

# Artificial Intelligence in the fight against Covid-19

Rushi Shah  
Student

Shah and Anchor Kutchhi Engineering  
College  
Mumbai, India  
rushishah@sakec.ac.in

Srikant Kodeboyina  
Founder & CEO  
Blue Eye Soft Corp  
Greer, SC 29650, United States  
sri@blueyesoft.com

Vatsal Gada  
Student

Shah and Anchor Kutchhi Engineering  
College  
Mumbai, India  
vatsal.gada@sakec.ac.in

Ramesh Menon  
Advisor  
Blue Eye Soft Corp  
CTO-S&T, IBM USA  
Greer, SC 29650, United States  
rmenon@us.ibm.com

Dr. Nilakshi Jain  
Associate Professor  
Shah and Anchor Kutchhi  
Engineering College  
Mumbai, India  
nilakshi.jain@sakec.ac.in

**Abstract** — As of late July 2020, COVID-19 has been creating havoc for more than seven months and has infected more than 16 million people and almost 650K fatalities. The IMF has predicted that the global economy will shrink by almost 3 percent in 2020. Over the past few years with the increase of available data and drastically more compute power, Artificial Intelligence (AI) has been helping and improving in all fields. The current COVID-19 pandemic poses a number of challenges. Researchers and scientists are using Artificial Intelligence (AI) to combat those problems that previously weren't possible using traditional technology and manpower. This paper presents an overview on a number of Artificial Intelligence (AI) applications in various fields. We touch on fields ranging from tracking, prediction, diagnosis, natural language processing and cure.

**Keywords**—Artificial Intelligence, Machine Learning, Deep Learning, COVID-19

## I. INTRODUCTION

As the world grapples with the current COVID-19 pandemic, researchers and scientists from all around the world are looking for ways to tackle problems in various fields. One of the most powerful tools in their arsenal is the use of Artificial Intelligence (AI). With the right use of Artificial Intelligence (AI), researchers and scientists can change the course of pandemic and how we fight against the virus. In this paper we look at the problems that are preventing us from using this technology to the fullest, areas where we are using Artificial Intelligence (AI) and achieving great results and the methods that are used to do so. The following sections provide a brief overview on difficulties faced while using Artificial Intelligence (AI) in tracking and predictions, using Artificial Intelligence (AI) medical image processing to diagnose the disease, using text mining and Natural language processing (NLP) to gather important data from the internet regarding COVID-19 and using Artificial Intelligence (AI) to repurpose old drugs and finding new ones to find a cure for the virus.

The paper is organized as follows:

The use of Artificial Intelligence (AI) in tracking and problems faced while doing so are discussed in section two. Section three and four discuss the detection and diagnosis of COVID-19 using Computed Tomography (CT) scans and Natural language processing (NLP) for news feature extraction. The fifth section covers the two possible ways in

which Artificial Intelligence (AI) is enabling researchers to find a possible cure for the virus.

## II. TRACKING

Artificial Intelligence (AI) can in theory help with tracking and prediction of how the virus will spread over time, but that has not been the case so far. Artificial Intelligence (AI) has not been useful in this field and there are a number of reasons contributing to this.

The first and the biggest by far is that data on COVID-19, that is required to train Artificial Intelligence (AI). 2015 Zika virus is one such example, where a dynamic neural network was trained to guess the spread. Since Zika is different from COVID-19 there is not enough data yet to develop Artificial Intelligence (AI) models at the time of writing, its prevalence can be tracked and estimated. The number of publications reporting on the use of Artificial Intelligence (AI) for diagnostic purposes has so far used small and perhaps biased data.[1] But a lot of organisations are taking initiatives to combat this situation by gathering new data and sharing existing data to train and help new Artificial Intelligence (AI) models. The World Health Organisation (WHO) has their own initiatives and also links to many other such initiatives. One of the notable ones being Global Initiative on Sharing All Influenza Data (GISAID) initiative. Another joint initiative between A12, Semantic Scholar, and other top tier silicon valley companies is working to make openly accessible COVID-19 Open Research Dataset (CORD-19). Kaggle, an online community of data scientists and machine learning practitioners, had organised a COVID-19 Open Research Dataset Challenge, which would help incentivise COVID-19 data research. Elsevier made their peer-reviewed research on COVID-19 and thousands of articles as well as mining techniques public. Lens has made all its patent information available online and free to use. And Chen et al. published CatchID-19 Twitter's first public data.[1]

Secondly the issue with using Artificial Intelligence (AI) in tracking and prediction of the spread of virus is not only the lack of training data but the inherent problems that come with using big data, example of which could be data harvested from social media. One such example of a failed attempt to use Artificial Intelligence (AI) in prediction of infectious disease is Google Flu Trends. The idea being people being sick with flu are going to search about it, providing google with instantaneous data of flu prevalence.



But in the 2013 flu season, Google Flu Trends (GFT) failed spectacularly, missing the peak season by 140% making it the poster child of the foibles of big data.[2] This is not to say that big data isn't useful, when used correctly it can be one of the most powerful tools in today's age. Another one of the problems that can be faced while approaching Artificial Intelligence (AI) techniques is data overload. Since the start of the pandemic numerous studies and data emerge almost everyday and shifting through the right data is a real issue, although this could be a place for data analytic tools to showcase their power.

To say the least is because of the above mentioned hurdles, Artificial Intelligence (AI) forecasts for COVID-19 aren't accurate or reliable. Most models used today aren't actually using Artificial Intelligence (AI) models to forecast the spread, instead epidemiology models are preferred by most forecasters.

### III. DIAGNOSIS

At present, various respiratory diseases are classified by RT-PCR i.e Reverse Transcription Polymerase Chain Reaction detection technique [3]. These strategies that are being used are, however, a bit too costly and take away a lot of time, their true positive rate turns out to be low, and generally require a set of equipment and instruments. Additionally, there has been a scarcity of testing kits throughout the globe because of the limitation on budget and techniques. While COVID-19 cases can be identified based upon chest Computed Tomography (CT) scans by radiologists and clinical specialists, their tasks are manual and take up a lot of time, particularly when there are a large number of cases to be solved. In the current situation, it is very important to constrain the spread of the disease, generate data to train Artificial Intelligence (AI) models and save as many lives as possible and this can only be done if we can achieve quick diagnosis and also more accurate ones.

Can Artificial Intelligence (AI) be as precise as humans? Can it spare radiologists' time? Can it help in performing a quicker and cheaper diagnosis than with standard tests for COVID-19? A survey by Bullock et al. [3] on Artificial Intelligence (AI) applications against COVID-19 argues that this is possible. Various Artificial Intelligence (AI) methods, particularly deep learning has greatly helped in improving the diagnosis performance for the radiologists and doctors and this has been possible by analyzing medical imaging data. Detection of the infected COVID-19 case is possible by using X-ray images and Computed Tomography(CT) scans as the input to deep learning models.

COVID19 detection neural network (COVNet) is a three dimensional deep learning method that is introduced in the paper: Artificial intelligence in the battle against coronavirus (COVID-19): a survey and future research directions.T. T, Nguyen. (2020) [4]. This neural network uses chest Computed Tomography (CT) scans to detect COVID-19. The model is fed with three kinds of Computed Tomography (CT) scans, which include non pneumonia cases, COVID-19 cases and community acquired pneumonia(CAP) to check its robustness and is illustrated in Fig. 1. The area under the receiver operating characteristic curve (AUC) is used to evaluate the detection method. COVNet is a convolutional ResNet-50 model which predicts the class labels of the Computed Tomography(CT) images by taking a series of Computed Tomography(CT) images as input. The AUC

value obtained is at 0.96, and hence the proposed model shows great ability to detect COVID-19 cases.

It was discovered that COVID-19 patients show abnormalities in chest radiography images. Motivated by this, a deep convolutional neural network(CNN) model was designed [5]. This model was trained on an open source dataset with 13,975 images of 13,870 patients. Based upon this training, it was observed that the model could achieve 93.3% accuracy.

Along with the studies portrayed above, there are several other papers that apply deep learning for COVID-19 diagnosis with the use of radiology images. The summary for the same is available in Table 1 [4] for comparisons.

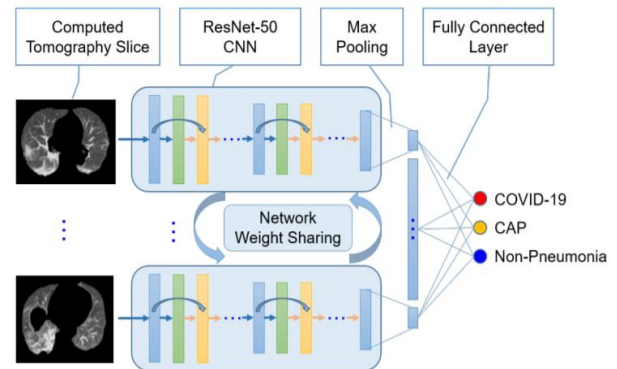


Fig. 1. Illustrative example of the COVNet model

Along with the studies portrayed above, there are several other papers that apply deep learning for COVID-19 diagnosis with the use of radiology images. The summary for the same is available in Table 1 [4] for comparisons.

### IV. NLP

Natural language processing (NLP) approaches can be utilized to answer key research questions related to COVID-19. These questions can be evaluated via various sources which include articles from the (CORD-19) which contains over 29,000 research articles about COVID-19, which is also the current largest open dataset available. Data from other sources, such as social media can be used to complete the textual data from scientific articles. For instance, COVID-19 TweetIDs and Covid19 Tweets are a few public Twitter datasets that have been released recently.

These datasets are maintained and updated based on coronavirus-related tweets. Natural language processing (NLP) strategies can be applied to create text mining tools using any of the sources mentioned above. This can prove of great help to the medical community to find answers to key scientific questions on COVID-19. Natural language processing (NLP) can also be employed for feature extraction from news such as control measures taken by governments to control this pandemic or disease prevention awareness of residents. The use of Natural language processing (NLP) and text mining tools can be done to analyse a multilanguage Twitter dataset to comprehend changing strategies and reactions of individuals to the COVID19 outbreak [4]. Since the outbreak of COVID-19 in late December 2019, governments of numerous countries have attempted to mitigate the spread of this virus by implementing various policies. Reactions of individuals may be gathered from social media platforms such as Twitter, related to the

pandemic and to the governmental policies. With the implementation of social distancing and other such strict policies throughout the nation, there has been a drastic change in people's lives and this can be seen and interpreted through people's reactions on various social media platforms. A detailed analysis report of this data is not only crucial for Government entities to navigate through this pandemic but also to frame any policies that can help with any such pandemics that may occur in the future. 3,000 COVID-19 related posts are classified into seven types of situational data using three machine learning methods which include support vector machine (SVM), naive Bayes and random forest. These posts were gathered from Sina Weibo, which is a Chinese social media platform. Identification of situational data helps the authorities to foresee its propagation scale, understand the situation better during the crisis and sense the mood of the public and thus, a proper response strategy can be created throughout the COVID-19 pandemic. China's hybrid Artificial Intelligence (AI) model extracts the relevant features of news information, examines the data related to COVID-19, sorts this data by date, province, and city, and filters out case reports and related foreign news. Feature extraction is only performed on the title and main content of each news text to obtain concise and robust features in practice. For each given news text in Chinese, a pretrained model of the BERT language model (RoBERTa) [6], which was designed by researchers at Facebook Artificial Intelligence (AI) and the University of Washington, is used to extract text features. This model combines the Chinese Whole Word Masking strategy, uses WordPiece segmentation to divide a complete word into several subwords, and also combines BERT. This model can achieve good feature extraction results with minimal training. The news titles and main content are separately obtained as the input to prevent overfitting and achieve efficient training, and the last hidden layer of the pretrained model is used to encode the text. Then, this model encodes 768-D title and 768-D text together to generate a 1536-D Natural language processing (NLP) feature vector, in which each vector corresponds to a piece of news. The dataset is divided into national and provincial datasets to achieve accurate daily predictions in different cities and regions and the entire country, where the national dataset contains news from all regions, and the provincial dataset contains news from each province. The news is classified by day to ensure the presence of at least one news per day, and the features of all news of the day are averaged as the Natural language processing (NLP) feature vector. Hence, Natural language processing (NLP) approaches specifically can be extremely helpful, given the immense amount of literary data that exists regarding COVID-19.

## V. CURE

Artificial Intelligence (AI) can serve two purposes in the medical field to find a cure. First is using Artificial Intelligence (AI) to find new medicines and vaccines and the other would be to use Artificial Intelligence (AI) for drug repurposing. It is not very likely that a cure for covid is going to be available in the near future, making a vaccine for such viruses require medical checks, trials and control to be approved and once approved, it needs to be screened. Then these vaccines will go in production. A rough estimate

provided by Relagado is that this process could take about 18 months. Hence a lot of researchers and scientists are looking into repurposing old drugs. Repurposing old drugs can cut a lot of cost as they omit a lot of initial stages and go directly to preclinical trials. Artificial Intelligence (AI) is used to predict how these drugs on their own or when used in combination of other techniques or drugs could affect covid virus. Artificial Intelligence (AI) can run through and process molecular structure at a rate no other team can work and hence using Artificial Intelligence (AI) in this field is essential. [1]

On the other hand Artificial Intelligence (AI) is also used in finding new medicines and vaccines. AlphaFold, a protein structure prediction system by Google Deepmind is able to predict protein structures associated with covid depending on their amino acid sequence. Google has also released these structures but further experimentation is needed on these structures. But the structures predicted can help with the development of vaccines or maybe even a deeper understanding of the virus itself. [7]

The Artificial Intelligence (AI) based method used in the chemical formulation to produce a novel chemical that can inhibit COVID-19. Cellular presentation to produce structures can be exploited using Machine learning models like manufacturing autoencoders and GANs which are then used through reinforcement learning methods. This is a work in progress as the authors have to sort and evaluate the molecules found.[8] However Artificial Intelligence (AI) methods are promising because they can exploit large amounts of drug-like chemical substances and automatically extract useful information from large-scale data. It is therefore capable of forming molecules without designing to study the relationship between molecular structures and its therapeutic properties. The proposed method is time-efficient and much less expensive than conventional methods and has the potential to produce a combination of novel drugs in COVID-19.

## VI. CONCLUSION

The COVID-19 pandemic has literally brought the world to a stand still. Artificial Intelligence (AI) innovations have been of great help to humans previously and will likely help researchers in this immense battle against COVID-19. This paper presents various applications of Artificial Intelligence (AI) ranging from data collection and prediction using ML models, detection and diagnosis using chest radiology images, use of Natural language processing (NLP) and text mining tools to understand public response and plan out better strategies, to data analysis for drug repurposing and discovery. There are relatively limited applications and contributions in this pandemic even though there have been several researches. This is majorly because of the limited amount of data resources available while Artificial Intelligence (AI) models typically require huge amounts of data to get trained well and acquire intelligence. But Artificial Intelligence (AI) does have a promising future. A set of problems related to COVID-19 along with the required types of data and potential Artificial Intelligence (AI) solutions that can be used to counter those problems are listed in Table 2.

TABLE I.

References	Paper Data	AI Methods	Result
[6]	4,356 chest CT exams from 3,322 patients from 6 medical centers: 1,296 exams for COVID-19, 1,735 for CAP and 1,325 for non pneumonia	A 3D convolutional ResNet-50, namely COVNet	AUC for detecting COVID-19 is of 0.96
[7]	170 X-ray images and 361 CT images of COVID-19 from 5 different sources	A new CNN and pre-trained AlexNet with transfer learning	Accuracy of 98% on X-ray images and 94.1% on CT images
[8]	1,341 normal, 1,345 viral pneumonia, and 190 COVID-19 chest X-ray images	AlexNet, ResNet-18, DenseNet201, SqueezeNet	Accuracy of 98.3%
[9]	Chest X-ray images of 50 normal and 50 COVID-19 patients	Pre-trained ResNet-50	Accuracy of 98%
[10]	618 CT samples: 219 from 110 COVID-19 patients, 224 CT samples from 224 patients with influenza-A viral pneumonia, and 175 CT samples from healthy people	Location-attention network and ResNet-18	Accuracy of 86.7%
[11]	1,065 CT images (325 COVID-19 and 740 viral pneumonia)	Modified inception transfer-learning model	Accuracy of 79.3% with specificity of 0.83 and sensitivity of 0.67
[12]	CT images obtained from 157 international patients (China and U.S.)	ResNet-50	AUC of 0.996
[13]	5,941 Posterior-anterior chest radiography images across 4 classes (normal: 1,583, bacterial pneumonia: 2,786, non-COVID-19 viral pneumonia: 1,504, and COVID-19: 68)	Drop-weights based Bayesian CNNs	Accuracy of 89.92%
[14]	Clinical data and a series of chest CT data collected at different times on 133 patients of which 54 patients progressed to severe/critical periods whilst the rest did not	Multilayer perceptron and LSTM	AUC of 0.954
[15]	970 CT volumes of 496 patients with confirmed COVID-19 and 1,385 negative cases	2D deep CNN	Accuracy of 94.98% and AUC of 97.91%
[16]	16,756 chest radiography images across 13,645 patient cases from two open access data repositories	A deep CNN, namely COVID-Net	Accuracy of 92.4%
[17]	CT images of 1,136 training cases (723 positives for COVID-19) from 5 hospitals	A combination of 3D UNet++ and ResNet-50	Sensitivity of 0.974 and specificity of 0.922

TABLE II.

Making drones and robots for disinfection, cleaning, obtaining patients' vital signs, distance treatment, and deliver medication	Simulation environments and demonstration data for training autonomous agents.	<ul style="list-style-type: none"> <li>- Safety must be guaranteed at the highest level.</li> <li>- Trust in autonomous systems.</li> <li>- Huge efforts from training agents to implementing them to real machines.</li> </ul>	Deep learning, computer vision, optimization and control, transfer learning, deep reinforcement learning, learning from demonstrations
Making chatbots to consult patients and communities, and combat misinformation (fake news) about COVID-19.	Medical expert guidelines and information.	<ul style="list-style-type: none"> <li>- Unable to deal with unsaved queries.</li> <li>- Require a large amount of data and information from medical experts.</li> <li>- Users are uncomfortable with chat-bots being machines.</li> <li>- Irregularities in language expression such as accents and mistakes.</li> </ul>	Deep learning and NLP tools, e.g. NLTK, ELMo, ULMFiT, Transformer, Google's BERT, Transformer-XL, XLNet, ERNIE, T5, BPT and OpenAI's GPT-2 and GPT-3
Track and predict economic recovery via, e.g. detection of solar panel installations, counting cars in parking lots.	Satellite images, GPS data (e.g. daily anonymized data from mobile phone users to count the number of commuters in cities).	<ul style="list-style-type: none"> <li>- Difficult to obtain satellite data in some regions.</li> <li>- Noise in satellite images.</li> <li>- Anonymized mobile phone data security.</li> </ul>	Deep learning, e.g. autoencoder models for feature extraction and dimensionality reduction, and CNN-based models for object detection

## VII. LIMITATION

Even though it seems that Artificial Intelligence (AI) is capable enough to be the only solution to major problems in the medical field, we're still far away from using AI as the only detection method for any condition. No authorities have a set regulation for the accuracy level of these models. It takes years of data and tuning to get these models to 100% accuracy and even then we have to look out for biases of various kinds. Hence, traditional practices and AI need to work hand in hand for the foreseeable future till algorithms and models can surpass accuracy of current systems.

## ACKNOWLEDGMENT

- Special thanks to my mentor Dr. Nilakshi Jain, Associate Professor, SAKEC, Mumbai for always encouraging me and constantly guiding me and giving her insightful inputs.
- Many thanks to Mr.Srikanth Kodeboyina, Founder & CEO, Blue Eye Soft Corp and Mr. Ramesh Menon, Advisor, Blue Eye Soft Corp for productive discussions and suggesting me to research in this domain.

## REFERENCES

- [1] Artificial intelligence vs COVID-19: limitations, constraints and pitfalls [published online ahead of print, 2020 Apr 28]. Naudé W. *AI Soc.* 2020;1-5. doi:10.1007/s00146-020-00978-0
- [2] Deep residual learning for image recognition. In IEEE Conference on Computer Vision and Pattern Recognition (pp. 770-778). Zhang, X., He, K., Sun, J., and Ren, S. (2016).
- [3] Mapping the Landscape of Artificial Intelligence Applications against COVID-19. Luccioni, & Bullock, Joseph & Pham, Katherine & Lam, Cynthia & Alexandra, & Luengo-Oroz, Miguel. (2020).
- [4] Artificial intelligence in the battle against coronavirus (COVID-19): a survey and future research directions. T. T. Nguyen. (2020).
- [5] Artificial Intelligence (AI) and Big Data for Coronavirus (COVID-19) Pandemic: A Survey on the State-of-the-Arts. 10.13140/RG.2.2.23518.38727. Huynh-The, Thien & C. Nguyen, Dinh & Hwang, won-Joo & Pham, Quoc-Viet & Pathirana, Pubudu. (2020).
- [6] Predicting COVID-19 in China Using Hybrid AI Model. IEEE Transactions on Cybernetics. PP. 1-14. 10.1109/TCYB.2020.2990162. Nanning & Du, Jianji & Zhang, Wenting & Kang, Tao & Lou, Bin & Chi, Zheng, He & Cui, Yuting & Long, Hong & Ma, Mei & Yuan, Qi & Zhang, Shupe & Zhang, Shaoyi & Wang, Dong & Ye, Zijian & Yang, Feng & Xin, Jingmin. (2020).
- [7] Application of Artificial Intelligence in COVID-19 drug repurposing. Diabetes & Metabolic Syndrome: Clinical Research & Reviews. 14. 10.1016/j.dsx.2020.06.068. Chandana & Swayamsiddha, Sweta & Harun Al Rashid, Mohanty, Mayank & Mohanty, Md & Mridul, Swati. (2020).
- [8] Artificial Intelligence (AI) applications for COVID-19 pandemic. Diabetes & Metabolic Syndrome: Clinical Research & Reviews. 14. 10.1016/j.dsx.2020.04.012.
- [9] Raju & Javaid, Vaishya, Mohd & Khan, Ibrahim & Haleem, Abid. (2020).