MEKATRONIKA
e-ISSN: 2637-0883
VOL. 2, ISSUE 1, 16 – 22
DOI: https://doi.org/10.15282/mekatronika.v2i1.6704

ORIGINAL ARTICLE

Effect of Image Compression using Fast Fourier Transformation and Discrete Wavelet Transformation on Transfer Learning Wafer Defect Image Classification

Jessnor Arif Mat-Jizat[1*], Ahmad Fakhri Ab. Nasir[2,3], Anwar P.P. Abdul Majeed[1,3] Edmund Yuen[4]

[1]Faculty of Manufacturing & Mechatronic Engineering Technology (FTKPM), Universiti Malaysia Pahang, 26600 Pahang, Malaysia.
[2]Faculty of Computing, Universiti Malaysia Pahang, 26600 Pahang, Malaysia.
[3]Centre for Software Development & Integrated Computing, Universiti Malaysia Pahang, 26600 Pahang, Malaysia
[4]Ideal Vision Integration Sdn Bhd, 02-25, Level 2, Setia Spice Canopy, Jln Tun Dr Awang, 11900 Bayan Lepas, Penang, Malaysia.

ABSTRACT – Automated inspection machines for wafer defects usually captured thousands of images on a large scale to preserve the detail of defect features. However, most transfer learning architecture requires smaller images as input images. Thus, proper compression is required to preserve the defect features whilst maintaining an acceptable classification accuracy. This paper reports on the effect of image compression using Fast Fourier Transformation and Discrete Wavelet Transformation on transfer learning wafer defect image classification. A total of 500 images with 5 classes with 4 defect classes and 1 non-defect class were split to 60:20:20 ratio for training, validating and testing using InceptionV3 and Logistic Regression classifier. However, the input images were compressed using Fast Fourier Transformation and Discrete Wavelet Transformation using 4 level decomposition and Debauchies 4 wavelet family. The images were compressed by 50%, 75%, 90%, 95%, and 99%. As a result, the Fast Fourier Transformation compression show an increase from 89% to 94% in classification accuracy up to 95% compression, while Discrete Wavelet Transformation shows consistent classification accuracy throughout albeit diminishing image quality. From the experiment, it can be concluded that FFT and DWT image compression can be a reliable method for image compression for grayscale image classification as the image memory space drop 56.1% while classification accuracy increased by 5.6% with 95% FFT compression and memory space drop 55.6% while classification accuracy increased 2.2% with 50% DWT compression.

ARTICLE HISTORY
Received: 25th May 2020
Revised: 2nd June 2020
Accepted: 5th June 2020

KEYWORDS
Disc. Wavelet Transformation
Fast Fourier Transformation
InceptionV3
Logistic Regression
Wafer Defect

INTRODUCTION

An automated optical inspection machine is key for a semiconductor company to remain competitive. This stem from manual human inspection accuracy can drop to between 70%-85% after 15 months of training due to several factors such as process advancement, increasing complexities due to product evolution and mental fatigue [1]. Thus, an automated optical inspection can offer an improvement in terms of inspection accuracy and less manpower usage. The core for an automated optical inspection machine is usually a machine learning model capable of learning various defect features and accurately predicting the outcome. However, adopting this automated inspection machine requires thousands of images to be stored for machine learning training purposes. Images captured by an automated inspection machine, usually on a large scale, to preserve several defects features. Thus, proper compression is required to preserve the defect features whilst maintaining an acceptable classification accuracy.

Transfer learning is a machine learning method that harnesses pre-trained models as a foundation to develop machine learning models, especially in computer vision and natural language processing. It reused the already developed model as a starting point to develop a model for another task, mostly on features extraction from images or signals. This simplifies the steps taken for developing a new model to suit various machine learning tasks. A myriad of transfer learning models had been used in various application fields such as sports[2], medical device [3], aquaculture [4], and agriculture [5].

RELATED WORK

Saqlain et al. [6] introduce a voting ensemble technique to increase classification accuracy for wafer pattern images. They were extracting features from the raw images and run four different classifiers. Density-based, geometric-based, and radon-based were among the features extracted. The voting ensemble allows the best component from each classifier to influence the final classification accuracy. Other researchers such as Ruifang et al. [7] classified 11 defect markers by using ZF-Net as features extraction or Y. S. Jeong [8] who uses Dynamic Time Warping and Support Vector Machine for feature extraction wafer defect classifier. Most of these works show that features were extracted from the RGB image with high quality to preserve the details defect markers. However, in this paper, the author wants to explore the classifiers with diminishing image quality. One of the diminishing qualities is from RGB to Grayscale conversion.

*CORRESPONDING AUTHOR | Jessnor Arif Mat Jizat | jessnor@ump.edu.my

Fast Fourier transformation (FFT) is a technique that transforms a spatial domain signal into a frequency domain. As an image can be considered a 2D signal in the spatial domain, transforming an image into a frequency domain can be done. Processing an image in a frequency domain enables a faster computational work time for image enhancement such as salient edges, or shadow smoothing compared to pixel to pixel processing in the time domain. After the transformation into the frequency domain, the transformed image consists of both quantized low and high-frequency coefficients [9]. These frequency coefficients have allowed for image compression as several quantized high-frequency coefficients had a value of near zero. By discarding this near-zero coefficient, the image had been compressed after the image reconstruction. However, the image reconstruction should retain most of the image quality as shown in the A. T. G and K. Vijayalakshmi [9] with their medical imaging or V. Cheepurupalli et al. [10] with their generic images.

An image that is compressed using Discrete Wavelet Transformation (DWT) in general will decompose the images into approximation, vertical, horizontal and diagonal detail components. The approximation detail component will display an approximation of the image at half of the image resolution, while the vertical, horizontal and diagonal detail component will display prominent vertical, horizontal and diagonal components of the image. This complete the first level of the transformation. This level can be repeated multiple times using the approximation detail component of each level. Figure 1 shows an example of 2 level DWT image decomposition. A wavelet is used to decompose the image into its components. They are several prominent wavelet families such as Haar, Daubechies, Symlets, Coiflets, and Biorthogonal.

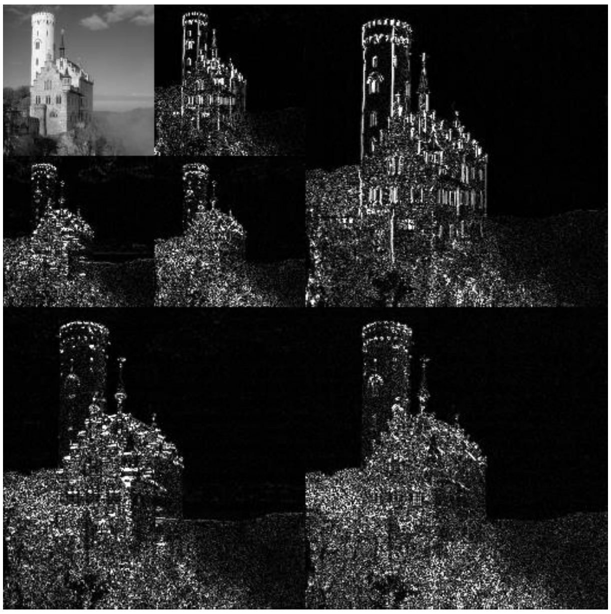


Figure 1. 2 level DWT image decomposition. From top left to right: Level 2 approximation, Level 2 vertical, Level 1 vertical, Level 2 horizontal, Level 2 diagonal, Level 1 horizontal, Level 1 diagonal. [Source:https://upload.wikimedia.org/wikipedia/commons/thumb/e/e0/Jpeg2000_2-level_wavelet_transform-lichtenstein.png/300px-Jpeg2000_2-level_wavelet_transform-lichtenstein.png, Retrieved: 22/7/2021 18:00]

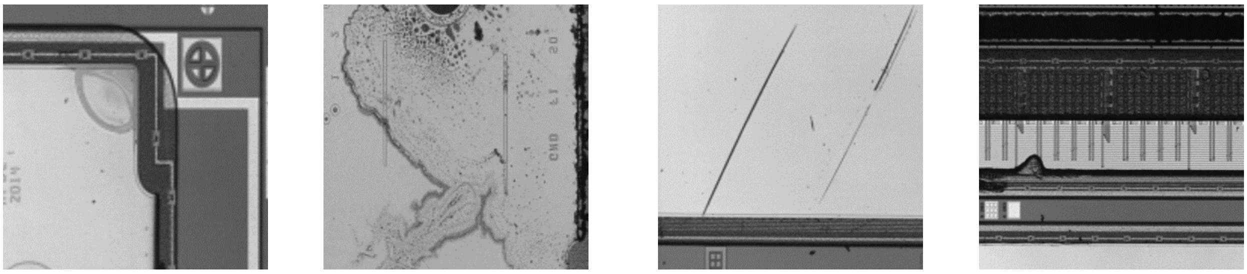
Lahiru D. Chamain and Zhi Ding [11] show that by using DWT-centric compression, they can achieve a more accurate classification of JPEG2000 images and faster and more accurate representation over limited bandwidth channel transfer. While Kutlu H and Avcı E.[12] using a Convolution Neural Network (CNN) -Discrete Wavelet Transform-Long-Short Term Memory (LSTM) pipeline to classify brain and liver tumours from Computer Tomography images. The DWT was used to compressed the features extracted by the CNN while the LSTM was the classifier. They compared their pipeline performance with the k-Nearest Neighbour and Support Vector Machine classifier. J. Sharma et al. [13] studies the effect of deeper level DWT decomposition on the classification accuracy. They were using satellite images to classify land used or land cover using Minimum Distance Classifier.

However, rather than evaluate the quality of the reconstructed images either from FFT or from DWT, this paper evaluated the effect of the compressed image as an input image for transfer learning models and image classifiers. This may be seen as a robustness challenge to the transfer learning model against the diminishing quality of the input image. How far can the image be compressed to see a significant classification accuracy drop?

METHODOLOGY

Dataset and Image Preprocessing.

The dataset, a total of 500 images with known defect features, is acquired from Idealvision Sdn Bhd, a machine vision company, using their own industrial machine vision platform Jaeger. There are four classes of defect features together with a set of the non-defect class involved in this study. The five classes are Burnt Mark (BM), Missing (MS), Scratch (SC), Contamination (CT) and Non-defect (GD). Figure 2 shows the defect features for each class. Each class consist of 100 images that were split into 60:20:20 ratio for Training, Validation and Testing. The images in this study were neither repeated nor augmented. different images were used for each Training, Validation and Testing category. All of the images were resized from [4096, 3072, 3] to [299, 299, 1]. The images were converted to grayscale using the average method where the value of red, green and blue components was added and divided by 3. The 299 x 299 pixels is chosen in preparation for the input requirement of the InceptionV3.



a) b) c) d)

Figure 2. Example of defect features from dataset a) burnt mark, b) contamination, c) scratch, d) missing.

Image Compression

In this study, there is 2 type of image compression used. Fast Fourier Transformation (FFT) and Discrete Wavelet Transformation (DWT). For both transformations, the images were transformed and their respective coefficient was converted into vectors and sorted from the largest to the smallest. From these sorted coefficient vectors, a threshold value is calculated according to the compression needed e.g 50%. Thus, from this threshold value, a mask was created whereby anything higher than the threshold equates to 1 whilst anything below this threshold value is equated to 0. Then, the mask was dot product with the transformed image to only keep the value that is above the threshold. Then the transformed image was inverse to the original images with the required compression. This technique is adopted from [14]. The compression percentage under investigation is 50%, 75%, 90%, 95%, 99% compression. The diminished quality of the images is shown in Figure 3.

In the case of FFT, the 2D FFT was used as images had always had pixels rows and pixels columns. The FFT runs in rows first followed by columns. While in DWT, Daubechies 4(db4) wavelets were used at four levels of transformation before the compression process was applied.

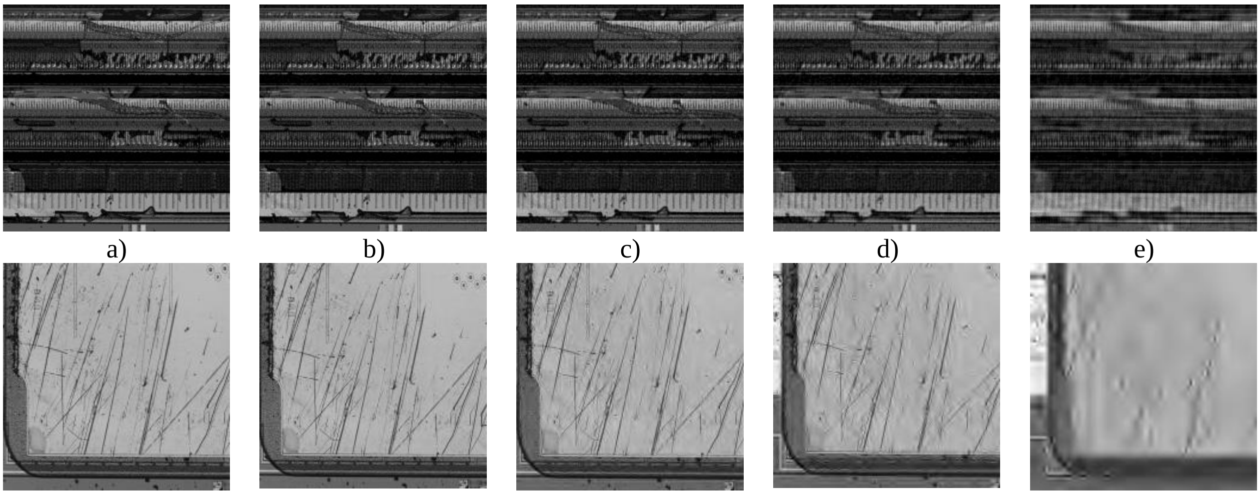


Figure 3. Examples of diminishing quality of the images at a) 50% FFT compression, b) 75% FFT compression, c) 90% FFT compression, d) 95% FFT compression, e) 99% FFT compression, f) 50% DWT compression, g) 75% DWT compression, h) 90% DWT compression, i) 95% DWT compression, j) 99% DWT compression,

Transfer Learning and Image Classification

In this study features extraction is using InceptionV3 with imageNet transfer learning model and Logistic Regression as image classifier model. This particular model is chosen due to its ubiquitous usage in transfer learning methods. However, in this study, the hyperparameter for the Logistic Regression is tuned in search of the best parameters to build the model. The tuned parameter was the solver, the regularization strength, and the regularization. “Liblinear” and “lbfgs” are the two options for the solver, where it has different algorithms to optimize the solution. The regularization consists of “Lasso Regression” and “Ridge Regression” and the regularization strength had 30 different numbers between 0.002 to 1000 in which smaller values specify stronger regularization. The best parameters from the hyperparameters tuning are Ridge Regression regularization with 60 regularization strength and “lbfgs” solver. The image training of the image was subjected to 5-fold cross-validation to avoid overfitting. This model, hence, yields 89% classification accuracy in the uncompressed image.

Primarily, the classification accuracy, precision and recall of testing classification results were used as evaluation criteria where classification accuracy indicates the capability of the classifier to accurately predict the images according to its class. Meanwhile, classification precision deals with the proportion of the correctly predicted images and the predicted images in their class. On the other hand, classification recall deals with the proportion of the correctly predicted images and the actual images in its class.

Hardware and Software

The training and testing were conducted using a desktop with Intel(R) Core(TM) i9-10900KF CPU @ 3.70GHz 3.70 GHz processor with 32.0 GB DDR3 RAM and NVIDIA GeForce RTX 2080 Ti graphic card. It was run using Spyder Anaconda, a python programming software using Scikit-learn, Keras and Tensorflow libraries.

Figure 4 shows an overview of the methodology process.

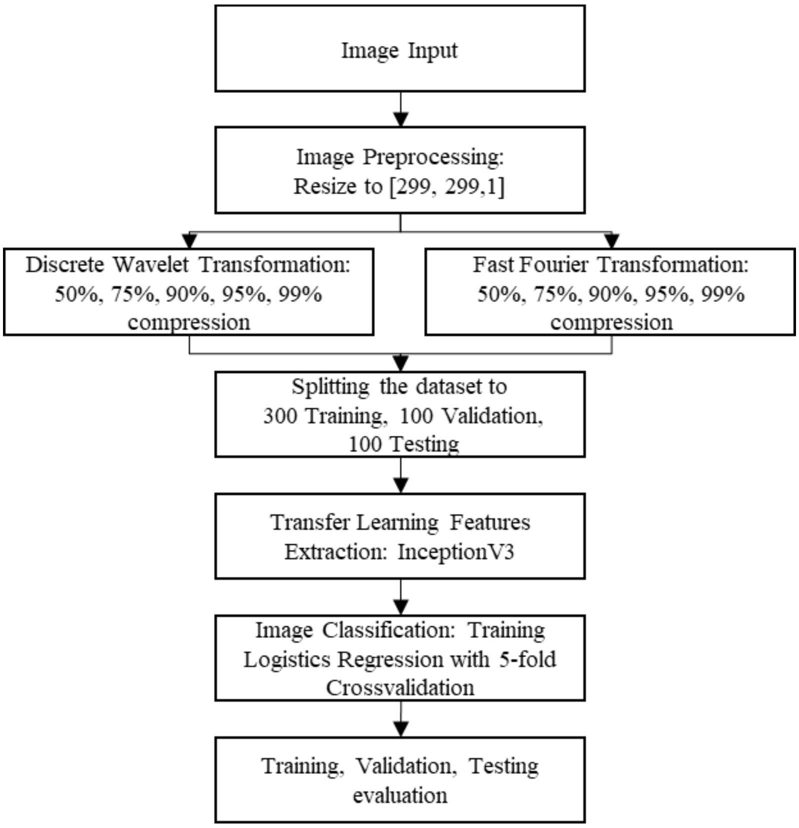


Figure 4. Workflow diagram of the methodology involved.

RESULTS AND DISCUSSION

Figure 5 shows the comparison of the classification accuracy performance for training, validation and testing of the FFT image compression as the compression gets higher from uncompressed image to 99% image compression. From the figure, an increasing trend can be noted for the testing classification accuracy. From 89% for the uncompressed image to 90%, 90%, 92% and 94% for FFT 50%, FFT 75%, FFT 90%, FFT 95% compression respectively with a drop to 86% for FFT 99% compression.

This trend indeed shows the robustness of the InceptionV3 transfer learning model paired with a Logistic Regression classifier. Even though the image had been compressed up to 95% of the original image, the model can yield a better accuracy from the original image. Furthermore, FFT is a known tool for denoising an image as it can manipulate and isolate certain frequency bands. This may have the defect feature become more prominent given the InceptionV3 transfer learning model hence the higher classification accuracy at 95% compression. A compression further than that causes more blurring on the original images, which diminish the defect features. Another advantage as compressed images requires less memory space, where in this case, the memory space required to store FFT95% is on average 56.1% less than the original grayscale image (Average memory space: FFT95% 79 kB, Uncompressed 180kb).

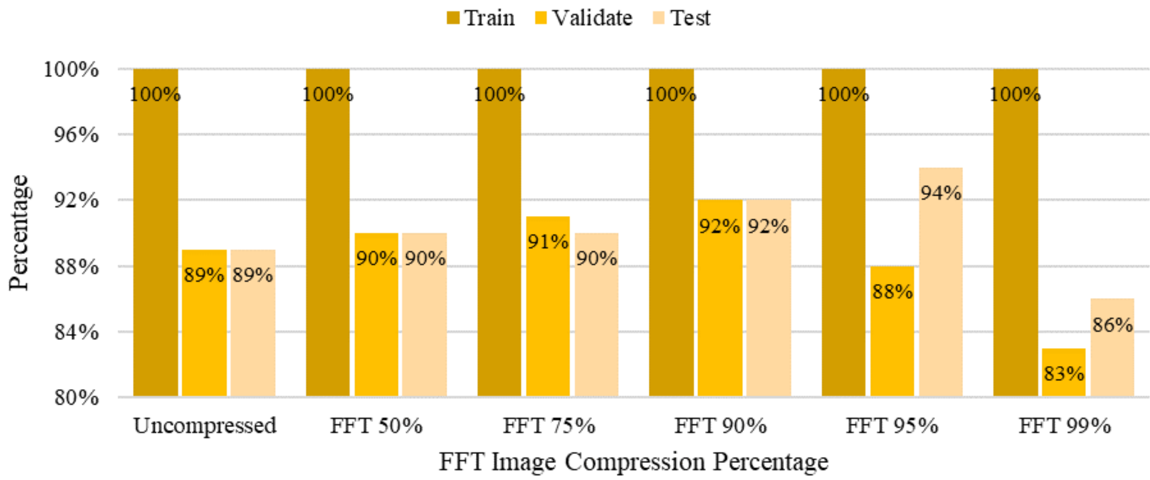


Figure 5. Classification accuracy performance for training, validation and testing of the FFT image compression

Furthermore, by referring to Table 1, class analysis can be done. Overall, all classes had precision and recall over 85%.except for Burnt mark class precision. Taking an average for precision and recall of all classes reveals that uncompressed images score 90% precision and 89% recall whilst FFT95% score 94% for both precision and recall. This shows that the classification accuracy is relevant with minimum or negligible bias. Burnt Mark class performed the worst with 75% precision and 90% recall with uncompressed images and 83% precision, and 95% recall with FFT95%. This may be caused by the image being less distinct from other classes.

Table 1. Comparison of Uncompressed and FFT95% Precision and Recall

Class	Uncompressed		FFT 95%	
	Precision	Recall	Precision	Recall
Burnt Mark	75%	90%	83%	95%
Missing	100%	85%	100%	95%
Scratch	100%	85%	100%	95%
Contamination	87%	100%	95%	100%
Good	89%	85%	94%	85%

Figure 6 shows the comparison of the classification accuracy performance for training, validation and testing of the DWT image compression as the compression gets higher from uncompressed image to 1% image compression. In terms of testing classification accuracy, the percentage seems to be neither increasing nor decreasing at about 90% despite higher image compression. The best classification accuracy is at 50% compression with 91% accuracy. Others are 89%, 90%, 90% 90% at DWT 75%, DWT 90%, DWT 95% DWT 99% compression respectively.

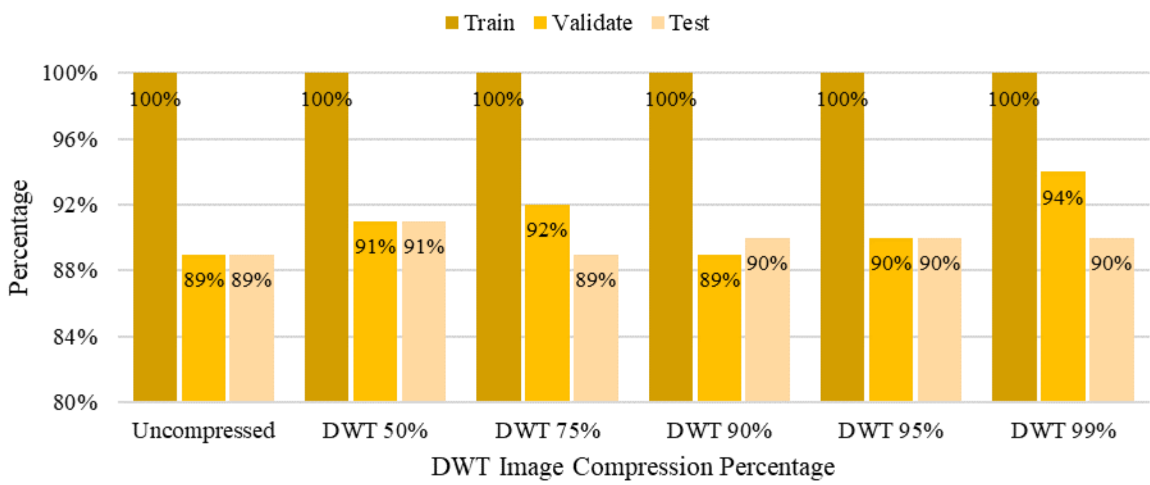


Figure 6. Classification accuracy performance for training, validation and testing of the DWT image compression

The robustness of the transfer learning and logistic regression model is again demonstrated here, albeit the testing accuracy is maintained at about 90% accuracy as fewer image data are present in the higher compression images. As in terms of memory space, DWT image compression offers better memory space saving compared to FFT image compression as they require on average of 80kB, 73kB, 72kB, 66kB, 44kB memory space for DWT 50%, DWT 75%, DWT 90%, DWT 95% DWT 99% compression respectively compared to 93kB, 91kB, 84kB 79kB 70kB memory space for FFT 50%, FFT 75%, FFT 90%, FFT 95% FFT 99% compression respectively. The lowest average memory space required is 44kB but still can yield more than 90% classification accuracy. However, for the best classification accuracy performance, the memory space saving is on average 55.8% less than the uncompressed image.

Table 2. Comparison of Uncompressed and DWT50% Precision and Recall

Class	Uncompressed		DWT 50%	
	Precision	Recall	Precision	Recall
Burnt Mark	75%	90%	78%	90%
Missing	100%	85%	100%	90%
Scratch	100%	85%	100%	90%
Contamination	87%	100%	91%	100%
Good	89%	85%	89%	85%

Consequently, class analysis of precision and recall is required to understand the bias. Table 2 are comparing the uncompressed image's precision and recall with the best performance in the DWT compression. From the table, overall the precision and recall percentage for both is above 85% except for uncompressed and DWT 50% compression precision. Both precision and recall are improving across all classes except for the good class. This improvement shows that DWT image compression can remove some bias from the data especially in the defect classes. Focusing on good class, the precision percentage is higher from recall. This can be interpreted as that the good class had a minor bias toward false negatives. In the case of wafer defect, a minor false negative where an actual good wafer was classified as bad wafer bias can be preferable compared to a false positive where a bad wafer is classified as good which may affect the product function down the manufacturing line.

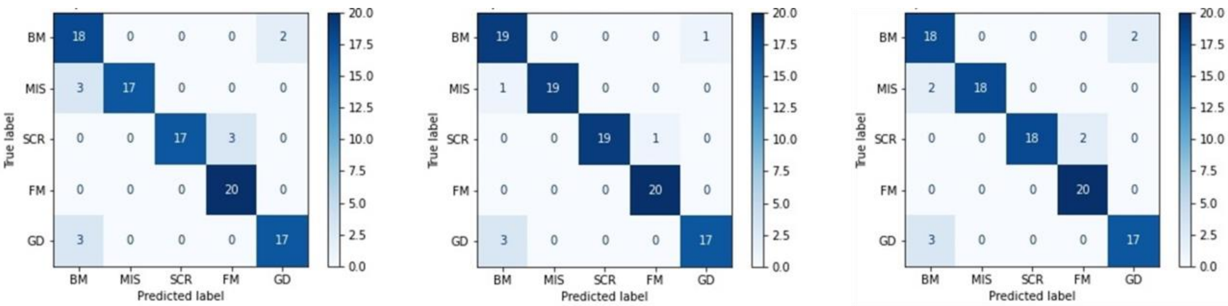


Figure 7. Confusion matrix for testing a) Uncompressed. b) FFT95%. c) DWT 50%
(BM = burnt mark, FM = contamination, SCR = scratch, MIS= missing GD = Non-defect.)

Looking at the confusion matrices of all best classification accuracy best performance shown in Figure 7. Confusion matrix for testing a) Uncompressed. b) FFT95%. c) DWT 50%Figure 7, we can see that the misclassification occurs in Burnt Mark, Missing, Scratch and Non-Defect classes. While all images in the contamination class had been correctly identified. In the uncompressed confusion matrix, two images from the burnt mark class were misclassified as a non-defect, three images from the missing class were misclassified as burnt mark class, three images from the scratch class were misclassified as contamination class and three images from the non-defect class were misclassified as burnt mark class. Both FFT and DWT increase the performance of the Missing and Scratch classes with FFT also increase the performance of the Burnt Mark class hence the increase in classification accuracy.

CONCLUSION

This paper reports on the effect of image compression using fast Fourier transformation and discrete wavelet transformation on transfer learning wafer defect image classification. From the experiment reported, fast Fourier transformation may be the best choice to classify wafer defects in grayscale as it can improve the accuracy as well as lower the memory space required to store the images with 95% FFT compression the image memory space drop 56.1% while classification accuracy increased by 5.6%. However, the discrete wavelet transformation too can provide a lower memory space requirement without compromising the classification accuracy as memory space drop 55.6% while classification accuracy increased 2.2% with 50% compression. This memory saving can be useful if ten of thousands of images need to be stored for the classification process. The experiment also shows the robustness of Inceptionv3 transfer learning paired with a Logistic Regression classifier to handle truncated image data whilst delivering more than 90% classification accuracy.

ACKNOWLEDGEMENT

The authors would like to thank IdealVision Sdn Bhd for providing the image dataset to make this evaluation possible as well as Universiti Malaysia Pahang for funding the study via UIC200815 and RDU202404.

REFERENCES

[1] E. Tuv, M. Guven, P. Ennis, and D. H. L. Lee, "IT@Intel White Paper: Faster, More Accurate Defect Classification Using Machine Learning," 2018. [Online]. Available: https://www.intel.com/content/dam/www/public/us/en/documents/best-practices/faster-more-accurate-defect-classification-using-machine-vision-paper.pdf.

[2] M. N. A. Shapiee *et al.*, "The Classification of Skateboarding Trick Manoeuvres Through the Integration of Image Processing Techniques and Machine Learning," in *Lecture Notes in Electrical Engineering*, 2020, vol. 632, doi: 10.1007/978-981-15-2317-5_29.

[3] J. L. M. Kumar *et al.*, "An Evaluation of Different Fast Fourier Transform-Transfer Learning Pipelines for the Classification of Wink-based EEG Signals," *Mekatronika*, vol. 2, no. 1, pp. 1–7, 2020.

[4] I. S. Isa, N. N. Norzrin, S. N. Sulaiman, N. A. Hamzaid, and M. I. F. Maruzuki, "CNN Transfer Learning of Shrimp Detection for Underwater Vision System," 2020, doi: 10.1109/ICITAMEE50454.2020.9398474.

[5] S. S. Rautaray, M. Pandey, M. K. Gourisaria, R. Sharma, and S. Das, "Paddy Crop Disease Prediction- A Transfer Learning Technique," *Int. J. Recent Technol. Eng.*, vol. 8, no. 6, 2020, doi: 10.35940/ijrte.f7782.038620.

[6] M. Saqlain, B. Jargalsaikhan, and J. Y. Lee, "A voting ensemble classifier for wafer map defect patterns identification in semiconductor manufacturing," *IEEE Trans. Semicond. Manuf.*, 2019, doi: 10.1109/TSM.2019.2904306.

[7] R. Ye, C. S. Pan, M. Chang, and Q. Yu, "Intelligent defect classification system based on deep learning," *Adv. Mech. Eng.*, 2018, doi: 10.1177/1687814018766682.

[8] Y. S. Jeong, "Semiconductor wafer defect classification using support vector machine with weighted dynamic time warping kernel function," *Ind. Eng. Manag. Syst.*, 2017, doi: 10.7232/iems.2017.16.3.420.

[9] A. T. G and K. Vijayalakshmi, "FFT Based Compression approach for Medical Images," *Int. J. Appl. Eng. Res.*, vol. 13, no. 6, pp. 3550–3567, 2018, [Online]. Available: https://www.ripublication.com/ijaer18/ijaerv13n6_54.pdf.

[10] V. Cheepurupalli, S. Tubbs, K. Boykin, and N. Naheed, "Comparison of SVD and FFT in image compression," 2016, doi: 10.1109/CSCI.2015.56.

[11] L. D. Chamain and Z. DIng, "Improving Deep Learning Classification of JPEG2000 Images over Bandlimited Networks," *ICASSP, IEEE Int. Conf. Acoust. Speech Signal Process. - Proc.*, vol. 2020-May, pp. 4062–4066, 2020, doi: 10.1109/ICASSP40776.2020.9053778.

[12] H. Kutlu and E. Avcı, "A Novel Method for Classifying Liver and Brain Tumors Using Convolutional Neural Networks, Discrete Wavelet Transform and Long Short-Term Memory Networks," *Sensors (Basel).*, vol. 19, no. 9, 2019, doi: 10.3390/s19091992.

[13] J. Sharma, R. Prasad, V. N. Mishra, V. P. Yadav, and R. Bala, "Land Use and Land Cover Classification of Multispectral Landsat-8 Satellite Imagery Using Discrete Wavelet Transform," *Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci.*, vol. XLII–5, no. November, pp. 703–706, 2018, doi: 10.5194/isprs-archives-xlii-5-703-2018.

[14] S. L. Brunton and J. N. Kutz, *Data-Driven Science and Engineering*. 2019.