TASK 3

Open-ended (graded)

You are in the group Deepgain consisting of auarnold (auarnold@student.ethz.ch (mailto://auarnold@student.ethz.ch)), habaumann (habaumann@student.ethz.ch (mailto://habaumann@student.ethz.ch)) and lscaglioni (lscaglioni@student.ethz.ch (mailto://lscaglioni@student.ethz.ch)).

1. READ THE TASK DESCRIPTION | 2. SUBMIT SOLUTIONS | 3. HAND IN FINAL SOLUTION

1. TASK DESCRIPTION

In this task, you will implement an image inpainting model that reconstructs missing regions in images.

We provide you with a dataset of MNIST digit images, where a central patch has been masked (blacked out). Your task is to predict the complete image, including the missing region. Sample masked images are shown below.

Your model should be able to reconstruct the complete images from their masked versions, as shown in these examples:

DATA DESCRIPTION

Download handout (/task3/download)

In the handout for this project, you will find the the following files:

- **train_data.npz** - this is the dataset that you will train on
- **test_data.npz** - this is the dataset for which you have to submit your predictions
- **template_solution.py** - a template file that will guide you through the implementation of the solution

- **sample.npz** - a template file of how your submission should look like

During training, you will corrupt the images by masking the center of the images. Then you will train your model to predict the complete image from the masked version. The images are 28x28 grayscale images. Data are corrupted by replacing the center 8x8 pixels with black pixels. In python this can be done by using the following code:

```
# Given an image of dimensions (28, 28)
# we set the center 8x8 pixels to black, i.e. mask them
image[10:18, 10:18] = 0
```

See the provided code for more details on how to load the data.

template_solution.py provides a starting template structure for how you can solve the task, by filling in the TODOs in the skeleton code. It is not mandatory to use this solution template but it is recommended since it should make getting started on the task easier. The template solution uses the PyTorch (https://pytorch.org/) deep learning framework. Although not mandatory, e advise you to use PyTorch for this task and on the following links you can get started with PyTorch:

- 60 Minute Blitz (https://pytorch.org/tutorials/beginner/deep_learning_60min_blitz.html)
- PyTorch with Examples (https://pytorch.org/tutorials/beginner/pytorch_with_examples.html)
- Tensors (https://pytorch.org/tutorials/beginner/former_torchies/tensor_tutorial.html)

The computational demands for this task should be moderate. Modern laptops with sufficient RAM (~8 GB) should be able to handle the task. If you are on Ubuntu and you find yourself with insufficient RAM, you can consider increasing swap memory. You can find a guide on how to do that here (https://www.digitalocean.com/community/tutorials/how-to-add-swap-space-on-ubuntu-22-04).

TASK DETAILS

Your task is the following:

- Implement a model that can reconstruct MNIST images where the center has been masked (replaced with black pixels)
- Train your model using the provided dataset
- Generate predictions for the test set images
- Submit your predictions

IMPORTANT NOTES

- Your submission should be a file containing the predicted complete images for the test set.
- You are not allowed to use any pre-trained models or additional data for this task.
- You may not train on the test data.

SUBMISSION FORMAT

Your submission should be a binary numpy file (.npz) containing the predicted complete images for the test set. The template solution provides code to properly format and save your predictions.

EVALUATION

Your submission will be evaluated by computing the Mean Squared Error (MSE) between your predicted images and the ground truth complete images:

$$\text{MSE} = \frac{1}{N \times H \times W \times C} \sum_{i=1}^{N} \sum_{c=1}^{C} \sum_{h=1}^{H} \sum_{w=1}^{W} (I_{i,c,h,w} - \hat{I}_{i,c,h,w})^2$$

Where:

- N is the number of test images
- C is the number of channels in the images, in our case 1, as we are using grayscale images
- H and W are the height and width of the images
- $I_{i,c,h,w}$ is the pixel value at position (h, w) in the ground truth image I
- $\hat{I}_{i,c,h,w}$ is the pixel value at position (h, w) in your predicted image $\hat{I}$

For the loss we are considering only the masked pixels, i.e. the center 8x8 pixels of the image. The MSE is computed only on these pixels.

Lower MSE indicates better reconstruction quality.

GRADING

This is a pass — fail project and you will not receive a grade for this task. You need to pass 3 / 4 projects in the IML course to be eligible for taking the exam. Your algorithm will make predictions on a held out test set. For each of your submissions, a score is computed.

When handing in the task, you need to select which of your submissions will get graded and provide a public link to a maximum one minute long video explaining your approach. Make sure that we can access the link till the first of October. Every student has to submit an individual 1 minute long video. This has to be done **individually by each member** of the team. For generating the video link, please use the ETH polybox (https://polybox.ethz.ch/index.php/login/) service, upload your video there and generate a shareable link (see this documentation (https://polyboxdoc.ethz.ch/share-with-people-outside-of-eth/) for more detail). Submissions that are not handed-in, do not have a video that we can access till the first of October or have a video exceeding the one minute threshold will not be graded.

We will compare your selected submission to our public baseline, which is visible on the public leaderboard. **You pass this project if your submitted solution beats our public baseline.** At the end of the semester, we will award one team with a certificate and prize for their performance on this task (excluding task 0). The selection criteria are based on the team's performance on our public and private leaderboards and the creativity of their solution. The prize will be disclosed at the end of the semester. The private leaderboards are based on a separate test score on an undisclosed test set. You only receive feedback about your performance on the public part in the form of the public score, while the private leaderboard remains secret. The purpose of this division is to prevent overfitting to the public score. Your model should generalize well to the private part of the test set. The creativity of your solution will be evaluated by our TAs based on your video submission.

⚠ Make sure that you properly hand in the task, otherwise you will fail this task.

PLAGIARISM

The use of open-source libraries is allowed and encouraged. However, we do not allow copying the work of other groups / students outside the group (including work produced by students in previous versions of this course). Publishing project solutions online is not allowed and use of solutions from previous years in any capacity is considered plagiarism. Among the code and the reports, including those of previous years, we search for similar solutions / reports in order to

detect plagiarism. Although not strictly forbidden, we discourage the use of Github Copilot or similar code/language generation tools for writing code. We expect that if such tools are used, this is clearly stated in the video submission explaining the solution. While it will have no effect on your grade or if a solution passes or fails, it may affect the awarding of prizes for best solutions. We discourage these tools because we feel that the best way to understand the material is to write the code yourself referring to just the lecture material, source papers and documentation of any libraries used. For the purposes of disclosing what generative AI tools you used to write code, we don't need you to disclose using e.g. basic code autocompletion such as those used in the default setup of Sublime Text 3. If we find strong evidence for plagiarism, we reserve the right to let the respective students or the entire group fail in the IML 2025 course and take further disciplinary actions. By submitting the solution, you agree to abide by the plagiarism guidelines of IML 2025.

FREQUENTLY ASKED QUESTIONS

⦿ WHICH PROGRAMMING LANGUAGE AM I SUPPOSED TO USE? WHAT TOOLS AM I ALLOWED TO USE?

You are free to choose any programming language and use any software library. However, **we strongly encourage you to use Python**. You can use publicly available code, but you should specify the source as a comment in your code.

⦿ WHAT TO DO IF I CAN'T RUN THE CODE/SETUP AN ENVIRONMENT ON MY PC?

If you are having trouble running your solution locally, consider using the ETH Euler cluster to run your solution. Please follow the Euler guide (/static/euler-guide.md). The setup time of using the cluster means that this option is only worth doing if you really can't run your solution locally.

⦿ AM I ALLOWED TO USE MODELS THAT WERE NOT TAUGHT IN THE CLASS?

Yes. Nevertheless, the baseline was designed to be solvable based on the material taught in the class up to the second week of each task.

⦿ IN WHAT FORMAT SHOULD I SUBMIT THE CODE?

You can submit it as a single file (main.py, etc.; you can compress multiple files into a .zip) having max. size of 1 MB. If you submit a zip, please make sure to name your main file as *main.py* (possibly with other extension corresponding to your chosen programming language).

⦿ WILL YOU CHECK / RUN MY CODE?

We will check your code and compare it with other submissions. We also reserve the right to run your code. Please make sure that your code is runnable and your predictions are reproducible (fix the random seeds, etc.). Provide a readme if necessary (e.g., for installing additional libraries).

⦿ SHOULD I INCLUDE THE DATA IN THE SUBMISSION?

No. You can assume the data will be available under the path that you specify in your code.

⦿ CAN YOU HELP ME SOLVE THE TASK? CAN YOU GIVE ME A HINT?

As the tasks are a graded (pass/fail) part of the class, **we cannot help you solve them**. However, feel free to ask general questions about the course material during or after the exercise sessions.

⦿ CAN YOU GIVE ME A DEADLINE EXTENSION?

⚠ We do not grant any deadline extensions!

⦿ CAN I POST ON MOODLE AS SOON AS I HAVE A QUESTION?

This is highly discouraged. Remember that collaboration with other teams is prohibited. Instead,

- Read the details of the task thoroughly.
- Review the frequently asked questions.
- If there is another team that solved the task, spend more time thinking.
- Discuss it with your team-mates.

⦿ WHEN WILL I RECEIVE THE PRIVATE SCORES? AND THE PROJECT GRADES?

We will publish all grades before the exam the latest.