

Deteksi Objek Pariwisata Sekitar Monas Menggunakan YOLO dan RetinaNet = Tourism Object Detection Around Monumen Nasional (Monas) Using YOLO and RetinaNet

Median Hardiv Nugraha, author

Deskripsi Lengkap: <https://lib.ui.ac.id/detail?id=20511004&lokasi=lokal>

Abstrak

Sektor pariwisata menjadi salah satu sektor yang memiliki banyak potensi pemasukan anggaran negara. Salah satu cara untuk meningkatkan pemasukan melalui sektor pariwisata adalah dengan memanfaatkan teknologi informasi agar dapat menarik lebih banyak wisatawan yang datang. Pemanfaatan teknologi tersebut adalah dengan menggunakan smart tourism. Implementasi smart tourism yang digunakan pada pariwisata di Indonesia, khususnya untuk objek wisata Monumen Nasional (Monas) adalah dengan memanfaatkan aplikasi telepon pintar berbasis Visual Question Answering (VQA) untuk memberikan informasi detail mengenai objek pariwisata yang sedang diamati dari kamera ponsel. Fokus dari skripsi ini adalah untuk menghasilkan model latihan dengan akurasi deteksi objek yang baik. Hasil dari proses latihan model akan dijadikan sebagai model untuk deteksi objek yang ada di sekitar Monas yang akan digunakan untuk melakukan VQA. Dataset yang digunakan dalam penelitian ini adalah gambar Monas beserta objek-objek sekitarnya sebanyak 600 gambar dengan label kelas sebanyak 25 kelas objek. Jaringan yang digunakan untuk melakukan deteksi objek adalah dengan menggunakan YOLO dan RetinaNet, dimana nantinya kedua jaringan ini akan dilakukan komparasi dengan mencari skor akhir dari hasil evaluasi kedua model yang telah dihasilkan. Dengan menggunakan dataset original, pada jaringan YOLO mean average precision (mAP) yang didapatkan dengan rentang nilai confidence level threshold 0,1 sampai 0,9 berkisar antara 60,77% sampai 71,99%, sedangkan untuk jaringan RetinaNet mAP yang didapatkan berkisar antara 72,18% sampai 92,98%. Dengan menggunakan dataset augmentasi, pada jaringan YOLO mAP yang didapatkan berkisar antara 52,51% sampai 93,72%, sedangkan untuk jaringan RetinaNet mAP yang didapatkan berkisar antara 23,8% sampai 56,19%. Untuk skor Area Under Curve (AUC) pada dataset original sebesar 0,99 dan 0,96 pada dataset augmentasi. Berdasarkan hasil eksperimen ini dapat disimpulkan model YOLO dapat mendekripsi lebih baik dibandingkan dengan RetinaNet dan dataset augmentasi dapat menghasilkan deteksi gambar lebih baik dibandingkan dengan dataset original.

.....

Tourism sector has become one of the most potential income for some countries. One of the ways to increase income from tourism sector is to implement information technology so it can attract more tourists to come. The technology that can be implemented is smart tourism. One of the smart tourism implementations for Indonesia tourism, especially for Monumen Nasional (Monas) tourism destination is mobile based Visual Question Answering (VQA) application that can provide detailed information about tourism object from mobile phone camera. Focus of this thesis is to produce training model with good detection accuracy. The result of the model training process will be used as model for object detection model that will be used for doing VQA. Dataset that will be used for this research are 600 pictures containing Monas and 25 surrounding objects called class. The networks that will be used for object detection is using YOLO and RetinaNet, where both of these networks will be compared each other by searching the accuracy from evaluation metric from both networks. By using original dataset, in YOLO network the mean average precision (mAP) score is

between 60.77% to 71.99% with 0.1 to 0.9 confidence level threshold range and in RetinaNet network the mAP score is between 72.18% to 92.98%. By using augmented dataset, in YOLO network the mAP score is between 52.51% to 93.72% and in RetinaNet network the mAP score is between 23.8% to 56.19%. The Area Under Curve (AUC) score for original dataset is 0.99 and 0.96 for augmented dataset using YOLO network. Based on the evaluation result, YOLO can detect objects better than RetinaNet and augmented dataset can produce better detection than original dataset.