Assignment
Course: Supervised Classification of Hyperspectral Images (HIP2)

Gabriel Dauphin

March 15, 2024

Specifications related to the proposal

We consider the hyperspectral image Pine and more specifically the classification of land cover as being soybean or not. The work consists in proposing a classification technique. This classification technique is to be precisely described in a report. The test set includes half of the pixels randomly chosen in the image, the training set and the validation set use the remaining samples. It would be appreciated that the proposed technique complies with the following constraints:

- It should take into account several bandwidths.
- It should take into account spatial context, to some extent.
- It should be derived from the minimization of a loss function (or be somewhat consistent with).
- The loss function should include the likelihood of a mixture of two Gaussians of equal standard deviation.

The numerical complexity of the proposal should be reasonable (it should be possible to do the training using an implementation in less than one hour). It should avoid overfitting and to this end, it is advised not to use more parameters than a twentieth of the number of samples available for each class in the training set.

Specifications related to the description of the proposal

The proposal should be clear enough for someone not having taken the course, to be able to implement it. It should be described using pseudocodes interacting together. The computations involved in these pseudocodes should be described in formulas. The notations need not be the same as in the lecture, however they must be precisely defined.

The report mentions how it complies with the constraints detailed previously. It should indicate the number of parameters to be learned with the training set. It should include the theoretical computations that lead to the formulas referred to in the pseudocodes.

The proposal may include tests described with pseudocodes that would give some relevant insights as to the correctness of the implementation and the appropriateness of the optimization technique selected in its ability to diminish the loss function. These tests could also measure the correctness of the probabilistic assumptions.

It is completely out of the scope of this assignment to implement the proposed method.

Expectations related to the report

It would be appreciated that the report complies with the following expectations listed by order of decreasing importance.

1. Formulas even if they happen to be wrong or contradictory, they should primarily make sense (for instance, a matrix should not be added to a scalar, a parameter used as a counter in a summation should not appear outside this summation).
2. Whereas in the literature, a single technique may actually be used in different ways in different proposals, the report should be unambiguous as to what exactly is being proposed.

3. What is claimed in the report in terms of complying with the constraints should be correct.

4. The computations included in the report should be correct (it is advised to make some simple numerical tests).

5. The technique used to fuse the different sources of information (spatial context, different bandwidths...) should be beneficial, in that performances should not be increased when one source of information is lacking.