Homework 5 : Overview
Single Image Super Resolution with GANs

Start Here

- **Collaboration policy:**
  - You are expected to comply with the University Policy on Academic Integrity and Plagiarism.
  - You are allowed to talk with / work with other students on homework assignments
  - You can share ideas but not code, you must submit your own code. All submitted code will be compared against all code submitted this semester and in previous semesters using MOSS.

1 Introduction

In this homework, we will work on the problem of Single Image Super Resolution, where the aim is to enhance the resolution of an input low resolution image. You will learn to implement and train your own GAN from scratch. You're expected to form teams of TWO people for this homework/project. Note that this is supposed to be a project-style homework. In 11785 projects, we usually see students struggling to find the right problem statement and workflow for the project. In this homework, we aim to streamline the learning process for a project by setting well defined milestone. But we expect you to explore and try out new ideas just as you would in a project.

The homework is divided into two phases. In the first phase you are expected to work individually and it will be worth 25% of your grade. In this phase, you're expected to learn the workings of GANs and Autoencoders so that you're ready to dive into building your own models. You're gradually expected learn about Autoencoders, GANs and Super Resolution. Your knowledge will be tested through quizzes on Canvas. The first phase ends with a part-1 style assignment where you have to code some common elements of any GAN architecture.

In the second phase, you have to implement your own architectures for Single Image Super Resolution from scratch. You will be working in teams of two for this phase (forming teams of two is a strict requirement, teams of 3 are NOT allowed). You will have to implement an Autoencoder and two baselines GANs architectures for comparison.

Part of your grade will be determined by your performance in a CodaLab (very similar to Kaggle) competition. The remaining portion of your grade will be determined by a report presented by your team where we expect you to perform rigorous analysis of your results, explain any new approaches tried by you and why they did or did not work. We also expect you to compare the three baseline architectures provided to you. A detailed template for the report will be provided to you for this.

2 Phase-1: Groundwork (25%)

The first phase of the homework is designed incrementally advance your knowledge in autoencoders and GANs. Especially because AEs and GANs haven't been covered in class yet, we expect you to go through previous year recitations, lectures and slides to work on the quizzes. The quizzes are not created to trick you and will be medium level difficulty, designed to be simpler than the weekly quizzes. The aim of the quizzes is to make sure that you have a firm grasp of the basics. The deadlines have been created to make sure you are not left behind in the project. No extensions are allowed phase-1 deadlines.

- Deadline 1 - April 7, 2021 : AutoEncoders and GANs Quiz (5%)

- Deadline 2 - April 14, 2021 : Super Resolution and GANs Quiz (5%)
- Deadline 3 - April 18, 2021 : Implementing components of a GAN (coding) (15%)

The quizzes will be released on the Monday before the deadline and you will have exactly 72 hours to work on it. The coding assignment for part-1 will be released shortly and will need to be submitted on Autolab.

For the first quiz, we expect you to have a basic understanding of how Autoencoders and GANs work. For reading up on Autoencoders, we recommend starting with Bhiksha's lecture from last semester. For GANs, we highly recommend watching last semester's recitation on GANs before watching the lectures (especially the first part of the recitation). Last year's GANs recitation was created to cater for an audience that has no background in GANs and we tried to keep it as math-free as possible. After that, we recommend watching last semester's lectures on GANs for a more complete experience with GANs.

For the Super Resolution quiz, you are expected to read through the papers of SRGAN (Ledig et al. [2017]) and ESRGAN (Wang et al. [2018]). Make sure you understand the different loss functions used for Super Resolution and the different evaluation metrics used to evaluate models. Also, we expect you to have gone through GANs lectures, so expect some questions about GANs as well.

The coding assignment for Phase -1 will be released on Autolab. Please download the assignment from there when released. We expect you to submit your assignments to Autolab.

3 Phase - 2: Research and Compete (75%)

The second phase of the homework is designed to be done in teams of two. Phase - 2 will be launched on April 15, 2021. You are expected to have formed teams by then. Teams of more than two are not allowed. The second phase deadlines will be very similar to 11785 deadlines and will be announced alongside 11785 final project deadlines.

With a good grasp of GANs, you're expected to implement your own GAN architectures. For baselines, you're expected to implement your own versions of autoencoders, SRGAN and ESRGAN architectures. You are expected to have a good understanding of the different loss functions used for GANS and Super Resolutions and pros and cons of different evaluation metrics for super resolution.

The CodaLab competition will judge the performance of your GAN architecture based on PSNR and SSIM metrics. We will provide you with an evaluation script to test your models against those metrics. The CodaLab competition will be worth 25% of your grade and we will release grade cut offs closer to the deadline. You are only expected to submit your best GAN-based architecture to the CodaLab competition, though you must include both SRGAN and ESRGAN comparison in your report.

The report will be worth 50% of your grade. You will be judged on your understanding of GANs and the Super Resolution problem statement. A detailed comparison between the three baselines models - autoencoders, SRGANs and ESRGANs is expected. We also expect you to come up with your own modifications over these three baselines architectures and will be on the lookout for any interesting ideas that you experimented with. We want to hear about all different architectures you tried, whether they worked or not. We expect you to present a nice ablation about the different methods tried.

A detailed template for the final report will be released closer to the deadline where we'll outline the expectations of the report.

3.1 Dataset

We will be using the DIV2K (Agustsson and Timofte [2017]) dataset for Phase-2. We will focus on x4 upsclaing. That means all models trained, either for the CodaLab competition or for the final report submission must be trained only on the DIV2K dataset for x4 upscaling. Here's the link to access the dataset : `https://data.vision.ee.ethz.ch/cvl/DIV2K/`.

We will provide you with the test set for the competition.

3.2 Individual Contributions

We will keep track of individual contributions for phase-2 to make sure one person is not bearing an unfair amount of burden for HW 5. The policies for checking individual contributions will be very similar to 11785 project.

4 Timeline

- **March 28 : HW5 Phase-1 Begins**
- April 7 : HW5 Quiz 1 (AE and GANs) Due on Canvas (5%)
- April 7: Team Formation Deadline
- April 7: Coding Assignment resleased
- April 14: HW5 Quiz 2 (SR and GANs) Due on Canvas (5%)
- **April 15 : HW5 Phase-2 Begins**
- April 18: HW5 Coding Assignment Due on Autolab (15%)
- May 5 (estimated): CodaLab Competition Closes (25%)
- May 6 (estimated): HW5 Report Submission Deadline (50%)

5 Conclusion

This homework is created to give you a directed pathway towards learning and implementing GANs. This is a project style homework. We expect you to come up with your own ideas to surpass performances achieved by the baselines architectures for super resolution, just as you would do in a project. So research and read papers, come up with your own ideas and Experiment! Not all your ideas are expected to work, but you are expected to try. As we said before, we'll be on the lookout for things you tried on top of the baseline, irrespective of whether they work or not.

Also, start early. Do not wait for the CodaLab competition to start implementing your baseline architectures. Begin as soon as you can. As always, feel free to ask any questions you have on Piazza.

References

Eirikur Agustsson and Radu Timofte. Ntire 2017 challenge on single image super-resolution: Dataset and study. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops, pages 126–135, 2017.

Christian Ledig, Lucas Theis, Ferenc Huszár, Jose Caballero, Andrew Cunningham, Alejandro Acosta, Andrew Aitken, Alykhan Tejani, Johannes Totz, Zehan Wang, et al. Photo-realistic single image super-resolution using a generative adversarial network. In Proceedings of the IEEE conference on computer vision and pattern recognition, pages 4681–4690, 2017.

Xintao Wang, Ke Yu, Shixiang Wu, Jinjin Gu, Yihao Liu, Chao Dong, Yu Qiao, and Chen Change Loy. Esrgan: Enhanced super-resolution generative adversarial networks. In Proceedings of the European Conference on Computer Vision (ECCV) Workshops, pages 0–0, 2018.