International Journal of Multidisciplinary Innovation and Research Methodology (IJMIRM)
Volume 1, Issue 1, October-December, 2022, Available online at: https://ijmirm.com

Comparative Analysis of Accuracy and Loss during Cataract Detection using Deep Learning

Sing Su

Associate Professor, Dept. of Computer Applications, The University of Hong Kong

ABSTRACT

Cataracts are a type of eye condition that affects people of all ages, although they are most prevalent in people who are between the ages of 40 and 50. Cataracts are one of the most common disorders that affect humans. If not treated, this condition might result in blindness. Early detection is key to preventing this sort of eye illness from progressing to a more serious stage. Therefore, we are going to present our model that makes use of Deep Learning to analyse Fundus Images in order to diagnose this condition. In comparison to other models, this one has much lower computational costs thanks to the inclusion of fully linked hidden layers that are equipped with cost functions and activation functions. Proper training and optimisation of the model are also possible thanks to these features. The model is given a total of 1130 photos that have been altered and 1130 images that have not been affected in order to ensure that it is trained correctly and does not become overfit.

Keywords: Cataract Detection, Deep Learning, Neural Network, and Fundus Images.

INTRODUCTION

A cataract is a disorder in which the lens of the eye becomes clouded, resulting in symptoms such as impaired vision, sensitivity to glare, and poor night vision. Cataracts are the most common cause of vision loss in older adults. It is a substantial contributor to the global problem of impaired vision and blindness. In its early stages, it does not have a significant impact, but if left untreated, it can eventually cause blindness in the affected eye. If it is diagnosed in its early stages, there is a chance that the patient will not require surgery and will not go blind. The World Health Organisation estimates that there are 285 million people living with some form of visual impairment around the globe. There are around 39 million people who have eyesight that is poor to extremely limited, and the remaining people are visually impaired. To put it another way, cataracts are responsible for 33 percent of cases of vision impairment and 51 percent of blindness cases.

It is projected that by the year 2025, the number of people who are blind would exceed 40 million. In addition, there has been a sharp increase in the quantity of cataract operations of every kind recently. According to the findings of several research, the proportion of female patients to male patients is higher. This addresses cataract surgery as well as cataracts in the nuclear and cortical layers ($p = 0{:}02\text{-}0{:}05$). In addition, the non-white community has a higher prevalence of it ($p = 0.001$) than the white community does. The cataracts are responsible for the majority of the instances. It is considered to be one of the most significant factors in causing blindness. Posterior sub capsular (PSC) cataract, nuclear cataract, and cortical cataract are the three primary subtypes of cataract that can be distinguished based on the location and mode of development of the condition. Early detection and treatment of cataracts can result in a significant reduction in the chance of developing cataract-related blindness. It is difficult to develop an automatic system for the identification of cataracts due to the fact that each human has a unique eye in terms of its size, shape, and the manner in which cataracts develop in their eyes.

Additionally, the fact that cataract development is dependent on factors such as age, gender, and eye type further complicates the matter. In recent years, researchers have focused their attention on automatic cataract identification through the use of a variety of imaging techniques. Mostly cataract detection model employs i) Slit lamp ii) retro-illumination iii) ultrasonography iv) Automatic cataract diagnosis and classification systems normally make use of four different kinds of images, one of which is called a fundus picture. Because fundus cameras are so straightforward to operate, fundus photography has emerged as a front-runner among these several imaging modalities as the modality that has garnered the most attention in this particular field. In order to perform the Slit-Lamp test, a trained ophthalmologist is required.

Therefore, it is essential to have an automated technique of diagnosing cataracts based on fundus photographs in order to streamline the process of early cataract screening. Despite the fact that various automatic cataract detection systems based on deep learning have been reported in research works, these systems continue to exhibit disadvantages such as inadequate detection accuracy, an excessive number of model-parameters, and large computing expenses. Our research addresses the issues that were raised earlier by classifying individuals into one of two categories, based on whether they have cataracts or other conditions. The proposed solution includes a number of novel components, the first of which is a reduction in the model's parameters, including its layers and weights, in order to improve the computational efficiency and reduce the costs. The second objective is to improve the reliability of the detection process by employing a cutting-edge structure known as a "deep neural network." As a consequence of this, the method that has been described enables the efficient screening of a wide population as well as the correct grading of cataracts.

The following are the most important contributions made by the article:

- The datasets used in this study include the HRF[14] (high resolution fundus) dataset , the fundus image registration (FIRE)[15] dataset , the ACHIKO-I fundus image dataset, the Indian diabetic retinopathy image (IDRiD) dataset, a colour fundus image database, and the digital retinal images for vessel extraction (DRIVE) [11] database. are all combined, reorganised, and preprocessed to create a cataract dataset. After that, a method of data augmentation is used to enlarge it to include a considerable number of photographs.
- In this research, a unique architecture for a deep learning neural network with 16 layers is presented for the purpose of cataract diagnosis. The purpose of the proposed network is to determine with high precision whether or not cataracts are present.
- In order to evaluate and demonstrate the usefulness of our proposed model, a total of five different C-NN models, including VGG16, VGG19, and Res Net-50, have been utilised.

The following categories make up the remaining sections of this paper:

- In Section II, a brief discussion of the pertinent recent works is presented.
- In Section III, a description of the suggested Model for cataract detection may be found.
- The experimental apparatus is discussed in Section IV.
- The results of the experiments are broken down and discussed in Section V of the report.
- In the final section of this study, which is part VI, we will discuss potential directions for further work.

RELATED WORKS

The technique of automatically detecting cataracts involves three steps: the extraction of features, the pre-processing of those features, and the classification of the results. Methods that are machine learning (ML)-based and methods that are deep learning (DL)-based are the two categories that can be created depending on the algorithms that are utilised throughout the stages of feature extraction or classification, respectively. Recent research [9]-[12] has discussed these methods in detail. In this part, we take a cursory look at a handful of the most significant works produced by both of these groups.

A. Previous Projects that Involved Machine Learning

There are now thirteen different approaches that have been proposed for the purpose of identifying cataracts. These approaches can either be used for general screening or as an intermediate step before cataracts are classified. In this study, the primary emphasis was placed on training the linear discriminant analysis (LDA) algorithm by making use of an enhanced texture feature. The findings of a clinical database investigation demonstrated an accuracy of 84.8 percent. Yang et al. devised a method for the automated detection of cataracts that consisted of three steps. The contrast between the foreground and background was amped up with the help of a top-to-bottom hat transition. It was considered that the luminance and the texture were the features. A back propagation neural network (BBNN) classifier was developed in order to categorise the cataracts according to the degree of their severity. Guo et al. developed an automatic classification of cataracts that was based on fundus photographs.

In order to complete the process of feature extraction, the wavelet transform and sketch-based techniques were utilised. Following that, a multiclass discriminant analysis method was utilised for the purpose of cataract diagnosis and classification. The results of the wavelet transform-based feature extraction were found to be 90.901% and 77.10%, respectively, while the results of the sketch-based feature extraction were found to be 86.0% and 74.0%, respectively. Fuadah et al[26]. "used the K-Nearest Neighbour (KNN) in conjunction with the dissimilarity", "used the K-Nearest Neighbour" in comparison, and "used the K-Nearest Neighbour" in uniformity. This technology was designed for use in mobile phones and boasts an accuracy rate of 97.5 percent.[8] Jagadale was able to determine the lens's centre and radius with the use of the Hough circle detection transform. After that, the statistical features were recovered and input into an SVM classifier so that an accurate identification could be made, which achieved a score of 90.25 percent. Sigit et al. [25] presented a method that is based on Android smartphones. For the purpose of classifying the data, a single-layer perceptron technique was utilised, and the accuracy achieved was 85.0%. In the recent past, an approach for cataract grading that is based on the hierarchical extraction of features was proposed in [24]. The original categorization of the cataract severity grading problem, which consisted of four classes, was divided into three adjacent classes consisting of two classes each. First, three distinct neural networks were utilised in order to merge these. This method achieved an accuracy of 94.83% when identifying cataracts and 85.98% when rating the severity of cataracts.

B. Different Models Using Deep Learning

Deep learning algorithms can be used to learn fundamental traits and include them into the model creation process in order to solve the limits of manually produced attributes and their application across a variety of medical imaging modalities. These constraints include the fact that manually designed attributes are limited in their applicability. Gao et al. [5] examined a deep learning-based technique for evaluating the severity of nuclear cataracts by analysing images obtained from slit-lamp examinations. They conducted an analysis of patches inside the input images using a convolutional neural network (CNN), which allowed them to derive local filters. In addition, a collection of recursive neural networks, also known as RNNs, was used in order to extract higher-order features. In the end, the grading of cataracts was performed using support vector regression, which made use of the previously taught features. The purpose of this comprehensive method was to improve both the accuracy and efficiency of the cataract examination. Zhang [20] designed the Deep CNN for the purpose of identifying and grading cataracts.

This network made use of the feature maps that were derived from the architecture's pooling layers. This method was speedy, and it had accuracy rates of 93.52% and 86.69%, respectively, for identifying cataracts and evaluating their severity. Ran et al. [23] presented a method for a six-tier cataract classifier that makes use of a combination of deep convolutional neural networks and random forests. The proposed deep convolutional neural network (DCNN) has three different modules that worked together to extract characteristics at many layers from fundus images. On the other hand, RF implemented a more complicated six-level cataract grading by making use of a feature dataset that DCNN had developed and was employing. The accuracy of this method is approximately 90.69% on average. Utilising this six-level grading system may facilitate improved comprehension of the patients' situations on the part of the medical experts. Pratap-Kokil [22] developed a computer-assisted method for classifying cataract severity ranging from mild to severe. The method was based on photographs of the fundus.

This method of automatically classifying cataracts used transfer learning in conjunction with a CNN that has previously been trained. The final classification was accomplished through the utilisation of feature extraction as well as a Support Vector Machine classifier that had a 4-stage CCR that was 92.91%. Jun et al. [40] proposed the use of a tournament-based ranking CNN system to score cataracts. This system would be comprised of binary CNN models and a tournament framework. Hossain [21] proposed a method for the automated diagnosis of cataracts that has an accuracy rate of 95.77 percent. DCNNs and a Residual NN classification model that has been optimised are utilised. Zhang's most recent research may be found at [12].

An attention-based Multi Model Ensemble method for automatically diagnosing cataracts on ultrasound photographs was found to have "the greatest accuracy (97.5%)" among the numerous deep learning-based methods that were published in the research papers. This was the case when compared to other methods. In this particular approach, the system as a whole was composed of a model ensemble module, three different grading networks, and an object detection network. The performance of this method was nevertheless remarkable, even taking into account the constraints imposed by the poor quality of the training data.

However, one of its major drawbacks was that it relied on evaluating the degree of blurriness in retinal images, which might possibly signal the existence of a variety of eye illnesses, such as cataracts, corneal edoema, and diabetes mellitus. This was a significant limitation. As a consequence of this, it is possible that this method will have difficulties distinguishing between the many kinds of eye diseases. Khan and colleagues have just lately achieved a level of accuracy that is almost equal (97.47%).

Different writers have developed and published a variety of models, each of which varies in terms of how accurately it represents reality. This research is innovative since the method that was recommended produces better results than other studies that have been done in the past.

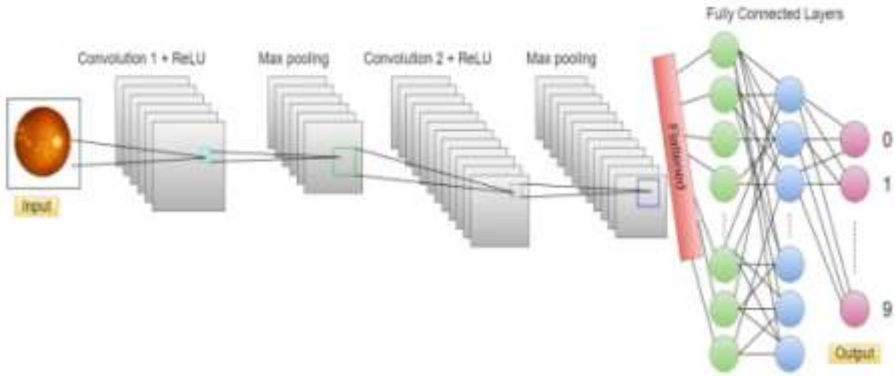

FIg 1: This is a architectural view of proposed model

PROPOSED ARCHITECTURE

A particular kind of ANN called as the convolutional neural network, also known as CNN, is a type of artificial neural network that is designed to handle input that has a grid-like structure, such as photos. Computer vision applications such as image identification commonly make use of CNNs as an important component.

Convolutional neural networks function by convolving the data that is fed into them using a variety of different algorithms. The term "convolution" refers to the process of altering the forms of two images. In the case of CNNs, the two functions that are important are the input data and a filter. The filter makes use of a very small number matrix in order to extract features from the data that is being input.

CNNs have the capability to learn features from images in a methodical and organized fashion. It is likely that they are able to do so given that they can acquire low-level features such as edges and corners and then use those features to learn higher-level features such as objects and faces. Because CNNs make use of hierarchical learning, they are exceptionally accurate when it comes to carrying out image identification tasks. When using methods that are based on deep learning, the characteristics of the images are retrieved and incorporated into the classification stages, however when using manual feature extraction approaches, these characteristics are separated. Our model that makes use of deep learning is provided in an effort to mitigate the limitations of manually extracting features and to reduce the amount of computational effort required.

The architecture of our proposed model is depicted in Figure 1, and it consists of SIXTEEN layers. FIFTY PERCENT of these layers are located in the first four blocks, and the other layers are used for grading. RGB pictures (224 by 224 pixels) and 32 filters with Ks (3 by 3 pixels) make up the first block's inputs.

The maximum number of pooling layers, with a stride of two, has been implemented. In order to make more efficient use of space. Blocks equipped with ReLu activation functions can be used. It is the same block that is used for the third block, except this time there are 64 filters rather than the 128 filters that are in the fourth block.

The output from all of these are sent to the remaining layers, which include flatten dense and drop out layers. All of these layers are fully connected to one another. Because this is a model for binary classification, we will be utilising a Sigmoid Function in this section. The formula is as follows:

$$\sigma(x) = 1 / 1 + e^{(-x)}$$

To investigate how the block numbers affect classification accuracy ,so cataract detection with three alternative models built on 3, 4, and 5 blocks with three distinct sets of parameters were utilised.

are created and assessed on the dataset, namely (16, 32, 64), (32, 32, 64, 128), and (32, 32, 64, 96, 128), respectively. The level of accuracy attained using these models is shown in Table 2. The efficiency of the model is enhanced by using four blocks instead of three. However, at 5-blocks, this efficiency was diminished. The four block(32,32,64,128) filters model did better than the rest.

No. of Blocks	Info about Filters	Accuracy%
3	**16,32,64**	**97.80%**
4	**32,32,64,128**	**99.13%**
5	**32,32,64,96,128**	**97.12%**

Fig 2: Shows the Efficiency achieved from these models

To know more and better about the model we should look into deeper aspects and specifically the total no. of blocks and cnn layer. We go deeper and look in to the effects of layers n the model as shown by Fig 2 which clearly shows that at layers 4 the models accuracy goes higher and at 5 it again start falling down.

The first layers extracts patterns such as corners, dots, edges...etc then these patterns are combined together and more expensive extraction of patterns takes places like circles and squares..etc. So adding more layers can increase the accuracy but not every time, this is shown in table above. Most used loss function in binary classification is log **loss function** = **-(y(log(py)) + (1-y) log(1-py)**, where y is real value (0 or 1) as per dataset and py is the predicted value and the log used here is normal log.

SETUP

A. Pre Processing

Datasets are collected from various sites and repositories. So here arises the problem, they are not same in sizes and many other characteristics are not identical, like their feature columns..etc. So Images are scaled to a single size (224 × 224 px). An important step here is to normalize the dataset, in order to have proper distribution, this helps in better and faster convergence rate. The formula we used for this is :

Xn = (X - X min) / (X max - X min)

Here, X max and X min :- max and min values of datasets, and Xn and X are of the same range.

B. Data Augmentation

The lack of a big training collection for medical imaging presents a challenging obstacle that impedes future development of deep learning. Therefore, in order to overcome the inadequacy of the dataset, four geometric modifications are applied to the training examples in order to enrich the dataset. These modifications are as follows: rescaling, rotating images (at random angles), zooming, and horizontal flipping. Not only does this provide us with data that had been abandoned, but it also prevents the model from becoming overfit to the data. This delivers as many data as four times the initial dataset. The training and testing ratio is eighty to twenty.

C. Implementation

All of the processes and experiments are carried out on our own computer, which has the following specifications: processor: Ryzen 5; GPU: 4GB NVIDIA GTX GEFORCE 1650 Ti; models are implemented via Python Keras and Tensorflow; learning rate: 0.0001; and ADAM is used to optimise the models. Which ultimately results in a productive and well-optimized output.

D. Evaluation

Accuracy alone is not sufficient when evaluating the usefulness of a model. In addition to that, we need to verify a wide range of other parameters, such as the F1 score recall, precision, and a lot of others.

We are able to view the performance of each model and verify its f1 score and precision, which are two of the most essential factors for determining whether or not a model will be successful.

One of the most important aspects of our model is that it allows us to determine the number of erroneous positive and false negative predictions that it makes. The Confusion Matrix will help us visualise this information.

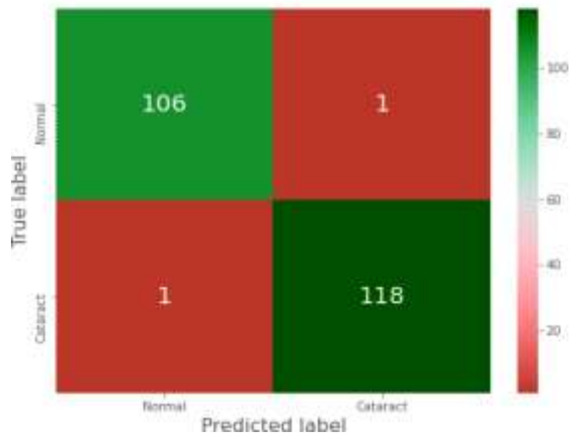

Fig 3: Confusion Matrix

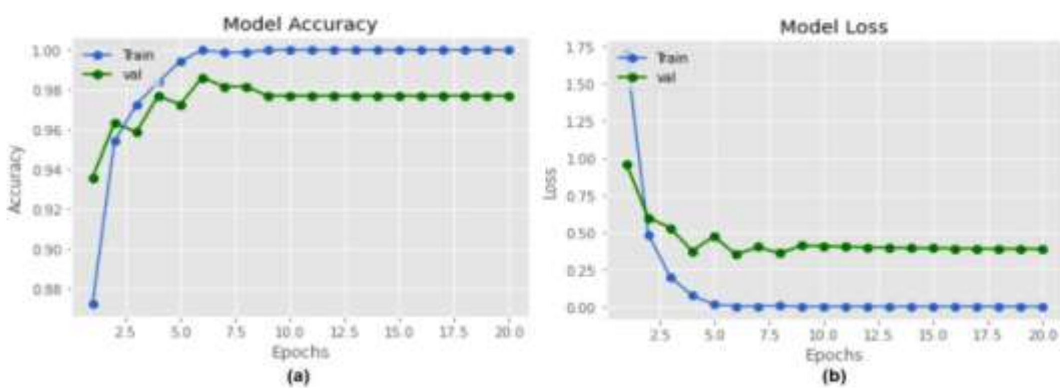

Fig 4: Accuracy and Loss Graph

CONCLUSION

The cataract can be successfully identified from the fundus images that have been provided by our proposed technique, which has a very high rate of accuracy which is also lightweight, meaning that it has a minimal cost in terms of computing and can be implemented more quickly. We enhanced the datasets so that they had more information in order to improve the training of the model.

Additionally, we experimented with the trial-and-error method for a variety of layers, activation functions, and loss functions and we utilised a great deal of optimisation algorithms in order to make our model quick and lightweight. We examined how well our model performed in comparison to other models. Nevertheless, our model is incapable of distinguishing between different age groups or pinpointing exactly where the cataract is located. Therefore, more work needs to be done on this.

REFERENCES

[1]. Mayo Clinic Staff, Cataracts

[2]. C. Szegedy, W. Liu, Y. Jia et al., "Going deeper with convolutions," in in 2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 1–9, 2015.

[3]. S. Sharma, Activation functions in neural cao networks, Towards Data Science, 2017

[4]. L. Cao, H. Li, Y. Zhang, L. Zhang, and L. Xu, ''Hierarchical method for cataract grading based on retinal images using improved Haar wavelet,'' Inf. Fusion, vol. 53, pp. 196–208, Jan. 2020.

[5]. X. Gao, D. W. K. Wong, T.-T. Ng, C. Y. L. Cheung, C.-Y. Cheng, and T. Y. Wong, ''Automatic grading of cortical and PSC cataracts using retro illumination lens images,'' in Proc. Asian Conf. Comput. Vis. Berlin, Germany: Springer, 2012, pp. 256–267

[6]. X. Gao, H. Li, J. H. Lim, and T. Y. Wong, ''Computer-aided cataract detection using enhanced texture features on retro-illumination lens images,'' in Proc. 18th IEEE Int. Conf. Image Process., Sep. 2011, pp. 1565–1568.

[7]. L. Guo, J.-J. Yang, L. Peng, J. Li, and Q. Liang, ''A computer aided healthcare system for cataract classification and grading based on fundus image analysis,'' Comput. Ind., vol. 69, pp. 72–80, May 2015.

[8]. A. B. Jagadale, S. S. Sonavane, and D. V. Jadav, ''Computer aided system for early detection of nuclear cataract using circle Hough transform,'' in Proc. 3rd Int. Conf. Trends Electron. Informat. (ICOEI), Apr. 2019, pp. 1009–1012.

[9]. H. Morales-Lopez, I. Cruz-Vega, and J. Rangel-Magdaleno, ''Cataract detection and classification systems using computational intelligence: A survey,'' Arch. Comput. Methods Eng., vol. 28, pp. 1–14, Jun. 2020.

[10]. H. E. Gali, R. Sella, and N. A. Afshari, ''Cataract grading systems: A review of past and present,'' Current Opinion Ophthalmol., vol. 30, no. 1, pp. 13–18, 2019.

[11]. J. Staal, M. D. Abràmoff, M. Niemeijer, M. A. Viergever, and B. van Ginneken, ''Ridge-based vessel segmentation in color images of the retina,'' IEEE Trans. Med. Imag., vol. 23, no. 4, pp. 501–509, Apr. 2004.

[12]. X. Zhang, Y. Hu, J. Fang, Z. Xiao, R. Higashita, and J. Liu, ''Machine learning for cataract classification and grading on ophthalmic imaging modalities: A survey,'' 2020, arXiv:2012.04830. [Online]. Available: http://arxiv.org/abs/2012.04830

[13]. X. Gao, H. Li, J. H. Lim, and T. Y. Wong, ''Computer-aided cataract detection using enhanced texture features on retro-illumination lens images,'' in Proc. 18th IEEE Int. Conf. Image Process., Sep. 2011, pp. 1565–1568.

[14]. A. Budai, R. Bock, A. Maier, J. Hornegger, and G. Michelson, ''Robust vessel segmentation in fundus images,'' Int. J. Biomed. Imag., vol. 2013, pp. 1–11, Dec. 2013.

[15]. Ocular Disease Recognition, Dataset, https://www.kaggle.com/andrewmvd/ocular-disease-recognition-odir5k

[16]. A. G. Howard, M. Zhu, B. Chen, D. Kalenichenko, W. Wang, T. Weyand, M. Andreetto, and H. Adam, ''MobileNets: Efficient convolutional neural networks for mobile vision applications,'' 2017, arXiv:1704.04861. [Online]. Available: http://arxiv.org/abs/1704.04861

[17]. C. Szegedy, V. Vanhoucke, S. Ioffe, J. Shlens, and Z. Wojna, ''Rethinking the inception architecture for computer vision,'' in Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR), Jun. 2016, pp. 2818–2826.

[18]. D. Kim, T. J. Jun, Y. Eom, C. Kim, and D. Kim, ''Tournament based ranking CNN for the cataract grading,'' in Proc. 41st Annu. Int. Conf. IEEE Eng. Med. Biol. Soc. (EMBC), Jul. 2019, pp. 1630–1636

[19]. C. Hernandez-Matas, X. Zabulis, A. Triantafyllou, P. Anyfanti,S. Douma, and A. A. Argyros, ''FIRE: Fundus image registration dataset,'' Model. Artif. Intell. Ophthalmol., vol. 1, no. 4, pp. 16–28,2017.

[20]. L. Zhang, J. Li, i. Zhang, H. Han, B. Liu, J. Yang, and Q. Wang, ''Automatic cataract detection and grading using deep convolutional neural network,'' in Proc. IEEE 14th Int. Conf. Netw., Sens. Control (ICNSC), May 2017, pp. 60–65.

[21]. Tilwani K., Patel A., Parikh H., Thakker D. J., & Dave G. (2022), "Investigation on anti-Corona viral potential of Yarrow tea", Journal of Biomolecular Structure and Dynamics, 1-13.

[22]. Parikh, H. (2021), "Diatom Biosilica as a source of Nanomaterials", International Journal of All Research Education and Scientific Methods (IJARESM), Volume 9, Issue 11

[23]. Parikh, H. (2021), "Algae is an Efficient Source of Biofuel", International Research Journal of Engineering and Technology (IRJET), Volume: 08 Issue: 11.