Computers and Electronics in Agriculture 163 (2019) 104874

Contents lists available at ScienceDirect

Computers and Electronics in Agriculture

journal homepage: www.elsevier.com/locate/compag

Original papers

Identification of haploid and diploid maize seeds using convolutional neural networks and a transfer learning approach

Yahya Altuntaş[a,*], Zafer Cömert[b], Adnan Fatih Kocamaz[a]

[a] İnönü University, Department of Computer Engineering Malatya, Turkey
[b] Samsun University, Department of Software Engineering, Samsun, Turkey

ARTICLE INFO

Keywords:
Haploid
Diploid
Maize
Convolutional neural network
Classification

ABSTRACT

Maize is one of the most significant grains cultivated all over the world. Doubled-haploid is an important technique in terms of advanced maize breeding, modern crop improvement and genetic programs, since this technique shortens the breeding period and increases breeding efficiency. However, the selection of the haploid seeds is a major problem of this breeding technique. This process is frequently conducted manually, and this unreliable situation leads to loss of time and labor. Inspired by the recent successes of deep transfer learning, in this study, we approached this problem as a computer vision task to provide a nondestructive, rapid and low-cost model. To achieve this objective, we adopted convolutional neural networks (CNNs) to recognize haploid and diploid maize seeds automatically through a transfer learning approach. More specifically, AlexNet, VVGNet, GoogLeNet, and ResNet were applied for this specific task. The experimental study was carried out using a new dataset consisting of 1230 haploid and 1770 diploid maize seed images. The samples in the dataset were classified considering a marker-assisted selection, known as the *R1-nj* anthocyanin marker. To measure the success of the CNN models, we utilized several performance metrics, such as accuracy, sensitivity, specificity, quality index, and F-score derived from the confusion matrix and receiver operating characteristic curves. According to the experimental results, the CNN models ensured promising results, and we achieved the most efficient results via VGG-19. The accuracy, sensitivity, specificity, quality index, and F-score of VGG-19 were 94.22%, 94.58%, 93.97%, 94.27%, and 93.07%, respectively. Consequently, the experimental results proved that CNN models can be a useful tool in recognizing haploid maize seeds. Furthermore, we conclude that this approach is significantly superior to machine learning-based methods and conventional manual selection.

1. Introduction

Maize (*Zea mays L.*) is one of the most significant agricultural products used as human food, animal feed and industrial raw materials (Cerit et al., 2016). A growing world population and climate change make it necessary to develop new maize varieties that are high-yield and resistant to biotic and abiotic stress conditions like all other cultivated plants. Achieving this goal is only possible through maize breeding programs. The first stage in maize breeding programs is to develop homozygote lines that will be parents to hybrid varieties (Cerit et al., 2016). Normally, the acquisition of homozygote lines takes a long time, approximately five to eight generations of self-cross mating by conventional methods, whereas this process can be achieved in about two to three generations in one year through the use of haploids (Prasanna et al., 2012). Haploids and doubled haploids (DH) have high importance in modern maize breeding, since this technique accelerates the breeding period and increases breeding efficiency (Chase and Nanda, 1969). These advantages of DH have led to increased interest in maize breeding and genetics in the last 20 years (Geiger, 2009).

A DH is a completely homozygous line produced by doubling the haploid chromosomes (Prasanna et al., 2012). Haploids are found in nature at a very small frequency of 0.1% (Geiger et al., 2013); therefore, they are not suitable for practical use (Charity et al., 2017). Haploids can be obtained at higher rates by using either *in vitro* or *in vivo* techniques. Most commercially available DH maize lines are obtained through the haploid technique *in vivo* while other techniques are reported to be less effective in the development of DH lines (Geiger, 2009). *In vivo* maternal haploid induction uses special genotypes, called inducers, as pollinators to obtain haploids at higher rates and has become the standard method (Charity et al., 2017). Due to currently

* Corresponding author at: Samsun University, Department of Software Engineering, Canik Yerleşkesi Gürgenyatak Mahallesi Merkez Sokak No:40-2/1, 55080 Canik, Samsun, Turkey.
E-mail addresses: yahyaaltuntas@gmail.com (Y. Altuntaş), zcomert@samsun.edu.tr (Z. Cömert), fatih.kocamaz@inonu.edu.tr (A.F. Kocamaz).

https://doi.org/10.1016/j.compag.2019.104874
Received 9 January 2019; Received in revised form 2 May 2019; Accepted 24 June 2019
Available online 26 June 2019

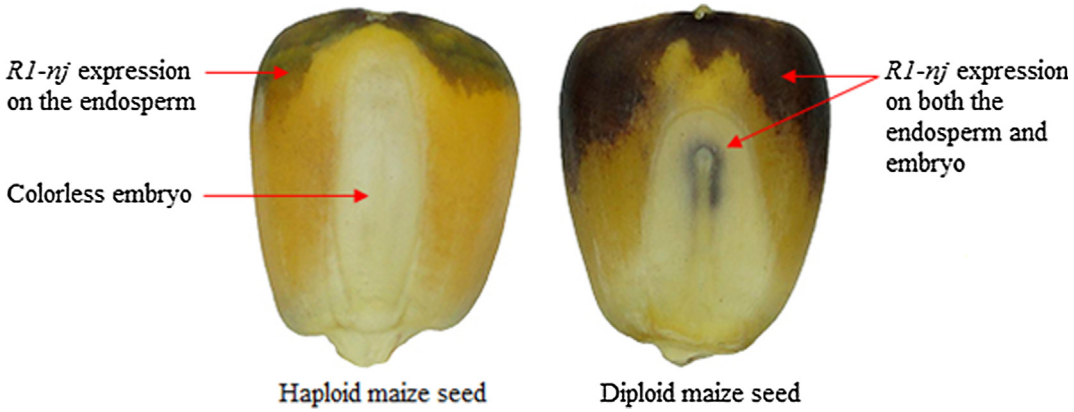

Fig. 1. Visual classification of haploid and diploid maize seeds according to the *R1-nj* color marker.

available inducers, the DH rate has been observed to be approximately 8% or higher (Melchinger et al., 2014). *In vivo* maternal haploid induction consists of four stages: (i) haploid induction, (ii) haploid identification, (iii) chromosome doubling and (iv) self-pollination to generate the final DH line.

Among the induced seeds before chromosome doubling, the haploids, which are a small proportion, need to be separated from the diploids because only haploids are used in chromosome doubling. This reality emphasizes the importance of the selection of haploid seeds quickly and precisely (Geiger et al., 2013). Although there are alternative selectable markers, such as the differences between the embryo weights (Smelser et al., 2014), the oil content levels (Melchinger et al., 2014) and red root (Chaikam et al., 2016), the most commonly and successfully used selectable marker is *R1-Navajo* (*R1-nj*) (Nanda and Chase, 1966). Dominant *R1-nj* gene expression leads to purple coloration of the aleurone and scutellum due to anthocyanin pigmentation (Melchinger et al., 2014). The different levels of *R1-nj* expression in the embryo and endosperm allow visual classification of haploid and diploid maize seeds as illustrated in Fig. 1. The seeds with red-purple coloration on both the embryo and endosperm are called diploids. Since these diploids contain chromosomes of both the source genotype and the inducer line, they have 2*n* chromosomes. Although *R1-nj* staining occurs on endosperms, seeds with colorless embryos are called haploids. Genetically, these seeds carry only the chromosomes of the donor plant and have a single *n* chromosome (Geiger, 2009). The selection of haploid and diploid seeds on the basis of *R1-nj* marker gene expression works well for most dent germplasms (Melchinger et al., 2014). Nonetheless, the expression of *R1-nj* varies from a small point to the entire seed. In addition, the darkness of *R1-nj* expression changes from very light to dark. This diversity in *R1-nj* expression leads to a high error rate in conventional manual selection. In addition, manual selection of haploids is a labor-intensive process that takes time. Therefore, there is a need for a system to automatically perform the identification of haploid seeds (Melchinger et al., 2013).

A sorting system of maize haploid seeds consisting of seed transmission, image acquisition and processing, sorting-unloading and a system control unit has been designed. This specific process has been realized considering the color features of maize embryos and the top of the maize endosperm. The system achieved a sorting speed of 500 seeds per minute (Song et al., 2010).

A method based on the least square error has been suggested to separate haploid from diploid maize seeds considering the oil content of the seeds measured using a nuclear magnetic resonance (NMR) analyzer. According to the results of the model, the least square error can quickly determine the thresholding value of oil content between haploid and diploid with a low number of samples (Li et al., 2016).

As a pattern recognition method, support vector machine (SVM) and visible spectroscopy (Vis) have been utilized for identifying maize haploid seeds. An experiment has been conducted on a balanced dataset, which consisted of 141 haploids and 141 hybrid kernels from nine genotypes. The authors have reported promising results (Liu et al., 2015).

A combination of a near-infrared spectroscopy technology and supervised virtual sample kernel locality preserving projection (SVSKLPP) has been proposed to separate haploid seeds from the hybrid seeds (L. Yu et al., 2018). The model ensured a strong classification performance due to its nonlinear structure, and it is superior to linear methods, such as principal component analysis, orthogonal linear discriminant analysis, and locality preserving projection.

In another work, the haploid and diploid maize seeds have been classified using five features extracted from the embryo and endosperm regions. The features were used to feed an SVM classifier. The average accuracy of the proposed model was 81.36% (Altuntaş et al., 2018a, 2018b).

Haploid and diploid maize seeds have also been identified using texture features obtained from a gray level co-occurrence matrix. These features have been used as input to decision trees (DT), k-nearest neighbors (kNN) and artificial neural networks (ANN) classifier. It was reported that the best performance was measured in DT with a classification success rate of 84.48% (Altuntaş et al., 2018a, 2018b).

An automatic sorting system based on the color features of the endosperm has been proposed to distinguish haploid maize seeds from cross-breeding seeds with Navajo label (Song et al., 2018). Similarly, a novel method using near-infrared hyperspectral imaging technology has been introduced to overcome the limitations of current automated haploid identification and to ensure more sensitive monitoring of haploid (Wang et al., 2018). Haploid and diploid maize seeds were sorted with the proposed fluorescence-based method. The authors have reported that the accuracy of the proposed method is greater than 80% in all seven genotypes used (Boote et al., 2016).

In another study conducted on six maize genotypes, the performance of third-party software was evaluated in the discrimination of haploid and diploid maize seeds via multispectral imaging. DNA markers were used to confirm the labels of the tested seeds. It was reported that the accuracy was over 50% for the majority of genotypes used (De La Fuente et al., 2017).

Recently, deep learning approaches are attracting more attention due to their success in practical applications (Salaken et al., 2019). As a kind of deep learning structure, convolutional neural networks (CNNs) are developed to learn features from data that come in the form of multiple arrays, such as images, video, text or sound (LeCun et al., 2015). The CNN models have achieved serious success on numerous practical applications, such as pattern recognition, image classification, speech recognition, etc. (Tajbakhsh et al., 2016; H. Yu et al., 2018; Yu et al., 2017). There are four key concepts under CNN models: local connections, shared weights, pooling and the use of many layers.

Various feature maps include distinctive motifs of the input data. Sharing weights among units at different locations are inclined to search the same pattern in different parts of the input data. Pooling layer merges semantically similar features into one, decreases the dimension of the representation and prevents overfitting (Yu et al., 2017). Due to the advances in deep learning algorithm, deeper models have been introduced, such as AlexNet (Krizhevsky et al., 2012), VVGNet (Simonyan and Zisserman, 2014), GoogLeNet (Szegedy et al., 2015), and ResNet (He et al., 2016).

In this article, we introduce a new dataset including 1230 haploid and 1770 diploid maize seed images. To ensure a novel haploid maize seed recognition model, we used CNN models, which are AlexNet, VVGNet, GoogLeNet, and ResNet and a transfer learning approach. The CNN models yielded excellent results on the classification task when compared to conventional machine learning-based methods.

The remainder of the paper is organized as follows: the materials and methods are given in Section 2. The experimental results and discussion are presented in Section 3. Finally, concluding remarks are given in Section 4.

2. Materials and methods

2.1. Description of the dataset

The haploid and diploid maize seeds used in this study were harvested in the summer of 2017 as part of a larger project carried out by the Maize Research Institute in Sakarya (Turkey) with coordinates 40 degrees 43 min 56 s North 30 degrees 22 min 40 s East. The seeds included a total of 3000 maize seeds as 1230 haploids and 1770 diploids that were produced by crossing the maternal haploid inducers RWS/RWK76 (Röber et al., 2005) with 107 source genotypes. The source genotypes are in 450–700 (FAO) maturity groups, and these are yellow dent and waxy grain types. The list of the genotypes can be found in the supplementary material. The seeds were selected to reflect the different expression levels of the *R1-nj* color marker (light-dark, dense-less).

Image Acquisition: An imaging system has been built for photographing maize seeds. A camera and a sufficient number of LEDs were installed in the ceiling of the imaging system. All seeds were photographed as 20 seeds (4 rows, 5 columns) in each image with embryo-up position and without contacting each other. A Sony SLT-A58 digital camera and Sony SAL 100 mm f/2.8 macro lens were used to capture the images. The camera was connected to the computer via a USB port and third-party software was used to capture the images. All images were taken in 'M' mode, with 100 ISO, 1/125 s shutter speed, f/9 aperture, 150 mm focal length, and 55 cm camera distance. The image format was JPEG and the resolution was 5456-by-3632 pixels.

Seed Segmentation: Since each image contained 20 maize seeds, as the first task, the maize seeds were segmented from the original images. For this task, maize seeds and background color distributions in the original images were examined in red (R), green (G), and blue (B) color channels. Fig. 2 shows the R, G and B color histograms of a sample maize seed. As can be understood from the histograms, the 140–160 value range in the B channel seems to be ideal for the segmentation process. The appropriate threshold value was searched in this range. It was observed that some embryos were segmented as background under the 150 threshold value, especially in haploid maize seeds. In addition, it was observed that, in some cases, the background pixels caused the noise around the maize seeds to be above the 150 threshold value. Therefore, the threshold value was empirically adjusted to 150.

According to Eq. (1), a binary image was obtained in which '1' represents the maize seed and '0' represents the background. Each pixel with its Blue component smaller than the threshold value has been designated as a maize seed and the remaining pixels were background.

$$pixel_{i,j} = \begin{cases} maizeseed, & B_{i,j} < threshold \\ background, & otherwise \end{cases} \quad (1)$$

In this equation, $pixel_{i,j}$ represents the value of the ith and jth pixel of the binary image. The $B_{i,j}$ parameter corresponds to the value of the i^{th} and jth pixel of the Blue channel of the original image.

Since some embryo pixels are threshed as background, a filling morphological process was applied to the obtained binary images. Then, a 9×9 median filter was used to reduce the noise. Boundary boxes were calculated by using the contour lines of the maize seeds. Each maize seed was stored as a separate image after removing its background. While the 150 original images were 424 MB in size, the 3000 segmented images were 48.9 MB in size. The process steps are shown in Fig. 3.

Image Resizing: The resolutions of the images in the dataset varied between 300-by-289 pixels and 610-by-637 pixels depending on the sizes of the seeds. The images were resized as 227-by-227 pixels resolution for AlexNet and 224-by-224 pixels resolution for the other CNN models. The bicubic interpolation method was applied to the long edge of the images for this process. Blank pixels were added to both sides of the short edge of the images to prevent distortion of the aspect ratio.

Maize Seeds Labeling: The maize seeds were labeled according to the *R1-nj* color marker in two stages, before and after the image acquisition process, by two field experts. Although a rigorous study has been conducted in the process of labeling maize seeds, it should not be overlooked that there may be mislabeled seeds in the dataset.

2.2. Convolutional neural networks

A CNN model is a deep learning architecture that consists of a set of successive layers, such as convolution, pooling, dropout, and fully connected layers that have different tasks in the architecture. The most basic layer in these architectures is the convolution layers providing local discriminative features by connecting each node to a small subset of spatially connected neurons in the input image channels (Tajbakhsh et al., 2016). The connection weights are shared among the nodes in the convolutional layers to allow searching for the same discriminative feature throughout the input channels. Each set of shared weights explains a convolution kernel. In this manner, a convolutional layer with

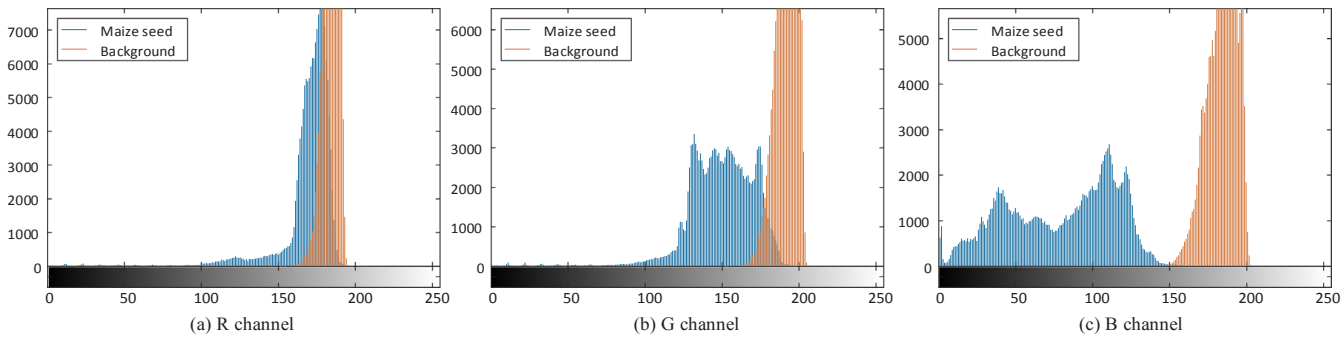

Fig. 2. R, G, and B color histograms of a sample maize seed.

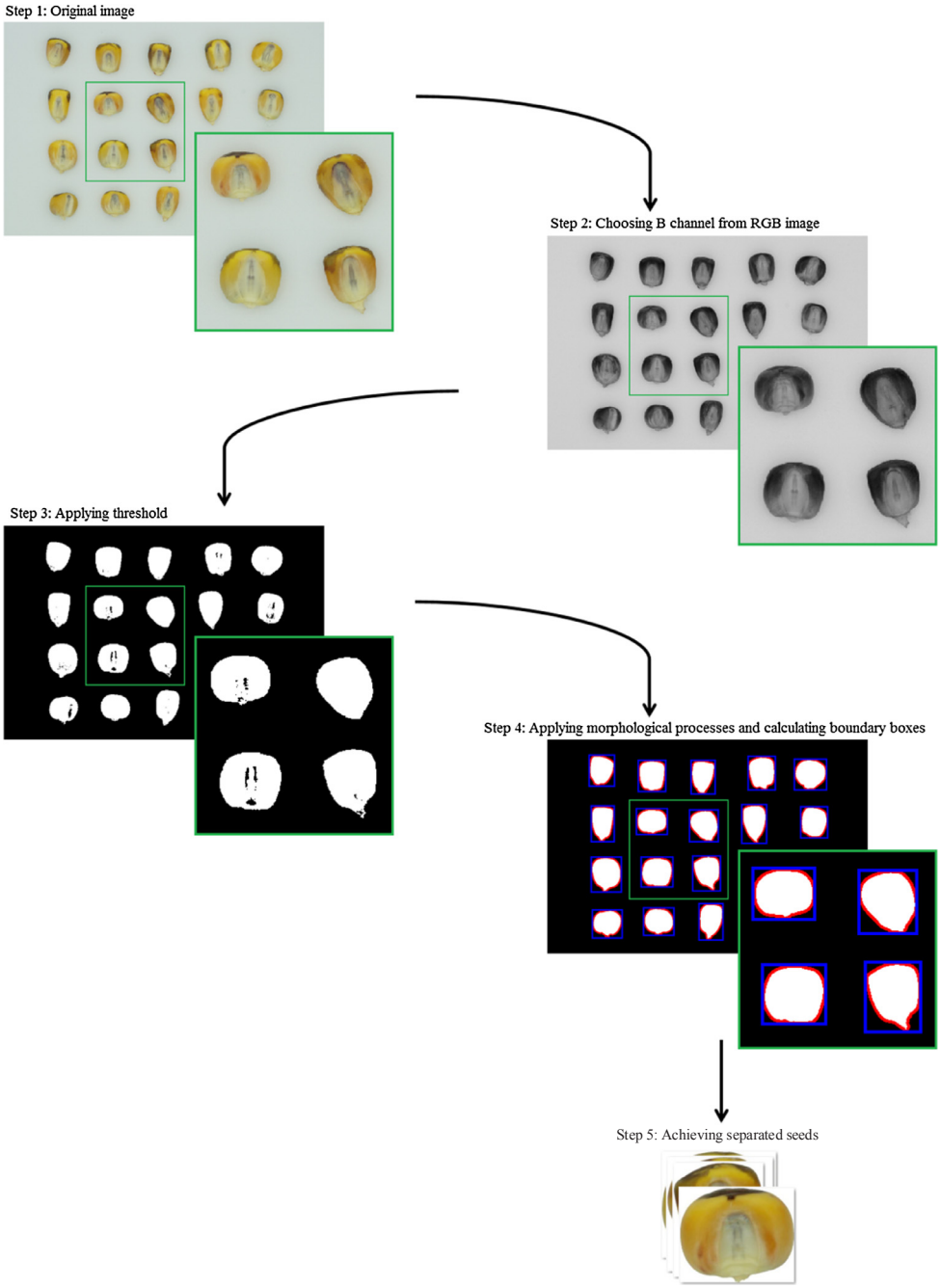

Fig. 3. The segmentation steps.

*n*kernels can learn to recognize *n* local discriminative features that are visible in the resulting *n* feature maps. To decrease the computational complexity, prevent overfitting and ensure a hierarchical set of image features, each sequence of convolution layers is followed by a pooling layer. The main task of the pooling layer is to simplify the spatial dimensions of the information derived from the feature maps. For this particular purpose, average pooling, L2-norm pooling, and especially max-pooling were utilized due to their speed and improved convergence (Traore et al., 2018). CNN models typically include the fully connected layer that desires a vector of numbers as input to connect each of them to outputs. Frequently, this is the last layer, after the last pooling layer in the CNN process. However, the number of fully connected layers may vary according to the networks' architecture. Finally, a softmax or regression layer is used to produce the desired outputs in the CNN architecture.

Back-propagation algorithms are frequently used in training CNNs in order to minimize the cost function in respect to the unknown weights as described below:

$$\mathcal{L} = -\frac{1}{|X|}\sum_{i}^{|X|} ln(p(y^i|\, X^i)) \tag{2}$$

Herein, $|X|$ shows the number of training images, X^i shows the i^{th} training image with the corresponding label y^i, and $p(y^i|\, X^i)$ shows the probability by which X^i is accurately matched. Normally, stochastic gradient descent (SGD) or stochastic gradient descent with momentum (SGDM) algorithms are utilized for minimizing this cost function (Cömert and Kocamaz, 2017). Let us assume that W_l^t shows the weights in the l^{th} convolutional layer at iteration t, and $\widehat{\mathcal{L}}$ indicates the cost over a mini-batch of size N, and then the equations given below are followed to update the weights of the CNN model for the next iteration:

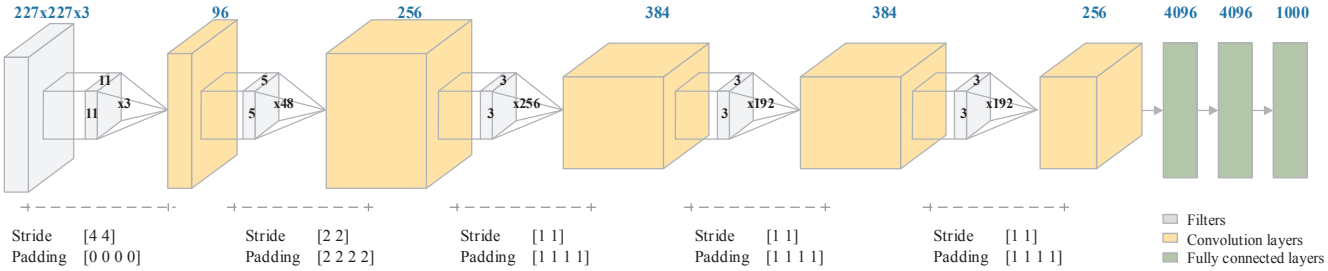

Fig. 4. The architecture of AlexNet.

$$\gamma^t = \gamma^{\lfloor tN/X \rfloor}$$
$$V_l^{t+1} = \mu V_l^t - \gamma^t \alpha_l \frac{\partial \widehat{\mathcal{L}}}{\partial W_l}$$
$$W_l^{t+1} = W_l^t + V_l^{t+1} \quad (3)$$

where α_l stands for the learning rate of the l^{th} layer, μ stands for the momentum that displays the contribution of the previous weight update in the current iteration, and γ stands for the scheduling rate used to reduce α at the end of each epoch.

So far, we have tried to summarize the common structures of CNN models. In the remainder of this section, we present an overview of the well-known CNN models.

AlexNet is one of the pioneering deep CNN algorithms that was introduced in the ImageNet Large Scale Visual Recognition Challenge (ILSVRC-2012), and it achieved a TOP-5 test accuracy of 84.6% (Krizhevsky et al., 2012). AlexNet has a simple architecture that consists of five convolution layers, some of which are followed by rectified linear unit (ReLU), norm, and the pooling layers as shown in Fig. 4. Additionally, there are three fully connected layers in the architecture (Cömert and Kocamaz, 2019).

VGG networks are one of the widely used CNN models. The main aim of this model is to investigate the effect of the convolutional network depth over its accuracy in large-scale image recognition setting. The model possesses two keywords that are design and its depth. The model provides a significant improvement by pushing the depth to 16–19 wt layers (Simonyan and Zisserman, 2014). The architecture of the VGG network is illustrated in Fig. 5.

Another notable deep CNN model is GoogLeNet introduced by Szegedy et al. in the ILSVRC-2014 challenge (Szegedy et al., 2015). The greatest hallmark of this architecture is a new block codenamed "inception module" that consists of a shortcut branch and a few deeper branches. The inception module ensures the improved utilization of the computing resources inside the network. The inception module with dimensionality reduction is shown in Fig. 6. Another noteworthy factor is that this network increases the depth and width of the architecture while keeping the computational cost constant. More specifically, GoogLeNet consists of 22 layers and has 7 million computational parameters. This means that this network uses 12 times fewer parameters compared to AlexNet and exhibits a higher performance (Garcia-Garcia et al., 2018).

The ResNet model puts forward a residual learning framework to ease the training of the networks. The focus of the model is the degradation problem (Lin et al., 2013). The novelty of the model arises with residual blocks and depth in its architecture. In a conventional deep learning model, stacked layers fit a desired underlying mapping, whereas the ResNet model permits these layers to fit a residual mapping (He et al., 2016).

Let us assume that a desired underlying mapping is shown $\mathcal{H}(x)$. ResNet lets the stacked nonlinear layers fit another mapping of $\mathcal{F}(x) := \mathcal{H}(x) - x$. The original mapping is recast into $\mathcal{F}(x) := \mathcal{F}(x) + x$. This hypothesis can be tested by feedforward neural networks with shortcut connections as shown in Fig. 7. In short, the intuitive idea under the model is that the sequential layer learns something new and different from what the input has already encoded (Garcia-Garcia et al., 2018).

As mentioned above, the CNNs have a deep architecture, and this architecture leads to a serious complexity and computational cost. The well-known CNN models are summarized with their properties in Table 1. As inferred from Table 1, AlexNet is the pioneer in this field due to its sample and base architecture. VGGNet has a deeper structure compared to AlexNet, and it has more than 2 times more computation parameters in its architecture. GoogLeNet ensures a proper training time and effective memory usage due to use of the inception module. Compared to the other CNNs, the number of computation parameters in this network is minimal. More and deeper architecture has come up with ResNet by introducing the residual learning strategy.

2.3. Transfer learning and fine-tuning

The CNN models ensure significant facilities by eliminating manual feature extraction, yielding state-of-the-art recognition results, and retraining for new specific tasks. These factors have made these models rather useful. There are two basic approaches for training a CNN model: (1) from scratch and (2) transfer learning. Furthermore, the features obtained from activation maps can also be used to feed machine-learning algorithms (Razavian et al., 2014). It is a fact that if large-scale data resources are available, CNN models can achieve easily generalizable results. However, in real-world applications, achieving large-scale labeled databases is a challenge, sometimes even impossible. Training a CNN model from scratch requires many samples of the objects to reveal the discriminative features. Although this approach gives us the most control over the network, the training times last too long. In addition, overfitting and convergence issues are the potential problems

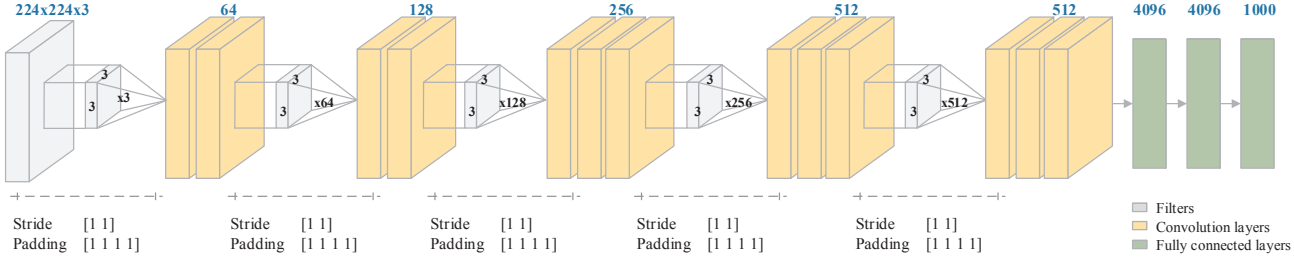

Fig. 5. The architecture of VGG networks.

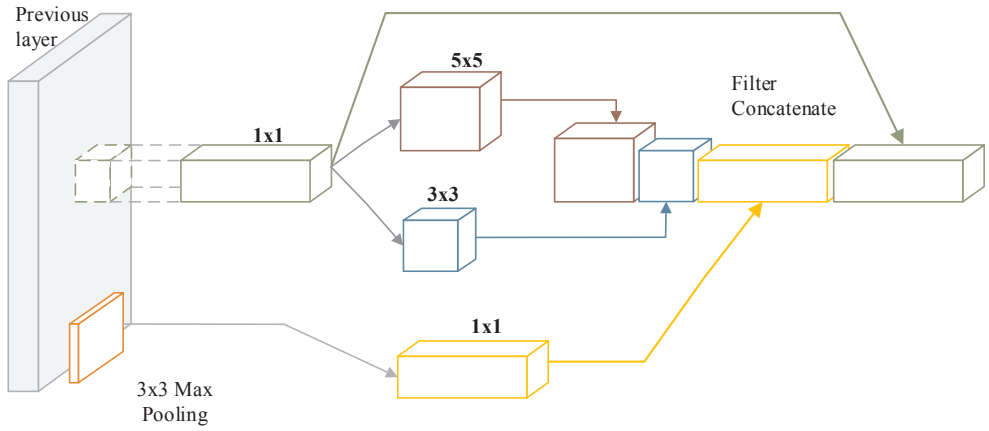

Fig. 6. Inception module with dimensionality reductione.

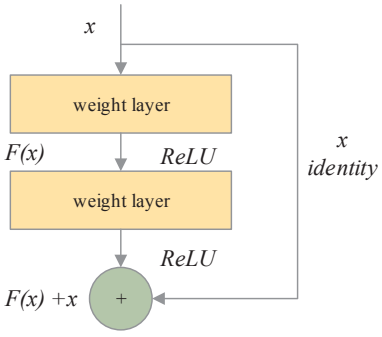

Fig. 7. Residual learning: a building block.

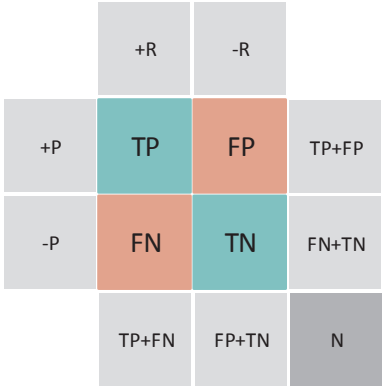

Fig. 8. Confusion matrix.

Table 1
The pretrained CNN models with properties.

Network	Depth	Parameters (Millions)	Image Input Size
AlexNet (Krizhevsky et al., 2012)	8	61	227-by-227
VGG-16 (Simonyan and Zisserman, 2014)	16	138	224-by-224
VGG-19 (Simonyan and Zisserman, 2014)	19	144	224-by-224
GoogLeNet (Szegedy et al., 2015)	22	7	224-by-224
ResNet-18 (He et al., 2016)	18	11.7	224-by-224
ResNet-50 (He et al., 2016)	50	25.6	224-by-224
ResNet-101(He et al., 2016)	101	44.6	224-by-224

Table 2
The parameters for fine-tuning the weights of the pretrained CNN models.

Parameters	Values
Max Epochs	64
Mini Batch Size	64
Initial Learn Rate	0.0001
Momentum	0.95
Weight Learn Rate Factor	10
Bias Learn Rate Factor	10
Learn Rate Schedule	Piecewise
Learn Rate Drop Factor	0.1
Learn Rate Drop Period	16

that may be faced in this process (Penatti et al., 2015).

To address these issues, CNN models can be retrained by using fine-tuning, and this approach is called transfer learning. Transfer learning is a quite convenient and effective method for knowledge adaptation

Table 3
The performance metrics for measuring the performance of the CNN models.

Metrics	Formulation	Short Description
Accuracy (Acc)	$\frac{TP+TN}{TP+FP+FN+TN}$	The overall efficiency of a model
Sensitivity (Se)	$\frac{TP}{TP+FN}$	The efficiency of a model on positive samples.
Specificity (Sp)	$\frac{TN}{TN+FP}$	The efficiency of a model on negative samples.
Quality Index (QI)	$\sqrt{Se * Sp}$	The geometric mean of Se and Sp.
F-score	$\frac{(2 * TP)}{(2 * TP + FP + FN)}$	The harmonic mean between precision and recall.
AUC	$\frac{1}{2}\left(\frac{TP}{TP+FN} + \frac{TN}{TN+FP}\right)$	The models' power to prevent misclassification.

Table 4
Data for training and testing.

Training set		Testing set	
Haploid	Diploid	Haploid	Diploid
861	1239	369	531

(Salaken et al., 2019). In a common transfer learning approach, convolutional layers are used as fixed features extractors, and only fully connected layers are set for a new specific task (Mazo et al., 2018). The fine-tuning starts by transferring the weights from a pretrained network to a new network that we want to train. First, the last layers, including the fully connected, softmax and classification layers of the networks are frequently separated from the network. Then, a new configuration is

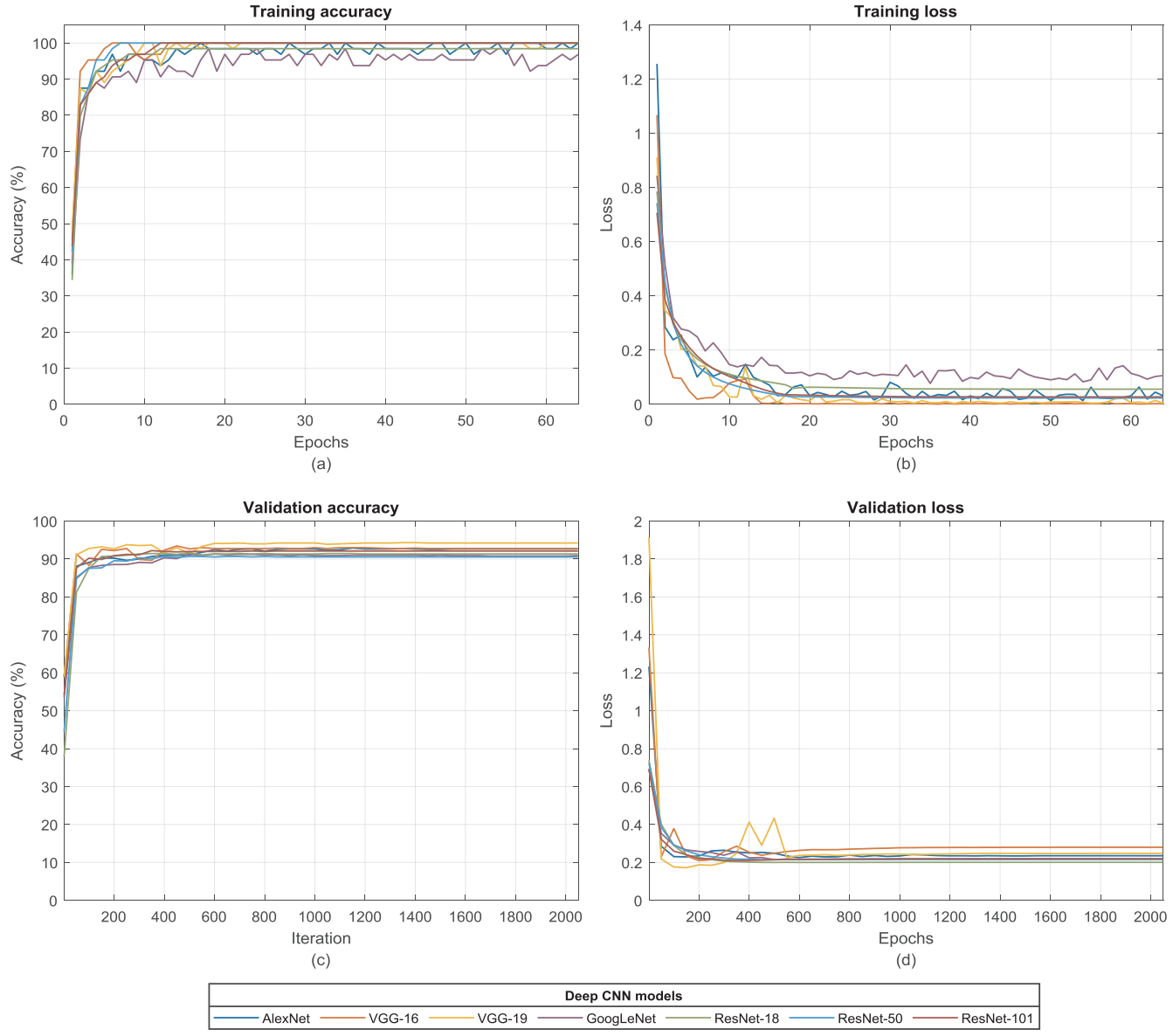

Fig. 9. (a) Training accuracy over the epochs of the CNN models. (b) Training loss over the epochs of the CNN models. (c) Validation accuracy over the iterations of the CNN models. (d) Validation loss over the iterations of the CNN models.

Table 5
The confusion matrices of the CNN models.

AlexNet		VGG-16		VGG-19		GoogLeNet		ResNet-18		ResNet-50		ResNet-101	
348[a]	45[b]	344[a]	40[b]	349[a]	32[b]	329[a]	42[b]	346[a]	45[b]	344[a]	57[b]	330[a]	41[b]
21[c]	486[d]	25[c]	491[d]	20[c]	499[d]	40[c]	489[d]	23[c]	486[d]	25[c]	474[d]	39[c]	490[d]

[a] TP – True Positive.
[b] FP – False Positive.
[c] False Negative.
[d] True Negative.

realized considering the structure of the new specific task (Salaken et al., 2019). In this study, we focused on haploid and diploid maize seeds (binary classification), so we had only two classes. For this reason, the new fully connected layer had two neurons. The details of the parameters used in the fine-tuning process are presented in Table 2.

The maximum number of epochs corresponds to a limit value for minimizing the loss function that depends on the minibatch size. The minibatch size is a subset of the training set that is used to evaluate the loss function gradient as well as to update the network weights. The network starts to update the weight with the initial learning rate. The learning rate schedule ensures decreasing the learning rate, and this depends on a certain number of drop periods and drop factors. In this manner, the training of the network is faster and more proper training time is ensured.

Table 6
The classification results of the CNN models.

	Acc (%)	Se (%)	Sp(%)	QI(%)	F-score (%)	AUC	Error Rate
AlexNet	92.67	94.31	91.53	92.91	91.33	0.9758	0.2352
VGG-16	92.78	93.23	92.47	92.85	91.37	0.9793	0.2800
VGG-19	**94.22**	**94.58**	**93.97**	**94.27**	**93.07**	**0.9852**	**0.2477**
GoogLeNet	90.89	89.16	92.09	90.61	88.92	0.9699	0.2170
ResNet-18	92.44	93.77	91.53	92.64	91.05	0.9795	0.2007
ResNet-50	90.89	93.23	89.27	91.22	89.35	0.9718	0.2160
ResNet-101	91.11	89.43	92.28	90.84	89.19	0.9750	0.2198

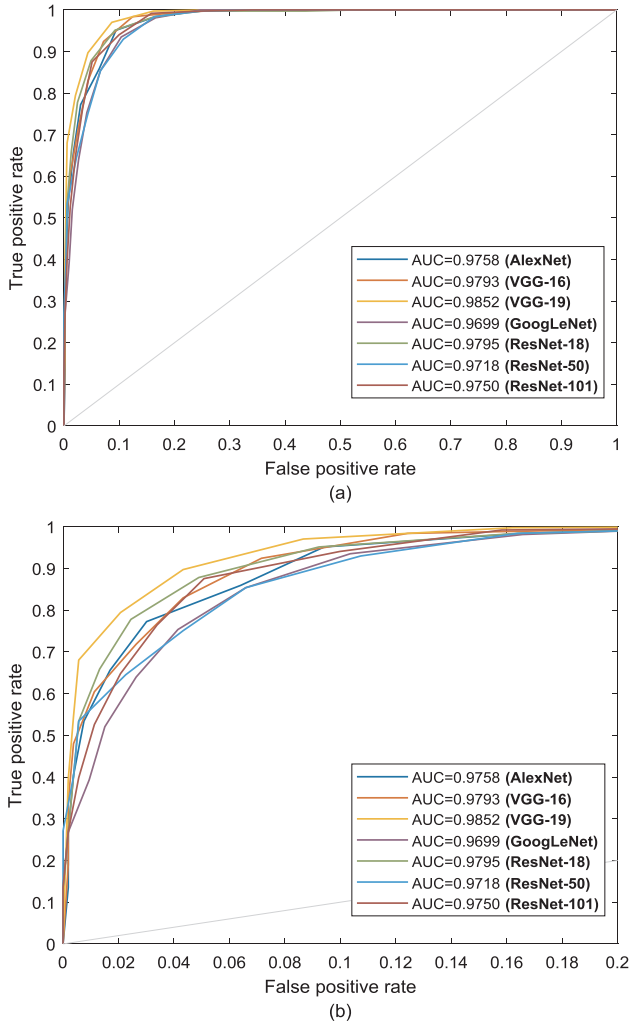

Fig. 10. (a) ROC curves of the CNN models. (b) The differences in the AUCs in the range of 0 and 0.2 false positive rates for all CNN models.

3. Results and discussion

3.1. Performance metrics

In this study, first a confusion matrix was utilized for predicting the performance of the CNN models. For a binary classification task, the confusion matrix included four indices that are true positive (TP), true negative (TN), false positive (FP) and false negative (FN) as shown in Fig. 8. In our case, the TP and TN matched the number of haploid and diploid maize seeds recognized correctly, whereas the FP and FN matched the number of haploid and diploid maize seeds recognized incorrectly. The number of real haploid and diploid maize seeds correspond to $+ R(TP + FN)$ and $- R(FP + TN)$, respectively, whereas the number of predicted haploid and diploid maize seeds correspond to $+ P(TP + FP)$ and $- P(FN + TN)$, respectively. Finally, N represents the total number of recordings.

Several performance metrics, such as accuracy (Acc), sensitivity (Se, recall), specificity (Sp), quality index (QI), and F-score, can be derived from the confusion matrix. The formulation of the performance metrics and their short descriptions are represented in Table 3. It should be noted that the importance of the QI and F-score increases in terms of a more true performance interpretation if the number of recordings among the classes is not equal (Cömert et al., 2018).

Receiver operating characteristic (ROC) curves are also a useful tool for measuring a model performance without considering class distribution or error costs. An ROC curve is generated by drawing all specific values versus correspondent sensitivity values (Cömert and Kocamaz, 2018). In this scope, the area under the curve (AUC) is a suitable metric for measuring the binary class tasks. It is desirable that the AUC value is as close to one as possible.

3.2. Experimental results

The experiments were implemented on MATLAB (R2018b) with a single NVIDIA Quadro P6000 GPU. The dataset was divided randomly into two parts as training and testing sets in the performance calculation of each CNN model. The rates of training and test sets were kept at 70% and 30%, respectively. The detailed settings are given in Table 4.

As mentioned before, we preferred to use the transfer learning approach with the given parameters in Table 2. The accuracy and loss over epochs are illustrated in Fig. 9 for the training and validation sets. We tested several experimental settings to find the most efficient parameters for all CNN models. As inferred from Fig. 9, all CNN models achieved high accuracy in the training phase. In addition, the models completed the process of convergence in approximately 20 epoch.

After completion of the training processes, the confusion matrices of the test sets of the CNN models were obtained as in Table 5. Additionally, the performance metrics derived from the confusion matrices are reported in Table 6. All CNN models could distinguish the haploid and diploid maize seeds with an accuracy of over 90%. In particular, the VGG-19 model ensured the best Se and Sp with 94.58% and 93.97%, respectively. In our case, the most important performance metric is Se because this metric represents the model success in recognizing haploid maize seeds. The best and worst sensitivity values were achieved by VGG-19 (Se = 94.58%) and GoogLeNet (Se = 89.16%). As mentioned before, a total of 369 haploid and 531 diploid seeds were used in the test set and VGG-19 could recognize 349 haploid and 499 diploids maize seeds correctly.

Due to the imbalanced data distribution among the classes, the ROC curves, QI, and F-score metrics become more significant. From this aspect, VGG-16 has the best values for the mentioned metrics. Furthermore, the AUCs were equal to 0.9758, 0.9793, 0.9852, 0.9699, 0.9795, 0.9718, and 0.9750 for AlexNet, VGG-16, VGG-19, GoogLeNet, ResNet-18, ResNet-50, and ResNet-101, respectively. The ROC curves of the CNN models are shown in Fig. 10.

Fig. 11 presents the comparison of the CNN model testing performances. As displayed in Fig. 11, the VGG-19 obtained the maximum performance of classification with respect to the others ignoring the consumed time for the training. AlexNet required the shortest training time (19.82 min), while ResNet-101 took the longest time for training (110.13 min). It should be noted that the training times of the networks depend on the hardware resources. Using advanced GPUs can ensure shorter training times for the CNNs.

3.3. Discussion

A one-to-one comparison among the related papers is not feasible due to the use of different methods, datasets, and division criterion. Nevertheless, we present a comparison in Table 7 considering several

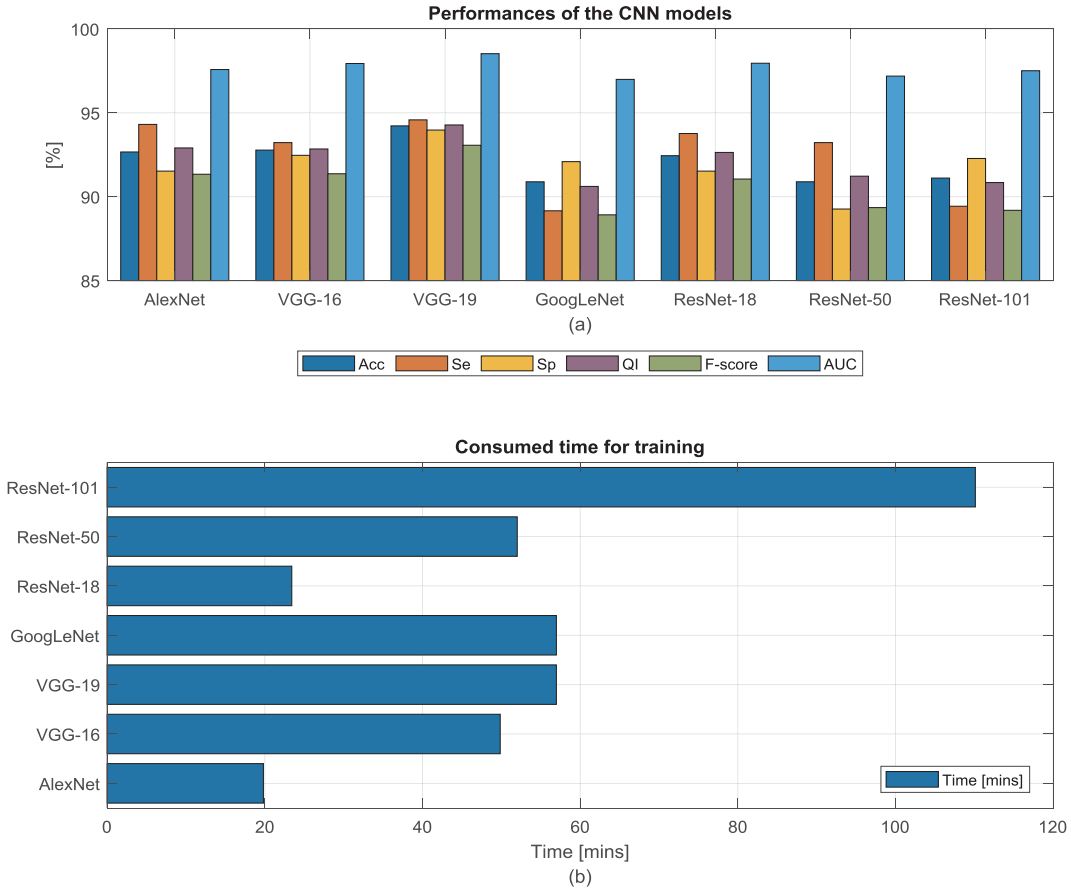

Fig. 11. (a) The performances of the respective CNN models. (b) The consumed time for training.

Table 7
Comparison of the proposed model and the related studies.

Author(s)	#G	Training set		Test set		Imaging method (sensor type)	Methods	Acc	Se	Sp
		#H	#D	#H	#D					
Boote et al. (2016)	7	–	–	59	N/A	Fluorescence imaging	Thresholding	N/A	93.2*	N/A
De La Fuente et al. (2017)	6	600	600	60	60	Multispectral imaging	3th party software	85.8*	76.6*	95.0*
Altuntaş et al. (2018a, 2018b)	150	87	326	87	326	Digital camera	SVM	81.3	94.2	77.9
Altuntaş et al. (2018a, 2018b)	150	87	326	87	326	Digital camera	DT, kNN, ANN	84.5	63.2	90.3
L. Yu et al. (2018)	1	100	100	100	100	Near-infrared spectroscopy	SVSKLPP	97.1	98.8	95.4
Song et al. (2018)	4	314	736	314	736	Digital camera	Thresholding	92.0*	91.4*	92.2*
Wang et al. (2018)	2	100	100	100	100	Hyperspectral imaging	BPR	N/A	N/A	N/A
This study	107	861	1239	369	531	Digital camera	CNN	93.4	95.1	92.2

#G: the number of genotypes, #H: the number of haploid seeds, #D: the number of diploid seeds. N/A: Information not available. BPR: Biomimetic pattern recognition.

* The metrics derived from the information given in the papers.

criteria, such as the number of genotypes, the number of samples, distribution of the samples among the classes, imaging methods, the methods used, and finally, the performance metrics.

While only a small number of source genotypes were used in similar studies, 107 source genotypes, which are yellow dent and waxy grain types, were used in this study. This high number of source genotypes is intended to increase the overall validity of the proposed method.

As the imaging equipment, associated studies have often used expensive devices, such as near-infrared spectroscopy (NIR), multispectral imaging (MDI), hyperspectral imaging (HIS) and fluorescence imaging (FI) hardware. In the proposed method, a low-cost digital camera was used. Thus, in terms of the aspect of the cost, a highly effective system has been developed.

The number of data are very variable in related studies as seen in Table 7. A low amount of data makes the reliability of the studies unreliable. As seen in Table 7, the total amount of data used in this study was almost 3 times higher than the highest amount of data used in the literature. In this case, the reliability level of the proposed study has been increased.

As we mentioned before, the seeds were labeled as haploid or diploid by two field experts according to their *R1-nj* expression. Although a rigorous study was carried out, there may be mislabeled seeds in the dataset. Although all CNN models yielded excellent results, possible mislabeled samples in the dataset may have negatively affected the learning of the CNN networks.

In previous studies, methods, such as thresholding and conventional machine learning, were used for the classification of maize seeds. In this study, a CNN method that does not need feature extraction with a high

classification ability has been used. In this method, in cases of sufficient data, classification performance can be increased to a high level.

4. Conclusion

In this paper, we adopted CNN models to recognize haploid and diploid maize seeds automatically through a transfer learning approach. Seven well-known CNNs architectures were trained to classify visual objects on natural images and were fitted to this specific task. Our transfer learning strategy relied on keeping and freezing the convolutional layers and updating the fully connected layers to recognize haploid and diploid maize seeds. Although training a deep CNN model needs a large-scale resource, we achieved excellent results on the identification of the haploid and diploid maize seeds using a transfer learning approach. Our approach was able to recognize approximately nine out of every ten maize seeds correctly using CNN models. This outcome emphasizes the major superiority of transfer learning and its ability to use the potential of deep learning on domains with a reduced number of training samples. Furthermore, we exploited the end-to-end learning advantages of the CNN models by eliminating the tedious manual feature extraction task. The experimental results prove that the CNN models ensure a new opportunity for distinguishing haploid maize seeds from diploids nondestructively, rapidly and at a low cost. Although all CNN models ensured promising results, the most effective model was determined to be VGG-19 for this particular task. Acc, Se, Sp, QI and F-score performance metrics of VGG-19 were achieved as 94.22%, 94.58%, 93.97%, 94.27%, and 93.07%, respectively. In addition, VGG-19 achieved a 0.9852 AUC value. Furthermore, we have introduced a new publicly available haploid and diploid maize seeds dataset consisting of 1230 haploid and 1770 diploid maize seed images. Finally, we conclude that the CNN model is significantly superior to machine learning-based methods and traditional manual selection. The results demonstrate that the use of CNN models in haploid and diploid maize recognition is a nondestructive, rapid and low-cost solution.

In the future, we will investigate the use of activation maps as a feature set for feeding shallow networks. In addition, a machine using deep CNN models will be designed in order to transfer the experimental knowledge to the field.

Funding

There is no funding source for this article.

Ethical approval

This article does not contain any data, or other information from studies or experimentation, with the involvement of human or animal subjects.

Data availability statement

The haploid and diploid maize seeds dataset can be downloaded freely on https://www.rovile.org/datasets/haploid-and-diploid-maize-seeds-dataset/.

Declaration of Competing Interest

The authors declare that there is no conflict to interest related to this paper.

Acknowledgment

We would like to thank Dr. Rahime Cengiz and Dr. Mesut Esmeray from Maize Research Institute, Sakarya, Turkey for providing and labeling the haploid and diploid maize seeds.

Appendix A. Supplementary material

Supplementary data to this article can be found online at https://doi.org/10.1016/j.compag.2019.104874.

References

Altuntaş, Y., Kocamaz, A.F., Cengiz, R., Esmeray, M., 2018a. Classification of haploid and diploid maize seeds by using image processing techniques and support vector machines. In: 2018 26th Signal Processing and Communications Applications Conference (SIU), pp. 1–4. https://doi.org/10.1109/SIU.2018.8404800.

Altuntaş, Y., Kocamaz, A.F., Cömert, Z., Cengiz, R., Esmeray, M., 2018b. Identification of haploid maize seeds using gray level co-occurrence matrix and machine learning techniques. In: 2018 International Conference on Artificial Intelligence and Data Processing (IDAP). Malatya, Turkey, pp. 1–5, https://doi.org/10.1109/IDAP.2018.8620740.

Boote, B.W., Freppon, D.J., De La Fuente, G.N., Lübberstedt, T., Nikolau, B.J., Smith, E.A., 2016. Haploid differentiation in maize kernels based on fluorescence imaging. Plant Breed. 135, 439–445. https://doi.org/10.1111/pbr.12382.

Cerit, İ., Cömertpay, G., Oyucu, R., Çakır, B., Hatipoğlu, R., Özkan, H., 2016. Melez Mısır Islahında In-Vivo Katlanmış Haploid Tekniğinde Kullanılan Farklı Inducer Genotiplerin Haploid İndirgeme Oranların Belirlenmesi. Tarla Bitk. Merk. Araştırma Enstitüsü Derg. 25, 52–57. https://doi.org/10.21566/tarbitderg.280162.

Chaikam, V., Martinez, L., Melchinger, A., Schipprack, W., Prasanna, B., 2016. Development and validation of red root marker-based haploid inducers in maize. Crop Sci. 56. https://doi.org/10.2135/cropsci2015.10.0653.

Charity, C., Faith, M., Joanah, M., Tendai, H., 2017. Production and use of haploids and doubled haploid in maize breeding: a review. African J. Plant Breed. 4, 201–213.

Chase, S.S., Nanda, D.K., 1969. Rapid inbreeding in maize. Econ. Bot. 23, 165–173. https://doi.org/10.1007/BF02860622.

Cömert, Z., Kocamaz, A.F., 2019. Fetal hypoxia detection based on deep convolutional neural network with transfer learning approach. In: Silhavy, R. (Ed.), Software Engineering and Algorithms in Intelligent Systems. Springer International Publishing, Cham, pp. 239–248. https://doi.org/10.1007/978-3-319-91186-1_25.

Cömert, Z., Kocamaz, A.F., 2018. Open-access software for analysis of fetal heart rate signals. Biomed. Signal Process. Control 45, 98–108. https://doi.org/10.1016/j.bspc.2018.05.016.

Cömert, Z., Kocamaz, A.F., 2017. A study of artificial neural network training algorithms for classification of cardiotocography signals. Bitlis. Eren. Univ. J. Sci. Technol. 7, 93–103. https://doi.org/10.17678/beuscitech.338085.

Cömert, Z., Kocamaz, A.F., Subha, V., 2018. Prognostic model based on image-based time-frequency features and genetic algorithm for fetal hypoxia assessment. Biol. Med. Comput. https://doi.org/10.1016/j.compbiomed.2018.06.003.

De La Fuente, G.N., Carstensen, J.M., Edberg, M.A., Lü bberstedt, T., 2017. Discrimination of haploid and diploid maize kernels via multispectral imaging. Plant Breed. 136, 50–60. https://doi.org/10.1111/pbr.12445.

Garcia-Garcia, A., Orts-Escolano, S., Oprea, S., Villena-Martinez, V., Martinez-Gonzalez, P., Garcia-Rodriguez, J., 2018. A survey on deep learning techniques for image and video semantic segmentation. Appl. Soft Comput. 70, 41–65. https://doi.org/10.1016/J.ASOC.2018.05.018.

Geiger, H.H., 2009. Doubled haploids. In: Bennetzen, J.L., Hake, S. (Eds.), Handbook of Maize. Springer New York, New York, NY, pp. 641–657. https://doi.org/10.1007/978-0-387-77863-1_32.

Geiger, H.H., Andrés Gordillo, G., Koch, S., 2013. Genetic correlations among haploids, doubled haploids, and testcrosses in maize. Crop Sci. 53, 2313–2320. https://doi.org/10.2135/cropsci2013.03.0163.

He, K., Zhang, X., Ren, S., Sun, J., 2016. Deep residual learning for image recognition. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Las Vegas, NV, USA, pp. 770–778. https://doi.org/10.1109/CVPR.2016.90.

Krizhevsky, A., Sutskever, I., Hinton, G.E., 2012. ImageNet classification with deep convolutional neural networks. In: Pereira, F., Burges, C.J.C., Bottou, L., Weinberger, K.Q. (Eds.), Proceedings of the 25th International Conference on Neural Information Processing Systems – Volume 1, NIPS'12. Curran Associates, Inc., USA, pp. 1097–1105.

LeCun, Y., Bengio, Y., Hinton, G., 2015. Deep learning. Nature 521, 436–444. https://doi.org/10.1038/nature14539.

Li, H., Li, W., Qin, H., Chen, S., Liu, J., Li, W., 2016. Classifying method of haploid and diploid based on least square error. Nongye Jixie Xuebao/Trans. Chinese Soc. Agric. Mach. 47, 259–264. https://doi.org/10.6041/j.issn.1000-1298.2016.06.034.

Lin, M., Chen, Q., Yan, S., 2013. Network in network. arXiv Prepr. arXiv1312.4400.

Liu, J., Guo, T.T., Li, H.C., Jia, S.Q., Yan, Y.L., An, D., Zhang, Y., Chen, S.J., 2015. Discrimination of maize haploid seeds from hybrid seeds using vis spectroscopy and support vector machine method. Guang pu xue yu guang pu fen xi 35, 3268–3274.

Mazo, C., Bernal, J., Trujillo, M., Alegre, E., 2018. Transfer learning for classification of cardiovascular tissues in histological images. Comput. Methods Programs Biomed. 165, 69–76. https://doi.org/10.1016/j.cmpb.2018.08.006.

Melchinger, A.E., Schipprack, W., Friedrich Utz, H., Mirdita, V., 2014. In vivo haploid induction in maize: identification of haploid seeds by their oil content. Crop Sci. 54, 1497. https://doi.org/10.2135/cropsci2013.12.0851.

Melchinger, A.E., Schipprack, W., Würschum, T., Chen, S., Technow, F., 2013. Rapid and accurate identification of in vivo-induced haploid seeds based on oil content in maize. Sci. Rep. 3, 2129. https://doi.org/10.1038/srep02129.

Nanda, D.K., Chase, S.S., 1966. An embryo marker for detecting monoploids of maize (Zea Mays L.). Crop Sci. 6, 213. https://doi.org/10.2135/cropsci1966.

0011183X000600020036x.
Penatti, O.A.B., Nogueira, K., dos Santos, J.A., 2015. Do deep features generalize from everyday objects to remote sensing and aerial scenes domains?. In: 2015 IEEE Conference on Computer Vision and Pattern Recognition Workshops (CVPRW), pp. 44–51. https://doi.org/10.1109/CVPRW.2015.7301382.
Prasanna, B.M., Chaikam, V., Mahuku, G., 2012. Doubled haploid technology in maize breeding: theory and practice. CIMMYT.
Razavian, A.S., Azizpour, H., Sullivan, J., Carlsson, S., 2014. {CNN} Features off-the-shelf: an Astounding Baseline for Recognition. CoRR abs/1403.6.
Röber, F.K., Gordillo, G.A., Geiger, H.H., 2005. In vivo haploid induction in maize – performance of new inducers and significance of doubled haploid lines in hybrid breeding. Maydica 50, 275–283.
Salaken, S.M., Khosravi, A., Nguyen, T., Nahavandi, S., 2019. Seeded transfer learning for regression problems with deep learning. Expert Syst. Appl. 115, 565–577. https://doi.org/10.1016/j.eswa.2018.08.041.
Simonyan, K., Zisserman, A., 2014. Very deep convolutional networks for large-scale image recognition. arXiv Prepr. arXiv1409.1556.
Smelser, A., Blanco, M., Lübberstedt, T., Schechert, A., Vanous, A., Gardner, C., 2014. Weighing in on a method to discriminate maize haploid from hybrid seed. Plant Breed. 134, 283–285. https://doi.org/10.1111/pbr.12260.
Song, P., Wu, K., Zhang, J., Li, W., Fang, X., 2010. Sorting system of maize haploid kernels based on computer vision. Nongye Jixie Xuebao/Trans. Chinese Soc. Agric. Mach. 41, 249–252. https://doi.org/10.3969/j.issn.1000-1298.2010.Supp. 054.
Song, P., Zhang, H., Wang, C., Luo, B., Zhang, J.X., 2018. Design and experiment of a sorting system for haploid maize kernel. Int. J. Pattern Recognit. Artif. Intell. 32. https://doi.org/10.1142/S0218001418550029.
Szegedy, C., Liu, Wei, Jia, Yangqing, Sermanet, P., Reed, S., Anguelov, D., Erhan, D., Vanhoucke, V., Rabinovich, A., 2015. Going deeper with convolutions. In: 2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). IEEE, pp. 1–9. https://doi.org/10.1109/CVPR.2015.7298594.
Tajbakhsh, N., Shin, J.Y., Gurudu, S.R., Hurst, R.T., Kendall, C.B., Gotway, M.B., Liang, J., 2016. Convolutional Neural Networks for Medical Image Analysis: Full Training or Fine Tuning? IEEE Trans. Med. Imaging. https://doi.org/10.1109/TMI.2016.2535302.
Traore, B.B., Kamsu-Foguem, B., Tangara, F., 2018. Deep convolution neural network for image recognition. Ecol. Inform. 48, 257–268. https://doi.org/10.1016/j.ecoinf.2018.10.002.
Wang, Y., Lv, Y., Liu, H., Wei, Y., Zhang, J., An, D., Wu, J., 2018. Identification of maize haploid kernels based on hyperspectral imaging technology. Comput. Electron. Agric. 153, 188–195. https://doi.org/10.1016/j.compag.2018.08.012.
Yu, H., Yang, Z., Tan, L., Wang, Y., Sun, W., Sun, M., Tang, Y., 2018a. Methods and datasets on semantic segmentation: a review. Neurocomputing 304, 82–103. https://doi.org/10.1016/j.neucom.2018.03.037.
Yu, L., Liu, W., Li, W., Qin, H., Xu, J., Zuo, M., 2018b. Non-destructive identification of maize haploid seeds using nonlinear analysis method based on their near-infrared spectra. Biosyst. Eng. 172, 144–153. https://doi.org/10.1016/j.biosystemseng.2018.05.011.
Yu, Y., Lin, H., Meng, J., Wei, X., Guo, H., Zhao, Z., 2017. Deep transfer learning for modality classification of medical images. Information 8, 91. https://doi.org/10.3390/info8030091.