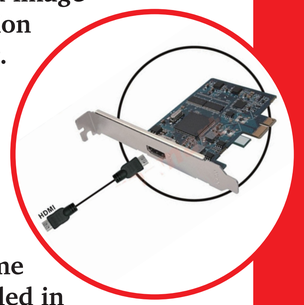
Automated Smart TV UI Performance Testing with Visual Recognition

Peter Li, Shim Young Lee

Motivation

Smart TV applications such as YouTube and Netflix have performance requirements. YouTube certification program specifies requirements regarding UI performance for system integrators and smart TV vendors that wish to ship their devices with YouTube application. Among those requirements is page transition performance, which refers to the time the transition between different types of pages takes.

In this project, we applied the current convolutional neural network based image classification algorithm to automation of Smart TV UI performance testing.



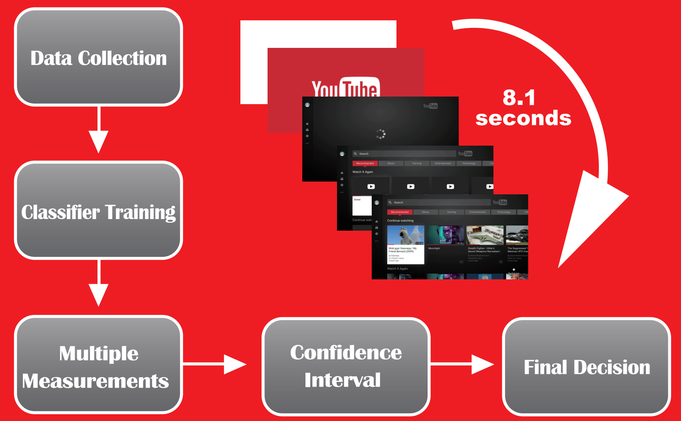
Data Acquisition

For image retrieval and collection, we used Decklink image capture card from BlackMagic. The captured images are 720P, with frame rate of 60Hz. The images were labeled in a semi-automated fashion: the images were first automatically assigned based on our prior knowledge about the duration of display for each page, and then each folder was manually checked and corrected.

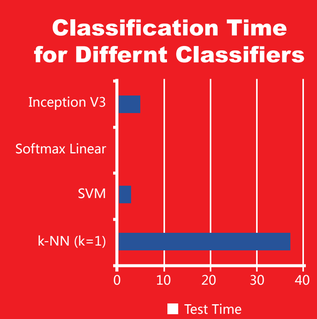
Classification

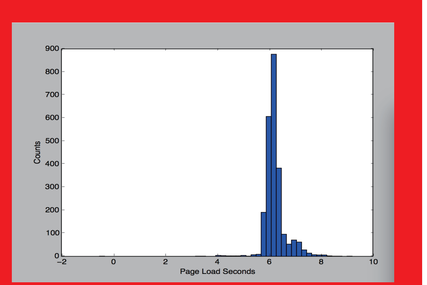
Several different convolutional neural network (CNN) architectures as well as other machine learning algorithms were tested, such as Inception V3, 1-convolution network, 1-hidden layer multi-layer perceptron, linear classfier (with softmax), SVM and k-NN.

Pipeline



Test Result





Distribution of the Launch time

- All CNN's, from Inception V3 to the simple 1-convolution network, achieved 100% validation accuracy.
- Linear classfier with softmax was the fastest classfier.
- Non-CNN classifiers could perform as well as the CNN's because of the little in-class variance of the images.

Measurement

$$\hat{f}(x) = 1\left\{\bar{x} + t^* \frac{s}{\sqrt{n}} < \gamma\right\}$$

$$s = \sqrt{\frac{\sum(x - \bar{x})^2}{n-1}}$$

For this specific dataset (YouTube), we have many classifiers that achieved 100% validation accuracy, so the page transition can be reliably measured based on the classification result at each time step.

The final statistic is given by the confidence interval (90%) based on the mean and the standard deviation calculated from the multiple measurements of the launch time. The decision rule for the performance test is whether the upper bound of the confidence interval is less than the required launch time specified by the smart TV applications.

Conclusion & Future work

We successfully applied convolutional neural network to the automation of Smart TV UI performance testing and was able to achieve a highly reliable performance with a generalizable pipeline.

Similar pipeline and training procedure can be applied not only to similar UI performance testing of different applications, but also to all UI related test automations that require similar computer vision capabilities. For example, similar system can be used to iOS and Android UI related test automation.