Bodies as Battleground

Feature catalogue (Relative Attributes)

Simon Polichinel von der Maase

November 2021

Contents

Introduction

This feature catalogue describes the features which we are going to extract from the Bodies as Battleground image-dataset as *relative attributes*. That is, features which are measured relative to each other across images (Won et al., 2017). These features will be extracted by letting student assistants annotate pair-wise comparisons denoting whether image A displays more of feature x than image B. The student assistants implement the annotation using a small program called ImageRanker[1], which was written for the task.

For each feature, the pairwise comparisons effectively creates a directed network of images. Given a network pertaining to some feature x, a directed edge between image A and image B denotes the relative ranking of these images with respect to x. For each features and each image, these relative rankings are turned into global scores using Bradley-Terry models (Won et al., 2017, 788). As such all relative attributes will be continues features $\in \mathbf{R}$.

If, given a specific feature, a comparison between two images is deemed to be non-sensible or invalid the student should simply leave the image pair un-annotated with respect to said feature. This maneuver, however, should only be used sparingly.

1   Emotions

The two features below both denote the level negative emotions displayed by the subject(s) depicting in a given image. The two features should not be understood as inverse, conflicting or mutually exclusive in any way.

1.1   Negative Emotions Type 1

This feature denotes what we call negative emotions type 1. Examples of negative emotions type 1 are anger, wrath, disgust and resentment. For examples of negative emotions *not* understood as type 1 see feature 1.2 regarding *negative emotions type 2*. Note that we are here referring to the emotions expressed by the subject(s) depicted in a given image and not the feelings experienced by the annotator

---

[1]https://github.com/Polichinel/imgRanker

while looking at said image.

The image which contains the subject(s) expressing highest degree of negative emotions type 1 should here be annotated (ticked in ImageRanker).

1.2   Negative Emotions Type 2

This feature denotes what we call negative emotions type 2. Examples of negative emotions type 2 are sadness, fear, despair, sorrow and grief. For examples of negative emotions *not* understood as type 2 see feature 1.1 regarding *negative emotions type 1*. Note that we are here referring to the emotions expressed by the subject(s) depicted in a given image and not the feelings experienced by the annotator while looking at said image.

The image which contains the subject(s) expressing highest degree of negative emotions type 2 should here be annotated (ticked in ImageRanker).

2   Mass protest

This feature denotes the level of mass protest depicted in the images. Mass protest is here understood as mobilization of crowds or masses such as demonstrations, large manifestations, riots and so forth. A mass protest does not need to exhibit violence, but events such as carnivals, parades, festivals and sports celebrations should not be considered mass protest.

The image which contains the highest degree of mass protest should here be annotated (ticked in ImageRanker).

3   Damaged property and infrastructure

This feature denotes the level of damage which property and infrastructure, depicted in a given image, have sustained. By property and infrastructure we here refer to buildings, roads, bridges, factories, culture or religious sites and so forth. The actual act of destruction is rarely expected to be depicted and as such the presence/absence of this act should not not influence this feature.

The image which contains the most damaged property and/or infrastructure should here be annotated (ticked in ImageRanker).

4 Scene

The features below all pertains to the general scene of the image. All examples given below are merely indicative heuristics and not hard guidelines.

4.1 Private/homely

This feature pertains to how private and/or homely the scene of a given image appear. As heuristic examples, a private home, a living room, and a kitchen should be considered more private/homely, while a mall, a courtroom, a plaza, or a press briefing should be considered less private/homely.

The image which appear to depict the most private/homely scene should here be annotated (ticked in ImageRanker).

4.2 Public

This feature pertains to how public the scene of a given image appear. As heuristic examples, a plaza, a city center, a government building, a demonstration, or a press briefing should be considered more public, while private homes, gardens and exclusive spaces should be considered less public.

The image which appear to depict the most public scene should here be annotated (ticked in ImageRanker).

4.3 Militarized

This feature pertains to how militarized the scene of a given image appear. As heuristic examples, military bases, checkpoints, and military escorts should be considered more militarized, while every day interactions solely between civilians should be considered less militarized.

The image which appear to depict the most militarized scene should here be annotated (ticked in ImageRanker).

4.4 Rural

This feature pertains to how rural the scene of a given image appear. As heuristic examples, farms, fields, and wilderness should be considered more rural, while city-scapes, apartment complex', and large infrastructure should be considered less rural.

The image which appear to depict the most rural scene should here be annotated (ticked in ImageRanker).

4.5 Urban

This feature pertains to how urban the scene of a given image appear. As heuristic examples, densely populated, congested or settled areas should be considered more Urban, while scarcely populated and remote areas should be considered less urban.

The image which appear to depict the most urban scene should here be annotated (ticked in ImageRanker).

4.6 Staged/formal

This feature pertains to how staged/formal the scene of a given image appear. As heuristic examples, press briefings, political speeches, inaugurations, and courtrooms are considered staged/formal scenes. People going about there daily lives are not considered formal or staged. Note that we a interested in staged events, and *not* whether the image as such was staged by the photographer.

The image which appear to depict the most staged/formal scene should here be annotated (ticked in ImageRanker).

References

Won, D., Steinert-Threlkeld, Z. C., and Joo, J. (2017). Protest activity detection and perceived violence estimation from social media images. In *Proceedings of the 25th ACM international conference on Multimedia*, pages 786–794. ACM.