Preprints.org

Article

Digital Identification of Different Varieties of Kazakhstan Apples Using Deep Learning Techniques

Jakhfer Alikhanov , Tsvetelina Georgieva , Eleonora Stefanova Nedelcheva , Aidar Moldazhanov ,
Akmaral Kulmakhambetova , Dmitriy Zinchenko [*] , Alisher Nurtuleuov , Zhandos Shynybay ,
Plamen Ivanov Daskalov [*]

Article

Digital Identification of Different Varieties of Kazakhstan Apples Using Deep Learning Techniques

Jakhfer Alikhanov [1], **Tsvetelina Georgieva** [2], **Eleonora Nedelcheva** [2], **Aidar Moldazhanov** [1], **Akmaral Kulmakhambetova** [1], **Dmitriy Zinchenko** [1], **Alisher Nurtuleuov** [1], **Zhandos Shynybay** [3] **and Plamen Daskalov** [2,*]

[1] Department of Energy Saving and Automation, Kazakh National Agrarian Research University, 8 Abay Str., Almaty, Kazakhstan

[2] Department of Automatics and Electronics, University of Ruse, 8 Studentska Str., 7017 Ruse, Bulgaria

[3] Department of Power Supply and Electric Drive, Almaty University of Power Engineering and Telecommunication Named Gumarbek Daukeev, Baitursynov Str. Office 126/1, Almaty, Kazakhstan

* Correspondence: daskalov@uni-ruse.bg

Abstract

Application of pre-trained CNN and transfer learning approach for identification of different varieties of Kazakhstan apples is presented in the paper. The study used the most popular Kazakhstan apple varieties as: Aport Alexander, Ainur, Sinap Almaty, Nursat and Kazakhskij Yubilejnyj. In this study, two pre-trained CNN networks, SqueezeNet and GoogLeNet, were fine-tuned using different values of the Initial Learning Rate (ILR) parameter of 0.0001, 0.0002, 0.00025, 0.0003, 0.00035, 0.0004, 0.0005 and three network tuning algorithms, respectively The Stochastic Gradient Descent with moment solver (Sgdm), the Adam optimization algorithm (Adam), and Root Mean Square Propagation (RMSprop). The performance of the networks was evaluated by analyzing the values of Training Accuracy (TA), Training Loss (TL), Validation Accuracy (VA), Validation Loss (VL), and Confusion Matrix. The training of the networks was performed in 1900 iterations and 30 epochs. The developed approach for the identification of five varieties of Kazakhstan apples using deep learning techniques achieved 100% correct classification of fruits for the Ainur variety with a GoogLeNet network, solver RMSprop and ILR= 0.0005. For the varieties Aport, Kazakhski Yubileinyi and Nursat, one of the three network evaluation indicators achieves 100% accuracy, and for Sinap Almatynski – all three indicators are with and above the value of 95%. For varieties Ainur, Aport, Kazakhski Yubileinyi and Sinap Almatynski as the best (with the highest optimal accuracy values) is the GoogLeNet network, with the following settings: solver RMSprop and ILR=0.0005. Only for variety Nursat is the SqeezeNet network suitable, with the following settings: solver Sgdm and ILR= 0.0003. The proposed approach for recognizing the varietal affiliation of apples using deep learning neural networks is suitable for the analyzed apple varieties and could be easily implemented and used in industrial conditions for sorting fruits. The achieved recognition accuracy meets the requirements in the field.

Keywords: apple; fruit; variety; identification; pre-trained CNN; deep learning; transfer learning

1. Introduction

Apples are considered one of the most common and popular fruits in the world. For about a hundred years, the Kazakh Research Institute of Fruits and Vegetables has been conducting selection work to create new varieties of fruit crops, with the aim of creating competitive varieties that combine in their genome a set of economically valuable qualities of varieties of modern world selection with excellent taste and appearance [1,2]. The State Register of Selection Achievements approved for use in the Republic of Kazakhstan includes 66 apple varieties. For example, the Aport apple variety occupies a special place in the fruit growing of Kazakhstan. Aport apples are in great demand among

the local population. The price of Aport fruits in the Almaty markets is twice as high as that of the most common varieties Golden Delicious, Starkrimson, Fuji, etc.

Each variety has its own characteristics that allow it to be identified by its appearance. The difference can be seen in the shape, color, size, texture and other properties of the fruit. The shape of the fruit is one of the first signs to pay attention to. It can be round, flattened or elongated. The color of the apple can vary from bright red to light green or yellow, and also has different shades and transitions. Some varieties also have specific colors, such as stripes or spots. The size of the apple varies from variety to variety and within one variety. It can be small, medium, large or even giant. The size is usually determined by the diameter of the fruit, with a certain range of values for each variety. The texture of the apple can be smooth or rough, soft or hard, and can also vary from firmer to juicier apples. There are other characteristics that can influence the determination of the apple variety, including the presence of bumps or pits on the surface of the fruit, wrinkles on the skin, and the presence of characteristics in the overall structure of the fruit.

Modern research in plant biology is actively addressing the problem of identifying different varieties of fruits and vegetables without the need to seek the help of experienced gardeners or experts. Science has developed a number of innovative approaches that allow obtaining a more accurate description of the differences between varieties with a diverse range of fruits [4–6]. One group of methods is based on the analysis of the physical characteristics of the fruit, such as shape, size, color, staining and the presence of spots on the peel [4,7]. Another group of methods focuses on the analysis of spectral characteristics [8], another equally accurate method for identifying varieties involves the use of molecular studies that can determine the genetic profile of fruit varieties [9]. In addition, modern research is also focused on the analysis of the chemical composition of apples by determining the content of various organic compounds [10], which can serve as indicators of a particular apple variety.

Algorithms and numerical methods for determining the parameters of apple fruits based on computer image processing have been developed, which increase the productivity and accuracy of quantitative assessment of weight, color and shape for automatic sorting of apples into commercial classes in accordance with the requirements of standards [4,11]. In most studies on determining the quality of apples using computer vision, fruits are classified by quality into categories according to their size [11,12], color [13,14] and shape, as well as for the presence of defects [15–18], but research on Kazakh apple varieties is very limited.

The remarkable growth in the use of artificial intelligence (AI) in many occasions and areas of life, including smart agriculture, has led to many researches focused on image identification and classification using deep learning methods [13,19,20]. Many researches addressed the approach of transfer learning [21–23], they use GoogLeNet, AlexNet, YOLO-V3, SqueezeNet, VGG and more.

Focusing on the use of transfer learning Ibarra-Pérez et al. [21] is analyzing different CNN architectures to identify the phenological stages of plants – bean. The perfor-mances of AlexNet, VGG19, SqueezeNet, and GoogleNet networks were compared by authors and GoogleNet was evaluated as best in their occasion reaching 96.71% accuracy.

Rady et al. [23] analyzed the ability of deep neural networks learned by transfer learning to classify the grade of staped cotton cultivars (egyptian cotton fibres). The au-thors used five Convolutional Neural networks (CNNs)—AlexNet, GoogleNet, SqueezeNet, VGG16, and VGG19 and concluded that AlexNet, GoogleNet, and VGG19 outperformed the others reaching F1-Scores ranging from 40.0–100% depending on the cultivar type.

Another research made by Yunong at al. [24] proposed the use of AlexNet, VGG, GoogleNet, and YOLO-V3 models for anthracnose lesion detection on apple fruits. Firstly, CycleGAN deep learning method is adopted to extract the features of healthy apples and anthracnose apples and to produce anthracnose lesions on the surface of healthy apple images. Compared with the traditional image augmentation methods, this method greatly enriches the diversity of training dataset and provides plentiful data for model training. Based on the data augmentation. DenseNet is adopted in their research to substitute the lower resolution layers of the YOLO-V3 model.

Yanfei et al. [25] researched apple quality identification and classification for grading apples into three quality categories from real images containing complicated disturbance information - background similar to the surface of the fruits. The authors developed and trained CNN-based identification architecture for apple sample images. They compared the overall performance of proposed CNN-based architecture, Google Inception v3 model, and HOG/GLCM + traditional SVM method, obtaining the accuracy of 95.33%, 91.33%, and 77.67%, respectively.

As mentioned earlier, each apple variety has its own unique taste and characteristics, but often the fruits have similar texture, color, and appearance to the naked human eye. Determining the exact apple variety is important for agronomists, gardeners, and farmers to properly care for the apple tree and take into account the growth and yield characteristics of each variety. Modern advances in computer image processing and artificial intelligence make it possible to solve this problem. In [30], a digital methodology for determining the main characteristics of apples through the analysis of digital images is presented, but a digital methodology for recognizing the varietal affiliation of apples is missing.

The aim of the study is to develop a method for apple variety recognition based on color imaging, machine vision techniques and deep neural networks using transfer learning, which will complement the digital methodology developed in [30] and determine whether the fruit belongs to a particular variety for further use in selection and identification of the correspondence of fruits to a particular variety when sold to consumers. The results of the study will be used for future research in the task of developing a machine for automatic sorting of apples into commercial varieties.

2. Materials and Methods

2.1. Apple Samples Collection and Digital Image Acquisition

The objects of the study are five varieties of apples of Kazakhstan selection: Aport Alexander, Aynur, Sinap Almaty, Nursat and Kazakhskij Yubilejnyj. The samples were collected in the pomological garden of the Talgar branch of the Kazakh Research Institute of Fruit and Vegetable Growing (GPS coordinates: 43.238949, 76.889709).

The sampling of apples was carried out using the stratified sampling method to ensure representativeness and coverage of the entire population variability. The selected apple varieties included typical fruit specimens, taking into account the color, size, weight and shape of each sample. Then, based on the main criteria of visual integrity, ripeness and lack of defects, sample fruits were selected. The total number of fruits examined was 250 specimens, 50 fruits from each variety, which corresponds to the required sample size to ensure statistical reliability at a significance level of α=0.25.

Digital images of the fruits were obtained under controlled lighting and background conditions, as shown in Figure 1.
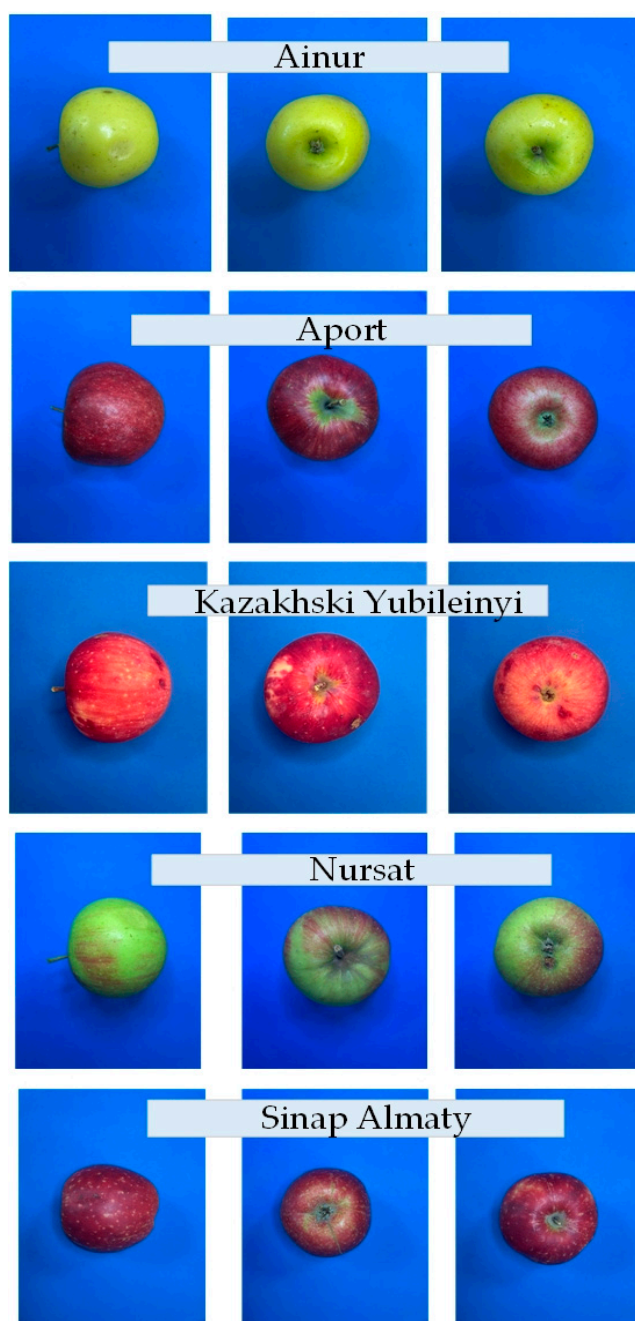
Figure 1. Image acquisition workstation for apple fruits.

To obtain high-quality images of the studied objects (4), a stationary vertical (top-down) photography setup was used in this work, shown in Figure 1. The setup is based on a Canon EOS 4000D digital camera (1), a tripod with a horizontal bar Benro SystemGo Plus (2) and a solid blue background (3), placed on a flat surface. This type of configuration is widely used in the construction of computer vision systems, digital sorting of agricultural products and in the preparation of training samples for subsequent analysis. The Canon EOS 4000D digital SLR camera with an EF-S 18–55mm f/3.5–5.6 III lens provides shooting with a resolution of 18 megapixels, using an APS-C CMOS matrix (22.3 × 14.9 mm). To minimize distortion and ensure high detail, shooting was performed at a focal length of 55 mm. The camera was fixed strictly vertically on a horizontal rod of the tripod using a ball head and a quick-change plate. The shooting mode was set to Manual, with manual focus (MF) on the central area of the object. The exposure parameters were selected experimentally: ISO sensitivity 100–200 units to reduce digital noise; shutter speed – from 1/60 to 1/125 second; aperture – f/8 to ensure uniform sharpness throughout the depth of the object. The white balance was set manually using gray cardboard or the preset "daylight" setting (Daylight, 5500K). The photos were taken in RAW format (for subsequent processing and analysis), and also duplicated in JPEG for quick viewing. The Benro SystemGo Plus tripod with a horizontal retractable rod allows you to fix the camera strictly above the object, ensuring stability and repeatability of the conditions. The adjustable height and rotation mechanism allow precise adjustment of the distance between the lens and the shooting surface, which in this case was about 50 cm. The tripod is equipped with a built-in level, which is used to correct the horizontal position of the camera and prevent distortion of the frame. A plain blue A4 background made of matte paper that does not create glare was placed on the working surface. The choice of blue color is due to its high contrast with the color of the fruits and the lack of intersections in the color spectrum with objects, which contributes to more accurate segmentation and subsequent processing of the image. The center of the frame was occupied by the studied object - in this case an apple fruit, located strictly in the center of the shooting area. Each fruit was positioned in the same orientation with symmetry control, which ensures standardization of the photographic material. The scene was illuminated using diffuse daylight or a pair of LED sources with a color temperature of 5500K and a color rendering index (CRI) of over 90. The light panels were positioned symmetrically on both sides at an angle of 45° to the surface, ensuring uniform illumination and minimizing shadows. This approach improves the visual highlighting of object contours and increases the accuracy of the parameters extracted during digital processing.

Each object was photographed serially. The camera was started using a two-second timer, which eliminates image blurring from pressing the button. After each frame, the photo was saved in the camera's memory and subsequently transferred to the computer using a card reader. All images were
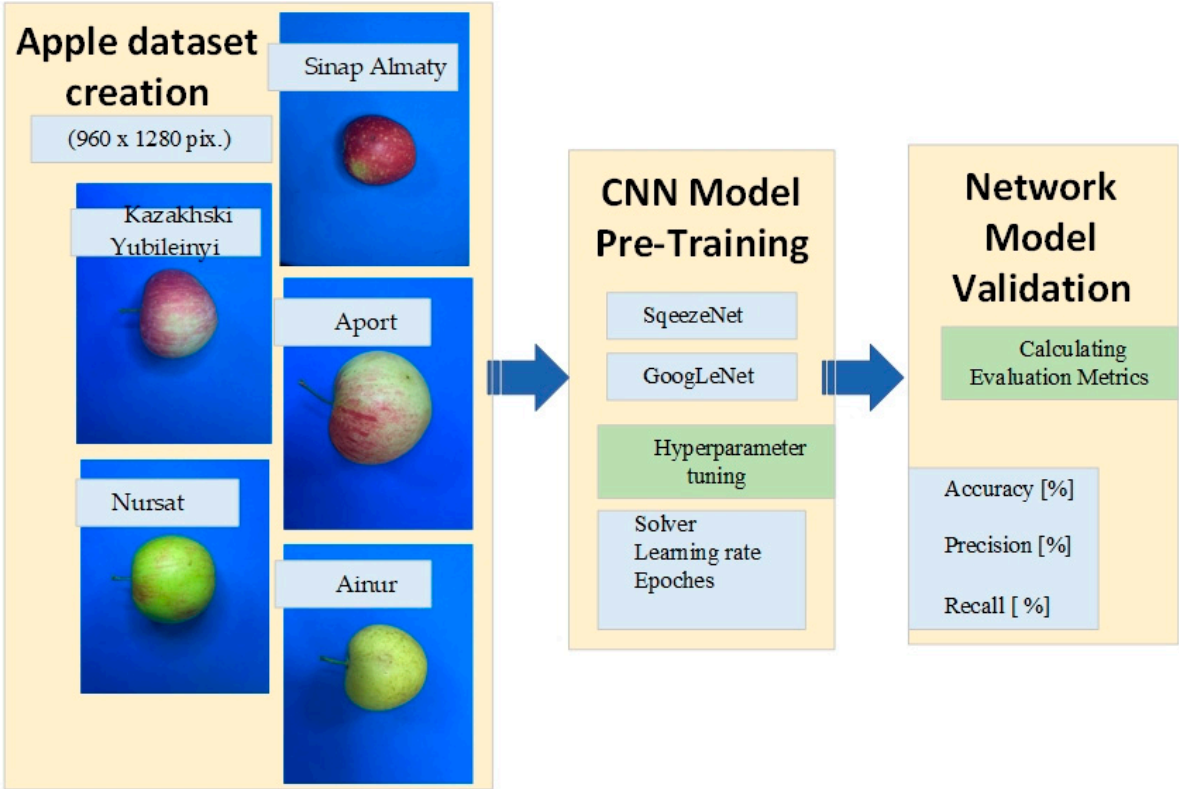
marked with the date and sample number and saved in a separate directory for easier inspection and analysis. Each apple was photographed in 3 different positions, as shown in Figure 2. The obtained images are in the RGB color space and have a resolution of 960 × 1280 pixels. The obtained images served as the basis for subsequent extraction of digital characteristics of the fruits. Image processing was performed using MATLAB software. Using the described setup allows standardization of the shooting process and ensures high data reproducibility, which is especially important in the context of scientific research, development of algorithms for automatic sorting and preparation of training samples for machine vision models.

Figure 2. Apple sample images for fruits from different variety, photographed in 3 different positions.

2.2. Algorithm for Digital Identification of Different Types of Apples by Deep Learning Techniques

The data processing algorithm has been developed and is shown in Figure 3.

Figure 3. Algorithm for digital identification of different types of apples by deep learning techniques.

The basic procedure for creating a model for digital identification of apple varieties is described in three stages (Figure 3). First, a sensor based on standard CCD technology is used to capture color images of apples with a resolution of 960x1280 pixels. The next stage involves building a test sample of images for each variety and training two types of deep learning networks, with different settings of the optimization algorithm - Solver and Initial Learning Rate, respectively. The third stage involves evaluating the performance of the networks and selecting the most suitable model. This is done on the basis of three basic metrics for network evaluation - Accuracy, Precision and Recall, which are calculated when identifying the variety of images from validation samples.

2.3. Deep Learning Model Parameters and Training Setting for Identification

In this study, two pre-trained CNN networks, SqueezeNet and GoogLeNet, were fine-tuned and trained using a transfer learning approach in Matlab to identify different types of apple varieties. The selection of an appropriate deep learning method was based on a review of studies conducted by other authors working on similar tasks [21].

The SqueezeNet network is based on AlexNet and has fewer parameters than GoogleNet and similar performance accuracy, achieved by introducing the fire module that uses 1×1 filters instead of 3×3 and reducing the number of input channels to 3×3 filters by using the fire module that contains 1×1 filters feeding into an augmented layer with a mixture of 1×1 and 3×3 filters [23,28]. The input image size according to the requirements for SqeezeNet is 227×227.

GoogleNet uses the Inception blocks technology, which integrates different convolutional algorithms and filter sizes into a single layer. This makes the model simpler, as the number of required computational parameters and processes is reduced, and shorter computational time is achieved. Compared to other CNN architectures that are available, such as AlexNet or VGG, this model has a significantly smaller total number of parameters [22,27]. The input image size for GoogLeNet is 224×224, and its architecture consists of 27 deep layers. This architecture is suitable for the task studied in the article.

For the needs of this study, certain elements of the SqueezeNet and GoogLeNet networks were modified to be able to perform recognition of 5 classes of objects corresponding to the five varieties of apples.

In deep learning, the optimizer, also known as a solver, is an algorithm used to update the parameters (weights and biases) of the model. For training of network models three different optimization algorithms were tested. The Stochastic Gradient Descent with moment solver (Sgdm), the Adam optimization algorithm (Adam solver), and Root Mean Square Propagation (RMSprop) were used consecutively, and a comparison of network performance is achieved.

Gradient Descent can be considered the most popular among the class of optimizers in deep learning. The SGD with heavy-ball momentum (SGDM) method have a wide range of applications due to its simplicity and great generalization. This solver has been widely applied in many machine learning tasks, and it is often applied with dynamic stepsizes and momentum weights tuned in a stagewise manner. This optimization al-gorithm uses calculus to consistently modify the values and achieve the local mini-mum.[29]

The Adam optimizer expands the classical stochastic gradient descent procedure considering the second moment of the gradients. The procedure calculates the uncen-tered variance of the gradients without subtracting the mean.

Root Mean Square Propagation, is an adaptive learning rate optimization algorithm used in training deep learning models. It's designed to address the limitations of basic gradient descent and other adaptive learning rate methods by adjusting the learning rate for each parameter based on the magnitude of recent gradients. This helps to stabilize training and improve convergence speed, particularly in cases with non-stationary ob-jectives or varying gradient magnitudes.

Other hyperparameters for training the convolutional neural networks (CNNs) in this study include an Initial Learning Rate of 0.0001, 0.0002, 0.00025, 0.0003, 0.00035, 0.0004, 0.0005 and a Learning Rate Drop Factor of 0.1.

The models were trained for 30 epochs, with a validation rate of 50. All images from all apple varieties were divided into training and validation sets, 70% and 30%, respectively, with the input image sets shuffled in each epoch with training and validation. Additionally, the input image sets were randomly augmented using the functions of the MATLAB Deep Network Designer application - Randomly rotate images in the range from -90 to +90 degrees and Randomly scale images in the range from 1 to 2, . Augmenting the input sets allows the trained networks to be invariant to distortions in the image data.

The output network was set to the best validation loss. Normalization of the input data was also included.

The retraining of the different CNNs was performed on a TREND Sonic computer running Windows 11 Pro 64-bit operating system with CPU: 13th generation Intel® Core™ i7-13700F/ 2.10 GHz; GPU: NVIDIA Ge-Force RTX 3070, 8GB; Memory: 64 GB. The implementation of transfer learning and subsequent classification tasks included in this study were performed using MATLAB R2023b a (MathWorks, Natick, MA, USA). The transfer training of networks was performed using Deep Learning Toolbox. For further analysis, the Experiment Manager in Matlab was also used.

2.4. Evaluation Metrics

Currently, a wide range of metrics are used in classification tasks to evaluate the performance of CNN models, which allow for numerical evaluation of the performance of the models. A number of true positives (TP), true negatives (TN), false positives (FP), and false negatives (FN) are used to calculate the performance of the models. These cases rep-resent the combinations of true and predicted classes in classification problems. A complex metric called the confusion matrix contains the total number of samples TP + TN + FP + FN and allows to calculate specific metrics such as: Accuracy, Recall and Precision [D31]. In this study, three metrics Accuracy, Recall and Precision were used as metrics for networks performance evaluation, calculated by following equations [21]:

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

Accuracy is the relation between the number of correct predictions and the total number of made predictions, as calculated by Equation (1).

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

Precision measures the proportion of correct predictions made by the model, more precisely, the number of items correctly classified as positive out of the total number of items identified as positive. Mathematically, precision is represented in Equation (2).

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

Recall calculates the proportion of correctly identified cases as positive from a total of true positives, as described in Equation (3).

In this study the trained CNN SqueezeNet и GoogLeNet were compared as a function of the metrics Accuracy, Precision and Recall. Additionally, the influence of the Initial Learning Rate (ILR) parameter and the network-tuning algorithm was investigated.

3. Results

In this study, two pre-trained CNN networks, SqueezeNet and GoogLeNet, were fine-tuned using different values of the Initial Learning Rate (ILR) parameter of 0.0001, 0.0002, 0.00025, 0.0003, 0.00035, 0.0004, 0.0005 and three network tuning algorithms, respectively The Stochastic Gradient Descent with moment solver (Sgdm), the Adam optimization algorithm (Adam), and Root Mean Square Propagation (RMSprop). The performance of the networks was evaluated by analyzing the values of Training Accuracy (TA), Training Loss (TL), Validation Accuracy (VA), Validation Loss (VL), and Confusion Matrix. The training of the networks was performed in 1900 iterations and 30 epochs.

Table 1 shows the minimum, maximum and average values for the indicators Training Accuracy, Training Loss, Validation Accuracy and Validation Loss when training and validating the SqueezeNet and GoogLeNet networks using the three Solver algorithms - Sgdm, Adam and RMSprop. When recognizing the five classes of objects corresponding to the five varieties of apples, a very high training accuracy was obtained for both tested networks. For the SqueezeNet network, Training Accuracy has values between 90.91% and 100% and Training Loss between 0.0002 and 0.1166. For GoogLeNet, the training accuracy is between 91.67% and 100% and Training Loss between 0.0002 and 0.0240. It can be concluded that the GoogLeNet network gives slightly better results in its training, since it has fewer losses.
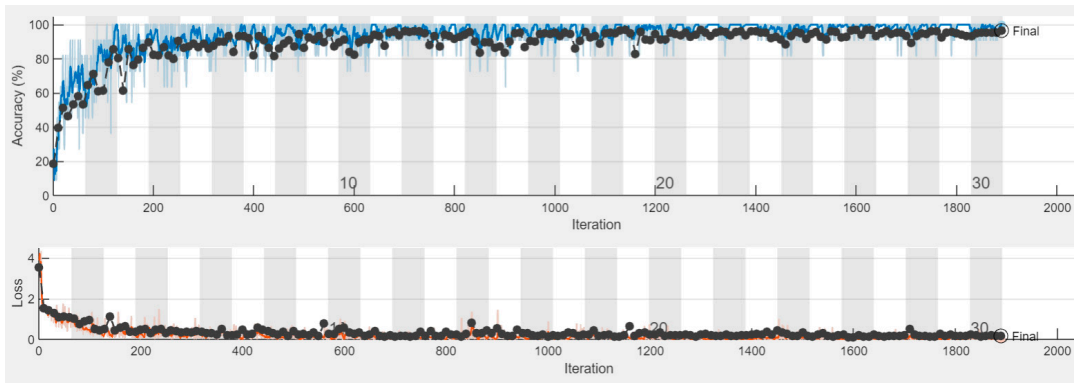
Table 1. Statistical values of network evaluation indicators.

Value	TA, %	TL	VA, %	VL	TA, %	TL	VA, %	VL
Solver Sgdm								
CNN	SqueezeNet	SqueezeNet	SqueezeNet	SqueezeNet	GoogLeNet	GoogLeNet	GoogLeNet	GoogLeNet
Min	90.91	0.0040	90.33	0.1373	100.00	0.0002	91.67	0.1351
Max	100.00	0.1166	96.67	0.3962	100.00	0.0108	96.00	0.3595
Average value	98.70	0.0343	94.05	0.2232	100.00	0.0032	93.14	0.2401
Solver Adam								
CNN	SqueezeNet	SqueezeNet	SqueezeNet	SqueezeNet	GoogLeNet	GoogLeNet	GoogLeNet	GoogLeNet
Min	100.00	0.0025	94.00	0.1453	100.00	0.0002	91.33	0.1251
Max	100.00	0.0304	97.00	0.3052	100.00	0.0240	97.33	0.2841
Average value	100.00	0.0113	95.43	0.2078	100.00	0.0048	95.00	0.2070
Solver RMSprop								

CNN	SqueezeNet	SqueezeNet	SqueezeNet	SqueezeNet	GoogLeNet	GoogLeNet	GoogLeNet	GoogLeNet
Min	100.00	0.0002	94.67	0.1325	91.67	0.0002	94.67	0.0948
Max	100.00	0.0748	96.33	0.2117	100.00	0.0069	98.00	0.1900
Average value	100.00	0.0111	95.33	0.1776	98.81	0.0027	96.62	0.1254

Regarding the validation accuracy, the obtained values are also very high, over 90%, they are in the range of 90.33% to 96.67% for SqueezeNet and 91.33% to 96.62% for GoogLeNet. From the results in Table 1, it can also be concluded that when using RMSprop Solver for both studied networks, the obtained Validation Accuracy is sufficiently high, with the smallest Validation Loss losses.

Figure 4 shows the Training plot SqueezeNet, solver Sgdm, at an average change in ILR - 0.00035, and Figure 5 shows the graph from the training of the GoogLeNet network at the same ILR settings. It can be seen that even after the 400th iteration, the Validation Accuracy for both networks retains values above 80% and the graphs have stable convergence. From the results obtained, it can be concluded that after the 20th epoch there are no significant changes in the accuracy and loss results, and 30 epochs of training and 1900 iterations are completely sufficient to properly train the networks for recognizing the 5 apple varieties from the present study.

Figure 4. Training Accuracy and Training Loss values for SqueezeNet, solver Sgdm at Initial Learning Rate 0.0003.
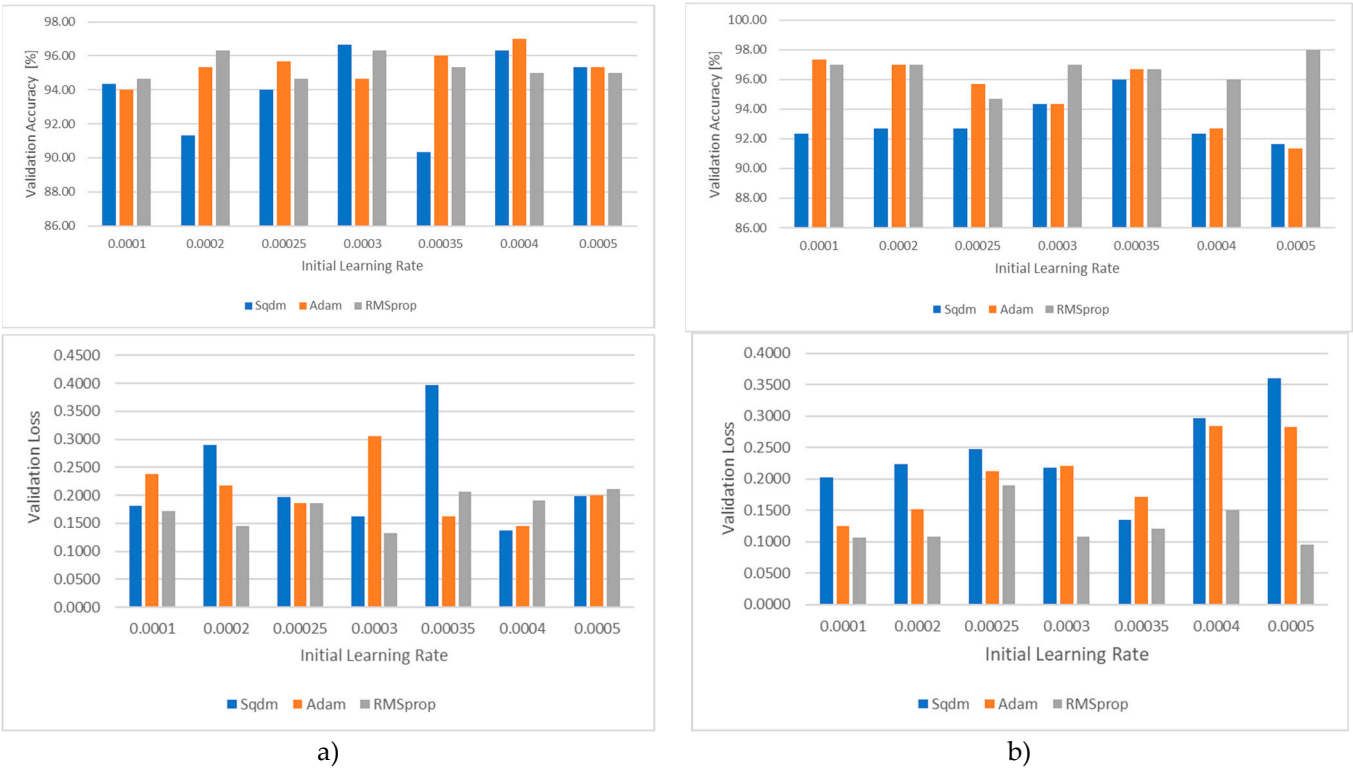

Figure 5. Training Accuracy and Training Loss values for GoogLeNet, solver Sgdm at Initial Learning Rate 0.0003.

Figure 6 graphically shows the values of the Validation Accuracy [%] and Validation Loss parameters for the SqueezeNet (Figure 7a) and GoogLeNet (Figure 7b) networks in more detail for each of the seven tested settings of the Initial Learning Rate parameter.
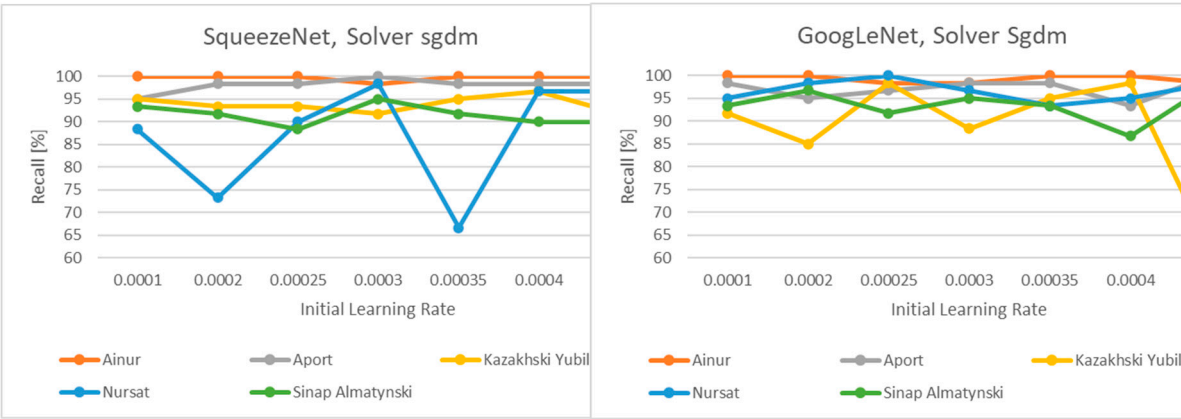
The obtained results show that Validation Accuracy varies from 90 to 98% for all values of ILR change for both networks. The results regarding Validation Loss are in the range from 0.1 to 0.4 for both networks.
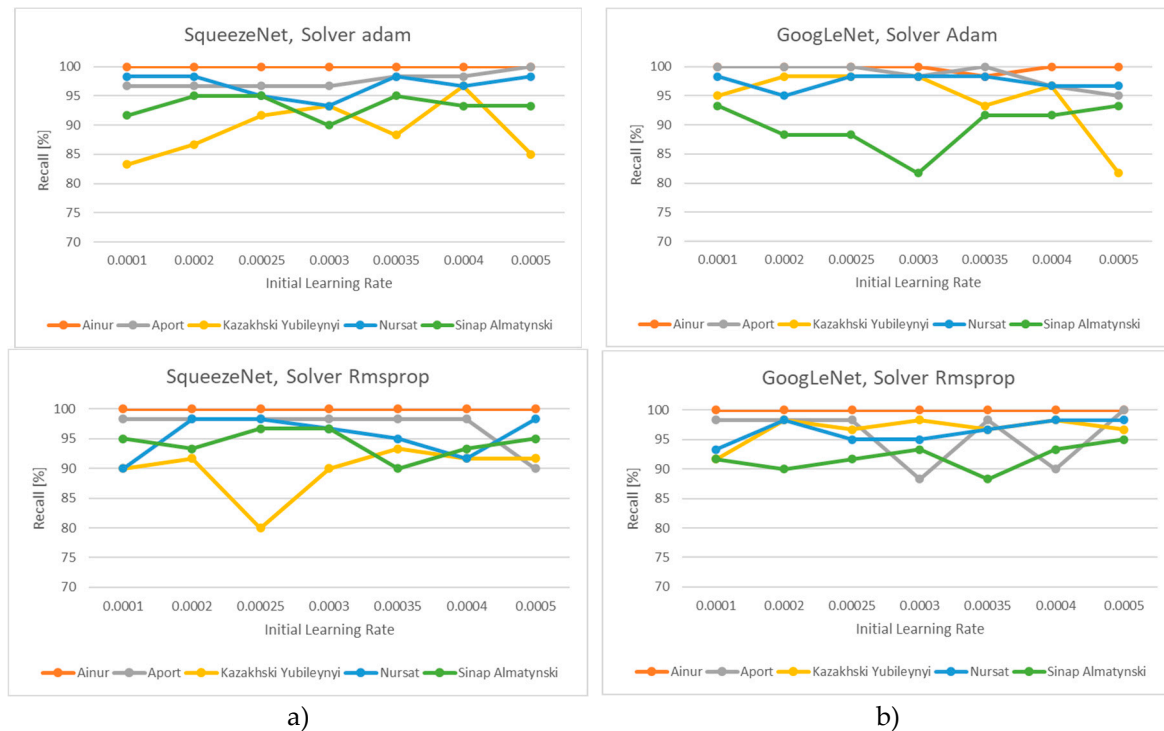
The values of Validation Accuracy for both algorithms Adam and RMSprop are close, regardless of the values of ILR change for the SqueezeNet network, while for the GoogLeNet network – this is observed only for the RMSprop algorithm. Regarding Validation Loss, all three solver algorithms show a dependence on the values of Initial Learning Rate.

Figure 6. Validation Accuracy and Validation Loss values for SqueezeNet (a) and GoogLeNet (b) at different Initial Learning Rates.

GoogLeNet (Figure 7b) networks. According to the Recall indicator for the SqueezeNet network, the varieties Ainur and Aport are not sensitive to changes in the ILR. These two varieties are also not significantly affected by the SqueezeNet network tuning algorithm. For the GoogLeNet network, Ainur, Aport and Nursat are the varieties that are less sensitive to changes in the ILR, while Kazakhski Yubileyinyi and Sinap Almaty are affected to a greater extent by changes in the ILR. For the GoogLeNet network, the tuning algorithm has a slightly more pronounced effect on the Recall parameter values overall.

Figure 7. Recall values for SqueezeNet (a) and GoogLeNet (b) under different network optimization algorithms.

Table 2 systematizes the evaluation indicators Accuracy (%), Precision (%) and Recall (%) of the networks from the validation sample. These indicators are calculated based on the number of samples TP- True Positive, TN-True negative, FP- False positive and FN – False negative, which are part of the known Confusion matrix for more than two classes of objects [31]. When using Confusion matrices in tasks with more than two classes (multiple classes), the concept of "positive" and "negative" classes from binary classification is replaced by the multiple classes. In a multi-class Confusion matrix, the results for the classified samples are interpreted in the following way: by the class to which each element is predicted to belong - by columns, and depending on its true class - by rows. Thus, the diagonal contains the elements that were correctly classified (the predicted class coincides with the true class - True Positive), while the off-diagonal elements show the number of incorrectly classified elements. In a multi-class Confusion matrix, it can be seen whether there are classes that are constantly confused with each other. In such a case, it would be advisable to train the network with additional samples from these classes in order to increase the overall accuracy.
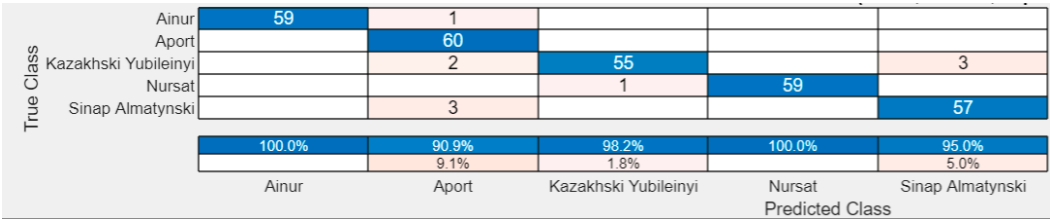
The Table 2 systematizes and shows the values of the evaluation indicators of the SqueezeNet and GoogLeNet networks using the selected best Initial Learning Rate settings for the three solver algorithms. In terms of Accuracy, the values vary between 97.67% and 100%. The highest accuracy is for the Ainur variety – 100%, with the GoogLeNet network and the Adam and RMSprop solvers, and the lowest – 97.67% is for the Sinap Almatynski variety, with the GoogLeNet network and the Sgdm solver. For the Precision indicator, the values range from 90.9 to 100%, with the highest values of 100% obtained for the Kazakhski Yubileyinyi variety, with four variants and settings of the two networks, and the lowest for the Aport variety. For the Recall indicator, the highest values of 100% were obtained for the Ainur variety with five variants for setting the networks, and the lowest for the Kazakhski Yubileyinu variety. In general, higher accuracy is demonstrated when using the GoogLeNet network for all varieties.

Table 2. Experimental results from the validation sample of network evaluation metrics.

DNN	Solver	Initial Learning Rate (ILR)	TP	*TN*	FP	FN	Accuracy, [%]	Precision, [%]	Recall, [%]
Ainur									
SqueezeNet	Sgdm	0.0003	59	240	0	1	99.67	100	98.3
	Adam	0.0004	60	239	0	1	99.67	98.4	100
	RMSprop	0.0003	60	239	1	0	99.67	98.4	100
GoogLeNet	Sgdm	0.00035	60	237	3	0	99.00	95.2	100
	Adam	0.0001	60	240	0	0	100.00	100	100
	RMSprop	0.0005	60	240	0	0	100.00	100	100
Aport									
SqueezeNet	Sgdm	0.0003	60	234	6	0	98.00	90.9	100
	Adam	0.0004	59	235	5	1	98.00	92.2	98.3
	RMSprop	0.0003	59	235	5	1	98.00	92.2	98.3
GoogLeNet	Sgdm	0.00035	59	235	5	1	98.00	92.2	98.3
	Adam	0.0001	60	235	5	0	98.33	92.3	100
	RMSprop	0.0005	60	236	4	0	98.67	93.8	100
Kazakhski Yubileynyi									
SqueezeNet	Sgdm	0.0003	55	241	1	3	98.67	98.2	91.7
	Adam	0.0004	58	239	1	2	99.00	98.3	96.7
	RMSprop	0.0003	54	240	0	6	98.00	100	90
GoogLeNet	Sgdm	0.00035	57	240	0	3	99.00	100	95
	Adam	0.0001	57	240	0	3	99.00	100	95
	RMSprop	0.0005	58	240	0	2	99.33	100	96.7
Nursat									
SqueezeNet	Sgdm	0.0003	59	240	0	1	99.67	100	98.3
	Adam	0.0004	58	240	0	2	99.33	100	96.7
	RMSprop	0.0003	58	239	1	2	99.00	98.3	96.7
GoogLeNet	Sgdm	0.00035	56	239	1	4	98.33	98.2	93.3
	Adam	0.0001	59	239	1	1	99.33	98.3	98.3
	RMSprop	0.0005	59	239	1	1	99.33	98.3	98.3
Sinap Almatynski									
SqueezeNet	Sgdm	0.0003	57	237	3	3	98.00	95	95
	Adam	0.0004	56	238	2	4	98.00	96.6	93.3
	RMSprop	0.0003	58	236	4	2	98.00	93.5	96.7
GoogLeNet	Sgdm	0.00035	56	237	3	4	97.67	94.9	93.3
	Adam	0.0001	56	238	2	4	98.00	96.6	93.3
	RMSprop	0,0005	57	239	1	3	98,67	98,3	95

Figures 8 and 9 show Confusion matrices for the two networks, which obtained the best accuracy indicators for the validation samples, respectively for the SqueezeNet network, Solver Sgdm and ILR=0.0003 (Figure 8) and for the GoogLeNet network with the RMSprop optimization algorithm and ILR=0.0005 (Figure 9).

Figure 8. Confusion matrix for validation set for SqueezeNet, Solver Sgdm and ILR= 0.0003.

Figure 9. Confusion matrix for validation set for GoogLeNet with RMSprop optimization algorithm and ILR=0.0005.

From the figures shown above, it can be seen that the most False Positive samples are recognized by the Aport variety, and in second place by the Sinap Almaty variety. The Ainur variety and the Nursat variety are distinguished well enough from the others, being recognized at 100%, with not a single incorrectly recognized sample with the SqueezeNet network, Solver Sgdm and ILR=0.0003, while with the GoogLeNet network with the RMSprop optimization algorithm and ILR=0.0005, 100% recognition is achieved by Ainur and Kazakhski Yubileynyi. The following Figures 10 and 11 show Confusion matrices for other networks, which obtained the lowest accuracy indicators of the validation samples. Figure 10 shows the matrix for the SqueezeNet network, Solver Sgdm and ILR=0.00035, where the Precision for two of the varieties – Ainur and Aport does not exceed 80 and 83.1%, respectively.
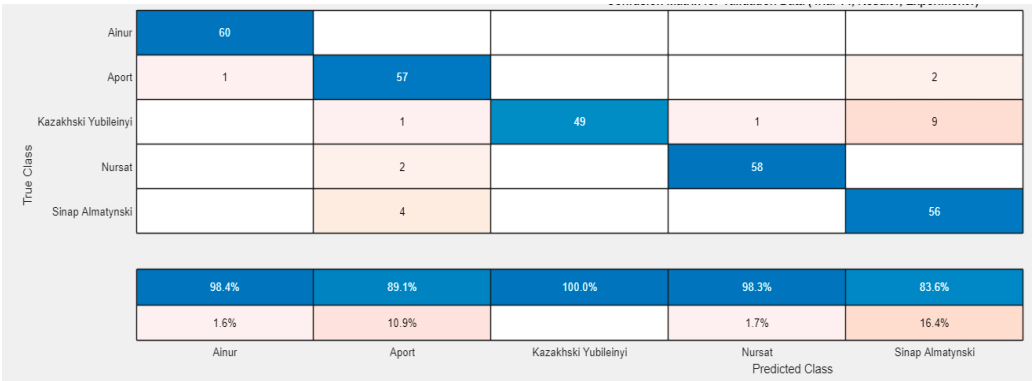

Figure 10. Confusion matrix for validation set for SqueezeNet, Solver Sgdm and ILR= 0.00035.

Figure 11. Confusion matrix for validation set for GoogLeNet with Sgdm optimization algorithm and ILR=0.0005.

In the GoogLeNet network with the Sgdm optimization algorithm and ILR= 0.0005. (Figure 11) the low validation Accuracy is mainly due to the poor recognition of the samples from the Sinap Almaty variety with only 73.10% Precision and from the Nursat variety with 92.20% Precision, while for the other three varieties the recognition is good. For the GoogLeNet network shown in Figure 12 GoogLeNet with the Adam optimization algorithm and ILR = 0.0005. (Figure 12) the Precision for four of the varieties is reduced, respectively for the Aynur variety - 98.40%, for the Aport variety - 89.10%, for the Nursat variety - 98.3% and for the Sinap Almaty variety - 83.6%.

Figure 12. Confusion matrix for validation set for GoogLeNet with Adam optimization algorithm and ILR=0.0005.

5. Discussion

Performance data of the best deep learning network models for classification of the five Kazakh apple varieties, with the corresponding solver tuning algorithms and Initial Learning rate values are presented in Table 5. The variety that showed the best recognition is the Aynur variety. For it, values of 100% were obtained for both Accuracy and Precision and Recall. The remaining four varieties are also recognized with a very high accuracy of over 98.67%.

The proposed approach for recognizing the varietal affiliation of apples using deep learning neural networks is suitable for the analyzed apple varieties and could be easily implemented and used in industrial conditions for sorting fruits. The achieved recognition accuracy meets the requirements in the field.

Table 5. Performance of the best deep learning network models for the five apple varieties.

Apple Variety	CNN model	Solver	ILR	Validation set		
				Accuracy [%]	Precision [%]	Recall [%]
Ainur	GoogLeNet	RMSprop	0.0005	100.00	100.00	100.00
Aport	GoogLeNet	RMSprop	0.0005	98.67	93.80	100.00
Kazakhski Yubileinyi	GoogLeNet	RMSprop	0.0005	99.33	100.00	96.70
Nursat	SqeezeNet	Sgdm	0.0003	99.67	100.00	98.30
Sinap Amaty	GoogLeNet	RMSprop	0.0005	98.67	98.30	95.00

6. Conclusions

Increased requirements for fruit quality and the growth of apple production in Kazakhstan require the use and implementation of new technologies in the classification of apples by varietal affiliation.

The analysis of the literature confirms that new deep learning techniques are precise, with high accuracy and are increasingly used in the assessment of quality and classification of agricultural products.

The developed approach for the identification of five varieties of Kazakh apples using deep learning techniques achieved 100% correct classification of fruits for the Ainur variety with a GoogLeNet network, solver RMSprop and ILR= 0.0005. For the varieties Aport, Kazakhski Yubileinyi and Nursat, one of the three network evaluation indicators achieves 100% accuracy, and for Sinap Almatynski – all three indicators are with and above the value of 95%. For varieties Ainur, Aport, Kazakhski Yubileinyi and Sinap Almatynski as the best (with the highest optimal accuracy values) is

the GoogLeNet network, with the following settings: solver RMSprop and ILR=0.0005. Only for variety Nursat is the SqeezeNet network suitable, with the following settings: solver Sgdm and ILR= 0.0003.

The conducted experimental studies showed that 30 training epochs and 1900 iterations are quite sufficient to properly train the networks for recognizing the 5 apple varieties from this study.

The proposed approach could be implemented in automated machines for sorting apples by variety, which will increase their productivity and process functionality.

The obtained additionally trained CNN networks can successfully complement the methodology developed in [30] for assessing the quality indicators of apples and serve as the basis for the development of a compact tool for assessing the quality and varietal affiliation of apples.

Author Contributions: Conceptualization, Plamen Daskalov; Project administration, Jakhfer Alikhanov; Data curation, Aidar Moldazhanov, Akmaral Kulmakhambetova and Zhandos Shynybay; Resources, Dmitriy Zinchenko and Alisher Nurtuleuov; Software and Formal analysis, Tsvetelina Georgieva; Writing—original draft preparation, Jakhfer Alikhanov, Tsvetelina Georgieva and Eleonora Nedelcheva, Visualization, Eleonora Nedelcheva , Writing – review & editing, Plamen Daskalov.

Data Availability Statement: The raw/processed data required to reproduce these findings cannot be shared at this time as the data also form part of an ongoing study. However, the data can be provided to readers when kindly requested.

Conflicts of Interest: The authors have no conflicts of interest to declare.

References

1. Robinson, T. Advances in apple culture worldwide, Revista Brasileira de Fruticultura **2011**, 33:37-47, https://doi,org/10,1590/S0100-29452011000500006

2. Morariu, P.; Mureşan, A.; Sestras, A.; Dan, C.; Andrecan, A.F.; Borsai, O.; Militaru, M.; Mureşan, V.; Sestras, R. E. The impact of cultivar and production conditions on apple quality, Notulae Botanicae Horti Agrobotanici Cluj-Napoca, 53, **2025** 14046. 10.15835/nbha53114046.

3. https://internet-law.ru/gosts/gost/66071/

4. Ma, J.; Sun, D.W.; Qu, J.H.; Liu, D.; Pu, H.; Gao, W.H.; Zeng, X.A. Applications of computer vision for assessing quality of agri-food products: A review of recent research advances, Crit. Rev. Food Sci. Nutr. **2016**, 56, 113–127.

5. Mizushima, A.; Lu, R. An image segmentation method for apple sorting and grading using support vector machine and Otsu's method, Comput. Electron. Agric. 94, **2013,** 29–37.

6. Mirbod, O.; Choi, D.; Heinemann, P.H.; Marini, R.P.; He, L. On-tree apple fruit size estimation using stereo vision with deep learning-based occlusion handling, Biosystems Engineering, 226, **2023**, 27–42.

7. Moallem, P.; Serajoddin, A.; Pourghassem, H. Computer vision based apple grading for golden delicious apples based on surface features, Inf Process Agric **2017**;4(1):33–40.

8. Roberts, C.A.; Workman, J.; Reeves, J.B. Near-Infrared Spectroscopy in Agriculture, (IM Publications, Boston, 2004) A,K, Bhatt, D, Pant, AI Soc, 30, 45, **2015**. https://doi.org/10.1007/s00146-013-0516-5

9. Nwosisi, S.; Dhakal, K.; Nandwani, D.; Raji, J.I. Krishnan, S.; Beovides-García, Y. Genetic Diversity in Vegetable and Fruit Crops, In: Nandwani, D, (eds) Genetic Diversity in Horticultural Plants, Sustainable Development and Biodiversity, vol 22, Springer, Cham, **2019**, https://doi.org/10.1007/978-3-319-96454-6_4

10. Campeanu, G.; Neata, G.; Darjanschi, G. Chemical Composition of the Fruits of Several Apple Cultivars Growth as Biological Crop Not, Bot, Hort, Agrobot, Cluj 37 (2) **2009**, 161-164.

11. Pilco, A.; Moya, V.; Quito, A.; Vásconez, J.P; Limaico, M. Image Processing-Based System for Apple Sorting, Journal of Image and Graphics, Vol. 12, No. 4, **2024**, 362-371, doi: 10.18178/joig.12.4.362-371

12. Miranda, J.; Arno, J.; Gene-Mola, J.; Lordan, J.; Asin, L.; Gregorio, E. Assessing automatic data processing algorithms for RGB-D cameras to predict fruit size and weight in apples, Computers and Electronics in Agriculture, 214, **2023**, 1-15.

13. Li, Y.; Feng, X.; Liu, Y. et al. Apple quality identification and classification by image processing based on convolutional neural networks, Sci Rep 11, 16618, **2021**, https://doi.org/10.1038/s41598-021-96103-2

14. Garrido-Novell, C.; Pérez-Marin, D.; Amigo, J.M.; Fernández-Novales, J.; Guerrero, J.E.; Garrido-Varo, A. Grading and color evolution of apples using RGB and hyperspectral imaging vision cameras, J. Food Eng, 113, **2012**, 281–288.

15. Li, Q.; Wang, M.; Gu, W. "Computer vision based system for apple surface defect detection," Computers and Electronics in Agriculture, vol. 36, no, 2-3, **2002**, 215–223.

16. Xiao-bo, Z.; Jie-wen, Z.; Yanxiao, L.; Holmes, M. In-line detection of apple defects using three color cameras system, Computers and Electronics in Agriculture, vol. 70, no. 1, **2010**, 129–134.

17. Lee, J.H.; Vo, H.T.; Kwon, G.J.; Kim, H.G.; Kim, J.Y. Multi-Camera-Based Sorting System for Surface Defects of Apples, Sensors **2023**, 23, 3968, https://doi.org/ 10.3390/s23083968

18. Amrutha, M.; Kousalya, S.; Ananya, R.; Harshitha, D. A. Grading of Apple Fruit Using Image Processing Techniques, International Journal of Advanced Research in Science, Communication and Technology (IJARSCT) Volume 8, Issue 1, **2021**, ISSN (Online) 2581-9429.

19. Kamilaris, A. Prenafeta-Boldú, F. X. Deep learning in agriculture: a survey, Computers and Electronics in Agriculture, vol. 147, **2018**, 70–90.

20. Zhu, L.; Spachos, P.; Pensini, E.; Plataniotis, K. N. Deep learning and machine vision for food processing: A survey, Current Research in Food Science, 4, **2021,** 233–249, https://doi.org/10.1016/j.crfs.2021.03.009

21. Ibarra-Pérez, T.; Jaramillo-Martínez, R.; Correa-Aguado, H.C.; Ndjatchi, C.; Martínez-Blanco, M.D.R.; Guerrero-Osuna, H.A.; Mirelez-Delgado, F.D.; Casas-Flores, J.I.; Reveles-Martínez, R.; Hernández-González, U.A. A Performance Comparison of CNN Models for Bean Phenology Classification Using Transfer Learning Techniques, AgriEngineering 6, **2024,** 841–857, https://doi.org/10.3390/agriengineering6010048

22. Yulita, I.N.; Rambe, M.F.R.; Sholahuddin, A.; Prabuwono, A.S. Convolutional Neural Network Algorithm for Pest Detection Using GoogleNet, AgriEngineering 5**, 2023**, 2366–2380, https://doi.org/10.3390/agriengineering5040145

23. Rady, A.; Fisher, O.; El-Banna, A.A.A.; Emasih, H.H.; Watson, N.J. Computer Vision and Transfer Learning for Grading of Egyptian Cotton Fibres, AgriEngineering **2025**, 7, 127, https://doi,org/10,3390/agriengineering7050127

24. Yunong, T.; Guodong, Y,; Zhe, W.; En, L.; Zize, L. Detection of Apple Lesions in Orchards Based on Deep Learning Methods of CycleGAN and YOLOV3-Dense, Journal of Sensors Volume **2019**, 13 pages https://doi.org/10.1155/2019/7630926

25. Yanfei, L.; Xianying, F.; Yandong, L.; Xingchang, H. Apple quality identification and classification by image processing based on convolutional neural networks, Scientific Reports, 11:16618, **2021** https://doi.org/10.1038/s41598-021-96103-2

26. https://unece.org/sites/default/files/2020-12/50_Apples.pdf

27. Yuesheng, F.; Jian, S.; Fuxiang, X.; Yang, B.; Xiang, Z.; Peng, G.; Zhengtao, W.; Shengqiao, X. Circular Fruit and Vegetable Classification Based on Optimized GoogLeNet, IEEE Access **2021**, 9, 113599–1135611.

28. Iandola, F.N.; Han, S.; Moskewicz, M.W.; Ashraf, K.; Dally, W.J.; Keutzer, K. SqueezeNet: AlexNet-level accuracy with 50x fewer parameters and< 0.5 MB model size, arXiv **2016**, arXiv:1602.07360.

29. Zeng, K.; Liu, J.; Jiang, Z.; Xu, D. A Scaling Transition Method from SGDM to SGD with 2ExpLR Strategy, Appl, Sci, **2022**, 12, 12023, https://doi.org/10.3390/app122312023

30. Alikhanov, J.; Moldazhanov, A.; Kulmakhambetova, A.; Zinchenko, D.; Nurtuleuov, A.; Shynybay, Z.; Georgieva, T.; Daskalov, P. Methodology for Determining the Main Physical Parameters of Apples by Digital Image Analysis, AgriEngineering **2025**, 7, 57, https://doi.org/10.3390/ agriengineering7030057

31. Qin, J.; Hu, T.; Yuan, J.; Liu, Q.; Wang,W.; Liu, J.; Guo, L.; Song, G. Deep-Learning-Based Rice Phenological Stage Recognition. Remote Sens. **2023**, 15, 2891. [CrossRef]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.