Jurnal Mantik, 7 (1) (2023), ISSN 2685-4236 (Online)

Published by: Institute of Computer Science (IOCS)

Jurnal Mantik

Journal homepage: www.iocscience.org/ejournal/index.php/mantik

Lightweight convolutional neural network for khat naskhi and riq'ah classification

Muhamad Taufiq Riza[1], Oddy Virgantara Putra[2], Taufiqurrahman[3]

[1,2,3]Informatic Engineering, University of Darussalam Gontor, Indonesia

ARTICLE INFO

Article history:

Received Mar 30, 2023
Revised Apr 8, 2023
Accepted Apr 13, 2023

Keywords:

Arabic Writing
Classification
Khat Naskhi
Khat Riq'ah
Lightweight CNN

ABSTRACT

Arabic writing has various types of khat that are complex and different from each other, so it requires proper classification to identify the type of khat used. This research uses the Lightweight Convolutional Neural Network (CNN) classsification method to recognize the types of khat naskhi and riq'ah on Arabic writing datasets. The evaluation results show that this classification model has an accuracy of 98.75% on training data and 100% on validation data, with a relatively fast processing time of 2s 375ms per step so that the model can be implemented well in systems that require high data processing speed and also devices that have limited resources. These results show that the classification model using the Lightweight CNN layer can be used as an effective alternative in classifying types of Arabic writing, especially in recognizing certain types of khat such as naskhi and riq'ah. Furthermore, this research can be developed using a larger and more diverse dataset, as well as evaluated and compared with other classification models to improve the performance of the model in recognizing more complex types of Arabic writing.

Corresponding Author:

Muhamad Taufiq Riza,
Informatic Engineering,
University of Darussalam Gontor,
Siman, Ponorogo, East Java, 63471, Indonesia.
Email: muhamadtaufiqriza@unida.gontor.ac.id

1. INTRODUCTION

Khat is an Arabic writing art that is very important in Middle Eastern culture and is spread all over the world. Khat naskhi and riqah are the two types of khat most commonly used in official writings, such as books, letters and other documents. Despite its importance, the classification between khat naskhi and riqah is still a challenge, especially for experts who are still new to this field, because the process of writing calligraphy khat is not something that is just random, it contains various meanings and rules in each of its creation (Fauzi & Thohir, 2021).

Khat is the art of writing letters and characters with certain styles and shapes that come from ancient traditions and cultures. In Indonesia, khat is widely used as a decoration on buildings, art objects, or as a material for making artworks. However, with the development of technology, khat can also be found in digital formats such as on apps, websites, and social media. In utilizing khat in digital format, one of the main challenges is how to classify types of khat automatically and efficiently. Manual calligraphic khat requires time, human resources, and considerable costs.

Therefore, the use of computational methods to automatically classify khat is becoming increasingly important. One of the computational methods used to classify khat is using Convolutional Neural Network (CNN), which is a type of Deep Learning that is usually used in image recognition tasks. However, CNN has the disadvantage of requiring considerable computational time and resources, making its use less efficient for khat applications on resource-constrained devices such as smartphones or tablets (Compression, 2021).

To overcome these problems, research on the use of Lightweight CNN for khat classification is becoming increasingly interesting. Lightweight CNN is a CNN specifically designed for image recognition tasks using a smaller number of parameters, so that it can produce smaller, lighter, and faster models in making predictions (He & Li, 2020). Therefore, the use of Lightweight CNN in khat classification is an attractive solution to produce efficient and fast models in classifying khat on devices with limited resources (He & Li, 2020).

This research is related to Offline Handwritten Chinese Text Recognition with Convolutional Neural Networks. In both studies, different neural network architectures and residual block configurations were explored to improve the performance and accuracy results of handwriting recognition, both in Chinese and Arabic. These results were achieved by exploring the best configuration of the basic VGG model and exploring an improved residual network architecture with the same dropout strategy. In the study, increasing the number of residual blocks in the network architecture decreased the character error rate (CER) and improved the accuracy of the model. The best residual model, ResNet-2451, reduced the CER from 10.37% to 9.41% and after being trained with the baseline and synthetic data, it reduced the CER by 2.6%. The best result of the study was a CER of 6.81%, which was also the best result published in the 2013 ICDAR competition set from CASIA-HWDB. Thus, this research makes an important contribution in developing handwritten character recognition technology using the Lightweight Convolutional Neural Network model that can be used in practical applications such as OCR and automatic checking of documents (Liu et al., 2020).

Previous research with the title Arabic Handwritten Characters Recognition using Convolutional Neural Network is also related to this research. which managed to achieve an accuracy of 97.2%. can be seen in the application of similar training and model evaluation techniques. In the research on Arabic handwritten character recognition using CNN, researchers evaluated their models by choosing different numbers of epochs to find the best accuracy. Just like the previous study, the results showed a significant increase in accuracy with the increase in the number of epochs. However, this study shows that some letters in Arabic have similar shapes that can confuse the model and result in miss-classification. This can be improved by adding more training data or considering the use of other techniques to improve recognition accuracy (Aljarrah et al., 2021). Related research also discusses classification with the title Fast Korean Syllable Recognition with Letter-based Convolutional Neural Networks, the dataset used in this study consists of images consisting of 2374 syllables and the test set consists of images consisting of 8798 syllables. In an attempt to test the capabilities of the model, several evaluation techniques were used, including per-letter model accuracy and per-syllable model accuracy measurements. The experimental results show that the model is able to recognize syllables from the training set with a fairly high accuracy, although there are still some letters that are difficult to recognize accurately. This research can serve as a basis for the development of more advanced Arabic handwriting recognition technology in the future (Zatsepin et al., 2019).

Therefore, this research aims to implement the Lightweight Convolutional Neural Network (LCNN) model architecture for the classification of khat naskhi and riqah. This classification method uses a deep learning approach to identify patterns contained in the writings of khat naskhi and riqah. In this research, a Lightweight CNN model will be built consisting of several layers of convolution and dimensionality reduction (Bulla, 2022). In addition, the training data and test data used in this research are obtained from a collection of khat naskhi and riqah data that has been collected from various sources.

The results of this research are expected to help experts in accelerating the classification process of khat naskhi and riqah with better accuracy. In addition, this research can also be a reference for other researchers who are interested in developing khat classification methods using Lightweight CNN architecture.

2. RESEARCH METHOD

The main objective of this research is to classify khat using lightweight CNN, to produce a khat classification model that is efficient and fast in making predictions on devices with limited resources such as smartphones or tablets.

2.1 Pre-processing

At this stage the first thing done in this research is the collection of datasets, which are taken in previous research (Putra et al., 2021), which has produced data in the form of 2 types of khat, namely Nakhi and Riq'ah, each of which each class has 100 image data, which is then divided into 80:20 for testing data.

The pre-processing stage is carried out to obtain maximum performance results in processing the initial data. In this stage, there are four stages, including image resizing, pixel intensity normalization, data standardization and data splitting. First of all, the image dataset is loaded and resized to the same size. Next, the image is converted to grayscale and the pixel intensity is normalized to improve consistency and reduce variability (Akagic & Buza, 2022), because in the previous data there was inconsistency in the background image data. The data is then standardized and divided into training data and test data to ensure the model can be generalized properly. After the image resize and pixel intensity normalization stages are carried out, the next stage is data standardization. This is done to ensure that the data has the same range of values. In data standardization, the average value of the data will be subtracted from each data and then divided by the standard deviation of the data (Han & Han, 2021). By standardizing the data, we can reduce the variability of the data and increase the stability of the model to be used. After that, the last step in data pre-processing is data splitting, which separates the data into training data and test data. Training data is used to train the model and test data is used to test the performance of the model that has been trained (Singh et al., 2023). As for the results of converting sample data that originally had different backgrounds as shown in Figure 1.

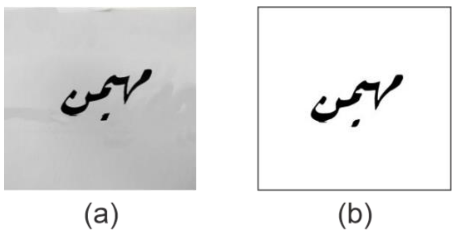

Figure 1. (a) Original data, (b) Background alteration data

The result of background equalization on the dataset aims to equalize the data between datasets 1 and 2, to provide balanced results between the two datasets.

2.2 Lighweight CNN Architecture Design

The design of the Lightweight CNN architecture is an important stage in building an efficient and accurate model. At this stage, key parameters such as kernel/filter size, number of convolution layers, and use of pooling layers need to be determined to process the images effectively. To obtain optimal performance, hyperparameters such as learning rate, number of epochs, and batch size need to be optimized. Regularization techniques such as dropout can also be used to avoid overfitting during model training (Sangeroki & Cenggoro, 2021). The performance of the model on the validation data can be monitored during the training process to adjust the hyperparameters to achieve optimal performance. Once the model is trained, the model performance can be tested on test data to measure accuracy. In the development of Lightweight CNN architecture, the main goal is to obtain an efficient and accurate model using few parameters and limited computing resources.

In this study, we implemented a model with multiple layers of convolution and pooling for image classification. The model consists of a convolution layer with a 3x3 kernel filter and relu activation, and a depthwise convolution that handles each channel separately (Truong et al., 2018). This depthwise convolution is followed by a regular convolution layer with a 1x1 kernel filter and relu activation. To reduce image size and speed up training, a MaxPooling2D layer with a pool size of 2x2 is used. In some layers, Dropout with a value of 0.25 is used to avoid overfitting (Compression, 2021).

After convolution and pooling layers, the result is then flattened to be given to a fully connected layer with relu activation and Dropout 0.5. The final output uses a dense layer with sigmoid activation and a neuron count of 2, indicating a binary class. This model can be used to perform image classification with two different classes. The architecture code utilizes the basic principles of Lightweight CNN, namely using a small kernel filter and depthwise convolution to reduce the number of parameters used and speed up model training, making it suitable for use on limited computing resources (Liu et al., 2020).

The total parameters used in this model are 3,904, all of which can be learned during the training process. This model will be used for data classification into 2 different classes. The Convolutional Layers stage can be seen in Figure 2.

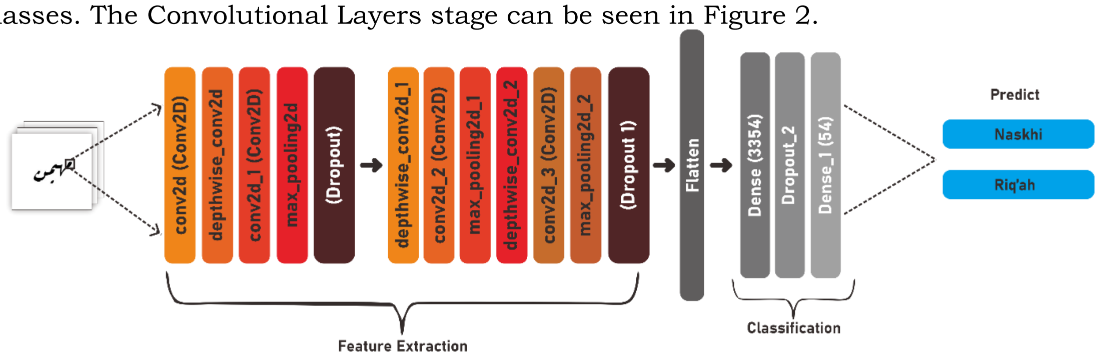

Figure 2. Lightweight CNN Architecture Design

The mechanism of this layer starts with conv2d, which performs convolution on the input image. Next, depthwise_conv2d is performed to deepen the layer by changing the convolution filter to one filter per channel. After that, conv2d_1 is used to process the result of depthwise_conv2d. Then, max_pooling2d is performed to extract important features and reduce the dimensionality of the input data (Veerappa, 2023). Furthermore, dropout is used to prevent overfitting of the data (Madduri et al., 2021). Depthwise_conv2d_1 is then applied to deepen the layer, followed by conv2d_2. Max_pooling2d_1 is then performed again to extract important features. Depthwise_conv2d_2 is then used to deepen the layer, and conv2d_3 is used to process the results from depthwise_conv2d_2.

After that, max_pooling2d_2 is performed again to extract the remaining important features. Dropout_1 is then performed again to prevent overfitting, and flatten is used to flatten the input data into a one-dimensional vector. Dense is used to create a connection between the previous layer and the output layer, followed by dropout_2 to prevent overfitting. Dense_1 is used as the output layer, with the same number of neurons as the number of classes to be classified.

This layer involves several layers consisting of conv2d, depthwise_conv2d, conv2d_1, max_pooling2d, dropout, depthwise_conv2d_1, conv2d_2, max_pooling2d_1, depthwise_conv2d_2, conv2d_3, max_pooling2d_2, dropout_1, flatten, dense, dropout_2, and dense_1. These layers successively perform convolution operations on the input image, deepen the layer by changing the convolution filter to one filter per channel, process the result of convolution, extract important features, prevent overfitting, and make connections between the previous layer and the output layer (Meng et al., 2022). Each layer is performed to optimize the image classification performance of the artificial neural network model. These layers were chosen because they have a good ability to extract important features from the input image and prevent overfitting of the data (Compression, 2021).

3. RESULTS AND DISCUSSIONS

Before testing, data totaling 200 images which are divided into 2 classes, are divided into 2, namely as tain data and test data, where the test divides with a scale of 80: 20 randomly. So that the train data amounted to 160 images and the test data amounted to 40 images. The test results that have been carried out by applying the Lightweight CNN algorithm are as follows:

3.1 Research Results

The process of all model training carried out in this study was executed on Google Colab using a GPU that has been provided by a virtual machine. The training graph can be seen in Figure 3.

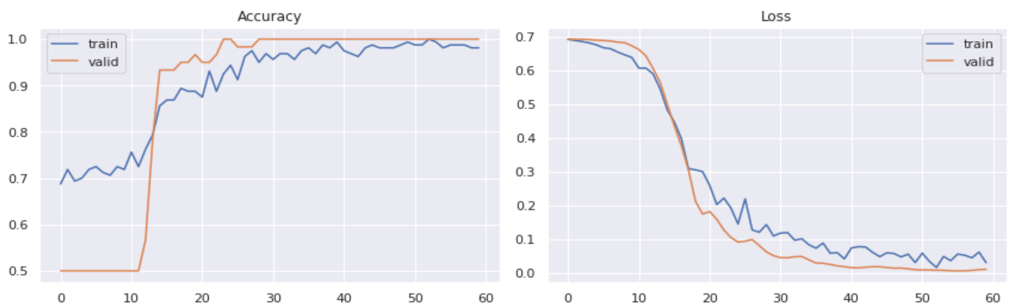

Figure 3. Training Graph Using Lightweight CNN

The results show that the use of the Lightweight CNN layer in classifying khat naskhi and riq'ah on the prepared dataset has excellent performance. The evaluation results show that the model has an accuracy value of 98.75% and a loss value of 0.0378 on training data, while on validation data, the model managed to achieve 100% accuracy with a loss value of 0.0014. From these results, it can be concluded that the classification model using the Lightweight CNN layer can be used as an effective alternative in classifying types of Arabic writing, especially in recognizing certain types of khat such as naskhi and riq'ah. Based on the model evaluation results, a confusion matrix is also created as follows:

Table 1. Confusion Matrix Results

Data	Naskhi Prediction	Riq'ah Prediction
Actual Naskhi	75	25
Actual Riq'ah	20	80

From the confusion matrix, it can be seen that the model successfully classified 75 khat naskhi images correctly and 80 khat riq'ah images correctly. However, there were 25 khat naskhi images that were incorrectly classified as khat riq'ah, and 20 khat riq'ah images that were incorrectly classified as khat naskhi. Although there are some cases where the model is wrong in performing the classification, overall the model performs very well and can be used as an effective alternative in classifying types of Arabic writing, especially in recognizing certain types of khat such as naskhi and riq'ah.

In addition, the model evaluation results also show good recall and precision values for both classes, which are 75% for the naskhi class and 80% for the riq'ah class. This shows that the model has a good ability to recognize both types of khat. In addition, the F1-score value produced by the model is also quite high, which is 0.7692 for the naskhi class and 0.8333 for the riq'ah class. From these results, it can be concluded that the classification model using the Lightweight CNN layer on the naskhi and riq'ah khat datasets is able to provide excellent and consistent results in recognizing both types of khat. The visualization of the confusion matrix can be seen in Figure 5.

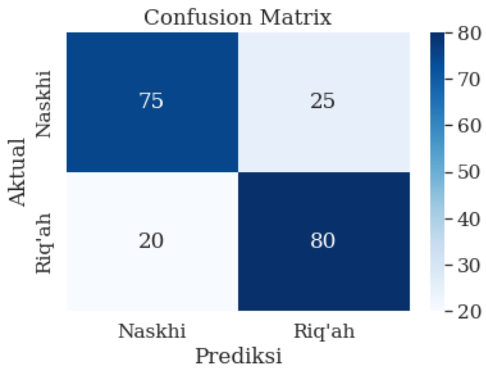

Figure 4. Visularization of Confusion Matrix

The confusion matrix visualization above illustrates the results of evaluating the performance of the classification model using precision and recall metrics. The confusion matrix has 4 cells, namely True Positive (TP) = 75, False Positive (FP) = 25, False Negative (FN) = 20, and True Negative (TN) = 80.

Table 2. Evaluation Model

Class	Recall	Precision	F1-Score
Naskhi	75%	75%	0.7692
Riq'ah	80%	80%	0.833

By using heatmaps from the seaborn library, the confusion matrix can be visualized clearly and easily understood. In the visualization, the x and y axes show the predicted and actual labels of the model, while the value of each cell is shown with annotations and the numeric format 'g'. The visualization also displays the plot title, x- and y-axis labels, and the label of each tick of the axis. The confusion matrix visualization above illustrates the results of evaluating the performance of the classification model using the precision and recall metrics. The confusion matrix has 4 cells, namely True Positive (TP) = 75, False Positive (FP) = 25, False Negative (FN) = 20, and True Negative (TN) = 80. The test results can be seen in Figure 4.

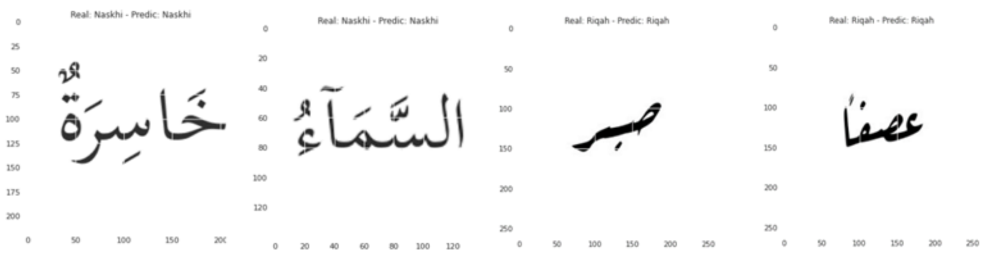

Figure 5. Prediction Results of 2 Khat Types (Naskhi and Riq'ah)

In addition, the results obtained also show that the model is able to perform classification with a relatively fast time, which is 2s 375ms per step, so the model can be implemented well in systems that require high data processing speed. With the results of this study, it is hoped that the classification model using the Lightweight CNN layer can provide an effective solution in recognizing types of Arabic writing, and can be applied in various applications that require classification of writing types such as in handwriting character recognition, vehicle license plate recognition, and so on.

4. CONCLUSION

Based on the results of the research and evaluation carried out, it can be concluded that the use of the Lightweight CNN layer in classifying khat naskhi and riq'ah on the prepared dataset provides excellent results, with accuracy reaching 98.75% on training data and 100% on validation data. In terms of time, the model is able to perform classification quickly, which is 2s 375ms per step. Therefore, the classification model using the Lightweight CNN layer can be used as an effective alternative in classifying types of Arabic writing, especially in recognizing certain types of khat such as naskhi and riq'ah. This research also provides limitations on the dataset used in this research. The dataset used in this study is only limited to certain types of Arabic writing, namely khat naskhi and riq'ah, so the results of this study may not be generalized to other types of Arabic writing. In addition, this research is also limited to the use of the Lightweight CNN layer as a classification method, not including the implementation on the product.

As further suggestions, further development can be carried out on this classification model by using a larger and more diverse dataset, so as to improve the performance of the model in recognizing more and complex types of Arabic writing. In addition, it is also necessary to evaluate and compare with other classification models to find out the better performance in classifying Arabic writing types. In addition, this model can also be implemented in more complex applications, such as handwriting recognition, so that it can provide greater benefits in various daily life applications.

ACKNOWLEDGEMENTS

Thus the journal "Lightweight Convolutional Neural Network for Khat Naskhi and Riq'ah Classification" from our research at Darussalam Gontor University has been published. We would like to express our deepest gratitude to all those who have participated in the completion of this journal. Thank you to the lecturers and academic staff who have guided and provided support during the research process. We would also like to thank our friends and family who have encouraged and motivated us to keep trying. Hopefully the results of this research can provide benefits for the academic world and the development of technology in the future. Thank you very much.

REFERENCES

Akagic, A., & Buza, E. (2022). LW-FIRE: A Lightweight Wildfire Image Classification with a Deep Convolutional Neural Network. *Applied Sciences (Switzerland)*, *12*(5). https://doi.org/10.3390/app12052646

Aljarrah, M. N., Zyout, M. M., & Duwairi, R. (2021). Arabic Handwritten Characters Recognition Using Convolutional Neural Network. *2021 12th International Conference on Information and Communication Systems, ICICS 2021*, 182–188. https://doi.org/10.1109/ICICS52457.2021.9464596

Bulla, P. (2022). Traffic Sign Detection and Recognition Based on Convolutional Neural Network. *International Journal on Recent and Innovation Trends in Computing and Communication*, *10*(4), 43–53. https://doi.org/10.17762/ijritcc.v10i4.5533

Compression, L. D. (2021). *Implementation of Lightweight Convolutional Neural*. 1–21.

Fauzi, M., & Thohir, M. (2021). Pembelajaran Kaligrafi Arab untuk Meningkatkan Maharah Al-Kitabah. *EL-IBTIKAR: Jurnal Pendidikan Bahasa Arab*, *9*(2), 226.

https://doi.org/10.24235/ibtikar.v9i2.6554

Han, Y., & Han, Y. (2021). A Deep Lightweight Convolutional Neural Network Method for Real-Time Small Object Detection in Optical Remote Sensing Images. *Sensing and Imaging*, *22*(1), 1–15. https://doi.org/10.1007/s11220-021-00348-0

He, Y., & Li, T. (2020). A lightweight CNN model and its application in intelligent practical teaching evaluation. *MATEC Web of Conferences*, *309*, 05016. https://doi.org/10.1051/matecconf/202030905016

Liu, B., Xu, X., & Zhang, Y. (2020). *Offline Handwritten Chinese Text Recognition with Convolutional Neural Networks*. http://arxiv.org/abs/2006.15619

Madduri, A., Adusumalli, S. S., Katragadda, H. S., Dontireddy, M. K. R., & Suhasini, P. S. (2021). Classification of Breast Cancer Histopathological Images using Convolutional Neural Networks. *Proceedings of the 8th International Conference on Signal Processing and Integrated Networks, SPIN 2021*, 755–759. https://doi.org/10.1109/SPIN52536.2021.9566015

Meng, Z., Jiao, L., Liang, M., & Zhao, F. (2022). A Lightweight Spectral-Spatial Convolution Module for Hyperspectral Image Classification. *IEEE Geoscience and Remote Sensing Letters*, *19*, 1–5. https://doi.org/10.1109/LGRS.2021.3069202

Putra, O. V., Musthafa, A., Nur, M., & Rido, M. (2021). Classification of Calligraphy Writing Types Using Convolutional Neural Network Method (CNN). *Procedia of Engineering and Life Science*, *2*(1), 2–8. https://doi.org/10.21070/pels.v2i0.1136

Sangeroki, B. A., & Cenggoro, T. W. (2021). A Fast and Accurate Model of Thoracic Disease Detection by Integrating Attention Mechanism to a Lightweight Convolutional Neural Network. *Procedia Computer Science*, *179*(2019), 112–118. https://doi.org/10.1016/j.procs.2020.12.015

Singh, V., Chug, A., & Singh, A. P. (2023). Classification of Beans Leaf Diseases using Fine Tuned CNN Classification of Beans Leaf Diseases using Fine Tuned CNN Model Model. *Procedia Computer Science*, *218*(2022), 348–356. https://doi.org/10.1016/j.procs.2023.01.017

Truong, T. D., Nguyen, V. T., & Tran, M. T. (2018). Lightweight deep convolutional network for tiny object recognition. *ICPRAM 2018 - Proceedings of the 7th International Conference on Pattern Recognition Applications and Methods*, *2018-January*(Icpram), 675–682. https://doi.org/10.5220/0006752006750682

Veerappa, N. (2023). Classification Classification Of Of Lung Lung Cancer Cancer Using Using Lightweight Lightweight Deep Deep Neural Neural Networks Networks. *Procedia Computer Science*, *218*, 1869–1877. https://doi.org/10.1016/j.procs.2023.01.164

Zatsepin, M., Vatlin, Y., Chulinin, I., & Zhuravlev, A. (2019). Fast Korean syllable recognition with letter-based convolutional neural networks. *2019 International Conference on Document Analysis and Recognition Workshops, ICDARW 2019*, *2019-January*, 10–13. https://doi.org/10.1109/ICDARW.2019.60124