Data Augmentation Techniques for Transfer Learning Improvement in Drill Wear Classification Using Convolutional Neural Network

Jarosław Kurek[1], Izabella Antoniuk[1], Jarosław Górski[2], Albina Jegorowa[2],
Bartosz Świderski[1], Michał Kruk[1], Grzegorz Wieczorek[1], Jakub Pach[1],
Arkadiusz Orłowski[1], Joanna Aleksiejuk-Gawron[3]
[1] *Institute of Information Technology*
[2] *Institute of Wood Sciences and Furniture*
[3] *Institute of Mechanical Engineering*
Warsaw University of Life Sciences – SGGW, Warsaw, Poland
jaroslaw_kurek@sggw.pl

Abstract. This paper presents an improved method for recognizing the drill state on the basis of hole images drilled in a laminated chipboard, using convolutional neural network (CNN) and data augmentation techniques. Three classes were used to describe the drill state: red – for drill that is worn out and should be replaced, yellow – for state in which the system should send a warning to the operator, indicating that this element should be checked manually, and green – denoting the drill that is still in good condition, which allows for further use in the production process. The presented method combines the advantages of transfer learning and data augmentation methods to improve the accuracy of the received evaluations. In contrast to the classical deep learning methods, transfer learning requires much smaller training data sets to achieve acceptable results. At the same time, data augmentation customized for drill wear recognition makes it possible to expand the original dataset and to improve the overall accuracy. The experiments performed have confirmed the suitability of the presented approach to accurate class recognition in the given problem, even while using a small original dataset.

Key words: convolutional neural networks, data augmentation, deep learning, tool condition monitoring.

1. Introduction

During a furniture manufacturing process the drill sharpness is one of many important aspects ensuring the quality of the final product. When the drill is not sharp enough it may have a negative impact on different furniture elements, which in turn may result in financial losses for the company. Therefore, determining the moment when an operator should replace the current drill with a new one is an important and genuine problem. Merely observing the drill state with an unarmed eye is not efficient enough, hence the research to automatize this stage and the use of a computer system for drill wear recognition is conducted.

When approaching this problem usually three classes of drills are considered: red, yellow, and green, to automatically asses the current state of the drill. The red class indicates that the drill should be replaced straight away, since its further use in the production process would probably result in too many damaged elements. The yellow class points to the drill which is suspected of being worn out. Such a drill should be evaluated and either disposed of, or used in further production, depending on the expert opinion. Final, the green state signifies the drill which is still in good condition and can be used in the production process without any changes.

Automatization process for this problem (often called tool condition monitoring or TCM) is not a new concept. Previous solutions, related to drill assessment, usually required a large collection of various sensors used for measuring different signals. Those signals were then used to define the diagnostic features for each drill state. Some of the most common signals are those related to noise, vibrations and acoustic emission, the cutting torque or feed force [4]. Although such setups can generate accurate results, they usually require numerous preprocessing stages before any considerable precision can be reached. Some of the prior steps may include an appropriate sensor and signal choice, the generation and selection of the best diagnostic features, or building the classification model. Such solutions are usually quite complex, time consuming and often expensive in preparation, not to mention that even simple mistakes at any given step may in result bring unacceptable and unusable products. In previous works [1, 2, 3] the authors took into account diverse features that were generated on the basis of different registered signals, yet the accuracy of presented solutions still did not exceed 90% for the recognition of the three classes defined.

One of the goals of this and the previous work (see [5]) is to decrease the complexity of the entire solution. Therefore, similarly as in the former paper, we relied on the assessment of images of drilled holes. In such an approach the only required external equipment is a camera used for taking pictures of drilled holes, which then are used to asses the state of the drill. Recently, the Convolutional Neural Network (CNN) is considered as one of the most effective solutions [6,9] which additionally does not require any specialized diagnostic features to be defined. It has also been proven effective in previous tests in a similar setup [5], therefore we decided to base the current solution on similar concepts.

In the previous solution [5] a CNN was applied to this problem on a limited set of training data (only 242 images, representing three classes: 102 samples for green, 60 samples for yellow and 80 samples for red class). Using such limited collection, the prepared algorithm was able to achieve 85% accuracy for a pretrained CNN algorithm, using AlexNet model created by Krizhevsky, Sutskever and Hinton [11, 12, 13]. Accuracy was improved to 93.4% by applying the Support Vector Machine (SVM) as an alternative

(a) Green (b) Yellow (c) Red

Fig. 1: Examples of holes produced by each drill wear class: Green, Yellow and Red.

to the final CNN layer. In another paper [6] the data were expanded by using simple image operations. CNN with classical deep learning approach was used to recognize two classes of drill wear (sharp enough or not sharp enough). In that case, 900 images were used (300 images for the first class and 600 images for second class), and the original system achieved 66.6% accuracy. The final precision was improved by increasing the number of original images through simple operations (rotation, scaling, adding noise), and resulted in 89% accuracy for 11700 images and 95.5% accuracy for 33300 images. Unfortunately, the learning process was lengthy in that case and lasted over 20 hours.

This paper proposes an approach combining the best elements of both solutions: usage of pretrained CNN applying the transfer learning methodology [11, 12, 14] in order to decrease the training time and expanding original the data set with artificially augmented data, to achieve better accuracy of final solution [6]. We will use three classes for drill wear determination: red (for worn out), yellow (as a warning for an operator) and green (to indicate good condition of the drill). The aim of the current work is to show that the application of transfer learning, even with a very small data set, expanded using data augmentation methods, can achieve high classification precision, without the need for costly equipment or time consuming CNN training (as in the classic approach). The initial data set, similarly as in [5], contains 242 samples of images which represent the three classes. The same dataset was used in our other paper in this volume [15], where we further extend our approach by applying an ensemble of classifiers.

2. Dataset preparation

The data samples containing images of drilled holes for all three classes were collected in cooperation with Faculty of Wood Technology, at Warsaw University of Life Sciences using standard Buselatto JET 100 CNC vertical machining centre. For the test purposes drilling process was performed on standard laminated chipboard (Kronopol U 511 SM), which is typically used in furniture industry. Dimension of the test piece were $150 \times 35 \times 18$ mm. For the drilling process a regular 12 mm FABA drill equipped with a tungsten carbide tip was used.

Used data set has the same structure as in [5]. Out of 242 samples, three subsets of original images were prepared: 102 samples for green class, 60 samples for yellow class and 80 samples for red class. Typical images for each of the specified classes are presented in Fig. 1a, 1b, 1c (images in this Figure and partly in Fig. 2 are similar to those used in our previous papers on drill wear classification due to that we used the same initial set of images).

To expand the initial data set, for each of the original images additional collection of augmented samples was produced. Total of 18 different operations were defined. After performing each of those procedures, the final data set consisted of 4598 images, including 1938 samples for green class, 1140 samples for yellow class and 1520 samples for red class. Operations used for data augmentation were as follows:

1. ColorToGrayscale – convert image to gray-scale values.
2. ColorBrightJitter – adjust brightness of image by random offset in range [-0.3, -0.1].
3. ColorContrast1 – adjust contrast of image by scale factor in range [1.2, 1.4].
4. ColorContrast2 – adjust contrast of image by scale factor in range [1.4, 1.6].
5. ColorHueJitter1 – change image hue by random offset in range [0.05, 0.15].
6. ColorHueJitter2 – change image hue by random offset in range [0.15, 0.30].
7. NoiseGauss – add Gaussian noise to the image.
8. Noise1 – add salt and pepper noise to the image, with strength factor 0.2
9. Noise2 – add salt and pepper noise to the image, with strength factor 0.4
10. Reflection – create a reflection that transforms original image by flipping it in each dimension with 50% probability.
11. Rotate30 – rotate image by random angle in range [0, 30].
12. Rotate45 – rotate image by random angle in range [30, 60].
13. Rotate90 – rotate image by random angle in range [60, 90].
14. Rotate120 – rotate image by random angle in range [90, 120].
15. Scale1 – scale image by random factor in range [1.2, 1.5]
16. Scale2 – scale image by random factor in range [0.8, 0.9]
17. Shear – shear transformation (horizontal), with angle selected randomly in [-30, 30].
18. Translate50 – translate image both vertically and horizontally by random number of pixels in range [-50,50].

Example results for each transformation are presented in Fig. 2. For a detailed description of the used transformations please refer to [20].

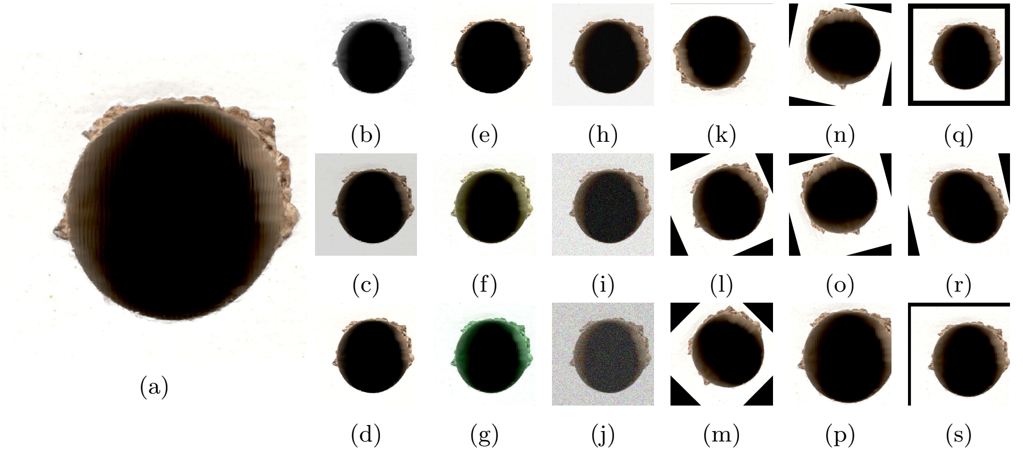

Fig. 2: Examples of augmented data set for a single original image (**a**): (**b**) ColorToGrayscale, (**c**) ColorBrightJitter, (**d**) ColorContrast1, (**e**) ColorContrast2, (**f**) ColorHueJitter1, (**g**) ColorHueJitter2, (**h**) NoiseGauss, (**i**) Noise1, (**j**) Noise2, (**k**) Reflection, (**l**) Rotate30, (**m**) Rotate45, (**n**) Rotate90, (**o**) Rotate120, (**p**) Scale1, (**q**) Scale2, (**r**) Shear, (**s**) Translate50

3. Deep learning approach using CNN

Deep learning approach is one of very popular solutions to classification problems, especially ones meant for images [7, 8, 9, 10]. Usually in such cases CNN would be used as a main function. Due to network structure, and manner in which elements of different type and size are processed, there is no need for the user to specify diagnostic features used to describe particular problem. Deep learning is capable of finding those features without any manual intervention of the user. What is even more important, features extracted in that way are more universal and can work on other types of images with reasonable accuracy.

Contrary to regular methods, deep learning analyses blocks of pixels with fixed sizes (i.e. 5×5, 10×10 or 15×15 for large images). Output signals are filtered, first with linear then nonlinear transformations, and finally use pooling operation to reduce the size of processed blocks. The most significant difference between classical machine learning and deep learning approaches lies in a way in which diagnostic features are extracted. In case of classical machine learning this process constitutes of separate stage of processing (usually consisting of various, often complicated steps). In case of deep learning, generating features is done automatically in hidden layers of neural network. Later on extracted features can be used as input data for softmax classifier or serve as input attributes for external classifiers. For more detailed description of used CNN solution refer to [5].

4. Pretrained CNN AlexNet

Since the database used in current approach is inadequate to start proper training of CNN from the scratch, the only possibility to achieve acceptable results requires using model that was earlier pretrained on large set of different images. Following the methodology of transfer learning [12, 13] we will use AlexNet model [11, 12, 13], that is implemented in Matlab [14]. It was pretrained using over a million images, with 1000 different classes [14, 16].

The AlexNet structure used in presented classification task is built of 9 layers composed of 25 sub-layers, and uses softmax classifier. This function takes u-vector and calculates i-th component of output vector, representing given number of classes: m according to equation:

$$\mathrm{softmax}(u)_i = \frac{\exp(u_i)}{\sum_{j=1}^{m} \exp(u_j)} \tag{1}$$

This function is normalized exponentially while its values are in the range [0, 1] and are treated as probabilities for each class to occur. Component with highest value indicates recognized class. For our application of three classes cross entropy has been used in calculations of softmax function.

Since AlexNet model requires for all images to have the same size ($227 \times 227 \times 3$), we additionally transform our input data to fit those requirements. In our solution first six trained layers were applied without any changes, taking mechanism for creating diagnostic features directly from classes of images used in pretraining stage of AlexNet. Last three, fully connected layers were adjusted during learning process to best fit our image classification problem.

5. Pretrained CNN AlexNet with augmented dataset

Since in previous approach pretrained CNN model achieved promising results, but evidently required more data than the original amount to achieve satisfactory results, in this approach we decided to artificially expand data set. We decided to use augmented data instead of collecting more samples, to minimize the amount of work user needs to perform, to achieve final results. Since taking and preparing images for further use in chosen model can take considerable amount of time it was decided, to instead use basic operations to generate new instances from the ones already collected. We performed 18 operations in total on original data set, choosing random values for most of them to achieve maximal diversity in generated set.

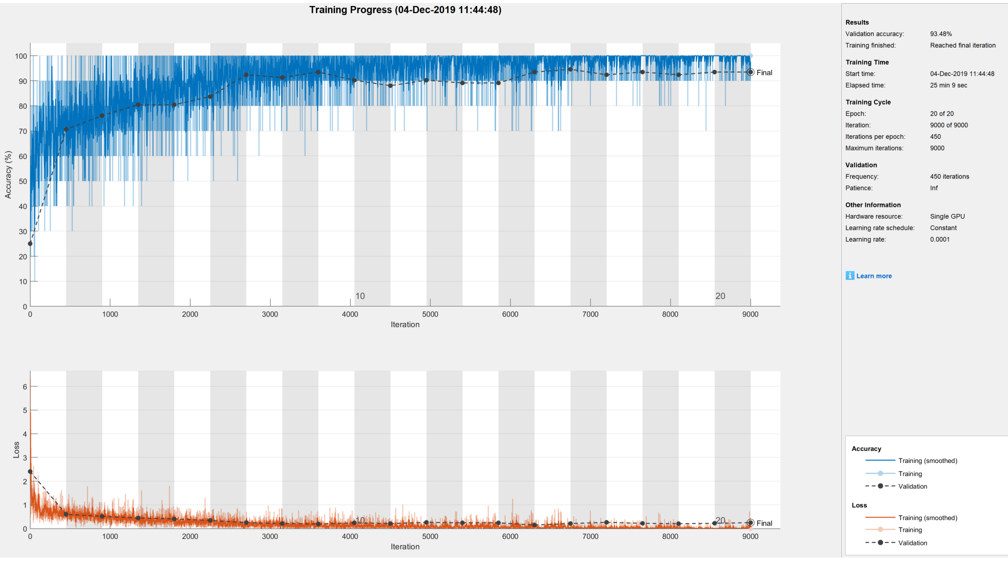

Fig. 3: Outline of training process for CNN model with the augmented data set.

While considering results of previous experiments and CNN characteristics in general, we assumed, that this approach would increase overall performance of our algorithm, without the need to use more complex solutions (like pretrained CNN model, with SVM as a classifier) or significant increase in computation time (like it was the case with classical CNN learning from scratch). At the same time, since additional images are generated, they remain close enough to original ones, to meet the required criteria, while improving overall data set diversity. For an overview of the learning process for this approach see Fig. 3.

6. Results

Original data set used in experiments contained 242 images representing three different classes. After performing data augmentation, extended set for final algorithm contained 4598 images, changed using basic operations such as rotation, translation, adding noise or others (see Fig. 2).

For first two algorithms set of 242 images was split randomly into 10 subsets, where 9 were used during training process, while the final one was used in testing mode. Performed experiments have been repeated 10 times in 10-fold cross validation mode, exchanging test part of the data for each iteration. For third algorithm similar approach

Tab. 1: Accuracy of chosen algorithms applied to the problem of drill condition classification with 3 classes (green, yellow and red). First two algorithms use 242 input images, while last algorithm used artificially expanded set of 4598 images.

Deep learning algorithm	**Accuracy** [%]
Standard CNN	35
Pretrained CNN	85
Pretrained CNN with augmented data	93

was used, but in that case 95% of the data in each iteration was used for training, while remaining 5% acted as testing set. Obtained statistical results in the form of the mean value of class recognition accuracy are presented in Table. 1. Subsequent results refer to standard CNN, which learned from scratch ("Standard CNN"), CNN using pretrained AlexNet model with softmax as a classifier ("Pretrained CNN") and the same pretrained AlexNet model, used on the augmented data set ("Pretrained CNN with augmented data"). Outline of training process for CNN algorithm with an augmented data set is shown in Fig. 3.

As in the previous paper [5], the advantage of using the pretrained CNN is clearly visible. Furthermore, by artificially expanding initial data set, presented approach was able to achieve acceptable accuracy of 93.48%. Such precision is even slightly higher than achieved by solution combining CNN with SVM (Support Vector Machine) used as the final classifier, which was best in previous work (93.4% accuracy). Since former approach is more complex than currently used one, final outcome of presented algorithm is more than acceptable.

7. Conclusion

In this paper we presented an improved method, that applied AlexNet CNN network to problem of recognising drill wear state, based on images of drilled holes. In our approach we used data augmentation, to artificially increase initial data set, and by using pretrained model, we achieved high class recognition accuracy. In previous works such precision was achieved either by using more complex solution or greatly increasing initial data set (in case of learning from scratch, without pretrained network), which in turn resulted in prolonged computations.

Presented results of numerical experiments confirm, that used approach was able to achieve acceptable level of accuracy when predicting states of the drill. The results were compared both to traditionally learned CNN and previous version of pretrained AlexNet CNN without the data augmentation. Final algorithm was able to achieve slightly higher accuracy than more complex solution, which used SVM as a classifier for AlexNet CNN network, while learning time did not exceed 30 minutes. Similarly as in previous approach, pretrained network required only minimal interference in the last layers of CNN. This process was relatively quick, while the resulting model had good generalization properties for the class recognition.

References

[1] K. Jemielniak, T. Urbański, J. Kossakowska, S. Bombiński. Tool condition monitoring based on numerous signal features. Int. J. Adv. Manuf. Technol., vol. 59, pp. 73-81, 2012

[2] S. S. Panda, A. K. Singh, D. Chakraborty, S. K. Pal. Drill wear monitoring using back propagation neural network. Journal of Materials Processing Technology, vol. 172, pp. 283-290, 2006.

[3] R. J. Kuo. Multi-sensor integration for on-line tool wear estimation through artificial neural networks and fuzzy neural network. Engineering Applications of Artificial Intelligence, vol. 13, pp. 249-261, 2000.

[4] J. Kurek, M. Kruk, S. Osowski, P. Hoser, G. Wieczorek, A. Jegorowa, J. Górski, J. Wilkowski, K. Śmietańska, J. Kossakowska. Developing automatic recognition system of drill wear in standard laminated chipboard drilling process. Bulletin of the Polish Academy of Sciences. Technical Sciences, vol. 64, pp. 633-640, 2016.

[5] J. Kurek, G. Wieczorek, M. Kruk, A. Jegorowa, S. Osowski. Transfer learning in recognition of drill wear using convolutional neural network. 18th International Conference on Computational Problems of Electrical Engineering (CPEE) (pp. 1-4). IEEE. September 2017.

[6] J. Kurek, B. Swiderski, A. Jegorowa, M. Kruk, S. Osowski. Deep learning in assessment of drill condition on the basis of images of drilled holes. Proc. SPIE 10225 Eighth International Conference on Graphic and Image Processing (ICGIP 2016), pp. 102251V, February 8, 2017.

[7] L. Deng, D. Yu. Deep Learning: Methods and Applications. Foundations and Trends in Signal Processing, vol. 7, pp. 3-4, 2014.

[8] Y. Bengio. Learning Deep Architectures for AI. Foundations and Trends in Machine Learning, vol. 2, no. 1, pp. 1-127, 2009.

[9] I. Goodfellow, Y. Bengio, A. Courville. Deep learning, MIT Press, 2016.

[10] J. Schmidhuber. Deep Learning in Neural Networks: An Overview. Neural Networks, vol. 61, pp. 85-117, 2015.

[11] A. Krizhevsky, I. Sutskever, G. Hinton. Image net classification with deep convolutional neural networks. Advances in Neural Information Processing Systems, vol. 25, pp. 1-9, 2012.

[12] O. Russakovsky, J. Deng, H. Su et al. ImageNet Large Scale Visual Recognition Challenge International Journal of Computer Vision (IJCV), vol. 115, no. 3, pp. 211-252, 2015.

[13] BVLC AlexNet Model. Online: `https://github.com/BVLC/caffe/tree/master/models/bvlc_alexnet`.

[14] *Matlab 2017a – User manual*, Natick, MA, USA: The MathWorks, Inc, 2017.

[15] J. Kurek, I. Antoniuk, J. Górski, A. Jegorowa, B. Świderski, M. Kruk, G. Wieczorek, J. Pach, A. Orłowski, and J. Aleksiejuk-Gawron. Classifiers ensemble of transfer learning for improved drill wear classification using convolutional neural network. *Machine Graphics & Vision*, 28(1/4):13–13, 2019. doi:10.22630/MGV.2019.28.1.2.

[16] ImageNet. Online: `http://www.image-net.org`.

[17] B. Scholkopf, A. Smola. *Learning with Kernels*, Cambridge: MIT Press, 2002.

[18] M. Kruk, B. Świderski, S. Osowski, J. Kurek, M. Słowińska, I. Walecka. Melanoma recognition using extended set of descriptors and classifiers. Eurasip Journal on Image and Video Processing, vol. 43, pp. 1-10, 2015.

[19] V. N. Vapnik. *Statistical Learning Theory*, New York: Wiley, 1998.

[20] Description of Matlab image transformations. Online: `https://www.mathworks.com/help/deeplearning/examples/image-augmentation-using-image-processing-toolbox.html`.