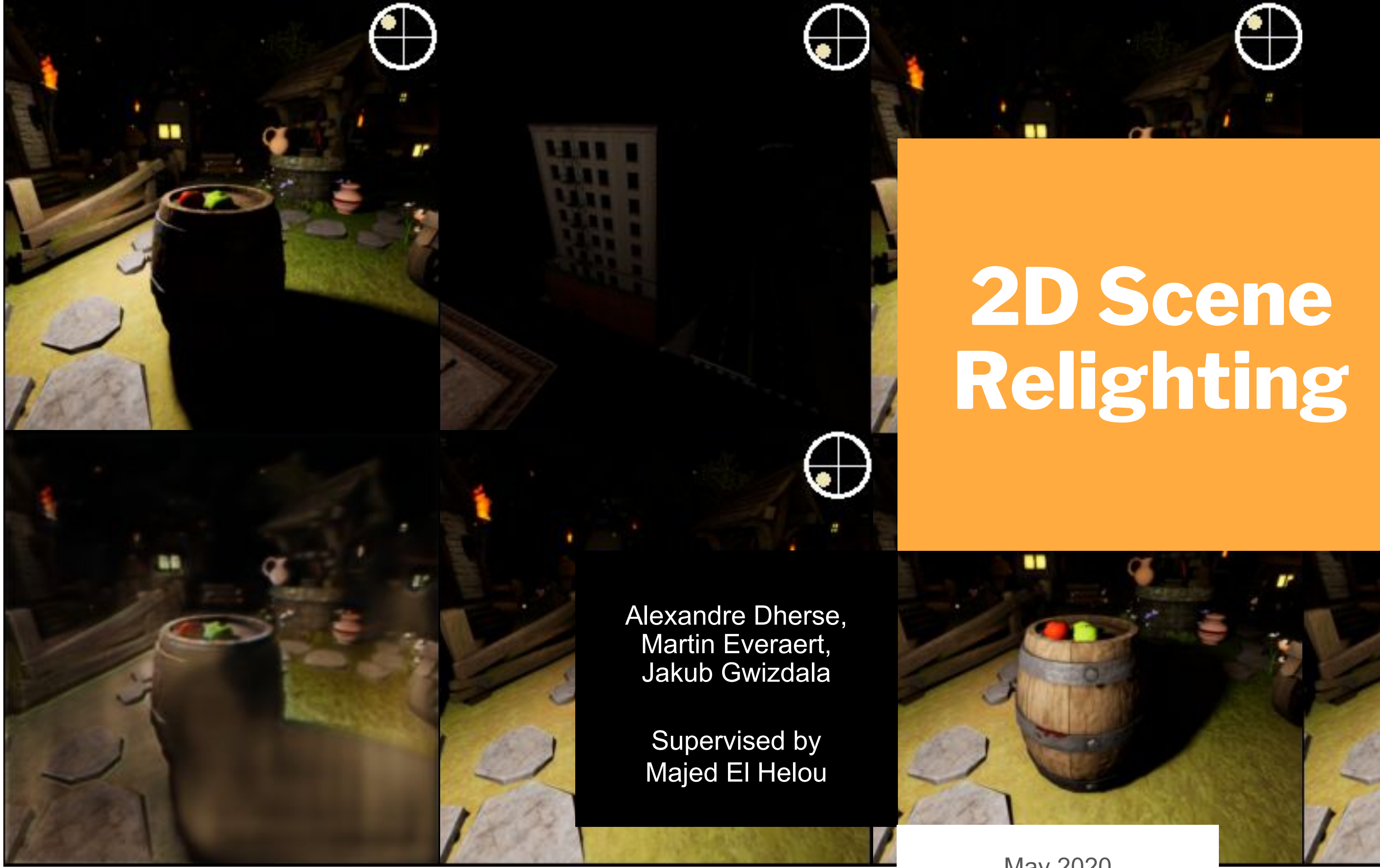


2D Scene Relighting

Alexandre Dherse,
Martin Everaert,
Jakub Gwizdala

Supervised by
Majed El Helou

May 2020

EPFL

CS-413
Computational
Photography

Introduction and motivation of the problem

Transfer the illumination from one image to another.

Having a method for general relighting of images with varied content

Alexandre Dherse, Martin Everaert, Jakub Gwizdala

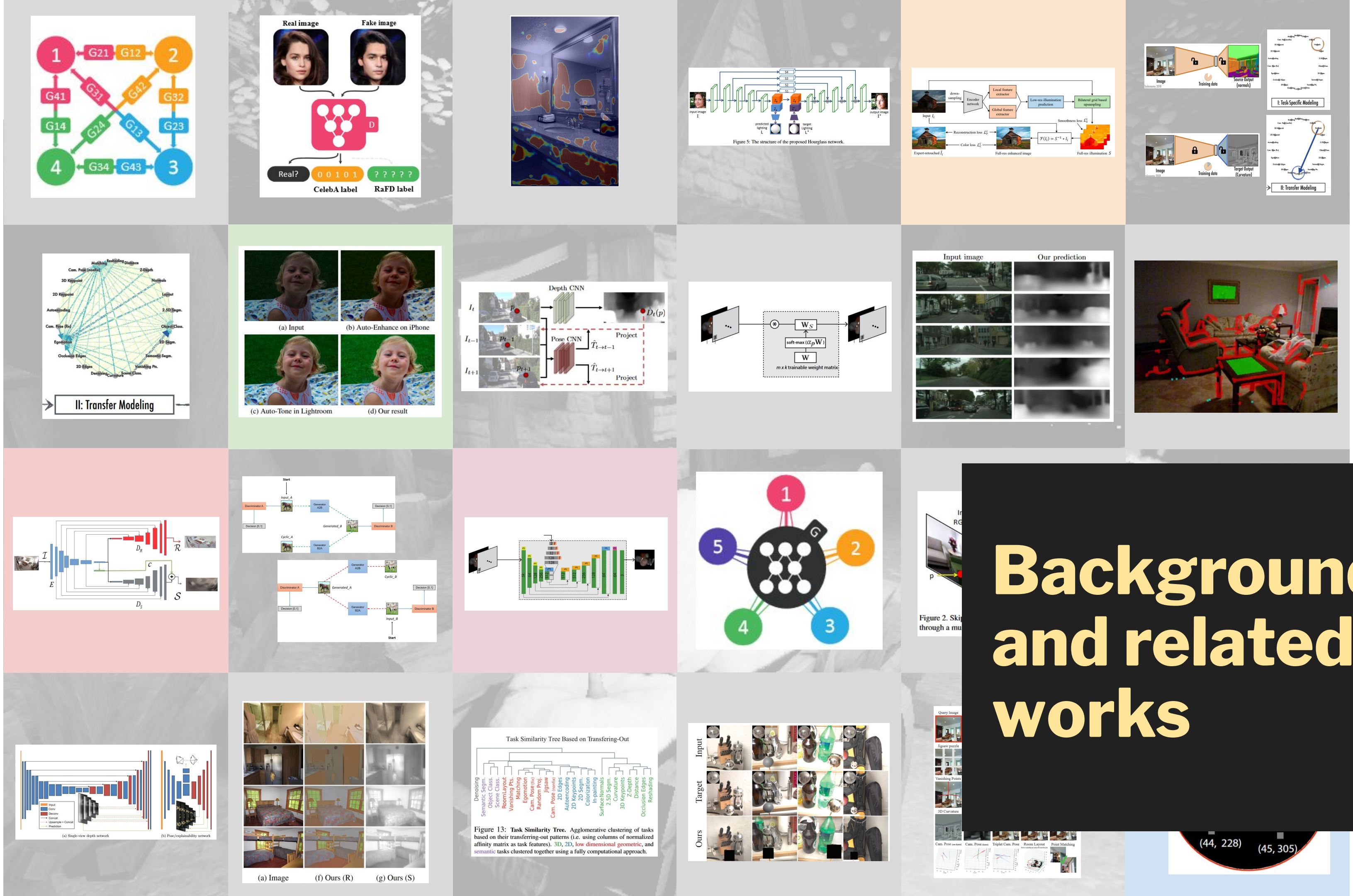
Introduction and motivation of the problem

- post-production tool for images (videos) to change original lighting (aesthetic reasons or eliminating the need of light setup equipment)

- perform data augmentation before training a network, allowing a wider variety of images for training (e.g. eliminate illumination bias)

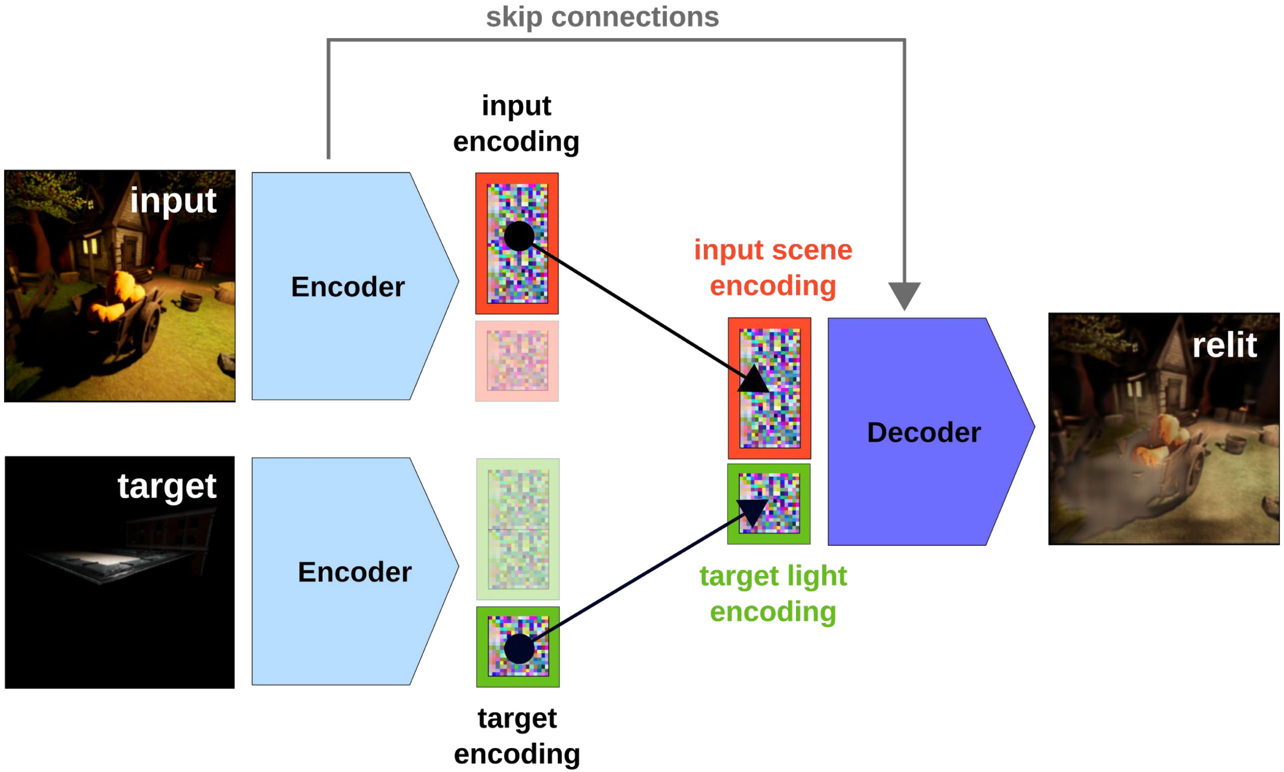
Alexandre Dherse, Martin Everaert, Jakub Gwizdala

Alexandre Dherse, Martin Everaert, Jakub Gwizdala



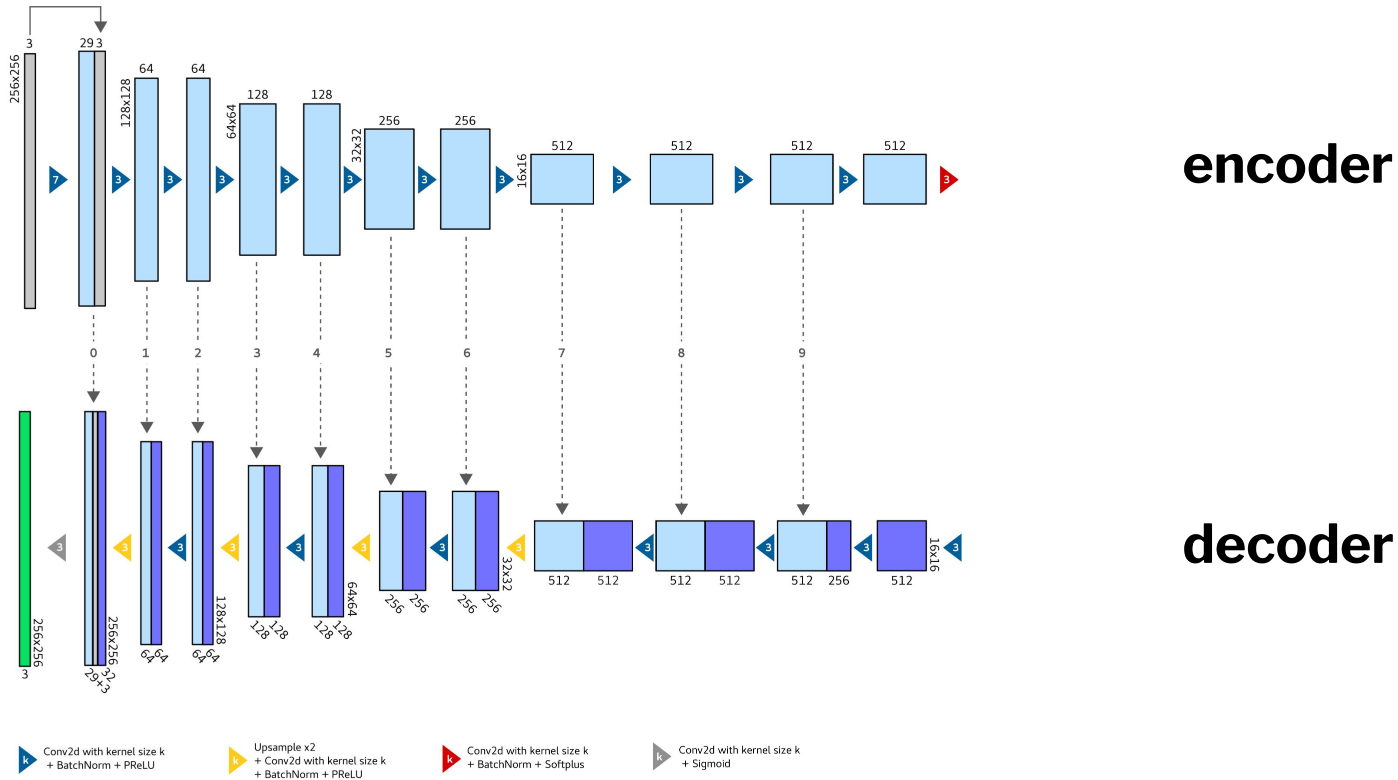
Background and related works

Overall architecture scheme



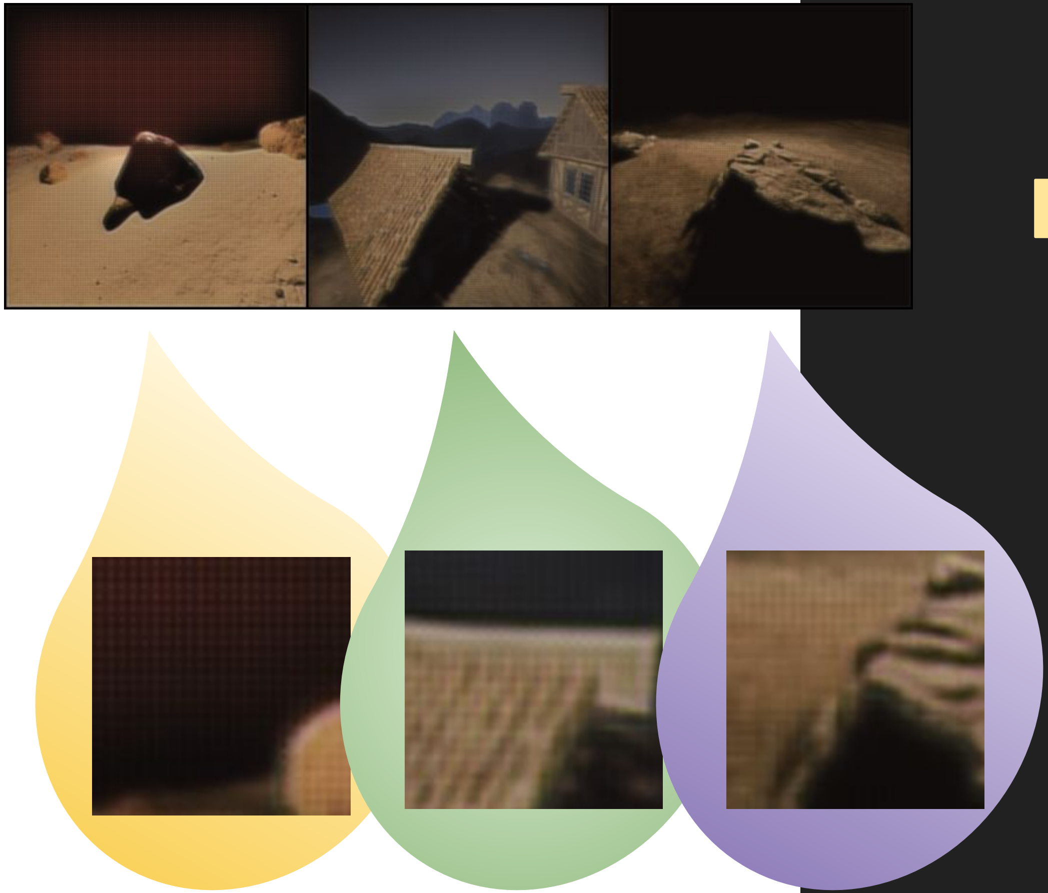
Alexandre Dherse, Martin Everaert, Jakub Gwizdala

Details of the solution @all



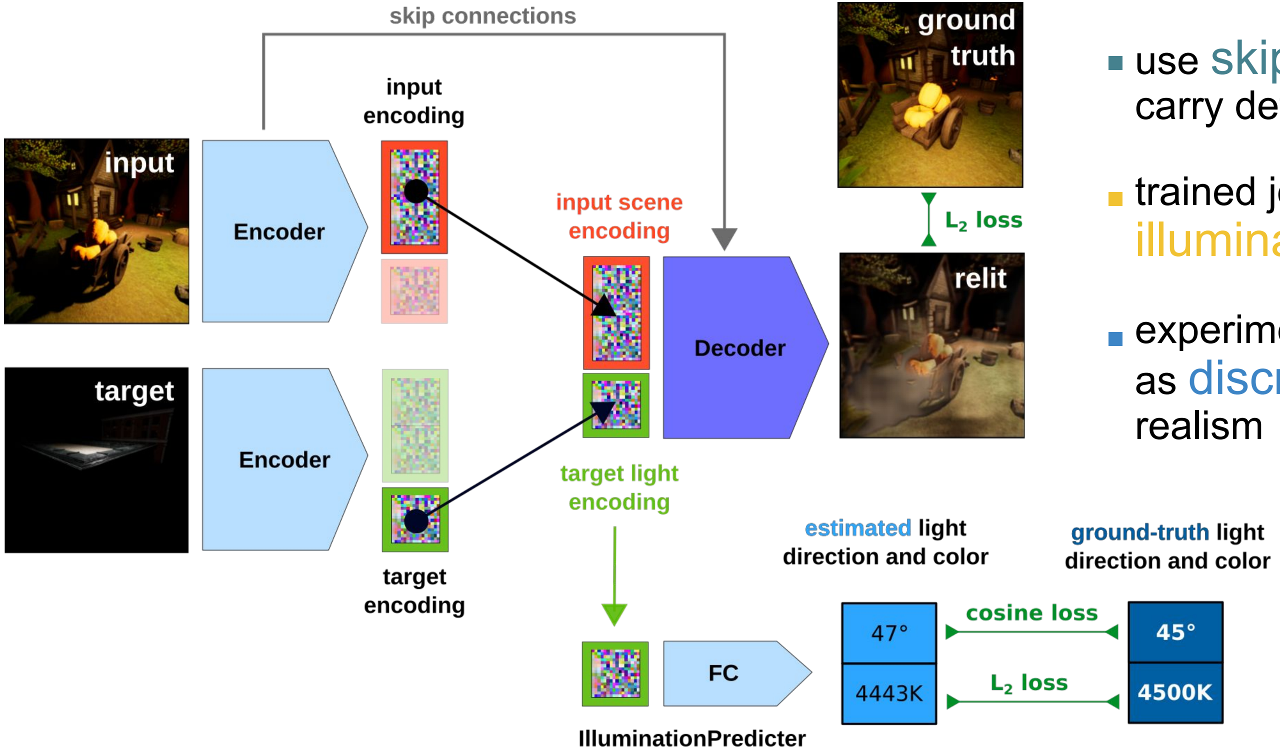
Alexandre Dherse, Martin Everaert, Jakub Gwizdala

Alexandre Dherse, Martin Everaert, Jakub Gwizdala



Implementation detail: checkerboard artifacts

Details of the solution @1

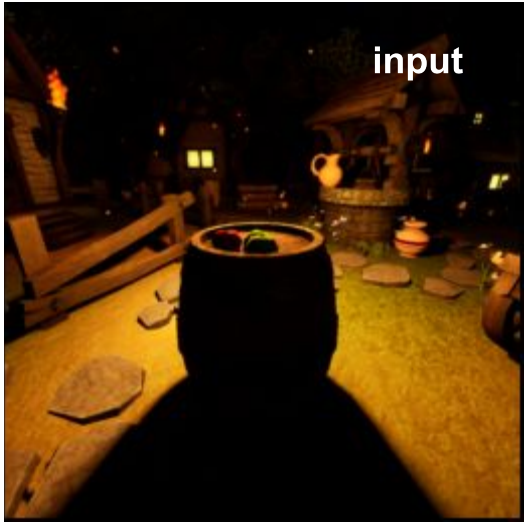


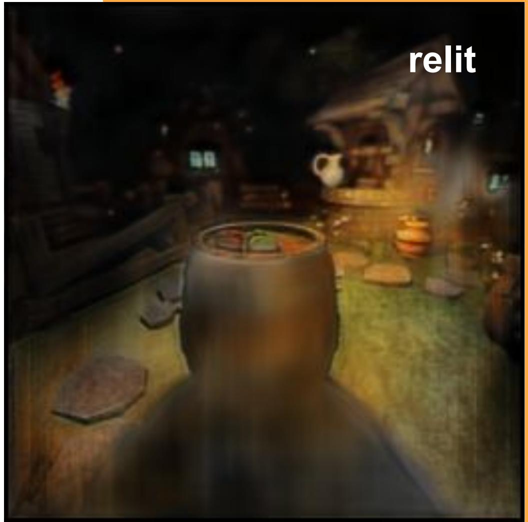
- split the latent variable into light and scene
- use skip-links from input to carry details
- trained jointly with an illumination predicter
- experiments with a PatchGAN as discriminator for more realism

Alexandre Dherse, Martin Everaert, Jakub Gwizdala

Experimental results @1 (eval w/o discriminator)

Alexandre Dherse, Martin Everaert, Jakub Gwizdala





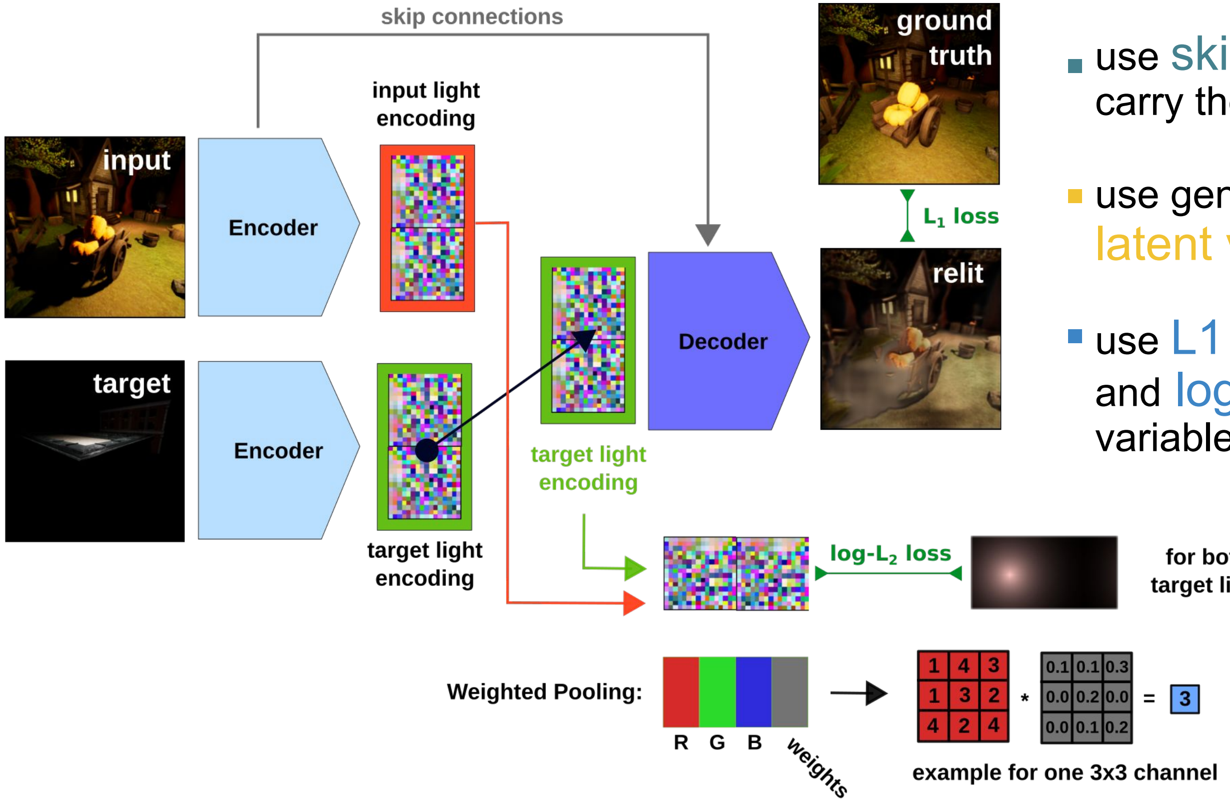
L_2 loss

Alexandre Dherse, Martin Everaert, Jakub Gwizdala

Experimental results @1 (eval w/ discriminator)

L_2 loss

Details of the solution @2

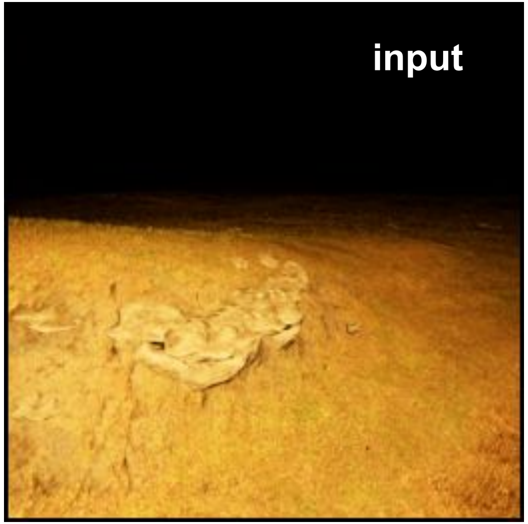


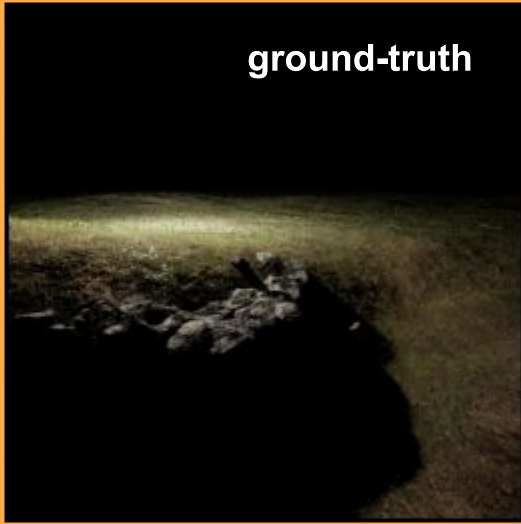
- latent space represents only light conditions
- use skip-links from input to carry the scene information
- use generated ground-truth latent variables
- use L1 reconstruction loss and log-L2 loss for latent variable

Alexandre Dherse, Martin Everaert, Jakub Gwizdala

Experimental results @2 (train)

Alexandre Dherse, Martin Everaert, Jakub Gwizdala





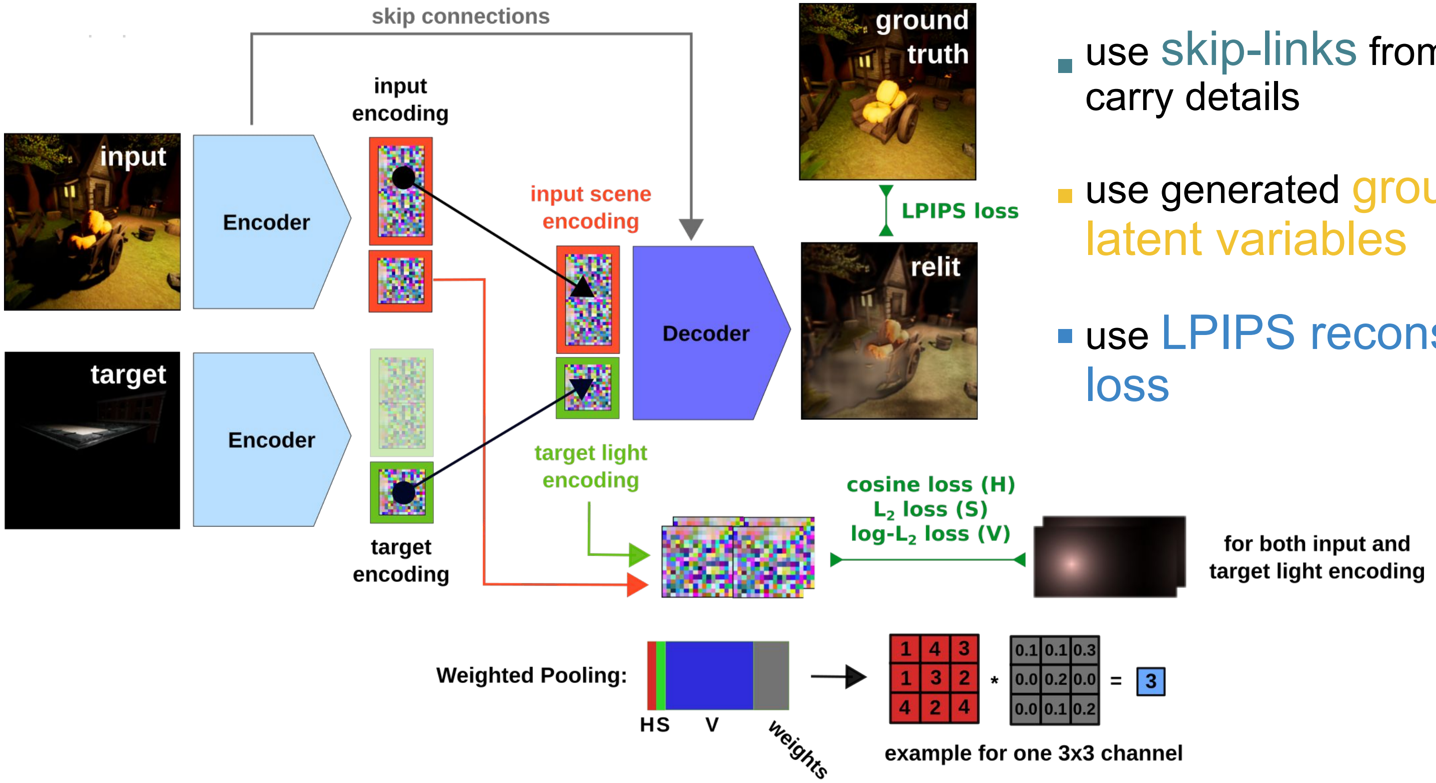
L_1 loss

Alexandre Dherse, Martin Everaert, Jakub Gwizdala

Experimental results @2 (eval)

L_1 loss

Details of the solution @3



- split the latent variable into light and scene
- use skip-links from input to carry details
- use generated ground-truth latent variables
- use LPIPS reconstruction loss

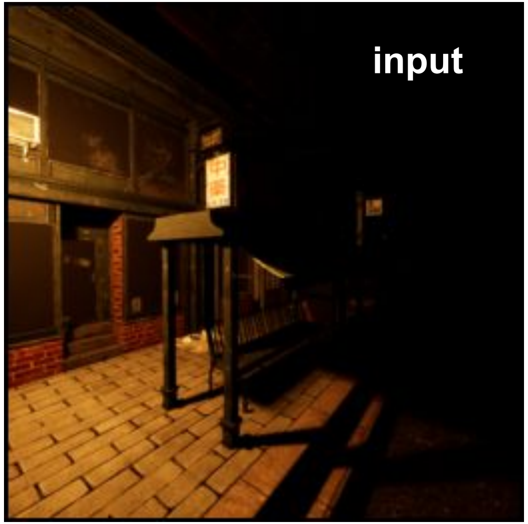
Alexandre Dherse, Martin Everaert, Jakub Gwizdala

Experimental results @3 (train)

LPIPS loss

Alexandre Dherse, Martin Everaert, Jakub Gwizdala

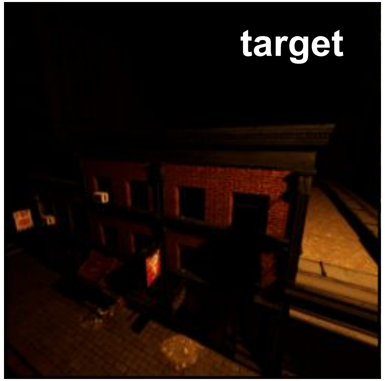
Experimental results @3 (eval)



LPIPS loss

EPFL

Experimental results @3 (eval)



LPIPS loss

Alexandre Dherse, Martin Everaert, Jakub Gwizdala

Experimental results

	@1	@2	@3
Metric name	IlluminationPredicter	Envmap	Envmap + scene
MSE	0.0238	**0.0219**	0.0254
SSIM	**0.3365**	0.1832	0.2988
PSNR	18.11 dB	**18.66** dB	18.10 dB
LPIPS	0.3268	0.2738	**0.2564**

Alexandre Dherse, Martin Everaert, Jakub Gwizdala

Pros and cons of the proposed method

- good results compared to identity mapping (relit = input) as well as for estimating light conditions (illumination predicter)
- color temperature much easier to estimate and transfer than light direction
- tendency to remove shadows instead of really change light direction (L_2 loss)
- poor realism

Alexandre Dherse, Martin Everaert, Jakub Gwizdala

Conclusion

Alexandre Dherse, Martin Everaert, Jakub Gwizdala

- experimenting with more variations: tune the network depth, latent variable size, regularization (as L_2-regularization), loss factors and functions (L_1,L_2, LPIPS, ...)

- for realism, conduct GAN experiments – e.g. use conditional GAN

EPFL

Thanks!

Special thanks to Majed El Helou, Ruofan Zhou and Sabine Süsstrunk for their supervision and for their help in this project

EPFL

Thanks!

Alexandre Dherse,
Martin Everaert,
Jakub Gwizdala

Supervised by
Majed El Helou

References

[1] I. Anokhin et al. "High-Resolution Daytime Translation Without Domain Labels". In: *arXiv preprint arXiv:2003.08791* (2020).

[2] S. Bell, K. Bala, and N. Snavely. "Intrinsic images in the wild". In: *ACM Transactions on Graphics (TOG)* 33.4 (2014), pp. 1–12.

[3] J. Bromley et al. "Signature verification using a "siamese" time delay neural network". In: *Advances in neural information processing systems*. 1994, pp. 737–744.

[4] Y. Choi et al. "Stargan: Unified generative adversarial networks for multi-domain image-to-image translation". In: *Proceedings of the IEEE conference on computer vision and pattern recognition*. 2018, pp. 8789–8797.

[5] F. Fleuret. *Transposed convolutions. Deep learning lecture handout for EE-559 at EPFL*. Accessed: 2020-05-26. URL: https://fleuret.org/ee559/materials/ee559-handout-7-1-transposed-convolutions.pdf.

[6] I. Goodfellow et al. "Generative adversarial nets". In: *Advances in neural information processing systems*. 2014, pp. 2672–2680.

[7] M. E. Helou et al. "VIDIT: Virtual Image Dataset for Illumination Transfer". In: *arXiv preprint arXiv:2005.05460* (2020).

[8] A. Hore and D. Ziou. "Image quality metrics: PSNR vs. SSIM". In: *2010 20th International Conference on Pattern Recognition*. IEEE. 2010, pp. 2366–2369.

References

[9] Y. Hu, B. Wang, and S. Lin. "Fc4: Fully convolutional color constancy with confidence-weighted pooling". In: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. 2017, pp. 4085–4094.

[10] P. Isola et al. "Image-to-Image Translation with Conditional Adversarial Networks". In: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. Code: https://github.com/junyanz/pytorch-CycleGAN-and-pix2pix. 2017, pp. 5967–5976.

[11] P. Isola et al. "Image-to-image translation with conditional adversarial networks". In: *Proceedings of the IEEE conference on computer vision and pattern recognition*. 2017, pp. 1125–1134.

[12] A. Jolicoeur-Martineau. "The relativistic discriminator: a key element missing from standard GAN". In: *arXiv preprint arXiv:1807.00734* (2018).

[13] B. Kovacs et al. "Shading annotations in the wild". In: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. 2017, pp. 6998–7007.

[14] Z. Li and N. Snavely. "Learning intrinsic image decomposition from watching the world". In: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. 2018, pp. 9039–9048.

[15] T. Mansencal et al. *Colour 0.3.15*. Version 0.3.15. Jan. 2020. DOI: 10.5281/zenodo.3627408. URL: https://doi.org/10.5281/zenodo.3627408.

References

[16] Martín Abadi et al. *TensorFlow: Large-Scale Machine Learning on Heterogeneous Systems*. Software available from tensorflow.org. 2015. URL: http://tensorflow.org/.

[17] M. Mirza and S. Osindero. "Conditional generative adversarial nets". In: *arXiv:1411.1784* (2014).

[18] L. Murmann et al. "A Dataset of Multi-Illumination Images in the Wild". In: *Proceedings of the IEEE International Conference on Computer Vision*. 2019, pp. 4080–4089.

[19] A. Paszke et al. "PyTorch: An Imperative Style, High-Performance Deep Learning Library". In: *Advances in Neural Information Processing Systems 32*. Ed. by H. Wallach et al. Curran Associates, Inc., 2019, pp. 8024–8035. URL: http://papers.neurips.cc/paper/9015-pytorch-an-imperative-style-high-performance-deep-learning-library.pdf.

Alexandre Dherse, Martin Everaert, Jakub Gwizdala

References

[20] E. Riba et al. *Kornia: an Open Source Differentiable Computer Vision Library for PyTorch.* 2020. URL: https://arxiv.org/pdf/1910.02190.pdf.

[21] W. Shi et al. "Real-time single image and video super-resolution using an efficient sub-pixel convolutional neural network". In: *Proceedings of the IEEE conference on computer vision and pattern recognition.* 2016, pp. 1874–1883.

[22] T. Sun et al. "Single image portrait relighting". In: *ACM Transactions on Graphics (Proceedings SIGGRAPH)* (2019).

[23] S. Uchida. *LPIPS implementation.* Accessed: 2020-05-28. URL: https://github.com/S-aiueo32/lpips-pytorch.

[24] R. Wang et al. "Underexposed photo enhancement using deep illumination estimation". In: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition.* 2019, pp. 6849–6857.

[25] Z. Wang et al. "Image quality assessment: from error visibility to structural similarity". In: *IEEE transactions on image processing* 13.4 (2004), pp. 600–612.

References

[26] Z. Xu et al. "Deep image-based relighting from optimal sparse samples". In: *ACM Transactions on Graphics (TOG)* 37.4 (2018), pp. 1–13.

[27] C. Yang, Y. Shen, and B. Zhou. "Semantic hierarchy emerges in deep generative representations for scene synthesis". In: *arXiv preprint arXiv:1911.09267* (2019).

[28] A. R. Zamir et al. "Taskonomy: Disentangling task transfer learning". In: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. 2018, pp. 3712–3722.

[29] L. Zhang et al. "Generating Digital Painting Lighting Effects via RGB-space Geometry". In: *Transactions on Graphics (Presented at SIGGRAPH)* 39.2 (2020).

[30] R. Zhang et al. "The unreasonable effectiveness of deep features as a perceptual metric". In: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. 2018, pp. 586–595.

[31] H. Zhou et al. "Deep Single-Image Portrait Relighting". In: *Proceedings of the IEEE International Conference on Computer Vision*. 2019, pp. 7194–7202.

[32] T. Zhou et al. "Unsupervised learning of depth and ego-motion from video". In: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. 2017, pp. 1851–1858.

[33] J.-Y. Zhu et al. "Unpaired image-to-image translation using cycle-consistent adversarial networks". In: *Proceedings of the IEEE international conference on computer vision*. 2017, pp. 2223–2232.