Light source presence in images

Filip Comșa

1 Problem Analysis

In order to classify an image as whether it contains or not a light source, we first need a clear definition for what a light source is.

In our case, by a light source we can mean either an artificial source of light, such as a bulb, or we can mean a source of natural light, such as the sun going through a window. What's more, we may also consider reflections as light sources. Our application should be able to recognize all of these sources.

Having this decided, I propose as a solution to this classification problem to create a computer vision machine learning model which would take as input an image and would output the class to which it belongs - "yes" or "no".

2 Comments

In order to accomplish what I have described above, I have decided to use the programming language "Python". The machine learning framework that I have used is **tensorflow**.

As there are no already existing datasets based on light sources, I have decided to create one myself to aid this classification. The dataset's size is rather small, containing about 450 pictures. 90% of this data is going to be used for training, and the rest of 10% will be used for validation purposes.

All of the data should go through a preprocessing and augmentation stage before being used in training with the model, or before the model should try to predict it's result. For this purpose, I have decided to augment all pictures as follows:

- Using a standardized resolution for all images: 250x250 pixels.
- Transforming the RGB interval from [0, 255] to [0, 1].
- Apply the luminance formula using dot product: P = R * 0.2126 + G * 0.7512 + B * 0.0722. The final image will look like a black and white image, but the lighting will be enhanced.
- Applying matrix multiplication multiple times. Applying this will make dark spots darker and lighter spots lighter, thus resulting in a feature enhancement - the features we seek being the lightest of spots.

The model's component will mostly consist of multiple sequences of convolution layers, having a max pooling applied after each sequence.