Horizon 2020 Framework Programme

Grant Agreement: 732328 – FashionBrain

Document Information

Deliverable number: D5.1

Deliverable title: Scalable Crowdsourced Social Media Annotation

Deliverable description: This deliverable consists of a publicly available website with data visualization functionalities. We demonstrate that we analysed hundreds of fashion blogs, instagram profiles and that we are able to constantly update the profiles with recently published images.

Due date of deliverable: M18

Actual date of deliverable: June 30th, 2018

Authors: Marko Jocic, Matthias Dantone

Project partners: Fashwell; USFD

Work package: 5

Work package Leader: Matthias Dantone

Change Log

Version	Date	Status	Author (Partner)	Description/Approval Level
0.1	29.05.18	Final	Fashwell	Public

Table of Contents

Summary. This document illustrates an overview of Deliverable 5.1. The deliverable is a public available website and this document gives a quick overview of the steps needed to create this demonstration.

1 Introduction

Social Media, image sharing platforms and bloggers have transformed the entire fashion discovery journey. Which fashion product is popular, trendy and a best-seller is defined by its presence on social media. Fashion bloggers are used by brands to distribute and influence this process. Bloggers are getting paid by brands and retailers to wear a specific product because they have a significant reach on the relevant social media platforms.

In this deliverable we crawled 100 fashion blogger Instagram accounts and identified all fashion products in the images. The images including the annotated products have been made available to the public on this website: link.

2 Demonstration Website

The demonstration website displays the images. The linked products can be seen by clicking on an image. The screenshot below shows the layout.
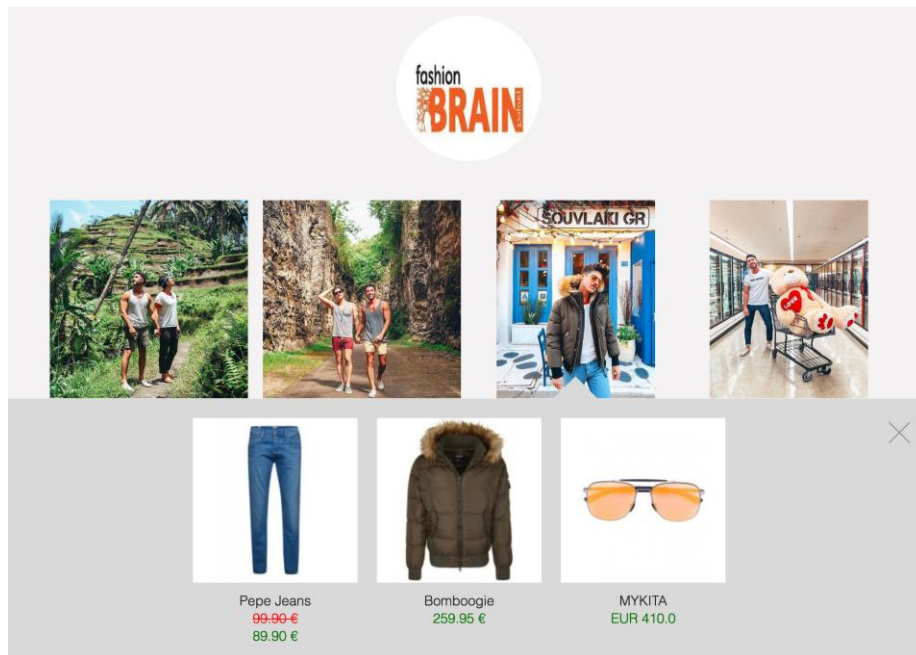
Screenshot of the publicly available website that shows the collected fashion blogger images and the manually annotated products.

The website is hosted on Google Cloud and was built using the following technology stack: Django, React.js and CSS.

3 Data

For this WP we downloaded around 100 images from 100 Instagram bloggers, which resulted in a total of 11'007 images. The list of the instagram users have been the output of the deliverable D6.3. The Instagram API [1] has been used to download the images and the corresponding metadata like number of likes, comments and title.

Some statistics about the images processed are shown below. An image classification algorithm has been used to analyse the content of the images.
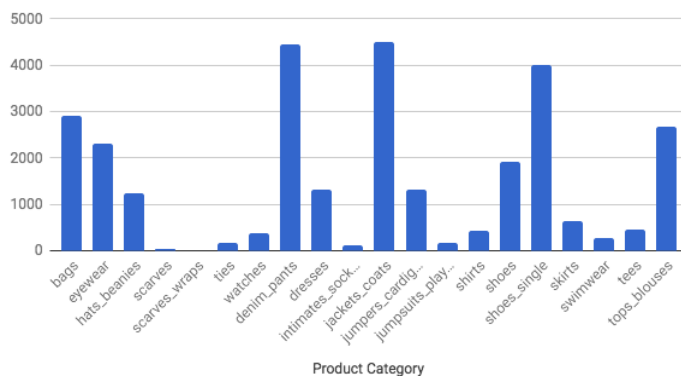

The most common category of products detected in the images were denim pants and jackets and coats. less popular categories were scarves, wraps and socks. A number of categories sat midway in popularity such as bags, eyewear and shoes.
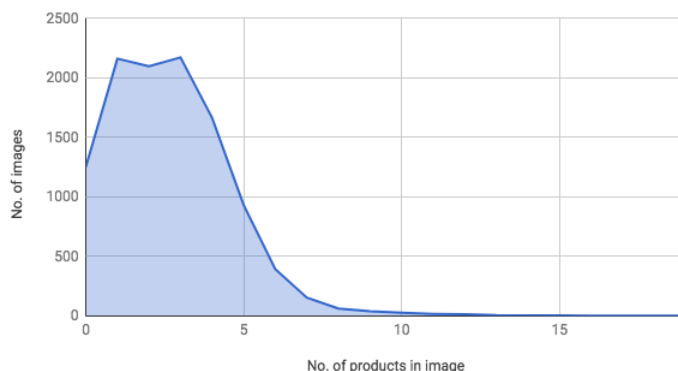
Most images had between 1 and 4 products detected, with very few images having more than 6 products. Some images did not have any products detected, this is due to the use of other post types such a text only posts or nature shots.
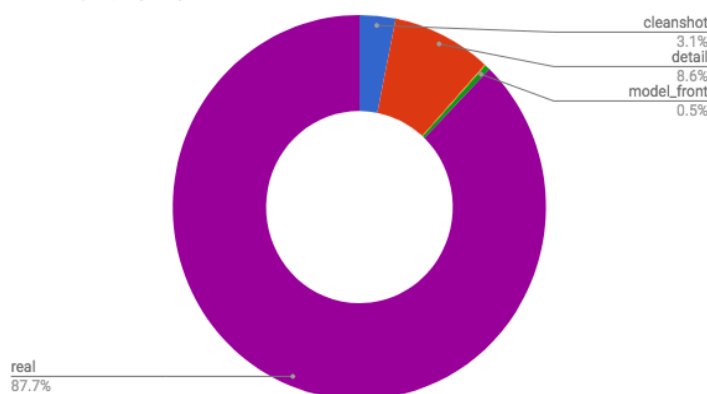
Fashwell proprietary image classification algorithms detected the style of photography. The vast majority of images were detected as "real" images, indicating an image with a non-studio background. The next most detected photographic style was detail shots, indicating an image that does not include all of the product. Clean shot images represent full product pictures. These were the least common.

4 Annotation

In this section, Fashwell quickly reviews the work presented in Deliverable 6.2 and Deliverable 3.1 and shows how the work developed in these two WP contributed to this deliverable.

All 11'007 fashion blogger images have been linked/annotated with the most similar products of the product database. The annotation process can be divided into two different tasks:

i) the automatic annotation and
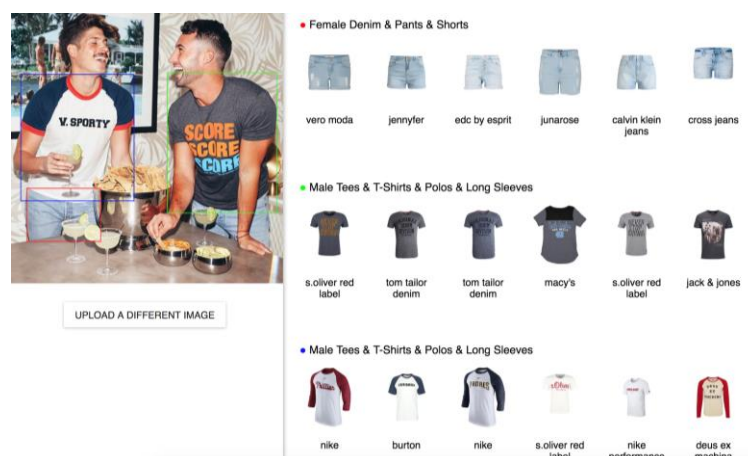
ii) refinement using crowdsourcing.

The automatic annotation takes advantage of state of the art computer vision and deep learning to find the most similar products in the database. The second task leverages the wisdom of the crowd to refine and verify the automatic annotation.

4.1 Automatic Annotation

In the Deliverable 6.2, we presented our work in the FashionBrain project in correspondence to the task of named entity linking using computer vision and machine learning. The presented approach outperforms all publicly available methods on the academic benchmark presented in (Liu et al. 2016 [2]) (details about the experiment can be found in D6.2).

The conceptual idea behind our method is to understand an embedding space where visually similar products are close to each other, whereas visually dissimilar products are further apart. Visual similarity can be described with many different features of a product such as the category, color, pattern, neck type, sleeve length, etc.

The presented approach localizes different fashion objects in the fashion blogger image and 'projects' the relevant product image patch into the embedding space. The distance in that embedding space is used to describe the similarity. Fashwell's approach outperforms current state-of-the-art methods, but there are still areas to improve. An example result of the methods can be found below.


Example of the automatic annotation. We see a clear mistake of the pants. The automatic algorithm recognized female shorts, but it's most likely not a pair of short and definitely not for women.

4.2 Annotation Refinement with Crowdsourcing

Since the automatic annotation approach is not perfect, we built an online tool where workers around the world could help us verify and improve the annotations. The design and concept of crowdsourcing refinement tool has already been discussed in the WP3.

A screenshot of the tool is shown below. The tool starts with the automatic results of named entity linking approach and uses computer vision and deep learning to quickly improve the annotations.
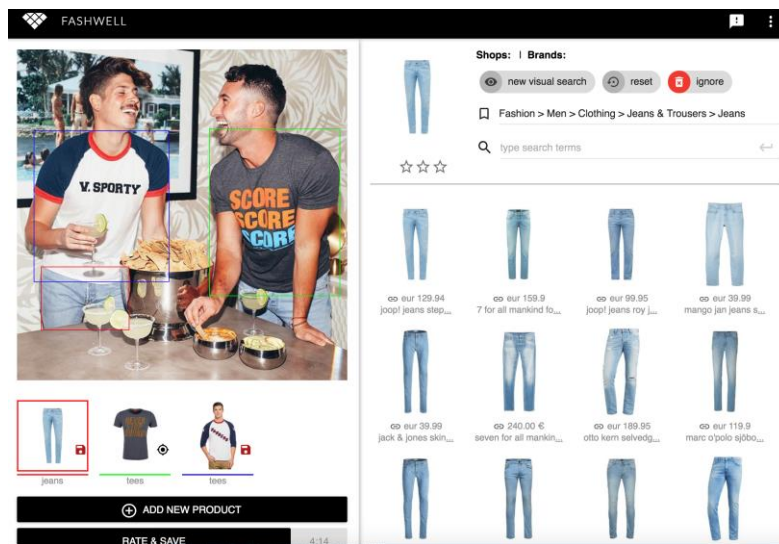
Our crowdsourcing tool allows works to refine the automatic annotations.

On average, a worker needs around 2 minutes to annotate an image perfectly. The time highly depends on the number of products in an image and on the image quality. The workers have been hired using a freelancing platform [3] and have been paid $3.33 an hour.

5 References

[1]   developers.instagram.com

[2]   Liu, Z., Luo, P., Qiu, S., Wang, X. and Tang, X., 2016. Deepfashion: Powering robust clothes recognition and retrieval with rich annotations. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (pp. 1096-1104).

[3]   www.upwork.com