

1 Multiple-Input Deep Convolutional Neural Network 2 Model for COVID-19 Forecasting in China

3 Chiou-Jye Huang¹, Yung-Hsiang Chen², Yuxuan Ma¹ and Ping-Huan Kuo^{3,*}

4 ¹ School of Electrical Engineering and Automation, Jiangxi University of Science and Technology, Ganzhou,
5 Jiangxi, 341000, China; cjhuang@jxust.edu.cn; 1103552453@qq.com

6 ² Department of Mechanical Engineering, National Pingtung University of Science and Technology, Pingtung
7 91201, Taiwan; yhchen@mail.npust.edu.tw

8 ³ Computer and Intelligent Robot Program for Bachelor Degree, National Pingtung University, Pingtung
9 90004, Taiwan; phkuo@mail.nptu.edu.tw

10 * Correspondence: phkuo@mail.nptu.edu.tw; Tel.: +886-87663800 ext. 32620

11

12 **Abstract:** COVID-19 is spreading all across the globe. Up until March 23, 2020, the confirmed cases in 173
13 countries and regions of the globe had surpassed 346,000, and more than 14,700 deaths had resulted. The
14 confirmed cases outside of China had also reached over 81,000, with over 3,200 deaths. In this study, a
15 Convolutional Neural Network (CNN) was proposed to analyze and predict the number of confirmed cases.
16 Several cities with the most confirmed cases in China were the focus of this study, and a COVID-19
17 forecasting model, based on the CNN deep neural network method, was proposed. To compare the overall
18 efficacies of different algorithms, the indicators of mean absolute error and root mean square error were
19 applied in the experiment of this study. The experiment results indicated that compared with other deep
20 learning methods, the CNN model proposed in this study has the greatest prediction efficacy. The feasibility
21 and practicality of the model in predicting the cumulative number of COVID-19 confirmed cases were also
22 verified in this study.

23 **Keywords:** Total confirmed forecasting; convolutional neural network; COVID-19

24

25 1. Introduction

26 In early December 2019, the first case of COVID-19 infection was discovered in Wuhan City in China's
27 Hubei Province [1]. In the following weeks, this disease broke out within China and continued to spread
28 extensively in other countries, causing worldwide panic. The pneumonia caused by this novel coronavirus was
29 officially named the Coronavirus disease 2019 (COVID-19) by the World Health Organization (WHO). The
30 2019 coronavirus epidemic was triggered by the Severe Acute Respiratory Syndrome Coronavirus 2 (SARS-
31 CoV-2), and it has, at the time of writing, spread in over 173 countries, causing more than 15,000 deaths. The
32 epidemic is still spreading, and the WHO has classified the global risk level of the disease as a "pandemic."
33 Currently, the earliest known case symptoms appeared on December 1, 2019, and the first case sought treatment
34 on December 8. On December 26, 2019, Wuhan City Respiratory and Intensive care doctor Zhang Jixian first
35 discovered this pneumonia with unknown cause and suspected that it is an infectious disease. Subsequently, the
36 disease broke out in Wuhan City. On January 20, 2020, Chinese academic Zhong Nanshan publicly announced
37 that the novel coronavirus pneumonia "definitely transmits between people." On January 23, 2020, the Wuhan
38 City Government announced the adoption of lockdown and quarantine measures in the infected areas, which
39 was the first case of lockdown in a major city (with a population of 11 million) in recent public health history.
40 Since January 13, the disease continued to spread to Thailand, Japan, and South Korea. On January 30, three
41 countries outside China were verified as having interpersonal propagation, and the pandemic was thus
42 designated as an international public health emergency event by the WHO.

43 According to reports on the epidemic, "the propagation speed is faster, and the virus propagation power
44 has increased." Infected people can infect others with the virus while exhibiting no symptoms; and the latency
45 from infection to the presentation of symptoms is as long as 14 days. These characteristics also increase the
46 difficulty of controlling the epidemic. As the epidemic continues, there is also a problem of the global
47 undersupply of surgical masks. Currently, no vaccine and remedy for the novel coronavirus have been
48 discovered. WHO assistant director-general Bruce Aylward stated that Remdesivir is currently the only drug

49 that is “considered to probably have real efficacy.” The China–Japan Friendship Hospital in China and the
50 American National Institute of Allergy and Infectious Diseases are starting to conduct clinical trials for the drug.
51 At present, no sufficient knowledge regarding the disease is known, and key factors, such as virus source, virus
52 birthplace, morbid mechanism, virus pathogenicity, and propagation power are still uncertain. The WHO stated
53 on March 3 that with the novel coronavirus epidemic, the world is in an “unknown state.”

54 Up until March 23, 2020, 345,297 people have been infected with the virus; among these people, 14,765
55 have died. The several European countries with the most severe epidemics have experienced the maximum
56 increase in confirmed cases in a single day on March 5: cases for Italy increased from 3,089 to 3,858, those for
57 Germany increased from 262 to 482, those for France increased from 285 to 423, and those for the Netherlands
58 increased more than twofold from 38 to 82. In the WHO report on February 26, the numbers of new cases in
59 China and Japan have decreased, but the numbers of new cases in Italy, Iran, and South Korea are still rising.
60 At the time of writing, the numbers of confirmed cases in Germany, France, and Spain have surpassed 1,000.

61 2. Related Works

62 Fig. 1 [2] presents the global map of COVID-19 occurrence according to data from the United States’
63 Centers for Disease Control and Prevention. The countries marked with a color are countries with confirmed
64 cases. At present, 173 of the 195 countries in the world have confirmed cases, underscoring the epidemic’s
65 astounding speed of propagation from its origin in China’s Hubei province. Most scholars have used
66 mathematical models to predict the spread of COVID-19. Roosa et al. [3] used the generalized logistic model,
67 the Richards model, and the sub-epidemic model to establish prediction models for the cumulative numbers of
68 confirmed cases in Guangdong Province and Zhejiang Province. They predicted that, in the next 5 days and 10
69 days from their time of writing, 65–81 and 44–354 new cases will appear in Guangdong and Zhejiang provinces,
70 respectively. In addition, Liu et al. [4] used mathematical models to simulate the propagation status of COVID-
71 19 in China and noted the importance of the government’s policy of limiting public activities and the movement
72 of infected people with no symptoms. Boldog et al. [5] adopted a Time-Dependent Compartmental Model of
73 the Transmission Dynamics to estimate the cumulative number of confirmed cases outside of Hubei Province.
74 Moreover, the Galton–Watson branch was used to analogize virus propagation, thus determining the
75 connectivity and transmissibility between China and the destination country. The disadvantage of this method
76 is that the uncertainty regarding the efficacy of control measures against the disease results in difficulties in
77 predicting the development trajectories of the disease. By comparison, Al-Qaness et al. [6] adopted the adaptive
78 neuro-fuzzy inference system (ANFIS), an improvement from flower pollination algorithm and salp swarm
79 algorithm, to predict the number of confirmed cases in the next 10 days. The comparison with the ANFIS
80 optimized with GA, PSO, ABC, and FPA indicated that the accuracy of the FPASSA-ANFIS model reached
81 0.97, and the predicted average increase in numbers of new cases in the next 10 days was 10% more than the
82 number of present cases. Jung et al. [7] used the rate of increase together with delayed distribution estimation
83 and statistical induction to establish a mathematical model, based on COVID-2019 case data reported before
84 January 24. The predicted cumulative number of confirmed cases for up until January 24 was 6,924, and the
85 predicted death ratio was 5.3%; these figures were within the 95% confidence interval. Compared with previous
86 methods for establishing mathematical models, Fan et al. [8] adopted a statistical method for predicting the
87 floating population in Wuhan City. The experiment indicated that the floating population in Wuhan region is
88 highly correlated with the number of daily confirmed cases. The residence time of the floating population for
89 local cases was longer than that for non-local cases, which results in a lower predicted number of confirmed
90 cases in the areas around Hubei Province. The prediction results indicate that approximately 80% of the
91 epidemic will be centralized in the top 30 districts.

92 However, applications of deep learning methods have mainly used the genome of the 2019-cCoV virus to
93 predict disease propagation. Yang et al. [9] used the Susceptible-Exposed-Infectious-Removed model to
94 integrate the epidemic curve, combining it with the long short-term memory model. The maximum value was
95 reached on February 28, and the curve gradually lowered in late April. Hu et al. [10] used an improved stacked
96 autoencoder and a cluster algorithm to group the instantaneous confirmed cases in every province; they noted
97 a high accuracy in AI-based methods for COVID-19 trajectory prediction. The epidemic was predicted to end
98 in mid-April. Guo et al. [11] used deep learning–based virus host prediction to compare the gene sequence of
99 2019-nCoVs with those of Severe Acute Respiratory Syndrome Coronavirus (SARS-CoV), bat SARS corona
100 virus, and Middle East respiratory syndrome–related coronavirus (MERS-CoV). The bat SARS coronavirus
101 was discovered to have a more similar mode of infection to COVID-19. Metsky et al. [12] used a nucleic acid

102 test based on clustered regularly interspaced short palindromic repeats (CRISPR) genome editing and selected
103 67 viruses and subspecies as well as subspecies that are highly similar to 2019-nCoV. Thus, 67 test methods
104 were designed in the ADAPT system to increase the sensitivity and speed of virus detection.

105 Riou et al. [13] simulated the early outbreak trajectories of 2019-nCoV and noted that the basic
106 reproduction number R of 2019-nCoV was approximately 2.2 (90% high density interval 1.4–3.8) and that
107 2019-nCoV has the potential for continual infection between people. Liu et al. [14] compared 2019-nCoV with
108 SARS and estimated the doubling time of 2019-nCoV and SARS as well as the basic reproduction number r_0
109 and time-varying instantaneous reproduction number r_t through data analysis. They discovered that the
110 infectivity of 2019-nCoV may be stronger than SARS, but that disease control efforts are effective. Ming et al.
111 [15] adopted an improved version of the Susceptible-Infectious-Recovered (SIR) model to predict the actual
112 number of infection cases of 2019-nCoV as well as the actual load on the intensive care units under different
113 diagnosis rates and public health intervention efficacies. They discovered that under a 50% diagnosis rate and
114 no public health intervention, the actual number of cases will be significantly higher than the reported number,
115 whereas under a 70% public health intervention, the load on the health system will decrease substantially. Zhao
116 et al. [16] used a probability prediction model to accurately predict the infection nodes from snapshots of
117 spreading. Fountain-Jones et al. [17] used machine learning to establish pathogen-risk models and compared
118 the prediction results of different machine learning methods and explained the results by using game theory.

119 Benvenuto et al. [18] adopted an autoregressive integrated moving average model to predict the epidemic
120 trend and morbidity of 2019-nCoV with a 95% confidence interval. They discovered that if the virus does not
121 mutate, the case number will reach a plateau. Li et al. [19] used a function to separately describe the daily
122 infection and death data of 2019-nCoV in Hubei and outside Hubei. The author thought that the inflection point
123 in the Hubei region was at February 6, 2020, and that the epidemic will end on March 10, 2020, with an
124 estimated infection rate of 39,000 people. However, the data did not include the data after February 12, where
125 such inclusion will increase the predicted figures by 1.4 times. Yang et al. [9] described the clinical
126 characteristics and imaging manifestation of the infected cases of COVID-19 that had been confirmed in the
127 hospitals in Wenzhou area. In addition, the patients with and without travel or residence history in Hubei were
128 compared. Tian et al. [20] analyzed the clinical and epidemiological characteristics of COVID-19 in the Beijing
129 area. The characteristics of severe confirmed cases and common confirmed cases were compared, and the
130 characteristics of COVID-19 and SARS were compared as well. Chen et al. [21] analyzed some cases of
131 pregnant women with COVID-19 and noted no evidence that COVID-19 causes intrauterine infection through
132 vertical transmission among women in late pregnancy. The major contribution of that study is its fast
133 establishment of a prediction model by using deep learning with a small dataset, making it an important
134 reference for other countries in their containment of the COVID-19 epidemic.

135 The structure of this study proceeds as follows. The first section introduces the entire study. The second
136 section is a literature review. The following third section details this study's proposed deep neural network
137 algorithm. The fourth section presents and discusses the experiment results in detail. Finally, the fifth section
138 concludes the entire study and underscores the contribution of this study and the predicted trend of the COVID-
139 19 epidemic.

140



141

142 **Fig. 1.** Global Map of Confirmed COVID-19 Cases as of 12:00 a.m. ET, March 21, 2020 [2].

143 **3. Artificial neural networks: A background**

144 An Artificial Neural Network (ANN) [22, 23] is a mathematical model that mimics how neurons operate
145 in organisms. It is a very powerful tool for establishing nonlinear models. An early ANN structure is the
146 Multilayer Perceptron (MLP), which uses the neural network of fully connected structures; it has decent efficacy
147 and numerous applications. However, if the data is too complex, simply using the MLP structure may cause the
148 model to be unable to effectively learn how to handle every situation. At present, ANN has already developed
149 into various new structures. The main structure of this study is the Convolutional Neural Network (CNN) [24,
150 25], and three other neural network structures are compared with it. These four neural network structures are
151 detailed in the following subsections.

152 *3.1. Convolutional neural network*

153 The 1D convolution operation [26] is illustrated in Fig. 2. The difference between CNN and MLP is that
154 CNN uses the concept of weight sharing. As illustrated in Fig. 2, the advantage of CNN is that its weight number
155 does not have to be as large as that for a fully connected structure. Consequently, the training is relatively easier,
156 and important characteristics are extracted more effectively. CNN is a feedforward neural network, which is an
157 improvement from MLP. CNN contains four levels in structure: an input layer, convolutional layer, pooling
158 layer, and fully connected layer.

159 The deep CNN structure proposed in this study is illustrated in Fig. 2, and the structure is composed of
160 CNN and the dropout layer. The input layer structure is constituted by six time sequences composed of factors
161 influencing the cumulative number of cases. The convolutional layer structure is divided into four layers, and
162 each convolution layer has 16, 32, 32, and 64 convolutional kernels. Going through every convolutional layer,
163 the convolutional kernels of the previous layer slide in the input data matrixes. The convolution process can be
164 expressed in Equation (1), in which X_{ij} is the i th and j th matrix in the row and column direction, respectively,
165 corresponding to the convolution kernels in the input matrix through smooth shifting. K is the convolution
166 kernel, and γ is the output matrix.

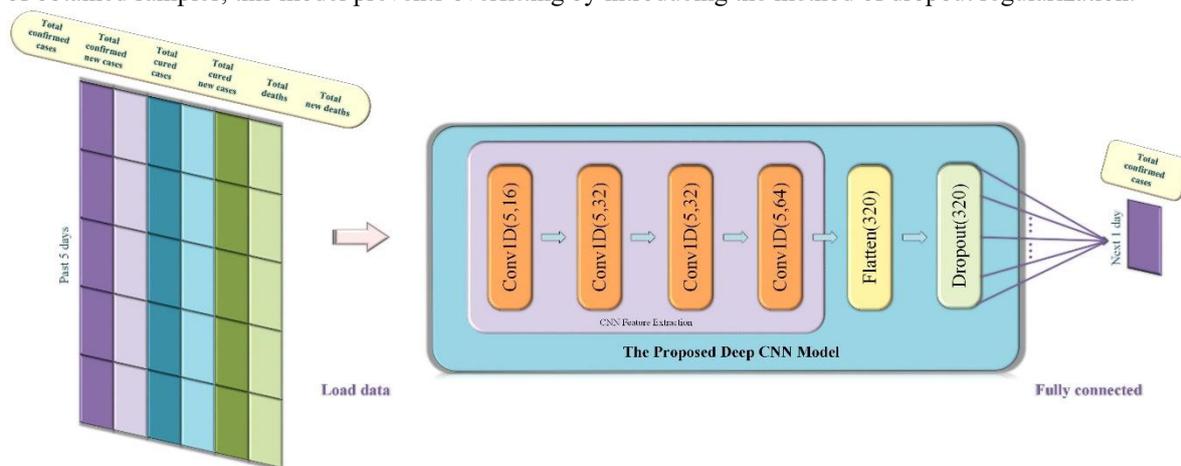
$$167 \gamma(i, j) = X_{ij} * K \quad (1)$$

168 The structure principle of the fully connected layer can be expressed by Equation (2):

$$169 \mathbf{o} = f(\mathbf{w}_f^T \mathbf{x} + \mathbf{b}) \quad (2)$$

170 In (2), \mathbf{o} is the vector composed of output values, \mathbf{x} is the vector composed of input values, \mathbf{w}_f^T is the
171 vector composed of weight values, \mathbf{b} is the vector composed of threshold values, and f is the activation
172 function. The output of the convolution layer can yield a 1D vector through Flatten expansion. Subsequently,
173 the vectors are connected to the fully connected layer (dense) to obtain a 1D output. At the time of writing, the

174 COVID-19 epidemic has continued for only 3 months, 1 week, and 3 days. Thus, considering the small number
 175 of obtained samples, this model prevents overfitting by introducing the method of dropout regularization.



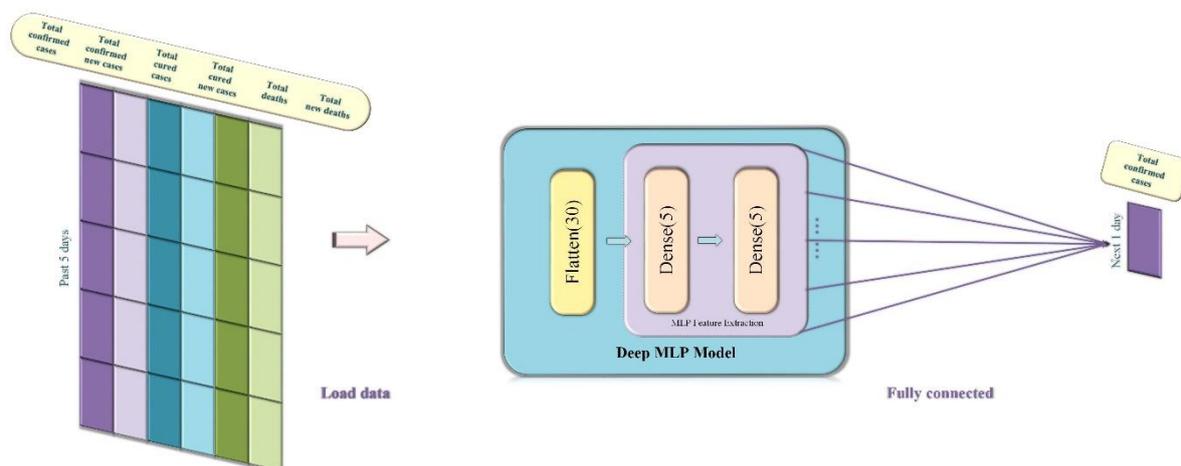
176
 177 **Fig. 2.** Structure of the proposed deep CNN model.

178 **3.2. Multilayer perceptron**

179 A MLP is an ANN model that includes the input layer, hidden layer, and output layer. The structure of the
 180 ANN is illustrated in Fig. 3. Different connection layers are all fully connected, and the principle is similar to
 181 that of MLP. Each neural unit is connected with a weight coefficient (w_i), and the combination of transmission
 182 result $w_i x_i$ and the deviation b_i is calculated using Equation (3). In addition, current neural unit output results
 183 are obtained through the activation function f .

$$184 \quad y_i = f(w_i x_i + b_i) \quad (3)$$

185 The structure of the proposed deep MLP is illustrated in Fig. 3. The structure is composed of two levels of
 186 MLP, and each MLP layer has 16 neural units. The input structure is constituted by six factors influencing the
 187 cumulative number of cases, with the time step of 5. Through a Flatten expansion, a 1×30 vector is derived,
 188 and the output is obtained after connecting the vector to the fully connected layer (dense).



189
 190 **Fig. 3.** Structure of the comparative deep neural network MLP adopted in this study.

191 **3.3. Long short-term memory neural network**

192 The long short-term memory model (LSTM) [27, 28] is an improvement from the recurrent neural network.
 193 The difference between the RNN and LSTM is that for LSTM, a cell state is added to store long-term states. In
 194 the neural unit model structure in Fig. 4 (a), the internal structure of LSTM can be divided into the input gate,
 195 forget gate, and output gate. The principle of the LSTM input gate is expressed in the following formulae.

196 Equation (4) is used to decide which piece of information is to be added by passing h_{t-1} and x_t through the
 197 sigmoid layer. Subsequently, Equation (5) is used to pass h_{t-1} and x_t through the tanh layer to obtain new
 198 information \tilde{C}_t . Equation (6) is used to combine the information of the current moment \tilde{C}_t and long-term
 199 memory C_{t-1} into a new memory state C_t .

$$200 \quad i_t = \sigma(W_i \cdot [h_{t-1}, x_t] + b_i) \quad (4)$$

$$201 \quad \tilde{C}_t = \tanh(W_c \cdot [h_{t-1}, x_t] + b_c) \quad (5)$$

$$202 \quad C_t = f_t C_{t-1} + i_t \tilde{C}_t \quad (6)$$

203 The forget gate of the LSTM uses a sigmoid layer and a dot product to allow information to pass through
 204 selectively. Equation (7) allows the LSTM to decide whether to forget the related information of the previous
 205 cell, at a certain probability, in which W_f is the weight matrix, and b_f is the offset term.

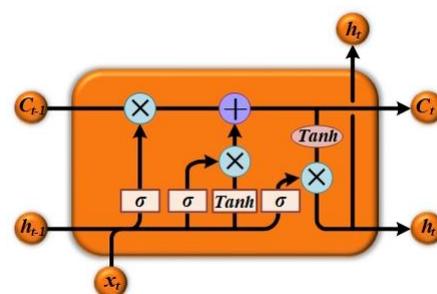
$$206 \quad f_t = \sigma(W_f \cdot [h_{t-1}, x_t] + b_f) \quad (7)$$

207 The output gate of LSTM decides which states are required to be maintained by the input h_{t-1} and x_t
 208 according to Equations (8) and (9). The final output results are obtained by passing the new information C_t
 209 through the tanh layer to multiply with state judgement vectors.

$$210 \quad O_t = \sigma(W_o \cdot [h_{t-1}, x_t] + b_o) \quad (8)$$

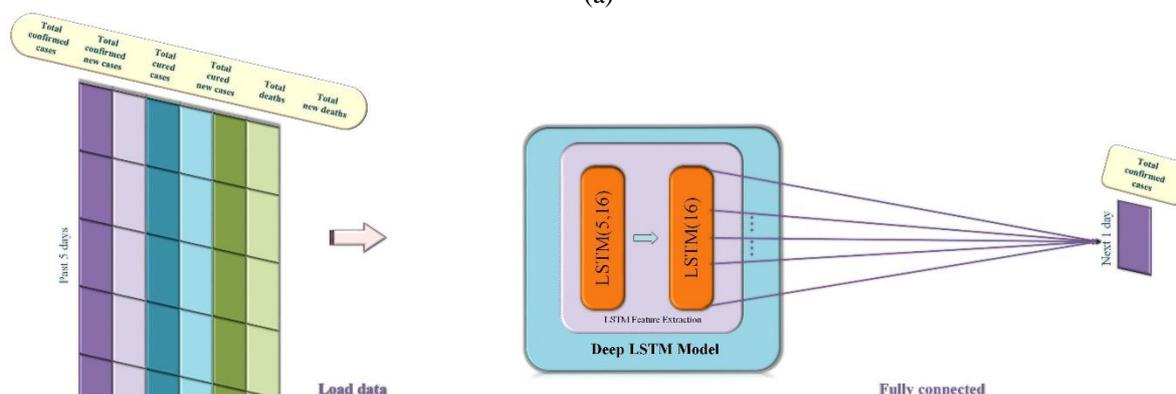
$$211 \quad h_t = O_t \tanh(C_t) \quad (9)$$

212 The deep LSTM structure proposed in this study is illustrated in Fig. 4 (b). The structure is composed of
 213 two LSTM layers, and each layer has 16 neural units. The input structure is constituted by six factors influencing
 214 the cumulative number of cases, with a time step of 5. A 1D output is obtained through the fully connected
 215 layer.



(a)

216
217



(b)

218
219

220 **Fig. 4.** Structure of the comparative deep neural network LSTM adopted in this study. (a) LSTM neural unit; (b) the
 221 deep neural network LSTM model with an input layer consisting six input signals.

222 3.4. Gate recurrent unit

223 The Gate Recurrent Unit (GRU) [29, 30] is a variation of the LSTM, as illustrated in Fig. 5 (a). GRU
 224 changes the input gate and forget gate of LSTM into the update gate. GRU retains the effect of LSTM and
 225 simplifies the internal structure of GRU model units. The principle of GRU's update gate is described in
 226 Equation (10). Specifically, the neural unit information of the previous moment h_{t-1} and the input data of the

227 current moment x_t are combined and input to the sigmoid layer to obtain the update information of the current
 228 moment z_t . W is the weight matrix.

229
$$z_t = \sigma(W_z \cdot [h_{t-1}, x_t]) \quad (10)$$

230 The principle of the GRU reset gate is described in Equation (11). The reset information r_t is reserved
 231 after h_{t-1} and x_t are separately combined with the weight matrix.

232
$$r_t = \sigma(W_r \cdot [h_{t-1}, x_t]) \quad (11)$$

233 The input information of the current moment x_t is maintained with the obtained r_t at the weight
 234 probability. The following Formula (12) is used to calculate the information of the current moment \tilde{h}_t .

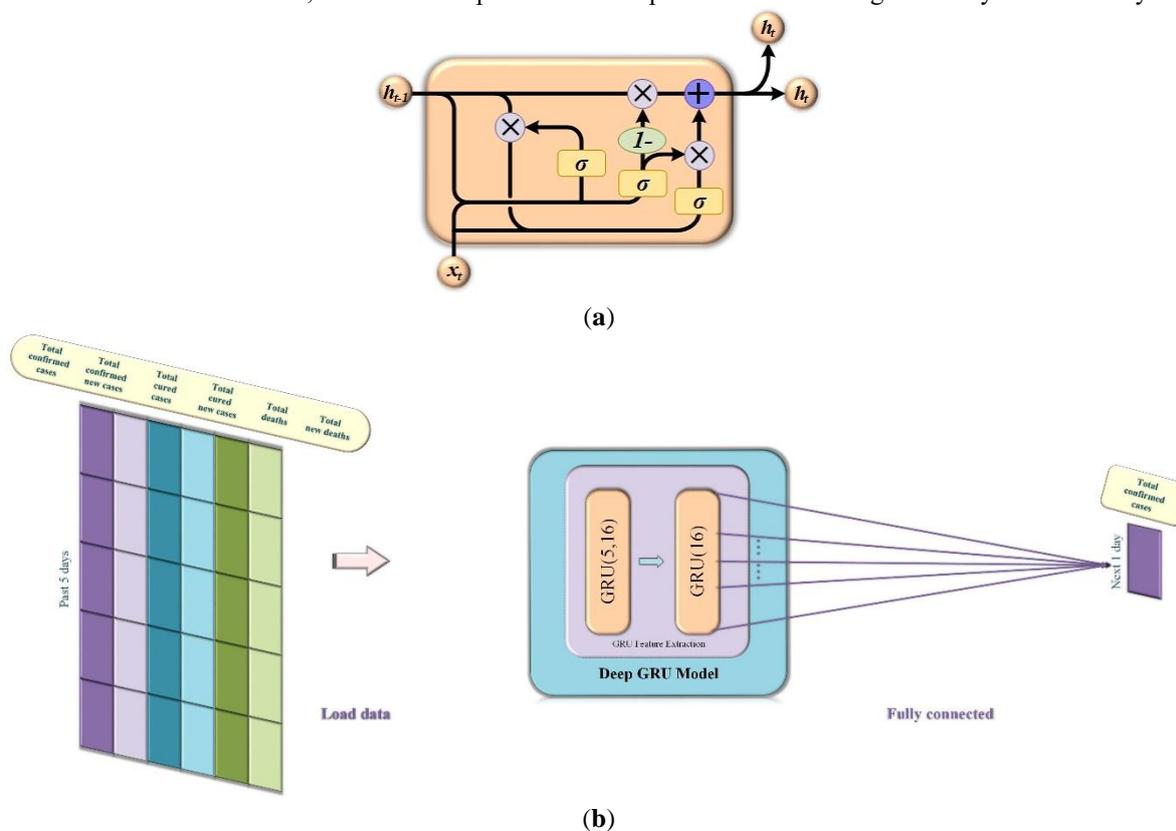
235
$$\tilde{h}_t = \tanh(W \cdot [r_t * h_{t-1}, x_t]) \quad (12)$$

236 Equation (13) combines the reserved long-term memory z_t and the information of the current moment
 237 \tilde{h}_t to obtain new information h_t .

238
$$h_t = (1 - z_t) h_{t-1} + z_t \tilde{h}_t \quad (13)$$

239 The structure of the proposed deep GRU is shown in Fig. 5 (b). The structure is composed of two GRU
 240 layers, and each layer has 16 neural units. The input structure is constituted by six factors influencing the
 241 cumulative number of cases, with a time step of 5. A 1D output is obtained through the fully connected layer.

242
 243



244
 245
 246

247 **Fig. 5.** Structure of the comparative deep neural network GRU adopted in this study. (a) GRU neural unit; (b)
 248 the deep neural network GRU model with an input layer consisting six input signals.

249 The pseudocode of the proposed CNN is described in Algorithm 1. Specifically, the dataset and test set
 250 are first loaded. The input data are then reconstructed into 3D matrixes, and the output data format is set to the
 251 2D matrix. The CNN model is constructed, and related CNN parameters are set. Before the training starts, the
 252 weighting value w is initialized, and the repeat operator a is initialized as 0. In the training process, the
 253 training samples are trained in batches, and the loss function is calculated once per batch to update the weight
 254 value w . After the training, the test set data are registered to the trained network. The errors are calculated, and
 255 the actual values and error values are compared.

256

Algorithm 1 Convolutional Neural Network

Input, target, and hyperparameter setting:

Load the training set and testing set.

Reshape the input into the corresponding three-dimensional matrix.

1. The first dimension is the sample index.
2. The second dimension is the time step.
3. The third dimension is the characteristic.

Reshape the output into a two-dimensional matrix.

1. The first dimension is the sample target index.
2. The second dimension is the sample target.

Set relevant parameters of the CNN model.

Training:

Initialize the weight w

$s \rightarrow$ Total number of training set samples

$c \rightarrow$ Batch size

$a \rightarrow 0$

for i to epoch then

for k to s then

$j = k \bmod c$

if $j = 0$ then

$j = c$

end

if $j - a > 0$ then

Calculate the output using Equations (1) and (2) for training set samples in the batch.

else

Calculate the loss using loss function, update the weight using optimizer, and calculate the output using Equations (1) and (2) for training set samples in the next batch.

end

$a = j$

end

end

Testing:

The established CNN model is used to predict the testing set. The real value is compared with the predicted value, and the prediction error is calculated.

257 4. Experimental Results and Discussion

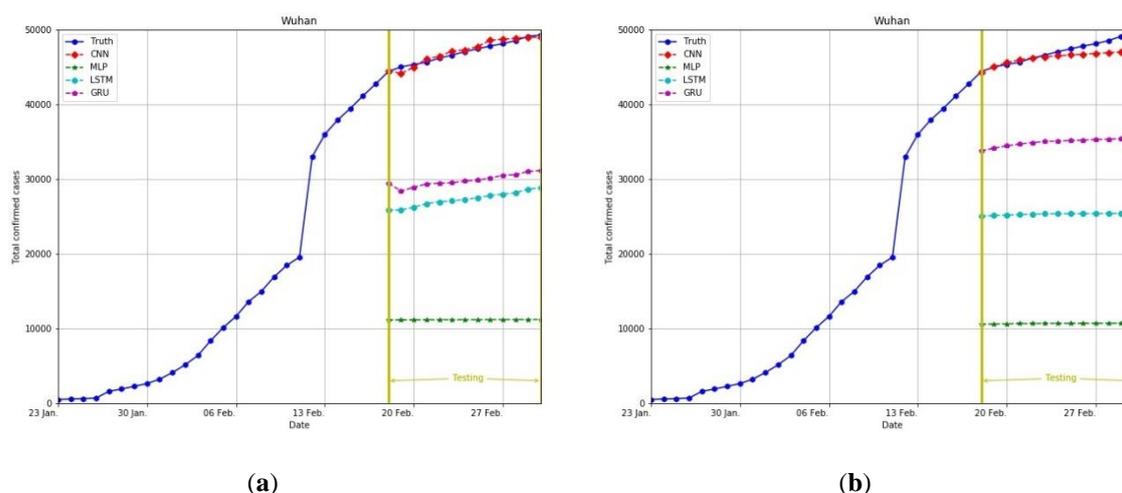
258 Data on confirmed cases of COVID-19 from January 23, 2020 to March 2, 2020, and from January
259 23, 2020 to March 2, 2020, were obtained from Surging News Network (a media outlet) [9] and WHO
260 [31], respectively. In this experiment, the two evaluation indexes of the mean absolute error (MAE) and root
261 mean square error (RMSE) were used; their formulae are presented in Equations (14) and (15). To conduct
262 comprehensive efficacy tests, the dataset for January 23 to February 17 was chosen as the training data, and the
263 dataset for February 18 to March 2 was used as testing data. Figs. 6 to 12 present the comparative prediction
264 results of all the algorithms. The detailed MAE and RMSE values are listed in Tables I, II, III, and IV. The

265 experiment results indicated that the GRU has decent efficacy, and that CNN is the best performing algorithm
 266 among its counterparts tested. The experiment demonstrated that the characteristic extraction of CNN is very
 267 helpful for predicting the number of confirmed cases of COVID-19.
 268

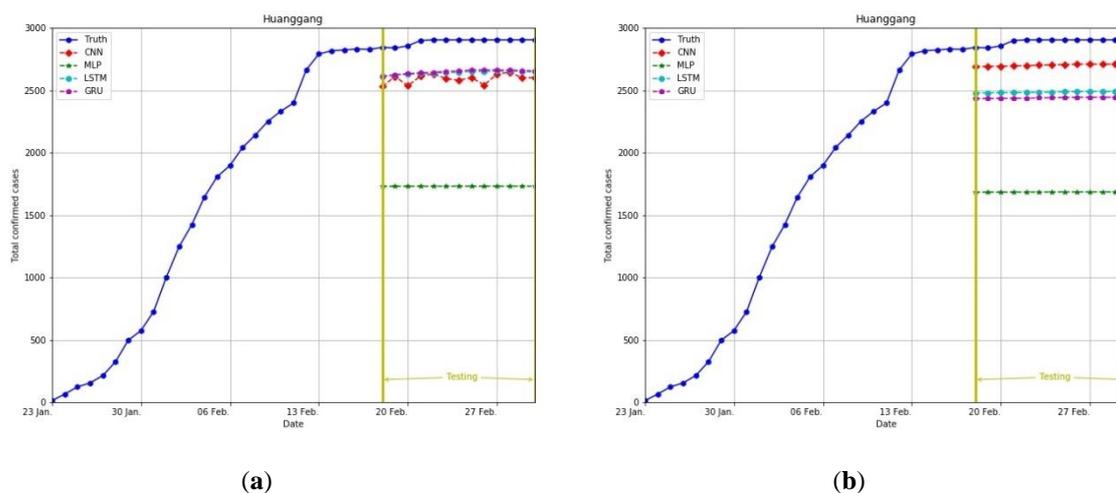
$$\text{MAE} = \frac{1}{N} \sum_{n=1}^N |y_n - \hat{y}_n| \quad (14)$$

$$\text{RMSE} = \sqrt{\frac{\sum_{n=1}^N (y_n - \hat{y}_n)^2}{N}} \quad (15)$$

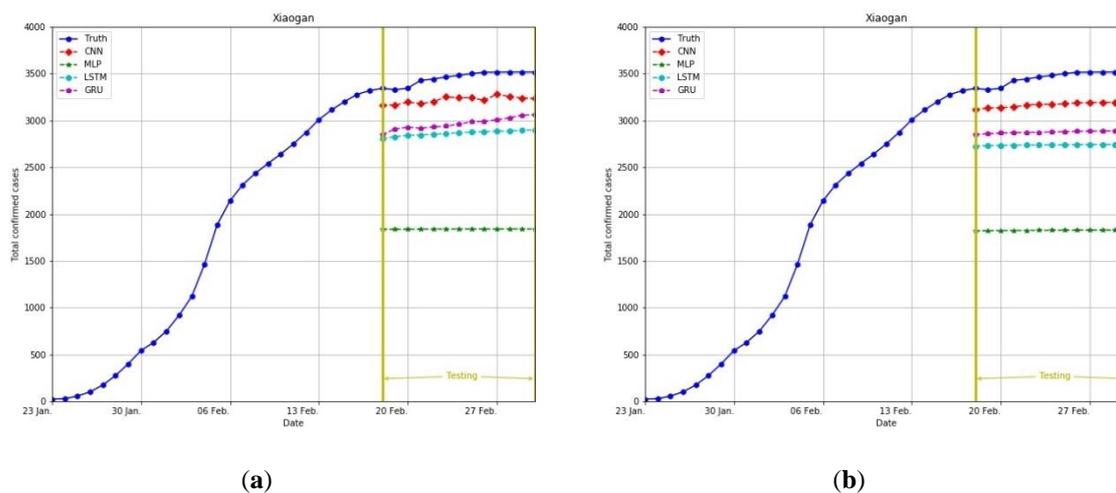
269



270 **Fig. 6.** Comparison of prediction results for the cumulative number of confirmed cases in Wuhan City, Hubei
 271 Province, China. The comparison used the deep neural networks CNN, MLP, LSTM, and GRU. (a) Six
 272 important factors were adopted as the input layer: new confirmed cases, new deceased cases, new cured cases,
 273 cumulative confirmed cases, cumulative deceased cases, and cumulative cured cases. (b) Only one important
 274 factor was adopted as the input layer: cumulative confirmed cases.

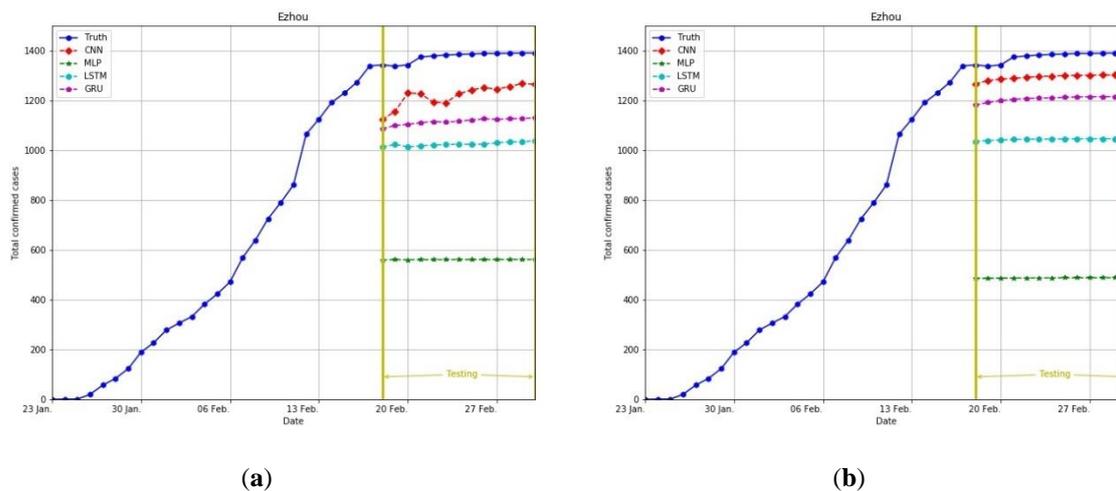


275 **Fig. 7.** Comparison of prediction results for the cumulative number of confirmed cases in Huanggang City,
 276 Hubei Province, China. The comparison used the deep neural networks CNN, MLP, LSTM, and GRU. (a) Six
 277 important factors were adopted as the input layer: new confirmed cases, new deceased cases, new cured cases,
 278 cumulative confirmed cases, cumulative deceased cases, and cumulative cured cases. (b) Only one important
 279 factor was adopted as the input layer: cumulative confirmed cases.



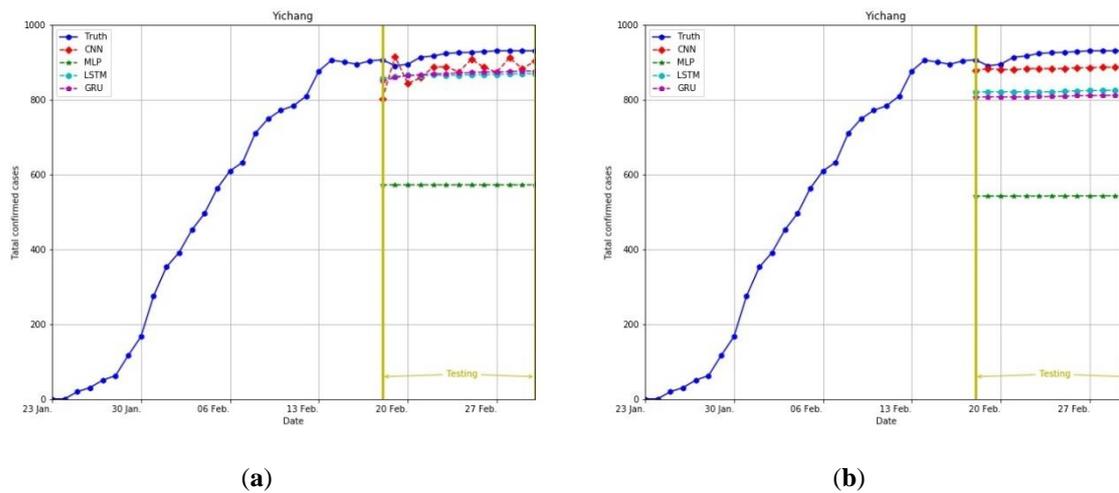
280
281
282
283
284

Fig. 8. Comparison of prediction results for the cumulative number of confirmed cases in Xiaogan City, Hubei Province, China. The comparison used the deep neural networks CNN, MLP, LSTM, and GRU. (a) Six important factors were adopted as the input layer: new confirmed cases, new deceased cases, new cured cases, cumulative confirmed cases, cumulative deceased cases, and cumulative cured cases. (b) Only one important factor was adopted as the input layer: cumulative confirmed cases.



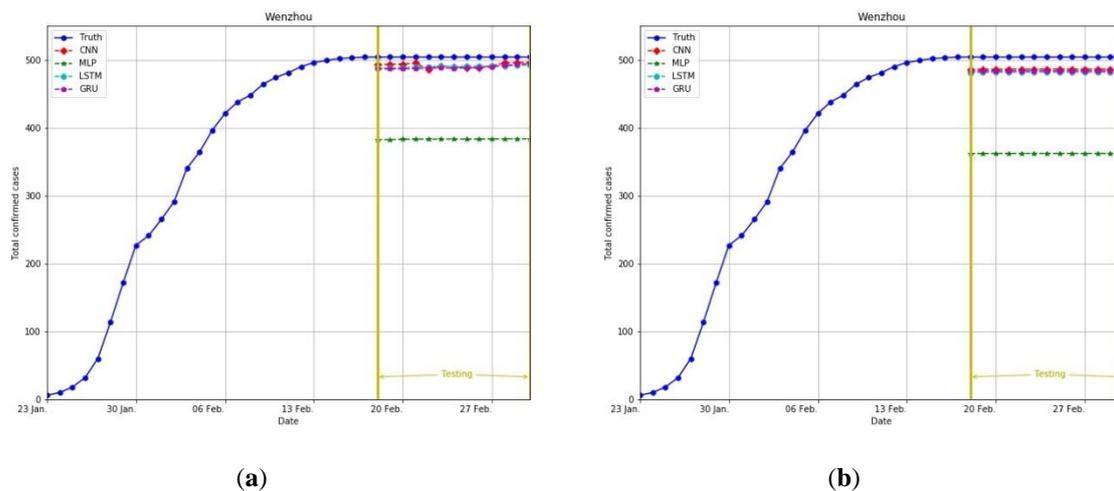
285
286
287
288
289

Fig. 9. Comparison of prediction results for the cumulative number of confirmed cases in Ezhou City, Hubei Province, China. The comparison used the deep neural networks CNN, MLP, LSTM, and GRU. (a) Six important factors were adopted as the input layer: new confirmed cases, new deceased cases, new cured cases, cumulative confirmed cases, cumulative deceased cases, and cumulative cured cases. (b) Only one important factor was adopted as the input layer: cumulative confirmed cases.



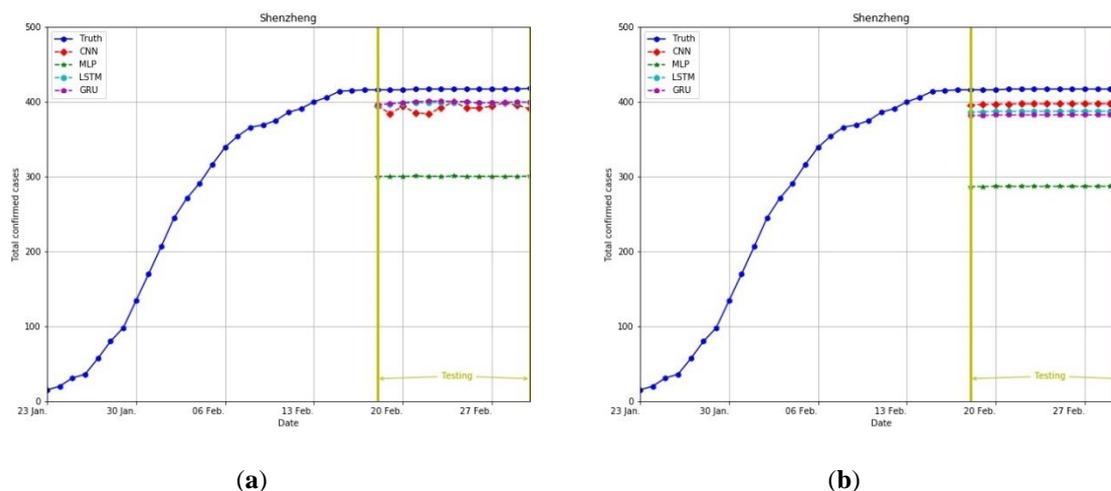
290
291
292
293
294

Fig. 10. Comparison of prediction results for the cumulative number of confirmed cases in Yichang City, Hubei Province, China. The comparison used the deep neural networks CNN, MLP, LSTM, and GRU. (a) Six important factors were adopted as the input layer: new confirmed cases, new deceased cases, new cured cases, cumulative confirmed cases, cumulative deceased cases, and cumulative cured cases. (b) Only one important factor was adopted as the input layer: cumulative confirmed cases.



295
296
297
298
299

Fig. 11. Comparison of prediction results for the cumulative number of confirmed cases in Wenzhou City, Zhejiang Province, China. The comparison used the deep neural networks CNN, MLP, LSTM, and GRU. (a) Six important factors were adopted as the input layer: new confirmed cases, new deceased cases, new cured cases, cumulative confirmed cases, cumulative deceased cases, and cumulative cured cases. (b) Only one important factor was adopted as the input layer: cumulative confirmed cases.



300 **Fig. 12.** Comparison of prediction results for the cumulative number of confirmed cases in Shenzhen City,
 301 Guangdong Province, China. The comparison used the deep neural networks CNN, MLP, LSTM, and GRU. (a)
 302 Six important factors were adopted as the input layer: new confirmed cases, new deceased cases, new cured
 303 cases, cumulative confirmed cases, cumulative deceased cases, and cumulative cured cases. (b) Only one
 304 important factor was adopted as the input layer: cumulative confirmed cases.

305 **TABLE I**
 306 **EXPERIMENTAL RESULTS IN TERMS OF MEAN ABSOLUTE ERROR OF THE INPUT LAYER OF SIX NEURONS**

City	GRU	LSTM	MLP	CNN
Wuhan	17119.101	19671.849	35797.372	426.179
Huanggang	243.399	250.142	1159.923	89.026
Xiaogan	489.004	592.220	1616.017	87.343
Ezhou	259.843	350.798	814.801	77.942
Yichang	50.219	53.632	346.853	31.688
Wenzhou	15.067	13.988	120.844	4.617
Shenzhen	17.414	18.200	116.240	3.808
Avg.	2599.150	2992.976	5710.293	102.943

307 **TABLE II**
 308 **EXPERIMENTAL RESULTS IN TERMS OF MEAN ABSOLUTE ERROR OF THE INPUT LAYER OF A SINGLE NEURON**

City	GRU	LSTM	MLP	CNN
Wuhan	12054.218	21669.078	36307.670	1131.627
Huanggang	450.634	405.468	1206.190	255.561
Xiaogan	580.525	717.547	1629.771	389.392
Ezhou	168.118	331.491	887.933	102.507
Yichang	109.966	96.664	376.636	63.651
Wenzhou	20.417	22.303	141.863	22.138
Shenzhen	34.452	29.589	129.942	27.136
Avg.	1916.904	3324.591	5811.429	284.573

310 **TABLE III**
 311

312 EXPERIMENTAL RESULTS IN TERMS OF ROOT MEAN SQUARE ERROR OF THE INPUT LAYER OF SIX NEURONS

City	GRU	LSTM	MLP	CNN
Wuhan	17140.879	19681.133	35829.022	456.910
Huanggang	243.715	250.594	1160.173	94.271
Xiaogan	490.381	593.989	1617.454	90.685
Ezhou	260.019	351.126	815.025	81.469
Yichang	50.960	54.844	347.112	32.855
Wenzhou	15.264	14.051	120.845	5.106
Shenzhen	17.446	18.217	116.241	4.779
Avg.	2602.666	2994.851	5715.125	109.439

313

314

315

TABLE IV

316

EXPERIMENTAL RESULTS IN TERMS OF ROOT MEAN SQUARE ERROR OF THE INPUT LAYER OF A SINGLE

317

NEURON

City	GRU	LSTM	MLP	CNN
Wuhan	12101.919	21715.728	36338.153	1416.082
Huanggang	451.136	406.029	1206.424	256.246
Xiaogan	583.509	720.488	1631.175	391.768
Ezhou	168.461	331.896	888.127	102.974
Yichang	110.629	97.439	376.867	64.649
Wenzhou	20.417	22.303	141.863	22.139
Shenzhen	34.454	29.591	129.943	27.139
Avg.	1924.361	3331.925	5816.079	325.857

318

Regarding the deep neural network prediction results for six and one-input factors, the MAE and RMSE values are listed in Tables 1 to 4. The MAE results from low to high are as follows: CNN (102.943, 284.573), GRU (2599.150, 1916.904), LSTM (2992.976, 3324.591), and MLP (5710.293, 5811.429). The RMSE results from low to high are as follows: CNN (109.439, 325.857), GRU (2602.666, 1924.361), LSTM (2994.851, 3331.925), and MLP (5715.125, 5816.079). The experiment results demonstrated that GRU and LSTM have decent efficacies and that CNN had the best performance. This experiment also demonstrated that the procedure of initial characteristic extraction through CNN and the subsequent input of characteristic values to the CNN structure greatly aid the prediction of the cumulative number of confirmed cases of COVID-19.

319

320

321

322

323

324

325



(a)

(b)

326

327 **Fig. 13.** MAE radar charts. (a) Six factors were adopted as the input layer. (b) Only one factor was adopted as the input
328 layer.

329 Fig. 13 (a) and (b) presents the radar charts of MAE values for the comparison between the output results
330 and real values for six and one factors, respectively. Regardless of whether one or six factors were used, the
331 prediction results, in terms of MAE, of the proposed CNN model were superior to those of its counterparts. For
332 the proposed CNN model, the use of six factors yielded better prediction results, in terms of MAE, in
333 comparison to the use of only one factor, particularly for Wuhan City.



334 **Fig. 14.** RMSE radar charts. (a) Six pieces of important information were adopted as the input layer. (b) Only one piece of
335 information was adopted as the input layer.
336

337 Fig. 14 (a) and (b) presents the radar charts of RMSE values for the comparison between the output results
338 and real values for six and one factors, respectively. The results further demonstrated the excellent predictive
339 performance of the proposed CNN model for the cumulative number of confirmed cases of COVID-19, making
340 the model a valuable reference for other countries in their establishment of country-specific COVID-19
341 prediction models. In general, the prediction performances of the models from most favorable to least favorable
342 were those for CNN, GRU, LSTM, and MLP.

343 5. Conclusions

344 A novel and multi-input CNN deep neural network model was proposed in this study to predict the
345 cumulative number of confirmed cases of COVID-19. The cumulative number of confirmed cases in the next
346 day is predicted according to the previous five days' total number of confirmed cases, total confirmed new cases,
347 total cured cases, total cured new cases, total deaths, and total new deaths. In the experiment, the datasets of
348 those seven Chinese cities with severe confirmed cases—from Hubei Province, Guangdong Province, and
349 Zhejiang Province—were used for the models' training and prediction. Because the COVID-19 epidemic is
350 ongoing, the algorithm proposed in this study can rapidly use small datasets to establish models with high
351 predictive accuracy, different from many other studies. In addition, with the algorithm, a deep learning network
352 prediction model for the number of confirmed cases of COVID-19 was established, and the verification and
353 comparison were conducted among different deep learning algorithms. The accuracy and reliability of the deep
354 learning algorithm were verified by having it predict the future trend of COVID-19. Furthermore, experiments
355 for multiple cities with more severe confirmed cases in China indicated that the prediction model of this study
356 had the lowest error rate among its counterparts tested. However, at the time of writing, countries other than
357 China have had COVID-19 outbreaks, such as Italy, South Korea, Iran, Spain, France, and Germany. In the
358 future, the establishment of more complete databases will lead to improvements to the proposed prediction
359 model. For example, deep learning networks with a mixed structure can be introduced to establish more accurate
360 models, which can be applied to more countries. The predicted trends can aid the containment of the COVID-
361 19 epidemic and extend the scope of application of artificial intelligence.
362

363 References

- 364 1. Zhu, N.; Zhang, D.; Wang, W.; Li, X.; Yang, B.; Song, J.; Zhao, X.; Huang, B.; Shi, W.; Lu, R.; Niu, P.;
365 Zhan, F.; Ma, X.; Wang, D.; Xu, W.; Wu, G.; Gao, G. F.; Tan, W.; China Novel Coronavirus, I.; Research,

- 366 T., A Novel Coronavirus from Patients with Pneumonia in China, 2019. *N Engl J Med* **2020**, *382*, (8), 727-
367 733.
- 368 2. Locations with Confirmed COVID-19 Cases Global Map Available online:
369 <https://www.cdc.gov/coronavirus/2019-ncov/locations-confirmed-cases.html>.
- 370 3. Roosa, K.; Lee, Y.; Luo, R.; Kirpich, A.; Rothenberg, R.; Hyman, J. M.; Yan, P.; Chowell, G., Short-term
371 Forecasts of the COVID-19 Epidemic in Guangdong and Zhejiang, China: February 13-23, 2020. *J Clin*
372 *Med* **2020**, *9*, (2).
- 373 4. Z. Liu, P. M., O. Seydi, G. Webb, Predicting the cumulative number of cases for the COVID-19 epidemic
374 in China from early data. **2020**.
- 375 5. Boldog, P.; Tekeli, T.; Vizi, Z.; Dénes, A.; Bartha, F. A.; Röst, G., Risk Assessment of Novel Coronavirus
376 COVID-19 Outbreaks Outside China. *Journal of Clinical Medicine* **2020**, *9*, (2).
- 377 6. Al-Qaness, M. A. A.; Ewees, A. A.; Fan, H.; Abd El Aziz, M., Optimization Method for Forecasting
378 Confirmed Cases of COVID-19 in China. *J Clin Med* **2020**, *9*, (3).
- 379 7. Jung, S.-m.; Akhmetzhanov, A. R.; Hayashi, K.; Linton, N. M.; Yang, Y.; Yuan, B.; Kobayashi, T.;
380 Kinoshita, R.; Nishiura, H., Real-Time Estimation of the Risk of Death from Novel Coronavirus
381 (COVID-19) Infection: Inference Using Exported Cases. *Journal of Clinical Medicine* **2020**, *9*, (2).
- 382 8. Fan, C.; Liu, L.; Guo, W.; Yang, A.; Ye, C.; Jilili, M.; Ren, M.; Xu, P.; Long, H.; Wang, Y., Prediction of
383 Epidemic Spread of the 2019 Novel Coronavirus Driven by Spring Festival Transportation in China: A
384 Population-Based Study. *Int J Environ Res Public Health* **2020**, *17*, (5).
- 385 9. Yang, W.; Cao, Q.; Qin, L.; Wang, X.; Cheng, Z.; Pan, A.; Dai, J.; Sun, Q.; Zhao, F.; Qu, J.; Yan, F., Clinical
386 characteristics and imaging manifestations of the 2019 novel coronavirus disease (COVID-19): A multi-
387 center study in Wenzhou city, Zhejiang, China. *J Infect* **2020**.
- 388 10. Zixin Hu , Q. G., Shudi Li , Li Jin and Momiao Xiong Artificial Intelligence Forecasting of Covid-19 in
389 China.
- 390 11. Guo, Q.; Li, M.; Wang, C.; Wang, P.; Fang, Z.; tan, J.; Wu, S.; Xiao, Y.; Zhu, H., Host and infectivity
391 prediction of Wuhan 2019 novel coronavirus using deep learning algorithm. **2020**.
- 392 12. Metsky, H. C.; Freije, C. A.; Kosoko-Thoroddsen, T.-S. F.; Sabeti, P. C.; Myhrvold, C., CRISPR-based
393 surveillance for COVID-19 using genomically-comprehensive machine learning design. **2020**.
- 394 13. Riou, J.; Althaus, C. L., PATTERN OF EARLY HUMAN-TO-HUMAN TRANSMISSION OF WUHAN
395 2019- NCOV. **2020**.
- 396 14. Liu, T.; Hu, J.; Xiao, J.; He, G.; Kang, M.; Rong, Z.; Lin, L.; Zhong, H.; Huang, Q.; Deng, A.; Zeng, W.;
397 Tan, X.; Zeng, S.; Zhu, Z.; Li, J.; Gong, D.; Wan, D.; Chen, S.; Guo, L.; Li, Y.; Sun, L.; Liang, W.; Song, T.;
398 He, J.; Ma, W., Time-varying transmission dynamics of Novel Coronavirus Pneumonia in China. **2020**.
- 399 15. Ming, W.-K.; Huang, J.; Zhang, C. J. P., Breaking down of healthcare system: Mathematical modelling
400 for controlling the novel coronavirus (2019-nCoV) outbreak in Wuhan, China. **2020**.
- 401 16. Zhao, N.; Wang, J.; Yu, Y.; Zhao, J.-Y.; Chen, D.-B., Spreading predictability in complex networks. **2020**.
- 402 17. Fountain-Jones, N.; Machado, G.; Carver, S.; Packer, C.; Mendoza, M.; Craft, M. E., How to make more
403 from exposure data? An integrated machine 1 learning pipeline to predict pathogen exposure **2019**.
- 404 18. Benvenuto, D.; Giovanetti, M.; Vassallo, L.; Angeletti, S.; Ciccozzi, M., Application of the ARIMA model
405 on the COVID-2019 epidemic dataset. *Data in Brief* **2020**, *29*.
- 406 19. Li, Q.; Feng, W.; Quan, Y. H., Trend and forecasting of the COVID-19 outbreak in China. *J Infect* **2020**.

- 407 20. Tian, S.; Hu, N.; Lou, J.; Chen, K.; Kang, X.; Xiang, Z.; Chen, H.; Wang, D.; Liu, N.; Liu, D.; Chen, G.;
408 Zhang, Y.; Li, D.; Li, J.; Lian, H.; Niu, S.; Zhang, L.; Zhang, J., Characteristics of COVID-19 infection in
409 Beijing. *J Infect* **2020**.
- 410 21. Chen, H.; Guo, J.; Wang, C.; Luo, F.; Yu, X.; Zhang, W.; Li, J.; Zhao, D.; Xu, D.; Gong, Q.; Liao, J.; Yang,
411 H.; Hou, W.; Zhang, Y., Clinical characteristics and intrauterine vertical transmission potential of
412 COVID-19 infection in nine pregnant women: a retrospective review of medical records. *The Lancet* **2020**,
413 395, (10226), 809-815.
- 414 22. Pham, T. A.; Ly, H.-B.; Tran, V. Q.; Giap, L. V.; Vu, H.-L. T.; Duong, H.-A. T., Prediction of Pile Axial
415 Bearing Capacity Using Artificial Neural Network and Random Forest. *Applied Sciences* **2020**, 10, (5).
- 416 23. Khalid, Z.; Abbas, G.; Awais, M.; Alquthami, T.; Rasheed, M. B., A Novel Load Scheduling
417 Mechanism Using Artificial Neural Network Based Customer Profiles in Smart Grid. *Energies* **2020**, 13,
418 (5).
- 419 24. Wang, Y.; Li, Y.; Song, Y.; Rong, X., The Influence of the Activation Function in a Convolution Neural
420 Network Model of Facial Expression Recognition. *Applied Sciences* **2020**, 10, (5).
- 421 25. Bai, T.; Pang, Y.; Wang, J.; Han, K.; Luo, J.; Wang, H.; Lin, J.; Wu, J.; Zhang, H., An Optimized Faster R-
422 CNN Method Based on DRNet and RoI Align for Building Detection in Remote Sensing Images. *Remote*
423 *Sensing* **2020**, 12, (5).
- 424 26. Abo-Tabik, M.; Costen, N.; Darby, J.; Benn, Y., Towards a Smart Smoking Cessation App: A 1D-CNN
425 Model Predicting Smoking Events. *Sensors (Basel)* **2020**, 20, (4).
- 426 27. Zhang, Q.; Gao, T.; Liu, X.; Zheng, Y., Public Environment Emotion Prediction Model Using LSTM
427 Network. *Sustainability* **2020**, 12, (4).
- 428 28. Zhang, M.; Geng, G.; Chen, J., Semi-Supervised Bidirectional Long Short-Term Memory and
429 Conditional Random Fields Model for Named-Entity Recognition Using Embeddings from Language
430 Models Representations. *Entropy* **2020**, 22, (2).
- 431 29. Jin, X. B.; Yang, N. X.; Wang, X. Y.; Bai, Y. T.; Su, T. L.; Kong, J. L., Hybrid Deep Learning Predictor for
432 Smart Agriculture Sensing Based on Empirical Mode Decomposition and Gated Recurrent Unit Group
433 Model. *Sensors (Basel)* **2020**, 20, (5).
- 434 30. Batur DİNler, Ö.; Aydın, N., An Optimal Feature Parameter Set Based on Gated Recurrent Unit
435 Recurrent Neural Networks for Speech Segment Detection. *Applied Sciences* **2020**, 10, (4).
- 436 31. Coronavirus disease (COVID-2019) situation reports:
437 <http://www.who.int/emergencies/diseases/novel-coronavirus-2019/situation-reports>.

438