Detection of the patient with COVID-19 relying on ML technology and FAST algorithms to extract the features

Seba Aziz Sahy¹, Sura Hameed Mahdi¹, Hassan Muwafaq Gheni², Israa Al-Barazanchi³

¹Institute of Medical Technology Al-Mansour, Middle Technical University, Baghdad, Iraq
²Department of Computer Techniques Engineering, Al-Mustaqlab University College, Hillah, Iraq
³Department of Computer Engineering Techniques, Baghdad College of Economic Sciences University, Baghdad, Iraq

ABSTRACT

COVID-19 is unquestionably one of the most hazardous health issues of our century, and it is a significant cause of mortality for both men and women throughout the globe. Even with the most advanced pharmacological and technical innovations, cancer oncologists, and biologists still have a substantial problem treating COVID-19. For patients with COVID-19, it is critical to offer initial, precise, and effective indicative procedures to increase their survival and minimize morbidity and mortality, which is currently lacking. A COVID-19 detection method has been presented in this paper for the initial identification of COVID-19 hazard factors. Features from accelerated segment test (FAST), a robust feature, was used to extract features in this suggested method. The experiments show that it is possible to identify FAST traits efficiently. A consequence was a high success rate (98%) for accuracy performance.

Keywords: COVID-19 Detection technique FAST descriptor Feature extraction

This is an open access article under the CC BY-SA license.

1. INTRODUCTION

The World Health Organization has confirmed the existence of COVID-19, a new coronavirus illness discovered in Wuhan, China, in December 2019 [1]. With no effective treatment, this virus spread quickly within and beyond China, causing severe acute respiratory syndrome (SARS) in people who got afflicted. At this point [2], a worldwide epidemic has become a pandemic. COVID-19 may be quantified and identified using various methods, including reverse transcription-polymerase chain reaction (RT-PCR), a chest computed tomography (CT) scan, and a chest x-ray, all of which have been employed. In the absence of other trustworthy biomarkers (such as blood routine and infection biomarkers), molecular analysis is the sole option. However, this method is both costly and time-consuming. They may also have serious adverse effects, such as secondary infection [3]. To validate COVID-19 illness, the commonly used RT-PCR test has the possibility of false negatives. Automated and reliable screening methods for COVID-19 patients are required because the RT-PCR test has a low sensitivity. For COVID-19 findings, medical imaging approaches such as chest x-ray and chest CT are non-invasive. COVID-19, on the other hand, might be difficult to detect in scans and photographs. Because of this, we urgently need advanced methods for predicting COVID-19 infection from medical images [4], [5]. As COVID-19's negative impacts spread around the globe, there is a cumulative reliance on artificial intelligence (AI) methods such as deep learning (DL) and machine learning (ML) to combat them on a variety of fronts. Given the correct input and unique algorithmic design, AI can recognize patterns, anticipate outcomes, aid in medical decision-making, and help
uncover valuable information from data [6]. Due to artificial intelligence's ability to process massive amounts of data in real-time. For example, ML and DL may be used to analyze very complex non-linear interactions in massive datasets to identify connections between predictors and outcomes. However, most standard statistical approaches cannot handle big data in many forms (e.g., texts and images). Capital investment in artificial intelligence-based medical imaging has skyrocketed to $6.6 billion by 2021 [7]. There was a correlation between a positive chest C.T. and an RT-PCR in 3% to 30% of patients with COVID-19 who had an initial negative RT-PCR test, as described in [8], [9]. The area of ML has grown substantially in the recent decade. Thanks to advancements in deep learning algorithms and computing power, the discipline has reached new heights. Because of this, ML is being used in a variety of industries. We believe that ML's most important use has been in combating COVID-19. Using ML, researchers have taken several approaches to tackle COVID-19, with varying degrees of success. Several ML applications [8] have been created to solve the infection's many problems. Here are the most recent results and accomplishments of the ML community's efforts to fight against the global epidemic. The following is an overview of the many studies that have been utilized in CT scans for COVID-19 categorization using features from accelerated segment test (FAST):

COVID-19 incidences may be detected from a patient's chest x-ray pictures using the inception (Xception) ML approach. Because of the time and money required to discover COVID-19 occurrences, they employed patient chest x-ray images to circumvent RT-PCR kit limits; their models outperformed comparable models [9]. Compare the performance of four prominent federal ML models, including ResNet18, mobile net, MobileNet-v2, and COVID-net, using the chest X-ray (CXR) patient chest imaging data set. ResNet18 was the fastest and most accurate, with a 91.26% accuracy rate and a 96.15% accuracy rate, after 100 rounds of testing.

The performance of unified learning models has been improved using a new integration-based learning strategy. Prior federal learning research has used virtual learning environments, which may lead to higher communication costs and poor performance when different consumer data. As a result, the system can recognize the interaction between consumers and servers, picking a client who participates in each round to get updates for his local model. The platform owner specified a maximum wait time for each customer during the server tour. They used this approach in four different models. This strategy outperformed the default option in terms of accuracy and was shown to lower the amount of time spent communicating with ResNet50, ResNet101, and training time, but not for GhostNet when applied to COVID-19 data sets [10], [11]. To avert the COVID-19 pandemic, it created a single model based on digital city twin principles to forecast the effects of different city prevention initiatives. It also used digital city twin systems to track the number of infections from various cities. Local models were created on each city's digital twin systems, which relayed model parameters or adjustments to suitable locations while ensuring data privacy. The federal model and the traditional method have been compared regarding anticipated accuracy and loss [12].

Automated federal learning automation (AutoFL) and the federated NAS (FedNAS) algorithm may be used to find optimal design parameters for local ML models that exchange model updates to improve their performance and effectiveness. Customers included mismatched and non-independent data, i.e., non-accounting clients, and the default parameters of local ML models did not fit the federal context. Experimental results show that FedNAS outperforms federated average (FedAvg) by 81.24% on the Canadian Institute for advanced research10 (CIFAR10) data set, compared to FedAvg's 77.78% [13].

COVID-capsule (COVID-CAPS), a modelling framework based on capsule networks, was created to address the constraints of CNN-based models in dealing with constrained data sets. It was found that COVID-CAPS beat the conventional network in terms of performance [14] by adjusting parameters in the model. An in-depth investigation of the current categories used to forecast COVID-19's future status yielded four distinct groupings. The authors compiled a detailed review to help ML practitioners improve their models and cope with challenges [15]. Their work proposed a COVID-19 classification system that uses deep X-ray learning to outperform the current methods [16]. Monitoring and detection system for COVID-19 patients in real-time. Seven machine learning algorithms were evaluated and compared during quarantine using an IoT device. According to them, five approaches have a prediction accuracy of higher than 90% [17]. To detect COVID-19 instances, a twisted neural network was used to analyze chest X-rays. There were 130 COVID-19 X-ray pictures and 130 standard images in the patient's chest X-ray data set, which was processed using convolutional neural network's (CNN's) ML method. The ML prediction model was dead on provided in [18]. Data analytics, data mining, and natural language processing are included in this look at the A.I. approach utilized in numerous COVID-19 applications natural language processing (NLP). The authors used COVID-19 AI algorithms on chest x-ray image sets to discover previous problems and treatments [19].

2. METHOD

FAST: in computer vision applications, feature points may be eliminated by employing the FAST angle detection approach. In 2006, Edward Ruston and Tom Drummond released the FAST angle detector... 

Detection of the patient with COVID-19 relying on ML technology and FAST algorithms ... (Seba Aziz Sahy)
for the first time [20]. Most noticeable is the computing efficiency of the FAST angle detector. Like the Gaussian (Mod) difference utilized by scale-invariant feature transform (SIFT), Susan, and Harris, this technique is quicker than several other well-known approaches for extracting functions. Methods that use ML may also shorten processing times and increase the efficiency of system resources. The FAST angle detector's high-speed capabilities make it perfect for real-time video processing [21]-[25].

2.1. The detector of segment tests

The quick angle detector uses a 16-pixel circle (Brenham circle of 3 radii) to assess whether the p filter point is an angle. In the circle, each pixel is labelled from one to sixteen clockwise [26]-[30]. Any N-pixel circle in which the combined brightness of all pixels is brighter or darker than the sum of the t-pixel threshold values for all its neighbours includes "a corner," denoted by the IP code. The following are the prerequisites:

− Condition 1: a collection of N continuous pixels S, display style for all Xin display style for all Xin S, the intensity of x > Ip + threshold, or display style 1 x>1 p+displaystyle I x>I p+t

− Condition 2: a collection of N continuous pixels S, display style for all Xin display style for all Xin display style for all Xin S, display style I xI p-t, I xI p-t I xI p-t I xI p-t I xI p-t I xI p-t I xI p-t I xI p-t I xI p-t I xI p-t I xI p-t I xI p-t I xI p-t I xI p-t I xI p-t I xI p-t

If a specific condition is satisfied, candidate p may be considered a corner. The threshold value of t and the number of continuous pixels N represent a compromise. The number of designated angle points should not be excessive but rather excessive when it comes to designated angle points. However, you should not sacrifice computational efficiency to get high performance. In the absence of ML advancements, N is often set at 12 k. A high-speed testing policy may be utilized to eliminate non-corner sites [31]-[35]. Figure 1 shows the detector of segment tests.

![Figure 1. The detector of segment tests](image)

2.2. Test at high speeds

Four sample pixels are inspected in the high-speed test to exclude non-angle points: 1, 9, 5, and 13. At least three of these four pixels must have a brighter or darker sample than the filter angle because there must be at least 12 pixels in a row that are brighter or darker than the selected angle. The p filter is not a corner if both I1 and I9 are within [Ip-t, Ip+t]. Additionally, pixels 5 and 13 are measured to see whether any three are brighter or darker than Ip+t. Pixels are evaluated to see whether three of them are brighter or darker than the rest. As stated in the inventor's first publication, [21] verifying the existence of the filtered angle pixels takes an average of around 3.8 pixels. Compared to 8.5 pixels per filter angle, a 3.8-pixel decrease per angle substantially improves performance.

3. RESEARCH METHOD

Pretreatment of COVID-19 models is applied and maintained in the database throughout both the training and testing stages, which are the two main phases of the proposed system. This vital feature is extracted using the FAST method, which has the same purpose as the SIFT technique: reducing processing time to improve detection and matching speed. Figure 2 shows the proposal system for COVID-19.

3.1. Dataset

The planned study would identify COVID-19 utilizing a CT scan of the lungs. Angelov and colleagues gathered actual CT scans of patients from hospitals in Sao Paulo, Brazil, to produce a collection of lungs CT pictures divided into two categories: COVID and non-COVID. You may find it on Kaggle at https://www.kaggle.com/plameneduardo/sarscov2-ctscan-dataset if you are interested. A total of 2,481 CT
images are utilized in the training and testing process. There are 1,252 CT scans of positive COVID-19 patients and 1,229 CT scans of negative COVID-19 patients in the SAR-CoV-2 CT data set, with the majority of the scans being positive. Figure 3 depicts a sample of C.T. scans of the lungs from the data set, which may be seen here. 80% of the database was utilized for training, and 20% was used for testing, resulting in a split of the database.

3.2. Pre-processing stage

Most image processing professionals know that the most critical operations will not begin until the necessary pre-processing has been completed beforehand. Its purpose is to enhance the target image to get better outcomes in subsequent therapy. Pre-processing on models is performed in the following ways: transcoding a picture from red, green, and blue (RGB) to hue, saturation, and value (HSV) format and storing it in a database.
3.3. Feature extraction using FAST algorithm

In COVID-19 photos, the application of the FAST algorithm has provided researchers the ability to detect places of interest and point characteristics with greater precision. Following criteria and a predetermined threshold, we choose the most noticeable qualities. After that, we bring back all the components that reflect the subject's core traits and original picture. As a further step, data carrier features, also known as feature descriptors, are mined, and processed from nearby pixels. A pixel is a discrete area on a computer display that represents a collection of linked properties. Finally, we can establish the centre position of nearby pixels. We may extract the most vital representation points for a lung segment of COVID-19 and describe the places called due to downloading information that can be identified and detected. This will be shown by the algorithm shown in the next section.

**Step 1:** Select a pixel “p” on the CT scan of the COVID-19 picture from the list below. Take for example the case when the intensity of this pixel is equal to IP. In other words, this is the pixel that decides whether a pixel is an intriguing point.

**Step 2:** Set the T value for the threshold intensity (say 20 percent of the pixel under test).

**Step 3:** Imagine a 16-pixel circle encircling the pixels.

**Step 4:** If a pixel is identified as a point of interest, “N” out of 16 adjacent pixels must be above or below the IP value T, where N equals 12.

**Step 5:** To make the process faster, compare the brightness of the circuit's pixels 1, 5, 9, and 13 to the IP. At least three of these four pixels must fulfil the point of interest presence threshold criteria.

**Step 6:** P is not a point of interest if at least three of the four-pixel values - I₁, I₅, I₉, I₁₃ - are not greater or lower than IP. + T. (angle). Ignore the pixel as a potential point of interest in this situation. If more than three pixels are above or below IP + T, examine all 16 pixels to see if 12 adjacent pixels meet the requirement.

**Step 7:** Repeat the technique for all pixels in the C.T. scan picture COVID-19.

3.4. Machine learning using decision tree

This will be explained via an algorithm that describes the job of ML in recognizing the COVID-19, which has been added to the FAST algorithm to make the algorithm more accurate in the case of the process of detecting the COVID19-infected area of the lung:

**Step 1:** Select a collection of CT-Scan of COVID-19 pictures to use for training purposes.

**Step 2:** In each picture, use the FAST approach to identify the areas of interest by evaluating each of the 16 pixels in the circle one at a time.

**Step 3:** Save the 16 pixels around each pixel "p" into a vector in your computer.

**Step 4:** Repeat the process for every pixel in each image. This is the P vector, which contains all of the data from the training session. For each value in the vector (for example, one of the 16 pixels, x), there may be three potential states,

**Step 5:** p can be made darker, lighter, or approximately the same shade.

\[
S_{p-x} = \begin{cases} 
    d, & I_{p-x} \leq I_p - t \\
    s, & I_p - t < I_{p-x} < I_p + t \\
    b, & I_p + t \leq I_{p-x} 
\end{cases} 
\]

(darker) (similar) (brighter)

**Step 6:** Depending on the circumstances, the whole vector P is partitioned into three subsets: Pd, Ps, and Pb.

**Step 7:** Set up a Kp variable that returns true or false depending on whether the given point of interest (p) is one or the other.

**Step 8:** Ask each subset for actual class knowledge using the decision tree classification and the Kp variable.

**Step 9:** The decision tree algorithm is based on the principle of entropy minimization. Use the 16 pixels to determine the proper class (interest point or not) feasible in the shortest amount of time. Instead, choose the pixel x with the information regarding the pixel in question, as shown in the equation below. The set P's entropy can be expressed mathematically using the following formula:

\[
H(P) = (C + \tilde{C})\log_2 (C + \tilde{C}) - C\log_2 C - \tilde{C}\log_2 \tilde{C}
\]

Where C = | {p | Kp is True } | number of corners
Where \( \tilde{C} = | {p | Kp is False } | \) number of non corner

**Step 10:** Recursively reduce entropy in each of the three subgroups.

**Step 11:** If the entropy of a subset is zero, the operation should be terminated.

**Step 12:** Using the decision tree has learned query sequence for quicker detection in additional photos is Step 12 of the process.
4. RESULTS AND DISCUSSION

In this section, the outcomes of the COVID-19 identification technique are explained. It began with a representation of the original raw picture. A description was generated using the areas centered on the observed features as a starting point. Feature descriptors are then extracted from photos at the places of interest and used to create a descriptor. This process begins with pre-processing. Figure 4 illustrates transforming a small area of pixels into a vector representation that may be compared regardless of COVID-19 orientation or size. Table 1 shows the feature vector for FAST algorithm.

![Figure 4. FAST algorithm's feature vector](image)

<table>
<thead>
<tr>
<th>Image</th>
<th>Threshold</th>
<th>Non-max suppression</th>
<th>Total keypoints without non-max suppression</th>
<th>Total keypoints with non-max suppression</th>
<th>Neighborhood</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>25</td>
<td>True</td>
<td>607</td>
<td>194</td>
<td>2</td>
</tr>
<tr>
<td>2</td>
<td>25</td>
<td>True</td>
<td>124</td>
<td>46</td>
<td>2</td>
</tr>
<tr>
<td>3</td>
<td>25</td>
<td>True</td>
<td>236</td>
<td>79</td>
<td>2</td>
</tr>
<tr>
<td>4</td>
<td>25</td>
<td>True</td>
<td>343</td>
<td>100</td>
<td>2</td>
</tr>
</tbody>
</table>

Table 1. Feature vector for FAST algorithm

Different performance measures based on the confusion matrix quantify rating performance for models trained on upgraded and undocumented picture data sets. Using CT scans, these performance metrics include F1 points, accuracy, accuracy, remembering, and selecting. This article’s main objective is to identify whether the individual has COVID-19. Negative (showing the patient has not been infected) or positive (indicating an infection) are possible outcomes (meaning that the patient has not contracted the virus). There is no guarantee that the test findings for any given patient will be classified in the same layer category as their actual layer. True positive (TP) patients who test positive for COVID-19 are appropriately recognized, while true negative (TN) refers to patients who test negative for COVID-19 identified in this amendment. Consequently, in (1), accuracy is measured by dividing the total number of items by the correct number of classifications: COVID-19-negative patients.

4.1. Accuracy

The number of accurate predictions (TP+TN) is distributed by the overall number of data groups (P+N) to arrive at the accuracy (ACC). In terms of accuracy, the maximum level is 1.0, and the lowest is 0.0. ERR may also be employed to measure it, as shown by (1).

Detection of the patient with COVID-19 relying on ML technology and FAST algorithms ... (Seba Aziz Sahy)
4.2. Sensitivity (recall or true positive rate)

TP predictions are divided by the total number of positives (P) to compute sensitivity (SN) recall (REC). The equation states that the best sensitivity is 1.0, while the poorest sensitivity is 0.0 (2).

\[
SN = \frac{TP}{TP + FN} = \frac{T.P.}{P}
\]

(2)

4.3. Specificity (true negative rate)

The term (true negative rate) is used to represent the method's specificity, which is defined as the ratio of the total number of negative predictions (N) to the number of TN predictions rue negative rate (TNR). Specificity is expressed in terms of one-to-one (0.0 to 1.0) (3). Table 2 shows the evaluation of COVID-19.

\[
S.N. = \frac{T.N.}{T.N. + F.P.} = \frac{T.N.}{N}
\]

(3)

5. CONCLUSION

In this paper, detection, and recognition of COVID-19 in the scene is based on a FAST algorithm, improving COVID-19 detection performance by selecting the most robust descriptor features. Our proposed COVID-19 detection and identification in CT image and calculating the corresponding result data set with three thresholds and tiny scales, COVID-19 can distinguish objects in a scene changing in terms of rotation and partial blockade, orientation, and illumination, among other things. FAST'S compatibility use makes the method strong compared to the previously confusing graph. The algorithm was found to achieve high accuracy using very few training examples. This shows how strongly we represent the carriers for features. Also, choosing lower thresholds while filtering FAST points and distinctive network carriers in the form enhances the overall algorithm's power for a spin and partial blockade changes.

REFERENCES


Detection of the patient with COVID-19 relying on ML technology and FAST algorithms ... (Seba Aziz Sahy)
BIOGRAPHIES OF AUTHORS

Seba Aziz Sahy received her bachelor’s degree Department of computer science from Ministry of Higher Education & scientific Research-University of Technology–Baghdad. Master’s degree of computer Science, from Ministry of Higher Education & scientific Research–Iraqi Commission for Computers and Informatics–Informatics Institute for Postgraduate Studies. Her research interest is network security, computer networking, software agent, and artificial intelligence. She can be contacted at email: saba.aziz@mtu.edu.iq.

Sura Hameed Mahdi nationality: Iraq, gender: female. Received her Bachelor Degree Department of computer science from Ministry of Higher Education & scientific Research-University of Technology–Baghdad. Master degree of computer Science, from Ministry of Higher Education and scientific Research–Iraqi-University of Technology–Baghdad. Her research interest is image processing and network security. She can be contacted at email: sura-hameed@mtu.edu.iq; surah.hameed@gmail.com; sura-hameed@mtu.edu.iq.

Hassan Muwafaq Gheni received his Bachelor (B.Sc) of Electrical and Electronic Engineering from Department of Electrical Engineering, Babylon university-Iraq-hilla in June 2016. In February 2018, he entered the master’s program at the Faculty of Electrical and Electronic Engineer, Universiti Tun Hussein Malaysia. He is a lecturer at Al-Mustaqbal University College/Department of Computer Techniques Engineering. His research interest is optical communication, IoT, wireless sensor network, communications, V2V system, and artificial intelligent. He can be contacted at email: hasan.muwafaq@mustaqbal-college.edu.iq.

Israa Al-Barazanchi received her Bachelor of Computer Science (BCS) from Department of Computer Science, Baghdad College of Economic Science University-Iraq-Baghdad in June 2002. In January 2010, she entered the master’s program at the Faculty of Information and Communication Technology, Graduate School of Computer Science (Internetworking Technology), Universiti Teknikal Malaysia. She is currently a student Doctor of Philosophy in Information and Communication Technology. She is a lecturer in Department of Computer Engineering Techniques. Editor-in-chief for international union of Universities journal. Member of Editor board in many Scopus journals for computer science field. Head of researcher group. Member in many conferences panel and communications events. She is a reviewer of various high impact factor journals. Her research activities are: WBAN, WSN, WiMAX, WiFi on vehicular Ad-Hoc networks (VANETs), communications, networking, signal processing, IoT, smart systems, and blockchain. She can be contacted at email: israa.albarazanchi@baghdadcollege.edu.iq; israa44444@gmail.com.