FHDR: HDR Image Reconstruction from a Single LDR Image using Feedback Network

Zeeshan Khan, IIT Gandhinagar
Mukul Khanna, IIT Gandhinagar
Shanmuganathan Raman, IIT Gandhinagar

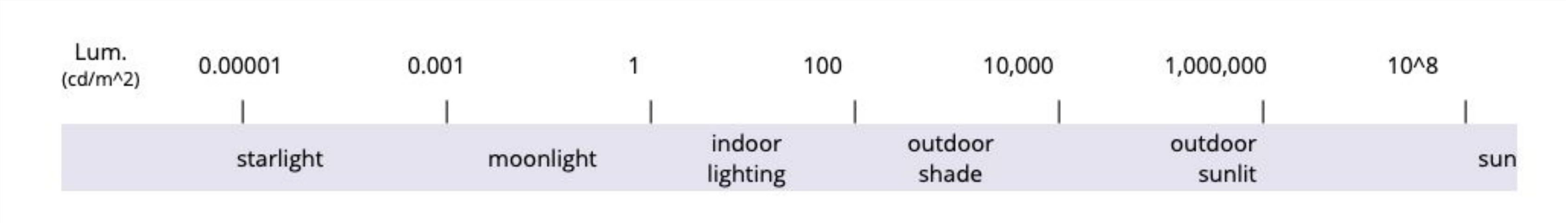
Introduction

Introduction

- Common digital cameras can not capture the wide range of light intensity levels in a natural scene.

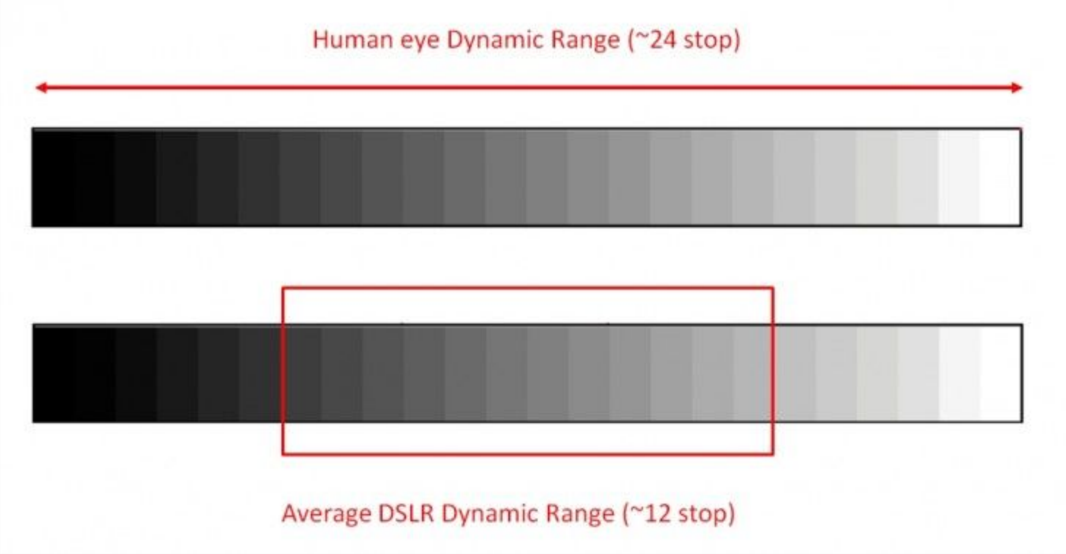
Introduction

- Common digital cameras can not capture the wide range of light intensity levels in a natural scene.



Introduction

- Common digital cameras can not capture the wide range of light intensity levels in a natural scene.



Introduction

- Common digital cameras can not capture the wide range of light intensity levels in a natural scene.
- This can lead to a loss of pixel information in under-exposed and over-exposed regions of an image, resulting in a low dynamic range (LDR) image.

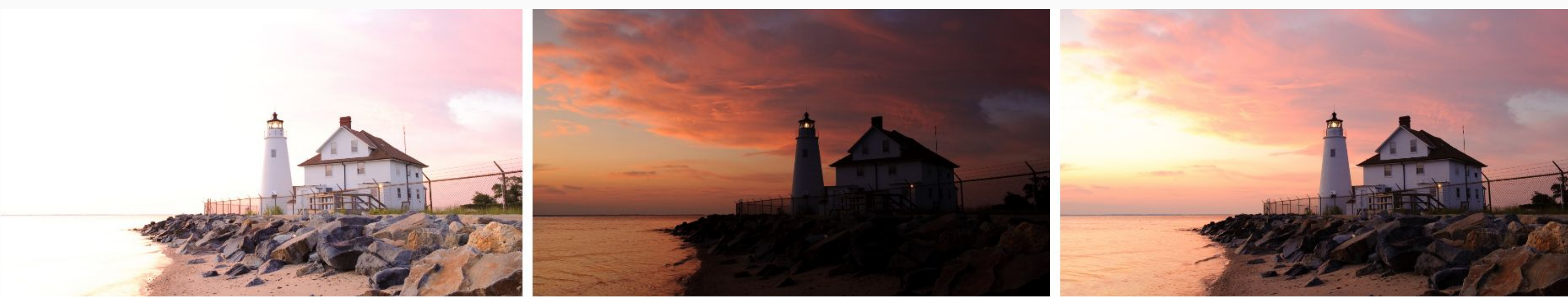
Introduction

- Common digital cameras can not capture the wide range of light intensity levels in a natural scene.
- This can lead to a loss of pixel information in under-exposed and over-exposed regions of an image, resulting in a low dynamic range (LDR) image.



Introduction

- To recover the lost information and represent the wide range of illuminance in an image, **High Dynamic Range (HDR)** images need to be generated.



Approaches - non learning based

- Conventionally, HDR images are developed by merging images captured at different exposures.

Approaches - non learning based

- Conventionally, HDR images are developed by merging images captured at different exposures.

Approaches - non learning based

- Conventionally, HDR images are developed by merging images captured at different exposures.

Approaches - non learning based

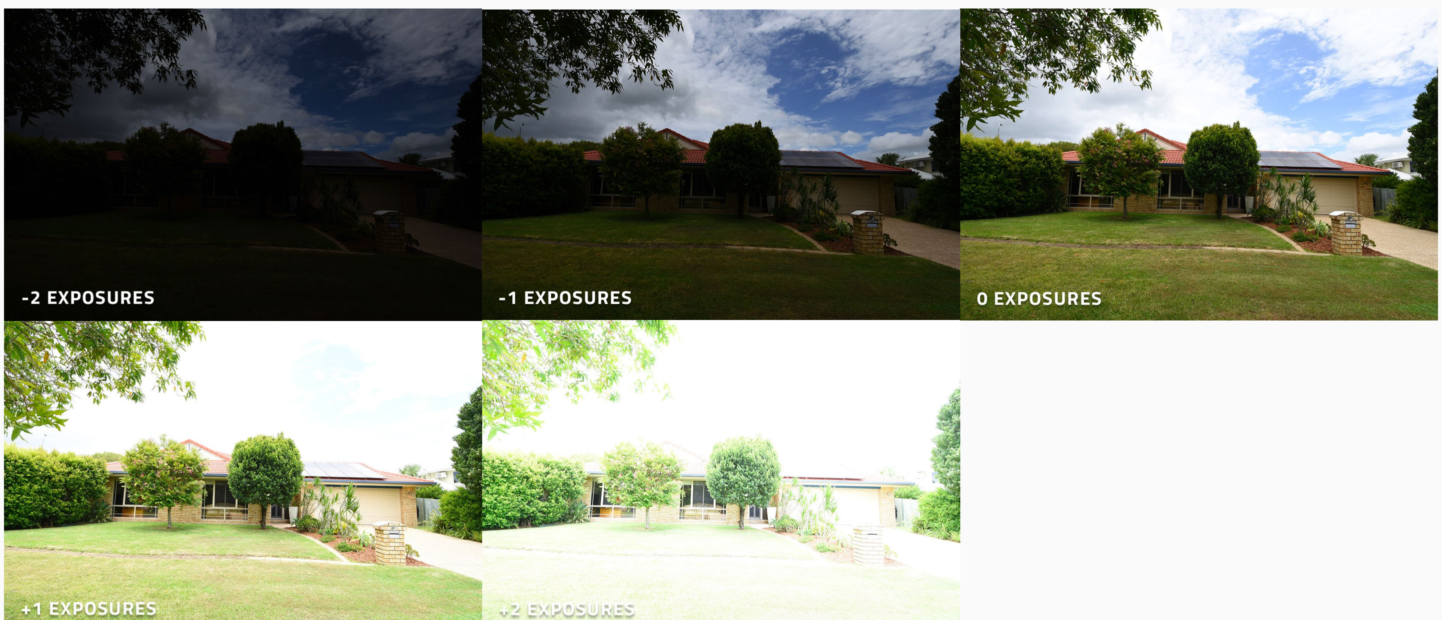
- Conventionally, HDR images are developed by merging images captured at different exposures.

Approaches - non learning based

- Conventionally, HDR images are developed by merging images captured at different exposures.

Approaches - non learning based

- Conventionally, HDR images are developed by merging images captured at different exposures.



Approaches - non learning based

- These images are merged using a software algorithm and are saved as a single HDR image, in a way that the best portions of each image make it to the final image.

Approaches - non learning based

- These images are merged using a software algorithm and are saved as a single HDR image, in a way that the best portions of each image make it to the final image.

Caveats

- Conventional approaches aren't robust enough when it comes to dynamic scenes with motion between the bracketed frames.

Caveats

- Conventional approaches aren’t robust enough when it comes to dynamic scenes with motion between the bracketed frames.
- They rely on Optical Flow to account for the motion between frames.

Caveats

- Conventional approaches aren't robust enough when it comes to dynamic scenes with motion between the bracketed frames.
- They rely on Optical Flow to account for the motion between frames.

Caveats

- Conventional approaches aren’t robust enough when it comes to dynamic scenes with motion between the bracketed frames.
- They rely on Optical Flow to account for motion between frames.
- But Optical Flow is not accurate.

Caveats

- Conventional approaches aren't robust enough when it comes to dynamic scenes with motion between the bracketed frames.
- They rely on Optical Flow to account for motion between frames.
- But Optical Flow is not accurate.
- This can result in ghosting artifacts in the final image.

Caveats

- Conventional approaches aren’t robust enough when it comes to dynamic scenes with motion between the bracketed frames.
- They rely on Optical Flow to account for motion between frames.
- But Optical Flow is not accurate.
- This can result in ghosting artifacts in the final image.

Approaches - learning based

- Learning based approaches harness the capabilities of deep neural network architectures as function approximators to learn LDR to HDR representations.

Approaches - learning based

- Learning based approaches harness the capabilities of deep neural network architectures as function approximators to learn LDR to HDR representations.

Approaches - learning based

- Learning based approaches harness the capabilities of deep neural network architectures as function approximators to learn LDR to HDR representations.
- Such networks can do better due to -
  - improved learning based flow mechanisms
  - hallucinating HDR content in saturated regions when LDR input is limited
  - optimised, quick, low-memory alternative

Approaches - learning based

- Learning based approaches can be broken down into two types -

Approaches - learning based

- Learning based approaches can be broken down into two types -
  - Single LDR input

Approaches - learning based

- Learning based approaches can be broken down into two types -
  - Single LDR input

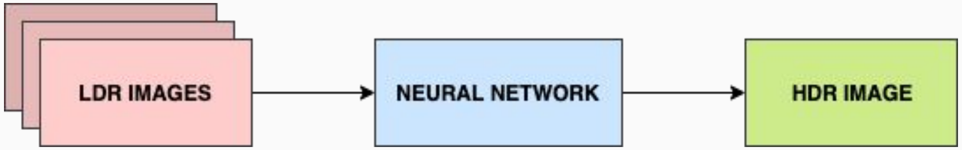
Approaches - learning based

- Learning based approaches can be broken down into two types -

  - Single LDR input
  - Multiple LDR inputs

Approaches - learning based

- Learning based approaches can be broken down into two types -
    - Single LDR input
    - Multiple LDR inputs



Learning based - multiple LDR inputs

- Multiple exposure input

Learning based - multiple LDR inputs

- Multiple exposure input

Learning based - multiple LDR inputs

- Multiple exposure input
- More dynamic range is provided to the network

Learning based - multiple LDR inputs

- Multiple exposure input
- More dynamic range is provided to the network
- Explicit mechanism is required for motion compensation

Learning based - multiple LDR inputs

- Multiple exposure input
- More dynamic range is provided to the network
- Explicit mechanism is required for motion compensation
- Better results

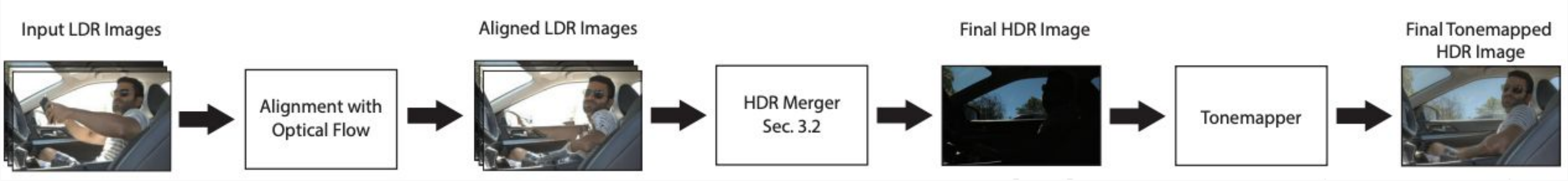
Learning based - multiple LDR inputs

- Multiple exposure input
- More dynamic range is provided to the network
- Explicit mechanism required for motion compensation
- Better results
- But input is a constraint

Related work
(multi-exposure input)

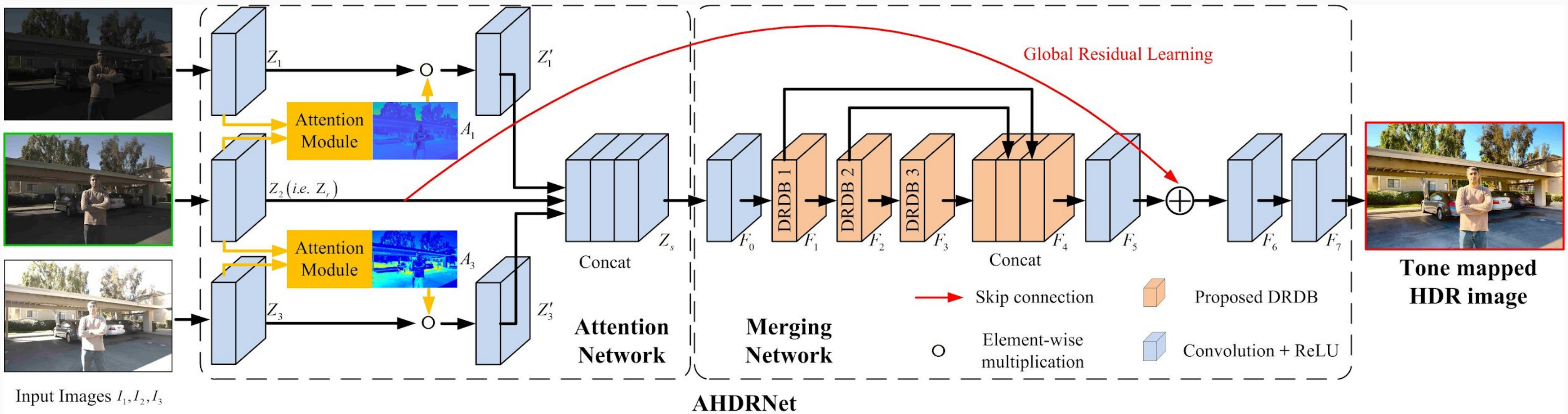
Deep HDR

N. K. Kalantari and R. Ramamoorthi, “Deep high dynamic range imaging of dynamic scenes.,” ACM Trans. Graph., vol. 36, no. 4, pp. 144–1, 2017.



AHDR-Net

Yan, Q., Gong, D. , Shi, Q., van den Hengel, A., Shen, C., Reid, I. and Zhang, Y.: Attention-guided Network for Ghost-free High Dynamic Range Imaging. The IEEE Conference on Computer Vision and Pattern Recognition (CVPR) 2019



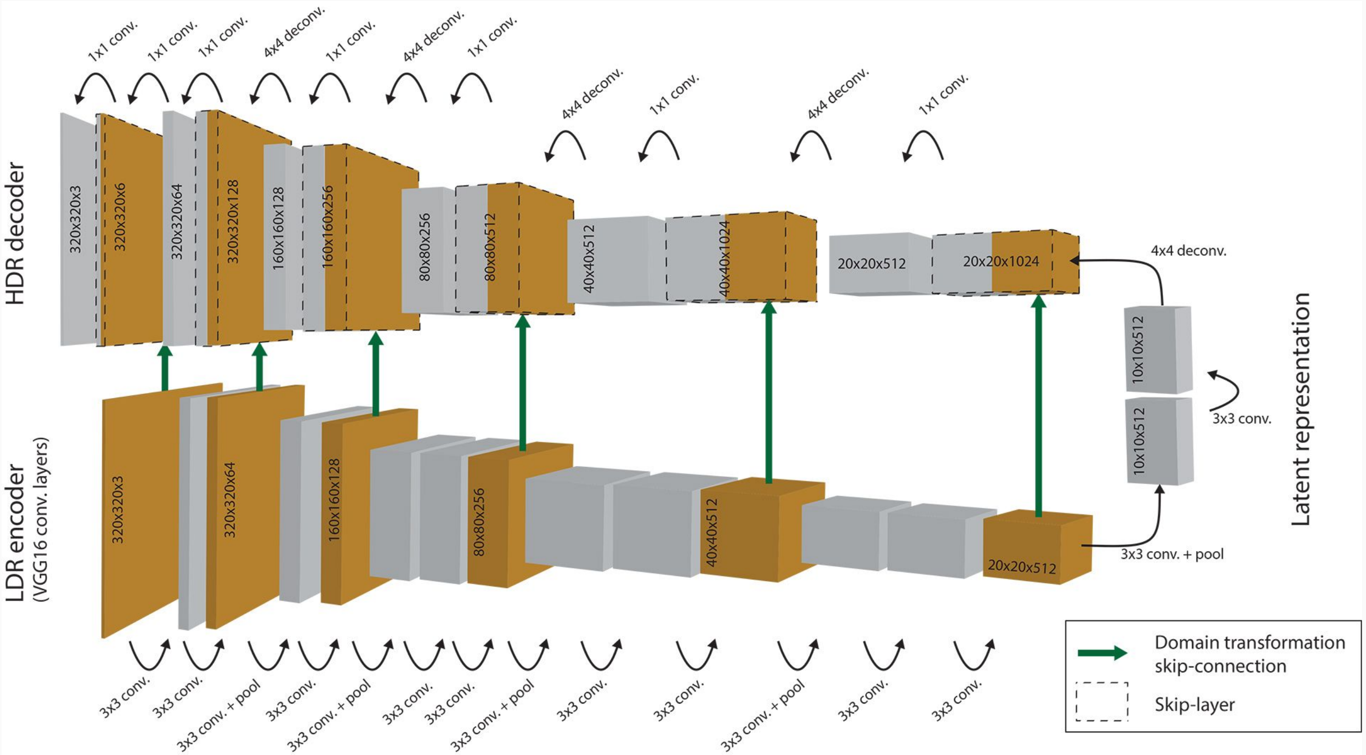
Learning based - single LDR input

- More challenging scenario
- Limited dynamic range information input
- More important for real life situations
- Heavily relies on ability of deep CNNs to hallucinate content in saturated image regions.

Related work
(single input)

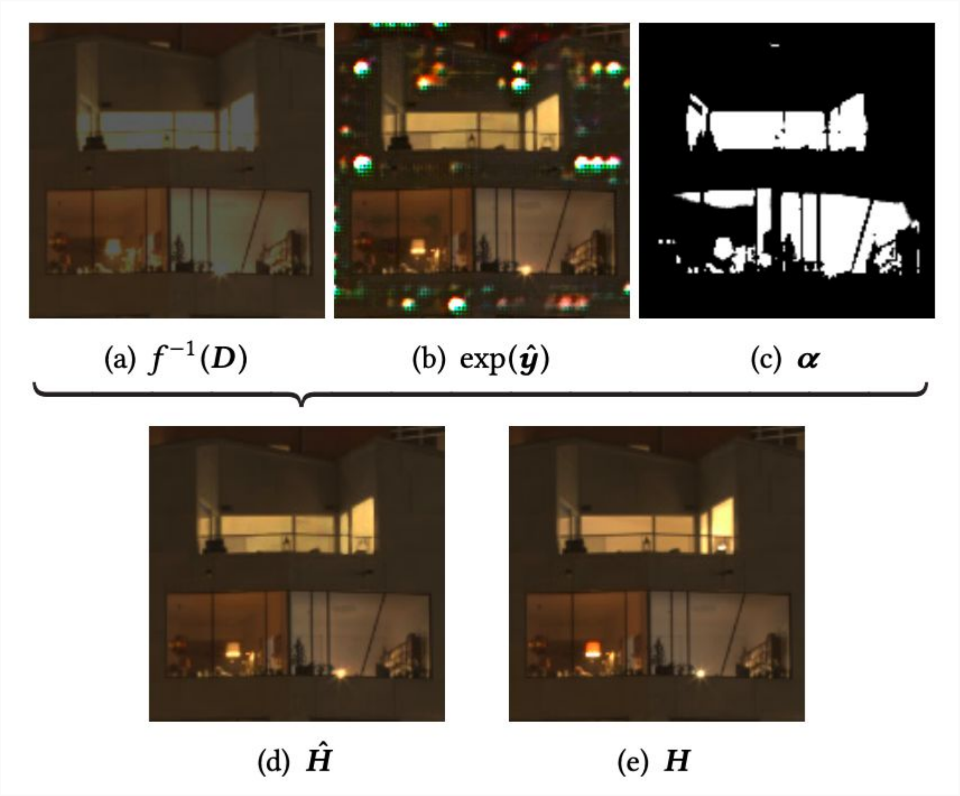
HDRCNN

G. Eilertsen, J. Kronander, G. Denes, R. K. Mantiuk, and J. Unger, “Hdr image reconstruction from a single exposure using deep cnns,” ACM Transactions on Graphics (TOG), vol. 36, no. 6, p. 178, 2017



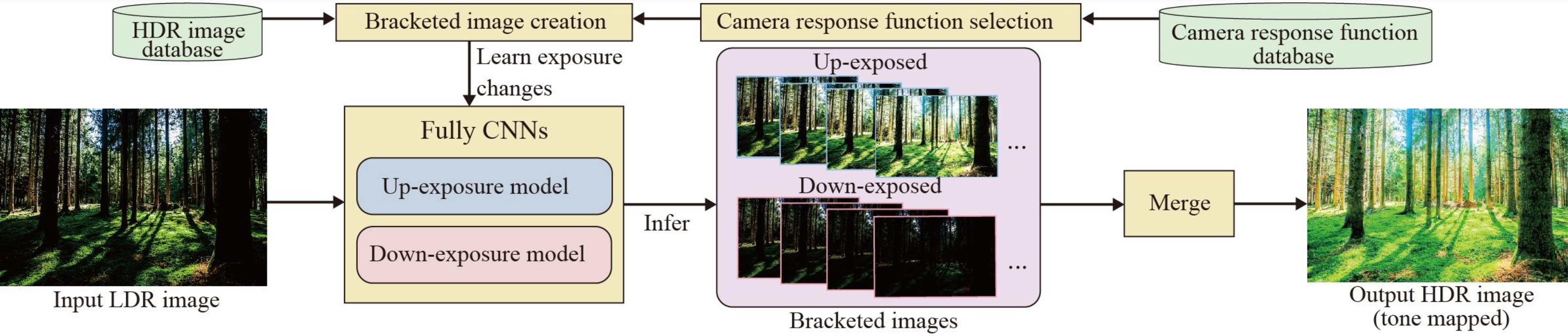
HDRCNN

G. Eilertsen, J. Kronander, G. Denes, R. K. Mantiuk, and J. Unger, “Hdr image reconstruction from a single exposure using deep cnns,” ACM Transactions on Graphics (TOG), vol. 36, no. 6, p. 178, 2017



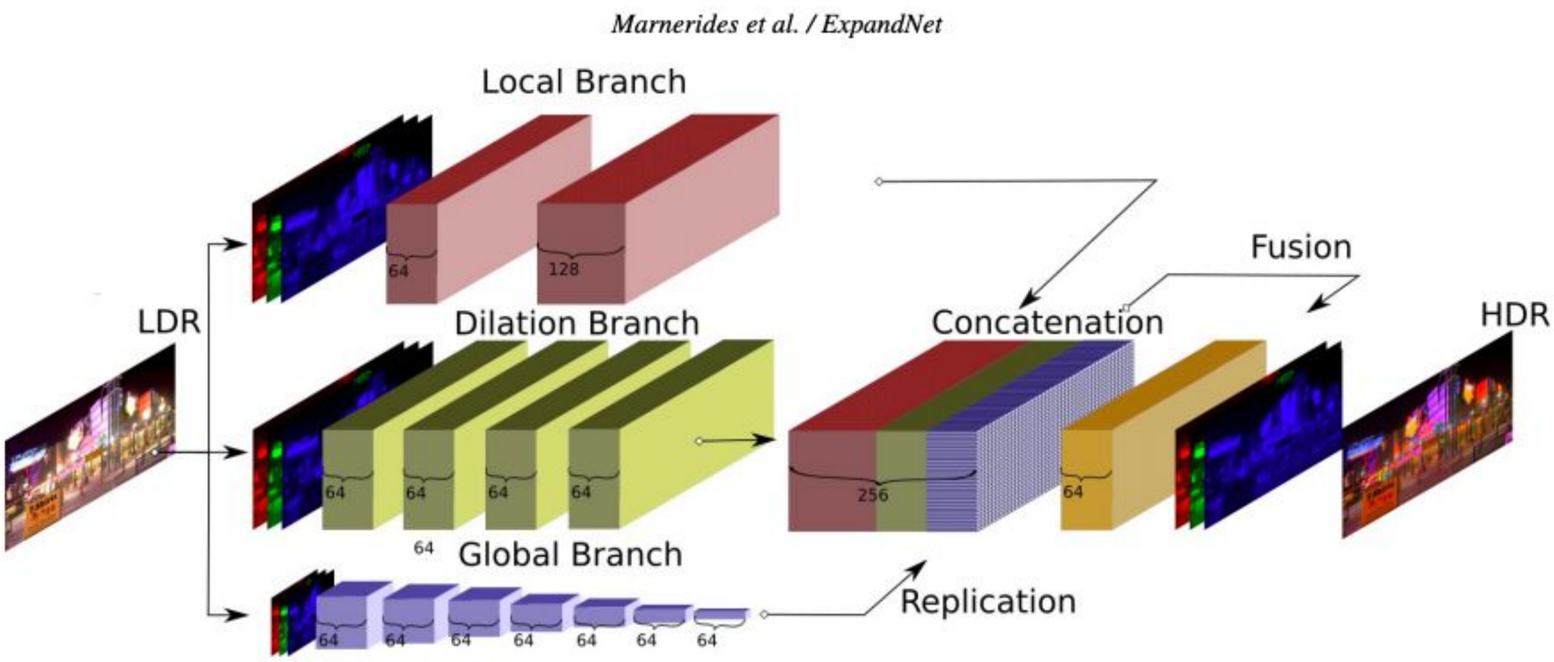
Deep reverse tone mapping

Y. Endo, Y. Kanamori, and J. Mitani, "Deep reverse tone mapping.," ACM Trans. Graph., vol. 36, no. 6, pp. 177–1, 2017.



ExpandNet

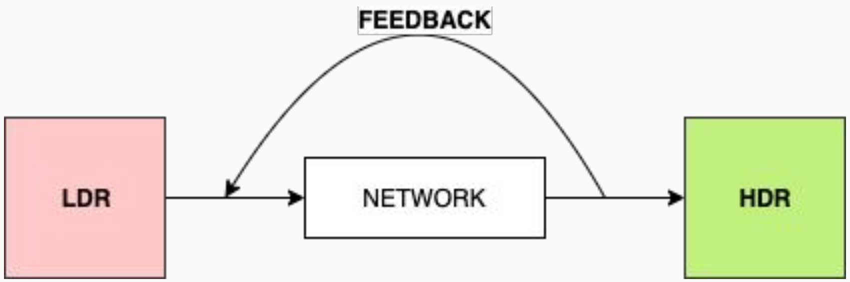
D. Marnerides, T. Bashford-Rogers, J. Hatchett, and K. Debattista, “Expandnet: A deep convolutional neural network for high dynamic range expansion from low dynamic range content,” in Computer Graphics Forum, vol. 37, pp. 37–49, Wiley Online Library, 2018.



Our approach

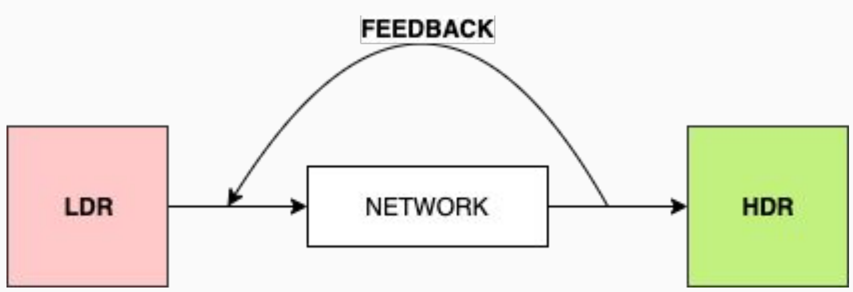
Feedback networks

- Feedback systems are adopted to influence the input based on the generated output.

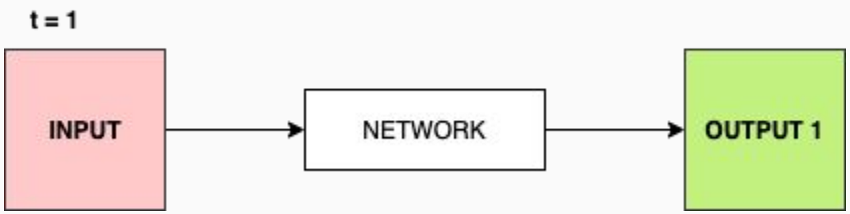


Feedback networks

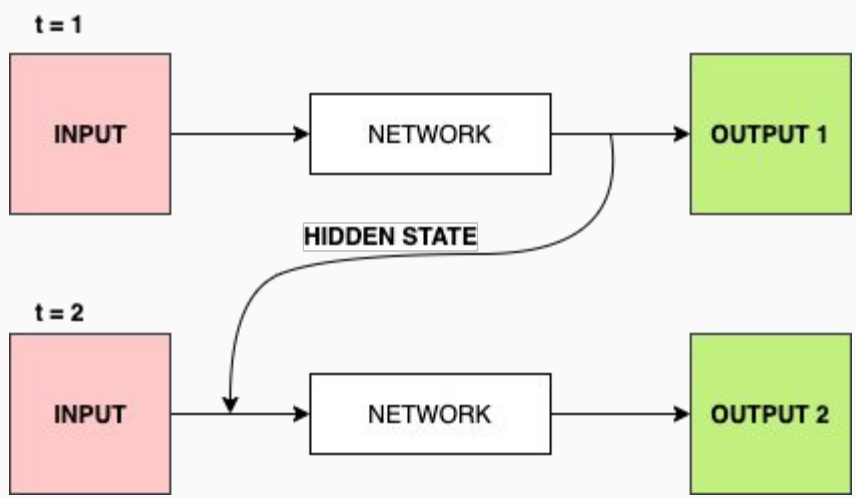
- Feedback systems are adopted to influence the input based on the generated output.
- Initial low level features are guided by the high level features using a hidden state of a Recurrent Neural Network over n iterations.



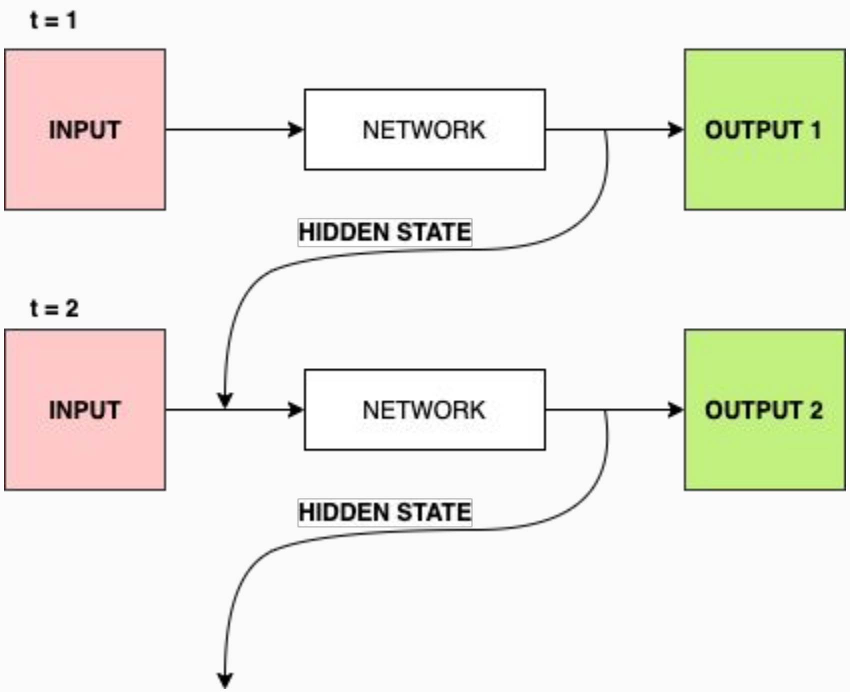
Feedback networks



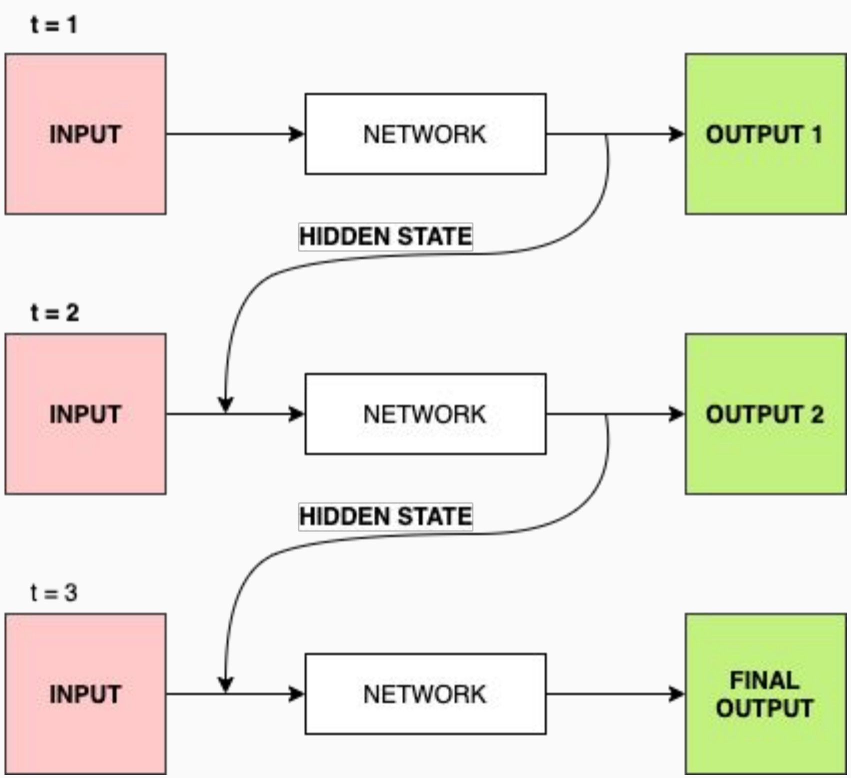
Feedback networks



Feedback networks

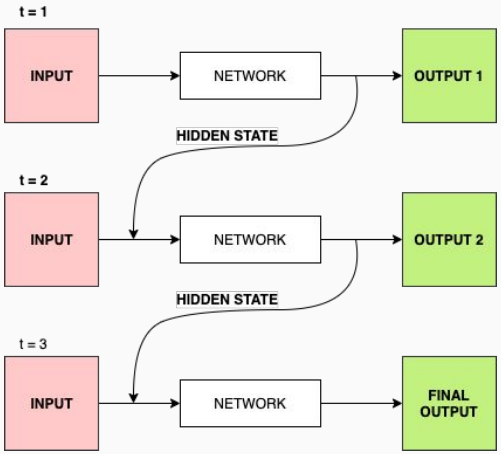


Feedback networks



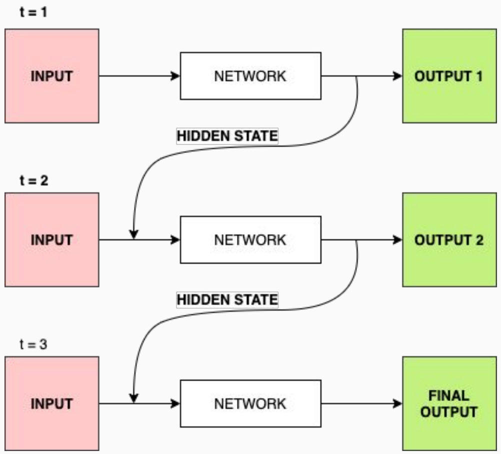
Feedback networks

- Feedback systems are adopted to influence the input based on the generated output.
- Initial low level features are guided by the high level features using a hidden state of a Recurrent Neural Network over n iterations.
- Backpropagation in time through an unfolded RNN

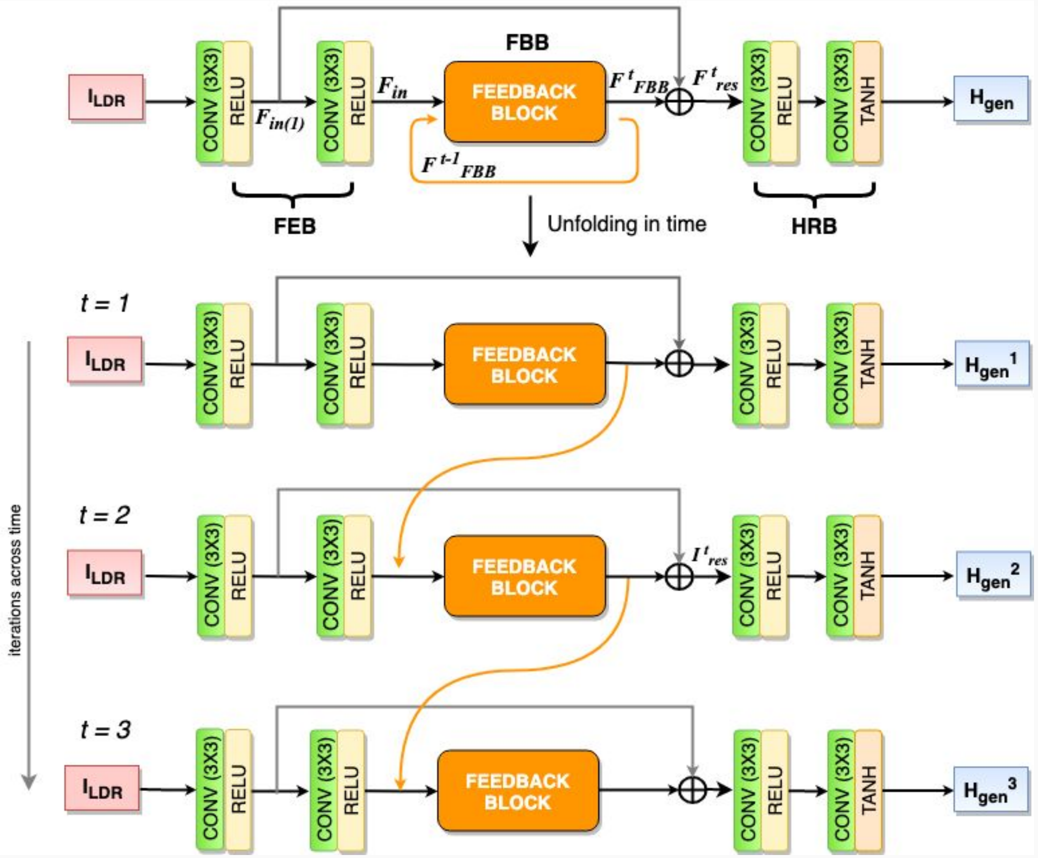


Feedback networks

- Feedback systems are adopted to influence the input based on the generated output.
- Initial low level features are guided by the high level features using a hidden state of a Recurrent Neural Network over n iterations.
- Backpropagation in time through an unfolded RNN
- Helps in learning coarse-to-fine representations



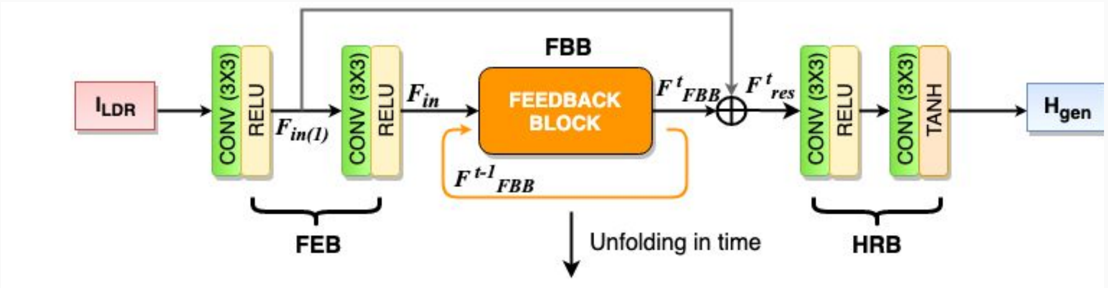
Model architecture



FHDR: HDR Image Reconstruction from a Single LDR Image using Feedback Network · Z. Khan, M. Khanna and S. Raman

Feature Extraction Block (FEB)

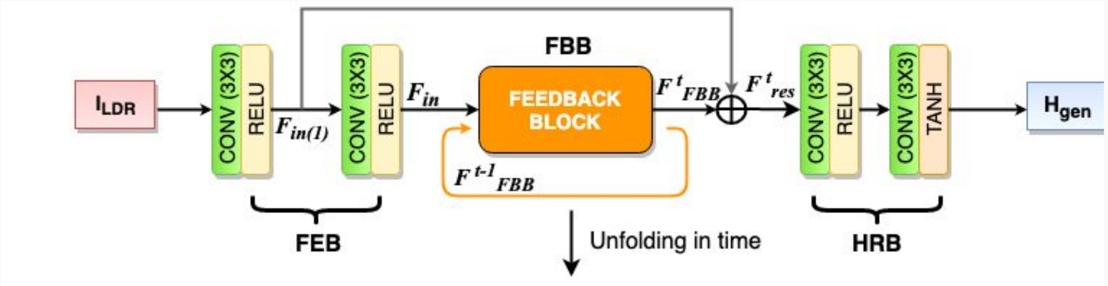
- The FEB is responsible for extracting the low-level feature information F_{in} from the input LDR image ILDR.



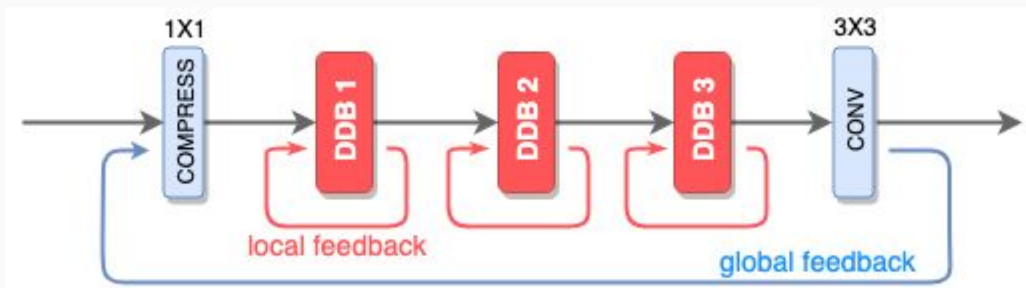
Feature Extraction Block (FEB)

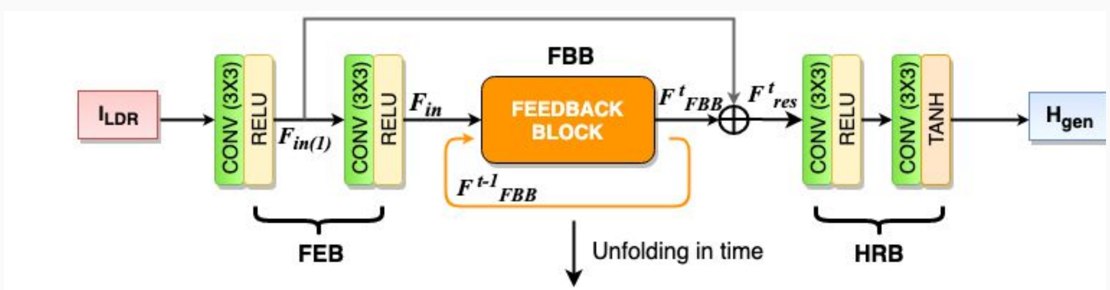
- The FEB is responsible for extracting the low-level feature information F_{in} from the input LDR image ILDR.

$$F_{in} = f_{FEB}(I_{LDR}).$$



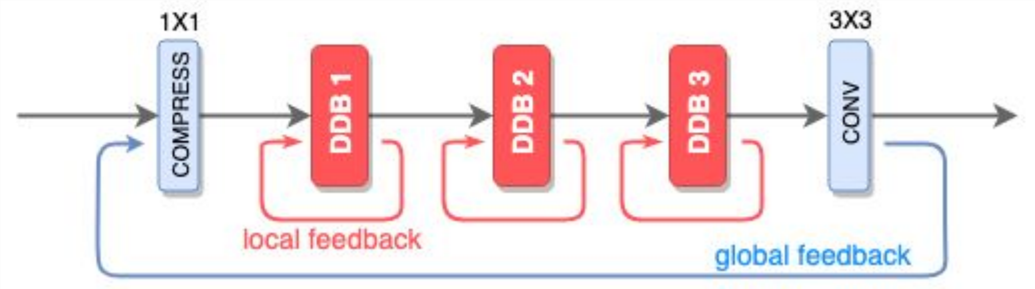
Feedback block





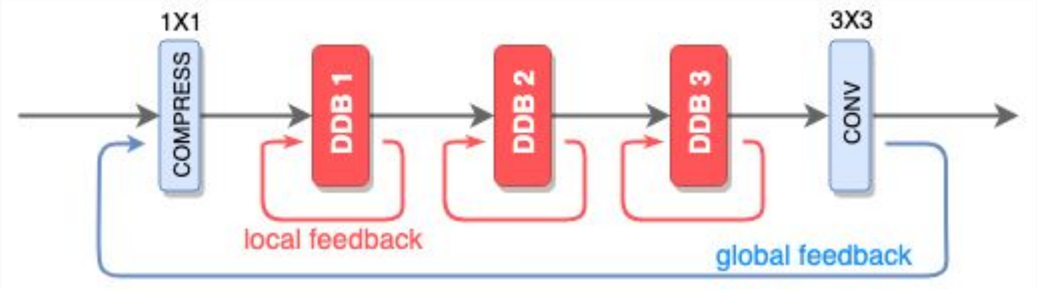
Feedback block

- Novel feedback block for the task of learning LDR-to-HDR mapping



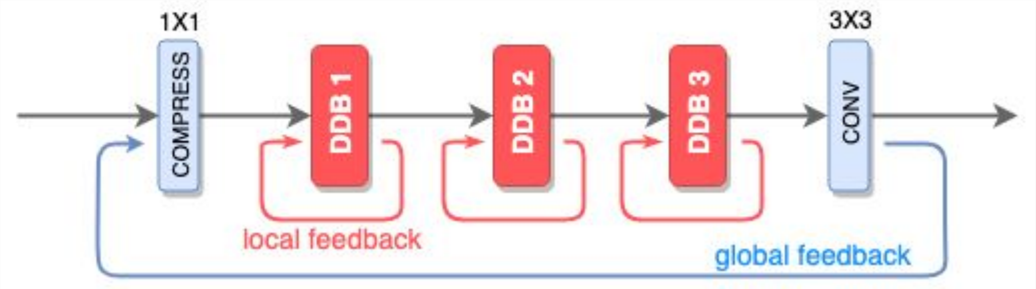
Feedback block

- Novel feedback block for the task of learning LDR-to-HDR mapping
- The basic unit of the feedback block is a Dilated Dense Block (DDB).



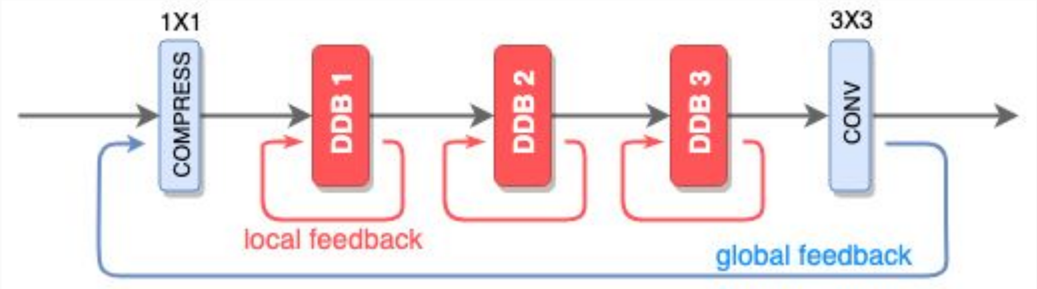
Feedback block

- Novel feedback block for the task of learning LDR-to-HDR mapping
- The basic unit of the feedback block is a Dilated Dense Block (DDB).
- A DDB helps in utilising all the hierarchical features from the input.

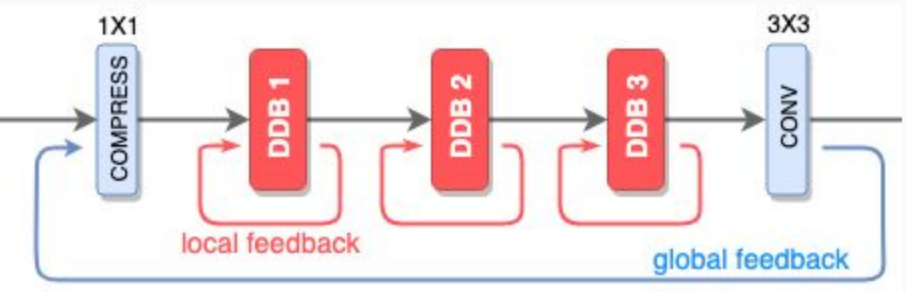


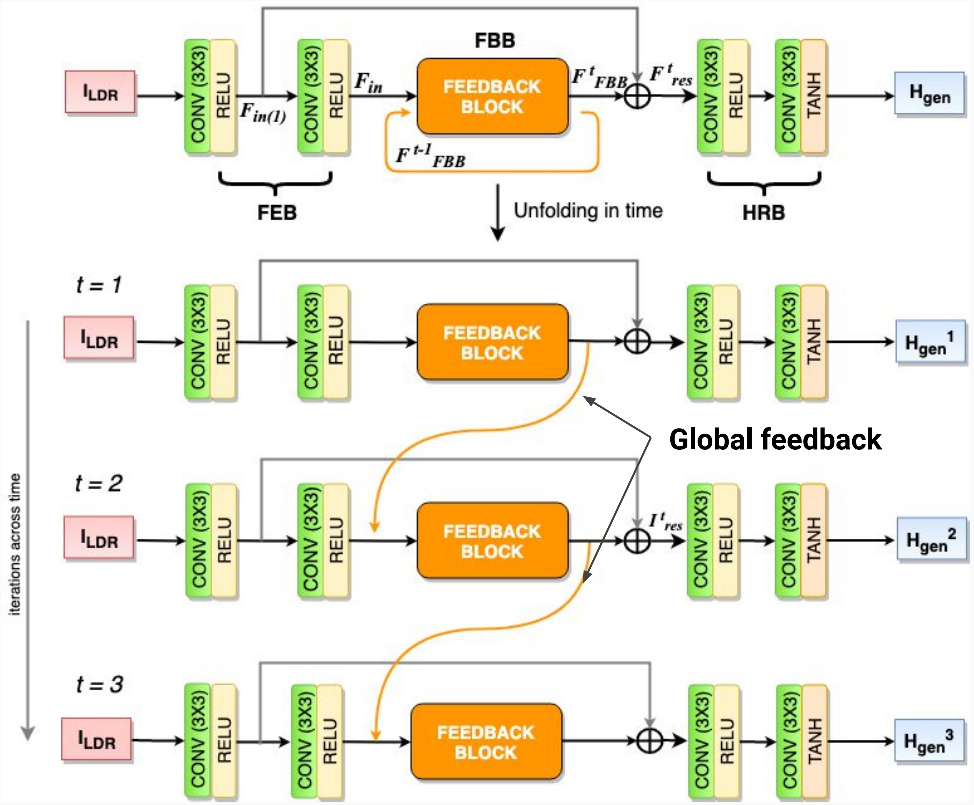
Feedback block

- Novel feedback block for the task of learning LDR-to-HDR mapping
- The basic unit of the feedback block is a Dilated Dense Block (DDB).
- A DDB helps in utilising all the hierarchical features from the input.
- Three DDBs are used for the feedback block of the network.



Feedback block - Global feedback

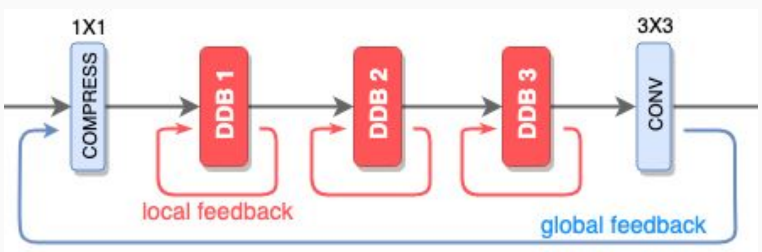


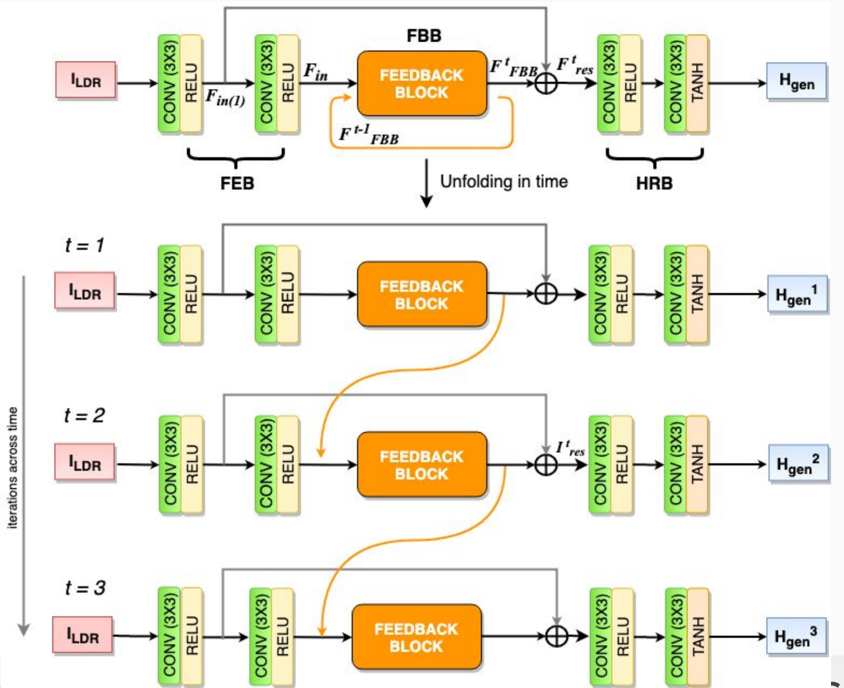


FHDR: HDR Image Reconstruction from a Single LDR Image using Feedback Network · Z. Khan, M. Khanna and S. Raman

Feedback block - Global feedback

- High level features are transferred from the output of the feedback block at the t − 1th iteration to its input at the tth iteration.

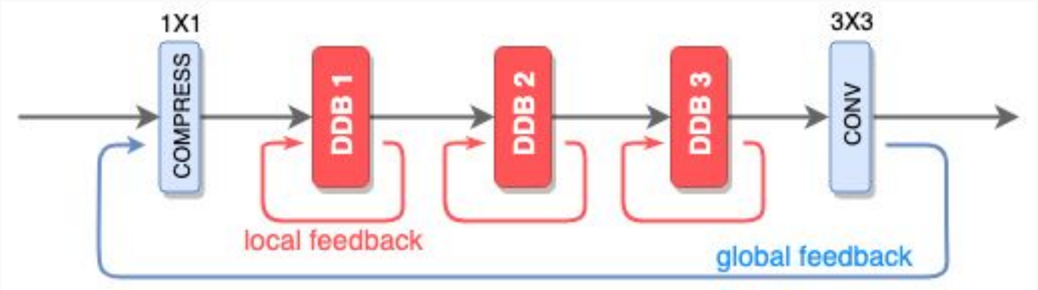




FHDR: HDR Image Reconstruction from a Single LDR Image using Feedback Network · Z. Khan, M. Khanna and S. Raman

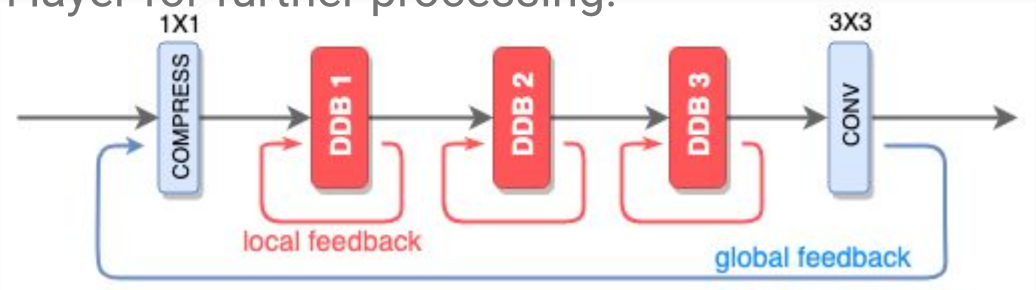
Feedback block - Global feedback

- High level features are transferred from the output of the feedback block at the $t-1^{th}$ iteration to its input at the t^{th} iteration.
- The hidden state is concatenated with the corresponding input, and a 1×1 compression convolution layer is applied for high-level and low-level feature fusion.



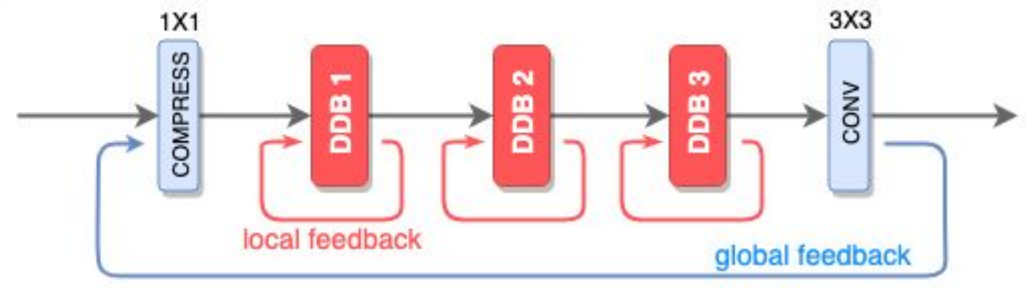
Feedback block - Global feedback

- High level features are transferred from the output of the feedback block at the $t - 1^{th}$ iteration to its input at the t^{th} iteration.
- The hidden state is concatenated with the corresponding input, and a 1×1 compression convolution layer is applied for high-level and low-level feature fusion.
- The fused features are passed to the DDBs, followed by a 3×3 convolution layer for further processing.



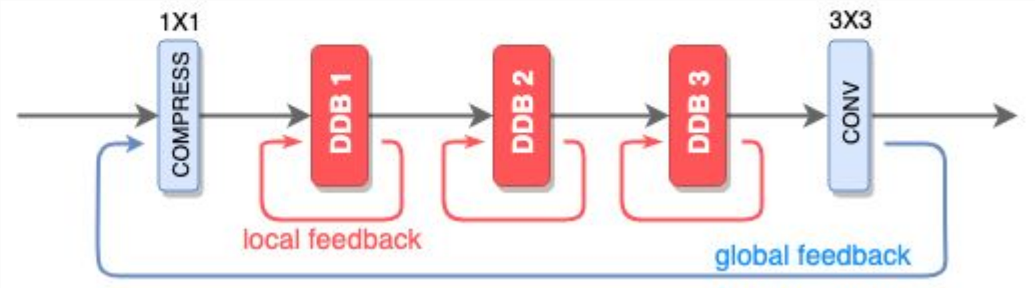
Feedback block - Local feedback

- We argue that a feedback connection is always beneficial.



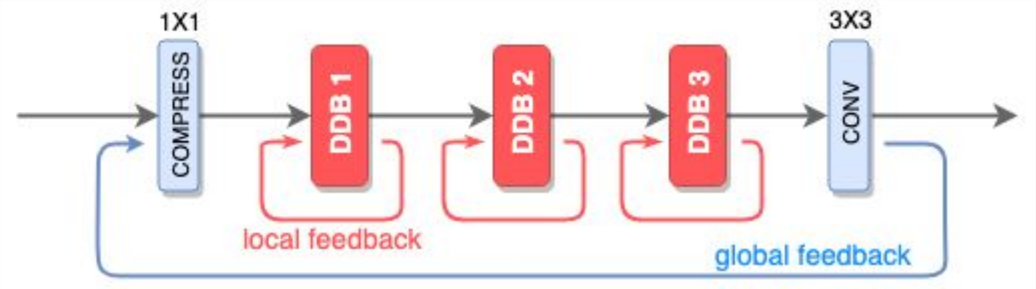
Feedback block - Local feedback

- We argue that a feedback connection is always beneficial.
- Helps to guide the low-level features which are in some way blind to the higher level features.



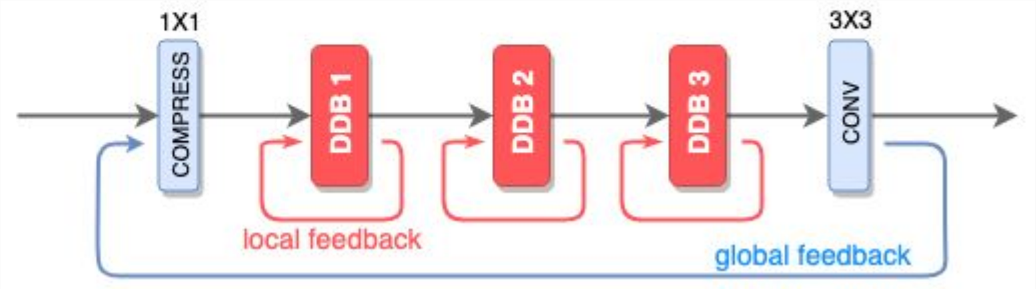
Feedback block - Local feedback

- We argue that a feedback connection is always beneficial.
- Helps to guide the low-level features which are in some way blind to the higher level features.
- Local feedback connections aim to improve the features generated locally.

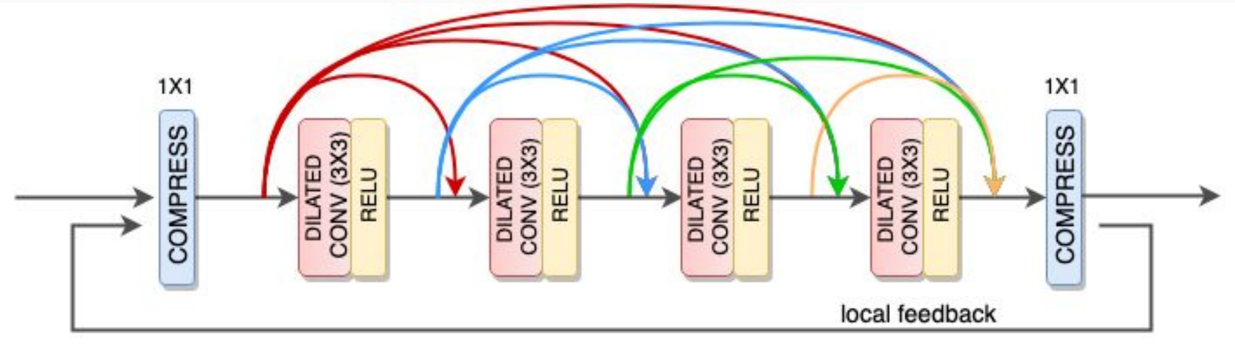


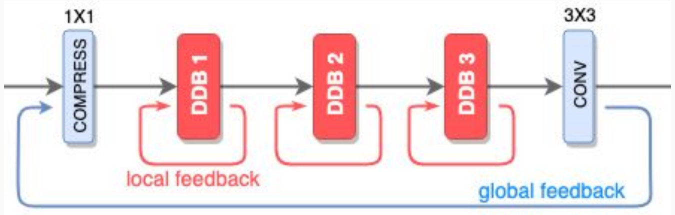
Feedback block - Local feedback

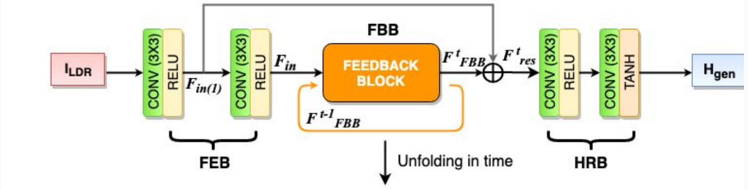
- We argue that a feedback connection is always beneficial.
- Helps to guide the low-level features which are in some way blind to the higher level features.
- Local feedback connections aim to improve the features generated locally.
- These connections run parallel to the global feedback connections and increase the overall effectiveness of the network.



Dilated Dense Block (DDB)

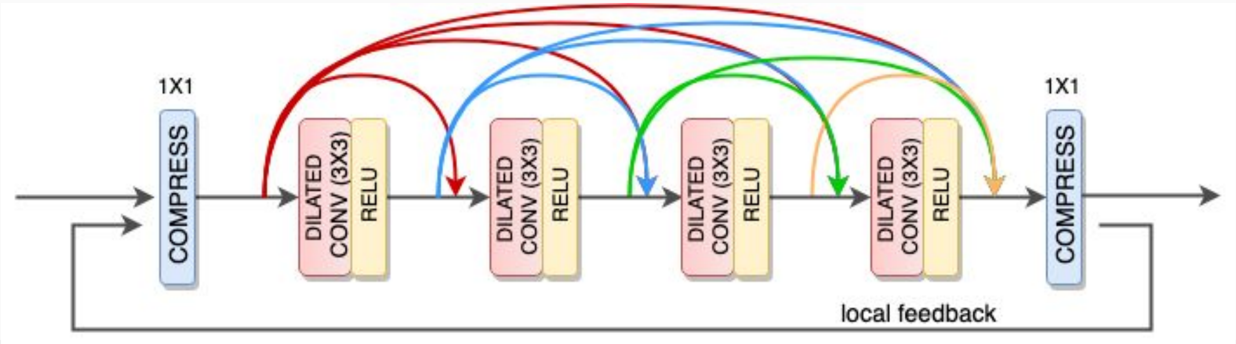






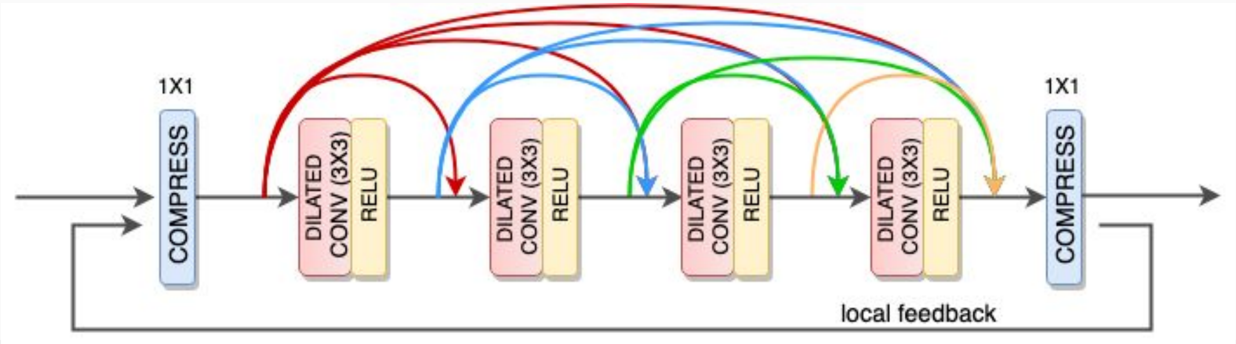
Dilated Dense Block (DDB)

- Basic unit of the feedback block



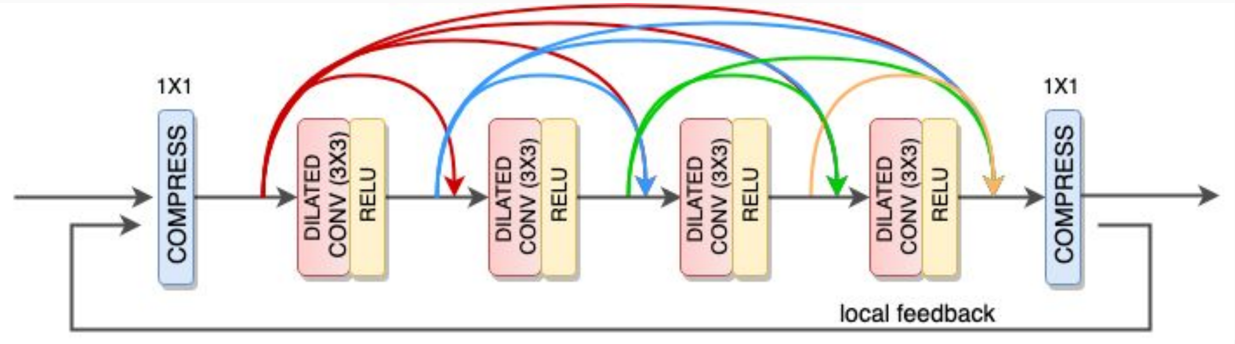
Dilated Dense Block (DDB)

- Basic unit of the feedback block
- Dilated convolutions help in increasing the receptive field of the network.



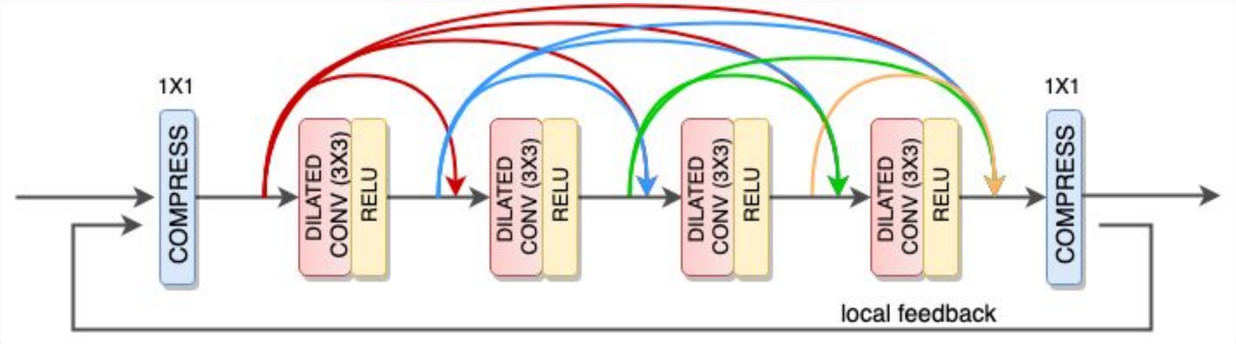
Dilated Dense Block (DDB)

- Basic unit of the feedback block
- Dilated convolutions help in increasing the receptive field of the network.
- Utilises all the hierarchical features from the input.



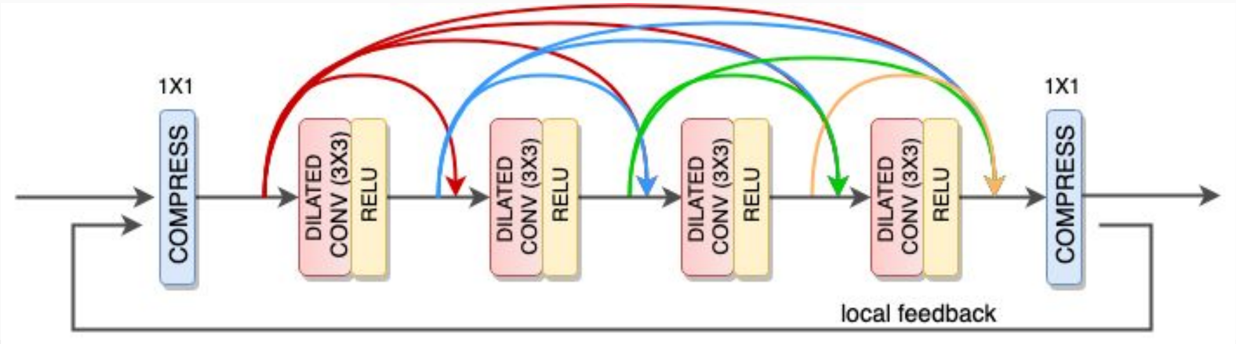
Dilated Dense Block (DDB)

- Basic unit of the feedback block
- Dilated convolutions help in increasing the receptive field of the network.
- Utilises all the hierarchical features from the input.
- Two 1 × 1 convolutional layers for feature compression + four dilated 3 × 3 dense convolutional layers

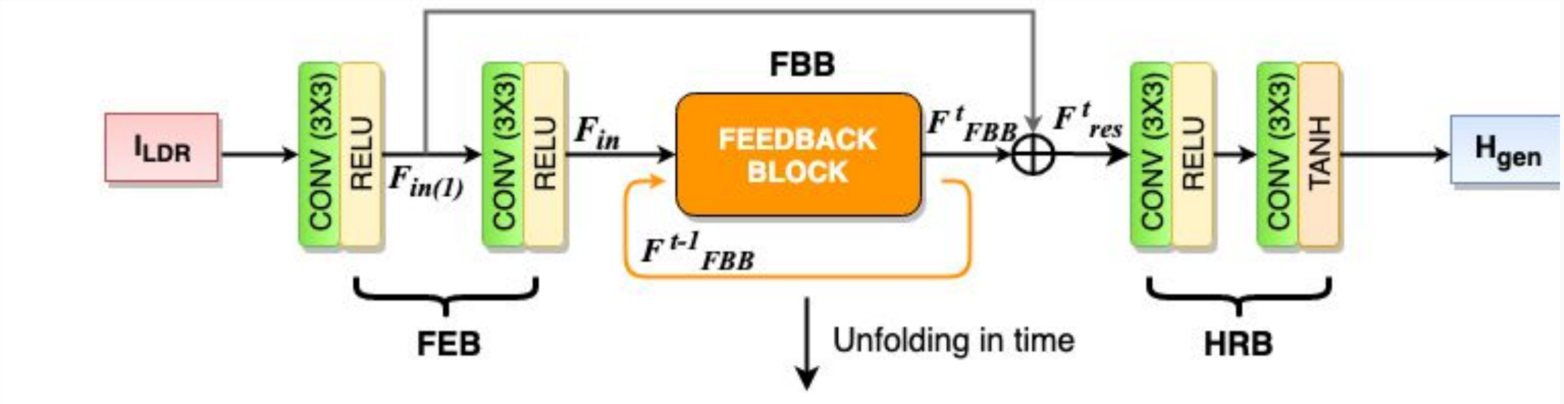


Dilated Dense Block (DDB)

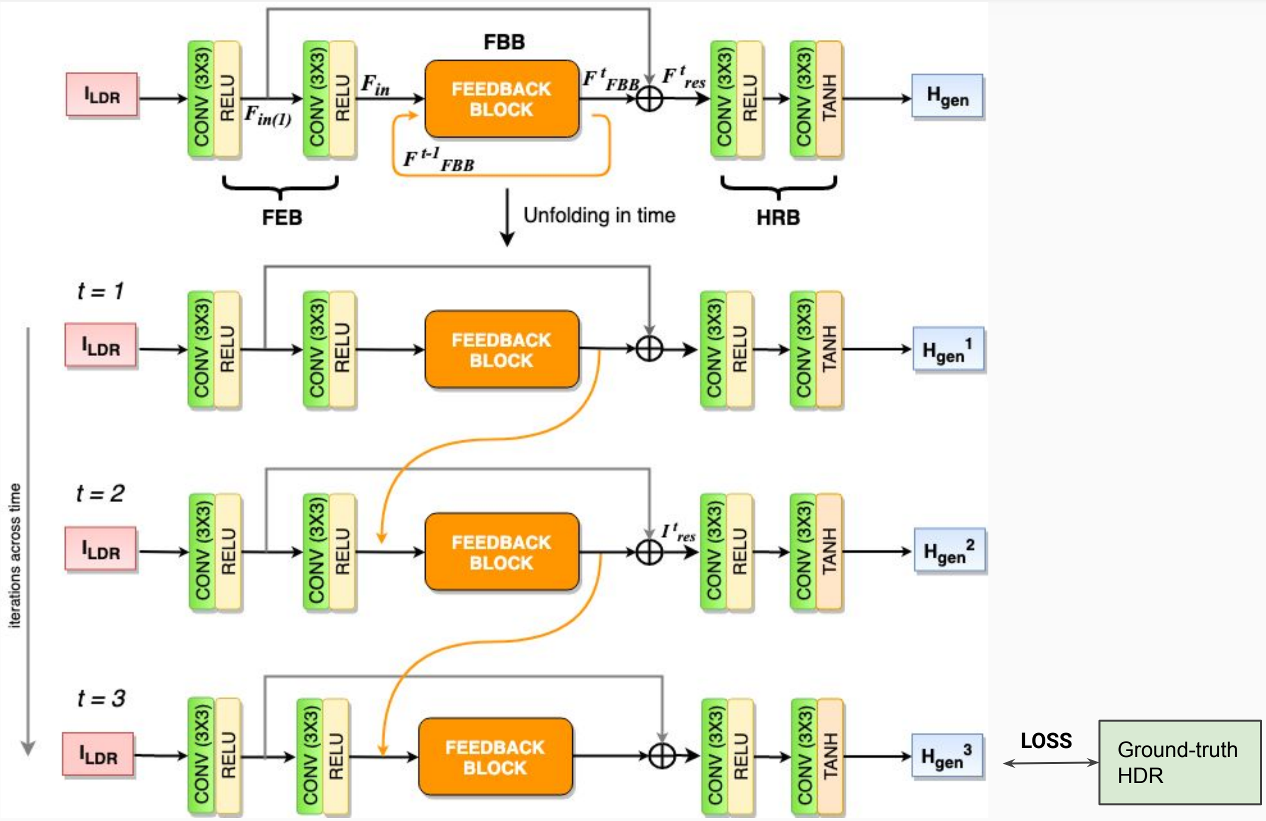
- Basic unit of the feedback block
- Dilated convolutions help in increasing the receptive field of the network.
- Utilises all the hierarchical features from the input.
- Two 1 × 1 convolutional layers for feature compression + four dilated 3 × 3 dense convolutional layers
- Feature reuse, reduced network parameters, improved learning ability.



HDR Reconstruction Block (HRB)



Loss function



Loss function

- Loss calculated directly on HDR images is misrepresented due to the dominance of high intensity values of images with a wide dynamic range.

Loss function

- Loss calculated directly on HDR images is misrepresented due to the dominance of high intensity values of images with a wide dynamic range.
- Therefore, we tonemap the generated and the ground truth HDR images to compress the wide intensity range before calculating the loss.

Loss function

- Loss calculated directly on HDR images is misrepresented due to the dominance of high intensity values of images with a wide dynamic range.
- Therefore, we tonemap the generated and the ground truth HDR images to compress the wide intensity range before calculating the loss.
- We use the μ-law for tonemapping -

$$T(H^t_{gen}) = \frac{log(1 + \mu H^t_{gen})}{log(1 + \mu)}$$

Loss function

- Loss calculated directly on HDR images is misrepresented due to the dominance of high intensity values of images with a wide dynamic range.
- Therefore, we tonemap the generated and the ground truth HDR images to compress the wide intensity range before calculating the loss.
- We use the μ-law for tonemapping.
- **L1 loss and Perceptual loss** ($\lambda = 0.1$)

$$\mathcal{L} = \mathcal{L}_p + \lambda \mathcal{L}_{L1}$$

Implementation

- Adam optimizer β1 = 0.5 and β2 = 0.999
- 200 epochs
- Batch size: 16 (CityScene dataset), 6 (Curated HDR dataset)
- Learning rate of 2×10^{-4} for first 100 epochs, decayed linearly over the next 100 epochs
- 2 RTX 2070 GPUs

Experiments

Datasets

The performance of the network was evaluated over two datasets-

Datasets

The performance of the network was evaluated over two datasets-

- CityScene dataset
  - 128 x 64 size
  - Training set - 39,460 LDR-HDR image pairs
  - Testing set - 1,672 pairs

Datasets

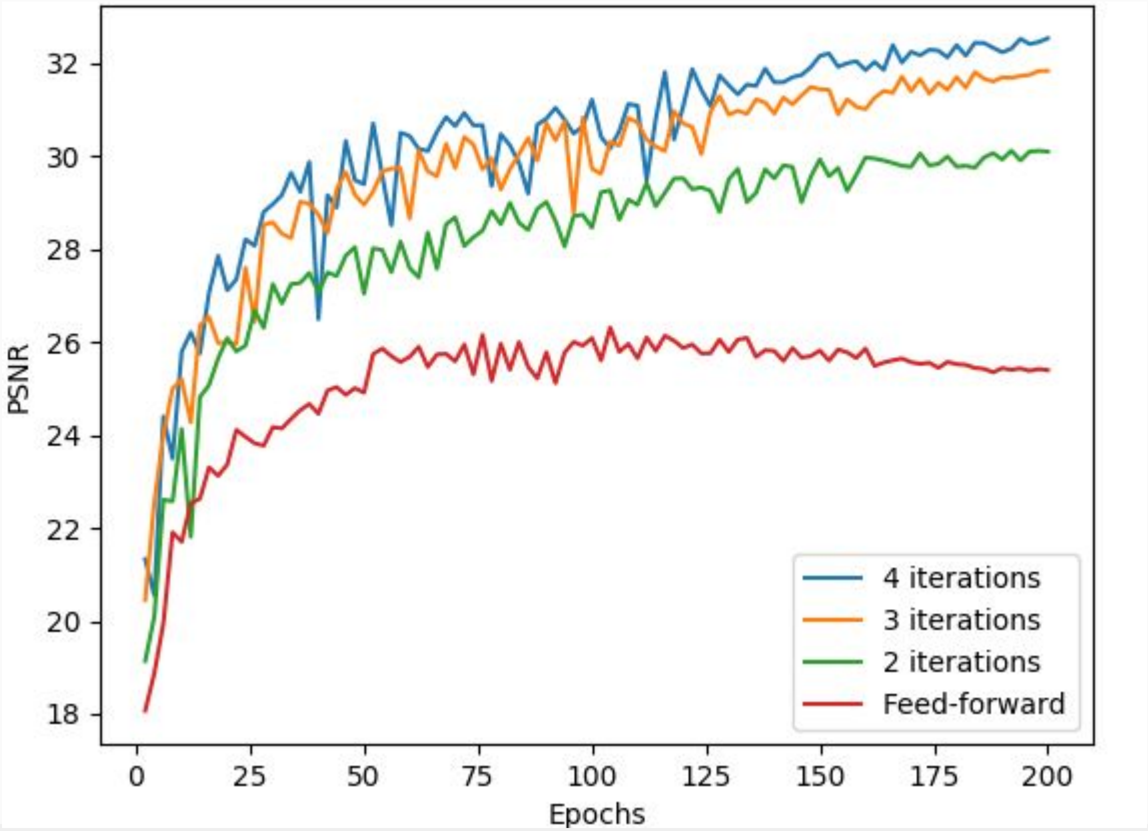
The performance of the network was evaluated over two datasets-

- Curated dataset
    - 256 x 256 size
    - Training set - 11,262 LDR-HDR image pairs
    - Testing set - 500 image pairs (512 x 512)

Evaluation metrics

- PSNR score (db) - Peak Signal-to-Noise Ratio
- SSIM score - Structural Similarity Index
- HDR-VDP2 Q-score

Feedback mechanism analysis



Results

Qualitative evaluation

LDR

GENERATED

GROUND TRUTH

Qualitative evaluation

LDR

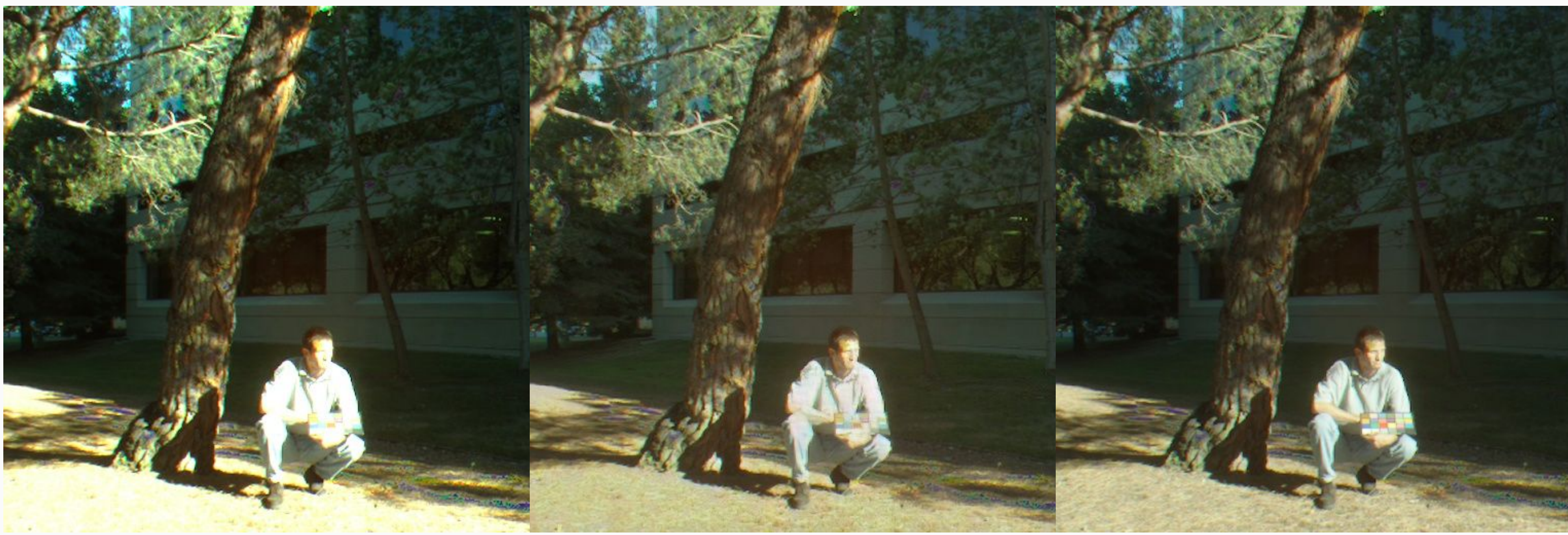
GENERATED

GROUND TRUTH

Qualitative evaluation

LDR GENERATED GROUND TRUTH

Qualitative evaluation



LDR GENERATED GROUND TRUTH

Qualitative comparisons

LDR | **DRTMO** | **FHDR** | **GROUND TRUTH**

Qualitative comparisons

LDR | HDRCNN | FHDR | GROUND TRUTH

Qualitative evaluation

LDR

GENERATED

GROUND TRUTH

Qualitative comparisons

LDR DRTMO FHDR GROUND TRUTH

Qualitative comparisons

LDR **HDRCNN** **FHDR** **GROUND TRUTH**

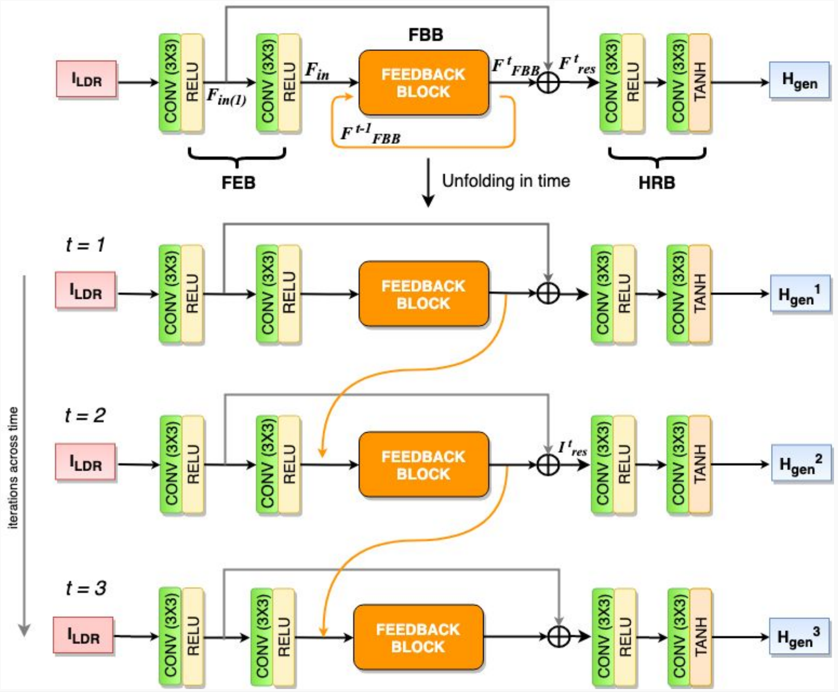
Quantitative evaluation

Methods	City Scene Dataset			Curated HDR Dataset		
	PSNR	**SSIM**	**Q-score**	**PSNR**	**SSIM**	**Q-score**
AKY [14]	15.35	0.44	35.40	9.58	0.20	33.47
KOV [15]	16.77	0.59	35.31	12.99	0.41	29.87
HDRCNN [1]	13.21	0.38	54.34	12.13	0.34	55.32
DRTMO [3]	-	-	-	11.4	0.28	58.85
DRHT [4]	-	0.93	61.51	-	-	-
FHDR/W	25.39	0.89	63.21	16.94	0.74	65.27
FHDR	**32.54**	**0.95**	**67.18**	**20.3**	**0.79**	**70.97**

Conclusion

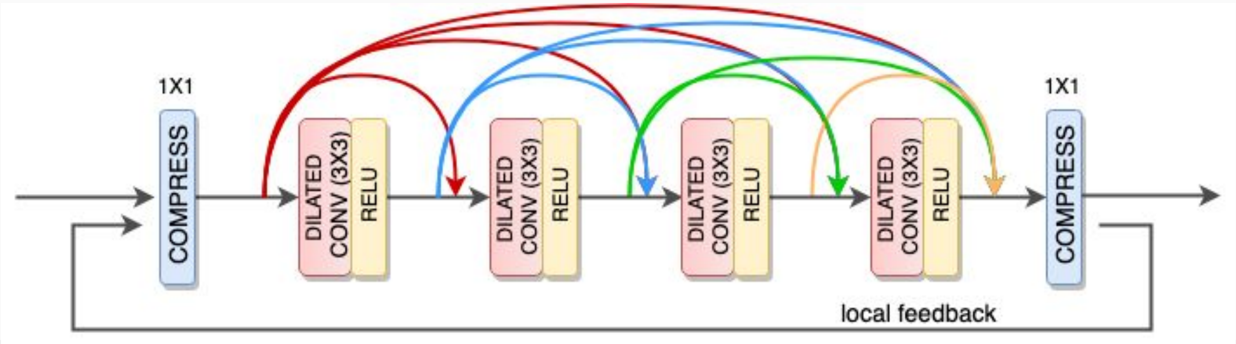
Conclusion

- Novel feedback network FHDR, to reconstruct an HDR image from a single exposure LDR image.



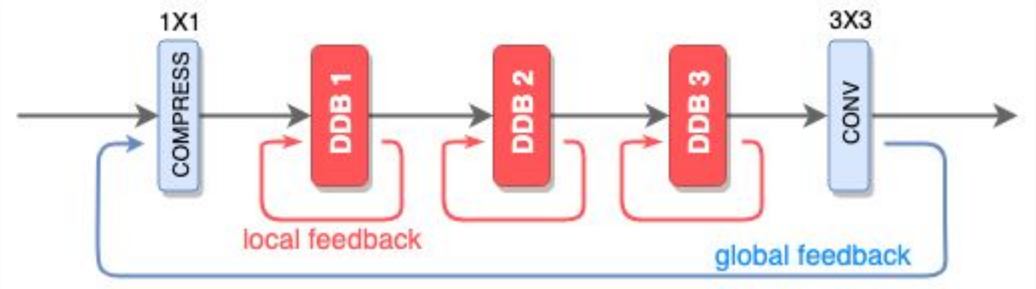
Conclusion

- Novel feedback network FHDR, to reconstruct an HDR image from a single exposure LDR image.
- The dense connections in the forward-pass enable feature-reuse, thus learning robust representations with minimum parameters.



Conclusion

- Novel feedback network FHDR, to reconstruct an HDR image from a single exposure LDR image.
- The dense connections in the forward-pass enable feature-reuse, thus learning robust representations with minimum parameters.
- Local and global feedback connections enhance the learning ability, guiding the initial low level features from the high level features.



Conclusion

- Novel feedback network FHDR, to reconstruct an HDR image from a single exposure LDR image.
- The dense connections in the forward-pass enable feature-reuse, thus learning robust representations with minimum parameters.
- Local and global feedback connections enhance the learning ability, guiding the initial low level features from the high level features.
- Iterative learning forces the network to create a coarse-to-fine representation which results in early reconstructions.

Conclusion

- Novel feedback network FHDR, to reconstruct an HDR image from a single exposure LDR image.
- The dense connections in the forward-pass enable feature-reuse, thus learning robust representations with minimum parameters.
- Local and global feedback connections enhance the learning ability, guiding the initial low level features from the high level features.
- Iterative learning forces the network to create a coarse-to-fine representation which results in early reconstructions.
- Extensive experiments demonstrate that the FHDR network is successfully able to recover the underexposed and overexposed regions outperforming state-of-the-art methods.

Thank you

Zeeshan Khan
zeeshank606@gmail.com

Mukul Khanna
mukul18khanna@gmail.com

Shanmuganathan Raman
shanmuga.r@iitgn.ac.in