Does Optical Character Recognition and Caption Generation Improve Emotion Detection in Microblog Posts?

June 22nd, 2017

Roman Klinger

Motivation

Outline

Motivation
○○

Emotion Analysis on Tweets
●○○

Methods & Experimental Setting
○○○○

Results
○○○○○

Conclusion & Discussion

What is Emotion Analysis?

Sentiment analysis

positive vs. negative
(neutral, mixed)

Subjectivity analysis

subjective vs. objective

Emotion analysis discrete (Ekman/Plutchik)

discrete emotion classes

Emotion analysis continuous (Posner/Russell/Peterson)

valence and arousal

Motivation
○○

Emotion Analysis on Tweets
○●○

Methods & Experimental Setting
○○○○

Results
○○○○○

Conclusion & Discussion

Emotion Models: Ekman
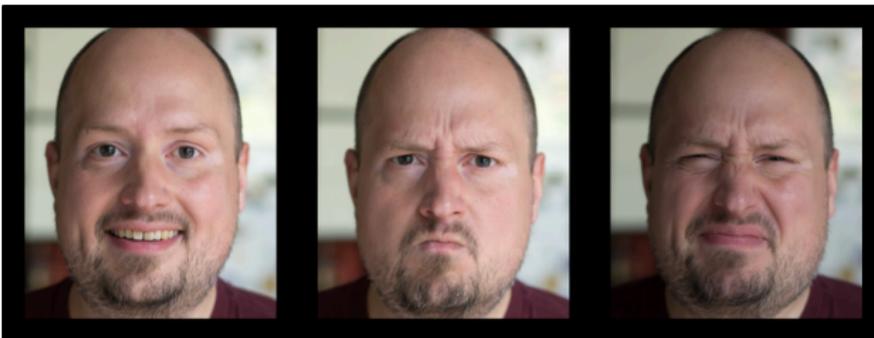
Joy   Anger   Disgust

Fear   Sadness   Surprise

Motivation
○○

Emotion Analysis on Tweets
○○●

Methods & Experimental Setting
○○○○

Results
○○○○○

Conclusion & Discussion
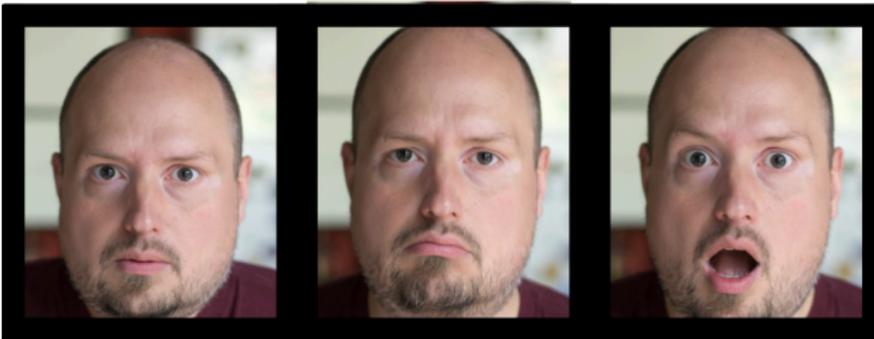
Task Description and Research Question

Task

Given a Tweet, assign an emotion from the set:
anger, *fear*, *joy*, *sadness*, *surprise*, *disgust*, *love*, *shame*, *trust*

Research Questions

- Which feature sets contribute to the task?
  - Tweet Text
  - Text extracted from attached image
  - Automatically generated caption for attached image
- Can we operationalize this with off-the-shelf tools?

Feature Extraction & Classifier

Tweet Text

- Bag of words, keep [^A-Za-z0-9#]
- Username blinding, emotion hashtags are ignored

Optical Character Recognition

- Tesseract 3.04.01
- Ignore output with less than 6 bytes

Caption Generation

- NeuralTalk 2
- Pretrained COCO data set model

- Classifier:
  linear maximum entropy classifier with L2 regularization

Motivation
○○

Emotion Analysis on Tweets
○○○

Methods & Experimental Setting
○●○○

Results
○○○○○

Conclusion & Discussion

Example for Caption Generation

Ivanka Trump ✔
@IvankaTrump

Happy Birthday to my incredible father
@realDonaldTrump. I hope that the year to come
is your best yet!
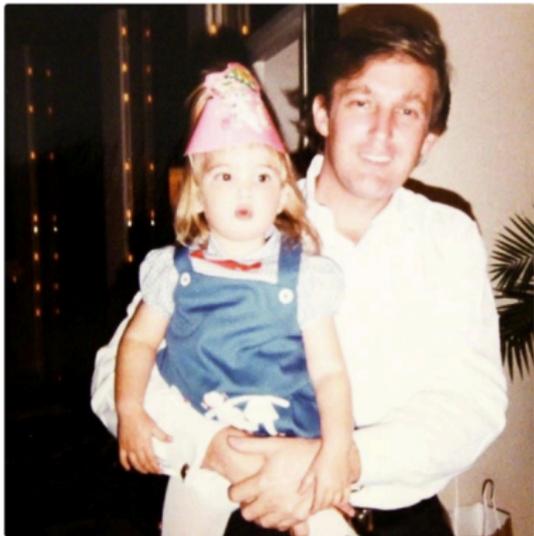
"a man and a woman
standing next to each other"

Example for Optical Character Recognition

Men Have Lupus
@menhavelupus
(Follow)

My wife says to let my #anger toward my
#Lupus #FuelMyFight. It works.
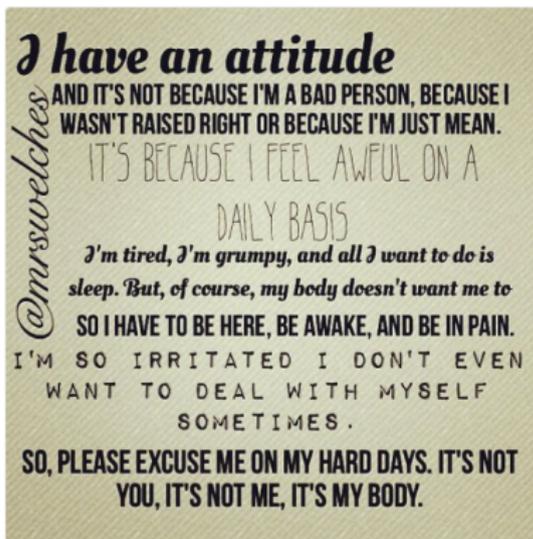
```
9 have an attitude
AND IT'S NDT BECAUSE I'M A BAD PERSON, BECAUSE I
WASN'T RAISED RIGHT CR BECAUSE I'M JUST MEAN.
OSOEENOOE 'EEEE OENO NN
J'm timd, J'm grumpy, and all 2 want tvdoiz
sleep. 3m, of course, my body doesn't want me to
SI] I HAVE Tl] BE HERE, BE AWAKE, AND BE IN PAIN.
I'M so IRRITATED I DON'T EVEN
WANT TO DEAL WITH MYSELF
SOMETIMES.
SD, PLEASE EXCUSE ME ON MY HARD DAYS. IT'S NOT
YDU, IT'S NOT ME, IT'S MY BODY.
```

"a picture of a man in a suit"

Corpus Generation

Data collection: self labeling

Crawl corpus from Twitter with hashtags for emotions

Sample three corpora of 200,000 Tweets

- D_{all} without constraints
- D_{OCR} from all instances with image and with OCR output
  - $\Rightarrow$ Tweets with images with text
- D_{Vis} from all instances with image and without OCR output
 - $\Rightarrow$ Tweets with images without text

  (captions are generated on all images)

Corpus Statistics

Emotion	D_{all}	w/ ϕ_{OCR}^{d}	w/ ϕ_{Vis}^{d}	D_{OCR}	D_{Vis}
joy	91,836	6,927 (8 %)	18,672 (20 %)	92,066	111,604
love	41,470	3,477 (8 %)	8,963 (22 %)	46,290	50,974
sadness	26,521	1,495 (6 %)	2,952 (11 %)	19,707	14,370
fear	12,721	1,490 (12 %)	2,299 (18 %)	19,400	7,925
anger	11,902	831 (7 %)	1,384 (12 %)	10,379	5,317
surprise	5,492	195 (4 %)	792 (14 %)	2,790	5,681
trust	5,170	561 (11 %)	886 (17 %)	7,274	3,151
shame	4,562	138 (3 %)	193 (4 %)	1,988	862
disgust	326	8 (2 %)	22 (7 %)	106	116
All	200,000	15,122 (8 %)	36,163 (18 %)	200,000	200,000

Experiment 1:
Tweet text features only on D_{all}, D_{OCR}, D_{Vis}
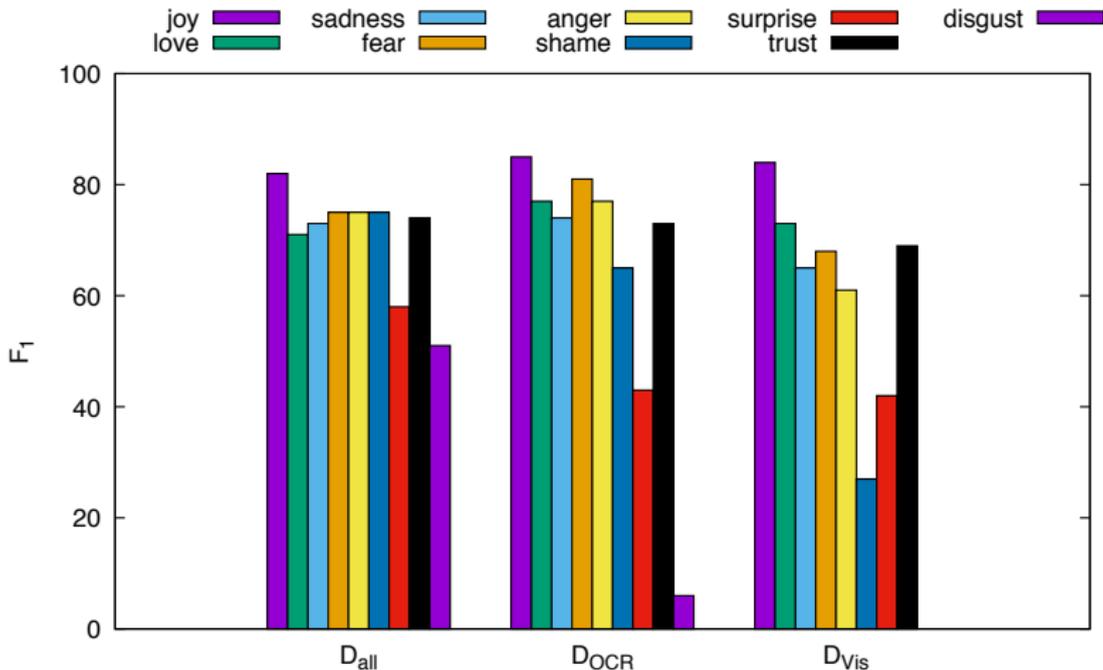
Macro-F: 71, 65, 54

Experiment 2:
OCR Features on D_{OCR}, Caption Features on D_{Vis}
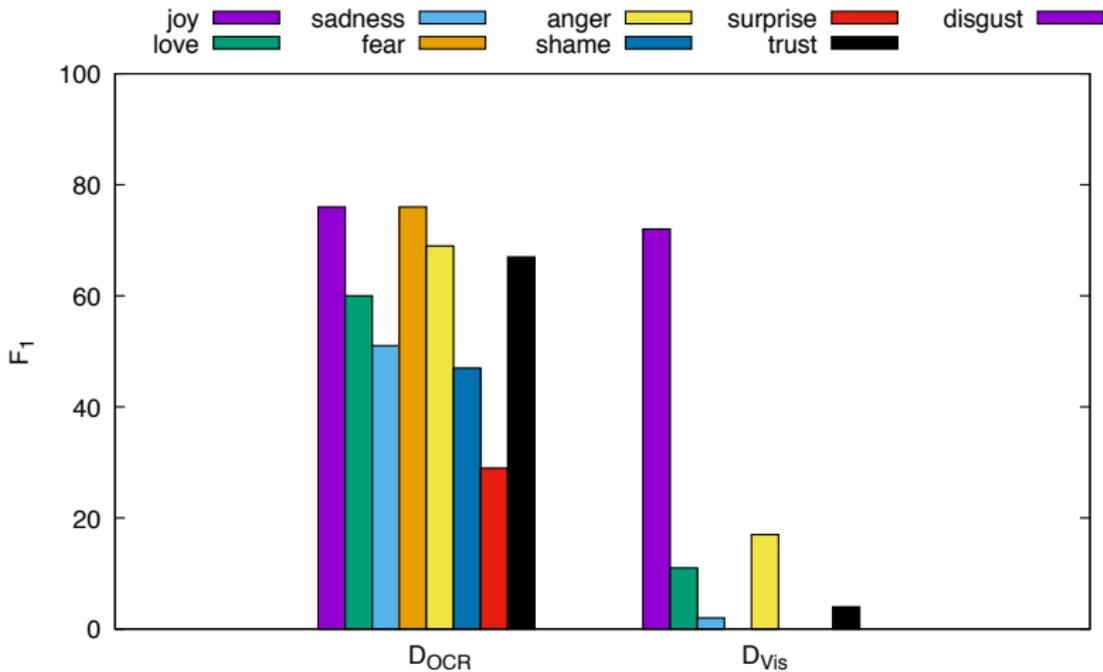
Macro-F: 53, 12

Feature Impact of OCR and Captions on $D_{\mathbf{OCR}}$

University of Stuttgart · Roman Klinger · June 22$^{\text{nd}}$, 2017 · 13 / 17
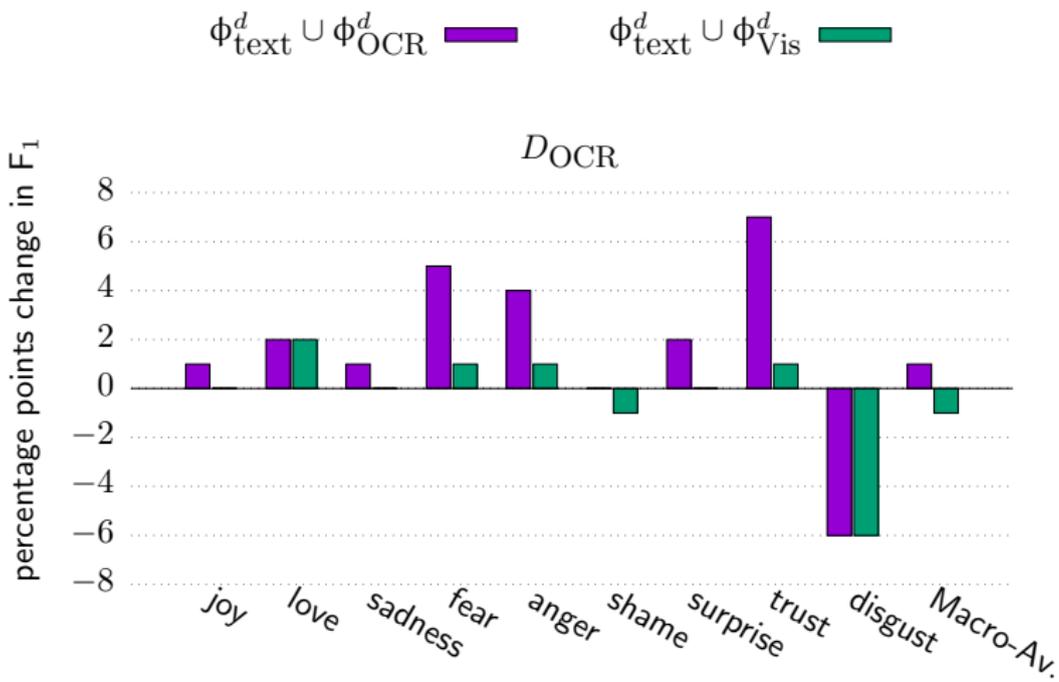
Feature Impact of OCR and Captions on $D_{\textbf{Vis}}$

$\phi_{\text{text}}^d \cup \phi_{\text{OCR}}^d$ ▬▬▬▬ $\phi_{\text{text}}^d \cup \phi_{\text{Vis}}^d$ ▬▬▬▬
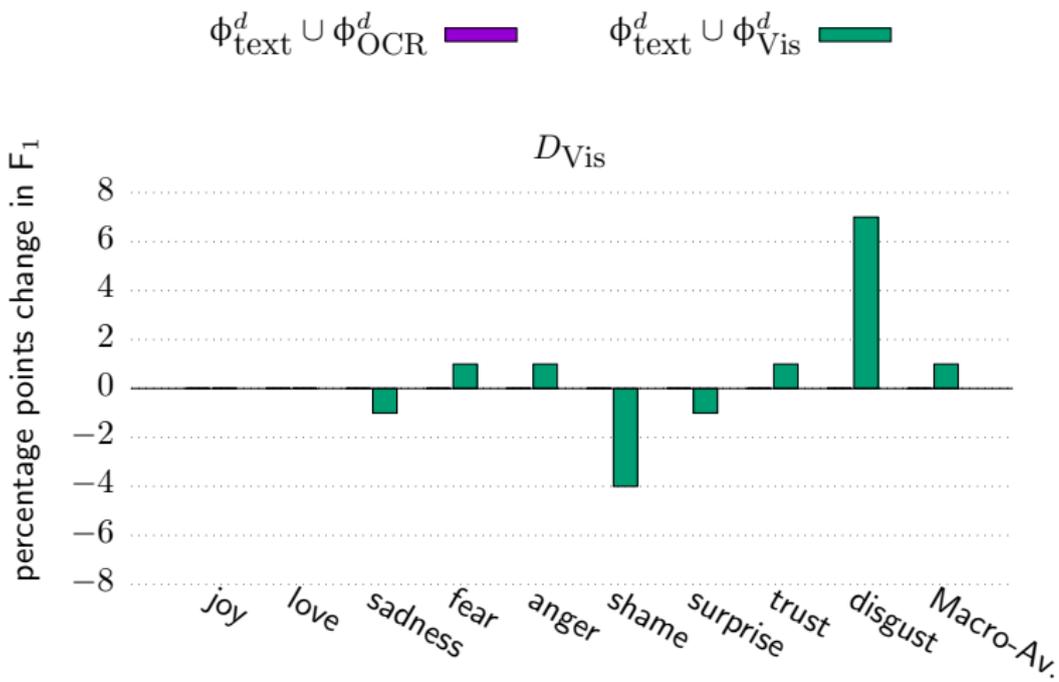
Conclusion & Discussion

- Optical character recognition improves emotion recognition

- Caption generation does not help

- Classification with images works surprisingly well based on Tweet text only
  (at least on frequent emotion classes)

Future Work

- Multiple classifiers for different object types
  (e.g. Caffe2 model zoo)
    - Facial emotion recognition
    - Age, gender
    - Landmarks
- Include deep learning emotion classifier taking the image as input directly
- Tweet generator instead of caption generator?