ISSN 2063-5346

DACTYLOGRAM TO IDENTIFY ACCIDENT VICTIMS A NEW DATA RECOVERY METHOD

Ahmed Abdul Rafay[1] Dr. Sunil Vijay Kumar Gaddam[2]

Article History: Received: 10.05.2023 Revised: 29.05.2023 Accepted: 09.06.2023

Abstract

The data facts of casualties are lacking as a result of concerns about accidents and incident cases not being updated efficiently in online databases, mainly by private and government-funded hospitals. It is challenging to trace the specifics of such fatalities from unforeseen mishaps and occurrences using the technique currently in use. To get around this, creating an integrated Dynamic Data Search Approach (DDSA) utilising the indiviual's dactylogram or thumbprint from which additional information like Aadhaar details and cellphone numbers are taken will allow the victim's details to be monitored with the pre-existing databases. In this project, we aim to explore two techniques for identification of dactylograms or fingerprints using (a) image processing and computer vision techniques like opencv and (b) machine learning and deep learning techniques. Since one would not have access to Aadhaar database, we have taken a sample Aadhaar like database which contains information like Name, Address, Phone No, Date of Birth, biometric details etc. The fingerprints are matched against the fingerprint database and using these details, personal information of victims is retrieved from the Aadhaar-like datasbase. We observe that both the techniques perform well on sample data of about 80 fingerprints. Machine Learning models have been created using Catboost, CNN and Random Forest algorithms and they have been evaluated for metrics like accuracy, precision, recall, F1-score and support and predictions have been made and the results are saved. The accuracy of methods is recorded to be about 96%. It can be observed that the first process is suitable for one-to-one match where as the second process is suitable for batch processing

Index Terms — Accidental; Reinforcement Learning; Finger impression;Automated Details Retrieval System(ADRS).

[1]Research Scholar, Dept. of Computer Science and Engineering, Lords Institute of Engineering & Technology, Hyderabad, Telangana

[2]Professor, Dept. of Computer Science and Engineering, Lords Institute of Engineering & Technology, Hyderabad, Telangana

DOI:10.48047/ecb/2023.12.9.81

I. INTRODUCTION

Modern The information related to those victims is far behind and sometimes they pass away in hospitals without having their deaths recorded in the security databases, according to reports on roads and safety, because countless deaths from accidents that happen every day are not automated and stored in databases. According to survey reports from the Ministry of Road Transport & Highways (MoRTH) and National Crime Records Bureau (NCRB) data, even if the information is stored, it takes a lot of manual labour to resolve the issue, which is a time-consuming process that can take anywhere from 1 week to 6 months [1] [2]. Particularly at private hospitals, if a victim is hospitalised, their information is briefly retained in computers before being rendered useless. Only the victim(s)' details will be retained for future examination if emergency action is required by government officials, which only happens in extremely rare circumstances. According to the most recent NCRB data, India saw 1,54,732 fatalities and 4,39,262 fatal injuries in 2019. The victim's identity will be registered as "Medico Legal Case" with no name when they are brought into a government hospital, and if they pass away later, an autopsy will be performed on their body while saving the information and assigning a special number. Most of the time, the government adheres to predetermined rules, such as transporting the body to a morgue and keeping it there for a predetermined amount of time. When a deadline is missed, the body is incinerated under police officers' watchful eyes. The government has discovered that the budget costs spent for the entire procedure are prohibited at the time leading up to the cremation of the body. The government spends millions of dollars each year on the upkeep and development of morgues, but if the identity of the deceased person could be located, this expenditure might be significantly reduced. Once the information was obtained, the deceased person's remains would have been sent to their relatives. However, the majority of these countries are now ignoring this issue. The government must make advantage of cutting-edge technical advancements that assist in locating victim information in order to cut down on the extra expenses placed on the government.

II. SYSTEM ANALYSIS

Problem Statement:

Identification of victims during accidents is very essential during mishaps as it would help in identification of persons. Deaths of accident victims are not being recorded as the bodies are not being identified. Having a fast and efficient data retrieval system based on biometrics would help in reducing the unnecessary cost incurred to the government.

Objective

In this project, we propose to build an integrated Dynamic Data Search Approach (DDSA) using victims Finger-print from which Aadhaar details and mobile numbers are extracted. This helps in resolving problems faced in identifying such victims from unexpected accidents and incidents and helps in tracing their details. Two techniques are using in the above mentioned proposed process for matching dactylograms: (a) Computer Vision Techniques like Image Processing and Open CV (b) Machine learning and Deep Learning techniques. We have used FVC2002 Fingerprint dataset in this project

Proposed System:

This DDSA utilizes two methods namely (a) Computer Vision Techniques and (b) Machine Learning concepts for finger print recognition and matching.

(a) Using Image processing and Open CV:

It first accepts the thumbprint as an input through a sensor, and then proceeds to extract the Whorls & Ridges of the finger(s) through a series of processing steps including Acquisition, Grayscale Transformation, Normalization, Segmentation, Skeletonization and Singularity Detection. Using this process, the fingerprint is compared against other fingerprints in the database and a match can be retrieved. The details of the matched fingerprint are extracted from Aadhaar like database.

(b) Using Machine Learning and Deep Learning Techniques:

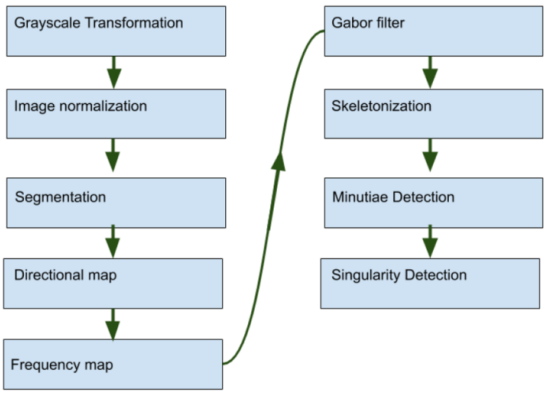
The fingerprint dataset has been taken and using for training and testing. The training dataset has accurate labels for identification of persons but the testing dataset has invalid labels. The fingerprints are converted to pixel information and fed to machine learning algorithms and deep learning algorithms for classification. Machine Learning models have been created using Catboost, CNN and Random Forest algorithms and they have been evaluated for metrics like accuracy, precision, recall, F1-score and support and predictions have been made and the results are saved.

Advantages of Proposed System

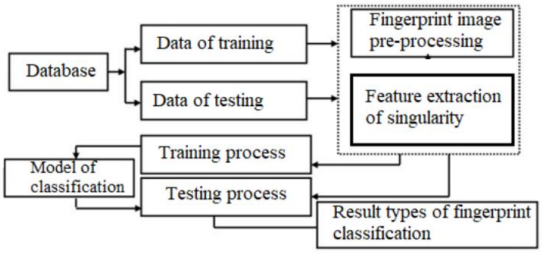
- By correctly identifying individuals and turning them over to the affected family and friends in situations of accidents, the DDSA will save the administration time and money while also assisting in the reduction of crimes that may employ biometric data.

Algorithm/ Technique Used

(a) Computer Vision Techniques



(b) Machine Learning Techniques



III. PROPOSED MODULAR IMPLEMENTATION

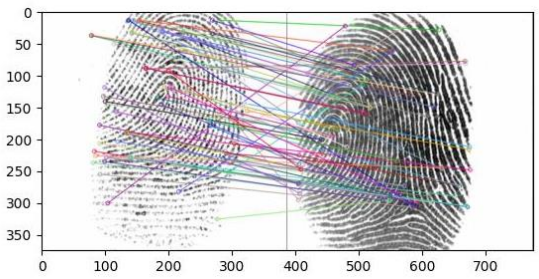
Algorithm for One to one matching:

1. Load the fingerprints into the application
2. User uploads a fingerprint to be matched with the fingerprint database
3. Iterate over the files in the fingerprint database and match the user uploaded fingerprint using the below process

a. Equalize the image
b. Normalise the image and find a ROI
c. find orientation of every pixel
d. find the overall frequency of ridges
e. create gabor filter and do the actual filtering
f. Set the threshold value for image segmentation
g. Perform skeletonization. Skeletonization is a process of reducing foreground regions in a binary image to a skeletal remnant that largely preserves the extent and connectivity of the original region while throwing away most of the original foreground pixels. In simpler

words, Skeletonization makes a BLOB very thin.

h. Find internal corners of the image using **Harris Corner detection. This** algorithm was developed to identify the internal corners of an image. The corners of an image are basically identified as the regions in which there are variations in large intensity of the gradient in all possible dimensions and directions.
i. Extract key points and compute descriptors
j. Match keypoint descriptors using BF matcher
k. Compute the threshold value
l. If the fingerprint has the match score greater than the threshold score, then it means that the fingerprint had matched.
m. Below is a sample match



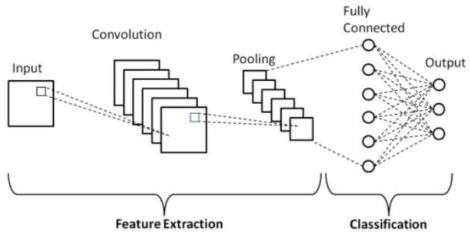
Algorithm using Deep Learning

1. Load Fingerprint dataset into the application
2. Create a data frame containing the file name of the data
3. Check if the image has been converted to pixels
4. Feed this pixel data frame to the below algorithms and the accuracies are recorded as follows:

Algorithm	Accuracy %
Catboost	98
CNN	98.21
Random Forest	98

CNN model is defined as below:

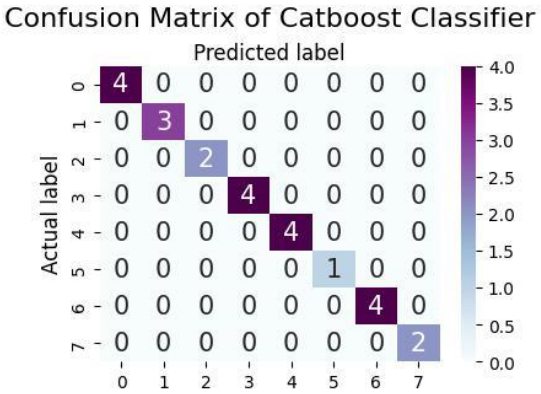
```
model = Sequential([
  Conv2D(nb_filters, kernel_size[0], kernel_size[1], padding='valid', input_shape=X_train.shape[1:],
      activation='relu'),
  Conv2D(nb_filters, kernel_size[0], kernel_size[1], activation='relu'),
  MaxPooling2D(pool_size=pool_size),
  Dropout(0.25),

  Flatten(),
  Dense(128, activation='relu'),
  Dropout(0.5),
  Dense(nb_classes, activation='softmax')
])
```



5. The results for test fingerprints are predicted and saved in csv format.

Mean Accuracy of Catboost Classifier: 0.98 (+/- 0.05)

Testing Accuracy of Catboost Classifier: 1.00



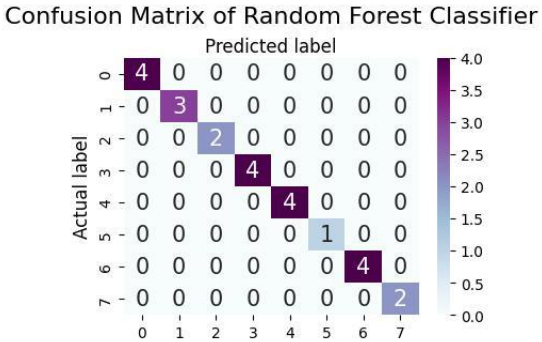
Catboost Classification Report:

	precision	recall	f1-score	support
1	1.00	1.00	1.00	4
2	1.00	1.00	1.00	3
3	1.00	1.00	1.00	2
4	1.00	1.00	1.00	4
5	1.00	1.00	1.00	4
6	1.00	1.00	1.00	1
7	1.00	1.00	1.00	4
8	1.00	1.00	1.00	2
accuracy			1.00	24

macro avg	1.00	1.00	1.00	24
weighted avg	1.00	1.00	1.00	24

Mean Accuracy of Random Forest Classifier: 0.98 (+/- 0.05)

Testing Accuracy of Random Forest Classifier: 1.00



Random Forest Classification Report:

	precision	recall	f1-score	support
1	1.00	1.00	1.00	4
2	1.00	1.00	1.00	3
3	1.00	1.00	1.00	2
4	1.00	1.00	1.00	4
5	1.00	1.00	1.00	4
6	1.00	1.00	1.00	1
7	1.00	1.00	1.00	4
8	1.00	1.00	1.00	2
accuracy			1.00	24
macro avg	1.00	1.00	1.00	24
weighted avg	1.00	1.00	1.00	24

Below is the proposed modular implementation of the project. It consists of two modules:

1. Admin
2. User

Admin Module:

The admin of the system is responsible for the activities like:

1. Uploading the fingerprint dataset
2. Data Pre-processing
3. Feature Extraction
4. Create the model.

User Module:

The user of the system can utilize the machine learning services that are offered like:

1. Logging into the system
2. Uploading the fingerprint
3. Search the database and retrieve the person's details based on the fingerprint.

IV. PROJECT EXECUTION

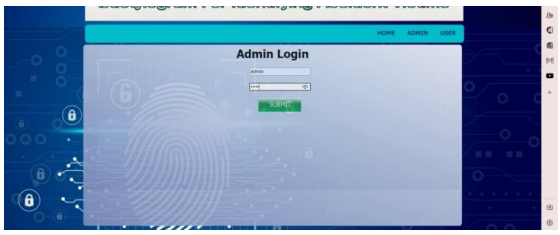
Home page:

This is the starting page of the application when the application is executed on Pycharm, the application is hosted on a web server and URL is generated to access the application once the user clicks on the URL the below page is opened on the browser.

Admin Login:

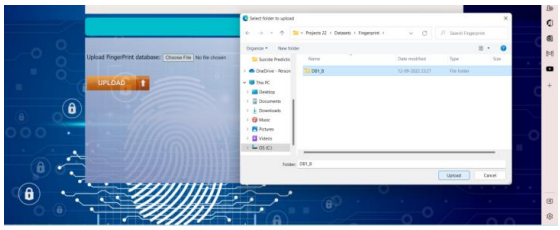
This is the login page for the admin module. The admin need to login into the system with his credentials in order to perform operations like uploading the dataset, Training the dataset, Exploratory data Analysis of the dataset, Feeding the dataset to different Machine learning Algorithms to find the Algorithm that can

meet the best accuracy and Create a model that can be hosted on the Flask Application to be used by the users.



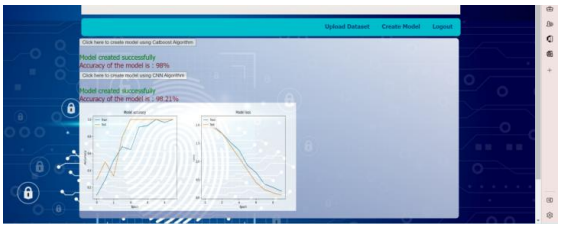
Upload Dataset:

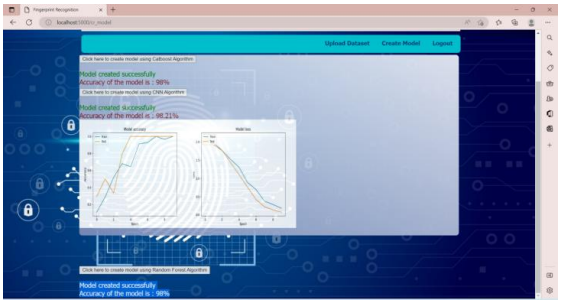
On this page, the administrator of the system can upload datasets that are used for training the machine learning models. The admin has to select the file by clicking on the Choose file button and click on the upload button to upload the file to the server. Once the upload is complete, a success message would be displayed that the file is successfully uploaded. For this project we are using diabetes.csv as a dataset.



Create Model:

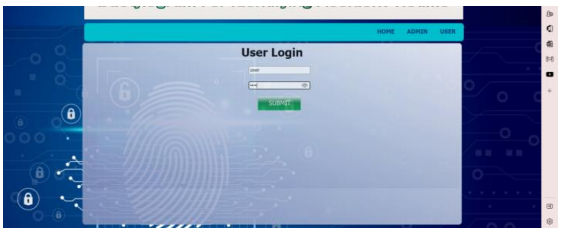
This screen shows the Accuracy of the Models about 98%.





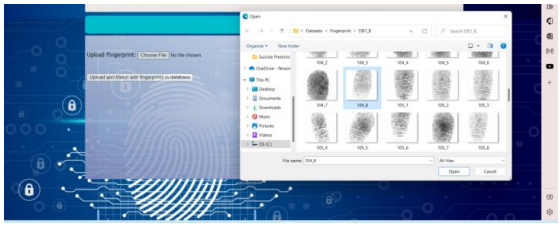
User Login:

This is the login page for the user module. The admin need to login into the system with his credentials in order to receive predictions based on health parameters and save those values into the database.



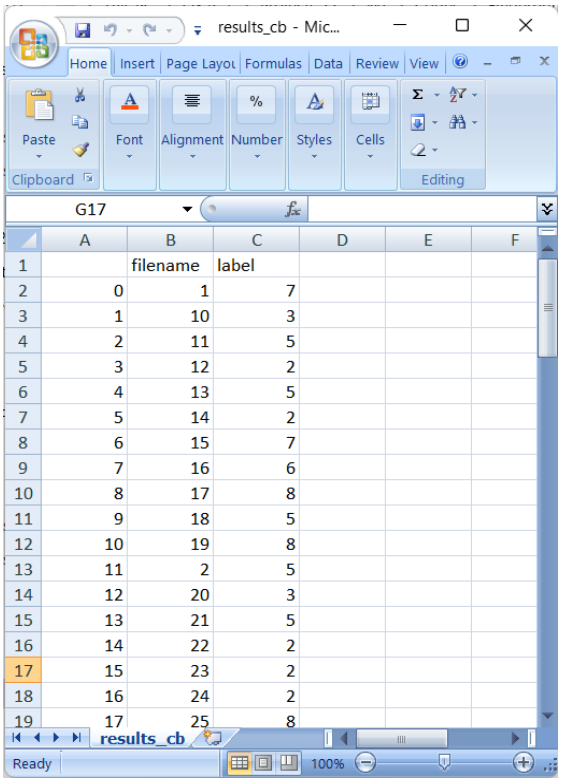
Predict using OpenCV:

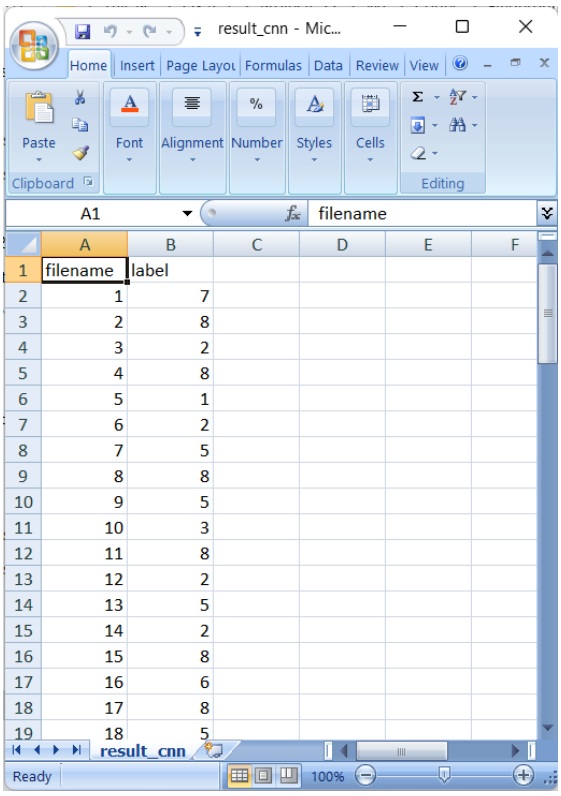
Once the user Logs into the system using this page, then the can enter the upload fingerprint and get the details of the person if the fingerprint matches with the one in databse.

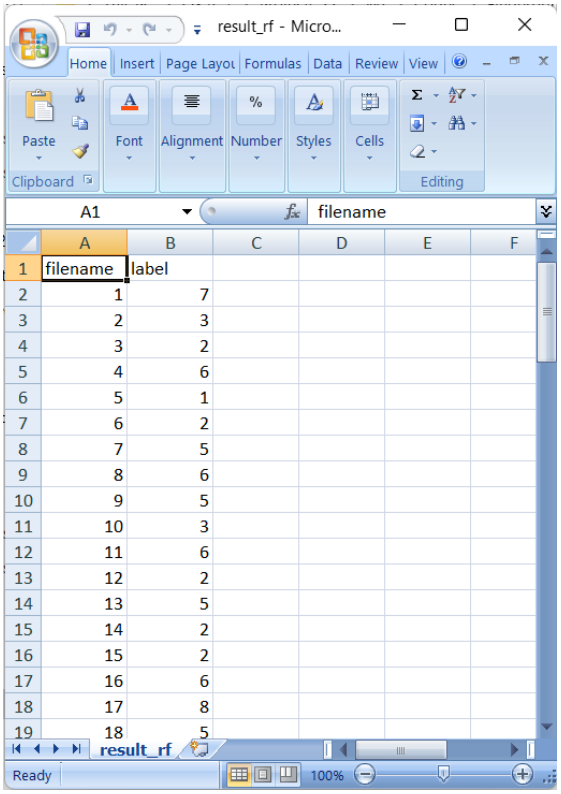


Prediction using ML Model:

Using this page, the user can predict the details associated with multiple fingerprints in a single attempt. It can be used for batch processing







	A	B
1	filename	label
2	1	7
3	2	3
4	3	2
5	4	6
6	5	1
7	6	2
8	7	5
9	8	6
10	9	5
11	10	3
12	11	6
13	12	2
14	13	5
15	14	2
16	15	2
17	16	6
18	17	8
19	18	5

CONCLUSION

In this project, we aim to explore two techniques for identification of dactylograms or fingerprints using (a) image processing and computer vision techniques like opencv and (b) machine learning and deep learning techniques. Since one would not have access to Aadhaar database, we have taken a sample Aadhaar like database which contains information like Name, Address, Phone No, Date of Birth, biometric details etc. The fingerprints are matched against the fingerprint database and using these details, personal information of victims is retrieved from the Aadhaar-like datasbase. We observe that both the techniques perform well on sample data of about 80 fingerprints. Machine Learning models have been created using Catboost, CNN and Random Forest algorithms and they have been evaluated for metrics like accuracy, precision, recall, F1-score and support and predictions have been made and the results are saved. The accuracy of methods is recorded to be about 96%. It can be observed that the first process is suitable for one-to-one match where as the second process is suitable for batch processing.

FUTURE WORK:

We intended to combine the information of the victims' facial and retinal (IRIS) (i.e., face and eyes) scans to further improve the system in order to address such challenges in the future. In addition to this, we also had plans for locating the victims' close family members and acquaintances by tracing and monitoring their mobile phone call histories through contacts with approved service providers of different cellular networks. Another significant future task is to improve our DDSA architecture by integrating capabilities that enable biometric thumbprint searching from global datasets for unique identification of a victim in the event that the victim's information cannot be correctly retrieved from the available repositories.

REFERENCES

[1] Dinesh Mohan, GeetamTiwari and KaviBhalla, (2020) “Road Safety in India: Status Report 2020”, Transportation Research & Injury Prevention Programme, Indian Institute of Technology Delhi. Link http://tripp.iitd.ac.in/assets/publication /Road_Safety_in_I ndia2018.pdf

[2] (Online) “Nation Crime Report Bureau Data”, May, 2020. link: https://ncrb.gov.in/

[3] (Online) “National Missing and Unidentified Persons (NamUS)”, database link: www.namus.gov/

[4] (Online) link: https://worldpopulationreview.com/sta te rankings/missing-persons-by-state.

[5] Mohammed MahmoodAli, KhajaMoizuddin Mohammed and Lakshmi Rajamani, (2014)

"Framework for Surveillance of Instant Messages in Instant messengers and Social neworking sites using Data Mining and Ontology," IEEE-Student‟s Technology Symposium, pp. 297-302.

[6] (Online) "The History of Fingerprints", link: https://onin.com/fp/fphistory.html

[7] KhinNandar win, Kenli Li, and et.al.,(2020) "Fingerprint classification and identification algorithms for criminal investigation: A survey", Elsevier, Future Generation Computer Systems Vol. 110, pp. 758-771.

[8] Arvind Kumar, Mahesh Kumar, Rishabh Kumar Singh, (2018) "Identification of unclaimed dead bodies- A possible aadhar based solution," International Journal of Forensic Medicine and Toxicological Sciences, October December, Vol. 3, Issue 4, pp. 74-76.

[9] Mohammed Mahmood Ali, "Datasets of Aadhaar details", link:https://www.dropbox.com/sh/decww48ffsmrw3z/AAB0QwkclLTY7sKhN7j9Nl0ta?dl=0

[10] Sera White law, Mamas A Mamas and et. al., (2020) "Applications of digital technology in COVID-19