

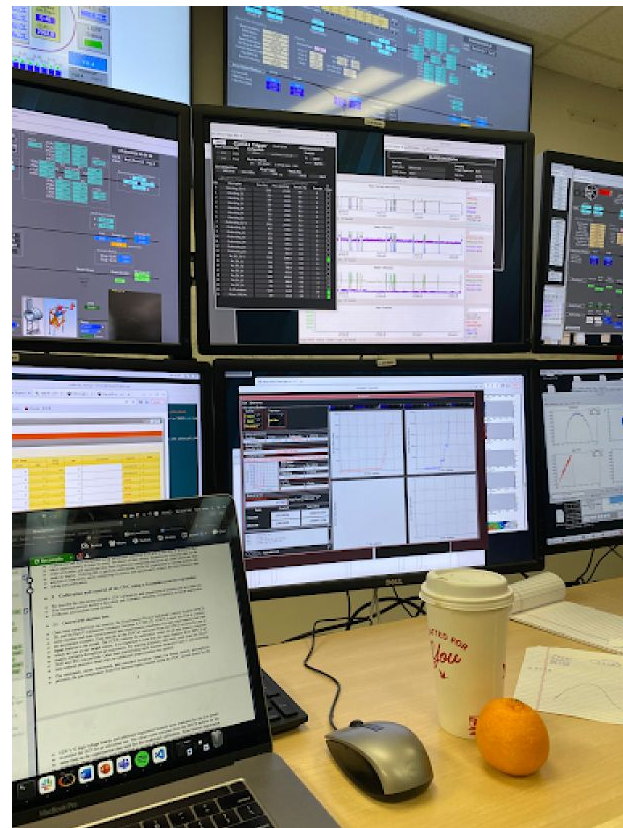
Jefferson Lab



Thomas Britton | tbritton@jlab.org



Developed and maintained by the Experimental Physics Software and Computing Infrastructure group





Services Layer (Python/Tensorflow):

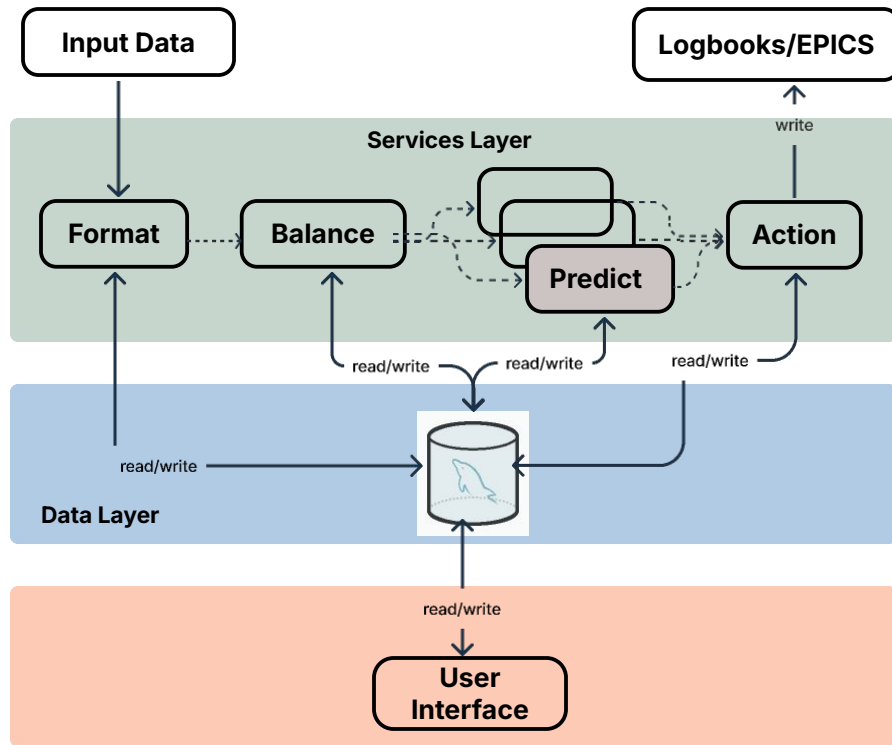
Convolutional networks (InceptionV3, VGG16, build your own), Siamese networks, Clustering Algorithms, Similarity Algorithms

Data Layer (MySQL):

Comprehensive MySQL database, optional Run Conditions Database integration. **Makes the data AI-ready.**

User Interface (React/Flask):

Web application for visualizing data, model classifications, labeling, performance metrics



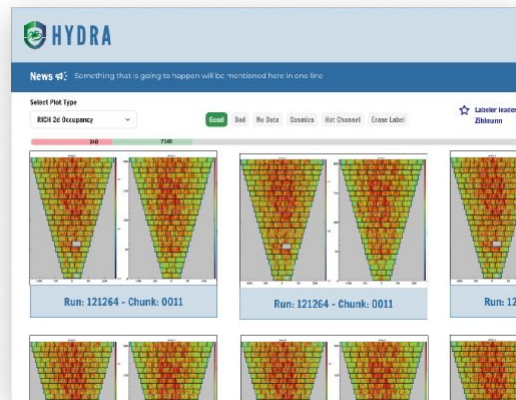
Primary Use Cases

The validation working group is using Hydra to develop an AI-ready image dataset for different stages of data quality monitoring for ePIC.



Visualization

View, sort, and filter images.



Offline Monitoring

Validate monitoring plots from simulation campaigns, calibration changes, etc.

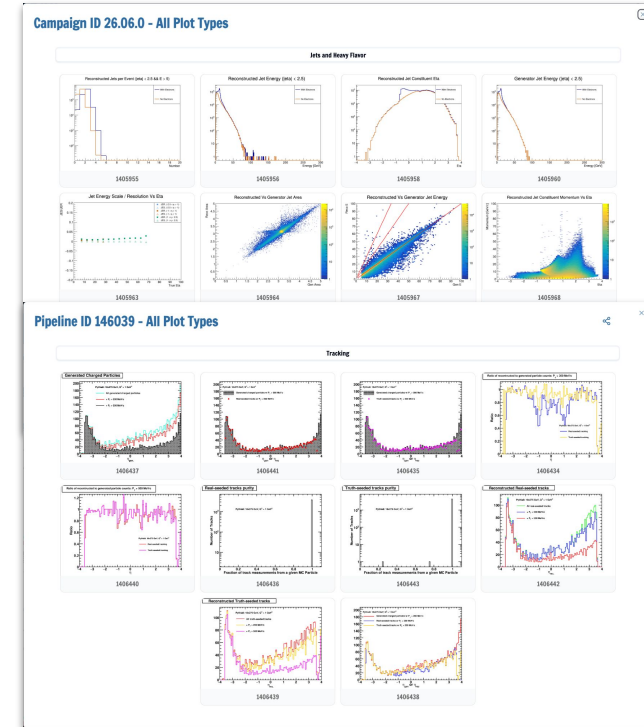


Online Monitoring

Use neural networks to classify monitoring images in near real time.

Plot Browser for Visualization

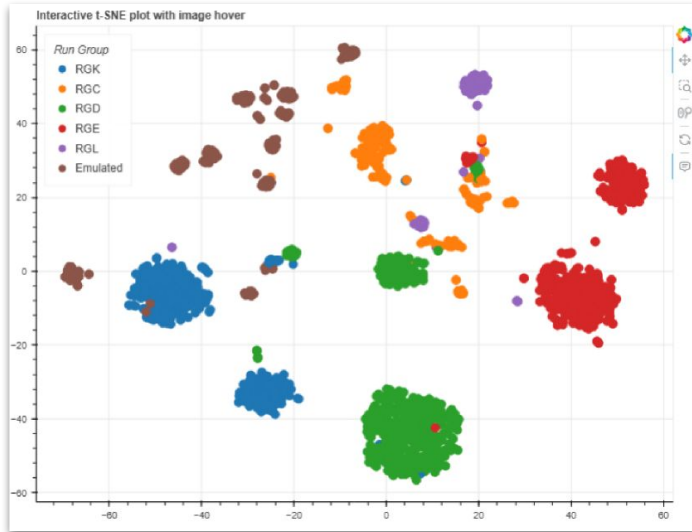
View, filter, and sort plots based on monitoring versions, working groups, and simulation campaigns.



Automated Validation of Images

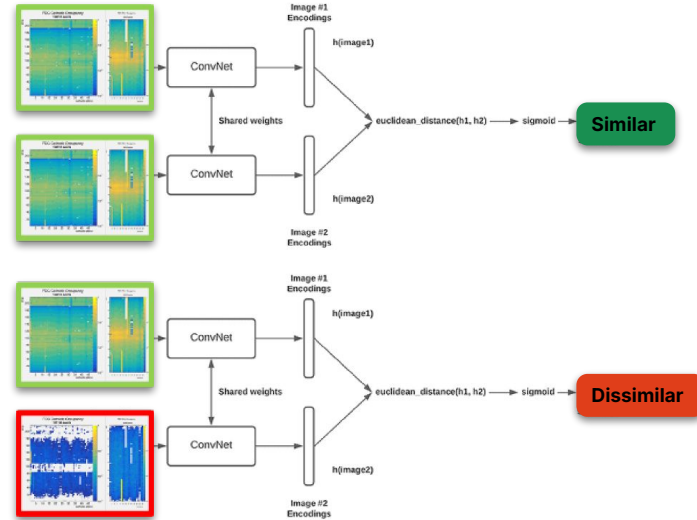
Hydra natively employs unsupervised and supervised learning techniques to cluster and classify images.

Cluster *Unlabeled* Images



t-distributed stochastic neighborhood embedding: similar objects are modeled by nearby points and dissimilar are modeled by distant points.

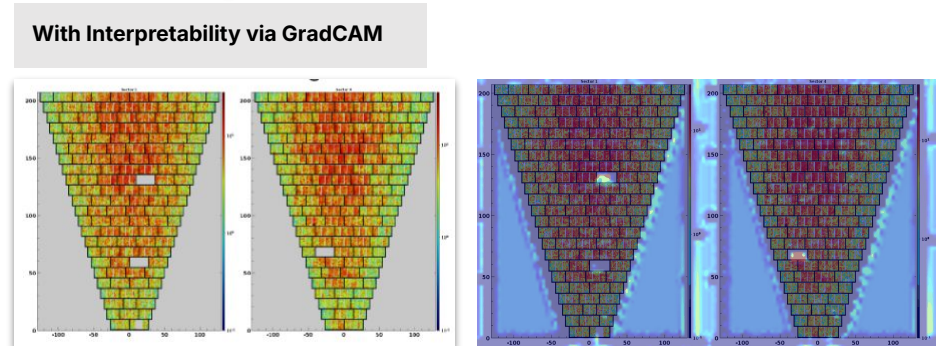
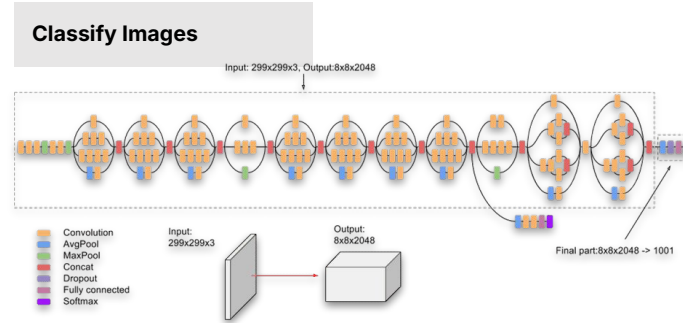
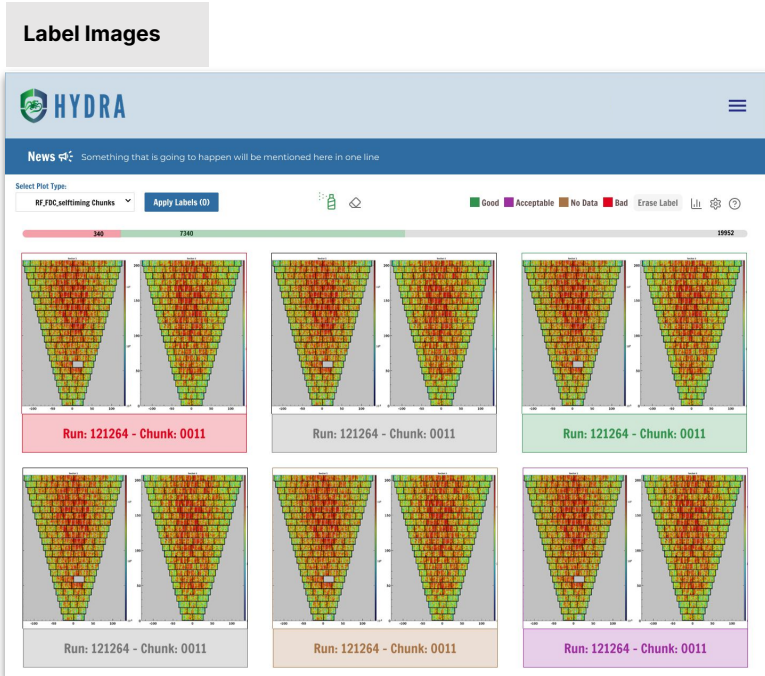
Identify Similar Images with Siamese Architectures



Hydra supports reference images that are used in Siamese architectures to evaluate the similarity of an image to a reference. *Foundational model for plot comparisons is in progress.*

Automated Validation of Images

Hydra's built in labeler provides a simple interface to efficiently label thousands of images. These can then be used in training convolutional neural networks to classify images in real time.



Gradient-weighted Class Activation Maps: Bright spots indicate important regions of the image for the particular classification.

Conclusions

Organize, store, visualize, and classify image data

The validation working group will work with you to get more images into Hydra.

AI-based Data Quality Monitoring

Hydra can validate simulation campaigns and artifacts from the CI system using supervised and unsupervised techniques.

Code is hosted on GitLab

code.jlab.org/epsci/hydra
hydrateam@jlab.org

Hydra team



David Lawrence

EPSCI Physics, CS, AI/ML



Thomas Britton

EPSCI Physics, CS, AI/ML



Torri Jeske

EPSCI Physics, Operations, Full
Stack Development



Nataliia Matsiuk

Information + Records
Containerization, QA



Raiqa Rasool

EPSCI CS, Full Stack
Development



Darren Upton

ODU Physics Graduate Student
MLOps, Data Augmentation



Jordan O'Kronley

UT-Knoxville Physics Graduate
Student Siamese Models

Sachin Gupta Tobias
Stockmanns
CBM experiment at GSI

Nathan Baltzell
Casey Morean
Naomi Jarvis
Christine Ploen
Jefferson Lab

