**CSE499A.10 - Research Paper Review**

**Tanvir Ahmed – 1410982042**

**CEREALS –** **Cost-Effective REgion-based Active Learning for Semantic Segmentation**

*Radek Mackowiak, Philip Lenz, Omair Ghori, Ferran Diego, Oliver Lange, Carsten Rother*

**Overview:**

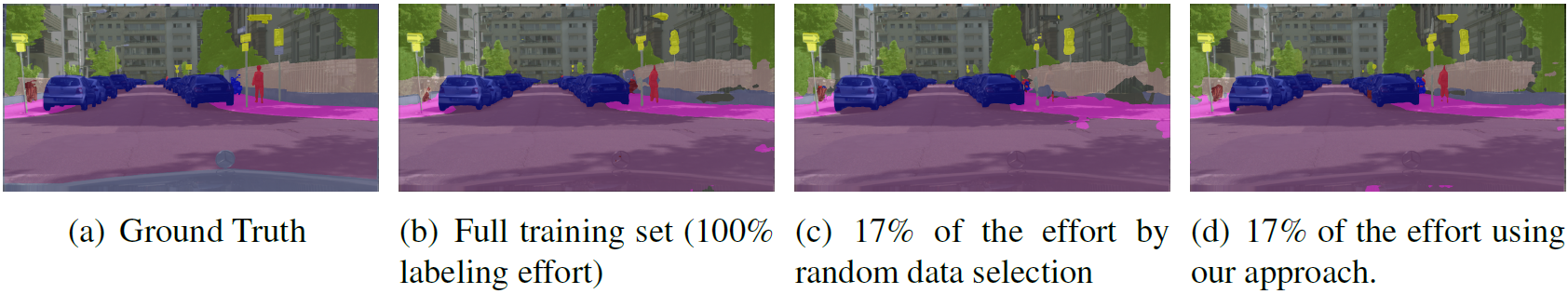
Semantic Image segmentation are trained in a supervised fashion using a large corpus (a collection of written texts) of fully labeled training images. However, gathering such a corpus is expensive, due to human annotation effort, in contrast to gathering unlabeled data.

The paper proposed an active learning-based strategy, called CEREALS, in which a human only has to hand-label a few, automatically selected, regions within an unlabeled image corpus.

The proposed framework reduces the labeling effort by

(i) Utilizing spatial estimates about annotation costs inferred from a learned cost prediction CNN and (ii) By focusing on image regions promising high information content and low annotation costs in a global context.

The performance of CEREALS is demonstrated on Cityscapes, where it was able to reduce the annotation effort to 17%, while keeping 95% of the mean Intersection over Union (mIoU) of a model that was trained with the fully annotated training set of Cityscapes.



**Method:**

They consider a pool-based AL scenario running in batch-mode. In such a setting a large unlabeled pool of data exists from which a small, randomly sampled subset, called the seed set, is initially extracted and labeled by an oracle.

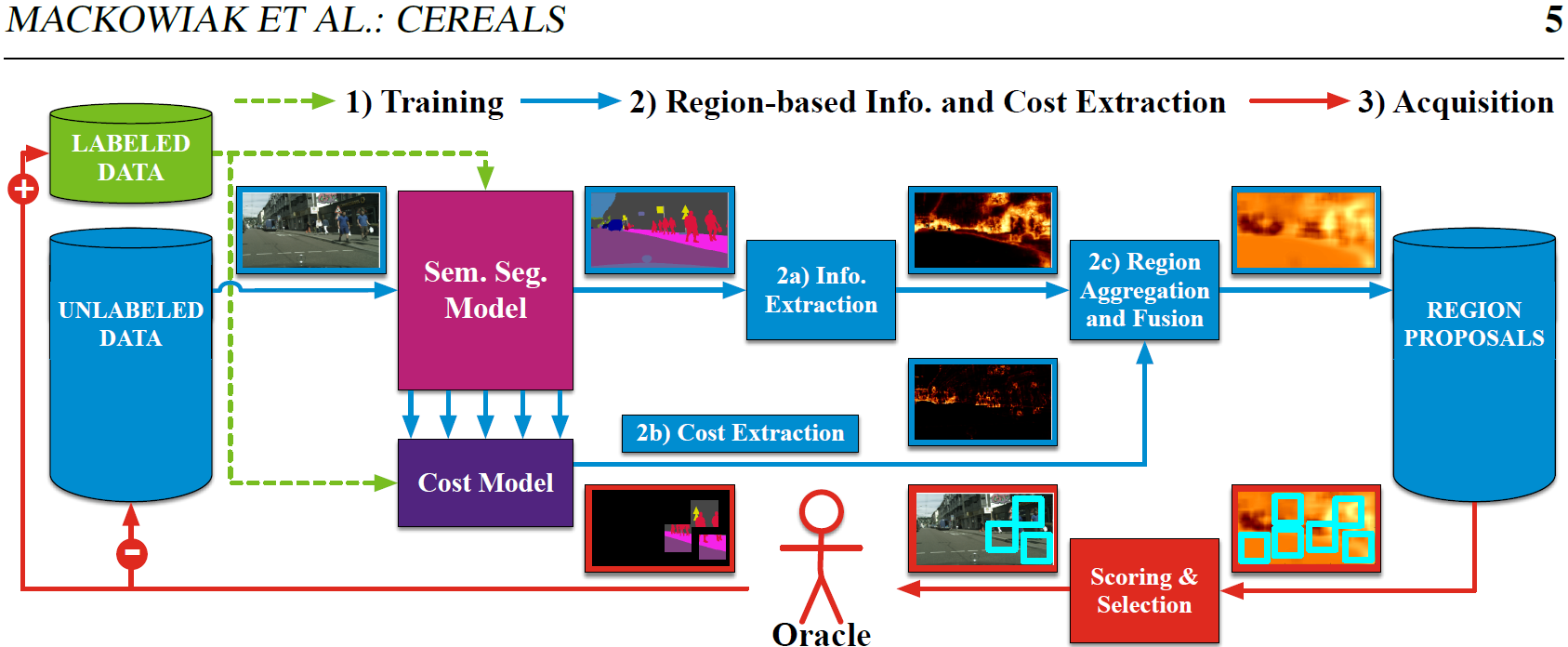
Using this seed set the algorithm works as follows:

First, a model is trained on the currently labeled pool.

Secondly, some measure of information on each individual unlabeled sample is being computed.

Thirdly, an acquisition function is applied. A subset of a pre-specified number of elements maximizing the acquisition function is annotated by an oracle. It is then added to the labeled pool.

The process is repeated until either a desired performance or labeling budget is reached. the stopping criterion is satisfied whenever the unlabeled pool becomes exhausted which is indicated by no further improvements after several acquisition steps.



**Conclusion:**

Their proposed method for cost effective active learning for semantic segmentation tailored to fully convolutional neural networks. They demonstrated their framework’s performance on Cityscapes, a highly diverse high definition dataset consisting of images of urban scenes captured in the wild. they showed that combining information content and cost estimates is a powerful approach for cost-effectively building new training datasets from scratch. With only 17% of the effort measured by the number of clicks which were executed for annotating the Cityscapes training set, it was able to achieve 95% of the full training set’s performance.