




# Simulating analogue film damage to analyse and improve artefact restoration on high-resolution scans

D. Ivanova<sup>1</sup>  and J. Williamson<sup>1</sup>  and P. Henderson<sup>1</sup> 

<sup>1</sup>University of Glasgow, UK

## Abstract

Digital scans of analogue photographic film typically contain artefacts such as dust and scratches. Automated removal of these is an important part of preservation and dissemination of photographs of historical and cultural importance. While state-of-the-art deep learning models have shown impressive results in general image inpainting and denoising, film artefact removal is an understudied problem. It has particularly challenging requirements, due to the complex nature of analogue damage, the high resolution of film scans, and potential ambiguities in the restoration. There are no publicly available high-quality datasets of real-world analogue film damage for training and evaluation, making quantitative studies impossible. We address the lack of ground-truth data for evaluation by collecting a dataset of 4K damaged analogue film scans paired with manually-restored versions produced by a human expert, allowing quantitative evaluation of restoration performance. We have made the dataset available at <https://doi.org/10.6084/m9.figshare.21803304>. We construct a larger synthetic dataset of damaged images with paired clean versions using a statistical model of artefact shape and occurrence learnt from real, heavily-damaged images. We carefully validate the realism of the simulated damage via a human perceptual study, showing that even expert users find our synthetic damage indistinguishable from real. In addition, we demonstrate that training with our synthetically damaged dataset leads to improved artefact segmentation performance when compared to previously proposed synthetic analogue damage overlays. The synthetically damaged dataset can be found at <https://doi.org/10.6084/m9.figshare.21815844>, and the annotated authentic artefacts along with the resulting statistical damage model at <https://github.com/daniela997/FilmDamageSimulator>. Finally, we use these datasets to train and analyse the performance of eight state-of-the-art image restoration methods on high-resolution scans. We compare both methods which directly perform the restoration task on scans with artefacts, and methods which require a damage mask to be provided for the inpainting of artefacts. We modify the methods to process the inputs in a patch-wise fashion to operate on original high resolution film scans.

## 1. Introduction

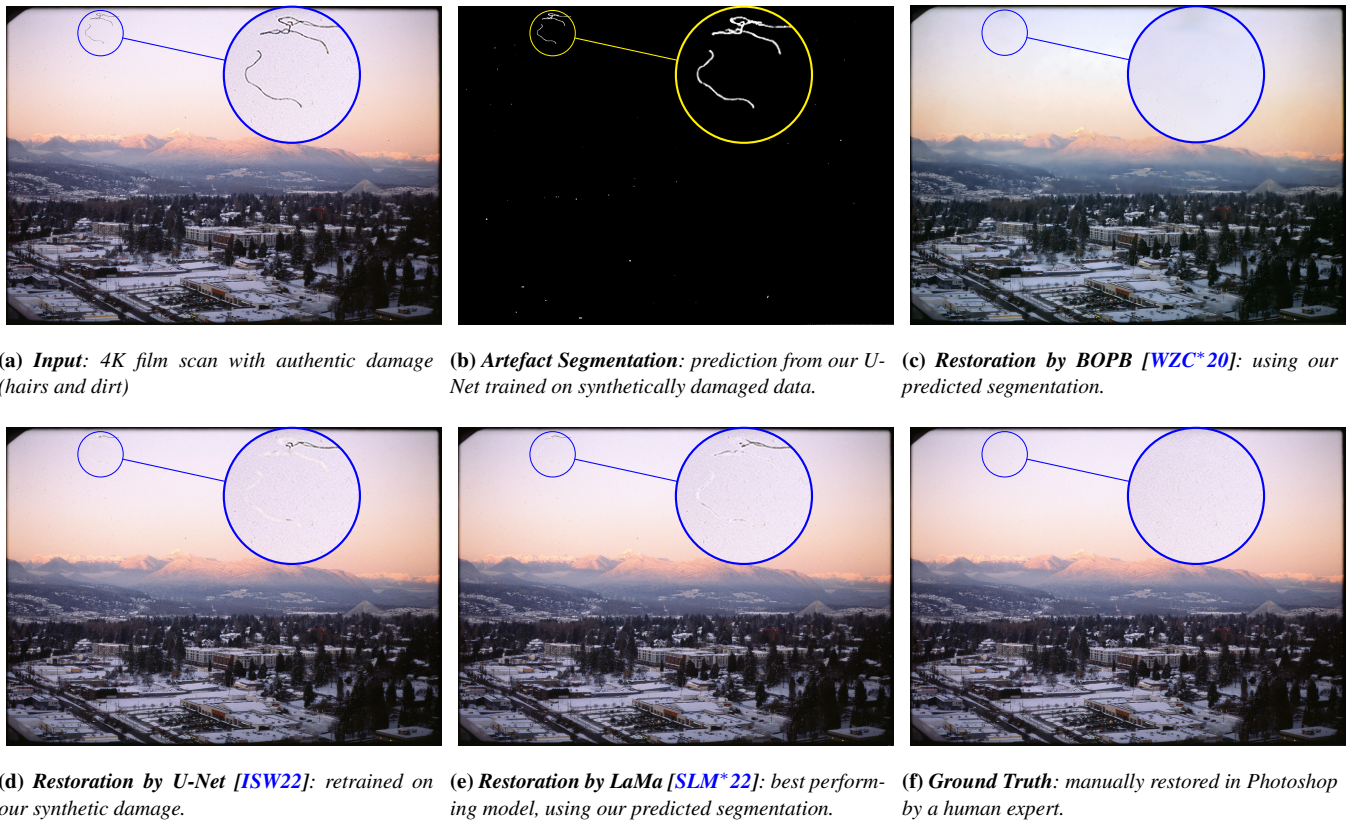
Photographs captured on film constitute a major part of our cultural heritage and historical record, and many photographers continue to shoot on film. While it has appealing imaging qualities, film emulsion is highly susceptible to various kinds of mechanical damage: scratches, dust, hairs and dirt which mar the image when scanned and persist in darkroom and digital prints, as well as in cinematographic release prints.

Isolating dust and scratches from natural image features in film scans and inpainting the damage are not difficult tasks for humans, but can be extremely time-consuming. **Automated film restoration** aims to localise and inpaint artefacts at a quality level comparable or exceeding that of human experts.

This is challenging: mechanical artefacts can have complex shapes and are non-uniformly distributed. They must be localised accurately when performing restoration, as it is crucial that restoration is applied only on the affected areas. It is also crucial to preserve the desirable qualities of analogue film, such as grain or

characteristic colour grading, avoiding over-smoothing or distorted colour distributions.

Automated film restoration can be broken down into two sub-tasks: artefact localisation (segmentation), and inpainting. Inpainting [LRS\*18, SLM\*22] and segmentation [RFB15, BKC17, MBP\*22] are well-studied in the literature, but relatively few models have been published that are suitable for professional-quality analogue film artefact restoration. Many existing general inpainting approaches are unsuitable as they operate at much lower resolutions than required for film scans (typically at least 4K), and rely on a mask being provided. Traditional solutions employed in commercial film scanners, such as Kodak's Digital ICE [SMW85, SBB\*99, Gan08], use a separate infrared illumination process to create an inpainting mask and apply simple nearest neighbour inpainting. This requires specialist hardware, only works on a limited number of emulsions, and cannot detect every form of artefact. We focus instead on a purely image-based automated film restoration process which works for all emulsion and damage types without additional imaging.



**Figure 1:** Input and ground truth from our authentic artefact damage dataset, along with restorations from some of the models we evaluated, presented at full resolution. Ours is the first public dataset allowing accurate quantitative evaluation of film damage restoration.

Large datasets with realistic film damage to train and evaluate on are not easily available [CSJHR06] and most ML-based film restoration instead relies on naïvely generated synthetic images to approximate analogue artefacts [ISS19, Mir20, WZC\*20, ISW22]; systems with for which authentic damaged images *have* been collected have typically not made their data publicly available [Mir20, WZC\*20]. Unlike tasks such as JPEG artefact removal, super-resolution or colourisation, there are no robust simulation models of analogue film damage to generate high-quality synthetic training samples. There is also no consensus as to what makes a “good” restoration, and therefore how a film restoration model should be evaluated, especially without human-restored ground-truth scans [CSJH05, Cha19].

### 1.1. Contributions

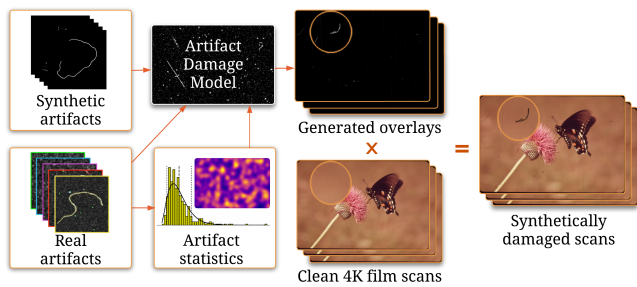
We make three contributions that address these issues:

- 1. We curate and release a dataset of real analogue film scans (Section 3);** the dataset consists of 4K scans of a variety of positive and negative colour emulsions, with damage artefacts such as dust, scratches, dirt and hairs present, paired with corresponding professionally manually restored versions. This is the first public dataset enabling quantitative evaluation of restoration models.

- 2. We construct a detailed statistical model of analogue film**

**damage, allowing us to create highly-realistic synthetic data for training (Section 4).** To estimate the parameters of this statistical model, we collect another dataset of scanned empty film emulsion which has been scratched and left to get dusty; each damage artefact in the scans is manually annotated, resulting in a set of over 12,000 unique artefacts. We extract statistics governing the shape, size and spatial density of different types of artefacts, which allows us to generate realistic synthetic analogue damage overlays to train restoration models. We validate the resulting statistical model through a human perceptual study. In addition, we release a dataset of analogue image scans paired with their synthetically damaged versions, which we use to fine-tune or re-train models.

- 3. We use these datasets to perform a detailed empirical comparison of seven state-of-the-art models on film artefact restoration (Section 5).** We include specialised film artefact restoration models, and general models that can be adapted for this purpose. We perform evaluation on our dataset of images with real analogue damage and manually restored ground truths. We address the low-resolution limitations of existing methods that make them unsuitable for professional-grade restoration and extend processing to analogue film scans at 4K resolution by applying them patch-wise. Our synthetically damaged dataset allows us to train a segmentation network to supply damage masks to inpainting models that explicitly require them.



**Figure 2:** Overview of our proposed approach to generating realistic damage overlays showing analogue film artefacts (Section 4). We extract statistics from 12 000 manually-annotated artefacts, and use these to build a probabilistic model of artefact damage that can be applied to clean film scans. The resulting damage overlays are highly realistic, and suitable for training image restoration models.

## 2. Background and related work

In this section we discuss definitions of film damage in the literature, and review existing machine learning approaches to film artefact damage restoration. Additionally, we consider state-of-the-art machine learning solutions to similar tasks which can be adapted to the film artefact restoration problems, and the challenges involved in doing so.

### 2.1. Film damage

Analogue film damage is a loose term covering a variety of degradation types. Chambah [Cha19] roughly categorises analogue damage into two groups: chemical degradations and mechanical degradations. Chemical degradations, such as colour dye fading, contrast saturation and vinegar syndrome, are usually spatially homogenous [Cha19, WZC\*20]; grain is also spatially homogenous, but while some works consider it to be another type of degradation [WZC\*20], others point to it as an example of the ill-defined difference between film damage and artistically valued properties of the medium [Cha19, ISW22]. On the other hand, mechanical degradations, for example abrasions to the emulsion (scratches) or dust specks or hairs stuck to the emulsion, are not spatially uniform [Cha19, WZC\*20].

As film is the source record for printing images, artifacts which affect the photochemical emulsion will be carried over to any prints made from it [Cha19]. Therefore, damaged film scans and the resulting damaged reproductions can be restored via the same techniques. However, further damage inflicted on the print itself, such as tears and folds, is independent of the emulsion. While some works consider film and print degradations interchangeable [WZC\*20], it is notable that they have distinct properties in terms of scale and spatial distributions with respect to the medium in which they originally occur. Furthermore, print degradations are not unique to prints of images captured on film, and as such are beyond the scope of this work.

Removal of any type of degradation can be reframed as an image-to-image translation problem, where some degradation operation has been applied to the original, non-damaged image, and

needs to be inverted, similar to denoising [LMH\*18, ZAK\*21, ZAK\*22, LCS\*21, CCZS22], superresolution [JAFF16, WYW\*18, LCS\*21], colourisation [ZIE16, IZZE17, KWK21, SCC\*22]. For these tasks, the conventional approach is to generate training data from clean images by applying a transform which is identical to the degradation process, e.g. with colourisation, grayscale versions of the training data images can be easily derived from color images. Yu et al. [YDLL18] propose a reinforcement learning approach to modelling more complex degradations as a dynamic mix of simpler types of damage, however, it is important to note that the most common degradations in the image denoising, restoration and inpainting tasks addressed in the literature happen entirely in the digital domain. Film damage, and especially mechanical film damage, in contrast, is a product of the physical properties of the film emulsion [Cha19], and is translated into the digital domain during the film scanning process. Due to their analogue nature, lack of spatial uniformity, variability in opacity, and randomness of shape and size, analogue artefacts such as scratches, dust specks and hairs are more challenging to model, and by extension, digitally simulate and restore [Cha19, WZC\*20, Mir20, ISW22].

Film damage has been simulated naïvely by directly compositing a small set of full frame damage textures over clean images [ISS19, WZC\*20, Mir20, ISW22]. Although transforms such as resizing, rotating, flipping and randomly cropping the damaged textures are employed in order to introduce more diversity in the synthetic data, to our knowledge, no approach in the literature varies the shape, location, and/or rotation of individual artefacts on the emulsion, nor their size with respect to the contents of the image and the film frame.

### 2.2. Inpainting

Inpainting is a task in which missing regions in the image, indicated by a binary mask, are filled in by estimation based on neighbouring pixel and global image context. The task lends itself to being reframed as part of the analogue artefact restoration problem: image areas obscured by the artefacts need to be suitably inpainted. However, there are some challenges: artefacts have arbitrary shape, size and location, which, in addition, are not known a priori. Recent advances in deep learning have enabled inpainting of larger image regions by generating semantically consistent content [SLM\*22]. The task has been tackled by various families of models, such as VAEs [HRCE18, PLXL21], GANs [YLY\*18], and diffusion models [SCC\*22, RBL\*21, LDR\*22]. While earlier approaches are constrained in the mask shape applied for inpainting, some recent models consider arbitrary masks. Another related task is single-image de-raining, which has been tackled by CNN-based approaches [RZH\*19], and more recently, by transformer-based ones, such as Restormer [ZAK\*22], which is currently state-of-the-art. Blind image inpainting is an extension of the inpainting task, where the mask indicating areas to be inpainted is not provided [WCTJ20, CSL\*17]. Hertz et al. [HFH\*19] propose a deep learning approach which attempts to address this problem in the context of watermark removal, by also predicting a mask to separate out the areas to be inpainted from the areas which are to be preserved. In line with this method, state-of-the-art inpainting approaches can be adapted to blind image inpainting via an additional

segmentation network tasked with predicting masks for the inpainting network to use.

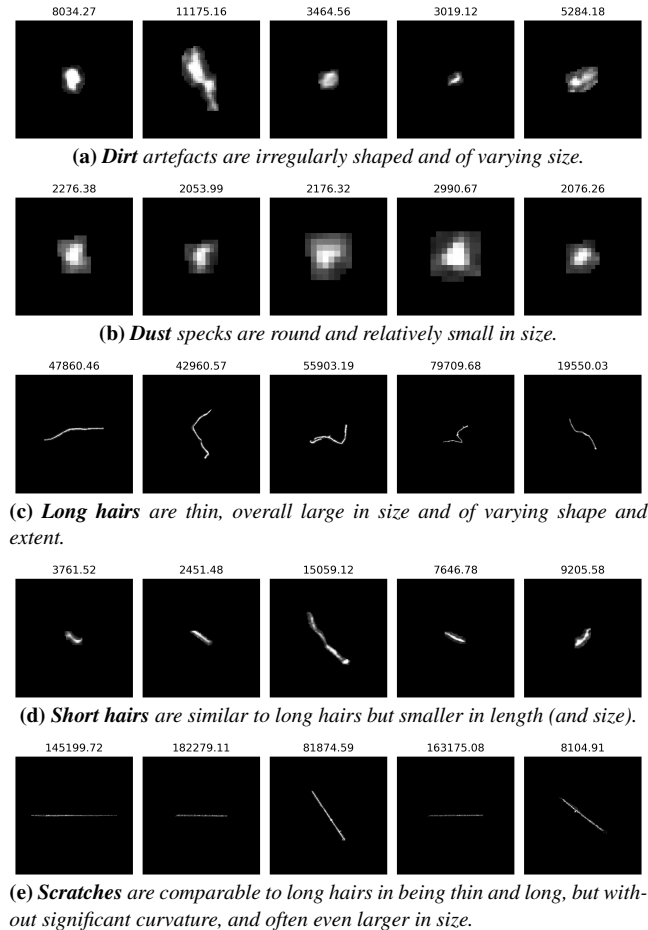
### 2.3. Processing of high-resolution data

Modern camera-equipped devices are able to capture images at very high resolution (i.e. mega- and even gigapixels). The exact definition of “high resolution” varies between applications [BZI22]; in film scanning, it is common to scan images at 3000DPI and at least 4K resolution. Many of the image-based tasks that have drawn attention in the machine learning community in the last few years do not require for the image input to be processed at full resolution – e.g. classification models are conventionally trained on cropped and downsampled images of size  $224 \times 224$  pixels to meet computational limitations. Talebi & Milanfar [TM21] even investigate learning the resizing operator on the input to improve network performance while maintaining the constraint for lower resolution image input. On the contrary, few state-of-the-art image restoration approaches, e.g. Restormer [ZAK\*22] and LaMa [SLM\*22] claim the ability to process high resolution image data. When applying machine learning approaches to problems in fields such as medical imaging, processing large medical image scans at their original resolutions is crucial. The standard approach is to split the images into patches, have the model process them, and stitch them back together [PW20]. Processing film scans poses a very similar problem.

### 2.4. Deep learning for analogue film restoration

Due to the challenges in modelling film damage and collecting relevant data, there are few approaches in the literature which set out to solve the specific problem of film artefact restoration. Strubel et al. [SMF19] train a SegNet model to remove dust and scratches, and provide a limited dataset of grayscale image scans for training and evaluation. Mironica [Mir20] proposes a GAN-based approach to film artefact restoration; to generate training data, they use a set of 100 synthetic overlays applied over a set of 2500 clean film scans, which are randomly cropped to patches of  $128 \times 128$  pixels. The qualitative results in the paper demonstrate some success in inpainting smaller artefacts, at the expense of overly smoothing grain. Similarly, Wan et al. [WZC\*20] address the problem of limited training data by applying synthetic damage overlays onto digital images, along with synthetic grain to mimic film scans and prints; they train a segmentation U-Net to predict damage masks, and further jointly train two VAEs to translate images between two latent spaces, corresponding to the domains of damaged and restored photos, respectively. The model is trained on cropped patches of size  $256 \times 256$  pixels. Since the training data is derived from digital images, the restorations produced by this approach have the effect of overly smoothing grain and shifting color in actual film scans from the test set, i.e. the network itself introduces loss of information to the input, including novel artefacts. A follow-up work by the same authors extends the approach to analogue video sequences [WZCL22]. Similarly, DeepRemaster [ISS19] also applies pre-rendered damage overlays to sequence frames to model analogue damage. Finally, Ivanova et al. [ISW22] propose a U-Net restoration network trained with a perceptual loss, which tackles the artefact detection and restoration tasks simultaneously. As with

previous works, film damage is again simulated by applying a set of overlays, modified by various simple transforms, over clean analogue film scans. While this restoration network is qualitatively and quantitatively shown to outperform the method of Wan et al., it is limited in the resolution of inputs which it can be applied to, due to being trained on images downsampled and cropped to  $256 \times 256$  pixels.

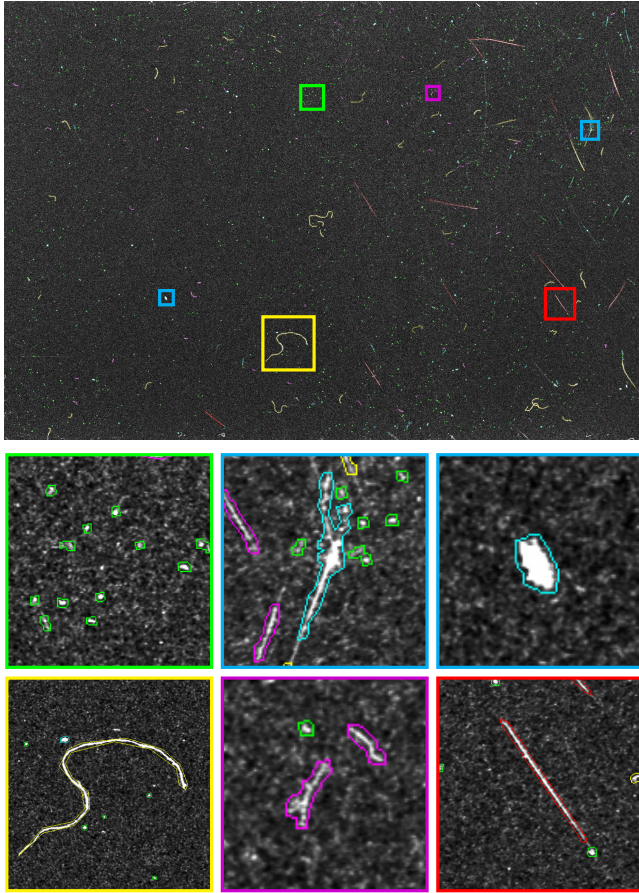


**Figure 3:** Examples of extracted and padded artefacts. The number above each is its area in  $\mu\text{m}^2$  (square microns).

### 3. Evaluation data for damage restoration

We curate a dataset of 35mm film scans at 4K resolution, with varying degree of authentic artefact damage in the form of dust, scratches, hairs and dirt. The dataset includes 44 images of various film emulsions, both positive (slide) and negative. The content of the images is also diverse, including landscapes, architecture shots and still lifes. The images have been shot, developed and scanned by Dmitri Tcherbadji of Analog.Cafe, and are used with his kind permission. In addition, each damaged image in the set has been paired with a ground-truth restoration via manual inpainting of the artefacts in Photoshop by the same expert. Original and manually-restored versions of an example photograph from the dataset are shown in Figure 1. To our knowledge, this is the only public dataset of high quality damaged film scans paired with ex-

pert restoration ground truths. We have made the dataset available at [www.doi.org/10.6084/m9.figshare.21803304](http://www.doi.org/10.6084/m9.figshare.21803304).



**Figure 4:** An example of an annotated scan. Full resolution (4944 by 3396 pixels) is shown at the top, along with zoomed-in regions around examples of each type of artefact below. Overlaid annotations for dust (green), dirt (blue), long hairs (yellow), short hairs (magenta), and scratches (red).

#### 4. Modelling analogue film damage

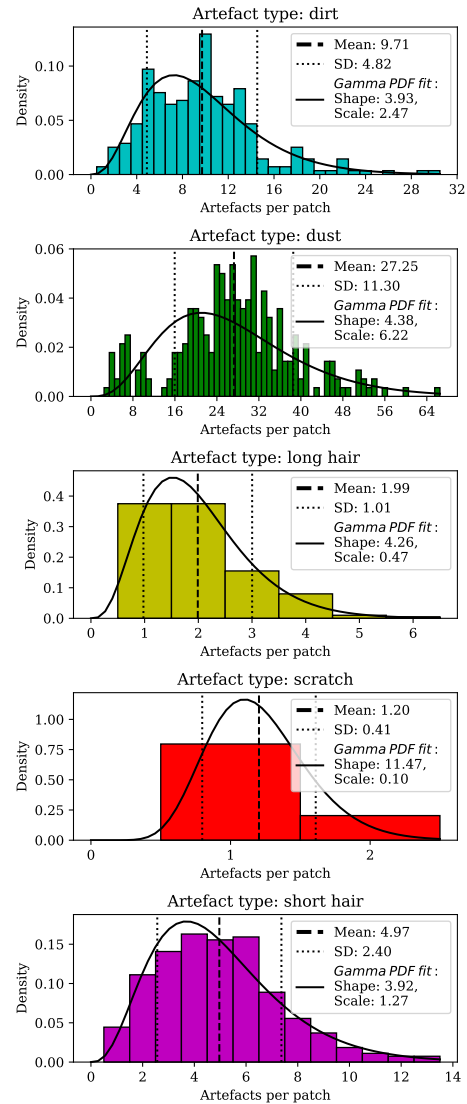
In this section, we describe our novel approach to generating synthetic analogue damage. Our overall approach is as follows:

1. Annotate, classify and extract over 12 000 individual real analogue film artefacts from heavily-damaged high resolution scans (Section 4.1).
2. Calculate statistics of the extracted artefacts, such as size, count, and spatial density (Section 4.2).
3. Build a probabilistic model to generate new artefact damage overlays using a combination of the extracted artefacts and synthetic ones, parameterised by the recorded statistics (Section 4.4).

We demonstrate the realism of damage overlays generated by our model, via a perceptual user study (Section 4.5). We further validate the damage by evaluating it in the context of artefact segmentation (Section 5.1) and artefact restoration. In both

cases, the models trained with data synthesised via our approach outperform the alternatives. The full damage synthesis pipeline, including the annotated analogue damage scans, is available at [www.github.com/daniela997/FilmDamageSimulator](http://www.github.com/daniela997/FilmDamageSimulator).

**Artefact types.** Prior works have classified damage as dust, scratches or hairs [Cha19, WZC\*20]. We extend this taxonomy by splitting hairs into two classes, short and long, due to their large variation in size; we also define an additional class, dirt, to capture artefacts of irregular shape which are larger than dust specks, but are neither hairs nor scratches.



**Figure 5:** Number of artefacts per  $256 \times 256$  pixel patch for each artefact type across all scans, and fitted Gamma distributions.

#### 4.1. Capturing real analogue film artefacts

To gather samples of real-life dust specks, scratches, hairs and dirt, we scan 10 heavily damaged, empty 35mm film frames of Lo-

mochrome Color Negative 400 ISO film using a Plustek OpticFilm 8100 dedicated film scanner and SilverFast 9. The resulting scans are saved at 4K resolution.

We manually annotate individual artefacts in the scans with bounding polygons, and classify each as dirt, dust, long hair, short hair or scratch (see Figure 4 for examples). We calculate the area of each polygon, and convert these to physical units of square microns based on the ratio between the scanned frame's size in pixels, and its size in millimeters – 35mm on the long edge and 24mm on the short edge. Finally, we extract each artefact, zero padded to square, to create a bank of isolated artefacts to sample from when generating new overlays; examples for each class are visualised in Figure 3. In total we have annotated 12135 artefacts across the 10 scanned frames.

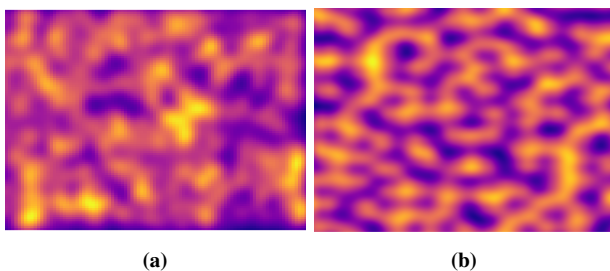
#### 4.2. Analysing artefact statistics

For each artefact class, we collect several statistics to ensure our generated overlays match the distribution of real film artefact damage. We measure individual artefacts (area in  $\mu\text{m}^2$ , square microns), as well as their distribution over the entire frame (counts, spatial frequency). Artefact counts and sizes are summarised in Table 1; Figure 7 displays the full distributions of sizes for each class.

Artefact type	Count	Avg. area ( $\mu\text{m}^2$ )	Std. dev. area ( $\mu\text{m}^2$ )
dirt	2700	8194	10304
dust	7631	3344	1637
long hair	398	53501	35524
short hair	1341	16365	11735
scratch	65	229660	521345

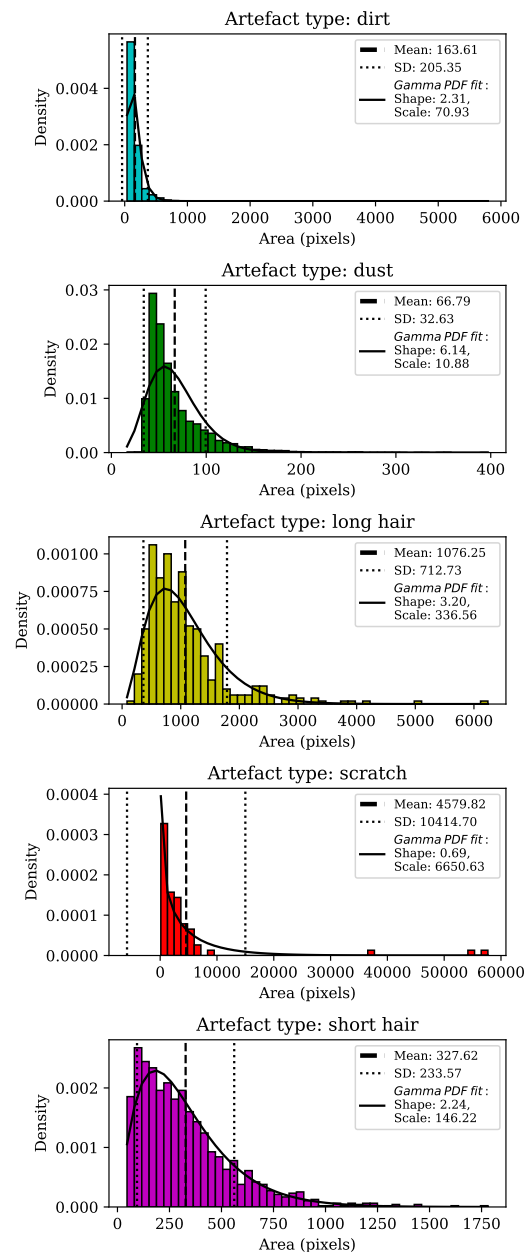
**Table 1:** Statistics of annotated artefacts – we report the total count and average area for each artefact class. See Figure 7 for the full distributions of areas.

**Counts.** For all 10 scans, we observe strong class imbalance in favor of dust and dirt, with scratches being very scarce. Moreover, the artefacts' spatial distribution is not uniform; we therefore split each scan into  $256 \times 256$  pixel patches (padding as required) and record the artefact counts for each class in each patch. The resulting distributions are shown in Figure 5.



**Figure 6:** Visual comparison between the true spatial distribution of short hairs (by kernel density estimation on annotation centroids) (a), and Perlin noise (b), which we use as an approximation when sampling new overlays. Each plot shows the distribution over a full 35mm frame, with brighter colors corresponding to higher density.

**Spatial frequencies.** We visualise the spatial distribution of artefacts via kernel density plots, for an example class in Figure 6a and for all classes in the supplementary material (Figure S5). While there is no obvious pattern to the distributions of artefact occurrences, we can observe that four artefact types are more frequent near the frame's upper and lower left corners; this could be related to the direction in which the film strip is advanced inside the camera and/or the scanner.



**Figure 7:** Artefact size histograms in pixels, and fitted Gamma distributions. While average sizes for each type of artefact vary, the sizes of artefacts are distributed in a similar way.

Familiarity	Participants	Resolutions											
		128 × 128		256 × 256		512 × 512		1024 × 1024		2048 × 2048		Overall	
		Mean	Std	Mean	Std	Mean	Std	Mean	Std	Mean	Std	Mean	Std
<b>Not Familiar</b>	49	52.4	12.7	52.6	11.1	50.6	9.6	52.5	12.4	50.3	13.2	51.5	6.9
<b>Somewhat Familiar</b>	81	52.9	10.1	54.1	12.0	51.7	10.0	54.4	13.6	51.1	15.9	52.9	7.0
<b>Very Familiar</b>	121	54.9	11.1	53.5	11.3	50.4	12.2	49.9	12.1	48.1	16.1	51.4	7.0
<b>Overall</b>	251	53.8	11.2	53.5	11.5	50.8	11.0	51.9	12.7	49.6	15.5	51.9	7.0

**Table 2:** Summary of participant scores (in %) based on responses to Survey 1. A score of 50% is optimal since it indicates that our synthetic damage is on average indistinguishable from real damage.

### 4.3. Synthetic artefacts.

In addition to the artefacts extracted from the 10 scans, we also use a set of 6100 synthetic dust, scratch, hair, lint, and dirt artefacts, which were manually drawn in Photoshop and kindly shared by Stefan Ringelschwandtner of Mononodes. We do this to further increase the diversity of rarer artefact classes. Every artefact in this set is of size  $400 \times 400$  pixels, and therefore must be rescaled to match the observed distribution of areas for real artefacts of the corresponding class. As the scratch class is particularly under-represented, we also programatically generate additional scratches matching the appearance of those found in real scans.

### 4.4. Generating damage overlays

We develop a probabilistic model based on the measured artefact properties, which allows us to generate new synthetic full-frame damage overlays with a realistic distribution of artefacts. Our generative process is as follows:

- 1. Sample the numbers of artefacts for each class:** For each artefact class, we sample a target count from a Gamma distribution fit by maximum likelihood estimation to the empirical counts in Figure 5 and round to the nearest integer. As the recorded counts are per  $256 \times 256$  pixel patch, we scale the sampled counts depending on the target overlay resolution.
- 2. Sample the artefact sizes:** Similarly, we sample artefact sizes from Gamma distributions fitted to the observed artefact sizes in Figure 7. The sampled sizes are further rescaled according to the target overlay size, using the relevant pixel-to-micron conversion ratio.
- 3. Sample the artefact appearances:** For each artefact, we randomly choose its appearance from among the real and synthetic artefacts of the relevant class. We ensure that the size of each (in pixels) is not divergent from the target sizes sampled in the previous step, in order to avoid excessive upsampling of small artefacts in classes with high variance in area (e.g. dirt).
- 4. Sample locations and rotations:** We approximate the spatial densities discussed earlier with Perlin noise [Per85]; we justify this choice by a visual comparison with the true artefact density aggregated over all scans (Figure 6). We sample the location of each artefact independently from this noise distribution; we also sample a rotation from the uniform distribution on  $[-\pi, \pi]$ .
- 5. Compose the final overlay:** Given the artefact appearances, sizes, locations and rotations, we alpha-composite them into the frame at the required output resolution. The overlay can be used

to simulate damage to film negatives (artefacts are white), or to developed slides (artefacts are black).

### 4.5. Human validation of generated artefact damage

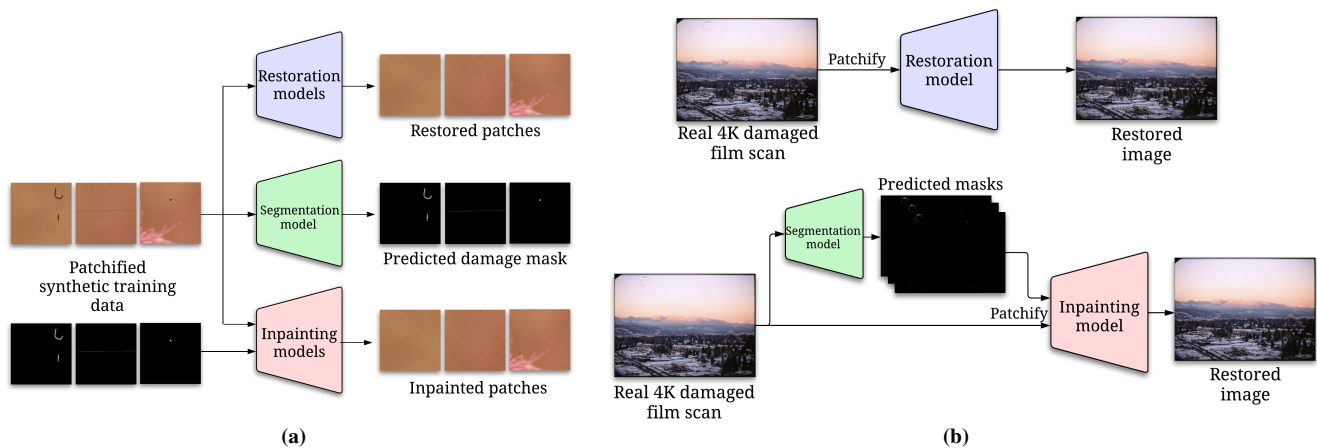
We conduct two user perceptual surveys to verify that our damage overlay generation process yields realistic results. We discuss the study design and summarise the results below. In both cases, we also ask each participant to self-report their familiarity level with analogue film and its associated artefacts, as ‘Not familiar at all’, ‘Somewhat familiar’ or ‘Very familiar’.

**Real vs. synthetic damage.** In the first survey, we compare real analogue damage to the synthetic damage produced by our model. The generated damage is applied to the ground truth restored images from our test set. We pair the synthetically damaged image scans with their real damaged versions. Each participant was shown 100 such pairs (see Figure S1 in the supplementary material for examples): 50 simulating negative film damage (i.e., white artefacts), and 50 simulating positive (slide) film damage (i.e. black artefacts), and is asked to choose one image per pair which they believe shows real analogue film damage. The images are cropped to five different target resolutions to reflect the relationship between artefacts and image features of varying size.

We collected 251 responses to this first survey. We calculate the percentage of pairs for which each participant chose the example with real artefacts.

We define the ‘score’ of a user as being the overall fraction of pairs for which they correctly selected the real damage. A score of 50% is ideal, as it would mean that the user is unable to distinguish real from synthetic damage better than chance. Participant score breakdown is summarised in Table 2. We find that unfamiliar participants scored on average 51.51% with standard deviation of 6.88%, moderately familiar participants scored on average 51.93% with standard deviation of 7.05%, and very familiar participants scored on average 51.43% with standard deviation of 7.0%. Thus, all groups of users found our synthetically damaged images to be indistinguishable from original damaged scans, and there was no statistically-significant variation among the groups.

**Ours vs. Ivanova et al. [ISW22].** In the second survey, participants are shown 30 pairs of film scans from the Documerica photographic collection [The77]. In each pair, one version of the image is damaged using the earlier approach of Ivanova et al. [ISW22], whereas the other image is damaged using our proposed approach.



**Figure 8:** Overview of (a) training and (b) evaluation workflows for restoration, segmentation and inpainting models. We retrain existing restoration models on our synthetic data where relevant, and train an additional segmentation model to adapt state-of-the-art inpainting methods to the film artefact restoration task.

Participants were asked which image in the pair showed more realistic damage.

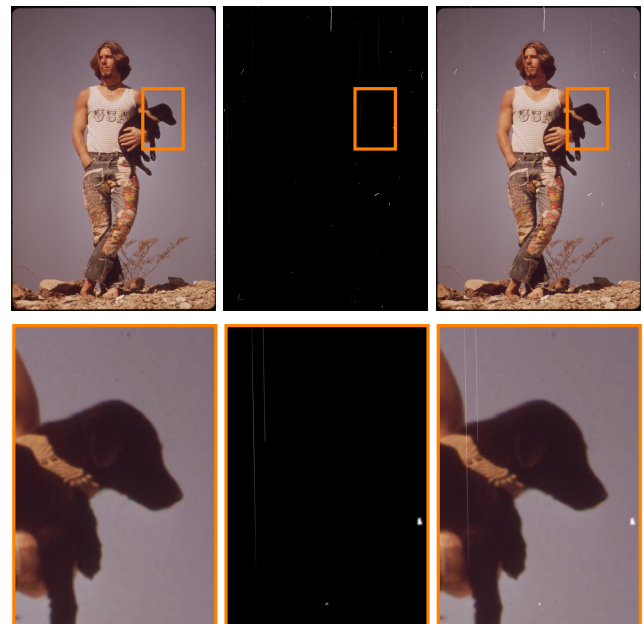
We collected 78 responses for the second survey. In this survey, unfamiliar participants preferred our damage on average in 63.88% of the examples with standard deviation of 23.17%, moderately familiar participants preferred our damage in 73.33% of the examples with standard deviation of 21.13%, and very familiar participants preferred our damage in 75.58% of the examples with standard deviation of 21.71%. These results indicate that all groups found our synthetic damage to be more realistic than that proposed in the prior work [ISW22]; expert participants favored our damage the most out of the three groups.

## 5. Experiments

Equipped with our damage simulator (Section 4) and our expertly curated test set of real analogue damage with hand-restored ground-truths (Section 3), we perform two sets of experiments. In Section 5.1, we directly compare our damage synthesis pipeline with those from two prior works [ISS19, WZC\*20]. In Section 5.2, we evaluate several approaches on our final goal of film restoration. The following sections discuss the chosen approaches (and how we adapt them to high resolution data), describe our experimental setup, and report the results of our experiments.

**Evaluation metrics for restoration.** Since our evaluation set has ground-truth restored images (Section 3), we can evaluate restoration quality directly using standard image similarity metrics. Specifically, we use peak signal-to-noise ratio (PSNR), structural similarity (SSIM) [WBSS04], and the learnt perceptual metric LPIPS [ZIE\*18]. These are calculated over full frames (not individual patches), and we report the average over the evaluation set.

**Evaluation metrics for segmentation.** We obtain approximate ground-truth segmentations of the authentically damaged dataset by subtracting the damaged images from the restored ones and binarising. We use the standard image segmentation metrics intersection-over-union (IoU) and F1 score. Since these metrics are sensitive to exact pixel-value overlap, which in turn is influenced by

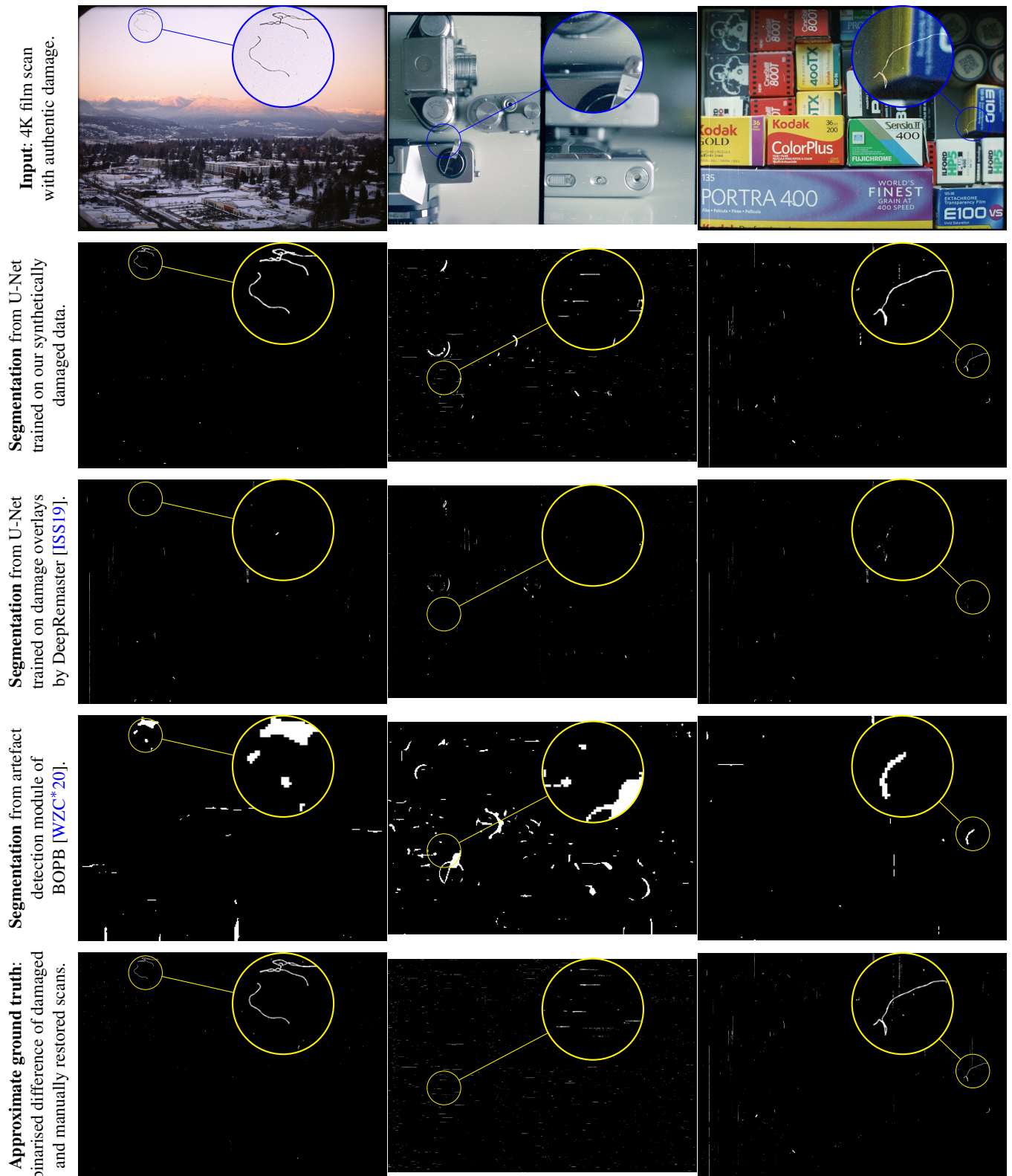


**Figure 9:** High resolution image, generated damage overlay and resulting synthetically damaged image.

the choice of binarisation threshold, we test several thresholds and choose the one which maximises the scores obtained by the baselines. We also borrow a popular point-cloud comparison metric, the earth-mover's distance (EMD) [ADMG18, FSG17], to provide an additional measure which is less sensitive to exact overlap, but more accurately compares proximity of long, narrow features such as hairs and scratches.

**Synthetically damaged training data.** We generate synthetic damage overlays using our proposed approach (Section 4) for 6228 clean 4K image scans from the Documerica collection [The77] (see Figure 9 for an example). We render the synthetic damage either





**Figure 10:** Qualitative comparison of segmentations of artefacts from our authentic damage dataset.

as negative or positive artefacts (corresponding to damage on film negatives or developed slides respectively), with equal probability for each image. The overlays are also binarised to produce segmentation maps indicating where damage was added. We split the set of 6228 images into training and validation sets in the ratio of 9:1. For the segmentation experiment, we use a variant of this dataset produced by damaging the same set of DocuMera images with damage overlays as provided by DeepRemaster [ISS19]. These datasets are used to retrain restoration models where relevant, as well as to train a segmentation model to detect artefact damage in our real artefact damage evaluation set. The data is available at [www.doi.org/10.6084/m9.figshare.21815844](http://www.doi.org/10.6084/m9.figshare.21815844).

### 5.1. Comparison against existing damage simulations

We train two U-Net [RFB15] segmentation networks to detect artefact damage in real damaged film scans – one using damage generated by our model, and one using the damage provided by DeepRemaster [ISS19]. We train both for 20 epochs on  $256 \times 256$  crops of the corresponding damaged training set with learning rate  $10^{-3}$  and a batch size of 16. We compare the predictions from the two segmentation networks directly to those from the artefact segmentation module of BOPB [WZC\*20] as our second baseline.

Segmentation model training	IoU $\uparrow$	F1 $\uparrow$	EMD $\downarrow$
<i>Our damage</i>	0.256	0.408	0.007
<i>DeepRemaster [ISS19] damage</i>	0.003	0.006	0.016
<i>BOPB [WZC*20] pre-trained module</i>	0.010	0.019	0.016

**Table 3:** Comparison between the predicted segmentations for the evaluated models trained on different simulated damage.

**Results.** Segmentations obtained from the U-Net trained on our synthetically damaged data are of much higher quality than those obtained from the same model trained on the damage from DeepRemaster, as well as those obtained from the segmentation module of BOPB [WZC\*20], as shown in Figure 10. Quantitatively, the model trained on our artefacts again outperforms BOPB and DeepRemaster, achieving lower EMD and higher IoU and F1 scores, summarised in Table 3.

### 5.2. Comparison of restoration models

We compare a diverse set of approaches to artefact restoration. First, we select three methods that directly perform damage restoration:

- *Bringing Old Photos Back to Life* (BOPB) [WZC\*20], an approach specifically targeting analogue damage, for which the authors have provided pre-trained weights. We compare variants using our segmentation model vs. theirs, and with fully-convolutional vs. patch-wise processing.
- A restoration U-Net trained with perceptual loss [ISW22], another approach which specialises in film artefact removal, for which the pre-trained weights are available. We also test a variant re-trained on crops of our synthetic 4K dataset, instead of the  $256 \times 256$  downsampled images used in the original. Lastly, we test a variant re-trained on the same crops of our synthetic dataset, damaged with overlays from DeepRemaster [ISS19].

- Adobe Photoshop’s *Dust & Scratch filter*, which is a commonly-used off-the-shelf solution for artefact restoration.

Second, we select three state-of-the-art methods for image inpainting. In order to apply these to the restoration task, we also train a model for segmenting artefacts (see below), which is used to determine which pixels must be inpainted:

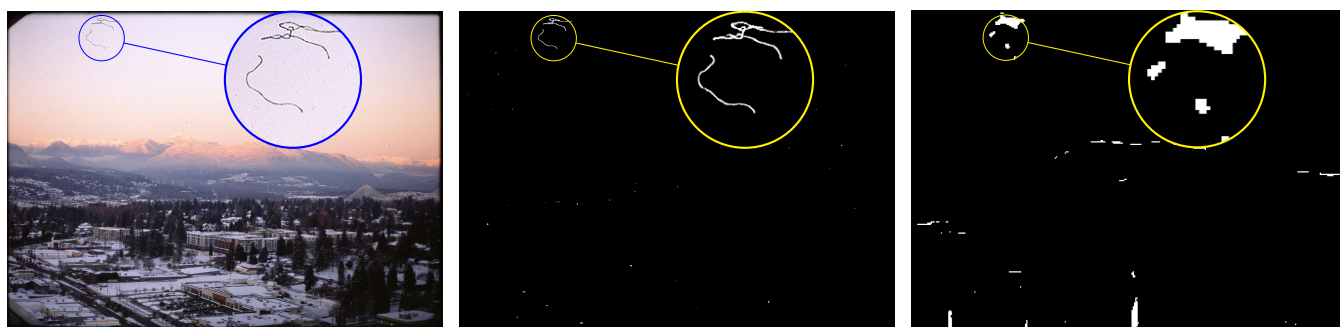
- *LaMa* [SLM\*22], a state-of-the-art approach for high resolution irregular hole inpainting, with a pre-trained model available.
- *Stable Diffusion* [RBL\*21], popular for its state-of-the-art text-to-image generation abilities, also supports image inpainting.
- *RePaint* [LDR\*22], another diffusion approach focusing on inpainting, proposing a resampling technique which conditions the inpainting on a pre-trained diffusion model.

Finally, we consider two blind inpainting approaches:

- *Blind Visual Motif Removal From a Single Image* (BVMR) [HFH\*19], which simultaneously detects the pixels representing the visual motif to be inpainted, and synthesises new content for the affected pixels. The provided pre-trained model is for semi-transparent emoji watermark removal, which we re-train for the artefact restoration task.
- *Restormer* [ZAK\*22], an approach which achieves state-of-the-art results in the tasks of blind denoising and deraining, and is applicable to large resolution inputs. Out of several models provided, we find that the closest to the task of film artefact removal is the one trained for deraining.

**Pre-segmentation of artefacts for inpainting methods.** LaMa, RePaint and Stable Diffusion rely on masks to be supplied for the inpainting task. As they cannot identify artefacts automatically, they cannot be applied to the film restoration task out-of-the-box. To remedy this, we employ the segmentation U-Net trained on our synthetically damaged data (described in Section 5.1) to predict damage masks on the test set of authentically damaged data. When evaluating LaMa, Stable Diffusion, and RePaint on damage restoration, we first pass the images through the trained segmentation U-Net, then inpaint the regions it indicates are damaged (i.e. we use the output of the segmentation model as the mask input for the inpainting model). In addition, while Wan et al.’s [WZC\*20] method predicts its own damage masks, we also evaluate a variant using the masks from our segmentation model (shown in Section 5.1 to produce artefact segmentations of superior quality).

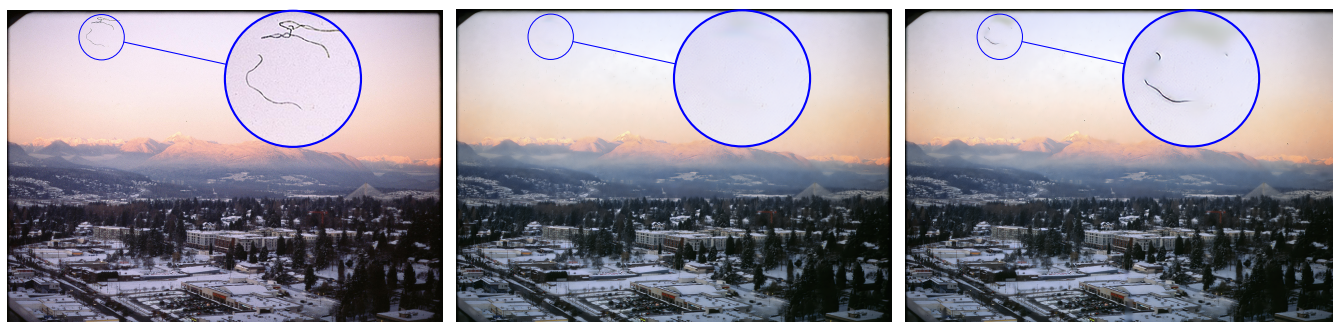
**Processing high resolution scans.** Most of the models are designed to operate on  $256 \times 256$  pixel inputs. We adapt these to process the 4K image scans in our test set without downsampling them first. During inference, we process the images in our dataset in patches of size  $256 \times 256$ , and we stitch back the predictions using 50% overlap and a smooth blending function, as proposed by Pielawski & Wählby [PW20]. For the inpainting models which require masks, we increase efficiency by only processing the patches for which the corresponding mask indicates the presence of artefacts; if no artefacts have been detected in the patch, we simply copy the input patch to the output tensor to be stitched with the rest of the restored image. This does not apply to the Photoshop Dust & Scratch filter, nor to LaMa, both of which can natively process 4K images.



(a) *Input: 4K film scan with authentic damage*

(b) *Artefact Segmentation: prediction from U-Net trained on synthetically damaged data.*

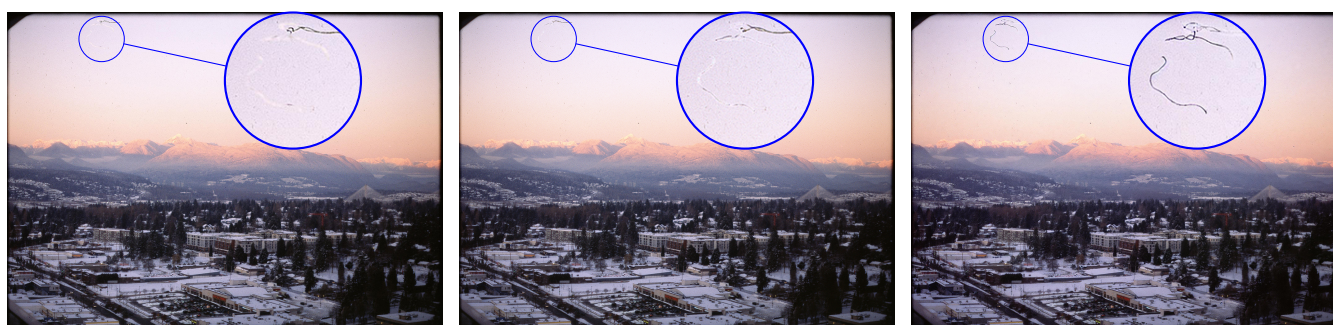
(c) *Segmentation by BOPB [WZC\*20].*



(d) *Restoration by U-Net + perceptual loss [ISW22]: using originally provided model weights.*

(e) *Restoration by BOPB [WZC\*20]: using our segmentation.*

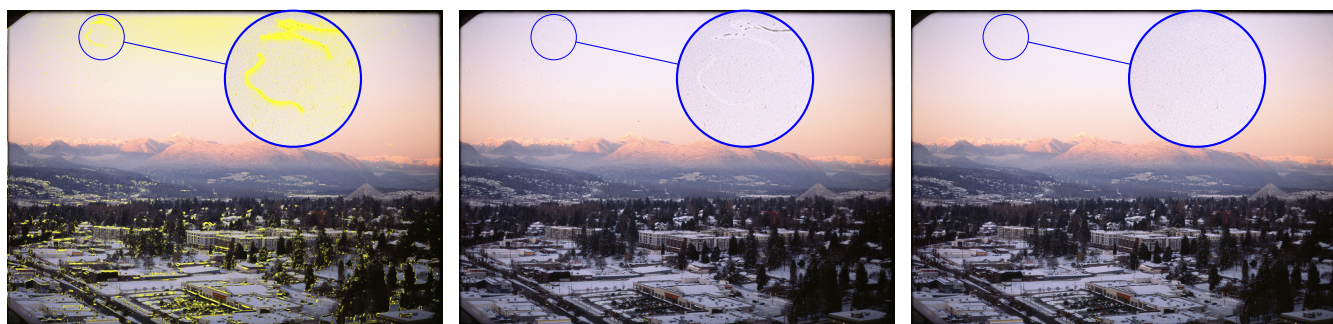
(f) *Restoration by BOPB [WZC\*20]: using their segmentation.*



(g) *Restoration by U-Net + perceptual loss [ISW22]: retrained on our synthetic damage.*

(h) *Restoration by LaMa [SLM\*22]: best performing model, using our segmentation.*

(i) *Restoration by Stable Diffusion [RBL\*21]: using our segmentation.*

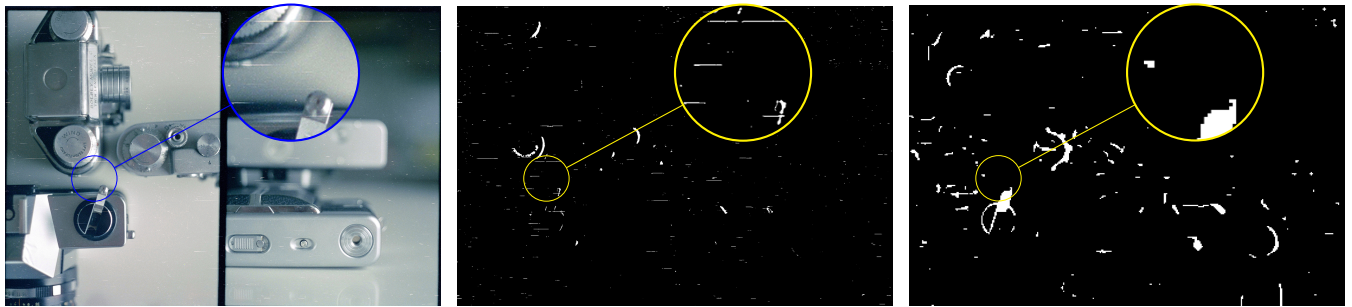
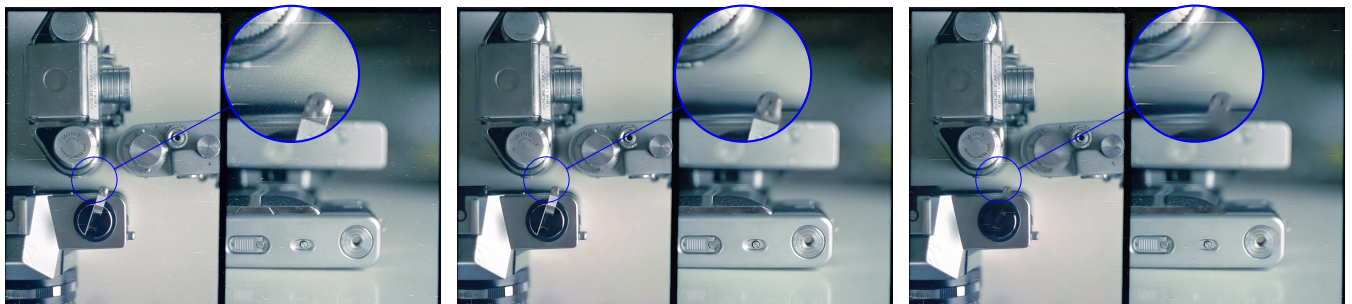
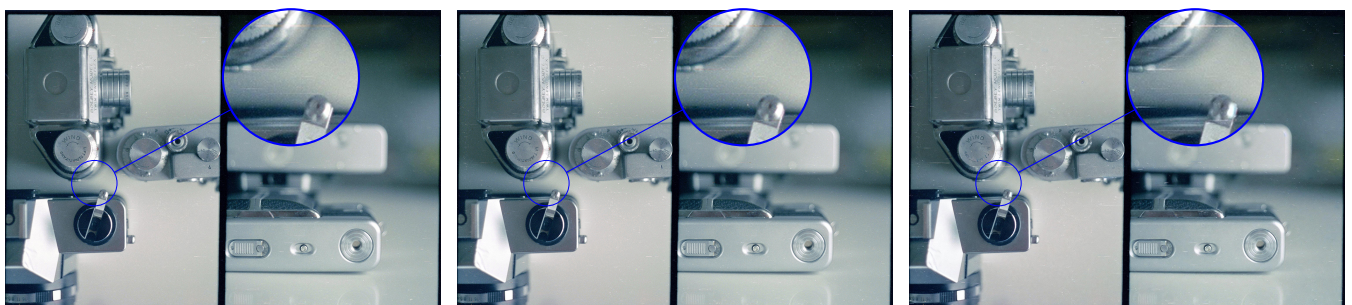
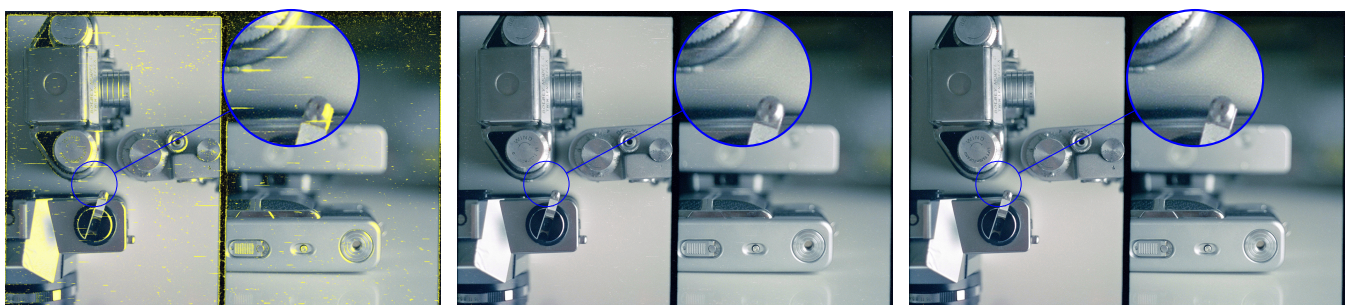


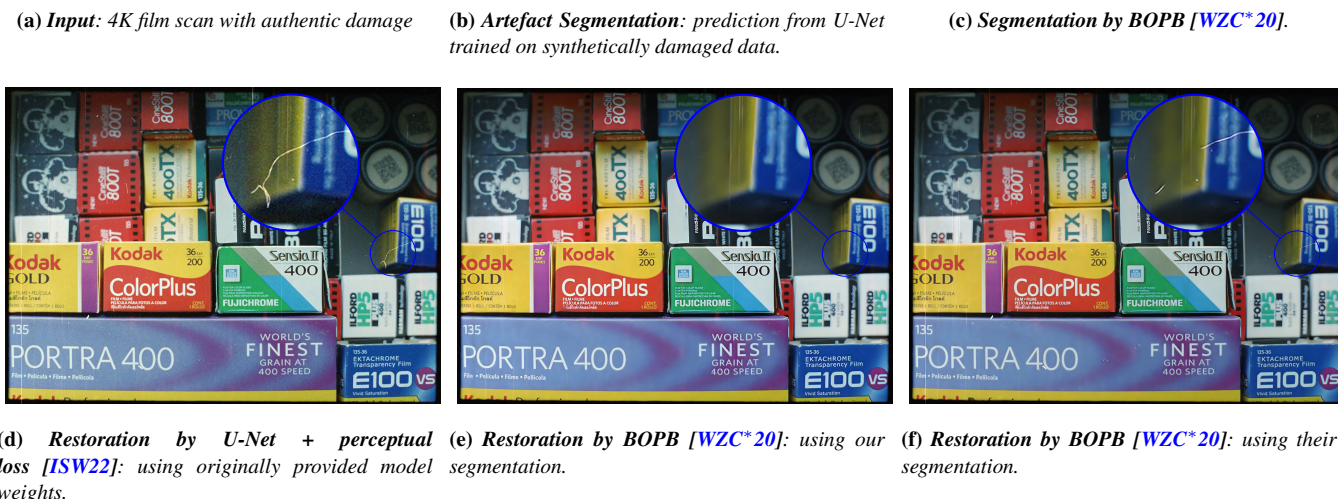
(j) *Restoration by BVMR [ISW22]: retrained on our synthetic damage.*

(k) *Restoration by RePaint [LDR\*22]: using our segmentation.*

(l) *Ground Truth: manually restored by human expert.*

**Figure 11:** *Input and ground truth from our authentic artefact damage dataset, along with chosen restorations.*

(a) *Input: 4K film scan with authentic damage*(b) *Artefact Segmentation: prediction from U-Net trained on synthetically damaged data.*(c) *Segmentation by BOPB [WZC\*20].*(d) *Restoration by U-Net + perceptual loss [ISW22]: using originally provided model weights.*(e) *Restoration by BOPB [WZC\*20]: using our segmentation.*(f) *Restoration by BOPB [WZC\*20]: using their segmentation.*(g) *Restoration by U-Net + perceptual loss [ISW22]: retrained on our synthetic damage.*(h) *Restoration by LaMa [SLM\*22]: best performing model, using our segmentation.*(i) *Restoration by Stable Diffusion [RBL\*21]: using our segmentation.*(j) *Restoration by BVMR [ISW22]: retrained on our synthetic damage.*(k) *Restoration by RePaint [LDR\*22]: using our segmentation.*(l) *Ground Truth: manually restored by human expert.***Figure 12:** *Input and ground truth from our authentic artefact damage dataset, along with chosen restorations.*



**Figure 13:** Input and ground truth from our authentic artefact damage dataset, along with chosen restorations.

Method	Notes	Patch-wise	PSNR (dB) ↑		SSIM ↑		LPIPS ↓	
			Mean	Std	Mean	Std	Mean	Std
BOPB [WZC*20]		Yes	14.8	3.14	0.756	0.063	0.561	0.060
BOPB [WZC*20]		No	24.0	2.80	0.904	0.048	0.432	0.062
BOPB [WZC*20]	w/ our masks	Yes	14.9	3.67	0.769	0.068	0.556	0.067
BOPB [WZC*20]	w/ our masks	No	24.4	2.34	0.917	0.043	0.425	0.425
<b>LaMa [SLM*22]</b>	<b>w/ our masks</b>	<b>No</b>	<b>39.6</b>	<b>5.24</b>	<b>0.991</b>	<b>0.012</b>	<b>0.042</b>	<b>0.038</b>
Photoshop Dust & Scratch Filter		No	28.9	3.09	0.965	0.022	0.277	0.088
U-Net + perceptual loss [ISW22]	trained on downsampled frames	Yes	31.8	2.62	0.979	0.015	0.224	0.047
U-Net + perceptual loss [ISW22]	re-trained on DeepRemaster damage [ISS19]	Yes	27.5	1.36	0.945	0.015	0.249	0.043
<b>U-Net + perceptual loss [ISW22]</b>	<b>re-trained on our synthetic damage</b>	<b>Yes</b>	<b>33.0</b>	<b>2.46</b>	<b>0.984</b>	<b>0.011</b>	<b>0.215</b>	<b>0.044</b>
Stable Diffusion [RBL*21]	w/ our masks	Yes	28.6	2.43	0.959	0.020	0.305	0.040
BVMR [HFH*19]	re-trained on artefact damage	Yes	17.8	4.08	0.798	0.108	0.392	0.153
Restormer [ZAK*22]	de-raining model	Yes	32.0	2.52	0.978	0.015	0.238	0.044

**Table 4:** Quantitative evaluation of restoration results on our high resolution dataset of authentic film damage.

### 5.3. Restoration results and discussion

Table 4 summarises our quantitative results, showing PSNR, SSIM and LPIPS metrics with respect to the ground-truth professionally-restored image set. All three metrics largely agree, with LaMa and the U-Net retrained with our damage as the best and second best performing models respectively. We can directly observe the substantial improvement that re-training the simple U-Net restoration model [ISW22] with our simulated damage brings to its restoration performance. On the other hand, re-training the same model with the damage from DeepRemaster [ISS19] results in deteriorated restoration performance.

BOPB works especially poorly in patch-wise mode as it introduces color distribution shift in each patch, leading to a very distorted result and correspondingly low scores. We attribute this to the fact that BOPB aims to remove film grain and modify color to match modern digital photographs, in addition to removing artefacts. This means it performs poorly at our specific task of damage removal (and is thus unsuitable for applications where the qualities of analogue photographs are desirable). Interestingly, the standard “Dust and Scratches filter” in Photoshop exceeds the performance of BOPB (patchwise and fully-convolutional modes), Stable Diffusion and BVMR. Restormer and the best U-Net model have comparable performance, but LaMa exceeds the performance of all of the competitive models on all three metrics, particularly as measured by LPIPS. This is in spite of LaMa being trained with inpainting masks that are significantly different in size and shape to typical analogue damage artefacts. Note that we do not include results for RePaint in Table 4. This model requires over 12 minutes for the restoration of a single  $256 \times 256$  pixel patch on a Tesla P100 GPU. Given that one 4K image from our test set is split into around 200 overlapping patches, about 40 hours are needed to process just one image. While this is clearly impractical in a real-life restoration scenario, we process a subset of 10 images from our test set for qualitative evaluation.

We present qualitative results in Figures 11, 12 and 13, along with additional examples in the supplementary material, Figures S2–4. BOPB performs significantly better when we substitute its predicted artefact segmentation masks (which are strongly

under-segmented) with ours. However, it still modifies colors and smoothes the film grain, significantly changing the character of images. Even when guided by our segmentation masks, Stable Diffusion fails to inpaint the damage. Similarly, BMVR, even when re-trained using our synthetic damage, struggles with the same issue as BOPB and injects additional damage. In general, we find that methods based on inpainting which do not modify regions away from the artefacts perform much better than BOPB, as they are less likely to introduce new damage or distortions. In line with our observations from the quantitative evaluation, retraining the U-Net model [ISW22] on our improved synthetic damage results in higher quality restorations. Visually, LaMa performs slightly better than [ISW22]; however both methods fail to inpaint part of the upper hair in Figure 1. RePaint’s performance, guided by our segmentation masks, is comparable to that of LaMa and [ISW22]. Despite producing restorations of relatively high visual quality, the model is limited in practical usefulness by the amount of time it takes to process images.

### 6. Conclusions

Automated restoration of mechanical damage in scanned film is a challenging task. Fully automated restoration would be transformative in improving the quality of imaging in the enormous archives of extant film held around the world. However, conventional inpainting and restoration processes are unsuitable for professional-quality restoration. All of the machine learning approaches we tested for film restoration perform well below the level required to be competitive with professional hand restoration. Models trained on low-resolution patches performed very poorly when applied patch-wise, but even state-of-the-art models like LaMa applied at native resolution were unable to adequately inpaint artefacts without introducing undue distortion. Many of the algorithms would also be computationally impractical for processing large image collections at high resolution.

Progress towards film restoration that can operate truly automatically at realistic scan resolutions requires better models of the film damage process. To that end, we have presented both a sophisticated statistical model for synthesising large quantities of real-

istic mechanical artefacts. Our extensive human validation of the synthetic artefact model suggests that even experts cannot reliably distinguish between our synthetic damage models and real damage at any level of zoom. Furthermore, we have demonstrated that training with damage generated by our model leads to a significant improvement in both artefact detection and end-to-end restoration tasks when tested on authentic damaged images. We have also published a high-quality baseline dataset to drive progress in restoration research.

Overall, we conclude that there remains significant work to achieve acceptable automatic restoration quality. We see the use of high-quality damage simulators, multi-scale approaches that can incorporate wide image context while operating at high resolution, and evaluation with challenging full-resolution image benchmarks as important directions to achieve these goals.

### Acknowledgements

This work was supported by the Engineering and Physical Sciences Research Council [grant number EP/R513222/1].

### References

- [ADMG18] ACHLIOPTAS P., DIAMANTI O., MITLIAGKAS I., GUIBAS L.: Learning representations and generative models for 3d point clouds. In *International Conference on Machine Learning* (2018), PMLR, pp. 40–49. 8
- [BKC17] BADRINARAYANAN V., KENDALL A., CIPOLLA R.: Segnet: A deep convolutional encoder-decoder architecture for image segmentation. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 39, 12 (2017), 2481–2495. 1
- [BZI22] BAKHTIARNIA A., ZHANG Q., IOSIFIDIS A.: Efficient high-resolution deep learning: A survey. *arXiv preprint arXiv:2207.13050* (2022). 4
- [CCZS22] CHEN L., CHU X., ZHANG X., SUN J.: Simple baselines for image restoration. *arXiv preprint arXiv:2204.04676* (2022). 3
- [Cha19] CHAMBAH M.: Digital film restoration and image quality. In *ICA-BELGIUM Colour Symposium* (Ghent, Belgium, 2019). URL: <https://hal.archives-ouvertes.fr/hal-02998573>. 2, 3, 5
- [CSJH05] CHAMBAH M., SAINT-JEAN C., HELT F.: Image quality evaluation in the field of digital film restoration. *Proceedings of SPIE - The International Society for Optical Engineering 5668* (01 2005). doi:10.1117/12.586738. 2
- [CSJHR06] CHAMBAH M., SAINT-JEAN C., HELT F., RIZZI A.: Further image quality assessment in digital film restoration. In *Image Quality and System Performance III* (2006), Cui L. C., Miyake Y., (Eds.), vol. 6059, International Society for Optics and Photonics, SPIE, p. 60590S. URL: <https://doi.org/10.1117/12.642414>, doi:10.1117/12.642414. 2
- [CSL\*17] CAI N., SU Z., LIN Z., WANG H., YANG Z., LING B. W.-K.: Blind inpainting using the fully convolutional neural network. *The Visual Computer* 33 (2017), 249–261. 3
- [FSG17] FAN H., SU H., GUIBAS L. J.: A point set generation network for 3d object reconstruction from a single image. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition* (2017), pp. 605–613. 8
- [Gan08] GANN R. G.: Image scanner and method for detecting a defect in an image to be scanned, Apr. 2008. (Google Patents Entry) <https://patents.google.com/patent/US7355159B2/>. 1
- [HFH\*19] HERTZ A., FOGEL S., HANOCKA R., GIRYES R., COHEN-OR D.: Blind visual motif removal from a single image. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition* (2019), pp. 6858–6867. 3, 10, 14
- [HRCE18] HAM C., RAJ A., CARTILLIER V., ESSA I.: Variational image inpainting. In *NeurIPS 2018 Workshops* (2018). 3
- [ISS19] IIZUKA S., SIMO-SERRA E.: DeepRemaster: Temporal Source-Reference Attention Networks for Comprehensive Video Enhancement. *ACM Transactions on Graphics (Proc. of SIGGRAPH ASIA)* 38, 6 (2019), 1–13. 2, 3, 4, 8, 9, 10, 14
- [ISW22] IVANOVA. D., SIEBERT. J., WILLIAMSON. J.: Perceptual loss based approach for analogue film restoration. In *Proceedings of the 17th International Joint Conference on Computer Vision, Imaging and Computer Graphics Theory and Applications - Volume 4: VIS-APP*, (2022), INSTICC, SciTePress, pp. 126–135. doi:10.5220/0010829300003124. 2, 3, 4, 7, 8, 10, 11, 12, 13, 14
- [IZZE17] ISOLA P., ZHU J.-Y., ZHOU T., EFROS A. A.: Image-to-image translation with conditional adversarial networks. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (2017), pp. 1125–1134. 3
- [JAFF16] JOHNSON J., ALAHI A., FEI-FEI L.: Perceptual losses for real-time style transfer and super-resolution. In *European Conference on Computer Vision* (2016), Springer, pp. 694–711. 3
- [KWK21] KUMAR M., WEISSENBORN D., KALCHBRENNER N.: Colorization transformer. *arXiv preprint arXiv:2102.04432* (2021). 3
- [LCS\*21] LIANG J., CAO J., SUN G., ZHANG K., VAN GOOL L., TIMOFTE R.: Swinir: Image restoration using swin transformer. In *Proceedings of the IEEE/CVF International Conference on Computer Vision* (2021), pp. 1833–1844. 3
- [LDR\*22] LUGMAYR A., DANELLJAN M., ROMERO A., YU F., TIMOFTE R., VAN GOOL L.: Repaint: Inpainting using denoising diffusion probabilistic models. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition* (2022), pp. 11461–11471. 3, 10, 11, 12, 13
- [LMH\*18] LEHTINEN J., MUNKBERG J., HASSELGREN J., LAINE S., KARRAS T., AITTALA M., AILA T.: Noise2Noise: Learning image restoration without clean data. In *Proceedings of the 35th International Conference on Machine Learning* (10–15 Jul 2018), Dy J., Krause A., (Eds.), vol. 80 of *Proceedings of Machine Learning Research*, PMLR, pp. 2965–2974. URL: <https://proceedings.mlr.press/v80/lehtinen18a.html>. 3
- [LRS\*18] LIU G., REDA F. A., SHIH K. J., WANG T.-C., TAO A., CATANZARO B.: Image inpainting for irregular holes using partial convolutions. In *Proceedings of the European conference on computer vision (ECCV)* (2018), pp. 85–100. 1
- [MBP\*22] MINAE S., BOYKOV Y., PORIKLI F., PLAZA A., KEHTARNAVAZ N., TERZOPOULOS D.: Image segmentation using deep learning: A survey. *IEEE Transactions on Pattern Analysis & Machine Intelligence* 44, 07 (jul 2022), 3523–3542. doi:10.1109/TPAMI.2021.3059968. 1
- [Mir20] MIRONICĂ I.: A generative adversarial approach with residual learning for dust and scratches artifacts removal. In *Proceedings of the 2nd Workshop on Structuring and Understanding of Multimedia heritAge Contents* (2020), pp. 15–22. 2, 3, 4
- [Per85] PERLIN K.: An image synthesizer. In *Proceedings of the 12th Annual Conference on Computer Graphics and Interactive Techniques* (New York, NY, USA, 1985), SIGGRAPH '85, Association for Computing Machinery, p. 287–296. URL: <https://doi.org/10.1145/325334.325247>, doi:10.1145/325334.325247. 7
- [PLXL21] PENG J., LIU D., XU S., LI H.: Generating diverse structure for image inpainting with hierarchical vq-vae. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition* (2021), pp. 10775–10784. 3
- [PW20] PIELAWSKI N., WÄHLBY C.: Introducing hann windows for

- reducing edge-effects in patch-based image segmentation. *PLoS one* 15, 3 (2020), e0229839. 4, 10
- [RBL\*21] ROMBACH R., BLATTMANN A., LORENZ D., ESSER P., OMMER B.: High-resolution image synthesis with latent diffusion models, 2021. [arXiv:2112.10752](https://arxiv.org/abs/2112.10752). 3, 10, 11, 12, 13, 14
- [RFB15] RONNEBERGER O., FISCHER P., BROX T.: U-net: Convolutional networks for biomedical image segmentation. In *International Conference on Medical image computing and computer-assisted intervention* (2015), Springer, pp. 234–241. 1, 10
- [RZH\*19] REN D., ZUO W., HU Q., ZHU P., MENG D.: Progressive image deraining networks: A better and simpler baseline. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition* (2019), pp. 3937–3946. 3
- [SBB\*99] STAVELY D. J., BLOOM D. M., BATTLES A. E., CAMPBELL D. K., E. O. R. H.: Film scanner with dust and scratch correction by use of dark-field illumination, Oct. 1999. (Google Patents Entry) <https://patents.google.com/patent/US5969372A/>. 1
- [SCC\*22] SAHARIA C., CHAN W., CHANG H., LEE C., HO J., SALIMANS T., FLEET D., NOROUZI M.: Palette: Image-to-image diffusion models. In *ACM SIGGRAPH 2022 Conference Proceedings* (2022), pp. 1–10. 3
- [SLM\*22] SUVOROV R., LOGACHEVA E., MASHIKHIN A., REMIZOVA A., ASHUKHA A., SILVESTROV A., KONG N., GOKA H., PARK K., LEMPITSKY V.: Resolution-robust large mask inpainting with fourier convolutions. In *Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision* (2022), pp. 2149–2159. 1, 2, 3, 4, 10, 11, 12, 13, 14
- [SMF19] STRUBEL D., MARC B., FOFI D.: Deep learning approach for artefacts correction on photographic films. In *Fourteenth International Conference on Quality Control by Artificial Vision* (Mulhouse, France, May 2019), Cudel, Bazeille C., Verrier S., N, (Eds.), vol. 11172 of *Proceedings of SPIE*, SPIE-INT SOC OPTICAL ENGINEERING, 1000 20TH ST, PO BOX 10, BELLINGHAM, WA 98227-0010 USA, p. 35. URL: <https://hal-univ-bourgogne.archives-ouvertes.fr/hal-02369128>, doi:10.1117/12.2521421. 4
- [SMW85] SANZ J. L. C., MERKLE F., WONG K. Y.: Automated digital visual inspection with dark-field microscopy. *J. Opt. Soc. Am. A* 2, 11 (Nov 1985), 1857–1862. URL: <http://opg.optica.org/josaa/abstract.cfm?URI=josaa-2-11-1857>, doi:10.1364/JOSAA.2.001857. 1
- [The77] THE US NATIONAL ARCHIVES: Documerica project, 1971–1977. (Digitised Flickr Collection), <https://www.flickr.com/photos/usnationalarchives/collections/72157620729903309/>. 7, 8
- [TM21] TALEBI H., MILANFAR P.: Learning to resize images for computer vision tasks. In *Proceedings of the IEEE/CVF International Conference on Computer Vision* (2021), pp. 497–506. 4
- [WBSS04] WANG Z., BOVIK A., SHEIKH H., SIMONCELLI E.: Image quality assessment: from error visibility to structural similarity. *IEEE Transactions on Image Processing* 13, 4 (2004), 600–612. doi:10.1109/TIP.2003.819861. 8
- [WCTJ20] WANG Y., CHEN Y.-C., TAO X., JIA J.: VCNet: A robust approach to blind image inpainting. In *European Conference on Computer Vision (ECCV)* (2020). 3
- [WYW\*18] WANG X., YU K., WU S., GU J., LIU Y., DONG C., QIAO Y., CHANGE LOY C.: Esrgan: Enhanced super-resolution generative adversarial networks. In *Proceedings of the European conference on computer vision (ECCV) workshops* (2018). 3
- [WZC\*20] WAN Z., ZHANG B., CHEN D., ZHANG P., CHEN D., LIAO J., WEN F.: Bringing old photos back to life. In *proceedings of the IEEE/CVF conference on computer vision and pattern recognition* (2020), pp. 2747–2757. 2, 3, 4, 5, 8, 9, 10, 11, 12, 13, 14
- [WZCL22] WAN Z., ZHANG B., CHEN D., LIAO J.: Bringing old films back to life. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition* (2022), pp. 17694–17703. 4
- [YDLL18] YU K., DONG C., LIN L., LOY C. C.: Crafting a toolchain for image restoration by deep reinforcement learning. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (2018), pp. 2443–2452. 3
- [YLY\*18] YU J., LIN Z., YANG J., SHEN X., LU X., HUANG T. S.: Generative image inpainting with contextual attention. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (2018), pp. 5505–5514. 3
- [ZAK\*21] ZAMIR S. W., ARORA A., KHAN S., HAYAT M., KHAN F. S., YANG M.-H., SHAO L.: Multi-stage progressive image restoration. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition* (2021), pp. 14821–14831. 3
- [ZAK\*22] ZAMIR S. W., ARORA A., KHAN S., HAYAT M., KHAN F. S., YANG M.-H.: Restormer: Efficient transformer for high-resolution image restoration. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)* (June 2022), pp. 5728–5739. 3, 4, 10, 14
- [ZIE16] ZHANG R., ISOLA P., EFROS A. A.: Colorful image colorization. In *European conference on computer vision* (2016), Springer, pp. 649–666. 3
- [ZIE\*18] ZHANG R., ISOLA P., EFROS A. A., SHECHTMAN E., WANG O.: The unreasonable effectiveness of deep features as a perceptual metric. In *CVPR* (2018). 8