

Implementation of Convolutional Neural Network Method in Identifying Fashion Image

Christian Sri Kusuma Aditya^{1*}, Vinna Rahmayanti Setyaning Nastiti², Qori Raditya Damayanti³,
Gian Bagus Sadewa⁴

^{1,2,3,4}*Informatika, Universitas Muhammadiyah Malang, Indonesia*

*corr_author: christianskaditya@umm.ac.id

Abstract - The fashion industry has changed a lot over the years, which makes it hard for people to compare different kinds of fashion. To make it easier, different styles of clothing are tried out to find the exact and precise look desired. So, we opted to employ the Convolutional Neural Network (CNN) method for fashion classification. This approach represents one of the methodologies employed to utilize computers for the purpose of recognizing and categorizing items. The goal of this research is to see how well the Convolutional Neural Network method classifies the Fashion-MNIST dataset compared to other methods, models, and classification processes used in previous research. The information in this dataset is about different types of clothes and accessories. These items are divided into 10 categories, which include ankle boots, bags, coats, dresses, pullovers, sandals, shirts, sneakers, t-shirts, and trousers. The new classification method worked better than before on the test dataset. It had an accuracy value of 95.92%, which is higher than in previous research. This research also uses a method called image data generator to make the Fashion MNIST image better. This method helps prevent too much focus on certain details and makes the results more accurate.

Keywords: Convolutional Neural Network; Fashion-MNIST; classification; augmentation

I. INTRODUCTION

In this industrial global era, the development of fashion in clothing is increasing. Over the past 30 years, the fashion industry has gone through big changes, leading to its growth and development. With the development of the fashion world, it makes it difficult for people to recognize the latest fashion variations because fashion has many types and variants, for example for tops there are variants such as sweaters, dresses, vests and many more. The many types of variants make it difficult for users to compare details of clothing types. In addition, the fashion industry has difficulty understanding customer tastes and directing sales to be better is a way to increase profits [1], [2]. The rise of the buying and selling process through digital

platforms has made people switch to buying fashion items through websites due to faster and easier access to technology. Introducing ways to make it easier for users to find things on these websites is very important [3], [4].

This problem can be solved by using object recognition. One way to do this is by using a method called Convolutional Neural Network to identify fashion images [5]. This method, which belongs to the category of Deep Learning techniques, is used to recognize and categorize an object in a digital image [6]. A significant problem in computer vision is image classification. It is used in many practical ways, like organizing and sorting images and videos. Even though it is easy for humans to identify objects in pictures, it is still a challenging problem for computers. Computer algorithms find it difficult to perform tasks as accurately as humans do.

Fashion classification is part of the broad task of categorizing its types [7], [8]. Creating image labels automatically to describe a product can help make the task of writing product descriptions easier. This kind of information can also aid in describing scenes and getting a clearer idea of the user's preferences, background, and financial situation [5], [9].

Fashion classification is a difficult task because it involves assigning labels to images that represent different types of fashion. The reason this multi-class fashion classification problem is hard is because there are many different fashion characteristics and a lot of different types of fashion to categorize. This complicated fashion sorting creates similarities in each label/class.

Deep Neural Networks have proven to be highly effective in solving a wide range of problems, exhibiting exceptional performance. One commonly utilized architecture in deep learning is the Convolutional Neural Network (CNN). CNNs are designed with multiple layers to detect patterns in data, making them particularly well-suited for image and pattern recognition tasks. They are trained using backpropagation, a technique that involves adjusting the network's parameters to minimize errors during training,

allowing them to learn from the data and improve their performance over time [10]. The Convolution Neural Network (CNN) has become very popular in image processing because it is great at classifying images [4] image recognition difficulties and significant improvements in accuracy results for machine learning. This technology has become a very strong and widely used method in machine learning [2].

The dataset that the authors use in this study is also used in previous studies [10]–[13]. The authors introduce the Fashion-MNIST dataset created from images found on Zalando, a large online fashion platform in Europe. Fashion-MNIST contains 70000 different items. Each item has a black and white image that measures 28x28 pixels. The items are divided into 10 categories: ankle boots, coats, dresses, pullovers, sandals, shirts, sneakers, t-shirts and trousers. The images are smaller versions of the images on their website. This dataset is available on keras and is typically used for testing Artificial Intelligence (AI) [5].

In [14], a research study focused on image classification using the CNN method with the VGG-11 model on the Fashion-MNIST dataset. They introduced a BatchNormalization layer after the PoolingLayer, achieving an accuracy result of 91.5%. Another research [12] employed the CNN4 model with 4 Convolutional Layers and Hyperparameter Optimization (HPO) and Regularization techniques, attaining an accuracy of 93.99%. Furthermore, in a different research on fashion article image classification using CNN with the CNN 2 Layer model, BatchNormalization, and Skip Connections, the study achieved an accuracy value of 92.54% [6].

In the research [15] the CNN regression model was applied in a Deep Learning approach to identify clothing classification in the clothing dataset. The dataset contains 25000 train data and 1472 test data. The purpose of the data augmentation process used in this study is to enhance the Deep Network's. The image augmentation used is rotation range 40, horizontal flip, and zoom range 0.2. The CNN Regression model is composed of two stacks of convolutional layers with max pooling layers and fully connected regression. The results of this study have a high level of accuracy (90%) using the CNN Regression method with additional.

Research [11] proposed two methods for the Fashion-MNIST classification case, namely HOG and Support Vector Machine (SVM). HOG is used as an extraction feature from an image to recognize fashion traits, the results of the extraction features are classified using the SVM method. Results of this research obtained the best accuracy of 86.53% by dividing the 4:4 image data..

Training that uses image data division with an unbalanced ratio will occur overfitting and even the author states that the right model has not been found from the given training data.

Based on the given description, this research utilizes the Convolutional Neural Network (CNN) as a specific type of Deep Learning algorithm for fashion image classification using the Fashion-MNIST dataset. CNN consists of multiple hidden layers, wherein each layer conducts mathematical computations involving input neurons, generating outputs based on weight, bias, and activation function values. One notable advantage of CNN is its ability to explore unsupervised predefined features, adding to the effectiveness of this method in image classification tasks. Furthermore, CNN incorporates preprocessing techniques during the convolution process to extract implicit characteristics from an image[16]. These advantages of CNN proved to be quite ideal for fashion image classification on the Fashion-MNIST dataset.

Based on the results of previous studies [6], [11], [12], [14], [15] that used classification methods, this research aims to compare the accuracy results with the accuracy results in previous studies. In addition, the author also wants to compare the accuracy results using augmented data and not using augmented data. Augmentation is done to enhance the performance of the Convolutional Neural Network. In addition, to determine how much impact augmentation use has on the accuracy of the results, it is carried out comparison of the accuracy results of the accuracy of architectural models using augmentation models and without using augmentation models[17]. And the author also wants to provide convenience for future research to make it easier to choose the most suitable classification method and so that it can get better accuracy results than this research.

II. METHOD

This research adopts the Convolutional Neural Network (CNN) architecture and leverages the Tensorflow library as the foundation of its methodology [18] to develop and train existing models in the application for cloud-based data generation, namely Google Collaboratory. The research methodology flow, illustrated in Fig. 1.

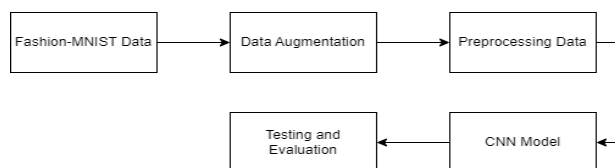


Fig. 1 Stage of research method

A. Dataset

In this section, the things that will be discussed are about the collection of information that will be used. The dataset used for this research is the Fashion-MNIST Dataset, which is the same as the dataset used in the previous research [11], [18]. This dataset is also available on the open repository site, which can be seen in detail in references [18].

The Fashion-MNIST dataset contains grayscale images of fashion objects, each measuring 28x28 pixels. This dataset has 4 files with labels and images. The files are divided into training and testing sets. You can find more information about what is inside the dataset in Table I. There are 60,000 images in the training set with labels, and the test set has 10,000 images. Fashion MNIST is a collection of various types of clothing images. There are 10 categories of clothing, and each category has a corresponding label and description explaining what it represents [11]. Detailed information is given in Fig. 2 [18].

B. Data Augmentation

The data augmentation process helps by adding more pictures to the collection of data. This facilitates preventing or reducing the problem of learning the model from only certain images and increases the accuracy of the model in classifying different images. In this research, augmentation is performed using the ImageDataGenerator, a preprocessing function class available in the Tensorflow library. The ImageDataGenerator is utilized with specific settings, including a rotation_range of 10 degrees, enabling horizontal_flip, and using 'nearest' as the fill_mode for image augmentation.

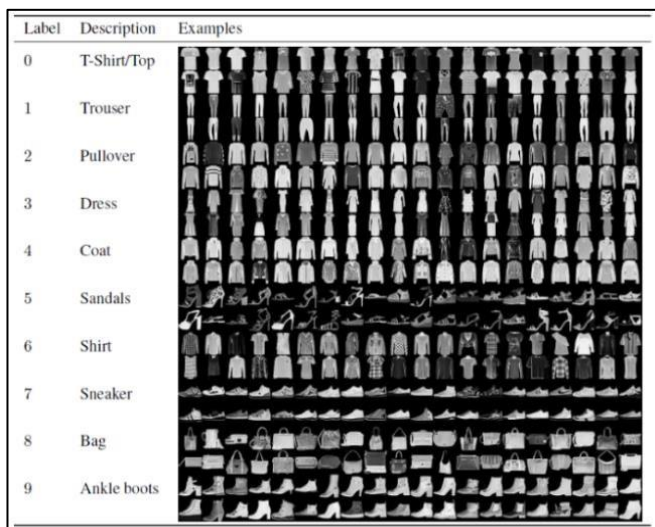


Fig. 2 Samples used along with Fashion-MNIST classes

TABLE I
FASHION-MNIST DATASET CLASS AND SIZE

Name	Description	Sample	Size
train-images-idx3-ubyte.gz	Training set images	60.000	46 MBytes
train-labels-idx1-ubyte.gz	Training set labels	60.000	10 KBytes
t10k-images-idx3-ubyte.gz	Test set images	10.000	8 MBytes
t10k-labels-idx1-ubyte.gz	Test set labels	10.000	10 MBytes

The dataset contains many variants in the form of image frames, which is necessary for adding arguments/parameters. In rotation_range=10, it means to rotate each image by 10 degrees randomly, to direct its unique pattern for better data augmentation. horizontal_flip=True, ImageDataGenerator provides to flip each dataset image horizontally or vertically means, in this case it is to flip each dataset image horizontally, to make the generated augmentation consistent from all images that are flipped fixedly instead of using the default randomly. Meanwhile, for fill_mode='nearest', it means that the default argument/parameter of fill_mode will fill the area with the nearest pixels and stretch it to handle inconsistent generated pixels which in this case, might come from the rotation range process and horizontal flip process.

When augmentation is not applied, the model's accuracy is lower compared to when augmentation techniques are used. The occurrence of overfitting is also a concern in the absence of augmentation. The distinction between these two scenarios is primarily attributed to the characteristics of the dataset, which play a crucial role in obtaining the optimal model accuracy. Data augmentation aims to adapt and enhance these characteristics by manipulating certain aspects of the dataset. This improvement can be observed in Table 5, where the comparison of accuracy values clearly demonstrates the positive impact of data augmentation on the model's performance.

C. Preprocessing

The dataset used in this research was obtained publicly [19]. Data preprocessing is carried out prior to the training phase with the objective of transforming raw data into a prepared, structured, and usable format [19].

The process performed during preprocessing is scaling the pixels between 0.0 - 1.0 and adding augmented images. Details of the preprocessing are shown in Table II.

D. Convolutional Neural Network

CNN is a category of neural network that possesses layers. It is considered to be a very popular and traditional framework used for deep learning [2]. CNN can be used to find and identify things in a image. CNN uses a method called convolution. It moves a filter of a certain size over an image. This process enables the computer to gather new information about the image by performing element-wise multiplication between specific parts of the image and the filter. CNN is a method that can easily recognize and group things without needing to do much beforehand. It can also understand visual pictures and easily pick out the important parts that make up a picture, such as the layers of information. [19].

Each layer in the Convolutional Neural Network performs distinct tasks with the input data. Within the convolutional layer, filters are applied to identify significant attributes within the. The pooling layer, on the other hand, performs either max pooling or average pooling, which involves finding the highest value or calculating the average value within specific regions of the image. Lastly, the fully connected layer gathers information from various features extracted from the image and plays a crucial role in determining its final classification [19]. Based on this research, using the CNN method that takes images from the Fashion-MNIST dataset as input and then processed in convolution and pooling layers, which produces output in the fully connected layer. Based on research [21] the parameters used result in an accuracy of 0.9399 where these results are quite good. But in this research the parameters used are as in Table III.

TABLE II
PREPROCESSING DATA

Data	Description	Total
Train	target_size = (28,28) batch_size = 1, subset = 'training'	48000
Validation	target_size = (28,28) batch_size = 1, subset = 'validation'	12000

TABLE III
DETAILS HYPERPARAMETER ON CNN MODEL

Layer	Output Shape	Parameter
BatchNormalization	(None, 28, 28, 1)	4
Conv2D	(None, 28, 28, 64)	1088
MaxPooling2D	(None, 14, 14, 64)	0
Dropout	(None, 14, 14, 64)	0
Conv2D	(None, 11, 11, 64)	65600
MaxPooling2D	(None, 5, 5, 64)	0
Dropout	(None, 5, 5, 64)	0
Flatten	(None, 1600)	0
Dense	(None, 256)	409856
Dropout	(None, 256)	0
Dense	(None, 64)	16448
BatchNormalization	(None, 64)	256
Dense	(None, 10)	650
Total params		493,902
Trainable params		493,772
Non-trainable params		130

Table III shows the architecture model and the layers that make up the architecture, including: (1) BatchNormalization, (2) Conv2D, (3) MaxPooling2D, (4) Dropout, (5) Flatten, (6) Dense. The Convolutional Neural Network (CNN) structure used as the architecture in this study uses a sequential model. In the proposed model, the initial step involves BatchNormalization, a technique used to mitigate the issue of covariance shift. It helps in equalizing the distribution of each input value, which tends to change due to variations in the previous layers during the process. The normalized image dimensions have an image format of a two-dimensional array (28 x 28 pixels). After normalization, the convolution process and Pooling Layer are performed. Conv2D is the hard layer used in this architecture. The model starts with a Conv2D layer with a filter parameter of 64. The pooling layer is used to decrease the dimensionality of the feature map, which causes its resize. Following the pooling layers, the dropout technique is applied. Dropout is used several times in this architecture. The first dropout is done after normalization, one Conv2D layer, and Pooling Layers with a dropout of 0.1. The second dropout of 0.3 is done after one Conv2D layer and Pooling Layers. The last dropout of 0.5 is done after Flatten and Dense. Flatten is used to convert 3D features into 1D vector features. Dense is done three times. The first and second Dense use parameters 256 and 64 and activation relu. Then normalization is carried out before the last solid is carried out using parameter 10, based on the number of classes of folders in the dataset. In the third solid layer,

the activation used is no longer relu but in the form of softmax.

E. Evaluation

In the final phase of classification for all classes, the model testing results are evaluated using metrics like Accuracy, Precision, Recall, and F1-Score to evaluate the performance on the Fashion-MNIST dataset and its respective classes during the validation process. Precision, Recall, and F1-Score are computed using the following calculation algorithm, which provides a reliable measure of the model's classification performance:

$$\text{Precision} = \frac{\text{True Positives}}{\text{True Positives} + \text{False Positives}} \quad (1)$$

$$\text{Recall} = \frac{\text{True Positives}}{\text{True Positives} + \text{False Negatives}} \quad (2)$$

$$\text{F1Score} = \frac{2 * \text{Precision} * \text{Recall}}{\text{Precision} + \text{Recall}} \quad (3)$$

III. RESULT AND DISCUSSION

This section talks about what happened during the discussion of the proposed method. It also discusses how the data is processed and ends with an evaluation using terms such as accuracy, recall, precision, and F1-Score. After evaluating, we will make a comparison between the accuracy value and earlier studies that used the same dataset. We will also compare the accuracy value when augmentation is applied compared to without augmentation.

A. Accuracy, Recall, F1-Score

Based on the references used as references in this study [6], [11], [12], [14], [15] as a comparison of accuracy results in classifying fashion using the Fashion-MNIST dataset. The specifics of the comparison are presented within Table IV.

TABLE IV
COMPARISON OF TEST ACCURACY BASED ON THE SAME DATASET

Model	Test Accuracy
HOG + SVM	0.8653
SVC	0.8970
EDEN	0.9060
CNN2	0.9117
VGG-11	0.9150
CNN Regression	0.9200
CNN2 + BN + Skip	0.9254
CNN4 + HPO + Reg	0.9399
CNN2 + Augmentation	0.9592

TABLE V
COMPARISON BETWEEN THE PROPOSED MODEL WITH AUGMENTATION AND THE PROPOSED MODEL WITHOUT AUGMENTATION

Use of Augmentation	Accuracy
With Augmentation	0.9592
Without Augmentation	0.9390

Table V displays the variation in outcomes with and without using the augmentation process. When we use augmentation, we get an accuracy of 95.92%. Without augmentation, the accuracy is 93.90%.

After performing various steps in the method we used before, the results of the performance calculation are displayed in Table VI. The average class achieves high accuracy values for precision, recall, and F1-Score. If in previous research [11] the class of 'shirts', 'pullover', and 'coat' had very low accuracy scores of 612, 765, and 796. In this reesearch, we found that the accuracy value of the class improved and gave better results. Specifically, the accuracy values were 783, 894, and 901. You can find more information in Table VI.

Generate Accuracy, Recall, and F1-Score results and test the results of the models that have been created in To assess the performance of machine learning models, accuracy, recall, and F1-Score measurements are carried out. The results obtained from testing the model with training data are saved in the historical variable. The Classification Report is used to determine the percentage of images that the model correctly classified in test data, providing insigh valuable about its classification performance [20]. To evaluate a classification task's performance, the Confusion Matrix, as shown in Fig. 3, serves as a valuable tool. This matrix provides a visual representation of the alignment between predicted values and actual values, offering insight into the accuracy and precision of the test data classification.

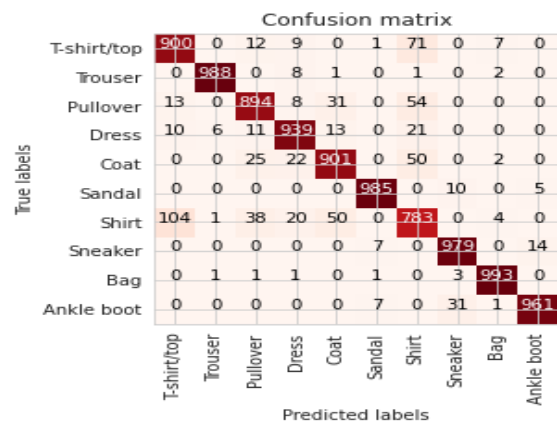


Fig. 3 Confusion matrix model

TABLE VI
PERFORMANCE DATA BY APPLYING AUGMENTATION AND WITHOUT AUGMENTATION

Class	Precision		Recall		F1-Score	
	Aug.	Non.Aug.	Aug.	Non.Aug.	Aug.	Non. Aug.
Class 0 (T-shirt/top)	0.90	0.87	0.85	0.91	0.88	0.89
Class 1(Trouser)	0.99	1.00	0.99	0.99	0.99	0.99
Class 2 (Pullover)	0.91	0.91	0.90	0.93	0.91	0.92
Class 3 (Dress)	0.94	0.93	0.94	0.96	0.94	0.94
Class 4 (Coat)	0.90	0.93	0.92	0.89	0.91	0.91
Class 5 (Sandal)	0.99	1.00	0.99	0.98	0.99	0.99
Class 6 (Shirt)	0.79	0.84	0.82	0.80	0.80	0.82
Class 7(Sneaker)	0.96	0.96	0.99	0.98	0.97	0.97
Class 8 (Bag)	0.99	0.99	0.99	0.99	0.99	0.99
Class 9 (Ankle boot)	0.99	0.98	0.96	0.97	0.98	0.97
Accuracy					0.94	0.94
Macro avg	0.94	0.94	0.94	0.94	0.94	0.94
Weighted avg	0.94	0.94	0.94	0.94	0.94	0.94

B. Evaluation Result Chart

In Fig. 4, the training data is depicted by a blue line, while the validation data is represented by an orange line. The graph clearly shows that the accuracy of the validation data begins at 87.90% and shows a consistent improvement over time. The peak accuracy of 96% is reached at epoch 94. With the progression of epochs, there is a noticeable improvement in the accuracy of the training data. By epoch 99, the highest achieved accuracy is 96.71%.

In Fig. 5, there is a graph that shows the loss of the model on the validation data. The orange line reaches a loss value of 0.3152 at the 1st epoch, and the best loss value of 0.1125 occurs at the 94th epoch. On the other hand, the loss on the train data is smaller and it reaches a minimum value of 0.1100 at the 60th epoch, with the best value of 0.0894 occurring at the 99th epoch. You can find this on the blue line in Fig. 5.

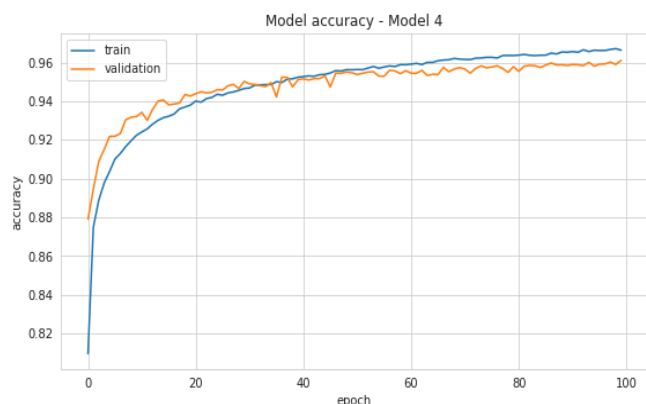


Fig. 4 Graph of model accuracy

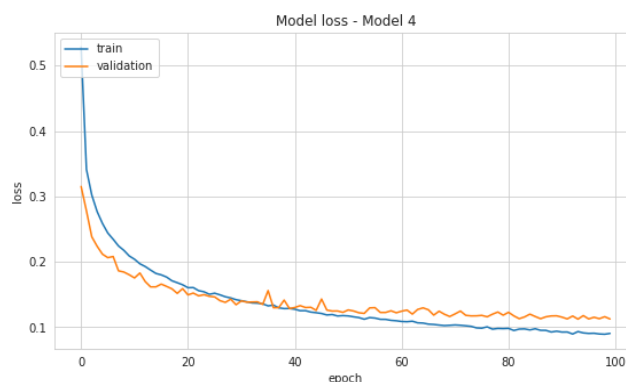


Fig. 5 Loss Model Graph

IV. CONCLUSION

Utilizing the CNN algorithm and data augmentation during the training dataset for fashion image classification significantly improves the accuracy in recognizing and categorizing images. This research used a set of pictures showing different fashion objects in black and white. There are 10 categories of objects, including ankle boots, bags, coats, dresses, pullovers, sandals, shirts, sneakers, t-shirts, and trousers. These pictures are used to train and test a computer program. Using the augmentation process can help reduce overfitting. By using the Image Data Generator and adding a Dense Layer called the Fully Connected Layer to our model, we can improve the accuracy of the training data by 96.43%. This also results in a testing data accuracy of 95.92% after 100 epochs. In the absence of augmentation, the accuracy of the testing data might decline to 93%. In more straightforward terms, this study yielded highly favorable outcomes, as the precision of the test data surpassed that of the cited publication. The

discrepancies in results can be attributed to the varying data analysis techniques employed in both studies.

REFERENCES

- [1] Y. Zhong and S. Mitra, "The role of fashion retail buyers in China and the buyer decision-making process," *J. Fash. Mark. Manag. An Int. J.*, vol. 24, no. 4, pp. 631–649, 2020, doi: <https://doi.org/10.1108/JFMM-03-2018-0033>.
- [2] M. Xin and Y. Wang, "Research on image classification model based on deep convolution neural network," *Eurasip J. Image Video Process.*, vol. 2019, no. 1, 2019, doi: [10.1186/s13640-019-0417-8](https://doi.org/10.1186/s13640-019-0417-8).
- [3] K. V. Madhavi, R. Tamilkodi, and K. J. Sudha, "An Innovative Method for Retrieving Relevant Images by Getting the Top-ranked Images First Using Interactive Genetic Algorithm," *Procedia Comput. Sci.*, vol. 79, pp. 254–261, 2016, doi: [10.1016/j.procs.2016.03.033](https://doi.org/10.1016/j.procs.2016.03.033).
- [4] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "ImageNet classification with deep convolutional neural networks," *Commun. ACM*, vol. 60, no. 6, pp. 84–90, 2017, doi: [10.1145/3065386](https://doi.org/10.1145/3065386).
- [5] A. S. Henrique *et al.*, "Classifying Garments from Fashion-MNIST Dataset Through CNNs," *Adv. Sci. Technol. Eng. Syst. J.*, vol. 6, no. 1, pp. 989–994, 2021, doi: [10.25046/aj0601109](https://doi.org/10.25046/aj0601109).
- [6] S. Bhatnagar, D. Ghosal, and M. H. Kolekar, "Classification of fashion article images using convolutional neural networks," *2017 4th Int. Conf. Image Inf. Process. ICIIIP 2017*, vol. 2018-Janua, no. 1, pp. 357–362, 2018, doi: [10.1109/ICIIP.2017.8313740](https://doi.org/10.1109/ICIIP.2017.8313740).
- [7] K. Yamaguchi, M. H. Kiapour, L. E. Ortiz, and T. L. Berg, "Parsing clothing in fashion photographs," in *2012 IEEE Conference on Computer vision and pattern recognition*, 2012, pp. 3570–3577, doi: [10.1109/CVPR.2012.6248101](https://doi.org/10.1109/CVPR.2012.6248101).
- [8] H. Chen, Z. J. Xu, Z. Q. Liu, and S. C. Zhu, "Composite templates for cloth modeling and sketching," in *2006 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'06)*, 2006, vol. 1, pp. 943–950, doi: [10.1109/CVPR.2006.81](https://doi.org/10.1109/CVPR.2006.81).
- [9] L. Bossard, M. Dantone, C. Leistner, C. Wengert, T. Quack, and L. Van Gool, "Apparel classification with style," in *Computer Vision—ACCV 2012: 11th Asian Conference on Computer Vision, Daejeon, Korea, November 5-9, 2012, Revised Selected Papers, Part IV 11*, 2013, pp. 321–335, doi: https://doi.org/10.1007/978-3-642-37447-0_25.
- [10] K. V. Greeshma *et al.*, "Classification of Garments from Fashion MNIST Dataset Using CNN LeNet-5 Architecture," *Proc. 2020 Int. Conf. Innov. Trends Commun. Comput. Eng. ITCE 2020*, vol. 8, no. 2019, pp. 357–362, 2019, doi: [10.1109/ICIIP.2017.8313740](https://doi.org/10.1109/ICIIP.2017.8313740).
- [11] K. V. Greeshma and K. Sreekumar, "Fashion-MNIST classification based on HOG feature descriptor using SVM," *Int. J. Innov. Technol. Explor. Eng.*, vol. 8, no. 5, pp. 960–962, 2019, Retrieval Number: E3075038519/19©BEIESP.
- [12] K. V. Greeshma and K. Sreekumar, "Hyperparameter optimization and regularization on fashion-MNIST classification," *Int. J. Recent Technol. Eng.*, vol. 8, no. 2, pp. 3713–3719, 2019, doi: [10.35940/ijrte.B3092.078219](https://doi.org/10.35940/ijrte.B3092.078219).
- [13] M. Kayed, A. Anter, and H. Mohamed, "Classification of Garments from Fashion MNIST Dataset Using CNN LeNet-5 Architecture," *Proc. 2020 Int. Conf. Innov. Trends Commun. Comput. Eng. ITCE 2020*, no. June, pp. 238–243, 2020, doi: [10.1109/ITCE48509.2020.9047776](https://doi.org/10.1109/ITCE48509.2020.9047776).
- [14] C. Duan, P. Yin, Y. Zhi, and X. Li, "Image Classification of Fashion-mnist Data Set Based on VGG Network," *Proc. 2019 2nd Int. Conf. Inf. Sci. Electron. Technol. (ISET 2019). Int. Informatiz. Eng. Assoc. Comput. Sci. Electron. Technol. Int. Soc.*, vol. 19, no. Iset, pp. 18–21, 2019.
- [15] G. Ngo, V. Gaurav, and T. Shibata, "Clothing extremity identification using convolutional neural network regressor," *2018 Jt. 7th Int. Conf. Informatics, Electron. Vis. 2nd Int. Conf. Imaging, Vis. Pattern Recognition, ICIEV-IVPR 2018*, pp. 436–441, 2019, doi: [10.1109/ICIEV.2018.8641046](https://doi.org/10.1109/ICIEV.2018.8641046).
- [16] D. Alamsyah and D. Pratama, "Implementasi Convolutional Neural Networks (CNN) untuk Klasifikasi Ekspresi Citra Wajah pada FER-2013 Dataset," *J. Teknol. Inf.*, vol. 4, no. 2, pp. 350–355, 2020, doi: <https://doi.org/10.36294/jurti.v4i2.1714>.
- [17] W. M. Pradnya D and A. P. Kusumaningtyas, "Analisis Pengaruh Data Augmentasi Pada Klasifikasi Bumbu Dapur Menggunakan Convolutional Neural Network," *J. Media Inform. Budidarma*, vol. 6, no. 4, p. 2022, 2022, doi: [10.30865/mib.v6i4.4201](https://doi.org/10.30865/mib.v6i4.4201).
- [18] H. Xiao, K. Rasul, and R. Vollgraf, "Fashion-MNIST: a Novel Image Dataset for Benchmarking Machine Learning Algorithms," pp. 1–6, 2017, [Online]. Available: <http://arxiv.org/abs/1708.07747>.
- [19] E. Rasywir, R. Sinaga, and Y. Pratama, "Analisis dan Implementasi Diagnosis Penyakit Sawit dengan Metode Convolutional Neural Network (CNN)," *J. Parad. Ubsi*, vol. 22, no. 2, pp. 117–123, 2020, doi: <https://doi.org/10.31294/p.v21i2>.
- [20] A. Taherkhani, G. Cosma, and T. M. McGinnity, "AdaBoost-CNN: An adaptive boosting algorithm for convolutional neural networks to classify multi-class imbalanced datasets using transfer learning," *Neurocomputing*, vol. 404, pp. 351–366, 2020, doi: [10.1016/j.neucom.2020.03.064](https://doi.org/10.1016/j.neucom.2020.03.064).
- [21] G. A. Sandag, J. Waworundeng, U. Klabat, J. A. Mononutu, and A. M. Utara, "Identifikasi Foto Fashion Dengan Menggunakan Convolutional Neural Network (CNN) Identify Fashion Images Using Convolutional

Neural Network (CNN),” vol. 7, no. 2, pp. 305–314, 2021, doi: <https://doi.org/10.31154/cogito.v7i2.340.305-314>.