Collaborative Learning of Semi-Supervised Clustering and Classification for Labeling Uncurated Data

Sara Mousavi[1], Dylan Lee[1], Tatianna Griffin[2], Dawnie Steadman[2], Audris Mockus[1]

1. Department of Electrical Engineering and Computer Science, 2. Department of Anthropology

Funded by National Institute of Justice

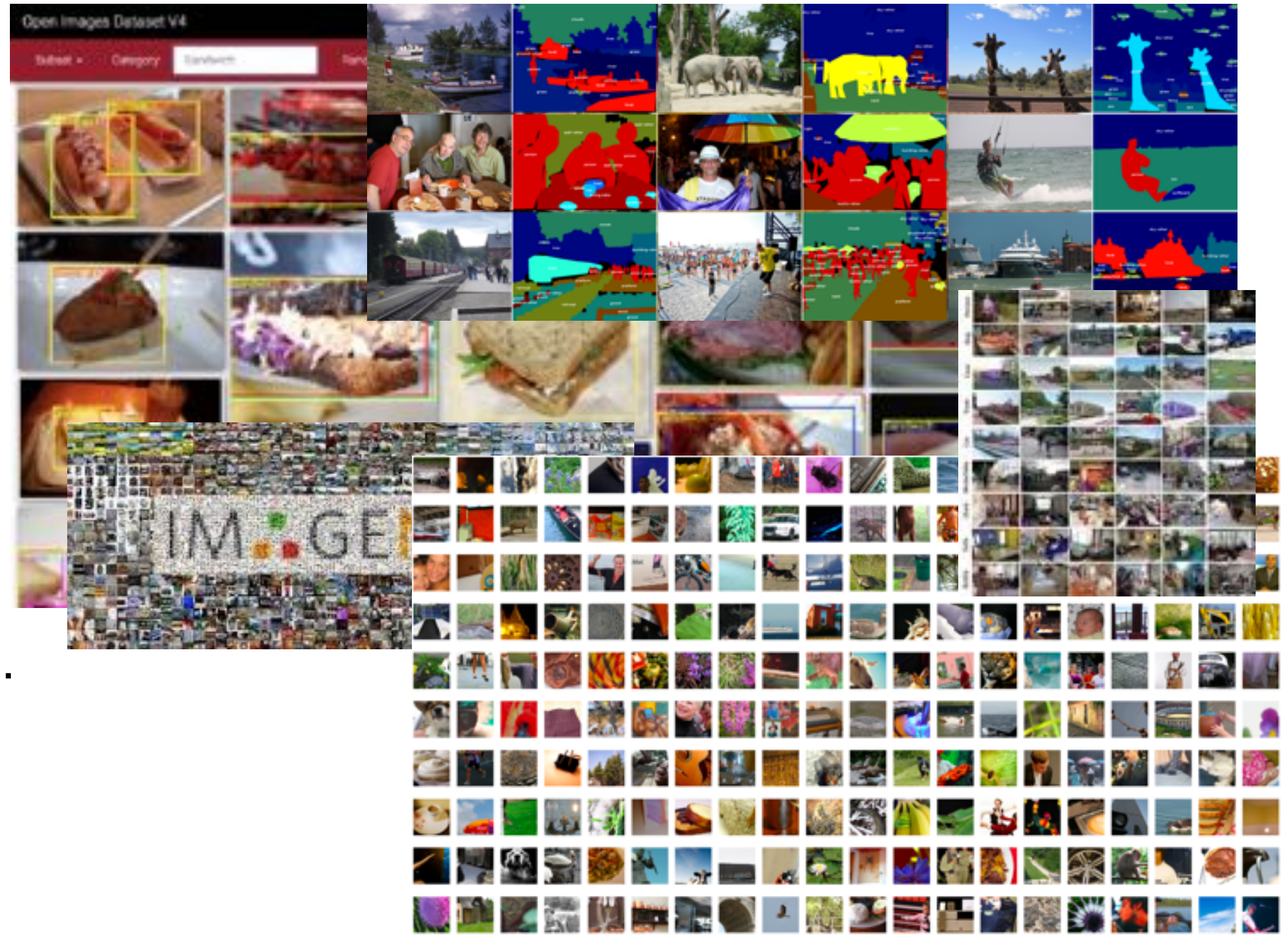
Preprint available at: https://arxiv.org/abs/2003.04261

July 2020

THE UNIVERSITY OF TENNESSEE KNOXVILLE

Large image collections are common

- Applications, e.g.,
  - Autonomous driving
  - Healthcare
  - Finance and banking
- Benchmarking datasets
  - ImageNet, COCO, Open Images, CIFAR, …



Specialized domains

- E.g., forensics, material science, biology, medical, ...
  - No pre-existing labels
  - Limited transferability of labeling efforts from other datasets
- Present potential value in various areas of science and business
- Require curation to answer research questions and search within the data
  - Expensive
  - Time consuming

Curation challenges

- Many datasets are too large for completely manual curation
- Privacy/proprietary/expertise concerns may preclude crowd-sourcing curation
  - Sensitive data (medical, personal, financial, …)
  - Requires domain expertise
- Conclusion: Efficient image labeling is an essential task needed to unlock valuable information in such image collections

Unsupervised and Supervised methods

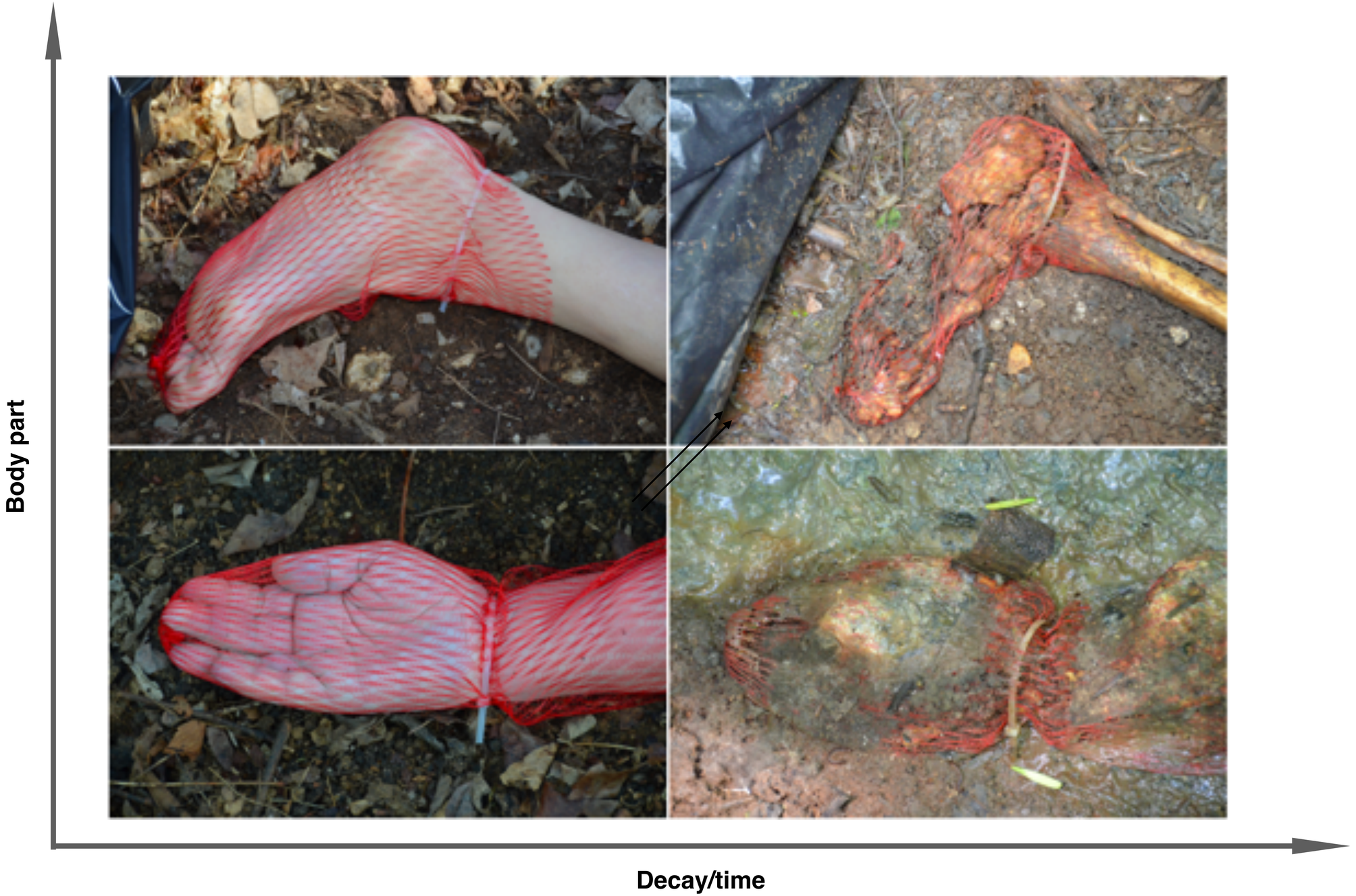
- Unsupervised methods do not depend on labeled data
  - Cluster image data using their feature representations
  - Good representations are hard to be obtained for domain specific data
  - Evaluation requires manual intervention
    - Time consuming
    - Expensive (requires domain expertise)
- Supervised methods depend on labeled data
  - Labeled data is not readily available for all domains

Plud: a Platform for Labeling Uncurated Data

- Human-machine collaboration (semi-supervised) for labeling data
  - Accelerates the labeling process to handle large amounts of uncurated data
  - Minimizes the labeling effort by experts to utilize the limited availability of experts
- A workflow consisting of unsupervised and supervised components

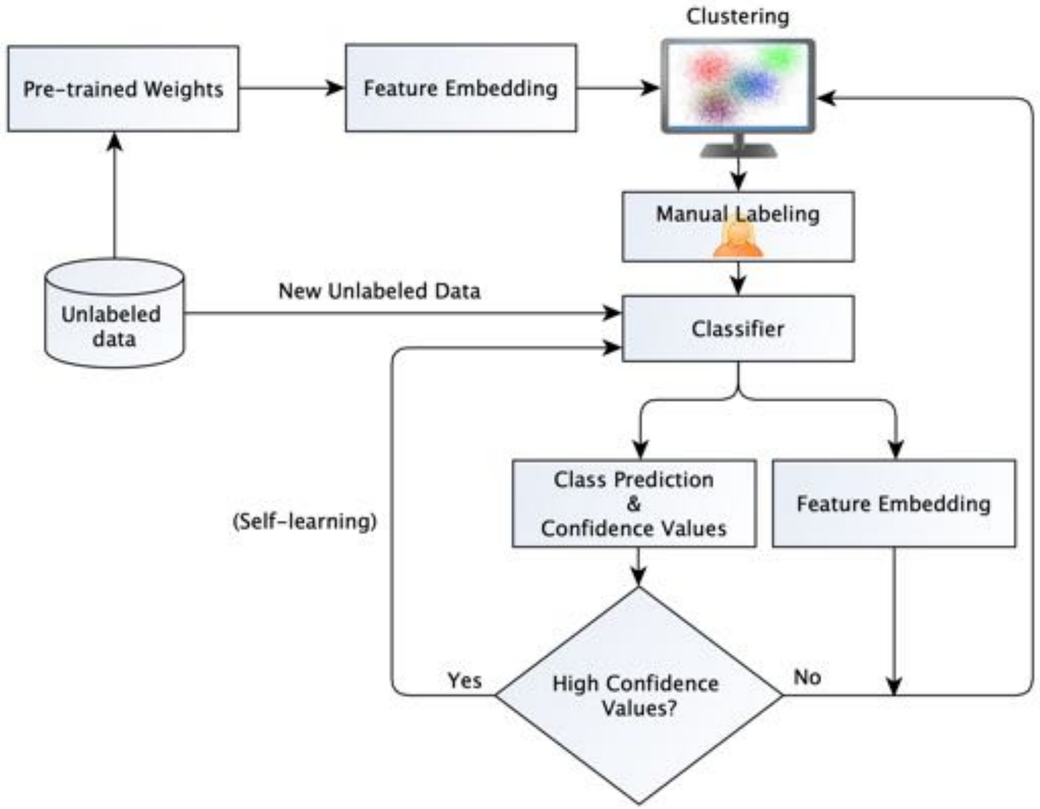
Human decomposition dataset

- Daily photos of ~500 subjects in an 8-year period
- Multiple images from various body part per subject
  (Arm, Hand, Foot, Legs, Full body, Head, Backside, Torso, Stake, Plastic)
- Various decomposition stages due to:
  - Weather changes
  - Time of death
  - Prior conditions of subjects


Body part
Decay/time

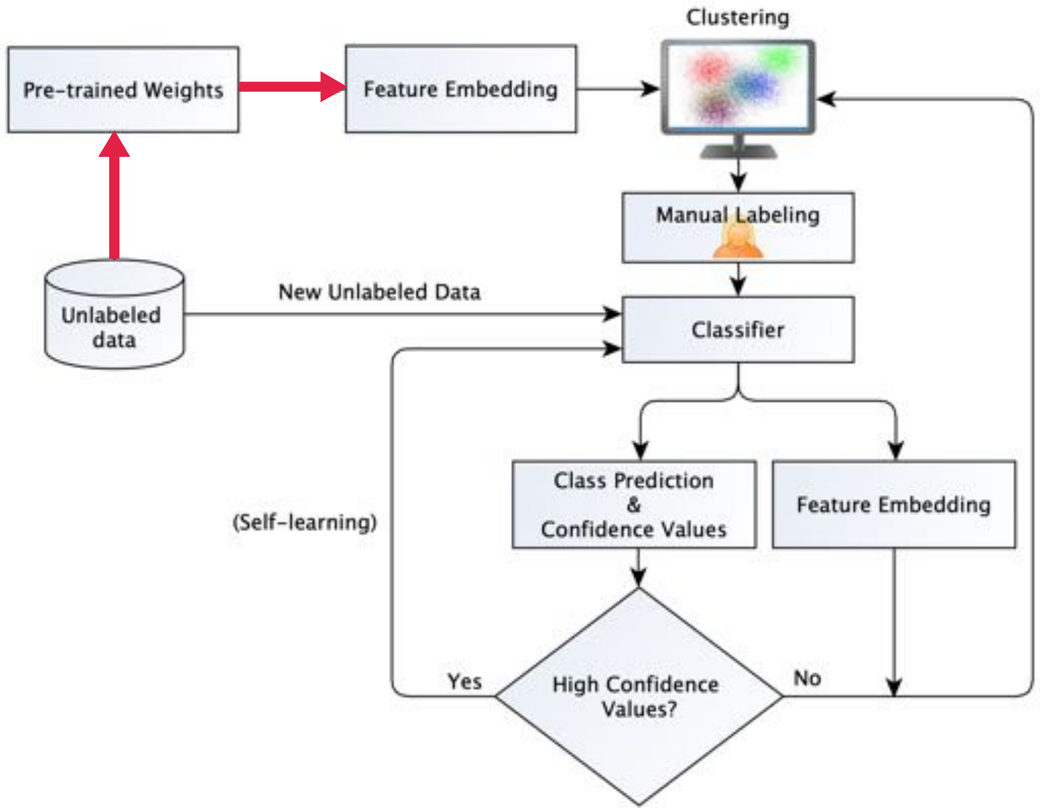
Plud: Machine-assisted labeling

- Iterates over
  - Clustering
  - Human supervision
  - Classification
- Objective: accelerate and simplify manual labeling



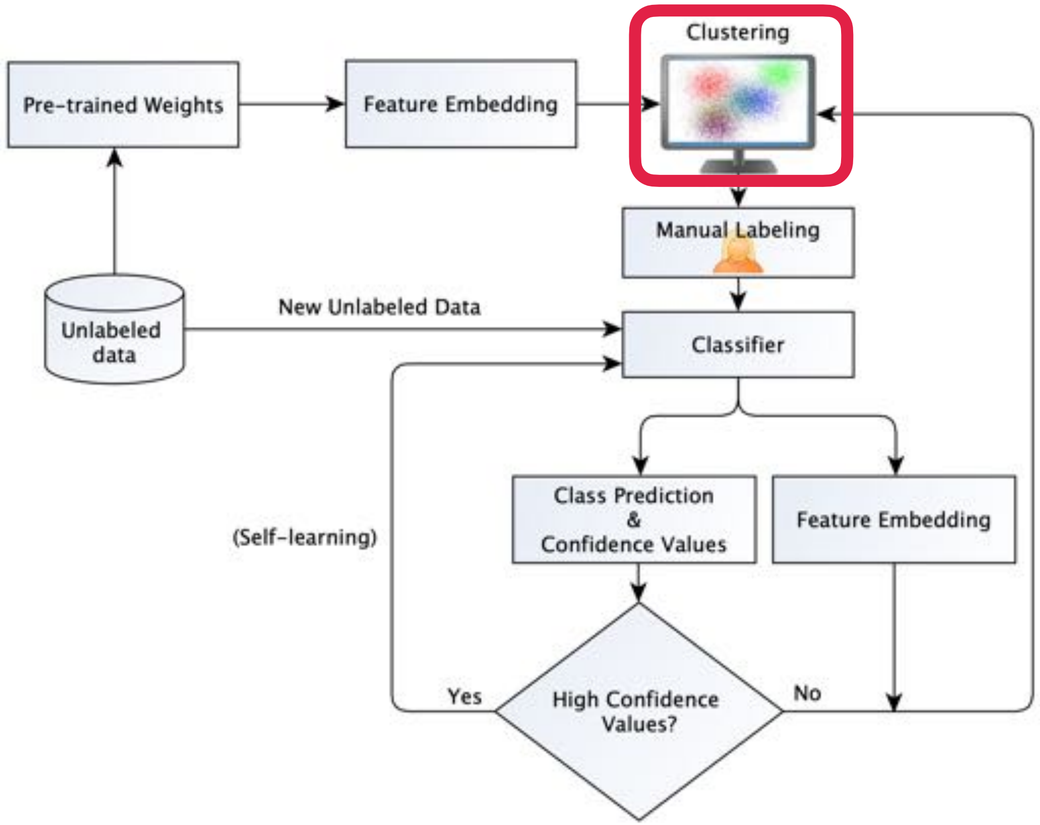
Plud: Machine-assisted labeling

- Iterates over
  - Clustering
  - Human supervision
  - Classification
- Objective: accelerate and simplify manual labeling



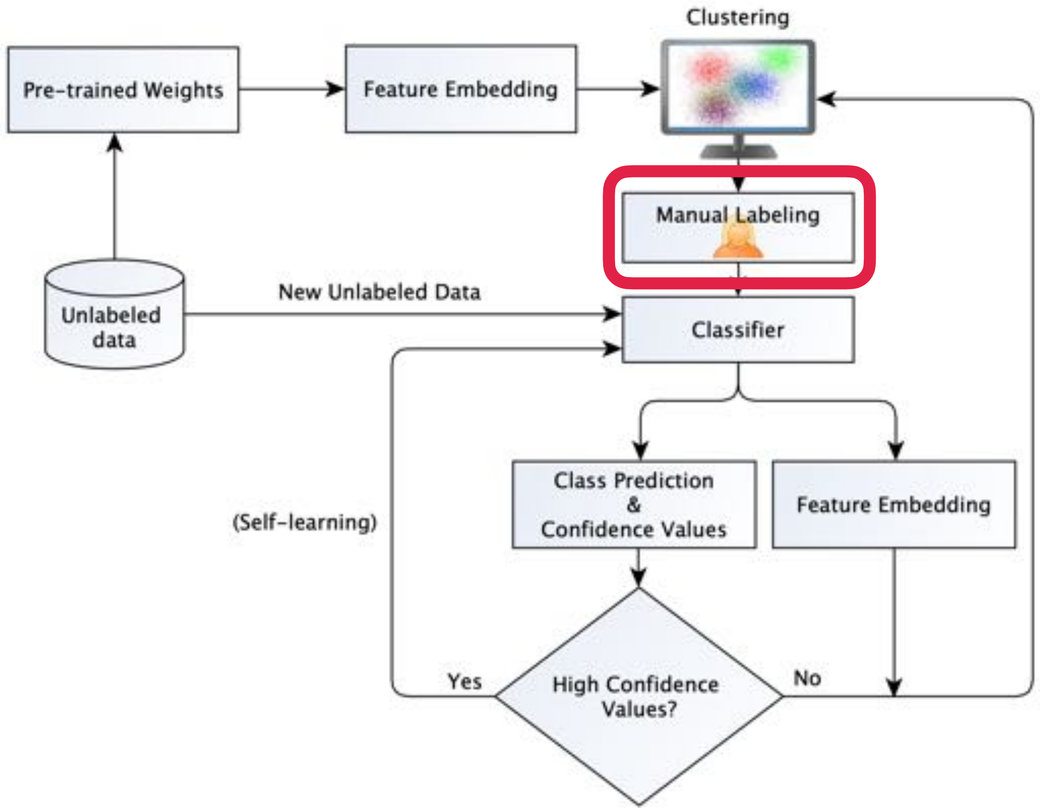
Plud: Machine-assisted labeling

- Iterates over
  - Clustering
  - Human supervision
  - Classification
- Objective: accelerate and simplify manual labeling



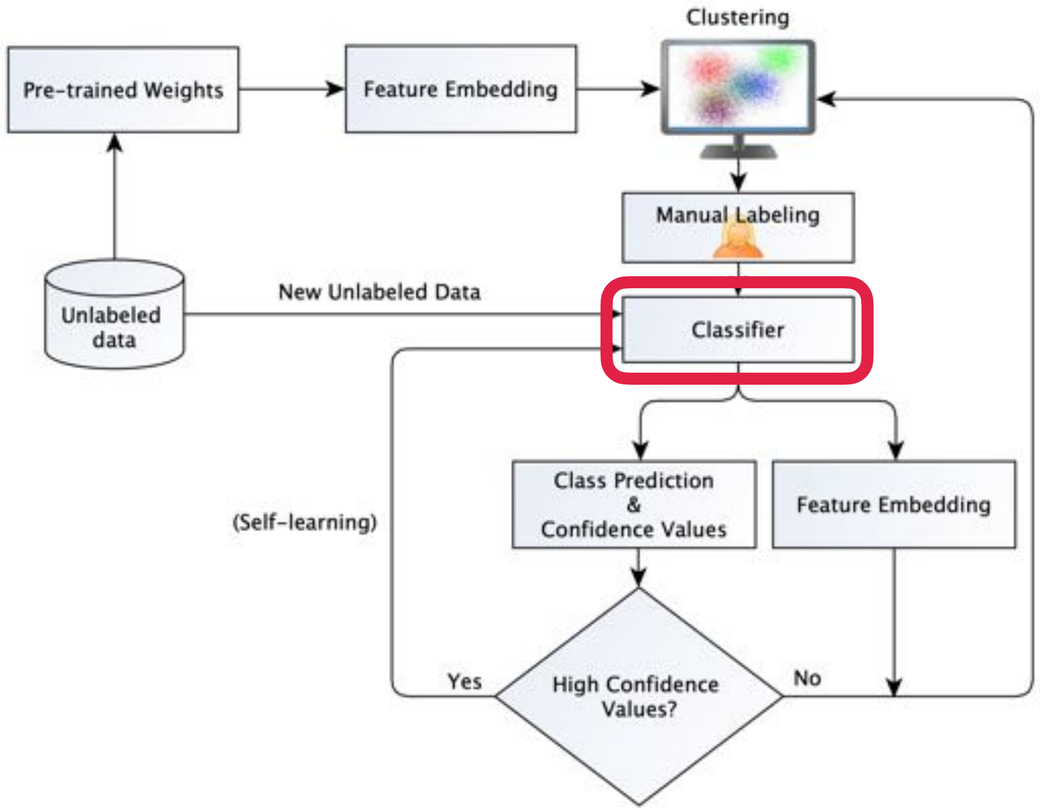
Plud: Machine-assisted labeling

- Iterates over
  - Clustering
  - Human supervision
  - Classification
- Objective: accelerate and simplify manual labeling



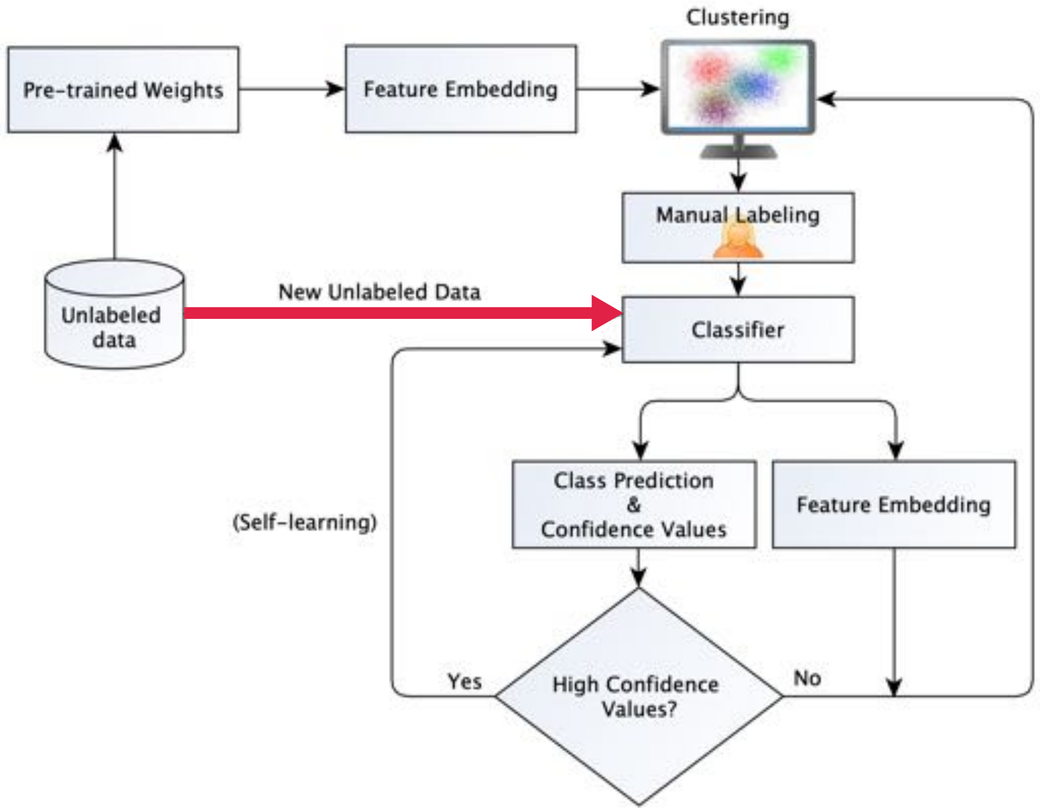
Plud: Machine-assisted labeling

- Iterates over
  - Clustering
  - Human supervision
  - Classification
- Objective: accelerate and simplify manual labeling



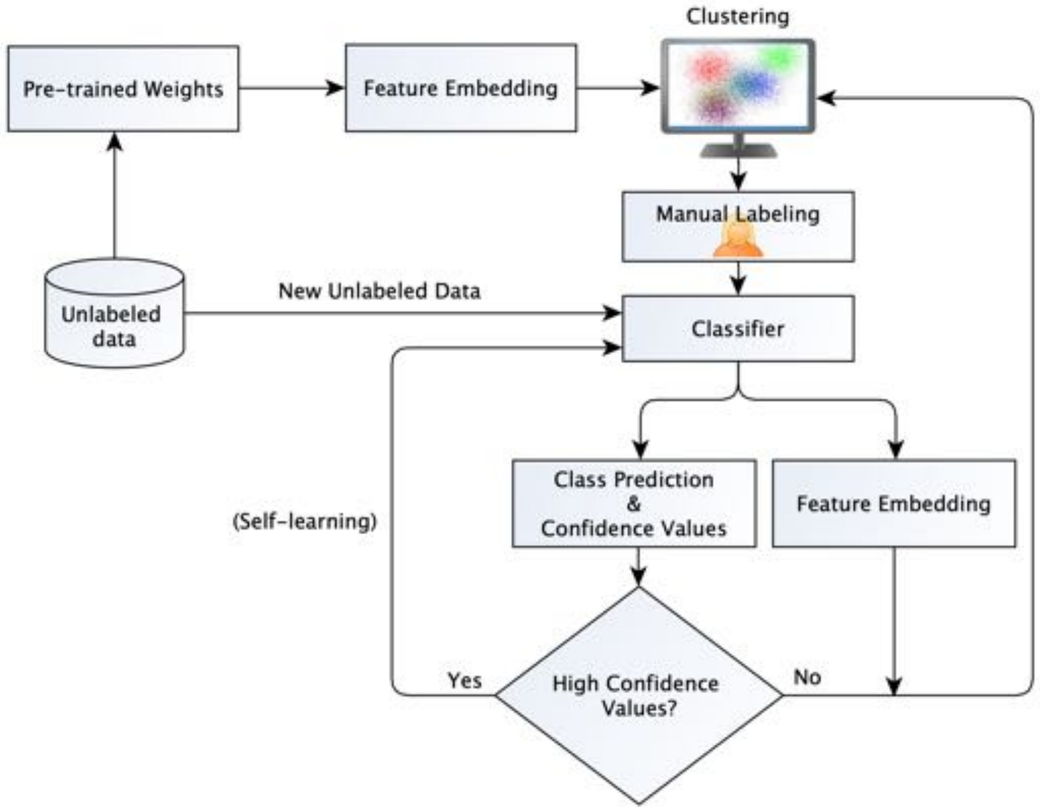
Plud: Machine-assisted labeling

- Iterates over
  - Clustering
  - Human supervision
  - Classification
- Objective: accelerate and simplify manual labeling



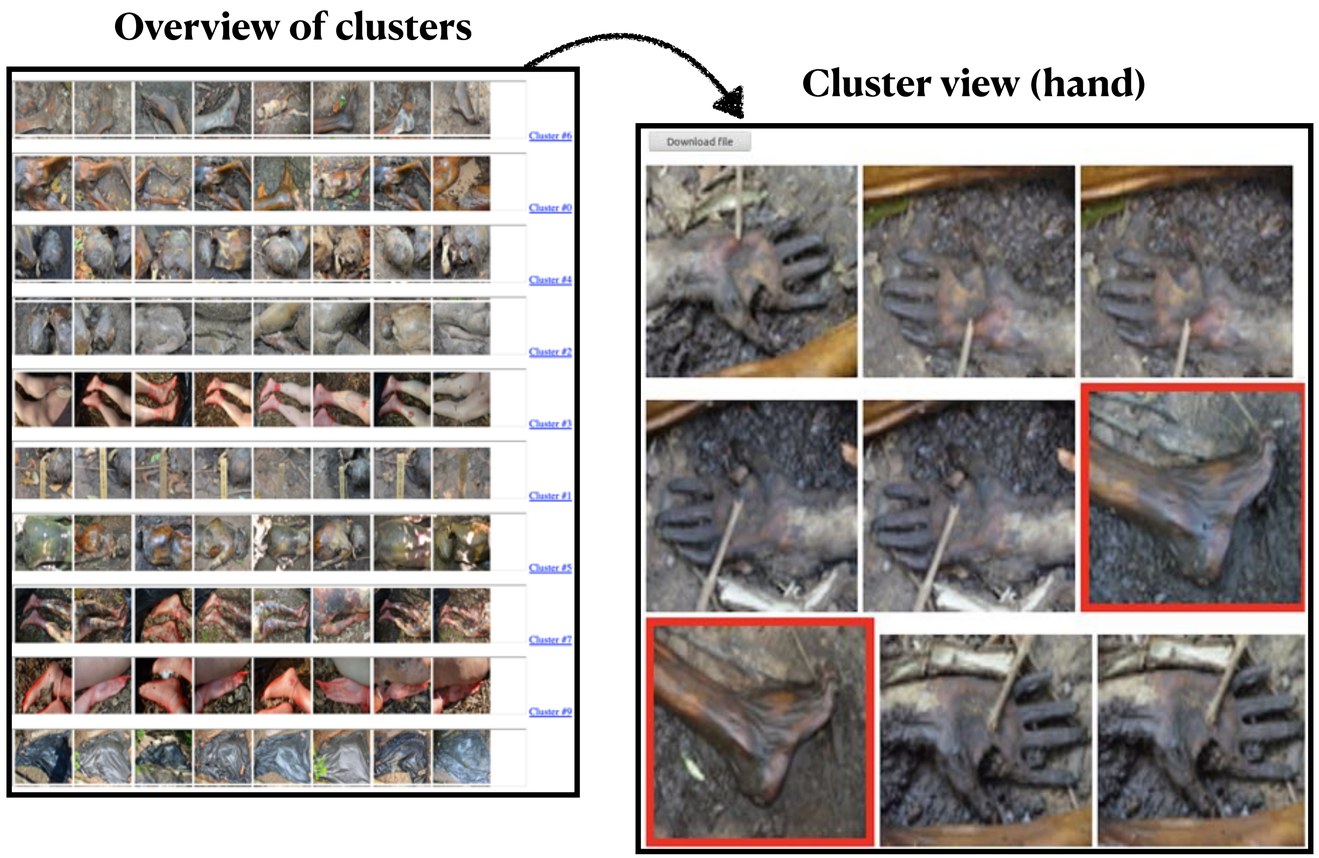
Plud: Machine-assisted labeling

- Iterates over
  - Clustering
  - Human supervision
  - Classification
- Objective: accelerate and simplify manual labeling



Cluster evaluation interface

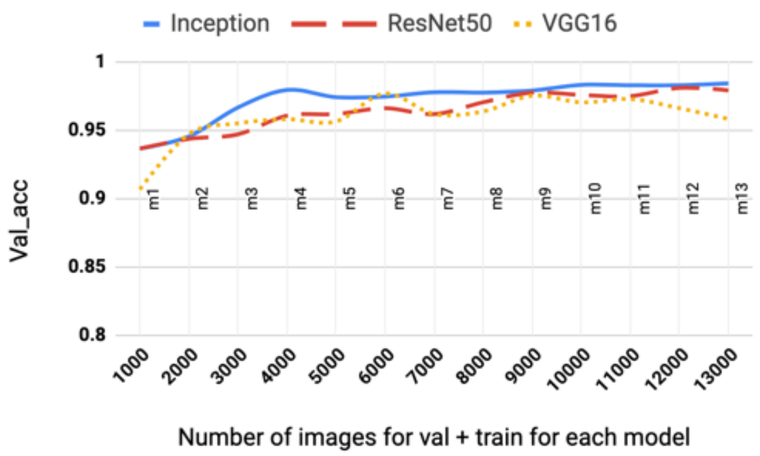
- Provides an overview of clusters
- Experts can remove mis-clustered images
- Experts label an entire cluster of images
- Labeling time and effort is reduced

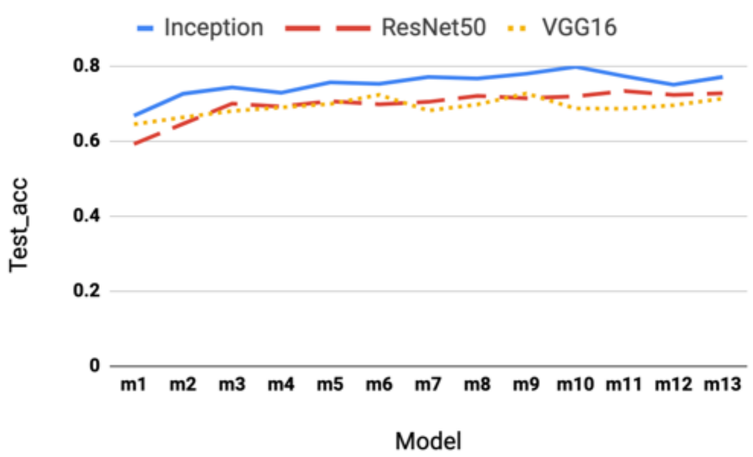


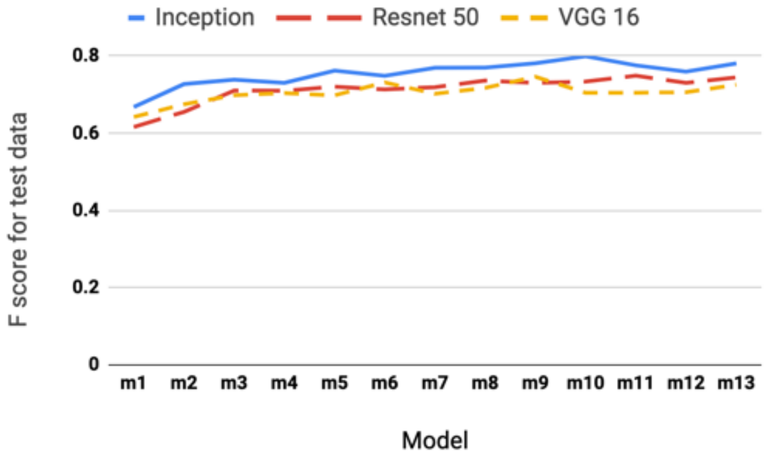
	150 images	300 images
Labeling	13m.22s	43m.52s
Labeling via Plud	4m.41s	11m.34s

Plud - F-Score

- 5500 manually labeled test images
- Inception outperformed RestNet50 and VGG16







Plud - Precision and recall

- Inception based classifier
- Tested on 5500 images of human decomposition

Model		Precision of Classes										AP
		Arm	Hand	Foot	Legs	Full Body	Head	Backside	Torso	Stake	Plastic	
Inception	Top 1	45.73	85.60	93.72	60.52	92.69	94.33	68.25	87.22	96.30	73.61	**79.80**
	Top 3	80.23	96.86	97.75	86.96	97.53	98.29	89.57	97.20	98.87	88.88	**93.21**

Model		Recall of Classes										AR
		Arm	Hand	Foot	Legs	Full Body	Head	Backside	Torso	Stake	Plastic	
Inception	Top 1	53.88	67.36	62.34	97.60	77.21	94.91	55.84	91.97	98.86	1	**80.00**
	Top 3	92.69	87.63	90.66	99.57	96.34	99.75	81.81	96.27	1.	1.	**94.47**

Conclusion

- Speeds up the labeling and curation process in large image collections when
  - No prior labeled data exists
  - Classes are vastly different from common datasets
  - Human supervision and expertise is required
- Enables fringe domains put their image data to use

Future work

- Provide label suggestions for the expert to validate
- Removing the human from iterations by developing an end-to-end method based on only a limited amount of domain expert input
- Expanding the image level labeling to image segmentation

Thank you!

mousavi@vols.utk.edu