Medical Device YOLOv3 Dataset Documentation

In this document, we report the description and construction of the dataset of images and labels used in order to train a YOLOv3 deep learning architecture to detect the presence of the Zoll vitals monitor. The final weights obtained after training for 500 epochs on this dataset can be found as **yolov3_training_last.weights**.

The overall unseparated dataset of all images is compressed within the **images.zip** file. Separate and categorized images are stored within the separately labeled folders. To preserve Google Drive storage, there are only a sample of representative images within each of the folders. Below, we describe the contents of each folder and how they were obtained.

Python Files: Within the folder, there are a few python files that were used to process labels and images or augment data.
- **annotate_negatives.py** creates blank labels for all .jpg images in the folder **no license**.
- **data_augment.py** creates new images by applying a random perspective transform to the images in the folder **Zoll**. It also distorts brightness and color for better training.

Online Datasets Used
- **Open Image Dataset - An online open source dataset containing images that are prelabeled to contain certain objects.**

  By Kuznetsova et. al in 2020, the Open Image Dataset "is a dataset of ~9M images annotated with image-level labels, object bounding boxes, object segmentation masks, visual relationships, and localized narratives." Using the OIDv4_ToolKit (https://github.com/EscVM/OIDv4_ToolKit), we obtained the images and labels in available classes that we figure the algorithm could falsely label as a license. In particular, the images and labels belonging to the categories "Accordion", "Bathroom cabinet", "Billboard", "Book", "Briefcase", "Calculator", "Door", "Mobile phone", "Paper towel", "Refrigerator", "Suitcase", and "Table" were obtained as a negative set.

  License: CC by 4.0 International
  Source: Kuznetsova, Alina, Hassan Rom, Neil Alldrin, Jasper Uijlings, Ivan Krasin, Jordi Pont-Tuset, Shahab Kamali, et al. 2020. "The Open Images Dataset V4." *International Journal of Computer Vision* 128 (7): 1956–81. doi:10.1007/s11263-020-01316-z.

- **Stanford Background Dataset - A lightweight scene understanding dataset**

  Originally by Gould, Fulton, and Koller, the Stanford Background Dataset contains images with labels that segment and label each pixel in the image. The images contain outdoor scenes, are approximately 320-by-240 pixels, contain at least one foreground object, and have the horizon position within the image. Because of the criteria for these images, we believe that they are representative of an environment wherein the YOLOv3

weights will be used. These images will both serve as control images, and as background images for data augmentation.

License: Creative Commons Attribution-ShareAlike 2.5 Generic License
Source: Gould, Stephen, Richard Fulton, and Daphne Koller. 2009. "Decomposing a Scene into Geometric and Semantically Consistent Regions." *2009 IEEE 12th International Conference on Computer Vision*. doi:10.1109/iccv.2009.5459211.

Categories of Images:

Category	Number of Images w/ Labels
No License	7209
Zoll	1800

- **No License:**

  These images were obtained from the Stanford Background Dataset and the Open Image Dataset and serve as negative training for the YOLOv3 algorithm. There are outdoor images, and images containing objects that we determined that the algorithm may confuse as licenses. Each image's label is empty.

  Examples:
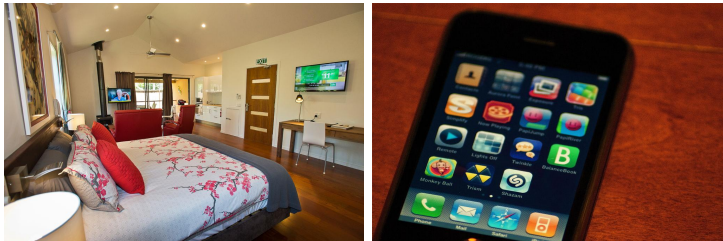
- **Zoll**

  These images contain views of the Zoll monitor and were obtained from feed from the Smart Glasses camera. Data augmentations, such as brightness and color adjustment and slight random homography were applied to increase the size of the dataset. **data_augment.py** creates new images by applying a random perspective transform to the images in the folder **JPG**. It also distorts brightness and color for better training.

  Examples: