Deep Learning-based System for Quality Control of Coatings in Recess Punch Manufacturing

Balint Newton Turcsanyi, Faisal Saeed and Emmett Cooper

School of Computing and Digital Technology, Birmingham City University, B4 7AP, Birmingham, UK

balint.turcsanyi@mail.bcu.ac.uk,faisal.saeed@bcu.ac.uk and emmett.cooper@bcu.ac.uk

Abstract. Increasing efficiency of the quality inspection process is an on-going pursuit in all manufacturing-related industries. The research was proposed by Tooling international ltd – a company situated in the UK – in an attempt to solve a decade-long problem faced when undertaking quality inspection of their coated products. The main objective of this research is to develop a model that detects faulty products with unsatisfactory coating. In this research, several convolutional neural network (CNN) architectures were tested in order to find the most suitable one for this particular task.

The best performing CNN model delivered 97.68% accuracy which exceeded the company's requirements, providing superior accuracy to when compared to current company methods. This study will be used to develop an automated quality inspection machine, thus enhancing the company's productivity, and will potentially be used as the foundation of further AI-based developments in similar manufacturing-related tasks.

Keywords: Computer Vision, Neural Network, Quality Control, Recess Punch Manufacturing.

1 Introduction

Artificial intelligence (AI) garners increasing attention throughout various fields from healthcare to manufacturing, due in-part to neural networks achieving human level performance and, in some cases even outperforming humans in specialised tasks [1]. The advantages of applying AI within manufacturing are indisputable, this has been recognised by industry as shown by the rapid growth of research literature on the topic of 'AI in manufacturing' [3].

AI-powered computer vision is a specific application of AI heavily used in myriad ways in manufacturing, from stock management [2], personal hygiene in factories [4], sales analysis and quality control [5]. Using AI for quality control can help factories increase their performance; by using them, human errors can be completely eliminated and the automated system benefits by being highly productive around the clock.

This research has been proposed by Tooling international ltd. company in the UK, in an attempt to resolve one of their longest ongoing problems with quality control; namely, the detection of an incomplete or damaged coating layer for their products. Their current methods of detecting damaged products is estimated to have 97% accuracy. The company is specialised in the manufacturing of recess punches used by the fastening industry. Tooling international ltd. is one of the over 400 companies in the Germany-based Würth Group.

Figure 1 shows two examples of recess punches, with the critical areas circled. The top portion of the parts with the hexagonal-like shape is the area where any flaws in the coating would significantly reduce the tool's life expectancy, as the coating is not only providing protection from oxidation, but also reduces the friction between the part and the surface of the screw on impact.

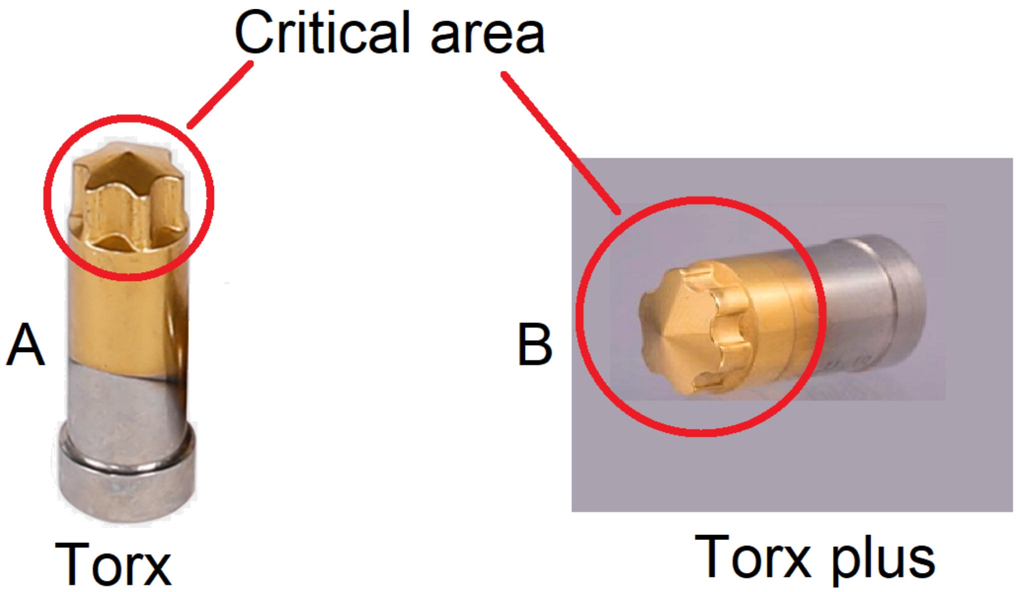

Fig. 1. product examples

2 Related works

When conducting a literature review some of the widely available databases have been used such as: Google Scholar, IEEE, ScienceDirect and MDPI, along with the printed resources of the Birmingham City University's library. When searching for related articles online keywords were used such as: AI and quality control, AI in manufacturing, deep learning in quality control, deep learning and computer vision in quality inspection, and articles mainly from the period of 2018 and 2022. The research is heavily based on computer vision-related topics as considering the nature of the problem, this is one of the proposed methods for detecting the incomplete or damaged coating layer of the company's products. In computer vision, the two types of technologies used are the RGB imaging (imaging in the visible light spectrum) and the hyperspectral imaging (imaging in the invisible spectrum). In light of the nature of the defects this research aims to resolve, hyperspectral analysis is not necessary as the colours of the coating (yellow) and the base material (silver) are completely different and easily distinguishable in the visible light spectrum, using hyperspectral imaging would introduce unnecessary extra financial burden to the company.

A study on object detection using DOTA and HRSC2016 datasets proposed the use of oriented R-CNN, creating oriented bounding boxes around objects rather than the old-fashioned horizontal ones [14]. The method used in this research is worth consideration for the research proposed in this study as it has produced good accuracy as a result of the rotated bounding boxes include less of the background regions and have less chance to include multiple objects (in our case coating faults).

[5] conducted a research on a very similar domain and examined different AI models on small part defect detection achieved 98% accuracy in detecting crocked shapes, 99% in the detection of length-, 97.8% in detection of endpoint-size errors and 79.4% at detecting wringing using single short detector network (SSD) and deep learning (DL) algorithm. The research compared YOLO V3, Faster-RCNN, FPN and the SSD algorithms and concluded with the SSD providing the highest accuracy. The data augmentation technique used by the authors was image rotation (0°, 22.5°, 45°, 67.5°, and 90°), this way the authors increased the size of the original dataset by roughly 50%.

A CNN-based research on small scale fault detection achieved remarkable results using for real-time micrometer scale detection of pits (defects on digital display LCD-films occurring during their manufacturing process, mainly caused by dust- and sand contamination) [15]. This research is important for the proposed research not only because the defects our research aims to detect is also on a small scale (borderline microscopic) but also because they developed a model that uses real-time image feed and, in the future, the same method will be examined and potentially used for this project when building the automated quality inspection machine (might not be possible, depending on the performance of the available hardware). Their model predicts boundary boxes around the candidate defects using an image processing algorithm optimised on pits, and after extracting those patches from a high-resolution image feed it applies a binary classification to detect defects in those patches.

Another similar research on this domain was focusing on the quality control in the food industry called suggested the use of hyperspectral imaging as it would increase the range of detectable defects [16]. This can be useful for future development in this particular domain, for example when a similar model is to be built for some slightly different defects such as the detection of cracking, inner holes etc., however it has been qualified as unnecessary by the authors for this project.

A study on deep learning (DL) computer vision system for the quality inspection of printing cylinders [13] achieved a 98.4% accuracy (only accuracy, the study does not mention other measures of the model's performance) and suggest that this result could be further improved by retraining the model using more data on the falsely classified faults.

The main objective of this study was to conduct a review of existing models, once the database is available. This review was focusing on different CNN models as based on several journals in the topic they are the ones most likely to deliver the highest accuracy scores in image classification tasks [9]. Then, as the image-based data is currently not available, the authors created a dataset by using an automated image generator. An explanation of how the image generator works is in the Methods section. The aim is to detect high accuracy rate of the faulty part. During this process, several data augmentation- and feature extraction techniques were tested. The final step was to evaluate and compare the results of all models to choose the one delivering the best accuracy and the least of loss to decide which model to be deployed. The models were evaluated with confusion matrix, measurement of accuracy, precision, recall and F1 score, as the most commonly used accuracy on its own is not enough to justify the model's performance [12].

3 Methods

This section is a detailed description of the methods and machines were be used to conduct this research from data collection to model evaluation.

3.1 Data Collection

As the image-based dataset is currently not available, it was the authors responsibility to generate it. For this task an automated image generator was built seen on figure 2. The image generator is consisted of a stepper motor powered by an Arduino and a standard industrial USB camera. The stepper motor with a small magnet attached to it ensured that the parts are held steadily in the same position and rotated with a uniform angle between taking the images. The angle of which the parts were rotated by was set to be 0.7 degrees. All together 3496 images were taken and that was divided into 2806 and 690 images for training and validation.

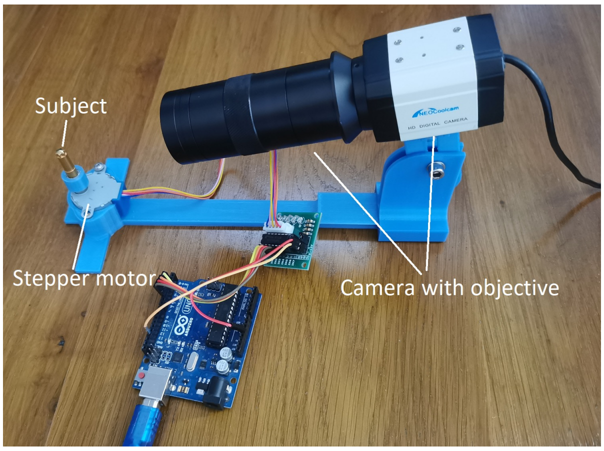

Fig. 2. The automated data gathering machine

3.2 The dataset

Figure 3 shows an example of a part with a faulty coating from the dataset. The resolution of the raw images was 1280 by 720.

Fig. 3: an example of the generated images from the dataset

3.3 Pre-processing and data augmentation

It is clear that the raw images contain a lot of unnecessary information therefore cropping has been applied to the dataset decreasing the resolution from 1280 by 720 to 810 by 538 resulting in a ~53% decrease of the amount of data (pixels) the models needed to process. Figure 4 shows an example of a cropped image compared to a raw one.

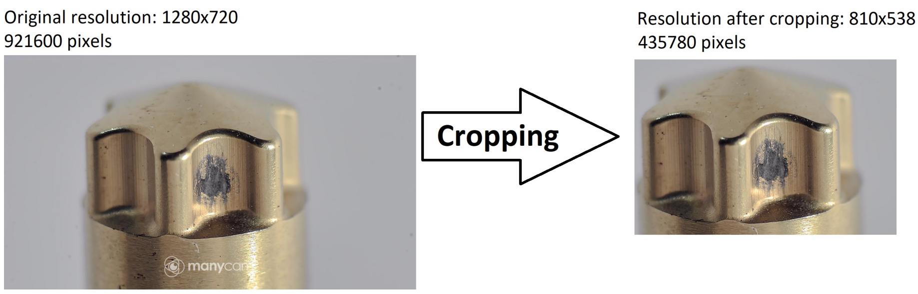

Fig. 4: Visualisation of cropping

During the experiments some data augmentation techniques has been applied to the data, such as rotation and flip of the images, converting them from RGB to grey scale and adding some random noise, this would increase the size of the dataset likely enhancing the overall accuracy of the model [8].

3.4 The Proposed AI Models

The most commonly used AI models for image classification is the convolutional neural network (CNN) [11]. Neural networks consist of an input layer, an output layer and one or more hidden layers. Convolutional neural networks are a type of neural networks that have specifically been designed to process images by having convolution kernel(s) and as a result, these types of models need image inputs. Convolution kernels are two dimensional matrixes with the weighted sum of the pixels of the image as values [7]. Figure 5 shows a diagram of a typical CNN model. CNNs are consisted of several convolutional layers with ReLU activation, pooling layers, a flattening function for feature extraction and this produces the input for the fully connected layer (NN) that performs the classification. The convolution layers pass their values through an activation function to the next layer. Rectifier linear activation function (ReLU) is a function that passes positive values through unchanged but passes 0 if it receives an input below 0. Pooling is a down-sampling step that is used to reduce dimensions and therefore the computational power requirements, by manipulating its value (maxpooling) overfitting can be reduced.

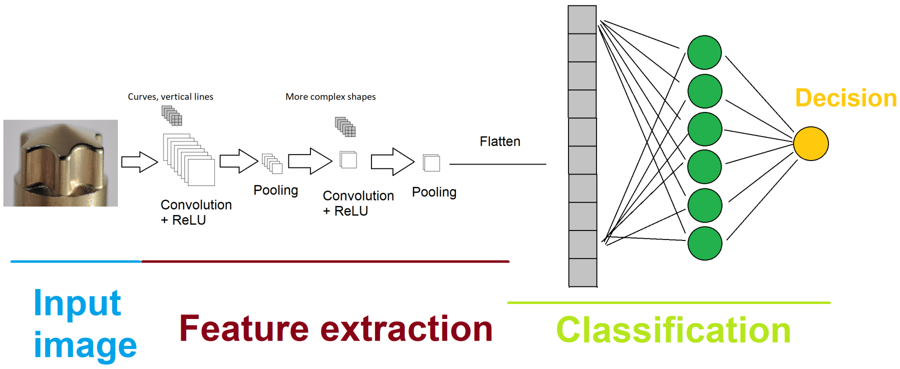

Fig. 5. The CNN Model

3.5 The Model Evaluation

In order to obtain a clear view on the accuracy of the model, a database of images were created of their products with acceptable- and unacceptable coating, this way a labelled database will be available in case should the company wish to test any other machine learning model for the same task in the future. To visualise the performance of the model, a confusion matrix and accuracy/loss graphs were used, and it is expected to achieve 97% and above in success rate in detecting faulty parts. Accuracy measurement is done by the following equation:

$$\frac{TP + TN}{TP + FP + TN + FN}$$

Where TP (true positive): The model successfully detected a faulty product as faulty. TN (true negative): The model successfully detected a not faulty product as not faulty. FP (false positive): The model falsely classified a good part as faulty. FN (false negative): The model failed detecting a faulty part. Out of these possibilities the false negative is the most damaging scenario as it allows faulty parts to proceed without detection, potentially allowing the faulty product to reach the customers.

Precision is the percentage of TP out of all positive findings. Measurement of the precision is done by the equation:

$$\frac{TP}{TP + FP}$$

Recall is the ratio of the true positive instances out of all positive ones, this gives an understanding of how many positive ones the model missed, it is calculated by the equation:

$$\frac{TP}{TP + FN}$$

F1 score is the harmonic mean average of the recall and the precision, it is calculated by the equation:

$$\frac{2 * precision * recall}{precision + recall}$$

F1 score can be useful, however as it gives equal importance to both precision and recall it does not always the best indicator, in some cases weighted F1 score is more suitable [12].

4 Results

Table 1 summarises the results of the 7 tested CNN models. The findings prove the strong viability of CNN in this domain as even the worst performing one achieved over 95%.

CNN model ID	Conv layers	Fully connected layers	Input shape	Validation loss	Accuracy
1	3 conv layers with kernel 7x7 in the first 3x3 in the rest	4 layers with 512, 512, 256, 256 neurons	300x300	0.075	97.68%
2	**3 conv layers with kernel 3x3 in all**	**3 layers with 512, 256 ,64 neurons**	**300x300**	**0.0362**	**97.68%**
3	3 conv layers with kernel 3x3 in all	2 layers with 32, 32 neurons	300x300	0.2895	95.94%
4	3 conv layers with kernel 5x5 in the first 3x3 in the rest	2 layers with 32, 32 neurons	300x300	0.0462	97.39%
5	3 conv layers with kernel 5x5 in the first 3x3 in the rest	2 layers with 64, 64 neurons	300x300	0.07	97.54%
6	3 conv layers with kernel 5x5 in the first 3x3 in the rest	2 layers with 64, 64 neurons	500x500	0.11	96.23%
7	2 conv layers with kernel 5x5 in the first 3x3 in the second	2 layers with 16, 16 neurons	300x300	0.1053	96.67%

Table 1: The summary of 7 CNN models tested during the study

The best performing model (model 2 on table 1) delivered 97.86% accuracy and it consisted of 3 convolutional layers, 7 by 7 kernel size in the first one and 3 by 3 in the rest and 3 fully connected layers with 512, 256, 64 neurons each followed by a maxpooling layer of 2 by 2.

As a last step of evaluation, a set of 215 new images have been created and the model was requested to deliver class prediction on them. Figure 6 shows the resulting confusion matrix where class 0 is the part with defect and class 1 is the good part.

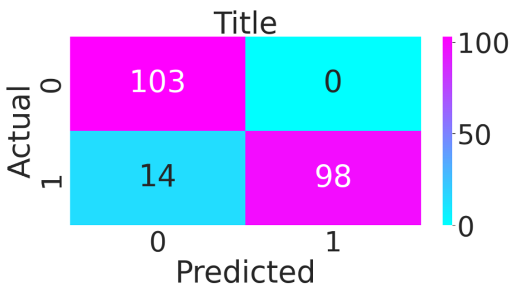

Fig. 6: The model's confusion matrix

The confusion matrix shows that the model was capable to detect the faulty parts effectively but classified 14 good parts out of 112 to be faulty.

Precision: 1.0
Recall: 0.875
F1 score: 0.93

5 Conclusion and future work

The study concluded that model 2 outperforms the currently used methods by the company and has strong viability to be the foundation of an automated quality inspection machine.

As this study aims to provide an algorithm to work in manufacturing, its speed is highly important, it is worth considering applying image segmentation as it is faster and less computationally expensive. To achieve this YOLOv5 model will be tested in future experiments. It is possible that YOLOv5 will deliver even higher accuracy than the current findings.

References

1. Thomsen, M., n.d. Microsoft's Deep Learning Project Outperforms Humans In Image Recognition [WWW Document]. Forbes, 2015. URL: https://www.forbes.com/sites/michaelthomsen/2015/02/19/microsofts-deep-learning-project-outperforms-humans-in-image-recognition/ last accessed 17.06.22.
2. Jobin Mathew, Binumon Joseph, 2021. A Study into the Use of Artificial Intelligence in E-Commerce Stock Management and Product Suggestion Generation for End Users, sections III.C and IV.D, https://doi.org/10.5281/ZENODO.5100977
3. Zeba, G., Dabić, M., Čičak, M., Daim, T., Yalcin, H., 2021. Technology mining: Artificial intelligence in manufacturing. Technological Forecasting and Social Change 171, 120971. https://doi.org/10.1016/j.techfore.2021.120971 last accessed 17.06.22.
4. Fujitsu Develops AI-Video Recognition Technology to Promote Hand Washing Etiquette and Hygiene in the Workplace - Fujitsu Global [WWW Document], 26.05.2020. URL: https://www.fujitsu.com/global/about/resources/news/press-releases/2020/0526-01.html last accessed 17.06.22.
5. Yang, J., Li, S., Wang, Z., Yang, G., 2019. Real-Time Tiny Part Defect Detection System in Manufacturing Using Deep Learning. IEEE Access 7, 89278–89291. https://doi.org/10.1109/ACCESS.2019.2925561
6. Boehm, B.W., 1988. A spiral model of software development and enhancement. Computer 21, 61–72. https://doi.org/10.1109/2.59
7. Lei, X., Pan, H., Huang, X., 2019. A Dilated CNN Model for Image Classification. IEEE Access 7, 124087–124095. https://doi.org/10.1109/ACCESS.2019.2927169
8. Li, W., Chen, C., Zhang, M., Li, H., Du, Q., 2019. Data Augmentation for Hyperspectral Image Classification With Deep CNN. IEEE Geoscience and Remote Sensing Letters 16, 593–597. https://doi.org/10.1109/LGRS.2018.2878773
9. Hasan, H., Shafri, H.Z.M., Habshi, M., 2019. A Comparison Between Support Vector Machine (SVM) and Convolutional Neural Network (CNN) Models For Hyperspectral Image Classification. IOP Conf. Ser.: Earth Environ. Sci. 357, 012035. https://doi.org/10.1088/1755-1315/357/1/012035
10. Pawar, D., 2020. Improving Performance of Convolutional Neural Network! Medium. URL https://medium.com/@dipti.rohan.pawar/improving-performance-of-convolutional-neural-network-2ecfe0207de7 last accessed 06.21.22.
11. Le, J., 2018. The 4 Convolutional Neural Network Models That Can Classify Your Fashion Images [WWW Document]. Medium. URL https://towardsdatascience.com/the-4-convolutional-neural-network-models-that-can-classify-your-fashion-images-9fe7f3e5399d last accessed on 06.21.22.
12. Nighania, K., 2019. Various ways to evaluate a machine learning models performance [WWW Document]. Medium. URL https://towardsdatascience.com/various-ways-to-evaluate-a-machine-learning-models-performance-230449055f15 last accessed 06.26.22.
13. Villalba-Diez, J., Schmidt, D., Gevers, R., Ordieres-Meré, J., Buchwitz, M., Wellbrock, W., 2019. Deep Learning for Industrial Computer Vision Quality Control in the Printing Industry 4.0. Sensors 19, 3987. https://doi.org/10.3390/s19183987
14. Xie, X., Cheng, G., Wang, J., Yao, X., Han, J., 2021. Oriented R-CNN for Object Detection, in: 2021 IEEE/CVF International Conference on Computer Vision

(ICCV). Presented at the 2021 IEEE/CVF International Conference on Computer Vision (ICCV), IEEE, Montreal, QC, Canada, pp. 3500–3509. https://doi.org/10.1109/ICCV48922.2021.00350

15. Ban, G., Yoo, J., 2021. RT-SPeeDet: Real-Time IP–CNN-Based Small Pit Defect Detection for Automatic Film Manufacturing Inspection. Applied Sciences 11, 9632. https://doi.org/10.3390/app11209632
16. De Ketelaere, B., Wouters, N., Kalfas, I., Van Belleghem, R., Saeys, W., 2022. A fresh look at computer vision for industrial quality control. Quality Engineering 34, 152–158. https://doi.org/10.1080/08982112.2021.2001828