Analysis Data Directory: Resized_image_data

Unit of Observation: Image file
- Houses raw image files

Not tabular and has no rows

Analysis Data File: combined_annotations.coco.csv

Unit of Observation: Image file (Images_id)

Variables

Category_id

Definition:
- An encoding of the different types of categories of people within the images
- 1: "Children Without Disability"
- 2: "Elderly Without Disability"
- 3: "With Disability"
- 4: "Non-Vulnerable"

Construction:
- This variable was prepacked in the dataset we used from
  https://github.com/devvansh1997/BGVP/tree/main/Dataset

Images_id

Definition:
- Assigns a unique Id to each image in the dataset

Construction:
- This variable was prepacked in the dataset we used from
  https://github.com/devvansh1997/BGVP/tree/main/Dataset

Images_file_name

Definition:
- filename of a corresponding image in the Resized_image_data directory

Construction
- https://github.com/devvansh1997/BGVP/tree/main/Dataset

Images_height

Definition:

- The height in pixels of each image

Construction:

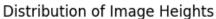- This variable was prepacked in the dataset we used from
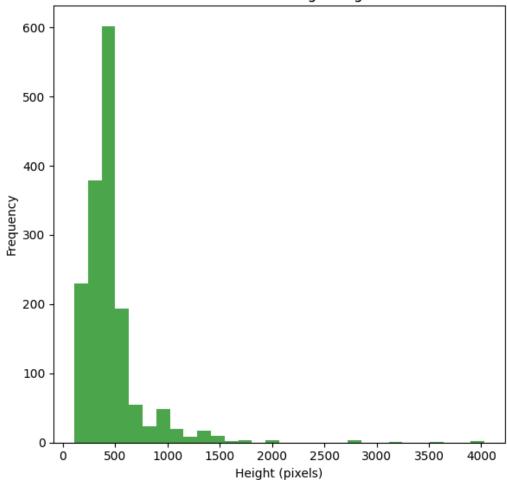 https://github.com/devvansh1997/BGVP/tree/main/Dataset

Total and missing values

- 1601 values, 0 missing.

Summary statistics

- Mean: 454.41
- Std: 315.38
- Min: 107.00
- Max: 4032.00

Distribution of Image Heights

Images_width

Definition:
- The width in pixels of each image

Construction:
- This variable was prepacked in the dataset we used from
  https://github.com/devvansh1997/BGVP/tree/main/Dataset

Total and missing values
- 1601 values, 0 missing.

Summary statistics
- Mean: 616.14
- Std: 371,82
- Min: 103.00
- Max: 6000.00

Distribution of Image Widths

Annotations__category_id

Definition:
- The classification of each image

Construction:
- This variable was prepacked in the dataset we used from
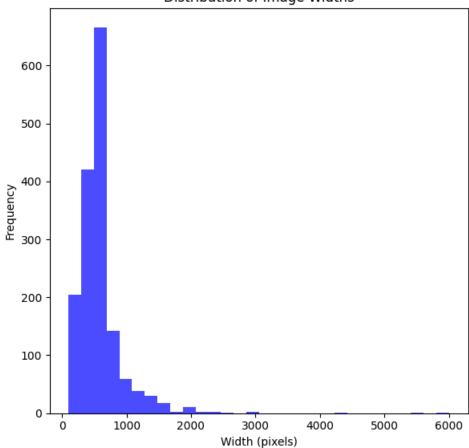 https://github.com/devvansh1997/BGVP/tree/main/Dataset

Total and missing values
- 1601 values, 0 missing

Summary statistics
- Counts for each classification
  - 1 (Children Without Disability): 456
  - 2 (Elderly Without Disability): 225
  - 3 (With Disability): 653
  - 4 (Non Vulnerable): 267

Proportion of Categories in the Dataset