A Technical Comparison of YOLO-Based Chest Cancer Diagnosis Methods

Juanlang Shi*

Department of Applied Physics and Applied Mathematics, Columbia University, 116th and Broadway, New York, New York, 10027, United States

Corresponding author email: js5952@columbia.edu

Abstract. Cancers have become one of the deadliest diseases in the world, and early diagnosis becomes vital for a patient's survival. As deep learning advances, YOLO has become an attractive tool as it supports real-time interactions. Thus, YOLO is expected to be applied in cancer diagnosis. A technical study of a YOLO-based computer aid diagnosis system for chest cancers is presented in the paper. Four kinds of the image in cancer diagnosis, histopathological images, mammograms, CTs, and Low-dose CTs, are introduced. Three issues of implementing a computer aid diagnosis system (CAD) are discussed and analyzed, including the usage of handcrafted features, the high false positive rate in clinical practice, and difficulty in detecting irregular nodules in spiral CTs. In discussion, the drawback of handcrafted features in the region of interest (ROI) extraction can be addressed by applying extra architectures like ResNet50 as extractors. A trained network can serve as a non-nodule filter to reduce the false positive rate in diagnosis. Image data can be categorized based on morphological features in data preprocessing to train a more sensitive model, then irregular-shape nodules can be detected by CAD.

Keywords: CNN, YOLO, Chest Cancer, Image Classification.

1. Introduction

In recent decades, modern medicine and pharmacy went through big progress, and many formerly deadly diseases have become curable. At the same time, chest cancer remains one of the deadliest diseases in the world. Due to the nature of cancer, the most challenging and important part of a patient’s survival is early diagnosis. Cancer diagnosis depends on images such as histopathological images, mammography, and computed tomography (CT). Diagnosis is hence drawn based on handcrafted features summarized in the prior experience. However, the high variation of appearance and subtle differences in features make it a tough job to classify or determine the lesions in an image.

As technology advances, convolutional neural networks (CNN) emerge and have proven their potential in image classification and feature detection. CNN can extract more hierarchical features in images, thus attracting a lot of attention with a promising application in cancer diagnosis. It is integrated into the computer-aided diagnosis (CAD) system, which aims to provide an essential second opinion for physicians. You Only Look Once (YOLO), as a new CNN architecture, has become a hit in cancer diagnosis since YOLO integrates object detection and classification in the same CNN architecture, and it supports real-time interactions. Nevertheless, several issues prevent applying YOLO-based CAD in actual practice.

Though CNN has proved its power in many image classification tasks, it encounters difficulties in the medical field due to the lack of available data, and CNN cannot extract hierarchical features as other tasks. In data preparation, image data are labeled based on handcrafted features summarized from prior experience, while more hierarchical features are hardly discovered. The high variance of appearance and subtle differences in handcrafted features make handcrafted features not useful in image classification. Although the YOLO algorithm can detect ROI in the image automatically, all ROIs are extracted manually based on handcrafted features in training. The uncertainty and complexity of image morphologies lead to low accuracy and reliability in the extraction of ROI.

In actual practice, CAD structures also suffer from multiple factors. In clinical practice, CAD is designed to detect all true positives, but it also leads to an increase in false positives and low accuracy in prediction. Though CAD only serves as a second opinion, it can mislead physicians. In lung nodule detection for lung cancer diagnosis, CAD suffers in detecting nodules with an irregular shape and thus has difficulties finding nodules on lung walls, which irregular shape is caused by the attachment between nodules and lungs.

The application cannot be promising without addressing these issues properly. In this paper, these drawbacks are discussed and analyzed. The potential solution includes applying extra architectures like AlexNet and ResNet to extract ROI before the YOLO detector, applying reduced IOU and non-nodule filter in architecture to reduce false positives, and training the model with generated nodule models.

The paper is organized as follows: first, the structure of CNN and steps for YOLO are introduced, then three kinds of medical image data are introduced, and the role of YOLO in processing these data is also presented. The challenge in the application is shown in section 2 and discussed deeper in section 3. Suggested solutions for these issues are presented in section 4. Finally, the conclusion of the work is in section 5.

2. Case Description

2.1. YOLO Architecture

YOLO is a newly emerged CNN structure that integrates object detection and object classification in the same deep learning structure. Conventional CNN structures stack several convolutional and pooling layers after the input, several fully connected layers following the convolutional blocks, and a softmax function output in a general image classification model. Fig 1 [1] shows a CNN structure using DarkNet Architecture and return a tensor with coordinates and class probabilities of detected objects.

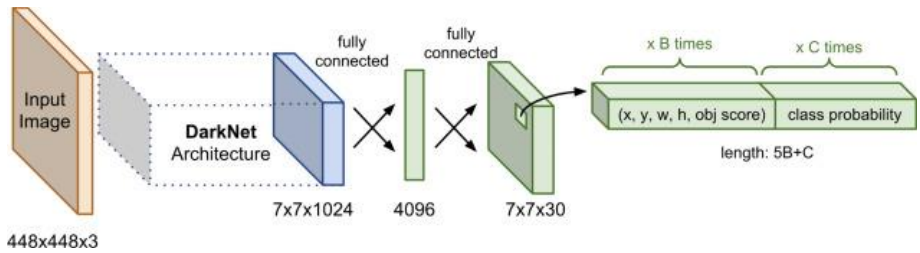


Figure 1. A CNN structure

YOLO architecture, as shown in Fig 2 from [2], enables CNN to detect and classify objects at a speed of more than 45 frames per second and hence supports real-time interactions. YOLO performs the following steps in image classification. In detection, first, run a sliding window on the image. The information in each window is turned into a tensor of information, including the coordinates of the window and probabilities of each class to be detected, then determine the coordinates of the bounding box, which contains the largest probability. The size of the bounding box is constrained by intersection over union (IOU). Then, CNN classifies the objects in the bounding boxes and returns the result of the classification.

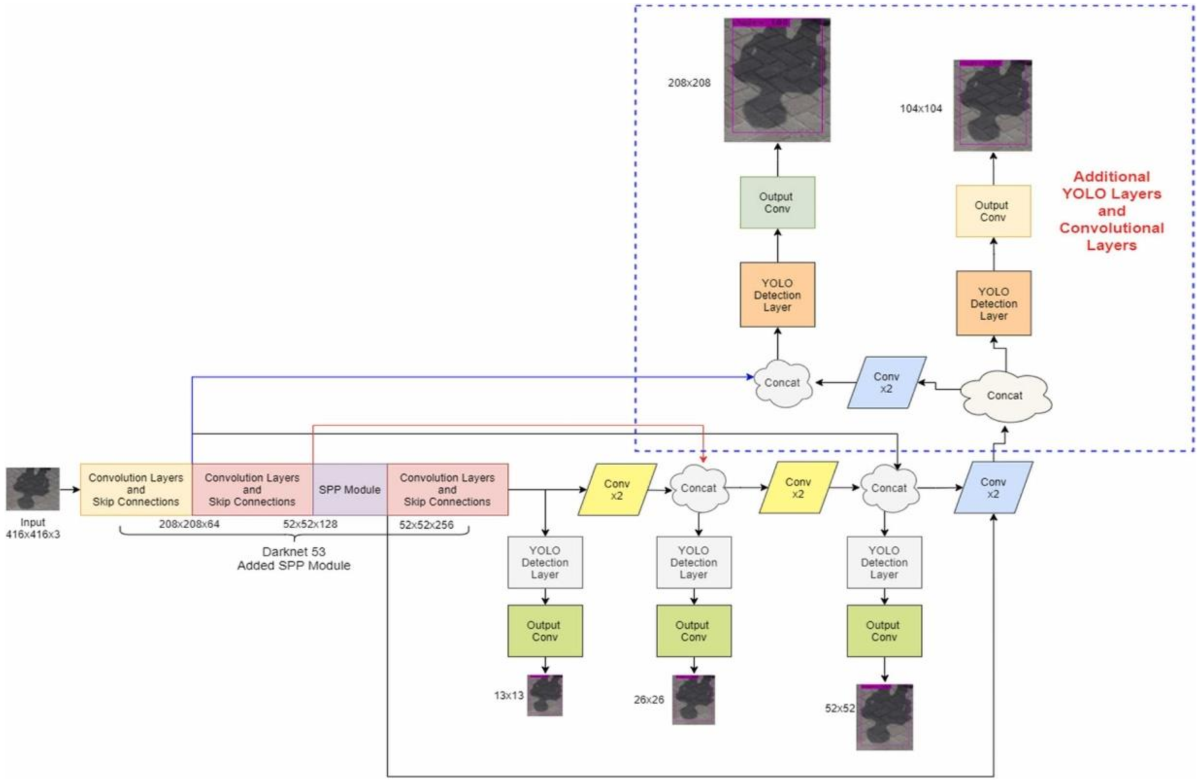


Figure 2. A YOLO Architecture

In medical practice, YOLO is usually assembled in CAD in chest cancer diagnosis. As a method of image detection and classification, YOLO has been applied to diagnose the existence of cancer from various data sources such as histopathological images, mammograms, CT, and low-dose CT scans (LDCT). With the huge advances in deep learning, CAD can capture and detect more hierarchical features in images, thus providing more reliable predictions. As the cancer rate increases, the work of diagnosis for physicians becomes overwhelming, and CAD provides an essential second opinion as a reference for physicians, optimizing the treatment plans.

2.2. Histopathological Images

Histopathological images are microscopic photos of tissue samples fixed and stained onto slides (Fig 3 from [3]). In the context of breast cancer, histopathological samples are collected by Fine Needle Aspiration Cytology (FNAC). Pathologists collect sample tissue from the breast using a narrow-gauge needle. The fixed and stained sample is then examined under the microscope and diagnosed by a cytopathologist. For CAD, YOLO detects and classifies the lesions in ROI.

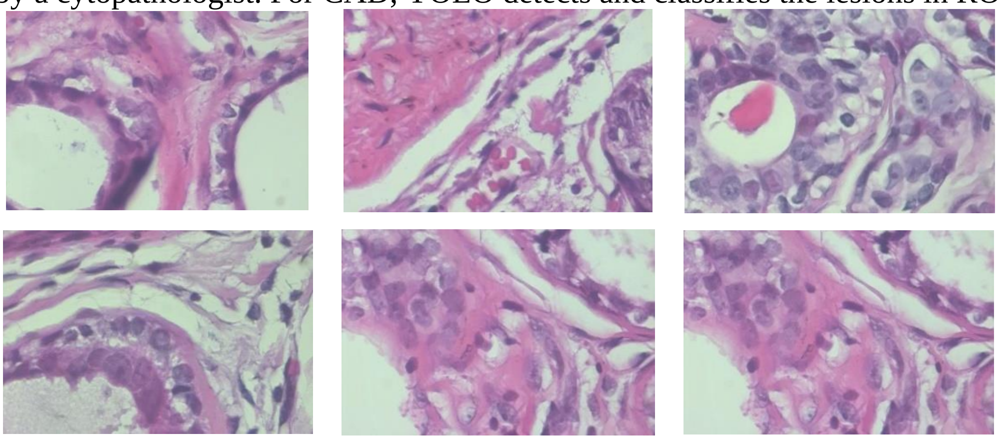


Figure 3. Histopathological Image [3]

2.3. Mammograms

Mammography is a common method of diagnosing breast cancer and has been widely used in breast screening and detecting masses at an early stage (Fig 4 from [4]). It involves a low-energy X-rays inspection of the breast to detect abnormal tumors. As the breast cancer rate increases and mammography is advocated in breast cancer screening tests, the work of breast cancer diagnosis becomes overwhelming for radiologists. YOLO is applied to detect tumors in mammograms and is also utilized to classify benign and malignant masses.

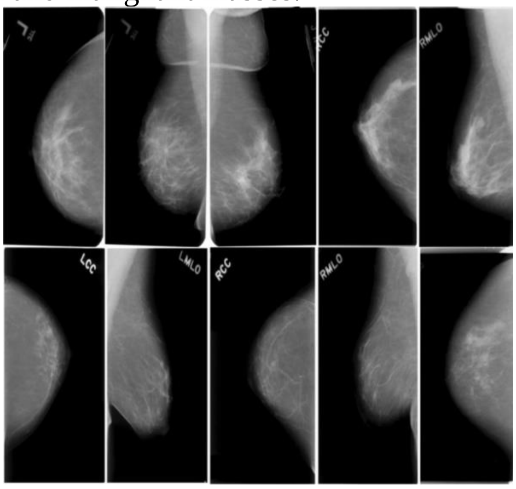


Figure 4. Mammogram

2.4. CT and LDCT

CT scans are the most common data source for cancer inspection (Fig 5 from [5]). People without any symptoms will include CT in the routine of physical examination. Meanwhile, CT is at low cost and available in most medical institution, making CT the first tool for cancer diagnosis. Chest CT scans are widely used in clinical examination to detect lung nodules and evaluate the risk of having lung cancer. YOLO algorithm serves as an object detection method and classifier for multiple lung lesions.

Like standard CT scans, LDCT is also a CT scan but applies lower-level X-ray when inspecting patients. LDCT is also widely used in screening detection of lung nodes and evaluating the malignancy of lesions. This method is more preferred by patients than standard CT because of less radiation exposure, which is beneficial to patients' health. YOLO serves as object detection and classifier as it does in regular CT. In lung lesion detection, LDCT has proven to be comparable to standard dose CTs [6].

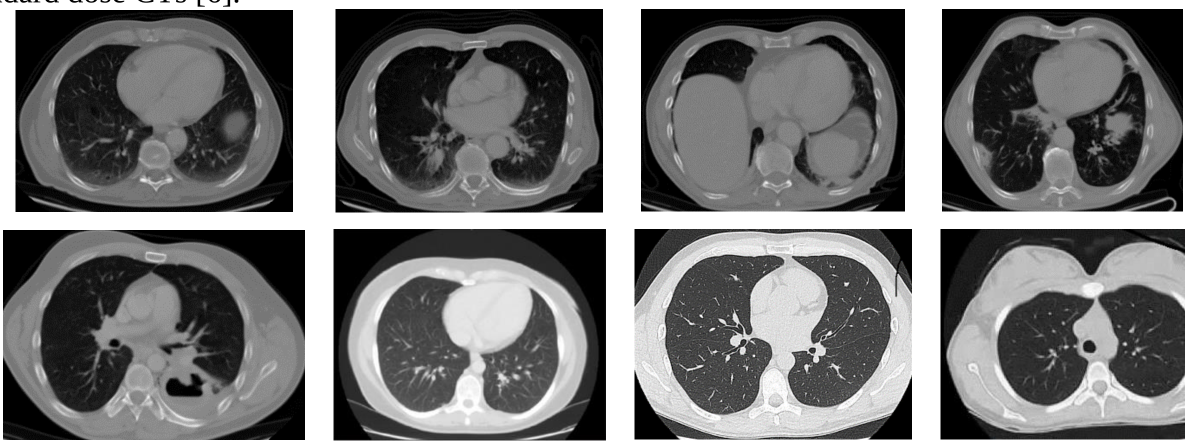


Figure 5. Computed Tomography (CT)

2.5. Challenges in Applying CAD

The application of CAD in chest cancer diagnosis improves efficiency and accuracy. However, the application of handcrafted features makes classification and extraction challenging tasks. In mammography, CAD suffered due to the similarity of breast tissues using only handcrafted features.

Other issues come from lung nodules detection. In CT lung nodules detection, it is hard to detect lung wall nodules. In clinical practice, nodule detection suffers a large number of false positives.

3. Analysis

For an automated diagnosis system, extraction, detection, and classification are three necessary stages for diagnosis, where classification and detection are usually integrated with YOLO architecture for YOLO-based CAD.

In extraction, handcrafted features limited the extraction of ROI. For instance, the common ROI of histopathological images is mitotic cells. Conventionally, in data preparation for training the model, ROIs are manually extracted based on handcrafted features by trained pathologists [7]. This work is tiresome and time-consuming. As a result, the size of the training data set is limited. What's more, in image data like histopathological images, the high variation in appearance and texture leads to a high risk of the wrong diagnosis. In detection and classification, a set of handcrafted or semi-automated features is needed to describe the difference between different classes of lesions or to classify the malignant and benign lesions. Handcrafted features are not applicable enough for classifying similar tissues in images, especially in mammography [8]. Thus, CAD suffered low accuracy.

In actual practice, CAD has to guarantee all true positives can be detected, but this also leads to increases in false positives. Though CAD only provides a second opinion for physicians, it can cause malpractice by misleading physicians into making the wrong diagnosis, which has been the most of medical malpractice according to a survey [9].

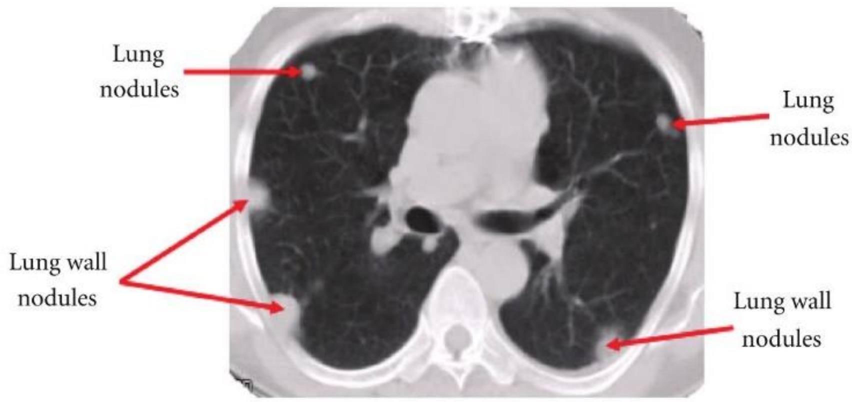
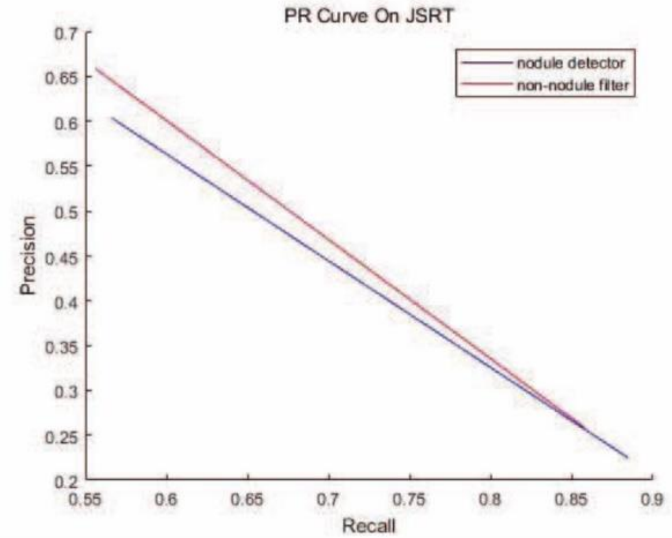


Figure 6. Lung Nodules and Lung Wall Nodules in CT

In lung nodule detection in spiral chest CTs, CNN uses convolutional blocks that assume lung nodules have circular, cylindrical, or spherical shapes, while most of the actual nodules have an irregular appearance. For nodules on the lung wall, the irregular shape can be caused by the attaching force between the nodule and pleural surface (Fig 6 from [10]). Hence, the CNN model is not good at detecting lung wall nodules, which reduces the accuracy of diagnosis.

4. Suggestion

The direct result of using handcrafted images is low accuracy in trained CNN models. Though the YOLO algorithm can detect and classify the lesion at the same time, applying an extra architecture as an extractor is more efficient in improving accuracy. ResNet50 and Inception can serve as an extractor for ROI extraction, and research comparing two architectures in extraction [7] has achieved an accuracy of 98.8% and 98.9%, respectively. Particularly, ROIs of histopathological images are usually mitotic cells. Nair [12] implemented YOLO v4 in mitotic nuclei detection and achieved an F1 score of 0.73 without nuclei segmentation. Zorgani [13] combined the YOLO v2 detector and ResNet50 extractor in mitotic cell detection and achieved an F1 score of 0.839.



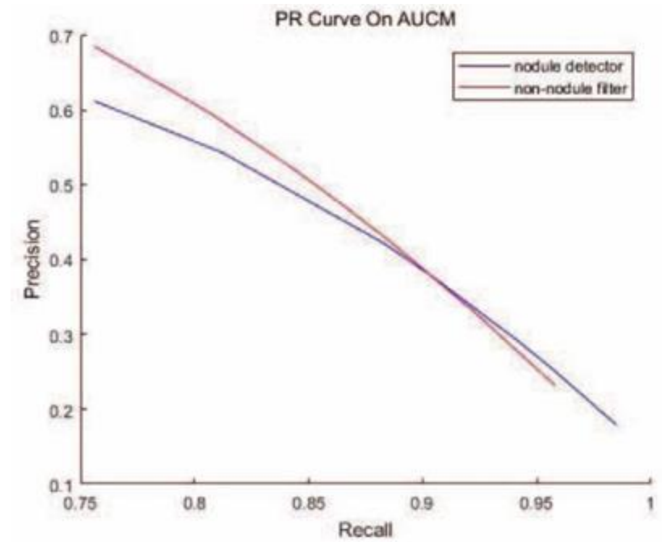


Figure 7. P-R Curve for Different Dataset Before and After Non-nodule Filter

For a later version of YOLO, a study [14] compare the performance of YOLO v3 and YOLO v4 and shows that YOLO v4 has enhanced the detection result to 8%. Though YOLO proves its ability to detect and classify objects, its low accuracy in classification is intolerable in medical diagnosis. In the meantime, architectures like Inception and ResNet50 show better performance in classification. Hence, models combine YOLO as a detector and other classifiers for better results.

The high volume of false positive predictions due to the guarantee of detecting all true positives can be addressed by integrating a YOLO architecture with a reduced IOU and non-nodule filter. A network applies to the output of the detection network. The network can be trained as a non-nodule filter for nodule detector [11] and has presented an effective enhancement in suppressing false negatives (Fig 7 from [11]). Multi-view CNN has proved its capability to reduce false positives [15] in the comparison of competition performance metric (CPM) [16] score among different models and provides more possibilities for the YOLO application in other fields [17-20].

The difficulty in detecting lung wall nodules is caused by the irregular shape of nodules, but nodules can be classified into four deformable templates to isolate abnormalities that have common morphological features. The genetic algorithm (GA) template matching method can be applied to model the shape of nodules into four categories. A study utilizing GA template matching presents overall accuracy of 82.3% after applying false positive removing techniques [10].

The suggested solutions of issues discussed are summarized in Table 1.

Table 1. Summary of Issues and Solutions

Issues	Solution
Using handcrafted features	Applying ResNet50 or Inception as an extractor for ROI extraction, YOLO as detector and Inception as a classifier
Too many false positive	Integrate Reduced IOU and Non-Nodule Filter
Difficulty in detecting irregular nodules	Model nodules into categories with similar morphological features and apply template matching algorithm.

5. Conclusion

The study presents the common image data categories for chest cancer diagnosis, including histopathological images, mammography, and CTs, and introduces the application and challenges of YOLO-based CAD in automatic diagnosis. Three issues are analyzed and discussed: low accuracy due to the usage of handcrafted features, overwhelming false positives, and difficulty in detecting irregular lung nodules in spiral chest CTs. Suggested solutions are presented for the above issues. The low accuracy in ROI extraction due to the handcrafted features is addressed by applying extra networks like Inception as an extractor or implementing later versions of YOLO with bounding boxes. False Positive can be suppressed by reducing IOU in the model and applying extra architecture as non-nodule detection. The difficulty in the detection of irregular-shape nodules is the limitation of CNN kernels. Classifying nodules into categories with similar morphological features can construct the template and apply a template matching algorithm, thus improving the sensitivity of detection.

The major limitation in optimizing YOLO-based CAD is the lack of available and applicable data. Object detection and classification tasks in daily life are not so difficult as in the medical field. In cancer diagnosis, image data are manually labeled by trained physicians and radiologists. Thus very limited labor can contribute to the dataset. Meanwhile, the highly variant appearance of image data and the subtle difference of handcrafted lesion features leads to the low efficiency and accuracy in data classification. Privacy protection also constrains the collection and usage of data. The study presents the analysis and solution to these problems under constraints of limited data and shows how well these methods addressed the drawbacks.

Due to the diversity in application, the paper analyzes three cases in the application of YOLO-based CAD, and most cases are not covered. Meanwhile, the research on each issue presented in the paper is not abundant. More optimizations can be researched for these issues, and more cases can be analyzed in future works.

References

[1] Salman, M. E., Çakar, G. Ç., Azimjonov, J., Kösem, M., and Cedimoğlu, İ. H. 2022 Automated prostate cancer grading and diagnosis system using deep learning-based Yolo object detection algorithm Expert Systems with Applications 201 117148

[2] Kaushal, M. 2022 Rapid-YOLO: A novel YOLO based architecture for shadow detection Optik 260 169084

[3] Parvin, F., and Hasan, M. A. M. 2020 A comparative study of different types of convolutional neural networks for breast cancer histopathological image classification 2020 IEEE Region 10 Symposium (TENSYMP) (IEEE) pp 945-948

[4] Yadav, R., Pande, S., and Khamparia, A. 2020 Breast Cancer Classification Using Convolution Neural Network (CNN) International Conference on Advanced Informatics for Computing Research (Singapore: Springer) pp 283-292

[5] Bhandary, A., Prabhu, G. A., Rajinikanth, V., Thanaraj, K. P., Satapathy, S. C., Robbins, D. E., ... and Raja, N. S. M. 2020 Deep-learning framework to detect lung abnormality–A study with chest X-Ray and lung CT scan images Pattern Recognition Letters 129 pp 271-278

[6] Kubo, T., Ohno, Y., Takenaka, D., Nishino, M., Gautam, S., Sugimura, K., ... and iLEAD study group 2016 Standard-dose vs. low-dose CT protocols in the evaluation of localized lung lesions: Capability for lesion characterization—iLEAD study European journal of radiology open 3 pp 67-73

[7] Bal, A., Das, M., Satapathy, S. M., Jena, M., and Das, S. K. 2021 Automated Diagnosis of Breast Cancer with RoI Detection Using YOLO and Heuristics International Conference on Distributed Computing and Internet Technology, ed Goswami, D. and Hoang, T.A. (Cham: Springer) pp 253-267

[8] Al-Antari, M. A., Al-Masni, M. A., and Kim, T. S. 2020 Deep learning computer-aided diagnosis for breast lesion in digital mammogram Deep Learning in Medical Image Analysis, ed Lee, G. and Fujita, H. (Cham: Springer) pp 59-72

[9] Whang, J. S., Baker, S. R., Patel, R., Luk, L., and Castro III, A. 2013 The causes of medical malpractice suits against radiologists in the United States Radiology 266(2) pp 548-554

[10] El-Baz, A., Elnakib, A., El-Ghar, A., Gimel'farb, G., Falk, R., and Farag, A. 2013 Automatic detection of 2D and 3D lung nodules in chest spiral CT scans International journal of biomedical imaging 2013 517632

[11] Liu, C., Wang, B., Jiao, Q., and Zhu, M. 2019 Reducing false positives for lung nodule detection in chest X-rays using cascading CNN 2019 14th IEEE Conference on Industrial Electronics and Applications (ICIEA) (IEEE) pp 1204-1207

[12] Nair, L. S., Prabhu, R., Sugathan, G., Gireesh, K. V., and Nair, A. S. 2021 Mitotic Nuclei Detection in Breast Histopathology Images using YOLOv4 2021 12th International Conference on Computing Communication and Networking Technologies (ICCCNT) (IEEE) pp 1-5

[13] Zorgani, A., Mohamed, M., Mehmood, I., and Ugail, H. 2021 Deep yolo-based detection of breast cancer mitotic-cells in histopathological images International Conference on Medical Imaging and Computer-Aided Diagnosis, ed Su, R., Zhang, YD., and Liu, H. (Singapore: Springer) pp 335-342

[14] Aly, G. H., Marey, M. A. E. R., El-Sayed Amin, S., Tolba, M. F. 2021 YOLO V3 and YOLO V4 for masses detection in mammograms with resnet and inception for masses classification International Conference on Advanced Machine Learning Technologies and Applications, ed Hassanien, AE., Chang, KC., and Mincong, T. (Cham: Springer) pp 145-153

[15] Setio, A. A. A., Ciompi, F., Litjens, G., Gerke, P., Jacobs, C., Van Riel, S. J., ... and Van Ginneken, B. 2016 Pulmonary nodule detection in CT images: false positive reduction using multi-view convolutional networks IEEE transactions on medical imaging 35(5) pp 1160-1169

[16] Niemeijer, M., Loog, M., Abramoff, M. D., Viergever, M. A., Prokop, M., and van Ginneken, B. 2010 On combining computer-aided detection systems IEEE Transactions on Medical Imaging 30(2) pp 215-223

[17] Beini, Z., Xuee, C., Bo, L., and Weijia, W. 2021 A new few-shot learning method of digital PCR image detection IEEE Access 9 pp 74446-74453

[18] Li, J., Gu, J., Huang, Z., and Wen, J. 2019 Application research of improved YOLO V3 algorithm in PCB electronic component detection Applied Sciences 9(18) 3750

[19] Zhang, B., Zhou, Z., Cao, W., Qi, X., Xu, C., and Wen, W. 2022 A New Few-Shot Learning Method of Bacterial Colony Counting Based on the Edge Computing Device Biology 11(2) 156

[20] Park, S. S., Tran, V. T., and Lee, D. E. 2021 Application of various yolo models for computer vision-based real-time pothole detection Applied Sciences 11(23) 11229