

ABSTRACT

DETECTION OF VULGARITY IN ANIME CHARACTERS: IMPLEMENTATION OF DETECTION TRANSFORMER METHOD

Oleh

Amalia Suciati

21102069

Anime currently experiencing its golden era with over 6.000 anime released and dominating 60% of the cartoon broadcasting industry. However, anime content often contains vulgar symbols or graphic elements, like mini clothing and the emphasis on sexually appealing body parts, causing harm such as addiction and overexposure. Vulgar content is not limited to sexual genres but also in general genre like action, adventure, and supernatural. This study focuses on the development of a model as an early-stage effort to prevent vulgar content in anime, utilizing Detection Transformer (DETR) as the primary approach, which adopts the attention mechanism from transformer. Experiments involved two backbone variant, ResNet-50 and ResNet-101, with the addition of dilated convolution to expand the receptive field. The next focus is a comparative analysis with another object detection approach, You Only Look Once (YOLO). Experimental results on DETR indicate two main outcomes: an increase in AP_s indicating improved small object detection, and a decrease in AP_m and AP_l due to the loss of local information and spatial density caused by early-stage dilation. DETR with ResNet-50 backbone and dilated convolution at stage 4 outperformed other configurations, achieving $AP_{50:95}$ 0.491, AP_{50} 0.875, AP_{75} 0.485, AP_s 0.2999, AP_m 0.491, and AP_l 0.518. While YOLOv9t outperformed DETR in all evaluation metrics across all experiments, the results were reversed during video inference DETR demonstrated superior performance in detecting vulgar objects, particularly in the chest and lower hip regions, as well as in identifying small objects, with fewer missed detections compared to the YOLO.

Keywords: anime, Detection Transformer, object detection, Transformer, vulgarity, You Only Look Once