### Motivation
There has been a significant amount of research done on the use of computer vision for autonomous on-road driving, but significantly less work focused on the off-road setting. In this domain, accurate terrain classification is crucial for safe and efficient operation. In this project, we attempt to build a CNN-based terrain classification system capable of processing images of both homogeneous and heterogeneous terrain.

### Method
1. Classify homogeneous image patches as one of 6 terrains / textures: bark, dirt, dry vegetation, green vegetation, grass, and pavement.
2. Process mixed terrain images by using the homogeneous terrain classifier and a sliding window algorithm to map each pixel to a terrain label.

### Data Collection & Processing
1. Took videos of the 6 homogeneous textures and images of mixed-terrain landscapes at Lake Lagunita & the Stanford Dish.
2. 12,000 frames were extracted from the videos and then down-sampled to 100x100 pixels.
3. Training set was augmented with transformations (e.g. rotations, contrast), yielding 72,150 training samples.
4. Produced 2,408 validation samples and 2,408 test samples.

### Baseline SVM Model & Results
- Bag of Visual Words (BoVW) model with 64-element Speeded Up Robust Features (SURF) descriptors.
- Mini-batch k-means to cluster SURF descriptors, generating a visual vocabulary of 125 words.
- Hessian threshold of 300.
- Radial basis function kernel, regularization parameter of 1.0.
- Implemented using scikit-learn.

### CNN Model & Results
- Learns filters preserving spatial relationship between pixels.
- Convolution layer: 32 feature maps, 5x5 filters.
- Max pooling layer: subsamples each 2x2 window.
- Dropout layer: reduces overfitting by excluding 20% of neurons.
- Flatten layer: vectorizes into a fully-connected layer with 128 neurons.
- Cross-entropy loss, Adam optimization, over 10 epochs.
- Implemented using Keras & Theano backend.

### Mixed-Terrain Results & Discussion
- The CNN's superior performance on the homogenous classification task appears to translate to the mixed-terrain processing task, generating smoother and more visually intuitive labeled images.
- Both classifiers have difficulty distinguishing between dry vegetation and dirt; the SVM additionally struggles to identify grass, likely due to the color insensitivity of SURF features.
- Future work includes training the homogeneous classifier on higher-resolution textures to enhance the distinctness of each terrain type, and more rigorous analysis of mixed-terrain categorization performance.

### References