How to use DeepLabCut

Now that you have installed DeepLabCut, you can start using it. Before we can create a new project we need to work on the video files from the behavioral test. So far we record them using Ethovision. Within this setup we have a "Delay Time" of 2 seconds. Meaning that the tracking of Ethovision only starts 2 seconds after it first recognized the mouse. However, the raw video that Ethovision saves, starts at the moment you hit the "start trial" button. Therefore, we need to cut the videos to the same length. You can do this manually in any video editing software or with a script. For this tutorial we will use an already trimmed video. You can download it at:

https://drive.google.com/file/d/1woqrpY1AQmCbgOlpM7yQYoemQMC_I2xF/view?usp=sharing

If you want to cut the video yourself with the script please follow these steps, otherwise proceed to "Creating a DLC-Project".

1. You need to install ffmepg first:
   a. For windows users please follow these instructions:
      https://www.geeksforgeeks.org/how-to-install-ffmpeg-on-windows/

   b. For Mac users install ffmpeg through Homebrew: Homebrew is a command-line package manager, which is similar to apt-get on popular Linux distributions. In order to use it, you need to install brew first, if you haven't already. Open the terminal and enter the following command:

```
/bin/bash -c "$(curl -fsSL https://raw.githubusercontent.com/Homebrew/install/master/install.sh)"
```

Follow the on-screen instructions. This will take a few minutes while it's installing the necessary developer tools for macOS. Then, run:

```
brew install ffmpeg
```

2. Copy the python script "video_cutting_script.py" to the folder where your videos are located.
3. Open terminal or anaconda command prompt and go to that folder using the command:
   cd ** full path of the folder**
4. Type python "video_cutting_script.py. The script will now ask you for the directory. Type it or copy the path and do not forget the "/" at the end. (**Note:** For Windows users use "/" instead of "\")
5. The script will now ask you how long the trial is. Type it in.
6. The script will now cut your videos according to your trial length. The videos will be saved in the same directory with an additional "_trimmed.mp4"

Creating a DLC-Project

Now you successfully installed DeepLabCut you want to start using it. For this you will need to create a new DLC Project and afterwards label your Data.

1. Create a folder for the tutorial files
2. Open the terminal/Anaconda prompt and navigate to that folder (cd **path_to_the_folder**)
3. Type in `git clone https://github.com/LW1989/DLC_tutorial.git`
4. Activate the DLC environment by typing:
   a. MacOS: `conda activate nameoftheenv` (eg. DEEPLABCUT)
   b. Windows: `activate nameoftheenv` (eg. DEEPLABCUT)
5. Type in: `jupyter notebook`
6. Open Demo_Notebook_1 and follow the instructions there

Creating a training data set

Now that you have your labeled data, you can start with creating the training dataset and the actual training. First you need to transfer your Project folder to the High-Performance Cluster (HPC) of the Uni. For that you need to send me (Lutz) your university account name so I can ask the Data Science Center of the university to give you access to the HPC. For connecting to the HPC you need to download putty at https://putty.org/ (Windows). Mac users can use the terminal. For the actual transfer of data you have two options on windows (Mac user, please check below):

1. Pscp → Only if you are comfortable with using the command prompt. It is included when you install putty. This guide will continue with option 2. For basic guide read: http://the.earth.li/~sgtatham/putty/0.60/htmldoc/Chapter5.html

2. Download and install WinSCP from https://winscp.net/eng/download.php

Open WinSCP and as the Server IP put in: 10.103.43.104 Port: 22
Enter your account name (without @uni-bremen.de) and password, save it and connect to the server (Note: It only works within the university network or if you use the university VPN). Look at the following photo to get an idea how it should look like.
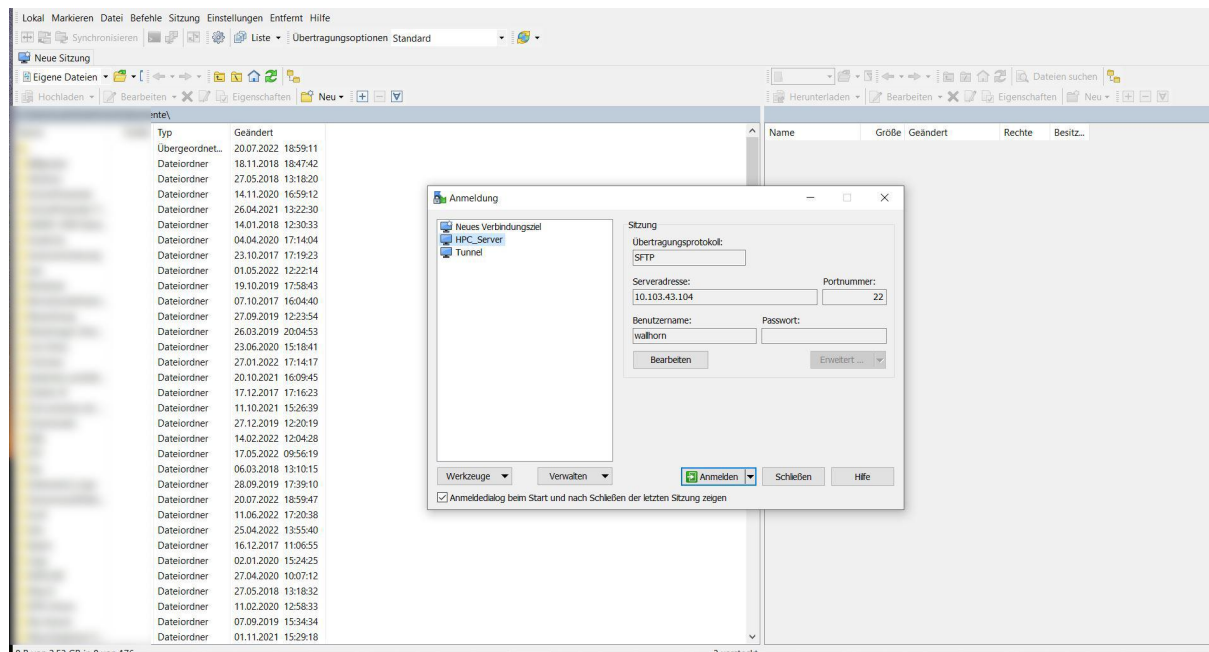
After you successfully connect to the server you can see your home directory. You can edit everything within winscp.

Mac users can use cyberduck. You can download this at https://cyberduck.io/
You can connect to the server similar to WinSCP:
1. Click the "+" to add a new connection
2. Change the type of the connection from FTP to SFTP (SSH)
3. Enter the following IP address: 10.103.43.104 Port: 22
4. Enter your university account and press enter
5. Now you can connect to the server

Now you can transfer your project folder to the HPC into a folder of your choosing.
Important: After you transferred the project folder you have to change the path in your config file, because the project was created on your local machine. You can do that with WinSCP. Just double-click the config and change the following parameters:

- project_path:
  → the new path should look like that: /home/wallhorn/OpenField-2022
- video_sets:
  → the new path should look like that: /home/wallhorn/OpenField-2022 /videos/

Next you need to connect to the server and install DeepLabCut:

1. Open Putty/Terminal and connect to the server (IP: 10.103.43.104)
2. Type in the following commands:
   a. module load python
   b. module load anaconda
   c. module load cuda

3. Create a folder where you want to install DLC: `mkdir **name_of_the_folder**`
4. Type the following command: `cd **Full path of the folder**`
5. Type: `git clone https://github.com/DeepLabCut/DeepLabCut.git`
6. Now, in Terminal go to the subfolder of the cloned repo. For example, If you cloned the repo onto your Desktop, the command may look like:
   `/DeepLabCut/conda-environments`

Tipp 1: To list every file and folder in the current directory type just type: `ls`
Tipp 2: To show the path of the current directory your are in type: `pwd`

7. Now, in the terminal run: `conda env create -f DEEPLABCUT.yaml`

8. You can now use this environment from anywhere on your computer (i.e. no need to go back into the conda- folder). Just enter your environment by running the following command within the terminal: `conda activate nameoftheenv` (eg. DEEPLABCUT)

9. Now you should see ("nameofenv") on the left of your terminal screen, i.e. (DEEPLABCUT) YourName-MacBook. NOTE: no need to run pip install deeplabcut, as it is already installed!!! :)

10. Run the following command (fixes some loading issue with a tensorflow library: `export LD_LIBRARY_PATH="$CONDA_PREFIX/lib"`

11. Now you are set to use DeepLabCut on the server if your are familiar with using python scripts, the terminal and Slurm (see https://slurm.schedmd.com/). Otherwise, keep following the tutorial.

For inexperienced users, we created additional jupyter notebooks to follow along. The nexts steps of the DeepLabCut pipeline need to be performed on the HPC server. Therefore, you need a tunnel connection in order to run jupyter notebooks on the server and interact with them in your browser. On windows please follow these steps.

1. Open PuTTY
2. Login into the server (IP: 10.103.43.104) with your university account
3. Close session by typing exit into the terminal
4. Save session
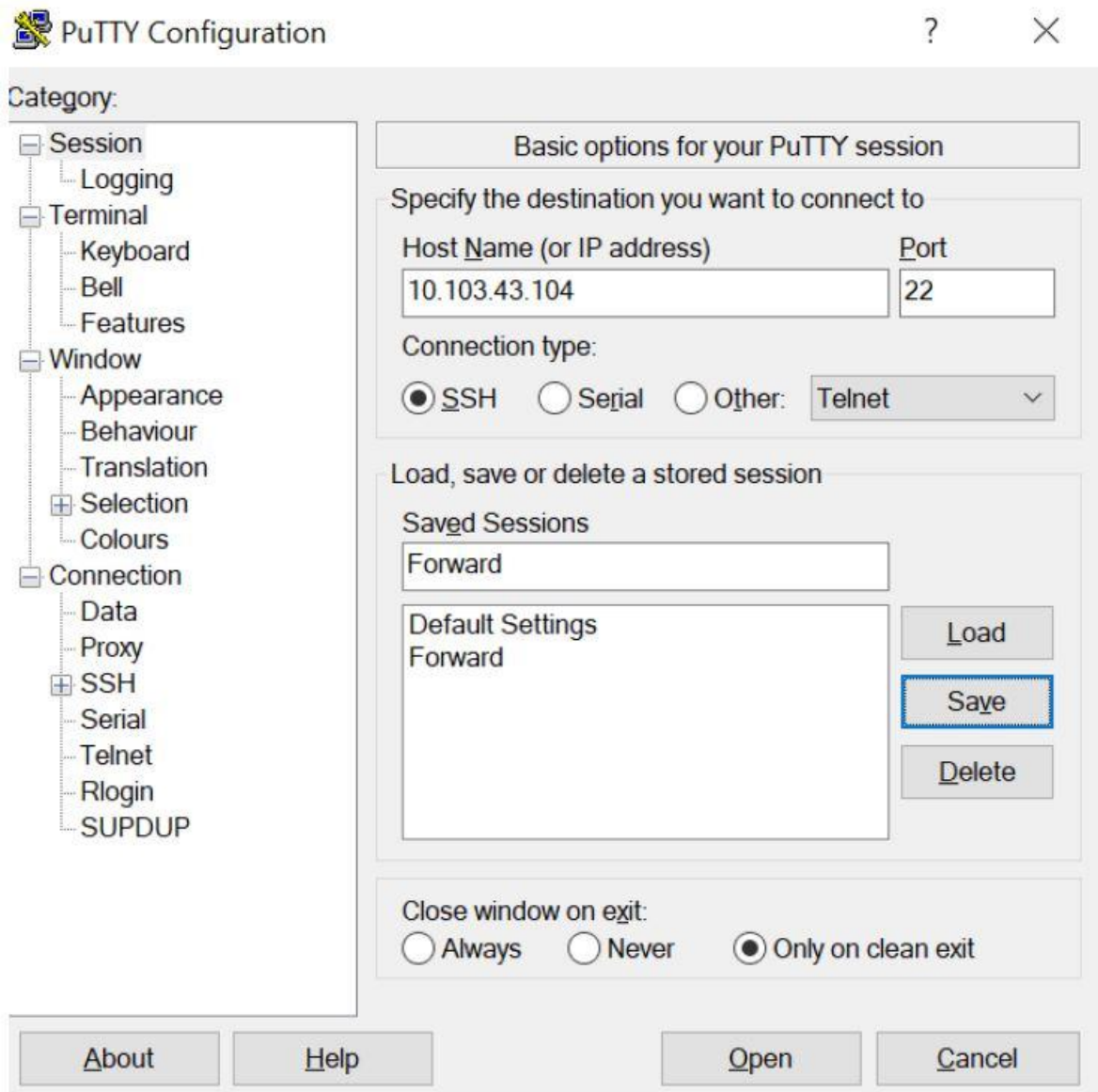
1. Go to connection → SSG → Tunnels
2. Enter 1238 at source port and 127.0.0.1:1238 at destination and click "add"
3. Go back to session and save
4. Open a session and type the following commands in that order:
   a. module load python
   b. module load anaconda
   c. module load cuda
   d. conda activate DEEPLABCUT
   e. `export LD_LIBRARY_PATH="$CONDA_PREFIX/lib"`
   f. `jupyter-notebook --no-browser --port=1238`
5. Copy one of the URLS at the bottom into your local browser
6. Now navigate to the Demo_Notebook_2 to create the training dataset

Mac users please follow these steps:
1. connect to the server via the terminal

2. enter the following command: `jupyter notebook --no-browser --port=8889`
3. Open a new terminal tab/window and type in the following command:

```
ssh -N -f -L localhost:8888:localhost:8889 university_account_name@hpc.dsc.uni-bremen.de
```

4. Go to your browser and enter "localhost:8888". Now, jupyter notebook should open
5. You might be asked to enter a token. Go to the terminal tab/window where you connected to the server. You should see something like that:
http://localhost:8889/?token=**a _lot_of_letters_and_numbers**
The part after "token=" is what you need
6. Now navigate to the Demo_Notebook_2 to create the training dataset

Training your DLC-Model on the High-Performance Cluster

Now that you created the training data set you are ready to finally train your network. For this you need a lot of GPU power. Therefore, all these parts are down on the HPC of the University. This server is managed by a framework called "Slurm" (https://slurm.schedmd.com/documentation.html). In simplified terms: It manages the different requests for processing time on the CPU as well as the GPU cluster. You submit a job that is put into a queue. Afterwards you can log out and check later that day or the next day if your job has already finished. You can interact with it through the terminal on the server or through the jupyter notebook. For the latter option you will need to install slurm-magic (https://github.com/NERSC/slurm-magic). To do that connect to the server via PuTTY/Terminal and within your base environment type:

```
pip install git+https://github.com/NERSC/slurm-magic.git
```

Note: For whatever reason (I do not have any idea why) the following steps need to be done within your base environment and NOT within the DEEPLABCUT environment. If you are still in your DEEPLABCUT environment type: `conda activate base`

Before you can execute this batch script you need to transfer the py_scripts folder from the tutorial directory to the server, the same way you did with the project folder (i.e. with WinSCP/Cyberduck)

The submitting of jobs is done by executing batch scripts. We will do that from within the jupyter notebook. these will look something like that:

```bash
#!/bin/bash
#
#SBATCH --job-name=DLC_training
#SBATCH --nodes=1
#SBATCH --ntasks=1
#SBATCH --cpus-per-task=48
#SBATCH --output=/home/wallhorn/DLC_Projects/job_reports/%j.out


module load cuda
module load anaconda

source /hpc/opt/apps/Anaconda/3-2021.11/etc/profile.d/conda.sh
conda init bash

conda activate DEEPLABCUT

export LD_LIBRARY_PATH="$CONDA_PREFIX/lib"
cd /home/wallhorn/DLC_Projects/py_scripts

srun python3 training_script.py
```

This script basically does what you did so far by hand:
- load the required modules
- activates the DEEPLABCUT environment
- Navigates to the directory of the py_scripts
- runs the "deeplabcut_training_scripts

After that you can run the squeue command to check if your job was successfully submitted.

Now you can proceed to Demo_Notebook_3