Sequential Restoration for Visual Recognition: an Empirical Analysis

VITA

Visual Informatics @Texas A&M

STRUCT

Spatial and Temporal Restoration, Understanding and Compression Team @PKU

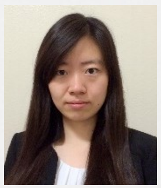
UG2 PRIZE CHALLENGE

TAMU-PKU Union (TPU)

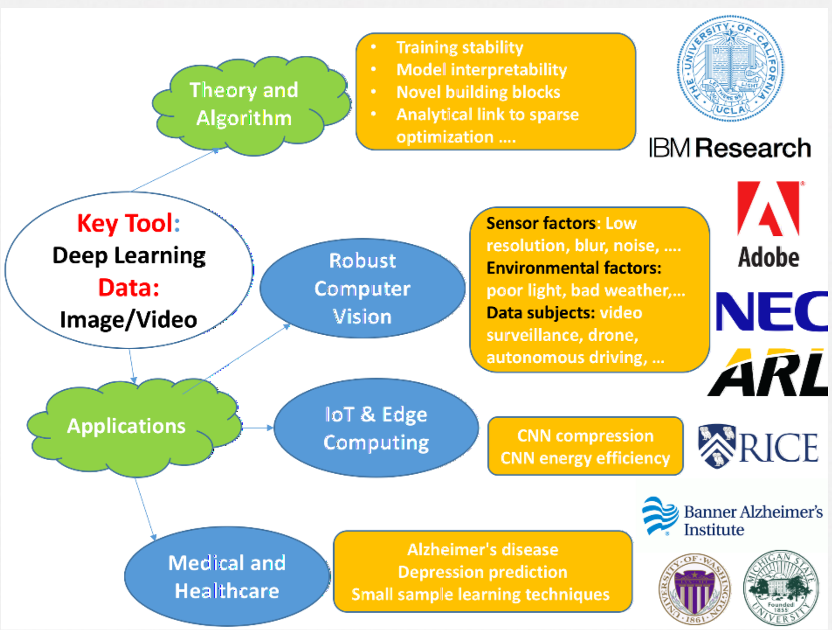
Code: https://github.com/yyvettey/TAMU-PKU-UG2

June 2018

Visual Informatics Group @TAMU (VITA)



VITA

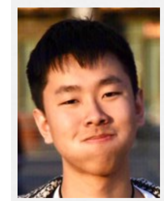


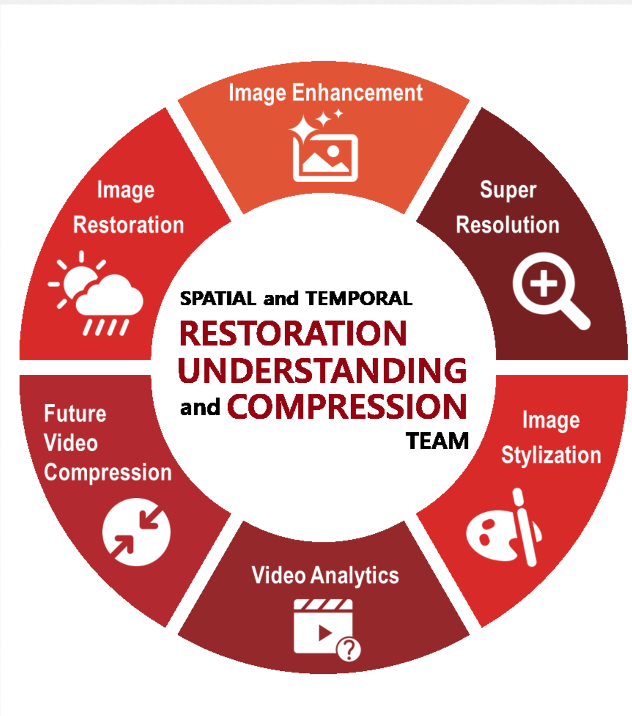
- http://www.atlaswang.com/group.html

STRUCT @PKU

Spatial and Temporal **Restoration**, **Understanding** and **Compression** Team

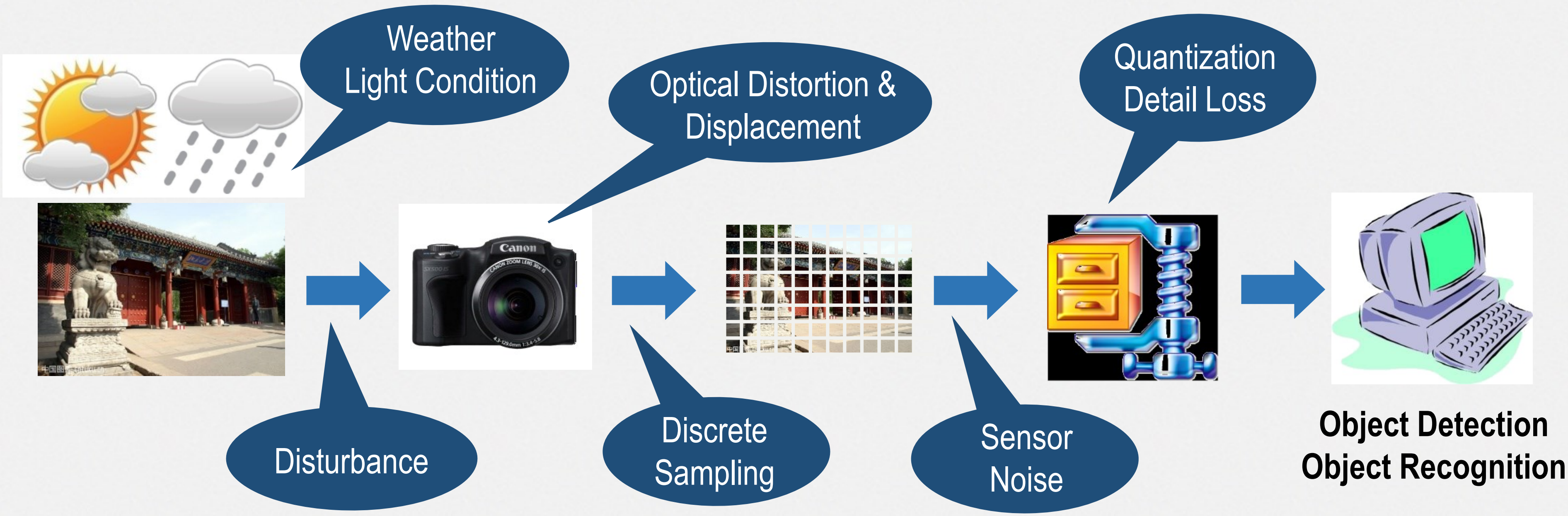




- http://www.icst.pku.edu.cn/struct/

Image Degradations

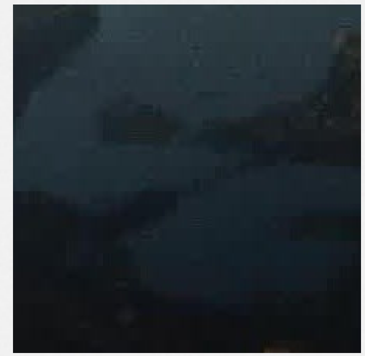


- Degradation types observed in the UG2 dataset
- Techniques used to restore image degradations
- Combination of techniques and restoration performance

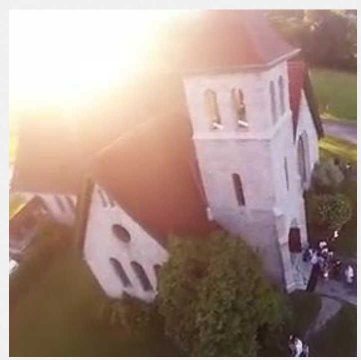
Degradations Observed in UG2 Dataset

- Low Resolution
- Low Light
- Noise
- Blocking
- Blurring
- Reflection
- Lens Flare
- Over Exposure
- Under Exposure
- Turbulence
- Annotation
- Rain Drop


Blurring


Low Light


Low Resolution


Over-Exposure


Blocking


Noise & Annotation

Existing Restoration Techniques

Atmosphere

- Flare removal
- Disturbance removal
- Light adjustment
- Haze removal
- Raindrop removal

Capturing

- Super-resolution
- Deblurring
- Shaking removal
- Digital removal

Post-Processing

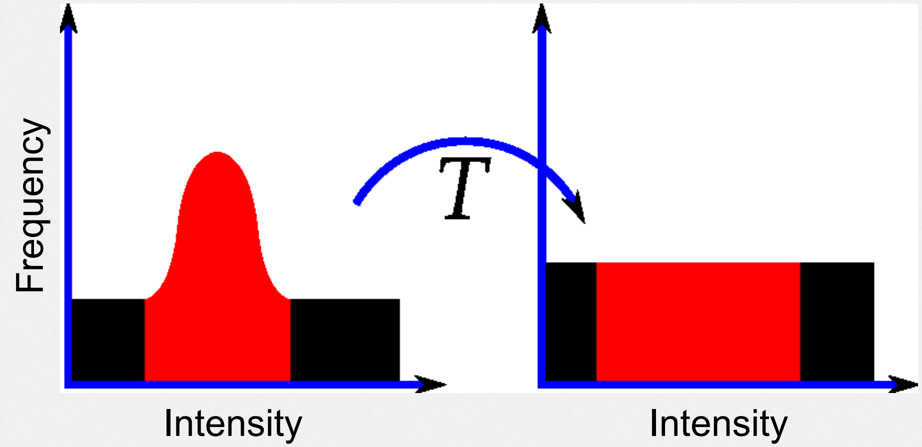
- Deblocking
- Denoising
- Contrast adjustment
- Sharpening
- High dynamic range

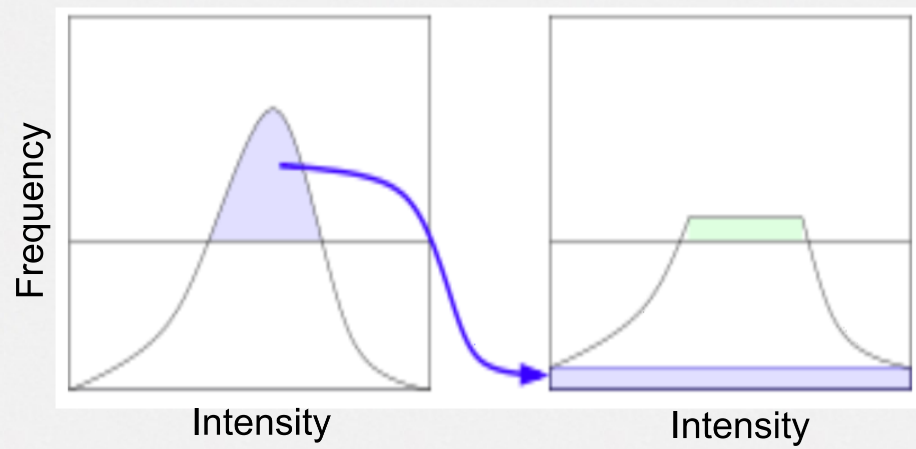
Our Choice

- Light adjustment, super-resolution, deblurring, denoising+SR, HDR, deblocking

Histogram Equalization

- **Vanilla Histogram Equalization**
  - A a technique for adjusting image intensities to enhance contrast
  - Find a mapping, such that the intensities can be evenly distributed on the histogram
- **Adaptive Histogram Equalization**
  - Mapping is based on local neighbourhood region rather than the whole image
- **Contrast Limited Adaptive Histogram Equalization (CLAHE)**
  - With limited contrast amplification





https://en.wikipedia.org/wiki/Adaptive_histogram_equalization

Histogram Equalization

- **Visual Results**

Without

With

Histogram Equalization

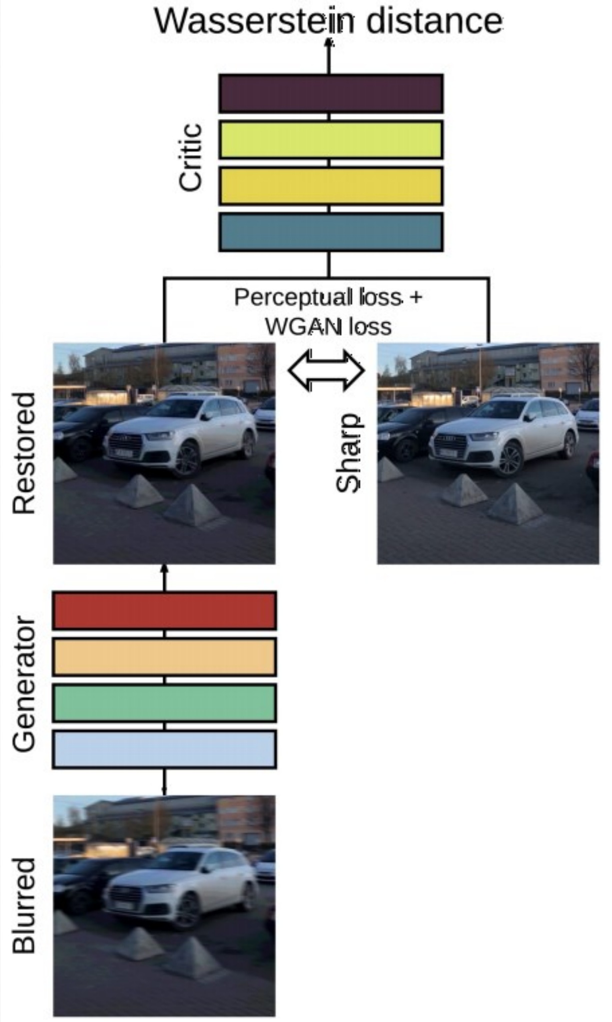
■ Results on Validation Set

M1	Baseline	HistEQ	Gain
UAV	0.27477	0.29284	0.01806
	0.18755	0.23147	0.04392
	0.22455	0.22640	0.00186
	0.23580	0.24137	0.00557
Glider	0.38176	0.35876	-0.02299
	0.38023	0.37399	-0.00624
	0.37673	0.39912	0.02238
	0.41922	0.38511	-0.03411
Ground	0.45025	0.49567	0.04542
	0.49486	0.49511	0.00025
	0.43297	0.43728	0.00431
	0.43555	0.49389	0.05834

M2	Baseline	HistEQ	Gain
UAV	0.04206	0.06557	0.02351
	0.02536	0.03427	0.00891
	0.03427	0.04590	0.01163
	0.01979	0.02301	0.00322
Glider	0.02406	0.01553	-0.00853
	0.05284	0.02223	-0.03061
	0.02436	0.03487	0.01051
	0.04431	0.02467	-0.01964
Ground	0.41492	0.48269	0.06777
	0.44452	0.43570	-0.00882
	0.39982	0.36712	-0.03269
	0.36996	0.42922	0.05925

DeblurGAN

- GAN based methods
  - Generator generate deblurred images
  - Discriminator distinguish between them real sharp image and deblurred image
- Generator Network
  - 9 ResBlocks
- Discriminator Network
  - PatchGAN Discriminator
- Trained on GOPRO Video Deblurring Dataset (2103 pairs)
- Perceptual Loss + Wasserstein GAN loss
- 5 times faster than competitor – DeepDeblur



Kupyn O, Budzan V, Mykhailych M, Mishkin D, Matas J. DeblurGAN: Blind Motion Deblurring Using Conditional Adversarial Networks. CVPR 2018

DeblurGAN

- **Visual Results**

Without

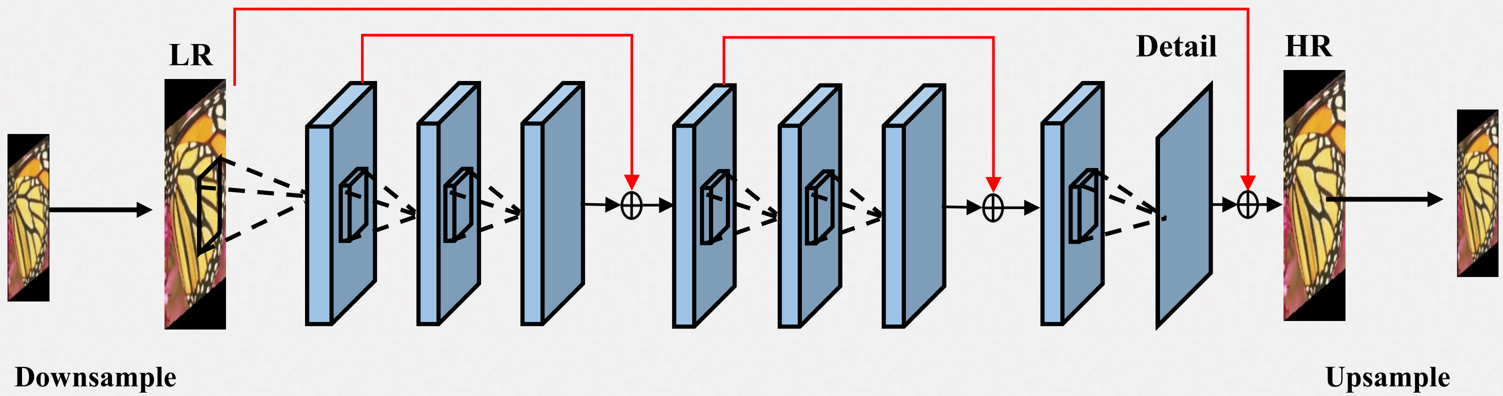
With

DeblurGAN

■ Results on Validation Set

M1	Baseline	DeblurGAN	Gain	M2	Baseline	DeblurGAN	Gain
UAV	0.27477	0.28282	0.00804	UAV	0.04206	0.04701	0.00495
	0.18755	0.20166	0.01410		0.02536	0.03142	0.00606
	0.22455	0.24508	0.02054		0.03427	0.04454	0.01027
	0.23580	0.24347	0.00767		0.01979	0.02326	0.00346
Glider	0.38176	0.36714	-0.01462	Glider	0.02406	0.03167	0.00761
	0.38023	0.35511	-0.02513		0.05284	0.04949	-0.00335
	0.37673	0.36744	-0.00929		0.02436	0.02650	0.00213
	0.41922	0.36881	-0.05040		0.04431	0.05025	0.00594
Ground	0.45025	0.44452	-0.00573	Ground	0.41492	0.39632	-0.01860
	0.49486	0.46140	-0.03345		0.44452	0.38654	-0.05799
	0.43297	0.40291	-0.03006		0.39982	0.36003	-0.03979
	0.43555	0.44665	0.01110		0.36996	0.38345	0.01348

Super-Resolution CNN



- **Super-Resolution**
  - Image resolution unchanged, we use Super-Resolution to enhance image details
- **6L ResNet**
  - 6-Layers Fully Convolutional Networks with Skip-Connection
  - Trained on standard image SR dataset with only 91 images, with 3X SR
  - Bicubic Down/Upsampling (smooth out image to remove artifacts)

Chao Dong, Chen Change Loy, Kaiming He, Xiaoou Tang. Learning a Deep Convolutional Network for Image Super-Resolution, ECCV 2014

Super-Resolution CNN

- **Visual Results**

Without

With

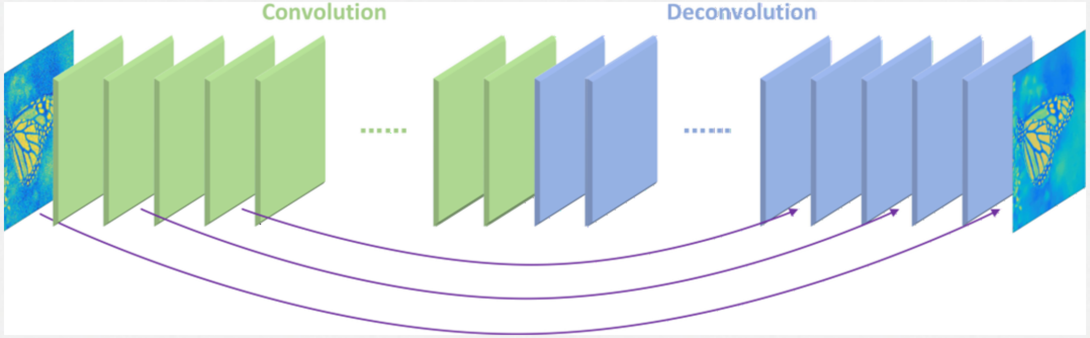
Super-Resolution CNN

▪ Results on Validation Set

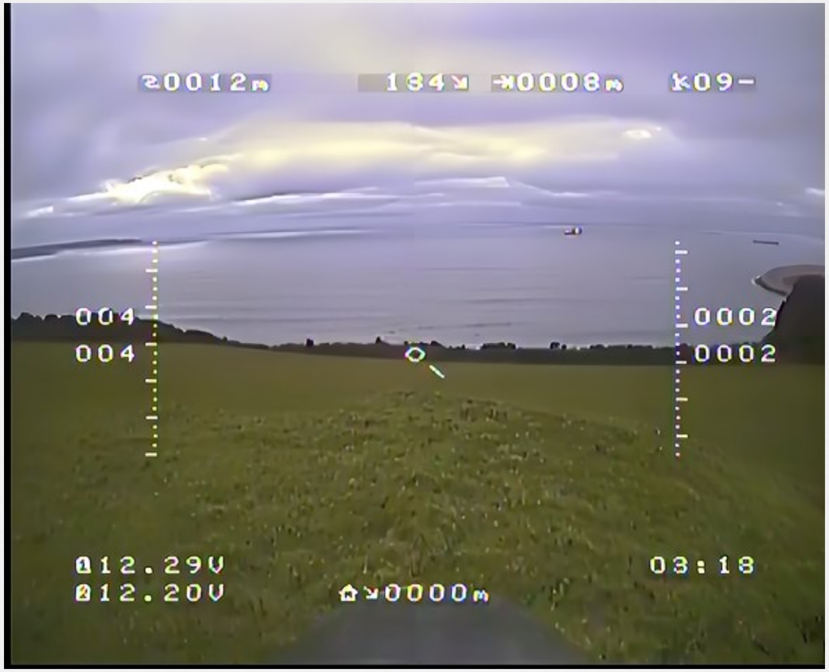
M1	Baseline	SRCNN	Gain
UAV	0.27477	0.27836	0.00359
	0.18755	0.18904	0.00148
	0.22455	0.22863	0.00408
	0.23580	0.23828	0.00247
Glider	0.38176	0.38282	0.00107
	0.38023	0.37490	-0.00533
	0.37673	0.36805	-0.00868
	0.41922	0.41160	-0.00761
Ground	0.45025	0.44843	-0.00182
	0.49486	0.49450	-0.00035
	0.43297	0.43322	0.00025
	0.43555	0.43702	0.00147

M2	Baseline	SRCNN	Gain
UAV	0.04206	0.04231	0.00025
	0.02536	0.02474	-0.00062
	0.03427	0.03365	-0.00062
	0.01979	0.02140	0.00161
Glider	0.02406	0.02695	0.00289
	0.05284	0.05193	-0.00091
	0.02436	0.02497	0.00061
	0.04431	0.04568	0.00137
Ground	0.41492	0.41031	-0.00461
	0.44452	0.43940	-0.00512
	0.39982	0.40017	0.00035
	0.36996	0.36783	-0.00213

Deblocking



- **RED-Net**
  - Finetuned using VGG-19 perceptual loss



Mao XJ, Shen C, Yang YB. Image restoration using convolutional auto-encoders with symmetric skip connections. *NIPS,* 2016.

Deblocking

■ Results on Validation Set

M1	Baseline	RED-Net	Gain
UAV	0.27477	0.18557	-0.08920
	0.18755	0.14029	-0.04726
	0.22455	0.17185	-0.05270
	0.23580	0.19720	-0.03860
Glider	0.38176	0.32953	-0.05223
	0.38023	0.33607	-0.04416
	0.37673	0.30836	-0.06837
	0.41922	0.37811	-0.04111
Ground	0.45025	0.39231	-0.05794
	0.49486	0.41797	-0.07689
	0.43297	0.41209	-0.02088
	0.43555	0.41360	-0.02195

M2	Baseline	RED-Net	Gain
UAV	0.04206	0.02387	-0.01819
	0.02536	0.01534	-0.01002
	0.03427	0.01200	-0.02227
	0.01979	0.01744	-0.00235
Glider	0.02406	0.02406	0.00000
	0.05284	0.05284	0.00000
	0.02436	0.02999	0.00563
	0.04431	0.06152	0.01721
Ground	0.41492	0.34492	-0.07000
	0.44452	0.34467	-0.09985
	0.39982	0.38000	-0.01982
	0.36996	0.35085	-0.01911

Denoising+SR

- **RED-Net for Mixed Degradations**

Original

DN0-60+SR1-3

DN0-60+SR2-4

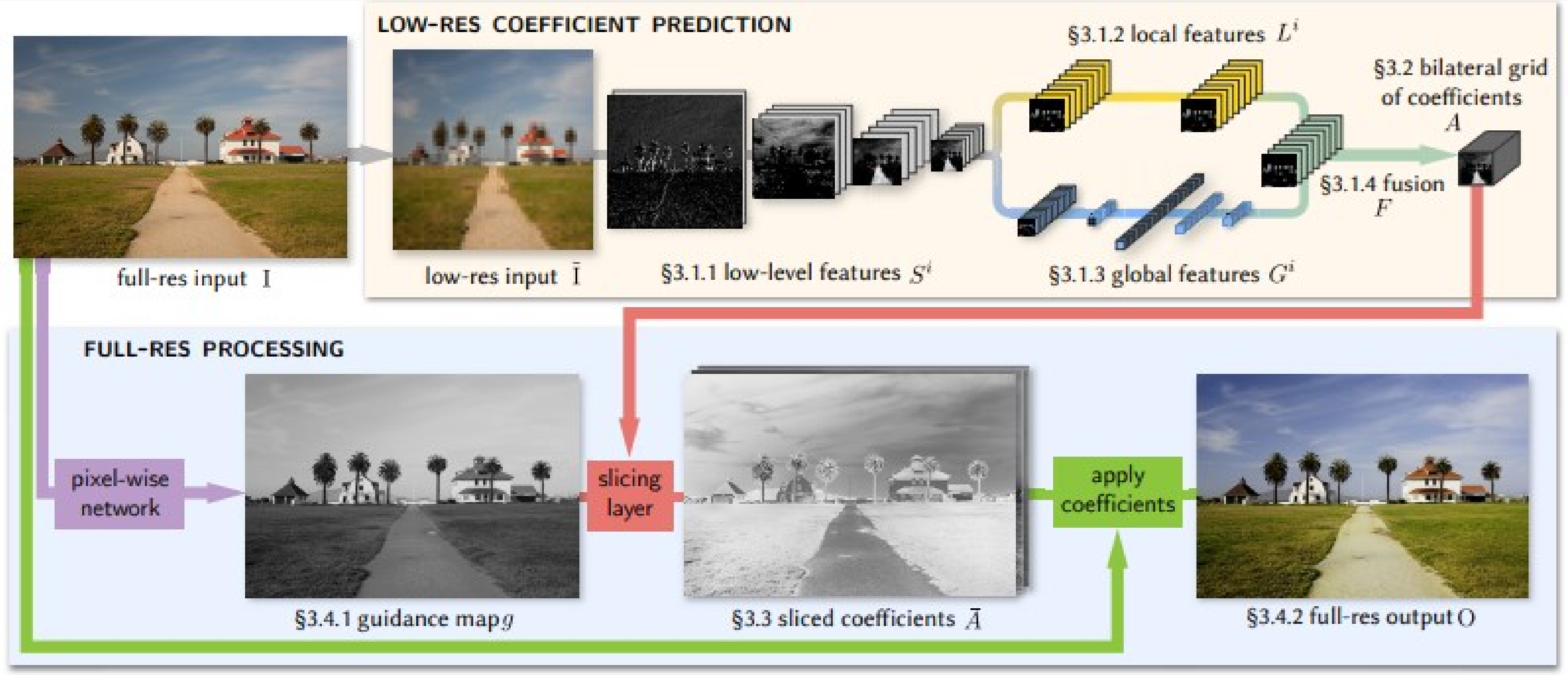
Top-3 predictions

Aircraft	0.28	Aircraft	0.38	Aircraft	0.15
Airship	0.21	Airship	0.17	Fish	0.14
Boat	0.08	Boat	0.06	Boat	0.03

Mao XJ, Shen C, Yang YB. Image restoration using convolutional auto-encoders with symmetric skip connections. *NIPS*, 2016.

HDRNet

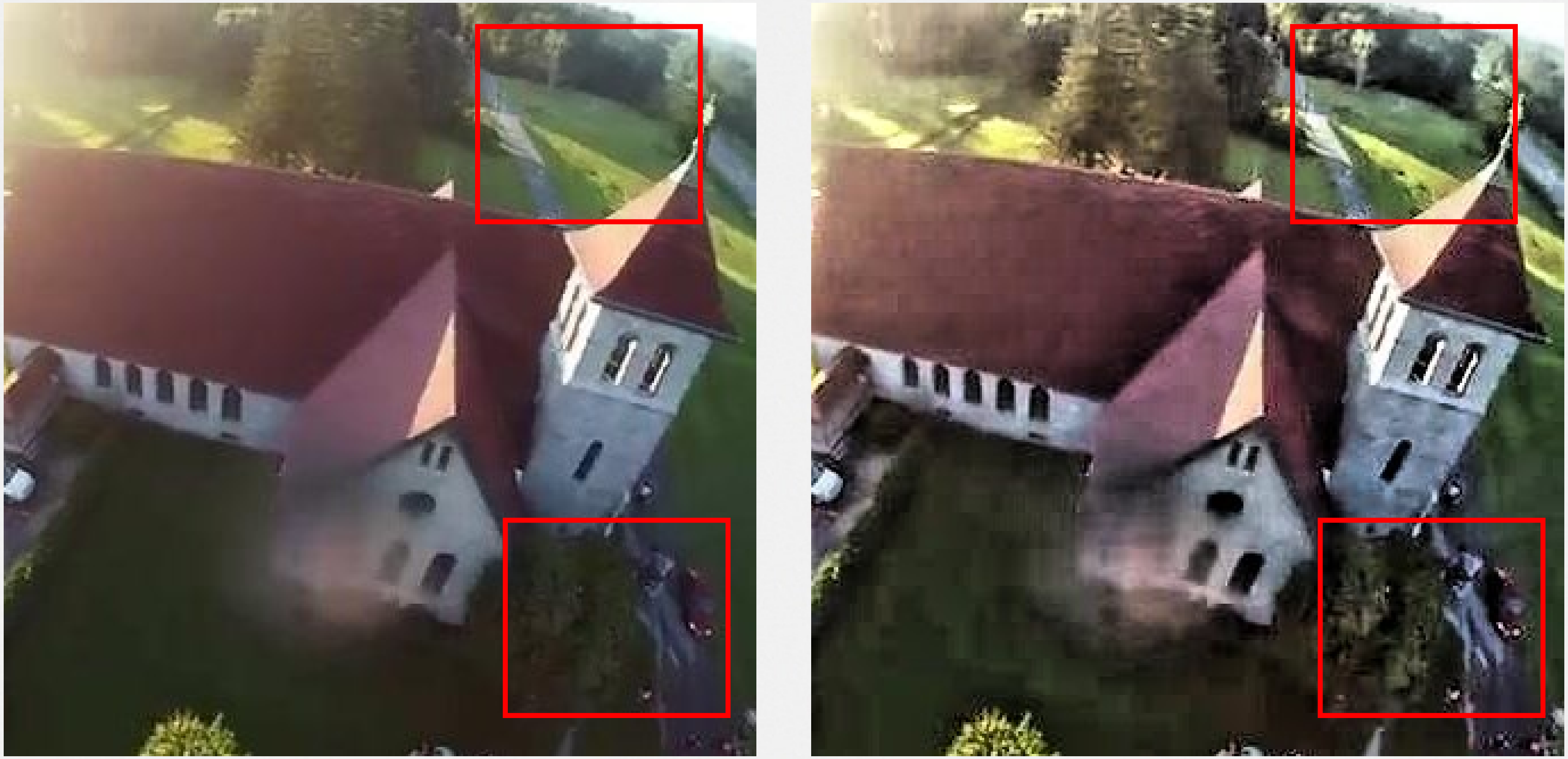
Local Laplacian Model



Gharbi M, Chen J, Barron JT, Hasinoff SW, Durand F. Deep bilateral learning for real-time image enhancement. ACM Transactions on Graphics (TOG). 2017

HDRNet

- **Visual Results**



HDRNet

- **Results on Validation Set**

M1	Baseline	HDRNet	Gain
UAV	0.27477	0.26560	-0.00920
	0.18755	0.23660	0.04900
	0.22455	0.24150	0.01700
	0.23580	0.24170	0.00590
Glider	0.38176	0.37380	-0.00800
	0.38023	0.40430	0.02410
	0.37673	0.41910	0.04240
	0.41922	0.39420	-0.02500
Ground	0.45025	0.46320	0.01290
	0.49486	0.50290	0.00800
	0.43297	0.41680	-0.01620
	0.43555	0.47300	0.03740

M2	Baseline	HDRNet	Gain
UAV	0.04206	0.04740	0.00530
	0.02536	0.02960	0.00420
	0.03427	0.04310	0.00880
	0.01979	0.02540	0.00560
Glider	0.02406	0.01720	-0.00690
	0.05284	0.04740	-0.00540
	0.02436	0.04020	0.01580
	0.04431	0.02300	-0.02130
Ground	0.41492	0.40830	-0.01070
	0.44452	0.41770	-0.02680
	0.39982	0.33480	-0.06500
	0.36996	0.40050	0.03050

Interesting Observation (show M1 gain only)

M1	HistEQ	DeblurGAN	SRCNN	DeblurGAN + HistEQ	DeblurGAN + SRCNN + HistEQ
UAV	0.0181	0.00804	0.00359	0.03316	0.03365
	0.0439	0.01410	0.00148	0.06730	0.06978
	0.0019	0.02054	0.00408	0.02709	0.02907
	0.0056	0.00767	0.00247	0.02054	0.02598
Glider	-0.0230	-0.01462	0.00107	-0.06396	-0.05954
	-0.0062	-0.02513	-0.00533	-0.04020	-0.03990
	0.0224	-0.00929	-0.00868	-0.02513	-0.02711
	-0.0341	-0.05040	-0.00761	-0.06822	-0.06989
Ground	0.0454	-0.00573	-0.00182	0.01490	0.01222
	0.0003	-0.03345	-0.00035	-0.10092	-0.10705
	0.0043	-0.03006	0.00025	-0.03700	-0.04445
	0.0583	0.01110	0.00147	0.05403	0.04856

- Combination of weak “positive model” results in a strong “positive model”
- “negative model” can also contribute
- Dataset dependent performance

Aggregation

■ Results on Validation Set (M1)

M1	Baseline	Model1	Gain	Model2	Gain
UAV	0.27477	0.30843	0.03365	0.13733	-0.13745
	0.18755	0.25733	0.06978	0.12805	-0.05951
	0.22455	0.25362	0.02907	0.14908	-0.07547
	0.23580	0.26178	0.02598	0.16504	-0.07077
Glider	0.38176	0.32222	-0.05954	0.18349	-0.19826
	0.38023	0.34034	-0.03990	0.22278	-0.15745
	0.37673	0.34963	-0.02711	0.22841	-0.14832
	0.41922	0.34932	-0.06989	0.28095	-0.13827
Ground	0.45025	0.46247	0.01222	0.53997	0.08972
	0.49486	0.38780	-0.10705	0.49830	0.00345
	0.43297	0.38851	-0.04445	0.52704	0.09407
	0.43555	0.48411	0.04856	0.55507	0.11952

Model1: DeblurGAN + SRCNN+ HistEQ
Model2: RED + P + HDRNET

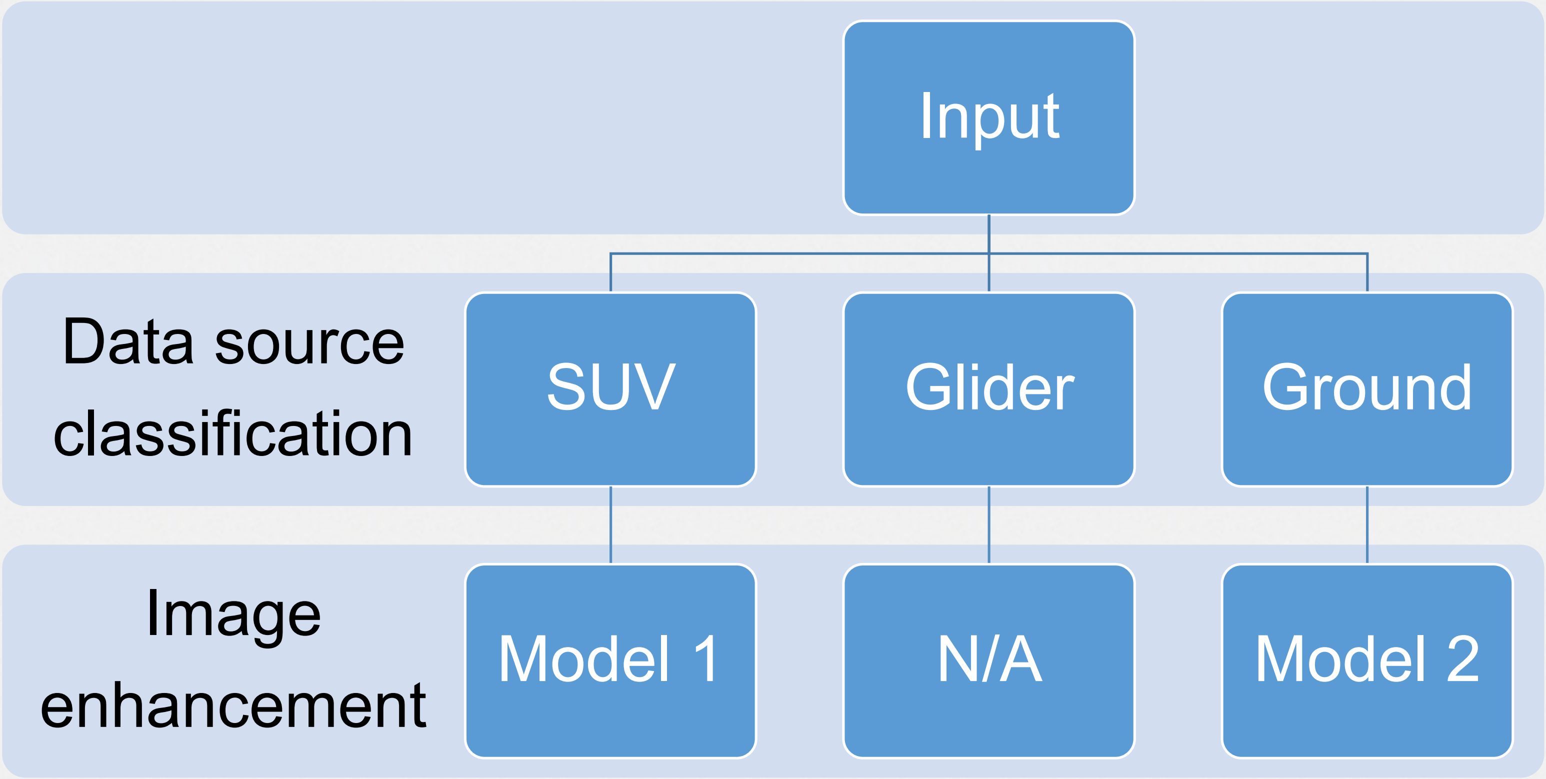
Aggregation

- **Results on Validation Set (M2)**

M1	Baseline	Model1	Gain	Model2	Gain
UAV	0.04206	0.07472	0.03266	0.01175	-0.03031
	0.02536	0.04342	0.01806	0.00965	-0.01571
	0.03427	0.05233	0.01806	0.00903	-0.02524
	0.01979	0.02474	0.00495	0.02004	0.00025
Glider	0.02406	0.01355	-0.01051	0.00213	-0.02193
	0.05284	0.02162	-0.03122	0.00883	-0.04401
	0.02436	0.03533	0.01096	0.02680	0.00244
	0.04431	0.02604	-0.01827	0.01188	-0.03243
Ground	0.41492	0.44538	0.03046	0.47732	0.06240
	0.44452	0.31471	-0.12981	0.43312	-0.01140
	0.39982	0.31375	-0.08607	0.48898	0.08916
	0.36996	0.41675	0.04678	0.47159	0.10163

Model1: DeblurGAN + SRCNN+ HistEQ
Model2: RED + P + HDRNET

Enhancement Pipeline



Model1: DeblurGAN + SRCNN+ HistEQ
Model2: RED + P + HDRNET

Classification Results on Validation Set

M1	Baseline	Model2	Gain
UAV	0.27477	0.29098	0.01621
	0.18755	0.24162	0.05406
	0.22455	0.24149	0.01695
	0.23580	0.25114	0.01534
Glider	0.38176	0.32283	-0.05893
	0.38023	0.34719	-0.03304
	0.37673	0.36090	-0.01584
	0.41922	0.36105	-0.05817
Ground	0.45025	0.53804	0.08779
	0.49486	0.49850	0.00365
	0.43297	0.53343	0.10046
	0.43555	0.55654	0.12099

M2	Baseline	Model3	Gain
UAV	0.04206	0.06854	0.02648
	0.02536	0.04355	0.01819
	0.03427	0.04924	0.01497
	0.01979	0.02388	0.00408
Glider	0.02406	0.01903	-0.00503
	0.05284	0.04203	-0.01081
	0.02436	0.01568	-0.00868
	0.04431	0.01797	-0.02634
Ground	0.41492	0.47539	0.06047
	0.44452	0.43332	-0.01120
	0.39982	0.49536	0.09554
	0.36996	0.47306	0.10310

Model3: combination of model 1 and model 2 with data source classification

Summary & Lessons

- Human visibility and machine perception goals can be (quite) unaligned
- Different platforms (sensors, flying conditions, etc.) matter a lot for what algorithm to choose
- The order of restoration algorithms in a streamlined pipeline is important

Thank you for listening!

Q & A

TAMU-PKU Union (TPU)

Code: https://github.com/yyvettey/TAMU-PKU-UG2

Key Questions and Framework

- **Core idea**
  - *Three cobblers combined makes a genius mind.*
- **Combinations of existing techniques**
  - Length of a pipeline
  - Pipeline order
  - Parallel or sequential
- **Uniformly or differently**
  - Classification
  - Aggregation

Some Issues

- **Why do not we use a completely end-to-end deep-learning framework?**
  - In many cases, we do not have coupled paired data for restoration
  - It is possible to utilize high-level labels for auxiliary supervision of restoration

Some Issues

- **Is the framework a general solution?**
  - Deployed devices usually include some simple pre-processing or post-processing.
  - Our work provides a practical way to further enhance the performance of these devices.
- **How do you use the training data?**
  - We only use the training data in the aggregation step.

To be Explored

- Bridge the gap from high-level labels to low-level priors.