Users Guide

# Functionality

This application is aimed to allow the users to:

1. Predict the location of alpha-helixes given a cryo-EM density maps of proteins, using a pre trained model.
2. Train a model to predict the location of alpha helixes given a labeled database of cryo-EM density maps.



Want to insert a picture from your files or add a shape or text box? You got it! On the Insert tab of the ribbon, just tap the option you need.

**User Interface**

The user interface is a command line interface.

The interface is as follows: python helix\_predictor\_UI *command*

Supporting the following commands:

1. --help: prints the interface of the supported commands.
2. -train output\_net\_path dataset\_path [evaluation\_results\_output]

output\_net\_path: The path in which the trained network will be saved.

dataset\_path: A path to a directory that contains both the protein density maps and their labeled helix maps. \*The expected format is specified under the following section.

evaluation\_results\_output: An optional argument, a path for the results of the evaluation.

1. -predict net\_path protein\_mrc\_path output\_mrc\_path

protein\_mrc\_path: a path of the protein density map.

output\_mrc\_path: a path to where the prediction will be saved at.

**In depth**

Training

The purpose of the training stage is to train a Convolutional Neural Network to predict the location of alpha helixes given a density map of a protein.

The training process requires a labeled database that consists of pairs of a protein density map and its labeled helix mask. The dataset is a directory, the pairs’ format should be “protein.mrc” for the protein density map and “protein\_helix.mrc” for its corresponding helix map.

A protein density map and its helix mask should be of the same dimensions.

To achieve the best possible performance, the model should be trained on similar proteins to the protein whose helices the user wants to predict.

Prediction

The application predicts the location of alpha helices given a protein density map and a trained neural network.

The input is protein density map in mrc file format and a neural network trained using the application in pt file format.

The output is a predicted helix mask in mrc file format.

## Pre-trained models

The application provides 6 pre-trained models:

1. Super family 4Å
2. Super family 6Å
3. All alpha class 4Å
4. All alpha class 6Å
5. Alpha and Beta (a/b) 4Å
6. Alpha and Beta (a/b) 6Å

All models were trained with grid size of 1Å