

Semi-supervised classification of disease prognosis using CR images with clinical data structured graph

Jun Bai*

jun.bai@uconn.edu

University of Connecticut

Department of Computer Science &
Engineering

Storrs, Connecticut, USA

Bingjun Li*

bingjun.li@uconn.edu

University of Connecticut

Department of Computer Science &
Engineering

Storrs, Connecticut, USA

Sheida Nabavi

sheida.nabavi@uconn.edu

University of Connecticut

Department of Computer Science &
Engineering

Storrs, Connecticut, USA

ABSTRACT

Fast growing global connectivity and urbanisation increases the risk of spreading worldwide disease. The worldwide SARS-COV-2 disease causes healthcare system strained, especially for the intensive care units. Therefore, prognostic of patients' need for intensive care units is priority at the hospital admission stage for efficient resource allocation. In the early hospitalization, patient chest radiography and clinical data are always collected to diagnose. Hence, we proposed a clinical data structured graph Markov neural network embedding with computed radiography exam features (CGMNN) to predict the intensive care units demand for COVID patients. The study utilized 1,342 patients' chest computed radiography with clinical data from a public dataset. The proposed CGMNN outperforms baseline models with an accuracy of 0.82, a sensitivity of 0.82, a precision of 0.81, and an F1 score of 0.76.

KEYWORDS

Graph Markov neural network, multimodal, deep learning, COVID

ACM Reference Format:

Jun Bai, Bingjun Li, and Sheida Nabavi. 2022. Semi-supervised classification of disease prognosis using CR images with clinical data structured graph. In *13th ACM International Conference on Bioinformatics, Computational Biology and Health Informatics (BCB '22)*, August 7–10, 2022, Northbrook, IL, USA. ACM, New York, NY, USA, 9 pages. <https://doi.org/10.1145/3535508.3545548>

1 INTRODUCTION

A new respiratory infectious disease COVID-19 caused by a novel severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) emerged in 2019, has spread worldwide [19]. According to Johns Hopkins University's statistical data, there is a total of 491,987,588 cases with 6,154,269 deaths among them till April 2022, which has massively burdened healthcare systems around the world [16]. The pandemic is projected to cost a 16 trillion dollars lost in the economy in the US alone [14]. According to studies around the globe, the spectrum of COVID-19 infection symptoms ranges from

moderate to risky [10, 22]. In a study conducted in Wuhan, China, 32% of patients were admitted to the intensive care unit (ICU), and 46% of patients admitted to ICU died [22]. Yet, recent national data from England illustrate an improvement in the survival rate of patients admitted into ICU due to the better distribution of resources [15]. Hence, early diagnosis of patient and prediction of severity of COVID-19 is crucial to allocate supportive medical supplies, estimate hospital admission, prognosticate the mechanical ventilation usage, and plan to transition to ICU [20].

In clinical practice, chest Computed Radiography (CR) is often included in the initial patient assessment as the primary exam to evaluate patient condition [26]. Additional clinical information is collected during the examination including patient symptoms. For hospitalized patients, certain laboratory tests data are gathered for diagnostic and prognosis prediction [19, 37]. In the review by Martin et al., the authors outlined the predictors of COVID-19 severity, including common symptoms, clinical predictors, laboratory tests, and end-organ dysfunction [19]. Among clinical predictors, body-mass index (BMI), diabetes and hypertension are strong risk factors. Furthermore, cancer and oxygen saturation level are other major components linked to worse COVID-19 prognosis [19]. Biomarkers of laboratory tests show great association with disease severity, including coagulation defects, cardiac dysfunction, low blood lymphocyte percentage, liver injury, and non-specific biomarkers of cellular injury. Specially elevated lactate dehydrogenase (LDH) are associated with disease severity. In addition, renal dysfunctions such as urine glucose and protein are suggested to be used for COVID-19 severity prediction. A fast and accurate prognosis prediction based on the CR exams and the collected clinical information at the hospital admission stage could not only help the health care system overcomes peak times of pandemic, but also give the medical support to the patients that need it most.

COVID-19 is not the only worldwide pandemic happened in the recent history. In the first two decades in the twenty-first century alone, we have already experienced six major epidemics, including swine flue pandemic in 2009, Zika virus disease epidemic in 2015 and more, which caused millions of tragic losses of lives and great economic losses all over the world [6]. Due to environmental changes, quick urbanization, and fast growing global connectivity the potential threat of another fast spreading infectious disease development in the near future is very high [6]. Therefore, an accurate disease prognostic prediction system at early stage that can classifies admitted patients for future need of ICU admission will significantly improve the efficiency of health care systems and reduce unnecessary labor cost. Such system is important for us

*Both authors contributed equally to this research.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than ACM must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from permissions@acm.org.

BCB '22, August 7–10, 2022, Northbrook, IL, USA

© 2022 Association for Computing Machinery.

ACM ISBN 978-1-4503-9386-7/22/08...\$15.00

<https://doi.org/10.1145/3535508.3545548>

to reduce patient mortality rate in current and possible future pandemic outbreaks.

The goal of this study is to integrate all the available data—image and non-image data—collected at the hospital admission stage, to predict if the disease will progress and if a patient needs ICU admission during the course of the disease. For early prognosis prediction, we propose a semi-supervised graph Markov neural network based classification model. The model utilizes patients' physical, symptoms, and laboratory data to construct the graph structure, and embeds chest radiography features to the graph. The major contributions of our study are i) developing a more accurate multi-modal prognostic classifier that integrates clinical information and CR images; and ii) introducing a graph Markov model that employs Expectation–Maximization (EM) algorithm and combines the learning power of Statistical Relational Learning (SRL) methods and the Graph Neural Network (GNN) to learn feature representations from similar patients and local label distribution from unbalanced datasets; iii) providing a practical early prognosis prediction tool for hospitals ICU admission prediction that the unlabeled patients (newly hospital admitted patients) can also contribute to the model as well as the labeled patients (previously admitted to the ICU) even in an unbalanced fashion.

2 RELATED STUDIES

In the past decades, deep learning—a sub-type of artificial intelligence (AI)—has achieved state-of-the-art performance in classifying diseases using medical images from different modalities [4, 5, 7, 29, 36] (image based approach). To further improve classification accuracy, recently, combining non-imaging data with imaging data has been studied [11]. Hence, multi-modal models have been proposed using deep learning algorithms. In the following we discuss studies that used image based and multi-modal approaches to analyze COVID-19 data.

2.1 Image based approaches

Since the beginning of the COVID-19 outbreak, several studies have employed deep learning, specially convolutional neural networks (CNNs) for early diagnostic and detection of COVID-19 using chest Computerized Tomography (CT) and chest CR exams. In the study by Alazab et al., a CNN model is developed to detect patients infected by COVID-19 using X-ray images [2]. Ardakani et al. studied 10 well known CNN models—including AlexNet, VGG16, VGG19, SqueezeNet, GoogleNet, MobileNet-v2, ResNet-18, ResNet50, ResNet101 and Xception—to detect COVID-19 infected patients using CT images [3]. The results showed that ResNet and Xception perform the best. The MobileNet, VGG16, and ResNet models have also been used to diagnose COVID-19 using CR and CT images in other studies [1, 9, 17, 18, 30, 38]. In the real world, large labeled high quality datasets of acute worldwide spread disease are not available. Therefore, many studies are seeking for a solution for developing deep learning models that can work properly with limited data. To address this problem, Zhao et al. employed transfer learning (model was pretrained on CIFAR-10, ILSVRC-2012 and ImageNet datasets) and trained a pretrained ResNet model using CT images to detect COVID-19 [38]. In the study by Li et al. [28], a pretrained siamese model (pretrained by the CheXpert dataset [23])

is used to predict the severity of the disease. The pretrained model is trained by pairing COVID-19 CR images with random pooled normal CR images from CheXpert.

2.2 Multi-modal based approaches

Multi-modal based models have been studied extensively in recent years to improve the deep learning model performance. Few multi-modal approaches have been proposed for COVID-19 diagnostic and prognostic [12, 20, 25–27, 33, 35]. In the study by Wang et al., an automatic multi-model prognosis system is developed to identify high-risk patients. The proposed model, first, employs DenseNet121 to segment lung areas from CT images as Regions of Interest (ROIs) and feeds the ROIs to a fully connected deep neural network, then, fuses the extracted features from the CT images with clinical features [35]. In another study by Shi et al., patients are grouped to severe and non-severe groups. In this study, CT images' features extracted by a V-Net model are combined with clinical laboratory information to predict disease severity [33]. Kwon et al., proposed a model that combines CR images' features generated by a CNN model with clinical variables as input of a fully connected layer to classify severity of COVID-19 [26]. In a study by Gong et al., CT derived biomarkers (total opacity ratio and consolidation ratio) are combined with health records and fed to a generalized linear model for prognosis prediction [20].

These studies show that combining clinical data with images improves prognostic accuracy. However, all the current models are supervised classifiers and are required to be trained on a fully labeled dataset. Mostly a large amount of data is required to train deep learning models to achieve a state-of-the-art performance. However, in real practice, collecting a large amount of data with labels requires tremendous time and resources. When acute world wide spread disease happens, it is unrealistic to leverage huge amounts of data with labels. Furthermore, the majority of the current COVID-19 analysis methods focus on detecting COVID-19 infected patients. Recent COVID-19 disease detection and diagnostic methods have shown promising results in identifying COVID-19 infected patients with high accuracy [1–3, 9, 17, 18, 30, 38]. However, developing accurate and effective automatic COVID-19 prognostic predictive models is lacking behind. Therefore, we intend to tackle this problem in this study and our goal is to identify COVID-19 patients who need to be admitted to ICU from small sample size unbalanced datasets. We propose a multi-modal semi-supervised prognostic graph Markov neural network, which integrates image and non-image data and incorporates distributions of similar patients' features and local labels.

3 METHOD

The intuition behind the proposed model is that regular GNN models only focus on extracting the node features via the graph structure but ignoring the local label distribution. The proposed model aims to combine strength of the extracted node features and the local label distribution to make better prediction. The overview of our model's structure is shown in Fig. 1. The blue table and cubes represent image features extracted from patients' CR exams and the corresponding node feature vectors, respectively. The yellow table and edges represent the patient's non-image information

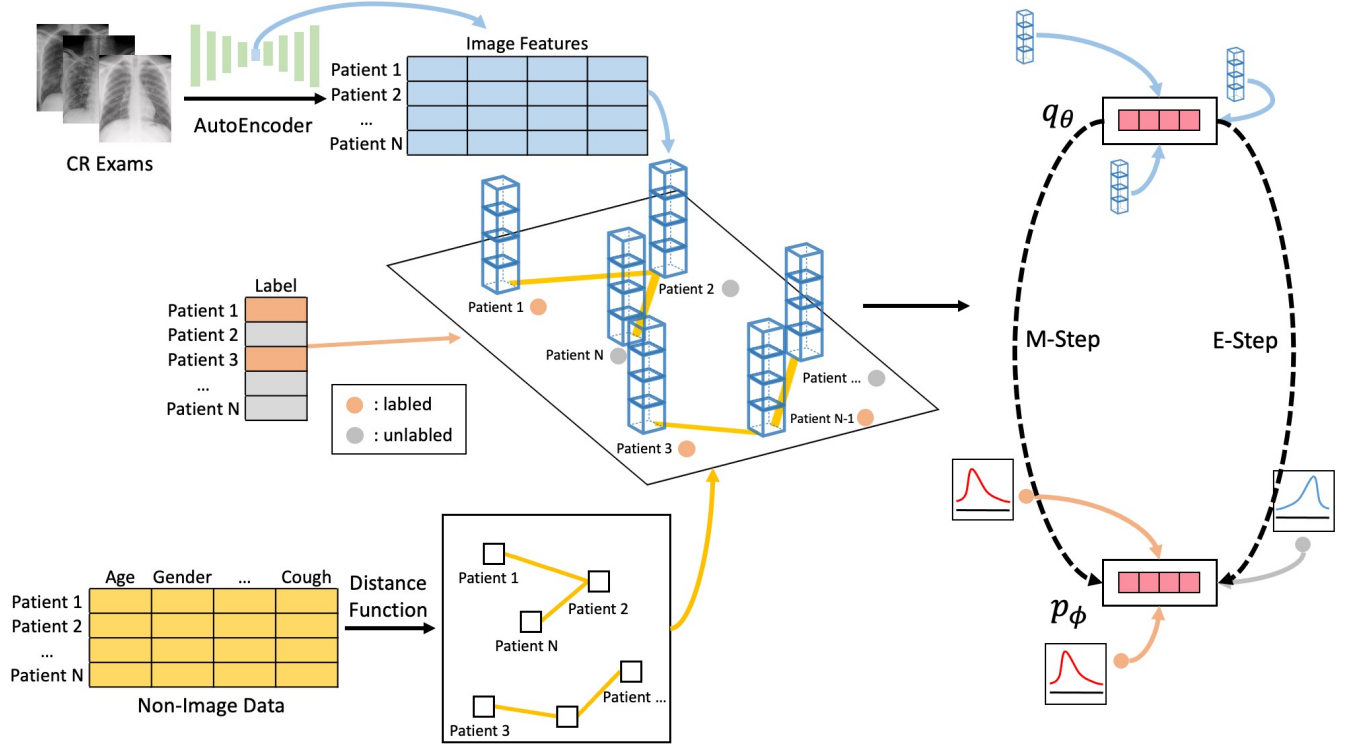


Figure 1: Overview of the model structure. A graph is built where patients are nodes and edges represent clinical information similarity between patients. An EM approach is used to estimate the label of unlabeled nodes. As shown in the top left, the chest CR exam’s features (blue blocks) are extracted by an autoencoder (green blocks) and are embedded in the graph as node features. The patients’ labels (orange for labeled patients, gray for unlabeled patients) are used in the graph as node labels, shown in the middle left. The patients’ non-image data (yellow blocks) are used to construct connections between patients (yellow edges) and used in the graph as edges between nodes, shown in the bottom left. The graph is fed into the EM framework to obtain graph representations (red blocks) for classification, shown in the right.

and the corresponding similarity network. The patients’ similarity network is used as the structure of the graph when nodes are patients (labelled and unlabeled). The image features are used as nodes’ attributes and the patients’ distances (similarity) are used as edges of the graph. The goal is to predict the label of the unlabeled patients. The proposed model extracts a lower-dimension representation for each node (patient) by learning both the node feature representation (node attributes) and the local (neighboring nodes) label distribution through the Expectation–Maximization (EM) algorithm.

The following sections will go through the details about obtaining the image features, constructing the node similarity graph, and formulating the proposed model.

3.1 Attribute Encoder & Network Construction

To extract the chest CR exams’ features, we implemented an unsupervised autoencoder model. The input dimension of the autoencoder model is 1024×1024 , and the output latent space is reduced to $z \in \mathbb{R}^{1 \times m}$, where the optimal value of m was determined by experiments. The autoencoder model contains six building blocks that each building block consists of a convolutional layer, a pooling

layer, and a normalization layer. To train the model, we use mean squared error between the reconstructed input and the original input as the loss function. The extracted features are used as the node feature matrix for the proposed model. Let’s denote the feature matrix, $\mathbf{x}_V$ as $\mathbf{x}_V = \{\mathbf{x}_{v1}, \mathbf{x}_{v2}, \dots, \mathbf{x}_{vn}\} \in \mathbb{R}^{n \times m}$, where n is number of patients, and m is the patient feature dimensions.

The edges of the proposed graph Markov model represent similarity between patients (nodes) that enable the model to aggregate information through neighboring nodes. We defined the edges based on the distances between patients in the clinical information space. The hypothesis is patients who share similar background, such as cancer history, age, Body Mass Index (BMI) and smoking history, would more likely have a similar physical response to the disease. We used the K nearest neighbors (KNN) to construct the patients’ similarity network. Given the patients’ non-image clinical data denoted as $\mathbf{CI} = \{\mathbf{CI}_1, \mathbf{CI}_j, \dots, \mathbf{CI}_n\}$, we used the Euclidean distance function as given below.

$$d_{i,j} = \sqrt{\sum_{c=0}^r (\mathbf{CI}_i^c - \mathbf{CI}_j^c)^2}, \quad (1)$$

where $\mathbf{CI}_i$ is patient i clinical data, i and j are two nodes (patients), and r is the dimension of non-image data. Then, we selected K nearest neighbors of each patient to connect and form edges. In the experiments we used $K = 3$ that gives the best performance.

3.2 Disease Graph Markov Neural Network (CGMNN)

The setup of our proposed model consists of a graph $G = (V, E, \mathbf{x}_V)$, where V is a set of nodes (patients), E is a set of edges between patients, and $\mathbf{x}_V$ is the patient feature matrix. We defined each patient as a node $v \in V$. The edges of graph were computed by KNN using the Euclidean distance of patients' clinical data as described above. The labels of the patients are defined as if they went into the ICU during their stay at hospital. The labeled patients are denoted as L with labels, $\mathbf{y}_L$. The goal of our model is to predict the labels of $\mathbf{y}_U$ for other unlabeled patients, $U = V \setminus L$.

The main idea of our proposed model, CGMNN, is to combine the strength of Statistical Relational Learning (SRL) and the Graph Neural Network (GNN) methods such that CGMNN can utilize both the power of learned feature representations of patients' chest CR exams and the local distribution of patients' labels [31]. CGMNN models the joint distribution of sample labels given sample attributes, $p_\phi(\mathbf{y}_V|\mathbf{x}_V)$ through a conditional random field, where ϕ is a model parameter. As the log-likelihood, $\log p_\phi(\mathbf{y}_V|\mathbf{x}_V)$ is difficult to optimize, we used a variational Expectation–Maximization (EM) framework to optimize the Evidence Lower Bound (ELBO) of the log-likelihood [8],

$$\log p_\phi(\mathbf{y}_L|\mathbf{x}_V) \leq \mathbb{E}_{q_\theta(\mathbf{y}_U|\mathbf{x}_V)} [\log p_\phi(\mathbf{y}_L, \mathbf{y}_U|\mathbf{x}_V) - \log q_\theta(\mathbf{y}_U|\mathbf{x}_V)], \quad (2)$$

where q_θ can be any distribution across $\mathbf{y}_U$.

In the E-step, p_ϕ is fixed as target and $q_\theta(\mathbf{y}_U|\mathbf{x}_V)$ is updated by a graph convolutional neural network (GCN) with the following objective function.

$$O_\theta = O_{\theta,U} + O_{\theta,L} = \sum_{n \in U} \mathbb{E}_{p_\theta(\mathbf{y}_n|\hat{\mathbf{y}}_{NB(n)}, \mathbf{x}_V)} [\log q_\theta(\mathbf{y}_n|\mathbf{x}_V)] + \sum_{n \in L} \log q_\theta(\mathbf{y}_n|\mathbf{x}_V). \quad (3)$$

In the M-step, q_θ is fixed instead and p_ϕ is updated by another GCN with the following objective function.

$$O_\phi = \sum_{n \in V} \log p_\phi(\hat{\mathbf{y}}_n|\hat{\mathbf{y}}_{NB(n)}, \mathbf{x}_V), \quad (4)$$

where $\hat{\mathbf{y}}_{NB(n)}$ is the estimated local labels for the neighbours of node n . Let's denote a neighbour node of node n , r . For a labelled node r , $\hat{y}_r$ is defined as the ground truth label, and for an unlabelled node r , $\hat{y}_r$ is defined as a sample from $q_\theta(y_r|\mathbf{x}_V)$. The optimized $q_\theta(\mathbf{y}_n|\mathbf{x}_V)$ is passed through a softmax layer for classification. The overall proposed model's structure is shown in Figure 1.

4 DATA & EXPERIMENT

4.1 Data

We used a public COVID-19 dataset provided by Stony Brook University [13, 32]. All cases in the dataset are confirmed COVID-19 patients. We selected 1,343 patients' (258 ICU admitted patients and 1085 no ICU history patients) CR exams with the corresponding physical data, symptoms, and laboratory data (shown in Table 1). Patients' physical data distributions are shown in Figure 2.

4.2 Baseline models

We compared the proposed model, CGMNN, with the following baseline models. We included models that employ only patients' CR exams, only patients' clinical information, and both patient's clinical and CR exams (multi-modal models):

ResNet. To compare the proposed model with a state-of-the-art image based model, we followed the study by Alazab et al. and trained ResNet model using only patients' CR images. We employed ImageNet pretrained ResNet50 to classify patients' ICU admission [24]. We used the binary-cross-entropy loss function.

RF. We compared the performance of the proposed model with that of the random forest (RF) model using only patients' clinical information (including laboratory findings, symptom findings, and physical findings). The depth of the random forest is 50, and the random state is set to 5. Since RF models perform well using balanced datasets, to achieve the best performance, we under-sampled the non-ICU samples to have a balanced dataset.

DNN. We trained a shallow fully connected neural network (DNN) model using only patients' clinical information — including laboratory findings, symptom findings and physical findings — (single modality model), and using concatenation of patients' CR features (extracted by the autoencoder) with the clinical information (multi-modal model). The single modality model contains two hidden layers (the first layer dimension is 32, and the second layer dimension is 16) and one output layer. The multi-modal model contains three layers (the first layer dimension is 64, and the second layer dimension is 32, and third layer dimension is 16). We used the binary-cross-entropy loss function.

GCN. We included semi-supervised graph convolutional neural networks (GCNs) as baseline models to compare their performance with that of the proposed model. To extensively study the performance of GCNs, we trained i) a GCN using physical information to form the graph structure and embedding node features using laboratory information and symptom information (single modality model); and ii) a GCN using physical information to form the graph structure and embedding node features using laboratory information, symptom information, and CR features extracted from autoencoder (multi-modal model). In experiments, we implemented two layers of GCN with one dropout layer. We used the binary-cross-entropy loss function.

ResDNN. We also compared the performance of the proposed model with that of an end to end deep learning multi-modal classification model using patients' CR image and clinical information (include patients' physical, symptom, and laboratory information).

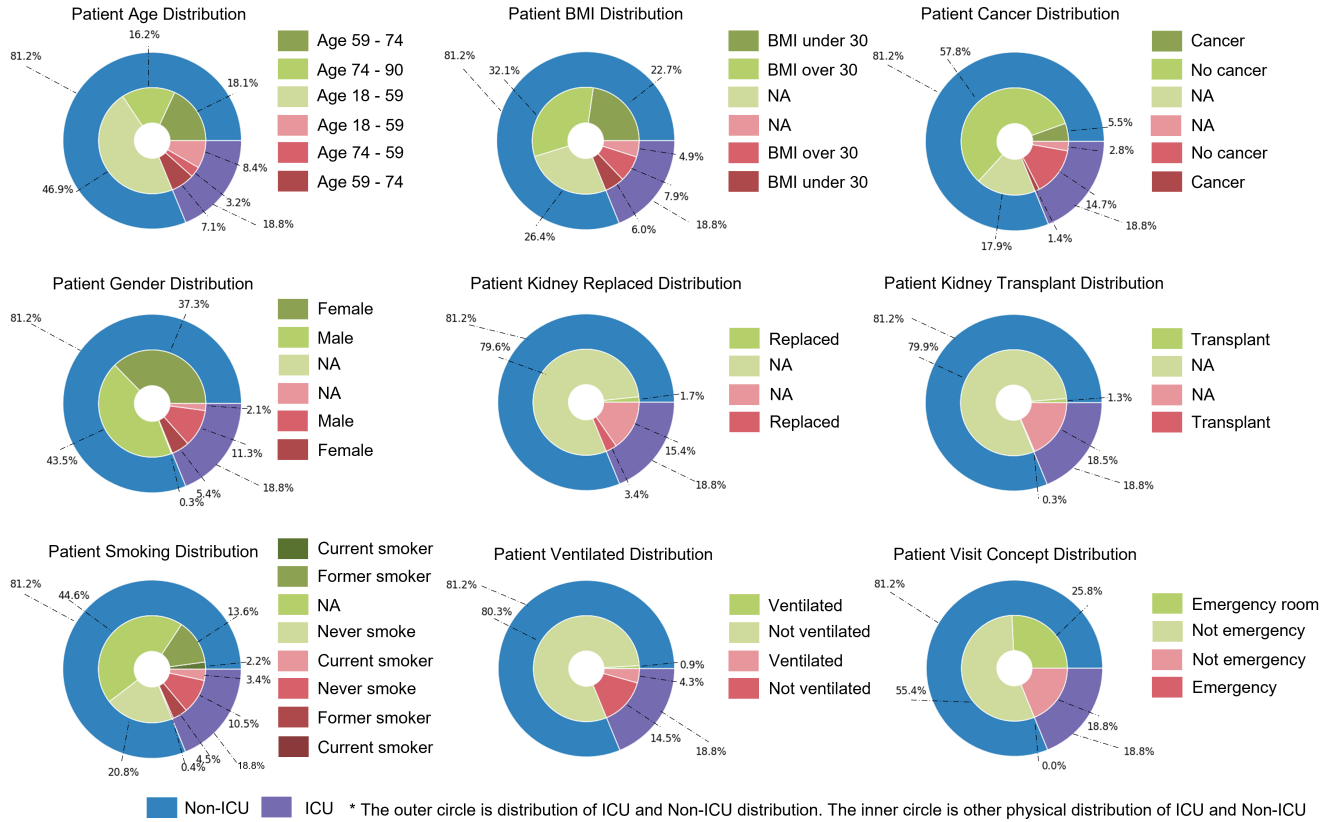


Figure 2: Patients' physical information distribution over ICU and Non-ICU admissions.

Table 1: Non-image data of patients

Clinical Information	Protein in urine, chronic obstructive pulmonary disease, lymphocytes, aspartate, alanine, A1C, sodium, potassium, chloride, bicarbonate, blood urea nitrogen, creatinine, estimated glomerular filtration rate, blood pH, troponin, D-dimer, erythrocyte sedimentation rate, microscopic hematuria, systolic blood pressure, mean arterial pressure, procalcitonin, ferritin, proteinuria
Symptom Information	Cough, dyspnea, nausea, vomiting, diarrhea, abdominal pain, fever, temperature, pulse, respiration, heart rate
Physical Information	age, gender, kidney replacement therapy, kidney transplant, cancer, smoking status, BMI, visit concept, ventilated

In this model patients' CR images are fed into ImageNet pretrained ResNet50 and patients' clinical information is fed into a shallow DNN. Features were late fused and fed to multi-layer perception (MLP) for classification. To avoid overfitting, we employed Gaussian Error Linear Unit activation function (GELU)[21] and normalization to smooth out the concatenation between CR features and clinical features. The loss function is binary-cross-entropy.

4.3 Experiment setup

We examined three variants of the proposed model as listed below.

- We computed similarity between patients and built the graph structure using only patients' physical information. CR image features combined with laboratory and symptom data are embedded as node features. We called this model CGMNN-PI.
- We computed similarity between patients and built the graph structure using patient's physical and symptom data. CR image features combined with laboratory data are embedded as node features. We called this model CGMNN-PISI.

Table 2: Performance of the methods using CR exams and clinical information

Model	Data	Graph Structure	Accuracy	Precision	Sensitivity	F1
ResNet	CR	-	0.75	0.81	0.52	0.63
RF	CI	-	0.71	0.83	0.71	0.74
DNN	CI	-	0.72	0.83	0.72	0.75
GCN	CI	PI	0.80	0.64	0.80	0.72
DNN	CR+CI	-	0.66	0.88	0.66	0.76
ResDNN	CR+CI	-	0.76	0.69	0.65	0.67
GCN	CR+CI	PI	0.80	0.64	0.80	0.72
CGMNN-PI	CR+CI	PI	0.81	0.82	0.80	0.74
CGMNN-PISILI	CR	PI+SI+CI	0.77	0.80	0.77	0.77
CGMNN-PISI	CR+CI	PI+SI	0.82	0.81	0.82	0.76

The bold font indicates the highest values

- We computed similarity between patients and built the graph structure using all the clinical information (physical, laboratory, and symptom information). CR image features are embedded as node features. We called this model CGMNN-PISILI.

In our study, we examined the performance of the proposed models with different CR feature dimension including 512, and 1024. The dimension of the CR features does not have a big effect on the performance and we reported the results for the CR feature size of 1024. A comprehensive evaluation metric has been used to monitor the bias-variance trade-off of the proposed model and baseline models. We used PyTorch in Python to implement the autoencoder, proposed CGMNN, and GCN models. The proposed model is trained with 100 epochs. The baseline ResNet, DNN, and ResDNN models are implemented with the Keras package in Python. RF model is implemented with scikit-learn package in Python. To overcome overfitting for the CNN-based models and feature extraction of CR images (ResNet, ResDNN, and autoencoder), we applied data augmentation including flip, rotation, and CLAHE [34] filtering to the training dataset.

Before applying the CNN-based models, including the autoencoder, we performed image pre-processing and normalization. To standardize model training for all the CNN-based models, CR images are trimmed and resized to 1024×1024 . For resizing, we employed bi-linear interpolation. For the baseline models, we used 70% of the data of each class for training, and the remaining 30% for validation and testing. To examine the effect of the percentage of unlabeled samples on the performance of the proposed model, we used different partition of the samples as labeled and unlabeled nodes. In addition, we studied the performance of CGMNN using different distance functions, including Euclidean distance, Minkowski distance, and Manhattan distance, to compute similarity between patients for building the graph. Moreover, in this study, we explored the effect of changing K for the KNN analysis to build the graph. The results are reported in section 5.2.

5 RESULTS

5.1 Overall results

We compared the performance of the proposed model to those of the five baseline models explained in section 4.2 in terms of Sensitivity,

Precision, Accuracy and F1 scores. The results of the proposed model and the baseline models are shown in Table 2. The baseline models can be divided into two major categories, single-modality data models and multi-modality data models. Single-modality data models consist of ResNet, RF and DNN that are based on either clinical information or CR image features alone. Multi-modality data models consist of GCN, DNN, and ResDNN that represent the multi-modality data in the form of early fusion using a supervised approach, or in the form of graph structure using a semi-supervised approach. As shown in Table 2, the proposed models outperform the baseline models in almost all the performance metrics.

In general, as expected, the multi-modal models outperform the single-modal models. Among the single-modal models the ResNet model with the CR image features performs the best in accuracy, 0.75, but the worst in all other three metrics. It indicates that the prediction from the ResNet model is very biased towards one of the classes, which is a common phenomenon for a model that doesn't adapt well to the high-dimensional unbalanced data. In addition, the results confirm our hypothesis that CR image features offer more information about the patient's prognosis compared to the clinical information, but a proper model is needed to utilize the power of both data. Note that for the RF model we used a balanced subset of data for training and test. Using unbalanced data results in a poor performance of the RF model (not shown here).

The multi-modal baseline models, GCN, DNN, and ResDNN, similar to our proposed models incorporate both clinical and image data. ResDNN and DNN models use early fusion as the integration method. Compared to the DNN model, the ResDNN model has a better performance in accuracy, 0.76, but the worse performance in all other metrics. Similar to the results for the single modality models, the high accuracy model suffers poor performance in precision, sensitivity and F1 score, which indicates the model doesn't adapt well to the unbalanced data. As shown in Table 2, the performance improvement gained from introducing the extra modality data in early fusion integration models is minor. Thus, a better way to integrate these two data modalities is needed. With the similar setup to build the graph as the proposed model, the GCN model uses the clinical information or a combination of the clinical information and the CR image features as the node features, and builds the graph network based on the physical information. Unlike

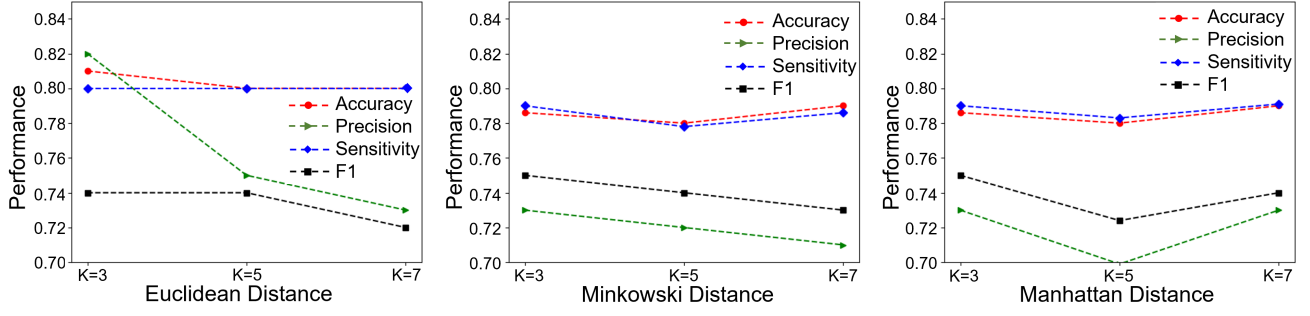


Figure 3: Ablation study for computing distance between pairs of nodes and selecting edges. K is the number of edges connected to a node. We studied Euclidean distance, Minkowski distance and Manhattan distance.

Table 3: Performance of baseline model and GCMNN on different partitions of the dataset. For example, in Case.No1, 30% of the dataset is labeled (training set for baseline models) and 70% of the dataset is unlabeled (test set for baseline models)

	Data	Graph Structure	Case.No1 30%		Case.No2 40%		Case.No3 50%		Case.No4 60%		Case.No5 70%	
			Accuracy	F1	Accuracy	F1	Accuracy	F1	Accuracy	F1	Accuracy	F1
ResNet	CR	-	0.454	0.51	0.494	0.54	0.611	0.64	0.685	0.70	0.749	0.63
RF	CI	-	0.336	0.33	0.411	0.24	0.471	0.30	0.601	0.59	0.711	0.74
DNN	CI	-	0.507	0.47	0.519	0.52	0.585	0.58	0.633	0.63	0.719	0.75
GCN	CI	PI	0.792	0.69	0.794	0.69	0.795	0.70	0.801	0.72	0.803	0.72
DNN	CR + CI	-	0.512	0.57	0.576	0.63	0.693	0.70	0.702	0.73	0.661	0.76
ResDNN	CR + CI	-	0.591	0.64	0.563	0.61	0.607	0.65	0.696	0.71	0.757	0.67
GCN	CR + CI	PI	0.802	0.72	0.801	0.71	0.793	0.70	0.801	0.72	0.803	0.72
GCMNN	CR + CI	PI	0.807	0.73	0.805	0.72	0.802	0.74	0.811	0.74	0.818	0.76

The bold font indicates the highest values

our proposed model, the GCN model employs graph convolution to aggregate information and represent features. While our model outperforms the GCN model, the GCN model shows the best accuracy, 0.80, among the baseline models, but it also shows a relatively low precision, 0.64, and F1 score, 0.72. We observed that the GCN model also suffers from the unbalanced data and provides a very biased prediction results towards the majority class in the data.

We trained the proposed model on three different graph networks built using: i) only physical information, called CGMNN-PI, ii) physical information plus symptom information, called CGMNN-PISI, and iii) all three kinds of non-image data, called CGMNN-PISILI. For the former two graph network (CGMNN-PI, CGMNN-PISI), a combination of clinical information and CR image features is used as the node features. For the later graph network (CGMNN-PISILI), only the CR image features are used as the node features. Overall, the proposed model with the graph network constructed from the physical information and symptom information performs the best in terms of accuracy, 0.82, and sensitivity, 0.81; and comparable in terms of precision and F1 score. The proposed models with three different settings of node feature and graph network perform relatively similar in terms of all the four performance metrics. Compared to the most accurate baseline models, the proposed model shows not only a better prediction accuracy, but also better scores

in precision, sensitivity and F1 scores, which indicates the proposed model achieves a good fit on the unbalanced data. It demonstrates that the proposed model performs better on the unbalanced data in all terms compared to all the baseline models.

5.2 Ablation Study

To further examine the sensitivity and robustness of the proposed model with respect to its hyperparameters, we conducted a series of experiments. We investigated the effect of using different percentages of labeled and unlabeled nodes, different values of K for the KNN algorithm in building graphs, and different distance functions in the KNN algorithms.

To examine the effect of using different percentages of labeled and unlabeled data on the performance of the proposed model, we constructed the model with different partitions of labeled and unlabeled nodes from 30% to 70% at a step of 10%. As shown in Table 3, the performance of the proposed model is the best in terms of accuracy and F1 score across all the different partition settings, and its performance shows very little variation in these two metrics at different partitions. This demonstrates the robustness and usefulness of the proposed model in real-world scenarios where not many labeled data are available. The proposed model shows the best performance at the partition of 70% of labeled and 30% unlabeled.

As expected, all the CNN-based models and RF show underperform when less labeled data are available for training. When using less than 60% of the data for training, CNN-based models failed to predict ICU demands. The RF model also show poor performance using less number of training data. The GCN-based semi-supervised models show robust performance with changing the portion of unlabeled data. However, the proposed model showed better results compared with the GCN-based models. The more robust performance of the semi-supervised graph-based models respect to the portion of unlabeled data can be due to the contribution of the unlabeled data to the learning process.

We also tested the performance of the proposed model using different distance functions to compute edges and different K values to build connections in the graph. The results are shown in Figure 3. The performance of the proposed model does not change significantly using different distance functions. However, the model with the network constructed using the Euclidean distance function achieves the best performance compared to the models constructed using the Minkowski distance or the Manhattan distance. The impact of the K value on the performance of the proposed model with a network constructed using Minkowski distance and Manhattan distance is negligible. In the case of employing the Euclidean distance function, precision decreases as the value of K increases and the optimal K value is 3. This can be caused by the KNN algorithm connects patients who aren't similar enough as K increases, which is equivalent to adding noise to the graph. Thus, the performance of the proposed model deteriorates with the increase in K .

6 CONCLUSION

In this study, we proposed a novel semi-supervised deep learning model to predict disease prognosis using partially labeled and unbalanced datasets. We experimentally investigated the proposed model for early prediction of ICU admission for COVID-19 infected patients. Our proposed deep learning model utilizes patients' CR exams and clinical information, including laboratory tested biomarkers, disease symptoms, and patients' physical information. The proposed model takes advantage of the great potential of statistical relational learning and GNN learning by employing graph Markov neural network, when clinical information is used to build the graph and CR features are embedded in the graph. Results show that employing a clinical data structured graph with CR features as the nodes' attributes improves the classification performance compared with employing only CR exams in a state-of-the-art deep learning model and employing only clinical information in a machine learning model. We show that, in addition to clinical and biomarker data, the CR exams can contribute to indicate COVID-19 severity and the need for ICU admission. Furthermore, the results indicate that the proposed model can perform well even when only a small portion of data is labeled, which is a more realistic scenario analyzing medical data. Our results suggest that relational learning deep learning based models can be employed to help overwhelmed healthcare systems to make faster and more accurate prognosis.

7 ACKNOWLEDGMENTS

This work is supported partially by a grant from the University of Connecticut Research Excellence Program, PIs: Nabavi and by the

National Science Foundation (NSF) under grant No. 1942303, PI: Nabavi.

REFERENCES

- [1] Talha Burak Alakus and Ibrahim Turkoglu. 2020. Comparison of deep learning approaches to predict COVID-19 infection. *Chaos, Solitons & Fractals* 140 (Nov. 2020), 110120. <https://doi.org/10.1016/j.chaos.2020.110120>
- [2] Moutaz Alazab, Albara Awajan, Abdelwaddood Mesleh, Ajith Abraham, Vansh Jatana, and Salah Alhyari. 2020. COVID-19 Prediction and Detection Using Deep Learning. *International Journal of Computer Information Systems and Industrial Management Applications* 12 (2020), 168–181.
- [3] Ali Abbasian Ardakani, Alireza Rajabzadeh Kanafi, U. Rajendra Acharya, Nazanin Khadem, and Afshin Mohammadi. 2020. Application of deep learning technique to manage COVID-19 in routine clinical practice using CT images: Results of 10 convolutional neural networks. *Computers in Biology and Medicine* 121 (June 2020), 103795. <https://doi.org/10.1016/j.combiomed.2020.103795>
- [4] Jun Bai, Annie Jin, Tianyu Wang, Clifford Yang, and Sheida Nabavi. 2022. Feature fusion siamese network for breast cancer detection comparing current and prior mammograms. *Medical Physics* n/a, n/a (2022). <https://doi.org/10.1002/mp.15598>. eprint: <https://onlinelibrary.wiley.com/doi/pdf/10.1002/mp.15598>
- [5] Jun Bai, Russell Posner, Tianyu Wang, Clifford Yang, and Sheida Nabavi. 2021. Applying deep learning in digital breast tomosynthesis for automatic breast cancer detection: A review. *Medical Image Analysis* 71 (July 2021), 102049. <https://doi.org/10.1016/j.media.2021.102049>
- [6] Rachel E. Baker, Ayesha S. Mahmud, Ian F. Miller, Malavika Rajeev, Fidisoa Rasambainarivo, Benjamin L. Rice, Saki Takahashi, Andrew J. Tatem, Caroline E. Wagner, Lin-Fa Wang, Amy Wesolowski, and C. Jessica E. Metcalf. 2022. Infectious disease in an era of global change. *Nature Reviews Microbiology* 20, 4 (April 2022), 193–205. <https://doi.org/10.1038/s41579-021-00639-z> Number: 4 Publisher: Nature Publishing Group.
- [7] Sweta Bhattacharya, Praveen Kumar Reddy Maddikunta, Quoc-Viet Pham, Thippa Reddy Gadekallu, Siva Rama Krishnan S, Chiranjil Lal Chowdhary, Mamoun Alazab, and Md. Jalil Piran. 2021. Deep learning and medical image processing for coronavirus (COVID-19) pandemic: A survey. *Sustainable Cities and Society* 65 (Feb. 2021), 102589. <https://doi.org/10.1016/j.scs.2020.102589>
- [8] Blei, David M., Kucukelbir, Alp, and McAuliffe, Jon D. 2017. Variational Inference: A Review for Statisticians. *J. Amer. Statist. Assoc.* 112, 518 (2017), 859–877. <https://doi.org/10.1080/01621459.2017.1285773>
- [9] Michaela Cellina, Marta Panzeri, and Giancarlo Oliva. 2020. Chest Radiography Features Help to Predict a Favorable Outcome in Patients with Coronavirus Disease 2019. *Radiology* 297, 1 (Oct. 2020), E238–E238. <https://doi.org/10.1148/radiol.2020202326>
- [10] Jasper Fuk-Woo Chan, Shuofeng Yuan, Kin-Hang Kok, Kelvin Kai-Wang To, Hin Chu, Jin Yang, Fanfan Xing, Jieliang Liu, Cyril Chik-Yan Yip, Rosana Wing-Shan Poon, Hoi-Wah Tsoi, Simon Kam-Fai Lo, Kwok-Hung Chan, Vincent Kwok-Man Poon, Wan-Mui Chan, Jonathan Daniel Ip, Jian-Piao Cai, Vincent Chi-Chung Cheng, Honglin Chen, Christopher Kim-Ming Hui, and Kwok-Yung Yuen. 2020. A familial cluster of pneumonia associated with the 2019 novel coronavirus indicating person-to-person transmission: a study of a family cluster. *The Lancet* 395, 10223 (Feb. 2020), 514–523. [https://doi.org/10.1016/S0140-6736\(20\)30154-9](https://doi.org/10.1016/S0140-6736(20)30154-9)
- [11] Wei Chen, Weiping Wang, Li Liu, and Michael S. Lew. 2021. New Ideas and Trends in Deep Multimodal Content Understanding: A Review. *Neurocomputing* 426 (Feb. 2021), 195–215. <https://doi.org/10.1016/j.neucom.2020.10.042>
- [12] Matteo Chierigato, Fabio Frangiamore, Mauro Morassi, Claudia Baresi, Stefania Nici, Chiara Bassetti, Claudio Bnà, and Marco Galelli. 2022. A hybrid machine learning/deep learning COVID-19 severity predictive model from CT images and clinical data. *Scientific Reports* 12, 1 (March 2022), 4329. <https://doi.org/10.1038/s41598-022-07890-1> Number: 1 Publisher: Nature Publishing Group.
- [13] Kenneth Clark, Bruce Vendt, Kirk Smith, John Freymann, Justin Kirby, Paul Koppel, Stephen Moore, Stanley Phillips, David Maffitt, Michael Pringle, Lawrence Tarbox, and Fred Prior. 2013. The Cancer Imaging Archive (TCIA): Maintaining and Operating a Public Information Repository. *Journal of Digital Imaging* 26, 6 (Dec. 2013), 1045–1057. <https://doi.org/10.1007/s10278-013-9622-7>
- [14] David M. Cutler and Lawrence H. Summers. 2020. The COVID-19 Pandemic and the \$16 Trillion Virus. *JAMA* 324, 15 (Oct. 2020), 1495. <https://doi.org/10.1001/jama.2020.19759>
- [15] John M. Dennis, Andrew P. McGovern, Sebastian J. Vollmer, and Bilal A. Mateen. 2021. Improving Survival of Critical Care Patients With Coronavirus Disease 2019 in England: A National Cohort Study, March to June 2020*. *Critical Care Medicine* 49, 2 (Feb. 2021), 209–214. <https://doi.org/10.1097/CCM.00000000000004747>
- [16] Ensheng Dong, Hongru Du, and Lauren Gardner. 2020. An interactive web-based dashboard to track COVID-19 in real time. *The Lancet Infectious Diseases* 20, 5 (May 2020), 533–534. [https://doi.org/10.1016/S1473-3099\(20\)30120-1](https://doi.org/10.1016/S1473-3099(20)30120-1)
- [17] Ezz El-Din Hemdan, Marwa A. Shouman, and Mohamed Esmail Karar. 2020. COVIDX-Net: A Framework of Deep Learning Classifiers to Diagnose COVID-19 in X-Ray Images. *arXiv e-prints* (March 2020), arXiv:2003.11055. <https://arxiv.org/abs/2003.11055>

- //ui.adsabs.harvard.edu/abs/2020arXiv200311055E/abstract
- [18] Muhammad Farooq and Abdul Hafeez. 2020. COVID-ResNet: A Deep Learning Framework for Screening of COVID19 from Radiographs. *arXiv e-prints* (2020), 5.
 - [19] Benjamin Gallo Marin, Ghazal Aghagholi, Katya Lavine, Lanbo Yang, Emily J. Siff, Silvia S. Chiang, Thais P. Salazar-Mather, Luba Dumenco, Michael C Savaria, Su N. Aung, Timothy Flanigan, and Ian C. Michelow. 2021. Predictors of COVID-19 severity: A literature review. *Reviews in Medical Virology* 31, 1 (2021), e2146. <https://doi.org/10.1002/rmv.2146>
 - [20] Kuang Gong, Dufan Wu, Chiara Daniela Arru, Fatemeh Homayounieh, Nir Neumark, Jiahui Guan, Varun Buch, Kyungsang Kim, Bernardo Canedo Bizzo, Hui Ren, Won Young Tak, Soo Young Park, Yu Rim Lee, Min Kyu Kang, Jung Gil Park, Alessandro Carriero, Luca Saba, Mahsa Masjedi, Hamidreza Talari, Rosa Babaei, Hadi Karimi Mobin, Shadi Ebrahimi, Ning Guo, Subba R. Digumarthy, Ittai Dayan, Mannudeep K. Kalra, and Quanzheng Li. 2021. A multi-center study of COVID-19 patient prognosis using deep learning-based CT image analysis and electronic health records. *European Journal of Radiology* 139 (June 2021), 109583. <https://doi.org/10.1016/j.ejrad.2021.109583>
 - [21] Dan Hendrycks and Kevin Gimpel. 2020. Gaussian Error Linear Units (GELUs). *arXiv:1606.08415 [cs]* (July 2020). <http://arxiv.org/abs/1606.08415> arXiv: 1606.08415.
 - [22] Chaolin Huang, Yeming Wang, Xingwang Li, Lili Ren, Jianping Zhao, Yi Hu, Li Zhang, Guohui Fan, Jiuyang Xu, Xiaoying Gu, Zhenshun Cheng, Ting Yu, Jiaan Xia, Yuan Wei, Wenjuan Wu, Xuelei Xie, Wen Yin, Hui Li, Min Liu, Yan Xiao, Hong Gao, Li Guo, Jungang Xie, Guangfa Wang, Rongmeng Jiang, Zhancheng Gao, Qi Jin, Jianwei Wang, and Bin Cao. 2020. Clinical features of patients infected with 2019 novel coronavirus in Wuhan, China. *The Lancet* 395, 10223 (Feb. 2020), 497–506. [https://doi.org/10.1016/S0140-6736\(20\)30183-5](https://doi.org/10.1016/S0140-6736(20)30183-5)
 - [23] Jeremy Irvin, Pranav Rajpurkar, Michael Ko, Yifan Yu, Silvana Ciurea-Ilcus, Chris Chute, Henrik Marklund, Behzad Haghighi, Robyn Ball, Katie Shpanskaya, Jayne Seekins, David A. Mong, Safwan S. Halabi, Jesse K. Sandberg, Ricky Jones, David B. Larson, Curtis P. Langlotz, Bhavik N. Patel, Matthew P. Lungren, and Andrew Y. Ng. 2019. CheXpert: A large chest radiograph dataset with uncertainty labels and expert comparison. In *33rd AAAI Conference on Artificial Intelligence, AAAI 2019, 31st Innovative Applications of Artificial Intelligence Conference, IAAI 2019 and the 9th AAAI Symposium on Educational Advances in Artificial Intelligence, EAAI 2019*. AAAI Press, 590–597. <https://jhu.pure.elsevier.com/en/publications/chexpert-a-large-chest-radiograph-dataset-with-uncertainty-labels>
 - [24] He Kaiming, Zhang xiangyu, Ren Shaoqing, and Sun Jian. 2015. Deep Residual Learning for Image Recognition. *arXiv:1512.03385 [cs]* (2015). <https://arxiv.org/pdf/1512.03385.pdf>
 - [25] Chris K. Kim, Ji Whae Choi, Zhicheng Jiao, Dongcui Wang, Jing Wu, Thomas Y. Yi, Kasey C. Halsey, Feysiope Eweje, Thi My Linh Tran, Chang Liu, Robin Wang, John Sollee, Celina Hsieh, Ken Chang, Fang-Xue Yang, Ritambhara Singh, Jie-Lin Ou, Raymond Y. Huang, Cai Feng, Michael D. Feldman, Tao Liu, Ji Sheng Gong, Shaolei Lu, Carsten Eickhoff, Xue Feng, Ihab Kamel, Ronnie Sebro, Michael K. Atalay, Terrance Healey, Yong Fan, Wei-Hua Liao, Jianxin Wang, and Harrison X. Bai. 2022. An automated COVID-19 triage pipeline using artificial intelligence based on chest radiographs and clinical data. *npj Digital Medicine* 5, 1 (Jan. 2022), 1–9. <https://doi.org/10.1038/s41746-021-00546-w> Number: 1 Publisher: Nature Publishing Group.
 - [26] Young Joon (Fred) Kwon, Danielle Toussie, Mark Finkelstein, Mario A. Cedillo, Samuel Z. Maron, Sayan Manna, Nicholas Voutsinas, Corey Eber, Adam Jacobi, Adam Bernheim, Yogesh Sean Gupta, Michael S. Chung, Zahi A. Fayad, Benjamin S. Glicksberg, Eric K. Oermann, and Anthony B. Costa. 2021. Combining Initial Radiographs and Clinical Variables Improves Deep Learning Prognostication in Patients with COVID-19 from the Emergency Department. *Radiology: Artificial Intelligence* 3, 2 (March 2021), e200098. <https://doi.org/10.1148/ryai.2020200098>
 - [27] Nathalie Lassau, Samy Ammari, Emilie Chouzenoux, Hugo Gortais, Paul Herent, Matthieu Devilder, Samer Soliman, Olivier Meyrignac, Marie-Pauline Talabard, Jean-Philippe Lamarque, Remy Dubois, Nicolas Loiseau, Paul Trichelair, Etienne Bendjebbar, Gabriel Garcia, Corinne Balleyguier, Mansouria Merad, Annabelle Stoclin, Simon Jegou, Franck Griscelli, Nicolas Tetelboum, Yingping Li, Sagar Verma, Matthieu Terris, Tasnim Dardouri, Kavya Gupta, Ana Neacsu, Frank Chemouni, Meriem Sefta, Paul Jehanno, Imad Bousaid, Yannick Boursin, Emmanuel Planchet, Mikael Azoulay, Jocelyn Dachary, Fabien Brulport, Adrian Gonzalez, Olivier Dehaene, Jean-Baptiste Schiratti, Kathryn Schutte, Jean-Christophe Pesquet, Hugues Talbot, Elodie Premier, Gilles Wainrib, Thomas Clozel, Fabrice Barlesi, Marie-France Bellin, and Michael G. B. Blum. 2021. Integrating deep learning CT-scan model, biological and clinical variables to predict severity of COVID-19 patients. *Nature Communications* 12, 1 (Jan. 2021), 634. <https://doi.org/10.1038/s41467-020-20657-4>
 - [28] Matthew D. Li, Nishanth T. Arun, Mishka Gidwani, Ken Chang, Francis Deng, Brent P. Little, Dexter P. Mendoza, Min Lang, Susanna I. Lee, Aileen O'Shea, Anushri Parakh, Praveer Singh, and Jayashree Kalpathy-Cramer. 2020. Automated Assessment and Tracking of COVID-19 Pulmonary Disease Severity on Chest Radiographs Using Convolutional Siamese Neural Networks. *Radiology: Artificial Intelligence* 2, 4 (July 2020), e200079. <https://doi.org/10.1148/ryai.2020200079>
 - [29] Geert Litjens, Thijs Kooi, Babak Ehteshami Bejnordi, Arnaud Arindra Adiyoso Setio, Francesco Ciompi, Mohsen Ghafoorian, Jeroen A. W. M. van der Laak, Bram van Ginneken, and Clara I. Sánchez. 2017. A survey on deep learning in medical image analysis. *Medical Image Analysis* 42 (Dec. 2017), 60–88. <https://doi.org/10.1016/j.media.2017.07.005>
 - [30] Yujin Oh, Sangjoon Park, and Jong Chul Ye. 2020. Deep Learning COVID-19 Features on CXR Using Limited Training Data Sets. *IEEE Transactions on Medical Imaging* 39, 8 (Aug. 2020), 2688–2700. <https://doi.org/10.1109/TMI.2020.2993291>
 - [31] Meng Qu, Yoshua Bengio, and Jian Tang. 2019. GMNN: Graph Markov Neural Networks. *arXiv e-prints* (2019). <https://arxiv.org/pdf/1905.06214.pdf>
 - [32] Joel Saltz, Mary Saltz, Prateek Prasanna, Richard Moffitt, Janos Hajagos, Erich Bremer, Joseph Balsamo, and Tahsin Kurc. 2021. Stony Brook University COVID-19 Positive Cases. <https://doi.org/10.7937/TCIA.BBAG-2923> Type: dataset.
 - [33] Weiya Shi, Xueqing Peng, Tiejie Liu, Zenghui Cheng, Hongzhou Lu, Shuyi Yang, Jiulong Zhang, Mei Wang, Yaozong Gao, Yuxin Shi, Zhiyong Zhang, and Fei Shan. 2021. A deep learning-based quantitative computed tomography model for predicting the severity of COVID-19: a retrospective study of 196 patients. *Annals of Translational Medicine* 9, 3 (Feb. 2021), 216. <https://doi.org/10.21037/atm-20-2464>
 - [34] Pizer Stephen M., Amburn E. Philip, Austin John D., Cromartie Robert, Geselowitz Ari, Greer Ari, ter Haar Romeny Bart, Zimmerman John B, and Karel Zuiderveld. 1987. Adaptive histogram equalization and its variations. *Computer Vision, Graphics, and Image Processing* 39 (Sept. 1987), 355–368. [https://doi.org/10.1016/S0734-189X\(87\)80186-X](https://doi.org/10.1016/S0734-189X(87)80186-X)
 - [35] Shuo Wang, Yunfei Zha, Weimin Li, Qingxia Wu, Xiaohu Li, Meng Niu, Meiyun Wang, Xiaoming Qiu, Hongjun Li, He Yu, Wei Gong, Yan Bai, Li Li, Yongbei Zhu, Liusu Wang, and Jie Tian. 2020. A fully automatic deep learning system for COVID-19 diagnostic and prognostic analysis. *European Respiratory Journal* 56, 2 (Aug. 2020). <https://doi.org/10.1183/13993003.00775-2020>
 - [36] Tianyu Wang, Jun Bai, and Sheida Nabavi. 2021. Single-cell classification using graph convolutional networks. *BMC Bioinformatics* 22, 1 (July 2021), 364. <https://doi.org/10.1186/s12859-021-04278-2>
 - [37] Tingyi Wanyan, Hossein Honarvar, Suraj K. Jaladanki, Chengxi Zang, Nidhi Naik, Sulaiman Somani, Jessica K. De Freitas, Ishan Paranjpe, Akhil Vaid, Jing Zhang, Riccardo Miotto, Zhangyang Wang, Girish N. Nadkarni, Marinka Zitnik, Ariful Azad, Fei Wang, Ying Ding, and Benjamin S. Glicksberg. 2021. Contrastive learning improves critical event prediction in COVID-19 patients. *Patterns* 2, 12 (Dec. 2021), 100389. <https://doi.org/10.1016/j.patter.2021.100389>
 - [38] Wentao Zhao, Wei Jiang, and Xinguo Qiu. 2021. Deep learning for COVID-19 detection based on CT images. *Scientific Reports* 11, 1 (July 2021), 14353. <https://doi.org/10.1038/s41598-021-93832-2>