Automated Rice Analysis for Sugar Content and Dietary Recommendation

Mokshil Deepal Tandel [1]
Shree Swaminarayan International School
Gujarat, India
mokshiltandel18@gmail.com

Reetu Jain [2]
Chief Mentor
On My Own Technology
Mumbai, India
reetu.jain@onmyowntechnology.com

Abstract: *Accurate rice sugar content analysis and dietary advice are essential for establishing good eating practices and avoiding chronic illnesses. However, the traditional approaches for determining the sugar content of rice are time-consuming, labor-intensive, and subject to human error. We created an automated rice analysis system that uses machine learning algorithms to precisely assess the sugar content and give individualized dietary advice in order to solve these issues. Additionally, we included nutritional suggestion algorithms in our system that took into account a person's age, sex, weight, and degree of exercise. Our approach offers consumers individualized advice on their rice intake, encouraging healthier dietary choices. It does this by integrating the anticipated sugar content of rice samples with personalized dietary recommendations. Extensive tests were performed utilizing a sizable dataset of rice samples with known sugar contents to assess the effectiveness of our neural network model. With a 90% average accuracy rate for predicting sugar content, the findings showed the precision and effectiveness of our automated rice analysis system. The time-consuming and prone to mistake character of conventional methods for rice sugar analysis has a practical answer in the automated rice analysis system. This technology has the ability to empower people in making knowledgeable decisions about their rice intake, thereby encouraging better lives and lowering the risk of chronic illnesses. It does this by precisely assessing sugar content and offering personalized dietary advice.*

Keywords: **Rice, Sugar content, Diabetes, Dietary, Insulin, Keras, Convolutional Neural Network**

1. Introduction

Rice is a cereal grain that is the staple food for over half of the world's population, particularly in Asia and Africa. Rice is the most important food crop with regard to human nutrition and caloric intake, providing more than one-fifth of the caloric consumed worldwide by humans. Rice comes in a variety of types, each with its own unique flavor and texture. Some of the most common types of rice are white, brown, jasmine, and basmati. White rice is the most widely consumed variety and is milled to remove the bran and germ, resulting in a mild, fluffy texture. Brown rice, on the other hand, retains the bran and germ layers, giving it a nutty flavor and chewy texture. While basmati rice has a nutty flavor and long slender grains that make it ideal for biryanis and pilafs.

However, rice can be harmful to people with diabetes because it is a high-carbohydrate food that can cause blood sugar levels to spike. Carbohydrates are broken down into glucose, which is then absorbed into the bloodstream, causing blood sugar levels to rise. White rice, in particular, is a highly refined carbohydrate that is quickly absorbed into the bloodstream, causing a rapid rise in blood sugar levels.

Brown rice is a better option for people with diabetes because it contains more fiber, which slows down the absorption of glucose into the bloodstream. However, even with brown rice, portion control is important for people with diabetes. It's important to monitor portion sizes and to balance rice intake with other foods that help stabilize blood sugar levels, such as protein, healthy fats, and non-starchy vegetables. Overall, rice can be a part of a healthy diet for people with diabetes, but it's important to be mindful of portion sizes and to choose the right type of rice to help regulate blood sugar levels. A registered dietitian can help develop a balanced meal plan that includes rice and other healthy foods for people with diabetes.

Due to variations in their sugar composition, fiber content, and processing methods, different varieties of rice, including white rice, brown rice, and parboiled rice, display variable glycemic reactions. In comparison to low GI rice, which takes longer to digest and absorb, rice with a high glycemic index (GI), which is quickly digested and absorbed, might cause a more noticeable rise in postprandial glucose levels. Consuming rice can have a major effect on a diabetic person's blood sugar levels. After ingesting high-GI rice, blood glucose levels rise quickly, making glycemic management difficult and maybe even resulting in hyperglycemia. To help diabetics make informed dietary decisions, it is essential to understand the sugar content of rice and its possible impact on blood glucose. This study aims to develop an automated rice analysis system capable of accurately detecting the sugar content in rice samples. By quantifying the sugar content, the system can provide personalized dietary recommendations to diabetes patients, indicating whether a specific rice variety is suitable for consumption or if alternative options should be considered. By tailoring dietary choices to individual health requirements, this system can contribute to improved glucose management and overall well-being for diabetic individuals.

Blood Glucose Chart			
Mg/DL	Fasting	After Eating	2-3 Hours After Eating
Normal	80-100	170-200	120-140
Impaired Glucose	101-125	190-230	140-160
Diabetic	126+	220-300	200+

Fig 2.Blood Glucose Chart(Image Credits:https://www.lark.com/resources/blood-sugar-chart)

Table 1. The nutrients content of several varieties of 100 g rice (USDA, 2011) (Image Credit: Rohman, Abdul & Helmiyati, S. & Penggalih, Mirza & Setyaningrum, Dwi. (2014). Rice in health and nutrition. International Food Research Journal. 21. 13-24.)

Rice	Glycemic Index (Gulcose=100)	Glycemic load Per Serving
White Rice (Oryza sativa), boiled	69 ± 15	30
White Rice, low amylose, boiled	17	7
White Rice, high amylose, boiled	39	15
Milled white rice, high amylose, boiled	61	26
Brown Rice, boiled	50 ± 19	17
Brown Rice, high amylose, boiled	39	16
Parboiled, low amylose rice	51	19
Parboiled. high amylose rice	32 ± 2	12

Table 2. The carbohydrate content of different branded rice samples (g/100 g)a.(Image Credit:Devindra, Shekappa & Longvah, Thingnganing. (2011). A New Approach for the Measurement of Digestible Carbohydrates in Different Food Samples with HPLC-RI. Journal of Agricultural Science and Technology B 1. 1216-1223.)

Rice	Soluble Starch	Insoluble Starch	Total Starch	Total Sugar	Starches and Sugar
No-64-Rice	12.51 ± 0.09	60.43 ±0.57	72.94±0.67	6.10±0.05	79.04
Kerala Rice	12.82 ±0.09	56.30 ±0.36	69.13±0.45	9.62±0.08	78.75
Natural label	13.27 ±0.06	53.60 ±0.33	68.63±0.39	5.69±0.00	74.32
Broken Rice	13.68 ±0.06	56.35 ±0.31	70.03±0.54	6.06±0.13	76.09
Hansa new	14.58 ±0.18	56.25 ±0.44	70.83±0.63	6.86±0.05	77.69
Suprime	14.98 ±0.12	56.38 ±0.34	71.36±0.46	9.36±0.05	80.75

2. Literature Review

The research conducted by Omar et el. [1] looked at five distinct varieties of rice samples that were consumed by the Kurdish community in the Kurdistan Region. The samples of rice's starch content ranged from 81.23% to 92.73%. The greatest starch percentage was found in sample 5, at 92.73%, while the lowest and most typical starch level was found in sample 1, at 81.23%. Using a phenol-sulfuric acid technique with different hydrolysis periods under constant acid concentration and temperature, the total sugar content of the rice samples was ascertained. The samples' physical characteristics, including length, breadth, thickness, surface area, sphericity, density, porosity, and thousand kernel weight, were also examined. These physical characteristics are essential for building machinery and storage facilities in the food processing sector for effective handling, shipping, and storage. Using information from the international Prospective Urban Rural Epidemiology (PURE) project, the Balaji Bhavadharini et el. study[2] explores the relationship between the consumption of white rice and the risk of getting diabetes. Previous studies on this subject have produced conflicting findings, but most of them were restricted to a small number of nations, mostly in Asia. The researchers examined data from a sizable sample of 132,373 people in 21 nations, ranging in age from 35 to 70, to close this gap. Consumption of white rice was divided into four categories based on cooked grammes: 150 g or less per day, 150 to 300 g/day, 300 to 450 g/day, and 450 g/day.The nutritional makeup of rice and its possible effect on diabetes treatment are both discussed in detail in the study by Pereira et el. [3] One of the most frequently grown and consumed cereals in the world, rice contains carbohydrates, hypoallergenic proteins, and a variety of bioactive substances with established nutritional benefits. The experiment conducted by Yaqiu Wang et el. [4] demonstrates that soluble sugar is a crucial metric to assess rice quality when it is being stored. In this study, the soluble sugars in rice were examined as they changed over 11 months at 0 °C, 10 °C, and room temperature (25 °C). The research done by Lin et el. [5] used a convolutional neural network for automatically extracting a number of rice kernel characteristics from the grey image. The findings of the study on convolutional neural networks detector have shown the capacity and promise of machine vision with well-trained convolutional neural network detectors for varietal kinds detection of rice grain samples with overall average accuracies of 99.52%. The accuracy of most models was more than 95% in the study conducted by Crisóstomo de Castro Filho et el. [6] which employed near-infrared (NIR) hyperspectral technology with traditional machine learning approaches to categorize the five different varieties of rice seed. Additionally, they searched for the band region that contributed the most to the data by visualizing each CNN model using the saliency map technique. In the study carried out by Murat Koklu et el. [7] there are several genetic variants of rice, one of the most extensively produced grain crops in the world. These variations are distinguished from one another by some of their characteristics. Classification procedures were carried out after models were developed using the (ANN) and (DNN) algorithms for the feature dataset and the (CNN) method for the image dataset. The confusion matrix values of the models were used to generate the statistical results of sensitivity, specificity, prediction, F1 score, accuracy, false positive rate, and false negative rate. The results for each model are presented in tables. The models' classification success rate for ANN was 99.87%.In the study carried out by Shizhuang Weng et el. [8] On high-quality rice, the phenomena of adulteration and subpar rice are a persistent problem that harms the interests of farmers, consumers, and merchants. Utilizing spectroscopy, texture, and morphology-based deep learning networks, hyperspectral imaging (HSI) was carried out to identify the type of rice.

3. Methodology

The section below demonstrates how to import data from the directory function and categorize photos of rice types using a Sequential model. It exemplifies the following ideas:

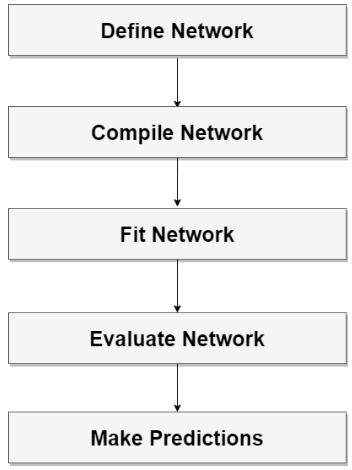


Fig. 4. Rice classification methods

3.1 Image Acquisition

A video-to-image conversion process is employed to get the 7000+ dataset pictures of the rice used in the study. The OpenCV and OS libraries are imported at the beginning of the Python module. The *cv2.VideoCapture()* method is then used to read a video file, which is then assigned to the vidcap variable. Next, it uses the *os.makedirs()* method to construct an output directory. It won't make a new directory if the existing one already exists. The module then uses the *video.read()* function to establish a loop that processes each frame of the movie in turn after setting the frame count to 0. Until there are no more frames to process, the loop will keep running.

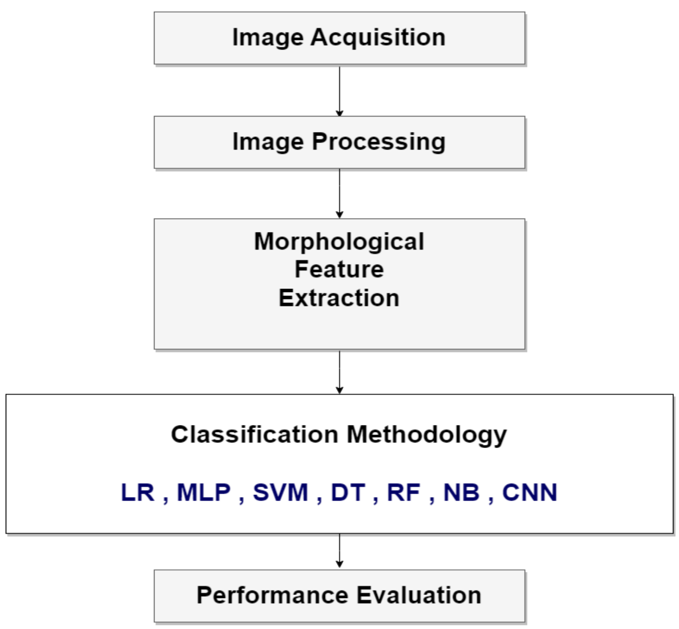


Fig. 5. Image classification phases

Visualize the data

Here are the first 16 images from the training dataset:

Fig. 6. Rice varieties sample set

The Matplotlib Python package is used in this part to display pictures from a dataset. This function takes the first batch of pictures and labels from the train_ds dataset, which is presumably made up of image data with associated labels. The code constructs a subplot with 3 rows and 3 columns for each image in the batch. There are 9 subplots in all because the I variable has a range of 0 to 8. The *plt.figure()*, specifies the size of the shown figure. This guarantees that the photographs are shown at a greater size, making it simpler to examine the images' features.

3.2 Data Acquisition

Data is preprocessed before a model is trained using the Python OpenCV package. The photographs go through a preparation procedure that involves scaling them to a standard size, making them grayscale, and normalizing the pixel values. In order to process a video file indicated by the path, a VideoCapture object must be created using the *cv2.VideoCapture()* function. The output directory is created using the *os.makedirs()* function, and frames taken from the video are stored there. The number of frames read and processed throughout a loop using the vidcap.read() function are recorded in the frame_count variable. The frames are normalized, turned to grayscale, then scaled to 640x480 resolution using cv2 techniques. Using *cv2.imwrite()*, the processed frames are saved as JPEG files with filenames determined based on the frame_count parameter. The final number of frame_count is shown after processing every frame, and the vidcap object is released using the *vidcap.release()* function to release system resources.

3.3 Machine learning model

The convolutional neural network (CNN) is a class of neural networks of deep learning. CNN is a major innovation in image recognition. They are most often used to analyze visual images and often work behind the scenes on the classification of images.

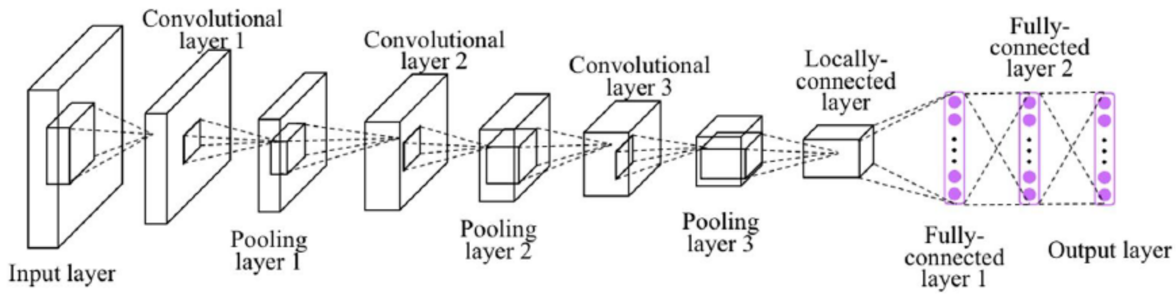


Fig. 7. CNN Model Visualization

A convolution retrieves blocks from the input resource map and applies filters to them to compute new resources, producing a map or converted resources (whose size and depth may be different from the input resource map).

The convolutions are defined by two parameters:

1. Size of the extracted blocks (usually 3x3 or 5x5 pixels).
2. The depth of the output resource map, which is the number of filters applied.

During the training, CNN "learns" the optimal values of the filter matrices to extract important resources (textures, borders, shapes) from the input resource map. As the number of filters (output resource mapping depth) applied to the input increases, the number of resources that can be retrieved by CNN also increases.

The Convolution layer : The image is entered into the convolutional neural network. The reading of the input matrix begins at the top left of the image. Next, the software selects a smaller matrix there, which is called a filter. The filter then produces convolution, moving along the input image. The filter's task is to multiply its values by the original pixel values. All these multiplications are summed up, resulting in one number. After passing the filter across all positions, a matrix is obtained, but smaller than the input matrix.

The Nonlinear layer: The activation function is added after each convolution operation. It introduces nonlinear properties to the network, which are essential for modeling the response variable. Without this property, the network would not be sufficiently intense and would not be able to learn complex relationships between the input and output variables.

The Pooling layer: The pooling layer works with the width and height of the image and performs a downsampling operation on them. This results in a reduction of the image volume, as the image is compressed to less detailed pictures.

Fully connected layer: After a series of convolutional, nonlinear, and pooling layers, it is necessary to attach a fully connected layer. This layer takes the output information from the convolutional networks and classifies the input image into different categories.

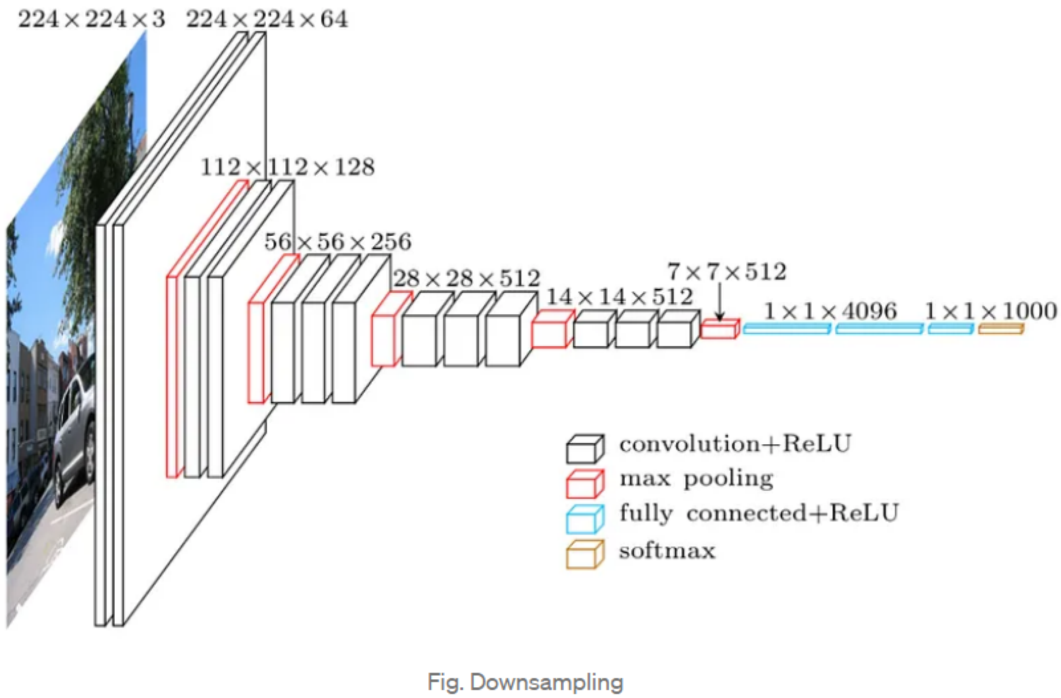


Fig. 8. CNN Model Downsampling

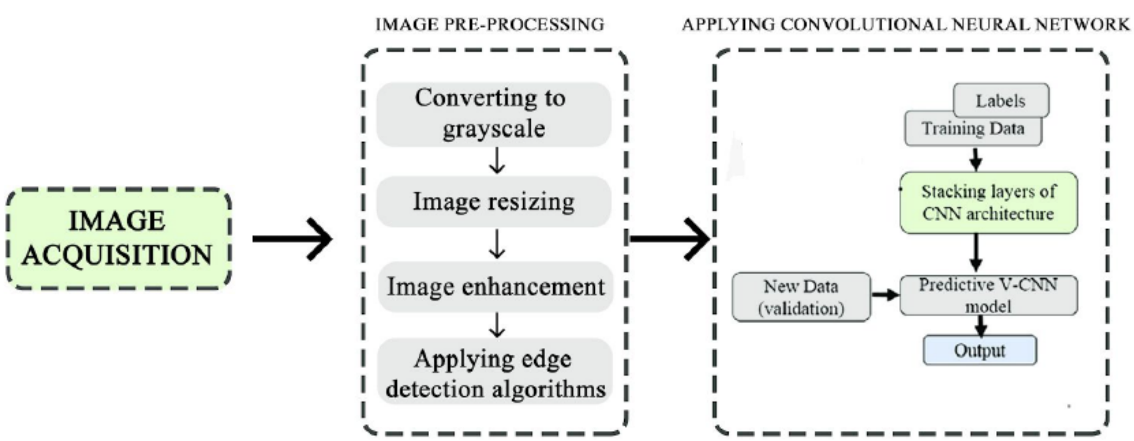


Fig. 9. Blocks representing Modules

a. Model Training:

We will train our model using the Adam optimizer and a categorical cross-entropy loss function. We will use a batch size of 32 and train the model for 15 epochs.

Fig. 10. Image Model Training Set

b. Evaluation:

We will evaluate the performance of our model on a test set consisting of 20% of the total dataset. We will calculate our model's accuracy, precision, recall, and F1-score.

c. Data Augmentation:

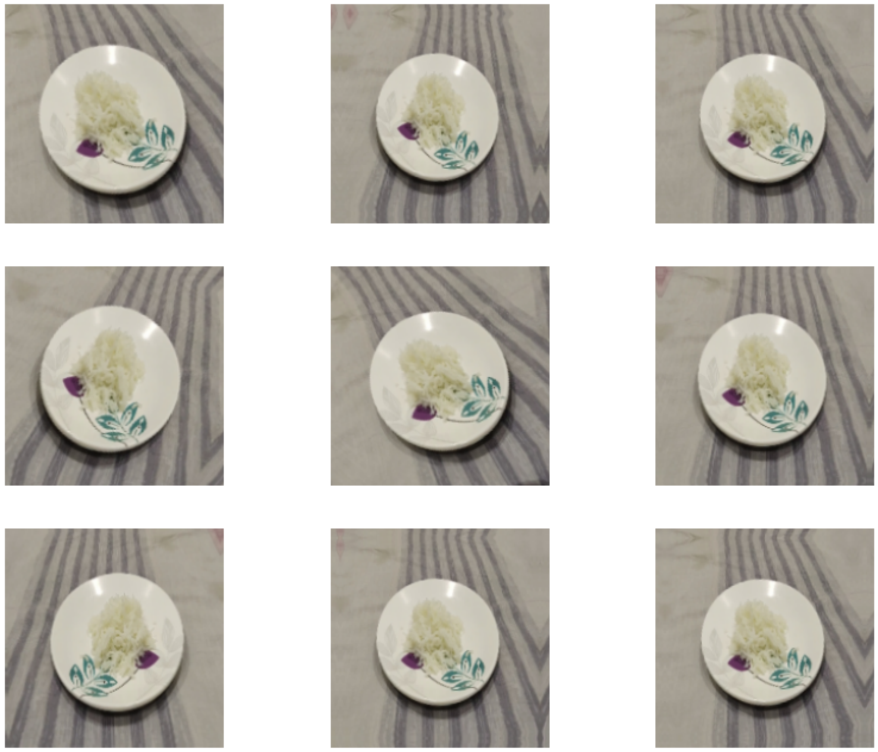


Fig. 11. Data Augmentation Samples

Training and Testing Neural Network Model

Different procedures for assessing a neural network's performance and enhancing its accuracy are used during testing and training. Creating a dataset by gathering and preprocessing the necessary data is the initial stage in training the neural network. The training set and validation set are then separated from this dataset. Next, an appropriate neural network design is selected based on the difficulty, size, and nature of the challenge. Predictions are obtained by passing training data through the network via forward propagation after initializing the model's parameters, such as weights and biases. Using an appropriate loss function, the projected and actual outputs are compared, and backpropagation is used to determine gradients. Then, the estimated gradients are used along with an optimization procedure, such as gradient descent, to update the model's parameters. Up until convergence or the required performance is attained, this process of forward propagation, loss computation, backpropagation, and parameter updating is repeatedly iterated. The neural network is tested once it has been trained. A separate test dataset is created that is not utilized for training or validation. To get predictions, the trained network is fed the test dataset via forward propagation. The model's performance is measured using a variety of performance measures, including accuracy, precision, recall, F1 score, and mean squared error. In order to assess how successfully the model generalizes to new data and to pinpoint areas for development, the outcomes of these metrics are examined. If the model's performance is subpar, modifications can be made by going back to the training phase and changing the model's architecture, hyperparameters, or amount of data. To get a fair assessment of the model's performance, it is essential to make sure that testing and evaluation are done on data that are not part of the training set. To improve the robustness of assessments and avoid overfitting, strategies like cross-validation and regularisation can be used. These stages can be used to test and train neural networks to perform at their best for a particular job.

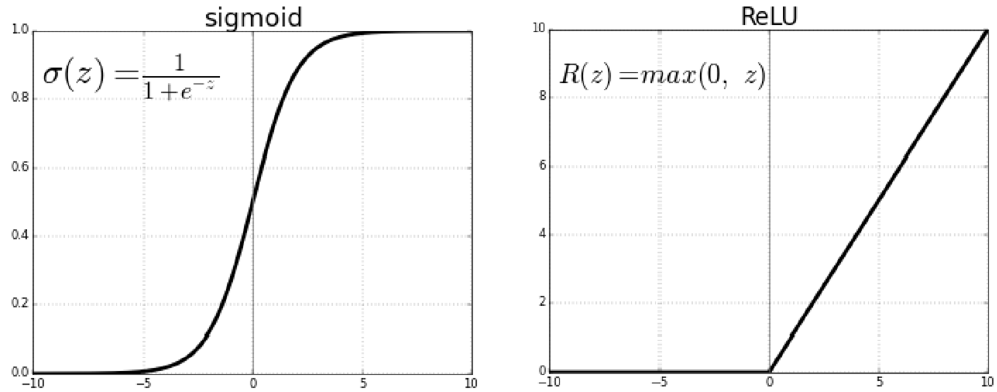


Fig. 12. ReLU Activation Function

a. **Pooling** Pooling progressively reduces the size of the input representation. It makes it possible to detect objects in an image no matter where they're located. Pooling helps to reduce the number of required parameters and the amount of computation required. It also helps control overfitting.
b. **Flattening** You flatten the pooled feature map into a sequential column of numbers (a long vector). This allows that information to become the input layer of an artificial neural network for further processing.
c. **Dropout** A dropout is an approach to regularization in neural networks which helps to reduce interdependent learning amongst the neurons. It prevents the overfitting of data.
d. **Effect of batch size** Higher batch sizes leads to lower asymptotic test accuracy. Using too large a batch size can have a negative effect on the accuracy of your network during training since it reduces the stochasticity of the gradient descent.

4. Results

Preliminary processing was conducted on the images obtained with CVS in order to get a total of 7000+ rice grains in order to categorize the rice kinds employed in our study.

```
image count =  7351
Cooked_basmati_rice_al COUNT = 417
Cooked_basmati_rice_sl COUNT = 456
Cooked_brown_rice_al COUNT = 441
Cooked_brown_rice_sl COUNT = 392
Cooked_masoori_rice_al COUNT = 401
Cooked_masoori_rice_sl COUNT = 466
Cooked_white_rice_al COUNT = 462
Cooked_white_rice_sl COUNT = 506
Uncooked_basmati_rice_al COUNT = 489
Uncooked_basmati_rice_sl COUNT = 424
Uncooked_brown_rice_al COUNT = 516
Uncooked_brown_rice_sl COUNT = 535
Uncooked_masoori_rice_al COUNT = 457
Uncooked_masoori_rice_sl COUNT = 430
Uncooked_white_rice_al COUNT = 465
Uncooked_white_rice_sl COUNT = 494
Found 7351 files belonging to 16 classes.
Using 5881 files for training.
Found 7351 files belonging to 16 classes.
Using 1470 files for validation.
```

Fig. . Importing Test Image and Training Data

```
*******************************************************************
************************* PREDICTION RESULTS **********************
This image most likely belongs to Cooked_basmati_rice_sl with a 100.00 percent confidence.
*******************************************************************
```

Fig. . Final Result validated on Cooked Basmati Rice sl rice variety

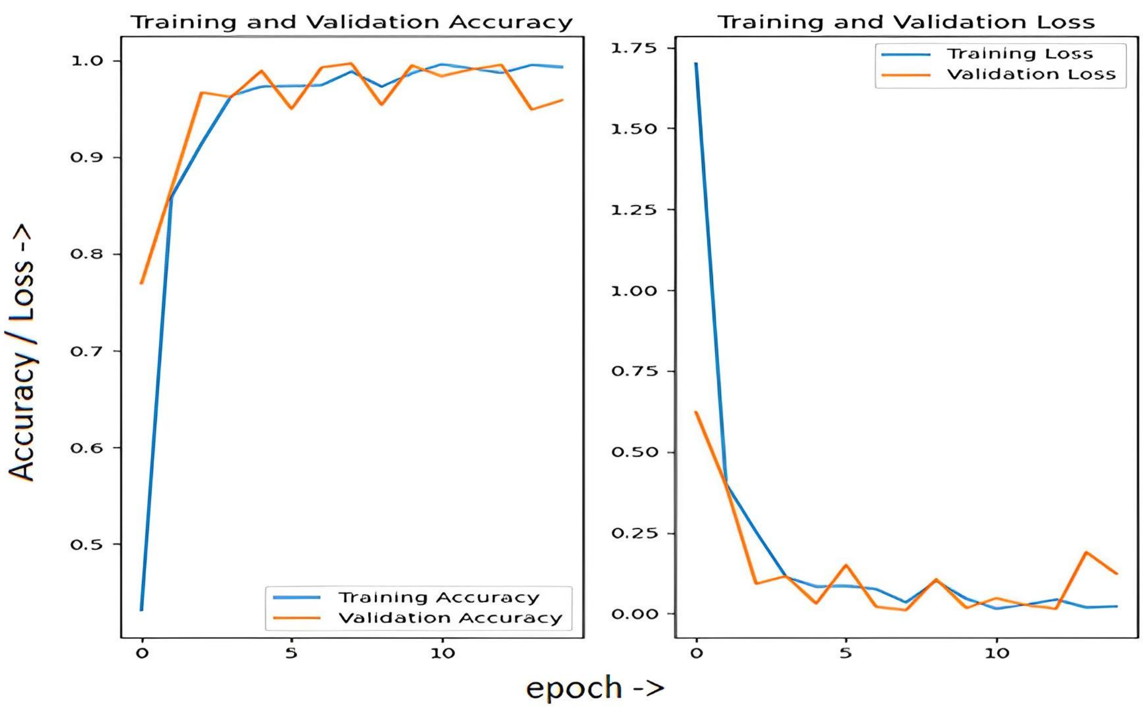
Confusion Matrix:

A confusion matrix, also known as an error matrix, is a specific table layout that allows visualization of the performance of an algorithm, typically a supervised learning one (in unsupervised learning it is usually called a matching matrix). Each row of the matrix represents the instances in a predicted class while each column represents the instances in an actual class (or vice versa). The name stems from the fact that it makes it easy to see if the system is confusing two classes (i.e. commonly mislabeling one as another).

DIAGRAM TO BE INCLUDED by RK

Plotting Accuracy Metrics:

We used the recorded history during training to get a plot of accuracy metrics. The following graph will plot the accuracy & loss on each epoch. We pick up the training data accuracy and the validation data accuracy for plotting. Similarly for the training data loss and the validation data loss is plotted in the second graph



.

5.Conclusion

We employed a convolutional neural network, which is a very effective and trainable neural network, in our research. It mostly serves to classify images. CNN's image classifications start with an input image, preprocess it, and then assign it to one of many categories. Despite all the problems, we ultimately reached an accuracy of about 95%. Additionally, while the majority of the current systems required a high-end GPU, the project is also optimized for low-end PCs, making it both efficient and cost-effective.

In this paper, we looked at how to classify rice using data collected using a TensorFlow Sequential model. With the help of picture scaling, grayscale conversion, and pixel value normalization, we have collected a dataset of more than 7000 photographs of cooked and uncooked basmati, brown, masoori, and white rice. Convolutional Neural Network (CNN) architecture was used to train our model in TensorFlow, and its performance was assessed using test data.

References

1. Omar, Karzan & Salih, Botan & Abdulla, Nahla & Hussin, Basi & Rassul, Shiraz. (2016). Evaluation of Starch and Sugar Content of Different Rice Samples and Study their Physical Properties. Indian Journal Of Natural Sciences. 6. 11084-11093.
2. Bhavadharini, B., Mohan, V., Dehghan, M., Rangarajan, S., Swaminathan, S., Rosengren, A., Wielgosz, A., Avezum, A., Lopez-Jaramillo, P., Lanas, F., Dans, A. L., Yeates, K., Poirier, P., Chifamba, J., Alhabib, K. F., Mohammadifard, N., Zatońska, K., Khatib, R., Keskinler, M. V., . . . Yusuf, S. (2020). White Rice Intake and Incident Diabetes: A Study of 132,373 Participants in 21 Countries. Diabetes Care, 43(11), 2643-2650. https://doi.org/10.2337/dc19-2335
3. Pereira C, Lourenço VM, Menezes R, Brites C. Rice Compounds with Impact on Diabetes Control. Foods. 2021; 10(9):1992. https://doi.org/10.3390/foods10091992
4. Yaqiu Wang, Yuanrong Zheng, Ran Zhou, Ming Ma, Kinetic studies on soluble sugar profile in rice during storage: Derivation using the Laplace transform, Innovative Food Science & Emerging Technologies, Volume 76,2022,102915, ISSN 1466-8564,https://doi.org/10.1016/j.ifset.2021.102915.
5. P. Lin, Y. Chen, J. He and X. Fu, "Determination of the Varieties of Rice Kernels Based on Machine Vision and Deep Learning Technology," 2017 10th International Symposium on Computational Intelligence and Design (ISCID), Hangzhou, China, 2017, pp. 169-172, doi: 10.1109/ISCID.2017.208.
6. Crisóstomo de Castro Filho, H.; Abílio de Carvalho Júnior, O.; Ferreira de Carvalho, O.L.; Pozzobon de Bem, P.; dos Santos de Moura, R.; Olino de Albuquerque, A.; Rosa Silva, C.; Guimarães Ferreira, P.H.; Fontes Guimarães, R.; Trancoso Gomes, R.A. Rice Crop Detection Using LSTM, Bi-LSTM, and Machine Learning Models from Sentinel-1 Time Series. Remote Sens. 2020, 12, 2655. https://doi.org/10.3390/rs12162655
7. Murat Koklu, Ilkay Cinar, Yavuz Selim Taspinar, Classification of rice varieties with deep learning methods, Computers and Electronics in Agriculture, Volume 187,2021,106285, ISSN 0168-1699,https://doi.org/10.1016/j.compag.2021.106285
8. Shizhuang Weng, Peipei Tang, Hecai Yuan, Bingqing Guo, Shuan Yu, Linsheng Huang, Chao Xu, Hyperspectral imaging for accurate determination of rice variety using a deep learning network with multi-feature fusion, Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy, Volume 234,2020,118237, ISSN 1386-1425,https://doi.org/10.1016/j.saa.2020.118237.
9. V. Truong Hoang, D. P. . Van Hoai, T. . Surinwarangkoon, H.-T. Duong, and K. Meethongjan, “A Comparative Study of Rice Variety Classification based on Deep Learning and Hand-crafted Features”, ECTI-CIT Transactions, vol. 14, no. 1, pp. 1–10, Jan. 2020.
10. M. Ramesh, S. Zakiuddin Ali, K.R. Bhattacharya, Structure of rice starch and its relation to cooked-rice texture, CarbohydratePolymers,Volume38,Issue4,1999,Pages337-347,ISSN0144-8617,https://doi.org/10.1016/S0144-8617(98)00125-8.