**Visual Listening In: Extracting Brand Image Portrayed on Social Media**

**Data and Code**

All data, codes, and model configuration files are located in the three folders: data/, src/, and model/. This README file provides detailed information of the files and how they were used to produce results reported in the “Visual Listening In: Extracting Brand Image Portrayed on Social Media” (Liu Liu, Daria Dzyabura, and Natalie Mizik, Marketing Science, 2020).

**Brief Overview of the Study**

1. Train BrandImageNet with Flickr Images (Section 3 in the paper)
   1. Download Flickr images
   2. Prepare image data for model training
   3. Set up and train BrandImageNet
   4. Evaluate BrandImageNet model performance on Flick data (Table 1 in the paper)
2. Apply BrandImageNet model to Instagram images (Section 4.1, 4.2 in the paper)
   1. Download Instagram images
   2. Predict brand attributes from Instagram images using BrandImageNet model
   3. Compute image-based brand image (IBBI) metric
   4. Misc
3. Assess BrandImageNet performance on Instagram images against human judges and online survey (Section 4.3 in the paper)
   1. Collect human labeled image data on Figure Eight
   2. Evaluate model performance on the labeled data set (Table 2, 3, 4 in the paper)
   3. Conduct brand perceptions survey with Instagram users
   4. Conduct correlation analysis of model-predicted IBBI of Instagram image data and survey-based measures of brand perceptions (Table 5, 6 in the paper)
4. Conduct a case study of underwear brands (Section 4.4 in the paper)
   1. Collect human labeled image data on Figure Eight
   2. Test the differences between IBBI scores across consumer and firm photos and across brands (Table 7 in the paper)

**Detailed Overview of Data, Programming Code, and Analyses in the Study**

1. Train BrandImageNet with Flickr Images (Section 3 in the paper)
   1. Download Flickr images
      1. Code for downloading images from Flickr: data/code\_to\_get\_flickr\_images/DownloadImage.java and data/code\_to\_get\_flickr\_images/download\_flickr.sh
   2. Prepare image data for model training
      1. Split data into training set, validation set, and test set

* src/Split Flickr Data to Train Val Test Data.ipynb
  + 1. Code and script for converting jpg image data to lmdb format for faster model training
* First convert image content to lmdb format with src/run\_convert\_imageset.sh and src/run\_convert\_imageset.s.
* Then generate multi-label for each image with create\_multilabel\_lmdb\_justlabel.py and src/run\_create\_multilabel\_lmdb\_justlabel.sh
  1. Set up and train BrandImageNet
     1. Model configuration files and script
* Model configuration files
  + model/deploy.prototxt
  + model/solver.prototxt
  + model/train\_val.prototxt
* Run src/train\_multilabel.s to train model
  + 1. Input to train BrandImageNet model
* Pre-trained model (bvlc\_reference\_caffenet) files: data/bvlc\_reference\_caffenet/\*
  + 1. Select the model iteration that has the best performance on the validation set
* Training log with model loss outputted during model training: data/brandimagenet/training\_log/\*
* Code for finding the best model snapshot: src/Choose BrandImageNet Model Best Snapshot and Report Performance on Test Data.ipynb
* Selected best model snapshot: data/brandimagenet/model/ imagenet\_tunelr0001\_step5000\_multilabel\_10000\_iter\_2800.caffemodel
  1. Evaluate BrandImageNet model performance on Flickr data (Table 1 in the paper)
     1. Model prediction on and labels for hold-out data
* Code and script for generating prediction on test data: src/predict\_flickr\_images.py and src/run\_predict\_imagenet\_multilabel\_80train\_10test.s
* Test data with predictions and labels: data/brandimagenet/prediction\_on\_flickr\_test\_set/imagenet\_tunelr0001\_multilabel\_80train10test.2800.prediction
  + 1. Code for evaluating out-of-sample model performance
* src/Choose BrandImageNet Model Best Snapshot and Report Performance on Test Data.ipynb

1. Apply BrandImageNet model to Instagram Images (Section 4.1, 4.2 in the paper)
   1. Download Instagram images
      1. Code and scripts for downloading Instagram images. **Note:** We used Instagram Public API to download all the data in 2016. Instagram has since closed their public API and provides only a limited access.

* Download images by hashtag: data/code\_to\_get\_instagram\_images/DownloadImage.java and data/code to\_get\_instagram\_images/download\_image.sh
* Download images by user account: data/code\_to\_get\_instagram\_images/DownloadImageByUser.java and data/code\_to\_get\_instagram\_images/download\_instagram\_by\_user.sh
  1. Predict brand attributes from Instagram images using BrandImageNet model
     1. Code and scripts for predicting brand attributes from Instagram images
* src/predict\_brand\_images.py
* src/run\_predict\_brand\_images\_user\_apparel.s
* src/run\_predict\_brand\_images\_user\_beverage.s
* src/run\_predict\_brand\_images\_official\_apparel.s
* src/run\_predict\_brand\_images\_official\_beverage.s
  + 1. Code and script for converting multi-label prediction to binary prediction per brand attribute (used later in the analysis)
* src/convert\_multilabel\_to\_binary\_prediction.py
* src/run\_convert\_multilabel\_to\_binary\_prediction.sh
  1. Compute image-based brand image (IBBI) metric
     1. Code and scripts for generating IBBI metric per brand
* src/compute\_ibbi.py
* src/run\_ibbi\_consumer.sh
* src/run\_ibbi\_firm.sh
  + 1. Generated model-predicted IBBI from consumer- and firm-created images for apparel and beverage categories
* data/brand\_perception\_metrics/apparel\_consumer.csv (Online appendix Table OA1 in the paper)
* data/brand\_perception\_metrics/apparel\_firm.csv (Online appendix Table OA3 in the paper)
* data/brand\_perception\_metrics/beverage\_consumer.csv (Online appendix Table OA2 in the paper)
* data/brand\_perception\_metrics/beverage\_firm.csv (Online appendix Table OA4 in the paper)
  1. Misc
     1. Brand name files needed for the above analysis. These files contain the hashtags and official account names of the brands used in the study
* data/brand\_hashtag\_and\_account/apparel\_brands.txt
* data/brand\_hashtag\_and\_account/beverage\_brands.txt
* data/brand\_hashtag\_and\_account/official\_apparel\_brands.txt
* data/brand\_hashtag\_and\_account/official\_beverage\_brands.txt
* data/brand\_hashtag\_and\_account/instagram\_hashtag\_to\_brand\_name.txt
* data/brand\_hashtag\_and\_account/official\_account\_to\_brand\_name.txt
* data/brand\_hashtag\_and\_account/apparel\_full\_brand\_names.txt
* data/brand\_hashtag\_and\_account/beverage\_full\_brand\_names.txt

1. Assess BrandImageNet performance on Instagram images against human judges and online survey (Section 4.3 in the paper)
   1. Collect human labeled image data on Figure Eight
      1. Create image data set for labeling

* src/create\_image\_sample\_for\_exp\_50\_per\_group.py
* src/run\_create\_image\_sample\_for\_exp\_50\_per\_group.sh
  + 1. Human-labeled image data from Figure Eight (including both images and human judgements)
* data/instagram\_validation\_data/glamorous/\*
* data/instagram\_validation\_data/rugged/\*
* data/instagram\_validation\_data/healthy/\*
* data/instagram\_validation\_data/fun/\*
  1. Evaluate model performance on the labeled data set (Table 2, 3, 4 in the paper)
     1. Code for evaluating model performance
* src/Instagram image labeling experiment result analysis.ipynb
  1. Conduct brand perceptions survey with Instagram users
     1. Raw survey data
* data/brand\_perception\_survey/Brand+Perception+Survey+-+Apparel\_July+23,+2019\_12.16.csv
* data/brand\_perception\_survey/Brand+Perception+Survey+-+Beverage\_July+23,+2019\_12.15.csv
  + 1. Compute survey-based brand perception score based on raw survey data
* src/Convert raw brand perception survey to perception score.ipynb
  + 1. Generated survey-based brand perception score
* data/brand\_perception\_metrics/apparel\_survey\_instagram.csv
* data/brand\_perception\_metrics/beverage\_survey\_instagram.csv
  1. Conduct correlation analysis of model-predicted IBBI of Instagram image data and survey-based measures of brand perceptions (Table 5, 6 in the paper)
     1. Code for computing correlations between all types of brand perception scores
* src/Correlation analysis with all brand perception scores.ipynb

1. Conduct a case study of underwear brands (Section 4.4 in the paper)
   1. Collect human labeled image data on Figure Eight

* data/underwear\_casestudy\_data/a1370681.csv
* data/underwear\_casestudy\_data/f1370681.csv
* data/underwear\_casestudy\_data/joeboxer-firm-images-filter.csv (This file contains text-only images. These images were included in the labeling task and the raw data files and were filtered out later in the analysis in 4.2 section below).
  1. Test the differences in IBBI scores across consumer and firm photos and across brands (Table 7 in the paper)
* src/Hanes-Joeboxer-Victoria case study analysis.ipynb