Visual Analytics in Marketing Research

Shunyuan Zhang

HARVARD | BUSINESS | SCHOOL

European Quant Marketing Workshop
April 2023

Agenda

Visual Data Analytics

- Why does visual data matter?
- How can visual data be useful?

Image Analytics: Frameworks and Examples

- To extract IV: Airbnb images, facial images, and ads. images etc.
- To construct DV or key metrics
- To develop a prediction model

Video Analytics: Working with Multi-modal Visual Data

- Steps & tools
- Two applications: YouTube influencers & in-store shopping video

Why visual data?

- Visual data is ubiquitous
  - Visual data make up more than **90%** of all consumer internet traffic (Cisco)
  - **86%** of businesses use video as a marketing tool (Wyzowl 2020)

- Owning customer (behavior) through mining visual data
  - As of 2022, an average person is predicted to spend **100 minutes per day** watching online videos (Zenith Media)

What can visual data do for us?

- Use visual data to **construct IV** that will be later used in your econometrics model
    - visual data for feature extraction

- Use visual data to **extract DV** that offers managerial implications
    - visual data to measure and/or to manipulate key metrics

- Use visual data to build a **prediction model**
    - visual data as additional (key) source of data

Image Analytics: extracting IV

- An econometrics model: *y~f(X)*
  - X=$\{X_1, X_2, \ldots, X_k, X_{visual}\}$

- Extracting X_{visual} from the visual data then ‘plug it in’ *X*
- Estimating the model *y* as you would normally do
- **The key is the step of feature extraction: X_{visual}**

General steps of extracting features

Is an off-the-shelf package (with good prediction accuracy) available?

Yes

Use existing API/packages

For common tasks:
Object detection/recognition, face recognition

Examples:
Microsoft Azure, Google Vision API

No

Is an existing training data available?

Yes

Using the existing data to train your ML model

Examples:
Predicating ethnicity of Airbnb hosts

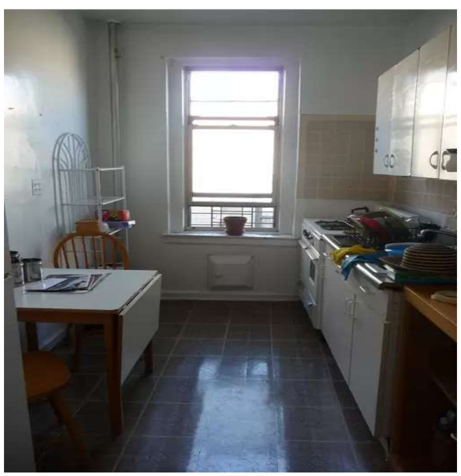
No

Constructing a training data

Using the existing data to train your ML model

Examples:
Predicting image quality

Example (1): image quality of Airbnb properties



Research question: what's the impact of high-quality images on the demand of Airbnb property?

Econometrics model: y~f(X)

- y: Airbnb property demand
- $\{X_1, X_2, \ldots, X_k,\}$: price, review, location, ….
- X_{visual}: image quality

Steps:

1. **Measuring image quality of property images**
2. Estimating a DiD model: exploiting within-variation via properties that have *changed* images

→ estimated coefficient of X_{visual}

Example (1): image quality of Airbnb properties

Creating a dataset to train an image quality classifier:

- Randomly selected 3,000 Airbnb property images and used Amazon Mechanical Turk to tag each image based on its quality → *training set*
- Build an image quality classifier using the training set (labeled images)
  - We applied VGG-16, a convolutional neural network that provided the state-of-the-art performance in image classification (Simonyan and Zisserman 2015)

Zhang et al. (2021) "What Makes a Good Image? Airbnb Demand Analytics Leveraging Interpretable Image Features." *Management Science*

Example of the AMT Aesthetic Quality Assessment Task

Instructions

We are interested in assessing the aesthetic quality of images. Give an aesthetic score to an image on a scale of 1-7 on how beautiful/pleasing/visually appealing you think the image is, where 1 is "very bad" and 7 is "excellent". Here is a short non-exhaustive guideline for judging images.

Reasons for a rating closer to 7:

- looks good
- visually pleasing
- well-lit
- clearly shows room/house features
- conveys well what the room/house may look like in person

Note 1: Most of our images are room/house images, however there are a few pitcures of outdoors (e.g., pictures of plants/gardens). The score should be given on the aesthetic quality of an image, regardless of its topic (i.e., whether it's a picture of indoor or a picture of outdoor)

Note 2: this is a subjective measure, but you can calibrate your standard by looking at our examples, to rate each image.

Examples of images with aesthetic scores of 1, 4, and 7:

	example score = 1	example score = 4	example score = 7
example 1			
example 2			

Give a score to an image on a scale of 1-7 on its aesthetic quality where 1 is "very bad" and 7 is "excellent".

1	2	3	4	5	6	7
very bad	bad	below average	average	good	very good	excellent

Example (1): image quality of Airbnb properties

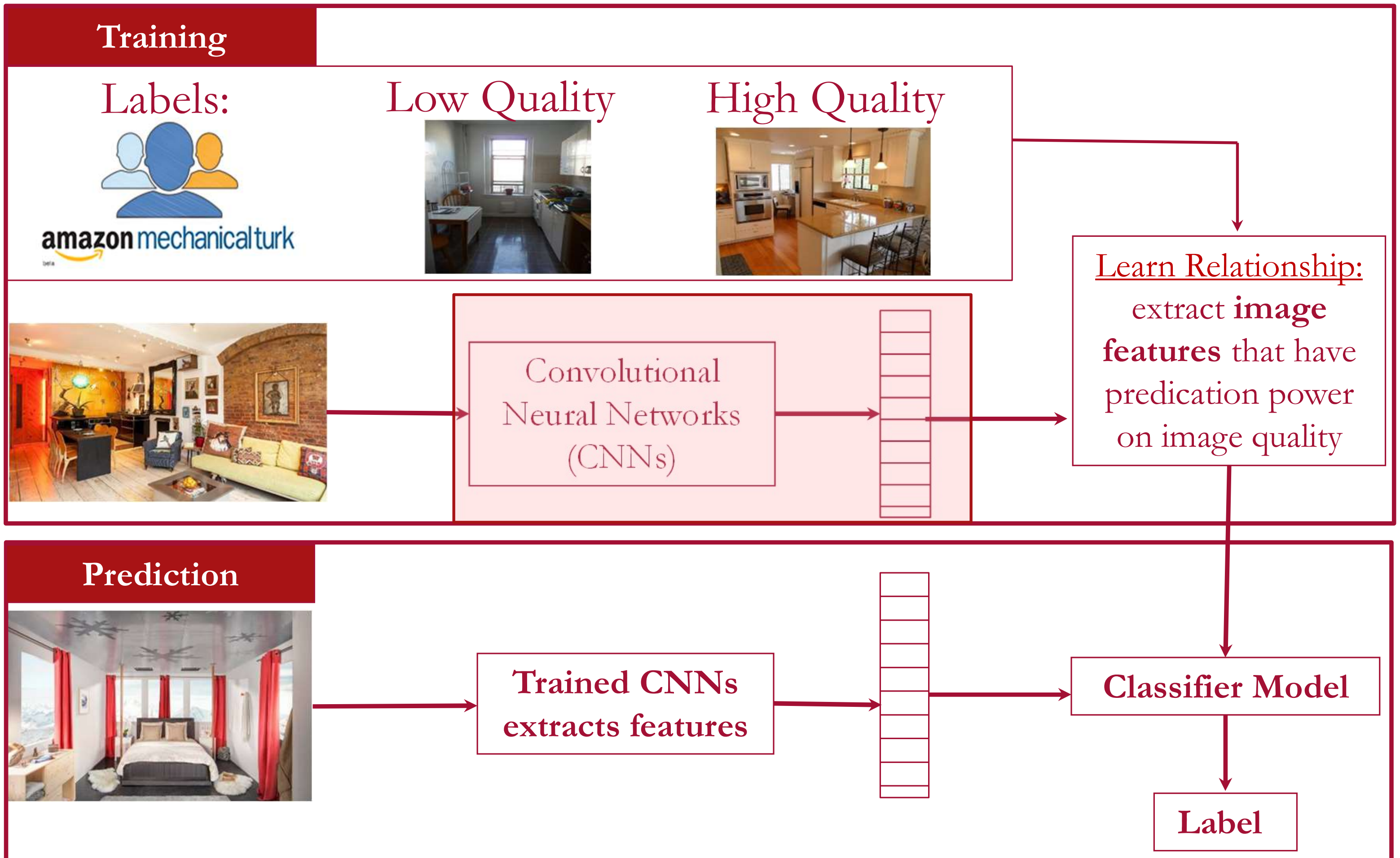
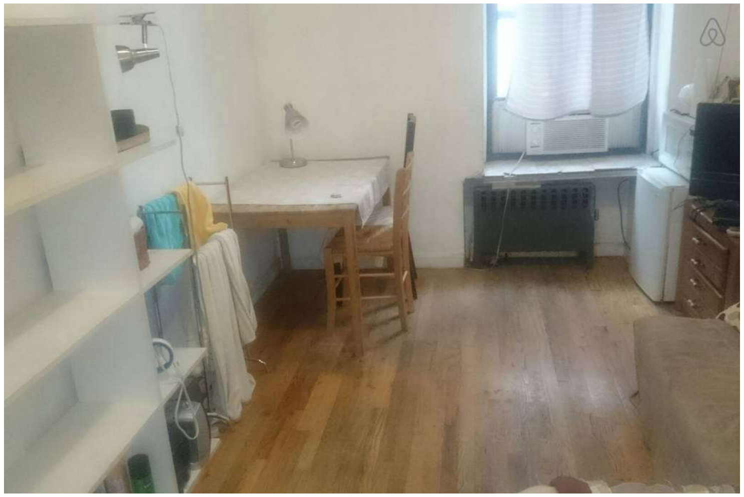


Image quality prediction examples (1)

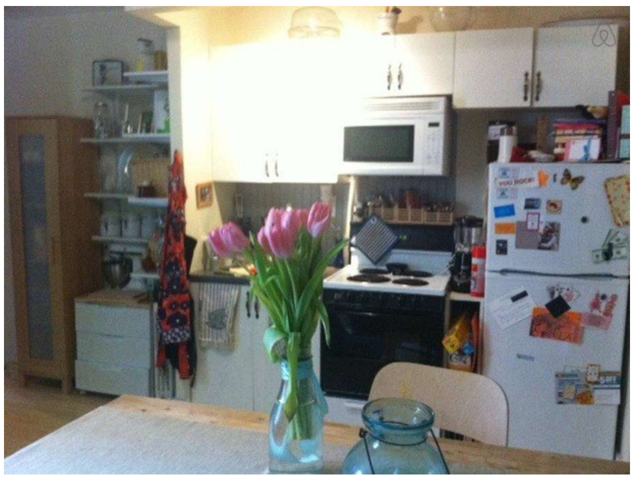
Low quality



High quality

Image quality prediction examples (2)

Low quality



High quality

Finally… incorporating predicted X_{visual} into your econometrics model

Treatment Group
- Properties with images amateur → professional

Control Group
- Properties with images stayed amateur

Empirical Analysis
- Propensity Score Weighting + Difference-in-Difference

Economic impact of high-quality images
- increase demand by 14%

Zhang et al. (2021) "What Makes a Good Image? Airbnb Demand Analytics Leveraging Interpretable Image Features." *Management Science*

General steps of extracting features

Is an off-the-shelf package (with good prediction performance) available?

- **Yes** → **Use existing API/packages**

  For common tasks:
  Object detection/recognition, face recognition

  Examples:
  Microsoft Azure, Google Vision API

- **No** → **Is an existing training data available?**

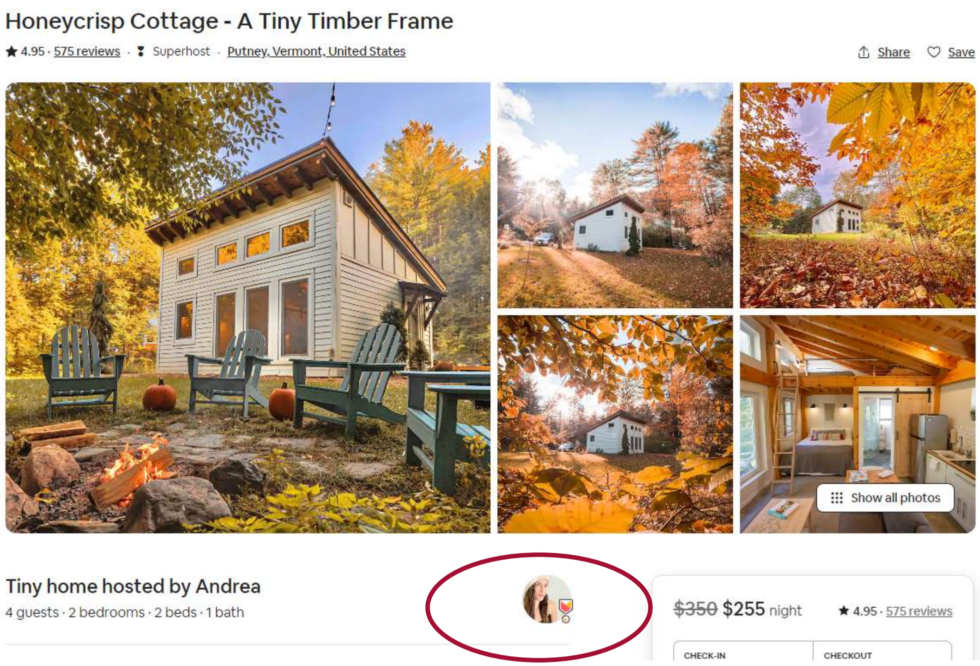
  - **Yes**
    Using the existing data to train your ML model

    Examples:
    Predicating ethnicity of Airbnb hosts

  - **No**
    Constructing a training data
    → Using the existing data to train your ML model

    Examples:
    Predicting image quality

Example (2): Ethnicity of Airbnb Hosts



Research question: What are the effects of using Smart Pricing algorithm on the revenue, for Airbnb hosts across different ethnic groups?

Econometrics model: y~f(X)

- y: Airbnb property's monthly revenue
- $\{X_1, X_2, \ldots, X_k,\}$: review, location, ….
- Key variable: ethnicity of an Airbnb host
  - ➢ How to obtain such info? Extracting from the Airbnb host profile picture

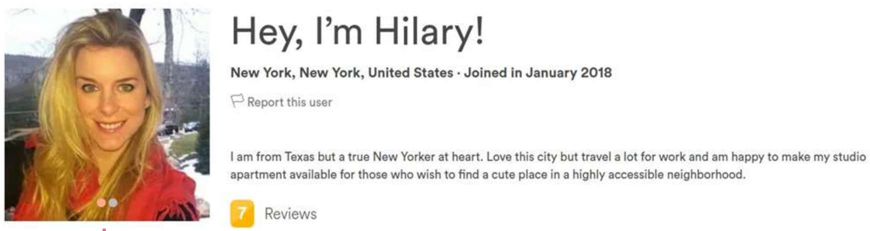
Zhang et al. (2021) "Can an AI Algorithm Mitigate Racial Economic Inequality? An Analysis in the Context of Airbnb." *Marketing Science*

Example (2): Ethnicity of Airbnb Hosts

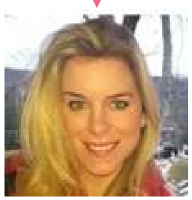
Creating a dataset to train an ethnicity classifier:

- Combine multiple public face databases with *ethnicity label*:
  - The color Facial Recognition Technology (FERET) Database collected by the National Institute of Standards and Technology (NIST),
  - Chicago Face Database (CFD) collected by the University of Chicago,
  - Face Place database collected by Brown University,
  - the IMDB-WIKI image database created by the Computer Vision Lab
- We then employ the ResNet-50 framework (Cao et al. 2018) to train an ethnicity (+ age) classifier on the consolidated training data.

ResNet-50: Architecture, Layer Operations, and Training Framework

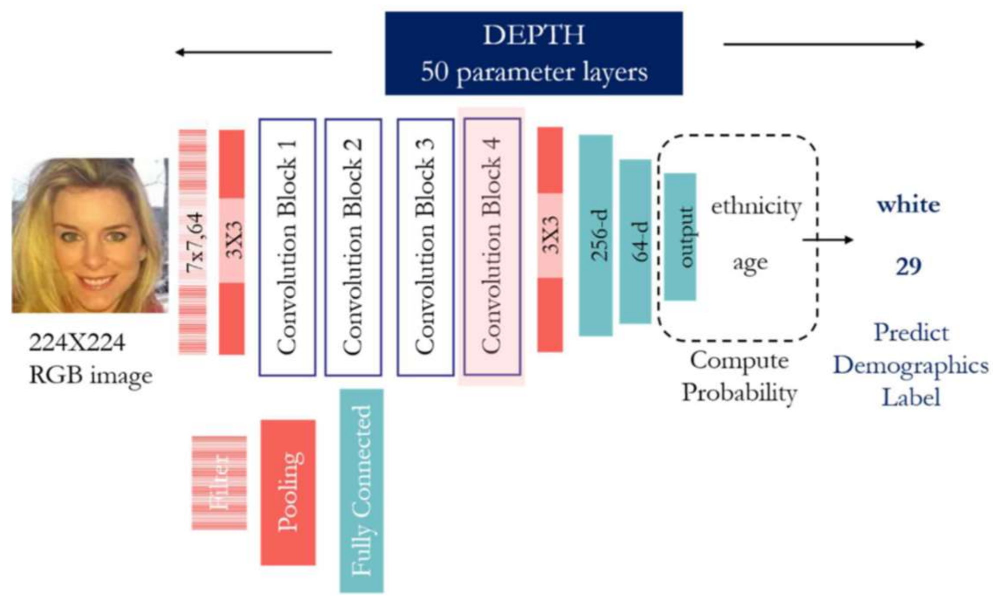


Detect & Extract Face



Crop & Resize the face

Feature Extraction & Demographic Prediction Architecture



Agenda

Visual Data Analytics

- Why does visual data matter?
- How can visual data be useful?

Image Analytics: Frameworks and Examples

- To extract IV: Airbnb images, facial images, and ads. images etc.
- **To construct DV or key metrics**
- To develop a prediction model

Video Analytics: Working with Multi-modal Visual Data

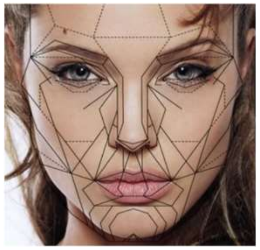
- Steps & tools
- Two applications: YouTube influencers & in-store shopping video

Image Analytics: extracting DV

If you are interested in measuring a key metrics that offers important managerial implications:

- Constructing y_{visual} from visual data
  - A 'follow-up' econometrics analysis is optional
- **The key is to make sure that:**
  - y_{visual} is indeed interesting and important
  - Constructing y_{visual} from visual data is effective (why is visual data useful?)
  - The extracted y_{visual} is aligned with human's perception/judgement of *y* (i.e., external validity, which applies to a broader setting)

Example (3): celebrity visual potential



FACE

Research objective: building a scalable model that extracts facial attributes from a person's face photo and scores the person on *Celebrity Visual Potential* (CVP).

Steps:

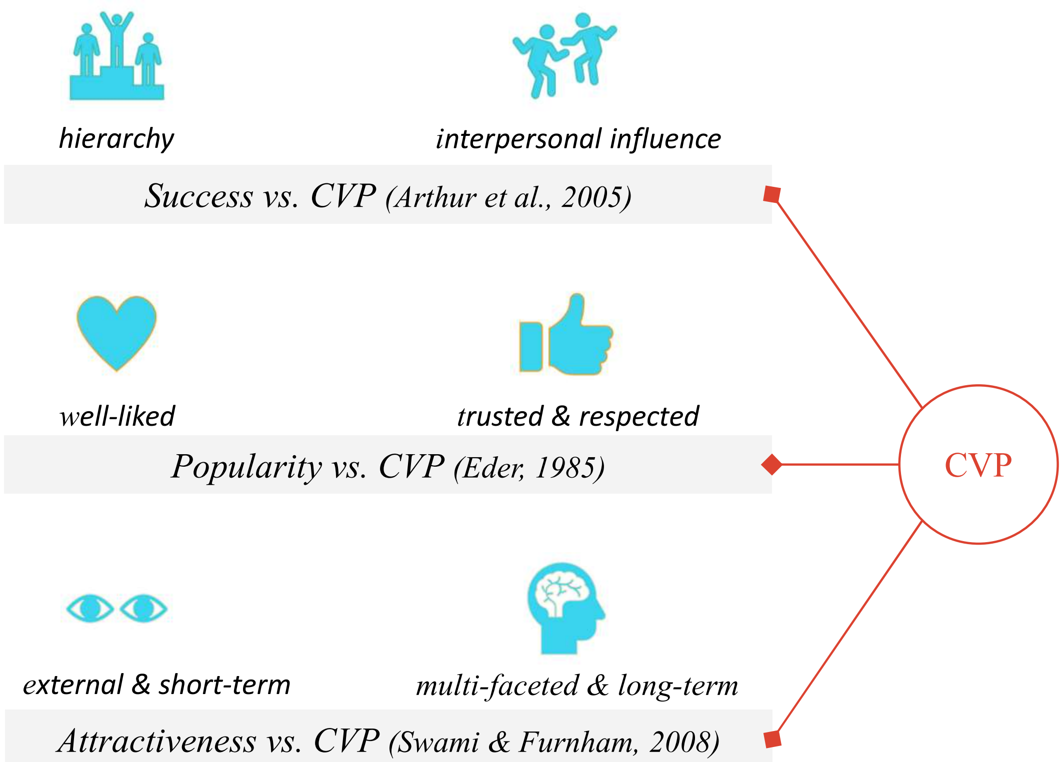
1. Defining CVP – what is CVP and why is it important
2. Constructing a training data that allows us to predict CVP
3. (optional) Exploring how CVP varies with one's face attribute
4. Verifying external validity
5. Application: how can one use extracted CVP?

Feng et al. (2022) "Beyond a Pretty Face: An AI Method to Score Celebrity Visual Potential," *Working paper*

Example (3): celebrity visual potential

What is CVP and Why CVP?

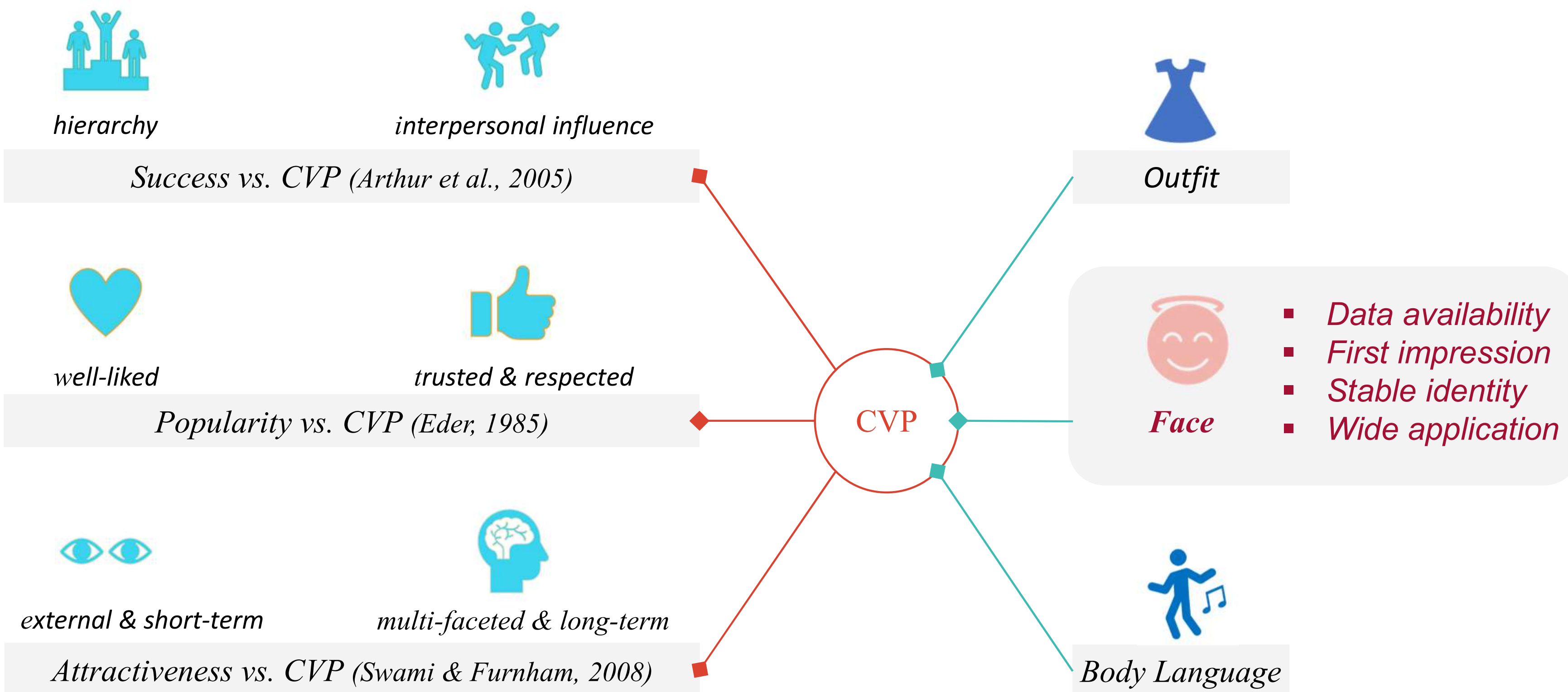
Celebrity Visual Potential is a Broad Concept



Example (3): celebrity visual potential

Why visual data (why face)?

Celebrity Visual Potential is a Broad Concept



Example (3): celebrity visual potential

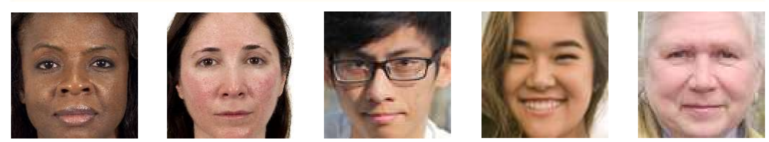
Data Construction

Celebrity Facial Images

- CelebFaces Attributes Dataset (CelebA)
(Liu et al., 2015)
- The IMDB-WIKI Dataset (IMDB-WIKI)
(Rothe, Timofte, and Gool, 2015)
- Labeled Faces in the Wild (LFW)
(Huang et al., 2007)

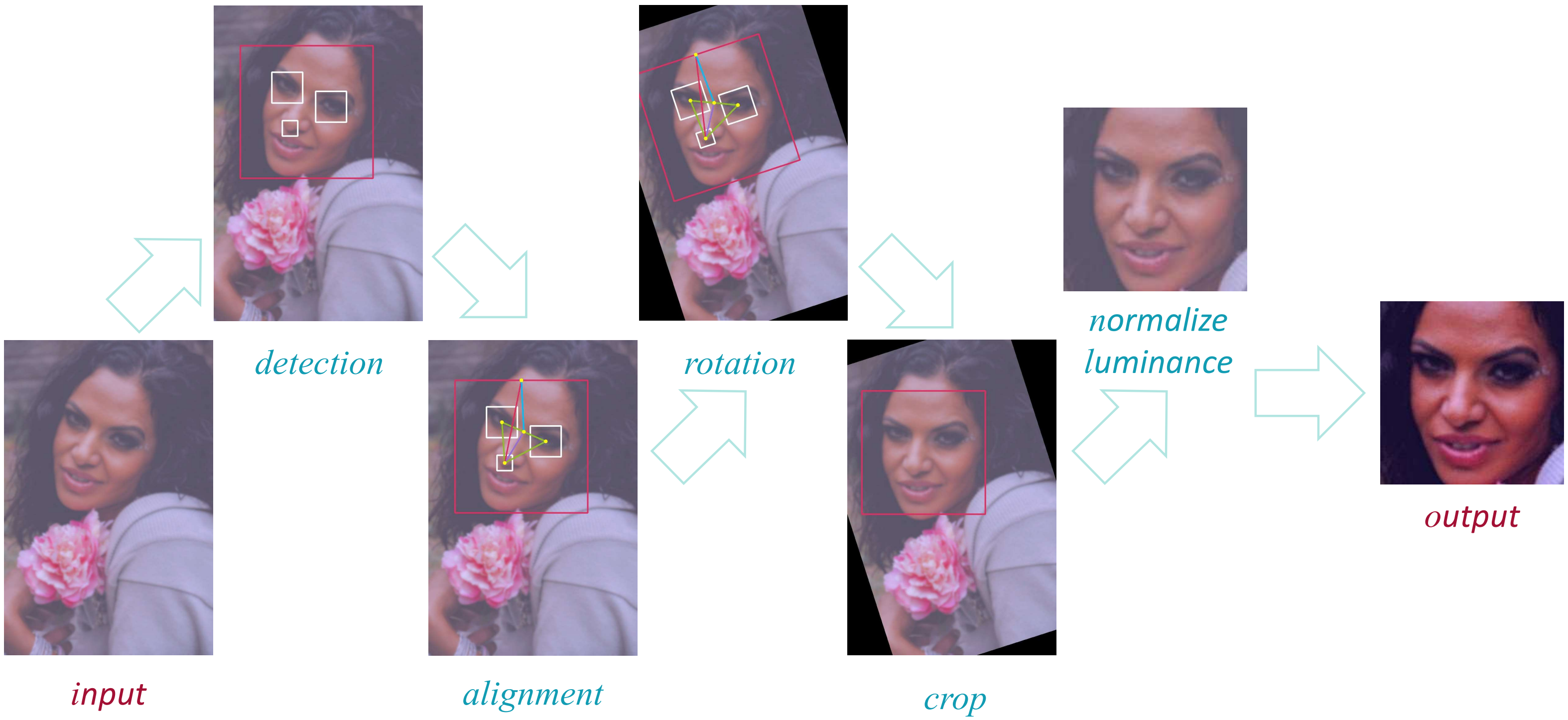
Non-Celebrity Facial Images

- Google Facial Expression Comparison (FEC)
(Vemulapalli and Agarwala, 2019)
- Chicago Face Database (CFD)
(Ma, Correl, and Wittenbrink, 2015)
- MPLab GENKI Database (GENKI-4K)
(Whitehill et al., 2009)
- Multitask Facial Landmark Dataset (MTFL)
(Zhang et al., 2014)
- Selfie Dataset (Selfie)
(Kalayeh et al., 2015)
- Flickr-Faces-HQ Dataset (FFHQ)
(Karras, Samuli, Aila, 2019)



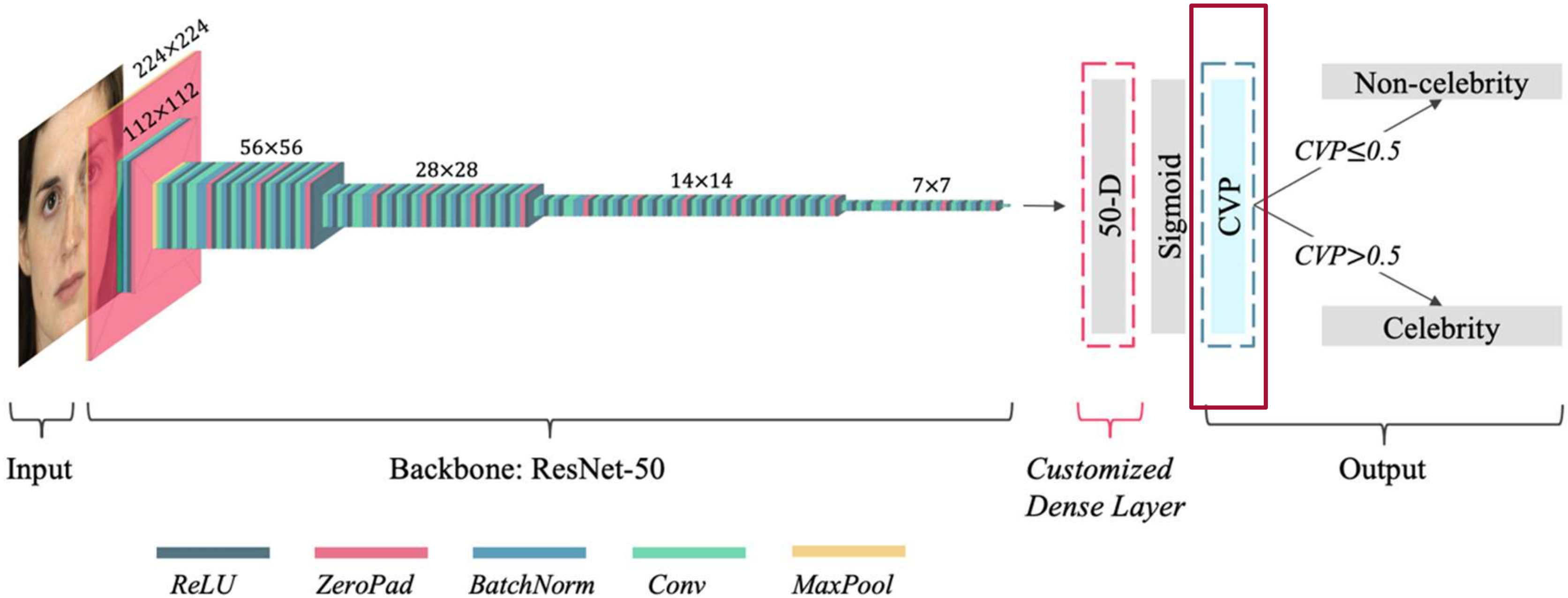
Example (3): celebrity visual potential

Data preprocessing & Training model



Example (3): celebrity visual potential

Data preprocessing & Training model



Example (3): celebrity visual potential

How good is the model prediction?

Fine Tuning and Model Selection

❖ Selection criterion:

Accuracy of classification

Stability of optimization curve

Model	Backbone	Preprocess	Optimizer	Stability	Accuracy
1	SE-ResNet-50	None	SGD	Low	0.9475
2	SE-ResNet-50	None	AdaDelta	High	0.8708
3	SE-ResNet-50	1+2+3	AdaGrad	High	0.9217
4	SE-ResNet-50	1+2+3	RMSprop	Low	0.9192
5	SE-ResNet-50	1+2+3	AdaDelta	High	0.8567
6	ResNet-50	1+2+3	SGD	Low	0.9458
7	ResNet-50	1+2+3	AdaDelta	High	0.9300
8	ResNet-50	1+2+3+4	AdaGrad	Low	0.9350
9	***ResNet-50***	***1+2+3***	***AdaGrad***	***High***	***0.9546***

Example (3): celebrity visual potential

(optional) Interpretation: what does the model tell us?

Interpretation of model prediction

Two groups of facial images with one diverging feature and all other features controlled to a certain level

(Theory motivated) Facial Features	Direction	t-stat
Facial width-to-height ratio	Negative	2.56
High cheekbones	Positive	63.04
(Dark) Color	Positive	36.80
Thin jaw	Negative	1.54
Mouth–nose distance	Negative	0.24
Large eyes	Positive	2.39
Sex dimorphism	Positive	19.19
Mouth–chin distance	Positive	1.58
Babyfaceness	Negative	1.79
Symmetry	Positive	14.57
Averageness	Negative	0.96

Example (3): celebrity visual potential

Validating model predicted *y* -- a controlled experiment

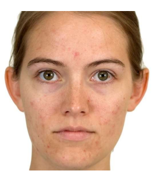
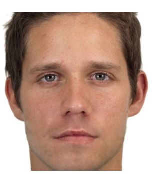
1. How likely do you think this person could become a celebrity?
2. If this person were on your favorite social media site, how likely you would subscribe to his/her content or follow him/her?
3. This person is active on social media. Relative to other people, how many followers or subscribers do you think s/he has?

Original Face Images → Model Prediction → Human Judgement → Validation

External validity

CVP manipulation
Across Race & Gender Groups



Model Prediction

Different CVP Levels

- *CVP Score ~ 0.0*
- *CVP Score ~ 0.2*
- *CVP Score ~ 0.4*
- *CVP Score ~ 0.6*
- *CVP Score ~ 0.8*

Human Judgement

Experiment 1: Rate
Experiment 2: Rank

Three Questions on CVP

Control Variables

- Race / Gender / Age
- Political Ideology
- Finance / Social

Validation

Regression Analysis

- Intercept
- *p*-Value
- Incremental R^2
- Moderation and Mediation Effects

Participant Response

Example (3): celebrity visual potential

External validity: real-world evidence

Validating CVP: LinkedIn Dataset

Randomly selected employees of the Fortune 500 on LinkedIn

- Profile images of 5 C-suite executives
- Profile images of 5 average employees

Calculated the average CVP for

- 150 C-suite executives
- 150 average employees

Mean score for c-suite executives: 0.845
vs. average employees: 0.197

Example (3): celebrity visual potential

External validity: real-world evidence

Validating CVP: Instagram Dataset

Data:

2,105 Instagram selfie posts

500 influencers

2016 ~ 2020

DV: Popularity of a post

Contextual variables:

- Influencer: gender, beauty
- Popularity: likes & comments
- Image aesthetics
- Text length, readability etc.
- Textual sentiment

Finding:

➢ The effect of CVP score is significant even if control for the contextual variables and attractiveness.

Example (3): celebrity visual potential

How might predicted *y* be useful?

Managerial implications of CVP

Advertising star selection
Persuading investors

Star style design

Celebrity Visual Potential

Scored Facial Features

Influencer marketing
Virtual influencer design

Leader selection
Election result prediction

Agenda

Visual Data Analytics

- Why does visual data matter?
- How can visual data be useful?

Image Analytics: Frameworks and Examples

- To extract IV: Airbnb images, facial images, and ads. images etc.
- To construct DV or key metrics
- **To develop a prediction model**

Video Analytics: Working with Multi-modal Visual Data

- Steps & tools
- Two applications: YouTube influencers & in-store shopping video

Image Analytics: prediction model

- A prediction model: **y~f(X)**
  - X={$X_1, X_2, \dots, X_k, X_{visual}$}

- Extracting X_{visual} from visual data then 'plug in' the prediction model
  - X_{visual} can be the *image*, if we are less interested in interpretability but more so in boosting the prediction accuracy

- **The key is to ensure that including visual information helps predict *y***
  - Ideally, please try to rationalize why does X_{visual} strengthen prediction power? What does it capture beyond what's captured by $X_1, X_2, \dots, X_k$?
  - Is including X_{visual} economically significant?

Example (5): predicting product return

Research objective: building a model that uses product images and traditional measures available prelaunch to predict individual item return rates.

Steps:
1. Obtaining a dataset that includes return, product image, and other measures
2. Extracting visual info. from product images
3. Comparing prediction models and show that a model that uses product image *outperforms* the other models
4. Validation: what's the economic **significance by using images**?

Dzyabura et al. (2022) “Leveraging the Power of Images in Managing Product Return Rates." *Working paper*

Example (5): predicting product return

1. Motivating the problem: is product return prediction (i.e., *y*) a big deal?
2. (Hopefully) yes → why should product image be included?
    - What information from product images should be included?
3. Train and compare various prediction models.
4. Circle back to question 1 and 2: role of product images.
    - The economic significance of including product images in the prediction model.

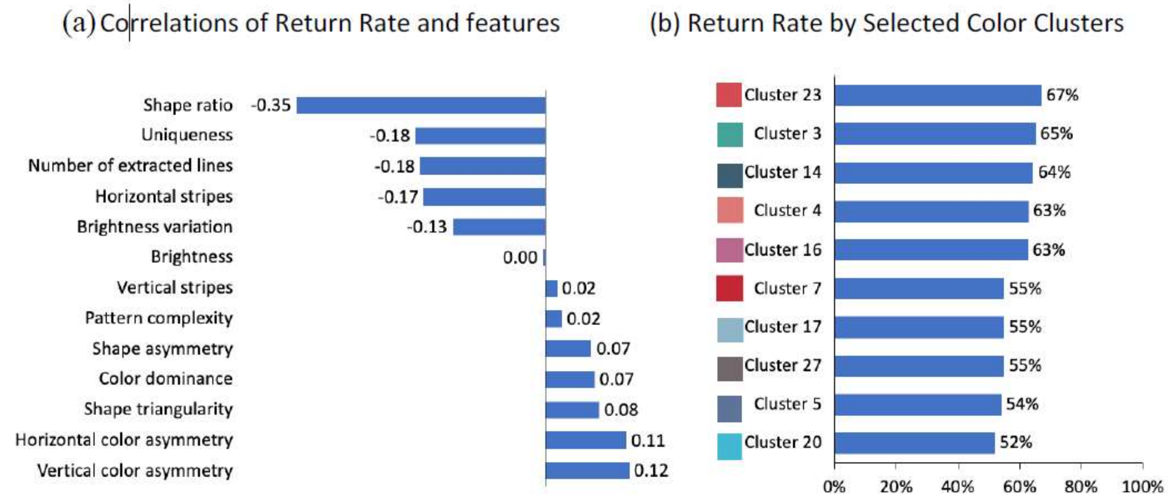
Dzyabura et al. (2022) "Leveraging the Power of Images in Managing Product Return Rates." *Working paper*

Example (5): predicting product return

1. Motivate the problem: is product return prediction a big deal?
2. (Hopefully) yes: why should product image be included?

“ …we observe that **return rates** for fashion items bought online range from 13% to 96%, with **an average of 53%** – many items are not profitable.”

Figure 11. Model-Free: Correlations of Return Rates and Interpretable Image-based Features



Dzyabura et al. (2022) “Leveraging the Power of Images in Managing Product Return Rates." *Working paper*

Example (5): predicting product return

3. What information from product images should be included?

Figure 4. Example RGB Color Histogram Encoding of an Apparel Item

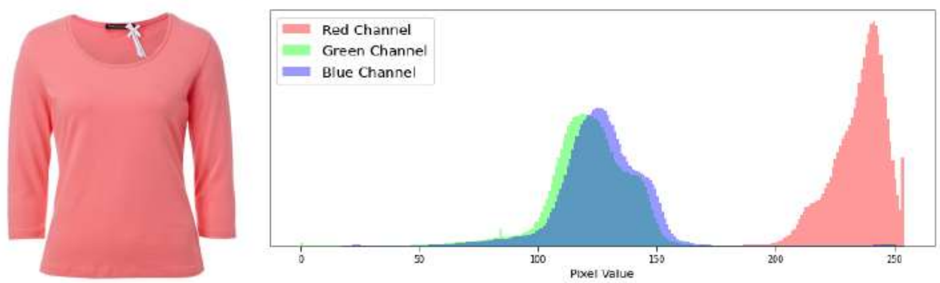
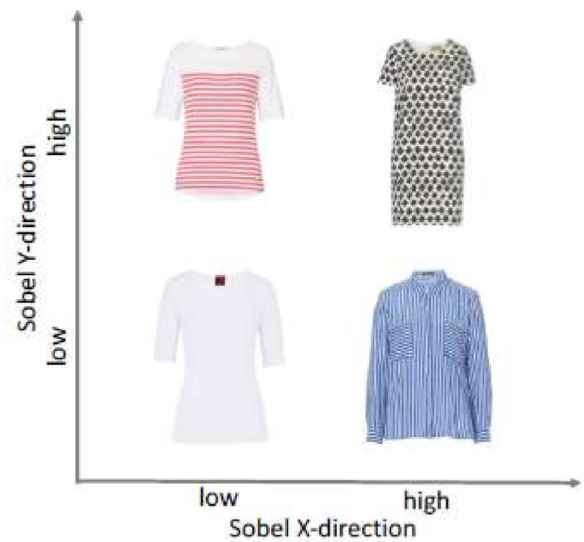


Figure 10. Illustration of Shape Ratio

Figure 8. Illustration of Pattern Direction (Sobel X- and Y- Directions)



Dzyabura et al. (2022) “Leveraging the Power of Images in Managing Product Return Rates." *Working paper*

Example (5): predicting product return

4. Train and compare various prediction models, i.e., model performance
5. Is including product images economically significant?

Table 4. Expected Profit Improvement Using Different Predictive Models

Model	Features	Percent Items not Launched	Profit Improvement vs. Launch All Items
Non-image baseline	Category, seasonality, and price	5.98% (0.11)	6.81% (0.18)
Color labels added to baseline	Category, seasonality, price, and color labels	6.26% (0.13)	7.16% (0.19)
CNN Features	Category, seasonality, price, CNN from image	7.13% (0.12)	8.29% (0.23)

Dzyabura et al. (2022) "Leveraging the Power of Images in Managing Product Return Rates." *Working paper*

Agenda

Visual Data Analytics

- Why does visual data matter?
- How can visual data be useful?

Image Analytics: Frameworks and Examples

- To extract IV: Airbnb images, facial images, and ads. images etc.
- To construct DV or key metrics
- To develop a prediction model

Video Analytics: Working with Multi-modal Visual Data

- Steps & tools
- Two applications: YouTube influencers & in-store shopping video

Video Data: Source

- Social media
- Company sites
- Live streaming
- MOOCs
- Autonomous vehicle
- Retail stores
- Parking lots
- Streets
- …

54% View videos every day

78% View videos every week

72% customers would rather learn about a product/service by way of video

10% Viewers retain messages when watch it in a **text**

95% Viewers retain messages when watch it in a **video**

Video Marketing Statistics in 2022 (summary of source: https://invideo.io/blog/video-marketing-statistics)

Video Data

Example: YouTube Video

- A social influencer promoting cosmetics products in a video sponsored by *Lancôme*.

UPDATED EVERYDAY MAKEUP & SIGNATURE NUDE & BROWN LIP LOOKS FOR FALL/AUTUMN

19K Dislike Share Save ...

Patricia Bright
2.87M subscribers

449,084 views • Aug 15, 2019

Hey guys so It's about time I did an update makeup routine, sharing my favourite every day makeup and nude lip looks I do regularly!

How would you incorporate this video into your model (analysis)?

How To Analyze A Video?

- A video is a sequence of three modalities of data
  - Image (frame)
  - Audio (voice)
  - Text (speech)
- Approach one: decomposing a video by data modality
  - Then, analyze each modality to extract image/audio/text features
  - Video is treated as a portfolio of (static) attributes

Video As Separate Sets of Data (by modality)

"It takes some time…"

"Next, I'm gonna…"

"… Lancôme foundation…"

"This is the classic…"

"This is actually a newer version…"

"… light coverage foundation"

How To Analyze A Video?

- A video is a sequence of three modalities of data
  - Images (frame)
  - Audio (voice)
  - Text (speech)
- Approach one: decomposing a video by data modality
  - Then, analyze each modality as if you were performing:
    - **image analysis**
    - **audio analysis**
    - **text analysis**
  - Video is treated as a portfolio of (static) attributes

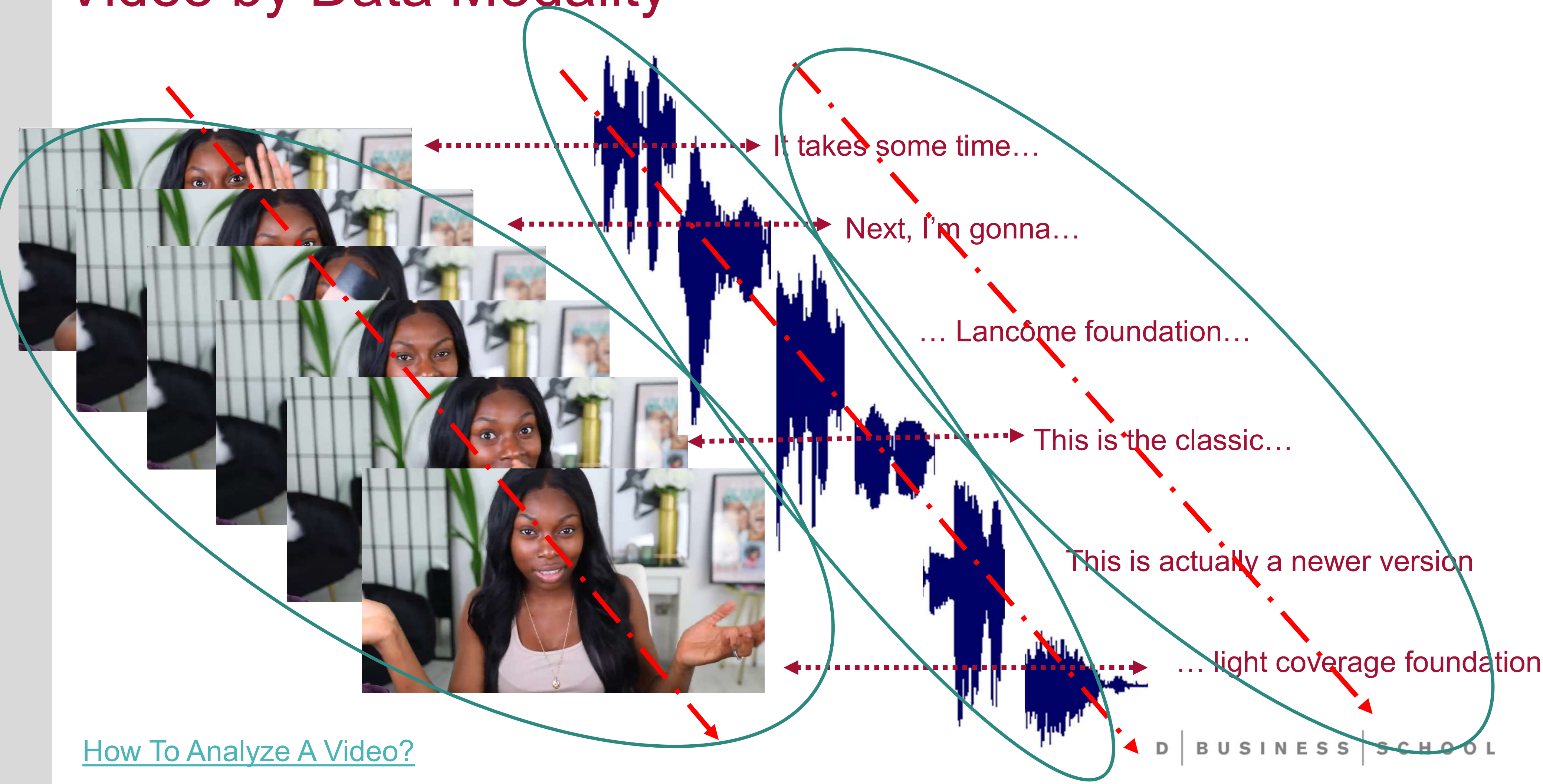
Example:
Videos posted by
Social Influencers

- Approach two: incorporating time-dependence
  - Dynamics in the video features
  - E.g., key variables are time-related

Example:
In-store shopping
videos

- Approach three: crosslinking the multi-modal data
  - E.g., coherence/complement/substitution among one's verbal and non-verbal cues
- ...

Video by Data Modality



How To Analyze A Video?

D | BUSINESS | SCHOOL

YouTube Videos by Influencers (Example 1)

UPDATED EVERYDAY MAKEUP & SIGNATURE NUDE & BROWN LIP LOOKS FOR FALL/AUTUMN

19K Dislike Share Save ...

Patricia Bright
2.87M subscribers

449,084 views • Aug 15, 2019

Hey guys so It's about time I did an update makeup routine, sharing my favourite every day makeup and nude lip looks I do regularly!

* Video is sponsored by Lancome*

1,105 Comments

Question:
How does a brand-sponsored video affect the influencer's reputation?

- **Key**: need to extract a set of video features that might affect how one is perceived.

Cheng and Zhang (2022) Reputation Burning: Analyzing the Impact of Brand Sponsorship on Social Influencers. *Working paper*

Example 1: What Might Matter in The YouTube Context?

UPDATED EVERYDAY MAKEUP & SIGNATURE NUDE & BROWN LIP LOOKS FOR FALL/AUTUMN

19K Dislike Share Save ...

Patricia Bright
2.87M subscribers

449,084 views • Aug 15, 2019

Hey guys so It's about time I did an update makeup routine, sharing my favourite every day makeup and nude lip looks I do regularly!

* Video is sponsored by Lancome*

1,105 Comments

"Face"

Kraut and Johnston (1985)

"Voice"

Hwang et al. (2021)

"Presentation"

Zhang et al. (2021)

Example 1: Extracting A List of Theory-Driven Features

- Five basic video properties that were found to affect the viewer's attention and engagement (Zhou et al. 2021).

- Aesthetic appeal of a video can affect viewers' preferences and satisfaction (Moorthy et al. 2010; Zhou 2021).

Basic Video Properties

Video Length

Speaking Rate

Scene Count

Average Scene Length

Sentiment

Visual Aesthetics

Motion

Warm Hue

Saturation

Brightness & Contrast

Clarity

Summary: Extracted Video Features

The Influencers' Voice Feature

Loudness

Loudness Variation

Pitch

Talking Duration

The Influencers' Face & Demographics

Emotion (facial expression)

Gender & Age

Attractiveness

Face Count

- Same person's voice may change across videos (Hwang et al., 2021).
- Vocal features affect perceived personal traits (e.g., attractiveness, dominance, capability) (Peterson et al. 1995).

- Face as a primary channel for the nonverbal communication (Ekman and Oster 1979).
- Appearance features impact perceived interpersonal relationship (Zhang et al. 2020).

Zhou, Mi, et al. "EXPRESS: Consumer Behavior in the Online Classroom: Using Video Analytics and Machine Learning to Understand the Consumption of Video Courseware." *JMR*
Moorthy, A. K., Obrador, P., & Oliver, N. (2010, September). Towards computational models of the visual aesthetic appeal of consumer videos. *ECCV*
Hwang, Serim and Liu, Xiao and Srinivasan, Kannan, Voice Analytics of Online Influencers—Soft Selling in Branded Videos (January 26, 2021).
Peterson, Robert A., Michael P. Cannito, and Steven P. Brown (1995), "An exploratory investigation of voice characteristics and selling effectiveness," *JPSP*

Example 1: Exploring Videos—Underlying Behavior?

Compared to the organic videos, <u>the influencers in sponsored videos</u>:

- Speak faster, less loud, less verbose, lower pitch

  Scripted Speaking? | More nervous? | More authority?

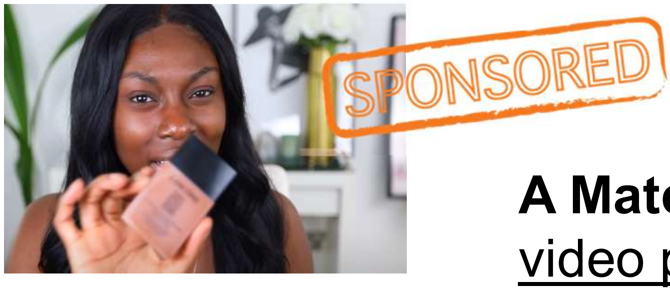
- Are more vibrant (involve more/stronger motions), smile less

  Using more gestures to 'make a point'? | Focusing on grabbing viewer's attention?

- Look more (facial) attractive

  Applying make-up?

Example 1: YouTube Influencers



A Matched Sample: constructing <u>similar influencer-video pairs</u>

- *Treated unit*: influencer 1 - sponsored video 1
- *Control unit*: influencer 2 - organic video 2
  - The influencers are “matched” on influencer characteristics
  - Videos are “matched” on video characteristics

Next, apply your econometrics method…

Example 2: Shopping Videos in A Retail Store

A Video Frame of Customers
Lining Up to Check Out

Question:
How does consumers' compliance of mask recommendation (and the motivation) affect their shopping behavior during the pandemic?

- **Key**: need to construct
  1) a customer's compliance behavior
  2) a customer's shopping trajectory and decision

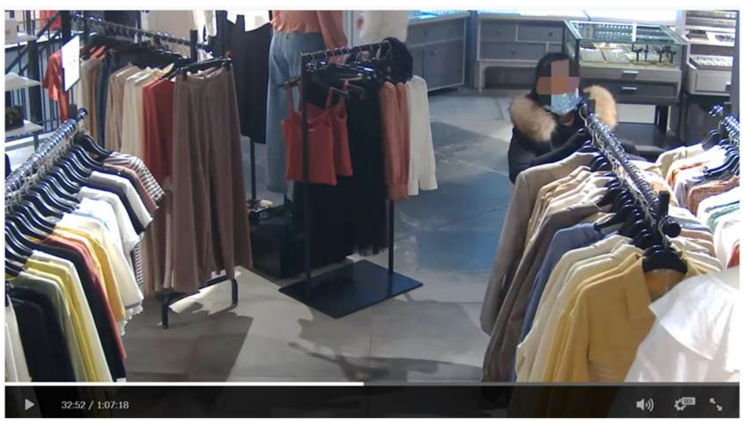
  Comparing: before vs during the pandemic periods.

Zhang et al. (2022) "Unmasking Social Compliance Behavior During the Pandemic," *Marketing Science*

Example 2: Shopping Videos in A Retail Store

Crosslinking Multiple Videos (Face Matching)

Store Entry



Store Exit (checkout)

Example 2: Shopping Videos in A Retail Store

Shopping Trajectory (entry-shopping-exit)

- **1) Construct the shopping trajectory for a customer**
  - Step 1: identify human face from all video frames (images)
  - Step 2: find matched faces → same customer appearing at different areas in the store
  - Step 3: incorporate time-dependence → a sequence of behavior
  - **Step 4: extract video features of interest**
    - Counting shoppers in videos → (dynamic) store crowdedness
    - Distances between shoppers (and the cashier)
    - How long did a shopper shop around?
    - What did the shopper buy?
    - …
- **2) Construct the compliance behavior of a customer**
  - Wearing a mask? What kind of mask?
  - Face coverage?
  - Social distancing?
  - Avoiding crowded areas?

Example 2: Shopping Videos in A Retail Store

Extracting compliance variables

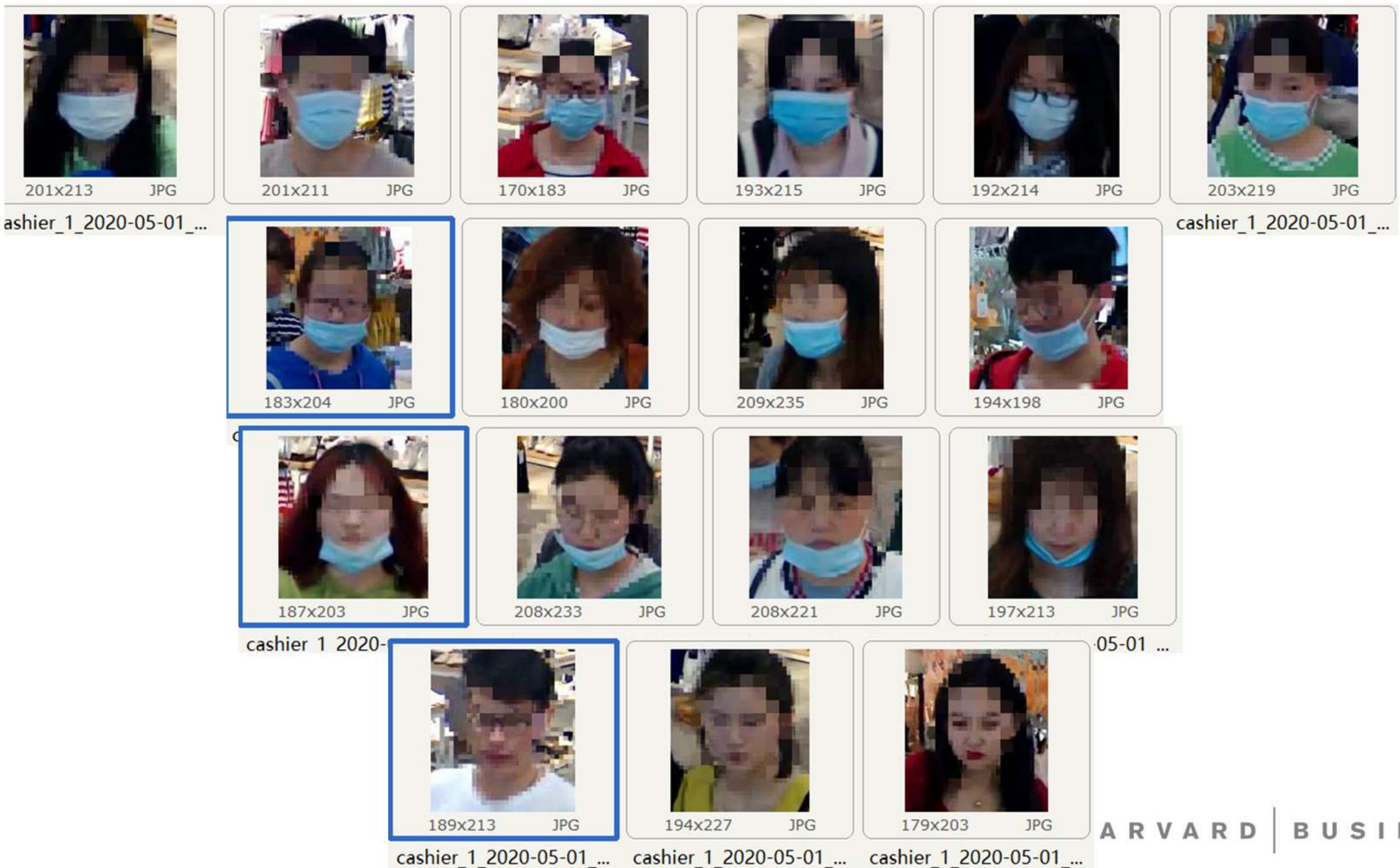
Social Distancing

Mask-wearing (Surgical vs N-95)

Mask coverage (full coverage vs nose uncovered)

Example 2: Shopping Videos in A Retail Store

Measuring Mask Fit



Example 2: Shopping Videos in A Retail Store

Inferring Motives From Videos

- Connecting shopper's compliance level to the *changes in the shopping behavior* before vs during the pandemic
  - Shopping duration; store visit frequency
  - Quantity and price of purchased items
  - Diversity in shopping (purchasing multiple categories/shelves, or restricted to only one category?)

Exploratory Analyses on Mask-wearing Motives

Fully-compliant

Wear a mask to **protect self** (sensitive to their own health risk)

- Shopped more quickly than pre-pandemic
- Practiced social-distancing
- Bought products with deeper discounts, high volume sales

Partially-compliant

Wear a mask primarily to comply to **social responsibility**

- Shopped slowly
- Did not avoid crowded areas
- More diverse shopping

Unmasked

Do not wear a mask at all

- Shop speed stayed unchanged during pandemic

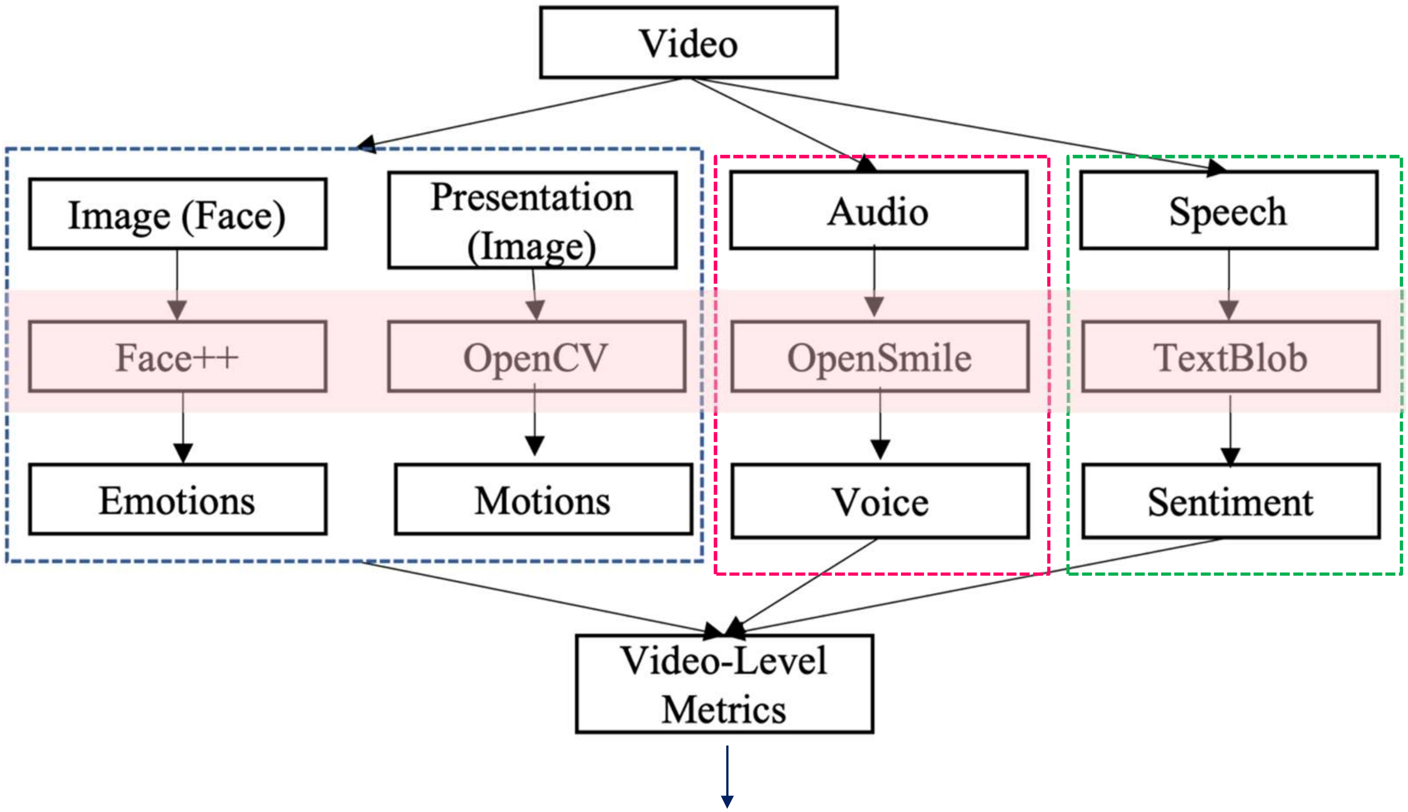
Thank you!

APPENDIX: A Few Technical Notes & Toolkits

- **Basic Video Properties:**
  - Video scene detection: PySceneDetect (https://pyscenedetect.readthedocs.io/en/latest/)
  - Speech to subtitle: AutoSub (https://github.com/agermanidis/autosub)
  - Subtitle sentiment analysis: TextBlob (https://github.com/sloria/TextBlob)
- **The Influencers' Voice Feature: *OpenSmile* (https://audeering.github.io/opensmile/get-started.html#default-feature-sets)**
- **The Influencers' Face & Demographics: Face++ (**SFace: An Efficient Network for Face Detection in Large Scale Variations**)**
- **Visual Aesthetics**
  - Dense Optical Flow in OpenCV (https://docs.opencv.org/master/d4/dee/tutorial_optical_flow.html)
  - Background Substraction in OpenCV (https://docs.opencv.org/3.4/d1/dc5/tutorial_background_subtraction.html

Processing Video Data: Steps and Tools

- STEP 1: Extract data modality
  - *PySceneDetect*: present the video as a sequence of **images** (scenes)
  - *VideoClip*: extract the **audio** file as .wav files
  - *Autosub:* obtain the speech (**text**) content of each video (or *Speech-to-Text* by Google)
- STEP 2: Processing each modality
  - E.g., facial expression, brand logos, voice loudness & pitch, use of language
- STEP 3: Aggregate the video-level measures
- STEP 4: Subsequent analyses



Incorporating videos into formal analysis, for example:

- How does smiling correlate with viewer engagement?
- Manipulate video features of interest in a controlled setting

https://github.com/Breakthrough/PySceneDetect
https://zulko.github.io/moviepy/ref/VideoClip/VideoClip.html#videoclip
https://github.com/agermanidis/autosub