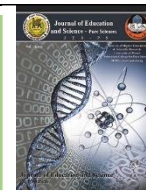
Journal of Education and Science (ISSN 1812-125X)

www.edusj.mosuljournals.com



Diagnosing Soft Tissue Tumors using Machine Learning Techniques: A Survey

N. Y. Ghanim*[1], J. S. M. Alneamy[2]

[1,2]Department of Software Engineering, College of Computer Sciences & Mathematics, University of Mosul, Mosul, Iraq

Article information

Article history:
Received January 10, 2023
Accepted April 16, 2023
Available online September 01, 2023

Keywords:
Deep Learning
Machine Learning
Colon Cancer
Lung Cancer
Breast Cancer

Correspondence:
Noor Y Ghanim
noor.21csp11@student.uomosul.edu.iq

Abstract

Cancer is the second-largest cause of death worldwide, constitute about one out of every six deaths. Diagnostic techniques have been developed for the early detection to determine benign and malignant ones and conduct appropriate treatment of the condition, which has led to a reduction in the incidence of death. Tumors can be detected and diagnosed by radiograph, magnetic resonance imaging, and ultrasound, and to confirm the type of tumor definitively, a biopsy is taken from the tumor, processed, and fixed on glass slides under a microscope and accurately identified. The explosive growth of artificial intelligence (AI) over the past ten years is the approved basis for making accurate decisions for diagnosing the type of tumor by building smart software based on machine learning (ML) and deep learning (DL), Which easier for specialists the access early detection of the type of tumor quickly. In this study, A study of previous works has been done for pathological conditions - breast, colon, and lung - which are the most common types of cancers, the accuracy of the diagnosis was studied for the type of tumor, benign or malignant , by using histological images by collecting biopsies from patients' tissues (histopathology) and characterizing them using the most recent convolutional neural networks (CNN), and researchers had to apply transfer learning techniques because Lack of explanations of the data histopathological dataset high-quality WSI (whole slice image), by training the network using a large computer vision data set (IMAGENET),in order to obtain a high diagnosis accuracy.

DOI: 10.33899/edusj.2023.137706.1316,

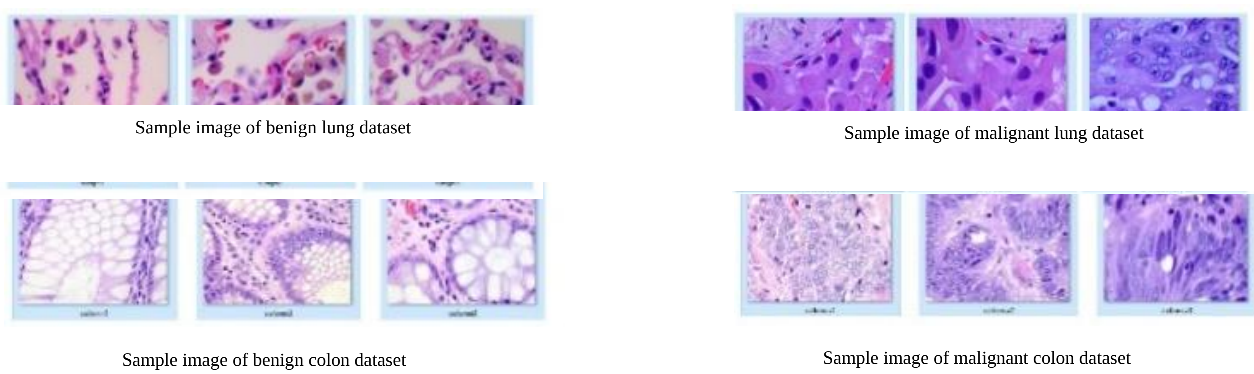
Introduction

Cancer has become a major public health problem in the modern world, causing 10 million deaths in 2020, as the number of cases reached (1.93 million cases of colon cancer, 2.21 million cases of lung cancer, and 2.26 million cases of breast cancer). The Cancer Research Agency of the World Health Organization (WHO) expects International) that by 2030 there will be 27 million cases of the disease [1][2]. One of the main causes is smoking, exposure to radon gas, arsenic and asbestos and to eating foods rich in preservatives Therefore, early detection of tumors reduces the risk of death [3]. Artificial intelligence (AI) has made significant contributions to the diagnosis of a variety of health problems, including cancer and new AI applications and methods will provide new insights in oncology [4][5]. The use of supervised machine learning and deep learning techniques in the medical field [6] speeds up the process of diagnosing cancer through patients' anatomy images taken from the patient's tumor biopsy [7][8]. These smart technologies interacted positively with bioinformatics in the diagnosis of various diseases [9][10]. Medical diagnosis using intelligent technologies is preceded by image pre-processing [11][12], feature extraction, analysis and classification into two types, benign and malignant [13][14][15]. Digital pathology has a significant impact on cancer diagnosis and prognosis [16][17]. To improve the performance of intelligent technologies and to obtain high accuracy in diagnosis, machine learning and deep learning experts can hybridize them with other intelligent algorithms, such as SVM (Support Vector Machine) KNN, LR (Logistic Regression) [14][18][19][20]. Intelligent systems used in diagnosis need to train the medical staff who will use them in their work to be qualified users and managers of modern technologies [21] [22]. Several intelligent systems for creating CAD models have been built by researchers to diagnose tumors for many applications including breast, lung, skin, prostate, colon, brain, cervix, bladder, and liver [23]. Note that cancer is divided according to the type of tissue and cell from which it arises into four main types: [24]

1- Carcinoma: Cancer that arises from the cells of the epithelial tissue that covers the surface of the body or its cavities (skin and intestines).
2- Sarcoma: Cancer that arises from connective tissue cells such as bones, muscles, and blood vessels.
3- Leukemia: cancer that arises from the primary cells (peregrinator) of different blood cells or immune systems in the case of leukemia
4- Lymphoma: Cancer that arises from the primary cells (peregrinator) of various blood cells or immune systems in the case of the immune system.
The topic of the research we are working on is within the first type

1. Datasets:

The previous work used data sets for different diseases, namely colon, lung, and breast. The data sets were used for no specific reason other than to be complete and show how well they can classify different medical images and were chosen from Kaggle because they contain enough images to train the model effectively and because the model's accuracy needs to be prepared. big data to validate the model, it is presented to the network to automatically divide it into 80% training data and 20% test data. All of them will be detailed in this section.
The first: Lung cancer begins in the lungs and spreads throughout the body and accounts for approximately 25% of all cancer deaths in the world and is the leading cause of death in both men and women in the United States and many other countries. Lung cancer is one of the most common types of cancer and results from a disruption of the body's basic unit of life and early detected and treatment are essential for a patient's recovery. Specialists use histopathology images of biopsy tissue from potentially infected areas of the lungs for diagnosis. Every year, about 150,000 people die of lung cancer. Histological images of lung tumors are shown in [13][25][23][26].
The second: Colon cancer is the third most public type of cancer worldwide in men and women, and it could be the reason environmental factors and genetic factors such as (obesity, red meat, tobacco, and alcohol) can be the cause. This type of cancer often occurs in people over the age of 50 and because early diagnosis of colon cancer is an important factor in shortening the treatment time. People should get a test for colon cancer at regular intervals. Its types are adenocarcinoma, fusiform, squamous, and undifferentiated. Histological images of colon tumors are shown in researches [13], [14], [27], and [28].



Sample image of benign lung dataset

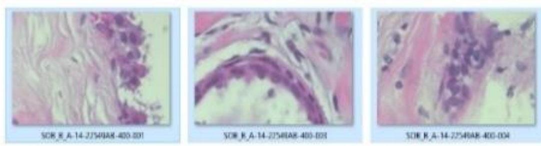
Sample image of malignant lung dataset

Sample image of benign colon dataset

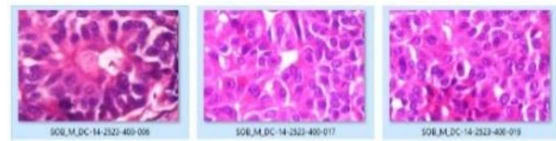
Sample image of malignant colon dataset

Figure 1: Sample images for each category in dataset [25]

The third: Breast cancer is the most prevalent cancer in women [2][29] and has high rates of death. Breast cancer is brought on by several factors, including estrogen and increased radiation exposure. Breast cancer occurs as a result of abnormal cell division that results from a mass called a tumor. Tumors can be classified as either cancerous (malignant) or non-cancerous (benign), while the case fatality rate was 14.7%, reports the WHO's IARC [1], and results from possible positive effects of the medical setting is the automatic classification of pathological breast cancer, in early detection methods in the analysis of pathological images by taking a histopathological biopsy. Histological images of the breast tumors are shown in [18], [21], [30], and [31].



Sample image of benign breast dataset



Sample image of malignant breast dataset

Figure 2: Sample images for each category in dataset [32]

Related work

In the past, ML models have shown particular potential for the detection and prognosis of cancer [33],[34] and Feature extraction is critical to the machine learning (ML) architecture [22][35]. Despite, these difficulties the BreaKHis[36] dataset were presented , which investigates the histopathology of breast cancer, were introduced, this dataset is not as large as ImageNet and includes fewer training examples; neural networks can be used to classify and evaluate it [37][38]. In identifying breast cancer using histology [16], pre-trained CNNs on ImageNet were utilized to extract features from histopathological images [15][19]. Transfer learning is used in this work [39][40], and the study showed that pre-trained CNNs are an excellent alternative to CNNs that are built from scratch [41].

The researcher (Talukder et al. 2022) suggested that the hybrid group extraction method was developed by researchers to effectively detect lung and colon cancer and reduce the probability of death. It combines deep feature extraction, group learning, and high-performance filtering. The model was evaluated on histological data sets of lung and colon (LC25000) with different magnification factors. The study results indicate that their hybrid model using deep and machine learning has accuracy rates for lung, colon, and colon cancer detection of 99.05%, 100%, and 99.30%, respectively. One of the weaknesses of this research is that preprocessing on authentic images did not increase accuracy [13].

The researcher (Abd Al-Ameer, Hussien, and Al Ameri 2022) proposed models for the automatic detection of lung cancer cells, including using Inception V3, Random Forest, and Convolutional Neural Network (CNN). The deep convolutional neural network was trained on a data set to extract the main factors that help in building the detection and diagnosis of lung cancer more efficiently. The results obtained were promising and satisfactory in terms of accuracy, accuracy, recall, F-score, and measurement of specificity in diagnosing lung cancer. It contains 178,000 images. The acquired accuracy was 97.31%, recall 97.09%, F-score measure, 96.88% specificity measure, and 97.09% accuracy. One of the weaknesses of this study is that the verification procedures are incorrect, which is more important than the training procedures [25].

The researcher (DSOUZA and ANSARI 2022) described a hybrid deep learning architecture for classifying histological images to identify the presence of cancer in a patient [42]. The proposed methodology is parallelized using the TensorFlow-GPU (Convolutional Networks) framework to accelerate the training of a deep CNN. To avoid over-equipping during the training phase, this study uses criteria for early stopping and transfer of learning during the training phase. Finally, this study concluded that the proposed ResNet-152 hybrid architecture performs the classification of pathological images very well with an accuracy of 92% and a data size of (7909) histological images with different magnification factors (×40, ×100, ×200, ×400) [43]

The researcher (Hatuwal and Thapa 2020) suggested that the detection of lung cancer using histological images of biopsy tissues from Convolutional Neural Networks (CNN) can identify and classify lung cancer types with an accuracy of 97.20% on dataset 15,000 histopathological images for diagnosis and in order to determine the correct treatment procedures for patients in a shorter time and their survival rate. One of its weaknesses is that the neural network tends to overfit in the case of a limited number of training data samples trained over a larger number of diagnosed periods, also related diagnosis of the types of lung cancer is error-prone and takes a long time [26].

The researcher (Mehmood et al. 2022) suggested the use of a large set of lung and colon disease images for training and validation. The dataset includes 25,000 histological images of colon and lung tissue which were evenly categorized into 5 groups [25] at different image magnification factors. Using a pre-trained neural network (AlexNet) by modifying its four layers, which resulting in an overall accuracy of 89%. The accuracy increased after using the proposed CSIP approach, reaching 98.8%, the proposed methodology not only outperformed the latest methods for lung and colon cancer detection in terms of accuracy but also reduced the computational time and cost. The image quality of the weak rather than the complete dataset is improved by transfer learning, and the proposed technique thus improved the accuracy and reliability of the model [27].

The researcher (Saxena, Shukla, and Gyanchandani 2020) proposed using pre-trained CNNs on datasets to extract features from the histopathology of breast cancer. Ten different diagnostic algorithms for histopathology with several classifiers were developed [39]. They arrived that each of ResNet50 and ResNet101 with the SVM classifier outperformed most of the other networks in the accuracy of diagnosis and had a data size of (7909) histological images of different magnifications, and the results showed that the performance varies with the magnification and, it was the accuracy of ResNet50_SVM = 90.12% and ResNet101_SVM = 90.05% with magnification (x). 200). One of the weaknesses of the study is that the image size of a data set BreakHis[28] is greater than the size of the input image of the pre-trained CNNs, and changing the image size will lead to the loss of the information effect [19].

The researcher (Vasal et al. 2018) proposed an approach based on transfer learning simple and effective in breast tissue image classification, submitted as part of the BACH 2018 Grand Challenge [41], and uses color corrections to tune convolutional neural networks (CNNs) from Google's Inception-V3 and ResNet50, both, pre-trained on the ImageNet database [29]. Network was outperformed (ResNet50) and Google's Inception-V3 on CNN trained network from the beginning [33], in terms of classification accuracy. The Inception-V3 network achieved an average test accuracy of 97.08%, in the data set of 400 histology pictures, the Inception-V3 network outperformed the ResNet50 network, with an average test accuracy of 97.08% as opposed to 96.66%. One of the weaknesses of this study is the wide color variance in histological image analysis, which results in a difference in the color responses of slides and raw materials [30].

The researchers (Yildirim and Cinar 2022) proposed a new method for detecting colon cancer image data by model MA_ColonNE consisting of 45 layers based on CNN for early detection and classification of colon cancer and used the data of 10,000 histological images from the Kaggle website with 8000 images to train MA_ColonNE and, after the training 2000 images to test the model based on CNN and the accuracy rate was 99.75%. According to the study, the suggested approach can identify colon cancer early, which will result in a more effective treatment process [28].

The researcher (Sarwinda et al. 2020) suggested extracting features by deep convolutional neural networks from (ResNet-50, and DenseNet-121) in the colon tissue data set from the Kaggle site, and then Colon cancer classification using all classifiers, and DenseNet-121 algorithm has gotten the highest level of accuracy and sensitivity. The specificity of ResNet-50 where the accuracy of DenseNet-121 classification is about 98.53% with the KNN classifier due to the depth of the network, but the features of ResNet-50 are twice the number of features of DenseNet-121.One of the weaknesses in the research is that the depth of the network can cause higher errors in training [14].

The researcher (Mahbod et al. 2018) suggested a final prediction by combining the results of two remaining neural networks (ResNet-101, ResNet-50) with different depths, using input images stained with hematoxylin and eosin. These networks were initially pre-trained to ImageNet images, then fine-tuned to breast histological images of a dataset of 200,000 histological images The approach method significantly outperformed when applied to the 2015 BioImaging dataset, achieving an accuracy of 97.22%. To improve the performance, they proposed an appropriate normalization technique, using more comprehensive training data, and integrating more accurate deep models to improve classification accuracy. One of the point weaknesses of this study is the resolution that obtained from extended images was lower quality due to the increased complexity of the images [31].

The idea of using DL structures to classify areas of colon cancer in histopathology data was proposed by the researcher (Hamida et al.) in 2021 as it appears as an attractive solution for tissue classification and segmentation in histopathological images, review and comparison were made for the latest Convolutional Neural Networks (CNN). They used transfer learning techniques to overcome the lack of datasets [40]. This method stands out because it uses a large computer vision dataset (IMAGENET) to train the network and generate many completely new features [15]. In the AiCOLO colon cancer dataset, they examined and compared the latest Convolutional Neural Networks (CNNs), and discovered reliable results for corrective tuning with an accuracy of up to 96.98% with ResNet for a dataset of up to 5000 images with factors different magnification (×40, ×100, ×200, × 400) and models have used CRC-5000 and NCT-CRC-HE-100K and combined datasets, were investigated and tested and have achieved RESNET respectively: 96.77%, 99.76%, and 99.98% for the three publicly available datasets. And have used a training network from scratch also, the Pixel Split Search feature is a more accurate tool for analyzing colon histological images and can deal with unbalanced and scattered data [33], and, weak points in this study, sparsely annotated histopathological dataset [44].

Table I. Summary of Cancer Detection Methods for Histopathological Dataset Size, Datatype, Techniques, And Accuracy

Reference no.	Method	Dataset size	Data type	Year	Accuracy
[13]	MobileNet	25000	lung	2022	99.05%
			colon		100%
			Lung + colon		99.3%
[25]	CNN- Inception V3 RF(Random Forest)	178000	Lung	2022	97.09%
[43]	ResNet152 LSTM	7909	breast	2022	92%
[26]	CNN	15000	Lung	2020	96.11% and 97.20%
[27]	CNN - AlexNet	25000	Lung+ colon	2022	98.8%
[19]	ResNet-50- SVM ResNet-101-SVM	7909	Breast	2020	ResNet50_SVM×200=90.12% ResNet101_SVM×200=90.05%
[28]	CNN MA_ColonNET model	10000	Colon	2021	99.75%
[30]	ResNet50	400	breast	2018	97.08%
[14]	DenseNet-121 KNN	10000	Colon	2021	98.53%
[31]	ResNet-50 ResNet-101	200000	Breast	2019	97.22%
[44]	ResNet101	5000	Colon	2021	96.98%

Conclusion

The data set used in this study showed (training and testing), including benign and cancerous cells. Because intelligent technologies have developed, deep learning and machine learning have recently greatly influenced medical image processing and high-accuracy diagnostics. The most important conclusions of this research were as follows:

1. Most of the previous studies used pre-trained convolutional neural networks (CNNs) to extract features from pathological anatomical images or CNNs trained from scratch and pre-trained transfer learning technology on a large data set, which reduces training effort and is faster and easier than building and training a network of zero.
2. A group of researches related to breast, lung, and colon tumors were reviewed and were within different Magnification factors (x40, x100, x200, and x400). Most of the research indicated that those within the Magnification factor x100 are the best, as the details of the tumor are clear with high accuracy.
3. The research used ready-made data from the Kaggle medical site, which makes it easier to compare the results to distinguish the best algorithm.
4. Most of the research had preliminary processing, which is necessary before using the smart algorithms of (ML, DL) technologies, and it included changing the image size to the required size, implementing the normalization process, increasing the data, applying the transfer learning method, converting all pixel values of the images to the range (0 ,1) to make convergence faster and use compression techniques to reduce computational costs.

5. Deep learning and machine learning techniques were used in research, and most of them were hybrid with other intelligence techniques to raise the accuracy of discrimination and obtain perfect accuracy. The algorithms used to obtain the discrimination percentage were as follows:

a- colon:

The CNN algorithm with the MA ColonNET model with a data size of 10000, preliminary processing, changing the size of the data, using Normalization, and converting the image from a multidimensional matrix format to a one-dimensional matrix format, and an accuracy of 100% was obtained.

B- Lung:

MobileNet algorithm with a data size of 25000 and preliminary processing, which is reducing the size of the input image, converting the image from a multidimensional matrix format to a one-dimensional matrix format, and then to NumPy, scaling within (0,1) values, and using Normalization, and an accuracy of 99.05% was obtained.

C- Breast:

ResNet50 algorithm hybridized with ResNet101 with a data size of 200000 and raw processing with Normalization, input image scaling and colorimetric contrast in tissue images, and an accuracy of 97.22% was obtained. Based on the previous study, we propose in the future to use a hybrid between two types of deep and machine learning with multiple initial treatments to reach 100% accuracy in distinguishing tumors for the three pathological conditions under study: breast tumors, lung tumors, and colon tumors.

Acknowledgment

This research was supported by university of Mosul hosted by the Department of Software in the college Computer Science and Mathematics, Mosul, Iraq.

Conflict of interest

The author has no conflict of interest.

Resources:

1. J. Ferlay *et al.*, “Cancer statistics for the year 2020: an overview,” *International Journal of Cancer*, vol. 149, no. 4, Apr. 2021, doi: https://doi.org/10.1002/ijc.33588.
2. F. Bray, J. Ferlay, and I. Soerjomataram, “Global Cancer Statistics 2018 : GLOBOCAN Estimates of Incidence and Mortality Worldwide for 36 Cancers in 185 Countries,” vol. 00, no. 00, pp. 1–31, 2018, doi: 10.3322/caac.21492.
3. A. M. Lewandowska, M. Rudzki, S. Rudzki, T. Lewandowski, and B. Laskowska, "Environmental risk factors for cancer - review paper," *Annals of Agricultural and Environmental Medicine*, vol. 26, no. 1. Institute of Agricultural Medicine, pp. 1–7, 2019. doi: 10.26444/aaem/94299.
4. F. Jiang *et al.*, "Artificial intelligence in healthcare: Past, present, and future," *Stroke and Vascular Neurology*, vol. 2, no. 4. BMJ Publishing Group, pp. 230–243, Dec. 01, 2017. doi: 10.1136/svn-2017-000101.
5. M. J. Iqbal *et al.*, "Clinical applications of artificial intelligence and machine learning in cancer diagnosis: looking into the future," *Cancer Cell International*, vol. 21, no. 1. BioMed Central Ltd, Dec. 01, 2021. doi: 10.1186/s12935-021-01981-1.
6. J. van der Laak, G. Litjens, and F. Ciompi, "Deep learning in histopathology: the path to the clinic," *Nature Medicine*, vol. 27, no. 5. Nature Research, pp. 775–784, May 01, 2021. doi: 10.1038/s41591-021-01343-4.
7. Rp. Kumar and Pb. KVTriveni, "HISTOPATHOLOGIC CANCER DETECTION," *IRJCS Int. Res. J. Comput. Sci.*, vol. VI, pp. 102–124, 2019, doi: 10.26562/IRJCS.2019.APCS10087.
8. M. Bakator and D. Radosav, "Deep learning and medical diagnosis: A review of literature," *Multimodal Technologies and Interaction*, vol. 2, no. 3. MDPI AG, Sep. 01, 2018. doi: 10.3390/mti2030047.
9. D. Komura and S. Ishikawa, "Machine Learning Methods for Histopathological Image Analysis," *Computational and Structural Biotechnology Journal*, vol. 16. Elsevier B.V., pp. 34–42, 2018. doi: 10.1016/j.csbj.2018.01.001.
10. Z. Hu, J. Tang, Z. Wang, K. Zhang, L. Zhang, and Q. Sun, "Deep learning for image-based cancer detection and diagnosis − A survey," *Pattern Recognit.*, vol. 83, pp. 134–149, Nov. 2018, doi: 10.1016/j.patcog.2018.05.014.
11. M. Thompson, R. C. Gonzalez, and P. Wintz, “Digital Image Processing,” *Leonardo*, vol. 14, no. 3, p. 256, 1981, doi: https://doi.org/10.2307/1574313.
12. N. Handa, Y. Kaushik, N. Sharma, M. Dixit, and M. Garg, “Image Classification Using Convolutional Neural Networks,” *Commun. Comput. Inf. Sci.*, vol. 1393, no. December, pp. 510–517, 2021, doi: 10.1007/978-981-16-3660-8_48.
13. M. A. Talukder, M. M. Islam, M. A. Uddin, A. Akhter, K. F. Hasan, and M. A. Moni, "Machine Learning-based Lung and Colon Cancer Detection using Deep Feature Extraction and Ensemble Learning," Jun. 2022, [Online]. Available: http://arxiv.org/abs/2206.01088
14. D. Sarwinda, A. Bustamam, R. H. Paradisa, T. Argyadiva, and W. Mangunwardoyo, "Analysis of Deep Feature Extraction for Colorectal Cancer Detection," in *ICICoS 2020 - Proceeding: 4th International Conference on Informatics and Computational Sciences*, Nov. 2020. doi: 10.1109/ICICoS51170.2020.9298990.
15. C. A. Ferreira *et al.*, "Classification of Breast Cancer Histology Images Through Transfer Learning Using a Pre-trained Inception Resnet V2," in *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 2018, vol.

10882 LNCS, pp. 763–770. doi: 10.1007/978-3-319-93000-8_86.
16. C. Wemmert, J. Weber, F. Feuerhake, and G. Forestier, "Deep Learning for Histopathological Image Analysis," *Deep Learn. Biomed. Data Anal.*, pp. 153–169, 2021, doi: 10.1007/978-3-030-71676-9_7.
17. S. Robertson, H. Azizpour, K. Smith, and J. Hartman, "Digital image analysis in breast pathology—from image processing techniques to artificial intelligence," *Translational Research*, vol. 194. Mosby Inc., pp. 19–35, Apr. 01, 2018. doi: 10.1016/j.trsl.2017.10.010.
18. M. A. Kahya, W. Al-Hayani, and Z. Y. Algamal, "Classification of Breast Cancer Histopathology Images based on Adaptive Sparse Support Vector Machine," online) Scienpress Ltd, 2017.
19. S. Saxena, S. Shukla, and M. Gyanchandani, "Pre-trained convolutional neural networks as feature extractors for diagnosis of breast cancer using histopathology," *Int. J. Imaging Syst. Technol.*, vol. 30, no. 3, pp. 577–591, Sep. 2020, doi: 10.1002/ima.22399.
20. Shallu and R. Mehra, "Breast cancer histology images classification: Training from scratch or transfer learning?," *ICT Express*, vol. 4, no. 4, pp. 247–254, Dec. 2018, doi: https://doi.org/10.1016/j.icte.2018.10.007.
21. A. Jalalian *et al.*, "Review article : FOUNDATION AND METHODOLOGIES IN COMPUTER-AIDED," pp. 113–137, 2017.
22. G. Kumar and P. K. Bhatia, "A detailed review of feature extraction in image processing systems," in *International Conference on Advanced Computing and Communication Technologies, ACCT*, 2014, pp. 5–12. doi: 10.1109/ACCT.2014.74.
23. S. Saxena and M. Gyanchandani, "Machine Learning Methods for Computer-Aided Breast Cancer Diagnosis Using Histopathology: A Narrative Review," *Journal of Medical Imaging and Radiation Sciences*, vol. 51, no. 1. Elsevier Inc., pp. 182–193, Mar. 01, 2020. doi: 10.1016/j.jmir.2019.11.001.
24. V. Kumar, A. Abbas, J. Aster, and J. Turner, *Robbins & Cotran Pathologic Basis of Disease Tenth Edition*. 2021.
25. A. A. Abd Al-Ameer, G. A. Hussien, and H. A. Al Ameri, "Lung cancer detection using image processing and deep learning," *Indones. J. Electr. Eng. Comput. Sci.*, vol. 28, no. 2, pp. 987–993, Nov. 2022, doi: 10.11591/ijeecs.v28.i2.pp987-993.
26. B. K. Hatuwal and H. C. Thapa, "Lung Cancer Detection Using Convolutional Neural Network on Histopathological Images," *Int. J. Comput. Trends Technol.*, vol. 68, no. 10, pp. 21–24, Oct. 2020, doi: 10.14445/22312803/ijctt-v68i10p104.
27. S. Mehmood *et al.*, "Malignancy Detection in Lung and Colon Histopathology Images Using Transfer Learning with Class Selective Image Processing," *IEEE Access*, vol. 10, pp. 25657–25668, 2022, doi: 10.1109/ACCESS.2022.3150924.
28. M. Yildirim and A. Cinar, "Classification with respect to colon adenocarcinoma and colon benign tissue of colon histopathological images with a new CNN model: MA_ColonNET," *Int. J. Imaging Syst. Technol.*, vol. 32, no. 1, pp. 155–162, Jan. 2022, doi: 10.1002/ima.22623.
29. W. Stewart, World Health Organization, A. For, M. De, and L. Cancer, *World cancer report*. Lyon: Iarc Press, 2014.
30. S. Vesal, N. Ravikumar, A. A. Davari, S. Ellmann, and A. Maier, "Classification of Breast Cancer Histology Images Using Transfer Learning," in *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 2018, vol. 10882 LNCS, pp. 812–819. doi: 10.1007/978-3-319-93000-8_92.
31. A. Mahbod, I. Ellinger, R. Ecker, Ö. Smedby, and C. Wang, "Breast Cancer Histological Image Classification Using Fine-Tuned Deep Network Fusion," in *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 2018, vol. 10882 LNCS, pp. 754–762. doi: 10.1007/978-3-319-93000-8_85.
32. F. A. Spanhol, L. S. Oliveira, C. Petitjean, and L. Heutte, "A Dataset for Breast Cancer Histopathological Image Classification," *IEEE Trans. Biomed. Eng.*, vol. 63, no. 7, pp. 1455–1462, Jul. 2016, doi: 10.1109/TBME.2015.2496264.
33. M. Masud, N. Sikder, A. Al Nahid, A. K. Bairagi, and M. A. Alzain, "A machine learning approach to diagnosing lung and colon cancer using a deep learning-based classification framework," *Sensors (Switzerland)*, vol. 21, no. 3, pp. 1–21, Feb. 2021, doi: 10.3390/s21030748.
34. J. Wiens, E. S. Shenoy, and H. Street, "Title: Machine Learning for Healthcare: On the Verge of a Major Shift in Healthcare Epidemiology," 2017.
35. A. Madabhushi and G. Lee, "Image analysis and machine learning in digital pathology: Challenges and opportunities," *Medical Image Analysis*, vol. 33. Elsevier B.V., pp. 170–175, Oct. 01, 2016. doi: 10.1016/j.media.2016.06.037.
36. Y. Benhammou, B. Achchab, F. Herrera, and S. Tabik, "BreakHis based Breast Cancer Automatic Diagnosis using Deep Learning: Taxonomy, Survey and Insights," *Neurocomputing*, Sep. 2019, doi: https://doi.org/10.1016/j.neucom.2019.09.044.
37. A. Krizhevsky, I. Sutskever, and G. E. Hinton, "ImageNet classification with deep convolutional neural networks," *Commun. ACM*, vol. 60, no. 6, pp. 84–90, Jun. 2017, doi: 10.1145/3065386.
38. D. A. E. Wagh, "Classification of Image using Convolutional Neural Networks," *Int. J. Res. Appl. Sci. Eng. Technol.*, vol. 9, no. 8, pp. 1212–1214, Aug. 2021, doi: 10.22214/ijraset.2021.37581.
39. J. H. Lee, S. Howland, J. Buckheit, L. A. Phillips, and C. D. Corley, "A DAPTIVE T RANSFER L EARNING :," no. Dl, pp. 1–10.
40. H. Chougrad, H. Zouaki, and O. Alheyane, "Multi-label transfer learning for the early diagnosis of breast cancer," *Neurocomputing*, vol. 392, pp. 168–180, Jun. 2020, doi: 10.1016/j.neucom.2019.01.112.
41. G. Aresta *et al.*, "BACH: Grand challenge on breast cancer histology images," *Medical Image Analysis*, vol. 56, pp. 122–139, Aug. 2019, doi: https://doi.org/10.1016/j.media.2019.05.010.
42. K. Tawheed, K. Salman, A. Farooqui, and S. Chavan, "Cancer Detection Using Deep Learning," *Int. Res. J. Eng. Technol.*, 2021, [Online]. Available: www.irjet.net
43. K. J. DSOUZA and Z. A. ANSARI, "HISTOPATHOLOGY IMAGE CLASSIFICATION USING HYBRID PARALLEL STRUCTURED DEEP-CNN MODELS," *Appl. Comput. Sci.*, vol. 18, no. 1, pp. 20–36, Mar. 2022, doi: 10.35784/acs-2022-2.
44. A. Ben Hamida *et al.*, "Deep learning for colon cancer histopathological images analysis," *Computers in Biology and Medicine*, vol. 136, p. 104730, Sep. 2021, doi: https://doi.org/10.1016/j.compbiomed.2021.104730.

تشخيص أورام الأنسجة الرخوة باستخدام تقنيات التعلم الآلي: دراسة

نور يحيى غانم و جمال صلاح الدين مجيد النعيمي

قسم البرمجيات، كلية علوم الحاسوب والرياضيات، جامعة الموصل، الموصل، العراق

الخلاصة

السرطان هو ثاني أكبر سبب للوفاة في جميع أنحاء العالم ، ويشكل حوالي واحد من كل ست وفيات. تم تطوير تقنيات التشخيص للكشف المبكر عن الأمراض الحميدة والخبيثة وإجراء العلاج المناسب للحالة ، مما أدى إلى انخفاض معدل حدوث الوفاة. يمكن الكشف عن الأورام وتشخيصها عن طريق التصوير الشعاعي والتصوير بالرنين المغناطيسي والموجات فوق الصوتية ، ولتأكيد نوع الورم بشكل نهائي ، يتم أخذ خزعة من الورم ومعالجتها وتثبيتها على شرائح زجاجية تحت المجهر وتحديدها بدقة. النمو الهائل للذكاء الاصطناعي (AI) على مدى السنوات العشر الماضية هو الأساس المعتمد لاتخاذ قرارات دقيقة لتشخيص نوع الورم من خلال بناء برمجيات ذكية تعتمد على التعلم الآلي (ML) والتعلم العميق (DL) ، مما يسهل على المتخصصين. الوصول المبكر إلى الكشف المبكر عن نوع الورم بسرعة. في هذه الدراسة تم عمل دراسة لأعمال سابقة للحالات المرضية - الثدي والقولون والرئة - وهي أكثر أنواع السرطانات شيوعاً ، وقد تمت دراسة دقة التشخيص لنوع الورم الحميد أو الخبيث عن طريق باستخدام الصور النسيجية عن طريق جمع الخزعات من أنسجة المرضى (التشريح المرضي) وتوصيفها باستخدام أحدث الشبكات العصبية التلافيفية (CNN) ، وكان على الباحثين تطبيق تقنيات تعلم النقل بسبب عدم وجود تفسيرات لبيانات مجموعة البيانات النسيجية المرضية عالية الجودة WSI (بأكملها) صورة شريحة) ، من خلال تدريب الشبكة باستخدام مجموعة بيانات رؤية حاسوبية كبيرة (IMAGENET) ، من أجل الحصول على دقة تشخيص عالية