Grenze International Journal of Engineering and Technology, Jan Issue

Wheat Variety Identification using Deep Lea

Lavanya N[1], Preethi R[2], Thejaswini U[3], Nesara[4] and Prerana[5]
[1]Assistant Professor, ATMECE, Mysuru
[2-5]UG Student, ATMECE, Mysuru

***Abstract*—India is the second largest producer of wheat in the world after China. Specifying the quality of wheat manually requires an expert judgment and is time consuming. To overcome this problem, machine algorithms can be used to classify wheat according to its quality. In this paper we have done wheat classification by using two machine learning algorithms, that is, Support Vector Machine (OVR) and Neural Network (LM). For classification, images of wheat grain are captured using digital camera and thresholding is performed. Following this step, features of wheat are extracted from these images and machine learning algorithms are implemented. The accuracy of Support Vector Machine is 86.8% and of Neural network is 94.5%. Results show that Neural Network (LM) is better than Support Vector Machine.**

Keywords— Deep learning techniques, Convolution Neural Network, Image classification**.**

I. INTRODUCTION

India is the world's greatest producer of pulses, dry fruits, dairy products, tea, spice products, and fibrous crops, as well as wheat, paddy, processed fruits and vegetables, sugarcane, cotton, and oilseeds. The country has achieved self-sufficiency in the production of pulses and coarse cereal grains. The worldwide market demand for cereals has developed an outstanding environment for Indian cereal exports. In addition to being the world's biggest cereal producer, India is also the largest cereal exporter. Wheat farming extends back more than 5000 years to the Indus Valley culture. With a projected production of 105 million tonnes in 2021-22, India is the second-largest wheat grower in the world. India has a rich diversity of around 371 local wheat varieties. These varieties are selected and evolved by farmers for specific human needs. The availability of numerous wheat varieties offers a challenge for the young generation to identify the varieties. Every variety has certain properties, specific purpose, and utility. The properties include grain quality (size, shape, weight, color, and texture), yield, tolerance to biotic and abiotic stresses, farm input requirements, maturity duration, end products, and the pricing are determined by the varietal characteristics. Therefore, varietal characterization and identification have become more relevant to the present situation. This is drawing the attention of farmers, breeders, millers, food processors, traders, technologists, seed grading, and certification agencies to address the issues and challenges in wheat variety identification and grading. Figure 1 depicts the images of various wheat varieties. The sort of wheat variety is a mark of its edible quality, genetic and agronomic attributes, and commercial esteem. These factors determine the grade and cost of wheat grains. However, their objective and quick measurements are complicated to obtain without the use of state-of-the-art technology. Some government laboratories have the expertise and technical devices to make measurements. The physical and visual methods are still the dominant techniques used in India today by the wheat trading industry. The widely used and accepted method in India is

Grenze ID: 01.GIJET.9.1.616

the classification based on the length-to-breadth ratio of wheat grains. This method is quite challenging and unreliable due to the inter varietal homogeneity of the wheat grains. In this context, the use of non-destructive, cost-effective, and advanced image-based systems for grain handling and quality control is the need of the day in India. The recent research and significant progress in the field of computer vision and image processing have gained wide acceptance in automating grain handling practices.

Figure 1: Varieties of Wheat

Machine learning (ML) is a type of artificial intelligence (AI) that allows software applications to become more accurate at predicting outcomes without being explicitly programmed to do so. Machine learning algorithms use historical data as input to predict new output values.Automatic pattern predictions based on behavior analysis. Deep learning is a class of machine learning algorithms that uses multiple layers to progressively extract higher-level features from the raw input. For example, in image processing, lower layers may identify edges, while higher layers may identify the concepts relevant to a human such as digits or letters or faces.In this paper we are identifying varieties of wheat available in market using deep learning techniques.

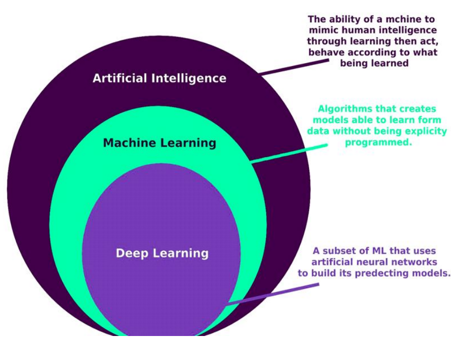


Figure 2: Introduction to deep learning

II. Literature Survey

A) Machine learning approaches for the classification for the corn seed using hybrid features:

Seed purity is an important indicator of crop seed quality. On the other side, corn is an important crop of the modern agricultural industry with more than 40% grain Worldwide production. The purpose of this study was to examine the feasibility of a machine learning (ML) approach for classifying different types of corn seeds. The seed digital images (DI) of six corn varieties were Desi Makkai, Sygenta ST-6142, Kashmiri Makkai, Pioneer P-1429, Neelam Makkai, and ICI 339. This was achieved through a digital camera in a natural environment without a complicated laboratory system. The acquired DI dataset converted to a hybrid feature dataset, which is the combination of histogram, texture, and spectral features. For each corn seed image, a total of fifty-five hybrid-features was acquired on every non-overlapping region of interest (ROI), sizes (75 × 75), (100 × 100), (125 × 125) and (150 × 150). The nine optimized features have been acquired by employing the correlation-based feature selection (CFS) technique with the Best First search algorithm. To build the classification models, Random forest (RF), BayesNet (BN), LogitBoost (LB), and Multilayer Perceptron (MLP) were employed using optimized multi-feature using (10-fold) cross-validation approach. A comparative analysis of four ML classifiers, the MLP performed outstanding classification accuracy (98.93%), on ROIs size (150 × 150). The accuracy values by MLP on six corn seed verities named Desi Makkai, Sygenta ST-6142, Kashmiri Makkai, Pioneer P-1429, Neelam Makkai, ICI- 339 was 99.8%, 97%, 98.5%, 98.6%, 99.9%, and 99.4%, respectively.

B) Crop pest classification using convolution neural network and transfer network:

The growth of most important field crops such as rice, wheat, maize, soybean, and sugarcane are affected due to attack of various pests and the crop production is reduced due to different types of insects. The classification and identification of all types of crop insects correctly is a difficult task for the farmers due to the similar appearance in the earlier stage of crop growth. To address this issue, Convolutional neural network (CNN) with deep architectures is being applied as it performs automatic feature extraction and learns complex high-level features in image classification applications. This study proposed an efficient deep CNN model to classify insect species on three publicly available insect datasets. The National Bureau of Agricultural Insect Resources (NBAIR) dataset used as first insect dataset that consists of 40 classes of field crop insect images, while the second and third dataset (Xie1, Xie2) contains 24 and 40 classes of insects respectively. The proposed model was evaluated and compared with pre-trained deep learning architectures such as AlexNet, ResNet, GoogLeNet and VGGNet for insect classification. Transfer learning was applied to fine-tune the pre-trained models. The data augmentation techniques such as reflection, scaling, rotation, and translation are also applied to prevent the network from overfitting. The effectiveness of hyperparameters was analysed in the proposed model to improve accuracy. The highest classification accuracy of 96.75, 97.47, and 95.97% was achieved in proposed CNN model for NBAIR insect dataset (40 classes), Xie1 (24 classes) insect dataset and Xie2 (40 classes) insect dataset respectively. The results demonstrated that the proposed CNN model is effective in classifying various types of insects in field crops than pretrained models and can be implemented in the agriculture sector for crop protection.

C) Soft computing for Wheat grain classification:

This paper attempts to identify varieties of wheat grains by using some prominent soft computing techniques. The methods primarily perform the identification based on certain geometric features of a given wheat grain. The performance of three soft computing methods including kNearest Neighbours (k-NN), Support Vector Machine (SVM) and Neural Networks (NN) are evaluated. With an accuracy of hundred percent, NNs have demonstrated the most promising results in comparison to the other methods

D) thenmoli and srinivasulu reddy(2019) proposed a cnn model for classifying insect speices in wheat crops using 3 sets of public insect data:

The present work proposes a deep learning-based approach that provides an accurate classification for wheat varietal level classification (VLC). Particularly, the Convolutional Neural network (CNN) was used to classify the wheat grain image into four varieties (Simeto, Vitron, ARZ, and HD). Furthermore, five standard CNN architectures were trained based on Transfer Learning to boost the classification performance. To assess theproposed models' quality, we used a dataset of 31,606 single-grain images collected from different Algeria regions, and their images were captured using different scanners.The results showed that the test accuracy ranged from 85% to 95.68% for varietal level classification. The best test accuracy was obtained with the DensNet201 architecture (95.68%), Inception V3 (95.62%), and MobileNet (95.49%). Hence, the proposed approach results are accurate and reliable, encouraging the deployment of such an approach in practice.

III. Existing System

To design and implement the proposed model for wheat classification, a comprehensive study of wheat variety and existing approaches presented in the literature to detect and classify the variety have been conducted. Including an analysis of common wheat variety, this part has been divided into four sub-sections, namely:

A) a background study on disease identification based on symptoms
B) existing variety identification systems
C) decision support systems for crop disease
D) Knowledge-based systems

IV. Proposed System

The present project proposes a deep learning-based approach that provides an accurate classification for wheat varietal level classification. Particularly, the Convolutional Neural Network (CNN) will be used to classify the wheat grain image into four varieties (V1, V2, V3, and V4). Leading deep learning models such as VGG16, ResNet-50, CapsNet, GoogleNet will be used for the classification of wheat varieties. A typical classification of images using deep learning model.

- In this proposed system we are collecting 4 type of wheat variety datasets and applying classification steps
- The datasets collected are pre processed using image pre processing techniques
- The Pre processed datasets are split into train and test
- The splitted datasets are passed as input to Convolution neural network which helps in training Model
- The trained model is compared with test data for accuracy Prediction.

V. Problem Statements And Objectives

Considering the potential of deep learning techniques in the classification of grains, the problem statement entitled “To develop an automated system for wheat variety identification using the deep neural networks” is defined with the following objectives.

i. To construct an image Dataset of wheat grains.
ii. To pre-processthe image to enhance the quality of images.
iii. To select a leading or prominent CNN/Deep learning model for Image Classification.
iv. Compare the performance ofselected CNN model with otherlearning

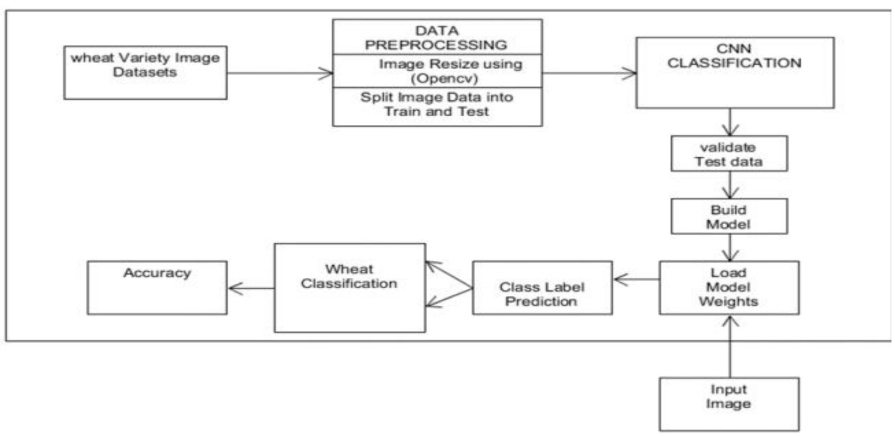


Figure 3 :Architecture of the propsed system

VI. Methodology

Collection datasets:

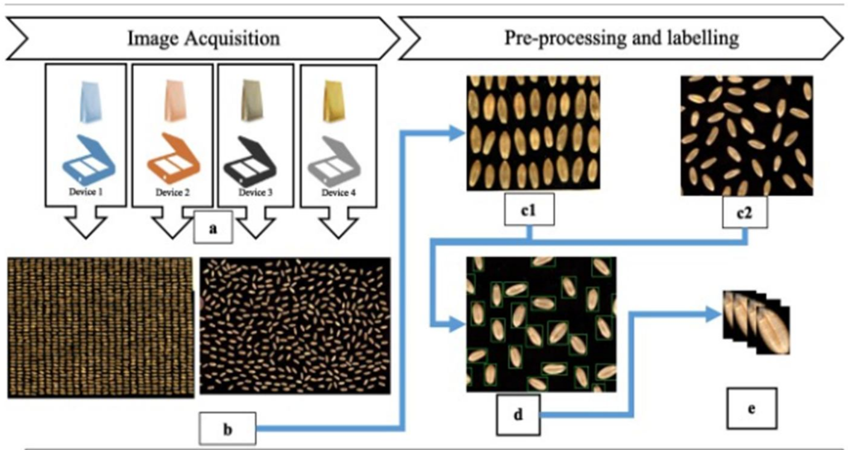
- We are going to collect datasets for the prediction from the kaggle.com
- The data sets consists of 3Classess

Data Pre Processing:

- In data pre-processing we are going to perform some image pre-processing techniques on the selected data
- Image Resize
- And Splitting data into train and test

Data Modelling:

- The splitted train data are passed as input to the CNN algorithm, which helps in training.
- The trained skin image data evaluated by passing test data to the algorithm
- Accuracy is calculated



a: Different models of flatbed scanner, **b :** principals RGB images (A4format), **c1**: Regular anteroposterior, **c2** : Random anteroposterior, **d**: Segmentation and cropping, **e**: single grain data set

The image acquisition step began with the acquisition of 65 principal images. All grains were scanned with a black background. Each main image included 500 to 1013 grains and had the following characteristics: depth of 24bit, taken in JPEG and TIFF format, with a resolution of 300 and 600 dots per inch (Dpi).

The image segmentation step was performed to obtain single grain images from the principal image, as shown in Fig.. The segmentation consisted of cropping the principal images into single grain ones without removing the background. We obtained 31 606 single grain RGB images with a random and regular anteroposterior direction.

Deep learning classification

We have used five standard CNN architectures: (InceptionV3, MobileNet, Xception, ResNet50, and DensNet201) trained using a transfer learning by fine-tuning. These architectures were pre-trained on the ImageNet dataset. The transfer learning strategy is chosen because it can improve accuracy while accelerating the training (. Transfer learning can provide important benefits for automated plant identification and can improve low-performance plant classification models.

Model structure

To build the model, we have replaced the top layers of standard architectures. The remaining part of the architecture is called “base model”. We replaced the original top layers with the following top layers:

The global Average Pooling layer calculates the global average of each base model's last layer's feature maps, which decreases the risk of overfitting.A fully connected layer of 1024 neurons is used as the first layer in the top layers.

The last fully connected prediction layer with SoftMax activation function. The number of neurons in this layer is four according to the number of classes (*Simeto, Vitron, ARZ, HD*).

Deep learning framework and machine specifications

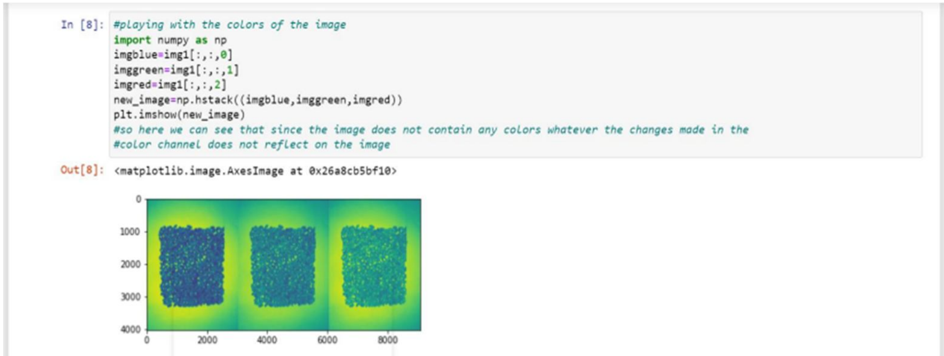
All our training and testing script are written in *Python 3* programming language. We have used python deep learning framework *Keras* , thanks to its simplicity. As *Keras* backend, we have used *Tensorflow 1*. Our experiments were performed on a machine having the following specifications:

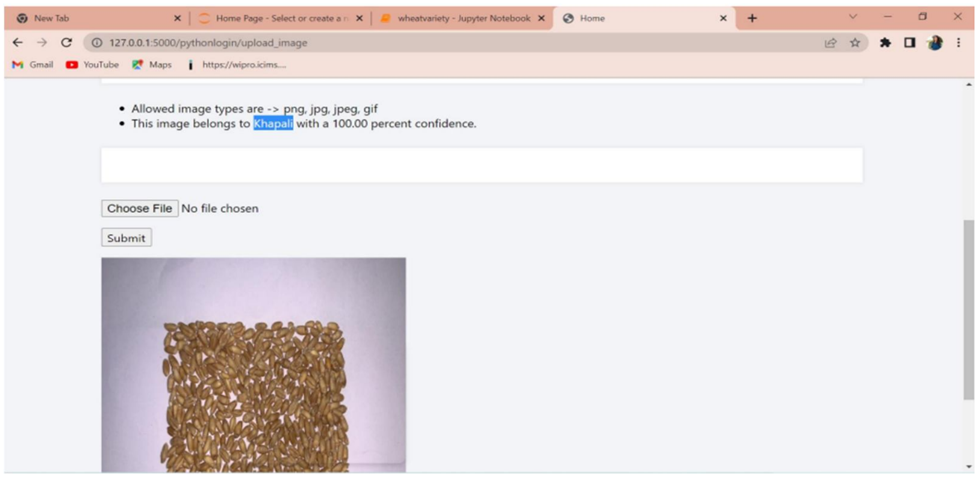
•Ram: 64 GB
•GPU: 4 × Nvidia Geforce 1080ti 11 Gb
•CPU: Intel(R) Xeon(R) CPU E5-2620 v4 @ 2.10 GHz

Training hyperparamaters

The batch size was set at 32, and the number of epochs was 100. The model optimization during the training is achieved using gradient descent algorithm Adam with learning rate 0.001. Moreover, we have used the early stopping regularization method that helps to avoid overfitting. In early stopping, the training is stopped when the loss function does not improve on the validation set; rather than finishing all the 100 epochs

Results





VII. APPLICATIONS

There are several food grain applications based on computer vision, in general, they can be grouped in:
i) quality control approaches
ii) grain variety classification. Some examples of each one of them are provided in the next subsections.

VIII. CONCLUSION

AI is at the centre of a new enterprise to build computational models of intelligence. The main assumption is that intelligence (human or otherwise) can be represented in terms of symbol structures and symbolic operations which can be programmed in a digital computer. There is much debate as to whether such an appropriately programmed computer would *be* a mind, or would merely *simulate* one, but AI researchers need not wait for the conclusion to that debate, nor for the hypothetical computer that could model all of human intelligence. Aspects of intelligent behaviour, such as solving problems, making inferences, learning, and understanding language, have already been coded as computer programs, and within very limited domains, such as identifying diseases of soybean plants, AI programs can outperform human experts. Now the great challenge of AI is to find ways of representing the commonsense knowledge and experience that enable people to carry out everyday activities such as holding a wide-ranging conversation, or finding their way along a busy street. Conventional digital computers may be capable of running such programs, or we may need to develop new machines that can support the complexity of human thought

REFERENCES

[1] M.Z. Alom, T.M. Taha, C. Yakopcic, S. Westberg, P. Sidike, M.S. Nasrin, M. Hasan, B.C. Van Essen, A.A. S. Awwal, V.K. Asari A State-of-the-Art Survey on Deep Learning Theory and Architectures

[2] M. Brahimi, K. Boukhalfa, A. Moussaoui Deep Learning for Tomato Diseases: Classification and Symptoms Visualization

[3] Chiara Delogu Validation of a new method: use of SDS-PAGE technique for the verification of Triticum and ×Triticosecale

[4] R. Choudhary, S. Mahesh, J. Paliwal, D.S. Jayas Identification of wheat classes using wavelet features from near infrared hyperspectral images of bulk sample**s**Biosyst. Eng., 102 (2009), pp. 115-127,

[5] Copeland, L.O., McDonald, M.B., 2001. Principles of Seed Science and Technology, Principles of seed science and technology. Springer US, Boston, MA. https://doi.org/10.1007/978-1-4615-1619-4.

[6] Davies, E.R., 2012. Computer and Machine Vision, Computer and Machine Vision, Texts in Computer Science. Elsevier, London. https://doi.org/10.1016/C2010-0-66926-4.

[7] E.R. Davies The application of machine vision to food and agriculture: a review Imaging Sci. J., 57 (2009), pp. 197-217

[8] Dolata, P., Reiner, J. 2018. Barley Variety Recognition with Viewpoint-aware Double-stream Convolutional Neural Networks. *Proceedings of the 2018 Federated Conference on Computer Science and Information Systems*. ACSIS, Vol. 15, pp. 101–105.

[9] C.J. Du, D.W. Sun Learning techniques used in computer vision for food quality evaluation: A review J. Food Eng., 72 (2006), pp. 39-55,

[10] S. Dutta Gupta, Y. Ibaraki Plant Image Analysi**s** Fundamentals and Applications (2014)