Deep Learning Algorithms as Potential Solutions to Challenges in Video Art Preservation

John H.F. Chow
Videotage

Sam H.C. Chan
Restituo

Abstract

Video artworks created in analog format present unique preservation challenges due to technological and personnel limitations. Current restoration approaches, including re-transferring from master tapes and digital image processing, are imperfect solutions. This paper explores the use of deep-learning algorithms to recover common defects in analog video art. It emphasizes the crucial need for a paradigm shift of concept in media art preservation for a possible way out from the long drastic debate on digital restoration measurement regarding the physical, aesthetical, and historical authenticity of preserving analog media artwork.

Keywords

Analog Media, Video Art, Deep Neural Network (DNN), Machine Learning, Media Art Archiving, Restoration

Introduction

The preservation of media art, particularly analog videotape, has become an increasingly pressing issue. The constant threat of technological obsolescence has aroused extensive discussions among conservators, curators, technicians, and community-driven initiatives over the past decade. The intricate nature of analog videotape and technological limitations have fostered a growing sense of confusion and evoked debates among practitioners. [1] This confusion arises not only from a lack of resources and knowledge for restoration efforts but also from the absence of common standards and a clear definition of a "perfect capture" of analog videotape for archival purposes. [2]

At present, there is no consensus or ideal solution for effectively preserving the physical, aesthetic, and historical integrity of this art form in a manner that is both sustainable and accessible for contemporary displays. [3] Despite this reality, patrons are inevitably swayed by the pervasive influence of the mainstream digital lobby. The transition from analog to digital is often portrayed and glorified as the ultimate panacea for all conservation challenges. This prevailing mindset has led to a significant surge in digitization efforts, as if the primary responsibility of conservators is simply to engage in mass digitization of all forms of analog artwork as a means to prevent their loss or deterioration. However, particularly within the realm of video art, there appears to be little interest in acknowledging that such digitization efforts do not genuinely preserve the original works. Instead, they merely create mediocre and highly fragile substitutes, at best serving as mere "viewing copies" of the artwork. [4] If this permissive and somewhat apathetic mindset persists, the risks of losing a substantial portion of media art history will inevitably emerge, especially in regions with fewer resources. The mechanical and chemical components of this unique art form will only deteriorate at an accelerated pace, further exacerbating the potential loss.

Challenges of Digitisation

Video artwork created before 1999 typically falls under the analog format, which includes media carriers like U-matic, VHS, and V8/Hi8 magnetic tapes. On the other hand, digital video tapes such as MiniDV store recorded sound and images as digital data. This allows for nearly identical transfers to contemporary devices. [5] Analog tapes record content as electromagnetic particles on a chemical binder, which are then read as an electrical signal. As a result, the quality of each playback heavily relies on the playback machines (e.g. VCRs), signal transferring wire, and display monitors (i.e. CRT monitors where images were created from accelerated electron hitting on phosphor-coated screen). This unique form of imaging method makes it extremely difficult to define and eliminate all defects during playback and to produce an objectively accurate representation of the work. [6] Current practices discourage restoring analog artwork digitally, such as by deinterlacing algorithmic-filters. [7] Instead, it is recommended that conservation technicians conduct forensic research on the analog material and ensure the quality control of the captured/digitized result. They must adapt all means to retain the analog playback equipment despite obsolescence. [8] Apart from the significant technical knowledge and resources needed in pursuing an identical digital representation of the artwork, smaller organizations often struggle to accurately identify the various types of video defects and the suitable methods to minimize them. [9] As a response, institutional definitions and community-driven initiatives have tried to identify a wide range of artifacts introduced by defective capture. They have narrowed these down to what is known as "common artifacts". [10]

In this paper, we propose possible considerations to digitally tackle these common artifacts. Instead of pursuing the vague objective of a perfect capture, we suggest exploring new options with digital image processing methods with the recent advancing application of deep learning algorithms in video restoration.

New Solution to Old Problems

In this section, we identify and categorize the common artifacts found in digital copies. General deep-learning-based approaches are suggested together with a handful of examples.

Examples are taken from our restoration project of Ellen Pau's video work *TV Game of the Year (1989)*. This project was initiated by the artist, with collaborative research effort from Videotage and Restituo. Artist's opinion and feedback are consulted to retain integrity of artistic intentions since the original master tape is no longer available.

Noises

Noises are identified as unwanted disturbances in an electronic signal. In the case of imaging, it was created due to contamination of the tape, or during faulty transmission of the signal. Common examples such as Video Dropout, manifest as the loss of one or more lines of the image. The losses are commonly caused by the deterioration of the magnetic binding on the tape known as "creasing", or it was contaminated by dust which causes Scratches or Tape Wears leading disruption to the smooth reading of the signal; and digital noise artifact such as Blocks, which is a common artifact appears to be a pixelated making the contour of image content in blurry; and SDI Spike usually appears to be a white line spikes across the whole frame. [10]

Prior studies have discussed extensively the use of Deep Neural Networks (DNN) to address specific "image cleaning" tasks such as denoising, deblurring, and reducing compression artifacts. [11] By adopting various pre-trained DNN models, we have achieved identifying and removing different noises from a faulty captured video signal, comparing results as demonstrated below. [12]

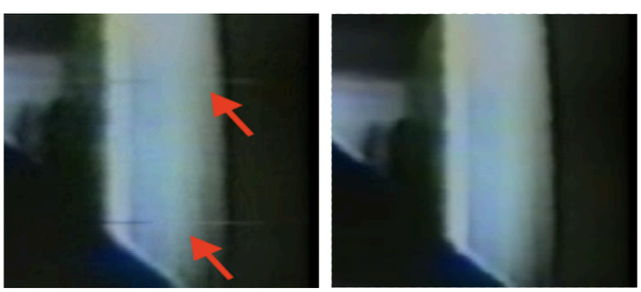

Fig 1. Example of clearing Video Dropout defect using CNN

Large Area Defects

Artifacts that cover a large portion of the image or disrupt the image sequence can be identified as Large Area Defects. The most common cause of the defect is Head Clogging (also known as a Dirty Head Drum), this issue introduces a wide horizontal line across the frame (indicated by the red arrow in Figure 2) which lasts for a certain duration, causing image distortion on the affected area. It occurs when the signal reading component on playback machines, known as the Head Drum, becomes contaminated, resulting in poor contact while reading the signal on the tape. [10] Other common causes of defect include Head Switch Noise and Magnetic Interference during playback capture.

Fig 2. Result of clearing Large Area Defects known as Head Clogging

The above restoration efforts were performed using the Image/Video Inpainting methods. [13] [14] [15] These methods fill in missing regions of a given video frame/sequence with content that is both spatially and temporally coherent. The replaced content data was generated in references to the same/consecutive video frame(s).

Chrominance Error

In analog tapes, both luma and chroma information is stored and transmitted within the same channel known as a composite signal. Where chrominance refers to the hue and saturation-related part of the signal. [16] This channel of data is susceptible to defects, usually caused by signal interruptions during transmission or degradation, or due to the limitation of Charge-Coupled Device (CCD) sensitivity in the initial recording occasion, faulty adjustment or condition of video processors, or muli-generation of information transfer. [17] Consequently, the color display on analog images is often unstable and inaccurate, as if deviating from the original creative intention.

Current restoration practice recommends monitoring the capturing signal through vectorscope and adjusting the playback operator based on the visualization of the frequency domain. [18] However, there is no effective and standard solution once the signals are digitized. Our proposed restoration approach consists of two parts: (1) Color stabilization, which addresses issues such as sudden color loss, shifting, leakages, noise and flashes; (2) Recolorization, which restores intended colors that were lost or altered due to defective signal, such as skin tones. We have adopted a combined approach that utilizes a colorization self-attention generative adversarial network (SA-GAN) and a Temporal Source-referencing Attention Network, while the former provides natural coloring suggestions and the latter uses them as references for generating stable colorization. [19] [20] These procedures lay the foundation for further color grading by maintaining a “healthy” color distribution of the image. This mixed approach allows for semi-interactive processing of numerous video restoration tasks in a single step, yielding consistent and stable recoloring results in long-duration video sequences.

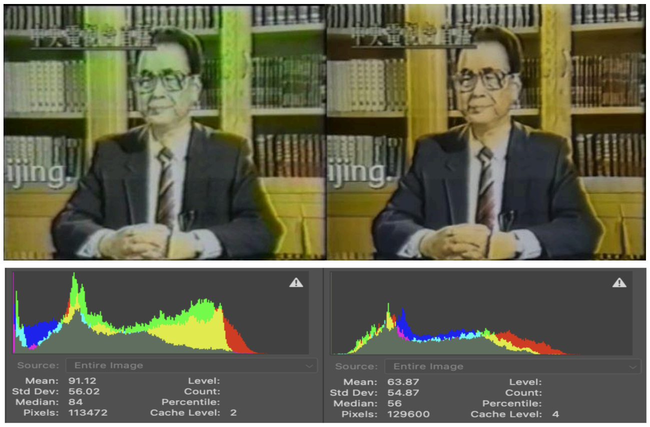

Fig 3. Restored color data by the combined method.

As shown in Figure 3, deep-learning-based methods have demonstrated the results of recovering and stabilizing large-scale signal-caused chrominance defects during digitalization. With the reconstruction of the color data profile, it also provides new possibilities for further color adjustment.

Display Limitation

Some proposed the digital equivalent to common VHS tape should be a 250x480 to 300x480 interlaced video stream. [21] In theory, any capturing or displaying device which exceeds this low resolution could potentially involve the use of interpolation algorithms (such as nearest-neighbor interpolation, bilinear/bicubic algorithms, etc.). [17] These might create unintended artifacts like Microblocks during playback. It is challenging to display analog videos on modern monitors while still retaining its unique aesthetic integrity since the modern market and manufacturing standard is full-HD and even going up to 4K resolution.

Video Super-Resolution (SR) has become a popular topic in machine learning and computer vision research recently. Along with this technological advancement, many commercial initiatives have started offering services based on deep-learning models. [5] Commercial media organizations, like broadcasting companies with large analog media collections, are beginning to trial these services. The aim is to upscale past collections, such as old TV programs and sports games. [22] We have tested with some of the popular video SR models, both Generative Adversarial Network(GAN) based and non-GAN-based approach, including Real-ESRGAN(2021), SwinIR (2021), BSRGAN (2021), EDSR (2017), SRGAN (2016) and SRCNN(2015). [23] [24] [25] [26] [27] [28] Digital artifacts were usually introduced during the processing of some GAN-based algorithms. These were particularly noticeable in detailed areas of the image, such as human faces, contours, and patterns, making them appear overly sharpened in an unnatural manner. Even for non-GAN-based methods, some of the results are not as favorable, as these methods are designed and trained with modern photographic images, which usually leads to over-sharpened or over-clean color palettes, as shown in fig 4.

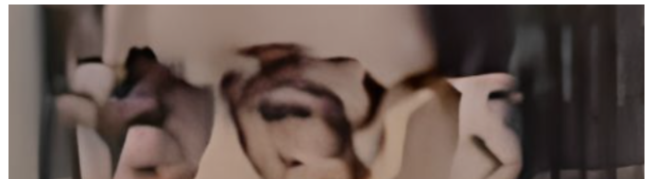
Fig 4. Example of unfavorable video super-resolution results using SRCNN(2015).

Some of the original SR methods work better in not overdoing the sharpness, though might still cause other problems like showing inconsistency within single images, as long as these methods divide images into smaller image patches and process them separately. In fact, similar satisfactory results can also be achieved using traditional computer vision methods like the Lanczos interpolation and resampling filter without creating the same defects. This resulted in more natural-looking images, especially in detailed and facial areas, as demonstrated in Figure 5.


Fig 5. Upscaling result (480i to 4K UHD) using Lanczos kernel.

Discussion

Misconceptions

In our old perception of art preservation, the sole objective is to protect and retain the authenticity of the work as it was initially collected. Conservators are trained and obligated to preserve the physical, aesthetic, and historical integrity of the artwork. The rules have appeared to be migrated and directed through such a rough transition to the digital environment, resulting in confusion about correct practices and often misconceptions in defining successful preservation of analog video artwork.

Pre-determined image processing algorithms

Back in the golden era of analog video, different commercial brands would compete in manufacturing the best quality of both the tapes and the player. This refers to the norms of the audience that would utilize a variety of built-in signal enhancement filters from different brands to adjust image quality. [29] Such phenomena open grounds for each audience to adjust the original image (i.e. blockbuster movie tape for rental) but only within the range of different offered brands (such as Sony or JVC). The employed algorithm or parameters are rather hidden behind the advertised "better color" or "sharper" but remain undisclosed to the public, which can be considered as a "black-box". Further extending this to the contemporary digital settings, the hidden and predetermined parameters still processed the video in every transfer, capture, or display activity (shown in Figure 6 below). Even for most image processing methods in the field of computer vision, they can be interpreted as a series of image convolution with manually designed kernels, as comparisons to the "trained" kernels through converging optimization in machine learning approaches. This further justifies the use of deep-learning algorithms in enhancing archival video quality as it shares no difference than all pre-determining image processing algorithms hidden involving hardwares and softwares in current preservation measures.

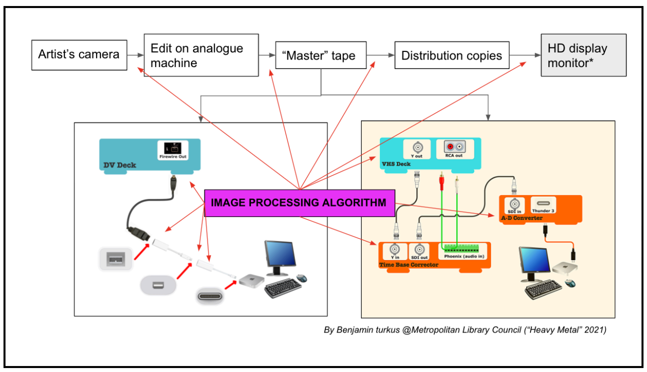

Fig 6. Infographic of Hidden Image Processing Algorithms.

Conclusion

In conclusion, the utilization of deep-learning algorithms could potentially represent a significant breakthrough in the restoration challenge of common defects found in video artwork from analog media formats. This advancement fills a crucial gap where previous remedies for rectifying faulty digitized files from videotape captures were ineffective. However, it is important to acknowledge that deep learning algorithms also possess inherent limitations. These limitations encompass the risk of producing unnatural and unintended outcomes and the challenge of establishing a widely accepted consensus among the proper use.

It is of paramount importance that policymakers and conservators recognize the potential of machine learning in offering fresh insights and inspiration for re-evaluating our existing practices and addressing prevailing issues. Additionally, fostering discussions among artists, memory institutions, historians, and conservation technicians regarding ethical and practical concerns is vital. Encouraging the documentation of innovative solutions will facilitate these discussions, ultimately leading to the refinement and redefinition of past preservation methodologies. Such endeavors will enable us to effectively adapt to the demands of our current digital world.

References

[1] "Ethics and Practices of Media Art Conservation, a Work-in-Progress (Version0.5) | Scart - a Website on Audiovisual Heritage by Packed," n.d. https://www.scart.be/?q=en/content/ethics-and-practices-media-art-conservation-work-progress-version05
[2] Bernard Serexhe, "Preservation or Reincarnation? At the Crossroads of Ethical Values," *Congreso Internacional Estéticas Híbridas De La Imagen En Movimiento: Identidad Y Patrimonio*, October 20, 2021. https://doi.org/10.4995/eshid2021.2021.14015
[3] Pip Laurenson, "Authenticity, Change and Loss in the Conservation of Time-Based Media Installations," *Tate Papers*, n.d. https://www.tate.org.uk/research/tate-papers/06/authenticity-change-and-loss-conservation-of-time-based-media-installations
[4] Perla Innocenti, "The Pursuit of Authenticity in Preserving Digital Art," *Digital Heritage*, September 1, 2015. https://doi.org/10.1109/digitalheritage.2015.7419540
[5] The Digital FAQ. "Guide to Understanding Video Sources, Part 2 – Capturing Videotapes," n.d. https://www.digitalfaq.com/guides/video/capture-understand-sources.htm#overscan
[6] Joanna Phillips, "Shifting Equipment Significance In Time-Based Media Art," *The Electronic Media Review 1*, 2012
[7] International Association of Sound and Audiovisual Archives, "Guidelines for the Preservation of Video Recordings." *IASA-TC*, no. 6 (March 2018).
[8] A. Bordina, and S. Venturini, "Operational Practices for Film and Video Preservation and Restoration Protocol," In *Preserving and Exhibiting Media Art. Challenges and Perspectives* (pp. 255-271). VU University of Amsterdam Press.
[9] Colin Post, "Preservation Practices of New Media Artists," *Journal of Documentation* 73, no. 4 (July 10, 2017): 716–32. https://doi.org/10.1108/jd-09-2016-0116
[10] Bay Area Video Coalition. "Common Artifacts " *AV Artifact Atlas*, n.d. http://www.avartifactatlas.com/tags.html#common-artifacts
[11] Claudio Rota et al., "Video Restoration Based on Deep Learning: A Comprehensive Survey," *Artificial Intelligence Review* 56, no. 6 (October 27, 2022): 5317–64, https://doi.org/10.1007/s10462-022-10302-5
[12] Satoshi Iizuka and Edgar Simo‐Serra, "DeepRemaster," *ACM Transactions on Graphics* 38, no. 6 (November 8, 2019): 1–13, https://doi.org/10.1145/3355089.3356570
[13] Wenbo Li et al., "MAT: Mask-Aware Transformer for Large Hole Image Inpainting," arXiv, 2022. https://doi.org/10.48550/arXiv.2203.15270
[14] Roman Suvorov et al., "Resolution-robust Large Mask Inpainting with Fourier Convolutions," arXiv, 2021. https://doi.org/10.48550/arXiv.2109.07161
[15] Kaidong Zhang, Jingjing Fu, and Dong Liu, "Flow-Guided Transformer for Video Inpainting," arXiv, 2022. https://doi.org/10.48550/arXiv.2208.06768
[16] "Care, Handling, and Storage of Audio Visual Materials - Collections Care - (Preservation, Library of Congress)," n.d. https://www.loc.gov/preservation/care/record.html
[17] AllTheVideoPeople, "PAR, SAR, and DAR: Making Sense of Standard Definition (SD) Video Pixels," BAVC Media, May 27, 2022, https://bavc.org/blog/par-sar-and-dar-making-sense-standard-definition-sd-video-pixels
[18] Video History Project. "Video Preservation - The Basics." Video History Project, n.d. https://www.videohistoryproject.org/video-preservation-basics
[19] Antoine Salmona, Lucía Bouza, and Julie Delon, "DeOldify: A Review and Implementation of an Automatic Colorization Method", *Image Processing On Line*, 12 (2022), pp. 347–368. https://doi.org/10.5201/ipol.2022.403
[20] Satoshi Iizuka, and Edgar Simo-Serra, "DeepRemaster: Temporal Source-Reference Attention Networks for Comprehensive Video Enhancement," arXiv. 2020. https://doi.org/10.48550/arXiv.2009.08692
[21] "How Sustainable Is File-Based Video Art? Exploring the Foundations for Best Practice Development – Electronic Media Review," n.d. https://resources.culturalheritage.org/emg-review/volume-4-2015-2016/bunz/
[22]Mike Tomkins, "Topaz Photo AI Review: Numerous ML-Powered Tools All in One App," *DPReview*, February 13, 2023. https://www.dpreview.com/reviews/topaz-photo-ai-review-numerous-ml-powered-tools-all-in-one-app
[23] Xintao Wang et al., "Real-ESRGAN: Training Real-World Blind Super-Resolution with Pure Synthetic Data," arXiv, 2021. https://doi.org/10.48550/arXiv.2107.10833
[24] Jingyun Liang et al., "SwinIR: Image Restoration Using Swin Transformer," arXiv, 2021. https://doi.org/10.48550/arXiv.2108.10257
[25] Kai Zhang et al., "Designing a Practical Degradation Model for Deep Blind Image Super-Resolution," arXiv, 2021. https://doi.org/10.48550/arXiv.2103.14006
[26] Bee Lim et al., "Enhanced Deep Residual Networks for Single Image Super-Resolution," arXiv, 2017. https://doi.org/10.48550/arXiv.1707.02921
[27] Christian Ledig et al., "Photo-Realistic Single Image Super-Resolution Using a Generative Adversarial Network," arXiv, 2016. https://doi.org/10.48550/arXiv.1609.04802
[28] Chao Dong et al., "Image Super-Resolution Using Deep Convolutional Networks," arXiv, 2015. https://doi.org/10.48550/arXiv.1501.00092
[29] Je H. Lee et al., "A Study on New DCT-Based Bit Rate Reduction Algorithm and Variable Speed Playback for a Home-Use Digital VCR," *IEEE Transactions on Consumer Electronics 38*, no. 3 (January 1, 1992): 236–42, https://doi.org/10.1109/30.156689

Bibliography

A. Bordina, and S. Venturini, "Operational Practices for Film and Video Preservation and Restoration Protocol," In *Preserving and Exhibiting Media Art. Challenges and Perspectives* (pp. 255-271). VU University of Amsterdam Press.

Annet Dekker, Collecting and Conserving Net Art: Moving beyond Conventional Methods, 2018. https://openlibrary.org/books/OL28922187M/Collecting_and_Conserving_Net_Art

Chao Dong et al. "Image Super-Resolution Using Deep Convolutional Networks." arXiv. 2015. https://doi.org/10.48550/arXiv.1501.00092

Johannes Gfeller, Agathe Jarczyk, Joanna Phillips, and Irene Schubiger. Kompendium Der Bildstörungen Beim Analogen

Video = Compendium of Image Errors in analog Video, 2012. http://ci.nii.ac.jp/ncid/BB15828918

Vinzenz Hediger, Cosetta Saba, Barbara Le Maître, and Julia Noordegraaf, Preserving and Exhibiting Media Art: Challenges and Perspectives, 2013. https://doi.org/10.26530/oapen_530353

Satoshi Iizuka, and Edgar Simo-Serra. "DeepRemaster: Temporal Source-Reference Attention Networks for Comprehensive Video Enhancement." arXiv. 2020. https://doi.org/10.48550/arXiv.2009.08692

Perla Innocenti, "The Pursuit of Authenticity in Preserving Digital Art," *Digital Heritage*, September 1, 2015. https://doi.org/10.1109/digitalheritage.2015.7419540

Pip Laurenson, "Authenticity, Change and Loss in the Conservation of Time-Based Media Installations," *Tate Papers*, n.d. https://www.tate.org.uk/research/tate-papers/06/authenticity-change-and-loss-conservation-of-time-based-media-installations

Christian Ledig et al. "Photo-Realistic Single Image Super-Resolution Using a Generative Adversarial Network." arXiv. 2016. https://doi.org/10.48550/arXiv.1609.04802

Je H. Lee et al., "A Study on New DCT-Based Bit Rate Reduction Algorithm and Variable Speed Playback for a Home-Use Digital VCR," *IEEE Transactions on Consumer Electronics 38*, no. 3 (January 1, 1992): 236–42, https://doi.org/10.1109/30.156689

Wenbo Li et al. "MAT: Mask-Aware Transformer for Large Hole Image Inpainting." arXiv. 2022. https://doi.org/10.48550/arXiv.2203.15270

Jingyun Liang et al. "SwinIR: Image Restoration Using Swin Transformer." arXiv. 2021. https://doi.org/10.48550/arXiv.2108.10257

Bee Lim et al. "Enhanced Deep Residual Networks for Single Image Super-Resolution." arXiv. 2017. https://doi.org/10.48550/arXiv.1707.02921

Joanna Phillips, "Shifting Equipment Significance In Time-Based Media Art," *The Electronic Media Review 1*, 2012

Charles Poynton, Digital Video and HD: Algorithms and Interfaces, *Elsevier*, 2003.

Claudio Rota et al., "Video Restoration Based on Deep Learning: A Comprehensive Survey," *Artificial Intelligence Review* 56, no. 6 (October 27, 2022): 5317–64, https://doi.org/10.1007/s10462-022-10302-5

Antoine Salmona, Lucía Bouza, and Julie Delon. "DeOldify: A Review and Implementation of an Automatic Colorization Method", *Image Processing On Line*, 12 (2022), pp. 347–368. https://doi.org/10.5201/ipol.2022.403

Tatja Scholte, and Glenn Wharton. Inside Installations: Theory and Practice in the Care of Complex Artworks, 2011.

Bernard Serexhe, "Preservation or Reincarnation? At the Crossroads of Ethical Values," *Congreso Internacional Estéticas Híbridas De La Imagen En Movimiento: Identidad Y Patrimonio*, October 20, 2021. https://doi.org/10.4995/eshid2021.2021.14015

Roman Suvorov et al. "Resolution-robust Large Mask Inpainting with Fourier Convolutions." arXiv. 2021. https://doi.org/10.48550/arXiv.2109.07161

Mike Tomkins, "Topaz Photo AI Review: Numerous ML-Powered Tools All in One App," *DPReview*, February 13, 2023.

Xintao Wang et al. "Real-ESRGAN: Training Real-World Blind Super-Resolution with Pure Synthetic Data." arXiv. 2021. https://doi.org/10.48550/arXiv.2107.10833

Kai Zhang et al. "Designing a Practical Degradation Model for Deep Blind Image Super-Resolution." arXiv. 2021. https://doi.org/10.48550/arXiv.2103.14006

Kaidong Zhang, Jingjing Fu, and Dong Liu. "Flow-Guided Transformer for Video Inpainting." arXiv. 2022. https://doi.org/10.48550/arXiv.2208.06768

Authors Biographies

John Chow Ho Fung currently serves as the project coordinator at the M+ museum in Hong Kong. He plays a pivotal role in the inaugural Asian Avant-Garde Film Circulation Library, an initiative focused on establishing the first-ever circulation platform for pioneering moving image works from the Asia-Pacific region in the 1960s-90s; He also actively involved in the M+ Restoration project for Hong Kong's feature-length classics. With a background as a former senior archivist at Videotage, John has shared his extensive experience in preservation and conservation practices for time-based media artwork in various platforms including ISEA, Art Basel HK, Photofair Shanghai, and ART Taipei.

Sam Chan Hin Chung studied Creative Media (New Media) at the City University of Hong Kong where he obtained his bachelor degree in 2016. This was followed by five years of research under the guidance of Hector Rodriguez and Prof Maurice Benayoun at the Center of Applied Computing and Interactive Media (ACIM) of City University of Hong Kong. He co-founds the professional restoration company Restituo and continues to lead the research. His major research interests include image/video processing and restoration utilizing spatio-temporal deep learning algorithms and attention-based transformers.