In this task, you will make decisions on food taste similarity based on images and human judgements.

We provide you with a dataset of images of 10000 dishes, a sample of which is shown below.



We also provide you with a set of triplets (A, B, C) representing human annotations: the human annotator judged that the taste of dish A is more similar to the taste of dish B than to the taste of dish C. A sample of such triplets is shown below.



You task is to predict for unseen triplets (A, B, C) whether dish A is more similar in taste to B or C.

**Data description**

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

* **dataset** - folder containing the dish images
* **train\_triplets.txt** - contains the training triplets. The entries of each triplet denote file names. For example, the triplet "00723 00478 02630" denotes that the dish in image "00723.jpg" is more similar in taste to the dish in image "00478.jpg" than to the dish in image "02630.jpg" according to a human annotator.
* **test\_triplets.txt** - the triplets you should make predictions for
* **sample.txt** - a sample submission file
* **template\_solution.py** - a template file that will guide you through the implementation of the solution
* **template\_solution.ipynb** - a template file in jupyter notebook format that will guide you through the implementation of the solution

You are free to use either jupyter notebook or the .py template file.

**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. We 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 are higher than the previous tasks. Modern laptops with sufficient RAM (~16-20 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). If you still find yourself requiring more computational power, you can use the Euler cluster. Though you should also consider the setup time when deciding if you need it. Below you can find a guide on how to set up a GPU-enabled environment on Euler.

Your task is the following: for each triplet (A, B, C) in **test\_triplets.txt** you should predict 0 or 1 as follows:

* 1 if the dish in image A.jpg is closer in taste to the dish in image B.jpg than to the dish in C.jpg
* 0 if the dish in image A.jpg is closer in taste to the dish in image C.jpg than to the dish in B.jpg

**Important**

* The order for your predictions for the triplets should agree with the order of the triplets in test\_triplets.txt and each prediction should be on a new line.
* You may (and probably should) use pretrained vision models for this task.

Using pretrained models is very common in current deep learning practice and a useful tool to have in your practical toolbox. PyTorch provides a number of pretrained models that you can use. You can find a list of the available models and how to use them [here](https://pytorch.org/vision/stable/models.html). To learn more about transfer learning and how to use pretrained models you can check the following link: [Transfer Learning Tutorial](https://pytorch.org/tutorials/beginner/transfer_learning_tutorial.html).

A screenshot of a computer

Description automatically generated

A close up of a box

Description automatically generated