Master's Degree in Artificial Intelligence, Pattern Recognition and Digital Imaging

Practical Session
Stable Diffusion Fine-Tuning Through Low-Rank Adaptation (LoRA)

Tools and Applications of Artificial Intelligence
Departament de Sistemes Informàtics y Computació
Universitat Politècnica de València

2023-2024

Abstract

This practical session aims to teach students how to fine-tune *Stable Diffusion XL* so it learns to generate a person or a certain style. For simplicity's sake, in the rest of this document we will focus on learning to generate a given person, but the same procedure can be applied to learning to generate a style. The session will be conducted using *Google Colab*.

Requirements

Prior to starting the session you will need:

- A *Google* account.
- 10 to 20 pictures of the subject person.
- A *Huggin Face*'s write token. You can obtain it at `https://huggingface.co/settings/tokens`.
- An *ngrok*'s token. You can obtain it at `https://dashboard.ngrok.com/get-started/your-authtoken`.

1 LoRA training

In this section, we will see how to fine-tune *Stable Diffusion XL*.

1.1 Notebook preparation

Prior to starting the training, we need to set up *Google Colab*. We will be using the following notebook[1], which is a custom version of *Huggin Face*'s `AutoTrain Advanced`[2].

The notebook contains a single cell of code which installs all requirements and starts up a session in *ngrok*[3], which we will be using for our user interface. Prior to running the code, we need to fill the *huggingface_token* and *ngrok_token* with the respective tokens that we had prepared in advanced. Then, we just need to run the notebook. Once the installation is complete, we need to click the url[4] that appears in the output.

1.2 Setting up the training

The previous link will redirect us to the *ngrok*'s session which, as mentioned, contains `AutoTrain`'s user interface (see Fig. 1).

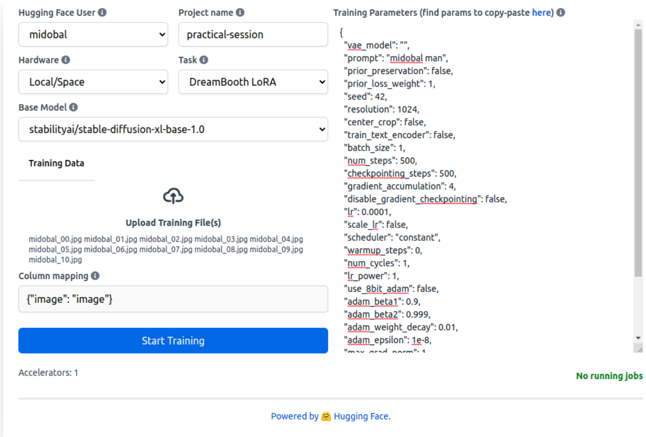

Figure 1: Example of *AutoTrain*'s user interface for training the model.

In this interface, we need to make the following changes:

1. On the **Task** drop menu, we need to select *DreamBoothLoRA*.
2. On the **Base Model** drop menu, we need to select *stabilityai/stable-diffusion-xl-base-1.0*.
3. Click on **Upload Training File(s)**. This will open a window which will allows us to select all the training images (which we should had prepared previous to starting the session).
4. From the training parameters (right side of the menu), we need to edit the **prompt**. This will be the set of words which will use to tell *Stable Diffusion XL* to generate our subject. It is recommended to use a unique identifier which *Stable Diffusion XL* has not seen during its training (e.g., our *UPV*'s username) plus an adjective that describing the person (woman, man, child, older woman, etc).

[1] `https://colab.research.google.com/github/huggingface/autotrain-advanced/blob/main/colabs/AutoTrain.ipynb`.

[2] `https://huggingface.co/docs/autotrain/index`.

[3] `https://ngrok.com/`.

[4] e.g., *https://b59c-34-125-231-223.ngrok-free.app/*.

Additionally, we can edit the *project name* and/or any other training parameter. When we are done, we just need to click on *Start Training*. This will pop up a message indicating that we will be charged for this action. Since we are running **AutoTrain** on a local hardware (*Google Colab*), we can ignore the message and click on *Yes, I'm sure*, at which point the training will start (the training id will be indicated in the bottom-right corner).

1.3 Model training

To see the model training we need to go back to *Google Colab*. The output of the execution will show us the log of the training process. Once is completed (with the default parameters, the process can take around 1 hour), we need to download the model. To do so, we need to click on the folder icon on the left menu and search for a file with the extension *.safetensors* (see Fig. 2). After that, we can stop the execution and close the notebook.

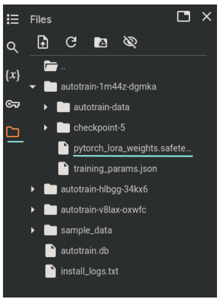

Figure 2: Example of how to locate the model.

2 Image generation

Once we have our new LoRA "model", we need to switch to the following notebook[5], which runs an instance of *Stable Diffusion XL* on Fooocus[6]. On this notebook, we run the cell and wait until the installing process is finished. Then, prior to clicking to the url that has been generated, we need to the file system (same menu from which we downloaded the model on the last step), look for the *loras* folder (inside *Fooocus/models*), right-click on it and select *Upload*. This will open up a window in which we need to select the model we downloaded at the previous step.

Once the model is at its correct location, we can click on the link[7] which will lead us to *Stable Diffusion XL*'s user interface we have just installed (see Fig. 3).

Now that we are on *Stable Diffusion XL*'s interface, the first thing we need to do is select our model. To do so, we need to click on *Advanced*, select the *Model* option on the right menu and select our model under *LoRA1* drop menu[8] Then, we need to adjust the weight of the model. The closer it is to 1, the more *Stable Diffusion XL* will take the

[5] `https://colab.research.google.com/github/lllyasviel/Fooocus/blob/main/fooocus_colab.ipynb`.
[6] `https://github.com/lllyasviel/Fooocus`.
[7] e.g., *https://054bc36b255597196a.gradio.live/*.
[8] If the model does not appear, we need to click on *Refresh All Files*.

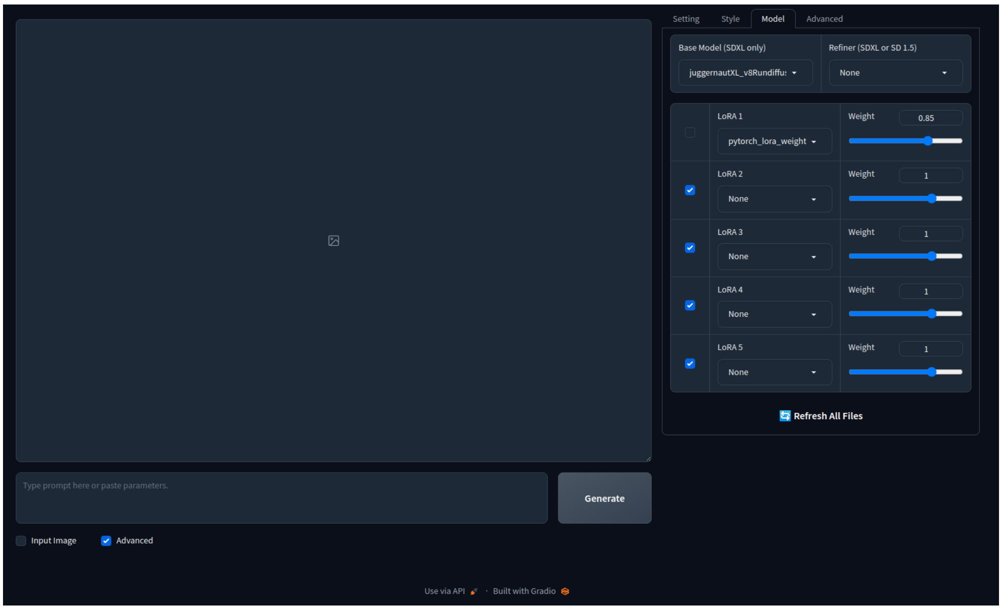

Figure 3: *Stable Diffusion XL*'s user interface.

model into account. A value of 0.85 should be reasonable for our purposes. Finally, if we select the *Style* option (from the same menu), we will find different clickable options. These are different styles that *Fooocus* has implemented in order to make the prompt construction more easily (making it transparent to the user).

Once we have the model set up, it's time to start writing prompts (which include the keyword used during the training process) and playing with the different styles! When you obtain an image that you like, remember to click on it and save it. Otherwise it will be deleted when the *Colab*'s session ends.