Building YoloV4 models for identification of rock minerals in thin section

by Muhammad Fatih Qodri

Building YoloV4 models for identification of rock minerals in thin section

To cite this article: B G Pratama *et al* 2023 *IOP Conf. Ser.: Earth Environ. Sci.* **1151** 012046

View the article online for updates and enhancements.

IOP Conf. Series: Earth and Environmental Science 1151 (2023) 012046 doi:10.1088/1755-1315/1151/1/012046

Building YoloV4 models for identification of rock minerals in thin section

B G Pratama[1], M F Qodri[2]* and O Sugarbo[2]

[1]Electrical Engineering Department, Faculty of Industrial Technology, Institut Teknologi Nasional Yogyakarta, Indonesia
[2]Geological Engineering Department, Faculty of Mineral Technology, Institut Teknologi Nasional Yogyakarta, Indonesia

*fatihqodri@itny.ac.id

Abstract. Rock mineral identification is a costly and time-consuming task using conventional methods of testing physical and chemical properties, especially in the petrographic laboratory. A comprehensive identification model for three rock minerals in sedimentary rocks based on the YoloV4 model is available as a solution. The models predict rock minerals by calculating the pixels and the weights that have been trained previously. First, the YoloV4 models and framework were built. Then, a total of 44 manually labelled thin section images (sedimentary rock thin section) were used to create the model to detect minerals accurately. The MAP and loss results showed that the parameters of the minerals detection model in PPL are 11% and 1.19, respectively. Meanwhile, The MAP and loss results of XPL are 19% and 1.18, respectively. Finally, Identification of rock minerals using deep learning algorithms is a very promising idea especially the YoloV4 model can build a comprehensive detection of rock samples in thin sections effectively.

1. Introduction

Conventional detection methods have clear physical implications. Generally, detection is performed based on the physical property's observation for identifying rocks mineral in thin sections. [1]. Identification models have been widely carried out along with the development of computer science, especially artificial intelligence [2]. On the other hand, deep learning studies are controlled by data [3].

Petrographic research using a microscope is an essential component of geological analysis ranging from rock determination for academic studies to exploration of the mining and petroleum industries. Computer-assisted image analysis techniques have made it easier to characterize the microscopic properties of rocks through digital thin-slice image analysis [4]. In a petrographic study, rock minerals in the thin section were identified with a polarization microscope. Petrographic analysis in both plane-polarized lights (PPL) and crossed polarized light (XPL) conditions is carried out to obtain mineral identification capabilities accurately according to the optical properties of minerals based on the image in thin-section photos (Figure 1).

Image recognition has been widely developed in the last decade using deep learning methods [5]. Studies have also been carried out to introduce minerals in the form of images [6]. Image development using deep learning methods will be influenced by the number of mineral datasets [7]. So far, studies that have been carried out on the application of deep learning methods in the identification of minerals in rocks have shown significant results and are more effective in terms of time [8]. [9] has conducted research using 45 thin slices of rock as a dataset which is then identified by color and produces results

IOP Conf. Series: Earth and Environmental Science 1151 (2023) 012046 doi:10.1088/1755-1315/1151/1/012046

with a high level of accuracy. [10] also conducted the same study with objects in the form of sedimentary rock thin sections. More development needs to be done, especially in developing a more accurate and objective mineral identification model.

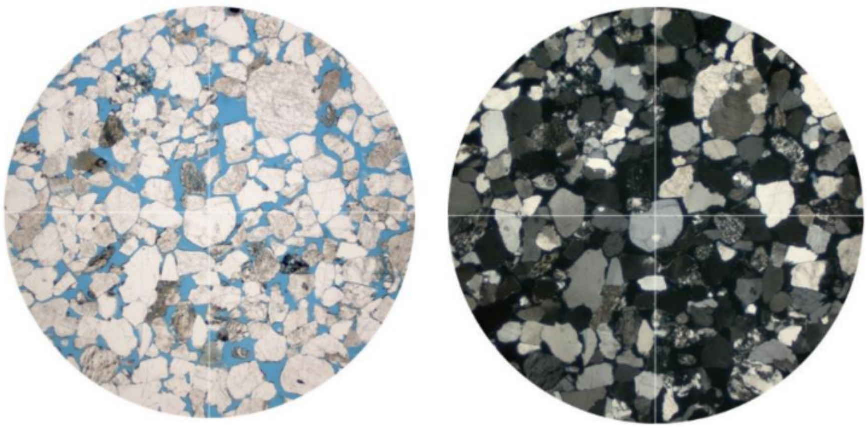

Figure 1. Rock Minerals Thin section in PPL and XPL [11]

Currently, many researchers have researched the design and implementation of deep learning frameworks. Identifying rock minerals images with deep learning methods must be carried out comprehensively using accurate datasets. In this study, the identification of rock minerals will be carried out entirely to solve problems related to accuracy and efficiency in real-time based on YoloV4. The YoloV4 object detection model has been improved to detect rock minerals more accurately and faster.

2. Methodology

In this section, proposed method is performed to build two models of petrographic concepts using YoloV4.

2.1. The architecture of YoloV4

In this study, the architecture of CNN used is YoloV4 [13]. YoloV4 has been shown to detect common objects with a mean average precision of 43.5% with a speed of 65 FPS by using Tesla V100. They can achieve good result of mean average precision by adding new features to CNN architecture, i.e., Weighted-Residual-Connection (WRC), Cross-Stage-Partial-Connection (CSP), Cross-mini-Batch-Normalization (CmBN), Self-Adversarial-Training (SAT), and Mish-Activation.

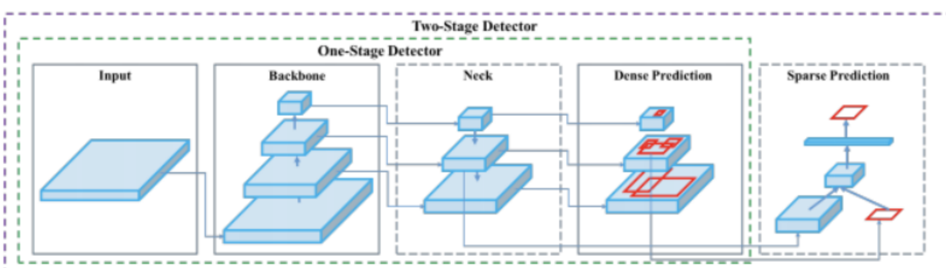

Figure 2. The Construction of object detector [13]

IOP Conf. Series: Earth and Environmental Science 1151 (2023) 012046 doi:10.1088/1755-1315/1151/1/012046

Input images are used by all object detectors, and features are compressed by a convolutional neural network backbone. These backbones serve as the network's core in picture categorization and may be used to forecast future results. The convolutional backbone's feature layers must be blended and held up in comparison to one another in order to create numerous bounding boxes around pictures for object identification and classification. In the neck, the layers of the backbone are combined. Since YOLO is a one-stage detector, its architecture does not make advantage of sparse prediction.

For backbone network, YoloV4 uses CSPDarknet53 to increase learning and decrease computing bottlenecks by transmitting an unaltered version of the future map. Meanwhile in neck network, YoloV4 selects PANet for the network’s feature aggregation and append SPP Block after backbone network to improves the receptive field and secures the major features from the backbone network. YoloV4 uses the same head as YoloV3 in dense prediction network.

There are two layers that are added in YoloV4, bag of freebies and bag of specials. Bag of freebies are employed to increase the performance of the network but do not add up the inference time. SAT is one of many types of bags of freebies which is used by YoloV4 in its architecture. The term of Bag of specials is coined because they boost performance tremendously while only slightly increasing inferencing time. YoloV4 utilizes mish activation function and CmBN for batch normalization.

2.2. Building The Dataset for Training and Validation

The subjects used in this research were minerals in sedimentary rocks, especially quartz, plagioclase, and lithic. The images used in this study were collected from a collective geological website. A python-based application is used called LabelImg [14] to create a dataset, as shown in Figure 3. The dataset created is based on images derived from thin-slice photography. A bounding box for each type of mineral in the photography is provided in each image,

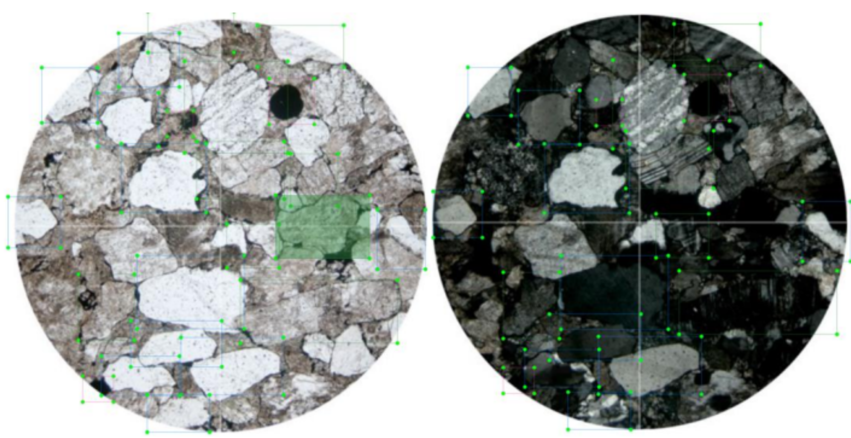

Figure 3. Labelling PPL and XPL image for Training the Models

The bounding box itself is formed from coordinate values (X, Y) of the box's left point, length (X - X'), and height (Y-Y') of the box. Each bounding box will be given a label related to the type of mineral in the bounding box. The number of images labeled is 22 for PPL and XPL, respectively. The total number of ideas combined is 44 images, and those images are accompanied by a txt file that stores the coordinates of the bounding box for each labeled mineral. When all the images have been labeled, the label format is changed into a TXT format that the YOLOv4 architecture can read for further training and create another text file to store the information about the types of minerals that will be used for training the models. The ratio proportion between training, validation, and test datasets of the overall dataset is 55% for the training dataset, 27% for the validation dataset, and 18% for the test dataset.

IOP Conf. Series: Earth and Environmental Science 1151 (2023) 012046 doi:10.1088/1755-1315/1151/1/012046

2.3. *Building The Models*

A desktop computer with the specifications of AMD R5 2600, 16 GB RAM, and RTX3060 used to train the models. By using RTX3060, Cuda cores could be implemented to increase deep learning performance when training or validating the dataset. NVidia jetson nano 4GB is used to test the dataset. Furthermore, the performance of the models using SBC and compare the desktop and SBC could be recognized.

The system was designed to produce a model that can predict mineral types in rock slices in 2 workflows: dataset creation and model training. The two workflows are shown in Figure 4. After creating a dataset for training the models from 3 datasets, 2 of 3 classes, training, and validation is used. Training and validation datasets were prepared so they could be recognized by YoloV4 [13] architecture by creating supporting files named class and object. Furthermore, for completing those supporting files, the parameter for the models needs to be set, such as how many epochs will be used for training, determining how many classes will be used to recognize the minerals, and the number of filters used corresponding to the number of classes. The test dataset will be used later for detecting the minerals to know the model's actual performance in the following work.

For every epoch, the system will calculate the best weight for the models. For every 100 epochs, the system will calculate the MAP to know the model's performance from the validation dataset. Every MAP calculated was plotted in the graph to know which weight was best at the given MAP. The Best MAP value, which corresponds to weights, was saved to the models. For every multiple of 1000 epochs, the system saved the weights into the models. After it reached more than 6000 epochs, the system kept the last weights and stored them in the models. The system produced many models to test each model freely which one was fit to the test dataset in the following work.

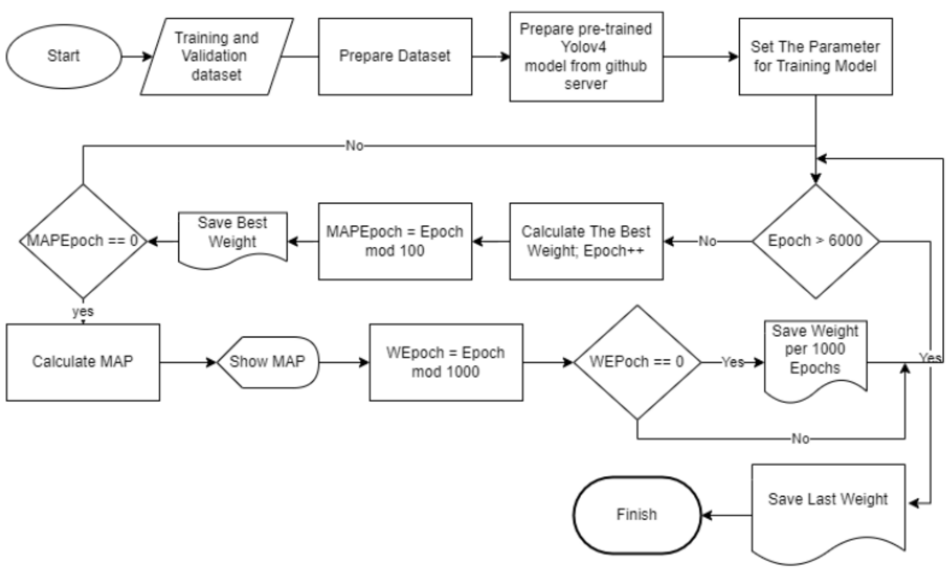

Figure 4. The Flowchart of Training the Models

3. Results and Discussion

YoloV4, which is deep learning, does not need to determine the type of feature extraction used because the architecture already regulates it. The dataset consists of 44 images of minerals in sedimentary rocks. Many factors make up a mineral identification model, including the amount and clarity of rock mineral images, background noise, and differences between mineral features. These factors will all affect the MAP. Thus, the increasing number of rock mineral images will affect the high and low MAP values.

IOP Conf. Series: Earth and Environmental Science 1151 (2023) 012046 doi:10.1088/1755-1315/1151/1/012046

Figure 5 shows the last value of MAP from training the models from the PPL dataset is 2%. MAP's value varied from the highest at 11% to the lowest at 1.8%. But the average loss value decreased as the system approached the final epoch of training. Before the system to train the model started, a pre-trained model from YoloV4 GitHub is developed. At an early stage of the epoch, the system calculated the weights. It generated a high value of the average loss, thus creating a straight blue line at the top of the chart because the value exceeded the maximum value of the y-axis range. As the training stage goes on, the average loss decreases. It can be seen at the 1000th epoch that the average loss value started to match with the range of the y-axis, and the final value of the average loss is 1.192. It shows that as the system calculated weights approaching the last epoch, the more converged the average loss value became.

In creating the XPL training model, as shown in Figure 6, MAP's value is better than PPL's. The highest number of MAP is 19%. Meanwhile, the lowest value is 7%. Another parameter that is better than PPL is average loss. The average loss of XPL is 1.185. The XPL model produces a better result than the PPL model because PPL observations is still not enough to distinguish minerals with many similar optical properties, especially those in one particular group. PPL observations are only preliminary observations that need to be validated by performing XPL observations.

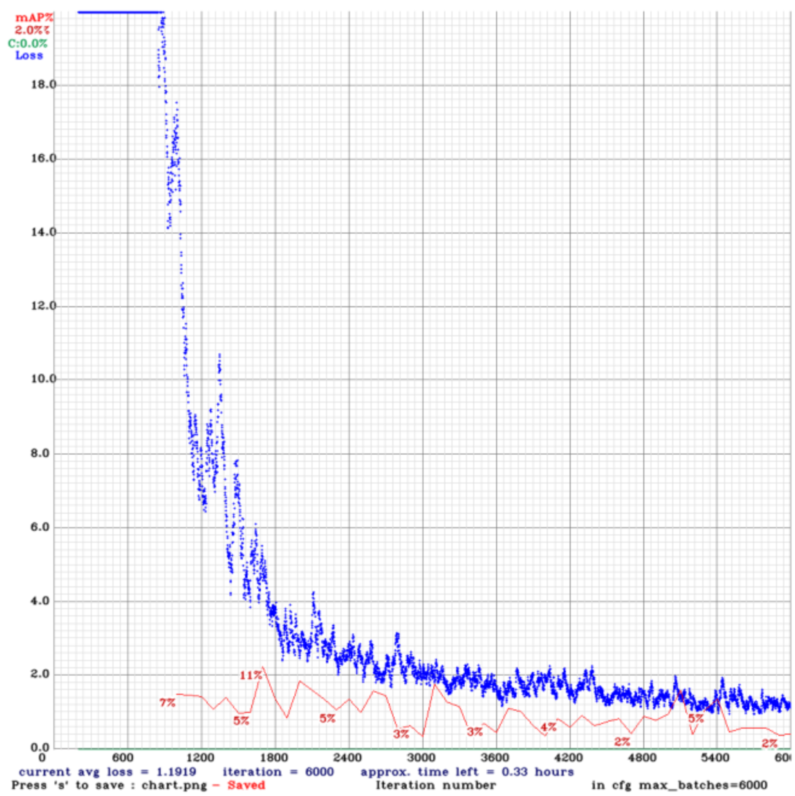

Figure 5. The Graph Showing The Result of Loss and MAP for PPL

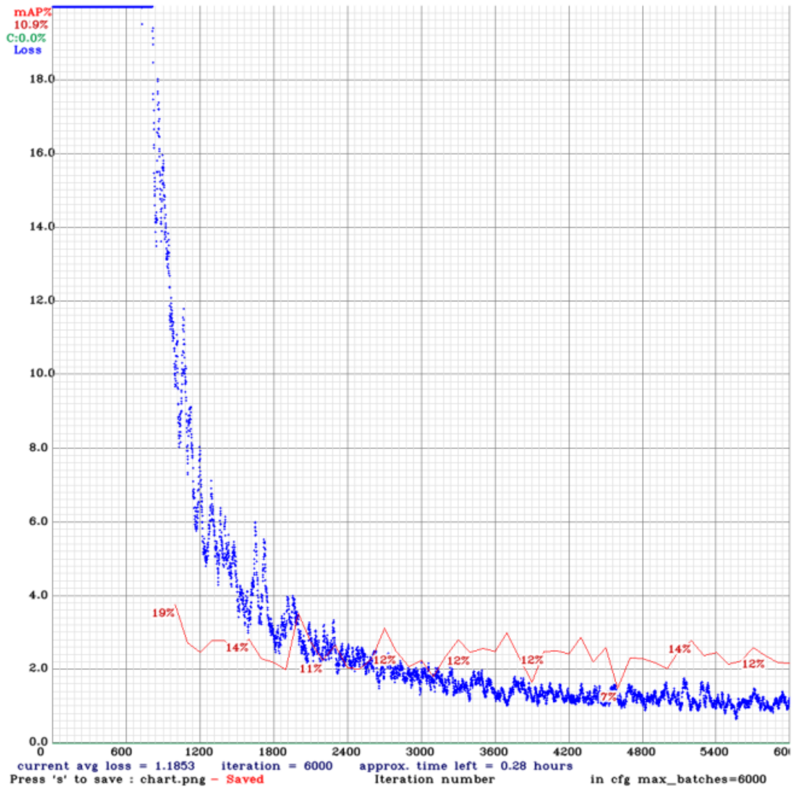

Figure 6. The Graph Showing The Result of Loss and MAP for XPL

4. Conclusion

This study's results indicate that identifying rock mineral thin sections in PPL and XPL with the proposed YoloV4 approach is a very efficient and promising idea. Based on the use of the YoloV4 algorithm, The MAP and loss results showed that the number of parameters of the minerals detection model in PPL is 11% and 1.19, respectively. Meanwhile, The MAP and loss results of XPL are 19% and 1.18, respectively. More comprehensive development is needed, especially in the YoloV4 approach. As for the following works, the model to the PPL and XPL images will be tested that correspond to each other in two kinds of devices, Nvidia Jetson Nano and Desktop PC with RTX3060, and calculate the confusion matrix to get the accuracy, sensitivity, and specificity.

References

[1] Yeshi K, Wangdi T, Qusar N, Nettles J, Craig S R, Schrempf M and Wangchuk P 2018 Geopharmaceuticals of Himalayan Sowa Rigpa medicine: Ethnopharmacological uses, mineral diversity, chemical identification and current utilization in Bhutan *J. Ethnopharmacol.* **223** 99-112

[2] Sadeghi B, Madani N and Carranza E J M 2015 Combination of geostatistical simulation and fractal modeling for mineral resource classification *J. Geochem. Explor.* **149** 59-73

[3] Sachindra D, Ahmed K, Rashid M, Shahid S and Perera B 2018 Statistical downscaling of precipitation using machine learning techniques *Atmos. Res.* **212** 240-258

[4] Srisutthiyakorn N, Hunter S, Sarker R, Hoffmann R and Espejo I 2018 Predicting elastic

properties and permeability of rocks from 2D thin sections *Lead. Edge* **37** 421-427

[5] Li Y, Chen C, Fang R and Yi L 2018 Accuracy enhancement of high-rate GNSS positions using a complete ensemble empirical mode decomposition-based multiscale multiway PCA *J. Asian Earth Sci.* **169** 67-78

[6] Wang W and Chen L 2015 Flotation bubble delineation based on harris corner detection and local gray value minima *Minerals* **5** 142-163

[7] Ślipek B and Młynarczuk M 2013 Application of pattern recognition methods to automatic identification of microscopic images of rocks registered under different polarization and lighting conditions *Geol. Geophys. Environ* **39** 373

[8] Młynarczuk M, Górszczyk A and Ślipek B 2013 The application of pattern recognition in the automatic classification of microscopic rock images *Comput. Geosci.* **60** 126-133

[9] Aligholi S, Lashkaripour G R, Khajavi R and Razmara M 2016 Automatic mineral identification using color tracking *Pattern Recognit. 65* 164-174

[10] Li N, Hao H, Gu Q, Wang D, Hu X 2017 A transfer learning method for automatic identification of sandstone microscopic images *Comput. Geosci.* **103** 111-121

[11] https://www.virtualmicroscope.org/sites/default/files/html5Assets/pblue/index2.html?specimen=/node/1430 accessed on July 11th 2022

[12] Saxena N, Day-Stirrat R J, Hows A and Hofmann R 2021 Application of Deep Learning for Semantic Segmentation of Sandstone Thin Sections *Computers & Geosciences* **152** 104778

[13] Bochkovskiy A, Wang C Y and Liao H Y M 2020 Yolov4: Optimal speed and accuracy of object detection arXiv preprint arXiv:2004.10934.

[14] Tzutalin LabelImg. Git code 2015 https://github.com/tzutalin/labelImg

Acknowledgment

This study was supported by Institut Teknologi Nasional Yogyakarta Research Fund. The authors are grateful to the Department of Electrical Engineering and Department of Geological Engineering, Institut Teknologi Nasional Yogyakarta, Indonesia for full supporting.

Building YoloV4 models for identification of rock minerals in thin section

ORIGINALITY REPORT

17%	14%	16%	10%
SIMILARITY INDEX	INTERNET SOURCES	PUBLICATIONS	STUDENT PAPERS

PRIMARY SOURCES

1	Submitted to Sogang University Student Paper	5%
2	coek.info Internet Source	2%
3	Michael Abebe Berwo, Zhipeng Wang, Yong Fang, Jabar Mahmood, Nan Yang. "Off-road Quad-Bike Detection Using CNN Models", Journal of Physics: Conference Series, 2022 Publication	2%
4	Bagus Gilang Pratama, Oni Yuliani. "Monitoring and Controlling Air Conditioner (AC) Using Improvised CNN and Internet of Things", 2021 International Conference on Converging Technology in Electrical and Information Engineering (ICCTEIE), 2021 Publication	2%
5	Dihua Wu, Shuaichao Lv, Mei Jiang, Huaibo Song. "Using channel pruning-based YOLO v4 deep learning algorithm for the real-time and accurate detection of apple flowers in natural	1%

environments", Computers and Electronics in Agriculture, 2020
Publication

6	Xiaoxin Chen, Zhansi Jiang, Hao Cheng, Hongxin Zheng, Yixian Du. "LF-YOLOv4: a lightweight detection model for enhancing the fusion of image features of surface defects in lithium batteries", Measurement Science and Technology, 2023 Publication	1 %
7	eprints.whiterose.ac.uk Internet Source	1 %
8	repository.itny.ac.id Internet Source	1 %
9	"Artificial Neural Networks and Machine Learning – ICANN 2018", Springer Science and Business Media LLC, 2018 Publication	1 %
10	Submitted to University of Surrey Student Paper	1 %
11	Xinrui Zhang, Lei Li, Bin Yan. "Material Decomposition of Dual-Energy CT Based on CBAM Attention U-Net", Journal of Physics: Conference Series, 2022 Publication	<1 %
12	pubmed.ncbi.nlm.nih.gov Internet Source	<1 %

13	link.springer.com Internet Source	<1 %
14	Changyu Jin, Kai Wang, Tao Han, Yu Lu, Aixin Liu, Dong Liu. "Segmentation of ore and waste rocks in borehole images using the multi-module densely connected U-net", Computers & Geosciences, 2022 Publication	<1 %
15	Submitted to Universitas Bangka Belitung Student Paper	<1 %

Exclude quotes	Off	Exclude matches	Off
Exclude bibliography	On		