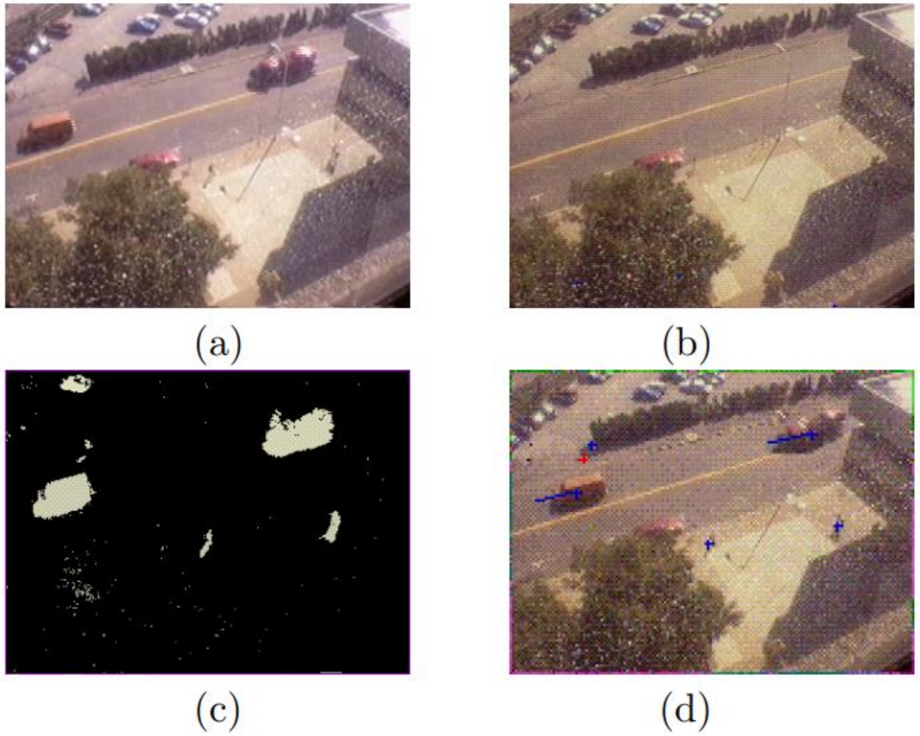
Pointly-supervised Scene Parsing with Uncertainty Mixture

Hao Zhao, Ming Lu, Anbang Yao, Yiwen Guo, Yurong Chen, Li Zhang

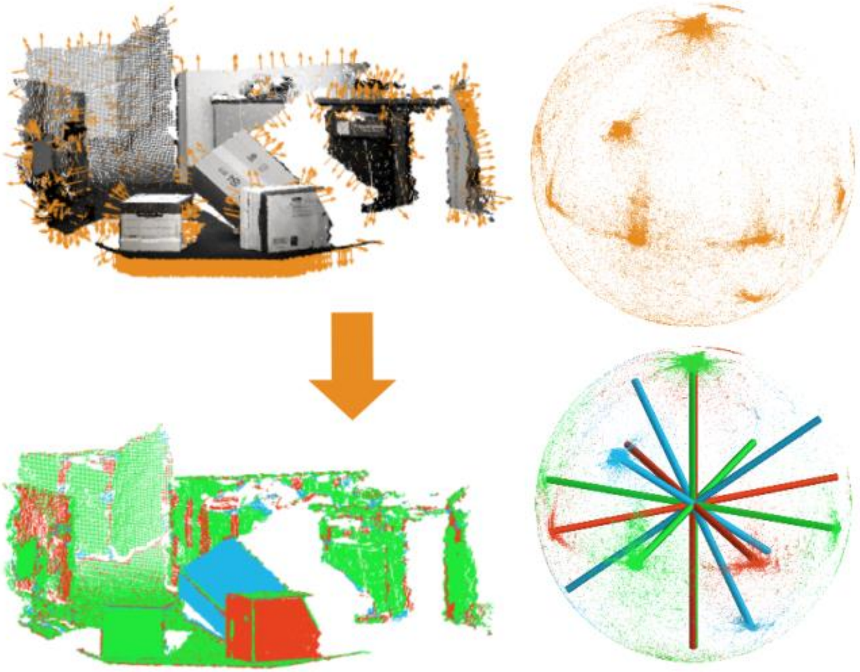
Contributions

- A Discovery, A Pipeline, A Model and An Evaluation
- We identify the existence of a previously unknown statistical phenomenon called uncertainty mixture.
- We propose a principled pipeline to harvest pseudo labels for pointly-supervised scene parsing, without the need of threshold tuning.
- We contribute a novel regularized Gamma mixture model.
- We achieve state-of-the-art results on PascalContext and ADE20k.

Finite Mixture Models in Computer Vision



Modelling pixel intensity in surveillance camera streams
Stauffer, CVPR 1999



Modelling normal in Manhattan indoor scences
Straub, CVPR 2014

In the DL era…

Does NN weights naturally emerge as mixture models?

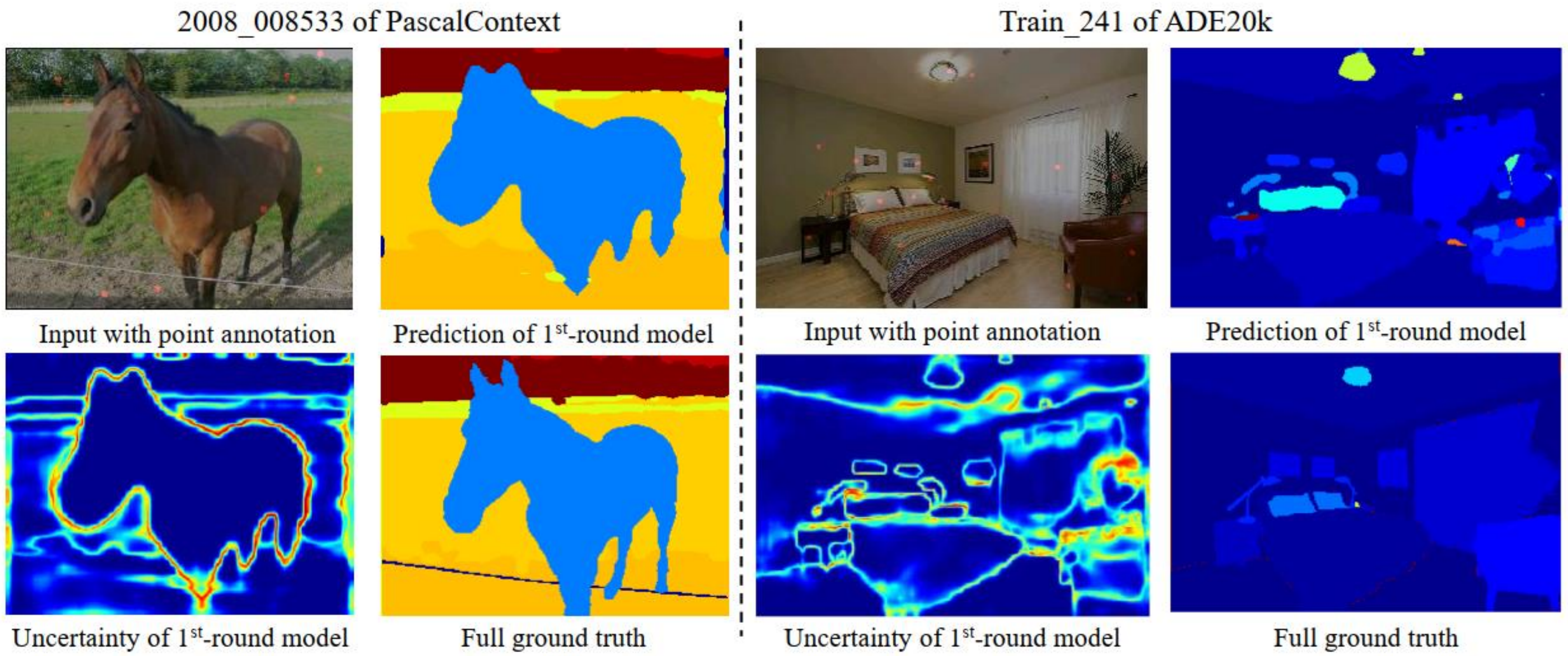
Seems not.

Does NN features naturally emerge as mixture models?

Seems not.

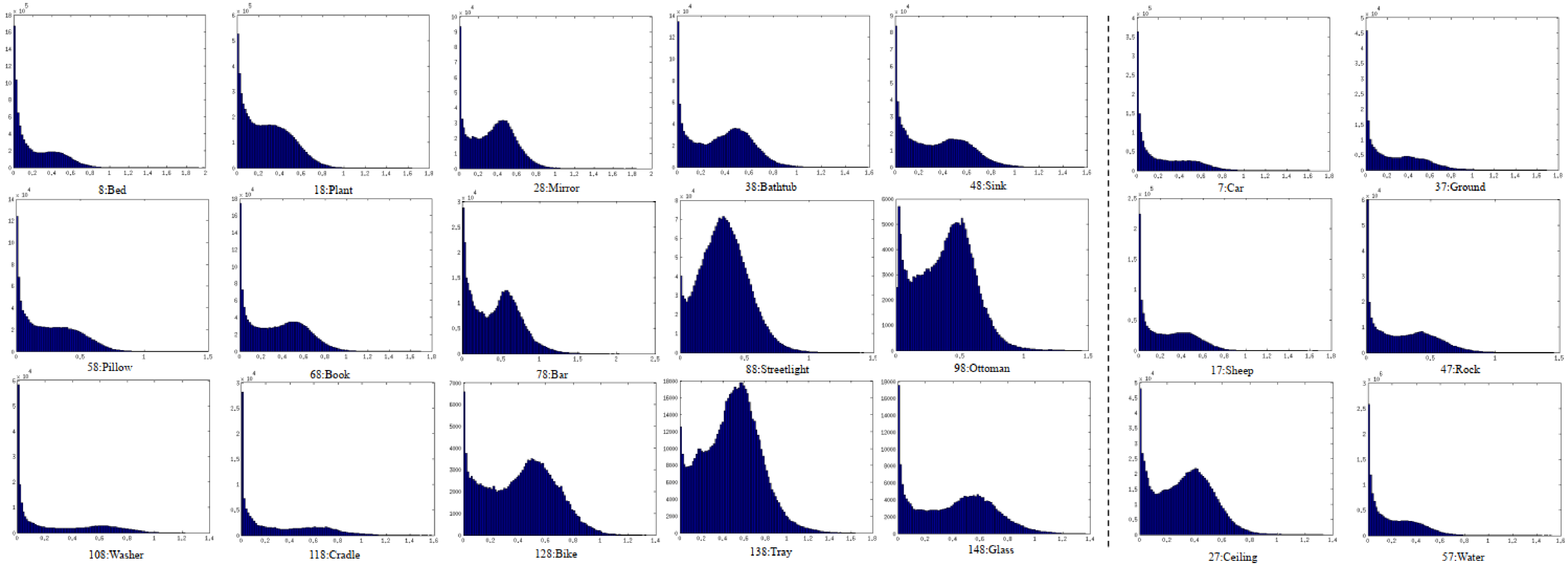
We identify the fact uncertainty measures (in the pointly-supervised scene parsing setting) emerge as Gamma mixtures

Pointly-supervised Scene Parsing



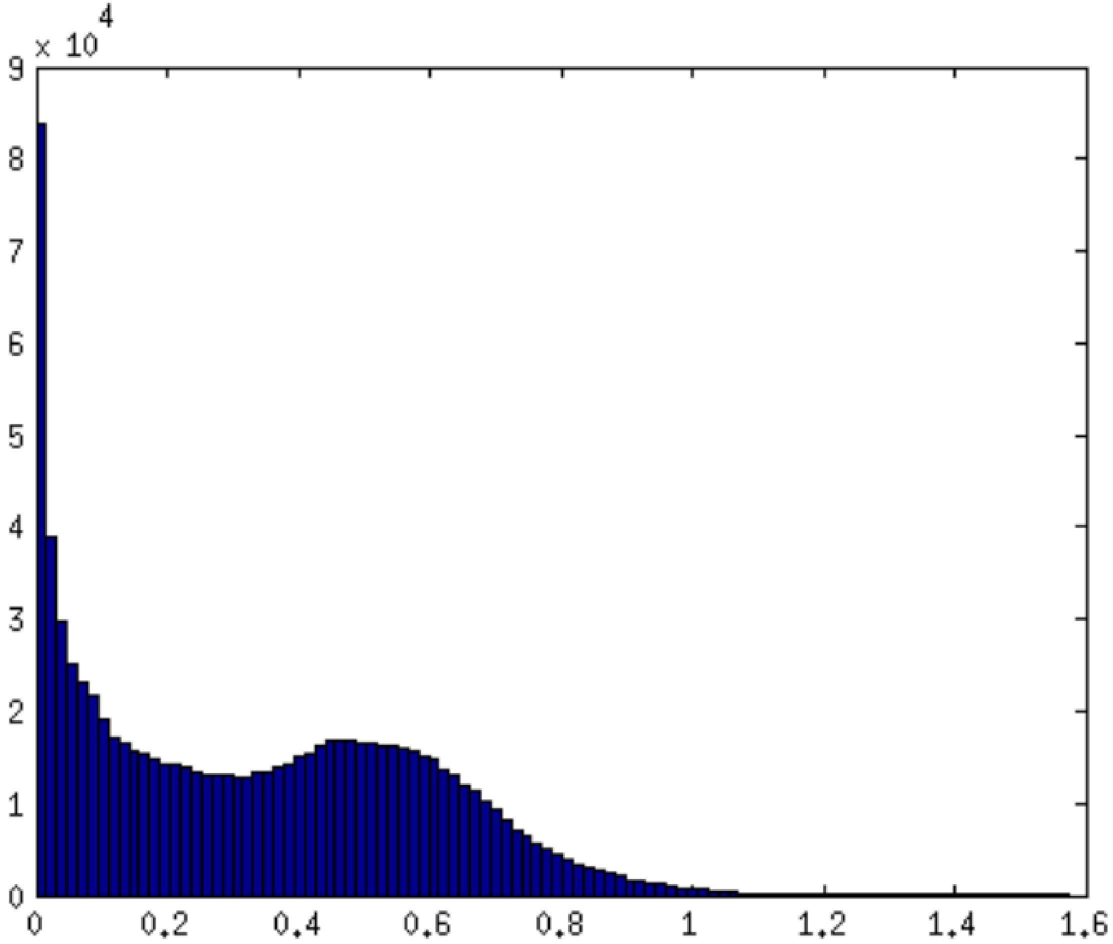
Pseudo Labels are noisy.
Harvesting pseudo labels with low uncertainty seems promising.
How to define 'low' uncertainty?

Category-wise uncertainty measures emerge as a two-peak mixture.



We claim the statistical phenomenon of uncertainty mixture exists as it is ubiquitously observed in large-scale datasets in ADE and PascalContext.

Modelling: Gaussian or Gamma?



LUC: Low-uncertainty component
HUC: High-uncertainty component

(1) Uncertainty measures are variance so that they are always positive; Gamma distribution support: $(0,+\infty)$

(2) LUC is naturally peaked near zero (just because it is the low component); problematic for Gaussian distribution

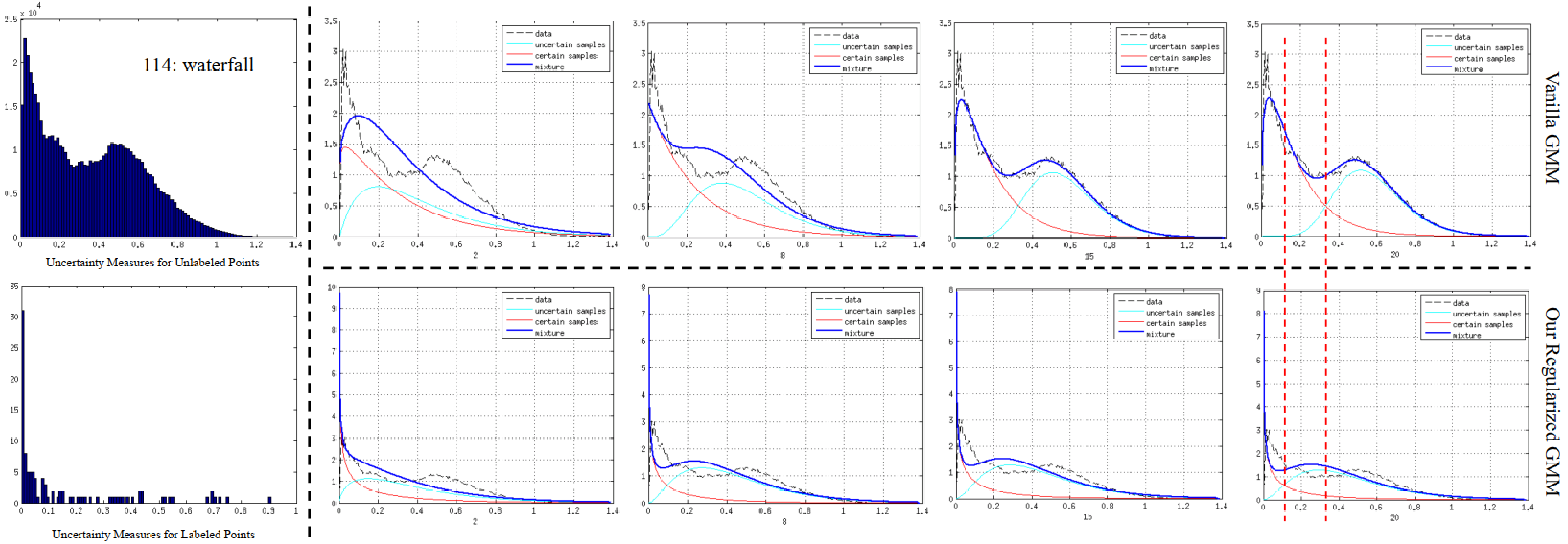
(3) HUC looks quite Gaussian? It is fine because Gamma approaches Gaussian when α grows to $+\infty$

Pipeline: Harvesting LUC labels and fine-tune the net.

(1) Train the first-round model using point supervision;
(2) Get pseudo labels and uncertainty measures on the training set;
(3) EM estimation for LUC and HUC, on a category-wise basis;
(4) Harvesting LUC labels;
(5) Finetune the first-round model.

A pipeline to harvest pseudo labels without manual thresholding.

A regularized Gamma mixture model



Idea: Assuming labelled points always belong to LUC

For EM convergence guarantee and analytical details, check the paper

Evaluation: Weakly-supervised learning and Transductive inference

	1st-round	Gamma	rGamma
22-layer	31.52	32.48 (+0.96)	**34.17 (+2.65)**
54-layer	32.63	33.52 (+0.89)	**35.44 (+2.81)**
105-layer	33.54	34.39 (+0.85)	**36.07 (+2.53)**

Table 1. Quantitative results on PASCALContext. All numbers are measured in the metric of mean intersection over union (mIoU, %).

	1st-round	Gamma	rGamma
22-layer	24.53	25.45 (+0.92)	**27.00 (+2.47)**
54-layer	25.20	26.29 (+1.09)	**27.19 (+1.99)**
105-layer	26.33	27.44 (+1.11)	**28.79 (+2.46)**

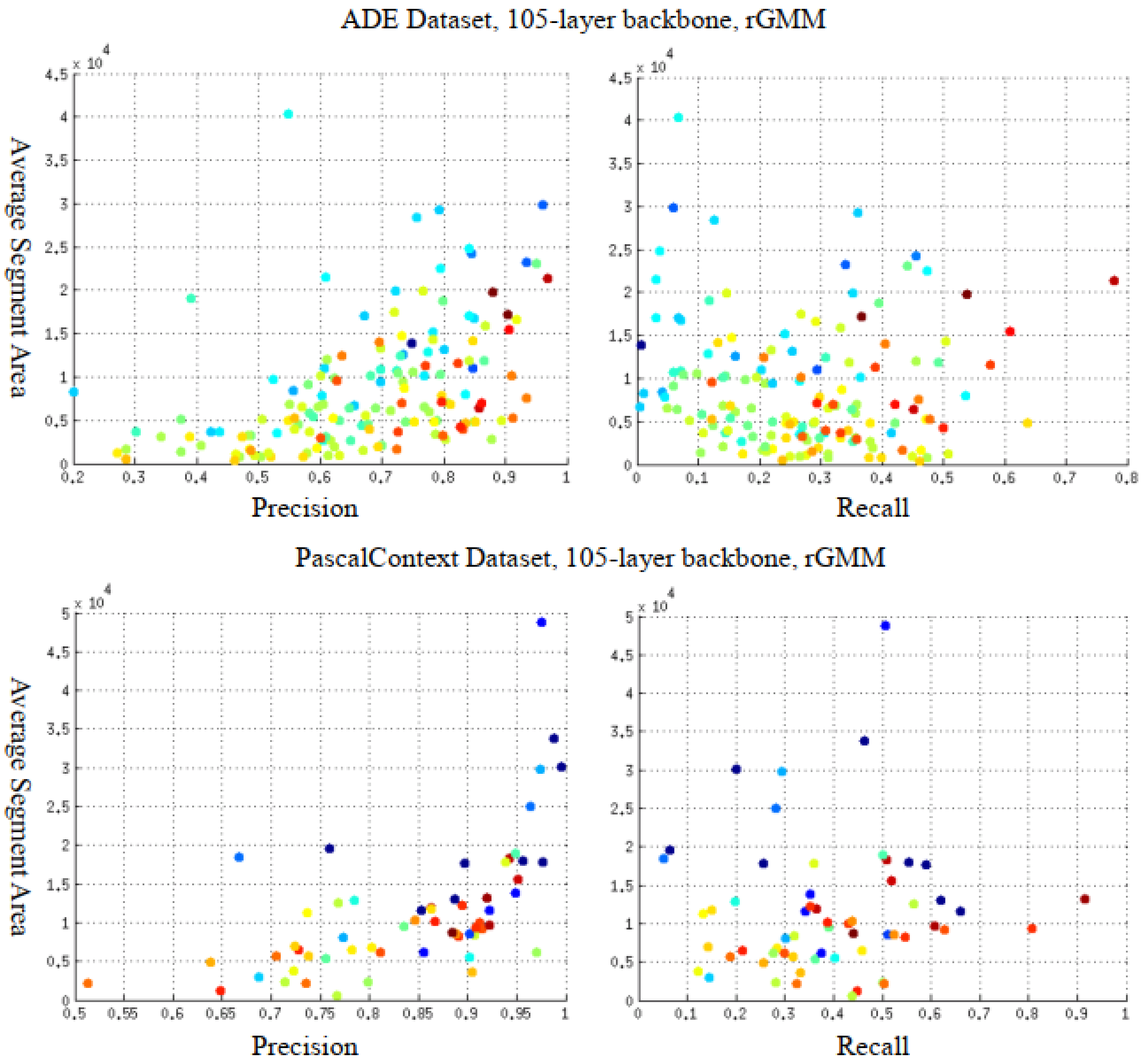
Table 2. Quantitative results on ADE20k. All numbers are measured in the metric of mean intersection over union (mIoU, %).

		ADE20k		PASCALContext	
	Arch	mR	mP	mR	mP
GMM	22	26.43	63.02	34.06	77.38
	54	26.87	63.76	37.19	79.91
	105	**27.64**	64.98	**39.03**	80.85
rGMM	22	25.45	65.72	33.28	81.28
	54	25.98	66.66	36.44	83.34
	105	26.89	**67.58**	38.26	**84.18**

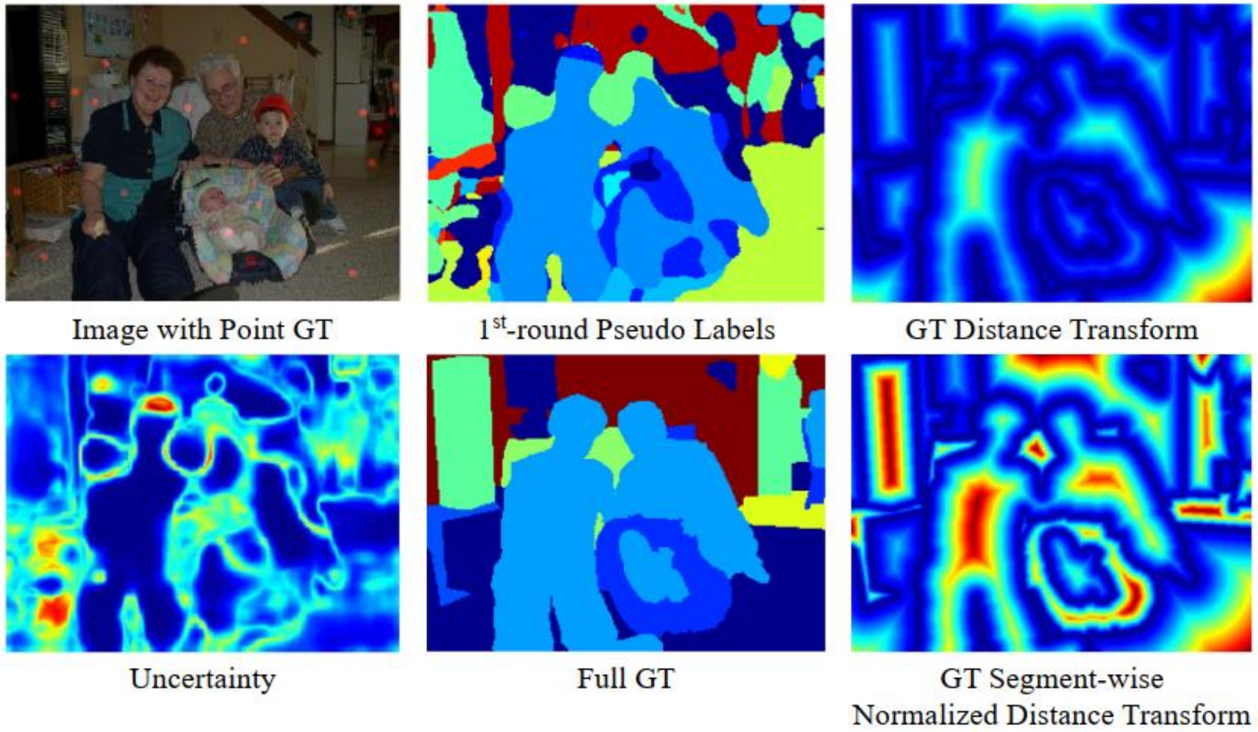
Table 5. Transductive inference performance for our method. Short names mP/mR are categorical mean values for precision/recall, which are measured in percentage (%). Arch stands for the depth of backbones.

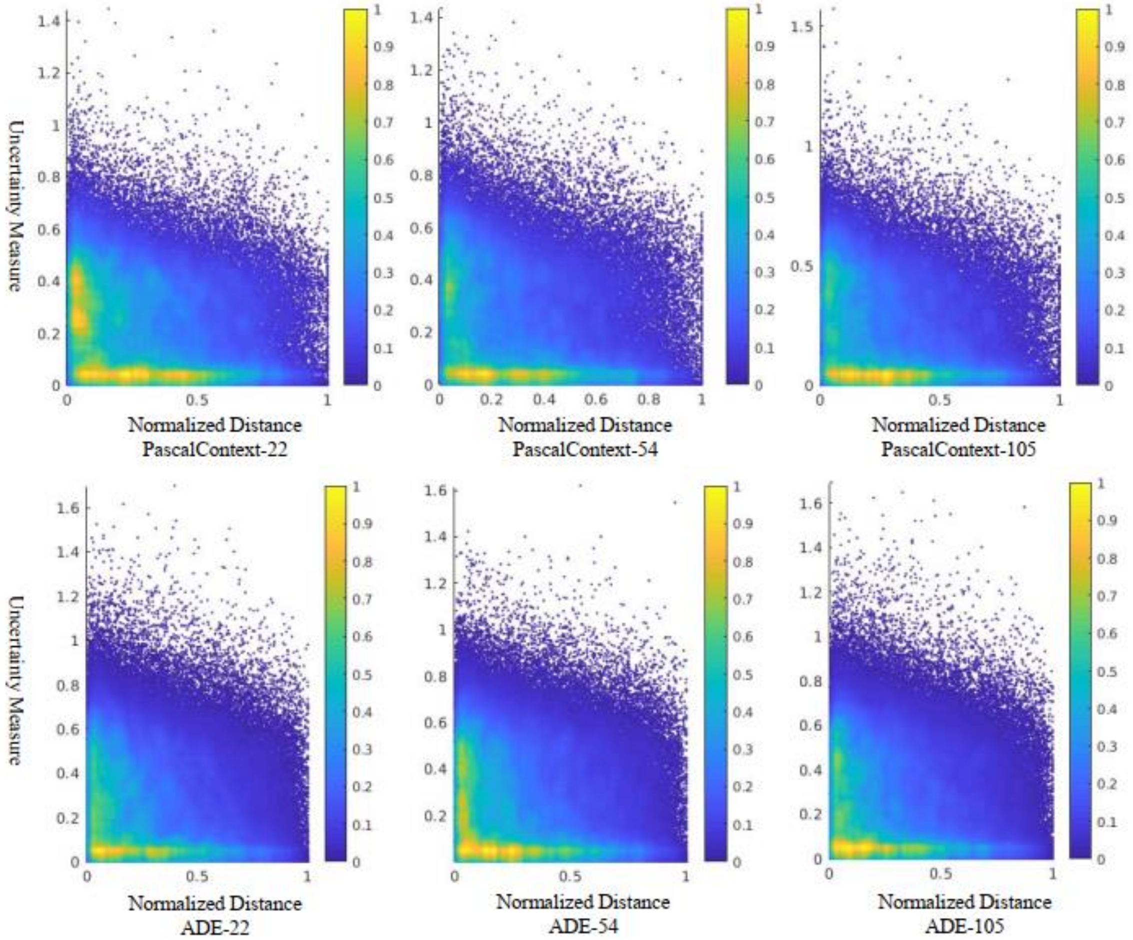
More insights:

Segment Area
Precision
Recall
Supervision Points



More insights: Drawing supervision points near boundaries?





Promising yet not ready

On-going research: Other uncertainty measures ?

- We use Drop-out uncertainty in our research.

Dropout as a bayesian approximation: Representing model uncertainty in deep learning

- We have checked random BN uncertainty, it worked too.

Bayesian uncertainty estimation for batch normalized deep networks

- We are investigating Gumbel-softmax uncertinaty, which works but weakens the baseline.

A Bayesian Neural Net to Segment Images with Uncertainty Estimates and Good Calibration

Contributions

- A Discovery, A Pipeline, A Model and An Evaluation
- We identify the existence of a previously unknown statistical phenomenon called uncertainty mixture.
- We propose a principled pipeline to harvest pseudo labels for pointly-supervised scene parsing, without the need of threshold tuning.
- We contribute a novel regularized Gamma mixture model.
- We achieve state-of-the-art results on PascalContext and ADE20k.