WEP Warwick Evans Publishing

International Journal of Computer Science and Information Technology

ISSN: 3005-9682 (Print), ISSN: 3005-7140 (Online) | Volume 1, Number 1, Year 2023
DOI: https://doi.org/10.62051/ijcsit.v1n1.10
Journal homepage: https://wepub.org/index.php/IJCSIT/index

Implementation of computer vision technology based on artificial intelligence for medical image analysis

Danqing Ma[1, *], Bo Dang[2], Shaojie Li[3], Hengyi Zang[4], Xinqi Dong[1]

[1] Computer Science, Independent research, Beijing, China
[2] Computer Science, Independent research, Fremont CA, USA
[3] Computer Technology, Independent Research, Beijing, China
[4] Big Data and Business Intelligence, Independent research, Beijing, China
*Corresponding Author E-mail: 3530761316qq@gmail.com

ABSTRACT

As one of the branches of machine learning, the deep learning model combined with artificial intelligence is widely used in the field of computer vision technology, and the image recognition field represented by medical image analysis is also developing. Its advantage is that it does not rely on human annotation, and the computer can recognize and process the feature information omitted by human beings during the model training process, so as to achieve or even exceed the accuracy of human processing. Based on the general lack of explain ability caused by the unknown data processing process in the deep model, the existing solutions mainly include the establishment of internal explain ability, attention mechanism interpretation of specific models, and the interpretation of unknowable models represented by LIME. The way to quantitatively assess interpretability is still being explored, especially in the interpretative assessment of both doctors and patients in medical decision-related models, several scales have been proposed for reference. The current research on the application of artificial intelligence deep learning models in medical imaging generally pays more attention to accuracy rather than explain ability, resulting in the lack of explain ability, and thus hindering the practical clinical application of deep learning models. Therefore, the need to analyze the development of medical image analysis in the field of artificial intelligence and computer vision technology, and how to balance accuracy and interpretability to develop deep learning models that both doctors and patients can trust will become the research focus of the industry in the future.

KEYWORDS

Medical image analysis; Deep learning model; Computer vision technology; Image segmentation.

1. INTRODUCTION

In recent years, deep learning technology based on artificial intelligence has begun to be applied in the field of image processing and assisted diagnosis of medical images. By using deep learning algorithms, it is possible to train medical image processing and focus analysis models with a high level of automation, which can quickly and accurately analyze and diagnose many types of diseases. The application of this technology can greatly improve the work efficiency and accuracy of doctors, while reducing medical errors and unnecessary medical treatment costs for patients [1]. As one of the main research directions in the computer field, artificial intelligence is based on computer programs and algorithms to simulate human learning, reasoning and decision-making behavior, so that the machine as an intelligent entity can imitate the human intelligence gate to the maximum extent. At present, machine learning is the main method to realize artificial intelligence, and its model training does not rely on rule programming, but calculates model weights through input training samples [2], which has been widely used in finance, medicine and other fields [3]. Traditional machine learning methods include naive Bayes, decision tree, support vector machine, etc., whose structure is simple and the decisions made are easy to be understood by humans. However, traditional machine learning relies on domain experts for display feature extraction, and has feature selection limitations, which has poor performance when scaling to large data sets. For the huge and rapidly evolving medical field, traditional machine learning does not make full use of the available data. The concept of deep learning based on deep neural networks provides a new way to improve the training and efficiency of traditional machine learning models.

However, intelligent analysis of multimodal medical image data remains a challenging task. Multimodal medical imaging technology refers to the use of different imaging techniques, different imaging sequences, different imaging angles to examine the same patient, in order to provide more comprehensive and accurate diagnostic information. Multimodal medical image data has a large and complex scale, and each modal image has unique characteristics and information, including different imaging parameters, different entry angles, and different physical meanings [4-5]. In the process of deep learning model training, how to use the prior knowledge of human body structure, disease causes or imaging principles to guide the training and analysis of deep learning models, improve training efficiency and prevent overfitting is one of the problems to be solved in the current research. Moreover, the image data of multimodal medical images tend to have more severe noise and interference, and the patient's Angle, position, and posture can also change during imaging, which makes intelligent analysis more challenging. Finally, compared with natural images, there is a problem of unbalanced sampling and classification in medical image data, which has caused some obstacles to the cross-domain application of deep learning methods. In this paper, the application space of computer vision technology in the acquisition, mining, fusion and interpretation of image information in multi-modal medical imaging is discussed, and the overfitting phenomenon in the cross-domain application of deep learning and the label-free evaluation method of depth model are discussed.

2. RELATED WORK

With the development of medical imaging technology, more and more imaging and detection methods have been applied in the early diagnosis of diseases, such as X-ray imaging (X-ray), Ultrasound[maging], Computed Tomography (Computed Tomography). CT), Magnetic Resonance Imaging (MRI), and Nuclear Imaging technology (Nuclear Imaging). [6-7]In clinical work, due to the limitations of different imaging principles and imaging equipment, the single form and single mode of medical imaging can not fully and accurately reflect the organ shape and lesion characteristics of patients.

2.1. Computer vision domain(CV)

Computer Vision (Computer Vision) refers to the use of computers to achieve human visual functions - perception, recognition and understanding of the three-dimensional scene of the objective world. This means that the research goal of computer vision technology is to make the computer have the ability to recognize three-dimensional environment information through two-dimensional images. Therefore, it is necessary not only to enable the machine to perceive the geometric information (shape, position, attitude, movement, etc.) of objects in the three-dimensional environment, but also to describe, store, recognize and understand them[8]. Computer vision can be considered different from the study of human or animal vision: it uses geometry, physics, and learning techniques to build models, and uses statistical methods to process data. The complete closed loop of artificial intelligence includes the process of perception, cognition, reasoning and then feedback to perception, in which vision occupies most of the perceptual process in our perceptual system. So studying vision is an important step in studying computer perception.

Therefore, you can follow the following steps to break it down[9]:

The input is a training set of N images with K categories, each of which is labeled as one of the categories.

The training set is then used to train a classifier to learn the external features of each class.

Finally, we predict the class labels of a new set of images and evaluate the performance of the classifier. We compare the class labels predicted by the classifier with the real class labels.

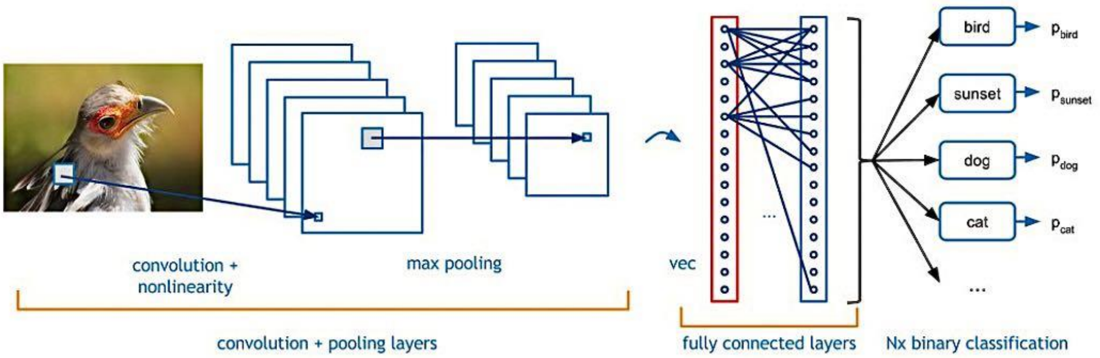

Figure 1. Computer vision image classification architecture is Convolutional Neural Network (CNN)

Currently, computer vision is one of the hottest research areas in deep learning. Computer vision is actually an interdisciplinary discipline that encompasses computer science (graphics, algorithms, theory, systems, architecture), mathematics (information retrieval, machine learning), engineering (robotics, speech, natural language processing, image processing), physics (optics), biology (neuroscience), and psychology (cognitive science), among others[10-11]. Many scientists believe that computer vision has paved the way for the development of artificial intelligence.

2.2. Medical image analysis

The combination of image omics and machine learning technology is a common method for AI image aided diagnosis. Imaging omics extracts and analyzes imaging features in high throughput from medical images such as CT, PET or MRI, and then realizes clinical applications such as diagnosis, evaluation and prediction of diseases based on machine learning algorithms such as image segmentation, extraction and screening features, and construction of prediction models[12]. Figure 2 shows the application of imaging omics and machine learning techniques in medical image analysis. In foreign countries, imaging omics is widely used to predict the prognosis and survival of cancer patients. For example, DERCLE et al. P used imaging omics and machine learning methods to predict the survival rate of patients with liver metastatic colorectal cancer based on CT images. DISSAUX et al. used a fuzzy local adaptive Bayesian algorithm to semi-automatically map tumors in PET images, extracted 184 image-omics features, and finally screened out 2 important features for local recurrence prediction in early non-small cell lung cancer patients[13-15]. AN et al. "Established a machine learning model for predicting prognosis of patients with enter opancreatic neuroendocrine tumors based on CT imaging and clinical data. In China, there are a large number of studies based on CT and MRI images and imaging omics methods for pathological grading lesion detection, disease diagnosis, metastasis prediction, prognosis prediction and other directions."

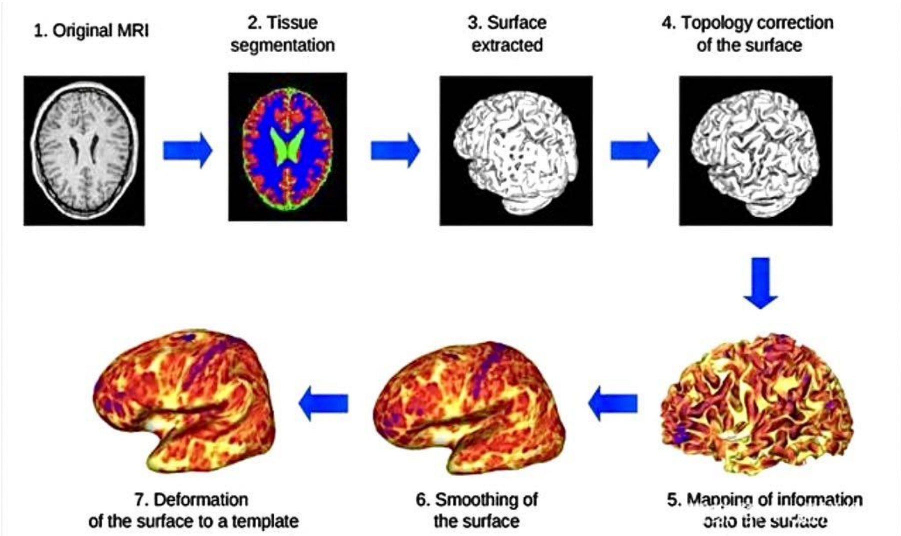

Figure 2:

Figure 2. Advanced medical image processing :MRI analysis

(1) The tumor is connected to the dura mater with a broad base, and the "cortical compression sign" (also known as "cortical collapse sign") can be seen in the inner margin, and there is a low signal boundary between the mass and the brain tissue or the formation of arachnoid cysts[16].

(2) T1WI or slightly lower signals, T2WI and T2 FLAIR magnitude or slightly higher signals, more uniform signals; The signal on DWI is uniform and slightly high to high, while the signal on sand is low

(3) Most of the tumors in the enhanced scan showed obvious homogeneous enhancement, and the "meningeal tail sign" was visible.

(4) Atypical signs: a few meningiomas may calcify throughout the tumor; Cystic meninges can be presented as cystic lesions, and the enhanced scan shows annular enhancement of the cyst wall. Some patients may present with multiple meningiomas, with imaging manifestations similar to single meningiomas.

3. METHODOLOGY

3.1. Image main processing flow

First, the infarct area is automatically segmented on the ADC map by applying normalized absolute thresholds. The second step is to construct quantile curves of the ADC intensity in the brain mask for each ADC plot of each subject, and to identify the intersection points between 2 tangents with maximum and minimum differential coefficients on each quantile intensity curve[17]. (This sentence does not understand, probably means that the account of automatic segmentation threshold is how to choose). The third step is to realize the normalization of ADC graph. In the fourth step, the threshold value of each normalized ADC map is the best absolute value 0.845, and finally, FLAIR images are jointly registered to the ADC map (as shown in Figure 3).

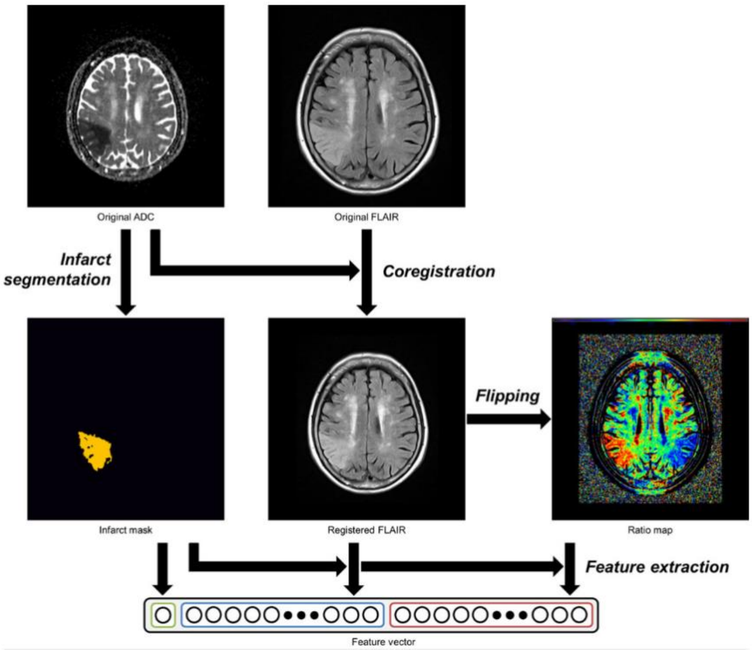

Figure 3. Medical image processing: Automatically segment the infarct area on the ADC map

3.2. Image extraction and generation

In the process of extracting image features, it is necessary to extract intensity, gradient and texture information from registered FLAIR images and FLAIR ratio graphs, and construct ratio graphs by reflecting images around the sagittal plane in FTD, so as to quantitatively compare the relative signals of the infarction area and the opposite side.

Table 1. Image Features Extracted Within the Infarct Regions

Feature Categories	Number of Features	Feature Names
Infarct volume		
Size	1	Volume
Features from the original image and the ratio map (44 features for each)		
Intensity	4	Mean, SD, skewness, kurtosis
Gradient	4	Mean, SD, skewness, kurtosis
GLCM	21	Energy, entropy, dissimilarity, contrast, ID, correlation, homogeneity, autocorrelation, cluster shade, cluster prominence, maximum probability, sum of squares, sum average, sum variance, sum entropy, difference variance, difference entropy, IOC1, IOC2, INN, IDN
GLRLM	11	SRE, LRE, GLN, RLN, RP, LGRE, HGRE, SRLGE, SRHGE, LRLGE, LRHGE
LBP	4	Mean, SD, skewness, kurtosis

GLCM indicates gray level co-occurrence matrix; GLN, gray level nonuniformity; GLRLM, gray level run-length matrix; HGRE, high gray run emphasis; ID, inverse difference; IDN, inverse difference moment normalized. INN, inverse difference normalized; 10C1, information of correlation 1[18-20]; I0C2information of correlation 2: LBP, local binary pattern; LGRE, low gray run emphasis; LRE, long run emphasis; LRHGE, long run high gray emphasis. LRLGE, long run low gray emphasis; RLN, run length nonuniformity; RP, run percentage; SRE, short run emphasis; SRHGE, short run high gray emphasis. And SRLGE, short run low gray emphasis.

3.3. Statistics and results

A. The single-factor T-test was used to screen 89 features, and the P-values were corrected using Bonferroni correction. (Here the researchers set the P-value at 0.2, which is very small if it is the conventional 0.05 and then divided by 89); If the number of features after correction is less than 5, the first 5 are selected according to the P-value.

B. Performance differences between machine learning models and human vision.

In the analysis of experimental results, the grouping situation (training set, test set), the distribution of baseline data and target are described first.

Naturally, there was little statistical difference between the two groups. Feature screening was performed by single factor analysis. A total of 34 features were screened for ML modeling (LR, SVM, RF). The best performers in each category were evaluated in the test set. Although RF had the largest AUROC, there was no statistical difference between the three.

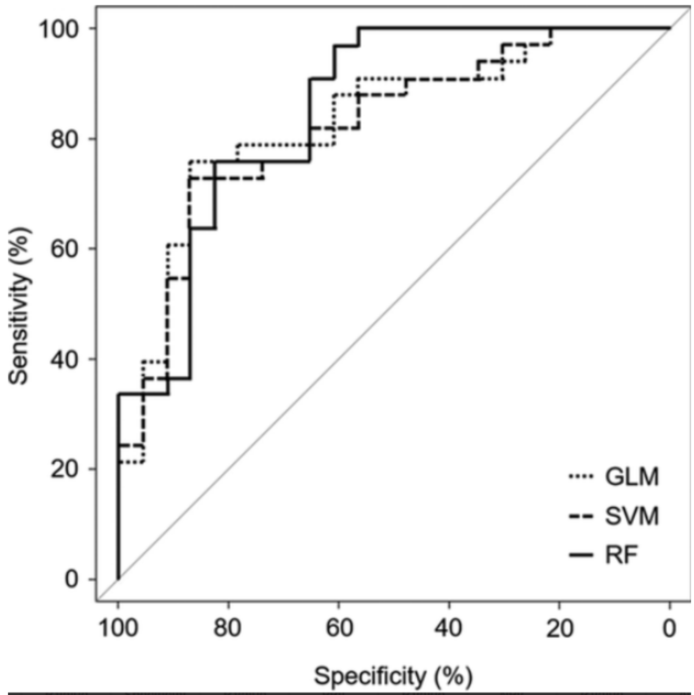

Figure 4. Filter data difference curves

The SVM model has two very important parameters C and gamma.Where C is the penalty coefficient, that is, the tolerance for error. The higher c is, the less error is tolerated and the easier it is to overfit. The smaller the C, the easier it is to underfit[21]. C is too large or too small, and the generalization ability becomes poor. gamma is an argument that comes with the RBF function after it is selected as the kernel. Implicitly determines the distribution after the number is mapped to the new feature space, the larger the gamma, the less the support vector, the smaller the gamma, the more the support vector. The number of support vectors affects the speed of training and prediction.

In summary, in this study, the analysis of consistency between the measurers also showed that the deep learning approach influenced clinical decisions for mpMRI [22]classification. Compared to the reference results, the deep learning method had a higher rate of evaluator agreement than the resident physicians and experienced imaging physicians. This shows that our method can not only greatly improve the diagnostic accuracy of interns, but also provide some help for the manual diagnosis of experts.

4. CONCLUSION

Currently, in the field of medical imaging, most of the existing interpretable deep learning models focus on the identification and interpretation of individual diseases. However, in practical medical decision-making, clinicians are often faced with the problem of combining various information to predict unknown diseases, and rarely deal with the problem of single disease judgment[23-26]. Developing interpretable depth models that integrate multiple medical disciplines in order to fit clinical needs is a challenge shared by clinicians and deep learning developers today. Therefore, the expansion of computer vision technology combined with artificial intelligence in all walks of life is the current trend, especially in medical image-based medical decision support has great potential, interpretability is a big help to promote practical applications[27]. This paper presents the different definitions and understandings of explain ability, and classifies and introduces the existing methods of realizing and evaluating explain ability based on the practical cases of medical image analysis. Finally, the present problems and possible development direction of the explainable model are analyzed[28]. By reviewing the latest research progress and practical application of explainable deep learning model in the field of medical imaging, it can promote the understanding of clinicians and artificial intelligence developers on interdisciplinary subjects. As the construction of explain ability of deep model requires both professional domain knowledge and computer literacy, training medical and computer interdisciplinary talents will become a major industry direction in the future.

ACKNOWLEDGEMENT

The research in this paper is largely inspired by Dr. Wenjian Sun's in-depth research and contributions in the field of electronic information engineering. Dr. Sun's advanced research in the application of artificial intelligence and computer vision technology to medical image analysis has provided valuable theoretical foundation and methodological guidance for our work. We would like to express our deep respect and gratitude to Dr. Wenjian Sun for his innovative work, which not only enriches the academic resources in the field, but also opens up new horizons for our research. We look forward to continuing to explore and expand the application potential of artificial intelligence in medical image analysis on this basis.

REFERENCES

[1] GILLIES R J,KINAHAN P E,HRICAK H. Radiomics:images are more than pictures, they are data[J]Radiology, 2016,278(2): 563-577.

[2] CAI J H, ZHENG J , SHEN J, et al. A radiomics model forpredicting the response to bevacizumab in brain necrosisafter radiotherapy[J]. Clin Cancer Res, 2020, 26(20): 1264.

[3] DERCLE L.LU L,SCHWARTZ L H,et al. Radiomicsresponse signature for identification of metastaticcolorectal cancer sensitive to therapies targeting EGFRpathway[J]. J Natl Cancer Inst, 2020,112(9): 902-912.

[4] Xinyu Zhao, et al. “Effective Combination of 3D-DenseNet’s Artificial Intelligence Technology and Gallbladder Cancer Diagnosis Model”. Frontiers in Computing and Intelligent Systems, vol. 6, no. 3, Jan. 2024, pp. 81-84, https://doi.org/10.54097/iMKyFavE.

[5] Shulin Li, et al. “Application Analysis of AI Technology Combined With Spiral CT Scanning in Early Lung Cancer Screening”. Frontiers in Computing and Intelligent Systems, vol. 6, no. 3, Jan. 2024, pp. 52-55, https://doi.org/10.54097/LAwfJzEA.

[6] [1] Liu, Bo & Zhao, Xinyu & Hu, Hao & Lin, Qunwei & Huang, Jiaxin. (2023). Detection of Esophageal Cancer Lesions Based on CBAM Faster R-CNN. Journal of Theory and Practice of Engineering Science. 3. 36-42. 10.53469/jtpes.2023.03(12).06.

[7] Yu, Liqiang, et al. “Research on Machine Learning With Algorithms and Development”. Journal of Theory and Practice of Engineering Science, vol. 3, no. 12, Dec. 2023, pp. 7-14, doi:10.53469/jtpes.2023.03(12).02.

[8] Xin, Q., He, Y., Pan, Y., Wang, Y., & Du, S. (2023). The implementation of an AI-driven advertising push system based on a NLP algorithm. International Journal of Computer Science and Information Technology, 1(1), 30-37.0

[9] Zhou, H., Lou, Y., Xiong, J., Wang, Y., & Liu, Y. (2023). Improvement of Deep Learning Model for Gastrointestinal Tract Segmentation Surgery. Frontiers in Computing and Intelligent Systems, 6(1), 103-106.6

[10] Implementation of an AI-based MRD Evaluation and Prediction Model for Multiple Myeloma. (2024). Frontiers in Computing and Intelligent Systems, 6(3), 127-131. https://doi.org/10.54097/zJ4MnbWW.

[11] Zhang, Q., Cai, G., Cai, M., Qian, J., & Song, T. (2023). Deep Learning Model Aids Breast Cancer Detection. Frontiers in Computing and Intelligent Systems, 6(1), 99-102.3

[12] Xu, J., Pan, L., Zeng, Q., Sun, W., & Wan, W. (2023). Based on TPUGRAPHS Predicting Model Runtimes Using Graph Neural Networks. Frontiers in Computing and Intelligent Systems, 6(1), 66-69.7

[13] Wan,Weixiang, et al. "Development and Evaluation of Intelligent Medical Decision Support Systems." Academic Journal of Science and Technology 8.2 (2023): 22-25.

[14] Tian, M., Shen, Z., Wu, X., Wei, K., & Liu, Y. (2023). The Application of Artificial Intelligence in Medical Diagnostics: A New Frontier. Academic Journal of Science and Technology, 8(2), 57-61.7

[15] Shen, Z., Wei, K., Zang, H., Li, L., & Wang, G. (2023). The Application of Artificial Intelligence to The Bayesian Model Algorithm for Combining Genome Data. Academic Journal of Science and Technology, 8(3), 132-135.2

[16] Zheng He, et al. “The Importance of AI Algorithm Combined With Tunable LCST Smart Polymers in Biomedical Applications”. Frontiers in Computing and Intelligent Systems, vol. 6, no. 3, Jan. 2024, pp. 92-95, https://doi.org/10.54097/d30EoLHw.

[17] Prediction of Atmospheric Carbon Dioxide Radiative Transfer Model based on Machine Learning. (2024). Frontiers in Computing and Intelligent Systems, 6(3), 132-136. https://doi.org/10.54097/ObMPjw5n

[18] Liu, Y., Duan, S., Shen, Z., He, Z., & Li, L. (2023). Grasp and Inspection of Mechanical Parts based on Visual Image Recognition Technology. Journal of Theory and Practice of Engineering Science, 3(12), 22-28.1

[19] Xinyu Zhao, et al. “Effective Combination of 3D-DenseNet’s Artificial Intelligence Technology and Gallbladder Cancer Diagnosis Model”. Frontiers in Computing and Intelligent Systems, vol. 6, no. 3, Jan. 2024, pp. 81-84, https://doi.org/10.54097/iMKyFavE.

[20] Liu, B. (2023). Based on intelligent advertising recommendation and abnormal advertising monitoring system in the field of machine learning. International Journal of Computer Science and Information Technology, 1(1), 17-23.

[21] Pan, Linying, et al. “Research Progress of Diabetic Disease Prediction Model in Deep Learning”. Journal of Theory and Practice of Engineering Science, vol. 3, no. 12, Dec. 2023, pp. 15-21, doi:10.53469/jtpes.2023.03(12).03.

[22] K. Tan and W. Li, "Imaging and Parameter Estimating for Fast Moving Targets in Airborne SAR," in IEEE Transactions on Computational Imaging, vol. 3, no. 1, pp. 126-140, March 2017, doi: 10.1109/TCI.2016.2634421.

[23] Zhou, H., Lou, Y., Xiong, J., Wang, Y., & Liu, Y. (2023). Improvement of Deep Learning Model for Gastrointestinal Tract Segmentation Surgery. Frontiers in Computing and Intelligent Systems, 6(1), 103-106.

[24] Liu, S., Wu, K., Jiang, C., Huang, B., & Ma, D. (2023). Financial Time-Series Forecasting: Towards Synergizing Performance And Interpretability Within a Hybrid Machine Learning Approach. arXiv preprint arXiv:2401.00534.

[25] Su, J., Nair, S., & Popokh, L. (2023, February). EdgeGym: A Reinforcement Learning Environment for Constraint-Aware NFV Resource Allocation. 2023 IEEE 2nd International Conference on AI in Cybersecurity (ICAIC), 1–7. doi:10.1109/ICAIC57335.2023.10044182

[26] Su, J., Nair, S., & Popokh, L. (2022, November). Optimal Resource Allocation in SDN/NFV-Enabled Networks via Deep Reinforcement Learning. 2022 IEEE Ninth International Conference on Communications and Networking (ComNet), 1–7. doi:10.1109/ComNet55492.2022.9998475

[27] Popokh, L., Su, J., Nair, S., & Olinick, E. (2021, September). IllumiCore: Optimization Modeling and Implementation for Efficient VNF Placement. 2021 International Conference on Software, Telecommunications and Computer Networks (SoftCOM), 1–7. doi:10.23919/SoftCOM52868.2021.9559076

[28] Bao, W., Che, H., & Zhang, J. (2020, December). Will_Go at SemEval-2020 Task 3: An Accurate Model for Predicting the (Graded) Effect of Context in Word Similarity Based on BERT. In A. Herbelot, X. Zhu, A. Palmer, N. Schneider, J. May, & E. Shutova (Eds.), Proceedings of the Fourteenth Workshop on Semantic Evaluation (pp. 301–306). doi:10.18653/v1/2020.semeval-1.