Deployable Critical Active Learning (DECAL) System Offers Better Characterization for Improved Diagnostics and Imaging

Existing active learning systems lack multi-modal data annotation capability

Conventional systems for active learning aid in decreasing the labor, time, and cost normally associated with manual data labeling and operate effectively with well-established data representation. These existing active learning systems, which are used for medical diagnostics, subsurface imaging, and other applications where annotations are expensive to obtain, rely primarily on natural image data and fail to incorporate useful and critical data from other modalities. In the medical field, the other modalities include clinical labels, demographic information, and biomarkers. In subsurface imaging, they include well logs and geology and reservoir characterizations. Data samples in such applications can vary extensively. For example, pathologies have visually diverse expressions across different patient types, and subsurface structures differ from one site to another.

This unbalanced dataset leads to an underdeveloped data representation space and poor model generalization. In medical diagnostics, this leads to an inadequate evaluation for disease manifestations. Physicians are presented with incomplete assessment results that prevent them from understanding, interpreting, and correcting model predictions—an unwanted and potentially dangerous situation. In subsurface imaging, a lack of diverse data affects accurate well potential and productivity assessment.

DECAL system enables large training set initialization for greater predictive performance

Incorporating an algorithm that utilizes subject awareness, this system overcomes the lack of multi-modal labeling present in current technologies. It utilizes DECAL, a framework within a bi-modal interface that integrates multi-modal data annotation. DECAL is a plug-in system that can be added to any underlying AI system. It is intended to initialize existing active learning algorithms and guide them in finding the most desirable data subset for labeling. The DECAL algorithm is unique in sampling important, unlabeled data. When utilized during the initialization period, it creates a framework that can be expanded and generalized to any data

method and application. Substantial labor, time, and cost savings are achieved in comparison to labeling an equal amount of data via manual methods.

When applied to a medical diagnosis application, it allows natural image–based algorithms to generalize more quickly and accurately, resulting in appropriate medical diagnosis and interventions.

When DECAL is applied in subsurface characterization, it is instrumental in guiding the system to appropriate sample selection and initialization for accurate well productivity assessments, exemplifying its wide-ranging capabilities.

This technology is related to [9098 "Multi-Modal Deep Learning Training Model"](#)

Summary Bullets

- The deployable critical active learning (DECAL) system's bi-modal interface enables better characterization with large training sets that generate more accurate output for higher precision.
- The algorithm used during the initialization period creates a framework that can be expanded and generalized to any data method and application.
- The DECAL algorithm samples uniquely important, unlabeled data to aid existing active learning systems in generalizing more quickly, providing faster diagnosis statistics.

Solution Advantages

- **More accurate:** DECAL's bi-modal interface enables better characterization of phenomena (e.g., disease states and well potentials) by using large training sets to generate more accurate output.
- **Faster:** The DECAL algorithm uses a bi-modal interface to improve generalization speeds in existing active learning systems.
- **Less expensive:** Labor, time, and cost savings are significant with DECAL labeling strategies in comparison to labeling an equal amount of data via manual methods.
- **Flexible:** When utilized during the initialization period, this training algorithm can be used for any type of data and application.

Potential Commercial Applications

- Medical Applications:
  - Medical image interpretation
  - Clinical disease detection
  - Personalized medicine
  - Deployable artificial intelligence for critical safety interventions
- Subsurface Applications:
  - Subsurface characterization
  - Carbon capturing
- Any applications having expensive and laborious annotations

Inventors

- Dr. Ghassan AlRegib
  Professor – Georgia Tech School of Electrical and Computer Engineering
- Mohit Prabhushankar
  Graduate Student – Georgia Tech School of Electrical and Computer Engineering

- Ahmad Mustafa
  PhD Student - Georgia Tech School of Electrical and Computer Engineering
- Yash-yee Logan
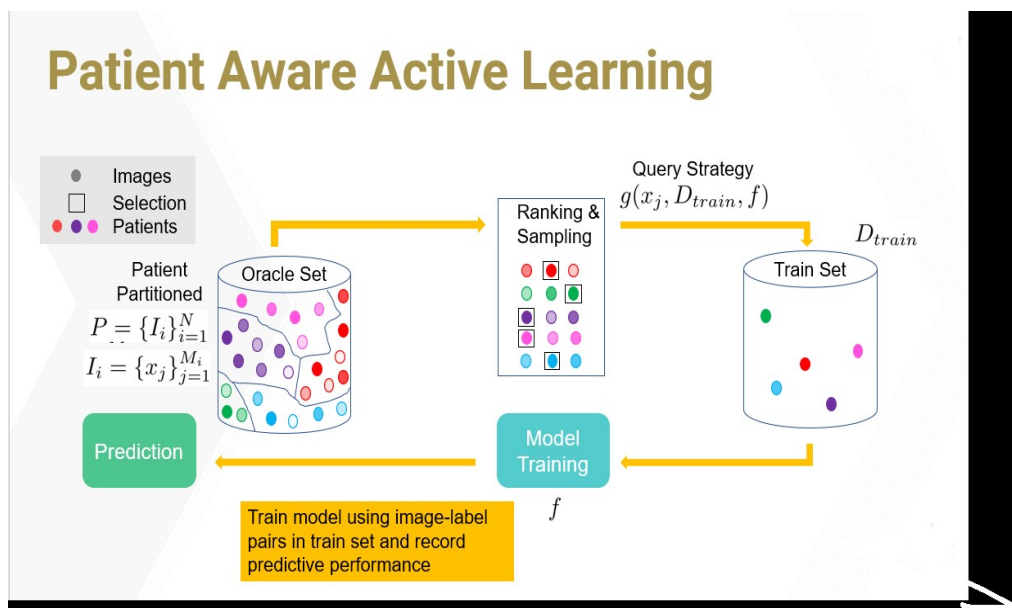 Graduate Research Assistant - Omni Lab for Intelligent Visual Engineering and Science

IP Status

<p>Patent application has been filed</p>: US63/384316

Publications

[Active learning with deep autoencoders for seismic facies interpretation](), GEOPHYSICS - August 2023

Images

This workflow shows the application of an active learning framework to the problem of subsurface characterization using a 3D seismic volume. The algorithm ranks all unlabeled images in the volume in order of increasing prediction uncertainty. It then samples the images with the highest uncertainty to be labeled by the expert.

Visit the Technology here:
[Deployable Critical Active Learning (DECAL) System Offers Better Characterization for Improved Diagnostics and Imaging]()

https://s3.sandbox.research.gatech.edu//print/pdf/node/4114