Image Classification of Room Tidiness Using VGGNet with Data Augmentation

Leni Fitriani^{1*}, Ayu Latifah², Moch. Rizky Cahyadiputra³

^{1,2,3} Department of Computer Science, Institut Teknologi Garut, Indonesia

*corr_author: leni.fitriani@itg.ac.id

Abstract - Tidiness becomes an essential aspect that everyone should maintain. Tidiness encompasses various elements, and one of the aspects closely related to it is the tidiness of a room. The tidiness of a room creates a comfortable and clean environment. The tidiness of a room is particularly crucial for individuals involved in businesses such as the hospitality industry. Therefore, a solution is needed to address this issue, and one of the approaches is to utilize Deep Learning for automatic room tidiness classification. One popular deep learning method for implementing image classification of room tidiness is the convolutional neural network (CNN), which creates a well-performing model for image classification with data augmentation. This research aims to develop an image classification model using CNN with the VGGNet architecture and data augmentation. This study is a reference for further development, with potential applications in the hospitality industry. The research results in a model that achieves an accuracy of 98.44% with a data proportion of 90% for training and validation, while the remaining 10% is used for testing purposes. The conclusion drawn from this study is that the CNN method, combined with data augmentation, can be utilized for image classification of room tidiness.

Keywords: Convolutional Neural Network, data augmentation, image classification, room tidiness, VGGNet

I. INTRODUCTION

Tidiness is the quality of an object, place, or situation being tidy, orderly, and clean. It encompasses various aspects, such as room cleanliness, layout, etc [1]. Ensuring tidiness involves several important factors, as mentioned in the book "The 5S: Five Keys to a Total Quality Environment." For example, individuals must sort out necessary and unnecessary items, and tidiness is assessed based on well-organized belongings [2]. By implementing these aspects, tidiness will create a comfortable environment and enhance productivity.

One fundamental aspect of creating a comfortable environment is the cleanliness and tidiness of a room. An article titled "What a Messy Room Says About You" highlights how a room's condition can impact an individual's health and mental well-being. A cluttered room can cause stress and anxiety and hinder productivity and creativity [3]. The tidiness of a room should be maintained at all times and in various settings, such as office spaces, hotel rooms, and individual rooms within homes. Moreover, maintaining cleanliness and tidiness is particularly important in smart buildings, where technological advancements and intelligent systems enable efficient and sustainable management to create a comfortable environment [4].

The issue of maintaining tidiness in places or rooms is not only relevant to individuals but also affects businesses. Business owners need to prioritize and emphasize the tidiness of their establishments. However, they may encounter challenges when managing multiple business locations, as they must constantly exert direct control to ensure each place remains hygienic, orderly, and clean. This challenge becomes even more complex for owners in the hospitality industry, where housekeeping personnel or business owners themselves need to inspect each hotel room individually to ensure they are clean and contribute to customer comfort [5]; additionally, during an interview session with a hotel team member at Fave Hotel, located at Jl. Cimanuk No. 338, Tarogong Kidul District, Garut Regency, it was revealed that checking room conditions is still manually. General cleaning or housekeeping staff must ensure that each room is tidy and clean.

Given the previous problem, particularly in the hotel industry, there is a need for a solution in the form of a model or system that can easily classify whether a room is tidy or not. One viable solution is leveraging technological advancements that have gained momentum in recent years, specifically Deep Learning (DL). Deep learning is a subfield of Machine Learning (ML) inspired by how the human brain processes information patterns. DL does not rely on explicitly designed rules but utilizes much available data to map the input to specific labels [6]. With DL, a previously manual task can be automated more effectively, and in

addressing the classification of room tidiness, one can employ a Convolutional Neural Network (CNN), a method within DL designed explicitly for image classification.

CNN is a family of algorithms and models within Deep Learning for object recognition. The development of CNN dates back to the 1990s when LeCun and his colleagues proposed a novel Neural Network (NN) architecture for classifying handwritten digits from images, which was published at the Neural Information Processing Systems (NeurIPS) conference[7]. CNN encompasses several architectures that can process image data, such as feature detection, LeNet Architecture, AlexNet Architecture, ResNet Architecture, DenseNet Architecture [8]. As a solution to the previous challenge, VGGNet is a widely popular architecture that has been extensively used for image classification and has proven highly effective [9].

Furthermore, in Deep Learning, the quantity of data plays a crucial role in learning patterns. If the utilized data is more considerable, the likelihood of the model experiencing overfitting significantly increases. Overfitting occurs when the model performs exceptionally well on the training data but fails to generalize well on testing data [6]. Data augmentation techniques can be employed to address the issue of overfitting. Data augmentation is a method used to increase the number of training samples by generating new variations of existing training data [10].

This research is based on previous studies. The first study focused on building a model for clean room detection using Support Vector Machine and Neural Network algorithms [11]. The second study performed a hotel room classification system using CNN with VGGNet architecture and achieved an accuracy of 92.92% [5]. The third study conducted image classification based on object categories, obtaining an accuracy of 93.57% with 18 layers [12]. The fourth study addressed detection in image classification for diabetic retinopathy using a Neural Network with VGGNet-16 architecture, resulting in an accuracy of 93.73% [13]. The fifth study conducted image classification for Garutan batik using CNN and data augmentation, resulting in the highest accuracy of 91.30% with the ResNet-50 architecture [14]. The last study discussed medical image classification for disease diagnosis using a Convolutional Neural Network with VGGNet architecture and achieved an accuracy of 92.3% [15].

This research aims to create a modeling framework to serve as a basis for further model development. The application of this model can be utilized in the hospitality industry to enhance job performance and implement the VGGNet architecture, which plays a significant role in the image classification model.

This research represents the original work and creativity of the researcher, including the utilization of Convolutional Neural Network, the implementation of Deep Learning methods for image processing, and the incorporation of data augmentation, which adds uniqueness to the study. The resulting model can be employed to develop more complex and practical models.

II. METHOD

Based on the previous problem, the researcher has established a framework to address the issue. The research framework is outlined as follows in Fig. 1. Fig. 1 can be explained as follows :

1) Initial Identification: In the initial identification phase, a literature review was conducted on relevant studies about image classification using Convolutional Neural Networks. During the literature review process, previous research was assessed, the issues addressed in prior studies were identified, and discussions were engaged with the supervisor to obtain feedback and suggestions regarding the proposed research. Additionally, an interview was conducted with a hotel industry team member to strengthen the research further.

2) Data Collection: In the data collection phase, the researcher gathered relevant data for the study and analyzed the collected dataset to enable classification. The data collection process focused on obtaining a dataset consisting of room images. Subsequently, an analysis was conducted, preparing for the subsequent pre-processing stage.

3) *Pre-processing:* The pre-processing stage involves the normalization and data augmentation of the dataset obtained during the data collection phase. The normalization process is carried out to transform the image data into a format easily interpretable by the machine or model. Once the normalization stage is completed, the data augmentation process introduces variations to the training data samples and prevents overfitting. The image data obtained was taken from several existing sources. Apart from that, the image data was obtained using electronic data collection:



Fig. 1 Stages of the method carried out

- Kaggle, neat and messy room image data obtained from one of the most extensive dataset providers, namely Kaggle with a dataset address at (https://kaggle.com/datasets/cdawn1/messy-vsclean-room), which consists of 10 testing data, 190 train data, and 20 validation data. Here are some images of the neat and messy room dataset taken from Kaggle by a user named Guanqiao Ding, which can be seen in Fig. 2.
- Historical data from several free image provider websites, namely Unsplash, Pixabay, Flickr, Wikimedia Commons, Rawpixel, and Pexels, were used. Some of these sources are image sources for neat and messy rooms. Some sites can be visited at the following addresses:

Unsplash (https://unsplash.com).

Pixabay (https://pixabay.com). Flickr (https://flickr.com). Wikimedia Commons (https://commons.wikimedia.org). Rawpixel (https://rawpixel.com). Pexels (https://pexels.com).

4) Classification: The researcher performed classification on the pre-processed dataset. Subsequently, the implementation of the VGGNet architecture in Fig. 3 and the training process were initiated. Implementing the VGGNet architecture is a crucial component of the Convolutional Neural Network (CNN), as shown in Fig. 4. Once constructed, the architecture underwent training, resulting in a model ready for testing and image classification.





Fig. 3 VGGNet-16 Architecture



Fig. 4 CNN Architecture for Image Classification

CNN serves as a fundamental architecture. By applying convolution filters, CNN is designed to automatically and adaptively learn hierarchies of features from input images. These filters detect various patterns, such as edges and textures. VGGNet, a specific CNN Architecture, gained prominence for its simplicity and effectiveness. It comprises a series of convolutional layers followed by poling layers, with increasing depth as the network progresses.

1) Testing: The final stage involves testing and evaluating the constructed model using a confusion matrix and several other metrics: recall, precision, and F1-score.

III. RESULT AND DISCUSSION

A. Initial Identification

In the initial identification phase, several processes were conducted to obtain insights into image classification using Convolutional Neural Network (CNN). It began with identifying the subject of research, which is the cleanliness of a room. Subsequently, the focus shifted to obtaining the research problem through a literature review, which allowed for a more targeted approach to the research topic.

After conducting a literature review, the researcher conducted an interview session with a team member of Fave Hotel located at Jln: Cimanuk No. 338, Tarogong Kidul sub-district, Garut district, namely Muhammad Bagja Sukriyansah. The interview concluded that checking rooms' cleanliness is still manually, where general cleaning and housekeeping staff need to ensure that each room is tidy and clean. Additionally, general cleaning staff must ensure that all objects and items, such as bathroom amenities and bedding accessories in the bedroom, are placed in their designated locations. Moreover, general cleaning staff need to inspect every corner of the room meticulously. After completing this process, they must also inform the front office that the room is ready.

B. Data Collection

1) Image Data Collection: The first process is data collection, which will be gathered according to the research theme of room tidiness. The collected data consists of data from various rooms, such as bedrooms, bathrooms, toilets, living rooms, kitchens, and workspaces. Two data types are to be collected: tidy rooms and messy rooms. The image data will be obtained from several sources, including Kaggle, Unsplash, Pixabay, Flickr, Wikimedia Commons, Rawpixel, and Pexels. Based on the previous sources, the total of image data obtained for data collection is 1130, 565 images categorized as "Tidy room" and 565 images categorized as "Messy room."

2) Data Analysis: In the data analysis phase, the obtained room image data will be divided into several subsets to train the image classification model. From the collected data, which consists of tidy and messy rooms with a total of 2 labels, the tidy room data contains 565 images. In contrast, the messy room data contains 565 images, resulting in a total of 1130 images when combined.

The data will be divided into several parts, namely training data, validation data, and testing data. This research adjusts the ratio used for Deep Learning according to the needs. The researcher has chosen a specific ratio for training and validation data, which is divided as follows: 90% for training and validation data, further divided into 60% for training data, totaling 610 images, and 40% for validation data, totaling 407 images. As for the testing data, it comprises 10% of the total data, which is 113 images. This ratio is visually represented in Fig.5.

C. Pre-processing

After the data analysis stage, the divided dataset will enter the pre-processing process, including data normalization and augmentation. Preparing the data before entering the model training phase is essential. The data normalization ensures that the data is transformed into a standardized format suitable for training. On the other hand, data augmentation generates variations of the training data to increase its diversity and prevent overfitting. A detailed description of each pre-processing step can be found in Fig. 6.



Fig. 5 Ratio for the division of image data



Fig. 6 Pre-processing process

1) Data Normalization: Data normalization standardizes the range of values for an attribute or feature. As depicted in the image, the pixels in each part of the image have values ranging from 0 to 255. By applying data normalization, each pixel value is transformed to have a consistent scale ranging from 0 to 1. Data normalization enables improved model performance. An example of the resulting normalized image data can be seen in Fig. 7.

2) Data Augmentation: The use of data augmentation techniques has a significant impact on training data. By employing these techniques, variations of the original images are created to generate new image data. This helps to increase the diversity and quantity of training data available to train the model, for example, by utilizing techniques such as rotation, zoom range, width shift range, height shift range, shear range, and horizontal flip. Therefore, using data augmentation techniques results in more variants of training data with various positions and conditions, which in turn helps improve the performance and generalization of the model in image classification tasks.



Fig. 7 Normalized data results

The rotation range is the first augmentation parameter, where the images are rotated. An example of the augmentation result from the rotation range can be seen in Fig. 8.

The zoom range is the second augmentation parameter, where the images are enlarged. An example of the augmentation result from the zoom range can be seen in Fig. 9. The width shift range is the third augmentation parameter, where the images undergo horizontal shifting. An example of the augmentation result from the width shift range can be seen in Fig. 10.

An example of the augmentation result from the height shift range can be seen in Fig. 11. The shear range is the fifth augmentation parameter, where the images are sheared clockwise, resulting in a parallelogram-like shape when used. An example of the augmentation result from the shear range can be seen in Fig. 12. Horizontal flip is the sixth augmentation parameter, where the images are horizontally flipped. An example of the augmentation result from the horizontal flip can be seen in Fig. 13.

Height shift range is the fourth augmentation parameter, where the images undergo vertical shifting.







Fig. 9 Rotation range





Fig. 10 Height shift range





Fig. 13 Shear range

D. Classification

After going through the pre-processing phase, it's time to move into the classification phase. The prepared data will be implemented using a Deep Learning architecture called Convolutional Neural Network (CNN), and one specific type of CNN architecture used is VGGNet. The workflow in the classification phase follows a sequential order from the beginning to the end—the step-by-step process for image classification, which is commonly used and can be illustrated in Fig 14.



Fig. 14 Classification workflow

1) VGGNet Architecture Implementation: In the implementation phase of the VGGNet architecture, the researcher utilized the VGGNet-16 architecture. The VGGNet architecture comes in various configurations, such as VGGNet-19 and VGGNet-16. The numbers represent the total number of layers in the architecture. Precisely, VGGNet-16 consists of 13 convolutional layers and three fully connected layers. An overview of the VGGNet-16 architecture can be seen in Fig. 15.

2) *Training*: The next step is the model training phase. In this phase, the constructed architecture undergoes a compilation process. Before entering the model fitting stage, some preparations are needed, including implementing callbacks. Callback functions notify the model to stop training when a specific target is achieved.

Next, move on to the model fitting phase, where the model undergoes the training process by the modeling engine. The binary cross-entropy activation function and Adam optimizer are used in this process. A comparison of the training results for each epoch of the model is presented in Table I.



Fig. 15 VGGNet-16 architecture

TABLE I									
MODEL TRAINING RESULTS									
	Train	ing	Validation						
Epoch	Accuracy	Loss	Accuracy	Loss					
1	0.6003	1.2280	0.8452	0.3395					
2	0.8374	0.3506	0.9410	0.1662					
3	0.8927	0.2555	0.9484	0.1496					
4	0.9170	0.2124	0.9656	0.0929					
5	0.9481	0.1457	0.9607	0.0952					
6	0.9619	0.1164	0.9656	0.0902					
7	0.9446	0.1306	0.9631	0.0930					
8	0.9654	0.1053	0.9533	0.1054					
9	0.9481	0.1187	0.9656	0.0767					
10	0.9446	0.1312	0.9607	0.1094					
11	0.9585	0.1010	0.9631	0.1001					
12	0.9671	0.1089	0.9582	0.0767					
13	0.9585	0.1099	0.9582	0.1157					
14	0.9550	0.1120	0.9582	0.0732					
15	0.9533	0.1316	0.9754	0.0700					
16	0.9689	0.0889	0.9484	0.1662					
17	0.9533	0.1141	0.9607	0.0851					
18	0.9533	0.1087	0.9558	0.1469					
19	0.9654	0.0814	0.9631	0.1084					
20	0.9429	0.1439	0.9631	0.0876					
21	0.9446	0.1444	0.9631	0.0961					
22	0.9481	0.1143	0.9214	0.2442					
23	0.9619	0.1146	0.9558	0.1589					
24	0.9637	0.0858	0.9607	0.0800					
25	0.9792	0.0745	0.9681	0.0988					
26	0.9827	0.0467	0.9509	0.1675					
27	0.9602	0.0975	0.9681	0.0810					
28	0.9723	0.0755	0.9730	0.0713					
29	0.9844	0.0527	0.9656	0.0799					

As shown in Table I, it can be observed that the model stopped at epoch 29 as it reached the predefined target specified by the callback. In each epoch, the model experienced an increase in accuracy and a decrease in loss. Higher accuracy indicates better prediction performance of the model, while lower loss signifies improved learning of patterns within the data. Minimizing the loss aims to reduce errors during the model training process. However, accuracy and loss values are not the sole criteria for evaluating the training results. Graphical representations of the training and validation data from the previous training session can be analyzed to assess the model's performance. The training results are depicted in Fig. 16.



Fig. 16 Evaluation model chart

In Fig. 16, it can be observed that the training results indicate optimal and stable performance. For instance, in terms of model accuracy, the gap between accuracy and validation accuracy across multiple epochs is small. This signifies that the model consistently performs well on the training and validation datasets. Similarly, the difference between loss and validation loss throughout the training epochs is slight for model loss and remains stable. This indicates that the model effectively minimizes errors and maintains a consistent performance.

E. Testing

Testing is the final phase in this research, where the trained model undergoes evaluation using several commonly used metrics to assess its performance. This evaluation determines whether the model is ready for further development or application and whether users can publish and use it.

Throughout testing, the model is presented with images or data samples from the test dataset that have not been encountered during the training phase. Subsequently, the model generates predictions or classifications based on these previously unseen data samples. These predictions are then juxtaposed against the test data's actual labels or ground truth values to assess the model's performance.

The first test involves using images of tidy rooms. The obtained results from the model predictions can be observed in Fig. 17. The second test involves using images of messy rooms. The obtained results from the model predictions can be observed in Fig. 18.



Fig. 17 Tidy room model prediction results



Fig. 18 Messy room model prediction results

After conducting direct testing on the image data, the final step is to examine the model's performance using a confusion matrix, which is one of the ways to evaluate the model. The evaluation of the model based on the confusion matrix can be represented in the form of Fig. 19.

In Fig. 19, it can be observed that the model accurately predicted all 113 data instances. Precisely, 57 images of tidy rooms were correctly predicted without any errors, and the remaining 56 images of messy rooms were also correctly predicted as messy rooms with no prediction errors. In addition to the confusion matrix, the model's performance is evaluated using precision, recall, and f1-score. The performance metrics, including accuracy, precision, recall, and f1-score, can be seen in Fig. 20.



Fig. 19 Confusion matrix

	Confusion Matrix								
	prec	ision	recall	f1-score	support				
	Clean	1.00	1.00	1.00	57				
	Messy	1.00	1.00	1.00	56				
	accuracy			1.00	113				
	macro avg	1.00	1.00	1.00	113				
	weighted avg	1.00	1.00	1.00	113				
	true negative: 57								
false positive: 0									
false negative: 0 true positif: 56									

Fig. 20 Classification report

In the tidy room category, the precision achieved 100%, recall achieved 100%, and the f1-score also reached 100%. Similarly, in the messy category, precision achieved 100%, recall achieved 100%, and the f1-score also reached 100%. These metrics were calculated manually and can be summarized as follows.

Accuracy =
$$\frac{\text{TN} + \text{TP}}{\text{TN} + \text{FP} + \text{FN} + \text{TP}} = \frac{57 + 56}{57 + 0 + 0 + 56} = 1.00$$

Messy room evaluation:

$$Precision = \frac{1P}{TP + FP} = \frac{56}{56 + 0} = 1.00$$

$$Recall = \frac{\Pi}{\mathrm{TP} + \mathrm{FN}} = \frac{30}{56 + 0} = 1.00$$

$$F_{I} = 2*\frac{(precision * recall)}{(precision + recall)} = 2*\frac{(1.00 * 1.00)}{(1.00 + 1.00)} = 1.00$$

Tidy room evaluation:

$$Precision = \frac{1P}{TP + FP} = \frac{37}{57 + 0} = 1.00$$
$$Recall = \frac{TP}{TP + FN} = \frac{57}{57 + 0} = 1.00$$

$$F_{I} = 2*\frac{(precision * recall)}{(precision + recall)} = 2*\frac{(1.00 * 1.00)}{(1.00 + 1.00)} = 1.00$$

The accuracy is calculated as the sum of True Negatives (TN) and True Positives (TP) divided by the total number of samples. It's 57 (TN) + 56 (TP) divided by 57 (TN) + 0 (FP) + 0 (FN) + 56 (TP), resulting in 113/113, which equals 1.00 or 100%. The accuracy score of 100% indicates that the model achieved good classification performance on the dataset used for evaluation.

Messy room evaluation. Precision: measures the proportion of correctly identified positive cases among all cases classified as positive. Precision is calculated to be 1.00 or 100%. This means all the samples classified as messy rooms are correctly identified as messy. Recall: Recall (also known as sensitivity) measures the proportion of actual positive cases the model correctly identified; recall is calculated as 1.00 or 100%. This means the model correctly identifies all the messy rooms in the dataset. F1-score: the harmonic mean of precision and recall, providing a single metric that balances precision and recall. The F1-score is calculated to be 1.00. This indicates a perfect balance between precision and recall, suggesting good performance in identifying messy rooms.

The results of the tidy room evaluation are similar to those of the messy room evaluation. Precision: signifies the accuracy of identifying tidy rooms among all cases classified as tidy. It's determined to be 1.00 or 100%, indicating that every sample classified as a tidy room is correctly identified. Recall: also known as sensitivity, assesses the model's ability to correctly identify actual positive cases, in this case, tidy rooms. With a calculated recall of 1.00 or 100%, the model accurately identifies all tidy rooms within the dataset. F1-score: representing the harmonic mean of precision and recall, offers a unified metric balancing both aspects. A calculated F1score of 1.00 suggests a harmonious blend of precision and recall, indicating strong performance in identifying tidy rooms.

IV. CONCLUSION

Based on the findings of this research, it can be concluded that the Convolutional Neural Network (CNN) method, combined with data augmentation techniques, can be used for image classification of room tidiness. Data augmentation plays a crucial role in this context, as the model requires a large amount of data to learn from, and the available data must be abundant. Therefore, data augmentation provides the model with increased

variations in the image data, preventing overfitting and leading to improved performance. This research has successfully produced a model capable of recognizing room tidiness using the CNN architecture, specifically VGG-16, with an accuracy of 98.44%. In summary, this research contributes to the development of image classification. The results can be further developed and applied in the hospitality industry, particularly in assisting housekeeping or general cleaning personnel in assessing the readiness of rooms for customers. However, it should be noted that this research still has room for improvement. The scope of this study was limited to the testing phase, and additional data is needed to enhance the model's ability to learn from a broader range of room tidiness patterns. Furthermore, future developments could include designing and implementing software for direct user interaction or exploring alternative CNN architectures beyond VGGNet.

REFERENCES

- [1] D. Collins, "The Health Benefits of Neatness and Order," *Linkedin*, Mar. 2020. .
- T. Osada, *The 5S : five keys to a total quality environment*. Tokyo : Asian Productivity Organization, [1991] ©1991, 1991.
- [3] K. Cherry, "What Does a Messy Room Say About You," *verywellmind*, Feb. 2021. .
- [4] A. Latifah, S. H. Supangkat, and A. Ramelan, "Smart Building: A Literature Review," in 2020 International Conference on ICT for Smart Society (ICISS), Nov. 2020, pp. 1–6, doi: 10.1109/ICISS50791.2020.9307552.
- [5] M. J. Dharmali, T. Lioner, and V. V. Susilo, "SISTEM KLASIFIKASI KERAPIHAN KAMAR HOTEL MENGGUNAKAN CONVOLUTED NEURAL NETWORK (CNN)," Comput. J. Comput. Sci. Inf. Syst., vol. 5, no. 2, p. 61, Dec. 2021, doi: 10.24912/computatio.v5i2.15175.
- [6] L. Alzubaidi, "Review of deep learning: concepts, CNN architectures, challenges, applications, future directions," *J. Big Data*, vol. 8, no. 1, p. 53, Mar. 2021, doi: 10.1186/s40537-021-00444-8.
- S. Raschka and V. Mirjalili, *Machine Learning and Deep Learning with Python, scikit-learn, and TensorFlow 2,* 3rd ed. Birmingham B3 2PB, UK: Packt Publishing Ltd., 2019.
- [8] K. Singh, D. Singh, and N. Mishra, "Review: Convolutional neural networks and its architecture," *Int. J. Health Sci. (Qassim).*, May 2022, doi: 10.53730/ijhs.v6nS1.7074.
- [9] U. Muhammad, W. Wang, S. P. Chattha, and S. Ali, "Pretrained VGGNet Architecture for Remote-Sensing Image Scene Classification," in 2018 24th International

Conference on Pattern Recognition (ICPR), Aug. 2018, pp. 1622–1627, doi: 10.1109/ICPR.2018.8545591.

- [10] C. Shorten and T. M. Khoshgoftaar, "A survey on Image Data Augmentation for Deep Learning," *J. Big Data*, vol. 6, no. 1, p. 60, Dec. 2019, doi: 10.1186/s40537-019-0197-0.
- [11] S. Suparni, H. Rachmi, and A. Al Kaafi, "Detection of Room Cleanliness Based on Digital Image Processing using SVM and NN Algorithm," *SinkrOn*, vol. 7, no. 3, pp. 777–783, Jul. 2022, doi: 10.33395/sinkron.v7i3.11479.
- [12] A. Mahajan and S. Chaudhary, "Categorical Image Classification Based On Representational Deep Network (RESNET)," in 2019 3rd International conference on Electronics, Communication and Aerospace Technology (ICECA), Jun. 2019, pp. 327–330, doi:

10.1109/ICECA.2019.8822133.

- [13] S. Wan, Y. Liang, and Y. Zhang, "Deep convolutional neural networks for diabetic retinopathy detection by image classification," *Comput. Electr. Eng.*, vol. 72, pp. 274–282, Nov. 2018, doi: 10.1016/j.compeleceng.2018.07.042.
- [14] L. Fitriani, D. Tresnawati, and M. B. Sukriyansah, "Image Classification On Garutan Batik Using Convolutional Neural Network with Data Augmentation," *JUITA J. Inform.*, vol. 11, no. 1, p. 107, 2023, doi: 10.30595/juita.v11i1.16166.
- [15] S. S. Yadav and S. M. Jadhav, "Deep convolutional neural network based medical image classification for disease diagnosis," *J. Big Data*, vol. 6, no. 1, p. 113, Dec. 2019, doi: 10.1186/s40537-019-0276-2.