Deep Learning from Corporate Presentations: Extracting Forward-Looking Visual Information in Capital Markets

Sean Cao, Yichen Cheng, Yusen Xia, Baozhong Yang*

Abstract

Unstructured information has become important for investors to understand capital markets in the rising era of high-performance computing. While numerous studies have attempted to extract value-relevant textual information from corporate news, valuable information contained in disclosed corporate pictures has been untapped due to the highly unstructured and high-dimensional nature of business images. This paper employs deep learning models to automatically extract forward-looking information from a novel image dataset of corporate presentation slides. We plan to compare several deep learning models and provide insights on the model construction techniques specific to analyzing business images. We validate the forward-looking visual information using market reaction and analysts' revisions. Overall, our study develops a deep learning framework/toolbox specifically designed for extracting value-relevant information from business images, highlights the potential value of identifying visual information on a large scale automatically, and calls for future studies in image analytics in capital markets.

* All authors are affiliated with the J. Mack Robinson College of Business at Georgia State University. Sean Cao is an Assistant Professor of Accounting (Email: scao@gsu.edu). Yichen Cheng is an Assistant Professor of Analytics (Email: ycheng11@gsu.edu). Yusen Xia is the Bradford and Patricia Ferrer Professor in Analytics (Email: ysxia@gsu.edu). Baozhong Yang is an Associate Professor of Finance (Email: bzyang@gsu.edu).

Project Description

Numerous studies have examined how to extract value-relevant information from textual analysis, e.g., sentiment analysis (Tetlock, 2007; Loughran and McDonald, 2011; Jegadeesh and Wu, 2013, Jiang et al. 2019) and competition information (Hanley and Hoberg, 2010; Hoberg and Phillips, 2016). However, one important information source for investors is missing in the literature: visual information disclosed by corporations. This study aims to extract value-relevant information from graphs and images in corporate disclosure, in particular, corporate presentations to investors and analysts.

In textual analysis, texts can be analyzed as unigrams, bigrams, or classified into topics. In contrast, images contain more unstructured and higher-dimensional features such as pixels, contrasts, colors, and objects that are difficult to analyze. In this study, we apply deep learning techniques to extract forward-looking features in corporate presentation images and classify them into five main categories: 1) performance summaries – pictures that summarize historical performance metrics; 2) performance prospects – pictures that contain projections of future performance metrics; 3) operations/product snapshots – pictures that describe operational activities or products; 4) operations/product prospects – pictures that describe future products, blueprints of construction, expansion maps, and construction sites of new facilities; and 5) slides with mostly text information. We provide examples of images in each category in the Appendix. In order to successfully classify presentation images, we also extract information about the color composition, contrasts, and other graphical features of the presentation.

For our analysis, the most interesting categories are categories 2 and 4 which contain forward-looking information about firms' performance and operations. It has been important for investors to identify forward-looking information in capital markets. Since the 1980's, investors have relied on managers' forecasts of forward-looking numerical earnings to assist with their investment decisions. In recent years, practitioners and scholars have used text mining to extract forward-looking information that is qualitative in nature (e.g., Li, 2010). This study advances this line of research and explores the rich and visual forward-looking information in business presentation images.

Given that forward-looking information conveys new information to investors, we hypothesize that images about performance and operations prospects contain more value-relevant information for investors than images about performance summaries and operations snapshots.[1] In addition, there are policy implications for examining forward-looking information. In the United States, the Private Securities Litigation Reform Act of 1995 provide safe harbors for forward-looking statements that cannot sustain itself as merely a historical fact. This law and later court rulings allow and encourage companies to disclose more speculative information to investors without fear of major legal repercussions. Our study provides unique evidence on the market consequences of such safe harbor regulations in the context of forward-looking graphical information.

The contribution of our study is three-fold. First, we develop a framework for future studies of business images. We plan to publish our codes after the completion of our study to help future academic and practitioner work specific for business images analytics. Second, we are the first to study the market implications of forward-looking visual information and other graphical features. The graphical features allow us to *separately* study forward-looking information related to performance metrics and operations/products, which is less operationalizable for text analytics. Third, corporate presentations are important yet understudied activities in capital markets (Bushee et al. 2011). We collect a large dataset of business presentation slides and images. Our study suggests that extracting and analyzing the visual features from corporate presentations are fruitful and can help to advance future studies.

[1] Since the focus of this paper is image analysis, we remove text slides from our main analysis. We plan to control for textual information.

Data and Methodology

We collect around 50,000 corporate presentations and 1.9 million pictures (including text images) from year 2002 to year 2018. We randomly select around 20,000 pictures to label and use those labelled pictures as our training set. Traditional image classification uses deep learning models such as convolutional neural network (CNN) (see, e.g., Goodfellow et al. 2016). However, CNN is not especially designed for understanding business pictures and usually requires a large training set. We explore two advanced machine learning methods especially designed for business pictures. Both methods use CNN as the building blocks and apply advanced machine learning methods including transfer learning and dimension reduction for further improved predictive performance. First, we explore a transfer learning technique (Pratt 1993; Rajat et al. 2006) to incorporate public labelled picture databases[2] into our deep learning models to improve the prediction accuracy and reduce the number of training samples needed (transfer CNN). The transfer learning approach will allow us to utilize existing deep neural nets trained using huge amounts of labelled data, while also enabling us to fine-tune the network to adapt to our problem. Second, we develop a two-stage feature extraction framework to achieve dimension reduction, noise removal, and image classification at the same time. We first extract common features of business pictures in the first stage and use those features as input when training a deep learning model in the second stage. Such novel dimension reduction procedures as a two-stage CNN will be crucial to improving the predictive accuracy because it will enable us to identify features that are important in business images and to remove other noisy information carried in the high dimensional data. We also consider the possibility of using both image and text information (CNN+RNN). We plan to show that these innovations help us beat the traditional deep learning models such as plain CNN models. We will use the Python programming language to train the deep learning models and perform image classifications. The proposed Table 1 below compares the performance of different machine learning approaches.

Table 1: Comparison of out-of-sample performance for different deep learning methods.

	CNN	Transfer CNN	Two-Stage CNN	CNN+RNN
Accuracy	xx%	xx%	xx%	xx%
Precision	xx%	xx%	xx%	xx%
Recall	xx%	xx%	xx%	xx%
F1-score	xx%	xx%	xx%	xx%

We then examine the information content of each picture category by investigating the market reactions and analyst revisions following the presentations, in an event-study setting. We also plan to control for corporate characteristics and the textual information contained in presentation transcripts.

Computing Requirements

Given the size of the dataset, we estimate that we need roughly 30,000 core-hours to finish one round of deep learning classification. We also need to run our procedures for many rounds for tuning of our models. The Institute for Insight at Georgia State University is equipped with a cluster with four GPUs with 3840 cores, with 192 gigabytes storage to meet this computational need. We also plan to use XSEDE and other high-performance computing resources if needed.

[2] ImageNet (ImageNet Large Scale Visual Recognition Challenge, 2015) and Oxford VGG (Very Deep Convolutional Networks for Large-Scale Image Recognition, 2014).

Proposed Timeline

Stage 1. Collection of presentation data and manual classification of training sample. Complete.

Stage 2. Training and tuning various deep learning models. March 2019 to May 2019.

Stage 3. Completing empirical analysis. May 2019 to August 2019.

References

Bushee, B.J., M. J. Jung, and G. S. Miller, 2011, Conference Presentations and the Disclosure Milieu, *Journal of Accounting Research* 49 (5), 1163-1192.

Hanley, K., and G. Hoberg, 2010. The Information Content of IPO Prospectuses, *Review of Financial Studies* 23, 2821–2864.

Hoberg, G., and G. Phillips, 2016. Text-Based Network Industries and Endogenous Product Differentiation, *Journal of Political Economy* 124(5), 1423-1465.

Goodfellow, I., Y., Bengio, and A., Courville. 2016. *Deep Learning*, MIT Press.

Jegadeesh, N., and D. Wu, 2013. Word Power: A New Approach for Content Analysis, *Journal of Financial Economics* 110, 712–729.

Jiang, F., Lee, J., Martin, X. and Zhou, G., 2019. Manager sentiment and stock returns. *Journal of Financial Economics*, 132(1), 126-149.

Li, F. 2010. The Information Content of Forward-Looking Statements in Corporate Filings—A Naïve Bayesian Machine Learning Approach. *Journal of Accounting Research* 48 (5), 1049- 1102.

Loughran, T., and B. McDonald, 2011. When Is A Liability Not A Liability? Textual analysis, dictionaries, and 10-Ks, *Journal of Finance* 66, 35–65.

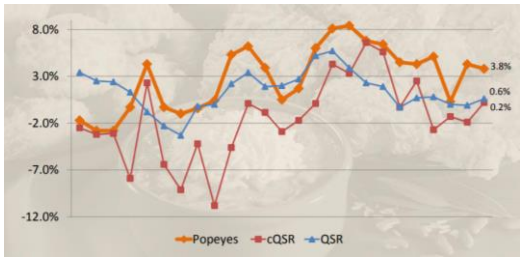
Pratt, L. Y., 1993. "Discriminability-based transfer between neural networks," *NIPS Conference: Advances in Neural Information Processing Systems* 5, Morgan Kaufmann Publishers, 204–211.

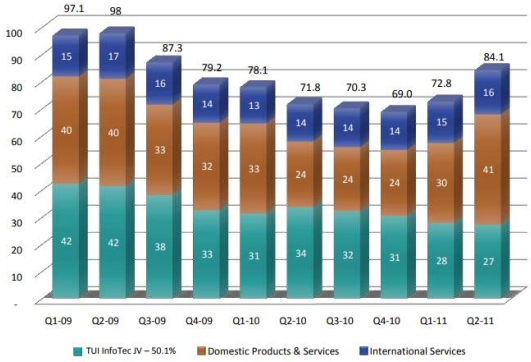
Rajat, R; A.Y., Ng, and D., Koller, 2006. "Constructing Informative Priors using Transfer Learning," *Twenty-third International Conference on Machine Learning.*

Tetlock, P., 2007. Giving Content to Investor Sentiment: The Role of Media in the Stock Market, *Journal of Finance* 62, 1139-1168.

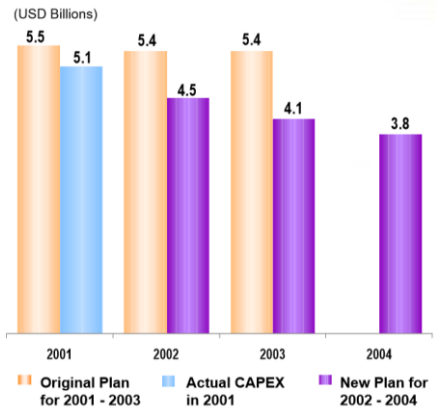
Appendix: Examples of Presentation Pictures in Each Category

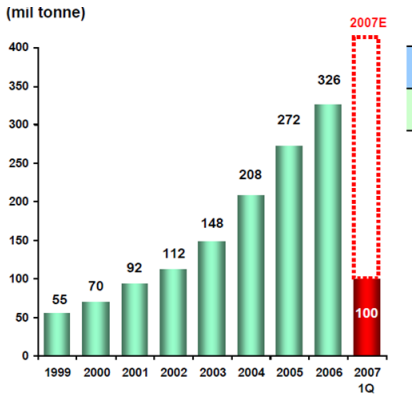
Category 1: Performance Summary – pictures that summarize historical performance metrics.

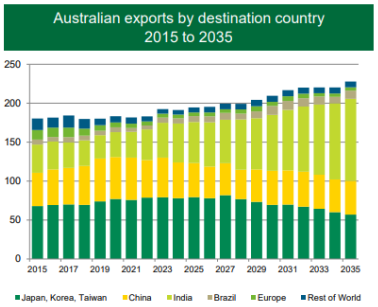


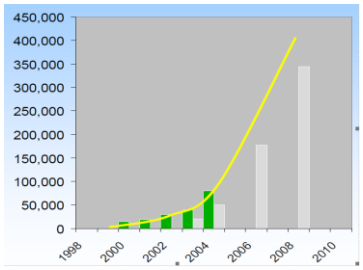


Category 2: Performance Prospects – pictures that contain projections of future performance metrics

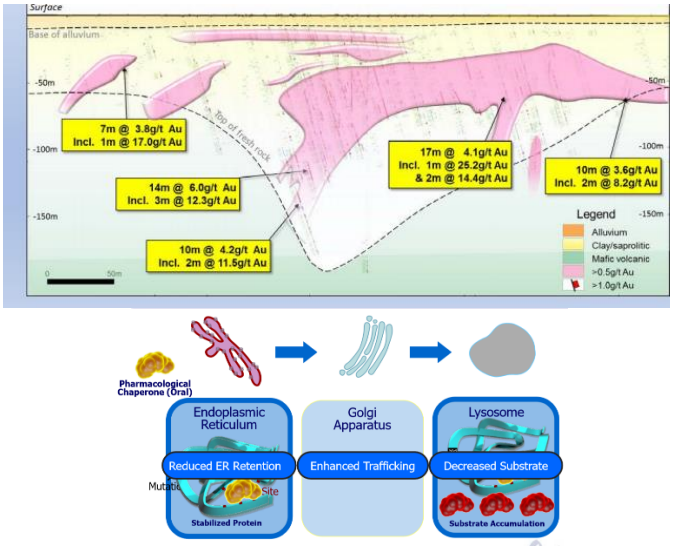


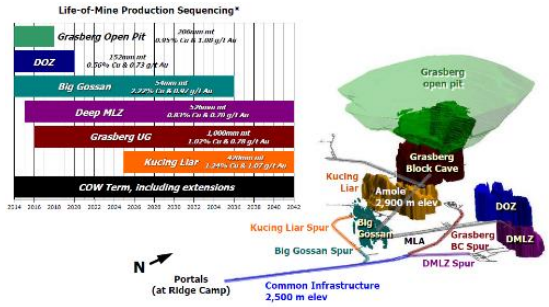






Category 3: Operations/Product Snapshots - pictures that describe operational or product activities

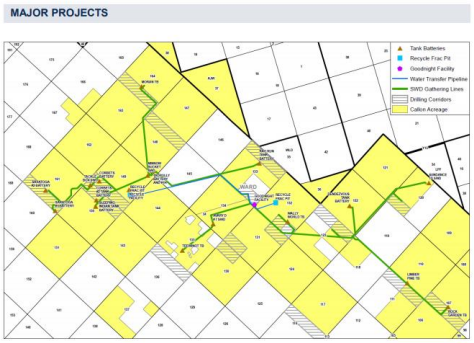


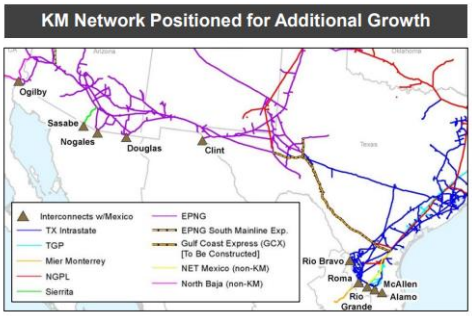




Category 4: Operations/Product Prospects – pictures that describe future products, blueprints of construction, expansion maps, and construction sites of new facilities.







Category 5: Slides with text information

Operational

Ready for next step in IT

- Succesful go-live next generation ERP
- Completion Office Automation in the Netherlands
- Strengthened leadership and organisation ICT
- Business process and ICT governance established
- Roadmap ICT aligned for coming 3 years across the organization
- ICT „Up to next level“ with One Neways (e.g., network, office automation)
- Further integration and improvement ICT organisations
- Streamlining architecture across Neways organisations