A Systematic Approach to Identify the Breast Cancer Grades in Histopathological Images Using Deep Neural Networks

S. H. S. Silva ^{#1}, T. M. K. K. Jinesena ^{#2}

[#]Department of Computer Science, University of Sri Jayewardenepura,
Gangodawila, 10250, Sri Lanka

¹hasinisandunika.usjp@gmail.com,

²kasun@sjp.ac.lk

Abstract- Breast cancer can be recognized as one of the most well-known and life-threatening cancers impacting women and this has been identified as the second most common cancer across the world. According to registered data, there were over 2 million newly reported cases in 2020. The Deep Convolutional Neural Network has been identified as one of the most dominant and powerful deep learning approaches involved in the analysis of visual imagination. There are many shreds of evidence that indicate the appropriateness of this in medical imaging including breast cancer detection, classification, and segmentation with higher accuracy rates. The main intent of the research is to develop an automated application that can determine the Nottingham Histologic Score of a given input histopathological image obtained from breast cancer or healthy tissues with DenseNet based architecture. Healthy or benign tissues are categorized as zero and cancerous tissues are categorized based on the grade obtained as one, two, or three. In this study, we were able to obtain more than 94% accuracy rates for each trained model including 2-predict, 3-predict, and 4-predict networks. Further, a desktop-based inference tool that allows us to perform breast cancer grading was also developed as a result of this study.

Index Terms- Breast Cancer Grading, Computer-Aided Diagnosis, Convolutional Neural Network, DenseNet, Histopathological Images, Nottingham Histologic Score

I. INTRODUCTION

Breast cancer is identified as the second most common cancer and this holds the highest mortality rate due to cancers among women worldwide. As per the UIUC organization, 2.3 million new cases were reported in 2020 [34]. Once the breast cancer cells are formed, that will lead to the spread of malignant cells to other parts of the body by making it life-threatening. Breast cancers are often found earlier when they are small and before they spread. Due to technical advancement, early detection, and procedures, the survival rate of breast cancer patients has significantly improved across the world within the last few years.

Breast cancer grade classification is a labor-intensive task and needs human experts and time to do the diagnoses. The grade of breast cancer tissue is calculated according to the way that cancer cells look under the microscope. For the laboratory breast cancer grade analysis purpose, it is best known that

specimens of the tissues with the affected areas are used. Histopathological images are extracted as a result of this. It is known that histopathological images can be extracted with different High-Power Field (HPF) values. But images with 40X HPF values are well defined and illustrate the key features. The classified grade is used to figure out what treatments work best.

Densely Connected Convolutional Network (DenseNet) is one of the discoveries in deep neural networks for visual object recognition. Deep Convolutional Neural Networks (DCNN) with DenseNet architecture can be identified as one of the most dominant and powerful deep learning approaches involved in the analysis of visual imagination that provide excellent performance in medical imaging including breast cancer detection, classification, and segmentation with higher accuracy rates of more than 90% [32] [33].

This research aims to introduce a new computer-aided diagnosis approach that can continue the existing laboratory grading procedure with a higher accuracy rate.

II. OBJECTIVE

The main intent of the research is to develop an application that can determine the grade of a given histopathological image of breast cancer or healthy tissue as 0-benign, 1, 2, or 3 using a DenseNet based DCNN.

III. RELATED WORKS

This explains the background and existing solutions related to the domain of breast cancer grade classification and past works related to transfer learning with DenseNet architecture in medical image analysis.

A. Breast Cancer Grade Classification

Cancers occur due to abnormal changes or mutations in the genes that are responsible for conducting the usual activities of the human body. As a result of this, cancer cells or malignant cells are formed. These created malignant cells can continue spreading without any control and this will lead to spreading cancer to other parts of the body. This indicates that breast cancer is caused due to uncontrolled growth of malignant cells inside the human body. Grading is a measurement that can be used to define the aggressiveness of breast cancer. There are different "scoring systems" introduced to ascertain the grade of breast cancer. The Nottingham Histologic Scoring system is considered

one of the best scoring systems among other scoring techniques [12]. Here the following features are considered by pathologists when determining the breast cancer grade;

1. The amount of gland formation
2. The nuclear features
3. The mitotic count

Above each feature is scored from 1-3, and these scores are added to obtain the final grade that the total score ranges between 3-9. The final total score is determined as follows;

1. Grade 1 tumors hold a total score of 3-5
2. Grade 2 tumors hold a total score of 6-7
3. Grade 3 tumors hold a total score of 8-9

H. Peiris et al. clarified the impact of Nottingham grade on breast cancer-specific survival and recurrence-free survival of operable breast cancer patients by evaluating the value of Nottingham grade in the Sri Lankan setting [20]. They have used Kaplan-Meier and Cox regression models for survival analysis. The results they have obtained were a total of 742 (grade 1-12%, grade 2-45%, grade 3-43%) patients, and their breast cancer-specific survival was 94% for grade 1, 80% for grade 2, and 72% for grade 3 with p -value <0.001 and also their recurrence-free survival was 86% for grade 1, 75% for grade 2, and 67% for grade 3 with p -value $=0.001$. Pathological prognostic factors in breast cancer, I. C. W. Elston et al. described that, since 1973 in the United Kingdom, histological grades were evaluated from 1831 breast cancer patients [21]. Here patients with grade 1 tumors had significantly better survival than those with grade 2 and 3 tumors with a p -value <0.0001 . S. Pal et al. analyzed the grade of breast carcinoma on cytology by using Robinson's grading system and correlating it with Elston's modified Bloom Richardson histological grading system [30]. They have achieved a concordance rate between cytological and histological grades of around 78%. According to the results, they were able to obtain a coefficient of correlation between cytological and histological grades of around 0.804 with a p -value <0.001 .

B. Transfer Learning with DenseNet Architecture in Medical Image Analysis

This includes some of the past research works that have been conducted involving transfer learning with DenseNet architecture in medical image analysis.

M. Talo has introduced two approaches with pre-trained ResNet-50 and DenseNet161 architectures to classify color and gray-scale histopathological images [23]. They have achieved more than 97% accuracy for both pre-trained models. X. Wu et al. introduced an approach that is applied to extract gait features obtained from Gait Energy Images [24]. They have extracted gait features through DenseNet based transfer learning and K nearest neighbor classifier (KNN) to classify and identify people. The average recognition rate was around 98%. Q. Cai et al. proposed a deep neural network called SE-DenseNet by combining both DenseNet architecture and the Squeeze-and-Excitation block [25]. They have used a twice fine-tuning method to classify the

breast mass. The introduced model achieved higher performance on the BCDR dataset with an accuracy rate of around 98%. S. Minaee et al. introduced an approach to identify COVID-19 disease by applying transfer learning on ResNet18, ResNet50, SqueezeNet, and DenseNet-121 pre-trained architectures [26]. To continue this, they have reserved 2000 chest X-ray images for the training purpose and 3000 chest X-ray images for the testing purpose. During the evaluation process, they were able to obtain a sensitivity rate of 98% ($\pm 3\%$) and a specificity rate of around 90% for most of the models. X. Xu et al. researched to analyze fundus images in two phases, training the DenseNet model from scratch, and in the second phase, they have applied transfer learning to obtain a fine-tuned network [27]. Y. Celik et al. presented a transfer learning approach for the detection of invasive ductal carcinoma by training both ResNet-50 and DenseNet-161 pre-trained architectures [28]. At the evaluation process, they achieved an F-score of 92.38%, a balanced accuracy value of 91.57% for the DenseNet-161 model, and an F-score of 94.11%, a balanced accuracy value of 90.96% for the ResNet-50 architecture. F. Imrie et al. combined DenseNet pre-trained architecture with a transfer learning approach that was used to create a composition of protein family-specific models [29]. S. H. Wang and Y. D. Zhang introduced a new transfer learning approach by comparing and parameter tuning for all DenseNet 121, 169, and 201 architectures for multiple sclerosis classification [31]. They observed that DenseNet201-D architecture achieved the best performance with 98.27 ± 0.58 sensitivity, 98.35 ± 0.69 specificities, and 98.31 ± 0.53 accuracy.

C. Breast Cancer Grade Classification in Medical Image Analysis

Some of the previous works related to the objectives are defined here.

T. Mahmood et al. introduced a multistage mitotic cell detection technique based on the Faster region convolutional neural network and DCNN [1]. They have used two separate datasets called ICPR 2012, ICPR 2014 and achieved a precision of 0.876, a recall of 0.841, an F1-score of 0.858 for the ICPR 2012 dataset, and a precision of 0.848, a recall of 0.583, an F1-score of 0.691 for the ICPR 2014 dataset. Another method for mitosis detection was presented by T. Mahmood et al. using the Support Vector Machine (SVM) approach to classify the candidate cells into mitotic and amitotic cells [3]. They have used the MITOS dataset and achieved a precision of 0.80, a recall of 0.90, an F-score of 0.85. H. Wang et al. implemented a new method by combining both convolutional neural networks and handcrafted features (morphology, color, and texture features) [6]. This detects the mitotic elements in histopathological images. M. A. Aswathy et al. introduced another approach to detect breast cancer using digital histopathology images by applying image processing techniques and discussed the future possibilities [4]. D. C. Cireşan et al. implemented a deep max-pooling convolutional neural network model to identify mitotic elements in histopathological images [2]. A. Tashk et al. contributed to constructing an algorithm including Complete Local Binary Patterns and other features to detect mitotic elements in histopathological images [5]. T. X.

Jian et al. mainly applied several classification techniques to distinguish mitotic cells from healthy normal cells [7]. S. Rao has introduced a method involving a region-based convolutional neural network to identify the mitotic elements in histopathological images [8]. A. Paul et al. presented an approach to detect mitotic elements from histopathological images whereas applying cell segmentation and classification of segmented cells as mitotic and amitotic using the random forest classifier [22].

After analyzing the previous works, the main contribution of the proposed solution is to fill the knowledge gap of the existing systems with a novel approach.

IV. METHODOLOGY

Below mentioned are the major tasks that had to be carried out during our research work.

- **Data preprocessing**
To complete the data preprocessing task, several mechanisms had to be applied and they are discussed under data preprocessing.
 - **Model building and evaluation**
After completing the dataset preparation, built three DCNN models with DenseNet architecture as,
 - 2-predict: classifies benign-0 and malignant-1
 - 3-predict: classifies grades into grade 1, grade 2, and grade 3
 - 4-predict: classifies as 0-benign, grade 1, grade 2, and grade 3
- Apart from that built another classical DCNN model to implement the 3- predict model.
- **Implementation of the inference tool**
To visualize the results obtained as well as to make the models available for use, an inference tool was built.

A. Data Preprocessing

Initially, we received a graded (1, 2, and 3) dataset [11] with different HPF values (4X, 10X, 20X, 40X), resolutions, and image formats. To complete the data preprocessing task, we used the image dataset with 40X HPF value and increased the received datasets by applying several data augmentation techniques including image cropping, rotating, etc. Later we were able to create an RGB image dataset for each grade with resolution 700*128*128*3. Again, we have used another dataset [14] to train the 2-predict model and after applying preprocessing techniques on the received benign 40X HPF dataset, we have created a dataset with image resolution 700*128*128*3. For 2-predict model implementation, the malignant dataset was created by applying the random sampling technique on the dataset of the 3-predict model. Dataset of the 4-predict model was created by combining both the benign image dataset of the 2-predict model and the dataset of the 3-predict model. Table 1 defines the resolutions of the training and testing datasets used in each model.

TABLE I: TABULAR REPRESENTATION OF THE DATASETS.

Model	Training Dataset	Testing Dataset
2-predict	600*128*128*3 for 0 600*128*128*3 for 1	100*128*128*3 for 0 100*128*128*3 for 1
3-predict	600*128*128*3 for 1 600*128*128*3 for 2 600*128*128*3 for 3	100*128*128*3 for 1 100*128*128*3 for 2 100*128*128*3 for 3
4-predict	600*128*128*3 for 0 600*128*128*3 for 1 600*128*128*3 for 2 600*128*128*3 for 3	100*128*128*3 for 0 100*128*128*3 for 1 100*128*128*3 for 2 100*128*128*3 for 3

B. Model Building and Evaluation

To build each model, the following key aspects were considered;

Grade classification was carried out by considering two scenarios. In the first scenario, a single neural network model was implemented to classify the histopathological images as benign-0, grade 1, grade 2, and grade 3. In the second scenario, two deep neural network models were trained to achieve the same objective as in the first scenario whereas implementing the initial model to classify the histopathological image as benign or malignant and at the second step, to classify the grade based on the prediction obtained at the first step. As in Fig. 1, the initial model classifies the input histopathological image as benign-0 or malignant-1 and if it is malignant, obtains the prediction from the second model as 1, 2, or 3 otherwise the grade will be remaining as zero.

Fig. 1. Logical Overview of the Proposed Solution.

As per the research objectives that were mentioned in previous sections, this research is targeted at building a DCNN with transfer learning to predict the grades (0-benign, 1, 2, and 3) of histopathological images. Below define the architectures of both classical, transfer learning models, and the process of the evaluation carried out to achieve the above-defined goal.

1. Classical DCNN Model Implementation

This was applied to implement the three predict model for 128*128 gray-scaled image dataset and achieved a less amount of test accuracy rate. Due to the lack of a higher accuracy rate and the other corresponding results obtained, we decided to implement the model involving the transfer learning technique. The classical DCNN model is a composition of,

- Two convolutional layers with ReLU operation
- Two max-pooling layers
- Flatten layer
- Fully connected layer with the tanh operation
- Fully connected layer with softmax operation (last layer)

Layer (type)	Output Shape	Param #
conv2d_4 (Conv2D)	(None, 42, 42, 10)	260
max_pooling2d_4 (MaxPooling2)	(None, 21, 21, 10)	0
conv2d_5 (Conv2D)	(None, 9, 9, 4)	1004
max_pooling2d_5 (MaxPooling2)	(None, 8, 8, 4)	0
flatten_2 (Flatten)	(None, 256)	0
dense_4 (Dense)	(None, 6)	1542
dense_5 (Dense)	(None, 3)	21
Total params: 2,827		
Trainable params: 2,827		
Non-trainable params: 0		

Fig. 2. The Architecture of the Classical DCNN Model.

Fig. 2 illustrates the architecture of the classical DCNN model. This was implemented with Google Collaboratory using Keras, TensorFlow, sklearn, and OpenCV.

2. Model Implementation with Transfer Learning

This was applied to implement 2-predict, 3-predict, and 4-predict models for 128*128 RGB image datasets and achieved better test accuracy rates than the classical model. This model is a composition of,

- Pre-trained DenseNet-201 weights
- 2D global average pooling layer
- Dropout layer with 0.5 probability
- Batch normalization layer
- Densely connected layer with softmax operation

Layer (type)	Output Shape	Param #
densenet201 (Functional)	(None, 4, 4, 1920)	18321984
global_average_pooling2d (Gl	(None, 1920)	0
dropout (Dropout)	(None, 1920)	0
batch_normalization (BatchNo	(None, 1920)	7680
dense (Dense)	(None, 2)	3842
Total params: 18,333,506		
Trainable params: 18,100,610		
Non-trainable params: 232,896		

Fig. 3. The Architecture of the Transfer Learning Model.

Fig. 3 illustrates the architecture of the transfer learning model. This was implemented with Spyder and Anaconda framework and the libraries, dependencies used are TensorFlow, Keras, sklearn, and OpenCV.

Apart from that, a Flask Application Program Interface with Anaconda framework was developed to obtain the prediction by considering each case.

3. Evaluation

This was performed on the reserved test datasets of each model. For each model, there were around 14% of allocated histopathological images for testing purposes out of all images.

The evaluation was carried out by changing parameters including batch size, dropout probability, and the number of epochs. But, all four models provided significant responses only for the varied number of epochs rather than batch size and dropout probability. Mainly we have performed evaluation operations on test datasets including test accuracy rate calculation, confusion matrix illustration, and classification reports.

C. Implementation of the Inference Tool

This was carried out to visualize the results obtained and to make the trained models available for use. The inference tool was implemented as a desktop graphical user interface application and has added some basic graphical user interface functions such as uploading an image, viewing results, saving an image with the classification, searching a record by ID, deleting a record by ID, converting a record into PDF format with a suitable database connection.

The summarized architecture of the proposed solution is illustrated in Fig. 4.

Fig. 4. Summarized Architecture of the Proposed Solution.

This was implemented with Apache NetBeans with Maven dependencies and MongoDB was used as the database. Apart from that, an Application Program Interface integration was carried out to obtain the predictions by considering each case as 1 and 2.

V. RESULTS AND DISCUSSION

To analyze the results generated by each model, the main method that followed was the test data-based results generation. The following includes the results obtained and the comparison between the obtained results.

A. Two-Predict Model

Fig. 5. Loss vs Epochs for Two-Predict Model

Fig. 5. shows the loss vs epochs received by applying transfer learning during the process of training for the two-predict model with 20 epochs. As per this, curves of both validation and training are getting converged with the number of epochs. We could see that the model attains to the minimum loss rate at around the 4th epoch and continues the same until the 20th epoch. This indicates that the model shows a good fit with the training dataset.

Fig. 6. Normalized Confusion Matrix for Two-Predict Model.

Fig. 6. shows the normalized confusion matrix received by applying transfer learning from the testing dataset for the two-predict model with 20 epochs.

```

in [13]: print(classification_report( true_y.argmax(axis=-1), y_predict))
          precision    recall  f1-score   support

     0       1.00      1.00      1.00       100
     1       1.00      1.00      1.00       100

 accuracy          1.00          1.00          1.00          200
 macro avg          1.00          1.00          1.00          200
 weighted avg          1.00          1.00          1.00          200
  
```

Fig. 7. Classification Report for Two-Predict Model.

Fig. 7. shows the classification report received by applying transfer learning from the testing dataset for the two-predict model with 20 epochs. According to this result, precision, recall, and F1-score have obtained the highest values as can be obtained. Also, the accuracy received for the test dataset is 100%.

B. Three-Predict Model

Fig. 8. Accuracy vs Epochs of Classical DCNN Model.

Fig. 8. shows the accuracy vs epochs received during the process of training for the classical DCNN three-predict model with 100 epochs. This was trained for a 128*128 gray-scaled image dataset and achieved the test accuracy rate of 54.667%. As per this illustration, the graph is getting diverged whereas increasing the number of epochs. This implies that the model does not fit the dataset as expected.

Fig. 9. Loss vs Epochs for Three-Predict Model.

Fig. 9. shows the loss vs epochs received by applying transfer learning during the process of training for the three-predict model with 30 epochs. This model was previously trained for 20 epochs. As per that, the trained model with 30 epochs tends to converge with the number of epochs rather than the model with 20 epochs. This indicates that the model with 30 epochs shows a better fit with the training dataset than the model with 20 epochs.

Fig. 10. Normalized Confusion Matrix for Three-Predict Model.

Fig. 10. shows the normalized confusion matrix received by applying transfer learning from the testing dataset for the three-predict model with 30 epochs.

```
In [47]: print(classification_report( true_y.argmax(axis=-1), y_predict))
```

	precision	recall	f1-score	support
0	0.98	0.93	0.95	100
1	0.96	0.94	0.95	100
2	0.92	0.98	0.95	100
accuracy			0.95	300
macro avg	0.95	0.95	0.95	300
weighted avg	0.95	0.95	0.95	300

Fig. 11. Classification Report for Three-Predict Model.

Fig. 11. shows the classification report received by applying transfer learning from the testing dataset for the three-predict model with 30 epochs. As this was trained for 20 epochs previously, the trained model with 30 epochs holds maximum values for precision, recall, and F1-score. The test

accuracy rates of trained models with 20 and 30 epochs are 94.33% and 94.99% respectively.

C. Four-Predict Model

Fig. 12. Loss vs Epochs for Four-Predict Model.

Fig. 12. shows the loss vs epochs received by applying transfer learning during the process of training for the four-predict model with 25 epochs. As this was previously trained for 20 and 30 epochs, the trained model with 25 epochs shows much convergence between validation and training curves. This indicates that the model with 25 epochs shows a good fit with the training dataset.

Fig. 13. Normalized Confusion Matrix for Four-Predict Model.

Fig. 13. shows the normalized confusion matrix received by applying transfer learning from the testing dataset for the four-predict model with 25 epochs.

```
In [61]: print(classification_report( true_y.argmax(axis=-1), y_predict))
```

	precision	recall	f1-score	support
0	1.00	0.94	0.97	100
1	0.94	0.96	0.95	100
2	0.97	0.92	0.94	100
3	0.91	0.99	0.95	100
accuracy			0.95	400
macro avg	0.95	0.95	0.95	400
weighted avg	0.95	0.95	0.95	400

Fig. 14. Classification Report for Four-Predict Model.

Fig. 14. show the classification report received by applying transfer learning from the testing dataset for the four-predict model with 25 epochs. As per previous results for 20 and 30 epochs, the trained model with 25 epochs holds maximum values for precision, recall, and F1-score. The test accuracy rates of trained models with 20, 25, and 30 epochs are 92.75%, 95.2499%, and 94.7499% respectively.

D. Model Type (Case 1 and 2)

```
Size of the output testing data set: (400, 4)
Size of the input testing data set: (400, 128, 128, 3)
13/13 - 63s - loss: 0.0602 - accuracy: 0.9525
True count: 372
False count: 28

Testing accuracy score for case 1: 95.24999856948853 %
Testing accuracy score for case 2: 93.0 %
```

Fig. 15. Test Accuracy Rates of Case 1 and 2.

Fig. 15 shows the test accuracy rates received by performing case 1 and 2 operations for the test dataset of the 4-predict model. This was implemented with the 2-predict model with 20 epochs, the 3-predict model with 30 epochs, and the 4-predict model with 25 epochs.

According to the results obtained,

- Two-predict model: When Fig. 5 is considered, we could see that the graph is getting converged whereas increasing the number of epochs. Fig. 6 and 7 also evidently express the appropriateness of the model by providing a 100% accuracy rate for the test dataset.
- Three-predict model: Fig. 8 shows that the proposed classical DCNN model is not enough to interpret the training dataset as expected. When Fig. 9 is considered, we could see that we can obtain a good fit by increasing the number of epochs up to 30. Fig. 10 and 11 also provide some evidence, providing a 94.99% accuracy rate for the test dataset for 30 epochs.
- Four-predict model: By considering Fig. 12, we can conclude that we can increase the fitness of the model by increasing the number of epochs up to 25. Fig. 13 and 14 also show a higher test data accuracy rate around 95.249% for 25 epochs.
- Model type (case 1 and 2) comparison: According to Fig. 15, can conclude that both approaches are good enough, and also both are applicable for the desktop graphical user interface prediction purpose.

By referring to the above results, we can conclude that it is appropriate to train the two-predict model up to 20 epochs, the three-predict model up to 30 epochs, and the four-predict model up to 25 epochs. Furthermore, both model types (case 1 and 2) are applicable for the desktop graphical user interface prediction purpose.

VI. CONCLUSION

In this research,

- Three DCNN based models with transfer learning have been implemented to classify the grades of breast cancer.
- These DCNN models have been trained and tested with histopathological image datasets with the 40X HPF value.
- 2-predict, 3-predict, and 4-predicts models were able to obtain test accuracy rates of 100%, 94.999%, and 95.2499% respectively.
- This research confirms the successfulness of using transfer learning with DenseNet architecture whereas providing test accuracy rates of more than 94% for all three trained models.
- Moreover, a desktop application was developed to infer the solution.

Ultimately, accuracy improvements, implementation of the developed system as a web-based application could be identified as works that could be carried out as further improvements of this research.

VII. REFERENCES

- [1] T. Mahmood, M. Arsalan, M. Owais, M. B. Lee, and K. R. Park, "Artificial intelligence based mitosis detection in breast cancer histopathology images using Faster R-CNN and deep CNNs," *J. Clin. Med.*, vol. 9, no. 3, p. 749, 2020.
- [2] D. C. Cireřan, A. Giusti, L. M. Gambardella, and J. Schmidhuber, "Mitosis detection in breast cancer histology images with deep neural networks," *Med. Image Comput. Comput. Assist. Interv.*, vol. 16, no. Pt 2, pp. 411–418, 2013.
- [3] T. Mahmood, S. Ziauddin, A. R. Shahid, and A. Safi, "Mitosis detection in breast cancer histopathology images using statistical, color and shape-based features," *J. Med. Imaging Health Inform.*, vol. 8, no. 5, pp. 932–938, 2018.
- [4] M. A. Aswathy and M. Jagannath, "Detection of breast cancer on digital histopathology images: Present status and future possibilities," *Inform. Med. Unlocked*, vol. 8, pp. 74–79, 2017.
- [5] A. Tashk, M. S. Helfroush, H. Danyali, and M. Akbarzadeh-jahromi, "Automatic detection of breast cancer mitotic cells based on the combination of textural, statistical and innovative mathematical features," *Appl. Math. Model.*, vol. 39, no. 20, pp. 6165–6182, 2015.
- [6] H. Wang et al., "Mitosis detection in breast cancer pathology images by combining handcrafted and convolutional neural network features," *J. Med. Imaging (Bellingham)*, vol. 1, no. 3, p. 034003, 2014.
- [7] T. X. Jian, M. Nazahah, M. M. Yusoff, and A. R. K. Shakir, "Mitotic cell detection in breast histopathology image: A review," *J. Phys. Conf. Ser.*, vol. 1372, p. 012026, 2019.
- [8] S. Rao, "Mitosis-rnn: Mitotic figure detection in breast cancer histopathology images using region based convolutional neural networks," 2018.

- [9] T. K. ten Kate, J. A. Beliën, A. W. Smeulders, and J. P. Baak, "Method for counting mitoses by image processing in Feulgen stained breast cancer sections," *Cytometry*, vol. 14, no. 3, pp. 241–250, 1993.
- [10] M. Veta et al., "Assessment of algorithms for mitosis detection in breast cancer histopathology images," *Med. Image Anal.*, vol. 20, no. 1, pp. 237–248, 2015.
- [11] H. Bolhasani, E. Amjadi, M. Tabatabaeian, and S. J. Jassbi, "A histopathological image dataset for grading breast invasive ductal carcinomas," *Inform. Med. Unlocked*, vol. 19, no. 100341, p. 100341, 2020.
- [12] "Staging & Grade," Jhu.edu. [Online]. Available: <https://pathology.jhu.edu/breast/staging-grade/>. [Accessed: 12-Jun-2021].
- [13] F. A. Spanhol, L. S. Oliveira, C. Petitjean, and L. Heutte, "Breast cancer histopathological image classification using Convolutional Neural Networks," in 2016 International Joint Conference on Neural Networks (IJCNN), 2016.
- [14] 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, 2016.
- [15] A. Fernando, U. Jayarajah, S. Prabashani, E. A. Fernando, and S. A. Seneviratne, "Incidence trends and patterns of breast cancer in Sri Lanka: an analysis of the national cancer database," *BMC Cancer*, vol. 18, no. 1, 2018.
- [16] Deep Learning by Ian Goodfellow, Yoshua Bengio and Aaron Courville published by MIT Press, 2016.
- [17] Stanford University's Course — CS231n: Convolutional Neural Network for Visual Recognition by Prof. Fei-Fei Li, Justin Johnson, Serena Yeung.
- [18] Y. Lecun, L. Bottou, Y. Bengio, and P. Haffner, "Gradient-based learning applied to document recognition," *Proc. IEEE Inst. Electr. Electron. Eng.*, vol. 86, no. 11, pp. 2278–2324, 1998.
- [19] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "ImageNet classification with deep convolutional neural networks," *Commun. ACM*, vol. 60, no. 6, pp. 84–90, 2017.
- [20] H. Peiris, L. Mudduwa, N. Thalagala, and K. Jayatilake, "The value of Nottingham grade in breast cancer re-visited in the Sri Lankan setting," *Malays. J. Pathol.*, vol. 39, no. 2, pp. 141–148, 2017.
- [21] C. W. Elston and I. O. Ellis, "Pathological prognostic factors in breast cancer. I The value of histological grade in breast cancer: experience from a large study with long-term follow-up," *Histopathology*, vol. 19, no. 5, pp. 403–410, 1991.
- [22] A. Paul and D. P. Mukherjee, "Mitosis detection for invasive breast cancer grading in histopathological images," *IEEE Trans. Image Process.*, vol. 24, no. 11, pp. 4041–4054, 2015.
- [23] M. Talo, "Automated classification of histopathology images using transfer learning," *Artif. Intell. Med.*, vol. 101, no. 101743, p. 101743, 2019.
- [24] X. Wu, T. Yang, and Z. Xia, "Gait recognition based on densenet transfer learning," *Ijset.net*. [Online]. Available: <https://www.ijset.net/journal/2458.pdf>. [Accessed: 12-Jun-2021].
- [25] Q. Cai, X. Liu, and Z. Guo, "Identifying architectural distortion in mammogram images via a SE-DenseNet model and twice transfer learning," in 2018 11th International Congress on Image and Signal Processing, BioMedical Engineering and Informatics (CISP-BMEI), 2018.
- [26] S. Minaee, R. Kafieh, M. Sonka, S. Yazdani, and G. Jamalipour Soufi, "Deep-COVID: Predicting COVID-19 from chest X-ray images using deep transfer learning," *Med. Image Anal.*, vol. 65, no. 101794, p. 101794, 2020.
- [27] X. Xu, J. Lin, Y. Tao, and X. Wang, "An improved DenseNet method based on transfer learning for fundus medical images," in 2018 7th International Conference on Digital Home (ICDH), 2018.
- [28] Y. Celik, M. Talo, O. Yildirim, M. Karabatak, and U. R. Acharya, "Automated invasive ductal carcinoma detection based using deep transfer learning with whole-slide images," *Pattern Recognit. Lett.*, vol. 133, pp. 232–239, 2020.
- [29] F. Imrie, A. R. Bradley, M. van der Schaar, and C. M. Deane, "Protein family-specific models using deep neural networks and transfer learning improve virtual screening and highlight the need for more data," *J. Chem. Inf. Model.*, vol. 58, no. 11, pp. 2319–2330, 2018.
- [30] S. Pal and M. L. Gupta, "Correlation between cytological and histological grading of breast cancer and its role in prognosis," *J. Cytol.*, vol. 33, no. 4, pp. 182–186, 2016.
- [31] S.-H. Wang and Y.-D. Zhang, "DenseNet-201-based deep neural network with composite learning factor and precomputation for multiple sclerosis classification," *ACM trans. multimed. comput. commun. appl.*, vol. 16, no. 2s, pp. 1–19, 2020.
- [32] Y. Yari, T. V. Nguyen, and H. T. Nguyen, "Deep learning applied for histological diagnosis of breast cancer," *IEEE Access*, vol. 8, pp. 162432–162448, 2020.
- [33] A. Rakhlin, A. Shvets, V. Iglovikov, and A. A. Kalinin, "Deep convolutional neural networks for breast cancer histology image analysis," in *Lecture Notes in Computer Science*, Cham: Springer International Publishing, 2018, pp. 737–744.
- [34] "Breast cancer," Uicc.org. [Online]. Available: <https://www.uicc.org/what-we-do/thematic-areas-work/breast-cancer>. [Accessed: 12-Jun-2021].