

Handwriting Analysis Tool v2.0 (HAT-2)

User Manual

Contact for support:

Mohammed, Hussein Adnan, M.Sc.

Scientific Service Project Z03 - CSMC

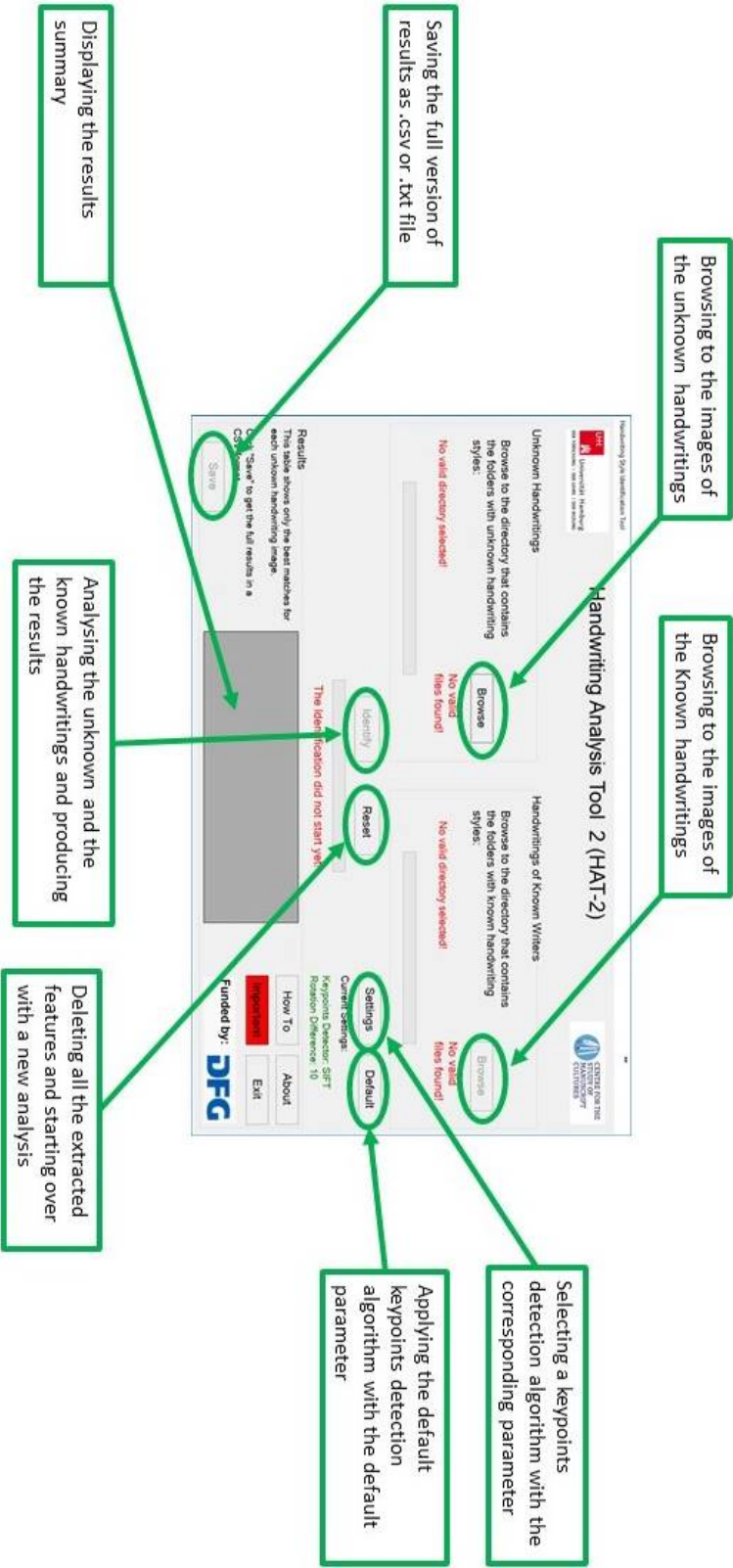
Address: Warburgstr. 26, 20354 Hamburg

Tel: +49 40 42838 - 9397

Fax: +49 40 42838 - 4899

Email: husein.adnan.mohammed@uni-hamburg.de

General Overview of HAT-2



What is HAT?

HAT-2 is a software tool that can be used to analyse handwriting styles. Several different handwriting styles (scribal hands) can be analysed concurrently and sorted according to their similarity to a questioned or unknown style (query). A similarity score will be calculated for each predefined style (scribal hand) to create a relative comparison between them with respect to an unknown style.

The handwriting styles do not need to be known, only discriminative labels must be used for each different style (hand).

This tool is developed by Hussein Mohammed, researcher in sub-project Z03 of CSMC at Universität Hamburg and is based on the method presented in the following publication:

This tool is developed by Hussein Mohammed, Msc from the Z03 sub-project for CSMC at Hamburg University, and it is based on the method presented in the publication:

H. Mohammed, V. Märgner, T. Konidakis, and H. S. Stiehl, "Normalised Local Naive Bayes Nearest-Neighbour Classifier for Offline Writer Identification", in Document Analysis and Recognition (ICDAR), 14th International Conference on. IEEE, 2017, Kyoto, Japan.

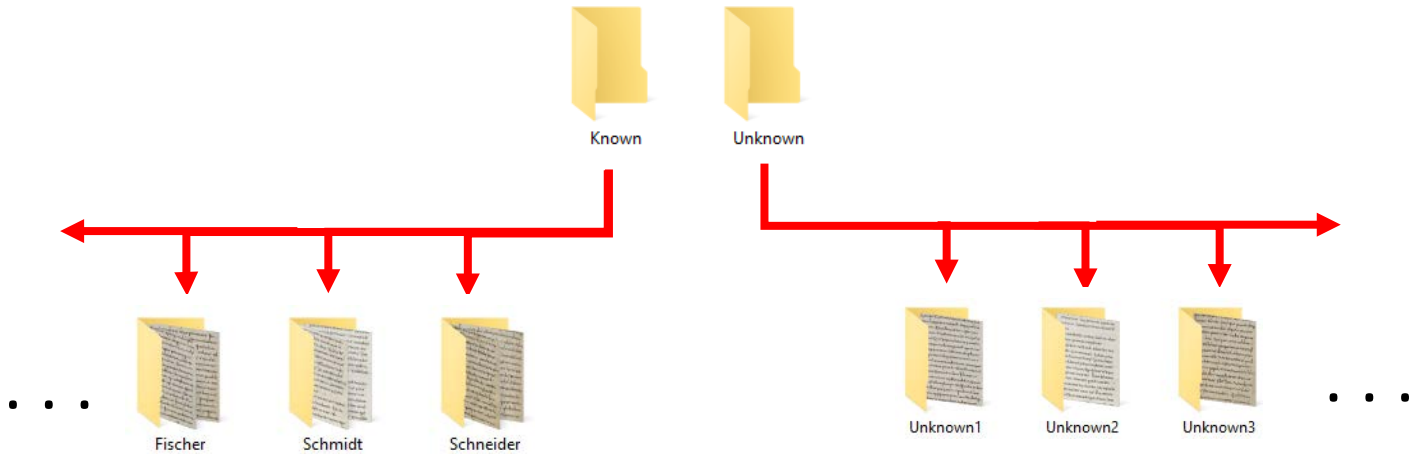
The recommendations in the software manual are partially based on the publication:

Accepted paper: H. Mohammed, V. Märgner, and H. S. Stiehl. Writer Identification in Historical Manuscripts: Analysis and Optimisation of a Classifier with an easy-to-use Implementation for Scholars from the Humanities. In 2017 16th IAPR International Conference on Frontiers in Handwriting Recognition (ICFHR).

HAT-2 is licensed under the Creative Commons Attribution-NonCommercial 4.0 International Public License



Required Directory Structure



- The name of the folders can be anything you like, as far as you can distinguish between them. Names used in the manual are arbitrary and for illustration purposes only.
- You can add as many subfolders as you want in both known and unknown handwritings directories. With as many images as you want in each subfolder.
- You can test several unknown handwritings simultaneously.
- In the directory of unknown handwritings, all images within the same subfolder will be treated as one query (one image). This is particularly useful when dealing with a heavily degraded or fragmented piece of handwriting; you can crop the best parts and save them as individual images, then you put them in the same subfolder as one query.

Usage Procedure (Users)

1. Browse to the folder that contains unknown handwriting styles (each style must be stored in a separated subfolder).
2. Browse to the folder that contains known handwriting styles (each style must be stored in a separated subfolder).
3. Click (Analyse) to analyse the known and unknown handwriting styles and produce results.
4. You can check the brief version of the results in the (Results) table, or you can save the complete version of the results to a file.

Usage Procedure (Experts)

1. Select the desired keypoints detection algorithm and enter the corresponding parameter from (Settings).
2. Browse to the folder that contains unknown handwriting styles (each style must be stored in a separated subfolder).
3. Browse to the folder that contains known handwriting styles (each style must be stored in a separated subfolder).
4. Click (Analyse) to analyse the known and unknown handwriting styles and produce results.
5. You can check the brief version of the results in the (Results) table, or you can save the complete version of the results to a file.

Settings (for expert users only)

Settings

You can select different keypoints and parameters from the settings:

Select Keypoints Detection Algorithm:

☐ SIFT Rotation-Difference Threshold: °

☐ FAST Percentage of Keypoints: %

From the settings, you can choose between two different keypoints detection algorithms. In addition, you can modify the corresponding parameters of the algorithms.

1. Scale Invariant Feature Transform (**SIFT**):

This keypoint detection algorithm can cope better with images of different resolutions; furthermore, you can specify the amount of rotation that can be tolerated between images. Rotation-Difference of 10 degrees is typically enough.

- Allowed values: Integers between (0 – 90)

2. Features from Accelerated Segment Test (**FAST**):

This keypoint detection algorithm can cope better with very low-contrast or very low-resolution images; furthermore, you can specify the percentage of keypoints to be considered in the analysis. Only the top XX% keypoints with the highest response will be considered. This option can greatly speed up the processing time, which could be of high importance when dealing with large collection of manuscripts. The recommended value for this parameter depends on the signal to noise ratio (S/N); in other words, it depends on the ratio of the relevant information to the irrelevant information with respect to the targeted handwriting. For handwritings on any heavily degraded non-contemporary writing materials such as parchments, the recommended value can be as low as 1%. In general, 10% is suitable for most manuscript images.

- Allowed values: Decimals between (1.0 – 100.0)

Results

- This tool produces two versions of the results:
 - Summary: a brief version of the results is displayed automatically as a table in the tool window. This summary shows only the best match for every unknown handwriting.
 - Full: a complete version of the results can be obtained by saving the results to a file. You can save it in a (.csv) or (.txt) format. In the saved file, you can find all the styles ranked according to their similarity to the unknown handwriting.
 - The scores are a relative measure of similarity. The tool calculates how similar an unknown style is to a given known style relative to the other known styles.
 - The scores of all different styles in the complete version of the results sums up to 100.

Example of summary results:

	File	Best Match	Score
▶	Unknown1	Fischer	71.3
	Unknown2	Schmidt	80.7
	Unknown3	Schneider	48.1
*			

Example of a Full results file:

	A	B	C
1	Results for Unknown1		
2	Rank	Directory	Score
3	1	Fischer	71.3
4	2	Schneider	15.3
5	3	Schmidt	13.2
6			
7	Results for Unknown2		
8	Rank	Directory	Score
9	1	Schmidt	80.7
10	2	Fischer	10.4
11	3	Schneider	8.7
12			
13	Results for Unknown3		
14	Rank	Directory	Score
15	1	Schneider	48.1
16	2	Fischer	31.8
17	3	Schmidt	19.9

Technical Considerations

- **Required system (platform):** Windows (x64 and x32)
- **Supported file extensions for input images:** jpg/jpeg, tif/tiff, png, bmp
- **Results formats:**
 - (.csv) files, use any spreadsheet application like Microsoft Excel
 - (.txt) files, in a plain text format
- The (**Known**) folder must contain at least 2 subfolders for handwritings from different styles (scribes)

Important Remarks

- In the directory of Known handwriting styles, the name of the subfolder will be used as the name of the style for the images in that subfolder.
- You can click (Reset) to start over, but make sure you save your results if you want to keep them!
- Remove irrelevant information when possible. This can be done simply by cropping and keeping the desired handwritings in the image only.