Hindawi
Discrete Dynamics in Nature and Society
Volume 2020, Article ID 6152041, 8 pages
https://doi.org/10.1155/2020/6152041

Research Article

Learning from Large-Scale Wearable Device Data for Predicting the Epidemic Trend of COVID-19

Guokang Zhu, Jia Li, Zi Meng, Yi Yu, Yanan Li, Xiao Tang, Yuling Dong, Guangxin Sun, Rui Zhou, Hui Wang, Kongqiao Wang, and Wang Huang

Huami Corporation, Hefei, China

Correspondence should be addressed to Kongqiao Wang; kongqiao.wang@huami.com

Received 27 March 2020; Accepted 15 April 2020; Published 5 May 2020

Guest Editor: Jinchang Ren

The coronavirus disease 2019 (COVID-19) pandemic has triggered a new response involving public health surveillance. The popularity of personal wearable devices creates a new opportunity for tracking and precaution of spread of such infectious diseases. In this study, we propose a framework, which is based on the heart rate and sleep data collected from wearable devices, to predict the epidemic trend of COVID-19 in different countries and cities. In addition to a physiological anomaly detection algorithm defined based on data from wearable devices, an online neural network prediction modelling methodology combining both detected physiological anomaly rate and historical COVID-19 infection rate is explored. Four models are trained separately according to geographical segmentation, i.e., North China, Central China, South China, and South-Central Europe. The anonymised sensor data from approximately 1.3 million wearable device users are used for model verification. Our experiment's results indicate that the prediction models can be utilized to alert to an outbreak of COVID-19 in advance, which suggests there is potential for a health surveillance system utilising wearable device data.

1. Introduction

Since the outbreak of the coronavirus disease 2019 (COVID-19) pandemic, more than 300,000 people have been infected in at least 127 countries as of March 23, 2020, according to the World Health Organization's (WHO's) report [1]. COVID-19 spreads easily from person to person and has killed thousands of people [2–5]. Since the beginning of the COVID-19 outbreak, several studies have been carried out to forecast the epidemic trend of COVID-19 in China [6–8]. For example, Wu et al. built a Susceptible-Exposed-Infectious-Recovered (SEIR) model to simulate the epidemics across the major cities in China [7]. Yang et al. applied the Long Short Term Memory (LSTM) model to predict the number of newly infected COVID-19 cases by utilizing data from the outbreak of Severe Acute Respiratory Syndrome (SARS) in 2003 [6]. Although the models used in those studies could simulate the outbreak trend of the disease, they relied heavily on officially reported statistics; therefore, the timeliness of the models could be affected. On the contrary, big data analysis, such as analysis of Internet data, may provide real-time surveillance and improve the timeliness of the forecasting [9–16]. For instance, Google invented the influenza epidemic prediction tool Google Flu Trend (GFT) to estimate the level of in-fluenza activity based on the individual web search queries from different regions [9–11]. They assumed that more individuals in a certain region might search online for the information about specific diseases if the influenza disease risk was higher in that certain region. Therefore, Google built a database containing 50 million of the most common web search queries on all influenza-related topics and constructed the risk prediction model GFT using this search query data as the input [9]. Google showed that GFT could help predict the influenza-like illness outbreak 7–10 days before the Centers for Disease Control and Prevention (CDC) report [10]. In fact, the surveillance report from CDC usually has a lag time of around 1-2 weeks. Therefore, the result from Google indicated that big data analysis could improve timeliness for public health surveillance. However, search queries can be

greatly influenced by social hotspots, which weakens the correlation between the search queries and the occurrence of influenza-like diseases [17].

With the rise in popularity of fitness band and smartwatch devices, physiological signs, such as heart rate, activity, sleep, etc., can be conveniently acquired from these wearable biosensors [18–20]. As of 2019, more than 100 million consumers owned Huami wearable devices, and the number continues to grow. In contrast with the big data from web search engines, data from wearable devices can provide more objective information on the health status of the users. For example, once users are infected with an influenza-like illness, their physiological signs would be altered. Radin et al. explored the relationship between the physiological anomaly rate from wearable device users and the influenza-like illness rate reported by the US CDC [21] to build the regression models for predicting the influenza-like illness cases within different states of America. They utilized the heart rate and sleep data from the wearable devices to improve upon the standard models. The prediction results have strong correlation with the official data. Li et al. also investigated the role of physiological changes measured with wearable devices on the diagnosis and analysis of disease [22]. The researchers established a personalized disease detection framework, which identifies abnormal physical signs, e.g., from Lyme disease and other inflammatory responses, from the longitudinal data of the individuals. All the studies mentioned above can inform the way wearable device data is used for public health surveillance.

According to clinical studies [23–25], the most common symptoms at the onset of COVID-19 are fever, cough, and fatigue, which are closely related to the physiological signs measured by the wearable devices. Therefore, a good method to predict the epidemic trend of COVID-19 may involve building a prediction model based on the wearable device data.

The main purpose of this study is to provide a novel framework for predicting the trend of COVID-19 outbreak within different countries and cities, using big data collected from wearable devices. There are two major contributions from this study: (1) a physiological anomaly detection method is developed and can identify the anomalous signs reflected by the physiological data from wearable sensors; (2) an online learning framework is proposed for public health emergency surveillance.

2. Methodology

2.1. Physiological Anomaly Detection. According to a study on fever and cardiac rhythm [26], heart rate increases by 8.5 beats per minute, on average, for every 1°C increase in body temperature, so an elevated resting heart rate (RHR) might be related to fever caused by COVID-19 or influenza-like illness. The basic anomaly detection method is based on the elevated RHR. Because shortened sleep length also causes an increase in RHR [27], we weaken the contribution of this factor in the physiological anomaly detection method.

RHR and sleep length are directly acquired with the corresponding sensors of Huami wearable devices. Both kinds of synchronized data from the accelerometer (ACC) sensor and the photoplethysmography (PPG) sensor are used to analyze sleep status (including sleep recognition and stage) for measuring sleep length. During sleep, the PPG data is used to compute the RHR. For each user, overall mean and standard deviation (SD) of RHR and sleep length throughout the entire period are calculated. A daily RHR is defined as an anomaly if it is larger than the average RHR plus 1.5 SD, and if in addition, the daily sleep is longer than the average sleep minus 0.5 SD. Considering that COVID-19 or influenza-like illness persist for several days, we define the detection standard of physiological anomaly as continuous anomaly measured for at least five consecutive days.

2.2. Online Prediction of COVID-19 Infection Rate. The physiological anomaly detected by our method is an indication of fever, which in fact can be caused by COVID-19 or other influenza-like illness. Thus, the key point for COVID-19 infection rate prediction is to distinguish an anomaly arising from COVID-19 from the wider category of physiological anomalies. To this end, as shown in Figure 1, a heterogeneous neural network [28] regression model combining sparse categorical features and dense numerical features (CDNet) is proposed.

CDNet concatenates 2 subnetworks: CatNN and DenNN. The inputs of the CatNN are sparse categorical features, i.e., holiday activity, season, and weather. The inputs of the DenNN are historical physiological anomaly rate, active user density, and historical officially reported COVID-19 rate, where the historically detected physiological anomaly rate is calculated with dividing the number of users detected with a physiological anomaly by the number of total active users. The output layer of CDNet normalized by a Sigmoid function outputs the predicted physiological anomaly rate. The detailed inputs and outputs are summarized in

$$R'_{t+1,k} = \mathrm{CDNet}\Big(\big\{R_{t-j,k}, r_{t-j,k}, C_{t-j,k}, c_{t-j,k},\ j = 0, 1, \ldots, 6\big\}, RC_{t,k}, D_{t,k}\Big), \tag{1}$$

where, for country or city k, the output $R'_{t+1,k}$ is the predicted physiological anomaly rate in the next period, $R_{t-j,k}$ is the physiological anomaly rate the j-th period earlier, $r_{t-j,k}$ is the physiological anomaly rate in the same period of $R_{t-j,k}$ last year, $C_{t-j,k}$ and $c_{t-j,k}$ are the corresponding categorical information with the same temporal definition as $R_{t-j,k}$ and $r_{t-j,k}$, respectively, $RC_{t,k}$ is the officially reported COVID-19 rate (ratio of confirmed COVID-19 patient number to the number of residents in the country or city) in the current period, $D_{t,k}$ is the current active user density (ratio of active user number to the number of residents in the country or city). To distinguish regional disparity, four different CDNet models are trained separately for North China, Central China, South China, and South-Central Europe.

In order to get the predicted anomaly rate caused by COVID-19 for the next period, the predicted physiological anomaly rate with ($R_{t+1,k}$) and without ($R'_{t+1,k}|RC_{t,k} = 0$) the supervision of officially reported data is calculated separately. As shown in Figure 2, the supervision is removed by

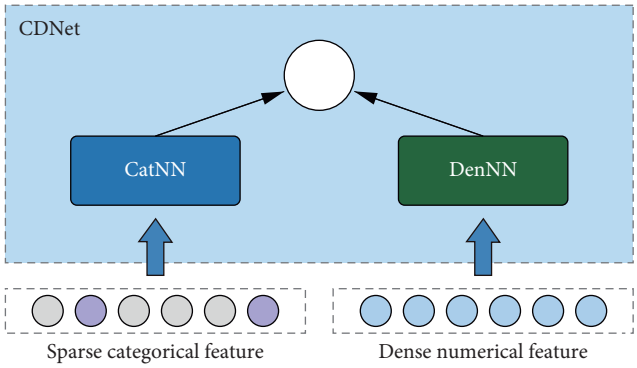

Figure 1: Diagram of the CDNet neural network architecture.

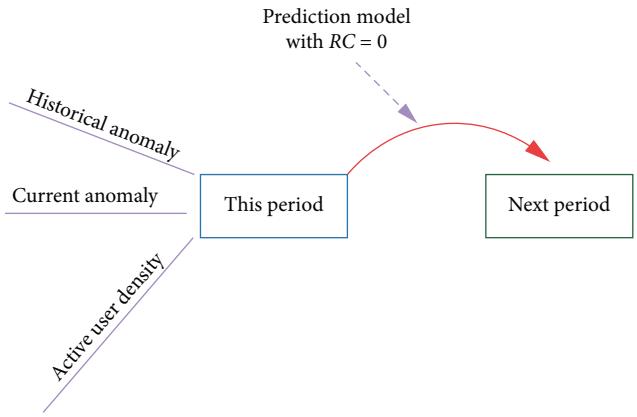

Figure 2: Illustration for the prediction stage of the model.

setting $RC_{t,k}$ as 0. Then, the predicted anomaly rate caused by COVID-19 for the next period $P'_{t+1,k}$ can be calculated as the difference between $R'_{t+1,k}$ and $R_{t+1,k} \mid RC_{t,k} = 0$:

$$\mathbf{P}'_{t+1,k} = \mathbf{R}'_{t+1,k} - \left\{\mathbf{R}'_{t+1,k} \mid \mathbf{RC_{t,k}} = 0\right\}. \quad (2)$$

To consecutively predict the epidemic trend of COVID-19, the CDNet model is trained in an online learning way. As shown in Figure 3, the initial CDNet model M_0 is trained with the input of $\{R_{t-j,k}, r_{t-j,k}, C_{t-j,k}, c_{t-j,k}, j = 1, 2, \ldots, 7\}$, and with the target as $R_{t,k}$. The weights of CDNet are updated step by step with the transmission of COVID-19, using the arriving data of newly officially reported COVID-19 rate and detected physiological anomaly rate. The step size of the sliding window for online learning is set as 1 week.

3. Experiments

3.1. Dataset. Anonymised sensor data of approximately 1.3 million users who wore Huami devices from July 1, 2017, to April 8, 2020 were obtained according to appropriate security control processes. All users are notified that their anonymised data could potentially be used for academic research under the Huami Privacy Policy.

All the users wore their Huami devices for at least 100 days throughout the entire period. Daily measures include RHR, activity, and sleep length, which are the bases of physiological anomaly detection. Data with missing RHR or sleep length were excluded. The daily COVID-19 infection rate data come from CDC of the corresponding countries.

We build separate models for different countries and cities listed in Table 1, according to the geographical segmentation considering the regional and lifestyle differences. Taking North China as an example, we utilized data from five representative cities (Beijing, Shijiazhuang, Jinan, Taiyuan, and Tianjin) for analysis and model building. The detailed summaries of the active user numbers are also listed in Table 1. The users enrolled in the study were chosen from 19 cities of Central, Southern, and Northern China and seven South-Central European countries to sufficiently reveal the regional disparity.

3.2. Analysis Result in China. The consecutive 3-year physiological anomaly rate curves in Wuhan together with the predicted physiological anomaly rate curves with and without the supervision of the officially reported COVID-19 infection rate in 2020 are illustrated in Figure 4. They are aligned by the time of the Chinese Spring Festival in the temporal axis. In the figure, all five curves peak around the time of Chinese Spring Festival. In addition, the predicted physiological anomaly rate with the supervision of official data in 2020 fits well with the rate calculated by the anomaly detection algorithm, which validates the prediction performance of the CDNet. Additionally, the physiological anomaly rate curve excluding COVID-19 in 2020 overlaps with both the predicted and the detected physiological anomaly rate curves including COVID-19 in 2020 before the outbreak of COVID-19, which verifies the basic reliability of the model. After that, all these three curves rise rapidly, which indicates that the outbreak of influenza-like illness is occurring alongside COVID-19. The predicted outbreak period aligns with the real-life situation. In addition, we also predicted the physiological anomaly rate curve from 2018 to 2019 with the prediction model and found that the predicted curve fits well with the total anomaly rate curve during the 2 years. This may indicate that the obvious separation happening around the Chinese Spring Festival between the predicted anomaly rate curves with and without the supervision of the officially reported COVID-19 infection rate in 2020 results from the outbreak of COVID-19.

Figure 5 illustrates the predicted COVID-19 infection rate across five Chinese cities and the officially reported accumulating COVID-19 infection cases in Wuhan. In the figure, there is a clear outbreak period in the predicted infection rate curve for each city, which may correspond to that of the newly confirmed cases. Taking Wuhan as an example, the predicted infection rate peaks around January 28, while the officially reported newly confirmed infection rate in Wuhan reached its highest on February 8 (the data after February 12 in Wuhan is omitted since the COVID-19 diagnostic criteria changed on that day, which causes a sudden sharp increase of 13,436 newly confirmed cases). The predicted disease peak is ahead of the officially reported peak by 11 days. The predicted earlier peak may indicate that health surveillance involving wearable sensors can play an

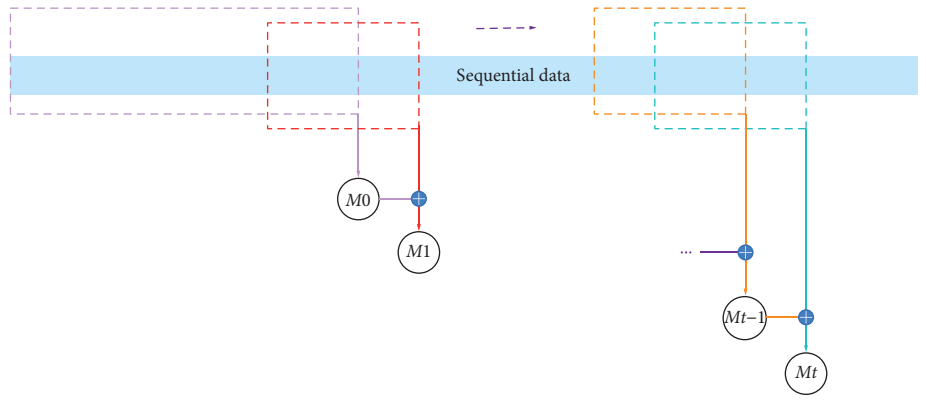

Figure 3: Work-flow of the online prediction framework for COVID-19 infection trend.

Table 1: Number of users enrolled in the study.

Region	Cities or countries	User number	Cities or countries	User number
North China	Beijing	126,575	Jinan	42,569
	Shijiazhuang	33,257	Taiyuan	16,753
	Tianjin	49,237		
Central China	Shanghai	153,711	Nanjing	76,204
	Hangzhou	78,840	Chengdu	64,436
	Wuhan	44,529	Hefei	33,257
	Nanning	19,641	Huanggang	3,412
	Xiaogan	2,358		
South China	Guangzhou	92,219	Shenzhen	71,669
	Foshan	20,229	Dongguan	29,146
	Fuzhou	23,061		
South-central Europe	Italy	68,494	Portugal	18,343
	Spain	187,788	Switzerland	4,431
	Germany	65,941	Greece	12,830
	France	34,711		

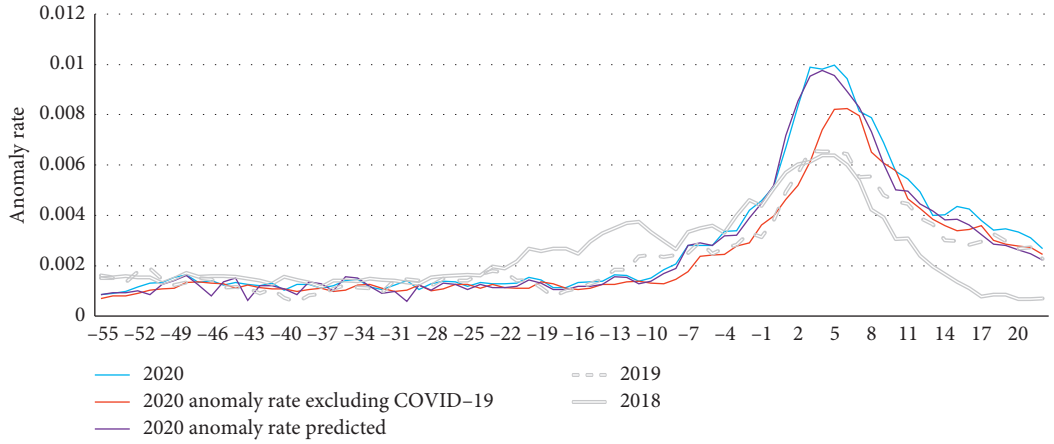

Figure 4: Physiological anomaly rate curves in Wuhan aligned by the time of Chinese Spring Festival. The two grey curves and the blue curve represent the physiological anomaly rate calculated by the anomaly detection algorithm from 2018 to 2020, respectively. The purple curve and the red curve represent the predicted physiological anomaly rate with and without the supervision of the official data, respectively.

important role in alerting to infectious disease outbreaks and in timely public health management. In fact, Wu and McGoogan also found there was a lag between the start of the illness and the diagnosis of COVID-19 by viral nucleic acid testing [2]. The newly infected cases actually peaked around January 28 if determined by the onset of the symptoms, which happens to be consistent with our findings. In addition, Figure 5 also shows that the predicted infection rate in Wuhan gradually decreases following January 28 and reaches a local minimum on February 1, which may

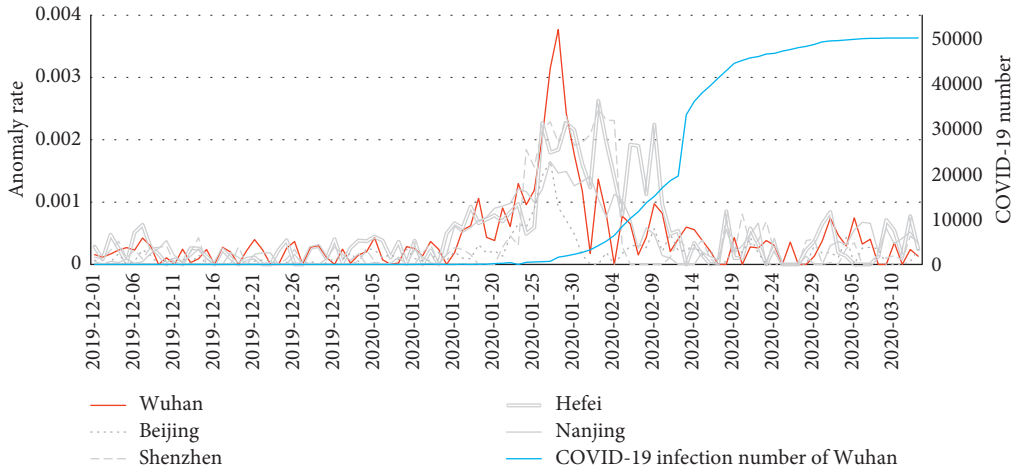

Figure 5: Predicted COVID-19 anomaly rate of 5 representative Chinese cities and the officially reported accumulating COVID-19 infection number of Wuhan.

correspond to the plateau in the officially reported accumulating infection curve that occurs after February 19. This result may indicate that the model can also predict the disease control outcome in advance. Moreover, Figure 5 shows Wuhan has the highest prediction disease peak among the five cities. This is also consistent with the fact that Wuhan is the most affected city in China.

3.3. Analysis Result in Italy and Spain. Figures 6(a) and 6(b) illustrate the predicted COVID-19 infection rate and the officially reported accumulating COVID-19 infection rate in Italy and Spain, respectively. The predicted infection rate in Italy rises rapidly from February 23, 2020, which coincides with the outbreak of COVID-19 in this country. As for Spain, the predicted infection rate starts to increase from February 29, which is 6 days later than Italy, and the predicted rate increases quickly following that. This is consistent with the real-life situation where the outbreak of COVID-19 was later in Spain.

As shown in Figure 6, the principal peak in the predicted COVID-19 infection curve of either Italy or Spain arrives as of April 8. In correspondence to the largest number of newly confirmed infection cases, which are reported officially by Italy on March 21 and Spain on March 25, the predicted principal peaks for the two countries occur around the time of March 13 and March 18, respectively. Both predicted principal peaks are ahead of the officially reported data by at least 1 week.

3.4. Correlation Analysis. To evaluate the appropriateness of predicting COVID-19 infection rate from physiological anomaly rate, we chose 19 Chinese cities to calculate the correlation between the officially reported COVID-19 infection rate and the detected physiological anomaly rate using Pearson's correlation coefficient shown in equation (3). In the equation, t_0 represents the start of the COVID-19 outbreak, t_1 stands for the end of the study period, and X, Y represent the officially reported COVID-19 infection rate and the physiological anomaly rate, respectively. The correlation analysis is performed in two steps. In the first step, we find the point, corresponding to the outbreak peak point of the officially reported COVID-19 infection curve, on the physiological anomaly rate curve. In the second step, we align the curves by the two points, and calculate the correlation coefficient.

$$\rho_{X,Y} = \frac{\sum_{t=t0}^{t1}(X_t - \overline{X})(Y_t - \overline{Y})}{\sqrt{\sum_{t=t0}^{t1}(X_t - \overline{X})^2}\sqrt{\sum_{t=t0}^{t1}(Y_t - \overline{Y})^2}}. \quad (3)$$

Pearson's correlation coefficients, ρ, for different cities in China are listed in Table 2. The average ρ value reaches around 0.68, which is strong correlation that further supports the opinion that physiological signs are useful for public health emergency alert. However, some cities do not show strong correlation, which may be due to the following reasons. Firstly, the officially reported cases of infection in some cities, e.g., Wuhan, were adjusted on certain days resulting in sudden changes. Secondly, the number of active users in some cities, e.g., Nanning, are relatively small which influences the performance of the model; therefore, the ρ value can be further improved when the number of active users increases. Finally, some cities, e.g., Beijing, have unstable user population and data noise due to the population shift.

3.5. Retention Effect. In the above correlation analysis, it is noticeable that there might be some retention effect in the detected physiological anomaly rate. To be specific, some people with anomalous measurements may continue to wear their devices so that they are calculated as anomalies on multiple days. This results in statistical error during the correlation analysis.

In order to analyze the impact, we calculate the retention rates of people detected as anomalies for several consecutive days. As shown in Figure 7, if a person is detected as anomaly on a certain day, the possibility of wearing the device is decreasing gradually from 3.5% down to 0.2% in the following 4 days. This indicates that the retention effect may have very limited influence on the correlation analysis.

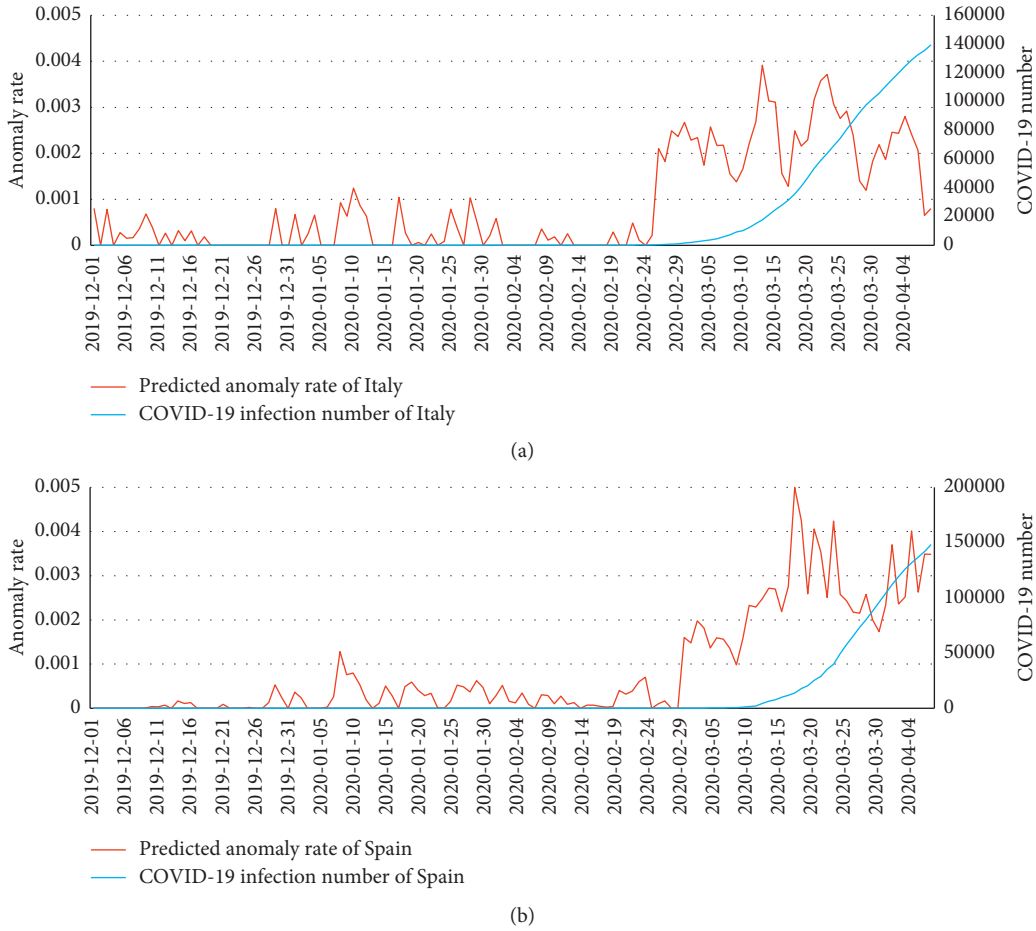

Figure 6: Predicted COVID-19 anomaly rate and the officially reported accumulating COVID-19 infection number of (a) Italy and (b) Spain.

Table 2: Correlation coefficient between the physiological anomaly rate and the officially reported number of newly confirmed COVID-19 cases.

City	ρ
Beijing	0.31
Shijiazhuang	0.58
Tianjin	0.53
Jinan	0.68
Taiyuan	0.55
Shanghai	0.51
Hangzhou	0.74
Wuhan	0.58
Nanning	0.52
Xiaogan	0.70
Nanjing	0.83
Chengdu	0.75
Hefei	0.84
Huanggang	0.87
Guangzhou	0.80
Foshan	0.81
Fuzhou	0.73
Shenzhen	0.82
Dongguan	0.75
Average	**0.68**

4. Discussion

In this study, a prediction model for COVID-19 epidemic trends has been realized using physiological data collected by wearable devices. The results show that prediction with dynamic physiological data may have an advantage in alerting to the infection outbreak in advance. However, the detection method for calculating the physiological anomaly rate has some limitations.

Firstly, on holidays, e.g., Chinese Spring Festival, Christmas, etc., transportation and population shift, social activities, and alcohol drinking might greatly influence the physiological signs of the users. For example, the elevated RHR due to heavy drinking on holidays might persist for several days and greatly influences the physiological anomaly rate to be detected. Especially for China, the outbreak of COVID-19 and influenza-like illness overlap with the Chinese Spring Festival. Thus, it is necessary to distinguish the elevated RHR cases induced by holiday activities from infection.

Secondly, the anomaly rate is the statistical description of wearable device users' physiological signs measured in the anomalous range. The validity of the statistical description depends on both the user scale and diversity. For example,

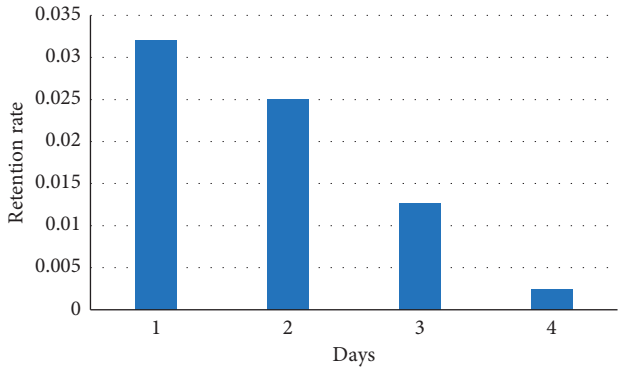

Figure 7: The retention rate across the timeline.

for a city with 0.1% officially reported infection rate of COVID-19, if the number of active users in the city is less than 10,000, there might be only 10 people infected among them. Such scale of data cannot support a convincing inference. Regarding the diversity, the prediction accuracy can be greatly improved if the distribution of active users is consistent with the natural distribution. For example, since elderly people and people with other diseases, e.g. cardiovascular disease (CVD), are more susceptible to COVID-19 [2, 3], the statistical performance of the model will be influenced if there is not enough coverage of such people.

Thirdly, although the current study provides a population evolution model for public health surveillance, it may be more meaningful for medical workers as well as individuals to take early precautions, if individualized health status prediction model is available. In the future, such prediction models based on wearable device data will be explored by incorporating more individual features, such as age, gender, body mass index (BMI), etc.

5. Conclusions

Public health emergencies can cause severe damage to the health and prosperity of our society. The popularity of wearable devices provides the opportunity for researchers to utilize big health data for public health emergency surveillance. In this study, a COVID-19 prediction framework using the health data from wearable devices was put forward. The proposed model could predict the epidemic trend of COVID-19 outbreak in various countries and cities. The results from the study may shed light on a nationwide solution for the infectious disease surveillance system.

Data Availability

The concerned sensor data cannot be shared due to user privacy. For academic purposes, anonymised region-level statistics can be shared under agreement.

Conflicts of Interest

The authors declare that there are no conflicts of interest regarding the publication of this paper.

Acknowledgments

This work was supported by the Huami Corporation.

References

[1] World Health Organization, *Coronavirus Disease (COVID-2019) Situation Reports*, World Health Organization, Geneva, Switzerland, 2020, https://www.who.int/emergencies/diseases/noveloronavirus-situation-eports.

[2] Z. Wu and J. M. McGoogan, "Characteristics of and important lessons from the coronavirus disease 2019 (COVID-19) outbreak in China," *JAMA*, vol. 323, 2020.

[3] W. Guan, Z. Ni, Y. Hu et al., "Clinical characteristics of 2019 novel coronavirus infection in China," *New England Journal of Medicine*, vol. 395, pp. 1708–1720, 2020.

[4] J. F.-W. Chan, S. Yuan, K.-H. Kok et al., "A familial cluster of pneumonia associated with the 2019 novel coronavirus indicating person-to-person transmission: a study of a family cluster," *The Lancet*, vol. 395, no. 10223, pp. 514–523, 2020.

[5] N. Chen, M. Zhou, X. Dong et al., "Epidemiological and clinical characteristics of 99 cases of 2019 novel coronavirus pneumonia in Wuhan, China: a descriptive study," *The Lancet*, vol. 395, no. 10223, pp. 507–513, 2020.

[6] Z. Yang, Z. Zeng, K. Wang et al., "Modified SEIR and AI prediction of the epidemics trend of COVID-19 in China under public health interventions," *Journal of Thoracic Disease*, vol. 12, 2020.

[7] J. T. Wu, K. Leung, and G. M. Leung, "Nowcasting and forecasting the potential domestic and international spread of the 2019-nCoV outbreak originating in Wuhan, China: a modelling study," *The Lancet*, vol. 395, no. 10225, pp. 689–697, 2020.

[8] S. Zhao, Q. Lin, J. Ran et al., "Preliminary estimation of the basic reproduction number of novel coronavirus (2019-nCoV) in China, from 2019 to 2020: a data-driven analysis in the early phase of the outbreak," *International Journal of Infectious Diseases*, vol. 92, pp. 214–217, 2020.

[9] J. Ginsberg, M. H. Mohebbi, R. S. Patel, L. Brammer, M. S. Smolinski, and L. Brilliant, "Detecting influenza epidemics using search engine query data," *Nature*, vol. 457, no. 7232, pp. 1012–1014, 2009.

[10] New York Times, *Google Uses Searches to Track flu's Spread*, New York Times, New York, NY, USA, 2008, http://www.nytimes.com/2008/11/12/technology/internet/12flu.html?_rp1.

[11] A. F. Dugas, Y.-H. Hsieh, S. R. Levin et al., "Google Flu Trends: correlation with emergency department influenza rates and crowding metrics," *Clinical Infectious Diseases*, vol. 54, no. 4, pp. 463–469, 2012.

[12] M. J. Paul, M. Dredze, and D. Broniatowski, "Twitter improves influenza forecasting," *PLoS Currents*, vol. 6, 2014.

[13] Q. Yuan, E. O. Nsoesie, B. Lv et al., "Monitoring influenza epidemics in China with search query from Baidu," *PLoS One*, vol. 8, no. 5, 2013.

[14] K. Liu, T. Wang, Z. Yang et al., "Using Baidu search index to predict Dengue outbreak in China," *Scientific Reports*, vol. 638040 pages, 2016.

[15] M. Santillana, A. T. Nguyen, M. Dredze et al., "Combining search, social media, and traditional data sources to improve influenza surveillance," *PLoS Computational Biology*, vol. 11, no. 10, 2015.

[16] Q. Xu, Y. R. Gel, L. L. R. Ramirez et al., "Forecasting influenza in Hong Kong with Google search queries and statistical model fusion," *PLoS One*, vol. 12, no. 5, 2017.

[17] D. Lazer, R. Kennedy, G. King, and A. Vespignani, "The parable of Google Flu: traps in big data analysis," *Science*, vol. 343, no. 6176, pp. 1203–1205, 2014.
[18] Y. Liu, H. Wang, W. Zhao, M. Zhang, H. Qin, and Y. Xie, "Flexible, stretchable sensors for wearable health monitoring: sensing mechanisms, materials, fabrication strategies and features," *Sensors*, vol. 18, no. 2, p. 645, 2018.
[19] D. Dias and J. Paulo Silva Cunha, "Wearable health devices-vital sign monitoring, systems and technologies," *Sensors*, vol. 18, no. 8, p. 2414, 2018.
[20] T. Arakawa, "Recent research and developing trends of wearable sensors for detecting blood pressure," *Sensors*, vol. 18, no. 9, p. 2772, 2018.
[21] J. M. Radin, N. E. Wineinger, E. J. Topol, and S. R. Steinhubl, "Harnessing wearable device data to improve state-level real-time surveillance of influenza-like illness in the USA: a population-based study," *The Lancet Digital Health*, vol. 2, no. 2, pp. e85–e93, 2020.
[22] X. Li, J. Dunn, D. Salins et al., "Digital Health: tracking physiomes and activity using wearable biosensors reveals useful health-related information," *PLoS Biology*, vol. 15, no. 1, Article ID e2001402, 2017.
[23] C. Huang, Y. Wang, X. Li et al., "Clinical features of patients infected with 2019 novel coronavirus in Wuhan, China," *The Lancet*, vol. 395, no. 10223, pp. 497–506, 2020.
[24] D. Wang, B. Hu, C. Hu et al., "Clinical characteristics of 138 hospitalized patients with 2019 novel coronavirus–infected pneumonia in Wuhan, China," *JAMA*, vol. 323, 2020.
[25] Z. Xu, L. Shi, Y. Wang et al., "Pathological findings of COVID-19 associated with acute respiratory distress syndrome," *The Lancet Respiratory Medicine*, vol. 8, 2020.
[26] J. Karjalainen and M. Viitasalo, "Fever and cardiac rhythm," *Archives of Internal Medicine*, vol. 146, no. 6, pp. 1169–1171, 1986.
[27] L. Faust, K. Feldman, S. M. Mattingly et al., "Deviations from normal bedtimes are associated with short-term increases in resting heart rate," *Npj Digital Medicine*, vol. 3, no. 1, pp. 1–9, 2020.
[28] G. Ke, Z. Xu, J. Zhang et al., "DeepGBM: a deep learning framework distilled by GBDT for online prediction Tasks," in *Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*, pp. 384–394, Anchorage, AK, USA, August 2019.