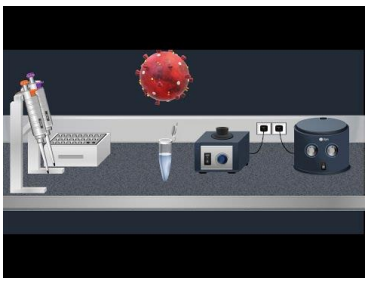
Classification scheme

Classifying thumbnails to investigate the influence of those thumbnails on the recommendation algorithm could be done in very different ways. There are many tutorials on YT about what is important in designing an effective thumbnail. The challenge in creating a classification scheme is to avoid overlap between the categories. I made four categories which should make a clear distinctions between the thumbnails.

1. **Thumbnail without any text or brand logo:**
   This categories only show an image or plain thumbnail background.
   Example:



2. **Thumbnail with text, without brand logo:**
   This category shows an image/background with some textual element(s).
   Example:

3. **Thumbnail without text, with logo:**
   This category shows an image without textual element(s), but with a brand logo. An example of a logo could be the logo of an news site.
   Example:

4. **Thumbnail with text and with brand logo:**
   This category shows an image with textual element(s) and a brand logo.
   Example:

Textual elements in a thumbnail gives the opportunity to give more information about a movie. Adding textual elements is highly recommended by a lot of the thumbnail tutorials. Also adding a brand logo could give more information to the user.

Discussion

The influence of textual elements and a brand logo in a thumbnail could be investigated with this classification scheme. A hypothesis on the influence of textual elements vs. without textual components in thumbnails can be: more information in a thumbnail will attract more views. Another hypothesis can be: giving more information, gives less room for the interpretation of the viewer, which has a negative effect on views.

As for the influence of logos its seems that this also can also have a positive influence on the number of views. To know who uploaded a video can give the viewer information about the content of the video. But, brand logos can also give a negative association which can negatively influence on the number of views.

The ResNet50 model is trained to classify 1000 different objects in images. Looking at the pretrained categories (https://gist.github.com/yrevar/942d3a0ac09ec9e5eb3a), text as object is not included. So this model doesn't look very suitable for this task. The Coco model from the syllabus doesn't fit our needs either. But, there is also a Coco-text model (https://vision.cornell.edu/se3/coco-text-2/). This one could probably used for classifying text in our thumbnails. To identify the logos in our thumbnails we could for example use the Google Object Detection API to download a logo dataset and train a model ourself (https://towardsdatascience.com/google-object-detection-api-to-detect-brand-logos-fd9e113725d8). A difficulty of this implementation is that a logo with text, easily will recognized as text by the coco-text model. This will of course bias our result.