Automatic Handwriting Analysis

Vincent Christlein, Andreas Maier

Pattern Recognition Lab, Friedrich-Alexander University of Erlangen-Nuremberg
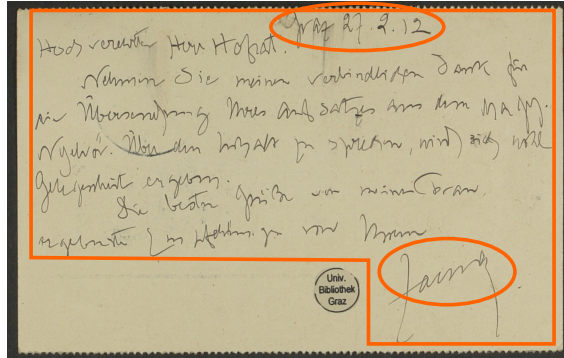
October 30th, 2018

Handwriting Analysis

Who?

Script Type?

Category?

When?

Where?

What?

Handwriting Analysis

Who?

When?

Script Type?

Where?

Category?

What?

Handwriting Analysis

Who?
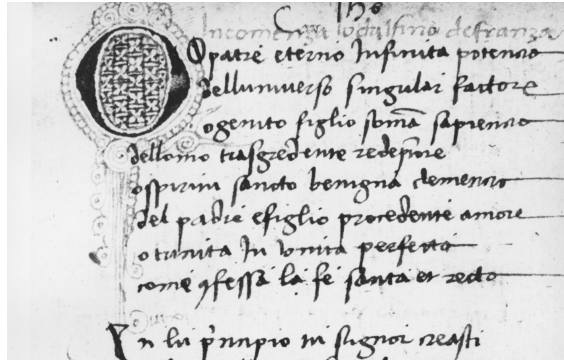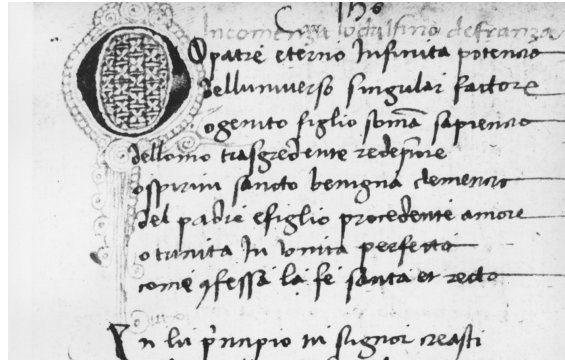
Script Type?

Category?

When?

Where?

What?

$\Rightarrow$ Need expert in history / paleography / forensics!

Handwriting Analysis

Who?

When?

Script Type?

Where?

Category?

What?

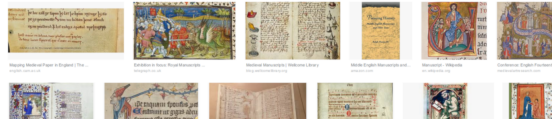$\Rightarrow$ Need expert in history / paleography / forensics!

$\Rightarrow$ Time and cost intensive

$\Rightarrow$ (Semi-)Automatic methods needed

Handwriting Analysis

Who?   When?

Script Type?   **Visual Image Retrieval/Classification**   Where?

Category?   What?

$\Rightarrow$ Need expert in history / paleography / forensics!

$\Rightarrow$ Time and cost intensive

$\Rightarrow$ (Semi-)Automatic methods needed

State of the Art in Writer Identification I

- ICDAR'13 benchmark dataset: 250 writers, 4 samples each
- Top 1: 99.6% accuracy
- All 3 among first 3 hits: 77% accuracy

V. Christlein and A. Maier, "Encoding CNN Activations for Writer Recognition", in *DAS*, 2018

State of the Art in Writer Identification II

- ICDAR'17 Historical-WI: 720 writers, 5 samples each
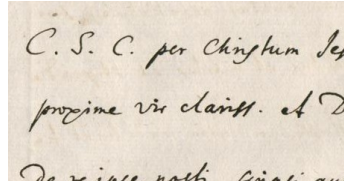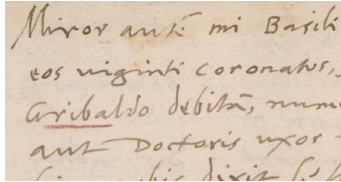
Legend: Christlein '17 | Christlein '15 | Mohammed '18 | Fiel '13 | ICDAR'17 Competition

Top-1 [%]

88.9 | 86.0 | 85.6 | 81.4 | 76.4 | 76.1 | 74.4 | 67.1 | 67.0 | 47.8
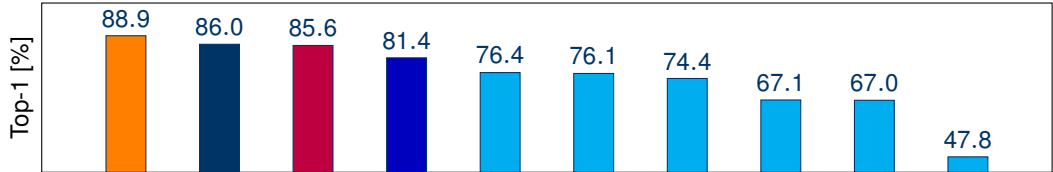
V. Christlein, M. Gropp, S. Fiel, *et al.*, "Unsupervised Feature Learning for Writer Identification and Writer Retrieval", in *ICDAR*, 2017

Classification of Latin Medieval Manuscripts (CLaMM)

- 2000 images, 12 script types (Uncial, Praegothica, Cursiva, . . .)
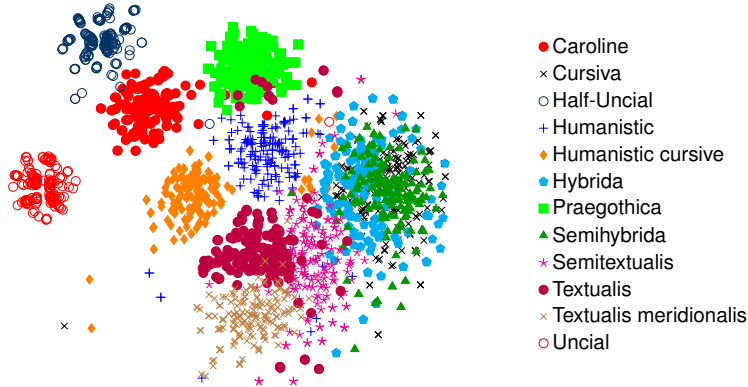
F. Cloppet, V. Églin, V. C. Kieu, *et al.*, "ICFHR2016 Competition on the Classification of Medieval Handwritings in Latin Script", in *ICFHR*, 2016

Script Type Classification

- Caroline
- Cursiva
- Half-Uncial
- Humanistic
- Humanistic cursive
- Hybrida
- Praegothica
- Semihybrida
- Semitextualis
- Textualis
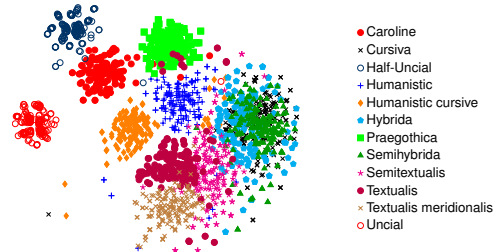- Textualis meridionalis
- Uncial

- Supervised: 85% accuracy

- Unsupervised: Model selection for OCR/HTR

M. Kestemont, V. Christlein, and D. Stutzmann, "Artificial Paleography: Computational Approaches to Identifying Script Types in Medieval Manuscripts", *Speculum*, vol. 92, 2017

Summary

Visual image retrieval/classification

- Large-scale analysis
- Meta-data acquisition
- Assisting the expert
- Model selection for OCR/HTR

- Caroline
- Cursiva
- Half-Uncial
- Humanistic
- Humanistic cursive
- Hybrida
- Praegothica
- Semihybrida
- Semitextualis
- Textualis
- Textualis meridionalis
- Uncial

Thank you for your attention!

References

References I

[1] F. Cloppet, V. Églin, V. C. Kieu, D. Stutzmann, and N. Vincent, "ICFHR2016 Competition on the Classification of Medieval Handwritings in Latin Script", in *ICFHR*, 2016.

[2] V. Christlein and A. Maier, "Encoding CNN Activations for Writer Recognition", in *DAS*, 2018.

[3] V. Christlein, M. Gropp, S. Fiel, and A. Maier, "Unsupervised Feature Learning for Writer Identification and Writer Retrieval", in *ICDAR*, 2017.

[4] M. Kestemont, V. Christlein, and D. Stutzmann, "Artificial Paleography: Computational Approaches to Identifying Script Types in Medieval Manuscripts", *Speculum*, vol. 92, 2017.