

TextonSVM

Automatic Particle Selection for Single Particle Analysis

version 1.0

GETTING STARTED GUIDE

CONTENTS

1. Introduction
2. Experimental Setup
3. Installation
4. Collecting Ground-truth
5. Detecting Particles
6. Visualizing Results

1. INTRODUCTION

The purpose of this guide is to provide a fast way to get started with our software package TextonSVM. The next sections describe in detail a full experiment of automated particle detection on a real Electron Microscopy dataset. By following the step-by-step directions on a standard Linux machine, you will be able to reproduce the full results for the Bacterioferritin dataset in our paper below, from downloading the code to the final performance evaluation.

By using the TextonSVM software you agree to all licensing terms in the LICENSE file distributed with the package. Scientific results produced using the TextonSVM software should acknowledge its use from the authors and cite the following article:

"Experimental Evaluation of Support Vector Machine-based and Correlation-based Approaches to Automatic Particle Selection".

**P. Arbeláez, B.G. Han, D. Typke, J. Lim, R.M. Glaeser, and J. Malik.
Journal of Structural Biology, 2011. doi:10.1016/j.jsb.2011.05.017**

For more details, please visit the project page:

<http://www.eecs.berkeley.edu/Research/Projects/CS/vision/bioimages/textonsvm/>

2. EXPERIMENTAL SETUP

We recommend the following experimental setup, used in our paper.

- Choose a train image: Choose an image of good quality, representative of the dataset and with enough good particles and declare it your **train image**. This image will be used to learn the appearance of the particle of interest and to evaluate results quantitatively.

- Collect Ground-truth: The first step is collecting the **ground-truth**, in which you mark manually *all* the particles you expect the algorithm to detect in the train image. This task is done with the help of the **input user interface**. Alternatively, you can collect the ground-truth using Eman boxer.

- Launch Automatic Detection: After ground-truth collection, the input user interface allows you to launch the core **backend** program. TextonSVM learns a model for the appearance of the particle, optimizes it with respect to the ground-truth and then proceeds to detect automatically particles in all the **test images** of the dataset.

- Visualize Results: TextonSVM produces a list of detected particles for each test image, sorted by detection confidence. The **output user interface** allows you to navigate and visualize the results, and to adjust the minimal detection confidence for each image individually.

The Bacterioferritin experiment below was tested on a standard single-core machine running Ubuntu Linux 12.04. Using default parameters, learning the model takes approximately 12h and detecting particles on each test image of size 13860x12000 takes 5 to 10 min.

3. INSTALLATION

3.1 *Open a terminal. Create a fresh directory called, e.g. “textonsvm”, and change directory to it*

```
mkdir textonsvm/
```

```
cd textonsvm/
```

3.2 *Download the code (or copy the file “textonsvm_v1.0.tgz” to the current directory)*

```
wget http://www.eecs.berkeley.edu/Research/Projects/CS/vision/bioimages/textonsvm_v1.0.tgz .
```

3.3 *Uncompress the file*

```
tar -xvzf textonsvm_v1.0.tgz
```

3.4 *Install the package*

```
cd textonsvm_v1.0
```

```
./easy_install > install.log 2>install.err
```

```
cd ..
```

TextonSVM and all required libraries will be installed. The installation script was tested on a standard Ubuntu Linux 12.04 distribution.

4. COLLECTING GROUND-TRUTH

4.1 *Download and uncompress the Bacteroferritin dataset*

```
wget enigma.lbl.gov/supplemental/TextonSVM/FER.tgz .
```

```
tar -xvzf FER.tgz
```

4.2 *Use the input user interface to manually collect exemplars in the train image*

```
source textonsvm_v1.0/textonsvm_setup
```

```
textonsvm_v1.0/bin/input_ui
```

4.3 *Click on “Open Image”, navigate and select with the mouse, or type the name of the train image file:*

```
FER/images/FER_006.tif
```

4.4 *A window called “Scale Options” pops out. Click on “OK”.*

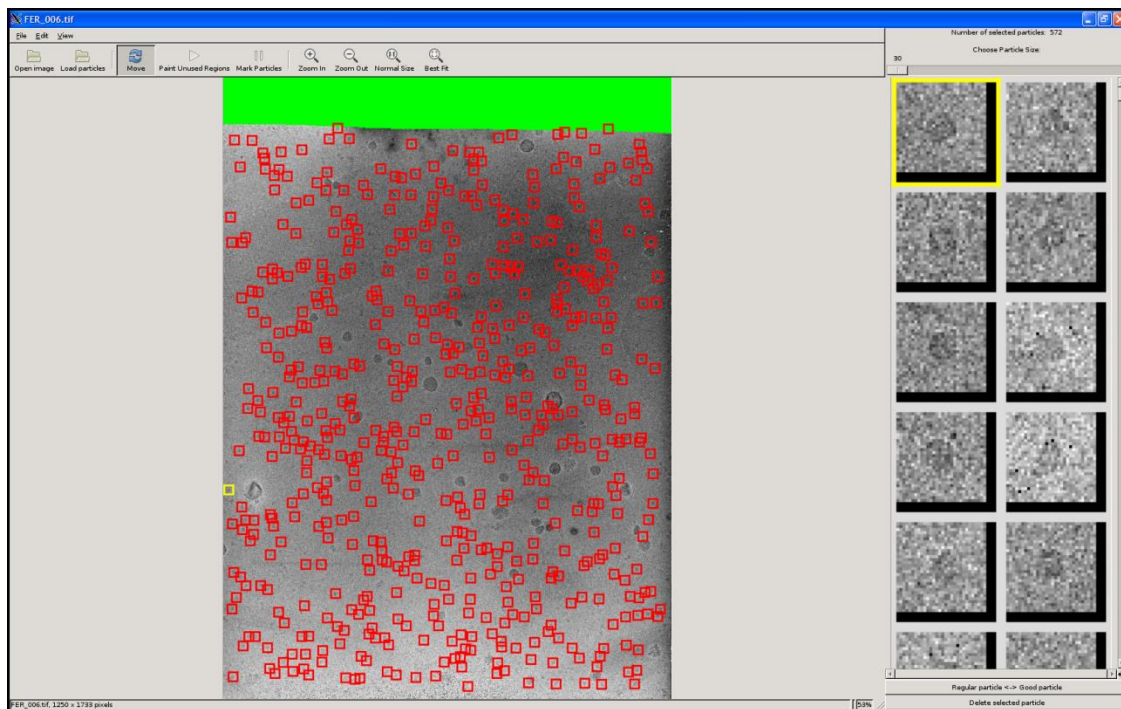
4.5 Click on “Paint Unused Regions”, paint with the mouse all the regions you want TextonSVM to ignore in the train image, e.g. the header or damaged parts.

4.6 Click on “Mark Particles” to start the manual collection of ground-truth.

The input user interface allows you to: zoom in to a comfortable resolution, navigate the image, select particles by clicking with the mouse, delete a selected particle, and adjust the size of the particle.

4.7 Once you have collected the ground-truth, you can store it by clicking on the menu “File” -> “Save Data” and selecting with the mouse, or typing the name of the dataset folder:

FER/images/FER_006.tif



Screenshot of input user interface

4.8 Close the input user interface by clicking on the menu “File” -> “Quit”

5. DETECTING PARTICLES

5.1 Remove any particles created in the previous step, download and uncompress the ground-truth collected for the experiment in our JSB paper

```
rm FER/images/*box FER/images/*mask
```

```
wget enigma.lbl.gov/supplemental/TextonSVM/FER_particles.tgz .
```

```
tar -xvzf FER_particles.tgz
```

5.2 Launch the input user interface. Diagnostic text output will be stored in a file called *FER.log*

```
source textonsvm_v1.0/textonsvm_setup
```

```
textonsvm_v1.0/bin/input_ui >FER.log
```

5.3 Go to the menu: “File” -> “Launch backend”

5.4 In the window “Choose files(s) to get good particles”, navigate and select with the mouse, or type the name of the train image file:

```
FER/images/FER_006.tif
```

5.5 In the window “Choose file(s) to train”, navigate and select with the mouse, or type the name of the train image file again:

```
FER/images/FER_006.tif
```

5.6 In the window “Choose file(s) to test”, navigate to *FER/images* and select with the mouse all the test images.

5.7 In the window “Choose directory to store results”, double click on the folder “FER”, type “results” in the text field and click on “OK”, or type:

```
FER/results
```

5.8 A window with the sign “Advanced Options” pops up. Click “OK”.

5.9 A window with the sign “Click OK to start pre-processing” pops up. Click “OK”.

TextonSVM will then pre-process the dataset, learn and optimize a model for the molecular structure of interest and detect similar particles in all the test images. Processing the Bacteroferritin dataset (19 images of size 13860x12000) with default parameters takes about 14h on a standard machine with a single core.

6. VISUALIZING RESULTS

The final output of TextonSVM is the following: for each test image, a file with the image name and extension “.det” is created, e.g., FER_001.det. This file is compatible with Eman boxer and contains detected particles sorted by confidence score. By default, the minimal confidence allowed to declare a detection is the same for all images. However, the user can navigate and visualize the results, and choose different thresholds for each image with the output user interface:

[source textonsvm_v1.0/textonsvm_setup](#)

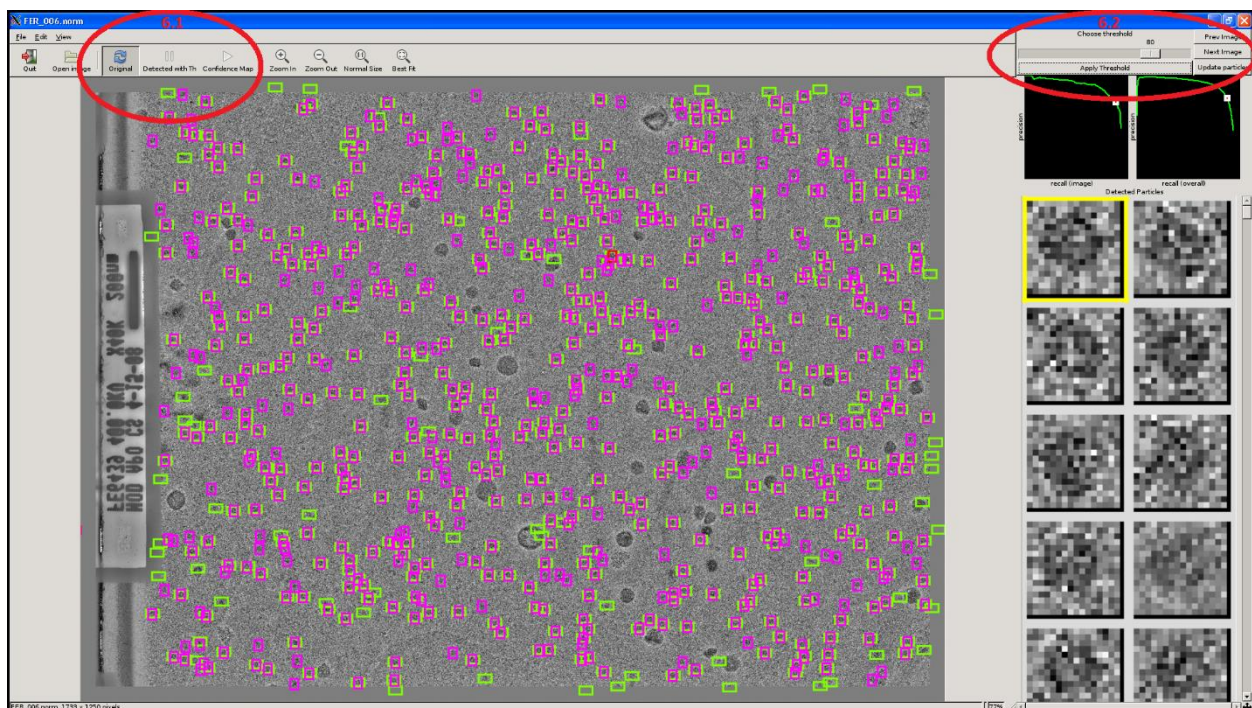
[textonsvm_v1.0/bin/output_ui >FER.log2](#)

6.1 All the results are stored in the folder FER/results/results/. Open one of the files by clicking on “Open Image” and its name:

[FER/results/results/FER_006.norm](#)

6.2 By clicking the buttons “Original” / “Detected with Thr” / “Confidence Map” you can select the type of output to visualize.

6.3 For each image, you can select a minimal confidence threshold in [0 100] with the scroll bar and clicking on “Apply Threshold”. If you want to store the particles with the new threshold, click on “Update Particles”. Buttons “Next Image” and “Prev Image” allow to navigate the dataset.



Screenshot of output user interface