Generate textual descriptions of STEM images to aid Visually-impaired

Abraham Daniel Williams, Kirthanaa Raghuraman, Raghuram Janakiraman

Problem to be Solved:

Classify infographics of science, technology, engineering, and mathematics and come up with descriptions to make them accessible to individuals with visual and print-reading disabilities.

Motivation:

Morash et al., have done a prior research to identify the best practice that can be used to describe STEM images to people with vision loss in an efficient way. Image accessibility in STEM is particularly important, because STEM topics often rely heavily on images to convey information that is not presented in accompanying text [Jayant et al. 2007; Ladner et al. 2005].

This paper shows a comparison of the FRID (Free Response Image description based on NCAM guidelines) and QID (Query Based Image Description) methods. It establishes that QID has better results in terms of consistency among descriptions, time taken to generate a description. QID refers to the method of classifying the images initially and generating a set of queries based on the image category pertaining to NCAM standards and crowd sourcing to answer the queries which can then be formulated into a textual description. Involving crowd sourcing incurs costs and unreliability in results depending on the understanding of the question posed by the labor. This is the motive for our project and we aim to go beyond this paper to classify the STEM images and then extract information from the image to fill in the template for textual description of the image rather than crowd sourcing which increases efficiency via cost-cutting.

Prior Research:

Manolis Savva et al., talk about classifying images into different categories of visualizations and automatically re-designing them to improve graphical perceptions. Their work involves two main steps:

1. Classifying images
2. Data and mark extraction.

They have used a traditional SVM classifier to classify the images. Post classification, they extract textual features from the charts using a series of steps. This step is broken down into mark and text extraction. The percentage of images from which they are able to extract useful information from decreases at every stage and consequently, the extraction was possible only for 60% of the images and they could not achieve a higher success rate in redesigning the info graphics.

Outline of steps:

The steps involved in achieving our end goal can be outlined as follows:

1. Understand the corpus of info graphics and come up with the categories of charts that we want our algorithm to identify.

2. Use a classifier (deep belief networks/ convolutional neural networks) to classify the info graphics.

3. Extract necessary features essential to identify information to describe the chart.

4. Come up with a textual description from the data extracted.

Timeline:

Date	**Goal**
Feb 15	Project Proposal
Feb 22	Research more about the project. Start implementation of classification.
Mar 14	Finish classification of charts.
Mar 21	Work on identifying features that will help describe the infographic for each chart type.
Mar 28	Finish writing mid term report.
Apr 15	Extract features to aid textual description.
Apr 28	Finish textual description module and fix issues.
May 5	Final Report

References:

1. Valerie S. Morash et al., "Guiding Novice Web workers in making image descriptions using templates".

2. Mannolis Salva et al., "ReVision: Automated Classification, Analysis and Redesign of Chart Images", UIST '11.