LONTAR_DETC: Dense and High Variance Balinese Character Detection Method in Lontar Manuscripts

NANIK SUCIATI1, (Member, IEEE), NI PUTU SUTRAMIANI1,2, AND DANIEL SIAHAAN1, (Member, IEEE)
1Department of Informatics, Faculty of Intelligent Electrical and Informatics Technology, Institut Teknologi Sepuluh Nopember, Surabaya 60111, Indonesia
2Department of Information Technology, Faculty of Engineering, Universitas Udayana, Badung 80361, Indonesia
Corresponding author: Nanik Suciati (nanik@if.its.ac.id)
This work was supported by the Directorate General of Higher Education, Ministry of Education and Culture Republic of Indonesia under Grant 3/E1/KP.PTNBH/2021.

ABSTRACT This paper proposed LONTAR_DETC, a method to detect handwritten Balinese characters in Lontar manuscripts. LONTAR_DETC is a deep learning architecture based on YOLO. The detection of Balinese characters in Lontar manuscripts is challenging due to the characteristics of Balinese characters in Lontar manuscripts. Balinese characters in Lontar manuscripts are dense, overlapping, have high variance, contain noise, and classes of these characters are imbalanced. The proposed method consists of three steps, namely data generation, Lontar manuscript annotation, and Balinese character detection. The first step is data generation, in which synthetic images of original Lontar manuscript images are generated with enhanced image quality. The second step is data annotation to build a new Lontar manuscript dataset. As a result, we also propose the Handwritten Balinese Character of Lontar manuscript (HBCL_DETC) dataset, a novel Balinese character in Lontar manuscripts dataset. HBCL_DETC contains 600 images that consists of more than 100,000 Balinese characters annotated by experts. Finally, the third step is training the YOLOv4 detection model using the HBCL_DETC dataset. We created this dataset specifically for the task of detecting Balinese characters in Lontar manuscripts. To evaluate the reliability of the dataset, we experimented with three scenarios. In the first scenario, the detection model was trained using original images of Lontar manuscripts, in the second scenario the detection model was trained with the addition of augmented grayscale images, and in the third scenario the detection model was trained using HBCL_DETC. Based on the experimental results, LONTAR_DETC can detect Balinese characters with high detection rate with mAP of 99.55%.

INDEX TERMS Balinese characters, dense, high variance, data generation, YOLOv4.

I. INTRODUCTION
Lontar manuscripts, a cultural heritage of Bali, are written in Balinese script and serves as a source of knowledge in Bali. In the past, Balinese people used Lontar as a guide for society. Lontar manuscripts are very sacred to Balinese people, but only certain people can understand their content. Many Lontar manuscripts are damaged due to age or improper maintenance. To avoid further damage, the provincial government of Bali have made efforts to preserve Lontar manuscripts through digitization. Transliteration is one of the processes in the digitization of Lontar manuscripts. In the transliteration process, the system is expected to translate the content of Lontar manuscripts into Latin script. The initial step of transliteration is to detect Balinese characters in images of Lontar manuscripts.

A unique characteristic of Lontar manuscripts is that the Balinese characters found within their content are very dense and no spaces occur between them. Due to this characteristic, several challenges arise in the detection of Balinese characters in Lontar manuscripts, such as the overlapping of characters, class imbalance, and the density of characters in an image. In addition, the availability of Lontar manuscripts datasets is currently limited, which makes this research very challenging.

In the context of Balinese characters in Lontar manuscripts, Sutramiani et al. [1] proposed an augmentation technique to synthesize and enhance images of Lontar manuscripts.
manuscripts and carried out recognition of Balinese characters using a convolutional neural network (CNN). However, it conducts the recognition using character images as input in which the number of recognized character classes was 18 characters classes. The proposed dataset also focuses on recognizing Balinese characters without a detection process.

Dewi et al. [2] used variations of Generative Adversarial Networks (GAN) to generate synthetic images for the purpose of object detection using data generation. The detected objects were traffic signs. The number of classes of traffic signs was four classes, namely no entry (Class P1), no stopping (Class P2), no parking (Class P3), and speed limit (Class P4). The detection method used in this study was YOLOv3 and YOLOv4. Another related research is digit detection using the YOLO algorithm, which was proposed by Kusetogullari et al. [3]. The study detects handwritten digits in historical documents.

Based on the studies above, this paper was compiled to detect dense, imbalanced, and large characters class in Lontar manuscripts images. This is the initial step of LONTAR manuscripts transliteration. In this paper, we propose LONTAR_DET, a method to detect Balinese characters in Lontar manuscripts. LONTAR_DET consists of three steps, namely data generation, Lontar manuscript annotation, and Balinese characters detection. The first step is data generation, which is carried out to synthesize images with enhanced image quality. The second step is data annotation for the purpose of producing a new Lontar manuscript dataset. Finally, the third step trains the detection model using YOLOv4. The proposed method is based on data generation following the characteristics of Balinese characters in Lontar manuscripts. Hence, it can improve detection performance.

The rest of the paper is organized as follows: Section 2 discusses related works. Section 3 describes the proposed Balinese characters detection in Lontar manuscripts. Section 4 present the experimental results and discussions. Finally, Section 5 presents the conclusion.

II. RELATED WORKS

Object detection and image recognition are often performed in image analysis. Object detection is a method for identifying objects and their positions in the image, while image recognition is classifying classes from an input image. Several papers have conducted the detection of various objects. Du et al. [4], detected moving objects, namely vehicles, blocked by complex environmental backgrounds from weak infrared camera aerial images using YOLOv4 as the detection model. Based on the experimental results, the average precision increased by 1.58%, and the F1 score increased by 0.48% which showed that the proposed method produced competitive and satisfactory results.

Dewi et al. [2] detected traffic signs using YOLOv3 and YOLOv4 to generate a detection model. In addition, this study used three variants of GAN, namely DCGAN, LSGAN, and WGAN, to increase the data and produce more realistic synthetic data variation. The experimental results revealed that the combination of the original images with the synthetic images from LSGAN resulted in the best detection performance of 84.9% on YOLOv3 and 89.33% on YOLOv4. Kumar et al. [5] proposed a novel face mask dataset due to the unavailability of appropriate datasets for face mask detection. The dataset was tested using YOLO to determine its effectiveness. Based on the experimental results, the original YOLOv4 produced the highest performance with a mAP value of 71.69%. In the medical domain, Albahli et al. [6] carried out melanoma lesion detection in dermoscopic images. The initial step of the research was removing irrelevant objects such as clinical marks and hairs by using morphological operations and sharpening the image. The next step was to detect the infected region using YOLOv4. The proposed approach achieved an average dice score of 1 and a Jaccard coefficient of 0.989. It was concluded from these results that YOLOv4 was reliable in detecting skin diseases.

Several studies have been conducted in the context of character detection and recognition. Khalil et al. [7] proposed a methodology to improve the Efficient and Accurate Scene Text Detector by adding new fully convolutional network branches for script identification. The study also proposed two e2e methods to train the model, namely multi-channel mask and multi-channel segmentation. Based on the experimental results, multi-channel segmentation outperformed existing methods with recall values of 54.34% and 81.13% for two different datasets. Yang and Hsieh [8] provided a solution to detect small and dense characters in large images and connected characters. The study adopted the CTPN or EAST character segmentation network as the main structure and proposed ResNet for feature extraction. Yang et al. [9] proposed a recognition guided detector (RGD) method that achieves tight Chinese character detection in historical documents. Santoso et al. [10] detected Kawi characters using the YOLO architecture. The experimental results of this study demonstrated that the proposed method achieved the highest accuracy compared to other methods. Research related to object detection in the cultural domain has been proposed by Darma et al. [11] on Balinese carving motif detection. The study evaluated the performance of YOLOv5 on Balinese carving images with a limited amount of data. Another study by Liu et al. [12] proposed a detection model that can distinguish coal and gangue coal based on the YOLOv4 model.

The process of image detection and recognition requires adequate training data for the training process. However, the limited amount of data is often a challenge. Several studies provide solutions by implementing data generation to overcome this problem. Pramanik and Bag [13] proposed a system to detect and correct skewness in handwritten Bangla and Devanagari words. This study used a transfer learning architecture based on CNNs. Ke et al. [14] proposed the Wasserstein GAN (ML-WGAN) data augmentation method. The proposed method can help reduce the over-fitting problem of CNN model training. Moreno-Barea et al. [15] proposed data augmentation to improve classification performance. This study used Variational Autoencoders (VAEs) and variants of GANs to generate synthetic samples. Furthermore, Qu et al. [16] proposed an augmentation method to improve the performance of in-air handwritten Chinese character recognition. This study combined global transformation with local distortion. The augmentation method
effectively enlarged the data set for training. In another related paper, Paulus et al. [17] proposed data generation to overcome data imbalance in an ancient Sundanese manuscript dataset. Synthetic data were tested using KNN classification and histogram of an oriented gradient. Based on the experimental results, the proposed data generation can improve the recognition performance up to 77%. Wu et al. [18] generated license plate images with PixTextGAN. The experimental results of license plate recognition using the ReId and CCPD datasets showed that the generation of synthetic data using the PixTextGAN can significantly improve the recognition performance. In addition to data generation, several studies have carried out image quality enhancement before the detection process. Akinbade et al. [19] used a gaussian adaptive thresholding algorithm to correctly group text characters from complex image backgrounds in the selected images. Based on the experiments, the proposed method can extract English character-based texts from images with complex backgrounds with an accuracy of 69.7% and 81.9% for word-level and character-level respectively. Sutramiani et al. [1] used data augmentation while also improving image quality of a Lontar manuscript dataset. This study proposed a multi augmentation technique for Balinese character recognition. The synthetic data was evaluated using five CNN architectures, namely VGG19, DenseNet169, InceptionResnetV2, ResNet152V2, and MobileNetV2. Based on the experimental results, the proposed method can significantly improve recognition performance of CNNs.

Several studies have been conducted on the recognition of Balinese characters. Kesiman [20] carried out word recognition for Balinese manuscripts on palm leaves. This study used the HoG feature with the NPW-Kirsch features to increase the recognition rate. Word recognition is conducted without character-by-character recognition but rather feature recognition on each segmented word image. Darma [21] carried out character recognition of Wrésastra script using Zoning and K-Nearest Neighbors (KNN). The experimental results showed that the proposed method can increase the recognition accuracy of Balinese script up to 97.5% by using a value of $K = 3$ and reference $= 10$. Darma and Sutramiani [22] segmented Balinese characters on images of Lontar manuscripts using a projection profile [22].
The experimental tests produced an accuracy value of 82.35% using a window value $= 70$ and a threshold value $= 0.05$. Sutramiani et al. [23] used transfer learning for Balinese character recognition. This study modified the number of parameters and three optimizers of the MobileNet architecture. Based on the test results, the best recognition accuracy obtained was 86.23% with the use of the SGD optimizer combined with 60% trainable parameters.

Several methods have been proposed to recognize Balinese characters based on related studies. However, the proposed recognition method only recognizes the character image; hence, it has not detected and recognized Balinese characters on Lontar manuscripts. In addition, there are many challenges in detecting Balinese characters in images of Lontar manuscripts. Balinese characters in Lontar manuscripts are handwritten, in which they are dense causing some characters to overlap, have high variance, contain noise, and character classes are imbalanced. Furthermore, the insufficient data of Lontar manuscripts hinders the detection of Balinese characters in Lontar manuscripts. In this study, we generated synthetic images with improved image quality and conducted data annotation to generate a new dataset which we call HBCL-DETC. The detection model used in this study is YOLOv4. With the use of the newly generated synthetic images to train the model, the detection performance of the model in detecting Balinese characters in Lontar manuscripts was enhanced.

### III. METHODOLOGY

This section explains the proposed method, which consists of three steps. The first step is data generation to add variety and improve image quality. The second step is data annotation to build a new Lontar manuscript dataset. Finally, the third step trains the detection model using YOLOv4. Fig. 1 shows the proposed method to detect Balinese characters in Lontar manuscripts.

### A. LONTAR MANUSCRIPTS

Lontar are ancient manuscripts made from dried palm leaves. Lontar manuscripts contain knowledge that is used as a way of life by Balinese people. Fig. 2 shows an example of a Lontar manuscript. Lontar manuscripts contain inscriptions in Balinese script written using a special knife called pengruppak. The final step of writing Lontar manuscripts is to rub the writing using roasted candle to clarify the writing.

There exists a rule in writing Balinese script called uger-uuger, in which Balinese script does not contain spaces between characters. This is a unique characteristic of Balinese scripts. Balinese script has several special script types, including wianjana script, suara script, pengangge script, gantungan, and gempelan. The characters often used in Lontar manuscripts are high-level scripts of swalalita. This script uses unique words, such as the name of God. Balinese characters on Lontar manuscripts are small and dense, making it challenging to detect individual characters.

The task of detecting Balinese characters in Lontar manuscripts is very challenging due to the overlapping of these characters, large and imbalanced classes of the characters, and the density of characters in one papyrus image. Overlapping characters always exist in every Lontar manuscript image as shown in fig. 2. Table 1 shows Balinese character classes with imbalance data. In addition, the availability of Lontar manuscripts datasets is currently limited, which makes this research very challenging.

### B. DATA GENERATION

Insufficient image data is one of the challenges in this study. In addition, Balinese characters in Lontar manuscripts have unique and challenging characteristics. Balinese characters in Lontar manuscripts are dense causing several characters to overlap each other and the classes of characters are imbalanced. Also, the images of Lontar manuscripts contain noise and have high variance.

We collected 200 Lontar manuscript images from the Lontar library in Bali to overcome this challenge. We applied the data generation method using one of our previous methods [1] and applied grayscale augmentation. Fig. 3 shows the augmentation results of two techniques, namely grayscale and adaptive gaussian thresholding. These methods produced new data variations with improved image quality. From the results of data augmentation, we generated 400 new synthetic images.

### C. LONTAR MANUSCRIPT ANNOTATION

The object detection task requires a dataset that has been labeled for ground truth. To build this dataset, we annotated all Balinese script characters in the images of Lontar manuscripts. We involved experts on Balinese character in annotating the characters as ground truth. The challenge in this process is the density of the characters in the ejection image. For example, one Lontar manuscript image has more than 100 characters. The second challenge is that the Balinese script characters are made up of 55 classes, requiring huge effort. Data annotation method using labelling.

We annotated Balinese characters on 600 Lontar manuscript images. A bounding box was used in this process to label each character. The annotation results store the coordinate information of each character in the image. The bounding box is represented by object class, object coordinates, height, and width. Fig. 4 shows the Balinese characters that has been labeled on a Lontar manuscript image.

The annotation of images of papyrus scripts produces a new Balinese script dataset for Balinese character detection tasks. To our knowledge, the dataset that we propose is a new dataset for detecting Balinese characters. The proposed dataset is very challenging because of the density of the characters, high number of classes, and imbalanced classes of characters. The new dataset of Balinese characters for the proposed detection task is called Handwritten Balinese Characters on Lontar Detection (HBCL-DETC) dataset.

### D. BALINESE CHARACTER DETECTION

The character detection step is carried out by training the model using the proposed HBCL-DETC dataset. However, this dataset is very challenging because of the unique characteristics of the Balinese characters. Therefore, an accurate
A detection model that can detect Balinese characters in a dense environment is needed.

We used YOLOv4 to build the detection model. YOLOv4 is a one-stage object detector [24]–[26]. Fig. 5 shows the YOLOv4 architecture that consists of a backbone, neck, and head. YOLOv4 uses CSPDarknet-53 as the backbone, spatial pyramid pooling (SPP) and path aggregation network (PANet) as the neck, and YOLOv3 head with anchor-based detection steps [27]. We trained the detection model by configuring the YOLOv4 architecture based on the Balinese script dataset. We use the mish activation function, channels = 3, width = 704, height = 128, momentum = 0.949, and decay = 0.0005. We use anchors = 37, 49, 65, 36, 46, 75,

---

**FIGURE 3.** Data generation result: (a) Grayscale and (b) Adaptive gaussian thresholding.

**FIGURE 4.** Annotated Balinese characters in a Lontar manuscript.

**FIGURE 5.** YOLOv4 Architecture for Balinese character detection in Lontar manuscript.
TABLE 1. Balinese character classes with imbalance data.

| ID | Character Classes | Number of Characters | ID | Character Classes | Number of Characters | ID | Character Classes | Number of Characters |
|----|-------------------|----------------------|----|-------------------|----------------------|----|-------------------|----------------------|
| 1  | Ha                | 848                  | 20 | Gantungan Na      | 139                  | 38 | Suku              | 2882                 |
| 2  | Na                | 1883                 | 21 | Gantungan Ca      | 14                   | 39 | Taleng            | 903                  |
| 3  | Ca                | 218                  | 22 | Gantungan Ra      | 429                  | 40 | Pepet             | 860                  |
| 4  | Ra                | 1108                 | 23 | Gantungan Da      | 214                  | 41 | Ceecek            | 4504                 |
| 5  | Ka                | 1267                 | 24 | Gantungan Ta      | 150                  | 42 | Surang            | 445                  |
| 6  | Da                | 575                  | 25 | Gantungan Sa      | 74                   | 43 | Bisah             | 850                  |
| 7  | Ta                | 1467                 | 26 | Gantungan Wa      | 408                  | 44 | Adeg-adeg         | 801                  |
| 8  | Sa                | 875                  | 27 | Gantungan La      | 127                  | 45 | Titik             | 419                  |
| 9  | Wa                | 1376                 | 28 | Gantungan Ma      | 105                  | 46 | A Kara            | 68                   |
| 10 | La                | 1325                 | 29 | Gantungan Ga      | 98                   | 47 | I Kara            | 47                   |
| 11 | Ma                | 1136                 | 30 | Gantungan Ba      | 85                   | 48 | U Kara            | 77                   |
| 12 | Ga                | 450                  | 31 | Gantungan Nga     | 59                   | 49 | Sa Saga           | 275                  |
| 13 | Ba                | 632                  | 32 | Gantungan Pa      | 128                  | 50 | Na Rambah         | 56                   |
| 14 | Nga               | 846                  | 33 | Gantungan Ja      | 9                    | 51 | Da Madu           | 5                    |
| 15 | Pa                | 799                  | 34 | Gantungan Ya      | 603                  | 52 | La Lenga          | 64                   |
| 16 | Ja                | 283                  | 35 | Gantungan Nya     | 1                    | 53 | Gantungan Da Madu | 45                   |
| 17 | Ya                | 928                  | 36 | Tedong            | 991                  | 54 | Gantungan Ra Repa | 212                  |
| 18 | Nya               | 87                   | 37 | Ulu               | 1722                 | 55 | Gantungan Ta Tawa | 145                  |

63, 58, 68, 85, 95, 73, 96, 105, 127, 303, 294, 137. An anchor box is a bounding box with various scales and aspect ratios centered on each pixel. The anchor box samples a large number of regions in the input image and determines whether these regions contain objects of interest and adjusts the region boundaries to predict with better accuracy the ground truth bounding box of the object.

We trained the model for 110,000 iterations. The iteration is calculated by multiplying the number of classes by 2,000. This calculation is the standard configuration of YOLOv4. The model was trained to detect 55 classes of Balinese characters. The detection process is carried out based on the bounding box of each annotated character. The density of Balinese characters in a single image is a challenge that is solved in this study.

E. DETECTION PERFORMANCE EVALUATION

We evaluated detection performance based on precision, recall, F1 score, and mean average precision (mAP). The evaluation was based on the standard Pascal VOC metrics using an intersection over union (IoU) threshold of 0.5, in which the performance was calculated by comparing the prediction results with the ground truth [28]. Precision, recall, F1 score, and mAP are defined as follows:

\[
\text{Precision} = \frac{TP}{TP + FP},
\]

\[
\text{Recall} = \frac{TP}{TP + FN},
\]

\[
\text{F1 Score} = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}},
\]

\[
\text{mAP} = \frac{\sum_{x=1}^{X} \text{AP}(x)}{X}
\]

In equation (1-2), TP is the outcome of a correctly predicted positive class, FP is the outcome of an incorrectly predicted positive class, and FN is the outcome of an incorrectly predicted negative class. Equation (3) is the weighted average of Precision and Recall. In equation (4), X is the number of queries. We measured the performance of the detection model based on these evaluation metrics.

The performance evaluation of the detection model was carried out on a computer with a 10-core 2.8 GHz processor, 12 GB RTX 3060 GPU, and 32 GB 2666 GHz DDR4 Memory specifications. The amount of time to train the model was 110 hours.

IV. EXPERIMENT RESULTS AND DISCUSSION

A. DATASET

The dataset used was the HBCL_DETC dataset composed of 600 images built through the data generation process using data augmentation. The dataset was split into 60% training data and 40% testing data. HBCL_DETC consists of 200 original images of Lontar manuscripts and...
400 synthetic images consisting of 200 images generated using adaptive gaussian thresholding and 200 images generated through grayscale data augmentation. The annotation process on these 600 images resulted in more than 144,000 labeled characters.

### B. BALINESE CHARACTER DETECTION RESULTS

We conducted experiments based on three scenarios. Table 2 shows a comparison of average precision of 55 Balinese character classes in the three scenarios. In the first scenario (S1), we trained the model on the original dataset without data augmentation. Based on the detection results of the 55 classes, the model in S1 obtained the highest average precision (AP) score in five classes in which the detection rate reached 100%. This indicates that the model in S1 can correctly detect characters based on the ground truth. However, there were two classes, namely the gantungan ja class and the da madu class, in which the model in S1 obtained the lowest AP score compared to the other two models. In the da madu class, the model in S1 only reached an AP score of 80%. This is due to class imbalance and high variation in the handwritten Balinese characters. The number of characters in the da madu class is far less than the other classes. Thus, the detection results are low.

In the second scenario (S2), we trained the model on the dataset with grayscale data augmentation. Based on the detection results of the 55 classes, the model in S2 has a
TABLE 3. Detection comparison on three experimental scenarios.

| Data Augmentation                                      | TP  | FP  | FN  | Precision | Recall | F1-score | mAP  |
|--------------------------------------------------------|-----|-----|-----|-----------|--------|----------|------|
| Original dataset (S1)                                  | 33658 | 1010  | 580  | 97%        | 98%    | 98%      | 96.53% |
| Original + Grayscale (S2)                              | 67733 | 1115  | 743  | 98%        | 99%    | 99%      | 99.10% |
| Original + Grayscale + Adaptive Gaussian Thresholding (S3) | 102283 | 721   | 431  | 99%        | 100%   | 99%      | 99.55% |

FIGURE 6. The challenging detection of the da madu class.

better detection performance compared to the model in S1. The model in S2 can detect 13 classes with an AP score of 100%. For the gantungan ja class and da madu class, the AP score of the model in S2 was 20% higher compared to AP score of the model in S1. This shows that by using data augmentation to produce synthetic images can improve detection performance in imbalanced classes.

In the third scenario (S3), we trained the model on the dataset with adaptive gaussian thresholding data augmentation. Based on the detection results of the 55 classes, the model in S3 exhibits a better detection performance than the models in the previous scenarios. The model in S3 obtained the highest AP score in 25 classes with an AP score of 100%. This shows that the model in S3 can outperform the models in the other two scenarios. Overall, the model in S3 can detect 55 classes with AP scores exceeding 95%. Improved detection performance is influenced by data augmentation that provides data variety and image quality improvement to overcome the challenges that arise in Balinese script detection.

C. PERFORMANCE EVALUATION

We evaluated detection performance of the model in the three scenarios using standard Pascal VOC metrics. This metric calculates the performance of a model using mAP at an IoU threshold of 0.5. The performance of a model is calculated by comparing the predicted results of the bounding box with the ground truth.

Based on the evaluation results, the performance of the model in S3 outperformed the models in S1 and S2, achieving a mAP score of 99.55%. Table 3 shows a comparison of the detection results of the three models in the three scenarios. Based on the experimental results, the data generation method using both the adaptive gaussian thresholding and grayscale data augmentation methods can improve detection rate by more than 3%. To elaborate, the detection rate of the da madu class increased by 20%. The da madu character rarely occurs in Lontar manuscripts. Fig. 6 shows the challenging detection of the da madu class. There are only five da madu characters in the entire Lontar manuscript which demonstrates the
challenge in detecting imbalanced classes. The rise in detection rate of the da madu class indicates that the proposed detection method can improve the detection of Balinese characters in Lontar manuscripts. Furthermore, the proposed method can overcome the problem of dense and high variance of Balinese characters. Fig. 7 shows a comparison of detection in the three scenarios. In the first scenario, the model was not able to detect the da madu character. However, in the second and third scenarios, the models were able to detect the da madu characters with a score of 0.86 and 0.88, respectively. Based on the experimental results, our proposed method can improve detection performance by applying the data generation method. These results indicate that the HBCL_DETC dataset can be used for the task of detecting Balinese scripts in Lontar manuscripts. Data generation through the augmentation method has succeeded in solving the problem of data imbalance, thereby increasing the performance of YOLOv4. In addition, the selection of anchor boxes based on the characteristics of objects in the dataset can improve detection performance.

To our knowledge, no research has been carried out to detect Balinese script characters, especially on Lontar manuscripts. Several studies have been conducted in relation to detecting Balinese script [18], [22], [27], but none have been able to detect characters of Balinese script in Lontar manuscripts. Furthermore, previous research have not detected Balinese scripts based on characters but rather based on isolated word images. Thus, the performance of previous studies only reached 44%. Our proposed method is able to obtain a better detection performance by implementing data generation to overcome challenges in detecting Balinese characters.

Our proposed method can detect Balinese characters with a performance of up to 99.55% based on the experimental results. To our knowledge, research related to the new Lontar image domain includes manual cropping and recognition based on character input images [1], [23], so it has not yet reached the Balinese character detection in Lontar manuscripts. Therefore, the evaluation matches the detection results with ground truth images that experts have validated.

V. CONCLUSION

In this paper, we proposed LONTAR_DETC, a method to detect Balinese characters in Lontar manuscripts. LONTAR_DETC consists of three steps, namely data generation, Lontar manuscript annotation, and Balinese character detection. We applied two augmentation techniques in the data generation step. In this study, we also proposed a new dataset named HBCL_DETC for the purpose of Balinese script detection. This dataset is a combination of original Lontar manuscript images and synthetic images generated from two data augmentation techniques. Based on the experimental results in three scenarios using YOLOv4 as the detection model, our proposed method can increase the detection performance of characters that rarely occurs and overcome the problem of class imbalance. Data generation through the augmentation method has succeeded in solving the problem of data imbalance, thereby increasing the performance of YOLOv4. In addition, the selection of anchor boxes based on the characteristics of objects in the dataset can improve detection performance. Furthermore, by evaluating detection performance using standard Pascal VOC metrics, the experimental results show that the proposed method a mAP score of up to 99.55% in detecting Balinese characters. These results indicate that our proposed method can detect Balinese scripts in Lontar manuscripts with high performance.

In future works, we will develop this research by transliterating the Balinese characters. This research is significant in an effort to conserve cultural heritage.
