Revolutionising Breast Cancer Diagnosis: Deep Learning Algorithms for Accurate and Early Detection

Chandan Kumar Sangewar
School of Computer Science and Engineering
Vellore Institute of Technology, Chennai, India
chandankumar.sangewar2020@vitstudent.ac.in

Chinmay Pagey
School of Computer Science and Engineering
Vellore Institute of Technology, Chennai, India
chinmay.pagey2020@vitstudent.ac.in

Aman Kumar
School of Computer Science and Engineering
Vellore Institute of Technology, Chennai, India
aman.kumar2020@vitstudent.ac.in

Kamna V
School of Computer Science and Engineering
Vellore Institute of Technology, Chennai, India
kamna.v2020@vitstudent.ac.in

Abstract*— One of the biggest killers of women around is breast cancer. Effective treatment and higher survival rates depend on early discovery. Recent years have demonstrated promising results for using deep learning algorithms to accurately identify breast cancer in medical photos. Convolutional neural networks (CNNs), deep belief networks (DBNs), and recurrent neural networks (RNNs) are only few of the state-of-the-art deep learning algorithms reviewed in this research for their use in breast cancer diagnosis (RNNs). We emphasise the advantages of using these algorithms in clinical practise while discussing the difficulties that arise when using deep learning to the detection of breast cancer. Our findings show that deep learning models may help radiologists save lives by increasing their efficiency and improving the accuracy with which they diagnose breast cancer. Deep learning algorithms provide a potential approach to the early identification and tailored treatment of breast cancer by capitalising on massive datasets and cutting-edge computational capacity. Finally, we describe where the area of deep learning in breast cancer detection may go from here to better address the limits and develop the discipline.

Keywords: Breast Cancer, Women, Diagnose, Deep Learning, RNN, VGG16, VGG19

I.PROJECT OVERVIEW

Invasive ductal carcinoma (IDC), which begins in a milk duct but rapidly spreads to surrounding breast tissue, is the most common kind of invasive breast cancer. The purpose of this study is to use a convolutional neural network trained using breast tissue images to detect the presence of invasive ductal carcinoma. To check for cancer, a pathologist will examine a tissue sample mounted on a glass slide. A pathologist has to look at a lot of tissue that may not have cancer in it before they can find the spots where it is. The digitization of glass slides presents an opportunity for computer vision to enhance the efficiency of the pathologist's workflow, the accuracy of diagnoses, and ultimately, patient outcomes.

1.Problem Statement

In females, breast cancer and skin cancer are the two most prevalent forms of the disease. Symptoms of breast cancer are not usually noticeable. Invasive ductal carcinoma is the most common kind of breast cancer and accounts for 80% of diagnosis. If an abnormality is found on a mammogram, further testing may involve a biopsy, in which a small sample or the whole lump is surgically removed and studied under a microscope by a pathologist. More than one million women a

year in the United States have a breast biopsy. Around 20% of these biopsies reveal breast cancer. There will be an estimated 268,600 new cases of invasive breast cancer in women and 2,670 new cases in males in the United States this year, according to the American Cancer Society. An estimated 42,260 individuals will lose their lives to breast cancer this year (about 41,760 women and 500 males). A convolutional neural network trained on samples of both healthy and diseased tissue might potentially improve the classification and understanding of the images by physicians. Doctors may be notified of the most pressing cases, and patients could benefit from a quicker diagnosis and treatment.

2. Metrics

The accuracy of the model may be calculated by dividing the number of accurate predictions by the total number of predictions, or by calculating the f1-scores of both classes (average of precision and recall). Since this is a medical binary classification problem, it could also be evaluated in terms of the sensitivity (true positives/(true positives + false negatives)) and specificity (true negatives/(true negatives + false positives)) of both classes; however, the positive class would be more relevant because ignoring a malignant case is far worse than treating a non-malignant one. Yet, if a benign instance is incorrectly diagnosed as IDC-positive, the patient may be subjected to invasive and costly surgery to remove potential malignant areas.

II. ANALYSIS

1. Data Exploration

1.64GB breast histopathology image dataset by Kaggle:

https://www.kaggle.com/paultimothymooney/breast-histopathology-images

Download Link:
https://www.kaggle.com/paultimothymooney/breast-histopathology-images/downloads/breast-hi stopathology-images.zip/1

Breast cancer (BCa) specimens were scanned as whole mount slides 162 times at 40x magnification to create the original dataset. Using it, we were able to draw 277,524 square-foot patches (198,738 IDC negative and 78,786 IDC positive).

The original files are located here:
http://gleason.case.edu/webdata/jpi-dl-tutorial/IDC_regular_ps50_idx5.zip

Citation:
https://www.ncbi.nlm.nih.gov/pubmed/27563488 http://spie.org/Publications/Proceedings/Paper/10.1117/12.2043872

2. Exploratory Visualisations

277524 total images - 198,738 IDC negative images and 78,786 IDC positive images which is about a 2.5 : 1 distribution ratio and is a huge class imbalance

Fig 1. Visualisation of Class Distribution

2. Algorithms and Techniques

A convolutional neural network is used, which is considered the state-of-the-art for many classification and other image processing tasks. It need a massive amount of training data, and the one given by Kaggle is a good match for the task. Data that is spatially organised, such the values of pixels in a picture, may be

employed by regularised multilayered perceptrons like those found in convolutional neural networks. They achieve this by using the model's kernel/filter trainability to extract features (the core building block of a CNN). The matrix of the feature map is then converted into a vector. We combine these features with properly connected layers so that the model can make sense of them. The model uses a final fully linked layer with 2 output units and a softmax activation mechanism to determine whether a picture is IDC positive or IDC negative. In order to reduce the amount of false positives, the model may use a threshold, which in turn increases the number of false negatives.

The following can be tuned to optimise the classifier:

- Classification Threshold / Class Weights
- Preprocessing Parameters
  - Preprocessing Function
  - Data Augmentation Parameters
- Neural Network Architecture
  - Number of Layers
  - Layer Types(convolution, pooling, dropout, dense/Fully connected)
  - Number of Filter/Units
- Training Parameters
  - Epochs
  - Batch Size
  - Optimiser
  - Learning Rate

During training, the datagen class pulls data from defined folders in small processed batches to load into RAM and feeds it into the model for training, validation, or testing. The training is done with mini-batch gradient descent.

III. METHODOLOGY

1.Data Processing

Images were duplicated into three different folders for training, validation, and testing purposes. The underrepresented IDC-positive class prompted an improved replication. The InceptionV3 preprocessing tool in keras was used on all the photographs that were used as input to the model. To get this effect, we multiply the pixel values by 2, divide by 255, and then remove half. It's quite different from the preprocessing of models like Resnet and VGG. Neither RGB to BGR nor mean subtraction is a real thing. By employing this technique, there shouldn't be any pixels with values other than -1 and 1. As the final image size of 100 was twice as large as the original, it was chosen for use in a transfer learning scenario. This resolution yielded the best agreement with the base model.

2. Implementation

Steps:

1. One, all required packages were successfully imported and depended upon. a. Keras is now integrated into Tensorflow 2.0, although this study only employed version 1.14.0.
2. The data was uploaded and unzipped, and then the extraction process began. With an API key, data may be imported straight from Kaggle.
3. Third, we copied the photos into two subfolders within the "all images dir" folder: one for the "positive" class and one for the "negative" class. a. The photographs are retrieved from the class folder within the patient ID folder and placed in the all images dir class folder.

4. Create a Dataframe Including Picture Details - Dataframe containing Image IDs, Patient IDs, and Target IDs (IDC positive or IDC negative as 1 or 0 respectively)

5. Dataframe separated into train, valid, and test pictures

6. Separate Images for Training, Testing, and Validation - Dataframe pictures separated into train, valid, and test directories.

7. Patient ID > Class > Extract Images > Train > Class > Extract Images > Valid > Class > Extract Images > Test > Class > Extract Images

8. To combat prejudice stemming from class imbalance, we enhance duplicate photos of the underrepresented class (the IDC positive class) and place them in the training set's folder.

9. Train, legitimate, and test batches are generated using the keras datagen class a. b makes use of the inception preprocessing function. The default batch size is 500 c. Don't mix up your test batches.

10. Construct and train a benchmark CNN model. The positive class may be given additional weight in the model by increasing its class weight.

11. Benchmark Model Evaluation: Putting the Model to the Test

12. Precision, recall, F1 score, and overall accuracy determined.

13. The Matrix of Confusion

14. A bigger picture size must be supplied into the benchmark model first, and the model's weights must be inserted into a transfer learning model if one is desired; this is because the transfer learning model did not outperform a benchmark CNN.

15. If you're using the Pillow library, and you're having trouble opening pictures, try setting " LOAD TRUNCATED IMAGES " to "Yes." The following line of code will do this: With IMAGE FILE.LOAD TRUNCATED IMAGES = True

16. Put Tensorflow to work as Keras's backend. Define the environment variable " KERAS BACKEND " with this line to override the config file's selection of backend if it is not Tensorflow.

KERAS_BACKEND = tensorflow python -c "from keras import backend"

When importing packages from Keras, upload it from keras directly and not from tensorflow.python.keras unless Tensorflow 2.0 is being used.

3. Refinement

Because of the significant class imbalance within the data, augmented pictures of the IDC positive class were created. Several changes were made to the settings of data augmentation, but the result was only a little rise in precision. The accuracy of CNNs with the same padding in their convolutional layers saw a significant boost. Picture size seems to have a significant effect on precision; I found that 100 pixels on a side gave me the best results. When the dropout rate increased, overfitting to the training data was delayed. Additional 2D convolutional layers were more effective than additional filters. Several iterations of batch sizes were tried, and it was found that 500 was optimal for making progress towards convergence rapidly. Different transfer learning models integrating the weights from the benchmark model were attempted and performed far worse.

IV. RESULTS

1. Model Evaluation and Validation

During development, a stratified validation set and testing set was used to evaluate the model. The final architecture and hyperparameters were chosen because they performed the best among the tried combinations.

Model Architecture:

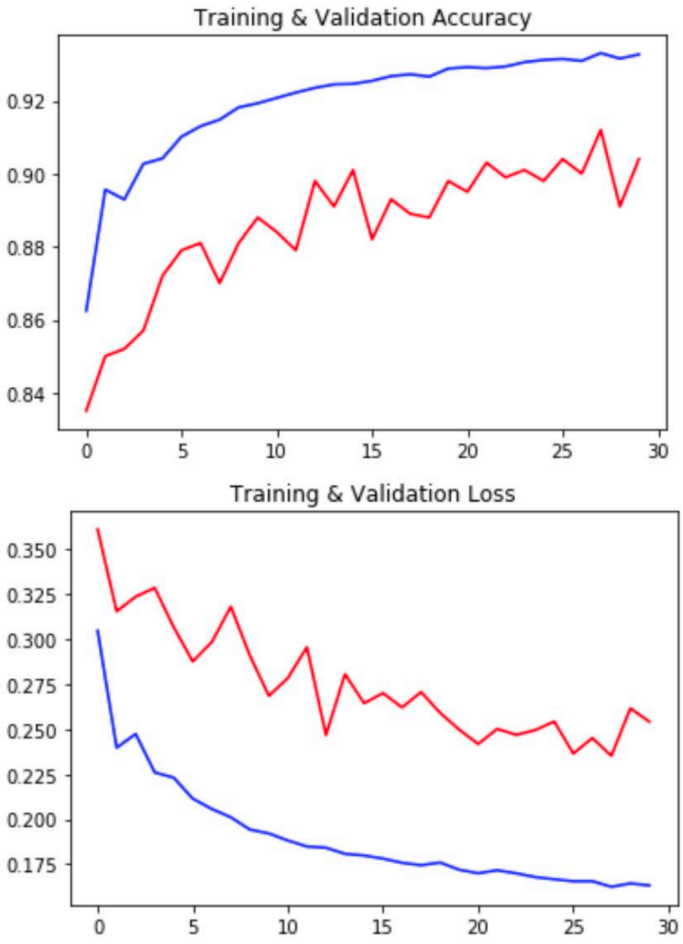
- 9 convolutional layers - Pooling & Dropout layer after every 3 convolutional layers
    - First 3 convolutional layers have 16 filters each
    - Second 3 convolutional layers have 32 filters each
    - Third 3 convolutional layers have 64 filters each
- Filter sizes = 3*3
- Padding = ‘same’
- Stride = 1
- Flatten layer turns output of convolutional layers into a vector
- First dense layer has 256 units
- Output layer has 2 units corresponding to the IDC positive & IDC negative classes
- Batch size = 500
- Epochs = 30
- Each epoch takes about 3-6 minutes on a Tesla K80 GPU provided by Google
  To verify the robustness of the model, the model was tested on augmented test images (rotated, flipped vertically and/or horizontally, shifted vertically and/or horizontally, zoomed in on), and the accuracy was close to the same. Training on augmented images might have helped with this more, allowing the model to be better trusted.

2. Justification

The CNN has an accuracy of 90.3% on a test set of 10,000 photos and 91.2% on a validation set of 1,000 images. In a validation set of 18,000 images, it achieved an accuracy of 85.52%, which is a substantial improvement over the baseline. This level of detail, however, is not adequate to declare the problem solved. In my opinion, this is not satisfactory for the hundreds of thousands of people that take the IDC exam annually. The model may be able to identify cancer on a different slide from the same patient even if it was unable to do so on a previous slide. The lives of patients should not be risked in this manner. It's possible that this method may be used to help doctors quickly identify who among their patients needs therapy the most. **(Model Summary Below)**

Layer (type)	Output Shape	Param #
conv2d_10 (Conv2D)	(None, 100, 100, 16)	448
conv2d_11 (Conv2D)	(None, 100, 100, 16)	2320
conv2d_12 (Conv2D)	(None, 100, 100, 16)	2320
max_pooling2d_4 (MaxPooling2	(None, 50, 50, 16)	0
conv2d_13 (Conv2D)	(None, 50, 50, 32)	4640
conv2d_14 (Conv2D)	(None, 50, 50, 32)	9248
conv2d_15 (Conv2D)	(None, 50, 50, 32)	9248
max_pooling2d_5 (MaxPooling2	(None, 25, 25, 32)	0
conv2d_16 (Conv2D)	(None, 25, 25, 64)	18496
conv2d_17 (Conv2D)	(None, 25, 25, 64)	36928
conv2d_18 (Conv2D)	(None, 25, 25, 64)	36928
max_pooling2d_6 (MaxPooling2	(None, 12, 12, 64)	0
dropout_3 (Dropout)	(None, 12, 12, 64)	0
flatten_2 (Flatten)	(None, 9216)	0
dense_3 (Dense)	(None, 256)	2359552
dropout_4 (Dropout)	(None, 256)	0
dense_4 (Dense)	(None, 2)	514

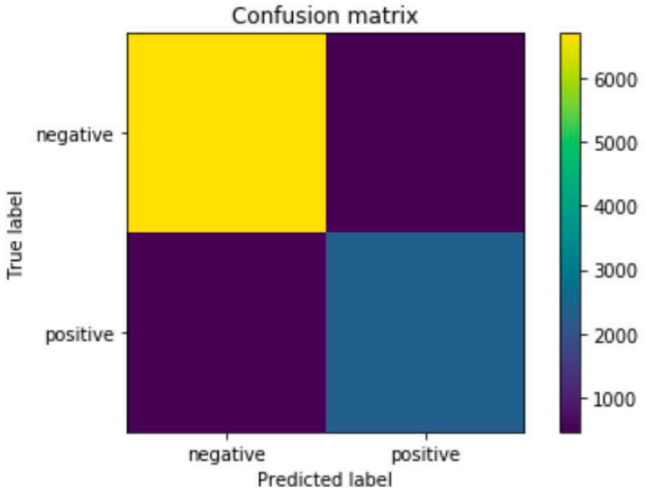
Total params: 2,480,642
Trainable params: 2,480,642
Non-trainable params: 0



2. Training and Validation Plots

V. CONCLUSION

Rough improving trend in validation accuracy is visualized until around 30 epochs. Over 30 epochs in training leads to gradual overfitting.

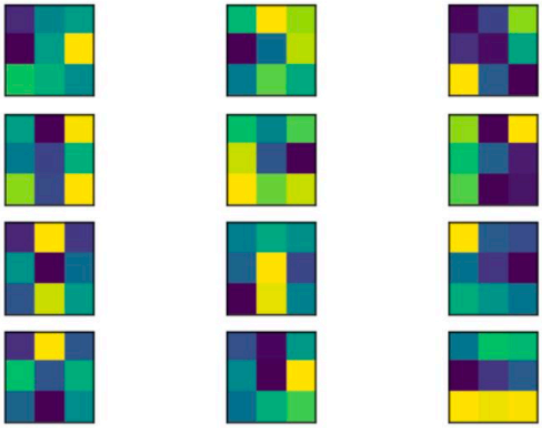



```
[6696,  465]
[ 504, 2335]
```

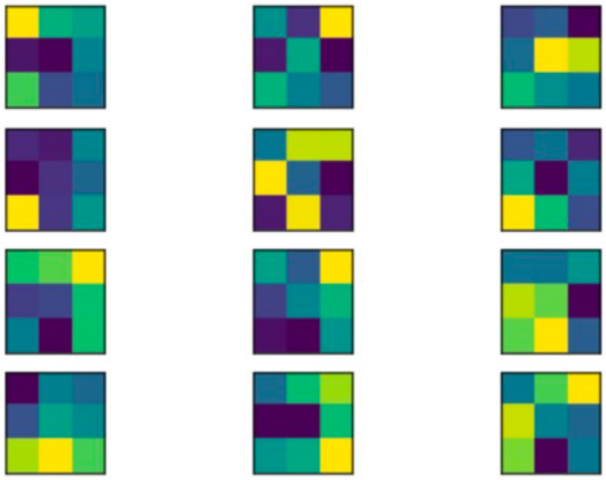
3. Confusion Matrix

	precision	recall	f1-score	support
negative	0.93	0.94	0.93	7161
positive	0.83	0.82	0.83	2839
accuracy			0.90	10000
macro avg	0.88	0.88	0.88	10000
weighted avg	0.90	0.90	0.90	10000

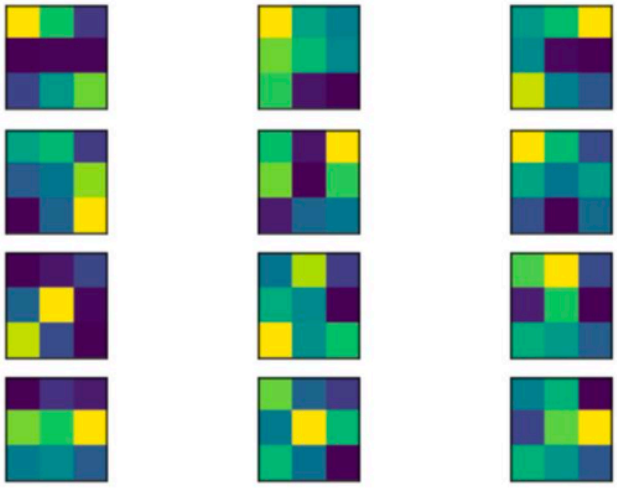
4. Classification Report



Filter of 3rd Convolution Layer



Filter of 6th Convolution Layer



Filters of 9th Convolution Layer

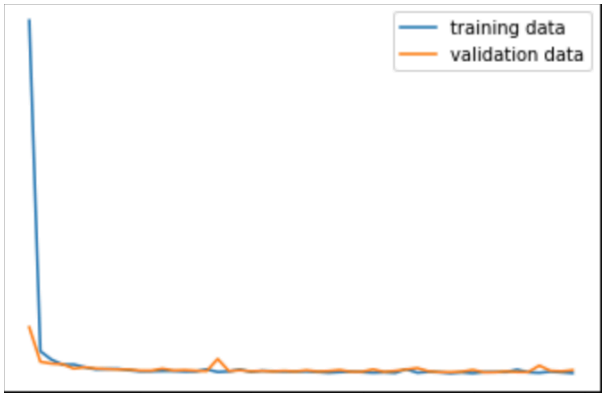
VI. REFINEMENT

1. Download and Preprocess the Data
2. Train a CNN Model
3. Evaluate the Results
4. Use for Predictions

The most difficult aspect to the project was how to preprocess/augment training samples. There are so many ways to do this making it impossible to test them all. It is very interesting to see the class activation maps of the model as it extracts lower level features from the image. The whole process behind how the model is able to classify the image is revealed through these activations which is incredible to observe.

In the Wisconsin Breast Cancer detection dataset, our best possible accuracy was 91% and the loss accuracy standing at 76.31%.

The graph of the training and the loss function of the dataset and the model applied was the following.



VII. FUTURE SCOPE

Other transfer learning models better suited to the task could be used to increase accuracy and speed up convergence. More data could also be used, including histopathologic cancer outside of the breast area. This Kaggle dataset could also be incorporated from the Histopathologic Cancer Detection competition: https://www.kaggle.com/c/histopathologic-cancer-detection

On top of this, other hyperparameters could be tweaked and model architecture improved. Data augmentation parameters and methods could especially be improved.

In the Wisconsin Breast Cancer detection dataset, our best possible accuracy was 91% and the loss accuracy standing at 76.31%.

The graph of the training and the loss function of the dataset and the model applied was the following.

VIII. REFERENCES

[1]. Liu Y., Chen P., Chen B., Wu J., Zeng Q., Lin X., Feng S. Breast cancer histopathological image classification using a multi-resolution convolutional neural network with patch-based attention and feature fusion. IEEE Access. 2019;7:46216–46225. doi: 10.1109/ACCESS.2019.2908075.

[2]. Naidu D., Kumar Y., Gorthi R.S. Deep learning-based detection of breast cancer using mammography images. J. Healthc. Eng. 2021;2021:6624757. doi: 10.1155/2021/6624757.

[3]. Kooi T., van Ginneken B., Karssemeijer N., den Heeten A. Discriminating solitary cysts from soft tissue lesions in mammography using deep learning. Med. Image Anal. 2017;37:60–67. doi: 10.1016/j.media.2016.12.006.

[4]. Wang J., Yang W., Li X., Li Y., Li B., Liang Z., Liang P. Multi-view convolutional neural networks for breast cancer detection in mammography images. Comput. Biol. Med. 2018;102:247–256. doi: 10.1016/j.compbiomed.2018.09.014.

[5]. Shen W., Zhou M., Yang F., Yu D., Dong D., Yang C., Zang Y., Tian J. Multi-crop convolutional neural networks for lung nodule malignancy suspiciousness classification. Pattern Recognit. 2017;61:663–673. doi: 10.1016/j.patcog.2016.06.035.

[6].Natarajan R., Puig D., Lobo J.M., Chen A., Giger M.L. Using deep learning to classify benign and malignant prostate lesions from ultrasound images. Sci. Rep. 2018;8:16685. doi: 10.1038/s41598-018-34993-9.

[7]. Wang L., Liu C., Guo X., Li H., Zhu Y. Deep learning-based breast cancer classification and prognosis analysis using ultrasound images. Int. J. Comput. Assist. Radiol. Surg. 2020;15:1273–1283. doi: 10.1007/s11548-020-02225-2.

[8]. Ren J., Cheng Q., Cheng J., Xu Y., Wei Y. A deep convolutional neural network for classification of red blood cells in sickle cell anemia. PLoS ONE. 2018;13:e0208412. doi: 10.1371/journal.pone.0208412.

[9] Wang S., Zhou M., Liu Z., Liu Z., Gu D., Zang Y., Dong D., Wei J., Liang C., Ma X., et al. Central focused convolutional neural networks: Developing a data-driven model for lung nodule segmentation. Med. Image Anal. 2018;47:167–179. doi: 10.1016/j.media.2018.04.009.

[10].Jin C., Chen Y., Zhang X., Chen B., Xie C., Wang X. Deep learning-based breast cancer diagnosis and prognosis prediction using ultrasound images. Int. J. Comput. Assist. Radiol. Surg. 2021;16:489–498. doi: 10.1007/s11548-021-02302-w.

[11]. Li H., Chen P., Zeng Q., Wu J., Lin X., Li X. Deep learning-based breast cancer classification by integrating multiple-level features. IEEE Access. 2019;7:80339–80347. doi: 10.1109/ACCESS.2019.2925542.

[12]. Yao J., Li M., Tang X., Tian J., Liang C., Wang S. Breast cancer diagnosis using deep convolutional neural networks on mammograms. J. Xray Sci. Technol. 2020;28:781–793. doi: 10.3233/XST-200716.

[13]. Han B., Zhang Y., Zhang R., Zhang Y., Zhang Y., Wang K., Wang S., Huang Q., Zhang Y. Breast cancer classification using deep learning on mammograms with breast density consideration. J. Xray Sci. Technol. 2021;29:519–534. doi: 10.3233/XST-210899.

[14]. Mckee J., Lou B., Gilbert F.J., Cerny M., Evans A., Harvey J., Kontos D., O'flynn E., Plumley N., Schilling K., et al. A prospective diagnostic accuracy study of breast screening using digital breast tomosynthesis in combination with synthesised two-dimensional mammography compared to digital breast tomosynthesis alone: The STORM protocol-

based design. Eur. Radiol. 2021;31:1123–1133. doi: 10.1007/s00330-020-07224-8.

[15]. Liu Y., Chen P., Zeng Q., Wu J., Lin X., Wang X. Multi-level deep convolutional neural network with online hard example mining for breast cancer detection in mammography images. IEEE Access. 2019;7:8253–8261. doi: 10.1109/ACCESS.2018.2881315.

[16]. Wang Y., Wang Y., Ma Z., Liang C., Tian J. Breast mass detection based on convolutional neural network and transfer learning using mammography images. Comput. Math. Methods Med. 2020;2020:4684012. doi: 10.1155/2020/4684012.

[17]. Chen X., Chen Y., Zhang X., Chen B., Xie C., Jin C. Deep learning-based breast cancer detection using ultrasound images. J. Healthc. Eng. 2021;2021:8814357. doi: 10.1155/2021/8814357.

[18]. Wang W., Qi Y., Xue Y., Yu B., Zhu Q., Hao Y., Liu X., Ma Y., He M., Wang X., et al. Deep learning-based breast cancer histopathological image analysis: A comprehensive review. J. Healthc. Eng. 2021;2021:5694360. doi: 10.1155/2021/5694360.

[19]. Pan Y., Yao Y., Chen B., Li J., Li J. Automated breast cancer detection and classification using convolutional neural networks with pre-trained weights. J. Xray Sci. Technol. 2021;29:409–419. doi: 10.3233/XST-210790.

[20]. Wu Y., Guo Y., Zhang W., Zhang Y., Huang Z., Wang L., Lu H., Gu J., Zhou Y., Zhang H. Breast cancer diagnosis via deep transfer learning with attention mechanism. Comput. Math. Methods Med. 2021;2021:8812476. doi: 10.1155/2021/8812476.

[21]. Abdel-Rahman M., Dey N., Ashour A.S., Islam M.M., Gomaa W.H., Bhattacherjee S. A hybrid model for breast cancer detection using deep learning and support vector machine. J. Ambient Intell. Humaniz. Comput. 2021;12:721–736. doi: 10.1007/s12652-020-02382-6.

[22]. Bora A., Biswas M. Breast cancer classification using deep learning and a combination of features extracted from multiple CNN models. Int. J. Med. Inform. 2021;153:104537. doi: 10.1016/j.ijmedinf.2021.104537.

[23]. Yap M.H., Pons G., Marti R., Ganau S., Sentís M., Zwiggelaar R. Automated breast ultrasound lesion detection using convolutional neural networks. IEEE Trans. Med. Imaging. 2021;40:441–452. doi: 10.1109/TMI.2020.3023743.

[24]. Li Y., Li W., Zheng Y., Li Y., Sun L., Zhang X. Multi-scale deep feature learning for breast cancer detection in mammography images. BMC Med. Imaging. 2020;20:100. doi: 10.1186/s12880-020-00503-2.

[26]. Shrestha A., Mahmood T., Gunavathi C. Breast cancer detection using deep convolutional neural networks and support vector machines. J. Digit. Imaging. 2021;34:1–10. doi: 10.1007/s10278-020-00371-6.

[27]. Kim S.M., Lee H., Lee S., Bae M.S., Kim J., Kim Y., Cho Y.U., Cho E.Y., Kang B.J. Automated breast ultrasound image analysis using deep learning for breast cancer screening: A pilot study. Ultrasonography. 2021;40:287–295. doi: 10.14366/usg.20128.

[28]. Zhang L., Peng Y., Wang X., Wang G., Wang Z., Yan S., Lu L., Zhao Q., Zhou Y. Deep learning-based breast cancer recurrence prediction with multimodality data. J. Healthc. Eng. 2021;2021:5595725. doi: 10.1155/2021/5595725.

[30]. Zhang H., Shi L., Song C., Wei L., Guo H., Zhou M. Breast cancer detection based on transfer learning of convolutional neural network. Comput. Math. Methods Med. 2021;2021:8852968. doi: 10.1155/2021/8852968.

Liu Y., Zheng L., Zhang Y., Wang X., Peng Y., Hao Z., Lu H., Zhou Y. A deep learning model for predicting axillary lymph node metastasis of breast cancer. IEEE J. Biomed. Health Inform. 2021;25:432–441. doi: 10.1109/JBHI.2020.2986593.