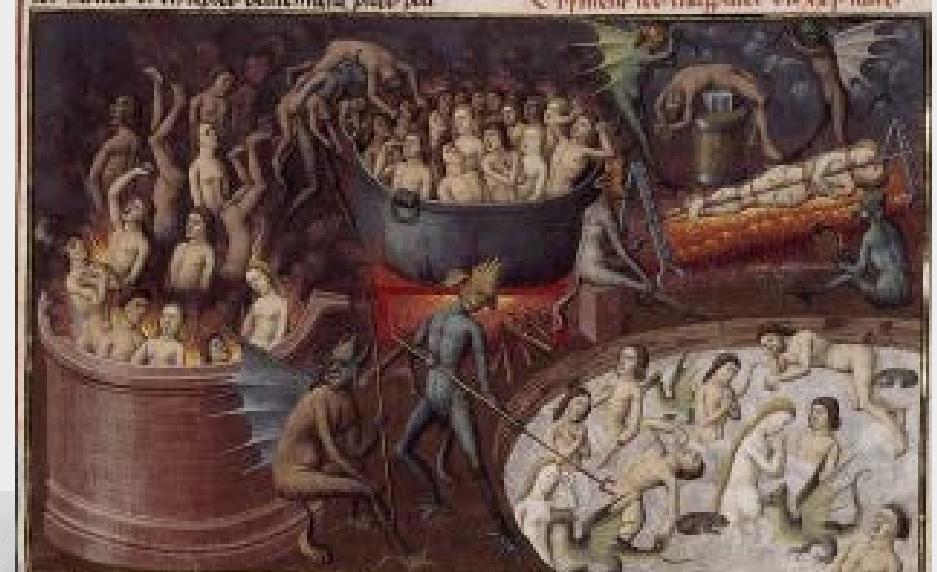
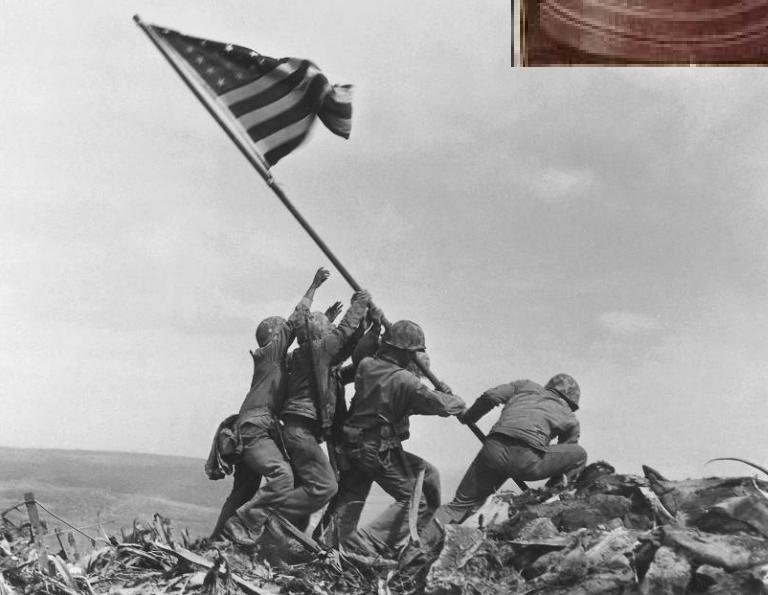




Machine Learning and History

exemplified by Image Analysis
of Heraldry

History. History? - The past ...



- political history
- cultural history
- social history
- economic history
- religious history
- history of mentality
- ...

History. History? – The questions ...

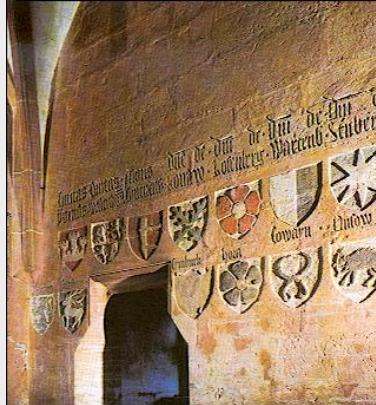
Understanding the past ... to better understand the present.

Not: how it was,

But: different interpretations and reconstructions of the past.

History. History? – The methods ...

- Historical critical method to interpret sources from the past.



Eis
Käse
Rhabarber
1 Bund Basilikum
1 reife Avocado
Pinienerne
1 Limette
Tomatenmark
Olivenöl
Toast
Aufbackbrötchen

History. History? - Problems ...

Vagueness / Uncertainty

- Incompleteness
 - ambigu: reading, identification etc. depending on the observer and his knowledge and presumptions
 - unstable: meaning may change depending on time and/or place
 - multi-perspective: statements instead of facts
- = Often not possible to decide.



Machine Learning in Historical Studies

The application of machine learning is still in its infancy:

- Computer vision
 - HTR / OCR
 - Digital Paleography (to analyze historical handwritings)
 - Object detection
- NLP – Areas of application
 - Named Entity Recognition
 - Topic Modelling
 - Stylometry
 - Text reuse detection

Examples

The screenshot shows the Transkribus software interface. On the left, there is a sidebar with various tools and options:

- Layout Analysis:** Set to "Ottob Advanced". Options include "Current page" (radio button selected), "Pages (0): 1-20", and "Only use uncorrected pages". Buttons for "Run" and "Configure..." are present.
- Text Recognition:** Set to "HTR (C)Tab HTR + PyTorch". Options include "Models...", "Train...", and "Run...".
- Compute Accuracy:** Reference: "17.02.20 15:34:28 - hannah.royd@mail.utoronto.ca - Ground Truth" and "Use current". Hypothesis: "17.02.20 15:34:28 - hannah.royd@mail.utoronto.ca - Ground Truth" and "Use current". Buttons for "Compare Text Versions..." and "Compare Samples..." are shown.
- Other Tools:** Includes "POfLA...", "Text2Image...", and "Document Understanding...".

On the right, the main window displays a handwritten document in Latin. The text is written in cursive Gothic script. A green rectangular box highlights the title "Beklungthorpe". Below the text, a transcription is provided with numbered footnotes at the bottom:

Beklungthorpe

clero de Othulveshoo pro homagio et servicio suo et pro triginta solidis argenti quos
nob̄ dedit in gersumam totam illam terram cum omnibus pertinentiis suis habaturis ibi plus
sive minus quam habuimus vel aliquo modo habere potuimus in cruta que vocatur
Holdelinecroft quam quidem crutam Iohannes Makewombe aliquando tenuit
in villa de Gestlinghorpe et totum pratum sive sit plus sive minus quod habuimus
vel nob̄ aliquo modo accidere potuit in prato de Oddewelle de prato quod quondam
medleyn cum omnib̄ p̄t̄ suis p̄p̄e domiñ Ric⁹ Binsayl. clero
p̄t̄ suo donaþ ammi redditus p̄cipionali de p̄do Ric⁹ Binsayl.
Innotescit p̄ quadam p̄t̄ p̄da p̄t̄ quam reginaldus p̄kelemeret
sic. Hend⁹ e tenend⁹ de nob̄ e hereditib⁹ nro vel de nro assign⁹
redib⁹ suis vel eni vel quando e m̄ quocunq; etiam fit oam
terram cum p̄t̄ videlic⁹ enim inveniat⁹. Engilas cepit⁹ fessatio
terre tse expectantib⁹ et dem p̄enni enim p̄t̄ suis omnib⁹ dage ven
sice legare vel assignare voluit exceptio donati iehanosis et hideris. libe amete

Footnotes:

- 1-1 clero de Othulveshoo pro homagio et servicio suo et pro triginta solidis argenti quos
- 1-2 nob̄ dedit in gersumam totam illam terram cum omnibus pertinentiis suis habaturis ibi plus
- 1-3 sive minus quam habuimus vel aliquo modo habere potuimus in cruta que vocatur
- 1-4 Holdelinecroft quam quidem crutam Iohannes Makewombe aliquando tenuit
- 1-5 in villa de Gestlinghorpe et totum pratum sive sit plus sive minus quod habuimus
- 1-6 vel nob̄ aliquo modo accidere potuit in prato de Oddewelle de prato quod quondam

Transkribus

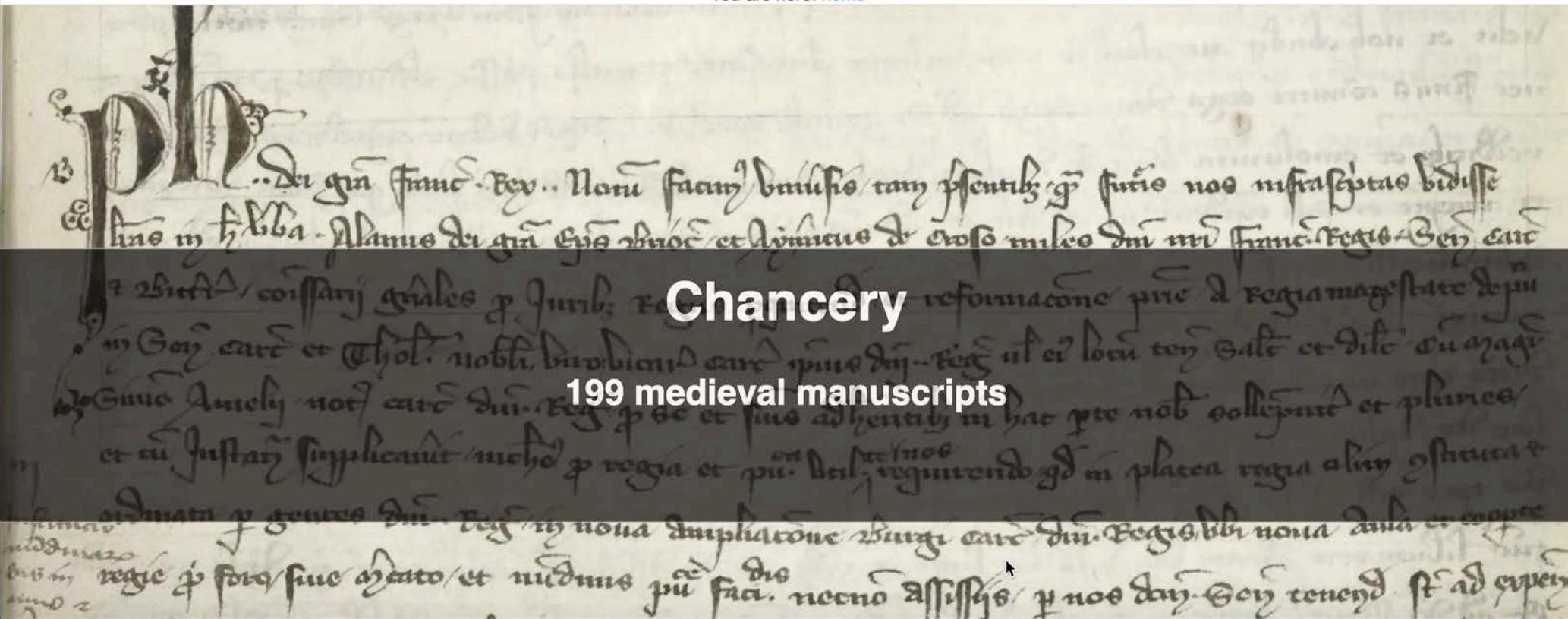
Examples

HIMANIS PRHLT search engine **BETA**

Search options Indexing details

Confidence: Max. results:

You are here: [home](#)



The image shows a page from a medieval manuscript. At the top left is a large, ornate initial 'P'. The text is written in two columns in a Gothic script. A prominent watermark or overlay text reads "Chancery reformacione prie a regia magestare depu". Below this, in a larger, bold font, is the text "199 medieval manuscripts".

[Example:] Image Analysis of Heraldry



Prof. Dr. Torsten Hiltmann
Digital History
Humboldt-Universität zu Berlin



Philipp Schneider, M.A.
Digital History
Humboldt-Universität zu Berlin



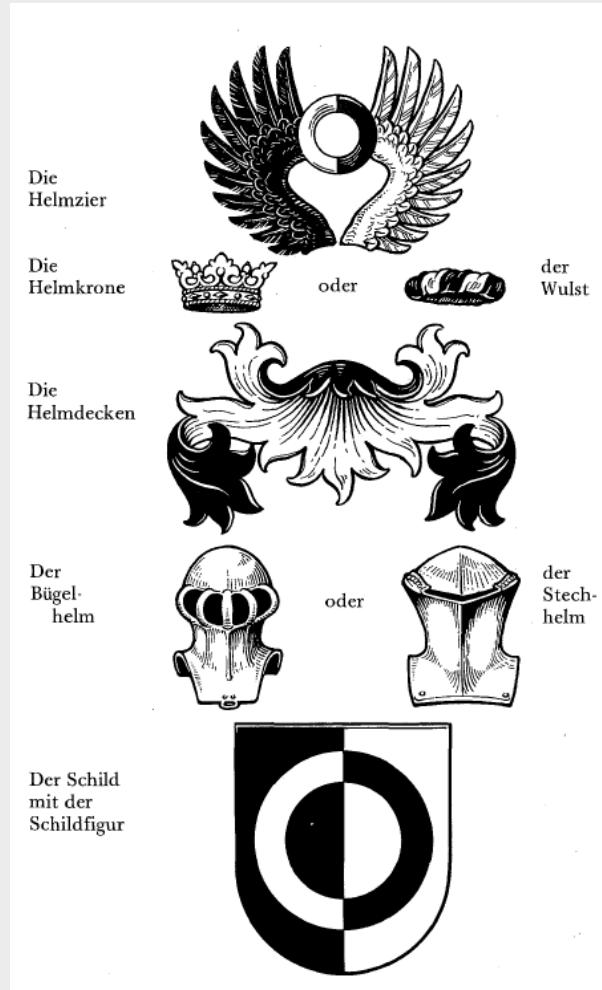
Jun.-Prof. Dr. Benjamin Risse
Computer Vision and Machine Learning
University of Münster



Sebastian Thiele, M.A.
Computer Vision and Machine Learning
University of Münster



[Example:] Image Analysis of Heraldry



Coat of arms



Heraldry and coats of arms today ...



Bundesministerium
für Bildung
und Forschung



BERLIN



Heraldry and coats of arms today and in the past



Federal President Steinmeier (2017)



Heinrich VI., Roman King and Emperor
(Codex Manesse, ca. 1300)

Heraldry and coats of arms today and in the past

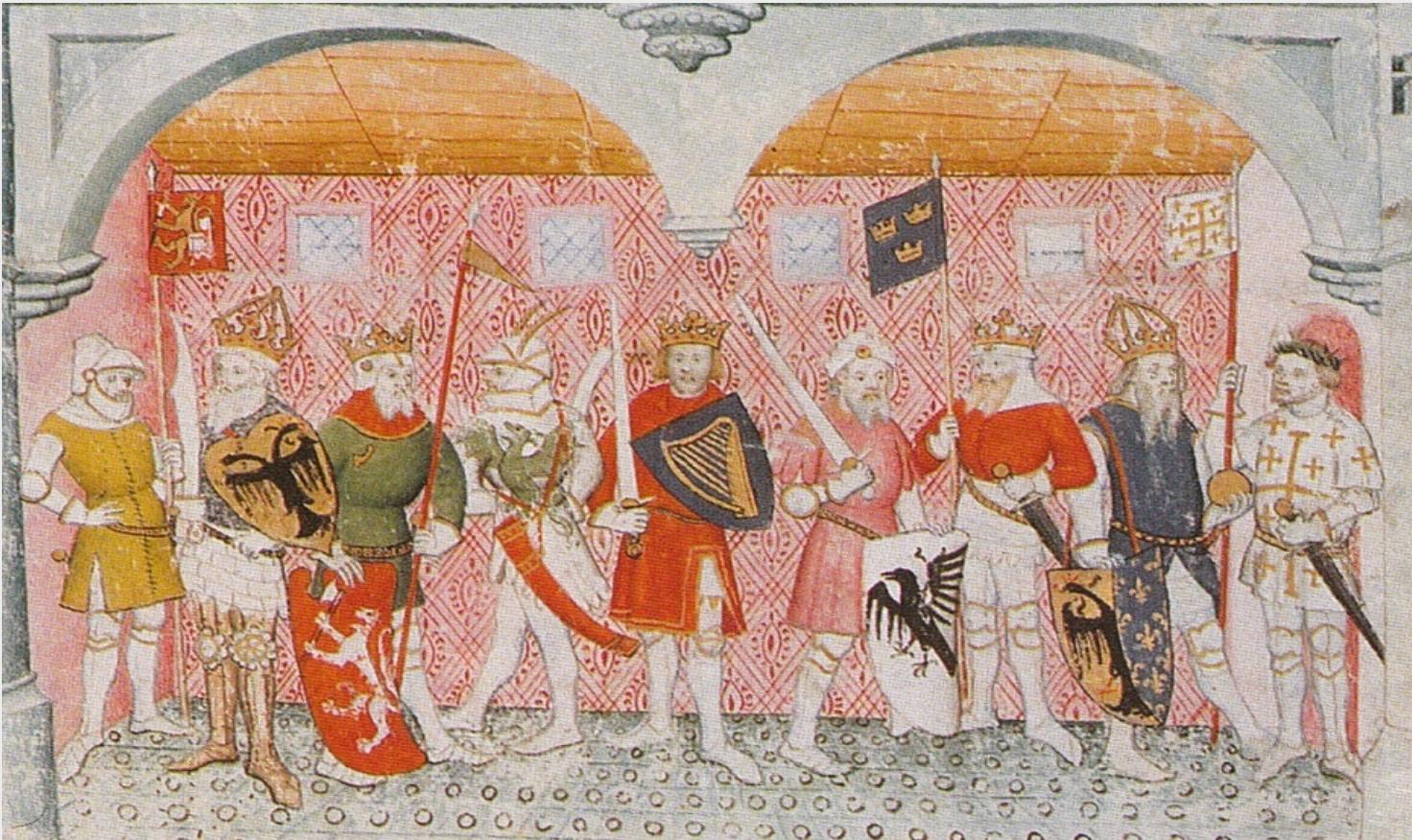


Coat of arms in the Middle Ages



early tournament depiction from the "Queste del saint graal",
Paris, BnF, fr. 342, fol. 102v (Artois, 13th c.)

Extension of use



Depiction of the Nine Worthies (Hector of Troy, Julius Caesar, Alexander the Great; Judas Maccabee, King David, Joshua ; King Arthur, Charlemagne, Gottfried von Bouillon)

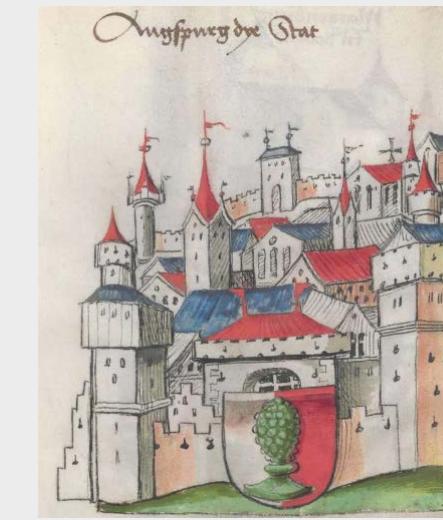
France, 1404



Peasant seal
(Normandy)



Carpenters' Guild
in Ghent



The city of Augsburg



Augsburg guildsmen

Extension of use



In the Middle Ages and early modern period, coats of arms could be depicted in almost all existing techniques on every conceivable medium, in the most private as well as in the most public space.

Significance for historical research



Ownership



Memoria



Origin,
dominion, kinship, status,
claims, ...



Abstract concepts (sins, death, Trinity)



Identity



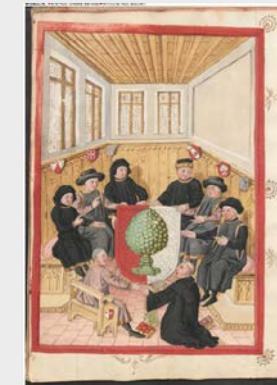
... expanded e.g. by
stories of origin



Political concepts



Historiographical concepts



Pragmatics - as part of actions



Quelle chose, lecteur ! mais véritablement digne d'un homme tel qu'il est.

State of the art and challenges

Coats of arms are a central means of communication in the Middle Ages and early modern period, the analysis of which allows extensive insights into pre-modern culture and society

Current state of research:

- Hardly researched so far
- Above all, there is a lack of comprehensive presentations and analyses.

Three (technical) challenges

1. Sheer quantity of the surviving sources
2. Heterogeneity of contexts of use and transmission
3. Complexity of the subject as such

Evidence (mentions in metadata):

- Clemensens Ordinary of Medieval Armorial: 80,000 coat of arms descriptions (from manuscripts)
- Bildindex Kunst und Architektur: 38,000 works
- Object catalogue of the GNM: 5650 objects
- Bibliothèque nationale de France (BnF): 800 Hss.
- Siebmacher: 130,000 entries

→ Computer vision for the acquisition
and analysis of visual sources

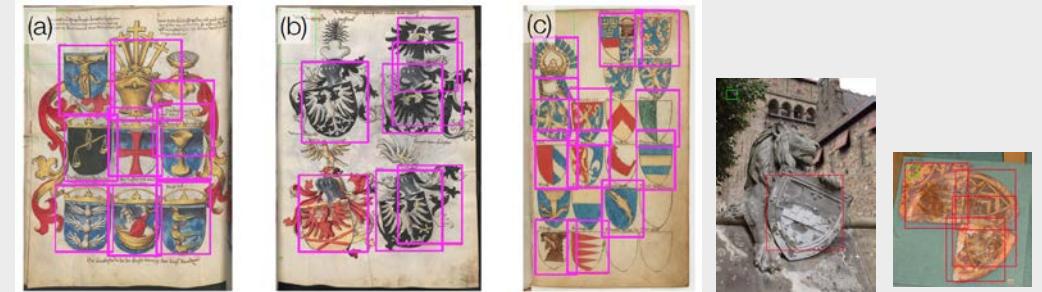
Different tasks for computer vision

1. Detection

Automatic annotation of heraldic representations in large image collections



Collection
of data



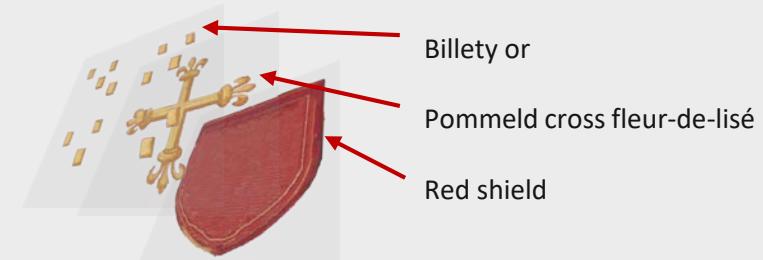
Different tasks for computer vision

2. Semantic segmentation

Division of the heraldic representation into its individual heraldic elements, for each pixel



Training with
existing data

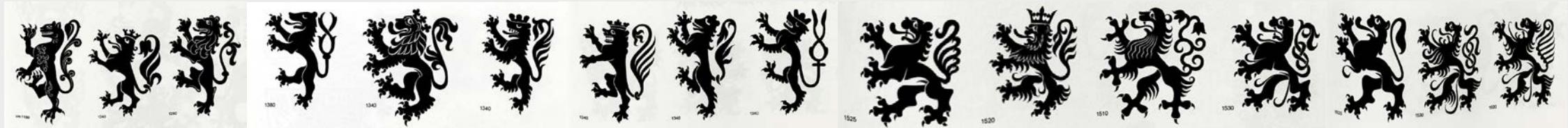


Different tasks for computer vision

3a Supervised classification : Assigning the elements in unknown coat of arms to the individual element groups and thus contribute to identify unknown coats of arms

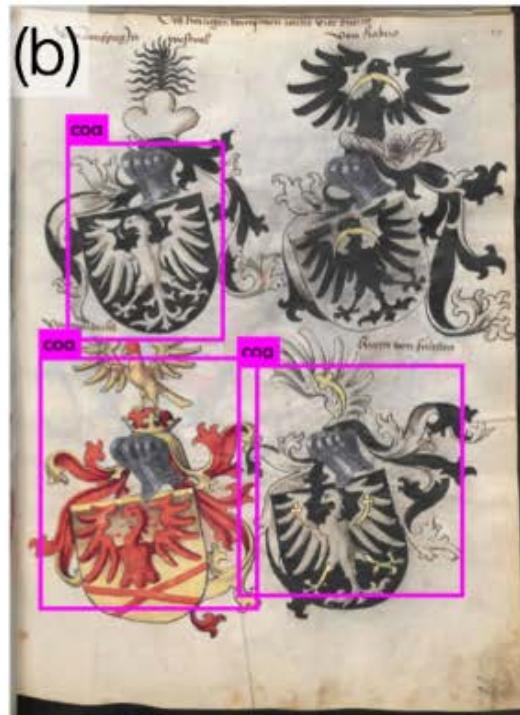
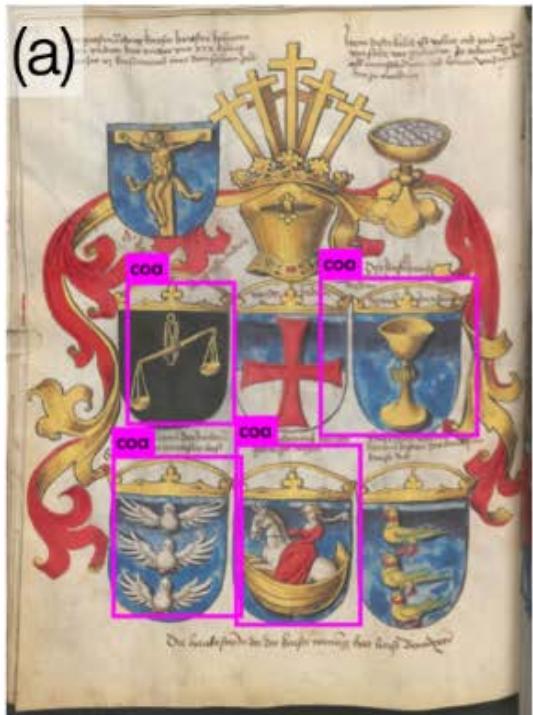
3b Nearest Neighbour: Find representations that are particularly similar to a new, previously unknown representation

3c Unsupervised classification: clustering of the recorded heraldic elements or coats of arms according to certain features in order to form (stylistic) groups and thus historical, i.e. spatial and temporal structures in the recorded representations and thus make e.g. stylistic developments, but also transfer processes, verifiable



Detection - Comparison of different network architectures (2019)

YOLO Ergebnisse



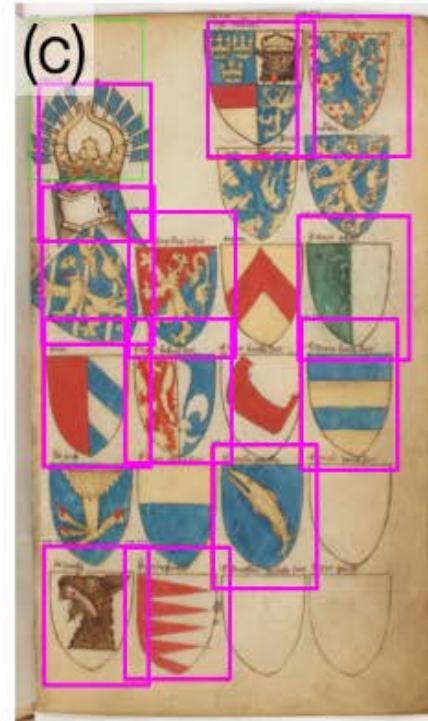
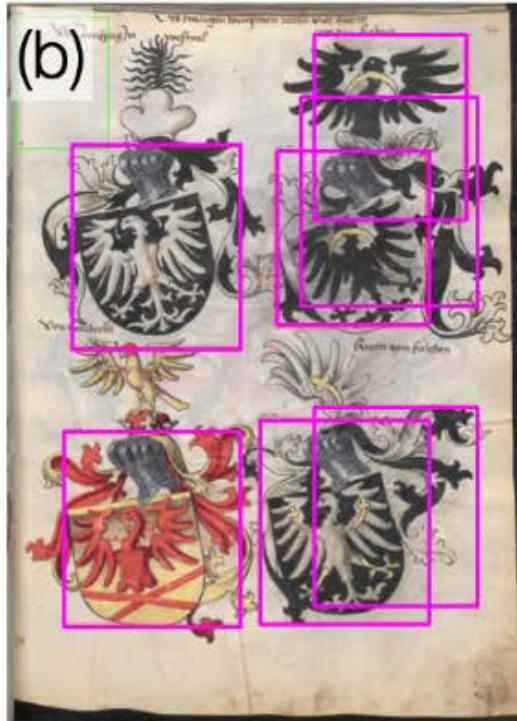
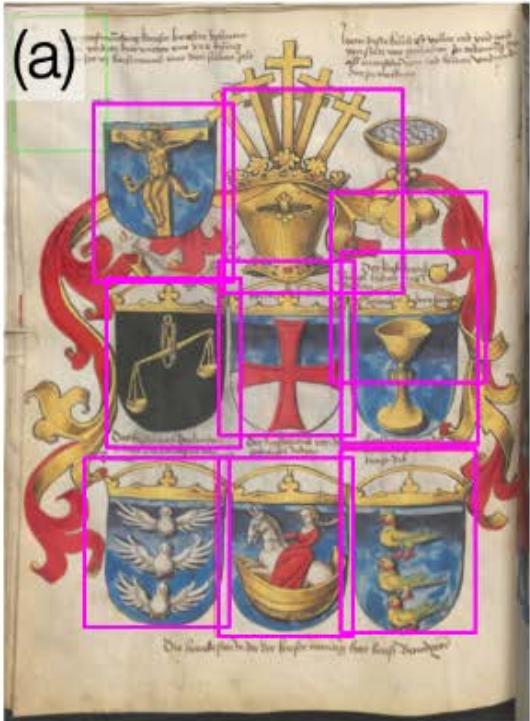
Precision: high

Recall: low

YOLO Detektionsergebnisse. **(a)** Validation Datensatz mit größter Ähnlichkeit zu den Trainingsdaten. **(b)** Datensatz mit ähnlicher Wappendichte und Wappengröße, jedoch erhöhtem Detailgrad. **(c)** Datensatz mit höherer Wappendichte und rotierten Wappen. YOLO zeigt grundsätzlich eine hohe Spezifität (erkannte Wappen sind tatsächlich Wappen), jedoch eine geringe Sensitivität (viele Wappen werden nicht erkannt).

Detection - Comparison of different network architectures (2019)

RetinaNet Ergebnisse



Precision: low

Recall: high

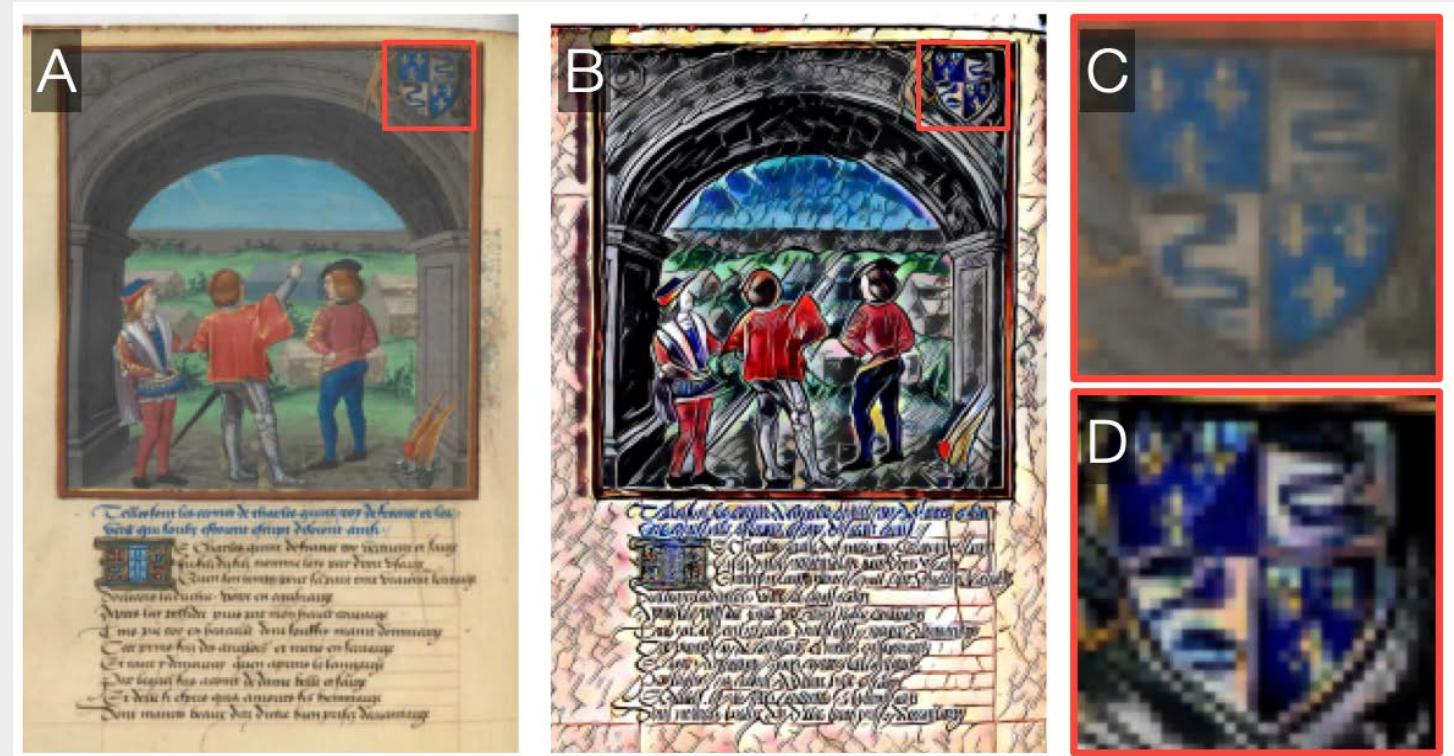
RetinaNet Detektionsergebnisse. (a) - (c) Datensätze analog zu YOLO Ergebnissen. Im Unterschied zu YOLO zeigt RetinaNet eine geringere Spezifität (nicht alle erkannten Wappen sind tatsächliche Wappen), weist dafür jedoch eine erhöhte Sensitivität auf (die meisten Wappen werden erkannt).

Problem: training data

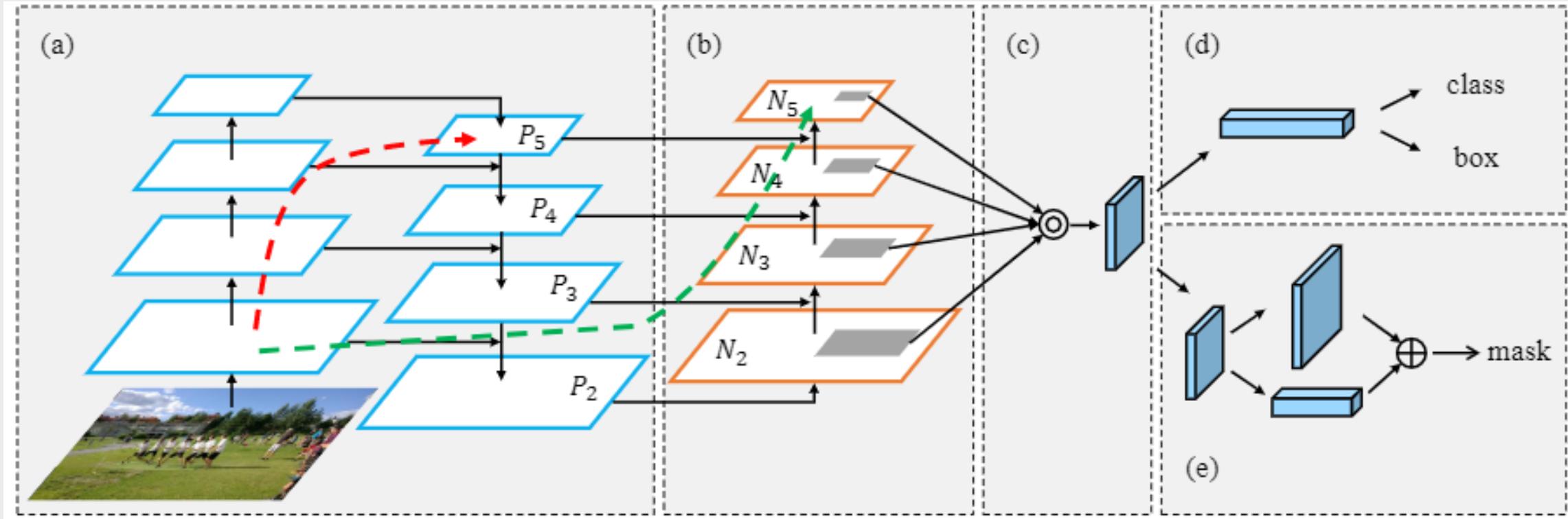
Individual labelling necessary

Data Augmentation

Iterative expansion and improvement of the training data



New Version: Yolo 4



Stand: 9.234 labeled coats of arms

Precision: 0.89

Recall: 0.80

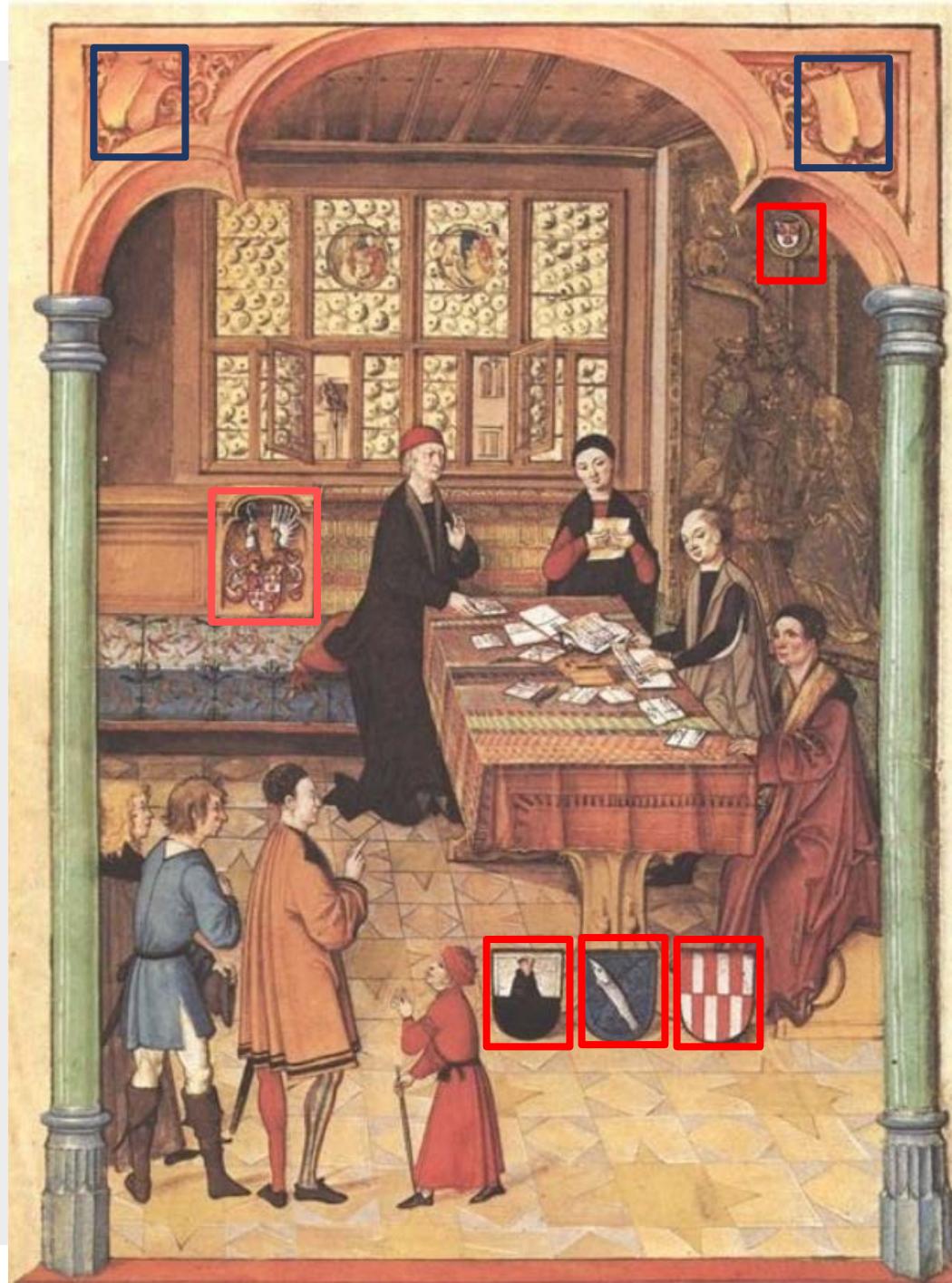
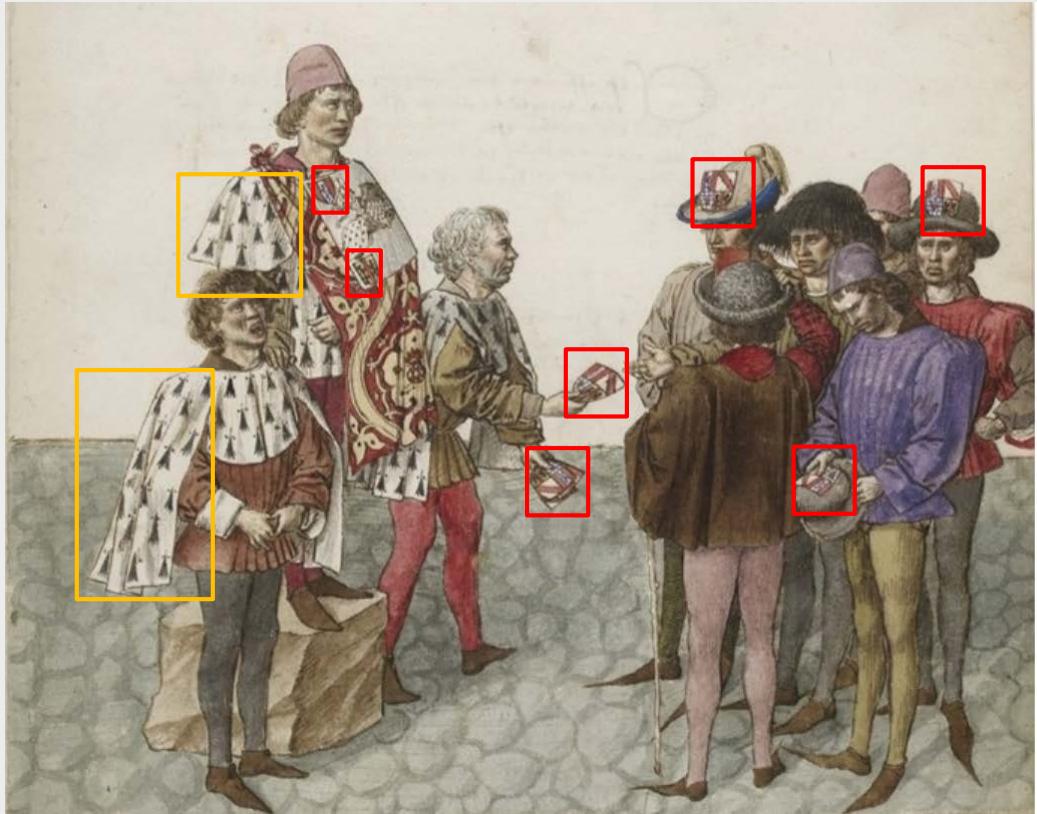
F1 score 0.84



still room for improvement!

interesting extensions

More complex depictions



interesting extension

To go beyond depictions in manuscripts



Current state of the project: segmentation

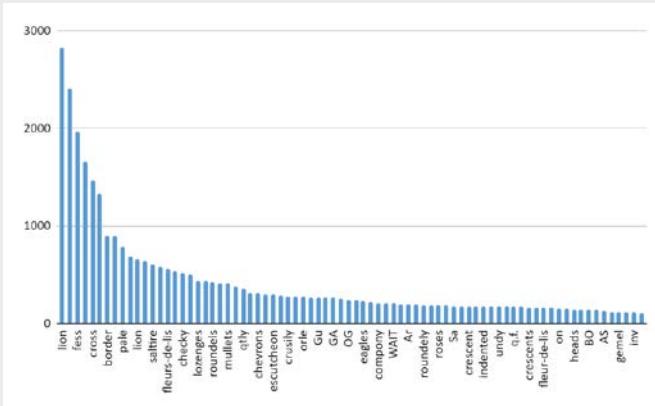


Training on
existing data



- Assignment of the individual pixels must normally be done by hand.
- However, there are now also automatic procedures if the data have already been labelled.

Problem: Imbalance / Long tail

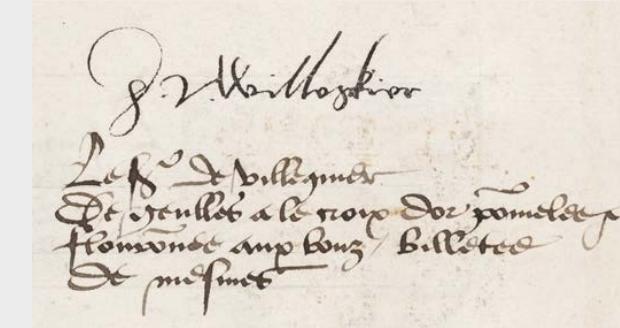


Number of heraldic elements in approx. 30,000 heraldic descriptions

3480 lion
2653 bend
1964 fess
1660 chief
1467 cross
from no. 7 under 1000
from no. 20 under 500
from no. 50 under 100

- Develop new procedures to deal with this problem

Coat of arms as abstract image



Le seigneur de Villequier
De gueules a le croix d'or pommelee
fleuronnee aux bouz billetee
et mesmes

Other possible applications of ML-Methods

Detection and identification of heraldic mentions in texts

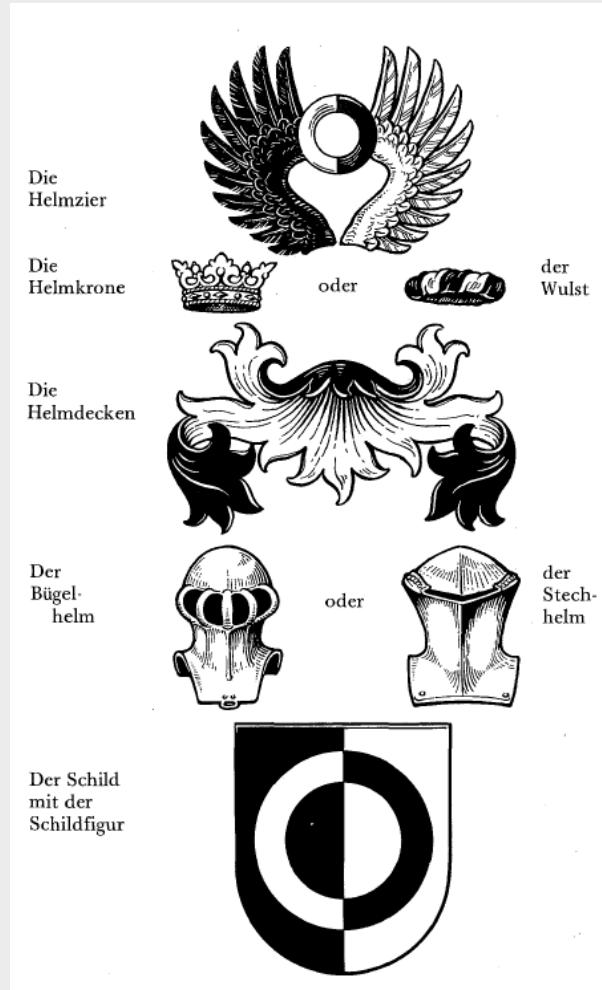
3902	igure et représentacion du conte de Rostelant 1 , aré de ses armes, et uog gi Enguerran de Monstrelet/Chronique de Monstrelet [01].pdf.txt	1
4286	comme les autres, et estaient tous leurs chevaux cons de blans drap de soye Enguerran de Monstrelet/Chronique de Monstrelet [01].pdf.txt	1
23272	et amour mutuelle telle comme scevent tous nobles. Et de rechef, pour plus Enguerran de Monstrelet/Chronique de Monstrelet [02].pdf.txt	1
23273	Et de rechef, pour plus grande confirmation de lae fraternité et société d'arm Enguerran de Monstrelet/Chronique de Monstrelet [02].pdf.txt	1
24061	.. partenir en aucune manière , estre aussi réputez indi- ·i%""· gnes de sa su Enguerran de Monstrelet/Chronique de Monstrelet [02].pdf.txt	1
24062	·i%""· gnes de sa succession , de son nom, de ses armes et de sa seigneur Enguerran de Monstrelet/Chronique de Monstrelet [02].pdf.txt	1
30015	quant il vivait, il donna à l'église de Nostre-Dame de Paris celle grande cloch Enguerran de Monstrelet/Chronique de Monstrelet [02].pdf.txt	1
30016	Paris celle grande cloche laquelle il fit nommer Ja, comme il appert par ses a Enguerran de Monstrelet/Chronique de Monstrelet [02].pdf.txt	1
48567	fu t baillé ung seel, ouquel estoit gravé l'ymage de la Royne , estant droict Enguerran de Monstrelet/Chronique de Monstrelet [03].pdf.txt	1
48568	Royne , estant droict , aians les deux bras teudans vers terre, et au costé d Enguerran de Monstrelet/Chronique de Monstrelet [03].pdf.txt	1
48569	vers terre, et au costé dextre estoit un escu des armes de France, .et au sen Enguerran de Monstrelet/Chronique de Monstrelet [03].pdf.txt	1
61875	des quatre qui menoient son dit chariot ouquel leroy estoit, avoit un calier où Enguerran de Monstrelet/Chronique de Monstrelet [04].pdf.txt	1
61876	estoit, avoit un calier où estoient peintes les anciennes armes d'Angleterre. A Enguerran de Monstrelet/Chronique de Monstrelet [04].pdf.txt	1
69147	la place, est assavoir sur le marché de BruxelJes. Et estoit grand noblesse d Enguerran de Monstrelet/Chronique de Monstrelet [04].pdf.txt	1
69148	Et estoit grand noblesse de les veoir. Car ils estoient tous moult richement h Enguerran de Monstrelet/Chronique de Monstrelet [04].pdf.txt	1
73988	messire Pierre Quiéret, seigneur de Ramencourt, avec pluiseurs aultres geo Enguerran de Monstrelet/Chronique de Monstrelet [04].pdf.txt	1
73989	pluiseurs aultres geotilz hommes, ses parens et amis, et séoit sur ung cheva Enguerran de Monstrelet/Chronique de Monstrelet [04].pdf.txt	1
74052	ijz en auroient le commandement du juge. Et leurs chevaux, qui estoient co Enguerran de Monstrelet/Chronique de Monstrelet [04].pdf.txt	1

Coat of arms in French: *armes*

Other expressions with *armes*: gens d'armes, armes de trait, hommes d'armes, prendre ses armes,

Great difficulty to produce training data, the different mentions require a lot of interpretation by the scholar himself.

[Example:] Image Analysis of Heraldry



Coat of arms

