



OPEN ACCESS

EDITED BY

Hermann Frieboes,
University of Louisville, United States

REVIEWED BY

Ģirts Brīģis,
Rīga Stradiņš University, Latvia
Haowei Wang,
Tongji University, China

*CORRESPONDENCE

Oluwasegun Julius Aroba
✉ aroseglinks@yahoo.co.uk
Mayowa O. Oyediran
✉ mo.oyediran@acu.edu.ng

RECEIVED 27 March 2024

ACCEPTED 11 September 2024

PUBLISHED 23 December 2024

CITATION

Oyediran MO, Ojo OS, Raji IA, Adeniyi AE and Aroba OJ (2024) An optimized support vector machine for lung cancer classification system. *Front. Oncol.* 14:1408199. doi: 10.3389/fonc.2024.1408199

COPYRIGHT

© 2024 Oyediran, Ojo, Raji, Adeniyi and Aroba. This is an open-access article distributed under the terms of the [Creative Commons Attribution License \(CC BY\)](#). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

An optimized support vector machine for lung cancer classification system

Mayowa O. Oyediran^{1*}, Olufemi S. Ojo², Ibrahim A. Raji³,
Abidemi Emmanuel Adeniyi⁴ and Oluwasegun Julius Aroba^{5,6*}

¹Department of Computer Engineering, Ajayi Crowther University, Oyo, Nigeria, ²Department of Computer Sciences, Ajayi Crowther University, Oyo, Nigeria, ³Department of Mathematical and Computer Sciences Education, Emmanuel Alayande University of Education, Oyo, Nigeria, ⁴Department of Computer Science, Bowen University, Iwo, Nigeria, ⁵Honorary Research Associate, Department of Operations and Quality Management, Durban University of Technology, Durban, South Africa, ⁶Centre for Ecological Intelligence, Faculty of Engineering and the Built Environment (FEBE), Electrical and Electronic Engineering Science, University of Johannesburg, Johannesburg, South Africa

Introduction: Lung cancer is one of the main causes of the rising death rate among the expanding population. For patients with lung cancer to have a higher chance of survival and fewer deaths, early categorization is essential. The goal of this research is to enhance machine learning to increase the precision and quality of lung cancer classification.

Methods: The dataset was obtained from an open-source database and was utilized for testing and training. The suggested system used a CT scan picture as its input image, and it underwent a variety of image processing operations, including segmentation, contrast enhancement, and feature extraction.

Results: The training process produces a chameleon swarm-based support vector machine that can identify between benign, malignant, and normal nodules.

Conclusion: The performance of the system is evaluated in terms of false-positive rate (FPR), sensitivity, specificity, recognition time and recognition accuracy.

KEYWORDS

chameleon swarm algorithm (CSA), lung cancer, support vector machine, optimization techniques, machine learning

1 Introduction

Due to its higher death rate than the combined deaths from breast, colon, and prostate cancer, lung cancer is one of the deadliest diseases in the world (1). The primary cause of carcinoma or lung cancer is cigarette smoking. Inhaling second-hand smoke, using pipes or cigars, being exposed to radon or asbestos at work or home, and having a family history of the disease are additional risk factors for lung cancer (1).

Malignant nodules in the lung that have the potential to spread quickly throughout the body after entering the bloodstream are referred to as lung cancer. Due to its rapid spread after creation, lung cancer poses a greater threat to life than other tumors (2). According to a previous study (3), exhaustion, chest discomfort, coughing, hemoptysis, sore throat, shortness of breath, weariness, weight loss, and chest infection are common signs of lung cancer.

Preventing the growth and spread of cancer cells is largely dependent on the early diagnosis process. While CT scan imaging can reveal both suspected and unsuspected lung cancer nodules, it is not the most reliable imaging technique for lung cancer diagnosis (4).

The clever behavior of chameleons in the wild served as the inspiration for the development of the chameleon swarm algorithm (CSA) in recent years (5). Many researchers have expressed interest in using it in various fields due to its simplicity and ease of implementation (6). It has been applied in image segmentation (7) and power engineering (8, 9). It has also been proven to be successful in resolving optimization issues in biology (10) and mathematics (11), respectively.

One machine learning approach used for classification is support vector machine (SVM). Assigning points to one of two disjoint half spaces is a common method of classifying points (12, 13). To classify data, SVMs project low-dimensional vectors into a high-dimensional space and create ideal hyperplanes that make it simple to classify fresh data points in the future (14). The linearly indistinguishable problem is converted into a linearly divisible problem in the high-dimensional space by selecting an appropriate kernel function (15). Hence, this study developed a model to select appropriate parameters for SVM using a chameleon swarm optimization technique.

A computer-aided diagnosis (CADx) system for classifying lung nodules was evaluated by (16). The main areas of focus were (i) the utility of the traditional CADx system (produced imaging feature + machine learning algorithm); (ii) a comparison of machine learning algorithms, namely, gradient tree boosting (XGBoost) vs. SVM; and (iii) the efficacy of parameter optimization using Bayesian optimization and random search. To calculate the feature vector, a local binary pattern (LBP) variation was employed. A feature vector and its matching label were used to train an SVM or XGBoost. SVM and XGBoost parameters were optimized using Bayesian techniques using the Tree Parzen Estimator (TPE). For TPE comparison, a random search was conducted. To optimize and assess the performance of our CADx system, we employed leave-one-out cross-validation. The receiver operating characteristic analysis' area under the curve (AUC) was used to assess performance. Compared to random search, Bayesian optimization of the SVM and XGBoost parameters was more effective. Two board-certified radiologists had AUC values of 0.898 and 0.822, according to an observer study. According to the findings, our CADx system's diagnostic precision in lung nodule classification was on par with that of radiologists.

The contribution of this study represents a novel approach to optimizing SVM parameters by using CSA to improve lung cancer classification performance through the CS-SVM model. Compared to conventional SVM, recognition accuracy, sensitivity as well as specificity, and recognition time are heightened with the introduction

of CS-SVM technique using CSA in this study. This demonstrates how well optimization works in practice.

Şekeröğlü and Emirzade (17) developed and implemented an SVM-based computer-aided diagnosis (CAD) system for the identification of lung cancer. Their goal in developing the CAD system was to improve diagnosis accuracy while cutting down on diagnosis time. The 271 recorded whole-lung CT scan images that make up the Lung Image Database Consortium (LIDC) subset of the public database were used in this investigation. To lessen impulse noise, a median filter was utilized during image preparation. The suggested CAD system used edge detection, morphological segmentation, and global thresholding in conjunction with the binarization procedure. Otsu's global image thresholding technique was utilized to convert the grayscale images into binary. Following the binarization procedure, the boundaries of the CT lung image were determined using the gradient detection approach. Then, extraneous perimeter lines were eliminated using the morphological operation. For every CT scan, statistical features were retrieved from the histogram and the gray-level co-occurrence matrix (GLCM) in four different orientations. Owing to the extracted features' high dimension, feature selection procedures were used to increase the classifier's accuracy. The SVM found the best-fit kernel for the classification by using its three most preferred kernels: linear, quadratic, and radial basis functions. In conclusion, the CAD system outperformed other suggested methods, achieving 97% accuracy, 92% sensitivity, and 97.3% specificity.

Sharma et al. (18) introduced a novel CAD system that uses image processing techniques to identify lung cancer nodules and classifies them using Convolutional Neural Network (CNN). The investigator utilized a CT-scan image as the input image and applied various image processing techniques, including segmentation, morphological operations, histogram equalization, and feature extraction. To identify if a nodule is malignant or not, the classifier was trained. The CAD system's overall performance was enhanced by this strategy, which yielded 98.08% accuracy, 96.6% specificity, and 90.39% sensitivity analysis.

To diagnose lung cancer more accurately, Singh and Gupta (19) developed a model in which they used neural networks and multilayer perceptron. The four criteria that they have employed are recall, accuracy, F1 score, and precision. A total of 6,910 benign and 8,840 malignant lung cancer photos were among the 15,750 clinical images that they used for testing and training. Finally, they developed an 88.55% accuracy multilayer perceptron, which proved to be a superior classifier.

Four classification techniques—SVM, Back Propagation Neural Network (BPNN), Probabilistic neural network (PNN), and k-means clustering—were compared by (20). High-boost filtering was utilized during the proposition stage. The Fuzzy C- Mean (FCM) clustering algorithm was then used for segmentation, and statistical techniques were employed for feature extraction. SVM, PNN, BPNN, and K-means clustering had accuracy ratings of 85%, 82%, 86%, and 81%, in that order. Local ternary co-occurrence patterns (LTCoP) and LBP feature extraction variations were carried out for 50 pictures from LIDC by Bruntha et al. (21). SVM was utilized for classification, whereas median filtering, intensity thresholding, and segmentation were employed in the preposition stage. It was 91.5% for LTCoP and 89.2% for LBP. In

addition, Karthiga and Rekha (22) conducted a comparison of the Fuzzy Multicategory Support Vector Machines Classifier (FMSVM), multilayer perceptron (MLP), K-Nearest Neighbor (KNN), Support Vector Machines (SVM), and I-Naive Bayesian Classifier (I-NBC) classifications; the corresponding accuracies were 98%, 52%, 46%, 96%, and 98%, respectively.

According to (23), machine learning algorithms including SVM, random forest, and artificial neural network (ANN) have been used to train the retrieved features. Then, to determine which technique has the best accuracy, a variety of factors like accuracy, precision, and recall were assessed, and, with the highest accuracy of 96%, the ANN was established as the best.

A method for identifying the condition using a classification technique called Content-Based Image Retrieval (CBIR)-based Bag of Features (BOF) and grouping pattern categories using K-means clustering was proposed by Bhatt and Soni (24). A testing accuracy of 98.56% and a training accuracy of 99% were achieved with this strategy. Yunianto et al. and Aroba et al. (25–29) examined a total of 120 CT scan image data, utilizing Otsu thresholding for segmentation and a median filter for placement. After that, data were retrieved using the GLCM feature, which had angle and direction fluctuations. Naïve Bayes was utilized for the classification procedure, and 88.33% accuracy was achieved.

2 Methodology

2.1 Image acquisition

The dataset acquired from open source database (www.kaggle.com) contained a total of 1,097 images. These images were grouped into three classes: normal, benign, and malignant, and the datasets for each of the classes were 416, 120, and 561, respectively. The CT scans were originally collected in DICOM format, and the whole dataset was divided into training and testing dataset using the K-fold validation method where $K = 10$.

2.2 Image pre-processing

This is a step that suppresses noise or other small fluctuations in an image and is also used to improve the interpretability or perception of information in an image to enhance better input when used in other image analysis techniques. In this study, contrast limited adaptive histogram equalization (CLAHE) was utilized to produce an enhanced version of the original image.

2.3 Segmentation

In this study, fuzzy C-means was used to perform segmentation by determining the cancer nodules in the lung (30). This phase helped to identify the regions of interest in the lung nodule for a better classification process.

2.4 Feature extraction

LBP is a simple yet influential algorithm used in this study for feature extraction. The algorithm works by analyzing the texture of an image, which is defined by the distribution of intensity variations in small regions of the image called local neighborhoods. The LBP algorithm works as follows:

- I. Select a pixel in the image and define a local neighborhood around it, typically a square or circular region of fixed size.
- II. Compare the intensity value of the central pixel with the intensity values of its neighbors. If a neighbor has a higher intensity value than the central pixel, then assign it a value of 1. Otherwise, assign it a value of 0.
- III. Concatenate the binary values of the neighbors into a binary string, creating a unique pattern for the local neighborhood.
- IV. Repeat this process for every pixel in the image, creating an LBP image where each pixel is replaced by its corresponding binary pattern.
- V. Calculate a histogram of the LBP patterns over the entire image. This histogram represents the texture of the image and can be used as a feature descriptor.

2.5 Proposed technique

This study used the strength chameleon swarm optimization algorithms to improve the performance of the SVM. The steps involved in achieving this optimization technique are as follows (Algorithm 1):

- I. Chameleon swarm initialization: Initialize a population of chameleon individuals with random parameter settings (for example, C and kernel parameters for SVM).
- II. Objective function evaluation: Evaluate the performance of each chameleon's parameter settings using cross-validation and the SVM objective function.
- III. Color-changing Behavior: Implement color-changing behavior to update chameleon parameters. For example, adjust the “ C ” value or kernel parameters based on the chameleon's performance.
- IV. Swarm movement: Allow chameleons to move within the parameter space, guided by the color-changing behavior and the optimization objectives.
- V. Fitness evaluation: Re-evaluate the fitness of chameleons based on their updated parameter settings.
- VI. Best chameleon selection: Select the best-performing chameleon as the optimal parameter setting for SVM.

INPUT: Input the SVM kernel functions

1. Set $Q_q \leftarrow 0.1$ (the position update probability)
2. r_1, r_2, r_3, r^i are the random numbers between 0 and 1
3. UB and LB are the upper and lower bounds of the search area
4. $N \leftarrow$ dimension of the problem

5. \tilde{y}_t^i is the center of the current position of SVM kernel function i at iteration t
6. yr_t^i is the rotating centered coordinates of SVM kernel function i at iteration t which can be defined using. $yr_t^i = m * yc_t^i$ where m is a rotation matrix that represents the rotation of SVM kernel function and yc_t^i represents the centering coordinates at iteration t
7. Randomly initialize the position of a swarm of n SVM kernel functions in the search space using equation $y^j = LB_j + r * (UB_j - LB_j)$
8. Initialize the velocity of dropping SVM kernel functions
9. Evaluate the position of the SVM kernel functions
10. while ($t < T$) do
11. Define the parameter $\lambda = \delta e^{(-\alpha t/T)^\beta}$ where t and T represent the current and maximum number of iterations respectively, δ , α , and β are three constant values used to control exploration and exploitation capabilities.
12. Define the inertia weight ω using equation $\omega = (1 - \frac{t}{T})^{(\rho\sqrt{t})}$.
13. Define the acceleration rate α using equation $\alpha = 2590 \times (1 - e^{-\log(t)})$
14. for $i = 1$ to M do
15. for $j = 1$ to N do
16. if $r^i \geq Qq$ then
17. $y_{t+1}^{ij} = y_t^{ij} + q_1 (Q_t^{ij} - G_t^j) \quad r_i + q_2 (G_t^j - y_t^{ij}) \quad r_2$
18. else
19. $y_{t+1}^{ij} = y_t^{ij} + \lambda ((UB^i - LB^i) \quad r_3 + LB_b^i) \text{sgn}(\text{rand} - 0.5)$
20. end if
21. end for
22. end for
23. for $i = 1$ to M do
24. $y_{t+1}^i = yr_t^i + \tilde{y}_t^i$
25. end for
26. for $i = 1$ to M do
27. for $j = 1$ to N do
28. $V_{t+1}^{ij} = \omega v_t^{ij} + c_1(G_t^j - y_t^{ij})r_1 + c_2(Q_t^{ij} - y_t^{ij})r_2$
29. $y_{t+1}^{ij} = y_t^{ij} + ((v_t^{ij})^2 - (v_{t-1}^{ij})^2)/(2\alpha)$
30. end for
31. end for
32. Adjust the SVM kernel function's positions according to UB and LB
33. Evaluate the new position of the SVM kernel functions
34. Update the position of the SVM kernel functions
35. $t = t + 1$
36. end while

OUTPUT: Output best SVM kernel function

Algorithm 1. Chameleon swarm support vector machine.

2.6 Proposed design

This proposed design included all processes involved in a lung cancer classification system from image acquisition and finally classification. Acquired images went through pre-processing stage and feature extraction using LBP, and the formulated CS-SVM algorithm was incorporated to perform feature selection and classification. The design was implemented in the Matrix Laboratory (R2020a) system specification of 2.60-Ghz processor, 500-GB HDD (hard disk drive), 4 GB of RAM and 64-bit operating system on a Windows 10 platform. Graphical user interface was designed using tool box in a MATLAB development environment for simulating classification system.

3 Results and discussion

3.1 Analyzing results with the malignant dataset

The outcomes of the SVM and CS-SVM techniques applied to the performance measures utilizing malignant datasets are shown in Table 1. The SVM technique yielded a 5.95% false-positive rate (FPR), 96.69% sensitivity, 94.05% specificity, and 95.90% accuracy in 59.89 s. Similarly, in 39.87 s, the CS-SVM approach produced 97.33% accuracy, 97.71% sensitivity, 96.43% specificity, and a false-positive rate of 3.57%. According to the results in Table 1, the CS-SVM approach performed better than the SVM technique in terms of recognition accuracy, sensitivity, specificity, and false-positive rate.

3.2 Analyzing results with the benign dataset

The findings for the SVM and CS-SVM approaches in terms of the performance metrics using Benign datasets are shown in Table 2. A false-positive rate of 11.11%, sensitivity of 91.67%, specificity of 88.89%, and accuracy of 90.83% at 59.92 s were all attained by the SVM technique. The CS-SVM method also produced results at 39.86 s with a false-positive rate of 5.56%,

TABLE 1 Findings using the SVM and CS-SVM methods on the malignant dataset.

Technique	FPR (%)	Specificity (%)	Sensitivity (%)	Recognition accuracy (%)	Recognition time (s)
SVM	5.95	94.05	96.69	95.90	59.89
CS-SVM	3.57	96.43	97.71	97.33	39.87

TABLE 2 Results obtained by the SVM and CS-SVM techniques with the benign dataset.

Technique	FPR (%)	Specificity (%)	Sensitivity (%)	Recognition accuracy (%)	Recognition time (s)
SVM	11.11	88.89	91.67	90.83	59.92
CS-SVM	5.56	94.44	94.05	94.17	39.86

TABLE 3 Results obtained by the SVM and CS-SVM techniques with the normal dataset.

Technique	FPR (%)	Specificity (%)	Sensitivity (%)	Recognition accuracy (%)	Recognition time (s)
SVM	9.60	90.40	94.85	93.51	61.74
CS-SVM	6.40	93.60	96.22	95.43	40.40

sensitivity of 94.05%, specificity of 94.44%, and recognition accuracy of 94.17%. According to the results in Table 2, the CS-SVM approach performed better than the SVM technique in terms of recognition accuracy, sensitivity, specificity, and false-positive rate.

3.3 Analyzing results with the normal dataset

The outcomes of the SVM and CS-SVM techniques applied to the performance measures utilizing the normal dataset are shown in Table 3. A false-positive rate of 9.60%, sensitivity of 94.85%, specificity of 90.40%, and accuracy of 93.51% at 61.74 s were attained by the SVM technique. Similar results were obtained at 40.40 s with the CS-SVM technique: a false-positive rate of 6.40%, sensitivity of 96.22%, specificity of 93.60%, and recognition accuracy of 95.43%. According to the results in Table 3, the CS-SVM approach performed better than the SVM technique in terms of recognition accuracy, sensitivity, specificity, and false-positive rate.

3.4 Comparative analysis of the results for all the datasets

The comparison analysis of the evaluation results obtained from both techniques shows that the developed CS-SVM outperformed SVM across all metrics. The false-positive rates for CS-SVM and SVM were 3.57% and 5.95% for malignant, 5.56% and 11.11% for benign, and 6.40% and 9.60% for normal, respectively. Also, the results of specificity for CS-SVM and SVM were 96.43% and 94.05% for malignant, 94.44% and 88.89% for benign, and 93.60% and 90.40% for normal, respectively. The CS-SVM and SVM produced sensitivity rates of 97.71% and 96.69% for malignant, 94.05% and 91.67% for benign, and 96.22% and 94.85% for normal, respectively. Furthermore, CS-SVM and SVM attained recognition accuracy of 97.33% and 95.90% for malignant, 94.17% and 90.83% for benign, and 95.43% and 93.51% for normal, respectively. In addition, recognition time for CS-SVM and SVM were 39.87 s and 59.89 s for malignant, 39.86 s and 59.92 s for benign, and 40.40 s and 61.74

s for normal, respectively. Table 4 shows the summary of the comparison of both techniques with respect to the datasets.

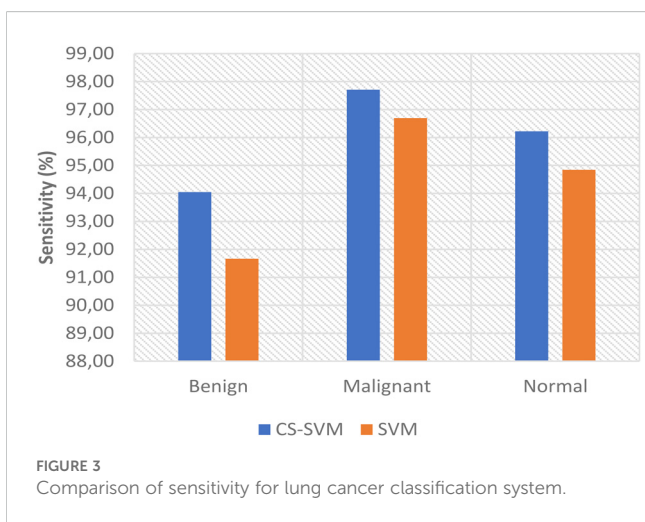
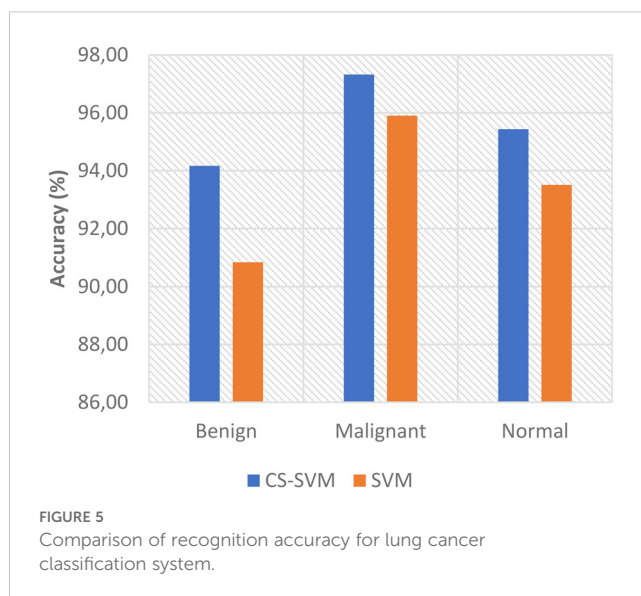
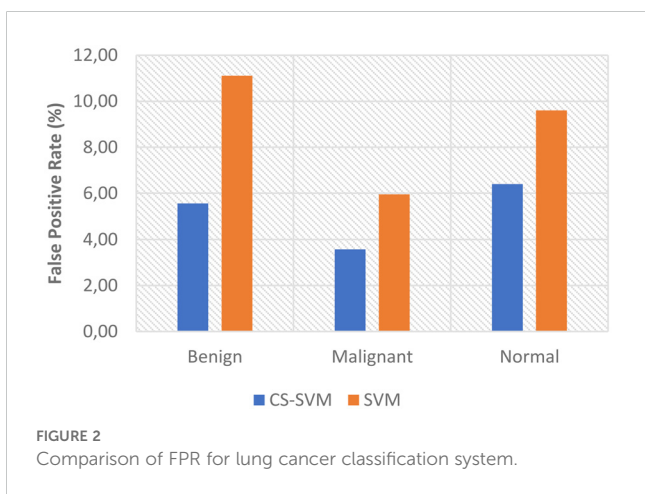
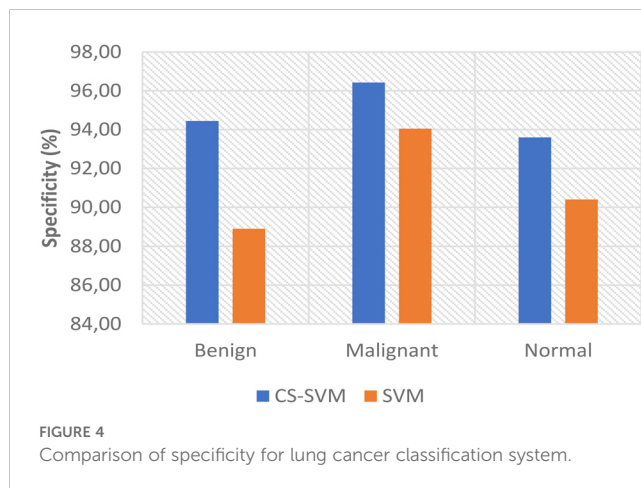
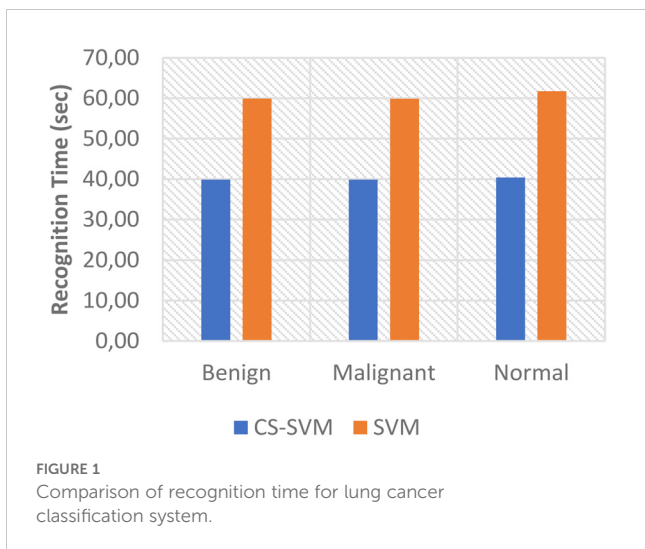
The charts showing the comparison of each of the evaluation metrics for the lung cancer classification system are as follows (Figures 1–5):

3.5 Limitation of the study

This study uses a restricted dataset that does not entirely mirror the variability observed within clinical settings across different contexts. Differences in image quality, patient demographics, or clinical conditions could influence how broadly this study’s findings can be applied. Despite having many pictures, the particular dataset has fewer pictures relative to the numerous variations present in actual clinical situations.

TABLE 4 Combined results for CS-SVM and SVM with respect to the datasets.

	Malignant	Benign	Normal	Average
FPR (%)				
SVM	5.95	11.11	9.60	10.22
CS-SVM	3.57	5.56	6.40	5.18
Specificity (%)				
SVM	94.05	88.89	90.40	91.11
CS-SVM	96.43	94.44	93.60	94.82
Sensitivity (%)				
SVM	96.69	91.67	94.85	94.40
CS-SVM	97.71	94.05	96.22	95.99
Recognition accuracy (%)				
SVM	95.90	90.83	93.51	93.41
CS-SVM	97.33	94.17	95.43	95.64
Recognition time (s)				
SVM	59.89	59.92	61.74	60.52
CS-SVM	39.87	39.86	40.40	36.71



3.6 Strength of the study

A robust methodology was used in the study, which employs advanced image processing techniques as well as the CSA used in optimizing parameters for SVMs. This strength is very important. The comparative analysis between SVMs and CS-SVMs is thorough when looking at different performance metrics. This is commendable and adds value to the study.

4 Conclusion

This research developed chameleon swarm–based SVM for lung cancer classification system. This conclusion explains why the developed technique performed better than the other method examined in this study in terms of recognition accuracy, false-

positive rate, sensitivity, recognition time, and specificity. This serves as evidence that enhancing the performance of systems for the classification of lung cancer disease can be achieved more successfully with the use of the CS-SVM technique. By leveraging the optimal selection of SVM parameters through the CS-SVM approach, the method significantly reduced false positive rates while achieving higher classification accuracy.

4.1 Future scope

More feature extraction and fusion methods may be included in future studies for better results and system performance investigation. Furthermore, higher convergence optimization algorithms can be incorporated for comparison and understanding of which technique returns the best results in the context of cancer classification.

Data availability statement

The original contributions presented in the study are included in the article/supplementary material. Further inquiries can be directed to the corresponding authors.

Author contributions

OA: Formal analysis, Project administration, Validation, Writing – original draft, Writing – review & editing. MO: Methodology, Supervision, Writing – original draft, Writing – review & editing. IR: Conceptualization, Investigation, Software, Writing – original draft, Writing – review & editing. OO: Conceptualization, Data curation, Formal analysis, Funding acquisition, Investigation, Methodology, Project administration, Resources, Software, Supervision, Validation,

Visualization, Writing – original draft, Writing – review & editing. AA: Conceptualization, Data curation, Formal analysis, Funding acquisition, Investigation, Methodology, Project administration, Resources, Software, Supervision, Validation, Visualization, Writing – original draft, Writing – review & editing.

Funding

The author(s) declare that no financial support was received for the research, authorship, and/or publication of this article.

Acknowledgments

The authors acknowledge the journal for the prompt review of our manuscript.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

References

- Neha C, Malviya AV. Image processing for detection of lung cancer: a review. *Int J Creative Res Thoughts (IJCRT)* (2021) 9(1):3064–70.
- Fransesco S, Ciccone G, Cappabianca S. Lung cancer: A review of pathogenesis, diagnosis, and treatment. *Cancer Treat Rev*. (2015) 41:547–61. doi: 10.3390/cancers15020351
- Bhandary A, Basthikodi M. Early diagnosis of lung cancer using computer aided detection via lung segmentation approach. *arXiv preprint arXiv*. (2021) 69(5). doi: 10.48550/arXiv.2107.12205
- Prajapati P, Handle V, Ingale A, Dwivedi S. Lung Cancer Detection and Classification Using SVM. *Journal of Emerging Technologies and Innovative Research (JETIR)*. Vol 6(5), pp60–65. Available online at: <https://www.jetir.org/papers/JETIRCS06015.pdf>.
- Braik MS. Chameleon Swarm Algorithm: A bio-inspired optimizer for solving engineering design problems. *Expert Syst Appl*. (2021) 174:114685. doi: 10.1016/j.eswa.2021.114685
- Mostafa RR, Ewees AA, Ghoniem RM, Abualigah L, Hashim FA. Boosting chameleon swarm algorithm with consumption AEO operator for global optimization and feature selection. *Knowledge-Based Syst*. (2022) 246:108743. doi: 10.1016/j.knsys.2022.108743
- Said M, El-Rifaie AM, Tolba MA, Houssein EH, Deb S. An efficient chameleon swarm algorithm for economic load dispatch problem. *Mathematics*. (2021) 9:2770. doi: 10.3390/math9212770
- Ogundokun RO, Awotunde JB, Sadiku P, Adeniyi EA, Abiodun M, Dauda OI. An enhanced intrusion detection system using particle swarm optimization feature extraction technique. *Proc Comput Sci*. (2021) 193:504–12. doi: 10.1016/j.procs.2021.10.052
- Dinh PH. Medical image fusion based on enhanced three-layer image decomposition and chameleon swarm algorithm. *Biomed Signal Process Control*. (2023) 84:104740. doi: 10.1016/j.bspc.2023.104740
- Zhang J, Zhang G, Kong M, Zhang T, Wang D, Chen R. CWOA: A novel complex-valued encoding whale optimization algorithm. *Mathematics Comput Simulation*. (2023) 207:151–88. doi: 10.1016/j.matcom.2022.12.022
- Oladipo ID, AbdulRaheem M, Awotunde JB, Bhoi AK, Adeniyi EA, Abiodun MK. Machine learning and deep learning algorithms for smart cities: a start-of-the-art review. *IoT IoE driven smart cities*. (2021), 143–62. doi: 10.1007/978-3-030-82715-1_7
- Awotunde JB, Ogundokun RO, Adeniyi AE, Abiodun KM, Ajamu GJ. Application of mathematical modelling approach in COVID-19 transmission and interventions strategies. In: *Modeling, Control and Drug Development for COVID-19 Outbreak Prevention*. Springer International Publishing, Cham (2021). p. 283–314.
- Huang S, Cai N, Pacheco PP, Narrandes S, Wang Y, Xu W. Applications of support vector machine (SVM) learning in cancer genomics. *Cancer Genomics Proteomics*. (2018) 15:41–51.

14. Ragab DA, Sharkas M, Marshall S, Ren J. Breast cancer detection using deep convolutional neural networks and support vector machines. *PeerJ*. (2019) 7:e6201. doi: 10.7717/peerj.6201
15. Fang H, Fu X, Zeng Z, Zhong K, Liu S. An improved arithmetic optimization algorithm and its application to determine the parameters of support vector machine. *Mathematics*. (2022) 10:2875. doi: 10.3390/math10162875
16. Nishio M, Nishizawa M, Sugiyama O, Kojima R, Yakami M, Kuroda T, et al. Computer-aided diagnosis of lung nodule using gradient tree boosting and Bayesian optimization. *PLoS One*. (2018) 13:e0195875. doi: 10.1371/journal.pone.0195875
17. Şekeroğlu B, Emirzade E. (2018). A computer aided diagnosis system for lung cancer detection using support vector machine, in: *Third international workshop on pattern*, American Journal of Applied Sciences: Science Pub. 7(12):1532–8.
18. Sharma S, Kaur M, Saini D. Lung cancer detection using convolutional neural network. *Int J Eng advanced Technol*. (2019) 8:3256–62. doi: 10.35940/ijeat.F8836.088619
19. Singh GAP, Gupta PK. Performance analysis of various machine learning-based approaches for detection and classification of lung cancer in humans. *Neural Computing Appl*. (2019) 31:6863–77. doi: 10.1007/s00521-018-3518-x
20. Bao J, Walliander M, Kovács F, Nagaraj AS, Hemmes A, Sarhadi VK, et al. SpARQ: an image analysis tool to visualise and quantify spatial phenotypes applied to non-small cell lung cancer. *Sci Rep*. (2019) 9:17613. doi: 10.1038/s41598-019-54038-9
21. Bruntha PM, Pandian SIA, Anitha J, Mohan P, Dhanasekar S. (2020). Local ternary co-occurrence patterns based lung nodules detection, in: *In 2020 6th International Conference on Advanced Computing and Communication Systems (ICACCS)*, Coimbatore, India: IEEE. pp. 489–92.
22. Karthiga B, Rekha M. Feature extraction and I-NB classification of CT images for early lung cancer detection. *Materials Today: Proc*. (2020) 33:3334–41. doi: 10.1016/j.matpr.2020.04.896
23. Banerjee N, Das S. (2020). Prediction lung cancer—in machine learning perspective, in: *2020 International Conference on Computer Science, Engineering and Applications (ICCSEA)*, Gunupur, India: IEEE. pp. 1–5.
24. Bhatt SD, Soni HB. (2021). Image retrieval using bag-of-features for lung cancer classification, in: *2021 6th International conference on inventive computation technologies (ICICT)*, Coimbatore, India: IEEE. pp. 531–6.
25. Yunianto M, Soeparmi S, Cari C, Anwar F, Septianingsih DN, Ardyanto TD, et al. Klasifikasi Kanker Paru Paru menggunakan Naïve Bayes dengan Variasi Filter dan Ekstraksi Ciri GLCM. *Indonesian J Appl Phys*. (2021) 11:256–68. doi: 10.13057/ijap.v11i2.53213
26. Aroba OJ, Xulu T, Msani NN, Mohlakoana TT, Ndlovu EE, Mthethwa SM. (2023). The adoption of an intelligent waste collection system in a smart city, in: *2023 Conference on Information Communications Technology and Society (ICTAS)*, Durban, South Africa: IEEE, pp. 1–6.
27. Aroba OJ. Improving node localization and energy efficiency for wireless sensor networks using hyper-heuristic optimization algorithms. *PhD diss*. (2022).
28. Aroba OJ. An ERP SAP implementation case study of the South African Small Medium Enterprise sectors. *Int J Computing Sci Res*. (2023) 7:2196–211. doi: 10.25147/ijcsr.2017.001.1.155
29. Aroba OJ. Professional leadership investigation in big data and computer-mediated communication in relation to the 11th Sustainable Development Goals (SDG) Global Blueprint global blueprint. *Int J Computing Sci Res*. (2024) 8:2592–611. doi: 10.25147/ijcsr.2017.001.1.177
30. Umamageswari A, Bharathiraja N, Irene DS. A novel fuzzy C-means based chameleon swarm algorithm for segmentation and progressive neural architecture search for plant disease classification. *ICT Express*. (2021) 9(2):160–7. doi: 10.1016/j.icte.2021.08.019