

Universitatea POLITEHNICA din București

Facultatea de Automatică și Calculatoare,

Departamentul de Calculatoare



LUCRARE DE DIPLOMĂ

Sistem de recunoaștere a siglelor

Conducător Științific:

Șl.dr.ing. Răzvan Deaconescu

Autor:

Flavius–Valentin Anton

București, 2015

University POLITEHNICA of Bucharest

Faculty of Automatic Control and Computers,
Computer Science and Engineering Department



BACHELOR THESIS

Pipelined Logo Recognition System

Scientific Adviser:

Șl.dr.ing. Răzvan Deaconescu

Author:

Flavius-Valentin Anton

Bucharest, 2015

Maecenas elementum venenatis dui, sit amet
vehicula ipsum molestie vitae. Sed porttitor
urna vel ipsum tincidunt venenatis. Aenean
adipiscing porttitor nibh a ultricies. Curabitur
vehicula semper lacus a rutrum.

Quisque ac feugiat libero. Fusce dui tortor,
luctus a convallis sed, lacinia sed ligula.
Integer arcu metus, lacinia vitae posuere ut,
tempor ut ante.

Abstract

Here goes the abstract about MySuperDuperProject. Lorem ipsum dolor sit amet, consectetur adipiscing elit. Aenean aliquam lectus vel orci malesuada accumsan. Sed lacinia egestas tortor, eget tristique dolor congue sit amet. Curabitur ut nisl a nisi consequat mollis sit amet quis nisl. Vestibulum hendrerit velit at odio sodales pretium. Nam quis tortor sed ante varius sodales. Etiam lacus arcu, placerat sed laoreet a, facilisis sed nunc. Nam gravida fringilla ligula, eu congue lorem feugiat eu.

Contents

Acknowledgements	i
Abstract	ii
1 Introduction	1
1.1 Motivation	1
1.2 Objectives	2
1.3 Use Cases	3
1.4 Thesis Summary	3
2 Existing Work	5
2.1 Image Based Analytics in Social Media	5
2.2 Object and Logo Recognition	6
2.2.1 Template Matching Approach	7
2.2.2 Learning Approach	8
3 Proposed Solution	9
3.1 Requirements and Technologies Used	9
3.1.1 Kafka – A Messaging System	9
3.1.2 Scala	9
3.2 System Architecture	9
3.2.1 Fetch Module	10
3.2.2 Detector Module	10
3.2.3 Annotated Data Storage Module	10
3.2.4 Queing System	10
3.2.5 Data Flow	10
3.3 Logo Recognition Algorithm	10
4 Testing and Evaluation	11
4.1 Algorithm Performance	11
4.2 Cost and Resource Consumption	11
4.3 Precision and Recall	11
4.3.1 Precision and Recall in Pattern Recognition	11
4.3.2 Precision and Recall for our Algorithm	11
4.4 Scalability and Portability	12
5 Further Work	13
5.1 System Architecture Enhancements	13
5.2 Better Logo Recognition Algorithm	13
5.3 Other Possible Social Media Image Processing	13
6 Conclusions	14

List of Figures

2.1	Random Logos	6
2.2	Sliding Template	8

List of Tables

Chapter 1

Introduction

TODO:

Talk about Social Media and analytics and importance of brand recognition

1.1 Motivation

Social Media has become an integral part of modern society, thus many activities in the software industry emerge around it. In the last few years, it has been observed that a social network like Facebook or Twitter started to overcome their initial purpose of easing access to information sharing or connecting people in the online environment. One relatively new purpose of social networks is using them as an advertisement platform. This *use case* is not that hard to imagine if we make a reference to the latest digits that social networks have to bear these days.

According to an online statistics portal^[2] Twitter has passed 300² million monthly active users whereas Facebook is just 50 million users under the 1.5 billion barrier, which is roughly one fifth of the total world population. Needless to say, Instagram, Google+ or even Tumblr are over the 200 million monthly active users barrier. Moreover, there are 350 million new photos uploaded every day on Facebook³ and 70 million on Instagram⁴. Twitter also faces more than 600 million tweets in a normal day of which approximately 172⁵ million contain images.

With these rapidly increasing numbers, naturally comes a new area of development and research called *social media analytics*, which is "the practice of gathering data from blogs and social media websites and analyzing that data to make business decisions. The most common use of social media analytics is to mine customer sentiment in order to support marketing and customer service activities"⁶. Understanding social media analytics gives companies a glimpse about how their products and services are perceived by clients and potential clients. For example, most of the people using a social network share content about their own personal life, their thoughts, images and videos from their everyday life. All this content can be used by companies to analyze how well they are performing and try to improve their services (i.e. offer better advertisements based on that user generated content).

That being said, in the last few years there was an increasing demand in the industry for tools of analyzing unstructured data found on social networks. This tools are specifically designed

²<http://www.statista.com/statistics/272014/global-social-networks-ranked-by-number-of-users/>

³www.businessinsider.com/facebook-350-million-photos-each-day-2013-9

⁴<https://instagram.com/press/>

⁵Numbers determined experimentally while working on this project

⁶<http://searchbusinessanalytics.techtarget.com/definition/social-media-analytics>

and built to consume as much data as they can from the social platforms, process it and offer a variety of representations (such as graphs, charts or tables) with aggregated and meaningful results. In this way, a client company not only can find out how much user reach it has, but also the general feeling about their services. For example, the presence of a company logo in a tweet or Facebook post is a strong indication that the user made a reference to a service or product offered by the respective company, such that an analytics tool that has a NLP (Natural Language Processing) component will feed that component with additional information (i.e. it is very likely that the text is related to our company) so it can be more effective.

1.2 Objectives

In this thesis we propose a system that focuses on logo mentions on social networks streams. Although the tool is designed to work and run independently, its major power would be paired with an existing social media analytics platform, like the one that uberVU via Hootsuite¹ offers. The project was designed and built from scratch in order to enrich their existing tools by adding statistics about logo mentions and was done in collaboration with engineers from Hootsuite.

One of the main objectives is to offer a very scalable system that accepts multiple data sources (social media streams), identifies posts with images and process those images in order to find logos, all in real-time. Before going next, let us be clear with two apparently exhaustive terms used here: *very scalable* and *real-time*. By a very scalable system we understand a software program that, if offered $N^2\%$ more machines (processing nodes), will perform approximately $N\%$ faster. Of course, we say *approximately* because we cannot ignore Amdahl's Law³. Probably not as intuitive, by *real-time* we understand, in this case, a system that can fetch and process data from a streaming endpoint of a social network without falling behind. To be even more clear, if Twitter offers 1% of total tweets, that gives us approximately 6 million tweets to process in a day. A simple math yields that this system has to process around 420 tweets per minute, or, more interestingly, 30 images per second only from Twitter.

Another objective of this project is to build a modular system with loosely coupled components, so that a new component can be inserted into the pipeline at any future time without too much effort. With this in mind, this system is designed as a chain of smaller self contained systems that can function independently and only require an input stream of data and produce an output stream of data. For example, one component can be the one that fetches data from the social networks and outputs only relevant posts, in our case the ones that contain images. The key objective here is that we can add a new component in-between two existing ones without changing them.

Coupled with the above statements, the system not only has to identify logos in downloaded images, but it can do any type of image processing. Although it may seem so, the primary focus of this project was not building a tool that only recognizes logos in social media streams, but a tool that can do any type of processing with social media streams, providing it has a dedicated component that does the work. Because we obviously needed a component of this type in order to have a fully functional pipeline and because it is indeed useful for social media analytics, a large part of this thesis concentrates on the logo detection component, but the idea here is that the bigger purpose is the actual pipelined infrastructure.

¹<https://hootsuite.com/products/social-media-analytics/ubervu-via-hootsuite>

²Where N is an arbitrarily chosen number

³<http://home.wlu.edu/~whaley/classes/parallel/topics/amdahl.html>

1.3 Use Cases

From the user perspective, we designed the system far from being verbose, the only visible output that it produces being a real time web dashboard that offers various statistics about detected logos built on top of annotated data stored in a database that gets polled once in a while. Nothing very complicated, all the magic happening all the way from fetching web streams to storing identified data in the database, but it is supposed to be like that, the system aiming to be a smaller part in a large dashboard for a complete analytics tool.

Another use case is, for example, to couple this system (or even add a separate component) with another one that does object or scene recognition, so that the upper layers of analytics have a better insight on the context that the logo was identified. It is clearly an advantage to know that Coca Cola's logo appeared on a 330ml can that takes 50% of the image or on some random guy's T-shirt that points a bended iPhone 6¹ to the camera and looks angry.

Along these lines, we can also imagine a company that recently launched a new product and wants to know how much buzz has created around that product, especially on social networks. It is true that there are many factors that matter here, like user sentiment analysis, for example, but if we only see a recent spike in the number of company logo occurrences on social media streams, it can be an indicator that people are talking about it, that they know about the product. Of course, it is not enough to have a decisive conclusion, but combined with other tools we can definitely have richer results.

From the technical point of view, we designed the system to have multiple, usually replicated, components that run as background jobs and do not do any interaction with a human user. Communication between components is done using JSON objects, such that the components are completely decoupled. This facilitates running the whole system in a cloud environment, like Amazon EC2². So, using the system consists in starting some jobs and using the graphical interface that polls a database that gets updated every time we have some new results (identified logos in our case).

1.4 Thesis Summary

We have seen in the previous three sections what was the motivation of building this project, what objectives we concentrated on and we also saw a couple of situations where it would be useful to have a logo recognition system.

In the next chapter, *Related Work*, we will talk about some interesting algorithms and ideas that are currently emerging in the logo recognition world. We will see that logos have very specific features that sometimes may seem useful (i.e. most of them have 2-3 colors and those are usually primary colors), but in fact are rather unhelpful.

In the third chapter we present the detailed implementation of the system, with deep insight in every component as well as some general facts about a messaging system used to pass around information between components. In addition, we will talk briefly about Scala, the programming language used in our implementation and we will offer a detailed description of the algorithm used for logo recognition.

The fourth chapter focuses on testing and evaluation of current solution. We will crunch some numbers regarding performance, machine costs, scalability or even a precision and recall comparison.

¹<http://goo.gl/2w3Lwz>

²<https://aws.amazon.com/ec2/>

The last two chapters present the conclusions and some ideas of future improvements to the present implementation, not only from the algorithmic point of view, but also regarding the underlying infrastructure.

Chapter 2

Existing Work

Like we saw in the previous chapter, social media analytics has become more and more important, therefore the businesses centred around it have gained a lot more attention because they produce a lot more impact. Over the past few years, whole companies like Hootsuite or Sprout Social² emerged around the social media ecosystem offering a large variety of tools not only for analytics, but for optimizing social media experience in general. Not unexpectedly, giants like Google or Facebook have built tools that help their users to better understand how they perform on their very own social network.

However, all these above solutions can be called rudimentary, at most, when it comes to logos. Most of them do not even take logos into account, while others have recently started to do some work in the domain. Of course, there is already some progress done in the industry and it seems quite professional, but its main disadvantage is that it is proprietary technology. The solution that Ditto Labs, Inc³ is promoting is one of the very few in the industry and seems promising offering visual search in social media streams, but it is not open source so we cannot say much about their approach.

In the next section we will explore one of the most innovative applications of image recognition in social media, while for the rest of this chapter we will concentrate on two techniques of logo recognitions, one of them featuring many written papers in the Computer Vision domain.

2.1 Image Based Analytics in Social Media

One of the new startups in the social media analytics industry is Curalate⁴, which, according to their claims, is the world's leading marketing and analytics suite for the visual web. "Their business started around Pinterest and tried to analyze Pins for their potential to expose campaigns and products. As they began to build the business, they realized many people do not use text on Pinterest as they do on other social media sites and therefore needed to figure out a way to specifically analyze images on their own"⁵.

Curalate's main idea is that people started to prefer communicating using images, especially on social networks dedicated for this, like Pinterest or Tumblr, and so they try to use object recognition techniques to offer better insight to their clients. Curalate CEO, Apu Gupta, said in an interview: "Before people buy, they save an image of a product and pin it on Pinterest

²<http://sproutsocial.com>

³<http://ditto.us.com>

⁴<http://www.curalate.com>

⁵<http://www.psfk.com/2014/06/curalate-social-image-analysis.html>

or post it on Tumblr. After they buy, they might take a selfie with their new sweater and post it on Instagram."

One true thing that stands for all the work done in the domain is that none¹ of the companies involved released an open source platform that allows multiple types of social media image processing, not only logos or objects.

2.2 Object and Logo Recognition

Let us dive now in the realm of object recognition as a stand alone activity, not paired up with social media or marketing purposes. Fortunately, the Computer Vision is a well studied and researched field and there is constant progress in the area. Before proceeding, let's be clear with the term *object recognition* which can be defined as the "process for identifying a specific object in a digital image or video. Object recognition algorithms rely on matching, learning, or pattern recognition algorithms using appearance-based or feature-based techniques. Common techniques include edges, gradients, Histogram of Oriented Gradients (HOG), Haar wavelets, and linear binary patterns."²

In the next sections, we will see some of the techniques used for object or logo recognition together with advantages and disadvantages. Before that, let's see what is different in recognizing a logo compared to recognizing an object, because if there were no differences we could simply use any type of object recognition approach and apply it in our case.

The first one that comes to mind, if we look at [Figure 2.1](#), may probably be the *color*. One of the most definitive characteristics of a logo is that its color changes very rarely, or, if it actually changes, we can almost consider it to be another logo. There are brands that change their logo color drastically over time (i.e. Apple), but that is not the common case. Another important fact about the color is that not only there are few or inexistent changes, but there are few colors that appear in the logo itself, to begin with. If we take the Twitter, Dropbox or Honda logos, we can clearly see that they are made of one single color, maybe with slight gradients.



Figure 2.1: Random Logos

¹To the best of our knowledge

²<http://www.mathworks.com/discovery/object-recognition.html>

Even the Metro Goldwyn Mayer logo, which seems very different compared to the other 7 logos, has one dominant color. Exceptions from this rule make Google and the old Apple logo, but even though they comprise of several colors, these are mostly primary and they do not overlap, so we can see a very abrupt transition from one color to another, one that do not usually occurs in a natural environment.

This observation brings us to the next major difference between a logo and a random object, the *shape*. While an object, despite being the same type of object, can have different shapes and colors (imagine that we are trying to do facial recognition), a brand's logo is always the same. It is one of the key ideas in using a logo that it never has to change in order for people to remember and recognize it easily, so we can also use it to our advantage. Therefore, we can assume that a specific logo, like the Google one, for example, will appear with the same font, the same aspect ratio, the same amount of space between its letters and so on and so forth. In some cases, it is even illegal to change the aspect of a logo, so we can rely on this and use it to our advantage.

That being said, we can also do another important observation regarding logos: they are very rarely, if at all, rotated. There are cases when we have to analyze an image with a Coca Cola can that is flipped on one side so the logo is also rotated 90 degrees, but these are situations that happen in less than 10% of the images used to build this project. With this in mind, we can safely state that considering logo rotation is not the biggest concern regarding the scope of this thesis.

To sum up the above observations, we have seen that three features of a logo usually stay the same across multiple occurrences: *color*, *shape* and *orientation*. Taking this into consideration, when we speak about logo recognition a natural idea should come to mind: *Template Matching*.

2.2.1 Template Matching Approach

This technique is one of the simplest, yet effective ones used in this area of study. In template matching, the query image is matched against an existing template according to a similarity measure, for example the Euclidean distance. "The templates can be represented as intensity or color images, when the appearance of the object has to be considered. Appearance templates are often specific and lack of generalization because the appearance of an object is usually subject to the lighting condition and surface property of the object. Therefore, binary templates representing the contours of the object are often used in object detection since the shape information can be well captured by the templates. Given a set of binary templates representing the object, the task of detecting whether an image contains the object eventually becomes the calculation of the matching score between each template and the image." [1]

One place where we can find a good implementation of the template matching algorithm is the open source computer vision library, OpenCV[3]. We will see later, in [Chapter 3](#) that we are actually using the OpenCV C++ library within the logo detection component. As [Figure 2.2](#) shows, the main idea here is to have a template smaller than the actual query image (in case we find ourselves in the trivial case where the template is bigger than the query image it means there is no possible match) and try and fit the template in every possible position on the original image. If the query image is big enough and the sliding box small compared to the former, it is not that hard to notice an obvious flaw of this approach: it can be very slow in some cases. Fortunately, in real life situations, there is usually no need for matching at high resolutions, so both images can be safely resized to much smaller sizes as long as we do not lose too many details, especially in the template image.

Given its nature, the template matching technique is mainly used in situations where the object characteristics are rather constant or where speed is more important than very high accuracy.



Figure 2.2: Sliding Template

If we remember the numbers stated in [Section 1.1](#) from [Chapter 1](#), it is not a surprise that this type of algorithm was the main choice in the implementation.

TODO:

Present this type of technique. Implementations. Usages. Advantages/Disadvantages. Performance blah blah

2.2.2 Learning Approach

TODO:

Present this type of technique. Implementations. Usages. Advantages/Disadvantages. Performance blah blah

Chapter 3

Proposed Solution

TODO:

Overview of this chapter. Focus on the fact that the system is algorithm agnostic. It can basically do any type of image computation.

3.1 Requirements and Technologies Used

TODO:

Explain what were the requirements, i.e. * the system has to be written in Scala (Hootsuite requirement) * has to be very easy to scale horizontally * has to offer a simple/fast way of introducing a new logo to the database * has to store the positive images to a database and annotate them with the found logo

3.1.1 Kafka – A Messaging System

TODO:

Talk about general Kafka architecture, why is it good in this situation

3.1.2 Scala

TODO:

General stuff about Scala. Advantages/disadvantages

3.2 System Architecture

TODO:

Prepare the next subsections

3.2.1 Fetch Module

TODO:

Explain how images are fetched from the Social Network, how do we store them. How are they passed to the next components etc.

3.2.2 Detector Module

TODO:

Explain how does a detector work. Where it pulls data from. Focus on the fact that detectors are algorithm agnostic

3.2.3 Annotated Data Storage Module

TODO:

Where is data stored. How is it stored. How do we pull data from the database

3.2.4 Queing System

TODO:

Explain how messages get passed around. Talk about message types. What data do we send. How do components fetch images without sending them around.

3.2.5 Data Flow

TODO:

Talk on a diagram that explains the complete dataflow

3.3 Logo Recognition Algorithm

TODO:

Explain the algorithm. Talk about the general idea. Why is it good in our situation.

Chapter 4

Testing and Evaluation

TODO:

Prepare the next sections. Explain why they are relevant for system evaluation

4.1 Algorithm Performance

TODO:

Present some numbers of how fast the algorithm is.

4.2 Cost and Resource Consumption

TODO:

Do an estimation of how many machines the system needs in order to be effective (i.e. do not fall behind of social media streaming endpoints). Resource consumption etc.

4.3 Precision and Recall

4.3.1 Precision and Recall in Pattern Recognition

TODO:

General stuff about precision and recall. Why are they a good benchmark

4.3.2 Precision and Recall for our Algorithm

TODO:

Do an evaluation based on a table with precision/recall for our algorithm

4.4 Scalability and Portability

TODO:

Talk about how easy to scale the system is. How portable it can be etc

Chapter 5

Further Work

5.1 System Architecture Enhancements

TODO:

What components can be improved. Fail recovery etc.

5.2 Better Logo Recognition Algorithm

TODO:

Some general ideas about what could have been done better regarding the algorithm

5.3 Other Possible Social Media Image Processing

TODO:

State some ideas of how one can use this system for other types of image processing, not only looking for logos

Chapter 6

Conclusions

Bibliography

- [1] Philip Ogunbona Duc Thanh Nguyen, Wanqinq Li. An improved template matching method for object detection. http://link.springer.com/chapter/10.1007%2F978-3-642-12297-2_19, 2010.
- [2] Statista Inc. The statistics portal. <http://www.statista.com>, 2015.
- [3] opencv dev team. Opencv documentation. <http://docs.opencv.org>, 2011-2015.