

## Article

# A Deep Learning Image System for Classifying High Oleic Sunflower Seed Varieties

Mikel Barrio-Conde <sup>1</sup>, Marco Antonio Zanella <sup>2</sup> , Javier Manuel Aguiar-Perez <sup>1</sup> , Ruben Ruiz-Gonzalez <sup>3</sup> and Jaime Gomez-Gil <sup>1,\*</sup> 

- <sup>1</sup> Departamento de Teoría de la Señal y Comunicaciones e Ingeniería Telemática, Universidad de Valladolid, ETSI Telecomunicación, Paseo de Belén 15, 47011 Valladolid, Spain
- <sup>2</sup> Agricultural Engineering Department, Federal University of Lavras, P.O. Box 3037, Lavras 37200-000, Brazil
- <sup>3</sup> Department of Electromechanical Engineering, Escuela Politécnica Superior, University of Burgos, Avda. Cantabria s/n, 09006 Burgos, Spain
- \* Correspondence: [jgomez@tel.uva.es](mailto:jgomez@tel.uva.es); Tel.: +34-(98)-3185556

**Abstract:** Sunflower seeds, one of the main oilseeds produced around the world, are widely used in the food industry. Mixtures of seed varieties can occur throughout the supply chain. Intermediaries and the food industry need to identify the varieties to produce high-quality products. Considering that high oleic oilseed varieties are similar, a computer-based system to classify varieties could be useful to the food industry. The objective of our study is to examine the capacity of deep learning (DL) algorithms to classify sunflower seeds. An image acquisition system, with controlled lighting and a Nikon camera in a fixed position, was constructed to take photos of 6000 seeds of six sunflower seed varieties. Images were used to create datasets for training, validation, and testing of the system. A CNN AlexNet model was implemented to perform variety classification, specifically classifying from two to six varieties. The classification model reached an accuracy value of 100% for two classes and 89.5% for the six classes. These values can be considered acceptable, because the varieties classified are very similar, and they can hardly be classified with the naked eye. This result proves that DL algorithms can be useful for classifying high oleic sunflower seeds.

**Keywords:** classification system; convolutional neural network; high oleic sunflower seed



**Citation:** Barrio-Conde, M.; Zanella, M.A.; Aguiar-Perez, J.M.; Ruiz-Gonzalez, R.; Gomez-Gil, J. A Deep Learning Image System for Classifying High Oleic Sunflower Seed Varieties. *Sensors* **2023**, *23*, 2471. <https://doi.org/10.3390/s23052471>

Academic Editors: Marco Santonico and Luca Vollerio

Received: 20 January 2023  
Revised: 17 February 2023  
Accepted: 21 February 2023  
Published: 23 February 2023



**Copyright:** © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction

Agriculture and the food industry are facing issues to meet the demand of future generations for higher food production and healthier food [1]. The global population is estimated to rise to 9.7 billion people by 2064 [2]. Oilseeds are extensively cultivated throughout the world because oilseed-derived products are used as inputs in several stages of the food supply chain. The worldwide harvest of oilseeds in 2021/2022 amounted to 619.17 million metric tons [3]. The main producers of sunflowers (*Helianthus annuus* L.) are Brazil, the United States, and China. The main oilseeds that are cultivated around the world are sunflower seeds, soybean, and rapeseed [3].

Edible oil extract from oilseeds contains compounds with many health benefits and quality attributes that are important for the food industry [4]. There are two main types of food oils: conventional oil and high oleic oil. Conventional oil contains high contents of linoleic acid, while high oleic oil has a content of oleic acid that is 80% or higher and very low levels of linoleic acid. High oleic oil is the most stable type of oil for industrial use [5]. Besides, there is evidence that high oleic oil is a healthy choice for use in food production [6,7]. The food industry needs a substitute for trans fatty acids (TFA), which are often used in fried snacks and bakery products and have been associated with health problems [8]. Zambelli concluded that high oleic sunflower oil is the optimal substitute for TFA, due to its fatty acid profile and good agronomic performance [9].

Seeds, oil, and oilcake derived from sunflowers have numerous components suitable for a healthy diet [10]. After the oil extraction process, the largest product obtained is the oilseed press cake, which can be processed for human consumption [11], and is used in animal feed [12]. Additionally, sunflower-derived products have important uses, such as energy generation and industrial production of detergents, adsorbents, paints, lubricants, and adhesives, among others [13–16].

Sunflower production has been increasing over recent years, in response to the demands of the food industry and as a useful cash crop for farmers. Plant breeding methods have been applied to develop sunflower varieties, improving oil content and quality for industrial purposes [17]. The agronomic characteristics of sunflower varieties have also been improved, so that varieties can be adapted to adverse growing environments [18,19]. There are sunflower varieties that can differ in oil quality, yield, and by-products, as a result of genetic modifications [20]. Sunflower seed varieties may be mixed during harvesting, transportation, and storage processes. Seed classification systems are therefore useful to trade intermediaries and the food industry, to ensure that sunflower seeds are suitable for industrial requirements.

The manual classification of seed varieties can be biased and time-consuming, since physical attributes of seeds are normally hard to differentiate with the human eye. Researchers have been applying conventional image processing to overcome this classification problem successfully. In general, conventional approaches include: (i) a preprocessing stage, where processes such as cropping, scaling, color conversion, noise filtering, and normalization are applied [21]; (ii) a segmentation stage, where thresholding or watershed techniques are applied [21,22]; (iii) a feature extraction stage, where shape, morphological, texture, and color features are extracted [23,24]; and (iv) a classification stage, where classical machine learning techniques are applied, in the form of algorithms such as Gaussian Naive Bayes, random forest (RF), support vector machines (SVM), k-nearest neighbors (kNN), and artificial neural networks (ANN) [23–27]. Nevertheless, those traditional approaches, based on computer vision and feature extraction, could not guarantee a high enough accuracy of classification in most cases.

Deep learning (DL) is a modern approach considered as a subset of machine learning. It uses ANN to model high-level abstractions of data as the data volume increases. DL algorithms can be divided into four categories: restricted Boltzmann machines (RBM), autoencoders, sparse coding, and convolutional neural networks (CNN), the last-mentioned being the most suitable algorithms in the classification of objects in images [28]. Various CNN models have been tested to classify seeds [29–32], and researchers have suggested modifications to some existing models [33,34]. Zhao et al. developed a system to detect soybean seed surface defects with DL [35]; Nie et al. and Qiu et al. applied DL to classify seeds using hyperspectral imaging [36,37]; Loddo et al. proposed a novel CNN model for seed species classification [38]; and Gao et al. developed a CNN model for the classification and recognition of wheat varieties, through the combination of wheat plant images in the tilling season, images of wheat plants in the flowering period, and images of the wheat seed [39].

DL has become popular in seed classification [22]. Its popularity is due to its easy application and high seed classification accuracy. It requires no further image processing despite preprocessing [40] and, in general, DL offers better results than classical machine learning approaches [41]. Lin et al. compared DL models with traditional hand-engineered approaches for rice kernel classification. The accuracy values of traditional approaches varied between 89.1% and 92.1%, while DL models achieved 95.5% [42]. Koklu et al. compared the performance of ANN models, deep ANN models, and DL models, for the classification of five varieties of rice, achieving classification accuracy values of 99.87%, 99.95%, and 100%, respectively [43]. Mukasa et al. tested multivariate machine learning and DL models in the discrimination of triploid watermelon seeds from diploid and tetraploid seeds. They reported accuracy values of 69.5% and 84.3% for multivariate machine learning, and 95.5% in DL models [44]. Finally, Przybyło and Jabłoński, using DL models for the

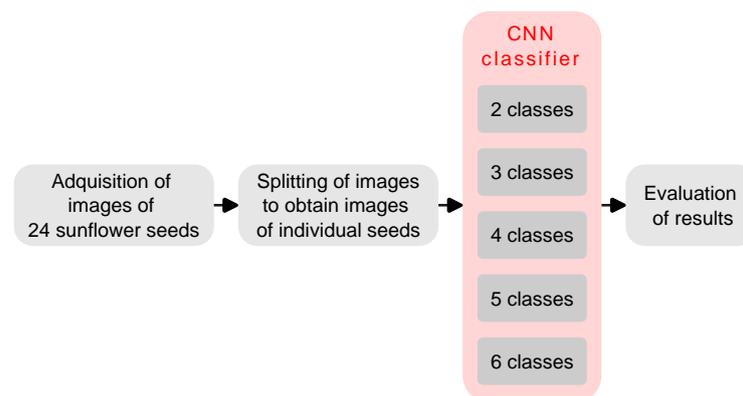
classification of oak acorns, achieved slightly higher accuracy values than those achieved by human experts making manual predictions [40].

Regarding sunflower seeds, to the best of our knowledge, there are two studies that applied CNN to classify sunflower seeds. Kurtulmuş, in 2021, achieved 95% accuracy using a CNN model to classify sunflower seeds of some varieties cultivated in Turkey [45]. Li et al. developed a custom CNN model to sort sunflower seeds from inert material (leaves, stones, and defective seeds), achieving accuracies between 97.33% and 99.56% [46].

The aim of the present study is to classify up to six different varieties of sunflower seeds using a CNN model, and later assess the accuracies of the classification model. The main novelty of this article with respect to the existing literature is the comparison of accuracy achieved when increasing the number of varieties from two to six. Moreover, the varieties chosen in this study are different from the ones chosen in other published articles, and those seeds are visually less similar than the ones employed in this article.

## 2. Materials and Methods

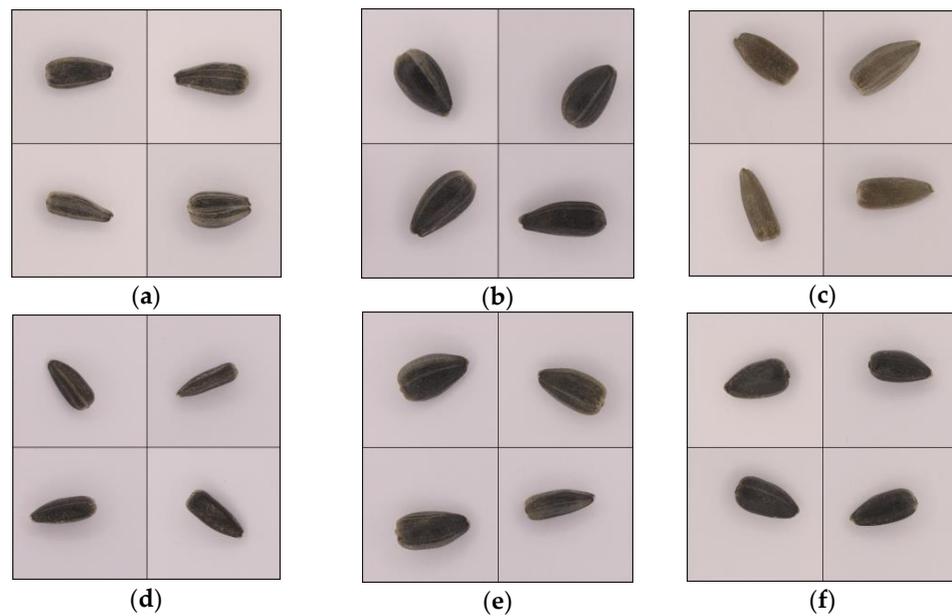
A sunflower seed classifier system using a CNN model was implemented. The following paragraphs describe the seeds to be classified, the image acquisition system, the CNN model that was implemented to classify the seed varieties, and the hardware equipment and software employed to perform all the underlying processing. Figure 1 shows a block diagram of the processing methodology applied in the present study.



**Figure 1.** Block diagram of the processing methodology.

### 2.1. Sunflower Seeds

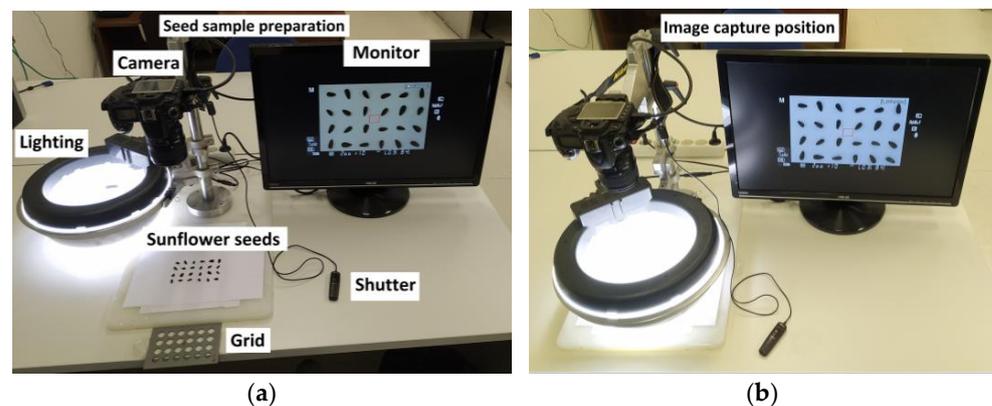
The following six sunflower (*Helianthus annulus* L.) seed varieties were classified: Talisman, from RAGT Seeds Group (Ickleton, UK); MAS830 and MAS89, from Maisadour Cooperative (Haut-Mauco, France); Kaledonia, from Caussade Semences Group (Caussade, France); Gibraltar, from Syngenta (Basel, Switzerland); and Orientes, from KWS SAAT SE & Co. KGaA (Einbeck, Germany). All the sunflower seeds employed in this article were harvested in the year 2020, and they were distributed to the authors by Sovena Oilseeds España (Brenes, Spain). The seeds were manually cleaned of foreign material such as dirt, char, leaf debris, dust, and stones. Immature and defective seeds were discarded before image acquisition and only ripe and flawless seeds were selected for further processing. Sunflower seeds were stored in plastic buckets and kept at room temperature (20–25 °C). Sample images of these varieties are shown in Figure 2.



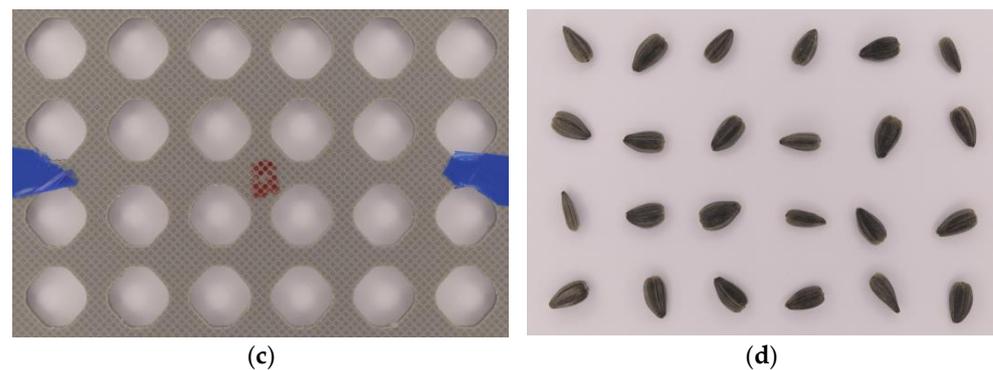
**Figure 2.** Example images of the six sunflower varieties used in the study: (a) Talissman, (b) MAS830, (c) MAS89, (d) Kaledonia, (e) Gibraltar, and (f) Orientes.

## 2.2. Image Acquisition Setup

The image acquisition system comprised: (i) a Nikon D90 camera equipped with a 60 mm AFD Micro-Nikkor lens  $f/2.8$  to acquire the images; (ii) a steel bracket to mount the camera; (iii) an LED ring light, 14 inches in diameter, with 180 LEDs and 36 watts of power, to illuminate the seeds; (iv) a circular reflective metal ring placed around the seeds for consistent lighting; (v) a remote shutter to avoid trepidation; and (vi) a grid to place groups of 24 sunflower seeds in a fixed position. The distance between the camera lens and the seeds was 32 cm. In the camera configuration: (i) the shutter speed was fixed to  $1/200$ , because lower speeds produced trepidation and results were no better at higher speeds; (ii) the working aperture was fixed at  $f:10$  because wider apertures gave a worse depth of field and some details of the object were lost, and narrow apertures yielded a worse resolution and softer details; (iii) the ISO sensitivity was fixed at 200, because higher ISO produced more noise and, at a lower ISO, the results were no better and could provoke trepidation; and (iv) the white balance was fixed in flash mode to avoid variations of color temperature. Images were saved in JPG format at  $4288 \times 2848$  resolution, employing the configuration in the camera that minimized the quality losses in the RAW to JPG conversion. The components of the image system acquisition in the position of seed sample preparation and also in the position of image capture are shown in Figure 3.



**Figure 3.** Cont.



**Figure 3.** (a) Image system acquisition in the position of seed sample preparation, showing the camera, monitor, and lighting together with the grid used for positioning the seeds, and the remote shutter employed to trigger a photo; (b) image system acquisition in the position of image capture; (c) grid employed to place the 24 seeds into their position for acquisition; and (d) final position of the seeds after using the grid for their placement.

A total of 42 photos, with each photo containing a total of 24 seeds (Figure 3d), were taken for each of the six sunflower seed varieties under study.

### 2.3. Splitting of Seed Images

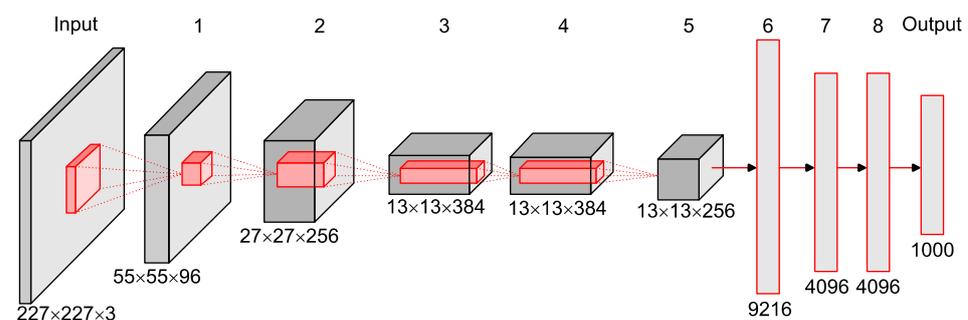
As explained in the previous section, the employed acquisition setup took a photograph of 24 seeds at the same time, placed in a grid of four rows and six columns (Figure 3d). Thus, it was necessary to split each seed into a different image. This process was made easier due to the use of a grid for the placement of the seeds, being achieved by using a constant mask for each cell, since the seeds were always placed at the same location within a bounding square. In this stage, the input images with  $4288 \times 2848$  resolution containing 24 sunflower seeds were split to obtain images of individual seeds, with  $712 \times 712$  pixels for each one. In addition, the size of each image was later reduced to a resolution of  $227 \times 227$  pixels to match the input size of the CNN classifier. The scaling of the images was done using the OpenCV resize function in Python, with a default interpolation method.

After this stage, a dataset containing 6000 seed images, 1000 from each variety, was generated and used in the subsequent classification process.

### 2.4. CNN Classifier

The CNN architecture employed for classifying sunflower seeds in this work was AlexNet, originally proposed by Krizhevsky et al. [47]. This particular model was chosen, after a careful review of the literature, due to it outperforming other existing models when tackling similar problems to the one in this work.

Figure 4 shows the architecture of AlexNet, with all involved layers and connections, while Table 1 shows the main parameters of each layer of this architecture.



**Figure 4.** Architecture of the AlexNet CNN model.

**Table 1.** Specific parameters of each layer of the AlexNet CNN architecture.

	Layer	Feature Map	Size	Kernel Size	Stride	Activation
Input	Image	1	$227 \times 227 \times 3$	-	-	-
1	Convolution	96	$55 \times 55 \times 96$	$11 \times 11$	4	relu
	Max pooling	96	$27 \times 27 \times 96$	$3 \times 3$	2	relu
2	Convolution	256	$27 \times 27 \times 256$	$5 \times 5$	1	relu
	Max pooling	256	$13 \times 13 \times 256$	$3 \times 3$	2	relu
3	Convolution	384	$13 \times 13 \times 384$	$3 \times 3$	1	relu
4	Convolution	384	$13 \times 13 \times 384$	$3 \times 3$	1	relu
5	Convolution	256	$13 \times 13 \times 256$	$3 \times 3$	1	relu
	Max pooling	256	$6 \times 6 \times 256$	$3 \times 3$	2	relu
6	FC	-	9216	-	-	relu
7	FC	-	4096	-	-	-
8	FC	-	4096	-	-	relu
Output	FC	-	1000	-	-	softmax

The CNN model that was implemented, AlexNet, was fitted with the parameters listed in Table 1, and trained to maximize the classification accuracy using the Adam optimizer, minimizing losses with the cross-entropy function. Six sunflower varieties were used to implement and test the classifier model, varying the number of classes applied from two to six: (i) two classes (MAS830 and MAS89); (ii) three classes (Kaledonia, MAS830, and MAS89); (iii) four classes (Kaledonia, MAS830, MAS89, and Talissman); (iv) five classes (Kaledonia, MAS830, MAS89, Orientes, and Talissman); (v) six classes (Gibraltar, Kaledonia, MAS830, MAS89, Orientes, and Talissman). Table 2 shows the specific parameters of the AlexNet CNN model used for each combination of classes.

**Table 2.** Specific parameters of the AlexNet CNN model.

Factor	Value
Image input size	$227 \times 227 \times 3$
Epochs	50
Optimizer	Adam
Loss function	Cross-entropy
Batch size	32
Validation size	30
Learning rate	0.00001
Dropout	0.5

The distribution of the dataset was randomly divided into three parts: training, validation, and test blocks. From the 1000 images of each variety, 700 images were used for training (70%), 100 for the validation (10%), and 200 for the testing (20%). This division of the dataset into training–validation–test was chosen following the same approach employed in other similar articles, for a more direct comparison and also for the sake of obtaining a unique set of parameters after training instead of several sets of parameters, which happens when using cross-validation.

### 2.5. Evaluation of Results

Based on the results, the following evaluation metrics were employed: accuracy, recall, precision, loss, F1-score, ROC (receiver operating characteristic) curve, AUC (area under ROC curve), and the normalized confusion matrix.

### 2.6. Hardware Equipment and Software

The hardware equipment for the processing was a workstation with Intel(R) Xeon(R) CPU at 2.00 GHz, and an NVIDIA Tesla T4 graphic card with 16 GB of memory.

All the software involved in the processing in this article was programmed in Python, language version 3. For programming all the software related to the AlexNet CNN model, the TensorFlow library was used, in version 2.9.2.

### 3. Results

Different experiments were performed to analyse the AlexNet CNN model with the configuration given in Table 2. These experiments were based on varying the number of classes applied to the model from two to six: (i) two classes (MAS830 and MAS89); (ii) three classes (Kaledonia, MAS830, and MAS89); (iii) four classes (Kaledonia, MAS830, MAS89, and Talissman); (iv) five classes (Kaledonia, MAS830, MAS89, Orientes, and Talissman); (v) six classes (Gibraltar, Kaledonia, MAS830, MAS89, Orientes, and Talissman).

Figure 5 shows five graphs, one per experiment, with both the accuracy and the loss evolution during the training and validation phases, depending on the epoch. The horizontal axis represents the number of epochs. The vertical axis represents the value of the accuracies (red and yellow) and the losses (green and blue) during the training and validation phases. The graph shows that: (i) initially, as the number of epochs increased, the loss decreased, and therefore the accuracy increased; and (ii) when the number of epochs reached four, the curves began to stabilize.

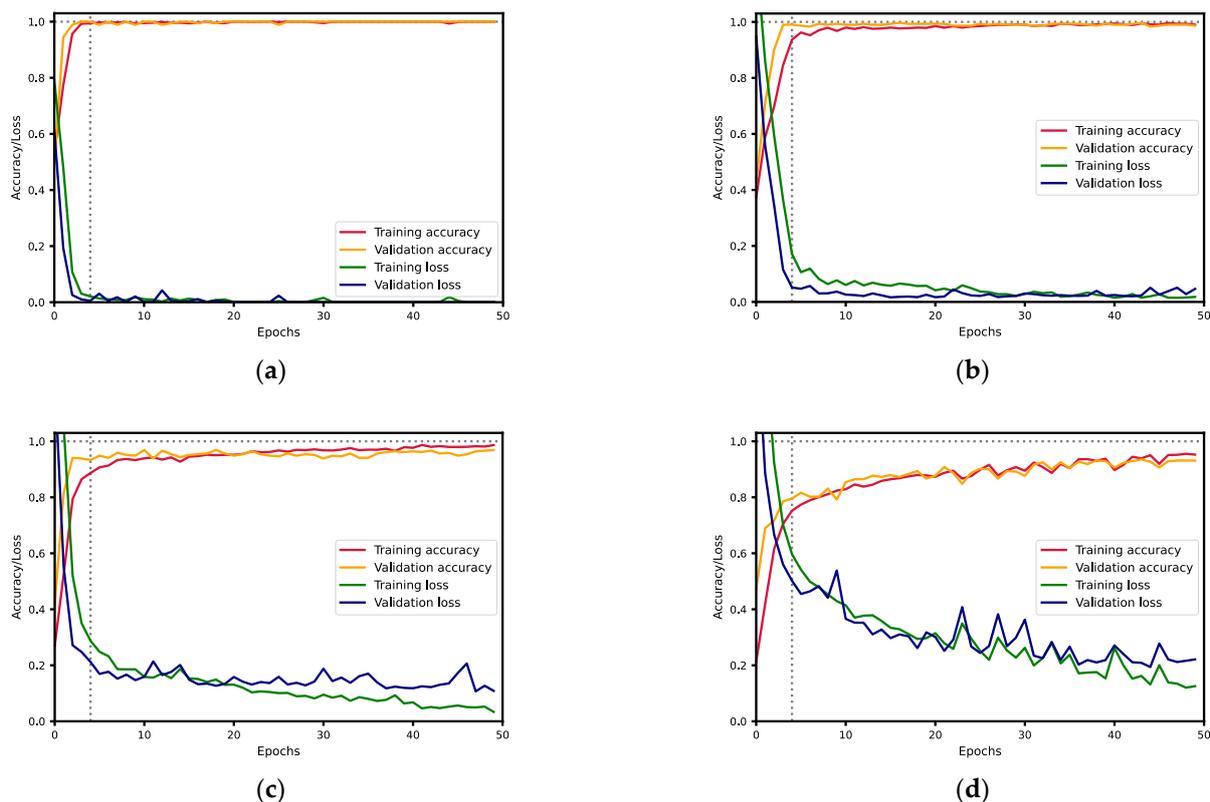
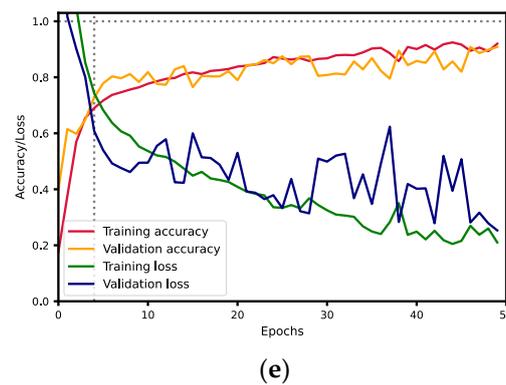


Figure 5. Cont.



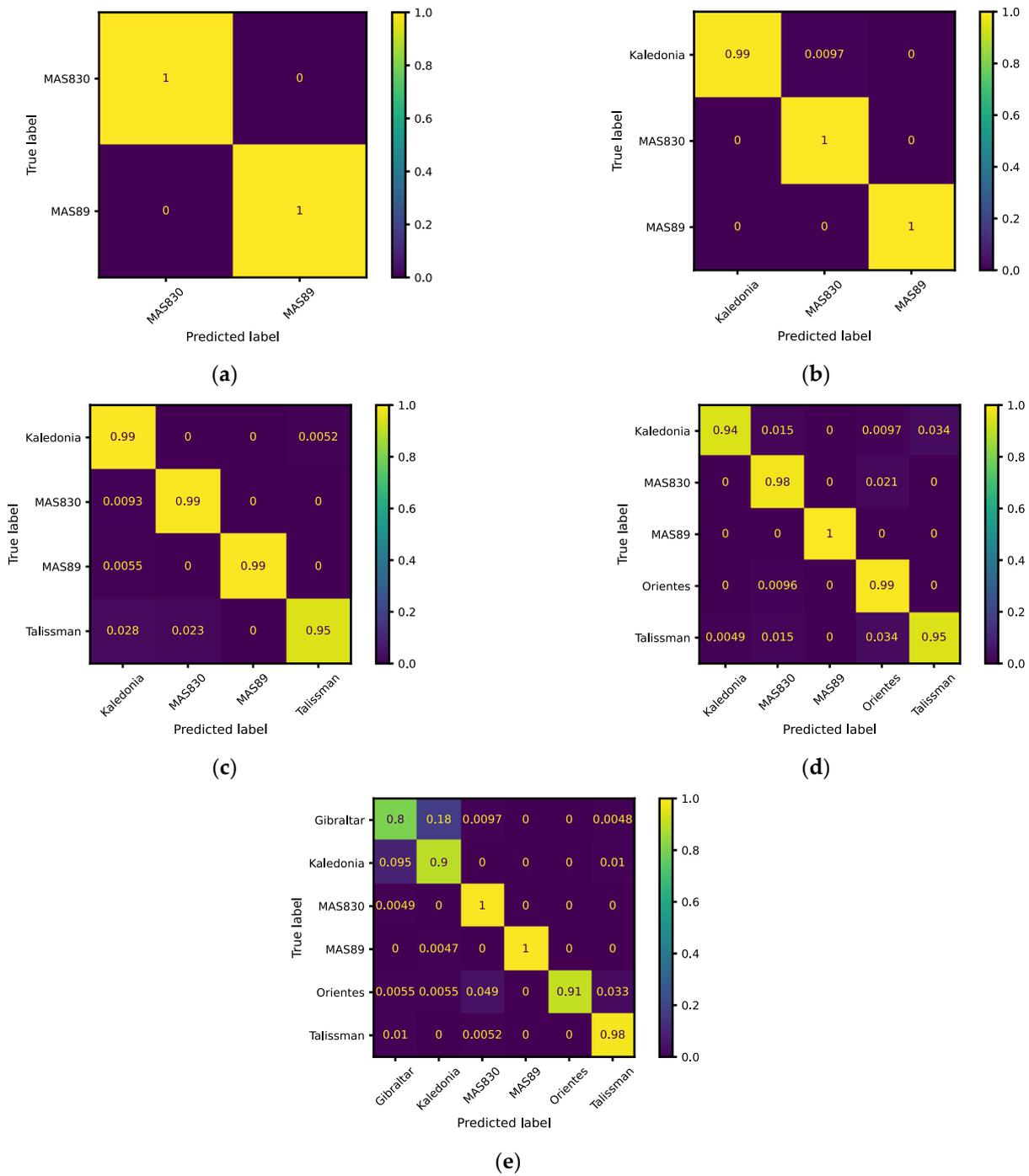
**Figure 5.** Curves of training and validation loss and accuracy for the different experiments: (a) two classes (MAS830 and MAS89); (b) three classes (Kaledonia, MAS830, and MAS89); (c) four classes (Kaledonia, MAS830, MAS89, and Talissman); (d) five classes (Kaledonia, MAS830, MAS89, Orientes, and Talissman); (e) six classes (Gibraltar, Kaledonia, MAS830, MAS89, Orientes, and Talissman).

The test dataset was used to verify the configuration of the model. The results of the model during this test phase are shown in Table 3 for the different experiments. These results include values of recall, precision, F1-score, AUC, accuracy, and loss.

**Table 3.** Classification results during the test phase.

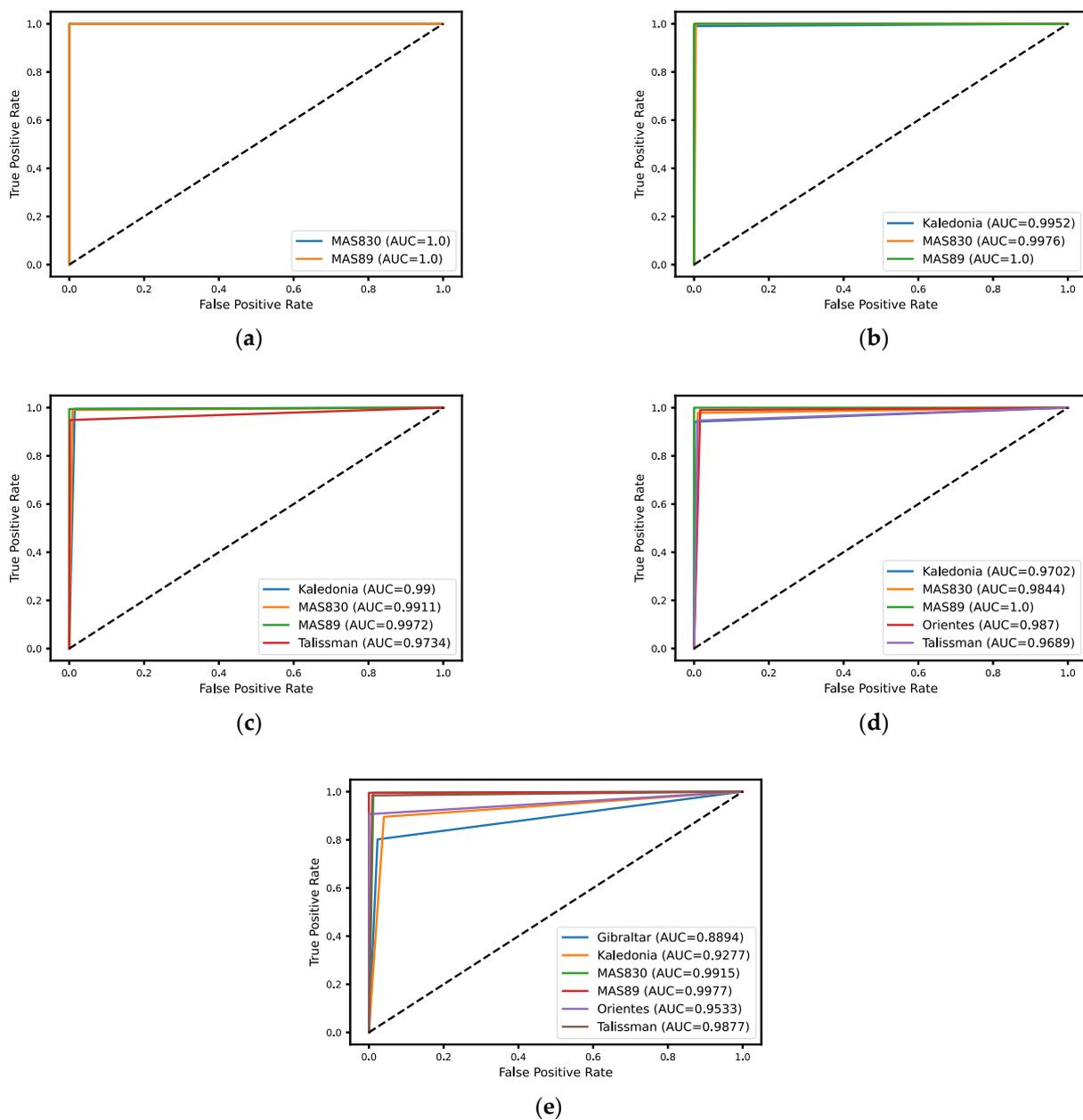
Classes	Varieties	Recall	Precision	F1-Score	AUC	Accuracy	Loss
2	MAS830	1	1	1	1	1	0.00005
	MAS89	1	1	1	1		
3	Kaledonia	0.9903	1	0.9951	0.9952	0.992	0.02415
	MAS830	1	0.9896	0.9948	0.9976		
	MAS89	1	1	1	1		
4	Kaledonia	0.9948	0.9548	0.9744	0.9900	0.972	0.07590
	MAS830	0.9907	0.9770	0.9838	0.9911		
	MAS89	0.9945	1	0.9972	0.9972		
	Talissman	0.9486	0.9951	0.9713	0.9734		
5	Kaledonia	0.9417	0.9949	0.9676	0.9702	0.940	0.18497
	MAS830	0.9786	0.9581	0.9683	0.9844		
	MAS89	1	1	1	1		
	Orientes	0.9904	0.9406	0.9649	0.9870		
	Talissman	0.9466	0.9653	0.9559	0.9689		
6	Gibraltar	0.8019	0.8783	0.8384	0.8894	0.895	0.29350
	Kaledonia	0.8955	0.8182	0.8551	0.9277		
	MAS830	0.9951	0.9444	0.9691	0.9915		
	MAS89	0.9953	1	0.9977	0.9977		
	Orientes	0.9066	1	0.9510	0.9533		
	Talissman	0.9843	0.9543	0.9691	0.9877		

Besides the results shown in Table 3, the normalized confusion matrix for the different experiments was computed, to obtain visual information on the test results, as shown in Figure 6. In this figure: (i) the horizontal axis is the predicted results; (ii) the vertical axis is the true values per class; (iii) the diagonal shows the classification accuracy per class of the model; and (iv) the other cells represent the classification errors per class.



**Figure 6.** Normalized confusion matrix for the different experiments: (a) two classes (MAS830 and MAS89); (b) three classes (Kaledonia, MAS830, and MAS89); (c) four classes (Kaledonia, MAS830, MAS89, and Talissman); (d) five classes (Kaledonia, MAS830, MAS89, Orientes, and Talissman); (e) six classes (Gibraltar, Kaledonia, MAS830, MAS89, Orientes, and Talissman).

Figure 7 shows ROC curves and the AUC for each variety in each experiment. The horizontal axis shows the false positive rate, and the vertical axis is the true positive rate.



**Figure 7.** The ROC curves and AUC scores of the models with the highest accuracies for the different experiments: (a) two classes (MAS830 and MAS89); (b) three classes (Kaledonia, MAS830, and MAS89); (c) four classes (Kaledonia, MAS830, MAS89, and Talissman); (d) five classes (Kaledonia, MAS830, MAS89, Orientes, and Talissman); (e) six classes (Gibraltar, Kaledonia, MAS830, MAS89, Orientes, and Talissman).

According to the results shown in Table 3, Figures 6 and 7, it can be seen that (i) the MAS89 variety was the best recognized seed; (ii) the Gibraltar variety was the worst recognized seed; and (iii) the accuracy decreases with increasing number of classes, from 100% for two classes to 89.5% for six classes.

#### 4. Discussion

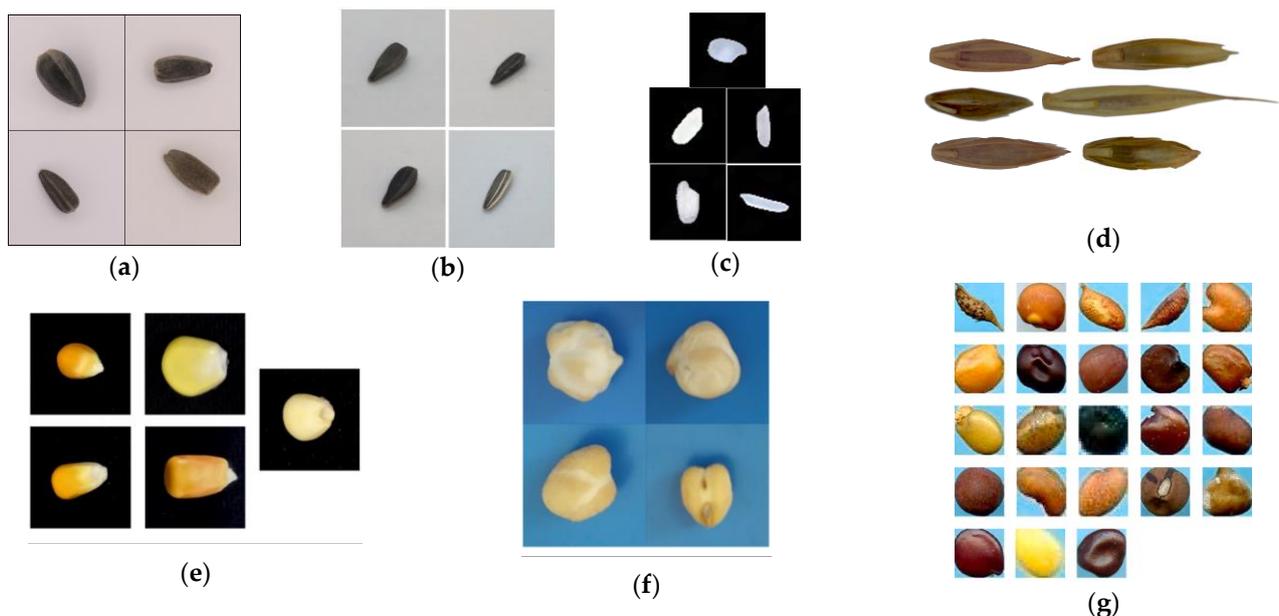
In our study, a seed classification accuracy of 89.5% was achieved in the classification with DL of six different varieties of high oleic sunflower seeds, Talissman, MAS830, MAS89, Kaledonia, Gibraltar, and Orientes, when the training, validation, and testing was performed with 1000 images of each variety, and the resolution of each image was

227 × 227 pixels. When less than six varieties were used for classification, accuracies reached values always higher than 94%, due to the Gibraltar variety being more difficult to distinguish from the rest of the seeds.

The accuracy of DL classification of varieties of seeds from plants such as sunflowers [45,46], chickpeas [34], maize [33], rice [43], grass [32], and mixed plants [48] has been evaluated in the literature.

There are, to the best of the authors' knowledge, only two studies in the literature on sunflower seed classification with DL [45,46].

In the first one, Kurtulmuş identified four sunflower seed varieties (Armada, Reina, Sirena, and Palanci) as shown in Figure 8b [45]. In that study, an approximate total of 1200 seeds per variety (a total of 4800 seed samples) were imaged. Different CNN architectures were trained and tested using three models (AlexNet, GoogleNet, and ResNet) and two learning algorithms (AdaDelta and SGD). In their experiments, there was not a significant difference between training and testing accuracies. All the models could obtain classification accuracies over 90%. The GoogleNet model yielded the highest accuracy, with a value of 95%, in distinguishing the sunflower seeds. AlexNet with AdaDelta, and AlexNet with SGD, also provided relatively good classification performances (accuracy of 93%, AUC = 0.96), requiring notably short training times (a little less than one hour) than the others (more than two hours). Moreover, no particular solver (AdaDelta and SGD) was more prominent in performance.



**Figure 8.** Images of varieties of seed classified in the literature. (a) Four high oleic sunflower seed varieties classified in our study. (b) Four sunflower seed varieties classified by Kurtulmuş [45]. (c) Five rice varieties classified by Koklu et al. [43]. (d) Six grass varieties classified by Eryigit and Tugrul [32]. (e) Five maize varieties classified by Xu et al. [33]. (f) Four chickpea varieties classified by Taheri-Garavand [34]. (g) Twenty-three seeds classified by Loddo et al. [48]. All figures have been reproduced from their original papers with previous consent from the copyright holder.

Our study, and the study by Kurtulmuş, reached very close high performance rates. In our study, the accuracy score achieved in discriminating between four varieties of seeds was 97.2% (AUC = 0.99%) and in the Kurtulmuş study, when using AlexNet models, was 93% (AUC = 0.96). There are some differences in the methodologies that can explain the small differences. Our optimizer, or solver, was Adam, in contrast with the solvers used by Kurtulmuş, which were Ada Delta and SGD. Another difference is that the four seed varieties of both studies are different, making the comparison between studies more difficult. Nevertheless, in both studies, the performance metrics observed from the experiments

were promising, with the fact that visually similar sunflower varieties could be identified by the methods proposed. Overall, both studies showed that the identification of sunflower seeds was feasible using deep learning techniques.

Figure 8a shows the sunflower seed varieties classified in our study and Figure 8b shows the sunflower seed varieties classified in the study of Kurtulmuş. As can be seen in both figures, the seeds of both studies look similar, and it is therefore coherent that both studies reached similar accuracies.

In the second study in the literature on sunflower seed classification, Li et al. conducted comparative experiments to compare the performance of the SeedSortNet network with six representative CNN models, that included AlexNet. They classified sunflower and maize seeds, though for the detection of abnormal seed images composed of leaves, stones, defective seeds, and normal sunflower seeds [46]. Experimental results show that AlexNet achieved accuracy rates of 99% and SeedSortNet of 99.56%, on sunflower seed dataset classification. Nevertheless, their study is not comparable with our study, because abnormal seed images are not sunflower seeds and therefore are easier to classify.

Some studies applying DL methods for classifying seeds reported accuracy values higher than the values achieved in our study. Kotlu et al. reported 100% accuracy in the classification of five rice varieties [43], Eryigit and Tugrul reported 99.4% accuracy for six grass varieties [32], and Xu et al. reported values between 96.4% and 99.7% for five maize varieties [33]. These higher accuracies could be due to a greater differentiation between the classes to be classified in their studies. That reasoning is coherent, because the images of the seeds in Figure 8c,d,f, visually differ in shape, size, color, and/or texture more than the images of the seeds shown in Figure 8a.

Other studies in the DL classification of seeds reported accuracies similar to the values found in our study, but lower than the aforementioned articles. Concretely, Taheri-Garavand et al. reported a classification accuracy of 94.2% for five chickpea varieties [34], and Loddo et al. reported accuracy values between 93.4% and 97.4% for 23 different varieties of seeds [48]. Those relatively lower accuracy values could be due to two reasons. The first one could be due to a smaller differentiation between the classes to be classified in the studies. That reasoning is coherent, because the images of the seeds shown in Figure 8f are more similar in shape, size, color, and texture than the images of the seeds shown in Figure 8a. The second one could be the higher number of seed varieties to be classified. That reasoning is also coherent, because twenty-three different seed classes are presented in Figure 8g, whereas only four are presented in Figure 8a.

One of the limitations of our study is the relatively small number of seed varieties employed, including only six sunflower varieties. Despite other studies also using only four varieties, like Kurtulmuş [45], there is room for improvement by incorporating additional varieties. In fact, the varieties used by Kurtulmuş were all different from the ones used by our study, and other existing varieties could be used too.

Regarding the main strength of our study, we considered seed varieties very similar in appearance, which made the differentiation of the classes more difficult. Moreover, the performance comparison of the behavior of the classifier when changing the number of classes showed the difficulties when adding new varieties, due to their similarity to one of the previously existing classes.

The classifier developed in the present study can be applied in industrial applications to rapidly analyze seed samples, since it uses a low-cost image acquisition system and is a non-destructive method. As sunflower varieties can differ in oil quality, yield, and by-products, the classifier can ensure that sunflower seeds meet the industrial requirements.

## 5. Conclusions

In summary, DL algorithms can be used to classify high oleic sunflower seeds with high accuracy. The accuracy values of classification decreased when adding more classes, varying from 100% to 89.5%. The largest decrease was observed with six classes, where the similarities within varieties increased, specially between the varieties Gibraltar and

Kaledonia. Future studies could increase classification accuracies with different approaches, to advance this research. Varying specific parameters of the CNN model, such as the number of varieties for classification and the size of the dataset, could be tested. Future studies could also employ multispectral or hyperspectral images of high oleic sunflower seeds.

**Author Contributions:** Conceptualization, J.M.A.-P. and J.G.-G.; methodology, J.M.A.-P. and J.G.-G.; software, M.B.-C.; validation, R.R.-G.; formal analysis, M.B.-C. and M.A.Z.; original draft preparation, M.B.-C. and M.A.Z.; writing—review and editing, J.G.-G., J.M.A.-P. and R.R.-G.; visualization, M.A.Z.; supervision, J.M.A.-P., R.R.-G. and J.G.-G. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research received no external funding.

**Institutional Review Board Statement:** Not applicable.

**Informed Consent Statement:** Not applicable.

**Data Availability Statement:** Not applicable.

**Acknowledgments:** The authors of this article wish to thank (i) José Alberto Gutierrez Vesga for his help constructing the device shown in Figure 3, (ii) Miguel Gutiérrez Hernández, from the Sovena Oilseeds España (Brenes, Spain) company, for the donated sunflower seed samples employed in our study, and (iii) Coordenação de Aperfeiçoamento de Pessoal de Nível Superior–Brazil (CAPES)–Finance Code 001 for the scholarship of the second author.

**Conflicts of Interest:** The authors declare no conflict of interest.

## References

1. Chávez-Dulanto, P.N.; Thiry, A.A.A.; Glorio-Paulet, P.; Vögler, O.; Carvalho, F.P. Increasing the Impact of Science and Technology to Provide More People with Healthier and Safer Food. *Food Energy Secur.* **2021**, *10*, 1–31. [CrossRef]
2. Vollset, S.E.; Goren, E.; Yuan, C.W.; Cao, J.; Smith, A.E.; Hsiao, T.; Bisignano, C.; Azhar, G.S.; Castro, E.; Chalek, J.; et al. Fertility, Mortality, Migration, and Population Scenarios for 195 Countries and Territories from 2017 to 2100: A Forecasting Analysis for the Global Burden of Disease Study. *Lancet* **2020**, *396*, 1285–1306. [CrossRef]
3. USDA Oilseeds: World Markets and Trade. Available online: <https://downloads.usda.library.cornell.edu/usda-esmis/files/tx31qh68h/k06999499/gm80jz36g/oilseeds.pdf> (accessed on 8 February 2022).
4. Zeb, A. A Comprehensive Review on Different Classes of Polyphenolic Compounds Present in Edible Oils. *Food Res. Int.* **2021**, *143*, 110312. [CrossRef]
5. Gupta, M.K. Sunflower Oil: History, Applications and Trends. *Lipid Technol.* **2014**, *26*, 260–263. [CrossRef]
6. Bowen, K.J.; Kris-Etherton, P.M.; West, S.G.; Fleming, J.A.; Connelly, P.W.; Lamarche, B.; Couture, P.; Jenkins, D.J.A.; Taylor, C.G.; Zahradka, P.; et al. Diets Enriched with Conventional or High-Oleic Acid Canola Oils Lower Atherogenic Lipids and Lipoproteins Compared to a Diet with a Western Fatty Acid Profile in Adults with Central Adiposity. *J. Nutr.* **2019**, *149*, 471–478. [CrossRef]
7. Raatz, S.K.; Conrad, Z.; Jahns, L.; Belury, M.A.; Picklo, M.J. Modeled Replacement of Traditional Soybean and Canola Oil with High-Oleic Varieties Increases Monounsaturated Fatty Acid and Reduces Both Saturated Fatty Acid and Polyunsaturated Fatty Acid Intake in the US Adult Population. *Am. J. Clin. Nutr.* **2018**, *108*, 594–602. [CrossRef]
8. Downs, S.M.; Bloem, M.Z.; Zheng, M.; Catterall, E.; Thomas, B.; Veerman, L.; Wu, J.H.Y. The Impact of Policies to Reduce Trans Fat Consumption: A Systematic Review of the Evidence. *Curr. Dev. Nutr.* **2017**, *1*, 1–10. [CrossRef]
9. Zambelli, A. Current Status of High Oleic Seed Oils in Food Processing. *JAACS* **2021**, *98*, 129–137. [CrossRef]
10. Petraru, A.; Ursachi, F.; Amariei, S. Nutritional Characteristics Assessment of Sunflower Seeds, Oil and Cake. Perspective of Using Sunflower Oilcakes as a Functional Ingredient. *Plants* **2021**, *10*, 2487. [CrossRef]
11. Arrutia, F.; Binner, E.; Williams, P.; Waldron, K.W. Oilseeds beyond Oil: Press Cakes and Meals Supplying Global Protein Requirements. *Trends Food Sci. Technol.* **2020**, *100*, 88–102. [CrossRef]
12. De Moraes Oliveira, V.R.; de Arruda, A.M.V.; Silva, L.N.S.; de Souza, J.B.F.; de Queiroz, J.P.A.F.; da Silva Melo, A.; Holanda, J.S. Sunflower Meal as a Nutritional and Economically Viable Substitute for Soybean Meal in Diets for Free-Range Laying Hens. *Anim. Feed. Sci. Technol.* **2016**, *220*, 103–108. [CrossRef]
13. Adeleke, B.S.; Babalola, O.O. Oilseed Crop Sunflower (*Helianthus Annuus*) as a Source of Food: Nutritional and Health Benefits. *Food Sci. Nutr.* **2020**, *8*, 4666–4684. [CrossRef] [PubMed]
14. Anastopoulos, I.; Ighalo, J.O.; Adaobi Igwegbe, C.; Giannakoudakis, D.A.; Triantafyllidis, K.S.; Pashalidis, I.; Kalderis, D. Sunflower-Biomass Derived Adsorbents for Toxic/Heavy Metals Removal from (Waste) Water. *J. Mol. Liq.* **2021**, *342*, 117540. [CrossRef]

15. Elkelawy, M.; Bastawissi, H.A.E.; Esmail, K.K.; Radwan, A.M.; Panchal, H.; Sadasivuni, K.K.; Suresh, M.; Israr, M. Maximization of Biodiesel Production from Sunflower and Soybean Oils and Prediction of Diesel Engine Performance and Emission Characteristics through Response Surface Methodology. *Fuel* **2020**, *266*, 117072. [[CrossRef](#)]
16. Jadhav, J.V.; Anbu, P.; Yadav, S.; Pratap, A.P.; Kale, S.B. Sunflower Acid Oil-Based Production of Rhamnolipid Using *Pseudomonas Aeruginosa* and Its Application in Liquid Detergents. *J. Surfactants Deterg.* **2019**, *22*, 463–476. [[CrossRef](#)]
17. Rauf, S.; Jamil, N.; Tariq, S.A.; Khan, M.; Kausar, M.; Kaya, Y. Progress in Modification of Sunflower Oil to Expand Its Industrial Value. *J. Sci. Food Agric.* **2017**, *97*, 1997–2006. [[CrossRef](#)]
18. Alberio, C.; Izquierdo, N.G.; Galella, T.; Zuil, S.; Reid, R.; Zambelli, A.; Aguirrezábal, L.A.N. A New Sunflower High Oleic Mutation Confers Stable Oil Grain Fatty Acid Composition across Environments. *Eur. J. Agron.* **2016**, *73*, 25–33. [[CrossRef](#)]
19. Attia, Z.; Pogoda, C.S.; Reinert, S.; Kane, N.C.; Hulke, B.S. Breeding for Sustainable Oilseed Crop Yield and Quality in a Changing Climate. *Theor. Appl. Genet.* **2021**, *134*, 1817–1827. [[CrossRef](#)]
20. Diovisalvi, N.; Calvo, N.R.; Izquierdo, N.; Echeverría, H.; Divito, G.A.; García, F. Effects of Genotype and Nitrogen Availability on Grain Yield and Quality in Sunflower. *Agron. J.* **2018**, *110*, 1532–1543. [[CrossRef](#)]
21. Krig, S. *Computer Vision Metrics: Survey, Taxonomy, and Analysis*; Springer: Berlin/Heidelberg, Germany, 2014; Volume 9781430259305, ISBN 9781430259305.
22. Velesaca, H.O.; Suárez, P.L.; Mira, R.; Sappa, A.D. Computer Vision Based Food Grain Classification: A Comprehensive Survey. *Comput. Electron. Agric.* **2021**, *187*, 106287. [[CrossRef](#)]
23. Sabanci, K.; Kayabasi, A.; Toktas, A. Computer Vision-Based Method for Classification of Wheat Grains Using Artificial Neural Network. *J. Sci. Food Agric.* **2017**, *97*, 2588–2593. [[CrossRef](#)]
24. Koklu, M.; Ozkan, I.A. Multiclass Classification of Dry Beans Using Computer Vision and Machine Learning Techniques. *Comput. Electron. Agric.* **2020**, *174*, 105507. [[CrossRef](#)]
25. Khatri, A.; Agrawal, S.; Chatterjee, J.M. Wheat Seed Classification: Utilizing Ensemble Machine Learning Approach. *Sci. Program.* **2022**, *2022*, 1–9. [[CrossRef](#)]
26. Shrestha, B.L.; Kang, Y.M.; Yu, D.; Baik, O.D. A Two-Camera Machine Vision Approach to Separating and Identifying Laboratory Sprouted Wheat Kernels. *Biosyst. Eng.* **2016**, *147*, 265–273. [[CrossRef](#)]
27. Gulzar, Y.; Hamid, Y.; Soomro, A.B.; Alwan, A.A.; Journaux, L. A Convolution Neural Network-Based Seed Classification System. *Symmetry* **2020**, *12*, 2018. [[CrossRef](#)]
28. Guo, Y.; Liu, Y.; Oerlemans, A.; Lao, S.; Wu, S.; Lew, M.S. Deep Learning for Visual Understanding: A Review. *Neurocomputing* **2016**, *187*, 27–48. [[CrossRef](#)]
29. Altuntaş, Y.; Cömert, Z.; Kocamaz, A.F. Identification of Haploid and Diploid Maize Seeds Using Convolutional Neural Networks and a Transfer Learning Approach. *Comput. Electron. Agric.* **2019**, *163*, 104874. [[CrossRef](#)]
30. Kozłowski, M.; Górecki, P.; Szczypiński, P.M. Varietal Classification of Barley by Convolutional Neural Networks. *Biosyst. Eng.* **2019**, *184*, 155–165. [[CrossRef](#)]
31. Laabassi, K.; Belarbi, M.A.; Mahmoudi, S.; Mahmoudi, S.A.; Ferhat, K. Wheat Varieties Identification Based on a Deep Learning Approach. *J. Saudi Soc. Agric. Sci.* **2021**, *20*, 281–289. [[CrossRef](#)]
32. Eryigit, R.; Tugrul, B. Performance of Various Deep-Learning Networks in the Seed. *Symmetry* **2021**, *13*, 1892. [[CrossRef](#)]
33. Xu, P.; Tan, Q.; Zhang, Y.; Zha, X.; Yang, S.; Yang, R. Research on Maize Seed Classification and Recognition Based on Machine Vision and Deep Learning. *Agriculture* **2022**, *12*, 232. [[CrossRef](#)]
34. Taheri-Garavand, A.; Nasiri, A.; Fanourakis, D.; Fatahi, S.; Omid, M.; Nikoloudakis, N. Automated in Situ Seed Variety Identification via Deep Learning: A Case Study in Chickpea. *Plants* **2021**, *10*, 1406. [[CrossRef](#)]
35. Zhao, G.; Quan, L.; Li, H.; Feng, H.; Li, S.; Zhang, S.; Liu, R. Real-Time Recognition System of Soybean Seed Full-Surface Defects Based on Deep Learning. *Comput. Electron. Agric.* **2021**, *187*, 106230. [[CrossRef](#)]
36. Qiu, Z.; Chen, J.; Zhao, Y.; Zhu, S.; He, Y.; Zhang, C. Variety Identification of Single Rice Seed Using Hyperspectral Imaging Combined with Convolutional Neural Network. *Appl. Sci.* **2018**, *8*, 212. [[CrossRef](#)]
37. Nie, P.; Zhang, J.; Feng, X.; Yu, C.; He, Y. Classification of Hybrid Seeds Using Near-Infrared Hyperspectral Imaging Technology Combined with Deep Learning. *Sens. Actuators B Chem.* **2019**, *296*, 126630. [[CrossRef](#)]
38. Loddo, A.; Loddo, M.; di Ruberto, C. A Novel Deep Learning Based Approach for Seed Image Classification and Retrieval. *Comput. Electron. Agric.* **2021**, *187*, 106269. [[CrossRef](#)]
39. Gao, J.; Liu, C.; Han, J.; Lu, Q.; Wang, H.; Zhang, J.; Bai, X.; Luo, J. Identification Method of Wheat Cultivars by Using a Convolutional Neural Network Combined with Images of Multiple Growth Periods of Wheat. *Symmetry* **2021**, *13*, 2012. [[CrossRef](#)]
40. Przybyło, J.; Jabłoński, M. Using Deep Convolutional Neural Network for Oak Acorn Viability Recognition Based on Color Images of Their Sections. *Comput. Electron. Agric.* **2019**, *156*, 490–499. [[CrossRef](#)]
41. Kamilaris, A.; Prenafeta-Boldú, F.X. Deep Learning in Agriculture: A Survey. *Comput. Electron. Agric.* **2018**, *147*, 70–90. [[CrossRef](#)]
42. Lin, P.; Li, X.L.; Chen, Y.M.; He, Y. A Deep Convolutional Neural Network Architecture for Boosting Image Discrimination Accuracy of Rice Species. *Food Bioproc. Tech.* **2018**, *11*, 765–773. [[CrossRef](#)]
43. Koklu, M.; Cinar, I.; Taspınar, Y.S. Classification of Rice Varieties with Deep Learning Methods. *Comput. Electron. Agric.* **2021**, *187*, 106285. [[CrossRef](#)]

44. Mukasa, P.; Wakholi, C.; Akbar Faqeerzada, M.; Amanah, H.Z.; Kim, H.; Joshi, R.; Suh, H.-K.; Kim, G.; Lee, H.; Kim, M.S.; et al. Nondestructive Discrimination of Seedless from Seeded Watermelon Seeds by Using Multivariate and Deep Learning Image Analysis. *Comput. Electron. Agric.* **2022**, *194*, 106799. [[CrossRef](#)]
45. Kurtulmuş, F. Identification of Sunflower Seeds with Deep Convolutional Neural Networks. *J. Food Meas. Charact.* **2021**, *15*, 1024–1033. [[CrossRef](#)]
46. Li, C.; Li, H.; Liu, Z.; Li, B.; Huang, Y. SeedSortNet: A Rapid and Highly Efficient Lightweight CNN Based on Visual Attention for Seed Sorting. *PeerJ. Comput. Sci.* **2021**, *7*, e639. [[CrossRef](#)]
47. Krizhevsky, A.; Sutskever, I.; Hinton, G.E. ImageNet Classification with Deep Convolutional Neural Networks. *Commun. ACM* **2017**, *60*, 84–90. [[CrossRef](#)]
48. Loddo, A.; di Ruberto, C. On the Efficacy of Handcrafted and Deep Features for Seed Image Classification. *J. Imaging* **2021**, *7*, 171. [[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.