

BAHÇEŞEHİR UNIVERSITY
FACULTY OF ENGINEERING AND NATURAL SCIENCES
DEPARTMENT OF SOFTWARE ENGINEERING
SEN4931 – SPECIAL TOPICS IN SE I
TERM PROJECT

Goal: To gain theoretical knowledge and applicational aspects of exploratory data analysis tools and algorithms by processing and analysing considerably big and multidimensional data.

Data: Hyperspectral image data. You should visit the following link and download two datasets and their corresponding ground truths. These names of the datasets are:

- Indian Pines
- Pavia Centre

[http://www.ehu.eus/ccwintco/index.php/Hyperspectral Remote Sensing Scenes](http://www.ehu.eus/ccwintco/index.php/Hyperspectral_Remote_Sensing_Scenes)

Instructions:

1. Acquire, provide and present necessary information about Hyperspectral Images (HSI) and their properties. This task is essential for performing further steps. Perform the following tasks for both datasets mentioned above using MATLAB.
2. Normalize all spectral signals.
3. Select 3 suitable bands at different wavelengths from the dataset, create an appropriate false-color image and show it.
4. Visualize the ground truth data. Assign a different color to each label to distinguish the classes properly in the image.
5. Present the number of elements belonging to each class using histogram plot. Exclude the 0 valued entities since they indicate background pixels.
6. Apply PCA to the HSI dataset and reduce the band number to 3, 5, 10, 20 and 25, respectively.
7. Generate another false color image for the dataset by using the 3-band version obtained using PCA.
8. Create similarity matrices and heat maps for each reduced data obtained in step 6 and determine which class is closer to which class.

9. Classification phase (Support Vector Machines). Use LIBSVM package to implement the following tasks:
 - a. Classify the dataset and calculate the overall accuracy (OA) and the average accuracy (AA).
 - b. Use radial basis function as SVM kernel.
 - c. Perform classification for training ratios 5%, 10% and 20%.
 - d. Sketch a graph to present the OA and AA values vs training ratio.
 - e. Find OA and AA values by classifying the reduced data in the 6th step at 10% training ratio. Sketch the corresponding accuracy values vs band number on a single graph.
10. Prepare a suitable report. Present your implementation details and results. Justify your observations.
11. You may benefit from the following references:
 - a. <https://ieeexplore.ieee.org/document/9258418>
 - b. <https://ieeexplore.ieee.org/abstract/document/6521421>
 - c. <https://ieeexplore.ieee.org/abstract/document/9609461>
 - d. <https://link.springer.com/article/10.1007/s10462-021-10018-y>
 - e. Other useful references are encouraged and welcome

Grading:

- 70% for fulfilling the instructions
- 30% for the preparation and presentation of the report

Due date:

- 23.01.2022, at 22:00

Submission:

- Through *itslearning* portal

Bahçeşehir University