Bismuth Vanadate (111) facet detection

Wanling Liu, Zixi Liu, Jiyao Yuan
{liuwl, liu1322, yuan999} @ stanford.edu

Introduction & Motivation

Nowadays, solving the energy challenge by harvesting the energy directly from sunlight through photosynthesis has become an attractive way. Bismuth Vanadate ($BiVO_4$) came out as the most promising material as its chemical reaction can be used as a photoanode that oxidizes solar water to O_2 in the photoelectrochemical cells. Research shows that higher percentage of (111) facet of $BiVO_4$ increases the efficiency of charge separation and helps save more energy. Subsequently, identification of (111) facet of $BiVO_4$ becomes important. Therefore, this project is focusing on the identification of (111) facet of $BiVO_4$ from SEM images.

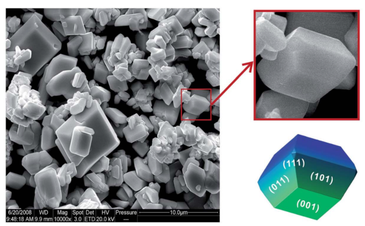


Fig 1. SEM image of $BiVO_4$

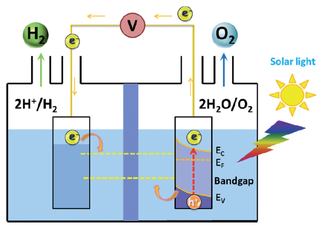
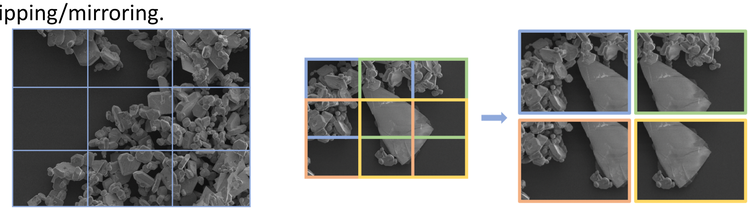


Fig 2. Theoretical progress in photoelectrochemical cell

Data

Our dataset contains around 3000 SEM images from experiments. Due to the relatively small dataset, data augmentation is an essential part of this project. Data augmentation methods that we mainly used include cropping, scaling and flipping/mirroring.



(a) (b)

Fig 3. Data Augmentation

In addition, due to the different magnifications of SEM images, we separated our dataset into three groups, and cropped them in different ways (see Fig.3).

Moreover, this project also tested the effect of edge detection, applying canny edge detection, probabilistic Hough and adaptive thresholding:

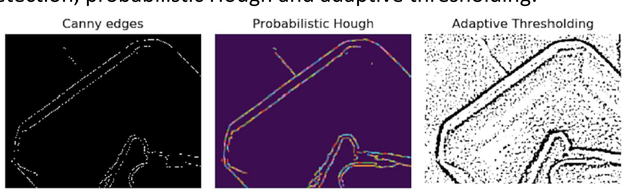


Fig 4. Edge Detection

Dataset was labeled manually by judgement from experts as either positive (image contains (111)) or negative (no observation of (111)). Data was then saved into either positive or negative folder so that the label of images could be read by the algorithms based on the name of the folder.

Models & Features

1. SVM:

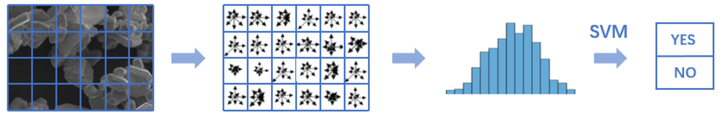


Fig 5. SVM with HOG

Input : 160 x 215-pixel, one-channel images;
HOG: 24 windows, 40 x 35 pixels in a window, 64 x 64 bins;
Features: vector with length 98304, represent the edge distribution;
SVM: linear kernel, penalty factor is 3.

2. ResNet:

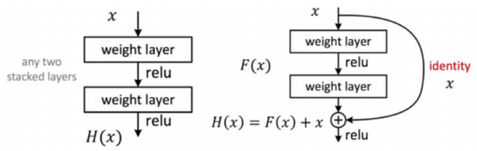


Fig 6. Standard vs. Residual

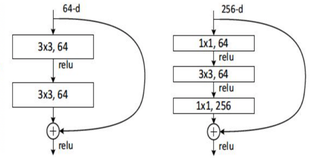


Fig 7. ResNet18 vs. ResNet50

Input: 160 x 215-pixel, three-channel images;
Number of layers: 18, 34, 50, 101, 152
Optimization algorithm: SGD, Adam
Loss function: Cross entropy loss
Addition: Transfer learning, image transformation (e.g. resizing, cropping)

3. Shallow CNN:

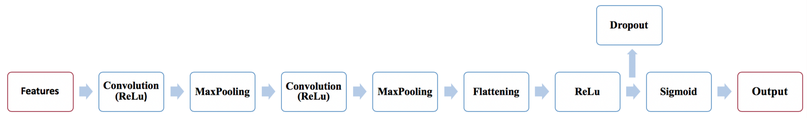


Fig 8. Shallow CNN Structure

Input: 160 x 215-pixel, three-channel images;
Structure: see Fig 8.

Results

There are around 3000 samples, including approximately 60% training data and 40% testing data. All inputs are images with 160 x 215 pixels.

Model	Training Accuracy	Test Accuracy
SVM (Sigmoid)	63.78%	64.10%
SVM with HOG	100.00%	78.74%
Shallow CNN	**99.79%**	**87.16%**
Shallow CNN (L2 Regularization)	98.76%	85.17%
ResNet18 (not pretrained)	63.49%	66.20%
ResNet18	70.34%	75.62%
ResNet50	72.86%	79.36%

Conclusion & Discussion

- Image preprocessing, such as image binarization and edge detection, has negligible effects on the results of our models, because after extracting edge information, features of (111) facets are less obvious compared to noise features.
- The predictions of SVM method are biased due to biased number of positive and negative training data. To be specific, the output of SVM is of higher probability of predicting negative, resulting in lower recall.
- SVM method has fewer parameters and simpler structure compared to neural networks. Thus, it computes much faster than neural networks, but the accuracy is much lower.
- The key of SVM is to find an appropriate feature descriptor that augments features from (111) facets and weakens unrelated features;
- We have tried several numbers of layers for ResNet and we find that prediction accuracy for our data does not benefit much from deeper network, which might be because we only have relatively limited number of data or the network structure is not very suitable for our data.
- Due to the relatively small dataset, we tried using transfer learning on pretrained models. As a result, the test accuracy was improved around 10% compared to not pretrained models.
- After tuning parameters, shallow CNN as described previously can reach the test accuracy of around 87%, which is the best accuracy we currently get. And thus, this structure might be more suitable for our data and we may further modify it and try its deeper version.
- Since we manually labeled the dataset, even supervised by experienced people in this area, there are still some ambiguous facets that are hard to determine whether they are (111) or not due to several reasons (for example, the quality of SEM images, the point of view and magnification of particles, and various shapes of particles), which might influence the accuracy of the models.

Future

In terms of applications:

- Extending (111) facet detection to other kinds of facets detection;
- Transferring classification problem to regression problem, i.e. determining the proportion of (111) facets in $BiVO_4$ SEM images;

In terms of methods:

- Considering other feature descriptors, such as SIFT, SURF, ORB;
- Tuning parameters such as learning rate, step sizes for better performances of neural networks;
- Further modifying the structure of neural networks;
- Continuing experiments to gain more high-quality data.

Reference

[1] Yiseul, P., Kenneth, J., McDonald, B., & Kyoung, S.C. (2013). Progress in bismuth vanadate photoanodes for use in solar water oxidation. Chemical Society Reviews, Issue 6.
[2] Li, G.L. (2017). First-principles investigation of the surface properties of fergusonite-type monoclinic BiVO4 photocatalyst. RSC Advances. Issue 15.
[3] He, K.M., Zhang,X.Y. Ren, S.Q., & Sun, J.(2016). Deep residual learning for image recognition. Proceedings of the IEEE conference on computer vision and pattern recognition. pages 770–778.