

Time-Travel Rephotography

XUAN LUO, University of Washington

XUANER (CECILIA) ZHANG, Adobe Inc.

PAUL YOO, University of Washington

RICARDO MARTIN-BRUALLA, Google Research

JASON LAWRENCE, Google Research

STEVEN M. SEITZ, University of Washington and Google Research



Fig. 1. Left: Antique film lacks red sensitivity, exaggerating wrinkles and darkening lips. Right: Our rendering of how Abraham Lincoln (c. 1863) would appear *rephotographed with a modern camera*. The input photo is from Mead Art Museum (public domain). Images in all figures are best viewed digitally, and zoomed in to 1024 × 1024 to see details.

Many historical people were only ever captured by old, faded, black and white photos, that are distorted due to the limitations of early cameras and the passage of time. This paper simulates traveling back in time with a modern camera to rephotograph famous subjects. Unlike conventional image restoration filters which apply independent operations like denoising, colorization, and superresolution, we leverage the StyleGAN2 framework to

Authors' addresses: Xuan Luo, xuanluo@cs.washington.edu, University of Washington; Xuaner (Cecilia) Zhang, Adobe Inc., cecilia77@berkeley.edu; Paul Yoo, University of Washington, yoosehy@cs.washington.edu; Ricardo Martin-Brualla, Google Research, rmbrualla@google.com; Jason Lawrence, Google Research, jdlaw@google.com; Steven M. Seitz, University of Washington and Google Research, seitz@cs.washington.edu.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than ACM must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from permissions@acm.org.

© 2021 Association for Computing Machinery.

0730-0301/2021/12-ART213 \$15.00

<https://doi.org/10.1145/3478513.3480485>

project old photos into the space of modern high-resolution photos, achieving all of these effects in a unified framework. A unique challenge with this approach is retaining the identity and pose of the subject in the original photo, while discarding the many artifacts frequently seen in low-quality antique photos. Our comparisons to current state-of-the-art restoration filters show significant improvements and compelling results for a variety of important historical people. Please go to time-travel-rephotography.github.io for many more results.

CCS Concepts: • Computing methodologies → Computational photography.

Additional Key Words and Phrases: Image restoration, colorization, super resolution, vision for graphics

ACM Reference Format:

Xuan Luo, Xuaner (Cecilia) Zhang, Paul Yoo, Ricardo Martin-Brualla, Jason Lawrence, and Steven M. Seitz. 2021. Time-Travel Rephotography. *ACM Trans. Graph.* 40, 6, Article 213 (December 2021), 12 pages. <https://doi.org/10.1145/3478513.3480485>

1 INTRODUCTION

Abraham Lincoln's face is iconic – we recognize him instantly. But what did he *really* look like? Our understanding of his appearance is based on grainy, black and white photos from well over a century ago. Antique photos provide a fascinating glimpse of the distant past. However, they also depict a faded, monochromatic world very different from what people at the time experienced. Old photos distort appearance in other less obvious ways. For example, the film of Lincoln's era was sensitive only to blue and UV light, causing cheeks to appear dark, and overly emphasizing wrinkles by filtering out skin subsurface scatter which occurs mostly in the red channel. Hence, the deep lines and sharp creases that we associate with Lincoln's face (Fig. 1) are likely exaggerated by the photographic process of the time.

To see what Lincoln really looked like, one could travel back in time to take a photo of him with a modern camera, and share that photo with the modern world. Lacking a time machine, we instead seek to simulate the result, by projecting an old photo into the space of modern images, a process that we call *time-travel rephotography*.

Specifically, we start with an antique photo as reference, and wish to generate the high resolution, high quality image that a modern camera would have produced of the same subject. This problem is challenging, as antique photos have a wide range of defects due both to the aging process (fading and dust) and to limitations of early cameras, film, and development processes (low resolution, noise and grain, limited color sensitivity, development artifacts). A common approach is to try to restore the image, by applying a sequence of digital filters that attempt to undo these defects, e.g., noise removal, image deblurring, contrast adjustment, super-resolution, and colorization. A challenge with this approach is that the properties of old film and the aging process haven't been fully characterized – hence, undoing them is an ill-posed problem.

Instead, we propose to project the antique photo into the space of modern images, using generative tools like StyleGAN2 [Karras et al. 2020a]. Unlike prior StyleGAN inversion methods [Abdal et al. 2019, 2020; Baylies 2019; Richardson et al. 2020; Tov et al. 2021; Zhu et al. 2020], our goal is not to reconstruct the input grayscale image, but to synthesize its missing colors, dynamic range and skin details while preserving the subject's identity. Our approach uses a physically-based *film degradation* operator that simulates properties of antique cameras and the film aging process. This includes modeling the film's chromatic sensitivity, addressing for the first time different antique photographic emulsions (*blue-sensitive*, *orthochromatic* and *panchromatic* [Adams 2018]). We also show that sharp photos with pleasing exposure and contrast can be achieved by explicitly modeling blur and non-linear camera response functions. Finally, to ensure natural color tones and high resolution face details, we synthesize a high-quality modern *sibling* image in the StyleGAN2 space to serve as an exemplar (Fig. 2). We transfer the color and skin details from the sibling to our output using a contextual loss [Mechrez et al. 2018] and a novel *color transfer loss* (Sec. 4).

We present results that show our approach, which addresses artifact removal, colorization, super-resolution, and contrast adjustment in one unified framework, consistently outperforms applying

a sequence state-of-the-art image restoration filters. We also demonstrate compelling time-travel portraits of many well-known historical figures from the last two centuries, including presidents, authors, artists, and scientists (Fig. 1, 10). Since we rely heavily on the face priors from StyleGAN2, our method is prone to the bias [Salminen et al. 2020] of StyleGAN2 and its training set. Such bias and our limitations are discussed in Sec. 6.

2 RELATED WORK

Our method integrates concepts from the mature image restoration literature with modern learning-based approaches for modeling the space of high-resolution images of human faces.

Image Restoration. Many prior methods address a *single* type of image degradation, including denoising [Buades et al. 2005; Dabov et al. 2007; Elad and Aharon 2006; Lefkimiatis 2017; Xie et al. 2012; Zhang et al. 2017b,c, 2018b,a], deblurring [Kupyn et al. 2018; Nah et al. 2017; Sun et al. 2015; Xu et al. 2014a,b], JPEG image deblocking [Dong et al. 2015a; Guo and Chao 2016; Wang et al. 2016], and super-resolution [Babacan et al. 2008; Dong et al. 2015b; Kim et al. 2016; Ledig et al. 2017; Menon et al. 2020b; Tai et al. 2017; Yang et al. 2010]. Some methods target specifically at faces, including deblurring [Hacohen et al. 2013; Pan et al. 2014] and super-resolution [Bulat and Tzimiropoulos 2018; Grm et al. 2019; Menon et al. 2020b; Ren et al. 2019; Shen et al. 2018]. To address restoring *multiple* artifacts, researchers have proposed using reinforcement learning [Yu et al. 2018] or attention-based mechanisms [Suganuma et al. 2019] to select the best combination of restoration operations to apply to each image. The work by Wan et al. [2020] also restores portraits suffering from multiple artifacts. However, it does not do super resolution and its restoration quality degrades at high resolution of 1024×1024 as evaluated in Sec. 5. The concurrent work by Wang et al. [2021] also leverages generative facial prior. They propose to incorporate such prior into joint face restoration and color enhancement using spatial feature transform layers. Nevertheless, none of the aforementioned techniques address colorization.

Colorization research can be categorized into scribble-based, exemplar-based and learning-based methods. Early work [Huang et al. 2005; Levin et al. 2004; Luan et al. 2007; Qu et al. 2006; Sýkora et al. 2009; Yatziv and Sapiro 2006] manually specify target colors for parts of the image via sparse *scribbles* drawn on the image. To reduce user involvement, an alternative is to transfer color statistics from a (manually-specified) reference image [Charpiat et al. 2008; Chia et al. 2011; Gupta et al. 2012; He et al. 2018; Ironi et al. 2005; Liu et al. 2008; Welsh et al. 2002]. Identifying a suitable reference image, however, is a research topic by itself. Automatic methods often involve a complicated image retrieval system [Chia et al. 2011]. We simplify this process by predicting a high-quality StyleGAN2 sibling as the reference. Most related are *fully automated* colorization methods that use machine learning on a large dataset [Cheng et al. 2015; Deshpande et al. 2015; Iizuka et al. 2016; Isola et al. 2017; Larsson et al. 2016; Zhang et al. 2016, 2017a; Zhao et al. 2018]. We compare with many of these methods in Sec. 5.

Despite the rapid progress in these individual areas, no prior works have addressed restoration, colorization and super-resolution

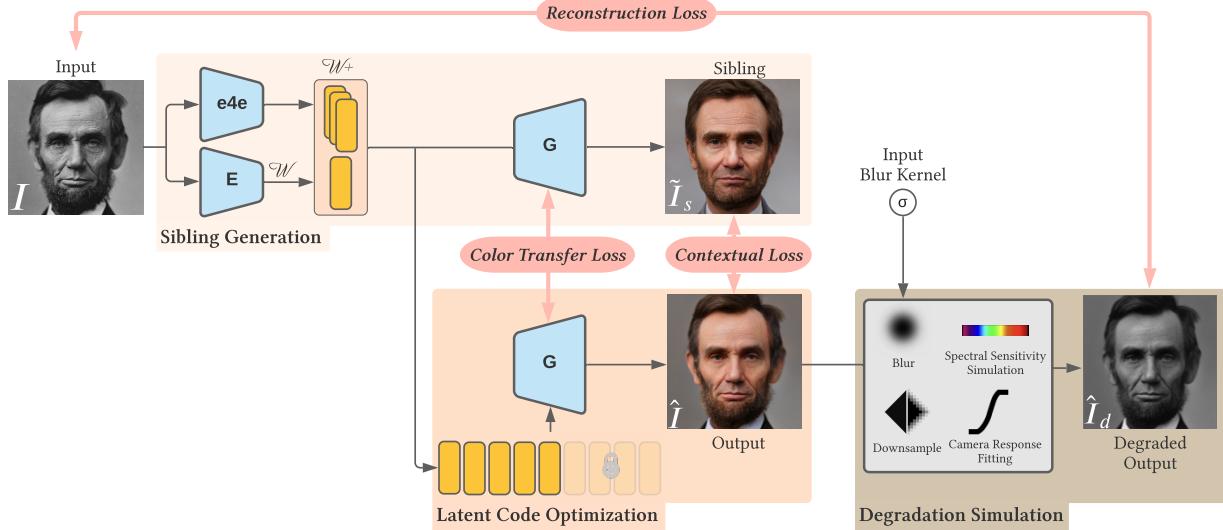


Fig. 2. Given an input antique photo, we generate a *sibling* in the StyleGAN2 latent space by *style mixing* [Karras et al. 2019] the predictions from two feed-forward encoders, one that models face identity (*e4e* [Tov et al. 2021]) and another for face color (*E*). We then optimize the latent code of the sibling to match the input, after passing through a degradation model that simulates antique images, guided by the color, contrast, and skin textures of the sibling.

in a single framework. We demonstrate that doing so produces better results than applying a sequence of state-of-the-art techniques.

Face embedding. Karras et al. [2019; 2020a] introduced the StyleGAN framework for synthesizing high resolution human faces from a latent space. Projecting or *embedding* a face image to the latent space is an active research topic. Current methods fall into three categories: 1) optimizing latent vectors to best fit the input image [Abdal et al. 2019, 2020], 2) training an *encoder network* that estimates a latent coordinate from an input image [Richardson et al. 2020; Tov et al. 2021], and 3) hybrid methods [Baylies 2019; Zhu et al. 2020] that use an encoder network to initialize an optimization, similar to our approach. In contrast to our work, all of these prior methods seek to reconstruct the input directly and do not address the restoration problem, i.e., they would seek embeddings that preserve an antique image's monochrome, blurry, and low-contrast properties. Directly extending these prior methods to overcome this limitation by inserting an image degradation operator into the loss function results in a poorly conditioned optimization problem that can quickly converge to a bad result (Fig. 3e), even with a good starting point [Baylies 2019; Zhu et al. 2020]. Our use of a sibling image is critical for obtaining a well conditioned optimization problem that reliably converges to a high-quality result. We note that Menon et al. [2020a] addressed the specific problem of upsampling. However, they focus on high-ratio upsampling where the identity of input images are hardly recognizable. Similarly, Yang et al. [2021] uses GAN priors to blindly restore images with extreme low resolutions without explicitly constraining the identity. Recently, Pan et al. [2020] exploits deep generative prior for independent image colorization and super resolution tasks, where they further finetune the generator together with the latent code to reduce the gap between the training and testing images. However, none of these works take into account the complicated real-world degradations of antique photos such as the

chromatic sensitivity and contrast adjustment that are critical for high-quality antique photo restoration.

3 PROBLEM STATEMENT

Our goal is to simulate traveling back in time and rephotographing historical figures with a modern camera. We call this *time-travel rephotography*, adapting the term *rephotography* which traditionally means “the act of repeat photography of the same site, with a time lag between the two images” [Wikipedia 2021c]. We mainly focus on portraits recorded a century ago, shortly after cameras were invented in the late 1800s, which are challenging to restore due to loss of quality through the aging process and limitations of early film, but our method operate on more recent photos as well.

Photographic film has evolved significantly since its invention. The first light-sensitive emulsions were sensitive only to blue and ultraviolet light [Newhall 1982]. *Orthochromatic* emulsions [Newhall 1982], introduced in 1873, provided sensitivity to both green and blue light. Photographic portraits from these eras rendered skin poorly, artificially darkening lips and exaggerating wrinkles, creases, and freckles, due to the lack of red sensitivity. In particular, they underestimate the effect of subsurface scattering, which gives skin its characteristic smooth appearance [Jensen et al. 2001]. *Panchromatic* film, sensitive to red, green, and blue first appeared in 1906, yet orthochromatic films remained popular through the first half of the 20th century [Wikipedia 2021b].

To simulate rephotographing historical people with a *modern camera*, we must account for these differences in color sensitivity of antique film, in addition to blur, fading (poor exposure, contrast), noises, low resolution, and other artifacts of antique photos.

4 METHOD

We seek to synthesize a *modern photo* of a historical person, using an antique black-and-white photo as reference. Our approach is

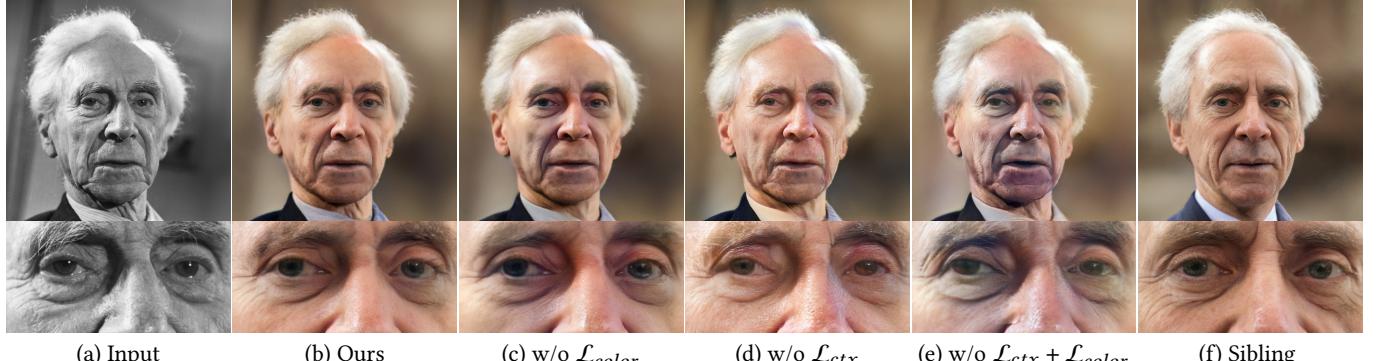


Fig. 3. Impact of each loss for transferring color and details from the sibling. In all these examples, the sibling code is used as an initialization and its fine codes (128 - 1024) are kept unchanged during optimization. Yet, low-frequency color artifacts appear without \mathcal{L}_{color} . Skin and eye details are poorly reconstructed without \mathcal{L}_{ctx} . Input image: Bertrand Russell (1872 - 1970) from BBC Photo Library.

based on the idea of *projecting* the antique photo into the space of modern high-resolution color images represented by the StyleGAN2 generative model [Karras et al. 2020b].

Similar to previous techniques [Baylies 2019], we optimize the latent space of StyleGAN2 to synthesize an image. However, directly fitting the antique image would reproduce a grainy black and white result, we instead reconstruct an image of the person without any artifact caused by antique negatives or the film aging process.

A first step is to convert the StyleGAN2 output to grayscale before comparing with the antique input image. This naive approach is poorly constrained, since multiple colors can correspond to the same grayscale output, and thus leads to unrealistic colorized results (Fig. 3e). Therefore, we employ an additional exemplar image as a reference, that has similar facial features as the input, yet contains high frequency details and natural color and lighting. Sec. 4.1 explains how we compute such an exemplar automatically. We call it the *sibling* image, as it resembles characteristics of the input while having a different identity. Sec. 4.2 introduces losses used in optimization that constrain our rephotographed output to retain the contrast, color and high-frequency details present in the sibling.

To further reduce the perceived identity gap between the input image and the modern portrait, we design reconstruction losses suited for antique images (Sec. 4.3). A key contribution is the proposed degradation module, that simulates the image formation model in antique photos, and is applied to the StyleGAN2 result before comparing it with the input antique photo. The degradation module accounts for different types of film substrate, camera response curves, image blur, and low resolution, which altogether leads to improved rephotography results. We provide details on the latent code optimization in Sec. 4.4 and a system overview in Figure 2.

4.1 Sibling Encoders

Given a low-resolution grayscale reference image as input I , we seek to generate a high-resolution color *sibling* image \hat{I}_s that has photo-realistic colors and preserves facial features in the original input. The state-of-the-art StyleGAN inversion method *e4e* [Tov et al. 2021] can embed an input image into the $\mathcal{W}+$ space [Abdal et al. 2019] (18 different 512-dimensional StyleGAN2 \mathcal{W} codes) with reasonably well-preserved identity. However, when applied to antique photos, such inversion methods will also preserve the artifacts, such as

blur and lack of color. An alternative approach is to train a feed-forward encoder E , aimed specifically at predicting a high-quality (i.e. realistic skin color and details) StyleGAN2 embedding from an antique image. Avoiding the transfer of artifacts is possible by limiting the expressiveness of the embedding, i.e. predicting an embedding in \mathcal{W} instead of $\mathcal{W}+$. As a result, this embedding is also worse at preserving the subject’s identity.

Our method combines the best of both worlds by *style mixing* [Karras et al. 2019] the predictions from *e4e* and from E . Specifically, we use *e4e*’s $\mathcal{W}+$ prediction for the first 10 coarse style codes and duplicate our \mathcal{W} prediction for the other 8 fine codes. The mixed $\mathcal{W}+$ codes are then converted to the sibling image using the pre-trained StyleGAN2 generator [Karras et al. 2020b]. Our encoder E is trained using random samples of StyleGAN2 \mathcal{W} latent codes and their corresponding images, downsampled to 256×256 and converted to grayscale based on the emulsion type (see details in Sec. 4.3).

4.2 Sibling Color And Detail Transfer

To further constrain colors and skin details to match the sibling (Fig. 3), we introduce a **color transfer loss** \mathcal{L}_{color} that enforces the distribution of the output colors from StyleGAN2’s ToRGB layers to be similar to those of the sibling image. We use a formulation inspired by style loss [Gatys et al. 2016], and apply to ToRGB layers outputs. In our implementation, we use the mean-squared distance between elements in the covariance matrices.

Although \mathcal{L}_{color} encourages matching the overall style of the image, details like skin texture were still not transferred properly (Fig. 3d). One could have a reconstruction loss between the sibling and the output to encourage detail synthesis, but such a loss would be very sensitive to misalignment between the sibling and the StyleGAN2 result, and encourage identity shift. We thus introduce a **contextual loss** \mathcal{L}_{ctx} [Mechrez et al. 2018] between the VGG features of \hat{I}_s and \hat{I} . \mathcal{L}_{ctx} compares each feature in \hat{I} to *all* features in \hat{I}_s , and minimizes the distance to only the most similar one. This allows for the transfer of high-frequency details without requiring precise alignment.



Fig. 4. Abraham Lincoln c. 1863, when the negatives were sensitive only to blue light. The panchromatic reconstruction has exaggerated wrinkles and unnatural colors. Input image: Mead Art Museum (public domain).



Fig. 5. Camera response function fitting (CRF-F) helps low-contrast images recover a wider dynamic range by making their contrast and exposure close to the sibling. Input image: Johann Strauss II (1899) from gallica.bnf.fr / BnF.

4.3 Reconstruction Losses for Antique Images

Rather than fitting the antique input image exactly, we seek a loss that helps preserve the input’s identity while being robust to defects of antique photos. We approach this by introducing a **reconstruction loss** \mathcal{L}_{recon} that applies a series of modifications to the StyleGAN2 result before comparing it to the antique input image. We define \mathcal{L}_{recon} as a loss between the input image I and a modified version of StyleGAN2 output $\hat{I}_d = \mathcal{D}(\hat{I})$, where \mathcal{D} is a degradation process, which attempts to make the generated image appear as if it were captured by an antique camera. In the following, we design \mathcal{D} to account for the spectral sensitivity of early negatives, different camera response functions, possible image blur, and resolution differences.

Antique Film Spectral Sensitivity. We convert StyleGAN2’s output \hat{I} to grayscale, denoted as $\hat{I}_g = \mathcal{G}(\hat{I})$ where \mathcal{G} is the grayscale conversion process. The grayscale conversion must accommodate for the unique sensitivity of early film which is far more sensitive to blue light than red. In particular, we extract the blue channel for *blue-sensitive* photos, average the blue and green intensities ($0.5 \cdot (G+B)$) to approximate *orthochromatic* photos [Geigel and Musgrave 1997], and use standard grayscale conversion ($0.299 \cdot R + 0.587 \cdot G + 0.114 \cdot B$) for *panchromatic* photos. As shown in Fig. 4, choosing the right film model affects the result quality. When multiple spectral sensitivity types are possible (e.g. based on the photo’s capturing time), the user can choose the one that produces the best result.

Camera Response Function. We model the unknown Camera Response Function (CRF) of the input old photo as $a + b\hat{I}_g^\gamma$ where a, b and γ are the bias, gain and gamma parameters to be optimized. During optimization, we initialize with $a = 0$ and $b = \gamma = 1$. To improve convergence, we pre-align the appearance of the input to be closer to the sibling using *histogram matching* [Burger and Burge 2016], i.e., we convert the sibling image to grayscale $\mathcal{G}(\tilde{I}_s)$, apply



Fig. 6. Impact of blur simulation in our reconstruction loss. Best viewed full screen. Input image: John McDouall Stuart (c. 1860) from the State Library of South Australia [2021]: B 501.

histogram transform to the input image to match the grayscale sibling (in the face region only) and produce I' . We then set I' instead of I to be the reconstruction target $\mathcal{L}_{recon}(I', \hat{I}_d)$. See the supplement for more details. We observe in Fig. 5 that when the input images suffer from poor exposure and contrast, CRF fitting helps avoid transferring these artifacts to the output image by bringing the exposure and contrast closer to those of the sibling.

Blur. We finally apply a Gaussian blur with a user-provided standard deviation σ to obtain the final degraded result \hat{I}_d . Values between 0 and 7 work well in our experiments. This blur accounts for the loss of details from the aging, scanning, and capture process (e.g., defocus, low film quality, etc.). Fig. 6 illustrates the benefit of simulating blur during optimization.

Reconstruction Loss. Using the degradation process \mathcal{D} outlined above, we now define our reconstruction losses. To capture the face identity, we *downsample* both I' and $\hat{I}_d = \mathcal{D}(\hat{I})$ from 1024×1024 to 256×256 and compute a perceptual loss between these downsampled images using a combination of VGG [Simonyan and Zisserman 2014] and VGG-Face [Parkhi et al. 2015] features. We add additional constraints to the eye region, which plays an essential part in human perception of faces. Specifically, we *downsample* \hat{I}_d to the original input resolution of I' , crop it and I' in the eye regions to get \hat{I}_d^{eye} and I'^{eye} . We apply VGG-based perceptual loss to reconstruct these eye crops. The complete reconstruction loss is:

$$\begin{aligned} \mathcal{L}_{recon} = & \lambda_{vgg} \mathcal{L}_{vgg}(f(I'), f(\hat{I}_d)) + \lambda_{face} \mathcal{L}_{face}(f(I'), f(\hat{I}_d)) \\ & + \lambda_{eye} \mathcal{L}_{vgg}(f^{eye}, \hat{I}_d^{eye}), \end{aligned} \quad (1)$$

where $f(\cdot)$ is the $4 \times$ downsampling operator to 256×256 .

4.4 Latent Code Optimization

Similar to previous methods [Baylies 2019; Zhu et al. 2016], we first initialize with an encoder output (in our case, the sibling’s latent codes), and then optimize the latent codes, the CRF parameters, and the StyleGAN2 per-layer noise maps. These are optimized to minimize the following loss:

$$\mathcal{L}_{recon} + \lambda_{color} \mathcal{L}_{color} + \lambda_{ctx} \mathcal{L}_{ctx} + \lambda_{noise} \mathcal{L}_{noise} \quad (2)$$

where \mathcal{L}_{noise} is a noise map regularization loss [Karras et al. 2020b]. In StyleGAN2 [Karras et al. 2020b], the latent codes in \mathcal{W}^+ are used at different scales from 4×4 to 1024×1024 . These codes roughly

correspond to different perceptual aspects of an image. The coarser spatial codes determine the overall structure of the face (identity, pose, expression etc.), whereas the finer layers encode aspects like skin tone, skin texture, and lighting. We leverage the expressiveness of W^+ space and optimize codes for layers up to 64×64 , which are sufficient to capture the identity and facial features. The finer spatial codes are copied from those of the sibling.

4.5 Implementation Details

Our method, including sibling computation and latent code optimization, takes about 10 minutes on one NVIDIA TITAN Xp GPU to produce a 1024×1024 result.

Latent Code Optimization. Rather than optimizing latent code layers (4-64) simultaneously, we achieve better results by first optimizing the coarse codes (4-32) for 250 iterations to obtain an intermediate result \hat{I}_{32} . Then, we set \hat{I}_{32} as our new sibling to be used in the color transfer and contextual losses, and jointly optimize latent codes of resolution 4 to 64 for another 750 iterations, producing the final output \hat{I} . Note that the color transfer loss is only enforced on the ToRGB layers of the latent codes being optimized. To navigate the latent space more comprehensively, we also add ramped-down noises to the latent codes as in the projection method of StyleGAN2 [Karras et al. 2020b]. We use the RAdam optimizer [Liu et al. 2020] with default parameters and learning rates 0.1 for the style codes and 0.01 for the camera response function parameters. The weights of each loss are $\lambda_{vgg}=1$, $\lambda_{face}=0.3$, $\lambda_{ctx}=\lambda_{eye}=0.1$, $\lambda_{color}=10^{10}$, $\lambda_{noise}=5 \times 10^9$. See the supplement for details on the specific layers used for all the losses.

Sibling Encoder (E). For each film type described in Section 4.3, we train a sibling ResNet18 encoder E [He et al. 2016] using 16, 128 StyleGAN2-generated samples that are converted to grayscale accordingly. We use an L1 loss between the predicted and ground truth latent codes. We apply color jitter, contrast and exposure augmentations during training. More details are in the supplement.

5 EXPERIMENTS

Dataset. We evaluate our methods on three sets of images. The first consists of a hand-picked set of photos that showcase the most interesting portraits of historical figures. We use this set for visualization purposes. For our second image set, to provide a fair and comprehensive evaluation, we propose a testing benchmark called the *Historical Wiki Face Dataset*. This dataset is collected in an objective manner by automatically crawling Wikipedia and removing unsuitable samples. This set is very diverse, covering various styles and ethnic groups of important historical people. We use this benchmark for testing purposes, i.e., the user study. In our third set, we aim to compare the restored images with ground truth color photos. We asked a professional photographer *Nick Brandreth* [2021] to reproduce a popular antique photographic process, the *gelatin dry plate* [Wikipedia 2021a], which is sensitive to both blue and UV light. He captured the same subjects under the same outdoor lighting in black and white with the dry plate and in color with the modern DSLR camera from similar viewpoints (Fig. 11). Please see our supplement for results from all three image sets.

Table 1. Quantitative comparisons between our approach and baseline pipelines composed of restoration, colorization and super-resolution methods. The NIQE score of the input images is also reported as a reference.

Input	DeOldify	Zhang	InstColorization	Zhang (FFHQ)	Ours
NIQE ↓	7.08	5.54	5.52	5.49	5.47 4.55

Historical Wiki Face Dataset. We collect a list of names from the “Significant people” section of the Wikipedia pages for “19th Century” and “20th Century”, and crawl down the main *page images* [MediaWiki 2021] on their Wikipedia pages. To filter out images unsuitable for our task, we remove the image if it belongs to any of the following categories: 1) The image is not a photo (e.g., a painting) or is in color. 2) The face is too small, i.e., less than 130 pixels. 3) The head is not fully visible. 4) Hands touch or occlude the face. 5) Face detection [King 2009] fails. 6) The photo is heavily retouched (e.g., manually colorized daguerreotypes). 7) The subject is politically controversial. The result of this filtering is 224 photos of unique historical figures dating from 19th century to 20th century, including people like Abraham Lincoln, Marie Curie, Winston Churchill, and Franz Kafka (e.g., Figs 1, 10). We cropped out the head region using the face alignment method of Karras et al. [2019] and resized them to have a maximum resolution of 1024×1024 (average 638×638 , min 133×133). Our test set covers a wide range of image quality, head poses, genders, ethnic groups, and historical fashion styles that often drastically differ from modern ones. Subjects in the early 19th century, for instance, often have big curly mustaches, long beards, and shaggy hair. Additionally, many antique accessories are uncommon in modern imagery, such as crowns, stand-up collars, and pince-nez. Fig. 1, 4, 5, 7, 12 all include samples from this test set.

Experiment Setup. In our experiments we manually select the input blur kernel. For film model, we use the blue-sensitive model for photos before 1873, manually select between blue-sensitive and orthochromatic for images from 1873 to 1906 and among all models for photos taken afterwards.

There is no published baseline method that performs the full complement of image restoration operators needed for antique photos, i.e., noise+blur removal, contrast adjustment, colorization, and super-resolution. We therefore compare our approach to sequentially applying state-of-the-art methods for each of these tasks. As a first step, we apply Wan et al. [2020], which was specifically designed to remove noise and artifacts in antique portrait photos at resolution 512×512 . We tried restoring at 1024×1024 but found the method produced blurrier and more noisy results compared to apply at 512×512 followed by a separate super-resolution technique (detailed later). As a second step, we colorize the image. We evaluated several colorization techniques including DeOldify [Antic 2019], InstColorization [Su et al. 2020], Zhang et al. [2017a]. All of these methods are designed for generic scenes and perform worse on antique portraits. We therefore retrained Zhang’s colorization network using the FFHQ dataset of face images [Karras et al. 2019], denoted *Zhang (FFHQ)*. We also augmented this training dataset by applying random Gaussian blur and noise to make their method more robust to antique imagery. As a final step, we use the SRFBN [Li et al. 2019] (BI model) to super-resolve ($2 \times$) the colorized output to our target resolution of 1024×1024 . For simplicity, we use the name of

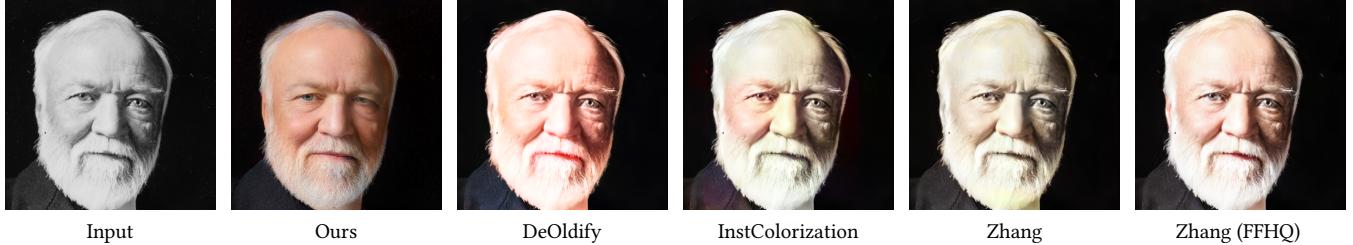


Fig. 7. Comparisons of our approach to a pipeline built from published techniques for restoring, colorizing, and super-resolving old photos. We evaluate against four baseline pipelines, each with a different colorization algorithm, detailed in Sec. 5. All of them fail to achieve the same realistic skin appearance and overall image quality as our approach. Input image: Andrew Carnegie (c. 1913) from Library of Congress.

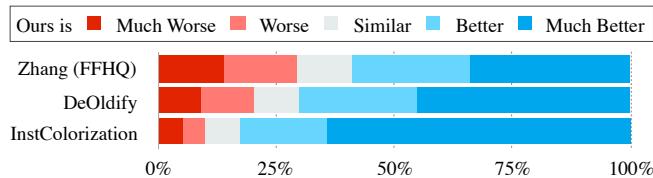


Fig. 8. User ratings of our approach compared to baselines composed of state-of-the-art restoration, colorization and super-resolution methods.

the colorization method to refer to the full pipeline baseline that is composed of Wan et al. [2020], one of the four colorization methods, and SRFBN [Li et al. 2019].

Qualitative Evaluation. Figure 7 shows a comparison between our method and the pipelines described above, including each of the four colorization methods. Our method outperforms the other baselines and produces much more photo-realistic skin texture and colors. Figure 10 shows a higher-resolution comparison between our method and our strongest baseline, *Zhang (FFHQ)*. None of the baseline methods are able to reproduce realistic skin appearance and sharp image details as well as ours.

Quantitative Evaluation. Table 1 quantitatively compares our method with the baseline pipelines on their abilities to resolve degradations in the luna component, such as denoising, deblurring, enhancing contrast and details. We extract the Y' channel from the Y'CrCb color space for all the color images, and measure the quality of the restored results using the no-reference metric NIQE [Mittal et al. 2012]. We also report the NIQE score for the input images as a reference. Our results are shown to produce more details and outperform all baseline methods in NIQE.

User Study. We conducted a user study over the *Historical Wiki Face Dataset*. 39 users participated in the study, where each participant was presented a random set of 56 or 112 pairwise comparisons of our result and one of the three top-performing baseline pipelines (*Zhang (FFHQ)*, *DeOldify* and *InstColorization*). The input image was also presented. The participants were asked to choose one of the two results and rated among “significantly better”, “slightly better”, or “similar”, when asked “Which of the two images is a higher-quality portrait?” We obtained answers for all 224 images in our dataset and all three baseline methods. The results of this study are presented in Fig. 8. Our approach is consistently and significantly preferred over *InstColorization* and *DeOldify*. Even compared with the strongest baseline, 58.9% of the time participants perceived our results to be

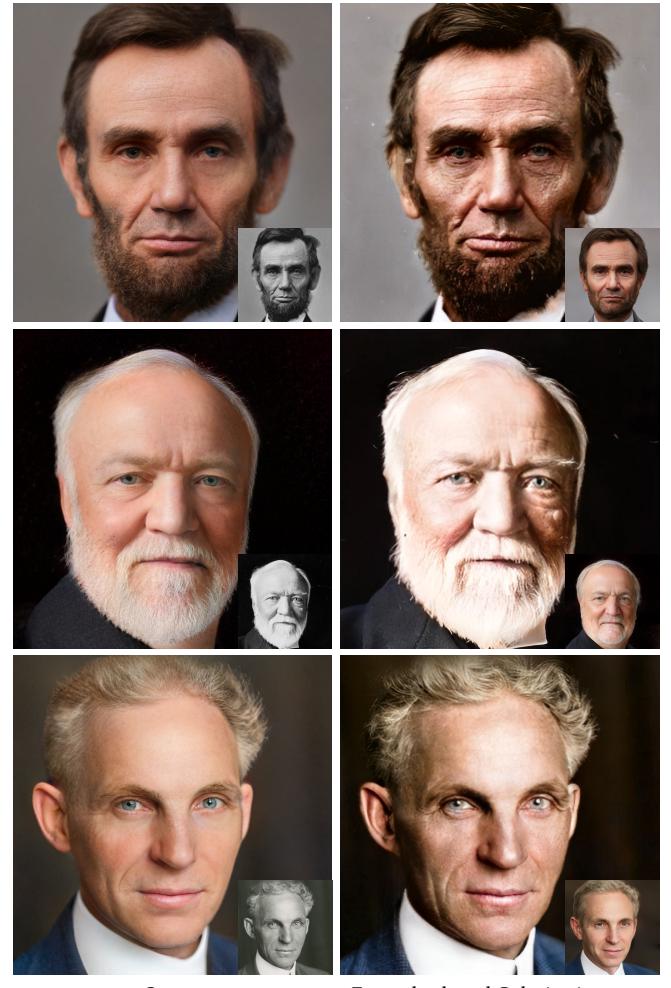


Fig. 9. Visual comparisons with the exemplar-based colorization technique by He et al. [2018] using sibling images as the exemplars. The input and sibling images are shown as insets in the first and second columns, respectively. Input images (top to bottom): Abraham Lincoln (1863) from Mead Art Museum, Andrew Carnegie (1913) from Library of Congress, Henry Ford (1863 - 1947) from the Collections of the Henry Ford. better compared to 29.4% for *Zhang (FFHQ)*. 11.7% expressed no strong preference.

Antique vs. Modern Photographic Processes. In Fig. 11, we use the gelatin dry plate photos as input, restore the images with the



Fig. 10. Visual comparison with the top-performing baseline, *Zhang (FFHQ)*. The input and sibling images are shown as insets in the first and second columns, respectively. Results are best evaluated at 1024×1024 . From top to bottom: Werner Heisenberg (1933) by BArch, Bild 183-R57262 / Unknown author / CC-BY-SA 3.0, Winston Churchill (1941) by Library and Archives Canada / flickr, and Clara Barton (c. 1904) by Library of Congress.

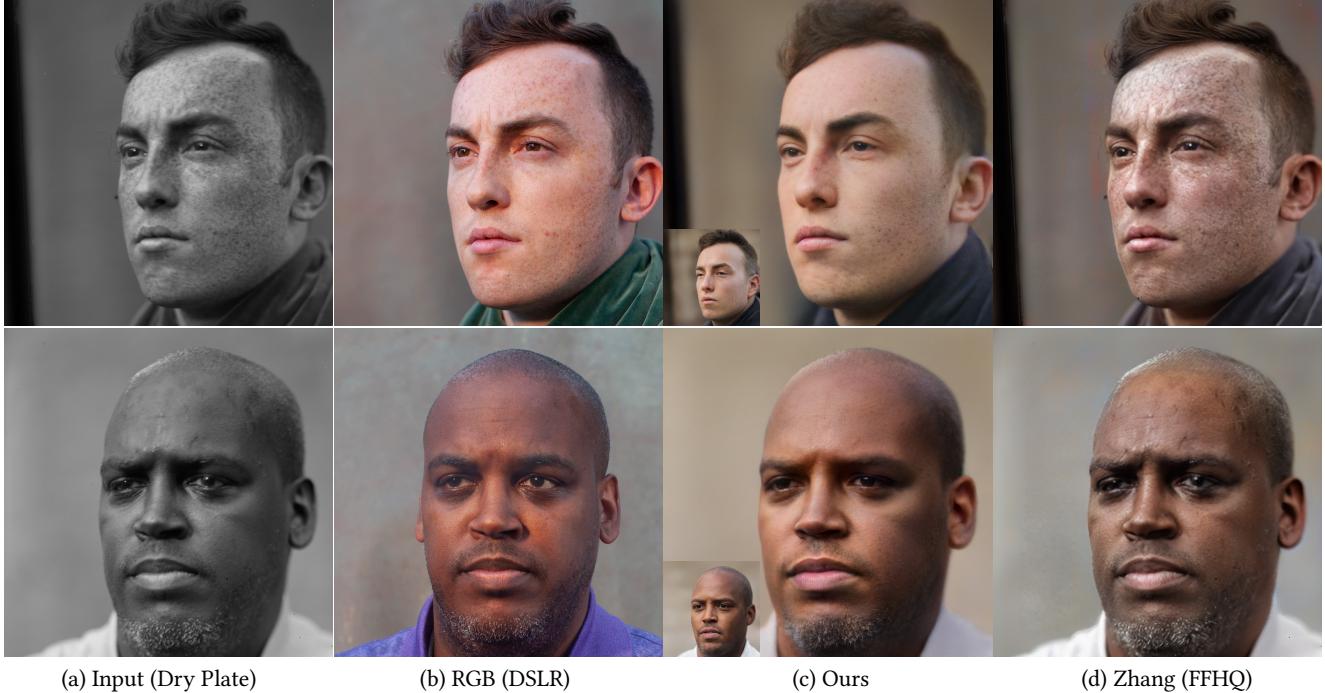


Fig. 11. Comparisons of restored images from an antique photographic process with the ground truth color photos. The same subject is captured under the same lighting with both (b) a modern DSLR camera and an antique photographic process – (a) *gelatin dry plate* [Wikipedia 2021a] ©Nick Brandreth, which is sensitive to both blue and UV light. We compare results from (c) our approach against the strongest baseline pipeline (d) *Zhang (FFHQ)*. Our results tone down the exaggerated pimples, freckles, and skin specularity due to blue sensitivity and produces a more natural skin texture.

strongest baseline, *Zhang (FFHQ)*, and compare with our method using the blue-sensitive model. In the top row, pimples and freckles are dramatically exaggerated, and skin appears more weathered and specular in the blue-sensitive gelatin dry plate. The results by *Zhang (FFHQ)* further exaggerate these features. While being slightly smooth, our results restore a more natural skin texture and specularity that better resembles the DSLR reference. Similarly, in the bottom row, the baseline exaggerates skin defects in the forehead and adds sharp specular effects that are reduced in our result.

Comparison with Exemplar-based Colorization. Fig. 9 compares our method with prior exemplar-based colorization method [He et al. 2018]. Similar to other baselines, We apply the same restoration [Wan et al. 2020] and super-resolution [Li et al. 2019] method but adopt the colorization method by He et al. [2018] using our sibling as the exemplar. Our sibling images as exemplars improves the overall color and tone, but the method still fails to remove artifacts in the luminance channel, nor does it account for the unique spectral response properties of antique negatives. As a result, our method produces more detailed and realistic results.

6 LIMITATIONS AND FUTURE WORK

As illustrated in Fig. 12, our method does not work equally well on all images, and inherits biases from the StyleGAN2 image generator. Historical hairstyles, accessories, and clothing that are very different from anything present in the StyleGAN2 training set are not reproduced well. Extreme head poses are also rare in the training data and harder to restore. In these cases, it is challenging to create a

StyleGAN2 sibling with such uncommon features (Fig. 12a-d), leading to inferior results. In some cases the sibling presents a different gender or ethnicity, which affects the synthesized result (Fig. 12). As shown by Salminen et al. [2020], StyleGAN’s generated images have strong biases towards younger and White people: 72.6% of generated images represent White, 13.8% Asian, 10.1% Black, and 3.4% Indian people. These biases in StyleGAN in turn lead to biases of our method towards predicting lighter skin tones for some inputs. Note that the brightness and contrast of the input image can also affect the predicted skin color. Addressing these gender and ethnicity shift problems is an important topic of future research. We believe this problem can be addressed by training StyleGAN2 with datasets of balanced race, gender, and age, such as the *FairFace* [Karkkainen and Joo 2021] dataset, which in turn can improve our method with a more balanced sibling encoder.

Another limitation of our method is inaccurately predicting skin texture from images with compressed intensity gamuts (Fig. 12). While camera response fitting alleviates poor exposure and contrast from the antique photos to some extent, handling more challenging cases requires hallucinating higher dynamic range. Our performance also degrades on images with severe noise.

In some cases, our method alters the shape of certain facial features such as the eyes (Kafka’s right eye in Fig. 10), wrinkles (Lincoln’s forehead wrinkles in Fig. 1), glasses (glasses of Gandhi in Fig. 1), etc. This is because we only optimize global $\mathcal{W}+$ style codes with a strong noise regularizer, making it hard to preserve local

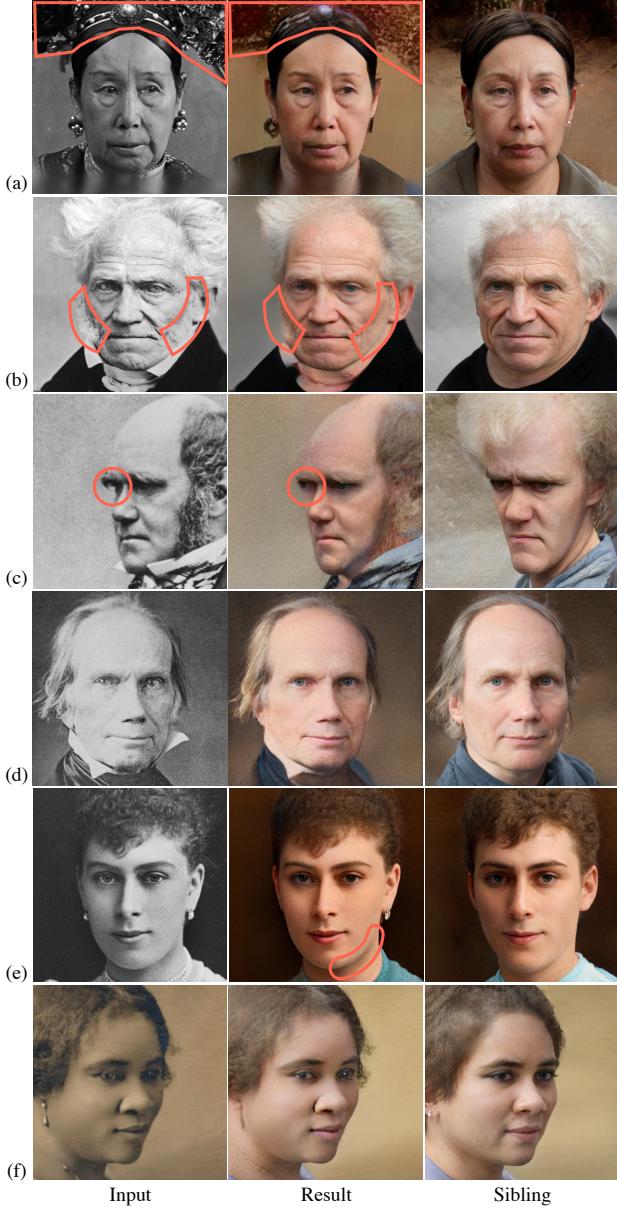


Fig. 12. Our approach can struggle with features that are not well represented in StyleGAN2 such as uncommon accessories, clothing, (facial) hairstyles, and extreme head poses (a-c). Extremely poor image quality or severely compressed intensity gamuts may also limit the quality of the result (d). The sibling image may also present a different gender or ethnicity than the input (e-f). Input images (top to bottom): Cixi, Empress Dowager of China (1835 - 1908), FSA A.13, Freer Gallery of Art and Arthur M. Sackler Gallery Archives, Arthur Schopenhauer (1859) from Goethe-Universität Frankfurt am Main, Charles Darwin (c. 1854) by GrrlScientist/flickr, Henry Clay (1848) by Jim Surkamp/flickr, Mary of Teck (1867 - 1953) from Royal Collection Trust/©Her Majesty Queen Elizabeth II 2021, and Madam C. J. Walker (c. 1914) from Smithsonian Institution, National Museum of American History.

image details. Integrating local features from the input image with StyleGAN2 face prior is another direction for future work.

Recovering the correct color for skin, eyes, or clothes, is challenging (Fig. 11) as many possible colors can correspond to the same degraded photo. One avenue of future work is to predict the distribution of likely color outputs, or guide the color prediction using references from paintings and textural descriptions.

7 CONCLUSION

We introduced *time-travel rephotography*, an image synthesis technique that simulates rephotographing famous subjects from the past using a modern high-resolution camera based on a black-and-white reference photo. Our basic approach is to project this reference image into the space of modern high-resolution images represented by the StyleGAN2 generative model [Karras et al. 2020a]. This is accomplished through a constrained optimization over latent style codes that is guided by a novel reconstruction procedure that simulates the unique properties of old film and cameras. We also introduce our *sibling* encoders that generate an image to recover colors and local spatial details in the result. Improving on applying a sequence of state-of-the-art techniques for image restoration, colorization, and super-resolution, our unified approach is able to render strikingly realistic and immediately recognizable images of historical figures.

ACKNOWLEDGMENTS

We thank Roy Or-El, Aleksander Holynski and Keunhong Park for insightful advice. This work was supported by the UW Reality Lab, Amazon, Facebook, Futurewei, and Google.

REFERENCES

- Rameen Abdal, Yipeng Qin, and Peter Wonka. 2019. Image2stylegan: How to embed images into the stylegan latent space?. In *Proceedings of the IEEE International Conference on Computer Vision*. 4432–4441.
- Rameen Abdal, Yipeng Qin, and Peter Wonka. 2020. Image2StyleGAN++: How to Edit the Embedded Images?. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. 8296–8305.
- Ansel Adams. 2018. *The negative*. Ansel Adams. 21–25 pages.
- Jason Antic. 2019. jantic/deoldify: A deep learning based project for colorizing and restoring old images (and video!). <https://github.com/jantic/DeOldify>
- S Derin Babacan, Rafael Molina, and Aggelos K Katsaggelos. 2008. Total variation super resolution using a variational approach. In *IEEE International Conference on Image Processing*. IEEE, 641–644.
- Peter Baylies. 2019. Stylegan encoder - converts real images to latent space. <https://github.com/pbaylies/stylegan-encoder>
- Nick Brandreth. 2021. Nick Brandreth. <https://www.nickbrandreth.com/>.
- Antoni Buades, Bartomeu Coll, and J-M Morel. 2005. A non-local algorithm for image denoising. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, Vol. 2. IEEE, 60–65.
- Adrian Bulat and Georgios Tzimiropoulos. 2018. Super-fan: Integrated facial landmark localization and super-resolution of real-world low resolution faces in arbitrary poses with gans. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. 109–117.
- Wilhelm Burger and Mark J Burge. 2016. *Digital image processing: an algorithmic introduction using Java*. Springer.
- Guillaume Charpiat, Matthias Hofmann, and Bernhard Schölkopf. 2008. Automatic image colorization via multimodal predictions. In *Proceedings of the European Conference on Computer Vision*. Springer, 126–139.
- Zechou Cheng, Qingxiang Yang, and Bin Sheng. 2015. Deep colorization. In *Proceedings of the IEEE International Conference on Computer Vision*. 415–423.
- Alex Yong-Sang Chia, Shaojie Zhuo, Raj Kumar Gupta, Yu-Wing Tai, Siu-Yeung Cho, Ping Tan, and Stephen Lin. 2011. Semantic colorization with internet images. *ACM Transactions on Graphics (TOG)* 30, 6 (2011), 1–8.
- Kostadin Dabov, Alessandro Foi, Vladimir Katkovnik, and Karen Egiazarian. 2007. Image denoising by sparse 3-D transform-domain collaborative filtering. *IEEE Transactions on Image Processing* 16, 8 (2007), 2080–2095.
- Aditya Deshpande, Jason Rock, and David Forsyth. 2015. Learning large-scale automatic image colorization. In *Proceedings of the IEEE International Conference on Computer Vision*. 567–575.

- Chao Dong, Yubin Deng, Chen Change Loy, and Xiaou Tang. 2015a. Compression artifacts reduction by a deep convolutional network. In *Proceedings of the IEEE International Conference on Computer Vision*. 576–584.
- Chao Dong, Chen Change Loy, Kaiming He, and Xiaou Tang. 2015b. Image super-resolution using deep convolutional networks. *IEEE transactions on pattern analysis and machine intelligence* 38, 2 (2015), 295–307.
- Michael Elad and Michael Aharon. 2006. Image denoising via sparse and redundant representations over learned dictionaries. *IEEE Transactions on Image Processing* 15, 12 (2006), 3736–3745.
- Leon A Gatys, Alexander S Ecker, and Matthias Bethge. 2016. Image style transfer using convolutional neural networks. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. 2414–2423.
- Joe Geigel and F Kenton Musgrave. 1997. A model for simulating the photographic development process on digital images. In *Proceedings of the annual Conference on Computer graphics and interactive techniques*. 135–142.
- Klemen Grm, Walter J Scheirer, and Vitomir Štruc. 2019. Face hallucination using cascaded super-resolution and identity priors. *IEEE Transactions on Image Processing* 29, 1 (2019), 2150–2165.
- Jun Guo and Hongyang Chao. 2016. Building dual-domain representations for compression artifacts reduction. In *Proceedings of the European Conference on Computer Vision*. Springer, 628–644.
- Raj Kumar Gupta, Alex Yong-Sang Chia, Deepu Rajan, Ee Sin Ng, and Huang Zhiyong. 2012. Image colorization using similar images. In *Proceedings of the ACM International Conference on Multimedia*. 369–378.
- Yoav Hocherman, Eli Shechtman, and Dani Lischinski. 2013. Deblurring by example using dense correspondence. In *Proceedings of the IEEE International Conference on Computer Vision*. 2384–2391.
- Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. 2016. Deep residual learning for image recognition. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. 770–778.
- Mingming He, Dongdong Chen, Jing Liao, Pedro V Sander, and Lu Yuan. 2018. Deep exemplar-based colorization. *ACM Transactions on Graphics (TOG)* 37, 4 (2018), 1–16.
- Yi-Chin Huang, Yi-Shin Tung, Jun-Cheng Chen, Sung-Wen Wang, and Ja-Ling Wu. 2005. An adaptive edge detection based colorization algorithm and its applications. In *Proceedings of the annual ACM International Conference on Multimedia*. 351–354.
- Satoshi Iizuka, Edgar Simo-Serra, and Hiroshi Ishikawa. 2016. Let there be color! Joint end-to-end learning of global and local image priors for automatic image colorization with simultaneous classification. *ACM Transactions on Graphics (ToG)* 35, 4 (2016), 1–11.
- Revital Ironi, Daniel Cohen-Or, and Dani Lischinski. 2005. Colorization by Example. *Rendering Techniques*. Citeseer, 201–210.
- Phillip Isola, Jun-Yan Zhu, Tinghui Zhou, and Alexei A Efros. 2017. Image-to-image translation with conditional adversarial networks. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. 1125–1134.
- Henrik Wann Jensen, Stephen R Marschner, Marc Levoy, and Pat Hanrahan. 2001. A practical model for subsurface light transport. In *Proceedings of the 28th annual conference on Computer graphics and interactive techniques*. 511–518.
- Kimmo Karkkainen and Jungseok Joo. 2021. FairFace: Face Attribute Dataset for Balanced Race, Gender, and Age for Bias Measurement and Mitigation. In *Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision*. 1548–1558.
- Tero Karras, Samuli Laine, and Timo Aila. 2019. A style-based generator architecture for generative adversarial networks. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. 4401–4410.
- Tero Karras, Samuli Laine, Miika Aittala, Janne Hellsten, Jaakko Lehtinen, and Timo Aila. 2020a. Analyzing and improving the image quality of stylegan. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. 8110–8119.
- Tero Karras, Samuli Laine, Miika Aittala, Janne Hellsten, Jaakko Lehtinen, and Timo Aila. 2020b. Analyzing and Improving the Image Quality of StyleGAN. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*.
- Jiwon Kim, Jung Kwon Lee, and Kyoung Mu Lee. 2016. Accurate image super-resolution using very deep convolutional networks. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. 1646–1654.
- Davis E. King. 2009. Dlib-ml: A Machine Learning Toolkit. *Journal of Machine Learning Research* 10 (2009), 1755–1758.
- Orest Kupyn, Volodymyr Budzan, Mykola Mykhailych, Dmytro Mishkin, and Jiří Matas. 2018. Deblurgan: Blind motion deblurring using conditional adversarial networks. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. 8183–8192.
- Gustav Larsson, Michael Maire, and Gregory Shakhnarovich. 2016. Learning representations for automatic colorization. In *Proceedings of the European Conference on Computer Vision*. Springer, 577–593.
- Christian Ledig, Lucas Theis, Ferenc Huszár, Jose Caballero, Andrew Cunningham, Alejandro Acosta, Andrew Aitken, Alykhan Tejani, Johannes Totz, Zehan Wang, et al. 2017. Photo-realistic single image super-resolution using a generative adversarial network. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. 4681–4690.
- Stamatios Lefkimiatis. 2017. Non-local color image denoising with convolutional neural networks. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. 3587–3596.
- Anat Levin, Dani Lischinski, and Yair Weiss. 2004. Colorization using optimization. In *ACM SIGGRAPH 2004 Papers*. 689–694.
- Zhen Li, Jinglei Yang, Zheng Liu, Xiaomin Yang, Gwanggil Jeon, and Wei Wu. 2019. Feedback network for image super-resolution. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. 3867–3876.
- Liyuan Liu, Haoming Jiang, Pengcheng He, Weizhu Chen, Xiaodong Liu, Jianfeng Gao, and Jiawei Han. 2020. On the Variance of the Adaptive Learning Rate and Beyond. In *Proceedings of the Eighth International Conference on Learning Representations (ICLR 2020)*.
- Xiaopei Liu, Liang Wan, Yingge Qu, Tien-Tsin Wong, Stephen Lin, Chi-Sing Leung, and Pheng-Ann Heng. 2008. Intrinsic colorization. In *ACM SIGGRAPH Asia 2008 papers*. 1–9.
- Qing Luan, Fang Wen, Daniel Cohen-Or, Lin Liang, Ying-Qing Xu, and Heung-Yeung Shum. 2007. Natural image colorization. In *Proceedings of the Eurographics Conference on Rendering Techniques*. 309–320.
- Roey Mechrez, Itamar Talmi, and Lihi Zelnik-Manor. 2018. The contextual loss for image transformation with non-aligned data. In *Proceedings of the European Conference on Computer Vision*. 768–783.
- MediaWiki. 2021. MediaWiki: PageImages. <https://www.mediawiki.org/wiki/Extension:PageImages>.
- Sachit Menon, Alex Damiani, McCourt Hu, Nikhil Ravi, and Cynthia Rudin. 2020a. PULSE: Self-Supervised Photo Upsampling via Latent Space Exploration of Generative Models. In *The IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*.
- Sachit Menon, Alexandru Damian, Shijia Hu, Nikhil Ravi, and Cynthia Rudin. 2020b. PULSE: Self-Supervised Photo Upsampling via Latent Space Exploration of Generative Models. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. 2437–2445.
- Anish Mittal, Rajiv Soundararajan, and Alan C Bovik. 2012. Making a “completely blind” image quality analyzer. *IEEE Signal processing letters* 20, 3 (2012), 209–212.
- Seungjun Nah, Tae Hyun Kim, and Kyoung Mu Lee. 2017. Deep multi-scale convolutional neural network for dynamic scene deblurring. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. 3883–3891.
- Beaumont Newhall. 1982. *The History of Photography: From 1839 to the Present* (5 ed.). The Museum of Modern Art.
- State Library of South Australia. 2021. Photo of John McDouall Stuart. <https://www.catalog.slsa.sa.gov.au/record=b2049594-S1>.
- Jinshan Pan, Zhe Hu, Zhixun Su, and Ming-Hsuan Yang. 2014. Deblurring face images with exemplars. In *Proceedings of the European Conference on Computer Vision*. Springer, 47–62.
- Xingang Pan, Xiaohang Zhan, Bo Dai, Dahua Lin, Chen Change Loy, and Ping Luo. 2020. Exploiting Deep Generative Prior for Versatile Image Restoration and Manipulation. In *European Conference on Computer Vision (ECCV)*.
- Omkar M. Parkhi, Andrea Vedaldi, and Andrew Zisserman. 2015. Deep Face Recognition. In *British Machine Vision Conference*.
- Yingge Qu, Tien-Tsin Wong, and Pheng-Ann Heng. 2006. Manga colorization. *ACM Transactions on Graphics (TOG)* 25, 3 (2006), 1214–1220.
- Wenqi Ren, Jiaolong Yang, Senyou Deng, David Wipf, Xiaochun Cao, and Xin Tong. 2019. Face video deblurring using 3D facial priors. In *Proceedings of the IEEE International Conference on Computer Vision*. 9388–9397.
- Elad Richardson, Yuval Alaluf, Or Patashnik, Yotam Nitzan, Yaniv Azar, Stav Shapiro, and Daniel Cohen-Or. 2020. Encoding in Style: a StyleGAN Encoder for Image-to-Image Translation. *arXiv preprint arXiv:2008.00951* (2020).
- Joni Salminen, Soon-gyo Jung, Shammur Chowdhury, and Bernard J Jansen. 2020. Analyzing demographic bias in artificially generated facial pictures. In *Extended Abstracts of the 2020 CHI Conference on Human Factors in Computing Systems*. 1–8.
- Ziyi Shen, Wei-Sheng Lai, Tingfa Xu, Jan Kautz, and Ming-Hsuan Yang. 2018. Deep semantic face deblurring. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. 8260–8269.
- Karen Simonyan and Andrew Zisserman. 2014. Very deep convolutional networks for large-scale image recognition. *arXiv preprint arXiv:1409.1556* (2014).
- Jheng-Wei Su, Hung-Kuo Chu, and Jia-Bin Huang. 2020. Instance-aware Image Colorization. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*.
- Masanori Suganuma, Xing Liu, and Takayuki Okatani. 2019. Attention-based adaptive selection of operations for image restoration in the presence of unknown combined distortions. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. 9039–9048.
- Jian Sun, Wenfei Cao, Zongben Xu, and Jean Ponce. 2015. Learning a convolutional neural network for non-uniform motion blur removal. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. 769–777.

- Daniel Sýkora, John Dingliana, and Steven Collins. 2009. Lazybrush: Flexible painting tool for hand-drawn cartoons. *Computer Graphics Forum* 28, 2 (2009), 599–608.
- Ying Tai, Jian Yang, Xiaoming Liu, and Chunyan Xu. 2017. Memnet: A persistent memory network for image restoration. In *Proceedings of the IEEE International Conference on Computer Vision*. 4539–4547.
- Omer Tov, Yuval Alaluf, Yotam Nitzan, Or Patashnik, and Daniel Cohen-Or. 2021. Designing an Encoder for StyleGAN Image Manipulation. *arXiv preprint arXiv:2102.02766* (2021).
- Ziyu Wan, Bo Zhang, Dongdong Chen, Pan Zhang, Dong Chen, Jing Liao, and Fang Wen. 2020. Old Photo Restoration via Deep Latent Space Translation. *arXiv:arXiv:2009.07047*.
- Xintao Wang, Yu Li, Honglun Zhang, and Ying Shan. 2021. Towards Real-World Blind Face Restoration with Generative Facial Prior. In *The IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*.
- Zhangyang Wang, Ding Liu, Shiyu Chang, Qing Ling, Yingzhen Yang, and Thomas S Huang. 2016. D3: Deep dual-domain based fast restoration of JPEG-compressed images. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. 2764–2772.
- Tomohisa Welsh, Michael Ashikhmin, and Klaus Mueller. 2002. Transferring color to greyscale images. In *Proceedings of the Conference on Computer graphics and interactive techniques*. 277–280.
- Wikipedia. 2021a. Dry Plate. https://en.wikipedia.org/wiki/Dry_plate.
- Wikipedia. 2021b. Wikipedia: Photographic film. https://en.wikipedia.org/wiki/Photographic_film.
- Wikipedia. 2021c. Wikipedia: Rephotography. <https://en.wikipedia.org/wiki/Rephotography>.
- Junyuan Xie, Linli Xu, and Enhong Chen. 2012. Image denoising and inpainting with deep neural networks. In *Advances in neural information Processing systems*. 341–349.
- Li Xu, Jimmy SJ Ren, Ce Liu, and Jiaya Jia. 2014a. Deep convolutional neural network for image deconvolution. In *Advances in Neural Information Processing Systems*. 1790–1798.
- Li Xu, Xin Tao, and Jiaya Jia. 2014b. Inverse kernels for fast spatial deconvolution. In *Proceedings of the European Conference on Computer Vision*. Springer, 33–48.
- Jianchao Yang, John Wright, Thomas S Huang, and Yi Ma. 2010. Image super-resolution via sparse representation. *IEEE transactions on image Processing* 19, 11 (2010), 2861–2873.
- Tao Yang, Ren Peiran, Xie Xuansong, and Lei Zhang. 2021. GAN Prior Embedded Network for Blind Face Restoration in the Wild. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*.
- Liron Yatziv and Guillermo Sapiro. 2006. Fast image and video colorization using chrominance blending. *IEEE transactions on Image Processing* 15, 5 (2006), 1120–1129.
- Ke Yu, Chao Dong, Liang Lin, and Chen Change Loy. 2018. Crafting a Toolchain for Image Restoration by Deep Reinforcement Learning. In *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition*. 2443–2452.
- Kai Zhang, Wangmeng Zuo, Yunjin Chen, Deyu Meng, and Lei Zhang. 2017b. Beyond a gaussian denoiser: Residual learning of deep cnns for image denoising. *IEEE Transactions on Image Processing* 26, 7 (2017), 3142–3155.
- Kai Zhang, Wangmeng Zuo, Shuhang Gu, and Lei Zhang. 2017c. Learning deep CNN denoiser prior for image restoration. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. 3929–3938.
- Kai Zhang, Wangmeng Zuo, and Lei Zhang. 2018b. FFDNet: Toward a fast and flexible solution for CNN-based image denoising. *IEEE Transactions on Image Processing* 27, 9 (2018), 4608–4622.
- Richard Zhang, Phillip Isola, and Alexei A Efros. 2016. Colorful image colorization. In *Proceedings of the European Conference on Computer Vision*. Springer, 649–666.
- Richard Zhang, Jun-Yan Zhu, Phillip Isola, Xinyang Geng, Angela S Lin, Tianhe Yu, and Alexei A Efros. 2017a. Real-time user-guided image colorization with learned deep priors. *ACM Transactions on Graphics (TOG)* 36, 4 (2017), 1–11.
- Yulun Zhang, Yapeng Tian, Yu Kong, Bineng Zhong, and Yun Fu. 2018a. Residual dense network for image super-resolution. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. 2472–2481.
- Jiajiao Zhao, Li Liu, Cees GM Snoek, Jungong Han, and Ling Shao. 2018. Pixel-level semantics guided image colorization. *arXiv preprint arXiv:1808.01597* (2018).
- Jiapeng Zhu, Yujun Shen, Deli Zhao, and Bolei Zhou. 2020. In-domain GAN Inversion for Real Image Editing. In *Proceedings of European Conference on Computer Vision (ECCV)*.
- Jun-Yan Zhu, Philipp Krähenbühl, Eli Shechtman, and Alexei A Efros. 2016. Generative visual manipulation on the natural image manifold. In *European conference on computer vision*. Springer, 597–613.