

Deep learning for predicting COVID-19 malignant progression

Fang et al.

Supplementary Information

Supplementary methods

Attention weights updating. In each mini-batch of the training process, the exponential moving average is used to update the attention weights in the training set. In the testing process, the average weights obtained in the training process are used to update the clinical data. The attention weights are updated as:

$$A_i = \begin{cases} A_c, & i = 1 \\ \alpha \cdot A_c + (1 - \alpha) \cdot A_{i-1}, & i > 1 \end{cases} \quad (1)$$

where A_i is the average attention weights; A_c is the current attention weights; the coefficient α represents the degree of weighting decrease, a constant smoothing factor between 0 and 1; i is the number of iterations in the training process. In our experiment α is 0.1.

Statistic metric. The following 4 metrics are used to evaluate the performance.

1. AUC:

$$AUC = \sum_{i=1}^n TPR_i \times (FPR_i - FPR_{i-1}) \quad (2)$$

where $T = [th_1, th_2, \dots, th_n]$, $0 \leq th_i \leq 1$, $th_i < th_{i+1}$, $TPR_i = \frac{TP}{TP+FN}$ and $FPR_i = 1 -$

$\frac{TN}{FP+TN}$ both with a threshold = $T[i]$.

2. accuracy:

$$accuracy = \frac{TP+TN}{TP+TN+FP+FN}, \text{ with a threshold} = 0.5 \quad (3)$$

3. sensitivity:

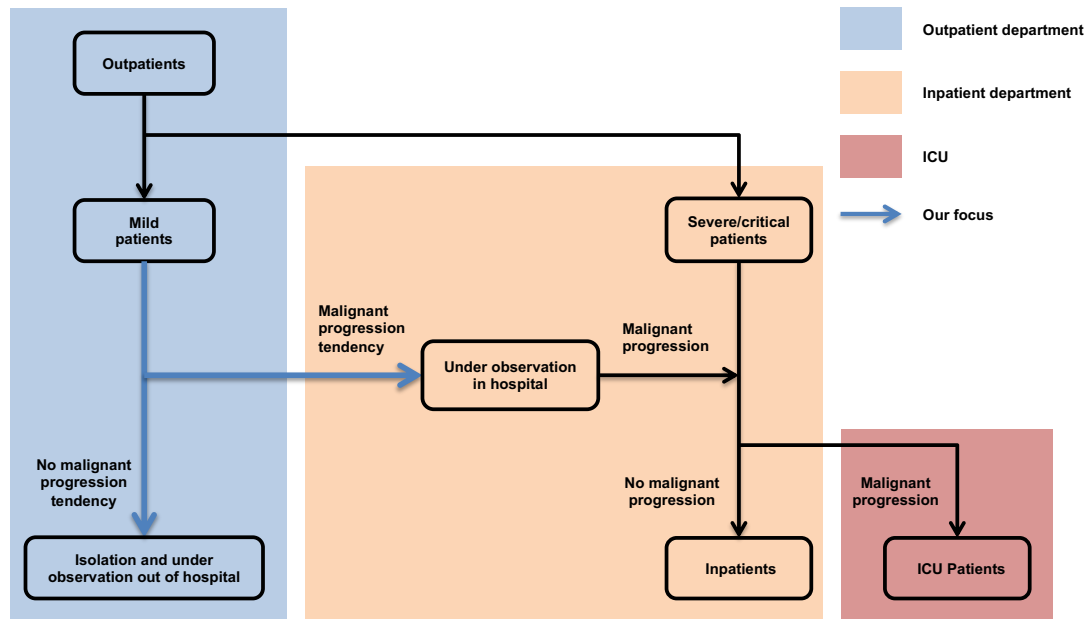
$$sensitivity = \frac{TP}{TP+FN}, \text{ with a threshold} = 0.5 \quad (4)$$

4. specificity:

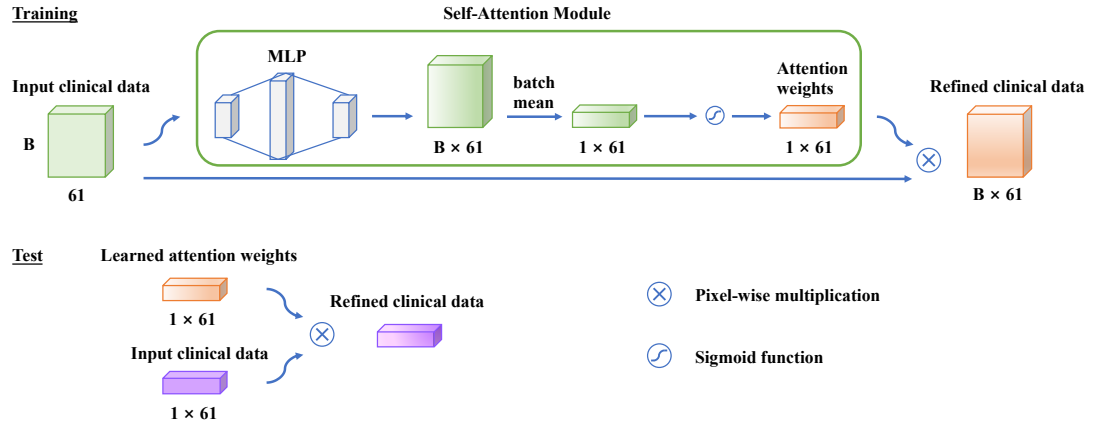
$$specificity = \frac{TN}{FP+TN}, \text{ with a threshold} = 0.5 \quad (5)$$

where TP is true positive, TN is true negative, FP is false positive and FN is false negative.

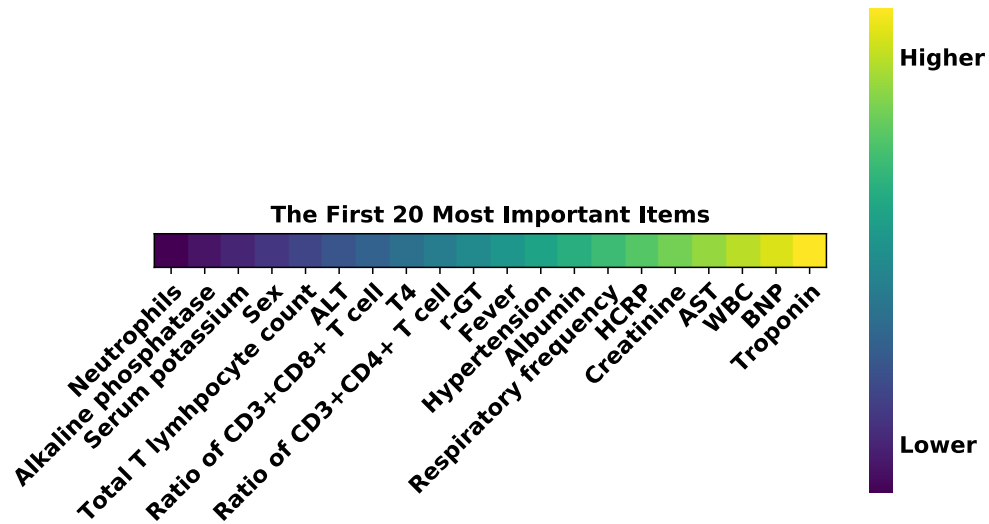
Supplementary Table 1. The detailed parameters of the 3D ResNet.			
Layer name	Operation	Input size	Output size
Conv1	$3 \times 3 \times 3$, 32, stride 2	$64 \times 64 \times 64 \times 1$	$32 \times 32 \times 32 \times 32$
Block1	$3 \times 3 \times 3$, 32, stride 2 $3 \times 3 \times 3$, 32, stride 1	$32 \times 32 \times 32 \times 32$	$16 \times 16 \times 16 \times 32$
Block2	$3 \times 3 \times 3$, 64, stride 2 $3 \times 3 \times 3$, 64, stride 1	$16 \times 16 \times 16 \times 32$	$8 \times 8 \times 8 \times 64$
Block3	$3 \times 3 \times 3$, 128, stride 2 $3 \times 3 \times 3$, 128, stride 1	$8 \times 8 \times 8 \times 64$	$4 \times 4 \times 4 \times 128$
	global average pooling	$4 \times 4 \times 4 \times 128$	$1 \times 1 \times 1 \times 128$



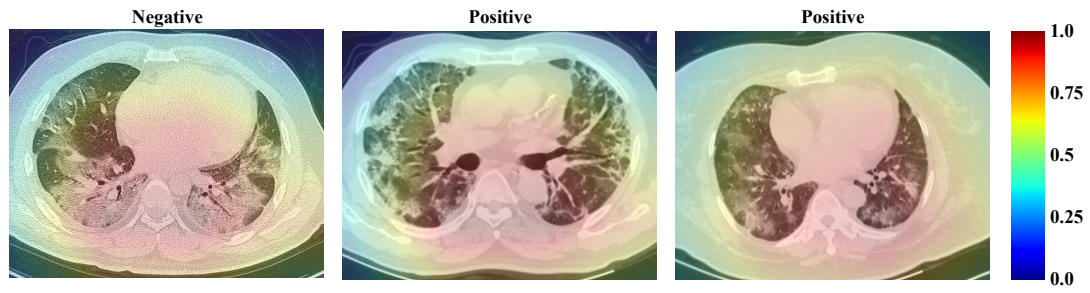
Supplementary Figure 1. Patient stratification from outpatients to ICU. Reasonable hierarchical management of COVID-19 patients is beneficial to optimizing the allocation of medical resources and improving the efficiency of diagnosis and treatment.



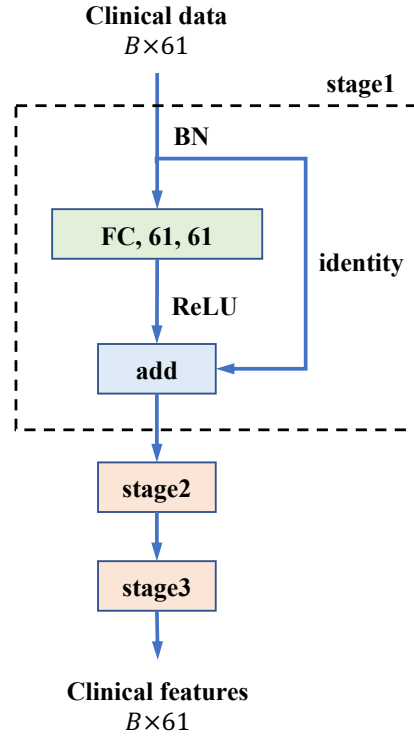
Supplementary Figure 2. Self-attention module for prognostic factors. MLP: multilayer perceptron, $B \times 61$ represents the batch size and the length of the vector.



Supplementary Figure 3. The top prognostic factors of clinical data. ALT: Alanine aminotransferase, r-GT: γ -Glutamyl transpeptidase, HCRP: Hypersensitive C-reactive protein, AST: Aspartate aminotransferase, WBC: white blood cell, BNP: Brain natriuretic peptide.



Supplementary Figure 4. Visualization of learned activation maps for COVID-19 patients with mild symptoms. Red regions correspond to high score for class, and our system localizes class-discriminative regions. Figure best viewed in color.



Supplementary Figure 5. Clinical data encoder. This encoder has three stages, each of which consists of a fully connected layer and an identity connection. BN: batch normalization, add: pixel-wise addition, identity: identity connection. FC, 61, 61 represents a fully connected layer, the size of input features, and the size of output features. $B \times 61$ represents the batch size and the length of the vector.