Feature Extraction/Machine Learning for Degradation Classification of Solar Modules

Benjamin Pierce[1,3], Ahmad Maroof Karimi[1,3], Justin S. Fada[1], JiQi Liu[1,2], Jennifer L. Braid[1,2], Mehmet Koyutürk[3], Roger H. French[1,2]

[1]SDLE Research Center, Case Western Reserve University, Cleveland, Ohio

[2]Department of Material Science Engineering, CWRU, Cleveland, Ohio

[3]Department of Electrical Engineering and Computer Science, CWRU, Cleveland, Ohio

Case Western Reserve University, EST. 1826, think beyond the possible™

SDLE Solar Durability and Lifetime Extension Center

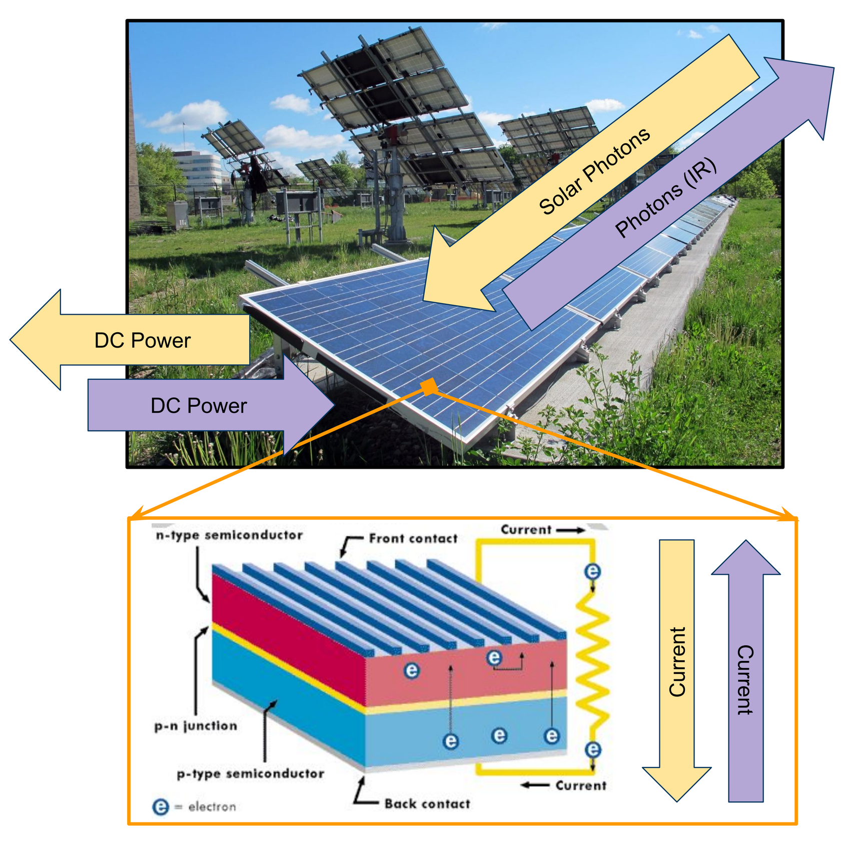
Objective

In order to identify degradation/faults in photovoltaic modules, a pipeline of electroluminescence imaging + machine learning has been developed

- **Electroluminescence (EL) images are taken in order to visually enhance/reveal faults in PV module**
- **Can identify these faults manually, but time consuming and expensive**
- **In order to automate this process, we combine classical computer vision techniques with machine learning**
- **The focus of this project is improving the feature extraction step**

Background

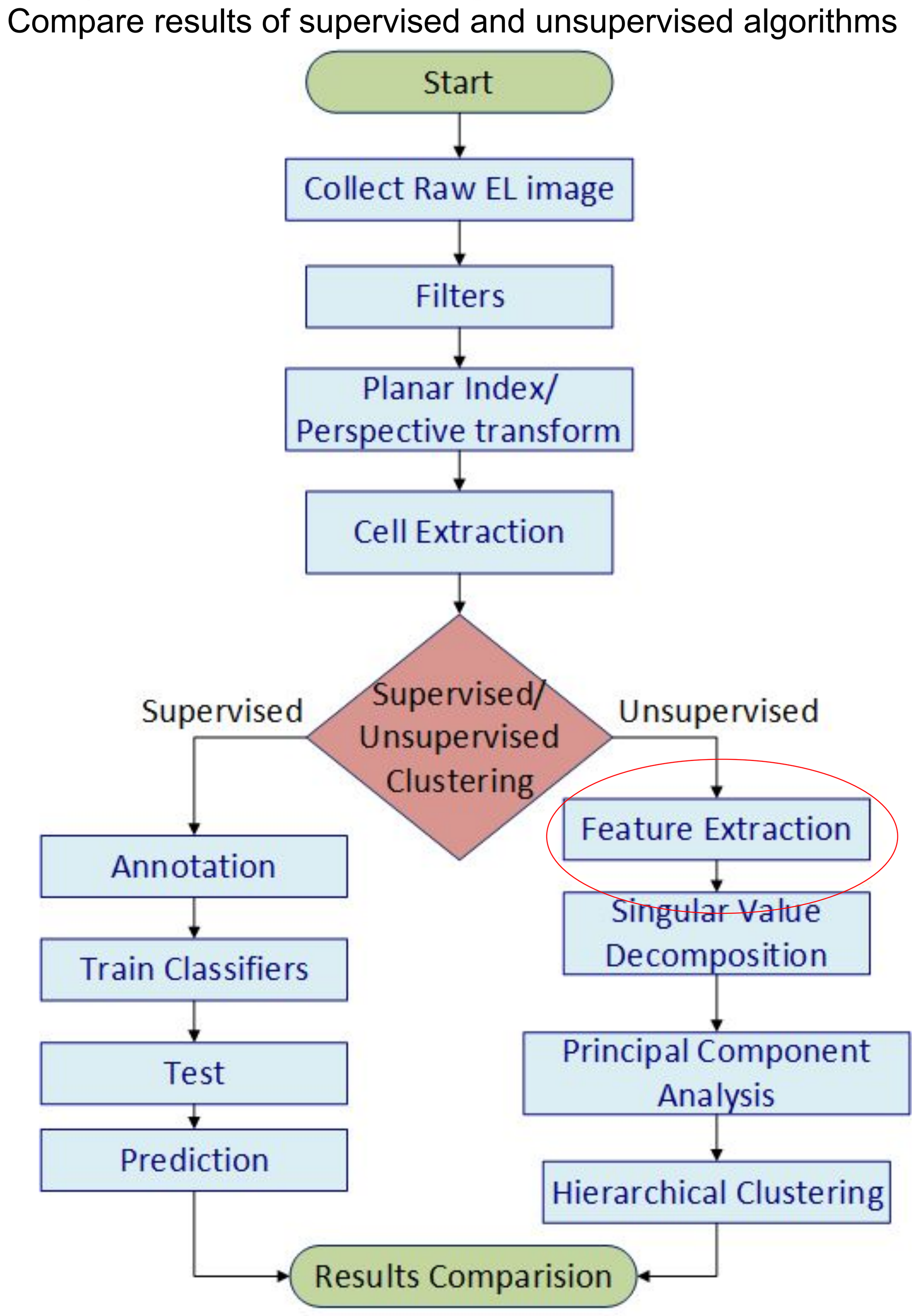
- PV generates electricity when light irradiates the surface
- PV cells emit infrared light, when powered with electric current
- Images are captured when current is fed into a cell
- Electric current generated is proportional to brightness of pixel
- Pixels intensity reflects the degradation of PV cell
- Process of Electroluminescence
  - Forward-bias→ IR LED
  - Current applied to pn-junction
  - Electrons recombine with holes
  - Non-optical emission
    - 1100 nm (infrared) for Si



- Leverage machine learning for pattern recognition on PV surface
- Study discernable textural features from EL images
  - Different orientation of textural pattern have different effect
  - Effect of size of the darkened pixels

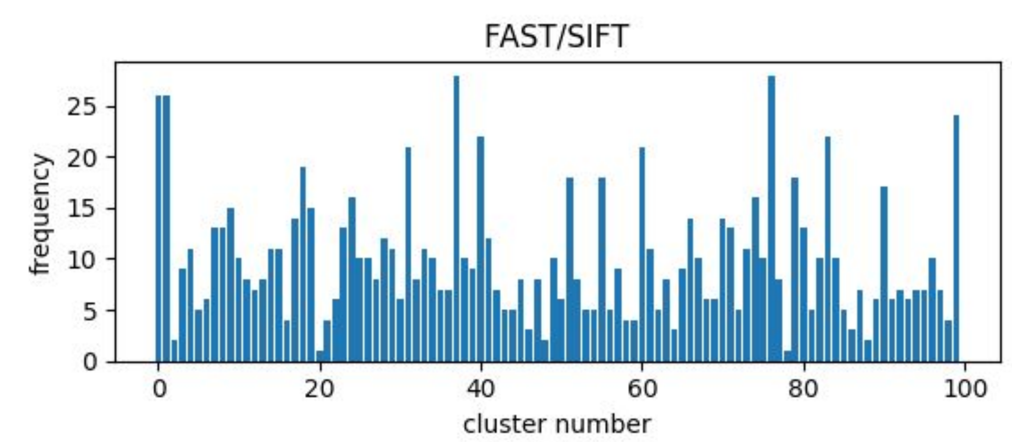
Pipeline

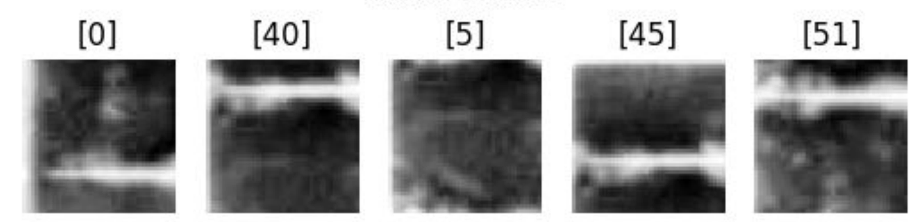
PVimage Python module major steps are:
- Image processing
- Supervised Machine Learning
- Unsupervised Machine Learning
- Compare results of supervised and unsupervised algorithms



Bag of Visual Words

- Can create a "visual dictionary" by clustering around feature vectors
  - K-means used to have a dictionary of consistent size
  - In example, n =100 cluster centers/ feature classifications

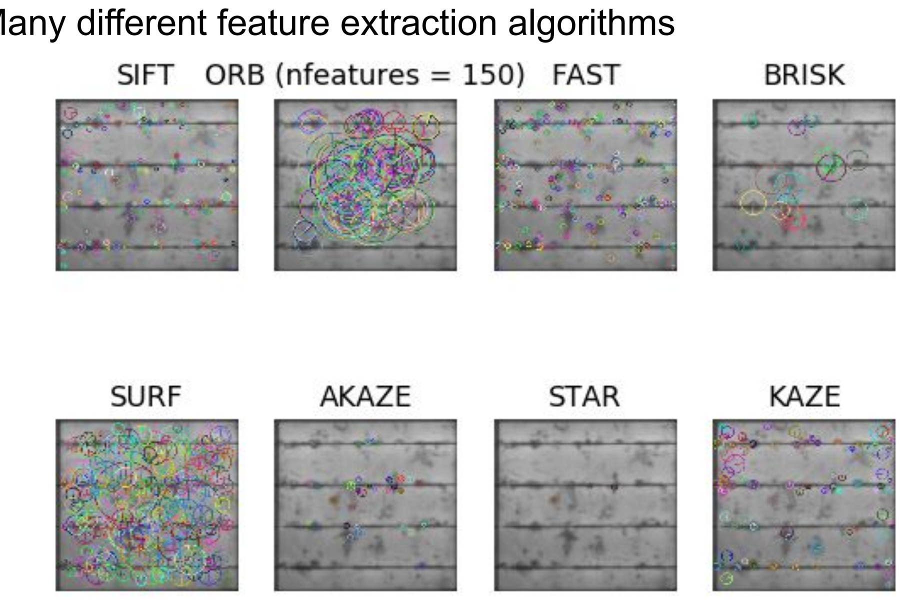




- Above is a dictionary from the combination of FAST keypoints and SIFT descriptors
- Below are examples of the features clustered around
  - All features in the cluster should be similar to each other
- We can create many combinations from the different algorithms and evaluate the combinations
  - For example, FAST/SIFT may be better than SIFT/SIFT
- Can then apply bag of visual words to classify features from other, similar images

Algorithms

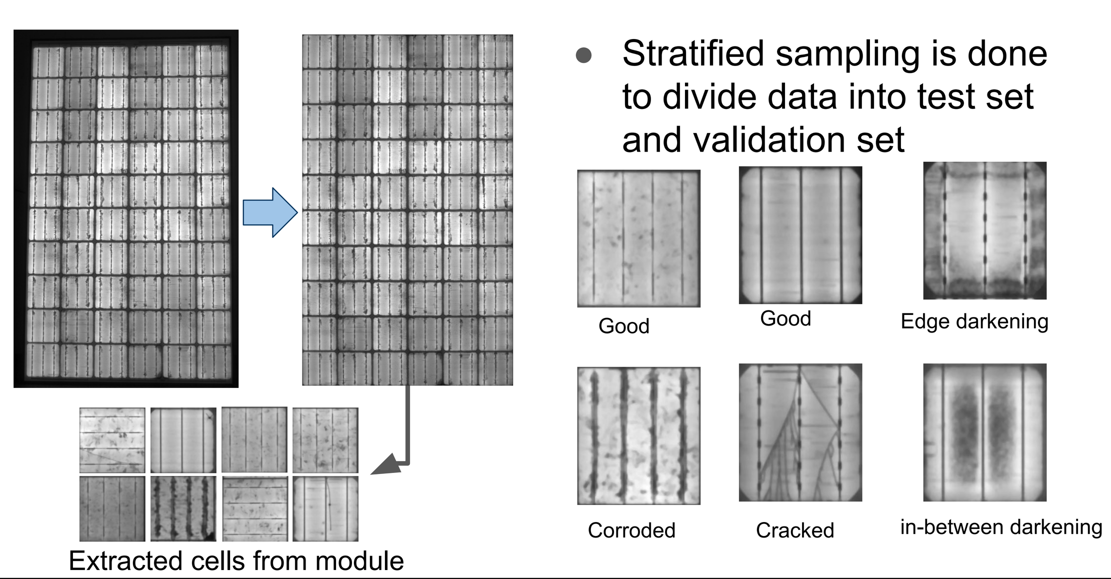
- Many different feature extraction algorithms



- Each method produces a feature vector/descriptor
  - Describes keypoint in terms of location, scale, rotation
  - For each image, vector is of dimension (#keypoints, 2^n)
    - n varies by algorithm, commonly 128
- SIFT/SURF
  - Earliest methods; viewed as stable
  - Uses Gaussian blurring
- KAZE/AKAZE
  - Based on nonlinear diffusion, not Gaussian space
    - To detect less distinct borders
- FAST/BRISK
  - Gener-ally used a corner detector
  - Good for detecting "sharp" shapes
- ORB
  - Uses FAST keypoints, but different descriptor
- STAR
  - SURF optimization for real-time use
  - Not used as it is over optimized for our purposes

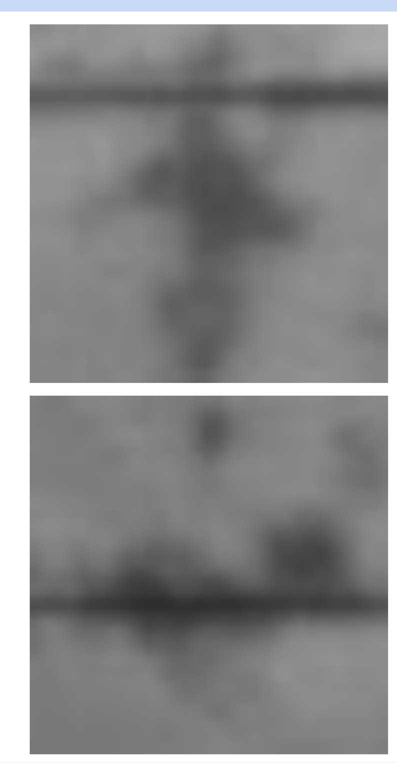
Cell Extraction and Annotation

- Functions applied
  - Planar index function
  - Cell extraction
- Labeled cells are input for Supervised machine learning classifiers
- Stratified sampling is done to divide data into test set and validation set



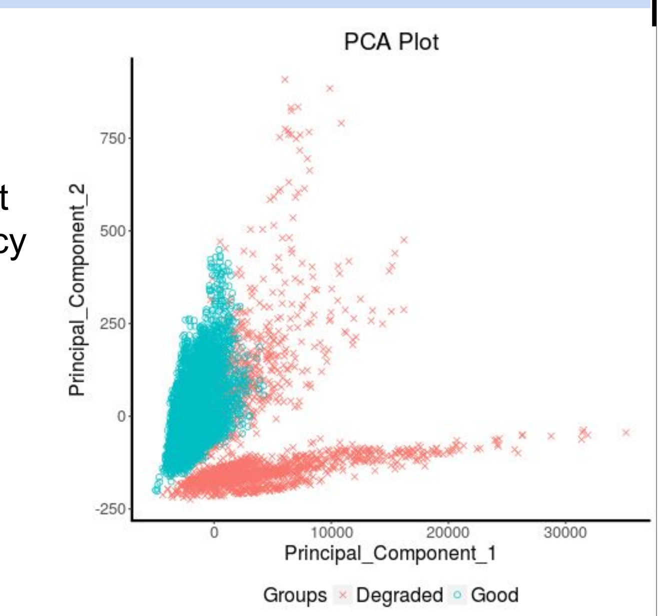
Motivation for Feature Extraction

- Machine learning on images is computationally expensive
  - EL images generally 500px x 500px
- Instead, use a subset of the image that is representative of important features
  - Reduces dimensionality vastly by discarding unimportant areas
- This process begins by recognizing key points and describing them quantitatively
  - Such as difference from background, presence of edge



Unsupervised Machine Learning

- Currently using 14 Haralick features
- Less accurate than CNN
- Motivation for this project
- is to improve the accuracy by improving feature recognition
- Singular Value Decomposition & PCA
  - used for dimensionality reduction
- Agglomerative clustering w/ Euclidean distance as similarity measure



Algorithm	Coherence
Hierarchical Clustering	78.2%

Dataset and Exposure Types

- 6264 cell images
  - Extracted from full size modules
- 5 brands
  - Distributed into 6 groups equally by all brands
- Six groups exposed to indoor accelerated exposure
- Accelerated exposure types:
  - Damp-heat
  - Thermal-cycling
  - Ultra-violet irradiance
  - Dynamic mechanical loading
  - Potential induced damage (PID) ±1000 V
- 80%/20% random test/train split for CNN

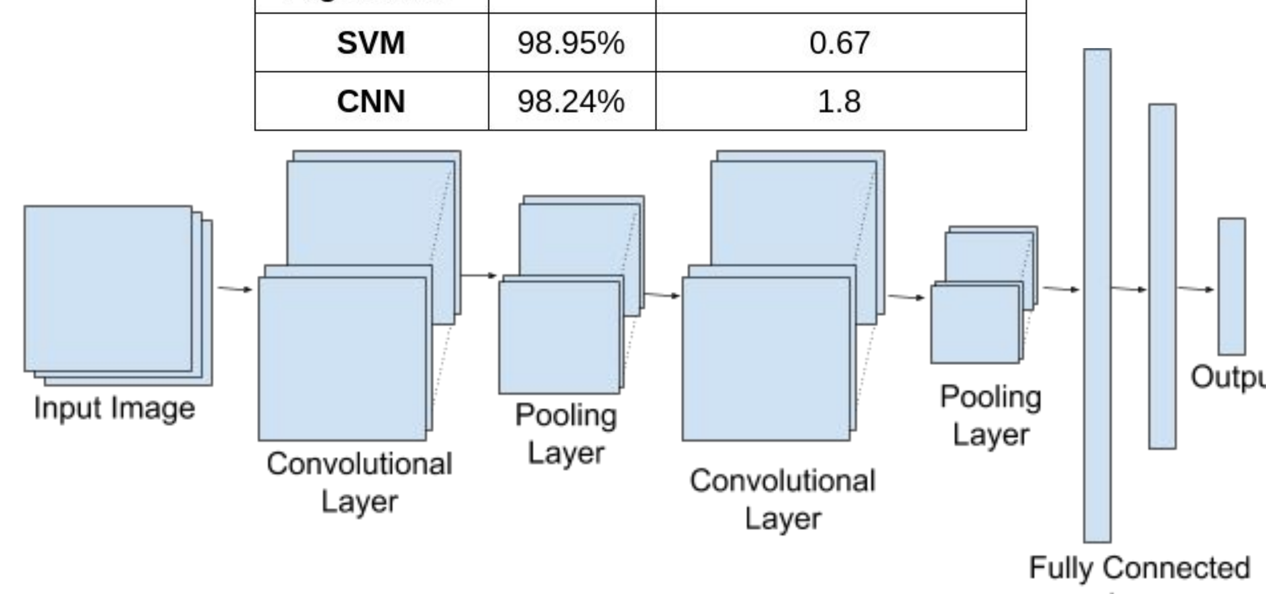
Mini-modules are also being tested after undergoing accelerated exposure

Supervised Machine Learning

Support Vector Machine (SVM)
- Kernel function:
  - Radial Basis function (RBF)

Algorithms	Accuracy	Standard Deviation
SVM	98.95%	0.67
CNN	98.24%	1.8

Convolutional Neural Network (CNN)
- Convolutional layer:
  - Set of kernels convolves on the image
  - Output of kernel is large matrix
  - It helps learn local spatial features
- Pooling layer
  - Max filter or average filter
- Dense layer/Fully connected layer
  - It helps learn global features



Structure of Convolutional Neural Network

References & Acknowledgment

- **References**
  - Fada et. al, "Democratizing an electroluminescence imaging apparatus and analytics project for widespread data acquisition in photovoltaic materials," Review of Scientific Instruments, vol. 87, no. 8, p. 085109, Aug. 2016.
  - J. S. Fada, M. A. Hussain, J. L. Braid, S. Yang, T. J. Peshek, and R. H. French, "Electroluminescent image processing and cell degradation type classification via computer vision and statistical learning methodologies," 44th PVSC, 2017.
  - R. M. Haralick, K. Shanmugam, and I. Dinstein, "Textural Features for Image Classification," IEEE Transactions on Systems, Man, and Cybernetics, vol. SMC-3, no. 6, pp. 610–621, Nov. 1973.

- **Acknowledgement**
  - Research was performed at the SDLE Research Center, which was established through funding through the Ohio Third Frontier, Wright Project Program Award tech 12-004.
  - This material is based upon work supported by the U.S. Department of Energy's Office of Energy Efficiency and Renewable Energy (EERE) under Solar Energy Technologies Office (SETO) Agreement Number DE-EE0007140.
  - This work made use of the Rider High Performance Computing Resource in the Core Facility for Advanced Research Computing at Case Western Reserve University.
  - This work was funded in part by the CWRU SOURCE SURES program.