Comparing ResNet34 and VGG-16 Models in Breast Cancer Detection and Diagnosis

Cyintia Aprilia Limmanto
School of Computer Science
Bina Nusantara University
Jakarta, Indonesia
cyintia.limmanto@binus.ac.id

Geocean Djohan
School of Computer Science
Bina Nusantara University
Jakarta, Indonesia
geocean.djohan@binus.ac.id

Lili Ayu Wulandhari, S. Si., M.Sc., Ph.D.
School of Computer Science
Bina Nusantara University
Jakarta, Indonesia
lili.wulandhari@binus.ac.id

Islam Nur Alam S. Kom., M. Kom.
School of Computer Science
Bina Nusantara University
Jakarta, Indonesia
islam.alam@binus.ac.id

Abstract **– Prevalence of breast cancer has risen each year for the last thirty years and has been one of the most common causes of cancer-related deaths in women. This is due to lifestyle changes and evolving risk factors. Artificial Intelligence, especially Deep Learning models such as VGG-16 and ResNet34 could assist medical professionals in detecting malignancy and speed up decision making to apply suitable medical intervention for affected patients. Early detection and management of breast cancer could significantly improve survival rates, patient outcomes, and quality of life.**

Keywords - Artificial Intelligence, Breast Cancer, Pre-trained Model, Pathology, Oncology, VGG16, ResNet34

I. INTRODUCTION

Breast cancer has been haunting the lives of so many women. It is one of the most common types of cancer diagnosed among women [1] and causing thousands of deaths worldwide each year [2]. Searching for a way to detect breast cancer early and accurately has been researched extensively over the past few decades. Early breast cancer detection helped in reducing the mortality rate as it enables effective treatment to be conducted [3].

Mammography has been one of the most utilized diagnostic methods to detect breast cancer. With the help of mammography, rates of mortality are greatly reduced by around 20% - 49%. But mammography reading caused a heavy workload for radiologists. This caused high percentages of false-positive and false-negative readings [4]. The remedy proposed to tackle this issue is integrating Artificial Intelligence to assist in mammography reading.

Technology has advanced significantly in recent years that leads to the utilization of Artificial Intelligence, specifically Deep Learning and Machine Learning in many fields, including breast cancer detection. Studies regarding Deep Learning integration for detecting breast cancer in lymph nodes showed that the accuracy level is comparable to that of flesh-and-bloods pathologists, eliminating time constraints associated with human diagnostics [5].

With this research, we aim to compare two different Deep Learning models for determining the malignancy of breast tumor tissue.

II. LITERATURE REVIEW

A. Breast Cancer

Breast cancer is a type of malignancy found in the breast tissue of men and women, but it usually affects women most. It is one of the leading causes of cancer-related deaths in the world. Incidence and death rates due to breast cancer have increased over the last three decades. This is caused by changes in lifestyle and evolving risk factors. Technological advancements also play a role in this, as better cancer detection also increases the rate of breast cancer diagnosis. About 80% of breast cancer patients currently are of populations over the age of 50 [6]. Survival rates vary widely according to the stage of the disease when it is first diagnosed, types of treatment given, treatment adherence, and molecular subtype. Treatment of breast cancer involves an array of medical interventions. These include a combination of surgery to remove malignant tissue or the whole affected breast, radiotherapy, chemotherapy, hormonal therapy, or targeted biological therapy. Prevention and treatment of breast cancer is currently in urgent need, since morbidity and mortality rates have increased significantly in recent years.

B. *Early Management of Breast Cancer and Effects on Patient Outcomes and Recovery*

When diagnosed and treated early, breast cancer patients have a greater chance of cure and survival. Early detection of breast cancer facilitates early treatment, which helps improve prognosis and avoid further spread and complication of the disease. Survival varies widely by stage at diagnosis. In 2012-2018, the 5-year relative survival rates for stage I breast cancer is more than 99%, 93% for stage II, 75% for stage III, and 29% for stage IV [7]. Later stages indicate that the disease has spread further away from the initial location. Cancer that has spread (metastasized) to distant organs is harder to cure, due to it affecting more organs in the body, higher complexity in its removal, and higher risk of toxicity and side effects from the medications used. Removal of the primary tumor in metastatic breast cancer also will not improve survival. This is the exact reason why early detection of breast cancer is crucial in improving survival chance, prognosis, and quality of life of patients, as it is easier to treat and manage cancer in its early stages.

C. Imaging techniques and diagnosis

There are several imaging techniques used to detect malignancies in the breasts. These include scans that utilize X-rays or ionizing radiation such as mammography, computed tomography (CT), positron emission computed tomography (PET), and single-photon emission computed tomography (SPECT), and scans that does not involve X-rays or the use of ionizing radiation, such as ultrasonography (USG) and magnetic resonance imaging (MRI). However, the most common techniques used are MG, USG, and MRI, due to high costs and poor practicability of other imaging techniques. CT, PET Scan, and SPECT can be used as secondary diagnostic methods for special cases of breast cancers, such as screening for metastatic breast cancer [8].

The use of imaging techniques shows the location and morphology of tumors clearly. This provides valuable information for pathologists and oncologists to diagnose and determine the severity and progression of the disease. Thus, treatment plans could be made.

D. Artificial Intelligence Assisted Breast Cancer Detection

The use of Artificial Intelligence in diagnosing breast cancer has seen a remarkable increase. Procedures previously done by humans now include AI algorithms and its derivatives, such as Deep Learning and Machine Learning. One example is the use of deep learning model convolutional neural networks (CNNs) in interpreting Digital Mammography images. This algorithm determines the presence of lesions and tumors with higher accuracy and speed than humans [9].

One study compared the accuracy and specificity of stand-alone artificial intelligence system in detecting malignancy as opposed to 101 breast radiologists. The results showed that the performance of the AI system was statistically superior to that of the average of the 101 radiologists. The AI system had an AUC higher than 62 of the 101 radiologists (61.4%) and higher sensitivity than 55 of the 95 radiologists (57.9%) [10].

In another study, the average sensitivity was shown to be significantly higher when using AI assisted detection. Similarly, the false-positive rate was decreased with the use of AI system [11]. These results clearly show that recent advances in AI technologies have narrowed the gap between human experts and computers in detecting breast malignancies in digital mammograms.

E. Previous Studies and Works

Wang et al. created a hybrid deep learning model using CNNs and GRU used on whole slide images available from PCam Kaggle dataset. The model achieved 86.21% accuracy. This accuracy is the highest compared to other ML/DL models[12].

Abunasser et al. proposed a Deep Learning Model using Xception to classify different types of breast cancer using BreaKHis dataset. The proposed model managed to get an average of 97.60% precision, 97.60% recall, and 97.58% F1-Score[13].

Shen et al. developed a Deep Learning Model using CNNs with Transfer Learning and applied it to 2 different datasets, CBIS-DDSM that consists of digitized film mammograms and INBreast database that consists of digital mammograms with ROI annotation. The best single model achieved an AUC of 0.88 and four model averaging managed to increase AUC to 0.95 on DDSM. For INBreast, the model achieved an AUC of 0.95 and with four models averaging increased it to 0.98. The study showed that deep learning model can obtain high accuracy on various mammography formats [14].

III. METHODOLOGY

A. Dataset

A collection of biopsy images from the BreaKHis (Breast Cancer Histopathological Image Classification) Dataset was collected and processed. BreaKHis is a dataset of microscopical images of biopsied benign and malignant breast tumor tissue. Samples were collected from 82 anonymous patients using the partial mastectomy method. The biopsied tissues were stained using hematoxylin and eosin (HE), and then captured into digital form. Expert pathologists diagnosed the individual slides and determined whether they are benign or malignant. This dataset contains 7909 images consisting of 2480 benign and 5429 malignant images [15], [16]. This dataset is available on Kaggle.

B. Image Preprocessing

The images were resized to size 224 x 224. Both the images used by the two models had 3 channels. The images were then transformed to Tensors and normalized before being used to train both the models.

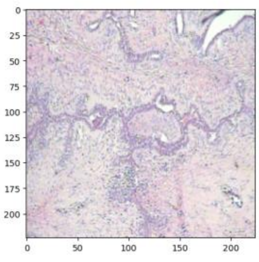

Fig. 2.1. Sample image

The images were then split into training, testing, and validation dataset. The splitting resulted in 6948 images for training, 2172 for testing, and 1738 for validation.

C. Models

The models used were VGG-16 and ResNet34.

VGG-16 pre-trained model is a version of the VGG-Net convolutional neural network. VGG16 consists of 13 convolutional layers and 3 fully connected layers. In this experiment, the top layers of the model were not included to implement transfer learning. The architecture of VGG-16 model is shown below:

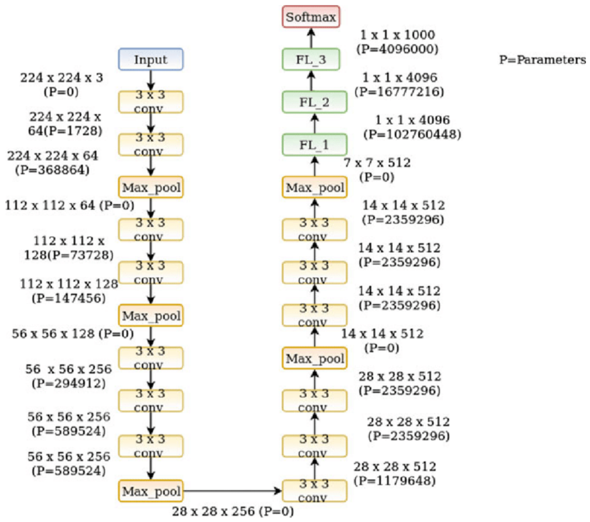

Fig. 2.3. Architecture of VGG-16 model

Another model used in this experiment is the ResNet34 pre-trained model. ResNet34 consists of 33 convolutional networks followed with a max-pooling layer, average pooling layer and a fully connected layer. In this experiment the classifier layer was not included and replaced with custom fully connected layer.

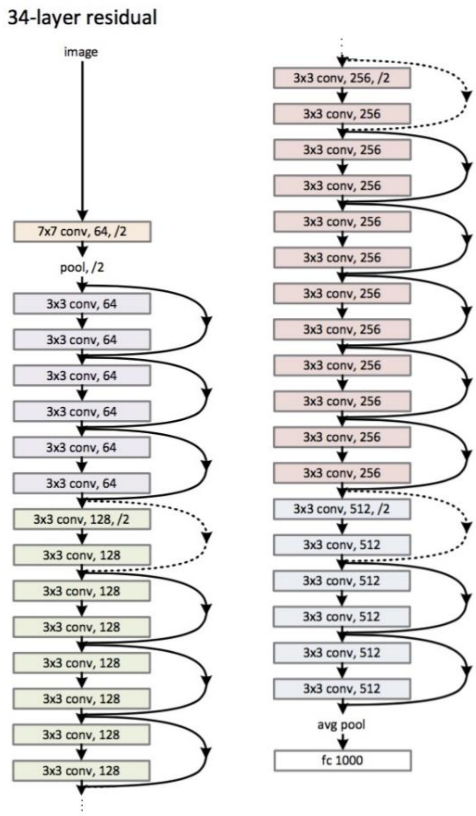

Fig. 2.4. Architecture of ResNet34 model

D. Deep Learning Framework

This research was conducted using PyTorch Stable version 2.3.1 framework on Windows Subsystem for Linux version 2, with Linux distribution ('distro') Ubuntu 22.04 LTS. Hardware acceleration was used, in this case an NVIDIA RTX 3050 Ti Laptop GPU with CUDA 11.8 compute-platform.

IV. RESULTS AND DISCUSSION

The fully connected layer of the ResNet34 expected 512 input features. The ResNet34 model has a training accuracy of 90.84% and the training loss is 0.21 The validation accuracy is 90.91% and the validation loss is 0.22.

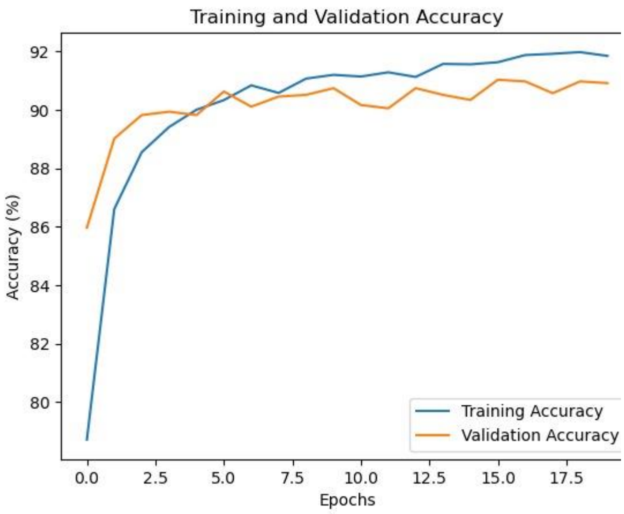

Fig. 3.1. Graph of ResNet34 Model Accuracy

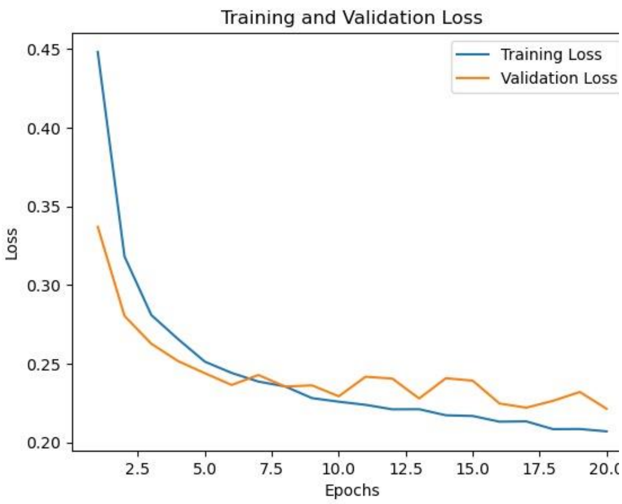

Fig. 3.2. Graph of ResNet34 Model Losses

The model was then used to predict images from the testing dataset. The results were 991 true positives, 972 true negatives, 125 false negatives, and 84 false positives.

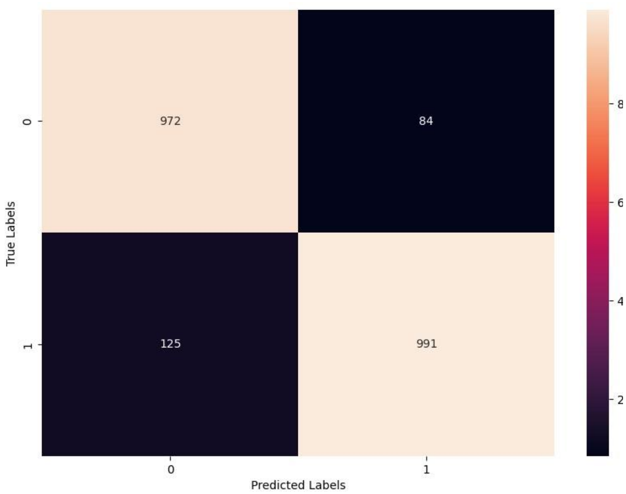

Fig 3.3. Confusion Matrix of ResNet34 Model

In contrast, the VGG-16 model's fully connected layer expected 25088 input features. The VGG-16 model has a training accuracy of 100% and the training loss is 0.0017. The validation accuracy is 93.38% and the validation loss is 0.24.

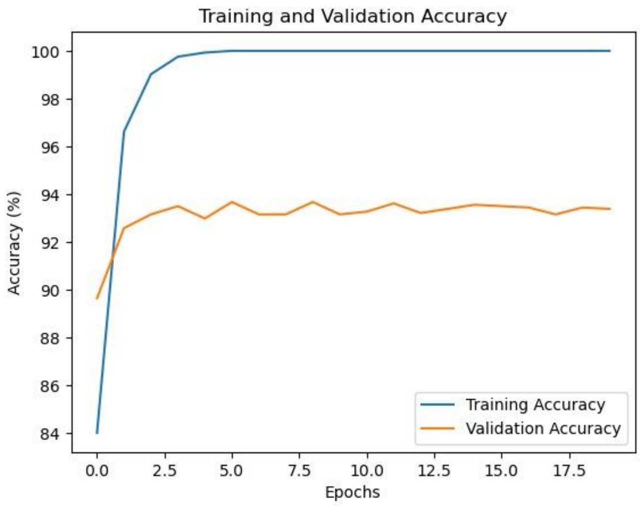

Fig. 3.4. Graph of VGG-16 Model Accuracy

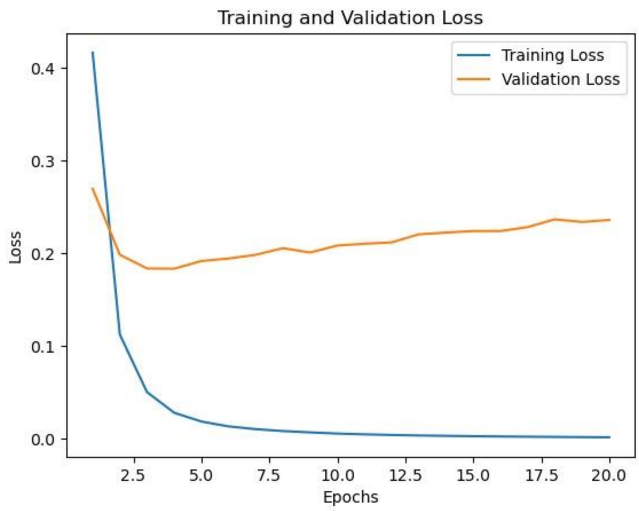

Fig. 3.5. Graph of VGG-16 Model Losses

The model was then used to predict images from the testing dataset. The results were 1036 true positives, 995 true negatives, 80 false negatives, and 61 false positives.

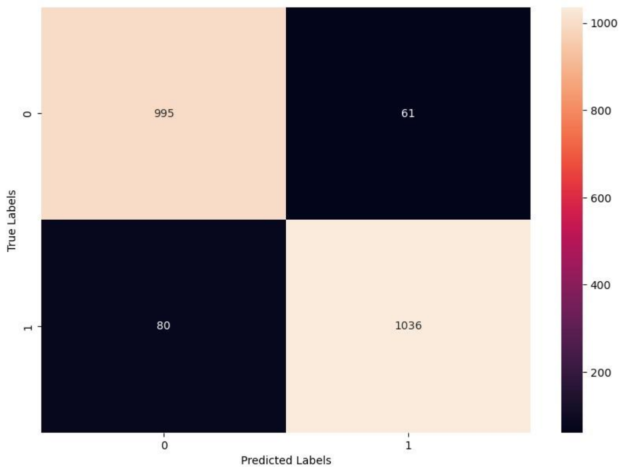

Fig. 3.6. Confusion Matrix of VGG-16 Model

V. CONCLUSION

This research looked at the performance of two pre-trained Deep Learning models in detecting malignancies in breast tumor biopsy images. The data extracted from the BreaKHis dataset needs to be preprocessed before being used to train the model.

The ResNet34 model showed significantly lower accuracy, higher losses, and higher false predictions of benign and malignant tumors present in breast tissue biopsy images compared to the VGG-16 model. This shows that using VGG-16 for determining the malignancy of breast tumor tissue is more effective than ResNet34 when it is used as-is, without further modifications and fine-tunings.

In this research, the difference in fully connected layers expected input features between VGG-16 and ResNet34 models are due to the difference in model architecture. Difference in number of input features between models impacts model performance. More features increase the complexity of the model, which could be beneficial due to more relevant information and better predictions. But more features also increase the computational cost. The VGG-16 model took significantly longer to compute compared to the ResNet34 model.

Utilizing deep learning models can assist pathologists and oncologists to make faster and more accurate diagnosis regarding whether malignancies are present in patients' breast tissue. This will shorten the time needed to treat patients with breast cancer, this in turn will improve the chance of survival and patient outcomes.

REFERENCES

[1] G. Dileep and S. G. Gianchandani Gyani, "Artificial Intelligence in Breast Cancer Screening and Diagnosis," *Cureus*, Oct. 2022, doi: 10.7759/cureus.30318.

[2] K. Freeman *et al.*, "Use of artificial intelligence for image analysis in breast cancer screening programmes: Systematic review of test accuracy," *The BMJ*, vol. 374, Sep. 2021, doi: 10.1136/bmj.n1872.

[3] Y. Shen *et al.*, "Artificial intelligence system reduces false-positive findings in the interpretation of breast ultrasound exams", doi: 10.1038/s41467-021-26023-2.

[4] S. Romero-Martín, E. Elías-Cabot, J. L. Raya-Povedano, A. Gubern-Mérida, A. Rodríguez-Ruiz, and M. Álvarez-Benito, "Stand-Alone Use of Artificial Intelligence for Digital Mammography and Digital Breast Tomosynthesis Screening: A Retrospective Evaluation," *Radiology*, vol. 302, no. 3, pp. 535–542, Mar. 2022, doi: 10.1148/radiol.211590.

[5] Y. Liu *et al.*, "Artificial Intelligence–Based Breast Cancer Nodal Metastasis Detection: Insights Into the Black Box for Pathologists," *Arch Pathol Lab Med*, vol. 143, no. 7, pp. 859–868, Jul. 2019, doi: 10.5858/arpa.2018-0147-OA.

[6] S. Łukasiewicz, M. Czeczelewski, A. Forma, J. Baj, R. Sitarz, and A. Stanisławek, "Breast Cancer—Epidemiology, Risk Factors, Classification, Prognostic Markers, and Current Treatment Strategies—An Updated Review," *Cancers (Basel)*, vol. 13, no. 17, p. 4287, Aug. 2021, doi: 10.3390/cancers13174287.

[7] A. N. Giaquinto *et al.*, "Breast Cancer Statistics, 2022," *CA Cancer J Clin*, vol. 72, no. 6, pp. 524–541, Nov. 2022, doi: 10.3322/caac.21754.

[8] Z. He *et al.*, "A review on methods for diagnosis of breast cancer cells and tissues," *Cell Proliferation*, vol. 53, no. 7. Blackwell Publishing Ltd, Jul. 01, 2020. doi: 10.1111/cpr.12822.

[9] D. Uzun Ozsahin, D. Ikechukwu Emegano, B. Uzun, and I. Ozsahin, "The Systematic Review of Artificial Intelligence Applications in Breast Cancer Diagnosis," *Diagnostics*, vol. 13, no. 1. 2023. doi: 10.3390/diagnostics13010045.

[10] A. Rodriguez-Ruiz *et al.*, "Stand-Alone Artificial Intelligence for Breast Cancer Detection in Mammography: Comparison With 101 Radiologists," *JNCI: Journal of the National Cancer Institute*, vol. 111, no. 9, pp. 916–922, Sep. 2019, doi: 10.1093/JNCI/DJY222.

[11] S. Pacilè, J. Lopez, P. Chone, T. Bertinotti, J. M. Grouin, and P. Fillard, "Improving breast cancer detection accuracy of mammography with the concurrent use of an artificial intelligence tool," *Radiol Artif Intell*, vol. 2, no. 6, pp. 1–9, Nov. 2020, doi: 10.1148/ryai.2020190208.

[12] X. Wang *et al.*, "Intelligent Hybrid Deep Learning Model for Breast Cancer Detection," *Electronics (Basel)*, vol. 11, no. 17, p. 2767, Sep. 2022, doi: 10.3390/electronics11172767.

[13] B. S. Abunasser, M. R. J. AL-Hiealy, I. S. Zaqout, and S. S. Abu-Naser, "Breast Cancer Detection and Classification using Deep Learning Xception Algorithm," *International Journal of Advanced Computer Science and Applications*, vol. 13, no. 7, 2022, doi: 10.14569/IJACSA.2022.0130729.

[14] L. Shen, L. R. Margolies, J. H. Rothstein, E. Fluder, R. McBride, and W. Sieh, "Deep Learning to Improve Breast Cancer Detection on Screening Mammography," *Sci Rep*, vol. 9, no. 1, p. 12495, Aug. 2019, doi: 10.1038/s41598-019-48995-4.

[15] H. Seo, L. Brand, L. S. Barco, and H. Wang, "Scaling multi-instance support vector machine to breast cancer detection on the BreaKHis dataset," *Bioinformatics*, vol. 38, no. Supplement_1, pp. 92–100, Jun. 2022, doi: 10.1093/bioinformatics/btac267.

[16] F. A. Spanhol, L. S. Oliveira, C. Petitjean, and L. Heutte, "A Dataset for Breast Cancer Histopathological Image Classification," *IEEE Trans Biomed Eng*, vol. 63, no. 7, pp. 1455–1462, Jul. 2016, doi: 10.1109/TBME.2015.2496264.