ASEAN Engineering Journal

Full Paper

IMAGE-BASED YOLOV4 ARCHITECTURE FOR DETECTING MINERAL IN SEDIMENTARY ROCKS THIN SECTION

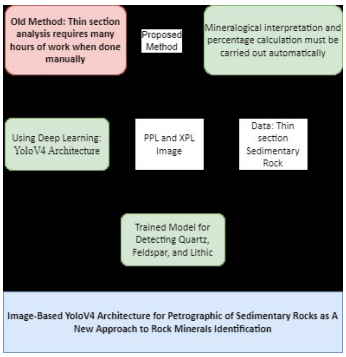
Bagus Gilang Pratama[a], Muhamad Fatih Qodri[b]*, Oky Sugarbo[b]

[a]Electrical Engineering Department, Faculty of Industrial Technology, Institut Teknologi Nasional Yogyakarta, Yogyakarta, 55281, Indonesia
[b]Geological Engineering Department, Faculty of Mineral Technology, Institut Teknologi Nasional Yogyakarta, Yogyakarta, 55281, Indonesia

Article history
Received
2 March2023
Received in revised form
24 October 2023
Accepted
26 October 2023
Published online
31 August 2024

*Corresponding author
fatihqodri@itny.ac.id

Graphical abstract



Abstract

Thin section analysis of sedimentary rocks is the basis for identifying minerals and textures. In general, quantitative analysis of thin sections of rock often requires many hours of work when done manually. In today's era, mineralogical interpretation and percentage calculations must be carried out automatically using more practical applications. The research method begins with the identification of 44 thin section samples in parallel plane polarized (PPL) and crossed polarized (XPL) conditions with thin section analysis then mineralogy detection is carried out using a computational approach, namely the use of image-based Deep Learning YoloV4 architecture with 2D RGB image objects from the thin section of sedimentary rock. The results of this study show the best values of Average Precision in Quartz, Feldspar, and lithic are 39.21% in the XPL model, 26.53% in the XPL model, and 15.75% in combined mode, according to the training and testing of YoloV4 Models for the identification of rock minerals in thin sections. Based on the complexity of the mineral types, the granularity of the detection, and the specific geological objectives, establishing a meaningful benchmark or baseline for comparison is always challenging. Additionally, consider discussing the trade-offs between precision and recall, as a higher precision may be more critical in some geological applications. It is expected that the application of this research can produce practical, fast and accurate interpretation of the determination of minerals in sedimentary rocks from all thin-section images of rocks and thus provide a complete understanding of geological views automatically.

Keywords: YOLOv4, Deep learning, Mineral, Sedimentary rock, Image

1.0 INTRODUCTION

Sedimentary rocks, a fundamental component of the Earth's crust, hold invaluable insights into the planet's geological history and environmental changes over time. The identification of minerals in clastic sedimentary rocks (sandstones) requires microscopic interpretation and statistical calculations that are collected in real-time on a thin section polarizing microscope [1]. In the last three decades, the classification of sedimentary rocks has been developed to objectively classify sandstones by taking into account the main material types in sedimentary rocks, namely quartz, feldspar and lithic (QFL) [2][3][4]. The classification has been developed according to the latest industry standards such as the oil and gas industry to assess the concept of risk in reservoirs [5].

The thin section study using a microscope is an important component of geological analysis from rock determination for academic studies to exploration of the mining and petroleum industries [6]. Advances in computer-assisted image analysis techniques have made the characterization of microscopic rock properties through digital thin section image analysis easier [7]. The ability to accurately identify and characterize minerals within these rocks is crucial for geologists, paleontologists, and various other scientific disciplines [8]. Traditionally, mineral identification in sedimentary rock thin sections has been a labor-intensive and time-consuming process, often requiring manual examination under microscopes. Absolutely, the need to automate some of these processes will be a challenge in geology.

Regarding the determination of rock composition, point counting has become the most commonly used method to determine the abundance of components in a sedimentary rock in thin sections [9]. This process is done manually to obtain the optical properties of a rock component. Although point counting can provide a numerical description of rock composition, calculations are not always like this and there is a high risk of errors in calculations.

According to the optical properties of minerals, mineral identification of thin sections image needs to be carried out for analysis in both planes polarized light (PPL) and crossed polarized light (XPL) conditions as shown in Figure 1. This needs to be done digitally, quickly and accurately. The image method has been done before and needs to be developed further [9][10].

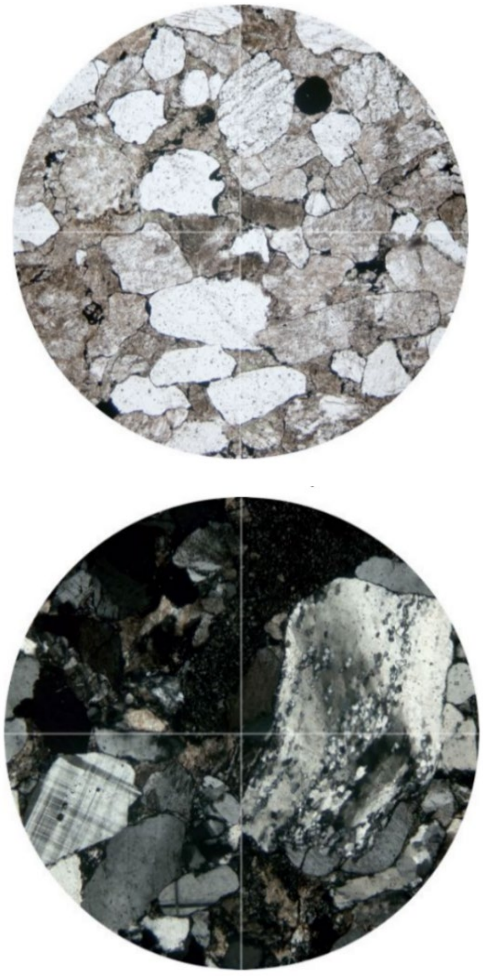


Figure 1 PPL and XPL Images

Deep learning has been used to automate tasks in many fields beyond geoscience, such as analysing handwriting and identifying symbol. Several attempts have been made to automate petrographic image analysis with deep learning. [11][12][13] used information from XPL conditions and image gradient information to make thin section segmentation and mineral identification. The novelty in this research is the use of Deep Learning architecture, named YOLOv4. YOLOv4 is an object detection model developed from YOLOv3 with several additional features such as Weighted-Residual-Connection (WRC), Cross-Stage-Partial-Connection (CSP), Cross-mini-Batch-Normalization (CmBN), Self-Adversarial-Training (SAT) and Mish-Activation [14]. With the addition of these features, YoloV4 offers fast and robust architecture to detect any kind of object from trained model [15]. Thus, researchers can use this method to build a model as a reference for the detection of quartz, feldspar, and lithic.

The application of YOLOv4 to mineral detection in sedimentary rocks addresses some of the inherent challenges in manual identification. Thin sections often contain a multitude of minerals, each with distinct optical properties and textural characteristics. Distinguishing these minerals requires expertise and a keen eye, and even then, human interpretation can be subjective and prone to error. The YOLOv4 model, on the other hand, is designed to recognize patterns and features that may elude human observers, providing a consistent and objective approach to mineral identification.

By training the YOLOv4 architecture on a diverse dataset of annotated thin section images, the model can learn to recognize the spectral signatures and spatial arrangements associated with different minerals. The resulting model can then be used to automatically identify and delineate minerals within new, unseen thin sections, significantly expediting the analysis process. Furthermore, YOLOv4's real-time capabilities enable rapid analysis of large datasets, allowing researchers to extract valuable insights more efficiently.

2.0 METHODOLOGY

In this study, one branch of Artificial Intelligence, namely Deep Learning will be utilized. The use of Deep Learning in mineral detection is not new. [13] identify minerals and textures in sedimentary rocks. Then, they also automate the mineralogical interpretation of thin sections. This research is carried out two thin section concepts mineral detection, namely the plane polarized light (PPL) and the crossed polarized light (XPL) which were much more accurate in mineral identification. This research is also shown mineral calculations (point counting) automatically according to the predictions of the system to be built. [14] conducted point counting and segmentation on quartz, feldspar, lithic, porosity, and density minerals. By using two types of microscopic images, PPL and XPL, they segmented them into 5 classes, namely porosity, rock fragments, feldspar, quartz, and density. The CNN architecture they use for segmentation and point counting is U-Net. In contrast to what researchers offer, researchers focus more on naming rocks and using the YoloV4 architecture in detecting the types of minerals present in sandstone [16].

This research is designed in system in 3 workflows, namely dataset creation, model training, and mineral detection testing as shown in the research flowchart in Figure 2 In this study, microscopic images in PPL and XPL are employed, then point counting is conducted to calculate the percentage of minerals for naming sedimentary rocks. The method used in this research is YoloV4 [17]. YoloV4 is an object detector based on Convolutional Neural Network (CNN) (Figure 3). Weighted Residual Connection (WRC), Cross-Stage-Partial-Connection (CSP), Cross-mini-Batch-Normalization (CmBN), Self-Adversarial-Training (SAT) and Mish-Activation are implemented on its architecture [15][18]. So that it produces an Average Precision of 43.5% on the MS COCO dataset with a speed of 65 FPS (Frames per Second) using a Tesla V100. The images used in this

study were collected from collective geological sites. Used a python-based application called LabelImg to create the dataset Tzutalin (2015).

The dataset created is based on images derived from the thin section photography containing various mineral types. A bounding box for each type of mineral in the photography is provided in each image. The dataset used in making the PPL, XPL, and the combination of the two models. It consists of 44 mineral images in sedimentary rocks. Many factors influence the identification model, including the number and clarity of elements in the rock image, the noise present in the image, and the different features present in the image. Several of these factors will affect the MAP (Mean Average Precision) value. So, increasing the number of rock images will affect the high and low of the MAP value [20][21].

The next step is annotating the dataset by labeling each image with bounding boxes around mineral regions. This could be a time-intensive step, as it requires domain expertise to accurately label minerals. The dataset was labelled using labelimg to provide the coordinates of mineral on the images (Figure 4).

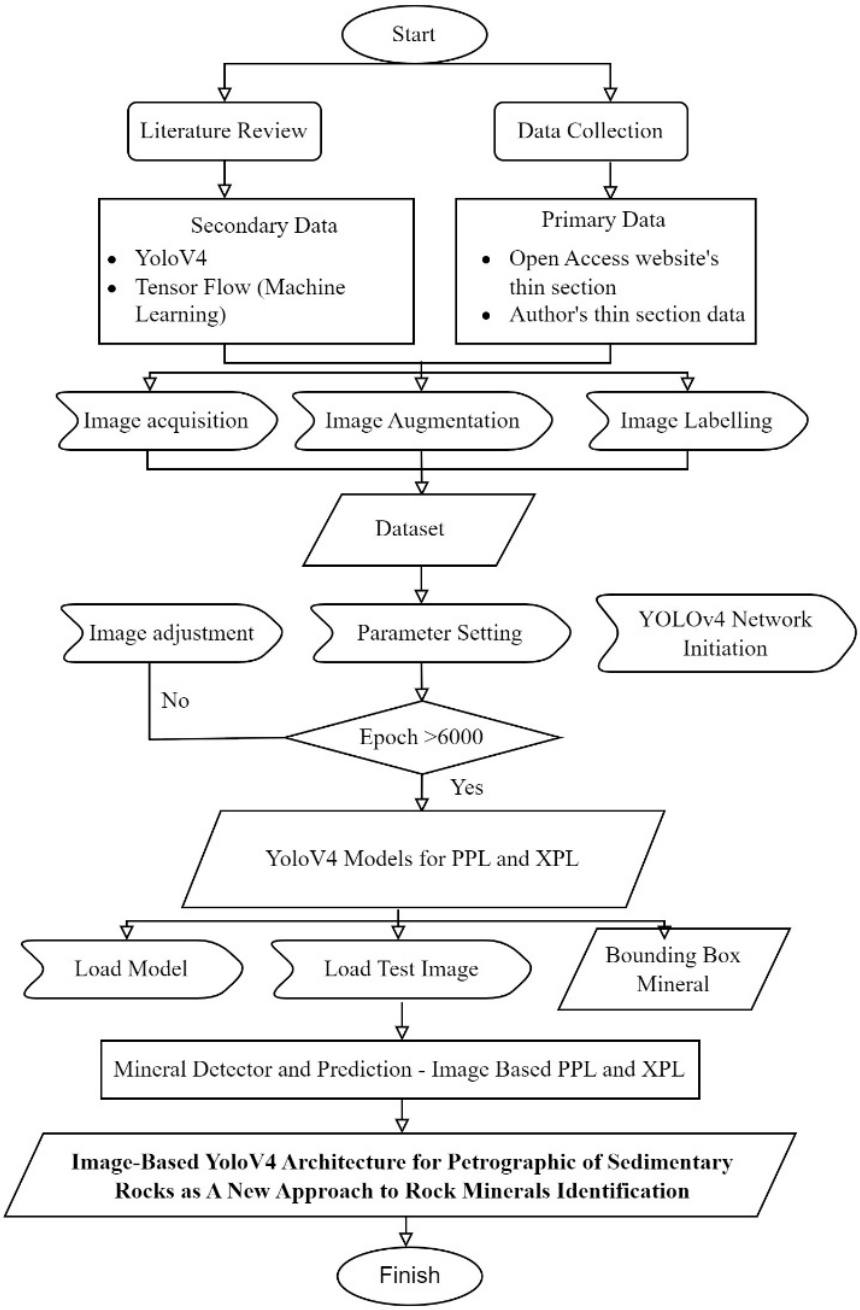


Figure 2 Research Flow

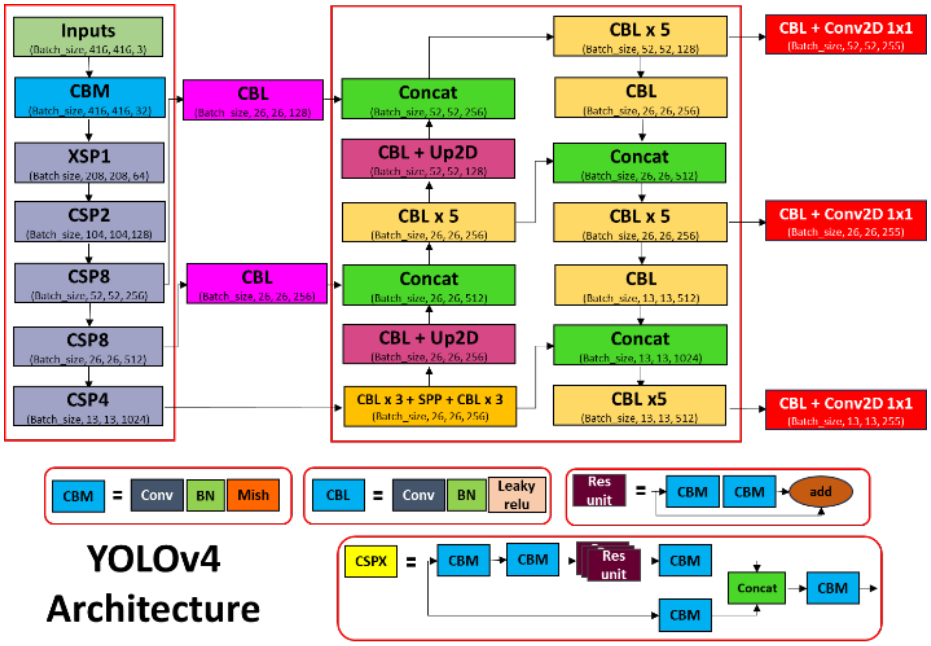


Figure 3 YoloV4 Architecture to Detect Minerals.

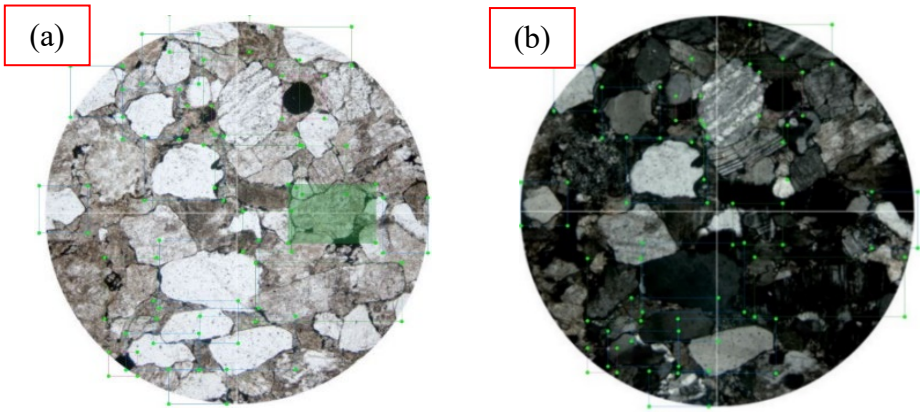


Figure 4 (a) Labelled PPL Image (b) Labelled XPL Image
(c) The Coordinates of Labelled Minerals (Pratama et al., 2023)

Three mineral classes were determined in this dataset, lithic, feldspar, and Quartz.

The PPL model will be trained relying solely on images from PPL, as well as the XPL model which will be trained on XPL images (Pratama et al., 202) by implementing the YOLOv4 architecture. In this study, the two models will be combined with XPL and PPL image training to see how far the performance of these models is compared to other models. The concept of testing the PPL model only uses PPL images, as well as XPL images, while the combined model will be tested by combining XPL and PPL thin section images. The parameters used in this study to configure YoloV4 are shown in Table 1

Table 1 The Parameters of YoloV4 Model for Minerals Detection

Parameters	Value
Input Size	609x608
Learning Rate	0.001
Batch	64
Classes	3
Max batches	6000

The testing model is built by using a desktop computer with AMD R5 2600 specifications, 16 GB of RAM, and an RTX 3060 graphics card. To evaluate three model's performance, four metrics, including precision, recall, mAP, and F1-score, were employed in this study. It is a true positive scenario where the IOU (Intersection Over Union) is greater than 0.5 and when IOU values below 0.5 indicate a false positive. The IOU is a false negative instance when it equals 0. The following equations (Equation 1- Equation 5) illustrate the results of the calculations for the IOU, precision, recall, mAP, and F1-Score. The mAP in this case represents the average score of the AP (Average Precision) during minerals detection; the greater the score, the better the minerals detection outcome.

$$\boldsymbol{IOU(R,R')} = \frac{|R \cap R'|}{|R \cup R'|} \quad \text{Equation 1}$$

$$\boldsymbol{precision} = \frac{TP}{TP+FP} * \mathbf{100\%} \quad \text{Equation 2}$$

$$\boldsymbol{recall} = \frac{TP}{TP+FN} * \mathbf{100\%} \quad \text{Equation 3}$$

$$\boldsymbol{mAP} = \frac{\Sigma_{C=1}^{C} AP(c)}{C} \quad \text{Equation 4}$$

$$\boldsymbol{F_1} = \mathbf{2}.\frac{precision\,.\,recall}{precision+recall} \quad \text{Equation 5}$$

Where R is the object's detected bounding box area and R' is the true bounding box area. The numbers of true positive, false positive, and false negative are denoted as TP, FP, and FN, respectively. Meanwhile C represents how many classes in this study which is three classes.

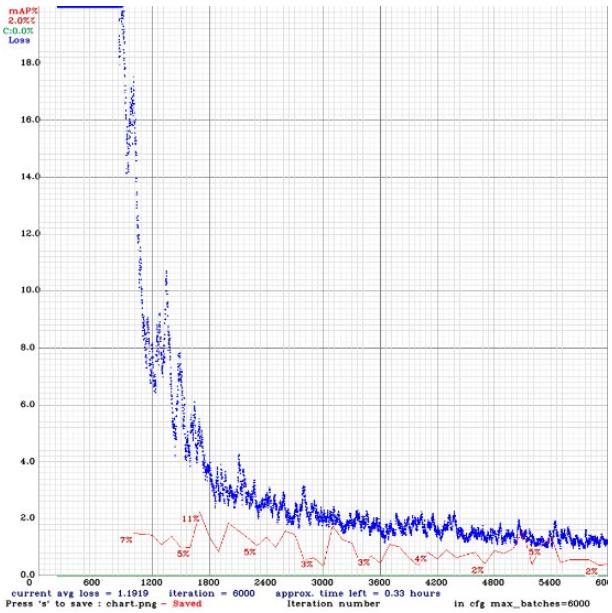


Figure 5 The result of Loss and MAP for PPL (Pratama et al., 2023)

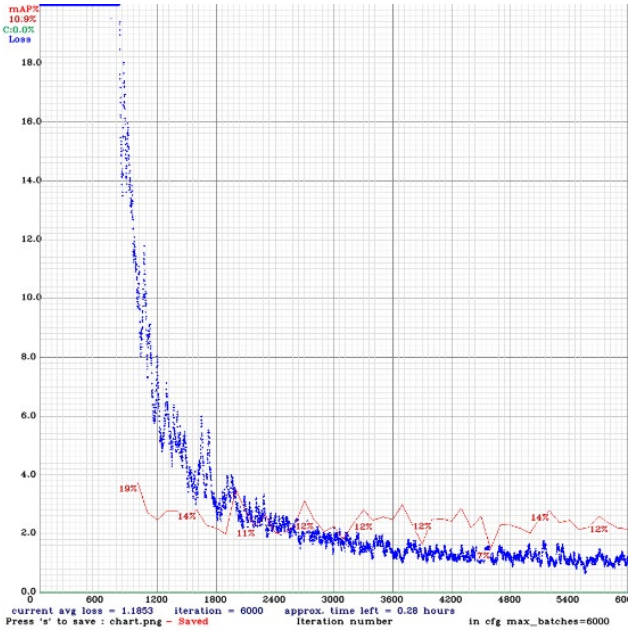


Figure 6 The result of Loss and MAP for XPL (Pratama et al., 2023)

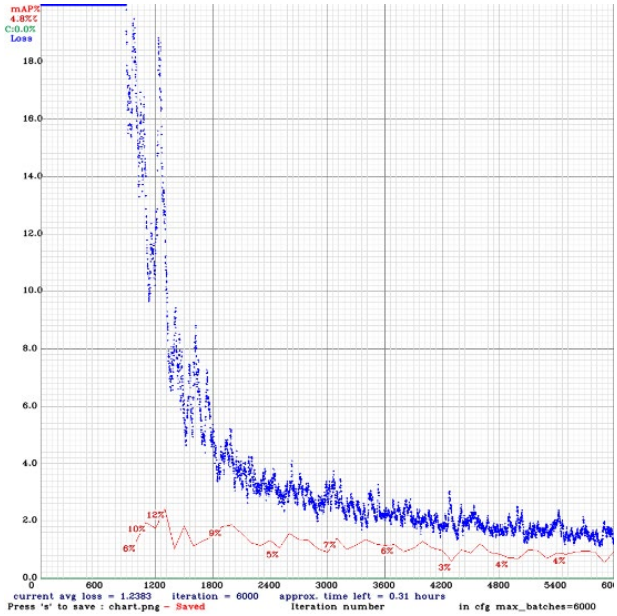


Figure 7 The result of Loss and MAP for XPL and PPL Combined

The mAP values obtained from training three models with 3 predetermined scenarios are as follows 11% in PPL, 19% in XPL, and 12% in both combinations as shown in Figure 5, Figure 6, and Figure 7. These three models will be tested later with a new image. These models will be tested to detect 3 types of composition, such as Feldspar, Quartz, and Lithic. The clast composition consisting of quartz and feldspar are characterized by being monomineralic. The lithic composition encompasses various grain types composed of mineral mixtures, including quartz, feldspar, clay, heavy minerals, volcanic groundmass, and carbonate. The test results are in the form of a confusion matrix which will then calculate the average precision value of each mineral class. These models will then be compared to see which model has the best detection capability.

3.0 RESULTS AND DISCUSSION

The number of images is used in the detection of minerals is five images. However, the total number of objects in the image is 83 objects which are divided into three classes of mineral types (Table 2).

Table 2 The Results of Mineral Detection Using YoloV4

Model	Precision (%)	Recall (%)	Average IoU (%)	mAP (%)	F1-Score
PPL	23	25	16.12	14.68%	0.24
XPL	29	45	20.40	24.79%	0.35
Combined	20	35	14.07	17.94	0.25

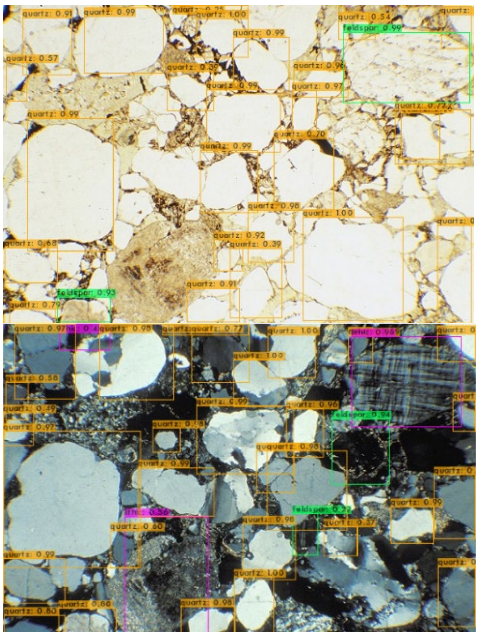


Figure 8 Testing the dataset using the PPL and XPL models, respectively

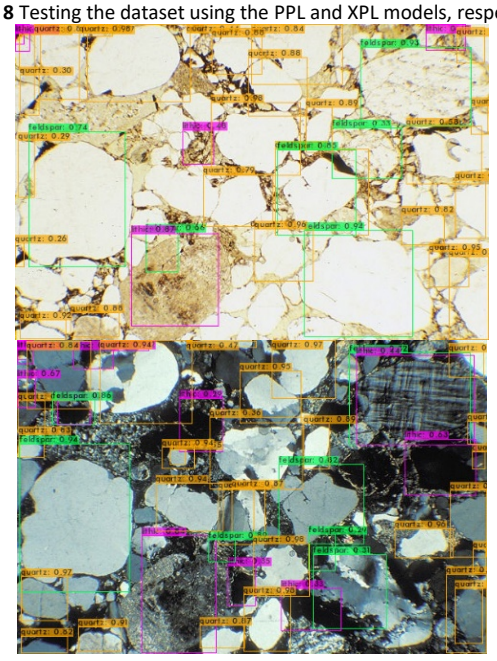


Figure 9 Testing the dataset using combination the PPL

Testing the dataset using the PPL model, researchers found a MAP value of 14.68% with details of an average precision of lithic minerals of 0%, quartz of 30.79%, and feldspar of 13.24% (Figure 8). The XPL model obtained better test results than the PPL model, the MAP value obtained was 24.79% with lithic details of 8.61%, quartz 39.21%, and feldspar 26.53% (Figure 9). The third model that combines images between XPL and PPL in one training data gets a MAP value of 17.94% with details of average precision lithic of 15.75%, quartz of 26.78%, and feldspar of 11.29% (Figure 9). Details for overall scores are shown in Figure 10.

The Mean Average Precision (MAP) value for the image-based YOLOv4 architecture's mineral detection in sedimentary rock thin sections should always be interpreted in context. The acceptable MAP value can vary depending on factors such as the complexity of the mineral types, the granularity of the detection, and the specific geological objectives. Based on those factors, establish a meaningful benchmark or baseline for comparison is always challenging. Additionally, consider discussing the trade-offs between precision and recall, as a higher precision may be more critical in some geological applications.

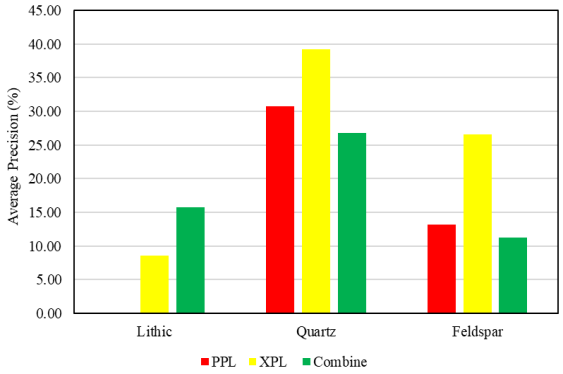


Figure 10 average precision results on each PPL, XPL and both observations

The XPL model outperforms the other two models at identifying quartz and feldspar in fresh photos because XPL images have the benefit of having image parameters with relatively distinct differences, like hue, twinning, and darkness. As a result, the XPL model can benefit from reduced glare and better object identification performance.

The combined model, on the other hand, performs lithic detection more effectively because the complimentary information from the two polarization models can assist the model better identify and categorize various types of lithic composition. While XPL photos can show internal structure and qualities, PPL images give information on surface textures and features. The model can use a wider range of attributes by integrating the two photos, creating a more robust depiction of the lithic compositions.

4.0 CONCLUSION

According to the training and testing of YoloV4 Models for identification of rock minerals in thin section, the results of this study show the Plane Polarized Light (PPL), Crossed Polarized

Light (XPL) and both datasets show a fairly good average precision for each mineral, namely 39.21% for the best value of quartz, 26.53% for the best value of feldspar and 15.75% for the best value of lithic. The model built from the PPL dataset has the lowest value in detecting lithic, whereas using the model built from XPL data in detecting Quartz and Feldspar minerals is better than the other three models. The last model, a combination of PPL and XPL thin section image, is better at detecting lithics than the other models. Aaccording to contextual factors like geological significance, detection requirements, and the potential impact of false positives/negatives should the guide of interpretation of the MAP value. This research provides nuanced perspective and accurately convey the practical implications of the result's performance.

Acknowledgement

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.

References

[1] Dutton, S.P., Loucks, R.G. and Day-Stirrat, R.J., 2012. Impact of regional variation in detrital mineral composition on reservoir quality in deep to ultradeep lower Miocene sandstones, western Gulf of Mexico. *Marine and Petroleum Geology*, 35(1): 139-153. https://doi.org/10.1016/j.marpetgeo.2012.01.006

[2] Folk, R.L., 1974. Petrology of sedimentary rocks: Austin, Texas.

[3] Pettijohn, F.J., 1975. Sedimentary rocks. 3: 628. New York: Harper & Row.

[4] Lander, R.H. and Walderhaug, O., 1999. Predicting porosity through simulating sandstone compaction and quartz cementation. *AAPG bulletin*, 83(3): 433-449.

[5] Nasseri, M.H.B. and Mohanty, B., 2008. Fracture toughness anisotropy in granitic rocks. I*nternational Journal of Rock Mechanics and Mining Sciences*. 45(2): 167-193. https://doi.org/10.1016/j.ijrmms.2007.04.005

[6] de Lima, R.P., Duarte, D., Nicholson, C., Slatt, R. and Marfurt, K.J., 2020. Petrographic microfacies classification with deep convolutional neural networks. *Computers & geosciences*, 142: 104481. https://doi.org/10.1016/j.cageo.2020.104481

[7] Srisutthiyakorn, N., Hunter, S., Sarker, R., Hofmann, R. and Espejo, I., 2018. Predicting elastic properties and permeability of rocks from 2D thin sections. *The Leading Edge*, 37(6): 421-427. https://doi.org/10.1190/tle37060421.1

[8] Stehli, F.G. and Webb, S.D. eds., 2013. T*he great American biotic interchange.* 4. Springer Science & Business Media.

[9] Lander, R.H., Larese, R.E. and Bonnell, L.M., 2008. Toward more accurate quartz cement models: The importance of euhedral versus noneuhedral growth rates. AAPG bulletin, 92(11): 1537-1563. https://doi.org/10.1306/07160808037

[10] Ciregan, D., Meier, U. and Schmidhuber, J., 2012, June. Multi-column deep neural networks for image classification. In *2012 IEEE conference on computer vision and pattern recognition*. 3642-3649. IEEE. https://doi.org/10.1109/CVPR.2012.6248110

[11] Budennyy, S., Pachezhertsev, A., Bukharev, A., Erofeev, A., Mitrushkin, D. and Belozerov, B., 2017, October. Image processing and machine learning approaches for petrographic thin section analysis. In *SPE Russian Petroleum Technology Conference*. OnePetro. https://doi.org/10.2118/187885-RU

[12] Ma, Z. and Gao, S., 2017, May. Image analysis of rock thin section based on machine learning. In International Geophysical Conference, Qingdao, China, 17-20 April 2017. 844-847. Society of Exploration Geophysicists and Chinese Petroleum Society. https://doi.org/10.1190/IGC2017-213

[13] Maitre, J., Bouchard, K. and Bédard, L.P., 2019. Mineral grains recognition using computer vision and machine learning. Computers & Geosciences. 130: 84-93. https://doi.org/10.1016/j.cageo.2019.05.009

[14] 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. https://doi.org/10.1016/j.cageo.2021.104778

[15] Tang, D.G., Milliken, K.L. and Spikes, K.T., 2020. Machine learning for point counting and segmentation of arenite in thin section. *Marine and Petroleum Geology*, 120: 104518. https://doi.org/10.1016/j.marpetgeo.2020.104518

[16] Miao, J., Hirakawa, T., Yamashita, T. and Fujiyoshi, H., 2021, February. 3D Object Detection with Normal-map on Point Clouds. In VISIGRAPP 5: 569-576. https://doi.org/10.5220/0010304305690576

[17] Pratama, B.G. and Yuliani, O., 2021, October. Monitoring and Controlling Thermal Comfort in Air Conditioner Using YoloV4 And Predicted Mean Vote. In *2021 7th International Conference on Electrical, Electronics and Information Engineering (ICEEIE)*. 460-465. IEEE. https://doi.org/10.1109/ICEEIE52663.2021.9616849

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

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

[20] Pratama, B. G., Qodri, M. F., & Sugarbo, O. (2023, March). Building YoloV4 models for identification of rock minerals in thin section. In *IOP Conference Series: Earth and Environmental Science.* 1151(1): 012046. IOP Publishing.

[21] Wu, D., Lv, S., Jiang, M. and Song, H., 2020. Using channel pruning-based YOLO v4 deep learning algorithm for the real-time and accurate detection of apple flowers in natural environments. *Computers and Electronics in Agriculture*. 178: 105742. https://doi.org/10.1016/j.compag.2020.105742