DOI: 10.54254/2755-2721/138/2025.21531

Fusing Multiple Exposure Images for HDR Images by Deep Learning

Longyao Wu[1,a,*]

[1]Ulster College, Shanxi University of Science & Technology, Xian, China
a. Wu_Longyao@outlook.com
**corresponding author*

Abstract: This paper explores the application of deep learning techniques in the fusion of high dynamic range (HDR) images, emphasizing its transformative impact on traditional HDR imaging methods. HDR images are renowned for capturing a broader range of luminosity; however, traditional methods face challenges such as camera shake and ghosting in dynamic scenes. The introduction of deep learning has automated and enhanced the HDR image generation process, particularly in image fusion, deblurring, and artifact correction. This paper reviews relevant deep learning algorithms and architectures, analyzes the strengths and limitations of current HDR imaging approaches, and suggests future research directions aimed at improving efficiency, accuracy, and applicability across various domains.

Keywords: High Dynamic Range (HDR), Deep Learning, Image Fusion, Neural Networks, Ghosting Correction

1. Introduction

High dynamic range (HDR) images are renowned for their ability to capture a broader range of luminosity, providing viewers with a richer and more immersive visual experience [1]. Consequently, achieving high-quality HDR images has become a focal point in the photographic industry.

Traditionally, photographers captured multiple frames of the same scene with varying exposures to preserve details in both the darkest and brightest areas. Typically, three photos are taken—one with low exposure, one with standard exposure, and one with high exposure—and then merged into an HDR image using professional image editing software. However, this process is not without challenges. Photographers must address issues like camera shake or moving objects, which can cause ghosting artifacts in the final image. Additionally, HDR images often exceed the display capabilities of standard monitors, requiring manual tone mapping to produce a visually appealing result [2-4]. These complex post-production steps demand a high level of skill and experience, making the process daunting for beginners.

The advent of deep learning has revolutionized HDR image production, fundamentally transforming the way these images are created and refined. Researchers are increasingly utilizing deep learning techniques to automate and enhance various stages of the HDR imaging workflow, including exposure fusion, deblurring, and the correction of common artifacts such as ghosting. By training models on suitable image datasets, deep learning algorithms can predict optimal exposure settings, and seamlessly fuse multiple images with varying exposures [5-14]. These advancements are democratizing HDR imaging, making it more accessible not only to professional photographers but also to enthusiasts and casual users who may lack the technical expertise required for traditional HDR methods. With continuous progress in this field, HDR image production is no longer reliant solely on manual editing. Instead, it benefits from the powerful computational capabilities of modern computers, streamlining the process and allowing for the creation of high-quality HDR images with greater ease.

This paper examines recent advancements in HDR image fusion methods driven by deep learning, highlighting key algorithms, architectures, and their contributions to the field. It also provides a critical analysis of the current state of deep learning in HDR imaging, identifying both the strengths and limitations of existing approaches. Finally, the paper offers recommendations for future research directions, with a focus on enhancing the efficiency, accuracy, and applicability of deep learning techniques in HDR image production across various domains.

2. Introduction to deep learning for HDR imaging

2.1. Traditional Techniques for Creating HDR Images

Exposure bracketing is a widely used technique that allows photographers to capture multiple frames of the same scene. This approach addresses the limitations of capturing the full dynamic range of a scene in a single frame, a challenge that the HDR fusion process is designed to overcome. To tackle these issues, Debevec & Malik (1997) and Drago et al. (2003) utilized this method to produce HDR images [15, 16].

However, HDR images often exceed the capabilities of standard display screens, requiring an additional processing step known as tone mapping to be properly displayed [17]. Traditional HDR generation methods typically rely on fixed algorithms or offer only a limited set of manually adjustable parameters, like [15]. These constraints can make it difficult to achieve optimal results, particularly in complex scenarios where a one-size-fits-all approach may not suffice. The lack of flexibility in traditional techniques can hinder the creation of customized, high-quality HDR images in more demanding situations.

2.2. Advantages of Deep Learning in HDR Imaging

The advent of deep learning, particularly the introduction of Convolutional Neural Networks (CNNs), has significantly advanced the field of high dynamic range (HDR) imaging. Kalantari and Ramamoorthi [5] were among the first to demonstrate the potential of CNNs for HDR image generation, especially in addressing the notorious issue of ghosting in dynamic scenes. Ghosting artifacts arise when there is motion between the different exposure frames, causing misalignment and blending issues in traditional HDR techniques. By leveraging CNNs, Kalantari and Ramamoorthi were able to develop a method that could intelligently predict and correct these misalignments, effectively eliminating ghosting and enabling HDR image generation even in complex, motion-filled scenes. This innovation was a major breakthrough in the field, allowing for more robust and reliable HDR processing under dynamic conditions.

Many subsequent studies [6-14, 23-25, 27-33] have built upon this foundational work [5]. In addition to benefiting from these advancements, more recent research has focused on reducing the artifacts that were prevalent in earlier methods. A key challenge in this regard is the misalignment of input photos. Optical flow has emerged as a popular technique for aligning images, with its conceptual foundations dating back to 1981 [18]. By the time of [19], a more mature optical flow algorithm had become practical for multi-frame HDR image generation.

DOI: 10.54254/2755-2721/138/2025.21531

3. Deep Learning Techniques for Combining Multiple Exposures

3.1. Neural Network Architectures for HDR Imaging

As the first work to use CNNs for HDR image generation [5], this approach marked a significant boost for the HDR imaging industry. CNNs allow networks to extract more abstract features from images, enabling more customized frame fusion, which helped researchers handle dynamic scenes more effectively. With the rise of GANs [20], many fields, including HDR imaging, began to adopt this novel framework. As a pioneer, [12] introduced adversarial networks to HDR imaging, allowing models to better reconstruct missing information in overexposed or underexposed areas and a well-designed generator largely mitigated misalignment issues during the fusion process.

Since its debut, the Transformer architecture [21] has been a standout in the deep learning field, and in 2020, it was successfully applied to computer vision [22]. Two years later, [23] developed a transformer-based model specifically for HDR image generation. This approach allows the model to learn both local and global image features, overcoming many of the challenges that previous methods encountered.

3.2. Training Deep Learning Models for HDR

To train an effective HDR-generating algorithm, a high-quality dataset is essential. [5] was the first to introduce a real-world dataset consisting of three LDR images and one HDR ground truth label for the HDR deghosting task. Expanding on this, [24] developed a larger multi-frame fusion HDR dataset using a similar approach to that of Kalantari et al. (2017) [5]. [25] sourced HDR images from [26] which featured challenging scenes for HDR video and synthesized corresponding LDR images which ensured the central medium frame aligned with the HDR ground truth. [27] introduced a dataset with more diverse subjects, complex environmental conditions, and larger motion variations.

Depending on the specific requirements of HDR imaging, different loss functions may be needed. For simpler scenarios, mean square error (MSE) can be sufficient to produce pleasing results. In more complex cases, L2 loss might be used as a stronger constraint. However, a more common approach is to design a custom combined loss function to balance various constraints in a given situation, often resulting in high-quality outputs.

4. Qualitative and Quantitative Evaluation

4.1. Qualitative Evaluation

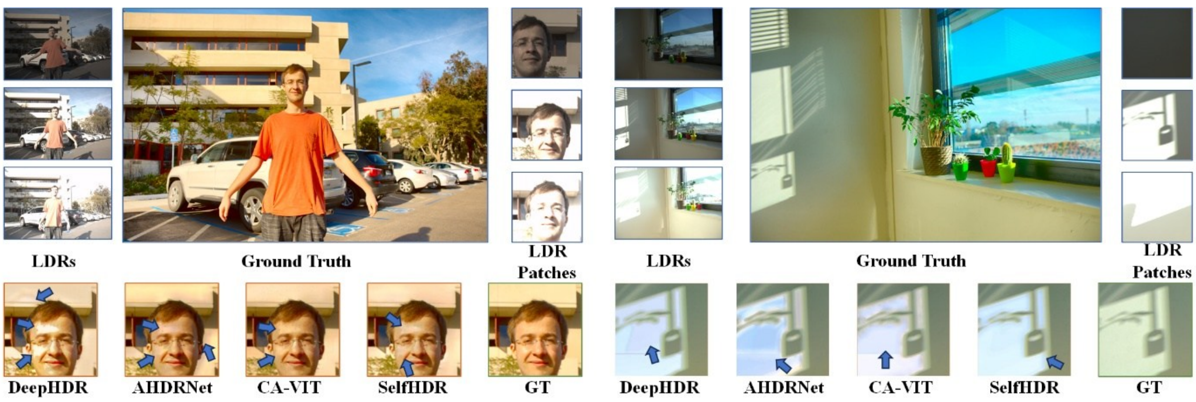

Figure 1: Visual representation of validation results for representative HDR imaging methods.

DOI: 10.54254/2755-2721/138/2025.21531

The bottom row in Figure 1 illustrates the progressive evolution of HDR imaging techniques. The first three methods—DeepHDR [34], AHDRNet [7], and CA-VIT [23]—represent significant milestones in HDR imaging research, while the final method, SelfHDR [28], showcases an impressive outcome achieved through self-supervised learning.

From these images, the advancements in the field become evident, particularly in minimizing ghosting artifacts and achieving visually realistic results. However, due to inherent limitations of real-world scenarios—such as overexposed areas, occluded objects, and other visual interferences—the true appearance of the scene cannot always be fully reconstructed. This limitation arises because even human observers cannot determine what lies behind these obstructions, reflecting the complexity of HDR imaging.

4.2. Quantitative Evaluation

Table 1 summarizes the quantitative evaluation results of several prominent HDR imaging methods. Performance metrics include PSNR (Peak Signal-to-Noise Ratio) and SSIM (Structural Similarity Index), computed in both tone-mapped (PSNR-μ, SSIM-μ) and linear (PSNR-l, SSIM-l) domains, compared to the ground truth. Additionally, HDR-DVP-2, a widely used metric for assessing human visual quality, is presented.

Table 1: Quantitative performance metrics for representative HDR imaging methods.

Method	PSNR-μ	SSIM-μ	PSNR-l	SSIM-l	HDR-DVP-2
DeepHDR	41.65	0.9860	40.88	0.9858	64.90
AHDRNet	43.63	0.9900	41.14	0.9702	64.61
CA-VIT	44.21	0.9918	42.17	0.9889	64.63
SelfHDR	43.95	0.9907	41.77	0.9889	64.77

5. Challenges and Future Directions

5.1. Practical Considerations

As HDR imaging has evolved, numerous methods have emerged to achieve high-quality results, some even claiming to be ghosting-free [7, 13, 29, 30]. Additionally, innovative learning strategies, such as self-supervised and few-shot learning [28, 31, 32], are gaining traction. However, there remains a critical need to enhance inference times, particularly for mobile devices. This paper suggests that finding a balance between speed and quality could be a promising direction for future research.

Moreover, recent advancements have introduced more complex architectures into low-level vision, paving the way for novel approaches. For instance, generative models like GANs and diffusion models have been successfully integrated into the HDR imaging process [12, 34]. These models excel in generating semantically coherent details by leveraging contextual information, and they allow for the incorporation of additional constraints to guide pixel generation.

5.2. Expanding HDR Imaging Beyond Photography

The demand for higher specifications in film and television is on the rise. Audiences increasingly seek immersive experiences that go beyond engaging storylines to include realistic visual elements. Fortunately, the HDR format is well-suited to deliver this enhanced visual enjoyment. The HDR imaging industry can capitalize on the rapid development of hardware, particularly for consumer-facing devices, and explore opportunities to integrate HDR formats with virtual reality (VR) and augmented reality (AR), which are gaining popularity in the context of Web 3.0.

6. Conclusion

In summary, this paper offers a comprehensive review of advancements in HDR imaging driven by deep learning techniques, highlighting the transformative impact of neural network architectures on the fusion of multiple exposure images. The evolution from traditional methods to contemporary deep learning approaches illustrates a significant leap in the quality and accessibility of HDR imaging, enabling both professionals and enthusiasts to produce stunning visuals with greater ease. Furthermore, the exploration of novel architectures, such as GANs and transformers, showcases their potential to address persistent challenges, including ghosting and alignment issues. Looking ahead, the integration of HDR imaging with emerging technologies such as virtual and augmented reality presents exciting opportunities for enhanced visual experiences across various domains, including film, gaming, and immersive storytelling. As the demand for high-quality visual content continues to rise, ongoing research aimed at improving efficiency, speed, and the applicability of these techniques will be crucial in shaping the future of HDR imaging. Ultimately, the fusion of deep learning and HDR imaging not only promises to elevate visual standards but also democratizes the art of photography, inviting a broader audience to engage with and appreciate the beauty of high dynamic range images.

References

[1] *McCann, J. J., & Rizzi, A. (2011). The art and science of HDR imaging. John Wiley & Sons.*
[2] *Debevec, P. E., & Malik, J. (2023). Recovering high dynamic range radiance maps from photographs. In Seminal Graphics Papers: Pushing the Boundaries, Volume 2 (pp. 643-652).*
[3] *Mann, S. (1993). Compositing multiple pictures of the same scene. In Proc. IS&T Annual Meeting, 1993 (pp. 50-52).*
[4] *Debevec, P., Reinhard, E., Ward, G., & Pattanaik, S. (2004). High dynamic range imaging. In ACM SIGGRAPH 2004 Course Notes (pp. 14-es).*
[5] *Kalantari, N. K., & Ramamoorthi, R. (2017). Deep high dynamic range imaging of dynamic scenes. ACM Trans. Graph., 36(4), 144-1.*
[6] *Wu, S., Xu, J., Tai, Y. W., & Tang, C. K. (2018). Deep high dynamic range imaging with large foreground motions. In Proceedings of the European Conference on Computer Vision (ECCV) (pp. 117-132).*
[7] *Yan, Q., Gong, D., Shi, Q., Hengel, A. V. D., Shen, C., Reid, I., & Zhang, Y. (2019). Attention-guided network for ghost-free high dynamic range imaging. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (pp. 1751-1760).*
[8] *Metwaly, K., & Monga, V. (2020, May). Attention-mask dense merger (attendense) deep hdr for ghost removal. In ICASSP 2020-2020 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP) (pp. 2623-2627). IEEE.*
[9] *Choi, S., Cho, J., Song, W., Choe, J., Yoo, J., & Sohn, K. (2020). Pyramid inter-attention for high dynamic range imaging. Sensors, 20(18), 5102.*
[10] *Yan, Q., Zhang, L., Liu, Y., Zhu, Y., Sun, J., Shi, Q., & Zhang, Y. (2020). Deep HDR imaging via a non-local network. IEEE Transactions on Image Processing, 29, 4308-4322.*
[11] *Ye, Q., Xiao, J., Lam, K. M., & Okatani, T. (2021, October). Progressive and selective fusion network for high dynamic range imaging. In Proceedings of the 29th ACM International Conference on Multimedia (pp. 5290-5297).*
[12] *Niu, Y., Wu, J., Liu, W., Guo, W., & Lau, R. W. (2021). Hdr-gan: Hdr image reconstruction from multi-exposed ldr images with large motions. IEEE Transactions on Image Processing, 30, 3885-3896.*
[13] *Liu, Z., Wang, Y., Zeng, B., & Liu, S. (2022, October). Ghost-free high dynamic range imaging with context-aware transformer. In European Conference on computer vision (pp. 344-360). Cham: Springer Nature Switzerland.*
[14] *Hu, T., Yan, Q., Qi, Y., & Zhang, Y. (2024). Generating content for hdr deghosting from frequency view. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (pp. 25732-25741).*
[15] *Debevec, P. E., & Malik, J. (1997) Recovering High Dynamic Range Radiance Maps from Photographs. In SIGGRAPH '97: Proceedings of the 24th annual conference on Computer graphics and interactive techniques. Association for Computing Machinery, Inc.*
[16] *Drago, F., Myszkowski, K., Annen, T., & Chiba, N. (2003, September). Adaptive logarithmic mapping for displaying high contrast scenes. In Computer graphics forum (Vol. 22, No. 3, pp. 419-426). Oxford, UK: Blackwell Publishing, Inc.*

[17] *Reinhard, E., Ward, G., Pattanaik, S., & Debevec, P. (2005). High Dynamic Range Imaging: Acquisition, Display, and Image-based Lighting, Chapter 2, Scetion 2.9: Gamma Display.*
[18] *Lucas, B. D., & Kanade, T. (1981, August). An iterative image registration technique with an application to stereo vision. In IJCAI'81: 7th international joint conference on Artificial intelligence (Vol. 2, pp. 674-679).*
[19] *Liu, C. (2009). Beyond pixels: exploring new representations and applications for motion analysis (Doctoral dissertation, Massachusetts Institute of Technology).*
[20] *Goodfellow, I., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., ... & Bengio, Y. (2020). Generative adversarial networks. Communications of the ACM, 63(11), 139-144.*
[21] *Vaswani, A. (2017). Attention is all you need. Advances in Neural Information Processing Systems.*
[22] *Dosovitskiy, A. (2020). An image is worth 16x16 words: Transformers for image recognition at scale. arXiv preprint arXiv:2010.11929.*
[23] *Liu, Z., Wang, Y., Zeng, B., & Liu, S. (2022, October). Ghost-free high dynamic range imaging with context-aware transformer. In European Conference on computer vision (pp. 344-360). Cham: Springer Nature Switzerland.*
[24] *Prabhakar, K. R., Arora, R., Swaminathan, A., Singh, K. P., & Babu, R. V. (2019, May). A fast, scalable, and reliable deghosting method for extreme exposure fusion. In 2019 IEEE International Conference on Computational Photography (ICCP) (pp. 1-8). IEEE.*
[25] *Pérez-Pellitero, E., Catley-Chandar, S., Leonardis, A., & Timofte, R. (2021). NTIRE 2021 challenge on high dynamic range imaging: Dataset, methods and results. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (pp. 691-700).*
[26] *Froehlich, J., Grandinetti, S., Eberhardt, B., Walter, S., Schilling, A., & Brendel, H. (2014, March). Creating cinematic wide gamut HDR-video for the evaluation of tone mapping operators and HDR-displays. In Digital photography X (Vol. 9023, pp. 279-288). SPIE.*
[27] *Tel, S., Wu, Z., Zhang, Y., Heyrman, B., Demonceaux, C., Timofte, R., & Ginhac, D. (2023). Alignment-free hdr deghosting with semantics consistent transformer. arXiv preprint arXiv:2305.18135.*
[28] *Zhang, Z., Wang, H., Liu, S., Wang, X., Lei, L., & Zuo, W. (2023). Self-supervised high dynamic range imaging with multi-exposure images in dynamic scenes. arXiv preprint arXiv:2310.01840.*
[29] *Song, J. W., Park, Y. I., Kong, K., Kwak, J., & Kang, S. J. (2022, October). Selective transhdr: Transformer-based selective hdr imaging using ghost region mask. In European Conference on Computer Vision (pp. 288-304). Cham: Springer Nature Switzerland.*
[30] *Yan, Q., Gong, D., Shi, J. Q., Van Den Hengel, A., Shen, C., Reid, I., & Zhang, Y. (2022). Dual-attention-guided network for ghost-free high dynamic range imaging. International Journal of Computer Vision, 1-19.*
[31] *Yan, Q., Zhang, S., Chen, W., Tang, H., Zhu, Y., Sun, J., ... & Zhang, Y. (2023). Smae: Few-shot learning for hdr deghosting with saturation-aware masked autoencoders. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (pp. 5775-5784).*
[32] *Prabhakar, K. R., Senthil, G., Agrawal, S., Babu, R. V., & Gorthi, R. K. S. S. (2021). Labeled from unlabeled: Exploiting unlabeled data for few-shot deep hdr deghosting. In Proceedings of the IEEE/CVF conference on computer vision and pattern recognition (pp. 4875-4885).*
[33] *Yan, Q., Hu, T., Sun, Y., Tang, H., Zhu, Y., Dong, W., ... & Zhang, Y. (2023). Towards high-quality hdr deghosting with conditional diffusion models. IEEE Transactions on Circuits and Systems for Video Technology.*
[34] *Wu, S., Xu, J., Tai, Y. W., & Tang, C. K. (2018). End-to-end deep HDR imaging with large foreground motions. In European Conference on Computer Vision (Vol. 1, No. 2).*