Image Inpainting Applied to Art
Completing Escher's Print Gallery

Lucia Cipolina-Kun [1] **Simone Caenazzo** [2] **Gaston Mazzei** [3][1] **Aditya Srinivas Menon** [4]

Abstract

This extended abstract presents the first stages of a research in inpainting suited for art reconstruction. We introduce M.C Escher's *Print Gallery* lithography as a use-case example. This artwork presents a void on its center and additionally, it follows a challenging mathematical structure that needs to be preserved by the inpainting method. We present our work so far and our future line of research.

1. Introduction

Completing missing sections of a painting poses several difficulties for the current state-of-the-art knowledge in AI. The model is required to learn how to paint like a certain artist, in a certain period, on a given painting and potentially continue it seamlessly as if the painter had a bigger canvas. We aim to develop a method that goes beyond learning the style of a given artist - see for example the *Style Transfer* models in Gatys et al. (Gatys et al., 2015) - and learn the idiosyncrasies of a single given artwork. This allow us to reconstruct the painting's missing pieces following the original style, content and context of the piece.

As an application to our model, we chose Escher's *Print Gallery* lithography as it poses an interesting blending of art and mathematics, as well as the above mentioned challenges of an inpainting model. Firstly, any attempt to complete the painting will have to preserve its unique mathematical properties. Secondly, this painting features a very unique style, effectively found only in a handful of works by Escher; there are therefore not enough samples to perform a straightforward fine-tuning of typical pre-trained models, such as *Image-GPT* (Chen et al., 2020) or *Taming Transformers* (Esser et al., 2020) which require thousands of training images.

[1]ML Collective [2]Riskcare Ltd., London [3]Departamento de Fisica, Universidad de Buenos Aires, Buenos Aires, Argentina [4]Indian Institute of Information Technology, Kottayam. Correspondence to: Lucia Cipolina-Kun <lcipolina@gmail.com>.

2. Related Work

The application of image completion and inpainting to fine art and painting reconstruction is still an incipient field, with very few examples to date like (Zeng et al., 2019) reconstructing Van Gogh's degraded painting or (Gupta et al., 2019) for restoration in digital art. In these works, the aim is not to create new content to fill-in missing sections of an image, but to restore existing content, degraded sections or lines.

3. Escher's Print Gallery

The work *Print Gallery* (Dutch: *Prententoonstelling*), shown in Figure 1, is a lithography made in 1956 by the Dutch artist M. C. Escher. It depicts a man looking at painting on a gallery placed in the Grand Harbour of Valletta (Malta).

At the center of the lithography, Escher left a blank space with his signature. The reason for this blank area is unknown: some authors, e.g. (Hofstadter), claim that he was not actually able to fill that space, while (Ernst, 1976) claims that it is not even possible to continue the spiral-like effect in a sensible way.

3.1. The Mathematical Structure Behind Print Gallery

Print Gallery poses an additional challenge for inpainting models as it has an underlying mathematical structure that needs to be preserved when continuing the lithography.

As explained in (Ernst, 1976) and (de Smit & Lenstra Jr, 2003), the lithography features a particular sequence of roto-homotheties known in art jargon as a *Droste* effect: i.e., the image contains a repetition of itself with a spiral-like motion inducing the idea of an infinite loop.

Additionally to this self-repeating structure, the lithography features a wave-like spatial distortion. To achieve this effect, Escher used a grid-like pattern based on curved lines, to enhance the idea of circular motion already embedded in the Droste effect. Figure 2 shows the grid structure used by Escher superimposed to the painting - additionally, in the center of the image, we show the corresponding Droste repetition as explained in (de Smit & Lenstra Jr, 2003).

Figure 1. M.C. Escher's lithography *Print Gallery* (*Prententen-toonstelling*), 1956.
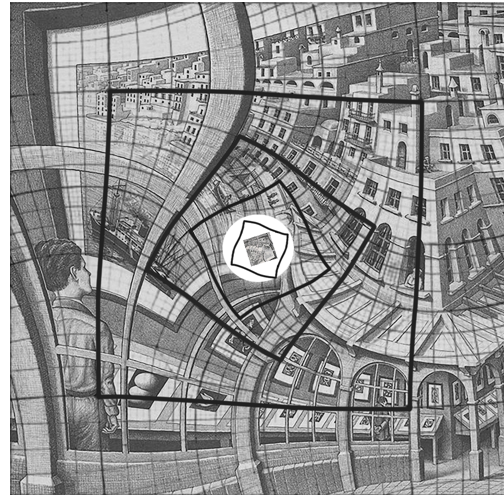

Figure 2. Escher's lithography with superimposed grid and Droste repetition in the middle

As shown above in Figure 2, the self-repetition placed in the center is not big enough to complete the entire center blank. The remaining void white area is due to the difference between the void circle and the lithography's own squared shape, as well as an expansion in the field of view [1]. Filling the remaining void white area is the aim of our inpainting model.

4. Methodology

Our general aim is to generate an inpainting model that can accurately complete a given painting - this is achieved not only by following the painter's generic style, but also by following the specific artwork's style and content.

In our *Print Gallery* application case, we need to train a model that will not only paint like Escher, but will specifically *paint like Escher when he worked on Print Gallery*. This poses a non-trivial challenge: in Machine Learning terms, we have a single sample from which to extract and/or create additional content.

To complete the remaining blank parts, we propose to use image synthesis/inpainting techniques to create content that follows the semantic structure of the lithography and its surrounding context in a consistent way.

In particular, we base our present work on the state-of-the-art *Large Scale Image Completion via Co-Modulated Generative Adversarial Networks-CoModGAN* model in (Zhao et al., 2021), as it was specifically developed to complete

---

[1] A smaller object gives the illusion of being placed farther apart from the viewer. This effect is equivalent to the viewer stepping back from a static object and thus, extending her field of view.

large regions of missing data on an image. This model implements a modulation of the conditional and unconditional image vectors into the traiditional Generative Adversarial Networks to generate content consistent with the image's semantics.

To enhance the capabilities of the base *CoModGAN* model for the present task, we additionally fine-tuned it using a database of 1,000 publicly-available images of Valletta's Grand Harbour and other Mediterranean urban architectures. To better suit the original lithography, all images were additionally converted to black/white and rendered as pencil sketches using the style-transfer model in (Li et al., 2019).

5. Preliminary Results

The below images show preliminary results of our tests with the *CoModGAN* model.

The overall approach is to first use *CoModGAN* to expand the image outwards (also knowns as *outpainting*, see Figure 3) and then fit the expanded image in the center (see Figure 4) following the Droste structure.

Figure 3 shows the performance of the model on a relatively small vertical region (highlighted in red) on the right of the original lithography. The model's results fit adequately within the context: the additional buildings on the top-right section follow naturally from the original group of buildings, and the gallery at the bottom is continued in an acceptable manner. Overall, for this small vertical section, the model is also able to reasonably follow the curvature of the original painting.

Figure 5 shows instead the performance of the model on a

larger outside region. In this case, the results are not satisfactory: the outpainted sections generally lack coherence with respect to the original painting in terms of both content and curvature. For this reason, the completion of the gap cannot be achieved in a straight-forward manner by applying the current model on a larger section. Other attempts, not shown here for brevity, at applying the model recursively (i.e. by filling small sections and reapplying the model on the resulting picture) and in an hybrid inpainting/outpainting capacity also showed incoherent results.
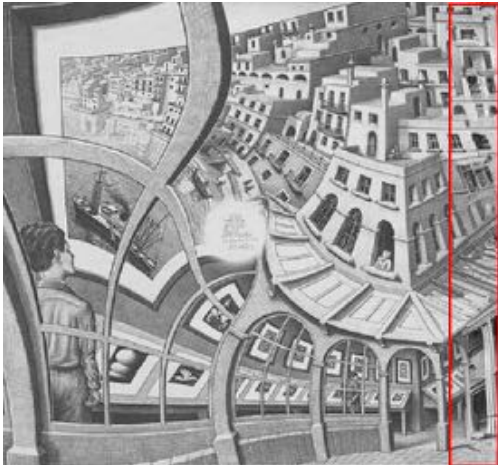

Figure 3. CoModGAN outpainting results on a large vertical strip section (highlighted in red).
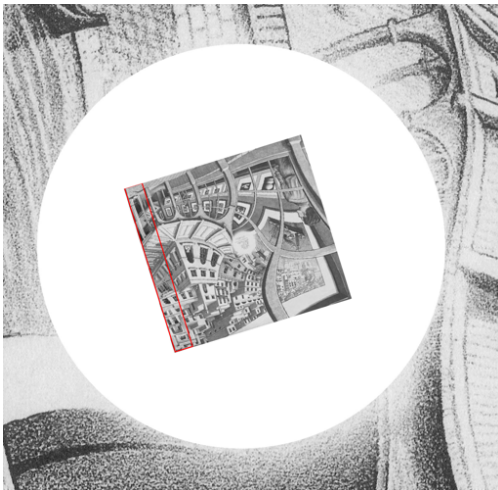

Figure 4. Enlarged center of the inpainting results after rotation and scaling. The red strip represents the area extended by the *CoModGAN* model.

Figure 5. CoModGAN outpainting results on a larger outer mask (gray) of *Print Gallery* (yellow).

6. Future Work

6.1. Enhanced fine-tuning

Fine-tuning of pre-trained models can greatly enhance their performance for specific tasks. In order to improve the results obtained on larger sections, we aim to produce a larger database of relevant pictures. For example, an improved fine-tuning database could feature archways, galleries, waterfronts and non-contemporary buildings to match more closely the content in Escher's original painting.

Potential future work also includes customizing the model further to the specifics of the painting, e.g. by changing the way the model selects images from the database for inpainting purposes. One possible route is to use smaller sections of the painting itself as an image database. In our example, the extra buildings needed could be generated by looking at the existing buildings on the lithography first, before searching for additional building images on the training database.

6.2. Curvature enforcing

In order to enforce a stricter adherence of the inpainted sections to the spiral-like grid that defines *Print Gallery*, we envisage a range of potential approaches:

1. **Direct modification of standard inpainting models, i.e. models that normally operate in "straight" 2D geometries**: for example, deskewing routines (such as (SKI)) could be inserted at strategic places in the model pipeline, to straighten small sections of the painting, augment them via inpainting, and then apply an in-

verse skewing to the augmented section to restore the curvature;

2. **Use *twisted* training pictures**: in addition to the style-transfer steps described in Section 4, the training database could be enriched with twisted and spiralized pictures, i.e. regular pictures that have undergone a wave-like spatial deformation that follows Escher's original grid. In this way, the model may learn the notion of rotation and deformation mimicking it in the generated content.

3. **Infuse the inpainting models with rotation/scaling parameters derived from Escher's original grid**: the grid can be parametrized as a discrete function of rotation vs. scaling by observing the reference square-like structures (highlighted with black lines in Figure 2). This function can then be expressed as a continuous curve by applying a smooth polynomial interpolation (e.g. a monotonic Hermite spline). This information could then be supplied to an additional layer/module of the inpainting model (to be developed), in order to deform the inpainted regions and generated content in accordance to the grid specifications.

6.3. Model evaluation

Evaluating the quality of an inpainting model is a non-trivial task, and is still an open field of research. Thus, once the chosen inpainting model(s) will achieve satisfactory results from an arbitrary/qualitative point of view, we envisage testing a selection of *quantitative* metrics. (Qureshi et al., 2017) provides a review of inpainting evaluation techniques, whilst (Amirkhani & Bastanfard, 2019) proposes a novel approach based on saliency maps. Turing tests would also be an interesting route, especially given the artistic dimension of the present inpainting exercise. For the curvature assessment, the model will produce a degree of deformation in the generated pictures that can be compared against the deformation induced by the grid. Lastly, the produced output can be compared against the previous work in (de Smit & Lenstra Jr, 2003).

7. Acknowledgements

References

Python library *deskew*. URL https://github.com/sbrunner/deskew.

Amirkhani, D. and Bastanfard, A. Inpainted Image Quality Evaluation Based on Saliency Map Features. In *2019 5th Iranian Conference on Signal Processing and Intelligent Systems (ICSPIS)*, pp. 1–6, 2019.

Chen, M., Radford, A., Child, R., Wu, J., Jun, H., Dhariwal, P., Luan, D., and Sutskever, I. Generative Pretraining from Pixels. 2020.

de Smit, B. and Lenstra Jr, H. W. The Mathematical Structure of Escher's Print Gallery. *Notices of the AMS*, 50(5): 446–451, 2003.

Ernst, B. De toverspiegel van M. C. Escher. 1976. Meulenhoff, Amsterdam, 1976; English translation by John E. Brigham: The Magic Mirror of M. C. Escher, Ballantine-Books, New York.

Esser, P., Rombach, R., and Ommer, B. Taming Transformers for High-Resolution Image synthesis, 2020.

Gatys, L. A., Ecker, A. S., and Bethge, M. A Neural Algorithm of Artistic style. *CoRR*, 2015. URL http://arxiv.org/abs/1508.06576.

Gupta, V., Sambyal, N., and Sharma, A. Restoration of artwork using deep neural networks. *Evolving Systems*, 2019.

Hofstadter, D. R. Gödel, Escher, Bach: an Eternal Golden Braid. *Basic Books*.

Li, Y., Fang, C., Hertzmann, A., Shechtman, E., and Yang, M.-H. Im2pencil: Controllable Pencil Illustration from Photographs. 03 2019.

Qureshi, M. A., Deriche, M., Beghdadi, A., and Amin, A. A critical survey of state-of-the-art image inpainting quality assessment metrics. *Journal of Visual Communication and Image Representation*, 49:177–191, 2017.

Zeng, Y., van der Lubbe, J., and Loog, M. Multi-scale convolutional neural network for pixel-wise reconstruction of Van Gogh's drawings. *Machine Vision and Applications: an international journal*, 30(7-8):1229–1241, October 2019.

Zhao, S., Cui, J., Sheng, Y., Dong, Y., Liang, X., Chang, E. I., and Xu, Y. Large Scale Image Completion via Co-Modulated Generative Adversarial Networks. In *International Conference on Learning Representations (ICLR)*, 2021.