Image Processing and Image Pattern Recognition A Programming Tutorial

Amit Chakraborty
Applied Data Tech
Bangalore , India
amit@appliedanalytics.tech

Abstract— **Image recognition is a major area of application of machine learning - evolving at a rapid pace with a number of programming platforms available to developers. While each platform has its own uniqueness, the methodology of image recognition consists of a sequence of image processing tasks, development of a classifier algorithm, training and testing followed by deployment. This tutorial will delve into the programming aspects of image processing including thresholding, contouring and template matching. In order to provide practical hands on programming this tutorial will closely look at three real life applications of image pattern recognition namely ALPR using Tesseract OCR and will touch upon using CNN for character detection. The tutorial will explain the algorithm, implementation of pseudocode through Python using two major platforms: OpenCV and Tensorflow.**

Scope of the Tutorial:
There are various ways of understanding image processing and image recognition, however in this tutorial we will follow the journey that we took as a development team at Applied Data Tech to build up our image processing stack.
During this tutorial, we will explain the programming techniques about processing images with the objective of Image recognition particularly for character recognition.
OCR has been around for decades with varying complexity and accuracy. We will take a look at two popular OCR technology has evolved: openCV and Tensorflow

Image Pre-processing:

- Binarization: converting RGB to Greyscale and then to black and white. Binarization helps reduce noise using various algorithms:
  Global fixed
  <code> <example input image > <example output image >
  Adaptive thresholding
  <code> <example input image > <example output image >
  Markov
  Otsu Binarization
  <code> <example input image >
  <example output image >

- Changing Colorspaces: used for identifying objects with similar color in an image. Especially helpful in identifying a colored object in multiple frames.
- Geometric transformation : translation, rotation and skewing.
- Smoothing : Blurring an image essentially passing through a LPF : 2D Convolution
  Averaging
  Gaussian Blurring
  Median Blurring
  Bilateral filtering
- Morphological transformation: Change in shapes in binary images. Mathematically it is a kernel that slides through the image performing convolution
  Erosion:
  Dilation:
  Opening
  Closing:
- Image gradients: Used for detecting edges and contours
  Sobel and Scharr derivatives : direction of derivatives : x or y
  Laplacian derivatives
- Edge detection: Edge detection is combination of reducing noise and then applying image gradients once in X and then in Y direction. OpenCV provides one single method Canny to accomplish this.

Character recognition:
Tesseract: Developed originally at HP and made open source in 2005 this is a versatile tool has been in use in production grade space for quite a while.
Using Tesseract from command line:
< code> <output>
pytesseract: A python wrapper over the CL tool using config argument. The image_to_string method of pytesseract invokes the CLI tool.
<code>

- Image Segmentations:

The methodologies and corresponding algorithms for segmenting an image based on various characteristics. These methodologies could utilise some of the pre processing techniques we have discussed earlier.
In this case we use the following processing techniques:

978-1-5386-9209-7/18/$31.00 ©2018 IEEE
DOI 10.1109/ai4i.2018.00040

<Original image> Otsu's Binarization <objects not yet segmented> Applying Erosion and Dilation -- > Distance transform → Thresholding -> Marker -> Watershed

<Final image properly segmented>

- Hough transforms:

When applied on an image this detects any shape that can be represented in a mathematical form. It works only for binary image therefore we need to apply thresholding or edge detection before applying this transform.
Most popular ones are detecting lines and circles. It also provides an internal thresholding which indicates a minimum vote dented by the minimum number of points that satisfy the mathematical formula for the shape.
<Image> <code for Hough Line transform> <output image>
<Image> <code for Hough Circle transform> <output image>

- Contour detection: drawing a curve joining all continuous points with the same color or Intensity.

Find contours
Drawing contours
<code> <image>
Template Matching

- Legacy Tesseract and LSTM:

Tesseract used a legacy algorithm earlier. Since Tesseract 4.0 Neural Net LSTM has been added that has improved efficiency.

<Code > <Output>

- Visual Attention based character recognition:
  With today's GPU based processing power fast CNN Tensorflow implementation on visual Attention based detection of characters or objects is becoming popular.

<Data Set Images>
<Training code>
<Tuning of the model>

- Application of OCR :
  ALPR is a well know application of OCR where we apply pre processing of image and then apply OCR techniques.
- Localization of Number Plate:
  The first step is to locate the number plate in the image of a car. The image is converted in grayscale and then Smoothing and media filter is applied to reduce noise. An image mask is applied followed Binary thresholding. This is followed by opening the image – a combination of erosion and dilation. This helps in locating the boundary box of the license plate.
  <code snippets> <images>
- Character Recognition:
- This rectangular image is now fed through Tesseract to obtain the characters.

<code> <image>

Post Processing:
However, there are several challenges that are faced in production grade deployment of ALPR.
<image> <text output>
<image> <text output>

- ALPR post processing involves:
- Validating characters based on the country format

<image> <text output>
<image> <text output>

- Multiple images and majority vote

<image> <text output>
<image> <text output>

- Exception handling for vanity number plates

<image> <text output>
<image> <text output>

Next Study:

- Techniques for object recognition using Tensorflow and OpenCV
- Specific applications in:
  - Vehicle detection
  - Pedestrian detection
  - Face detection

References:

[1] Maya R. Gupta□, Nathaniel P. Jacobson, Eric K. Garcia, OCR binarization and image pre-processing for searching historical documents

[2] Dirk Walther1 , Laurent Itti2 , Maximilian Riesenhuber3 , Tomaso Poggio3 , and Christof Koch1, Attentional Selection for Object Recognition – a Gentle Way

[3] University Library, University of Illinois Uraban Champagne, Introduction to OCR and Searchable PDFs: Using Tesseract
http://guides.library.illinois.edu/c.php?g=347520&p=4121426

[4] How we tuned Tesseract to perform as well as a commercial OCR package.http://vbridge.co.uk/2012/11/05/how-we-tuned-tesseract-to-perform-as-well-as-a-commercial-ocr-package/