A Review Comparative Mammography Image Analysis on Modified CNN Deep Learning Method

Siti Ramadhani
Departement of Informatic Engineering, Faculty of Science and Technology
Universitas Islam Negeri Sultan Syarif Kasim Riau
Email: siti.ramadhani@uin-suska.ac.id

Article Info

Article history:
Received Oct 03rd, 2020
Revised Jan 16th, 2021
Accepted March 04th, 2021

Keyword:
Deep Learning
CNN
Datasets
Mammography
Breast Cancer

ABSTRACT

This study aims to review the classification of breast abnormality accuracy on deep learning using comparative CNN development of concepts and models in various cases and implementation. The CNN-based breast mass detection approach localizes and classifies the masses on the images as benign or malignant simultaneously by exploring all major types of medical image modalities collected on dataset and hospital this CNN method modified to R-CNN and SD-CNN based on modification on feature extraction to improve accuracy level. R-CNN adopts RPN and ROI for Feature extraction. The model's purpose is to be design, train, and evaluate the achieved detection of accuracy. R-CNN's This model has been designed to get a detection accuracy level of up to 91.86%, the sensitivity level is 94.67%, and the AUC-ROC level is 92.2%. SD-CNN study the two-fold applicability of CNN to improve breast cancer diagnosis. This method combines images from CEDM for analysis of breast abnormalities using the Deep-CNN method with virtual feature images. The experiments produced features from the LE images at an accuracy of 0.85 and AUC of 0.84. When the recombination imaging feature was added, the model's performance increased to an accuracy of 0.89 with an AUC of 0.91 until 0.92.

Corresponding Author:
Departement of Information Engineering,
Faculty of Science and Technology,
Universitas Islam Negeri Sultan Syarif Kasim Riau,
155 HR. Soebrantas Rd Panam Km. 15, Tuah Madani, Kec. Tampan, Kabupaten Kampar, Riau 28293
Email: siti.ramadhani@uin-suska.ac.id

DOI: http://dx.doi.org/10.24014/ijaidm.v4i1.10891

1. INTRODUCTION

Cancer is a disease caused by changes that occur in cells that spread uncontrollably. Most cancer cells form a lump called a tumor as the part of the body where it grows. Breast cancer generally affects women cancers and causes the death of the woman who suffered it. About 1 in 8 of U.S.women (~12%) will potentially have malignant breast cancer during their lifetime based on the U.S. Breast Cancer Statistics, 2018)[1]. Deaths from breast cancer have significantly decreased since the implementation of population-based breast cancer screening programs in the late 1970s due to improved early detection of cancer and cancer treatment methods. Based on data of life expectancy of sufferers, the number of new cases in 2018 was approximately 18,078,957 and 9,555,027 deaths (52.85%). Breast cancer cases totaled 2,088,849 (11.55%) and an estimated 626,679 deaths (6.56%).

In general, breast tumors have two types, benign and malignant. Benign is a non-invasive(non-cancerous), while malignant is an invasive (cancerous) type of tumor. A benign or malignant disorder depends on its invasive stage. Both tumors have further subtypes that need individually diagnosed. Each symptom can cause different symptoms and different treatment plans as well. Breast cancer usually does not produce pain at an early stage that is still easy to treat, so screening is essential for early detection. The lack of early detection caused thousands of women to feel pain, leading to lower survival rates and surgical scars due to surgery.

Survival rates vary according to the stage cancer is detected[2]. Early detection of breast cancer has shown a decrease in mortality rates between 38% and 48%. However, manual analysis of mammograms and the reinterpretation leads to a 10% - 30% misdiagnosis rate.

Medical imaging is usually done manually by one or more physicians (radiologists, sinologists, or pathologists) and facing three significant problems[3]. First, more than one pathologist in one place is usually not available in developing countries. Second, the image analysis procedure for multi-classification brc (breast cancer) is complicated and time consuming for pathologists. Therefore, pathologists may experience a grueling state that leads to a deterioration in the interpretation quality during image analysis. Finally, identifying a reliable BrC subtype depends on the professional experience and domain knowledge of the pathologist. This problem can lead to misdiagnosis, especially in the early stages of BrC. During this time interval, the disease can reach an uncontrolled stage before the cancer is positive, leading to lower life expectancy levels.

Mammography is a low-dose x-ray procedure that can visualize images of the internal structure of the breast. Mammography is a standard imaging modality used to detect breast abnormalities at an early stage. Hence, microcalcification and mass are early signs of breast cancer that can only notice using imaging modalities[4]. Among the existing imaging modalities, there are other imaging modalities, namely Full Field Digital Mammography (FFDM), as the only clinically acceptable imaging modality for population-based breast cancer screening. Simultaneously, Ultrasound (USA) and Magnetic Resonance Imaging (MRI) are the additional imaging modalities for mammography image detection for specific subgroups in certain women. In clinical breast imaging (AS, MRI, FFDM, and CEDM), reading and interpreting images remains a difficult task for radiologists.

Computer-aided detection (CADe) and diagnosis scheme (CADx) is a scheme that has been developed and demonstrates the clinical potential to be used as a "second reader" to assist radiology performance in diagnostic activities. Computer-assisted diagnosis systems (CAD) can serve as a second opinion for solving brc multi-classification problems. CAD systems are an affordable, available, fast, and reliable source of early[7] diagnosis. The system helps radiologists and doctors identify abnormalities using various imaging modalities to reduce 30 to 70%[5]. Therefore, a system like this can affect humans, increase diagnosis rates, and reduce overall treatment costs as it reduces false positive and negative predictions (FN). Therefore, this issue is a challenge, and much research uses Machine Learning and Deep Learning learning methods.

Deep Learning is a machine learning technique in which computer models perform classification tasks directly by learning from text, images, or sounds. The model train on many CNN datasets and architectures containing many layers[6]. In medical imaging, deep learning use to detecting cancer cells automatically. In its recent research, due to a large amount of data, the high computing power of the Graphics Processing Unit (GPU), deep Learning has shown promising success in Natural Language Processing (NLP), object detection, and medical image analysis. Deep learning-based methods are sensitive to image acquisition settings, scanner types, and applied image e-processing.

CNN's application use in medical imaging since the 1990s when the classification method detected in digital mammography. "Transferability" is one crucial aspect of CNN, embedded in CNN preprocessing. Research shows that transfer learning in the field of medical imaging categorizes into two groups[7]. First, to use network preprocessing to extract features from particular network layers, they train a new pattern classifier.

The Convolutional Neural Network (CNN), based on a breast mass detection approach, simultaneously localizes and classifies mass into benign or malignant abnormalities by exploring all kinds of medical image modalities collected on the dataset[8]. The validation process and test methods take from different sites, such as various hospitals, clinic datasets, and Mammography datasets available on the internet.

The review of CNN's application in the study presented a systematic review of CNN-based CAD systems for Brc image classification of five aspects: BrC imaging modality, dataset, image processing, CNN workings, and performance accuracy measurement[9]. This review adopts a systematic review methodology for finding and selecting studies from well-known sources to ensure the authenticity and quality of selected literature[10]. Besides, this review provides a critical analysis of CNN's performance on its common dataset. Finally, this review presents 3-way research in Brc image classification using CNN's method and modification.

2. RESEARCH METHOD

Analysis of the literature that has been done, among others, obtained results related to the development of concepts and models in the use of CNN in various cases and implementation environments as in some related studies. The review showed that the BrC classification consisted of several unique medical imaging modalities and their combination known as multimodality. Distribution of various modalities imaging Table 1. Imaging modalities can be colored images and gray images.

Table 1. Various Medical Imaging Modalities

Type	Short description
Mammogram (Mg)	Mammograms are found in three forms, such as screen film mammograms (SFMs), digital mammograms (DMs), and digital breast tomography (DBT). SFMs and DMs are grayscale 2D, but DBT provides several 2D grayscale image frames that appear like black-and-white videos.
ULTRASOUND (US)	The US is also known as Sonograms. US images are used in three combinations: simple 2D grayscale US images, US images along with additional additive features of shear-wave elastography (SWE) color images, and US images along with Nakagami color images.
Magnetic Resonance Imaging (MRI)	MRI is used with pre and post contrast (DynamicContrast-enhanced (DCE) to diagnose BrC. Post contrast images are colored images but are usually converted to grayscale to feed to ANN
Histopathology (HP) Images	HP Images is a color image stained with H&E and divided into two categories: the entire slide image (WSI) and the image patch extracted from WSI by the pathologist.
Multimodality	Some studies use a combination of two grayscale image modalities named as multimodality antara MG for brc classification.
Full-field digital mammography (FFDM)	This imaging modality is a clinically acceptable image for population-based breast cancer screening, while Ultrasound (USA) and Magnetic Resonance Imaging (MRI) are also used as additional imaging modalities for mammography for certain female subgroups..
Contrast-Enhanced Digital Mammography (CEDM)	The latest development of digital mammography uses intra-venous injection of iodinated contrast agents in conjunction with mammography examination and combines FFDM and MRI.

The classification image features are divided into two classes (normal and abnormal) and three classes (normal, benign, and malignant). Despite the limitations of Mg-based classifications (two or three class labels), HP images play an essential role in solving multi-class problems (up to eight subtypes) in the BrC classification. This model tested on public datasets is more reliable than models tested on complete datasets, independent of database types (exclusive or shared) at an abstract, grayscale (e.g., Mg, AS, and MRI), or colored images (e.g., HP images) used for BrC classification. Most studies do binary type, and very few studies focus on multi-class issues for BrC classification.

CNN needs a massive amount of data to train and achieves high accuracy. The literature data set was collected from hospitals using reports of pathology documents and datasets available on the internet. Due to the large availability of datasets, training and testing play an essential part in the most common datasets available on the internet, such as those shown in table 2.

Table 2. Available class data sets and labels

#	Data Set Name	Pictures	Class labels
1	BCDR	1734	3 : Normal, Benign, Ferocious
2	CBIS-DDSM	4067	2: Benign, Ferocious
3	DDSM	10480	2: Benign, Ferocious
4	INBreast	419	2 or 3 : 2: Benign, Malignant or Normal, Benign, Malignant
5	Mias / Mini-MIAS	322	2: Benign, Ferocious
6	BICBH	269	4: Normal, Benign, Carcinoma in situ, Carcinoma
8	BreakHis (In Close)	7909	2 or 8

The CNN modifications implemented in the CNN study discussed this time are in two forms, namely Shallow-Deep CNN (SD-CNN). Improvements include preprocessing techniques adopted in the BrC classification process. In general, BrC image processing tasks involve augmentation, ROI extraction, scaling, image normalization, cropping, stain normalization, feature reduction, denoising, and image reputation. The importance of preprocessing due to raw images (without preprocessing) usually shifts the classification model's focus and can lead to high classification error rates.

3. RESULT AND ANALYSIS

The MIAS Mammograms dataset has a total of 322 images already on a grayscale. SGDM training, using the stochastic gradient reduction momentum. Optimal results obtaining by setting parameters such as base learning rate, mini-batch dan max epochs—the original data is 1024-by-1024. The information is divided into two classes: regular and abnormal classes, 150 for training data and 100 for test data, then subdivided into

70:30 comparisons. The data resized to 224-by-224, noise is removed by applying morphological operations such as binarization and masking to extract Region of Interest (ROIs)[10]. Dataset is divided into seven sub-classes, 6 of which include various types of abnormalities and 1 class containing only typical images. The implementation process using Matlab 2017a dan Figure 1. demonstrate the methodology for the destination system.

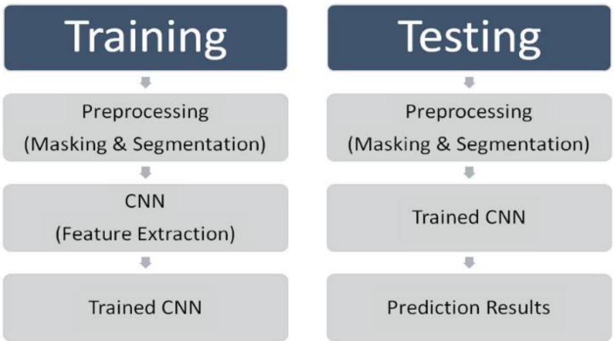


Figure 1. Conceptual block diagrams of training and testing procedures[8]

The segmentation process carries out on the raw image as an input in which the morphological closure is applied for noise removal. Morphological closure performs element structure settings. It helps in the removal of small stains and removes small holes. Existing components are connected in binary images. Among all connected areas extracted, the largest related site is selected for masking. Masking is applied at the end to set the background pixel value as zero resulting in the image as shown in figure 2.

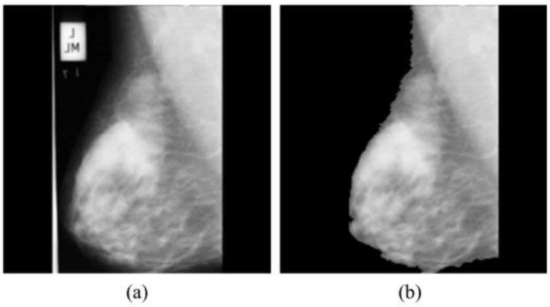


Figure 2. Morphological surgery used in data. (a) a raw image example of a MIAS dataset (b) ROI segmentation results using morphological closing and masking operations[8].

There are two methods for Training and testing; firstly, Dividing the dataset into two classes: normal and abnormal, secondly, further sub-divisions of abnormal classes, which include six types of abnormalities found in the breast such as asymmetry, calcification, spiculated masses, circumscribed masses, architectural distortions, and others. Training and testing are carried out on original and preprocessed data. Preprocessing is done to get better performance and faster learning. Different filter sizes and preprocessing techniques are used in the original data to eliminate noise factors that can lower network analysis's overall accuracy. Detection accuracy on MIAS datasets is 65%. Furthermore, please note that proper segmentation is mandatory for efficient extraction and classification of features. Masking and segmentation based on morphological surgery can significantly improve classification results.

MG Images on R-CNN research collected from several local hospitals in Ethiopia. The collected images show the results of screening and diagnosis of patients. The document report results are based on pathology confirmation and Breast Imaging-Reporting Data System (BI-RADS). More than 5000 x-ray mammogram images were diagnosed between 2016 and 2018. 1588 MG images contain selected mass abnormalities and are then annotated by professional radiologists with labels or labeling.

Furthermore, this dataset is randomly divided into 80% for training, 10% for validation, and 10% for testing. At the preprocessing stage, different imaging formats such as DICOM medical image formats are converted to .png image format. The noise is removed, next to breast regions are extracted from the background, then patient information is deleted, next artifacts and other unwanted objects are cleaned. Gaussian, medium, and bilateral filters with sizes 3x3 and 5x5 are used to eliminate noise and evaluate denoised results using MSE. Two filter size, which is considered one with a length of 3x3 is used. Besides, CLAHE is used to enhance denoised MG images, after which the breast area is extracted and unwanted artifacts removed using OTSU and morphological surgery. Four different threshold values, such as T=1 (100%), 0.75 (75%), 0.5

(50%), and 0.25 (25%) used in experiments for Intersection over Union (IoU) overlap on the ground truth bound box and predictive bound box on RPN.

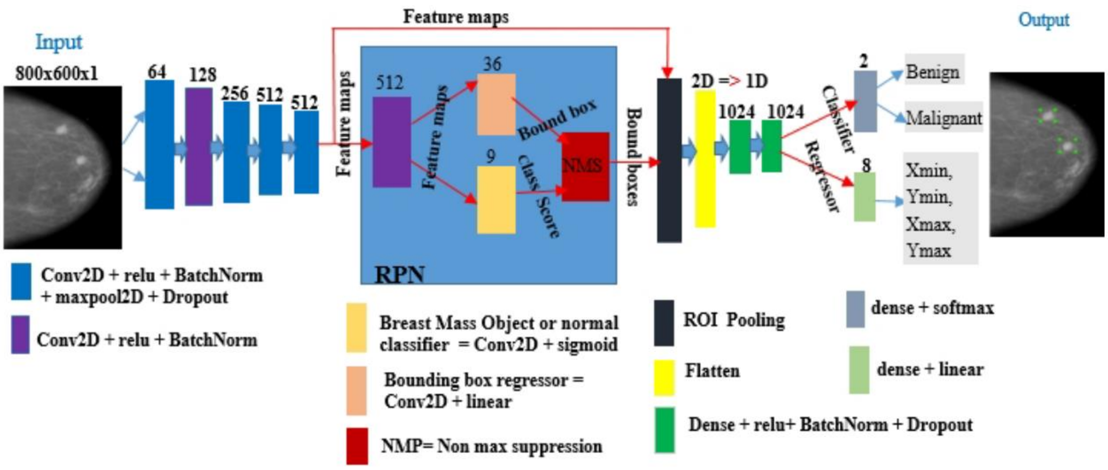
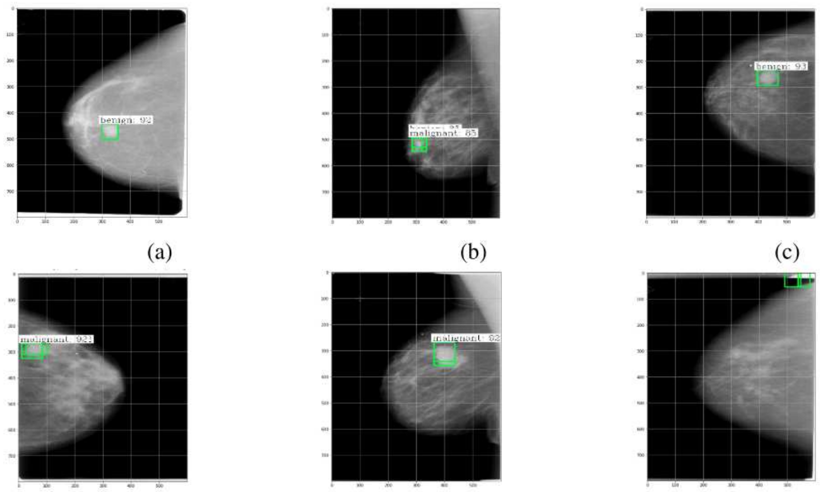


Figure 3. Breast Cancer Detection architectural flow structure. RPN and ROI detection sections in Merging adapted from Faster R-CNN[10]



(a) (b) (c)

(d) (e) (f)

Figure 4. Mass abnormality samples (a) to (b) detected by faster VGG-based R-CNN models, (c) to (e) detected and (f) incorrectly detected by the proposed to model [11].

The model produces abnormality images displayed in images 4 (a) and (b), and (c) for (e), while image 4 (f) is one of the detection errors in the MG image. The box is green in shape, and the numbers indicate a detectable mass abnormality. Each detected mass abnormality contains a border box with green color, class name, and trust score. This model is trained and evaluated the accuracy of up to 91.86%, the sensitivity of 94.67%, and AUC-ROC 92.2%.

P's next research is about FeiGao's use of Shallow -CNN et al. using the In breast dataset of a full-field digital mammographic database (FFDM), one online dataset [7]. This method is applied to FFDM images to create recombinant photos that are "virtual." A new digital image technology that is CEDM (Contrast-enhanced digital mammography), this image is low energy and bears a resemblance to FFDM and MRI or FFDM and MRI. However, since this image service can not be obtained easily in the health center, FFDM images with advanced CNN processing can be combined with augmentation or virtual images. The recombined with FFDM images will produce pictures with CEDM quality and increased accuracy in detecting malignant and benign cancers.

Working with FFDM images, a trained CNN model developed and implemented, is Deep CNN. Deep CNN is a CNN training method with large amounts of layer data set on ResNet with acceptable tuning parameters. Furthermore, CNN's capabilities are also in synthetic image rendering, where images are divided into some small sections inserted into CNN as inputs and outputs are synthesized images. The input and output images are not as linear as from successful applications from regular X-rays to bone-suppressed recombined X-rays. These two potentials are then put together in SHALLOW-Deep CNN (SD-CNN) as a new CAD to work with FFDM combined with virtual images to have capabilities such as CEDM images.

The In breast dataset consists of FFDM images with a pixel size of 70 mm (microns) and 14-bit contrast resolution. The process starts with Pre-processing damper with four steps. The first step is to identify the minimum area limit box of the tumor area. Specifically, each tumor has a list of border points with coordinates in pairs (x, y). The result is two images because it uses CC and MLO (two contour ways to recombine the image as an extraction feature) and then cloned to a new image. The limiting box's size varies per case due to the various tumor sizes (ranging from 65 × 79 to 1490 × 2137). It is further enlarged by 1.44 times (width 1.2 times and height 1.2 times). In the second step, this 'enlarged' rectangle is extracted and saved as a single image. The third step is to normalize the image's intensity to between 0 and 1 using max-min normalization. The normalized image is resized to 224×224 to retrieve the trained ResNet model results fully in the last step.

After pre-processing, the next step for Shallow-CNN is Virtual Image Rendering. This step fills through different layers (convolutional, merging, fully connected). CNN 4-layer implemented latent model relationships between LE images (patches) and re-combined images (patches). This model is then used to render-the combined "virtual" image (patch) of the FFDM (patch) image. The next proses are Deep-CNN Feature Generation. ResNet output blocks take the final classification result and the initial input (short-cut) when updating parameters. Adopting ResNet-50 consists of four types of building feature blocks. The output feature extract from the finish layer. The n-calculated average use to represent the entire feature map, which has 3840 (256 + 512 + 1024 + 2048) total features. The final stage is classification to improve Machine-Learning to reduce bias and variance using Gradient Boosting Trees (GBT). Build decisions tree and minimize the noise in tree decisions. The tree-making process serves as a feature selection and classification.

One contribution from the SD-CNN study was to investigate the recombination of FFDM into such a CEDM in aiding breast diagnosis using the Deep-CNN method. CEDM is a more promising imaging modality, which provides information from standard FFDM combined with improved characteristics associated with neoangiogenesis (similar to MRI). Using advanced ResNet in trained as a feature generator in LE image feature classification modeling, it can achieve 0.85 and AUC 0.84 accuracy, which adds re-combined imaging features, improved model performance to 0.89 accuracies, and AUC 0.91. The contribution of CEDM in developing SD-CNN to improve breast cancer diagnosis using FFDM images expected in the medical world. Shallow Deep-CNN can be applied to create "virtual" re-combined images from FFDM images to make them like CEDM images for improved breast cancer diagnosis accuracy. The experiment was conducted on 89 FFDM datasets using trained ResNet and achieved 0.84 accuracies with AUC 0.87. Furthermore, if imaging features a combined "virtual," the model's performance increases its accuracy to 0.90 with AUC 0.92.

Mei-Ling et al. also conducted research using the INBreast dataset[9]. They covered 106 images of mammography with lumps or cancer masses from a database that initially had 410 images. Preprocessing uses the CLAHE method with eight density categories on the image for the classification process. In CLAHE, augmentation doing by refining the image with a value of 11 times vertically and horizontally. Augmentation is also applied to increase accuracy, thus increasing the number of images to 7632 pieces. Image sizes are reset from 3328x4084 and 2560x3328 to 224x224 with ShuffleNEt and DenseNet and 227x227 with Alexnet. The comparison with him with the research done shown in the following table 3.

Table 3. Two different treatments on the Inbreast dataset

#	Subject	Simon Hadush Nrea, et al.	Mei-ling Huang, et al
1	Preprocessing	- Gausian filtering, median and bilateral filtering - increased image contrast with CLAHE	- Clahe - Original, rotating, flipping
2	Model Trainning	- extraction features with fivelayers of convoluted filters (64, 128, 256, 512,512) - Each convoluted layer is followed by the: o Relu activation layer, o batch normalization, o maxpooling layer, and o dropout except.	
3	Activation	- Relu, - Batch normalization, - Maxpooling layer	
4	Implementation environment	Python and hard, Tensorflow as backend	Matlab R2019a

#	Subject	Simon Hadush Nrea, et al.	Mei-ling Huang, et al
5	Category klasification	2 : Benign and ferocious	2 : Benign and ferocious
6	Proposed models and augmentation results	SD-CNN,	CNN produced 7632 number of augmentation images with 8 categories.
7	Accuracy	91,86%	

4. CONCLUSION

Modification of detection methods is processing by modifications of the pre-processing and excretion features process. This deep learning technique uses several datasets and data collected from local hospitals, such as MIAS mammograms datasets, MIAS hospital datasets, and mammographic datasets from the internet. Different filter sizes as well as different pre-processing as well are worn in native source to eliminate noises that can reduce overall accuracy. Also, note that actual segmentation is a require step for effective extraction and classification of features. R-CNN adopts RPN and ROI for Extraction Features. Models design, trained, and evaluated to achieve the expected detection accuracy. The models proposed on the R-CNN achieve detection exactness become 91.86%, the sensitivity of 94.67%, and AUC-ROC 92.2%. On SD-CNN, the application of CNN occurred in double measures to enhance breast cancer result by recombines the method from CEDM using the Deep-CNN method. CEDM is a promising imagery method that provides knowledge of basic FFDM coupled with improved feature associated with neoangiogenesis (comparabe to MRI). Experiments showed that the LE images' features could achieve 0.85 accuracies and AUC 0.84, include to combined imagining features and presentation improved accuracy by 0.89 with AUC 0.91. This model also reunites images with virtual features and improved performance to 0.90 accuracies with AUC 0.92.

REFERENCES

[1] R. Agarwal, O. Díaz, M. H. Yap, X. Lladó, and R. Martí, “Deep learning for mass detection in Full Field Digital Mammograms,” *Comput. Biol. Med.*, vol. 121, no. April, pp. 1–10, 2020, doi: 10.1016/j.compbiomed.2020.103774.

[2] M. L. Huang and T. Y. Lin, “Dataset of breast mammography images with masses,” *Data Br.*, vol. 31, p. 105928, 2020, doi: 10.1016/j.dib.2020.105928.

[3] H. R. Fajrin, H. A. Nugroho, and I. Soesanti, “Ekstraksi Ciri Berbasis Wavelet Dan Glcm Untuk Deteksi Dini Kanker Payudara Pada Citra Mammogram,” in *SNST ke-6*, 2015, pp. 47–52.

[4] P. Kaur, G. Singh, and P. Kaur, “Erratum: Intellectual detection and validation of automated mammogram breast cancer images by multi-class SVM using deep learning classification (Informatics in Medicine Unlocked (2019) 16, (S2352914818301813), (10.1016/j.imu.2019.01.001)),” *Informatics Med. Unlocked*, vol. 16, no. August, p. 100239, 2019, doi: 10.1016/j.imu.2019.100239.

[5] M. A. Mazurowski, M. Buda, A. Saha, and M. R. Bashir, “Deep learning in radiology: An overview of the concepts and a survey of the state of the art with focus on MRI,” *J. Magn. Reson. Imaging*, vol. 49, no. 4, pp. 939–954, 2019, doi: 10.1002/jmri.26534.

[6] T. Fujioka, K. Kubota, M. Mori, Y. Kikuchi, L. Katsuta, and K. U. Mizuki, Kimura; Emi, Yamaga; Mio, Adachi; Goshi, Oda; Tsuyoshi, Nakagawa; Yoshio, “Efficient Anomaly Detection with Generative Adversarial Network for Breast Ultrasound Imaging Tomoyuki,” *Diagnostics*, vol. 10, no. 456, pp. 1–10, 2020.

[7] G. Murtaza *et al.*, “Deep learning-based breast cancer classification through medical imaging modalities: state of the art and research challenges,” *Artif. Intell. Rev.*, vol. 53, no. 3, pp. 1655–1720, 2019, doi: 10.1007/s10462-019-09716-5.

[8] S. Charan, M. J. Khan, and K. Khurshid, “Breast cancer detection in mammograms using convolutional neural network,” in *2018 International Conference on Computing, Mathematics and Engineering Technologies: Invent, Innovate and Integrate for Socioeconomic Development, iCoMET 2018 - Proceedings*, 2018, pp. 1–5, doi: 10.1109/ICOMET.2018.8346384.

[9] F. Gao *et al.*, “1803.00663,” *Comput. Med. Imaging Graph.*, vol. 70, no. Desember, pp. 53–62, 2018.

[10] M. Hasan, S. Das Barman, S. Islam, and A. W. Reza, “Skin cancer detection using convolutional neural network,” in *ACM International Conference Proceeding Series*, 2020, pp. 254–258, doi: 10.1145/3330482.3330525.

[11] Shaoqing Ren, Kaiming He, Ross Girshick, and Jian Sun. Faster r-cnn: Towards real-time object detection with region proposal networks. In Advances in neural information processing systems, pp. 91–99, 2015.

BIBLIOGRAPHY OF AUTHORS

Lecturer of Informatic Engineering in Sains and Tecnology Faculty of Universitas Islam Negeri Sultan Syarif Kasim. Reach Computer Science Master degree from Universitas Putra Indonesa "YPTK" Padang, Undergraduate of Informatic Enginering STAIN Bukittinggi and Diploma from Sekolah Tinggi Teknik Payakumbuh.