



OPEN ACCESS

EDITED BY

Chuhan Wu,
Tsinghua University, China

REVIEWED BY

Tao Qi,
Tsinghua University, China
Peiru Yang,
Tsinghua University, China

*CORRESPONDENCE

Zhuo Li
lizhuo@cqupt.edu.cn
Yifei Shen
sam.yifei.shen@gmail.com

SPECIALTY SECTION

This article was submitted to
Infectious Diseases - Surveillance,
Prevention and Treatment,
a section of the journal
Frontiers in Public Health

RECEIVED 30 June 2022

ACCEPTED 12 October 2022

PUBLISHED 22 November 2022

CITATION

Li Z, Xu R, Shen Y, Cao J, Wang B,
Zhang Y and Li S (2022) A multistage
multimodal deep learning model for
disease severity assessment and early
warnings of high-risk patients of
COVID-19.

Front. Public Health 10:982289.

doi: 10.3389/fpubh.2022.982289

COPYRIGHT

© 2022 Li, Xu, Shen, Cao, Wang,
Zhang and Li. This is an open-access
article distributed under the terms of
the [Creative Commons Attribution
License \(CC BY\)](#). The use, distribution
or reproduction in other forums is
permitted, provided the original
author(s) and the copyright owner(s)
are credited and that the original
publication in this journal is cited, in
accordance with accepted academic
practice. No use, distribution or
reproduction is permitted which does
not comply with these terms.

A multistage multimodal deep learning model for disease severity assessment and early warnings of high-risk patients of COVID-19

Zhuo Li^{1,2*}, Ruiqing Xu³, Yifei Shen^{2*}, Jiannong Cao²,
Ben Wang⁴, Ying Zhang⁵ and Shikang Li⁵

¹School of Computer Science and Technology, Chongqing University of Posts and Telecommunications, Chongqing, China, ²Department of Computing, The Hong Kong Polytechnic University, Kowloon, Hong Kong SAR, China, ³Department of Computer Science, The University of Sheffield, Sheffield, United Kingdom, ⁴Baoding No. 2 Central Hospital, Baoding, China, ⁵Chongqing Public Health Medical Center, Chongqing, China

The outbreak of coronavirus disease 2019 (COVID-19) has caused massive infections and large death tolls worldwide. Despite many studies on the clinical characteristics and the treatment plans of COVID-19, they rarely conduct in-depth prognostic research on leveraging consecutive rounds of multimodal clinical examination and laboratory test data to facilitate clinical decision-making for the treatment of COVID-19. To address this issue, we propose a multistage multimodal deep learning (MMDL) model to (1) first assess the patient's current condition (i.e., the mild and severe symptoms), then (2) give early warnings to patients with mild symptoms who are at high risk to develop severe illness. In MMDL, we build a sequential stage-wise learning architecture whose design philosophy embodies the model's predicted outcome and does not only depend on the current situation but also the history. Concretely, we meticulously combine the latest round of multimodal clinical data and the decayed past information to make assessments and predictions. In each round (stage), we design a two-layer multimodal feature extractor to extract the latent feature representation across different modalities of clinical data, including patient demographics, clinical manifestation, and 11 modalities of laboratory test results. We conduct experiments on a clinical dataset consisting of 216 COVID-19 patients that have passed the ethical review of the medical ethics committee. Experimental results validate our assumption that sequential stage-wise learning outperforms single-stage learning, but history long ago has little influence on the learning outcome. Also, comparison tests show the advantage of multimodal learning. MMDL with multimodal inputs can beat any reduced model with single-modal inputs only. In addition, we have deployed the prototype of MMDL in a hospital for clinical comparison tests and to assist doctors in clinical diagnosis.

KEYWORDS

COVID-19, disease severity assessment, disease progression prediction, sequential stage-wise learning, multimodal feature fusion

1. Introduction

Since December 2019, a novel viral pneumonia caused by severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2), also known as coronavirus disease 2019 (COVID-19) (1–3), first occurred in Wuhan, Hubei Province, China (4), then swept the globe very quickly. As of Sep 1, 2022, data from the World Health Organization (WHO) revealed more than 600 million infections confirmed worldwide with approximately 6.45 million deaths since the outbreak of COVID-19 (5). In view of its strong infectivity and high mortality, WHO declared the pandemic as a Public Health Emergency of International Concern (6).

In practice, the clinical manifestations of COVID-19 vary diversely from asymptomatic, mild infection to severe symptoms (4, 7–9). According to clinical statistics, the majority of COVID-19 cases are mild, and only approximately 5% of the total patients (a part of severe cases) require admission to ICU (10, 11). One of the serious problems we are facing is that the surge of COVID-19 infections leads to rapid depletion of the limited medical resources. The fact is that most of the mild patients can heal without supportive treatment (2, 3), and only a small proportion of them will progress toward severe illness. However, patients whose condition subsequently deteriorate are more prone to be older adults with comorbidities of diabetes, hypertension, cardiac disease, obesity etc. (9, 12). Once the illness changes for the worse, the mortality rate increases significantly, moreover, treating critical patients consumes more medical resources and takes longer treatment courses.

During COVID-19 treatment, doctors perform clinical examinations and laboratory tests on patients every few days. Hence, in every round of the tests, massive multimodal (i.e., various types or categories) clinical data are generated, including the patient demographics, clinical manifestation, laboratory outcomes, the use of drugs and medication, etc. Naturally, it is of great importance that we quickly and accurately distinguish mild and severe patients on admission, then identify those mild cases who are at high risk of turning for the worse in the future based on clinical data analysis and modeling. As a result, early intervention can be taken to prevent mild patients from deterioration.

In the past decade, AI and big data technologies have been widely applied in healthcare and medication and made remarkable achievements (13), which also play an important role in COVID-19 prevention and containment, including screening, testing, contact tracing, treatment and vaccination, and drug development (14–16). So far various forecasting models have been developed for the diagnosis and prognosis of COVID-19 (17–19), which leveraged X-ray and CT images (20–22), clinical characteristics (23, 24), blood test results (25), etc., for model development.

Most of the existing literature for the diagnosis and prognosis of COVID-19 simply makes use of one or two modalities of clinical data, which fails to explore the complementary information provided by multimodal sources.

Moreover, the prognostic model is purely based on a single round of lab test results and cannot track the disease progression since onset. To address the characteristics of the consecutive rounds of multimodal clinical test data, in this paper, we propose a multistage multimodal deep learning (MMDL) model to (1) first assess the disease severity, and (2) identify those who are at an early stage of illness and are likely to grow worse. In MMDL, we conceive and implement a sequential stage-wise learning architecture, which abandons the classic structure of RNN (26, 27)/LSTM (28). It is because most patients take no more than five rounds of exams and lab tests before they recover from COVID-19 and are discharged from the hospital, so if we insist on using the RNN/LSTM (Recurrent Neural Network/Long Short-Term Memory) model, the input time step of RNN/LSTM is too few to forecast the future. The design philosophy of MMDL is motivated by the sequence-to-sequence (seq2seq) model (29, 30) in contextual sequence prediction, which extracts the latent feature of one sequence (encoder) and turns it to another sequence (decoder), then the decoded word in a sentence is based on the output from its previous contexts. Concretely, the embodiment of sequential stage-wise learning incorporates the input of the latest round of multimodal clinical data and the past information, and higher weights are given to recent inputs because it has direct influences on the final assessment and prediction results. In each round, to extract the feature of the multimodal clinical data, we design a two-layer multimodal feature extractor: in the **1st**-hierarchy, we build multiple separate fully-connected multi-layer perceptron (MLP) neural networks sharing the same network architecture, and each MLP extracts the intra-modal latent feature of an independent modality of clinical data; in the **2nd**-hierarchy, the extracted latent features of all modalities are concatenated, then input to another similar MLP for cross-modal feature fusion.

Extensive experiments are conducted on a dataset consisting of 216 patients diagnosed with COVID-19, which has passed the review of the medical ethics committee and can be used for research purposes only. These patients were admitted to the Public Health Medical Center in Chongqing, China, and received intensive medical care. The experimental results of the prognostic study show the advantage of sequential stage-wise learning of MMDL over conventional single-stage learning. In addition, the results also prove that MMDL with multimodal inputs can surpass the reduced model with any single-modal clinical data input by a large margin, particularly for the severe group in disease severity assessment and the mild-to-severe incidence group in the disease progression prediction.

2. Dataset description

2.1. Patient demographics

We retrospectively review the medical records of 216 patients with COVID-19 who were admitted to the Public Health Center in Chongqing, China from January 24, 2020,

to February 16, 2020. These patients were admitted fulfilling the following criteria: (1) tested positive with two consecutive nucleic acid tests; (2) showed distinct characteristics of pneumonia in CT images.

Figure 1 shows the demographics of these admitted patients that 103 cases (47.69%) out of the total number were male patients while female patients occupied 52.31% (113 cases). Depending upon the patient's severity of symptoms, 186 cases (86.11%) and 30 cases (13.89%) are diagnosed with mild and severe symptoms, respectively. By age, patients aged between 41 and 50 are the largest group with 50 cases (23.15%), which is followed by the group 31–40 and the group 60+ accounting for 20.83% (45 cases) each. Patients under the age of 18 and aged 19–30 only make up 5.09% (9 cases) and 11.57% (25 cases), respectively. Moreover, Figure 1d shows the average duration of onset of symptoms to hospital admission. A total of 76 and 20% of the patients were admitted in the first and the second week, respectively, since the onset of the disease. The remaining 4% of the patients developed symptoms after 2 weeks.

Figure 2 reveals the top 10 clinical manifestations of the 216 patients with COVID-19 infections. As it shows, cough [135 cases (62.50%)], fever [108 cases (50.00%)], and expectoration [68 cases (31.48%)] are reported as the most typical symptoms. It records 53 cases (24.54%) of fatigue and 38 cases (17.59%) of shortness of breath, which are another two common clinical manifestations. In addition, about 35 patients show no symptoms on admission.

2.2. Multimodal clinical data

During the COVID-19 treatment, numerous patients' clinical data are produced including patients' vital signs, laboratory test results, CT image findings, medical experts' diagnoses, and corresponding treatment plans. Among these clinical data, laboratory test results comprise 11 different categories, which is termed "multimodal" in the context of big data and machine learning. Specifically, the 11 modalities are named blood test, flow cytometry, inflammation, liver function, renal function, blood lipids, glucose, electrolyte, myocardial zymogram and heart failure indicator, coagulation, and arterial blood gas. Each modality contains many laboratory test items. For example, the blood test modality consists of white blood cell count (WBC), red blood cell count (RBC), neutrocyte count (NEUT#), monocytes count (MONO#), lymphocyte count (LYMPH#), etc., and the inflammation modality contains erythrocyte sedimentation rate (ESR), C-reactive protein (CRP) and hypersensitive C-reactive protein (hs-CRP), and procalcitonin (PCT).

Table 1 describes the statistical results of the 11 modalities of laboratory tests of the mild group, the severe group, and the total population below:

3. Summary of notations

All the notations used in this paper are summarized below:

\mathcal{F}_Θ	MMDL model for disease severity assessment with network parameters Θ ;
\mathcal{F}_Φ	MMDL model for disease progression prediction with network parameters Φ ;
$\mathbf{X}^{(n)}$	A matrix. $\mathbf{X}^{(n)} = [X_1^{(n)}, \dots, X_k^{(n)}, \dots, X_K^{(n)}]$ is the input multimodal clinical data of the n -th stage;
$X_k^{(n)}$	A vector. The k -th modality clinical input data of the n -th stage;
$X_k'^{(n)}$	A vector. The extracted latent feature of k -th modality of the n -th stage;
$X_{\text{CAT}}^{(n)}$	A matrix. $X_{\text{CAT}}^{(n)} = [X_1'^{(n)}, \dots, X_k'^{(n)}, \dots, X_K'^{(n)}]$ is the concatenation of the extracted intra-modal latent features of all different modalities of the n -th stage;
N	A scalar. The total rounds (stages) of the performed clinical examination and lab tests;
K	A scalar. The total number of input modalities, $K = 13$, including the patient demographics, clinical manifestation and laboratory test results (e.g., blood test, inflammation, liver function, renal function, blood lipids, etc.);
\mathbf{D}	A vector. The extracted latent feature of patient demographics modality;
$\mathbf{Z}^{(n)}$	A vector. The extracted cross-modal feature representation of the n -th round clinical manifestation modality and other 11 laboratory test modalities;
$\mathbf{S}^{(n)}$	A vector. The intermediate learning outcome of the n -th stage;
\mathbf{W}_D	A matrix. The weighting matrix multiplying with \mathbf{D} ;
$\mathbf{W}_Z^{(n)}$	A matrix. The weighting matrix multiplying with $\mathbf{Z}^{(n)}$;
$\mathbf{W}_S^{(n)}$	A matrix. The weighting matrix multiplying with $\mathbf{S}^{(n)}$;
$\mathbf{b}_S^{(n)}$	A vector. The bias vector added to the computed results at stage n ;
α	A scalar. $\alpha \in [0, 1]$ is a decay factor, by multiplying with which the learning outcome of the previous stage $\mathbf{S}^{(n)}$ is attenuated every round;
\bar{y}	A vector. The output vector for computing \bar{y} ;
\bar{y}	A scalar. The obtained result of either disease severity assessment ($\bar{y} \in \{\text{mild}, \text{severe}\}$) or disease progression prediction ($\bar{y} \in \{\text{not-develop-severe}, \text{develop-severe}\}$);
y	A scalar. The corresponding ground truth label;
MLP	The fully connected multilayer perceptron neural network;
$ReLU$	The rectified linear unit activation function;

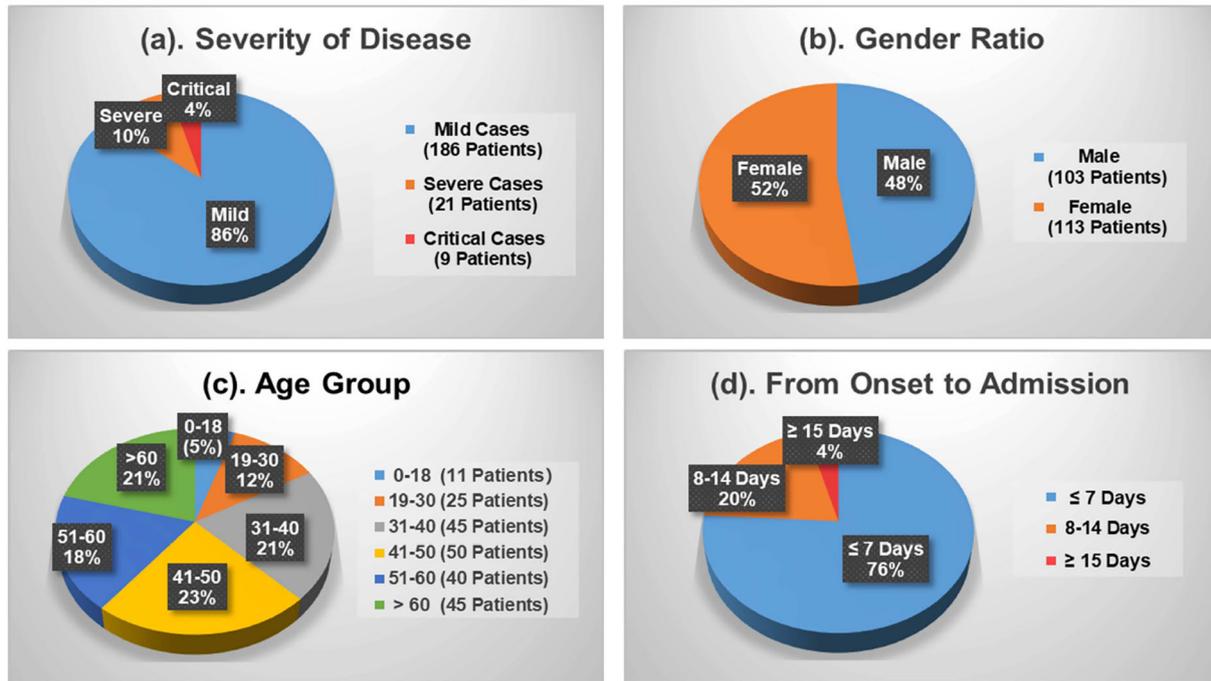


FIGURE 1 Demographics of the COVID-19 patients contained in the dataset. (a) Severity of disease. (b) Gender ratio. (c) Age group. (d) From onset to admission.

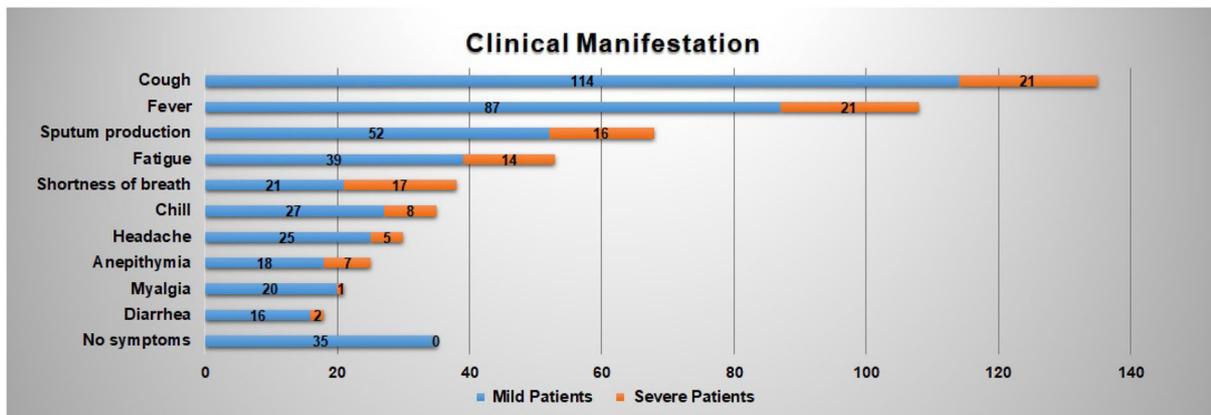


FIGURE 2 Clinical manifestation of the COVID-19 patients contained in the dataset.

Softmax The softmax multi-class classifier;
 $\mathcal{L}_{\text{cross-entropy}}$ The cross-entropy loss function.

4. Problem formulation

Given patients infected by COVID-19 take N rounds of clinical examination and laboratory tests in total during the treatment. In each round, K different modalities of clinical

data (e.g., clinical manifestation, blood test, inflammation, liver function, etc.) are collected for disease assessment and prediction. Take the n -th round as an example, the notation $\mathbf{X}^{(n)}$ is used to denote the stage-wise multimodal input:

$$\mathbf{X}^{(n)} = [X_1^{(n)}, X_2^{(n)}, \dots, X_k^{(n)}, \dots, X_K^{(n)}], \quad (1)$$

where $1 \leq n \leq N$ and $1 \leq k \leq K$.

By leveraging the stage-wise multimodal clinical data, $\mathbf{X}^{(1)}, \mathbf{X}^{(2)}, \dots, \mathbf{X}^{(n)}, \dots, \mathbf{X}^{(N)}$, our goal is to develop a model \mathcal{F} to: (1) assess the disease severity of patients diagnosed with COVID-19, and (2) forecast mild cases who have a high risk of progressing to critical illness. The two tasks share the same network architecture but are trained separately with two different sets of network parameters.

Mathematically, in the disease severity assessment task, it can be expressed as:

$$\bar{y}_{\bar{y} \in \{mild, severe\}} = \mathcal{F}_{\Theta}(\mathbf{X}^{(1)}, \mathbf{X}^{(2)}, \dots, \mathbf{X}^{(n)}, \dots, \mathbf{X}^{(N)} | \Theta), \quad (2)$$

where $\bar{y} \in \{mild, severe\}$ is used to denote the obtained result of assessment, \mathcal{F}_{Θ} represents the MMDL model with the parameter Θ that maps the multistage input to the output \bar{y} .

Similarly, in the disease progression prediction task, it can be written as:

$$\bar{y}_{\bar{y} \in \{not-develop-severe, develop-severe\}} = \mathcal{F}_{\Phi}(\mathbf{X}^{(1)}, \mathbf{X}^{(2)}, \dots, \mathbf{X}^{(n)}, \dots, \mathbf{X}^{(N)} | \Phi), \quad (3)$$

Likewise, $\bar{y}_{\bar{y} \in \{not-develop-severe, develop-severe\}}$ is the predicted results and \mathcal{F}_{Φ} is the corresponding prediction model with network parameters Φ .

5. Multistage multimodal deep learning model

In this section, we introduce the multistage, multimodal deep learning (MMDL) model in detail. We first illustrate the sequential stage-wise learning framework, then present the feature extraction of the multimodal clinical data at each stage, and finally come to the end-to-end training of MMDL.

5.1. Sequential stage-wise learning

Sequential stage-wise learning and sequence prediction share some common ground, although they are different in some respects. The sequence-to-sequence (seq2seq) model (29, 30) is one of the classical benchmarks in contextual sequence prediction. It transforms one sequence into another sequence, and the context of the decoded sentence is based on the output from its previous contexts.

Motivated by this, we propose the sequential stage-wise learning architecture of MMDL, which is illustrated in Figure 3. As we can see, it meticulously joins the extracted cross-modal latent feature of the previous stage and the current stage, then concatenates the result with the extracted multimodal feature of the next stage sequentially for further processing.

In the first stage, the model takes the patient demographics \mathbf{D} and the initial examination and laboratory test results $\mathbf{Z}^{(1)}$

when admitted to the hospital as the model input. It should be noted that \mathbf{D} is the extracted latent feature of the patient demographics modality only, and \mathbf{Z} is the merged multimodal feature representation across all different modalities of lab test results. How \mathbf{Z} is extracted and merged will be justified in the next subsection in detail.

\mathbf{D} and \mathbf{Z} first multiply with \mathbf{W}_D and $\mathbf{W}_Z^{(1)}$ respectively, and add together, then pass through the *ReLU* activation function.

$$\mathbf{S}^{(1)} = \text{ReLU}(\mathbf{W}_D \mathbf{D} + \mathbf{W}_Z^{(1)} \mathbf{Z}^{(1)} + \mathbf{b}_S^{(1)}), \quad (4)$$

where $\mathbf{S}^{(1)}$ is the output of the first stage. \mathbf{W}_D and $\mathbf{W}_Z^{(1)}$ are the weight matrices, and $\mathbf{b}_S^{(1)}$ is the bias term. *ReLU* is short for rectified linear units (31), which can be expressed as $\text{ReLU}(x) = \max(0, x)$ and is a non-linear activation function.

Then $\mathbf{b}_S^{(1)}$ multiplies with the weight $\mathbf{W}_S^{(1)}$ accordingly and adds to the result of the multiplication of $\mathbf{Z}^{(2)}$ and $\mathbf{W}_Z^{(2)}$. Similarly, we can derive the expression of the learning process of the second stage:

$$\mathbf{S}^{(2)} = \text{ReLU}(\mathbf{W}_Z^{(2)} \mathbf{Z}^{(2)} + \alpha \mathbf{W}_S^{(1)} \mathbf{S}^{(1)} + \mathbf{b}_S^{(2)}), \quad (5)$$

in which $\mathbf{Z}^{(2)}$ is the extracted cross-modal latent feature of the second stage and $\mathbf{W}_Z^{(2)}$ is the corresponding weight. It should be mentioned that $\mathbf{W}_S^{(1)} \mathbf{S}^{(1)}$ is attenuated by multiplying with a decay factor $\alpha \in [0, 1]$ because the new round of test results has a direct impact on the predicted results while the influence of the test result obtained long ago weakens as time passes.

Note that the modality of patient demographics is incorporated into the model in the initial stage only since the patient demographics modality contains patients' basic information, like gender and age, that does not change in every round of test.

For the n -th stage ($2 \leq n \leq N$), a more general form can be written as:

$$\mathbf{S}^{(n)} = \text{ReLU}(\mathbf{W}_Z^{(n)} \mathbf{Z}^{(n)} + \alpha \mathbf{W}_S^{(n-1)} \mathbf{S}^{(n-1)} + \mathbf{b}_S^{(n)}), \quad (6)$$

Finally, the learned representation of the last stage $\mathbf{W}_S^{(N)}$ is further fused to get the output vector \bar{Y} :

$$\bar{Y} = \mathbf{W}_S^{(N)} \mathbf{S}^{(N)} + \mathbf{b}'_S, \quad (7)$$

and pass it through a multi-class Softmax classifier to get the predicted outcome \bar{y} .

$$\bar{y} = \text{Softmax}(\bar{Y}), \quad (8)$$

\bar{y} is the output scalar, which is either *mild* or *severe* in the first task and *not-develop-severe* or *develop-severe* in the second.

TABLE 1 Characteristics of multimodal lab test results of the COVID-19 patients contained in the dataset.

Characteristics	All patients N = 216	Mild patients N = 186 (86.11%)	Severe patients N = 30 (13.89%)
Blood test			
White blood cell (WBC), $\times 10^9/L$	6.03	5.32	6.47
Neutrophils (NEUT), $\times 10^9/L$	4.05	3.32	4.98
Lymphocyte (LYMPH), $\times 10^9/L$	1.34	1.43	0.86
Monocytes (MONO), $\times 10^9/L$	0.40	0.42	0.36
Eosinophils (EO), $\times 10^9/L$	0.06	0.06	0.06
Basophils (BASO), $\times 10^9/L$	0.02	0.02	0.02
Red blood cell (RBC), $\times 10^{12}/L$	4.20	4.25	4.00
Hemoglobin (HGB), g/L	128	130	124.3
Hematocrit (HCT), L/L	38.80	39.10	37.57
Mean corpuscular hemoglobin concentration (MCHC), g/L	331.6	332	331.2
Platelet (PLT), $\times 10^9/L$	229	216	244.8
Mean platelet volume (MPV), fL	9.40	9.30	9.62
Platelet hematocrit (PCT), (%)	0.21	0.20	0.23
Flow cytometry			
Absolute CD3+ T lymphocyte, $cells/\mu L$	729	829	461
Absolute CD4+ T lymphocyte, $cells/\mu L$	420	451	259
Absolute CD8+ T lymphocyte, $cells/\mu L$	281	316	145
CD4+ / CD8+ ratio	1.43	1.38	1.47
Inflammation			
Erythrocyte sedimentation rate (ESR), mm/h	46.6	38.2	71.3
C-reactive protein (CRP), mg/L	26.2	17.3	65
Hypersensitive C-reactive protein (hsCRP), mg/L	32.5	21.9	71.1
Procalcitonin (PCT), ng/L	0.096	0.040	0.331
Liver function			
Prealbumin (PA), $\mu g/dL$	221	227	200
α -L-Fucosidase (AFU), U/L	27.8	27.4	29.7
Alanine aminotransferase (ALT), U/L	34	28.1	59.3
Aspartate aminotransferase (AST), U/L	27.8	24.6	41.5
Alkaline phosphatase (ALP), IU/L	58.9	57.2	66.1
Gamma-glutamyltransferase (GGT), IU/L	43.3	32.2	90.7
Lactate dehydrogenase (LDH), IU/L	235	209	343
Total protein (TP), g/L	66.7	67.3	64.5
Albumin (ALB), g/L	40.2	41.2	35.5
Globulin (GLB), g/L	27.3	26.8	29.5
A/G Ratio	1.40	1.50	1.24
Total bilirubin (TBIL), $\mu mol/L$	15.6	15.8	15.1
Total bile acid (TBA), $\mu mol/L$	3.00	3.10	2.34
Renal function			
UREA, $mmol/L$	3.83	3.71	4.10
Creatinine (CREA), $\mu mol/L$	66.9	67.2	65.9
Uric acid (UA), $\mu mol/L$	303	320	231
Beta 2-microglobulin ($\beta 2 - M$), mg/L	2.19	2.16	2.38
Cystatin C (CysC), mg/L	0.99	0.96	1.16

(Continued)

TABLE 1 (Continued)

Characteristics	All patients N = 216	Mild patients N = 186 (86.11%)	Severe patients N = 30 (13.89%)
Glucose modality			
Glucose (hexokinase (HK) method), <i>mmol/L</i>	6.41	6.11	7.73
Blood lipids			
Triglyceride (TG), <i>mmol/L</i>	2.23	2.11	2.69
Total cholesterol (CHOL), <i>mmol/L</i>	4.39	4.28	4.26
High-density lipoprotein (HDL), <i>mmol/L</i>	1.02	1.03	0.97
Low-density lipoprotein (LDL), <i>mmol/L</i>	2.47	2.42	2.38
Electrolyte			
Potassium (K), <i>mmol/L</i>	4.21	4.24	4.00
Sodium (NA), <i>mmol/L</i>	138.2	138.4	137.8
Chlorine (CL), <i>mmol/L</i>	102.7	103.0	101.5
Calcium (CA), <i>mmol/L</i>	2.24	2.26	2.13
Phosphorus (P), <i>mmol/L</i>	1.08	1.11	0.98
Magnesium (MG), <i>mmol/L</i>	0.89	0.88	0.93
Coagulation function			
Prothrombin time (PT), <i>seconds</i>	11.77	11.76	11.81
International normalized ratio (INR)	1.00	1.01	0.96
Activated partial thromboplastin time (APTT), <i>seconds</i>	38.50	38.40	39.25
Thrombin time (TT), <i>seconds</i>	14.90	14.70	15.15
Fibrinogen (FIB), <i>g/L</i>	4.25	4.20	4.62
D-dimer, <i>mg/L</i>	0.68	0.44	1.64
Myocardial zymogram & Heart failure			
Adenosine deaminase (ADA), <i>U/L</i>	14.10	13.85	14.36
Creatine kinase (CK), <i>U/L</i>	100	88	151
α -Hydroxybutyrate dehydrogenase (α -HBDH), <i>IU/L</i>	179	160	262
5'-Nucleotidase (5'-NT), <i>U/L</i>	4.38	3.94	6.25
Cholinesterase (CHE), <i>U/L</i>	7960	8409	6699
Arterial blood gas			
Arterial blood pH	7.42	7.41	7.45
Partial pressure of oxygen (PaO ₂), <i>mmHg</i>	85	85	86
Partial pressure of carbon dioxide (PaCO ₂), <i>mmHg</i>	41	41	40
Bicarbonate (HCO ₃ ⁻), <i>mEq/L</i>	26.10	26.00	26.95
Oxygen saturation (SaO ₂), (%)	96.90%	97.30%	96.15%

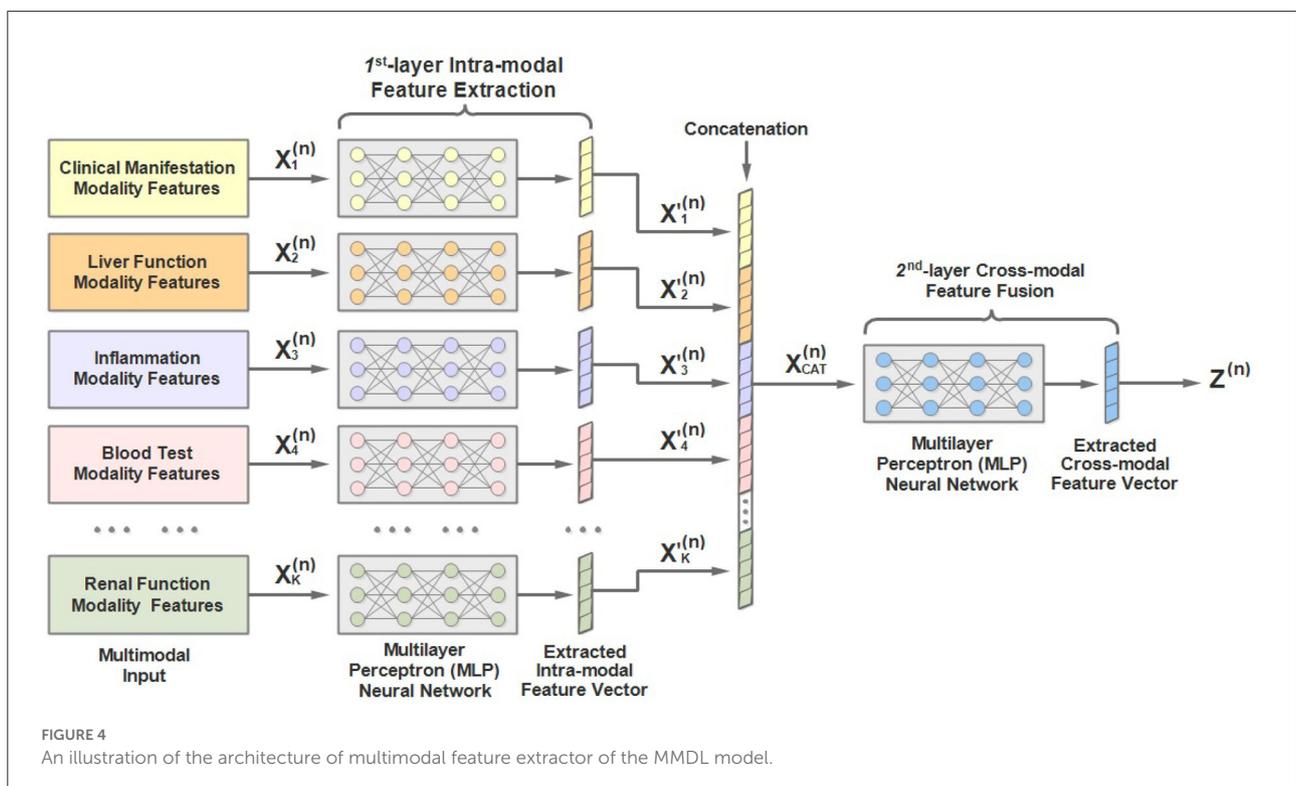
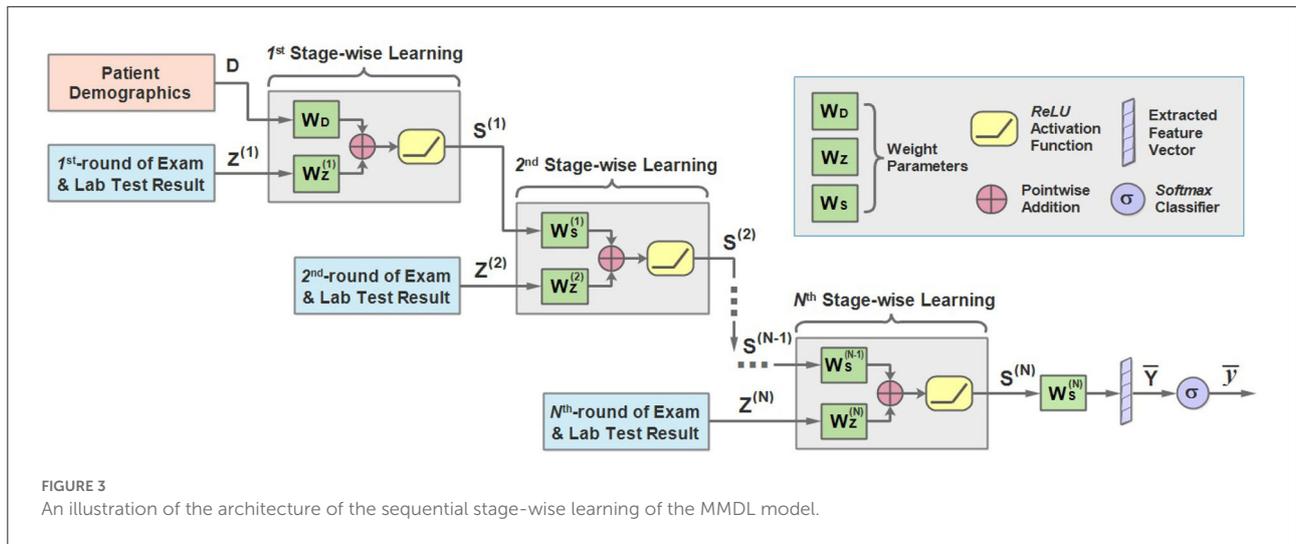
5.2. Multimodal feature extraction

As mentioned above, the input to each stage is multimodal lab test results. To address the characteristics of the multimodal input data, we would like to introduce a two-layer multimodal feature extractor conceived by us and the architecture of which is shown in Figure 4. As we can see, the 1st-hierarchy aims to perform intra-modal feature learning and extraction, while the 2nd-hierarchy attempts to perform cross-modal feature fusion.

Concretely, in the 1st hierarchy, we build up K ($K = 13$) independent multilayer perceptron (MLP) neural networks. Each MLP is responsible for extracting the latent feature of a separate input modality, including the patient demographics, clinical manifestation, and 11 other modalities of laboratory test data.

For example, the extracted feature of the k -th modality of the n -th round test can be expressed as:

$$X_k'^{(n)} = MLP(X_k^{(n)}), \quad (9)$$



where $X_k'^{(n)}$ denotes the extracted latent feature vector of the k -th modality ($1 \leq k \leq K$), and MLP represents the multi-layer fully-connected neural network.

Afterwards, the extracted intra-modal feature vectors of all modalities $X_1'^{(n)}, X_2'^{(n)}, \dots, X_k'^{(n)}, \dots, X_K'^{(n)}$ are concatenated:

$$X_{CAT}^{(n)} = [X_1'^{(n)}, X_2'^{(n)}, \dots, X_k'^{(n)}, \dots, X_K'^{(n)}]. \quad (10)$$

Finally, the concatenated feature vectors $X_{CAT}^{(n)}$ are further processed by another MLP to obtain the fused cross-modal

feature representation $Z^{(n)}$, which is then taken as the input of the n -th stage of MMDL:

$$Z^{(n)} = MLP(X_{CAT}^{(n)}), \quad (11)$$

5.3. Model training

Before training the model, we have to define the loss function in the first place, which gives the learning objective

during the training process. The loss function compares the difference between the predicted results \bar{y} and the ground truth labels y given by medical experts, and a smaller value of \mathcal{L} means the model's performance is better. Either the patient disease assessment or the disease progression prediction can be regarded as a classification problem, hence we choose cross-entropy as the loss function, which is widely used in multi-class classification problems:

$$\mathcal{L}_{\text{cross-entropy}} = - \sum_j y_j \cdot \log(p(\bar{y}=j)), \quad (12)$$

where j represents the predicted class, and $j = \{\text{mild, severe}\}$ for the disease severity assessment and $j = \{\text{not-develop-severe, develop-severe}\}$ for disease progression prediction. $p(\hat{y}=j)$ is the predicted probability of the class j using *Softmax*, i.e.,

$$p(\bar{y}=j) = \text{Softmax}(Y_j) = \frac{e^{Y_j}}{\sum_{i=1}^{\#class} e^{Y_i}}. \quad (13)$$

The notation *#class* represents the number of classes, and *#class* = 2 in our settings because there are two results for both the assessment and the prediction tasks.

During the training process, the end-to-end supervised learning is used to train MMDL. Adam optimizer is adopted to backpropagate the calculated loss to the input layer of the model, and all network parameters (weights and biases) are updated through iterative optimization. MMDL is trained two times separately to learn two different sets of network parameters, i.e., Θ and Φ , one for the disease severity assessment and the other for the disease progression prediction.

6. Experiments

In this section, we will present the experimental part of MMDL in detail. We first introduce how we set up the experiment, then the evaluation metrics, and finally present the comparison results for both tasks.

6.1. Experiment setup

In the experiment, we first assess the severity of illness of patients using different numbers of consecutive stages of multimodal inputs (*#Multistage Input*), then forecast whether patients with mild symptoms will progress to severe illness or not with different prediction steps (*Prediction Step*). To start with, we assess the severity of illness of patients using the initial exam and lab test data on admission, then identify patients diagnosed with mild symptoms who are prone to develop severe symptoms. Subsequently, we extend it to the scenario of the disease severity assessment using multistage input, i.e., use multiple successive rounds of clinical test data to assess the disease severity.

In addition, in view of the limited samples contained in the dataset, 10-fold cross validation is adopted, that is, in each round of training, 10% of the cases are randomly selected for testing and the remaining 90% cases are used for training, while in another round, another 10% cases are selected as the test set.

6.2. Evaluation metric

We use a group of evaluation metrics to evaluate the classification performance of MMDL, including accuracy, error rate, precision, recall, and F1 score, which are computed as follows:

$$\left\{ \begin{array}{l} \text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN}, \\ \text{Error Rate} = 1 - \text{Accuracy}, \\ \text{Precision} = \frac{TP}{TP + FP}, \\ \text{Recall} = \frac{TP}{TP + FN}, \\ \text{F1} = \frac{2}{\frac{1}{\text{Precision}} + \frac{1}{\text{Recall}}} = \frac{2 \times TP}{2 \times TP + FP + FN}, \end{array} \right. \quad (14)$$

where TP, FP, TN, and FN represent True Positive, False Positive, True Negative, and False Negative samples, respectively. The higher the value obtained, the better performance is achieved for all evaluation metrics but the error rate.

6.3. Results

6.3.1. Disease severity assessment with different numbers of multistage input

To show the advantage of learning with multistage data, we compare the performance of MMDL using a single stage's inputs (i.e., *#Multistage Input* = 1) and multiple successive stages' inputs (i.e., *#Multistage Input* > 1) on disease severity assessment (i.e., *Prediction Step* = 0).

Table 2 shows the obtained results. Accuracy is the proportion of the correctly classified samples (i.e., TP + TN) to the total number of samples, so the mild and the severe groups have the same accuracy, which increases from 96.26% using the current stage inputs only to 98.10% using five consecutive stages' inputs. Precision is the correct predictions (i.e., TP) out of all patients predicted to be infected (i.e., TP + FP), which grows from 97.52% with *#Multistage Input* = 1 to 98.69% with *#Multistage Input* = 5, respectively. Recall, which represents the percentage of truly predicted infections (i.e., TP) among all infections (i.e., TP + FN), goes from 98.33 to 99.60%. F1 is the harmonic mean of precision and recall, which is a more balanced evaluation metric to reflect the overall classification results. Moreover, we plot the curve depicting the change of F1 as the increase of the numbers of used multistage input in Figure 5.

TABLE 2 Performance comparison of MMDL model with different numbers of multistage inputs.

	#Multistage Input data	Prediction Step	Accuracy	Error Rate	Precision	Recall	F1 Score
Mild group	1	0	96.26%	3.73%	97.52%	98.33%	0.9792
	2	0	96.98%	3.01%	97.74%	98.86%	0.9830
	3	0	97.74%	2.25%	98.03%	99.40%	0.9871
	4	0	98.09%	1.90%	98.60%	99.50%	0.9887
	5	0	98.10%	1.89%	98.69%	99.60%	0.9890
Severe group	1	0	96.26%	3.73%	84.12%	77.94%	0.8091
	2	0	96.98%	3.01%	90.56%	82.75%	0.8648
	3	0	97.74%	2.25%	95.52%	86.48%	0.9078
	4	0	98.09%	1.90%	95.16%	92.18%	0.9365
	5	0	98.10%	1.89%	95.71%	93.05%	0.9436

We can see that in the training phase, the obtained results are all 100% for both the mild and severe groups, which reveals that MMDL fits the training set perfectly. In the testing phase, the F1 score of the mild and the severe groups increases from 0.9792 and 0.8091 by simply taking a single stage's inputs to 0.9890 and 0.9436 by considering all five successive stages' multimodal data, respectively.

6.3.2. Prediction of disease progression with different prediction steps

In this subsection, we would like to forecast progression from mild to severe COVID-19. First, it gives a brief introduction to the labels of the dataset. Medical experts assess patients' status after every round of the exam and lab test, which is treated as the ground truth labels of the prediction task of that stage. We set *Prediction Step* = 1 if we want to predict the patient's condition after the next round's test, and *Prediction Step* = 4 if we predict the patient's situation four stages ahead.

Table 3 describes the predicted results as the increase of the *Prediction Step*. The accuracy, precision, and recall of the mild group are 96.26, 97.52, and 98.33%, respectively, when *Prediction Step* = 0, then gradually decrease to 93.44, 95.74, and 97.12% when predicting the state of the illness of patients four stages away from now (*Prediction Step* = 4). For the severe group, these figures start from 96.26, 84.12, and 77.94%, then drop rapidly and finally stop at 93.44, 48, and 36.36%, respectively. Figure 6, left, right plot the curves of the F1 score of the mild-to-mild and mild-to-severe progressions as the increase of *Prediction Step*. In the testing phase, the F1 of the mild-to-mild incidence decreases from 0.9792 (*Prediction Step* = 0) to 0.9653 (*Prediction Step* = 4) gradually. But the situation worsens when predicting progression from mild to severe COVID-19 that F1 begins at 0.8091 (*Prediction Step* = 0), then declines dramatically to 0.5957 (*Prediction Step* = 1) and 0.5098

(*Prediction Step* = 2), then continues to decrease and finally stops at 0.4137 for *Prediction Step* = 4.

6.3.3. The ROC and AUC of disease severity assessment and prediction of disease progression

Figure 7 shows the receiver operating characteristic (ROC) and the area under the curve (AUC) of disease severity assessment (left) with #Multistage Input = 1 and *Prediction Step* = 0 and prediction of disease progression (right) with #Multistage Input = 1 and *Prediction Step* = 4. An ROC curve is a graph showing the performance of a classification model at different classification thresholds. The *x*-axis is FPR (False Positive Rate) which is calculated as $FPR = \frac{FP}{FP+TN}$, and the *y*-axis is TRP (True Positive Rate), also known as recall, which is computed as $TRP = \frac{TP}{TP+FN}$. Lowering the classification threshold classifies more items as positive, thus increasing both False Positives (FP) and True Positives (TP).

As Figure 7 (left) illustrates, the blue and red curves represent the ROC of the mild group and the severe group, respectively, and the green curve is the arithmetical average of them. The blue curve approaches the top-left corner, which means the MMDL model is a good classifier for distinguishing mild cases out of all infections. The classification result of severe cases is worse than that of the mild cases as the red curve is not as steep as the blue one in the beginning until FPR equals around 0.3, then the red curve approaches the blue one. In the right subfigure of Figure 7, it shows the ROC and AUC of patients' progression from mild to severe infection (denoted by the light blue curve) and of the ones that deteriorate (denoted by the red curve). As the figure shows, the achieved results of the MMDL model for predicting disease progression are not as good as those for assessing the patient's status. We can observe a distinct plateau region in Figure 7, right where both the blue and red curves do not go up. It means the MMDL model has difficulty

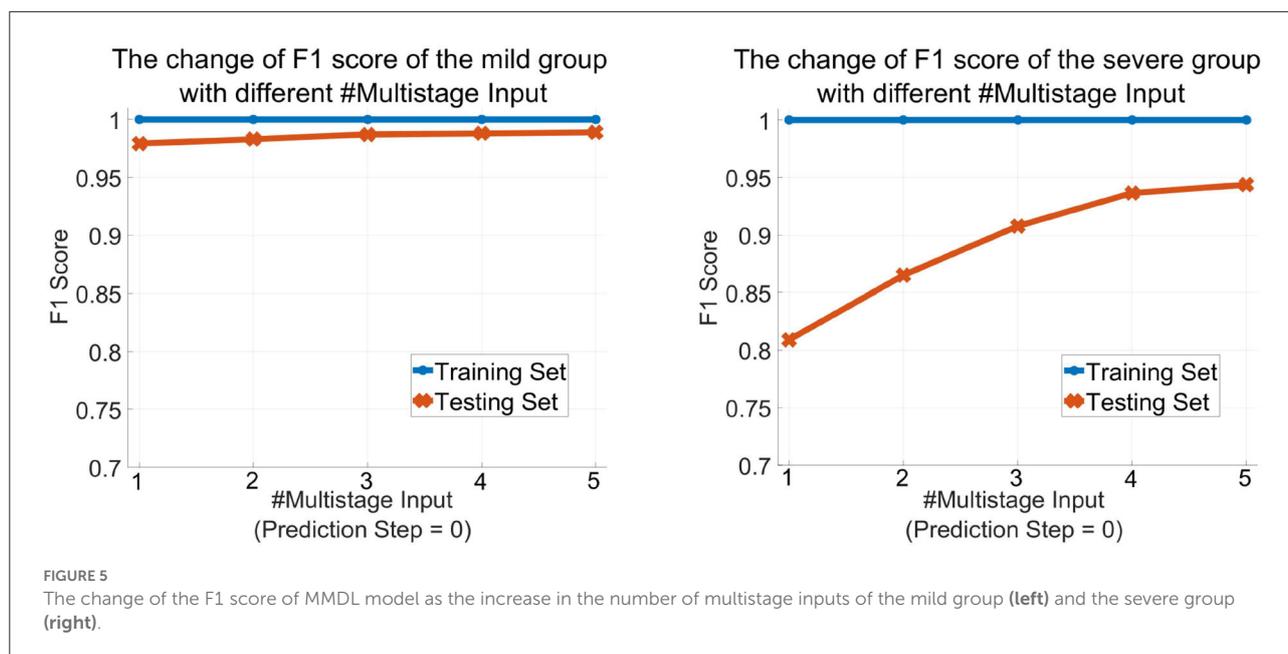


TABLE 3 Performance comparison of MMDL model with different prediction step.

	#Multistage Input data	Prediction Step	Accuracy	Error Rate	Precision	Recall	F1 Score
Mild group	1	0	96.26%	3.73%	97.52%	98.33%	0.9792
	1	1	95.63%	4.30%	96.86%	98.52%	0.9769
	1	2	94.43%	5.50%	96.25%	97.85%	0.9704
	1	3	93.61%	6.30%	96.00%	97.22%	0.9661
	1	4	93.44%	6.55%	95.74%	97.12%	0.9653
Severe group	1	0	96.26%	3.73%	84.12%	77.94%	0.8091
	1	1	95.63%	4.30%	70.00%	51.85%	0.5957
	1	2	94.43%	5.50%	59.09%	44.82%	0.5098
	1	3	93.61%	6.30%	50.00%	40.62%	0.4482
	1	4	93.44%	6.55%	48.00%	36.36%	0.4137

distinguishing between the one developing and not developing severe symptoms when $0.2 \leq FPR \leq 0.7$.

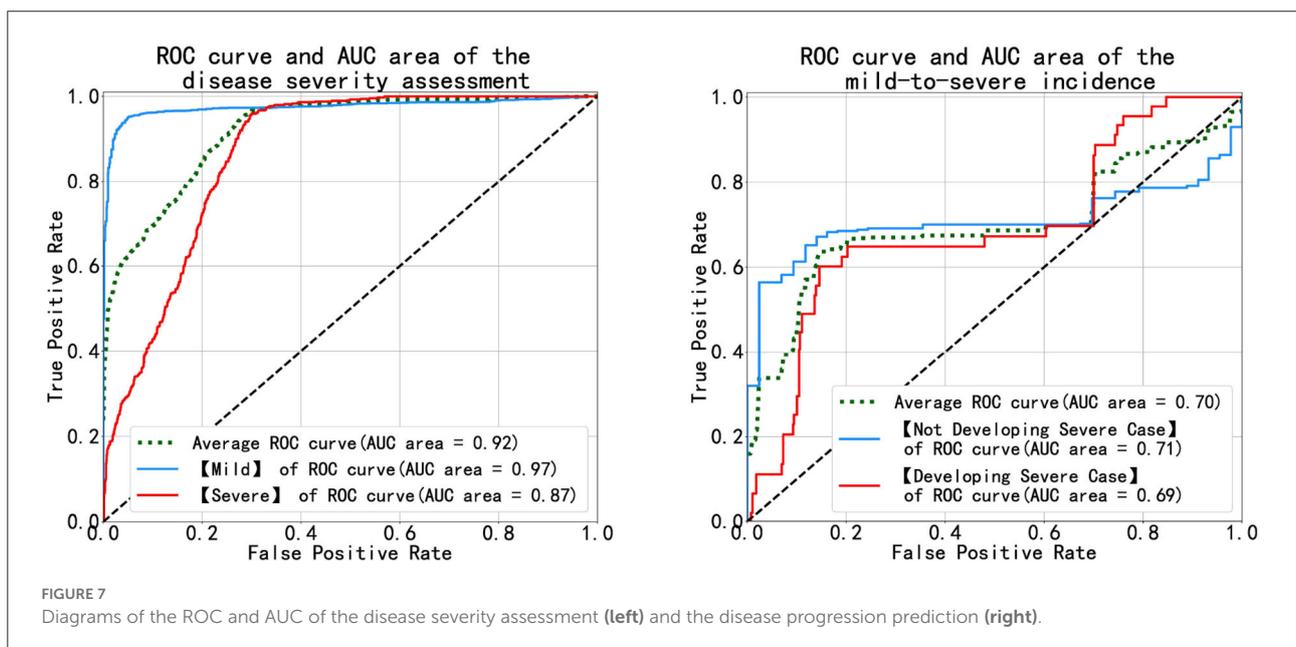
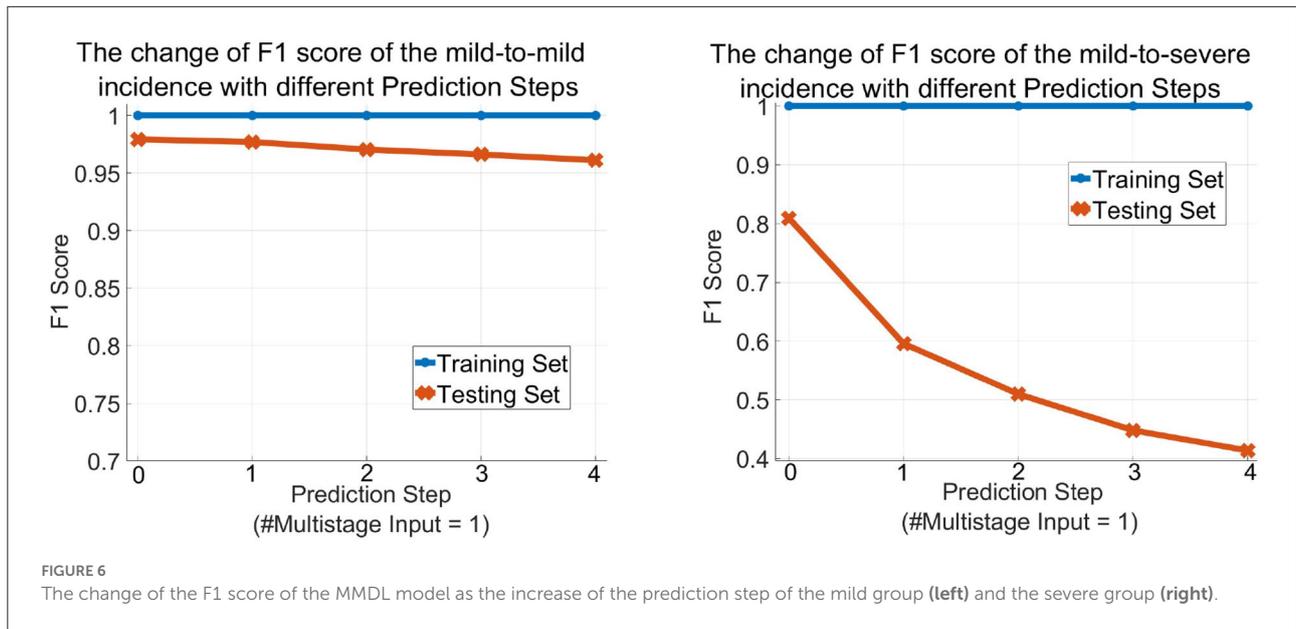
6.3.4. The impact of multimodal deep learning on the performance of the MMDL model

To show the advantages of multimodal learning for feature extraction and fusion across different modalities of clinical data, we compare the performance of the complete MMDL model with multimodal inputs and the reduced models with separate single-modal inputs only in the testing phase.

Figure 8 (upper) and (middle) compare MMDL using the latest round of exam and lab test results as the model input (*Multistage Input* = 1 and *PredictionStep* = 0) and leveraging the last five consecutive rounds of test data (*Multistage Input* = 5 and *PredictionStep* = 0). As we can see, the overall

performance of the MMDL model for assessing the mild group exceeds that of the severe group by a large margin regardless of using single-stage or multistage inputs. Noted that the performance gain is limited for the mild group, particularly for *#Multistage Input* = 1, but significant for the severe group, which grows at least 15% for both *#Multistage Input* = 1 and *#Multistage Input* = 5.

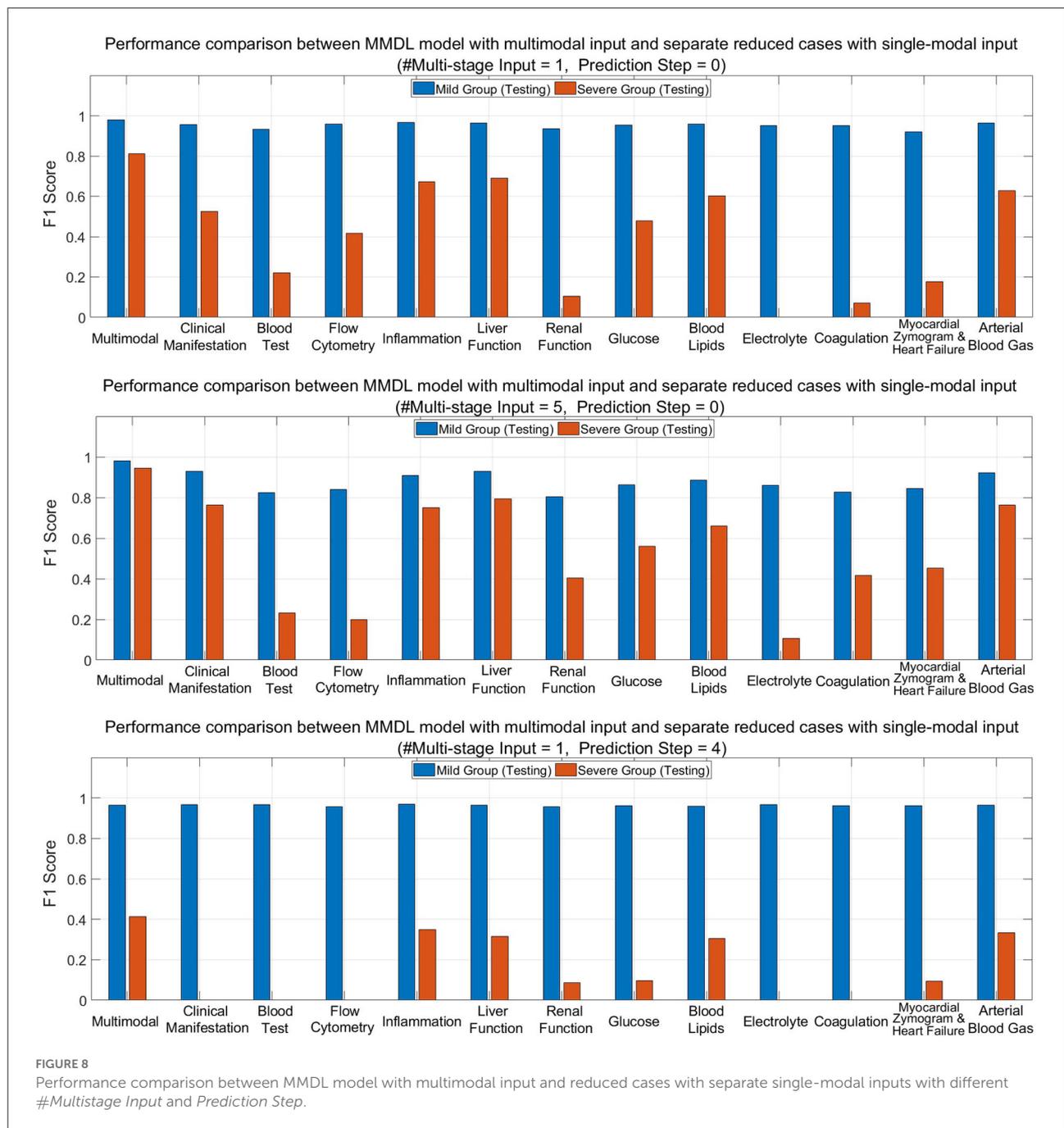
Figure 8 (lower) depicts the bar chart of the F1 score for forecasting a patient's condition four stages away from now (*Prediction Step* = 4). From the diagram, we have the following observations: the achieved F1 score is good for mild-to-mild but, to some extent, terrible for mild-to-severe incidence prediction. It reveals the false-negative rate of the mild-to-severe incidence prediction is high, that is, samples are more prone to be classified as not developing severe symptoms. Nevertheless, MMDL predicting with multimodal inputs outperforms reduced



models using any single-modal clinical data, which validates the superiority of multimodal learning. Furthermore, it is worth pointing out that among all modalities, inflammation, liver function, blood lipids, and arterial blood gas reach much higher F1 than any other modality. Hence, further explorations need to be conducted to discern the effective biomarkers within these modalities, which can be treated as signs to discriminate between mild cases developing and not developing severe symptoms.

7. Discussion

We notice that in the disease severity assessment task, MMDL's classification performance in the severe group is not as good as the mild group irrespective of using single-stage or multistage inputs. Similar observations are made in the disease progression prediction task as well that the prediction results of the mild-to-severe incidence fall far behind the mild-to-mild



incidence. From our perspective, the reasons for this are twofold: (1) The patient samples contained in the dataset are quite limited for clinical data analysis and model development. We are only authorized to use these 200+ samples legally that pass the review of the ethics committees (RECs). However, ethics and compliance are extremely important in clinical research, and samples that fail to pass RECs are strictly forbidden to use; (2) What makes the situation even worse is the imbalanced distribution of patient samples (only 30 severe cases). As a result,

it is insufficient to learn the characteristics of the patients of the severe group and the transition from mild to severe symptoms.

Besides the small sample learning, another challenge is that patient samples contained in the dataset were collected during the first wave of the pandemic, and the pre-trained model may no longer take effect as the virus has evolved to the Omicron variant in 2022. To address the challenge, we have deployed the prototype of MMDL in Chongqing Public Health Center, China, to validate the effectiveness of MMDL when facing new variants

of COVID-19. Alongside model testing, we also collect new patient samples and attempt to train MMDL using new samples incrementally. Furthermore, to test the MMDLs availability in other chronic diseases, we are extending it to epilepsy prediction characterized by many follow-ups.

Another observation is that MMDL using multiple sequential stages' exam and lab test data outperforms the current stage's data in disease severity assessment. In particular, the latest three rounds' inputs dominate the assessment results, and history long ago has little influence on the model's output. Moreover, in predicting the disease progression, we can observe prediction results deteriorate as *Prediction Step* increases. It is because, according to our point of view, biomarkers show no significant abnormality to discriminate whether patients will turn for the worse in the distant future.

Also, experimental results validate multimodal feature extraction and fusion can provide complementary information to single-modal feature learning. Another interesting finding reveals that either in assessment or prediction, merely leveraging the modality of inflammation, liver function, or blood lipids data, etc., overwhelms any other single-modal input. It suggests that some test items in the inflammation modality and the liver function modality, such as C-reactive protein (CRP), hypersensitive C-reactive protein (hsCRP), γ -glutamyltransferase (GGT), and Albumin (ALB), are potential biomarkers in distinguishing COVID-19 infections.

8. Conclusion

In this paper, we have conceived and implemented a multistage, multimodal deep learning (MMDL) model to assess the disease severity and forecast the disease progression of patients with COVID-19. In summary, the novelty of MMDL embodies sequential stage-wise learning with multimodal inputs. MMDL shows the advantage of studying whole courses of the disease compared to single-stage learning. Also, mining the multimodal clinical data can provide significant performance gains over using single-modal data only. Some potential biomarkers have been identified in the control experiment, such as C-reactive protein (CRP) and hypersensitive C-reactive protein (hsCRP) of the inflammation modality, and γ -glutamyltransferase (GGT) and Albumin (ALB) of the liver function modality. A strong correlation is seen between these potential biomarkers and the assessment/prediction results. In addition, we have deployed the prototype of the MMDL model in Chongqing Public Health Center, China, to test MMDLs robustness to the new variants of COVID-19 and collect more clinical data for further incremental training.

Data availability statement

The datasets presented in this article are not readily available because according to China's COVID-19 regulation and policy, any data related to patients with COVID-19 is not allowed to be disclosed publicly. Requests to access the datasets should be directed to cliff.zhao.li@gmail.com.

Author contributions

ZL conceived of the presented idea, designed the model, performed the implementation wrote the manuscript, and also encouraged. RX and YS to conduct the data preprocessing work, supervised the findings of this work and were mainly responsible for data cleaning, preprocessing, experiment setup, doing extensive comparison experiments, and analyzed the obtained results. JC contributed to the idea of overall model design and proofreading of the final version of the manuscript. BW provided insights on analyzing the clinical data and finding potential hall markers of infection of COVID-19 from the perspective of a physician. YZ and SL helped collect the multistage multimodal clinical data of examination and lab test results, facilitated the deployment of the prototype of the prediction model, and handled other routine work. All authors discussed the results and contributed to the final manuscript.

Funding

This research is supported by the Ph.D., Scientific Research Sharing Foundation of the Chongqing University of Posts and Telecommunications (Project No. E012A2022026) and the Chongqing Key Special Program for Technology Innovation and Application Development (Project No. cstc2020jscx-fyzzX0023).

Conflict of interest

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

Publisher's note

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

References

- Zhu N, Zhang D, Wang W, Li X, Yang B, Song J, et al. A novel coronavirus from patients with pneumonia in China 2019. *N Engl J Med.* (2020) 382:727–33. doi: 10.1056/NEJMoa2001017
- Zhou P, Yang X, Wang X, Hu B, Zhang L, Zhang W, et al. A pneumonia outbreak associated with a new coronavirus of probable bat origin. *Nature.* (2020) 579:270–3. doi: 10.1038/s41586-020-2012-7
- Lu R, Zhao X, Li J, Niu P, Yang B, Wu H, et al. Genomic characterisation and epidemiology of 2019 novel coronavirus: implications for virus origins and receptor binding. *Lancet.* (2020) 395:565–74. doi: 10.1016/S0140-6736(20)30251-8
- Guan W, Ni Z, Hu Y, Liang W, Ou C, He J, et al. Clinical characteristics of coronavirus disease 2019 in China. *N Engl J Med.* (2020) 382:1708–20. doi: 10.1056/NEJMoa2002032
- World Health Organization (WHO). *WHO Coronavirus Disease (COVID-19) Dashboard.* (2022). Available online at: <https://covid19.who.int/> (accessed on July 1, 2022).
- World Health Organization (WHO). *Statement on the Second Meeting of the International Health Regulations 2005 Emergency Committee Regarding the Outbreak of Novel Coronavirus (2019-nCoV).* (2021). Available online at: [https://www.who.int/news-room/detail/30-01-2020-statement-on-the-second-meeting-of-the-international-health-regulations-2005-emergency-committee-regarding-the-outbreak-of-novel-coronavirus-\(2019-ncov\)](https://www.who.int/news-room/detail/30-01-2020-statement-on-the-second-meeting-of-the-international-health-regulations-2005-emergency-committee-regarding-the-outbreak-of-novel-coronavirus-(2019-ncov)).
- Chen N, Zhou M, Dong X, Qu J, Gong F, Han Y, et al. Epidemiological and clinical characteristics of 99 cases of 2019 novel coronavirus pneumonia in Wuhan, China: a descriptive study. *Lancet.* (2020) 395:507–13. doi: 10.1016/S0140-6736(20)30211-7
- Huang C, Wang Y, Li X, Ren L, Zhao J, Hu Y, et al. Clinical features of patients infected with 2019 novel coronavirus in Wuhan, China. *Lancet.* (2020) 395:497–506. doi: 10.1016/S0140-6736(20)30183-5
- Wang D, Hu B, Hu C, Zhu F, Liu X, Zhang J, et al. Clinical characteristics of 138 hospitalized patients with 2019 novel coronavirus-infected pneumonia in Wuhan, China. *JAMA.* (2020) 323:1061–9. doi: 10.1001/jama.2020.1585
- Wu Z, McGoogan JM. Characteristics of and important lessons from the coronavirus disease 2019 (COVID-19) outbreak in China. *JAMA.* (2020) 323:1239–42. doi: 10.1001/jama.2020.2648
- Raoof S, Nava S, Carpati C, Hill NS. High-flow, noninvasive ventilation and awake (nonintubation) proning in patients with COVID-2019 with respiratory failure. *Chest.* (2020) 158:1992–2002. doi: 10.1016/j.chest.2020.07.013
- Simonnet A, Chetboun M, Poissy J, Raverdy V, Noulette J, Duhamel A, et al. High prevalence of obesity in severe acute respiratory syndrome coronavirus-2 (SARS-Co-2) requiring invasive mechanical ventilation. *Obesity.* (2020) 28:1195–9. doi: 10.1002/oby.22831
- Granter SR, Beck AH, Papke DJ. AlphaGo, deep learning, and the future of the human microscopist. *Arch Pathol Lab Med.* (2017) 141:619–21. doi: 10.5858/arpa.2016-0471-ED
- Whitelaw S, Mamas E, Mamas A, Topol, Van Spall HGC. Applications of digital technology in COVID-19 pandemic planning and response. *Lancet Digit Health.* (2020) 2:E435–40. doi: 10.1016/S2589-7500(20)30142-4
- Lalmuanawma S, Hussain J, Chhakchhuak L. Applications of machine learning and artificial intelligence for Covid-19 (SARS-CoV-2) pandemic: a review. *Chaos Solitons Fractals.* (2020) 139:59. doi: 10.1016/j.chaos.2020.110059
- Budd J, Miller BS, Manning EM, Lampos V, Zhuang M, Edelstein M, et al. Digital technologies in the public-health response to COVID-19. *Nat Med.* (2020) 26:1183–92. doi: 10.1038/s41591-020-1011-4
- Rahman A, Sultan K, Naseer I, Majeed R, Musleh D, Gollapalli M, et al. Supervised machine learning-based prediction of COVID-19. *Comput. Mater. Continua.* (2021) 69:21–34. doi: 10.32604/cmc.2021.013453
- Zoabi Y, Deri-Rozov S, Shomron N. Machine learning based prediction of COVID-19 diagnosis based on symptoms. *NPJ Digit Med.* (2021) 4:3. doi: 10.1038/s41746-020-00372-6
- Wynants L, Van Calster B, Collins GS, Riley RD, Heinze G, Schuit E, et al. Prediction models for diagnosis and prognosis of COVID-19: systematic review and critical appraisal. *BMJ.* (2020) 26:1183–92. doi: 10.1136/bmj.m1328
- Narin A, Kaya C, Pamuk Z. Automatic detection of coronavirus disease (COVID-19) using X-ray images and deep convolutional neural networks. *Pattern Anal Appl.* (2021) 24:1207–20. doi: 10.1007/s10044-021-00984-y
- Wang S, Zha Y, Li W, Wu Q, Li X, Niu M, et al. A fully automatic deep learning system for COVID-19 diagnostic and prognostic analysis. *Eur Respir J.* (2020) 56:2000775. doi: 10.1183/13993003.00775-2020
- Ozturk T, Talo M, Yildirim EA, Baloglu UB, Yildirim O, Rajendra Acharya U. Automated detection of COVID-19 cases using deep neural networks with X-ray images. *Comput Biol Med.* (2020) 121:103792. doi: 10.1016/j.combiomed.2020.103792
- Yadaw AS, Li Yc, Bose S, Iyengar R, Bunyanich S, Pandey G. Clinical features of COVID-19 mortality: development and validation of a clinical prediction model. *Lancet Digital Health.* (2020) 2:E516–25. doi: 10.1016/S2589-7500(20)30217-X
- Jehi L, Ji X, Milinovich A, Erzurum S, Rubin BP, Gordon S, et al. Individualizing risk prediction for positive coronavirus disease 2019 testing: results from 11,672 patients. *Chest.* (2020) 158:1364–75. doi: 10.1016/j.chest.2020.05.580
- Wu J, Zhang P, Zhang L, Meng W, Li J, Tong C, et al. Rapid and accurate identification of COVID-19 infection through machine learning based on clinical available blood test results. *medRxiv.* (2020) doi: 10.1101/2020.04.02.20051136
- Elman JL. Finding structure in time. *Cogn Sci.* (1990) 14:179–211. doi: 10.1207/s15516709cog1402_1
- Ian G, Yoshua B, Aaron C. 6.2.2.3 Softmax Units for Multinoulli Output Distributions. In: *Deep Learning.* MIT Press (2016). p. 180–4.
- Hochreiter S, Schmidhuber J. Long short-term memory. *Neural Comput.* (1997) 9:1735–80. doi: 10.1162/neco.1997.9.8.1735
- Cho K, van Merriënboer B, Gulcehre C, Bahdanau D, Bougares F, Schwenk H, et al. Learning phrase representations using RNN encoder-decoder for statistical machine translation. In: *Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing.* EMNLP'14 (2014). p. 1724–34.
- Sutskever I, Vinyals O, Le QV. Sequence to sequence learning with neural networks. In: *Proceedings of the 27th International Conference on Neural Information Processing Systems.* Vol. 2 of NIPS'14. Montreal, QC: MIT Press (2014). p. 3104–12.
- Nair V, Hinton GE. Rectified linear units improve restricted boltzmann machines. In: *Proceedings of the 27th International Conference on International Conference on Machine Learning.* ICML'10. Haifa: Omnipress (2010). p. 807–14.