

Methodology

COVID-19 Hotspot Trend Prediction Using Hybrid Cellular Automata in India

Pokkuluri Kiran Sree^{1*} , Smt. S. S. S. N Usha Devi. N² 

¹Department of Computer Science & Engineering, Shri Vishnu Engineering College for Women, Bhimavaram, West Godavari, Andhra Pradesh, India

²Department of Computer Science & Engineering, University College of Engineering, Kakinada, East Godavari, Andhra Pradesh, India
Email: drkiransree@gmail.com

Received: 25 August 2020; **Revised:** 20 October 2020; **Accepted:** 26 October 2020

Abstract: The coronavirus disease 2019 (COVID-19) is an infectious disease identified at Wuhan, China, in December 2019 caused by new Coronavirus. The Indian government has taken many initiatives to mitigate the effect of COVID by encouraging the standard mechanisms of social distancing, the use of masks, and various safety parameters. COVID-19 hotspot identifies regions in India where COVID-19 severity is very high. We propose a novel hybrid cellular automata classifier for predicting the trend of various Hotspots in India, processing different parameters including infection control, virus reproduction rate, critical correlation, safety parameters, and social distancing. The proposed classifier was named Hybrid Cellular Automata-Hotspot (HCA-HS), predicts the number of hotspots in various districts of states, and also gives the status of each city marked either as Totally Safe or Marginally Safe or Unsafe. This will alert the state authorities to take necessary action to mitigate the COVID effect and help the people for possibly refraining from going to the infected areas, i.e., hotspots. The data sets were collected from Kaggle and the local Indian database for more adaptability. The accuracy of the predictions of Hotspots is reported as 91.58%, which is considerable at this moment. The developed classifier is compared with Support Vector Mechanism (SVM), K-Means, Decision Tree, and HCA-HS has reported an accuracy of 10.69% higher than the existing literature.

Keywords: COVID-19, hybrid cellular automata, infection control, reproduction rate, social distancing, usage of masks, safety parameters

1. Introduction

The rate change of the infections happening because of COVID-19 are increasing for the last six months. The death, infected, and recovery rate of COVID-19 is changing dynamically with time. Many researchers have focused on predicting the trends of COVID-19 using many mechanisms such as machine learning, data analytics, and image processing. Since this is a new and dynamic, open problem, none of the developed mechanism can provide a reliable, accurate, adaptable and robust method. There is a lot of scope for a new mechanism that can process diverse parameters and predict the variations of COVID-19 with reasonable accuracy.

We have studied various measures to reduce the effect of COVID-19 on India in various states. Dr. Kiran et al. [1] have proposed a novel Cellular Automata Classifier for predicting the trends of COVID19 [2]. Arab et al. [3] have

reported the hotspot variations in Iran, and this research helped plenty of travellers. Eckerle et al. [4] have analyzed the impact of hotspots. Cohen et al. [5] worked on collecting image data pertaining to COVID-19 so that the X-rays can be the benchmark for comparison. Many machine learning algorithms/mechanisms are evolved to address the problem of COVID-19 [6]-[8]. Several algorithms are proposed to process the X-rays of infected COVID patents to monitor the practice of social distancing and use of masks [9]-[13].

2. Design of HCA-HS

Cellular Automata (CA) is a set of cells on a given Grid. CA is an innate versatile classifier where the working of the CA is shown in Figure 1.

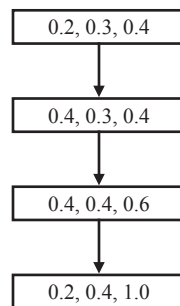


Figure 1. Design of HCA-HS with an example
Rule Applied $\langle Ti + 1 + Ti, Ti + 1, Ti + Ti + 1 \rangle$ (Basin: 0.2, 0.3, 1.0)

Cellular Automata is employed in several terms of rules, where these rules tell the system about the dependability of the corresponding states. For example, Rule 1 depends on its state and its neighbor.

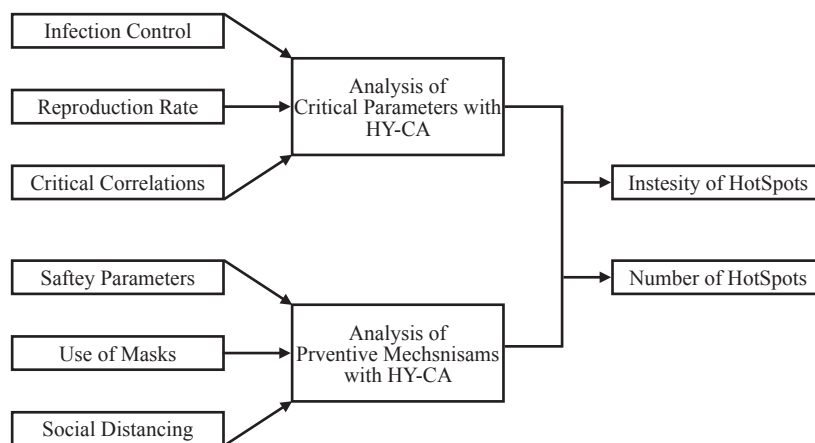


Figure 2. Overall Design of HCA-HS with Output

Rule 2 states it has to depend on the next state, and Rule 3 states it has to depend on its state and its neighbor. The transitions are represented as per the rules to get 0.2, 0.4, as a basin.

We have designed the entire research into two modules, as shown in Figure 2. The first module is trained and tested to analyze critical parameters like infection control, reproduction rate, critical correlation like (learning from

the mistakes) to give the intensity of the hotspots of various areas. The second module is trained and tested to analyze preventive parameters like the implementation of safety parameters, usage of masks, and social distancing. The output that is expected is bifurcated into totally safe, marginally safe, and unsafe in relation to the number of hotspots.

Table 1. Rules used in HY-CA

	111	110	101	100	011	010	001	000
250	0	1	1	0	1	0	1	0
248	1	1	1	1	1	1	0	0
228	0	1	1	0	1	1	1	0
182	1	1	0	0	1	1	0	0
155	1	1	1	0	0	0	0	0
160	1	0	1	0	1	0	1	0
120	1	1	1	1	1	0	1	0
0	0	0	0	0	0	0	0	0

Algorithm 1

Input: Training Sequence $S_n = \{S_1, S_2, \dots, S_p, \dots, S_l\}$ with m classes.

Output: HCA-HS tree.

Partition (S, m)

Step 1: Generate an HCA-HS with n attractor states.

Step 2: Allocate the sequence S into these n attractor states.

Step 3: Verify all the elements in the attractor basins.

Step 4: If all the elements indicate S_1 belongs to an attractor basin, then label the sequence with the class associated with the attractor basin.

Step 5: If S_1 belongs to more than one class, then Partition (S_1, m) till S_1 belong to an attractor basin.

Step 6: Stop.

The rules of form referred in Table 1. We have taken six different rules and their operations in this prediction. The simple algorithms shows the method of prediction.

3. Results and comparisons

Datasets were collected from Kaggle [2] and the local Indian database for more adaptability. We have extracted 2,02,689 various records for developing a robust, reliable classifier.

The research work has considered all the major cities in India with a population of more than 500000 for the analysis, but the analysis of the critical parameters was reported only for the cities Mumbai (MUM), New Delhi (ND), Hyderabad (HYD), Kakinada (KKD), Vijaywada (VIJ) and Visakhapatnam (VSP). As the results obtained in Figure 3, Vishakapatnam shows a better infection control, considerable reproduction rate, and proper critical correlation parameters.

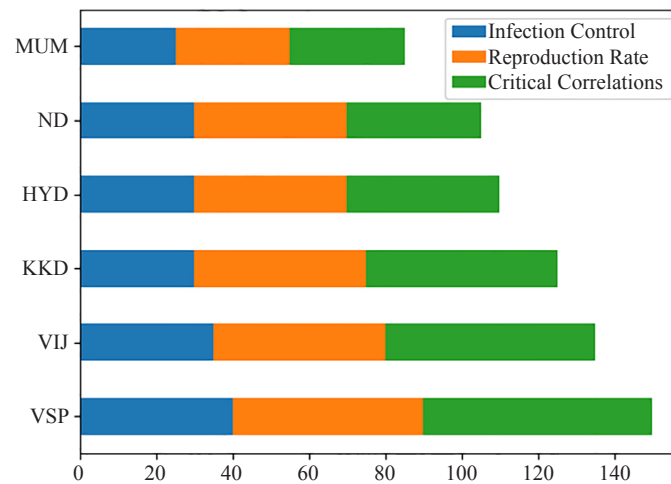


Figure 3. Analysis of Critical Parameter as on 20-07-2020

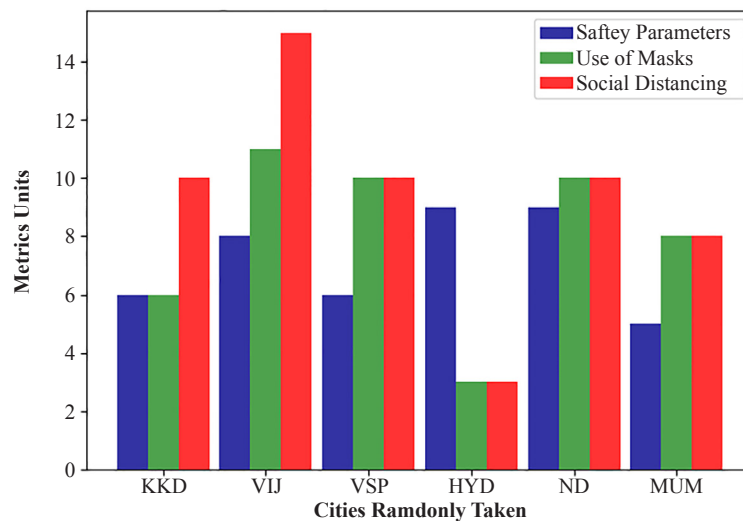


Figure 4. Analysis of Preventive Mechanisms as on 20-07-2020

Figure 4 shows that Vijaywada (VIJ) adopts good social distancing norms and stands first in implementing the usage of masks. Hyderabad and New Delhi have reported considerably better safety parameters.

HCA-HS predicts the number of hot spots in various states of India using the two modules designed. Figure 5 shows that Tamil Nadu (TN) reports the highest number of hotspots and Orrissa (OD) reports very few. The other states reported in the graph are Kerala, Madhya Pradesh (MP), Maharashtra (MH), Pondicherry (PU), Rajasthan (RAJ), Telangana (TE), Uttar Pradesh (UP) and West Bengal (WB).

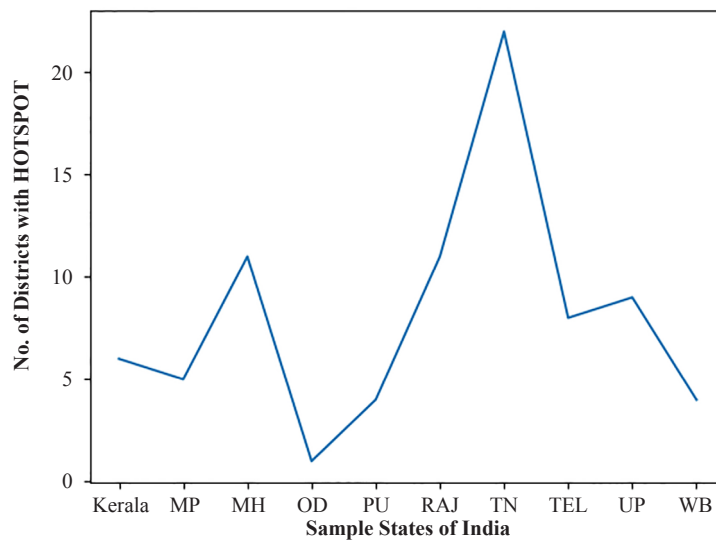


Figure 5. Number of Districts with Hot Spots in few States of India as on 20-07-2020

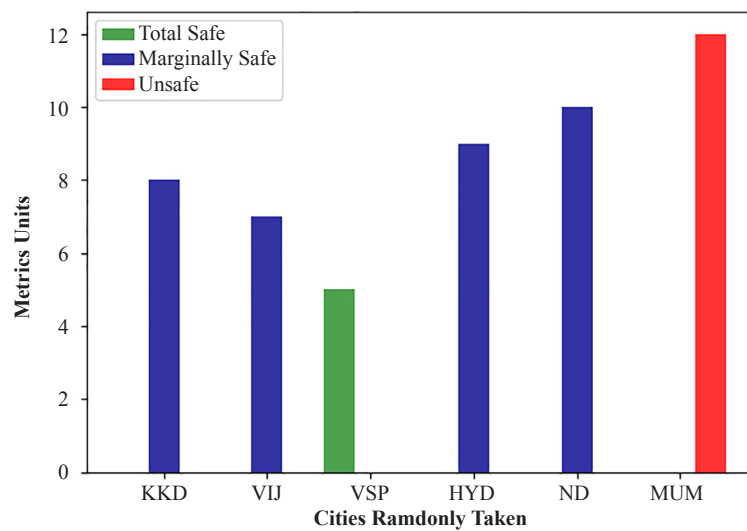


Figure 6. Intensity of the Hotspots in India as on 20-07-2020

Figure 6 shows that Vishakapatnam(VSP) is considerably doing better, which was designated as Totally Marginally Safe, and Mubai is designated as Unsafe at this moment.

We have compared the performance of our classifier HCA-HS with other standard mechanisms SVM, K-Means, and Decision Tree algorithm, and HCA-HS has performed considerably well compared to the other existing literature. The performance of the developed classifier is reported as 91.58%, which is 10% more accurate when compared with standard models. The result depicted is, as per the data collected at that point, can be considered only to conduct research and do not contain any legal validity.

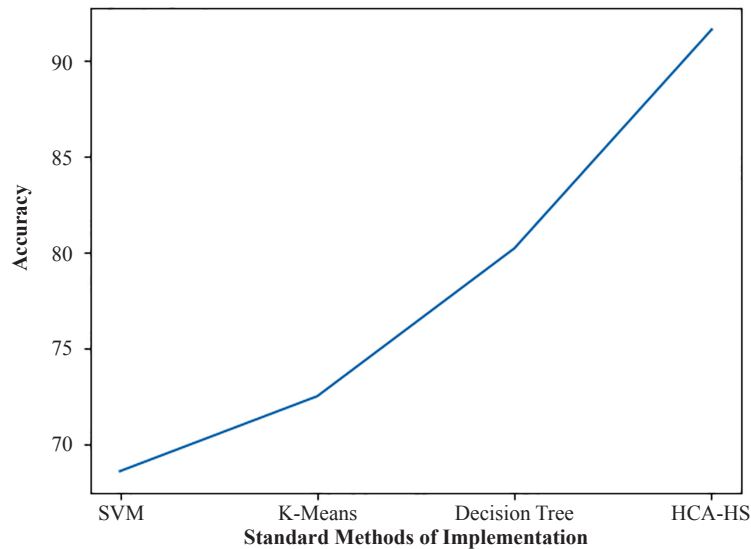


Figure 7. Comparison of the Performance of HCA-HS with SVM, K-M, and DT

4. Conclusion

We have successfully developed a novel, unique, robust, adaptable, and most accurate classifier HCA-HS to predict the number of COVID-19 hotspots in India. This work can be used by the administrative authorities of the government to keep track and monitor the growing number of hotspots in India. The public and travelers will also utilize this work for updating their COVID knowledge and act accordingly. HCA-HS processes live data and report dynamic changes. The accuracy reported is 91.58%, which is the best in the existing literature.

Conflict of interest

The authors declare that there is no conflict of interest regarding the publication of this paper.

References

- [1] K. S. Pokkuluri and S. U. Devi Nedunuri, "A novel cellular automata classifier for COVID-19 prediction," *Journal of Health Sciences*, vol. 10, no. 1, pp. 34-38, 2020.
- [2] S. Rajkumar, "COVID-19 in India," *kaggle.com*, 2020. [Online]. Available: <https://www.kaggle.com/sudalairajkumar/covid19-in-india>. [Accessed October 2, 2020].
- [3] Z. A. Mazar, R. Sah, A. A. Rabaan, K. Dhama and A. J. Rodriguez-Morales, "Mapping the incidence of the COVID-19 hotspot in Iran-implications for travellers," *Travel Medicine and Infectious Disease*, vol. 34, pp. 101630, 2020.
- [4] I. Eckerle and B. Meyer, "SARS-CoV-2 seroprevalence in COVID-19 hotspots," *The Lancet*, vol. 396, no. 10250, pp. 514-515, 2020.
- [5] J. P. Cohen, P. Morrison, L. Dao, K. Roth, T.Q. Duong and M. Ghassemi, "Covid-19 image data collection: Prospective predictions are the future," *ArXiv preprint*, 2020.
- [6] P. Mehta, D. F. McAuley, M. Brown., E. Sanchez, R. S. Tattersall, J. J. Manson and HLH Across Speciality Collaboration, "COVID-19: Consider cytokine storm syndromes and immunosuppression," *The Lancet*, vol. 395, no. 10229, pp. 1033-1034, 2020.
- [7] G. S. Randhawa, M. P. Soltysiak, H. El Roz, C. P. de Souza, K. A. Hill and L. Kari, "Machine learning using intrinsic genomic signatures for rapid classification of novel pathogens: COVID-19 case study," *Plos One*, 2020.

Available: <https://doi.org/10.1371/journal.pone.0232391>. [Accessed April 24, 2020].

- [8] A. Alimadadi, S. Aryal, I. Manandhar, P. B. Munroe, B. Joe and X. Cheng, "Artificial intelligence and machine learning to fight COVID-19," *Physiological Genomics*, vol. 52, no. 4, pp. 200-202, 2020.
- [9] M. Barstugan, U. Ozkaya and S. Ozturk, "Coronavirus (covid-19) classification using ct images by machine learning methods," ArXiv preprint, 2020.
- [10] M. A. Elaziz., K. M. Hosny, A. Salah, M. M. Darwish, S. Lu and A. T. Sahlol, "New machine learning method for image-based diagnosis of COVID-19," *Plos One*, vol. 15, no. 6, June, 2020. [Online serial]. Available: <https://journals.plos.org/plosone/article?id=10.1371%2Fjournal.pone.0235187>. [Accessed October 6, 2020].
- [11] S. F. Ardabili, A. Mosavi, P. Ghamisi, F. Ferdinand, A. R. Varkonyi-Koczy, U. Reuter and P. M. Atkinson, "Covid-19 outbreak prediction with machine learning," *MedRxiv*, 2020. Available: <https://doi.org/10.1101/2020.04.17.20070094>.
- [12] S. Tuli, S. Tuli, R. Tuli and S. S. Gill, "Predicting the growth and trend of COVID-19 pandemic using machine learning and cloud computing," *Internet of Things*, vol. 1, pp. 100222, 2020.
- [13] P. K. Sree, "Automatic social distancing mechanism for preventing covid-19 using deep learning," *International Journal of Recent Development in Computer Technology & Software Applications*, vol. 4, no. 1, pp. 2581-6276, 2020.