You are what you tweet…pic! Gender prediction based on semantic analysis of social media images

Michele Merler, Liangliang Cao and John R. Smith
IBM T.J. Watson Research Center, USA

mimerler@us.ibm.com
Outline

- Motivation & Research Questions

- Proposed Approach
  - Visual and Textual Analytics
  - Fusion Strategies
  - Experimental Results

- Conclusions & Future Directions
Social Media is a Goldmine of for Multimedia Research

Event Discovery and Summarization

Attributes
Discovery

Gender, age, location, education, political preferences, job, etc.

Training Data for Visual Classification

Sentiment Analysis
Previous Work

- **Text-based Gender estimation from social media**, *language dependent, performance ceiling*
  - Tweets [Burger et al. EMNLP11] [Pennacchiotti et al. ICWSM11]
  - First name [Liu and Ruths AAAI13]
  - Hashtags [Totems14]
  - Psycho-linguistic features [System U] [Kokkos et al. FM14]
  - Topic modeling on boards [Chang et al. CSCW14]

- **Non-text based Gender estimation from social media**
  - Collaborative Filtering (who you are friends with, who you follow) [Ito et al. ASONAM13] [Ludu CORR14]
    - *limited performance*
  - Profile Picture face analysis, *not always available/reliable*
  - Page Colors [Alowibdi et al. CASNAM13], *limited performance, not always available*
  - Whole Feed Images [Ma et al. IWCMASM14], *small set of ad hoc classifiers, no use of profile pictures, extremely limited generalization power*

- **Combinations Gender estimation from social media**
  - Text + Images [Sakaki et al. ICCL14]
    - *small preliminary study, over-simplicistic fusion method*: $\alpha p(\text{text}) + (1-\alpha)p(\text{visual})$, *limited performance*

- **Multimodal Fusion (not for gender prediction)**
  - Extensive literature, early fusion, late fusion, *general fusion strategies vs proposed specific filtered fusion* (see experimental results)

A need remains for a system that derives user gender using an effective multimodal combination of visual and non-visual cues
Research Questions

- Is there a **correlation** between **gender** and the content of the **images** that people post on social media?
  
  *yes*

- If so, can we **predict** a social media user's gender based on a semantic analysis of those images?
  
  *yes*

- Does the visual insight provide **complementary** information with respect to others (text)?
  
  *yes*
Summary of Invention – Extracted Information

Multimodal Cues

- **Textual**
  - Name
  - Description
  - Tweets (text)

- **Visual**
  - Profile picture
  - Header picture
  - Profile colors
  - Feed images/videos

![Multimodal Cues Diagram](image-url)
Summary of Invention

**User Profile**
- Feed Images/Videos
- Header picture
- Profile colors
- Profile picture

**Non-visual information**
- Feed Text (description+200 tweets)
- Name

**Visual Information**
- V1 Classifier
  - Based on generic visual classifiers
- V2 Classifier
  - Based on face gender analysis
- V3 Classifier
  - Based on low level descriptors

**T1 Classifier**
- Based on name/gender frequency analysis
- Based on standard text BOW, n-gram classifier

**T2 Classifier**

**Filtered Fusion**
- Fusion Classifier
  - Normalization
- First Level Gender Classifiers
  - V1 Classifier
  - V2 Classifier
  - V3 Classifier

**Second Level Gender Classifiers**
- Agree?
  - yes
  - no

**Confidence check**
- p1
- p2

**Gender Classification**
- Male
- Female
Textual Analytics

- Profile Name
- Text from tweets
Visual Analytics

- Profile Picture
- Color
- Analysis of Collection of Posted Images
What’s in a Twitter Profile Picture?
What’s in a Twitter Profile Picture?

the good (looking)

the bad

the cute

…and the weird?
Limitations of Profile Picture Face Analysis

- Misleading Clothing
- Occlusion
- Cloning
- Interesting angles
- Time travel
- Celebrity swap
- Multiple people
- Non-human pictures
Limitations of Profile Picture Face Analysis

- Race, glass, smiling
- Age, gender

Profile Picture : proposed approach

- Face++ detector

- Concept Detectors for 25 categories
  - Adult, Animal, Baby, Beach, Boy, Brand Logo, Building, CGI, Car, Cat, Child, Dog, Elderly Man, Elderly Person, Elderly Woman, Female Adult, Girl, Horse, Human Portrait View, Human, Icon, Male Adult, Motorcycle, Nature, Two People

  - Train SVM on top of Semantic Model Vector of concept detectors
Linear Profile SVM weights by category
Images From the Entire Feed

- Same Semantic Model Vector Approach
  - SMV 51 (subset of IMARS Taxonomy)
  - SMV 717 (subset of IMARS Taxonomy)
  - SMV Deep (from Caffe, 1K ImageNet categories)

- Aggregation Strategies
  - Model on Images directly
  - Simple Prediction Scores Aggregations (avg, max)
  - Statistical Count scores (threshold + count)
Twitter Gender Dataset Examples

Male
Twitter Gender Dataset Examples

Female
Linear SVM weights by category (SMV 51)
Top 14 weighted categories by gender (SMV51)

**Male**
- Stadium
- Bike
- Gym
- Zoo
- Parade
- American Football
- Airport
- Sport
- Campsite
- Market
- Monument
- Airplane
- Sports Activity
- Car

**Female**
- Baby
- Wedding
- Dog
- Sandwich/Food
- Celebration
- Computer Room
- Cat
- Store
- Bathroom
- Water Scene
- Tool
- Human
- Flower
- Greenery
Top 20 weighted categories by gender (Deep ImageNet)

**Male**
- black_stork
- plastic_bag
- trimaran
- gas_pump
- wire-haired_fox_terrier
- lifeboat
- ambulance
- Saint_Bernard
- castle
- assault_rifle

**Female**
- European_gallinule
- leaf_beetle
- loggerhead
- comic_book
- slide_rule
- whistle
- Dandie_Dinmont
- hand-held_computer
- American_black_bear
- bikini
- bolete
- canoe
- maillot
- otter
- aircraft_carrier
- obelisk
- nematode
- coucal
- hummingbird
- hard_disc
- dome
- king_penguin
- window_screen
- carousel
- purse
- amphibian
- tiger_beetle
- hen-of-the-woods
- bath_towel
Use Profile Page Color Info

Jalal S. Alowibdi, Ugo A. Buy and Philip Yu, Language Independent Gender Classification on Twitter, 2013 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining

- Background color
- Text color
- Link color
- Sidebar Fill color
- Sidebar Border color

512 Quantized colors (RGB with 3 bits each)
71% Accuracy when using all 5

http://www.twitteraccountsdetails.com/
Use Profile Page Color Info

512 bins

729 bins
Experimental Setup and Proposed Approach

- Public Annotated Dataset of 10K Twitter users¹

- 10 Training/Test random splits, each test split with 400 male and 400 female users

Proposed Approach: **Fusion** of 4 Multimodal Cues

- First Name (when available): associated with frequency in Male/Female populations²

- Text: standard bag of words from ~200 tweets, analyzed with LibShortText³ library

- Profile Picture: Semantic Model Vector (25 concepts) and gender inference from Face++

- *Stream Pictures*: Semantic Model Vector (51, 717, 1K) aggregated over user pictures

- Page colors
  - Background and Header picture

- Gender modeling was conducted using SVM with RBF kernel, with kernel parameters estimated via grid search

---

   From paper: **Liu and Ruths, What’s in a Name? Using First Names as Features for Gender Inference in Twitter, AAAI 2013**


Experimental Results

- Public Annotated Dataset of 10K Twitter users
- 10 Training/Test random splits, each test split with 400 male and 400 female users, rest used for training
- Gender modeling was conducted using SVM with RBF kernel, kernel parameters estimated via grid search

Image/video collection (visual feed) different pooling strategies

<table>
<thead>
<tr>
<th>Method</th>
<th>Accuracy</th>
</tr>
</thead>
<tbody>
<tr>
<td>Max Pooling</td>
<td>67.53</td>
</tr>
<tr>
<td>Avg Pooling</td>
<td>69.43</td>
</tr>
<tr>
<td>Avg Top-Quarter Pooling</td>
<td>69.56</td>
</tr>
<tr>
<td>Threshold-count</td>
<td>70.82</td>
</tr>
<tr>
<td>Avg Prediction Pooling</td>
<td>71.38</td>
</tr>
</tbody>
</table>

Individual Performance of Different Approaches

<table>
<thead>
<tr>
<th>Method</th>
<th>Accuracy</th>
</tr>
</thead>
<tbody>
<tr>
<td>Background SMV717</td>
<td>60.11</td>
</tr>
<tr>
<td>Header SMV717</td>
<td>64.41</td>
</tr>
<tr>
<td>color</td>
<td>66.18</td>
</tr>
<tr>
<td>Visual Feed SMV51</td>
<td>66.67</td>
</tr>
<tr>
<td>Visual Feed SMV717</td>
<td>71.38</td>
</tr>
<tr>
<td>Visual Feed SVMDeep1000</td>
<td>75.40</td>
</tr>
<tr>
<td>Profile SMV25</td>
<td>69.11</td>
</tr>
<tr>
<td>Profile Face++</td>
<td>74.90</td>
</tr>
<tr>
<td>First Name</td>
<td>71.22</td>
</tr>
<tr>
<td>First Name Frequency</td>
<td>69.58</td>
</tr>
<tr>
<td>LibText 200 Tweets</td>
<td>83.37</td>
</tr>
</tbody>
</table>

Fusion Strategies

<table>
<thead>
<tr>
<th>Method</th>
<th>Accuracy</th>
</tr>
</thead>
<tbody>
<tr>
<td>Visual Feed Early Fusion</td>
<td>75.58</td>
</tr>
<tr>
<td>Visual Feed Late (avg) Fusion</td>
<td>74.34</td>
</tr>
<tr>
<td>Visual Feed Late (SVM) Fusion</td>
<td>75.6</td>
</tr>
<tr>
<td>Profile Late (avg) Fusion</td>
<td>77.85</td>
</tr>
<tr>
<td>Profile Late (SVM) Fusion</td>
<td>78.63</td>
</tr>
<tr>
<td>Profile Filtered Fusion</td>
<td>79.05</td>
</tr>
<tr>
<td>All Visual Late(SVM) Fusion</td>
<td>80.08</td>
</tr>
<tr>
<td>All Visual Late(SVM) Filtered Fusion</td>
<td>83.36</td>
</tr>
<tr>
<td>Textual Early Fusion</td>
<td>84.08</td>
</tr>
<tr>
<td>Textual Feed Late (avg) Fusion</td>
<td>84.53</td>
</tr>
<tr>
<td>Textual Late (SVM) Fusion</td>
<td>84.67</td>
</tr>
<tr>
<td>Textual Filtered Fusion</td>
<td>85.72</td>
</tr>
<tr>
<td>Visual+Text Early Fusion</td>
<td>84.07</td>
</tr>
<tr>
<td>Visual+Text Late (SVM) Fusion</td>
<td>85.97</td>
</tr>
<tr>
<td>Visual+Text Late (SVM) Filtered Fusion</td>
<td>88.01</td>
</tr>
</tbody>
</table>

From paper Liu and Ruths, What’s in a Name? Using First Names as Features for Gender Inference in Twitter, AAAI 2013
Experimental Results

- Public Annotated Dataset of 10K Twitter users\(^1\)
- 10 Training/Test random splits, each test split with 400 male and 400 female users, rest used for training
- Gender modeling was conducted using SVM with RBF kernel, kernel parameters estimated via grid search

Gender Prediction Accuracy

<table>
<thead>
<tr>
<th>Method</th>
<th>Accuracy</th>
</tr>
</thead>
<tbody>
<tr>
<td>Visual Feed</td>
<td>75.6</td>
</tr>
<tr>
<td>Profile Picture</td>
<td>79.05</td>
</tr>
<tr>
<td>Visual Fusion</td>
<td>83.36</td>
</tr>
<tr>
<td>First Name</td>
<td>71.225</td>
</tr>
<tr>
<td>LibText 200 Tweets</td>
<td>83.375</td>
</tr>
<tr>
<td>Text Fusion</td>
<td>85.72</td>
</tr>
<tr>
<td>Visual + Text Fusion</td>
<td>88.0125</td>
</tr>
</tbody>
</table>

\(^1\)From paper *Liu and Ruths, What’s in a Name? Using First Names as Features for Gender Inference in Twitter, AAAI 2013*

*Different splits
Conclusions and Future Directions

Conclusions

- There is a correlation between the content of images posted on social media and the users’ gender, which can be exploited for gender prediction.
- Visual and textual information can be combined to boost gender prediction performance.
- Not all sources of information are equal. Filtered fusion provides best results.

Future Directions

- Improve/tailor visual classifiers.
- Cross media comparisons (Twitter vs Instagram vs Facebook…).
- Use framework to predict other attributes (age, hobbies, life events…).
Questions?

#UnaGeek
@UnaGeek

Ah, so that’s the age of Ultron.
References