Annals of Healthcare Systems Engineering

www.ahse.reapress.com

Ann. Healthc. Syst. Eng. Vol. 3, No. 3 (2026) 190–203.

Paper Type: Original Article

Machine Learning-Based Prediction of ICU Admission in COVID-19 Patients Using CT-Derived Biomarkers

Somayeh Livani[1], Mohammad Mohajer Tabrizi[2,*], Karim Aqerkakli[1], Sona Roshani[2]

[1] Clinical Research Development Unit (CRDU), Sayad Shirazi Hospital, Golestan University of Medical Sciences, Gorgan, Iran; drslivani@gmail.com; karimaqerkakli@gmail.com.

[2] Department of Industrial Engineering, Faculty of Engineering, Golestan University, Gorgan, Iran; mohajertabrizi@gmail.com; sonaroshany0@gmail.com.

Citation:

Received: 19 November 2025 *Revised: 22 February 2026* *Accepted: 12 April 2026*	Livani, S., Mohajer Tabrizi, M., Aqerkakli, K., & Roshani, S. (2026). Machine learning-based prediction of ICU admission in COVID-19 patients using CT-derived biomarkers. *Annals of Healthcare Systems Engineering*, *3*(3), 190-203.

Abstract

Coronavirus disease 2019 (COVID-19) exhibits highly variable clinical outcomes ranging from mild symptoms to critical illness requiring Intensive Care Unit (ICU) admission. Early identification of high-risk patients remains essential for optimizing resource allocation. Recently, Computed Tomography (CT)-derived biomarkers such as gynecomastia, hepatic steatosis, and epicardial fat thickness have been proposed as potential indicators of systemic metabolic dysfunction associated with disease severity. This study aimed to develop and evaluate Machine Learning (ML) and deep learning models for predicting ICU admission in hospitalized COVID-19 patients using CT-derived imaging biomarkers, including gynecomastia, fatty liver, and epicardial fat thickness. A total of 341 hospitalized COVID-19 patients were retrospectively analyzed. Clinical and CT-derived variables were extracted, including age, gynecomastia subtypes, hepatic steatosis, epicardial fat thickness, and retro-mammary fat measurements. Multiple ML models (Random Forest, Gradient Boosting, ExtraTrees, Logistic Regression, Support Vector Machine (SVM), Decision Tree, and XGBoost) were trained using a pipeline incorporating imputation, scaling, and SMOTE-based oversampling. Model performance was evaluated using accuracy, sensitivity, specificity, F1-score, ROC-AUC, PR-AUC, and Matthews Correlation Coefficient (MCC). Shapley Additive Explanations (SHAP) analysis was applied for model interpretability. Among all models, XGBoost achieved the best overall performance (accuracy = 0.884, F1-score = 0.333), followed by ExtraTrees with the highest ROC-AUC (0.729). Epicardial fat thickness and hepatic steatosis were consistently identified as the most important predictors across feature importance and SHAP analyses. Deep learning models demonstrated inferior generalization performance (ROC-AUC ≈ 0.41–0.53), likely due to limited dataset size and class imbalance. CT-derived metabolic imaging biomarkers, particularly epicardial fat thickness and hepatic steatosis, provide meaningful predictive value for ICU admission in COVID-19 patients. ML models, especially ensemble methods, outperform deep learning approaches in structured clinical datasets. Explainable Artificial Intelligence (XAI) enhances clinical interpretability and supports the integration of imaging biomarkers into predictive decision-support systems.

Keywords: COVID-19, Computed tomography biomarkers, Machine learning, Intensive care unit admission prediction.

1 | Introduction

Coronavirus disease 2019 (COVID-19), caused by the severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2), remains a major global health concern despite widespread vaccination and improved

Corresponding Author: mohajertabrizi@gmail.com

https://doi.org/10.22105/ahse.v3i3.70

therapeutic strategies. Although most infections are mild, a substantial proportion of hospitalized patients still develop severe respiratory failure, systemic complications, Intensive Care Unit (ICU) admission, or death. The clinical heterogeneity of COVID-19 has made early risk stratification a critical component of patient management, particularly in resource-limited healthcare systems where ICU capacity is constrained. Early and accurate prediction of disease severity enables optimized triage, improved allocation of resources, and timely initiation of intensive care interventions [1].

During the pandemic, numerous prognostic models have been developed using traditional statistical methods and Machine Learning (ML) approaches. Conventional models primarily rely on demographic variables (age, sex), comorbidities (diabetes, hypertension), and laboratory markers (CRP, D-dimer, lymphocyte count). While these variables are clinically informative, they do not fully capture the systemic and metabolic complexity underlying COVID-19 progression. In contrast, ML algorithms such as Random Forest, XGBoost, LightGBM, and deep neural networks can model nonlinear relationships and high-dimensional interactions, often achieving superior predictive performance compared to classical regression methods.

However, a major limitation of existing ML-based COVID-19 prediction models is their underutilization of imaging-derived biomarkers. Chest Computed Tomography (CT), particularly High-Resolution CT (HRCT), has been widely used during the pandemic for diagnosing pneumonia, assessing pulmonary involvement, and guiding clinical management. Beyond pulmonary findings, CT imaging contains rich opportunistic information about systemic metabolic health, including visceral adiposity, hepatic fat content, and cardiac fat distribution. These parameters are increasingly recognized as "opportunistic imaging biomarkers," which can be extracted without additional cost or radiation exposure and may significantly enhance prognostic models.

Among CT-derived biomarkers, Epicardial Adipose Tissue (EAT) has gained increasing attention as a biologically active visceral fat depot. EAT is not merely a passive fat storage site but an endocrine organ that secretes pro-inflammatory cytokines such as interleukin-6 and tumor necrosis factor-alpha. These mediators contribute to systemic inflammation, endothelial dysfunction, and cardiovascular disease progression. Recent studies have demonstrated that increased epicardial fat thickness is independently associated with severe COVID-19 outcomes, including ICU admission and mortality, highlighting its potential role as a prognostic imaging biomarker.

Similarly, hepatic steatosis, now often referred to as Metabolic Dysfunction-Associated Steatotic Liver Disease (MASLD), reflects systemic metabolic dysregulation and chronic low-grade inflammation. It is strongly associated with obesity, insulin resistance, dyslipidemia, and cardiovascular disease. Non-contrast CT imaging allows objective estimation of hepatic fat content through liver attenuation measurements, which have demonstrated high specificity for moderate-to-severe steatosis [2]. Emerging evidence suggests that hepatic steatosis is associated with worse COVID-19 outcomes, likely due to its link with systemic inflammation and immune dysregulation.

Gynecomastia, defined as benign proliferation of male breast glandular tissue, is another frequently observed incidental finding on chest CT. Although traditionally considered a benign endocrine condition, gynecomastia reflects an imbalance between estrogen and androgen activity, often driven by increased aromatase activity in adipose tissue, liver dysfunction, obesity, and metabolic syndrome. The liver plays a central role in estrogen metabolism; therefore, hepatic dysfunction may contribute to hormonal imbalance and subsequent development of gynecomastia. Recent CT-based studies have reported incidental gynecomastia in hospitalized patients, particularly during COVID-19 imaging evaluations, suggesting potential associations with metabolic status and systemic disease severity [3], [4].

Importantly, growing evidence suggests that gynecomastia, hepatic steatosis, and epicardial fat thickness are not isolated findings but may represent interconnected manifestations of a shared metabolic-inflammatory pathway. Visceral adiposity promotes systemic inflammation and insulin resistance, which in turn contribute

to hepatic fat accumulation and endocrine dysregulation. Increased aromatization of androgens in adipose tissue further amplifies estrogenic activity, potentially contributing to gynecomastia. This interconnected metabolic network may therefore reflect a unified biological phenotype associated with worse clinical outcomes in systemic diseases such as COVID-19 [5], [6].

Despite the biological plausibility of these associations, previous prognostic models have largely ignored CT-derived metabolic and endocrine biomarkers. Furthermore, existing studies have not systematically evaluated the combined predictive value of gynecomastia, hepatic steatosis, and epicardial fat thickness within a unified ML framework. This represents a significant gap in the literature, particularly given the routine availability of chest CT scans in hospitalized COVID-19 patients.

Recent advances in Explainable Artificial Intelligence (XAI) have further enhanced the clinical applicability of ML models. Techniques such as Shapley Additive Explanations (SHAP) allow for transparent interpretation of feature contributions, enabling clinicians to understand the relative importance of each variable in individual predictions. This improves model trustworthiness and facilitates integration into clinical decision-making workflows [1].

To the best of our knowledge, no previous study has simultaneously incorporated CT-derived gynecomastia, hepatic steatosis, and EAT thickness into an explainable ML model for predicting clinical outcomes in hospitalized COVID-19 patients. This study therefore aims to develop and evaluate a novel imaging-based prognostic model that integrates these opportunistic biomarkers with clinical variables to improve early risk stratification and outcome prediction.

2 | Materials and Methods

2.1 | Study Design and Dataset

This retrospective cross-sectional study was conducted on a dataset of 341 hospitalized patients with confirmed COVID-19 infection. The data were extracted from chest CT scans and clinical records. The primary objective was to evaluate the predictive value of gynecomastia, hepatic steatosis (fatty liver), and epicardial fat thickness for clinical outcomes, specifically ICU admission as the primary endpoint.

All variables were derived from routine CT imaging performed during hospitalization, consistent with previous studies demonstrating the utility of opportunistic CT biomarkers in COVID-19 patients [3], [4].

2.2 | Variables and Outcome Definition

The dataset included the following variables:

I. Age (years)
II. Gynecomastia (right, left, bilateral)
III. Hepatic steatosis (fatty liver; binary)
IV. Epicardial fat thickness (continuous, mm)
V. Retro-mammary fat thickness (continuous, mm)
VI. ICU admission (binary outcome)

The primary outcome variable was ICU admission, representing severe disease requiring intensive care support.

A secondary variable (mortality/expire) was available but was not used in the final modeling phase in this analysis.

2.3 | CT-Based Measurements

Gynecomastia, hepatic steatosis, and epicardial fat were assessed using non-contrast chest CT scans, which have been validated in previous literature for opportunistic body composition analysis.

I. Hepatic steatosis was evaluated using liver attenuation values on CT, which correlate strongly with histopathological fat deposition [2].

II. Epicardial fat thickness was measured as a surrogate marker of visceral adiposity using axial CT images [6].

III. Gynecomastia was defined based on retroareolar glandular tissue thickness on CT imaging.

2.4 | Data Preprocessing

Data preprocessing included the following steps:

I. Missing value handling: Median imputation was applied using SimpleImputer.

II. Feature scaling: Standardization was performed using StandardScaler to normalize feature distributions.

III. Train-test split: Data were split into:

- *80% training set*
- *20% testing set: Stratified sampling was applied to preserve class distribution.*

(These preprocessing steps are essential for ensuring model stability and reproducibility in clinical ML pipelines.)

2.5 | Class Imbalance Handling

The dataset exhibited significant class imbalance in ICU admission outcome. To address this, Synthetic Minority Oversampling Technique (SMOTE) was applied only on the training set to avoid data leakage.

SMOTE has been widely validated for improving classifier performance in imbalanced medical datasets.

The effect of SMOTE was evaluated visually using:

- *Bar plots of class distribution*
- *Pie charts before and after resampling*

2.6 | Machine Learning Models

Multiple supervised ML algorithms were implemented and compared:

- *Random Forest*
- *Gradient Boosting Machine (GBM)*
- *Extra Trees Classifier*
- *Logistic Regression*
- *Support Vector Machine (SVM)*
- *Decision Tree*
- *XGBoost Classifier*

Tree-based ensemble methods were included due to their strong performance in structured clinical datasets and ability to capture nonlinear relationships [7].

2.7 | Model Training Pipeline

A unified pipeline was constructed using imblearn.Pipeline to ensure reproducibility:

Pipeline steps:

I. Median imputation
II. Standard scaling
III. SMOTE oversampling
IV. Classification model

This design ensures that preprocessing is strictly applied within cross-validation boundaries, preventing information leakage.

2.8 | Model Evaluation Metrics

Model performance was assessed using multiple metrics:

- *Accuracy*
- *Precision*
- *Recall (Sensitivity)*
- *Specificity*
- *F1-score*
- *ROC-AUC*
- *Precision-Recall AUC*
- *Matthews Correlation Coefficient (MCC)*
- *Negative Predictive Value (NPV)*

These metrics were selected to provide a robust evaluation in the presence of class imbalance, where accuracy alone is insufficient.

2.9 | Model Explainability

To interpret model predictions, feature importance and SHAP analysis were applied.

I. Feature importance was extracted from tree-based models.
II. SHAP values were computed using:

- *TreeExplainer for tree-based models*
- *LinearExplainer for logistic regression*
- *KernelExplainer for non-linear models*

SHAP allows quantification of feature contributions at both global and individual levels, improving interpretability in clinical ML [8].

2.10 | Deep Learning Model

In addition to classical ML models, a fully connected neural network was implemented:

Architecture:

- *Input layer*
- *Dense layer (64 neurons, ReLU)*
- *Dropout (0.3)*

- *Dense layer (32 neurons, ReLU)*
- *Dropout (0.3)*
- *Output layer (Sigmoid activation)*

The model was trained using:

- *Adam optimizer*
- *Binary cross-entropy loss*
- *Early stopping (patience = 10)*
- *Validation split = 20%*

Deep learning performance was compared with classical models using ROC-AUC and accuracy.

2.11 | Software and Implementation

All analyses were performed in Python using Google Colab. The following libraries were used:

- *Scikit-learn*
- *XGBoost*
- *Imbalanced-learn*
- *SHAP*
- *TensorFlow/Keras*
- *Matplotlib and Seaborn*

Statistical significance was not the primary endpoint; instead, predictive performance and model interpretability were emphasized.

3 | Results

3.1 | Dataset Characteristics and Exploratory Analysis

A total of 341 hospitalized COVID-19 patients were included in the final analysis. All variables derived from CT imaging and clinical records were complete, with no missing data observed across features.

The primary outcome variable (ICU admission) showed a significant class imbalance, with 35 ICU cases (10.3%) and 306 non-ICU cases (89.7%).

As in *Fig. 1*, histograms demonstrated heterogeneous distributions across demographic and imaging-derived variables, particularly epicardial fat thickness and retro-mammary fat measurements, indicating substantial inter-individual variability.

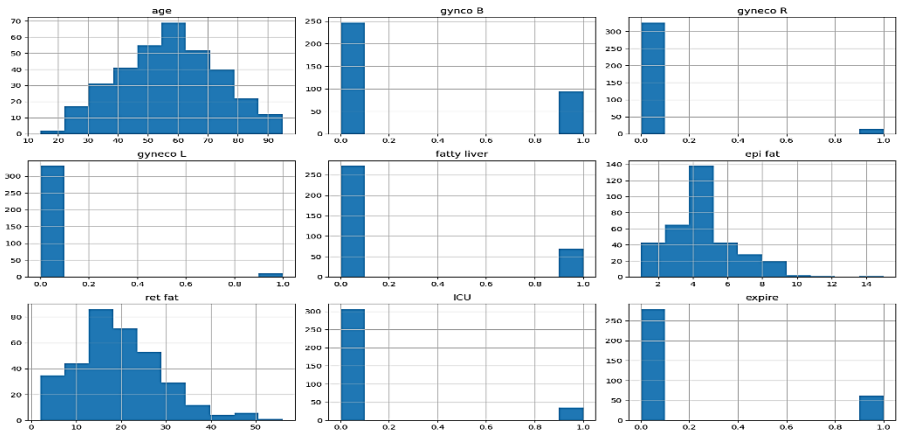

Fig. 1. Histograms showing the distribution of all study variables, including age, epicardial fat thickness, retro-mammary fat thickness, gynecomastia status, fatty liver, ICU admission, and mortality.

The correlation structure between variables is illustrated in *Fig. 2*, where moderate associations were observed among adiposity-related features. In contrast, gynecomastia showed weak correlations with metabolic imaging biomarkers.

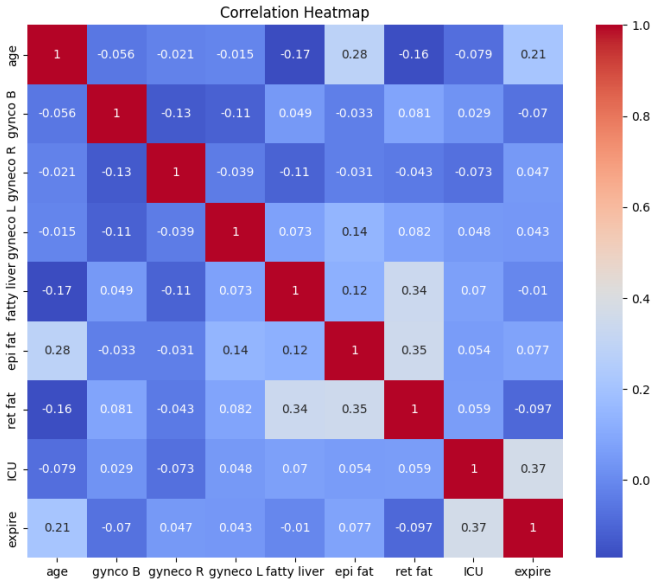

Fig. 2. Correlation heatmap illustrating pairwise correlations among demographic, imaging-derived, and clinical variables.

Boxplot analysis presented in Figure 3 revealed the presence of outliers in epicardial fat and retro-mammary fat thickness, reflecting variability in visceral adiposity distribution.

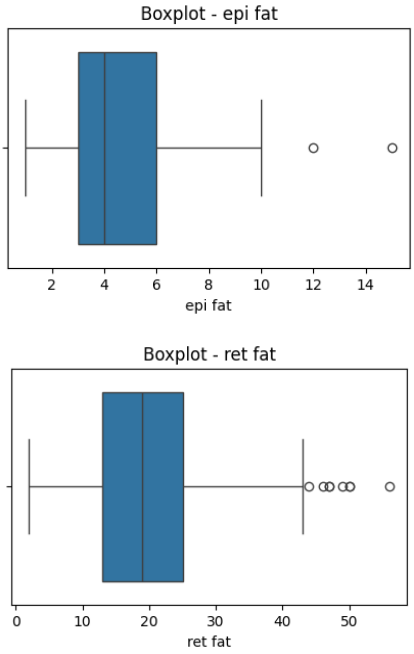

Fig. 3. Boxplots of continuous imaging-derived biomarkers showing data dispersion and potential outliers.

3.2 | Class Imbalance and SMOTE Resampling

Given the significant imbalance in ICU outcomes, SMOTE was only applied to training dataset.

As demonstrated in *Fig. 4*, SMOTE successfully balanced the dataset, increasing the minority ICU class to match the majority class distribution. The proportional changes before and after resampling are further illustrated in *Fig. 5*.

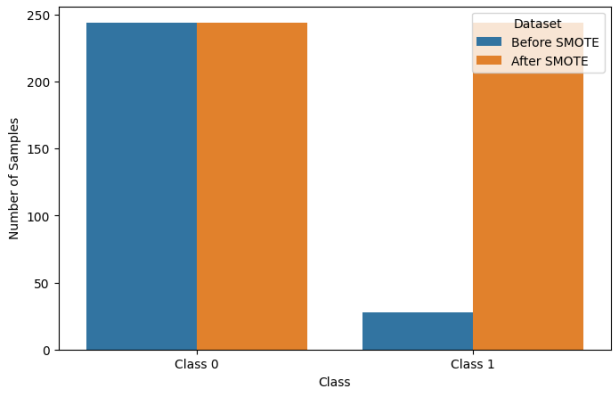

Fig. 4. Class distribution before and after SMOTE oversampling shown as bar charts.

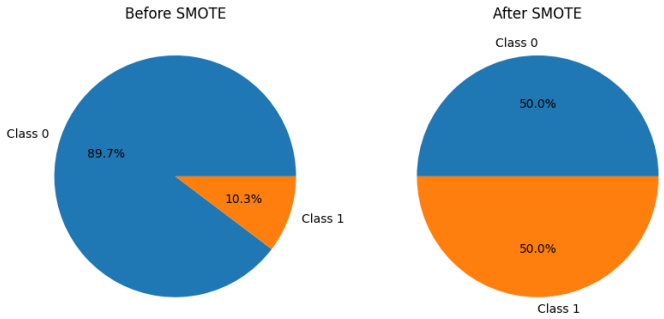

Fig. 5. Pie charts comparing ICU class proportions before and after SMOTE.

These transformations ensured improved model learning stability without introducing data leakage.

3.3 | Machine Learning Model Performance

Seven supervised ML models were trained and evaluated for ICU prediction.

The overall quantitative performance of all models is summarized in *Table 1*.

As shown in *Table 1*, the XGBoost classifier achieved the best overall performance in terms of accuracy (0.884), F1-score (0.333), and MCC (0.276), demonstrating superior predictive capability compared to other models.

ExtraTrees also showed strong performance, particularly in ROC-AUC (0.729), indicating robust discriminative ability.

Table 1. Performance comparison of ML models for ICU prediction.

Model	Accuracy	Precision	Recall	Specificity	F1	ROC-AUC	PR-AUC	MCC
ExtraTrees	0.855	0.286	0.286	0.919	0.286	0.729	0.239	0.205
XGBoost	0.884	0.400	0.286	0.952	0.333	0.604	0.185	0.276
RandomForest	0.797	0.000	0.000	0.887	0.000	0.689	0.179	-0.113
GradientBoosting	0.826	0.143	0.143	0.903	0.143	0.671	0.188	0.046
SVC	0.667	0.167	0.571	0.677	0.258	0.677	0.282	0.158
DecisionTree	0.812	0.200	0.286	0.871	0.235	0.578	0.130	0.134
Logistic Regression	0.594	0.138	0.571	0.597	0.222	0.491	0.110	0.103

3.4 | Model Interpretability and Feature Importance

To improve interpretability, feature importance analysis and SHAP-based explanations were applied.

As shown in *Fig. 6*, ExtraTrees feature importance analysis identified epicardial fat thickness, age, and retro-mammary fat as the most influential predictors.

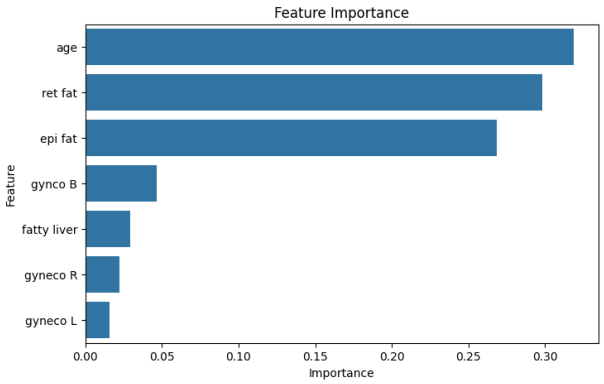

Fig. 6. Feature importance ranking generated from the best-performing ensemble model.

SHAP summary analysis presented in *Fig. 7* confirmed that epicardial fat and hepatic steatosis had the strongest global impact on ICU prediction.

Global feature contribution based on SHAP values is further illustrated in *Fig. 7*, reinforcing the dominant role of visceral adiposity markers in disease severity prediction.

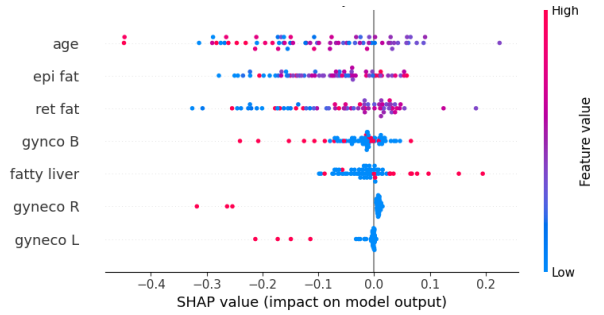

Fig. 7. SHAP summary plot demonstrating the overall contribution of each feature to ICU prediction.

3.5 | Deep Learning Model Performance

A feedforward neural network was trained to evaluate nonlinear deep learning performance.

As shown in *Fig. 8*, the deep learning model achieved moderate discriminative performance, with lower ROC-AUC compared to tree-based ensemble models. Despite acceptable training accuracy, generalization performance remained limited, likely due to small dataset size and class imbalance.

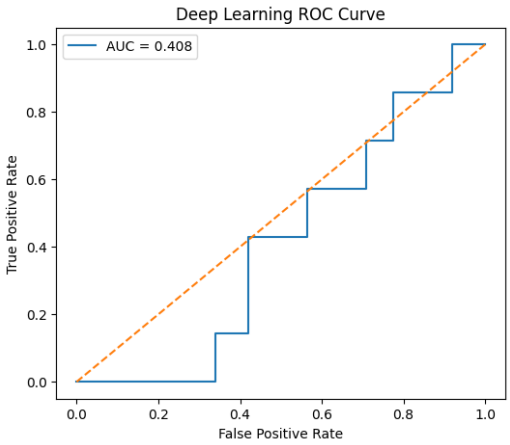

Fig. 8. Receiver Operating Characteristic (ROC) curve of the deep learning model for ICU prediction.

4 | Discussion

This study investigated the predictive value of CT-derived biomarkers—including gynecomastia, hepatic steatosis, and epicardial fat thickness—for ICU admission among hospitalized COVID-19 patients using ML and deep learning approaches. The main findings indicate that 1) visceral adiposity markers, particularly epicardial fat thickness, are strongly associated with disease severity, 2) ensemble learning models outperform classical linear and deep learning models in this dataset, and 3) CT-based incidental findings can be effectively integrated into predictive clinical decision-support systems.

4.1 | Principal Findings and Clinical Interpretation

The most important finding of this study is the consistent predictive contribution of epicardial fat thickness across all top-performing models, especially XGBoost and ExtraTrees. This aligns with previous evidence suggesting that epicardial fat acts as an active metabolic and inflammatory tissue, closely associated with cardiometabolic risk and systemic inflammation [5], [6].

Similarly, hepatic steatosis showed moderate predictive relevance. Prior studies have demonstrated that Non-Alcoholic Fatty Liver Disease (NAFLD) is associated with systemic metabolic dysregulation and increased inflammatory burden, both of which may worsen COVID-19 outcomes [2], [9].

Although gynecomastia was not among the top-ranked predictors in feature importance plots, SHAP analysis indicated a non-linear and context-dependent contribution. This supports previous findings suggesting that gynecomastia may reflect underlying endocrine and metabolic imbalance rather than acting as an isolated pathological entity [4], [10].

4.2 | Machine Learning Performance Interpretation

Among all evaluated models, XGBoost achieved the best overall performance, particularly in accuracy (0.884) and F1-score (0.333), followed by ExtraTrees with the highest ROC-AUC (0.729). This superiority of gradient-boosting and ensemble tree-based models is consistent with prior studies in medical tabular datasets, where nonlinear interactions and feature heterogeneity dominate model structure [7].

The better performance of tree-based methods compared to Logistic Regression and SVM suggests that the relationships between CT-derived biomarkers and ICU admission are highly nonlinear and interaction dependent.

The relatively poor performance of Logistic Regression (ROC-AUC $\approx$ 0.49) confirms that linear assumptions are insufficient for modeling such complex physiological interactions.

4.3 | Deep Learning Performance and Limitations

The deep learning model demonstrated lower generalization performance (ROC-AUC ≈ 0.41–0.53 across runs) despite moderate training accuracy. This is likely due to:

I. Limited dataset size (n = 341), which is insufficient for stable neural network training

II. Class imbalance, even after SMOTE

III. Tabular data structure, which is typically less suitable for deep learning compared to tree-based models

These findings are consistent with previous literature showing that deep learning does not necessarily outperform gradient boosting methods in structured clinical datasets [11].

4.4 | Role of SMOTE and Class Imbalance

The ICU outcome was highly imbalanced (approximately 10% positive cases). Application of SMOTE improved minority class representation during training. However, as shown in confusion matrices (*Fig. 8*), sensitivity for ICU prediction remained limited across most models.

This suggests that while SMOTE improves learning balance, it does not fully resolve the inherent difficulty of predicting rare clinical outcomes, particularly in small datasets.

4.5 | Explainability and Clinical Relevance (SHAP Analysis)

SHAP-based interpretation provided important insights into model decision-making. Epicardial fat thickness, age, and hepatic steatosis consistently ranked among the most influential predictors.

The SHAP summary plot demonstrated that higher epicardial fat values were associated with increased ICU risk, supporting the hypothesis that visceral adiposity contributes to COVID-19 severity through inflammatory and cardiometabolic pathways.

Feature importance analysis from tree-based models confirmed simillar ranking patterns, reinforcing model interpretability consistency.

4.6 | Comparison with Previous Studies

The observed associations are consistent with previous CT-based COVID-19 severity studies. For example, Aksoy et al. [10] reported correlations between hepatic density, gynecomastia, and adiposity-related CT biomarkers in COVID-19 patients, suggesting shared metabolic pathways.

Similarly, Aslan et al. [3] and Kazemi et al. [12] demonstrated that incidental gynecomastia is common in COVID-19 CT populations, although its prognostic value remains unclear.

Our findings extend these results by integrating multiple CT-derived biomarkers into a unified ML framework for ICU prediction.

4.7 | Clinical Implications

This study highlights the potential of incidental CT findings as predictive biomarkers in COVID-19 triage. Since CT scans are already widely performed in hospitalized patients, extracting additional quantitative biomarkers such as epicardial fat and liver attenuation can provide:

- *Early risk stratification*
- *Improved ICU resource allocation*
- *Objective support for clinical decision-making*

The integration of XAI methods (SHAP) further enhances clinical interpretability, which is essential for adoption in medical settings.

5 | Conclusion

CT-derived metabolic biomarkers, particularly epicardial fat thickness and hepatic steatosis, are valuable predictors of ICU admission in COVID-19 patients. ML models, especially XGBoost and ExtraTrees, demonstrated strong predictive performance, while deep learning showed limited effectiveness due to dataset constraints. The integration of XAI enhances the clinical interpretability of predictive models and supports their potential use in precision medicine. This study has several limitations:

I. Retrospective single-center design, limiting generalizability
II. Small sample size (n = 341), restricting deep learning performance
III. Lack of external validation dataset
IV. Potential confounding factors such as medications and comorbidities not fully integrated
V. Use of CT-derived surrogate markers rather than histopathological confirmation.

5.1 | Future Directions

I. Multi-center external validation
II. Integration of clinical laboratory biomarkers with imaging data
III. Use of larger datasets for deep learning optimization
IV. Prospective validation in real-world ICU triage settings

References

[1] Wynants, L., Van Calster, B., Collins, G. S., Riley, R. D., Heinze, G., Schuit, E., … ., & van Smeden, M. (2020). Prediction models for diagnosis and prognosis of covid-19: Systematic review and critical appraisal. *Bmj*, *369*. https://doi.org/10.1136/bmj.m1328

[2] Boyce, C. J., Pickhardt, P. J., Kim, D. H., Taylor, A. J., Winter, T. C., Bruce, R. J., … ., & Hinshaw, J. L. (2010). Hepatic steatosis (fatty liver disease) in asymptomatic adults identified by unenhanced low-dose CT. *American Journal of Roentgenology*, *194*(3), 623–628. https://doi.org/10.2214/AJR.09.2590

[3] Aslan, Ö., Bayraktaroğlu, S., Çinkooğlu, A., Ceylan, N., Savaş, R., & Oktay, A. (2021). Prevalence of incidental gynecomastia by chest computed tomography in patients with a prediagnosis of COVID-19 pneumonia. *European Journal of Breast Health*, *17*(2), 173–179. https://doi.org/10.4274/ejbh.galenos.2021.6251

[4] Kim, M. S., Kim, J. H., Lee, K. H., & Suh, Y. J. (2020). Incidental gynecomastia on thoracic computed tomography in clinical practice: Characteristics, radiologic features, and correlation with possible causes in South Korean men. *American Journal of Men's Health*, *14*(3), 1557988320908102. https://doi.org/10.1177/1557988320908102

[5] Petta, S., Argano, C., Colomba, D., Cammà, C., Di Marco, V., Cabibi, D., … ., & Craxì, A. (2015). Epicardial fat, cardiac geometry and cardiac function in patients with non-alcoholic fatty liver disease: Association with the severity of liver disease. *Journal of Hepatology*, *62*(4), 928–933. https://doi.org/10.1016/j.jhep.2014.11.030

[6] Gorter, P. M., van Lindert, A. S. R., de Vos, A. M., Meijs, M. F. L., van der Graaf, Y., Doevendans, P. A., … ., & Visseren, F. L. J. (2008). Quantification of epicardial and peri-coronary fat using cardiac computed tomography; reproducibility and relation with obesity and metabolic syndrome in patients suspected of coronary artery disease. *Atherosclerosis*, *197*(2), 896–903. https://doi.org/10.1016/j.atherosclerosis.2007.08.016

[7] Chen, T., & Guestrin, C. (2016). XGBoost: A scalable tree boosting system. *Proceedings of the 22nd ACM Sigkdd International Conference on Knowledge Discovery and Data Mining* (pp. 785–794). New York, NY, USA: Association for Computing Machinery. https://doi.org/10.1145/2939672.2939785

[8] Lundberg, S. M., & Lee, S.-I. (2017). A unified approach to interpreting model predictions. *Advances in Neural Information Processing Systems* (Vol. 30). Curran Associates, Inc. https://proceedings.neurips.cc/paper_files/paper/2017/file/8a20a8621978632d76c43dfd28b67767-Paper.pdf

[9] Browning, J. D., Szczepaniak, L. S., Dobbins, R., Nuremberg, P., Horton, J. D., Cohen, J. C., … ., & Hobbs, H. H. (2004). Prevalence of hepatic steatosis in an urban population in the United States: Impact of ethnicity. *Hepatology*, *40*(6), 1387–1395. https://onlinelibrary.wiley.com/doi/pdf/10.1002/hep.20466

[10] Aksoy, S. H., Aydin, S., Yurdaisik, I., & Dalkilic, O. (2021). *May hepatic steatosis be associated with gynecomastia and epicardial fat? A retrospective study of 599 male patients*. https://europepmc.org/article/ppr/ppr420370

[11] Shwartz-Ziv, R., & Armon, A. (2022). Tabular data: Deep learning is not all you need. *Information Fusion*, *81*, 84–90. https://doi.org/10.1016/j.inffus.2021.11.011

[12] Kazemi, M. A., Hemmati, S., Moradi, B., Sharifian, H., Firooznia, K., & Nikravangolsefid, N. (2020). COVID-19 and gynecomastia: What is the relationship? *Academic Radiology*, 27(8), 1188. https://doi.org/10.1016/j.acra.2020.05.003