



Machine Vision and Machine Learning

This article is part of the Digitalization Applications 101 learning module, which provides a comprehensive understanding on the basic concepts of digitalization terminologies, technologies and its applications in the steel industry.

The course was developed by the Digitalization Applications Technology Committee as an introductory course to educate industry personnel in digitalization



Authors

Grzegorz Miebs (top row, left)
Data Scientist, PSI Poland, Advanced Analytics Team, Poznań, Poland
gmiebs@psi.pl

Rafał Bachorz (top row, right)
Line Manager, PSI Poland, Advanced Analytics Team, Poznań, Poland
rbachorz@psi.pl

Luc Van Nerom (bottom)
Deputy Managing Director, PSI Metals, Brussels, Belgium
lvannerom@psi.de

Machine vision plays an important role in a variety of applications these days. The rapid development of machine-learning (ML) technologies allows for designing and executing predictive models that can analyze very complex data. Due to the commonly available computational capabilities — mainly based on graphical processor units (GPUs) — and mature machine learning libraries, extensive neural network (NN) architectures can be exposed to a huge amount of production image data. This results in significant predictive capabilities to make a precise decision based on complex input image data.

This paper presents an application of convolutional neural network (CNN) architecture for detecting defects on steel surfaces. The proposed approach is based on composite ML architecture containing two independent NNs. Based on input data in the form of a steel surface image, the solution precisely segments the picture showing the locations of probable defects.

Definition

Machine vision technology is the set of methods used to automatically extract information from an image.⁸ The response from an algorithm can have very different forms. It may be basic like a binary classification of the provided image. On the other hand, a machine vision system can return very complex information about an object present on an image, such as position, size, type and orientation. Also, the image is not limited to typical photography. The subject of a machine vision can be a normal photo as well as an x-ray, ultrasound image, infrared image, 3D image and any other possible way of imaging. This makes machine vision a wide area with application possibilities in a variety of industries.

Within this field, there is a combination of technologies, software, hardware, integrated systems and human expert knowledge to solve real-world problems meeting the

requirements of industrial automation and similar application areas. The information returned by a system can be used for applications such as automatic inspection and robot and process guidance in industry, for sorting goods based on their appearance, for security monitoring, vehicle guidance, and many others.

Such models can easily be introduced in the steel production management process; for example to detect defect types from production images, to identify scrap type from truck or wagon pictures before unloading or identify the scrap grip contents while preparing baskets, unblock material decisions, identify complex plate pile topology, etc.

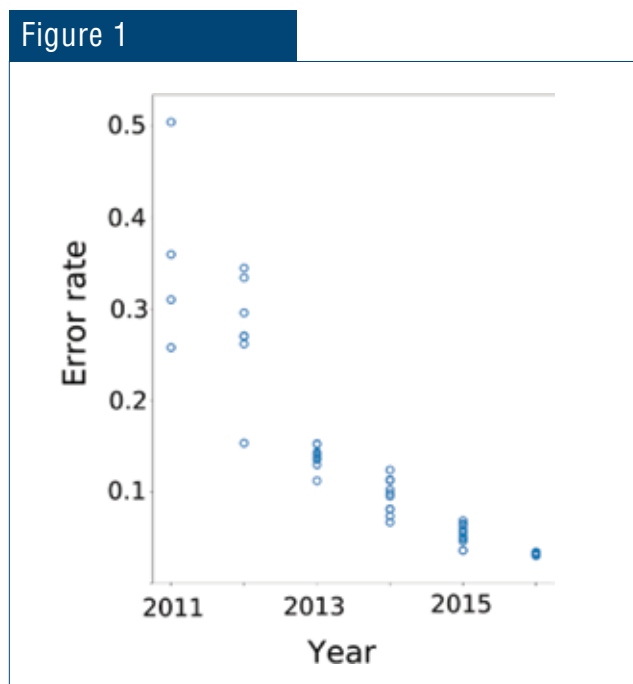
History

Machine vision has a long and prominent history. It is noteworthy to mention the pioneer paper of David Hubel and Torsten Viesel¹ describing the core response properties of visual cortical neurons. After some years, this led Japanese computer scientist Kunihiko Fukushima to invent the idea of Neocognitron,² which can be understood as a prototype of modern CNN architecture. The idea of Neocognitron resulted in a variety of CNN-based architectures solving various problems of machine vision.

ImageNet^{3,4} is a project and challenge that gave much to the area of CNN application in a machine vision. Their database contains more than 14 million hand-annotated images. Additionally, for at least 1 million of the images, the bounding boxes related to the annotations are also provided. The task is to predict which object from more than 20,000 categories is present on an image. Since 2010, the ImageNet project runs an annual software contest in which software programs compete to correctly classify and detect objects and scenes. Each year, contestants introduce more and more complex NN architectures, which achieve constantly improved results. Fig. 1 shows a plot presenting the results, expressed as an average error rate, achieved in this competition over the years.

The community can benefit from this project in many ways. First, the database is a trusted source of annotated images and can be used to train new models. Secondly, one can copy the best architectures, which are tested and proved to work very well. What is more, the technique called transfer learning can be applied here. Both the architecture and trained weights can be taken and considered as an initial pre-trained predictive model. In order to incorporate the nature of steel defect pictures, this model should be only retrained on the final data set.

Technologies



ImageNet competition results – error rate of best models per year.

Nowadays, since there are much more publicly available data, machine learning and especially neural processing are becoming increasingly important in the area of machine vision. In this approach, the more training data related to a target, the better the results. The libraries supporting applications involving neural networks are also very scalable. Whenever there is a need to increase the number of processed images or to increase the architecture size, it is enough to add another GPU. The next important advantage of this solution is there is no need for expert domain knowledge in the given field. One skilled machine learning engineer can solve problems in a variety of areas. Additionally, one can use amazing libraries to build a neural network on a high level of abstraction without worrying about details or performance.

On the other hand, there are plenty of specialized algorithms in the machine vision field. Using them, one can build from scratch a fully functional machine vision system with significantly less training data available. This includes such methods as morphological operations, multiple filters, edge and shape detection, optical character recognition (OCR), etc. For such approaches, there are also efficient libraries that allow users to use all these methods in a fast and easy way.

Both approaches can also be mixed. One can pre-process and post-process an image using image processing algorithms while the core part will be done using CNN. There are also solutions based on NNs that can be used as a part of traditional processing, e.g., facial recognition.

Example Application

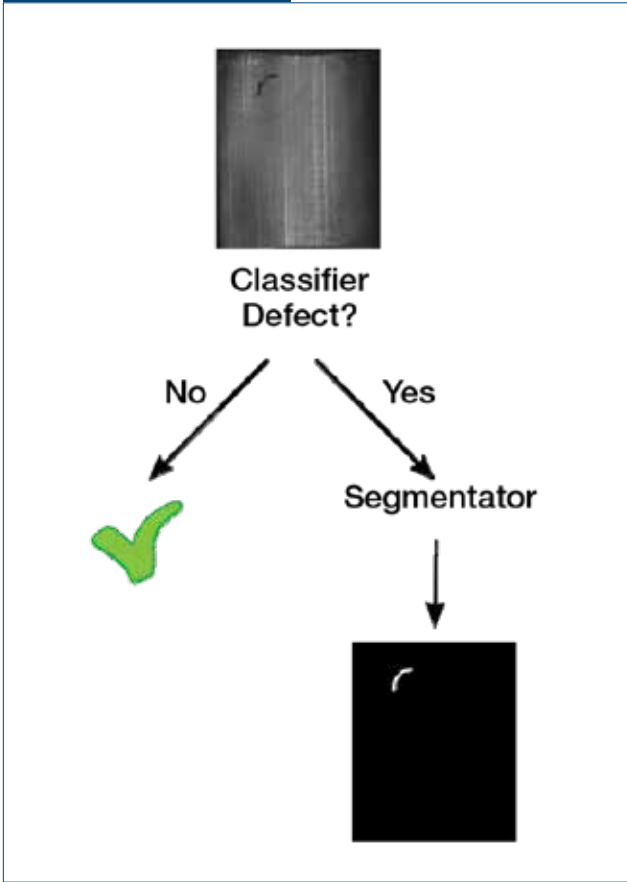
Machine vision in the steel industry can be used for surface defect detection. It allows us to find defects much faster and more precise than any human being. Within a fraction of a second, a modern solution can analyze hundreds of steel surface pictures and determine whether there is a defect as well as the defect type and the position of it.

To solve a steel surface defect detection problem, a composite approach consisting of two predictive models (see Fig. 2) is proposed. The first one — a classifier — can detect whether there is a defect and which defect type it is. If this model returns positive information, then the image is passed to the second model — a segmentator — which detects the position of the defects.

It is worth mentioning that introducing the classifier part allows a significant reduction of false positives.

In order to build a model that can achieve good results on the relatively small data set,⁶ transfer learning is applied twice. First, an NN architecture called EfficientNetB0 is used, trained on the ImageNet data set, and slightly customized last layers to adjust

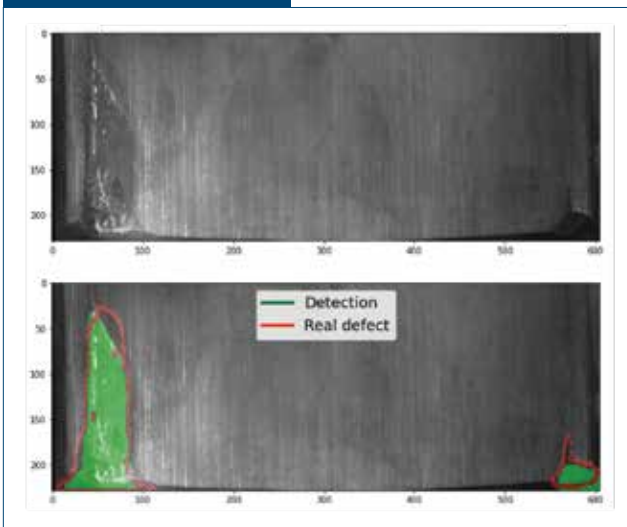
Figure 2



The predictive model composite architecture: classifier and segmentator.

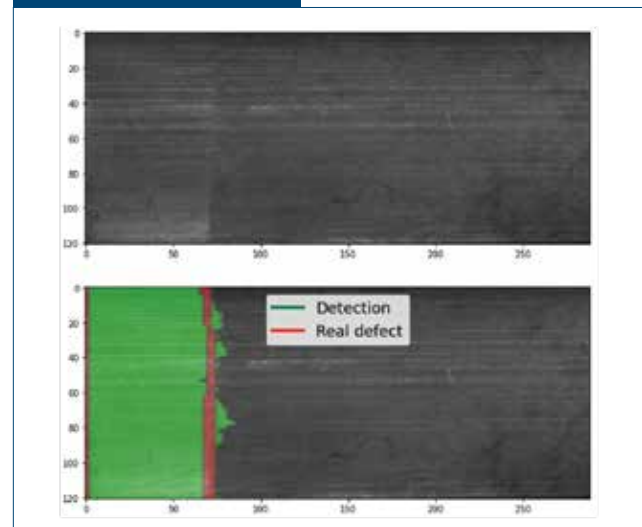
architecture to a given problem — classification or segmentation. Then, these models were retrained on a larger database containing the proprietary images with steel defects. At this level, the models are capable of detecting defects on a steel surface, but only like those from this training data. Then transfer learning is used one more time — copying these models and again retraining them on a new set of images. This time an open data set was applied, containing six common magnetic tile defects.⁶ In the case of the first data set used, there were approximately 50,000 pictures, among which about one-fourth were related to the defects. The second data set contains 1,345 pictures, where almost 400 were related to the magnetic tile defects. All pictures were scaled to a common resolution of 120 × 800 pixels. In order to artificially expand the size of the data set, image augmentation techniques were applied — in particular, rotation, reflection, zooming and blurring. On the top of each picture, the histogram equalization technique was applied. On a separate channel of the picture, the binary bitmap was also provided, containing detected edges reflecting the most important shapes and motifs. Of the pictures, 75% are always for learning and the rest are used for test/validation purposes. The training was carried out in a standard manner by using Keras library with a TensorFlow back end.^{7,8} In order to avoid overfitting, the loss function calculated on top of the validation data set was monitored. The neural network optimization was stopped when the monitored metric was not improving within 5 epochs.

Figure 3



Steel defect detection result: fray defect.

Figure 4



Steel defect detection result: uneven defect.

Usually a couple hundred epochs were enough to converge the neural network. The exemplary results can be seen in Figs. 3 and 4. It is noteworthy that the predictive model can provide the qualitative and quantitative description of the defects of various metallurgical natures. The composite design of the network successfully alleviates the false positive ratio. The classifier part of the predictive model has excellent quality metrics of 0.96, 0.95 and 0.97 in terms of accuracy, precision and recall, respectively. The quality of the segmentator expressed as a dice coefficient, which can be interpreted as a pixel-wise agreement between the prediction and the ground truth, is equal to 0.75. Thus the resulting predictive model is robust and reliable.

There are also many other applications of machine vision in the steel industry.⁹ One of them is a measurement of rhomboidity and ovality in a continuous casting shop. The machine vision system measures the length of diagonals and, based on the difference between them, decides whether the product can be rolled. The others are the use of OCR to read coil details that are printed on it to determine the path of further processing or vision-guided robotics where the output of the machine vision system is used to determine the movement of robotic arms.

Summary

Machine vision is still a growing field of computer science; however, it already can be found in many areas of the steel industry. Further deployments can be expected in nearly all processes as the models are becoming better and more reliable. They can efficiently support human employees or even replace them at some level, making the steel production process more automatized, hence faster, safer and cheaper. An application of machine vision paradigms have been shown here in the context of a particular use case, namely surface defect detection. The currently available machine learning tools allow for creating very effective, robust and reliable applications that can be successfully applied in the context of automated recognition of steel surfaces.

References

1. Hubel, D.H., and Wiesel, T.H., "Receptive Fields of Single Neurones in the Cat's Striate Cortex," *J. Physiol.*, Vol. 148, No. 3, 1959, pp. 574–591.
2. Fukushima, K., "Neocognitron: A Self-Organizing Neural Network Model for a Mechanism of Pattern Recognition Unaffected by Shift in Position," *Biol. Cybernetics*, Vol. 36, 1980, pp. 193–202.
3. Russakovsky, O.; Deng, J.; Su, H.; Krause, J.; Satheesh, S.; Ma, S.; Huang, Z.; Karpathy, A.; Khosla, A.; Bernstein, M.; Berg, A.C.; and Fei-Fei, L., "ImageNet Large Scale Visual Recognition Challenge," *International Journal of Computer Vision*, 2015.
4. Wikipedia, ImageNet, 2020.
5. Steger, C.; Ulrich, M.; and Wiedemann, C., *Machine Vision Algorithms and Applications*, Weinheim: Wiley-VCH, 2018.
6. Huang, Y.; Qiu, C.; Guo, Y.; Wang, X.; and Yuan, K., "Surface Defect Saliency of Magnetic Tile," *IEEE 14th International Conference on Automation Science and Engineering*, 2018.
7. Chollet, F., et al., Keras, <https://keras.io>, 2015.
8. Abadi, M., et al., TensorFlow: Large-Scale Machine Learning on Heterogeneous Systems, <http://www.tensorflow.org>, 2015.
9. Mallik, K.; K., J.; and R., K., "Innovative Machine Vision Applications in Steel Industry," *Steel India*, 2011. ◆