APRIL 28, 2019

IMAGE CONTENT ANALYSIS

RESEARCH PAPER

MATTHEW LYDON
COMMUNICATIONS FOR IT PROFESSIONALS

Abstract

With an increasing number of digital cameras and sensors being used to collect image data computers need a way to sort through all the information to make decisions and solve problems.

This paper presents the results of a library research project which studied the topic of Image Content Analysis. It's broken down into two main sections, each an important step for preforming modern Image Content Analysis.

- Edge Detection
- Machine Learning

Data Scientists use Image Content Analysis as a tool to help understand the raw data collected from an image. This research report serves as a brief introduction to the techniques used to perform Image Content Analysis.

Image Content Analysis

How do we use technology to gain a better understanding of our physical world? One of the most effective ways is through imaging, taking pictures or video, and giving that information to a computer as an input. This research paper will discuss the field of Image Content Analysis, which is what a program must do in order to provide an output based upon an image input. Image Content Analysis has been used in practice for decades to solve problems and provide image services, however two technological developments have prompted research institutions and companies to dive deeper into this area.

With the rapid increase in digital camera technology, the quality of images has improved significantly. While there have been advanced digital camera sensors available for quite some time now, the increased production and lower costs has resulted in much greater access to this technology. Simultaneously, the field of machine learning has also made significant breakthroughs. By pairing these two fields together, improved camera hardware and machine learning software, the field of Image Content Analysis has grown significantly. Researchers and developers are finding ways to use Image Content Analysis to solve more advanced problems and create exciting new products.

There are many similar fields and alternative names for Image Content Analysis, some of which include Computer Vision, Machine Vision, Computational Image Analysis, and Computing Imaging. Many of the topics covered in this paper refer to each of these areas, and differences only occur at a very advanced level. To keep things consistent and provide clarity, during the remainder of this paper the term *Image Content Analysis* will be used in place of and encompass these fields.

Many of the Image Content Analysis advancements being made occur when it is being used to solve research question. For example, "[in] modern astrophysics, machine learning has increasingly gained popularity with [researchers through] its incredibly powerful ability to make predictions or calculated suggestions for large amounts of data" (Bai, Liu, Wang, & Yang, 2018). In the pursuit to solve a completely different challenge, Image Content Analysis is being used to detect lane markings and external hazards to give automobiles the ability to drive autonomously (Vacek, Schimmel, & Dillmann).

Edge Detection

In order for a program to analyze the different objects within an image, it must be able to distinguish where one object ends, and another begins. Two methods have been key to solving this problem, Edge Detection and Bounding Boxes.

A report from researches at U.C. Berkeley and the University of British Columbia defined three key questions each team must ask when performing Image Content Analysis: "What counts as an object? [...] Which objects are easy to recognize, and which are hard? [...] and which objects are indistinguishable [...]?" (Duygulu, Barnard, de Freitas, & Forsyth, 2002). Edge detection begins when a computer model takes an image input and analyzes the **differences** between pixels. One of the most commonly used models is Nearest Neighbor detection. In this model, differences between the pixels *nearest* to each other are determined, and areas with the biggest discrepancies are used to define the edge. The second step of edge detection is **smoothing**, which takes the jagged edges produced earlier and uses a number of mathematical decisions to create a set of vectors. Smoothing comes with a positive and negative tradeoff. On one hand smoothing can "reduce noise and [...] ensure robust edge detection", however it also results in a loss of detail (Ziou & Tabbone, 2010). The best edge detection systems find "a favorable compromise between edge detection and noise reduction" (Ziou & Tabbone, 2010).

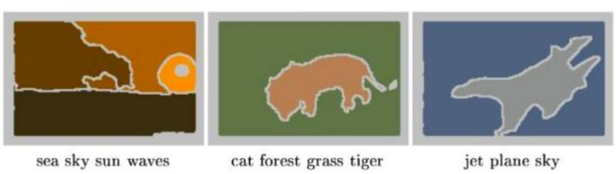

Figure 1 | Demonstration of how a computer views images using edge detection (Duygulu, Barnard, de Freitas, & Forsyth, 2002).

For a computer to make a understand what the outline is displaying, it must be linked to a database where a statistical model makes a prediction. Translating images into information begins in the following section when this paper discusses machine learning.

Machine Learning

To better understand how machine learning is playing its part in Image Content Analysis, we'll take a moment to briefly describe the machine learning process. Machine learning occurs when data scientists 'feed' large amounts of information into a computer program in a process called ***training***. This information contains both the x-variable (the input), and the y-variable (the correct output). For the purposes of this paper, we will focus on how machine learning takes pictures as it's form of input. The machine learning program takes this information and creates a statistical model based upon the relationship between the x and y-variables. After training the program and receiving a statistical model, researchers ***test*** the model to make sure it works properly and find areas which need improvement. To do this, data scientists have a set of data which the program has never seen before, which contains both the x and y-variables. The data scientists then load only the x-variable data (images) into the program, and check to see if the model is any good at guessing the correct y-variables.

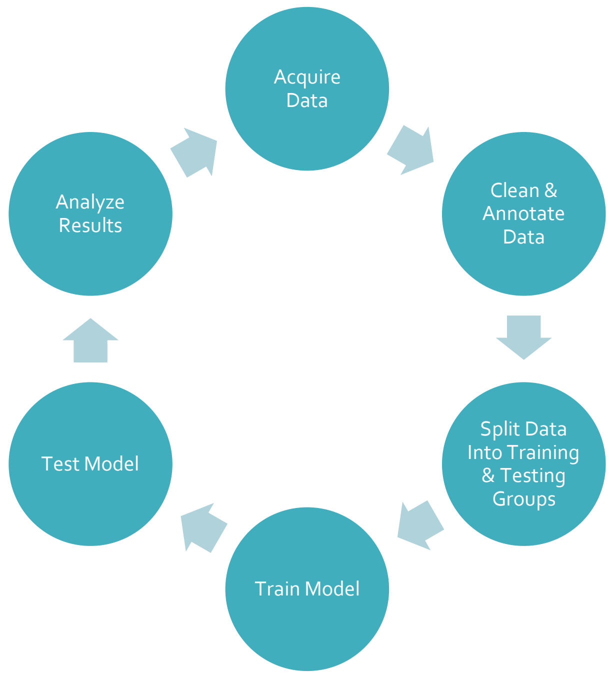

Figure 2 | Machine learning process

To get a better understanding of this process, we can evaluate how a feature available on Google was created in response to a research challenge using Stanford's Dog Dataset. The goal: create a model that uses Image Content Analysis to determine the breed of a dog based on a picture. The Sandford Dog Dataset contains 20,580 images of 120 different dog breeds (Khosla, Jayadevaprakash, Yao, & Fei-Fei). Each image (the x-variable) was annotated with the breed of dog in the image (the y-variable). A number of these images were set aside for testing to make sure that the soon-to-be created model functions properly. Data scientists and programmers then worked to create a statistical model by providing machine learning software with the Stanford Dog Dataset. Once the data was processed, they tested the model's accuracy by showing it images of dogs from the dataset that was set aside. The machine learning program uses its generated model to guess the breed of each dog in the image and returns its guesses. Next, data scientists determine if the model's answers fall within an acceptable margin of error. The process is repeated with additional data to make the model more accurate.

The key driver to success in machine learning is data. With more data comes greater opportunity to perform the final step of the machine learning process, ***repeat***. This is why research institutions and companies with access to larger sets of data have better success at creating machine models. The reason Google is so successful at the example above is because they have access to many pictures of people's dogs. Not only do they have many pictures of dogs, but those pictures vary. If all the pictures given to the machine learning program were of poodles in dark rooms, the model would be no good and detecting pictures of golden retrievers in bright sunlight. Giving a machine learning program too much of the same thing leads to bias, also called ***overfitting***. The more variance of data given to a machine learning program, the more accurate the final model will be.

One company taking advantage of their access to large amounts of data is Tesla Motors, who recently demonstrated their machine learning model used for autonomous driving. Every Tesla vehicle on the road has cameras capturing pictures of the road (x-variables) and a log of the actions being performed by their drivers (y-variables). Tesla explained how "hundreds of thousands of cars driving all over the world, [provide] a large, varied, and real dataset" (Gurskiy, 2019). The advantage of having more data is crucial to any machine learning problem, and thanks to improvements in camera quality and reduced cost, image data is more available than ever before.

Almost all modern methods of Image Content Analysis involve some form of machine learning to classify and predict the relationship between images and annotated databases. Machine learning is a powerful tool being used to conduct Image Content Analysis, and Image Content Analysis is one of the leading causes for advancements in machine learning tools

Conclusion

Each of the techniques discussed in this paper are used together to perform Image Content Analysis. By developing more advanced techniques for Edge Detection and Machine Learning, software can become more adapt at comprehending the natural world. This research report was conducted to gain a base-level understanding of techniques being used by researchers and companies who perform Image Content Analysis. Hopefully it will leave readers with the desire to dive deeper into this very relevant subject

Works Cited

Bai, Y., Liu, J., Wang, S., & Yang, F. (2018). Machine Learning Applied to Star–Galaxy–QSO Classification and Stellar Effective Temperature Regression. *The Astronomical Journal*, 157:9.

Duygulu, P., Barnard, J., de Freitas, & Forsyth, D. (2002). Object Recognition as Machine Translation: Learning a Lexicon for a Fixed Image Vocabulary. *European Conference on Computer Vision*, 97-112.

Gurskiy, D. (2019, April 25). *Tesla Autonomy Day Recap – Tesla FSD.* Retrieved from EV Bite: https://evbite.com/tesla-autonomy-day-recap/

Khosla, A., Jayadevaprakash, N., Yao, B., & Fei-Fei, L. (n.d.). *Stanford Dogs Dataset.* Retrieved from Stanford University: http://vision.stanford.edu/aditya86/ImageNetDogs/

Vacek, S., Schimmel, C., & Dillmann, R. (n.d.). *Road-marking analysis for autonomous vehicle guidance.* Institute for Computer Science and Engineering, University of Karlsruhe, Karlsruhe, Germany.

Ziou, D., & Tabbone, S. (2010). *Edge Detection Techniques - An Overview.* Québec: Université de Sherbrooke.