Self-Supervised Representation Learning for Digital Agriculture

Sudhir Sornapudi

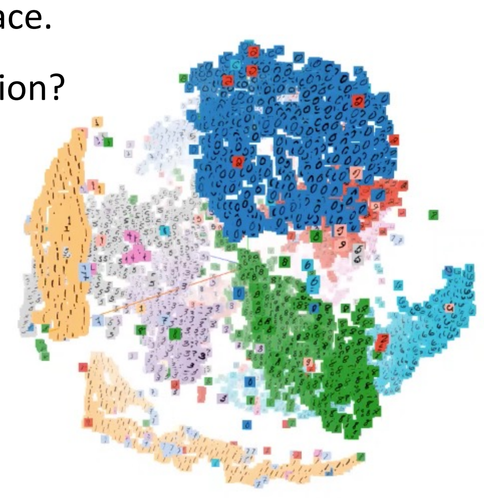
Farming Solutions & Digital, Research & Development, Corteva Agriscience, Johnston, USA

Problem Statement

- Computer vision in agriculture is game changing, with applications in plant phenotyping, remote sensing, bioinformatics, aerial imaging and scanning, harvesting with machines, grading and sorting, etc.
- Bottleneck: Rely on the availability of large annotated datasets for machine learning and lack of efficient labelling approaches.
- Solution: Develop and evaluate novel approaches with self-supervised learning to learn meaningful feature representations from raw image data.

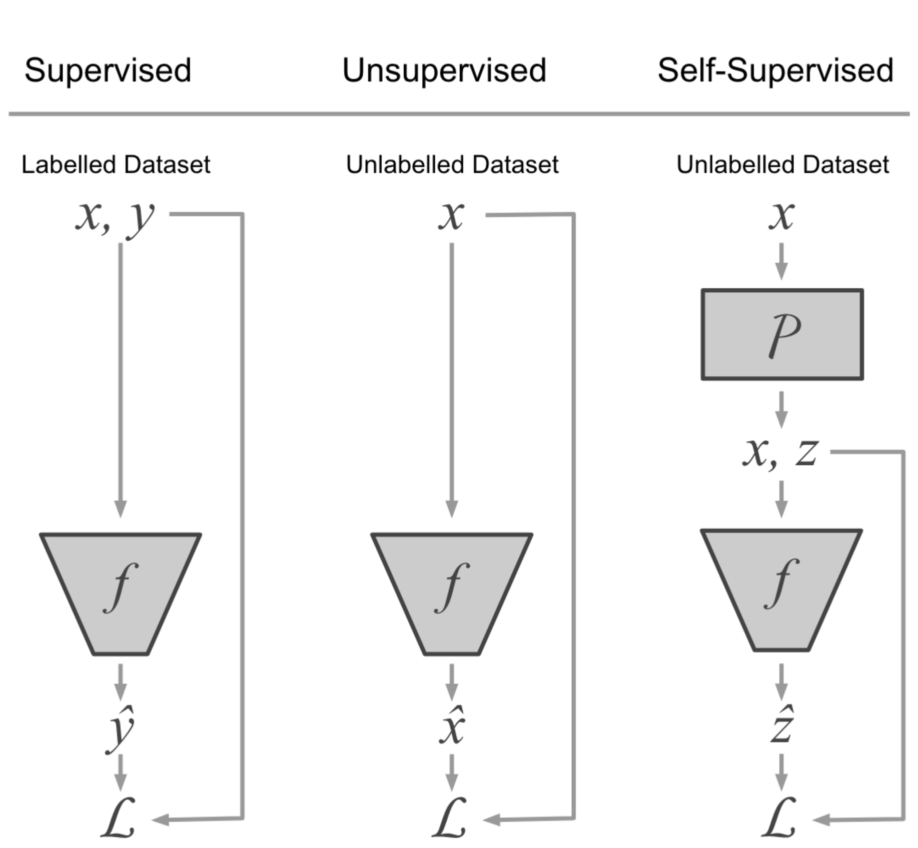
Representation Learning

- Representation learning is concerned with training machine learning algorithms to learn useful features.
- Deep neural networks can be considered representation learning models that typically encode information which is projected into a different subspace.
- What makes a good representation?
  - Low dimensional
  - Spatially coherent
  - Reusable across tasks
  - Disentangled
  - Hierarchical and meaningful

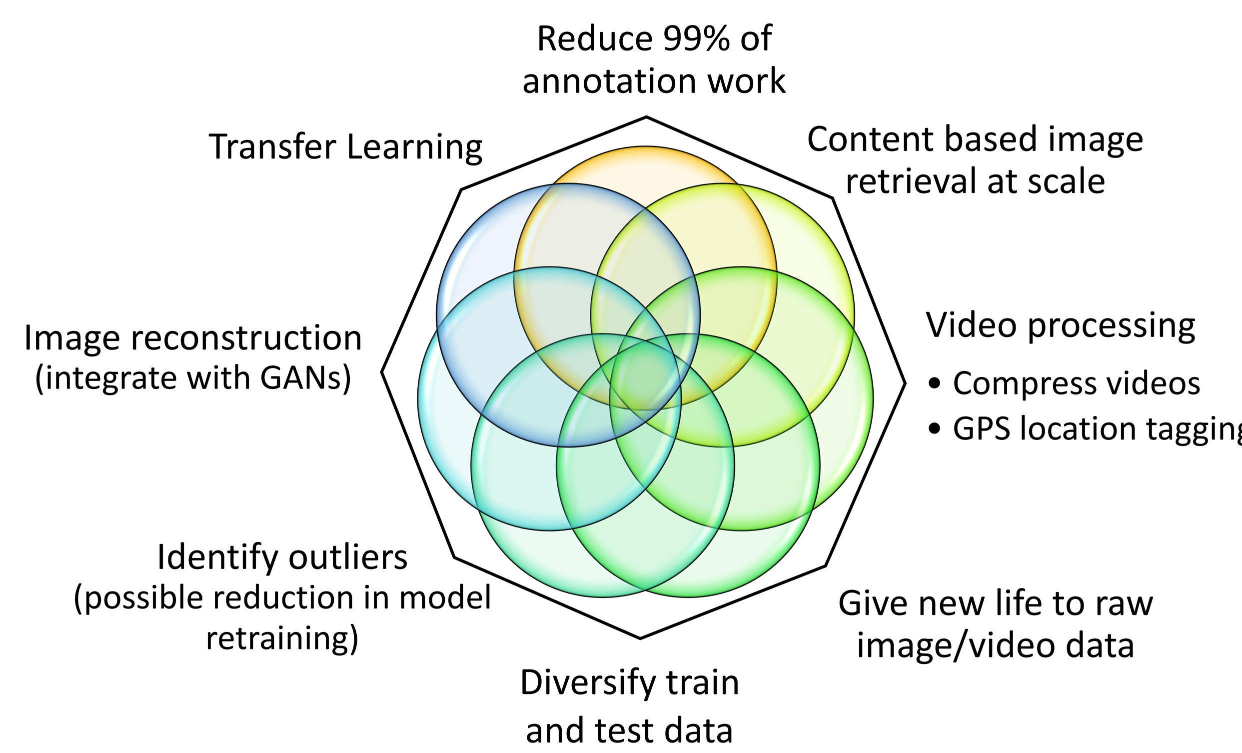


Self-Supervised Learning

- A special case of unsupervised learning.
- Self-supervised methods rely on pretext tasks to generate pseudo labels. This enables training in a supervised manner.
- General purpose representations are learnt during this training process that provide high performance and data-efficient learning of downstream tasks.
- The efficacy of self-supervised models approach and sometimes surpass fully supervised pre-training alternatives.
- We have utilized contrastive learning to create a pretext task while designing Contrastive Learning Representations for Agriculture Field Images (AgCLR) model.
- Contrastive learning is a self-supervised approach that enables model to learn attributes by contrasting samples against each other without the use of labels.



Potential Benefits

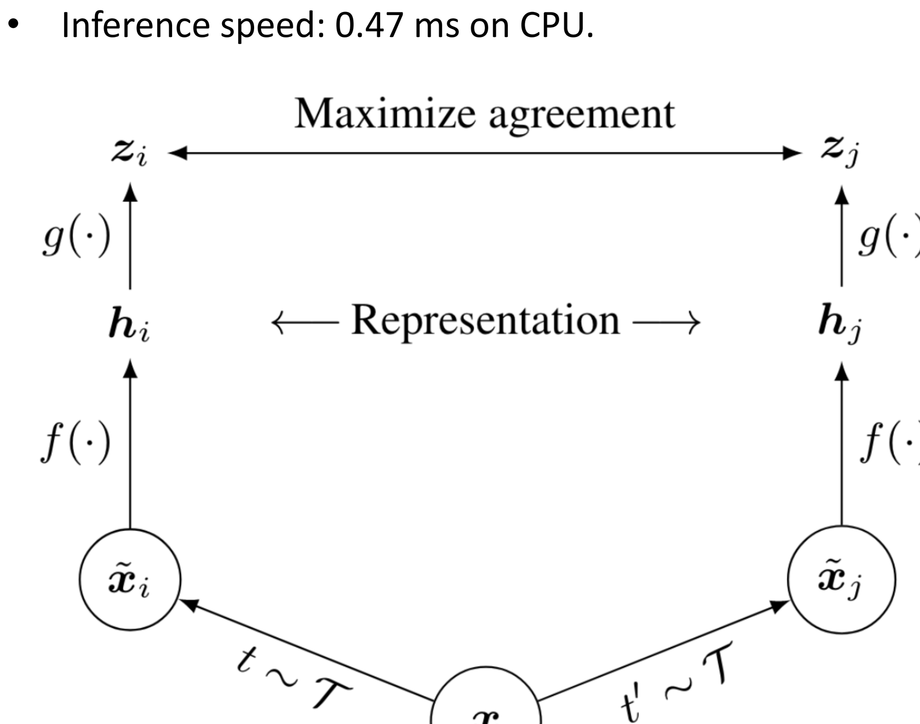


Data

- Real-world agriculture field data – 776,377 images including:
  - Corn field
  - Corn cob
  - Corn leaf
  - Soy field
  - Satellite view screenshot
  - Harvested field
  - Unknown field

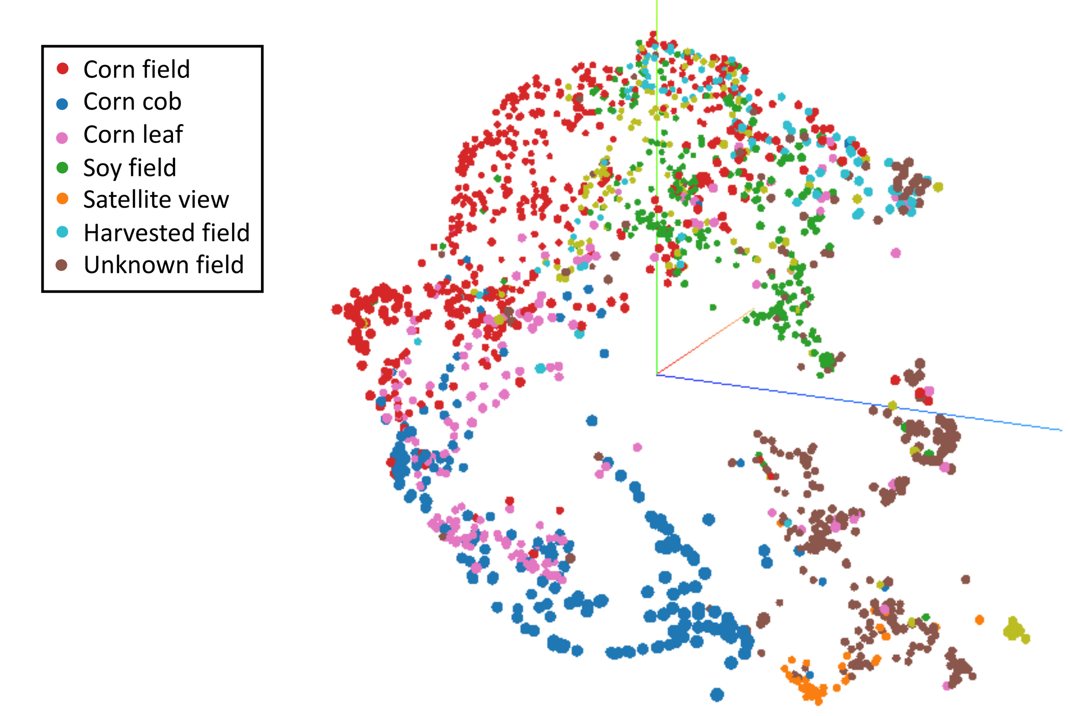
AgCLR Model

- AgCLR leverages SimCLRv2 framework to learn representations by maximizing the agreement between differently augmented views of same sample.
- Training phase:
  - Pretrained on entire dataset.
  - Fine-tuned on 1% of labelled data.
  - Critical enablers:
    - Mixed-precision
    - Distributed parallel computing
    - Hardware: TPU v3 8-core 128 GB.
- Testing phase:
  - 1864 images - labeled unseen test data.
  - Inference speed: 0.47 ms on CPU.



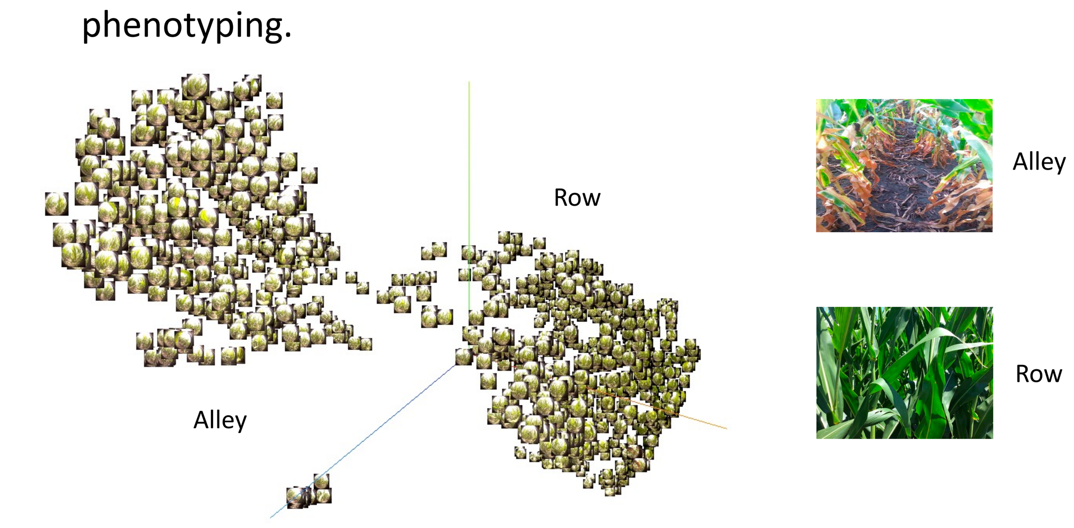
Results & Summary

- Achieved 80.2% accuracy on test data.
- The figure represents 3D point cloud of 1864 test samples. Each point is a 2048 feature embedding vector representing its input image.
- Cluster formations based on ground truth indicate the model's capability to learn the good representations and serve as an efficient automated labeling tool.
- As future work, we want to
  - experiment with deeper convolutional neural network (CNN) and vision transformer backbone networks.
  - explore the potential benefits and aim to reduce 99% of annotation work that saves time, money, workforce and speeds up model development and deployment process.

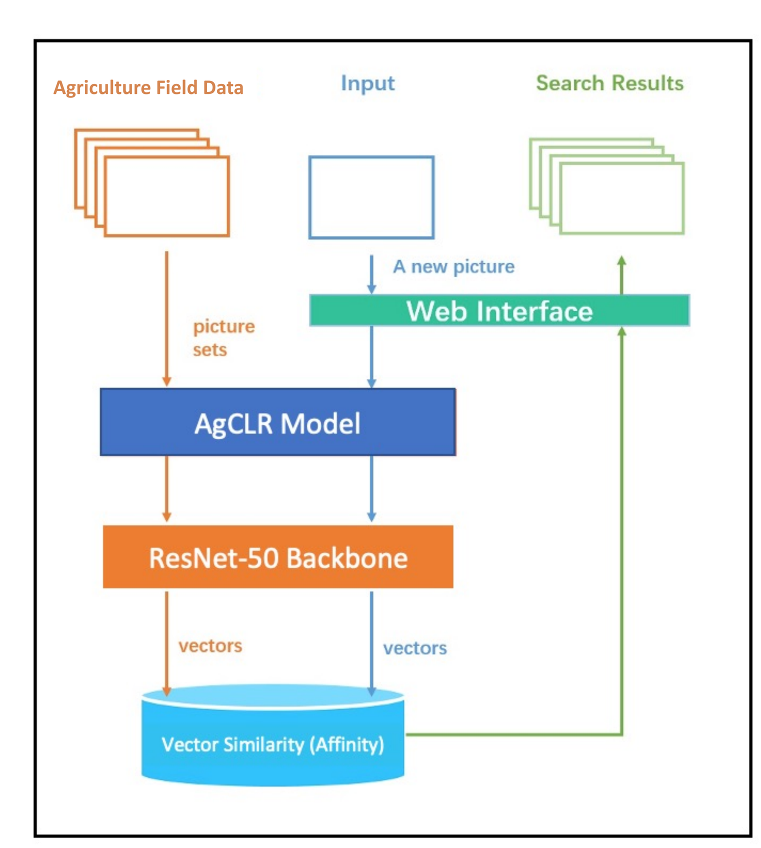


Applications

- Alley vs row determination in corn field videos for corn phenotyping.



- Pixel Affinity - Content based image retrieval tool.



References

Ericsson et. al., "Self-Supervised Representation Learning: Introduction, Advances and Challenges," IEEE SPM 2022.

Chen et. al., "Big Self-Supervised Models Are Strong Semi-Supervised Learners", NIPS 2020.

Abstract:
Published in NAPPN 2023 Abstracts.
DOI: 10.22541/au.167573772.22946655/v1

SCAN ME