Classification Fashion Image Using Local Binary Pattern and Artificial Neural Network Multi Layer Perceptron

Muhamad Fajar Rivaldy[1], Febryanti Sthevanie[2], Anditya Arfianto[3]

[1,2,3]Fakultas Informatika, Universitas Telkom, Bandung
[1]fajarrivaldy@students.telkomuniversity.ac.id, [2]sthevanie@telkomuniversity.ac.id,
[3]anditya@telkomuniversity.ac.id,

Abstrak

Busana merupakan kebutuhan pokok yang selalu manusia pakai setiap hari. Pakaian digunakan manusia sebagai pelindung dan sebagai penutup tubuh. Namun seiring perkembangan kehidupan, pakain dijadikan simbol kedudukan, status, jabatan dan sebagai pakain untuk beraktivitas. Karenanya jenis busana bermacam-macam dimulai dari pakaian untuk tidur sampai pakain untuk bekerja. Sistem klasifikasi dibutuhkan dalam untuk bisa mengkategorikan busana terutama bagi e-commerce dan lain-lain. Pada sistem yang kami rancang, Terdiri dari tiga tahapan antara lain: Preprocessing, Fitur Ekstraksi dan Klasifikasi. Pada tahap Preprocessing data akan di resize kemudian dikonversi menjadi grayscale. Kemudian pada tahap Fitur ekstraksi gambar akan di diproses dengan LBP atau Local Binary Pattern setelah itu citra akan di menjadi 2 yaitu citra training dan citra untuk testing. Citra untuk Training akan dimasukan terhadap model JST, model JST akan belajar dari data training untuk menemukan pola pada data tersebut. Setelah tahapan training selesai citra akan di tes untuk mengetahui akurasi terhadap data yang belum pernah di dipelajari dengan menggunakan data test sebagai pengujian citra yang akan diidentifikasi. Mencapai nilai akurasi 96.38%

Kata kunci :Image Processing, Jaringan Saraf Tiruan, Classification, Fashion Classification, Local Binary Pattern

Abstract

Clothing Fashion is a basic need that humans always use every day. Clothing is used by humans as a protector and as a body covering. But along with the development of life, Fashion are used as symbols of position, status, position and as clothes for activities. Therefore various types of clothing ranging from clothing for sleeping to clothes for work. A classification system is needed in order to be able to categorize clothing especially for e-commerce and others. In the system that we designed, it consists of three stages including: Preprocessing, Feature Extraction and Classification. At the Preprocessing stage the data will be resized and then converted to grayscale. Then in the image extraction feature will be processed with LBP or Local Binary Pattern after that the image will be in 2, namely training images and images for testing. The image for the training will be added to the ANN model, the ANN model will learn from the training data to find patterns in the data. After the training stage is completed the image will be tested to determine the accuracy of the data that has never been studied by using test data as an image test that will be identified. Reach 96.38% accuracy.

Keywords: Image Processing, Artificial Neural Network, Classification, Clothing Classification, Local Binary Pattern

1. Introduction

Background

Fashion in various cultures reflects the characteristics of age, social status, lifestyle and gender. Fashion clothing is also one of the important descriptions in identifying humans as an example of ”men who wearing an orange jacket.” or ”women who wear yellow high heels”. Clothing classification has many roles in its application, for example, predicting fashion details in unlabeled pictures that can facilitate finding similar items [9] in e-commerce database. Similarly, the classification of images that are preferred by users can encourage automatic image management, which will provide recommended fashion clothing recommendations based on the style desired by the user. fashion classification can also automatically annotate images with tags or descriptions related to images.

Based on a particular fashion classification system, system will focus on image prediction, with the aim of removing tags from images that have similarities or are identical to the data in the training data. Making this system will be useful for classifying and making a recommendation system. The way this system works is to examine each pixel in an image and identify the closest pixel. In the system clothing classification created using Local Binary Pattern image processing approach and machine learning Artificial Neural Networks Multi Layer Perceptron. In previous tests the ANN method can classify the image with the results of 89 % accuracy [11].

In the approach using machine learning, there are several methods that can be done to classify clothing one method that will be used is Artificial Neural Networks [12].

Topic and Boundary

In the background in this study, the topic of the problem being solved is how to classify clothing by method Local Binary Pattern with classifier Neural Network.

Limitation problems in this study are as follows: dataset taken from www.kaggle.com [1], classes used in the data set are 8 classes, the data used has a white background with objects in the middle of the image, the size of the image used is 80x60 pixels.

Purpose

In designing this system, it is expected to be able to build models and implement Artificial Neural Networks and Local Binary Pattern extraction features in classifying and producing tags on images using the *Local Binary Pattern* and Artificial Neural Networks method.

Writing Organization

Organization of writing in this thesis include: Identification of Problems where the problem will be identified about how to classify clothing images using local binary patterns and artificial neural networks. Literature Study, looking for references related to fashion as a theoretical basis in providing solutions to problems that occur. Data Collection, collecting data as a data set and test data. Design Design, describes how the workflow processes the classification of fashion images. Analysis of Research Results on whether the output is in accordance with the objectives. Report Writing of the stages that have been carried out as a result of the solution of the problem.

2. Related Study

2.1 Previous Research

Brian Lao and other conducted research about Fashion Classification using CNN, to annotate image and discover similar fashion items to a fashion item in a query image [8]. obtained accuracy 74.5%.

Muhamed Kayed and other conducted research about Classification of Garments from Fashion MNIST Dataset Using CNN LeNet-5 Architecture [7] and has simillar dataset, Kayed comparing accuracy result with other neural network architecture and other classifier. Using MLP accuracy reach 78.33%, using SVM kernel rbf reach 89.70%.

Raynaldi Fatih Amanullah and other conducted research about ”Detection batik motive using features Extraction and Artificial Neural Network” it use LBP and ANN reach accuracy of 74% [2].

Tasnuva Hassan and other conducted research about ”Handwritten Bangla Numeral Recognition usingLocal Binary Pattern” Tasnuva Hassan use KNN as classifier and get result accuracy of 96.70% [4].

2.2 The Need for a Fashion Image Classification System

Fashion clothing is a term of goods worn by the body. Based on the information obtained it has been concluded that clothes have been known since 42,000-72,000 years ago [10].

The need for a system that can recommend to the desires of the user is increasingly evidenced by advertising that seems to be able to guess the desires of users in fashion products. The classification of clothing types is a matter of predicting multiple classes of a single label as described in an image [3]. Clothing classification can also provide recommendations on the types of goods contained in the history that has been recorded.

2.3 Change the RGB image to Grayscale

Converting an RGB image to grayscale is needed in order to improve performance and speed [14]. Because in this case what is needed is the state of the geometry in an image.

Conver RGB into Grayscale image using formula in equation 1 below :

$$Grayscale = ((0.3XR) + (0.59XG) + (0.11XB)) \tag{1}$$

using formula1 will generate grayscale image.

2.4 Local Binary Pattern (LBP)

Local Binary Pattern or LBP is a simply texture descriptor yang and efficient using threshold on every pixel neighbor and consider resulting biner value [6].

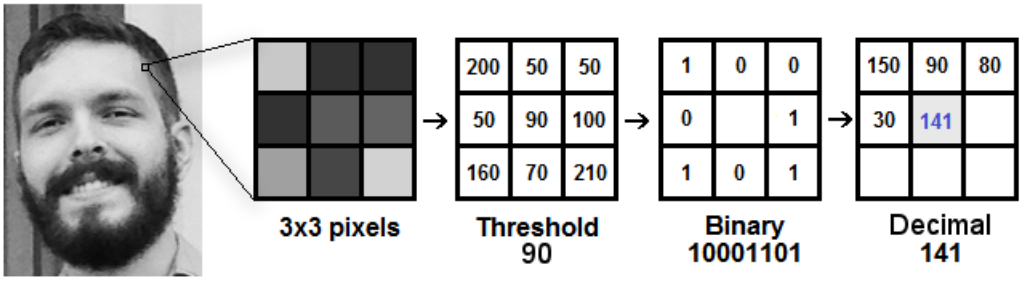

Figure 1. Ilustration LBP

The image used in LBP is a grayscale image with a 3x3 neighbor, where the pixel value in the middle position is made into a threshold. Can be obtained in the following way :

$$LBP_{pr} = \sum_{p-0}^{p-1} s(g_c - g_p)2^p, s(x)\{1 \quad x \geq 00 \quad x < 0 \tag{2}$$

The pixels found in the neighbor are then compared to the threshold, according to the equation 2, if the value of the neighbor pixel is greater than the threshold. Then, the resulting value will be 1 if not, the resulting value will be 0. After that the binary value for each pixel will be multiplied by 2n, where n is the sequence of neighboring pixels from the top left corner around the pixel in the middle position in the direction from the left top to bottom right in the order starting from 0 - 7, the results of the multiplication are then added together.

2.5 Normalization

Normalization is a process that changes the range of pixel intensity values [3], before entering the dataset model first from RGB to Grayscale aims to speed up training in the model, producing an array of one-dimensional pixels so that the pixels have integer values from 0 to 255. Then the pixels normalization divided by 255 to produce a decimal value between 0 and 1.

2.6 Artificial Neural Network

Artificial Neural Networks are defined as a processing system that is inspired by the workings of the human nervous system [5, 13]. This system was created as a generalization of a mathematical model of human understanding. Artificial neural networks are known as black boxes or are not transparent because ANN cannot explain how an outcome is obtained. ANN consists of input layer, hidden layer and output layer. In the ANN system that is built will be used Multi Layer Perceptron or MLP architecture. MLP architecture is considered better at handling large amounts of data. Can be seen from the Figure 2.

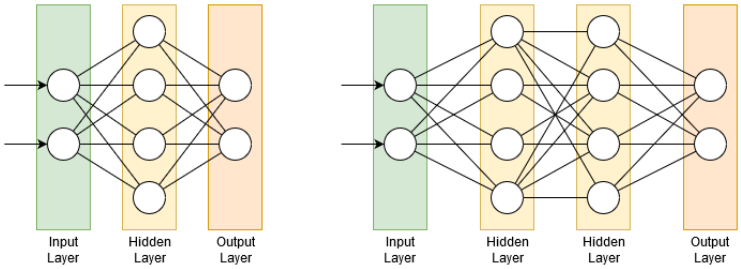

Figure 2. Architecture ANN and MLP [11]

1. Activation Function

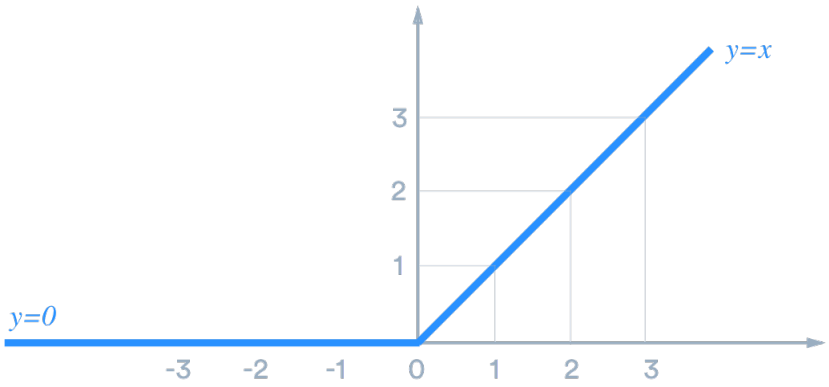

Figure 3. ReLU

Activate neurons in the Multi Layer Perceptron model that will determine whether a hidden neuron will be active or inactive in receive input. Activation function that will be used in hidden neurons is Rectified Linear Unit (ReLU), ReLU produce a linear value for input with a Positive value and 0 for negative value input, can be seen in the ReLU graph in the equation 3. ReLU can be used with formulas :

$$f(x) = max(0,x) \tag{3}$$

3. System Design

3.1 System Plan

In the design of development system to classify clothing utilizing digital image processing. there are 3 main part such as preprocessing, feature extraction and classification. Flowchart show in Figure 4. Image will go through preprocessing to resize and convert RGB to Grayscale image, image after grayscaling will input to feature extraction in order to extract features and then use it as input to classifier. In classifier section image after extraction feature will classify to determine the class of the image throug traing process.

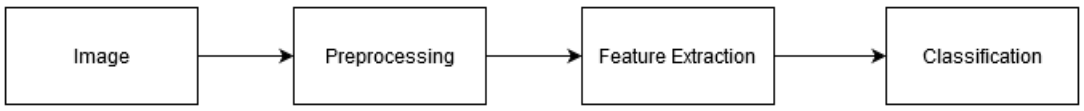

Figure 4. Flow Chart Sistem

3.2 Preprocessing

Before the training phase the data to be input to the preprocessing stage first, there are several stages in preprocessing including:

1. Image Resizing

   This resizing process aims to make the image the same size so that the image can be used in the training stage. The image will be resized to 80x60 pixels for all images.

2. Changing RGB Image into Grayscale

   The process of converting an RGB image into Grayscale aims to change the dimensions of the image matrix and improve performance at the processing stage because the matrix in grayscale has a dimension of 1 dimension.

3.3 Feature Extraction

feature extraction stage *Local Binary Pattern* the image will be divided into blocks, then the pixel in the middle will become the threshold to be used as a reference in LBP calculations so that if the neighbor is greater than the threshold, the resulting value will be 1 and if not the resulting value will be 0 [6], then the results of the calculation will be multiplied by 2n then the results of the calculation will be added together.

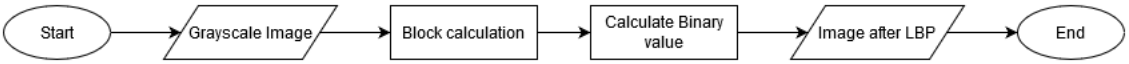

Figure 5. LBP data flow

3.4 Normalization

In normalization data after feature extraction stage every pixel will divided by 255 to aim speed up training in the model. and the result is a decimal value between 0 and 1.

3.5 Classification

The training method used on this system is the Artificial Neural Network method. Inputs that have been previously processed at the ANN stage will be entered into the ANN model. Each pixel matrix in the image will be input against the input layer. Input layer is obtained by multiplying image length and image width. Each data matrix that has been inputted will be calculated on the hidden layer with the calculation that was discussed in the previous chapter. After the calculation phase in the hidden layer, the results of the input will go to the output layer, the output layer has the same number of classes in the dataset in the training data, therefore, the output layer will produce the possibilities of the classes in the training data.

The results of these outputs will be compared with the actual outputs, i.e. the actual class images. The resulting error will be used again to do backpropagation which will correct it to produce a smaller error. This process is carried out on all data entered. Then it will be repeated again starting from the initial data to the final data based on the epoch specified at the training stage.

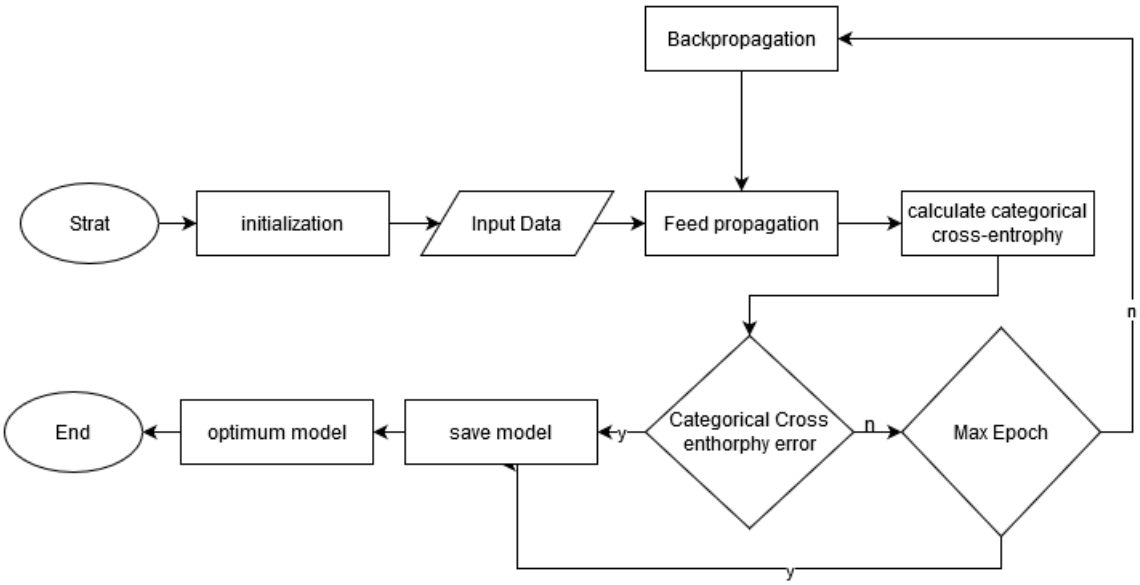

Figure 6. ANN Train Flow Diagram

After optimum model generated by training section, then model need to be test with data that never included to model. The flow of testing is different from training because only required feed forward from model already generate from training section.

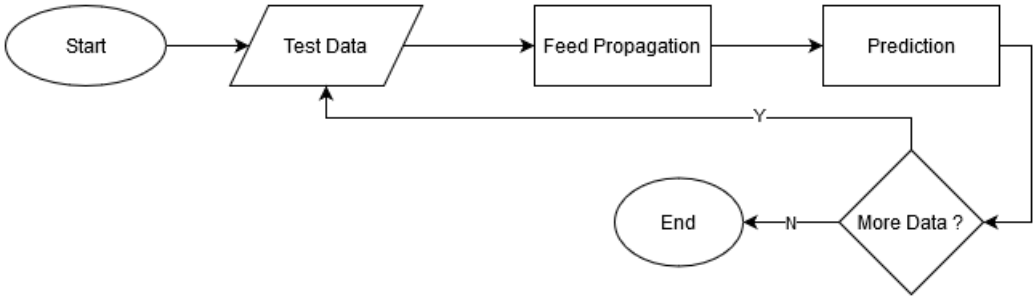

Figure 7. ANN Test Flow Diagram

4. Evaluation

4.1 System Perfomance Measurement

To measure the accuracy of this system, it is done by the method of confusion matrix or error matrix where the confusion matrix is a specific table that allows visualizing the performance of the algorithm itself.

Table 1. Table Confusion Matrix

		Actual	
		Positive	Negative
Prediction	Positive	TP	FP
	Negative	FN	TN

Confusion matrix using the value of TP (True Positive), FP (False Positive), FP (False Negative), TN (True Negative) values obtained from accuracy. Accuracy is obtained by using equation 4.

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} x100 \quad (4)$$

Accuracy can measure how many predictions are correct by the classifier with a distance between 0-100% of the guesswork from the classifier.

4.2 Test Result

Concept of ANN is more trained the accuracy better. Therefore, Selecting best hyperparameter is required in order obtained best combination of hyperparameter. train Testing of neural network parameters. In a system built using artificial neural networks as a classifier using Multi Layer Perceptron (MLP) architecture. In the test scenario carried out by changing value of radius, sampling point, hidden neuron and learning rate.

Testing is executed by changing the feature extraction parameters and parameters in the classifier in order to get better accuracy in the system testing. Any changes to the feature extraction parameters and classifier are presented in tabular form. Sampling Point and Radius on LBP. In the first scenario, to determine the effect of parameters on feature extraction Local Binary Pattern done by changing parameter sampling point and the radius found on LBP. Executed by 10 epoch in order to found best hyperparameter. with learning rate 0.0001 and amount of node in hidden layer 1 as much 32 nodes and nodes in hidden layer 2 as much 32 nodes. Obtained the accuracy results shown in table 2 using the training image and test image as follows:

Table 2. Scenario testing of radius and Sampling point parameters on LBP

No	Radius	Sampling Point	Accuracy(%)
1	1	4	85,09
1	1	8	43,08
2	2	4	79,83
2	2	16	12,50

Then selecting hyperparameter of LBP is sampling point and radius scheme tested in using certain value such as radius 1,1,2,2 and sampling point 4,8,4,16. In Table 2 showed best accuracy at scheme radius 1 and sampling point 4. choosing LBP scheme is affect accuracy obtained.

Testing of ANN Learning Rate parameters. In this test carried out by changing the learning rate parameter on artificial neural network classifier with a value between 0 and 1, in this test if the value of the learning rate parameter is large, the learning distance from the classifier will be greater and vice versa the smaller the value of the learning rate will be the smaller the distance of learning distance makes the classifier more careful in studying images.

The purpose of this test is to find the most optimal accuracy value so that it gets a good accuracy model. Results of experiments that have been carried out shown in the following Table 3. parameters used in testing learning rate such as, radius 1 and sampling point 4, epoch 10, layer 1 has 32 node and layer 2 has 32 node. the combination of higher accuracy model will be use to next scenario.

Table 3. Result of hyperparameter learning rate

No	Learning Rate	Accuracy(%)
1	0,01	59,94
2	0,001	85,00
3	0,0001	78,96
4	0.00001	34,37

Testing of hidden nueron ANN parameters testing use with parameters sampling point 4, radius 1, epoch 10 and learning rate 0.001 In this MLP scenario two hidden layers are used with each hidden layer having its own number of nodes. the results obtained with accuracy in the table below.

Experiments with the second scenario obtained with the results shown in Table 4, with the highest accuracy obtained by experiments with layers 1 and 2 with the number of layer 1 128 nodes and layer 2 256 nodes. Based on Table 4, the number of nodes or neurons that is small or large does not make a good classifier needs to be done trial and error. Scheme 10 have best accuracy than other scheme in Table 4 used as hyperparameter.

Table 4. Result of Hidden Neuron

No	Nodes in layer 1	Node in layer 2	Accuracy(%)
1	8	8	44,08
2	8	16	35,96
3	16	16	63,74
4	16	32	75,68
5	32	32	87,82
6	32	64	89,78
7	64	64	91,14
8	64	128	91,11
9	128	128	91,95
10	128	256	93,33
11	256	256	92,51

4.3 Result Test Analysis

Result test with cross validation method on 1272 images using best LBP parameter 1 radius, 4 sampling point and best ANN scheme using learning rate 0.001, hidden neuron 1 128 and hidden neuron 2 256. Also parmeters k is 10 in cross validation through mean accuracy calculation in every fold. Accuracy obtain with train the model until model cannot developed again. Details accuracy showed on Table 5, obtained Train accuracy and accuracy test 96,38 %.

Table 5. Result using test data

No	Accuracy(%)	Test Accuracy(%)
1	97,87	96,38

From test data tested by model there are 8 classes such as, Bags, Footwear, Bottom Wear, wallet, Eyewear, Topwear, Belts and Watches each of those has 159 pieces of images. Based on Tabel 6, there are some data that is wrongly predicted, there 10 wrong predict as bags, 3 images as belt, 8 images as bottomwaer, 1 image predicted as Footwear, 12 images predicted as Topwear, 11 images predicted as wallet and 3 predicted as watch. For the detail showed in Table 6.

Table 6. Cofussion Matrix

Class	Bags	Belt	Bottomwear	Eyewear	Footwear	Topwear	Wallets	Watches
Bags (159)	149	2	3	0	0	1	4	0
Belt (159)	0	157	2	0	0	1	0	0
Bottomwear (159)	6	1	150	0	0	2	0	0
Eyewear (159)	0	0	0	159	0	0	0	0
Footwear (159)	0	0	1	0	158	0	0	0
Topwear (159)	2	0	4	0	0	148	0	5
Wallets (159)	8	2	1	0	0	0	148	0
Watches (159)	0	0	3	0	0	0	0	156

Wrong predicting in machine learning is it should happen, in image dataset there image has same shape pattern because Local Binary Pattern extract the shape and texture so that make system assume that has same class, most predictably wrong is Wallet and bags because there has similarity. Then, height of image is affected prediction in general lengthwise image will predicted as watch because watch image has more length than other images.

Table 7. Sample wrong predicted class

No	Original Image	feature Visualisation	Actual	Predicted
1			Topwear	Bags
2			Bags	Wallet
3			Wallet	Bags

From the data that has been analyzed the system will be wrong to predict if it has a similar shape because the system is designed to recognize the pattern of shapes in the image. Therefore, in table 7 the image with the bag class model is wrongly predicted and predicted as a wallet. Likewise with predictions that should be predicted as belts due to similar shapes then predicted as a wallet.

5. Conclusion

In the research that has been executed, the classification of clothing classification systems uses Local Binary Pattern and Artificial Neural Networks. Based on research and systems that have been made, choosing parameter of feature extraction and classifier is affected the accuracy, this system can classify the types of clothing by using the Local Binary Pattern extraction feature and Artificial Neural Networks with an accuracy obtained of 96.38%, with certain hyper parameter such as Sampling Point is 4, Radius is 8, Learning Rate is 0.001 and hidden neuron layer 1 is 128 nodes and layer 2 is 256 nodes.

For further research, the system can handle background images, the number of classes is more varied and can detect image objects so that the output on the system is more varied.

References

[1] Fashion product images (small) — kaggle. https://www.kaggle.com/paramaggarwal/fashion-product-images-small, 2019. (Accessed on 07/21/2020).

[2] Raynaldi Fatih Amanullah, Ade Pujianto, Bayu Trisna Pratama, and Kusrini Kusrini. Deteksi motif batik menggunakan ekstraksi tekstur dan jaringan syaraf tiruan. *CSRID (Computer Science Research and Its Development Journal)*, 10(2):69–79, 2019.

[3] C Rafael Gonzalez. Richard eugene woods. *Digital image processing*, 85, 2007.

[4] Tasnuva Hassan and Haider Adnan Khan. Handwritten bangla numeral recognition using local binary pattern. In *2015 International Conference on Electrical Engineering and Information Communication Technology (ICEEICT)*, pages 1–4. IEEE, 2015.

[5] Arief Hermawan. Jaringan saraf tiruan teori dan aplikasi. *Andi, Yogyakarta*, 2006.

[6] Di Huang, Caifeng Shan, Mohsen Ardabilian, Yunhong Wang, and Liming Chen. Local binary patterns and its application to facial image analysis: a survey. *IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews)*, 41(6):765–781, 2011.

[7] Mohammed Kayed, Ahmed Anter, and Hadeer Mohamed. Classification of garments from fashion mnist dataset using cnn lenet-5 architecture. pages 238–243, 2020.

[8] Brian Lao and Karthik Jagadeesh. Convolutional neural networks for fashion classification and object detection. *CCCV 2015: Computer Vision*, pages 120–129, 2015.

[9] Si Liu, Zheng Song, Guangcan Liu, Changsheng Xu, Hanqing Lu, and Shuicheng Yan. Street-to-shop: Cross-scenario clothing retrieval via parts alignment and auxiliary set. In *2012 IEEE Conference on Computer Vision and Pattern Recognition*, pages 3330–3337. IEEE, 2012.

[10] E Money-Kyrle, Roger. The psychology of clothes: By jc flügel, ba, d. sc.(the international psycho-analytical library. the hogarth press, 1930. *International Journal of Psycho-Analysis*, 12:98–101, 1931.

[11] Panji Novia Pahludi, Achmad Hidayatno, and R Rizal Isnanto. *Klasifikasi Citra Berdasarkan Tekstur Menggunakan Jaringan Saraf Tiruan Perambatan Balik*. PhD thesis, Jurusan Teknik Elektro Fakultas Teknik Undip, 2011.

[12] Zhiguang Wang, Weizhong Yan, and Tim Oates. Time series classification from scratch with deep neural networks: A strong baseline. In *2017 International joint conference on neural networks (IJCNN)*, pages 1578–1585. IEEE, 2017.

[13] Priyo Wibowo. Pengaruh perbedaan jumlah hidden layer dalam jaringan syaraf tiruan terhadap prediksi kebutuhan captopril dan paracetamol pada rumah sakit. *Media Aplikom*, 11(2), 2019.

[14] Yiting Xie and David Richmond. Pre-training on grayscale imagenet improves medical image classification. In *Proceedings of the European Conference on Computer Vision (ECCV)*, pages 0–0, 2018.