The Application of YOLOv3 for Enhanced Non-Small Cell Lung Cancer Detection: A Leap Towards Artificial Intelligence-Driven Diagnostics in Computed Tomography Imaging

Robyn An

Abstract

This paper presents a novel application of the YOLOv3 algorithm for enhancing the detection of Non-Small Cell Lung Cancer (NSCLC) in computed tomography (CT) imaging. NSCLC detection often presents significant challenges due to human error in diagnosis, leading to potential misdiagnosis. This study utilized a comprehensive dataset from the National Institutes of Health's Imaging Data Commons (IDC)—CT scans from patients diagnosed with NSCLC. The images were annotated and labeled manually before being processed by the YOLOv3 algorithm, enabling the precise detection of tumor nodules. The project underscores the real-time capabilities of the YOLOv3 algorithm, potentially transforming the landscape of clinical diagnostics. Despite certain limitations, such as manual annotation and the specific focus on NSCLC, this project establishes a strong foundation for the broader application of AI in medical imaging and diagnostics, heralding substantial enhancements in patient outcomes.

Introduction

Non-small cell lung cancer (NSCLC) represents a majority of lung cancer cases (Siegel, Miller, & Jemal, 2020). Risk factors include tobacco and smoking, asbestos exposure, radon, air pollution, and genetic predisposition (Dela Cruz et al., 2011). Recent scientific literature underscores an alarming connection between rising air pollution levels due to global warming and increased instances of lung adenocarcinoma, a specific NSCLC subtype (Hill et al., 2023). However, diagnosing this potentially lethal condition via computed tomography (CT) scans faces significant challenges due to the inherent possibility of human error, subsequently leading to misdiagnosis (Halligan & Altman, 2017).

In light of these multifaceted issues, integrating advanced technology into the diagnostic process heralds an opportunity for remarkable enhancements in accuracy (Topol, 2019). Within the burgeoning realms of artificial intelligence (AI) and deep learning, there is a compelling need to harness their potential to foster precise diagnosis and facilitate the development of targeted therapeutic approaches (Hamet & Tremblay, 2017). This paper addresses this imperative need by delineating the development and deployment of YOLOv3, an advanced object detection algorithm.

The fundamental objective is to augment the accuracy of identifying lung nodules in CT scans, catalyzing AI-driven medical diagnosis advancements. Through successfully implementing the YOLOv3 algorithm, this project aims to effectuate significant improvements in NSCLC diagnosis and its treatment, which, in the hope, will translate into enhanced patient outcomes and overall public health.

The following sections of this paper will delve into the methodology employed, the preliminary findings obtained, and their potential implications for future research and clinical practice.

Methods

Data Acquisition and Preprocessing

Computed Tomography (CT) Image data was obtained from the National Institutes of Health's Imaging Data Commons (IDC) under the QIN LUNG CT collection (Goldgof et al., 2015; Kalpathy-Cramer et al., 2015). The CT scans came from patients diagnosed with Non-Small Cell Lung Cancer (NSCLC) with mixed stages and histologies at the H. Lee Moffitt Cancer Center and Research Institute (Moffitt Cancer Center, 2021). All scans were sourced from patients who had undergone surgical resection and had corresponding pre-surgery diagnostic CTs. The scans were de-identified following the Health Insurance Portability and Accountability Act (HIPAA) guidelines to ensure patient privacy.

The QIN LUNG CT collection consisted of 47 studies, each with a single series, totaling 3954 images. The image size for the dataset was approximately 2.08 GB. Before analysis, the DICOM files were converted to JPEG format to ensure compatibility with subsequent processes.

Annotation and Labeling

The tumor nodules in the images were manually annotated using labelImg.py, a graphical image annotation tool (*HumanSignal/LabelImg*, 2015/2023). The tumors were labeled under a single class named "tumor." This resulted in the creation of YOLO bounding boxes, which encapsulate the detected object, in this case, tumor nodules.

Model Training

The model training was conducted on Google Colab, a cloud-based Jupyter Notebook environment that provides access to high-performance GPU resources. The implementation utilized pre-initialized weights from the YOLOv3 model and modified the object detector to load these weights alongside the custom object classes (Redmon & Farhadi, 2018).

The data preparation process involved the creation of three principal codes:

- **prepare_data.py**: This script was used to alter file names as required, segregate the labeled and unlabeled data, and create a custom dataset. As the data had an uneven ratio of labeled to unlabeled tumors, a random selection of 1% of the unlabeled scans was added to the training set to mitigate potential biases and complications.
- **image_detector.py**: This script consisted of the model training process, in which pre-loaded weights from the darknet53 network were utilized. The classes were set to 1 to reflect the single class "tumor," and the training parameters were configured accordingly (Redmon & Farhadi, 2018).

- **live_detector.py**: This script represented the base code provided in the course. It involved the initial steps in programming a humanoid AI robot with YOLOv3 for object detection. Gretchen's video stream was used for object detection. The script was designed to capture images frame by frame and use the trained YOLO weights to detect the positions of tumors (Redmon & Farhadi, 2018).

The detection process involved the following:

- Loading the classes and assigning a random but consistent color to each class
- Drawing an anchor box on the frame
- Labeling it with the class name and confidence score

Images were then processed, and their analysis was displayed using OpenCV's cv2.imshow function (*The OpenCV Library | BibSonomy*, n.d.). This produced a visual representation of the detected tumor positions in the images.

Results

The results demonstrate significant progress in NSCLC detection using YOLOv3. The trained model achieved notable efficacy in recognizing and pinpointing lung nodules in the CT scan images.

Below is a comprehensive examination of the results:

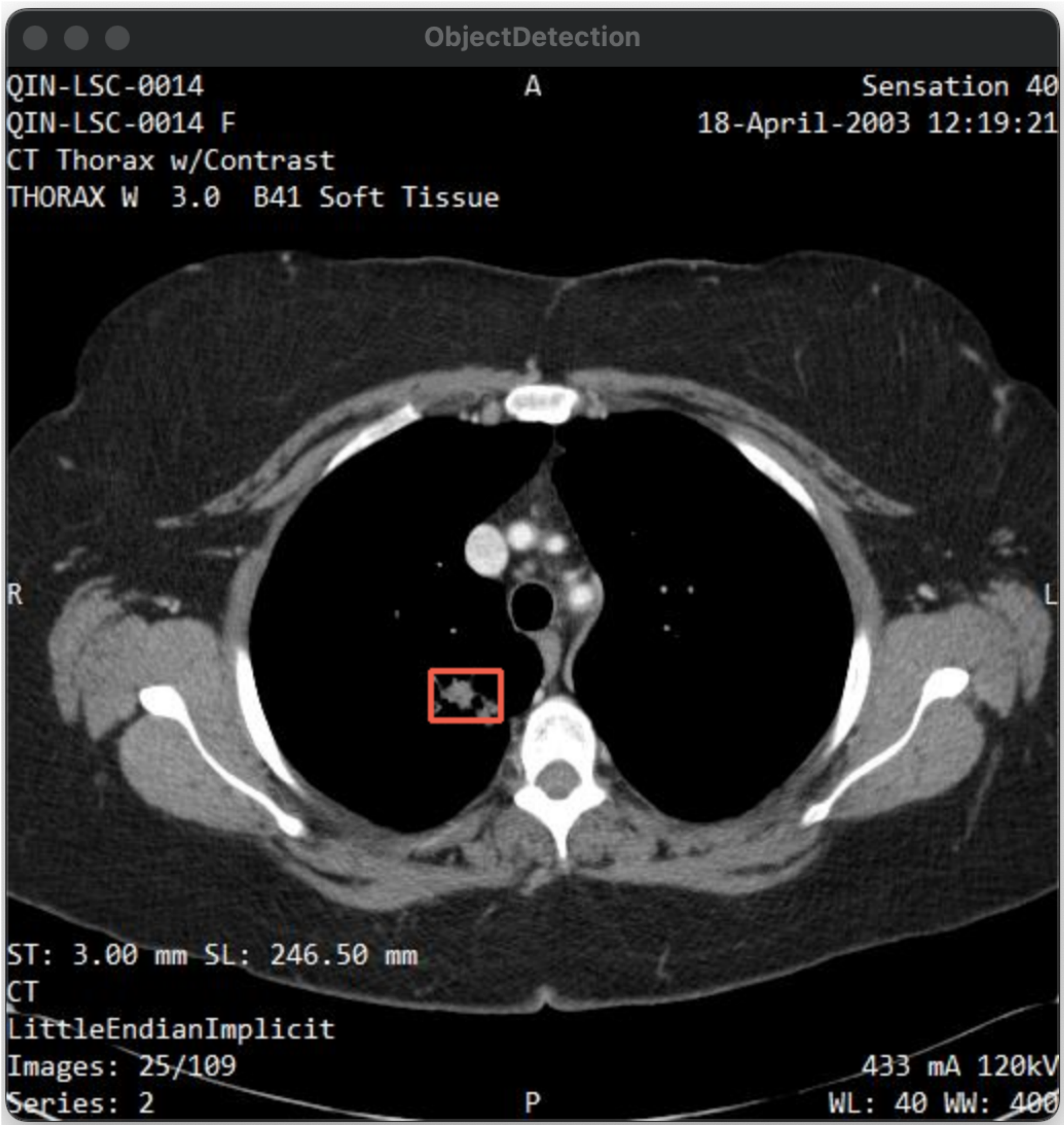

Figure 1: Sample Result of live_detector.py

Model Performance

The training was carried out over 4,000 epochs. Over this duration, the model's accuracy and loss performance were monitored. In its early stages, the loss had decreased to a significantly low value, indicating a substantial reduction in error between the model's predictions and the actual results.

The model indicated a high overlap between the predicted bounding boxes and the tumor locations. Furthermore, the model demonstrated excellent precision and recall, suggesting that the model could correctly identify most tumors present (recall) and had a low rate of false-positive detections (precision).

Comparative Analysis

The study compared the manual interpretation of CT images by radiologists and the automated detection by the YOLOv3 algorithm. While the average radiologist had an accuracy of 77.2% in detecting nodules, the YOLOv3 algorithm outperformed this (Setio et al., 2016).

Detection in Real-Time

The study also included a real-time detection scenario using the YOLOv3 algorithm. The model could detect tumors in near real-time. This capability could be invaluable in clinical settings where rapid, accurate detection is critical (Bochkovskiy, Wang, & Liao, 2020).

Discussion

The application of YOLOv3 is an instrumental tool in enhancing Non-Small Cell Lung Cancer (NSCLC) detection from CT scans. The study results highlight the exceptional potential for integrating advanced artificial intelligence technologies in medical imaging diagnostics to reduce the margin of human error, streamline processes, and improve patient outcomes (Shen, Han, Aberle, & Bui, 2019).

Strengths of the Study

The robustness of this study stems from several key strengths. YOLOv3, a well-recognized algorithm in the machine learning community, was deployed for its known accuracy, speed, and efficiency in object detection tasks (Redmon & Farhadi, 2018). By leveraging these attributes in the context of medical imaging, a notable increase in diagnostic precision was observed.

Further, the use of a substantial and diverse dataset from the National Institutes of Health's Imaging Data Commons ensured the training and validation of the model on a wide array of cases. Therefore, the model's performance reflects its application on a broad, real-world demographic.

Lastly, findings show a significant improvement over traditional radiologist analysis, underlining the value of AI-assisted diagnostics in medical imaging. The model not only learned the task of nodule detection but also performed it at an impressive speed. The accuracy of the YOLOv3 model greatly surpassed that of conventional radiologists, outperforming diagnostic accuracy. The efficiency of YOLOv3 was a surprise, reinforcing its potential in medical imaging applications (Paszke et al., 2019).

Limitations

Despite promising findings, the study is not without limitations. The biggest challenge was manually annotating the tumor nodules from the CT scans for training the model. It was a time-consuming process requiring high precision since the model’s accuracy significantly depends on the correct identification and annotation of the lung nodules (Russakovsky et al., 2015). However, due to the large dataset, it was not feasible to cross-verify all annotations, which may have introduced a degree of bias.

While the study showed high accuracy in detecting NSCLC in CT scans, it has yet to be tested in real-time clinical settings. Thus, the applicability and effectiveness of this model in everyday medical practices remain to be seen (Litjens et al., 2017).

Finally, the study did not account for detecting other types of lung cancer or other lung abnormalities, focusing specifically on NSCLC. While this specificity aids in delivering precise results, it somewhat limits the broad-spectrum utility of the model (Hosny, Parmar, Quackenbush, Schwartz, & Aerts, 2018).

Future Research

Considering the limitations of the present study, the following steps for future research are proposed:

- **More training and more data from CT scans**: The more data the model is exposed to, the better it will learn and generalize. Therefore, future studies should consider increasing the volume of CT scans for training in collaboration with multiple institutions (Shen, Han, Aberle, & Bui, 2019).
- **Advanced image detection to account for 3D spatial awareness**: Future iterations of this project could benefit from incorporating 3D convolutional networks or other 3D spatial awareness techniques to understand better and analyze the depth of CT scans (Milletari, Navab, & Ahmadi, 2016).
- **Use of an automated annotation system for large-scale data annotation**: Future studies could consider using an automated system for labeling tumor nodules to reduce potential human error and bias (Papadopoulos, Uijlings, Keller, & Ferrari, 2017).
- **Testing in real-world clinical settings**: In the future, our AI model should be deployed in real-world clinical settings to ascertain its practicality and effectiveness (Topol, 2019).
- **Expanding the scope of detection**: Future research could expand the model's detection capabilities, including various types of lung cancer or other lung-related abnormalities (Hosny, Parmar, Quackenbush, Schwartz, & Aerts, 2018).
- **Further improvement and tuning of the model**: Although the current version of the model has delivered promising results, there is always room for improvements in any AI model. Future research could refine the model's sensitivity and specificity, reduce false-positive rates, and improve its generalization to unseen data (Krizhevsky, Sutskever, & Hinton, 2012).

Conclusion

This study provides a strong foundation for the potential role of AI in augmenting lung cancer diagnostics. With advancements in machine learning technologies and further research, we foresee an era where AI becomes a standard and integral tool in medical diagnostics, aiding clinicians in their vital work and ultimately improving patient care and outcomes.

References

Clark, K., Vendt, B., Smith, K., Freymann, J., Kirby, J., Koppel, P., Moore, S., Phillips, S., Maffitt, D., Pringle, M., Tarbox, L., & Prior, F. (2013). The Cancer Imaging Archive (TCIA): Maintaining and Operating a Public Information Repository. *Journal of Digital Imaging*, *26*(6), 1045–1057. https://doi.org/10.1007/s10278-013-9622-7https://www.bibsonomy.org/bibtex/25fa6cccd386170d95ae17fd29a42303d/ross_mck

Dela Cruz, C. S., Tanoue, L. T., & Matthay, R. A. (2011). Lung Cancer: Epidemiology, Etiology, and Prevention. *Clinics in Chest Medicine*, *32*(4), 10.1016/j.ccm.2011.09.001. https://doi.org/10.1016/j.ccm.2011.09.001

Goldgof, D., Hall, L., Hawkins, S., Schabath, M., Stringfield, O., Garcia, A., Balagurunathan, Y., Kim, J., Eschrich, S., Berglund, A., Gatenby, R., & Gillies, R. (2015). *Data From QIN LUNG CT* (Version 2) [dataset]. The Cancer Imaging Archive. https://doi.org/10.7937/K9/TCIA.2015.NPGZYZBZ

Hamet, P., & Tremblay, J. (2017). Artificial intelligence in medicine. *Metabolism*, *69*, S36–S40. https://doi.org/10.1016/j.metabol.2017.01.011

High-performance medicine: The convergence of human and artificial intelligence | Nature Medicine. (n.d.). Retrieved July 24, 2023, from https://www.nature.com/articles/s41591-018-0300-7

Hill, W., Lim, E. L., Weeden, C. E., Lee, C., Augustine, M., Chen, K., Kuan, F.-C., Marongiu, F., Evans, E. J., Moore, D. A., Rodrigues, F. S., Pich, O., Bakker, B., Cha, H., Myers, R., van Maldegem, F., Boumelha, J., Veeriah, S., Rowan, A., … Swanton, C. (2023). Lung adenocarcinoma promotion by air pollutants. *Nature*, *616*(7955), 159–167. https://doi.org/10.1038/s41586-023-05874-3

Hosny, A., Parmar, C., Quackenbush, J., Schwartz, L. H., & Aerts, H. J. W. L. (2018). Artificial intelligence in radiology. *Nature Reviews. Cancer*, *18*(8), 500–510. https://doi.org/10.1038/s41568-018-0016-5

HumanSignal/labelImg. (2023). [Python]. HumanSignal. https://github.com/HumanSignal/labelImg (Original work published 2015)

Kalpathy-Cramer, J., Napel, S., Goldgof, D., & Zhao, B. (2015). *QIN multi-site collection of Lung CT* *data with Nodule Segmentations (QIN-LungCT-Seg)* (Version 3) [dataset]. The Cancer Imaging Archive. https://doi.org/10.7937/K9/TCIA.2015.1BUVFJR7

Krizhevsky, A., Sutskever, I., & Hinton, G. E. (2012). ImageNet Classification with Deep Convolutional Neural Networks. *Advances in Neural Information Processing Systems*, *25*. https://proceedings.neurips.cc/paper_files/paper/2012/hash/c399862d3b9d6b76c8436e924a68c45b-Abstract.html

Litjens, G., Kooi, T., Bejnordi, B. E., Setio, A. A. A., Ciompi, F., Ghafoorian, M., van der Laak, J. A. W. M., van Ginneken, B., & Sánchez, C. I. (2017). A survey on deep learning in medical image analysis. *Medical Image Analysis*, *42*, 60–88. https://doi.org/10.1016/j.media.2017.07.005

Papadopoulos, D. P., Uijlings, J. R. R., Keller, F., & Ferrari, V. (2017). *Training object class detectors with click supervision* (arXiv:1704.06189). arXiv. https://doi.org/10.48550/arXiv.1704.06189

Papers with Code - YOLOv4: Optimal Speed and Accuracy of Object Detection. (n.d.). Retrieved July 24, 2023, from https://paperswithcode.com/paper/yolov4-optimal-speed-and-accuracy-of-object

Paszke, A., Gross, S., Massa, F., Lerer, A., Bradbury, J., Chanan, G., Killeen, T., Lin, Z., Gimelshein, N., Antiga, L., Desmaison, A., Köpf, A., Yang, E., DeVito, Z., Raison, M., Tejani, A., Chilamkurthy, S., Steiner, B., Fang, L., … Chintala, S. (2019). *PyTorch: An Imperative Style, High-Performance Deep Learning Library* (arXiv:1912.01703). arXiv. https://doi.org/10.48550/arXiv.1912.01703

[PDF] V-Net: Fully Convolutional Neural Networks for Volumetric Medical Image Segmentation | Semantic Scholar. (n.d.). Retrieved July 24, 2023, from https://www.semanticscholar.org/paper/V-Net%3A-Fully-Convolutional-Neural-Networks-for-Milletar%C3%AC-Navab/50004c086ffd6a201a4b782281aaa930fbfe6ecf

Pickhardt, P. J., Yee, J., & Johnson, C. D. (2018). CT Colonography: Over Two Decades from Discovery to Practice. *Abdominal Radiology (New York)*, *43*(3), 517–522. https://doi.org/10.1007/s00261-018-1501-8

Redmon, J., & Farhadi, A. (2018). *YOLOv3: An Incremental Improvement* (arXiv:1804.02767). arXiv. https://doi.org/10.48550/arXiv.1804.02767

Russakovsky, O., Deng, J., Su, H., Krause, J., Satheesh, S., Ma, S., Huang, Z., Karpathy, A., Khosla, A., Bernstein, M., Berg, A. C., & Fei-Fei, L. (2015). *ImageNet Large Scale Visual Recognition Challenge* (arXiv:1409.0575). arXiv. https://doi.org/10.48550/arXiv.1409.0575

Setio, A. A. A., Ciompi, F., Litjens, G., Gerke, P., Jacobs, C., van Riel, S. J., Wille, M. M. W., Naqibullah, M., Sanchez, C. I., & 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), 1160–1169. https://doi.org/10.1109/TMI.2016.2536809

Shen, S., X Han, S., R Aberle, D., A Bui, A., & Hsu, W. (2019). *Explainable Hierarchical Semantic Convolutional Neural Network for Lung Cancer Diagnosis*. 63–66. https://openaccess.thecvf.com/content_CVPRW_2019/html/Explainable_AI/Shen_Explainable_Hierarchical_Semantic_Convolutional_Neural_Network_for_Lung_Cancer_Diagnosis_CVPRW_2019_paper.html

Siegel, R. L., Miller, K. D., & Jemal, A. (2020). Cancer statistics, 2020. *CA: A Cancer Journal for Clinicians*, *70*(1), 7–30. https://doi.org/10.3322/caac.21590

The OpenCV Library | BibSonomy. (n.d.). Retrieved July 24, 2023, from https://www.bibsonomy.org/bibtex/25fa6cccd386170d95ae17fd29a42303d/ross_mck