

Comparative Analysis of Deep Learning Models for Early Prediction and Subtype Classification of Ovarian Cancer: A Comprehensive Study

Kokila R. Kasture¹, Dr. Wani V. Patil², Dr Amalraj Shankar³

Submitted: 03/10/2023 Accepted : 22/11/2023 Accepted: 02/12/2023

Abstract This study compares the performance of two state-of-the-art deep convolutional Neural Network architectures, AlexNet and VGG-19, for predicting and classifying the sub-types of ovarian cancer from histopathological images. The dataset consisted of 500 images, augmented to generate 24,742 images, which were used to train both models. The results showed that VGG-19 outperformed over AlexNet, achieving an accuracy of 90% compared to 70% for AlexNet. Other performance metrics, such as precision, recall, F1-score, and AUC-ROC, were also analyzed. This study provides valuable insights into the use of computer-aided diagnosis for accurately Predicting the diagnosis and subtype of ovarian cancer, which can lead to early detection and treatment.

Keywords: AlexNet, Convolution neural network, Ovarian cancer, VGGNet.

1. Introduction

Ovarian cancer is a leading cause of cancer death among women and often goes undetected until symptoms such as bloating, pelvic pain, appetite loss, or abdominal swelling appear [1]. Unfortunately, by this point, the cancer has often spread to other parts of the body, making treatment difficult. Women who are post-menopausal or have a family history of ovarian cancer are at a higher risk [2]. Early detection is challenging, and it is the most persistent cause of death compared to other gynecological cancers. Various imaging techniques and serum markers have been researched to improve early detection, but they have disadvantages such as missed detections, time-consuming procedures, and requiring skilled clinicians. Serum Carbohydrate Antigen 125 (CA125) serves as a commonly employed biomarker in the detection of ovarian tumors. Elevated concentrations of CA125 are observed in roughly 50% of women during the early stages of ovarian cancer (OC), and this percentage increases to over 80% during the later stages [3].

Magnetic resonance imaging (MRI), Ultrasound imaging, and positron emission tomography (PET) are frequently employed imaging modalities for the identification and characterization of tumors associated with ovarian cancer.

However, machine learning algorithms such as linear support vector machine (SVM), logistic regression, ensemble SVM, random forest, and boosting have demonstrated inadequate classification accuracy [4]. To increase the chances of early detection and improve patient survival rates, a combined approach using biomarkers and machine learning algorithms may be effective. Previous research has predominantly focused on manual feature extraction and the utilization of supervised machine learning algorithms to classify images as either cancerous or non-cancerous. For example, Chen et al. [5] used a generic support vector machine with textural and pathological features to classify thyroid nodules, while Chang et al. [6] incorporated ultrasound images to detect Graves' disease using SVM.

2. Literature Review

In their study, Albarqouni et al. [7] introduced a novel multi-scale CNN AggNet that incorporates crowdsourcing as an additional layer for data aggregation. Although deep learning methods were used for labelling the ground truth from non-expert crowd annotation, the influence of the computational aggregation method was found to be relatively small. Minig et al. [6], on the other hand, compared the outcomes of laparoscopy and open surgery in patients with ovarian cancer and showed that laparoscopic surgery had similar surgical and oncological outcomes. Finally, Sirinukunwattana et al. [7] developed a deep learning approach for detecting and classifying nuclei in colorectal cancer images that were sensitive to the local neighborhood. This method involved creating a spatially constrained CNN for nucleus detection and a nearby

¹ Research Scholar, Department Electronic and Telecommunication, G H Raisoni University, Amravati, Maharashtra, India.
koki.thakur@gmail.com

² Assistant Professor, Department Electronic and Telecommunication, G H Raisoni University, Amravati, Maharashtra, India.
wani.patil@raisoni.net

³ Assistant Professor, Department Electronic and Telecommunication, G H Raisoni University, Amravati, Maharashtra, India.
shankaramalraj@gmail.com

ensemble predictor for classification, resulting in a systematic analysis of tissue morphology that improved our understanding of the tumor microenvironment.

Xu. J et al. [8] proposed an innovative Deep CNN approach for the segmentation and classification of epithelial and stromal regions in breast and colon cancer. By analyzing the complex features using a data-driven approach, their method outperformed traditional methods that relied on handcrafted features, achieving higher classification accuracy rates across various applications.

H. Sharma et al. [9] utilized CNN-based deep learning algorithms for automated classification to diagnose malignancy and detect necrosis in tissues. They used the widely used AlexNet framework to address categorization issues, resulting in impressive results with a cancer classification accuracy of 0.699 and necrosis classification accuracy of 0.8144.

W. Sun et al. [10] compared deep learning techniques with traditional CADx systems that relied on handcrafted features to automatically extract features from lung images. The approach considered features such as morphology, density, and texture, and outperformed standard CADx algorithms with a larger dataset and fine-tuned parameters.

H. Sharma et al. [12] utilized CNN architecture-based deep learning algorithms to automatically classify malignancy and detect necrosis in tissues. The widely used AlexNet framework was used to analyze categorization issues, which produced impressive results. This CNN framework classified cancer with an accuracy of 0.699 and necrosis with an accuracy of 0.8144.

W. Sun et al. [13] compared deep techniques with traditional CADx systems that rely on hand-crafted features to automatically extract features from lung images. This approach incorporated morphology, density, and texture features, outperforming current standard CADx algorithms with a larger dataset and fine-tuned parameters. However, the method did not investigate the optimal input size and only employed a few layers.

A. Das et al. [14] used CT liver images to develop a watershed Gaussian approach based on deep learning algorithms to detect cancer. Cancer regions were segmented using a Gaussian mixture model, and various textural features were extracted and used as input by the DNN classifier to determine the type of malignancy. The study achieved a 99.38 percent classification accuracy using 200 epochs and a validation loss of 0.062. Its main drawback was the lack of assessment of the lesion's volumetric size.

S. Wang et al. [15] developed a non-invasive prediction model for detecting cancer recurrence by deriving

biomarkers from CT images using a unique DL technique. The decision curve analysis confirmed the high effectiveness of this recurrence prediction model, and DL characteristics showed a better predictive value than clinical ones. However, the model could only recover intrinsic traits, and its use is limited to Cox-PH for additional research.

Y. Feng et al. [16] used 3D convolutional processes to extract spatial and temporal features and obtain dynamic information for cancer detection using the perfusion process. The model was trained and validated using CEUS images obtained by two contrast agents, and the deep learning model outperformed prior methods by 91 and 90 percent, respectively, in terms of specificity and accuracy.

S. Marciauskas et al. [17] developed a novel selection approach for the proteome study of ovarian cancers that incorporated various normalization processes, as well as univariate statistics, naive Bayes, and logistic tree classifiers. This model aimed to achieve superior outcomes for verifying the selected proteins for early-stage ovarian cancer diagnosis.

In their research, A. Dascalu and E.O. David [18] explored the influence of image quality on the accuracy of skin cancer diagnosis using a skin magnifier with polarized light (SMP). The acquired SMP images were subjected to deep learning techniques for processing, and the sonification outcomes were employed for diagnosing the results. Interestingly, the study revealed that image quality did not impact the accuracy of skin cancer diagnosis. The autonomous detection capability of this equipment proved advantageous for healthcare professionals.

Z. Alyafeai and Ghouti [19] developed an autonomous deep learning model for detecting the cervix region and classifying cervical cancer. Our system used two pre-trained DL techniques for automatic detection and classification. The detection model was considerably faster than traditional data-driven techniques, and the classification model employed self-extracted features using two lightweight CNN models. The proposed deep learning classifier demonstrated superior performance compared to similar models, excelling in terms of factors such as speed and classification accuracy. Consequently, it is well-suited for deployment on mobile phones.

P. Kaur, G. Singh, and P. Kaur [20] developed a system based on Multi-Support Vector Machine (MSVM) with DL K-mean clustering, which provided superior results to the decision tree model. The effectiveness of this DL model was analyzed and validated using only a small dataset.

Z. Liu et al. [21] studied the efficacy of medical IoT-enabled CAD applications for deep reinforcement learning and lung cancer treatment. This model offered several deep reinforcement learning applications, which were effective

in locating lung cancer and improving treatment outcomes.

Pham et al. [22] developed a two-stage DL model to enhance the detection of malignancy and reduce false-positive predictions. The initial stage of the DL algorithm eliminated noncancerous areas that were mistakenly labeled, followed by a DL classifier for cancer cell identification. The two-stage approach reduced errors by 36.4% to 89%, resulting in a higher detection rate and fewer false-positive outcomes. However, some limitations of this approach, such as the small dataset used and high sensitivity with low specificity, were observed.

Shakeel et al. [23] employed the Improved Profuse Clustering Technique (IPCT) and DL Instantaneously Trained Neural Networks (DITNN) to predict lung cancer. They improved the image quality using the weighted mean algorithm and segmented the affected regions based on the estimated pixel similarity value after improving image quality. The trained features achieved a classification accuracy of 98.42% and a minimal classification error of 0.038.

Authors Contributions:

- I. Collection of Dataset (Ovarian Cancer)
- II. Pre-Processing of Data
- III. Dataset Augmentation
- IV. Fine Tuning of SOTA Architecture

3. Methods and Material

A. Image Dataset

In this particular investigation, the collection process involved acquiring 85 images devoid of cancerous cells from a total of 42 patients receiving medical treatment at Smt. Kashibai Navale Medical College & General Hospital located in Pune, India. The challenge encountered in this endeavor stemmed from the scarcity of patients in the hospital's database who had undergone treatment for types of ovarian cancer (OC) other than Serous carcinoma. Proficient pathologists were responsible for the procurement of patient samples, subsequent slide preparation in their laboratory, and utilization of a Leica ICC50 microscopic camera to capture stained cell samples. Significantly, this cutting edge 5-megapixel camera enables the real-time transmission of high-definition images to smartphones and laptops. Correspondingly, carcinoma images representative of each OC subtype were obtained from a publicly accessible resource known as The Cancer Repository, as depicted in Table 1. It is important to highlight that registration on their respective websites is required for open access to the Cancer Repository. In order to align the dataset with the latest state-of-the-art (SOTA) models used for ovarian cancer prediction, the RGB images underwent uniform resizing as an essential

preprocessing step. In its entirety, the dataset consists of images containing carcinomas associated with OC, encompassing malignant tumors across four distinct subtypes. Additionally, it includes non-carcinoma images containing normal tissues and benign images of non-tumor glandular tissues. The primary goal of this dataset is to enable the automated classification of OC using histopathology images, with a particular focus on the subtypes characterized as carcinomas.

Table I : Image Dataset details

Class	Original Images	Augmented Images
Serous	175	5640
Mucinous	100	5223
Endometrioid	60	4353
Clear Cell	80	4999
Non-Cancerous	85	4527
Total	500	24742

B. Data Augmentation

Data augmentation refers to a collection of methods that increase the quantity and improve the quality of training datasets for Deep Learning, thereby enabling the creation of superior deep learning models.

The recent advancements in deep learning technology have been propelled by improvements in deep network architectures, high-performance computing, and access to large datasets. Deep convolutional neural networks (CNNs) have exhibited exceptional accomplishments across a range of computer vision tasks, encompassing image classification, object detection, and image segmentation. However, the generalizability of deep learning models remains a challenging issue, which refers to the performance gap between a model's performance on training data and unseen testing data. Models with poor generalizability are typically overfitted to the training data. To overcome this challenge and build effective deep learning models, Data Augmentation has emerged as a powerful technique that can minimize the gap between the training and testing sets by providing a more extensive set of possible data points and reducing overfitting.

Types of Image Data Augmentation techniques are Vertical shift, Horizontal shift, Vertical flip, Horizontal flip, Rotation, Brightness adjustment, and Zoom In/Out.

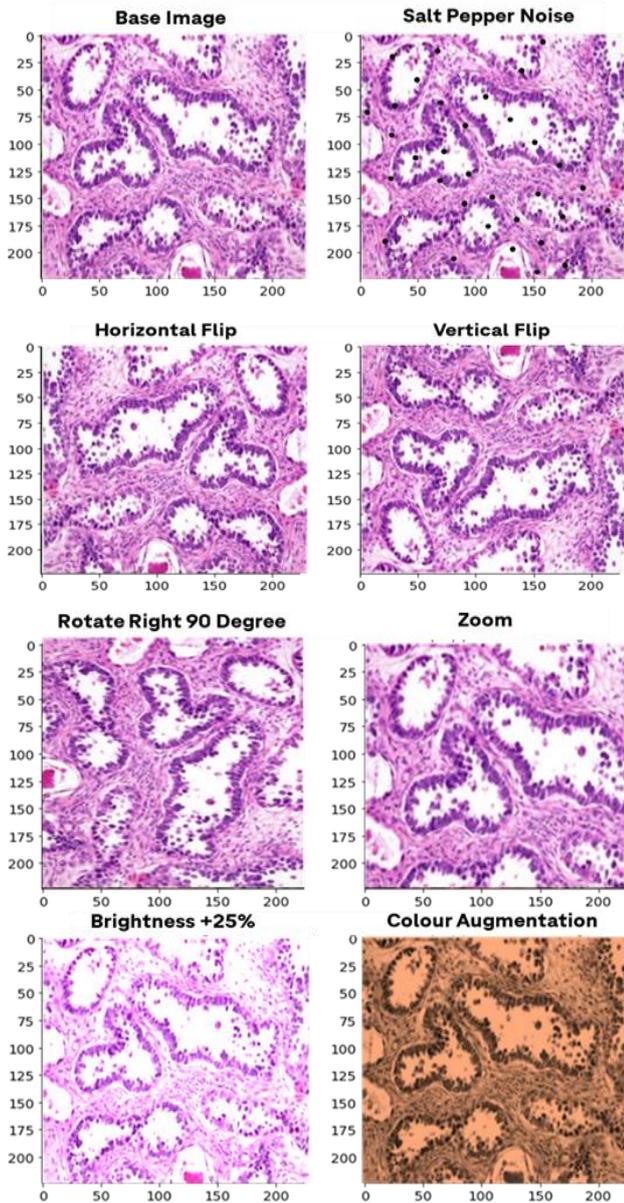


Fig. 1 : Data augmentation techniques

4. Implementation of Deep Learning Models

Our dataset was utilized to evaluate the performance of state-of-the-art methods, namely AlexNet and VGG-19, and a comparative analysis of the results was conducted. Now, let's delve into a discussion and comparison of the architectural aspects of these two cutting-edge models, AlexNet and VGG-19.

A. Architecture of AlexNet

Table II Summary of AlexNet model

	Layers	Feature Map	Size	Kernel Size	Stride	Activation Function
Input	RGB Image	1	227x227x3	-	-	-
1	Convolution Layer	96	55x55x96	11x11	4	ReLU
	Maxpooling Layer	96	27x27x96	3x3	2	ReLU

In 2012, a pioneering DCNN architecture known as AlexNet was unveiled. This architecture, created by Alex Krizhevsky for the ImageNet Large Scale Visual Recognition Challenge (LSVRC), consisted of five convolutional layers, three max-pooling layers, two fully connected layers, and one softmax layer. The model was developed with an extensive parameter count of 60 million to address the classification task involving a vast collection of over 1.2 million high-resolution images from the ImageNet LSVRC-2010 dataset, encompassing more than 1000 distinct categories.

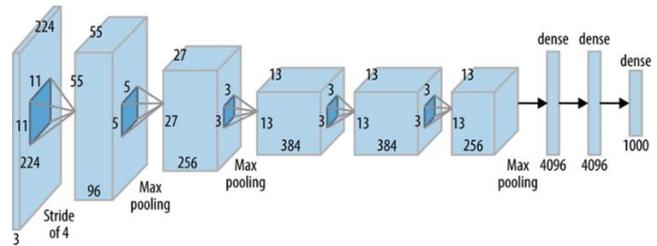


Fig. 2 : The architecture of AlexNet

B. Layers of AlexNet Architecture

Convolutional Kernels, known for their ability to employ multiple filters, play a crucial role in extracting valuable features from images. Typically, a single convolutional layer encompasses multiple kernels of identical size. As an illustration, in the case of AlexNet, the initial Convolutional Layer comprises 96 kernels, each with dimensions of 11x11x3. Here, the width and height of the kernel are typically equal, while its depth corresponds to the number of channels present.

In the architecture, Overlapping Max Pooling layers follow the first two Convolutional layers. The third, fourth, and fifth convolutional layers are directly connected. Subsequently, an Overlapping Max Pooling layer is introduced after the fifth convolutional layer, and its output is then directed into two fully connected layers. The second fully connected layer serves to transmit 1000 class labels into a SoftMax classifier.

Throughout the network, ReLU nonlinearity is applied after each convolutional and fully connected layer. In particular, ReLU nonlinearity is implemented following the first and second convolutional layers, while a local normalization step is conducted prior to pooling.

2	Convolution Layer	256	27x27x256	5x5	1	ReLU
	Maxpooling Layer	256	13x13x256	3x3	2	ReLU
3	Convolution Layer	384	13x13x384	3x3	1	ReLU
4	Convolution Layer	384	13x13x384	3x3	1	ReLU
5	Convolution Layer	256	13x13x256	3x3	1	ReLU
	Maxpooling Layer	256	6x6x256	3x3	2	ReLU
6	Fully Connected Layer	-	9216	-	-	ReLU
7	Fully Connected Layer	-	4096	-	-	ReLU
8	Fully Connected Layer	-	4096	-	-	ReLU
Output	Fully Connected Layer	-	1000	-	-	SoftMax

As mentioned above, the model was trained using 24,742 images, and the validation set comprised 1000 images. Our neural network architecture has 62,422,549 (62.4M) parameters of which 19,600 were non-trainable. We achieved a training accuracy of 70% and validation Accuracy of 55.5%.

C. Architecture of VGG-19

VGG19, an adapted iteration of the VGG model, consists of a comprehensive architecture with a total of 19 layers. These layers include 16 convolution layers, 3 fully connected layers, 5 max-pooling layers, and 1 softmax layer. Moreover, there are other variations of the VGG model, such as VGG11, VGG16, and several others. In terms of computational complexity, VGG19 involves approximately 19.6 billion floating-point operations (FLOPs).

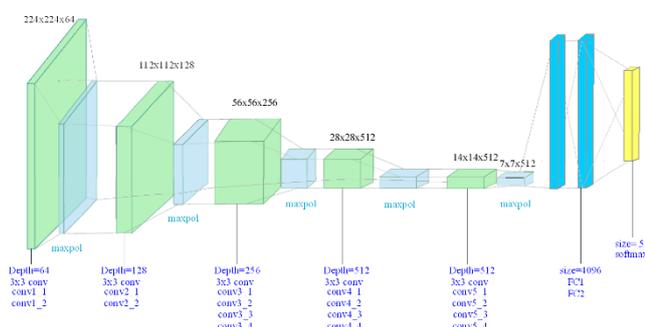


Fig. 3 : Architecture of VGG-19

- 1) This network accepted a fixed-size (224 * 224) RGB image as input, resulting in a matrix with a shape of (224,224,3).
- 2) The sole preprocessing step conducted involved calculating the average RGB value for each pixel

within the entirety of the training set.

- 3) The use of (3 * 3) sized kernels with a stride size of 1 pixel allowed for the entire image to be covered.
- 4) The spatial padding was utilized to preserve the image's spatial resolution.
- 5) Max pooling was applied by using a stride size of 2 over a window of 2x2 pixels.
- 6) To enhance both classification accuracy and computational efficiency, a Rectified Linear Unit (ReLU) was employed to introduce non-linearity to the model. This approach proved significantly more effective compared to previous models that relied on sigmoid or tanh functions.
- 7) The implemented model consisted of three fully connected layers, where the initial two layers had a size of 4096, while the subsequent layer contained 1000 channels dedicated to classification employing the 1000-way ILSVRC. The third layer employed a softmax function.

Table III Summary of VGG-19 model

Layer	Patch Size	Input Size
Convolutional Layer×2	3×3/1	3×224×224
Maxpooling Layer	2×2	64×224×224
Convolutional Layer×2	3×3/1	64×112×112
Maxpooling Layer	2×2	128×112×112

Layer	Patch Size	Input Size
Convolutional Layer×4	3×3/1	128×56×56
Maxpooling Layer	2×2	256×56×56
Convolutional Layer×4	3×3/1	256×28×28
Maxpooling Layer	2×2	512×28×28
Convolutional Layer×4	3×3/1	512×14×14
Maxpooling Layer	2×2	512×14×14
Fully Connected	128	512x7x7
Fully Connected	Softmax	5

5. Result Analysis and Discussion

While Training Model, the accuracy keeps increasing per epoch and the loss keeps decreasing as seen in Fig. 4. Also, early stopping at best accuracy has limited the number of epochs to 10 while the manual configuration was to run for 15 epochs.



Fig. 4 : Various graphs of AlexNet model training process

From above Fig. 4 in graph (a) shown Epoch vs training Accuracy/Training Loss, ideally training Loss should be zero. AlexNet shows 0.075 loss at the accuracy of 0.70 and validation accuracy of 0.025.

In Fig. 4(b) shows Epoch vs validation Accuracy /validation loss.as the Epoch increases validation accuracy increase and validation accuracy loss decreases. while training the model, at epoch 9 validation loss is 0.02616 at a validation accuracy of 0.02544.

In Fig. 4(c) shows Epoch VS Training loss and validation loss from graph 4(c) it can be observed that validation loss is approximately equal to the Training Loss. Fig. 4 (d) shows Epoch VS Training Accuracy and validation Accuracy. Fig. 4 graph (d) shows that as Training Accuracy increases validation accuracy also increases so model performance also increases.

A.VGG-19

From above Fig. 5 in graph (a) shown Epoch vs training Accuracy/Training Loss, ideally training Loss should be zero. VGG-16 shows 0.0226 loss at the accuracy of 0.90 and validation accuracy of 0.87.

In Fig. 5 (b) shows Epoch vs validation Accuracy /validation loss.as the Epoch increases validation accuracy increase and validation accuracy loss decreases. while training the model, at epoch 11 validation loss is 0.0217 at a validation accuracy of 0.91.

In Fig. 5 (c) shows Epoch VS Training loss and validation loss from graph (c) it can be observe that validation loss is approximately equal to the Training Loss. In Fig. 5 (d) shows Epoch VS Training Accuracy and validation Accuracy. Fig. 5 graph (c) shows that as Training Accuracy increases validation accuracy also increases so model performance also increases.



Fig. 5 : Various graphs of VGG-19 model training process

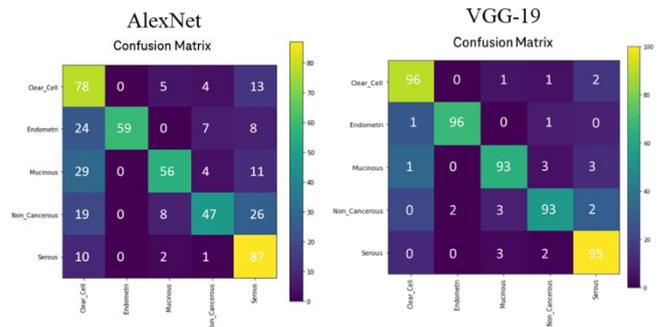


Fig. 6 : Confusion matrix of VGG-19 and AlexNet models

A confusion matrix is a valuable tool for assessing the effectiveness of a deep learning model. It provides a tabular representation that presents the count of accurate and inaccurate predictions made by the model, enabling the identification of misclassified classes. This information can be leveraged to enhance the model's overall accuracy. As depicted in Figure 6, VGG-19 exhibits a true positive rate exceeding 90% when compared to AlexNet.

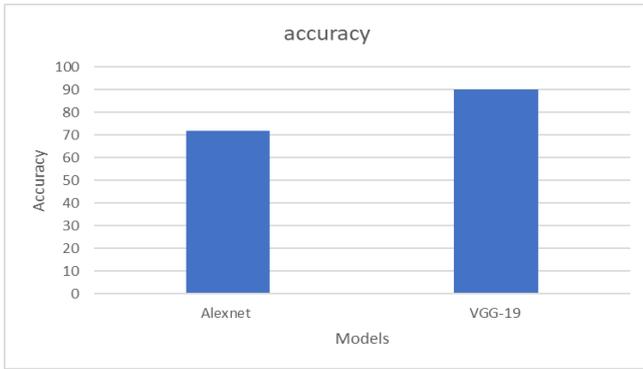


Fig. 7 : Comparison of Accuracy for AlexNet and VGG-19

Accuracy is a commonly employed performance metric in deep learning, serving as an indicator of a model's capacity to make accurate predictions on a given dataset. It is determined by calculating the ratio of accurate predictions to the total number of predictions made. The comparative results depicted in Figure 7 highlight the superior accuracy of VGG-19 in comparison to AlexNet.

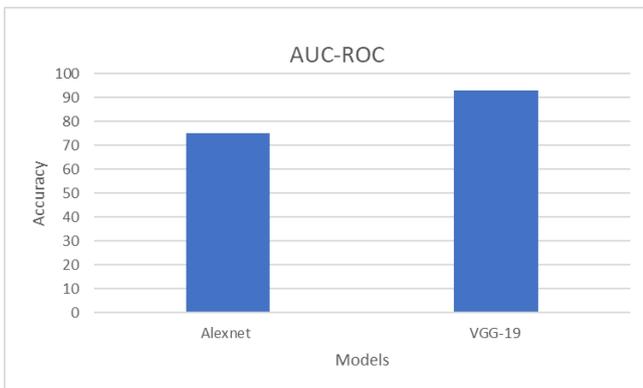


Fig. 8 : Comparison of AUC-ROC curve for AlexNet and VGG-19 models

Figure 8 displays the AUC-ROC curve of both AlexNet and VGG-19 architectures. This curve serves as a visual representation of the performance of a deep learning multiclass classification model. It provides a metric for evaluating the accuracy of the model in distinguishing between positive and negative cases. A higher AUC-ROC score indicates better performance. Fig. 8 shows the comparative analysis of AlexNet and VGG-19.

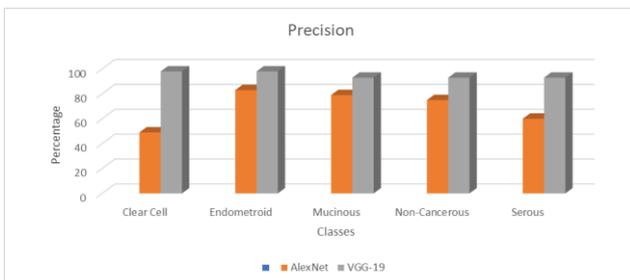


Fig. 9 : Comparison of Precision of VGG-19 and AlexNet models

Figure 9 showcases a comparative analysis of the Precision

metric, which evaluates the model's ability to generate accurate positive predictions. Precision is determined by dividing the number of true positive predictions by the total number of positive predictions made by the model. A higher precision score for VGG-19 indicates more precise positive predictions, while a lower precision score suggests a higher occurrence of false positive predictions, as illustrated in Figure 9.

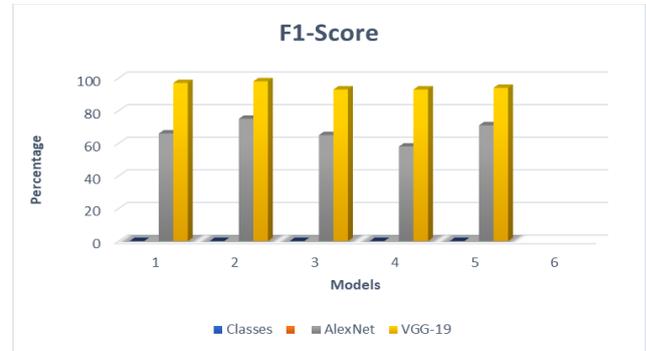


Fig. 10 : Comparison of F1-Score of VGG-19 and AlexNet models

Figure 10 presents the F1 score comparison between the AlexNet and VGG-19 models. The F1 score combines precision and recall to provide an evaluation metric ranging from 0 to 1, with 1 representing optimal performance. A higher F1 score indicates better precision and recall, indicating a lower rate of false positives and false negatives, respectively. In essence, a higher F1 score signifies that the model can accurately predict both positive and negative cases. Thus, achieving a high F1 score is desirable as it demonstrates the model's accuracy and reliability. Notably, Figure 10 illustrates VGG-19 outperforming the AlexNet model across all five classes.

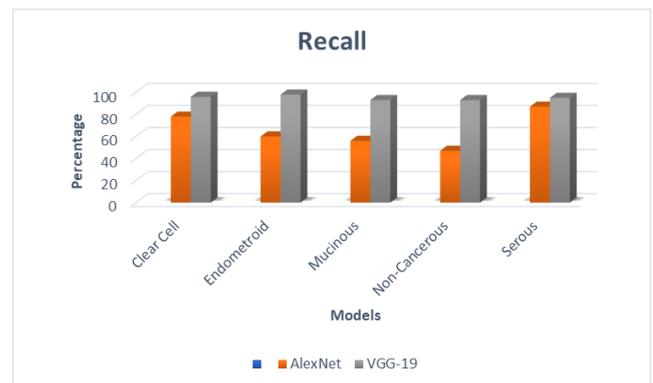


Fig. 11 : Comparison of Recall of AlexNet and VGG-19 models

Figure 11 depicts the Comparison of Recall between the AlexNet and VGG-19 models. Recall is a crucial performance metric in deep learning that measures the model's effectiveness in correctly identifying all relevant instances within a dataset. It is computed by dividing the number of true positive predictions by the total number of actual positive cases. As illustrated in the figure, VGG-19

demonstrates a notable recall score across all classes. A high recall score is desirable as it indicates the model's capability to accurately identify a significant proportion of relevant cases.

6. Conclusion

This study demonstrates the effectiveness of deep learning models in predicting and classifying subtypes of ovarian cancer from histopathological images. Specifically, we compared the performance of two popular deep convolutional neural network architectures (DCNN), AlexNet and VGG-19, and found that VGG-19 outperformed AlexNet with an accuracy of 90% compared to 70%. This indicates that VGG-19 is a better model for accurately diagnosing and classifying ovarian cancer subtypes.

Apart from accuracy, the authors conducted an analysis of additional performance metrics including recall, precision, F1-score, and AUC-ROC, which serve to affirm the superiority of VGG-19 in comparison to AlexNet. These metrics offer a comprehensive evaluation of the models' performance, highlighting the potential of deep learning to enhance medical diagnosis.

References

- [1] "Cancer," World Health Organization, <https://www.who.int/news-room/factsheets/detail/cancer> (accessed June 2, 2022).
- [2] "Colorectal cancer facts & figures 2020-2022 - American cancer society," <https://www.cancer.org/content/dam/cancer-org/research/cancer-facts-and-statistics/colorectal-cancer-facts-and-figures/colorectal-cancer-facts-and-figures-2020-2022.pdf> (accessed May 2, 2023).
- [3] L. A. Torre *et al.*, "Ovarian cancer statistics, 2018," *CA: A Cancer Journal for Clinicians*, vol. 68, no. 4, pp. 284–296, Jul. 2018, doi: 10.3322/caac.21456.
- [4] G. Chornokur, E. K. Amankwah, J. M. Schildkraut, and C. M. Phelan, "Global ovarian cancer health disparities," *Gynecologic Oncology*, vol. 129, no. 1, pp. 258–264, Apr. 2013, doi: 10.1016/j.ygyno.2012.12.016.
- [5] W. Sun, B. Zheng, and W.-J. Qian, "Automatic feature learning using multichannel ROI based on deep structured algorithms for computerized lung cancer diagnosis," *Computers in Biology and Medicine*, vol. 89, pp. 530–539, Oct. 2017, doi: 10.1016/j.compbiomed.2017.04.006.
- [6] A. Das, U. R. Acharya, S. R. Panda, and S. Sabut, "Deep learning based liver cancer detection using watershed transform and Gaussian mixture model techniques," *Cognitive Systems Research*, vol. 54, pp. 165–175, May 2019, doi: 10.1016/j.cogsys.2018.12.009.
- [7] S. Wang *et al.*, "Deep learning provides a new computed tomography-based prognostic biomarker for recurrence prediction in high-grade serous ovarian cancer.," *Radiotherapy and Oncology*, vol. 132, pp. 171–177, Mar. 2019, doi: 10.1016/j.radonc.2018.10.019.
- [8] Y. Feng *et al.*, "A Deep Learning Approach for Targeted Contrast-Enhanced Ultrasound Based Prostate Cancer Detection," *IEEE/ACM Transactions on Computational Biology and Bioinformatics*, vol. 16, no. 6, pp. 1794–1801, Nov. 2019, doi: 10.1109/tcbb.2018.2835444.
- [9] S. Marcišauskas, B. Ulfenborg, B. Kristjansdottir, S. Waldemarson, and K. Sundfeldt, "Univariate and classification analysis reveals potential diagnostic biomarkers for early stage ovarian cancer Type 1 and Type 2," *Journal of Proteomics*, vol. 196, pp. 57–68, Mar. 2019, doi: 10.1016/j.jprot.2019.01.017.
- [10] A. Dascalu and E. O. David, "Skin cancer detection by deep learning and sound analysis algorithms: A prospective clinical study of an elementary dermoscope," *EBioMedicine*, vol. 43, pp. 107–113, May 2019, doi: 10.1016/j.ebiom.2019.04.055.
- [11] Z. Alyafeai and L. Ghouti, "A fully-automated deep learning pipeline for cervical cancer classification," *Expert Systems With Applications*, vol. 141, p. 112951, Mar. 2020, doi: 10.1016/j.eswa.2019.112951.
- [12] P. Kaur, G. Singh, and P. Kaur, "Intellectual detection and validation of automated mammogram breast cancer images by multi-class SVM using deep learning classification," *Informatics in Medicine Unlocked*, vol. 16, p. 100151, Jan. 2019, doi: 10.1016/j.imu.2019.01.001.
- [13] Z. Liu, C. Yao, H. Yu, and T. Wu, "Deep reinforcement learning with its application for lung cancer detection in medical Internet of Things," *Future Generation Computer Systems*, vol. 97, pp. 1–9, Aug. 2019, doi: 10.1016/j.future.2019.02.068.
- [14] H. V. Pham, M. Futakuchi, A. Bychkov, T. Furukawa, K. Kuroda, and J. Fukuoka, "Detection of Lung Cancer Lymph Node Metastases from Whole-Slide Histopathologic Images Using a Two-Step Deep Learning Approach," *American Journal of Pathology*, vol. 189, no. 12, pp. 2428–2439, Sep. 2019, doi: 10.1016/j.ajpath.2019.08.014.
- [15] P. M. Shakeel, M. A. Burhanuddin, and M. I. Desa, "Lung cancer detection from CT image using

improved profuse clustering and deep learning instantaneously trained neural networks,” *Measurement*, vol. 145, pp. 702–712, Oct. 2019, doi: 10.1016/j.measurement.2019.05.027.

- [16] Gayther, S. A., Russell, P., Harrington, P., Antoniou, A. C., Easton, D. F., & Ponder, B. A. (1999). The contribution of germline BRCA1 and BRCA2 mutations to familial ovarian cancer: no evidence for other ovarian cancer-susceptibility genes. *American journal of human genetics*, 65(4), 1021–1029. <https://doi.org/10.1086/302583>
- [17] La Vecchia, C., Tavani, A., Franceschi, S., Levi, F., Corrao, G., & Negri, E. (1997). Epidemiology and prevention of oral cancer. *Oral oncology*, 33(5), 302–312. [https://doi.org/10.1016/s1368-8375\(97\)00029-8](https://doi.org/10.1016/s1368-8375(97)00029-8)
- [18] Ebell, M. H., Culp, M. B., & Radke, T. J. (2016). A Systematic Review of Symptoms for the Diagnosis of Ovarian Cancer. *American journal of preventive medicine*, 50(3), 384–394. <https://doi.org/10.1016/j.amepre.2015.09.023>
- [19] Hawkins, Thomas R. Jr. Symptoms of Ovarian Cancer. *Obstetrics & Gynecology* 98(6):p 1150, December 2001.
- [20] Roett, M. A., & Evans, P. (2009). Ovarian cancer: an overview. *American family physician*, 80(6), 609–616.
- [21] Epithelial ovarian cancer. Hoffman B.L., & Schorge J.O., & Bradshaw K.D., (2016). *Williams Gynecology*, 3e. McGraw Hill. <https://accessmedicine.mhmedical.com/content.aspx?bookid=1758§ionid=118173184>
- [22] Kumari, Suchitra. (2018). Serum Biomarker Based Algorithms in Diagnosis of Ovarian Cancer: A Review. *Indian Journal of Clinical Biochemistry*. 33. 1-5. [10.1007/s12291-018-0786-2](https://doi.org/10.1007/s12291-018-0786-2).
- [23] Van Gorp, T., Cadron, I., Despierre, E., Daemen, A., Leunen, K., Amant, F., Timmerman, D., De Moor, B., & Vergote, I. (2011). HE4 and CA125 as a diagnostic test in ovarian cancer: prospective validation of the Risk of Ovarian Malignancy Algorithm. *British journal of cancer*, 104(5), 863–870. <https://doi.org/10.1038/sj.bjc.6606092>.