

# Readme

## Introduction

The HOGM\_CoMatch program finds the common pattern between two datasets and outputs the association matrix between them. Let  $D_1$  be a  $n \times m_1$  matrix [ $m_1$  features,  $n$  samples] and  $D_2$  be a  $n \times m_2$  matrix [ $m_2$  features,  $n$  samples]. HOGM\_CoMatch returns the association matrix  $X_{HOGM}$  of size  $m_2 \times m_1$ .

The following instructions are for running the program in iOS. However, it can be easily adapted for Windows as well.

The code in HOGM\_CoMatch is a modified version of the HOGMMNC program introduced in the paper:

Chen, H. Peng, G. Han, H. Cai, and J. Cai, "HOGMMNC: a higher order graph matching with multiple network constraints model for gene–drug regulatory modules identification," *Bioinformatics*, vol. 35, no. 4, pp. 602–610, Feb. 2019, doi: 10.1093/bioinformatics/bty662.

## Input

It requires 2 files as input:

1. D1.csv
2. D2.csv

The name and format of your files should match those of the provided sample files. That is the names of the files need to be "D1" and "D2". Each file should have both row and column names and should be of ".csv" type.

## Running the program

Use the following steps to run the code.

1. Download the folder HOGM\_CoMatch.
2. Edit files *mexComputeFeature\_angle\_distance\_sample.cpp* and *mexComputeFeature\_angle\_distance\_sample.h* files in the HOGM\_CoMatch/mexSource/ folder.
  - a. Open file *mexComputeFeature\_angle\_distance\_sample.h* in Matlab.
  - b. Edit line 29 within function computeFeatureSimple() with the number of samples in your D1 or D2 data.
  - c. Open file *mexComputeFeature\_angle\_distance\_sample.cpp* in Matlab.
  - d. Edit lines 18 and 19 within function mexFunction() with the number of samples in your D1 or D2 data.
  - e. Run *mex mexComputeFeature\_angle\_distance\_sample.cpp* in Matlab.

- f. Place the newly created *mexComputeFeature\_angle\_distance\_sample.mexmaci64* file in HOGM\_CoMatch/mex/ folder.
3. Copy your D1.csv and D2.csv to the HOGM\_CoMatch/input/ folder.
4. Run the file *do\_our\_experiment.m* in Matlab.
5. Your results will be generated in the HOGM\_CoMatch/result/ folder.

You can customize the program as needed. The following are some simple examples:

1. Change the name of the directory where results will be stored in line 3 of *do\_our\_experiment.m*.
2. Instead of using “HOGM\_CoMatch/input/” folder, you can specify the full path of your own folder containing your input files in line 3 of *do\_our\_experiment.m*.

## Output

The program generates 4 output files:

1. D1.jpg: Visual representation of D1.csv
2. D2.jpg: Visual representation of D2.csv
3. X\_HOGM.csv: Association matrix between D1 and D2
4. X\_HOGM.jpg: Visual representation of the transpose of the matrix in X\_HOGM.csv

We use the output X\_HOGM.csv in subsequent steps of the CoMatch algorithm.