

Project title:

Improving an algorithm to identify and count books

Project abstract:

Our project aims to identify and count the number of books contained in the images to be submitted by respondents to a web survey. An algorithm for these purposes already exists, but it needs to be improved as it presents several problems (no identification of books, or assignment of only one box to two or more books). The aim of the thesis is to overcome these problems by improving the algorithm.

Extended abstract:

Surveys in social science have repeatedly used the question on the number of books at home to characterize dimensions such as people's cultural and economic capital, or to explain academic performance. This question is usually answered using response categories with intervals. Consequently, there is no information on the exact number of books. In addition, another relevant dimension such as the title of the books is not asked because it would represent a huge burden for the respondents. In this scenario, this question could be improved in order to obtain higher accuracy and more insights, and to reduce the respondents' burden.

One way to improve this question is to ask respondents for photos of the books they have at home. Through the classification of such images, it would be possible to count the number of books and, if the quality of the photos is good enough, it would also be possible to detect the title of the books.

Ideally, this classification could be performed automatically by an algorithm that allows identifying and counting the books. State-of-the-art detection methods based on deep learning present certain problems to identify some of the books (missing some of them), especially when they are in horizontal position or have an object in front of them, and to assign the correct area for a book (sometimes, the algorithm assigns a single box for several books). Both types of errors are presented in the image below.



Thus, we would like to work with a student interested in improving a state-of-the-art detection algorithm based on Mask R-CNN. It is expected that a) techniques to avoid repetition of items (in case a book appears in two or more images sent by the same respondent) will be implemented, and that b) the problems explained above will be solved to a large extent. Other improvements are also welcome, such as implementing Optical Character Recognition on the areas where books were identified in order to identify the title of the books.

This master dissertation will be developed as part of the project "New opportunities to enhance or extend (mobile) web survey data and get better insights", conducted by Dr. Melanie Revilla from the Institut Barcelona d'Estudis Internacionals (IBEI).

Academic supervisor Name: Dr. Gloria Haro

Academic Supervisor e-mail: gloria.haro@upf.edu

Academic Supervisor Institution: UPF