Needle in a Haystack

Carlos Ruiz & Zhengyi Wang

Content

- Introduction
- Our solution
- Feature Extraction
  - Color
  - Line
- Training Dataset
- Result
- Conclusions
- Future work

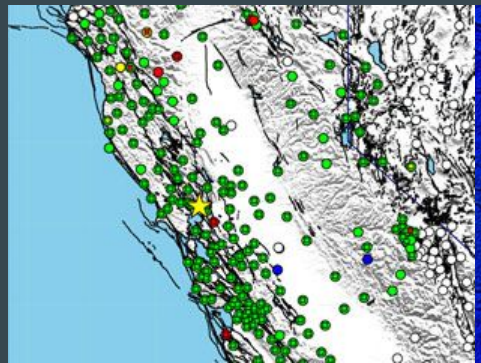
Carlos

Zhengyi

Introduction

Our project idea comes from an USGS research project:

- Probable locations of a big earthquake’s epicenter
- Want to analyze consequences/risk for each possibility
- Knowing the location of energy supplies would be very helpful



What is the problem?

Electrical companies don’t want to provide exact location of their power plants and electrical substations (potential terrorist attacks)

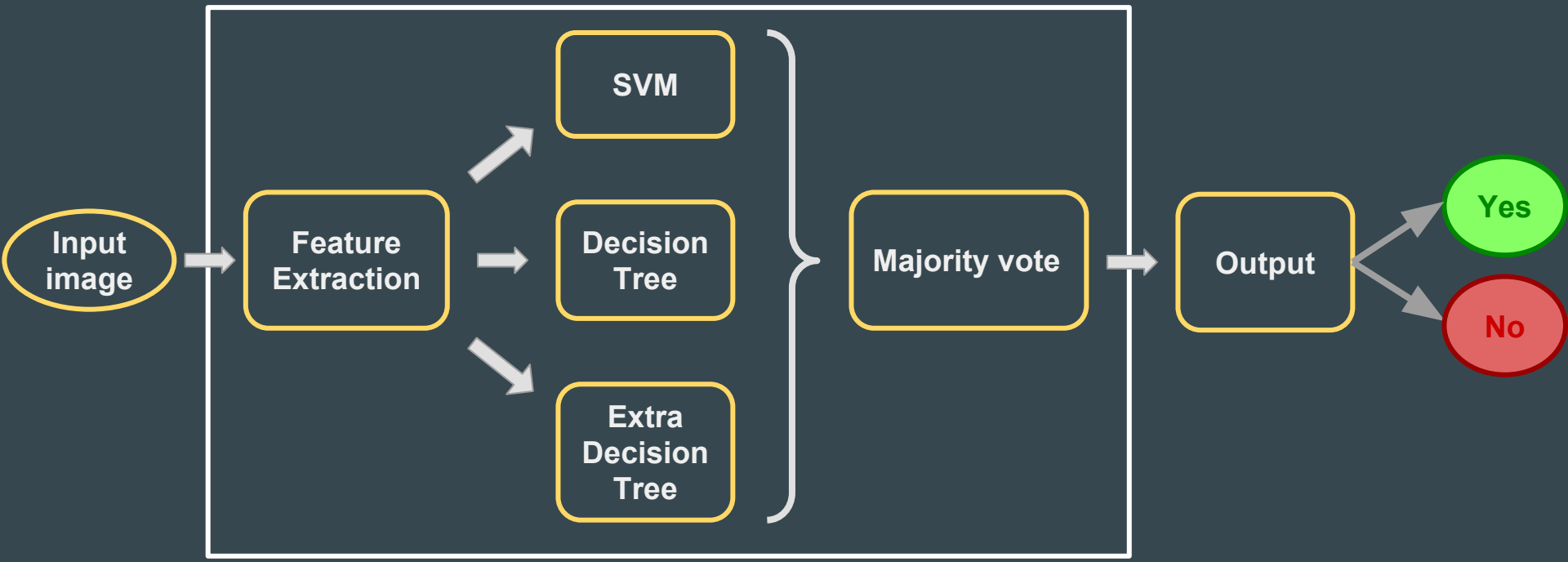
… So what can we do?

Substation List - LCR

Local Capacity Area Substation List
Based on August 21, 2014

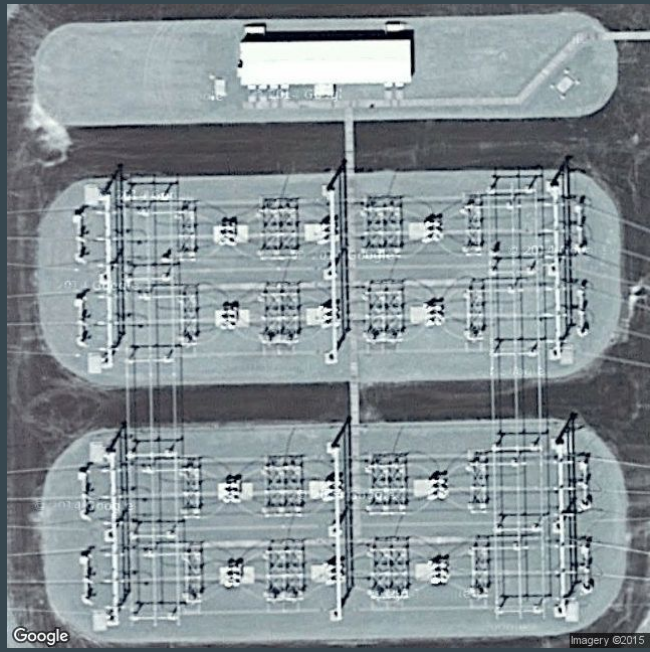
SAP Substation Name	LCR Area	City served	Division	Owner
BRENTWOOD SUB	Greater Bay Area	Bethel Island	DIABLO	PGE
BRITTON SUB	Greater Bay Area	Sunnyvale	DE ANZA	PGE
BRITTON SUB	Greater Bay Area	Santa Clara	DE ANZA	PGE
BROOKSIDE SUB	Greater Bay Area	San Pablo	EAST BAY	PGE
BROOKSIDE SUB	Greater Bay Area	Richmond	EAST BAY	PGE
BRYANT SUB	Greater Bay Area	Orinda	DIABLO	PGE
BURLINGAME SUB	Greater Bay Area	Hillsborough	PENINSULA	PGE
BURLINGAME SUB	Greater Bay Area	Burlingame	PENINSULA	PGE
BURNS SUB	Greater Bay Area	Santa Cruz	CENTRAL COAST	PGE
CALERO SUB	Greater Bay Area	San Jose	SAN JOSE	PGE
CAROLANDS SUB	Greater Bay Area	San Mateo	PENINSULA	PGE
CAROLANDS SUB	Greater Bay Area	Hillsborough	PENINSULA	PGE
CAROLANDS SUB	Greater Bay Area	Burlingame	PENINSULA	PGE
CASTRO SUB	Greater Bay Area	San Francisco	SAN FRANCISCO	PGE
CASTRO VALLEY SUB	Greater Bay Area	Union City	MISSION	PGE
CASTRO VALLEY SUB	Greater Bay Area	Hayward	MISSION	PGE
CASTRO VALLEY SUB	Greater Bay Area	Fairview	MISSION	PGE
CASTRO VALLEY SUB	Greater Bay Area	Cherryland	MISSION	PGE
CASTRO VALLEY SUB	Greater Bay Area	Castro Valley	MISSION	PGE
CASTRO VALLEY SUB	Greater Bay Area	Ashland	MISSION	PGE
CAYETANO SUB	Greater Bay Area	Livermore	MISSION	PGE
CHERRY SUB	Greater Bay Area	San Leandro	MISSION	PGE
CHRISTIE SUB	Greater Bay Area	Hercules	DIABLO	PGE
CLAYTON SUB	Greater Bay Area	Walnut Creek	DIABLO	PGE
CLAYTON SUB	Greater Bay Area	Concord	DIABLO	PGE
CLAYTON SUB	Greater Bay Area	Clayton	DIABLO	PGE
CLAYTON SUB	Greater Bay Area	Alamo	DIABLO	PGE
CONCORD SUB	Greater Bay Area	Concord	DIABLO	PGE
CONTRA COSTA PP SUB	Greater Bay Area	Antioch	DIABLO	nonPGE
CONTRA COSTA SUB	Greater Bay Area	Oakley	DIABLO	PGE
CONTRA COSTA SUB	Greater Bay Area	Brentwood	DIABLO	PGE
CONTRA COSTA SUB	Greater Bay Area	Antioch	DIABLO	PGE
COOLEY LANDING SUB	Greater Bay Area	Palo Alto	PENINSULA	PGE
CROCKETT COGEN SW STA SUB	Greater Bay Area	Crockett	EAST BAY	PGE

Our solution



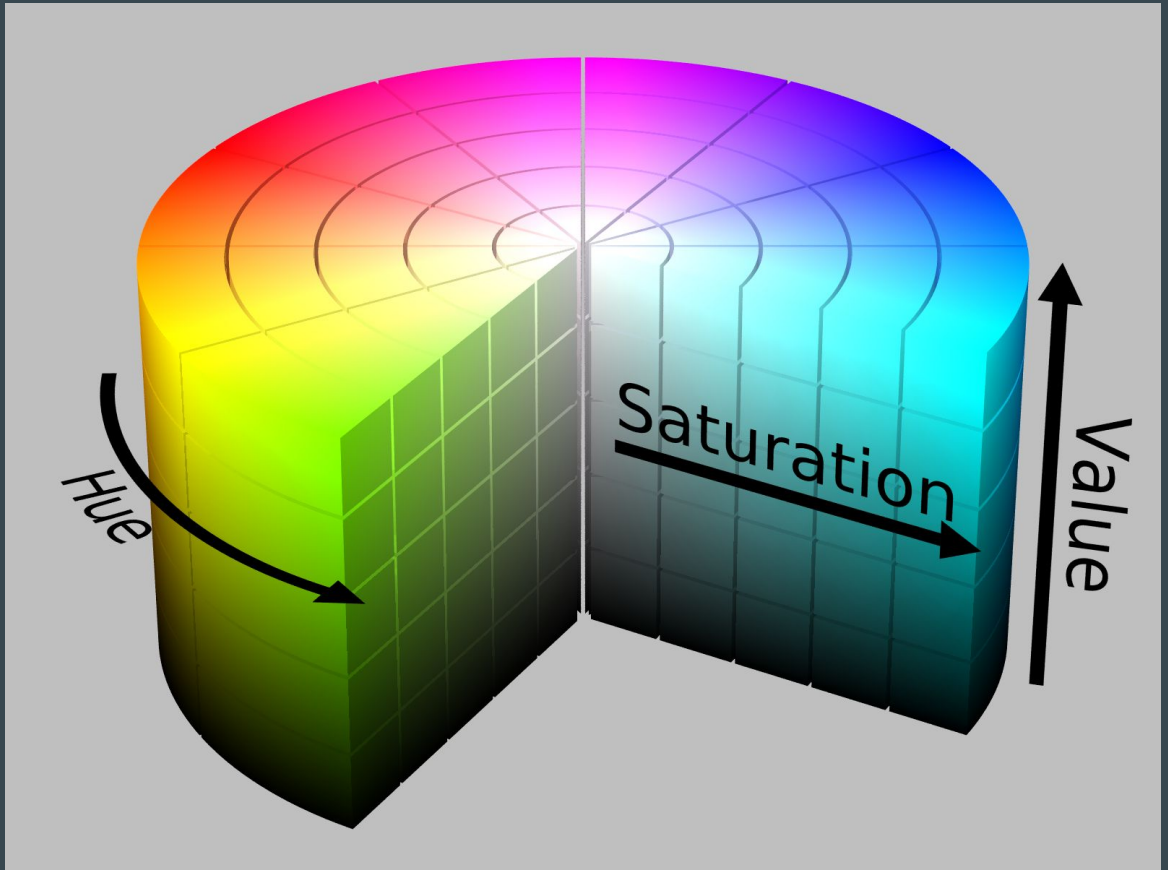
Feature Extraction

- Color
  - HSV colorspace Histogram
  - *Spacial Histogram
- Lines
  - Number of lines
  - Spatial distribution of angle: perpendicularity
  - Histogram
- …

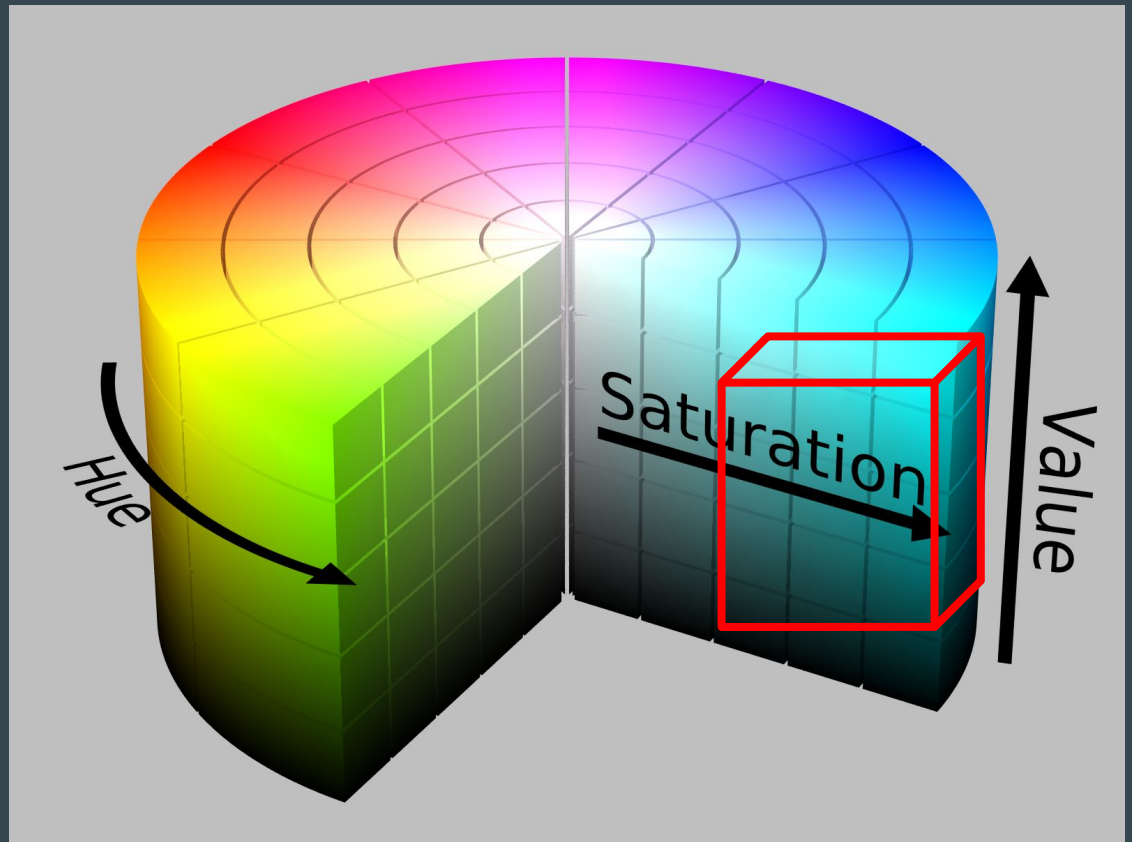


Feature - Color

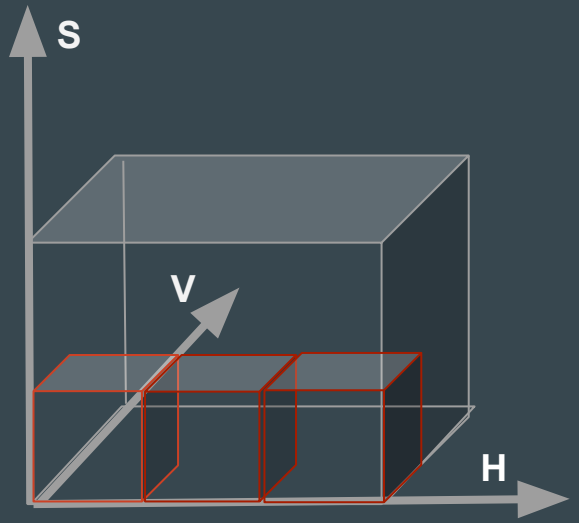
HSV Colorspace

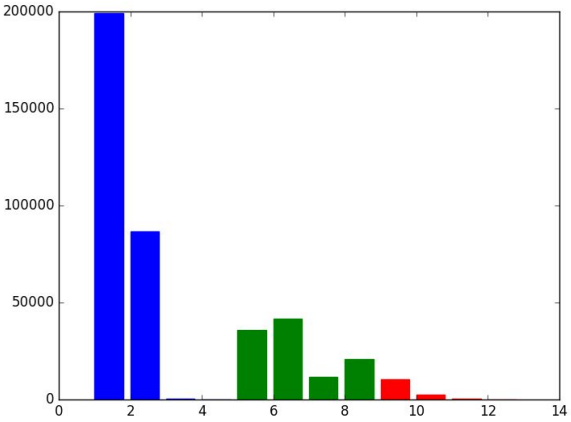


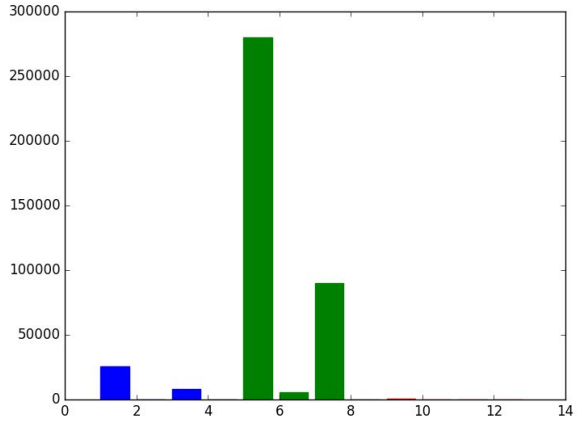
Feature - Color



Feature - Color

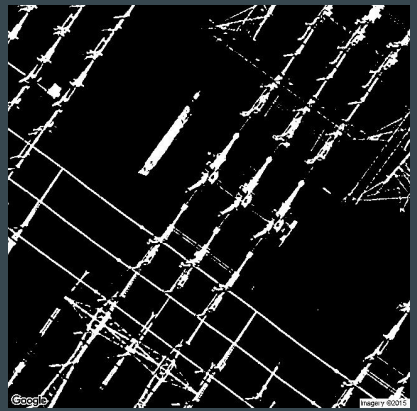


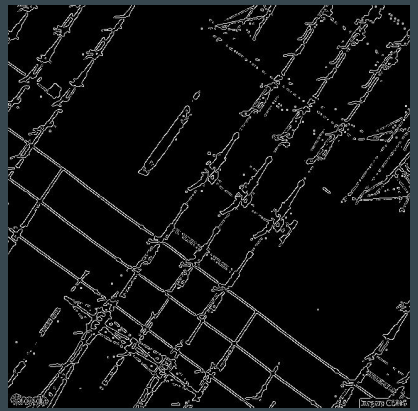


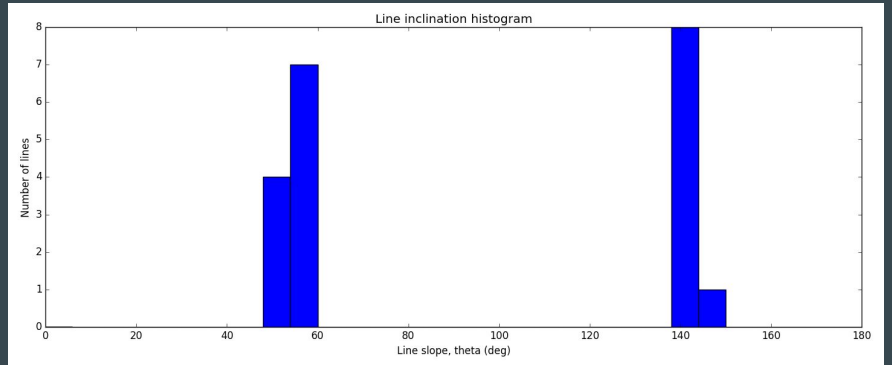


Feature - Lines

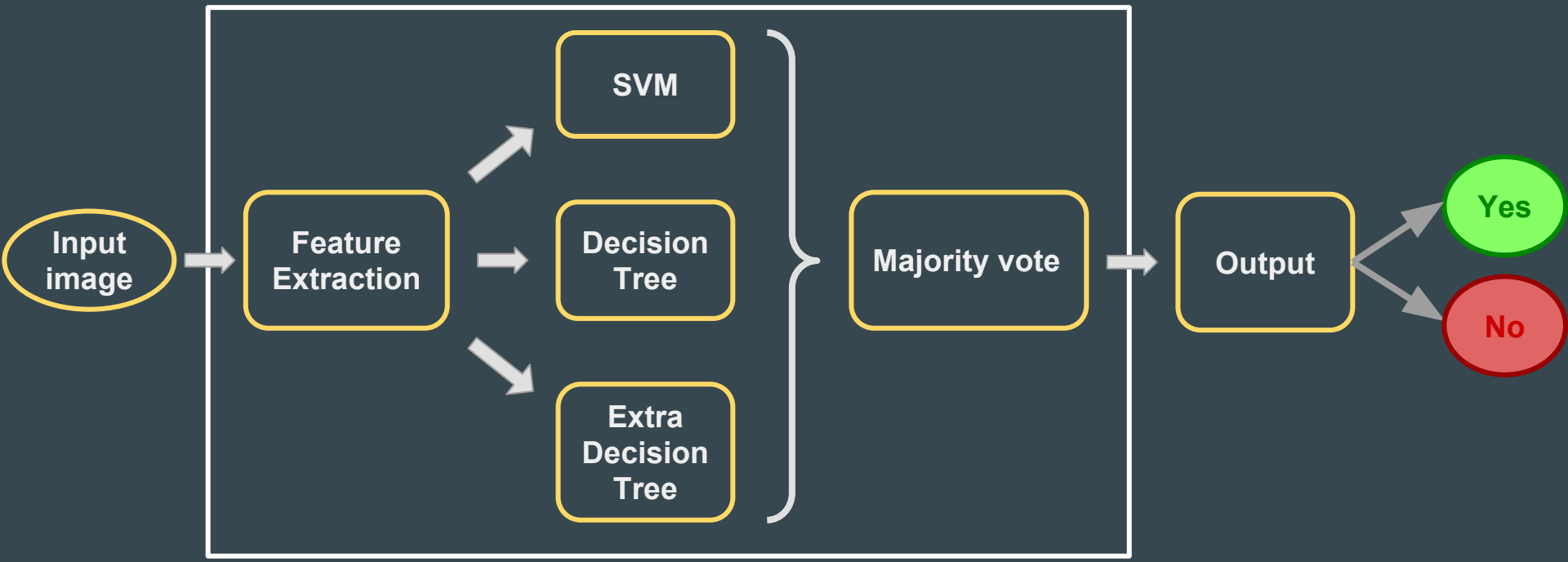
Original image -> Threshold filtering -> Edge Detection -> Line Detect







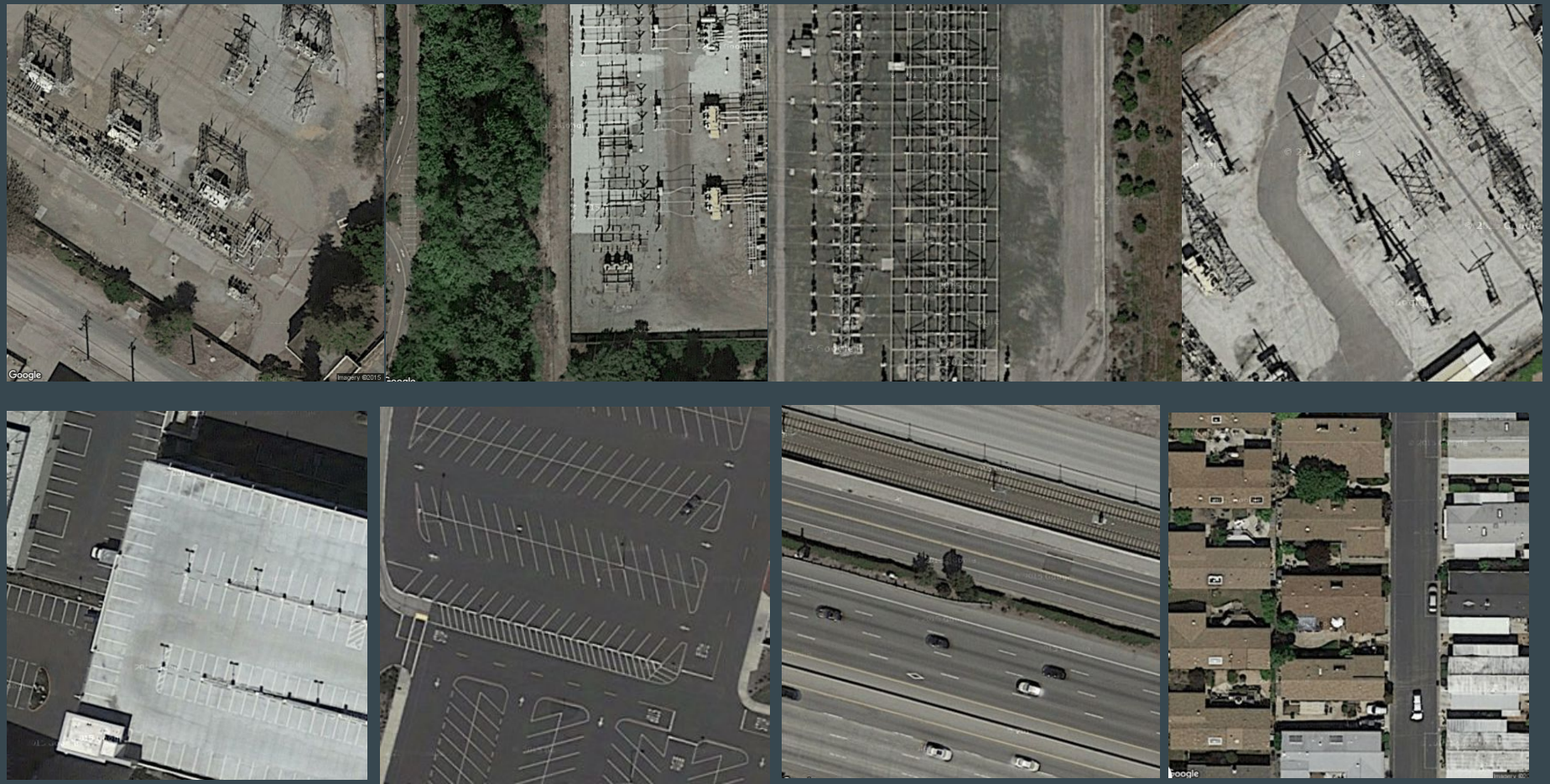
Proposed solution

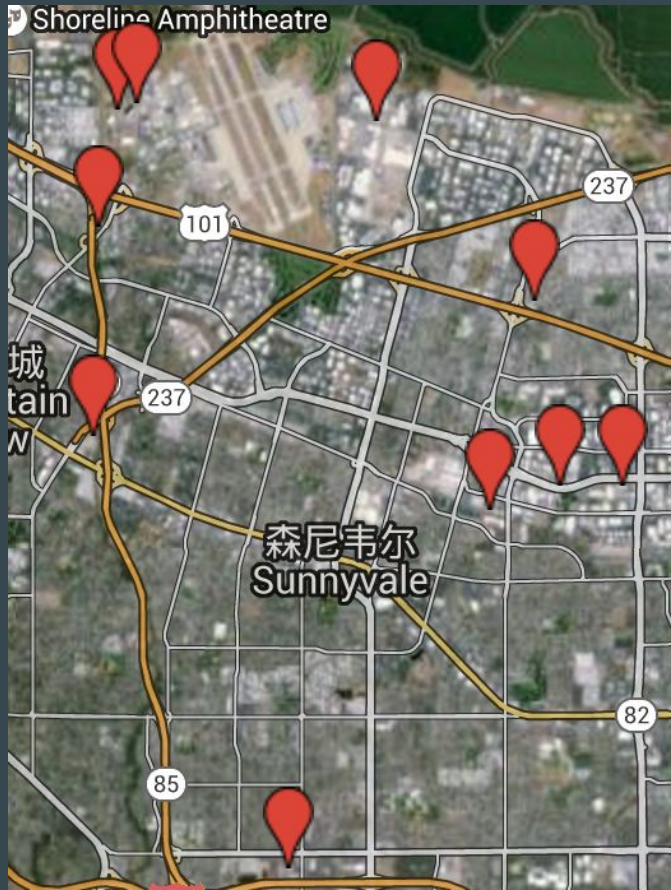


Dataset

- Difficulty - lack of ground truth
- Manually find substations +
  Script to download positive samples
  - 20 substations (60 images)
  - Most substations are not marked in Google Maps
  - Hard to learn from limited info
- Iterations to improving our algorithm
  - SVM, Decision Tree, Extra Decision Tree, Vote
  - Number of Bins for colorspace
  - Downloaded Sunnyvale region ~15000 pics
  - Found all 10 in Sunnyvale (pdf document)
  - Next: Retrain to test in San Fransisco and Berkeley
  - …

Result



Conclusion

- Result
  - False Negative ~0%
  - False Positive ~95%
  - ~100 ms / image (vs. ~3sec/image old code)
- Future Work
  - More feature - Spacial Histogram
  - More training data