

ABSTRAK

Penerapan metode deteksi objek pada tugas analisis tata letak dokumen (*Document Layout Analysis* atau DLA) masih menghadapi tantangan, terutama ketika diaplikasikan pada lembar jawaban matematika yang tidak berformat. Keterbatasan ketersediaan dataset dan minimnya penelitian pada domain tersebut memperumit proses ekstraksi informasi secara otomatis. Selain itu, penelitian DLA masih terfokus pada peningkatan akurasi dan mengesampingkan efisiensi. Padahal, model deteksi yang kompleks seringkali membutuhkan sumber daya komputasi tinggi dan konsumsi memori yang besar. Hal ini dapat memperlambat sistem saat dipakai untuk proses penilaian massal. Oleh karena itu, penelitian ini mengusulkan sebuah model deteksi objek yang efisien untuk tugas DLA pada lembar jawaban matematika dengan memodifikasi arsitektur YOLOv8. Modifikasi melibatkan penggantian modul *Cross-Stage Partial* (C2f) pada *backbone* YOLOv8 dengan *Channel Attention Module* (CAM). Modifikasi ini membuat model dapat fokus dan selektif terhadap fitur relevan sekaligus meningkatkan efisiensi dalam DLA. Penelitian ini menggunakan dataset berupa lembar jawaban mata pelajaran matematika yang ditulis tangan oleh siswa SMA di Indonesia. Dataset ini terdiri atas 4.686 lembar yang telah dianotasi ke dalam enam jenis elemen: teks matematika, angka, gambar, penanda identitas, penanda jawaban, dan teks salah. Evaluasi performa difokuskan pada metrik *mean Average Precision* (mAP), jumlah parameter, (*Giga Floating Point Operations*) GFLOPs, dan ukuran model. Hasil penelitian menunjukkan peningkatan efisiensi model dengan penurunan ukuran model sebesar 21%, jumlah parameter 21,8%, dan GFLOPs 12,2%. Uji statistik dilakukan dengan metode *Wilcoxon signed-rank test* untuk dua jenis hasil pengukuran, yaitu mAP dan *Optimization Score*. Hasil menunjukkan bahwa mAP model CAM tidak mengalami penurunan yang signifikan dari model *basic*, sementara *Optimization Score* model tersebut mengalami peningkatan signifikan. Penelitian ini berkontribusi pada pengembangan sistem penilaian otomatis yang efisien dan adaptif untuk lembar jawaban matematika tulisan tangan. Integrasi CAM pada YOLOv8 terbukti efektif dalam mengurangi kompleksitas model sambil meningkatkan performa deteksi terkhusus pada kelas gambar yang memiliki fitur visual yang kuat. Akan tetapi, tantangan dalam mendeteksi kelas-kelas lain, seperti kelas teks salah, masih memerlukan penelitian lebih lanjut.

Kata kunci—YOLOv8, *Channel Attention Module* (CAM), Lembar Jawaban, *Document Layout Analysis* (DLA), Deteksi Objek

ABSTRACT

The application of object detection methods in *Document Layout Analysis* (DLA) tasks still faces challenges, especially when applied to unstructured mathematics answer sheets. The limited availability of datasets and the lack of research in this domain complicate the process of automatic information extraction. In addition, DLA research is still focused on improving accuracy while neglecting efficiency. In fact, complex detection models often require high computational resources and large memory consumption. This can slow down the system when used for mass grading processes. Therefore, this study proposes an efficient object detection model for DLA tasks on mathematics answer sheets by modifying the YOLOv8 architecture. The modification involves replacing the *Cross-Stage Partial* (C2f) module in the YOLOv8 backbone with a *Channel Attention Module* (CAM). This modification allows the model to focus and be selective on relevant features while improving efficiency in DLA. This study uses a dataset consisting of handwritten mathematics answer sheets from Indonesian high school students. The dataset consists of 4,686 sheets that have been annotated into six element types: mathematical text, numbers, images, identity markers, answer markers, and incorrect text. Performance evaluation focuses on the metrics of *mean Average Precision* (mAP), number of parameters, GFLOPs, and model size. The results show an increase in model efficiency with a reduction in model size by 21%, number of parameters by 21.8%, and GFLOPs by 12.2%. Statistical testing was conducted using the *Wilcoxon signed-rank test* on two types of measurement results: mAP and *Optimization Score*. The results indicate that the mAP of the CAM model does not experience a significant decrease from the basic model and its *Optimization Score* shows a significant improvement. This research contributes to the development of an efficient and adaptive automatic scoring system for handwritten mathematics answer sheets. The integration of CAM in YOLOv8 has proven effective in reducing model complexity while enhancing detection performance, particularly in the image class, which possesses strong visual features. However, challenges such as detecting other classes still require further research.

Keywords—YOLOv8, *Channel Attention Module* (CAM), Answer Sheet, *Document Layout Analysis* (DLA), Object Detection