Digital Manufacturing Technology
http://ojs.wiserpub.com/index.php/DMT/

Research Article

Detection and Segmentation of Defects in CNC Machine Inserts Using Transfer Learning with Dataset Similarity Evaluation

Chunling Du*, **Gnanaprakasam Naveen**, **Zhenbiao Wang**

Advanced Remanufacturing & Technology Centre (ARTC), Agency for Science, Technology and Research (A*STAR), 3 Cleantech Loop, #01/01 CleanTech Two, Singapore 6371431
E-mail: du_chunling@artc.a-star.edu.sg

Received: 23 October 2023; **Revised:** 4 December 2023; **Accepted:** 4 January 2024

Abstract: The transfer of knowledge from one product to another product has been a highly demanded technique in industrial domains, since it does not need a large training dataset which is often costly available. However, this technique performance may not be always satisfying, due to several issues such as target dataset size or large difference between the source and target datasets. In this paper, we perform transfer learning for segmentation of CNC machine inserts to detect defective inserts using a pre-trained segmentation network. DeepLabv3 framework is adopted for the segmentation task with a modified loss function to speed up its training. A transfer learning strategy with pre-trained model backbone fixed and classifier fine-tuned is applied and the transfer learning performance is investigated on how it relates to the properties such as dataset size. A similarity measure between datasets is proposed to determine which source dataset is the most appropriate for transfer learning on a target dataset.

Keywords: defect detection, deep learning, automated surface inspection, CNC machine insert, defect detection, semantic segmentation, transfer learning

1. Introduction

Transfer learning is a popular approach in deep learning [1] in recent years because it makes it costly and efficient for implementing deep learning models with limited dataset. In the transfer learning method, a pre-trained model is used as a baseline to be transferred to a new dataset. New dataset, which is limited and different, yet similar, is to be trained on the basis of the pre-trained model. This is significant for execution in industrial practices as it reduces much cost consumption of labour and resources in data collection and computation background from manufacturing employees. Therefore, effective and efficient methods are needed for training the new model when adapting to a new dataset. The method of transfer learning is the one that can be applied, using much less newly-collected data for modelling while remaining accuracy.

Transfer learning approach has been extensively applied to manufacturing field. For classification task, it is used in [2] for defect classification of laser welding in battery manufacturing and in [3] for product defect detection in semiconductor manufacturing. Refs. [4, 5] use transfer learning for segmentation task, where [4] is about the detection of metal casting defects with X-ray images as inputs, and [5] reports the application of transfer learning strategies to detect defective PCBs. In terms of semantic segmentation, in [6], transfer learning is performed with a pre-trained fully convolutional

Copyright ©2024 Chunlin Du, et al.
DOI: https://doi.org/10.37256/dmt.4120243812
This is an open-access article distributed under a CC BY license
(Creative Commons Attribution 4.0 International License)
https://creativecommons.org/licenses/by/4.0/

network for segmentation of slums using satellite images with different characteristics, and in [7], the authors have explored transfer learning for semantic segmentation of off-road driving environments, where a pre-trained segmentation network is transferred to a light-weight network and fine-tuned.

In this paper, transfer learning is implemented with three datasets from three different products of CNC machine inserts. The aim is to study the segmentation performance when transferring the pre-trained model trained on one dataset to another dataset. On the other hand, for transfer learning implementation, how to choose appropriate source dataset is currently still an open question [8]. Thereby, we also propose a method to evaluate dataset similarity, and investigate how to use the similarity to select source data to achieve best transfer learning performance.

For the study reported in this paper, we choose DeepLab [9, 10] for segmentation task, a publicly available framework, as it is advantageous in speed, accuracy and simplicity [11, 12]. The DeepLab system is composed of a cascade of two well-established modules: DCNN (deep convolutional neural network) such as VGG16 and ResNets [13], and fully connected CRF (conditional random field). This eases the implementation of transfer learning strategies.

The study therefore includes three portions:

1. Develop transfer learning approach for segmentation task to detect defects in CNC machine inserts. DeepLabv3 [10] framework is adopted with a modified loss function.

2. Investigate how the transfer learning performance relates to the properties such as dataset size.

3. Investigate whether the proposed dataset similarity measure can be used to determine which source dataset is the most appropriate for transfer learning on a target dataset.

2. Dataset description and objective

2.1 *Description of datasets*

In this paper, three datasets of a CNC machine insert are studied, which include images and labelled defects of three batches with different colours. In these datasets, images are paired with labelled defects, i.e., masks. The image and the mask size is 512×512. Figures 1–3 show some examples in the three datasets. The data amount for training and validation is listed in Table 1.

Table 1. Data amount for training and validation for each dataset.

Datasets	1	2	3
Number of training data	3911	1046	1715
Number of validation data	945	235	426

2.2 *Objective and task*

As previously stated, the objective of the paper is to study transfer learning from one dataset to another dataset, that can be accomplished by transferring the feature of pre-trained model while fewer data for training is required.

The task is to detect the defects in images through image segmentation for defects as shown in masks in Figures 1–3. Segmentation is well known for a very long time in the domain of computer vision and image processing. A number of deep learning techniques on the basis of convolutional neural networks have been applied to the task of image segmentation. One of these network architectures for deep learning is DeepLabv3 [10]. In this paper, we shall focus on how to use a pre-trained DeepLabv3 network trained on one dataset for another dataset through transfer learning approach. Transfer learning results from efforts that have been made to be able to train models from limited data, since annotation is tedious and time-consuming for large amounts of data in most practical applications.

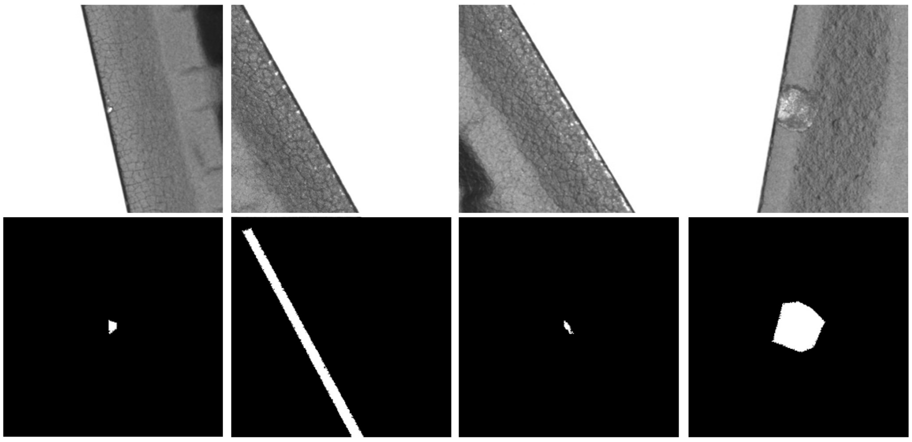


Figure 1. Examples in dataset 1(1st row: images; 2nd row: masks, i.e., ground truth).

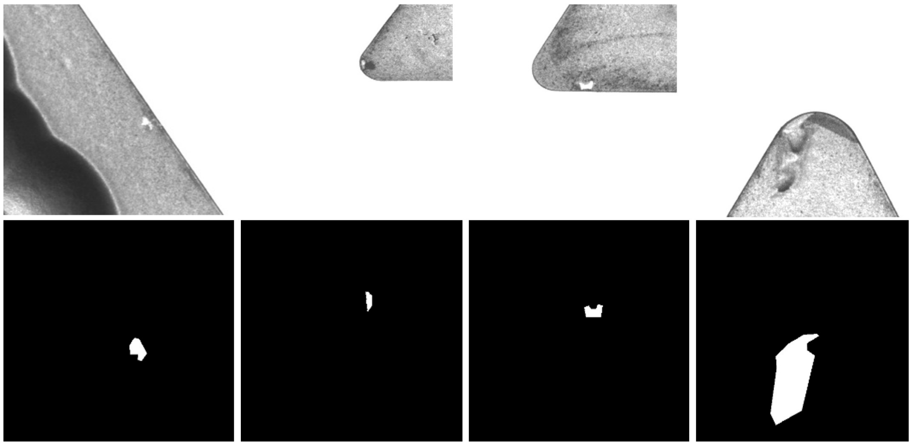


Figure 2. Examples in dataset 1(1st row: images; 2nd row: masks, i.e., ground truth).

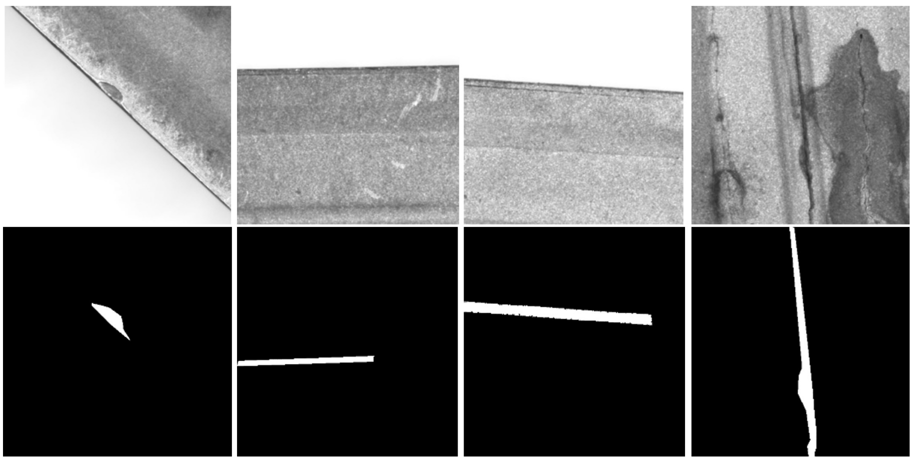


Figure 3. Examples in dataset 1(1st row: images; 2nd row: masks, i.e., ground truth).

3. Method

3.1 *Modified loss function for DeepLabv3 to speed up training*

In this paper, the publicly available framework DeepLab [9, 10] is selected for the segmentation task, as it is advantageous in speed, accuracy and simplicity [11, 12], and composed of a cascade of two well-established modules: deep neural network module such as VGG and ResNets [13], and fully connected CRF module. The cascaded-module structure eases the implementation of transfer learning strategies.

Specifically, DeepLab [9] model is composed of a backbone network and a classifier network, as shown in Figure 4, where the backbone generates feature that is passed onto the classifier. For the application in this paper, we employ DeepLabv3 [10], which significantly outperforms previous DeepLab versions and is comparable with other state-of-art models for semantic image segmentation. We use Resnet50 [13] as the backbone for the DeepLabv3 model in this paper.

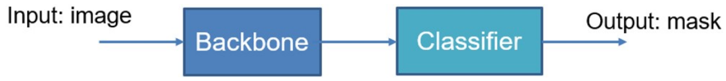


Figure 4. Structure of DeepLabv3.

In the DeepLabv3 original model, the loss function is the sum of cross-entropy terms for each spatial position in the CNN output map. To improve the shape similarity of the target and the predicted masks, the generalised Dice loss [14, 15] is added onto the original cross-entropy loss.

The Dice loss is based on Dice similarity coefficient which is widely used metric in computer vision community to calculate the similarity between two images. It is given by

$$\mathrm{DiceLoss}(y,\hat{y}) = 1 - \frac{2y\hat{y}+1}{y+\hat{y}+1} \tag{1}$$

where $\hat{y}$ is the predicted value of y by the prediction model.

3.2 *Transfer learning approach*

A transfer learning approach is proposed to help ease the training of the convolutional neural network (CNN) algorithm.

The diagram of the transfer learning approach adopted in this paper is shown in Figure 5. Datasets A and B refer to any two datasets in Table 1.

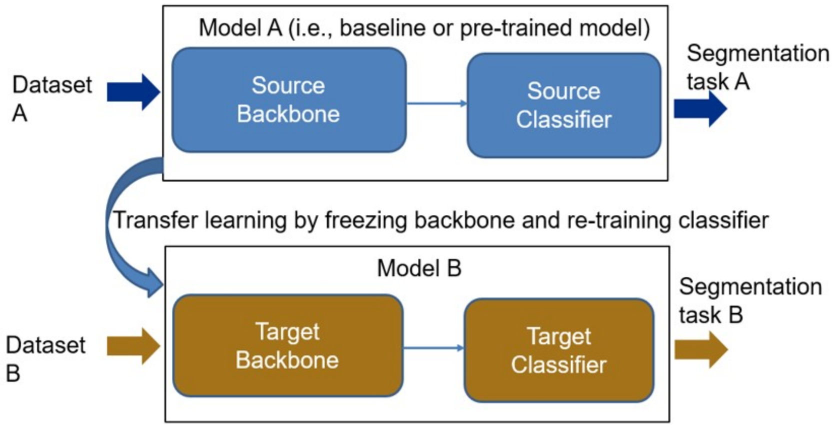


Figure 5. Diagram of adopted transfer learning approach.

In the approach, the backbone of the pre-trained (or baseline) model on one dataset is re-used for training on another dataset using less amount of data. In doing so, before training on a new dataset, say dataset B, all backbone network weights are fixed with values gained from a dataset, say dataset A. Only the classifier is retrained on dataset B. The idea is that the backbone network for image features calculation on one dataset is applicable to a similar dataset, under the assumption that in the segmentation solution spaces, image features are similar for both datasets. In view of this, datasets similarity is to be evaluated before carrying out the transfer learning, as shown in Figure 6.

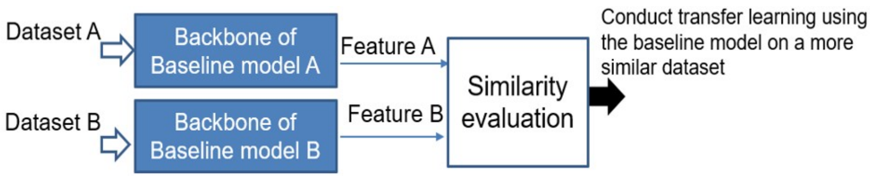


Figure 6. Diagram of transfer learning with datasets similarity evaluation.

The similarity evaluation of two datasets is based on a statistical analysis on Cosine similarity of two vectors u and v, which is defined as the dot product after l2 normalization, and given by

$$\mathrm{sim}(u,v) = \frac{u^T v}{||u||\,||v||} \tag{2}$$

The statistical analysis uses Boxplot method [16]. The features A and B as seen in Figure 6 are converted to two vectors to be used in Equation (2).

All methods proposed in this paper are implemented by using Python and Pytorch [17], open-source platform with libraries that facilitate statistical analysis and the creation of various advanced neural networks.

4. Experiment details and results

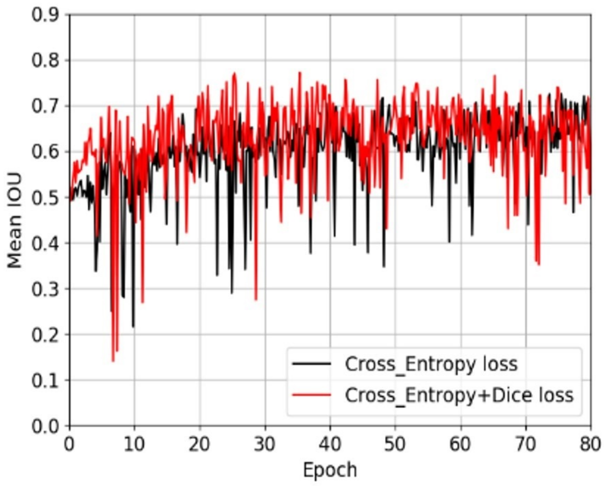
4.1 *Baseline model training*

Training on the three datasets was separately conducted on a GPU with RTX 3080 TI. The batch size was four. The initial learning rate was 0.1, and PolyLR was used as the learning rate schedular policy. The used optimizer was stochastic gradient descent (SGD) optimizer. Images were normalised to have pixel values within $[-1, 1]$ and masks were in [0, 1]. The images and masks were randomly rotated and flipped horizontally at training time.

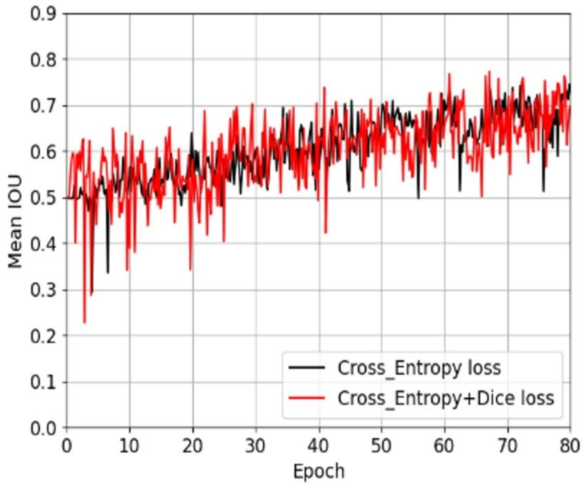
The segmentation performance is measured in terms of pixel intersection of union (IOU). The IoU is the ratio of the intersected area to the combined area of prediction and ground truth. Mean IOU refers to the IOU averaged over two classes (background and defect).

Total training epoch was 80. Validation was conducted at certain iterations during training. During training and validation, mean IOU calculated on the validation dataset, and the model corresponding to the highest mean IOU over the training epochs was saved as the best model and used for performance visualization and analysis. The saved best model was then noted as the baseline model to be used as the pre-trained model in transfer learning.

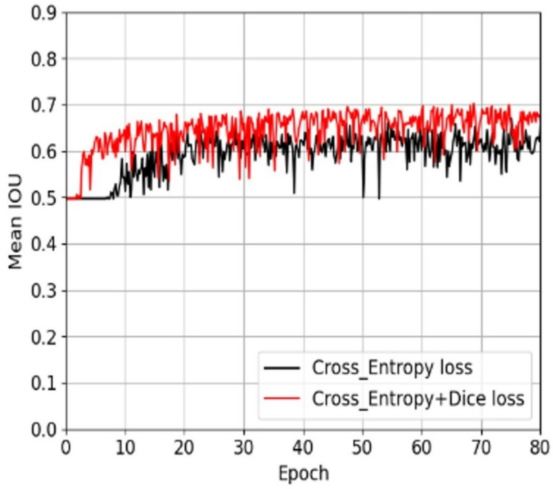
Figure 7 shows the plots of mean IOU on validation datasets. As mentioned previously, the trained model leading to the highest mean IOU on the validation dataset over the training epochs was used as the pre-trained model for the transfer learning. In Figure 7, it is noticed that for all three datasets, the highest mean IOU over training epochs was higher when the new loss function, i.e., cross-entropy plus dice loss, was used. The achieved mean IOU and defect IOU values are listed in Table 2 for the three datasets.



(a) Dataset 1



(b) Dataset 2



(c) Dataset 3

Figure 7. Mean IOU on validation datasets obtained by using different loss functions.

Table 2. Segmentation performance comparison between using two loss functions.

Dataset No.	1		2		3	
IOU	Mean IOU	Defect IOU	Mean IOU	Defect IOU	Mean IOU	Defect IOU
Cross-entropy loss	0.75	0.51	0.74	0.5	0.67	0.35
Cross-entropy plus Dice loss	0.79	0.58	0.77	0.56	0.71	0.42

4.2 *Datasets' similarity evaluation*

The evaluation of similarity of two datasets is based on the trained baseline model on individual dataset, while only the backbone part is used, as seen in Figure 6. The baseline model backbone contains the information from both images and masks for image segmentation tasks, and thus it is reasonable and suitable to similarity analysis of two datasets for transfer learning. The model backbone output is calculated with the images as input and represents the image features. The calculated image features corresponding to randomly selected images (same amount) from each full dataset are then used to obtain similarity metric by Equation (2) and the similarity can be quantitatively assessed. Note that several iterations are needed in order to cover more image subset combinations. The obtained similarity values are randomly distributed and should be analysed statistically. The statistical analysis method BoxPlot is thereby used. Cosine Similarity defined in Equation (2) is adopted as the metric of the similarity level. The boxplots are shown in Figure 8, which reads that the similarity between datasets 1 and 2 is much higher than the other two.

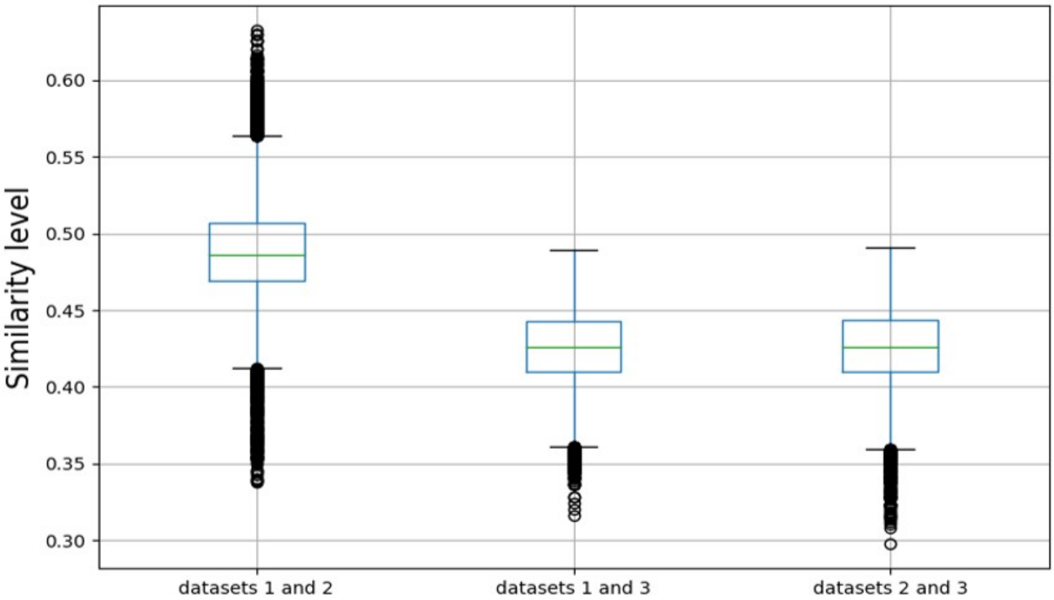


Figure 8. Similarity level Boxplots between datasets.

4.3 *Transfer learning*

Since datasets 1 and 2 are more similar, as analysed in Section 4.2, the transfer learning conducted between datasets 1 and 2 should be more preferred than in the other two. Therefore, to carry on the transfer learning, the baseline model from dataset 2 was used as the pre-trained model, which was to be transferred to dataset 1 and dataset 3 for comparison.

In the training scheme of the transfer learning, the same training process as in Section 4.1 was repeated but the initial learning rate was set to be 0.01, the feature extractor (or backbone) of the pre-trained model was fixed and the classifier was re-trained on the new dataset.

In the transfer learning, the randomly selected dataset with different percentages such as 5, 10, 30, 50, 70, and 100 of full training dataset was taken for training, and mean IOU was obtained on the same validation dataset. The mean IOU values are plotted in Figure 9, for the two transfer learnings of dataset 2 to 1 and to 3. The mean IOU value decreases as less training data is used. It is understood that the performance is well kept even using 50% of the full training dataset. The performance of mean IOU obtained by the transfer learnings of dataset 2 to 1 is always better than that of dataset 2 to 3, which agrees well with the similarity evaluation results in Section 4.2. To better understand the results, some examples are given in Figure 10, with comparison of the predicted defects by using three models with the ground truth. Even using 50% of the whole dataset for the training is still able to give comparable results with the baseline model and the model by using 70% data for the training through transfer learning.

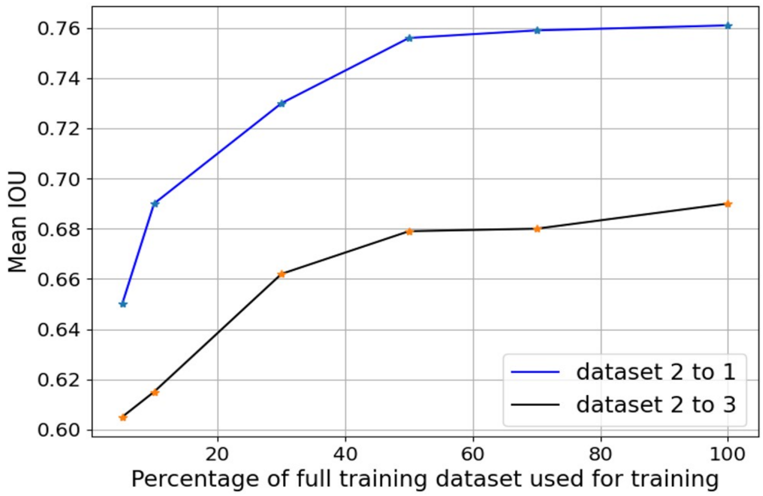


Figure 9. Mean IOU versus percentage of full training dataset used for training for two transfer learnings.

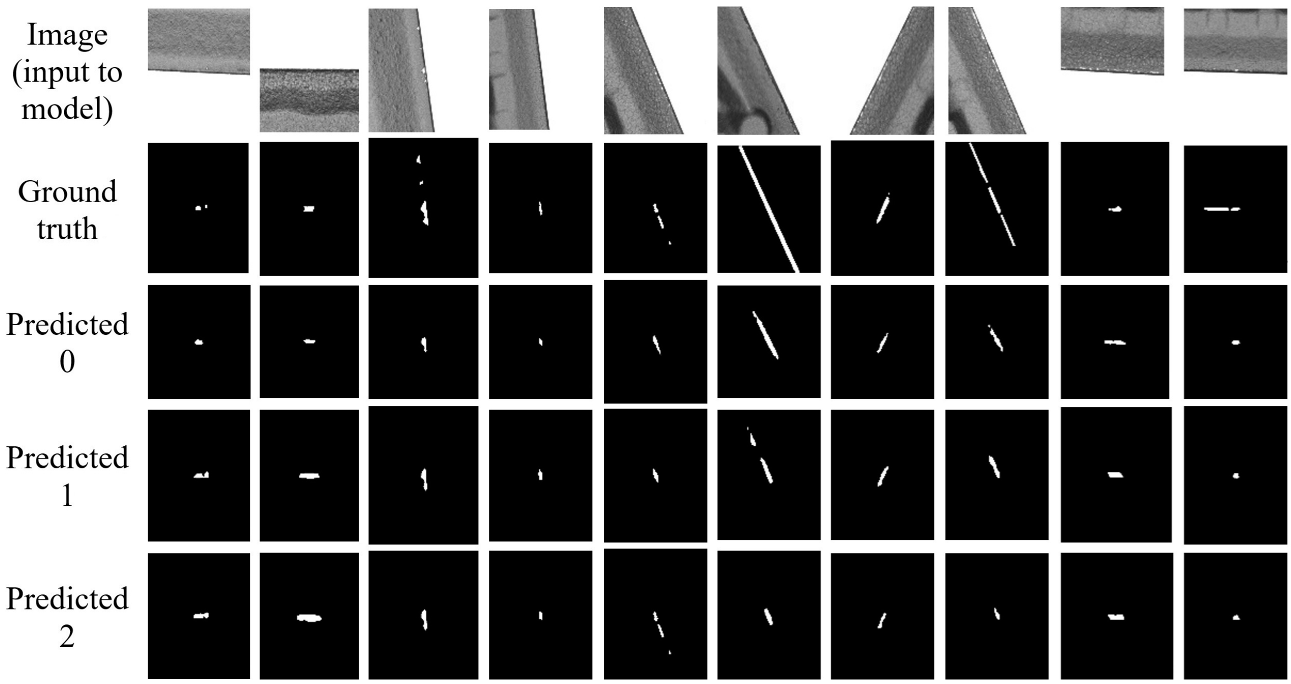


Figure 10. Examples of predicted defects compared with ground truth. Predicted 0: Baseline model; Predicted 1: Transfer learning with 70% training data used for training; Predicted 2: Transfer learning with 50% training data used for training.

5. Conclusions

Using a pre-trained segmentation network, transfer learning has been performed for segmentation of CNC machine inserts to detect defects. The publicly available DeepLabv3 framework has been adopted for the segmentation task. The existing loss function of the DeepLabv3 has been added by the Dice loss in order to expedite the training. The used transfer learning strategy includes the pre-trained model backbone fixed and the classifier fine-tuned on the target dataset. The segmentation performance by using the transfer learning has been investigated on how it relates to the target dataset size used for training. It has been shown that using 50% and 70% of the full dataset for the training through the transfer learning are still able to give comparable results with the baseline model. Furthermore, the similarity measure between datasets is proposed to determine which source dataset is the most appropriate to implement the transfer learning on a target dataset. As a future work, it is interesting to study different strategies of transfer learning such as freezing internal certain layers of the backbone. Moreover, in certain industrial scenarios when more than three channels of the images are available, the developed method can be directly applied as the parameters of the deep neural network are easily adapted. In addition, the proposed transfer learning approach can be extended to other various industrial applications that need defect detection through image segmentation.

Confilict of Interest

There is no conflict of interest for this study.

References

[1] Arinez JF, Chang Q, Gao RX, Xu C, Zhang J. Artificial intelligence in advanced manufacturing: Current status and future outlook. *Journal of Manufacturing Science and Engineering*. 2020; 142(11): 110804. Available from: https://doi.org/10.1115/1.4047855.

[2] Yang Y, Pan L, Ma J, Yang R, Zhu Y, Yang Y, Zhang L. A high-performance deep learning algorithm for the automated optical inspection of laser welding. *Applied Sciences*. 2020; 10(3): 933. Available from: https://doi.org/10.3390/app10030933.

[3] Imoto K, Nakai T, Ike T, Haruki K, Sato Y. A CNN-based transfer learning method for defect classification in semiconductor manufacturing. *IEEE Transactions on Semiconductor Manufacturing*. 2019; 32(4): 455-459.

[4] Ferguson MK, Ronay AK, Lee YT, Law KH. Detection and segmentation of manufacturing defects with convolutional neural networks and transfer learning. *Smart and sustainable manufacturing systems*. 2018; 2: 10.

[5] Silva LH, Azevedo GO, Fernandes BJ, Bezerra BL, Lima EB, Oliveira SC. In *the 2019 IEEE Latin American Conference on Computational Intelligence (LA-CCI)*. 2019; pp 1-6.

[6] Wurm M, Stark T, Zhu XX, Weigand M, Taubenböck H. Semantic segmentation of slums in satellite images using transfer learning on fully convolutional neural networks. *ISPRS journal of photogrammetry and remote sensing*. 2019; 150: 59-69.

[7] Sharma S, Ball JE, Tang B, Carruth DW, Doude M, Islam MA. Semantic segmentation with transfer learning for off-road autonomous driving. *Sensors*. 2019; 19(11): 2577.

[8] Brandt IV, Fok F, Mulders B, Vanschoren J, Cheplygina V. Cats, not CAT scans: a study of dataset similarity in transfer learning for 2D medical image classification. *arXiv*. 2017. Available from: https://doi.org/10.48550/arXiv.2107.05940.

[9] Chen LC, Papandreou G, Kokkinos I, Murphy K, Yuille AL. Deeplab: Semantic image segmentation with deep convolutional nets, atrous convolution, and fully connected crfs. *IEEE transactions on pattern analysis and machine intelligence*. 2017; 40(4): 834-848.

[10] Chen LC, Papandreou G, Schroff F, Adam H. Rethinking atrous convolution for semantic image segmentation. *arXiv*. 2017. Available from: https://doi.org/10.48550/arXiv.1706.05587.

[11] Liu L, Cao D, Wu Y, Wei T. Defective samples simulation through adversarial training for automatic surface inspection. *Neurocomputing*. 2019; 360: 230-245.

[12] Barth R, Hemming J, Van Henten EJ. Optimising realism of synthetic images using cycle generative adversarial networks for improved part segmentation. *Computers and Electronics in Agriculture*. 2020; 173: 105378.

[13] He K, Zhang X, Ren S, Sun J. Deep residual learning for image recognition. In Proceedings of *the IEEE conference on computer vision and pattern recognition*. 2016; pp 770-778.

[14] Sudre CH, Li W, Vercauteren T, Ourselin S, Jorge Cardoso M. Generalised dice overlap as a deep learning loss function for highly unbalanced segmentations. In *Deep Learning in Medical Image Analysis and Multimodal Learning for Clinical Decision Support: Third International Workshop, DLMIA 2017, and 7th International Workshop, ML-CDS 2017, Held in Conjunction with MICCAI 2017*. 2017; pp 240-248.

[15] Jadon S. A survey of loss functions for semantic segmentation. In *the 2020 IEEE conference on computational intelligence in bioinformatics and computational biology (CIBCB)*. 2020; pp 1-7.
[16] DuToit SH, Steyn AG, Stumpf RH. *Graphical exploratory data analysis*. Berlin: Springer Science & Business Media; 2012.
[17] Paszke A, Gross S, Massa F, Lerer A, Bradbury J, Chanan G, Killeen T, Lin Z, Gimelshein N, Antiga L, Desmaison A. Pytorch: An imperative style, high-performance deep learning library. *Advances in neural information processing systems*. 2019; 32.