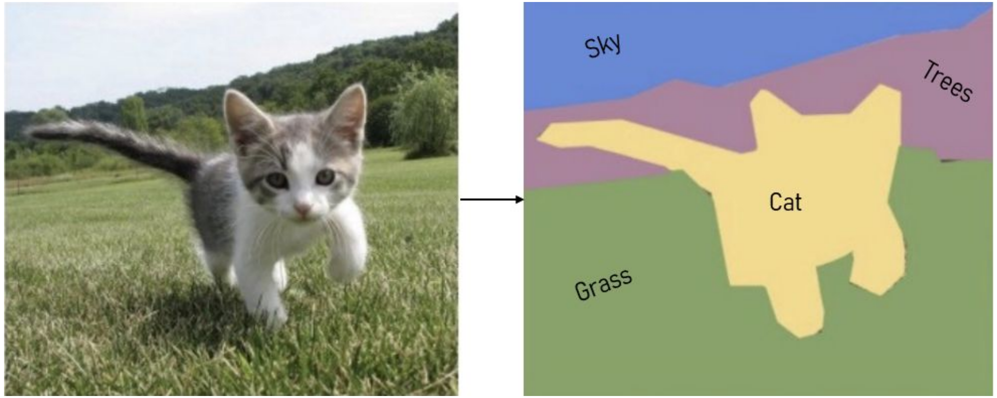
Active Learning in Semantic Image Segmentation

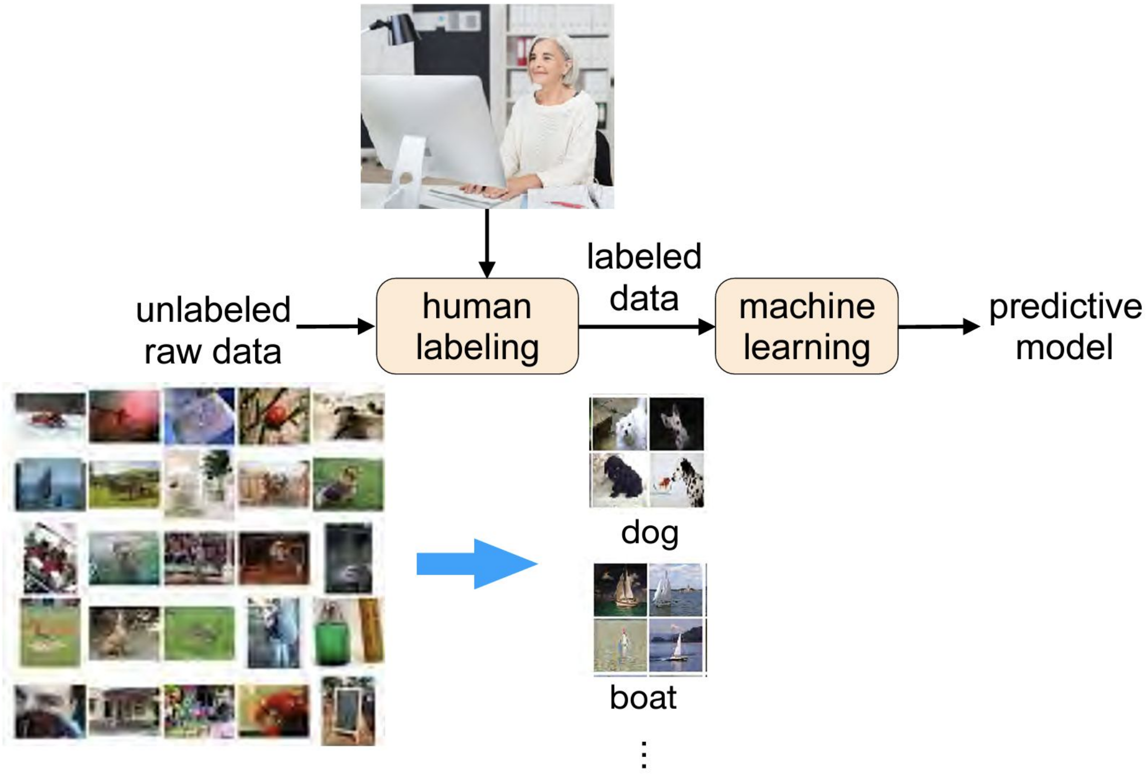
MLRG 2022 Summer
Helen Zhang



Outline

- Introduction and Motivation
- AL Query Strategies
- Applications
  - Revisiting Superpixels for Active Learning in Semantic Segmentation with Realistic Annotation Costs
  - All you need are a few pixels: semantic segmentation with PIXELPICK

Conventional (passive) Machine Learning



Motivation

In many scenarios we will actually have access to a lot of data, but it will be infeasible to annotate everything.

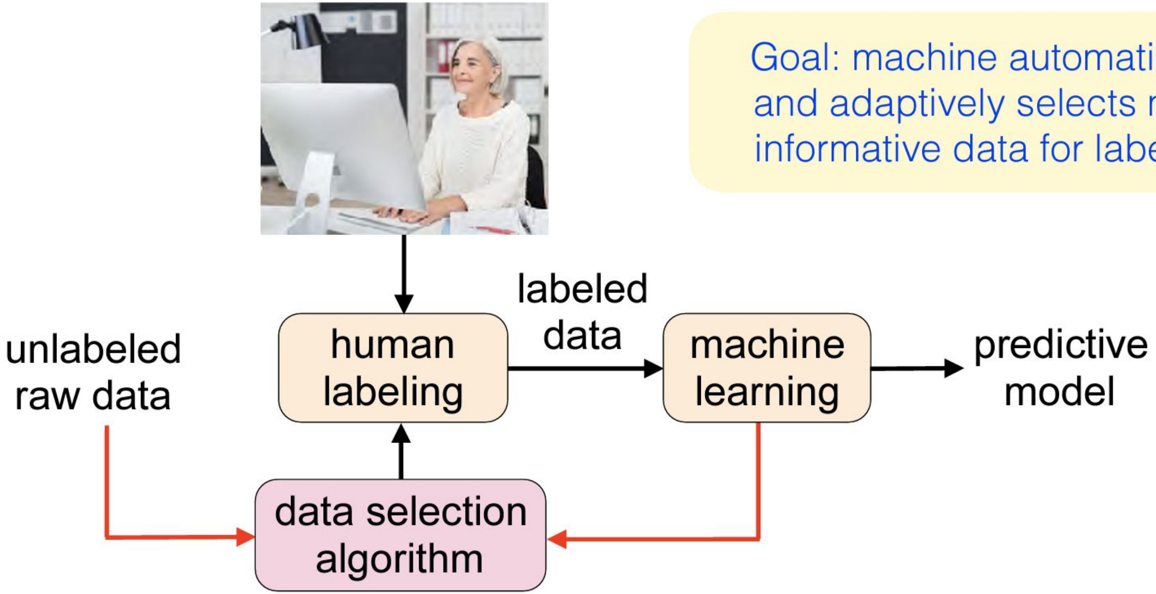
Can we train machines with less labeled data and less human supervision?

- *Semi-supervised Learning*. Exploit the unannotated data to get better feature representations and improve the algorithms learned on the annotated data.
- *Active Learning*. Choose the data that is going to be annotated.
- *Reinforcement Learning*.

Active Machine Learning

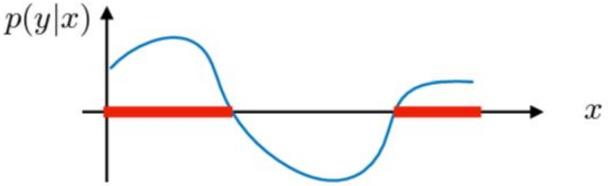
(also called “query learning,” or sometimes “optimal experimental design” in the statistics literature)

Goal: machine automatically and adaptively selects most informative data for labeling

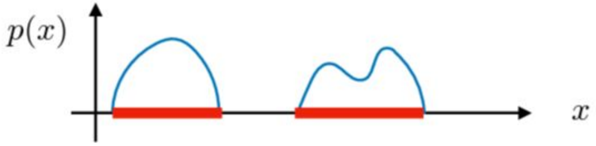


What and Where Information

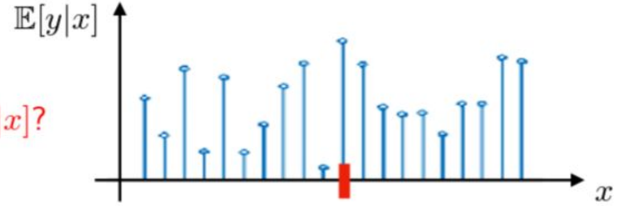
Density estimation: What is $p(y|x)$?
Classification: Where is $p(y|x) > 0$?



Density estimation: What is $p(x)$?
Clustering: Where is $p(x) > \epsilon$?

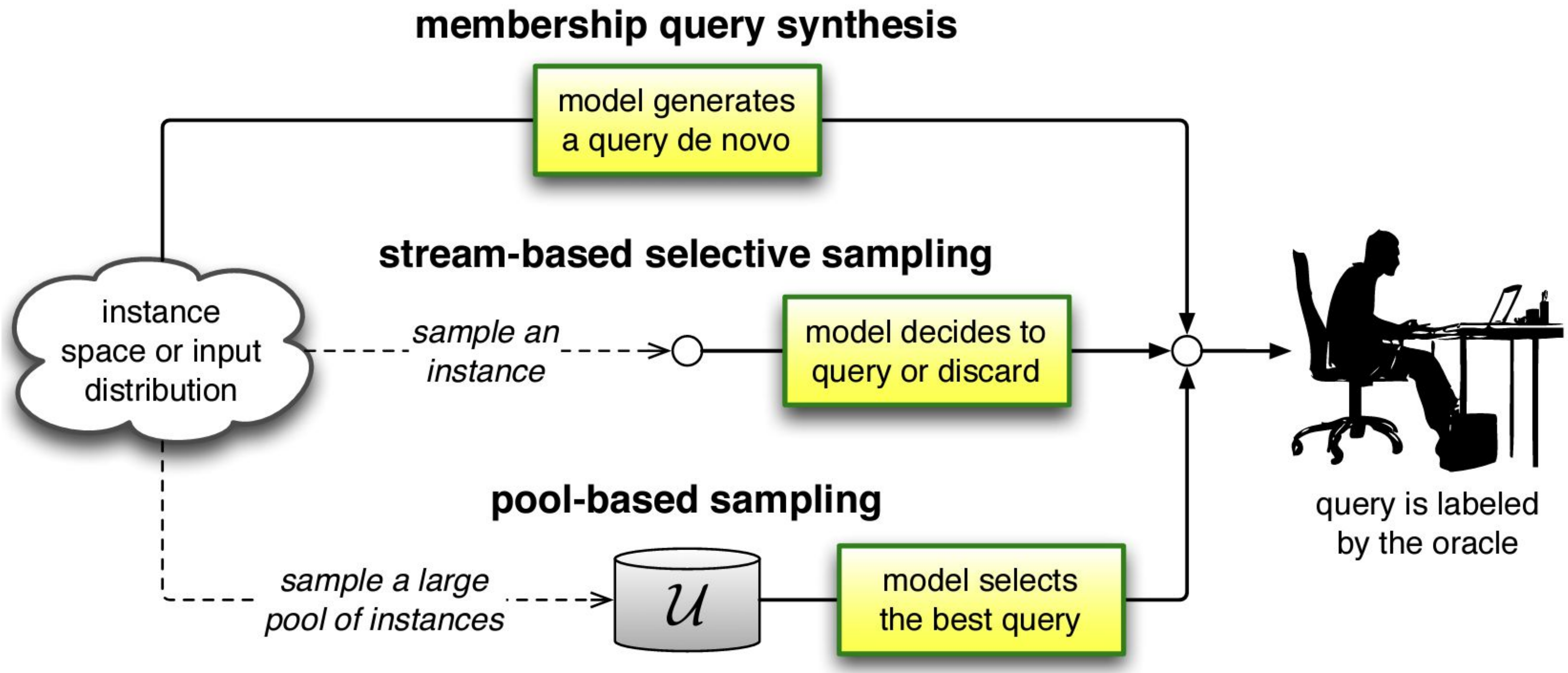


Function estimation: What is $\mathbb{E}[y|x]$?
Bandit optimization: Where is $\max_x \mathbb{E}[y|x]$?

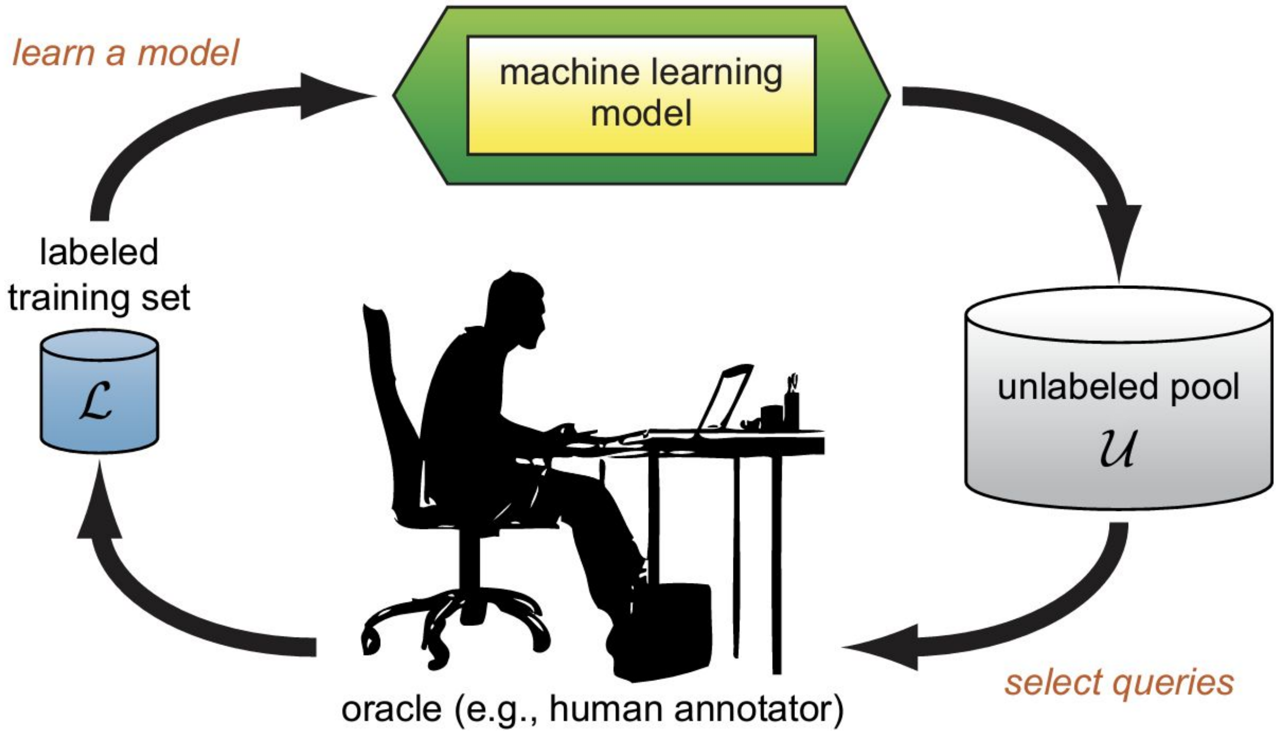


Active learning is more efficient than passive learning for localized “where” information

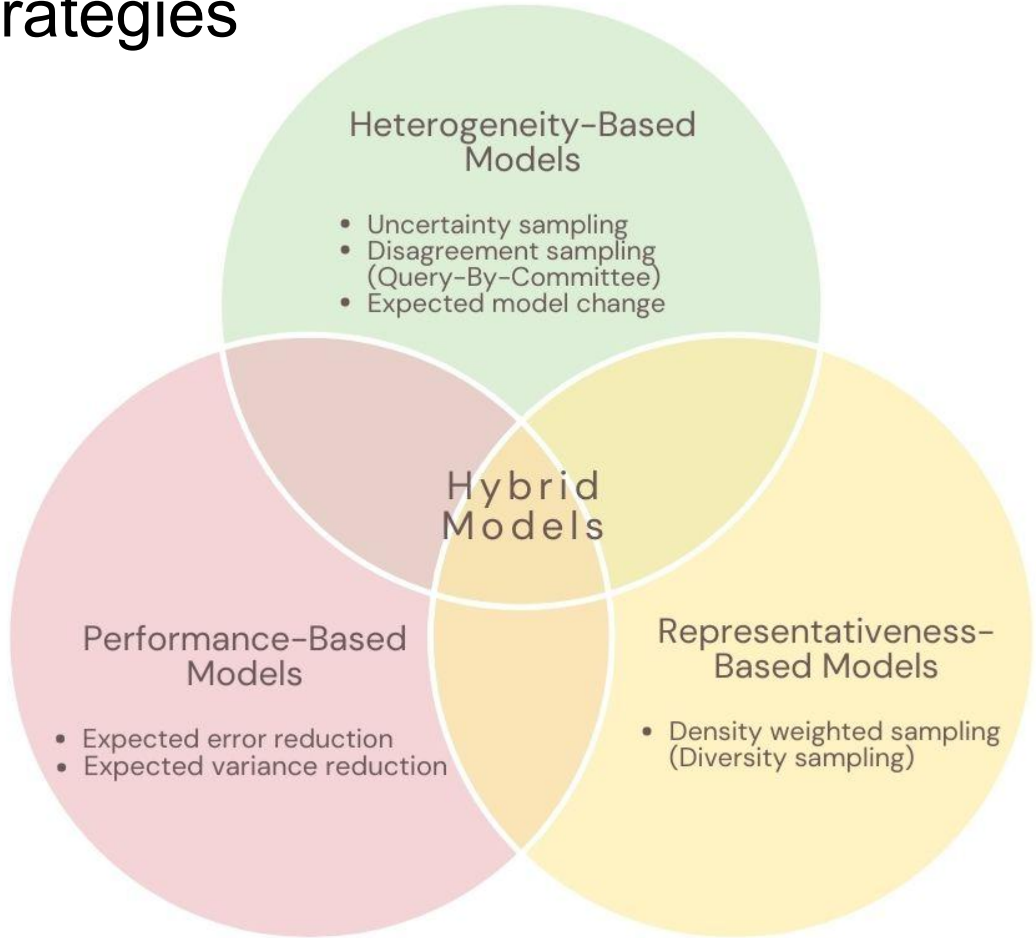
Three Main AL Scenarios



Pool-Based Scenario



Querying Strategies



Uncertainty Sampling

Acquisition function: map data to uncertainty rankings

- Entropy $H(p) = -\sum p_i Log_2(p_i)$
  - Entropy function is maximized the all it’s inputs are equal, meaning model is completely confused between the categories

$$H([0.5, 0.5]) = 1.0$$

$$H([1.0, 0.0]) = 0.0$$

- Variation Ratio: proportion of cases which are not in the mode category
- Best-versus-second best(BvSB) margin:ratio between the posteriors of the two most confident classes

Disagreement Sampling (Query-By-Committee)

Instead of measuring the uncertainty of a single model, we can train an ensemble of many different models that are consistent with labeled data.

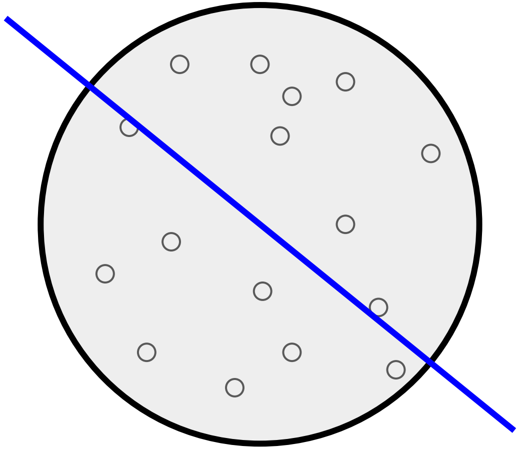
The most informative query is the data about which they most disagree.

Measure of disagreement:

- Vote Entropy
- Kullback-Leibler (KL) divergence

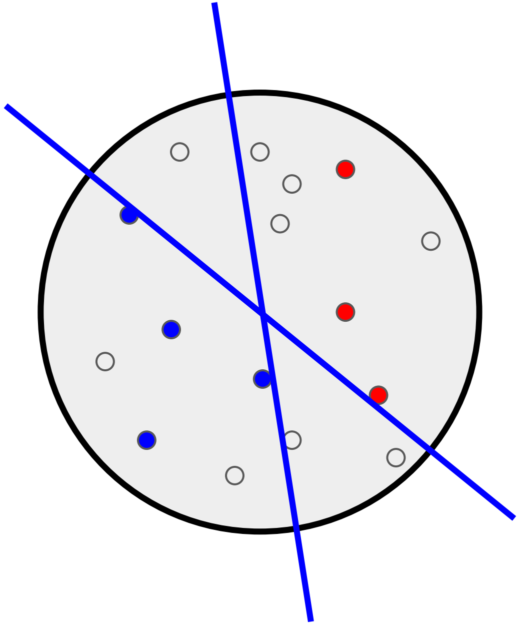
Disagreement Sampling

Consider points uniform on unit ball and linear classifiers passing through origin



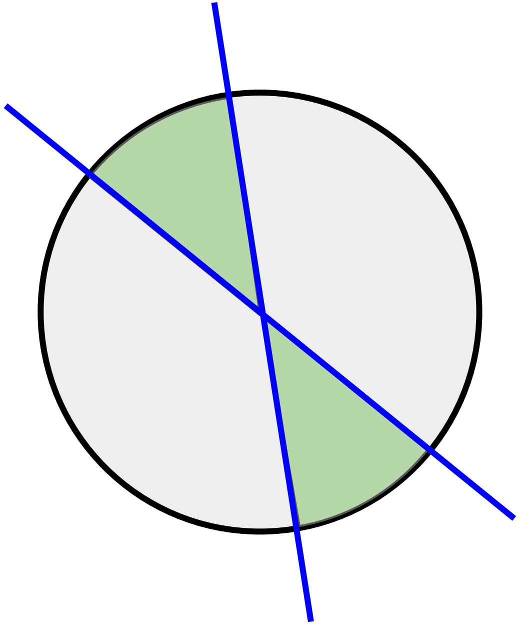
Disagreement Sampling

Consider points uniform on unit ball and linear classifiers passing through origin



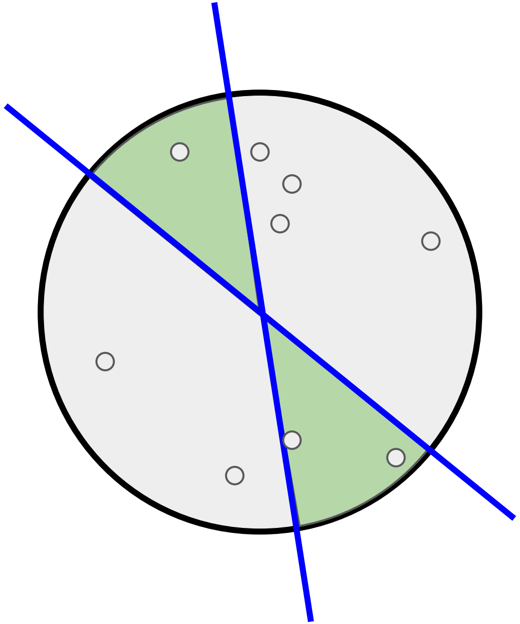
Disagreement Sampling

Consider points uniform on unit ball and linear classifiers passing through origin



Disagreement Sampling

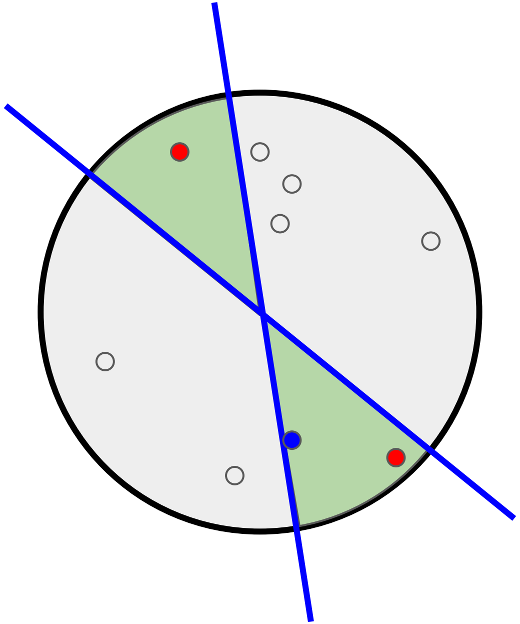
Consider points uniform on unit ball and linear classifiers passing through origin



- Only label points in the region of disagreement D

Disagreement Sampling

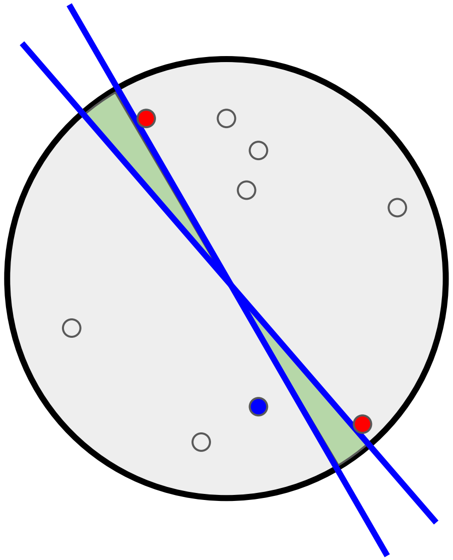
Consider points uniform on unit ball and linear classifiers passing through origin



- Only label points in the region of disagreement D

Disagreement Sampling

Consider points uniform on unit ball and linear classifiers passing through origin



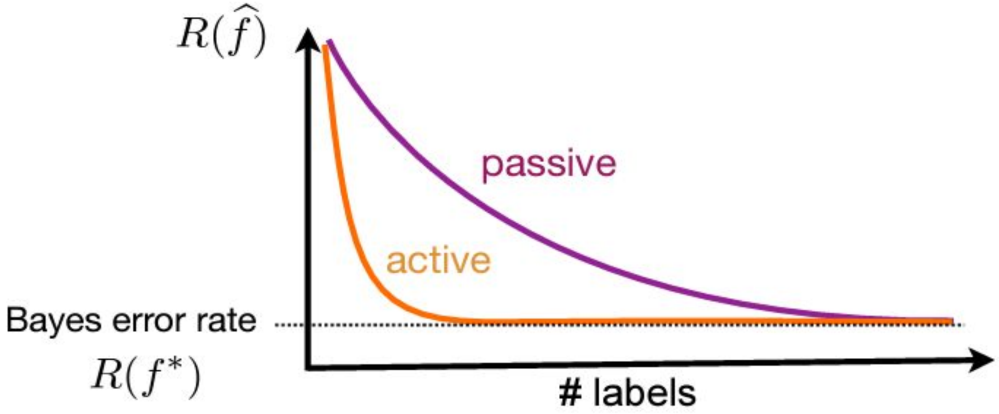
- Only label points in the region of disagreement D

Active Binary Classification

$$\epsilon = R(\widehat{f}) - R(f^*)$$

passive $\epsilon \sim \dfrac{d}{n}$ parametric rate

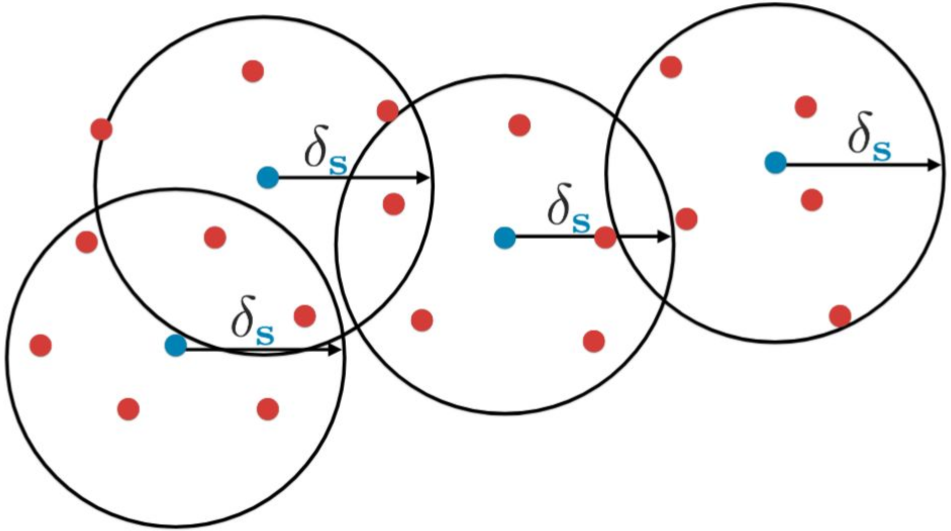
active $\epsilon \sim \exp\left(-c\,\dfrac{n}{d}\right)$ exponential speed-up



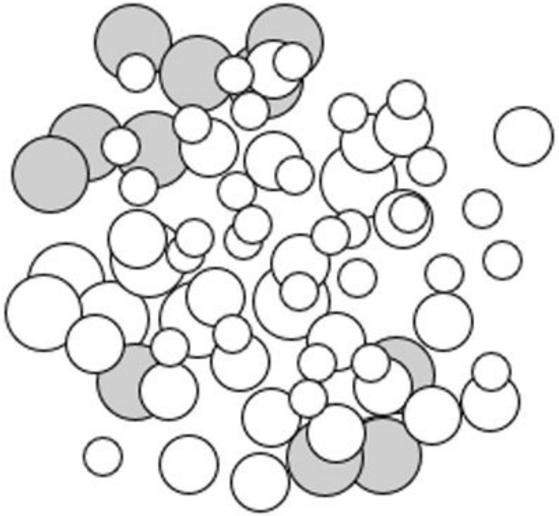
Diversity Sampling

Minimize the distance between unlabeled data and it’s closest labelled data.

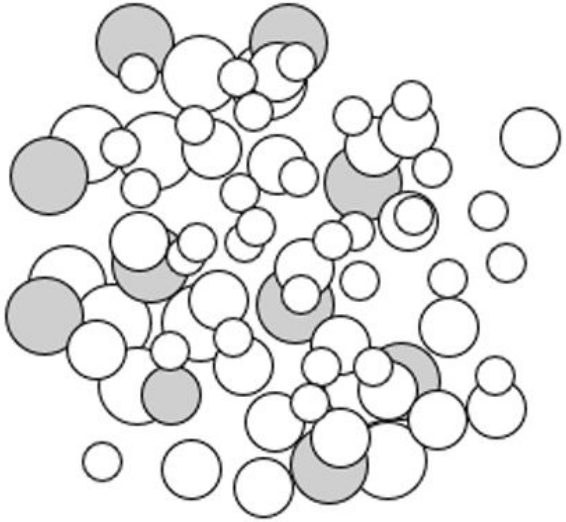
Greedy approximation: Find data x with the largest distance from the training set D, add x to D. Repeat.



Hybrid Methods

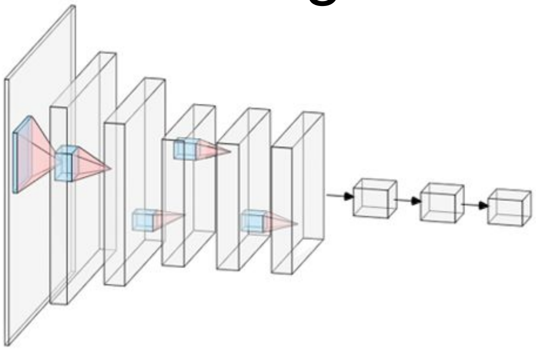


(a) Batch query strategy considering only the amount of information.

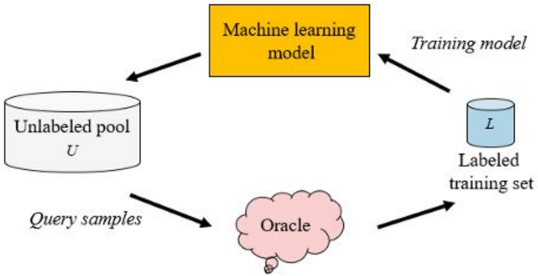


(b) Batch query strategy considering both information volume and diversity.

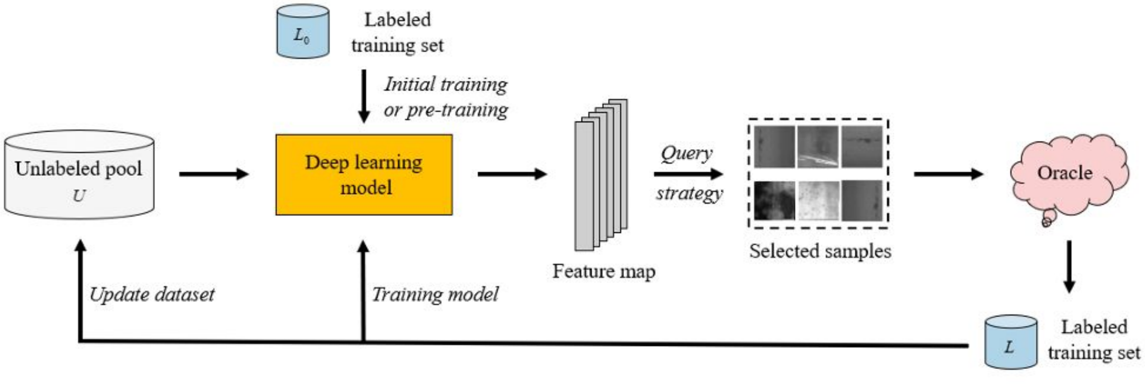
Deep Active Learning



(a) Structure diagram of convolutional neural network.



(b) The pool-based active learning cycle.



(c) A typical example of deep active learning.

Challenges Combining DL and AL

- Model uncertainty in DL is not clear; The output from the final softmax layer tends to be over confident.
  - Bayesian neural networks with Monte Carlo (MC) Dropout with can be used to obtain posterior uncertainties over network predictions
- Need acquisition function that works for batch setting
  - Naive method: choose top k
  - How to capture diversity and correlation between data?

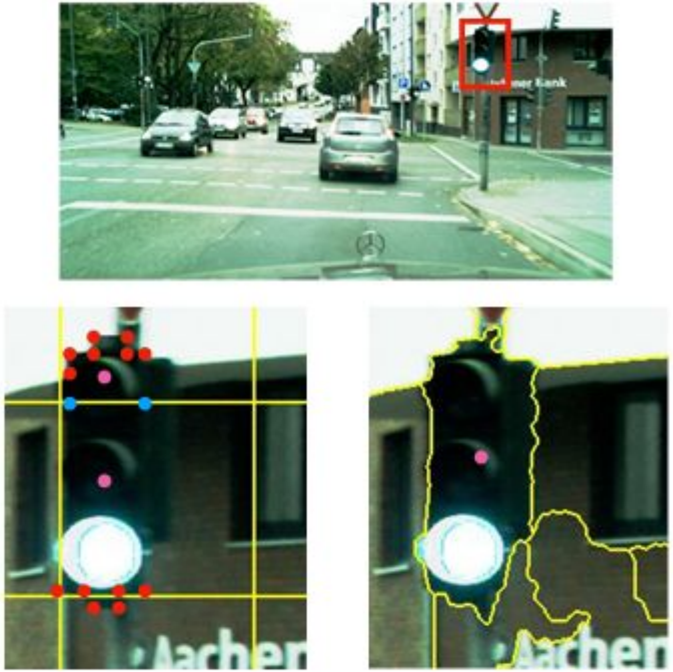
Application

- Visual Data Processing
  - Image classification and recognition
  - Object detection and semantic segmentation
  - Video processing
- Natural Language Processing
  - Machine translation
  - Text classification
  - Semantic analysis
  - Information extraction
  - Question-answering
- Gene expression, Robotics, Social Networking…

AL for semantic image segmentation

- Image-based
- Region-based
    - Regular shapes: rectangles, polygons…
    - Superpixels
    - Pixels

Applications in digital pathology, remote sensing and autonomous driving.

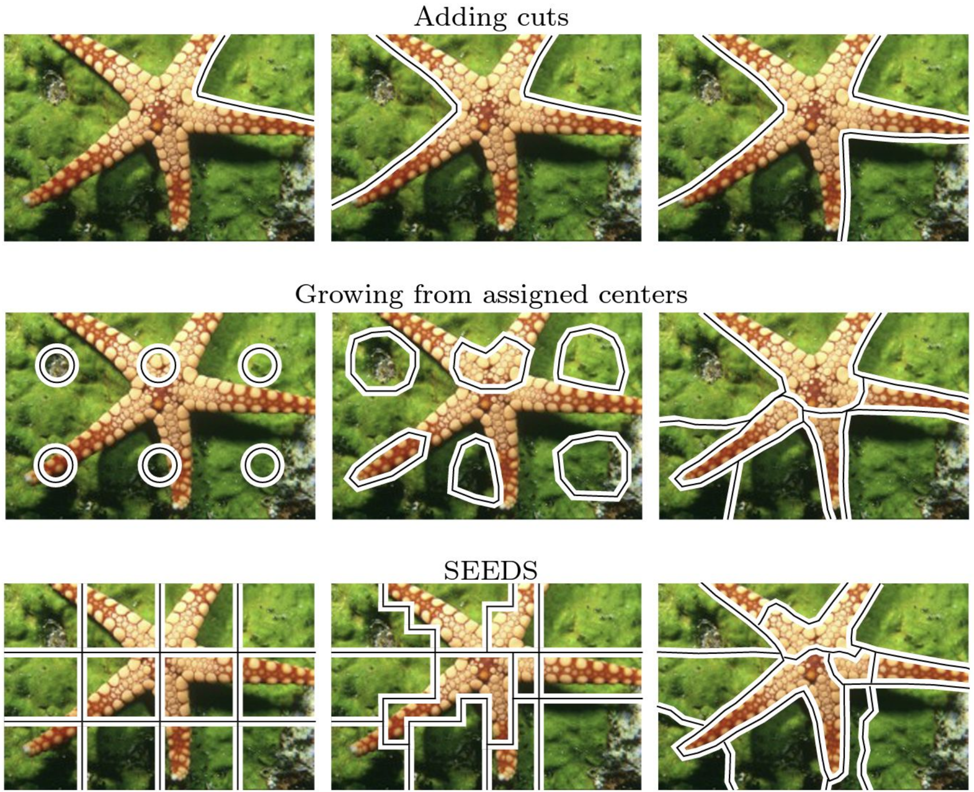


Revisiting Superpixels for Active Learning in Semantic Segmentation with Realistic Annotation Costs

Lile Cai[1], Xun Xu[1], Jun Hao Liew[2], Chuan Sheng Foo[1]
[1]Institute for Infocomm Research, Singapore
[2]National University of Singapore

CVPR 2021

Superpixel Generation



Methodology

Algorithm 1: Batch-Mode Active Selection

Input : unlabeled set of regions $\mathcal{U}_t$, labeled set of regions $\mathcal{L}_{t-1}$ selected in previous batches, model M_t trained on $\mathcal{L}_{t-1}$, annotation budget of K clicks for batch t

Output: Output selected set of regions $\mathcal{B}_t$

$\mathcal{B}_t = \emptyset$;

$total\_cost = 0$;

while $total\_cost < K$ **do**

$\quad s^* = \arg\max_{s \in \mathcal{U}_t} a(s, M_t)$;

$\quad \mathcal{B}_t = \mathcal{B}_t \cup s^*$;

$\quad \mathcal{U}_t = \mathcal{U}_t \setminus s^*$;

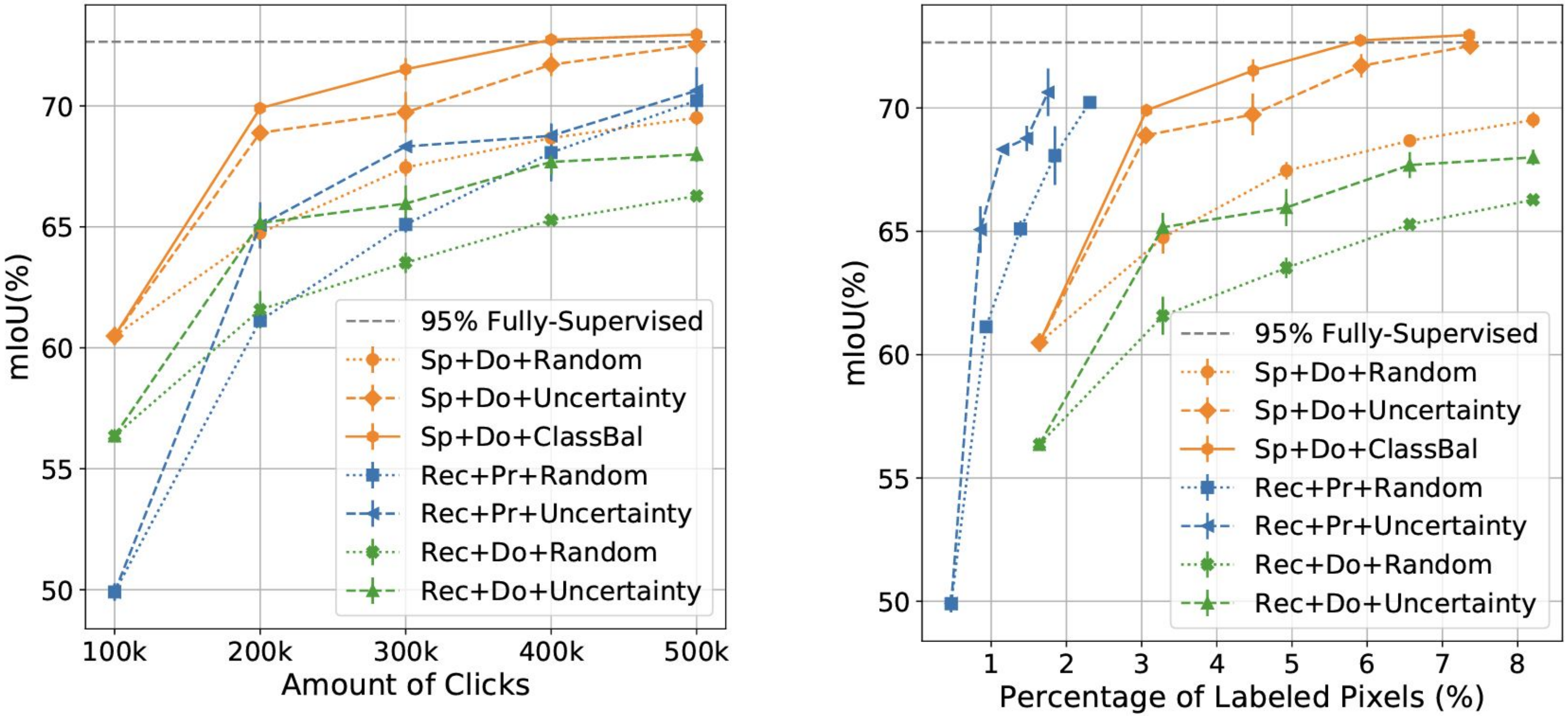
$\quad total\_cost = total\_cost + cost(s^*)$;

end

- Acquisition function: Best-versus-Second Best margin
- Class-balanced Sampling: assign weights to the uncertainty measure to favour samples from the under-represented classes.

Annotation Cost Measurement

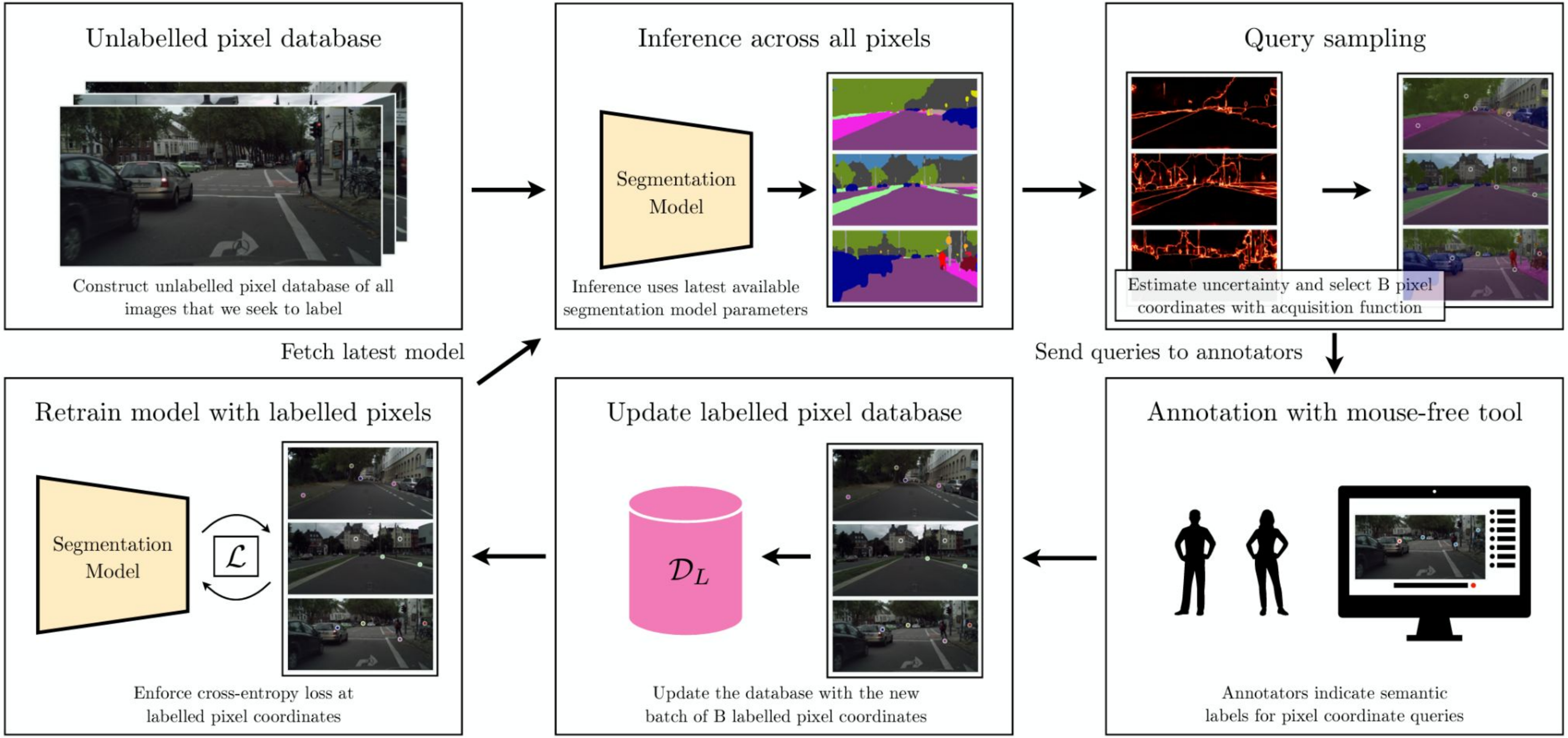
$$IoU = \frac{target \cap prediction}{target \cup prediction}$$



All you need are a few pixels: semantic segmentation with PixelPick

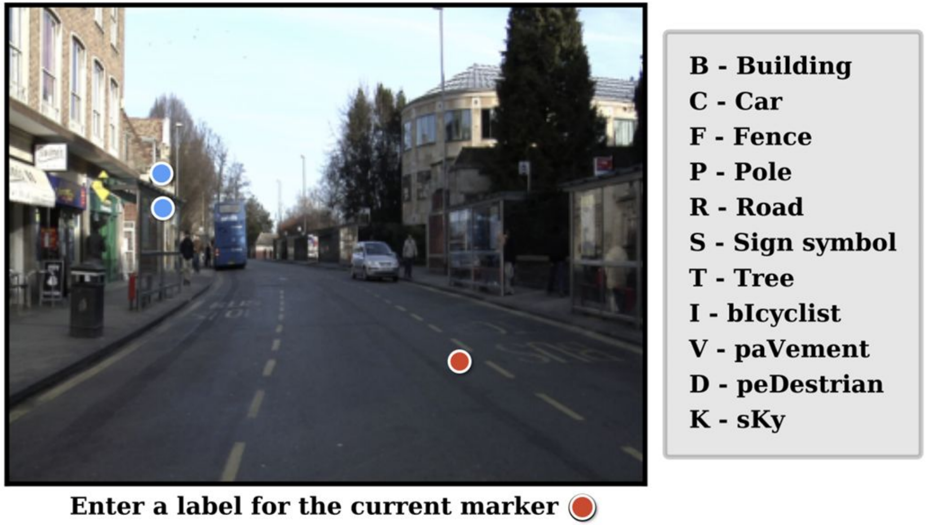
Gyungin Shin Weidi Xie Samuel Albanie
Visual Geometry Group, Department of Engineering Science
University of Oxford, UK

Best paper at ICCV 2021 ILDAV workshop

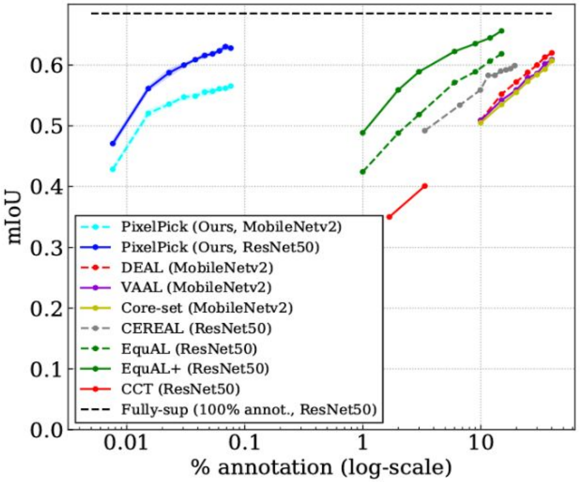

Unlabelled pixel database
Construct unlabelled pixel database of all images that we seek to label
Inference across all pixels
Segmentation Model
Inference uses latest available segmentation model parameters
Query sampling
Estimate uncertainty and select B pixel coordinates with acquisition function
Send queries to annotators
Annotation with mouse-free tool
Annotators indicate semantic labels for pixel coordinate queries
Update labelled pixel database
$\mathcal{D}_L$
Update the database with the new batch of B labelled pixel coordinates
Retrain model with labelled pixels
Segmentation Model
$\mathcal{L}$
Enforce cross-entropy loss at labelled pixel coordinates
Fetch latest model

Sampling Pixel Coordinates

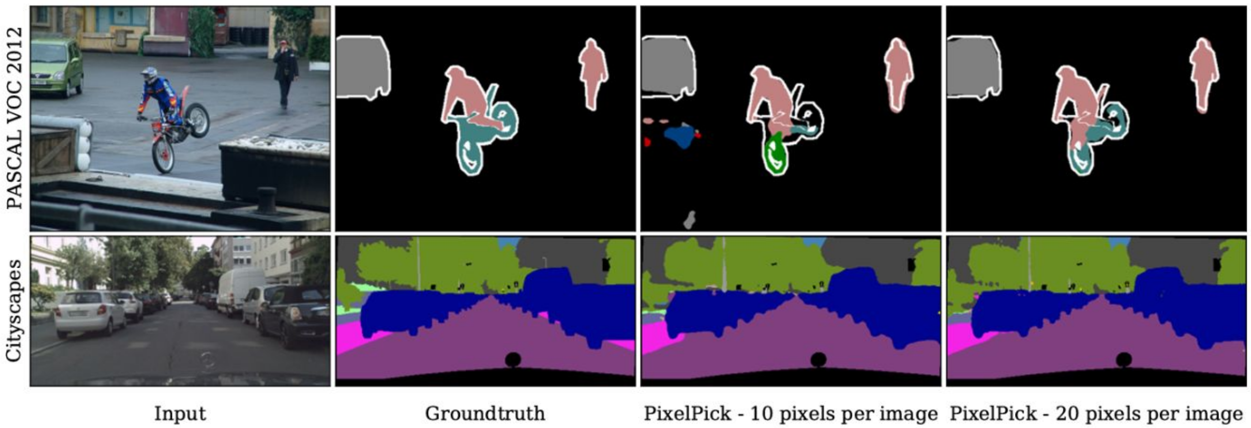
- From a “localise and classify” task to a “classify” task (single key-press).
- Leverage inductive biases provided by ConvNet to capture spatial dependencies.
- Robust to annotators errors.



Results



(a) Comparison to prior work on CITYSCAPES.



(b) Qualitative results for models trained with PIXELPICK on VOC12 (top) and CITYSCAPES (bottom).

Reference

- Bergh, Michael Van Den, et al. ‘Seeds: Superpixels Extracted via Energy-Driven Sampling’. INTERNATIONAL JOURNAL OF COMPUTER VISION, 2015, pp. 298–314.
- Burr Settles. Active Learning Literature Survey. Computer Sciences Technical Report 1648, University of Wisconsin–Madison. 2009.
- Cai, Lile, et al. ‘Revisiting Superpixels for Active Learning in Semantic Segmentation with Realistic Annotation Costs’. 2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), IEEE, 2021, pp. 10983–92. DOI.org (Crossref), https://doi.org/10.1109/CVPR46437.2021.01084.10988–10997. Print.
- Hanneke, Steve, and Robert Nowak. “[ICML 2019] Active Learning from Theory to Practice.” YouTube, 5 Oct. 2019, https://youtu.be/_Ql5vfOPxZU.
- Gal, Yarin et al. "Deep Bayesian Active Learning with Image Data." (2017)
- Gildenblat, Jacob. “Overview of Active Learning for Deep Learning.” Jacob Gildenblat, 21 Feb. 2020, https://jacobgil.github.io/deeplearning/activelearning.
- Jordan, J. (2018, December 18). Evaluating image segmentation models. Jeremy Jordan. Retrieved June 22, 2022, from https://www.jeremyjordan.me/evaluating-image-segmentation-models/
- Ren, Pengzhen, et al. ‘A Survey of Deep Active Learning’. CoRR, vol. abs/2009.00236, 2020, https://arxiv.org/abs/2009.00236
- Shin, Gyungin, et al. ‘All You Need Are a Few Pixels: Semantic Segmentation with PixelPick’. 2021 IEEE/CVF International Conference on Computer Vision Workshops (ICCVW), IEEE, 2021, pp. 1687–97. DOI.org (Crossref), https://doi.org/10.1109/ICCVW54120.2021.00194.