

Application of Artificial Intelligence Machine Learning Models in Cancer Diagnosis

Qizhen Guo

Westlake High School, Westlake Village, California, USA

qizhenguo000@gmail.com

Abstract. In the medical domain, artificial intelligence (AI) pertains to the utilization of machine learning models and algorithms for the analysis of intricate medical data. This paper demonstrates the application of learning models in cancer diagnosis through two innovative approaches: the Hybrid U-Net model and an improved MobileNet-V3 model, as well as a novel Multiple Instance Learning (MIL) model. The combination of these models has achieved significant improvements in cancer diagnosis accuracy, interpretability, and performance. The integration of the Hybrid U-Net with MobileNet-V3 effectively enhances image segmentation and feature extraction efficiency, particularly in analysis of medical imaging tasks such as skin cancer detection. Through hyperparameter optimization, these models perform exceptionally well with complex data, significantly improving performance metrics such as accuracy, PPV, NPV, and AUC. On the other hand, the MIL model, with its innovative attention mechanism, has shown considerable promise in improving the diagnosis of lung cancer by providing greater interpretability and aiding clinical understanding. However, despite the impressive performance of these models, the research also highlights limitations such as small sample sizes and reliance on existing segmentation methods, which affect the stability and broad applicability of the models.

Keywords: Artificial intelligence, machine learning, cancer diagnosis.

1. Introduction

Cancer is a kind of disease caused by uncontrolled abnormal cells and the survival rate of cancer is low. It has become a global burden. According to WHO (World Health Organization), in 2022, there were an anticipated 20 million new cases of cancer and 9.7 million deaths from the disease globally. Once someone is diagnosed with cancer, they have a low possibility of living more than 5 years. Cancer is predicted to strike 1 in 5 people at some point in their lives. It is likely to be developed in both men and women [1]. However, studies have found that the earlier the cancer has been diagnosed, the higher survival rate the one has. Early diagnosis of cancer can save patients from deaths and suffering. Because early diagnosis can stop the spread of cancer from early stages which is much easier than treating patients from late stage[2]. Currently, scientists have already developed several methods of cancer diagnosis. These include Positron Emission Tomography (PET), Computed Tomography (CT), Magnetic Resonance Imaging (MRI), and Magnetic Resonance Spectroscopy (MRS). These imaging techniques allow for the early detection and monitoring of cancer, which is crucial for improving patient outcomes and survival rates [3,4]. Nowadays, Cancer diagnosis has been

a popular topic among the medical field. More accurate and efficient methods are under development, among these, artificial intelligence learning models based strategies draw great attention as its remarkable speed and moderate precision.

Machine learning models are becoming increasingly useful and are helping to transform various sectors of our current society. In machine learning, algorithms process large amounts of data to uncover patterns and correlations. This technology has become an essential tool for a multitude of companies, including e-commerce, banking, and healthcare, radically altering the way that decisions take place and data is assessed [5]. Examples from the healthcare industry include risk assessment instruments, patient monitoring programs, and diagnostic support systems, illustrating how machine learning may give doctors useful information to enhance patient outcomes [6]. Machine learning models have been specifically applied to the prediction of the response of chimeric antigen receptor (CAR) T treatment. Daniels et al. created a deep learning algorithm that makes use of signaling motifs to assess a specific CAR's antitumor activity. The CAR motif sequence is fed into their deep learning architecture, which then uses two convolutional neural network (CNN) layers, one long short-term memory (LSTM) network layer, and seven fully connected neural network (FCNN) layers to propagate the encoded sequence. By using motif combinations specifically created for CAR T cells, this method allows for the direct prediction of tumor stemness and cytotoxicity, which in turn directs the engineering of CAR signaling domains for CAR T therapies [7].

At present, machine learning algorithms have also been widely applied in cancer diagnosis and prognosis, and have demonstrated improved performance over traditional methods. Machine learning-based computer-assisted detection (CAD) has demonstrated encouraging outcomes in helping identify cancer, outperforming traditional diagnostic methods [8]. Since this paper focuses on the application of artificial intelligence machine learning model in cancer diagnosis, these advancements highlight the potential of these techniques to enhance cancer detection and management.

2. AI Models in Cancer Diagnosis

2.1. hybrid U-Net and improved MobileNet-V3

2.1.1. What kind of cancer does this models used to perform diagnosis.

With an expected 1.5 million new cases in 2022, skin cancer is the most common type of cancer diagnosed globally[9]. Convolutional neural networks (CNNs) have shown a great deal of promise in terms of increasing skin cancer detection accuracy. These deep learning-based algorithms, particularly CNNs, have shown good classification and segmentation capabilities for dermatoscopic skin lesion images[14].The ability of CNNs to process and analyze these images effectively has made them an invaluable tool for improving the diagnosis and detection of skin cancer[10]. Some deep learning frameworks and methodologies are being invented as a better tool for detecting and classifying skin cancer. For example, Convolutional Neural Networks (CNN), Recurrent Neural Networks (RNN), Decision Trees (DT) and Deep Neural Networks (DNN).

2.1.2. What kind of method are used to train this model.

In order to enhance efficacy and accuracy, deep learning models—like the Dual Optimization Utilizing Deep Learning Network (DUODL) for skin cancer identification—often require sophisticated strategies during training [11].

Transfer learning is a primary method used, where a previously trained model, such as MobileNet-V3, is adapted for the new task. This approach reduces training complexity, speeds up convergence, and minimizes the need for extensive data. MobileNet-V3, trained on large datasets, provides valuable feature extraction from dermoscopic images, which is then utilized by the DUODL network [12].

For accurate segmentation, the DUODL network also uses the U-Net design. Skin lesion segments more precisely because to U-Net's encoder-decoder structure with skip connections, which enhances the features recovered by MobileNet-V3.

The model's performance is enhanced by adjusting hyperparameters via Bayesian optimization. This systematic approach optimizes the model's settings, contributing to the DUODL network's notable precision of 98.76% in skin cancer diagnosis.

2.1.3. Data for the model training

To prepare data for training their skin disease classification model, the researchers performed various preprocessing steps. Among these were image resizing, pixel value normalization, and the use of data augmentation methods like rotation, flipping, and scaling. The researchers sought to enhance the model's capacity for generalization by diversifying the training data via various augmentation techniques [13].

The 'HAM-10000' collection, which is accessible to the public, comprises 10,000 pictures of various skin lesions, such as nevi, melanoma, and other related disorders. This dataset, referred to as 'Human against the machine', is commonly used in skincare research for classifying and diagnosing skin diseases [15].

The HAM-10000 dataset was divided into validation, training, and test sets by the researchers in order to evaluate the model's performance on unobserved data. The researchers were able to assess the model's capacity to generalize and provide precise diagnoses on fresh, real-world skin lesion photos, as opposed to merely the training set of data, by analyzing the model's predictions on the held-out test set [13].

The images in the dataset include a wide variety of skin lesion types, including benign nevus, actinic keratosis, melanoma, and seborrheic keratosis. The dataset also provides ground truth annotations for the lesion areas and their corresponding diagnoses, which were used to train the model's segmentation and classification capabilities [13].

2.1.4. the accuracy of this model in test data

The experimental results of the testing dataset for 100 epochs for both the current and proposed models are presented in Table 1 [16]. The U-Net and MobileNetV3 models, when combined, produced excellent classification results for skin cancer. The model demonstrated validation outcomes with an accuracy of 98.86%, with Recall, F1-score, and precision all exceeding 95%. The model's ability to distinguish between malignant and benign tumors is reflected by its ROC-AUC score of 98.45%, indicating outstanding performance. It consistently achieved strong results on the test set, maintaining an accuracy of 98.45% [13].

Table 1. Test findings obtained after 100 epochs of experimentation, modified from [13]

Model	Precision%	Recall%	F-1 Score%	Accuracy%	ROC-AUC%
MobileNet	91.25	89.18	88.34	92.45	96.25
Resnet-152v2	90.34	88.57	87.65	92.57	95.32
VGG-16	90.78	89.27	88.95	92.89	94.21
MobileNet-V2	92.36	90.78	89.36	94.47	91.78
VGG-19	93.65	91.47	90.65	94.89	93.54
Proposed model	97.84	95.27	97.32	98.86	98.45

According to empirical findings, the proposed optimized hybrid MobileNet-V3 model surpasses existing skin cancer detection and segmentation strategies in terms of precision (97.84%), sensitivity (96.35%), accuracy (98.86%), and specificity (97.32%). This research's improved performance may result in faster and more precise diagnoses, potentially saving lives and lowering healthcare costs. The constructed model has high clinical relevance, as evidenced by its precision and recall scores, making it a valuable tool for dermatologists in the early diagnosis of skin cancer. The increased accuracy provides a possible avenue for reducing misdiagnoses and improving patient outcomes.

2.2. Multiple instance learning

2.2.1. Application of Multiple Instance Learning model in lung cancer

Globally, lung cancer is the most prevalent type of cancer. It ranks second in women's cancer cases and the most frequently seen in men's. In 2022, there were 2,480,675 new instances of lung cancer [17].

Advances in medical imaging, particularly computed tomography (CT) scans, have enabled the identification of these small lung lesions [18]. Even so, it still can be difficult to determine whether lung nodules are benign or cancerous, and correct radiological interpretation is frequently needed.

Researchers have looked into using machine learning methods, including multiple instance learning (MIL), to classify lung nodules in CT scans in order to overcome this difficulty. MIL is a well-suited approach for this task, as it can effectively handle the inherent ambiguity and uncertainty in the nodule-level annotations.

The successful application of attention mechanisms in multiple instance learning (MIL) to cancer diagnosis is becoming increasingly feasible.

2.2.2. Method

When it comes to MIL, nodules are categorized into "bags of instances" (i.e., many nodules in a single chest CT scan). Therefore, the objective is to diagnose the subject as a whole. It is not necessary to label each individual nodule discovered in the subject; just the diagnoses at the subject level, or the bag labels, are needed. [19] Because the subject-level diagnosis is much more widely available than annotations on each nodule, this approach is therefore better suited for real-world data mining in lung cancer.

The MIL approach is used to address the inherent labeling ambiguity in medical imaging tasks, where each CT scan or cytological image represents a "bag" of instances (individual nodules or regions of interest). Instead of classifying individual instances, the model is trained to classify the entire bag.

The training procedures leverage the strengths of MIL, such as learning features directly from the data, incorporating instance-level information, and achieving high performance even with limited labeled datasets[20]. The reported results demonstrate the effectiveness of these MIL-based approaches for lung nodule classification, lung pathophysiology detection, and thoracic disease diagnosis from medical images.

The wide approaches of MIL include MI - SVM,[21] mi - Graph, [22] miVLAD, [23] and MI - Net.

2.2.3. Data

There is a dataset to help train this model and also feature extraction is used to modify this model

2.2.3.1. Information Set

The Lung Image Database Consortium (LIDC-IDRI) dataset, an open-access collection of clinical chest CT exams, served as the main source of data for this investigation. The dataset was retrieved in May 2020 under a Creative Commons Attribution Non-Commercial 3.0 Unported (CC BY-NC) license obtained from The Cancer Imaging Archive (TCIA) [24].

The LIDC-IDRI dataset includes 1018 CT scans from seven different institutions, as seen in figure 1. Independent radiologists analyzed the scans, yielding 7371 annotations in all, 2669 of which were consensus nodules. As per existing diagnostic criteria, the researchers excluded participants with nodules less than 3 mm in diameter and those with unreported or uncertain diagnoses. As a result, 310 eligible nodules came from 110 distinct subjects in the final dataset [25].

An XML file with the binary masks for the nodules was supplied during the data collection. Figure 1 [26] lists the number of patients and the nodules that were omitted, along with the rationale. As seen

in Table 2, the majority of participants (75%) and lung nodules (77%) in the dataset tested positive for lung cancer.

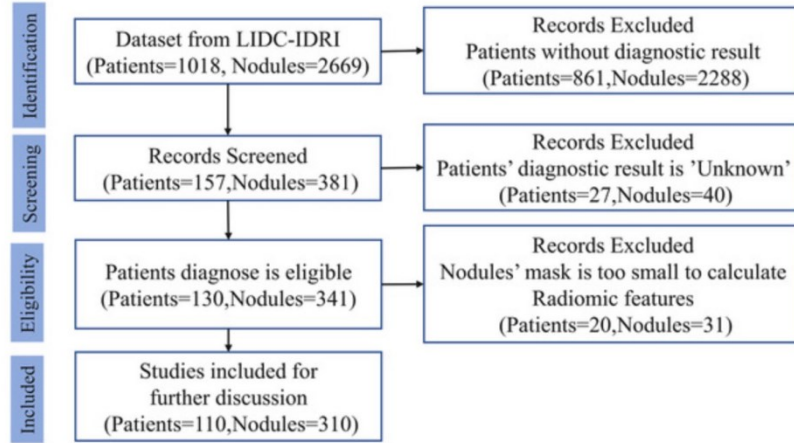


Figure 1. Flowchart of Sample Selection Illustrating the Count of Subjects and Nodules Chosen for Analysis [26].

Table 2. Distribution of actual lung cancer cases among patients and nodules group, modified from [26]

	Lung cancer	Not Lung cancer	Total
Numbers of (% of total) patients	82 (75%)	28 (25%)	110
Numbers of (% of total nodules)	239 (77%)	71 (23%)	310

2.2.3.2. Extraction of features

Radiomics features were extracted using the pyRadiomics library (v2.2.0), an open-source Python tool. Before extraction, 36 images were resampled to 2 mm isotropic voxels. In total, 103 features were extracted, including 13 morphology (shape) features, 17 first-order intensity-histogram features, and 73 texture (Haralick) features. Binary masks for the gross tumor volume (GTV) were generated from XML files in the LIDC-IDRI dataset using the open-access pylidc library. Additionally, SimpleITK (v1.2.4) was employed to convert DICOM CT images into 3D images, which streamlined the pyRadiomics feature extraction process. The online documentation provides the mathematical definitions for each feature.

2.2.4. The accuracy of this model in test data

Different MIL techniques are compared in Table 3[27]. The best PPV was reached by MI-Net, and the maximum recall was attained by the attention-based MIL model without oversampling. Overall, the attention-based MIL model performed the best in terms of PPV, AUC, and accuracy. It performed below the best outcomes for recall and NPV (Wilcoxon test, $p = 0.02$ and $p < 0.01$, respectively), but it significantly beat other techniques in PPV and AUC (Wilcoxon test, $p < 0.01$). Furthermore, in comparison to its non-oversampling counterpart, attention-based MIL with oversampling shown a significant improvement in all measures ($p = 0.02$ for recall and $p < 0.01$ for accuracy, PPV, NPV, and AUC) in all but recall.

Table 3. Results of the attention-based deep MIL approach with class imbalance correction, compared to other MIL methods (attention-based MIL w/o oversampling, MI-SVM, mi-graph, miVLAD, and MI-Net), modified from [27]

Methods	Attention-based MIL	Attention-based MIL w/o oversampling	MI-SVM	mi-graph	miVLAD	MI-Net	Traditional MIL
Recall	0.870±0.061	0.889±0.061	0.756±0.084	0.777±0.048	0.871±0.087	0.835±0.109	0.850±0.099
Accuracy	0.807±0.069	0.768±0.059	0.703±0.080	0.749±0.055	0.782±0.063	0.727±0.050	0.748±0.065
PPV	0.928±0.078	0.842±0.071	0.560±0.199	0.772±0.042	0.835±0.059	0.522±0.265	0.835±0.070
NPV	0.591±0.155	0.483±0.209	0.810±0.080	0.713±0.229	0.675±0.160	0.838±0.069	0.478±0.233
NUC	0.842±0.071	0.696±0.108	0.625±0.099	—	—	0.662±0.093	0.681±0.080

Because source code from the LAMDA lab at Nanjing University was used, the AUC values for mi-graph and miVLAD are not available[28]. This implementation outputs only classification labels, not probabilities, preventing the calculation of AUC values.

3. Discussion

In tasks pertaining to skin cancer diagnosis, the hybrid U-Net and Improved MobileNet-V3 models perform better than the state-of-the-art designs, including MobileNet, VGG-16, and ResNet-152v2. The U-Net architecture excels in image segmentation, crucial for distinguishing skin tumors from surrounding tissues, while MobileNet-V3 is more appropriate for feature extraction. By combining these strengths, the hybrid model generates accurate segmentation masks and extracts meaningful features, resulting in enhanced model performance.

The proposed model's integration of two complex designs may result in significant computational demands during both training and inference, which is an important limitation to consider, especially in resource-constrained environments. Additionally, the interpretability of deep learning models like the one presented can be a common challenge. In clinical settings, despite the model's strong performance, understanding the precise factors underlying its classification decisions may be difficult. To address these limitations, future studies should focus on improving the model's computational efficiency through techniques like model compression or lightweight architectures. Addressing the interpretability issue is also crucial, as employing explainable AI approaches could help healthcare providers better understand the decision-making procedure of the model. This balanced approach, acknowledging both strengths and limitations, contributes to the ongoing discussion in automated skin cancer detection. Future research ought to integrate further clinical data, devise interpretability techniques customized for healthcare practitioners, and confront the moral ramifications of employing such AI models. Maintaining innovation remains central to making a meaningful impact in dermatology. However, it is imperative to acknowledge that although the model has exhibited impressive performance on the HAM-10000 dataset, additional validation is imperative to evaluate its efficacy in novel medical scenarios encountered in real-world settings[13].

With regard to the new attention-based MIL model for classifying lung cancer, it demonstrates the success in improving the model's interpretability through attention weights. The method showed promising results compared to traditional MIL techniques, achieving enhanced performance metrics such as accuracy, PPV, NPV, and AUC. Notably, the use of minority oversampling significantly improved these metrics, although some drop in recall was observed.

However, the study identified several limitations. First, it relied on existing lung nodule detection and segmentation methods, which may introduce interobserver variability, especially for small nodules. The relatively small sample size also impacted model stability, and no external dataset was available for independent validation. Additionally, the approach's dependence on manually created radiomic characteristics from low-dose CT scans may also generate noise, making it harder to apply to different datasets.

To address these issues in future work, the study suggested automating nodule detection and segmentation, improving the reliability of radiomic features, conducting large-scale evaluations of the attention mechanism, and performing external validation with newer CT technologies to better understand the model's robustness and generalizability.

4. Conclusion

The research outlined in this section highlights the significant advancements in deep learning models for the detection and diagnosis of two prevalent cancers—skin and lung cancer. The hybrid U-Net combined with the improved MobileNet-V3 model demonstrated exceptional diagnostic capabilities in skin cancer, surpassing established models such as MobileNet, VGG-16, and ResNet-152v2. With an accuracy of 98.86%, precision of 97.84%, and recall of 95.27%, this model showcases strong potential for clinical application. Its ability to accurately differentiate between malignant and benign cases could significantly enhance dermatological practices, enabling quicker and more reliable diagnoses, ultimately improving patient outcomes.

Similarly, the application of attention-based multiple instance learning (MIL) for lung cancer classification yielded promising results, particularly in comparison to traditional MIL methods. The use of attention processes not only enhanced the interpretability of the model but also improved its accuracy, positive predictive value, and area under the curve performance metrics. This study does, however, also highlight a number of significant difficulties, including the manual creation of radiomic characteristics, a small dataset, and the reliance on currently available nodule identification methods. Subsequent investigations ought to concentrate on mechanizing nodule examination, improving radiomic feature extraction, and carrying out extensive verifications using more varied and up-to-date CT imaging datasets.

In conclusion, while the deep learning models presented in this research have demonstrated considerable promise, further refinement, validation, and real-world testing are essential for their widespread adoption in clinical environments. By addressing current limitations and optimizing these models for interpretability and computational efficiency, the research community can accelerate the integration of AI-driven solutions into cancer diagnostics. Ultimately, these advancements could lead to earlier detection, more accurate diagnoses, and personalized treatment options, contributing to significantly improved outcomes in cancer care.

References

- [1] <https://www.who.int/news/item/01-02-2024-global-cancer-burden-growing--amidst-mounting-need-for-services#:~:text=In%202022%2C%20there%20were%20an,women%20die%20from%20the%20disease.>
- [2] Ott, J. J., Ullrich, A., & Miller, A. B. (2009). The importance of early symptom recognition in the context of early detection and cancer survival. *European journal of cancer* (Oxford, England : 1990), 45(16), 2743–2748. <https://doi.org/10.1016/j.ejca.2009.08.009>
- [3] Pulumati, A., Pulumati, A., Dwarakanath, B. S., Verma, A., & Papineni, R. V. L. (2023). Technological advancements in cancer diagnostics: Improvements and limitations. *Cancer reports* (Hoboken, N.J.), 6(2), e1764. <https://doi.org/10.1002/cnr2.1764>
- [4] Tempany, C. M., Jayender, J., Kapur, T., Bueno, R., Golby, A., Agar, N., & Jolesz, F. A. (2015). Multimodal imaging for improved diagnosis and treatment of cancers. *Cancer*, 121(6), 817–827. <https://doi.org/10.1002/cncr.29012>
- [5] Yaqoob, A., Musheer Aziz, R. & verma, N.K. Applications and Techniques of Machine Learning in Cancer Classification: A Systematic Review. *Hum-Cent Intell Syst* 3, 588–615 (2023). <https://doi.org/10.1007/s44230-023-00041-3>
- [6] <https://www.foreseemed.com/blog/machine-learning-in-healthcare#:~:text=Machine%20learning%20in%20healthcare%20examples%20include%20diagnostic%20support%20systems%2C%20risk,insights%20derived%20from%20vast%20datasets.>

- [7] Liu, L., Ma, C., Zhang, Z., Witkowski, M. T., Aifantis, I., Ghassemi, S., & Chen, W. (2022). Computational model of CAR T-cell immunotherapy dissects and predicts leukemia patient responses at remission, resistance, and relapse. *Journal for immunotherapy of cancer*, 10(12), e005360. <https://doi.org/10.1136/jitc-2022-005360>
- [8] Silva, H. E. C. D., Santos, G. N. M., Leite, A. F., Mesquita, C. R. M., Figueiredo, P. T. S., Stefani, C. M., & de Melo, N. S. (2023). The use of artificial intelligence tools in cancer detection compared to the traditional diagnostic imaging methods: An overview of the systematic reviews. *PloS one*, 18(10), e0292063.
- [9] <https://www.iarc.who.int/cancer-type/skin-cancer/#:~:text=Introduction&text=Skin%20cancers%20are%20the%20most,in%20men%20than%20in%20women.>
- [10] Mohakud, R., & Dash, R. (2023). A Hybrid Model for Classification of Skin Cancer Images After Segmentation. *International Journal of Image and Graphics*, 2550022. <https://doi.org/10.1142/S0219467825500226>
- [11] Sharma, G., & Chadha, R. (2023). An Optimized Predictive Model Based on Deep Neural Network for Detection of Skin Cancer and Oral Cancer. <https://doi.org/10.1109/INOCON57975.2023.10101118>
- [12] Behara, K., Bhero, E., & Agee, J. T. (2023). Skin Lesion Synthesis and Classification Using an Improved DCGAN Classifier. *Diagnostics (Basel, Switzerland)*, 13(16), 2635. <https://doi.org/10.3390/diagnostics13162635>
- [13] Kumar Lilhore, U., Simaiya, S., Sharma, Y. K., Kaswan, K. S., Rao, K. B. V. B., Rao, V. V. R. M., Baliyan, A., Bijalwan, A., & Alroobaea, R. (2024). A precise model for skin cancer diagnosis using hybrid U-Net and improved MobileNet-V3 with hyperparameters optimization. *Scientific reports*, 14(1), 4299. <https://doi.org/10.1038/s41598-024-54212-8>
- [14] <https://arxiv.org/pdf/1703.04197>
- [15] Sayed, G. I., Soliman, M. M., & Hassanien, A. E. (2021). A novel melanoma prediction model for imbalanced data using optimized SqueezeNet by bald eagle search optimization. *Computers in biology and medicine*, 136, 104712. <https://doi.org/10.1016/j.compbiomed.2021.104712>
- [16] Kumar Lilhore, U., Simaiya, S., Sharma, Y. K., Kaswan, K. S., Rao, K. B. V. B., Rao, V. V. R. M., Baliyan, A., Bijalwan, A., & Alroobaea, R. (2024). A precise model for skin cancer diagnosis using hybrid U-Net and improved MobileNet-V3 with hyperparameters optimization. *Scientific reports*, 14(1), 4299. <https://doi.org/10.1038/s41598-024-54212-8>
- [17] <https://www.wcrf.org/cancer-trends/lung-cancer-statistics/>
- [18] Bhattacharjee, K., Pant, M., Zhang, Y.-D., & Satapathy, S. C. (2020). Multiple Instance Learning with Genetic Pooling for medical data analysis. *Pattern Recognition Letters*, 133, 247-255. <https://doi.org/https://doi.org/10.1016/j.patrec.2020.02.025>
- [19] Bhattacharjee, K., Pant, M., Zhang, Y.-D., & Satapathy, S. C. (2020). Multiple Instance Learning with Genetic Pooling for medical data analysis. *Pattern Recognition Letters*, 133, 247-255. <https://doi.org/https://doi.org/10.1016/j.patrec.2020.02.025>
- [20] Frade, J., Pereira, T., Morgado, J., Silva, F., Freitas, C., Mendes, J., Negrão, E., de Lima, B. F., Silva, M. C. D., Madureira, A. J., Ramos, I., Costa, J. L., Hespanhol, V., Cunha, A., & Oliveira, H. P. (2022). Multiple instance learning for lung pathophysiological findings detection using CT scans. *Medical & biological engineering & computing*, 60(6), 1569–1584. <https://doi.org/10.1007/s11517-022-02526-y>
- [21] Andrews, S., Tsochantaridis, I., & Hofmann, T. (2002). Support vector machines for multiple-instance learning *Proceedings of the 15th International Conference on Neural Information Processing Systems*.
- [22] Zhou ZH, Sun YY, Li YF. Multi - instance learning by treating instances as non - iid samples. In: *Proceedings of the 26th Annual International Conference on Machine Learning* ; 2009: 1249 - 1256.

- [23] Wei, X. S., Wu, J., & Zhou, Z. H. (2017). Scalable Algorithms for Multi-Instance Learning. *IEEE transactions on neural networks and learning systems*, 28(4), 975–987. <https://doi.org/10.1109/TNNLS.2016.2519102>
- [24] Armato, S. G., 3rd, McLennan, G., Bidaut, L., McNitt-Gray, M. F., Meyer, C. R., Reeves, A. P., Zhao, B., Aberle, D. R., Henschke, C. I., Hoffman, E. A., Kazerooni, E. A., MacMahon, H., Van Beeke, E. J., Yankelevitz, D., Biancardi, A. M., Bland, P. H., Brown, M. S., Engelmann, R. M., Laderach, G. E., Max, D., ... Croft, B. Y. (2011). The Lung Image Database Consortium (LIDC) and Image Database Resource Initiative (IDRI): a completed reference database of lung nodules on CT scans. *Medical physics*, 38(2), 915–931. <https://doi.org/10.1118/1.3528204>
- [25] Dhara, A. K., Mukhopadhyay, S., Dutta, A., Garg, M., & Khandelwal, N. (2016). A Combination of Shape and Texture Features for Classification of Pulmonary Nodules in Lung CT Images. *Journal of digital imaging*, 29(4), 466–475. <https://doi.org/10.1007/s10278-015-9857-6>
- [26] Chen, J., Zeng, H., Zhang, C., Shi, Z., Dekker, A., Wee, L., & Bermejo, I. (2022). Lung cancer diagnosis using deep attention-based multiple instance learning and radiomics. *Medical physics*, 49(5), 3134–3143. <https://doi.org/10.1002/mp.15539>
- [27] Hancock, M. C., & Magnan, J. F. (2016). Lung nodule malignancy classification using only radiologist-quantified image features as inputs to statistical learning algorithms: probing the Lung Image Database Consortium dataset with two statistical learning methods. *Journal of medical imaging (Bellingham, Wash.)*, 3(4), 044504. <https://doi.org/10.1117/1.JMI.3.4.044504>
- [28] Learning and Mining from Data . [WWW document]. URL <http://210.28.132.67/Data.ashx>. [accesses on 4 Jun 2021], 2021.