

A Deep Convolutional Neural Network Approach for Detecting Malignancy of Ovarian Cancer Using Densenet Model

Gitanjali Wadhwa

¹Assistant Professor

Dept. of Computer Engineering
KPR Institute of Engineering and Technology,
Coimbatore, Tamil Nadu, India.
Gitanjali.wadhwa1994@gmail.com

Dr. Jayanthi N

Assistant Professor(Sr.G)

Dept. of Computer Engineering
KPR Institute of Engineering and Technology,
Coimbatore, Tamil Nadu, India.
jayanthi.n@kpriet.ac.in

Mansi Mathur

Research Scholar

Dept. of Computer Engineering
J.C. Bose University of Science and Technology, YMCA,
Faridabad, Haryana
mansimathur66@gmail.com

Abstract: Ovarian malignant development has a helpless perseverance rate since general analysis and improved methods are needed for its underlying revelation. Ovarian danger is the sixth most characteristic infection in ladies, causing 152,000 decrease by and large yearly. To lessen this rate there is a requirement for an early detection of the infection. We plan a deep learning approach for essential recognition of ovarian malignancy utilizing histopathological pictures. Convolutional Neural Network (CNN) is utilized for the process. Feature extraction is finished by utilizing DenseNet-201 model of CNN. Malignant and benign cancer types are classified. In the characterization cycle, the PLCO dataset is utilized with most elevated ac-curacy accomplished is 94.73, exactness is 0.91, review and f1-score is 0.90 and 0.95 separately. Conditional outcomes and assessment of other earlier work explains truly trustworthy execution and the profitability of proposed work.

Keywords: Ovarian Cancer, CNN, Machine Learning, Deep Learning, Histopathological image

I. INTRODUCTION

Ovarian threat is quite possibly the most notable gynaecologic malignant growths [4]. Precise classification of ovarian infection types (mucous carcinoma, endometrioid cancer, serous carcinoma, transparent cell cancers) is a key part in different finding. Computer aided diagnosis (CADx) can give accommodating appeal to the specialists and pathologists to analyze the features [4][21]. In our examination, we used a Deep Convolutional Neural Networks (DCNN) for the early recognition of ovarian malignant growth from histopathology pictures.

Ovaries are important part of female reproductive system. The importance of these tiny glands is derived from the production of female sex hormones and female gametes. The location of these ductless almond shaped small glandular organs is on just opposite sides of uterus attached with ovarian ligament. When cells in the ovaries going to develop abnormally, they became tumors-

can be cancerous / non-cancerous [22]. If the cancerous cells/tumors couldn't be detected early; the cancer cells can spread throughout the ovary and pelvic region and continue to spread into abdominal area and into other organs. Determining the stage of ovarian cancer is very important in developing a treatment plan. There are four stages of ovarian cancer which have approximately maximum 5 years of survival rate of a woman who is infected. At stage 1- there is 82-92% chance of surviving the cancer. In this stage the tumor is limited to one or both ovaries, it originates from ovarian surface. At stage 2- there is 51-69% survival rate. In this the tumor invades one or both the ovaries in extension to the pelvic region but without spreading to the abdomen. At stage 3- the survival chance is 17-39% as tumor extends beyond the pelvic into the abdominal organs. Finally, in stage 4, only 11.5% chances of life because of distant metastasis to the lung, liver or lymph nodes in the neck. Ovaries consists of 3 types of cells. The first cell is Epithelial cell [20]. Most common cancers among them is epithelial tumors. About 85-90% are epithelial and are more aggressive cancers. The second type is Stromal cell which is about 2-3% only. Stromal tissues are the connected tissues of the ovaries, most of them are benign. And the third type is Germ cell tumors are actually the producers of ovum/the egg. They are also less frequent and pretend to happen more in teenage or in younger women. But majority of them are benign and if they are not benign then they are very treatable by just removing their ovary and they can bear fertility as well because they usually don't undergo big surgeries.

Ovarian tumors are classified in two type of tumors namely benign and malignant [4][20][21]. In Benign Tumors some originates from surface ovarian epithelium, such as serous cystadenomas, which are often bilateral and filled with a watery fluid, and mucinous cystadenomas, which tend to be multicular and contains a mucus-like fluid and the other originates from ovarian germ cells, such as mature cystic teratomas, also called dermoid cysts [23]. These are most common ovarian tumours in young women and usually contain a heterogenous mix of mature tissue that come from two or three of the germ cells layers. The other classification is Malignant Tumors, they are like serous or mucinous cystadenocarcinomas, which develop from surface ovarian epithelium, just like their benign counterparts.

Computer-aided diagnosis System [17-23] are mainly used for recognizing the cancer using image analysis. With the enhancement in deep learning technique CAD systems are improved and are able to produce more reliable results in the analysis of ovarian cancer with the help of histopathology images. CAD systems add to match the finding of the pathologist, disregard inter-pathologist variants in diagnosis, and highly expand the efficiency of detection. There are various ML techniques [3][7][24] other than CNN such as RF (Random forest) [20], SVM (Support Vector Machine) [5], KNN, Decision Tree [20] which are also used in CAD system to diagnose ovarian cancer. Classification methods are the cultured methods that are used to classify medical images of ovary as cancerous or non-cancerous shown in Fig. 1.:

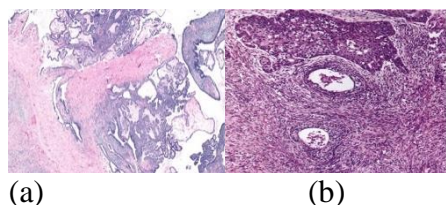


Fig. 1. Illustration of images: (a) Benign (b) Malignant

As shown in Fig. 1., minuscule depictions are outsized, cytological images [15-16], and the incalculable changes among these pictures that differs from patients to patients, which expend high intra-class dissimilarities and low inter-class dissimilarities [11]. For the characterization deep Convolutional neural system is mostly utilized and we are using DenseNet multi-network

feature model for ovarian cancer detection.

In DenseNet-201 framework, an intensive thick segment is anticipated, which joins each layer with all further layer in a feed-forward routine. It releases the slope vanishing problematic, fortifies element transmission, support up include reprocess and altogether wounds down the number of imperatives. DCNN (Deep convolutional neural system) [12] is notable model for clinical pictures in end to end architecture that settle the issue of overfitting [17]. These deep learning models have high computational expense when contrasted with different models utilized for other reason used for classification purpose [10]. Our purposed technique removes the duplicate features that are extracted using these models for feature extraction and also select only the important features also SoftMax layer of DCNN is used for boosting up the classification performance [2].

In this research, we design a method for the classification of ovarian cancer by using medical histopathological images of ovaries implemented with the help of multi-network feature extraction model i.e. DenseNet-201 and Deep learning classifier Convolutional Neural Network.

The main features of the proposed work are as follows: -

- Multi-layered component extraction model is castoff to remove the features from the Haematoxylin and eosin stain pictures (H & E stain).
- The PLCO dataset for Classification is reserved of 256 clinical biopsy pictures of patients experiencing ovarian malignant growth. Tumour tissue of patients are made out of 48 patients with different amplifying features.

This paper is managed in a certain pattern as followed. In next Section I.e. Related work, explanation of brief reviews related to our effort of the analysis of ovarian cancer. Then in the next Section III, i.e. Methodology, we define our projected technique in details. Experimental outcomes are specified in Section IV tracked by results in Section V. The discussions are in Section VI. Finally, the conclusions are in Section VII.

II. Related Work

Sakshi Srivastava [1] proposed a fine-tuned model using traditional VGG-16 deep neural network trained on ImageNet dataset comprised of 240 images from ultrasound of various ovaries of different females to detect whether they have cysts or not. The proposed algorithm gave 92.11% accuracy. Future works includes classification of different types of cysts including functional, HOC (hemorrhagic ovarian cyst), PCOS (polycystic ovarian syndrome) and dermoid.

Zheng Zhang [3] presented a machine learning based framework for detection of ovarian cancer by taking their picture content for classification task in convolutional neural network technique. Basically, they proposed a single methodology characterization(classification) algorithm. They tested the performance of the proposed model named ML-CNN-LR classifier which abstracts obstetric tumor images, on the basis of two factors: - precision ratio and recall rate.

Wu, Miao, et al. [4]utilized a Deep Convolutional Neural Networks (DCNN) in view of AlexNet to naturally characterize the various kinds of ovarian diseases from cytological pictures. The DCNN comprises of five convolutional layers, three max pooling layers, and two full reconnect layers. At that point we prepared the model by two gathering input information independently, one was unique picture information and the other one was enlarged picture information including picture improvement and picture rotation. The testing results are acquired by the technique for 10-overlap cross-approval, indicating that the exactness of order models has been improved from 72.76 to 78.20% by utilizing enlarged pictures as preparing information. The created conspire was valuable for grouping ovarian tumors from cytological pictures.

Komura and Ishikawa [5] they proposed a technique to analyze histopathology images with the help of computer-aided diagnosis system to detect cancer at early stage. Their proposed algorithm performs quite good and provide an accuracy of 92.7%.

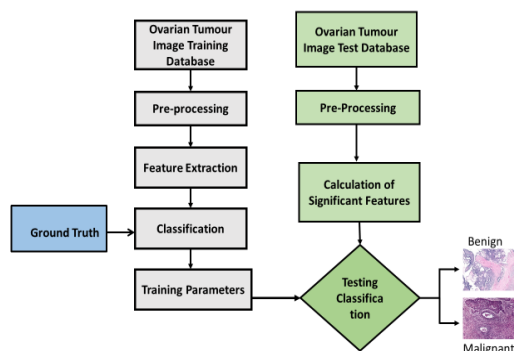
Wang et al. [6] they assess the projected process on the open ICIAR2018 Challenge dataset of medical histology images of cancer patients and they achieved a highest classification accuracy of 94.70%. and they experimented on DCNN with E-SVM i.e. Ensemble Support Vector Machine with four pre-trained models (e.g. ResNet-50, multi-level InceptionV3, DenseNet-121, and multi-level VGG-16) their technique can attain fairly auspicious results and beat state-of-the-art approaches. They train their E-SVM classifier for merged feature mining and chosen procedure to expand the classification.

Eiryō Kawakami et al. [19] concluded that AI frameworks can give basic prognostic and diagnostic expectation for patients with Epithelial Ovarian Cancer before beginning medication, and the utilization of prescient calculations may encourage customized treatment alternatives through pre-treatment delineation of patients. They combined the weak decision tree like random forest (RF), conditional random forest (CRF) and gradient boosting machine (GBM) with ensemble methods to achieve the best results. The resultant accuracy and area under the ROC curve (AUC) for isolating Epithelial Ovarian Cancer (EOC) from generous ovarian tumours with RF were 92.4% and 0.968, individually. The most elevated precision and AUC for anticipating clinical stages with RF were 69.0% and 0.760, individually.

III. Methodology

We design a technique for the identification of OC by utilizing clinical Histopathology pictures executed with the assistance of multi-network feature extraction model for example DenseNet-201 and Deep learning classifier Convolutional Neural Network.

This Fig. 2. shows the system of proposed method which comprises of three stages that are: -



1. Data (dataset input which has histopathological pictures of two classes benign and malignant),
2. Feature Extraction model (highlights have been mined from the histology pictures with the assistance of extraction model – DenseNet 201), and
3. Classification (In this progression either the patient is experiencing malignant growth or not).

3.1 Dataset

In this framework the performance is calculated by using PLCO dataset which is an open source dataset available online. It contains H&E stain images (Hematoxylin and eosin stain) stained medical images of ovarian slides. There are 4 dimensions of images in that dataset x40, x100, x200 and x400 and their pixel size is 0.49µm, 0.20µm, 0.10µm and 0.05µm respectively. These images have 24-bit TrueColor color space. This dataset consists of histopathological images of 48 patients and there is some diagnostically significant RIO (Region of Interest) present as well.

Cloudy as well as low focused images are removed and there are 20-23 images available for each patient. Some images are of borderline cases of ovarian cancer so those were placed in ductal carcinoma as well as in lobular carcinoma.

There are 2 categories of images present in PLCO dataset as shown in Fig. 3.:-

- a) *Benign*
- b) *Malignant*

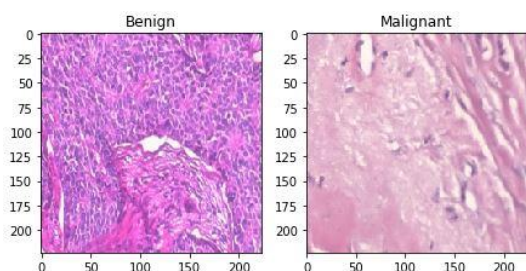


Fig. 2. Graphical notations of Benign and Malignant types

3.2 Feature Extraction Using DenseNet201

Effectiveness of Deep Learning strategy relies upon size of information test for preparing and testing. The proportions of training and testing is 8:2 individually. To benefit high evaluating parameters like accuracy, precision, recall and so forth., we need enormous training set of clinical pictures which is once in a while accessible, limited training set can cause overfitting. So, to comprehend this issue we utilize this DenseNet-201 model-Dense Convolutional Network (Image Classification) [12] shown in Fig. 4.

We train our framework with PLCO dataset and instate with pre-prepared loads for the learning reason. In DenseNet each layer has extra information layer to pass feature data. In this method, model link is utilized for concatenation. Every layer has consolidated data from all procedure layer. Check of channels will be less a direct result of continuing layer. This model has high memory productivity just as high computational proficiency. The pace of learning is 0.0001 for 20 epochs. GlobalAveragePooling layer is likewise utilized with half dropouts which help us in diminishing over-fitting [8][11].

To sort the outcomes in two ways as: benign and malignant; we utilize thick layer and cluster standardization utilizing 2 neurons as to give binary input, SoftMax is utilized as an activation function [18]. Binary cross-entropy is utilized as loss function and we use Adam to be an optimizing agent.

```
Model: "sequential_1"
```

| Layer (type) | Output Shape | Param # |
|---|--------------------|----------|
| densenet201 (Model) | (None, 7, 7, 1920) | 18321984 |
| global_average_pooling2d_1 ((None, 1920) | | 0 |
| dropout_1 (Dropout) | (None, 1920) | 0 |
| batch_normalization_1 (Batch (None, 1920) | | 7680 |
| dense_1 (Dense) | (None, 2) | 3842 |

Total params: 18,333,506
 Trainable params: 18,100,610
 Non-trainable params: 232,896

Fig. 3. Summary of Dense Convolutional model used for implementation purpose

ModelCheckpoint and ReduceLROnPlateau these two are likewise utilized. This Model-Checkpoint is use to spare the best aftereffect of the usage when there is no improvement in results considerably after numerous cycles(iterations). This Reduce-LR-On-Plateau work utilized as Reduce Learning Rate for the metric when it isn't improving [6][16].

3.3 Classification

In this subsection we are utilizing CNN classifier for the classification reason for that we train the classifier by using features to analyze ovarian cancer. We train this framework with solid vigor and with high exactness as in precision [23]. We train CNN classifier derived from DenseNet model as to offer their features with one another layers as they linked to their features [13].

Along these lines the characterization is finished utilizing CNN and the model consequently recognizes whether the patient from her histopathology images are benign or malignant.

IV. Experimental Results

We figure out the performance of our anticipated strategy to analyze ovarian malignant growth utilizing medicinally recolor pictures of PLCO dataset evaluation is assessed as ACC (exactness), ROC (Receiver Operating Characteristics), AUC (Area Under Curve), Recall, PRE (accuracy) and F1-score [14], [15], [19],[21].

$$ACC = \frac{TP + TN}{TP + FN + TN + FP}$$

$$PREC = \frac{TP}{TP + FP}$$

$$RECALL = \frac{TP}{TP + FN}$$

$$F1 = \frac{2 * (RECALL * PREC)}{RECALL + PREC}$$

Notations: -

ACC = Accuracy,
PREC = Precision,
TP = True_Positive,
TN = True_Negative
FP = False_Positive,
FN = False_Negative

V. Results

We execute DenseNet-201 by utilizing histology images of cropped size of x100 and x200 pixels. This classifier is made and trained to perform cancer diagnostic to obtain good results that are shown in Fig. 5.,

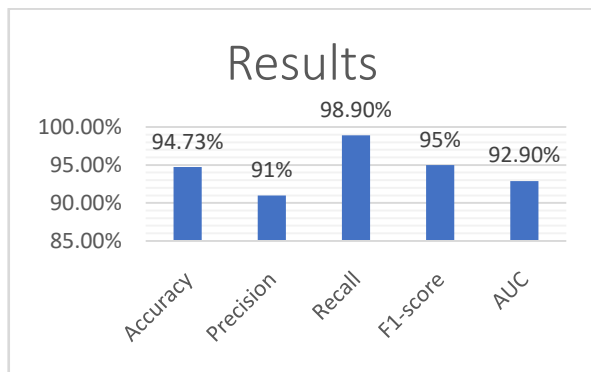


Fig. 4. Experimental Results of framework proposed for classification using H&E Stained images

The accuracy obtained is 94.73%. distinctive cropped size picture has various features. So, we combined the two sizes pictures for better outcomes.

Table 1.Result table showing achieved values of all the Performances Metrics

| Performances Matrices | Results |
|-----------------------|---------|
| Accuracy | 94.73% |
| Precision | 91 |
| Recall | 98.90 |
| F1-score | 95 |
| AUC | 92.90 |

While preparing the model, sample dataset further partitioned into two inside sub groups. One of them is utilized for training routine and the another is utilized for approval that occurred after each epoch (in our case, we have epoch=20). There are two terms in acc and val_accuracy, acc is a term utilized for Accuracy of trained samples and val_accuracy characterized as validation sub group. These val_accuracy are the sample sets that doesn't appeared on the system during training. However, they gave us how our model will be like in general. Diagram below shows val_accuracy and acc chart.

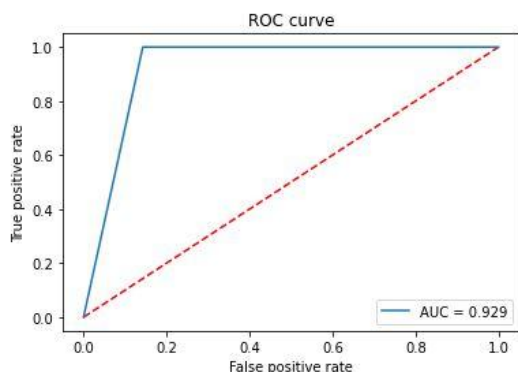


Fig. 5. ROC-AUC curve

Higher AUC shows how reliable the model is. 1 is the highest point that one can have as it is the best result because at that point right angle triangle will be formed. For debugging the model ROC curve is very helpful.

VI. Discussion

In this research, we propose a plan to analyze ovarian malignant growth using clinical histopathology pictures, in this we are utilizing Deep Convolutional Neural Network model and PLCO database to train the model and to survey the evaluation, assessment is done on different parameters like Accuracy, Precision, Recall, F1-score and Area Under the Curve - Receiver Operating Characteristics. To include pictures arbitrarily 20 pictures are trimmed and afterward further their features were mapped by utilizing the last layer of convolutional layer.

At the point when we increment the depth of the network, structure likewise increments and afterward we get detailed features like surface, shading, limit and shape. These highlights at that point are used to train the classifier.

Results of the experiments are shown in Table 1. It is seen that best quality pictures give the best outcomes. There are many scientists who used diverse methods and techniques for classification of ovarian cancer which is as shown in Table 2.

Table 2. Experimental results of Classification Methods used by different researchers from related work.

| Classification Methods | Accuracy |
|------------------------------|----------|
| Proposed Work | 94.73% |
| Pinsky, Paul F., et al. [25] | 90% |
| Wu, Miao, et al. [4] | 78.20% |
| Zhu, Claire S., et al. [26] | 89.20% |
| Wang et. al. [6] | 94.70% |
| Buys, Sandra S et al. [27] | 94% |

The graph in Fig. 7. is showing the comparison of the experimental results of existing classification methods used by researchers from the related work with our proposed method on the basis of accuracy as a parameter.

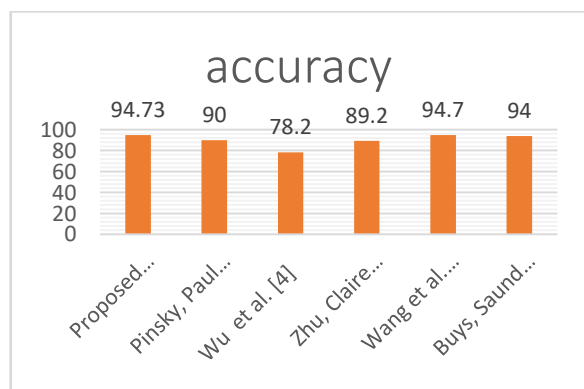


Fig. 6. Comparison of results of various researchers on the basis of accuracy

VII. Conclusion

In this research paper, we defined a proficient and viable framework for the determination of ovarian malignant growth utilizing H&E recolored clinical histopathology pictures. Features are being separated with the assistance of Dense convolutional system and afterward help the model

to prepare the framework with highlights to improve the arrangement method [9]. With the assistance of proposed method, we obtain high precision and all other evaluation parameters.

This proposed procedure is to be utilized clinically by the specialists to productively analyze ovarian malignant growth at an early stage, which will assist with expanding the death pace of disease casualties. Our technique is able to attain 94.73% accuracy.

The detection of ovarian cancer tumor should be done at an early age to reduce maternal and perinatal mortality rate. Once a woman is caused with ovarian cancer her survival chances decreases and it is difficult to cope up as normal beings. If this happens in early age then it can be cured by removing the ovary so that with time their body will regenerate a new one. There are two methods to be taken as precaution for the detection of ovarian tumor is biopsy procedures: - Liquid and mass. There are many factors on the basis of which anyone can have ovarian cancer but if someone has family history then they should get check-ups on time to avoid the tumor and to be safe.

In future, we will speed up the working of our proposed strategy. To improve the precision of the determination procedure we will going to concentrate on further component extraction matters from malignant growth pictures. To look at the particular areas for feature work to do.

References

1. Srivastava, Sakshi, et al. "Detection of Ovarian Cyst in Ultrasound Images Using Fine-Tuned VGG-16 Deep Learning Network." *SN Computer Science* 1.2 (2020): 1-8.
2. Zhu, Xiaofeng, Heung-Il Suk, and Dinggang Shen. "Low-rank dimensionality reduction for multi-modality neurodegenerative disease identification." *World Wide Web* 22.2 (2019): 907-925.
3. Z. Zhang and Y. Han, "Detection of Ovarian Tumors in Obstetric Ultrasound Imaging Using Logistic Regression Classifier With an Advanced Machine Learning Approach," in *IEEE Access*, vol. 8, pp. 44999-45008, 2020.
4. Wu, Miao, et al., "Automatic classification of ovarian cancer types from cytological images using deep convolutional neural networks." *Bioscience reports* 38.3, 2018.
5. D. Komura and S. Ishikawa, "Machine learning methods for histopathological image analysis," *Computational and structural biotechnology journal*, vol. 16, p. 34-42, 2018.
6. Y. Wang, B. Lei, A. Elazab, E.-L. Tan, W. Wang, F. Huang, X. Gong and T. Wang, "Breast Cancer Image Classification via Multi-Network Features and Dual-Network Orthogonal Low-Rank Learning," *IEEE Access*, vol. 8, p. 27779-27792, 2020.
7. K. Doi, "Computer-aided diagnosis in medical imaging: historical review, current status and future potential," *Computerized medical imaging and graphics*, vol. 31, p. 198-211, 2007.
8. D. Shen, G. Wu and H.-I. Suk, "Deep learning in medical image analysis," *Annual review of biomedical engineering*, vol. 19, p. 221-248, 2017.
9. Madabhushi, Anant, and George Lee. "Image analysis and machine learning in digital pathology: Challenges and opportunities." (2016): 170-175.
10. A. Pimkin, G. Makarchuk, V. Kondratenko, M. Pisov, E. Krivov and M. Belyaev, "Ensembling neural networks for digital pathology images classification and segmentation," in *International Conference Image Analysis and Recognition*, pp. 877-886, 2018.
11. Otálora, Sebastian, et al. "Image magnification regression using densenet for exploiting histopathology open access content." *Computational pathology and ophthalmic medical image analysis*. Springer, 148-155, 2018.
12. Celik, Yusuf, et al. "Automated invasive ductal carcinoma detection based using deep transfer learning with whole-slide images." *Pattern Recognition Letters*, March ,2020.
13. Y. Zheng, Z. Jiang, H. Zhang, F. Xie, Y. Ma, H. Shi and Y. Zhao, "Histopathological whole slide image analysis using context-based CBIR," *IEEE transactions on medical imaging*, vol. 37, p. 1641-1652, 2018.
14. G. Litjens, T. Kooi, B. E. Bejnordi, A. A. A. Setio, F. Ciompi, M. Ghahfarooi, J. A. Van Der

- Laak, B. Van Ginneken and C. I. Sánchez, "A survey on deep learning in medical image analysis," *Medical image analysis*, vol. 42, p. 60–88, 2017.
15. X. Zhu, J. Yao, F. Zhu and J. Huang, "Wsis: Making survival prediction from whole slide histopathological images," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2017.
 16. G. Huang, Z. Liu, L. Van Der Maaten and K. Q. Weinberger, "Densely connected convolutional networks," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 4700-4708, 2017.
 17. K. He, X. Zhang, S. Ren and J. Sun, "Deep residual learning for image recognition," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 770-778, 2016.
 18. Chu, Carlton, et al. "Does feature selection improve classification accuracy? Impact of sample size and feature selection on classification using anatomical magnetic resonance images." *Neuroimage* 60.1: 59-70, 2012.
 19. Kawakami, Eiryō, et al. "Application of Artificial Intelligence for Preoperative Diagnostic and Prognostic Prediction in Epithelial Ovarian Cancer Based on Blood Biomarkers." *Clinical Cancer Research* 25.10: 3006-3015, 2019.
 20. Yassin, Nisreen IR, et al. "Machine learning techniques for breast cancer computer aided diagnosis using different image modalities: A systematic review." *Computer methods and programs in biomedicine* 156: 25-45, 2018.
 21. Trinidad, Camille V., et al. "Reducing Ovarian Cancer Mortality Through Early Detection: Approaches Using Circulating Biomarkers." *Cancer Prevention Research* 13.3: 241-252, 2020.
 22. John DT, "Classification of Ovarian Cysts Using Artificial Neural Network", *International Research Journal of Engineering and Technology (IRJET)*, Volume: 03 Issue: 06 | June-2016.
 23. F. Gao, T. Wu, J. Li, B. Zheng, L. Ruan, D. Shang and B. Patel, "SD-CNN: A shallow-deep CNN for improved breast cancer diagnosis," *Computerized Medical Imaging and Graphics*, vol. 70, p. 53–62, 2018.
 24. M.Navya, M.RamakrishnaMurty, et al, "A comparative analysis of breast cancer data set using different classification methods" *International Conference and published the proceedings in AISC, Springer, SCI-2018*.
 25. Pinsky, Paul F., et al. "Potential effect of the risk of ovarian cancer algorithm (ROCA) on the mortality outcome of the Prostate, Lung, Colorectal and Ovarian (PLCO) trial." *International journal of cancer* 132.9: 2127-2133, 2013.
 26. Zhu, Claire S., et al. "A framework for evaluating biomarkers for early detection: validation of biomarker panels for ovarian cancer." *Cancer Prevention Research* 4.3: 375-383, 2011.
 27. Buys, Sandra S et al. "Effect of screening on ovarian cancer mortality: the Prostate, Lung, Colorectal and Ovarian (PLCO) Cancer Screening Randomized Controlled Trial." *JAMA* vol. 305,22: 2295-303, 2011. doi:10.1001/jama.2011.766.