Literature review:

Borji – Human vs. computer in scene and object recognition:

Comparing human and computer performance allows us to learn cases where humans fail and computers succeed and vice versa. This will enable us to build better algorithms or understand human vision better. Also, using human performance as a baseline is a good way to create computer vision algorithms that replicate the human visual system.

5 tests -

1 & 2 regard scene categorization w/ color photographs and line drawings

3 regards invariance properties of models on animal vs. non-animal recognition

4 local vs. global information -- jumbled scenes

5 object recognition

**Eberhardt – Deep learning and animal recognition:**

**This study looks at the depth of specific neural networks and analyzes which one’s are more accurate and which ones emulate human vision better. As the depth increases, the accuracy increases. But intermediate depths were more reflective of human performance on the task (animal vs. non-animal in 500 ms). It also depends how long the images are shown for. It seems deeper neural networks reflect deeper analysis in the human visual system of information (deeper processing in higher stages in the visual system?). Could also be top-down processing with more time so maybe recurrent neural networks are better?**

**Parikh – representing jumbled images**

**Low resolution images represent only global information, jumbled images represent local information. Contains a good intro in to what has been done on studying role of local and global information in distorted images. Study: block an image and each image is associated with a scene and then the total labeling of an image is voted on using the bag of words model. Bag of words model performed around the same for humans and computers. Only one block size was used.**

**Vogel – Categorization of Natural Scenes**

**Has a discussion of ways to select images that are representative. A couple of experiments:**

**Experiment 1: obtaining ground truth data on whole images**

**Experiment 2: scrambled images only local information**

**Experment 3: blurred images – only global information**

**Experiment 4: Gray-scaled images – no color present scrambled and blurred**

**Also, compares 2 computer vision techniques to human performance to see role of local and global information**

**Yokosawa – Does Disruption of a Scene impair change detection**

**Talks about how scrambling images and blacking some squares out affects how we can detect change from one image to the next. Although not totally pertinent, I thought it was good in talking about the distortion techniques and why they were chosen.**

**Renninger – When is Scene Identification just Texture Recognition?**

**Talks about how texture is important in rapid scene identification. Is there a way to distort the texture of an image is an important question I should think about.**

**Zhou – Learning Deep Features for Scene Recognition**

**Discussion of the neural network we’ve been using and a way to compare image data sets. Comparison of image data sets not too important, but there is a discussion of the receptive fields of the neural network and what they could possibly be looking for.**

**Greene – Visual Scenes are categorized by function**

**Argues that a scene’s identification category may be determined by its function, not the objects and features that make it up, or the actions one can perform in the scene. Ask people to identify if two scenes are in the same category. Function-based models performed better. I think this would be an interesting research question to test. By using distorted images, maybe we could get a better answer to this question.**

**Oliva – Building the gist of a scene: the role of global image features in scene recognition**

**Michelle Greene – Recognition of Natural Scenes from global properties: seeing the trees without representing the forest:**