

Convolutional Neural Network for Cataract Maturity Classification Based on LeNet

Radimas Putra Muhammad Davi Labib¹, Sirojul Hadi², Parama Diptya Widayaka³, Irmalia Suryani
Paradisa⁴

^{1,4}Electrical Engineering, National Institute of Technology, Indonesia

²Information Technology, Universitas Bumigora, Indonesia

³Electrical Engineering, Surabaya State University, Indonesia

Article Info

Article history:

Received : 12-21-2021

Revised : 01-04-2022

Accepted : 04-11-2022

Keywords:

CNN;
Classification;
Cataract maturity;
LeNet.

ABSTRACT

The eyes are one of the vital organs owned by humans. One of the common eye diseases is cataracts. This disease is characterized by clouding of the lens of the eye and can interfere with vision. Worst case, sufferers can experience blindness. Cataract maturity can be divided into four categories, namely incipient, immature, mature, and hypermature. Cataracts can be removed through surgery when the cataract is in the mature or hypermature phase. Cataract examination is usually done using a slit lamp. The lack of hospitals that have this equipment can cause delays in the healing process for cataract sufferers. This study created an image processing algorithm for the maturity classification process of cataracts using the Convolutional Neural Network method with LeNet network architecture. The algorithm that has been built is capable of classifying the maturity of cataracts with an accuracy rate of 93.33%.



Accredited by Kemenristekdikti, Decree No: 200/M/KPT/2020

DOI: <https://doi.org/10.30812/varian.v5i2.1629>

Corresponding Author:

Radimas Putra Muhammad Davi Labib,
Electrical Engineering, Institut Teknologi Nasional
Email: radimas@lecturer.itn.ac.id

This is an open access article under the [CC BY-SA](https://creativecommons.org/licenses/by-sa/4.0/) license.



A. INTRODUCTION

The eyes are vital organs owned by humans that work to see the surrounding environment. The organ consists of a complex optical system that collects light from the surrounding environment (Atchison, 2018). Disorders of these organs can occur in anyone, especially in people with old age. One of the common eye disorders is cataracts. A cataract is an eye disease characterized by cloudiness in the eye's lens which can interfere with the process of entering light into the eye. This can result in blurred vision or even blindness. The disease can be caused by various factors such as age, diabetes, hypertension, and smoking habits (Harun et al., 2020). There are six types of cataracts, namely senile, congenital, traumatic, complicated, toxic, and secondary cataract. Cataract maturity is divided into four, namely incipient, immature, mature, and hyper mature. This disease can be cured by means of surgery, but the process is only when the cataract is in the mature and hyper mature phase, so it is necessary to classify the cataract maturity before surgery (Astari, 2018). Cataract examination is usually done using a slit lamp. This equipment has a very high price, so that not all health infrastructure has it. This results in delays in the healing process in cataract patients and also causes the number of cataract sufferers to increase.

The development of artificial intelligence system methods is very rapid at this time. One of the branches of artificial intelligence methods that have been widely developed is the Convolutional Neural Network (CNN). At the beginning of its appearance, this method was used as a handwriting recognition system based on images with a neural network architecture called LeNet. In its

development, the architecture can be used for the general object recognition process. In 2017, there was research on the use of LeNet architecture to perform the recognition process in handwritten Arabic numerals based on digital images (Sawy et al., 2017). There was a study in 2018 on an electronic nose gas identification system using the CNN method with LeNet as the architectural model (Wei et al., 2019). The research produced output in the form of a gas identification system with an accuracy rate of 98.67%. This research proves that LeNet is not only used for handwriting recognition but can also be used as a general classification algorithm. Research in 2019 on a sleep apnea detection system based on electrocardiogram signals using the CNN method with the LeNet architectural model (Wang et al., 2019). The study produced output in the form of a sleep apnea detection system with an accuracy rate of 97.1%.

There was a similar study in 2016, research conducted by Hariyanto and his team (Hariyanto et al., 2016) regarding the process of classifying cataracts based on pathological abnormalities using the Learning Vector Quantization (LVQ) algorithm. In this study, the classification process was carried out using one of the techniques in an artificial neural network system with a dataset in medical records from patients with cataracts. This study resulted in an accuracy of 99% cataract determination. This research focuses on the implementation of LVQ for cataract classification based on medical record datasets. In 2017, a study was conducted by Purba and his team (Purba et al., 2017) regarding the cataract diagnosis system using the concept of the retrograde method. The study used an expert system algorithm to diagnose based on the symptoms felt by the sufferer. This research produces an output in the form of a final diagnosis system regarding the severity of cataracts suffered in the form of a percentage. However, in this study, the accuracy of cataract eye detection was not explained. In 2019, Risma and her team (Risma et al., 2019) researched performance analysis of cataract detection systems based on digital images using Discrete Cosine Transform as a feature extraction technique and artificial neural networks as a classification technique. The research produced output in the form of a simulation that can detect and classify cataracts with an accuracy rate of 87.66%. In addition, a study conducted by Gifran and his team (Gifran et al., 2019) regarding the cataract classification process based on digital images using the Discrete Wavelet Transform method as a feature extraction technique and Support Vector Machine as a classification technique. The research produced output in a cataract classification system algorithm with an accuracy rate of 80%.

The purpose of this study is to be able to classify cataract maturity cheaply and have high accuracy. Based on previous research, the classification of cataract maturity was carried out using medical record data, the classification would require a longer time because it had to enter medical record data into the system to produce a cataract maturity classification. In this paper, the classification of cataract maturity is based on image data. The system was built using the CNN method with LeNet architecture, which can produce a high-accuracy classification process without having to use other methods to obtain certain features from digital images as in previous studies.

B. LITERATURE REVIEW

1. Convolutional Neural Network (CNN)

Convolutional Neural Network (CNN) is an artificial neural network that performs classification with high accuracy by inputting raw images (Suniantara et al., 2020). Before CNN, various feature extraction methods were used to describe the object in the picture. The resulting features were only suitable for describing specific objects, so they cannot be used to describe various objects universally. CNN provided a scalable approach by utilizing the principles of linear algebra to identify features in objects (Albawi et al., 2017). CNN can be used to describe various objects universally, but it requires high computational resources. There are three main types of layers, namely the convolutional layer, the pooling layer, and the fully connected layer (Bau et al., 2020).

The convolutional layer is the core of a CNN where most of the computation is done (Liu and Guo, 2019) (Suniantara et al., 2020). Generally, this layer is followed by an additional convolutional layer or it can also be followed by a pooling layer. The convolutional layer requires several components such as input data, kernel, and feature map. For example, the input data is an image with a Red Green Blue (RGB) color space. This means that the input is a three-dimensional matrix consisting of height, width, and depth. Height and width represent pixel data, while depth is a color channel. In addition to the input, there is also a feature detector (also known as the kernel) which will move through the receptive plane of the image to check for the presence of the required features. This process is known as convolution. A feature detector or kernel is a two-dimensional array that represents a part of the image. The size of the kernel can vary but is usually a 3×3 matrix. The kernel is then applied to the image area and the dot product calculation process is carried out between the input pixels and the kernel. The result of the dot product calculation is then entered into the output array. Then, the kernel is shifted and the process is repeated until it covers the entire image area. The final output result of a series of dot product calculations between the input and the kernel is known as a feature map, activation map, or convolved feature. Each output value in the feature map does not have to be related to the pixel

value of the input or image. The output value only needs to correspond to the receptive field to which the kernel is applied. In general, the convolution process can be seen in Figure 1.

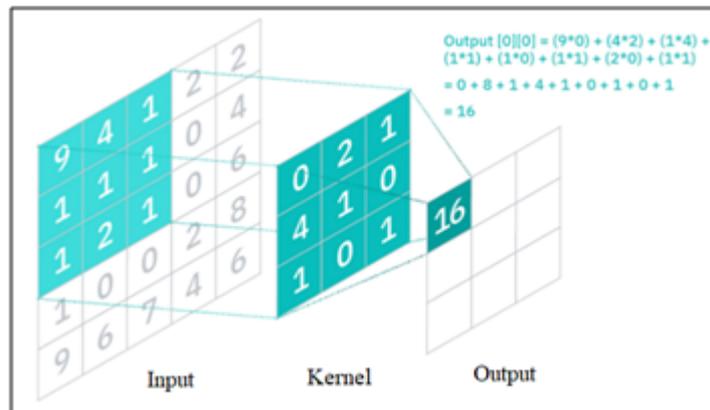


Figure 1. Convolution Process

After convolution operation, an activation function in the form of a Rectified Linear Unit (ReLU) transformation is applied to the feature map. It aims to introduce non-linearity to the model. Rectifier Linear Unit or commonly abbreviated as ReLU is a linear activation function. The ReLU activation function began to emerge in the context of visual feature extraction in neural networks in the late 1960s. ReLU has become the default activation function in many types of neural networks because it produces a model that is easy to train and has a strong biological motivation and mathematical justification and so often gives excellent performance (Agarap, 2018). ReLU activation function is mathematically described by Equation (1). The ReLU activation function is used to pass the input directly if it is positive. If the input is negative, this activation function will produce an output of zero.

$$y = \max(0, x) \tag{1}$$

The second layer in the Convolutional Neural Network is the pooling layer. This layer is also known as the down sampling process which has the main goal of reducing the dimensions of the input so that its parameters are reduced (Gholamalizhad and Khosravi, 2020). The working principle of this layer is similar to the Convolutional Layer, which is to apply a filter to the input, but the kernel in this layer does not have a weight value. Instead, the kernel at this layer applies an aggregation function to the receptive field and the result is entered into an output array. There are two types of pooling, namely max pooling and average pooling. However, in this study only the max-pooling type was used, the way it works is shown in Figure 2.

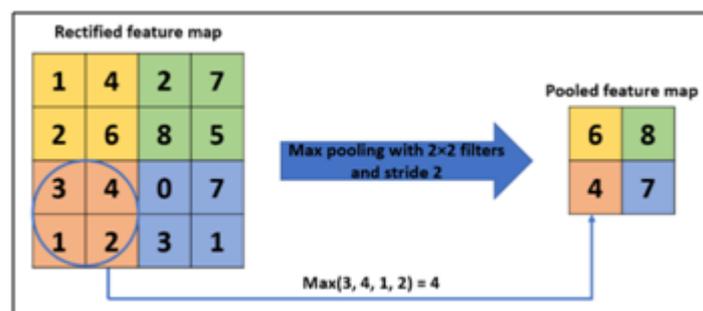


Figure 2. Max-Pooling

The third layer in the Convolutional Neural Network is the fully connected layer. After getting some features map from the convolutional layer and the pooling layer, the matrices are smoothed into a vector and put into a fully connected layer (Zhou et al., 2017). An example of a fully connected layer can be seen in Figure 3.

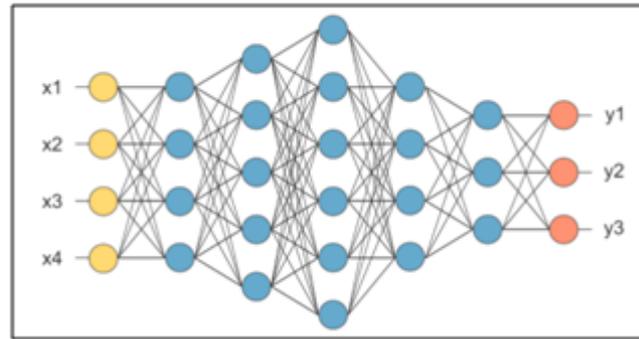


Figure 3. Fully Connected Layers

In Figure 3, the feature map matrix will be converted into a vector (x_1, x_2, x_3, x_4) . With a fully connected layer, the system can combine these features to produce a model. The output classification process is carried out by applying the softmax activation function. The softmax activation function or also known as the normalized exponential activation function is a generalization of the logistic function for many dimensions (Kanai et al., 2018). Softmax is used in multinomial logistic regression and is often used as the last activation function in neural networks to normalize the network output to a probability distribution over the predicted output class. Mathematically, the softmax activation function is described in Equation (2).

$$y_i = \frac{\exp(x_i)}{\sum_j \exp(x_j)} \quad (2)$$

2. LeNet Network Structure

LeNet is one of the earliest convolutional neural network structures (Rongshi and Yongming, 2019). In 1989, Yann LeCun and his team implemented a combined convolutional neural network that was trained to use backward propagation to read handwriting. The neural network structure succeeded in identifying the zip code numbers written by hand. This structure is a prototype of what is hereinafter referred to as LeNet. The classification system in this study uses LeNet as a CNN architecture with a network structure as shown in Figure 4.

In this study, the input used is an image of an eye with RGB color space and dimensions of 28×28 pixels. Next, a convolution process will be carried out with a 5×5 kernel to get a line of feature maps measuring 24×24 with a depth of 20 feature maps. After getting the first line of feature maps, then the max-pooling process is carried out with a 2×2 kernel to produce a line of feature maps measuring 12×12 with a depth of 20 feature maps. Then a second convolution process is carried out with a 5×5 kernel to get a line of feature maps measuring 8×8 with a depth of 50 feature maps. Then the max-pooling process is carried out with a 2×2 kernel to produce a row of 4×4 feature maps with a depth of 50 feature maps. The next process is flattened which serves to convert the last row of feature maps into vectors. The vector will be inserted into the fully connected layer to get the prediction results.

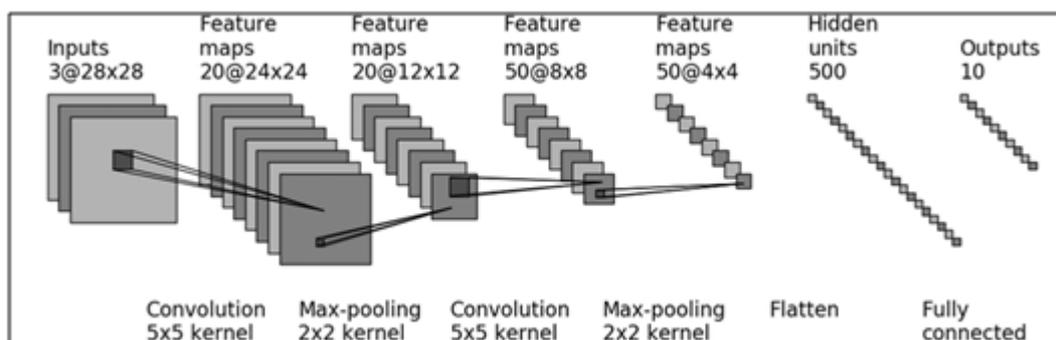


Figure 4. LeNet Network

C. RESEARCH METHOD

1. Training Algorithm Design

Artificial neural networks are designed to imitate the work of the human brain through a combination of input data, weights, and biases. These elements work together to identify, classify, and accurately describe objects. The stages of machine learning can be seen in Figure 5.

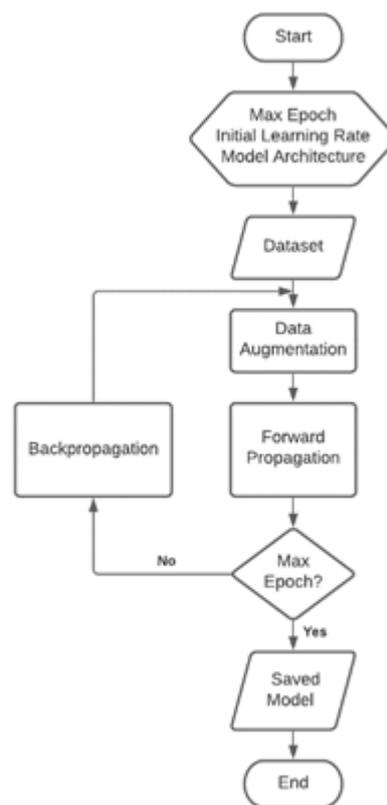


Figure 5. Training Algorithm Design

A deep neural network consists of several layers of interconnected nodes. The artificial neural network model requires a learning process based on datasets that have been prepared to get prediction results with a high level of accuracy. There are two main processes in a machine learning system on an artificial neural network, namely forward propagation and backward propagation. Forward propagation is a computational process through each layer of the network to get the prediction results from the classification system and get the loss-accuracy value. Backpropagation is an evaluation process to improve the elements in each layer of the neural network to improve the accuracy of the prediction process. Both processes are carried out internally by the Tensorflow module.

2. Prediction Algorithm Design

After doing the training process, an artificial neural network model will be obtained which is ready to be used for the prediction process. Before carrying out the prediction process, preprocessing is carried out on the input data to adjust the input standards for the neural network. After that, the prediction process will be carried out through the forward propagation process using an artificial neural network model generated by the learning process. Broadly speaking, the stages of the prediction system can be seen in Figure 6.

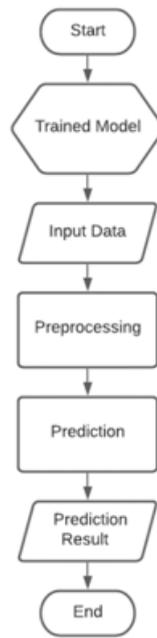


Figure 6. Algorithm of Prediction Process Design

3. Dataset

The dataset is used in the training process on an artificial neural network. In this study, a dataset of 37 images was used with details of 10 images of eyes with immature cataracts, 17 images of eyes with mature cataracts, and 10 images of normal eyes . The dataset used is obtained from Google Images. The dataset is shown in Figure 7.



Figure 7. Dataset

D. RESULTS AND DISCUSSION

1. Machine Learning Result

The learning process was carried out three times with a different number of epochs in each experiment. The first experiment used 25 epochs with a training graph as shown in Figure 8 and a validation graph as shown in Figure 9. In the training process, there is a change in the value of accuracy and loss at each epoch. Although the change does not always increase, the final results show that the accuracy value is greater than in the previous epochs. This training process gives the final results in the form of an accuracy value of 81.82% and a loss value of 0.3957.



Figure 8. Training process on the first experiment

The validation process is carried out after getting the results of each epoch of the training process. The validation process of the model that has been trained in the last epoch produces an accuracy value of 73.33% and a loss value of 0.4085. Based on these results, the model generated can be used to classify cataract maturity because it has an accuracy rate above 50%.

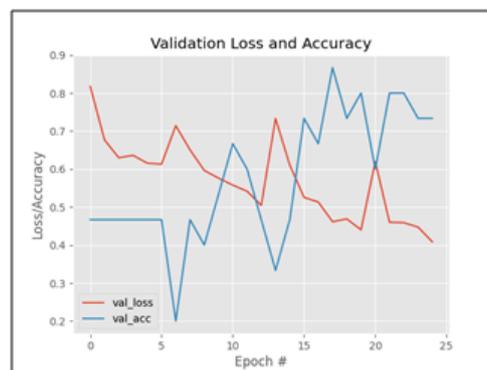


Figure 9. Validation process on the first experiment

The second experiment used 50 epochs with a training graph as shown in Figure 10 and a validation graph as shown in Figure 11. During the training process, the accuracy value in the last epoch was not the best. However, it shows that the accuracy value is still above 50%. This training process gives the final results in the form of an accuracy value of 86.36% and a loss value of 0.3092.



Figure 10. Training process on the second experiment

In the validation process, the final result is an accuracy value of 73.33% and a loss value of 0.4025. Based on these data, it can be shown that the loss value decreased slightly compared to the previous experiment. In addition, there is no difference in the accuracy obtained when compared with the previous experiment. However, the model generated in this experiment is also able to carry out the process of classifying cataract maturity.

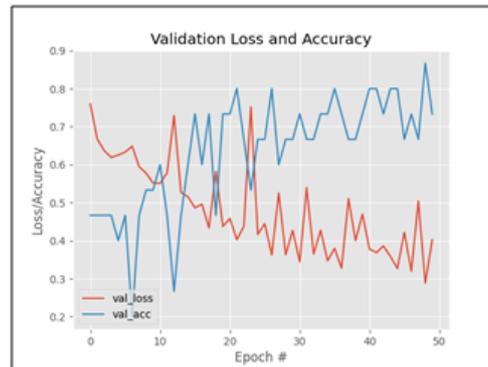


Figure 11. Validation process on the second experiment

The third experiment used 100 epochs with a training graph as shown in Figure 12 and a validation graph as shown in Figure 13. The training process gives the final results in the form of an accuracy value of 95.45% and a loss value of 0.1424. It shows that the training results obtained are much better than the first and second experiments.

In the validation process, the resulting accuracy value is 93.33% and the resulting loss value is 0.2705. Based on these data, it can be shown that the validation results obtained are much better than the first and second experiments. In addition, the resulting accuracy rate is above 90% for the results of the training and validation, so that the developed model is capable of classifying the maturity of cataract s with very high accuracy. The value of accuracy and loss in all machine learning that has been carried out can be shown in Table 1.



Figure 12. Training process on the third experiment

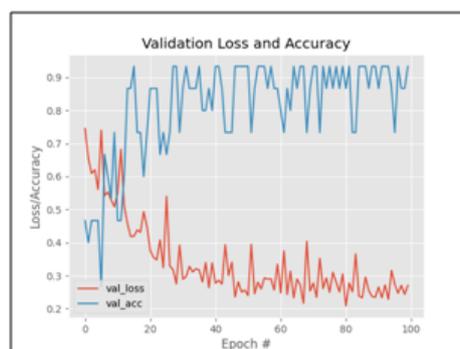


Figure 13. Validation process on the third experiment

Table 1. Results of training and validation on cataract maturity

Epoch	Training		Validation	
	Accuracy	Loss	Accuracy	Loss
25	81.82%	0.3957	73.33%	0.4085
50	86.36%	0.3092	73.33%	0.4025
100	95.45%	0.1424	93.33%	0.2705

2. Prediction System Result

The prediction system uses an artificial neural network model generated by the learning process in the third experiment. The experiment has a higher accuracy value compared to the learning process in the first and second experiments. The results obtained in the prediction system can be shown in Figure 14.

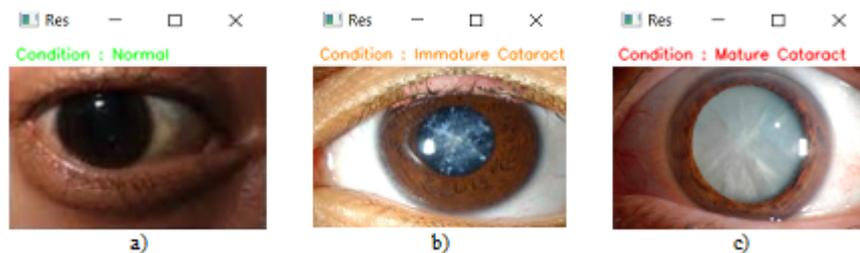


Figure 14. Prediction result: a) Normal eye; b) Immature Cataract; c) Maturation Cataract

In detail, the prediction process produces an output value for each label. From all these output values, the label with the highest value was chosen as a predictor of cataract maturity. The process can be shown in Table 2. In the table, the output value written in bold is the highest value. So the label with the highest value will be used as a prediction result.

Table 2. Cataract maturity prediction system result

Input	Output Value			Prediction
	Normal	Immature	Maturation	
	0.6812226	0.3176026	0.0011747	Normal
	0.0039405	0.9110546	0.0850049	Immature
	0.0293672	0.2084033	0.7622294	Maturation

E. CONCLUSION AND SUGGESTION

Based on the experiments that have been carried out, it can be concluded that the image processing algorithm based on the Convolutional Neural Network designed can carry out the maturity classification process of cataracts. Unlike in previous studies, the algorithm that has been designed does not require additional methods to extract certain features. Based on the results of experiments that have been carried out, this algorithm can perform the classification process with an accuracy rate of 93.33%. This proves that the use of CNN with LeNet architecture for the classification process is more efficient and provides more accurate classification results

ACKNOWLEDGEMENT

We would like to thank Institut Teknologi Nasional Malang for funding our research.

REFERENCES

- Agarap, A. F. (2018). Deep Learning using Rectified Linear Units (RELU). <http://arxiv.org/abs/1803.08375>.
- Albawi, S., Mohammed, T. A. M., and Alzawi, S. (2017). Layers of A Convolutional Neural Network. *IEEE*, page 16.

- Astari, P. (2018). Katarak: Klasifikasi, Tatalaksana, dan Komplikasi Operasi. *Cermin Dunia Kedokteran*, 45(10):748–753.
- Atchison, D. A. (2018). Optics of The Human Eye. *Encyclopedia of Modern Optics*, 1-5:43–63.
- Bau, D., Zhu, J. Y., Strobelt, H., Lapedriza, A., Zhou, B., and Torralba, A. (2020). Understanding The Role of Individual Units in A Deep Neural Network. *Proceedings of the National Academy of Sciences of the United States of America*, 117(48):30071–30078.
- Gholamalinezhad, H. and Khosravi, H. (2020). Pooling Methods in Deep Neural Networks, A Review.
- Gifran, N. A., Magdalena, R., and Fuadah, R. Y. N. (2019). Klasifikasi Katarak Menggunakan Metode Discrete Wavelet Transform dan Support Vector Machine Classification of Cataract Using Discrete Wavelet Transform and Support Vector Machine. *e-Proceeding of Engineering*, 6(2):4170–4177.
- Hariyanto, R., Basuki, A., and Hasanah, R. N. (2016). Klasifikasi Penyakit Mata Katarak Berdasarkan Kelainan Patologis dengan Menggunakan Algoritma Learning Vector Quantization. *Jurnal Ilmiah NERO*, 2(3):177–182.
- Harun, H. M., Abdullah, Z., and Salmah, U. (2020). Pengaruh Diabetes, Hipertensi, Merokok dengan Kejadian Katarak di Balai Kesehatan Mata Makassar. *Jurnal Kesehatan Vokasional*, 5(1):45.
- Kanai, S., Yamanaka, Y., Fujiwara, Y., and Adachi, S. (2018). Sigsof Tmax: Reanalysis of The Softmax Bottleneck. *Advances in Neural Information Processing Systems*, 2018-Decem(NeurIPS):286–296.
- Liu, G. and Guo, J. (2019). Bidirectional LSTM with Attention Mechanism and Convolutional Layer for Text Classification. *Neurocomputing*, 337:325–338.
- Purba, W., Aisyah, S., and Tamba, S. P. (2017). Perancangan Sistem Pakar Diagnosa Penyakit Mata Katarak Menggunakan Konsep Metode Runut Mundur. *JUSIKOM PRIMA (Jurnal Sistem Informasi Ilmu Komputer Prima)*, 1(1).
- Risma, H. A., Patmasari, R., and Magdalena, R. (2019). Analisis Performansi Sistem Pendeteksi Katarak Menggunakan DCT (Discrete Cosine Transform) dan Jaringan Saraf Tiruan Backpropagation (JST Backpropagation). *e-Proceeding of Engineering*, 6(1):364–371.
- Rongshi, D. and Yongming, T. (2019). Accelerator Implementation of Lenet-5 Convolution Neural Network Based on FPGA with HLS. *3rd International Conference on Circuits, System and Simulation, ICCSS 2019*, pages 64–67.
- Sawy, A. E., El-Bakry, H., and Loey, M. (2017). CNN for Handwritten Arabic Digits Recognition Based on LeNet-5. *Advances in Intelligent Systems and Computing*, 533:565–575.
- Suniantara, I. K. P., Suwardika, G., and Soraya, S. (2020). Peningkatan Akurasi Klasifikasi Ketidaktepatan Waktu Kelulusan Mahasiswa Menggunakan Metode Boosting Neural Network. *Jurnal Varian*, 3(2):95–102.
- Wang, T., Lu, C., Shen, G., and Hong, F. (2019). Sleep Apnea Detection From A Single-Lead ECG Signal with Automatic Feature-Extraction Through A Modified LeNet-5 Convolutional Neural Network. *PeerJ*, 2019(9):1–17.
- Wei, G., Li, G., Zhao, J., and He, A. (2019). Development of A LeNet-5 Gas Identification CNN Structure for Electronic Noses. *Sensors (Switzerland)*, 19(1).
- Zhou, Y., Song, S., and Cheung, N. M. (2017). On Classification of Distorted Images with Deep Convolutional Neural Networks. *ICASSP, IEEE International Conference on Acoustics, Speech and Signal Processing - Proceedings*, (January):1213–1217.