Hybrid approach towards the assessment of a drill condition using deep learning and the Support Vector Machine

Jaroslaw Kurek
Faculty of Applied Informatics and Mathematics
Warsaw University of Life Sciences
Warsaw, Poland
jaroslaw_kurek@sggw.pl

Abstract— **This paper describes an application of a novel method which relies on applying the fusion of two sets of features obtained from deep learning and hand crafted features. Moreover, in terms of deep learning, transfer learning in deep learning has been applied. A hybrid method has been compared to classical deep learning approach, and let us state that the new approach can obtain more accuracy than regular methods. Transfer learning in deep learning should be used in case of lack of data to train the use of deep learning algorithms. We have used this method for the classification of drill wear state on the basis of drilled hole images. The specialists divided the whole data set into three classes: red, yellow, and green, which correspond to 3 stages of drill wear. The red class corresponds to a drill which is worn out and should be replaced; the yellow class should send a warning message to an operator for a manual check of the state of a drill; and the last one, the green one, is connected with the state of a drill which should be further used in production (it is still sharp enough). The important advantage of this approach is that it collects features from two sources: one set is automatically extracted by means of deep learning and the other is manually extracted by a researcher. The next advantage of this approach is training classification models only on the basis of a small portion of data, which in case of popular deep learning methods is too small to achieve reasonable accuracy. Hence, in order to get around this issue connected with a small portion of training data, transfer learning in deep learning has been applied. The achieved results confirm the fact that this approach can be applied in this situation.**

Keywords—hybrid approach, deep learning, support vector machine, convolutional networks, drill condition, classification

I. INTRODUCTION

The assessment of a drill state in furniture manufacturing industries is an important aspect of business losses. A drill that is not sharp enough can have bad impact on produced furniture elements in the production. It can lead to increased budget losses for a company. It is important to catch the moment when an operator should replace a drill with a new one. Still, watching a drill is not an effective way, and one is still researching methods to automatize this stage by machine without human participation. Specialists say that three classes should be considered: red, yellow, and green. The red class means that a drill should be replaced immediately and should not be used for drilling process, because it will generate company losses (too many damaged furniture elements). The yellow state corresponds to a suspect drill state and manual drill state assessment should be carried out by an expert to validate the current drill state. Manual retrospection by an operator will answer the question whether a drill should be replaced or it is only a falsely positive message generated by the machine. The last state, the green one, can leave production process onward without any time consuming by an operator. Dedicated sensors are used in the regular assessment of a drill state very often, and it is called tool condition monitoring (TCM).

The drill wear process can be assessed by many features, and these features are well known. These features are usually derived from the online registration of different signals in a production stage. The goal of TCM is to achieve the improved and cost effective product quality [1,2,3]. TCM approach is a classical method; it is quite difficult to apply and generates high costs. Many different sensors should be attached to monitor the drill state in the production state.

To build an application which will assess the current state of a drill, many diagnostic features should be generated by registered signals. The most commonly used signals are the following [4,13]:

- feed force,
- cutting torque,
- noise,
- vibration,
- acoustic emission.

The regular approach (TCM) to monitoring the condition of a drill usually consists of several stages:

1) the selection and attachment of dedicated sensors
2) the registration of signals
3) the generation of diagnostic features
4) the selection of the best features for a classifier
5) building a model based on chosen features and classifiers.

The papers [1,2,3] have taken into accounts many features generated on the basis of aforementioned registered signals. The accuracy presented in these papers does not exceed the level of 90%.

The usage of many dedicated sensors for the online assessment of a drill condition generated complexity and costs. To decrease these features (complexity and costs), we suggest using images of drilled holes. Only a camera is needed to take a picture of a hole after the drilling process. On the basis of a drilled hole image, we try to assess a drill condition. The best idea in case of images as input of model is deep learning method. This lets

978-1-5386-8164-0/18/$31.00 ©2018 IEEE

us not focus on diagnostic features and lets the algorithm choose the best on its own.

In paper [5], we have used the deep learning approach for a similar problem, but we have only 2 classes; and we artificially expand a data set several times to prepare the sufficient number of samples for deep learning training purposes. We recall a number from paper [5]. We have 900 images (300 images of the first class, and 600 of the second class) on basis of which we generated 11700, and then 33300 images obtaining accuracy on level of 66.6%, 89% and 95.5% respectively.

In this, we take into consideration the new data set split into 3 classes called the red, yellow, and green one according to traffic lights in production. The data set will be significantly smaller than presented in paper [5]. It means that the problem will be harder than previously; one class more and significantly fewer samples (242 samples).

II. Data acquisition

The samples in form of drilled hole images have been collected thanks to cooperation with Faculty of Wood Technology at Warsaw University of Life Sciences using standard Buselatto JET 100 CNC vertical machining Centre. This acquisition center is depicted in Figure 1.

Fig. 1. The CNC machine used to collect the experimental data.

With the mentioned acquisition center, we have collected some images of holes after the drilling process. Material used for this purpose was a standard laminated chipboard (Kronopol U 511 SM) with the dimensions of 150x35x18mm. This is the same material which is used in furniture industry. Figures (2, 3) below show what a standard laminated chipboard looks like.

Fig. 2.A view of a standard laminated chipboard.

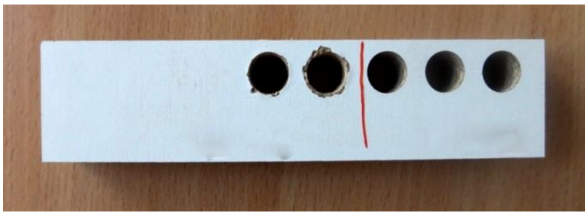

Fig. 3. Examples of holes after the drilling process.

The acquisition of images has been performed by means of a regular drill, used in furniture industry very often -“FABA” - Poland, 12mm in diameter, equipped with a tungsten carbide tip (Fig. 4).

Fig. 4.A two-edged drill (FABA WP-01) with the HW cutting edge with a 12mm diameter.

The database consists of two subsets of data. The first subset is connected with the green class of samples where a drill is sharp enough to continue the drilling process (Figure 5a). After 102 drills (102 images of holes), a drill has been manually blunted by the experts using a microscope (Figure 5b). Then, the further acquisition has been performed and these samples have been assigned to the second subset called yellow (60 images of holes). The last subset, called red, has been retrieved when the experts have blunted a drill enough to leave more damage during the drilling process. This class represents the case when a drill should be immediately replaced with a new one to reduce company’s losses and damage in a standard laminated chipboard (80 images of holes).

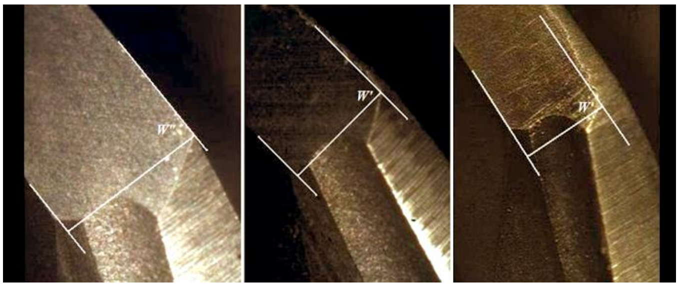

Fig. 5.A drill blunt level analysis under a microsope.

To summarize, the database consists of 242 samples (hole images) with the following distribution:

1) Green class- 102 samples/images
2) Yellow class – 60 samples/images
3) Red class – 80 samples/images

Figure 6 depicts the examples of images of all three classes.

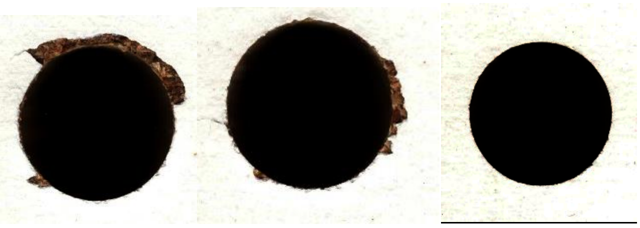

Fig. 6. The examples of images belong to three classes, respectively from left to right: Red, Yellow, Green.

III. Deep Learning Approach

Deep learning is currently a very popular classifier chosen especially, when images are the input. Deep learning is sometimes called deep structured learning or deep machine learning [6,7,8,9]. Deep learning is a group of algorithms having multiple hidden layers with complex structures. Unlike regular methods where pixel by pixel is analysed, a deep learning algorithm analyses a block of pixels filtering the subsequent regions of the size of 5x5 pixels. The result of high complexity and image context analysis rather pixel by pixel is avoidance of using dedicated diagnostic features. Deep learning algorithms try to find features themselves during a training process without manual generation and selection features.

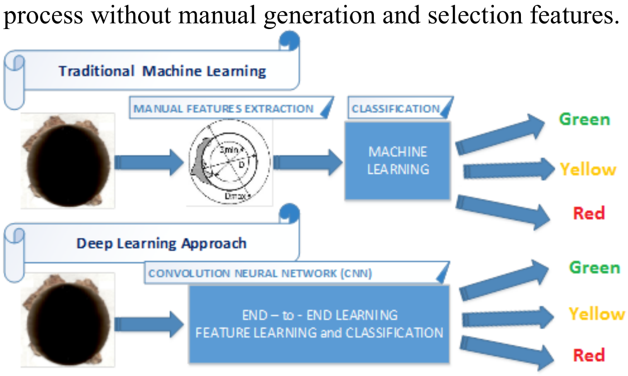

Fig. 7. Comparison of traditional and deep learning approach

IV. Pretrained AlexNet Convolutional Neural Network

In this paper, we have only 242 images divided into 3 classes, which is significantly too small a size of a database to start training a deep learning algorithm. In such cases, when there is a small portion of data to train and test, a special pretrained model can be applied [12,14]. This approach is well-known as transfer learning [12,14].A very common use solution is pretrained AlexNet model [12,14]. This model is pretrained using about 1 milion images of 1000 different types of classes [12,14]. The AlexNet structure used in our task consists of 9 layers composed of 25 sub-layers, which might be presented in the following form [14]:

1. The input layer represents the input of the size of 227x227x3 RGB images with 'zerocenter'; normalization should be applied.
2. The first convolution layer: 96 filtering neurons of the reception field of 11x11x3, stride [4 4], padding [0 0]
   Activation function: ReLU
   Cross channel normalization with 5 channels per element
   Pooling: 3x3 max pooling, stride [2 2] and padding [0 0].
3. The second convolution layer: 48 filtering neurons of the reception field of 5x5, stride [4 4] and padding [0 0]
   Activation function: ReLU
   Cross channel normalization with 5 channels per element
   Pooling: 3x3 max pooling, stride [2 2] and padding [0 0].
4. The third convolution layer: 256 filtering neurons of the reception field of 3x3, stride [1 1] and padding [1 1]
   Activation function: ReLU.
5. The fourth convolution layer: 256 filtering neurons of the reception field of 3x3, stride [1 1] and padding [1 1]
   Activation function: ReLU.
6. The fifth convolution layer: 192 filtering neurons of the reception field of 3x3, stride [1 1] and padding [1 1]
   Activation function: ReLU.
   Pooling: 3x3 max pooling, stride [2 2] and padding [0 0]
7. The first fully connected layer of 4096 neurons
   Activation function: ReLU
   Dropout of 50% chosen randomly.
8. The second fully connected layer of 500 neurons
   Activation function: ReLU
   Dropout 50% chosen randomly.
9. The output layer – the layer of 3 neurons fully connected with the previous layer
   Softmax classifier.

To apply pretrained AlexNet above, images should have the same size as AlexNet first layer of 227x227x3. Moreover, we had to extract all the layers except the last three sublayers (23,24,25) from the pretrained network and add 3 new ones to last layers' structure- appropriate to our issue.

The softmax function takes the vector u and calculates the ith component of an output vector representing M classes in the form

$$\mathrm{softmax}(\mathbf{u})_i = \frac{\exp(u_i)}{\sum_{j=1}^{M}\exp(u_j)} \qquad (1)$$

This function is a normalized exponent and takes the values from 0 to 1. They are treated as the probability of a particular class. The highest value of the component indicates the recognized class. In calculation of softmax function, the cross entropy has been used. In our application for three classes M=3.

V. Feature Generation

A. *Automatically extracted features based on deep learning method*

The advantage of deep learning approach is automatic feature generation based on built-in mechanisms. A researcher has impact on how many features we would like to achieve. In our case, we have set 500 features on one of last sublayers. It is difficult to say what features represent exactly because deep learning is treated as a black box. So we have finally a matrix with the dimensions of 242x500.

B. Hand crafted extracted features

Despite the fact that we have automatically generated features by means of deep learning approach, we can sometimes extend a set of features using manual feature generation. These manually generated features are well-known as hand-crafted features. The authors have generated 21 hand-crafted features which are as follows [11]:

1) The radius of the inscribed circle
2) The radius of the described circle
3) The difference between the maximum diameter of the hole and the diameter of the drill.
4) The area of the hole
5) The convex area
6) The perimeter of the hole
7) The major axis length
8) The minor axis length
9) The solidity factor
10) The extent specifies the ratio of pixels in the region to pixels in the total bounding box computed as the Area divided by the area of the bounding box.
11) Eccentricity specifies the eccentricity of the ellipse that has the same second-moments as the region.
12) Ten descriptors based on distance from the center of gravity on the basis of contours

C. The fusion of two sets of extracted features

Finally, we have merged two sets of features to have 521 features in total which can be applied in numerical experiments. But before that, we had to select a subopptimal set of feature by feature selection to choose only these features which model this phenomenon best. The application of 521 features in comparison to the number of 242 attempts can be inappropriate from the statistical point of view.

VI. Feature selection

To decrease the number of features in comparison to the number of training attempts, the feature selection algorithm has been applied in the form of a sequential feature selection algorithm. This algorithm provides the best class discriminative set of features [12].

This approach detects a subset of features that predicts the classes by selecting features sequentially until there is no further improvement in class prediction accuracy. Starting from an empty feature set, the feature selection creates candidate feature subsets by adding and removing each of the features not chosen yet sequentially. Each candidate feature subset is checked in a 10-foldcross validation by repeating the prediction process with different training and testing subsets of observations [12].

After applying the selection algorithm, we have obtained only 25 features from the whole set of 521 features extracted. The algorithm has chosen 20 features from a deep learning feature set (automatically extracted) and 5 features from hand-crafted generated one (manual approach).

VII. Numerical results of experiments

The data set taken into consideration in numerical experiments consists of 242 images. We have split this data set in a random way into two subsets: 90% for the train process and the other for the test purpose.

TABLE I. The comparison of numerical experiments

Deep learning algorithm	**Feature selection**	**Accuracy [%]**
Standard CNN	Automatically extracted one based on deep learning	35%
Pretrained CNN	Automatically extracted one based on deep learning	85%
Pretrained CNN+SVM (C=1000, gamma=0.01)	Automatically extracted one based on deep learning	93.4%
Hybrid approach: Pretrained CNN+SVM (C=1000, gamma=0.01)	Fusion of automatically extracted one based on deep learning and hand crafted one	95.9%

VIII. Conclusions

Table 1 presents the result of drill condition classification drill into one of 3 classes: green, yellow, and red one, which corresponds to the level of drill damage. The presented results show us huge impact on accuracy in case when we use hybrid feature generation approach in applying pretrained CNN, especially when we have a small portion of data which is insufficient to train with deep learning algorithms. It means that the fusion of features extracted automatically by pretrained deep learning and manually generated features (hand-crafted) can contribute to increasing model accuracy.

IX. References

[1] K. Jemielniak, T. Urbański, J. Kossakowska and S. Bombiński, “Tool condition monitoring based on numerous signal features”, Int J Adv Manuf Technol 59, 73–81, (2012).

[2] S. S. Panda, A. K.Singh, D. Chakraborty and S. K. Pal S.K., “Drill wear monitoring using back propagation neural network”, Journal of Materials Processing Technology 172, 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 Intelligence13, 249-261, (2000).

[4] R. Gonzalez and R. Woods, “Digital image processing”. : Prentice Hall,New Jersey (2008).

[5] Matlab user manual – Image processing toolbox, MathWorks, Natick, USA, (2012).

[6] L. Deng, and D. Yu, "Deep Learning: Methods and Applications", Foundations and Trends in Signal Processing 7: 3–4 (2014).

[7] Y. Bengio, "Learning Deep Architectures for AI", Foundations and Trends in Machine Learning 2(1), 1–127 (2009).

[8] J. Schmidhuber, "Deep Learning in Neural Networks: An Overview", Neural Networks 61, 85–117, (2015).

[9] A. Krizhevsky, I. Sutskever and G. Hinton, “Image net classification with deep convolutional neural networks”, Advances in Neural Information Processing Systems 25,1-9 (2012).

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

[11] M. Kruk, A. Jegorowa, J. Kurek, S. Osowski, J. Gorski, „ Automatic recognition of drill condition on the basis of images of drilled holes”, Computational Problems of Electrical Engineering (CPEE), 2016 17th International Conference, IEEE, 1-4, 2016/11/10

[12] Matlab 2017a, user manual, The MathWorks, Inc. Natick, MA, USA, 2017.

[13] J. Kurek, M. Kruk, S. Osowski, P. Hoser, G. Wieczorek, A. Jegorowa, J. Górski, J. Wilkowski, K. Śmietańska and 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.

[14] ImageNet. http://www.image-net.org