This document is a manuscript of:
Michał Kozłowski and Piotr M. Szczypiński.
"Barley defects identification by convolutional neural networks."
***International Conference on Information Technologies in Biomedicine*.**
Springer, Cham, 2019.

https://doi.org/10.1007/978-3-030-23762-2_17
https://itib.polsl.pl/

Barley defects identification by convolutional neural networks

Michał Kozłowski[1] and Piotr M. Szczypiński[2]

Abstract

The appropriate choice of ingredients, particularly barley, is a key issue in the malting and brewing industry. Nowadays, controlling barley quality involves visual inspection to identify defective or infected kernels. It requires evaluation and is labour-intensive. Computer vision solutions sequentially applying attribute extraction and classification algorithms tend to be inaccurate. Deep learning networks combine the two aspects together to enable their mutual adjustment and to increase classification ability. We use this technique to identify the most common defects of malting barley. Two ways of data presentation, two implementations of convolutional neural networks and a handcrafted-features-based method, are examined. The classification results are presented, compared and discussed.

Introduction

Malting is a process of partial germination of grains, which causes to development of carbohydrates. It is suspended by drying when a maximum amount of sugar is gained. Grains should be intact and contain the germ, which is necessary for the germination. The existence of kernels with fungal infections is unaccepted able due to the possible contamination of the final product with the toxins and, in case of beer, an extensive gushing. The utilization of low-quality cereal for malting results in lower quality of the final product and in economic losses. Therefore, assuring superior quality of barley grains for the malting process is crucial. When a new supply of cereal, usually several tons, is acquired by a malting house, its utility for malting is assessed from a sample of 100g. The evaluation of barley is typically carried out by an expert, who visually identifies kernels with spots of fungal infection, the grains which are already sprouted, undeveloped or without the germ. This kind of visual evaluation is a tedious and time-consuming task, while its reliability depends on the skills, observation, and experience of the evaluator. The inferior quality is indicated by the existence of infected kernels, contribution of foreign matter, undeveloped or mechanically damaged kernels. The details of the evaluation procedure and acceptable amounts of defective kernel fractions are specified by the Polish industry standard BN-87/9131-13.

1 Department of Mechatronics and Technical and IT Education, Faculty of Technical Science, University of Warmia and Mazury, Słoneczna 46A, 10-710 Olsztyn, Poland, michal.kozlowski@uwm.edu.pl
2 Institute of Electronics, Łodź University of Technology, Wólczańska 211/215, 90-924 Łodź, Poland, pms@p.lodz.pl

The attempts to apply computer vision and machine learning techniques to automatize the evaluation procedure were already made. The [13] introduced an ontology, which formalizes the human expert knowledge, and an expert system based on the ontology to classify barley kernels. The system extracts selected visual attributes by applying computer vision methods. However, the assessment of more complex or abstract attributes requires human involvement – therefore, the solution is not fully automatic. In [11] the algorithm is presented, which in sequence applies image processing, segmentation, morphology analysis, extraction of texture and colour attributes, supervised learning and classification. The procedures were validated on a small dataset and enabled classification of defective kernels with the accuracy of 91-97%. The deep learning networks have gained a huge interest recently. They join the attribute extraction (in convolution layers), the feature selection (max pooling) and the classification (in the fully connected layer) in one body. Thus, the machine learning procedure enables joint and mutual optimization of these components, which was not feasible in traditional approaches. Therefore, we expect that the deep learning networks can gain higher accuracy ratios in the detection of defective barley kernels than the traditional algorithms.

In this paper, a machine vision approach to identify barley defects using convolutional neural networks is presented. A fully trained and size-tailored model of a network is compared with the pretrained model commonly used in the image classification tasks. For reference, we apply the hand-engineered-features-based method presented in [11]. Moreover, we present two approaches to the image data presentation. In one of them, the single side of the kernel is analysed. In the other, both sides, the ventral and the dorsal, are jointly presented to the networks. The classification results obtained by each of the examined solutions are presented, compared and discussed.

Related works

In the literature, there are several approaches to the identification of the grains using image analysis. The researches usually use hand-engineered methods to identify and extract the significant features from the images. Such features are the statistics or the model parameters used to characterize colour, texture, or shape of the cereal kernels. Most solutions use machine learning for selection of the most important discriminative features and at the final stage of the classification.

Dollawat Ngampak et al. [7] explored the possibilities of classifying images of broken rice grains using least-square support vector machine (LS-SVM) with radial basis function (RBF) kernel. In [5] the greyscale histogram, statistical features have distinguished eight classes of wheat varieties. The classification of four types of cereal grains was performed with use of the morphological and colour features in [6].

Fig. 1: Example images: a) broken, b) infected or sprouted, c) with missing germ, d)green or undeveloped and e) normal

There were several approaches to use perceptron neural network for classification. An input for the classifiers were feature vectors extracted from the cereal images. An equalized colour component extraction, computation of morphological features, and the application of multilayer perceptron neural network were used to analyse bulk samples of three barley varieties of [9]. In [14] four different architectures of neural classifiers were examined, namely back propagation network (BPN), Ward network, general regression neural network (GRNN) and probabilistic neural network (PNN). The quantitative characteristics of shape, colour, and texture computed for every kernel served as an input to the classier, capable of recognizing grain types, varieties, defective kernels or foreign objects. It enabled detection of defective kernels in [15]. All the above methods assumed that the features are extracted from images presenting only one side of the kernel. In [12] it was presented that classification accuracy rises if two sides of the kernel, the dorsal and the ventral ones, are analysed concurrently. Recently, an innovative automatic acquisition system has appeared [4, 8], which records images of both sides of kernels. It enabled development of a novel workflow for classification of barley, in which features extracted from both the images are combined and jointly presented to the classifier [11].

Recently, convolutional neural networks (CNNs) have made an impact on many vision-based problems. CNNs are capable of learning meaningful image features directly from data and in some applications, they significantly outer from traditional computer vision methods [2]. The CNN applied for classification of 8 barley varieties was presented in [1]. Two separate convolution layers analyse the images of dorsal and ventral sides, respectively. The information is merged at the stage of fully connected layers. The network was trained on the relatively small set of 200-500 cases per class and gained the classification accuracy of 97%.

Following the above works, we examine CNN application in detection and classification of defective barley kernels. To our knowledge, this issue has not been the subject of scientific publication yet. We compare two different architectures of CNNs and verify whether the analysis of single or both sides of kernels has a significant impact on the classification performance

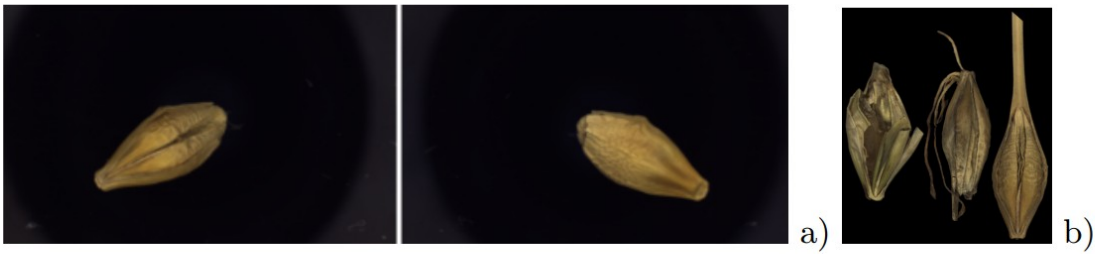

Fig. 2: a) Input images presenting two opposite sides of kernel, b) outliers

Materials

Barley grains were obtained from selected farms in Poland. The experimental material consisted of the images of four classes of barley defects and one reference class of healthy grains (Fig. 1). The classes of defects included kernels which were a) broken, b) infected or sprouted, c) without a germ, d) green or not fully developed. The initial classification of all grains was carried out by an expert of Słodownia Soufflet Polska Sp. z o.o. malt house.

The grains were photographed by a two-camera acquisition system [4]. The device drops kernels on the flat transparent surface and enables image acquisition of their both sides. However, the dorsoventral orientation of each kernel is random, and the system does not determine which of the cameras takes an image of which side. An additional problem is a random orientation of anteroposterior axis, as kernels can rotate on the flat surface (Fig. 2 a). The data set is based on biological material, which makes it highly diversified within classes and at the same time kernels belonging to different classes may look similar. The exception is the class of broken grains, which includes grain pieces highly distinctive with the morphological attributes.

Another issue is numerous difficult or unusual cases which significantly differ from other kernels (Fig. 2 b). These are kernels which were deformed by other than mechanical causes, sprouted kernels, kernels with awns, which in most other cases are detached, and kernels covered with a husk sticking out. All these cases highly differ with their attributes from the other kernels and become outliers, which may cause problems in machine learning. Nevertheless, the outliers were deliberately kept present in the training and the test sets. This enabled assessment of the algorithm's immunity to the difficult cases. It also makes the experiment conditions closer to the real life situation. Moreover, if the object does not fit within the image frame, the algorithm analyses its visible part only.

The entire dataset of images was randomly divided into training, validation and test sets, containing 80%, 10% and 10% of samples respectively. The detailed number of cases used in each of the three sets and in every class is listed in Table 1.

Tab. 1: Dataset

		Total 100%	Training 80%	Validation 10%	Test 10%
ID	Class name	29714	23772	2972	2972
1	broken	6046	4836	604	604
2	infected	8808	7046	880	880
3	missing germ	5258	4206	526	526
4	green	5260	4208	526	526
5	normal	4342	3474	434	434

Methods

Preprocessing

Since the brightness of the kernel is higher than the brightness of the background, the initial step in image processing is greyscale thresholding to find the region of interest. Next, the binary image is median-filtered to smooth the contour of the region. The connected set of the highest intensity area is selected as a mask of a kernel. Next, the contour of every mask is approximated by an ellipse. The longer diameter of the ellipse designates the kernel main axis. Then, the width of the kernel is estimated along the axis to determine the germ-brush orientation. This information is used to correct the anteroposterior orientation of the kernel image to set the germ side upward [11]. Moreover, the background of the original image outside the mask is replaced with a uniform black colour.

The images are then prepared to fit to the inputs of CNNs. If a single side of the kernel is analysed, the resolution of a single image is reduced to fit in the 80×170 (the proposed CNNs configuration) or in the 227×227 (pretrained AlexNet model) pixel window. If the goal is to analyse pairs of the images show ing the opposite sides of a single kernel, the images of both sides are combined in a single frame. One of the images is located on the left-hand side and the other on the right-hand side, next to each other. The joint image is resized to fit in the 170x170 or in the

227x227 pixel window for respective CNNs. The Fig. 3 presents an average of all the images belonging to the training set, preprocessed in the way explained above. The image's average is used during training to properly bias neurons belonging to the input layer.

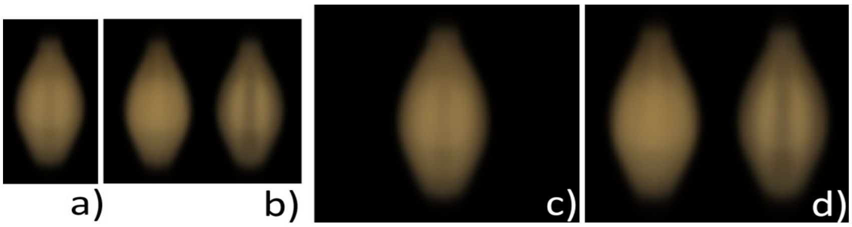

Fig. 3: Averages of training samples: a) 80x170 (CNN), b) 170x170 (CNN), c) and d) 227x227 (AlexNet)

Convolutional neural network

In the last decade, CNNs have become state-of-the-art tools for many images classification and recognition tasks. CNNs are made from many layers that come arranged one after another. During the classification, the entire layer is displayed on the input layer. As opposed to multi-layer perceptron (MLP) neural networks which are usually composed of 3 fully connected layers. CNNs usually consist of more layers that have shaped the concept of deep networks.

Our experiments were based on a simplified architecture that was adapted to the problem under study. The known AlexNet model was used as the reference method [3]. This model was trained using 1.5 million natural images and won the ImageNet Large Scale Visual Recognition Competition (ILSVRC) 2010 and 2012). For this reason this architecture has become state of art models. AlexNet contains eight layers, the first five are convolutional layers (11x11, 5x5, 3x3), and the last three are fully connected (2x2048 and 1000 neurons). It implements a non-saturating rectified-linear layer (RELU) the activation function, which outperforms hyperbolic tangent or sigmoid functions.

In addition to AlexNet, there are more models that are much more complex and achieve better classification results (when classifying objects for thousands of classes). It was noticed that the smaller 3×3 filters in the convolutional layers are better at distinguishing the important features from the image. This information was confirmed using the VGG [10] and ResNet [2] models. By increasing the depth of the model (the number of convolutional layers), the classification results were improved at the expense of the speed of operation dictated by a large number of calculations.

The motivation of our work is to find the optimal solution between the speed of model operation and high classification. This is dictated by the practical application in the malt house, where the speed of classification is important. The quoted models were designed for classification tasks based on natural images in a thousand classes, as evidenced by the number of exits of the last layer of FC (Fully-Connected Layer) in AlexNet. Our problem is different. The image acquisition conditions are repeatable and the number of classes is limited to five. We suggest that basing on proven solutions create a model suited to the classification task.

A deep learning framework, Caffe was used to build a new architecture suited to the problem being studied. It contains 2 convolutional layers and 2 fully connected layers. The input layer is adapted to the objects of interest and has a resolution of 80×170 (one grain) or 170×170 (two grains in one image). This reduced the size of the final model and influenced on the speed of the learning process. The number of convolutional layers (CONV) was also limited. The first consisting of 64 filters with a size of 3×3 and the second with 128 filters of the same size. We assume that the simplification of the convolutional layers will reduce the computational requirements, and the important features identified will be sufficient for proper classification. After each CONV layer, the RELU activation function was used, which speeds up the training process and maintains the same level of the accuracy [3]. The amount of data from CONV or RELU layers are large. Further spatial downsampling of information is performed by the max pooling (POOL) layer. The max operator is applied across the local neighbourhood of the previous layer outputs, with predefined strides. To prevent overfitting, a dropout was used before the first fully connected (FC) layer. It is worth noting that the dropout of 0.5 extends the training time. We used two FC layers. The first is 1024 neurons and the number of neurons in the second one is equal to 5, which is the number of classes.

Two approaches were compared (Fig. 4). The difference between them is how to enter data on the first layer of a convolutional neural network. In the first approach for the CNN model, the images are read one at a time in 80×170 resolution. Combining a pair of images on both sides of the grain increases the resolution of the input image to 170×170. Higher resolution images (227×227) were used in the reference method, which was the previously trained AlexNet model.

Caffe contains a so-called Zoo model that allows the use of a model trained on a large dataset like ImageNet. This technique is called transfer learning and uses the experience gained in a very advanced classification task to shorten significantly the training time and usually may improve the classification results. In order to adopt such a model to the new problem, it is enough to modify the FC layer by changing the number of outputs to match the number of classes.

In the AlexNet model that we use, the number of FC layer outputs was reduced to 5. This model revealed excessive overfitting characteristics during the initial tests. This is due to the large capacity of the AlexNet model, which is able to remember all the samples. Instead of distinguishing features – they are taught by heart. To deal with this problem, the number of neurons in the first two FC layers was halved, which brought the intended effect. Finally, the architecture of both models is presented in the (Fig. 4).

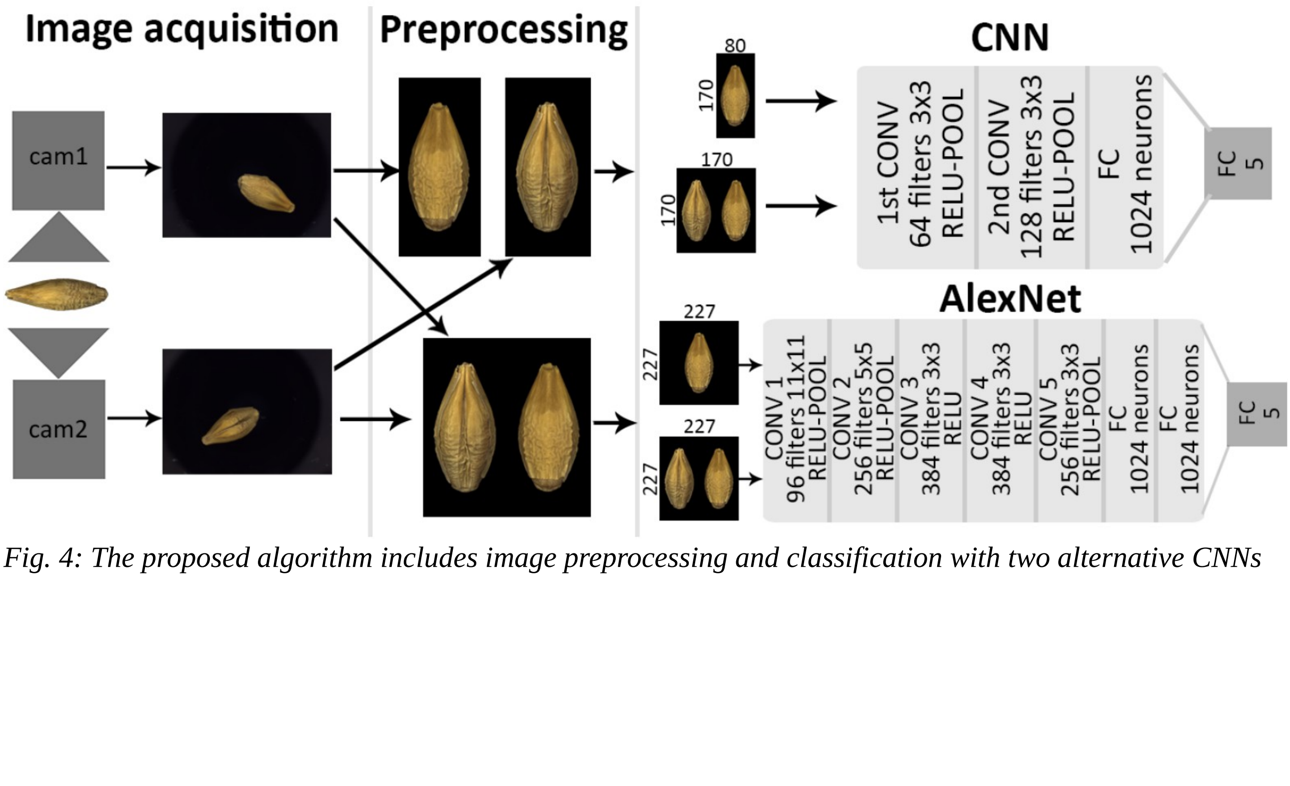

Fig. 4: The proposed algorithm includes image preprocessing and classification with two alternative CNNs

Training

During training, the network was validated on the valuation data and the classification accuracy was estimated. The training was continued until the accuracy value was no longer increasing. Finally, the network was validated on the test set to compute the credible value of the classification accuracy. The goal of learning is to find a set of weights that minimizes the loss function. The experiments were run on Linux Ubuntu computer with Intel Core i7-4930K CPU @ 3.40GHz with 16 GB RAM and NVIDIA GeForce GTX 780 Ti with 3GB memory. There were four configurations of CNNs examined and undergoing the complete training.

The loss function was calculated using the Softmax Loss Layer, which is a combination of Multinomial Logistic Loss Layer and Softmax Layer. The pro posed model was trained using a modified stochastic gradient descent function loss with momentum. The main hyperparameters were: the highest stable initial learning speed $\alpha = 0.001$, momentum $\mu = 0.8$. The batch size that fits into the available GPU memory. The batch size was 50 (for images with a resolution of 80×170) and was reduced to 8 (170×170).

For the reference method, the hyperparameters were as follows: $\alpha = 0.001$, $\mu = 0.9$, batch size = 35. We set the maximum number of iterations to 200000. A models snapshot is taken every 1000 iterations. Finally, the networks were validated on the test set to compute the credible value of classification accuracy.

Reference method

The method for detection of defective barley kernels was presented in [11]. Fol lowing this publication, we compare the proposed neural network with the image analysis procedure implemented in QMaZda software. The software computes quantitative features to characterize every kernel in terms of morphology (shape), colour and texture. The morphological features are computed from bi nary masks, and they estimate area, height, width, perimeter, minimum and maximum diameters, slenderness, compactness, corrugation, circularity, elongation, Danielson index, Blair-Bliss ratio, Malinowska ratio and Hu moments. The attributes of brightness distribution, texture and colour are extracted from the image fragments bounded by the contours of the masks.

Colour and brightness distribution features are statistics computed from his tograms of colour components. They include components of RGB, YUV, YIQ, HSB, CIE XYZ and CIE Lab colour models. The texture is described in terms of second order statistics derived from the grey-level co-occurrence matrix and the grey-level run-length matrix, magnitudes of Haar, Fourier and Gabor transform components, parameters of an autoregressive model, local binary patterns, and histograms of oriented gradients. Altogether, if the single side of the kernel is considered, we compute over 750 attributes to characterise an individual kernel. Optionally, if the feature vector combines information extracted from the pair of images, both sides, the number of attributes amounts to over 1500.

Not all the attributes carry information relevant for identification of defective kernels. Moreover, applying the 1500-dimensional vectors for training would cause overfitting problems. Therefore, a subset of most discriminative features is selected before training of the classifier. The goal is to establish a feature subset or less dimensional subspace which would enable the best discrimination of every class from each other. The criterion for the selection is a Fisher discriminant resulting from linear discriminant analysis. This procedure leads to the feature space dimensionality reduction from 750 or 1500 to 50. The 50-dimensional feature vectors are used for training the support vector machines classifier. We use the classifier with a 3rd order polynomial kernels.

The entire dataset of image pairs was split into training and test sets containing 80% and 20% of cases, respectively. The procedure of feature selection and classifier training was performed with the training set images. Finally, the classification ability was established on the test set.

Results

The results (Fig. 5) are presented in the form of a normalized confusion matrices. Each matrix presents classification counts for one of six examined methods. The three matrices at the top present results obtained by the analysis of one side of the kernels. The bottom row present results generated for the joint images. Matrices on the left relate to the analysis of the proposed CNN architecture, in the middle the AlexNet, whereas the right-hand side ones present the results of the QMaZda reference method. The values presented on the diagonals express the percentage of correctly classified cases, whereas the values out of the diagonal indicate error rates. We use a balanced accuracy to quantitatively compare the results obtained by the four approaches. The balanced accuracy is computed as a sum of diagonal elements divided by the sum of all the elements of the particular matrix.

		Predictions [%]														
		CNN (one side)					AlexNet (one side)					QMaZda (one side)				
ID	True label	1	2	3	4	5	1	2	3	4	5	1	2	3	4	5
1	broken	98,9	0,4	0,4	0,2	0,2	99,1	0,2	0,5	0,0	0,3	98,0	0,3	0,7	0,5	0,5
2	infected	0,0	90,3	0,5	7,9	1,3	0,2	88,4	1,8	7,5	2,1	0,0	91,2	1,5	5,2	2,2
3	missing germ	0,2	1,3	96,7	0,9	0,9	0,0	1,0	94,8	2,6	1,6	0,0	1,5	93,9	1,1	3,4
4	green	0,6	20,5	1,6	75,2	2,1	0,0	20,2	2,7	74,8	2,3	0,0	19,0	2,1	73,4	5,6
5	normal	0,0	5,0	0,9	1,8	92,2	0,0	6,3	1,7	3,9	88,0	0,0	4,4	6,2	2,8	86,7
		CNN (two sides)					AlexNet (two sides)					QMaZda (two sides)				
ID	True label	1	2	3	4	5	1	2	3	4	5	1	2	3	4	5
1	broken	99,3	0,0	0,3	0,3	0,0	99,7	0,0	0,3	0,0	0,0	98,0	0,2	1,2	0,2	0,5
2	infected	0,0	92,3	0,7	5,7	1,4	0,0	89,1	0,7	9,9	0,2	0,0	90,7	0,8	7,7	0,8
3	missing germ	0,0	0,4	97,9	0,4	1,2	0,7	0,4	97,5	1,1	0,4	0,2	1,9	95,2	1,0	1,7
4	green	0,0	13,5	0,8	80,8	4,9	0,4	15,5	1,9	78,9	3,4	0,0	16,5	2,1	78,5	2,9
5	normal	0,4	1,3	0,9	1,7	95,7	0,0	2,3	1,4	5,6	90,7	0,0	1,2	2,3	0,9	95,6

100%

50%

0%

Fig. 5: Normalized confusion matrix comparing the classifications of the defects grains

Classification accuracy, error rates and duration of training and classification are compared and summarised in Table 2. The proposed size-tailored CNN model generates about 2% fewer errors than the AlexNet. It is also evident that classification accuracy increases by 1.5-2.5% when the two sides of the kernel are analysed concurrently. The duration of the training is a time from the start of the training until reaching the highest accuracy of classification estimated on the validation set. It can be noticed that the use of the transfer learning significantly reduces the training time in favour of AlexNet (over 3 times).

Tab. 2: Training and classification ratings

	Test samples	Errors	Accuracy [%]	Test duration on CPU [s]	Test duration on GPU [s]	Training [hh:mm:ss]
CNN (one side)	2972	274	90,80	484,3	8,9	05:13:49
CNN (two sides)	1486	102	93,14	506,0	9,6	05:35:49
AlexNet (one side)	2972	312	89,49	3615,3	29,2	01:35:22
AlexNet (two sides)	1486	131	91,18	1421,8	16,4	01:37:27
QMaZda (one side)	2972	350	88.2	224.0	N/A	41:00:05
QMaZda (two sides)	1486	124	91.6	111.7	N/A	41:00:05

Combining two images into one to present both sides of the kernel resulted in different duration of the classification process. It extended the duration slightly in the proposed CNN. In AlexNet the test duration decreases significantly by 1.8-2.5 times. Caffe framework can use either CPU or GPU for testing (classification). The use of the graphics processing unit (GPU) resulted in faster computations. The time was reduced by 50-130 times when compared to the central processing unit (CPU) application.

The reference method enabled classification accuracy of 88.2 for analysis of one side of the kernels and 91.6 for analysis of the both-sides images. It places this method between the proposed network configuration and the AlexNet solution. The computation of feature vectors for 23772 cases required 40 hours. The feature selection process lasted 1 hour and the classifier training took less than 5s. The table 2. Presents the total time of all the stages of the analysis. However, it should be noted that once calculated, feature vectors can be reused many times in various selection and training experiments. Computation of 50 features required for the classification of a single kernel takes 0.037s on average. Therefore, feature extraction from 2972 kernels takes 112s. The classification time is negligible, and it was estimated as 0.2s for the entire test set.

Conclusions

In this article, we compared four approaches to classify defects of malting bar ley using a computer vision and convolutional neural network. We proposed the simple CNN model with reduced number and sizes of architectural layers. In comparison with the state-of-the-art AlexNet solution, despite the lower resolution images, the model was able to classify the examined objects with the accuracy higher by 2%. The experiment confirmed that concurrent analysis of images presenting both sides of kernels increases the accuracy of classification by 1.5-2.5%. The overall accuracy obtained by the simple CNN on the images combining views of the two sides of the kernel gained 93%. This result is satisfactory considering the in-class complexity and diversity of the input data. The highest classification errors can be seen in the identification of green and infected cases. These classes are the most often confused by the classifier. From 13% to 20% of green or undeveloped kernels are identified as infected. It can be noticed that an analysis of two-side images significantly reduces these errors. In contrast, the defects originating from mechanical causes, such as broken or missing germ kernels, are discriminated fairly easily, with errors below 2.1% (in the most effective classifier). The highest error in missing germ detection occurred in AlexNet fed with single-side images, and it exceeded 5%. The broken kernels and the kernels with missing germ significantly differ with shape from the kernels belonging to the other classes. This shows, that the CNN can cope with the classification problems requiring the analysis of morphological characteristics. The classification time is 5-10 times shorter in the simplified CNN. This gives to the proposed method the advantage over the pretrained AlexNet. The applications of GPU technology enabled reduction of the classification time by 50-100 times when compared with the usage of the CPU. With the GPU, the classification rate of 300 kernels per second was achieved, which is more than satisfactory in the quality assessment applications. On the other hand, the training of the proposed model took over 5 hours and was 4 times longer than in case of the AlexNet solution. However, we still find this time acceptable in practical applications. We have compared two CNNs, the proposed one required a complete training of the convolutional layers and the AlexNet was pretrained. It can be observed, that the training based on the public image database may be insufficient for analysis of specific biological material. It was demonstrated that the network in which the convolutional layers were trained from scratch, and on the dedicated data, gained a higher classification efficiency. Moreover, the reduction of the number of layers and of their sizes resulted in better generalization capabilities of the proposed network. The classification accuracy of the reference method ranks in between the compared neural networks. Computationally, it is the most efficient when using a CPU. However, complex algorithms for handcrafted feature extraction and nonlinear decision boundaries of the classifiers can be hardly implemented on GPUs. Therefore, having a modern computer system with an efficient GPU, the proposed CNN configuration would be a method of choice. On other systems, methods based on feature extraction, selection, and data classification would prevail.

The results of this study indicate that the recognition of individual defects of barley grains can be achieved by CNN with satisfactory accuracy. The proposed model gains an advantage over AlexNet by shortening the classification time, and it improved the classification accuracy.

Acknowledgements

This work was supported by the National Center for Research and Development (NCBR) in Poland, grant no. PBS3/A8/38/2015.

References

1. Dolata, P., Reiner, J.: Barley variety recognition with viewpoint-aware double stream convolutional neural networks. In: 2018 Federated Conference on Computer Science and Information Systems (FedCSIS), pp. 101–105 (2018)

2. He, K., Zhang, X., Ren, S., Sun, J.: Deep residual learning for image recognition. In: Proceedings of the IEEE conference on computer vision and pattern recognition, pp. 770–778 (2016)

3. Krizhevsky, A., Sutskever, I., Hinton, G.E.: Imagenet classification with deep con volutional neural networks. In: Advances in neural information processing systems, pp. 1097–1105 (2012)

4. Lampa, P., Mrzygłód, M., Reiner, J.: Methods of manipulation and image acqui sition of natural products on the example of cereal grains. Control & Cybernetics 45(3), 339–354 (2016)

5. Manickavasagan, A., Sathya, G., Jayas, D., White, N.: Wheat class identification using monochrome images. Journal of cereal science 47(3), 518–527 (2008) 6. Mebatsion, H., Paliwal, J., Jayas, D.: Automatic classification of non-touching ce real grains in digital images using limited morphological and color features. Com puters and electronics in agriculture 90, 99–105 (2013)

7. Ngampak, D., Piamsa-nga, P.: Image analysis of broken rice grains of khao dawk mali rice. In: 2015 7th International Conference on Knowledge and Smart Tech nology (KST), pp. 115–120 (2015). DOI 10.1109/KST.2015.7051471

8. Ni, C., Wang, D., Vinson, R., Holmes, M., Tao, Y.: Automatic inspection ma chine for maize kernels based on deep convolutional neural networks. Biosystems Engineering 178, 131–144 (2019)

9. Pazoki, A., Pazoki, Z., Sorkhilalehloo, B.: Rain fed barley seed cultivars identifica tion using neural network and different neurons number. World Applied Sciences Journal 22(5), 755–762 (2013)

10. Simonyan, K., Zisserman, A.: Very deep convolutional networks for large-scale image recognition. arXiv preprint arXiv:1409.1556 (2014)

11. Szczypiński, P.M., Klepaczko, A., ek, M.K.: Barley defects identification. In: Pro ceedings of the 10th International Symposium on Image and Signal Processing and Analysis, pp. 216–219 (2017). DOI 10.1109/ISPA.2017.8073598

12. Szczypiński, P.M., Zapotoczny, P.: Computer vision algorithm for barley kernel identification, orientation estimation and surface structure assessment. Computers and electronics in agriculture 87, 32–38 (2012)

13. Szturo, K., Szczypiński, P.M.: Ontology based expert system for barley grain clas sification. In: Signal Processing: Algorithms, Architectures, Arrangements, and Applications (SPA), 2017, pp. 360–364. IEEE (2017)

14. Visen, N., Paliwal, J., Jayas, D., White, N.: Automation and emerging technologies: specialist neural networks for cereal grain classification. Biosystems Engineering 82(2), 151–159 (2002)

15. Zapotoczny, P., Zielinska, M., Nita, Z.: Application of image analysis for the varietal classification of barley:: Morphological features. Journal of Cereal Science 48(1), 104–110 (2008)