ÇUKUROVA UNIVERSITY
INSTITUTE OF NATURAL AND APPLIED SCIENCES

MSc. THESIS

Harun-Ur-RASHID

ANALYSIS OF BREAST CANCER CLASSIFICATION ROBUSTNESS WITH RADIOMICS FEATURE EXTRACTION AND DEEP LEARNING TECHNIQUES

DEPARTMENT OF ELECTRICAL AND ELECTRONICS ENGINEERING

ADANA-2021

ABSTRACT

MSc. THESIS

ANALYSIS OF BREAST CANCER CLASSIFICATION ROBUSTNESS WITH RADIOMICS FEATURE EXTRACTION AND DEEP LEARNING TECHNIQUES

Harun-Ur-RASHID

ÇUKUROVA UNIVERSITY
INSTITUTE OF NATURAL AND APPLIED SCIENCES
DEPARTMENT OF ELECTRICAL AND ELECTRONICS ENGINEERING

Supervisor : Prof. Dr. Ulus ÇEVİK
Co-Supervisor : Prof. Dr. Turgay IBRIKCI
Year: 2021, Pages: 71
Jury : Prof. Dr. Ulus ÇEVİK
: Assoc. Prof. Dr. Sami ARICA
: Assist. Prof. Dr. Erkut TEKELI

The diagnostic of breast cancer and breast imaging procedures are typically carried out using a variety of imaging modalities, including mammography, MRI, and Ultrasound. However, Ultrasound and mammography have limitations. MRI is better than other procedures. Recent computational approaches, such as the Radiomics applied to image analysis, have shown remarkable progress for removing diagnostic difficulties. This thesis analyzed the robustness of breast tumor classification with features extraction (radiomics) and featureless method (deep learning). It contains two stages: the first stage introduced and explored radiomics based steps. A total of 111 tumor lesions were used to derive 74 radiomic features consisting of shape, first-order, and three separate second-order metrics. Four separate associations of features were used to classify tumor lesions with four different kernels from support vector machine algorithm. Second-order defined data split showed better cross-validation performance with highest accuracy of 96.17%, where all feature combinations data split showed 96.08% accuracy. In the confusion matrix analysis, the SVM-RBF kernel developed optimal diagnostic efficiency with a maximum test accuracy of 97.06% on two separate combination data group analysis. The second stage developed with deep learning techniques (InceptionV3 and CNN-SVM). A total of 2998 images were used to create the models. In this portion, the CNN-SVM model achieved the highest accuracy, 95.28%, with an AUC of 0.974, where the pre-trained InceptionV3 achieved an AUC of only 0.932. Finally, the obtained result in both stages was discussed together and other related studies.

Keywords: Breast Tumor Classification, Radiomic Features, Deep Learning.

ÖZ

YÜKSEK LİSANS TEZİ

RADYOMİK ÖZELLİK EKSTRAKSİYONU VE DERİN ÖĞRENME TEKNİKLERİ İLE MEME KANSERİ SINIFLANDIRMA SAĞLAMLIĞININ ANALİZİ

Harun-Ur-RASHID

ÇUKUROVA ÜNİVERSİTESİ
FEN BİLİMLERİ ENSTİTÜSÜ
ELEKTRİK-ELEKTRONİK MÜHENDİSLİĞİ ANABİLİM DALI

Danışman : Prof. Dr. Ulus ÇEVİK
Ikinci Danşman : Prof. Dr. Turgay İBRİKCİ
Year: 2021, Sayfa: 71
Jüri : Prof. Dr. Ulus ÇEVİK
: Doç. Dr. Sami ARICA
: Dr. Öğretim Üyesi Erkut TEKELİ

Göğüs kanserinin teşhisi ve göğüs görüntüleme işlemlerinde tipik olarak, mamografi, MRI ve Ultrason gibi çeşitli görüntüleme araçları kullanılarak gerçekleştirmektedir. Ancak, Ultrason ve mamografinin görüntülemelerinin sınırlamaları vardır. MRG ise diğer kullanılan yöntemlere göre daha iyidir. Görüntü analizine uygulanan Radiomics gibi son zamanların hesaplama yaklaşımları, teşhis zorluklarını ortadan kaldırmak için dikkate değer bir ilerleme göstermiştir. Bu tez, meme tümörü sınıflandırmasının sağlamlığını özellik çıkarma (radyomik) ve özelliksiz yöntem (derin öğrenme) ile analiz etmektedir. İki ana bölümden meydana gelmektedir: ilk bölüm, radyomik temelli adımlar araştırıldı ve on uygun uygulamalar gerçekleştirilmişti. Görüntü üzerinden, birinci derece ve üç ayrı ikinci derece metrikten oluşan 74 radyomik özelliklerini türetebilmek için toplam 111 tümör lezyon görüntüsü kullanılmıştır. Destek vektör makine algoritmasından dört farklı çekirdek ile tümör lezyonlarını sınıflandırmak için dört ayrı özellik ilişkilendirmesi kullanılmıştır. İkinci dereceden tanımlanmış veri bölümleme analizi ile %96,17'lik en yüksek doğrulukla daha iyi sonuçlar ortaya koymuştur. Burada tüm özelliklerin kombinasyonları veri bölümleme ile de %96,08 doğruluk başarısı göstermiştir. Karışıklık matrisi analizinde, SVM-RBF çekirdeği, iki ayrı kombinasyon veri grubu analizinde maksimum %97.06 doğrulukla optimum tanı verimliliğini geliştirdi. İkinci bölümde de ise, derin öğrenme tekniklerinden InceptionV3 ve CNN-SVM kullanılmıştır. Modelleri oluşturmak için toplam 2998 görüntü kullanmıştır. Bu bölümde, CNN-SVM modeli, önceden eğitilmiş InceptionV3'ün yalnızca 0,932'lik bir AUC elde ettiği, 0,974'lük bir AUC ile en yüksek doğruluğu %95,28'e ulaşmıştır.

Anahtar Kelimeler: Meme Tümörü Sınıflandırması, Radyomik, Derin Öğrenme

EXTENDED SUMMARY

Breast cancer is the most common type of cancer in women. According to a report by the American Cancer Society, over 2 million women were diagnosed with breast cancer in 2018. Although the number of cancer deaths among American women has increased over the past nine decades, breast cancer was the second leading cause of death among women from 1930 to 2016, but the death rate has dropped by 40 percent in the last 40 years. Early detection is responsible for this drop. The diagnosis of cancer is a highly complex procedure in medical research, and numerous tests are needed for appropriate diagnosis. Recent advancements in medical imaging technology, particularly in artificial intelligence-based image analysis, can significantly help areas of medical practice by providing real-world difficulties in detecting cancer, evaluating treatment, and tracking disease progression. The ultimate goal of medical imaging is to see the disease as early as possible, detect tumors, predict effects on the patient's physical well-being, and better control.

Typically, diagnostic and imaging operations are performed utilizing a range of imaging modalities, including mammography, magnetic resonance imaging, and ultrasound. However, there are drawbacks to both mammography and ultrasonic imaging methods. Breast MRI offers the highest effectiveness for detecting breast cancer of any contemporary clinical imaging modality and is crucial in breast imaging operations. While T1-weighted contrast-enhanced imaging is the foundation for breast MRI, T2-weighted and diffusion-weighted imaging are utilized to characterize lesions further. The MR evaluation of breast tumors enables accurate differentiation of benign from malignant lesions. The approach to customized medicine naturally leads to quantitative analysis of medical imaging. Advances in artificial intelligence for the interpretation of massive amounts of figurative data from various imaging technologies are critical in identifying cancer. Three distinct

significant techniques have been discovered under this paradigm: Radiomics, Machine Learning, and Deep Learning.

The Radiomics quantitative characteristics extraction method is designed to get a number of different images from a defined area of interest (ROI). This process is divided into several stages. At first, the process requires the gathering of image data. After the preprocessing phases are completed, the images are segmented to determine the location of the tumor. The ROIs have been found by a semi-automatic segmentation methodology known as thresholding-based segmentation. A series of radiological images, initially taken for conventional purposes, is then used to locate areas of interest (ROIs) and their functionality. Once the ROIs have been found, radiomics features are retrieved from the ROIs for statistical evaluation. In the last few years, an increasing number of user-friendly and open-access software is available to extract radiomics features from medical images. The most prominent examples include MaZda, PyRadiomics, TexRAD, LIFEx, MIM, ONCOradiomics, and 3DSlicer. The features, in this case, were extracted by setting up the PyRadiomics setup.

The computational radiomics properties are most frequently classified as Shape, First-Order Statistics, Second-Order Statistics, and Higher-Order Statistics. Shape, First-Order Statistics, and three Second-Order Statistics characteristics (GLCM, GLDM, and NGTDM) were employed in this analysis to differentiate between benign and malignant breast lesions. The extracted features are studied using a variety of statistical models to machine learning methods. Traditionally, the majority of radiomics research has been conducted using traditional classification methods such as Bayesian (BY) techniques, Boosting (BST), Decision trees (DT), Discriminant analysis (DA), or support vector machine (SVM). This thesis employed the SVM technique to verify classification accuracy using four different kernels (Linear, RBF, Polynomial, and Sigmoid).

More recently, deep learning has helped many areas enhance their accuracy. Due to deep learning's success in various medical applications, this work suggested

two deep learning-based systems for breast tumor classification. The thesis implemented convolutional neural networks (CNNs) followed by a support vector machine-based classifier. Additionally, it examined an InceptionV3 pre-trained model. In this instance, the dataset was used entirely imbalanced. Thus, a balance data generator with an up-sampling approach and data augmentation approach was created in combination to address the issue of data imbalance. The image preprocessing and segmentation phases were performed using an open-source image processing application called Fiji (ImageJ). The complete evaluation process was written and implemented on the Google Colab Cloud platform, designed for Python code development.

Finally, the radiomics with machine learning and deep learning models were compared to other state-of-the-art investigations. After studying the literature related to both processes, the thesis concluded that the research was sufficiently acceptable in the database that used, and the model strategies was adopted.

GENİŞLETİLMİŞ ÖZET

Meme kanseri kadınlarda en sık görülen kanser türüdür. Amerikan Kanser Derneği tarafından hazırlanan bir rapora göre, 2018'de 2 milyondan fazla kadına meme kanseri teşhisi kondu. Amerikalı kadınlar arasında kanser sebepli ölümlerin sayısı son doksan yılda artarken, meme kanseri 1930'dan 2016'ya kadar ölüm nedenlerinde ikinci sırada yer aldı. Ancak erken teşhis sayesinde ölüm oranı son 40 yılda yüzde 40 düştü. Kanser teşhisi son derece karmaşık bir işlemdir ve doğru teşhis için çok sayıda test gereklidir. Tıbbi görüntüleme teknolojisinde özellikle yapay zeka tabanlı görüntü analizindeki son gelişmeler; kanseri tespit etmede, tedaviyi değerlendirmede ve hastalığın ilerleme sürecinin takibinde zorlukların üstesinden gelmeye önemli ölçüde yardımcı olur.

Tipik olarak tanı ve görüntüleme işlemleri, mamografi, manyetik rezonans görüntüleme ve ultrason dahil olmak üzere bir dizi görüntüleme yöntemi kullanılarak gerçekleştirilir. Bununla birlikte, hem mamografi hem de ultrasonik görüntüleme yöntemlerinin dezavantajları vardır. Meme MRI, herhangi bir çağdaş klinik görüntüleme modalitesinin meme kanserini saptamak için en yüksek etkinliği sunar ve bu nedenle meme görüntüleme operasyonlarında çok önemlidir. T1 ağırlıklı kontrastlı görüntüleme meme MRI'nin temelini oluştururken, lezyonları daha iyi karakterize etmek için T2 ağırlıklı ve difüzyon ağırlıklı görüntüleme kullanılır. Meme tümörlerinin MR değerlendirmesi, iyi huylu ve kötü huylu lezyonların doğru bir şekilde ayırt edilmesini sağlar. Kişiye özel tıbbî yaklaşım, doğal olarak tıbbi görüntülemenin nicel analizine yol açar. Çeşitli görüntüleme teknolojilerinden elde edilen büyük miktarlardaki figüratif verilerin yorumlanması için yapay zekadaki gelişmeler, kanserin tanımlanmasında kritik öneme sahiptir. Bu paradigma altında Radyomik, Makine Öğrenimi ve Derin Öğrenme dahil olmak üzere üç farklı ana teknik keşfedilmiştir.

Radiomics nicel karakteristikleri çıkarma yöntemi, tanımlanmış bir ilgi alanından (ROI) bir dizi farklı görüntü elde etmek için tasarlanmıştır. Bu süreç birkaç

aşamaya ayrılmıştır. İlk başta, süreç görüntü verilerinin toplanmasını gerektirir. Ön işleme aşamaları tamamlandıktan sonra, tümörün yerini belirlemek için görüntüler segmentlere ayrılır. ROI'ler, eşiklemeye dayalı segmentasyon olarak bilinen yarı otomatik bir segmentasyon metodolojisi ile bulunmuştur. Başlangıçta geleneksel amaçlarla alınan bir dizi radyolojik görüntü daha sonra ilgi alanlarını (ROI'ler) ve bunların işlevselliğini belirlemek için kullanılır. ROI'ler belirlendikten sonra, istatistiksel değerlendirme için ROI'lerden radyomik özellikler alınır. Son birkaç yılda, tıbbi görüntülerden radyomik özellikleri çıkarmak için artan sayıda kullanıcı dostu ve açık erişimli yazılım mevcuttur. En belirgin örnekler MaZda, PyRadiomics, TexRAD, LIFEx, MIM, ONCOradiomics ve 3DSlicer'dir. Bu durumdaki özellikler, PyRadiomics kurulumu ayarlanarak çıkarıldı.

Hesaplamalı radyomik özellikler en sık Şekil, Birinci Dereceden İstatistikler, İkinci Dereceden İstatistikler ve Yüksek Dereceden İstatistikler olarak sınıflandırılır. İyi huylu ve kötü huylu meme lezyonlarını ayırt etmek için bu analizde Şekil, Birinci Derece İstatistik ve üç İkinci Derece İstatistik özelliği (GLCM, GLDM ve NGTDM) kullanıldı. Çıkarılan özellikler, makine öğrenimi yöntemlerine yönelik çeşitli istatistiksel modeller kullanılarak incelendi. Geleneksel olarak radyomik araştırmalarının çoğu, Bayesian (BY) teknikleri, Artırma (BST), Karar ağaçları (DT), Diskriminant analizi (DA) veya destek vektör makinesi (SVM) gibi geleneksel sınıflandırma yöntemleri kullanılarak yürütüldü. Tezimizde, dört farklı çekirdek (Doğrusal, RBF, Polinom ve Sigmoid) kullanarak sınıflandırma doğruluğunu ölçmek için SVM tekniğini tekniği kullanıldı.

Son zamanlarda, derin öğrenme birçok çalışma alanında doğruluğun artmasına yardımcı oldu. Derin öğrenmenin çeşitli tıbbi uygulamalardaki başarısı nedeniyle, bu tezde meme tümörü sınıflandırması için iki derin öğrenme tabanlı sistem önerdik. Evrişimli sinir ağları (CNN'ler) ve ardından bir destek vektörü makine tabanlı sınıflandırıcı bu çalışmada kullanıldı. Ek olarak, InceptionV3 adlı önceden eğitilmiş bir modeli inceliyoruz. Bu örnekte, veri seti tamamen dengesiz kullanılmıştır. Böylece, veri dengesizliği sorununu ele almak için bir yukarı-

örnekleme yaklaşımı ve veri büyütme yaklaşımına sahip bir denge veri oluşturucu bir arada oluşturulmuştur. Görüntü ön işleme ve segmentasyon aşamaları, Fiji (Image) adlı açık kaynaklı bir görüntü işleme uygulaması kullanılarak gerçekleştirilmiş ve tüm değerlendirme süreci Python ile kod geliştirme için tasarlanmış Google Colab platformunda yazılıp uygulanmıştır.

Son olarak, makine öğrenimi ve derin öğrenme modellerine sahip radyomik, diğer son teknoloji araştırmalarla karşılaştırıldı. Her iki süreçle ilgili literatürü inceledikten sonra, kullandığımız veri tabanında ve benimsediğimiz model stratejilerinde aramalarımızın yeterince kabul edilebilir olduğu sonucuna varıldı.

ACKNOWLEDGEMENTS

I want to thank Prof. Dr. Ulus ÇEVİK for his support and motivation during the academic year and for his approval to be a member of jury. I am very grateful to him for his direction, motivation, constructive advice with this thesis.

I also extend my deepest gratitude to Prof. Turgay İBRİKCİ. I cannot say thank you enough for his tremendous support and help. I am very grateful to him for his supervision, direction, motivation, constructive advice, and continued faith in me with this thesis. I feel motivated and encouraged every time I attend his meeting. Without his encouragement and guidance this thesis would not have materialized.

I want to thank Assoc. Prof. Dr. Sami ARICA, as well for his consent to be a member of the jury.

I owe extra appreciation to Prof. Semra PAYDAŞ for sharing her knowledge and experience on breast cancer. Also, for supporting me a great deal by sharing the original dataset.

I would like to express my gratitude to Assist. Prof. Dr. Erkut TEKELİ for his willingness to participate on the jury as well.

I want to share my sincere thanks and gratitude to Turkish Government, Presidency for Turks Abroad and Related Communities (YTB) and Turkish people for giving me this opportunity and for their hospitality.

Finally, I wish to express my heartfelt thankfulness to my family for their patience, encouragement, and continuous moral support.

CONTENTS **PAGE**

LIST OF TABLES **PAGE**

LIST OF FIGURES **PAGE**

LIST OF ABBREVIATIONS

AI	: Artificial Intelligence
AT	: Acquisition Time
AUC	: Area Under Curve
BCDR-F03	: Breast Cancer Digital Repository
BI-RADS	: Breast Imaging Reporting and Database System score
CAD	: Computer Aided Design
CART	: Classification And Regression Trees
CBIS-DDSM	: Enterprise Computer-Based Information System
CNN	: Convolutional Neural Network
DBT	: Digital Breast Tomosynthesis
DCNN	: Deep Convolutional Neural Network
DDSM	: Digital Database for Screening Mammography
DICOM	: Digital Imaging and Communications in Medicine
DCE	: Dynamic Contrast Enhanced
DWI	: Diffusion-Weighted Imaging
ET	: Repetition Time
FFDM	: Full Field Digital Mammograms
FPR	: False Positive Rate
GLCM	: Gray-Level Co-Occurrence Matrix
GLDM	: Gray Level Dependence Matrix
GPU	: Graphics Processing Unit
GUI	: Graphical User Interface
GLRLM	: Gray Level Run Length Matrix
IoT	: Internet of Things

JPEG	: Joint Photographic Experts Group
KNN	: k-Nearest Neighbors
LDA	: Linear Discriminant Analysis
LSTM	: Long Short-Term Memory
MIAS	: Mammographic Image Analysis Society
MLP	: Multi-Layer Perceptron
MRI	: Magnetic Resonance Imaging
MVFF	: Multi-View Feature Fusion
NA	: Number of Average
NGTDM	: Neighborhood Gray-Tone Difference Matrix
NRRD	: Nearly Raw Raster Data
OMI-DB	: OPTIMUM Mammography Image Database
PCA	: Principal Component Analysis
QDA	: Quantitative Descriptive Analysis
PPPOV	: Percent Phase Field of View
RAM	: Random Access Memory
RBF	: Radial Basis Function
ReLU	: Rectified Linear Activation Function
ResNet	: Residual Network
ROC	: Receiver Operating Characteristic
ROI	: Region Of Interest
STIR	: Short Tau Inversion Recovery
SVM	: Support Vector Machine
T2W-TSE	: T2 Weighted-Turbo Screen Echo
T2W-SPAIR	: T2 Weighted-Spectral Attenuated Inversion Recovery
TPR	: True Positive Rate

TPU	: Tensor Processing Unit
VOI	: Volume Of Interest
VGG16	: Visual Geometry Group
XGBoost	: Extreme Gradient Boosting

1. INTRODUCTION

1.1. Background and Research Motivation

Breast cancer is one of the leading causes of female mortality. Breast cancers that are diagnosed early have a slightly higher chance of having a favorable clinical outcome. Medical imaging has long been recognized as a reliable tool for the early detection of cancer and monitoring patients during and after chemotherapy or surgery. Artificial intelligence advances in analyzing vast amounts of interpretive image data produced by various imaging methods contribute significantly to tumor detection. Three fundamental approaches to goal-oriented research can be differentiated in this context: radiomics, machine learning, and deep learning. The increase in science journals is shown in Figure 1, which shows that the number of publications using the radiomics, MRI, and deep learning techniques discussed in this study has increased annually since 2016.

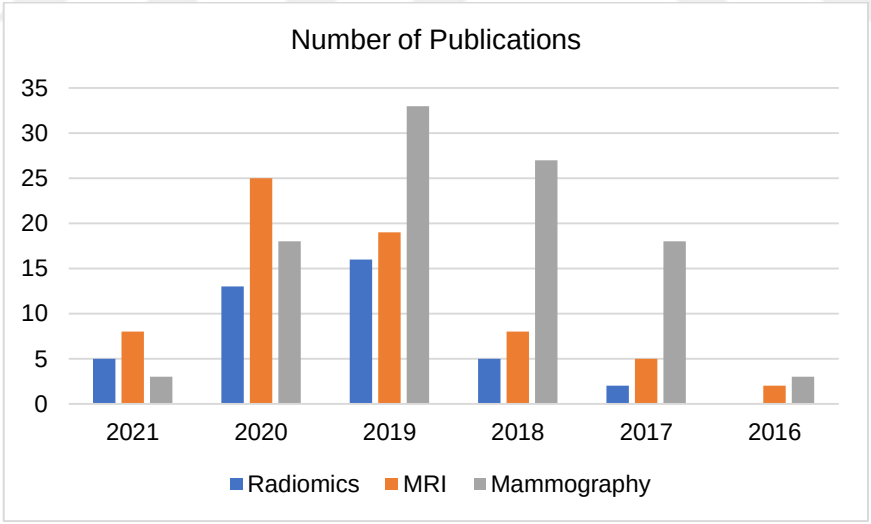

Figure 1.1. Number of publications per year in PubMed containing “machine learning”, “breast cancer”, “classification” and one of the three modality keywords (MRI, mammography, radiomics) from 2016 to 2021. (Queried: April 9, 2021)

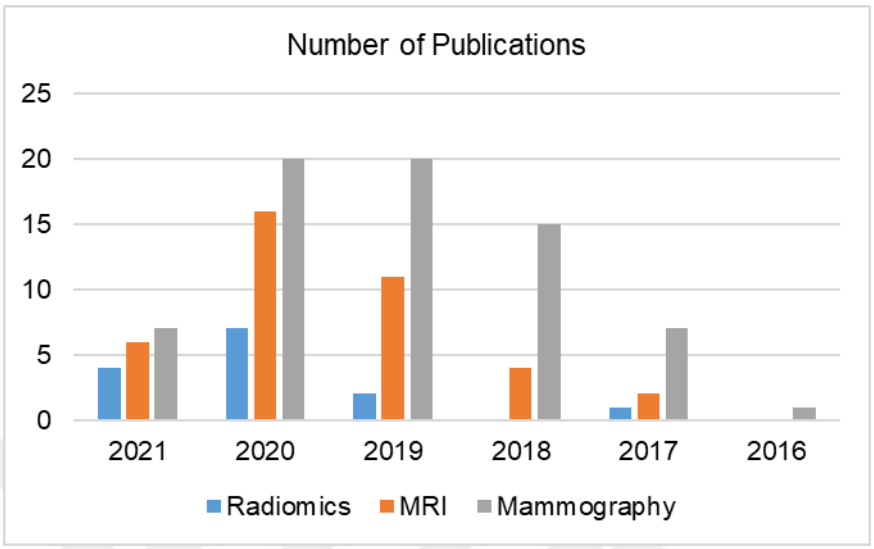

Figure 1.2. Number of publications per year in PubMed containing “deep learning”, “breast cancer”, “classification” and one of the three modality keywords (MRI, mammography, radiomics) from 2016 to 2021. (Queried: April 9, 2021)

Radiomics is a rapidly growing discipline of medical image analysis field used to investigate tumor information. In breast cancer classification radiomics is a new field. To develop radiomics process it includes different steps. At the beginning original DICOM image used for tumor identification. Next steps it includes tumor segmentation and feature extraction process. Finally extracted feature data are organized and develops a dataset for analysis.

After finishing the work, the breast cancer classification (benign or malignant) will be facilitated, and care will begin early. A comprehensive idea to find out the robustness of classification accuracy based on deep learning and radiomics will provided.

1.2. Objectives of the Thesis

The principal objectives of this study are to classify the breast tumor classes (benign and malignant) using radiomics method and also the deep learning method. After that, the classification robustness of above-mentioned method will be analyzed briefly.

- To concentrate on the uses of radiomics features in the classification of breast tumor lesion.
- To explore the suitability of radiomics in distinguishing malignant from benign breast tumor lesions.
- To enable comparisons with more established image-based techniques.
- To compare the classification robustness between radiomics and deep learning-based cancer classification.

After completing the proposed work, it is expected that the accuracy of diagnoses and therapy responses assessment will be improved. It also increases the clinical applicability of artificial intelligence. Classification robustness also ensure the best method to classify the breast tumor. This will provide new approaches to handling the problems in large data management.

It believes that radiomics is rapidly moving beyond the realm of specialized study and is establishing itself as a translational technology. The purpose of this paper is to familiarize a broad audience of practicing clinicians, including radiologists, with the practice of radiomics.

1.3. Proposed Idea of the Thesis

At present days, most of the clinical diagnosis are implemented by smart solutions like internet of things (IoT) and artificial intelligence. Therefore, thesis shows the ability to use radiomics feature for tumor lesion analysis and also deep

learning techniques for cancer class differentiation. This work shows the way to develop radiomics process and also the deep learning process and their robustness analysis. A block diagram represents the whole idea to develop and implement this thesis.

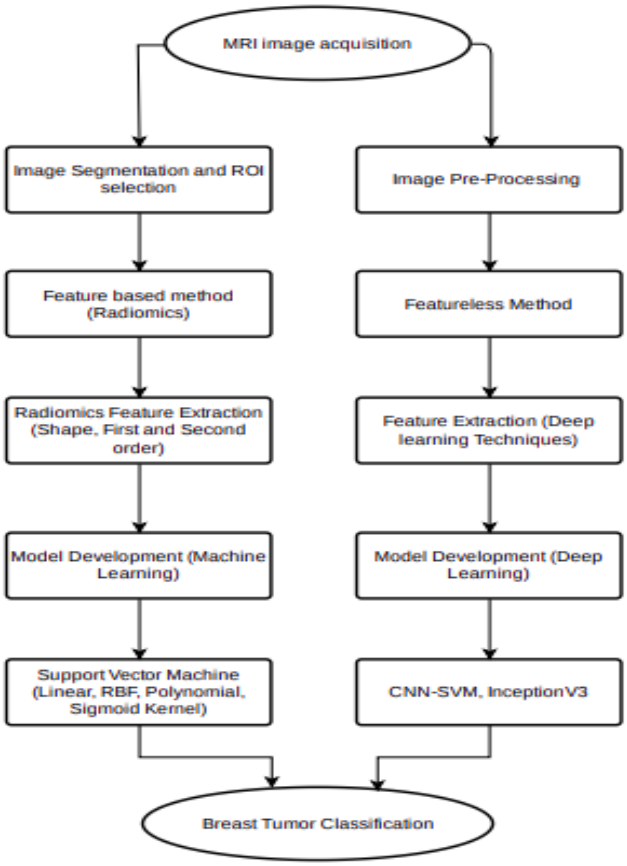

Figure 1.3. A block diagram of proposed work

This work is separated by two different steps. In the first step, ROI is selected by segmentation in the selected images from the raw dataset. These ROIs are then used to extract tumor features from the original images. Extracted information are analyzed with machine learning algorithm for classification purposes. This process formed as Feature Based analysis.

In the second step, the images in the image dataset are input directly into the two different deep learning model. Deep learning analysis automatically extracts features so there is no need to save new datasets separately. Again, the extracted information is analyzed automatically, that means no separate algorithm is needed. This is why this step is called featureless method.

Therefore, this studies aim is to determine the robustness of cancer classification by radiomics features extraction and deep learning with respect to variability in semi-automatic breast tumor segmentation on MRI.

1.4. Outline of the Thesis

Chapter 1 presents an overview of the motivation and problems of the present day's research summary about breast cancer classification and the major objectives of this thesis. This chapter also contains a summary representation of the proposed work. **Chapter 2** gives a literature review of related studies with similarities, gaps, limitations, and concepts. The datasets that are used in this thesis are discussed in **Chapter 3.** The mathematics of radiomics features, feature extraction process, and algorithms for both machine learning and deep learning are also represented in **Chapter 3**. In **Chapter 4**, the obtained results are mentioned. **Chapter 5** contains a brief discussion about proposed techniques. Finally, the limitation and some advice for how to overcome these limitations are addressed in **Chapter 6**.

2. LITERATURE REVIEW

In this chapter, its covered most of the recent studies that are related to radiomics based breast lesion classification and Magnetic Resonance image analyzed deep learning-based classification. It is found that the majority of these results may contain some intriguing possibilities that must be checked and compared to this dissertation. Though this thesis has analyzed two different methods and compared them, so that the related studies of these methods were also presented sequentially.

Over the last few years, a great deal of work has been done on medical image processing field where deep learning, and radiomics coming to be two of the most common techniques. Radiomics is a relatively new technique that employs a set of sophisticated mathematical operations to convert imaging data to a high-dimensional feature vector (Aerts et al., 2014; Yip & Aerts, 2016). Many researchers have studied the correlation between such features and clinical variables as well as treatment outcome and tumor type classification (benign vs. malignant) to identify potential mechanisms of radiomics (Keek et al., 2018). Breast cancer has also been the subject of extensive studies in radiomics (Conti et al., 2020; Tagliafico et al., 2020).

Radiomics (Gillies et al., 2016) is an imaging area that is increasingly flourishing and advancing in recent years. It can generalize image data and extract high-dimensional computational features based on shape, strength, and texture to provide information about the tumor phenotype and surroundings (Aerts et al., 2014; Kumar et al., 2012). These features are computational results of different image-processing and data optimization algorithms of the first, second or higher-order statistics.

Radiomic features are intended to extract the distinctive details of benign or malignant nature by quantifying various texture properties (Lambin et al., 2012). More precisely, first-order features are calculated on a Region of Interest (ROI) or Volume of Interest (VOI) utilizing voxel intensity distribution, whereas second-

order radiomic features (GLCM, GLDM) calculate the perceptual representation in texture image of voxel intensity values (Velichko et al., 2020).

Several researchers applied radiomics methods to identify the different forms of cancers using imaging techniques. For example, MRI-based radiomics research helps to differentiate subtle distinctions between various tumor phenotypes and has been conducted for clinical manifestations of breast and lung tumor type (Ardakani et al., 2015; Lang et al., 2019; Nam et al., 2019).

MRI has several advantages for identifying breast cancer lesions. MRI is capable of multi-planar scanning and 3D reconstruction, which allows for a more accurate display the different features of a selective lesion like size, shape, breast lesions' position (Honda et al., 2016). However, MRI can distinguish benign from malignant breast cancer in approximately 72% of cases (Rankin, 2000). Indeed, its specificity is restricted by a number of image-quality-related factors including magnetic field and gradient power, coil efficiency, contrast agent effectiveness, and menstrual cycle (Zhou et al., 2015).

Kaya et al. (Kaya et al., 2017) used brain MRI to segment the tumor and classified it via PCA which resulted in the inexplicit identification of tumor shape assisting to the risk of human life. Image sequences involved in classical procedure such as T2-weighted, Short-Tau-Inversion-Recovery (STIR), sdyn-eTHRIVE and Weighted Diffusion (DW) can satisfactorily be used (Thakran et al., 2018) for this purpose.

Pareckh et al. (V. S. Parekh & Jacobs, 2017) developed a feature map by extracting features from Magnetic Resonance (MR) images and analyzed it with the IsoSVM model where the classifier expressed 93% sensitivity, 85% specificity and 91% AUC. Similarly, Bickelhaupt et al. (Bickelhaupt et al., 2017) also collected statistical features from Magnetic Resonance (MR) images (DWI-MR and DCE-MR) and continued by Lasso-supervised methods to classify them. This model has helped separate the malignant tumor from benign lesions with an AUC of 84.2% (DCE-MRI) and 85.1% (DWI-MR).

Whitney et al. (Whitney et al., 2019) have used DCE-MRI-derived radiomic features to help distinguish benign lesions from luminal A breast tumors. This classification exercise showed how features such as entropy and irregularity proved valuable by achieving an AUC rate of 72.9% with a linear discriminant analysis (LDA) algorithm.

Mao et al. (Mao et al., 2019) revealed the characteristics of quantitative radiomics coupled with mammographic images and used machine learning algorithm to extract features. Using logistic regression on 51 features of radiomics, they found 88.6% accuracy, a better outcome than experienced radiologists.

Recently, researchers proposed that radiomics retrieved features derived from two distinct methods of image collection, most likely ultrasound imaging and digital breast tomosynthesis (DBT), were successful at differentiating malignant from benign lesions (Tagliafico et al., 2018; Q. Zhang et al., 2017). Specifically, Zhang et al. (Q. Zhang et al., 2017) used hierarchical clustering to pick seven important characteristics and monitored the testing accuracy, sensitivity and precision where the results were 88.0%, 85.7%, and 89.3% respectively. Tagliafico et al. (Tagliafico et al., 2018) identified three radiomic features (energy, entropy, and dissimilarity) that significantly distinguished cancerous from typical breast tissue of patients. However, researchers' findings indicated that these characteristics cannot be used in isolation to analyze breast lesion well. They achieved very poor AUC 56.7%.

Another study was reported that authors extracted three different features and analyzed them with and without making combination. Individual analysis of different feature and their combinational analysis model was developed by Park et al. (B. E. Park et al., 2016) where support vector machine algorithm were used for recognizing a rotator cuff supraspinatus tendon tear. The author considering five (5) first order, 52 second-order features (GLCM-40 and GLRLM-12). From the results, first order and second order (GLCM and GLRLM) features give 95%, 85%, and

100% accuracy, respectively. The association of all features (First order, GLCM and GLRLM) achieved highest 100% accuracy.

In comparison to radiomics, which use pre-defined (manual extraction) features, deep convolutional networks learn and retrieve excellent information from training data for specific classification tasks. However, a significant barrier to the success of deep convolutional networks is a scarcity of sufficiently large datasets (Marentakis et al., 2021). The extensive use of deep learning in medical image processing and classification has recently emerged as the most widely used AI strategy in medical science (Bengio et al., 2013).

Recently, different academic groups such as the Visual Geometry Community and Google researcher used sophisticated deep learning methods to design and implement the VGG-16, ResNet, GoogLeNet and so on models (Nahid & Kong, 2018). Among these sophisticated modeling methods focused on convolutional neural networks (CNNs), which are used to increase the performance of breast cancer diagnosis (Toğaçar et al., 2020). Consider the following, Khan et al. (S. Khan et al., 2019) analyzed two distinct dataset (public and self-collected) by utilizing fine-tuned DCNN architectures VGGNet, GoogleNet, and ResNet-50. They categorized the fused features using the average pooling approach and achieved an accuracy of 97.67%. Furthermore, Li et al. (Li et al., 2017) developed a procedure using 3D CNNs to differentiate between malignant and benign breast tumors. They used their own dataset which belongs 143 patients, 66 of whom had benign and 77 of whom had malignant tumors. The accuracy (Acc), sensitivity (Sens), and specificity (Spec) of 3D CNN model were 0.781, 0.744, and 0.823, respectively. But their (Li et al., 2017) model accuracy were comparatively low.

In the same way, Khan et al. (H. N. Khan et al., 2019) analyzed two different public dataset DDSM and CBIS-DDSM for the labeling of mammograms. For this purpose, they suggested four deep learning architectures (VGG-16, VGG-19, GoogLeNet, and ResNet50) based on multi-view feature fusion (MVFF). They found the classification accuracy 96.66% and the AUC 93.4%.

Wu et al. (Wu et al., 2020) introduced a pre-trained deep learning model (ResNet) for classifying breast cancer screening exams that was trained and tested on over 200000 exams (over 1000000 images), with the network pre-trained on screening BI-RADS data. When tested on the sampling audience, the described model produced an AUC of 89.5% in forecasting the existence of breast cancer.

A fully automated method for mass detection in Full-Field Digital Mammograms (FFDM) using a Faster Region-based Convolutional Neural Network (Faster-RCNN) was addressed by Agarwal et al. (Agarwal et al., 2020). First, author developed the model by using OMI-DB Database and acquired a True Positive Rate (TPR) of 0.93 was obtained with a False Positive Rate (FPR) of 0.78 per image. Then, another public dataset INbreast, was tested on same model and obtained a TPR of 0.91± 0.06.

Jiang et al. (Jiang et al., 2017) classified breast tumor lesions in a new-found BCDR-F03 dataset with two different deep learning algorithm GoogLeNet and AlexNet. They received 88% and 83% AUC, respectively. However, Antropova et al. (Natasha Antropova et al., 2018) also achieved a good AUC 85% while classifying 703 MRI image data. This was achieved by another deep learning model named as long short-term memory network (LSTM).

Recently, the success of the deep convolutional neural network-based method for breast lesion classification was compared to that of radiologists and conventional artificial intelligence techniques. A CNN-based ResNet model was developed by Truhn et al. (Truhn et al., 2018). The model used for classifying the benign and malignant lesion from a private dataset which contains 447 patients with 787 malignant and 507 benign lesions from MRI image data. Finally, CNN achieved an AUC of 0.88, which is superior to that of a radiologist.

The area of deep learning-based transfer learning has been intensively investigated for the task of classifying breast cancers. Song et al. (Spanhol et al., 2016) offered a GoogLeNet, Inception-v2, and Inception-based CAD scheme for classifying three classes: normal, benign, and malignant. In final, fully connected

layer was replaced by XGBoost. They worked on a public DDSM dataset and original DCNN model achieved an accuracy of 82.84%; however, when the DCNN model was fine-tunned with XGBoost the classification accuracy was improved at 92.8%.

Mendel et al. (Mendel et al., 2019) used pre-trained VGG-19 and SVM to isolate and classify the features of 78 mammogram lesions. The AUC value obtained was 81%. According to Ekici et al. (Ekici & Jawzal, 2020), convolutional neural networks (CNNs) optimized with the Bayes algorithm also able to identify early symptoms of the disease and classifies the breast images as normal or suspected. To this end, they used 140 instances of thermal breast images and obtained an 98.95% accuracy. The main drawback of Ekici et al. (Ekici & Jawzal, 2020) is the limited number of training set. The number of training and testing were very small.

The deep learning algorithm incorporated significant bias into the final classification result, despite the limited size of the dataset. So, different method also applied by the researchers to overcome the data lacking problems. Data augmentation techniques were applied by Zhang et al. (X. Zhang et al., 2017). A transfer learning based neural network models employed for the classification and whole mammography images were inputted. The AUC was found 73%.

Maicas et al. (Maicas et al., 2017) present a deep reinforcement learning Q-method for the automatic identification of breast tumor lesions from a dataset containing 117 images from DCE-MRI module and authors found sensitivity 80% with an operating time 9221s.

3. MATERIAL AND METHODS

This chapter was covered data preparation, platforms of training and testing, image processing tools, Radiomics feature extraction methods mathematical expression, machine learning, and deep learning algorithms. Lastly, it represented the metrics, which were used to measure the performance of algorithms.

3.1. Data Preparation

3.1.1. Patient Population

MR images were collected systematically from number of patients with imaging conducted using the identic scanner and image procedure to eliminate interscan heterogeneity correlated with variations in imaging ingredient or the imaging arrangement. The age range of the patients was 39 to 85 years. Here we used data from 35 patients, of whom 10 had benign and 25 had malignant tumors in their breasts. A total of 111 tumor locations identified in the images obtained. The Department of Oncology at the Cukurova University Hospital and the Department of Electrical and Electronics Engineering accepted this systematic research, and patient confidentiality was strictly ensured under the committee of ethics rules during the whole work.

3.1.2. Training, Testing Platform and Tools

3.1.2.1. Fiji (ImageJ)

The images that were obtained using MRI machines were processed in open-source image analysis application Fiji, which incorporates plugins and macros.

Fiji combines fully featured libraries with a diverse set of scripting languages to rapidly prototype image-processing algorithms. Fiji simplifies the process of converting numerical algorithms into ImageJ plugins that can be distributed to end users via an incorporated upgrade scheme (Abramoff et al., 2004).

It extends its functionality by continuing to support a broad range of programming languages (Jython, Python, JavaScript, JRuby, and Bean shell), offering the most extensive set of programming tools available on both open-source platforms. These languages are not only accessible to those with no prior knowledge of Java, but also serve as a useful contrast to the macro language in terms of programming syntax, in the case of newcomers being at an adequate level.

It is cited in over 1000 peer-reviewed works. Fiji supports a huge number of various image formats and a big number of pixel-based actions. Additionally, images with bit depths larger than 8 or 16 bits per channel are supported. It can also be used as an image-conversion tool. Many image formats can be read natively by Fiji, and with the help of a plugin, many proprietary formats can be opened.

3.1.2.2. Google Colaboratory

We used Google Colaboratory, a powerful Python development environment. It is a free cloud service built on Jupyter notebooks for developing machine and deep learning projects. Colab is also integrated with Google Drive, allowing users to share and upload notebook and associated materials to the account.

Colab provides a runtime optimized for deep learning with unrestricted access to a robust GPU/TPU. The only constraint on Colab is that the runtime is interrupted after 12 hours of continuous operation. Occasionally, it can disconnect after an hour of inactivity. Additionally, we found that Colab comes with 12GB of RAM and can be expanded to 25GB during runtime if required. The internet connection in Colab is up to 120MB/s. We uploaded the original and mask images to Google Drive and then mounted them on Colab. Due to RAM crashes at the data augmentation stage, we exercised some restraint when running the codes. Colab comes pre-installed with nearly all of the required Python libraries, including TensorFlow, Keras, Pandas, Numpy, and Matplotlib.

3.1.3. MR Imaging

The MR imaging was conducted at the Cukurova University Hospital, Turkey, prior to a typical imaging procedure containing six different sequences. They are T2W-TSE (T2-Weighted Turbo Spin Echo), T2W-SPAIR (T2-Weighted Spectral Attenuated Inversion Recovery), STIR (Short Tau Inversion Recovery), DWI (Diffusion Weighted Imaging), 7dyn-eTHRIVE (7dynamic-Enhanced T1 High Resolution Isotropic Volume Excitation), and sdyn-eTHRIVE (Subtraction of Dynamic Enhanced T1 High Resolution Isotropic Volume Excitation). MRI was performed on a 3.0T MR scanner (Philips healthcare, Ingenia, 2016, Netherlands) with a bore of 70 cm with independent receive channel. The T2W-TSE, T2W-SPAIR, STIR and DWI series sequenced images were gathered with a 3mm slice thickness, but the sdyn-eTHRIVE and 7dyn-eTHRIVE slice thickness were 2mm. **Table 3.1** contains the above-mentioned sequence parameters such as matrix, echo time, number of averages, repetition time, and other values.

Table 3.1. MR image series parameters

Parameters	AT(s)	ET (ms)	RT (ms)	NA	Matrix	PPFOV (mm)
T2W	152102.86	120	4444	1	350×245	125
T2W-SPAIR	155013.72	70	4086	2	340×258	121.35
STIR	153945.20	65	4626	2	248×205	131.96
DWI	152627.18	93.31	13191	2	155×153	121.21
Sdyn	153327.76	2.14	4.3161	1	339×338	121.21
7dyn	153552.06	2.14	4.3161	1	339×338	121.21

A gadolinium-based contrast agent (gadobutrol, Gadovist; Bayer Healthcare, Berlin) which are relatively large molecules, can easily extravasate from such vessels, and therefore rapidly accumulate breast cancer stroma and was used 0.1 mmol/kg of body weight by intravenous.

3.1.4. Image Segmentation

All of the patient's MR images were converted from DICOM to NRRD format. Additionally, the surrounding areas around the breast region were cropped and saved in PNG format for deep learning analysis.

A semi-automated segmentation method has been used in the breast tumor segmentation process. The segmentation stage was conducted to contextualize each benign and malignant breast tumor's tumor in fragments of 2D and 3D MR images. An experienced radiologist from Cukurova University Hospital worked initially on the breast MR images to detect tumors. Figure 3.1 represented a segmented output. An open-source dedicated medical image processing program called ImageJ-based FIJI (Schindelin et al., 2012) has been used to build the process on breast MR image slices. Initially, a new stack has developed for each tumor lesion, distinguishing the initial tumor images from the images identified by a radiologist. Each wound-carrying stack has been transformed into an 8-bit grayscale image. Tumors located in these stacks have been controlled by changing the contrast and brightness values to represent them better. Different threshold methods are applied to stack images.

The OTSU threshold technique has shown an appropriate outcome in certain situations, and so did the MaxEntorpy. Then the outlier's modification has been added to eliminate the noise where the outlier's structure in white and black layers is radius = 0.5-2.0 pixels, threshold = 50; A mask has generated by designing these phases such that only the tumor area may be correctly found. Then the standard mask has transformed into a binary mask. Binary eroding and dilation methods have often employed in certain situations to classify the tumor correctly. Median filters have been applied to create all masks with a radius value of 1.0-2.0 pixels. The masks have been able to establish the exact position and boundaries of the tumor because of the overall work.

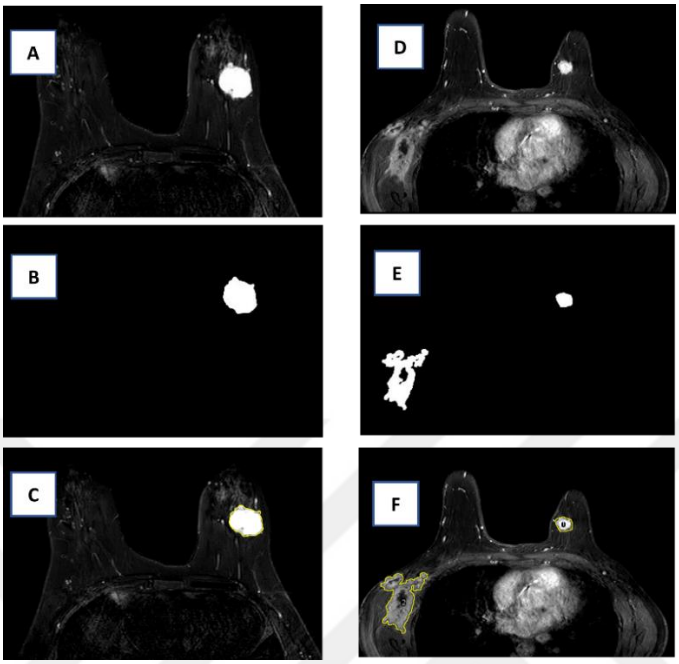

Figure 3.1. T2 weighted (T2W) image (left) and 7dynamic-Enhanced (7dyn) T1 High Resolution Isotropic Volume Excitation MR image (right); (A, D) original MR images; (B, E) segmented mask images, (C, F) Selected ROI

3.2. Radiomics Feature Extraction

Radiomics as a field seeks to derive quantitative, and preferably reproducible, information from medical images, which may include complex patterns that are difficult to recognize or quantify with the human open eye. (Mayerhoefer et al., 2020).

Semantic and agnostic features can be categorized under Radiomics. (Lewis et al., 2021). A brief overview of the various radiomics types mentioned in the *Table 3.2*. Semantic features are qualitative characteristics that are commonly used in radiology reports, such as the size, shape, position, and necrosis of lesions. Each of the agnostic features is described by a sophisticated mathematical method and can be classified as morphological or statistical. (Yang et al., 2018).

The morphological characteristics of a segmented volume characterize its form and physical structure. Additional classifications for statistical features include first order, second order, and higher order features. Without taking spatial relationships into account, first-order statistical features explain the distribution of pixel intensities within VOIs. (Just, 2014).

Table 3.2. Classification of radiomics features

Main Group	Subgroup	Name of techniques
Semantic	-	Size, shape, location, necrosis
Agnostic	Morphological	-First order (diameter, volume) -Higher order (Fractional dimension)
	Statistical	-Histogram based first order statistics, (energy, mean, kurtosis, skewness, entropy) -Second order texture (GLCM, GLDM, GLRLM, NGTDM, GLSZM) -Higher order (Wavelet, Gabor, Fourier, and Laplacian transforms)

According to (Conti et al., 2020), In particular, the second-order statistics that are texture features take into account spatial interdependence and co-occurrence of information through voxels close to each other. This gray-level co-occurrence matrix (GLCM) also alluded to as a gray tone spatial dependency matrix is among the second-order features most widely implemented in radiomics studies.

However, according to (Lewis et al., 2021), numerous texture analysis techniques generate hundreds of potential texture properties that can be incorporated into radiomics analysis. (Lewis et al., 2021; V. Parekh & Jacobs, 2016). Higher order radiomics techniques incorporate filters to obtain characteristics from images.

3.2.1. Shape Features

Shape description or representation is a significant issue in image analysis for object recognition and classification. Shape features are crucial because they offer a means of defining an object without relying on its most salient characteristics, thus reducing the amount of data stored.

Let,

The number of voxels is N_v within ROI

The number of faces (triangles) are N_f defining the Mesh.

The volume of the mesh V in mm^3 ,

The surface area of the mesh A in mm^2

Table 3.3. Mathematical explanation of the features of the shape features

Feature	**Description**
1. Mesh Volume	$Mesh\ Volume = \sum_{i=1}^{N_f} V_i$ where, $V_i = \frac{Oa_i.(Ob_i \times Oc_i)}{6}$
2. Voxel Volume (V_{voxel})	$V_{voxel} = \sum_{k=1}^{N_v} V_k$
3. Surface Area	$Surface\ Area, A = \sum_{i=1}^{N_f} A_i$ where, $A_i = \frac{1}{2}\lvert a_i b_i \times a_i c_i \rvert$
4. Surface to Volume Ratio	$surface\ to\ volume\ ratio = \frac{A}{V}$
5. Sphericity	$sphericity = \frac{\sqrt[3]{36\pi V^2}}{A}$
6. Maximum 3D diameter	described as the maximum pairwise Euclidean distance between the vertices of the tumor surface mesh. Additionally, referred to as Feret Diameter.
7. Maximum 2D diameter (Slice)	described as the maximum pairwise Euclidean distance in the row-column plane between tumor surface mesh vertices.
8. Maximum 2D diameter (Column)	described as the maximum pairwise Euclidean distance in the row-slice plane between tumor surface mesh vertices.
9. Maximum 2D diameter (Row)	described as the maximum pairwise Euclidean distance in the column-slice plane between tumor surface mesh vertices.
10. Major Axis Length	major axis=$4\sqrt{\lambda_{major}}$
11. Minor Axis Length	minor axis=$4\sqrt{\lambda_{minor}}$
12. Least Axis Length	least axis=$4\sqrt{\lambda_{least}}$
13. Elongation	elongation=$\sqrt{\frac{\lambda_{minor}}{\lambda_{major}}}$
14. Flatness	flatness=$\sqrt{\frac{\lambda_{least}}{\lambda_{major}}}$

3.2.2. First-order Radiomics

First-order features (Conti et al., 2020) are obtained from statistical instances of the image intensity histogram, such as kurtosis and skewness, which are characterized as the measure of the distribution of image voxel values "peakedness' and asymmetry around the mean value, respectively (Traverso et al., 2018).

Let, A set of N_p voxels are X included in ROI

The first order histogram with N_g discrete intensity levels are $P(i)$, where, the number of non-zero bins are N_g, equally spaced from 0 with a width.

The normalized first order histogram is $p(i) = \frac{P(i)}{N_p}$

Table 3.4. Mathematical explanation of features of the first order features

Feature	Description
1. Energy (E)	$E = \sum_{i=1}^{N_p}(X(i)+c)^2$
2. Total Energy (TE)	$TE = V_{voxel}\sum_{i=1}^{N_p}(X(i)+c)^2$
3. Entropy (Ep)	$Ep = -\sum_{i=1}^{N_g} p(i)log_2(p(i)+\epsilon)$
4. Minimum (Min)	$\text{Min} = \min(X)$
5. 10Percentile	The 10th percentile of X
6. 90Percentile	The 90th percentile of X
7. Maximum (Max)	$Max = \max(X)$
8. Mean	$mean = \frac{1}{N_p}\sum_{i=1}^{N_p} X(i)$
9. Median	The ROI exhibits the median gray-level intensity
10. Interquartile Range (IR)	$IR = P_{75} - P_{25}$
11. Range	$Range = \max(X) - \min(X)$
12. Mean Absolute Deviation (MAD)	$MAD = \frac{1}{N_p}\sum_{i=1}^{N_p}\lvert X(i) - \bar{X}\rvert$
13. Robust Mean Absolute Deviation (rMAD)	$rMAD = \frac{1}{N_{10-90}}\sum_{i=1}^{N_{10-90}}\lvert X_{10-90}(i) - \bar{X}_{10-90}\rvert$
14. Root Mean Squared (RMS)	$RMS = \sqrt{\frac{1}{N_p}\sum_{i=1}^{N_p}(X(i)+c)^2}$
15. Skewness	$\text{Skewness} = \frac{\mu_3}{\sigma^3} = \frac{\frac{1}{N_p}\sum_{i=1}^{N_p}(X(i)-\bar{X})^3}{\left(\sqrt{\frac{1}{N_p}\sum_{i=1}^{N_p}(X(i)-\bar{X})^2}\right)^3}$
16. Kurtosis (krts)	$krts = \frac{\mu_4}{\sigma^4} = \frac{\frac{1}{N_p}\sum_{i=1}^{N_p}(X(i)-\bar{X})^4}{\left(\frac{1}{N_p}\sum_{i=1}^{N_p}(X(i)-\bar{X})^2\right)^2}$
17. Variance	$variance = \frac{1}{N_p}\sum_{i=1}^{N_p}(X(i)-\bar{X})^2$
18. Uniformity	$uniformity = \sum_{i=1}^{N_g} p(i)^2$

3.2.3. Second Order Features

3.2.3.1. Gray Level Co-occurrence Matrix (GLCM) Features

Haralick and Shanmugam established the GLCM in the 1970s (B. E. Park et al., 2016), and it is now one of the most widely used statistical methods for obtaining texture features from images data. The GLCM was used to expose the MR image's texture features. In the original image, it occurs as a pair of gray-level frequencies. The grey level co-occurrence matrix (GLCM) (Zwanenburg et al., 2020) is a matrix that describes the distribution of discretized intensities (grey levels) of neighboring pixels or voxels in a 3D volume along one of the image directions.

The dimension of the GLCM is given by the $N_g \times N_g$ form, which denotes a second order joint probabilities of the image region bordered by the mask. The mask is described as $P(i,j|\delta,\theta)$ where $(i,j)^{th}$ components indicate the number of instances when two levels of i and j appear in two adjacent pixels in the image and those pixels are separated by a distance of δ pixels along angle θ. The distance δ measured in terms of the infinity norm. In 3D, when $\delta = 1$, it provides two neighbors for each of the 13 angles (26-connectivity); when $\delta = 2$, it provides 98 connections for 49 unique angles.

PyRadiomics currently uses symmetrical GLCM calculations by default.

Let:

An arbitrarily small positive number is ϵ ($\approx 2.2 \times 10 - 16$)

The co-occurrence matrix for an arbitrary δ and θ is $P(i,j)$

The normalized co-occurrence matrix is $p(i,j) = \frac{P(i,j)}{\sum P(i,j)}$

The number of discrete intensity levels are N_g

The marginal row probabilities are $p_x(i) = \sum_{j=1}^{N_g} P(i,j)$

The marginal column probabilities are $p_y(i) = \sum_{i=1}^{N_g} P(i,j)$

The mean gray level intensity of p_x is μ_x and denoted as $\mu_x = \sum_{i=1}^{N_g} p_x(i)i$

The mean gray level intensity of p_y is μ_y and denoted as $\mu_y = \sum_{j=1}^{N_g} p_y(j)j$

The standard deviation of p_x is σ_x and the standard deviation of p_y is σ_y

$p_{x+y}(k) = \sum_{i=1}^{N_g}\sum_{j=1}^{N_g} p(i,j)$, where $i+j=k$, and $k = 2,3,\ldots,2N_g$

$p_{x-y}(k) = \sum_{i=1}^{N_g}\sum_{j=1}^{N_g} p(i,j)$, where $|i-j|=k$, and $k = 0,1,\ldots,N_g-1$

$HX = -\sum_{i=1}^{N_g} p_x(i)log_2(p_x(i)+\epsilon)$ is the entropy of p_x

$HY = -\sum_{j=1}^{N_g} p_y(i)log_2\big(p_y(j)+\epsilon\big)$ is the entropy of p_y

$HXY = -\sum_{i=1}^{N_g}\sum_{j=1}^{N_g} p(i,j)\,log_2(p(i,j)+\epsilon)$ is the entropy of $p(i,j)$

$HXY1 = -\sum_{i=1}^{N_g}\sum_{j=1}^{N_g} p(i,j)\,log_2\big(p_x(i)p_y(j)+\epsilon\big)$

$HXY2 = -\sum_{i=1}^{N_g}\sum_{j=1}^{N_g} p_x(i)p_y(j)\,log_2\big(p_x(i)p_y(j)+\epsilon\big)$

Table 3.5. Mathematical explanation of the features of the GLCM

Feature	**Description**
1. Autocorrelation:	$autocorrelation = \sum_{i=1}^{N_g}\sum_{j=1}^{N_g} p(i,j)ij$
2. Joint Average (JA):	$JA = \mu_x = \sum_{i=1}^{N_g}\sum_{j=1}^{N_g} p(i,j)i$
3. Cluster Prominence (CP):	$CP = \sum_{i=1}^{N_g}\sum_{j=1}^{N_g}\big(i+j-\mu_x-\mu_y\big)^4 p(i,j)$
4. Cluster Shade (CS):	$CS = \sum_{i=1}^{N_g}\sum_{j=1}^{N_g}\big(i+j-\mu_x-\mu_y\big)^3 p(i,j)$
5. Cluster Tendency (CT):	$CT = \sum_{i=1}^{N_g}\sum_{j=1}^{N_g}\big(i+j-\mu_x-\mu_y\big)^2 p(i,j)$
6. Contrast:	$contrast = \sum_{i=1}^{N_g}\sum_{j=1}^{N_g}(i-j)^2 p(i,j)$
7. Correlation:	$correlation = \frac{\sum_{i=1}^{N_g}\sum_{j=1}^{N_g} p(i,j)ij-\mu_x\mu_y}{\sigma_x(i)\sigma_y(j)}$
8. Difference Average:	$DA = \sum_{k=0}^{N_g-1} kp_{x-y}(k)$
9. Difference Entropy (DE):	$DE = \sum_{k=0}^{N_g-1} p_{x-y}(k)log_2(p_{x-y}(k)+\epsilon)$
10. Difference Variance (DV):	$DV = \sum_{k=0}^{N_g-1}(k-DA)^2 p_{x-y}(k)$
11. Joint Energy (JE):	$JE = \sum_{i=1}^{N_g}\sum_{j=1}^{N_g}\big(p(i,j)\big)^2$
12. Joint Entropy (JEp):	$JEp = \sum_{i=1}^{N_g}\sum_{j=1}^{N_g} p(i,j)log_2(p(i,j)+\epsilon)$
13. Informational Measure of Correlation (IMC1)	$IMC1 = \frac{HXY-HXY1}{max\{HX,HY\}}$ IMCA assesses the correlation between the probability distribution of *i* and *j*, using joint info $I(x,y)$:

	$I(i,j) = \sum_{i=1}^{N_g} \sum_{j=1}^{N_g} p(i,j) log_2 \left(\frac{p(i,j)}{p_x(i) p_y(j)} \right)$ $= \sum_{i=1}^{N_g} \sum_{j=1}^{N_g} p(i,j) \big(log_2(p(i,j) - log_2(p_x(i) p_y(j)) \big)$ $=$ $\sum_{i=1}^{N_g} \sum_{j=1}^{N_g} p(i,j) log_2(p(i,j) - \sum_{i=1}^{N_g} \sum_{j=1}^{N_g} p(i,j) log_2(p$ $= -HXY + HXY1$ The numerator is defined as $HXY - HXY1$, and is therefore ≤ 0. At uniform distribution with complete dependence, joint info will be equal to $log_2(N_g)$. Finally, $HXY - HXY1$ is divided by the maximum of the 2 marginal entropies, where in the latter case of complete dependence. it will result in $IMC1 = -1, as\ HX = HY = I(i,j).$
14. Informational Measure of Correlation (IMC2)	$IMC2 = \sqrt{1 - e^{-2(HXY2 - HXY)}}$ Here, $HXY1 = HXY2$ and that $HXY2 - HXY \geq 0$ represents the joint info of the 2 distributions and the range of $IMC2 = [0,1]$
15. Maximal Correlation Coefficient (MCC):	$MCC = \sqrt{second\ largest\ eigenvalue\ of\ Q}$ $Q(i,j) = \sum_{k=0}^{N_g} \frac{p(i,k) p(j,k)}{p_x(i) p_y(k)}$ The Maximal Correlation Coefficient is a measure of complexity of the texture and 0≤MCC≤1. In a flat region, each GLCM matrix has shape (1,1), resulting in just 1 eigenvalue. Here, an arbitrary value of 1 is returned.
16. Inverse Difference Moment Normalized (IDMN):	$IDMN = \sum_{k=0}^{N_g - 1} \frac{p_{x-y}(k)}{1 + \left(\frac{k^2}{N_g^2} \right)}$
17. Inverse Difference (ID):	$ID = \sum_{k=0}^{N_g - 1} \frac{p_{x-y}(k)}{1+k}$
18. Inverse Difference Normalized (IDN):	$IDN = \sum_{k=0}^{N_g - 1} \frac{p_{x-y}(k)}{1 + \left(\frac{k}{N_g} \right)}$
19. Inverse Variance (IV)	$IV = \sum_{k=0}^{N_g - 1} \frac{p_{x-y}(k)}{k^2}$, k=0 is avoided.
20. Maximum Probability (MP):	$MP = \max\big(p(i,j)\big)$
21. Sum Average (SA):	$SA = \sum_{k=2}^{2N_g} p_{x+y}(k) k$
22. Sum Entropy (SE):	$SE = \sum_{k=2}^{2N_g} p_{x+y}(k) log_2\big(p_{x+y}(k) + \epsilon\big)$
23. Sum Squares (SS):	$SS = \sum_{i=1}^{N_g} \sum_{j=1}^{N_g} (i - \mu_x)^2 p(i,j)$

3.2.3.2. Gray Level Dependence Matrix (GLDM) Features

A Gray Level Dependence Matrix provides a quantitative description of the quantity of gray level dependencies in an image by quantifying the presence of dependency levels. A gray level dependency is calculated as the quantity of related voxels within a specific distance δ that are dependent on the core voxel. If $|i - j| \leq \alpha, j$ is nominated as dependent gray level on central voxel with respect to gray level i. $P(i,j)$ describes as dependence matrix, where the $(i,j)th$ component indicates the voxel number of instances with two levels of i and j voxels in its neighborhood appears through the image.

Let: the total number of discrete levels of intensity are N_g

The total number of discrete dependency sizes are N_d

The total number of dependency zones are $N_z = \sum_{i=1}^{N_g} \sum_{j=1}^{N_d} P(i,j)$

The dependence matrix is $P(i,j)$

The normalized dependence matrix is $p(i,j) = \frac{P(i,j)}{N_z}$

Table 3.6. Mathematical explanation of the features of the GLDM

Feature	Description
1. Small Dependence Emphasis (SDE)	$SDE = \frac{\sum_{i=1}^{N_g}\sum_{j=1}^{N_d}\frac{P(i,j)}{i^2}}{N_z}$
2. Large Dependence Emphasis (LDE)	$LDE = \frac{\sum_{i=1}^{N_g}\sum_{j=1}^{N_d}P(i,j)j^2}{N_z}$
3. Gray Level Non-Uniformity (GLN)	$GLN = \frac{\sum_{i=1}^{N_g}\left(\sum_{j=1}^{N_d}P(i,j)\right)^2}{N_z}$
4. Dependence Non-Uniformity (DN)	$DN = \frac{\sum_{j=1}^{N_d}\left(\sum_{i=1}^{N_g}P(i,j)\right)^2}{N_z}$
5. Dependence Non-Uniformity Normalized (DNN)	$DNN = \frac{\sum_{j=1}^{N_d}\left(\sum_{i=1}^{N_g}P(i,j)\right)^2}{N_z^2}$
6. Gray Level Variance (GLV)	$GLV = \sum_{i=1}^{N_g}\sum_{j=1}^{N_d}P(i,j)(i-\mu)^2$ where, $\mu = \sum_{i=1}^{N_g}\sum_{j=1}^{N_d}ip(i,j)$
7. Dependence Variance (DV)	$DV = \sum_{i=1}^{N_g}\sum_{j=1}^{N_d}P(i,j)(j-\mu)^2$ where, $\mu = \sum_{i=1}^{N_g}\sum_{j=1}^{N_d}jp(i,j)$
8. Dependence Entropy (DE)	$DE = -\sum_{i=1}^{N_g}\sum_{j=1}^{N_d}p(i,j)log_2(p(i,j)+\epsilon)$
9. Low Gray Level Emphasis (LGLE)	$LGLE = \frac{\sum_{i=1}^{N_g}\sum_{j=1}^{N_d}\frac{P(i,j)}{i^2}}{N_z}$
10. High Gray Level Emphasis (HGLE)	$HGLE = \frac{\sum_{i=1}^{N_g}\sum_{j=1}^{N_d}P(i,j)i^2}{N_z}$
11. Small Dependence Low Gray Level Emphasis (SDLGLE)	$SDLGLW = \frac{\sum_{i=1}^{N_g}\sum_{j=1}^{N_d}\frac{P(i,j)}{i^2j^2}}{N_z}$
12. Small Dependence High Gray Level Emphasis (SDHGLE)	The mutual distribution of low dependency with higher gray-level values is measured.
13. Large Dependence High Gray Level Emphasis (LDHGLE)	$LDHGLE = \frac{\sum_{i=1}^{N_g}\sum_{j=1}^{N_d}P(i,j)i^2j^2}{N_z}$
14. Large Dependence Low Gray Level Emphasis (LDLGLE)	$LDLGLE = \frac{\sum_{i=1}^{N_g}\sum_{j=1}^{N_d}\frac{P(i,j)i^2}{i^2}}{N_z}$

3.2.3.3. Neighboring Gray Tone Difference Matrix (NGTDM) Features

A Neighboring Gray Tone Difference Matrix evaluates the variance between a gray value and the average gray value of its immediate neighbors within distance δ. The matrix stores the total of the absolute errors for gray level i. Let X_{gl} is a group

of segmented voxels and $x_{gl}(j_x, j_y, j_z) \in X_{gl}$ is the gray level of a voxel at position (j_x, j_y, j_z), then the average gray level of the neighborhood is:

$$\bar{A}_i = \bar{A}(j_x, j_y, j_z)$$
$$= \frac{1}{W}\sum_{k_x=-\delta}^{\delta}\sum_{k_y=-\delta}^{\delta}\sum_{k_z=-\delta}^{\delta} x_{gl}(j_x + k_x, j_y + k_y, j_z + k_z),$$

where, $k_x, k_y, k_z \neq (0,0,0)$ and $x_{gl}(j_x + k_x, j_y + k_y, j_z + k_z) \in X_{gl}$

Here, W contains the voxel number in X_{gl} neighborhood

Let, n_i is the voxel number with gray level i in X_{gl}

Total number of voxels are $N_{v,p}$ in X_{gl}

which is equal to $\sum n_i$ (at least one neighbor with a valid region)

$N_{v,p} \leq N_p$ where N_p contains the total number of voxels in the ROI.

The gray level probability is denoted as $p_i = \frac{n_i}{N_v}$

gray level i, $s_i = \begin{cases} \sum^{n_i} |i - \bar{A}_i| \; for \; n_i \neq 0 \\ 0 \; for \; n_i = 0 \end{cases}$ the sum of absolute differences

The number of discreet gray levels denoted as N_g

The number of gray levels $N_{g,p}$ where $p_i \neq 0$

Table 3.7. Mathematical explanation of the features of the NGTDM

Feature	Description
Busyness	$Busyness = \frac{\sum_{i=1}^{N_g} p_i s_i}{\sum_{i=1}^{N_g}\sum_{j=1}^{N_g} \lvert i p_i - j p_j \rvert}$ where, $p_i \neq 0, p_j \neq 0$
Coarseness	$Coarseness = \frac{1}{\sum_{i=1}^{N_g} p_i s_i}$
Complexity	$complexity = \frac{1}{N_{v,p}}\sum_{i=1}^{N_g}\sum_{j=1}^{N_g} \lvert i - j \rvert \frac{p_i s_i + p_j s_j}{p_i + p_j}$ where, $p_i \neq 0, p_j \neq 0$
Contrast	$Contrast = \left(\frac{1}{N_{g.p}(N_{g.p}-1)}\sum_{i=1}^{N_g}\sum_{j=1}^{N_g} p_i p_j (i-j)^2\right)\left(\frac{1}{N_{v,p}}\sum_{i=1}^{N_g} s_i\right), \; p_i \neq 0$
Strength	$strength = \frac{\sum_{i=1}^{N_g}\sum_{j=1}^{N_g} (p_i + p_j)(i-j)^2}{\sum_{i=1}^{N_g} s_i}$ where, $p_i \neq 0, p_j \neq 0$

3.2.4. Implementation of PyRadiomics for Feature Extraction

The extraction of the radiomic features refers to the measurement of features as an ultimate evaluation phase, where descriptors of features are used to explain the properties of the grey values inside the ROI/VOI (Zwanenburg et al., 2020). The radiomic features extraction will eventually be done after image processing and segmentation. PyRadiomics (Griethuysen et al., 2017) conducted radiomics feature extraction based on ROIs from the sequence images on each tumour lesion. The PyRadiomics implementation phases presented here to illustrate the properties of breast lesions. The following steps are taken to extract the features:

First step: As an input to PyRadiomics, MR images and segmented mask images (e.g., performed by ImageJ) were loaded. Both input images were converted to the format of NRRD (nearly raw raster data). The initial handling of MR and mask images was handled using SimpleITK, which delivers the popularly used open-source Insight Toolkit (ITK) with a simplified GUI.

Second step: Features on the original image ROI were computed based on a set of several integrated filters such as Gaussian Laplacian, wavelet, logarithm, exponential filters. The SimpleITK library was used mainly for filter execution, and the remaining filters were implemented using NumPy.

Third step: A group of shape, first-order statistics feature, and second-order (GLCM, GLDM, NGTDM) features were derived from input original and segmented mask images (Zwanenburg et al., 2020). Each feature was indicated by a specific name containing the filter, the feature classes, and the element's name. The function extractor module was also developed to describe the extract pipeline and the connection to the other modules in the platform to increase compatibility.

Finally, the determined characteristics were stored in an ordered database and released. This dictionary provides additional information on processing, including library versions, settings, original and masks image information, and filters (Griethuysen et al., 2017). Pandas data frame was used to pre-visualize the output

after the function had been extracted. Then, the created features were saved for analysis in the class-based CSV (Comma-separated values) files.

3.3. Analysis Algorithms

3.3.1. Machine Learning Algorithm

The variety of methods used to analyze radiomic features is huge, ranging from statistical models (J. E. Park et al., 2019; J. E. Park & Kim, 2018) to machine learning techniques. To differentiate benign breast lesions from malignant with radiomics tasks and develop a more effective and reliable prediction model for the identification of tumor class with Support Vector Machine (SVM) kernels utilized.

3.3.1.1. Kernels of Support Vector Machine (SVM)

A Support Vector Machine (SVM)(Chagas et al., 2020) supervised binary classifier selects an optimal hyperplane $f(w,x) = w \cdot x + b$ for optimum margin differentiation of the classes (benign or malignant) in a given dataset with features $x \in \mathbb{R}^m$. Essential ingredients of the training samples called support vectors have optimally calculated this margin. The various kernel functions could be used to transform vectors to a higher-dimensional space wherein the inputs could be segregated linearly, as these groups are non-linearly separable. SVM takes into matter the sign of a function f(x) to categorize an input feature vector, given by:

$$\text{Final decision function, } f(x) = \text{sign}\left(\sum_{i=1}^{v} y_i \alpha_i N(x_i, x_j) + b\right) \qquad (3.3.1.1)$$

where, v = Support vectors with the parameters of a model y_i

b = parameter for bias, $N(x_i, x_j)$ = kernel function

α_i = Laplacian coefficients from the concurrent problem of optimization

Linear Kernel: It is a convenient and widely used mathematical term that the data can be ordered linearly, where single line is used to divide the class. It is just the usual dot product,

$$N(x_i, x_j) = x_i . x_j \quad (3..1)$$

Radial basis function: It is one of the most popular kernels used in SVM. When no previous knowledge of the data exists, it can use.

$$N(x_i, x_j) = exp\left(-\gamma \|x_i - x_j\|^2\right) \quad (3.2)$$

Polynomial Kernel: The polynomial kernel is a type of kernel function that is frequently used in support vector machines (SVM). It computes the training samples in a feature space defined by the initial variables' polynomials, allowing for non-linear model learning.

$$N(x_i, x_j) = \left(x_i^T x_j + 1\right)^d \quad (3.3.1.13)$$

where, d is the degree of the polynomial, e.g., $d = 2$ for quadratic.

Sigmoid kernel: Sigmoid Kernel is another name for the Hyperbolic Tangent Kernel. Due to its origins in neural network theory, this kernel is very common with support vector machines. (Lin & Lin, n.d.).

$$N(x_i, x_j) = tanh\left(\beta\left(x_i^T x_j + a\right)\right) \quad (3.3.1.1.4)$$

The sigmoid kernel has two customizable parameters, β and the intercept constant a.

The k-fold cross-validation method was used to classify breast cancer using the SVM classifier. Initially, all the features' data in the database were randomly selected and stored in 5-fold form.

The four-fold folding data were used during the SVM training process, and the remaining fold was asked to check. This method was repeated more than five times in exchange for completing the SVM classifier training and testing process of all folds.

3.3.2. Deep Learning Techniques

This proposed framework used two different deep learning models for transfer learning, as well as the use of an established (trained) model from the original dataset to train the new dataset. The fully connected dense layer of CNN-SVM and InceptionV3 model was modified and fine-tuned to fit classification requirements. The models were optimized using the Adam Optimizer, and the binary cross-entropy was used to quantify the loss in the InceptionV3 model. Data augmentation was applied, and the model has been assessed with the evaluation of ROC (receiver operating characteristic), AUC (area under curve) curve by calculating the sensitivity, precision, and F-score. In addition, the ROC-AUC values of the test data in both models were also measured to indicate further acceptability of the model's performance.

3.3.2.1. Data Augmentation

The training set is augmented artificially by creating new instances of each training example, a process referred to as data augmentation. It has a variety of applications in computer vision, enabling the most significant generalizations, increasing the number of training instances by classifying in the initial case, and enabling faster classification. Data augmentation results in a reduction in overfitting and an improvement in inaccuracy.

The most widespread method called “On The Fly Data Augmentation” (Rosebrock, 2019) was applied in this work. It is this method that the Image Data Generator class implements. In this case, the model does not train on the initial instance but only on the generated instance, which changes each epoch to allow the network to generalize better.

In this regard, this process applied on training and validation set except on testing set. Each image was resized with a fixed size of 128X128, the image rotates as 10°, height width and zoom ranged at 10%.

3.3.2.2. Balanced Data Generator

It is so difficult to collect a large number of balanced data for deep learning analysis in real-world problems. Data augmentation techniques are mainly used to increase the dataset and improve variance to deal with these circumstances. However, data augmentation does not affect the distribution of labels in the original dataset, which results that an unbalanced position of the dataset is unchanged in all cases, and the algorithm will be biased to predict only the most frequent class.

Many techniques are available to overcome this unbalanced data problem on the evolution of the deep learning model. In this work, the oversampling method was used, which consists of re-sampling less frequent samples to adjust their amount compared to predominant samples.

Here, the oversampling module and Keras-based Image Data Generator module were mixed to balance the dataset, named Balanced Data Generator (Gualberto, 2020). This technique has given enough steps per epoch. The number of samples per class will follow a uniform distribution of batch size.

3.3.2.3. InceptionV3

InceptionV3 is a convolutional neural network that was developed as a module for GoogLeNet. It specializes in image analysis and object detection. It is

the third version of Google's Inception family in CNN architecture, which was first implemented as part of the ImageNet Recognition Challenge.

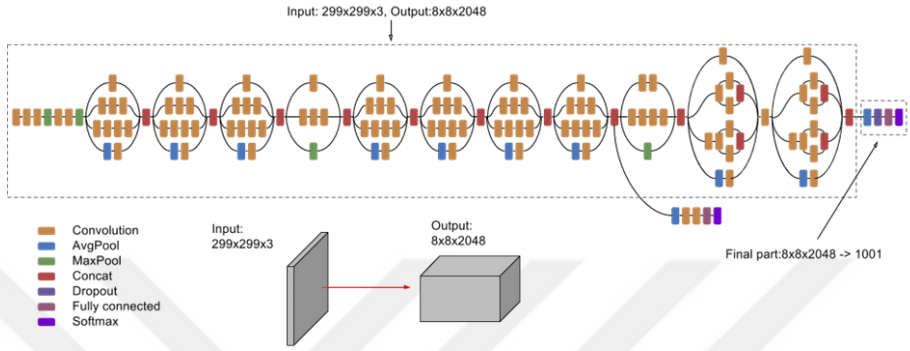

Figure 3.2. A schematic view of the Inception V3 model architecture *(Szegedy et al., 2015)*

The core concept of the Inception architecture is to calculate an ideal local sparse configuration in a convolutional neural network, and to then use readily available dense components to approximate and cover that configuration.

3.3.2.4. CNN-SVM

Convolutional Neural Networks (CNNs) are a form of deep feed-forward artificial neural networks that are frequently used to solve medical image classification problems (Agarap, 2019). Convolutional layers, Maximum pooling layers, drop out or batch normalization, and non-linear activation functions such as tanh, sigmoid, and ReLu are all components of a CNN network.

An activation function ReLU (Figure *3.3*) is used here to incorporate non-linearities into the model evaluation. The model can only learn line mappings in the absence of an activation function.

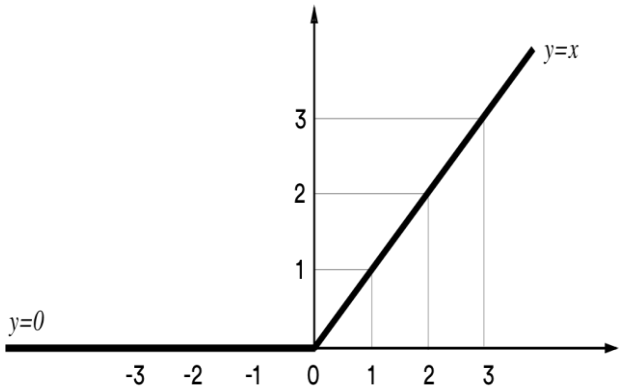

Figure 3.3. Graphical view of Rectified Linear Unit (ReLU) function *(Liu, 2017)*

Compared to the other two functions, ReLU is activated by simply thresholding matrix values to zero (Equation 3.3.2.4.6) and significantly accelerates stochastic gradient descent convergence.

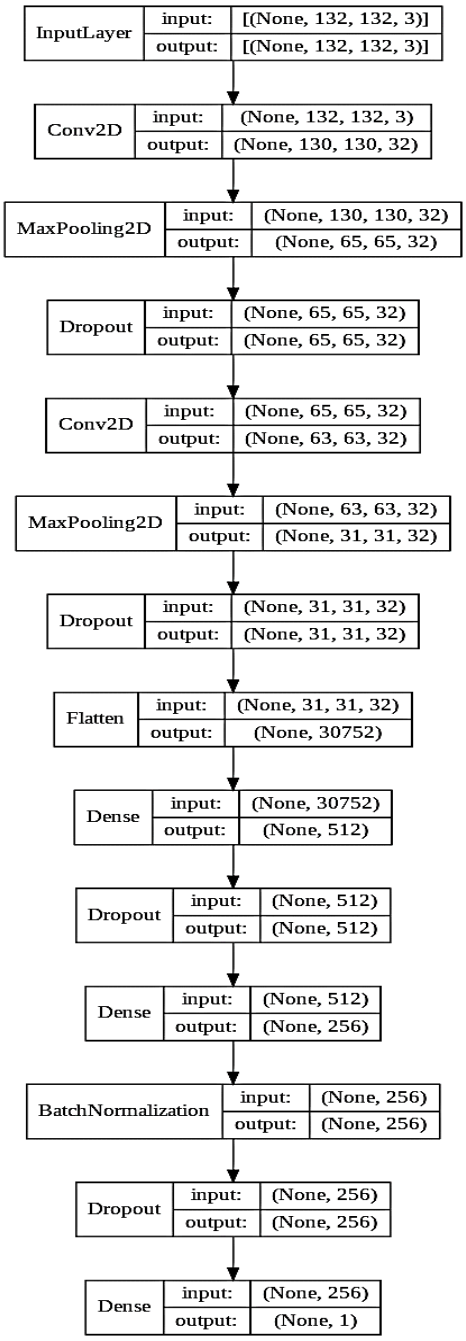

Figure 3.4. A detail view of modified CNN-SVM model architecture

$$f\big(h_\theta(x)\big) = h_\theta(x)^+ = \max\big(0, h_\theta(x)\big) \quad (3.3.2.45)$$

In the final dense layer of the CNN, rather than using a SoftMax or sigmoid function with loss function (binary or categorical), the regularizer L2-SVM is used. That is, the performance shall be interpreted as follows: $y \in \{-1, +1\}$, and the loss is calculated in the following Eq. 8. Adam is then used to determine the weight parameters.

The regularizer L2-SVM learns value of w by resolving an optimization problem (Eq. 3.3.2.46).

$$\min \frac{1}{p} w^T w + C \sum_{i=1}^{P} \max\big(0, 1 - \acute{y}_i (w^T x_i + b)\big) \quad (3.3.2.46)$$

where $w^T w$ is the L1 norm (also identified as Manhattan norm), C is the penalty parameter, y'_i is the actual label, and $w^T x + b$ is the predictor function.

Equation 3.3.2.47 is identified as L1-SVM, with the traditional hinge loss. It is distinctive equivalent, L2-SVM (Eq. 3.3.2.48), delivers additional durable outcomes.

$$\min \frac{1}{p} \|w\|_2^2 + C \sum_{i=1}^{P} \max\big(0, 1 - \acute{y}_i (w^T x_i + b)\big)^2 \quad (3.3.2.47)$$

where $\|w\|_2$ is the L2 norm (additionally referred to as Euclidean norm), with the squared hinge loss.

3.4. Performance Assessments

This is needed to find out the model's accuracy after the training phase by analysis algorithms. So, the test set was employed to evaluate the model. The test set

includes samples that have never been seen before by the algorithm. If the model performs well in predicting, it can be assumed that the model is generalizing well.

Confusion matrix: The efficient classification model is capable of summarizing the model's effectiveness at classifying examples into distinct categories. The expected label is identical to the real label on one side of the confusion matrix. On the other hand, the real label matches the predicted label.

TP: True Positive: Benign correctly identified as Benign
TN: True Negative: Malignant correctly identified as Malignant
FP: False Positive: Benign incorrectly identified as malignant
FN: False Negative: Malignant incorrectly identified as Benign

Accuracy: The accuracy based on the confusion matrix is proportional to positive occurrences accurately and positive occurrences inaccurately and is inversely comparable to negative occurrences accurately and negative occurrences inaccurately rates.

$$\textbf{Accuracy} = \frac{TP+TN}{(TP+TN+FP+FN)}$$

Precision: Precision is the ratio of positive occurrences accurately predicted, divided by sum of positive occurrences and negative occurrences predicted accurately.

$$\textbf{Precision} = \frac{TP}{(TP+FP)}$$

Recall or Sensitivity or True Positive Rate (TPR): Recall is the ratio of positive occurrences accurately predicted, divided by positive and negative

occurrences accurately and incorrectly predicted. The low false negative occurrence rate is known to be a significant recall.

$$\textbf{Recall} = \frac{\text{TP}}{(\text{TP+FN})}$$

F-Score: F1-score is the weighted average value of precision and recall. It is usually more beneficial than accuracy, mostly when the dataset classes is unbalanced.

$$\textbf{F-score} = \frac{2\times\text{Recall}\times\text{Precision}}{(\text{Recall} + \text{Precision}}$$

ROC (Receiver operating characteristic): An ROC curve is a graphical plot used to demonstrate the ability of a binary class classification system to differentiate between two categories as its distinction threshold is varied. The ROC curve illustrates the trade-off between sensitivity (TPR) and specificity (1 – FPR). Classifiers that mean better results result in curves that are closer to the top-left corner. The closer the curve approaches the 45-degree diagonal of the ROC space, the less accurate the measurement.

4. EXPERIMENTAL RESULTS

The main task of this thesis is to use radiomics and deep learning methods to classify breast tumor lesions by handling magnetic resonance imaging (MRI) image data. Radiomics is an emerging field in healthcare, especially in tumor analysis. The main objectives of this works were to check out the classification robustness by using two popular and reliable methods radiomics and deep learning. Same performance analysis matrices are developed to utilize these models.

4.1. Radiomics Feature Analysis

Investigations have been done on the original MR image dataset, which had been obtained at the Department of Medical Oncology and Department of Radiology, Cukurova University Hospital in Turkey.

At the evaluation phase, 70% of the data is used to train the model. In the beginning, second-order features with different matrices, namely co-occurrence matrix, dependence matrix, and difference matrix, were given to SVM-kernel based model, and classification accuracies were calculated.

Later, shape with second order and shape with first-order features have been analyzed. Finally, the shape, first-order, and second-order feature combined dataset has been analyzed with SVM kernels. The 5-fold mean accuracy result of cross-validation is shown in Table 4.1. In every combination of data split, the SVM-linear kernel in the cross-validation model evaluation stage has shown the highest accuracy.

In 5-fold cross-validation, the original sample is randomly partitioned into five equal-sized groups. One group was removed from the five groups as a hold-out set, and the remaining groups used the training data. The SVM model then fit on the training data and evaluated on the hold-out set. This procedure is five times so that all groups have served exactly one as the hold-out set.

Table 4.1. Cross-validation mean accuracy of feature data split with SVM-kernels

Kernel	**GLCM + GLDM + NGTDM (%)**	**Shape + GLCM + GLDM + NGTDM (%)**	**First order + GLCM + GLDM + NGTDM (%)**	**Shape+ First order + GLCM + GLDM + NGTDM (%)**
Linear	96.17	96.17	96.08	96.08
RBF	94.92	94.92	94.92	93.67
Polynomial	91.00	91.00	91.00	90.92
Sigmoid	89.75	89.75	89.83	85.83

Two combinations of features data split have shown the highest 5-fold mean accuracy of 96.17% on SVM-linear kernel analysis. All the feature data split performed similar mean accuracy at SVM-RBF kernel-based model development except a combination of all features split. The polynomial kernel showed 91% mean accuracy with all splits. The worst result found in SVM-Sigmoid based analysis.

The SVM-kernel approach's overall findings are summarized in *Figure 4.1*, *Figure 4.2*, *Figure 4.3* and *Figure 4.4*, which illustrated the proposed approach's performance against previously described matrices.

Figure 4.1 showed that, combine second order features (GLCM, GLDM, NGTDM) split and shape with second order data split performed the same accuracy, which means that the shape features did not have any impacts for getting 88.24% accuracy. When the first-order features combined with second order features, the model performance increased. Meanwhile, the highest accuracy was found on all feature combination splits and provided 94.12% test accuracy on the SVM-linear kernel.

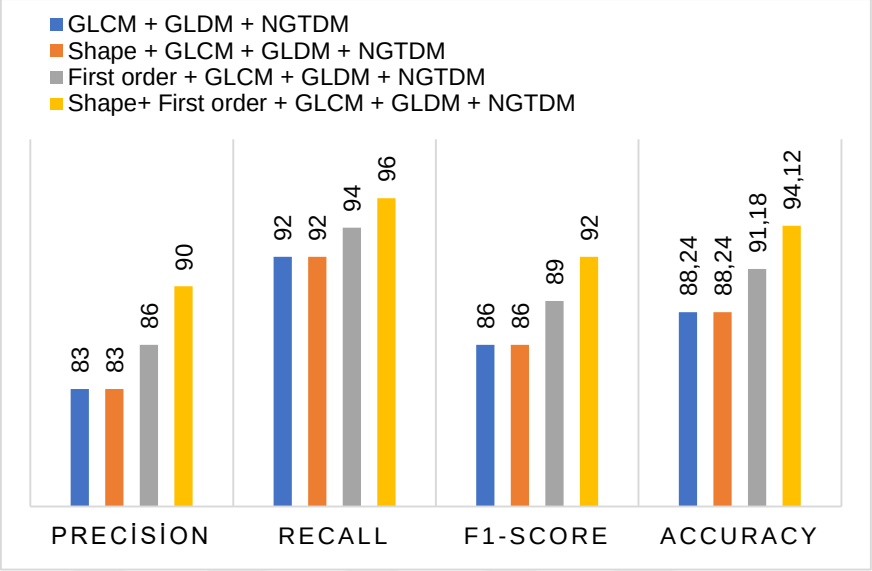

Figure 4.1. SVM-Linear Kernels Confusion Matrix Output

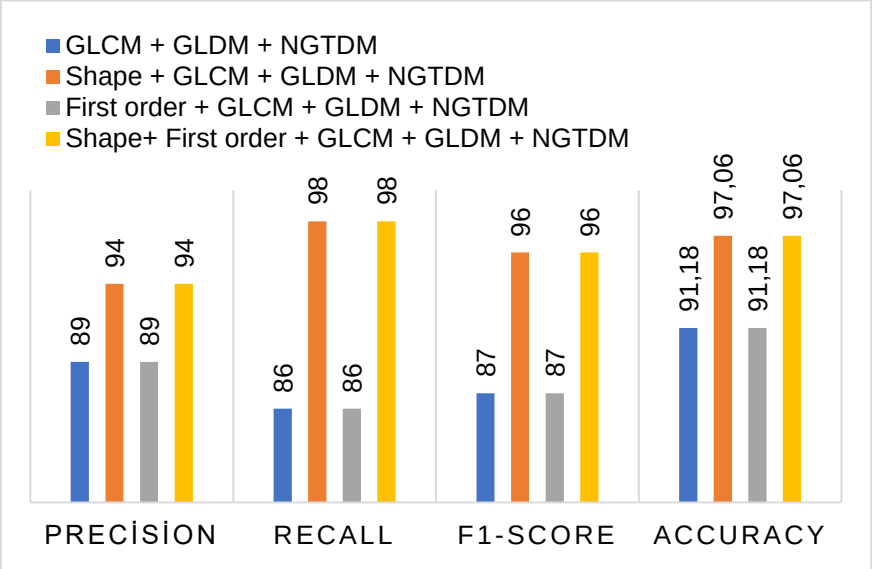

Figure 4.2. SVM-RBF Kernels Confusion Matrix Output

In SVM-RBF kernel analysis, second order features combination and first with second order features combination achieved 91.18% accuracy. That means the

combination of first-order features did not enough influence for getting this output. Another features data split shape, second order, and all features combination provided highest 97.06% accuracy. This analysis aims to understand that the combination of shape and second order feature data split is enough to get the maximum performance of the SVM-RBF model. The all-features combination analysis on the SVM-RBF kernel is not efficient.

The lowest performance was found in the combined data group of the first order and 2nd Order features. However, if we look from the beginning, it can be seen that the highest accuracy of 91.18% was found in Group-1. Since then, the accuracy has not increased even after adding the shape feature separately to the 2nd order feature. The same accuracy has been found by adding all the features together. This means that combining the second order feature with the shape or the 1st order feature does not affect its model performance analysis. On the other hand, the performance of polynomial kernels (Figure 1.3) is not better than the two types of kernels mentioned earlier.

At this stage, the SVM-sigmoid kernel (Figure 1.4) analysis shows that the first three feature combinations exhibit the same accuracy. That is, maximum accuracy is found in second order feature analysis. Later, by adding shape and first-order feature with the 2nd order feature and analyzing it separately, the model performance did not increase. On the contrary, in the combined analysis of all the features, the accuracy has decreased compared to the past. This means that the maximum accuracy is only available in the 2nd order feature, so other features do not play any role in this kernel analysis. At the same time, it is found with certainty that, compared to the polynomial kernel with the other three kernels, it is showing the lowest accuracy.

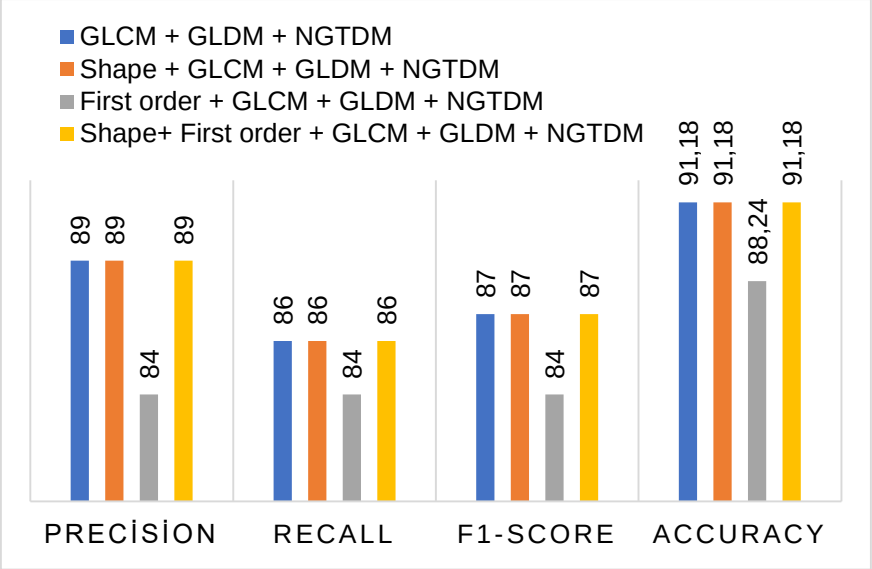

Figure 4.3. SVM-Polynomial Kernels Confusion Matrix Output

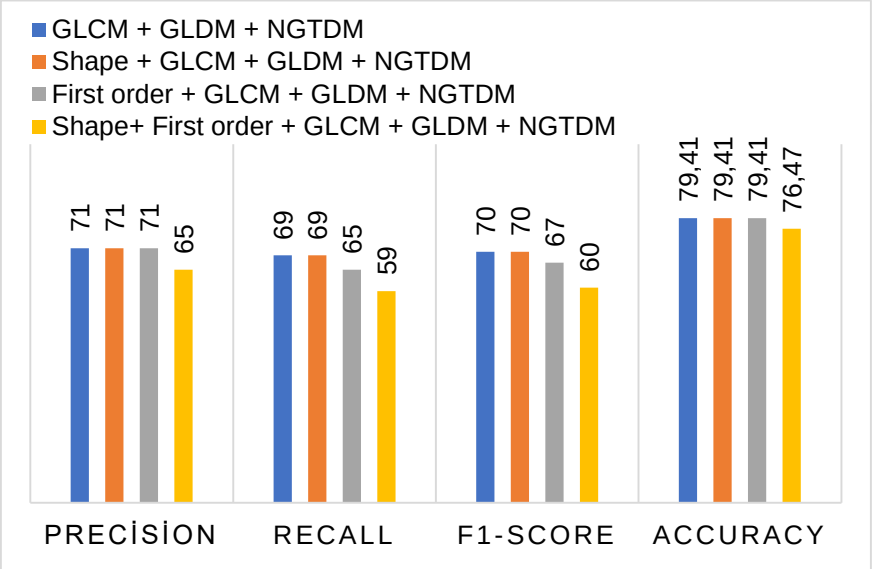

Figure 4.4. SVM-Sigmoid Kernels Confusion Matrix Output

To conclude the evaluation metrics analysis, in kernel-based feature analysis, the SVM-RBF kernel has achieved the highest accuracy, 97.06%, in

second-order features and the combined analysis of all features. However, since the highest results were obtained in a single analysis of the 2nd order features, the shape, first-order feature is of no importance in obtaining maximum results. That is, shape, first-order features can be ignored in SVM-RBF kernel analysis. Finally, It can say that the second order feature combinational analysis with the SVM-RBF model is best for radiomics to feature generated breast tumor classification and the obtained result also ensured that this data split is perfectly separable radially.

4.1.1. Parameter Optimization for Selecting best value of SVM-kernels

Parameter optimization helps to select the best model configuration to ensure the highest accuracy. Choosing the best parameter values in SVM-kernels, a well-known method Grid Search CV was used which is somewhat tricky due to some ambiguous results from the model. Three main variables of kernel function were tested on our SVM-kernels framework. These variables are 'C', 'gamma' and 'degree'. In a default value of 'C' and 'gamma' is 1 where C is denoted as regularization parameter. The parameter 'C' also compromising the classification accuracy for scalability. This parameter corrects the over-reliance on simplicity. Larger values of the parameter of 'degrees' appear to overfit the results.

Table 4.2. Parameters range of each SVM kernel to be evaluated

Kernels	**Parameter value**
Linear	*'**C**': [0.0001, 0.001, 0.01, 1.0, 10.0, 100.0, 1000.0]*
RBF	*'**C**': [0.0001, 0.001, 0.01, 1.0, 10.0, 100.0, 1000.0], '**gamma**': [0.0001, 0.001, 0.1, 1, 2]*
Polynomial	*'**C**': [0.0001, 0.001, 0.01, 1.0, 10.0, 100.0, 1000.0], '**gamma**': [0.0001, 0.001, 0.1, 1, 2], **'degree'**: [1, 2, 3, 4, 5]*
Sigmoid	*'**C**': [0.0001, 0.001, 0.01, 1.0, 10.0, 100.0, 1000.0], '**gamma**': [0.0001, 0.001, 0.1, 1, 2]*

The identical value of cross-validation K-Fold was applied to evaluate the SVM. Table *4.2* represents the set of parameters with four SVM kernels.

Table 4.3. Every SVM kernel best results on binary classification

Data	Kernel	Parameter	Mean Accuracy
GLCM + GLDM + NGTDM	Linear	*'C': 1.0*	96.17**%**
	RBF	*'C': 100, 'gamma': 0.001*	**97.42%**
	Polynomial	*'C': 1, 'degree': 1, 'gamma': 1*	96.17**%**
	Sigmoid	*'C': 1000, 'gamma': 0.001*	96.17**%**
Shape + GLCM + GLDM + NGTDM	Linear	*'C': 1.0*	96.17**%**
	RBF	*'C': 100, 'gamma': 0.001*	**97.42%**
	Polynomial	*'C': 1, 'degree': 1, 'gamma': 0.1*	**97.42%**
	Sigmoid	*'C': 100, 'gamma': 0.001*	**97.42%**
First order + GLCM + GLDM + NGTDM	Linear	*'C': 1.0*	96.08**%**
	RBF	*'C': 100, 'gamma': 0.001*	**97.42%**
	Polynomial	*'C': 1, 'degree': 1, 'gamma': 2*	96.08**%**
	Sigmoid	*'C': 100, 'gamma': 0.001*	**97.42%**
Shape+ First order + GLCM + GLDM + NGTDM	Linear	*'C': 1.0*	96.08**%**
	RBF	*'C': 100, 'gamma': 0.001*	**97.42%**
	Polynomial	*'C': 1, 'degree': 1, 'gamma': 0.1*	96.08**%**
	Sigmoid	*'C': 1000, 'gamma': 0.001*	96.08**%**

The bold values show the higher results that were obtained in the analysis, for SVM-kernels.

Table *4.3* showed the best configurations of three parameters of proposed four kernels at each data group. The configuration was formed using accuracy as the metric for optimization. Notably, the RBF kernel achieves the best results, implies that the feature space is radially separable. Here, all kernels had similar scores, proving the robustness of the radiomics feature depends on classification analysis with the SVM model.

4.2. Deep Learning based Analysis

Two deep learning techniques were chosen for comparison, according to their performances. A total of 2998 individual images were used for analyzing the deep learning models. For training purposes, 2098 were selected where validation and testing images contained at 540 and 360, respectively. All images were resized

as 128×128, and benign lesions contained images defined as 0 class, and malignant lesions contained images defined as 1 class.

The proposed models also outperformed compared to other researchers' analysis with much fewer parameters, as summarized in **Table *4.4***

As the primary outcome, CNN-SVM obtained perfect accuracy with almost two times fewer parameters than the other pre-trained InceptionV3 deep learning architectures.

Table 4.4. Number of parameters of the evaluated models

Models	**Total Parameters**
InceptionV3	21,810,977
CNN-SVM	14,888,865

4.2.1. InceptionV3 Results

InceptionV3 is the third generation of the Inception family of convolutional neural network architectures. It is being fine-tuned so that it can distinguish between two classes rather than 1000. Additionally, the model is composed of symmetric and asymmetric building blocks, such as convolutions, average pooling, max pooling, concatenation, dropouts, and fully connected dense layers. The Batch normalization process is widely used in the model and is applied to activation inputs. Since the classification is binary class, the loss was computed using the Sigmoid function.

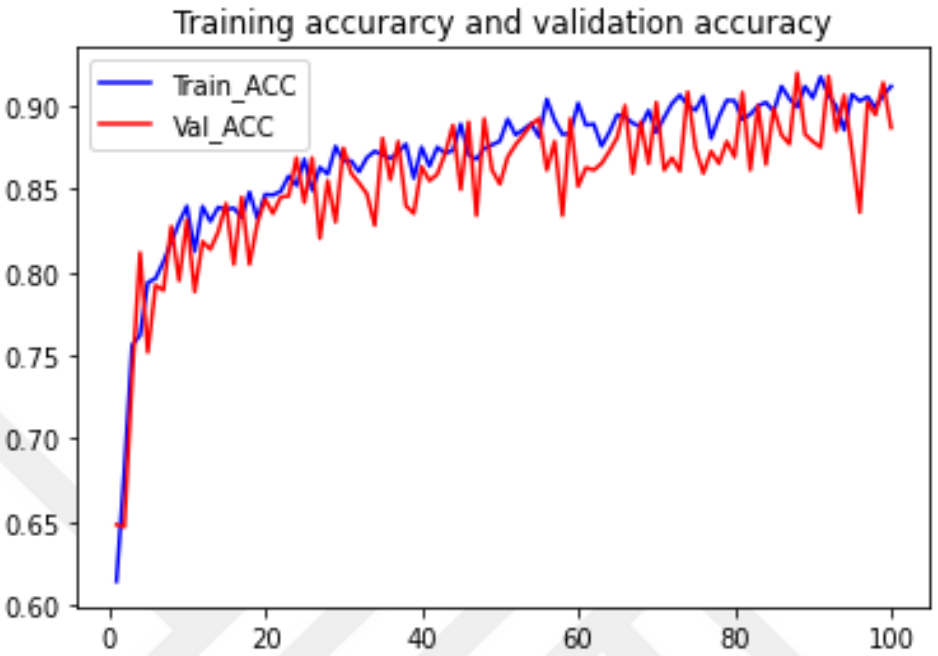

Figure 4.5. Accuracy of Training and Validation Phase of InceptionV3 model

The accuracies of the InceptionV3 model's training and validation phases are depicted in Figure *4.5*. Both accuracy levels started at 60% during the early stages of the epoch. At the end of the 20-epoch count, the model achieved nearly 85% accuracy. After 60 epochs, the machine is capable of achieving greater than 90% accuracy. The model reached the highest degree of accuracy, which is close to 90%, throughout approximately 70-100 epochs.

Volatility is moderate but has fluctuated around 86±5%. As shown in Figure *4.6*, as the validation loss decreased, the training loss decreased as well. With the training loss reaching zero, it is possible to reduce the validation loss significantly.

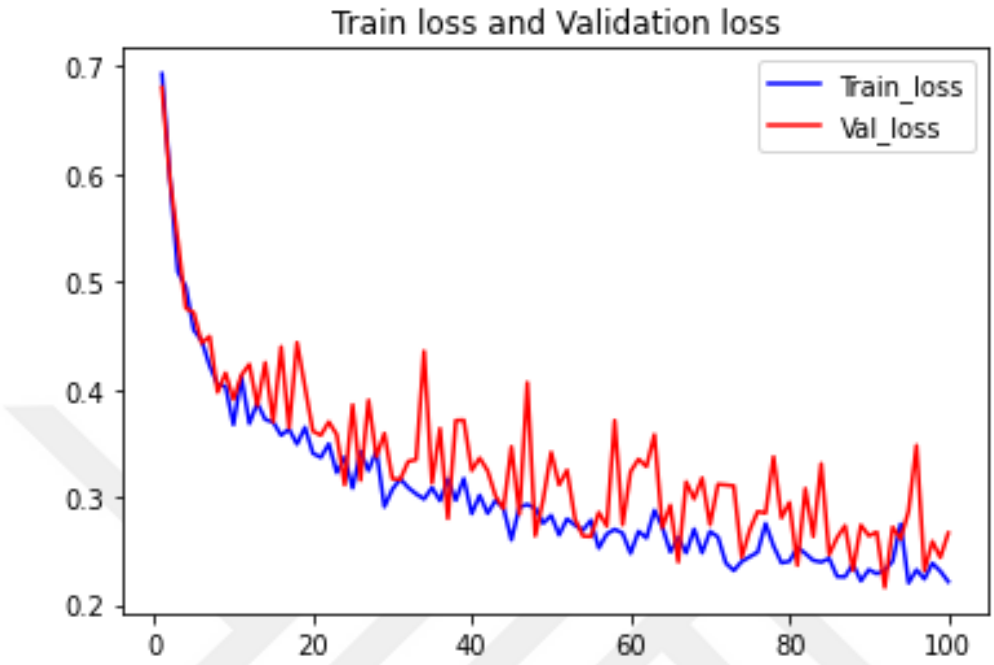

Figure 4.6. Loss Plot of Training and validation Phase of InceptionV3 model

Table *4.5* also contains information about the validation's accuracy. The testing accuracy was 95.28%, according to the assessment metrics review. The precision values were 99.05%, indicating a high level of classification accuracy for the benign class. The InceptionV3 had a high F1 score, which established a more accurate classification on this image dataset.

Table 4.5. Evaluation metrics report on InceptionV3 model

Technique	**Validation Accuracy**	**Overall Test Accuracy**	**Precision**	**Recall**	**F1 Score**
InceptionV3	91.96%	95.28%	99.05%	95.73%	97.36%

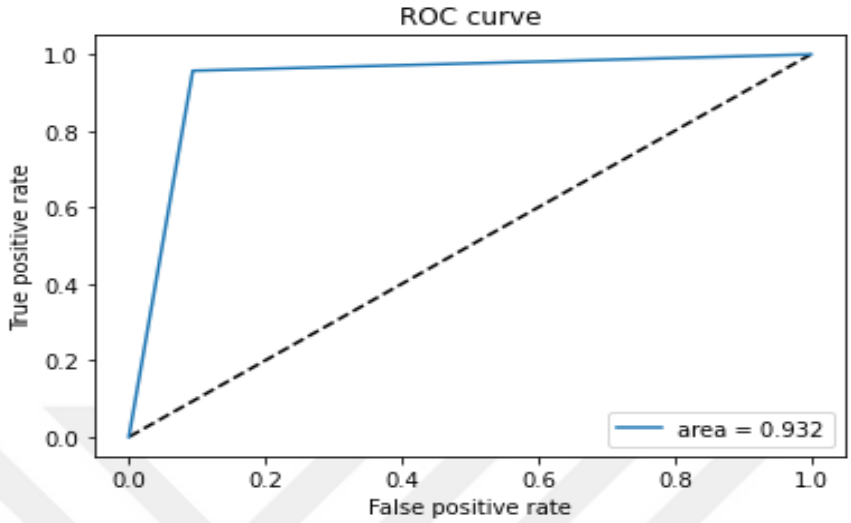

Figure 4.7. ROC-AUC Plot of InceptionV3 with test data

The AUC area of the InceptionV3 model was 0.932, indicating that the model performed well in distinguishing tumor class from the test dataset.

4.2.2. CNN-SVM Results

The CNN-SVM model designed with two convolutional layers, two Maximum Pooling layers, one Flatten layer, three Dense layers, and dropout in all layers. In a fully connected dense layer, batch normalization is also implemented with L2-SVM at the end. The entire architecture of the CNN - SVM model is shown in Figure *3.4*.

The accuracy of training and validation phase and loss curve are shown in Figure *4.8* and Figure *4.9*, respectively. Figure *4.10* represents the ROC-AUC score according to the testing data. According to Figure *4.8*, the training and validation phase began at the accuracy of 50% on first initial epochs. After 20 epochs, the accuracies increased by an average of 80%. Between 70 and 100 epochs, the accuracies exhibited almost constant values.

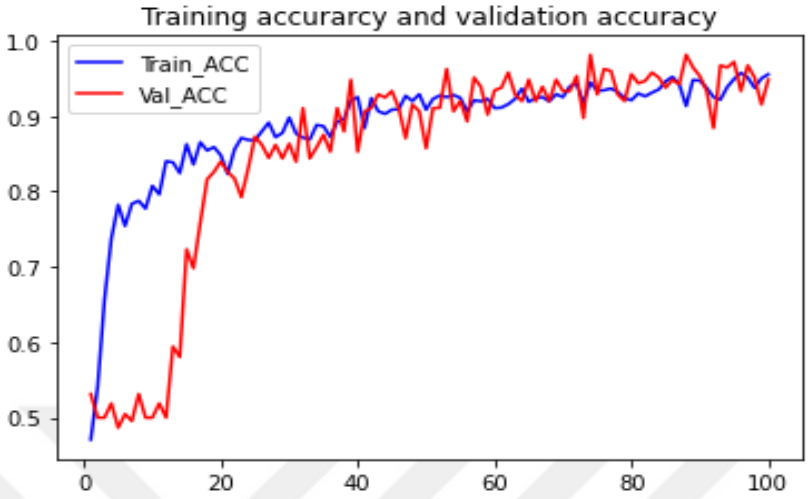

Figure 4.8. Accuracy of Training and Validation Phase of CNN-SVM model

Though training and validation accuracies vary slightly, overall accuracies remain between 93-98%. Simultaneously, the training and validation losses decreased represents at Figure *4.9*, and the losses in the last 30 epochs were found to be near 0.1%.

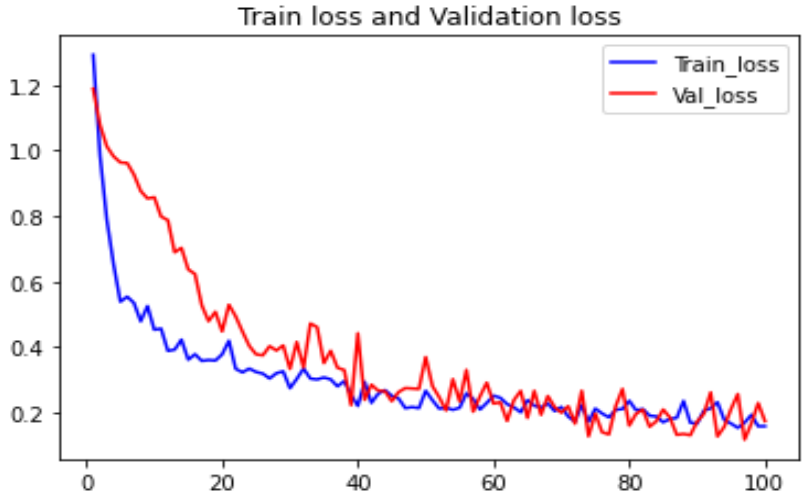

Figure 4.9. Loss Plot of Training and Validation Phase of CNN-SVM model

Table 4.6. Evaluation metrics results on CNN-SVM model

Model	Validation Accuracy	Overall Test Accuracy	Precision	Recall	F1 Score
CNN-SVM	98.83%	95.28%	100%	94.82%	97.34%

As a result of this research, it was discovered that test accuracy was 95.28%, which is equivalent to the results found in the InceptionV3 model. The CNN-SVM model had 100% precision, which was an excellent classifier for the positive class, reflecting the benign class in this task. Additionally, the CNN-SVM model obtained a high F1 score, suggesting that the model correctly detected the tumor lesion even though the dataset was unbalanced.

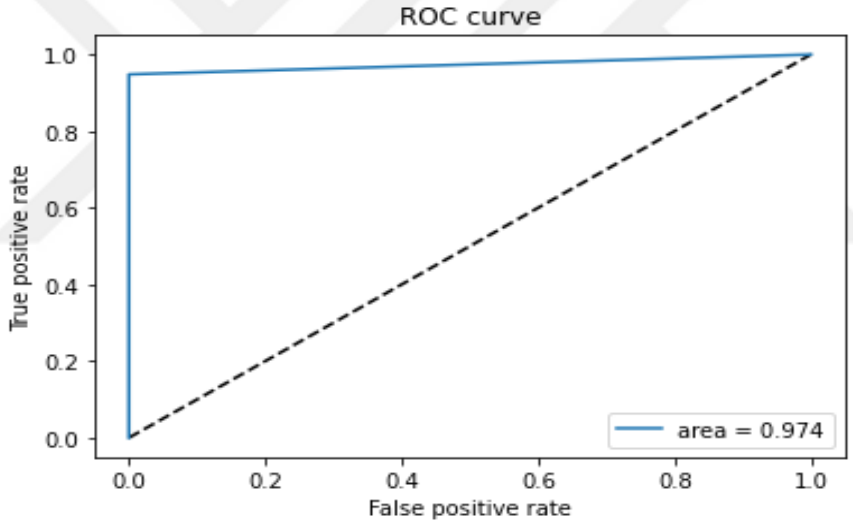

Figure 4.10. ROC-AUC Plot of CNN-SVM with test data

Finally, the CNN-SVM model achieved the highest AUC value of 0.974, showed in Figure *4.10*, indicating the model's highest degree of efficiency in reliably distinguishing benign and malignant classes.

4.3. Analysis of deep learning models

A comparison between two deep learning techniques has been utilized to classify breast cancer with the analysis of evaluation metrics. **Table *4.7*** and **Table *4.8*** shows the results of deep learning methods. Though the dataset is unbalanced, random up sampling with data augmentation techniques were applied to overcome the imbalance problem. In the testing stage, the original images were tested; those images are not used before in training and validation steps.

Table 4.7. Confusion metrics parameters of deep learning models

Models	True Positive	False Positive	False Negative	True Negative
InceptionV3	29	3	14	314
CNN-SVM	32	0	17	311

A total of 100 epochs were monitored while developing the model. By using callbacks in Keras, the validation accuracy was controlled. It is a monitoring system to control the model development process by looking at accuracy and loss. In this case, after 70 epochs, almost stable results were found. The batch size was 32 in the data augmentation and model developing phase because a big batch size would lead to poor generalization and lower test accuracy. It is found that the CNN-SVM model showed the highest validation accuracy with 98.83%, where InceptionV3 showed 91.96%. Area Under Curve (AUC) found 97.4% for the CNN-SVM model, representing the best efficient model.

Table 4.8. Performance results of the deep learning models

Models	Validation Accuracy	Test Accuracy	Precision	Recall	F1 Score	AUC
InceptionV3	91.96%	95.28%	99.05%	95.73%	97.36%	0.932
CNN-SVM	98.83%	95.28%	100%	94.82%	97.34%	0.974

5. DISCUSSIONS

Many computational methods have been described in the context of breast lesion classification, each of which incorporated a diverse set of features (textural features, morphological features, etc.). In this work, shape, first, and second order contained GLCM, GLDM, NGTDM features, and their combination was used in breast tumor classification. High accuracy was achieved with all metrics to classify the benign and malignant classes of breast cancer with a SVM algorithm.

The proposed radiomics feature analyzed solution was compared to five well-known state-of-the-art methods, the outcomes of which are presented in **Table *5.1***. As shown in Table *5.1*, the proposed method achieves a higher rate of classification accuracy than other strategies. Additionally, Table *5.1* demonstrates that the accuracy of the methods described in (Spanhol et al., 2016; Sayed & Hassanien, 2017; Al-Salihy & Ibrikci, 2017; Carvalho et al., 2020; Khan et al., 2021) is 80–85%, 92.99%, 88-97%, 95%, 96.3% respectively, while the proposed method obtained 97.06% accuracy, indicating the proposed method's prominence in comparison to other techniques.

Table 5.1. A comparison with other feature-based machine learning analysis

Authors	**Attributes**	**Algorithms**	**Accuracy**
(Spanhol et al., 2016)	Texture features	K-NN, QDA, SVM	80-85%
(Sayed & Hassanien, 2017)	NS, MFO	Decision tree -CART	92.99%
(Al-Salihy & Ibrikci, 2017)	Various	Decision Tree	88-97%
(Carvalho et al., 2020)	Texture features	MLP, XGBoost, SVM	95%
(S. U. Khan et al., 2021)	GLCM	SVM	96.3%
Proposed method	**Shape, GLCM, GLDM, NGTDM**	**Four SVM-Kernel**	**97.06%**

Based on achieved results from radiomics feature analysis, it can be concluded with confidence that this new approach of machine learning-based combinational analysis of radiomics features from magnetic resonance images can effectively differentiate the breast tumors.

Finally, as seen in **Table 5.2**, the proposed deep learning strategies were contrasted to applicable state-of-the-art frameworks to demonstrate their efficacy. The findings indicate that the proposed techniques outperform other deep learning techniques.

When compared to the performance of various deep convolutional neural network architectures, it was clear that the proposed techniques achieved excellent results. (Hepsağ et al., 2017 & Tan et al., 2017) gained 68% and 85.85% accuracy, respectively, while using a public dataset with CNN architectures.

Table 5.2. Comparison with other deep learning techniques

Authors	**Dataset**	**Algorithms**	**Accuracy**	**AUC**
(Hepsağ et al., 2017)	MIAS	DCNN	68.00%	-
(Tan et al., 2017)	MIAS	CNN	85.85%	-
(Natalia Antropova et al., 2017)	Private	AlexNet		0.91
(Zheng et al., 2018)		CNN, LSTM	84.7%	-
(Jiang et al., 2017)	BCDR-F03	GoogLeNet, AlexNet		0.88
(Hadad et al., 2017)	Private	VGG, Transfer Learning	93%	
(Truhn et al., 2018)	Private	ResNet		0.88
(Ragab et al., 2019)	CBIS-DDSM	DCNN-AlexNet, SVM	87.20%	0.94
(S. Khan et al., 2019)	Microscopic Images	VGG, GoogLeNet, ResNet	97.67%	-
(Song et al., 2020)	DDSM	GoogLeNet, InceptionV2, XGBoost	92.80%	-
Proposed Work	**Private dataset**	**CNN-SVM InceptionV3**	**95.28%**	**0.974 0.932**

An AUC of 91% and 88% achieved by (Natalia Antropova et al., 2017, Jiang et al., 2017 & Truhn et al., 2018), respectively, using the AlexNet deep learning

architecture where (Truhn et al., 2018) used a ResNet pre-trained model. However, they validated their methods using a variety of different datasets. (Hadad et al., 2017 & Zheng et al., 2018) classified the tumor using the VGG and CNN models, respectively, and achieved an acceptable level of accuracy.

Additionally, in 2020, (Song et al., 2020) extracted and categorized in-depth features from GoogLeNet and InceptionV2 using the XGBoost classifier, achieving a 92.80% accuracy. Similarly, there was a minor discrepancy in the accuracy obtained by this method and the work performed by (S. Khan et al., 2019). They incorporated VGG, GoogLeNet, and ResNet DCNN's in-depth features from microscopic samples.

According to the above analysis, the results obtained in the deep learning analysis showed that deep learning models had given high performance in this task with all kinds of evaluation metrics and ROC-AUC values. On the other hand, **Table 5.2** mentioned studied in the private dataset where researchers used a few images to develop their models. Several others chose ROC-AUC instead and obtained low performance.

By analyzing radiomics feature-based and feature-less (deep learning) results, it can be expressed that the proposed combinational analysis of radiomics features is excellent for breast tumor classification and can be efficiently executed for tumor diagnosis.

6. CONCLUSIONS

Breast cancer is the second most common form of cancer in females after lung cancer. Early detection has significantly increased the survival rate of breast cancer. Thus, automatic detection can be used to provide the required solutions in this area. The explosive growth of radiomics and deep learning-based medical image analysis makes it possible to extract the features and take full advantage of big data for early detection and enrich cancer diagnostics.

This work provided an overview of radiomics feature extraction and their machine learning-based analysis without feature extracting based deep learning analysis for breast tumor classification. It is structured around the dataset, model design, and evaluation of the models.

According to the previous researcher's study, radiomics features with distinct metrics were studied separately, with very few studies focusing on their combinational analysis. Apart from the deep learning-based review, another significant problem is the lack of an image dataset.

In this study, a combined application of various radiomics features with statistical features was investigated. Support Vector Machines with four different kernels demonstrated exceptional results. This analysis suggested an accurate and feasible method for classifying benign and malignant tumors using the proposed classification scheme based on radiomics features. It performs astonishingly well when several features were combined into a classification scheme. The application of the algorithm described in this thesis can increase tumor classification accuracy and decrease misdiagnoses.

To make this analysis more credible, appropriate, and accurate, we will collect additional images and compile a large image dataset. We will investigate how to build these models in future work so that clinical data and derived feature data can be combined. Additionally, we will assist in determining the tumor type, which is critical for initiating early diagnosis. We hope that the collaboration with different

researchers such as radiologists, biomedical, and informatics researchers to advance the fields of radiomics that are both depends on feature-based and feature-less (deep learning).

REFERENCES

Abramoff, M. D., Magalhães, P. J., & Ram, S. J. (2004). Image processing with ImageJ. Biophotonics International, 11(7), 36–42.

Aerts, H. J. W. L., Velazquez, E. R., Leijenaar, R. T. H., Parmar, C., Grossmann, P., Carvalho, S., Cavalho, S., Bussink, J., Monshouwer, R., Haibe-Kains, B., Rietveld, D., Hoebers, F., Rietbergen, M. M., Leemans, C. R., Dekker, A., Quackenbush, J., Gillies, R. J., & Lambin, P. (2014). Decoding tumour phenotype by noninvasive imaging using a quantitative radiomics approach. Nature Communications, 5, 4006. https://doi.org/10.1038/ncomms5006

Agarap, A. F. (2019). An Architecture Combining Convolutional Neural Network (CNN) and Support Vector Machine (SVM) for Image Classification. ArXiv:1712.03541 [Cs, Stat]. http://arxiv.org/abs/1712.03541

Agarwal, R., Díaz, O., Yap, M. H., Lladó, X., & Martí, R. (2020). Deep learning for mass detection in Full Field Digital Mammograms. Computers in Biology and Medicine, 121, 103774. https://doi.org/10.1016/j.compbiomed.2020.103774

Al-Salihy, N. Kh., & Ibrikci, T. (2017). Classifying breast cancer by using decision tree algorithms. Proceedings of the 6th International Conference on Software and Computer Applications, 144–148. https://doi.org/10.1145/3056662.3056716

Antropova, Natalia, Huynh, B. Q., & Giger, M. L. (2017). A deep feature fusion methodology for breast cancer diagnosis demonstrated on three imaging modality datasets. Medical Physics, 44(10), 5162–5171. https://doi.org/10.1002/mp.12453

Antropova, Natasha, Huynh, B., & Giger, M. (2018). Recurrent neural networks for breast lesion classification based on DCE-MRIs. Medical Imaging 2018: Computer-Aided Diagnosis, 10575, 105752M. https://doi.org/10.1117/12.2293265

Ardakani, A. A., Gharbali, A., Saniei, Y., Mosarrezaii, A., & Nazarbaghi, S. (2015). Application of Texture Analysis in Diagnosis of Multiple Sclerosis by Magnetic Resonance Imaging. Global Journal of Health Science, 7(6), 68–78. https://doi.org/10.5539/gjhs.v7n6p68

Bengio, Y., Courville, A., & Vincent, P. (2013). Representation Learning: A Review and New Perspectives. IEEE Transactions on Pattern Analysis and Machine Intelligence, 35(8), 1798–1828. https://doi.org/10.1109/TPAMI.2013.50

Bickelhaupt, S., Paech, D., Kickingereder, P., Steudle, F., Lederer, W., Daniel, H., Götz, M., Gählert, N., Tichy, D., Wiesenfarth, M., Laun, F. B., Maier-Hein, K. H., Schlemmer, H.-P., & Bonekamp, D. (2017). Prediction of malignancy by a radiomic signature from contrast agent-free diffusion MRI in suspicious breast lesions found on screening mammography. Journal of Magnetic Resonance Imaging, 46(2), 604–616. https://doi.org/10.1002/jmri.25606

Carvalho, E. D., Filho, A. O. C., Silva, R. R. V., Araújo, F. H. D., Diniz, J. O. B., Silva, A. C., Paiva, A. C., & Gattass, M. (2020). Breast cancer diagnosis from histopathological images using textural features and CBIR. Artificial Intelligence in Medicine, 105, 101845. https://doi.org/10.1016/j.artmed.2020.101845

Chagas, P., Souza, L., Araújo, I., Aldeman, N., Duarte, A., Angelo, M., dos-Santos, W. L. C., & Oliveira, L. (2020). Classification of glomerular hypercellularity using convolutional features and support vector machine. Artificial Intelligence in Medicine, 103, 101808. https://doi.org/10.1016/j.artmed.2020.101808

Conti, A., Duggento, A., Indovina, I., Guerrisi, M., & Toschi, N. (2020). Radiomics in breast cancer classification and prediction. Seminars in Cancer Biology. https://doi.org/10.1016/j.semcancer.2020.04.002

Ekici, S., & Jawzal, H. (2020). Breast cancer diagnosis using thermography and convolutional neural networks. Medical Hypotheses, 137, 109542. https://doi.org/10.1016/j.mehy.2019.109542

Gillies, R. J., Kinahan, P. E., & Hricak, H. (2016). Radiomics: Images Are More than Pictures, They Are Data. Radiology, 278(2), 563–577. https://doi.org/10.1148/radiol.2015151169

Griethuysen, J. J. M. van, Fedorov, A., Parmar, C., Hosny, A., Aucoin, N., Narayan, V., Beets-Tan, R. G. H., Fillion-Robin, J.-C., Pieper, S., & Aerts, H. J. W. L. (2017). Computational Radiomics System to Decode the Radiographic Phenotype. Cancer Research, 77(21), e104–e107. https://doi.org/10.1158/0008-5472.CAN-17-0339

Gualberto, A. (2020, August 9). How to apply data augmentation to deal with unbalanced datasets in 20 lines of code. Medium. https://medium.com/analytics-vidhya/how-to-apply-data-augmentation-to-deal-with-unbalanced-datasets-in-20-lines-of-code-ada8521320c9

Hadad, O., Bakalo, R., Ben-Ari, R., Hashoul, S., & Amit, G. (2017). Classification of breast lesions using cross-modal deep learning. 2017 IEEE 14th International Symposium on Biomedical Imaging (ISBI 2017), 109–112. https://doi.org/10.1109/ISBI.2017.7950480

Hepsağ, P. U., Özel, S. A., & Yazıcı, A. (2017). Using deep learning for mammography classification. 2017 International Conference on Computer Science and Engineering (UBMK), 418–423. https://doi.org/10.1109/UBMK.2017.8093429

Honda, E., Nakayama, R., Koyama, H., & Yamashita, A. (2016). Computer-Aided Diagnosis Scheme for Distinguishing Between Benign and Malignant Masses in Breast DCE-MRI. Journal of Digital Imaging, 29(3), 388–393. https://doi.org/10.1007/s10278-015-9856-7

Jiang, F., Liu, H., Yu, S., & Xie, Y. (2017). Breast mass lesion classification in mammograms by transfer learning. Proceedings of the 5th International Conference on Bioinformatics and Computational Biology, 59–62. https://doi.org/10.1145/3035012.3035022

Just, N. (2014). Improving tumour heterogeneity MRI assessment with histograms. British Journal of Cancer, 111(12), 2205–2213. https://doi.org/10.1038/bjc.2014.512

Kaya, I. E., Pehlivanlı, A. Ç., Sekizkardeş, E. G., & Ibrikci, T. (2017). PCA based clustering for brain tumor segmentation of T1w MRI images. Computer Methods and Programs in Biomedicine, 140, 19–28. https://doi.org/10.1016/j.cmpb.2016.11.011

Keek, S. A., Leijenaar, R. T., Jochems, A., & Woodruff, H. C. (2018). A review on radiomics and the future of theranostics for patient selection in precision medicine. The British Journal of Radiology, 91(1091), 20170926. https://doi.org/10.1259/bjr.20170926

Khan, H. N., Shahid, A. R., Raza, B., Dar, A. H., & Alquhayz, H. (2019). Multi-View Feature Fusion Based Four Views Model for Mammogram Classification Using Convolutional Neural Network. IEEE Access, 7, 165724–165733. https://doi.org/10.1109/ACCESS.2019.2953318

Khan, S., Islam, N., Jan, Z., Ud Din, I., & Rodrigues, J. J. P. C. (2019). A novel deep learning based framework for the detection and classification of breast cancer using transfer learning. Pattern Recognition Letters, 125, 1–6. https://doi.org/10.1016/j.patrec.2019.03.022

Khan, S. U., Islam, N., Jan, Z., Haseeb, K., Shah, S. I. A., & Hanif, M. (2021). A machine learning-based approach for the segmentation and classification of malignant cells in breast cytology images using gray level co-occurrence matrix (GLCM) and support vector machine (SVM). Neural Computing and Applications. https://doi.org/10.1007/s00521-021-05697-1

Kumar, V., Gu, Y., Basu, S., Berglund, A., Eschrich, S. A., Schabath, M. B., Forster, K., Aerts, H. J. W. L., Dekker, A., Fenstermacher, D., Goldgof, D. B., Hall, L. O., Lambin, P., Balagurunathan, Y., Gatenby, R. A., & Gillies, R. J. (2012). Radiomics: The process and the challenges. Magnetic Resonance Imaging, 30(9), 1234–1248. https://doi.org/10.1016/j.mri.2012.06.010

Lambin, P., Rios-Velazquez, E., Leijenaar, R., Carvalho, S., Stiphout, R. G. P. M. van, Granton, P., Zegers, C. M. L., Gillies, R., Boellard, R., Dekker, A., & Aerts, H. J. W. L. (2012). Radiomics: Extracting more information from medical images using advanced feature analysis. European Journal of Cancer, 48(4), 441–446. https://doi.org/10.1016/j.ejca.2011.11.036

Lang, N., Zhang, Y., Zhang, E., Zhang, J., Chow, D., Chang, P., Yu, H. J., Yuan, H., & Su, M.-Y. (2019). Differentiation of spinal metastases originated from lung and other cancers using radiomics and deep learning based on DCE-MRI. Magnetic Resonance Imaging, 64, 4–12. https://doi.org/10.1016/j.mri.2019.02.013

Lewis, S., Hectors, S., & Taouli, B. (2021). Radiomics of hepatocellular carcinoma. Abdominal Radiology, 46(1), 111–123. https://doi.org/10.1007/s00261-019-02378-5

Li, J., Fan, M., Zhang, J., & Li, L. (2017). Discriminating between benign and malignant breast tumors using 3D convolutional neural network in dynamic contrast enhanced-MR images. Medical Imaging 2017: Imaging Informatics for Healthcare, Research, and Applications, 10138, 1013808. https://doi.org/10.1117/12.2254716

Lin, H.-T., & Lin, C.-J. (n.d.). A Study on Sigmoid Kernels for SVM and the Training of non-PSD Kernels by SMO-type Methods. 32.

Liu, D. (2017, November 30). A Practical Guide to ReLU. Medium. https://medium.com/@danqing/a-practical-guide-to-relu-b83ca804f1f7

Maicas, G., Carneiro, G., Bradley, A. P., Nascimento, J. C., & Reid, I. (2017). Deep Reinforcement Learning for Active Breast Lesion Detection from DCE-MRI. In M. Descoteaux, L. Maier-Hein, A. Franz, P. Jannin, D. L. Collins, & S. Duchesne (Eds.), Medical Image Computing and Computer Assisted Intervention − MICCAI 2017 (pp. 665–673). Springer International Publishing. https://doi.org/10.1007/978-3-319-66179-7_76

Mao, N., Yin, P., Wang, Q., Liu, M., Dong, J., Zhang, X., Xie, H., & Hong, N. (2019). Added Value of Radiomics on Mammography for Breast Cancer Diagnosis: A Feasibility Study. Journal of the American College of Radiology, 16(4), 485–491. https://doi.org/10.1016/j.jacr.2018.09.041

Marentakis, P., Karaiskos, P., Kouloulias, V., Kelekis, N., Argentos, S., Oikonomopoulos, N., & Loukas, C. (2021). Lung cancer histology classification from CT images based on radiomics and deep learning models. Medical & Biological Engineering & Computing, 59(1), 215–226. https://doi.org/10.1007/s11517-020-02302-w

Mayerhoefer, M. E., Materka, A., Langs, G., Häggström, I., Szczypiński, P., Gibbs, P., & Cook, G. (2020). Introduction to Radiomics. Journal of Nuclear Medicine: Official Publication, Society of Nuclear Medicine, 61(4), 488–495. https://doi.org/10.2967/jnumed.118.222893

Mendel, K., Li, H., Sheth, D., & Giger, M. (2019). Transfer Learning From Convolutional Neural Networks for Computer-Aided Diagnosis: A Comparison of Digital Breast Tomosynthesis and Full-Field Digital Mammography. Academic Radiology, 26(6), 735–743. https://doi.org/10.1016/j.acra.2018.06.019

Nahid, A.-A., & Kong, Y. (2018). Histopathological Breast-Image Classification Using Local and Frequency Domains by Convolutional Neural Network. Information, 9(1), 19. https://doi.org/10.3390/info9010019

Nam, K. J., Park, H., Ko, E. S., Lim, Y., Cho, H.-H., & Lee, J. E. (2019). Radiomics signature on 3T dynamic contrast-enhanced magnetic resonance imaging for estrogen receptor-positive invasive breast cancers: Preliminary results for correlation with Oncotype DX recurrence scores. Medicine, 98(23), e15871. https://doi.org/10.1097/MD.0000000000015871

Parekh, V., & Jacobs, M. A. (2016). Radiomics: A new application from established techniques. Expert Review of Precision Medicine and Drug Development, 1(2), 207–226. https://doi.org/10.1080/23808993.2016.1164013

Parekh, V. S., & Jacobs, M. A. (2017). Integrated radiomic framework for breast cancer and tumor biology using advanced machine learning and multiparametric MRI. Npj Breast Cancer, 3(1), 1–9. https://doi.org/10.1038/s41523-017-0045-3

Park, B. E., Jang, W. S., & Yoo, S. K. (2016). Texture Analysis of Supraspinatus Ultrasound Image for Computer Aided Diagnostic System. Healthcare Informatics Research, 22(4), 299–304. https://doi.org/10.4258/hir.2016.22.4.299

Park, J. E., & Kim, H. S. (2018). Radiomics as a Quantitative Imaging Biomarker: Practical Considerations and the Current Standpoint in Neuro-oncologic Studies. Nuclear Medicine and Molecular Imaging, 52(2), 99–108. https://doi.org/10.1007/s13139-017-0512-7

Park, J. E., Park, S. Y., Kim, H. J., & Kim, H. S. (2019). Reproducibility and Generalizability in Radiomics Modeling: Possible Strategies in Radiologic and Statistical Perspectives. Korean Journal of Radiology, 20(7), 1124–1137. https://doi.org/10.3348/kjr.2018.0070

Ragab, D. A., Sharkas, M., Marshall, S., & Ren, J. (2019). Breast cancer detection using deep convolutional neural networks and support vector machines. PeerJ, 7, e6201. https://doi.org/10.7717/peerj.6201

Rankin, S. C. (2000). MRI of the breast. The British Journal of Radiology, 73(872), 806–818. https://doi.org/10.1259/bjr.73.872.11026854

Rosebrock, A. (2019, July 8). Keras ImageDataGenerator and Data Augmentation. PyImageSearch. https://www.pyimagesearch.com/2019/07/08/keras-imagedatagenerator-and-data-augmentation/

Sayed, G. I., & Hassanien, A. E. (2017). Moth-flame swarm optimization with neutrosophic sets for automatic mitosis detection in breast cancer histology images. Applied Intelligence, 47(2), 397–408. https://doi.org/10.1007/s10489-017-0897-0

Schindelin, J., Arganda-Carreras, I., Frise, E., Kaynig, V., Longair, M., Pietzsch, T., Preibisch, S., Rueden, C., Saalfeld, S., Schmid, B., Tinevez, J.-Y., White, D. J., Hartenstein, V., Eliceiri, K., Tomancak, P., & Cardona, A. (2012). Fiji: An open-source platform for biological-image analysis. Nature Methods, 9(7), 676–682. https://doi.org/10.1038/nmeth.2019

Song, R., Li, T., & Wang, Y. (2020). Mammographic Classification Based on XGBoost and DCNN With Multi Features. IEEE Access, 8, 75011–75021. https://doi.org/10.1109/ACCESS.2020.2986546

Spanhol, F. A., Oliveira, L. S., Petitjean, C., & Heutte, L. (2016). A Dataset for Breast Cancer Histopathological Image Classification. IEEE Transactions on Biomedical Engineering, 63(7), 1455–1462. https://doi.org/10.1109/TBME.2015.2496264

Szegedy, C., Vanhoucke, V., Ioffe, S., Shlens, J., & Wojna, Z. (2015). Rethinking the Inception Architecture for Computer Vision. ArXiv:1512.00567 [Cs]. http://arxiv.org/abs/1512.00567

Tagliafico, A. S., Piana, M., Schenone, D., Lai, R., Massone, A. M., & Houssami, N. (2020). Overview of radiomics in breast cancer diagnosis and prognostication. The Breast, 49, 74–80. https://doi.org/10.1016/j.breast.2019.10.018

Tagliafico, A. S., Valdora, F., Mariscotti, G., Durando, M., Nori, J., Forgia, D. L., Rosenberg, I., Caumo, F., Gandolfo, N., Houssami, N., & Calabrese, M. (2018). An exploratory radiomics analysis on digital breast tomosynthesis in women with mammographically negative dense breasts. The Breast, 40, 92–96. https://doi.org/10.1016/j.breast.2018.04.016

Tan, Y. J., Sim, K. S., & Ting, F. F. (2017). Breast cancer detection using convolutional neural networks for mammogram imaging system. 2017 International Conference on Robotics, Automation and Sciences (ICORAS), 1–5. https://doi.org/10.1109/ICORAS.2017.8308076

Thakran, S., Gupta, P. K., Kabra, V., Saha, I., Jain, P., Gupta, R. K., & Singh, A. (2018). Characterization of breast lesion using T1-perfusion magnetic resonance imaging: Qualitative vs. quantitative analysis. Diagnostic and Interventional Imaging, 99(10), 633–642. https://doi.org/10.1016/j.diii.2018.05.006

Toğaçar, M., Özkurt, K. B., Ergen, B., & Cömert, Z. (2020). BreastNet: A novel convolutional neural network model through histopathological images for the diagnosis of breast cancer. Physica A: Statistical Mechanics and Its Applications, 545, 123592. https://doi.org/10.1016/j.physa.2019.123592

Traverso, A., Wee, L., Dekker, A., & Gillies, R. (2018). Repeatability and Reproducibility of Radiomic Features: A Systematic Review. International Journal of Radiation Oncology*Biology*Physics, 102(4), 1143–1158. https://doi.org/10.1016/j.ijrobp.2018.05.053

Truhn, D., Schrading, S., Haarburger, C., Schneider, H., Merhof, D., & Kuhl, C. (2018). Radiomic versus Convolutional Neural Networks Analysis for Classification of Contrast-enhancing Lesions at Multiparametric Breast MRI. Radiology, 290(2), 290–297. https://doi.org/10.1148/radiol.2018181352

Velichko, Y. S., Mozafarykhamseh, A., Trabzonlu, T. A., Zhang, Z., Rademaker, A. W., & Yaghmai, V. (2020). Association Between the Size and 3D CT-Based Radiomic Features of Breast Cancer Hepatic Metastasis. Academic Radiology. https://doi.org/10.1016/j.acra.2020.03.004

Whitney, H. M., Taylor, N. S., Drukker, K., Edwards, A. V., Papaioannou, J., Schacht, D., & Giger, M. L. (2019). Additive Benefit of Radiomics Over Size Alone in the Distinction Between Benign Lesions and Luminal A Cancers on a Large Clinical Breast MRI Dataset. Academic Radiology, 26(2), 202–209. https://doi.org/10.1016/j.acra.2018.04.019

Wu, N., Phang, J., Park, J., Shen, Y., Huang, Z., Zorin, M., Jastrzębski, S., Févry, T., Katsnelson, J., Kim, E., Wolfson, S., Parikh, U., Gaddam, S., Lin, L. L.

Y., Ho, K., Weinstein, J. D., Reig, B., Gao, Y., Toth, H., … Geras, K. J. (2020). Deep Neural Networks Improve Radiologists' Performance in Breast Cancer Screening. IEEE Transactions on Medical Imaging, 39(4), 1184–1194. https://doi.org/10.1109/TMI.2019.2945514

Yang, F., Ford, J. C., Dogan, N., Padgett, K. R., Breto, A. L., Abramowitz, M. C., Dal Pra, A., Pollack, A., & Stoyanova, R. (2018). Magnetic resonance imaging (MRI)-based radiomics for prostate cancer radiotherapy. Translational Andrology and Urology, 7(3), 445–458. https://doi.org/10.21037/tau.2018.06.05

Yip, S. S. F., & Aerts, H. J. W. L. (2016). Applications and limitations of radiomics. Physics in Medicine and Biology, 61(13), R150–R166. https://doi.org/10.1088/0031-9155/61/13/R150

Zhang, Q., Xiao, Y., Suo, J., Shi, J., Yu, J., Guo, Y., Wang, Y., & Zheng, H. (2017). Sonoelastomics for Breast Tumor Classification: A Radiomics Approach with Clustering-Based Feature Selection on Sonoelastography. Ultrasound in Medicine and Biology, 43(5), 1058–1069. https://doi.org/10.1016/j.ultrasmedbio.2016.12.016

Zhang, X., Zhang, Y., Han, E. Y., Jacobs, N., Han, Q., Wang, X., & Liu, J. (2017). Whole mammogram image classification with convolutional neural networks. 2017 IEEE International Conference on Bioinformatics and Biomedicine (BIBM), 700–704. https://doi.org/10.1109/BIBM.2017.8217738

Zheng, H., Gu, Y., Qin, Y., Huang, X., Yang, J., & Yang, G.-Z. (2018). Small Lesion Classification in Dynamic Contrast Enhancement MRI for Breast Cancer Early Detection. In A. F. Frangi, J. A. Schnabel, C. Davatzikos, C. Alberola-López, & G. Fichtinger (Eds.), Medical Image Computing and Computer Assisted Intervention – MICCAI 2018 (pp. 876–884). Springer International Publishing. https://doi.org/10.1007/978-3-030-00934-2_97

Zhou, Z., Qutaish, M., Han, Z., Schur, R. M., Liu, Y., Wilson, D. L., & Lu, Z.-R. (2015). MRI detection of breast cancer micrometastases with a fibronectin-targeting contrast agent. Nature Communications, 6(1), 7984. https://doi.org/10.1038/ncomms8984

Zwanenburg, A., Leger, S., Vallières, M., & Löck, S. (2020). Image biomarker standardisation initiative. Radiology, 295(2), 328–338. https://doi.org/10.1148/radiol.2020191145

BIOGRAPHY

He completed his B.Sc. in Electrical and Electronics Engineering degree from the University of Dhaka, Bangladesh in 2017. In 2018, he received Turkish Govt. scholarship for postgraduate studies in Cukurova University, Turkey. His research interests include Data Analysis, Machine Learning, IoT, Deep Learning in Healthcare.