PROVIDING AN EFFICIENT COVID-19 PREDICTION MODEL TO IMPROVE CLASSIFICATION USING DEEP LEARNING AND MACHINE LEARNING CLASSIFIER.

Chander Deep Singh[1*], Dr. Dinesh Kumar Garg[2], Vishali Sharma[3]

Abstract:
An outbreak of unidentified infections known as COVID-19 primarily affected the circulatory tract. The illness spreads throughout the entire world and eventually gets worse significantly declining population. The well developed & perfected Covid-19 prediction method is still debatable. In this article, we suggested using affordable X-ray scans to diagnose COVID-19 patients. The majority of healthcare facilities have access to X-ray imaging contrasted to other imaging techniques. To use machine learning and deep learning models CNN with the goal of analyzing its regular exponential behaviour as well as making predictions about the COVID-2019's potential reach across countries by using real-time data. On the premise of accuracy, class specifications TP rate, FP rate, precision, recall, as well as F-measure, the suggested CNN is validated to other existing methods Naive Bayes, Support Vector Machine, Random Forest, as well as decision tree. The suggested method's accuracy is 99.07 percent.

Keywords: Covid-19 Prediction, Deep Learning, Machine Learning, Covid-19 Testing, CNN.

[1*]Research Scholar, Department of Computer Science and Engineering, Sri Sai College of Engineering & Technology, Badhani, Pathankot, Punjab (India)
[2]Associate Professor and Head, Department of Computer Science and Engineering, Sri Sai College of Engineering & Technology, Badhani, Pathankot, Punjab (India)
[3]Assistant Professor, Department of Computer Science and Engineering, Sri Sai College of Engineering & Technology, Badhani, Pathankot, Punjab (India)

***Corresponding Author:** Chander Deep Singh
[*]Research Scholar, Department of Computer Science and Engineering, Sri Sai College of Engineering & Technology, Badhani, Pathankot, Punjab (India)

DOI: 10.48047/ecb/2023.12.si5a.0341

I.INTRODUCTION

Earlier times were mostly marked by epidemics. Following a number of investigations, the first case of the Corona virus was found in 1960. Up until 2003, there were 500 recorded instances, which were all handled as the common flu. Many researchers believe the corona virus originated in bats, but it is uncertain how it was transmitted to humans. Hubei Province received notification from the first case in China on November 17, 2019. The virus began to spread slowly to large populations before spreading to the entire world [1].

There are approximately six different varieties of corona viruses, to be more precise. Severe Acute Respiratory Syndrome (SARS), one of six forms of respiratory illnesses that have been identified in humans to date, is well-known for raising body temperatures in its victims and frequently ending in death from rapid or progressive failure of the respiratory and gastrointestinal systems. The virus has a negative impact on the planet. More than 6.2 million people had been exposed to this virus as of June 5. With more than 105,000 deaths, the USA has the highest rate in all of America. Europe too took a huge impact. In the middle of June, Spain had the most patients. More than 6.2 million people had been exposed to this virus as of June 5. With more than 105,000 deaths, the USA has the highest rate in all of America. Europe too took a huge impact. In the middle of June, Spain had the most patients. The new wave of the Corona virus began by mostly affecting India in the Sub-Continent in western America by the middle of September and into October. Inaccurate viral forecasts had reduced the lockdown's effectiveness. Lack of widespread testing & delayed results in inaccurate forecasts[2]. Additionally, 37% of the results frequently turn out to be inaccurate. Accurately predicting the future day's viral load with precise timings and quickly identifying COVID are essential to stopping the spread of this virus.

On the other hand, DL makes medical optimizations worse. While a lot of user instructions are needed for ML, only a few instructions are needed for DL. First, a novel strategy has been developed to address the predicting problem. The COVID patients' registration forms had been gathered. Timings had already been observed precisely. Most likely, when a patient notices problems, they will come in for testing.

The organization of the paper is arranged as follows. In section II describes the Local Binary pattern. Section III contain the GLCM. Section IV gives the literature Survey. Section V contains the problem formulation and objectives. Section VI contains the results and discussion part. Finally, section VII ends the paper with conclusion followed by references.

II.LOCAL BINARY PATTERN

In a number of applications, such as texture classification and segmentation, image retrieval, or surface inspection, LBP features have excelled. By thresholding the 3-by-3 neighborhood surrounding every pixel with its centre pixel value and employing the result as a binary form, the original LBP operator identifies the pixels of a picture. A LBP computation example is shown in Figure 1.

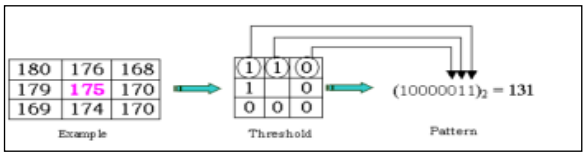

Figure 1. Example of LBP calculation [3]

An picture's tags generated across a 256-bin distribution could be utilized as a texture descriptor. Every histogram bin (LBP code) could be thought of as a little text on. These containers codify many kinds of curved edges, spots, flat surfaces, and other local primitives. Illustrations are shown in Figure 2.

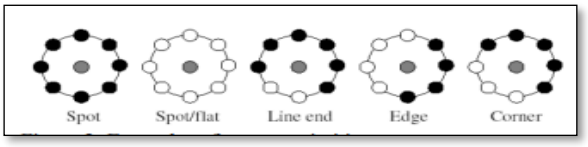

Figure 2. Examples of texture primitives [4]

Varying neighbor sizes are now taken into consideration by the LBP operator. As an illustration, the operators LBP4, 1 employs four neighbors, while LBP16, 2 takes into account sixteen neighbors on a radius of two. A neighborhood dimension of P evenly spaced pixel on a circle of radius R that constitutes a collection of circularly symmetric neighbors is generally referred to by the operators LBPP, R. In accordance with the 2P distinct binary patterns that can be created by the P pixels in the neighbor set, LBPP, R generates 2P distinct correct output. It has been demonstrated that some bins have a high data density than others. As a result, only a portion of the 2P LBPs can be used to represent the textured images. These basic patterns were described by [4]

having a limited amount of bitwise transitions from 0 to 1 and vice versa. For instance, the transition between 00000000 & 11111111 is zero, but the transition between 00000110 and 01111110 is two, and so on. An LBP description is produced by grouping patterns with much more than two transitions into a single bin. The toleration for monotonic illumination changes and computation efficiency of LBP characteristics are its key characteristics.

III. Gray-Level Co-Occurrence Matrix (GLCM)

The technique for analyzing textures has advanced significantly in recent years and could be divided into four groups: statistical, geometrical, model-based, and signal analysis. Despite various statistical techniques, GLCM has been demonstrated to be one of the most successful methods to texture analysis[5]. The spatial distribution of an image's grey value and the frequency of grey occurrence pattern at particular angles and distances are both described by GLCM. GLCM can be used for texture categorization and identification. Haralick[4] developed the GLCM for texture analysis. N is the grayscale significant limitations, and N is the length of the GLCM matrix. The values of the matrix elements on (i,j) are each determined using the source grayscale image's numerical pattern, which must include a specific pattern. The supplied pattern provides grey pixels I & j next to each other at angle & length d. The variable d denotes the separation between two adjacent grayscale pixels, while the variable denotes the length of the discrete angle (00, 450, 90 0, 1350). the usual angle value for GLCM.

IV.LITERATURE SURVEY

Tahir et al.,(2021) features two layers. First Dimension: Using the registration slips from COVID-19 patients, a new technique of prediction has been devised. The native data set of COVID-19 patient registration slips has been used with ResNet-101. 5003 E-registration slips with precise timings are included in the database. In terms of timing, forecasting accuracy was 82%. It was decided how many COVID-19 positive instances would occur the following day. A framework for errors was also devised, using the juxtaposition of MOE and MAE. The COVID-19 chest X-ray prediction is part of the second Dimension. Chest X-ray data from 8009 was gathered locally. Three neural networks—Faster R-CNN, Mask-CNN, or ResNet-50—were suggested to be at fault. The best accuracy is displayed by faster R-CNN at 87%. Mask RCNN's accuracy came up at 83%, whereas resNet-50's was 72%. ACC, PRC, or RCL were used as performance criteria. The addition of the batch normalizing method enhanced SVM performance.

Ayesh et al., (2022) centered on using DL techniques to distinguish between COVID-19 pneumonia, normal pneumonia, or pneumonia without COVID from chest X-rays. To determine COVID-19 using Posteroanterior (PA) and Anteroposterior (AP) view Chest X-rays, we created two DL models. 300 chest X-rays were divided into two datasets for the PA and AP perspectives. A brand-new CNN was created as the initial DL model. Then, transfer learning models based on VGG16, VGG19, and ResNet50 were applied. Eventually, by adding additional layers on top of the current model, the transfer learning models were expanded. Using PA view X-rays, we were able to achieve 98% overall accuracy for the first phase of this investigation, as well as 98% precision, 99% recall, and 99% f1-score for the COVID-19 class. Using AP view X-rays, we achieved 79% overall accuracy, 96% precision, 83% recall, & 89% f1-score for the COVID-19 class in the second section. The areas that assisted the suggested expanded VGG19 model in recognizing COVID-19 were then shown using gradient-based class activation maps. This study demonstrated that extended transfer learning models were able to identify COVID-19 with excellent performance. The proposed extended ResNet50 model outperformed the competition in AP view X-rays whereas the suggested extended VGG-19 models excelled in PA view X-rays.

Nour et al.,(2021) From patients' chest X-ray pictures, DL algorithms have been utilised to identify COVID-19 sick patients (CXR). ResNet, GoogleNet, or AlexNet, three well-known CNNs, were used to remove features from the database. Additionally, the principle component analysis (PCA) method was applied, which further decreased the dataset's dimensionality. To detect COVID-19 disease from the photos, the collected features were fed into classifiers like SVM and KNN. With KNN & SVM, the suggested technique had an accuracy level of 97.7% & 98.1%, respectively. The model's sensitivity and specificity were calculated to be 97% and 98%, correspondingly, demonstrating the model's effectiveness in accurately diagnosing the condition.

Rohan et al.,(2022) establish how COVID19 chest X-ray or CT pictures are quickly or accurately identified by ML, image processing, image segmentation, or feature extraction. In this article, we evaluated the effectiveness of ML algorithms on chest X-ray images or CT scans to COVID-19 diagnosis. The effectiveness of the model was evaluated using pertinent classification metrics, including accuracy, precision, recall, and F1 score. The system has a training accuracy of 100% for COVID-19 patient identification from CXRs. We think that this Computer-Aided Diagnosis (CAD) method, which is very accurate & relatively quick, could be very helpful in containing the epidemic.

Samiul et al.,(2021) examined how well COVID-19 diagnosis using SSL from chest X-ray pictures worked. In order to train our SSL algorithm in the teacher/student typology, we looked at a prepossessing method that involved extracting and combining local phases picture features into multi-feature images. Our research has demonstrated that the SSL method achieved 93.45% accuracy while using 17.0% of the total dataset for training. We also offer metrics that contrast the SSL method to other fully supervised methods.

Duaa F. Eljamassi et al.,(2020) provided a classification approach that uses the chest X-ray pictures to identify the diseased state. Images from chest x-rays of healthy individuals, patients with pneumonia caused by SARS, streptococci, or pneumococcus, as well as additional COVID-19 patients, were gathered in a database. The extraction of visual features is performed using a histogram of oriented gradients (HOG). SVM, RF, KNN are then used to classify the photos, with classification rates of 98.14%, 96.29%, & 88.89%, etc. These findings might be useful in accurately diagnosing COVID-19 disease.

V.PROPOSED WORK

The year 2020 will certainly be remembered in human history as the year in which humans faced a global pandemic that drastically affected every living soul on planet earth. The COVID-19 pandemic certainly had a massive impact on human's social and daily lives. The economy and relations of all countries were also radically impacted. Due to such unexpected situations, healthcare systems either collapsed or failed under colossal pressure to cope with the overwhelming numbers of patients arriving at emergency rooms and intensive care units. The COVID -19 tests used for diagnosis were expensive, slow, and gave indecisive results. Unfortunately, such a hindered diagnosis of the infection prevented abrupt isolation of the infected people which, in turn, caused the rapid spread of the virus. As per the literature review, author proposed the use of cost-effective X-ray images in diagnosing COVID-19 patients. Compared to other imaging modalities, X-ray imaging is available in most healthcare units. Deep learning was used for feature extraction and classification by implementing a multi-stream convolutional neural network model. The model extracts and concatenates features from its three inputs, namely; grayscale, local binary patterns, and histograms of oriented gradients images. Major drawback of traditional detection system is that features were extracted by using the Histogram of oriented gradients (HOG), which no doubt is providing good results. However, HOG has a major disadvantage that they are highly sensitive to image rotations. Due to this reason, HOG feature extraction is not recommended for classification phases where images rotate.

Objectives

1. To propose a Hybrid LBP and GLCM feature extraction technique for enhance pre-processing.
2. To implement an effective Machine Learning and deep learning method to improve classification.
3. To perform an analysis and comparison of the proposed scheme

VI.RESULTS

MALTAB is used because it has been highly validated simulation software by the networking research community.

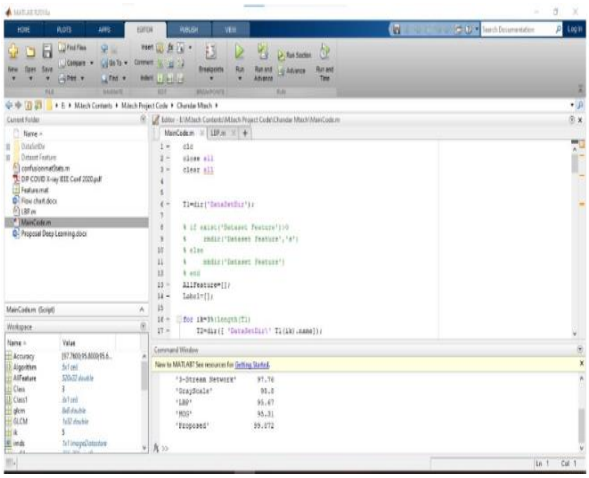

Figure 3: MATLAB

In figure 4, We use a original image of chest x-ray for the research.

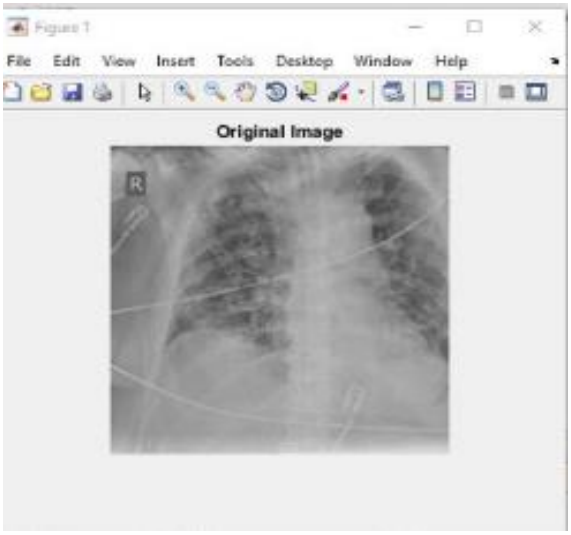

Figure 4: Original Image

The Local Binary Pattern (LBP) texturing operator labels every pixel in an image by thresholding its immediate surroundings and treating the result as a binary number. The LBP texture operator has gained popularity as a strategy in many areas due to its discriminative power & computational simplicity.

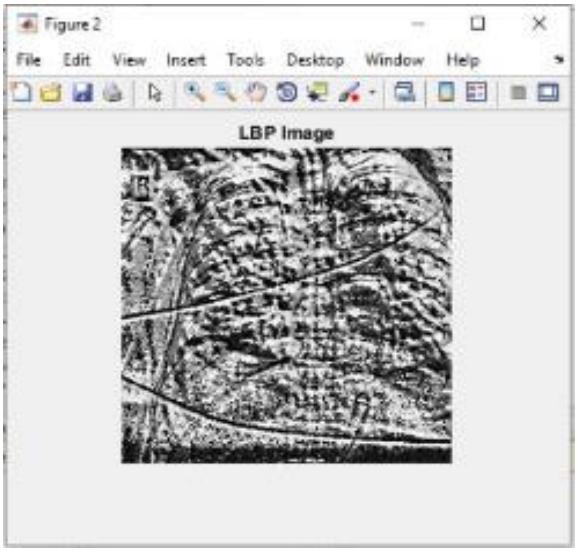

Figure 5: LBP Image

Accuracy: The most popular measure for classification performance is accuracy. Accuracy is calculated by dividing the set of correctly classified instances by the total number of samples, whereas error rate is calculated by dividing the set of incorrectly classified instances by the set of correctly classified instances.

Accuracy = (True Positive + True Negative)/(True Positive + False Positive + True Negative + False Negative)

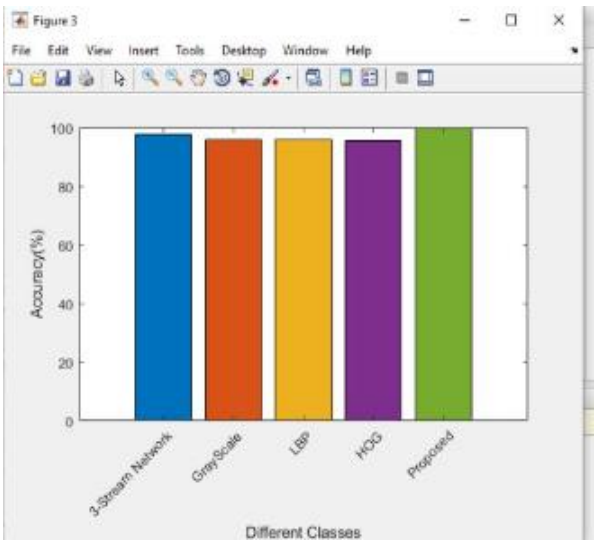

Figure 6: Accuracy

It is clear from figure 6 that the proposed technique has a higher accuracy 99.07 % as compared to other existing techniques.

Precision & recall are two metrics commonly used to evaluate performance in text classification as well as other text mining fields such as information retrieval. These variables are utilized to determine preciseness as well as correctness.

Precision = (True Positive)/(True Positive +False Positive)
Recall = (True Positive)/(True Positive + False Negative)

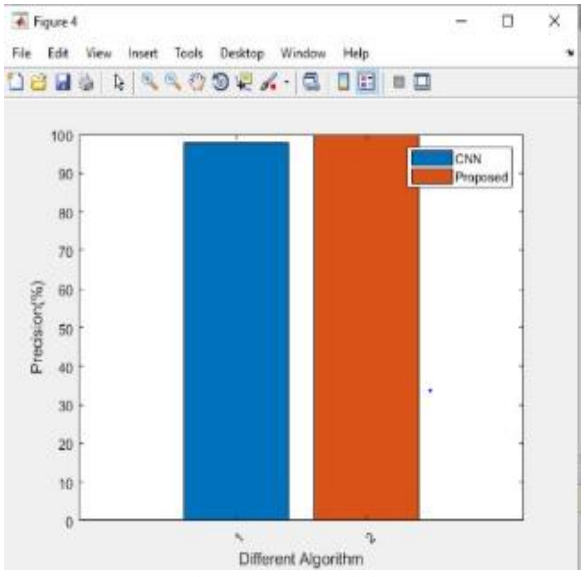

Figure 7:Precision

From figure 7 and 8 shows the proposed approach has precision 100 % and highest recall as compare to the other technique.

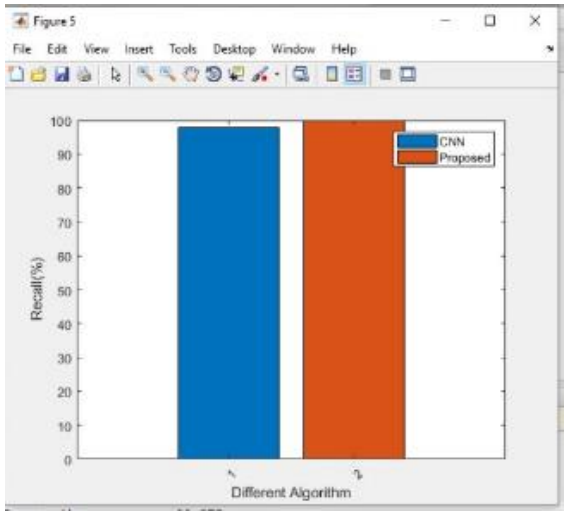

Figure 8:Recall

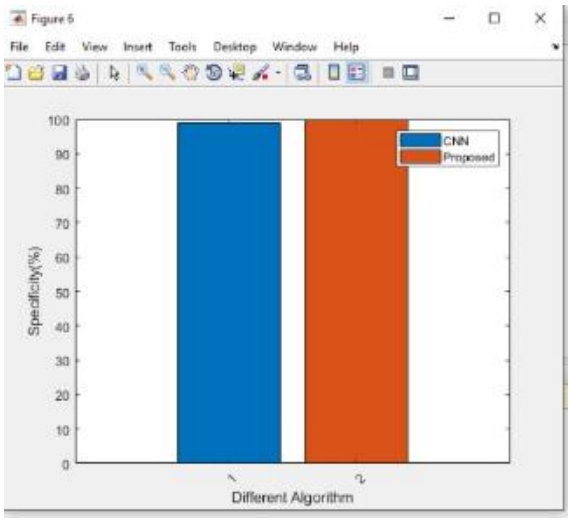

Figure 9: Specificity

From the above results show that the proposed approach as far better as compared to other techniques.

VII.CONCLUSION

A speedy & accurate identification of COVID-19 is made possible by effective SARS-CoV-2 screening, which can also lessen the strain on medical systems. There have been built forecasting model that assess the likelihood of infection by combining a number of parameters. These are intended to help medical professionals worldwide prioritize patients, particularly when there are little healthcare resources available.

This result paper proposes a deep learning and machine-learning method for determining a positive or negative Covid infection. On the premise of accuracy, class specifications TP rate, FP rate, precision, recall, as well as F-measure, the suggested CNN is validated to other existing methods Naive Bayes, Support Vector Machine, Random Forest, as well as decision tree. The suggested method's accuracy is 99.07 percent. The findings obtained demonstrate that the suggested approach outperforms the other methodologies across all performance variables.

For future work, more optimization technologies will be integrated to decrease the time complexity in addition to reducing the classifier construction time. Certain ensemble methods, like Adaboost, Bagging, as well as Stacking, could also be used to improve the outcomes.

REFERENCES

1. Zoabi, Y., Deri-Rozov, S., & Shomron, N. (2021). Machine learning-based prediction of COVID-19 diagnosis based on symptoms. Npj Digital Medicine, Vol. 4, No. 1,2021.
2. Gozes, O. et al. Rapid AI development cycle for the coronavirus (COVID-19) pandemic: initial results for automated detection & patient monitoring using deep learning CT image analysis. arXiv e-prints, 2020.
3. Lahdenoja, O., Poikonen, J., & Laiho, M., " Towards understanding the formation of uniform local binary patterns.", ISRN Machine Vision, 2013.
4. Hu, Y., Zhao, C. X., & Wang, H. N., "Directional analysis of texture images using gray level co-occurrence matrix", In Computational Intelligence and Industrial Application, PACIIA'08. Pacific-Asia Workshop on (Vol. 2, pp. 277-281). IEEE,2008.
5. Tahir, H., Iftikhar, A., & Mumraiz, M. (2021). Forecasting COVID-19 via Registration Slips of Patients using ResNet-101 and Performance Analysis and Comparison of Prediction for COVID-19 using Faster R-CNN, Mask R-CNN, and ResNet-50. International Conference on Advances in Electrical, Computing, Communication and Sustainable Technologies (ICAECT),2021.
6. Ayesh Meepaganithage; Mgnas Fernando, "Detecting COVID-19 Pneumonia using Chest X-rays through Deep Learning Techniques", 2nd International Conference on Advanced Research in Computing (ICARC),2022.
7. Nour Haj Hammadah; Nilima R. Das; Mamata Nayak; Tripti Swarnkar, "A hybrid approach of Deep Learning Algorithms for Identification of COVID-19 disease using Chest X-Ray Images", International Conference in Advances in Power, Signal, and Information Technology (APSIT),2021.
8. Rohan Tomar, Abhilasha Sharma, "Analysis & Detection of COVID-19 on Chest X- Ray

Images based on Support Vector Machines", International Conference on Sustainable Computing and Data Communication Systems (ICSCDS),2022.

9. Samiul Haque; Mohammad Akidul Hoque; Mohammad Ariful Islam Khan; Sabbir Ahmed, "COVID-19 Detection Using Feature Extraction and Semi-Supervised Learning from Chest X-ray Images", IEEE Region 10 Symposium (TENSYMP),2021.
10. Duaa F. Eljamassi; Ashraf Yunis Maghari, "COVID-19 Detection from Chest X-ray Scans using Machine Learning", 2020 International Conference on Promising Electronic Technologies (ICPET),2020.