Iraqi Journal for Computer Science and Mathematics

Manuscript 1251

Advancing Lymphoma Diagnosis in Histopathology Image Classification Using Multi Deep Learning Models

Ahmed Elaraby

Andrey Nechaevskiy

Aymen Saad

IRAQI JOURNAL FOR COMPUTER SCIENCE AND MATHEMATICS 2025;6:180–187

ORIGINAL STUDY

Advancing Lymphoma Diagnosis in Histopathology Image Classification Using Multi Deep Learning Models

Ahmed Elaraby [a,b], **Andrey Nechaevskiy** [c], **Aymen Saad** [d,e,*]

[a] Department of Computer Science, Faculty of Computers and Information, South Valley University, Qena, 83523, Egypt
[b] Researcher, Academy of Scientific Research & Technology (ASRT), Cairo, Egypt
[c] Meshcheryakov Laboratory of Information Technologies, Joint Institute for Nuclear Research, Joliot-Curie 6, Dubna, 141980, Russia
[d] Department of Information Technology, Management Technical College, Al-Furat Al-Awsat Technical University, Kufa, 54003, Iraq
[e] School of Electrical Engineering, Universiti Teknologi Malaysia, 81310 UTM Skudai, Johor, Malaysia

ABSTRACT

Deep learning's rapid development is generating significant interest in its potential to improve medical imaging. It has shown promising results in detecting malignant lymphoma in histopathology medical images. Image classification methods are widely used to aid in making diagnoses from medical images. In recent years, deep learning methods have achieved high performance in detecting malignant lymphoma in histopathology images. This study proposes a novel approach to improving lymphoma diagnosis in histopathology images called the Lightweight Convolutional Neural Network (LWCNN). The proposed LWCNN model comprises multiple deep learning architectures, including a convolutional neural network (CNN) that has been trained to classify lymphoma subtypes based on histopathology images using ResNet50 and MobileNetV2. The LWCNN model aggregates the predictions of disparate architectures to arrive at a definitive diagnosis, leveraging the unique capabilities of each. A comprehensive dataset of annotated lymphoma histopathology images was assembled for the purpose of training the multi-deep learning model. To ensure a representative and diverse training set, the dataset was meticulously curated to encompass various subtypes of lymphoma. Performance evaluation of the proposed deep learning model for lymphoma classification using standard metrics revealed the following accuracies: LWCNN (97.34% training, 86.71% testing), ResNet50 (88.76% training, 86.37% testing), and MobileNetV2 (88.47% training, 86.60% testing). These results indicate that the LWCNN model significantly surpasses existing approaches in diagnostic accuracy.

Keywords: Deep learning, Lightweight Convolutional Neural Networks (LWCNN), Image processing, Histopathology, Lymphoma diagnosis

1. Introduction

Lymphoma, characterized by the abnormal proliferation of lymphocytes, is a heterogeneous group of malignancies affecting the lymphatic system [3]. This cancer poses significant diagnostic challenges due to its diverse subtypes, histopathological variations, and overlapping features with other hematological disorders. Accurate diagnosis of lymphoma is essential to guide appropriate treatment strategies and optimize patient outcomes [21].

As lymphocytes have different physiological immune functions depending on their origin and stage of differentiation, the classification of lymphomas arising from these normal populations of lymphoid cells is difficult. In medicine, imaging is a convenient and important tool for diagnosing and deciding on treatment in clinical practice [25]. Pileri et al., [18]

Received 20 December 2024; revised 15 March 2025; accepted 16 March 2025.
Available online 9 May 2025

* Corresponding author.
E-mail addresses: ahmed.alaraby@svu.edu.eg (A. Elaraby), nechav@jinr.ru (A. Nechaevskiy), aymen.abdalameer@atu.edu.iq (A. Saad).

https://doi.org/10.52866/2788-7421.1251

highlights the need for accurate diagnosis to select the optimal treatment and predict the outcome of the disease. They note that modern diagnostic approaches make it possible to achieve high accuracy and reliability. Conventional clinical methods of diagnosis, such as tumor biopsy and percutaneous puncture aspiration, have disadvantages, including high costs and insufficient classification accuracy [17].

Deep learning methods are used to accurately diagnose lymphoma based on the analysis of whole-slide histopathological images [21]. The authors demonstrate the effectiveness of these methods in recognizing different types of lymphoma based on histological examination data. Special attention is paid to issues of diagnostic quality and accuracy, as well as the possibilities of integrating such systems into clinical practice. The results show the promising application of deep learning in medicine, especially in the context of diagnosing complex diseases such as lymphoma.

It is important to highlight the value of AI in automating and enhancing the accuracy of diagnostic procedures for screening, predicting, and managing diabetic retinopathy [2]. This can significantly improve the quality of patient care [8]. The application of deep neural networks to classify skin cancers has demonstrated high accuracy and the potential to reduce diagnostic time and the burden on clinicians [4]. Machine learning (ML) [12, 13] and deep learning [15] have shown good results in detecting malignant lymphoma.

Artificial intelligence (AI) offers a robust tool for enhancing the efficacy and precision of medical diagnosis, enabling physicians to prioritize the most critical elements of treatment and elevate the overall standard of care. From the perspective of lymphoma diagnosis, deep learning offers opportunities to enhance the accuracy, efficiency, and reproducibility of histopathological image analysis, thereby improving clinical practice [11, 19].

2. Related works

Automatic classification of lymphoma pathology images is essential for clinical management. The lack of publicly available datasets and the complex characteristics of pathological images make it difficult to advance research in automatic classification technology. Recently, many solutions have been proposed for the automatic classification of lymphoma pathology images. Despite their effectiveness, classical image processing methods often face challenges when processing the increasing volume and complexity of medical image data.

These efforts have led to the development of new algorithms capable of accurately distinguishing different subtypes of lymphoma and identifying subtle morphological features indicative of disease progression. Deep learning (DL) [23] is a machine learning (ML) technique that utilizes a large amount of data to train a deep neural network for pattern recognition and the acquisition of predictive properties. One of the DL architectures, the convolutional neural network (CNN), enhances the effectiveness of deep learning in image-related fields [20]. Deep learning models based on CNNs have already been developed for various medical image classification tasks, such as the classification of melanoma [4], breast cancer, diabetic retinopathy [6, 7, 22], and lymphoma [21].

Miyoshi et al. [15] proposed using deep learning to diagnose malignant lymphoma, achieving high accuracy at different magnification levels of histopathological images. Similarly, Zhang et al. [24] compared the performance of neural network architectures and demonstrated the superiority of a neural network optimized by a genetic algorithm in classifying histopathological images of malignant lymphoma.

The integration of predictive models based on deep learning into clinical practice opens up significant opportunities for improving patient outcomes. Hashimoto et al. [10] applied a deep learning model based on multifactorial learning to diagnose subtypes of malignant lymphoma and demonstrated increased accuracy in diagnosing typical cases compared to atypical cases. Other researchers have explored the use of surface-enhanced Raman scattering (SERS) and convolutional neural networks (CNNs) to diagnose malignant lymphoma, achieving notable success in distinguishing between normal and malignant DNA samples. These advances pave the way for more accurate patient stratification and treatment decisions, ultimately leading to improved clinical outcomes.

Gaidano et al. [5] evaluated several machine learning systems for the classification of B-cell non-Hodgkin lymphoma (B-NHL), achieving high accuracy rates and demonstrating the potential of artificial intelligence in oncology. Similarly, Lisson et al., [14] compared ML/DL approaches to detect mantle cell lymphoma (MCL), highlighting the superior performance of enhanced 3D CNN networks in diagnosing pathological images of MCL.

Recent research emphasizes the importance of using deep learning models to improve lymphoma diagnosis by classifying histopathological images. By integrating multiple neural network architectures and employing ensemble learning techniques, researchers have demonstrated improved classification efficiency and generalization across different patient

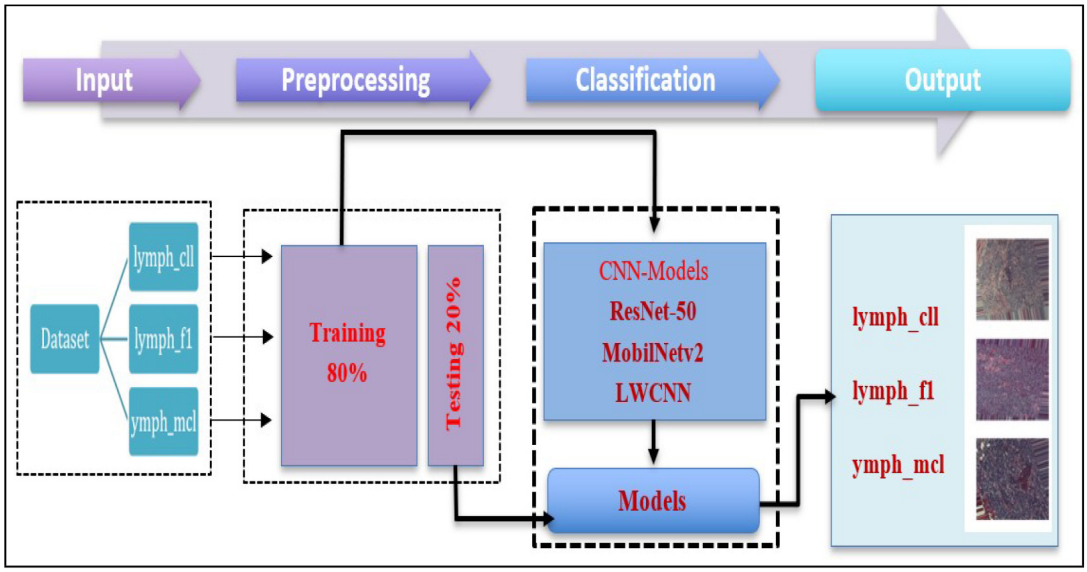

Fig. 1. Overview of methodology.

populations. The synergy among various deep learning models enables comprehensive feature extraction and reliable decision-making, ultimately leading to more dependable diagnostic results. Additionally, the integration of deep learning methods with advanced image processing techniques has further enhanced the ability to diagnose lymphoma from histopathological images. Methods such as transfer learning, data augmentation, and attention mechanisms have been employed to increase the reliability of the model, reduce overfitting, and improve the interpretability of the classification results. These studies illustrate the broad application of advanced methods and approaches based on deep learning to achieve greater accuracy and effectiveness in lymphoma diagnosis.

3. Dataset and methodology

In this paper, four stages are adopted as a methodology: obtaining the dataset, preprocessing, classification, and output. The steps of the adopted methodology are presented in detail in Fig. 1. The implemented framework for lightweight deep learning-based malignant lymphoma image classifiers is executed using Python 3 with computer specifications, 16 GB DDR4 of RAM, CPU AMD Ryzen 5 3550H with Radeon Vega Mobile GFX 2.10 GHz and operating system Windows 11.

3.1. Datasets

In this study, we evaluated the systems using biopsies that were mixed with hematoxylin and eosin (H&E) solutions on liquid-based cytology (LBC) slides. These slides were subsequently converted into whole slide images (WSI) under a microscope. The dataset comprises a total of 15,000 WSI images, meticulously divided into three categories of malignant lymphoma: Follicular lymphoma (FL), Chronic lymphocytic leukemia (CLL), and Mantle cell lymphoma (MCL) [9]. Researchers and interested parties can access this dataset on Kaggle, where it is provided in JPEG format at a resolution of 512 × 512 pixels. Fig. 2 illustrates examples from the dataset, showcasing three randomly selected classes to highlight the diversity and characteristics of the images.

Table 1. Characteristics of the dataset.

Types of class	Abstract Name	Training	Testing	Sum
lymphomas_CLL	Lymph_CLL	4000	1000	5000
lymphomas_FL	Lymph_FL	4000	1000	5000
lymphomas_MCL	Lymph_MCL	4000	1000	5000
Total of images		12000	3000	15000

For the training and evaluation of the systems, the dataset was partitioned such that 80% of the images were allocated for training and testing purposes, while the remaining 20% were reserved specifically for system testing, following an 80:20 split. Given the balanced nature of the dataset, each class contains 5,000 images. Consequently, the systems divided each class into 4,000 images for training and 1,000 images for testing, as summarized in Table 1.

3.2. Our model (LWCNN)

We used five hidden layers in this model. Each layer contains layers (Batch Normalization, Leaky

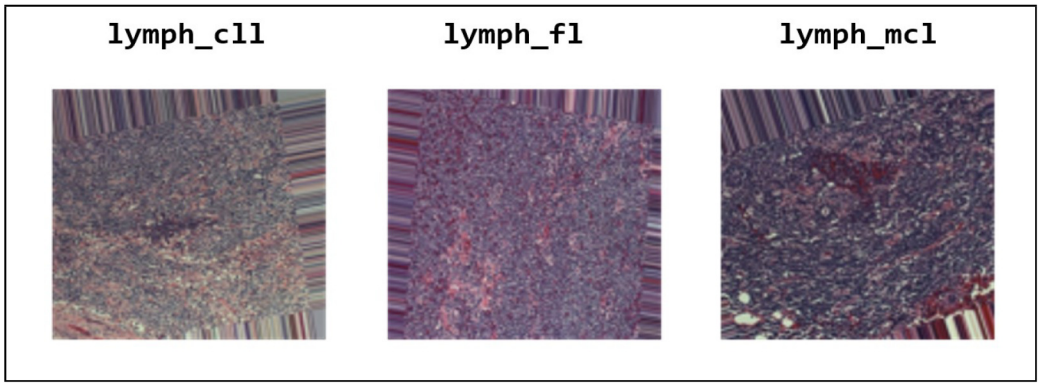

Fig. 2. Sample of datasets.

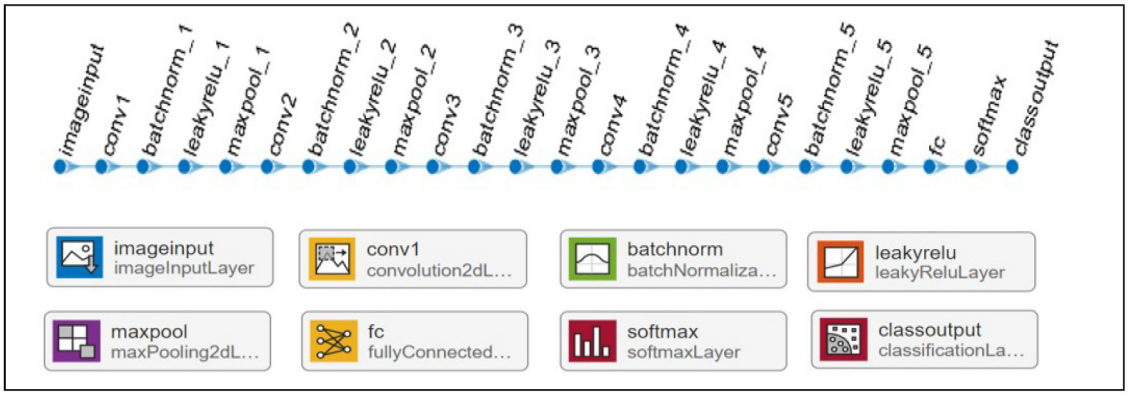

Fig. 3. LWCNN architecture.

Relu, Max Pooling layers) to deal with weight scattering and to extract important features from the pictures. The CNN containing these parts of layers is shown in Fig. 3 and its details are presented below in Table 2.

In this paper, we present enough to overcome the disadvantages of databases of feature extraction and high resolution. For this purpose, we use two pretrained models such as ResNet-50 and MobileNet V2. We develop an LWCNN model also. In all three experiments, all realized datasets are from the original size to $224 \times 224 \times 3$ to be performed using the suggested CNN model. Then, each class of the dataset is divided into 80% for training and 20% for testing. Finally, the dataset is used to train our CNN model.

4. Experimental results

This section presents the training results obtained using ResNet50, MobileNetV2 and LWCNN models. The total number of 15000 samples were extracted from the training dataset, where each class derived about 4000 images and 1000 images for the validation step besides 32 minimum batch sizes. The low

Table 2. Details of the proposed LWCNN model.

Name of layer	Decimation	# Of Filter	Padding	Stride
Input	224 224 3			
Conv1	33	8	same	
batch normalization				
leaky Relu	0.01	1		
max-pooling	22			22
Conv2	33	16	same	
batch normalization				
leaky Relu	0.01			
max-pooling	22			22
Conv3	33	32	same	
batch normalization				
leaky Relu	0.01			
max-pooling	22			22
Conv4	33	16	same	
batch normalization				
leaky Relu	0.01			
max-pooling	22			22
Conv5	33	8	same	
batch normalization				
leaky Relu	0.01			
max-pooling	22			22
Fully Connect	3			
Softmax	0 or 1 or 2			
Classification	Lymphomas_CLL OR lymphomas_FL OR lymphomas_MCL			

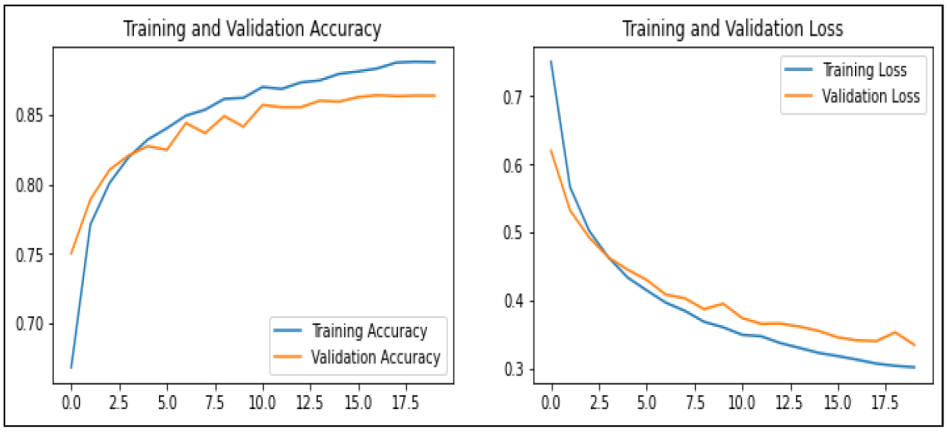

Fig. 4. Model-1-ResNet50.

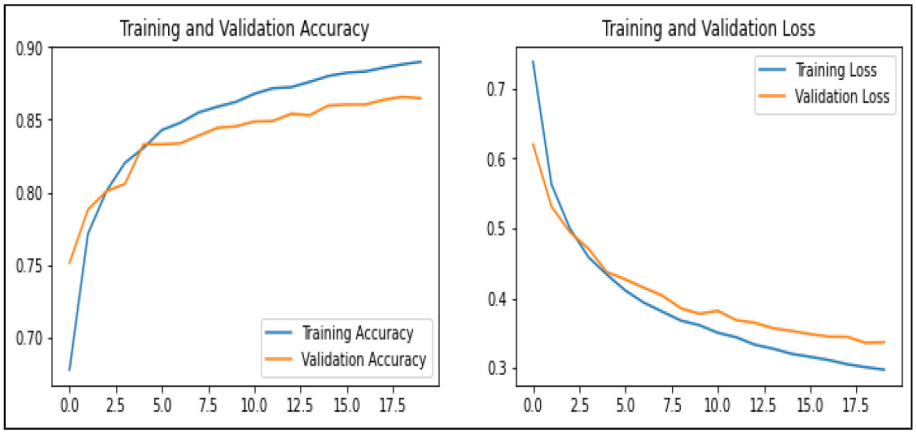

Fig. 5. Model-2-MobilNetV2.

Table 3. Illustrate results of models.

Model	AC-Training	AC-Validation	Time (min)
ResNet50	88.76	86.37	5.20
MobilNetV2	88.47	86.60	5.15
OUR_Proposal	97.34	86.71	4.30

complexity CNN model was trained on these images. In the CNN model, 20 epochs are used for training, which means that in the training phase, each type of data set is processed twenty times, as shown in Table 3.

For the ResNet50, MobileNetV2 and LWCNN models, accuracy and loss are observed over the course of training as shown in Figs. 4 to 6, respectively. Accuracy increases with the epoch until reaching a saturation level where the imprecision is very low and varies around a certain level. The loss declines, but decreases as the epochs decrease until saturated. To put it simply, accuracy determines goodness, and loss determines badness. Typically, high accuracy and low loss will characterize a good model.

In the validation dataset, the performance of the model is shown by a confusion matrix that can accurately predict the class used in the supervised learning process. Validating the proposed model involves predicting each of the two classes. Each column of the matrix represents the number of predictions for each class, while each row represents real class instances. The results are plotted in the confusion matrix as shown below in Figs. 7 to 9, ResNet50, MobileNetV2 and LWCNN models, respectively.

The performance measures used in this study are the most widely used metrics, shown below in Eqs. (1) to (5), are accuracy, sensitivity, specificity, and precision [1] which are given as:

$$\text{Accuracy} = \frac{TN + TP}{TP + FP + TN + FN} \tag{1}$$

$$\text{Sensitivity} = \frac{TP}{TP + FN} \tag{2}$$

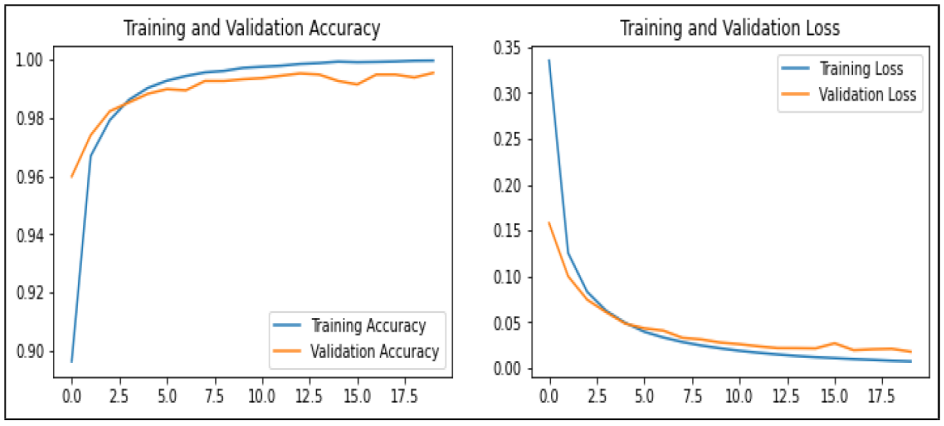

Fig. 6. Model-3-OUR_Proposed.

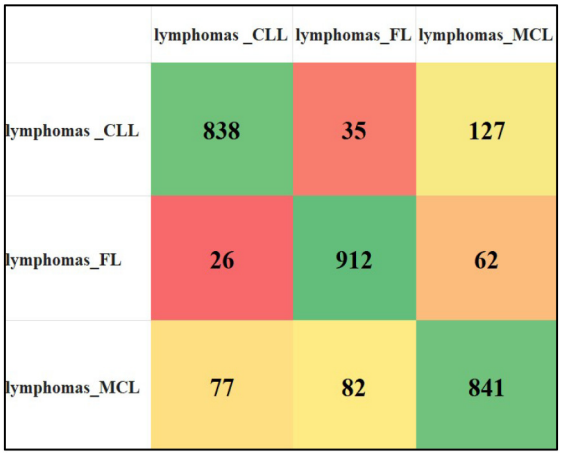

Fig. 7. Model-1-ResNet50.

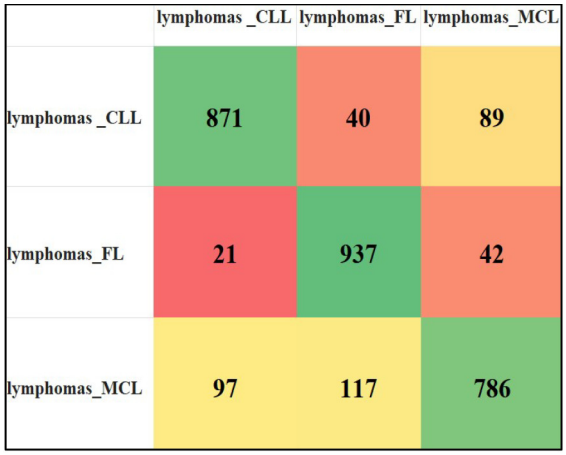

Fig. 8. Model-2-MobilNetV2.

$$\text{Specificity} = \frac{T\,N}{T\,N + F\,P} \tag{3}$$

$$\text{Precision} = \frac{TP}{TP + FP} \tag{4}$$

Table 4. Training results of experiment models and proposed framework.

Name model	Name of class	Ac	Pr	Se	Sp	Fm
ResNet50	lymphomas_CLL	88	82	84	90	83
	lymphomas_FL	95	89	97	93	93
	lymphomas_MCL	91	89	84	94	86
MobileNet-v2	lymphomas_CLL	91	86	86	93	86
	lymphomas_FL	93	86	97	91	91
	lymphomas_MCL	91	88	87	94	88
OUR_Proposal	lymphomas_CLL	90	80	90	89	85
	lymphomas_FL	97	94	97	97	95
	lymphomas_MCL	92	86	90	92	88

$$F\_measure = \frac{2\,(precision * sensitivity)}{precision\ +\ sensitivity} \tag{5}$$

Where, True positive (TP), False-positive (FP), True-negative (TN) and False-negative (FN).

Table 4 presents the performance metrics of three different deep learning models (ResNet-50, MobileNet-v2, and "OUR Proposal") in classifying four types of lymphomas: CLL (Chronic Lymphocytic Leukemia), FL (Follicular Lymphoma), and MCL (Mantle Cell Lymphoma). The metrics included are accuracy (Ac), precision (Pr), sensitivity (Se), specificity (Sp), and F-measure (Fm). The results show that the "OUR Proposal" model generally outperforms the other two pre-trained models, ResNet-50 and MobileNet-v2, across most of the performance metrics and lymphoma types. Specifically, the "OUR Proposal" model achieves the highest accuracy (92%) and F-measure (88%) for the MCL lymphoma type, indicating its superior ability to correctly classify this particular lymphoma subtype compared to the other models. The table provides a comprehensive comparison of the model performances, which can be useful for researchers and clinicians working on lymphoma classification tasks using deep learning techniques.

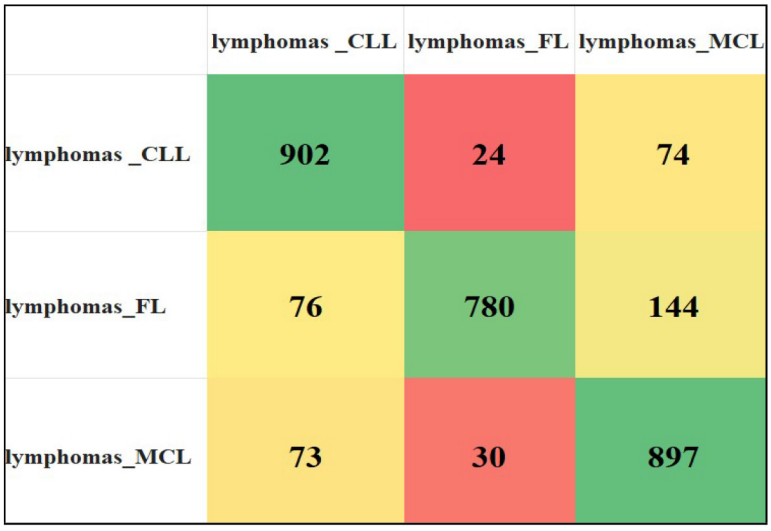

Fig. 9. Model-3-Proposed method.

5. Conclusion

Accurate diagnosis of lymphoma, a cancer of the lymphatic system, is essential for effective treatment planning and patient management. Deep learning techniques for various medical imaging tasks, including histopathology image classification, have shown great potential in recent years. In this study, we present a novel approach to improving lymphoma diagnosis in histopathology images by using a multi-deep learning model. Our proposed model, which is trained to classify different lymphoma subtypes based on histopathology images, incorporates multiple deep learning architectures, including convolutional neural networks (CNNs). By leveraging the complementary strengths of these architectures, the model combines their predictions to make a final diagnosis. We assembled a large dataset of annotated lymphoma histopathology images to train the multi-deep learning model. To ensure the dataset is representative and robust for training purposes, it was carefully curated to include various lymphoma subtypes. Evaluation using standard performance measures showed that the proposed LWCNN model achieved a remarkable 97.34% training accuracy and 86.71% testing accuracy in lymphoma classification, exceeding the performance of ResNet50 (88.76% training, 86.37% testing) and MobileNetV2 (88.47% training, 86.60% testing). The study highlights the potential of our novel deep learning model to improve lymphoma diagnosis in histopathological imaging. Harnessing the power of deep learning and combining multiple architectures can enhance the accuracy and efficiency of lymphoma classification, leading to better treatment decisions and patient outcomes. Future studies could focus on expanding the dataset to include a more diverse range of lymphoma subtypes and patient demographics. This would enhance the model's robustness and generalizability across different populations and variations of the disease. Collaborating with multiple institutions to gather a larger, multicentric dataset could be particularly beneficial.

Acknowledgments

The authors would like to thank ASRT Egypt (Academy of Scientific Research and Technology) for Supporting this work under ASRT-JINR Research Project Fund (JINR Research Project_202324618).

References

1. E. Alabdulkreem, H. Elmannai, A. Saad, I. S. Kamil, and A. Elaraby, "Deep learning-based classification of melanoma and non-melanoma skin cancer," *Traitement du Signal*, vol. 41, No. 1, pp. 213–223, 2024, https://doi.org/10.18280/ts.410117.
2. T. Ali Abdalkareem, K. A. Zidan, & A. S. Albahri, "A systematic review of adversarial machine learning and deep learning applications," *Al-Iraqia Journal for Scientific Engineering Research*, vol. 3, No. 4, pp. 14–40, 2024, https://doi.org/10.58564/IJSER.3.4.2024.275.
3. H. E. Achi, T. Belousova, L. Chen, A. Wahed, I. Wang *et al.*, "Automated diagnosis of lymphoma with digital pathology images using deep learning," *Ann Clin Lab Sci.*, vol. 49, no. 2, pp. 153–160, 2019 Mar, PMID: 31028058.
4. A. Esteva, B. Kuprel, R. A. Novoa *et al.*, "Dermatologist-level classification of skin cancer with deep neural networks," *Nature*, vol. DXXIV, no. 7639, pp. 115–118, 2017, https://doi.org/10.1038/nature21056.
5. V. Gaidano, V. Tenace, N. Santor *et al.*, "Clinically applicable approach to the classification of B-cell non-hodgkin

lymphomas with flow cytometry and machine learning," *Cancers*, vol. 12, pp. 1684, 2020.
6. R. Gargeya, and T. Leng, "Automated identification of diabetic retinopathy using deep learning," *Opthamology*, vol. CXXIV, no. 22, pp. 962–969, 2017, doi: https://doi.org/10.1016/j.ophtha.2017.02.008.
7. V. Gulshan, L. Peng, and Coram M. "Development and validation of deep learning algorithm for detection of diabetic retinopathy in retinal fundus photographs," *Journal of the American Medical Association*, vol. CCCXVI, no. 22, pp. 2402–2410, 2016, doi: https://doi.org/10.1001/jama.2016.17216.
8. D. V. Gunasekeran, D. S. W. Ting, G. S. W. Tan, T. Y. Wong, "Artifcial intelligence for diabetic retinopathy screening, prediction and management," *Curr Opin Ophthalmol*, vol. 31, no. 5, pp. 357–65, 2020.
9. M. Hamdi, E. M. Senan, M. E. Jadhav *et al.*, "Hybrid models based on fusion features of a CNN and handcrafted features for accurate histopathological image analysis for diagnosing malignant lymphomas," *Diagnostics*, vol. 13, pp. 2258, 2023, https://doi.org/10.3390/diagnostics13132258.
10. N. Hashimoto, K. Ko, T. Yokota, K. Kohno *et al.*, "Sub-type classification of malignant lymphoma using immunohistochemical staining pattern," *Int. J. Comput. Assist. Radiol. Surg*, vol. 17, pp. 1379–1389, 2022.
11. I. F. Jassam, A. A. Mukhlif, A. A. Nafea, M. A. Tharthar, and A. I. Khudhair, "A review of breast cancer histological image classification: Challenges and limitations, Iraqi journal for computer science and mathematics Iraqi," *Journal for Computer Science and Mathematics*, vol. 6, no. 1, Article 10, 2025.
12. H. Im, D. Pathania, P. J. McFarland, A. R. Sohani, I. Degani, M. Allen *et al.*, "Design and clinical validation of a point-of-care device for the diagnosis of lymphoma via contrast-enhanced micro-holography and machine learning," *Nat Biomed Eng*, vol. 2, no. 9, pp. 666–74, 2018.
13. D. Li, J. R. Bledsoe, Y. Zeng, W. Liu, Y. Hu, K. Bi *et al.*, "A deep learning diagnostic platform for difuse large B-cell lymphoma with high accuracy across multiple hospitals," *Nat Commun*, vol. 11, no. 1, pp. 6004, 2020.
14. C. S. Lisson, C. G. Lisson, M. F. Mezge, D. Wolf, S. A. Schmidt *et al.*, "Deep neural networks and machine learning radiomics modelling for prediction of relapse in mantle cell lymphoma," *Cancers*, vol. 14, no. 2008, 2022.
15. H. Miyoshi, K. Sato, Y. Kabeya, S. Yonezawa, H. Nakano, Y. Takeuchi *et al.*, "Deep learning shows the capability of high-level computer-aided diagnosis in malignant lymphoma," *Lab Investig*, vol. 100, no. 10, pp. 1300–10, 2020.
16. Multi Cancer Dataset Kaggle [Electronic resource]: https://www.kaggle.com/datasets/obulisainaren/multi-cancer (accessed on 8 May 2023).
17. P. M. Mwamba, W. O. Mwanda, N. Busakhala, R. M. Strother, P. J. Loehrer, S. C. Remick, "AIDS-related non-Hodgkin's lymphoma in subSaharan Africa: Current status and realities of therapeutic approach," *Lymphoma*, 2012.
18. S. Pileri, S. Ascani, L. Leoncini, E. Sabattini, P. Zinzani, P. Piccaluga *et al.*, "Hodgkin's lymphoma: the pathologist's viewpoint," *J Clin Pathol*, vol. 55, no. 3, pp. 162–76, 2002.
19. R. S. Shahadh, and B. Al-Khateeb, "Double dual convolutional neural network (D2CNN): A deep learning model based on feature extraction for skin cancer classification," *Iraqi Journal for Computer Science and Mathematics Iraqi Journal for Computer Science and Mathematics*, Vol. 6, no. 1, Article 10, 2025.
20. F. Sultana, A. Sufian, and G. Hinton, "Advancements in image classification using convolutional neural network," *Proc. - 2018 4th IEEE Int. Conf. Res. Compt. Intell. Commun.Networks, ICRCICN 2018*, pp. 122–129, 2018.
21. C. Syrykh, A. Abreu, N. Amara, A. Siegfried *et al.*, "Accurate diagnosis of lymphoma on whole-side histopathology images using deep learning," *npj Digital Medicine*, vol. III, no. 63, 2020, doi: https://doi.org/10.1038/s41746-020-0272-0.
22. D. S. W. Ting, C. Y. Cheung, G. Lim, and G. Tan, "Development and validation of deep learning system for diabetic retinopathy ad related eye diseases using retinal images from multiethnic population," *Jornal of the American Medical Association*, vol. CCCVIII, no. 7, pp. 962–969, 2017, doi: https://doi.org/10.1001/jama.2017.18152.
23. L. Yann, Y. Bengio, and G. Hinton, "Deep learning," *Nature*, no. 521, pp. 436-444, 2015, doi: https://doi.org/10.1038/nature14539.
24. X. Zhang, K. Zhang, M. Jiang, and L. Yang, "Research on the classification of lymphoma pathologi-cal images based on deep residual neural network," *Technol. Health Care*, vol. 29, pp. 335–344, 2021.
25. Q. Zheng, L. Yang, B. Zeng, J. Li, K. Guo, Y. Liang *et al.*, "Artificial intelligence performance in detecting tumor metastasis from medical radiology imaging: a systematic review and meta-analysis," *EClinicalMedicine*, vol. 31, 2021.