Demonstrator: Image Retrieval and Visualization

Deliverable D2.6

Version FINAL

Negotiating Olfactory and Sensory Experiences in Cultural Heritage Practice and Research

The Odeuropa project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 101004469. This document has been produced by the Odeuropa project. The content in this document represents the views of the authors, and the European Commission has no liability in respect of the content.

Grant Agreement No.	101004469
Project Acronym	ODEUROPA
Project full title	Negotiating Olfactory and Sensory Experiences in Cultural Heritage Practice and Research
Funding Scheme	H2020-SC6-TRANSFORMATIONS-2020
Project website	http://odeuropa.eu/
Project Coordinator	Prof. Dr. Inger Leemans KNAW Humanities Cluster Email: inger.leemans@huc.knaw.nl
Document Number	Deliverable D2.6
Status & version	FINAL
Contractual date of delivery	31 December 2023
Date of delivery	30 December 2023
Type	Demonstrator
Security (distribution level)	Public
Number of pages	14
WP contributing to the deliverable	WP2
WP responsible	WP2
EC Project Officer	Hinano Spreafico
Authors: Mathias Zinnen[1] and Vincent Christlein[1]	
Internal reviewers: Inna Novalija[2] and Sara Tonelli[3]	
Affiliations (1) FAU, (2) JSI, (3) FBK	
Keywords: computer vision, image processing, demonstrator	

Abstract: This document reports on the actual deliverable D2.6: two demonstrator tools to explore the visual smell references extracted by the image processing team of the Odeuropa project (WP2). The Images Nosebook facilitates the exploration of visual information extracted from nearly 100k artworks from various museum collections. It serves as a user-friendly gateway for scholars and laymen alike, promoting the analysis of visual smell references. The Image Processing Live Demo complements this by providing users with the ability to test the detection systems on their own images, offering a swift understanding of the underlying image processing systems. With the combination of these tools, we aim to open our results to a broad public and bridge the gap between technical and non-technical users, allowing scholars of various disciplines to engage with the fascinating history of smells as conveyed through visual culture.

Table of Revisions

Version	Date	Description and reason	By	Affected sections
0.1	December 14, 2023	Draft	Mathias Zinnen	All
0.2	December 15, 2023	First review	Inna Novalija	All
0.3	December 15, 2023	Revision after Review	Mathias Zinnen	Abstract, Executive Summary, Images Nosebook
0.4	December 18, 2023	Second review	Sara Tonelli	All
0.5	December 18, 2023	Revision after Review	Mathias Zinnen	Images Nosebook, Image Processing Live-Demo
0.6	December 20, 2023	Final Revision	Vincent Christlein, Mathias Zinnen	all
1.0	December 30, 2023	Final check and approval by project manager	Marieke van Erp	-

Executive Summary

Challenges & Barriers

- We wanted the demonstrators to be accessible to non-technical audiences while retaining the flexibility to programmatically change the functionality. We tried to meet this requirement by implementing the Images Nosebook using the interactive Jupyter environment and using abstractions over the code. Finding the right balance between the needs of technical and non-technical users remains challenging.
- Image Data has a low information density when compared to texts or tabular data which leads to large memory requirements. However, to explore the visual data generated during the project, it is necessary to store nearly 100k images on the hardware on which the Images Nosebooks are executed (Colab instance). As Colab instances do not provide persistent storage, the images have to be downloaded before running the notebook. To this end, we implemented an initial download from Zenodo[1], a storage service for publicly available datasets. By compressing the images we were able to decrease the time needed for the initial download to 10-15 minutes. Simultaneously, we display links to the source collection with each artwork to enable users to explore the images in full resolution. While we would have preferred a faster solution, it is still the best possible way to operate with the image data under the current circumstances.

Practices & Guidelines

- The envisioned functionality of the Nosebook was collaboratively designed with multiple work packages from the Odeuropa project. While the design process was inclusive, recognising the diverse perspectives of a wider group of scholars could have further enriched the application. In hindsight, initiating the process by sending out questionnaires to potential users at the start of development and soliciting their feedback would have been a proactive and effective strategy to gather valuable feedback on specific functions. Looking ahead, establishing a streamlined process for users to share their input on desired functionalities would be a beneficial enhancement for future use.
- One successful approach was presenting the audience at the final event with the key functionalities of the demonstrators and offering them the opportunity to engage with the demonstrators independently after the event. The exposure to the technical capabilities already sparked numerous questions and discussions. We anticipate that these discussions can be continued, either virtually between members of the PastScent network or in the context of academic publications and conferences.

Layman Summary This report presents deliverable 2.6: two demonstrator tools for the image processing strand of the Odeuropa project:

- *The Images Nosebook* to explore the visual information extracted from artworks during the project, and
- *the Image Processing Live-Demo* to try out the visual smell recognition systems.

Subsequently, functionalities and design decisions for both tools are reported.

[1] https://zenodo.org

Contents

1 Introduction

Interdisciplinary and broad accessibility are among the main aims of the Odeuropa project. Thus, we want the image analysis tools and visual olfactory references we generated during the project period to be accessible to a broad public, including humanities-based scholars without technical backgrounds or interested laypeople. The image demonstration tools described in this deliverable were designed to adhere to this goal, while at the same time providing as much functionality and flexibility as possible. In the following, we will describe two demonstrators:

- *The Images Nosebook*[2] which enables users to query nearly 100k artworks from various museum collections processed by our detection systems, and
- *the Image Processing Live-Demo* which lets users upload arbitrary images to try out the detection systems on their own.

Both tools will be introduced and further discussed in the following sections.

2 Images Nosebook

2.1 Deployment

The Images Nosebook is a jupyter-based notebook which provides an accessible interface to the visual olfactory references extracted during the project. It is available online via https://bit.ly/odeuropa-images.

Jupyter[3] is a web-based framework to interactively run Python code in the form of notebooks in a browser. We could have provided a similar functionality by implementing a web app using javascript or python frameworks such as gradio[4] or flask[5] but favoured the jupyter-based solution for the following reasons:

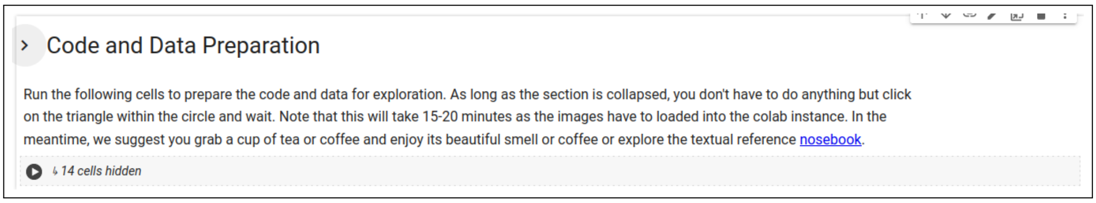
- *Code accessibility*: While jupyter notebooks allow the implementation of user-friendly interface elements that allow usage without touching any code, they still provide an easy way to read and modify the code on which these abstractions are based. Figure 1 shows how these different layers of abstraction can be accessed in the demonstrator interface.
- *Simplicity & Uniformity*: As most of our code base was already written in Python (and, in some cases, even in the form of Jupyter Notebooks), keeping the programming language consistent avoids errors that could be introduced during framework or language transition.

More specifically, the demonstrator runs on Google Colab,[6] an online environment to write, host, and publish Jupyter notebooks, which are executed on hardware provided by Google (Colab instances). Each time a user accesses the notebook, their own local copy will be created. This prevents them from changing other users' application state. While the Colab solution has the disadvantage that it requires a Google account to use the demonstrator, we still favour this approach over exposing the notebooks on our own servers for scalability and security reasons.

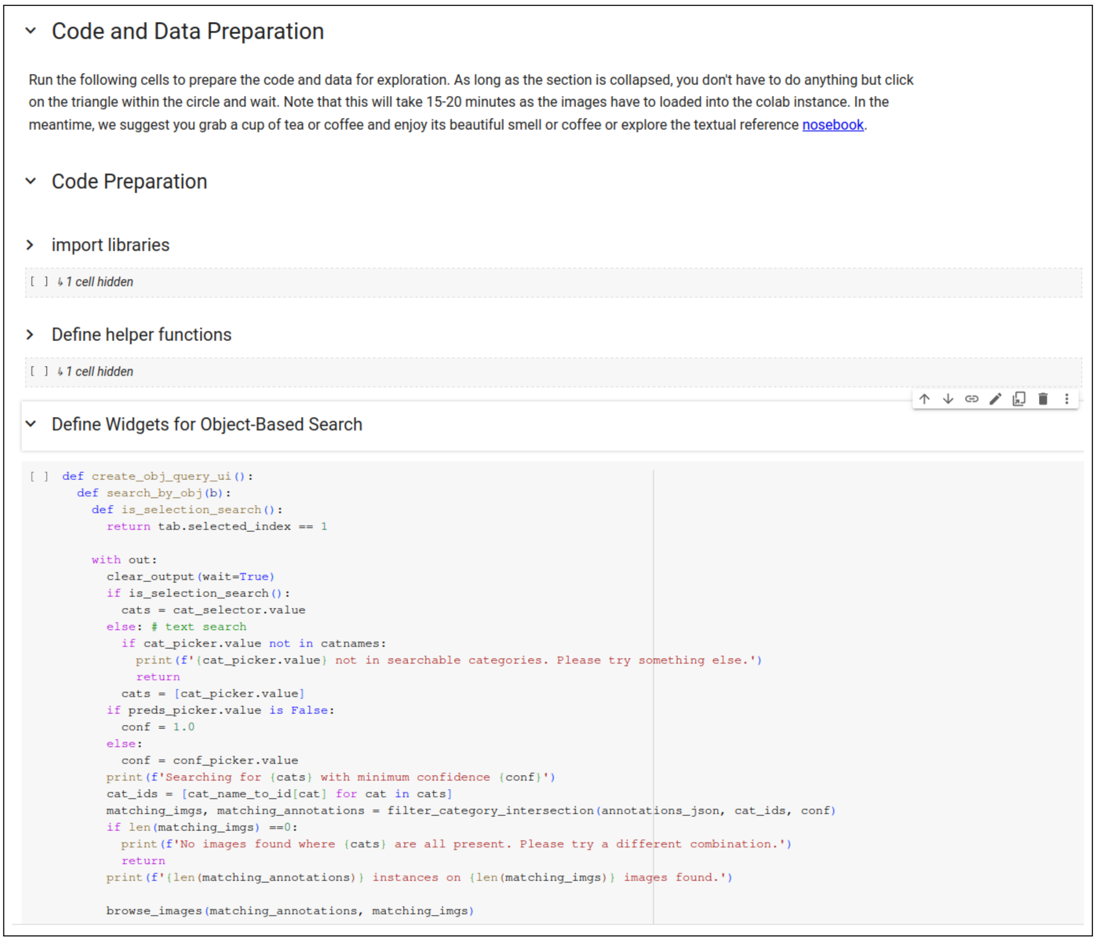
2.2 Functionality

The notebook consists of three chapters : (1) Code and Data preparation, (2) Query Image Data, and (3) Quantitative Analysis.

[2]portmanteau word from nose and notebook
[3]https://jupyter.org
[4]https://gradio.app
[5]https://flask.palletsprojects.com
[6]https://colab.research.google.com

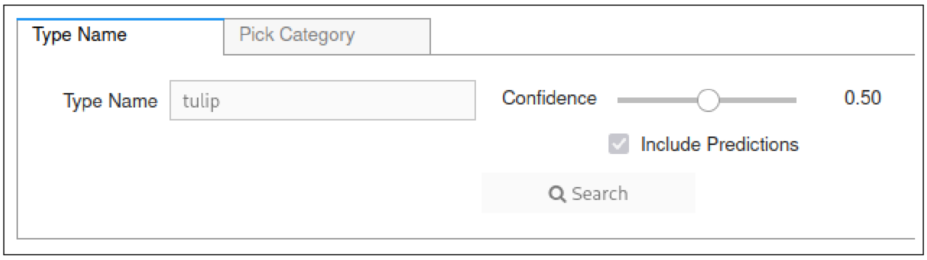


(a) Collapsed Cells

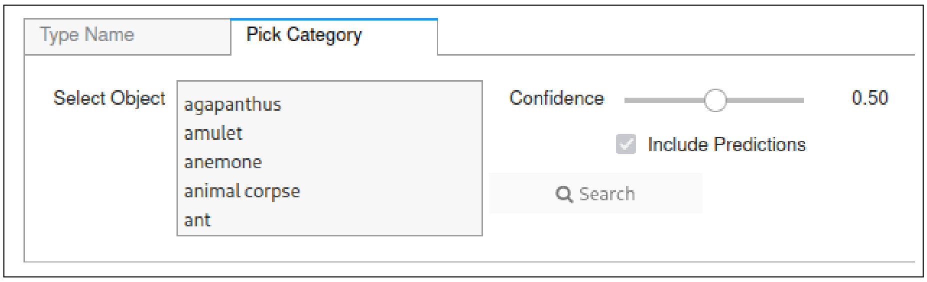


(b) Expanded Cells

Figure 1: Screenshot from the Images Nosebook illustrating different levels of code-abstraction. In Fig. 1a the code cells are hidden from the user, and the data preparation can be performed by simply clicking on the circle next to the hidden code cells. Figure 1b shows the result of expanding the code cells. Coding-affine users can expand code cells and investigate technical details or modify code according to their needs.



(a) Type Name tab of the object query widget



(b) Pick Category tab of the object query widget

Figure 2: Screenshot from the object query widget of the Images Nosebook illustrating the different query modes.

Code and Data Preparation This section provides specific technical information regarding data preparation. Here, preparatory steps like downloading the data to the Colab instance, defining functions, and creating data structures are performed. The data that is being downloaded consists of the images, collected from various museum collections as described in [Zinnen and Christlein, 2022a, Zinnen, 2022], and the results of the visual smell reference extraction system described in [Zinnen and Christlein, 2022b]. It is collapsed by default, meaning that all of its internal functionality can be executed by a simple click on the circular `run cell` button next to the cell (Fig. 1a). However, as described above, interested users are invited to expand the code cells and explore the technical details (Fig. 1b).

Query Image Data In the *Query Image Data* chapter, users can explore the data in two ways: either following smell-related objects and gestures that we found in the artworks, or by the image meta-data provided by the source collections.

Figure 2 displays the various query modes supported in object-based search. To select the objects to search for, they can either be entered in a textbox (Fig. 2a), or selected from a complete list of processed objects (Fig. 2b). The latter mode also allows to query for artworks with a combination of objects where all selected objects must be present. Furthermore, users can define a minimum confidence using the `confidence`-slider. The confidence score represents the models' certainty of having found the query object. Note that these scores are not consistent across categories, i. e., a high score for an object that was rare in the training set might still not guarantee reliable predictions. Some users might not want to rely on automatically found objects at all and only want to see manual annotations. This can be achieved by unticking the `include predictions` checkbox.

Figure 3 shows how the results matching the defined parameters are displayed once the `search` button is clicked.

Additionally, users can query the Odeuropa image data based on image-level metadata provided by the respective source collections.[7] For these queries, no image processing re-

[7]For a complete listing of available metadata fields, please refer to Deliverable 2.4 [Zinnen, 2022, p. 7-8].

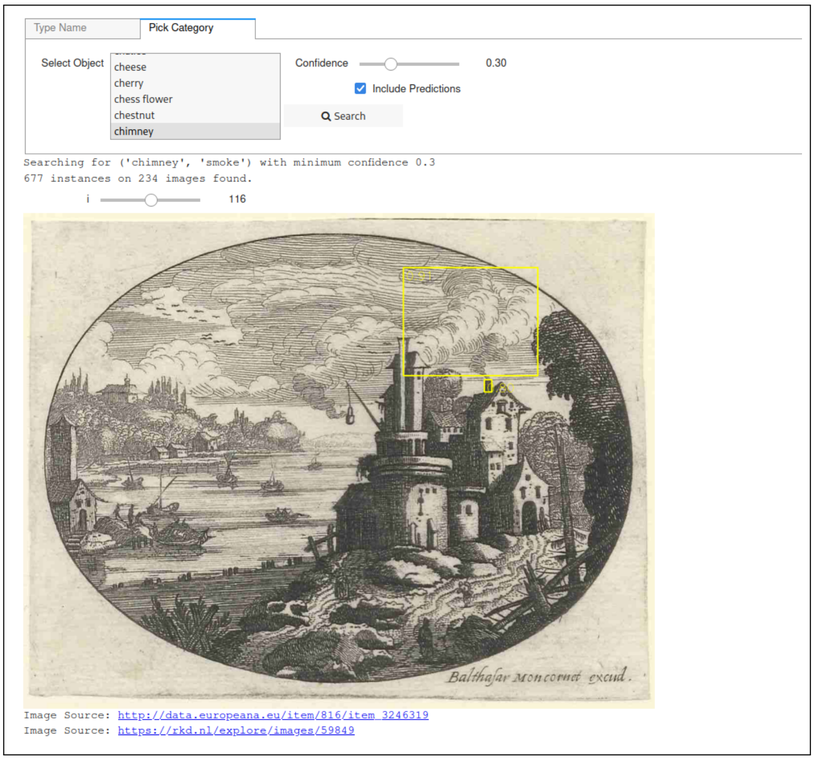


Figure 3: Screenshot from the object query widgets displaying the results of a query. The slider at the bottom can be used to navigate through the matched artworks.

sults are employed. Figure 4 shows the user interface of the metadata query. By switching between tabs at the top of the interface, the images can be filtered according to various metadata fields that were collected during the image gathering process described in Deliverable 2.2 [Zinnen and Christlein, 2022a] and Deliverable 2.4 [Zinnen, 2022]. Search by description and by title is based on substring matching, i. e., for an image to be returned, the search string must be a substring of the respective metadata field. Search by keyword, by artist, and by iconclass code [Couprie, 1983], on the other hand, requires exact matches of the query string and the metadata field. To simplify usage, users will get suggestions as they start typing.

Quantitative Analysis In addition to using the data for querying and conducting manual analysis of the query results, the extracted smell references can serve as a foundation for exploring quantitative approaches and research questions. Exemplary, the demonstrator provides interfaces and visualisations for two different types of statistical analysis that could be conducted on the basis of the collected smell references:

(1) *Object Co-Occurrences*. To gain a deeper understanding of the extracted smell references, one effective approach involves analyzing the co-occurrences of objects. Figure 5 shows the interface where users can display co-occurrences between objects and gestures from the extracted references. To generate a visual representation of these co-occurring objects, users can input up to three object categories in the designated fields. The ensuing Venn diagram depicts the frequency with which these objects appear together in images as opposed to their individual occurrences, offering a visual guide to their interconnections.

(2) *Distribution over Time*. Another possibility of exploring visual smell references quantitatively is by visualising their appearances over time. Figure 6 illustrates the widget which displays

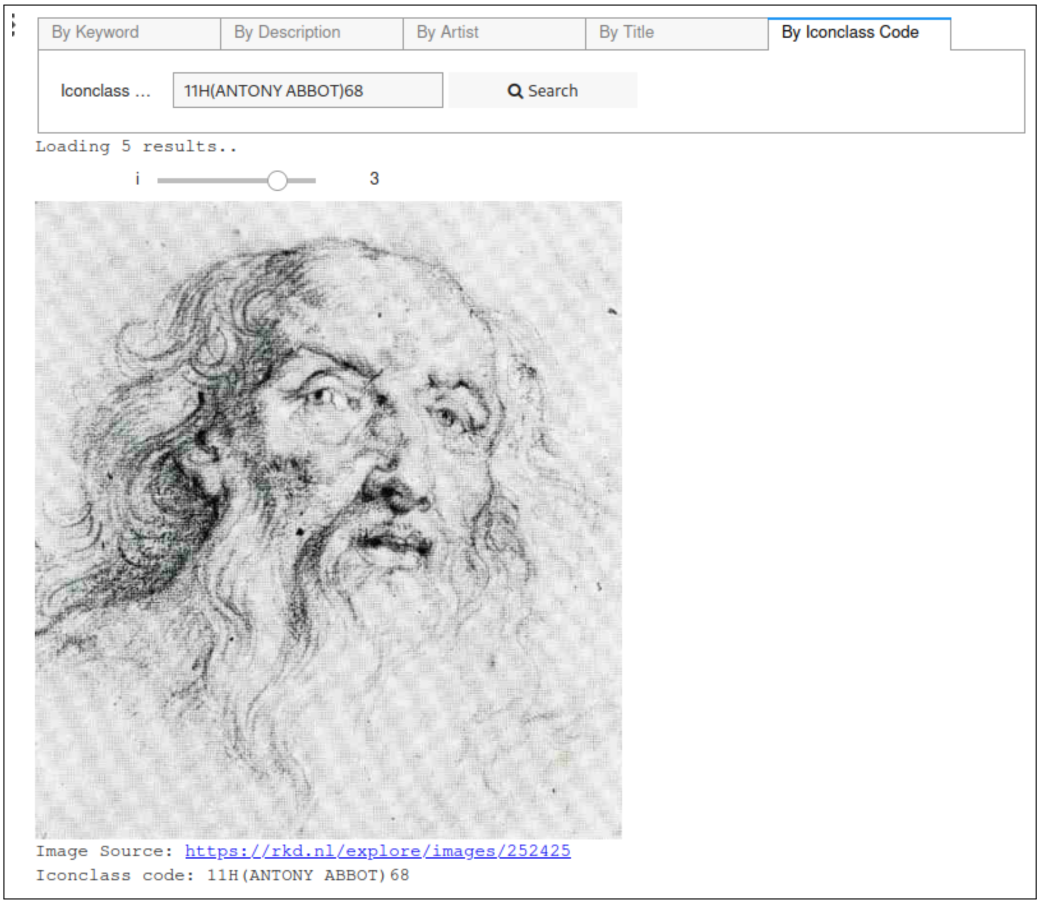


Figure 4: Screenshot of the metadata query widget. Different query modes can be obtained by selecting the tabs at the top of the interface.

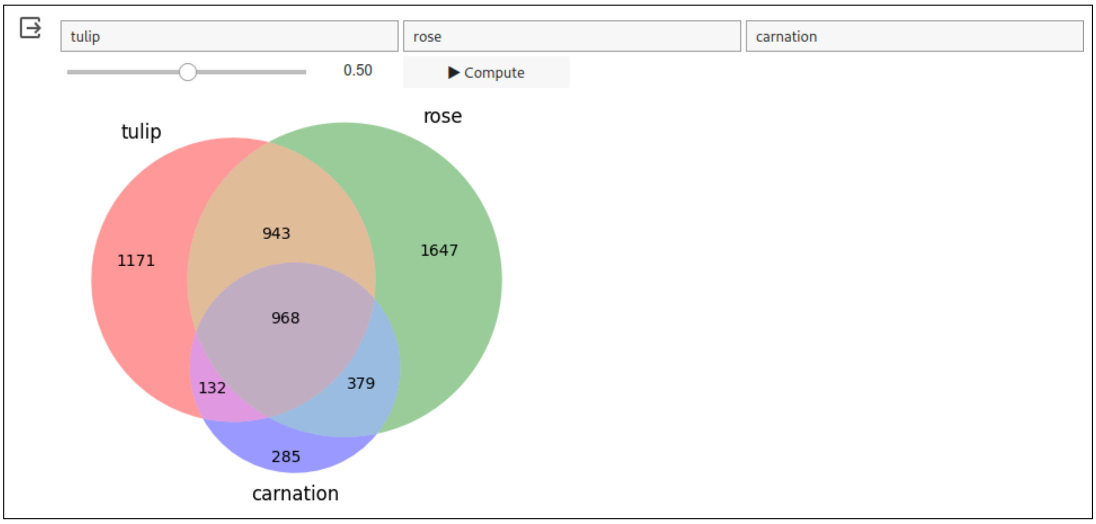


Figure 5: Screenshot from the object co-occurrences widget. Users can input up to three objects and display how often they occur together in the same image.

object occurrences throughout the time frame of the Odeuropa project image analysis (1600-1920). Users can type the object of interest into the `Object` text input field. Additionally, they can visualise the temporal distribution of one comparison object by specifying it in the `Reference` field. As the analyzed images are not equally distributed over time, the visualisation will likely be biased towards time periods for which we have more images in our dataset. To mitigate this bias, users can enable normalisation by `All Instances` or by `Supercategory`. Checking these options displays the percentage of annotations compared to all annotations, or annotations of a similar type, respectively.

3 Image Processing Live-Demo

3.1 Deployment

The Image Processing Live Demo is based on a neural network trained to detect smell-related objects and gestures on historical artworks. Using the gradio framework,[8] we implemented a user-friendly interface to the detection functionality of the network and uploaded the app on Hugging Face.[9] Similar to our reasoning for the Nosebook, we rely on a large provider of infrastructure to ensure scalability and security of the application. The deployed app is available online at https://huggingface.co/spaces/ThRi/FAU_PR_Detrex and can be accessed in the browser without the need for a local installation.

3.2 Functionality

Figure 7 displays the UI element which is presented to users entering the application. By clicking on the input field or by dragging an image into it, users can upload their own images for processing. Furthermore, they can select one of the example images to get a quick demonstration of the system without having to upload anything. Note that we currently only have one example image defined but plan to extend this in the future.

[8] https://gradio.app
[9] https://huggingface.co

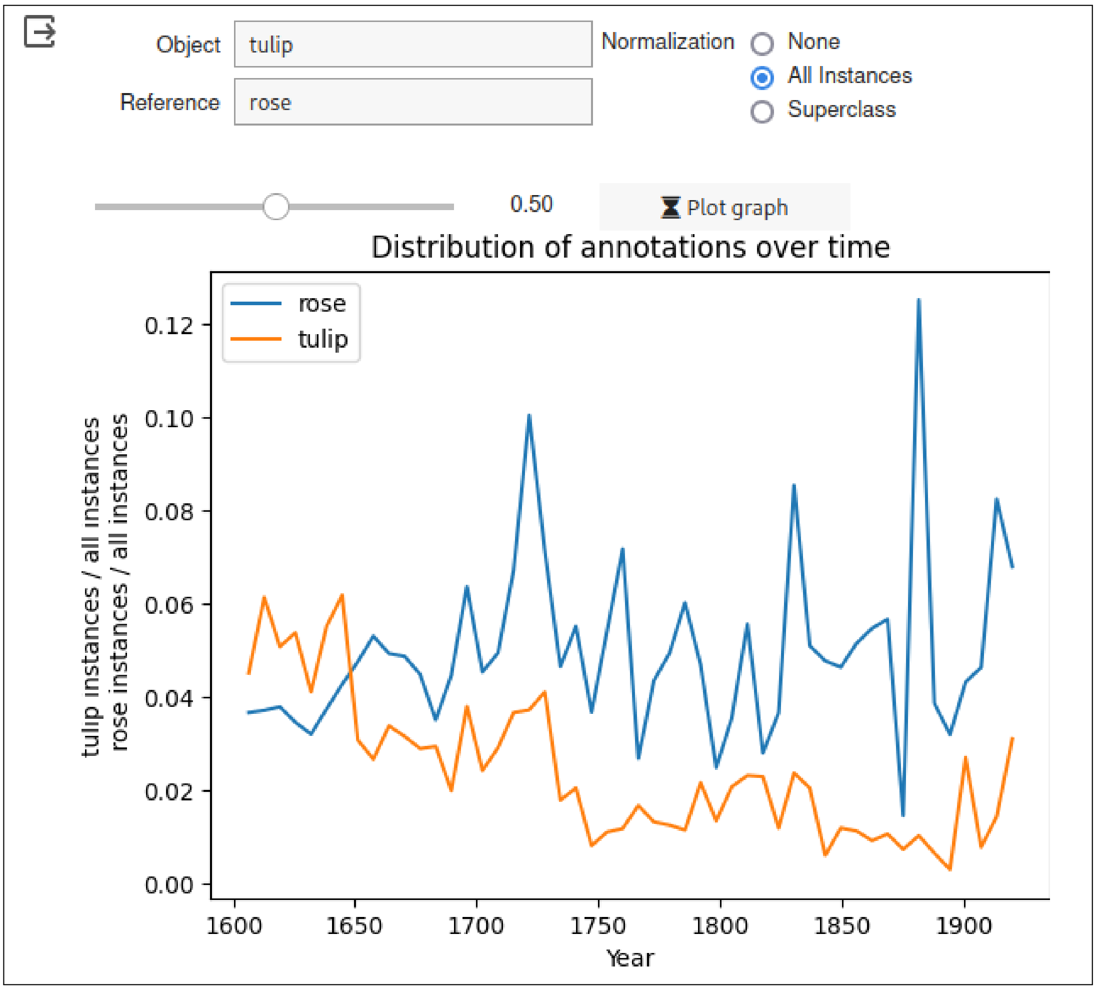


Figure 6: Screenshot from the object appearances over time widget. Users can visualise the occurrences of objects found in our data over the time frame of the Odeuropa project.

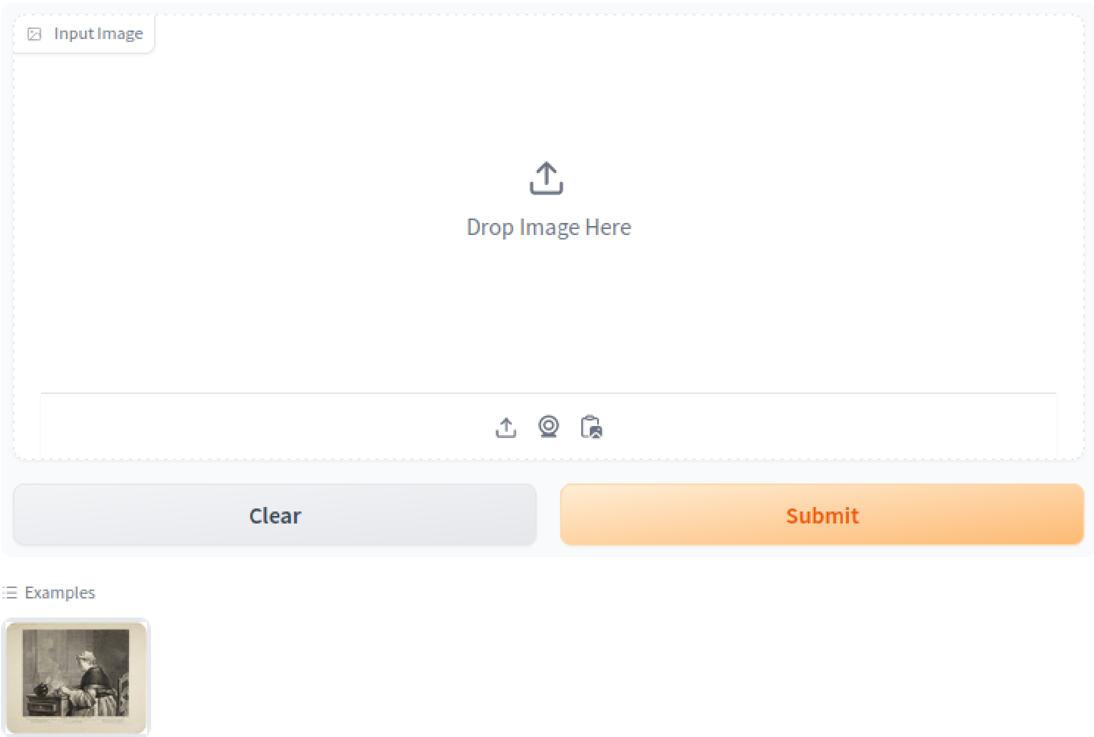


Figure 7: Screenshot from the Image Processing Demonstrator application displaying the functionality to upload images for further processing.

After submission, the image processing is triggered and smell-related objects and gestures are detected on the image. For further details about the image processing systems, please refer to Deliverable 2.3 [Zinnen and Christlein, 2022b]. Depending on the image size, this step can take between one and three minutes.

Figure 8 showcases the output image, featuring visualizations of detected objects through bounding boxes, along with a display of categories and confidence scores.

The Image Processing Demo serves as a valuable tool for gaining an intuitive understanding of how the data was generated which can be analysed in the Image Nosebook.

4 Conclusion

The Images Nosebook stands as a user-friendly gateway, inviting scholars from diverse backgrounds and expertise levels to delve into the analysis of the visual smell references generated within the Odeuropa project. It offers an intuitive user interface that is aimed at users without programming knowledge and at the same time offers the flexibility to make extensions and changes by more technically experienced users. The Image Processing Live Demo, on the other hand, provides a swift and concise intuition into the workings of the image processing systems. It thus helps estimate the provenance and certainty of the data which can be analysed with the Nosebook. Most significantly it helps to assess to what extent specific categories among the automatically extracted smell references can be trusted.

By combining those two demos, we aim to make our results accessible to a broad audience, fostering new approaches in data-driven research in smell heritage. Furthermore, we hope the demos serve as an entry point for those not yet interested in the fascinating history of smells as conveyed through visual culture.

Figure 8: Screenshot from the Image Processing Demonstrator application showing the visualiation of detected objects and confidence scores.

References

[Couprie, 1983] Couprie, L. D. (1983). Iconclass: an iconographic classification system. *Art libraries journal*, 8(2):32–49.

[Zinnen, 2022] Zinnen, M. (2022). Annotated image data version 2. Technical report, Odeuropa. Deliverable D2.4.

[Zinnen and Christlein, 2022a] Zinnen, M. and Christlein, V. (2022a). Annotated image data version 1. Technical report, Odeuropa. Deliverable D2.2.

[Zinnen and Christlein, 2022b] Zinnen, M. and Christlein, V. (2022b). Object detection/image analysis. Technical report, Odeuropa. Deliverable D2.3.