International Journal of Multidisciplinary Research and Literature
IJOMRAL
Vol. 3, No. 2, March 2024 pp. 172-183
Journal Page is available at http://ijomral.esc-id.org/index.php/home

MODELING PISTON DAMAGE DETECTION USING A CONVOLUTIONAL NEURAL NETWORK BASED ON DIGITAL IMAGE

Lega putri utami[1*], Desmarita Leni[2]
[1]Mechanical Engineering, Faculty of Engineering, Universitas Andalas, Indonesia
[2]Mechanical Engineering, Faculty of Engineering, Universitas Muhammadiyah Sumatera Barat, Indonesia
Email: *lega@eng.unand.ac.id[1]*

Abstract

Product inspection is a crucial component of product quality control, aiming to evaluate and ensure that products meet predefined standards. In this research, the modelling of piston damage detection is conducted using a Convolutional Neural Network (CNN). The dataset employed consists of images of pistons categorized into three groups: Defected1, Defected2, and Normal. Two hundred eighty-five images are utilized as training data, with the data distribution percentages for Defected1, Defected2, and Normal being 30.9%, 34.4%, and 34.7%, respectively. The model is validated using newly generated data through augmentation techniques, resulting in 60 images. The CNN model uses a sequential Keras architecture comprising convolutional layers, pooling layers, fully connected layers, and softmax activation. The Adam optimizer with a learning rate 0.0001 is employed for model training, with validation using a 5-fold cross-validation. The model is evaluated using the Loss, Accuracy, and Confusion Matrix, achieving a training accuracy of 0.722 and a validation accuracy of 0.689. An early stopping function is applied to halt training when there is no improvement in validation accuracy. The confusion matrix results indicate that the model adequately classifies data with Accuracy, Recall, and Precision values of 69%, 69%, and 70%, respectively.

Keywords: Modeling, Damage, Piston, CNN

INTRODUCTION

Quality control in the manufacturing industry of pistons is a crucial process due to its significant contribution to a company's success in meeting consumer needs, maintaining its reputation, and increasing revenue (De Vitis et al., 2020). Product quality control involves the supervision and control of product quality from the early stages of production to the point where the product is ready for market release (Benbarrad et al., 2021a). In the piston production process, errors may occur, resulting in defects that render the piston ineffective and poor piston quality. In manufacturing industries, poor-quality products can cause various issues, such as substantial costs, decreased customer satisfaction, reduced sales, loss of market share, and higher production costs due to the need to discard or repair damaged or defective products (Gunasekaran et al., 2019).

Product inspection is a crucial component of product quality control, aiming to evaluate and ensure that products meet predefined standards (Kim et al., 2021). Piston inspection is vital in maintaining product quality by examining and selecting abnormal piston shapes, such as defects, cracks, and rust. Many companies still need to rely on manual piston inspection processes that involve human observation, which becomes ineffective when dealing with large-scale piston production. Detecting defects in pistons is a repetitive and monotonous task, and this process can reduce an individual's effectiveness if carried out over an extended period. Additionally, it requires a considerable workforce and entails significant costs (Benbarrad et al., 2021b).

Machine learning is a branch of artificial intelligence (AI) that enables computers or machines to learn from provided data, and these models can enhance their performance as more training data is incorporated into the dataset (Leni et al., 2023). Machine learning can automatically undergo the learning process through data processing, making predictions, and decision-making based on patterns in the training data. Convolutional Neural Network (CNN) is one popular and highly effective machine learning algorithm in recognizing patterns in visual data such as images and videos (Hendri Candra Mayana & Desmarita Leni, 2023). CNN can classify photos by breaking them into smaller parts (filters) and examining each part to determine the desired category. It allows machine learning modelling to perform repetitive and monotonous tasks, such as the piston inspection.

Numerous research studies and applications have utilized CNN in manufacturing industries, such as inspecting metal casting products; capable of automatically detecting defects like blow holes, chipping, cracks, and washing, with a high accuracy rate of 98% (Leni & Yermadona, 2023). Additionally, CNN is employed to detect defects in electronic products with limited data, demonstrating fairly good Accuracy (Leni, 2023). Another study by Park et al. focuses on defect detection on the surface of machine transmissions using a CNN-based method with a single RGB camera. Detectable defects include craters, pores, foreign substances, and fissures, and the proposed system shows excellent performance in all conducted experiments.

Based on the results of the studies above, it is evident that CNN exhibits excellent performance in classifying visual data such as photos. This research aims to develop a machine learning model based on the Convolutional Neural Network (CNN) algorithm for detecting damage in pistons. The data used in this study consists of three categories of pistons: defective pistons, oil-stained pistons, and regular pistons. The proposed method in this research involves image processing with segmentation techniques to separate piston images into specific parts, facilitating the feature recognition process. Subsequently, CNN is employed to classify piston images based on damage categories, and the model's performance is evaluated using evaluation metrics such as Accuracy, loss, and confusion matrix.

METHOD

This research designs a machine-learning model for detecting piston damage using a Convolutional Neural Network (CNN). The data utilized in this study consists of three piston categories: (1) Defected1, representing images of defective pistons; (2) Defected2, depicting pistons stained with oil; and (3) Normal, representing regular pistons. The study employs an 80% training data and 20% testing data split. Model evaluation is performed using confusion metrics. The stages of this research are outlined as follows:

1. Data Preprocessing

   In this stage, the data is prepared to suit the input of the designed CNN model. The preprocessing process involves several steps, including converting the image format to RGB, resizing images to the exact dimensions, and normalizing pixel intensity. This stage also includes the division of the dataset into training and testing data. The training data is used to train the CNN model, while the testing data is employed to evaluate the performance of the generated CNN model. Maintaining consistent image sizes can facilitate the training and data processing processes, enhancing the CNN model's ability to recognize normal and damaged piston images.

2. Convolutional Neural Network (CNN) Modeling

   The proposed modelling architecture in this research is illustrated in Figure 2, where the CNN architecture consists of convolutional layers, pooling layers, fully connected layers, and softmax activation.

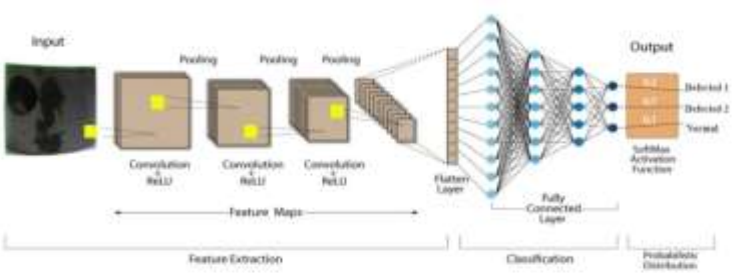

Figure 2. Proposed Model Architecture

   The Convolutional Layer is used to create features that can distinguish between images of damage and standard images. In contrast, the pooling layer is used to reduce the size of the images and enable the network to eliminate unnecessary details (Zhang et al., 2017). After passing through several convolutional and pooling layers, the model will enter the Fully Connected Layer, which aims to combine the previous layers' features into more complex ones. Softmax activation is the activation function used in this modelling, which is applied to the output layer in a neural network that performs multi-class classification. The softmax function transforms the numeric values at each neuron in the output layer into probability values, ensuring that the total probability across all classes sums up to 1 (Peng et al., 2017).

3. Model Training

Model training aims to generate a model that can predict desired outcomes based on the given inputs. The model training process involves optimizing model parameters to minimize prediction errors and enhance the model's ability to recognize patterns in the data.

4. Model Validation

   This study employs cross-validation with k-fold to measure the model's performance on limited data, where the data is divided into several folds or partitions, each serving as the testing data once. In contrast, the remaining data is used as training data. It ensures that each data point is used alternately as testing and training data (Leni et al., 2022).

5. Model Evaluation

   The model evaluation uses a confusion matrix to calculate Accuracy, precision, recall, and F1-score (Kareem et al., 2021).

   a. Accuracy is the ratio between the number of correctly classified data (true positive and true negative) and the total number of data. Accuracy depicts how often the model can correctly classify data. Accuracy is expressed as a percentage; the higher the accuracy value, the better the model's performance. Accuracy can be calculated using Equation 1.

   $$\text{Accuracy} = \frac{TP+TN}{TP+TN+FP+FN}$$

   b. Precision: The ratio between the number of actual positive instances and the total number of instances classified as positive by the model (true positive and false positive). It depicts how accurately the model classifies data as positive. Precision can be calculated using Equation 2.

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

   c. Recall: It is the ratio between the number of valid positive instances and the total number of actual positive instances (true positive and false negative). Recall depicts how sensitive the model is in classifying data that is positive. Recall can be calculated using Equation 3.

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

   d. F1-score: It is the harmonic mean between precision and recall. This indicator illustrates the balance between precision and recall. F1-score can be calculated using Equation 4.

   $$F1 - Score = \frac{(Recall\ x\ Presisi)}{(Recall+Presisi)}$$

   Using TP, TN, FP, and FN indicates True Positive, True Negative, False Positive, and False Negative, respectively.

RESULTS AND DISCUSSION

The data utilized in this research consists of images of pistons categorized into three distinct categories: (1) Defected1, representing pistons with defects or damage in specific areas; (2) Defected2, depicting pistons with oil stains or lubricant fluid adhering to specific parts; and (3) Normal, representing pistons in a normal condition without defects and oil stains, as observed in Figure 3.

Figure 3. Piston dataset, (a) Defected 1, (b) Defected 2, (c) Normal

1. Data Preprocessing

    In this stage, the acquired data is cleaned from invalid and unclear data. Subsequently, the piston data is organized into folders based on the type of damage. The percentage distribution of data for each piston category includes 30.9% for Defected 1, 34.4% for Defected 2, and 34.7% for Normal, as depicted in Figure 4. The difference in the number of data points for each category is insignificant, with Defected 1 having the most minor data points compared to the other two categories.

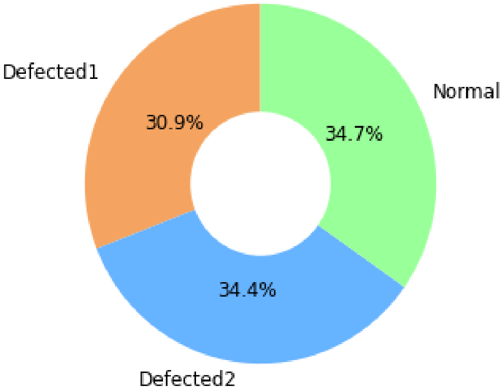

Figure 4. Percentage of Piston Dataset

    The next step is to prepare the training and testing data. In this study, the piston image data undergo normalization before input into the model. This normalization aims to transform image pixel values into a smaller range, facilitating better learning for the model (Borhanuddin et al., 2019). The normalization process is done by dividing each pixel value by the maximum possible value, 255, for RGB images with a value range of 0-255. In addition to normalization, augmentation techniques are employed to expand the testing dataset and increase variation among the images. Augmentation techniques include various image transformations, such as rotation, flipping, and scaling. Augmentation techniques can enhance the model's performance, and overfitting can be avoided (Han et al., 2018), (Frid-Adar et al., 2018). The training dataset for the model in this research comprises 285 images, as shown in Figure 5.

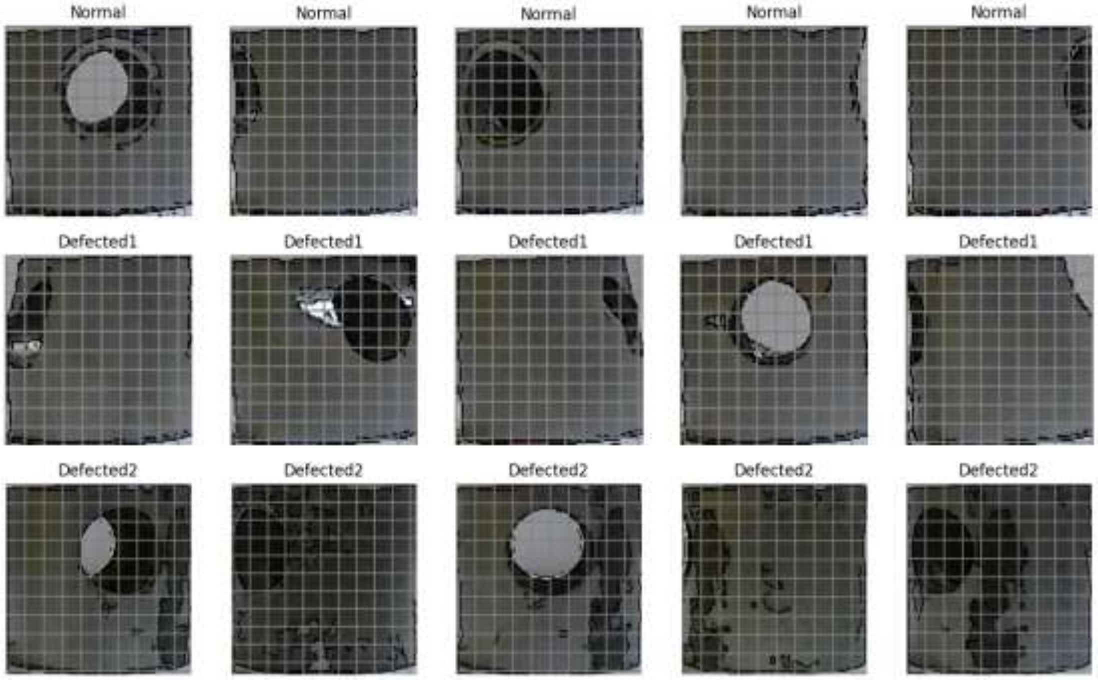

Figure 5. Model Training Dataset

The augmentation configuration utilized in this research encompasses various data transformation techniques to create a more diverse testing dataset and train the model to understand variations in the data. The augmentation techniques employed include rotation, zoom, flipping, cropping, and resizing. Rotation aims to rotate images within a specific angle range, introducing perspective and object orientation variations. Zoom is used to magnify or shrink images within a specific range, allowing for variations in object scale. Flipping is applied to horizontally or vertically flip images, introducing variations in object orientation. Cropping involves randomly cutting a portion of the image, introducing variations in object composition. At the same time, resizing is applied to change the image size to smaller or larger dimensions, introducing variations in object size. According to Liat al (Li et al., 2019), transformations such as rotation, flipping, and cropping can enhance the variation among dataset images, thereby improving model performance and preventing overfitting.

Meanwhile, resize and zoom transformations help address the dataset's varying image sizes by adjusting image sizes consistently. Probability controls how often a transformation is applied to dataset images. The higher the probability value, the more frequently the transformation is applied to the images.

2. Convolutional Neural Network (CNN) Modeling

The model is constructed with a CNN architecture using the Keras Sequential, consisting of several layers, to perform image classification on the augmented piston dataset. The model comprises multiple Conv 2D layers that extract features from images, followed by Maxpooling 2D layers to extract the most significant features. Subsequently, a flattened layer is employed to flatten the output before being fed into several Dense layers for classification. This architecture includes three Dense layers, with the final layer utilizing the softmax activation function for classifying into three distinct classes. The model uses the Adam optimizer with a learning rate 0.0001, a batch size 32, and the mean squared error loss function to measure prediction errors.

Additionally, the model employs various image augmentation techniques, such as rotation, shifting, and flipping, implemented through an Image data generator. The model is also trained using a learning rate scheduler and callbacks like early stopping, which halts training if there is no improvement in validation accuracy. The model architecture can be observed in Figure 6.

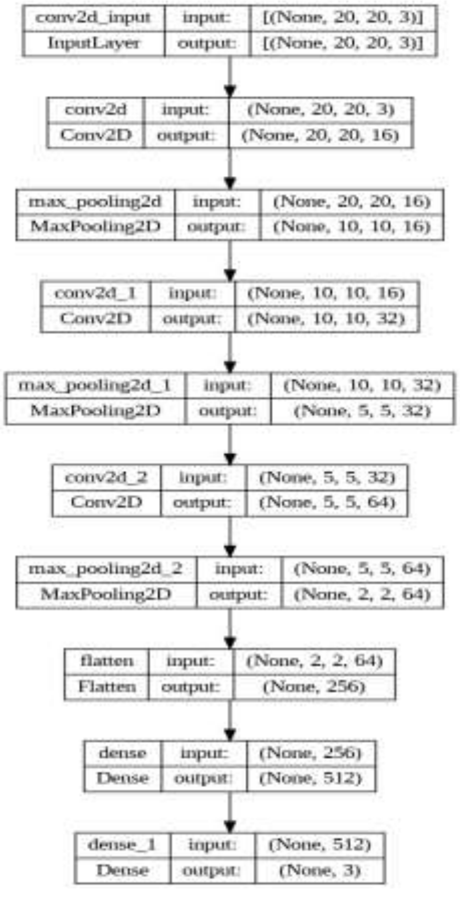

Figure 6. CNN Model Architecture

This CNN modelling employs the k-fold 5 cross-validation method to evaluate model performance. The k-fold cross-validation method divides the dataset into k equally sized groups, where each group functions as the testing dataset once, and the other groups are used as the training dataset (Fushiki, 2011). This study's training dataset comprises 285 images, and the testing dataset comprises 60 images. Using k-fold cross-validation with k=5, the dataset is divided into 5 equally sized groups, each containing 57 images (285/5). Subsequently, iterations are performed k times, where in each iteration, one group serves as the testing dataset, and the other 4 groups are used as the training dataset. Thus, each image is tested once and is learned in four different learning instances. The average Accuracy from these five iterations is then used as the model performance measure. For a more precise illustration of using k-fold 5 validation, refer to Figure 7.

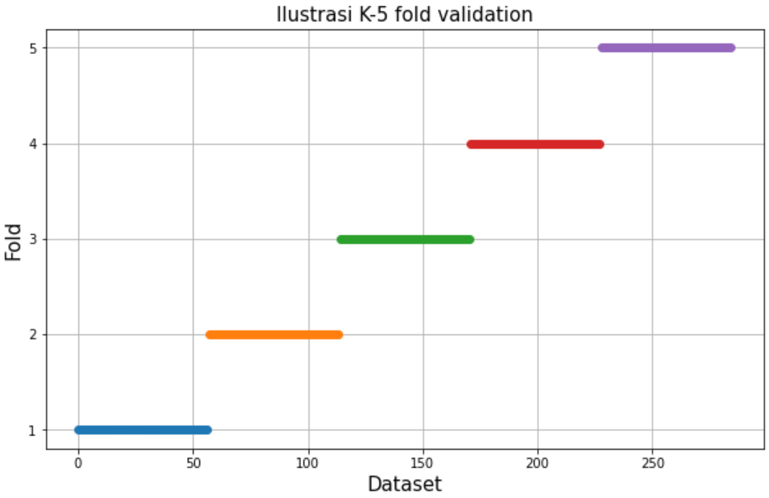

Figure 7. Illustration of k=5-fold validation.

The k-fold cross-validation method is employed to prevent overfitting in the model. Using k-fold cross-validation, the entire dataset is utilized as testing and training data simultaneously, allowing the model to learn more variations. In this study, the k-fold cross-validation method is used to evaluate the performance of the image classification model for pistons, which has been constructed using augmentation techniques to increase the testing dataset.

3. Model Evaluation

In this study, model evaluation is conducted by observing the Loss and Accuracy results during training and validation. The training Loss for this model is 0.151, and the validation Loss is 0.156. The Accuracy metric is also utilized to evaluate the model, with training Accuracy resulting in 0.722 and validation Accuracy in 0.689. These evaluation results indicate that the difference in loss values between the training and validation datasets is relatively small, suggesting that the model is not experiencing overfitting or underfitting. The training Accuracy of 0.722 demonstrates that the model can correctly classify data during training.

In contrast, the validation Accuracy of 0.689 indicates that the model still has limited capability in classifying new, unseen data. This modelling incorporates the early stopping function, which halts training if there is no improvement in validation accuracy. Thus, it can be observed from Figure 8 and Figure 9 that the model stopped at Epoch 25.

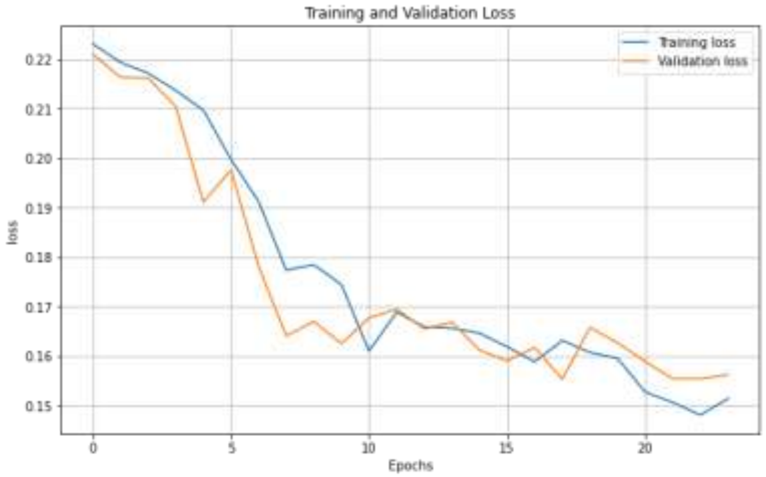

Figure 8. Comparison of training loss and validation loss results

Figure 9. Comparison of training accuracy and validation accuracy results

This model is also evaluated using confusion metrics with 60 test images, each category having 20 test data to obtain accurate overall model results. Based on the obtained confusion matrix results, it can be seen that the Accuracy value is 0.69, recall is 0.69, and precision is 0.70. Accuracy indicates the percentage of correctly classified overall data by the model, where in this study, an accuracy value of 0.69 was obtained, indicating that around 69% of the data has been correctly classified by the model. Recall shows how well the model recognizes and classifies data belonging to the positive class (Defected1, Defected2, and Normal). A recall of 0.69 indicates that the model can recognize 69% of all data belonging to the positive class. Precision indicates how much data classified as the positive class by the model belongs to the positive class. A precision of 0.70 indicates that out of all the data classified by the model as positive, 70% belong to the positive class. The model can effectively classify images of defective, oil-stained, and standard piston wells. However, the validation results of the model using new data, such as augmentation techniques, still need to be improved. For a clearer view, refer to Figure 10.

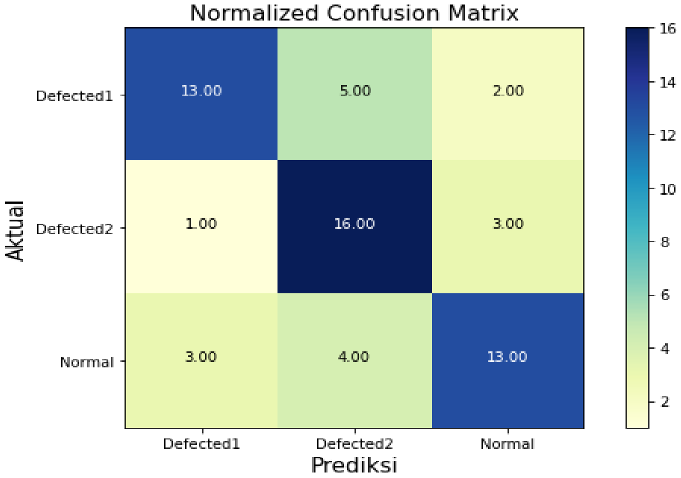

Figure 10. Confusion matrix results using the CNN model

CONCLUSION

This study designs a machine learning model using the Convolutional Neural Network (CNN) algorithm to detect piston damage. The data used in this study consists of piston images categorized into three categories: (1) Defected1, which is a piston with defects or damage in certain parts; (2) Defected2, which is a piston with oil stains or lubricant fluid adhering to certain parts, and (3) Normal, which is a piston in a normal condition without defects and oil stains. The data used in this study amounted to 285 images for training data, with the percentage of Defected1 data being 30.9%, Defected2 34.4%, and Normal 34.7%. The validation data uses new data created with augmentation techniques, with a total of 60 data, each piston category with 20 images. The model is created with a CNN architecture using Keras Sequential, consisting of several layers such as a convolutional layer, pooling layer, Fully Connected Layer, and activation softmax. This model uses the Adam optimizer with a learning rate of 0.0001, and k-fold 5 validation is used to validate the model. The model is evaluated using the Loss, Accuracy, and Confusion matrix, where the training loss is 0.151, and the validation loss is 0.156. The Accuracy results obtained from this modelling are 0.722 for training and 0.689 for validation. This modelling incorporates the EarlyStopping function, which stops training if there is no improvement in validation accuracy. Based on the results obtained, the model stops at epoch 25. The confusion matrix results in this study show an Accuracy value of 0.69, recall of 0.69, and precision of 0.70, indicating that the model can classify data quite well but is still limited in classifying new data. Although the results are satisfactory, this study suggests continuing model development by fine-tuning and experimenting with other parameters to improve model performance and enhance the ability to classify images of difficult-to-recognize damage classes.

REFERENCES

Benbarrad, T., Salhaoui, M., Kenitar, S. B., & Arioua, M. (2021a). Intelligent Machine Vision Model for Defective Product Inspection Based on Machine Learning. *Journal of Sensor and Actuator Networks*, *10*(1), 7. https://doi.org/10.3390/jsan10010007

Benbarrad, T., Salhaoui, M., Kenitar, S. B., & Arioua, M. (2021b). Intelligent Machine Vision Model for Defective Product Inspection Based on Machine Learning. *Journal of Sensor and Actuator Networks*, *10*(1), 7. https://doi.org/10.3390/jsan10010007

Borhanuddin, B., Jamil, N., Chen, S. D., Baharuddin, M. Z., Tan, K. S. Z., & Ooi, T. W. M. (2019). Small-Scale Deep Network for DCT-Based Images Classification. *2019 4th International Conference and Workshops on Recent Advances and Innovations in Engineering (ICRAIE)*, 1–6. https://doi.org/10.1109/ICRAIE47735.2019.9037777

De Vitis, G. A., Foglia, P., & Prete, C. A. (2020). Row-level algorithm to improve the real-time performance of glass tube defect detection in the production phase. *IET Image Processing*, *14*(12), 2911–2921. https://doi.org/10.1049/iet-ipr.2019.1506

Frid-Adar, M., Diamant, I., Klang, E., Amitai, M., Goldberger, J., & Greenspan, H. (2018). GAN-based synthetic medical image augmentation for increased CNN performance in liver lesion classification. *Neurocomputing*, *321*, 321–331. https://doi.org/10.1016/j.neucom.2018.09.013

Fushiki, T. (2011). Estimation of prediction error by using K-fold cross-validation. *Statistics and Computing*, *21*(2), 137–146. https://doi.org/10.1007/s11222-009-9153-8

Gunasekaran, A., Subramanian, N., & Ngai, W. T. E. (2019). Quality management in the 21st-century enterprises: Research pathway towards Industry 4.0. *International Journal of Production Economics*, *207*, 125–129. https://doi.org/10.1016/j.ijpe.2018.09.005

Han, D., Liu, Q., & Fan, W. (2018). A new image classification method using CNN transfer learning and web data augmentation. *Expert Systems with Applications*, pp. *95*, 43–56. https://doi.org/10.1016/j.eswa.2017.11.028

Hendri Candra Mayana & Desmarita Leni. (2023). Deteksi Kerusakan Ban Mobil Menggunakan Convolutional Neural Network dengan Arsitektur ResNet-34. *Jurnal Surya Teknika*, *10*(2), 842–851. https://doi.org/10.37859/jst.v10i2.6336

Kareem, S., Hamad, Z. J., & Askar, S. (2021). An evaluation of CNN and ANN in prediction weather forecasting: A review. *Sustainable Engineering and Innovation*, *3*(2), 148–159. https://doi.org/10.37868/sei.v3i2.id146

Kim, T.-H., Kim, H.-R., & Cho, Y.-J. (2021). Product Inspection Methodology via Deep Learning: An Overview. *Sensors*, *21*(15), 5039. https://doi.org/10.3390/s21155039

Leni, D. (2023). The Influence of Heatmap Correlation-based Feature Selection on Predictive Modeling of Low Alloy Steel Mechanical Properties Using Artificial Neural Network (ANN) Algorithm. *Journal of Energy, Material, and Instrumentation Technology*, *4*(4), 152–162. https://doi.org/10.23960/jemit.v4i4.203

Leni, D., Earnestly, F., Sumiati, R., Adriansyah, A., & Kusuma, Y. P. (2023). Evaluasi sifat mekanik baja paduan rendah bedasarkan komposisi kimia dan suhu perlakuan panas menggunakan teknik exploratory data analysis (EDA). *Dinamika Teknik Mesin*, *13*(1), 74. https://doi.org/10.29303/dtm.v13i1.624

Leni, D., Kusuma, Y. P., & Sumiati, R. (2022). *Perbandingan Alogaritma Machine Learning Untuk Prediksi Sifat Mekanik Pada Baja Paduan Rendah*. *5*(2), 167–174.

Leni, D., & Yermadona, H. (2023). Pemodelan Inspeksi Kerusakan Ban Mobil Menggunakan Convolutional Neural Network (CNN). *Jurnal Rekayasa Material, Manufaktur Dan Energi*, *6*(2). https://doi.org/10.30596/rmme.v6i2.16198

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*(4), 593–597. https://doi.org/10.1109/LGRS.2018.2878773

Peng, H., Li, J., Song, Y., & Liu, Y. (2017). Incrementally Learning the Hierarchical Softmax Function for Neural Language Models. *Proceedings of the AAAI Conference on Artificial Intelligence*, *31*(1). https://doi.org/10.1609/aaai.v31i1.10994

Zhang, R., Zheng, Y., Mak, T. W. C., Yu, R., Wong, S. H., Lau, J. Y. W., & Poon, C. C. Y. (2017). Automatic Detection and Classification of Colorectal Polyps by Transferring Low-Level CNN Features From Nonmedical Domain. *IEEE Journal of Biomedical and Health Informatics*, *21*(1), 41–47. https://doi.org/10.1109/JBHI.2016.2635662