



# Seeing History

## *The Visual Side of the Digital Turn*

@MelvinWevers

@ThomasSmits

Dr. Melvin Wevers - DHLab - KNAW HUC, Amsterdam

Thomas Smits - Radboud University, Nijmegen

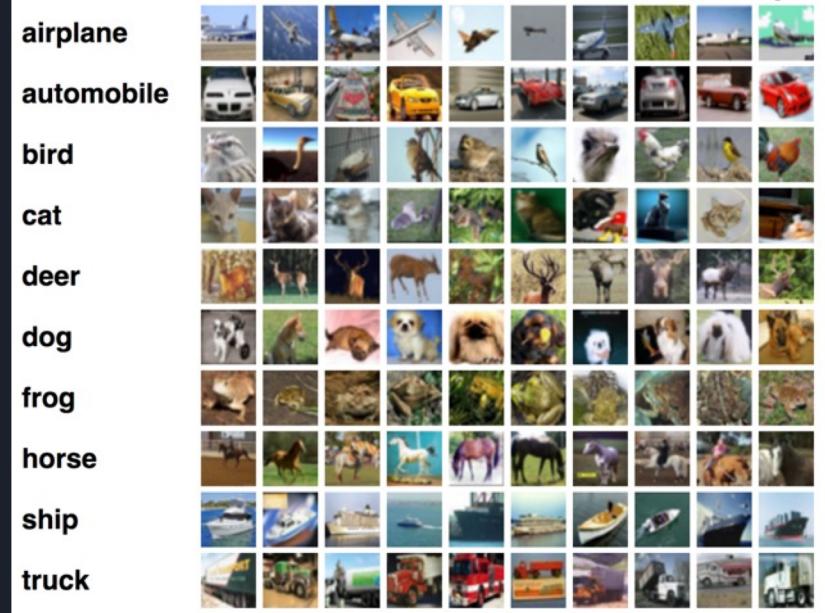
# Introduction

- The Digital Humanities are too text-heavy
- Large collections of visual material, limitations of searching with OCR
- Researcher-in-Residence at National Library of the Netherlands
- How can computers help us to explore and analyze large collection of historical visual material?



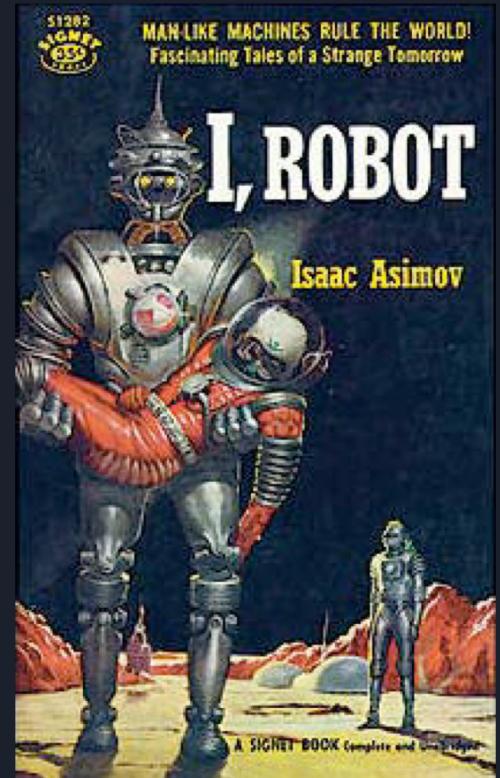
# Computer Vision

- Computer Vision - gain high-level understanding of images
- Object detection
- Convolutional neural networks

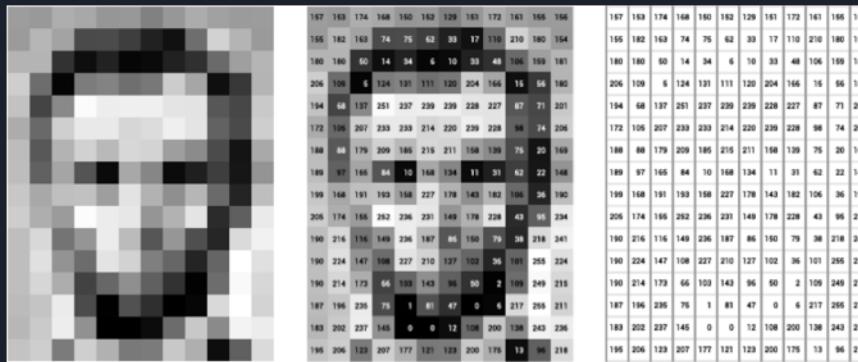


# History of Computer Vision

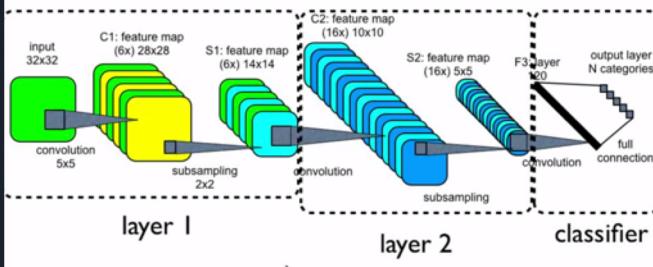
- Goes back to late 1950s.
- Mostly military projects.
- Robotics/Artificial intelligence.
- Facial recognition tasks
- Some of the algorithms (signal processing) were developed earlier in the 1940s For example.



# From an image to a neural network



## Convolutional Neural Networks





# Convolutional neural networks on historical visual material

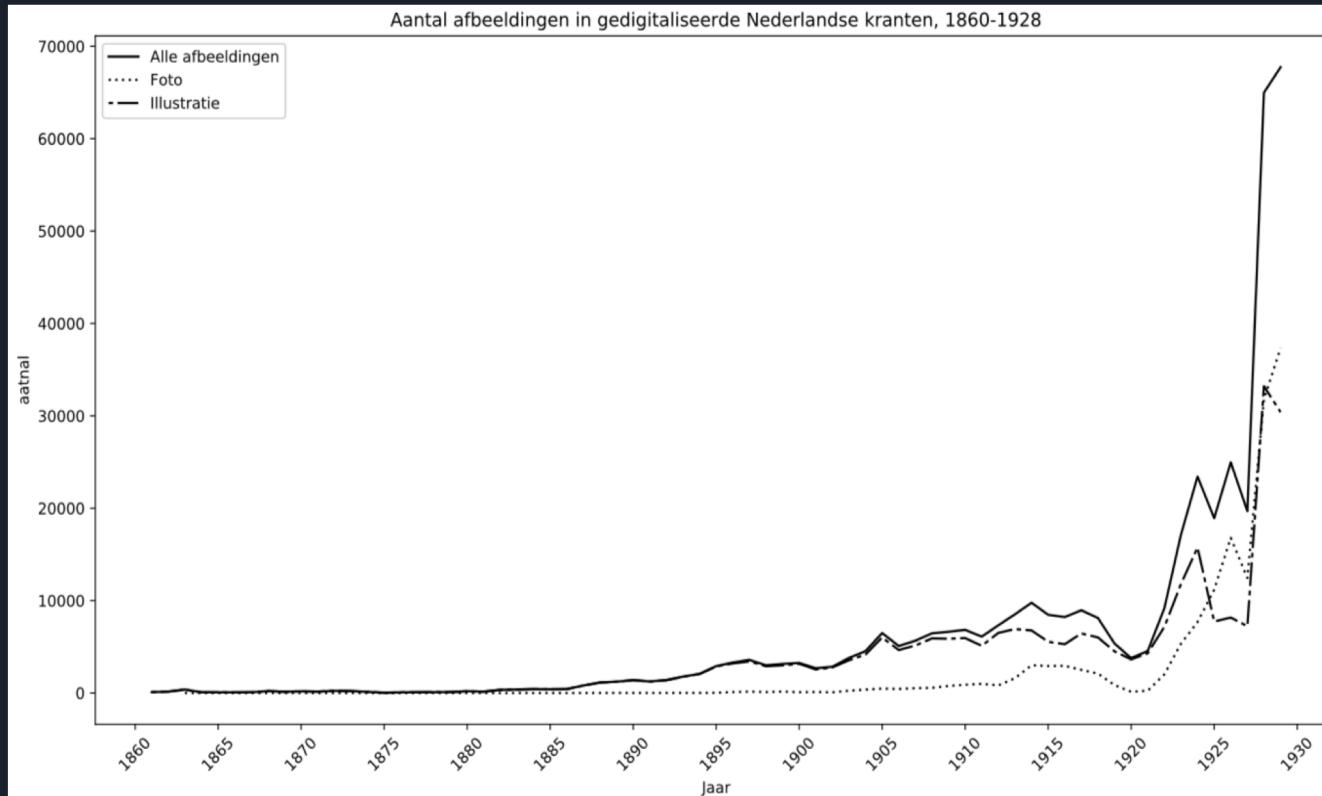
- Two datasets extracted from Delpher
  - CHRONIC (452,543 images of the news 1860-1930)
  - SIAMESET (426,000 historical advertisements 1945-1995)
- Three approaches
  - Detecting medium-specific features (separating photographs from illustrations)
  - querying images based on abstract visual aspects (clustering visually similar advertisements)
  - Training a neural network based on visual categories developed by domain experts

# Approach I: Medium-specific characteristics

- Research the transition between the use of illustrations and photographs by newspapers to visualize the news
- Classify images of CHRONIC as either illustration or photograph (F1-score: 0.9)

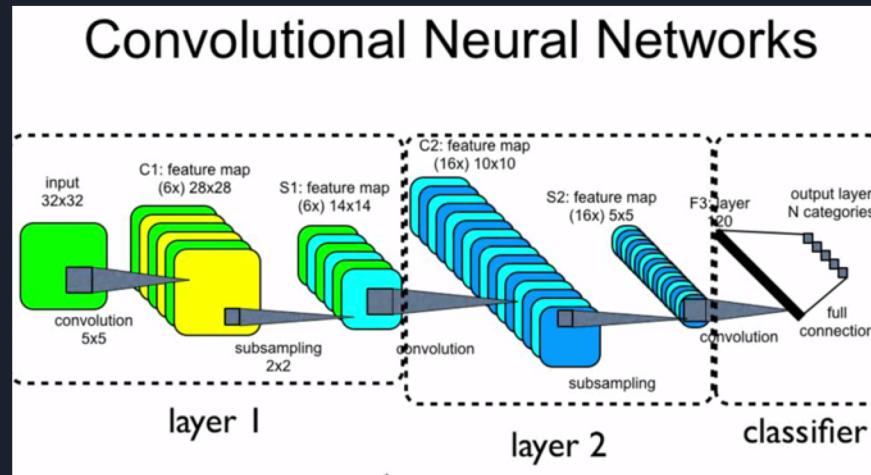


# Approach I: Medium-specific characteristics



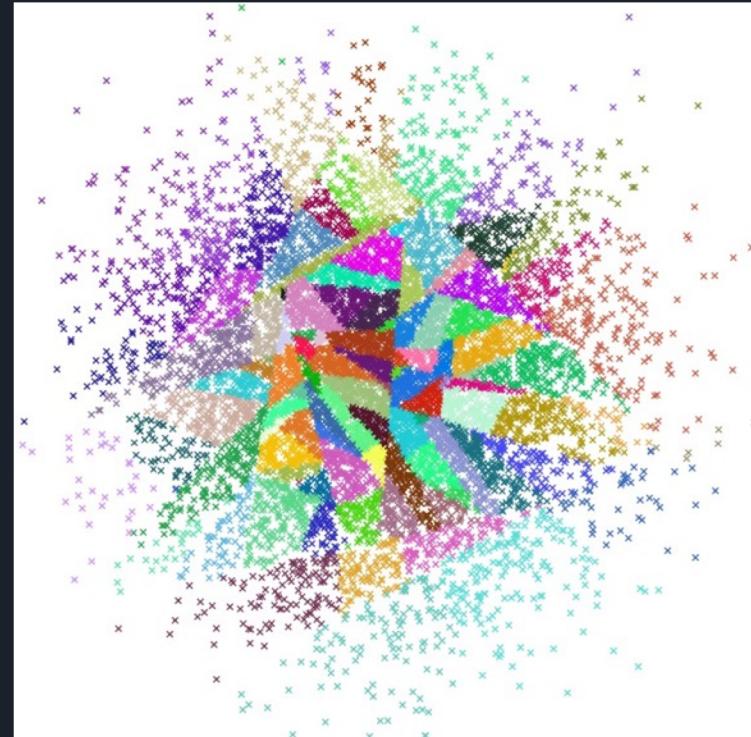
# Approach II: SIAMESE

- Can we use convolutional neural networks to trace visual change in historical advertisements?
- Object detection for historical images is sub-optimal



## Approach II: Cluster on visual similarity

- Select image in penultimate layer
- Cluster in multidimensional space based on 2,048 visual aspects
- Find nearest neighbors in clustered space



# Approach II: Style of advertising



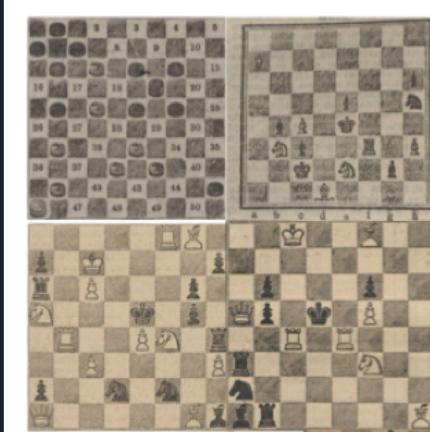
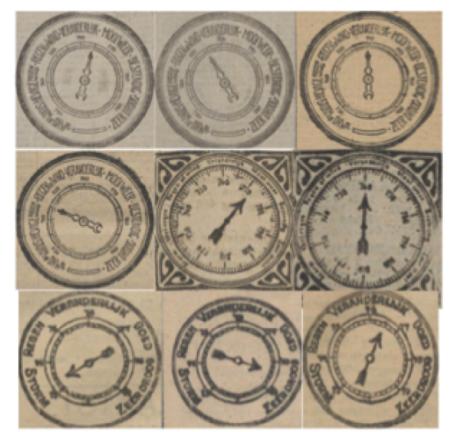
Top ten nearest neighbors



Nearest neighbor examples

# Approach III: Building your own classifiers

- Recognize nine relevant categories: buildings, cartoons, chess, crowds, logos, maps, schematics, sheet music, and weather reports
- Similar to OCR → provides direct access to visual content
- Visual similarity ≠ stylistic similarity ≠ conceptual similarity



# Approach III: Building your own classifiers

The screenshot shows a web browser window titled "KB ImageSearchQueryBuilder" on the "www.kbresearch.nl/query\_builder/" page. The interface includes a date range selector from "1860-01-01" to "1937-01-31", a "Photo or drawing" radio button selected for image type, and an "Any" dropdown for image category. A text input field contains the URL "http://www.kbresearch.nl/xportal?image\_type=\*\*\* and date within \*1860-01-01T00:00:00Z 1937-01-31T00:00:00Z\*". Below the URL is a "go" button and a W3C XHTML 1.0 validation logo.

QuickTime Player Archief Wijzig Weergave Venster Help

KB ImageSearchQueryBuilder

Meest bezocht Digital archives Tools Volgen DH Volgen Gerechten Overig ILLU PM

www.kbresearch.nl/query\_builder/ Zoeken

Image search query builder

Select date-range:

1860-01-01 1937-01-31

Select image type:

Photo  Drawing  Photo or drawing

Select image category:

Any

Text in article:

http://www.kbresearch.nl/xportal?image\_type=\*\*\* and date within \*1860-01-01T00:00:00Z 1937-01-31T00:00:00Z\*

go

W3C XHTML 1.0

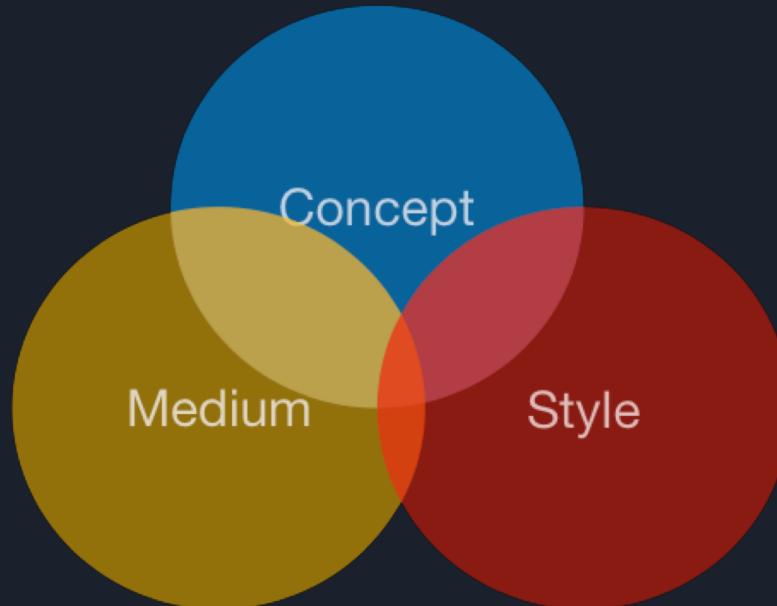


# Conclusion: Opportunities

- CNNs offer opportunities for:
  - collection specialists
  - (digital) humanities researchers
- Explore and analyze large collections of visual sources



## Conclusion: 1) Structures of Visual Similarity



## Conclusion: 2) Historicity of images



# Acknowledgements / Data



Martijn Kleppe



Willem-Jan Faber



Juliette Lonij



Leonardo Impett

**Tools and data for CHRONIC and SIAMESE: <http://lab.kb.nl/>**

@MelvinWevers

@ThomasSmits