УДК: 004.665.6

МЕТОДЫ ГЛУБОКОГО ОБУЧЕНИЯ ДЛЯ ДИАГНОСТИКИ COVID-19: МОДЕЛИ, НАБОР ДАННЫХ И ПРОБЛЕМЫ

ХУДОЙКУЛОВ З.Т., ИСЛОМОВ Ш.З., ДАВРОНОВА Л.У., РУСТАМОВА С.Р.

Ташкентский университет информационных технологий имени Мухаммада ал-Хоразмий
(Ташкент, Узбекистан)

Аннотация. Эта статья посвящена методам обнаружение covid-19 на основе компьютерной томографии и рентгеновских снимков с использованием методов глубокого обучения. Кроме того, в работе приведены краткие описания искусственного интеллекта применяемых в здравоохранении, пандемии Covid-19, наборов данных, моделей и сетей глубокого обучения, проблем. Описание наборов данных и моделей глубокого обучения представлено в виде таблицы.

Ключевые слова: Covid-19, глубокое обучение, набор данных, рентгеновский снимок, компьютерная томография, проблема, ошибка.

Abstract. This paper is directed to methods of detection of covid-19 pandemic by CT and X-ray images using Deep learning techniques. Also, here is given brief description about Artificial intelligence in health care, Covid-19 pandemic, datasets, Deep learning models and networks, challenges. Description of datasets and Deep learning models are presented as a table.

Keywords: Covid-19, deep learning, dataset, X-ray image, CT, problem, error.

Introduction

Artificial intelligence has been inspired by the functioning of biological neurons and includes the basics of sensing, recognition, and object recognition to enable machines to perform as good as or even better than humans. Machines can be more precise, reliable, and comprehensive and have relatively lower risk of bias; however, they still lack the elements of trust and empathy

Safety in health care implies the reduction or minimization of risks and uncertainty of harmful events. Machine learning applications are largely classified as type A (eg, medical diagnosis) and type B applications (eg, speech transcription systems), depending on safety and risk minimization. Although safety is of paramount importance in type A applications, risk minimization is the focus in type B applications. Machine learning has gained importance in the prevention, diagnosis, and management of various disease conditions.

Artificial intelligence plays an important role in augmenting knowledge and improving outcomes in health care. Artificial intelligence has widespread applications for the prediction and diagnosis of disease, handling of large quantities of data and synthesis of insights, and maximizing efficiency and outcomes in medical management of disease states [1].

Artificial intelligence has several applications in diagnosis and decision support. Artificial intelligence enables decision makers to access the right and up-to-date information to help make better decisions in real time. Application of artificial intelligence has brought about an evolutionary change in radiological diagnosis by improving the value and accuracy of image analysis [2]. Designs based on deep learning have enabled digital image analysis for the early detection of breast pathologies with precision [3]. In another example, an ML software library has been trained to detect changes in Parkinson’s disease by DaTscan image analysis. This library can be a useful adjunct to clinical diagnosis [4].

COVID-19 pandemic

The COVID-19 pandemic, also known as the coronavirus pandemic, is an ongoing pandemic of coronavirus disease 2019 (COVID-19) caused by the transmission of severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2), which was first identified in December 2019 in Wuhan, China [5]. The outbreak was declared a Public Health Emergency of International Concern in January 2020, and a pandemic in March 2020. As of 22 October 2020, more than 41.2 million cases have been confirmed, with more than 1.13 million deaths attributed to COVID-19. Common symptoms include fever, cough, fatigue, breathing difficulties, and loss of smell. Complications may include pneumonia and acute respiratory distress syndrome. Pneumonia is an inflammatory condition of the lung primarily affecting the small air sacs known as alveoli. Identifying the responsible pathogen can be difficult. Diagnosis is often based on symptoms and physical examination [6]. Chest X-rays, blood tests, and culture of the sputum may help confirm the diagnosis. A chest radiograph, called a chest X-ray (CXR), or chest film, is a projection radiograph of the chest used to diagnose conditions affecting the chest, its contents, and nearby structures. Chest radiographs are the most common film taken in medicine. X-ray images are sent to analyze to Pneumonia. Analyzing is provided by special tools which constructed by image analyzing methods. After becoming Covid-19 virus is developed several projects by developers and scientists.

Two types of images are used to diagnose to Covid-19: CT scans, X-Ray images.

In figure 1 is illustrated a typical methodology of a DL based detection system, where the system uses a deep learning algorithm to predict whether the X-ray images of suspected patient's lung is normal or having COVID-19 pneumonia.

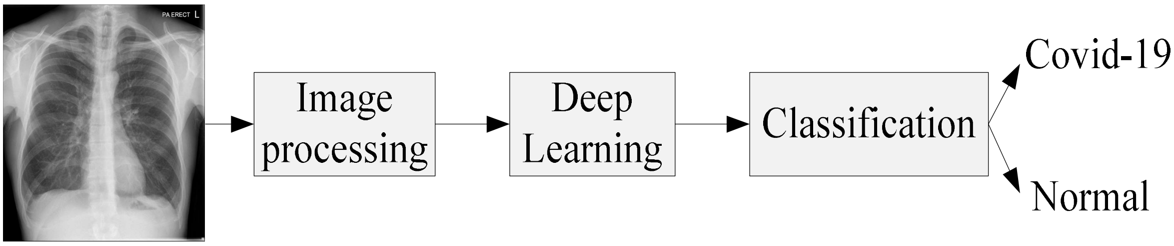

Fig. 1. A overall block diagram of a Deep learning-based detection system

Image datasets

Also, to trained this deep learning, large datasets as well as powerful computing resources are required. For a new pandemic, data insufficiency and it's variation over different geographic regions is a huge problem, so here Deep Transfer Learning would be effective as it learns from one task and could apply in another task after required fine-tuning. IoT, Webcam, Drone, Intelligent Medical Equipment, Robot are very useful tools in any pandemic situation.

Table 1. Covid-19 Image datasets

Organization	**Dataset**	**Types of Data**						**Number of images**
		X-Ray	CT	MRT	Ultrasound	Meta data	Case review	
Allen Institute for Artificial intelligence and partners [7]	CORD-19	Y	Y	N	N	Y	N	50 000
Institute for Reproducible Research [8]	COVID-19 Image Data Collection	Y	Y	U	U	Y	Y	195
Vision and Image Processing Lab [9]	COVID-Net Open Source Initiative	Y	N	N	N	N	N	14 000
Dep. for Biosystems Science and Engineering [10]	COVID-19 Pocus Ultrasound Dataset	N	N	N	Y	Y	N	60
Italian Society of Radiology [11]	SIRM COVID-19 Database	Y	Y	U	U	Y	Y	115
British Institute of Radiology [12]	COVID-19 BSTI Imaging Database	Y	Y	N	U	Y	Y	-
Kaggle [13]	COVID-19 Chest X-Ray Database	Y	N	N	N	Y	N	2905
	COVID-19 Chest X-Ray	Y	N	N	N	Y	N	371
medicalsegmentation.com [14]	COVID-19 CT Segmentation Dataset	N	Y	N	N	N	N	100
University of California San	COVID-CT	N	Y	N	N	Y	N	349

Diego [15]								
Dep. of Biomedical Systems & Informatics Engineering [16]	Augmented COVID-19 X-ray Images Dataset	Y	N	N	N	N	N	912
Radiology Artificial Intelligence [17]	coronacases.org	N	Y	N	U	Y	Y	-
European Society of Radiology [18]	eurorad.org	Y	Y	U	U	Y	Y	16037
Radiopaedia [19]	radiopaedia.org	Y	Y	U	U	Y	Y	-

Some challenges on detection of Covid-19

There are following problems and challenges on detection of Covid-19:

Quality of X-ray images. Quality of images is main parameter in deep learning based approaches. Therefor X-ray analyzing need to provide on efficiency medical diagnostic tools.

Size of Dataset or number of images. Deep learning nets are directed to train and test under images. Number of images per object marks accuracy of detection. Infected and not infected lung X-ray images divided in to different categories truly. Also, collection of large dataset is time consuming process.

Variety of datasets. This COVID-19 virus has mutating itself over different geographic regions, environments, and time [20]. Therefore, the pandemic dataset collected from one region for this region peoples, for other regions are used other datasets. Also, Most of the medical data comes from China and European countries which may lead to selection bias when applied in other countries. As a result, the practice of diagnosing a patient with COVID-19 using Artificial intelligence/Machine learning is very rare. Moreover, it is yet to be investigated if Artificial intelligence/Machine learning can detect COVID-19 before its symptoms appear in other laboratory methods to justify its practice [21].

High computational resources. High computational resources required for a DL network and model.

FAR and FRR errors. The FAR - false acceptance rate is the probability of cases for which a system inaccurately returns a positive detection. The FRR - false rejection rate is the probability of cases for which a system inaccurately returns a no-match detection. These types of errors mark accuracy of image based diagnosis system.

Most of the data and code on COVID-19 analysis is closed source. Whatever data is available, it is limited for applications of deep learning methods.

Relation between Deep learning specialist and Covid-19 medical doctor. In deep learning, specialists, doctors and radiologists must know which features distinguish a COVID-19 case from non-COVID-19. the probability of error needs to be estimated and communicated with the practitioners and patients [22].

Representative works

After becoming Covid-19 various CT-scanning automated approaches have been proposed. A lot of approaches and performances of the computer vision CT-based disease diagnosis were developed and here have been selected some recent representative works that provide an overview of their effectiveness. In table 2 is presented representative works for CT and X-ray image based COVID-19 analysis.

Table 2. Representative works for CT and X-ray image based COVID-19 analysis

Method and model	**Accuracy (%)**	**Specification**
For CT images		
UNet++ [23]	95,24	106 patients with 51 confirmed COVID-19 pneumonia. 46,096 CT images.
Modified inception [24]	82,9	Modified inception with transfer learning. 453 CT images of pathogen confirmed COVID-19 with 99 patients.
Combination of Two	86,7	Combination of Two CNN three-dimensional classification

CNN [25]		models and VNET based segmentation model. 618 CT samples. 110 patients with COVID- and 224 patients with Influenza-A viral pneumonia.
DRE-Net + ResNet50 [26]	86	DRE-Net + ResNet50 with Feature Pyramid Network. 777 CT images from 88 patients diagnosed with the COVID-19 and 86 healthy persons.
2D deep CNN based on Resnet-50 [27]	95	U-net architecture for image segmentation. 56 patients with confirmed COVID-19 diagnosis.
VB-Net [28]	91,6	VB-Net to segment COVID-19 infection regions in CT scans. 249 COVID-19 patients, and validated using new COVID-19 patients.
Grey Level Matrix [29]	99,68	Grey Level Matrix Discrete Wavelet Transform + SVM. 150 CT images
COVNet [30]	96	COVNet was developed to extract visual features from volumetric chest CT RESNET50. U-Net is used for segmentation. 4356 chest CT images were collected from 6 hospitals and 3,322 patients.
DeCoVNet [31]	95,9	3D deep CNN to Detect COVID-19 from CT volumes. U-Net is used for segmentation. 540 patients from Union Hospital.
Transfer learning on ResNet-50 [32]	92,2	Segmentation model as 3D U-Net++. Used 1,136 training cases (723 positives for COVID-19) from five hospitals.
For X-ray images		
U-Net+ CLAHE [33]	97,5	Combining with U-Net+ adversarial+ Contrast Limited Adaptive Histogram Equalization. 247 images from Japanese+Shenzhen dataset contains a total of 662 chest X-rays
ResNet+ DeTraC [34]	95,12	CNN features of pre-trained models on ImageNet and ResNet+ Decompose, Transfer, and Compose for the classification of COVID-19 chest images. I80 samples of normal CXRs.
ResNet50 [35]	97	ResNet50 InceptionV3. open source X-Ray Images.
COVID-Net [36]	92,4	Lightweight residual projection expansion projection-extension design pattern. 16,756 chest radiography images across 13,645 patient.

Feature works and methods not related to lung images

Using Deep learning image processes methods is detected Covid-19. Also, there are following methods to diagnose this types of viruses:

Masked Face Recognition approach [37]. In this approach is used e lot of masked (Covid and no covid) face images to analyzing. Because, healthy and ill faces are differ from each other.

Infrared thermography [38]. This approach is recommended as an early detection strategy for infected people, especially in crowns like passengers on an airport, underground, bus, markets and shopping centers. To detect temperature is used special scanners, there are methods which directed analyzing thermography infected and healthy faces. It is useful for distance detection infections and detection speed high. Also, may be detection error shows high degree.

Pandemic drones [39]. It uses remote sensing and digital imagery, which were recommended for identifying infected people.

Conclusion

Modern-era largely depends on image processing Artificial Intelligence including Data Science - Deep Learning is one of the current flag-bearer of these techniques. Therefore, these techniques could also assist to mitigate COVID-19 virus and dangerous pandemics in terms of stop spread, diagnosis of the disease, drug & vaccine discovery, treatment, and many more. Artificial intelligence based Machine learning and Deep learning methods and the available datasets, resources, and results in the fight against COVID-19. In this paper is provided the Deep learning models, covid-19 lung image datasets and challenges also, health community a comprehensive overview of the current state-of-the-art methodologies and applications with details of how Deep learning and data can improve the status of COVID-19, and further studies to stop the COVID-19 outbreak. 2019-2020 years were developed a lot of works to detect lung and these works are presented as a table with model, specifications and accuracy degree. After becoming Covid-19 infected and healthy lung datasets spread over world. At the end of this work is presented feature approaches and other methods to detect Covid-19 infections.

References

1. Ellahham S., Ellahham N., Simsekler M. C. E. Application of artificial intelligence in the health care safety context: opportunities and challenges //American Journal of Medical Quality. – 2020. – Т. 35. – №. 4. – С. 341-348.
2. Mayo RC, Leung J. Artificial intelligence and deep learning—radiology's next frontier? Clin Imaging. 2018;49: 87-88
3. Robertson S, Azizpour H, Smith K, Hartman J. Digital image analysis in breast pathology-from image processing techniques to artificial intelligence. Transl Res. 2018;194: 19-35.
4. Zhang YC, Kagen AC. Machine learning interface for medical image analysis. J Digit Imaging. 2017;30: 615-621.
5. Novel Coronavirus – China". World Health Organization (WHO). Retrieved 9 April 2020
6. How Is Pneumonia Diagnosed?". NHLBI. 1 March 2011. Archived from the original on 7 March 2016. Retrieved 3 March 2016
7. Wang L. L. et al. CORD-19: The Covid-19 Open Research Dataset //ArXiv. – 2020.
8. Cohen J. P. et al. Covid-19 image data collection: Prospective predictions are the future //arXiv preprint arXiv:2006.11988. – 2020.
9. Kalkreuth R., Kaufmann P. COVID-19: a survey on public medical imaging data resources //arXiv preprint arXiv:2004.04569. – 2020.
10. Born J. et al. POCOVID-Net: automatic detection of COVID-19 from a new lung ultrasound imaging dataset (POCUS) //arXiv preprint arXiv:2004.12084. – 2020.
11. Ko H. et al. COVID-19 pneumonia diagnosis using a simple 2D deep learning framework with a single chest CT Image: Model Development and Validation //Journal of Medical Internet Research. – 2020. – Т. 22. – №. 6. – С. e19569.
12. Caetano A. P. et al. Development of a Portuguese COVID-19 Imaging Repository and Database: Learning and Sharing Knowledgeduring a Pandemic //Acta Médica Portuguesa. – 2020. – Т. 33. – №. 13
13. COVID-19 Chest X-Ray Database. https://www.kaggle.com/tawsifurrahman/covid19-radiography-database. Accessed: 2020-10-15
14. COVID-19 CT Segmentation Dataset. http://www.medicalsegmentation.com/covid19/. Accessed:2020-10-15
15. Zhao J. et al. COVID-CT-Dataset: a CT scan dataset about COVID-19 //arXiv preprint arXiv:2003.13865. – 2020.
16. Alqudah A. M., Qazan S. Augmented COVID-19 X-ray Images Dataset //Mendeley Data, v4http://dx. doi. org/10.17632/2fxz4px6d8. – 2020. – Т. 4
17. Coronacases. https://coronacases.org. Accessed:2020-10-15
18. Eurorad. https://www.eurorad.org/. Accessed: 2020-10-15
19. Radiopedia. https://www.radiopaedia.org. Ac-cessed: 2020-10-15
20. Shen Z. et al. Genomic diversity of SARS-CoV-2 in Coronavirus Disease 2019 patients //Clinical Infectious Diseases. – 2020.
21. Nguyen T. T. Artificial intelligence in the battle against coronavirus (COVID-19): a survey and future research directions //Preprint, DOI. – 2020. – Т. 10.
22. Naudé W. Artificial intelligence vs COVID-19: limitations, constraints and pitfalls //Artificial intelligence & Society. – 2020. – С. 1.
23. Chen C. et al. SARS-CoV-2–positive sputum and feces after conversion of pharyngeal samples in patients with COVID-19 //Annals of internal medicine. – 2020.
24. Wang S. et al. A deep learning algorithm using CT images to screen for Corona Virus Disease (COVID-19) //MedRxiv. – 2020.

25. Butt C. et al. Deep learning system to screen coronavirus disease 2019 pneumonia //Applied Intelligence. – 2020. – С. 1.

26. Song Y. et al. Deep learning enables accurate diagnosis of novel coronavirus (COVID-19) with CT images //medRxiv. – 2020.

27. Gozes O. et al. Rapid Artificial intelligence development cycle for the coronavirus (covid-19) pandemic: Initial results for automated detection & patient monitoring using deep learning ct image analysis //arXiv preprint arXiv:2003.05037. – 2020.

28. Shan F. et al. Lung infection quantification of covid-19 in ct images with deep learning //arXiv preprint arXiv:2003.04655. – 2020.

29. Barstugan M., Ozkaya U., Ozturk S. Coronavirus (covid-19) classification using ct images by machine learning methods //arXiv preprint arXiv:2003.09424. – 2020.

30. Li L. et al. Artificial intelligence distinguishes COVID-19 from community acquired pneumonia on chest CT //Radiology. – 2020.

31. Shi H. et al. Radiological findings from 81 patients with COVID-19 pneumonia in Wuhan, China: a descriptive study //The Lancet Infectious Diseases. – 2020.

32. Jin S. et al. Artificial intelligence -assisted CT imaging analysis for COVID-19 screening: Building and deploying a medical Artificial intelligence system in four weeks //medRxiv. – 2020.

33. Gaál G., Maga B., Lukács A. Attention u-net based adversarial architectures for chest x-ray lung segmentation //arXiv preprint arXiv:2003.10304. – 2020.

34. Abbas A., Abdelsamea M. M., Gaber M. M. Classification of COVID-19 in chest X-ray images using DeTraC deep convolutional neural network //arXiv preprint arXiv:2003.13815. – 2020.

35. Narin A., Kaya C., Pamuk Z. Automatic detection of coronavirus disease (covid-19) using x-ray images and deep convolutional neural networks //arXiv preprint arXiv:2003.10849. – 2020.

36. LINDA W. A tailored deep convolutional neural network design for detection of covid-19 cases from chest radiography images //Journal of Network and Computer Applications. – 2020.

37. Wang Z. et al. Masked face recognition dataset and application //arXiv preprint arXiv:2003.09093. – 2020.

38. Lahiri B. B. et al. Medical applications of infrared thermography: a review //Infrared Physics & Technology. – 2012. – Т. 55. – №. 4. – С. 221-235.

39. Al-Naji A. et al. Life signs detector using a drone in disaster zones //Remote Sensing. – 2019. – Т. 11. – №. 20. – С. 2441.

Information about the authors

Khudoykulov Zarif Turakulovich, PhD, assistant professor, head of the department at Tashkent University of Information Technologies named after Muhammad al Khwarizmi

Islomov Shaxboz Zokir ugli, PhD student Tashkent University of Information Technologies named after Muhammad al-Khwarizmi

Davronova Lola Uktamovna, senior teacher at Tashkent University of Information Technologies named after Muhammad al-Khwarizmi

Rustamova Sanobar Ravshanovna, student Tashkent University of Information Technologies named after Muhammad al-Khwarizmi

Сведения об авторах

Худойкулов Зариф Туракулович, кандидат технических наук, доцент, заведующий кафедрой Ташкентского университета информационных технологий имени Мухаммада ал-Хоразмий

Исломов Шахбоз Зокир угли, кандидат технических наук, доцент Ташкентского университета информационных технологий имени Мухаммада Аль-Хорезми

Давронова Лола Uktamovna, старший преподаватель в Ташкентском университете информационных технологий имени Мухаммада ал-Хоразмий

Рустамова Санобар Равшановна, студент Ташкентского университета информационных технологий имени Мухаммада Аль-Хорезми

Address for correspondence

100000, Uzbekistan, Tashkent, Amir Temur-108, Tashkent University of Information Technologies named after Muhammad al-Khwarizmi

phone. +998 99 842 06 01

e-mail: shaxboz4044@gmail.com

Islomov Shahboz Zokir ugli

Адрес для корреспонденции

100000, Узбекистан, Ташкент, Амир Темур-108, Ташкентский университет информационных технологий имени Мухаммеда Ал-Хоразмий

телефон. +998 99 842 06 01

e-mail: shaxboz4044@gmail.com

Исломов Шахбоз Зокир угли