



## INTISARI

# PENGENALAN WAJAH DARI POTONGAN BAGIAN WAJAH DENGAN MASK R-CNN

Oleh

Rahmat Budiarsa

20/468177/SPA/00743

Pengenalan wajah, domain penting dalam *computer vision*, memerlukan pendeteksian dan identifikasi atribut wajah. Studi ini berfokus pada kumpulan data dengan objek wajah teroklusi yang terdiri dari wajah-wajah orang Indonesia, dimana 'oklusi' menunjukkan data wajah yang tidak memiliki visibilitas lengkap yang mencakup kejadian dimana objek memiliki tampilan *angle* yang tidak memperlihatkan beberapa bagian wajah atau terpotong sebagian. Di tengah kondisi ini, metode Mask R-CNN adalah pendekatan yang menonjol. Meskipun penelitian sebelumnya sebagian besar mempelajari penyempurnaan *loss function* dan penyempurnaan deteksi objek atau wajah, pengenalan dan identifikasi wajah dengan data yang tidak sempurna masih relatif belum dieksplorasi.

Tujuan utama metode Mask R-CNN adalah untuk memastikan kehandalan dalam mencapai pengenalan akurat untuk input wajah yang tidak jelas atau wajah teroklusi. Investigasi ini melibatkan eksperimen dengan mengganti *Intersection over Union* (IoU) dengan *Generalized Intersection over Union* (GIoU) dari *Region Proposal Network* (RPN), selain itu dilakukan modifikasi arsitektur pada bagian RPN dan *Fully Convolutional Neural Network* (FCN) dengan kombinasi berbagai penggunaan *pooling* agar sesuai dengan objek wajah Indonesian teroklusi, sehingga meningkatkan akurasi dalam tugas pengenalan dan segmentasi. Inovasi dalam modifikasi GIoU terletak pada pemanfaatan strategis *anchors*, yang dapat menggunakan anchor yang berada di luar batas *ground-truth box* untuk mengoptimalkan efisiensi komputasi. Sedangkan, kombinasi penggunaan *pooling* dapat mempengaruhi akurasi identifikasi dan segmentasi.

Hasil eksperimen dalam penelitian ini sangat mengejutkan; Model GIoU yang dibangun dengan dataset *training* wajah teroklusi, menghasilkan peningkatan kinerja akurasi dibandingkan dengan pendekatan Mask R-CNN konvensional dengan peningkatan 4,65% untuk masing-masing nilai mAP, mAR, dan F1-score. Lebih lanjut, kombinasi Mask R-CNN dengan GIoU dan FCN *pooling* P2 dan P3 adalah model terbaik dari variasi 12 model arsitektur yang dibangun dengan peningkatan mAP, mAR, dan F1-score. Sedangkan pada tahapan *testing* menunjukkan penggunaan model GIoU yang digunakan pada mask R-CNN konvensional dapat meningkatkan nilai *macro* F1-score bahkan pada model terbaik yang diusulkan yaitu Mask R-CNN dengan GIoU dan FCN *pooling* P2 dan P3 peningkatan yang diperoleh sebesar 18,03%.

**Kata Kunci:** *Mask R-CNN, Face Recognition, Image Recognition, Deep Learning, Generalized Intersection over Union, Region Proposal Network, Fully Convolutional Network.*



## ABSTRACT

# FACE RECOGNITION FROM FACE PARTS WITH MASK R-CNN

By

Rahmat Budiarsa

20/468177/SPA/00743

Face recognition, an important domain in computer vision, requires detection and identification of facial attributes. This study focuses on a dataset with occluded facial objects consisting of Indonesian faces, where 'occlusion' denotes facial data that does not have complete visibility which includes instances where the object has an angle view that does not show some parts of the face or is partially cut off. Amidst these conditions, the Mask R-CNN method is a prominent approach. Although previous research has mostly studied loss function refinement and object or face detection refinement, face recognition and identification using imperfect data remains relatively unexplored.

The main objective of the Mask R-CNN method is to ensure reliability in achieving accurate recognition for input faces that are unclear or occluded faces. This investigation involves experimenting with replacing the Intersection over Union (IoU) with the Generalized Intersection over Union (GIoU) of the Region Proposal Network (RPN), in addition to architectural modifications to the RPN and Fully Convolutional Neural Network (FCN) parts with the combination of various pooling usages to suit occluded Indonesian face objects, thus improving accuracy in recognition and segmentation tasks. The innovation in the GIoU modification lies in the strategic utilization of anchors, which can use anchors that are outside the boundary of the ground-truth box to optimize computational efficiency. Meanwhile, the combined use of pooling can affect identification and segmentation accuracy.

The experimental results in this study are striking; the GIoU model built with an occluded face training dataset, resulted in improved accuracy performance compared to the conventional Mask R-CNN approach with an improvement of 4.65% for each of the mAP, mAR, and F1-score values. Furthermore, the combination of Mask R-CNN with GIoU and FCN pooling P2 and P3 is the best model from a variety of 12 architectural models built with improved mAP, mAR, and F1-score. Meanwhile, the testing stage shows that the use of the GIoU model used in conventional R-CNN masks can increase the macro F1-score value even in the best proposed model, namely Mask R-CNN with GIoU and FCN pooling P2 and P3, the increase obtained is 18.03%.

**Keywords:** *Mask R-CNN, Face Recognition, Image Recognition, Deep Learning, Generalized Intersection over Union, Region Proposal Network, Fully Convolutional Network.*