Hybrid CNN-SVM Based Seed Purity Identification and Classification System

Suganthi M[1] and Dr. J. G. R. Sathiaseelan[2]

Affiliated to Bharathidasan University, Bishop Heber College (Autonomous), Tamilnadu, India

Abstract

Manual seed classification challenges can be overcome using a reliable and autonomous seed purity identification and classification technique. It is a highly practical and commercially important requirement of the agricultural industry. Researchers can create a new data mining method with improved accuracy using current machine learning and artificial intelligence approaches. Seed classification can help with quality making, seed quality controller, and impurity identification. Seeds have traditionally been classified based on characteristics such as colour, shape, and texture. Generally, this is done by experts by visually examining each model, which is a very time-consuming and tedious task. This approach is simple to automate, making seed sorting far more efficient than manually inspecting them. Computer vision technologies based on machine learning (ML), symmetry, and, more specifically, convolutional neural networks (CNNs) have been widely used in related fields, resulting in greater labour efficiency in many cases. To sort a sample of 3000 seeds, KNN, SVM, CNN and CNN-SVM hybrid classification algorithms were used. A model that uses advanced deep learning techniques to categorise some well-known seeds is included in the proposed hybrid system. In most cases, the CNN-SVM model outperformed the comparable SVM and CNN models, demonstrating the effectiveness of utilising CNN-SVM to evaluate data. The findings of this research revealed that CNN-SVM could be used to analyse data with promising results. Future study should look into more seed kinds to expand the use of CNN-SVMs in data processing.

Keywords:

Convolutional Neural Networks (CNNs), Support Vector Machines (SVM), Seed Purity classification.

1. Introduction

The agriculture sector employs a large portion of the world's population. Agriculture is vital to the economies of many emerging and impoverished countries. With the rise in the world's population, this industry has seen numerous changes [1]. The need to boost global agricultural output and sustainability is strongly tied to human population growth. Agriculture was first introduced to technology more than a century ago, and various research have been undertaken since the 1990s to improve production efficiency [1]. Artificial intelligence (AI)-based advanced industrial technologies have recently been used in agriculture to improve productivity, environmental impact, food security, and sustainability. Before tackling any issues in this field, it is necessary to have a basic understanding of agriculture's environment and the fundamental necessities of farming. This is one of the most difficult research disciplines, and technology has enormous potential to be integrated into it in order to raise the quantity and quality of agricultural goods. Artificial intelligence, specifically deep learning perceptions, can be utilised to better the agricultural business [2].

Crop growing in the agriculture sector is primarily reliant on seeds. There is no way to produce or harvest any crops without seeds. For many years, the human population has been rapidly rising. As a result of this population growth, agricultural area is dwindling, resulting in a decrease in food production. Crop production rates should be enhanced to balance consumption and production rates. People have begun to grow crops and vegetables in their homes in this regard. However, not everyone held the necessary knowledge. Cultivation can only be done by someone who is skilled at identifying seeds. To eliminate this reliance, an automated system is required to aid in the identification and classification of various types of seeds. Several research have been undertaken in which AI approaches have been used to handle various difficulties linked to seeds, ranging from simple object classification to complicated texture and pattern detection. Machine learning approaches have been used more frequently in recent studies to classify seeds from various crops, fruits, and vegetables. The majority of these research focused on a single seed type (e.g., weed seeds [3], cottonseeds [4], rice seeds [5,6], oat seeds [7], sunflower seeds [8], tomato seeds [9], and corn [10,11]), with various objectives.

Based on a hybrid deep learning CNN model and the use of symmetry, this study provided an efficient strategy for seed purity identification and classification. The model is trained using transfer learning, which is a promising machine learning technology that focuses on transferring information across domains. It's a useful machine learning (ML) tool for dealing with challenges like insufficient training data, as it encourages the idea of not having to train the model from scratch, which cuts down on training time [12]. The seed classification technique was carried out using a dataset of symmetric images of different seeds in this research. To train the proposed model, several symmetric images from each

Manuscript received October 5, 2022
Manuscript revised October 20, 2022
https://doi.org/10.22937/IJCSNS.2022.22.10.35

seed class were taken and subsequently supplemented. In order to achieve successful seed identification and classification [13]. However, the model was altered by the addition of new layers, with the goal of improving classification accuracy and lowering error rates. This change to the hybrid model improves the classification process and, as a result, helps the average person in detecting and classifying different types of seeds.

The goal of this study was to use X-Ray imaging and a convolutional neural network to determine seed purity. The particular goals were to: (1) construct deep learning models to detect seed purity; (2) compare the results of hybrid deep learning approaches to SVM and CNN; and (3) assess the impact of the amount of training samples on SVM, CNN, and CNN-SVM.

Kantip Kiratiratanapruk [2020] [06] designed and tested a quality inspection system for identifying 14 rice cultivars from a database of over 5,000 seeds obtained from a variety of planting locations. They developed a strategy for dealing with a huge number of seeds in order to properly prepare data quality before using it as input for machine learning algorithms and to increase categorization abilities. In a statistical experiment, they found that when the PCA dimension number was increased to 600, SVM produced the best classification with an accuracy of 83.9 percent. In a deep learning experiment, they discovered that the InceptionResNetV2 model using validation data had the lowest loss value and the best classification accuracy (95.14 %). The results revealed that the deep learning method outperformed the traditional method by up to 11.24 %.

Tongyun Luo [2021] [07] proposed a recognition scheme for the intelligent classification of a variety of weed seed species, as well as a technique that is more probable to have broad economic and technological implications. To begin, single weed seed photos were segmented rapidly and comprehensively. Six popular and unique deep CNN models are then used to compare the best technique for intelligently detecting 140 species of weed seeds. The quantitative study revealed that AlexNet and GoogLeNet were the best. AlexNet has a high level of classification accuracy and efficiency (low time consumption), while GoogLeNet has the highest level of classification accuracy.

Rashidah Ruslan [2022] [14] shows how to distinguish weedy rice seed variations and farmed rice seeds using image processing and machine learning approaches. The RGB picture used the logistic regression (LR) model to produce the most optimal model, which obtained 85.3 percent sensitivity, 99.5 percent specificity, 97.9 percent accuracy, and 92.4 percent average accurate classification utilising all 67 characteristics.

Seema Shedole [2022] [15] Image processing methods and Deep neural networks are included in the proposed paper's implementation phases. Good grain, Damaged Grain, Broken Grain, and Foreign Particles are the four groups evaluated in this model. A convolutional system is utilised to categorise wheat seeds and identify the state of four different classes.

Aqib Ali's [2020] [16] goal was to determine if a machine learning (ML) approach could be used to classify different varieties of corn seeds. Six corn types' seed digital images (DI). LogitBoost (LB), Random Forest (RF), BayesNet (BN) and Multilayer Perceptron (MLP) were recycled to develop classification models utilising an optimal multi-feature (10-fold) cross-validation technique. In a comparison of four machine learning classifiers, the MLP had the highest classification accuracy (98.93%) on ROIs of 150 x 150 pixels. MLP's accuracy values for six maize seed varieties named Kashmiri Makkai, Desi Makkai, Neelam Makkai, ICI339 Sygenta ST-6142 Pioneer P-1429were 99.8%, 97.8%, 98.5 %, 98.6%, 99.9%, and 99.4 %, respectively.

Venkat Margapuri[2021] [17] proposed a seed classification background based on deep learning and artificial picture datasets for seed phenotyping. The possibility of low-altitude UAV images and synthetic datasets to meet the abundance of training examples required for training neural network architectures is demonstrated using a prototypical seed phenotyping scenario. Furthermore, the research presents a seed classification background as a proof-of-concept employing Oxford's VGG-16, Microsoft's ResNet-100 and VGG-19 convolutional neural networks. An ensemble model is created to improve the framework's classification accuracy, resulting in a total accuracy of 94.6 %. Rough rice, soy, Canola sorghum, and wheat are the five dissimilar types of seeds that are classified.

Bashier [2021] [18] created the CNN model as well as VGG16, VGG19, Resnet50, Resnet101, and Resnet152, which are all ready-made models. The package includes 1,695 images of sesame leaves divided into three categories: diseased leaves, healthy leaves, and diseases currently impacting the Sesame in Sudan. The leaves were photographed in several arenas in Gadarif State. With a testing accuracy of 88.5 %, and training accuracy of 90.77 % the generated model produced the best results. This model's future development and potential upgrades were also considered.

Using a MATLAB graphical user interface, Peng Xu [2021] [19] proposed extracting 16 key features (12 dimensions and 4 of form) from maize seed images, and a consumer interface was constructed (GUI).). The multilayer perceptron (MLP), linear discrimination (LDA), decision tree (DT), Naive Bayes (NB), k-nearest neighbours (KNN), support vector machine (SVM), and AdaBoost algorithms were used to create the varietal classification model. MLP, DT, LDA, NB, SVM, KNN, and AdaBoost had overall classification accuracy of 96.26, 94.95, 95.97, 93.97, 96.46, 95.59, and 95.31 %, respectively. The maximum accuracy was achieved by the SVM algorithm for the BaoQiu, ShanCu, XinNuo, LiaoGe, and KouXian types, which were 93.06, 98.94, 96.14, 89.64, and 99.21reduce percentages, respectively.

A solution for supporting farmers in crop optimization was proposed by Belal A [2019] [20]. They designed and deployed a two-class classifier that takes plant seedling photographs from 12 distinct species as input, creates a model using deep learning convolutional neural networks, and utilises this model to identify the kind of (unseen) plant seedling images. The projected process yields talented results, including an accuracy rate of 99.48%.

2. METHODOLOGY

A) CNN

A CNN is a model of an artificial neural network (ANN). The input layer, hidden layer, and output layer make up this type of network in general. The output of one layer is typically used as an input for the following layer in this type of network. The CNN model comprises four phases and is used for image analysis: convolution layer, pooling layer, flatten layer, and fully connected layer. However, a sample glance at alternative architectures will reveal that the amount of layers, as well as the type of layers, will vary.

The convolution layer is the first layer in the CNN model to extract a feature from an image; inside this, the property of each pixel and the relationship between neighbour pixels will be extracted using specific mathematical operations. After extracting features, by avoiding unnecessary data, the pooling layer will consider the most important information, this process is known as subsampling, and It can be (Max pooling, Avg pooling, Sum Pooling) and it tries to lower the size of the data map without losing imported data, when the pooling layer is used and the all-important feature is mapped, the feature mapping can sometimes result in overfitting, the flatten layer 2D arrays to 1D arrays before applying. Fully connected is typically the last layer of a network, and because all networks are properly connected, the output of fully connected is the final outcome.

B) SVM

The Support Vector Machine, created by Vapnik and Cortes, is a powerful discriminative classifier. For many pattern classification/recognition tasks, it has been frequently used with positive results. Because of its parsimony, flexibility, prediction capacity, and global optimal character, it is considered the state-of-the-art tool for addressing linear and non-linear classification problems. The structural risk minimization, rather than the empirical risk minimization utilised in artificial neural networks, is the foundation of their formulation. Even when the data are linearly inseparable, SVM is utilised to establish an appropriate separation hyper-plane (equation 1) or decision surface by embracing a revolutionary technique based on mapping the sample points into a high-dimensional feature space and categorising it using a nonlinear transformation Φ. By solving a quadratic programming problem using regularisation parameters, the optimal hyper-plane can be found. Kernel functions such as linear, radial basis function, sigmoid, and polynomial kernel types performed this transformation.

The linear kernel: $K(x, y) = x \times y$

- The polynomial kernel: $K(x, y) = [(x \times y) + 1]^d$
- The Sigmoid kernel: $K(x, y) = tanh\,(\beta_0 x\, y + \beta_1)$
- RBF kernel (Radial Basis Function): $K(x, y) = exp(-\gamma\, ||x - y||^2)$

With d, β_0, β_1, and γ are parameters that will be determinate empirically.

$$f(x) = W^T \Phi(x) + b \; (1)$$

Where $W \in R^n, b \in R$ and $\Phi(x)$ is a feature map.

In this work, because the feature space is linearly inseparable, we applied a transformation by mapping the input data (x_i, y_i) into a higher dimensional feature space by using a nonlinear operator $\Phi(x)$. As a result, the optimal hyper-plane can be defined as:

$$f(x) = sgn(\,\Sigma\, y_i \alpha_i K(x_i, x) + b\,) \; (2)$$

Where $K(x_i, x) = exp(-\gamma||x_i - x||^2)$ is the kernel function founded on a radial basis function (RBF), and sgn(.) is the sign function. To conduct out seed purity classification, the RBF kernel SVM classifier model is added to replace the last output layers of the CNN architecture.

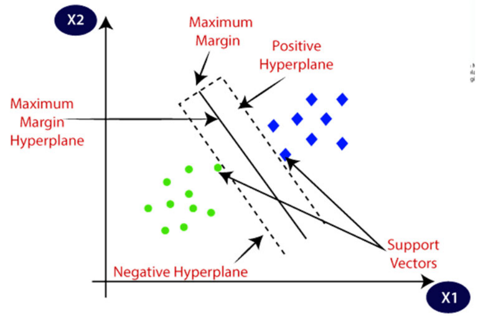

Fig.1.SVM Model

C) K-Nearest Neighbours (KNN) algorithm

The function is only approximated locally in k-NN classification, and all computation is postponed until the function is evaluated. Normalizing the training data can considerably improve the accuracy of this method because it relies on distance for classification. An effective strategy for both classification and regression are to assign weights to the contributions of the neighbours, so that the closer neighbours contribute more to the average than the farther neighbours. Giving each neighbour a weight of 1/d, where d is the distance between them, is a popular weighing approach. The neighbours are chosen from a set of objects that have a known class (in k-NN classification) or object property value (in k-NN regression). This can be regarded the algorithm's training set, even if no formal training phase is required.

The k-nearest neighbour classifier can be thought of as giving the k closest neighbours a weight of 1/k and the rest of the neighbours a weight of 0. This applies to weighted closest neighbour classifiers as well. That is, where the ith nearest neighbour is assigned a weight $\sum_{i=1}^{n} w_{ni} = 1$. A similar finding holds for the strong consistency of weighted closest neighbour classifiers.

Let C_n^{wnn} denote the weighted nearest classifier with weights $\{w_{ni}\}_{i=1}^{n}$. Subject to regularity conditions on the class distributions the excess risk has the following asymptotic expansion.

$$R_R(C_n^{wnn}) - R_R(C^{Bayes}) = (B_1 s_n^2 + B_2 t_n^2)\{1 + 0(1)\},$$

for constants B_1 and B_2 where $s_n^2 = \sum_{i=1}^{n} w_{ni}^2$ and $t_n = n^{\frac{-2}{d}} \sum_{i=1}^{n} \{ i^{1+2/d} - (i-1)^{1+2/d} \}$

D) Hybrid Model (CNN+SVM)

The combined CNN+SVM model's structure will be designed by substituting the last layer of CNN with the SVM model; hence, the output of the fully connected layer of CNN will be used as an input for SVM to improve the classification. The key rationale for combining the benefits of CNN and SVM is that the CNN model has always been important since it allows for quick access to hidden layers, which has led to the extraction of features while also improving accuracy and performance. However, among other algorithms, SVM offers excellent efficiency and speed due to its unique qualities in extracting features. With this combination, we can get the best seed image extraction outcomes with X-Ray images. Fig. 2 demonstrates the structure of the hybrid model of CNN and SVM. In this manner, the CNN model applies different convolution and subsampling and simultaneously convolution layer on basis of 28 × 28 feature map with having 5 × 5 convolution, and in the same way feature map 14 × 14 with having 2 × 2 convolution. This function seeks to speed up the extraction of information from the train and test processes by using it. Following that, the SVM model takes the output of a fully connected layer as input and improves the feature vectors, classification, and decision-making processes.

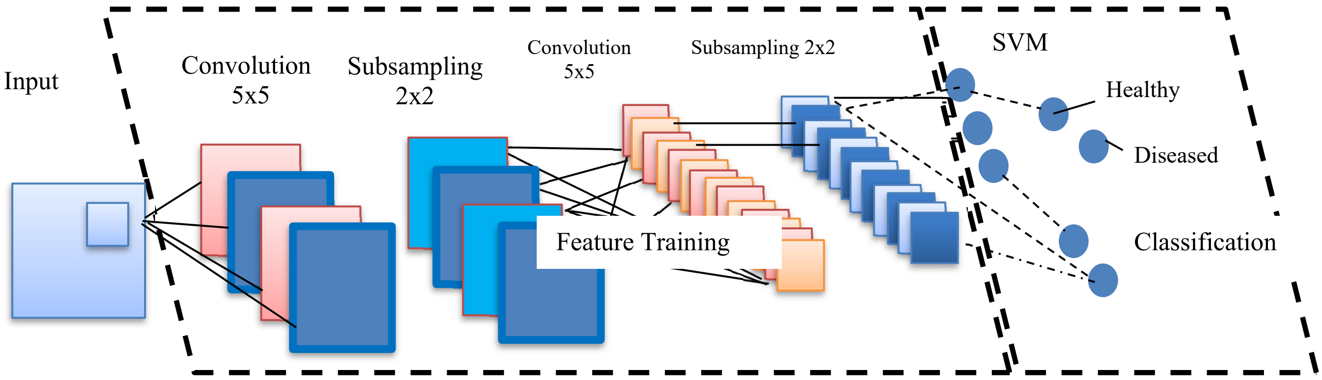

Fig.2 Architecture of proposed Hybrid CNN-SVM Classifier

E) X-RAY SEED PURITY TEST

We proposed a method for investigating internal tissues using X-ray images because seed surface profile can be affected but important internal regions of seeds cannot be reached. A badam plant was chosen as a model species because it is a multipurpose crop with global economic significance.

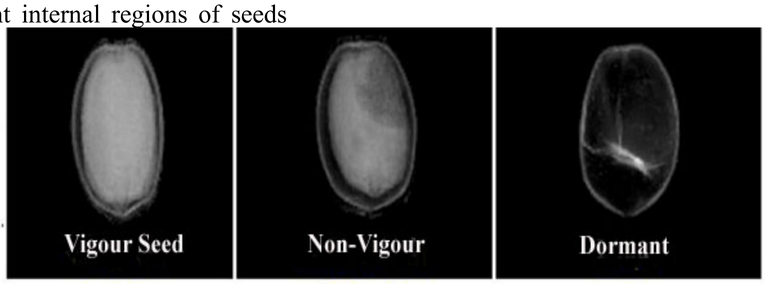

Fig. 3. Based on germination capacity, X-Ray pictures of the seeds' ventral and dorsal surfaces were acquired.

The tegument of the badam seed is thick and black. Figure 4 shows X-Ray images of the seeds' ventral and dorsal surfaces, as well as corresponding reflectance images collected at 940 nm (Wavelength Range) images, based on germination capacity.

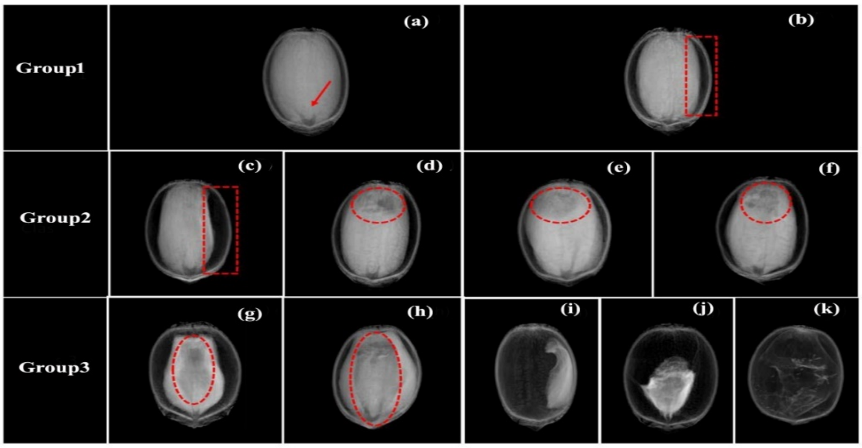

Fig. 4 Badam seeds were divided into three classes based on tissue integrity in X-ray images: Group 1–(a) tissues completely filling the seed [arrow points to the embryonic axis region], (b) slight empty spaces (≤ 1.23 mm) between the endosperm and the seed coat; Group 2–(c) large empty spaces (> 1.24 mm) between the endosperm and the seed coat; (d–f) deteriorated tissues without reaching the embryonic axis; Group 3– (g, h) deteriorated tissues reaching the embryonic

Based on seed tissue integrity and seed performance in the germination test, radiographic images from all seed lots were divided into three groups (Fig. 3). Because soft tissues are connected with damaged tissues and absorb the X-ray beam less as it passes through them, these areas look black on radiographic images. Meanwhile, high grey intensity areas suggest increased X-ray penetration, which is linked to higher tissue density (healthy tissues). Seeds in Group 1 were entirely filled or had small gaps between the endosperm and the seed coat (≤ 1.23 mm), and these seeds mostly produced normal seedlings. Group 2 seeds had huge empty spaces (> 1.24 mm) between the endosperm and the seed coat, as well as degraded tissues that did not reach the embryonic axis, and they produced largely aberrant seedlings. Finally, seeds with damaged tissues reaching the embryonic axis, as well as deformed and empty seeds, were all linked to dead seeds. As illustrated below, the weight of pure seed is divided by the overall weight of the working sample.

A purity percentage is calculated as:

Purity (%) = Weight of pure seeds (g) x 100 /Total weight of working sample (g)

F) GERMINATION

1. Definitions

Seed germination is a physiological process in which a latent seed with low moisture content shows an increase in its general metabolic activity and initiates seedling production from an embryo. The process by which a seed can generate an organism / healthy Seedling that led to a healthy plant is known as seed germination. Seed germination is the return of active embryonic growth that results in the seed coat rupture and the emergence of a young plant under favourable conditions. Or Seed germination is the process of a seed's embryo becoming activated and growing into a new seedling.

2.Calculation and Reporting results

The average of 4x100 seed replicates is used to compute the germination test result. It is calculated as a proportion of the total number of normal seedlings. To the nearest full number, the percentage is calculated. In the same approach, the percentage of aberrant seedlings, hard, fresh, and dead seeds is calculated. These should be entered in the proper space on the certificate analysis. If any of these categories have a 'nil' outcome, it will be reported as '0,'

The sum total %age of all the category of seeds (normal, abnormal, dead, hard, fresh ungerminated) should be equal to 100.

$$Germination\ (\%) = \frac{\text{Number of normal seedlings}}{No.\ Of\ Seed\ placed\ for\ test} \times 100$$

3. RESULTS AND DISCUSSION

Several extensive experiments were performed to fairly evaluate the performance and prove the effectiveness of our suggested method for processing seed purity identification and classification over a broad variety of seeds. The proposed work is implemented in Python, a scientific programming language on the Windows 10 operating system. An i7 processor with 16GB RAM is included in the system setup. The suggested model took 67 hours to train using the stated parameters. An android device operating on Android Pi with 6GB RAM and a 12-megapixel camera was utilised to collect images. White light was used for lighting, and the seeds were placed on a white background. Many conventional machine learning classifiers were used for classification, but these three hybrid classifiers named KNN, SVM, CNN and CNN-SVM performed well on verity seed hybrid features image mining datasets as well.

Model training has three distinct features: image capture, decayed learning rate, and model check pointing. In general, the trained models perform well on the training set but fail miserably when exposed to the test set. Overfitting is one of the main reasons for this. An image mining technique was used to avoid overfitting and improve the generalisation capability of machine learning models. Various images were artificially created during seed image collection using various processing methods or a combination of multiple processing methods, such as random rotation and shifts. Figure 5 depicts a sample of colour and X-Ray images collected from all seed classes considered.

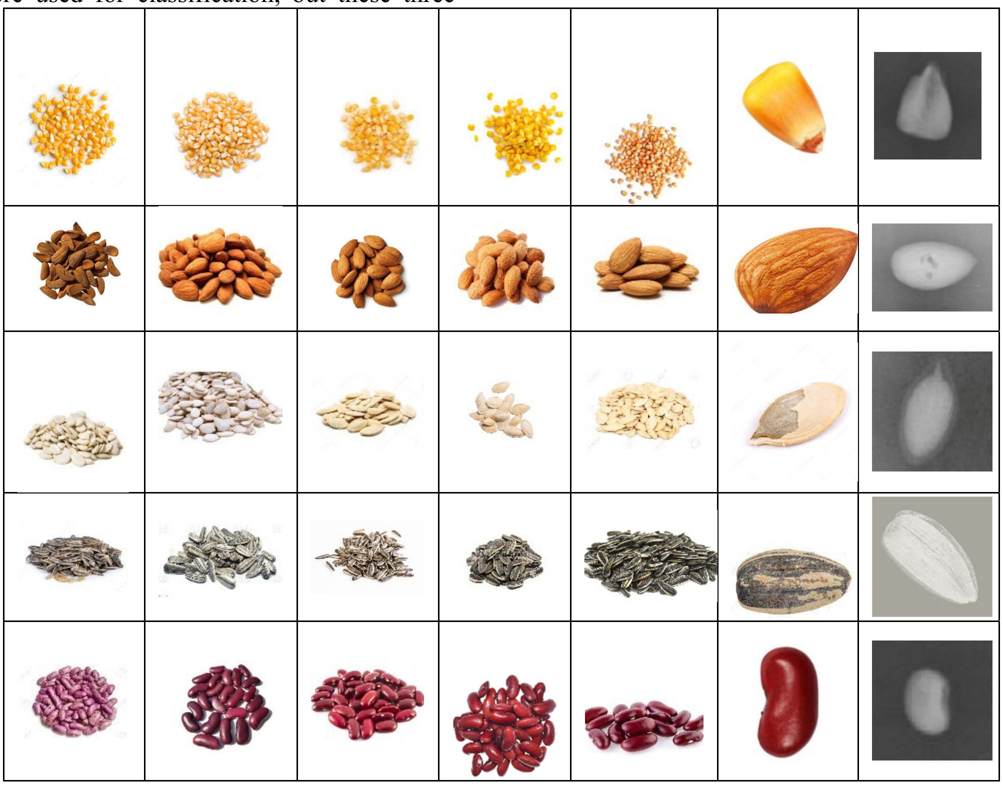

Fig 5 Sample of Collected seed images (Including X-Ray images)

A. CLASSIFICATION RESULTS OF DIFFERENT MODELS

Discriminant models were built by KNN, SVM, CNN and CNN-SVM models. As per the proposed methodology, after pre-processing, the extracted feature with con-sideration of 30% training and 70% testing is going to process and set in specific tests and train category as shown in Fig.2. In this process, as shown in Fig.2, the CNN model, extract the features, and then use the activation function to move to a fully connected layer in various steps. The CNN classifier will be trained using the extracted picture features. The SVM model, on the other hand, uses the CNN's fully linked output as an input to train each image. Finally, the accuracy of the proposed hybrid CNN-SVM is assessed using the assessment specified parameter.

(a)Accuracy: The accuracy of seed image classification is a percentage that represents the total quantity of correctly classified pixels divided by the total number of pixels in the image. It calculates the total number of properly pixeled pixels in an image.

$$Accuracy = \frac{TP + TN}{(TP + TN + FP + FN)}$$

(b)PPV:In addition to accuracy, the true positive (TP) indicates the likelihood of correctly categorised pixels.

$$Positive\ Predictive\ Value = \frac{TP}{(TP + FP)}$$

(c) FPV: True negative (TN) is the possibility pixel verification that rationally reveals normal but is categorised as an abnormal feature in image analysis due to underuse of accuracy and TP.

$$False\ Predictive\ Value\ = \frac{FP}{(FP + TP)}$$

B. RESULT ANALYSIS

The proposed technique got better outcomes compared with other models. In Table1, as shown the accuracy of existing techniques compared with the proposed model. The classification based on seed with KNN received 59.86 percent, SVM received 74.77 percent, CNN received 97.67 percent, and the hybrid model received 98.62 percent. If we look at the graph in Fig.8, we can observe that the accuracy level of CNN and SVM independently does not exhibit good development, which is supported by the information in Table1. At the same time, the proposed hybrid CNN and SVM accuracy degrees both show significant progress.

However, FPV is another crucial element for determining whether or not an activity is being performed correctly. Figure 6 indicates that when CNN and SVM were employed separately, they did not produce a superior outcome, but the suggested technique did so with greater accuracy.

Simultaneously, PPV, which indicates the degree of correct age classification. In fact, a detailed examination of CNN and SVM reveals that the output is unsatisfactory and requires further refinement. The proposed model has made a positive stride, as indicated in Fig.7, and it has improved at a faster rate.

Table 1: Comparison of other models with proposed Hybrid CNN and SVM.

Model	KNN	SVM	CNN	Hybrid (SVM+CNN)
Accuracy	59.86	74.77	97.67	98.62
F1 score	0.623	0.782	0.981	0.989

Along with the suggested technique with a specific evaluation parameter, as demonstrated in Fig. 5, 6, 7, other aspects such as FPV and PPV in the hybrid model have been considerably superior outcomes in comparison to independently CNN and SVM. Furthermore, as shown in Fig.8, when comparing the suggested research job to existing research works, the proposed work has a high rate, and we may have higher classification output.

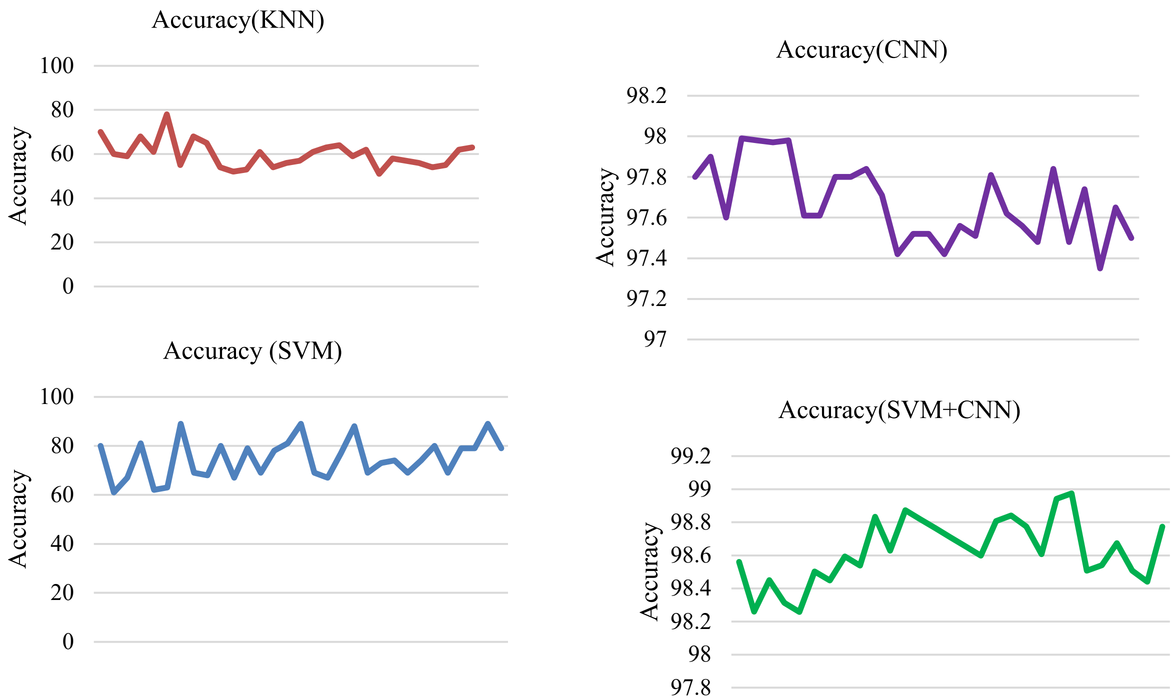

Fig.5.Accuracy comparison of KNN, SVM, CNN, Hybrid (CNN+SVM) model.

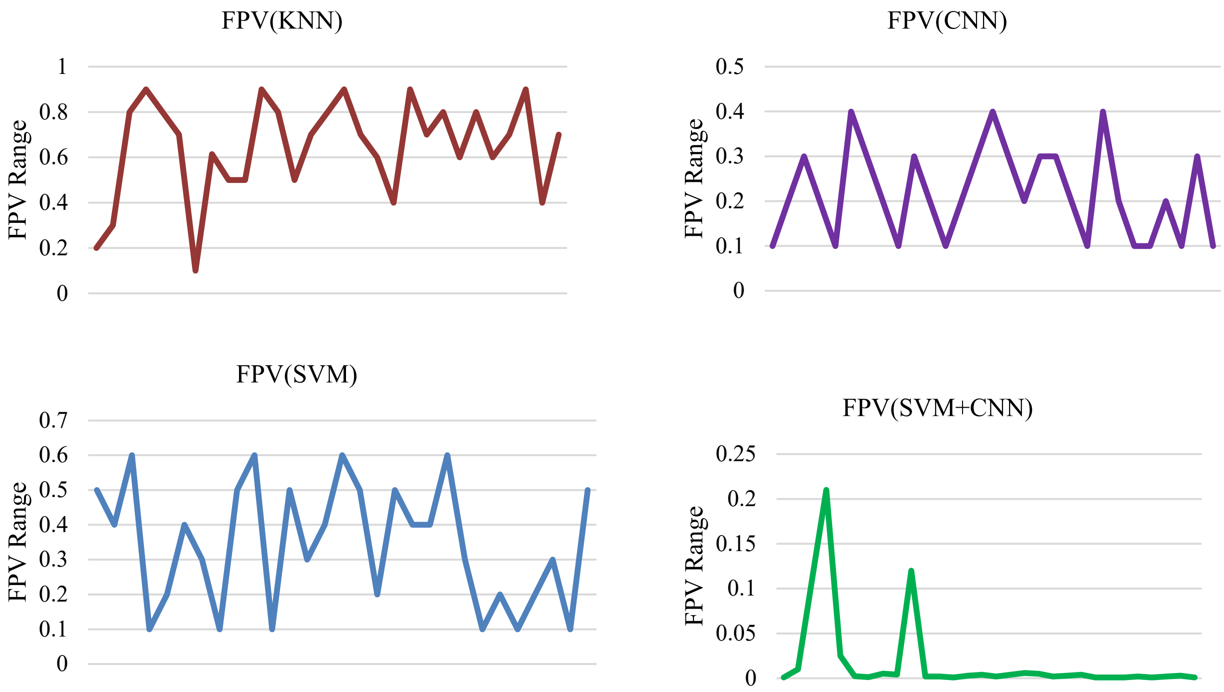

Fig 6.FPV comparison on KNN, SVM, CNN, Hybrid (CNN+ SVM) model.

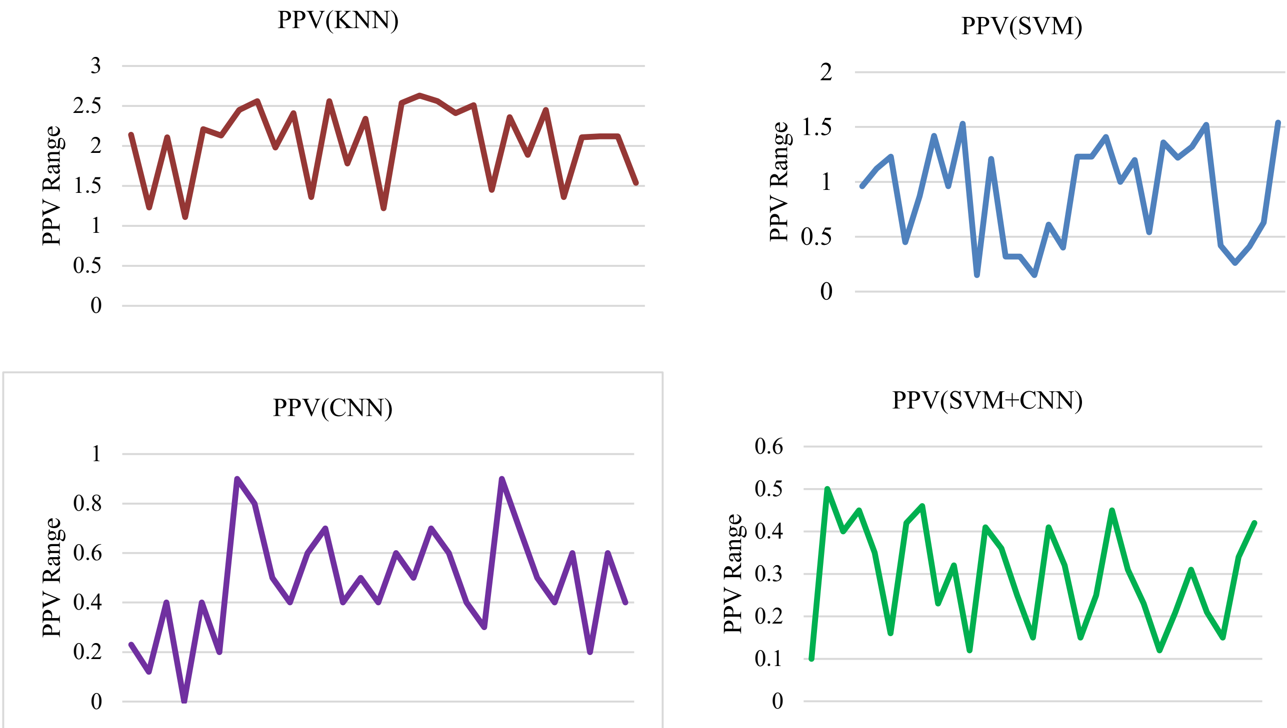

Fig 7. PPV comparison on KNN, SVM, CNN, Hybrid (CNN+ SVM) model

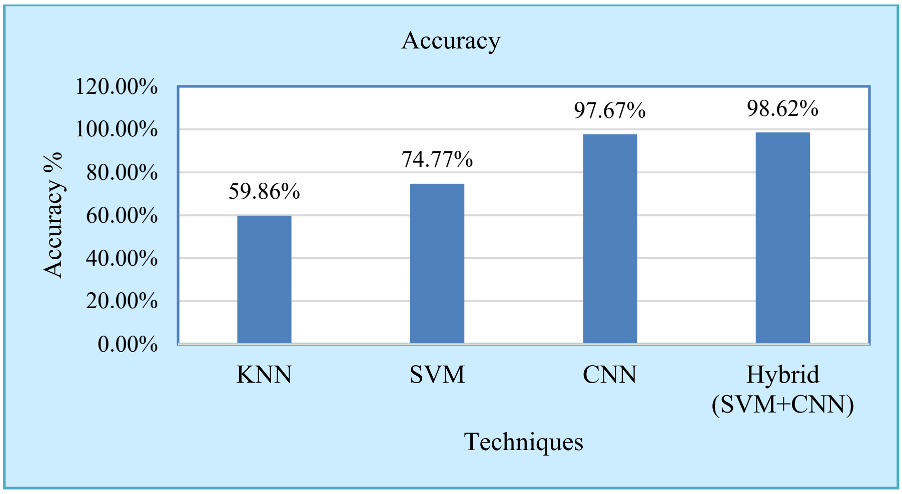

Fig 8. Comparison of classification accuracy with another model

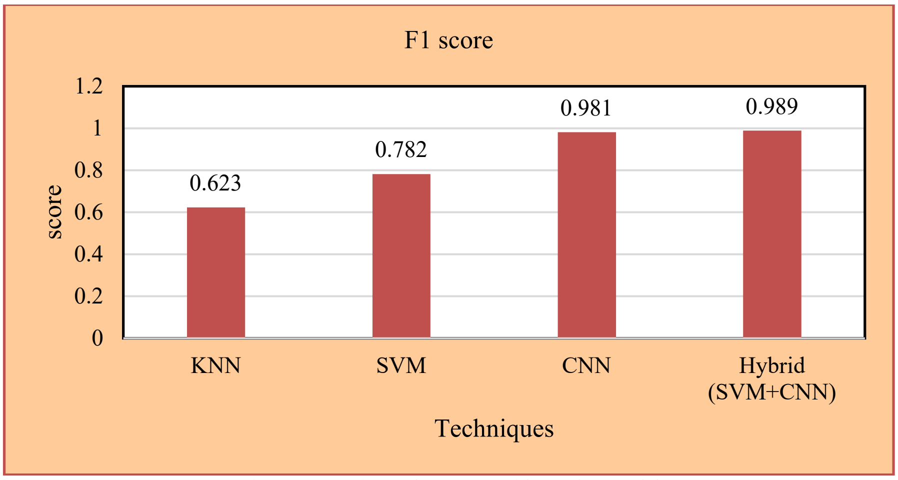

Fig 9. Comparison of F1 score with another model

4. CONCLUSIONS AND FUTURE WORK

This paper developed an effective model for identifying and classifying seed purity. The model is trained using images of seed variants from a dataset. The seed images were reduced in size to 224 x 224 x 3. Performance of three machine learning methods, namely KNN, SVM, CNN and CNN-SVM were tested. The influence of the number of training samples was also studied. As the size of training set increase, CNN-SVM models outperformed the other two models. The hybrid model reached 98.62 percent properly categorised with PPV and the lowest with FPV, and it achieved a total of 59.86 % correctly classified with KNN, 74.77 % with SVM, and 97.67 % with CNN. With a brief look came to know that the proposed hybrid model provides more effective and improvement techniques for classification. In future work, we intend to create a mobile-based application that will use a larger number of different seeds (than the 14 types considered in this study), resulting in a broader range of seed classification. This application would benefit people with limited knowledge of seeds because it is difficult to obtain appropriate advice from agriculture experts or experienced farmers to manually identify, sort, or classify the seeds.

References

[1] Santos, Luís, Filipe N. Santos, Paulo Moura Oliveira, and Pranjali Shinde. "Deep learning applications in agriculture: A short review." In Iberian Robotics conference, pp. 139-151. Springer, Cham, 2020.

[2] Barman, Utpal, Ridip Dev Choudhury, Diganto Sahu, and Golap Gunjan Barman. "Comparison of convolution neural networks for smartphone image based real time classification of citrus leaf disease." Computers and Electronics in Agriculture 177 (2020): 105661.

[3] Hu, Kun, Zhiyong Wang, Guy Coleman, Asher Bender, Tingting Yao, Shan Zeng, Dezhen Song, Arnold Schumann, and Michael Walsh. "Deep Learning Techniques for In-Crop Weed Identification: A Review." arXiv preprint arXiv:2103.14872 (2021).

[4] Morelli-Ferreira, Francielle, Nayane Jaqueline Costa Maia, Danilo Tedesco, Elizabeth Haruna Kazama, Franciele Morlin Carneiro, Leticia Bernabe Santos, Getulio Freitas Seben Junior, Glauco Souza Rolim, Luciano Shozo Shiratsuchi, and Rouverson Pereira Silva. "Comparison of machine learning techniques in cotton yield prediction using satellite remote sensing." (2021).

[5] Qian, Yan, Qianjin Xu, Yingying Yang, Hu Lu, Hua Li, Xuebin Feng, and Wenqing Yin. "Classification of rice seed variety using point cloud data combined with deep learning." International Journal of Agricultural and Biological Engineering 14, no. 5 (2021): 206-212.

[6] Kiratiratanapruk, Kantip, Pitchayagan Temniranrat, Wasin Sinthupinyo, Panintorn Prempree, Kosom Chaitavon, Supanit Porntheeraphat, and Anchalee Prasertsak. "Development of paddy rice seed classification process using machine learning techniques for automatic grading machine." Journal of Sensors 2020 (2020).

[7] Luo, Tongyun, Jianye Zhao, Yujuan Gu, Shuo Zhang, Xi Qiao, Wen Tian, and Yangchun Han. "Classification of weed seeds based on visual images and deep learning." Information Processing in Agriculture (2021).

[8] Luan, Zhengguang, Chunlei Li, Shumin Ding, Miaomiao Wei, and Yan Yang. " Sunflower seed sorting based on convolutional neural network." In Eleventh International Conference on Graphics and Image Processing (ICGIP 2019), vol. 11373, p. 113731K. International Society for Optics and Photonics, 2020.

[9] Tatsumi, Kenichi, Noa Igarashi, and Xiao Mengxue. "Prediction of plant-level tomato biomass and yield using machine learning with unmanned aerial vehicle imagery." (2021).

[10] Taylor, Jared, Chien-Ping Chiou, and Leonard J. Bond. "A methodology for sorting haploid and diploid corn seed using terahertz time domain spectroscopy and machine learning." In AIP Conference Proceedings, vol. 2102, no. 1, p. 080001. AIP Publishing LLC, 2019.

[11] Ali, Aqib, Salman Qadri, Wali Khan Mashwani, Samir Brahim Belhaouari, Samreen Naeem, Sidra Rafique, Farrukh Jamal, Christophe Chesneau, and Sania Anam. "Machine learning approach for the classification of corn seed using hybrid features." International Journal of Food Properties 23, no. 1 (2020): 1110-1124.

[12] Sarker, Iqbal H. "Deep learning: a comprehensive overview on techniques, taxonomy, applications and research directions." SN Computer Science 2, no. 6 (2021): 1-20.

[13] Li, Zewen, Fan Liu, Wenjie Yang, Shouheng Peng, and Jun Zhou. "A survey of convolutional neural networks: analysis, applications, and prospects." IEEE Transactions on Neural Networks and Learning Systems (2021).

[14] Ruslan, Rashidah, Siti Khairunniza-Bejo, Mahirah Jahari, and Mohd Firdaus Ibrahim. "Weedy Rice Classification Using Image Processing and a Machine Learning Approach." Agriculture 12, no. 5 (2022): 645.

[15] Shedole, Seema, and B. J. Sowmya. "A Convolution Neural Network-Based Wheat Grain Classification System." Journal of Scientific Research 66, no. 2 (2022).

[16] Margapuri, Venkat, Niketa Penumajji, and Mitchell Neilsen. "Seed Classification using Synthetic Image Datasets Generated from Low-Altitude UAV Imagery." arXiv preprint arXiv:2110.02846 (2021).

[17] Bashier, Israa Hassan, Mayada Mosa, and Sharief Fadul Babikir. "Sesame Seed Disease Detection Using Image Classification." In 2020 International Conference on Computer, Control, Electrical, and Electronics Engineering (ICCCEEE), pp. 1-5. IEEE.

[18] Xu, Peng, Ranbing Yang, Tiwei Zeng, Jian Zhang, Yunpeng Zhang, and Qian Tan. "Varietal classification of maize seeds using computer vision and machine learning techniques." Journal of Food Process Engineering 44, no. 11 (2021): e13846.

[19] Ashqar, Belal AM, Bassem S. Abu-Nasser, and Samy S. Abu-Naser. "Plant seedlings classification using deep learning." (2019).

Ms. M. Suganthi, received M. Phil. Degree in Computer Science from Bharathidasan University, Trichy, Tamilnadu, India in 2014. She completed M.C.A. Degree in Computer Application from Bishop Heber College, Bharathidasan University, Trichy, Tamilnadu, India in 2013. Now, she is doing her Ph.D. in Computer Science from Bishop Heber College, Bharathidasan University, Tiruchirappalli. Her area of research includes Image Mining, Machine Learning, and Deep Learning. Email: sugan.swt09@gmail.com.

Dr. J. G. R. Sathiaseelan is the Associate Professor and Head, Department of Computer Science at Bishop Heber College, Trichy. He has 30 years of teaching experience. He has presented more than 35 research papers in international conference publication which are published in ACM, IEEE, Springer, Elsevier, and reputed journals. He has authored a book entitled as, “Programming in C#, DotNet”, which was published in PHI, New Delhi, in 2009. His research areas include Data Mining, Web services security, Image Processing and Big Data Analytics. E-mail: jgrsathiaseelan@gmail.com.