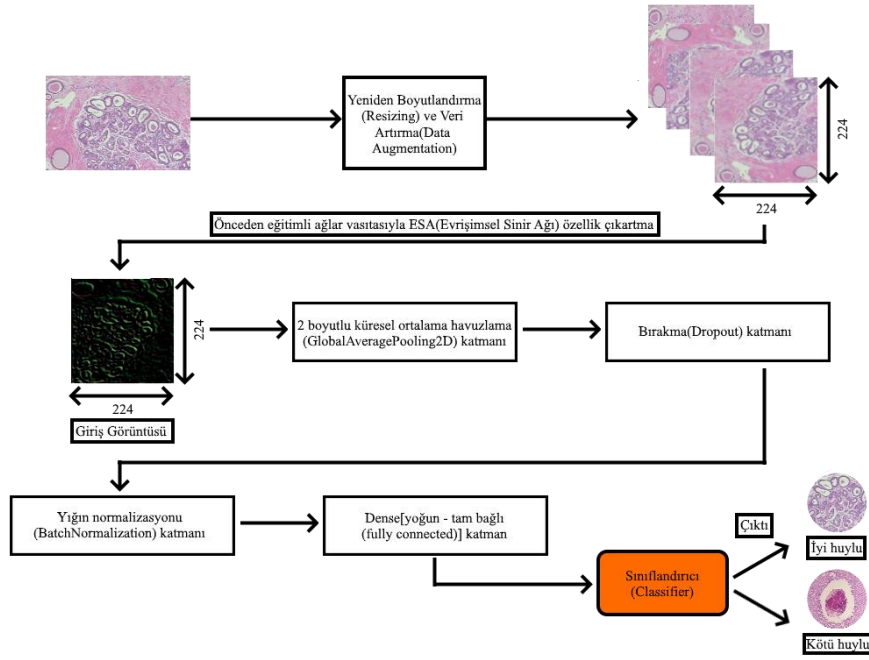
4. II. ARA RAPOR DÖNEMİNDE YAPILAN ÇALIŞMALAR

Projenin II. ara rapor dönemine kadar olan süreçte literatür tarama çalışmalarına devam edilmiş, kullanılan histopatolojik görüntü veriseti genişletilmiş, temin edilen malzemeler (bilgisayar) ile global veri setleri üzerinde Derin Öğrenme yöntemleri ve ön-eğitilmiş ağlar kullanılarak test çalışmaları artırılmış ve yüksek sınıflandırma sonuçları elde edilmiştir.

4.1. Metodoloji

Proje kapsamında bu ara rapor döneminde histopatolojik görüntüler üzerinde meme kanserinin tespiti için önerilen derin sinir ağlarına dayalı modelin genel blok diyagramı Şekil 22'de gösterilmiştir.

Önerilen modelde görüntü alındıktan sonra ilk olarak, derin sinir ağları için oldukça önem arz eden yeniden boyutlandırma işlemi gerçekleştirilmektedir. Burada derin sinir ağlarının aldığı girdilerin aynı boyutta olması gerekmektedir ve bu nedenle girdiler CNN (Convolutional Neural Network) mimarisine girmeden önce sabit bir boyuta getirilir. Bunun yanında yapılan çalışmaya göre yüksek çözünürlüğe sahip bir görüntü bellekte daha fazla yer kaplar ve daha büyük bir sinir ağı kullanılmasına da neden olur [33]. Yine aynı çalışmaya göre eğer görüntünün çözünürlüğü çok fazla küçültülürse detaylar ve görüntü de bulunan özellikler kaybolabilir. Bu gibi sebeplerden ötürü bu çalışmada yeniden boyutlandırma işlemi yatayda ve düşeyde 224x224 olarak uygulanmıştır. Bunun yanında, veri artırma ile veriseti daha fazla genişletilerek ağa daha fazla örnek sunulmuş ve ağın eğitimi yürütülmüştür. Sonraki aşamada görüntüye uygulanan özellik çıkartma ile ayırt edici özellikler belirlenmektedir. Havuzlama (Pooling) işlemi yapılarak sınıflandırmadaki görüntünün şekli herhangi bir şekilde değişse dahi, sistemin bunu tanıyabilmesi ve mevcut görüntünün özelliklerini kaybetmeden daha küçük boyutta bir matris elde edilmesi amaçlanmıştır (Tolias, Sicre, & Jégou, 2015). Ek olarak, ezberleme (Over-fitting) işleminden kaçınmak amacıyla bırakma (dropout) işlemi gerçekleştirilmiştir. Yığın normalizasyonu (batch normalization) derin sinir ağlarının hızını, performansını ve sürekliliğini arttıran bir yöntemdir [34] [35]. Son olarak görüntüler sınıflandırılmak için tam bağlantılı katmandan geçer ve iyi huylu veya kötü huylu meme kanseri sınıflandırma işlemi tamamlanır.



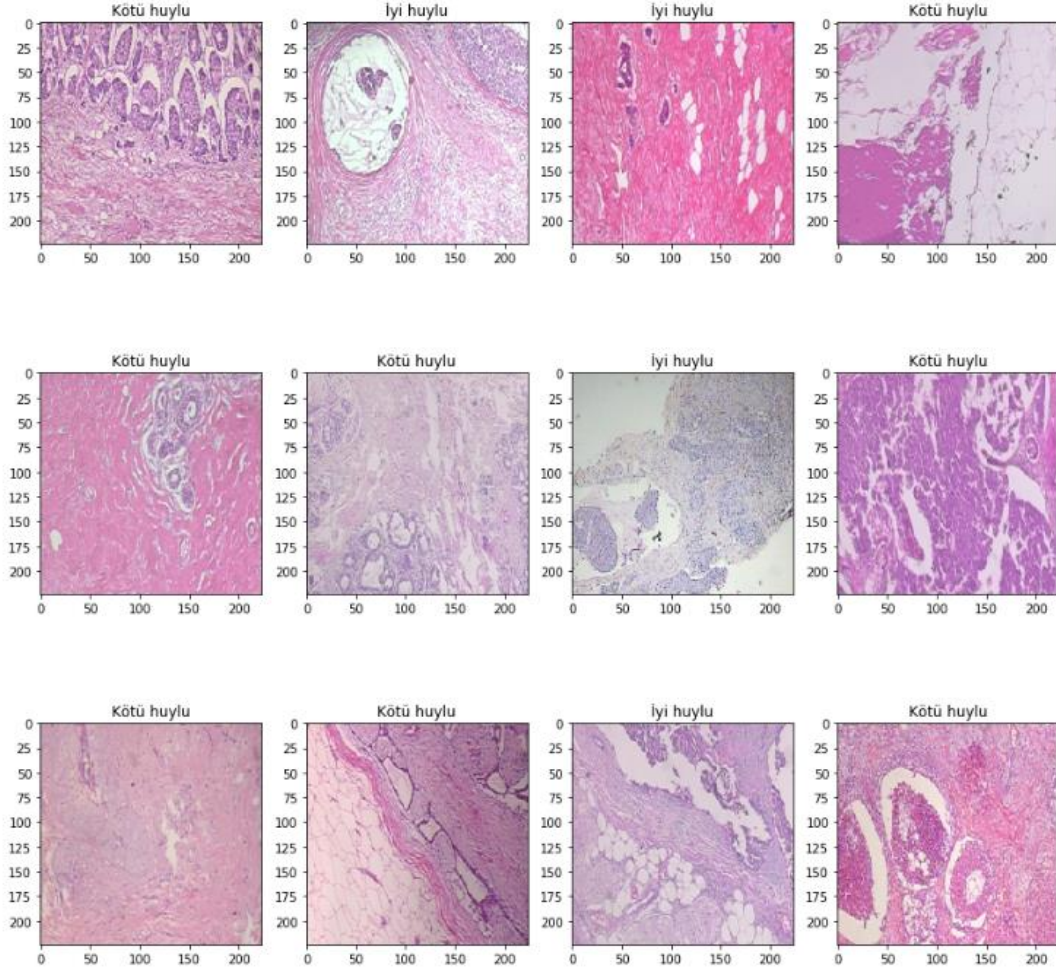
Şekil 22. Histopatolojik görüntülerde meme kanserinin tespiti ve sınıflandırılması için önerilen derin sinir ağları tabanlı modelin blok diyagramı

4.2. Veri Setinin Hazırlanması

Proje kapsamında önerilen modelin performansını değerlendirmek amacıyla oldukça fazla ve kaliteli histopatolojik görüntü içeren BreakHis veriseti kullanılmıştır (Fabio A Spanhol, Oliveira, Petitjean, & Heutte, 2015). Bu verisinde 82 hastadan toplanan toplamda 7909 mikroskopik görüntü bulunmaktadır. 2480 iyi huylu ve 5429 kötü huylu görüntüden oluşan bu veriseti 700x460 çözünürlüğe ve 3 kanal RGB – 8 bit derinliğe sahip PNG formatındadır. Bu veri seti P&D laboratuvarı-Patolojik Anatomi ve Sitopatoloji tarafından Parana-Brezilya’da oluşturulmuştur. Veriseti Tablo 2’de gösterildiği üzere 40X, 100X, 200X ve 400X büyütme oranlarına sahip görüntülerden oluşmaktadır. Şekil 23’te bu veri setinde bulunan iyi huylu ve kötü huylu görüntülerin 224x224 çözünürlükteki bazı örnekleri gösterilmiştir (Fabio Alexandre Spanhol, Oliveira, Petitjean, & Heutte, 2016) (Fabio A Spanhol, Oliveira, Cavalin, Petitjean, & Heutte, 2017).

Tablo 2. BreakHis veri setindeki yakınlaştırma oranları ve görüntü sayıları

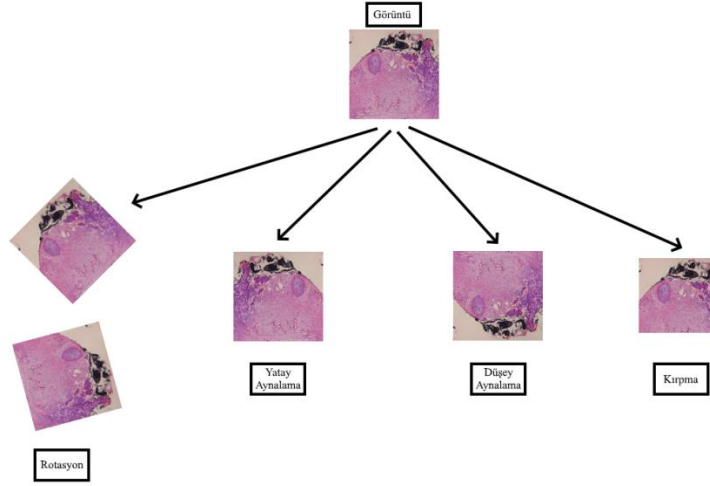
Yakınlaştırma Oranı	İyi Huylu Görüntü Sayısı	Kötü Huylu Görüntü Sayısı	Toplam
40X Oranında	652	1370	1995
100X Oranında	644	1437	2081
200X Oranında	623	1390	2013
400X Oranında	588	1232	1820
Görüntü Toplamı	2480	5429	7909



Şekil 23. BreakHis verisetinden iyi huylu ve kötü huylu örnek histopatolojik görüntüler

4.3. Veri Artırma(Data Augmentation) İşlemi

Derin öğrenme ve makine öğrenmesi algoritmalarında karşılaşılan en büyük sorun ağı eğitecek yeterince veri olmamasıdır. Yeterince veri bulunmaması bu algoritmalarda sık sık ortaya çıkan ve büyük bir sorun olan ezberleme (over-fitting) problemini meydana getirir. Bu olay kısaca ağın gördüğü eğitim verilerini (görüntü, text vb.) ezberlemesi ve eğitim verileri dışında bir girdi ile karşılaştığında başarı gösterememesidir. Bu sorundan kurtulmanın en önemli yöntemlerinden birisi veri artırımı yapılmasıdır. Bu yöntem eğitim setine uygulanır ve elde bulunan verilerin özellikleri değiştirilerek yapay (artificial) olarak daha fazla görüntü elde edilir. Proje çalışması kapsamında kullanılan verisinde de veri artırımı yapılarak görüntüler yatay ve dikey olarak döndürülmüş, görüntülere rastgele yakınlaştırma değerleri uygulanmıştır [36] [37]. Şekil 24’te uygulanan veri artırma işlemlerinden bazıları şekilsel olarak gösterilmiştir. Veri artırma işleminin test setine de uygulanması işlemine Test-zamanı-artırması(Test time augmentation - TTA) ismi verilmiştir [38].

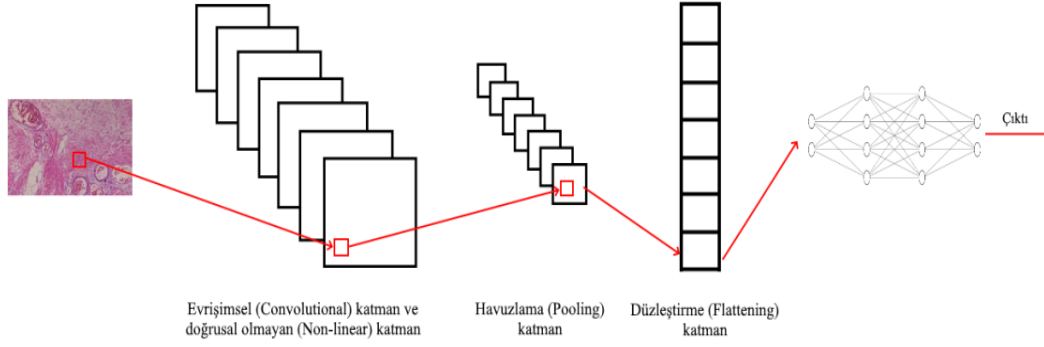


Şekil 24. Örnek bir histopatolojik görüntüye veri artırma (data augmentation) işlemlerinin uygulanması

4.4. Evrişimsel Sinir Ağı (CNN – Convolutional Neural Network) ve Ön-Eğitilmiş Ağlar

Derin öğrenmede ki temel amaç yüksek-boyutlu (high-dimensional) verilerden bazı özellikler çıkartmak ve girdiler ile çıktılarını bu özellikler vasıtasıyla ilişkilendirerek bir sonuca varmaktır. Karmaşık özellikleri çıkartabilmek için derin öğrenme algoritmaları genelde çok katmanlı bir yapıya sahiptir.

Evrişimsel sinir ağları derin öğrenmede kullanılan bir ağ modelidir. Şekil 25'te gösterildiği üzere görüntüdeki pikseller öncelikle evrişimsel katman vasıtasıyla özellik çıkartma (feature extraction) işlemine tabi tutulur. Özellik çıkartmada belirli bir filtre (maske) tüm görüntü üzerinde dolaştırılır ve gerçekten iyi huylu veya kötü huylu görüntünün pikseli ile oluşturulan yeni görüntü arasında belirli bir olasılıkta benzerlik kurulur. Daha sonra sisteme doğrusal olmayanlık (non linear – çıktıdaki değişimin girdinin değişimi ile orantılı olmadığı sistemler) gösterilir. Doğrusal olmayan katman aktivasyon katmanı olarak da anılır çünkü burada bir aktivasyon fonksiyonu kullanılır. Doğrusallığın problem olmasının nedeni basitçe tüm ağın tek bir algılayıcı (perceptron) gibi davranmasının önüne geçilmesidir. Bu çalışmada, aktivasyon fonksiyonu olarak softmax kullanılmıştır. Havuzlama katmanında ağırlık sayısı azaltılır, uygunluk kontrol edilir. Havuzlama işlemindeki amaç mevcut resmin özelliklerini kaybetmeden daha küçük boyutta bir matris (görüntü) elde etmektir. Havuzlama işleminde de özellik çıkartmadakine benzer bir filtreleme uygulanır. Düzleştirme katmanında ise klasik sinir ağı yapısı için veriler hazırlanır ve sunulur. Bu katmandaki temel amaç verileri tam bağlı katman (fully connected layer) için hazırlamaktır. Son olarak tam bağlı katmanda ise yapay sinir ağları vasıtasıyla öğrenme sağlanır [39] [40] [41] [42].



Şekil 25. Çalışmada kullanılan evrişimsel sinir ağı mimarisinin genel yapısı

Ön-eğitilmiş ağlar CNN modellerinde öğrenmenin daha hızlı sağlanması, işlemlerin daha hızlı yapılması ve doğruluk başarımının artırılması için oldukça önemlidir. Bu ağlar benzer bir problem üzerinde önceden kullanılarak buna bağlı ağırlıklar alan modeller olarak bilinir. Her bir problem için tam olarak en uygun çözüm olmasalar dahi aynı işlemleri tekrar tekrar yapmanın önüne geçerler. Bu ağlar genellikle ImageNet ismi verilen ve görsel nesne tanıma üzerine kurulmuş çok büyük bir veritabanı üzerinde önceden kullanılmıştır. Bu çalışmada da ön-eğitilmiş olan ResNet50, DenseNet201, InceptionV3 ve Xception ağları kullanılmıştır.

Residual Network (ResNet) artık değerlerin (residual value) ve artık blokların (residual block) modele eklenmesiyle meydana gelir. Her iki katmanda bir sisteme farklı bir değer ekleyerek gradyan kaybolmasının (vanishing gradient) önüne geçer [43].

ResNet'e benzer bir ön-eğitilmiş ağ olan DenseNet'te ResNet'ten farklı olarak 2 katmanda bir sisteme eklenen değer kendinden sonraki tüm katmanlara eklenmiştir. Bu sistem modele birçok avantaj kazandırmıştır. Bu sayede ağ daha kolay eğitilebilir hale gelmiş ve yüksek katmanlı ağlarda çoğu katmanın işlevsiz olma sorununu optimize etmiştir [44].

Inception ağı, ağ içinde ağ olarak tanımlanabilir. Evrişim katmanlarındaki filtreleme işlemlerinin ve havuzlama işleminin eş zamanlı olarak gerçekleştirilmesi temeline dayanır. Modüller halinde işlemler yapmaktadır. InceptionV3 ağı ise ek modüllerin küme normalizasyonu (batch normalization) ve full connected layeri'da yardımcı sınıflayıcı olarak ağın içinde kullanır [45] [30].

Xception ağı ise, temelde InceptionV3 ağının üzerine ekleyerek gelişen bir ağ olmuştur. Diğer ağlara göre başarıyı yakalamasındaki en önemli etken ise evrişimsel (convolutional) kısımdaki yaptığı farklılıklardır. Normal bir ağ evrişimsel bölümde bir filtreyi en, boy ve derinlik gibi çok boyutlu matrisler üzerinde gezdirerek işlemlerini meydana getirmektedir.

Xception ağı bu kısımda InceptionV3'teki gelişmelere ek olarak 2 farklı yaklaşım sunmaktadır. Bunlar, akıllı derinlik evrişimi (depthwise convolution) ve akıllı nokta evrişimi (Pointwise convolution)'dir. Akıllı derinlik evrişimi bölümünde her kanalda değilde sadece tek bir kanalda işlem yaparak sonuca ulaşır. Örneğin RGB değerlerine sahip bir görüntüde 3 kanalı değil 1 kanalı kullanır. Bu birçok özellik kaybına ve oldukça başarısız sonuçlara neden olacağından akıllı nokta evrişiminde 1 kanal üzerinde işlem yapılarak elde edilen görüntü üzerinde klasik evrişim işlemi $1 \times 1 \times \text{Kanal Sayısı}$ olarak uygulanır ve sonuca ulaşılır [46].

4.5. Bırakma(Dropout) İşlemi

Derin öğrenme metodlarında modellerin birleştirilerek (kombinasyon) kullanılması pekçok durumda daha iyi bir sonuç verir. Bununla birlikte büyük ağların eğitimi çok fazla hesaplama gerektirir ve en iyi sonuca ulaşmak için gerekli parametreleri bulmak oldukça zorlu bir işlemdir. Ayrıca gelişmiş ağların eğitilmesi için oldukça fazla veriye ihtiyaç duyulmaktadır ve farklı ağların farklı veri setleri ile eğitilmesi için yeterli düzeyde veri bulunamayabilir. Bir şekilde bu sorunlar çözülerek modeller birleştirilerek kullanılsa dahi hepsini test anında kullanmak hızın önemli olduğu durumlarda mümkün değildir. Veri artırma işlemi modellere uygulanmasına karşın ezberleme problemi de kimi durumlarda tam olarak çözülememektedir. Bu problemler ile başa çıkmak için bırakma yöntemi önem arz etmektedir. Bırakma işleminde temel işlem eğitim sırasında nöronları (bağlantıları ile birlikte) modelden (sinir ağı) rastgele olarak düşürmektir. Bu işlem yalnızca eğitim safhasında uygulanır test aşamasında uygulanmaz. Ayrıca bırakma işlemi pekçok farklı modele uygulanabilmektedir [47].

4.6. Proje kapsamında yapılan deneysel çalışmalar

Bu proje döneminde ResNet50, DenseNet201, InceptionV3 ve Xception olmak üzere dört farklı ön-eğitilmiş ağ kullanılmış ve gerçekleştirilen deneysel çalışmalar ile her birisi ile elde edilen sonuçlar kendi arasında karşılaştırılmıştır. Ulaşılan sonuçlar histopatolojik görüntüler üzerinde meme kanserinin tespiti ve sınıflandırılmasında Xception ağının daha başarılı olduğunu göstermektedir. Nadiren de olsa, DenseNet201 ağı Xception ile benzer ve hatta iyi sonuçlar elde etse de bu durum bir süreklilik arz etmemektedir.

Bu çalışma da önerilen model eğitilirken Proje kapsamında temine dilen bilgisayar kullanılmıştır. Yapılan çalışma da kullanılan görüntülerin %80'i eğitim için ve kalan %20'si

test için kullanılmıştır. Deneysel çalışmalar Python programlama dili ve Keras kütüphanesi ortamı kullanılarak gerçekleştirilmiştir.

Kullanılan veri setine göre görüntünün gerçek niteliğinin (GGN) kötü huylu ve model tarafından tahmin edilen niteliğin (MTTEN) de kötü huylu olduğu durum Doğru Pozitif (DP), GGN iyi huylu ve MTTEN kötü huylu olduğu durum Yanlış Pozitif (YP), GGN kötü huylu MTTEN iyi huylu olduğu durum Yanlış Negatif (YN) ve GGN iyi huylu MTTEN’de iyi huylu olduğu durum Doğru Negatif (DN) olarak isimlendirilir. Başarım ölçümleri için kullanılan doğruluk, hassasiyet, hatırlama ve F1 skoru performans ölçütleri aşağıda gösterilmiştir [48] [49]. Burada dikkat edilmesi gereken durum hassasiyet ve hatırlama değerlerinin her bir sınıf için hesaplanması gerektiğidir. Bu çalışma da iyi huylu ve kötü huylu olarak 2 sınıf bulunmaktadır.

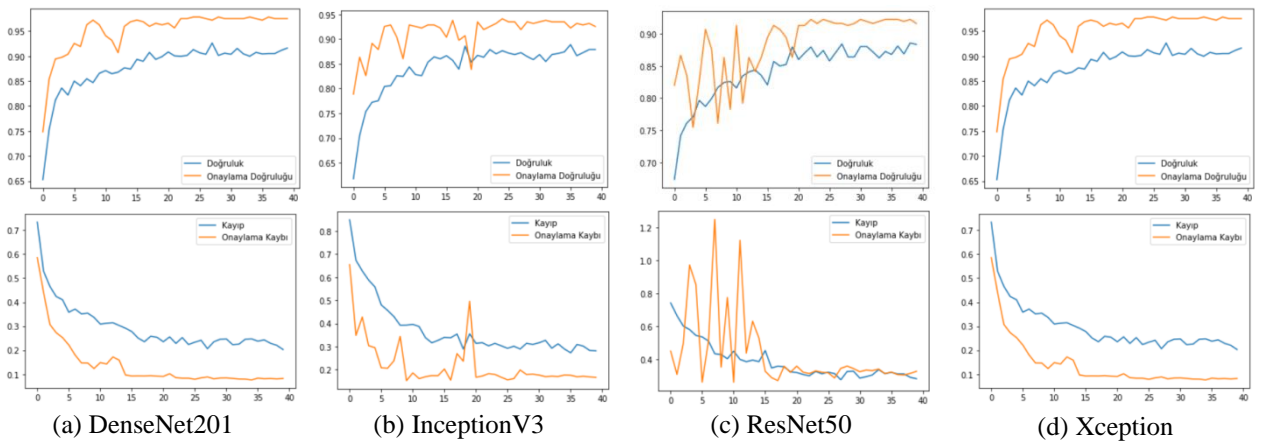
$$\text{Doğruluk(Accuracy)} = \frac{DP + DN}{(DP + DN + YP + YN)}$$

$$\text{Hassasiyet(Precision)} = \frac{DP}{(DP + YP)}$$

$$\text{Hatırlama(Recall)} = \frac{DP}{(DP + YN)}$$

$$\text{F1 Skor} = \frac{2 * (\text{Hassasiyet} * \text{Hatırlama})}{(\text{Hassasiyet} + \text{Hatırlama})}$$

Çalışmada kullanılan tüm ağlar 40 epok boyunca eğitilmiş, doğruluk, onaylama doğruluğu, kayıp ve onaylama kaybı değerlerini gösteren grafikler Şekil 26’da gösterilmiştir. Bu grafikler veri setindeki 200X yakınlaştırma oranına sahip görüntülerin sonuçlarını göstermektedir. Ağların tüm yakınlaştırma oranlarındaki detaylı analizleri ise Şekil 26’da gösterilmiştir. Elde edilen verilere göre Xception ağı ile diğer ağlara göre daha başarılı sonuçlar elde edilmiştir.



Şekil 26. DenseNet201, InceptionV3, ResNet50 ve Xception önceden eğitilmiş ağlarının 200X yakınlaştırma oranındaki başarı grafikleri

Tablo 2’de sunulan sonuçlar, bu çalışma da kullanılan tüm önceden eğitilmiş ağların 40X, 100X, 200X ve 400X yakınlaştırma oranlarına sahip histopatolojik görüntüler üzerindeki eğitimlerinin sonuçlarını özetler niteliktedir. Doğruluk skoru değerleri ağların genel olarak başarımlarını göstermektedir. Bu değerlere göre ağların farklı yakınlaştırma değerlerindeki doğruluk skorları oldukça farklılık göstermektedir. Örneğin ResNet50 ağı 200X yakınlaştırma oranında 95.54’lük bir başarı elde ederken 400X yakınlaştırma oranında bu değer 90.43’e gerilemiştir. Bütün ağların doğruluk skorları genel olarak karşılaştırıldığında ise Xception ağının göstermiş olduğu başarı açık bir şekilde görülmektedir. Özellikle 200X yakınlaştırma oranında Xception 98.01 yani %98.01’lik bir başarı ortaya koymuştur. Bu tablodaki makro ortalama ve ağırlıklı ortalama değerleri her bir performans ölçütü için ayrı ayrı hesaplanır. Eşitlik (1), Eşitlik (2) ve Eşitlik (3)’te sırasıyla hassasiyet değeri, hatırlama değeri ve bu iki değerlerin harmonik ortalaması olan F1 skor değerinin makro ortalamaları gösterilmiştir. Eşitlik (4), Eşitlik (5) ve Eşitlik (6)’da ise hassasiyet değeri, hatırlama değeri ve F1 skor değerinin ağırlıklı ortalamaları gösterilmiştir. Eşitliklerdeki n değeri veri setindeki sınıf sayısını; P_1, P_2, \dots, P_n dizisi ise her bir sınıf için hesaplanan hassasiyet değerini; R_1, R_2, \dots, R_n ise her bir sınıf için hesaplanan hatırlama değerini göstermektedir [50] [51] [52] [53].

$$\text{Hassasiyet Değeri (Makro Ortalama)} = \frac{\sum_{i=1}^n P_i}{n} \quad (1)$$

$$\text{Hatırlama Değeri (Makro Ortalama)} = \frac{\sum_{i=1}^n R_i}{n} \quad (2)$$

$$\text{F1 Skor Değeri (Makro Ortalama)} = \frac{2}{\frac{1}{\frac{\sum_{i=1}^n P_i}{n}} + \frac{1}{\frac{\sum_{i=1}^n R_i}{n}}} \quad (3)$$

$$\text{Hassasiyet Değeri (Ağırlık. Ort.)} = \frac{\sum_{i=1}^n |y_i| \frac{dp_i}{dp_i + yp_i}}{\sum_i^n |y_i|} \quad (4)$$

$$\text{Hatırlama Değeri (Ağırlık. Ort.)} = \frac{\sum_{i=1}^n |y_i| \frac{dp_i}{dp_i + yn_i}}{\sum_i^n |y_i|} \quad (5)$$

$$\text{F1 Skor Değeri (Ağırlık. Ort.)} = \frac{\sum_{i=1}^n |y_i| \frac{2dp_i}{2dp_i + yn_i + yp_i}}{\sum_i^n |y_i|} \quad (6)$$

Tablo 3’te gösterilen F1 skor ve doğruluk skor değerleri ağların başarımlarının değerlendirilmesinde kullanılan ölçütlerdir. Bu ölçütlerin hesaplanma yöntemleri birbirinden farklılık gösterebilir dahi ikisi de performans ölçütü olarak kullanılabilir ve F1 skor, hassasiyet, hatırlama ve doğruluk değerlerinin hesaplanma yöntemleri dikkate alındığında doğruluk değerinin gerçek pozitif ve gerçek negatif değerleri önem arz ettiğinde F1 skor değerinin ise yanlış negatif ve yanlış pozitif değerlerinin önem arz ettiği durumlarda kullanılmasının daha uygun olduğu söylenebilir. Yine Tablo 3’deki doğruluk ve F1 skor değerleri dikkate alındığında bu ağlar arasındaki kıyaslama Xception -> DenseNet201 -> InceptionV3 ve ResNet50 şeklinde sıralanabilir. Bu tablodan çıkarılabilecek bir diğer husus ağların yakınlaştırma oranlarına göre verdikleri sonuçların durumlara göre farklılık göstermesidir. Örneğin InceptionV3 ağı 400X yakınlaştırma oranında 92.34 gibi nispeten

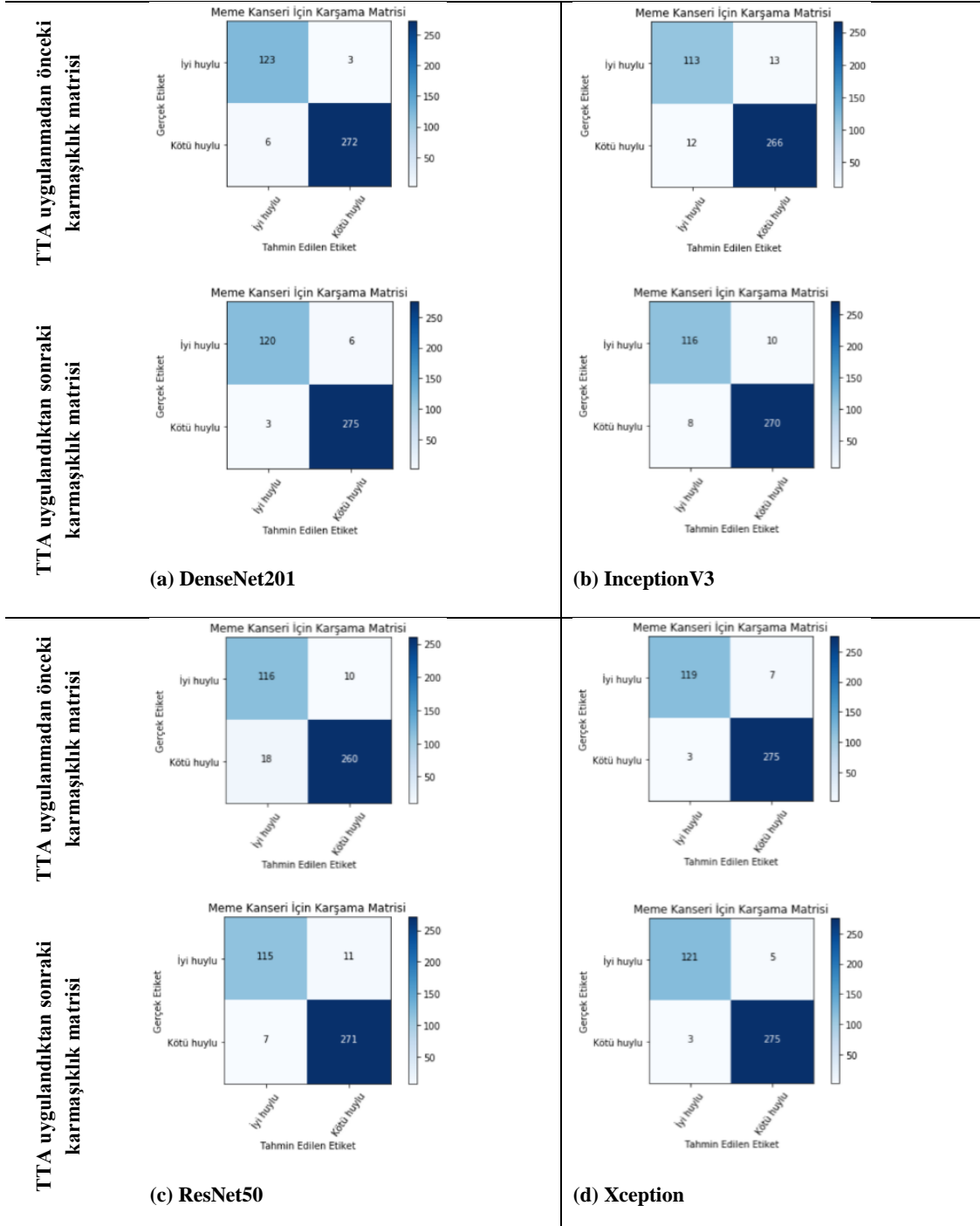
düşük bir doğruluk skoru değerine ulaşırken 200X değerinde 95.54 gibi oldukça yüksek bir değere ulaşmıştır. Bu da göstermektedir ki veriseti (farklı yakınlaştırma oranları), seçilen ağ, kullanılacak parametreler bunların tamamı başarımlarına sonucuna etki etmekte ve buna dayalı olarak seçimler yapılmalıdır.

Tablo 3. Bütün ağlarda elde edilen başarımların karşılaştırılması (D.S.: Doğruluk Skoru, Has. D.: Hassasiyet Değeri, Hat.D.: Hatırlama Değeri, M.O.: Makro Ortalama, A.O.: Ağırlıklandırılmış Ortalama)

Ağ Adı	Yakınlaştırma	D.S.	Has.D. (M.O.)	Hat.D. (M.O.)	F1 Skoru (M.O.)	Has.D. (A.O.)	Hat. D. (A.O.)	F1 Skoru (A.O.)
DenseNet201	40X	97.50	96.92	97.31	97.11	97.51	97.50	97.50
	100X	94.49	94.41	92.63	93.51	94.48	94.49	94.49
	200X	97.77	97.71	97.07	97.39	97.77	97.77	97.77
	400X	95.08	94.20	94.59	94.40	95.10	95.08	95.09
InceptionV3	40X	94.75	93.55	94.45	93.99	94.83	94.75	94.79
	100X	95.21	94.96	93.78	94.37	95.19	95.21	95.20
	200X	95.54	94.98	94.59	94.79	95.53	95.54	95.53
	400X	92.34	92.15	90.13	91.13	92.32	92.34	92.33
ResNet50	40X	93.25	93.66	90.57	92.09	93.31	93.25	93.28
	100X	93.30	94.28	90.07	92.13	93.50	93.30	93.40
	200X	95.54	95.18	94.37	94.77	95.52	95.54	95.53
	400X	90.43	90.90	86.94	88.88	90.54	90.43	90.48
Xception	40X	97.25	97.11	96.49	96.80	97.24	97.25	97.24
	100X	94.97	94.39	93.82	94.10	94.95	94.97	94.96
	200X	98.01	97.89	97.47	97.68	98.01	98.01	98.01
	400X	95.35	94.16	95.46	94.80	95.49	95.35	95.42

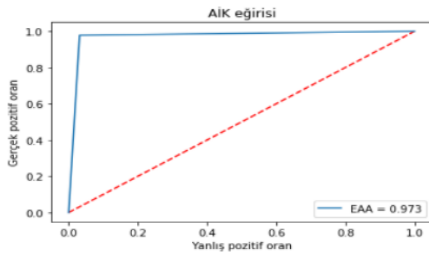
Xception ağının en başarılı olduğu 200X yakınlaştırma oranı için Xception ve diğer ağların karmaşıklık matrisi TTA işlemi uygulanmadan önce ve uygulandıktan sonra Şekil 27’de gösterilmiştir. Karmaşıklık matrislerinde ağların veri setinde yapılan ayırım sonucu görüntülere verdikleri tahminler ve görüntülerin gerçek değerleri gösterilmektedir. Xception ağının karmaşıklık matrisi incelenecek olursa (TTA işleminden sonra) toplam 404 görüntü üzerinde gerçekten iyi huylu olan 126 görüntü ve gerçekten kötü huylu olan 278 görüntü bulunmaktadır. Xception ağı iyi huyluların 121 tanesini doğru (iyi huylu tahmini yapmış) ve 5 tanesini yanlış (kötü huylu tahmini yapmış) tahmin etmiş, kötü huyluların ise 275 tanesini doğru (kötü huylu tahmini yapmış) ve 3 tanesini yanlış (iyi huylu tahmini yapmış)

değerlendirmiştir. Diğer ağların karmaşıklık matrisleri de bu şekilde yorumlanabilir ve bu tahminler arasında Xception ağının çok daha başarılı olduğu görülür. Öte yandan TTA işleminin ağlara olan etkileri incelendiğinde toplam hatalı tahmin sayısı: Xception için 10'dan 8'e, ResNet50 için 20'den 18'e, InceptionV3 için 25'ten 18'e düşmüş ve DenseNet201 ağı için de olumlu veya olumsuz bir sonuç doğurmamıştır. Bu bilgiler ışığında TTA işlemi üç ağı olumlu yönde etkilemiş bir ağı ise ne olumlu ne de olumsuz olarak etkilememiştir. Buna göre TTA işleminin genel anlamda ağların başarımını olumlu yönde etkilediği söylenir.

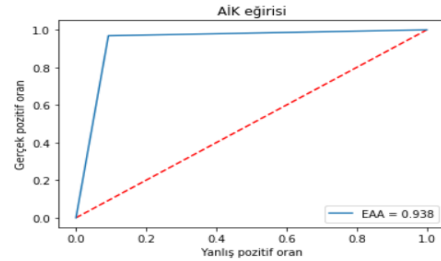


Şekil 27. 200X yakınlaştırma oranındaki görüntülerde ağların TTA'lı ve TTA'sız karmaşıklık matrisleri

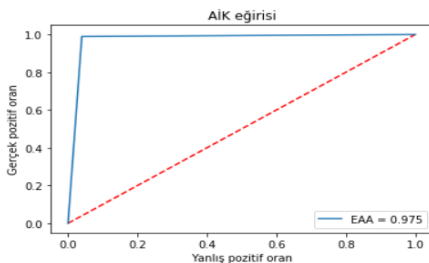
Makine öğrenmesi uygulamalarında performans değerlendirmesi için en yaygın olarak kullanılan yöntemlerden birisi de AİK [Alıcı İşletim Karakteristiği – ROC (Receiver Operating Characteristic)] eğrileridir. AİK grafikleri modelin tahminlerinde ne kadar başarılı olduğunu gösterir. AİK eğrisinin altında kalan alan ise EAA [Eğri altındaki alan – AUC (Area under curve)] olarak adlandırılır. EAA değeri bir alan kapsamıyla gösterilir ve bu alan ne kadar geniş ise modelin sınıflandırma da elde ettiği başarı da o kadar yüksek olarak tanımlanır. EAA değeri için optimum değer 1 olarak ifade edilir [54] [55]. Şekil 28’de DenseNet201 ağıının 40X ve Xception ağıının 100X, 200X ile 400X yakınlaştırma oranlarındaki AİK grafikleri ve EAA değerleri gösterilmiştir. EAA değeri, doğruluk değeri, f1 skor değeri hepsi ağların başarımlarının değerlendirilmesi için kullanılan ölçütlerdir. Kimi çalışmalara göre EAA değeri doğruluk değerine göre daha güvenilir ve ağıın başarımını ölçmede doğruluk değerine göre daha doğru sonuçlar vermektedir [56] [57]. Şekil 28’de yer alan değerlere göre DenseNet201 ağıının 40X yakınlaştırma oranındaki EAA değeri (0.973) Xception ağıının EAA değerine (0.965) göre daha yüksektir. Ayrıca 40X’te F1 skor (A.O.) değerinde de DenseNet201 daha başarılı bir sonuç ortaya koymuştur. Bu da sınıflandırma da DenseNet201’in daha başarılı olduğu anlamına gelmekte ve hatta doğruluk değerine göre de DenseNet201 ağı daha başarılıdır. Bu durumda 40X yakınlaştırma oranı özelinde DenseNet201 ağı bu dört ağ arasındaki en başarılı ağıdır. Buna karşın 40X harici bütün yakınlaştırma oranları incelendiğinde diğer tüm yakınlaştırma oranlarında Xception ağıının EAA ve doğruluk değerinin daha yüksek olduğu görülmektedir. Sonuç olarak 40X yakınlaştırma oranında DenseNet201, 100X, 200X ve 400X yakınlaştırma oranında Xception ağı diğer ağlara oranla daha başarılıdır. Bu iki ağ arasında ise Xception ağıının daha başarılı olduğu söylenir.



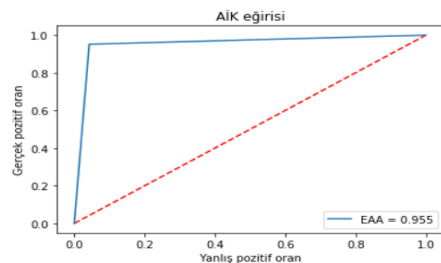
(a) DenseNet201 40X



(b) Xception 100X

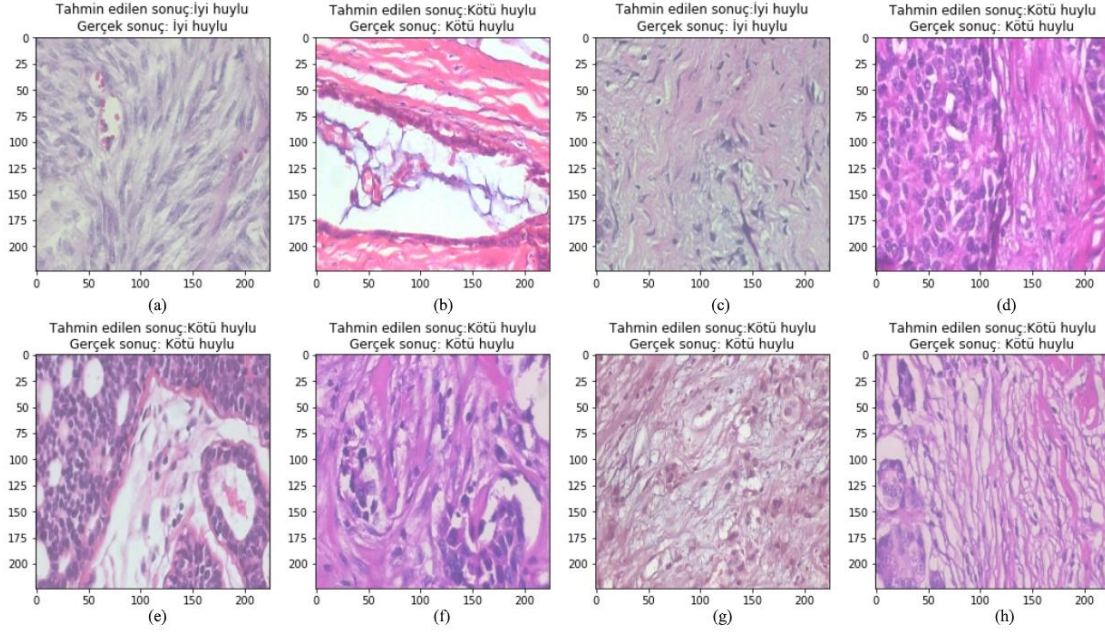


(c) Xception 200X



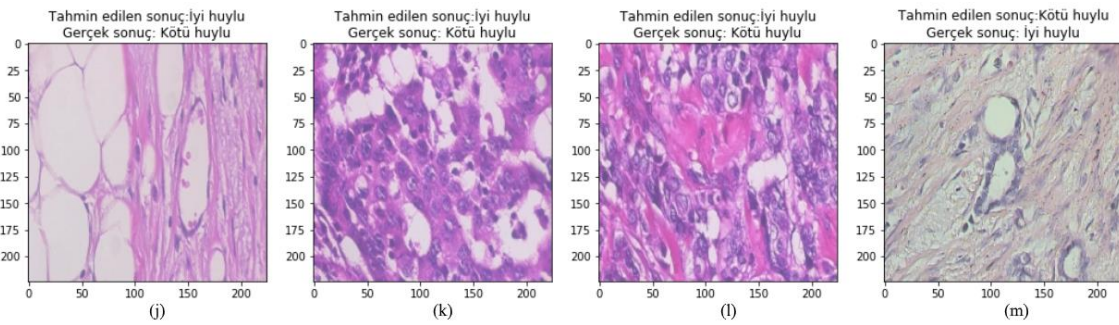
(d) Xception 400X

Şekil 28. Yakınlaştırma oranlarına göre ağların AİK eğrilerini ve EAA değerleri



Şekil 29. Xception ağının 200X yakınlaştırma oranındaki tahminleri ve gerçek sonuçlar.

Şekil 29’da rastgele seçilen 8 adet 200X yakınlaştırma oranındaki histopatolojik görüntü için Xception ağının yapmış olduğu tahminler ve bu görüntülerin gerçek değerleri gösterilmektedir. Bu görüntüler resizing(yeniden boyutlandırma) işlemi ile 224x224 konumuna getirilmiş ve ağın bu görüntüler için bir tahmin yapması istenmiştir. Görüldüğü üzere b, d, e, f, g ve h’de gerçek sonuçlar kötü huylu ve ağın tahmin ettiği sonuçta kötü huyludur. a ve c’de ise gerçek sonuçlar iyi huylu ve ağın tahmin ettiği sonuçta iyi huyludur. Bu 8 resim içinde Xception ağı doğru tahmin yapmıştır. Buna karşın Şekil 30’da gösterildiği üzere j, k ve l’de gerçek sonuç kötü huylu iken Xception iyi huylu tahmini yapmıştır. Son olarak m’de ise gerçek sonuç iyi huylu iken Xception kötü huylu tahmini yapmıştır. Bu 4 resimde ise Xception yanılmıştır.



Şekil 30. Xception ağının 200X yakınlaştırma oranındaki tahminleri ve gerçek sonuçlar.

Bu proje dönemindeki çalışmalar kapsamında verisetinde dört farklı yakınlaştırma oranlarında yapılan deneysel çalışmalarda, ağlar yakınlaştırma oranından etkilenmiş ve yapı aynı olsa dahi farklı başarımlarına ulaşmışlardır. Bu da göstermektedir ki veri setlerinin farklı olması bir kenara aynı görüntüde ki farklı yakınlaştırma oranları dahi ağların başarımlarını etkilemektedir. Eğitim sırasındaki başarımlar daima onaylama aşamasındakine göre daha düşük kalmıştır. Bu başlangıçta kötü bir durummuş gibi gözükse de aslında oldukça faydalı

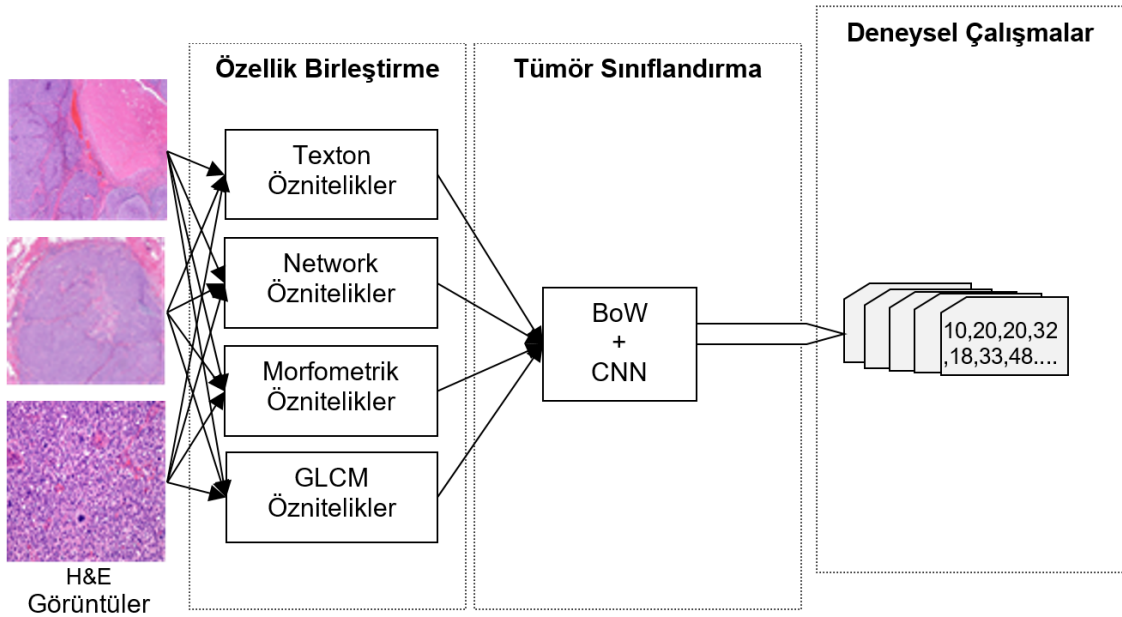
hatta neredeyse her durumda istenilen bir şeydir çünkü ağın son hali onaylama işlemi sırasında ki ağıdır ve son hali o parametrelerde bulunan ağ olacaktır. Bunun yanı sıra eğitim sırasındaki başarı oranının daha düşük olmasının sebebi ezberleme işlemine de engel olan bırakma değeridir. Bırakma değeri eğitim sırasında bazı birimleri yok sayarak parametrelerin gelişmesini sağlayarak onaylama sırasında daha sağlam ve daha doğru kararlar alabilen bir model ortaya koymaktadır bu nedenle onaylama aşamasındaki başarı oranı daha yüksektir. Yapılan analizler ve karşılaştırmalara göre Xception ağı ve DenseNet201 ağı her bir yakınlaştırma oranına sahip veri setinde yüksek performans metrik değerleri elde ederek diğer 2 ağa göre bu veri setinde daha başarılı olduklarını göstermişlerdir. Bu 2 başarılı ağ arasında bir karşılaştırma da ise Xception ağı çok az farkla da olsa daha başarılıdır. Ulaşılan bir diğer sonuç ise TTA işleminin sonuçlara olumlu yönde etki ettiği olmuştur.

Yapılan çalışmalarda kullanılan görüntüler ağa girdi olarak vermeden önce aynı boyuta getirilmelidir. Bu çalışma da tek bir veri seti kullanıldığı için bu durum bir sorun olmamıştır; ancak birden fazla veri setinden görüntüler alındığı durumda da veya aynı veri setinde ama farklı boyutlarda görüntüler kullanıldığında bunların mutlaka aynı boyuta getirilmesi gerekmektedir. Model eğitiminin çok uzun süreler almaması için güçlü bir donanıma sahip bilgisayar kullanılmalıdır.

5. III. ARA RAPOR DÖNEMİNDE YAPILAN ÇALIŞMALAR

Histopatolojik görüntüler üzerinde göğüs tümör alt tiplerinin sınırlı tedavi olanaklarından dolayı araştırmacılar tarafından daha sık çalışılmaktadır. Bazal benzeri tümörler ayırt edici morfolojik, genetik, immünofenotipik ve klinik özellikler ile karakterize edilmesine rağmen bu tümörün alt tipinin tanımlanması konusunda kabul görmüş bir fikir birliği olmadığı gibi sistematik olarak sınıflandırılması ile ilgili bir yol tanımlanmamıştır. Sınıflandırma için immünohistokimyasal belirteçler kullanılmakta fakat bu belirteçlerin başarımları ise %60 seviyelerinde kalmaktadır. Projenin üçüncü rapor döneminde histopatolojik görüntülerden elde edilen özellikler birleştirilerek, DNN (Deep Neural Networks, CNN) ve BoW (Bag of Words) sınıflandırıcıları ile sistemin başarımlarının artırılması sağlanmıştır.

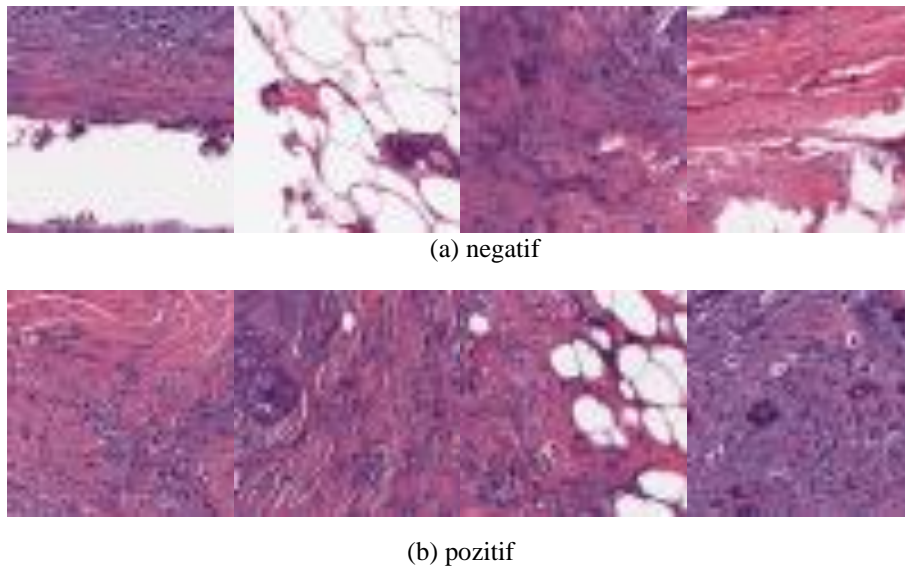
Bu proje döneminde, histopatolojik görüntülerde zamanla bazal benzeri meme tümörlerinin gelişimini incelemek ve bazal benzeri meme kanseri tespiti için histopatolojik görüntüleri analiz etmek için otomatik bir sistem geliştirmiştir. Yapılan çalışmada ilk önce tümörlü ve tümörlü olmayan bölgeler arasında doğru sınıflandırmayı destekleyen kullanışlı özelliklerin çıkartılması sağlanmıştır. Sonraki çalışma adımlarında ise, gürbüz piksel tabanlı ve parçalı tabanlı segmentasyon elde etmek amacıyla da derin öğrenme temelli teknikler kullanarak sınıflandırma işlemi gerçekleştirilmiştir. Geliştirilen sistemi test etmek amacıyla, daha önceki proje dönemlerinde olduğu gibi herkesin kullanımına açık olan veri setleri kullanılmıştır. Proje dönemi boyunca yapılan olan işlemlere ait blok diyagram Şekil 31’de gösterilmektedir.



Şekil 31. Özellik birleştirmeye dayalı önerilen sistemin çalışma şekli

5.1.1. Veriseti

Projenin üçüncü ara rapor döneminde deneysel çalışmalar Breast Histopathology Images veriseti [26] meme kanseri histopatoloji görüntüleri üzerinde yürütülmüştür. Görüntüler orijinal olarak 40x taranan 281 meme kanseri (BCa) örneğinin 281 slayt görüntüsünden çıkartılan 277.524 adet 50×50 parçadan oluşmaktadır. Elde edilen görüntülerin 198.738 adetine IDC negatif, 78.786 adetine ise IDC pozitif olarak tanı konulmuştur. Şekil 2’de Breast Histopathology Images veriseti içerisinde bulunan IDC negatif örnek görüntüler görülmektedir. Şekil 32’de ise aynı veri seti içerisinde bulunan IDC negatif ve pozitif örnek görüntüler görülmektedir.



Şekil 32. Breast histopathology veritinde (a) pozitif ve (b) negatif görüntü örnekleri

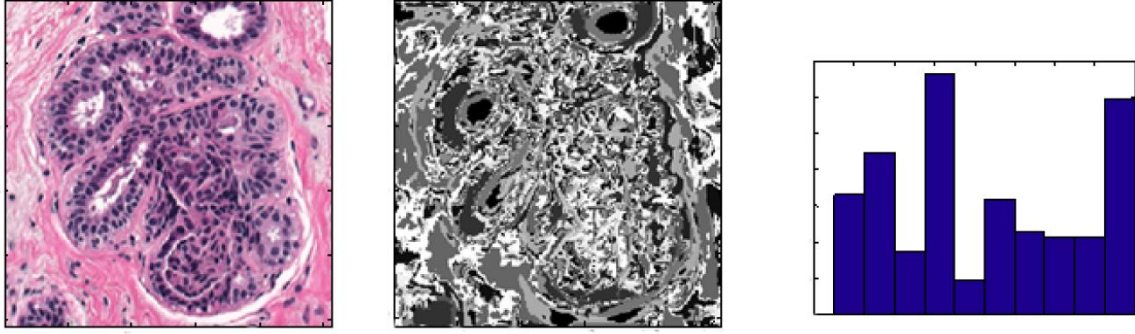
5.2. Özellik Çıkarımı ve Özellik Birleştirme

Proje çalışmasının bu döneminde, histopatolojik görüntüler üzerinde texton özellikler, network özellikleri, morfometrik özellikler ve GLCM özellikleri çıkarılarak, BoW (Bag of Words) ve DNN yöntemleri ile özellik birleştirmeye dayalı, bazal benzeri göğüs kanseri sınıflandırması gerçekleştirilmiştir.

5.2.1. Texton (Texture/Dokusal) özellikler

Daha önceki çalışmalarda, Khurd vd. [58] çalışmalarında, prostat kanserinin evrenlenmesinde dokusal (texton) bir sınıflandırma sistemi önermiştir. Bu özellik çıkarımında, temel seviyedeki dokusal elementlerin belirlenmesi için kümeleme tabanlı bir filtreleme uygulandığından texton olarak isimlendirilmektedir [59]. Texton özelliklerin çıkarılması için

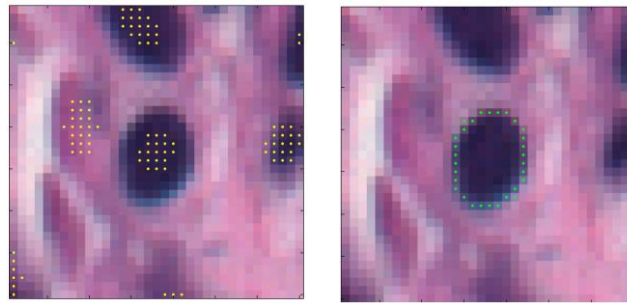
her piksel düzeyinde değişmez bir filtre bankası kullanılmaktadır. Bu özellik çıkarımı için, Varma ve Zisserman tarafından sunulan Maximum Response (MR) filtre bankasının özellikleri elde edilmiştir. Bu filtre bankaları, MR8 filtresi, Gaussian filtresi ve Laplacian of Gaussian filtresidir. Bir histopatolojik görüntüden çıkarılan texton özellik ve onun histogramı Şekil 33'te gösterilmiştir.



(a) Orijinal görüntü (b) Texton haritası (c) Texton histogramı
Şekil 33. MR8 filtresi, Gaussian filtresi ve Laplacian of Gaussian filtrelerinden Texton özelliklerin çıkarılması

5.2.2. Morfometrik özellikler

Çalışma kapsamında histopatolojik görüntüler için belirlenen morfometrik özellikler, patoloğlar tarafından belirlenen parametrelerle uyumlu olarak, görüntü üzerindeki hücre çekirdeğinin boyutu ve şekli ile ilgili değişiklikleri yakalamaktadır [60]. Bu özellik çıkarım metodu ile 3 farklı türde morfometrik özellik araştırılmıştır. Bunlar; Hessian matrisinden çıkarılan bilgi, Fourier Şekil Tanımlayıcılarından gelen bilgiler ve çekirdeğin çevresinden çıkarılan kanal yapısı ile ilgili kodlama ile oluşturulan özel bir özelliktir [59]. Şekil 34'te çıkarılan bazı morfometrik özellikler sunulmuştur.



Şekil 34. Morfometrik özelliklerin çıkarılması

5.2.3. Network özellikleri

Birçok görüntü tanıma probleminde olduğu gibi, bir görüntüde üzerinde nesnelere oluşturan nokta çiftleri arasındaki kenar bağlantı biçimleri oldukça önemlidir. Bu çalışmada kullanılan network özellikleri, histopatolojik görüntülerden tespit edilen hücre çekirdeklerinin komşuluk graf ilişkilerine dayalı olarak Urquhart graflarından [61] elde edilmektedir. Çalışma

kapsamında, Urquhart grafları ve ekstra hücrel matrisler kullanılarak, özel işaretleri yakalamak için network istatistiklerinde dayalı olarak network devirleri (network cycles) kullanılmıştır. Deneysel çalışmalar kapsamında farklı devirlerin ağırlıklı ve ağırlıksız uzunluklarına bağlı olarak kullanılan özelliklerin listesi Tablo 4’te verilmiştir.

Tablo 4. Histopatolojik görüntülerden çıkarılan network özellikleri

Network özellikleri	
Özellik adı	Açıklaması
number of cycles with length greater than 3	Uzunluğu 3’ten büyük devir sayısı
average non-triangular cycle length	Üçgensel olmayan ortalama devir uzunluğu
maximum non-triangular cycle length	Üçgensel olmayan maksimum devir uzunluğu
number of vertices	Tepe sayısı
number of graph components	Graf bileşenlerinin sayıları
clustering coefficients	Kümeleme katsayıları
Fiedler values	Fiedler (vertex Laplacian and the edge Laplacian)
average shortest path length	Ortalama en kısa yol uzunluğu

5.2.4. GLCM(gray-level co-occurrence matrix) özellikleri

Gri seviyeli bir dokunun istatistiksel karakteristikleri ilk olarak Haralick tarafından önerilen GLCM (gray-level co-occurrence matrix)' den türetilir [62, 63]. GLCM metodu farklı gri düzeyli pikseller arasındaki ilişkiyi gösterir. Görüntünün iki boyutlu (2D) olması durumunda farklı açı yönlerinden özellikler çıkarılabilir. Metot i ve j piksellerinin belli bir açı ve uzaklıktaki piksel yoğunluklarını hesaplayarak bir GLCM matrisi oluşturur. Bu çalışmada histopatolojik görüntüler üzerinden çıkarılan GLCM özelliklerinin listesi Tablo 5’te sunulmuştur. H&E görüntülerden bu özellikler özellikleri 0^0 , 45^0 , 90^0 ve 135^0 açıları doğrultularından $d=2$ uzaklığında çıkarılmıştır.

Tablo 5. Histopatolojik görüntülerden çıkarılan GLCM özellikleri

GLCM özellik adı	Özellik sırası
Angular Second Moment	1
Entropy	2
Dissimilarity	3
Contrast	4
Inverse Difference	5
Correlation	6
Homogeneity	7
Autocorrelation	8
Cluster Shade	9
Cluster Prominence	10
Maximum probability	11
Sum of Squares	12
Sum Average	13
Sum Variance	14
Sum Entropy	15
Difference Variance	16
Difference Entropy	17
Information measures of correlation-1	18
Information measures of correlation2	19
Maximal correlation co-efficient	20
Inverse difference normalized	21
Inverse difference moment normalized	22
Toplam özellik sayısı	88

5.2.5. Özellik Birleştirme

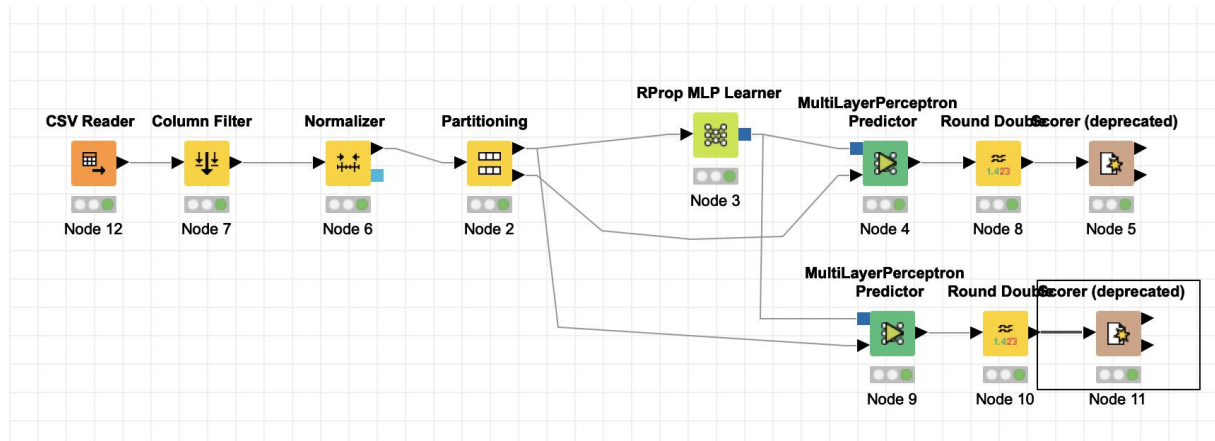
Proje çalışması kapsamında, histopatolojik görüntüler üzerinden çıkarılan texton özellikler, network özellikleri, morfometrik özellikler ve GLCM özelliklerine, Şekil 35'te görüldüğü gibi özellik birleştirme (feature fusion) bazal benzeri göğüs kanser görüntüleri BoW ve DNN yöntemleri ile sınıflandırılmıştır.

imgName	glcm	texton	perimeter	kpHarris	kpHarrisLaplace
"10253_idx...	1x88 dou...	8x2500 d...	1x24 double	[487.5928;...	[48,49,27.8628]
"10253_idx...	1x88 dou...	8x2500 d...	1x24 double	[556.1460;...	[49,48,40.1225]
"10253_idx...	1x88 dou...	8x2500 d...	1x24 double	[663.7469;...	[49,48,40.1225]
"10253_idx...	1x88 dou...	8x2500 d...	1x24 double	[685.4202;...	[48,49,40.1225]
"10253_idx...	1x88 dou...	8x2500 d...	1x24 double	[664.8374;...	[47,48,11.1974]
"10253_idx...	1x88 dou...	8x2500 d...	1x24 double	[561.2599;...	[49,48,13.4369]
"10253_idx...	1x88 dou...	8x2500 d...	1x24 double	[497.4432;...	[49,49,40.1225]
"10253_idx...	1x88 dou...	8x2500 d...	1x24 double	[560.0729;...	[48,49,40.1225]

Şekil 35. Texton özellikler, network özellikleri, morfometrik özellikler ve GLCM özelliklerinin birleştirilmesi

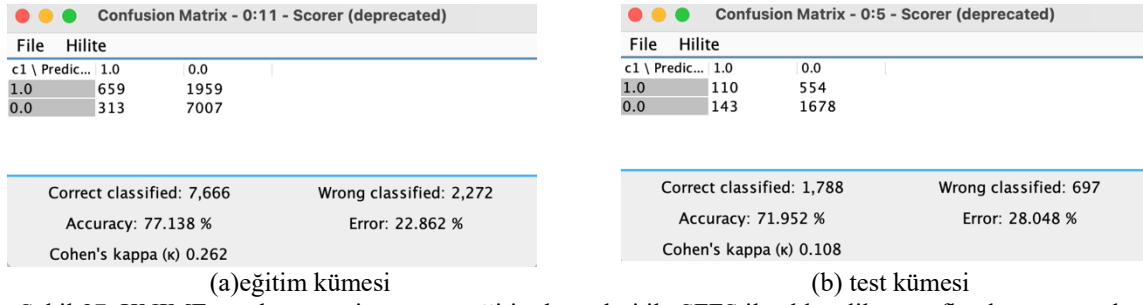
5.2.6. Özellik Seçimi

Proje çalışması kapsamında, histopatolojik görüntüler üzerinden çıkarılan texton özellikler, network özellikleri, morfometrik özellikler ve GLCM özelliklerinin toplam sayısı çok fazla olduğundan, hepsinin aynı anda kullanımının tespit başarımını düşürdüğü görülmüştür. Bu nedenle birleştirilen özellikler içerisinde en uygun özelliklerin seçimi sağlanmıştır. Deneysel çalışmalar kapsamında özellik seçimi KNIME yazılımı [64] ile Sequential Floating Forward Selection (SFFS) yöntemi ile gerçekleştirilmiştir. SFFS genel amaçlı bir özellik çıkarım yöntemi olup, özelliklerin düzey önem derecelerine göre geniş özellik matrislerinden en etkili özelliklerin seçimi için kullanılmaktadır [65]. Özellik seçimi için KNIME yazılımında yapay sinir ağları ile Şekil 36'da görülen mimari blok diyagram tasarlanmıştır.



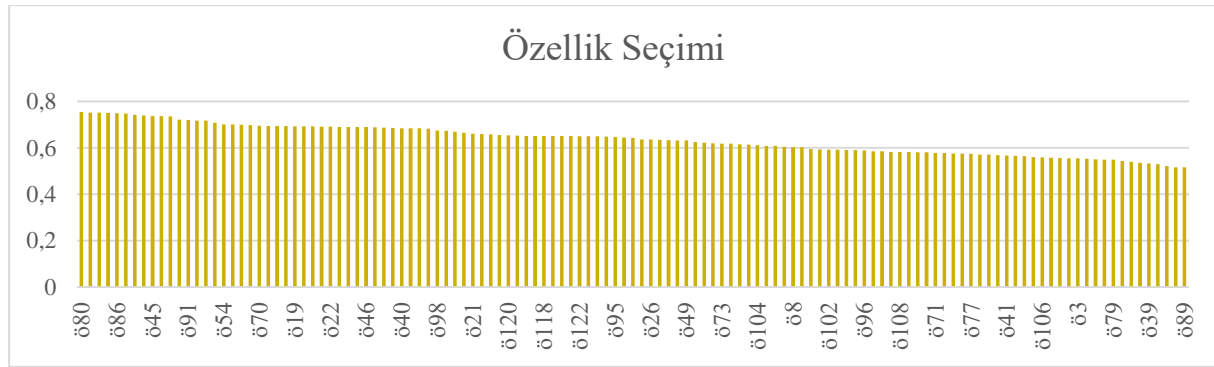
Şekil 36. Özellik seçimi için KNIME ortamında tasarlanan blok diyagram

KNIME yazılımı üzerine test ve eğitim kümeleri ile SFFS ile elde edilen sınıflandırma sonuçları Şekil 37’de sunulmuştur.



Şekil 37. KNIME yazılımı üzerine test ve eğitim kümeleri ile SFFS ile elde edilen sınıflandırma sonuçları

Özellik seçimi sonucunda özelliklerin her birisinin önemine göre sıralanmış hali Şekil 38’de gösterilmiştir. Buradan BoW ve DNN için en uygun özellikler seçilerek histopatolojik görüntülerin sınıflandırılması sağlanmıştır.



Şekil 38. Özelliklerin önem derecelerine göre sıralanması

5.3. Kullanılan Sınıflandırma Yöntemleri

Proje çalışması kapsamında, histopatolojik görüntüler üzerinden çıkarılan texton özellikler, network özellikleri, morfometrik özellikler ve GLCM özellikler kullanılarak, BoW (Bag of Word) ve DNN (Deep Neural Network) yöntemleri ile özellik birleştirmeye dayalı, bazal benzeri göğüs kanseri sınıflandırması gerçekleştirilmiştir.

DNN (Deep Neural Network), CNN(Convolutional Neural Network) mimarisinin farklı bir uygulaması olup, sınıflandırma problemlerinde yaygın olarak tercih edilmektedir. DNN yapısı aynı zamanda 1D-CNN olarak da literatürde birçok çalışmada kullanılmaktadır [66]. Çalışma kapsamında, histopatolojik görüntüler üzerinde çıkarılan özelliklerin sınıflandırılması için, Phyton programlama dilinde TensorFlow ortamında tasarlanan DNN mimarisi Şekil 39’da gösterilmiştir. Tasarlanan DNN mimarisinde optimizer olarak ‘ADAM’ fonksiyonu tercih edilmiştir.

```

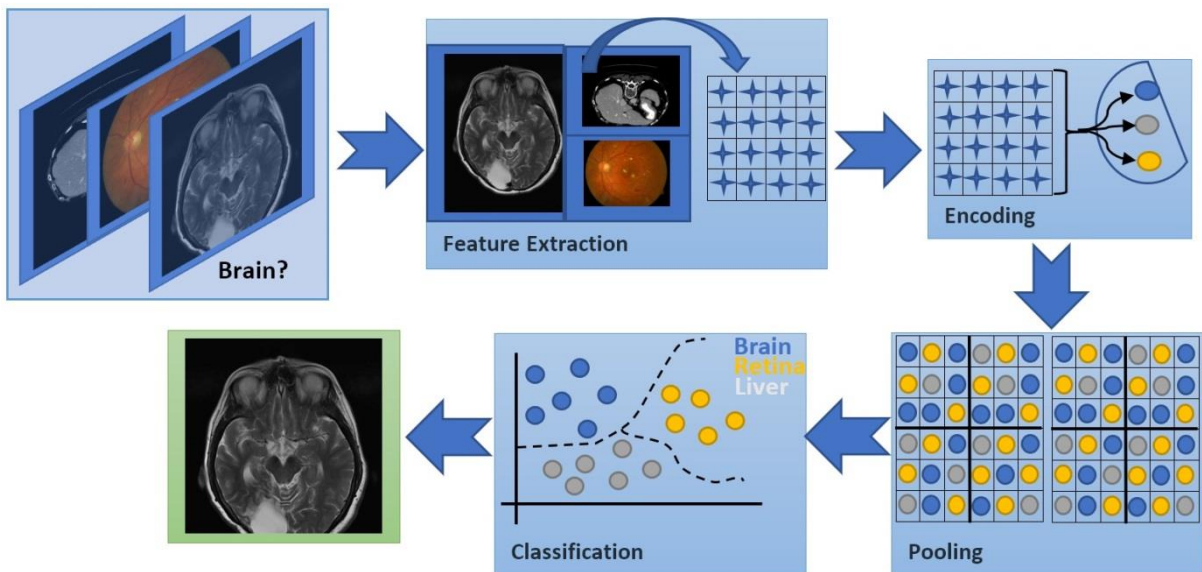
model= Sequential();
model.add(Dense(64,input_dim=125))
model.add(Activation('relu'))
model.add(Dense(128))
model.add(Activation('relu'))
model.add(Dense(256))
model.add(Activation('relu'))
model.add(Dense(128))
model.add(Activation('relu'))
model.add(Dense(64))
model.add(Activation('relu'))
model.add(Dense(1))
model.add(Activation('sigmoid')) #k

```

Şekil 39. Histopatolojik görüntülerden çıkarılan özelliklerin sınıflandırılması için tasarlanan DNN mimarisi

Kelime Çantası olarak bilinen BoW modeli doğal işlemede sıklıkla kullanılan bir modeldir. Doğal dil işleme ve bilgi geri alma (Information retrieval)'dan türetilmiştir [67]. Günümüzde Bag of words, görüntü işleme alanında yaygın olarak kullanılmaktadır. BoW, görüntüyü bir dizi özellik olarak temsil etmektedir. Görüntü üzerindeki özellikler temel noktalardan ve tanımlayıcılardan oluşmaktadır. Bir görüntüdeki özelliklerin frekansı histogramı çıkartılarak sınıflandırma yapılmaktadır.

Görüntü sınıflandırma, görüntü özelliklerini kullanarak görüntüleri sınıflandıran bir işlemdir [68]. Temel olarak, öğrenen sınıf görüntülerinin ana özelliğini çıkarır ve bu bilinmeyen sınıfların görüntülerini temsil edecek en iyi sınıfı bulmak için bilinmeyen sınıfların görüntülerinin özellikleriyle karşılaştırılmak üzere kullanılacak bir kod kitabına dönüştürür [69]. BoW algoritmasının çalışma yapısını gösteren blok diyagram Şekil 40'da gösterilmiştir.



Şekil 40. BoW modelinin çalışma yapısını gösteren blok diyagram

5.4. Deneysel Çalışmalar

Proje çalışması kapsamında üçüncü rapor döneminde yürütülen deneysel çalışmalarda, histopatolojik görüntüler üzerinden çıkarılan texton özellikler, network özellikleri, morfometrik özellikler ve GLCM özellikler BoW (Bag of Word) ve DNN (Deep Neural Network) yöntemleri ile özellik birleştirme yapılarak, bazal benzeri göğüs kanseri sınıflandırması gerçekleştirilmiştir. Böylece sadece BoW ve sadece DNN ile elde edilen tekil sonuçlardan, özellik birleştirme ile daha başarılı sonuçlar elde edilmiştir. Şekil 41 ve Şekil 42’de görülebileceği gibi, Bow ve DNN yöntemleri ile özellik birleştirme yapılmadan elde edilen sınıflandırma sonuçları sırasıyla %67.0 ve %74.16 olarak hesaplanmıştır.

KNOWN	PREDICTED	
	class0	class1
class0	0.67	0.33
class1	0.32	0.68

* Average Accuracy is 0.67.

Şekil 41. BoW yöntemi ile elde edilen sınıflandırma sonuçları

	precision	recall	f1-score	support
0	0.74	0.99	0.85	1844
1	0.48	0.02	0.04	641
accuracy			0.74	2485
macro avg	0.61	0.51	0.44	2485
weighted avg	0.68	0.74	0.64	2485
[[1831 629]				
[13 12]]				
accuracy is	0.7416498993963783			

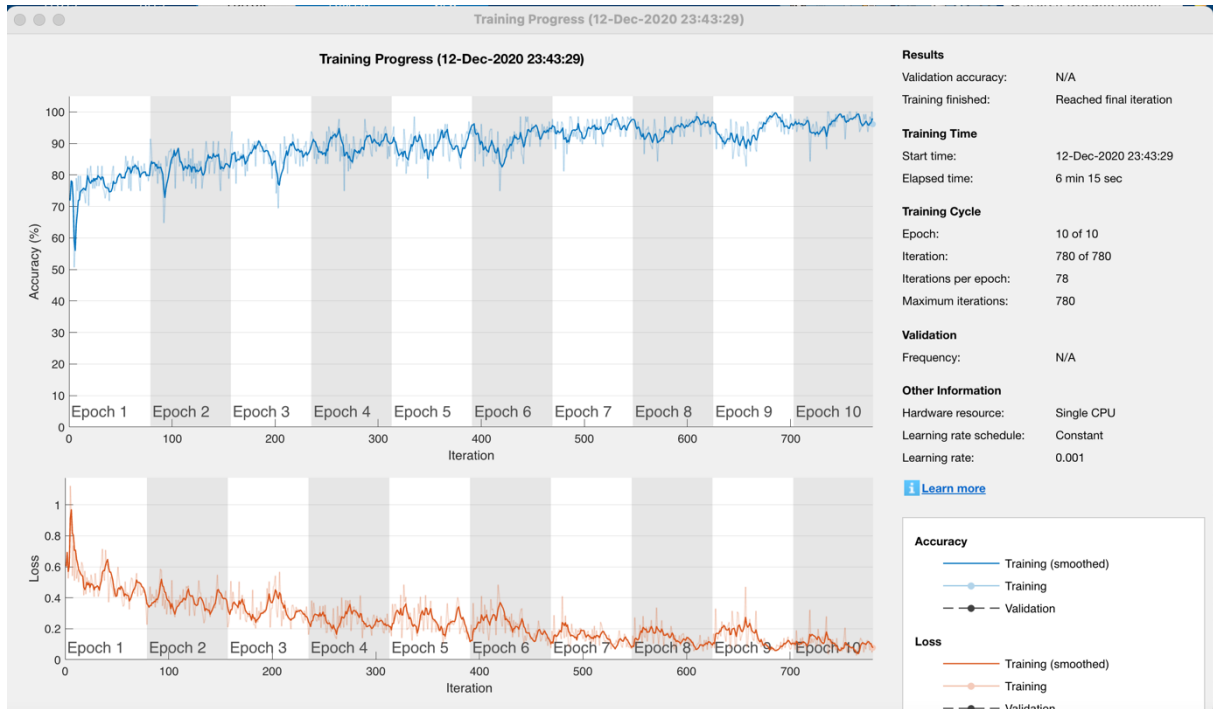
Şekil 42. DNN yöntemi ile elde edilen sınıflandırma sonuçları

Çalışma kapsamında daha yüksek sınıflandırma başarımına ulaşmak için özellik birleştirme yapılarak BoW ve DNN yöntemleri ile hibrit bir yöntemin geliştirilmesi sağlanmıştır. Özellik birleştirme sonrası BoW+CNN hibrit yöntemi ile yapılan eğitim aşamasına ait ulaşılan bazı parametre değerleri Şekil 43’te gösterilmiştir. Ayrıca önerilen hibrit ağın eğitimdeki doğruluk ve Loss değerlerinin değişimini gösteren board ekranı ise Şekil 44’te sunulmuştur.

Initializing input data normalization.

Epoch	Iteration	Time Elapsed (hh:mm:ss)	Mini-batch Accuracy	Mini-batch Loss	Base Learning Rate
1	1	00:00:03	62.50%	0.6984	0.0010
1	50	00:00:10	65.62%	1.0063	0.0010
1	100	00:00:17	68.75%	0.6766	0.0010
1	150	00:00:24	53.12%	0.7867	0.0010
1	200	00:00:31	81.25%	0.4991	0.0010
1	250	00:00:38	90.62%	0.3725	0.0010
1	300	00:00:45	90.62%	0.3293	0.0010
2	350	00:00:53	81.25%	0.5234	0.0010
2	400	00:00:59	84.38%	0.3232	0.0010
2	450	00:01:06	53.12%	0.8251	0.0010
2	500	00:01:12	87.50%	0.3832	0.0010
2	550	00:01:19	81.25%	0.3502	0.0010
2	600	00:01:26	90.62%	0.2876	0.0010
3	650	00:01:33	75.00%	0.4139	0.0010
3	700	00:01:41	75.00%	0.5071	0.0010
3	750	00:01:48	71.88%	0.7431	0.0010
3	800	00:01:54	96.88%	0.1748	0.0010
3	850	00:02:01	87.50%	0.2805	0.0010
3	854	00:02:02	90.62%	0.2567	0.0010

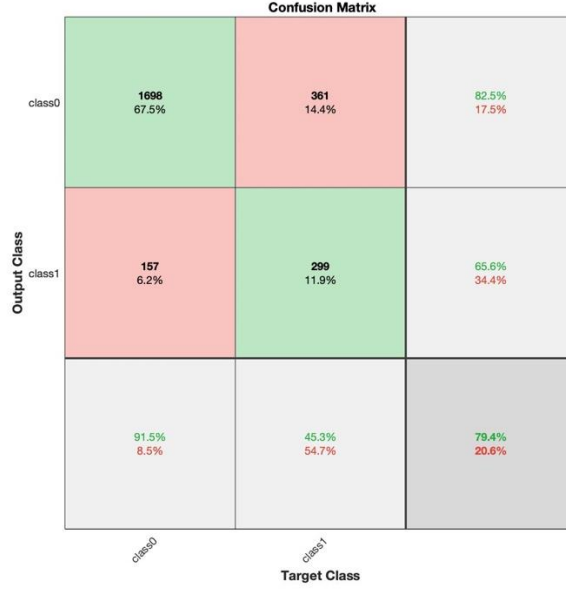
Şekil 43. BoW+CNN hibrit yöntemi ile yapılan eğitim aşamasına ait ulaşılan bazı parametre değerleri



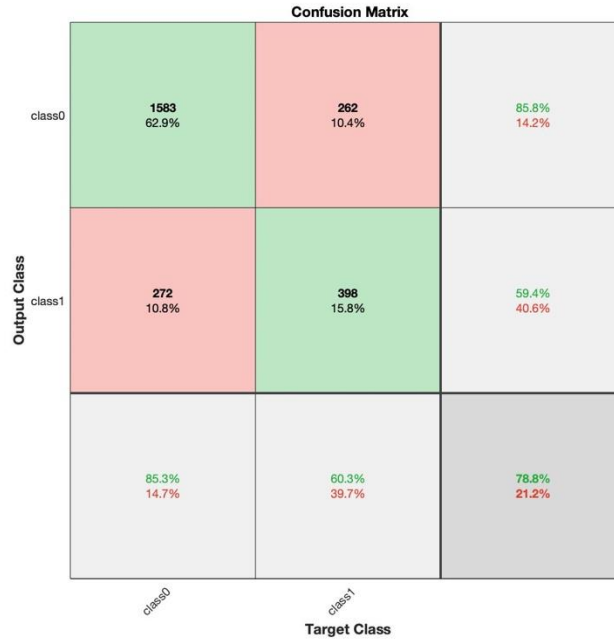
Şekil 44. Önerilen hibrit ağı eğitimdeki doğruluk ve Loss değerlerinin değişimini gösteren board ekranı

Eğitim aşaması tamamlandıktan sonra hibrit yöntem ile sırasıyla eğitim ve test kümeleri üzerinde elde edilen sınıflandırma sonuçlarını gösteren karmaşıklık matrisleri ise Şekil 45 ve Şekil 46'da gösterilmiştir. Bu karmaşıklık matrislerinden de görüleceği üzere, eğitim veri kümesi ile %79.4 sınıflandırma başarımına ulaşılırken, test kümesi ile %78.8 sınıflandırma başarımına ulaşılmıştır. Böylece, histopatolojik görüntüler üzerinde bazal benzeri göğüs

kanserinin sınıflandırılmasında özellik birleştirmeye dayalı BoW+CNN hibrit yöntemi ile başarılı sonuçlara ulaşıldığı doğrulanmıştır.



Şekil 45. Hibrit yöntem ile sırasıyla eğitim kümesi üzerinde elde edilen sınıflandırma sonuçlarını gösteren karmaşıklık matrisi ve sınıflandırma sonuçları



Şekil 46. Hibrit yöntem ile sırasıyla eğitim kümesi üzerinde elde edilen sınıflandırma sonuçlarını gösteren karmaşıklık matrisi ve sınıflandırma sonuçları

6. SONUÇLAR VE DEĞERLENDİRMELER

Projenin üçüncü ve son kapanış döneminde, derin öğrenme ile tümör sınıflandırılması ve elde edilen sınıflandırmaların değerlendirilmesi işlemlerinin yapılacağı ve edilen veri setlerinin yaygınlaştırılması ve deneysel çalışmaların artırılması ile başarı oranlarının yükseltilmesi kapsamında çalışmalar yürütüleceği planlanmıştı. Geline nokta istenilen hedeflere ulaşıldığı görülmektedir. Projenin mevcut uygulama planına göre ayrıca üçüncü ara rapor döneminde çalışmada kullanılacak olan veri setleri belirlenerek temin edilmiştir. Elde edilen veri seti kullanılarak öznitelik çıkarımı ve öznitelik tanımlayıcılarının birleştirilmesi işlemleri yapılmıştır.

Eğitim aşaması tamamlandıktan sonra hibrit yöntem ile sırasıyla eğitim ve test kümeleri üzerinde elde edilen sınıflandırma sonuçlarını gösteren karmaşıklık matrislerinden de görüleceği üzere, eğitim veri kümesi ile %79.4 sınıflandırma başarımına ulaşıırken, test kümesi ile %78.8 sınıflandırma başarımına ulaşılmıştır. Böylece, histopatolojik görüntüler üzerinde bazal benzeri göğüs kanserinin sınıflandırılmasında özellik birleştirmeye dayalı BoW+CNN hibrit yöntemi ile başarılı sonuçlara ulaşıldığı doğrulanmıştır.

Projede bundan sonraki aşamalarda, kapanış raporu sunulmuş olsa bile, farklı derin öğrenme modeli ile histopatolojik görüntülerde meme kanserinin sınıflandırılması çalışmalarında devam edilecek, verisetinin genişletilerek yeni verisetleri üzerinde yeni deneysel çalışmalar yapılacak ve proje çalışmalarından akademik makalelerin üretilmesi sağlanacaktır.

7. PROJE KAPSAMINDA YAPILAN AKADEMİK ÇALIŞMALAR

Bu proje kapsamında, şu ana kadar 3 akademik çalışma tamamlanmış ve yayınlanmıştır. Bu akademik çalışmalar aşağıda listelenmiştir. Ayrıca yeni yayınlar için akademik çalışmalar devam etmektedir.

Kitap Bölümü:

ÇEVİK, K.K., DANDIL, E.,UZUN, S.,YILDIRIM, M.S,SELVİ, A. O., **Detection of Breast Cancer Using Deep Neural Networks with Transfer Learning on Histopathological Images.** AIDDMD-2020: Artificial Intelligence for Data-Driven Medical Diagnosis. Yayınevi : Walter de Gruyter GmbH (*Basım Aşamasında*)

Makale (TR Dizin):

Dandıl, E, Serin, Z., **Breast Cancer Detection on Histopathological Images Using Deep Neural Networks.** European Journal of Science and Technology (EJOSAT).

Uluslararası Konferans (Özet Metin):

Dandıl, E, Serin, Z., **Derin Sinir Ağları Kullanarak Histopatolojik Görüntülerde Meme Kanseri Tespiti.** 2nd International Congress on Human-Computer Interaction, Optimization and Robotic Applications, ss.52, 26-27 Haziran2020, Ankara, Turkey.

KAYNAKLAR

- [1] Siegel, R.L., Miller, K.D., and Jemal, A.: 'Cancer Statistics, 2015', *CA: a cancer journal for clinicians*, 2015, **65**, (1), pp. 5-29
- [2] Stewart, B. and Wild, C.P.: 'World Cancer Report 2014', 2014,
- [3] Han, Z., Wei, B., Zheng, Y., *et al.*: 'Breast Cancer Multi-Classification from Histopathological Images with Structured Deep Learning Model', *Scientific reports*, 2017, **7**, (1), pp. 1-10
- [4] Siegel, R.L., Miller, K.D., and Jemal, A.: 'Cancer Statistics, 2019', *CA: a cancer journal for clinicians*, 2019, **69**, (1), pp. 7-34
- [5] Dogan, B.E., Gonzalez-Angulo, A.M., Gilcrease, M., *et al.*: 'Multimodality Imaging of Triple Receptor–Negative Tumors with Mammography, Ultrasound, and Mri', *American Journal of Roentgenology*, 2010, **194**, (4), pp. 1160-1166
- [6] Wang, D., Khosla, A., Gargeya, R., *et al.*: 'Deep Learning for Identifying Metastatic Breast Cancer', *arXiv preprint arXiv:1606.05718*, 2016,
- [7] Kumar, A., Singh, S.K., Saxena, S., *et al.*: 'Deep Feature Learning for Histopathological Image Classification of Canine Mammary Tumors and Human Breast Cancer', *Information Sciences*, 2020, **508**, pp. 405-421
- [8] Öztürk, Ş. and Akdemir, B.: 'Hic-Net: A Deep Convolutional Neural Network Model for Classification of Histopathological Breast Images', *Computers & Electrical Engineering*, 2019, **76**, pp. 299-310
- [9] Rakhlin, A., Shvets, A., Iglovikov, V., *et al.*: 'Deep Convolutional Neural Networks for Breast Cancer Histology Image Analysis', *International Conference Image Analysis and Recognition*, 2018, pp. 737-744
- [10] Jones, C.E., Maben, J., Lucas, G., *et al.*: 'Barriers to Early Diagnosis of Symptomatic Breast Cancer: A Qualitative Study of Black African, Black Caribbean and White British Women Living in the Uk', *BMJ open*, 2015, **5**, (3)
- [11] Ibrahim, F., Thio, T.H.G., Faisal, T., *et al.*: 'The Application of Biomedical Engineering Techniques to the Diagnosis and Management of Tropical Diseases: A Review', *Sensors*, 2015, **15**, (3), pp. 6947-6995
- [12] Badve, S., Dabbs, D.J., Schnitt, S.J., *et al.*: 'Basal-Like and Triple-Negative Breast Cancers: A Critical Review with an Emphasis on the Implications for Pathologists and Oncologists', *Modern Pathology*, 2011, **24**, (2), pp. 157-167
- [13] Sertel, O., Lozanski, G., Shana'ah, A., *et al.*: 'Computer-Aided Detection of Centrioblasts for Follicular Lymphoma Grading Using Adaptive Likelihood-Based Cell Segmentation', *IEEE Transactions on Biomedical Engineering*, 2010, **57**, (10), pp. 2613-2616
- [14] Öztürk, Ş. and Akdemir, B.: 'Application of Feature Extraction and Classification Methods for Histopathological Image Using Glcm, Lbp, Lbglcm, Glrlm and Sfta', *Procedia computer science*, 2018, **132**, pp. 40-46
- [15] Mikhaylov, V. and Bakhshiev, A.: 'The System for Histopathology Images Analysis of Spinal Cord Slices', *Procedia Computer Science*, 2017, **103**, (C), pp. 239-243
- [16] Azar, A.T. and El-Said, S.A.: 'Probabilistic Neural Network for Breast Cancer Classification', *Neural Computing and Applications*, 2013, **23**, (6), pp. 1737-1751
- [17] Abdel-Zaher, A.M. and Eldeib, A.M.: 'Breast Cancer Classification Using Deep Belief Networks', *Expert Systems with Applications*, 2016, **46**, pp. 139-144
- [18] Nahid, A.-A., Mehrabi, M.A., and Kong, Y.: 'Histopathological Breast Cancer Image Classification by Deep Neural Network Techniques Guided by Local Clustering', *BioMed research international*, 2018, **2018**

- [19] Mohammed, M.A., Al-Khateeb, B., Rashid, A.N., *et al.*: 'Neural Network and Multi-Fractal Dimension Features for Breast Cancer Classification from Ultrasound Images', *Computers & Electrical Engineering*, 2018, **70**, pp. 871-882
- [20] Wang, P., Xu, S., Li, Y., *et al.*: 'Feature-Based Analysis of Cell Nuclei Structure for Classification of Histopathological Images', *Digital Signal Processing*, 2018, **78**, pp. 152-162
- [21] Khameneh, F.D., Razavi, S., and Kamasak, M.: 'Automated Segmentation of Cell Membranes to Evaluate Her2 Status in Whole Slide Images Using a Modified Deep Learning Network', *Computers in biology and medicine*, 2019, **110**, pp. 164-174
- [22] Budak, Ü., Cömert, Z., Rashid, Z.N., *et al.*: 'Computer-Aided Diagnosis System Combining Fcn and Bi-Lstm Model for Efficient Breast Cancer Detection from Histopathological Images', *Applied Soft Computing*, 2019, **85**, p. 105765
- [23] Hekler, A., Utikal, J.S., Enk, A.H., *et al.*: 'Deep Learning Outperformed 11 Pathologists in the Classification of Histopathological Melanoma Images', *European Journal of Cancer*, 2019, **118**, pp. 91-96
- [24] Sudharshan, P., Petitjean, C., Spanhol, F., *et al.*: 'Multiple Instance Learning for Histopathological Breast Cancer Image Classification', *Expert Systems with Applications*, 2019, **117**, pp. 103-111
- [25] Yan, R., Ren, F., Wang, Z., *et al.*: 'Breast Cancer Histopathological Image Classification Using a Hybrid Deep Neural Network', *Methods*, 2020, **173**, pp. 52-60
- [26] Cruz-Roa, A., Basavanahally, A., González, F., *et al.*: 'Automatic Detection of Invasive Ductal Carcinoma in Whole Slide Images with Convolutional Neural Networks', *Medical Imaging 2014: Digital Pathology*, 2014, pp. 904103
- [27] Aksac, A., Demetrick, D.J., Ozyer, T., *et al.*: 'Brecahad: A Dataset for Breast Cancer Histopathological Annotation and Diagnosis', *BMC research notes*, 2019, **12**, (1), pp. 1-3
- [28] Aresta, G., Araújo, T., Kwok, S., *et al.*: 'Bach: Grand Challenge on Breast Cancer Histology Images', *Medical image analysis*, 2019, **56**, pp. 122-139
- [29] He, K., Zhang, X., Ren, S., *et al.*: 'Deep Residual Learning for Image Recognition', *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2016, pp. 770-778
- [30] Szegedy, C., Vanhoucke, V., Ioffe, S., *et al.*: 'Rethinking the Inception Architecture for Computer Vision', *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2016, pp. 2818-2826
- [31] Gonzalez, T.F., *Handbook of Approximation Algorithms and Metaheuristics*, (CRC Press, 2007)
- [32] Iandola, F.N., Han, S., Moskewicz, M.W., *et al.*: 'Squeezenet: Alexnet-Level Accuracy with 50x Fewer Parameters and < 0.5 Mb Model Size', *arXiv preprint arXiv:1602.07360*, 2016,
- [33] Hashemi, M.: 'Enlarging Smaller Images before Inputting into Convolutional Neural Network: Zero-Padding Vs. Interpolation', *Journal of Big Data*, 2019, **6**, (1), p. 98
- [34] Ioffe, S. and Szegedy, C.: 'Batch Normalization: Accelerating Deep Network Training by Reducing Internal Covariate Shift', *arXiv preprint arXiv:1502.03167*, 2015,
- [35] Collis, J., Glossary of Deep Learning: Batch Normalisation', (Jun, 2017)
- [36] Perez, L. and Wang, J.: 'The Effectiveness of Data Augmentation in Image Classification Using Deep Learning', *arXiv preprint arXiv:1712.04621*, 2017,
- [37] Taylor, L. and Nitschke, G.: 'Improving Deep Learning Using Generic Data Augmentation', *arXiv preprint arXiv:1708.06020*, 2017,
- [38] Moshkov, N., Mathe, B., Kertesz-Farkas, A., *et al.*: 'Test-Time Augmentation for Deep Learning-Based Cell Segmentation on Microscopy Images', *Scientific reports*, 2020, **10**, (1), pp. 1-7

- [39] Ide, H. and Kurita, T.: 'Improvement of Learning for Cnn with Relu Activation by Sparse Regularization', 2017 International Joint Conference on Neural Networks (IJCNN), 2017, pp. 2684-2691
- [40] Paasio, A. and Dawidziuk, A.: 'Cnn Template Robustness with Different Output Nonlinearities', *International Journal of Circuit Theory and Applications*, 1999, **27**, (1), pp. 87-102
- [41] Ma, X., Dai, Z., He, Z., *et al.*: 'Learning Traffic as Images: A Deep Convolutional Neural Network for Large-Scale Transportation Network Speed Prediction', *Sensors*, 2017, **17**, (4), p. 818
- [42] Kalchbrenner, N., Grefenstette, E., and Blunsom, P.: 'A Convolutional Neural Network for Modelling Sentences', *arXiv preprint arXiv:1404.2188*, 2014,
- [43] Tai, Y., Yang, J., and Liu, X.: 'Image Super-Resolution Via Deep Recursive Residual Network', Proceedings of the IEEE conference on computer vision and pattern recognition, 2017, pp. 3147-3155
- [44] Huang, G., Liu, Z., Van Der Maaten, L., *et al.*: 'Densely Connected Convolutional Networks', Proceedings of the IEEE conference on computer vision and pattern recognition, 2017, pp. 4700-4708
- [45] Szegedy, C., Ioffe, S., Vanhoucke, V., *et al.*: 'Inception-V4, Inception-Resnet and the Impact of Residual Connections on Learning', Thirty-first AAAI conference on artificial intelligence, 2017, pp.
- [46] Chollet, F.: 'Xception: Deep Learning with Depthwise Separable Convolutions', Proceedings of the IEEE conference on computer vision and pattern recognition, 2017, pp. 1251-1258
- [47] Srivastava, N., Hinton, G., Krizhevsky, A., *et al.*: 'Dropout: A Simple Way to Prevent Neural Networks from Overfitting', *The journal of machine learning research*, 2014, **15**, (1), pp. 1929-1958
- [48] Davis, J. and Goadrich, M.: 'The Relationship between Precision-Recall and Roc Curves', Proceedings of the 23rd international conference on Machine learning, 2006, pp. 233-240
- [49] Powers, D.M.: 'Evaluation: From Precision, Recall and F-Measure to Roc, Informedness, Markedness and Correlation', 2011,
- [50] Behera, B. and Kumaravelan, G.: 'Performance Evaluation of Deep Learning Algorithms in Biomedical Document Classification', 2019 11th International Conference on Advanced Computing (ICoAC), 2019, pp. 220-224
- [51] Vani, S. and Rao, T.M.: 'An Experimental Approach Towards the Performance Assessment of Various Optimizers on Convolutional Neural Network', 2019 3rd International Conference on Trends in Electronics and Informatics (ICOEI), 2019, pp. 331-336
- [52] Goutte, C. and Gaussier, E.: 'A Probabilistic Interpretation of Precision, Recall and F-Score, with Implication for Evaluation', European Conference on Information Retrieval, 2005, pp. 345-359
- [53] Powers, D.M.: 'What the F-Measure Doesn't Measure: Features, Flaws, Fallacies and Fixes', *arXiv preprint arXiv:1503.06410*, 2015,
- [54] Bradley, A.P.: 'The Use of the Area under the Roc Curve in the Evaluation of Machine Learning Algorithms', *Pattern recognition*, 1997, **30**, (7), pp. 1145-1159
- [55] Hajian-Tilaki, K.: 'Receiver Operating Characteristic (Roc) Curve Analysis for Medical Diagnostic Test Evaluation', *Caspian journal of internal medicine*, 2013, **4**, (2), p. 627
- [56] Ling, C.X., Huang, J., and Zhang, H.: 'Auc: A Better Measure Than Accuracy in Comparing Learning Algorithms', Conference of the canadian society for computational studies of intelligence, 2003, pp. 329-341

- [57] Huang, J. and Ling, C.X.: 'Using Auc and Accuracy in Evaluating Learning Algorithms', *IEEE Transactions on knowledge and Data Engineering*, 2005, **17**, (3), pp. 299-310
- [58] Khurd, P., Bahlmann, C., Maday, P., *et al.*: 'Computer-Aided Gleason Grading of Prostate Cancer Histopathological Images Using Texton Forests', 2010 IEEE International Symposium on Biomedical Imaging: From Nano to Macro, 2010, pp. 636-639
- [59] Chekkoury, A., Khurd, P., Ni, J., *et al.*: 'Automated Malignancy Detection in Breast Histopathological Images', *Medical Imaging 2012: Computer-Aided Diagnosis*, 2012, pp. 831515
- [60] Bloom, H. and Richardson, W.: 'Histological Grading and Prognosis in Breast Cancer: A Study of 1409 Cases of Which 359 Have Been Followed for 15 Years', *British journal of cancer*, 1957, **11**, (3), p. 359
- [61] Andrade, D.V. and de Figueiredo, L.H.: 'Good Approximations for the Relative Neighbourhood Graph', *CCCG*, 2001, pp. 25-28
- [62] Haralick, R.M., Shanmugam, K., and Dinstein, I.H.: 'Textural Features for Image Classification', *IEEE Transactions on systems, man, and cybernetics*, 1973, (6), pp. 610-621
- [63] Clausi, D.A.: 'An Analysis of Co-Occurrence Texture Statistics as a Function of Grey Level Quantization', *Canadian Journal of remote sensing*, 2002, **28**, (1), pp. 45-62
- [64] Knime, <https://www.knime.com/>, accessed: 15/11/2020
- [65] Pudil, P., Novovičová, J., and Kittler, J.: 'Floating Search Methods in Feature Selection', *Pattern recognition letters*, 1994, **15**, (11), pp. 1119-1125
- [66] Eren, L., Ince, T., and Kiranyaz, S.: 'A Generic Intelligent Bearing Fault Diagnosis System Using Compact Adaptive 1d Cnn Classifier', *Journal of Signal Processing Systems*, 2019, **91**, (2), pp. 179-189
- [67] Zhang, Y., Jin, R., and Zhou, Z.-H.: 'Understanding Bag-of-Words Model: A Statistical Framework', *International Journal of Machine Learning and Cybernetics*, 2010, **1**, (1-4), pp. 43-52
- [68] Ali, N.M., Karis, M.S., Abidin, A.F.Z., *et al.*: 'Traffic Sign Detection and Recognition: Review and Analysis', *Jurnal Teknologi*, 2015, **77**, (20)
- [69] Li, T., Mei, T., Kweon, I.-S., *et al.*: 'Contextual Bag-of-Words for Visual Categorization', *IEEE Transactions on Circuits and Systems for Video Technology*, 2010, **21**, (4), pp. 381-392