Transparent Similarity Estimation of Medieval Pen Flourishing via Local Visual Patterns

Florian Kibler[1], *Monica Appelaniz-Portos, Max Theisen*[2], *Victor Adriel de Jesus Oliveira*[1], *Martin Haltrich*[2], *Matthias Zeppelzauer*[1], *Markus Seidl*[1]

[1] St. Pölten University of Applied Sciences

[2] Center for Cultural Heritage Studies, Klosterneuburg Abbey

Source: Klosterneuburg Abbey

Team

St. Pölten University of Applied Sciences

Florian Kibler
Computer Vision

Matthias Zeppelzauer
Computer Vision

Monica Appelaniz-Portos
Computer Vision

Markus Seidl
Computer Vision

Victor Adriel de Jesus Oliveira
Usability

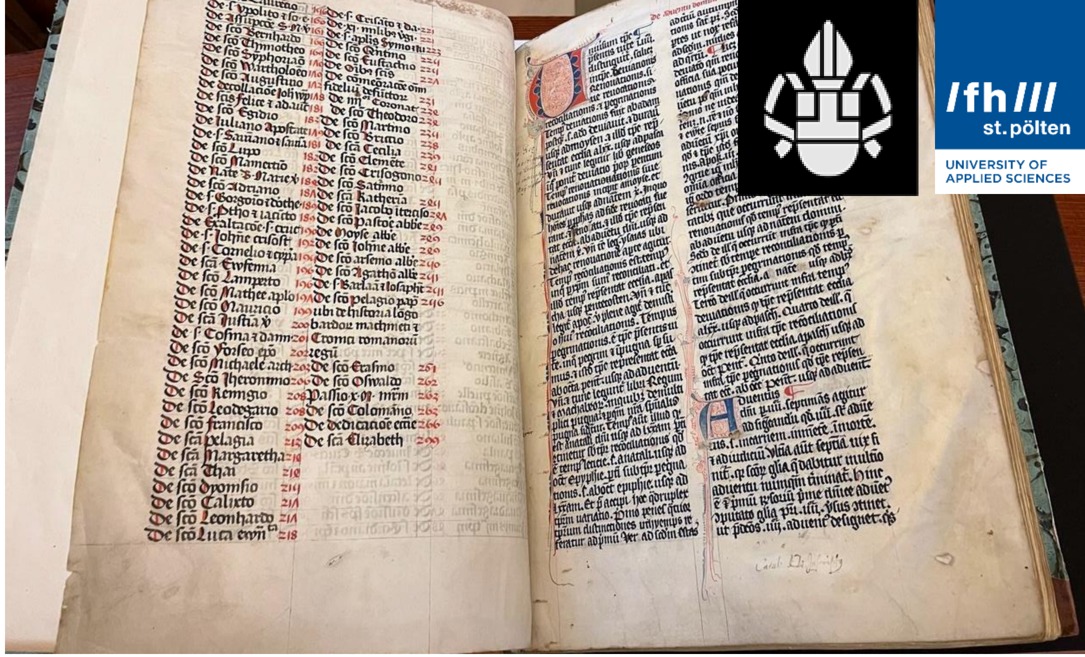
Center for Cultural Heritage Studies Klosterneuburg Abbey

Max Theisen
Cultural Heritage

Martin Haltrich
Cultural Heritage

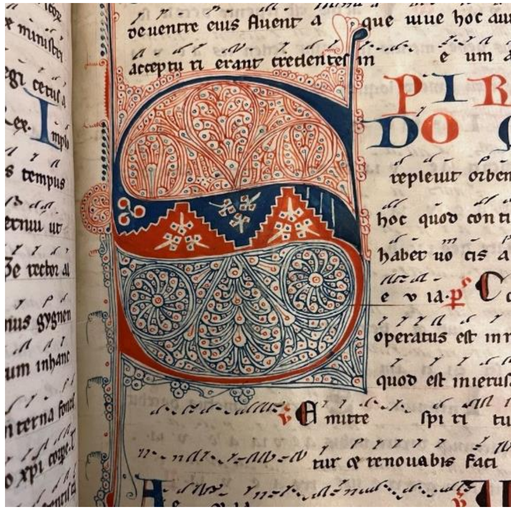
Medieval Manuscripts and Decoration

- Medieval manuscripts are central cultural assets that not only transmit texts but also communicate through their visual design [1,2].
- Decorations acted as navigational and aesthetic markers, guiding reading and emphasizing hierarchy or meaning [1,2].
- Decoration included *initials*, *miniatures*, or ***pen flourishing***.



Source: Klosterneuburg Abbey

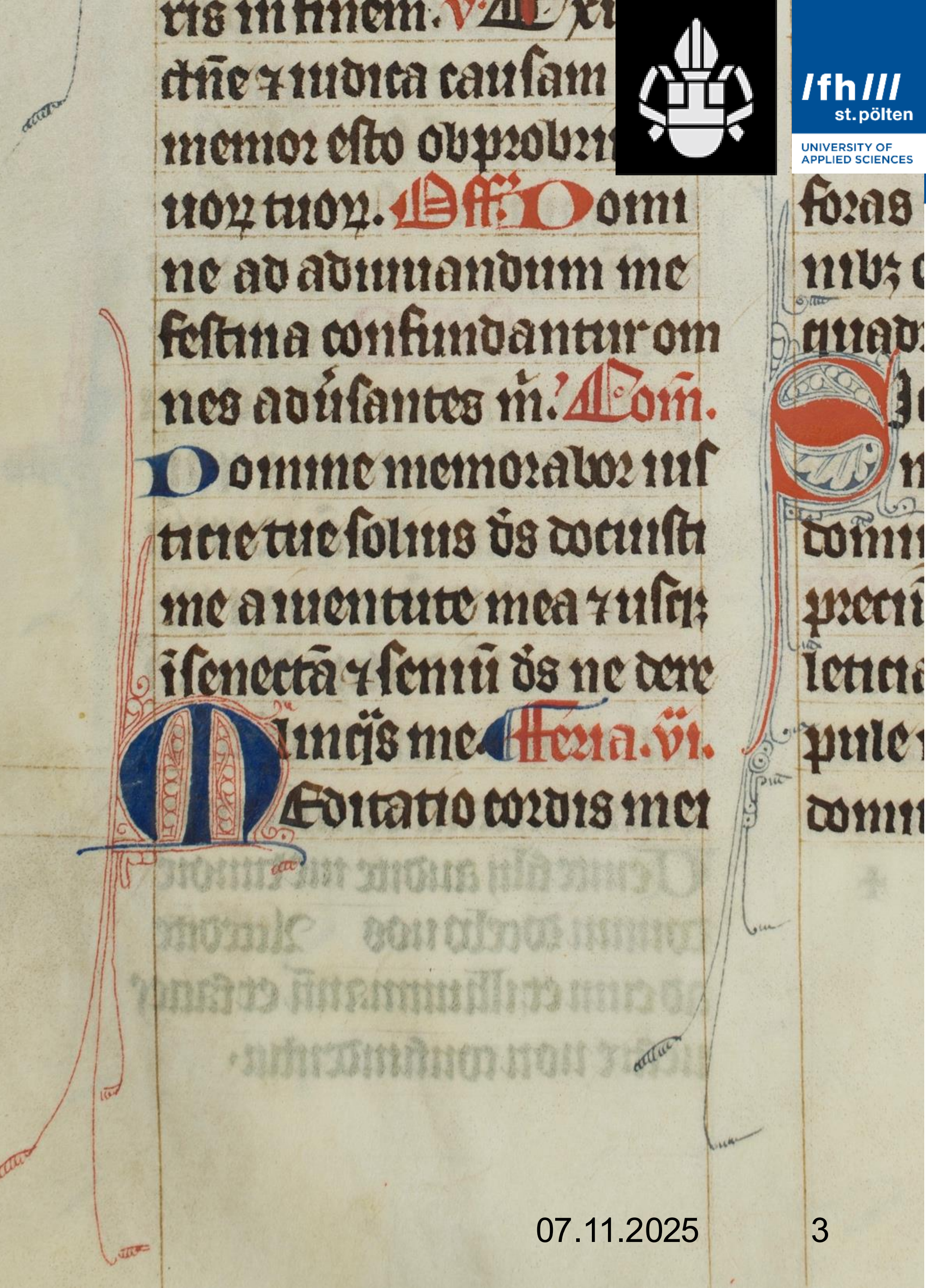
Source: Clementinae, and other text(s) von Clemens V. - Leiden University Libraries, Netherlands - Public Domain.



Source: Klosterneuburg Abbey

What is Pen Flourishing?

- Decorative ink line ornamentation (often red/blue) added after text was written [1,2, 3].
- Pen flourishing can take highly complex, interwoven shapes
- Created by rubricators (specialized scribes/artists) [4,5,6].
- Carried no narrative meaning, but expressed stylistic identity and workshop practice [1,2].
- Domain experts (art historians) use execution details of the flourishing to identify *rubricator groups* [4,5,6].
- Grouping rubricators by shared stylistic features allows for insight into medieval social structures and trading networks [1,2].



Source: Klosterneuburg Abbey (CCl 71 fol. 32v (Group B)

What is Pen Flourishing?

- Decorative ink line ornamentation (often red/blue) added after text was written [1,2, 3].
- Pen flourishing can take highly complex, interwoven shapes
- Created by rubricators (specialized scribes/artists) [4,5,6].
- Carried no narrative meaning, but expressed stylistic identity and workshop practice [1,2].
- Domain experts (art historians) use execution details of the flourishing to identify regional *rubricator groups* [4,5,6].
- Grouping rubricators by shared stylistic features allows for insight into medieval social structures and trading networks [1,2].

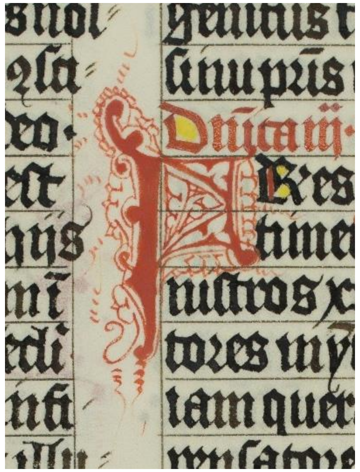
Rubricator Group A

CCl 71 fol. 108v

CCl 71 fol. 5r

CCl 71 fol. 4r

Rubricator Group D



CCl 71 fol. 217r

CCl 71 fol. 221r

CCl 71 fol. 240v

Source: Klosterneuburg Abbey

Research Context

Problem and Motivation

- Many manuscripts available, however …
- Few labeled examples - little ground truth to learn from.
- Hardly any studies on which groups belong together.
- Domain experts face too much data, need help exploring it.
- A tool is needed that is **explorative**, **explainable** to foster trust in experts.

Source: Klosterneuburg Abbey

State of the Art

- **Historical manuscript analysis**
  - **Early**: hand-crafted stroke/contour features + Bag-of-Features + SVM for writer dating/ID [7].
  - **Recent**: CNNs (AlexNet, Inception-ResNet), hybrids with transformers + keypoints, and self-/metric-learning (e.g., Online-Bag-of-Words) for stronger embeddings [8, 9, 10].
- **Illumination and decoration detection**
  - Region proposals + dense SIFT/VLAD + template matching → later Convolutional Neural Networks / Feature Pyramid Network pipelines and similarity-based retrieval improve detection [11, 12, 13].
- **Unsupervised art-style analysis (for paintings)**
  - Clustering & retrieval with deep embeddings (Deep Embedded Clustering / Deep Convolutional Embedded Clustering / DELIUS), multi-style feature aggregation, adversarial/contrastive training [14 ,15, 16, 17, 18].

State of the Art: Status for pen flourishing

- Little dedicated attention in prior work; usually folded into general illumination studies [19, 20, 21].
- Pen flourishing can take highly complex, interwoven shapes; style cues live in the strokes and micro-motifs [1].
- Global appearance is secondary for stylistic analysis - local, stroke-level features are primary cues for similarity.
- Methods emphasize global image features → limited interpretability for stroke-level comparisons.
- Hierarchical models (e.g., Feature Pyramid Networks, hierarchical transformers) offer finer multi-scale processing, but still lack transparency and interpretable region-level attributions.
- **Need for local modeling of motifs of arbitrary shape, transparent similarity attribution (region-level), and approaches that work with limited labels while aligning with expert reasoning.**

Gap / Our Contribution

- **Motif-level representation for pen flourishing**: Patch-based system that focuses on local stroke motifs rather than whole-page cues.
- **Transparent similarity & explainability**: Region/feature type attributions and visual overlays that show *why* two items match at the stroke level.
- **Label-efficient learning**: Works with scarce ground truth via self/metric learning and unsupervised clustering steps; robust with few labels.

Approach

Approach

Data

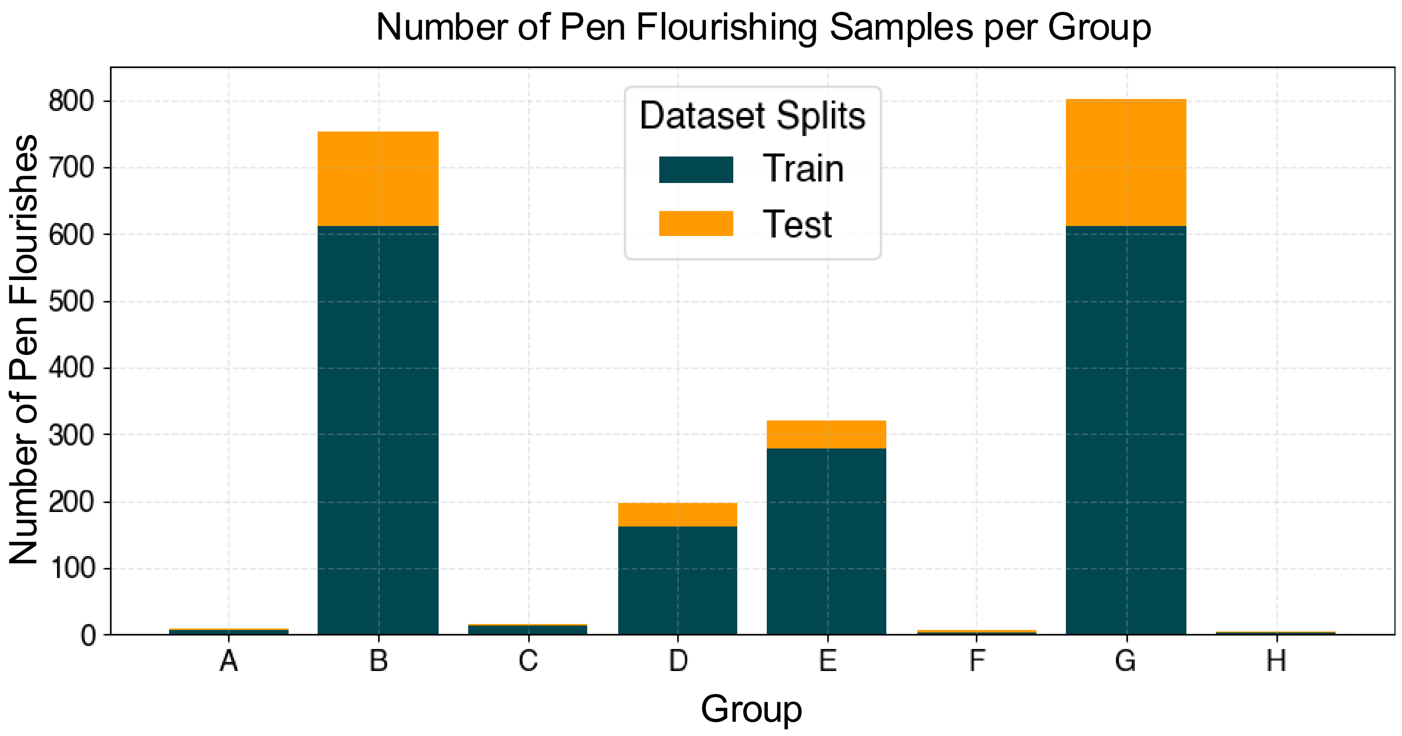
Total Number of Rubricator Groups	Total Number of Pen Flourishes
8	**2,102**

Total Number of Rubricator Groups (Used)	Total Number of Pen Flourishes (Used)
4	**788**

Source: Klosterneuburg Abbey

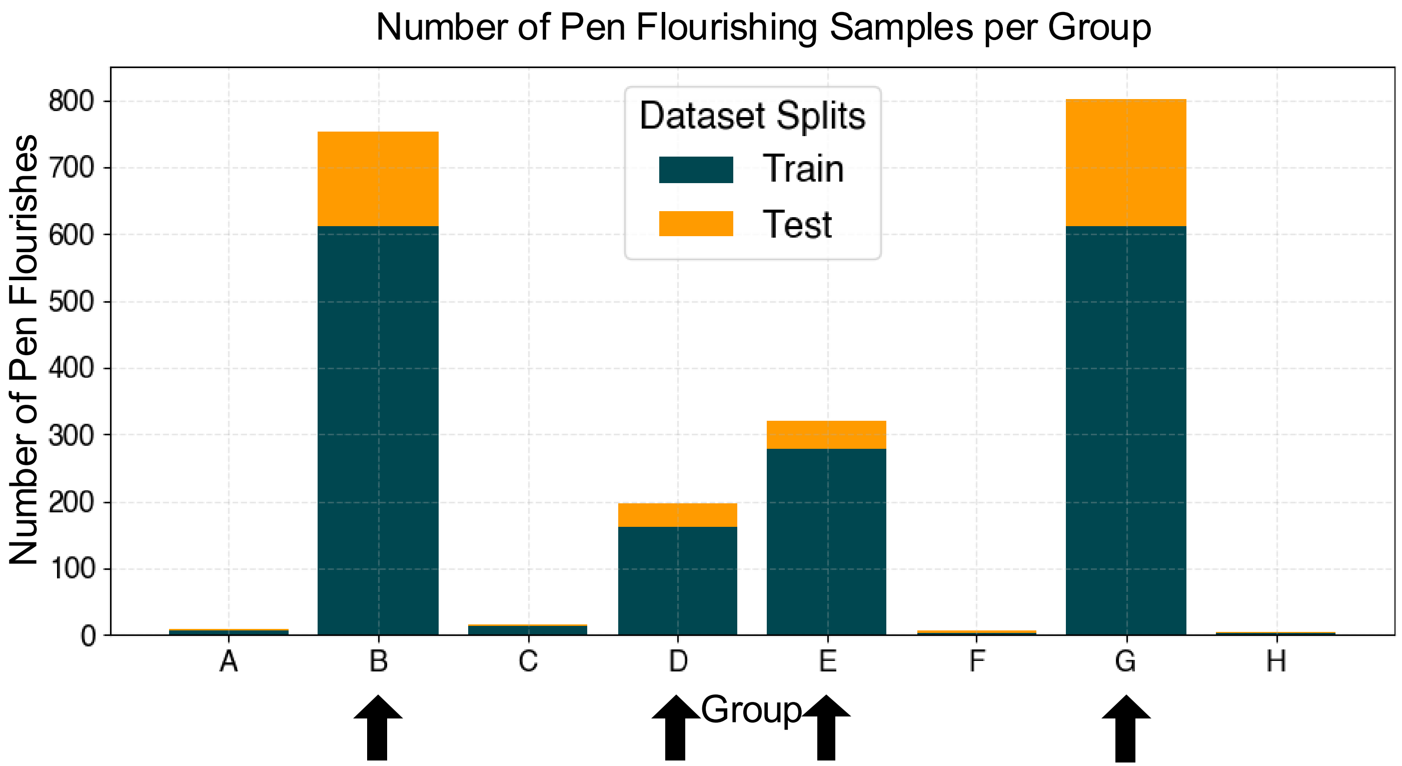
Data

Total Number of Rubricator Groups **8**	Total Number of Pen Flourishes **2,102**
Total Number of Rubricator Groups (Used) **4**	Total Number of Pen Flourishes (Used) **788**



Data

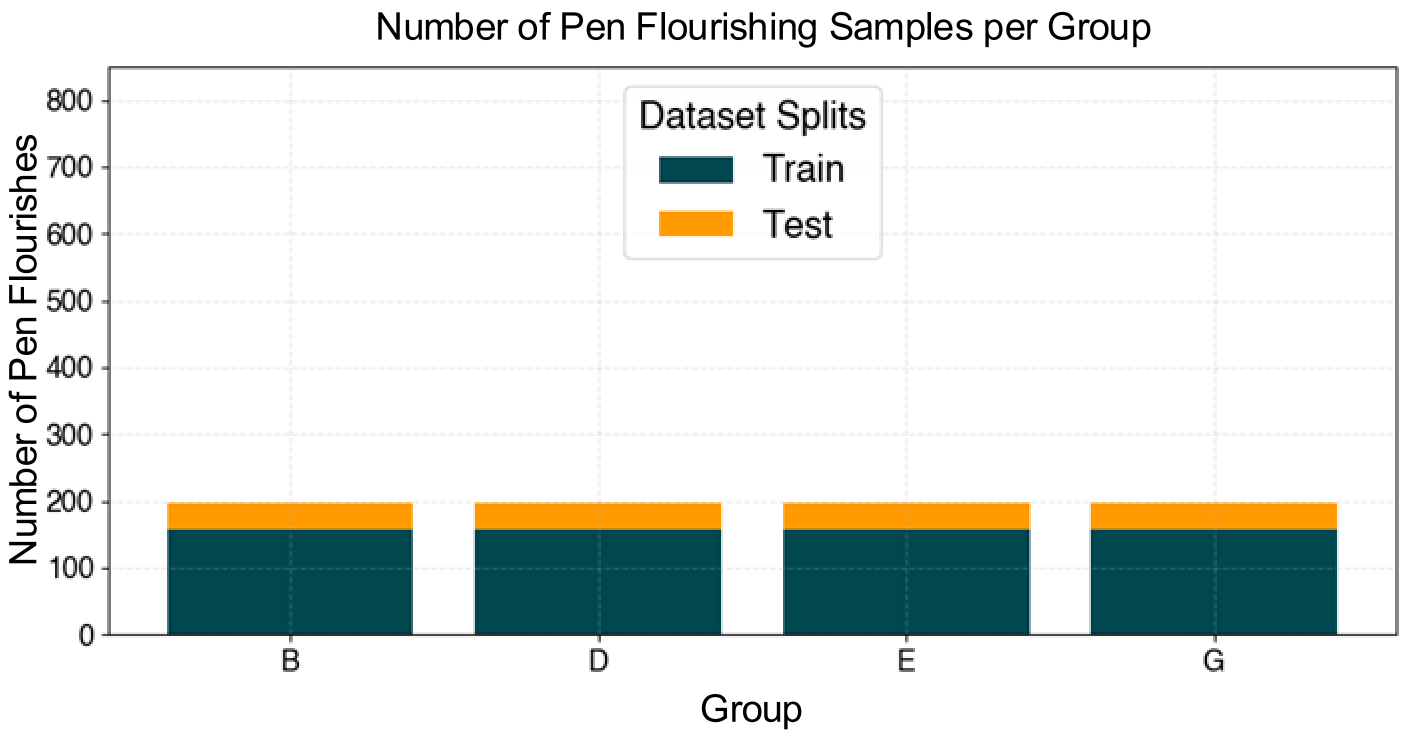
Total Number of Rubricator Groups **8**	Total Number of Pen Flourishes **2,102**
Total Number of Rubricator Groups (Used) **4**	Total Number of Pen Flourishes (Used) **788**



Approach

Data

Total Number of Rubricator Groups **8**	Total Number of Pen Flourishes **2,102**
Total Number of Rubricator Groups (Used) **4**	Total Number of Pen Flourishes (Used) **788**



Data

Total Number of Rubricator Groups	Total Number of Pen Flourishes
8	2,102
Total Number of Rubricator Groups (Used)	**Total Number of Pen Flourishes (Used)**
4	788

Pen Flourishing Example per Group

Group B CCl 71 fol. 23v3 | Group D CCl 71 fol. 226v1 | Group E CCl 73 fol. 43r2 | Group G CCl 73 fol. 158v3

Image Source: Klosterneuburg Abbey

Architecture

Segmentation

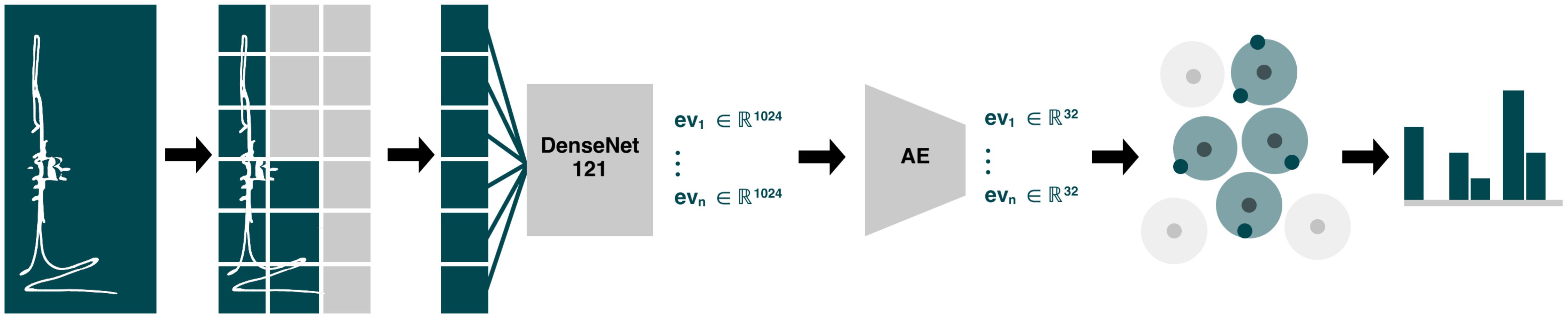
Dense Patching

Deep Feature Embedding

Dimensionality Reduction

Codebook Generation (K-Means)

Descriptor Encoding



Architecture

- Scanned in high-resolution (4435px × 5975px)
- Pen flourishing masks created by automated segmentation
- Gaussian blur applied to non-flourishing areas
- Pages are split by flourishes

CCl 71 fol. 17v3 (B)

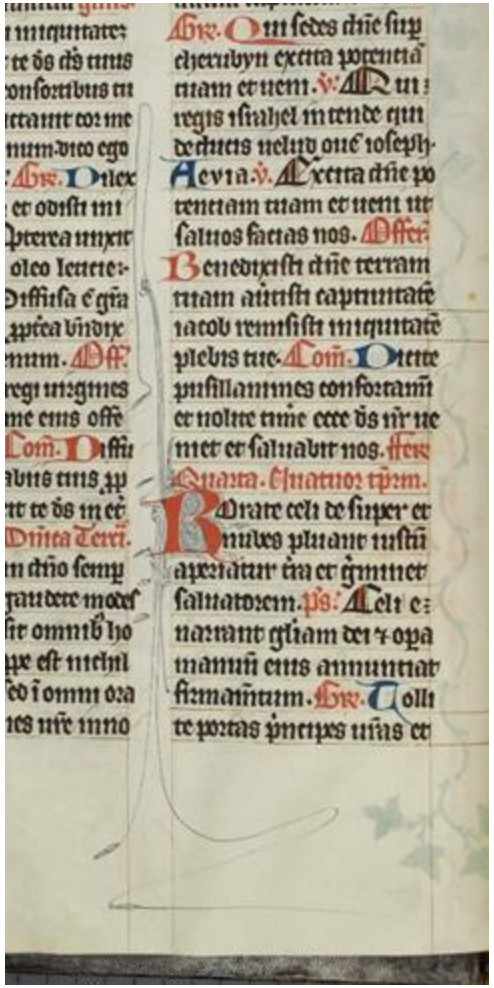
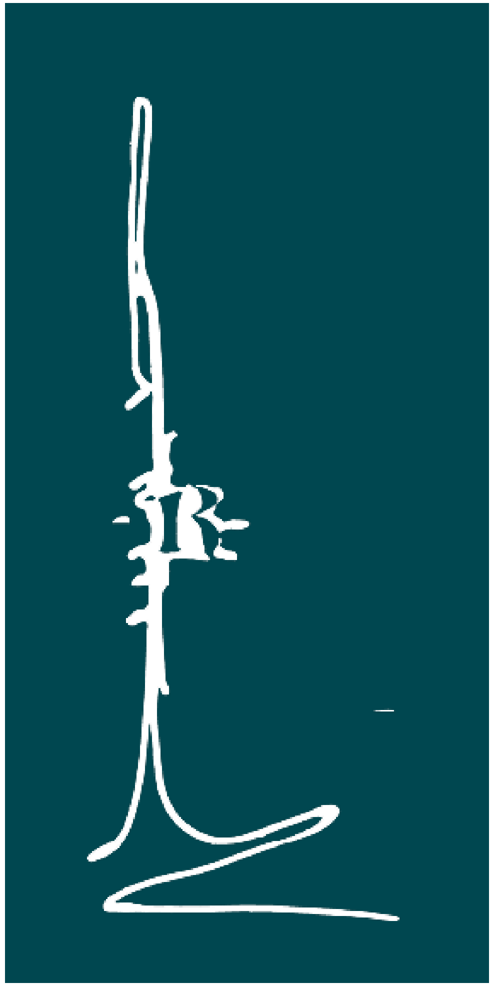


Image Source: Klosterneuburg Abbey

Architecture

- Scanned in high-resolution (4435px × 5975px)
- Pen flourishing masks created by automated segmentation
- Gaussian blur applied to non-flourishing areas
- Pages are split by flourishes

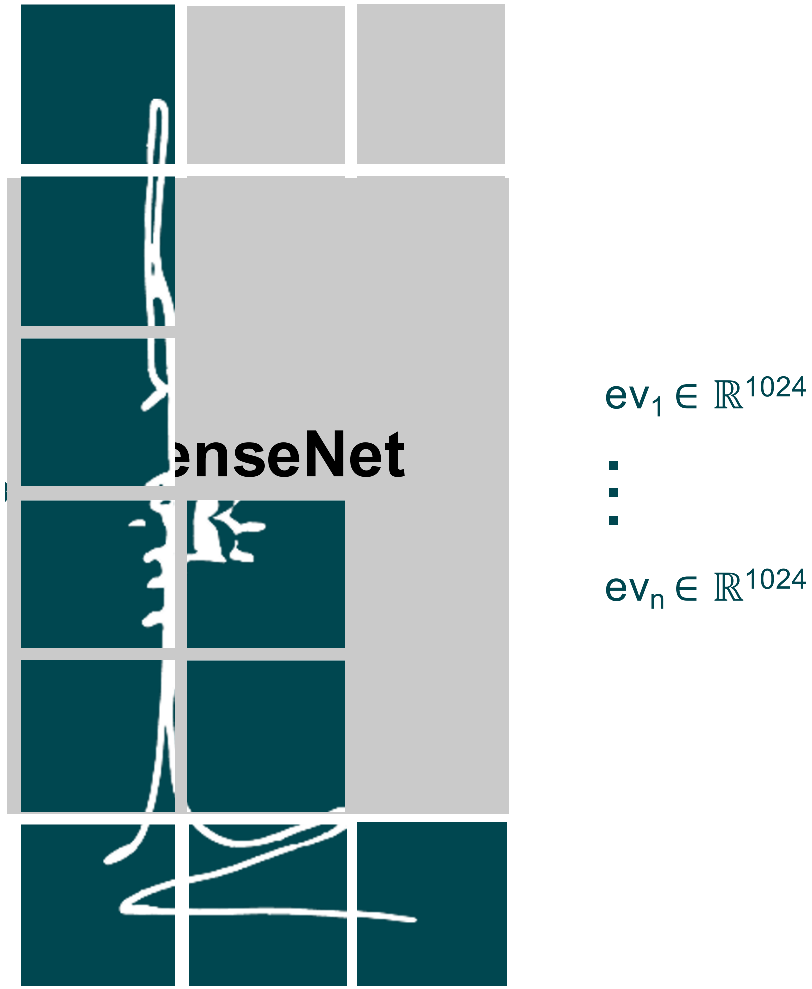


Architecture

- Patches extracted from a regular overlapping grid
  - Patch size set to 224px × 224px
  - Overlap of 1/3 of the patch size (74px)
  - Ensures independence from pen flourishing shape
  - Focusses on local features
- Patches with >20% pen flourishing foreground (masks) are retained

Architecture

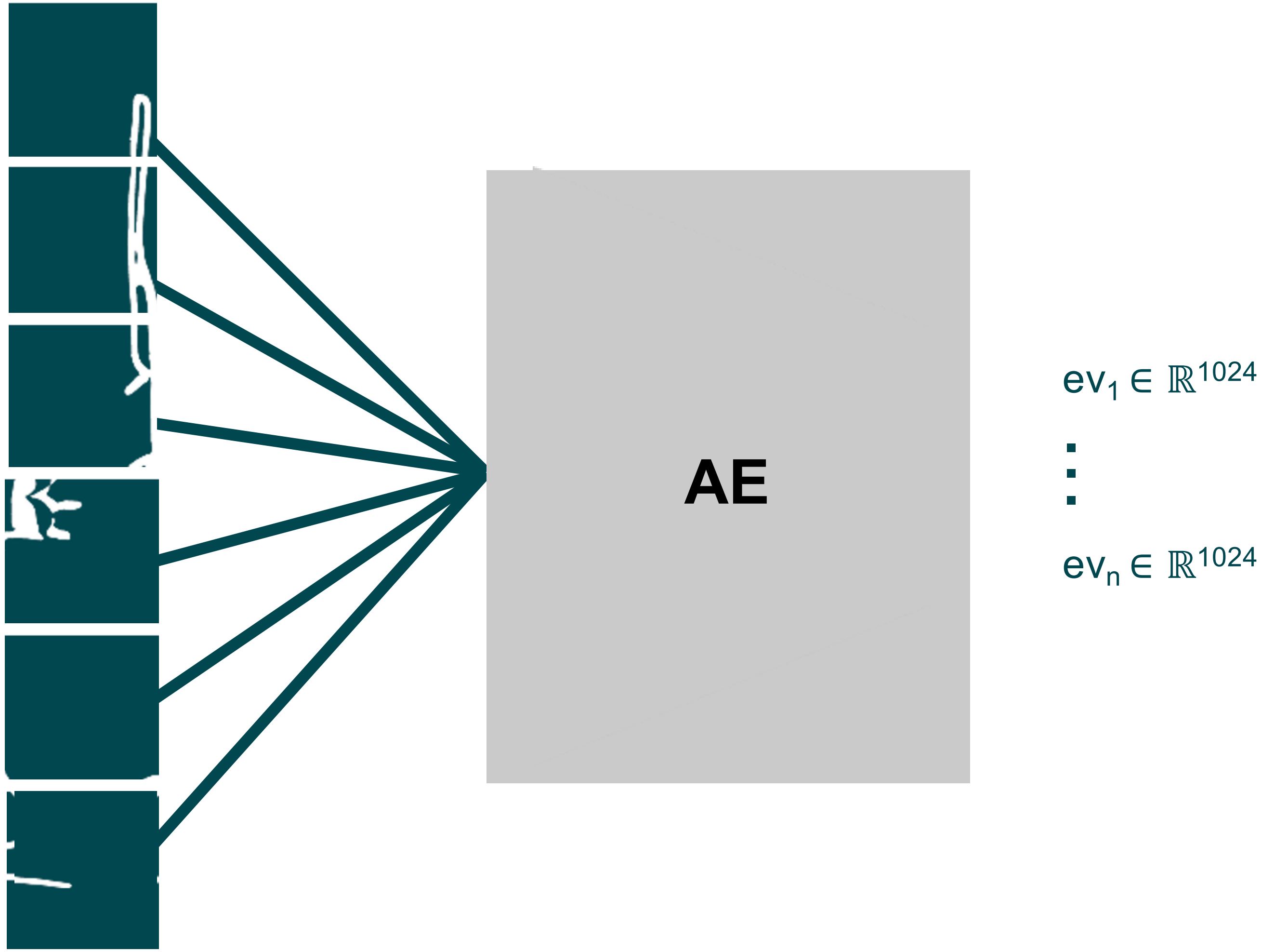
- DenseNet-121 [22] used to embed patches
  - Cut off at the global average pooling layer
- Subsequent non-linear dimensionality reduction with fully connected autoencoder
  - Four-layered architecture: 500-500-2000-32 (3.57M parameters)
  - ReLU activations for non-linearity
  - Trained with mean squared error objective



Approach

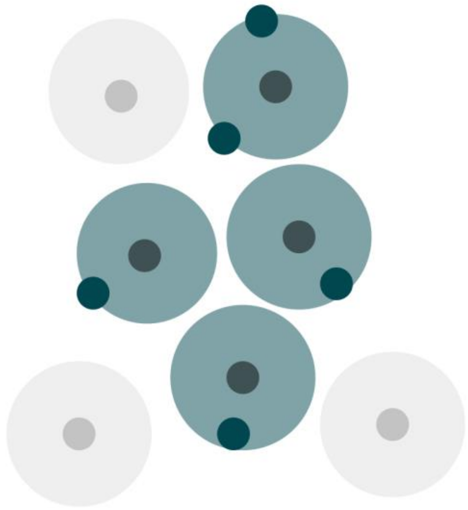
Architecture

- DenseNet-121 [22] used to embed patches.
  - Cut off at the global average pooling layer
- Subsequent non-linear dimensionality reduction with fully connected autoencoder.
  - Four-layered architecture: 500-500-2000-32 (3.57M parameters)
  - ReLU activations for non-linearity.
  - Trained with mean squared error objective

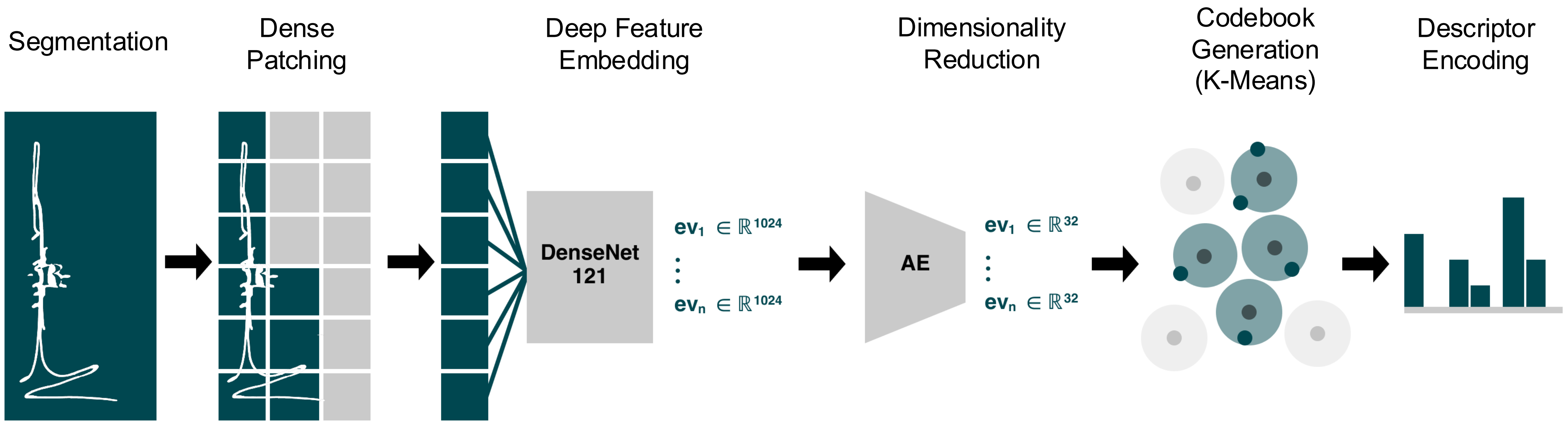


Architecture

- Computation of Bag-of-Features [23] codebook of size 100 with K-Means [24] clustering.
- Computation of global histogram vectors based on codebook.
- Root and L_2 normalization applied to the histogram vectors.

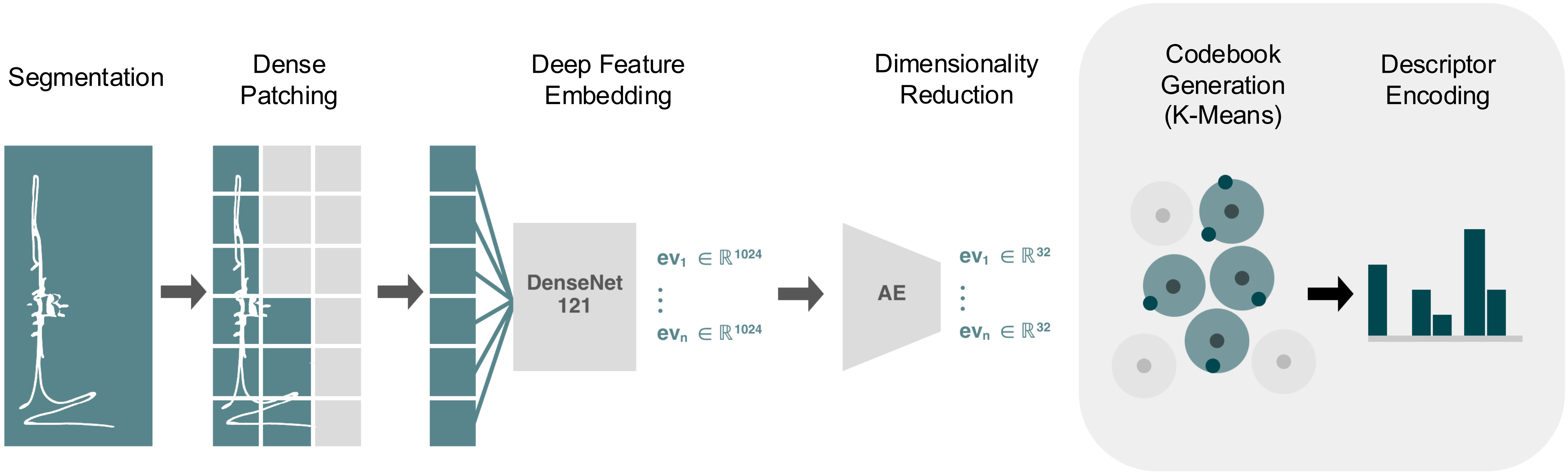


Architecture



Approach

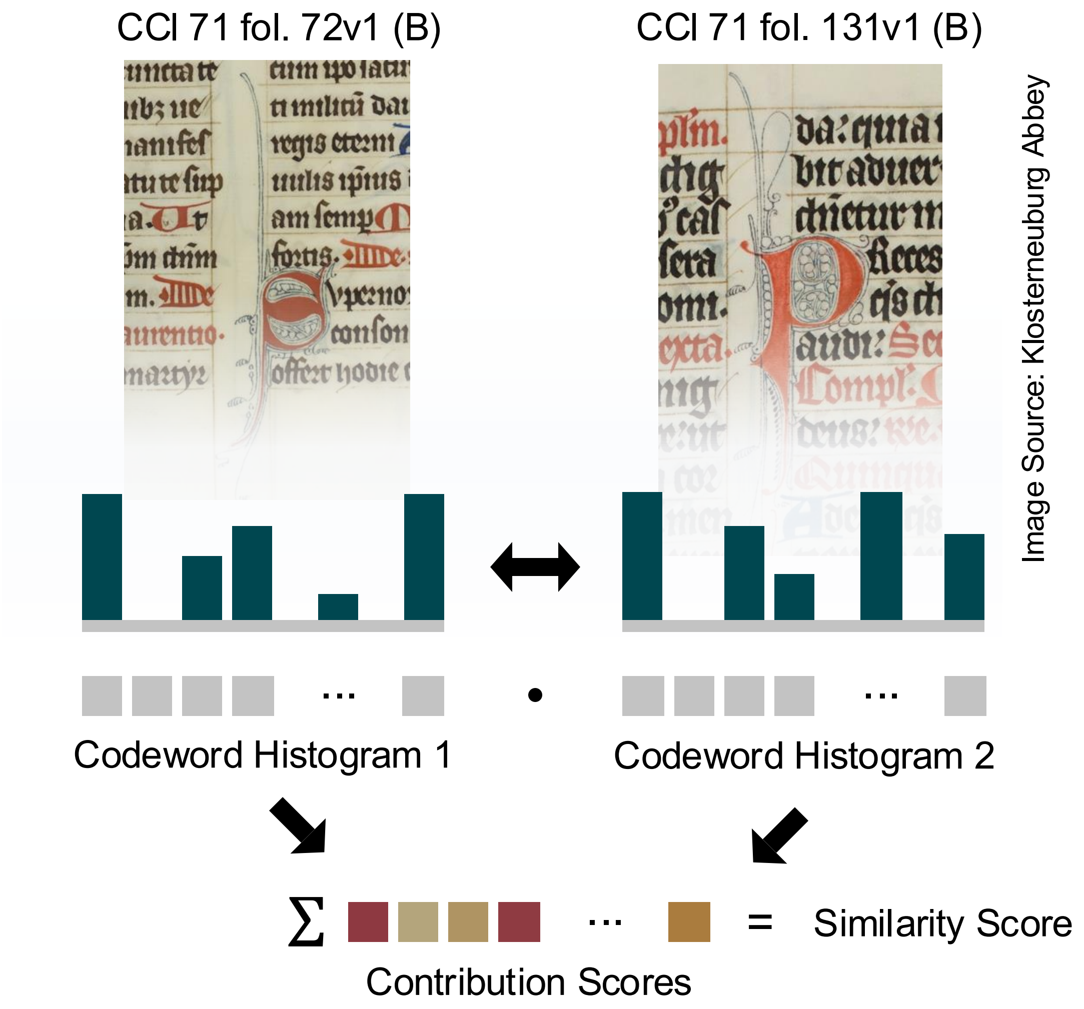
Model Transparency



Similarity described by single scalar

Approach

Model Transparency



1 Similarity-Level Transparency

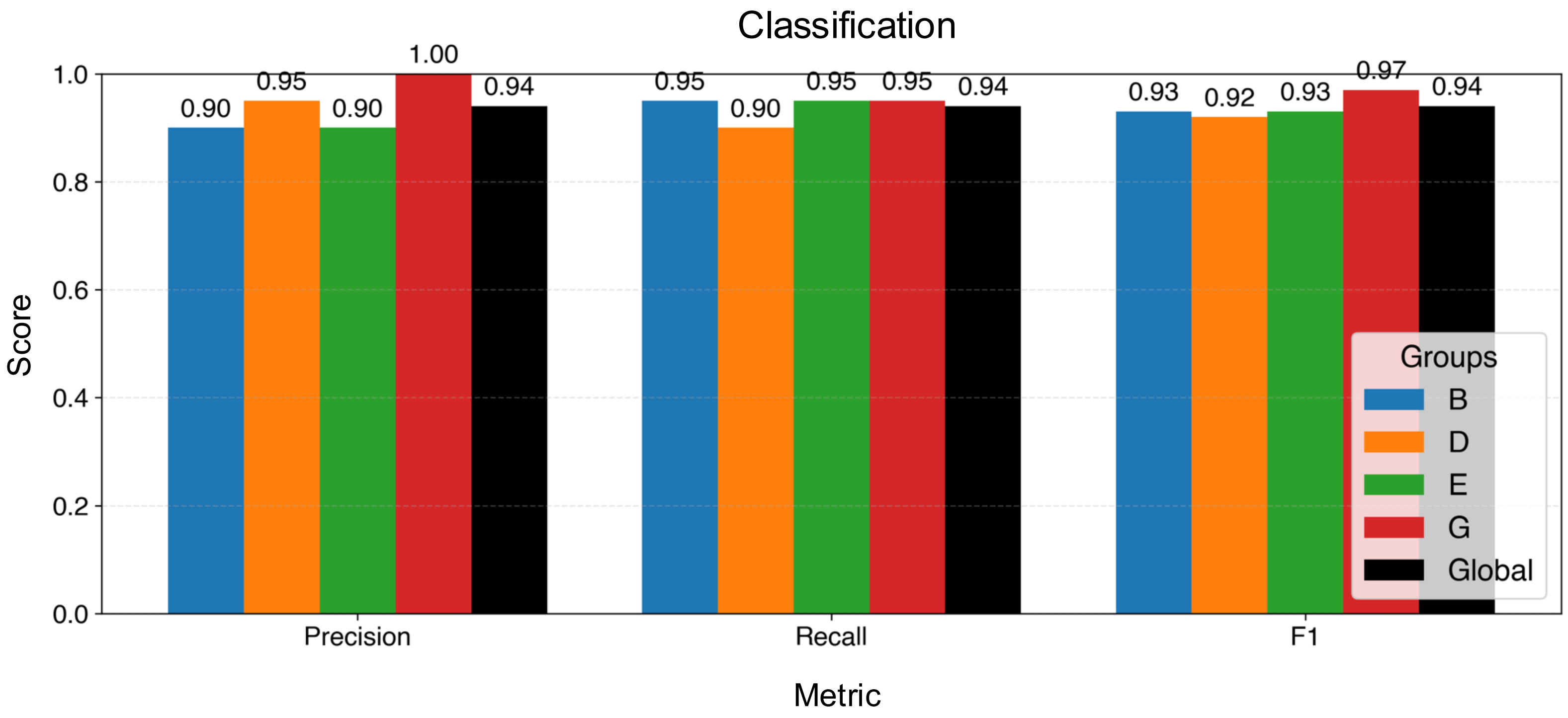
Why are flourishes similar?

2 Cluster-Level Transparency

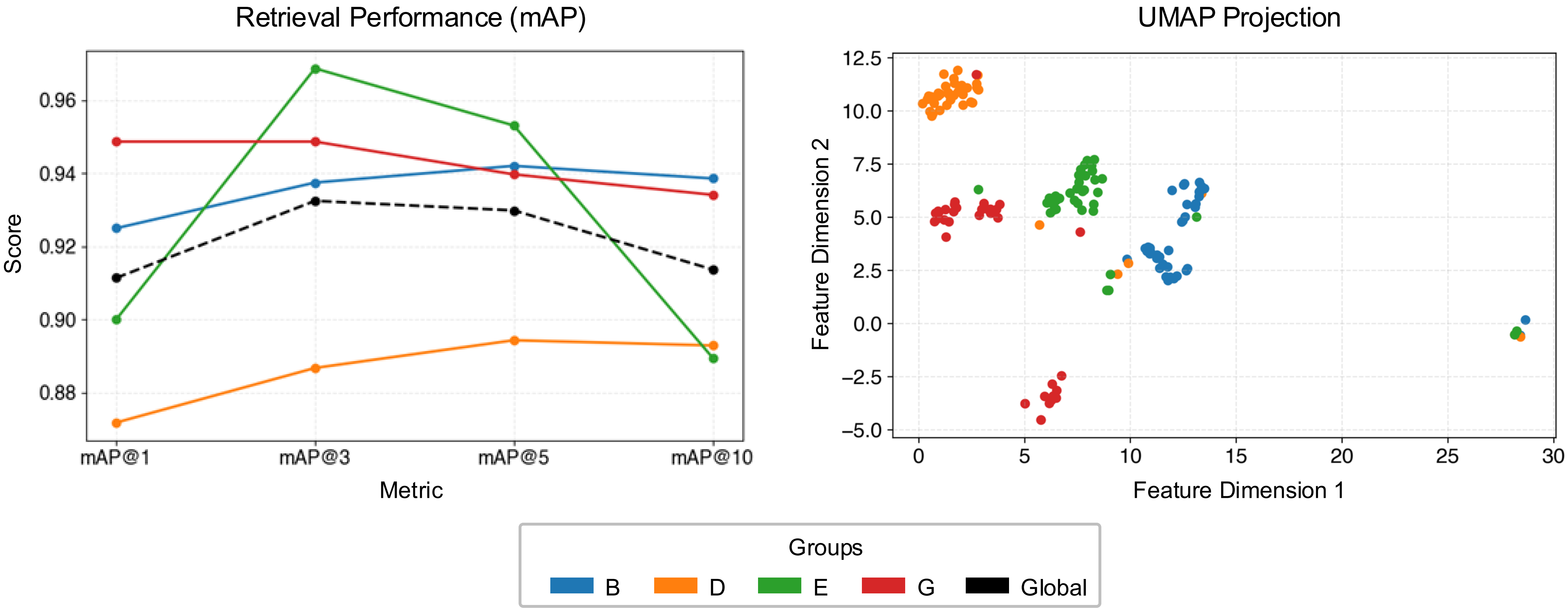
Which features are most characteristic for each group?

Classification

- Classification performance measured by a k-Nearest neighbor classifier [25] trained on top of the codeword histograms.



Feature Space



Similarity-Level Transparency

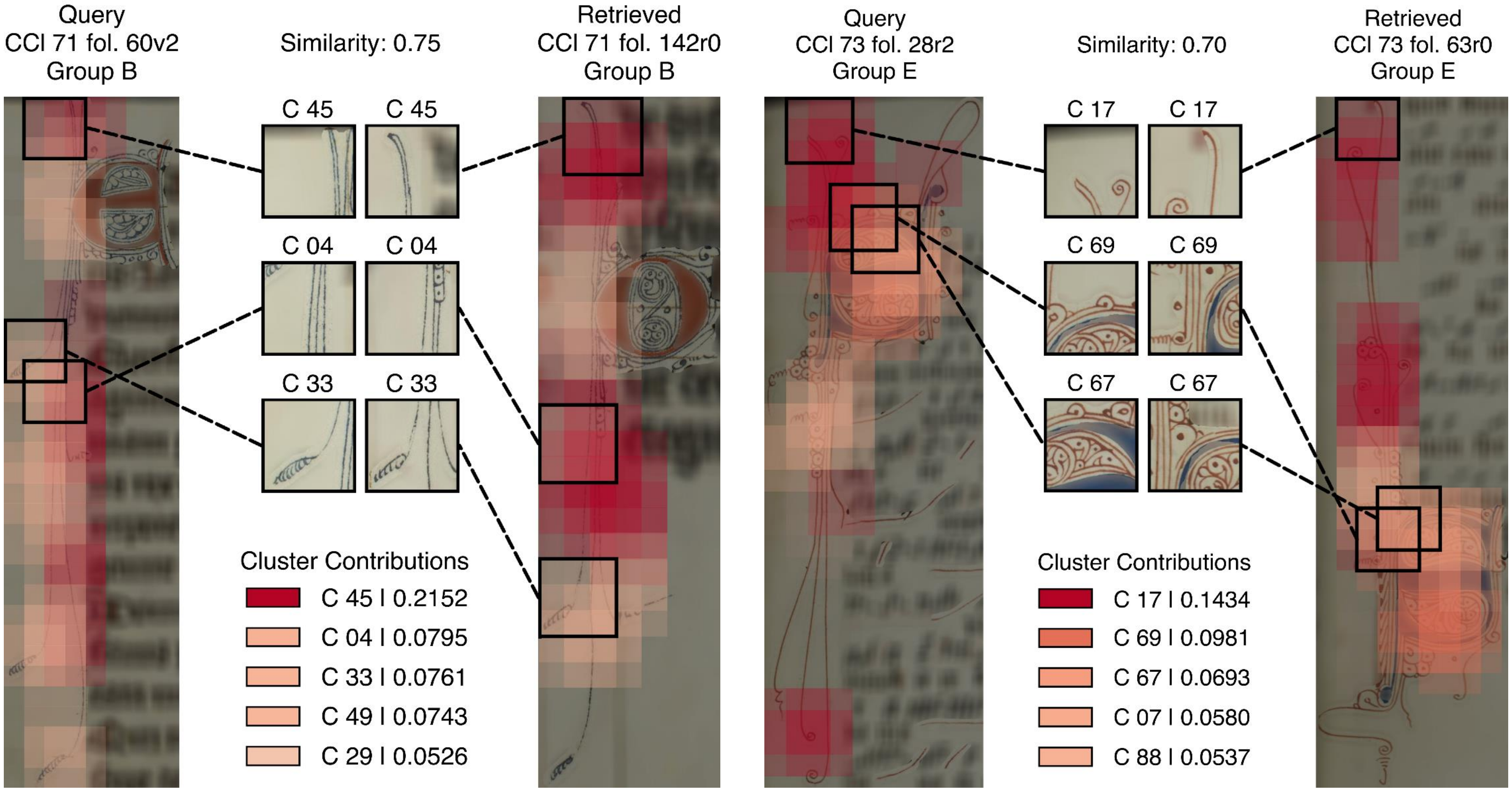


Image Source: Klosterneuburg Abbey

Similarity-Level Transparency

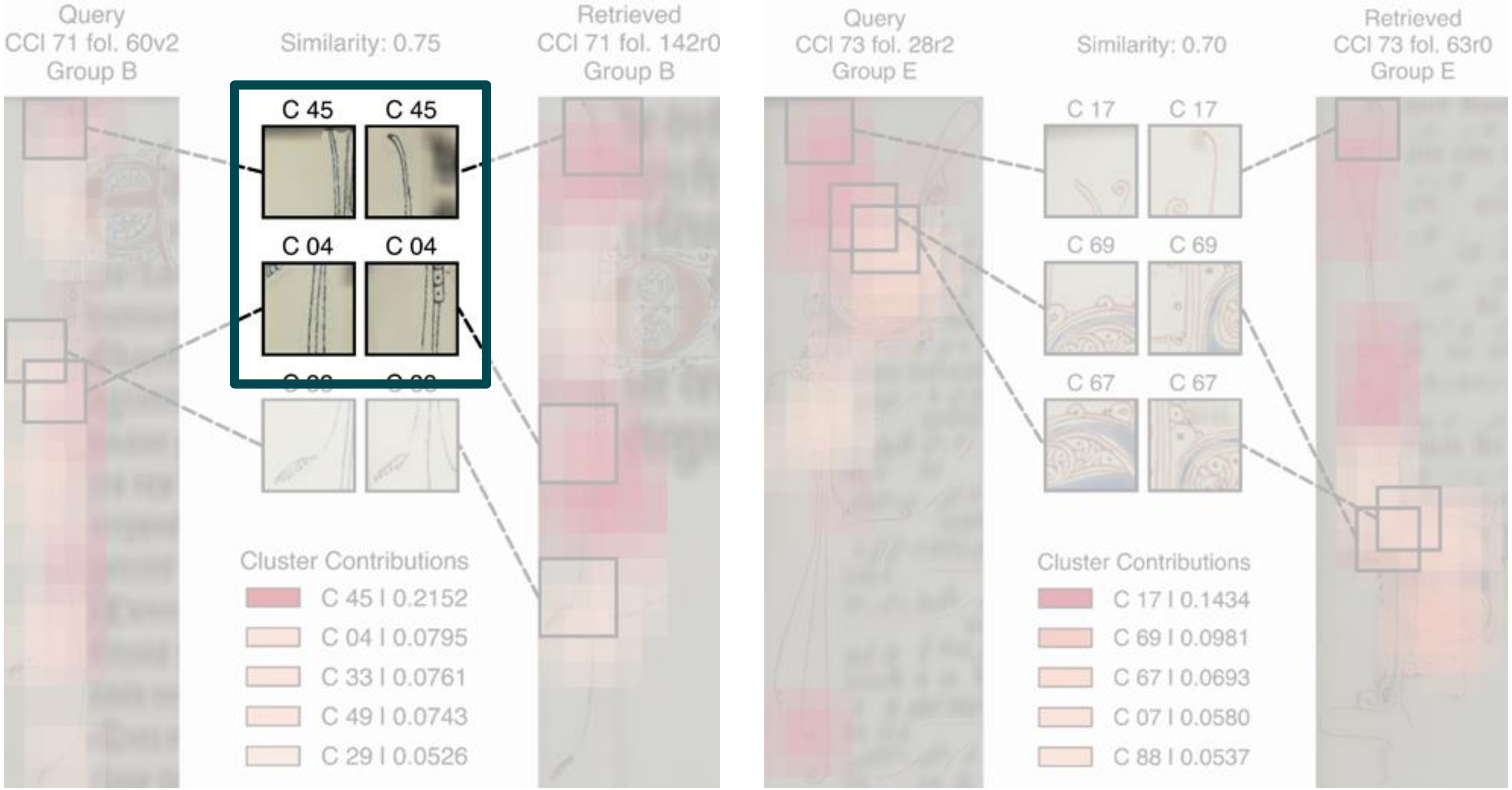


Image Source: Klosterneuburg Abbey

Similarity-Level Transparency

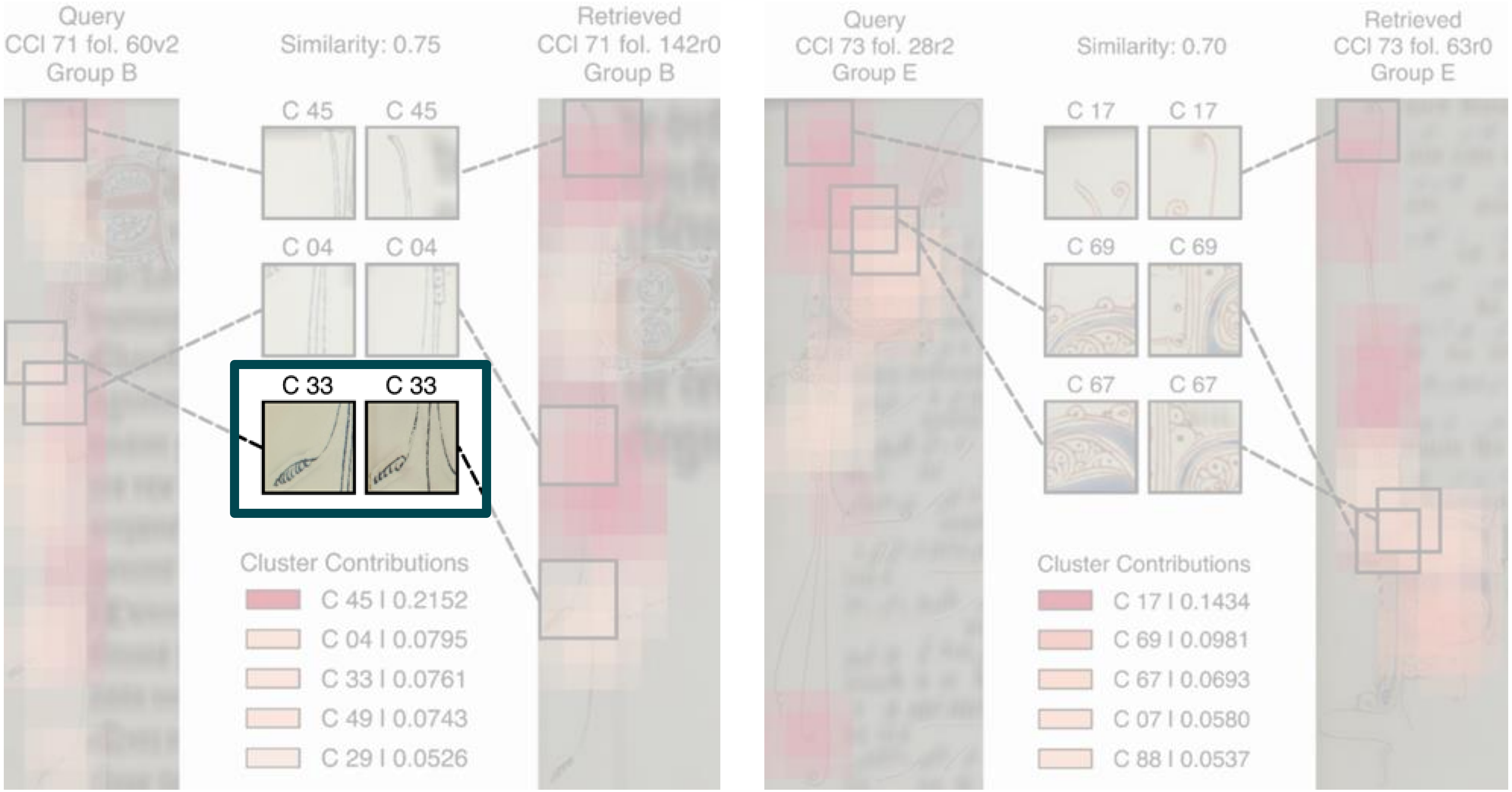


Image Source: Klosterneuburg Abbey

Evaluation

Similarity-Level Transparency

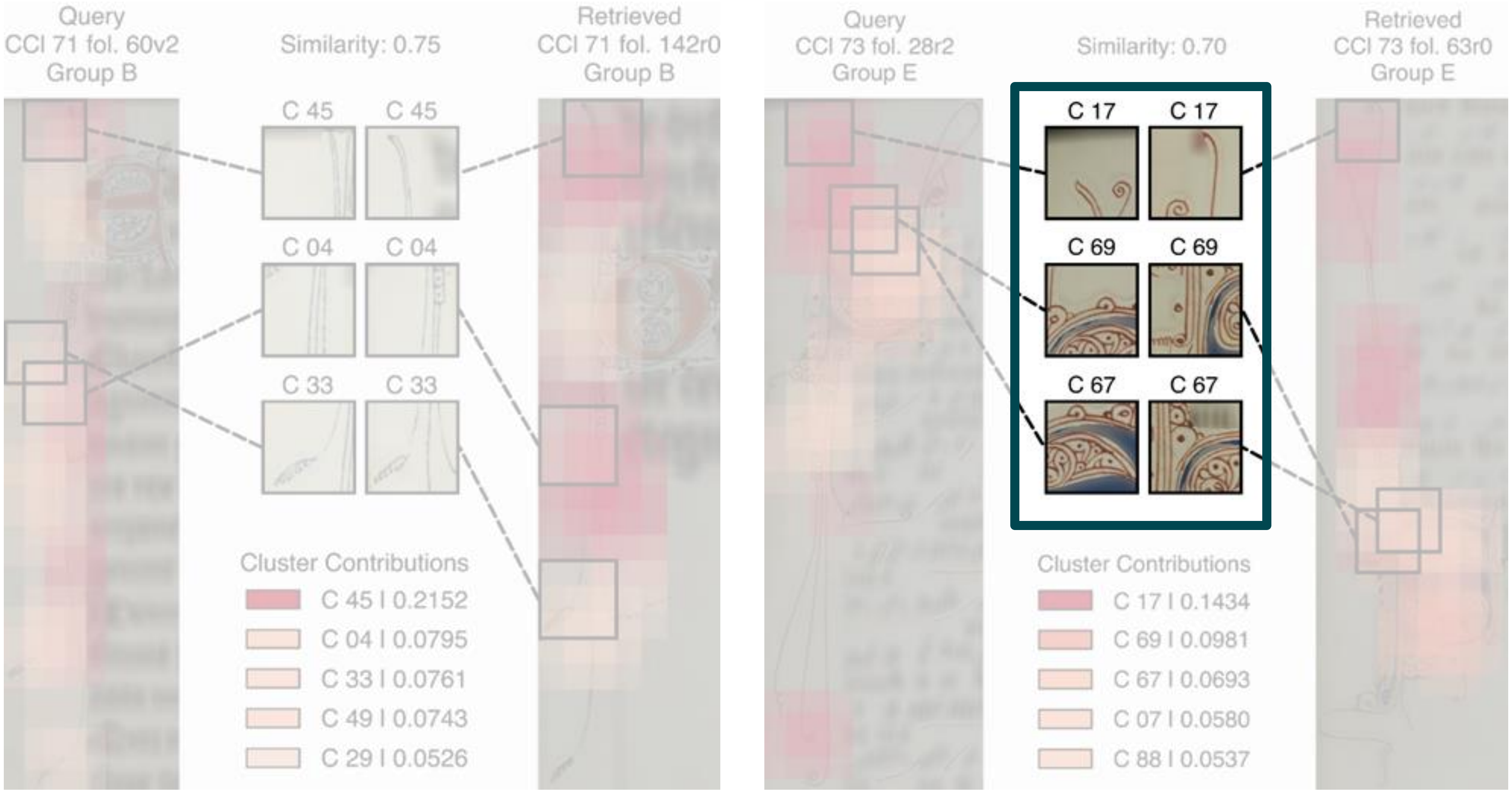


Image Source: Klosterneuburg Abbey

Evaluation

Cluster-Level Transparency

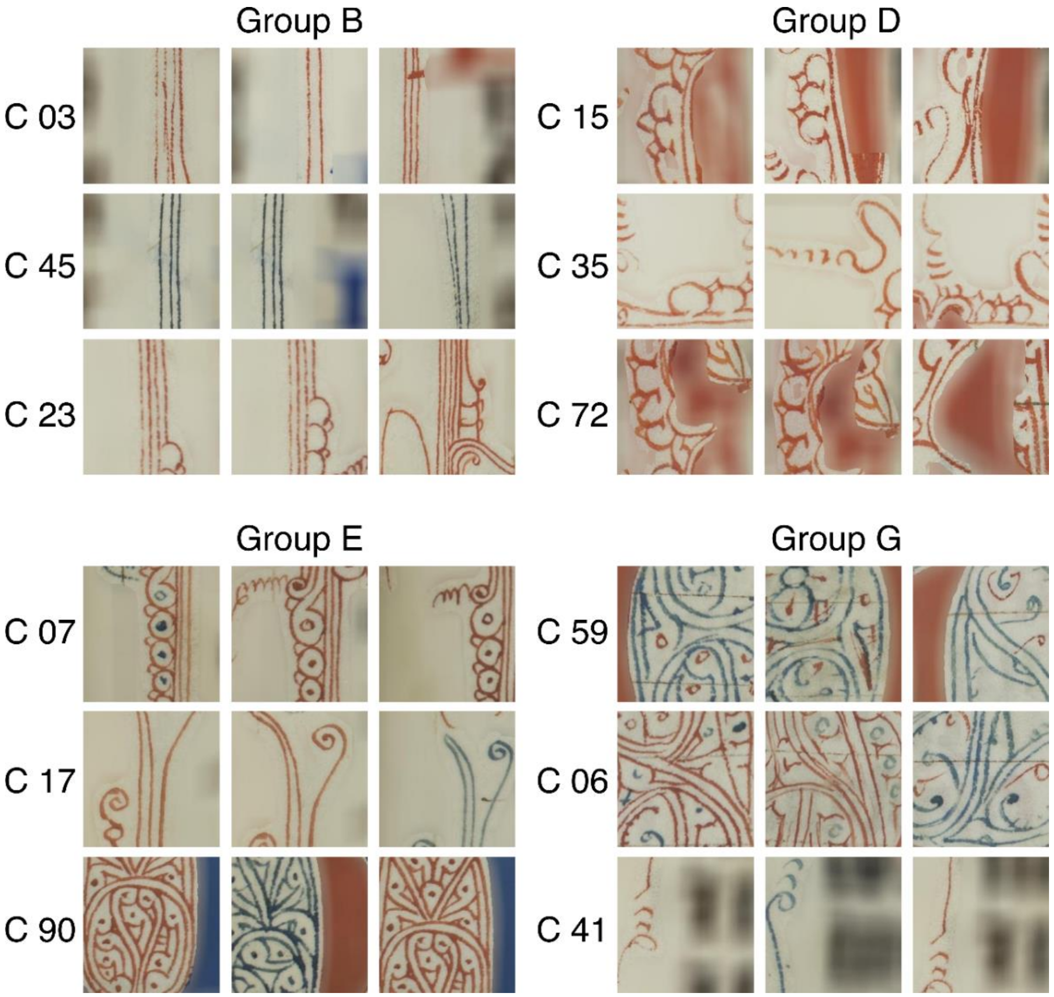


Image Source: Klosterneuburg Abbey

Cluster-Level Transparency

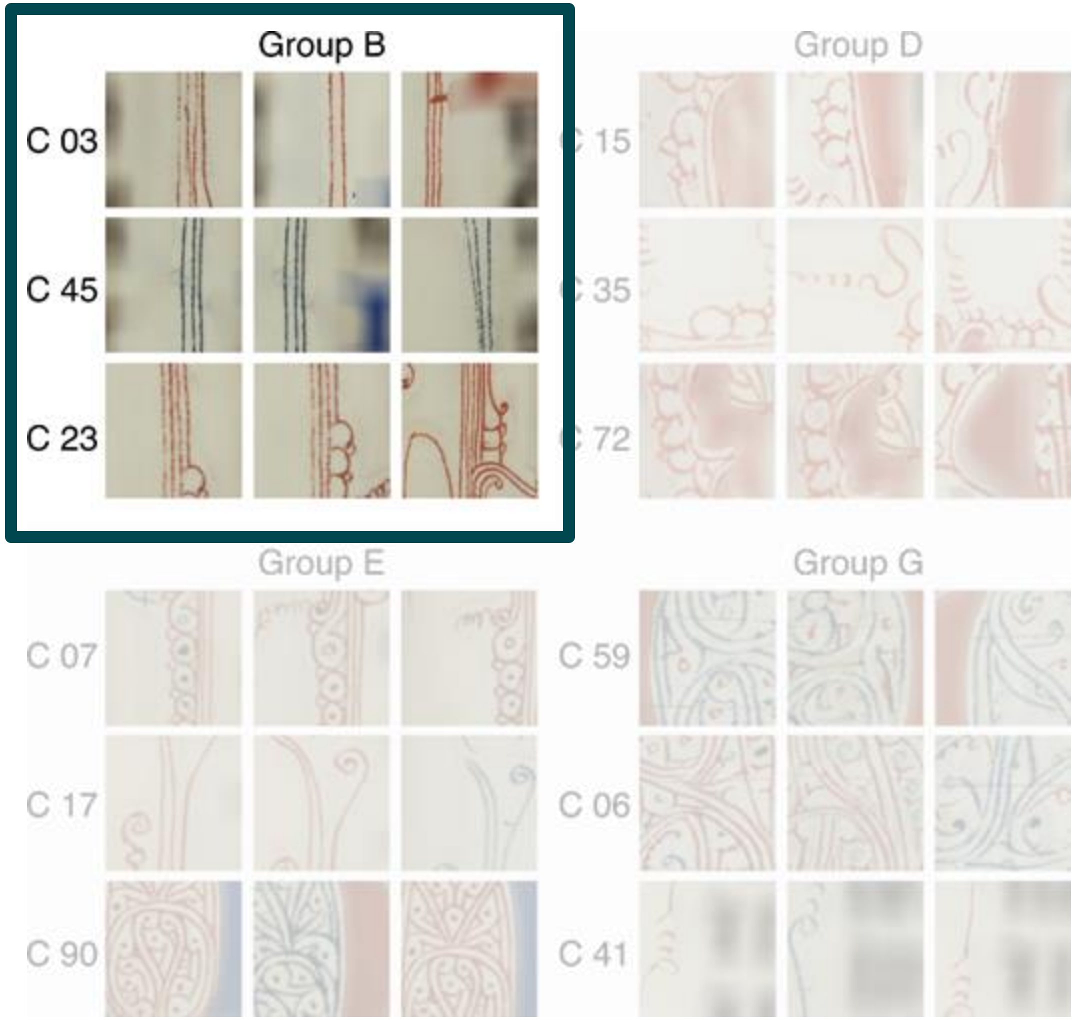


Image Source: Klosterneuburg Abbey

Cluster-Level Transparency

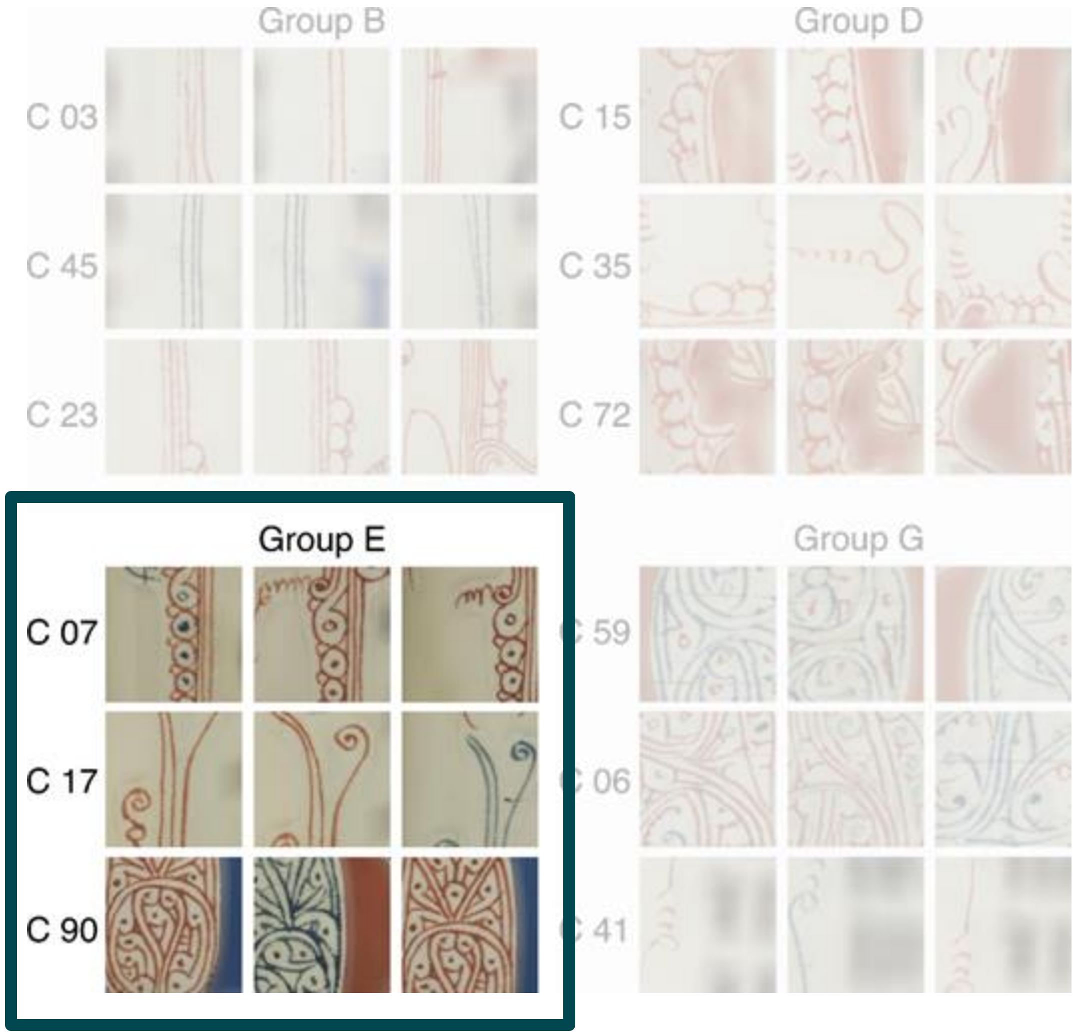


Image Source: Klosterneuburg Abbey

Domain Expert Annotations

Our Approach | Expert Annotations

CCl 71 folio 120v0 (C)

Our Approach | Expert Annotations

CCl 71 fol. 214r3 (E)

Image Source: Klosterneuburg Abbey

Domain Expert Annotations

Our Approach | Expert Annotations

CCl 71 folio 120v0 (C)

Our Approach | Expert Annotations

CCl 71 fol. 214r3 (E)

Image Source: Klosterneuburg Abbey

Domain Expert Annotations

Our Approach | Expert Annotations

CCl 71 folio 120v0 (C)

Our Approach | Expert Annotations

CCl 71 fol. 214r3 (E)

Image Source: Klosterneuburg Abbey

Takeaways

1 We presented a Bag-of-Features approach which allows both high quantitative performance and qualitative interpretability in visual style analysis of pen flourishing.

2 A codeword-contribution-based transparency layer reveals which visual motifs drive similarity - enabling explainable, expert-aligned results.

3 We identified that the *discriminative visual vocabulary* corresponds closely to art-historical motifs.

Outlook

1. **Scale up**: Extend dataset to more manuscripts, periods, and stylistic schools for broader validation.
2. **Benchmarking**: Compare against end-to-end architectures to better contextualize model performance.
3. **Interactive Tool**: Develop an expert-in-the-loop interface for similarity search, visual exploration, and motif comparison.

Thank you!

Acknowledgements

We thank the art historians of our “Ornament & Algorithm Workshop” and acknowledge Gesellschaft für Forschungsförderung Niederösterreich (GFF) for funding.

Code and dataset

Research project

Sources

1. J. J. G. Alexander, Ed., Medieval illuminators and their methods of work, eng. New Haven: Yale Univ. Press, 1992, ISBN: 978-0-300-06073-7 978-0-300-05689-1.
2. Otto Mazal, Buchkunst der Gotik, ger. Graz: Akademische Druck - und Verlagsanstalt, 1975, ISBN: 978-3-201-00949-2.
3. Christine Jakobi-Mirwald, “Flourish and blotts or how we name salient ornament features in manuscripts,” Textes et Etudes du Moyen Âge, vol. 94, pp. 183–205, 2019.
4. Wolfgang Augustyn, Christine Jakobi-Mirwald, Christine Sauer, and Martin Roland, “Fleuronné,” Reallexikon zur Deutschen Kunstgeschichte, vol. 9, pp. 1114–1196, 1997.
5. Evelyn Theresia Kubina, “Für recht und ordnung. fleuronnée-initialen in handschriften und urkunden des vierten viertels des 13. jahrhunderts in wien und niederösterreich,” in Europäische Bild- und Buchkultur im 13. Jahrhundert, Wien, Köln, Weimar, 2020, pp. 167– 177.
6. Martin Roland, Buchschmuck in Lilienfelder Handschriften. Von der Gründung des Stiftes bis zur Mitte des 14. Jahrhunderts (Studien und Forschungen aus dem Niederösterreichis- chen Institut für Landeskunde 22). Wien: NÖ Institut für Landeskunde, 1996.

Sources

7. Sheng He, Petros Samara, Jan Burgers, and Lambert Schomaker. 2016. Image-based historical manuscript dating using contour and stroke fragments. PatternRecognition 58 (2016), 159–171. doi:10.1016/j.patcog.2016.03.032

8. Anmol Hamid, Maryam Bibi, Momina Moetesum, and Imran Siddiqi. 2019. DeepLearning Based Approach for Historical Manuscript Dating. In 2019 InternationalConference on Document Analysis and Recognition (ICDAR). 967–972. doi:10.1109/ICDAR.2019.00159

9. Merouane Boudraa, Akram Bennour, Mohammed Al-Sarem, Fahad Ghabban,and Omair Ameer Bakhsh. 2024. Contribution to historical manuscript dating:A hybrid approach employing hand-crafted features with vision transformers.Digital Signal Processing 149 (June 2024), 104477. doi:10.1016/j.dsp.2024.104477

10. Francimaria Nascimento, Stephen Smith, and Márjory Costa-Abreu. 2022. Ex-ploring Medieval Manuscripts Writer Predictability: A Study on Scribe andLetter Identification. Digital Studies / Le champ numérique 12 (01 2022).doi:10.16995/dscn.8096

11. David G. Lowe. 2004. Distinctive Image Features from Scale-Invariant Keypoints.International Journal of Computer Vision 60, 2 (Nov. 2004), 91–110. doi:10.1023/B:VISI.0000029664.99615.94

Sources

12. Sovann En, Caroline Petitjean, Stéphane Nicolas, Laurent Heutte, and FrédéricJurie. 2016. Pattern localization in historical document images via templatematching. In 2016 23rd International Conference on Pattern Recognition (ICPR).2054–2059. doi:10.1109/ICPR.2016.7899938

13. Ignacio Úbeda, Jose M. Saavedra, Stéphane Nicolas, Caroline Petitjean, andLaurent Heutte. 2020. Improving pattern spotting in historical documents using feature pyramid networks. Pattern Recognition Letters 131 (2020), 398–404.doi:10.1016/j.patrec.2020.02.002

14. Giovanna Castellano and Gennaro Vessio. 2021. Deep Convolutional Embeddingfor Digitized Painting Clustering. In 2020 25th International Conference on PatternRecognition (ICPR). 2708–2715. doi:10.1109/ICPR48806.2021.9412438

15. Giovanna Castellano and Gennaro Vessio. 2022. A Deep Learning Approach to Clustering Visual Arts. International Journal of Computer Vision 130, 11 (Aug.2022), 2590–2605. doi:10.1007/s11263-022-01664-y

16. Nikolai Ufer, Max Simon, Sabine Lang, and Björn Ommer. 2021. Large-scale interactive retrieval in art collections using multi-style feature aggregation. PLOSONE 16, 11 (Nov. 2021), e0259718. doi:10.1371/journal.pone.0259718

Sources

17. Catherine Sandoval, Elena Pirogova, and Margaret Lech. 2021. AdversarialLearning Approach to Unsupervised Labeling of Fine Art Paintings. IEEE Access9 (2021), 81969–81985. doi:10.1109/ACCESS.2021.3086476

18. Lucas David, Luis Liborio, Guilherme Paiva, Marianna Severo, Eduardo Valle,Helio Pedrini, and Zanoni Dias. 2023. Adversarial Feature Hallucination in aSupervised Contrastive Space for Few-Shot Learning of Provenance in Paintings.In 2023 IEEE Latin American Conference on Computational Intelligence (LA-CCI).1–6. doi:10.1109/LA-CCI58595.2023.10409405

19. Christofer Meinecke, Estelle Guéville, David Joseph Wrisley, and Stefan Jänicke.2022. Is Medieval Distant Viewing Possible? : Extending and Enriching Annota-tion of Legacy Image Collections using Visual Analytics. doi:10.48550/ARXIV.2208.09657 Version Number: 2.

20. Fouad Aouinti, Victoria Eyharabide, Xavier Fresquet, and Frédéric Billiet. 2022.Illumination Detection in IIIF Medieval Manuscripts Using Deep Learning. DigitalMedievalist 15, 1 (Aug. 2022). doi:10.16995/dm.8073

21. Ying Yang, Ruggero Pintus, Enrico Gobbetti, and Holly Rushmeier. 2015. Automated color clustering for medieval manuscript analysis. In 2015 Digital Heritage,Vol. 2. 101–104. doi:10.1109/DigitalHeritage.2015.7419462

Sources

22. Gao Huang, Zhuang Liu, Laurens van der Maaten, and Kilian Q. Weinberger. 2016.Densely Connected Convolutional Networks. doi:10.48550/ARXIV.1608.06993Version Number: 5

23. J. Sivic, B.C. Russell, A.A. Efros, A. Zisserman, and W.T. Freeman. 2005. Dis-covering objects and their location in images. In Tenth IEEE InternationalConference on Computer Vision (ICCV’05) Volume 1, Vol. 1. 370–377 Vol. 1.doi:10.1109/ICCV.2005.77

24. Pang-Ning Tan, Michael Steinbach, and Vipin Kumar. 2014. Introduction to DataMining (new internat. edition ed.). Pearson, Harlow.

25. Padraig Cunningham and Sarah Delany. 2007. k-Nearest neighbour classifiers.Mult Classif Syst 54 (04 2007). doi:10.1145/3459665

Appendix

Patching

CCl 71 fol. 169r1 (B)

CCl 73 fol. 143v2 (G)

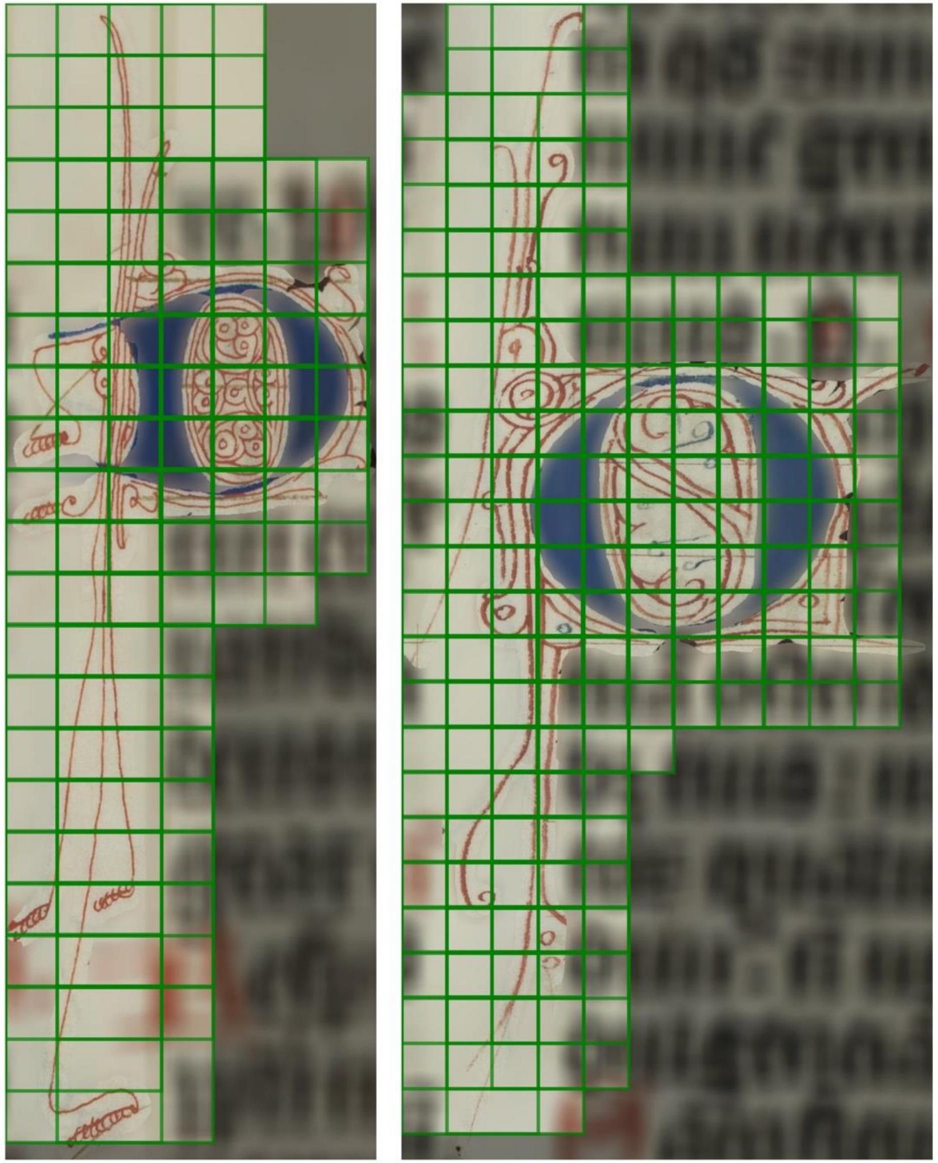


Image Source: Klosterneuburg Abbey

Similarity-Level Transparency

$$\mathbf{h}^{(i)} = \left[h_1^{(i)},\ h_2^{(i)},\ \ldots,\ h_K^{(i)}\right]^\top, \quad h_d^{(i)} \geq 0, \quad \|\mathbf{h}^{(i)}\|_2 = 1$$

$$S\big(\mathbf{h}^{(i)}, \mathbf{h}^{(j)}\big) = \langle \mathbf{h}^{(i)}, \mathbf{h}^{(j)} \rangle = \sum_{d=1}^{K} h_d^{(i)}\, h_d^{(j)}$$

$$c_d = h_d^{(i)}\, h_d^{(j)}$$

$$\sum_{d=1}^{K} c_d = S\big(\mathbf{h}^{(i)}, \mathbf{h}^{(j)}\big)$$

$\mathbf{h}^{(i)}$	Histogram for the i-th sample
S	Similarity of two histograms
K	Total number of codewords
d	Codeword (dimension)
c	Contribution score

Cluster-Level Transparency

$$\mathcal{I}_g = \{\, i : x_i \text{ has label } g\}, \quad \mathcal{I}_{\neg g} = \{\, i : x_i \text{ has label } \neq g\}$$

$$\mathcal{P}_g = \{(i,j) : i < j,\ i,j \in \mathcal{I}_g\}, \quad \mathcal{Q}_g = \{(i,j) : i \in \mathcal{I}_g,\ j \in \mathcal{I}_{\neg g}\}$$

$$\bar{c}_d^{\text{within}}(g) = \frac{1}{|\mathcal{P}_g|} \sum_{(i,j)\in\mathcal{P}_g} c_d\big(\mathbf{h}^{(i)}, \mathbf{h}^{(j)}\big)$$

$$\bar{c}_d^{\text{between}}(g) = \frac{1}{|\mathcal{Q}_g|} \sum_{(i,j)\in\mathcal{Q}_g} c_d\big(\mathbf{h}^{(i)}, \mathbf{h}^{(j)}\big)$$

$$\delta_d(g) = \bar{c}_d^{\text{within}}(g) \ - \ \bar{c}_d^{\text{between}}(g)$$

I_g, I_{-g}	Set of Images with / without group label *g*
P_g, Q_g	Within and between-group pairs
$c_d^{within}(g)$	Within-group contribution for codeword *d*
$c_d^{between}(g)$	Between-group contribution for codeword *d*
d	Codeword
$\delta_d(g)$	Discriminative score for codeword *d* and group *g*
$\mathbf{h}^{(i)}$	Histogram for the i-th sample

Codeword Assignment Balance