JURNAL INFORMATIKA, Vol.11 No.2 October 2024, Page 65-72
ISSN (Print) 2355-6579 | ISSN (Online) 2528-2247
DOI: https://doi.org/10.31294/inf.v11i2.20328

Implementation of Convolutional Neural Network in Recognize Lontara Text

Wahyu Arthanugraha[1,2]*, Zahir Zainuddin[3], Abdul Latief Arda[4]

[1,4] Magister Sistem Komputer, Universitas Handayani Makassar
Jl. Adhyaksa Baru, Makassar, Sulawesi Selatan, Indonesia

[2] Teknik Informatika, Universitas Muhammadiyah Parepare
Jl. Jl. Jend. Ahmad Yani Km. 6, Parepare, Sulawesi Selatan, Indonesia

[3] Magister Teknik Informatika, Universitas Hasanuddin
Jl. Perintis Kemerdekaan Km.10 Tamalanrea, Makassar, Sulawesi Selatan, Indonesia

Correspondence e-mail: wahyusaja212@gmail.com

Submission:	Revision:	Acceptance:	Available Online:
15-05-2024	28-06-2024	11-07-2024	20-08-2024

Abstract

The introduction of the Lontara text that is applied is still conventional in that it comes from books or teaching materials and there is no application for a translator of the Lontara text, making it difficult for students, especially those from outside the Bugis tribe. This research aims to measure the ability and apply the CNN method in the context of introducing the Lontara text. The input that use in this research is words taken from the original text or writing manuscripts of the ancient literature of the Bugis tribe "Sureq Maqkelluqna Nabittaq", a book that was used as a tool to spread Islam in 1611. The model used is the VGG-16 architecture. This research uses the R&D (Research and Development) method which is used to create products and test their level of effectiveness. As for the research results obtained, the system accuracy obtained in classifying the Lontara text using the VGG16 architecture was 97,65%. In addition, the system is also able to display the translation of the Lontara text according to the database input.

Keywords: Lontara Text, CNN, VGG16

1. Introduction

The Bugis heritage represents a significant aspect of the identity of Muhammadiyah University of Parepare (UMPAR). In accordance with this, the Muhammadiyah University of Parepare has developed a Rencana Program dan Kegiatan Pembelajaran Semester (RPKPS) with a course on local wisdom. The Local Wisdom course is a subject that facilitates the study of the cultural richness of the archipelago. This course serves as a conduit for the investigation of noble local thoughts and as a forum for the critical examination of the prospects for the continued existence of these local thoughts in the context of cultural development and global influence. It is hoped that all students will gain an understanding of the cultural heritage of the archipelago, with a particular focus on the Bugis culture and its Lontara text. The method of introducing the Lontara text at UMPAR remains conventional, relying on books or teaching materials. There is currently no Lontara text translator application, which presents a challenge for students, particularly those from outside the Bugis tribe.

The Lontara text, which represents the Bugis language, is manifested in the form of distinctive characters. It is important to note that the use of the Lontara text is primarily associated with two major ethnic groups in South Sulawesi: the Bugis and the Makassar tribes. In addition to serving as a writing system, the Lontara text also symbolizes regional identity and acts as a means to pass on valuable cultural values or local knowledge. The Lontara text is a cultural asset with the potential to be developed as a regional cultural tourism attraction, while also contributing to the development of national culture.

Based on related research that investigates pattern recognition in writing, specifically, the pattern recognition in the Sundanese text using the "Modified Direction Feature method" to extract images in 2017. The results of the tests conducted found that the best accuracy was 78.67% (Riansyah et al., 2017).

Related research discussing the recognition of Lontara text was conducted by Disa Ainun Safitri using the “Optical Character Recognition” method with the Manhattan algorithm, achieving the highest accuracy of 60.87% for pattern recognition with new data (Safitri, 2018).

A study by Yamashita et al. (2018) shows that CNN is capable of achieving high accuracy in handwriting recognition. "Handwritten character recognition using convolutional neural networks has shown remarkable performance improvements, achieving high accuracy rates on various datasets" (Yamashita et al., 2018).

Zhou et al. (2019) investigated the use of CNNs for Chinese character recognition and found that this method is highly effective. "CNNs have demonstrated superior performance in recognizing Chinese characters, significantly outperforming traditional methods" (Zhou et al., 2019).

According to Al-Hajj et al. (2020), CNNs are also effective in recognizing Arabic characters. "The use of CNNs for recognizing handwritten Arabic characters has resulted in significant accuracy improvements, handling diverse handwriting styles efficiently" (Al-Hajj et al., 2020).

Simonyan and Zisserman (2014) introduced the VGG16 architecture, which emphasizes the use of deeper convolutional layers for better pattern recognition. "The VGG16 architecture, with its deep convolutional layers, has proven to be highly effective in various image recognition tasks, achieving top performance on benchmarks such as ImageNet" (Simonyan & Zisserman, 2014)

Research by Xiao et al. (2017) showed that VGG16 can be used for handwriting recognition with excellent results. "VGG16, when fine-tuned on handwritten text datasets, demonstrates remarkable accuracy and robustness, making it suitable for practical handwriting recognition applications" (Xiao, Y., et al., 2017).

CNN, or more commonly referred to as deep learning, is an approach that employs convolutional features to extract characteristics from images. In light of previous research that has yielded satisfactory accuracy rates, CNN was selected as the method for identifying Lontara text from an image.

The objective of this study is to assess the efficacy of the convolutional neural network (CNN) method for the recognition of Lontara text. The input consists of words taken from manuscripts or ancient literary works of the Bugis tribe, specifically from Sureq Maqkelluqna Nabittaq, a book used as a tool for spreading Islam in 1611, as a substitute for the book Berazanji at that time (Arifin et al., 2018) The image of the Lontara text word that is input will be translated into Latin letters and then into Indonesian.

2. Research Methods

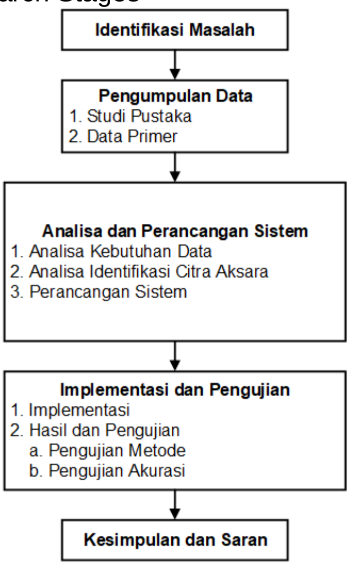
2.1. Types of Research

This study uses the Research and Development (R&D) method aimed at developing and validating new products or systems. The R&D process includes several main stages:

1. Data Collection and Needs Analysis: Identifying existing problems or needs through surveys, interviews, or literature studies.
2. Initial Product Development: Designing prototypes or initial models based on the collected data.
3. Testing and Evaluation: Testing the initial product on a small scale to gather feedback and assess effectiveness.
4. Revision and Improvement: Improving and refining the product based on test results and evaluations.
5. Field Testing: Implementing the product on a larger and more realistic scale to observe how it performs in real-world conditions.
6. Final Implementation and Dissemination: Distributing the refined and validated product for widespread use.

The research and development (R&D) process typically follows an iterative cycle, whereby each stage can be repeated several times until the developed product meets the desired quality standards. The R&D method guarantees that the resulting product is not only innovative but also pertinent and efficacious in addressing the identified needs.

2.2. Research Stages



Source: Writer (2023)

Figure 1. Research Stage

Here is the translation of the research stages as described in Figure 1:

1. Problem Identification

The initial step in the process of gathering information about the application of image processing and artificial neural networks is the problem identification stage. This stage is undertaken by researchers who have previously conducted similar studies. This stage also entails the search for data pertaining to the recognition of Lontara text through the use of Convolutional Neural Network (CNN) methodologies.

2. Data Collection

This research utilizes two methods for data collection:

Literature Review: For further information, please refer to studies on image processing using Convolutional Neural Network (CNN) methods and other related research.

Primary Data: The data set comprises 1,071 images of Lontara text words from the manuscript Sureq Maqkelluqna Nabittaq. The Lontara text data was captured by photographing the images from above using a mobile phone camera.

3. System Analysis and Design

The training and testing process utilising the CNN method involves the division of the collected dataset into training and testing data sets. In order to achieve results, it is necessary to conduct training and testing. Once the training phase has been completed, the subsequent step is to calculate the loss and accuracy values derived from the testing results. The accuracy of the testing data is evaluated in relation to the results obtained from the training data. The determination of epochs in the testing phase is conducted in a single session in order to obtain errors and achieve optimal results. Following the labeling of the data and its separation into training and testing sets, the dataset is trained using the Convolutional Neural Network (CNN) algorithm to form a network architecture, with the objective of obtaining accuracy results.

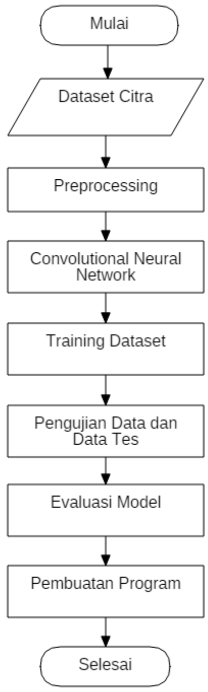
4. Implementation and Testing

The implementation stage involves The execution of the analyzed and designed data within the system. Whitebox testing is employed in the application to ascertain whether the convolutional neural network (CNN) preprocessing algorithm source code aligns with the intended design and output. The testing process entails the utilisation of a confusion matrix, in conjunction with uniform learning rate testing, with the objective of determining the accuracy level of Lontara text classification.

5. Conclusion and Recommendations

In this stage, the results of the previous research are analysed in depth in order to draw comprehensive conclusions. The objective of this stage is to ascertain whether the system constructed using Convolutional Neural Network (CNN) methodologies functions as intended. Furthermore, recommendations are provided for the findings of this research to be further developed by future researchers.

2.3. System Design



Source: Writer (2023)

Figure 2. Flowchart System

The initial stage of the process involves the collection of image datasets from the Sureq Maqkelluqna Nabittaq manuscript. The dataset consists of images of Lontara text words that will be used for the purposes of training and testing images. The subsequent stage is pre-processing, which encompasses normalization, rotation, zoom, and flip. Subsequently, the next step is to construct the convolutional neural network (CNN) architecture and training dataset, which constitutes the training process. The subsequent phase entails the assessment of the CNN model's precision through a testing phase. The model is evaluated using a number of sample images from the dataset. Following these steps, the result is the translation of the Lontara text. This information will be displayed on a mobile device.

3. Results and Discussion

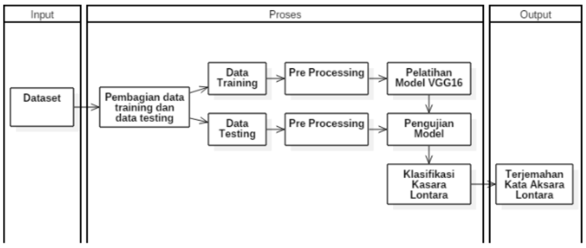
3.1. Research Dataset

The dataset used consists of images obtained from the Sureq Maqkelluqna Nabittaq manuscript, comprising 51 words with a total of 1071 image data. Out of these, 80% (816 images) are used for training data, and 20% (255 images) are used for validation data. The Lontara text word images obtained have different image sizes with RGB (Red, Green, and Blue) color channels from the total 1071 images. This dataset division aims to manage the training and testing process of the model on the dataset used. The following examples illustrate several datasets in the Lontara script:

Source: writer (2023)

Figure 3. File dataset

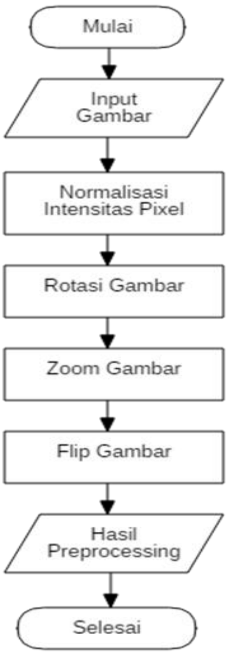
3.2. Algorithm Implementation



Source: writer (2023)

Figure 4. Model Testing

Figure 4 depicts the model testing process, which is comprised of three stages: input, process, and output. The initial stage of the process entails the preparation of the dataset, which will be utilized in the study. The objects utilized are words extracted from the manuscript Sureq Maqkelluqna Nabittaq. Once the images have been acquired, a series of preprocessing steps are employed with the objective of enhancing the quality of the images, thereby facilitating and improving the system's ability to identify objects. Figure 5 shows the preprocessing process.

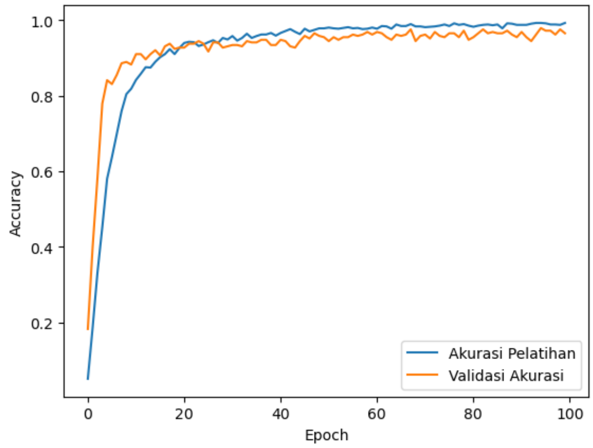


Source: Writer (2023)

Figure 5. *Flowchart* of *preprocessing Stage*

In the preprocessing stage, normalization is performed by resizing the images to a size of 224x224 pixels. The next step involves scaling the images to change the pixel range from [0, 255] to [0, 1]. The data augmentation methods applied in this study include rotating the image by 20 degrees. The next augmentation process involves zooming in the image by 20%. Additionally, the image will be flipped horizontally to increase data variation, allowing the program to recognize different data. The output of the preprocessing process is an image that has pass the resizing and augmentation.

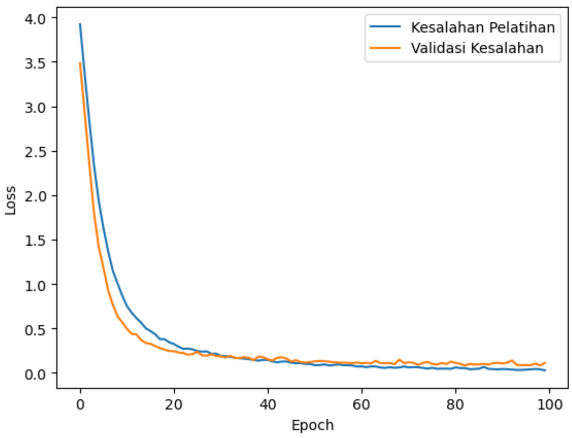
After completing the preprocessing stage, the next step is to train the VGG-16 model, where the weights of this VGG-16 model have been previously trained using the dataset ImageNet. In this training stage, the process involves convolution, ReLU activation function, Max Pooling, and Softmax. Next, the model is tested to produce classification results. Testing is conducted on the data that has been divided into training and testing datasets, using the previously mentioned dataset consisting of 1071 lontara character images. The accuracy of the model in classifying lontara characters is evaluated based on the probability values obtained for each class. The model is considered accurate when the probability value approaches 1, which can range from 0 to 1. The accuracy results are shown in Figure 6 below:



Source: Writer (2023)

Figure 6. Accuracy and Validation Graph

The graphs in Figures 5 and 6 illustrate the accuracy and error results on the training and validation data during the model creation process. The term "accuracy" refers to the comparison of correct predictions, both positive and negative, against the overall data for each class. Concurrently, the error function enables the model to assess the accuracy of the predictions. The primary objective of the model development process was to achieve a low error value.



Source: Writer (2023)
Figure 7. Training Error and Validation Error Graph

In this model development process, error evaluation was measured using cross-entropy loss, a common method used in classification. The model creation process was built with several parameters including epoch = 100, batch size = 32, optimizer = Adam and the ReLU activation function.

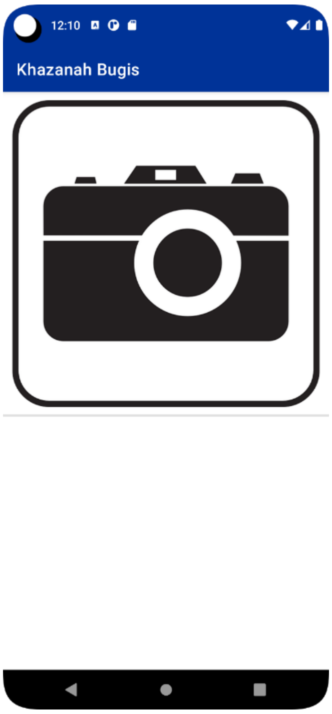
After classifying the Lontara text images, the process continues with translation.

Source: Writer (2023)
Figure 8. Translation Process Program Flowchart

3.3. System Implementation

Figure 9 shows the Application Splash Screen. When the application is first launched, a splash screen appears for 4 seconds before displaying the main menu of the Lontara text translation application on Android devices.

Source: Writer (2023)
Figure 9. *Splash Screen*

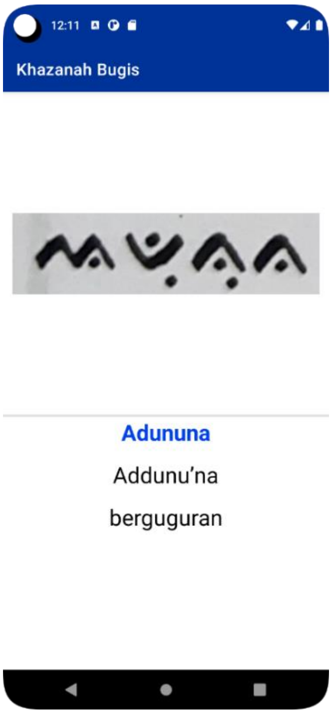


Source: Writer (2023)
Figure 10. Main Menu of the Application

Figure 10 shows the main menu. In this main menu, there is a camera icon that has function to capture an image using the camera or to select an image from the gallery, which will be detected.

Source: Writer (2023)
Figure 11. Image Selection Screen

Figure 11 shows the screen after an image is selected from the camera or gallery. The image can be cropped according to the Lontara text word that will be detected.



Source: Writer (2023)
Figure 12. Detection Result Screen

Figure 12 shows the detection result screen where the input is a Lontara text word, and the output includes Latin letters, pronunciation, and the translation into Indonesian from the inputted Lontara text.

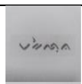
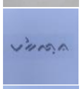
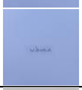
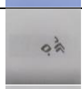
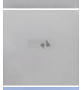
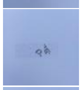
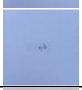
3.4. System Testing

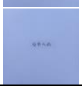
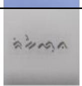
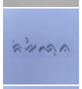
In this stage, testing will be conducted on the data and system to evaluate the system's capability in detecting Lontara text. The object detection results are tested by comparing different camera resolutions which are:

a. Samsung A6 Plus camera with a 16 MP resolution.
b. Vivo Y95 camera with a 13 MP resolution.

The testing results can be seen in the following table:

Table 1 Object detection testing results.

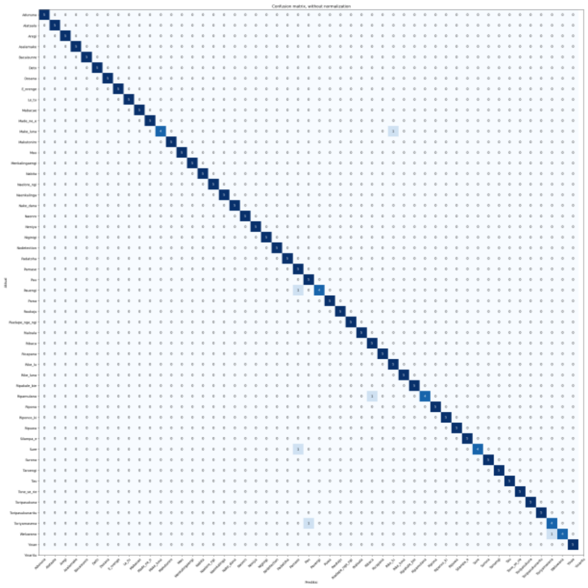
No	Figure	Device	distance (cm)	Actual	Info
1		Samsung A6 Plus	10	Make_luna	True
			20	Make_luna	True
			30	Make_luna	True
		Vivo Y95	10	Make_luna	True
			20	Make_luna	True
			30	Make_luna	True
2		Samsung A6 Plus	10	Sure	True
			20	Sure	True
			30	Sure	True
		Vivo Y95	10	Sure	True
			20	Sure	True
			30	Sure	True
3		Samsung A6 Plus	10	Puwa	True
			20	Puwa	True
			30	Puwa	True
		Vivo Y95	10	Puwa	True

No	Figure	Device	distance (cm)	Actual	Info
			20	Puwa	True
			30	Puwa	True
4		Samsung A6 Plus	10	Surona	True
			20	Surona	True
			30	Surona	True
		Vivo Y95	10	Surona	True
			20	Surona	True
			30	Surona	True
5		Samsung A6 Plus	10	Rike_luna	True
			20	Rike_luna	True
			30	Make_luna	False
		Vivo Y95	10	Rike_luna	True
			20	Rike_luna	True
			30	Rike_luna	True

Source: Writer (2023)

In Table 1, a comparison of object detection results using different cameras and distances can be observed. At distances of 10cm and 20cm, the Samsung A6 Plus camera successfully detected the Lontara text word "Rike_luna" with the translation "Utusannya" but at 30cm, it incorrectly detected the intended Lontara text word as "Rike_luna" The Vivo Y95 camera only successfully detected the Lontara text at a distance of 10cm. The resolution of the camera used is one of the factors affecting the accuracy of Lontara text detection, where the 16MP camera has better accuracy compared to the 13MP camera. The distance at which the images are captured also affects the accuracy, with the effective range for capturing images not exceeding 20cm.

As for the accuracy testing of the system using a confusion matrix, the results are as follows:



Source: Writer (2023)

Figure 13. *Confusion Matrix*

accuracy			0.98	255
macro avg	0.98	0.98	0.98	255
weighted avg	0.98	0.98	0.98	255

Accuracy: 97.65%

Source: Writer (2023)

Figure 14. *Report* Classification

Based on the evaluation of the test data, an accuracy of 97.65% was achieved.

4. Conclusion

From the research findings, Lontara text classification was successfully conducted using the Convolutional Neural Network method with the Visual Geometry Group 16 (VGG-16) architecture. The model was trained with additional parameters including epoch = 100, batch size = 32, optimizer = Adam, and ReLU activation function. Based on the evaluation of the data test, an accuracy of 97.65% was achieved. The system is capable of displaying Latin letters, pronunciation, and translations of Lontara text from input images. During the training process, several factors influenced the accuracy, including camera resolution, where a 16MP camera showed better accuracy compared to 13MP. The distance at which images were captured also affected accuracy, with an effective range not exceeding 20cm. Further research could explore various CNN algorithms to achieve even higher accuracy, such as comparing VGG-16 with VGG-19.

Reference

Al-Hajj, R., Mokbel, C., Likforman-Sulem, L., & El-Beze, M. (2020). Combining Slanted-Frame Classifiers for Improved HMM-Based

Arabic Handwritten Word Recognition. Pattern Recognition, 51, 21-32.

Alom, M. Z., Sidike, P., Hasan, M., Taha, T. M., & Asari, V. K. (2018). Handwritten Bangla Character Recognition Using the State-of-the-Art Deep Convolutional Neural Networks. Computational Intelligence and Neuroscience, 2018. https://doi.org/10.1155/2018/6747098

Arifin, S., Surahman, & Mursalim. (2018). Analisis Naskah Sureq Makelluqna Nabittaq Di Tinjau dari Aspek Filologi. *Jurnal ILmu Budaya*, *2*(2).

James, A., Manjusha, J., & Saravanan, C. (2018). Malayalam handwritten character recognition using AlexNet based architecture. Indonesian Journal of Electrical Engineering and Informatics, 6(4). https://doi.org/10.11591/ijeei.v6i4.518

K. S. Younis. 2017. Arabic Handwritten Character Recognition Based On Deep Convolutional Neural Networks. Jordanian Journal of Computers and Information Technology (JJCIT).

Khan, S. H., Khan, S. H., & Raza, M. A. (2020). Handwritten Digit Recognition Using Convolutional Neural Networks. International Journal of Computer Applications, 175(2), 25-28.

Khandokar, I., Hasan, M., Ernawan, F., Islam, S., & Kabir, M. N. (2021). Handwritten character recognition using convolutional neural network. Journal of Physics: Conference Series, 1918(4). https://doi.org/10.1088/1742-6596/1918/4/042152

Nur, R., Andyana, S., & Winarsih. (2018). Pengenalan Aksara Lontara Tulis Tangan Menggunakan Metode Convolutional Neural Networks Berbasis Android. In Winarsih Jurnal Informatika (Vol. 18, Issue 1).

Riansyah, R. R., Nurhasanah, Y. I., & Dewi, I. A. (2017). Sistem Pengenalan Aksara Sunda Menggunakan Metode Modified Direction Feature dan Learning Vector Quantization. *Jurnal Teknik Informatika Dan Sistem Informasi*, *3*(1). https://doi.org/10.28932/jutisi.v3i1.563

Sabri, M., Nurhayati, & Syahrir. (2020). Aplikasi Pembelajaran Aksara Lontara Bugis Makassar Berbasis Mobile. Prosiding Seminar Nasional Teknik Elektro Dan Informatika (SNTEI) 2020.

Safitri, D. A. (2018). Pengenalan Pola Aksara Lontara Dengan Metode Optical Character Recognition. *Skripsi*.

Simonyan, K., & Zisserman, A. (2014). *Very Deep Convolutional Networks for Large-Scale Image Recognition*. http://arxiv.org/abs/1409.1556

Wardhani, N. (2018). Aplikasi Pembelajaran Aksara Lontara Dengan Algoritma Linear Congruent Method Berbasis Android Studi Kasus: SLTP Unggulan 13 Makassar. Instek, 3(3).

Yamashita, R., Nishio, M., Do, R. K. G., & Togashi, K. (2018). Convolutional neural networks: an overview and application in radiology. In *Insights into Imaging* (Vol. 9, Issue 4, pp. 611–629). Springer Verlag. https://doi.org/10.1007/s13244-018-0639-9

Zhou, J., Wu, Q., Zhang, X., & Zhu, Y. (2019). Chinese Handwriting Recognition Using Convolutional Neural Networks. Journal of Visual Communication and Image Representation, 59, 367-375.