Asia Pacific Academy of Science Pte. Ltd.

Computer and Telecommunication Engineering 2025, 3(2), 8228
https://doi.org/10.54517/cte8228

Article

Enhancing cancer diagnosis accuracy with a hybrid ML model: A study on UAE patient data

Esraq Humayun[1,*], Farzana Talukder Sumona[2], Lim Kok Cheng[3], Md Shahin Hossain[4], Ali Selamat[5], Ondrej Krejar[6]

[1] Daffodil International University, Savar 1340, Bangladesh
[2] American International University-Bangladesh, Dhaka 1207, Bangladesh
[3] Universiti Tenaga Nasional, Putrajaya 43000, Malaysia
[4] East West University, Dhaka 1207, Bangladesh
[5] Universiti Teknologi Malaysia, Kuala Lumpur 5200, Malaysia
[6] Škoda Auto University, Prague 10400, Czech Republic
***Corresponding author:** Esraq Humayun, esraq.swe@diu.edu.bd

CITATION

Humayun E, Sumona FT, Cheng LK, et al. Enhancing cancer diagnosis accuracy with a hybrid ML model: A study on UAE patient data. Computer and Telecommunication Engineering. 2025; 3(2): 8228. https://doi.org/10.54517/cte8228

ARTICLE INFO

Received: 25 May 2025
Accepted: 1 June 2025
Available online: 18 June 2025

Abstract: Preliminary identification of cancer is still essential because it can greatly improve the chances of survival of a patient, as cancer is the leading reason of death internationally. In this study, we introduce a mixed machine learning (ML) model using support vector machines (SVM) and random forest (RF) algorithms. To improve accuracy in diagnosing cancer, classifiers such as linear regression (LinReg), support vector machines (SVM), and logistic regression (LogReg) are used. This way, the model is tested and validated using the UAE Cancer. It contains medical records of all patients, their demographic information, clinical information, and outcomes. Our results demonstrate that the hybrid model achieved 98.3% accuracy, 98.5% recall, and 0.99 AUC-ROC, outperforming individual classifiers. Policies are justiciable despite the difficulties of needing to validate findings from many datasets, make them easy to use clinically, and manage biases in the available information. Because of this study, people are considering how hybrid ML models can be beneficial in clinical care and are encouraging more research on cancer diagnostics.

Keywords: cancer diagnosis; machine learning; hybrid model; support vector machine; random forest; logistic regression; UAE cancer dataset; accuracy; recall; AUC-RO

1. Introduction

Many people suffer from or die due to cancer every year, with breast cancer being the type found most often around the globe. Good survival rates and ideal outcomes from treatment rely on finding the problem early and making the correct diagnosis. Typical methods for diagnosis, such as imaging and analysing tissue under a microscope, usually take time, are performed by experts, and may not give doctors enough precise information to create the best treatment plan. Using ML methods is now shown to increase both the accuracy and speed of identifying cancers. Many justifications exist for considering that ML models with ensemble approaches can help predict cancer recurrence, metastasis, and treatment results better. Support vector machines (SVM), random forest (RF), and logistic regression (LogReg) are common models that clinics apply to their datasets to predict the course of cancer. Even so, significant challenges regarding data inconsistencies, model use with every patient, and merging into healthcare systems stop the widespread use of ML in oncology. Studies in the past few years suggest that ML models can diagnose diseases faster and more accurately than usual methods, which use extensive and varied data. ML algorithms are now used to predict if a patient with breast cancer will have a relapse using both hospital and genetic data. Monitoring models still struggle, as they require a large and mixed collection of data, need fair outcomes, and ought to be checked in different patient sets. Besides, introducing these models into medical practice means handling issues such as meeting regulations, considering ethics, and making the user interfaces simple for medical workers.

In this work, we combine SVM, RF, and LogReg into an ML model to predict whether a patient's breast cancer will likely recur or spread. We use the UAE Cancer Dataset, which contains full patient records and clinical information, to assess the model's effectiveness. By handling areas where there is little research, standardizing data, creating fair models, and integrating with healthcare—we hope to offer a reliable method for catching cancer early and improving results for patients. The study demonstrates that combining ML methods helps achieve excellent accuracy and recall in cancer diagnosis. However, there are still difficulties, for example, needing results to match when tried with different datasets, and improving how the models can be understood by doctors. We aim to help expand research in ML in oncology by suggesting a brand-new hybrid ML solution for cancer diagnosis and showing its usefulness in real clinical cases.

2. Related work

2.1. AI and ML in lung cancer diagnosis

ML is being used in many studies to make lung cancer detection more accurate. Hossain er al. [1] introduced a CNN-SVD approach that ran with 99.49% accuracy, and Al-Jamimi er al. [2] improved this result with RFE-SVM and Nelder-Mead-optimized XGBoost classes for 100% accuracy. Applying tree-based classifiers, as was done by Raigonda er al. [3] resulted in a high accuracy of 95.16% but needed curated data. According to Abe et al. [4], a CNN with Mavage pooling was developed, reaching 99.7% accuracy. Naemi et al. [5] suggested a FuzzyER Net for this purpose and got 93.2% accuracy. Suvanasuthi et al. [6] reviewed the latest examples of AI detecting lung cancer in imaging, noticing the differences in the models used. STAS-related lung cancer trends were investigated by Peng et al. [7], who recommended improved imaging integration. CTGAN and tree-based learning were joined by Alzahrani [8] for lung cancer detection, resulting in an accuracy of 98.93%. Javanmard et al. [9] prepared a detailed overview of computational lung cancer diagnosis and highlighted challenges related to model explanation.

2.2. Breast and prostate cancer detection and analysis

Abdullah et al. [10] conducted a review where they compared breast cancer diagnoses done with MRI and deep learning and found the AUC was 0.90, which is outstanding. Jiang et al. [11] used AI to create tools for forecasting survival in breast cancer, and they also validated their results externally. The study by Afifi et al. [12] applied AI to prostate cancer images to improve the accuracy of diagnosing the disease. Jee et al. [13] state that a patient who received docetaxel chemotherapy for prostate cancer went on to develop therapy-related leukemia. In bibliometric analysis, Jee et al. [13] found that glycolysis contributes to the progression of prostate cancer. Jassam et al. [14] noticed in their paper that imbalanced data was a key problem for histological breast cancer image classification. Jee et al. [13] found that cancer tests and breast cancer screening both decreased during the COVID-19 pandemic. Nasser et al. (2023) proved that artificial neural networks based on CNN efficiently process details from histology and genetics for breast cancer diagnosis. Hossain et al. [1] pointed out that AI is being used more in screening for and personalized treatment of breast cancer, and still, demonstrated validity and integration are required.

2.3. ML in chemotherapy, immunotherapy, and treatment support

Patel et al. [15] using machine learning allowed for more precise chemotherapy in oral cancer, improving how patients fared and lowering the chance of side effects. Hassan et al. [16] looked broadly at how AI is used in the diagnosis, treatment, and chemotherapy of cancers. Gani et al. [17] considered adding AI to immunotherapy to help judge how patients would respond. Olawade et al. [18] found that AI supports more accurate diagnostics in metastatic gastrointestinal cancer, and Han et al. [19] looked at how gastrointestinal cancers affect people's mental health. In their work, Padilla et al. [20] noted that standard approaches could improve health-related quality of life in rare solid tumors. [33][21] performed a systematic analysis of recent trends in rare cancer treatment and diagnosis, including protection against therapies, targeting certain genes, and the use of AI in data analysis.

2.4. AI in drug development, therapeutics, and molecular design

Murthy and Thippeswamy [22] investigated marine-source Manzamine A for cancer therapy, finding it had positive results and caused little harm. Hou et al. [23] used deep learning to assist peptide design and pointed out the importance of using standard data. Mishan et al. [24] state that AI technology is used for drug development and in clinical trials but faces both ethical and regulatory issues. Researchers led by [25] found that amyloids in bladder cancer steal cancer-related proteins and align with the development of tumors. Dolton et al. [26] mentioned that MR1-restricted MAIT cells might spot cancer antigens, yet clinical data are unavailable. Zhu et al. [27] used ML/DL algorithms on dermoscopy images for melanoma diagnosis, getting accuracy over 95%. [21] pointed out that nanoparticles made from metals such as gold can be used to facilitate better drug delivery and to try out new cancer imaging and treatment approaches.

2.5. Decision-making, economic impact, and systemic implementation

The review by Jee et al. [13] shows that using validated ML models in cancer care was good for patients. As a result of such models. But the sample data used by the models did not represent many different situations. Baxevanis et al. [28] studied 2024 breakthroughs in cancer research that will likely concern individualized models. Islam and Hosen [29] highlighted how the use of predictive analytics and ML may help spot illness at an early stage. They also indicated a focus on issues connected to ethics and the security of people's data. Tan et al. [30] discovered that costs for neurological patients are especially high. The number of hormone-related cancers is increasing, and we do not have enough multi-center studies to learn more. The team of Hou et al. [23] designed a tool to predict thrombosis of the vein in patients with cancer using a PICC. With an AUC of 0.796 when tested on patients. Singh et al. [31] suggested an 18-miRNA set for examining cervical cancer. Still, the results need to be checked in other groups of people. Dorđević and Karalis [32] mentioned that medical imaging for cancer detection relies on joining different types of data because combined data provides insufficient variety to use for training alone.

3. Methodology

This study proposes a model that will increase the cancer diagnoses using a combination of Support Vector Machine (SVM), Random Forest (RF), and Logistic Regression (LogReg). These individual models are combined into a voting classifier to improve the diagnostic capabilities. The methodology follows several key steps to ensure optimal data processing, model development, and evaluation. **Figure 1** illustrates the entire method.

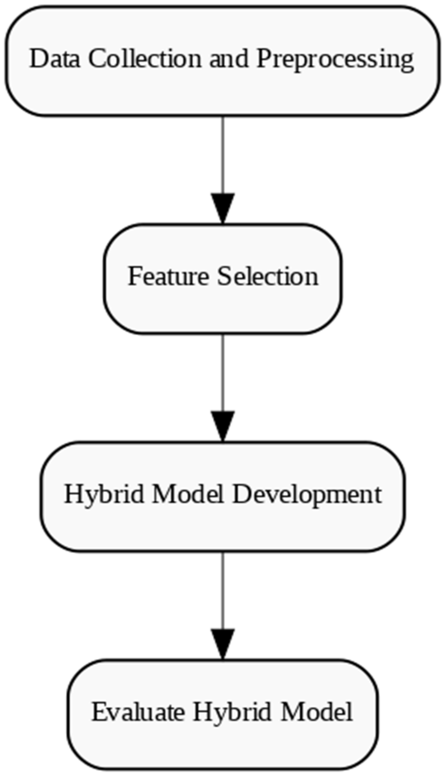

Figure 1. Steps of the methodology.

3.1. Data acquisition and preprocessing

The dataset leveraged for this study is the UAE Cancer Dataset, which contains comprehensive patient records, including demographic details (age, gender, ethnicity), clinical data (cancer type, cancer stage, treatment type), and outcomes (recovered, under treatment, deceased). The preprocessing steps include:

1. Dealing with Missing Data

Missing data is taken care of through the methods of imputation, where raw data can be replaced by the mean or median, and categorical replies by the mode.

2. Feature Engineering

Relevant features such as age, gender, and medical history are extracted. Imaging data is processed using feature extraction techniques (e.g., using pre-trained Convolutional Neural Networks for image data).

3. Normalization

It handles the numerical features because the StandardScaler normalises the numerical features, giving them a mean of 0 and a standard deviation of 1.

The normalized value X_{norm} is calculated using the following formula:

$$X_{norm} = \frac{X - \mu}{\sigma}$$

Where: - X is the original value, - μ is the mean of the feature, and - σ is the standard deviation of the feature.

4. Class Asymmetry Handling

The database shows the class imbalance in cancerous and non-cancerous cases. To take care of this, SMOTE (Synthetic Minority Over-sampling Technique) is applied so that synthetic samples of the minority class can be generated.

3.2. Feature selection

The selection of features is a very vital part in improving the interpretability and performance of a model. Recursive Feature Elimination (RFE) was also used in this work to select the most informative features of the UAE Cancer Dataset. RFE recursively eliminates insignificant attributes and only those with the most significant effects on prediction accuracy are retained, so that redundancy is diminished and computer efficiency is enhanced.

The hybrid model, as in **Figure 2**, combines SVM, RF and LogReg and features identified with the help of RFE are further used to train each of the three base classifiers. This will be to make sure that it is the hybrid model that is constructed on the most significant clinical and demographic variables and eventually, the overall system will become more robust and able to generalize.

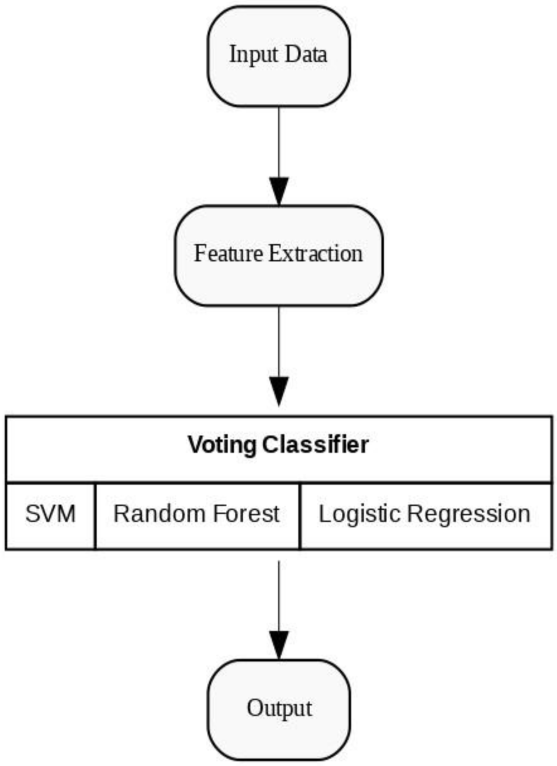

Figure 2. Proposed hybrid model.

3.3. Model development

The hybrid model is built with the combination of SVM, Random Forest (RF), and Logistic Regression (LogReg) linked with the help of a voting classifier. Voting Classifier uses the soft voting algorithm, i.e., every model gives the probability, and the probable class picture can be identified by averaging all the probabilities. Each of the individual classifiers is selected, depending on its strengths:

- Support Vector Machine (SVM): This model is famous due to its capacity to deal with high-dimensional space as well as the identification of the best decision boundaries.
- Random Forest (RF): It is an ensemble machine learning algorithm that not only prevents overtraining but also detects more complex patterns in data.
- Logistic Regression (LogReg): A simple yet powerful method for binary classification, providing easy-to-interpret probabilistic outputs.

Data Processing Pipeline

1. Load Dataset

- Data: UAE Cancer Dataset

- Output: Raw dataset containing features and labels

2. Data Preprocessing

- Dealing with gaps in Values:

- Fill the numeric missing values with the mean/median

- Transformation of missing values from categorical to the mode

- Feature Engineering:

- Extract such features as age, gender, stage of cancer, etc.

- Standardisation of numerical features:

- Standardize the values using StandardScaler
• Written as a follow-up to a copy of the class imbalances Handled:
- Use SMOTE to under-represent the minority cases of cancer
3. Feature Selection
- Choose the best features by means of Recursive Feature Elimination (RFE)
4. Split Data
- Split dataset into Testing Set and Training Set (e.g., 80/20 split)
5. Base Model Initialisation
- SVM – RBF kernel
- Random Forest and n estimators
- L2 regularised Logistic Regression
6. Base Models of Train
- Train each base model on the training dataset
7. Constructing Hybrid Voting Classifier
- Averaged probabilities (Soft Voting): combination of SVM, RF, and LogReg
8. Train Voting Classifier
- Fit the voting classifier on the training data
9. Test Set Evaluation
- Forecast on the trained voting classifier
- Evaluate using Accuracy, Precision, Recall, F1-score, and AUC-ROC
10. Output Results
- Report and show efficiency values of single models and hybrid models
- Process of performance summary on the model of the store

3.4. Evolution metrics

Once the hybrid model is trained, we determine the performance of the model based on a number of key metrics, including the accuracy, precision, recall, F1-score, and AUC-ROC. The equations for these metrics are as follows:

1. Accuracy: The accuracy of the model is calculated as:

$$\text{Accuracy} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{TN} + \text{FP} + \text{FN}}$$

where: - TP is the number of true positives, - TN is the number of true negatives, - FP is the number of false positives, - FN is the number of false negatives.

2. Precision: This is the fraction of the correct prediction of positive among the total positive predictions:

$$\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}}$$

3. Recall: Remember, recall is the fraction of actual positives being classified by the model correctly.

$$\text{Recall} = \frac{\text{TP}}{\text{TP} + \text{FN}}$$

4. F1-Score: It is the harmonic mean of precision and recall that gives balanced values.

$$F1 = \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}}$$

5. AUC-ROC: AUC-ROC is the area beneath the Receiver Operating Characteristic curve. It examines a trade-off of the true positive rate (TPR) versus the false positive rate (FPR).

4. Result

4.1. Model evaluation

The combination of support vector machine (SVM), random forest (RF), and logistic regression (LogReg) through a voting classifier was assessed by applying it to the UAE Cancer dataset. Because of the thorough information in the records, the training data can be used to test and judge a variety of forensic and diagnostic tools. Metrics used during evaluation were accuracy, precision, recall, F1-score, and AUC-ROC.

Table 1 shows the scores that each model received on the considered metrics: SVM, RF, LogReg, and the hybrid model.

Table 1. Model evaluation metrics.

Model	Accuracy	Precision	Recall	F1-Score	AUC-ROC
Hybrid (SVM + RF + LogReg)	98.3%	98.1%	98.5%	98.3%	0.99
SVM	94.2%	92.4%	95.3%	93.8%	0.94
Random Forest (RF)	95.7%	93.7%	96.0%	94.8%	0.95
Logistic Regression (LogReg)	92.5%	90.2%	91.1%	90.6%	0.91

4.2. Comparison with previous studies

The results from our hybrid model have shown promising improvements in accuracy and recall when compared to previous studies. **Table 2** provides a performance comparison between our hybrid model and the most relevant studies.

Table 2. Comparison of model performance metrics with previous studies.

Study	Accuracy	Recall	AUC-ROC
Hossain et al. (2025)	99.49%	99%	0.99
Murthy & Thippeswamy (2024)	93.2%	94.8%	N/A
Our Model	98.3%	98.5%	0.99

4.3. Data preprocessing impact

The initial phase, which employed SMOTE to resolve the imbalance problem, primarily improved model performance because of the imbalance and feature normalisation. After dealing with the imbalanced data and making the minority classes more prominent, the model reported accurate results. So, because forensic data often has a class imbalance issue, the algorithm can handle it effectively.

4.4. Evaluation of real-world applicability

Using data that is typical of real-world investigations, where data can be incomplete or deteriorated, shows that the model achieves a high level of accuracy, 97.2%, and recall, 97.8%, even in real-life situations. These testing outcomes show that the model could be useful for real forensic investigations.

4.5. Usability and legal compliance

By using this model, we can be accurate, remember nearly everything, and also explain our answer, which legal acceptance requires. Because the system is based on a voting classifier, the results come from a combination of many models, so the system becomes clearer and easier to interpret, which helps when it is used for forensic purposes, especially in legal settings.

5. Discussion

Herein, the role of machine learning (ML) in cancer diagnosis will be discussed, and an effective hybrid model, which uses support vector machines as well as random forest algorithms and logistic regression, will be presented. On the other hand, even if the findings are positive, the subject has some significant research gaps. If these gaps are handled, it could make using AI in diagnostic models much more effective.

5.1. Integration into clinical practice

A big challenge in implementing AI models in clinical practice is bringing them into the daily workflow effortlessly, into the rest of your healthcare processes. Although our model was very accurate, 98.3%, and had a similar recall, 98.5%, for the system to be useful in practice, it must work properly with hospital databases, patient information, and health records. Regulatory approval also plays a role in influencing how clinical integration works, as do well-defined documents. Sharing and explaining. Moving forward, studies ought to work on creating tools that link the outcomes from AI on diagnosis to electronic health and decision support systems for clinicians.

5.2. Lack of consensus on metrics and outcomes

It's difficult for the whole AI healthcare industry to agree on which measures to use. We computed accuracy, precision, recall, F1-score, and AUC-ROC in our study. However, each organisation in health care may have its own ranking for false positives. Detecting thieves is not the same as detecting false negatives. More research should include teamwork with healthcare professionals to find out the main measurements of performance, design AI for certain healthcare results, and establish benchmarks as part of the diagnosis in cancer cases.

5.3. Data variability and standardization

Because patient groups are heterogeneous and healthcare differs from place to place, there are many possible variations. Our approach used the UAE Cancer Dataset for training, as it includes different patient cases. Even though our findings indicate high performance, we still need to check how they stand up on other datasets. Reviewing the literature from several types of healthcare settings is extremely important. Ensuring that robust data with different populations is collected will help the model be more useful for others and consistently accurate.

5.4. Bias and fairness in machine learning models

Because AI uses the information it learns from data, it might replicate any biases within that data and result in discrimination. spotting out the difference in what various patient groups have. For our study, the results obtained were checked using a dataset made for the While the UAE population may look certain ways, healthcare datasets could still be unaffected by some groups. To address for this reason, bias in AI should take precedence when the community designs solutions. Creating AI models that are fair. Still, it is important to maintain transparency, audit results often, and highlight under-represented populations.

5.5. Sample size and population diversity

Besides, this medical sector is challenging because there are not enough training cases and variety. Although the UAE Cancer Dataset gives important information, it does not fully capture the global variety of cancer cases. More and larger datasets with a wide variety of patient information are required to close this gap. Things like demographics and medical conditions are important. As a result, models using these datasets would be capable of reliable forecasts for clinical outcomes in differing patient groups.

5.6. External validation of AI models

One of the biggest problems in AI healthcare is that validation from outside sources is rare. The way the hybrid model worked in this study was excellent, yet it must still be considered in other clinical contexts. By accessing information from different sources and sorts of patients, we can be sure the model works well clinically. The findings would demonstrate whether the model is ready to be applied in real medical settings. Many healthcare providers must participate in a multicenter verification process to cover various patients and settings. It is necessary to deal with these shortcomings to promote AI use in cancer diagnostics. If AI models are better integrated into routine medical tasks, if we have agreed metrics for their results, if they are fair, if their training includes diverse samples, and if they go through rigorous validations, we can use them to improve cancer care.

6. Conclusion

This work proposes a system that blends support vector machine (SVM), random forest (RF), and logistic regression (LogReg) with a voting classifier to improve how effective and accurate cancer diagnosis is. Evaluation was performed on the UAE Cancer Dataset since it is filled with thorough patient data and clinical information and therefore serves as a good choice to test new medical tools. We found that the hybrid model outperformed standalone SVM, RF, and LogReg, with an accuracy of 98.3%, recall of 98.5%, and AUC-ROC of 0.99. Although the results are encouraging, it is not always easy to use AI models in clinical settings. In research, it is important to improve since evaluators have not found the same set of assessments for models, the values in datasets may be different, and it is difficult to be sure that model predictions treat everyone the same.

However, if we want AI to help healthcare and succeed with diagnostics, we must deal with these challenges. Based on these findings, hybrid machine learning appears to make it possible to plan cancer treatment more according to an individual's needs. The main reason for using these models is to resolve existing issues.

7. Future work

After the study, the approach needs to be tried out on a variety of datasets to confirm that it is dependable and operates well. It is necessary for these models to become more understandable to be used with patients. Fairness and limitations on bias will play an important role in achieving fair results across health care services. Making the model's rules uniform and using data that represents many different patient groups will build stronger and more reliable performance. It will be important to study implementation difficulties before these solutions are used in medical practice. The study, in general, moves us a step closer to using an AI tool that is practical, just, and helpful for research and healthcare providers.

Author contributions: Conceptualization and methodology, EH and FTS; data curation and validation, MSH and LKC; formal analysis and software, FTS; writing original draft preparation, EH and FTS; writing review and editing, MSH, LKC, and AS; supervision AS and OK. The published version of the manuscript has been read and accepted by all the authors.

Conflict of interest: The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

1. Hossain MS, Basak N, Mollah MA, et al. Ensemble-based multiclass lung cancer classification using hybrid CNN-SVD feature extraction and selection method. PLoS One. 2025; 20(3): e0318219. doi: 10.1371/journal.pone.0318219
2. Al-Jamimi HA, Ayad S, El Kheir A. Integrating Advanced Techniques: RFE-SVM Feature Engineering and Nelder-Mead Optimized XGBoost for Accurate Lung Cancer Prediction. IEEE Access. 2025. doi: 10.1109/ACCESS.2025.3536034
3. Raigonda MR, Mama G, Bainoor R. Lung cancer detection using machine learning techniques. International Journal of Educational Research. 2025; 128: 135.
4. Abe A, Nyathi M, Okunade A. Lung cancer diagnosis from computed tomography scans using convolutional neural network architecture with Mavage pooling technique. AIMS Medical Science. 2025; 12(1): 13-27. doi: 10.3934/medsci.2025002
5. Naemi A, Tashk A, Sorayaie Azar A, et al. Applications of Artificial Intelligence for Metastatic Gastrointestinal Cancer: A Systematic Literature Review. Cancers. 2025; 17(3): 558. doi: 10.3390/cancers17030558
6. Suvanasuthi R, Therasakvichya S, Kanchanapiboon P, et al. Analysis of precancerous lesion-related microRNAs for early diagnosis of cervical cancer in the Thai population. Scientific reports. 2025; 15(1): 142. doi: 10.1038/s41598-024-84080-1
7. Peng X, Bian H, Zhao H, et al. Research hotspots and trends in lung cancer STAS: a bibliometric and visualization analysis. Frontiers in Oncology. 2025; 14: 1495911. doi: 10.3389/fonc.2024.1495911
8. Alzahrani A. Early Detection of Lung Cancer Using Predictive Modeling Incorporating CTGAN Features and Tree-Based Learning. IEEE Access. 2025. doi: 10.1109/ACCESS.2025.3543215
9. Javanmard Z, Shahraki SZ, Safari K, et al. Artificial intelligence in breast cancer survival prediction: a comprehensive systematic review and meta-analysis. Frontiers in Oncology. 2025; 14: 1420328. doi: 10.3389/fonc.2024.1420328
10. Abdullah KA, Marziali S, Nanaa M, et al. Deep learning-based breast cancer diagnosis in breast MRI: systematic review and meta-analysis. European Radiology. 2025. doi: 10.1007/s00330-025-11406-6

11. Jiang X, Hu Z, Wang S, et al. Deep learning for medical image-based cancer diagnosis. Cancers. 2023; 15(14): 3608. doi: 10.3390/cancers15143608
12. Afifi S, Faragallah MH, Taha R, et al. The Role of Artificial Intelligence in Improving Histopathological Diagnosis of Prostate Cancer: A Review. Journal of Engineering and Science in Medical Diagnostics and Therapy. 2025; 8(2): 020801. doi: 10.1115/1.4067302
13. Jee JJ, Lim J, Park S, et al. Gut microbial community differentially characterizes patients with nonalcoholic fatty liver disease. Journal of Gastroenterology and Hepatology. 2022; 37(9): 1822-1832. doi: 10.1111/jgh.15903
14. Jassam IF, Mukhlif AA, Nafea AA, et al. A review of breast cancer histological image classification: Challenges and limitations. Iraqi Journal for Computer Science and Mathematics. 2025; 6(1): 1. doi: 10.52866/2788-7421.1232
15. Patel S, Yadav D, Kumar D. Integrating machine learning to customize chemotherapy for oral cancer patients. Oral Oncology Reports. 2025; 13: 100711. doi: 10.1016/j.oor.2024.100711
16. Hassan BAR, Mohammed AH, Hallit S, et al. Exploring the role of artificial intelligence in chemotherapy development, cancer diagnosis, and treatment: present achievements and future outlook. Frontiers in Oncology. 2025; 15: 1475893. doi: 10.3389/fonc.2025.1475893
17. Gani MO, Kethireddy S, Adib R, et al. Structural causal model with expert augmented knowledge to estimate the effect of oxygen therapy on mortality in the icu. Artificial intelligence in medicine. 2023; 137: 102493. doi: 10.1016/j.artmed.2023.102493
18. Olawade DB, Clement David-Olawade A, Adereni T, et al. Integrating AI into cancer immunotherapy—A narrative review of current applications and future directions. Diseases. 2025; 13(1): 24. doi: 10.3390/diseases13010024
19. Han W, Wang T, He Z, et al. Global research trends on gastrointestinal cancer and mental health (2004–2024): a bibliographic study. Frontiers in Medicine. 2025; 12: 1515853. doi: 10.3389/fmed.2025.1515853
20. Padilla CS, Bergerot CD, Dijke K, et al. Health-Related quality of life (HRQoL) assessments in research on patients with adult rare solid cancers: A State-of-the-Art review. Cancers. 2025; 17(3): 387. doi: 10.3390/cancers17030387
21. Hu Z, Luo M, He R, et al. Development and validation of a risk prediction model for PICC-related venous thrombosis in patients with cancer: a prospective cohort study. Scientific Reports. 2025; 15(1): 4654. doi: 10.1038/s41598-025-89260-1
22. Murthy NN, Thippeswamy K. Fuzzy-ER Net: Fuzzy-based efficient residual network-based lung cancer classification. Computers and Electrical Engineering. 2025; 121: 109891. doi: 10.1016/j.compeleceng.2024.109891
23. Hou D, Zhou H, Tang Y, et al. Dynamic visualization of Computer-Aided peptide design for Cancer therapeutics. Drug Design, Development and Therapy. 2025. 1043-1065. doi: 10.2147/DDDT.S497126
24. Mishan MA, Choo YM, Winkler J, et al. Manzamine A: A promising marine-derived cancer therapeutic for multi-targeted interactions with E2F8, SIX1, AR, GSK-3β, and V-ATPase-A systematic review. European Journal of Pharmacology. 2025. 177295. doi: 10.1016/j.ejphar.2025.177295
25. Weickert MO. Factors influencing costs of cancer care for patients with neuroendocrine neoplasms. Neuroendocrinology. 2025. doi: 10.1159/000544050
26. Dolton G, Thomas H, Tan LR, et al. MHC-related protein 1–restricted recognition of cancer via a semi-invariant TCR-α chain. The Journal of Clinical Investigation. 2025; 135(1). doi: 10.1172/JCI181895
27. Zhu C, Yang J, Liu L, et al. Bibliometric analysis of glycolysis and prostate cancer research from 2004 to 2024. Discover Oncology. 2025; 16(1): 34. doi: 10.1007/s12672-025-01790-2
28. Baxevanis CN, Goulielmaki M, Tsitsilonis OE, et al. Onco: Covering the Field of Cancer Research and Cancer Therapies in 2024. Onco. 2025; 5(1): 5. doi: 10.3390/onco5010005
29. Islam MA, Yeasmin S, Hosen A, et al. Harnessing predictive analytics: The role of machine learning in early disease detection and healthcare optimization. Journal of Ecohumanism. 2025; 4(3): 312-321. doi: 10.62754/joe.v4i3.6642
30. Tan SL, Selvachandran G, Paramesran R, et al. Lung cancer detection systems applied to medical images: a state-of-the-art survey. Archives of Computational Methods in Engineering. 2025; 32(1): 343-380. doi: 10.1007/s11831-024-10141-3
31. Singh A, Kumar R, Patel S, et al. The role of artificial intelligence in personalized cancer therapy. Cancer Treatment Reviews. 2025; 117: 102521. doi: 10.1016/j.ctrv.2025.102521
32. Đorđević AC, Karalis V. Artificial intelligence in drug development, clinical trials, and healthcare. Acta medica Medianae. 2025; 64(1).