International Journal of Applied Sciences and Smart Technologies

Volume 2, Issue 2, pages 179–196

p-ISSN 2655-8564, e-ISSN 2685-9432

Backpropagation Neural Network for Book Classification Using the Image Cover

I Putu Budhi Darma Purwanta[1,*], Cyprianus Kuntoro Adi[1], Ni Putu Novita Puspa Dewi[1]

[1]*Department of Informatics, Faculty of Science and Technology, Sanata Dharma University, Yogyakarta, Indonesia*

[*]*Corresponding Author: budhidarmap@gmail.com*

(Received 14-06-2020; Revised 26-08-2020; Accepted 26-08-2020)

Abstract

Artificial Neural Networks are known to provide a good model for classification. The goal of this research is to classify books in Bahasa (Bahasa Indonesia) using its cover. The data is in the form of scanned images, each with the size of 300 cm height, 130 cm width, and 96 dpi image resolution the research conducted features extraction using image processing method, MSER (Maximally Stable Externally Regions) to identify the area of book title, and Tesseract Optical Character Recognition (OCR) to detect the title. Next, features extracted from MSER and OCR are converted into a numerical matrix as the input to the Backpropagation Artificial Neural Network. The accuracy obtained using one hidden layer and 15 neurons is 63.31%. Meanwhile, the evaluation using 2 hidden layers with a combination of 15 and 35 neurons resulted in accuracy of 79.89%. The ability of the model to classify the book was affected by the image quality, variation, and number of training data.

Keywords: classification, image processing, MSER, tesseract, text processing, backpropagation artificial neural network

Volume 2, Issue 2, pages 179–196
p-ISSN 2655-8564, e-ISSN 2685-9432

1 Introduction

Book classification is a very common research topic. There are, however, not so many papers doing classification using book cover to represent book content. Variety of backgrounds, fonts, and locations of book titles, as well as similar titles to represent different content, makes classification approach difficult [1]. Iwana, et al. (2016) conducted classification using book cover. They employed features such as font types, color characteristics, color contrast, image characteristics, and writing characteristics [1]. Meanwhile, this study focuses more on book title as a mean to classify book content using Artificial

Backpropagation is a widely used to train feed-forward neural networks for supervised learning. Backpropagation method computes the gradient of the loss or error function with respect to each weight by the chain rule. It calculates the gradient one layer at a time, iterating backward from the last layer to avoid redundant calculations of the intermediate terms in the chain rule. This way of computing similar to what dynamic programming do [2]. Various papers employ back-propagation approach for their researches. Karlissa, et al. (2018), for example, developed backpropagation-based autonomous control system for three-wheeled robot [3]. Maneesukasem and Pintavirooj (2012) used feed forward backpropagation to segment urine sediment image to identify crystals, casts, red blood cells, white blood cells and bacteria of yeast in urine sediment [4].

Backpropagation processed information retrieval data to look for error tolerance in human information. Backpropagation showed its ability to handle heterogeneous data such as handling features using different words [5]. Mandl (2000), meanwhile, analyzed model-based human-computer interaction that integrate human knowledge into retrieval process in Cognitive Similarity Learning in Information Retrieval (COSIMIR). It applies backpropagation to information retrieval, and integrates human-centered and tolerant computing to retrieval process [5].

The question is, then, how to extract features in data, specifically image data, before processing in backpropagation neural networks? Some previous works show different types of image transformations using region detector methods, such as MSER, Harris-

Volume 2, Issue 2, pages 179–196
p-ISSN 2655-8564, e-ISSN 2685-9432

Affine, Hessian-Affine, Intensity-extreme-based regions, Edge Based Regions, and Salient [6–10]. It seemed that MSER performed well on images that contained homogeneous regions with distinctive boundaries [6]. In addition, for hand-writing studies, MSER method had been improved using additional canny edge detection and custom threshold upon data [11-14].

For image data such as book cover, a relevant feature representation to know the contents of the book through its cover is the title of the book [1]. However, the location and font type of book title are various. Some object detection techniques, such as Optical Character Recognition (OCR) [14-17] can be utilized to capture book title. The OCR method developed in Keras Deep Learning and Tesseract is worth to consider. Keras Deep Learning, as an example, was able to identify characters with an error rate 3.72% [14]. Meanwhile, Tesseract – Google open source optical character recognition – was capable of identifying Roman and Chinese characters with 90% accuracy [15–17].

The difficult part of classifying books using their title lies on two things, namely, how to relate book title to its content, and the complexity of word structures. First, book title versus its content. A book with title "Investasi Rohani" (English: "Spiritual Investment") or "Matematika Pahala" (English: "Mathematics of Reward"), for example, neither discuss business nor mathematical problems. Instead, they are about spiritual or religious matter. Secondly, complex word structures in Bahasa Indonesia.

Bahasa has many combinations in suffixes and prefixes of its word structures [18-19]. It needs some methods to identify its root words in order to get clear information on the book title. Some words change its meaning after additional suffix or prefix. The word "ajar" (English: "teach"), for example, gets suffix "ber-" becomes "belajar" (English: "learning") instead of "berajar" (English: "teaching"). Both of those words have different meanings. So, it needs an extra effort to create a dictionary that able to distinguish the words which have already changed due to additional suffix or prefix.

Motivated by the recent success of method in image processing and artificial neural networks, the goal of this study is to classify books in Bahasa using the scanned image of its cover. This paper combines the technique of MSER, Tesseract, and ANN backpropagation to propose better result in classification. The contributions of this paper are mainly in the text processing of Bahasa and classification using the

Volume 2, Issue 2, pages 179–196
p-ISSN 2655-8564, e-ISSN 2685-9432

Backpropagation Neural Networks. Hopefully, this classification approach helps librarians and person in charged of bookstores finding books easier and faster. The second part of this paper discusses backpropagation method of classification, followed by its result and discussion in third part. This paper highlights its main contribution and conclusion in the fourth part.

2 Methodology

Figure 1 shows the diagram of the research. There are two important steps, namely, feature extraction and classification. Feature extraction includes data preprocessing to highlights area of the text and removes unnecessary background, optical character recognition (OCR), and feature extraction of the OCR result. The classification step builds classifier using data training and its label. The performance of classifier is evaluated using data testing. The optimal model is determined through k-fold combination of data training and data testing.

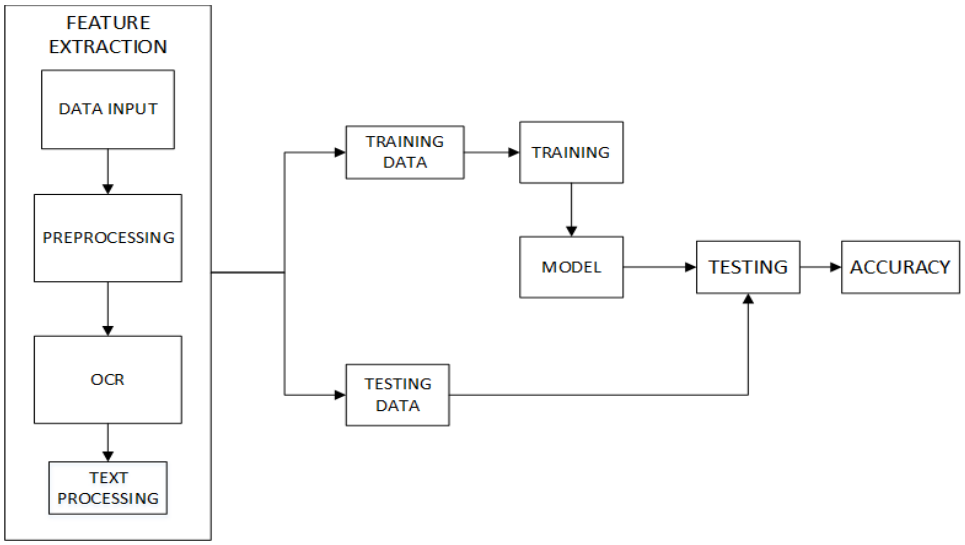


Figure 1. Block diagram of classification method

2.1 Feature Extraction

Feature analysis or extraction in this research consists of three processes, namely: preprocessing, character detection with OCR and text processing to identify words that formed book title. After explanation of data, each process is presented.

Volume 2, Issue 2, pages 179–196

p-ISSN 2655-8564, e-ISSN 2685-9432

2.1.1 Data

Data of scanned image book-covers are from The Kanisius Yogyakarta Publisher, categorized into 3 classes: 53 of philosophy books, 101 of religious books, and 200 of education books. Each image has size of 300 cm height and 130 cm width with 96 dpi image resolution. Every image has a label which assigned to its correct class. The labeling process is done manually. In order to compare the performance of the classification model, the training process is done using 2 types of data: the first is data of the book title recognized by OCR and the second is data without OCR. The labelling process of the data took the same process of the image preprocessing data.

2.1.2 Image Prepocessing

As a book title might have different color and gradation from its background, the grayscale image shown in Figure 2 is used as the input of the process of text detection by MSER. MSER captured the object on book cover using parameter thresholds of 12 and 20 and 1200 of region area. The result of MSER was the coordinate value of the object on the book cover. The detected object may contain the title and the background. After getting the object coordinates, then the process changes the background by modifying the value of each point of 0 (for non-title text) and 1 for text title. Figure 3 shows the detection result of MSER, and Figure 4 shows the result of the changing background of book title.

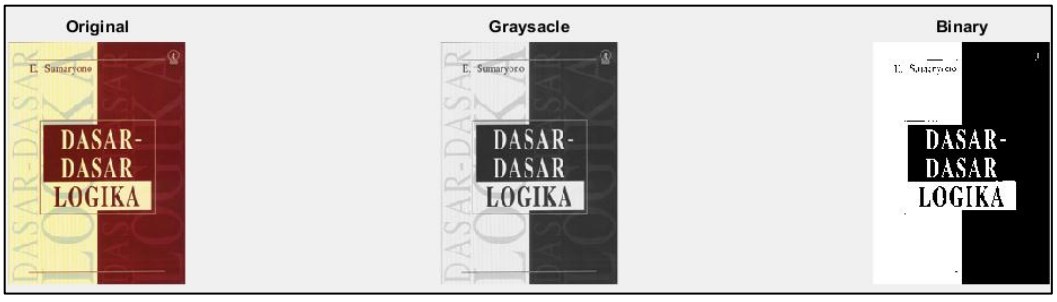


Figure 2. Book cover preprocessing: original - grayscale - binary form.

International Journal of Applied Sciences and Smart Technologies

Volume 2, Issue 2, pages 179–196

p-ISSN 2655-8564, e-ISSN 2685-9432

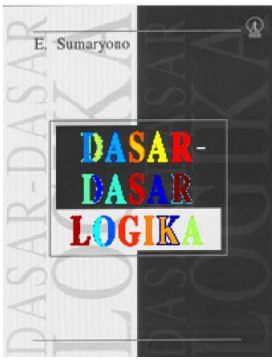


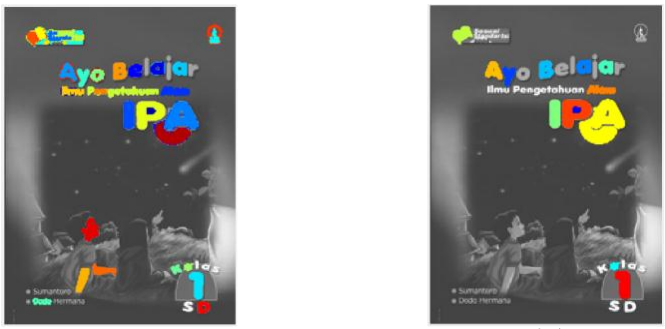
Figure 3. Detection with MSER.



Figure 4. Modification of book cover background (value 0) and book title (value 1)

The MSER capabilities can be improved by adding canny edge detection in the process. This research tunes some parameters to find out the optimal MSER detection. It starts with threshold number of 12 for area filtering, and sets the region area between 20 and 1200. This tuning is able to handle more images, but in some cases it removes some part of the book title. The different setup using a smaller number of 5 for the threshold, and region area between 20 and 800 results in better MSER detection as shown in Figure 5.

Finally this paper used the combination of 2 values. In the first run, it sets up the threshold into 12 and the Region Area was between 20 and 1200. If the value was less than 1, the threshold was set up into 5 and the region area was set up between 20 and 800. As Figure 5 showed, when it sets up a smaller value, the title detection result were getting better, but other objects outside the title of the book were also included as part of the book title.



(a) **(b)**

Figure 5. (a) Using region area of 20-1200 and 12 threshold delta; (b) Using region area of 20-800 and 10 threshold delta.

2.1.3 Optical Character Recognition (OCR)

Image preprocessing with MSER allocates book title area for further recognition. Tesseract OCR, then converts image to texts that are part of a book title. Variation of

Volume 2, Issue 2, pages 179–196

p-ISSN 2655-8564, e-ISSN 2685-9432

image sizes show the performance of Tesseract OCR to recognize word inside the image. Table 1 provides minimum image size Tesseract OCR able to identify. Image with minimum size of 70px seems a better candidate for character recognition. The words or terms identified in OCR process then become input for next process, namely: text processing.

Table 1. Comparison of Text images Recognition

Image	Explain	Result
	Image with height 17px and line weight 4px	**Identified** Height and line weight are identified
	Image with height 14px and line weight 4px	**Identified** Height and line weight are identified
	Image with height 10px and line weight 2px	**Unidentified** Height is unidentified
	Image with height 11px and line weight 1px	**Identified** line weight is unidentified

2.1.4 Text Processing

Text processing simply means to bring text into a form that is analyzable for certain task. There are different ways to process text. To find out words from certain text, this research uses steps as follows: case-folding (lowercasing), tokenization (separate words, characters from a text), stop-word removal (removing low information words from text) and stemming (removing prefix and suffix, and reducing inflection in word to its root form). The result is a bag of words modeling - word database that store unique words retrieved from book cover. Table 2 illustrates how text data is represented in numeric vector based on how many terms or words are in the text. Table 3 shows a guideline for cutting off word prefix in Bahasa. Table 4 shows how words or terms extracted from book cover is represented in a database or term matrix. Matrix size depends on how many words are in the database.

Table 2. Illustration of transform string data into numeric data

Database / Data	"One"	"Two"	"Five"
"One Two One"	2	1	0
"Two Three"	0	1	0
"Five One Four"	1	0	1
"Two Five"	0	1	1

International Journal of Applied Sciences and Smart Technologies

Volume 2, Issue 2, pages 179–196

p-ISSN 2655-8564, e-ISSN 2685-9432

Table 3. List of prefix change [19]

Prefix	Phoneme or terms	Changing
meng-	*/r, l, m, n, w, y, ng, ny/*	*me-*
	/p, b, f, v/	*mem-*
	/t, d, c, j, z, sy/	*meng-*
	Phrase< more than one words	*menge-*
peng-	*/r, l, m, n, w, y, ng, ny/*	*pe-*
	/p, b, f, v/	*pem-*
	/t, d, c, j, z, sy/	*peng-*
	Phrase< more than one words	*menge-*
ber-	*/r/*	*be-*
	/ajar/	*bel-*
per-	Affinity	*pe-*
	/ajar/	*pel-*
ter-	*/r/*	*te-*

Table 4. Data input illustration

No	Words Database / Title	'PAHARGYAN'	'BOJANA'	'KURBAN'	'RAKA'	…	'MANUSIA'
1	'PAHARGYAN BOJANA KURBAN'	1	1	1	0	…	0
2	'RAKA AGUNG SEBUAH RENUNGAN'	0	0	0	1	…	0
3	'KURBAN UNTUK ALLAH'	0	0	1	0	…	0
…	…	…	…	…	…	…	…
354	'FILSAFAT MANUSIA'	0	0	0	0	…	1

2.2 Classification with Backprpagation Neural Network

This section describes the classification model architecture and the experimental setup applied to the classification model training and testing process.

International Journal of Applied Sciences and Smart Technologies

Volume 2, Issue 2, pages 179–196

p-ISSN 2655-8564, e-ISSN 2685-9432

2.2.1 Classification with Backprpagation Neural Network

The neural network architecture for classification is shown in Figure 6. The input features are word vectors extracted from book cover images. The size of the input matrix is (P x Q), where P is the number of word features (in this case, 489 features) and Q is the number of image data. This research employs two hidden layers with two neuron output. The optimal architecture is found through combination number of neuron inside the hidden layer. The neuron would be varied from 5 to 40; with log-sigmoid transfer function (equation (1)) inside the hidden layer and pure-linear transfer function (equation (2)) for the output, where:

$$a = \frac{1}{1 + e^{-n}}, \tag{1}$$

$$a = n. \tag{2}$$

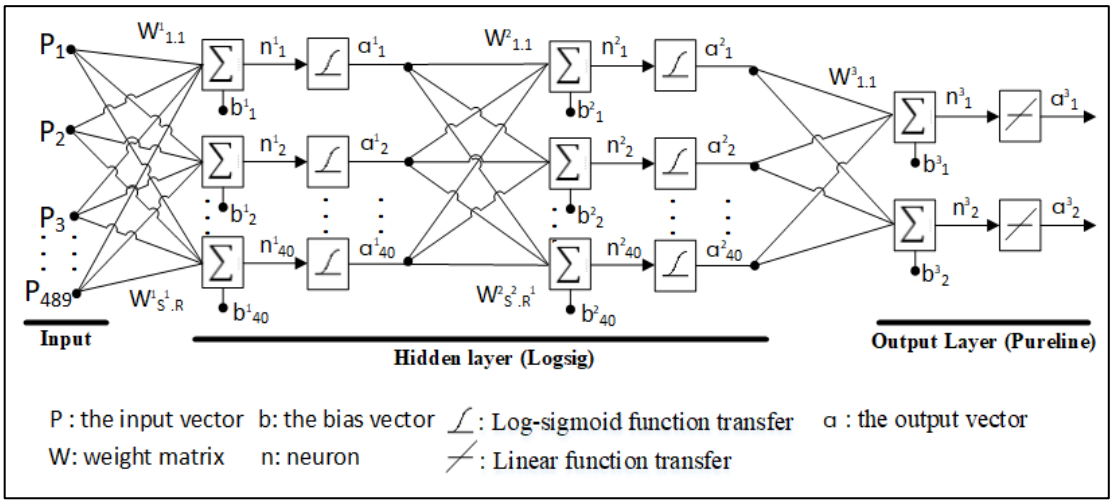


Figure 6. Network Architecture for Training and Testing [20]

2.2.2 Experimental Setup

Backpropagation artificial neural network in this experiment builds a classifier model of a set of data that contains both the inputs and the desired outputs. The data, known as training data, consists of a set of training examples. Each training example is represented by an array of vector, sometimes called a feature vector, and the training data is represented by a matrix. In this study, the data matrix consists of 348 samples with 489 features and a class label. Three fold cross validation approach for obtaining best model accuracy is employed as follows: 53 philosophical books are divided into 18

Volume 2, Issue 2, pages 179–196
p-ISSN 2655-8564, e-ISSN 2685-9432

data for training, 18 for validations, and 17 for data testing; 101 religious books are separated into 34 samples for training, 34 for validation, and 33 for testing; meanwhile 200 education books are divided into 67 sample training, 67 data validation and 66 data testing (see Table 5). Three fold data training and label training build three classifier model. Each model is evaluated using its corresponding testing data. The overall accuracy is taken from its average. The experiment started with one hidden layer neural network architecture with variation of number of neuron in the hidden layer, followed by two hidden layer neural network architecture as a way to find out an optimal architecture.

Table 5. Data Composition for training and testing the classification model

Label Class	Training Data	Validation Data	Testing Data
Religious	34	34	33
School Education	67	67	66
Philosophical	18	18	17

3 Results and Discussion

This section presents the experimental results. The experimental results present the performance of the proposed model on using 1 and 2 hidden layers. In regard to all the experiments done, we also presented the optimal model. To ensure the reliability of the model in performing classification, the proposed model will be compared with 2 other models.

3.1 Model with One Hidden Layer

This research trained the classifier model using a combination of neurons (5, 10, 15, 20, 25, 30, 35 and 40) in the first hidden layer. The classifier model is build using training data and is evaluated using testing data that are preprocessed with OCR and data that are not preprocessed with OCR. Figure 7 shows that classifier model trained with data preprocessed without OCR outperformed model trained with OCR. It seemed that image preprocessing and OCR are not able to offer good features for classifier model building. On average, the performance are 10% below classifier model that is based on non-OCR feature (data label are identified manually).

International Journal of Applied Sciences and Smart Technologies

Volume 2, Issue 2, pages 179–196

p-ISSN 2655-8564, e-ISSN 2685-9432

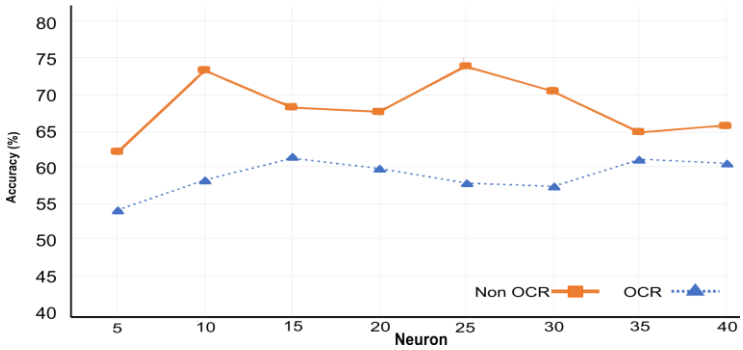


Figure 7. Comparison of the results between OCR data and non-OCR using 1 hidden layer

3.2 Model with Two Hidden Layer

The second model is developed from the previous section by adding one more hidden layer. This classifier model is trained and tested using both data preprocessed with OCR and data not processed with OCR. Figure 8 shows better performance of the model with 2 hidden layers, trained with OCR data and 15 neurons in the first layer, where the accuracy of the model increased and outperformed the model with only 1 hidden layer. The highest accuracy is 63.31%.

This research utilized non-OCR data into account to observe its performance. This two-hidden-layer neural networks consist of 25 neurons in the first layer and various number of neurons in the second hidden layer. Figure 8 shows that the highest accuracy of 79.89% is obtained using 15 neurons in the second hidden layer. The lowest accuracy of 67.82% is obtained at 5 neurons in the second hidden layer. The gap between the lowest and the highest accuracy is 12%. The results implied that there is an increase in the performance of the model trained with non-OCR data with 2 hidden layers neural network architecture. As can be seen from Figures 8, the non-OCR data has reach higher accuracy compared to classifier model trained with OCR data.

International Journal of Applied Sciences and Smart Technologies

Volume 2, Issue 2, pages 179–196

p-ISSN 2655-8564, e-ISSN 2685-9432

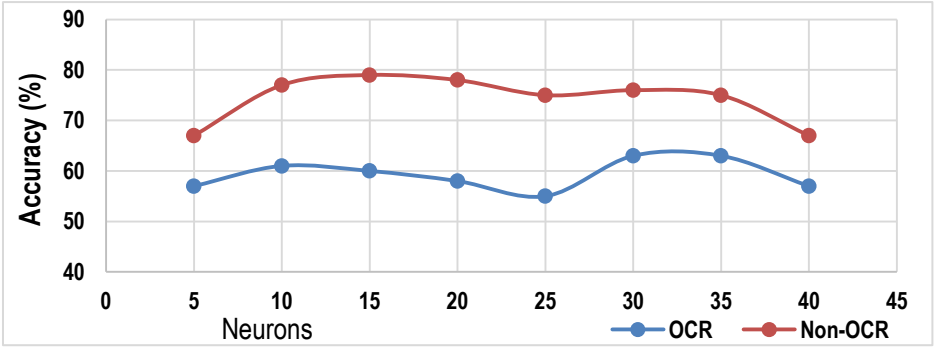


Figure 8. Model performance of two-hidden-layer backpropagation neural networks

3.3 Optimal Model

From the experiment done, this research obtained the optimal model with two hidden layer neural network architecture as seen in Figure 9. This optimal model employed 489 dimension feature input, with log-sigmoid activation function (equation 1) both in the first and second hidden layer, 15 neurons in the first hidden layer and 35 neurons in the second hidden layer. The output layer used a linear activation function (equation 2) and 2 neurons to represent the class output. Table 6 presents one of the results of the confusion matrix of 3 fold cross-validation applied to the most optimal model. Accuracy calculation of the model (table 6, equation 3) where sum of true positive (TP) and true negative (TN) divide by sum of TP, TN, false positive (FP) and false negative (FN).

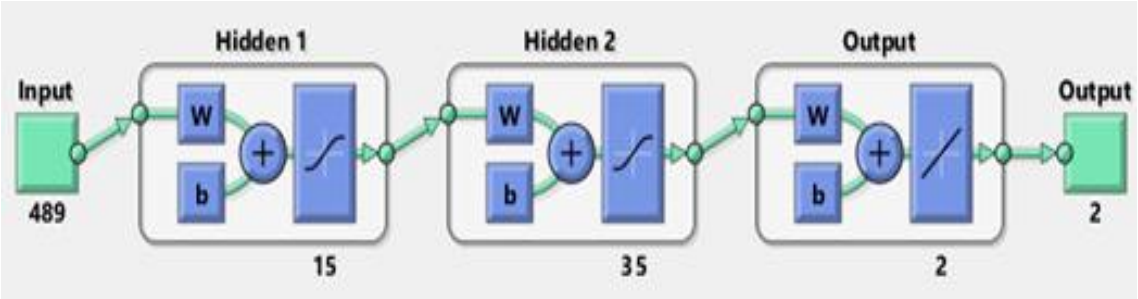


Figure 9. Architecture of the optimal model

$$\text{Accuracy} = \frac{TP + TN}{(TP + TN + FP + FN)}$$

$$= \frac{12 + 55 + 13}{(12 + 11 + 10 + 4 + 55 + 7 + 2 + 2 + 13)} \quad (3)$$

$$= \frac{80}{116} = 0.689655172$$

International Journal of Applied Sciences and Smart Technologies

Volume 2, Issue 2, pages 179–196

p-ISSN 2655-8564, e-ISSN 2685-9432

Table 6. Confusion Matrix

Predict / True	Religion	Education	Philosophy
Religion	12	11	10
Education	4	55	7
Philosophy	2	2	13

3.4 Single Data Testing

We evaluated the proposed model with new data testing. Tables 7 and 8 show the classification results. Based on the results, no model reached 90% of accuracy. Some different books are classified as the same category. Book with part of the title contained term "Katolik", for example, may belong to class Religion or Education. The model, however, classified them as similar category: Religion book. It is possible that the classifier had difficulties in identifying those two classes due to high similarity of books title, or limited data training to model the classifier.

3.5 Comparison on Other Classification Methods

The proposed model then compared with two different classification methods, namely Naïve Bayes and Support Vector Machines (SVM) using the same data. A comparison of the average accuracy result is shown in Table 9. These results confirmed that the proposed method has outperformed the accuracy of other models. It has a gap of 2% with the Naïve Bayesian Probabilistic Classifier and a gap of 12% with SVM using a polynomial kernel.

3.6 Discussion

The highest accuracy of this research obtained using 2 hidden layer models with combinations of 15 and 35 neurons was 63%. Due to several factors (i.e. low image resolution), as many as 32 of 354 images failed to be detected by MSER. In these 32 images, bias occurred between the font and its background. Some titles also blended with the background image of the cover so that MSER is unable to detect it clearly. In the OCR processing, there are a number of problems in the identification, especially those related the size and thickness of the font. Smaller and thin fonts are difficult to identify. As much as 61% of data testing was successfully detected by Tesseract, 27% was incorrectly detected, and 12% was not able to detect. The text processing

Volume 2, Issue 2, pages 179–196
p-ISSN 2655-8564, e-ISSN 2685-9432

contributed in the increasing of 16% correct classification. As mentioned above, when the model met data that contained similar word from different classes, it had difficulties in identifying the difference due to high similarity of books title, or limited data training to model the classifier.

This research then did some data preprocessing without using MSER and Tesseract preprocessing. It seemed that the accuracy increase to 83.33%. The proposed model performed well especially in handling data failed to detect with former model.

Table 7. Result of Single Data Testing using Model with One Hidden Layer

Number of Neuron on First Hidden Layer	ACC (%)
5	50
10	50
15	66.67
20	50
25	50
30	50
35	66.67
40	50

Table 8. Result of Single Data Testing using Model with Two Hidden Layer

Number of Neuron on First Hidden Layer	Number of Neuron on Second Hidden Layer	ACC (%)
15	5	83.33
15	10	83.33
15	15	66.67
15	20	66.67
15	25	66.67
15	30	83.33
15	35	83.33
15	40	66.67

Table 9. Comparison of Model Accuracy

Method	Average Accuracy Model (%)	Average of Misclassified data
Backpropagation	63.31	129
Naïve Bayesian	61.29	137
SVM	51.44	171

4 Conclusion

Backpropagation artificial neural network classifier were able to correctly identify book class or category based on title written in the book cover. Modeling using data without preprocessed with OCR offered better performance than data preprocessed with OCR. The highest accuracy of 63.31% was obtained for model trained with OCR using two hidden layers with 15 neurons in the first hidden layer and 35 neurons in the second hidden layer. Meanwhile, the accuracy of 79.89% was obtained for classifier model build without OCR preprocessed data, using two hidden layer with 25 neurons in the first layer and 15 neurons in the second hidden layer. This approach outperformed two other different models, namely Naïve Bayes and Support Vector Machines, for the same data.

The ability of the model to carry out the classification depends on the image quality, data variation, and the number of training data. To increase the performance of the classifier, some additional observations such as image quality and variation, different image preprocessing method, vocabulary variations, and different modeling approaches are worth to consider.

Declaration of interest

The authors report no conflicts of interest. The authors themselves are responsible for the content and writing of this article.

Acknowledgements

The authors acknowledge The Kanisius Yogyakarta Publisher for their support and encouragement especially for providing the access for book cover images.

References

[1] B. K. Iwana, S. T. R. Rizvi, S. Ahmed, A. Dengel, and S. Uchida, "*Judging a Book by its Cover*." 2016.

[2] I. Goodfellow, Y. Bengio, and A. Courville, *Deep Learning*. MIT Press, 2016.

[3] K. Priandana, I. Abiyoga, W. Wulandari, S. Wahyuni, M. Hardhienata, and A. Buono, "Development of Computational Intelligence-based Control System using Backpropagation Neural Network for Wheeled Robot." in *IEE International*

Conference, **17**, pp. 101–106, 2016.

[4] W. Maneesukasem and C. Pintavirooj, "Urine Sediment Image Segmentation based on Feedforward Backpropagation Neural Network." in *The 5th Biomedical Engineering International Conference (BMEiCON)*, pp. 3–6, 2012.

[5] T. Mandl, "Tolerant Information Retrieval with Backpropagation Networks." *Neural Comput. Appl.*, **9** (4), pp. 280–289, 2000.

[6] K. Mikolajczyk *et al.*, "A Comparison of Affine Region Detectors." *Int. J. Comput. Vis.*, **65** (1), pp. 43–72, 2005.

[7] J. Matas, O. Chum, M. Urban, and T. Pajdla, "Robust wide-baseline stereo from maximally stable extremal regions." *Image Vis. Comput.*, **22** (10), pp. 761–767, 2004.

[8] D. Nistér and H. Stewénius, "Linear Time Maximally Stable Extremal Regions." in *Forsyth D., Torr P., Zisserman A. (eds) Computer Vision – ECCV*, **5303**, Berlin, Heidelberg: Springer, pp. 183–196, *2008.*

[9] X. Shen, G. Hua, L. Williams, and Y. Wu, "Dynamic hand gesture recognition: An exemplar-based approach from motion divergence fields." *Image Vis. Comput.*, **30** (3), pp. 227–235, 2012.

[10] Q. Zhang, Y. Wang, and L. Wang, "Registration of images with affine geometric distortion based on Maximally Stable Extremal Regions and phase congruency." *Image Vis. Comput.*, **36**, pp. 23–39, 2015.

[11] O. Noboyuki, "A Tlreshold Selection Method from Gray-Level Histograms." **20** (1), pp. 62–66, 1979.

[12] H. Chen, S. S. Tsai, G. Schroth, D. M. Chen, R. Grzeszczuk, and B. Girod, "Robust text detection in natural images with edge-enhanced Maximally Stable Extremal Regions." in *18th IEEE International Conference on Image Processing*, pp. 2609–2612, 2011.

[13] M. R. Islam, C. Mondal, M. K. Azam, and A. S. M. J. Islam, "Text detection and recognition using enhanced MSER detection and a novel OCR technique." in *5th International Conference on Informatics, Electronics and Vision (ICIEV)*, pp. 15–20, 2016.

Volume 2, Issue 2, pages 179–196

p-ISSN 2655-8564, e-ISSN 2685-9432

[14] Z. Zhang, K. Qi, K. Chen, C. Li, J. Chen, and H. Guan, "A Novel System for Robust Text Location and Recognition of Book Covers." in *In: Zha H., Taniguchi R., Maybank S. (eds) Computer Vision – ACCV 2009. Lecture Notes in Computer Science*, **5995**, Berlin, Heidelberg: Springer, pp. 608–609, 2010.

[15] R. Smith, "An Overview of the Tesseract OCR Engine." in *Ninth International Conference on Document Analysis and Recognition (ICDAR)*, **2**, pp. 629–633, 2007.

[16] R. Smith, D. Antonova, and D. Lee, "Adapting the Tesseract Open Source OCR Engine for Multilingual OCR." in *Proceedings of the International Workshop on Multilingual OCR* , pp. 1–8, 2009.

[17] Google, "Tesseract Open-Source OCR." [Online]. Available: https://opensource.google.com/projects/tesseract, 2018.

[18] C. D. Manning, P. Raghavan, and H. Schutze, *An Introduction to Information Retrieval*. Cambridge, United Kingdom: Cambridge University Press, 2009.

[19] M. Mustakim, *Seri Penyuluhan Bahasa Indonesia: Bentuk dan Pilihan Kata*. Jakarta: Pemasyarakatan Pusat Pembinaan dan Bahasa, Badan Pengembangan dan Pembinaan Kementerian Pendidikan dan Kebudayaan, 2014.

[20] M. T. Hagan and M. H. Beale, *Neural Network Design*, 2nd ed. Oklahoma, 2014.

International Journal of Applied Sciences and Smart Technologies

Volume 2, Issue 2, pages 179–196

p-ISSN 2655-8564, e-ISSN 2685-9432

This page intentionally left blank