Using Deep Learning Techniques to Classify Photographs

My task will be to build a model to assign photographs to one of a pre-determined number of categories. For example, the categories could be landscape, wildlife, people, etc. The potential application for this type of model could be a stock photography website. This website could take photographs and automatically place them into appropriate categories. This would remove the need for humans to take the time to classify the photos manually. By removing the human-labeling element, it would also help to standardize the categories that their photographs fall into. With enough photographs and enough categories, an extension of this application could be to auto-generate hashtags for social media sites such as Instagram and Twitter.

The photographs that I will use for the training data will come from my personal archive of photos. I currently have an archive of over 10,000 color and black and white photographs to choose from for training. The size of the dataset will depend on the ability of the model to predict the categories accurately. I will self-label the photos used for training and testing the model. If my personal photographs are not sufficient to build a robust enough training set, I can reach out to a handful of fellow photographers who have already indicated a willingness to provide photographs for training my model. An added benefit of obtaining photographs from other photographers is that they may have a different artistic style than me. One danger of using just my photographs is that the model may learn my personal style and not be able to correctly predict the work of other photographers.

To solve this problem, I will use transfer learning methods on an already established image classification neural network model. I intend to research a handful of different models to determine the one that is best suited to help with this task. I will then strip off some number of the hidden layers and then re-train the model on my set of photographs and labels. The type of model, i.e. fully connected, convolutional, recurrent, will depend on the existing model that I determine best fits my needs. If more than one model seems to be appropriate, I will explore the option of retraining more than one model to determine if one is more efficient or more successful than the other. I will also need to explore re-sizing and re-shaping photographs so that I can feed a photograph of any dimension into the model for prediction.

The deliverables for this project will be my code, a report detailing the model, and a slide deck presentation. Included in the report will be details on how I classified the photographs, what steps I did to reshape each image, which existing image classification models I explored and why I chose the model(s) that I did, as well as any unforeseen obstacles that I was forced to overcome.