

# Using Computer Vision to Study Visual Trends in Historical Advertisements

---

Melvin Wevers

March 8, 2018

Digital Humanities Lab - KNAW Humanities Cluster

1. Advertisements as a Historical Source
2. Studying Images using Computer Vision
3. Back to Historical Advertisements: SIAMESE

# Advertisements as a Historical Source

---

# Background

- Ads offer “insight into the **ideals and aspirations** of past realities. They show the **state of technology**, the **social functions of products**, and provide **information on the society** in which a product was sold” (Marchand, 1985)
- What does the **career/genealogy of a concept** tell us about today’s dominant ways of understanding social change? (James & Steger, 2014)
- “**Consumption** as a privileged site for the fabrication of self and society, of culture and identity ...” (Camoroff et al, 2001)

- Dutch Digitized Newspapers archive (Delpher) contains  $\approx 20$  million advertisements (1890-1990)
- How can we use computation to detect/study cultural-historical changes in large corpora of digitized ads?

- Metadata
- Text
- Image



Figure 1: *Limburger Koerier*, April 13, 1938

# METADATA: SIZE

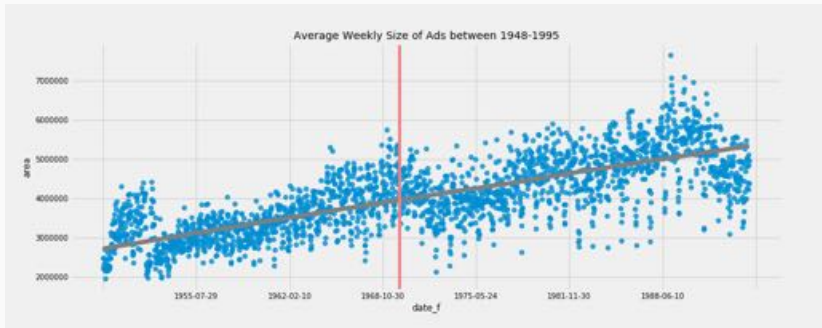


Figure 2: Average Size in *Algemeen Handelsblad* & *NRC*

# IMAGE-TEXT RATIO

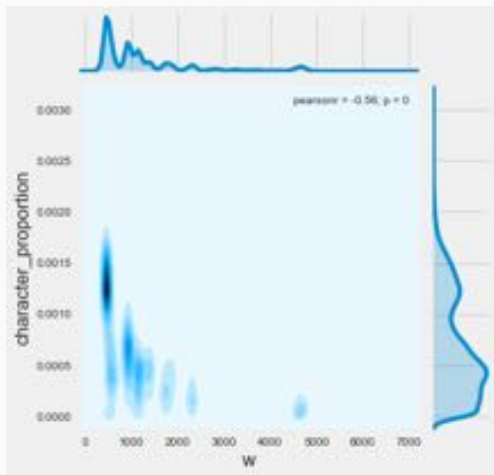


Figure 3: Image-Text Ratio vs Width



# TEXT: SEARCHING FOR WORDS IN ADS

The screenshot shows the Delpher search interface. At the top, the search query is 'sigaret AND camel'. The left sidebar contains filters for 'Soort bericht' (Advertentie: 487), 'Krantentitel', 'Plaats van uitgave', 'Herkomst', and 'Toevoegingen in Delpher'. The main area displays three search results, each featuring a thumbnail of a newspaper page and a star icon for favoriting. The results are as follows:

Advertentie	Krantentitel	Datum
<p>sigaret rokers, CAMEL, dankt zijn bereikbaarheid en succes aan zijn geheel eigen smaak. Gerust ook van CAMEL...</p>	Algemeen Handelsblad	27-01-1967
<p>sigaret rokers, CAMEL, dankt zijn bereikbaarheid en succes aan zijn geheel eigen smaak. Gerust ook van CAMEL...</p>	Algemeen Handelsblad	09-05-1967
<p>sigaret rokers, CAMEL, dankt zijn bereikbaarheid en succes aan zijn geheel eigen smaak. Gerust ook van CAMEL...</p>	Algemeen Handelsblad	26-09-1967

Figure 4: Searching Delpher for 'Sigaret' AND 'Camel'

# TEXT: BURSTY WORDS



# TEXT: CO-OCCURENCE ANALYSIS

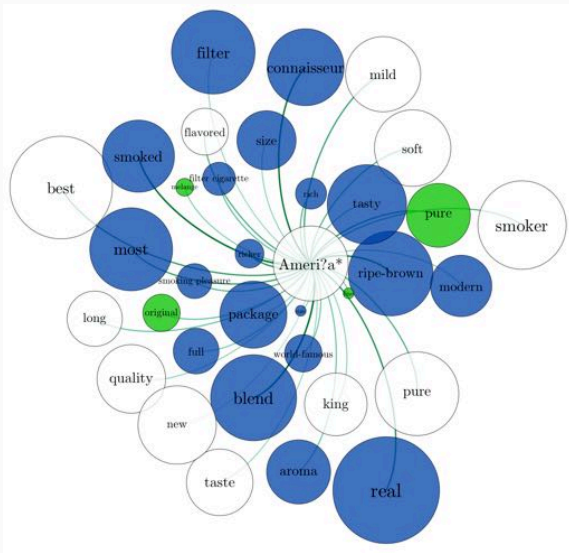


Figure 6: Characteristics of the American cigarette

# LIMITATIONS OF TEXT ANALYSIS

**Cotton Shadow**  
Een voorbeeld van onze nieuwste ontwerperij, een stijlvolle combinatie van de Amerikaanse samenleving en de beste kwaliteitsstoffen. De prijs is 14,95 per stuk, met gratis verzending. De prijs is 14,95.

**Twilight**  
Op voor de grote modie heeft C&A ontworpen een prachtige avondjurk. De prijs is 19,95 per stuk, met gratis verzending. De prijs is 19,95.

**Miss Young America**  
Een voorbeeld van onze nieuwste ontwerperij, een stijlvolle combinatie van de Amerikaanse samenleving en de beste kwaliteitsstoffen. De prijs is 24,75 per stuk, met gratis verzending. De prijs is 24,75.

**Missy Faith**  
Dit voorbeeld heeft een stijlvolle combinatie van de Amerikaanse samenleving en de beste kwaliteitsstoffen. De prijs is 49,75 per stuk, met gratis verzending. De prijs is 49,75.

**High Noon**  
Een voorbeeld van onze nieuwste ontwerperij, een stijlvolle combinatie van de Amerikaanse samenleving en de beste kwaliteitsstoffen. De prijs is 29,75 per stuk, met gratis verzending. De prijs is 29,75.

**Missy Faith**  
Dit voorbeeld heeft een stijlvolle combinatie van de Amerikaanse samenleving en de beste kwaliteitsstoffen. De prijs is 29,75 per stuk, met gratis verzending. De prijs is 29,75.

De nieuwe eigen Amerikaanse modie op een voorbeeld van de beste kwaliteitsstoffen. De prijs is 29,75 per stuk, met gratis verzending. De prijs is 29,75.

Figure 7: Advertisement C&A

# Studying Images using Computer Vision

---

- Goal: Using computers to gain high-level understanding of images/video
- ImageNET: 1.2m images, 1000 classes
- Precision improved with popularization of neural networks

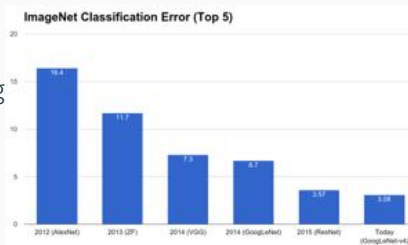


Figure 8: <https://www.quora.com/What-is-the-state-of-the-art-today-on-ImageNet-classification-In-other-words-has-anybody-beaten-Deep-Residual-Learning>



# POSSIBLE CONVOLUTIONS








Operation	Filter	Convolved Image
<b>Identity</b>	$\begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}$	
<b>Edge detection</b>	$\begin{bmatrix} 1 & 0 & -1 \\ 0 & 0 & 0 \\ -1 & 0 & 1 \end{bmatrix}$	
	$\begin{bmatrix} 0 & 1 & 0 \\ 1 & -4 & 1 \\ 0 & 1 & 0 \end{bmatrix}$	
	$\begin{bmatrix} -1 & -1 & -1 \\ -1 & 8 & -1 \\ -1 & -1 & -1 \end{bmatrix}$	
<b>Sharpen</b>	$\begin{bmatrix} 0 & -1 & 0 \\ -1 & 5 & -1 \\ 0 & -1 & 0 \end{bmatrix}$	
<b>Box blur</b> (normalized)	$\frac{1}{9} \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix}$	
<b>Gaussian blur</b> (approximation)	$\frac{1}{16} \begin{bmatrix} 1 & 2 & 1 \\ 2 & 4 & 2 \\ 1 & 2 & 1 \end{bmatrix}$	

Figure 10: <https://ujjwalkarn.me/2016/08/11/intuitive-explanation-convnets/>



# NEURAL NETWORK ARCHITECTURE FOR IMAGE CLASSIFICATION

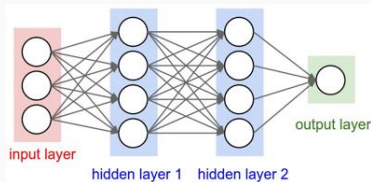


Figure 11: Neural Network. Visualization:  
<http://scs.ryerson.ca/~aharley/vis/conv/>

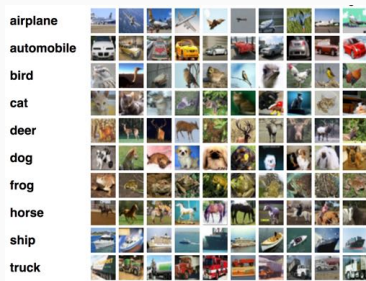


Figure 12: Output Layer

# CLASSIFYING IMAGES



Figure 13: Precision of Classification

# Back to Historical Advertisements: SIAMESE

---

# BACKGROUND SIAMESE PROJECT



Figure 14: Screenshot of KB Lab website

- Developed during KB Researcher-in-Residency
- Worked on this project together with Juliette Lonij
- Goal: get insights into visual trends
- To what extent can out-of-the-box neural networks be applied to historical advertisements?

- Raw data:  $\approx$  1.6m advertisements from two national newspapers between 1948-1995
- Data cleaning:
  - Small ads
  - Classified ads
  - Ads with predominantly text
- Result: 426,777 ads between 1945-1995

# TASKS

- Select image representation in penultimate layer
- Cluster in multidimensional space based on 2,048 (abstract) visual aspects
- Find nearest neighbors in clustered space)
- Time line with most similar images per year

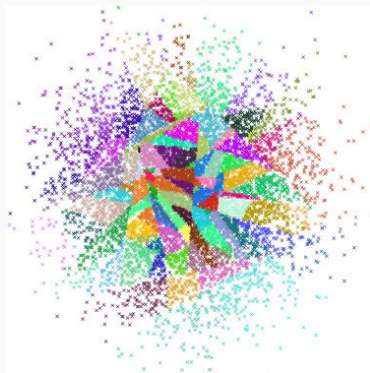


Figure 15: Annoy: Approximate Nearest Neighbor

# SIMILAR ADVERTISEMENT SEARCH



Figure 16: Query cars



Figure 17: Query fashion

- Creating more labeled data
- Retraining last layers of network
- Combining textual and visual analysis
- Seeking collaboration between Computer Science and the Humanities



# Question/Comments

melvinwevers@icloud.com  
<http://www.melvinwevers.nl>