DOCUMENT VALIDATION FOR COOPERATION AGREEMENT DOCUMENTS AT THE UNDIKSHA COOPERATION AND PUBLIC RELATIONS AGENCY (BADAN KERJA SAMA DAN KEHUMASAN) USING LOCAL BINARY PATTERN (LBP) AND YOLOV5 METHODS

Komang Jepri Kusuma Jaya[1], Made Windu Antara Kesiman[2],
I Made Dendi Maysanjaya[3]

[1,2,3]Virtual Vision, Image, and Pattern Research Group, Faculty of Engineering and Vocational, Universitas Pendidikan Ganesha

email: jepri@undiksha.ac.id[1], antara.kesiman@undiksha.ac.id[2], dendi.ms@undiksha.ac.id[3]

Abstract

A cooperation agreement document managed by BKK Undiksha is a conventionally managed document. With the implementation of the Kampus Merdeka - Merdeka Belajar curriculum in 2021, the number of incoming cooperation agreement documents has increased rapidly, making the collection of document data much longer and inefficient. Therefore, there is a need for a scheme that can automatically validate collaboration documents. The document validation scheme was developed using the Local Binary Patterns and YOLOv5 methods. The data used in the development is primary data from BKK Undiksha in 2021, where two agencies collaborate. The development result is the document validation model collaboration using the Local Binary Pattern and YOLOv5 methods with the best accuracy results of 95.36%, and the ability of the model to detect the number of seal, stamp and signature components is 90.73%.

Keywords : document validation, cooperation agreement document, BKK Undiksha, Local Binary Pattern, YOLOv5

INTRODUCTION

Digital documents are the result of technological advances that are increasingly being used. Digital documents are documents that are stored and published digitally. The output of digital documents can be printed or digital images. Based on the initial creation process, digital documents can be divided into two groups, namely the group of digitally created documents (born digital documents) and the group of physical/printed documents that are digitised (non-born digital documents).

Posters are included in the born digital document because posters are usually designed digitally from scratch and are managed, stored and published digitally. Not only posters, but also textual documents that are created, managed and published digitally can be called born digital documents. The advantage of a born digital document is that the metadata of the document still exists, so extracting a born digital document can be done simply by reading its metadata. Meanwhile, documents that are digitally created but printed are included in non-born digital

documents, such as proclamation text documents or other scanned results. Of course, because they are derived from physical documents, non-born digital documents do not have metadata. Therefore, digital image processing is required to extract non-born digital documents.

In Indonesian universities, cooperation agreement documents are one of the digital documents that are managed. At Ganesha University of Education, cooperation agreement documents are managed by the Undiksha Cooperation and Public Relations Agency (BKK). Based on the results of the initial observations and interviews, the cooperation agreement documents managed by BKK Undiksha are digital documents that include non-born digital documents. The cooperation agreement documents require conventional signatures, stamps and seals, in other words, the cooperation agreement documents prepared by typing are first printed to be signed, stamped and sealed. The document is then scanned and

stored digitally and submitted online to Undiksha's BKK management.

Based on the results of an interview with Mr Made Windu Antara Kesiman, S.T., M.Sc., Ph.D. as the head of BKK Undiksha, it was found that the data collection on the cooperation agreement documents was done manually. When a new cooperation agreement document arrives, the BKK Undiksha management has to check the completeness of the document, which includes a signature, agency stamp and seal. If the new document is complete, it is considered valid and the administrator enters the document data into Excel. If not, the document is returned to the sending agency. Figure 1 shows the number of cooperation documents managed by BKK Undiksha between 2017 and 2021 (21 December 2021)[1]
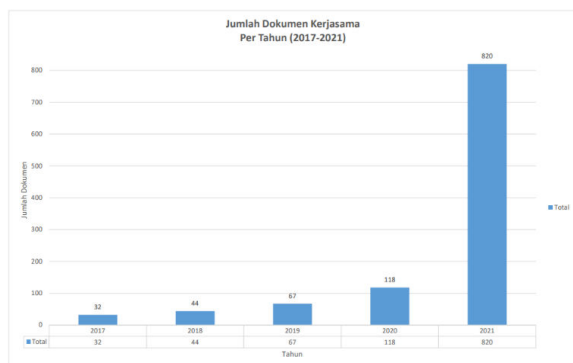

Figure 1. Graph of The Number of Cooperation Documents per Year

Referring to the Figure 1, the number of incoming cooperation agreement documents has increased rapidly in line with the implementation of the Kampus Merdeka – Merdeka Belajar curriculum in 2021. Naturally, this increase makes the document data collection process much more time consuming and inefficient, so it needs to be optimized in the form of creating a digital scheme to automatically check the validity of BKK Undiksha cooperation documents. The scheme must be able to determine whether the cooperation agreement document is valid or not by detecting the number of signatures, seals and stamps. The biggest challenge in making this scheme is that the scanned document has noise and the possibility of a stamp on the cooperation document which is not clearly visible.

Noise is anything that interferes with the image [2]. In some cases, unwanted brightness or signal variations may occurs [3]. There are many types of noise, one of which is Gaussian noise, which is a statistical noise that has a Probability Density Function (PDF) equivalent to the normal distribution, or often called the Gaussian distribution [4].

In a study titled *Implementasi Local Binary Pattern untuk Deteksi Keaslian Mata Uang Rupiah*, it was found that Local Binary Patterns (LBP) can classify genuine or counterfeit *rupiah currency* with up to 95% accuracy using the KNN method [5]. In this study, LBP plays a role in extracting features from the input image in the form of an image of rupiah banknotes with some images containing noise. Then, in the analysis of document images using local binary patterns [6], LBP is used to remove the noise in the input image. As a result, LBP is able to classify pixel noise with interior pixels in images with a noise level of up to 50%.

On the other hand, one of the machine learning architectures for object detection is You Only Look Once (YOLO). The YOLO architecture is widely used because of its high performance and short learning time. In the case of signature and logo detection in a study titled Signature and Logo Detection using Deep CNN for Document Image Retrieval [7], it was found that the YOLOv2 architecture was able to achieve a Mean Average Precision (mAP) of 0.778 or 77.8% over time. process per image 0.130 seconds. To date, the YOLO architecture continues to evolve and improve with its latest architecture, YOLOv5. A comparison between YOLOv3, YOLOv4 and YOLOv5 can be seen in a study by Nepal [8] with a case study of Autonomous Landing Spot Detection in Faulty UAVs which produced a mAP on YOLOv3 of 0.46, YOLOv4 of 0.607 and YOLOv5 of 0.0633.

Based on the problems described and related studies, this research will focus on developing a scheme in the form of a digital image model that is capable of checking the number of signatures, stamps and seals that are then used as the validity of the collaborative document. The model developed uses the Local Binary Pattern (LBP) as the feature extraction stage and the YOLOv5 architecture to detect signatures, stamps and seals.

METHOD
The development of the BKK Undiksha document validation system using the AI Project Cycle method [9]. The development flow is shown in the flowchart in Figure 2.
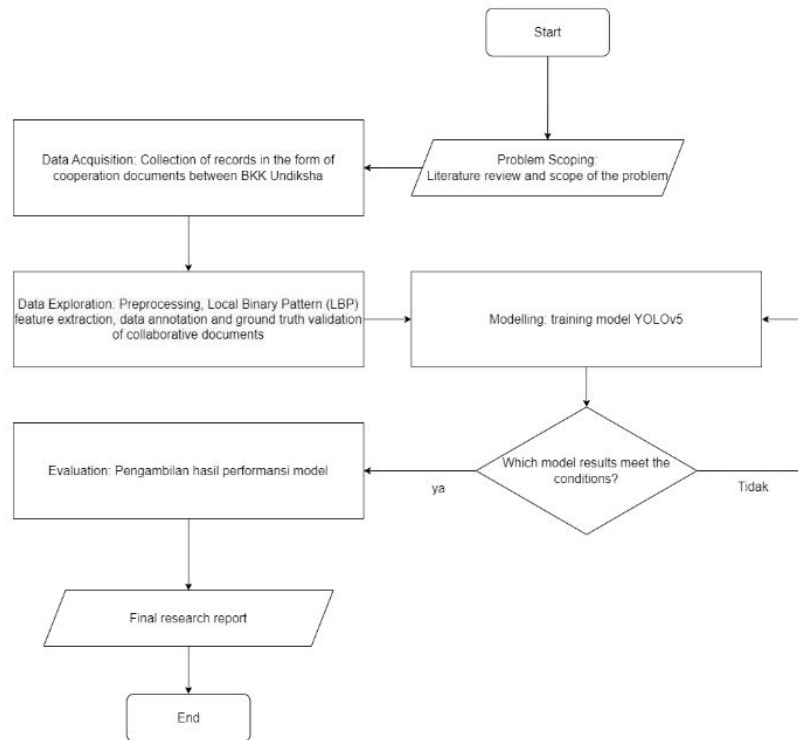
Figure 2. AI Project Cycle Flowchart

With reference to the work flow chart in Figure 2, the flow of project work in this study can be described as follows.

Problem Scoping

Problem scoping is done to identify the problems that need to be solved using AI [9]. In problem scoping, the detailing is done in 4W, namely what, who, why and where. Table 1 is a problem scoping in this study.

Table 1. Problem Scoping

4w	Description
What	The BKK Undiksha cooperation document validation process is slow.
Who	The Undiksha Cooperation and Public Relations Agency (BKK)
Where	Ganesha University of Education (Undiksha)
Why	Because the process of validating collaboration documents is still done in the traditional way, with management reviewing each document and recording it in an Excel file.

Data Acquisition

The dataset used in this study is primary data, that is cooperation agreement documents obtained from BKK Undiksha with a total of 857 data with PDF extensions. The data is then subjected to data cleansing to identify the documents used as training data and test data. Below are the data that were deleted in the data cleansing stage.

- The collaboration document has more than two collaborators
- Collaboration documents that are truncated (especially in the signature section)

The total number of cooperation documents obtained after data cleansing was 619. To perform model learning, 75% (468 documents) of the total data is used as training data. Meanwhile, 25% (151 documents) are used as test data. The characteristics of the data used are shown below.

- Collaborative documents are scanned documents or unborn digital documents.
- The collaborative document has two parties working together

- There are no truncated collaborative documents (especially in the signature section)
- There are documents scanned using a smartphone, resulting in noise or distorted scan results.
- There are documents that have poor contrast, especially on the stamp.

Data Exploration

The input document is converted into a set of images. Then the process of data pre-processing and image normalisation is carried out, which is the initial processing carried out to maximise the model training process. After that, the feature extraction process is performed using the Local Binary Pattern (LBP) method. The representation of the LBP results is carried out by the annotation process and the creation of ground truth. Each process is described in more detail as follows.

A. Preprocessing

There are several parts to the pre-processing stage, including the input image is greyscale, sharpening, gaussian blur and the last is image normalisation. In more detail, each stage includes:

- Grayscale is the process of converting the input RGB image into a grey scale or grey image.
- Sharpening is a process of sharpening the line structure of the greyscale image by adding the edge image to the original image.
- Gaussian blur is a process to reduce noise in the image.
- Image Normalisation is the process of equalising the size of all the images to 640 x 640. Image normalisation is very important because the YOLOv5 model uses the same input image size.

B. LBP Extraction

Local Binary Pattern (LBP) is a method that can be used to extract texture features from greyscale images. LBP was first introduced in 1996 by Timo Ojala. The LBP method is a texture descriptor that uses a comparison of the grey values of neighbouring pixels [10]. The image is extracted from the pre-processing results by extracting the image features [11]. The feature extraction process at this stage uses the LBP method. The LBP method is an extraction method based on binary values in the image that searches for each centre point with a size of 3x3 neighbouring pixels, then calculates each neighbouring pixel with the centre to obtain the pixel value of each window. This process produces binary values which are converted to LBP values as a result of the feature extraction process.

C. Data Annotation and Ground Truthing Modeling

After identifying the data set from the corpus of BKK Undiksha, next step is labelling or annotation. The model learning process requires data annotations or ground truth as a learning medium. Figure 3 shows an example of the data annotations used. The annotations used are in the form of placing labels in the signature, stamp and seal areas using the labelimg application. The label is then stored in a TXT format file.
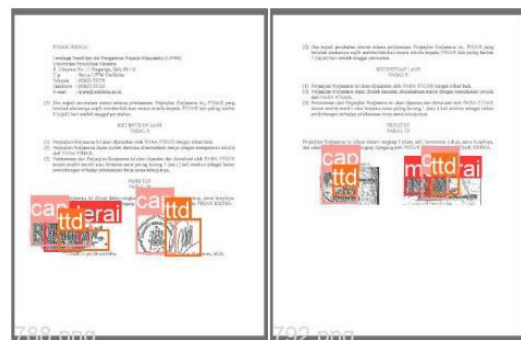

Figure 3. Data Anotations

The annotations from the table are then grouped based on the location of the components (left or right area), which is used to determine the ground truth ROI of the data. On the other hand, in order to determine the performance in checking the validation of cooperation documents, a ground truth component is needed in the form of the number of signature, stamp and seal components by counting the number of classes in the annotation. As in the example in Table 2, the ground truth is in Excel format with the contents of the columns being the number of signatures, seals and stamps on the document.

Table 2. Ground Truth Extraction ff The Number of The Component

File Id	Seals	Stamps	Signatures
1	1	2	2
2	0	1	2
3	1	1	2

Modeling

You Only Look Once (YOLO), proposed in 2016, is an algorithm that takes a single image as input without dividing the image into regions [8]. The YOLO algorithm uses the idea of regression in image recognition, resulting in a

simple and fast process [12,13]. The use of a single stage neural network to complete the classification and localisation of detected objects results in high speed. YOLO can view the entire image during training and search for objects in the global information contained in the object detection case [14]. YOLOv5 is the fifth generation of the YOLO object recognition algorithm, to be released in April 2020. YOLOv5 is written in the Python programming language. The architecture used in YOLOv5 is generally not significantly different from the previous generation of YOLO [8].

The model and the source code for the training system are taken from the official YOLOv5 website via the github url https://github.com/ultralytics/yolov5. The modelling process is divided into two parts, the ROI model and the component model. The ROI model aims to find ROI areas and the component model aims to detect seals, stamps and signatures in each ROI. When training the ROI model and the component models, several configurations are used, as shown in Table 3.

Table 3. Model Configuration

Parameter	Value
Image size	640
Batch Size	16
Epoch	200
Data source	data-document.yaml
Weight	YOLOv5x.pt

Evaluation

Testing is performed by inputting test data, namely collaboration document files with PDF extensions. The scheme first performs pre-processing and feature extraction using LBP. Stamp, signature and seal detection is then performed on the model created. The result is the number of stamps, signatures and seals on each document stored in the EXCEL file. The results obtained are then compared with the document validation ground truth to obtain the performance of the model.

RESULT AND DISCUSSION
Results of The Schematic Creation
A.   Preprocessing and Data Normalitation

Documents obtained from PDF to Image extraction have different sizes, so pre-processing and data normalisation are required. The pre-processing stage consists of several processes, including greyscale, sharpening and gaussian blurring. An example of the final result of this step is shown in Figure 4.

Figure 4. Example of Preprocessing and Normalitation Result

B.   Local Binary Pattern Exctraction

The images that have passed through the pre-processing and data normalisation stages are then subjected to feature extraction using the Local Binary Pattern (LBP) method. This study uses a kernel or window size of 3 x 3. The window moves through the coordinates [0, 0] to the last coordinate. Each window is calculated for the mean value of the window. The results of each LBP calculation are rearranged into a new image. Figure 5 shows an example of an image resulting from LBP.
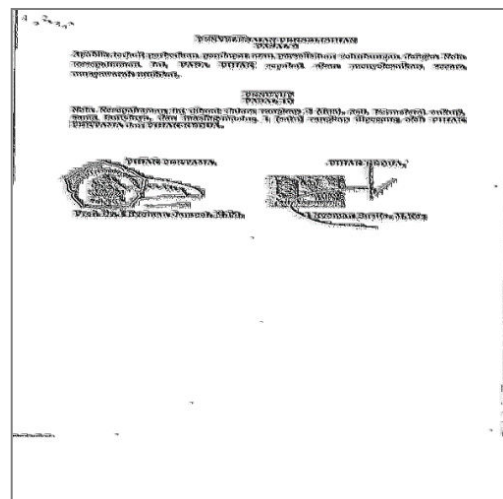

Figure 5. Data Anotations

C.   Region of Interest Extraction

The Region of Interest (ROI) is an area suspected of containing signatures, stamps and seals for Party 1 and Party 2. By first determining the ROI before searching for signature, stamp and seal components, the possibility of misdetecting the model, such as

being detected as a logo stamp, is reduced. The results of the ROI are used later for the signature, stamp and seal component detection areas. The ROI detection was performed using the YOLO model.

The input data used for this training model uses LBP image results from the train data. The training data is divided into 80% (374) as data training and 20% (94) as data validation. The batch size is the number of data samples used in an iteration. The model was trained for 200 iterations, resulting in the model performance shown in Figure 6.
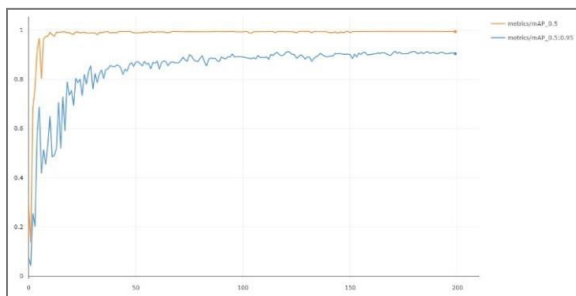

Figure 6. Graph of the mAP ROI extraction model

Model performance is obtained by comparing validation data during the training process. The graph in Figure 6 shows an increase in the Mean Average Precision (mAP) during the training process. The orange line shows the mAP value with an IOU threshold of 0.50, where the results for all classes are 0.995. Meanwhile, the blue line is the mAP with an IOU threshold of 0.50 - 0.95, with the results for all classes being 0.914. The precision of all classes is 0.989 and the recall is 0.989. From the mAP, precision and recall values it can be seen that the model is performing well. When the model is run, it will generate coordinate values for the ROI area along with confidence values. When these coordinates are described, it will produce a bounding box. In addition, ROI extraction or cropping is performed based on the ROI coordinates. ROI extraction is performed on pre-processed images to obtain new images that still contain detail. ROI extracted images are later used for component detection. The Figure 7 is an example of an ROI extracted image.
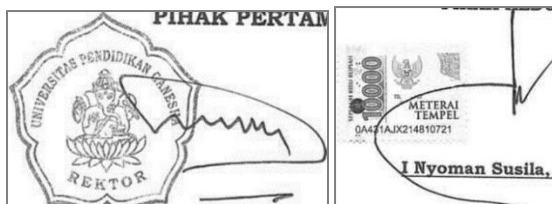

Figure 7. ROI Extraction Final Results

D. Components Extraction

Component extraction aims to obtain the signatures, stamps and marks contained in each ROI section. Therefore, the component extraction stage was performed using the detection model with YOLO. The training of the component extraction model uses input data in the form of extracted images from the ROI stage (as shown in Figure 7) in the training data. There are 908 images extracted from the ROI train, divided into 80% (727) training data and 20% (181) validation data. The Figure 8 is the performance obtained from the component model training process.
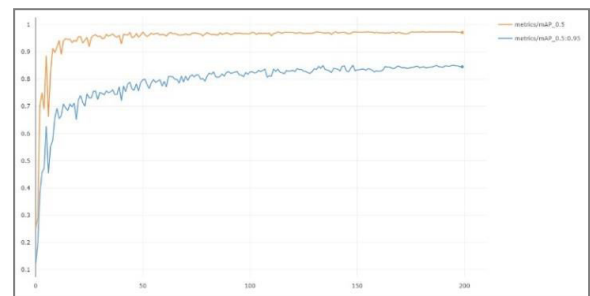

Figure 8. Graph of the mAP Component Extraction Model

Referring to Figure 8, mAP with an IOU threshold of 0.50 is higher than mAP with an IOU threshold of 0.50 - 0.95. The value of mAP 50 for all classes is 0.974 and mAP 0.50 - 0.95 for all classes is 0.835. The precision and recall for all classes in Figure 8 are 0.965 and 0.959 respectively. The lowest precision, recall and mAP were found in the signature class because signatures have inconsistent shapes and are more likely to be in the centre of the stamp or seal. Nevertheless, the model still has a good performance.

The model generates coordinates that are similar to ROI coordinates. The coordinates of the components resulting from the model are relative coordinates that can adapt to the height and weight of the image. Three class codes are defined in the model. 0 is the class code for the seal, 1 is the class code for the stamp and 2 is the class code for the signature.

The component model will work in the ROI area so that there will generally be two stages of component detection in a drawing, includes on Party 1 and Party 2. The two are then combined in one drawing to determine the final coordinates of the stamp, stamp and signature components along with their numbers. Figure 9 shows an example of the final coordinates of the combined component models.
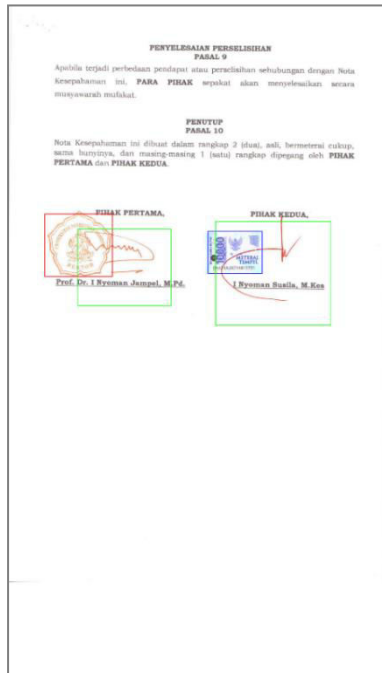
Figure 9. Example of component extraction
Coordinate end result

Result of Evaluation

A. Evaluation of ROI Coordinates

The evaluation was carried out using test data from 151 data points with the ROI extraction model. Coordinates are considered valid if they have an Intersection Over Union (IOU) value greater than 0.50. Thus, the confusion matrix can be defined as follows [15].

- True Positive (TP) is the correct detection. Detection with an IUO value > 0.50
- False Positive (FP) is an incorrect detection. Detection with IUO value ≤ 0.50
- False Negative (FN) is a ground truth that is not detected.
- True Negative (TN) includes all detection results that do not depend on the ground truth, then TN is not used because there are many possible predictions that should not be detected.

Based on these rules, each class is searched for the confusion matrix and stored in a table containing information on the document ID, class ID, confusion matrix and confidence of each predicted document. Next, the results are sorted based on the highest confidence value and the number of TP and FP are calculated from the previous data. The results of the number of TP and FP are used to calculate the precision and recall of each id. Finally, a simplification is performed using interpolated precision with finding the precision value for

recall 0.0 to 1.0, where the precision value used is the largest precision value to the right of the recall. Figure 10 is a plot of the interpolated precision results.
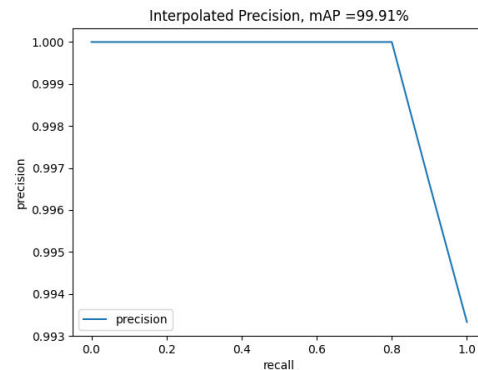

Figure 10. Interpolated Precision Result for ROI Model

Each precision value on the precision interpolated graph is summed and divided by the number of recalls used to produce the Average Precision (AP) for each class. The ROI has only one class, so the AP is the mAP of the ROI model. Based on Figure 10, the AP and mAP obtained are 99.91%. Therefore, so the ROI model have a good performance.

B. Evaluation of Component Coordinates

The evaluation of the component coordinates was carried out using test data from 151 previously extracted ROIs. The evaluation stage is the same as the ROI coordinate evaluation stage, i.e. first, if the IOU value is greater than 0.5, then the coordinates are declared valid. Next, the confusion matrix is determined for each document and the obtained precision and recall values are calculated. Then, to obtain the AP value for each class, the graph is simplified with interpolated precision. Figure 11 is the result of the interpolated precision graph of the component extraction model.
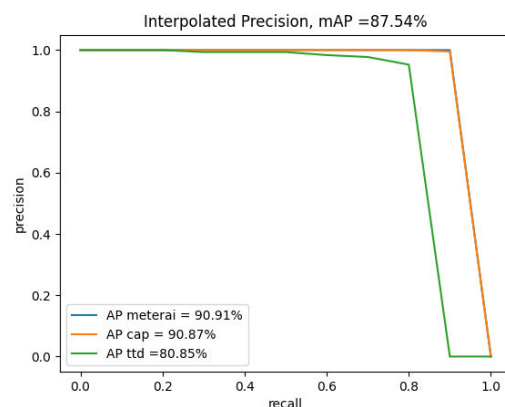
Figure 11. Interpolated Precision Result for Components Model

Based on the graph in Figure 11, the AP obtained from signatures, stamps and seals are 90.91%, 90.87% and 80.85% respectively. The average of the three APs or mAP obtained is 87.54%, so the component extraction have a good performance.

C. Evaluation of Total Components

The number of components evaluation is an evaluation that aims to check the performance of the model in detecting the number of components without taking into account the coordinates of these components. The evaluation is performed using the ground truth components. The prediction is declared correct if the number of components in the predicted result matches the number of components in the ground truth. The components evaluated are the number of postage components, stamps, signatures and combinations of these three components. The data used for the evaluation are test data of 151 data. Figure 12 is a graph of the evaluation results for the number of components.
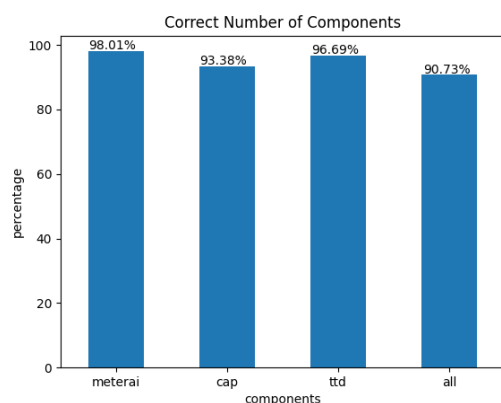

Figure 12. Graph of Number of Components

Based on the graph in Figure 12, the postage stamp component is the component with the closest number of predictions to the ground truth with an average number of 98.01%. The stamp component has 93.38% correct recognition and 96.69% correct signature. Meanwhile, the data for which the three components are correct (correct seal, correct stamp and correct signature) is 90.73%.

D. Evaluation of Component Validition Discussion

The document validity evaluation aims to check the performance of the model in classifying documents as valid or invalid. The provisions of documents that can be declared valid are defined as follows.

- A valid domestic cooperation document is a cooperation document that has at least 1 seal, 2 stamps and 2 signatures.
- A valid foreign cooperation document is a cooperation document with at least 2 stamps and 2 signatures.

Based on this valid definition, a ground truth is created that describes the type of document (domestic or foreign cooperation document) and the validity of the document (the document is considered valid or invalid). The data used are test data of 151 data, with the ground truth results being 3 foreign documents and 149 domestic documents, with 89 documents being valid documents and 62 documents being invalid documents.

The next step is to compare the predicted data with the ground truth. This produces the confusion matrix shown in Figure 13.
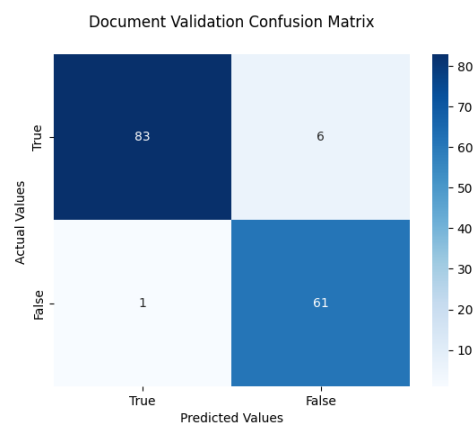

Figure 13. Confusion Matrix for Document Validation

Referring to the Confusion Matrix in Figure 13, 83 valid cooperation documents are detected as valid and 6 valid cooperation documents are detected as invalid. Meanwhile, only 1 invalid document was successfully detected as invalid. Based on the Confusion Matrix, calculations are then performed to find the values of Precision, Recall and Accuracy. The graph in Figure 14 is the result obtained.

The graph in Figure 14 shows that the model achieved a precision value of 98.81%, a recall value of 93.26% and an accuracy value of 95.36%. Therefore, it can be concluded that the model has a qualified performance in classifying the validation of work documents.
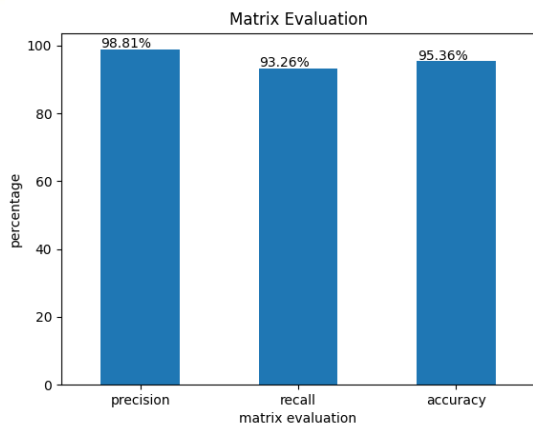
Figure 14. Graph of Precision, Recall and Accuracy of Document Validation

Discussion

The validation model for the Undiksha BKK collaboration document was successfully developed using YOLOv5, which was supported by LBP extraction for ROI detection. The model stages are divided into 4 stages, which are divided into pre-processing and data normalisation, LBP extraction, ROI extraction and component extraction. ROI extraction is a very crucial stage, if the model detects ROI incorrectly, component extraction will not run optimally. Without ROI extraction, there is a possibility that the component model will detect the wrong component. As shown in Figure 15, the logo on the document is detected as a postage stamp because it uses a shape similar to a postage stamp. Without the use of ROI, the results of the component count evaluation will be poor with a percentage of only 49.68%. Therefore, before a component is detected, the correct ROI must first be determined.
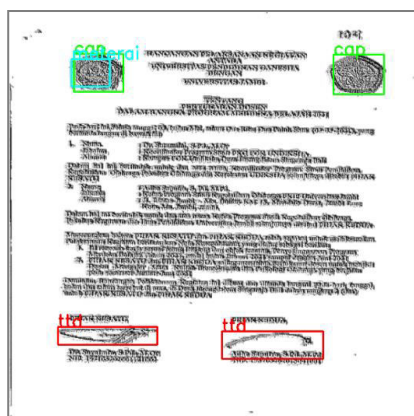

Figure 15. Example Without ROI Model

A qualified ROI model is required to minimise errors such as those shown in Figure 15. In this study, a ROI model was developed using YOLOv5x with input data in the form of LBP extracted images. The use of LBP helps the YOLO model to classify ROI areas well. However, when the input data is in the form of images without LBP extraction, the performance of the model decreases, so that there are several documents with undetected ROI. This is described in more detail in Table 4, which compares the evaluation results of the ROI model using input data in the form of images with and without LBP.

Table 4. Comparison of ROI Models With/Without LBP

No	Input Image	Model	Classes	Epoch	mAP
1	With No LBP	YOLOv5s	2	200	53.83%
2	With LBP	YOLOv5s	2	200	90.81%
3	With LBP	YOLOv5s	1	200	93.88%
4	With LBP	YOLOv5x	1	200	99.91%

The results of the comparative evaluation of the ROI model in Table 4 describe that the input data with LBP has a significant performance in detecting ROI with mAP, which is more than 90%, while the results with input data without LBP are only 53.83%. This significant result is due to the ability of LBP to group components in the ROI region so that the ROI region has a different texture from other areas. On the other hand, without LBP extraction, the model has problems in determining the ROI, especially for documents with only a signature and a stamp on the document. Therefore, it is concluded that LBP extraction can optimise the YOLO model in detecting ROI. The ROI model with 2 classes (Party 1 class and Party 2 class) has a lower performance than the model using 1 class (ROI class). In general, the ROI area for party 1 and party 2 is the same (has at least one metari, stamp or signature), by dividing into 2 classes causes several classes to be detected upside down (party 2 is detected as party 1), otherwise class simplification causes prediction results to become more optimal and training data to become more. Therefore, the model with 1 class has the best performance, namely 99.91%.

Different than the ROI model, the component model uses input data in the form of image processing results without LBP extraction. The resulting image is a greyscale image, i.e. it ignores the original colour of the stamp, postmark and signature components. Although some components have different

colours, some documents have almost the same stamp and signature colours, and some documents are also greyscale to begin with. Therefore, the colour feature is not effective in helping the model to recognise documents. On the other hand, the stamp, postmark and signature components are usually stacked on top of each other, so when LBP is extracted, the texture on the components is lost and the model has difficulty in recognising the components. A comparison of the results of the component model with image input data from LBP and without LBP is shown in Table 5.

Table 5. Comparison of ROI Models With/Without LBP

| No | Input Image | Model | Evaluation Results (%) | | | |
			AP Seal	AP Stamp	AP TTD	mAP
1	With LBP	YOLOv5x	81.16	81.45	55.92	72.84
2	Without LBP	YOLOv5x	90.91	90.87	80.85	87.54

LBP feature extraction is less able to describe the texture of stacked components. Comparing the evaluation results in Table 5, the model with input data in the form of LBP images has a poor performance, namely 72.84% mAP compared to input data in the form of images without LBP with 87.54% mAP. The signature is the component that has the lowest AP among the other components, caused by the position of the signature stacked between the stamp and the seal. The use of LBP results in the loss of some of the signature textures, so that the AP obtained is only 55.92%. Based on the evaluation results, the pre-processed input image without LBP extraction was used for component extraction.

Stacked components make component extraction more complex than ROI extraction. LBP is not optimal in helping the model detect stacked components, so an extension of the YOLO model is needed to maximise the results of component detection. YOLOv5s is the YOLO version with the simplest architecture compared to YOLOv5x, which results in YOLOv5x having higher performance in detecting more complex images. Therefore, the YOLOv5x weight model is very suitable to be used in the detection of each component of the seal, stamp and signature. The performance of YOLOv5s and YOLOv5x is shown in Table 6.

Table 6. Comparison of ROI Models With/Without LBP

| No | Model | Percentage of Number of Components | | | |
		Seal	Stamp	Sign	Combine
1	YOLOv5s	94.7%	93.38%	86.09%	78.15%
2	YOLOv5x	96.69%	92.05%	92.72%	86.75%

Table 6 shows the performance of the YOLOv5s and YOLOv5x models in detecting the number of seal, stamp and signature components. The two models have almost the same performance in detecting seals and stamps, but have different performance in detecting signatures. The position of the signature makes it the most difficult component to detect. The YOLOv5x model successfully detects signatures with a difference of about 12% and when combined (the sum of all detected components is correct) it reaches 90.73%.

Based on the comparison of the evaluation results described, LBP extraction has the ability to group textures, but is less than optimal in separating textures from overlapping components. Therefore, LBP can be used maximally to help determine ROI and less optimally in component extraction models.

Meanwhile, the YOLOv5x model, which has a more complex architecture, is used to optimise component extraction. The end result is that the LBP and YOLOv5 methods can detect the number of components by 90.73% and classify valid and invalid documents with an accuracy of 95.36%.

CONCLUSION

The document validation model using the local binary pattern and YOLOv5 methods has an excellent performance with a mAP ROI of 90.91%, a component mAP of 87.54%, component number detection results of 86.75% and a document validity classification accuracy of 94.01%. The definition of a valid document in this study depends only on the number of stamp, seal and signature components without considering the location of these components. Therefore, future research can also consider the location of each component to broaden the definition of a valid document.

REFERENCES

[1] BKK Undiksha, *Data dan Statistik Dokumen Kerja Sama*. 2021. [Online]. Available: https://www.bertelsmann-stiftung.de/fileadmin/files/BSt/Publikationen/GrauePublikationen/MT_Globalization_Report_2018.pdf%0Ahttp://eprints.lse.ac.uk/43447/1/India_globalisation%2Csociety and inequalities%28Isero%29.pdf%0Ahttps://www.quora.com/What-is-the

[2] A. Prasetio and P. M. Hasugian, "Improving the Quality of Digital Images Using the Median Filter Technique to Reduce Noise," *SinkrOn*, vol. 4, no. 1, p. 143, Oct. 2019, doi: 10.33395/sinkron.v4i1.10155.

[3] O. Joshua, J. S. Owotogbe, T. S. Ibiyemi, and B. A. Adu, "A Comprehensive Review On Various Types of Noise in Image Processing," *Article in International Journal of Scientific and Engineering Research*, 2019, [Online]. Available: http://www.ijser.org

[4] R. K. Riad Al-Taie, B. Jumaa Saleh, and H. A. Abu-Alsaad, "A REVIEW PAPER: DIGITAL IMAGE FILTERING PROCESSING." [Online]. Available: www.techniumscience.com

[5] M. Miladiah, R. Umar, and I. Riadi, "Implementasi Local Binary Pattern untuk Deteksi Keaslian Mata Uang Rupiah," *Jurnal Edukasi dan Penelitian Informatika (JEPIN)*, vol. 5, no. 2, p. 197, 2019, doi: 10.26418/jp.v5i2.32721.

[6] N. Anusha and M. Siva Sindhuri, "Analysis of Document Images through Local Binary Patterns," *Indian J Sci Technol*, vol. 10, no. 20, pp. 1–6, 2017, doi: 10.17485/ijst/2017/v10i20/73915.

[7] N. Sharma, R. Mandal, R. Sharma, U. Pal, and M. Blumenstein, "Signature and logo detection using deep CNN for document image retrieval," *Proceedings of International Conference on Frontiers in Handwriting Recognition, ICFHR*, vol. 2018-Augus, pp. 416–422, 2018, doi: 10.1109/ICFHR-2018.2018.00079.

[8] U. Nepal and H. Eslamiat, "Comparing YOLOv3, YOLOv4 and YOLOv5 for Autonomous Landing Spot Detection in Faulty UAVs," *Sensors*, vol. 22, no. 2, 2022, doi: 10.3390/s22020464.

[9] CBSE Academic, *AI Curriculum Handbook*. 2019.

[10] I. P. N. Syair, Ayu Sabrina., Muh Ihsan., "Identifikasi citra tanda tangan menggunakan Local Binary Pattern (LBP) pada Smartphone berbasis Android," *semanTIK*, vol. 5, no. 1, pp. 17–24, 2019.

[11] F. Cruz, N. Sidere, M. Coustaty, V. P. D'Andecy, and J. M. Ogier, "Local Binary Patterns for Document Forgery Detection," in *Proceedings of the International Conference on Document Analysis and Recognition, ICDAR*, IEEE Computer Society, Jan. 2018, pp. 1223–1228. doi: 10.1109/ICDAR.2017.202.

[12] Š. Hubálovský, P. Trojovský, N. Bacanin, and V. K, "Evaluation of Deepfake Detection Using YOLO with Local Binary Pattern Histogram," *PeerJ Comput Sci*, vol. 8, p. e1086, Sep. 2022, doi: 10.7717/peerj-cs.1086.

[13] J. Yao, J. Qi, J. Zhang, H. Shao, J. Yang, and X. Li, "A real-time detection algorithm for kiwifruit defects based on yolov5," *Electronics (Switzerland)*, vol. 10, no. 14, 2021, doi: 10.3390/electronics10141711.

[14] Jupiyandi Saniputra, F. R. Pratama, and Yoga Dharmawan, "Pengembangan Deteksi Citra Mobil Untuk Mengetahui Jumlah Tempat Parkir Menggunakan Cuda Dan Modified Yolo," *Jurnal Teknologi Informasi dan Ilmu Komputer (JTIIK)*, vol. 6, no. 4, pp. 413–419, 2019, doi: 10.25126/jtiik.201961275.

[15] R. Padilla, S. L. Netto, and E. A. B. Da Silva, "A Survey on Performance Metrics for Object-Detection Algorithms," *International Conference on Systems, Signals, and Image Processing*, vol. 2020-July, pp. 237–242, 2020, doi: 10.1109/IWSSIP48289.2020.9145130.