



## Research article

# Predicting survival of Iranian COVID-19 patients infected by various variants including omicron from CT Scan images and clinical data using deep neural networks

Mahyar Ghafoori<sup>a</sup>, Mehrab Hamidi<sup>c,d</sup>, Rassa Ghavami Modegh<sup>b,c,d</sup>,  
Alireza Aziz-Ahari<sup>a</sup>, Neda Heydari<sup>a</sup>, Zeynab Tavafizadeh<sup>a</sup>, Omid Pournik<sup>a</sup>,  
Sasan Emdadi<sup>d</sup>, Saeed Samimi<sup>d</sup>, Amir Mohseni<sup>d,c</sup>, Mohammadreza Khaleghi<sup>a</sup>,  
Hamed Dashti<sup>d</sup>, Hamid R. Rabiee<sup>b,c,d,\*</sup>

<sup>a</sup> Radiology Department, Hazrat Rasoul Akram Hospital, School of Medicine, Iran University of Medical Sciences, Hemmat, Tehran, 14535, Iran

<sup>b</sup> Data science and Machine learning Lab, Department of Computer Engineering, Sharif University of Technology, Azadi, Tehran, 11365-8639, Iran

<sup>c</sup> BCB Lab, Department of Computer Engineering, Sharif University of Technology, Azadi, Tehran, 11365-8639, Iran

<sup>d</sup> AI-Med Group, AI Innovation Center, Sharif University of Technology, Azadi, Tehran, 11365-8639, Iran

## ARTICLE INFO

## Keywords:

Computer-aided diagnosis  
Computerized tomography imaging  
Deep neural networks  
Omicron variant  
Multimodal analysis

## ABSTRACT

**Purpose:** The rapid spread of the COVID-19 omicron variant virus has resulted in an overload of hospitals around the globe. As a result, many patients are deprived of hospital facilities, increasing mortality rates. Therefore, mortality rates can be reduced by efficiently assigning facilities to higher-risk patients. Therefore, it is crucial to estimate patients' survival probability based on their conditions at the time of admission so that the minimum required facilities can be provided, allowing more opportunities to be available for those who need them. Although radiologic findings in chest computerized tomography scans show various patterns, considering the individual risk factors and other underlying diseases, it is difficult to predict patient prognosis through routine clinical or statistical analysis.

**Method:** In this study, a deep neural network model is proposed for predicting survival based on simple clinical features, blood tests, axial computerized tomography scan images of lungs, and the patients' planned treatment. The model's architecture combines a Convolutional Neural Network and a Long Short Term Memory network. The model was trained using 390 survivors and 108 deceased patients from the Rasoul Akram Hospital and evaluated 109 surviving and 36 deceased patients infected by the omicron variant.

**Results:** The proposed model reached an accuracy of 87.5% on the test data, indicating survival prediction possibility. The accuracy was significantly higher than the accuracy achieved by classical machine learning methods without considering computerized tomography scan images ( $p$ -value  $\leq 4E-5$ ). The images were also replaced with hand-crafted features related to the ratio of infected lung lobes used in classical machine-learning models. The highest-performing model reached an accuracy of 84.5%, which was considerably higher than the models trained on mere

\* Corresponding author at: Data science and Machine learning Lab, Department of Computer Engineering, Sharif University of Technology, Azadi, Tehran, 11365-8639, Iran.

*E-mail addresses:* [mahyarghafoori@gmail.com](mailto:mahyarghafoori@gmail.com) (M. Ghafoori), [mehrabhamidi78@gmail.com](mailto:mehrabhamidi78@gmail.com) (M. Hamidi), [ghavami@ce.sharif.edu](mailto:ghavami@ce.sharif.edu) (R.G. Modegh), [azizahari.ar@iums.ac.ir](mailto:azizahari.ar@iums.ac.ir) (A. Aziz-Ahari), [Nedaheydari90@yahoo.com](mailto:Nedaheydari90@yahoo.com) (N. Heydari), [zaynabtavafi@gmail.com](mailto:zaynabtavafi@gmail.com) (Z. Tavafizadeh), [pournik.o@iums.ac.ir](mailto:pournik.o@iums.ac.ir) (O. Pournik), [sasan.emdadi@gmail.com](mailto:sasan.emdadi@gmail.com) (S. Emdadi), [sadsamm95@gmail.com](mailto:sadsamm95@gmail.com) (S. Samimi), [amir.mohseni@sharif.edu](mailto:amir.mohseni@sharif.edu) (A. Mohseni), [Mkh712003@yahoo.com](mailto:Mkh712003@yahoo.com) (M. Khaleghi), [dashtih@ce.sharif.edu](mailto:dashtih@ce.sharif.edu) (H. Dashti), [rabiee@sharif.edu](mailto:rabiee@sharif.edu) (H.R. Rabiee).

<https://doi.org/10.1016/j.heliyon.2023.e21965>

Received 12 October 2022; Received in revised form 26 October 2023; Accepted 1 November 2023

Available online 8 November 2023

2405-8440/© 2023 The Author(s). Published by Elsevier Ltd. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

clinical information (p-value  $\leq 0.006$ ). However, the performance was still significantly less than the deep model (p-value  $\leq 0.016$ ).

**Conclusion:** The proposed deep model achieved a higher accuracy than classical machine learning methods trained on features other than computerized tomography scan images. This proves the images contain extra information. Meanwhile, Artificial Intelligence methods with multimodal inputs can be more reliable and accurate than computerized tomography severity scores.

## 1. Introduction

In late 2021, the Omicron variant of Coronavirus emerged in South Africa, leading to the spread of the disease to a wide range of countries. This unknown virus affected a significant population, overwhelming hospitals that lacked sufficient capacity to accommodate the high number of patients. Consequently, individuals at higher risk were unable to access hospital beds, Intensive Care Unit (ICU) rooms, or ventilator support, increasing mortality rates. However, due to vaccination efforts, many individuals are now considered outpatients and do not require hospitalization for treatment. Hospitals need a reliable method to determine which patients need hospitalization to optimize resources and reduce the mortality rate.

COVID-19 is not just a pulmonary infection. It can affect multiple organs, including the central or peripheral nervous system [1,2], cardiovascular system [3], and cutaneous manifestation [4]. There are various symptoms reported, including fever in 85.6%-88.7% of patients [5,6] to rare ones like viral myositis [7]. Also, radiologic findings and lung computerized tomography (CT) scan patterns vary from case to case, ranging from ground-glass opacity with the highest prevalence [5,8,9] to the rare ones like pneumothorax, or pneumomediastinum [10]. The ultimate outcome of COVID-19 is influenced by several factors, including individual risk factors, and underlying medical conditions.

COVID-19 patients demonstrate divergent clinical trajectories, characterized by a spectrum of severity in their prognostic outlooks. There is no single clinical, laboratory, imaging, or routine statistical analysis that can provide a reliable assessment of a patient's prognosis. The complex network interactions between these factors can cause clinicians to allocate resources in an inappropriate manner. Higher-risk patients can receive more intensive care if a fast and efficient method for this prediction is developed, while lower-risk patients may be managed outpatient or with minimal inpatient care. As a result of this method, hospital and health resource utilization will be maximized, and the number of mortalities resulting from the lack of required facilities will be reduced.

Artificial intelligence (AI) has made tremendous progress in recent years. AI-based systems have excelled over humans in many fields, such as games, and achieved competitive performance in others. Consequently, AI-based techniques have been used in a variety of real-life applications, such as autonomous vehicles and medical analysis. The machines are fast, fatigue-free, and produce consistent results regardless of the time of the day. In pandemics, they may provide assistance in reducing the load and managing the conditions [11–13]. Many studies have concentrated on utilizing AI-driven systems during the pandemic for different objectives. Several studies have focused on distinguishing COVID-19 patients from patients with other pulmonary diseases and healthy people from computerized tomography (CT) scan images of the chest [14–16,13,17–19]. Some of the introduced models can also calculate the percentage of infected regions of the lungs [13]. Another group of studies has focused on diagnosing COVID-19 from chest X-ray images to be utilized as a screening stage [10,20–29]. Other than single-person-based analysis, some studies have focused on large-scale predictions using the time series to help in large-scale management decisions, e.g., the number of COVID-19 cases and the mortality rate of COVID-19 [30–33]. AI was also used in other applications, such as vaccine development and drug discovery for COVID-19 disease [34].

Recent studies have used AI to predict patients' outcomes using different features. In [35], they tried to predict short-term outcomes (favorable or adverse) from clinical information, including sex, age, WBC, Lymph, etc., into the volume of infected lungs of patients from their CT scans. They achieved 88% accuracy, 90% sensitivity, and 87% specificity by applying a support vector machine (SVM) algorithm on 106 patients. In [36], Li and his colleagues tried to predict COVID-19 patient mortality using clinical symptoms and CT scan images of 98 older patients. They achieved 87.5% sensitivity and 70.6% specificity. Furthermore, CT scans and their extracted features, clinical information, serology, and hematology tests of patients demonstrated promising results in predicting adverse outcomes in COVID-19 patients [37–39]. In previous studies, various methods have been used to predict mortality and hospitalization of patients, including the CT-Severity Index [40–42]. One of the drawbacks of this model is the unequal volume of the pulmonary lobes despite the equal pulmonary involvement score. Another problem in these studies is the prominent role of the interpreter. The radiologists' experience and accuracy are essential in determining the score. These factors may explain why these scoring and predictive systems are less commonly used in clinical practice.

Previous studies have addressed the question of predicting survival, however no studies have predicted survival in a particular condition. In this study, we attempted to predict outcomes based on patient health conditions at the time of hospitalization and specific treatment conditions (number of hospitalization days, number of ICU days, and number of ventilator days). Thus, it is possible to identify the minimum needed facilities for the patient to recover, or if even the best facilities are not sufficient to ensure survival. Axial CT scan images of the patient's lung were used along with clinical characteristics, including age and sex, and simple blood tests containing WBC, Hb, Plt, X.Neut, and X.Lymph. Using the above features, a deep network model was developed to predict the probability of survival. We also compared the results of the proposed method with simple machine learning methods trained on all the features of CT scan images to demonstrate their added value.

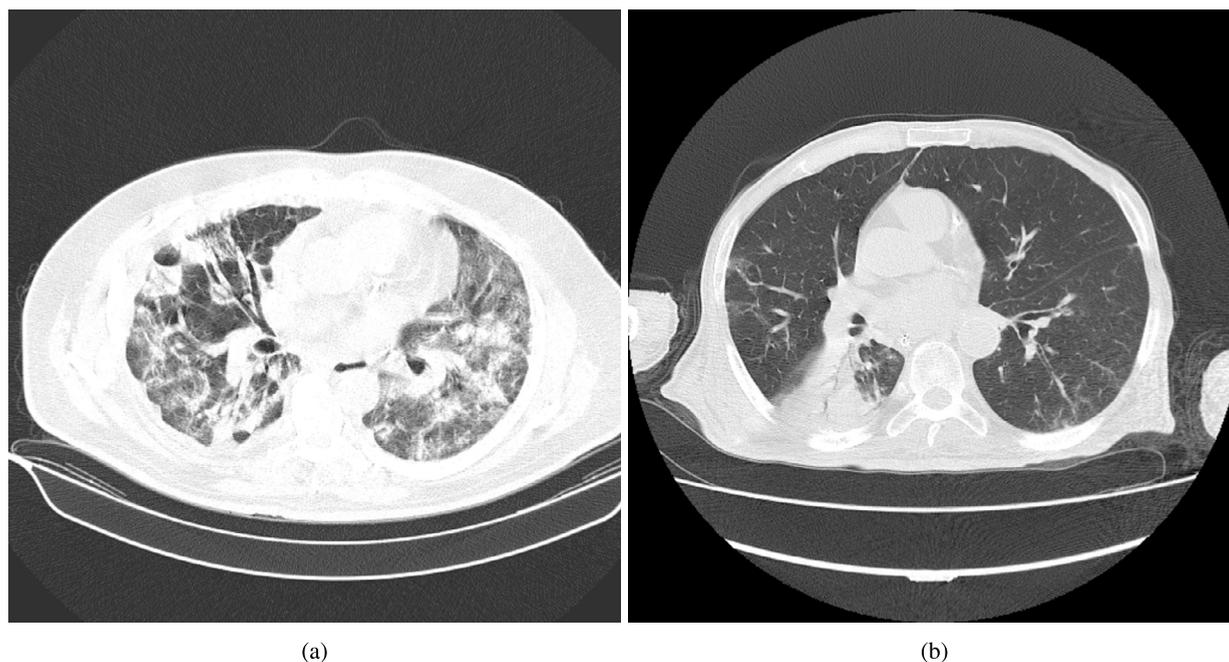


Fig. 1. Two middle slices from the CT scan of lung of (a) a survivor infected with the Omicron variant (b) a patient who passed away after 4 days.

## 2. Materials and method

### 2.1. Dataset

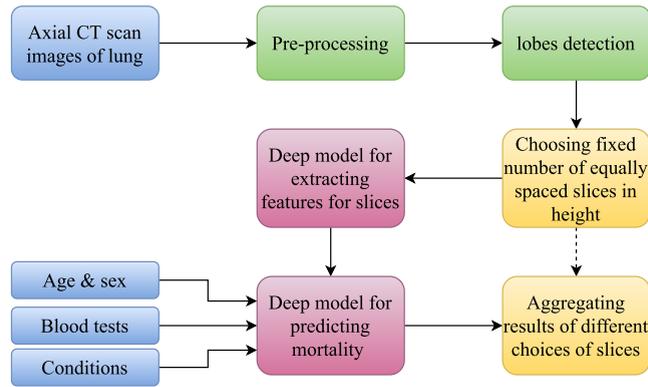
A total of 640 patients (506 survived and 134 died) were studied at the Rasoul Akram Hospital in Tehran. We used 430 samples (336 survivors) for training, 68 samples (54 survivors) for validation, and 142 samples (116 survivors) for testing. Fig. 1a and Fig. 1b show CT scan images of two patients from this dataset. In addition to CT scan images of the patients, clinical information including age, sex, WBC, Hb, Plt, X.Neut, and X.Lymph, as well as treatment conditions including the number of hospitalization days, the number of days in the intensive care unit, and receiving ventilator support, were gathered from the medical records. The following data was collected for all patients: age, sex, WBC, Hb, Plt, and treatment conditions (X.Neut and X.Lymph were not recorded for less than 150 patients). A linear regression model was trained on fully known data in order to predict missing information. Training was conducted using the completed information.

### 2.2. Workflow

By analyzing clinical data such as age and gender of patients, blood test results such as WBC, Hb, Plt, X.Neut, and X.LYmph, as well as axial lung CT scan images, the deep learning pipeline in Fig. 2 was used to predict the mortality status of COVID-19 patients upon arrival at the hospital. The model takes into account the conditions under which the patient is to be treated and predicts their mortality status. The number of hospitalization days, the number of days in the ICU, and the number of days receiving ventilator support are among these conditions. In the first step of the pipeline, CT scan images are preprocessed, and lobes are detected and extracted from images of slices using the method described in [13]. In order to obtain a general understanding of the condition of the lungs, a fixed number of slices is selected along with a sample height to cover different lung sizes with nearly equal distances. To extract features for each of the slices, images of the left and right lobes are fed independently into a deep convolutional neural network. In order to predict the patient's mortality, the extracted features, clinical features, blood test results, and treatment conditions are fed into another deep model. Finally, the results of different packs of slices are aggregated, and the highest probability of mortality in each group is assigned to the entire sample. In the preceding subsections, each part of the pipeline is described in detail.

### 2.3. Choosing pack of slices for each sample

This step involves selecting a fixed number of evenly spaced slices along the lung's height for each sample. As a result of this sampling, GPU RAM is used less during training, allowing for larger batches to be used. Additionally, it ensures that the conditions for samples with different numbers of slices are equal, since the subsequent slices in the batch will have a nearly equal distance from one another. The use of different slices for each sample also prevents overfitting due to the same observations being made for each sample.



**Fig. 2.** Flowchart of the pipeline for predicting survival of COVID-19 patients. The pipeline uses clinical information (age, sex), blood test results, and axial lung computerized tomography scan images as inputs. It also considers additional conditions such as hospitalization days, ICU days, and ventilator support. Computerized tomography scan images are preprocessed, and lobes are extracted. A fixed number of slices is chosen to cover different lung sizes. Deep convolutional neural networks are used to extract features from the left and right lobes of the chosen slices. Another deep model is then used to predict patient mortality using these features, clinical information, and treatment conditions. The results from different packs of slices are aggregated, and the maximum probability of mortality in the groups is assigned to the entire sample.

Some CT-scan images may contain fewer slices than the number selected. The problem is fixed by repeating some sample slices until they become equal to the selected number. By choosing this method over interpolating, artificial effects were avoided. Using the collection of points calculated by Equation (1), all slices are divided into partitions of nearly equal length. In this equation,  $n_{sample\_slices}$  refers to the number of slices of the sample,  $n_{chosen\_slices}$  refers to the fixed number of slices chosen, and  $seq$  represents the sequence of numbers from the beginning to the end.

$$Points = Rounded\left\{\frac{n_{sample\_slices}}{n_{chosen\_slices}} \times seq(start = 0, end = n_{chosen\_slices})\right\} \quad (1)$$

Each pack of slices is selected by adding an offset to the reference set of points. The offset can range from zero to the maximum partition length, which is the maximum difference between two consecutive points. The offset is selected at random during the training phase. During the test phase, all possible offsets are used to make the packs, and the mortality probability assigned by the network to the packs will be applied to the entire sample.

#### 2.4. Extracting features for images of lobes

Fig. 3 illustrates the network structure used to extract features from lobe images. It is a part of the architecture employed in [13] for predicting diseased lung CT-scan images. Three consecutive slices of images related to one lobe are received by the network as a  $256 \times 256 \times 3$  array. The middle slice is used to calculate the features. Images related to the previous and next slices provide additional information regarding the continuity of the white material in the middle slice. The inputs are fed into a convolutional subnetwork (Table 1) with an output size of  $32 \times 32 \times 256$ , corresponding to a  $32 \times 32$  mesh of neurons with 256 features for each. In this network, each neuron's receptive field corresponds to a  $36 \times 36$  patch in the input image, so the extracted features for each neuron are related to the  $36 \times 36$  patch in the input image that the neuron is viewing. To add additional information from the vicinity of each patch, the output of the previous subnetwork is fed into a U-Net-style encoder-decoder (Table 1), in which features are extracted from a larger receptive field in the encoder part and the decoder part can increase the resolution to the input size. The distance between each patch and the lung peripheral is calculated as the minimum Manhattan distance between each pixel in the patch and the nearest peripheral pixel using the Breadth-First Search (BFS) algorithm. In order to determine the final features for each patch, we concatenate the features extracted for each patch, those extracted from its vicinity, and the minimum distance of each patch from the lung peripheral. The features for each patch are then fed into a fully connected subnetwork (Table 1) to calculate a bounded attention weight between 0 and 1 for that patch. Based on the computed weights, the final features for the lobe are calculated by using the weighted average of the patches in the lobe. Considering that this network operates only on patches, the final decision can easily be mapped to the responsible patches, and the logic can be readily understood. The network is trained in conjunction with the model for predicting mortality described in section 2.5. Section 2.6 discusses the training procedure in detail.

#### 2.5. Predicting mortality

At this point, the features of the chosen slices' lobes have been calculated. As illustrated in Fig. 4, the left and right lobes of each slice are concatenated to the relative position of the slice in the lung's height, the clinical characteristics, the blood test results, and the conditions of interest for the patient's treatment. The expanded features of slices are fed sequentially to a slice iterator subnetwork (Table 2). According to the history of the observed slices, this subnetwork calculates the updated features based on the characteristics of subsequent slices. The updated features of the last slice are then fed into a fully connected subnetwork (Table 2) to determine the probability of the patient's mortality.

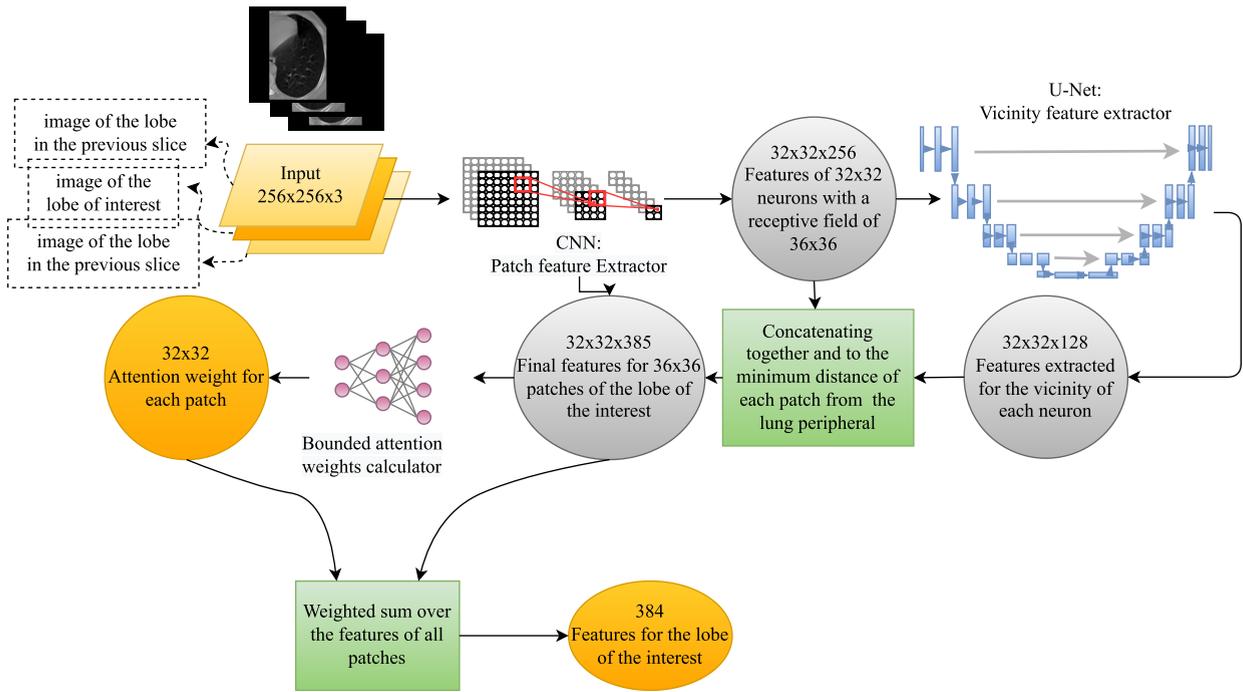


Fig. 3. The network structure used for extracting features from lobe images. The network calculates features for the middle slice after receiving images of one lobe in three consecutive slices. In order to extract features from a larger receptive field, a convolutional subnetwork and a U-Net encoder-decoder are applied. Additionally, the minimum distance between each patch and the lung peripheral is calculated. Final features are formed by concatenating the features from each patch, the surrounding vicinity, and the minimum distance. A bounded attention weight is then calculated for each patch based on these features. By taking the weighted average of the patch features, we determine the final features for the lobe. The network is trained in conjunction with a model for predicting mortality.

Table 1  
Layer specifications used in each subnetwork of the model for extracting features from lobe images.

Subnetworks	Layers
Patch feature extractor	conv(64x3x3,s=1), relu, conv(64x3x3,s=1), relu, MaxPool(2x2,s=2), [Name of the output: PFE1]; conv(128x3x3,s=1), relu, conv(128x3x3,s=1), relu, MaxPool(2x2,s=2) [Name of the output: PFE2]; conv(256x3x3,s=1), relu, conv(256x3x3,s=1), relu, MaxPool(2x2,s=2)
Vicinity feature extractor (Encoder)	conv(64x1x1,s=1) [Name of the output: VFE1]; conv(64x3x3,s=2), conv(64x3x3,s=1), relu [Name of the output: VFE2]; conv(64x3x3,s=2), conv(64x3x3,s=1), relu
Vicinity feature extractor (Decoder)	TransposeConv(64x2x2,s=2), conv(64x3x3,s=1), relu; UNet concatenation with VFE2; TransposeConv(64x2x2,s=2), conv(64x3x3,s=1), relu; UNet concatenation with VFE1
Bounded attention weights calculator	Conv(64x1x1,s=1), relu, Dropout(0.5), Conv(2x1x1,s=1), Softmax

Table 2  
Layer specifications for each subnetwork in the mortality prediction model.

Subnetworks	Layers
Slice Iterator	LSTM(hidden = 512, layers = 5, (with learnable initial hidden and memory), Dropout(0.5) after each layer), BatchNorm1d(512), elu, Dropout(0.5)
Mortality decider	Linear(256), BatchNorm1d(256), elu, Dropout(0.5), Linear(64), Linear(2), Softmax

For iterating over slices, this submodule uses a Long Short Term Memory (LSTM) model rather than a fully connected layer connecting all the slices' features. Fully connected layers would have had many more parameters, resulting in overfitting.

## 2.6. Training model

### 2.6.1. The loss function

Cross-entropy loss is commonly used for classification problems. Based on the predicted probability of mortality for each patient and the ground truth about the patient's status, the same loss was used as the primary loss. Deep networks have a known gradient flow problem, which increases with the depth of the network, and an LSTM network is considered a deep network because its depth

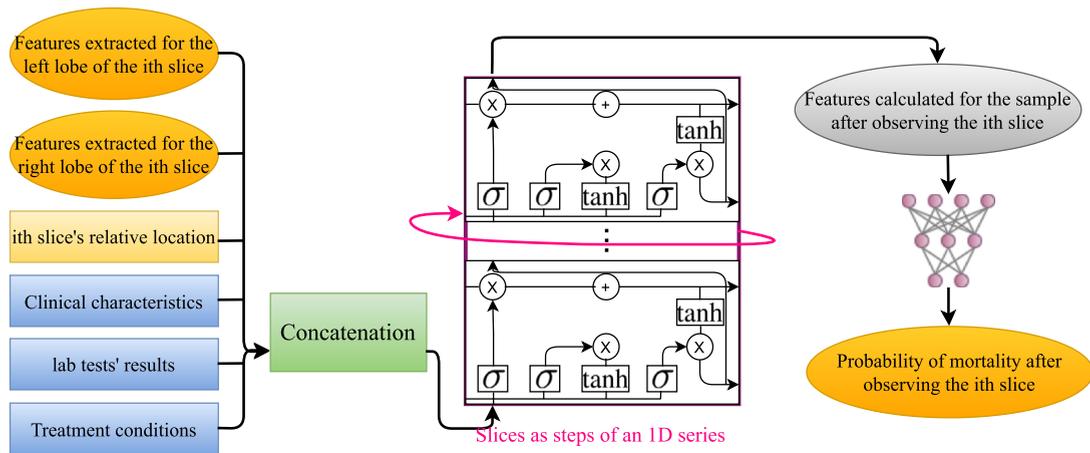


Fig. 4. The model used for predicting mortality for a specific treatment condition. The features extracted from the left and right lobes of the model along with clinical characteristics, blood test results, and treatment conditions are passed to a Long Short Term Memory (LSTM) network. Each time the LSTM network receives a slice from a computerized tomography scan image, it updates the features extracted from the image. To calculate the probability of mortality, the extracted features are passed to a multi-layer, fully connected neural network. The final decision is made after observing all slices.

affects the length of the sequence it receives. To solve the problem, an auxiliary loss was used to inject a gradient through the LSTM model's steps. Each slice's updated features were fed independently into the same decider network that was used for the last slice's features for calculating the auxiliary loss. For each of them, a mortality probability was calculated. The cross-entropy loss was then calculated using the ground truth label. We considered the auxiliary loss as the estimated likelihood for each slice and the average of each slice's losses. The LSTM model was also made to predict the mortality as soon as possible, not only based on all slices, which would result in a more robust prediction.

The loss calculated for the  $i$ th slice of the  $j$ th sample is illustrated in Equation (2). In this equation,  $Loss_i^j$  is the cross-entropy loss estimated based on the ground truth of the  $j$ th sample,  $GT^j$  and the probability of mortality determined by the network based on the sequence of the slices from the beginning till observing the  $i$ th slice,  $P_i^j$ .  $GT^j$  was considered 1 for mortality and 0 otherwise. The auxiliary loss,  $Aux\_Loss^j$ , was determined by averaging the cross-entropy losses of all the slices as in Equation (3). In this equation,  $n\_chosen\_slices$  stands for each batch sample's fixed number of slices. The primary loss was considered the cross-entropy loss of the last slice. The weighted sum of two loss terms calculated the total loss as in Equation (4). These factors were selected based on multiple trials. The total loss for each batch was calculated as the average loss of the samples in the batch.

$$Loss_i^j = -1 \times ((1 - GT^j) \times \ln(1 - P_i^j) + GT^j \times \ln(P_i^j)) \tag{2}$$

$$Aux\_Loss^j = \frac{1}{n\_chosen\_slices} \sum_{i=1}^{n\_chosen\_slices} Loss_i^j \tag{3}$$

$$Loss^j = Main\_Loss^j + 0.1 \times Aux\_Loss^j \tag{4}$$

### 2.6.2. Hyperparameters of training

In order to evaluate the model, a part of the dataset was separated as the test set. To help select hyperparameters and choose models, the remaining samples were also severed as the validation set. We trained the model using the remaining samples. The weights of the final model trained in [13] were utilized to initialize the model for extracting lobe features. During the training, the parameters of the model were frozen from the attention segment to prevent overfitting. In addition, the remaining parts of the model as well as the model used to calculate the probability of mortality were trained using the Adam optimization algorithm [43] with default parameters and an initial learning rate of  $1e-4$ . To prevent the model from becoming biased toward the class with more samples, batches of size 8 with four positive and four negative samples were used to train the model. For each sample, forty slices were selected. A number of trials have been conducted in order to determine this number. The model was trained for 500 epochs, and the epoch model with the highest accuracy for the validation data was selected as the final model.

To expand the network's data for training, the treatment conditions were randomized in each batch. Considering the availability of more facilities, the randomized treatment conditions were superior to the original treatment. If a patient survives at a specific level of illness, he will also survive at a higher level. Nevertheless, due to the limited number of facilities, the randomized conditions were worse than the original treatment, meaning the patient had died due to receiving the worst treatment possible.

### 2.7. Evaluation metrics

We used the most common metric to assess the performance of a binary classifier. The survival distribution is highly skewed towards the negative class in the problem of survival prediction. Models were evaluated using sensitivity (true positive rate) and specificity (true negative rate) to show error rates for each class. For the final model selection and comparison, accuracy was selected

**Table 3**  
Results of evaluating different models on test data.

	Accuracy (%)	Sensitivity (%)	Specificity (%)	AUC (%)
ML-F1	72.9	61.5	75.0	74.55
ML-F2	84.5	26.9	97.4	69.83
Deep Model	87.5	53.8	94.8	82.42

**Table 4**  
Machine learning methods and their hyperparameters.

Model	Set of the hyperparameters (The unspecified parameters are the defaults used by scikit-learn)
SVM	kernel: ['linear', 'poly', 'rbf', 'sigmoid']
MLP	sizes of hidden layers: [(i, j) for i in range(1, 202, 10) for j in range(2, 6)] activation: ['identity', 'logistic', 'tanh', 'relu'] solver: ['lbfgs', 'sgd', 'adam'] learning rate: ['constant', 'invscaling', 'adaptive']
Random Forest	max depth: range [1-24] criterion: ['gini', 'entropy'] N estimators: [20, 40, 60, 80, 100]
Logistic regression	C = [1.0, 3.6, 12.9, 46.4, 166.8, 599.5, 2154.3, 7742.6, 27825.6, 1.e05] penalty = ['l2', 'l1']

as a metric for representing the error rate in the distribution. Furthermore, Receiver Operating Characteristic (ROC) curves were computed to demonstrate the models' general performance.

### 3. Experimental results and discussion

According to previous reports, the predominant CT imaging pattern in COVID-19 was GGO with occasional consolidation. The majority of patients in the gathered dataset exhibit the same pattern [44]. Bilateral lung involvement, mainly in the lower lobes, has been reported in all studies. In this study, right lung involvement was slightly more prevalent than left lung involvement, which contradicts some other studies. Pleural effusion was also significantly higher in cases with a severe form of the disease [45,44].

#### 3.1. Evaluations

The results of the trained model are presented in Table 3. A test accuracy of 87.5% (sensitivity of 53.8% and specificity of 95.8%) indicates that survival can be predicted from CT scan images and simple blood tests. In order to assess the added value of CT scan images, the proposed model's results were compared with classical machine learning classifiers trained over all features except CT scan images (ML-F1). As extra information, the percentage of the infected volume of the whole lung and the left and the right lobes, the volume of the entire lung and the left and the right lobes, and the average percentage of infected volume in the top 20% and also in the top 4 lobes were calculated by the infection mask and the lung mask of the model trained in [13] and added to the clinical feature set (ML-F2). With these features, machine learning models such as SVM, perceptron, random forest, and logistic regression classifiers were trained. A 5-fold cross-validation was performed over both the training and validation sets in order to choose the hyperparameters for these models. In this study, scikit-learn [46] was utilized to implement models and set hyperparameters. The hyperparameters for these models are presented in Table 4. Within the first feature set (ML-F1), the perceptron model performed best in terms of accuracy (accuracy of 72.5%, sensitivity of 61.5%, and specificity of 75%). Among the models trained using the 2nd feature set (ML-F2), Random Forest had the highest test accuracy (84.5%), sensitivity (24.9%), and specificity (97.3%). The significance of the differences in error ratios between the proposed deep pipeline and each of the other models was calculated using McNemar's test [47], which is the only test with acceptable Type I error for algorithms that can only be executed once, like deep models [48]. The proposed deep framework had a significantly lower error ratio than the other models (p-value of 0.00004 for the ML-F1 model and p-value of 0.0156 for the ML-F2 model). The difference between the ML-F1 model and the ML-F2 model was also significant (p-value = 0.006). Significant improvements in accuracy can be attributed to information related to lung volume and infection percentage. Furthermore, the proposed deep pipeline reaches a higher accuracy than both models, suggesting that CT scan images may provide more information about mortality. The comparison results are shown in Table 6.

In order to demonstrate the validity of the model under the treatment conditions, the survivors should also survive under better treatment conditions, and the dead patients should die under worse treatments. The percentage of 4 random possible better conditions for the true negative patients that are also predicted as negative, as well as the percentage of 4 random possible worse conditions for the true positive patients which are also predicted as positive, is calculated for training, validation, and test sets, within the range of 0-60 days of hospitalization. This model almost always works correctly, except in some exceptional cases, as shown in Table 5.

**Table 5**  
Percentage of valid augmentations in true positive and true negative samples.

	Percentage of better conditions for true negative samples predicted as negative	Percentage of worse conditions for true positive samples predicted as positive
Train	100%	100%
Validation	100%	100%
Test	100%	99.03%

**Table 6**  
Comparison of current and previous studies.

Study	Size of Dataset	Accuracy	AUC
[49]	181	83.3%	0.756
[50]	5,766	84.4%	0.844
[51]	2,670,000	89.98%	0.93
[52]	383	87.1%	0.887
Current Study	640	87.5%	0.824

#### 4. Limitations of the study

This study has the following limitations:

- This study could benefit from more positive cases (the unsurvived patients). A larger dataset would have provided more information regarding the patterns that contribute to mortality. Because of the low mortality rate associated with COVID-19 disease in this study, collecting the dataset required more resources than were available.
- More details from the treatment process, such as the medicines used during hospitalization, in addition to the clinical histories of the patients, such as previous diseases and medicines taken for other ailments, would have resulted in a more accurate mortality prediction model. However, this study demonstrated that there were common patterns resulting in mortality within the set of data used.
- The results of this study would be enhanced by using a multi-center dataset, as this would ensure that they are not biased by the specific conditions at each site. Nevertheless, obtaining acceptable results from one center is a crucial step before investing resources in a larger multi-center study.

#### 5. Conclusion

Based on the conditions at the time of the patient's admission to the hospital, this study proposed a deep learning model to predict the survival and severity of COVID-19 patients in order to guide treatment options ranging from outpatient care to ICU admission and ventilator support. Based on age, sex, and simple blood tests of WBC, Plt, Hb, X.Neut, and X.Lymph with axial CT scans of the lung, the proposed model achieved an accuracy of 87.3%. This model had a specificity of 94.8%, which is more significant than sensitivity, which means that the patients who survived were correctly predicted. Thus, each patient's survival probability can be predicted based on all the possible conditions, and the minimum facilities required for survival can be determined. As a result, limited facilities will be used more efficiently, and fewer patients will die as a result of a lack of resources. In previous studies, machine learning models such as support vector machines and deep neural networks have been used to predict mortality using clinical signs such as fever and cough severity. When hospitals are overloaded with patients, predicting mortality is useful. As part of this study, we also included treatment conditions in order to predict mortality based on the planned treatment. By using the proposed model, hospitals can manage their facilities and use them for patients whose survival depends on them. The proposed model was also compared with classical machine learning models trained over all the features, CT scan images and even the classical models trained over all the mentioned features were compared, as well as some features related to lung volume and the percentage of infected volume calculated with another deep model from CT scan images. Although the classical models that used the features extracted from CT scan images as input reached a higher accuracy than the first group, indicating that the percentage of infection and the size of the lung lobes affect survival. As a result of the proposed deep model's higher accuracy, it is evident that lung CT scan images contain more information about survival. It is crucial to study CT scan images of lungs better, as there may be signs in the lungs of the patients that prevent them from receiving more intense treatment, which may result in a better treatment for COVID-19 patients. Although CT severity scores are used to predict a patient's condition, they are not accurate or sufficient to assess the patient's condition. Therefore, AI can improve the accuracy and reliability of mortality assessment methods when combined with other clinical inputs. Artificial intelligence systems would perform better, be more reliable, and be more generalizable if they used a larger dataset with vast diversity, included the patient's history of diseases, and performed extra clinical tests.

#### Ethical approval

The study has been approved by the Iran University of Medical Sciences (IUMS) and the Ethical Code IR.IUMS.REC.1399.008 authorized the use of CT-scan images. The patients consented to have their internal scans (e.g. x-rays, MRIs, CTs, ultrasound) images published.

## Abbreviations

Computerized tomography (CT) Long Short Term Memory (LSTM) Intensive Care Unit (ICU) Artificial Intelligence (AI) Area Under the Curve (AUC) Long Short Term Memory (LSTM) Support Vector Machine (SVM) Multi-Layer Perceptron (MLP)

## CRediT authorship contribution statement

**Mahyar Ghafoori:** Conceptualization, Data curation, Formal analysis, Methodology, Supervision. **Mehrab Hamidi:** Methodology, Validation, Writing – original draft. **Rassa Ghavami Modegh:** Formal analysis, Methodology, Validation, Writing – original draft. **Alireza Aziz-Ahari:** Data curation, Supervision. **Neda Heydari:** Data curation, Resources. **Zeynab Tavafizadeh:** Data curation, Resources. **Omid Pournik:** Conceptualization, Methodology, Supervision. **Sasan Emdadi:** Data curation. **Saeed Samimi:** Data curation. **Amir Mohseni:** Data curation. **Mohammadreza Khaleghi:** Data curation, Resources. **Hamed Dashti:** Conceptualization, Methodology, Writing – original draft. **Hamid R. Rabiee:** Conceptualization, Formal analysis, Methodology, Supervision, Writing – review & editing.

## Declaration of competing interest

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests:

Hamid R. Rabiee reports financial support was provided by Iran National Science Foundation. If there are other authors, they declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

## Data availability

The datasets associated with this study have not been deposited into a publicly available repository. However, these datasets are available from the corresponding authors upon reasonable request.

## Acknowledgements

This study was partially funded by IR National Science Foundation (INSF) Grant No. 96006077 and ISTI grant number 11/41701 (Hamid R. Rabiee was the recipient of both grants).

## References

- [1] L. Mao, M. Wang, S. Chen, Q. He, J. Chang, C. Hong, Y. Zhou, D. Wang, Y. Li, H. Jin, B. Hu, Neurological manifestations of hospitalized patients with COVID-19 in Wuhan, China: a retrospective case series study, 2020.
- [2] Y. Wu, X. Xu, Z. Chen, J. Duan, K. Hashimoto, L. Yang, C. Liu, C. Yang, Nervous system involvement after infection with Covid-19 and other coronaviruses, *Brain Behav. Immun.* 87 (2020) 18–22.
- [3] Y.-Y. Zheng, Y.-T. Ma, J.-Y. Zhang, X. Xie, Covid-19 and the cardiovascular system, *Nat. Rev. Cardiol.* 17 (2020) 259–260.
- [4] G. Casas, A. Catalá, G.C. Hernández, P. Rodríguez-Jiménez, D. Fernández-Nieto, A.R.-V. Lario, I.N. Fernández, R. Ruiz-Villaverde, D. Falkenhain-López, M.L. Velasco, J. García-Gavín, O. Baniandrés, C. González-Cruz, V. Morillas-Lahuerta, X. Cubiró, I.F. Nart, G. Selda-Enriquez, J. Romaní, X. Fustà-Novell, A. Melian-Olivera, M.R. Riesco, P. Burgos-Blasco, J.S. Ortigosa, M.F. Rodriguez, I. García-Doval, Classification of the cutaneous manifestations of COVID -19: a rapid prospective nationwide consensus study in Spain with 375 cases, *Br. J. Dermatol.* 183 (2020) 71–77.
- [5] A. Lovato, C. De Filippis, Clinical presentation of Covid-19: a systematic review focusing on upper airway symptoms, *Ear, Nose, Throat J.* 99 (2020) 569–576.
- [6] A.J. Rodriguez-Morales, J.A. Cardona-Ospina, E. Gutiérrez-Ocampo, R. Villamizar-Peña, Y. Holguin-Rivera, J.P. Escalera-Antezana, L.E. Alvarado-Arnez, D.K. Bonilla-Aldana, C. Franco-Paredes, A.F. Henao-Martinez, A. Paniz-Mondolfi, G.J. Lagos-Grisales, E. Ramírez-Vallejo, J.A. Suárez, L.I. Zambrano, W.E. Villamil-Gómez, G.J. Balbin-Ramon, A.A. Rabaan, H. Harapan, K. Dhama, H. Nishiura, H. Kataoka, T. Ahmad, R. Sah, Clinical, laboratory and imaging features of COVID-19: a systematic review and meta-analysis, *Trav. Med. Infect. Dis.* 34 (2020) 101623.
- [7] Q. Zhang, K.S. Shan, A. Minalyan, C. O'Sullivan, T. Nace, A rare presentation of coronavirus disease 2019 (Covid-19) induced viral myositis with subsequent rhabdomyolysis, *Cureus* 12 (2020).
- [8] S. Salehi, A. Abedi, S. Balakrishnan, A. Gholamrezaezhad, Coronavirus disease 2019 (Covid-19): a systematic review of imaging findings in 919 patients, *Am. J. Roentgenol.* 215 (2020) 87–93.
- [9] D. Caruso, M. Zerunian, M. Polici, F. Pucciarelli, T. Polidori, C. Rucci, G. Guido, B. Bracci, C.D. Dominici, A. Laghi, Chest CT features of COVID-19 in Rome, Italy, *Radiology* 296 (2020) E79–E85.
- [10] W. Wang, R. Gao, Y. Zheng, L. Jiang, COVID-19 with spontaneous pneumothorax, pneumomediastinum and subcutaneous emphysema, *J. Travel Med.* 27 (2020).
- [11] S.K. Zhou, H. Greenspan, C. Davatzikos, J.S. Duncan, B. Van Ginneken, A. Madabhushi, J.L. Prince, D. Rueckert, R.M. Summers, A review of deep learning in medical imaging: imaging traits, technology trends, case studies with progress highlights, and future promises, *Proc. IEEE* 109 (2021) 820–838.
- [12] R.G. Modegh, A. Salimi, S. Ilami, A.H. Dehqan, H. Dashti, S.H. Javanmard, H. Ghanaati, H.R. Rabiee, Covid-19 diagnosis with artificial intelligence, in: *The Science Behind the COVID Pandemic and Healthcare Technology Solutions*, Springer, 2022, pp. 353–378.
- [13] R.G. Modegh, M. Hamidi, S. Masoudian, A. Mohseni, H. Lotfalinezhad, M.A. Kazemi, B. Moradi, M. Ghafoori, O. Motamedi, O. Pournik, K. Rezaei-Kalantari, A. Manteghinezhad, S.H. Javanmard, F.A. Nezhad, A. Enhesari, M.S. Kheyrikhah, R. Eghtesadi, J. Azadbakht, A. Aliasgharzadeh, M.R. Sharif, A. Khaleghi, A. Foroutan, H. Ghanaati, H. Dashti, H.R. Rabiee, Accurate and rapid diagnosis of Covid-19 pneumonia with batch effect removal of chest ct-scans and interpretable artificial intelligence, 2020.
- [14] L. Li, L. Qin, Z. Xu, Y. Yin, X. Wang, B. Kong, J. Bai, Y. Lu, Z. Fang, Q. Song, et al., Using artificial intelligence to detect Covid-19 and community-acquired pneumonia based on pulmonary ct: evaluation of the diagnostic accuracy, *Radiology* 296 (2020) E65–E71.

- [15] H. Gunraj, A. Sabri, D. Koff, A. Wong, COVID-net ct-2: enhanced deep neural networks for detection of Covid-19 from chest ct images through bigger, more diverse learning, *Front. Med.* 8 (2022) 3126.
- [16] S.A. Harmon, T.H. Sanford, S. Xu, E.B. Turkbey, H. Roth, Z. Xu, D. Yang, A. Myronenko, V. Anderson, A. Amalou, et al., Artificial intelligence for the detection of Covid-19 pneumonia on chest ct using multinational datasets, *Nat. Commun.* 11 (2020) 4080.
- [17] X. Wang, X. Deng, Q. Fu, Q. Zhou, J. Feng, H. Ma, W. Liu, C. Zheng, A weakly-supervised framework for Covid-19 classification and lesion localization from chest ct, *IEEE Trans. Med. Imaging* 39 (2020) 2615–2625.
- [18] H. Gunraj, L. Wang, A. Wong, Covidnet-ct: a tailored deep convolutional neural network design for detection of Covid-19 cases from chest ct images, *Front. Med.* 7 (2020) 608525.
- [19] C. Jin, W. Chen, Y. Cao, Z. Xu, Z. Tan, X. Zhang, L. Deng, C. Zheng, J. Zhou, H. Shi, et al., Development and evaluation of an artificial intelligence system for Covid-19 diagnosis, *Nat. Commun.* 11 (2020) 5088.
- [20] D. Das, K. Santosh, U. Pal, Truncated inception net: Covid-19 outbreak screening using chest x-rays, *Phys. Eng. Sci. Med.* 43 (2020) 915–925.
- [21] S. Basu, S. Mitra, N. Saha, Deep learning for screening covid-19 using chest x-ray images, in: 2020 IEEE Symposium Series on Computational Intelligence (SSCI), IEEE, 2020, pp. 2521–2527.
- [22] X. Li, C. Li, D. Zhu, COVID-mobilexpert: on-device Covid-19 screening using snapshots of chest x-ray, 2020.
- [23] I. Castiglioni, D. Ippolito, M. Interlenghi, C.B. Monti, C. Salvatore, S. Schiaffino, A. Polidori, D. Gandola, C. Messa, F. Sardanelli, Artificial intelligence applied on chest x-ray can aid in the diagnosis of Covid-19 infection: a first experience from lombardy, Italy, *MedRxiv* 2020 (2020-04).
- [24] T. Hu, M. Khishe, M. Mohammadi, G.-R. Parvizi, S.H.T. Karim, T.A. Rashid, Real-time Covid-19 diagnosis from x-ray images using deep cnn and extreme learning machines stabilized by chimp optimization algorithm, *Biomed. Signal Process. Control* 68 (2021) 102764.
- [25] C. Wu, M. Khishe, M. Mohammadi, S.H. Taher Karim, T.A. Rashid, Evolving deep convolutional neural network by hybrid sine-cosine and extreme learning machine for real-time Covid19 diagnosis from x-ray images, *Soft Comput.* (2021) 1–20.
- [26] K. Shankar, S.N. Mohanty, K. Yadav, T. Gopalakrishnan, A.M. Elmisery, Automated Covid-19 diagnosis and classification using convolutional neural network with fusion based feature extraction model, *Cogn. Neurodyn.* (2021) 1–14.
- [27] D. Shome, T. Kar, S.N. Mohanty, P. Tiwari, K. Muhammad, A. Altameem, Y. Zhang, A.K.J. Saudagar, Covid-transformer: interpretable Covid-19 detection using vision transformer for healthcare, *Int. J. Environ. Res. Public Health* 18 (2021) 11086.
- [28] A. Saffari, M. Khishe, M. Mohammadi, A. Hussein Mohammed, S. Rashidi, Dcnn-fuzzzywoa: artificial intelligence solution for automatic detection of Covid-19 using x-ray images, *Comput. Intell. Neurosci.* (2022).
- [29] C. Cai, B. Gou, M. Khishe, M. Mohammadi, S. Rashidi, R. Moradpour, S. Mirjalili, Improved deep convolutional neural networks using chimp optimization algorithm for Covid19 diagnosis from the x-ray images, *Expert Syst. Appl.* 213 (2023) 119206.
- [30] S. Sah, B. Surendiran, R. Dhanalakshmi, S.N. Mohanty, F. Alenezi, K. Polat, Forecasting Covid-19 pandemic using prophet, arima, and hybrid stacked lstm-gru models in India, *Comput. Math. Methods Med.* 2022 (2022).
- [31] N. Sharma, S. Yadav, M. Mangla, A. Mohanty, S. Satpathy, S.N. Mohanty, T. Choudhury, Geospatial multivariate analysis of Covid-19: a global perspective, *GeoJournal* (2021) 1–15.
- [32] S. Dash, S. Chakravarty, S.N. Mohanty, C.R. Pattanaik, S. Jain, A deep learning method to forecast Covid-19 outbreak, *New Gener. Comput.* 39 (2021) 515–539.
- [33] S. Satpathy, M. Mangla, N. Sharma, H. Deshmukh, S. Mohanty, Predicting mortality rate and associated risks in Covid-19 patients, *Spat. Inf. Res.* 29 (2021) 455–464.
- [34] W. Hwang, W. Lei, N.M. Katritis, M. MacMahon, K. Chapman, N. Han, Current and prospective computational approaches and challenges for developing Covid-19 vaccines, *Adv. Drug Deliv. Rev.* 172 (2021) 249–274.
- [35] J. Matos, F. Paparo, I. Mussetto, L. Bacigalupo, A. Veneziano, S.P. Bernardi, E. Biscaldi, E. Melani, G. Antonucci, P. Cremonesi, M. Lattuada, A. Pilotto, E. Pontali, G.A. Rollandi, Evaluation of novel coronavirus disease (Covid-19) using quantitative lung CT and clinical data: prediction of short-term outcome, *Eur. Radiol. Exp.* 4 (2020).
- [36] Y. Li, Z. Yang, T. Ai, S. Wu, L. Xia, Association of “initial CT” findings with mortality in older patients with coronavirus disease 2019 (Covid-19), *Eur. Radiol.* 30 (2020) 6186–6193.
- [37] R. Zhang, H. Ouyang, L. Fu, S. Wang, J. Han, K. Huang, M. Jia, Q. Song, Z. Fu, CT features of SARS-CoV-2 pneumonia according to clinical presentation: a retrospective analysis of 120 consecutive patients from Wuhan city, *Eur. Radiol.* 30 (2020) 4417–4426.
- [38] M. Yuan, W. Yin, Z. Tao, W. Tan, Y. Hu, Association of radiologic findings with mortality of patients infected with 2019 novel coronavirus in Wuhan, China, *PLoS ONE* 15 (2020) e0230548.
- [39] D. Colombi, F.C. Bodini, M. Petrini, G. Maffi, N. Morelli, G. Milanese, M. Silva, N. Sverzellati, E. Michieletti, Well-aerated lung on admitting chest CT to predict adverse outcome in Covid-19 pneumonia, *Radiology* 296 (2020) E86–E96.
- [40] M. Francone, F. Iafra, G.M. Masci, S. Coco, F. Cilia, L. Manganaro, V. Panebianco, C. Andreoli, M.C. Colaiacomo, M.A. Zingaropoli, et al., Chest ct score in Covid-19 patients: correlation with disease severity and short-term prognosis, *Eur. Radiol.* 30 (2020) 6808–6817.
- [41] Z. Feng, Q. Yu, S. Yao, L. Luo, W. Zhou, X. Mao, J. Li, J. Duan, Z. Yan, M. Yang, et al., Early prediction of disease progression in Covid-19 pneumonia patients with chest ct and clinical characteristics, *Nat. Commun.* 11 (2020) 1–9.
- [42] A. Aziz-Ahari, M. Keyhanian, S. Mamishi, S. Mahmoudi, E.E. Bastani, F. Asadi, M. Khaleghi, Chest ct severity score: assessment of Covid-19 severity and short-term prognosis in hospitalized Iranian patients, *Wien. Med. Wochenschr.* 172 (2022) 77–83.
- [43] D.P. Kingma, J. Ba Adam, A method for stochastic optimization, *arXiv preprint arXiv:1412.6980*, 2014.
- [44] R. Yang, X. Li, H. Liu, Y. Zhen, X. Zhang, Q. Xiong, Y. Luo, C. Gao, W. Zeng, Chest CT severity score: an imaging tool for assessing severe Covid-19, *Radiology* 2 (2020) e200047.
- [45] D. Sun, H. Li, X.-X. Lu, H. Xiao, J. Ren, F.-R. Zhang, Z.-S. Liu, Clinical features of severe pediatric patients with coronavirus disease 2019 in Wuhan: a single center’s observational study, *World J. Pediatr.* 16 (2020) 251–259.
- [46] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, A. Müller, J. Nothman, G. Louppe, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. Passos, D. Cournapeau, M. Brucher, M. Perrot, Édouard duchesnay, scikit-learn: machine learning in python, *J. Mach. Learn. Res.* 12 (2011) 2825–2830.
- [47] Q. McNemar, Note on the sampling error of the difference between correlated proportions or percentages, *Psychometrika* 12 (1947) 153–157.
- [48] T.G. Dietterich, Approximate statistical tests for comparing supervised classification learning algorithms, *Neural Comput.* 10 (1998) 1895–1923.
- [49] J.S. Zhu, P. Ge, C. Jiang, Y. Zhang, X. Li, Z. Zhao, L. Zhang, T.Q. Duong, Deep-learning artificial intelligence analysis of clinical variables predicts mortality in Covid-19 patients, *J. Am. Coll. Emerg. Physicians Open* 1 (2020) 1364–1373.
- [50] X. Li, P. Ge, J. Zhu, H. Li, J. Graham, A. Singer, P.S. Richman, T.Q. Duong, Deep learning prediction of likelihood of icu admission and mortality in Covid-19 patients using clinical variables, *PeerJ* 8 (2020) e10337.
- [51] M. Pourhomayoun, M. Shakibi, Predicting mortality risk in patients with Covid-19 using machine learning to help medical decision-making, *Smart Health* 20 (2021) 100178.
- [52] J.J. Näppi, T. Uemura, C. Watari, T. Hironaka, T. Kamiya, H. Yoshida, U-survival for prognostic prediction of disease progression and mortality of patients with Covid-19, *Sci. Rep.* 11 (2021) 9263.