Defect Data Augmentation Method for Robust Image-based Product Inspection

Youngwoon Choi[1], Hyunseok Lee1, and Sang Won Lee[2]

[1] *Department of Mechanical Engineering, Sungkyunkwan University, Suwon-si, Gyeonggi-do, 16419, Republic of Korea*
woonathome@g.skku.edu
ddsa2210@g.skku.edu

[2] *School of Mechanical Engineering, Sungkyunkwan University, Suwon-si, Gyeonggi-do, 16419, Republic of Korea*
sangwonl@skku.edu

ABSTRACT

In this paper, we develop a model for detecting defects in fabric products based on an object segmentation algorithm, including a novel image data augmentation method to enhance the robustness. First, a vision-based inspection system is established to collect image data of the fabric products. The three types of fabric defects, such as a hole, a stain, and a dyeing defect, are considered. To enhance defect detection accuracy and robustness, a novel image data augmentation method, referred to as the defect-area cut-mix, is proposed. In this method, the shapes that are the same as each defect are extracted using the masks, and then they are added to non-defective fabric images. Second, an ensemble process is implemented by combining the results of two models, one with high sensitivity in defect diagnosis and the other with lower sensitivity. The results demonstrated that the model trained on the augmented dataset exhibits improved metrics such as intersection over union and classification accuracy in defect detection on the test dataset.

1. INTRODUCTION

Recently, there has been a surge in demand for automation in the product manufacturing and inspection processes across various sectors within the manufacturing industry. However, due to the considerable time and cost, product inspection by small and medium manufacturers is usually done manually with the human eye. Meanwhile, human inspections can lead to inconsistent test results based on the examiner's skill level and fatigue. Recently, thanks to artificial intelligence (AI) technology, manufacturing companies have actively applied automated inspection processes that can be adapted to different types of products (Jung et al., 2021). In order to conduct automated inspections of products with diverse geometries, it is necessary to develop a robust image-based inspection algorithm and an imaging system that is robust to external factors, including lighting conditions.

Research on image-based product inspection algorithms can be divided into many approaches: statistical approach, AI model-based approach and hybrid approach (Hanbay et al., 2016). Among various approaches, the statistical approach involves using image processing techniques (such as frequency decomposition or filtering) to extract features from images, while the hybrid approach combines statistical methods and modeling techniques to leverage the strengths of both. Therefore, these can be recategorized as follows: those based on combined image processing and those based on deep learning models such as convolutional neural networks (CNN) (Bhatt et al., 2021). For combined image processing algorithms that involve a mix of image processing steps, it is possible to achieve high inspection accuracy only for specific products with the same geometries. However, this approach has a drawback: if the product type or capturing environment changes, we must adjust the algorithm's parameters or create a new algorithm from scratch. On the other hand, algorithms based on deep learning models can respond to various products and environments depending on the training dataset and can achieve high inspection accuracy. However, since the performance of the inspection model greatly depends on the quality of the dataset (number of data points, diversity, etc.), developing a robust model requires investing time and effort to collect a large number of images of defective products (Russakovsky et al., 2015).

Combined image processing algorithms are primarily used when the shooting environment is consistent and the variety of inspected products is limited. Tong et al. (2016) presented an optimal Gabor filter for inspecting woven fabrics. Zhou (2019) focused on inspecting defects in semiconductor wafers. They applied a median denoising process to images to extract favorable features for defect detection and proposed algorithms using machine learning techniques such as KNN and SVM to classify images. Deep learning model-based algorithms require significant computing resources and longer processing times compared to combined image processing algorithms. However, they offer the advantage of creating robust models that are resilient to variations in

shooting environments and product types, depending on the quality of the training dataset. Tabernik et al. (2020) presented a novel segmentation network and decision network trained on the Kolektor Surface-Defect-Dataset (KolektorSDD), containing product surface defects, to infer surface defects with higher accuracy. Ho et al. (2021) proposed a step-by-step algorithm where an object detection model locates defects, and an instance segmentation model infers the shape of defects in woven fabrics. Deep learning model-based algorithms are generally evaluated to have higher robustness compared to combined image processing algorithms.

However, if the feature distribution of the images in the training dataset (such as brightness, types of products, defects, etc.) is not sufficiently diverse, the model may only accurately infer data within the feature distribution of the training data. For instance, assuming a specific defect in the training dataset has a radius of 0.1~0.2mm and appears darker than its surroundings, the detection may become challenging if the defect is larger than 0.2mm in radius or if the product's color is darker than the training images. This is due to the new data displaying features not seen in the training. To address this issue, collecting more data would be one solution. However, as mentioned earlier, for small and medium-sized enterprises, investing significant time and resources without immediate productivity gains can be challenging. Therefore, there is a need to generate unseen data using observed feature distributions within the dataset.

Therefore, in this work, we present a novel defect data augmentation method, referred to as the defect-area cut-mix, to improve the accuracy and robustness of deep learning-based fabric defect detection models from the perspective of dataset quality. In addition, an ensemble process is applied by combining high and low sensitivity models in the fabric defect diagnosis. Figure 1 shows the schematic diagram depicting the research methodology in this paper.

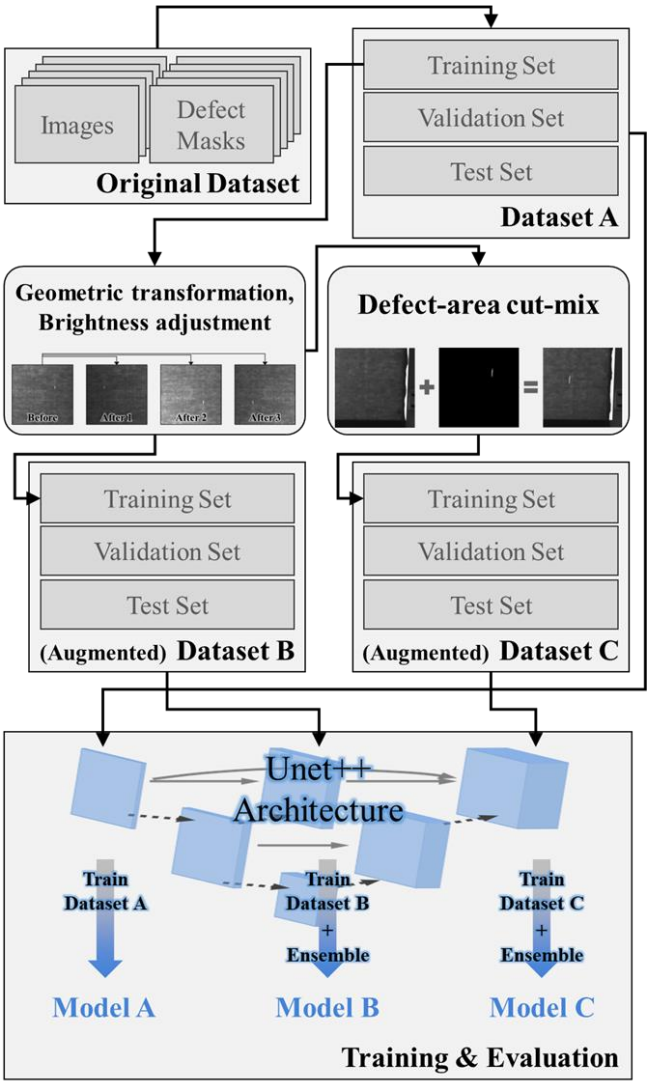


Figure 1. Schematic overview of the research methodology

2. MODEL CONSTRUCTION AND DATA AUGMENTATION

2.1 Image Data Collection and Model Construction

In this section, the research methodology for image collection and defect detection model construction is explained. There are three steps, as follows:

Step 1. Collection of fabric defect image data using vision cameras in the fabric inspection machine.

To collect the fabric defect image data, two single-channel machine vision cameras were installed on the fabric inspection machine, as shown in Figure 2. The fabric inspection machine was designed to inspect the fabric rolls while they were continuously rotating. The three fabric types were black denim, blue denim, and light blue fleece. All images were collected while moving the fabric at a speed of approximately 30 cm per second, like in the actual fabric inspection environment.

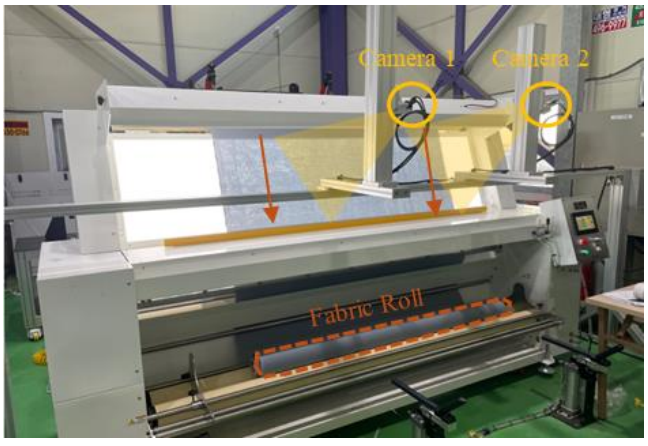


Figure 2. Image data collection using fabric inspection machine

Step 2. Construction of the original datasets and defect-augmented datasets through image patching and masking of defect areas. (Masking and labeling are conducted using CVAT, which is a well-known open-source annotation tool.) (OpenCV et al. n.d.)

The collected original images have dimensions of 3000 pixels in height and 4080 pixels in width. Due to their large size, it is inefficient for deep learning models to process them directly, necessitating resizing or patching. When resizing the original images directly, as shown in Figure 3, small defect areas may become significantly reduced and thus may not be detected. To avoid this issue, we performed patching in a grid of 3 (vertical) by 4 (horizontal) patches and resized each patch to 320x320 pixels. Subsequently, masking was applied to all patches containing defects to construct the dataset.

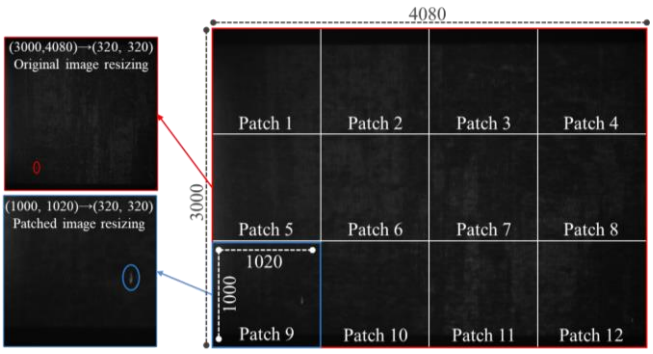


Figure 3. Image patching and resizing

Step 3. Design and construction of the defect detection model architecture for training.

This dataset was then divided into training, validation, and testing sets, as shown in Table 1. Both the training and validation datasets were composed entirely of patches containing defects. For the testing dataset, we included both patches with defects and patches without defects to assess the tendency for false positives in detecting defects on normal (non-defective) patches.

Table 1. The number of images of original patched dataset

Defect type	Without defect	Hole	Stain	Dyeing	Total
Training	0	468	109	71	648
Validation	0	465	95	95	655
Testing	6312	80	45	3	6440

Deep learning-based defect inspection commonly uses object detection and instance or semantic segmentation models. In this study, we selected a semantic segmentation model to reflect the characteristic of quantitatively calculating the area of defects during the quality assessment of fabric products.

We designed our own Unet++ architecture, which has recently demonstrated strong segmentation performance in the biomedical field, to build the model as depicted in Figure 4. The model specifically takes an image input size (320,320) and outputs masks (320,320,4) corresponding to non-defective regions and each defect. Additionally, the input image undergoes four rounds of down-sampling and up-sampling, with skip connections applied between all feature maps (Zhou et al. 2019). To ensure performance in both pixelwise classification and defect classification during model training, we employed the BCE Dice loss function, calculated as Eq. (1).

By using both BCE (Binary Cross-Entropy) and DICE loss functions, it is possible to accommodate diversity in loss calculations while leveraging the stability provided by BCE. In semantic segmentation tasks, x_n and y_n are both binary images (masks), representing the ground truth and the predicted mask, respectively.

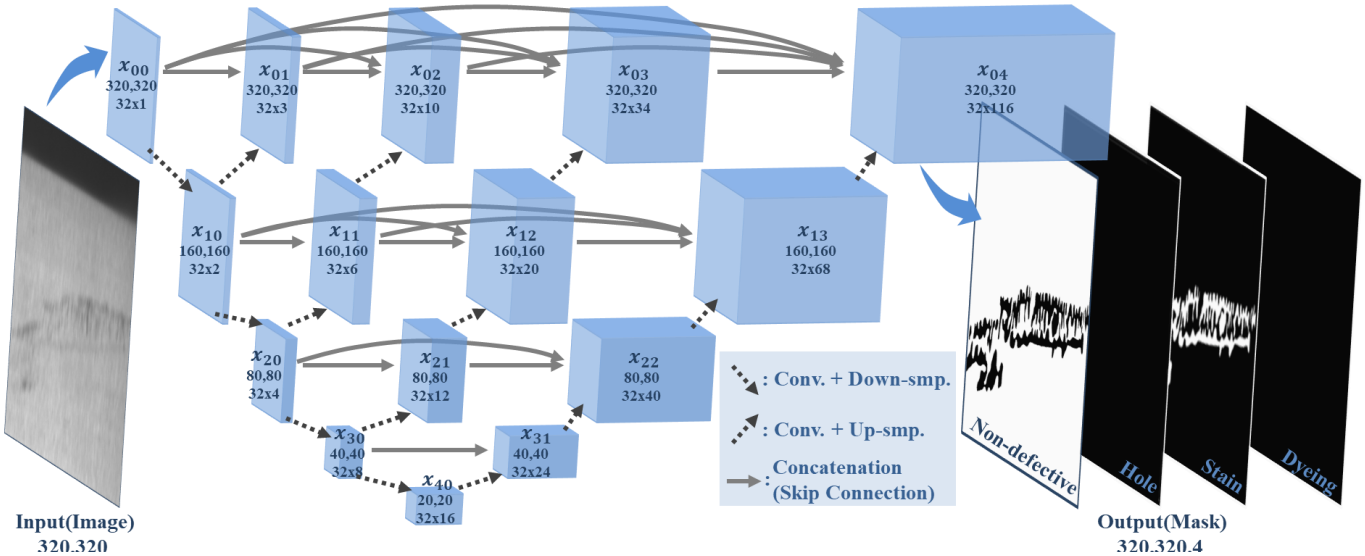


Figure 4. Architecture of fabric defect detection model.

$$\text{BCE Dice Loss} = \text{BCE loss} + 0.5 \times \text{Dice loss}$$

$$\text{BCE Loss} = -\sum(y_n \log \sigma(x_n) + (1 - y_n) \log(1 - \sigma(x_n))) \quad (1)$$

$$\text{Dice loss} = \sum\left(1 - \frac{2 \times y_n \times x_n}{x_n^2 + y_n^2 + 10^{-5}}\right)$$

2.2 Defect Data Augmentation

The number of images within a dataset to ensure the basic performance of CNN-based deep learning models can vary depending on several factors (Luo et al., 2018; Israel et al., 2021). However, models used in industrial applications often face limitations on the available data within their own manufacturing environments. Thus, transfer learning using pre-trained weights on custom datasets is a commonly used strategy to ensure robustness (Redmon et al. 2016). However, even in such cases, it is generally recommended to have at least 1000 images per class (Cho et al., 2015). Therefore, it is necessary to apply a proper data augmentation approach to ensure enhanced performance of the defect detection model.

In the dataset created by patching the collected original images in this study, the numbers of defect instances are highly imbalanced, as shown in Table 1. Additionally, the absolute number of images is insufficient. The most frequently observed defect in both the training and validation datasets is "hole," occurring more than four times as frequently as "stain" and "dyeing" defects combined. Even when combining the training and validation datasets, the total number of images does not exceed 1500. Therefore, image data augmentation is essential, and it is necessary to reflect the characteristics of each defect in the augmentation process. Stain defects exhibit various shapes and intensities within the defect areas, while dyeing defects primarily appear as broad horizontal stripes. Hole defects typically manifest as long, uniformly shaped vertical openings. Figure 5 shows examples of images depicting each type of defect.

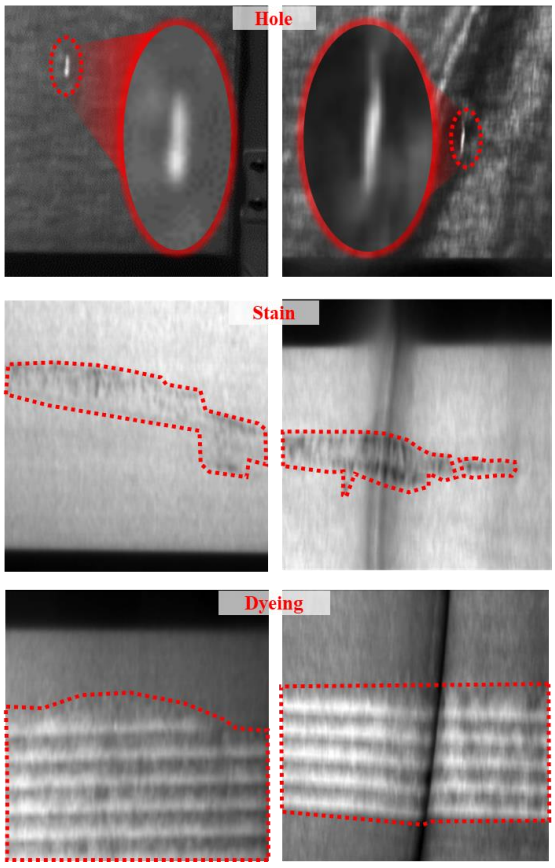


Figure 5. Example images of each type of defects.

Commonly used methods for image augmentation include geometric transformations (such as flipping, rotating, and affine/perspective transformations) and brightness adjustment (such as histogram equalization). These methods are advantageous because the operations applied to the

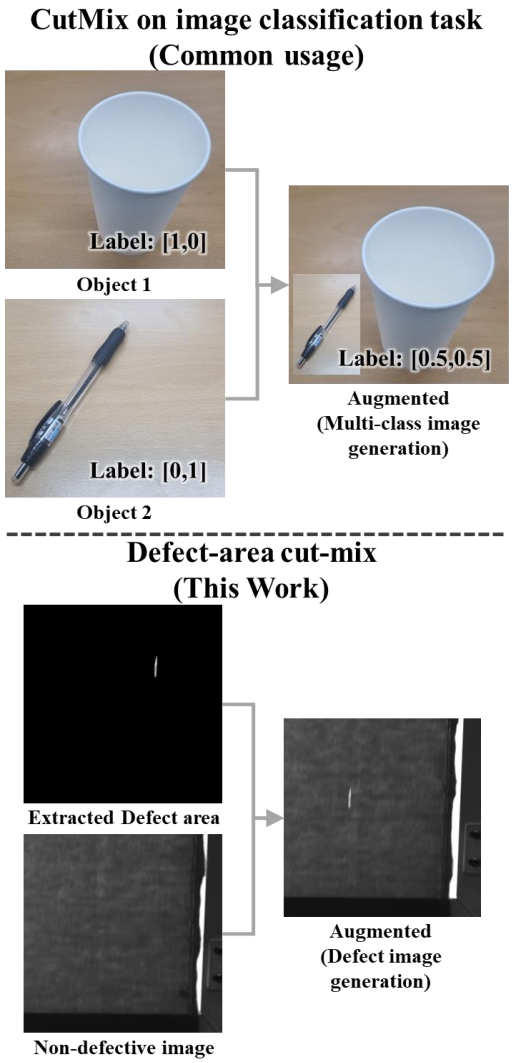


Figure 6. Comparison of conventional cut-mix and defect-area cut-mix.

images are simple, allowing for the rapid generation of new images. However, these methods alter the shape and brightness information of defects along with the surrounding background area, making it challenging to reflect cases where the model inputs unseen shapes of defects.

Therefore, in this study, to overcome the limitations of such data transformations, the cut-mix augmentation technique was applied specifically to defect areas. The cut-mix augmentation is a technique proposed to improve the generalization performance of image classification models by encouraging the model to better learn the local features of each class (Yun et al., 2019). Figure 6 illustrates the difference between the conventional cut-mix augmentation method and the approach used in this study, referred to as "defect-area cut-mix." While the original cut-mix augmentation mixes rectangular regions of different class images directly, in this study, only the defect areas were cut out and overlaid onto images without defects to create new images with defects. This approach allows for the assumption of situations where the shape and location of defects may change from the dataset's perspective.

– Extract defect areas using defect masks and images.

– Select images without defects and overlay defect areas directly or randomly, depending on the characteristics of the defects. Here, the characteristics of the defects refer to the area or shape of the defect areas.

Specifically, for hole defects (see Figure 7), where the defect size is relatively small, the defect areas can be randomly overlaid in portions of the fabric that are not occupied by the fabric. However, manual overlay position selection is required for stain and dyeing defects, where the defect regions' forms and areas vary considerably and may occupy a significant portion of the image.

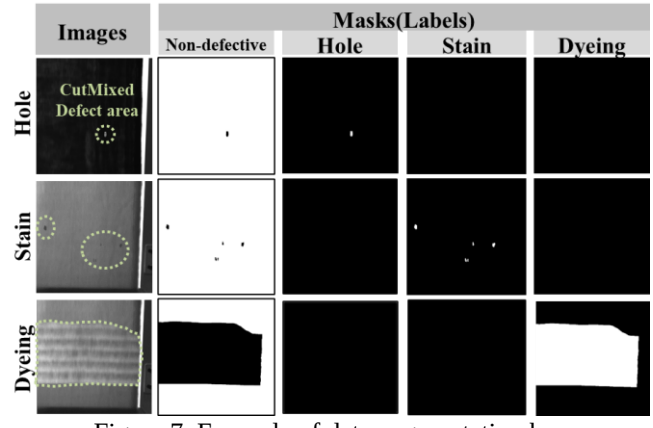


Figure 7. Example of data augmentation by defect-area cut-mix.

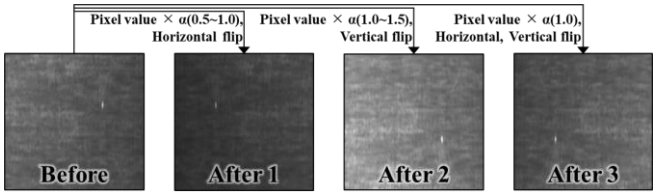


Figure 8. Example of data augmentation by geometric transformation and brightness adjustment.

After applying the defect-area cut-mix as described above, geometric and brightness adjustments were applied to the dataset. The following three image processing techniques were independently applied with a certain probability: brightness adjustment to reflect changes in external lighting conditions during image capture (random alpha values within the range of 0.5 to 1.5 multiplied by the entire image), horizontal flip, and vertical flip. Figure 8 illustrates examples of data augmentation.

To assess the performance difference caused by augmentation, the composition of the final datasets constructed by applying augmentations to the original patched dataset in Table 1 is presented in Table 2 and Table 3 (datasets B and C). To validate the effectiveness of defect-area cut-mix augmentation, augmentations depicted in Figure 7 without defect-area cut-mix were applied to dataset B. On the other hand, every augmentation containing geometric transformation, brightness adjustment, and defect-area cut-mix was applied to dataset C. All augmentations were applied only to the training dataset. However, for stain defects, it was empirically confirmed that applying the defect-area cut-mix technique was not effective due to the diverse shapes of the defects. Therefore, this augmentation was applied to the hole and dyeing defect only. Additionally, to avoid misdiagnosing defects on images without defects, images without defects and with temporary wrinkles (which are not considered defects) were added to the training and validation datasets. Compared to the original patched dataset, the augmented datasets partially alleviated the class imbalance between defect types.

Table 2. Composition of the augmented dataset B

Defect type	Without defect	Hole	Stain	Dyeing	Total
Training	930	549	491	186	2156
Validation	7200	465	95	95	7855
Testing	6312	80	45	3	6440

Table 3. Composition of the augmented dataset C

Defect type	Without defect	Hole	Stain	Dyeing	Total
Training	930	702	907	314	2853
Validation	7200	465	95	95	7855
Testing	6312	80	45	3	6440

3. PERFORMANCE EVALUATION

Through a series of steps, three datasets were constructed before and after augmentation (naturally, the original patched dataset is a subset of the augmented dataset). Each of the three different datasets was trained on our self-designed architecture of the Unet++ model, as depicted in Figure 3. For convenience in description, the model trained on the pre-augmented data will be referred to as Model A, the model trained on the augmented dataset without defect-area cut-mix will be referred to as Model B, and the one trained on the augmented dataset containing defect-area cut-mix will be referred to as Model C. By training a strict model and a less strict model independently, and then using an ensemble method, we were able to avoid incorrectly classifying defect-free fabric data as defective in Model B and Model C. To produce the segmentation results, the ensemble applied to Model B and Model C used soft voting on the pixel-wise classification probabilities of two distinct models.

The identical testing dataset was used to test Models A, B, and C. Masks for the defect and non-defective areas were inferred by the models. The mean Intersection over Union (IOU) score for each patch of each image was used to assess the performance metrics for defect identification, with a 0.7 threshold. This measure evaluates each model's ability to identify defects and distinguish defective regions from non-defective ones.

Furthermore, the confusion matrix depicted in Figure 8 was used to calculate the models' precision, recall, accuracy, and F1 score in order to assess their performance in defect identification. Note that false positives represent instances where non-defective regions were incorrectly predicted as defective, while false negatives represent instances where defects were present but not detected. These metrics provide insights into the classification performance of the models in both non-defective and defective regions. Classification among different types of defects showed a 100% accuracy for all models. This high accuracy can be attributed to the clear characteristics of each type of defect, as described earlier.

Overall performance metrics for Model A, B and C are presented in Table 4. Model C, trained on the augmented dataset with defect-area cut-mix, demonstrated superior performance across all metrics compared to Model A and Model B. Among the performance metrics for defect detection, Recall exhibited the smallest change, indicating that both models excelled at detecting actual defects with minimal variation.

The effects of geometric transformation and brightness adjustment augmentation, as well as additional training on defect-free images, were evident in the difference between models A and B. Comparing the confusion matrix of Model B with that of Model A shows that all metrics improved. Notably, the total number of false positives decreased significantly from 1252 to 181. This indicates a substantial

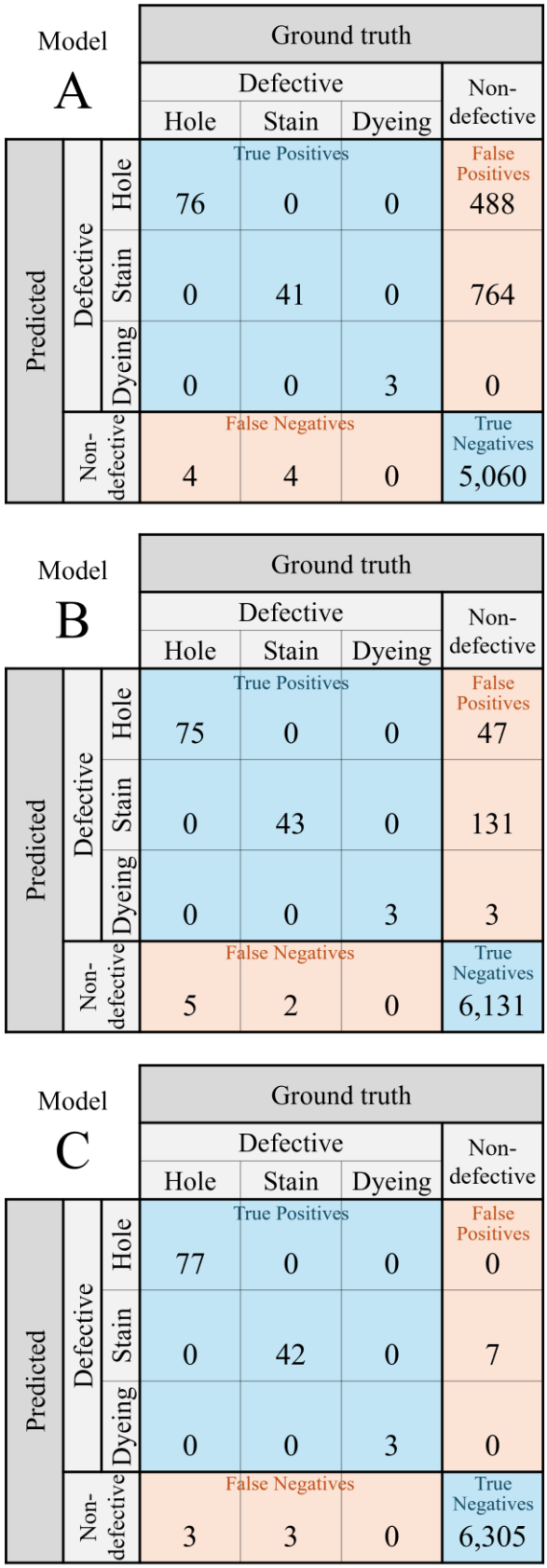


Figure 8. Confusion matrices of Model A, B and C

reduction in the model's tendency to incorrectly classify normal regions as defects. Moreover, the Mean IOU improved significantly from 0.375 to 0.704.

In the case of Model C, trained on a dataset where defect-area cut-mix was applied, the number of false positives decreased significantly compared to the other two models, with only 7 false positives in the Stain defect category. (It is interesting to note that the cut-mix was not applied to the Stain defect.) Furthermore, Model C achieved a more robust test result, with a mean IoU of 0.902, compared to the other two models.

In summary, the augmentation proposed in this study and ensemble techniques led to a significant improvement in defect detection performance compared to model without these enhancements. This demonstrates that the proposed defect-area cut-mix technique can enhance the robustness of deep learning models in image-based defect detection tasks.

Table 4. Metrics comparison of model A, B and C

Model	A	B	C
Mean IOU	0.375	0.704	0.902
Precision	0.086	0.401	0.946
Recall	0.938	0.945	0.953
F1 score	0.160	0.281	0.950
Accuracy	0.804	0.971	0.998

4. CONCLUSION

This study proposed the novel image data augmentation method, referred to as the defect-area cut-mix, for enhancing defect detection accuracy and robustness of the deep learning-based fabric inspection system. In the defect-area cut-mix method, the defect shapes that are the same as actual fabric defects (hole, stain and dyeing defect) were extracted using the masks, and they were added to the non-defective fabric images for an augmentation. To demonstrate the effectiveness of the proposed defect-area cut-mix augmentation method, three data sets were prepared, such as the original dataset with augmentation (dataset A), that with conventional geometrical augmentation and brightness adjustments (dataset B), and that with defect-area cut-mix, geometrical augmentation, and brightness adjustments (dataset C). Then, the ensemble approach combining the deep-learning models with high and low sensitivity was applied to datasets B and C. Finally, it was found that the fabric defect diagnosis model with the dataset C and ensemble approach showed the best performance in terms of mean IOU, precision, recall, F1 score, and accuracy.

ACKNOWLEDGEMENT

This work was supported by the Technology development Program (S3303060, 2022R1A2C3012900) funded by the Ministry of SMEs and Startups (MSS, Korea) and the National Research Foundation of Korea (NRF) grant funded by the Korea government (MSIT).

REFERENCES

Jung, W. K., Kim, D. R., Lee, H., Lee, T. H., Yang, I., Youn, B. D., ... & Ahn, S. H. (2021). Appropriate smart factory for SMEs: concept, application and perspective. *International Journal of Precision Engineering and Manufacturing*, *22*, 201-215. DOI 10.1007/s12541-020-00445-2

Hanbay, K., Talu, M. F., & Özgüven, Ö. F. (2016). Fabric defect detection systems and methods—A systematic literature review. *Optik, 127(24)*, 11960-11973. DOI 10.1016/j.ijleo.2016.09.110

Bhatt, P. M., Malhan, R. K., Rajendran, P., Shah, B. C., Thakar, S., Yoon, Y. J., & Gupta, S. K. (2021). Image-based surface defect detection using deep learning: A review. *Journal of Computing and Information Science in Engineering, 21(4)*, 040801. DOI 10.1115/1.4049535

Russakovsky, O., Deng, J., Su, H., Krause, J., Satheesh, S., Ma, S., ... & Fei-Fei, L. (2015). ImageNet large scale visual recognition challenge. *International journal of computer vision*, *115*, 211-252. DOI 10.1007/s11263-015-0816-y

Tong, L., Wong, W. K., & Kwong, C. K. (2016). Differential evolution-based optimal Gabor filter model for fabric inspection. *Neurocomputing*, *173*, 1386-1401. DOI 10.1016/j.neucom.2015.09.011

Zhou, Y. (2019). Research on image-based automatic wafer surface defect detection algorithm. *Journal of Image and Graphics*, *7*(1), 26-31. DOI 10.18178/joig.7.1.26-31

Tabernik, D., Šela, S., Skvarč, J., & Skočaj, D. (2020). Segmentation-based deep-learning approach for surface-defect detection. *Journal of Intelligent Manufacturing*, *31*(3), 759-776. DOI 10.1007/s10845-019-01476-x

Ho, C. C., Chou, W. C., & Su, E. (2021). Deep convolutional neural network optimization for defect detection in fabric inspection. *Sensors*, *21*(21), 7074. DOI 10.3390/s21217074

OpenCV. (n.d.). CVAT. Retrieved [2024.05.10], from https://www.cvat.ai/.

Zhou, Z., Siddiquee, M. M. R., Tajbakhsh, N., & Liang, J. (2019). Unet++: Redesigning skip connections to exploit multiscale features in image segmentation. *IEEE transactions on medical imaging*, *39*(6), 1856-1867. DOI 10.1109/TMI.2019.2959609

Luo, C., Li, X., Wang, L., He, J., Li, D., & Zhou, J. (2018, November). How does the data set affect cnn-based image classification performance?. In *2018 5th international conference on systems and informatics (ICSAI)* (pp. 361-366). IEEE. DOI 10.1109/ICSAI.2018.8599448

Israel, I. M., Israel, S. A., & Irvine, J. M. (2021, October). Factors influencing CNN performance. In 2021 IEEE Applied Imagery Pattern Recognition Workshop (AIPR) (pp. 1-4). IEEE.
DOI 10.1109/AIPR52630.2021.9762112

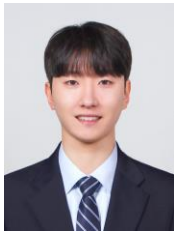
Redmon, J., Divvala, S., Girshick, R., & Farhadi, A. (2016). You only look once: Unified, real-time object detection. *In Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 779-788).

Cho, J., Lee, K., Shin, E., Choy, G., & Do, S. (2015). How much data is needed to train a medical image deep learning system to achieve necessary high accuracy? *arXiv preprint arXiv:1511.06348*. https://arxiv.org/abs/1511.06348
DOI 10.48550/arXiv.1511.06348

Yun, S., Han, D., Oh, S. J., Chun, S., Choe, J., & Yoo, Y. (2019). Cutmix: Regularization strategy to train strong classifiers with localizable features. In *Proceedings of the IEEE/CVF international conference on computer vision* (pp. 6023-6032).

BIOGRAPHIES

Youngwoon Choi is now Ph.D. candidate in the Sustainable Design and Manufacturing Laboratory, Department of Mechanical Engineering, Sungkyunkwan University. He obtained his bachelor's degree (Mechanical Engineering) from Sungkyunkwan University in 2020 and completed his master's degree in 2022. His research interest is vision AI-based prognostics and health management for smart manufacturing.



Hyunseok Lee is now master candidate in the Sustainable Design and Manufacturing Laboratory, Department of Mechanical Engineering, Sungkyunkwan University. He obtained his bachelor's degree (Mechanical Engineering) from Sungkyunkwan University in 2023. His research interest is vision AI-based prognostics and health management for smart manufacturing.

Sang Won Lee is now professor in the school of Mechanical Engineering, Sungkyunkwan University. He obtained his bachelor's degree in 1995 and master's degree in 1997 (Mechanical Design and Production Engineering) from Seoul National University. He obtained his Ph.D. degree (Mechanical Engineering) from University of Michigan in 2004. His research interest includes prognostics and health management (PHM), cyber-physical system (CPS), additive manufacturing, and data-driven design.