

# Pattern Analysis Software Tools (PAST) for Written Artefacts

Cluster of Excellence: Understanding Written Artefacts

Hamburg, Germany

Hussein Mohammed, Agnieszka Helman-Wazny, Claudia Colini, Wiebke Beyer, and Sebastian Bosch

06.06.22

# Pattern Analysis Software Tools (PAST)

## Currently Available Tools

- Handwriting Analysis Tool (HAT)
- Visual-Pattern Detector (VPD)
- Line Detection Tool (LDT)
- X-ray Fluorescence Data Analysis Tool (XRF-DAT)
- Artefact-Features Analysis Tool (AFAT)



[www.csmc.uni-hamburg.de/publications/software.html](http://www.csmc.uni-hamburg.de/publications/software.html)



# Pattern Analysis Software Tools (PAST)

## *Technical Details*

### Development details:

- Offline Web Applications
- Responsive design
- Self-contained environment
- Creative Commons Attribution-NonCommercial 4.0 International Public License

### Design details:

- Easy-to-use GUI
- Step-wise design with instructions for each step
- “How To” Section for general guidelines
- Tested and validated by experts in the CSMC

Parameter Selection

Select Keypoints Detection Algorithm

FAST

Apply

Select percentage of keypoints

10

Apply Changes

---

Currently Selected settings:

Keypoints Detection Algorithm: FAST

Main Keypoint Parameter: 10

Analyse

Progress: **Waiting to start...**

Current files: 0

1

Upload

Accepted extensions are: ".jpeg", ".jpg", ".tiff", ".tif", ".png" and ".bmp"

No name has been assigned yet !

2

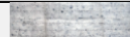
Name of Handwriting Style: Unlabelled\_2

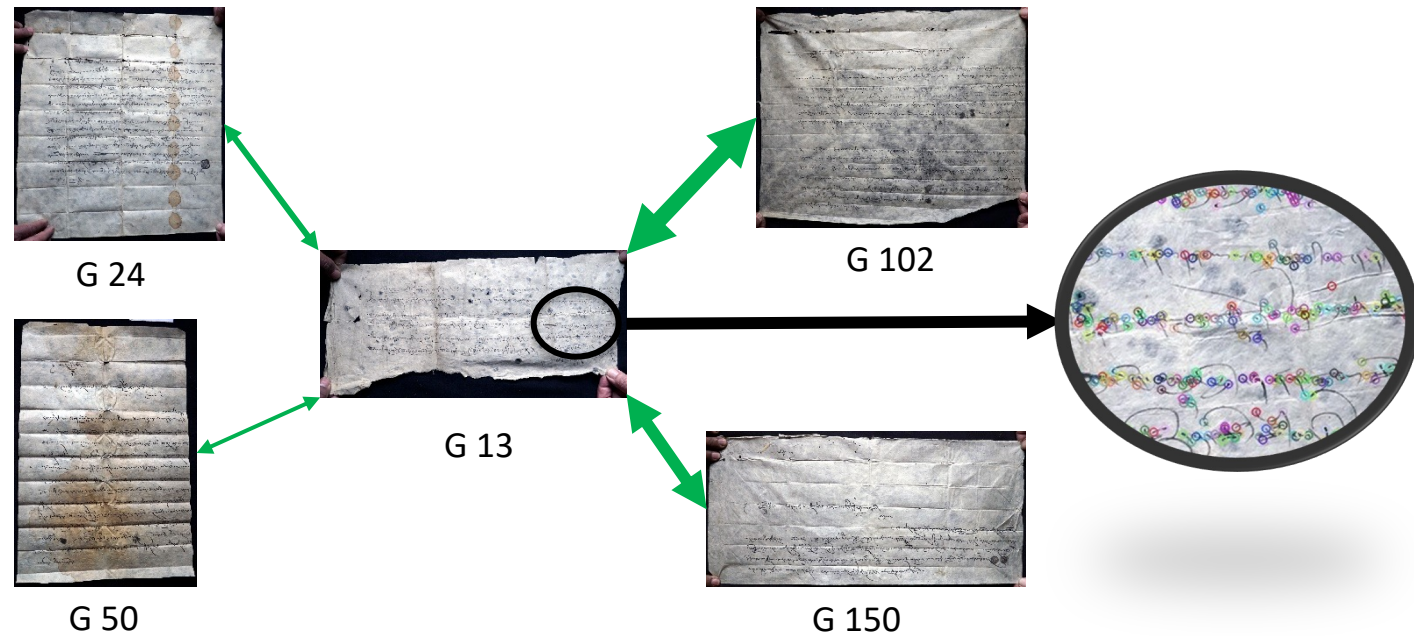
Assign

3

Add Handwriting Style

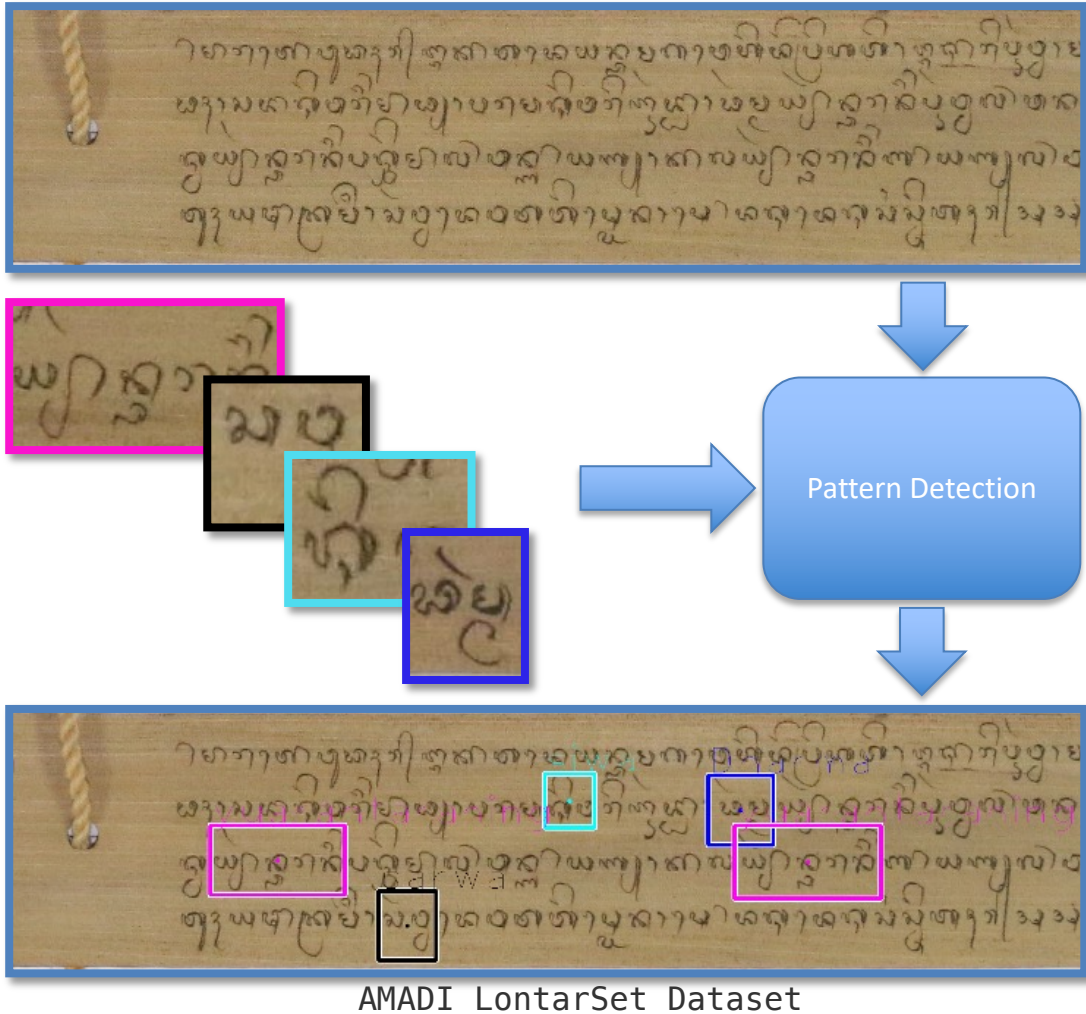
Next: Adding Labelled Images

Name of Handwriting Style	Number of Uploaded Files
G 13_KMI2876 (Example images)	1 <span style="font-size: 8px; color: #007bff;">Delete</span>
<div style="text-align: center;">             G 13_KMI2876.jpg         </div>	



- Writer Identification for Historical Manuscripts: Analysis and Optimisation of a Classifier as an Easy-to-Use Tool for Scholars from the Humanities, ICFHR, New York, USA, 2018.
- Normalised local naïve bayes nearest-neighbour classifier for offline writer identification, ICDAR, Kyoto, Japan, 2017.

# Visual-Pattern Detector (VPD)



Current files: 0      No name has been assigned yet!


1 **Upload**

Accepted extensions are: ".jpeg", ".jpg", ".tiff", ".tif", ".png" and ".bmp"

Pattern Name:       2 **Assign**


3 **Add Pattern**

**Next: Adding Data Images**

Pattern Name	Example	Number of Uploaded Images	
Pattern_1		1	Delete




**Apply Changes**




Drawing Progress: **Waiting to start...**

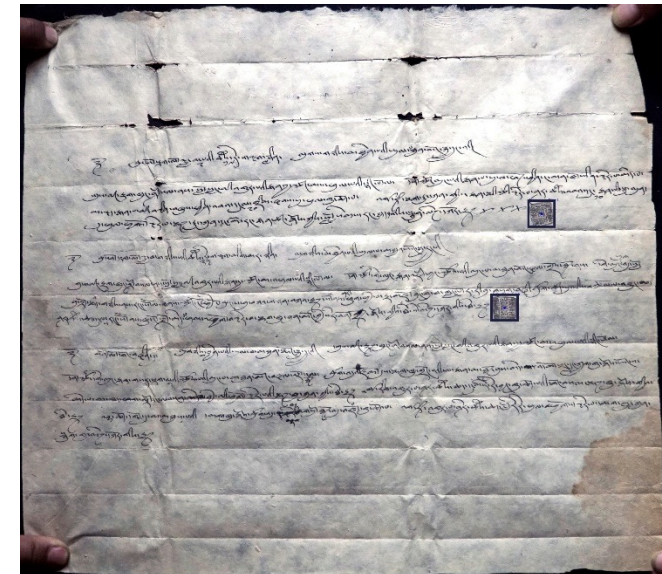
**Pattern: Pattern\_1** 

Number of Considered Detections:  **Change**

Total Detections: 1932      Considered Detections: 6

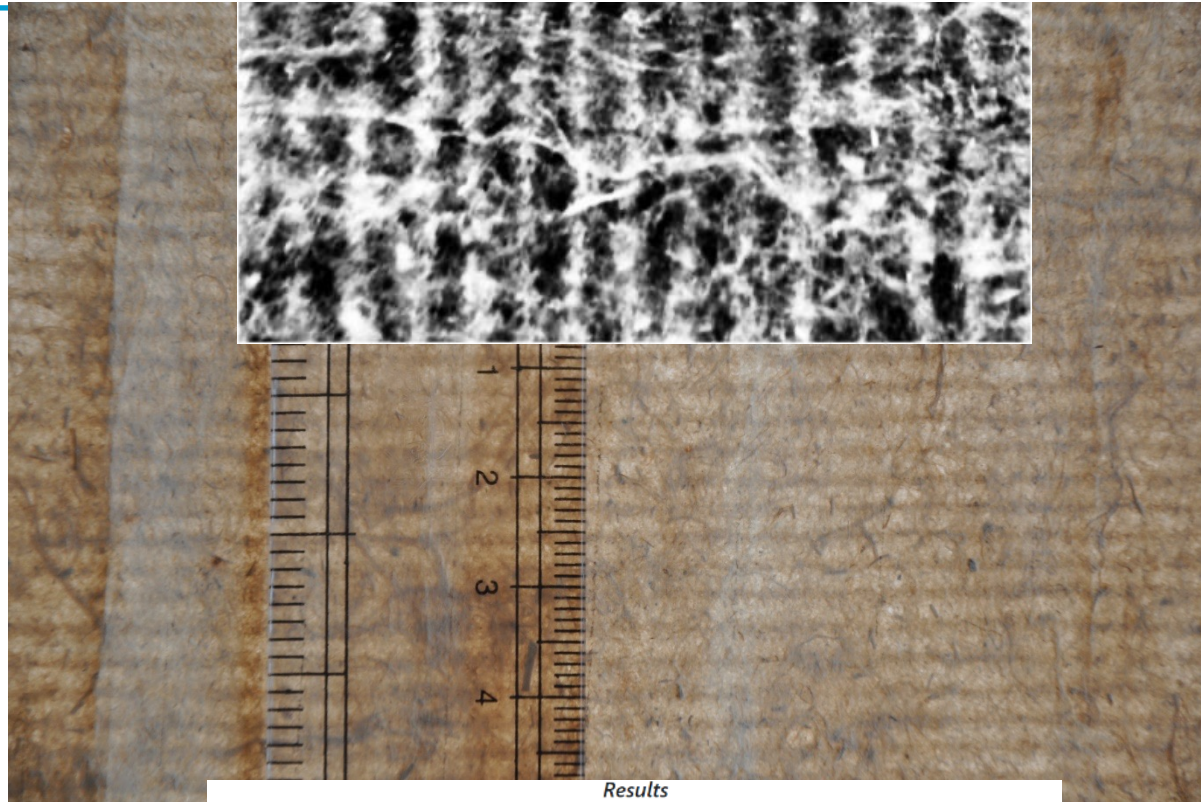
Best Detections:   

Worst Detections:   



Learning-free pattern detection for manuscript research: An efficient approach toward making manuscript images searchable, IJDAR, Lausanne, Switzerland, 2021.

# Line Detection Tool (LDT)



*Results*

Image ID	Provided Number	Estimated Number	Estimated Density	Min Spacing	Max Spacing	Accuracy	
4cm_ChinesePaper	13	13	3.25	2.48	3.14	100	<a href="#">Details</a>

Chinese Silk Roads Papers, S1731-07, Agnieszka Helman-Wazny

Average accuracy: 86.67

[Download Results \(Separated by Comma\)](#)

[Download Results \(Separated by Semicolon\)](#)

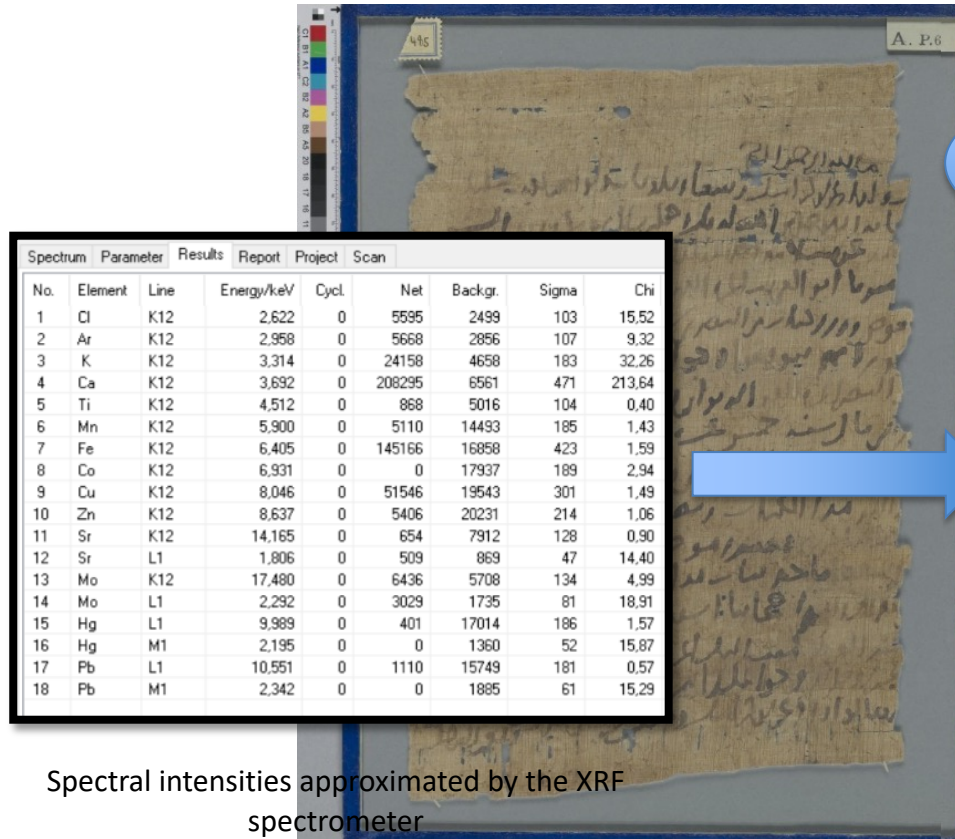
Images from sub-project RFA06, CSMC

# XRF-Data Analysis Tool (XRF-DAT)

*XRF = X-Ray Fluorescence*



XRF spectrometer

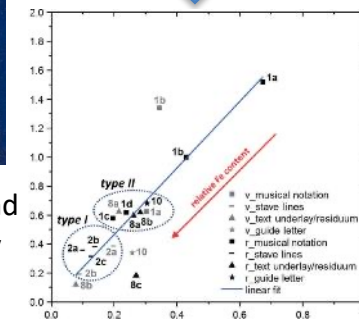


Spectral intensities approximated by the XRF spectrometer

The fragment P.Hamb.Arab.6. Copyright of Staats- und Universitätsbibliothek Hamburg Carl von Ossietzky

Preparation  
Normalisation  
Filtration

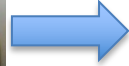
Automated  
Data Analysis



0.0034 0.0203 1.0 0.0052 0.4248 0.0343971 0.0014 0.0059



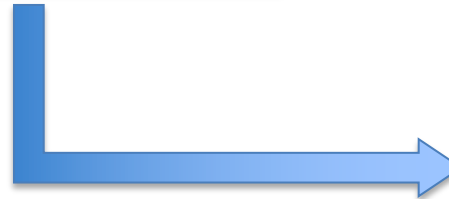
## Pattern Analysis of Sign Occurrences in Cuneiform Tablets



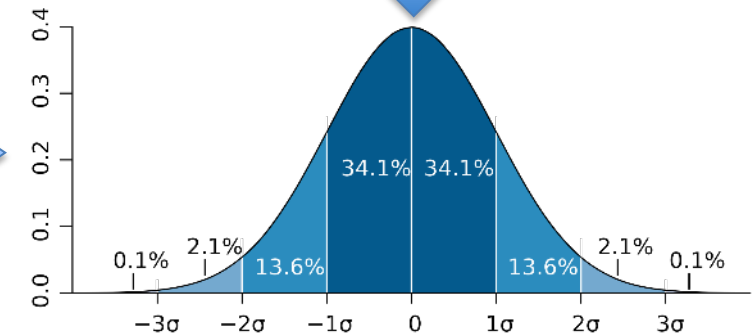
	A	B	C	D	E	F
1		Feature1	Feature2	Feature3	Feature4	Feature5
2	Artefact1	2b	1		2a	1
3	Artefact2	1a	2c	4a		1
4	Artefact3	2b	2a	1b	2a	2b
5	Artefact4	2a	1	4e	2a	1
6	Artefact5	3		4a	2a	2a



Automated Pattern Analysis



Results for Style4					
Artefact	Mean Distance	Standard Deviation	Average Mean	Average SD	Outliers
LA236	0.16	0.1	0.24	0.09	LA245
LA244	0.16	0.1			LA533
LA235	0.22	0.09			
LA245	0.34	0.08			
LA237	0.25	0.12			
LA247	0.21	0.08			
LA533	0.34	0.1			



Reproduced from: Larsen, M. T. 1. (2010). *The archive of the Šalim-Aššur family*. Ankara: Türk Tarih Kurumu Basımevi.

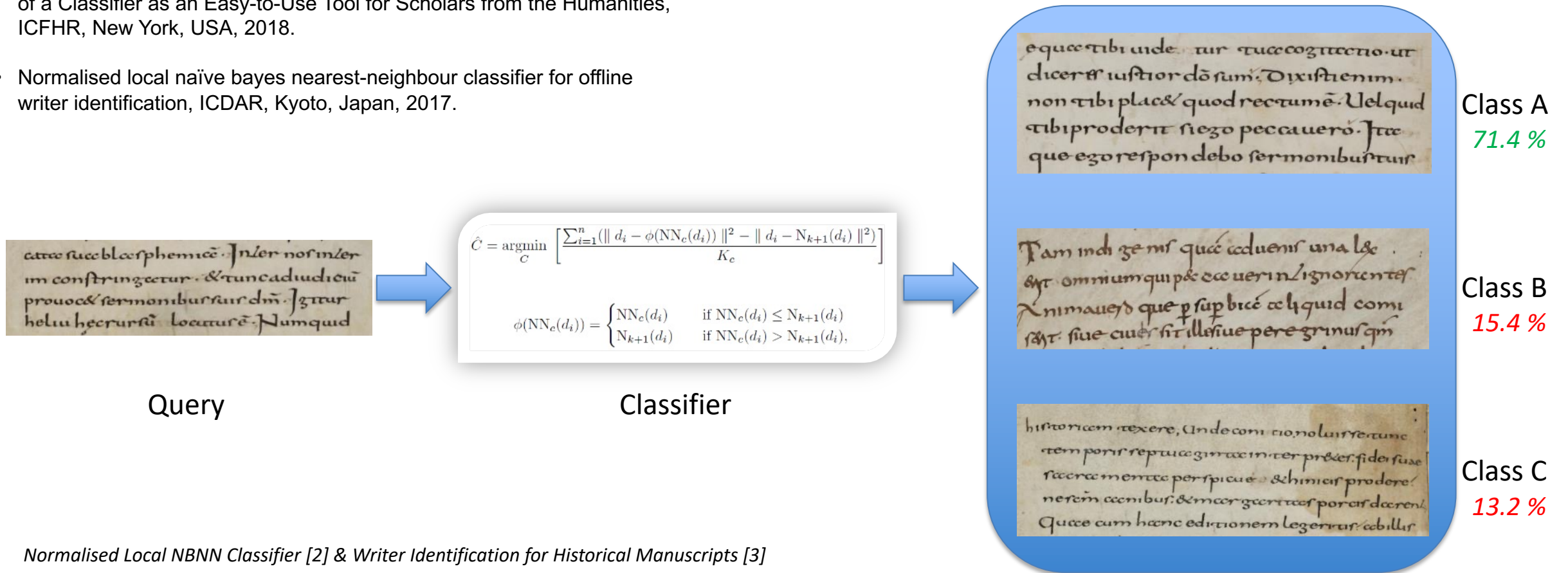




*Thank you*

# Additional Slides

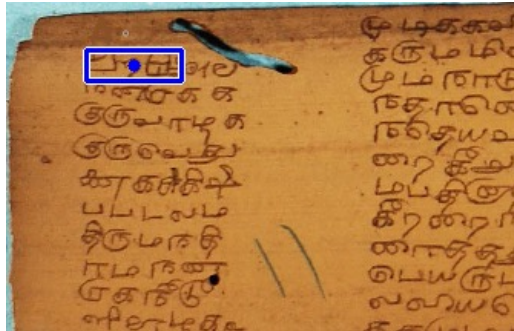
- Writer Identification for Historical Manuscripts: Analysis and Optimisation of a Classifier as an Easy-to-Use Tool for Scholars from the Humanities, ICFHR, New York, USA, 2018.
- Normalised local naïve bayes nearest-neighbour classifier for offline writer identification, ICDAR, Kyoto, Japan, 2017.



Normalised Local NBNN Classifier [2] & Writer Identification for Historical Manuscripts [3]

# Visual-Pattern Detector (VPD)

Learning-free pattern detection for manuscript research: An efficient approach toward making manuscript images searchable, IJDAR, Lausanne, Switzerland, 2021.



École française d'Extrême Orient [EFEO]



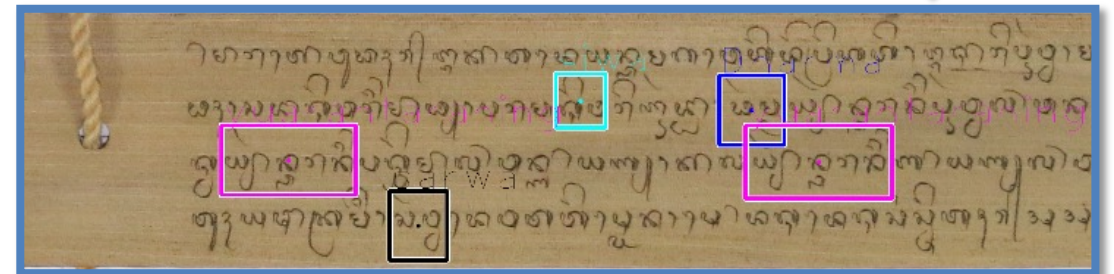
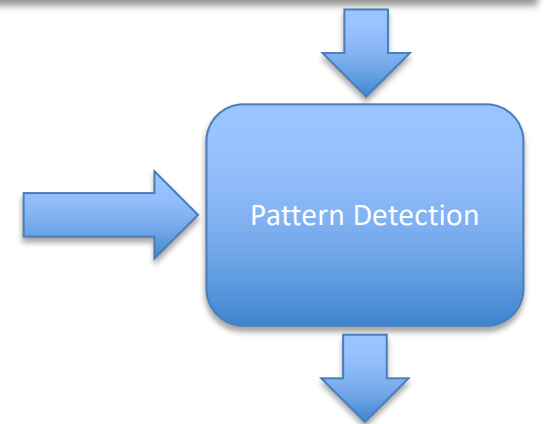
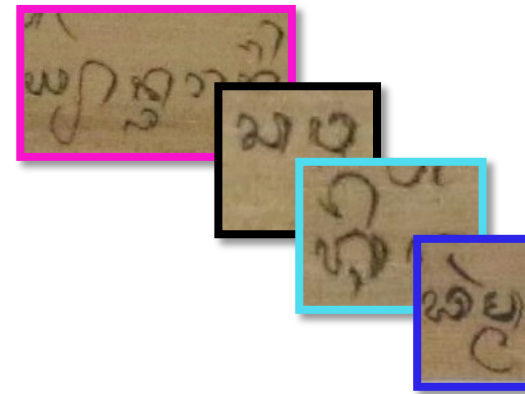
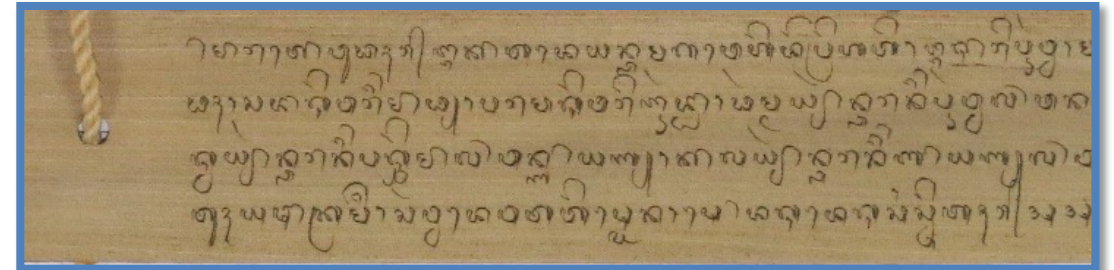
British Library: Oriental Manuscripts



DocExplore dataset

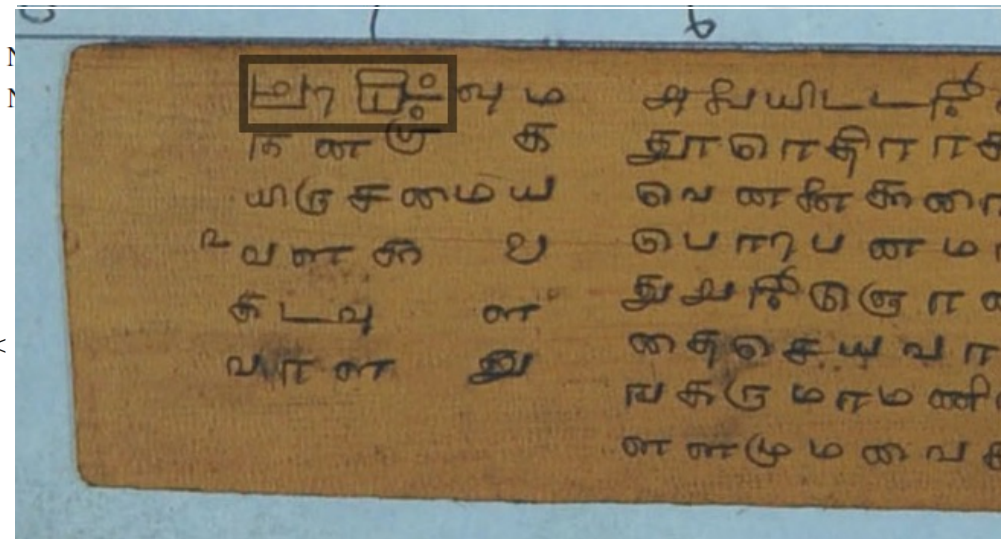
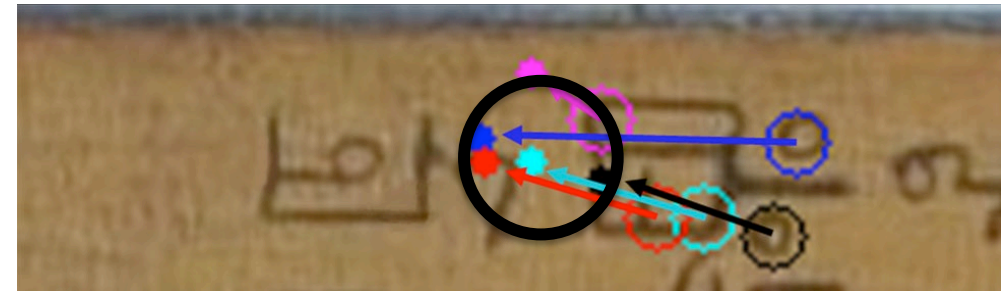
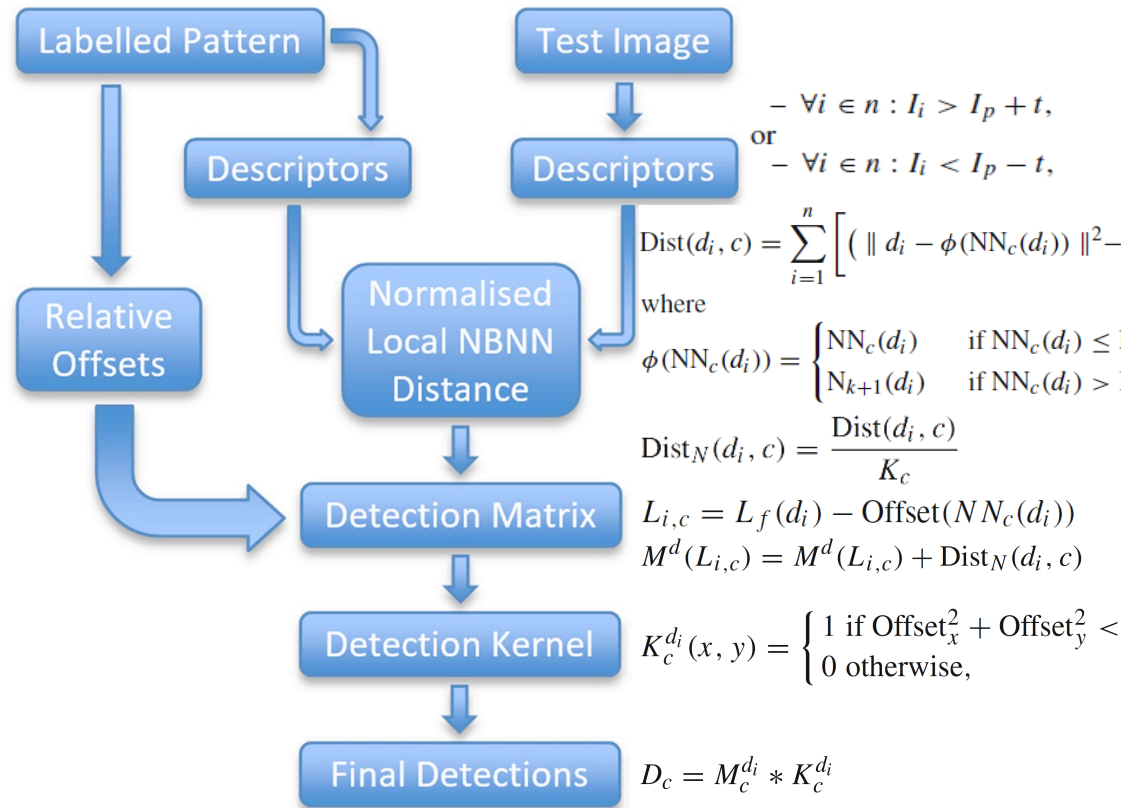


DocExplore dataset



AMADI LontarSet Dataset

# Training-Free Pattern Detection



Learning-Free Pattern Detection for Manuscript Research [6]

HAT3.5 [How To](#) [Contact](#) [About](#)



HAT3.5



## Handwriting Analysis Tool v3.5 (HAT3.5)

[Start the Analysis](#)

This software tool has been developed by [Dr. Hussein Mohammed](#) as a part of sub-project RFA05: [Pattern Recognition in 2D Data from Digitised Images and Advanced Acquisition Techniques](#).

The main goal of this software tool is to analyse handwritten documents (or manuscripts) in order to measure the similarities between different handwriting styles. A similarity score will be produced for each of the different handwriting styles (scribes) so that the user can have a relative comparison between the similarities of handwriting styles with respect to specific handwriting styles (possibly from unknown scribes). The analysis results can be useful for:

- Providing quantitative similarity measurements as a supporting information.
- Providing additional support for the work-flow of scholars in manuscript research.
- Identifying certain scribes, handwriting styles and schools of writing.
- Dating manuscripts based on the gradual change in handwriting styles.

© 2021 - HAT3.5

Visual-Pattern Detector [How To](#) [Contact](#) [About](#)



VPD



## Visual-Pattern Detector v1.3 (VPD1.3)

[Start Detecting Patterns](#)

This software tool has been developed by [Dr. Hussein Mohammed](#) as a part of sub-project RFA05: [Pattern Recognition in 2D Data from Digitised Images and Advanced Acquisition Techniques](#).

The main goal of this software tool is to automatically recognise and allocate visual patterns (such as words, drawings and seals) in digitised manuscripts. The recall-precision balance of detected patterns can be controlled visually, and the detected patterns can be saved as annotations on the original images or as cropped images depending on the needs of users.

Please pay attention to the following limitations:

- The required computational resources depend on your data. Detecting several patterns concurrently can exhaust your computer memory, and searching in huge datasets is limited by your storage capacity and your browser upload settings.
- This version provides best results of pattern detection when applied to datasets with relatively similar image qualities, such as resolution, degradation level, rotation, etc. an examples of such datasets are manuscripts that have been digitised under similar conditions and using similar equipments. Future releases will provide higher tolerance levels with respect to these aspects.
- This version has a limited capacity of benefiting from a large number of examples for the same pattern. Future releases will generate pattern models adaptively so that it gets better with every additional pattern example.

© 2021 - VPD