Benchmarking Framework for Bad Weather Distortion in Semantic Segmentation in the Context of Autonomous Driving

Iris KREMER (337635), Nicolas d'ARGENLIEU (276507)
CS-503 Project Progress Report 2

I. Proposal Statement Reminder

We chose to develop a **Benchmarking Framework for Bad Weather Distortion in Semantic Segmentation in the Context of Autonomous Driving**. Our goal is to develop an accessible framework used to uniformly assess the performances of any bad weather de-noising model, e.g. de-raining or de-fogging. We plan on creating an evaluation pipeline (*see Milestone 1 report for details*).

The pipeline will apply the **candidate de-noising model** to samples of the dataset. It will then pass the raw and de-noised images through a **SOTA semantic segmentation model**. Finally, a **scoring module** is going to assess the performance of the de-noising model through the relative variation of semantic segmentation results.

II. Milestones Overview

This section summarizes the work achieved up to milestone 2 and details the next steps. Until now, the following has been achieved :

- **Milestone 1**
  - Selection of a sample candidate de-raining model and of the SOTA semantic segmentation model, see [1] and [2].
  - Selection of the dataset images, see [3] and [4].
- **Milestone 2**
  - Test of the de-noising model [1] and SOTA semantic segmentation model [5] on sample images.
  - Research and selection of the most relevant performance metrics to evaluate the Semantinc Segmentation model performance.
  - Creation of the basic modular structure in Docker to run a single model.
  - Tests and exploration of the `Cityscapes` dataset for the normal [3] and the weather-distorted [6] [7] versions.
  - Selection of sub-datasets for the pipeline's tests.

We also outline the actions to reach completion until the project's deadline:

- **Final Milestone**
  - Develop scoring module.
  - Link all aforementioned components into the pipeline.
  - Documentation creation.

For the final version of the framework, the user will be prompted to place his candidate de-noising framework at a *specified location* relative to the code. The code will necessitate to expose the model through some *standardized functions* called by the framework. The final framework should include:

- framework's reduced dataset (images and labels),
- SOTA model code,
- scoring module code,
- pipeline's Dockerfile,
- experiments configuration file (`.yaml`),
- README documentation.

Some **optional developments** have been thought through following the Milestone 2 feedback. However as previously reported, these will depend on time constraints.

III. Milestone Progress

We detail here some aspects of the Milestone 2.

Regarding the scoring metrics, various options have been explored [8].The **Intersection over Union** (IoU) and **F1-score** (Dice Coefficient in the context of image segmentation) were selected against the **pixel accuracy** metric, judged insufficiently indicative as it can return high percentages on poor segmentation results. However these metrics are defined for binary classification. Since more than two classes are predicted, we will evaluate the performance for the specified class against all other classes.

Next, we proceeded to run the selected models. The de-raining model had to be adapted to run on the Windows OS, but worked locally on the provided rain images. Running the model on real-world images forced us to switch to Google Colab. Similarly, the SOTA model experienced local GPU memory issues for the Cityscapes dataset. We thus run the models on the SCITAS cluster from now on.

Then we created our first Dockerfiles. In order to experiment, we wrote one file for each model separately. This enabled us to identify the required environments and packages for our models. However we encountered a major issue with the access to the host's GPU resources. We are currently in the process of troubleshooting the issue.

IV. Discussion & Next Steps

The main problem we encountered until now is **unexpected configuration bugs** related to the use of the

models with sometimes *light* documentations. Without compromising to much our planning, we continue to apply our *essentials-first* approach, i.e. implementing a functional pipeline before designing proper experiments in order to anticipate eventual coding problems.

Another major problem that we ran into is the **access to the machine's resources from the Docker container**. We are still investigating a solution currently with further tests to be made on the SCITAS cluster. However our fallback is to implement a non-dockerised version. This would incur a heavier configuration for the user but would allow us to finish the project.

Finally, the next steps are developing the scoring module, linking the pipeline's elements together, and creating the documentation, as detailed in the **Milestone's Overview**.

REFERENCES

[1] R. Yasarla, V. A. Sindagi, and V. M. Patel, "Syn2real transfer learning for image deraining using gaussian processes," in *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, 2020, pp. 2726–2736.

[2] K. Sun, Y. Zhao, B. Jiang, T. Cheng, B. Xiao, D. Liu, Y. Mu, X. Wang, W. Liu, and J. Wang, "High-resolution representations for labeling pixels and regions," *CoRR*, vol. abs/1904.04514, 2019. [Online]. Available: http://arxiv.org/abs/1904.04514

[3] M. Cordts, M. Omran, S. Ramos, T. Rehfeld, M. Enzweiler, R. Benenson, U. Franke, S. Roth, and B. Schiele, "The cityscapes dataset for semantic urban scene understanding," *CoRR*, vol. abs/1604.01685, 2016. [Online]. Available: http://arxiv.org/abs/1604.01685

[4] F. Yu, H. Chen, X. Wang, W. Xian, Y. Chen, F. Liu, V. Madhavan, and T. Darrell, "Bdd100k: A diverse driving dataset for heterogeneous multitask learning," in *IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, June 2020.

[5] Xie Jingyi, Ke Sun, Jingdong Wang, "High-resolution networks and Segmentation Transformer for Semantic Segmentation," 7 2020. [Online]. Available: https://github.com/HRNet/HRNet-Semantic-Segmentation

[6] Raoul de Charette, J-F. Lalonde, "Rain Rendering for Evaluating and Improving Robustness to Bad Weather," 7 2020. [Online]. Available: https://github.com/cv-rits/rain-rendering/

[7] Christos Sakaridis, "Fog Simulation on Real Scenes for Partially Synthetic Foggy Data," 10 2018. [Online]. Available: https://github.com/sakaridis/fog_simulation-SFSU_synthetic

[8] Ekin Tiu, "Metrics to Evaluate your Semantic Segmentation Model," 2019. [Online]. Available: https://towardsdatascience.com/metrics-to-evaluate-your-semantic-segmentation-model-6bcb99639aa2