ASME-CIE Hackathon: Patterns in the Noise, Exploring the AFRL FactoryNet Dataset

FactoryNet

Problem Statement:
The most successful machine vision AI systems require large amounts of accurately labeled images for training. The AFRL FactoryNet dataset aims to build the highest quality image dataset for the manufacturing domain. By sourcing images from web scraping, photographing machine shops and factories, and receiving contributions from manufacturing partners FactoryNet has a large volume of images with wide ranging subject matter. Building this into the most useful dataset possible comes with many challenges in defining the scope, organization, and curation strategies. One such challenge, and the current development stage of the dataset, is collecting and organizing the human labels for the images. Where does the most value lie in a dataset of images with multiple unstructured labels? One way to assess the emerging value and gain insight into the patterns and trends that will guide dataset development/expansion as well as ontology design is to test the images ability to inform a classifier. In the case of a dataset with an open ended amount of labels, the classes can be defined in many ways and finding classes that stand out as well defined and successfully labeled will reveal strengths and weaknesses in the dataset.**The challenge presented is to organize/consolidate freeform labels and create classifiers that show the image recognition capabilities enabled by this dataset.** Participants should aim to sanitize and propose structure to the image label data in an intuitive way that can enable image classification. Post-organization the participants should aim to show that classes yielded are useful and accurate by demonstrating they provide enough information to train and validate a classification model. The topic areas of this problem are data sanitization, dataset design, and creation/use of AI/ML models.

Goal:
Successful submissions will, firstly, sanitize and structure human generated image labels. Secondly, the submission will create one or more image classification models and report results that meaningful and accurate classes have been curated.

Challenges:
Several challenges lie ahead when curating the labels to select classes and choosing how to prove the efficacy of the class choices with image classification trials.
Data organization challenges:
- Specificity: Some labels may be broad and some may be specific. Specific labels may be more consistent across the images they tag and broader labels may be too general to create a consistent image class. Finding this balance is key when choosing which to inform models as a ground truth.
- Classes may overlap: Since each image is tagged with many labels some classes may

contain the same images. In these cases showing the classes as mutually exclusive will be impossible. Strategies for leaving out images that have mutual classes or using the number of times the tag was submitted may be ways to show the classes contain meaningful differences.

Classification model challenges:
- Open endedness: We are not specifying a model type, the goal is to show that a class label is able to be used successfully with the emphasis on proving the label is significantly specific and contains enough images. This creates the challenge of choosing complexity vs efficiency.
- Subselection and Comparisons: Choosing how to compare the target classes is vital. It is an option to compare the class to all other images in the dataset, images outside the dataset, other classes you have determined already, etc. The type of comparison in the model will determine how strong the case for them being a "meaningful class" will be.

Strategies:

Data Sanitization:
- Label data will be as the users input it; this means typos, case discrepancies, and synonymous labels may be present. Try to clean the data in ways that lead to more meaningful classes and not separate similar classes.
- Choose wise targets and use an ontology or inference. Look and see if there are easy labels that are synonymous and can be grouped together, feel free to create rules that assign new labels that are informed by the existing labels or image data.
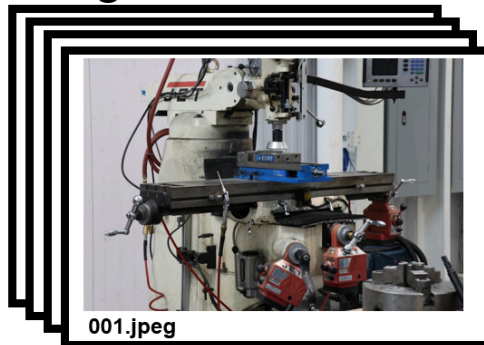
Image Classification:
- Transform the images to be consistent before modeling. Resolution and dimensions may need to be adjusted as well as filetype. Don't forget to rightsize the information for your chosen model.
- Each class needs to be proven meaningful. This can be done with models that compare the class against a subselection of random images from the dataset, the entire dataset, a different class etc. Complex multiclass models may not be ideal and too cumbersome, try smaller comparisons to make the case for significance and don't forget to avoid overfitting!
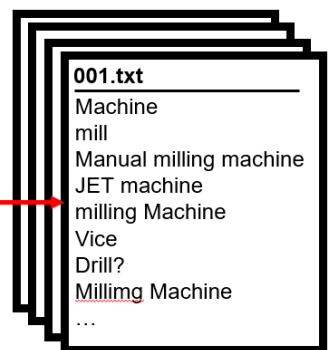
Dataset:

For this challenge the current state of the factorynest dataset will be provided upon launch. The Dataset will have at least 20,000 images and each image will be accompanied by all the human submitted labels it has been tagged with at the time of the challenge as plain text.

Image files

001.jpeg

Label data

001.txt
Machine
mill
Manual milling machine
JET machine
milling Machine
Vice
Drill?
Millimg Machine
…

Submission:
Teams will submit a presentation summarizing their methods and results as well as .CSV or other comparable data structure summarizing their quantitative results. Any code written for the challenge along with model weights and any other information needed for reproduction of results should be accessible via github.

Evaluation Criteria

Criteria	Description	Scoring
Data Organization Milestones **20% of Total Score**	The amount of classes elucidated with a significant amount of images. Scored relative to competitors	100% Most Classes Identified 90% Top several teams 80% Middle Rank 70% or less - Poor effort
Classification Model Milestones **20% of Total Score**	The quantity of classes proven accurate to a significant extent against a control. >75% accuracy min. Scored relative to competitors	100% Most Classes Identified 90% Top several teams 80% Middle Rank 70% or less - Poor effort
Methodology Choices **40% of Total Score**	Proficiency and creativity in class identification. Appropriate image classification models. Sufficient proof of accuracy: Appropriate comparisons and validation splits.	100% Creative, Accurate, and Communicated Well 90% Sufficient choices to accomplish task 80% Singular errors but decent effort 70% or less - Multiple errors
Presentation and Documentation **20% of Total Score**	Clear and effective presentation of the solution, including documentation and visuals.	100% Perfectly convey process and results 90% Small errors 80% or less - Several errors