The Seventh International Workshop on

Computer Modeling and Intelligent Systems CMIS-2024

3 of May 2024, Zaporizhzhia, Ukraine

Organized by
National University
"Zaporizhzhia Polytechnic",
Department of Software Tools

Web-page:
https://www.fb.com/groups/cmis.workshop/

COLLECTION OF PAPER PRESENTATIONS

2024

Convolutional-capsule model for wheat classification

Yaroslav A. Kulyk, Bogdan P. Knysh, Oleh V. Bisikalo, Volodymyr M. Papinov, Vladyslav V. Kabachii

Vinnytsia National Technical University, Vinnytsia, Ukraine

CMIS-2024: Seventh International Workshop on Computer Modeling and Intelligent Systems, May 3, 2024, Zaporizhzhia, Ukraine

kulyk.y.a@vntu.edu.ua (Y. A. Kulyk); tutmos-3@i.ua (B. P. Knysh); obisikalo@gmail.com (O. V. Bisikalo); vnpapinov@gmail.com (V. M. Papinov); vkabachiy@gmail.com (V. V. Kabachii)

0000-0001-8327-8259 (Y. A. Kulyk); 0000-0002-6779-4349 (B. P. Knysh); 0000-0002-7607-1943 (O. V. Bisikalo); 0009-0000-4411-9327 (V. M. Papinov); 0009-0001-3158-2889 (V. V. Kabachii)

Problem statement

Wheat recognition models are important for agriculture in general and for precision farming in particular. These models provide the opportunity for the automation and optimization of this agricultural growth.

It is especially actual when requirements for the productivity and quality of agricultural products are constantly increasing.

The importance of a wheat recognition model in agriculture is its ability to increase the precision of wheat quality assessment, conduct the identification of seed variety and class, and improve the process of crop management.

The use of models significantly facilitates and accelerates the control of compliance with crop cultivation standards. Using recognition models, it is possible to detect in time signs of diseases and pests that allow us to take operational measures for their control, prevent their spread, and model, and assess the influence of different agents on the efficiency of wheat growing.

The application of models helps to decrease expenditures, improve product quality, and initiate new possibilities for agriculture establishments.

Literature review

Deep Fourier neural network (FFCDNN).

The FFCDNN model reaches the general precision of 92.12%, overwhelming another base model such as BIT-DNN. This model shows reliability and generalization during the detection of plant stress with significant improvements in precision and interpretation in comparison with classic models. However, this model cannot assess the degree of vegetative growth of wheat.

R-CNN

The R-CNN shows possibilities of the wheat spikes detection. Despite the difficult conditions of image capturing, R-CNN models reached the high precision of detection in the range of 88%-94% showing the efficiency of deep learning in the analysis of wheat spikes. This model works well when the distance to wheat is small.

CNN-based

The research presents the method for the recognition and classification of many wheat diseases. The proposed CNN model reaches 98.84% precision in the classification of healthy wheat crops and various categories of diseases. However, this model has problems with transferring knowledge to other models.

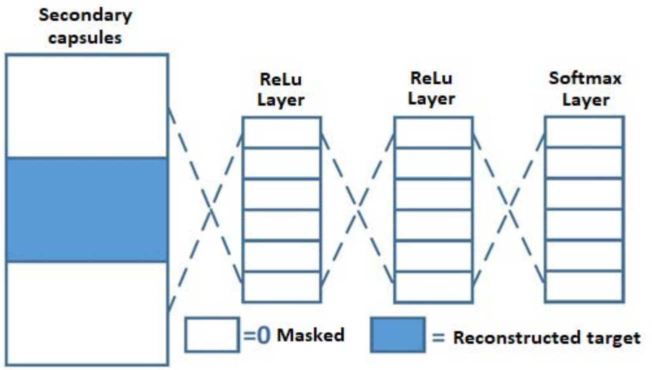
CapsNets shows stable learning processes with a lower sensitivity to knowledge transferring in comparison with tested CNN underlining their reliability and stability during learning in case of complicated datasets with limited training data. Preprocessing methods such as normalization and oversampling can work differently on CapsNet in comparison with CNN and in some scenarios, CapsNet is more affected by preprocessing methods.

In this manner, the combination of capsular networks and convolutional neural networks opens new possibilities for the improvement of precision, reliability, and efficiency of image classification tasks. Such factors as dataset complexity, learning time, and preprocessing techniques play a crucial role in the definition of the general productivity of such hybrid architectures.

The review shows that no model would be ideal for use in wheat classification tasks and works immediately. That is why there is a need to develop own model.

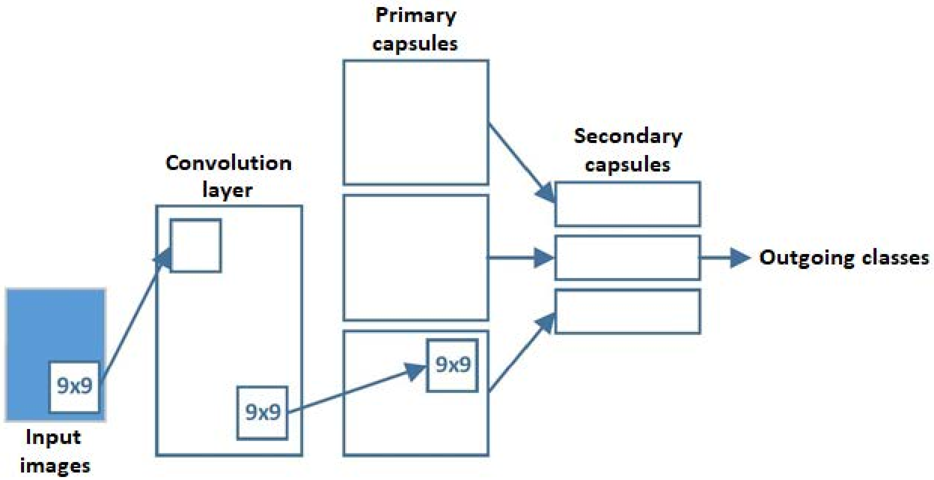
Materials and methods

The CNN-CapsNet architecture consists of two parts namely CNN and CapsNet. Convolutional layers are used to obtain the initial maps of functions. Then maps of functions are loaded into the CapsNet model to carry out the final classification. CNN model uses CNN-CapsNet architecture formed on the base of four convolutional layers.
Each convolution level uses 64, 64, 128, and 128 cores, respectively. This is "3×3". These values are used by a significant part of neural networks with a similar architecture.



The activation function for layers is “ReLU”. After the second layer of convolution, there is a layer of intermediate connection.

At the moment, two models have been developed. One to identify cases based on the binary designation of “ripe wheat" or "unripe wheat". The second model conducts classification based on the three categories of “ripe wheat”, “unripe wheat”, and “diseased/damaged wheat”.



Materials and methods

The goal of computational experiments is the development and estimation of the ability of two classifiers to generalize, especially in the case of CNN and CapsNet. The efficiency of each model will be estimated from such indicators as classification precision, recall, and F-score which are the crucial parameters of measurements of the learned model productivity.

All images are used to calculate the parameters of efficiency. The researched dataset of images consists of training subsets, validation subsets, and testing subsets.

$$precision = \frac{T^P}{T^P + F^P}, \quad (1)$$

$$recall = \frac{T^P}{T^P + F^N}, \quad (2)$$

$$Fscore = \frac{2 \times recall \times precision}{recall + precision}, \quad (3)$$

where:

T^N – negative – when the detection model predicted a negative and the actual value is false.

F^P – false positive - is a type I error where the detection model predicted true, but the actual value is negative.

T^P - positive – when the identification model predicted the true and the actual value is positive.

F^N - false negative - is a type II error where the actual value is positive and the detection model predicted a false

Materials and methods

The following parameters were used in the learning process: optimization algorithm - Adam, gradient drag coefficient β_1 - 0.9, quadratic gradient loading coefficient β_2 - 0.999, training duration (number of epochs) - 30, minimum batch size - 256, initial learning rate α - 0.001, coefficient ε – 1e-8.

At an initial learning rate of 0.001, the model agrees with Adam's optimizer. The marginal loss function is used to determine the characteristics of the class. Its equation:

$$L_k = T_k \max(0, m^+ - \|v_k\|)^2 + \alpha(1 - T_k) \max(0, m^+ - \|v_k\| - m^-)^2 \ , \quad \textbf{(4)}$$

where T_k - class feature indicator, T_k = 1 if class feature k exists,

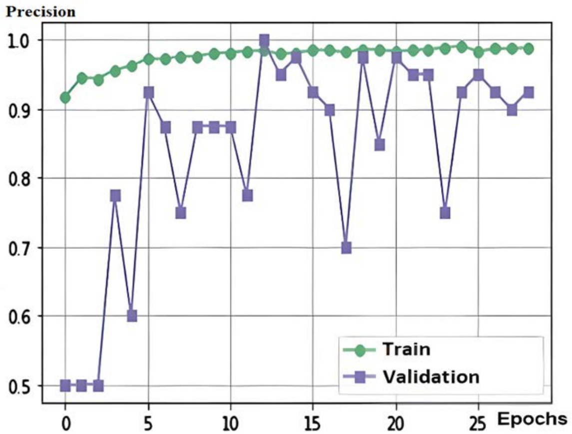
α – initial step size,

m^+ - compute left bias-corrected first-moment estimate, m^+=0.1 to confirm that the length of the vector will remain within the practical limits of m^++=0.9,

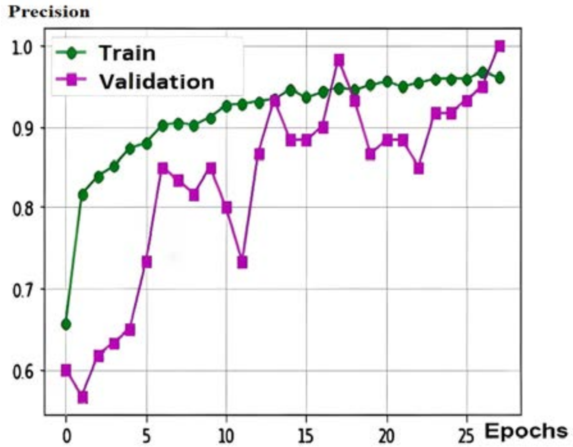
m^- - compute right bias-corrected first moment estimate,

v_k - compute bias-corrected second raw moment estimate.

Experiments and results



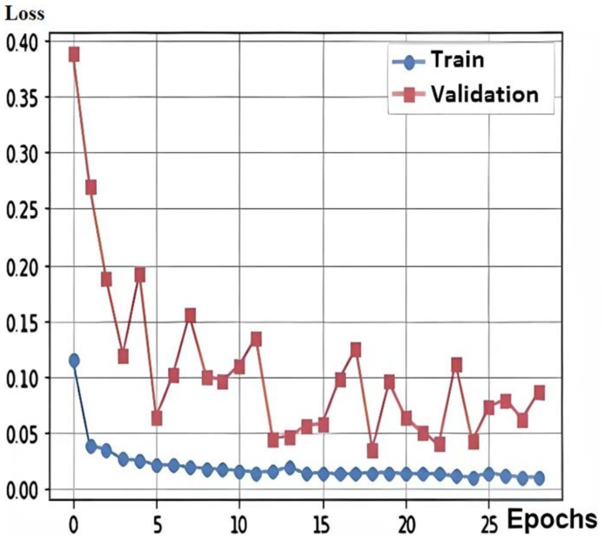
The precision of the CNN-CapsNet model for the binary classification task



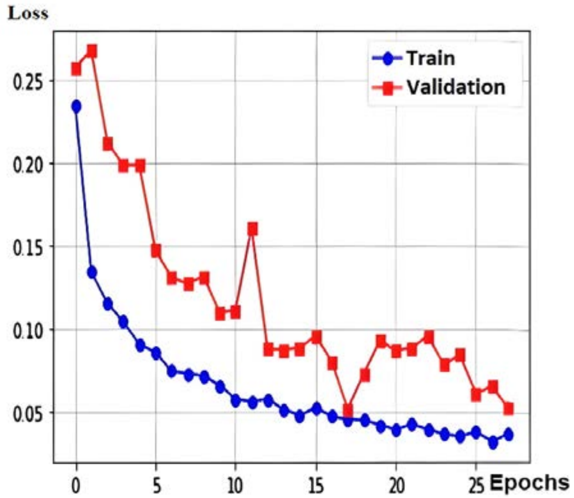
The precision of the CNN-CapsNet model for the ternary classification tasks

Figures shows the learning precision and testing of the convolutional model of the CNN-CapsNet neural network for the binary and ternary classification problem.

Figures show a graph of the classification precision and losses during training and validation of the CNN-CapsNet neural network model for the binary and ternary classification problem. After thirty epochs the learning stopped. After learning, the precision of the model on the validation dataset exceeds 0.9.



Losses of CNN-CapsNet model for binary classification tasks

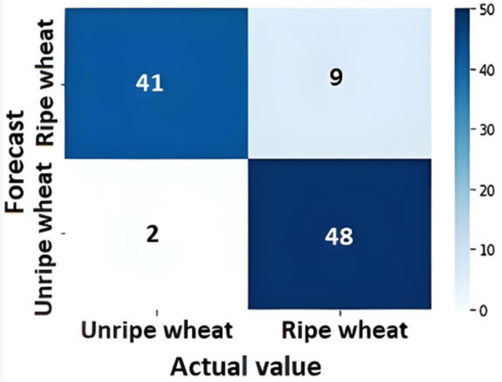


Losses of CNN-CapsNet model for the ternary classification tasks

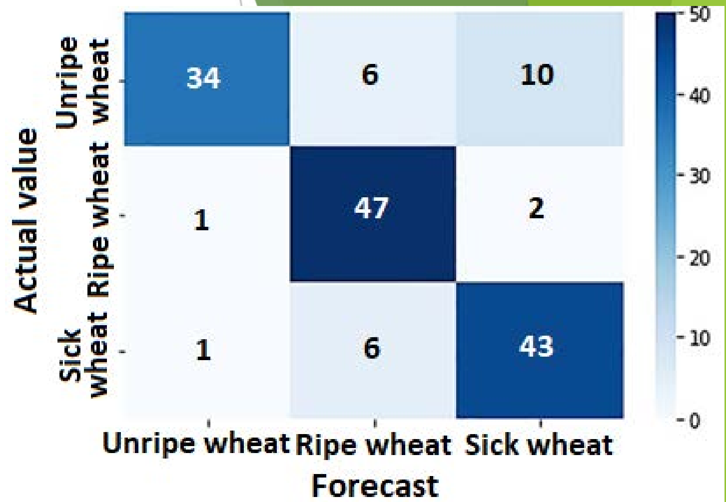
The Hyperspectral Library of Agricultural Crops (USGS) dataset was used to train this model. The image set was developed for all major world cultures and was collected by the US Geological Survey (USGS) which contains 6988 images. Also used dataset “Global Wheat Head Detection 2021” which contains 6500 images of 1024x1024 pixels and 275,000 wheat heads collected from several countries around the world at different growth stages with a wide range of genotypes.

The distribution of the dataset to the training and validation set was in the ratio of 80% to 20% and was carried out by random mixing.

Experiments and results



The dataset Wheat-Ears-Detection-Dataset has 236 high-resolution images (6000*4000) with 30729 ears identified in 20 contrasted genotypes with 6 replicated growths in two environments.



To analyze other characteristics of the 2-class model, in particular (1)-(3), test recognition was performed on 100 images from the dataset Wheat-Ears-Detection-Dataset. The test image set is divided into 50 images of ripe wheat and 50 images of unripe wheat.

Based on the recognition results, a confusion matrix was built, which is shown in Figure.

The confusion matrix compares the actual values of the wheat classes and the model forecast and allows for the calculation of I, and II errors, based on which, according to (1)-(3), it is possible to calculate precision, recall, and F-score.

To analyze other characteristics of the 3-class models, in particular (1)-(3), test recognition was performed on 150 images from the dataset Wheat-Ears-Detection-Dataset. The test image set is divided into 50 images of ripe wheat, 50 images of unripe wheat, and 50 images of sick wheat.

The confusion matrix of the CNN-CapsNet convolutional capsular neural network model for the three-class problem is shown in Figure.

Table 1
Performance results of the CNN-CapsNet model when solving the problem of binary classification

Class	Precision	Recall	F-Score
Ripe wheat	0.90	0.81	0.85
Unripe wheat	0.82	0.90	0.85
Sick wheat	0.85	0.87	0.84
Precision	0.86		

Table 2
Performance results of the CNN-CapsNet model when solving the problem of ternary classification

Class	Precision	Recall	F-Score
Ripe wheat	0.88	0.80	0.84
Unripe wheat	0.84	0.91	0.86
Precision	0.85		

Experiments and results

To check the efficiency of the model, a comparison of this model with others was made according to the precision.

For comparison, models with a similar architecture were used, which combined CNN and CapsNet architectures: FFCDNN, grain-CNN, CapsNet, and Faster-RCNN.

Test images were taken from the Wheat-Ears-Detection-Dataset. The results are aggregated in Table 3.

Table 3

Comparison of precision of model application for wheat classification

Model	FFCDNN	grain-CNN	CapsNet	Faster-RCNN	CapsNet+CNN
Precision	87	85	89	85	90

Example of the classification of wheat

Conclusions

- This article proposed the models of the neural network CNN+CapsNet with improved architecture which combines the corresponding architectures of neural networks and allows us to use the advantages of both architectures. The CNN-CapsNet architecture consists of two parts namely CNN and CapsNet. Convolutional layers are used to obtain the initial maps of functions. Then maps of functions are loaded into the CapsNet model to carry out the final classification. CNN model uses CNN-CapsNet architecture formed on the base of four convolutional layers.
- The model is used to classify wheat into 2 classes (“ripe”/” unripe”) and 3 classes (“ripe”/” sick”/” unripe”). The evaluation of the characteristics of the model is carried out according to the criteria of precision of classification, recall, and F-score based on errors of type I and II. For the precision of the model, confusion matrices were constructed. Precision values mostly rate 0.8-0.9 which can be considered a good result. Errors in classifications do occur, although not often.
- Graphs of the model’s precision and loss demonstrate good values on a validation dataset. The graphs show that the model has no significant overfitting.
- This model can be used for forecasting the quality of the wheat harvest and an independent assessment of a class of wheat, which are important for planning the business processes of agriculture enterprises and for precise farming. The results of comparison with other models allow us to assert its adequacy.