Classification of Book Genres

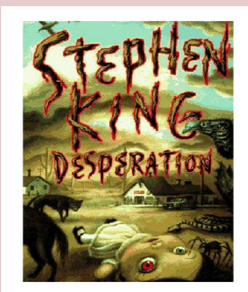
Holly Chiang, Yifan Ge, and Connie Wu
CS 229 Fall 2015
Stanford University

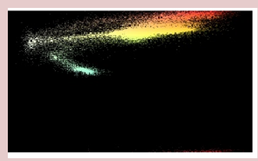
Background

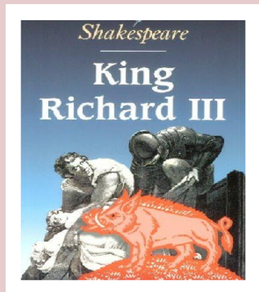
There is a saying that you shouldn't judge a book by its cover, but for most people the cover forms their first impression of a book. Identifying the genre of a book is not a trivial problem, since books within a genre can span a spectrum of designs, but with machine learning we aim to find some patterns between the genres. We attempt to classify book covers into one of five genres: business, fantasy, history, science-fiction, or romance. The **oracle** is human selection of genre (which is surprisingly not that perfect). The **baseline** is where the genre is randomly chosen.

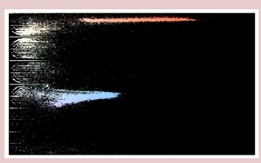
Color Histogram

In one of our earlier attempts we tried using image color information to predict the genre. This includes the color histogram distribution and image complexity estimation. In the histogram distribution, we use HSV color space to classify the pixels to 16 bins and record 1 if the number of pixels in the bin is more than 2% of the total pixels (two examples show on the right). Image complexity is estimated through spatial information of an image, which is calculated by Sobel filtered results.



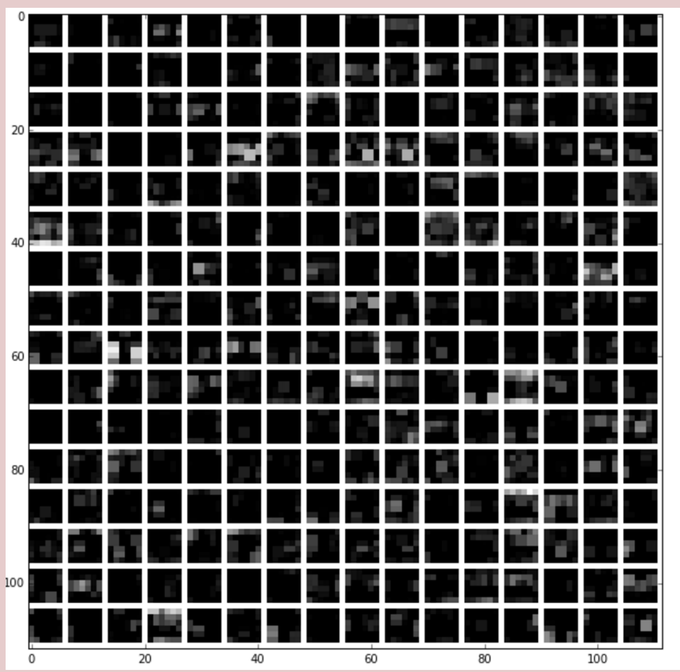






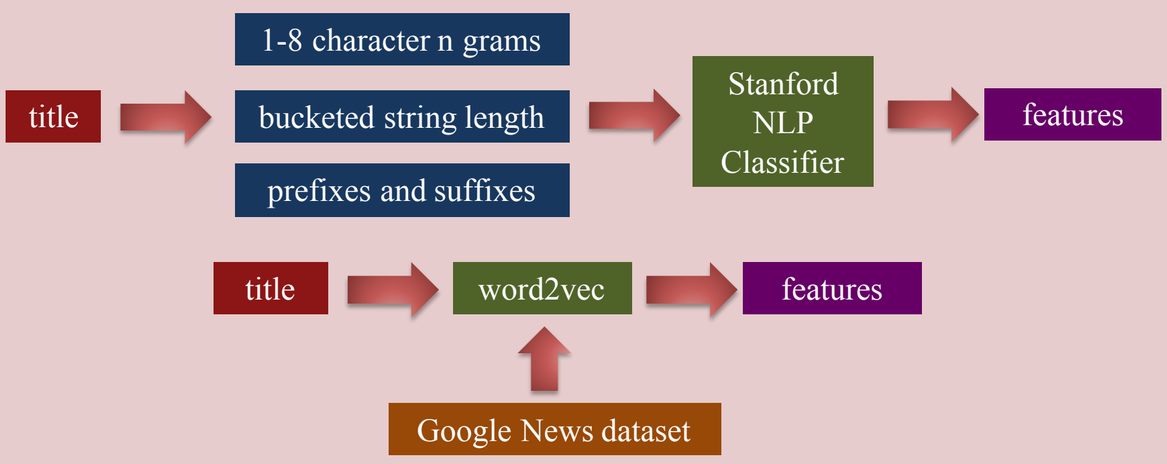
Transfer Learning

We extracted features from a fully-connected 'fc7' layer of a convolutional neural network pre-trained on ImageNet from Krizhevsky et al., which contains 1.2 million images with 1000 categories. Then we matched those features with our genre labels.



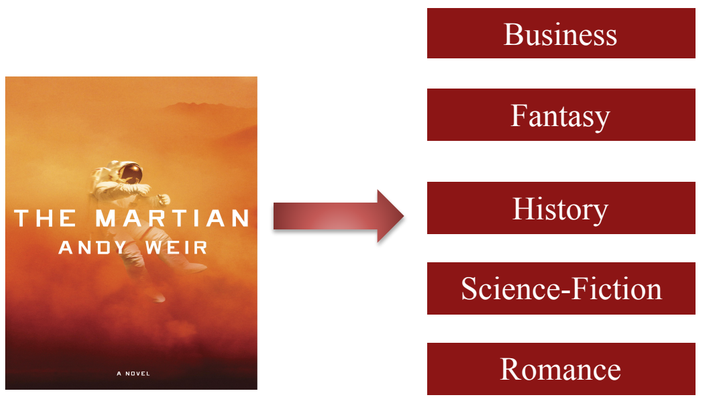
Natural Language Processing

We extracted features from book covers using two different NLP classifiers. The first classifier is the Stanford NLP Classifier, which is a probabilistic softmax classifier. The second classifier is word2vec, which produces word vectors from a text corpus using continuous bag of words.



Data Set

The original book cover images are from the OpenLibrary API. We used image processing techniques to remove empty or plain text book covers for a higher quality dataset.



Classification with Multi-Class SVM

We fed in the features extracted from all three approaches into a multi-class SVM with the following structure : n samples with n_classes * (n_classes - 1) / 2) with radial basis function kernel

$$\min_{\alpha} \frac{1}{2}\alpha^\top y_i y_j K(x_i, x_j)\alpha + e^\top \alpha$$

$$\text{subject to } y^\top \alpha = 0$$

$$0 \le \alpha_i \le C, i = 1, ..., n$$

$$K(x_i, x_j) = \exp(-\gamma |x_i - x_j|^2)$$

Evaluation Metrics for Success

To define the success of the genre prediction, we evaluate the precision and recall of each of our approaches. We used a training set of 5000 images and a test set of 1000 images. The color histogram approach had the worst performance and had no impact on the combined precision and recall metrics, and so is not shown below.

TRANSFER LEARNING					
Genres	TP	FN	FP	Precision	Recall
Business	157	43	63	71%	79%
Fantasy	111	89	104	52%	56%
Sci-Fi	92	108	89	51%	46%
History	121	78	61	66%	61%
Romance	131	69	70	65%	66%

NLP					
Genres	TP	FN	FP	Precision	Recall
Business	152	45	37	80%	77%
Fantasy	88	108	107	45%	45%
Sci-Fi	87	109	116	43%	44%
History	117	79	67	64%	60%
Romance	125	72	86	59%	63%

TRANSFER LEARNING AND NLP					
Genres	TP	FN	FP	Precision	Recall
Business	159	38	60	73%	81%
Fantasy	109	87	97	53%	56%
Sci-Fi	93	103	84	53%	47%
History	120	76	59	67%	61%
Romance	132	65	69	66%	67%

Results

The image-based transfer learning and title-based NLP approaches had similar precision, which suggests that the two metrics are equally representative of a book's genre. This is surprising considering that two different modalities resulted in similar categorization on the test set. Some challenges include distinguishing between science-fiction and fantasy genres, since there are a significant number of books that belong to both. Consequently, that makes it hard for even a human (oracle) to tell the difference. Overall we do much better than the baseline and show that it is possible to train to predict genre from cover and title.