



CLASSIFICATION OF FIVE DIFFERENT RICE SEEDS GROWN IN TURKEY WITH DEEP LEARNING METHODS

Bulent TUGRUL¹

¹Department of Computer Engineering, Ankara University, Ankara, TURKEY

ABSTRACT. The increase in the world population and harmful environmental factors such as global warming necessitate a change in agricultural practices with the traditional method. Precision agriculture solutions offer many innovations to meet this increasing need. Using healthy, suitable and high-quality seeds is the first option that comes to mind in order to harvest more products from the fields. Seed classification is carried out in a labor-intensive manner. Due to the nature of this process, it is error-prone and also requires a high budget and time. The use of state-of-the-art methods such as Deep Learning in computer vision solutions enables the development of different applications in many areas. Rice is the most widely used grain worldwide after wheat and barley. This study aims to classify five different rice species grown in Turkey using four different Convolutional Neural Network (CNN) architectures. First, a new rice image dataset of five different species was created. Then, known and widely applied CNN architectures such as Visual Geometry Group (VGG), Residual Network (ResNet) and EfficientNets were trained and results were obtained. In addition, a new CNN architecture was designed and the results were compared with the other three architectures. The results showed that the VGG architecture generated the best accuracy value of 97%.

1. INTRODUCTION

Since cereals are a source of protein, carbohydrates and fat, they have a crucial role in continuing of human life. Rice (*Oryza sativa* (Asian rice) or *Oryza glaberrima* (African rice)), which is one of the most fundamental food sources in the world, has an essential place in the food culture of many countries, especially far eastern countries [1]. It is estimated that there are thousands of rice types grown by farmers. There are 33 registered rice species in Turkey, but only 15 of them are widely cultivated [2]. Some species have more commercial value in terms of their nutritional value, taste, appearance and structure. The most cultivated species is

Keywords. Convolutional neural networks, VGG, ResNet, EfficientNet.

✉ btugrul@eng.ankara.edu.tr-Corresponding author

ORCID 0000-0003-4719-4298.

Osmancık due to its high yield. Some other rice species produced in Turkey are Ergene, İpsala, Altinyazi, Yatkın, Cameo, Demir, Sarıklık and Özgür.

The commercial value of rice depends on its genetic characteristics and quality. Currently, the classification of seeds is carried out visually by experienced employees. This process, which is both laborious and slow, can cause financial losses because it cannot be done with adequate precision. Furthermore, even a skilled person can provide accurate and reliable quality control on only a few common rice varieties. If this problem can be solved with an automated process based on a computer vision system, it will provide a more accurate and stable seed classification. Thus, the incomes of both farmers and countries will increase with the increase in the quality of the crops to be harvested. Besides, the natural variety of seeds makes the process of classification with computer vision a challenging problem.

Computer vision gives computers the ability to recognize the world around them through visual inputs. Thus, the necessary actions can be taken. Basically, a computer vision system has two components. The first component is a sensing device that captures an image or video. The other is called an interpreting device that processes input and generates an output. Collaboratively, they both try to understand and interpret the objects and the events that occur around them.

Computer vision has been improving rapidly in parallel with the great developments in Artificial Intelligence (AI) and Deep Learning (DL) in recent years [3]. It finds widespread application areas in many fields, from autonomous vehicles to disease detection. Moreover, DL has reached the capability to create new things like new human faces that didn't exist before [4]. A traditional computer vision pipeline necessitates the following steps; visual data acquisition, pre-processing, feature extraction and model building. DL, on the contrary, eliminates the feature extraction phase from the traditional computer vision pipeline. DL methods can learn representations through a series of data transformations in their successive layers.

The main objective of this study is to create a new rice dataset that did not exist before and to compare the performances of various CNN architectures on this dataset. For this purpose, seeds of five different types of rice, which are widely grown and consumed in Turkey, were obtained from distinct locations. The images of the seeds were captured using a digital microscope. A database was created by randomly selecting approximately a thousand images of each type.

1.1. Related Studies. Before the widespread use of DL methods, the seed classification problem was solved with traditional computer vision methods. Huang and Chien extracted the shape features of rice seeds and trained the collected dataset using a Multilayer Neural Network to classify them [5]. The highest accuracy value was achieved in the Tainan-11 type with 97.35%. Ali et al. derived all morphological, texture, and color features of six different rice species and applied fuzzy logic methods [6]. According to their results, the lowest accuracy value is 94.2% and the highest value is 98.9%.

There are also numerous comparative studies using traditional machine learning and DL methods. Qui et al. captured hyperspectral images of four rice species at two different spectral ranges. KNN, SVM and CNN models were chosen as a classifier [7]. As expected, the CNN model gave the best results over the other methods. They also concluded that more accurate CNN models could be built by increasing the number of samples. Kiratiratanapruk et al. classified 14 rice varieties with four statistical machine learning methods (LR, LDA, k-NN, and SVM) and five different CNN architectures (VGG16, VGG19, Xception, InceptionV3, and InceptionResNetV2) [8]. The SVM method gives the highest accuracy value among machine learning methods with 90.61%. On the other hand, InceptionResNetV2 model produces a 95.15% accuracy value much more successful than other models.

In parallel with the developments in DL methods and technology, various CNN architectures have started to be preferred in solving the seed classification problem. Hoang et al. classified six rice types using eleven different CNN architectures [9]. DenseNet with 121 layers has a 99.05% accuracy value. Despite this, MobileNet gave the worst results, training fewer parameters than other CNN models. Gilanie et al. proposed their CNN model called RiceNet to classify Pakistani grown seven different rice seeds [10]. They also trained VGG-19, ResNet50, and GoogleNet(Inception-V3) to compare the results with their model. RiceNet has succeeded in accurately classifying all types of rice with 100% success.

2. CONVOLUTIONAL NEURAL NETWORKS

Recent developments in Artificial Intelligence and Deep Learning have allowed computer vision systems to build more accurate and reliable classifier systems than a human can do. DL methods, such as CNN, enhance the accuracy of classification models by using a large data set and sufficient computing power. CNN consists of a variety of layers, such as input, convolution, pooling, fully connected, softmax and output. In a DL architecture, the first layer that receives raw pixels of input data learns how to represent simple features. Each successive layer learns the complex features of the previous layer as they collect and recombine the features of the previous layers. Adding more layers to the network allows it to handle higher-dimensional data.

Basically, a CNN architecture has three types of layers that can be seen in almost any model.

- Convolutional layer: Extracts features from visual inputs.
- Pooling layer: Reduces the number of parameters passed to the next layer.
- Fully connected layer: A Multi-layer Perceptron (MLP) network that takes features learned from the previous layer as input and returns class labels as output.

2.1. Visual Geometry Group (VGG). Although VGG is not the winner of the 2014 ImageNet Large-Scale Visual Recognition Challenge (ILSVRC) competition,

it has been used in many applications due to its modular and simpler architecture [11]. Unlike previous CNNs, VGG prefers smaller kernel sizes. This property causes VGG to increase its expressability. VGG architecture has several different variations according to the number of layers. The smallest one (VGG-11) has 11 layers and 133 million parameters. The largest one (VGG-19) consists of 19 layers and requires 144 million parameters to be tuned. The architecture of the VGG is displayed in Figure 1.

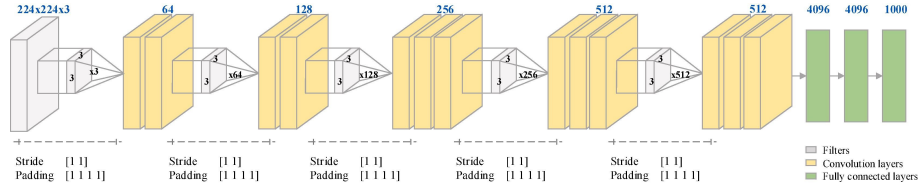


FIGURE 1. The architecture of VGG [12].

2.2. Residual Network (ResNet). ResNet, which won all five categories of the ImageNet Challenge in 2016, was proposed by Microsoft Research Team [13]. The top-5 error rate of ResNet is 3.57%. With the addition of more layers to deep learning models, the vanishing gradient problem arises. ResNet tackles this problem by devising a residual module with a skip connection. This approach allows them to train networks with 50, 101, and 152 layers without increasing the number of parameters.

2.3. EfficientNets. EfficientNets suggests a new scaling technique that scales all dimensions using a compound coefficient [14]. Previous CNN models prefer to scale only one of the three dimensions. They designed seven different versions of the model called EfficientNets. Top-1 accuracy of EfficientNet-B7 is 84.3% on ImageNet. They also conducted experiments using five different learning data sets, including CIFAR-100. The results show that the proposed method can perform better using fewer parameters. The architecture of EfficientNet-B0 is given in Figure 2.

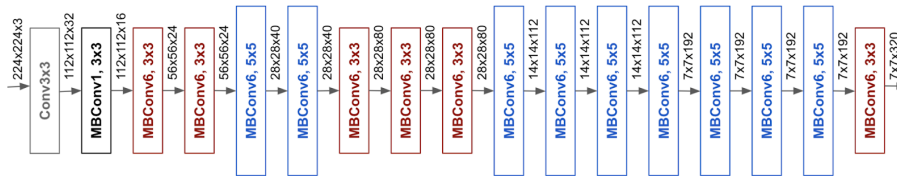


FIGURE 2. The architecture of EfficientNet-B0.

2.4. The Custom CNN Model. In addition to the three popular CNN architectures preferred in computer vision problems, a new CNN model was designed. The new model consists of a total of 16 layers. The first layer is the input layer used to feed the network. Images with a size of 66x64x3 will be fed through this layer. Then there are successive convolutional and pooling layers. Finally, there are the fully connected, softmax and output layers. The architecture of the custom CNN architecture is given in Figure 3. In addition, a detailed description of the custom architecture can be found in Table 1.

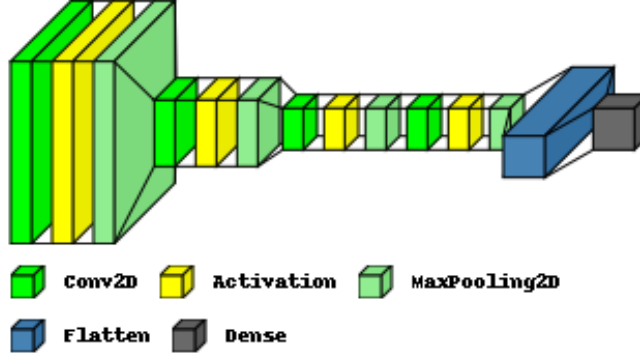


FIGURE 3. The architecture of the custom CNN model.

3. METHODOLOGY

3.1. Image Acquisition. The rice seed images were acquired using a digital microscope with a spatial resolution of 0.01 mm/pixel and a color depth of 24 bit. The images were captured under natural lighting conditions. There are 6833 images of Turkey’s five most extensively grown rice seed species. These species are Osmançık, Cameo, Özgür, Sarıkılçık and Yatkın. Between 703 and 1736 images of each type were randomly selected to ensure that the dataset was balanced proportionately. An expert on rice seed labelled all seed images according to their class. A sample of each rice type is displayed in Figure 4.

3.2. Performance Metrics. The success of classification models should be evaluated using various criteria. Thus, the results of different architectures will be compared and the most successful one will be selected. Accuracy is the most common and frequently used criterion. The accuracy formula is as follows;

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \quad (1)$$

TABLE 1. Description of the custom CNN model.

Layer #	Operation layer		Stride value	Activation
1	Input Image			64x64x3
2	Convolutional Layer	Convolutional	3x3	22x22x64
3		ReLU		22x22x64
4	Pooling Layer	Max Pooling	1x1	22x22x64
5	Convolutional Layer	Convolutional	3x3	8x8x256
6		ReLU		8x8x256
7	Pooling Layer	Max Pooling	1x1	8x8x256
8	Convolutional Layer	Convolutional	3x3	3x3x256
9		ReLU		3x3x256
10	Pooling Layer	Max Pooling	1x1	3x3x256
11	Convolutional Layer	Convolutional	3x3	1x1x256
12		ReLU		1x1x256
13	Pooling Layer	Max Pooling	1x1	1x1x256
14	Inner Product Layer	Fully Connected		1x1x5
15	Softmax Layer	Softmax		1x1x5
16	Classification Layer	Classification Output		5

TABLE 2. Description of rice seed image dataset.

Rice Species	Number of Images
Cameo	1736
Osmancık	1246
Özgür	703
Sarıklıçık	1456
Yatkın	1692

Accuracy alone is not a sufficient criterion in some special cases. In addition, Precision and Recall values are also preferred to evaluate the performance and robustness of classification models. F1-Score is the geometric mean of both Precision and Recall.

$$Precision = \frac{TP}{TP + FP} \quad (2)$$

$$Recall = \frac{TP}{TP + FN} \quad (3)$$

$$F1 - Score = \frac{2 * Precision * Recall}{Precision + Recall} = \frac{2 * TP}{2 * TP + FP + FN} \quad (4)$$

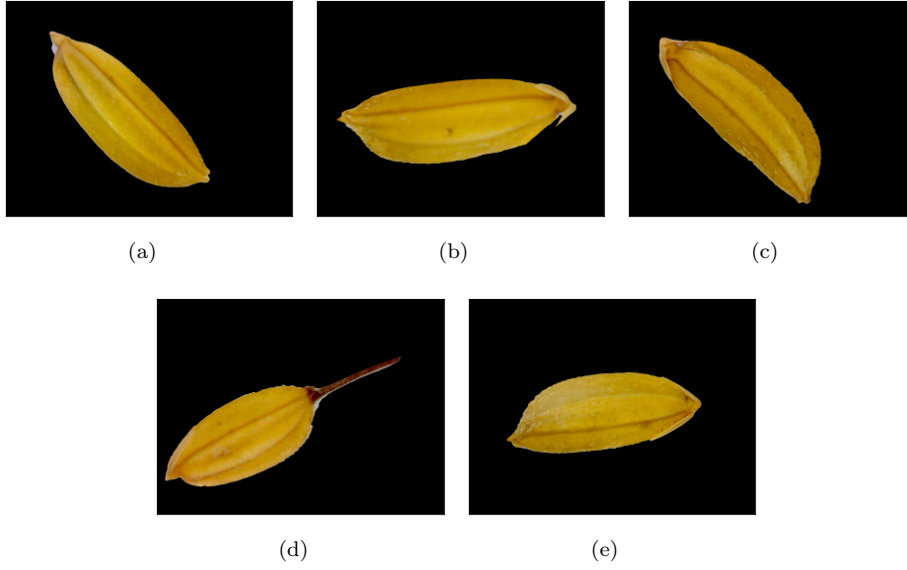


FIGURE 4. Sample images of different species of *Rice* (a) *Cameo* (b) *Osmancık* (c) *Özgür* (d) *Sarıkılçık* (e) *Yatkın*.

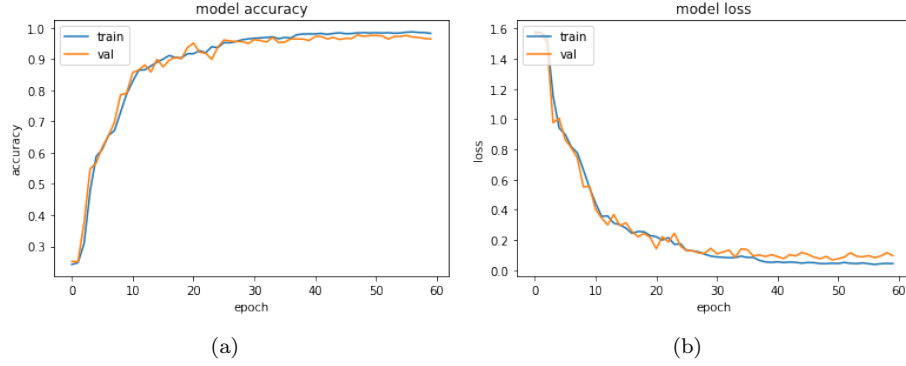


FIGURE 5. VGG.

4. RESULTS AND DISCUSSION

Classification of rice seeds is carried out by traditional methods that require heavy human labor. However, rice seed types are very similar in terms of shape, texture and color, so the probability of making mistakes is high. Recently, many studies

have been published showing that DL methods successfully classify different types of seeds. For this purpose, an image data set of five different species grown in Turkey was created. This dataset was trained with three different popular and custom CNN architecture. The accuracy and loss graphics of each CNN architecture's training and validation data are displayed in Figures 5 , 6, 7 and 8. The confusion matrix and overall performance of the CNN models are presented in Tables 3 and 4, respectively.

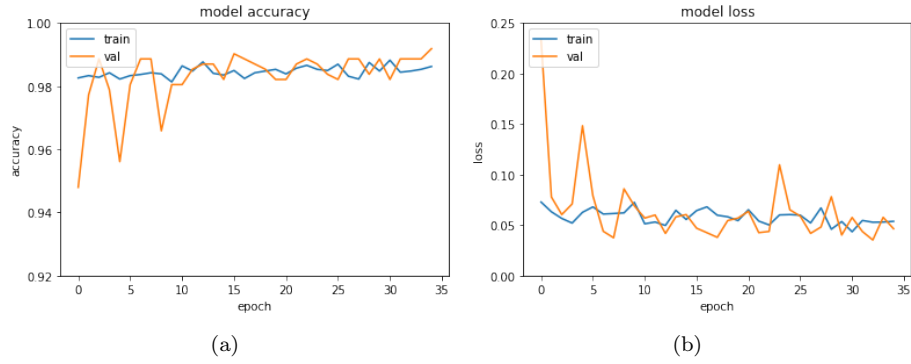


FIGURE 6. ResNet.

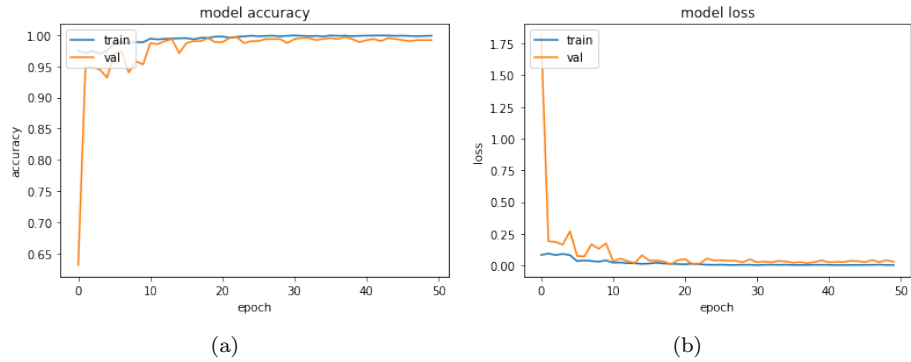


FIGURE 7. EfficientNet.

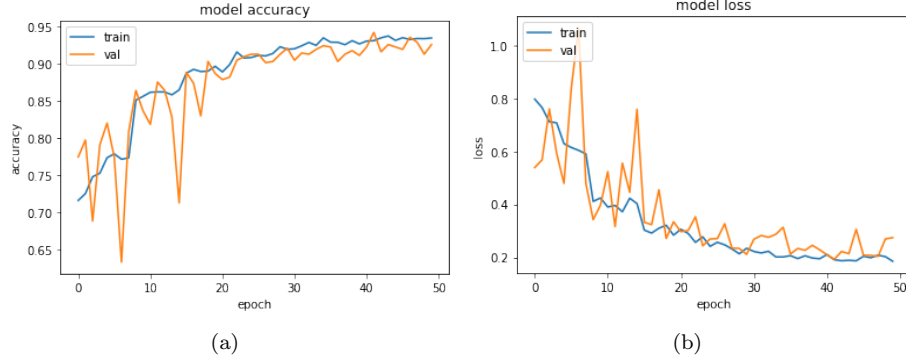


FIGURE 8. The Custom CNN model.

TABLE 3. Confusion matrix of CNN architectures.

Rice Species	CNN Architectures	Cameo	Osmancık	Özgür	Sarıklıçık	Yatkın
Cameo	VGG	165	0	1	0	3
	ResNet	169	0	0	0	0
	Efficient	168	0	0	1	0
	Custom	159	0	0	4	6
Osmancık	VGG	0	128	2	0	0
	ResNet	1	113	16	0	0
	Efficient	0	97	33	0	0
	Custom	0	127	3	0	0
Özgür	VGG	0	2	63	1	0
	ResNet	7	0	59	0	0
	Efficient	2	0	63	1	0
	Custom	3	2	59	2	0
Sarıklıçık	VGG	0	0	0	145	0
	ResNet	8	0	0	137	0
	Efficient	1	0	0	144	0
	Custom	0	0	0	145	0
Yatkın	VGG	9	0	0	0	165
	ResNet	41	0	0	0	133
	Efficient	84	0	0	0	90
	Custom	8	0	0	0	166

TABLE 4. Overall performance of CNN architectures.

CNN Architectures	Accuracy	Precision	Recall	F1-Score
VGG	0.97	0.97	0.97	0.97
ResNet	0.89	0.91	0.89	0.89
EfficientNet	0.82	0.88	0.82	0.82
Custom	0.96	0.96	0.96	0.96

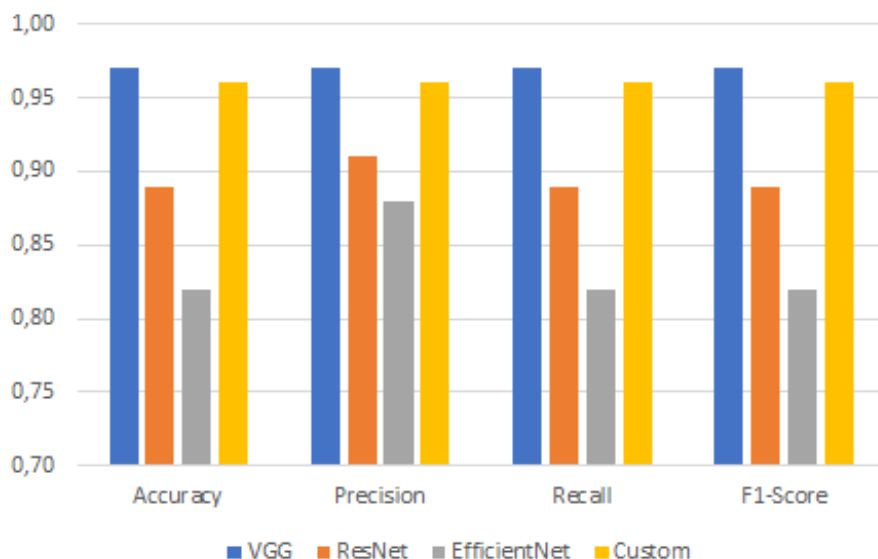


FIGURE 9. Overall performance of CNN architectures.

As shown from Tables 3 and 4, VGG and the custom CNN architecture generated successful and robust results. The best performance belongs to VGG and has an accuracy value of 0.97. The custom model is more successful than both ResNet and EfficientNet. Even though it is worse than VGG, its performance is promising as a classifier.

5. CONCLUSIONS AND FUTURE WORKS

The main objective of this study is to classify the rice seeds data set with three different and a new custom CNN architecture. For this purpose, a data set consisting of five different types of rice grown in Turkey was created. According to the results, it was observed that the CNN architecture with the best accuracy value was VGG. The other architectures have also generated acceptably close results. Thus, it was shown that the classification of rice seeds could be carried out more successfully with advanced computer vision methods. This study used only images of five different seed types and four different CNN architectures. In future studies, we aim to broaden the data set by increasing both the number of images and the number of rice species. We also plan to get results using the Vision Transformers (ViT).

Declaration of Competing Interests The author declare that he has no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

REFERENCES

- [1] Chen, C., He, W., Nassirou, T. Y., Nsabayumva, A., Dong, X., Adedze, Y. M. N., Jin, D., Molecular characterization and genetic diversity of different genotypes of *Oryza sativa* and *Oryza glaberrima*, *Electron. J. Biotechnol.*, 30 (2017), 48–57, <https://doi.org/10.1016/j.ejbt.2017.08.001>.
- [2] Taşlıgil, N., Şahin, G., Türkiye’de çeltik (*Oryza sativa* L.) yetiştiriciliği ve coğrafi dağılımı, *Adıyaman Üniversitesi Sosyal Bilimler Enstitüsü Dergisi*, 6 (2011), 182–203, <https://doi.org/10.14520/adyusbd.105>.
- [3] Voulodimos, A., Doulamis, N., Doulamis, A., Protopapadakis, E., Deep learning for computer vision: A brief review, *Comput. Intell. Neurosci.*, 2018 (2018), <https://doi.org/10.1155/2018/7068349>.
- [4] Creswell, A., White, T., Dumoulin, V., Arulkumaran, K., Sengupta, B., Bharath, A. A., Generative adversarial networks: An overview, *IEEE Signal Process Mag.*, 35 (1) (2018), 53–65, <https://doi.org/10.1109/MSP.2017.2765202>.
- [5] Huang, K., Chien, M., A novel method of identifying paddy seed varieties, *Sensors*, 17 (4) (2017), 809–816, <https://doi.org/10.3390/s17040809>.
- [6] Ali, T., Jhandhir, Z., Ahmad, A., Khan, M., Khan, A. A., Choi, G. S., Detecting fraudulent labeling of rice samples using computer vision and fuzzy knowledge, *Multimed. Tools Appl.*, 76 (23) (2017), 24675–24704, <https://doi.org/10.1007/s11042-017-4472-9>.
- [7] Qiu, Z., Chen, J., Zhao, Y., Zhu, S., He, Y., Zhang, C., Variety identification of single rice seed using hyperspectral imaging combined with convolutional neural networks, *Appl. Sci.*, 8 (2) (2018), 212–223, <https://doi.org/10.3390/app8020212>.
- [8] Kiratiratanapruk, K., Temniranrat, P., Sinthupinyo, W., Prempre, P., Chaitavon, K., Porntheeraphat, S., Prasertsak, A., Development of paddy rice seed classification process using machine learning techniques for automatic grading machine, *J. Sens.*, 2020 (2020), <https://doi.org/10.1155/2020/7041310>.
- [9] Hoang, V. T., Van Hoai, D. P., Surinwarangkoon, T., Duong, H. T., Meethongjan, K., A comparative study of rice variety classification based on deep learning and hand-crafted features, *ECTI Transactions on Computer and Information Technology (ECTI-CIT)*, 14 (1) (2020), 1–10, <https://doi.org/10.37936/ecti-cit.2020141.204170>.
- [10] Gilanie, G., Nasir, N., Bajwa, U. I., Ullah, H., RiceNet: Convolutional neural networks-based model to classify Pakistani grown rice seed types, *Multimed. Syst.*, 27 (5) (2021), 867–875, <https://doi.org/10.1007/s00530-021-00760-2>.
- [11] Simonyan, K., Zisserman, A., Very deep convolutional networks for large-scale image recognition, *arXiv preprint arXiv:1409.1556*, (2014), <https://doi.org/10.48550/arXiv.1409.1556>.
- [12] Altuntaş, Y., Cömert, Z. Kocamaz, A. F., Identification of haploid and diploid maize seeds using convolutional neural networks and a transfer learning approach, *Comput. Electron. Agric.*, 163 (2019), 104874, <https://doi.org/10.1016/j.compag.2019.104874>.
- [13] He, K., Zhang, X., Ren, S., Sun, J., Deep residual learning for image recognition, *Proceedings of the IEEE conference on computer vision and pattern recognition*, (2016), 770–778.
- [14] Tan, M., Le, Q., Efficientnet: Rethinking model scaling for convolutional neural networks, *International conference on machine learning*, (2019), 6105–6114.