4 Quantitative Comparison of Unsupervised Extraction Methods

In the following, all experiments carried out with unsupervised labeling approaches and synthetic data and their results are described.

4.1 Evaluation Metrics

For benchmarking the different labeling methods the question arises how to compare the output of different extraction methods to the ground truth of the synthetic data. Therefore, we first have to understand the predicted output of both PCA-ICA and CNMF-E method. The methods are predicting weighted masks for all predicted neurons, and their calcium traces. The masks are weighted with respect to how much each pixel contributes to the signal of the cell. For each neuron a threshold of one third multiplied by the maximum weight of the neuron is set. This is then applied to the mask such that only those pixel that hold a weight in the upper two third part margin remain. This is applied in the same matter to both output methods such that we end up with "$\frac{2}{3}$ binary masks". Now for each trace the task of extracting the spiking events remains. Although a peak finder labeling tool as shown in Section 2.1.1 is implemented in the PCA-ICA method, in case of much easier to interpret synthetic traces it might not achieve much better results than a simple hypertuned peak finder. Due to time-constraints and technical constraints with respect to the peak finder labeling tool the labeling of spiking events of the PCA-ICA detected traces is skipped. Instead, the function `find_peaks()` of the `scipy` library is used for both the PCA-ICA traces and the CNMF-E method. Additionally, a clearer comparison of the predictions of both methods is possible. Finally, the extracted post-processed results are in the form of binary masks and extracted events for each mask.

A common metric for segmentation tasks is Intersection over Union (IoU) which is defined as the area of overlap divided by the area of union. A graphical interpretation of the IoU metric can be seen in Figure 4.1. To match each neuron of the ground truth to a detected neuron of each of the to be tested extraction method, we loop over all possible detection and with a threshold of 0.5 match the detected neuron with the highest IoU score. The threshold of 0.5 is applied such that a predicted neuron ROI is not falsely matched with a GT neuron ROI. The predicted and GT neuron's ROI's are divided in true positive (tp), false positive (fp) and false negative (fn) which are used to determine the ROI F1-Score which is defined as the following:

$$F_1 - Score = \frac{tp}{tp + \frac{1}{2}(fp + fn)} \tag{1}$$

After a predicted neuron was matched successfully, the detected events are matched with a given tolerance of + / - 5 frames. The GT and predicted

events are analogously divided into tp, fp and fn and the event F1-Score is determined following equation 1. The events of not detected neuron shapes are not considered.
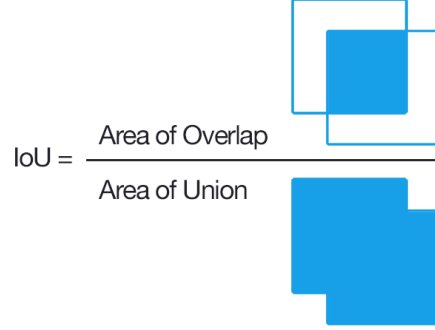

$$IoU = \frac{\text{Area of Overlap}}{\text{Area of Union}}$$

Figure 4.1: Graphical Explanation of Intersection over Union. Figure reproduced from [24].

4.2 PCA-ICA Method

First, the PCA-ICA method is examined. Therefore, the synthetic datasets described in Chapter 3 are given as input to the unsupervised extraction method as shown in figure 4.2. The extracted cellmaps with the GT and predicted ROI's for both datasets are shown in Figure 4.3. As one can see in a close up shown in Figure 4.4 many of the coordinate shapes of the PCA-ICA method are very small or have uneven edges. This is due to the fact that extracted $\frac{2}{3}$ binary masks have artifacts in the form of missing pixel values because of the weighted mask operation. Since some pixels do not have a high weight even though they are located in the center of the cell, this pixel is set to 0. When loaded in the GUI labeling tool described in the work preceding this one [17], the effect is taken care of by `OpenCV` morphological operations such that the shape is closed. Note, that while this does occur on real data as well the problem is not as prominent as on synthetic data. This has to be taken into account when considering the ROI scores shown in Table 2.