



PDF Download
3744925.pdf
14 January 2026
Total Citations: 0
Total Downloads: 280

Latest updates: <https://dl.acm.org/doi/10.1145/3744925>

RESEARCH-ARTICLE

Deep Chroma Compression of Tone-Mapped Images

XENIOS MILIDONIS, CYENS - Centre of Excellence, Nicosia, Cyprus

ALESSANDRO ARTUSI, CYENS - Centre of Excellence, Nicosia, Cyprus

FRANCESCO BANTERLE, Italian National Research Council, Rome, RM, Italy

Published: 22 November 2025

Online AM: 04 July 2025

Accepted: 12 June 2025

Revised: 08 May 2025

Received: 06 February 2025

[Citation in BibTeX format](#)

Open Access Support provided by:

CYENS - Centre of Excellence

Italian National Research Council

Deep Chroma Compression of Tone-Mapped Images

XENIOS MILIDONIS and ALESSANDRO ARTUSI, DeepCamera group, CYENS Centre of Excellence, Nicosia, Cyprus

FRANCESCO BANTERLE, Istituto di Scienza e Tecnologie dell'Informazione (ISTI), Consiglio Nazionale delle Ricerche (CNR), Pisa, Italy

Acquisition of High Dynamic Range (HDR) images is thriving due to the increasing use of smart devices and the demand for high-quality output. Extensive research has focused on developing methods for reducing the luminance range in HDR images using conventional and deep learning-based tone mapping operators to enable accurate reproduction on conventional 8- and 10-bit digital displays. However, these methods often fail to account for pixels that may lie outside the target display's gamut, resulting in visible chromatic distortions or color clipping artifacts. Previous studies suggested that a gamut management step ensures that all pixels remain within the target gamut. However, such approaches are computationally expensive and cannot be deployed on devices with limited computational resources. We propose a generative adversarial network for fast and reliable chroma compression of HDR tone-mapped images. We design a loss function that considers the hue property of generated images to improve color accuracy and train the model on an extensive image dataset. Quantitative experiments demonstrate that the proposed model outperforms state-of-the-art image generation and enhancement networks in color accuracy, while a subjective study suggests that the generated images are on par or superior to those produced by conventional chroma compression methods in terms of visual quality. Additionally, the model achieves real-time performance, showing promising results for deployment on devices with limited computational resources.

CCS Concepts: • **Computing methodologies** → **Computational photography**; *Image processing*; *Image and video acquisition*;

Additional Key Words and Phrases: High Dynamic Range, Deep Learning, Tone Mapping, Chroma Compression, Generative Adversarial Networks

ACM Reference format:

Xenios Milidonis, Alessandro Artusi, and Francesco Banterle. 2025. Deep Chroma Compression of Tone-Mapped Images. *ACM Trans. Multimedia Comput. Commun. Appl.* 21, 12, Article 346 (November 2025), 17 pages. <https://doi.org/10.1145/3744925>

The research work of Dr. Xenios Milidonis and Dr. Alessandro Artusi has been partially funded by the European Union's Horizon 2020 Research and Innovation Programme under grant agreement No. 739578 and by the Government of the Republic of Cyprus through the Deputy Ministry of Research, Innovation and Digital Policy.

Authors' Contact Information: Xenios Milidonis, DeepCamera group, CYENS Centre of Excellence, Nicosia, Cyprus; e-mail: x.milidonis@cyens.org.cy; Alessandro Artusi, DeepCamera group, CYENS Centre of Excellence, Nicosia, Cyprus; e-mail: a.artusi@cyens.org.cy; Francesco Banterle (corresponding author), Istituto di Scienza e Tecnologie dell'Informazione (ISTI), Consiglio Nazionale delle Ricerche (CNR), Pisa, Italy; e-mail: francesco.banterle@cnr.it.



This work is licensed under [Creative Commons Attribution-ShareAlike International 4.0](https://creativecommons.org/licenses/by-sa/4.0/).

© 2025 Copyright held by the owner/author(s).

ACM 1551-6865/2025/11-ART346

<https://doi.org/10.1145/3744925>

1 Introduction

High Dynamic Range (HDR) imaging is an important technology targeting the acquisition, display, and storage of high-fidelity images that contain a considerably wider luminance spectrum than conventional images [9]. HDR cameras [60], modern setups [3], and smartphones [26] can acquire a scene with intensities approaching the range perceivable by human vision, reliably capturing shadows and sunlight in the same image. However, due to the limited capacity of traditional monitors to display HDR images, range reduction techniques are employed to generate **Standard Dynamic Range (SDR)** images that can be accurately reproduced [4]. These techniques correct primarily the luminance, but subsequent color adjustment may be performed to remove any remaining chromatic distortions or color clipping artifacts (Figure 1). Most techniques do not consider color correction. The norm in HDR management on SDR display systems is the use of a **Tone-Mapping Operator (TMO)** to adjust the luminance alone [35, 48], which can often generate pixel values lying outside the target display’s gamut. When SDR images are displayed, such pixels are clipped to the gamut boundaries and appear over-saturated, affecting the overall conveyed color appearance of the original HDR image.

Šikudová et al. [59] were the first to develop a unified tone and chroma management framework that ensures pixels lie within the target gamut, while requiring no calibrated input or tunable parameters. This technique can be used in conjunction with any existing TMO and employs compression algorithms to correct both lightness and chroma channels. Despite its advantages, the technique is computationally expensive, hampering its deployment and use on devices with limited computational resources. Current trends in digital technology are pushing toward higher resolution displays with faster and more efficient processing capabilities, calling for a further step up in computational efficiency.

Deep learning offers a unique opportunity to tackle this problem as it can replace complex conventional image processing pipelines with simple fine-tuned models that often match their accuracy while significantly decreasing computational costs. In this work, we propose the use of an image-to-image generative model tailored for fast and reliable gamut management of tone-mapped HDR images [44]. Following extensive experiments with **State-of-the-Art (SOTA)** image generation and enhancement models, our approach exploits the seminal **Generative Adversarial Networks (GANs)** [30]. We correct undesirable chromatic effects produced by the prior step of tone mapping through careful network optimization and design a custom loss function to reduce hue distortions. For training, we developed a large dataset augmented by multiple TMOs. The main contributions of this work are as follows:

- We present the first deep learning model dedicated to chroma compression, compatible with all common TMOs.
- We use a novel loss function that combines the GAN and standard $L1$ losses with a hue-based loss to improve chromatic accuracy.
- Our model offers significant improvements in computational speed compared to the conventional framework.
- The extensive dataset of HDR and SDR images, code, and trained weights are publicly available.¹

2 Related Work

2.1 Chroma Compression

A display system’s gamut can be considered as a volume containing the colors that can be reproduced by that display [4]. If compared with the standard xy -chromaticities diagram of the overall colors

¹<https://github.com/DeepCamera/HDR-chroma-compression>.



Fig. 1. Accurate and efficient chroma compression is important for the reproduction of HDR images on standard displays with limited computational resources. This study proposes a deep learning-based method for chroma compression that is up to three orders of magnitude faster than the conventional approach, while generating pixel values that similarly fall within the target display's gamut (here *sRGB*). The scatter plot shows the chroma values (x axis) of pixels within a hue slice ($h = 60^\circ$) against their lightness values (y axis). Tone mapping cannot compress all pixels within the gamut boundaries, a problem that chroma compression is able to solve. HDR, High Dynamic Range.

perceived by the human visual system, a display's gamut is only a subset of it. Tone-mapped images often contain colors lying outside of this gamut, which need to be processed with dedicated chroma correction techniques to be accurately reproduced on displays [48]. Chroma correction has been explored for over 2 decades. Initial work attempted to match the viewing conditions between a display and the real-world scene using complex models of the human vision that partially adapted the luminance in the images [50]. Akyüz and Reinhard [2] later utilized color appearance models for preserving chromatic appearance across various scenes and display environments, but models required manual tuning of multiple parameters which hampered automation and adaptability to variable lighting conditions. Other work tried to develop simple color correction formulas using fewer parameters, focusing either on reducing out-of-gamut colors or enhancing the details in the images but with an impact on the naturalness of the scenes [39]. Pouli et al. [51] presented a parameter-free method for chroma compression with improved visual quality and adaptability to various scenes and TMOs. The same group later extended the algorithm by adding a gamut correction step to further reduce hue shifts and luminance distortion [5].

In parallel, Šikudová et al. [59] developed a unified framework combining color correction with a gamut management process that reliably handles out-of-gamut pixels. This method first compresses the luminance of an HDR image using any available TMO, corrects the chroma using a novel hue-specific approach that models the target gamut, and finally both luminance and chroma are processed to ensure that all pixels remain within the gamut boundaries. The authors showed that this approach is superior to the aforementioned methods in handling all pixels while avoiding desaturated or over-saturated results and preserving the details and naturalness of the scene. For these reasons, this method is widely considered among the SOTA in chroma compression, despite its development in the previous decade. However, its algorithmic complexity and reliance on both the HDR and tone-mapped SDR images for calculations lead to a substantial computational burden impacting its deployment potential.

2.2 Deep Learning in HDR Imaging

Deep learning methods for processing HDR content have gained considerable attention in recent years, due to the significant improvement in computational speed they offer. Various neural network architectures can learn contextual information in the images and produce realistic outputs in under- or over-saturated regions, something that conventional processing pipelines are lacking [64].

Nevertheless, the vast majority of methods target tone mapping or inverse tone mapping (the reconstruction of HDR from SDR data) [48, 64], and no published work has tackled the important problem of chroma compression for accurate image visualization.

A considerable effort was put into the reconstruction of HDR images from multi-exposure SDR images [31] or even from a single-exposure image [18, 42] using **Convolutional Neural Networks (CNN)**. Alternatively, Liu et al. [38] modeled the conventional HDR-to-SDR pipeline and developed a sequence of CNNs to learn the inverse process of HDR reconstruction. One of the scarce models that targeted HDR color enhancement using CNNs is the popular HDRNet [24], a multi-scale encoder–decoder that learns local affine transformations from a low-resolution version of the input image without the need to perform full image generation. Color transfer models represent another class of learning-based methods that aim to shift the color distribution of an image to match that of a reference, often based on histogram alignment or neural representations [7]. While effective for stylization, these models are not trained to enforce gamut constraints, and their reliance on target images makes them ill-suited for fast, consistent, device-specific color correction.

GANs have also been used for HDR image processing due to their ability to learn the true data distribution [25]. The highly successful implementations for paired and unpaired image-to-image translation, Pix2pix [30] and CycleGAN [72], respectively, have recently been adapted for unsupervised tasks in HDR imaging other than color enhancement [29, 47, 49, 69]. The adversarial loss in GANs plays an important role in the production of realistic outputs that follow the input data distribution, while the addition of a cycle consistency loss in the case of CycleGAN [72] allows the preservation of the semantic content present in the original images, which could represent the details and intensity levels in the context of HDR imaging. Wang et al. [65] later developed a multi-scale extension of Pix2pix suitable for high-resolution images. However, to our knowledge, no GAN has yet focused on chroma compression.

Driven by the inability of the standard $L1$ and $L2$ metrics to account for high-level image features and their typical use on non-perceptually uniform color spaces, studies in HDR imaging often use perceptual losses to ensure the production of images with improved visual quality [45, 56, 64]. However, while realistic and visually pleasing images are generally desirable, perceptually guided image generation risks producing out-of-gamut pixels similar to tone mapping techniques. Alternatively, others added a cosine similarity term in the $L1$ loss to improve color stability during SDR-to-HDR translation, measured in the input RGB color space [42].

3 Method

3.1 Chroma Compression Framework

To train the proposed deep learning-based chroma compression method, we considered the unified gamut management framework by Šikudová et al. [59] (Figure 2). In this framework, luminance and color are separated according to the standard $CIELCh$ color space assuming a specific target gamut; here, we used the $sRGB$ with $D65$ white point, the boundaries of which take the form of a triangular cusp for each hue value in $CIELCh$ coordinates (Figure 1). The chroma is compressed by linearly aligning the input and target chroma cusps while maintaining lightness and hue. Fine details in the image are preserved by splitting each hue slice into base and details layers, of which only the base is compressed. Finally, the remaining pixels lying outside the target gamut are clipped along both luminance and chroma channels. The framework was implemented in MATLAB 2022b.

3.2 Dataset

We created a dataset covering a diverse range of scenes and lighting conditions. The dataset contains 1,467 HDR images at a resolution of 512×512 pixels, captured by the authors or collected from

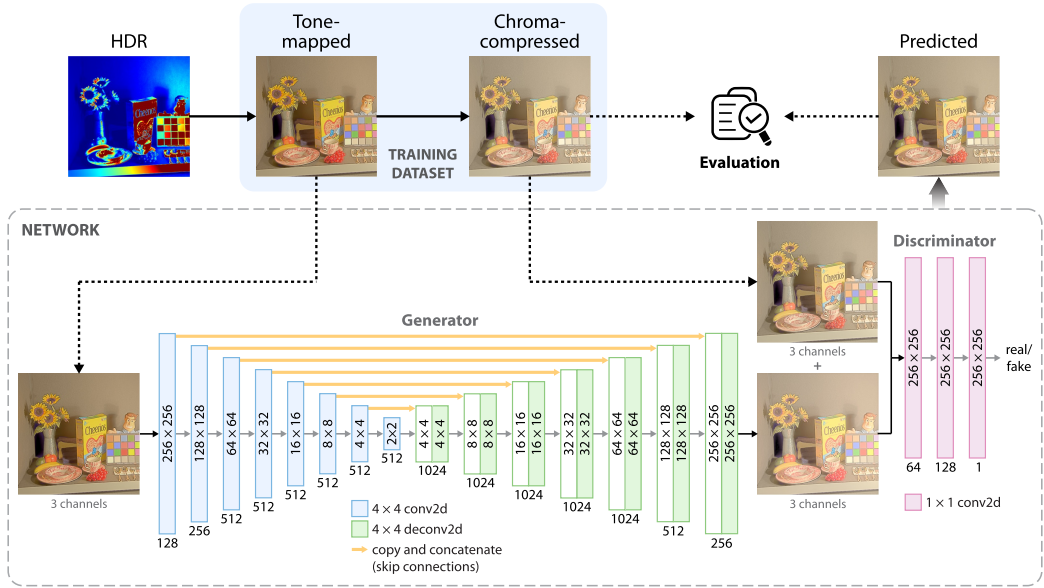


Fig. 2. Overview of method and evaluation. HDR images were tone-mapped and then chroma-compressed into SDR images. Our GAN learns the same process, using a U-Net generator and a pixel-based discriminator. Since the U-Net is fully convolutional, it can process images with any resolution divisible by 2^n , where n is the number of downsampling steps (here we use 8). A relatively large number of generator filters are used in the first convolutional layer (128), which capture more complex patterns and help improve sharpness. The lightweight pixel-based discriminator focuses on pixel-wise fidelity, in contrast to typical patch-based architectures that have more convolutional layers. For performance evaluation, the final images predicted by the trained network were compared against the reference chroma-compressed images.

different publicly available HDR repositories [1, 8, 19, 22, 23, 27, 28, 46, 68]. We carefully inspected images to exclude identical scenes or images with very low native resolution. From HDR videos, we selected a frame every 3 seconds to avoid duplicating the scene. Furthermore, we removed black and white borders from panoramic or other images.

Rather than using a single TMO to generate the reference SDR images, we used a large number of existing TMOs to augment the dataset and generalize the model (see Table 3 for the full list). Tone mapping was performed using a MATLAB toolbox [9]. Images were then chroma-compressed with the chroma compression framework [59]. The final dataset contained a total of 42,543 pairs of reference tone-mapped and chroma-compressed images (Figure 2) and was split into training, validation, and test subsets at an 8:1:1 ratio (34,017, 4,263, and 4,263 images, respectively). TMO versions of the same image were not shared between subsets, as previously suggested [10].

3.3 Model Implementation and Training

We propose the use of a GAN architecture, which has proven effective for an extensive range of image-to-image translation applications [30]. We have extensively evaluated the performance of this and alternative base models (compared below) and found that it performs more accurately for this task following architecture and hyperparameter optimization. The model consists of a generator and a discriminator, which compete in an adversarial manner to produce an image and then determine whether it is fake or real, respectively, driving the optimization of the model weights (Figure 2). Isola et al. [30] used (i) a conditional GAN objective that the generator tries to minimize

against an adversarial discriminator that tries to maximize it, and (ii) a generator $L1$ loss to ensure generated images are visually close to the reference images. In this study, we chose a least-squares GAN objective due to its increased stability during the learning process, leading to the production of higher-quality images compared to standard GANs [41]. The GAN loss is the MSE. The model's generator uses a U-Net architecture and a pixel-based rather than a more common patch-based discriminator to reduce tiling artifacts while keeping the details unaffected. A total of 128 generator and 64 discriminator filters are used in the last and first convolutional layers, respectively.

For training, we used the Adam solver with a momentum term of 0.5 and an initial learning rate of 10^{-4} , with a linear policy and 50 decay iterations. Instance normalization was performed to improve image quality and due to its compatibility with the U-Net architecture [62]. Orthogonal weight initialization was used to speed up the convergence. No data pre-processing was applied so that the images were loaded and used for training at their native resolution.

3.4 Hue-Based Loss Function

$L1$ and $L2$ loss functions are based on the absolute difference between the reference and predicted image pixel values, and they have proven critical in image generation and restoration tasks [64]. However, since these losses are traditionally used on non-perceptually uniform color spaces, they consider all pixels equally and do not give weight to regions that are perceptually more important. On the other hand, perceptual metrics alone or in combination with $L1$ and $L2$ were shown to improve the visual quality of images [71], but are less useful in tasks requiring chromatic consistency (i.e., constraining rather than expanding the images' gamut), as the one concerned in this study. Since the task of color correction focuses primarily on the image's hue, we propose the use of an improved loss function that shifts the attention to the hue channel to improve the chromatic accuracy of the result. The predicted and reference images are first converted from RGB to $CIELCh$ space through $CIELab$. The hue channel H is computed from the $CIELab$ components a and b as:

$$H = \arctan(b, a).$$

The hue loss \mathcal{L}_H is defined as the mean squared angular difference between the predicted (\hat{H}) and reference (H) hue channels:

$$\mathcal{L}_H = \frac{1}{mn} \sum_{i,j} \left(\hat{H}_{ij} - H_{ij} \right)^2,$$

where m and n denote the image height and width, respectively. \mathcal{L}_H is measured in an $L2$ sense due to a small improvement in performance when used in this context, compared to the $L1$ norm. All color space operations are implemented with differentiable layers to preserve gradient flow during backpropagation [55]. The loss is combined with the conditional GAN objective to capture global image statistics and perceptual realism to enhance the generalization capabilities of the network, as well as the standard $L1$ loss to improve pixel-wise accuracy. The final objective is given below:

$$G^* = \arg \min_G \max_D \mathcal{L}_{cGAN}(G, D) + \lambda_1 \mathcal{L}_{L1}(G) + \lambda_H \mathcal{L}_H(G).$$

Weights λ_1 and λ_H were fine-tuned independently and in combination; the optimal values were found to be 100 and 10, respectively, which encourage the model to favor pixel-wise accuracy over its generative ability.

4 Evaluation

In all experiments described below, unless otherwise stated, the same test dataset was used for an unbiased evaluation. Reference and predicted images were converted to double-precision floating-point format and normalized between $[0, 1]$. The performance of the models was evaluated using

standard metrics (**Peak Signal-to-Noise Ratio (PSNR)** and **Structural Similarity Index (SSIM)**) [48], metrics that focus on color accuracy (mean absolute error of the hue channel based on the *CIE LCh* color space (MAE_H), and root MSE of the color channels in the *CIE Lab* color space ($\Delta E_{CIE Lab }$) [58], and a recent color-aware visual quality metric (**ColorVideoVDP (CVVDP)**) [40]. CVVDP has **Just-Objectable-Difference (JOD)** units with a maximum possible value of 10 indicating no difference. Model performance at the local level was evaluated using the CVVDP distortion map, which overlays a heatmap with pixel-wise JOD differences between 0 and 1 over a grayscale version of the image [40]. Models were trained and tested on a GPU cluster with eight nodes, each with eight NVIDIA RTX A5000 (24 GB on-board GPU memory), running Rocky Linux 8.5, and using Python 3.10 and PyTorch 1.13. Test metrics and processing times for the various models are reported using the mean and were compared using one-way analysis of variance. Statistical significance was set at $p < 0.05$.

4.1 Comparison against SOTA Models

There are no published neural networks for chroma compression of SDR images. Therefore, the performance of the proposed model was compared against well-established image-to-image translation models, including three GANs and an HDR-focused transformation model: Pix2pix [30], CycleGAN [72], Pix2pixHD [65], and HDRNet [24]. All models were trained from scratch and tested using the same SDR images as our method, and using the parameters described in the corresponding papers. Training was performed for 200 epochs for all models; HDRNet converges faster than the GANs, but it was found that the same number of epochs provides a performance improvement over fewer epochs.

4.2 Component Contribution

In this experiment, the proposed model was trained sequentially by replacing critical components with other common techniques to determine their impact on performance. A model trained without an adversarial loss (L_1 only) to focus exclusively on pixel-wise accuracy was also considered.

4.3 Leave-One-Out TMO Evaluation

To evaluate the generalizability of the model, leave-one-out experiments were performed by training the model on SDR images tone-mapped with all except one TMO. The model was then tested on images tone-mapped with the excluded TMO. This was repeated for five randomly selected TMOs.

4.4 Subjective Evaluation

To evaluate the perceived visual quality of the images generated by the proposed method compared to dedicated conventional chroma-compression methods [5, 39, 51, 59], we designed a two-alternative forced-choice experiment. Ten images were drawn from the test subset, covering both indoor and outdoor scenes captured in either daylight or nighttime and encompassing a large dynamic range of luminance and chromaticity (Figure 3(a)). The study compared the images produced by each method in pairs shown in successive trials (a total of 100 pairs covering all possible combinations between ours and four existing methods). Each pair of images was displayed side-by-side over a dark gray background (Figure 3(b)), and between trials, a light gray background was shown for 3 seconds. Both the order of the two images in each pair and the order of trials were randomized. In each trial, participants were asked to select the image that they preferred based on perceived realism/naturalness and overall visual quality. Before the experiment, participants were given written instructions and informed consent was obtained. The experiment was performed on a 4K HDR hardware calibration LCD monitor (Eizo ColorEdge CG319X), located in a dark room. The monitor was set to the standard SDR mode, the *sRGB* color space with $D65$ white point, and

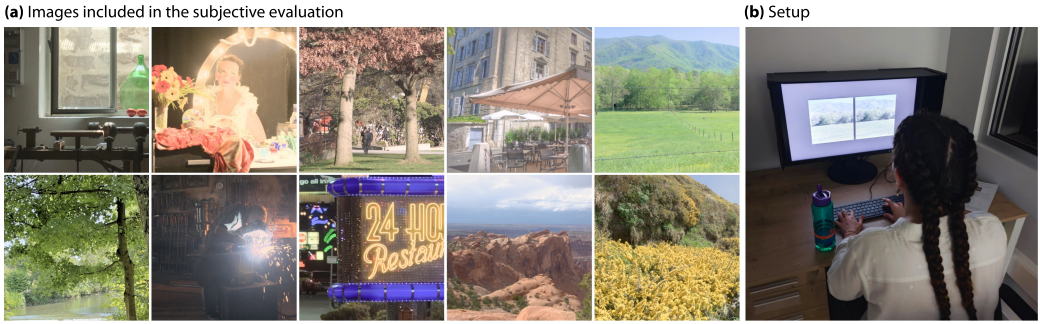


Fig. 3. (a) Selected images for the subjective evaluation. (b) The experimental setup. The room is shown here with dim lights on for visualization purposes.

a peak luminance of 100 cd/m^2 . Votes for each method were assessed using paired comparison standard analysis [15], which has previously been used in computer graphics and is equivalent to the analysis of variance [12, 35]. Methods were grouped using the separating $R = 90$ value at significance level $\alpha = 0.05$.

4.5 Computational Performance

The inference time during testing of the proposed model was measured to evaluate its computational performance in comparison to the reference conventional framework [59]. A batch size of 1 was used to process each image in series with a single GPU. For the conventional approach, both CPU and CUDA implementations of the chroma compression process were considered.

5 Results

Figure 4 and Table 1 compare the performance of the proposed model against common image generation and dedicated HDR image enhancement models. Our model significantly outperforms considered methods based on all metrics in Table 1 ($p < 0.001$). Among the previous models, HDRNet was found to have the next best performance in chroma compression, followed by Pix2pix. It should be emphasized that efforts to improve the performance of these models suggested that an optimized GAN architecture leads to higher overall pixel-wise accuracy; this could be because GANs target both global patterns and fine-grained local features more effectively than traditional CNN-based methods, making them more compatible with global and local TMOs [24]. At the local level, CVVDP distortion maps in Figure 4 demonstrate the performance of all models in the production of images with accurate colors. The standard Pix2pix generally performs well, but it often generates tiling artifacts that can be missed by qualitative evaluation (Figure 5). Despite being an unsupervised model, CycleGAN can preserve the clarity and colorfulness of the images. However, our results indicate that CycleGAN and Pix2pixHD lack consistency in the production of color-accurate results (Figure 4). Our simple GAN architecture significantly improves pixel-wise accuracy with considerably fewer chromatic distortions, while being effective in suppressing any artifacts generated by other models.

Various components contribute to the increase in model performance. As shown in Table 2, the pixel-based discriminator increases PSNR by 1.7 dB compared to a 70×70 PatchGAN discriminator, instance normalization increases SSIM by 0.005 compared to batch normalization, and orthogonal weight initialization improves MAE_H by 0.024 compared to normal initialization. Additionally, the hue-based loss function improves MAE_H by 0.006 compared to the standard loss ($\text{GAN} + L1$).

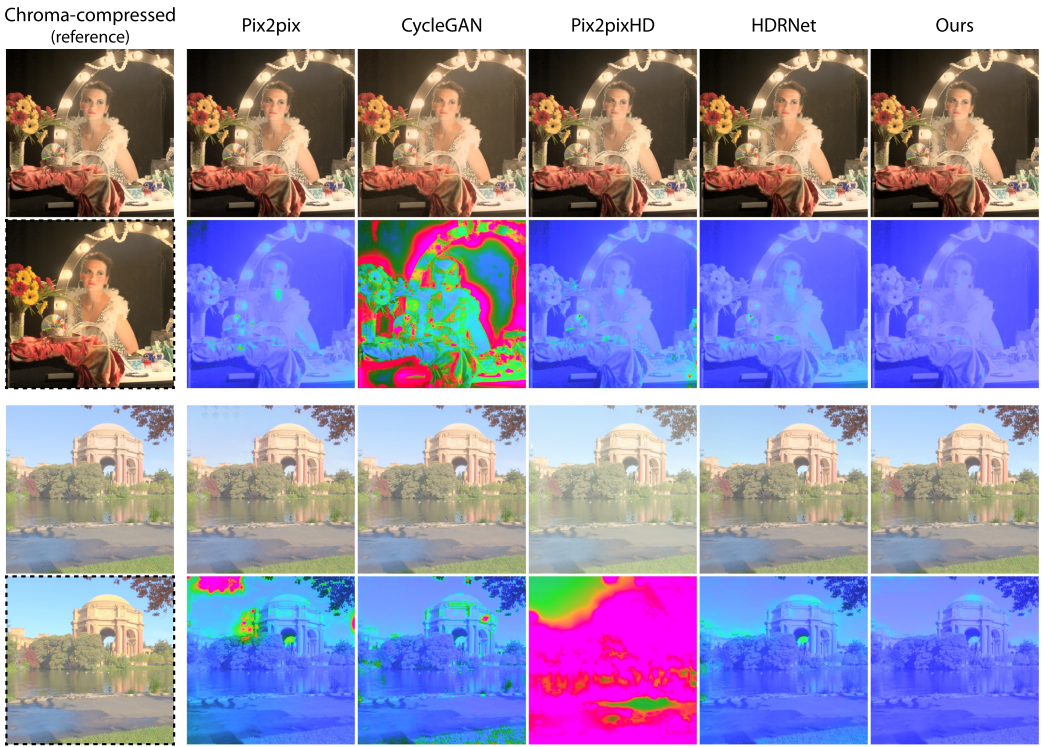


Fig. 4. An example image tone-mapped (dashed frame) and subsequently chroma-compressed by the reference method, various SOTA models for image restoration, and our method. Corresponding CVVDP distortion maps are also shown (blue indicates smaller distortion and red/pink indicates larger distortion).

Table 1. Model Comparisons

Model	PSNR (dB) \uparrow	SSIM \uparrow	MAE _H \downarrow	ΔE_{CIELab} \downarrow	CVVDP (JOD) \uparrow
Pix2pix	35.7	0.941	0.463	2.72	9.63
CycleGAN	28.6	0.921	0.489	4.00	9.28
Pix2pixHD	27.5	0.898	0.533	4.61	9.14
HDRNet	40.6	0.982	0.250	1.60	9.82
Ours	43.6	0.993	0.168	1.06	9.91

The best scores are highlighted in bold.

A vanilla loss function without an adversarial or hue-based component ($L1$) provides similar performance across all metrics, but it is still inferior overall to our full hue-based loss.

Figure 5 shows the impact that the size of the discriminator’s receptive field has on the generated images. A standard patch-based discriminator [30] often generates tiling artifacts in the images, particularly in bright areas and usually small in size. These artifacts are removed by the use of a discriminator evaluating all individual pixels (i.e., 1×1 patches), without affecting the colorfulness or spatial sharpness of the results. Figure 6 demonstrates the effect of the hue-based loss function. Despite not having an effect on image sharpness, generated images tend to have improved chromatic accuracy in challenging regions with high hue or saturation values.

Table 3 summarizes the performance of the proposed model on images tone-mapped by individual TMOs. The proposed method handles images tone-mapped by various TMOs with high accuracy,

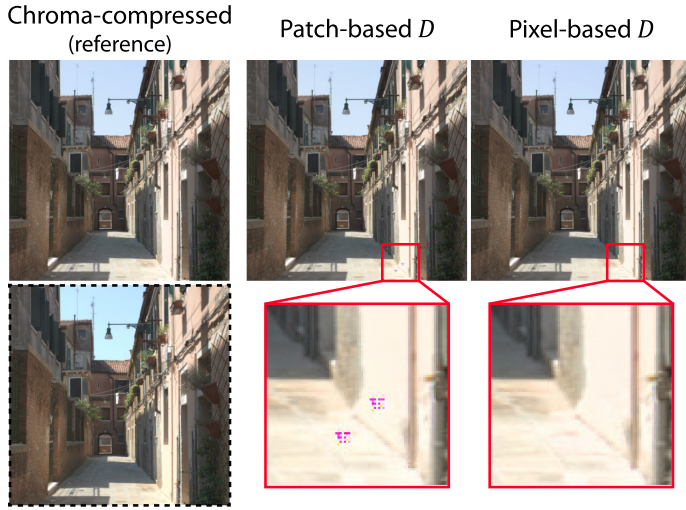


Fig. 5. The effect of the size of the discriminator’s receptive field. The 70×70 patch-based discriminator occasionally generates tiling artifacts that the pixel-based discriminator can suppress. The original tone-mapped image has a dashed frame.

Table 2. Contribution of Model Components

Component	PSNR (dB) \uparrow	SSIM \uparrow	MAE _H \downarrow	ΔE_{CIELab} \downarrow	CVVDP (JOD) \uparrow
Patch-based D	41.9	0.991	0.175	1.23	9.86
Image-based D	43.1	0.992	0.175	1.10	9.90
64 G filters	42.6	0.991	0.190	1.18	9.89
Batch norm.	42.4	0.988	0.183	1.21	9.89
Normal init.	43.0	0.990	0.192	1.16	9.89
Cross-entropy obj.	43.5	0.992	0.174	1.08	9.91
(GAN + $L1$) loss	43.5	0.992	0.174	1.07	9.90
$L1$ loss	43.6	0.992	0.170	1.06	9.91
Ours	43.6	0.993	0.168	1.06	9.91

In each row, we replaced a component of our model, as described in Section 3.3, with alternative common techniques. All other model parameters in each row remain fixed. The best scores are highlighted in bold.

with PSNR ranging between 40.2 dB and 47.5 dB, SSIM between 0.983 and 0.997, MAE_H between 0.108 and 0.261, ΔE_{CIELab} between 0.65 and 1.55, and CVVDP between 9.81 JOD and 9.95 JOD. The results of leave-one-out experiments are given at the bottom of Table 3. All performance metrics suggest that each model’s accuracy is on par with the model trained on images from all TMOs and generalizes well to the unseen TMOs.

In the subjective study, we recruited 21 participants (mean age 33 [24–58] years, 15 men and 6 women), all having a normal or corrected-to-normal vision and no color blindness. Table 4 shows the results of paired comparison standard analysis [15]. The coefficients of consistency and agreement are 0.810 (possible range [0, 1]) and 0.093 (possible range [−1, 1]) respectively, suggesting that there is high consistency overall in voting and a positive agreement among voters [15]). Participants agree on voting at a statistically significant level based on a X^2 of 204.76 (critical value at $\alpha = 0.05$ is 18.31), and differences in scores were significant based on a D_n of 197.93 (critical value at $\alpha = 0.05$ is

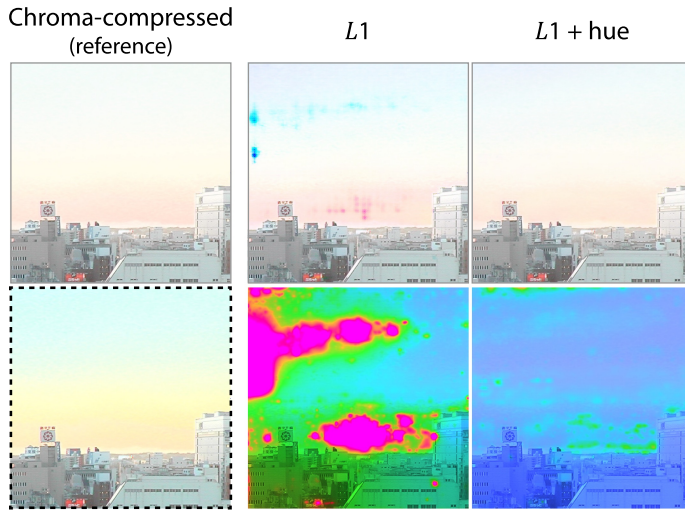


Fig. 6. The effect of the hue-based loss function on model performance. Adding a hue loss in the final objective can improve color accuracy in challenging high-hue or saturated regions. Corresponding CVVDP distortion maps are also shown (blue indicates smaller distortion and red/pink indicates larger distortion). The original tone-mapped image has a dashed frame. This image is courtesy and copyright of Frédéric Drago [16].

9.45). The analysis revealed that our neural network is equivalent to Pouli’s and Šikudová’s analytic methods in producing visually pleasing images and was selected significantly more often than Artusi’s and Mantiuk’s methods (Figure 7). Pouli’s method was chosen less consistently than our and Šikudová’s methods (standard deviation is 0.82 versus 0.44 and 0.35, respectively).

Computational performance at inference time is significantly better for the proposed model compared to the conventional chroma compression framework at all tested resolutions ($p < 0.001$; Figure 8). For example, for the smallest resolution of 512×512 (0.26 MPixel), our approach takes 15 ms per image, compared to 627 ms for the conventional CUDA and 3.8 s for the conventional CPU approaches (41-fold and 247-fold improvement, respectively). Real-time compression is maintained up to 1 MPixel with 25 frames per second, bringing the performance of chroma compression to the level of optimized tone mapping techniques [48]. At the highest tested resolution of 9.4 MPixel, the model can process images significantly more efficiently (375 ms, compared to 7.1 s for the conventional CUDA and 133.0 s for conventional CPU approaches). The computational advantage stems from the streamlined U-Net architecture with a pixel-based discriminator, avoiding the potentially increased overhead of patch-based or multi-scale GANs. In contrast, the conventional method by Šikudová et al. involves complex gamut modeling and iterative processing in $CIELCh$ space that significantly increase computational burden [59].

6 Conclusion

In this work, we proposed a method for fast, accurate, and generalizable chroma compression of tone-mapped images. To the best of our knowledge, this is the first deep learning approach performing chroma compression, offering excellent performance in terms of pixel-wise accuracy, color accuracy, and imaging artifacts. Importantly, the model’s performance based on PSNR and SSIM is superior to that reported for a variety of other tasks in HDR imaging [24, 64]. Despite being of secondary importance, the proposed method also matches [51, 59] or surpasses [5, 39] the visual quality of the results of conventional SOTA frameworks for chroma compression. The reference framework by Šikudová et al. [59] was chosen as the foundation due to its superiority

Table 3. Performance for Individual TMOs

TMO	PSNR (dB) \uparrow	SSIM \uparrow	MAE _H \downarrow	ΔE_{CIELab} \downarrow	CVVDP (JOD) \uparrow
Ashikhmin [6]	43.5	0.995	0.142	1.02	9.91
Banterle [11]	42.9	0.994	0.143	1.11	9.90
Bruce [13]	42.1	0.993	0.155	1.23	9.89
Chiu [14]	42.6	0.986	0.184	1.29	9.88
Drago [16]	42.6	0.992	0.162	1.17	9.89
Durand [17]	41.9	0.993	0.151	1.26	9.88
Fattal [20]	42.3	0.992	0.154	1.19	9.89
Ferwerda [21]	45.5	0.993	0.190	0.82	9.93
Kim–Kautz [32]	44.3	0.995	0.145	0.95	9.93
Krawczyk [33]	43.4	0.992	0.171	1.05	9.90
Kuang [34]	45.3	0.995	0.108	0.83	9.95
Li [36]	43.4	0.997	0.133	1.01	9.92
Lischinski [37]	44.7	0.995	0.130	0.92	9.93
Mertens [43]	47.5	0.994	0.217	0.65	9.95
Raman [52]	41.1	0.983	0.261	1.40	9.88
Reinhard [54]	44.0	0.995	0.134	1.00	9.92
Reinhard–Devlin [53]	41.6	0.990	0.184	1.34	9.88
Schlick [57]	44.9	0.994	0.158	1.02	9.92
Tumblin [61]	43.3	0.992	0.159	1.12	9.90
Tumblin (YP)	41.6	0.990	0.190	1.33	9.88
van Hateren [63]	44.0	0.990	0.176	1.07	9.91
Ward [67]	45.0	0.996	0.162	0.84	9.93
Ward global [66]	44.6	0.993	0.175	0.93	9.91
Ward global (YP)	44.4	0.993	0.161	0.95	9.92
Best exposure*	44.9	0.994	0.153	0.90	9.92
Exponential*	40.2	0.985	0.226	1.55	9.81
Logarithmic*	44.2	0.992	0.175	1.00	9.91
Normalize*	46.0	0.994	0.221	0.74	9.94
Photographic*	43.6	0.994	0.142	1.05	9.91
<i>Chiu</i>	40.7	0.990	0.150	1.08	9.91
<i>Krawczyk</i>	41.1	0.994	0.142	0.87	9.92
<i>Reinhard</i>	43.9	0.998	0.102	0.68	9.95
<i>Van Hateren</i>	43.5	0.989	0.167	0.85	9.93
<i>Logarithmic*</i>	44.0	0.995	0.146	0.72	9.95

The YP methods use Yee and Pattanaik’s luminance adaptation [70]. Five methods using basic luminance operations were also considered (*). Below, the metrics for the five leave-one-out experiments (in *italics*) are provided (values for the same TMOs may differ also due to the different size of test subsets).

over competing methods, which can generate similarly visually pleasing results (e.g., [51]) but are more susceptible to hue distortion artifacts and out-of-gamut pixels, as previously demonstrated [59]. Therefore, our model, trained to be a significantly faster alternative to the conventional framework, shows great promise for applications where efficient and accurate color reproduction is of paramount importance. Our work offers a simple solution that could be implemented on devices with limited computational resources and software employing any TMO for handling HDR data, such as smartphones and tablets, reducing chromatic artifacts in captured images and videos. The proposed model is currently tailored for the *sRGB* color gamut, which remains the most widely adopted standard for SDR content across consumer displays and web platforms, making it relevant to practical deployment scenarios. Adapting the method to support wider color gamuts was beyond

Table 4. Analysis of the Subjective Evaluation

	Coeff. Consistency	Coeff. Agreement	X^2	D_n
Overall	0.810	0.093	204.76	197.93
Groups				
Overall	M(288) A(305)		S(477) O(513) P(517)	

The top row provides the coefficients of consistency and agreement, as well as the X^2 value for the coefficients of agreement and the D_n value for the test of equality to determine the significance of our data. The bottom row reports the ranking of the methods with total votes in parentheses. We group together methods that are statistically the same using the critical value $R = 90$ at significance level $\alpha = 0.05$. A, Artusi; M, Mantiuk; O, Our Model; P, Pouli; S, Šikudová.

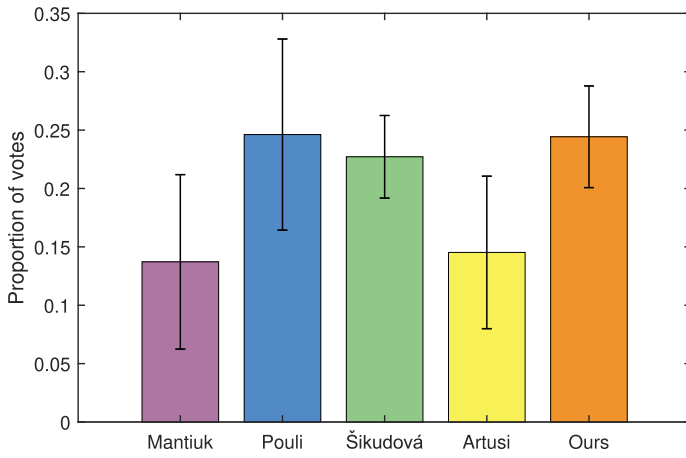


Fig. 7. Results of the subjective evaluation. The graph shows the proportion of times each method was selected by participants. Error bars indicate the standard deviation.

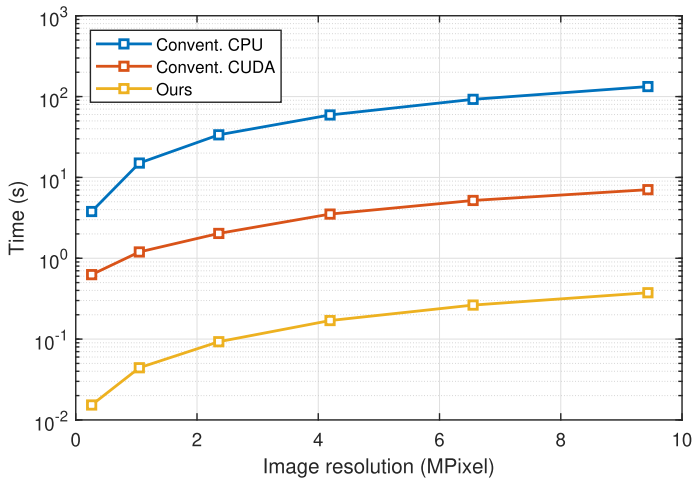


Fig. 8. Mean inference time per image for the conventional chroma compression algorithm (CPU and CUDA versions) and our approach.

the scope of our study and could be the subject of future work. Finally, while the proposed method is theoretically adaptable to video processing, future research could incorporate it into a pipeline for real-time HDR video management.

References

- [1] Ahmet Oğuz Akyüz, Roland Fleming, Bernhard E. Riecke, Erik Reinhard, and Heinrich H. Bühlhoff. 2007. Do HDR displays support LDR content?: A psychophysical evaluation. *ACM Transactions on Graphics* 26, 3 (2007), 38. DOI: <https://doi.org/10.1145/1276377.1276425>
- [2] Ahmet Oguz Akyüz and Erik Reinhard. 2006. Color appearance in high-dynamic-range imaging. *Journal of Electronic Imaging* 15, 3 (2006), 033001. DOI: <https://doi.org/10.1117/1.2238891>
- [3] Masheal Alghamdi, Qiang Fu, Ali K. Thabet, and Wolfgang Heidrich. 2021. Transfer deep learning for reconfigurable snapshot HDR imaging using coded masks. *Computer Graphics Forum* 40, 6 (2021), 90–103. DOI: <https://doi.org/10.1111/CGF.14205>
- [4] Alessandro Artusi, Francesco Banterle, Tun Ozan Aydn, Daniele Panozzo, and Olga Sorkine-Hornung. 2016. *Image Content Retargeting: Maintaining Color, Tone, and Spatial Consistency*. A. K. Peters, Ltd.
- [5] Alessandro Artusi, Tania Pouli, Francesco Banterle, and Ahmet Oğuz Akyüz. 2018. Automatic saturation correction for dynamic range management algorithms. *Signal Processing: Image Communication* 63 (2018), 100–112. DOI: <https://doi.org/10.1016/j.image.2018.01.011>
- [6] Michael Ashikhmin. 2002. A tone mapping algorithm for high contrast images. In *Proceedings of the 13th Eurographics Workshop on Rendering (EGRW '02)*. Eurographics Association, Goslar, DEU, 145–156.
- [7] Mor Avi-Aharon, Assaf Arbelle, and Tammy Riklin Raviv. 2023. Differentiable histogram loss functions for intensity-based image-to-image translation. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 45, 10 (2023), 11642–11653. DOI: <https://doi.org/10.1109/TPAMI.2023.3278287>
- [8] Maryam Azimi, Amin Banitalebi-Dehkordi, Yuanyuan Dong, Mahsa T. Pourazad, and Panos Nasiopoulos. 2018. Evaluating the performance of existing full-reference quality metrics on high dynamic range (HDR) video content. arXiv:1803.04815. Retrieved from <https://arxiv.org/abs/1803.04815>
- [9] Francesco Banterle, Alessandro Artusi, Kurt Debattista, and Alan Chalmers. 2017. *Advanced High Dynamic Range Imaging* (2nd ed). AK Peters (CRC Press), Natick, MA.
- [10] Francesco Banterle, Alessandro Artusi, Alejandro Moreo, Fabio Carrara, and Paolo Cignoni. 2023. NoR-VDPNet++: Real-time no-reference image quality metrics. *IEEE Access* 11 (2023), 34544–34553. DOI: <https://doi.org/10.1109/ACCESS.2023.3263496>
- [11] Francesco Banterle, Alessandro Artusi, Elena Sikudova, Thomas Bashford-Rogers, Patrick Ledda, Marina Bloj, and Alan Chalmers. 2012. Dynamic range compression by differential zone mapping based on psychophysical experiments. In *Proceedings of the ACM Symposium on Applied Perception (SAP '12)*. ACM, New York, NY, 39–46. DOI: <https://doi.org/10.1145/2338676.2338685>
- [12] Francesco Banterle, Patrick Ledda, Kurt Debattista, Marina Bloj, Alessandro Artusi, and Alan Chalmers. 2009. A psychophysical evaluation of inverse tone mapping techniques. *Computer Graphics Forum* 28, 1 (2009), 13–25.
- [13] Neil D. B. Bruce. 2014. ExpoBlend: Information preserving exposure blending based on normalized log-domain entropy. *Computers and Graphics* 39 (2014), 12–23. DOI: <https://doi.org/10.1016/j.cag.2013.10.001>
- [14] Kenneth Chiu, Michael Herf, Peter Shirley, S. Swamy, C. Wang, and K. Zimmerman. 1993. Spatially nonuniform scaling functions for high contrast images. In *Proceedings of Graphics Interface '93 (GI '93)*. Canadian Human-Computer Communications Society, Toronto, Ontario, Canada, 245–253. Retrieved from <http://graphicsinterface.org/wp-content/uploads/gi1993-34.pdf>
- [15] H. A. David. 1988. *The Method of Paired Comparisons* (2nd ed). Oxford University Press, Oxford University Press, Northamptonshire, UK.
- [16] Frederic Drago, Karol Myszkowski, Thomas Annen, and Norishige Chiba. 2003. Adaptive logarithmic mapping for displaying high contrast scenes. *Computer Graphics Forum* 22, 3 (2003), 419–426. DOI: <https://doi.org/10.1111/1467-8659.00689>
- [17] Frédo Durand and Julie Dorsey. 2002. Fast bilateral filtering for the display of high-dynamic-range images. *ACM Transactions on Graphics* 21, 3 (July 2002), 257–266. DOI: <https://doi.org/10.1145/566654.566574>
- [18] Gabriel Eilertsen, Joel Kronander, Gyorgy Denes, Rafał K. Mantiuk, and Jonas Unger. 2017. HDR image reconstruction from a single exposure using deep CNNs. *ACM Transactions on Graphics* 36, 6 (November 2017), Article 178, 1–15. DOI: <https://doi.org/10.1145/3130800.3130816>
- [19] Mark D. Fairchild. 2007. The HDR photographic survey. In *International Conference on Communications in Computing*. IS&T Publications, 233–238.

- [20] Raanan Fattal, Dani Lischinski, and Michael Werman. 2002. 2002. Gradient domain high dynamic range compression. *ACM Transactions on Graphics* 21, 3 (July 2002), 249–256. DOI : <https://doi.org/10.1145/566654.566573>
- [21] James A. Ferwerda, Sumanta N. Pattanaik, Peter Shirley, and Donald P. Greenberg. 1996. A model of visual adaptation for realistic image synthesis. In *Proceedings of the 23rd Annual Conference on Computer Graphics and Interactive Techniques (SIGGRAPH '96)*. ACM, New York, NY, 249–258. DOI : <https://doi.org/10.1145/237170.237262>
- [22] Jan Froehlich, Stefan Grandinetti, Bernd Eberhardt, Simon Walter, Andreas Schilling, and Harald Brendel. 2014. Creating cinematic wide gamut HDR-video for the evaluation of tone mapping operators and HDR-displays. In *Digital Photography X*. Nitin Sampat, Radka Tezaur, Sebastiano Battiato, and Boyd A. Fowler (Eds.), Vol. 9023, International Society for Optics and Photonics, SPIE, 90230X. DOI : <https://doi.org/10.1117/12.2040003>
- [23] Marc-André Gardner, Kalyan Sunkavalli, Ersin Yumer, Xiaohui Shen, Emiliano Gambaretto, Christian Gagné, and Jean-François Lalonde. 2017. Learning to predict indoor illumination from a single image. *ACM Transactions on Graphics* 36, 6 (November 2017), Article 176, 1–14. DOI : <https://doi.org/10.1145/3130800.3130891>
- [24] Michaël Gharbi, Jiawen Chen, Jonathan T. Barron, Samuel W. Hasinoff, and Frédo Durand. 2017. Deep bilateral learning for real-time image enhancement. *ACM Transactions on Graphics* 36, 4 (2017), 1–12.
- [25] Ian Goodfellow, Jean Pouget-Abadie, Mehdi Mirza, Bing Xu, David Warde-Farley, Sherjil Ozair, Aaron Courville, and Yoshua Bengio. 2020. Generative adversarial networks. *Communications of the ACM* 63, 11 (October 2020), 139–144. DOI : <https://doi.org/10.1145/3422622>
- [26] Samuel W. Hasinoff, Dillon Sharlet, Ryan Geiss, Andrew Adams, Jonathan T. Barron, Florian Kainz, Jiawen Chen, and Marc Levoy. 2016. Burst photography for high dynamic range and low-light imaging on mobile cameras. *ACM Transactions on Graphics* 35, 6 (November 2016), Article 192, 1–12. DOI : <https://doi.org/10.1145/2980179.2980254>
- [27] HDR Labs. 2013. sIBL Archive Free HDRI Sets for Smart Image-Based Lighting. Retrieved from <https://hdrlabs.com>
- [28] Yannick Hold-Geoffroy, Akshaya Athawale, and Jean-François Lalonde. 2019. Deep sky modeling for single image outdoor lighting estimation. In *IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*. IEEE, 6920–6928.
- [29] Yifei Huang, Sheng Qiu, Changbo Wang, and Chenhui Li. 2021. Learning representations for high-dynamic-range image color transfer in a self-supervised way. *IEEE Transactions on Multimedia* 23 (2021), 176–188. DOI : <https://doi.org/10.1109/TMM.2020.2981994>
- [30] Phillip Isola, Jun-Yan Zhu, Tinghui Zhou, and Alexei A. Efros. 2017. Image-to-image translation with conditional adversarial networks. In *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 5967–5976. DOI : <https://doi.org/10.1109/CVPR.2017.632>
- [31] Nima Khademi Kalantari and Ravi Ramamoorthi. 2017. Deep high dynamic range imaging of dynamic scenes. *ACM Transactions on Graphics* 36, 4 (July 2017), Article 144, 1–12. DOI : <https://doi.org/10.1145/3072959.3073609>
- [32] Min H. Kim and Jan Kautz. 2008. Consistent tone reproduction. In *Proceedings of the 10th IASTED International Conference on Computer Graphics and Imaging (CGIM '08)*. ACTA Press, 152–159.
- [33] Grzegorz Krawczyk, Karol Myszkowski, and Hans-Peter Seidel. 2005. Lightness perception in tone reproduction for high dynamic range images. *Computer Graphics Forum* 24, 3 (2005), 635–645. DOI : <https://doi.org/10.1111/j.1467-8659.2005.00888.x>
- [34] Jiangtao Kuang, Garrett M. Johnson, and Mark D. Fairchild. 2007. iCAM06: A refined image appearance model for HDR image rendering. *Journal of Visual Communication and Image Representation* 18, 5 (2007), 406–414. DOI : <https://doi.org/10.1016/j.jvcir.2007.06.003>
- [35] Patrick Ledda, Alan Chalmers, Tom Troscianko, and Helge Seetzen. 2005. Evaluation of tone mapping operators using a high dynamic range display. *ACM Transactions on Graphics* 24, 3 (2005), 640–648.
- [36] Yuanzhen Li, Lavanya Sharan, and Edward H. Adelson. 2005. Compressing and companding high dynamic range images with subband architectures. *ACM Transactions on Graphics* 24, 3 (July 2005), 836–844. DOI : <https://doi.org/10.1145/1073204.1073271>
- [37] Dani Lischinski, Zeev Farbman, Matt Uyttendaele, and Richard Szeliski. 2006. Interactive local adjustment of tonal values. *ACM Transactions on Graphics* 25, 3 (July 2006), 646–653. DOI : <https://doi.org/10.1145/1141911.1141936>
- [38] Yu-Lun Liu, Wei-Sheng Lai, Yu-Sheng Chen, Yi-Lung Kao, Ming-Hsuan Yang, Yung-Yu Chuang, and Jia-Bin Huang. 2020. Single-image HDR reconstruction by learning to reverse the camera pipeline. In *2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*. IEEE, 1648–1657. DOI : <https://doi.org/10.1109/CVPR42600.2020.00172>
- [39] R. Mantiuk, R. Mantiuk, A. Tomaszewska, and W. Heidrich. 2009. Color correction for tone mapping. *Computer Graphics Forum* 28, 2 (2009), 193–202. DOI : <https://doi.org/10.1111/j.1467-8659.2009.01358.x>
- [40] Rafal K. Mantiuk, Param Hanji, Maliha Ashraf, Yuta Asano, and Alexandre Chapiro. 2024. ColorVideoVDP: A visual difference predictor for image, video and display distortions. *ACM Transactions on Graphics* 43, 4 (July 2024), Article 129, 1–20. DOI : <https://doi.org/10.1145/3658144>
- [41] Xudong Mao, Qing Li, Haoran Xie, Raymond Y. K. Lau, Zhen Wang, and Stephen Paul Smolley. 2017. Least squares generative adversarial networks. In *Proceedings of the IEEE International Conference on Computer Vision*, 2794–2802.

- [42] Demetris Marnerides, Thomas Bashford-Rogers, Jonathan Hatchett, and Kurt Debattista. 2018. ExpandNet: A deep convolutional neural network for high dynamic range expansion from low dynamic range content. In *Computer Graphics Forum*. Vol. 37. Wiley Online Library, 37–49. DOI: <https://doi.org/10.1111/cgf.13340>
- [43] Tom Mertens, Jan Kautz, and Frank Van Reeth. 2007. Exposure fusion. In *15th Pacific Conference on Computer Graphics and Applications (PG '07)*, 382–390. DOI: <https://doi.org/10.1109/PG.2007.17>
- [44] Xenios Milidonis, Francesco Banterle, and Alessandro Artusi. 2024. Deep chroma compression of tone-mapped images. arXiv:2409.16032. Retrieved from <https://arxiv.org/abs/2409.16032>
- [45] Kenta Moriwaki, Ryota Yoshihashi, Rei Kawakami, Shaodi You, and Takeshi Naemura. 2018. Hybrid loss for learning single-image-based HDR reconstruction. arXiv:1812.07134. Retrieved from <https://arxiv.org/abs/1812.07134>
- [46] Hiromi Nemoto, Pavel Korshunov, Philippe Hanhart, and Touradj Ebrahimi. 2015. Visual attention in LDR and HDR images. In *9th International Workshop on Video Processing and Quality Metrics for Consumer Electronics (VPQM)*. Springer-Verlag. Retrieved from <http://infoscience.epfl.ch/record/203873>
- [47] Yuzhen Niu, Jianbin Wu, Wenxi Liu, Wenzhong Guo, and Rynson W. H. Lau. 2021. HDR-GAN: HDR image reconstruction from multi-exposed LDR images with large motions. *IEEE Transactions on Image Processing* 30 (2021), 3885–3896. DOI: <https://doi.org/10.1109/tip.2021.3064433>
- [48] Yafei Ou, Prasoon Ambalathankandy, Shinya Takamaeda, Masato Motomura, Tetsuya Asai, and Masayuki Ikebe. 2022. Real-time tone mapping: A survey and cross-implementation hardware benchmark. *IEEE Transactions on Circuits and Systems for Video Technology* 32, 5 (2022), 2666–2686. DOI: <https://doi.org/10.1109/TCSVT.2021.3060143>
- [49] Vaibhav Amit Patel, Purvik Shah, and Shanmuganathan Raman. 2018. A generative adversarial network for tone mapping HDR images. In *Computer Vision, Pattern Recognition, Image Processing, and Graphics*. Renu Rameshan, Chetan Arora, and Sumantra Dutta Roy (Eds.), Springer Singapore, Singapore, 220–231.
- [50] Sumanta N. Pattanaik, James A. Ferwerda, Mark D. Fairchild, and Donald P. Greenberg. 1998. A multiscale model of adaptation and spatial vision for realistic image display. In *Proceedings of the 25th Annual Conference on Computer Graphics and Interactive Techniques (SIGGRAPH '98)*. ACM, New York, NY, 287–298. DOI: <https://doi.org/10.1145/280814.280922>
- [51] Tania Pouli, Alessandro Artusi, Francesco Banterle, Ahmet Oğuz Akyüz, Hans-Peter Seidel, and Erik Reinhard. 2013. Color correction for tone reproduction. In *CIC21: 21st Color and Imaging Conference*. Society for Imaging Science and Technology (IS&T), 215–220. Retrieved from <http://vcg.isti.cnr.it/Publications/2013/PABASR13>
- [52] Shanmuganathan Raman and Subhasis Chaudhuri. 2009. Bilateral filter based compositing for variable exposure photography. In *Eurographics 2009—Short Papers*, P. Alliez and M. Magnor (Eds.), The Eurographics Association. DOI: <https://doi.org/10.2312/egs.20091034>
- [53] Erik Reinhard and Kate Devlin. 2005. Dynamic range reduction inspired by photoreceptor physiology. *IEEE Transactions on Visualization and Computer Graphics* 11, 1 (January 2005), 13–24. DOI: <https://doi.org/10.1109/TVCG.2005.9>
- [54] Erik Reinhard, Michael Stark, Peter Shirley, and James Ferwerda. 2002. Photographic tone reproduction for digital images. *ACM Transactions on Graphics* 21, 3 (July 2002), 267–276. DOI: <https://doi.org/10.1145/566654.566575>
- [55] E. Riba, D. Mishkin, D. Ponsa, E. Rublee, and G. Bradski. 2020. Kornia: An open source differentiable computer vision library for PyTorch. In *Winter Conference on Applications of Computer Vision*. Retrieved from <https://arxiv.org/pdf/1910.02190.pdf>
- [56] Marcel Santana Santos, Tsang Ing Ren, and Nima Khademi Kalantari. 2020. Single image HDR reconstruction using a CNN with masked features and perceptual loss. *ACM Transactions on Graphics* 39, 4 (August 2020), Article 80, 1–10. DOI: <https://doi.org/10.1145/3386569.3392403>
- [57] Christophe Schlick. 1995. Quantization techniques for visualization of high dynamic range pictures. In *Photorealistic Rendering Techniques*. Georgios Sakas, Stefan Müller, and Peter Shirley (Eds.), Springer Berlin, Berlin, 7–20.
- [58] Gaurav Sharma. 2003. *Digital Color Imaging: Handbook*. CRC Press.
- [59] Elena Šikudová, Tania Pouli, Alessandro Artusi, Ahmet Oguz Akyüz, Francesco Banterle, Zeynep Miray Mazlumoglu, and Erik Reinhard. 2016. A gamut-mapping framework for color-accurate reproduction of HDR images. *IEEE Computer Graphics and Applications* 36, 4 (2016), 78–90. DOI: <https://doi.org/10.1109/MCG.2015.116>
- [60] Michael D. Tocci, Chris Kiser, Nora Tocci, and Pradeep Sen. 2011. A versatile HDR video production system. *ACM Transactions on Graphics* 30, 4 (2011), 1–10. DOI: <https://doi.org/10.1145/2010324.1964936>
- [61] Jack Tumblin, Jessica K. Hodgins, and Brian K. Guenter. 1999. Two methods for display of high contrast images. *ACM Transactions on Graphics* 18, 1 (January 1999), 56–94. DOI: <https://doi.org/10.1145/300776.300783>
- [62] Dmitry Ulyanov, Andrea Vedaldi, and Victor Lempitsky. 2016. Instance normalization: The missing ingredient for fast stylization. arXiv:1607.08022. Retrieved from <https://arxiv.org/abs/1607.08022>
- [63] Johannes Hendrik van Hateren. 2006. Encoding of high dynamic range video with a model of human cones. *ACM Transactions on Graphics* 25, 4 (October 2006), 1380–1399. Retrieved from <http://www.rug.nl/informatica/organisatie/overorganisatie/iwi>

- [64] Lin Wang and Kuk-Jin Yoon. 2021. Deep learning for HDR imaging: State-of-the-art and future trends. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 44, 12 (2021), 8874–8895.
- [65] Ting-Chun Wang, Ming-Yu Liu, Jun-Yan Zhu, Andrew Tao, Jan Kautz, and Bryan Catanzaro. 2018. High-resolution image synthesis and semantic manipulation with conditional GANs. In *2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 8798–8807. DOI: <https://doi.org/10.1109/CVPR.2018.00917>
- [66] Greg Ward. 1994. A contrast-based scalefactor for luminance display. In *Graphics Gems*. Paul S. Heckbert (Ed.), Elsevier, 415–421. DOI: <https://doi.org/10.1016/b978-0-12-336156-1.50054-9>
- [67] Gregory Ward Larson, Holly Rushmeier, and Christine Piatko. 1997. A visibility matching tone reproduction operator for high dynamic range scenes. In *ACM SIGGRAPH 97 Visual Proceedings: The Art and Interdisciplinary Programs of SIGGRAPH '97 (SIGGRAPH '97)*. Association for Computing Machinery, New York, NY, 155. DOI: <https://doi.org/10.1145/259081.259242>
- [68] Feng Xiao, Jeffrey M. DiCarlo, Peter B. Catrysse, and Brian A. Wandell. 2002. High dynamic range imaging of natural scenes. In *The 10th Color Imaging Conference: Color Science and Engineering Systems, Technologies, Applications (CIC '02)*. Society for Imaging Science and Technology, 337–342. DOI: <https://doi.org/10.2352/CIC.2002.10.1.ART00062>
- [69] Zhiguang Yang, Youping Chen, Zhuliang Le, and Yong Ma. 2020. GANFuse: A novel multi-exposure image fusion method based on generative adversarial networks. *Neural Computing and Applications* 33, 11 (2020), 6133–6145.
- [70] Yangli Hector Yee and Sumanta N. Pattanaik. 2003. Segmentation and adaptive assimilation for detail-preserving display of high-dynamic range images. *The Visual Computer* 19, 7-8 (2003), 457–466.
- [71] Hang Zhao, Orazio Gallo, Iuri Frosio, and Jan Kautz. 2017. Loss functions for image restoration with neural networks. *IEEE Transactions on Computational Imaging* 3, 1 (2017), 47–57. DOI: <https://doi.org/10.1109/TCI.2016.2644865>
- [72] Jun-Yan Zhu, Taesung Park, Phillip Isola, and Alexei A. Efros. 2017. Unpaired image-to-image translation using cycle-consistent adversarial networks. In *2017 IEEE International Conference on Computer Vision (ICCV)*, 2242–2251. DOI: <https://doi.org/10.1109/ICCV.2017.244>

Received 6 February 2025; revised 8 May 2025; accepted 12 June 2025