No-Reference Metrics for Standard Dynamic Range and High Dynamic Range Image Content

Francesco Banterle, Alessandro Artusi, Alejandro Moreo,
Fabio Carrara, Paolo Cignoni

Why Do We Need Metrics?

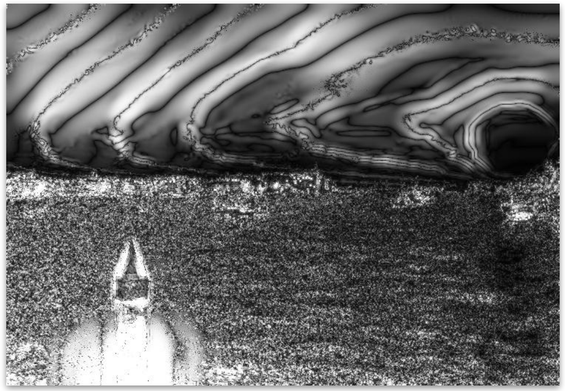
- In HDR/SDR Imaging, we need to determine and to understand what is happening during different steps of the pipeline:
  - **Acquisition**: we want to understand if there are artifacts due to acquisition or single image reconstruction;
  - **Compression**: we want small file size at maintaining high-quality;
  - **Tone mapping**: we want to adapt content for different display while keeping quality as it was “scene-referred”.

Reference Metrics

Reference Image

Distorted Image

Reference Metric



Probability Map

Q = 42.0

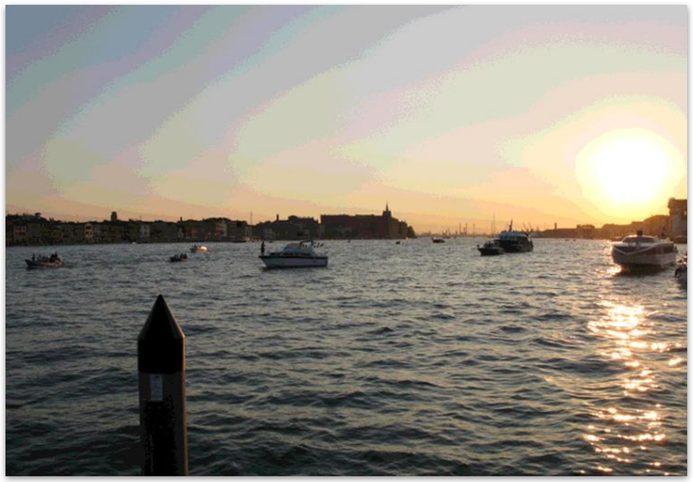
Quality Value

Reference Metrics: Current Limitations

- These models are very complex:
  - Difficult to port to GPUs with ease.
- They are computationally expensive; e.g., minutes of computations for a full HD image.
- Do we need a distortion map?
  - For most tasks we just need **a single value**!

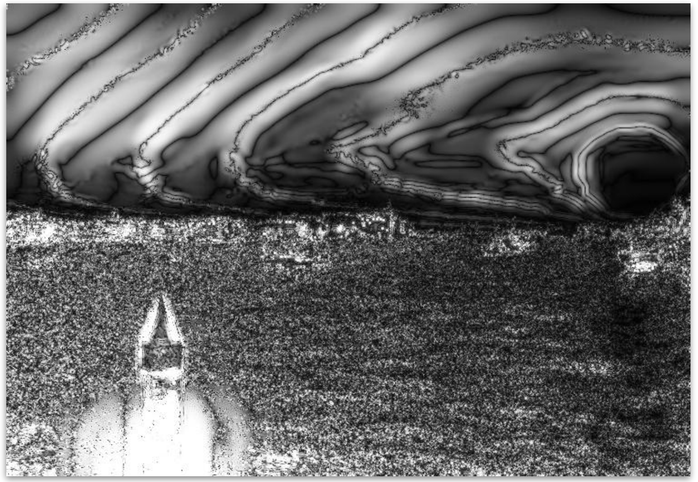
Going No-Reference

No-Reference Metrics



Distorted Image

No-reference Metric



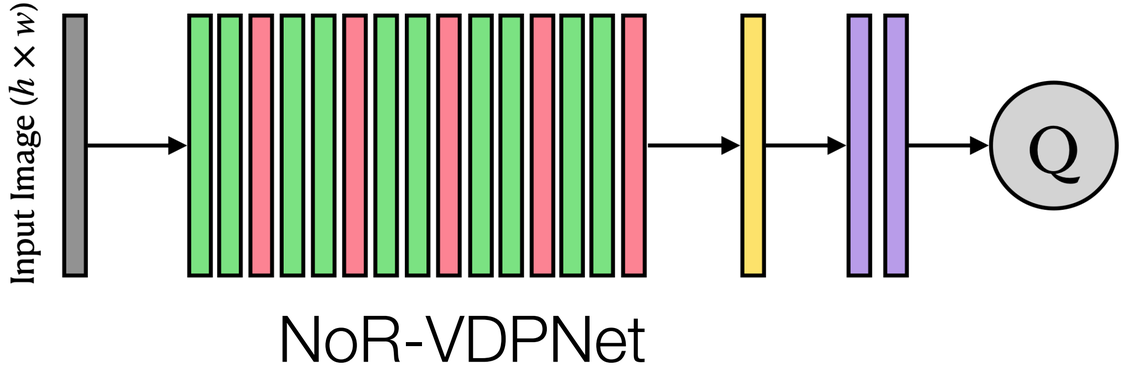
Probability Map

Q = 42.7

Quality Value

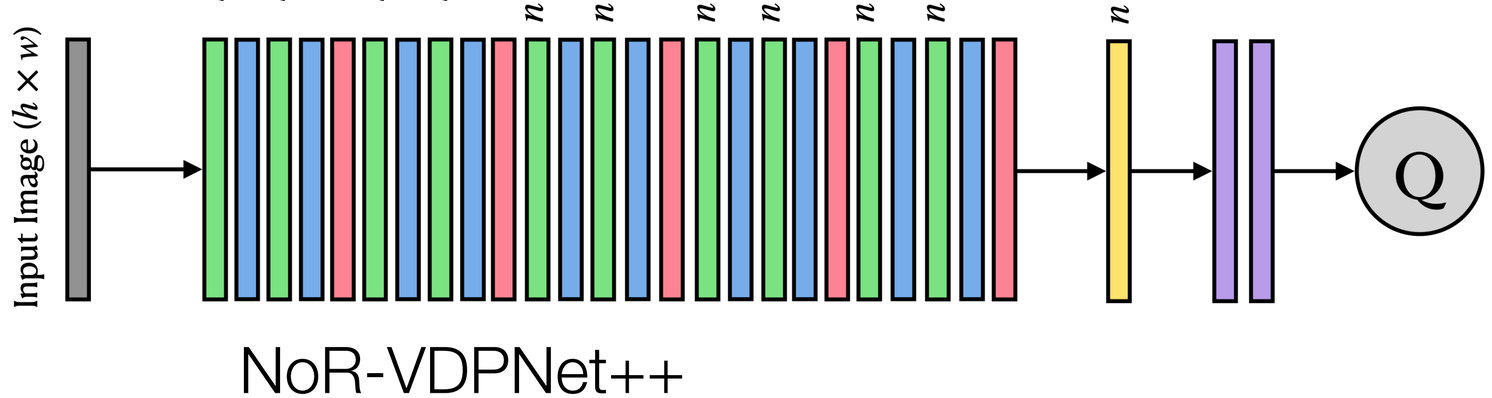
NoR-VDPNet(++): Architecture

Distorted Image

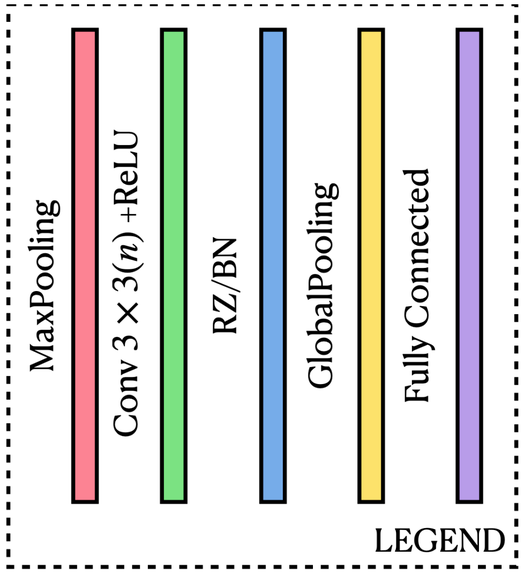


NoR-VDPNet

Distorted Image



NoR-VDPNet++



Training Set

Distorted Image

Reference Image

FR Metric

Q=49

Quality Value

TRAINING SAMPLE

Q=49

Input
Distorted Image

Target
Quality Value

NoRVDPNet(++): HDR-VDP2.2/TMQI Datasets

	TRAINING SET	VALIDATION SET	TEST SET	TOTAL
HDR-C (HDR-VDP2.2)	49.602	6.216	6.216	62.034
SDR-D (HDR-VDP2.2)	80.244	10.025	10.044	100.313
TMO (TMQI)	106.290	13.320	13.320	132.930
ITMO (HDR-VDP2.2)	106.290	13.320	13.320	132.930

NoR-VDPNet(++): SDR-D Dataset

REFERENCE SDR IMAGE

BLUR DISTORTION

NOISE DISTORTION

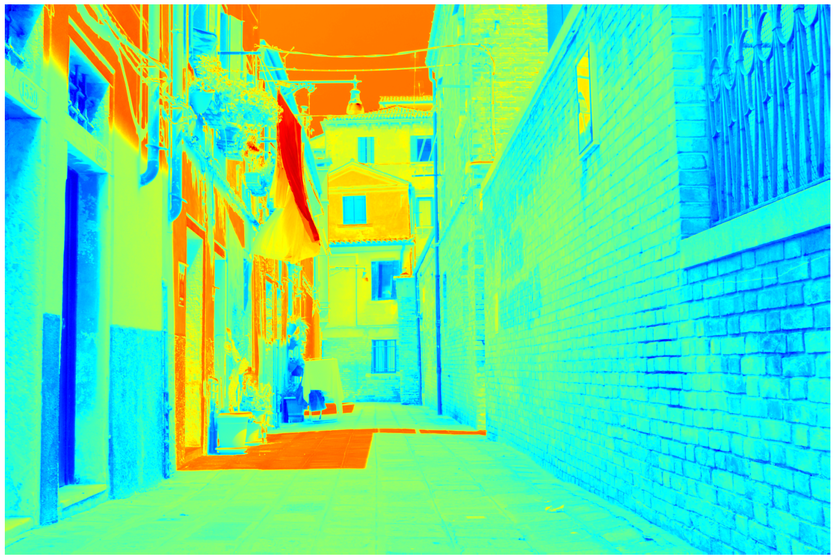
NoR-VDPNet(++): SDR-D Dataset

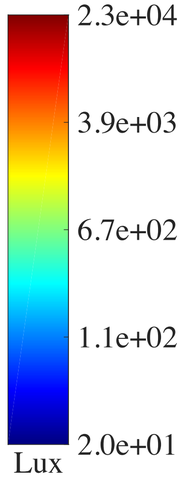
REFERENCE SDR IMAGE

QUANTIZATION DISTORTION

SIN GRATE DISTORTION

NoR-VDPNet(++): HDR-C Dataset





HDR Image

JPEG-XT:

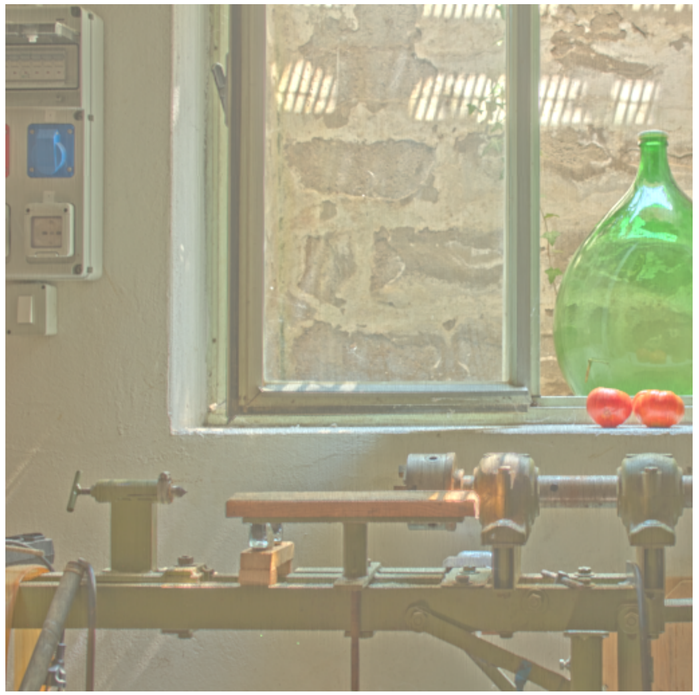
- Random Profile
- Random Residual Compressione

8-bit Layer

METADATA

NoRVDPNet(++): TMO Dataset

Drago et al. 2003



Durand and Dorsey 2002

Reinhard et al. 2002

18 tone mapping operators from the HDR-Toolbox: https://github.com/banterle/HDR_Toolbox/

NoRVDPNet(++): ITMO Dataset

Input SDR Image

Eilertsen et al. 2017
(tonemapped)

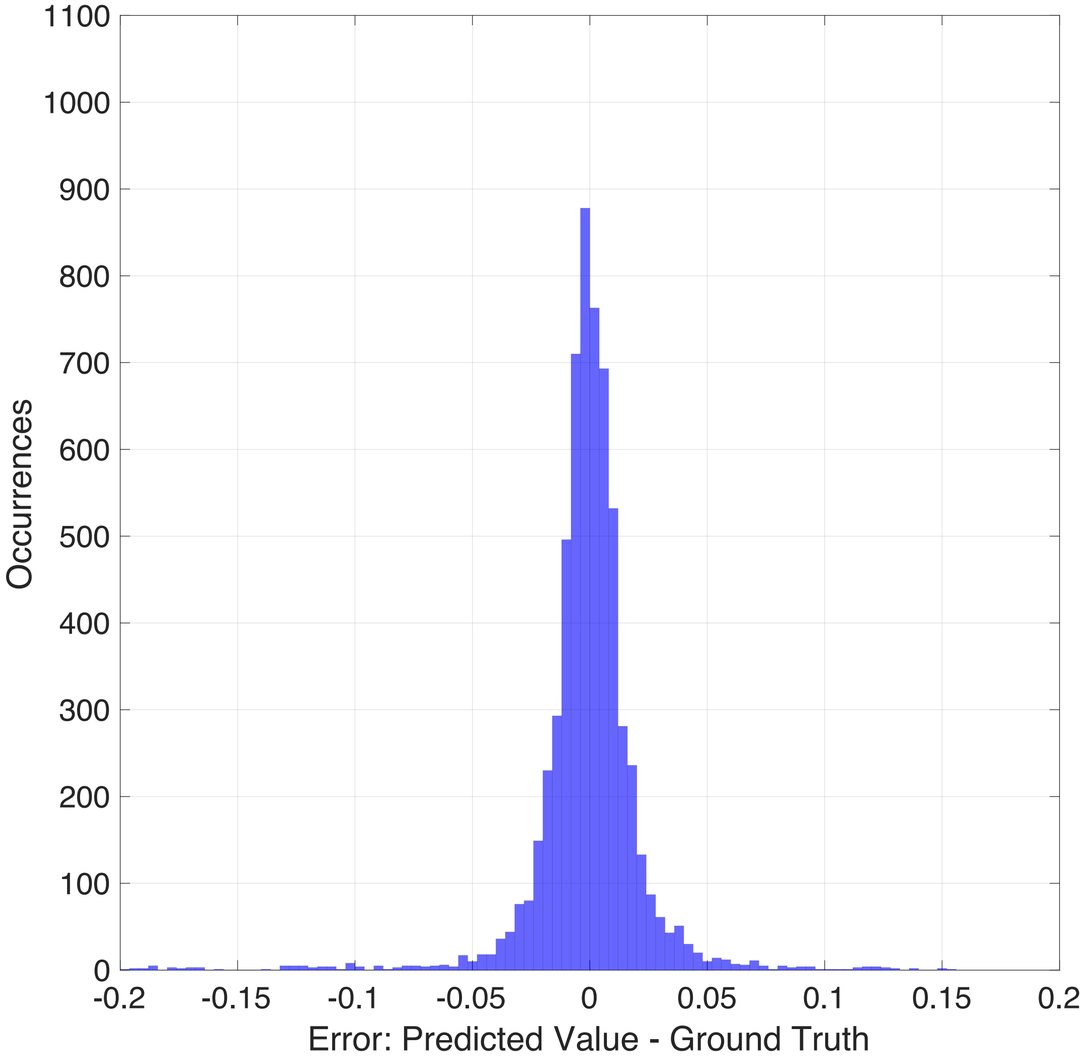
Santos et al. 20202
(tonemapped)

6 inverse tone mapping operators 4 available in the HDR-Toolbox: https://github.com/banterle/HDR_Toolbox/

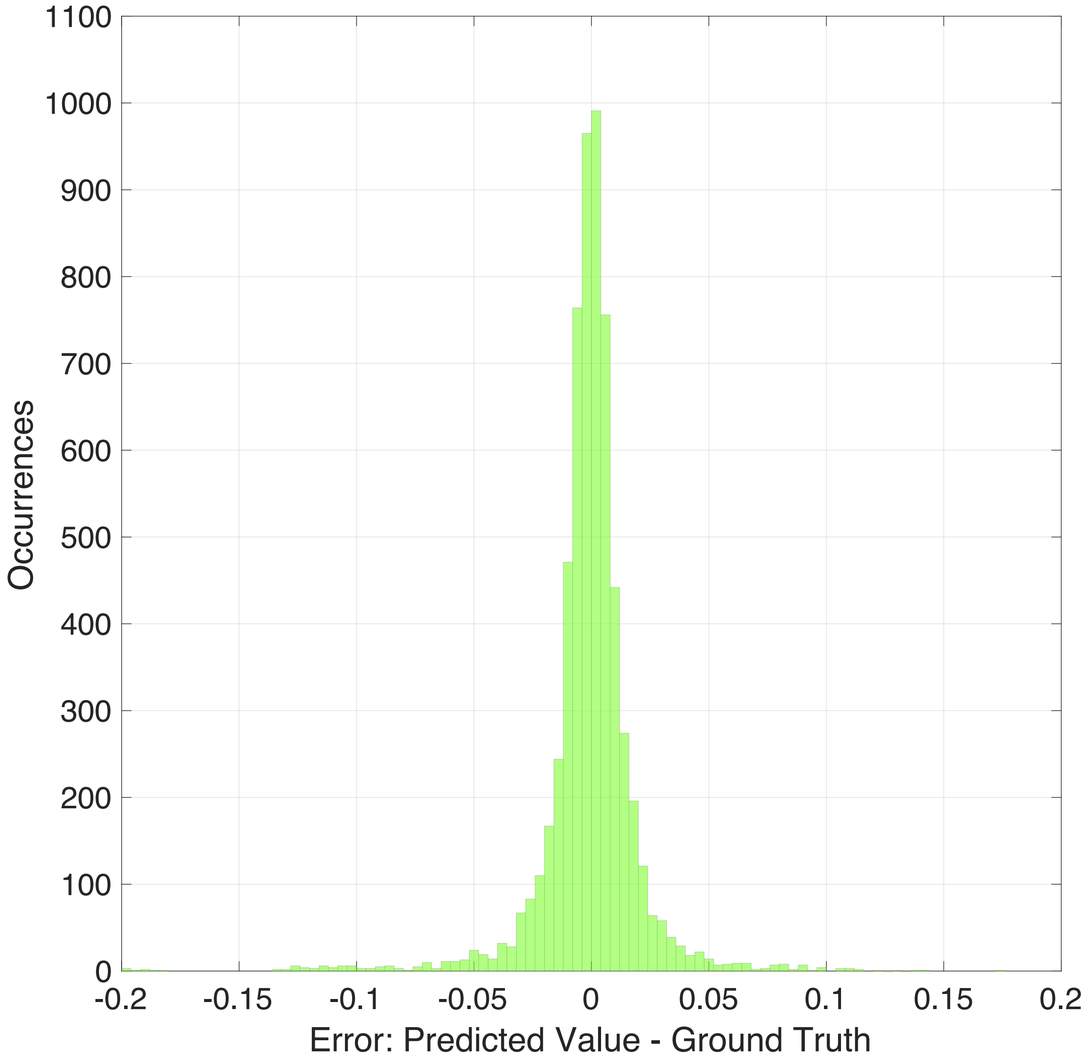
NoR-VDPNet(++): Loss and Encoding

- Loss is a classic MSE; it works well for predicting quantitative values:
- Encoding:
  - SDR Images: linear scaling to fit the range $[0,1]$
  - HDR Images: $\log_{10}(x+1)$

Results: HDR-C Test Set

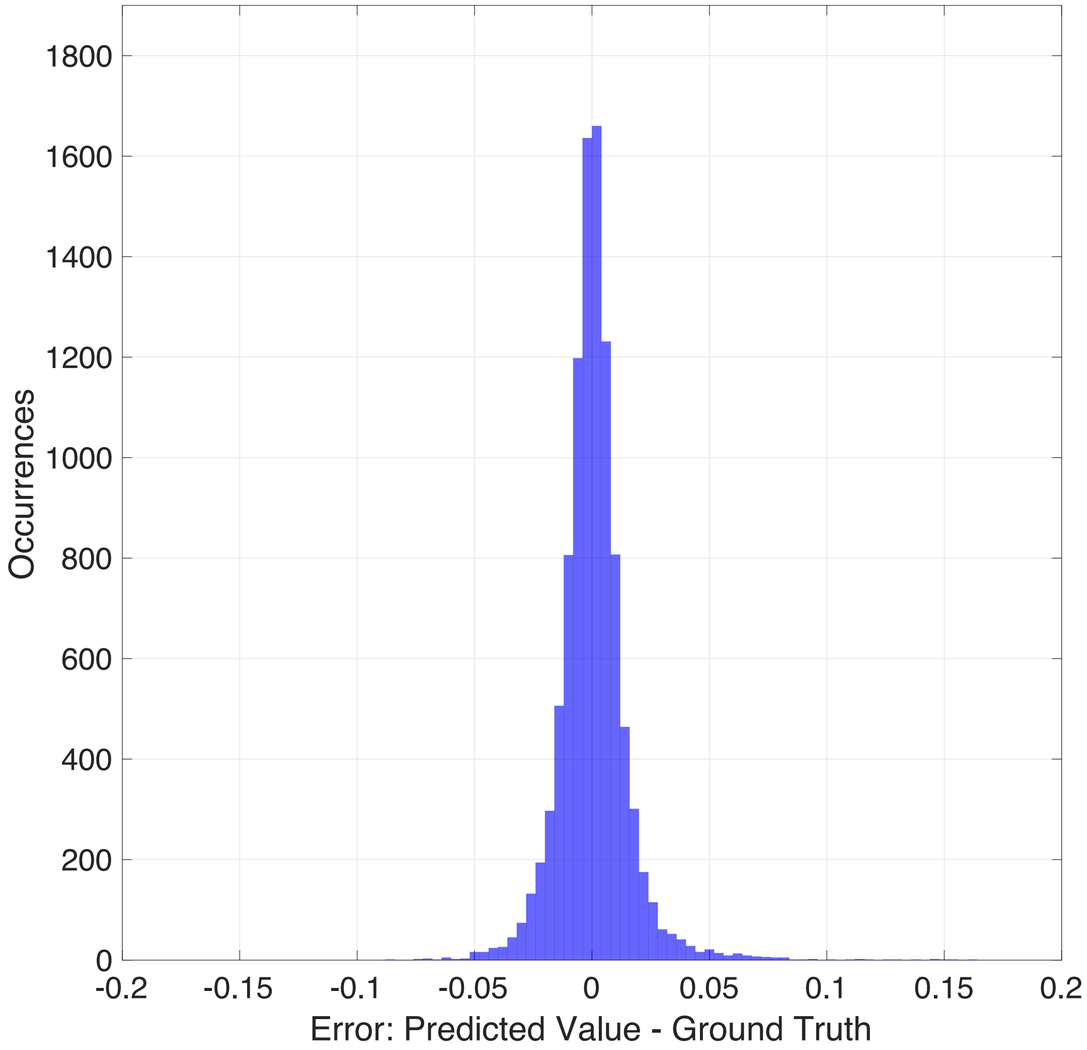


NoRVDPNet

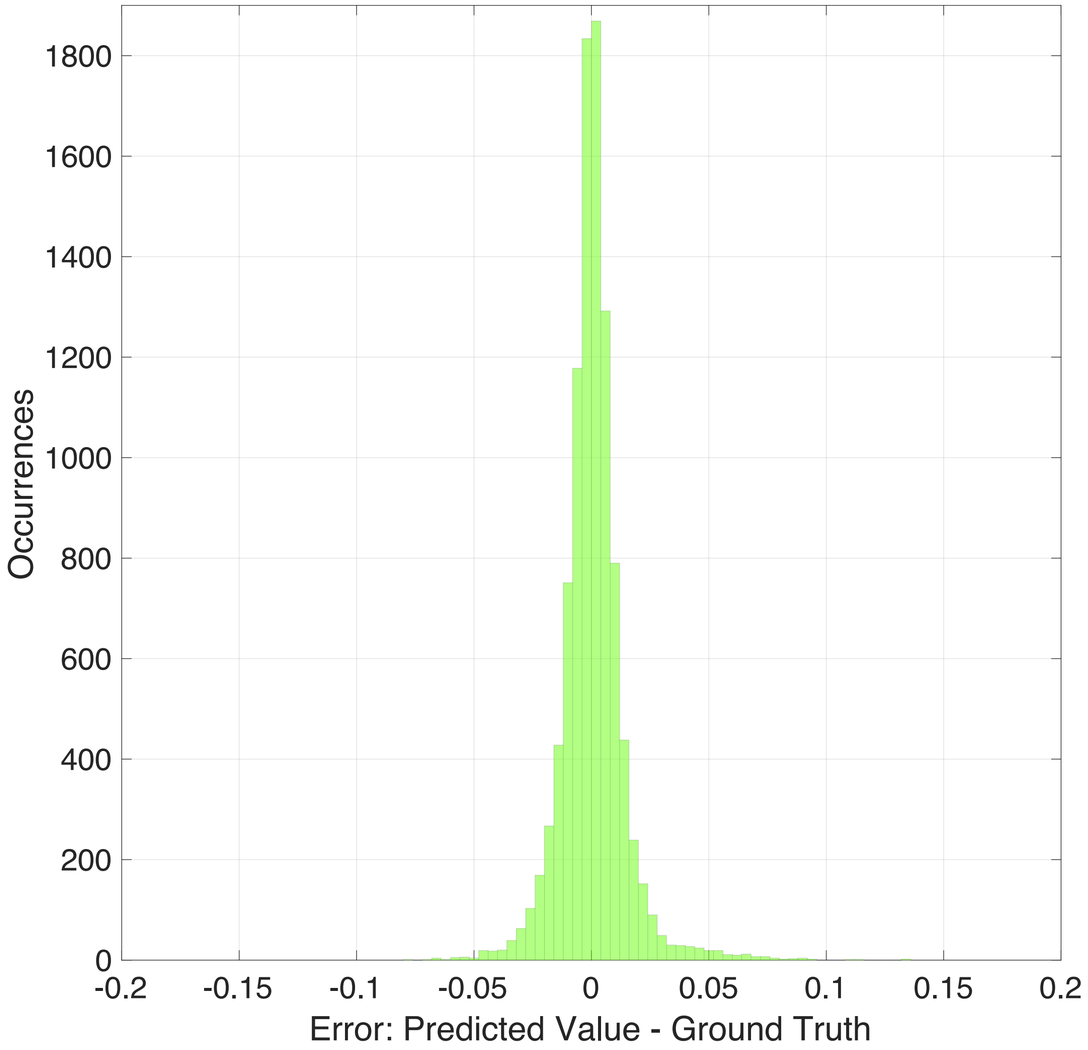


NoRVDPNet++

Results: SDR-D Test Set

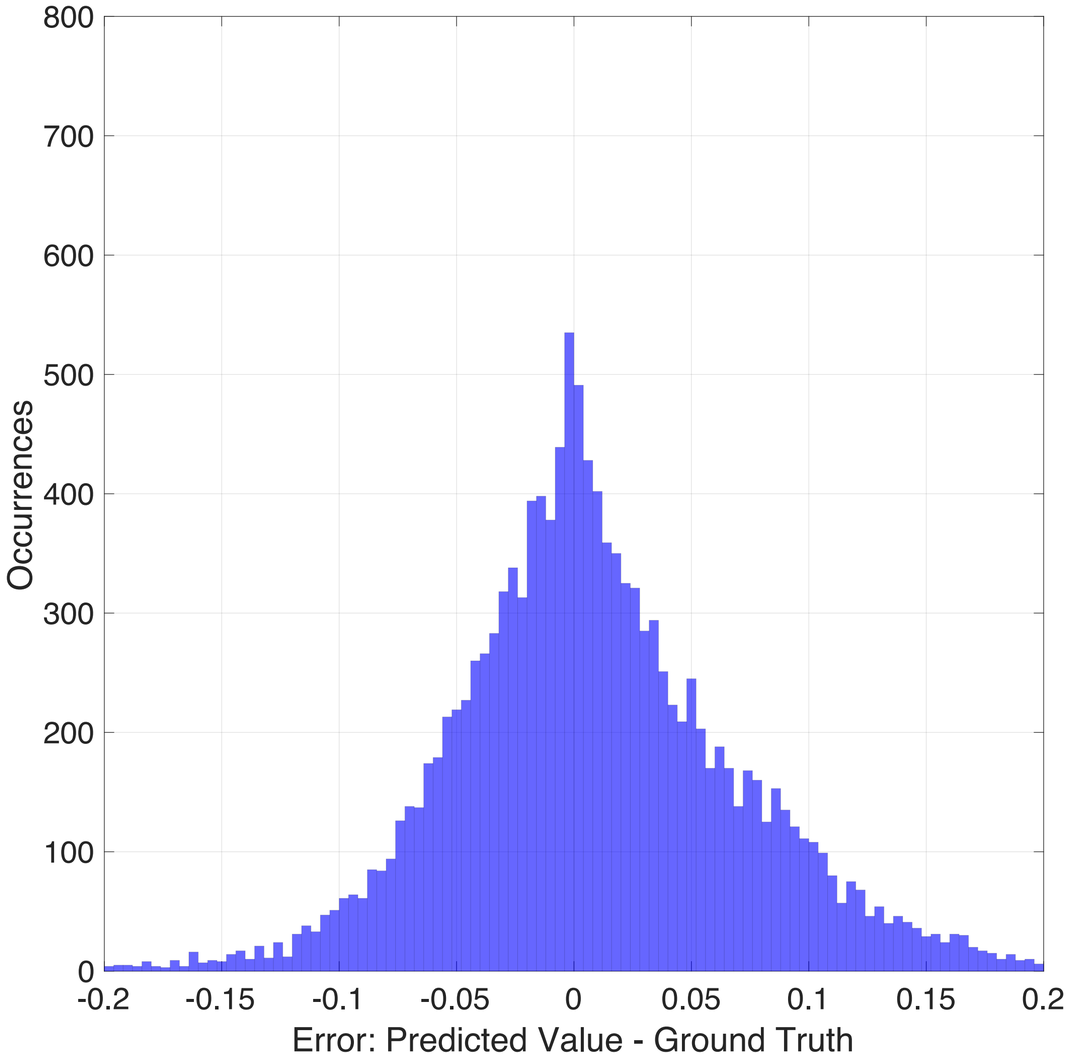


NoRVDPNet

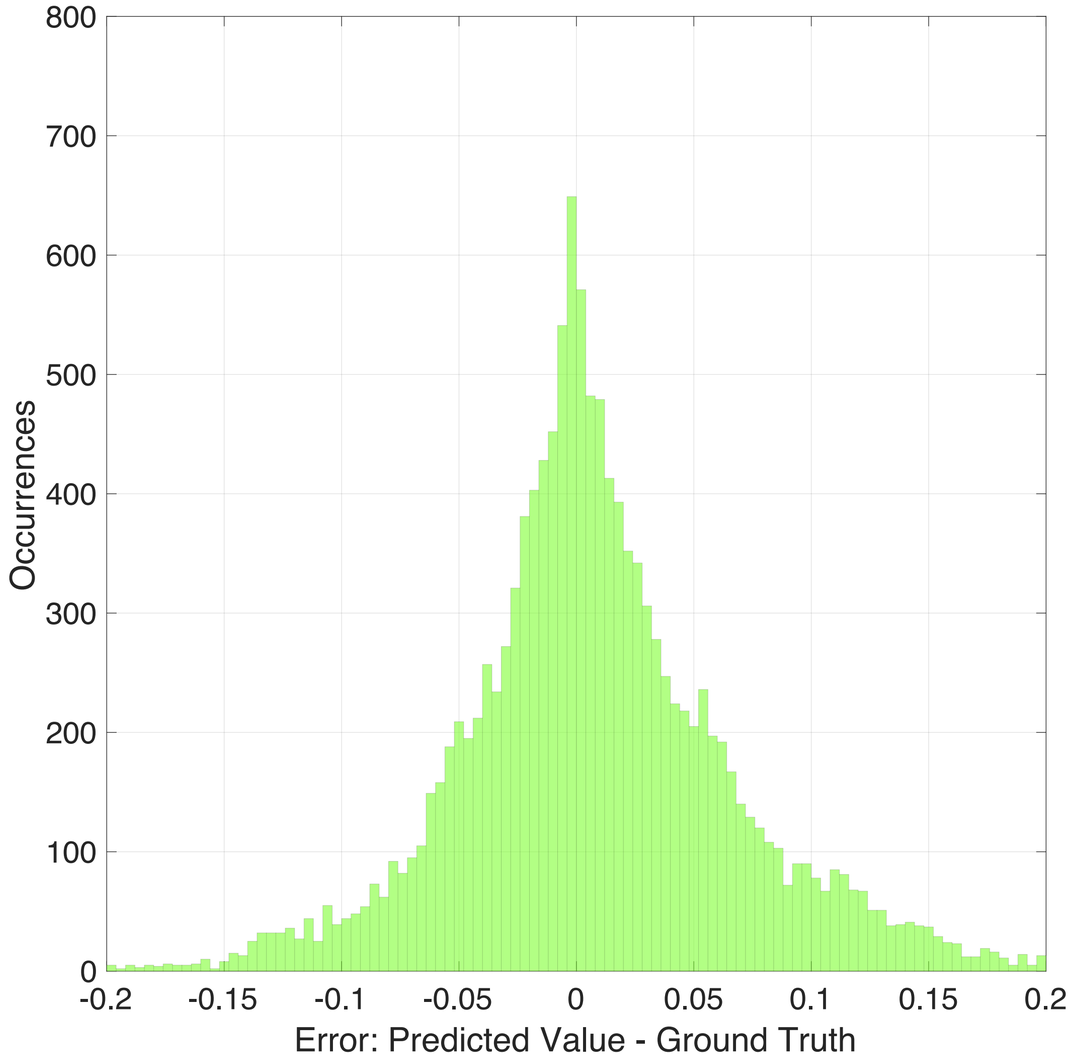


NoRVDPNet++

Results: ITMOS Test Set

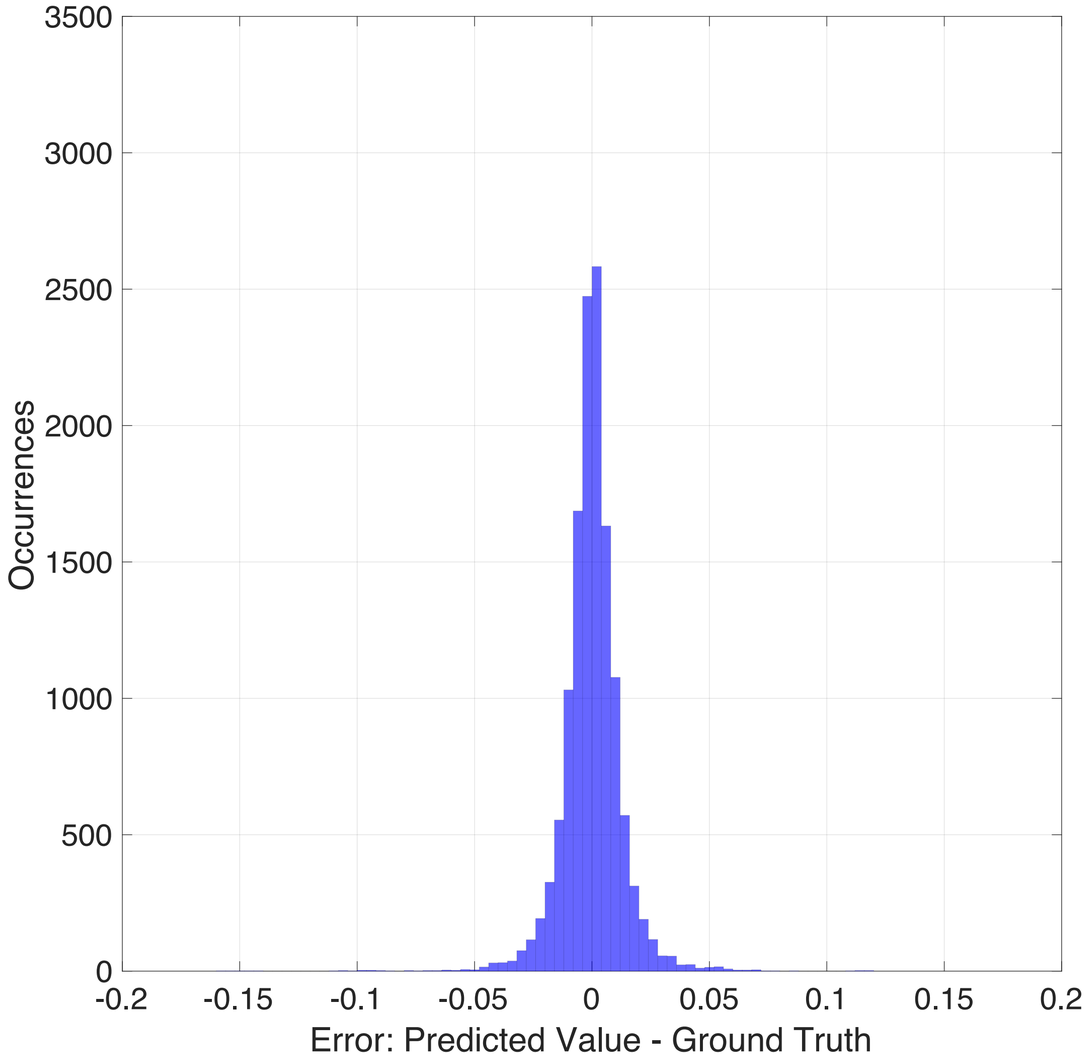


NoRVDPNet

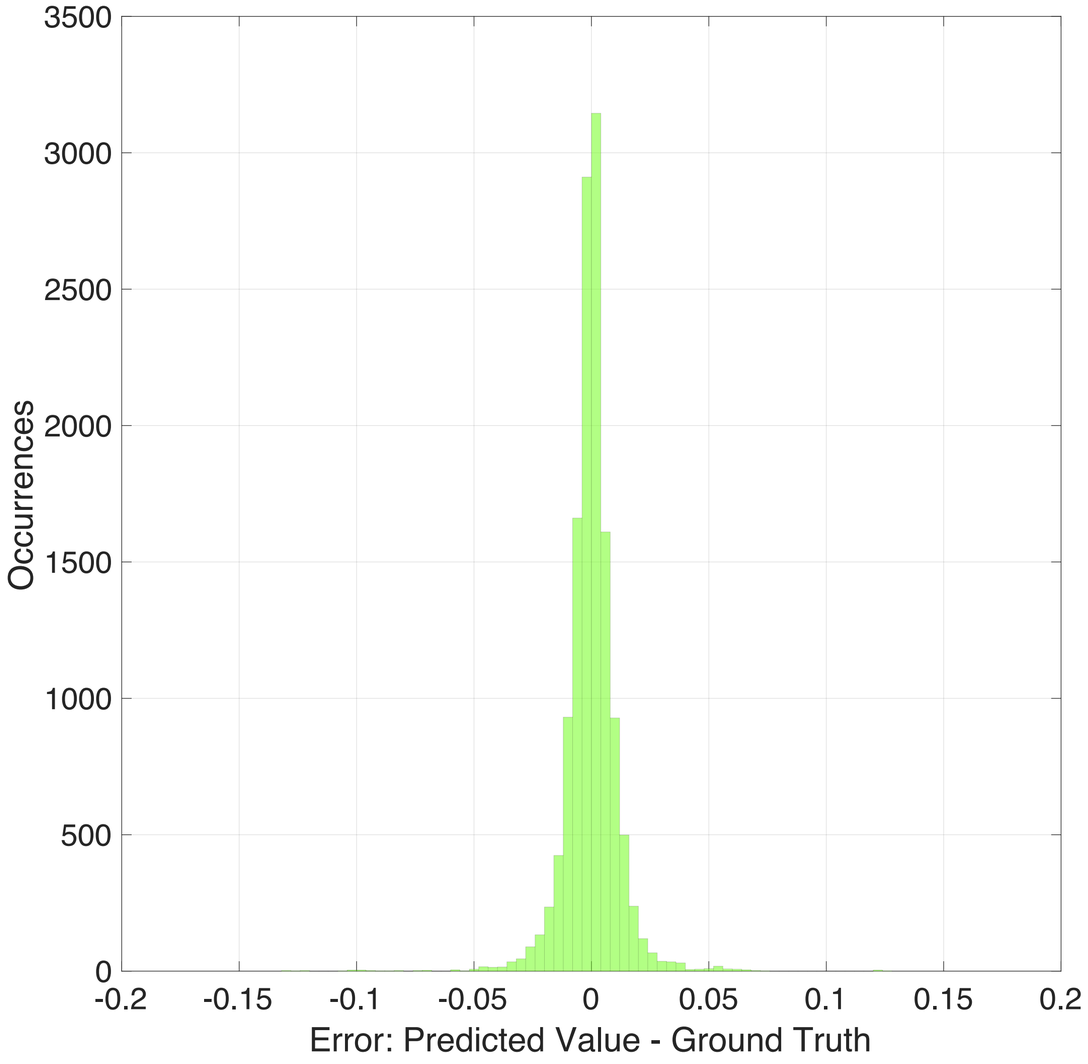


NoRVDPNet++

Results: TMOS Test Set

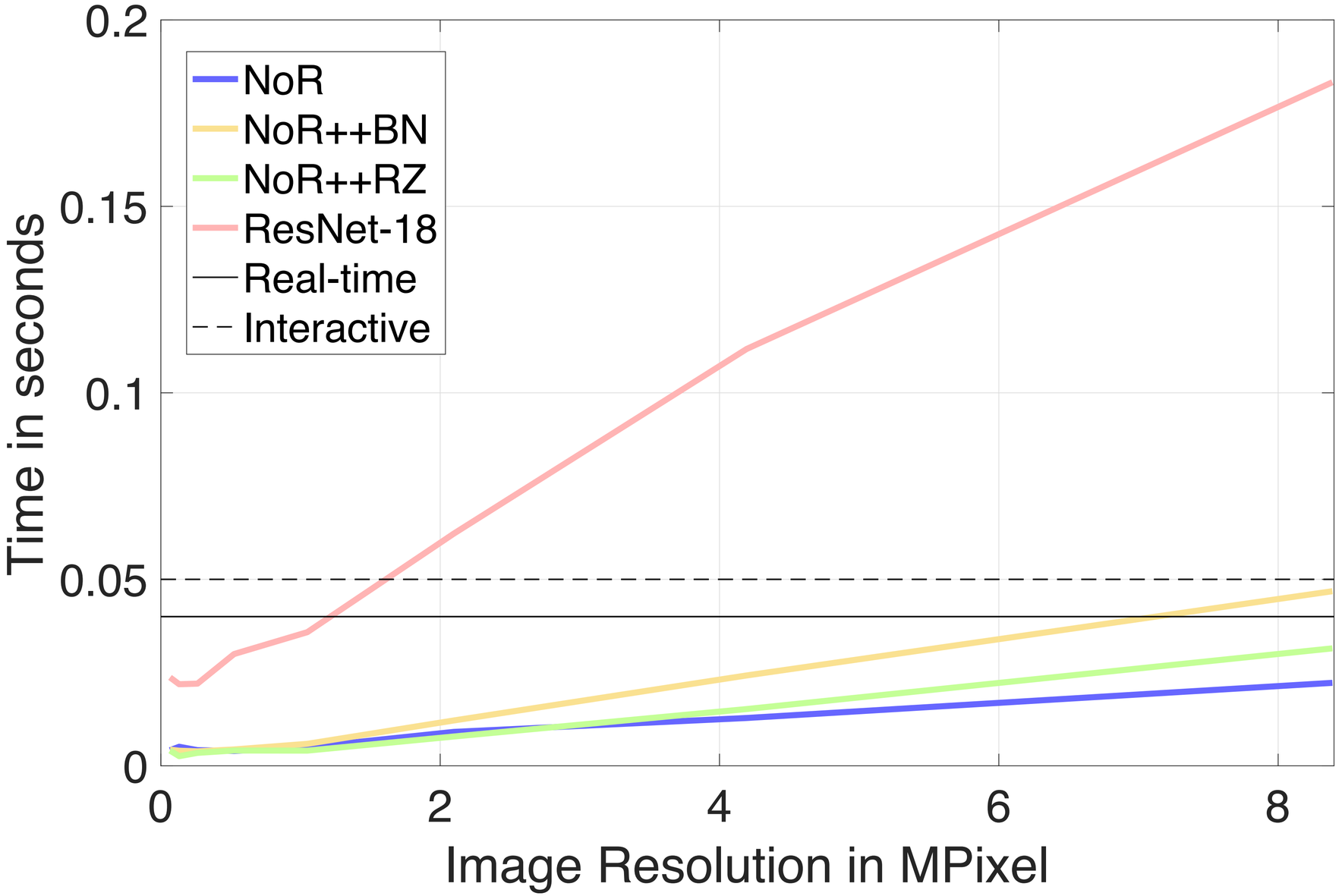


NoRVDPNet



NoRVDPNet++

Timings


0.2
0.15
0.1
0.05
0
Time in seconds
NoR
NoR++BN
NoR++RZ
ResNet-18
Real-time
Interactive
0
2
4
6
8
Image Resolution in MPixel

NoR-VDPNet(++): Conclusions

- We can go from reference to no-reference;
- When we model several distortions we have a larger error than a single distortion;
- Layer normalization increases quality;
- This scheme works for TMQI-I (SSIM-based);
- Still real-time performance.

Applications

Applications: TMO Optimization Task

Input HDR image

TMO

Tone Mapped Image

DEEP NR-METRIC

Q

Q > Thr

NO

YES

Tone Mapped Image

Applications: Optimized TMO

TMO without optimized parameters

TMO with optimized parameters

Video Courtesy of Jan Fröhlich - Stuttgart HDR Video Dataset

Application: Optimized TMO

(a) $\hat{Q} = 0.903$ /$Q = 0.885$

(b) $\hat{Q} = 0.906$/$Q = 0.930$

(c) $\hat{Q} = 0.933$ /$Q = 0.914$

(d) $\hat{Q} = 0.918$/$Q = 0.903$

(e) $\hat{Q} = 0.902$/$Q = 0.889$

(f) $\hat{Q} = 0.841$/$Q = 0.771$

(g) $\hat{Q} = 0.951$/$Q = 0.831$

(h) $\hat{Q} = 0.875$/$Q = 0.909$

(i) $\hat{Q} = 0.951$/$Q = 0.967$

(j) $\hat{Q} = 0.958$/$Q = 0.974$

(k) $\hat{Q} = 0.967$/$Q = 0.976$

(l) $\hat{Q} = 0.997$/$Q = 0.979$

Applications: JPEG-XT Compression Task

Input HDR image

TMO

Tone Mapped
+
Metadata

Reconstructed
HDR Image

DEEP
NR-METRIC

Q

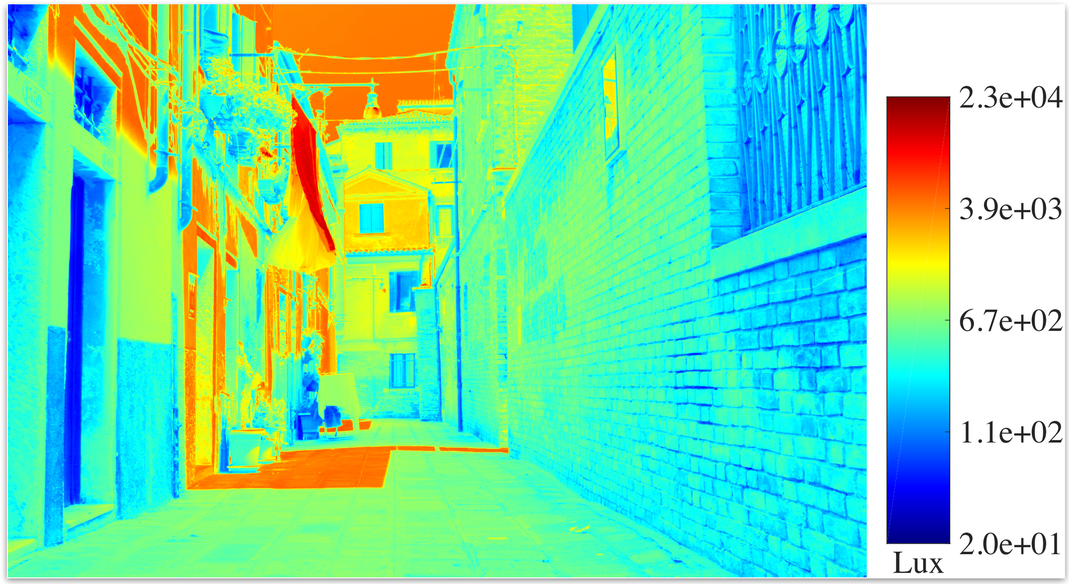
Q > Thr

NO

YES

Tone Mapped
+
Metadata

Applications: Results JPEG-XT Compression



Input HDR image

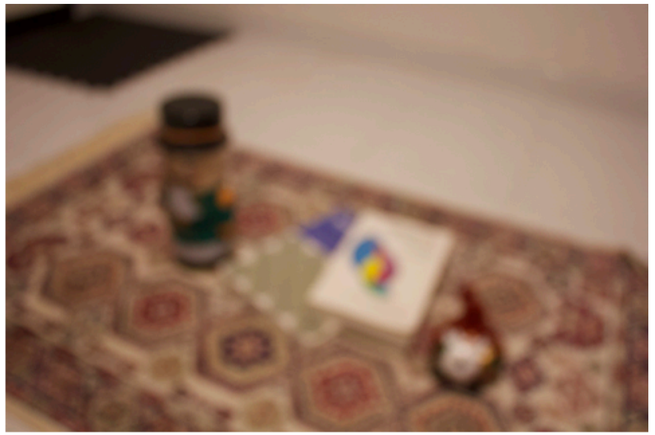
Reinhard et al.'s TMO
optimized with NoRVDPNet

Tone Mapped HDR image
for JPEG-XT

Applications: Photo Selection

Q=86.99

Q=86.92



Q=56.46

Q=91.39

Q=76.26

Q=59.9

Applications: Photo Selection

Q=86.99

Q=86.92

Q=56.46

Q=91.39

Q=76.26

Q=59.9

Future Directions

Future Directions

- Going in the temporal domain.
- Extend approaches to perceptual uniform domains.
- Mix perceptual experiments results and metrics.

Thank you for your attention!

Please contact us at:

a.artusi@cyens.org.cy francesco.banterle@isti.cnr.it

or visit us:

https://deepacamera.org.cy http://vcg.isti.cnr.it

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No. 739578 and No. 820434, and from the Government of the Republic of Cyprus through the Deputy Ministry of Research, Innovation and Digital Policy.

This project is co-financed by the European Regional Development Fund and the Republic of Cyprus through the Research and Innovation Foundation (POST-DOC/0916/0034).