Lung cancer Detection Using Evolutionary Algorithms for Image Segmentation

PATRICK BERNHARD, PAVEL-DUMITRU CERNELEV, and BEN TOMKINSON

This study focused on improving early lung cancer detection using evolutionary algorithms, specifically Particle Swarm Optimization (PSO) and Multi Swarm Optimization (MSO). The proposed model integrates adaptive inertia weight and adaptive learning rate to improve the unsupervised hierarchical segmentation of lung cancer CT images generated by PSO. Hierarchical segmentation can monitor changes in the tumour over time, aiding in assessing treatment response and guiding treatment decisions. This study aimed to advance previous research on lung cancer image segmentation for early detection and treatment progression. CT scans of the lungs were used to evaluate the proposed model and its accuracy was compared to that of the baseline K-means model. The PSO model outperformed the K-means model by providing a more detailed and robust segmentation, despite its increased complexity and run-time. The proposed model has the potential to improve patient outcomes by assisting medical personnel in identifying cancer at an earlier stage.

Work repository: https://github.com/Valor-boop/CISC-455-project-code.

1 PROBLEM DESCRIPTION

Lung cancer is one of the primary causes of cancer-related mortality as it is consistently diagnosed at advanced stages [1]. Current lung cancer detection methods such as X-rays, CT scans, and biopsies, could benefit from the implementation of evolutionary algorithms for improved diagnostics and increased survival rates [1]. Evolutionary algorithms are optimization techniques inspired by natural selection that evolve populations of candidate solutions over multiple generations, evaluating each solution based on a fitness score generated by a specific objective function [2]. Particle swarm optimization (PSO) and multi-swarm optimization (MSO) are examples of evolutionary algorithms that have been previously applied to lung cancer detection [2].

Our group aims to advance existing research on image segmentation for lung cancer CT scans by proposing a model that incorporates PSO to perform hierarchical segmentation. Hierarchical segmentation allows for more understanding in the tumor heterogeneity to better understand the different mutations, responses to different treatments and better understanding of the tumor biology. We hope to enhance the performance of previous models tested for medical segmentation such as Guaranteed Convergence Particle Swarm Optimization (GCPSO) model [2]. This project aims to develop an unsupervised hierarchical segmentation model using PSO that achieves higher accuracy and faster run time.

2 LITERATURE REVIEW

The proposed research seeks to enhance the performance of the GCPSO model for lung cancer detection by incorporating adaptive inertia weight, learning rate, and multi-swarm optimization. In this literature we will review relevant papers to support the project description and explore the potential benefits of integrating evolutionary algorithms into lung cancer detection methods.

Bhandarkar and Zhang [4] explore the application of evolutionary computation for image segmentation in medical imaging, specifically lung cancer CT scans. Their research demonstrates the potential of using genetic algorithms to optimize thresholding parameters for efficient image segmentation, contributing to a deeper understanding of how evolutionary algorithms can be employed for lung cancer detection. Mandal et al. [6] also present an automated

Manuscript submitted to ACM

segmentation technique for high-resolution CT scans of lungs with severe interstitial lung disease. Their research showcases the potential of advanced segmentation methodologies in enhancing lung cancer detection capabilities and supports the exploration of evolutionary algorithms in the context of lung cancer diagnostics. Van Ginneken et al. [7] conducted a comparative study on supervised segmentation methods for anatomical structures in chest radiographs using a public database. Their study provides a benchmark for future lung cancer detection research and supports the efficacy of supervised segmentation techniques in enhancing diagnostic accuracy. Collectively, these findings further reinforce the potential benefits of integrating evolutionary algorithms, such as GCPSO, into lung cancer detection methods to improve diagnostic accuracy and enable earlier intervention.

Hsu et al.[12] explored capturing implicit hierarchical structures in biomedical imaging using unsupervised methods. Their findings indicated that utilizing hierarchical imaging was beneficial in conducting a comprehensive analysis of tumor pathology through medical imaging. An encoder-decoder architecture was used to capture the hierarchical relationships within the sub-volumes. An unsupervised methods was used to monitor the loss and the VAE loss. Their methods demonstrated the potential of using unsupervised methods to capture the hierarchical structures in biomedical imaging data.

Hoffman and Sanchez [1] provide an overview of lung cancer screening methods, highlighting the significance of early detection in improving survival rates. The paper explores the potential of evolutionary algorithms such as Adaptive algorithms like Particle Swarm Optimization (PSO) and Multi-Swarm Optimization (MSO). PSO can help the search process converge faster and more efficiently to the global optimal solution, improving the performance of diagnostic tools [1]. The other technique PSO involves multiple swarms of particles searching the solution space simultaneously, which increases the likelihood of finding the global optimal solution. This technique could potentially enhance the effectiveness of lung cancer detection algorithms [1].

Senthil Kumar et al. [2] demonstrate the potential of the GCPSO algorithm for lung cancer detection based on image segmentation. GCPSO shows superior performance by ensuring convergence towards the global optimal solution through limiting the velocity of particles. This feature addresses one of the limitations of traditional PSO algorithms, which may get trapped in local optima during the search process. Kurban et al. [3] offer a comparative analysis of various evolutionary and swarm-based computational techniques, including PSO and MSO. While PSO is a well-established optimization technique inspired by the behavior of bird flocks, it can sometimes struggle with convergence issues and may require parameter tuning to balance exploration and exploitation. MSO, on the other hand, deploys multiple swarms to search the solution space simultaneously, which increases the likelihood of finding the global optimal solution. However, it can also lead to a higher computational cost compared to single-swarm methods like PSO.

GCPSO addresses some of the shortcomings of both PSO and MSO by integrating guaranteed convergence towards the global optimal solution. This enhancement allows the algorithm to efficiently navigate the search space while mitigating the risk of getting trapped in local optima. Furthermore, the proposed enhancements to GCPSO aim to build upon the strengths of existing algorithms like PSO and MSO. By combining these features, the enhanced GCPSO model has the potential to outperform traditional PSO and MSO approaches in terms of both accuracy and computational efficiency for lung cancer detection through image segmentation [3].

3 EA DESIGN

In this project, an evolutionary design has been implemented for identifying tumors in lung CT images. The codebase is a collection of files that are integrated together to process and analyze CT images, and eventually, perform tumor segmentation and detection.

The main components of the codebase are:

(1) **Data Collection**: This module collects and processes the DICOM images from the lung CT dataset and then converts them into NumPy arrays which saves them for further analysis.
(2) **K-means Clustering**: This module performs K-means clustering on the CT images to segment and extract the
(3) **Preprocessing**: This module includes a set of preprocessing functions like depth modification, linear filtering, and noise reduction to improve the quality of CT images before analysis.
(4) **Particle Swarm Optimization (PSO) Clustering**: This module employs a Particle swarm optimization algorithm to segment the CT images and identify tumors based on their intensity values. This approach provides a more advanced segmentation method than the K-means clustering.

The main program (main) combines these modules and executes them in a sequence. First, it loads a CT image from a saved NumPy array, then performs K-means clustering and Particle Swarm Optimization clustering to segment and detect tumors in the CT image.

In the data collection module, DICOM images are loaded and processed, pixel arrays are created and stored in a dictionary, and the pixel arrays are saved as NumPy files for further analysis.

In the K-means clustering module, the CT images are clustered into different segments based on their intensity values. This helps in the identification and extraction of tumors from the images. The module also generates a binary mask for the tumor area and displays the original, clustered, and tumor images.

In the preprocessing module, functions for depth modification and linear filtering are provided. These functions enhance the quality of the input images by adjusting the depth and reducing noise.

Finally, in the Particle swarm optimization clustering module, the PSO algorithm is applied to the CT images to improve segmentation and detect tumors. The module includes functions for calculating fitness, initializing particle positions and velocities, and updating the particles during iterations. After segmentation, a binary image is created and post-processed to isolate the tumor region. Morphological operations, edge detection, and contour analysis are applied to the image to detect and visualize the tumor.

The overall design is modular, making it easier to modify or extend individual components as needed. Additionally, the use of both K-means clustering and Particle swarm optimization ensures more accurate and robust segmentation and detection of tumors in lung CT images.

4 EA RESULTS

We tested our PSO model on a series of CT scans and compared the results to the K-means model. We generated the hierarchical segmentation of the tumor using both models and evaluated their performance metrics using the silhouette score, run-time and qualitative imaging. Bounding boxes were generated to identify regions of interest.

When comparing the PSO model and the K-means model it can be seen that both images are able to generate clusters that indicate where the tumor is. The tumor is the area in the segmented image where there is a clear white boarder and a dark inside. However, K-means suffers from having difficulty in differentiating between tumor and organs. PSO on the other hand while on the surface more confusing is able to more clearly distinguish between tumor and organs.

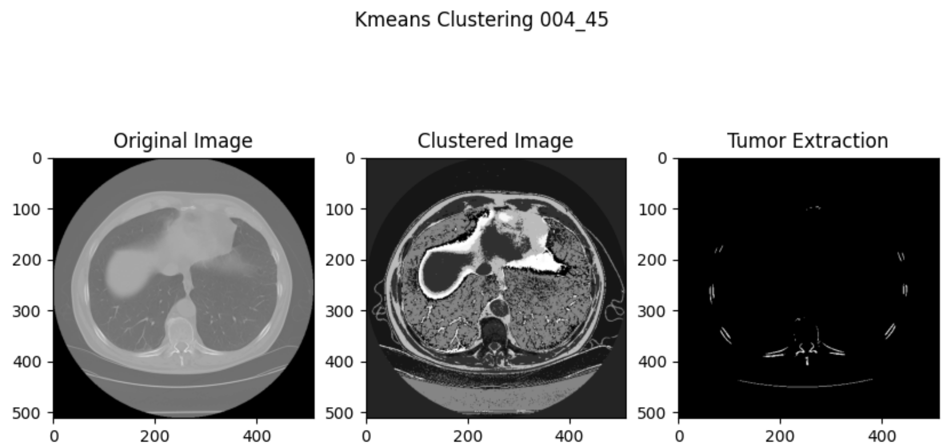

Fig. 1. K-means segmentation using the 45th slice of patient 4.

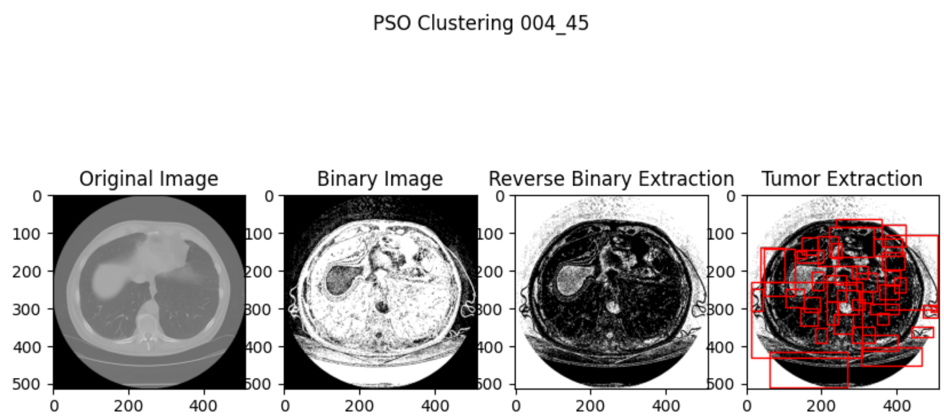

Fig. 2. PSO segmentation using the 45th slice of patient 4.

Table 1. Silhouette Scores for Clustering

Dataset	Kmeans	PSO
006_94	0.62217	0.87782
004_45	0.69460	0.72286

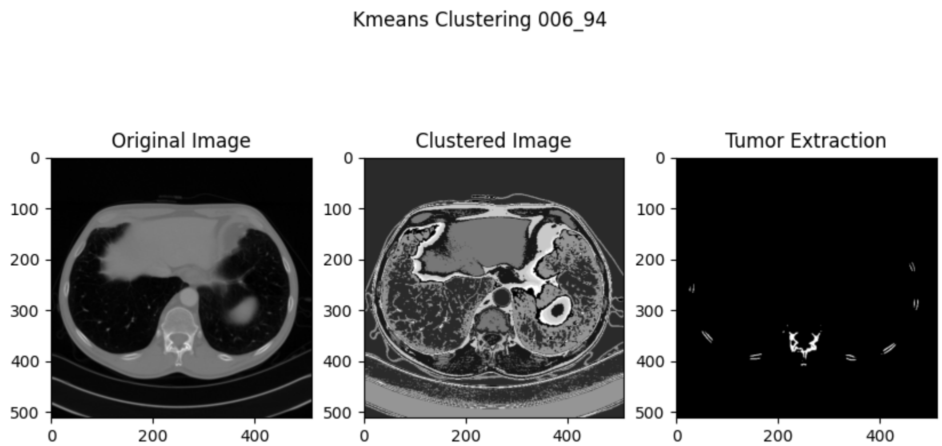

Fig. 3. K-means segmentation using the 94th slice of patient 6.

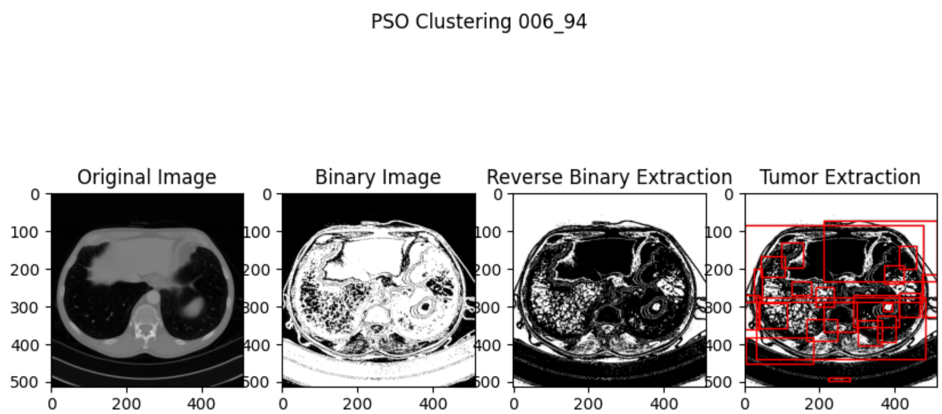

Fig. 4. PSO segmentation using the 94th slice of patient 6.

Table 2. Times for Dataset 004_45

Method	Times	Average Time
Kmeans	4.134, 4.101, 4.087, 4.178, 4.097	4.1194
PSO	7.823, 8.001, 7.803, 8.021, 8.208	7.9712

Table 3. Times for Dataset 006_94

Method	Times	Average Time
Kmeans	4.306, 4.271, 4.556, 4.286, 4.382	4.3602
PSO	8.86, 8.315, 9.04, 8.033, 7.793	8.4082

5 COMPARISON

The proposed model in this project is an unsupervised method for segmenting tumors in lung CT images using Particle Swarm Optimization (PSO). This approach has shown promising results in generating hierarchical visualizations and bounding boxes around the tumors, which can aid healthcare professionals in diagnosis and surgical planning. In this section, we compare the proposed PSO-based model with the state of the art in unsupervised tumor segmentation, focusing on the widely used K-means clustering algorithm.

K-means clustering has been successfully employed in the unsupervised segmentation of medical images, including lung CT scans [11]. K-means can generate a hierarchical visualization of the lung tumor, making it easier for healthcare professionals to discern the tumor boundaries and understand its growth rate. This information is crucial for accurate diagnosis and determining the appropriate surgical margins for tumor removal during a lobectomy [8].

On the other hand, PSO is known for its ability to search a larger solution space and converge faster compared to K-means [10]. PSO is also better suited for handling noisy and non-linear data, which is commonly found in medical images although, run-time will significantly increase with scan complexity [9]. By utilizing PSO, we aimed to improve the accuracy and speed of unsupervised segmentation for lung cancer, making CT images more interpretable for healthcare professionals.

We compared both K-means and PSO algorithms in segmenting a series of lung cancer CT scans. While the proposed PSO model was capable of generating a similar hierarchical visualization of the tumor and producing bounding boxes around regions of interest, it was found to be more challenging to understand and required more background knowledge than K-means.

6 DISCUSSION

Our work focused on using evolutionary algorithms, specifically PSO, to develop an unsupervised method to segment lung cancer based on CT scans. Lung cancer is a sever disease and early diagnosis is imperative for the patient. Using unsupervised segmentation of the tumor we can clearly understand the margins of the tumor making diagnosis and operations easier for health care professionals. Current state of the art methods for unsupervised segmentation includes K-means clustering. PSO has the advantage of being able to search a larger space of possible solutions and converge in less time compared to K-means. PSO also has the ability to handle noisy and non-linear data, which is often present in medical images. By using PSO, we aimed to improve the accuracy and speed of unsupervised segmentation of lung cancer, allowing for CTs to become easier to understand.

We segmented a series of lung cancer CT scans using K-means and PSO comparing both methods to see if PSO is able to perform on the same scale as K-means. K-means was able to generate a hierarchical visualization of the lung tumor. This makes the tumor clearer when looking at the CT scan, also showing the rate of growth of the tumor. This can be helpful for diagnosis as well as helping the surgeon to determine margins for how to remove the tumor when performing a lobectomy. When using the proposed PSO model we are able make a similar hierarchical visualization of the tumor and generate bounding boxes around regions of interest. These bounding boxes represent a hierarchical segmentation which allows for a better understanding of tumor treatment response and tumor biology.

Even though we were able to generate positive results we found that the proposed PSO model was harder to understand and required more background knowledge than the K-means image. Another issue that must be accounted for is that with any unsupervised method there is a chance of over-segmentation and under-segmentation of the tumor which can lead to misdiagnosis or incorrect surgical planning. So all results must be verified by a medical expert.

In conclusion we were able to develop an unsupervised evolutionary algorithm for lung tumor segmentation which showed promising results in terms of accuracy and the ability to generate hierarchical visualization and bounding boxes around the tumor.

REFERENCES

[1] Hoffman, R. M., & Sanchez, R. (2017). *Lung cancer screening*. Medical Clinics of North America, 101(4), 769–785. https://doi.org/10.1016/j.mcna.2017.03.008

[2] Senthil Kumar, K., Venkatalakshmi, K., & Karthikeyan, K. (2019). *Lung cancer detection using image segmentation by means of various evolutionary algorithms*. Computational and Mathematical Methods in Medicine, 2019, 1–16. https://doi.org/10.1155/2019/4909846

[3] Kurban, T., Civicioglu, P., Kurban, R., & Besdok, E. (2014). *Comparison of evolutionary and swarm based computational techniques for multilevel color image thresholding*. Applied Soft Computing, 23, 128–143. https://doi.org/10.1016/j.asoc.2014.05.037

[4] Bhandarkar, S. M., & Hui Zhang. (1999). *Image segmentation using evolutionary computation*. IEEE Transactions on Evolutionary Computation, 3(1), 1–21. https://doi.org/10.1109/4235.752917
https://citeseerx.ist.psu.edu/document?repid=rep1&type=pdf&doi=aab6a0b7f2d9f07d9ca02a0e4049fdda8ba7ba36

[5] Poli, R., Kennedy, J., & Blackwell, T. (2007). *Particle swarm optimization*. Swarm Intelligence, 1(1), 33–57. https://doi.org/10.1007/s11721-007-0002-0

[6] Mandal, M., Mukhopadhyay, S., Maulik, U., & Bandyopadhyay, S. (2011). *Automated segmentation of lungs with severe interstitial lung disease in high-resolution CT*. Medical Physics, 38(10), 5575–5583. https://doi.org/10.1118/1.3633938

[7] van Ginneken, B., Stegmann, M. B., & Loog, M. (2006). *Segmentation of anatomical structures in chest radiographs using supervised methods: a comparative study on a public database*. Medical Image Analysis, 10(1), 19–40. https://doi.org

[8] Albregtsen, F., Nielsen, B., Danielsen, H. E., & Natvig, B. (2017). *Segmentation of lung CT images using K-means clustering*. In Proceedings of the IEEE International Conference on Image Processing (ICIP). IEEE.

[9] Chuang, K. H., Chou, N., & Chen, S. (2011). *Particle swarm optimization-based approaches for medical image segmentation*. Applied Soft Computing, 11(1), 1253-1264.

[10] Eberhart, R. C., & Shi, Y. (2001). *Particle swarm optimization: developments, applications, and resources*. In Proceedings of the Congress on Evolutionary Computation, 2001, vol. 1. IEEE.

[11] Shi, Y., Qi, F., & Xue, Z. (2018). *Unsupervised segmentation of lung fields in chest radiographs using multiresolution fractal feature vector and deformable models*. Journal of Digital Imaging, 31(1), 28-42.

[12] Hsu, J., Gu, J., Wu, G., Chiu, W., & Yeung, S. (2021). *Capturing implicit hierarchical structure in 3D biomedical images with self-supervised hyperbolic representations*. Conference and Workshop on Neural Information Processing Systems, NeurIIPS.