

INTISARI

Pewarnaan Otomatis Citra Sketsa Kartun Menggunakan Metode Pix2Pix Conditional Generative Adversarial Network Dengan Augmentasi Data

Oleh

NASHIRA OKSANI ARDINE SANTOSA

20/462188/PA/20160

Proses pewarnaan citra akan lebih efektif bila dapat dilakukan secara otomatis dengan memanfaatkan model *deep learning*. Rodrigues et al. (2022) menyebutkan dataset citra karakter non-humanoid masih sangat terbatas, sehingga menjadi tantangan dalam pengembangan model untuk tugas spesifik tersebut. Pada penelitian ini akan dibangun model pewarnaan citra sketsa kartun non-humanoid secara otomatis menggunakan Pix2Pix dengan menerapkan augmentasi data untuk meningkatkan jumlah dataset. Dataset yang digunakan adalah citra kartun Pokemon, terdiri dari 830 pasang citra *training set* dan 20 pasang citra *test set*. *Training set* akan diaugmentasi menjadi 2490 pasang citra dengan metode transformasi tradisional, seperti rotasi, translasi, *zoom*, *shear*, dan *flip*.

Model Pix2Pix umumnya mempelajari data tanpa melibatkan representasi manusia. Citra output yang dihasilkan cenderung berpaku pada citra target. Sifat ini kurang cocok diimplementasikan pada tugas pewarnaan citra yang membutuhkan unsur seni dan kreativitas. Pada penelitian akan digunakan metrik *Mean Opinion Score* (MOS) sebagai representasi manusia dalam pelatihan model. MOS didapatkan melalui survei kepada responden dan digunakan sebagai penentu bobot *L1 loss* untuk setiap interval pelatihan 20 *epochs*.

Tiga model akan dibandingkan, terdiri dari model tanpa augmentasi milik Adler (2019), model dengan augmentasi, serta model dengan augmentasi dan MOS hasil penelitian. Evaluasi ketiga model dengan metrik MOS dan SSIM menunjukkan augmentasi transformasi tradisional dapat meningkatkan performa pewarnaan model. Model augmentasi mendapat skor MOS dan SSIM sebesar 4.20 dan 67.76%. Model augmentasi dengan MOS mendapat skor MOS dan SSIM sebesar 3.98 dan 63.57%. Model tanpa augmentasi mendapat skor MOS dan SSIM sebesar 1.97 dan 47.92%. Penggunaan MOS pada iterasi pelatihan mendorong model untuk menghasilkan citra output yang lebih bervariasi, tidak hanya berpaku pada citra target.

Kata-kata kunci : pewarnaan citra, citra sketsa, Pix2Pix, augmentasi data.

ABSTRACT

Automatic Cartoon Sketch Image Colorization Using Pix2Pix Conditional Generative Adversarial Network and Data Augmentation

By

NASHIRA OKSANI ARDINE SANTOSA

20/462188/PA/20160

The image colorization process would be more efficient if it could be done automatically using deep learning models. Rodrigues et al. (2022) mentioned that the dataset of non-humanoid character images is still very limited, thus posing a challenge in developing models for such specific tasks. This study aims to build an automatic image colorization model for non-humanoid cartoon sketches using Pix2Pix by applying data augmentation to increase the dataset size. The dataset used consists of 830 pairs of training images and 20 pairs of test images of Pokemon characters. The training set will be augmented to 2490 pairs of images using traditional transformation methods, such as rotation, translation, zoom, shear, and flip.

Pix2Pix models generally learn data without involving human representations. The generated output images tend to closely resemble the target images. This nature is less suitable for implementation in image colorization tasks that require artistic elements and creativity. In this study, the Mean Opinion Score (MOS) metric will be used as a human representation in the model training process. MOS is obtained through surveys conducted with respondents and used as a determinant of the weight of L1 loss for each 20 epochs of training interval.

Three models will be compared, including the non-augmented model from Adler (2019), the model trained with augmentation, and the model trained with augmentation along with MOS. Evaluation of the three models using MOS and SSIM metrics indicates that traditional transformation augmentation can improve the performance of the coloring model. The augmented model achieves MOS and SSIM scores of 4.10 and 67.76%, respectively. The augmentation model with MOS achieves MOS and SSIM scores of 3.98 and 63.57%, respectively. The non-augmented model achieves MOS and SSIM scores of 1.97 and 47.92%, respectively. The utilization of MOS in the training iterations encourages the model to produce more varied output images, not just closely resembling the target images.

Keywords : image colorization, sketch image, Pix2Pix, data augmentation.