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

Sara Mousavi¹, Dylan Lee¹, Tatianna Griffin², Dawnie Steadman², Audris Mockus¹

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

Funded by National Institute of Justice


July 2020
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
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

<table>
<thead>
<tr>
<th></th>
<th>150 images</th>
<th>300 images</th>
</tr>
</thead>
<tbody>
<tr>
<td>Labeling</td>
<td>13m.22s</td>
<td>43m.52s</td>
</tr>
<tr>
<td>Labeling via Plud</td>
<td>4m.41s</td>
<td>11m.34s</td>
</tr>
</tbody>
</table>
• 5500 manually labeled test images
• Inception outperformed RestNet50 and VGG16
Plud - Precision and recall

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

<table>
<thead>
<tr>
<th>Model</th>
<th>Precision of Classes</th>
<th>Recall of Classes</th>
<th>AP</th>
<th>AR</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Arm</td>
<td>Hand</td>
<td>Foot</td>
<td>Legs</td>
</tr>
<tr>
<td>Inception Top 1</td>
<td>45.73</td>
<td>85.60</td>
<td>93.72</td>
<td>60.52</td>
</tr>
<tr>
<td>Inception Top 3</td>
<td>80.23</td>
<td>96.86</td>
<td>97.75</td>
<td>86.96</td>
</tr>
</tbody>
</table>
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