Hussein Mohammed, Mahdi Jampour, Jost Gippert

# Inpainting with Generative AI: A Significant Step towards Automatically Deciphering Palimpsests

**Abstract:** Palimpsests are manuscripts that have been scraped or washed for reuse, typically as another document. Recovering the undertext of these manuscripts is of significant interest to scholars in the humanities. Therefore, scholars often employ multispectral imaging (MSI) techniques to render the undertext of palimpsests visible. Nevertheless, this approach may not be sufficient in many cases, given that the undertext in the resulting images remains obscured by the overtext. Recent advances in the field of generative artificial intelligence present unprecedented opportunities to discern patterns in highly complex visual data and reconstruct them accordingly. Hence, we propose framing this challenge as an inpainting task in computer vision, aiming to enhance the readability of the undertext through generative image inpainting. To achieve this objective, we have devised a novel approach for generating a synthetic multispectral image dataset of palimpsests, thereby providing a substantial number of training examples without requiring manual annotation. Furthermore, we employed this dataset in fine-tuning a generative inpainting model to improve the legibility of palimpsested undertext. The efficacy of this methodology is demonstrated using coloured and MSI images of Georgian palimpsests with Caucasian Albanian undertexts from Mount Sinai.

## 1 The need for enhancing the readability of palimpsests

There is a general need to reconstruct missing or damaged portions of text or other visual elements within the field of manuscript studies to enhance our comprehension of these historical artefacts. These components are often compromised due to the degradation of the artefacts themselves or the inferior quality of their digitisation. In the instance of a palimpsest, part or the entirety of the textual content has been deliberately removed, thereby impeding efforts to fully understand these artefacts.


https://doi.org/10.1515/9783111552170-019

Scholars often employ the technique of multispectral imaging (MSI), which renders the undertext of a palimpsest visible by capturing light within a specific range of wavelengths, dependent upon the optical properties of the ink used in inscribing this undertext. Nevertheless, the visibility of the undertext does not inherently equate to its readability, as both layers can overlap. Significant parts of the undertext might remain unreadable as a result. Frequently, even when using MSI, the undertext remains a formidable challenge to decipher. In certain instances, it may remain entirely indecipherable. This challenge arises from various factors, encompassing portions of the undertext being obscured by the overtext and other visual elements. Nonetheless, tackling these challenges paves the way for innovative solutions, one of which is the image inpainting technique.

## 2 Inpainting approaches and reconstructing the undertext

Image inpainting, a technique used in computer vision, holds the potential to reconstruct missing regions in an image.[1] This versatile tool finds application in various tasks, including object removal and image restoration. A mask is defined as an image with the same size as the target image, where the pixels within the parts of the image that we intend to 'inpaint' (remove and regenerate) are labelled as 'zeros', while the pixels we want to preserve in the original image are labelled as 'ones'.

Indeed, by viewing the overtext as an undesired visual element, we can effectively frame the task as an object removal problem. In this context, the focus shifts towards eliminating the overtext – as an unwanted entity – while simultaneously reconstructing the hidden undertext and the very surface on which it is written. This amalgamation of techniques and goals underscores the intricate nature of reviving palimpsested manuscripts and uncovering their obscured narratives.

In the experimentations of our work, we defined the pixel labels in the mask so that each pixel belonging to the upper text was to be inpainted, and new pixel values were to be generated so that we could reconstruct the undertext. Figs 1a–d shows an illustrative example of this approach, where (a) is a palimpsest page, (b) shows a processed MSI image of this page, (c) is the mask automatically created for this image, and (d) is the generated results of our image inpainting approach.

**1** See Xiang et al. 2023.

a b c d

**Figs 1a–d:** Illustration of the proposed approach applied to a palimpsest image: (a) original palimpsest (Sinai, St Catherine's Monastery, georg. NF 13, fol. 4v); (b) the outcome of MSI techniques; (c) mask image; and (d) the enhanced undertext using generative image inpainting.

In certain palimpsest specimens, no ink traces remain from the undertext that can be recovered by imaging techniques, such as X-ray fluorescence (XRF) and transmissive light. This means that some parts of the undertext will not be readable, despite not being obscured by any overtext. To reconstruct regions other than those covered by the overtext, it would be necessary to employ other techniques including natural language processing and image registration.

# 3 The Caucasian Albanian palimpsests

The palimpsests chosen for developing our approach are peculiar, given that they represent, in their undertexts, the only manuscripts preserved in the language and script of the so-called Caucasian 'Albanians', an ethnic group of the southern Caucasus whose literacy emerged with the Christianisation of the region in about the fifth century CE.[2] After being erased, the manuscripts were overwritten in Georgian, probably in the tenth century; they are part of the so-called New Finds of St Catherine's Monastery on Mount Sinai, stored in the monastery's library under the shelf numbers georg. NF 13 and NF 55. A first edition of their contents, based upon ultraviolet and multispectral images, was published in 2008,[3] including a rendering of the (preserved or reconstructed) undertexts in the original

**2** See Gippert and Dum-Tragut 2023 for general information on the Caucasian 'Albanians'.

**3** Gippert et al. 2008.

script, for which a special typeface was developed using images of the characters as they appear in the palimpsests themselves.[4]

The Albanian materials revealed in the palimpsests stem from two original manuscripts, one containing the Gospel of John[5] and one representing a collection of other passages from biblical texts, mostly of the New Testament, which were read as lections in divine services.[6] The readability of the two originals was extremely divergent: while the lection passages were discernible up to nearly 100%, the text of the Gospel of John remained uncertain for about two-thirds. This discrepancy meanwhile was overcome, at least to a certain extent, by new imaging methods developed in the course of the Sinai Palimpsests Project,[7] especially with the method of transmissive light imaging, which increased the readability rate to approximately 75%.[8] As a result of this project, we now have access to a large amount of MSI data for the palimpsests, consisting of false-colour and transmissive light images;[9] these were used as the input for our present AI-based approach.

# 4 The proposed approach

We conceptualised the task of enhancing the readability of undertexts as an image inpainting challenge and implemented this approach using a generative artificial intelligence (AI) technique. To this end, we introduced a novel method for generating synthetic multispectral images of palimpsests and used this dataset to fine-tune a generative inpainting model, aiming to improve the readability of palimpsest undertexts. The refined model is subsequently employed in an automated

4 A draft of the font was designed in 2005 by Jost Gippert on the basis of ultraviolet and multispectral images taken on Mount Sinai in 1999–2004 for the first edition of the palimpsests. The final version of the typeface, also used for the Unicode Code Chart of Caucasian Albanian, was developed by Andreas Stötzner in 2007; see <https://unicode.org/charts/PDF/U10530.pdf>.

5 See Gippert 2023a, 105–119 for this part of the palimpsests.

6 See Gippert 2023a, 119–141 as to the 'lectionary' part of the palimpsests.

7 The project (see <http://sinaipalimpsests.org/>) was directed by Michael Phelps and Claudia Rapp and supported by Arcadia Foundation from 2012–2017.

8 See Gippert 2023b for an account of the progress made. A new edition of the Albanian palimpsests is presently underway.

9 The images, available via the Sinai Manuscripts Digital Library at the University of California, Los Angeles (see <https://sinaimanuscripts.library.ucla.edu/>) were produced by Keith T. Knox and kindly provided by the Sinai Palimpsests Project (<https://sinai.library.ucla.edu>, a publication of St Catherine's Monastery of the Sinai in cooperation with the UCLA Library and the Early Manuscripts Electronic Library (EMEL).

pipeline and applied to actual palimpsest samples, thereby generating images with enhanced undertext visibility.

Several state-of-the-art models have been proposed in recent years for the task of image inpainting. Among these, the mask-aware transformer (MAT) has outperformed other models in several critical benchmarks. This model has proven to be exceptionally well-suited for our problem for multiple reasons, including its high performance, the availability of pretrained models, and its capability to learn from the global context of the image. We encourage readers to explore more technical details about this model in a recent article.[10] In the following subsections, we provide further details on the key components of this approach.

## 4.1 The selection of a pretrained model

The MAT model is a novel approach designed for the complex task of image inpainting, combining the advantages of transformers and convolutional techniques to efficiently handle high-resolution images. It excels in modelling long-range interactions, utilising a dynamic mask to focus on valid image portions for high-quality reconstruction. The MAT model's initial training was on a widely accessible dataset, specifically the Places365 dataset. This dataset is designed for scene recognition and encompasses an impressive collection of 10 million images spanning 434 distinct scene categories.[11] Training a model entirely from scratch on such an extensive dataset can be a laborious and computationally demanding endeavour. Therefore, the prevailing approach is to use pretrained models as a foundation and then fine-tune them for particular tasks. Our comprehensive evaluations, encompassing both qualitative and quantitative assessments, have convincingly established that a MAT model, pretrained on the Places365 dataset, consistently delivers the most promising results.

## 4.2 Creating a synthetic dataset

Generative inpainting methods have demonstrated cutting-edge performance,[12] but they require a substantial corpus of training data, comprising authentic images and meticulously annotated pixel-level masks. The generation of such training

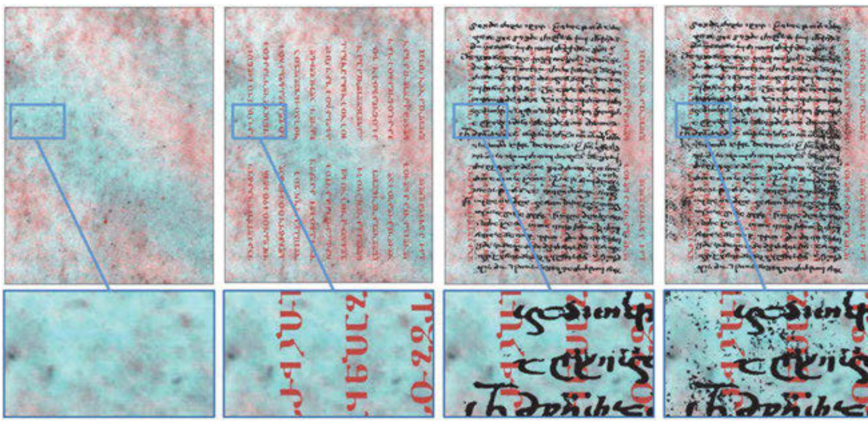
**10** Jampour, Mohammed and Gippert 2024.
**11** See Zhou et al. 2018.
**12** Li et al. 2022; Guo et al. 2021

data is a labour-intensive task and, in many instances, infeasible owing to the unavailability of such information. To address these challenges, we propose the creation of synthetic MSI images of palimpsests. This innovative approach involves a four-step process. Firstly, we initiate the synthesis with a background texture reminiscent of the processed MSI images. Then, we incorporate undertext, using a font design that resonates with the original text and merging it with the background texture. Following this, we introduce random overtext and amalgamate it with the image produced in the preceding stage. Finally, to replicate the customary texture found in MSI images, we add random noise patterns.

To develop a background texture similar to MSI in the training images, we selected portions of processed MSI images of our palimpsests where no text was present. We then applied well-established post-processing techniques, including stitching, scaling, and refining. The undertext is rendered using the font derived from the palimpsest images themselves, as explained in Section 3. To this end, we automatically generated text using the Caucasian Albanian typeface to be overlaid on the automatically generated background, enabling the model to learn the letters' shapes. As a result, the generative part of the process is guided by the characteristics of this typeface.

Following this step, we introduce automatically generated overtext, which is then incorporated into the synthetic images. Within these images, the textual content of the overtext is selected at random. However, efforts are made to align the orientation, line spacing, and proportions with those observed in the genuine palimpsest images. In the final step, we introduce additional visual elements to mimic the unpredictable anomalies commonly found in MSI images of palimpsests. It is important to emphasise that the quantity, distribution, and placement of these anomalies are entirely arbitrary. The aforementioned process of generating synthetic samples is demonstrated in Figs 2a–d.

Through this process, we compiled a dataset containing 1000 synthetic MSI images of the Georgian Albanian palimpsests. Each sample in the dataset is paired with both a mask and a ground-truth image. The ground-truth image serves as the benchmark for evaluating the quality of the generated image using the inpainting model. The mask image outlines the areas requiring inpainting, including the overtext and other noise-related anomalies. For visual reference, Figs 3a–c provides an example of a synthetic sample along with its corresponding mask and ground-truth images. It is worth noting that all images within the dataset adhere to uniform dimensions, measuring 2800 × 2100 pixels, and are stored in full-colour PNG format. In support of academic research and collaboration, this dataset is publicly available for research purposes and can be accessed at our Research Data Repository (RDR).[13]



a b c d

**Figs 2a–d:** Synthesising MSI images of palimpsests: (a) generated background; (b) rendered under-text using the Caucasian Albanian typeface on generated background; (c) randomly generated over-text added on top; (d) randomly generated noise added to mimic the typical texture in our MSI images. The second row shows zoomed-in regions in the images.

a b c

**Figs 3a–c:** A synthesised sample from the proposed dataset. Left to right: (a) ground-truth; (b) mask; and (c) synthesised MSI image.

13 See the dataset in Jampour, Mohammed and Gippert 2023.

## 4.3 Fine-tuning the pretrained model

The MAT model, pretrained on the Places365 dataset, underwent additional fine-tuning using the above-described synthetic dataset. This fine-tuning process aimed at enhancing the model's performance and achieving superior quality when reconstructing the underlying text. The dataset, comprising 1000 samples, was partitioned into subgroups – 800 for training, 100 for validation, and 100 for testing – to facilitate this refinement. During fine-tuning, the model learned to inpaint regions specified by the mask image, and the resulting inpainted images were compared with the corresponding ground-truth. Over time, we expect the model to acquire knowledge of letter shapes through this process. Consequently, when applied to the test dataset, the model, having already internalised the letter shapes, is anticipated to provide accurate predictions for the pixel values specified in a given mask image. Further technical details and evaluation results can be found in our 2024 article.[14]

## 4.4 Reconstructing the undertext

Following the fine-tuning process, the inpainting model is deployed to effectively eliminate the superimposed text and analogous visual elements present in MSI images of palimpsests. Subsequently, it generates the concealed portions of the image lying beneath, which could pertain to either the undertext or the writing support. The identification of overtext and noise areas within the image is performed automatically, relying on the numerical characteristics of their pixel values. A band-pass filter is implemented by defining a minimum and maximum threshold in order to automatically select pixels belonging to the masked regions. Fig. 4 shows an example of the generated results from the proposed generative image inpainting approach.

14 Jampour, Mohammed and Gippert 2024.

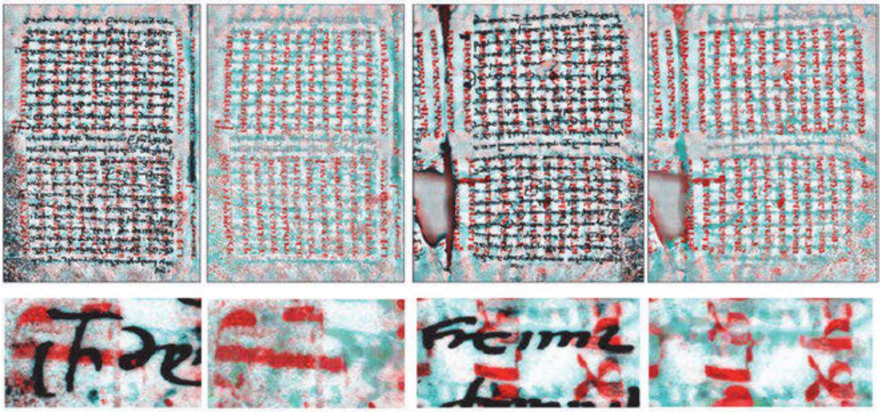


**Fig. 4:** Results of generative inpainting on two MSI images of palimpsests (Sinai, St Catherine's Monastery, georg. NF 13, fols 4[v] and 78[r]).

## 5 Discussion and conclusions

Generative AI holds significant potential for various facets of manuscript studies, notably in reconstructing textual and visual information. In this article, we have demonstrated the capacity of the presented techniques to enhance the readability of undertext in palimpsest manuscripts. The broader application of similar methods to other restoration challenges seems to be promising and offers intriguing scientific challenges for interdisciplinary research.

The effectiveness of the proposed approach is evident in reconstructing the undertext and enhancing its readability, thereby helping to reveal more about its history. However, the quality of this performance relies partially on the quality of the MSI images of the palimpsest. This image quality may not necessarily align with what human vision considers 'better' or more visible, as computer vision systems process images based on various visual features encapsulated within the pixel values. Therefore, working with raw MSI data might improve inpainting performance and further enhance the readability of the undertext – a prospect we are eagerly pursuing.

While the model used has been trained on a typeface of the Caucasian Albanian language, the data preparation, image pre-processing, and results post-processing are generic approaches applicable to any palimpsest. Moreover, the image inpainting model can be fine-tuned for any typeface in any other language. Additionally, this approach can serve as a pre-processing step for further processing in other computer vision systems, such as handwriting style analysis or handwritten text recognition (HTR) systems.

In many cases, the undertext is damaged and fragmented, regardless of whether it is occluded by overtext. Such cases require extensive restoration beyond inpainting only the regions covered by overtext. Such restoration can be achieved through combining approaches from computer vision (CV) and natural language processing (NLP). One possibility is to use image registration techniques for every letter, guided by an NLP model, which must be trained on the same script used in a given palimpsest.

Furthermore, the typefaces used to train the model understandably only generally approximate the visual appearance of the actual letters in the undertext. Both the handwriting style and the scale of letters can vary in each palimpsest sample. This variation can negatively affect the performance of this approach, to a degree influenced by the extent of deviation from the used typeface. More advanced training strategies could mitigate this influence, which can be explored in future work.

## References


Gippert, Jost (2023a), 'The Textual Heritage of Caucasian Albanian', in Gippert and Dum-Tragut 2023, 95–166 <https://doi.org/10.1515/9783110794687-003>.

Gippert, Jost (2023b), 'New Light on the Caucasian Albanian Palimpsests of St. Catherine's Monastery', in Claudia Rapp, Giulia Rosetto, Jana Grusková and Grigory Kessel (eds), *New Light on Old Manuscripts: The Sinai Palimpsests and Other Advances in Palimpsest Studies* (Veröffentlichungen zur Byzanzforschung, 45; Österreichische Akademie der Wissenschaften, Philosophisch-historische Klasse, Denkschriften, 547), Vienna: Austrian Academy of Sciences Press, 187–215 <https://doi.org/10.1553/978OEAW91575s187>.

Gippert, Jost and Jasmine Dum-Tragut (eds) (2023), *Caucasian Albania: An International Handbook*, Berlin: De Gruyter, <https://doi.org/10.1515/9783110794687>.

Gippert, Jost, Wolfgang Schulze, Zaza Aleksidze and Jean-Pierre Mahé (2008), *The Caucasian Albanian Palimpsests of Mount Sinai*, 2 vols (Monumenta Palaeographica Medii Aevi, Series Ibero-Caucasica, 2), Turnhout: Brepols.

Guo, Qing, Xiaoguang Li, Felix Juefei-Xu, Hongkai Yu, Yang Liu and Song Wang (2021), 'JPGNet: Joint Predictive Filtering and Generative Network for Image Inpainting', in Heng Tao Shen, Yueting Zhuang and John R. Smith (eds), *Proceedings of the 29th ACM International Conference on Multimedia (MM '21)*, New York: Association for Computing Machinery, 386–394 <https://doi.org/10.1145/3474085.3475170>.

Jampour, Mahdi, Hussein Mohammed and Jost Gippert (2023), *Synthetic MSI Images of Georgian Palimpsests* (SGP Dataset) (Version 1.0) [Dataset], <https://doi.org/10.25592/uhhfdm.13378>.

Jampour, Mahdi, Hussein Mohammed and Jost Gippert (2024), 'Enhancing the Readability of Palimpsests Using Generative Image Inpainting', in Modesto Castrillón-Santana, Maria De Marsico and Ana Fred (eds), *Proceedings of the 13th International Conference on Pattern Recognition Applications* *and Methods ICPRAM (February 24–26, 2024, in Rome, Italy)*, Setúbal: Science and Technology Publications (Scitepress), vol. 1, 687–694 <https://doi.org/10.5220/0012347100003654>.

Li, Wenbo, Zhe Lin, Kun Zhou, Lu Qi, Yi Wang and Jiaya Jia (2022), 'MAT: Mask-Aware Transformer for Large Hole Image Inpainting', in *2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, New Orleans, LA: Computer Vision Foundation, 10748–10758 <https://doi.org/10.1109/CVPR52688.2022.01049>.

Xiang, Hanyu, Qin Zou, Muhammad Ali Nawaz, Xianfeng Huang, Fan Zhang and Hongkai Yu (2023), 'Deep Learning for Image Inpainting: A Survey', *Pattern Recognition*, 134 <https://doi.org/10.1016/j.patcog.2022.109046>.

Zhou, Bolei, Agata Lapedriza, Aditya Khosla, Aude Oliva and Antonio Torralba (2018), 'Places: A 10 Million Image Database for Scene Recognition', *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 40/6: 1452–1464 <https://doi.org/10.1109/TPAMI.2017.2723009>.