Image-to-image translation

Computer Vision II

23/24

Overview

The goal of this project is the use of image-to-image translation models, such as Pix2Pix [1], to generate realistic images from input label maps in diverse application domains. Pix2Pix employs adversarial training to learn a mapping function from an input image domain to an output image domain. This process involves simultaneously training a generator network to produce realistic images and a discriminator network to distinguish between generated and real images.

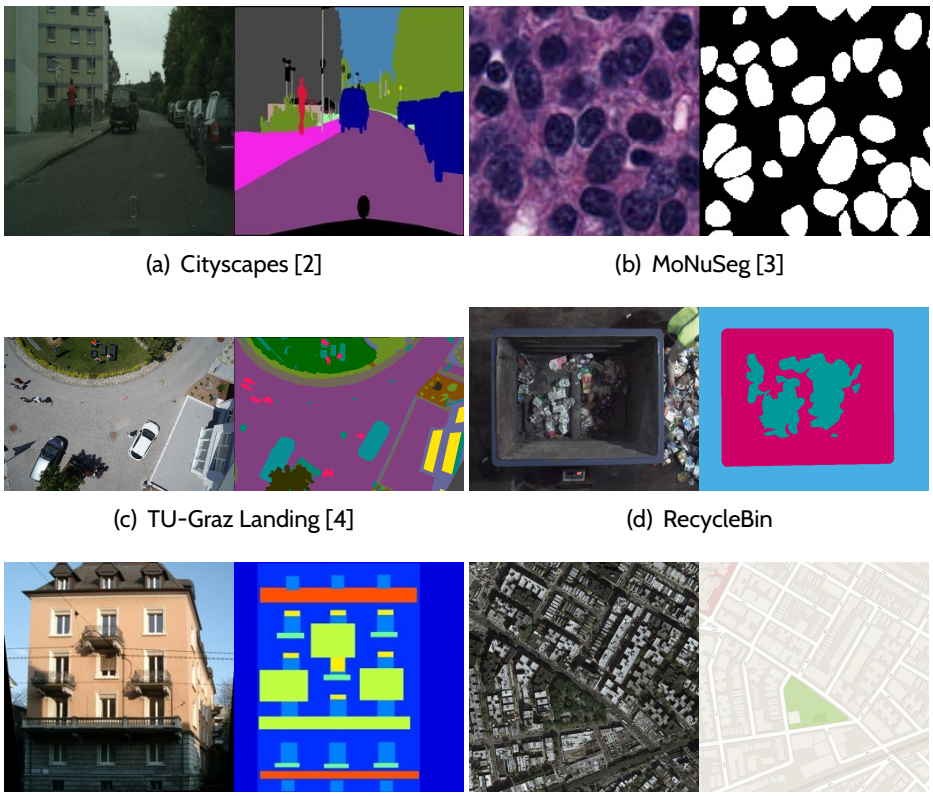
Materials

Six datasets from different application domains are provided for this project[1]. Table 1 summarizes the general information and Figure 1 displays an example of image from each dataset. Each image in the datasets is a paired sample composed of a real image (left side) and its corresponding label map (right side).

Dataset	Images	Size
Cityscapes	500	256×256
MoNuSeg	500	200×200
TU-Graz Landing	400	400×600
RecycleBin	217	480×600
Facades	500	256×256
Maps	500	600×600

Table 1: Summary of general information for each dataset.

[1] Each student/team will be randomly assigned one of the available datasets.



(a) Cityscapes [2] (b) MoNuSeg [3]

(c) TU-Graz Landing [4] (d) RecycleBin

(e) Facades [5] (f) Maps [6]

Figure 1: Samples of paired data from each dataset where the left side displays the real image and the right side its corresponding label map. (a) Semantic labels for urban street scenes. (b) Nuclei cell labels for organ tissues. (c) Aerial semantic labels for safe UAV Landing. (d) Occupancy labels for recycling containers (e) Architectural labels for build facades. (f) Geographic labels for satellite images.

Methods

The objective of this assignment is the development of an image-to-image translation method that, given a label map provides a realistic image.

Task1: Baseline approach

- Design and implement a baseline version of the Pix2Pix model.
- Train the model on the assigned dataset (use training, validation, and testing splits.).
- Compute performance metrics.

Task2: Improvements

- Propose potential improvements or select at least one of the following:
  - Explore different architectures.
  - Explore different loss functions.
  - Use data augmentation.
  - Use transfer learning.
- Train and repeat Task 1
- Compute performance metrics and compare with Task 1

Submission

The due date for this project is **May 10th, 2024**. This project can be solved individually or in pairs (the grades will be individual). Each team should submit before the deadline a compressed file with:

- The code written in `Python`
- A report describing and justifying the proposed approach (a PDF created with an IEEE conference template [2], max. 4 pages). The following is a suggested structure for the report:
  - Title, authors

[2] `https://www.ieee.org/conferences/publishing/templates.html`

- Abstract
- Introduction to the problem and the proposed approach.
- Technical approach
- Results
- Conclusion
- References

The submission will be through the UDC Virtual Campus, or for students without access, via email to *l.ramos@udc.es*. Teams will have to discuss their approach in a Teams meeting with an instructor after the submission.

References

[1] Phillip Isola, Jun-Yan Zhu, Tinghui Zhou, and Alexei A Efros. Image-to-image translation with conditional adversarial networks. *CVPR*, 2017.

[2] Marius Cordts, Mohamed Omran, Sebastian Ramos, Timo Rehfeld, Markus Enzweiler, Rodrigo Benenson, Uwe Franke, Stefan Roth, and Bernt Schiele. The cityscapes dataset for semantic urban scene understanding. In *Proc. of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2016.

[3] Neeraj Kumar, Ruchika Verma, Sanuj Sharma, S. K. Bhargava, Abhishek Vahadane, and Amit Sethi. A dataset and a technique for generalized nuclear segmentation for computational pathology. *IEEE Transactions on Medical Imaging*, 36:1550–1560, 2017.

[4] Safe-UAV-Landing. `https://www.tugraz.at/index.php?id=22387`, 2021. [Accessed 20-March-2024].

[5] Radim Šára Radim Tyleček. Spatial pattern templates for recognition of objects with regular structure. In *Proc. GCPR*, Saarbrucken, Germany, 2013.

[6] Pix2Pix Datasets. `http://efrosgans.eecs.berkeley.edu/pix2pix/datasets/`, 2018. [Accessed 20-March-2024].