SPIE 2020
International Society for Optics and Photonics

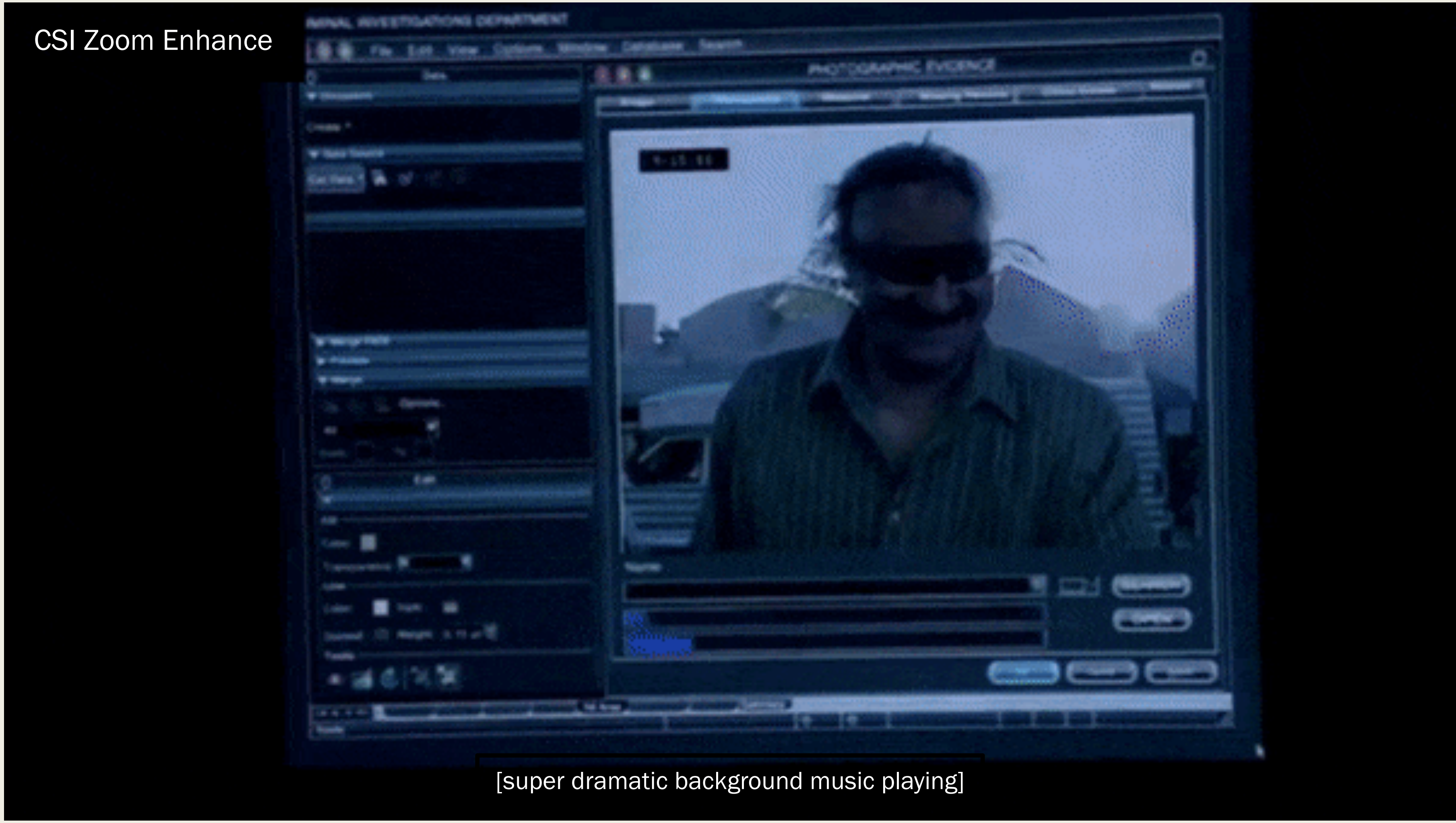
TRAINING SET EFFECT ON MARITIME SUPER RESOLUTION FOR AUTOMATED TARGET RECOGNITION

Matthew Ciolino, David Noever, Josh Kalin, Dominic Hambrick
PeopleTec Inc

Why Super Resolution

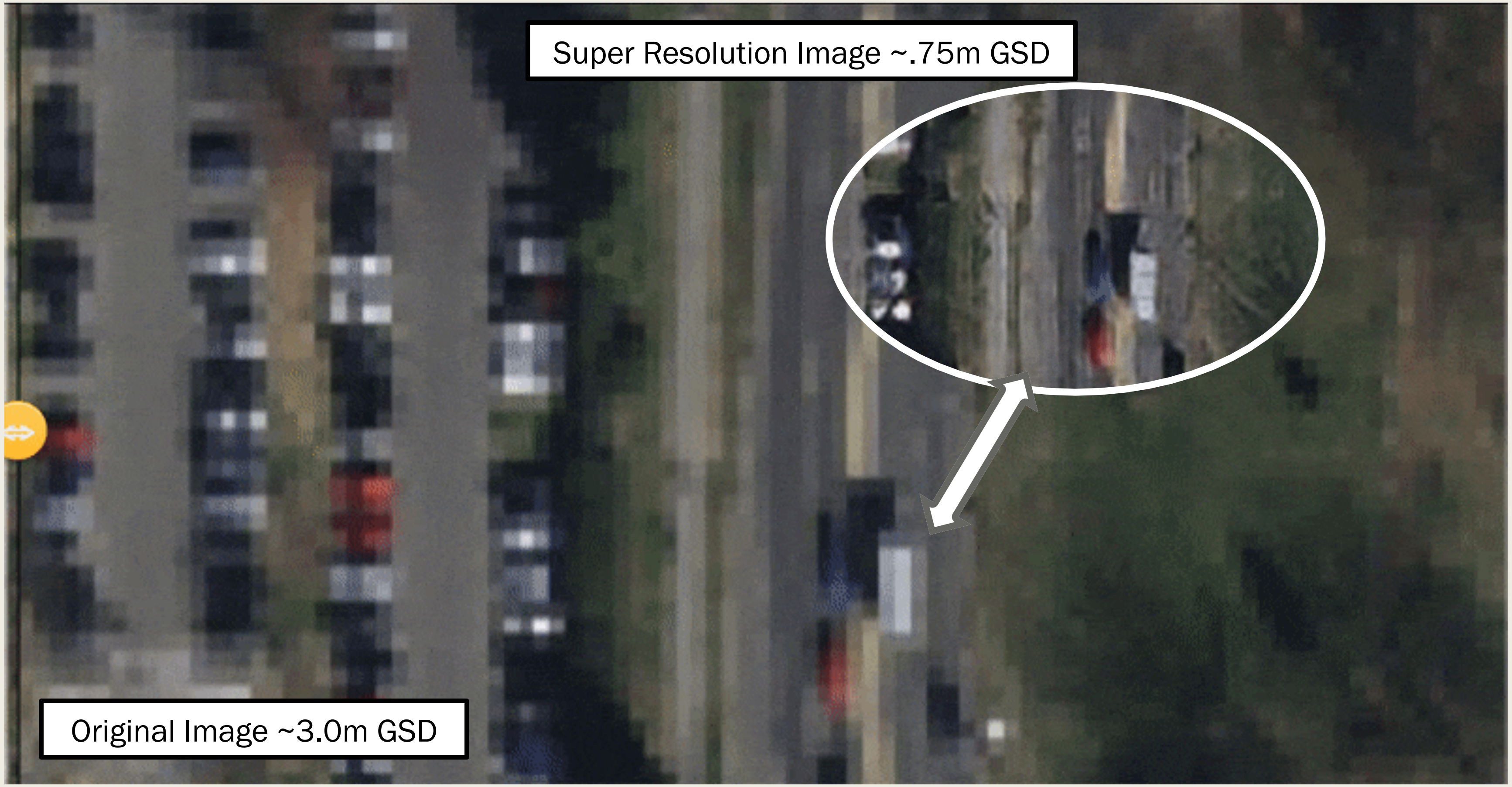
Enhance !!

- CSI Enhance scene showing the general perception of Super Resolution



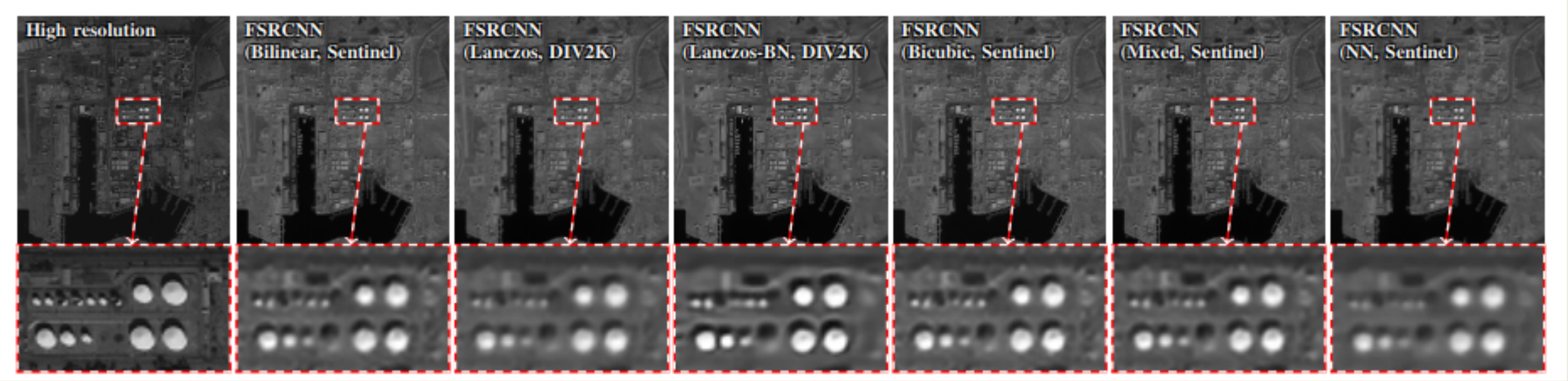
What is Super Resolution

- Convert low resolution imagery into a high-resolution image
- Find pattern between a down-sampled image and that original image
- Better image similarity than scaling algorithms



Related Works

- Changing down sampling method (bicubic, bilinear) to see how it affects the super resolution model



Michal Kawulok, Szymon Piechaczek, Krzysztof Hrynczenko, Pawel Benecki, Daniel Kostrzewa, and Jakub Nalepa. On training deep networks for satellite image super-resolution. arXiv preprint arXiv:1906.06697, 2019.

- Training super resolution model on a dataset and running inference on different datasets

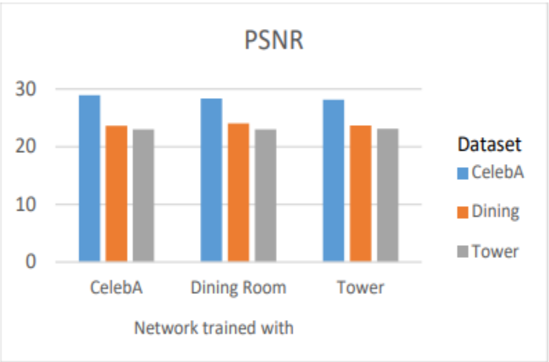


Chart 1: Pixel-wise Euclidean Distance Metrics – PSNR. The higher the value, the better.

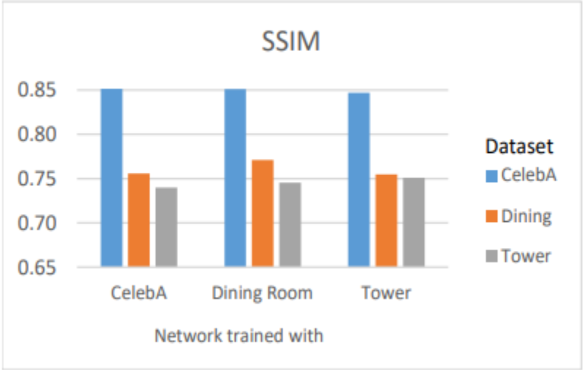


Chart 2: Pixel-wise Euclidean Distance Metrics – SSIM. The higher the value, the better.

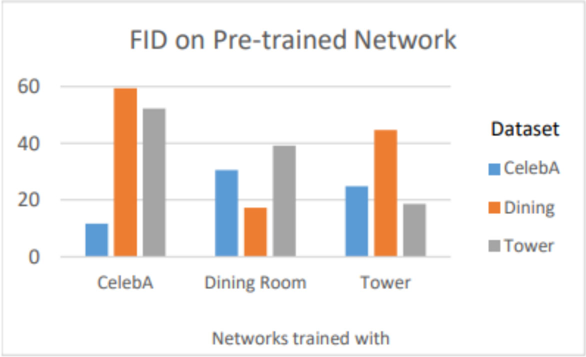


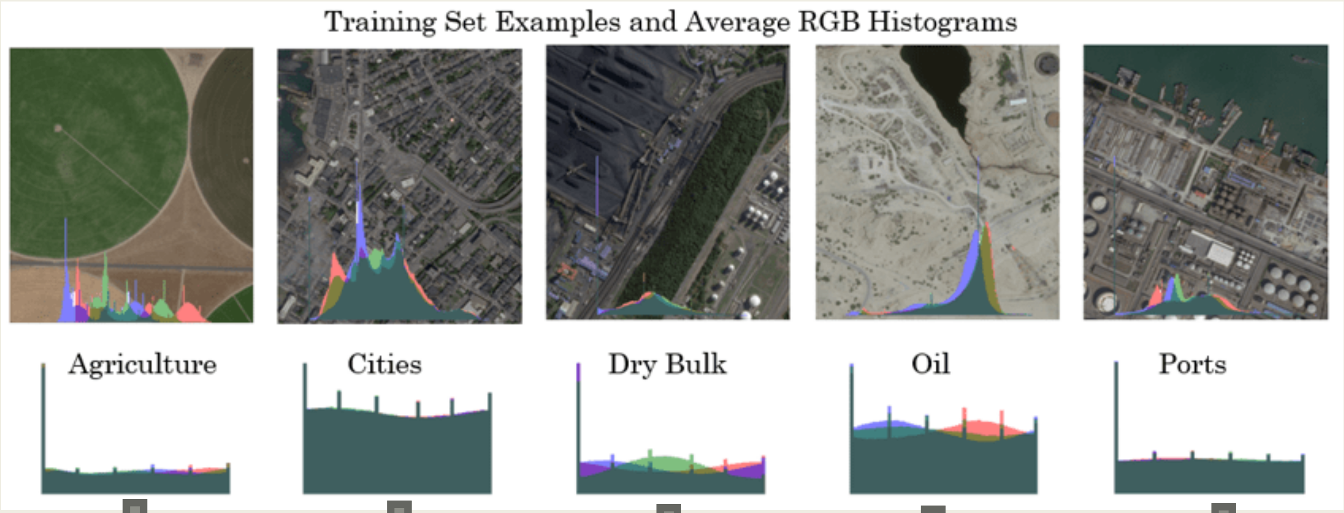
Chart 3: Each pre-trained network that has been trained with CelebA, Dining Room, and Tower is tested against three datasets. The lower the value is, the better.

Nao Takano and Gita Alaghband. Srgan: Training dataset matters. arXiv preprint arXiv:1903.09922, 2019.

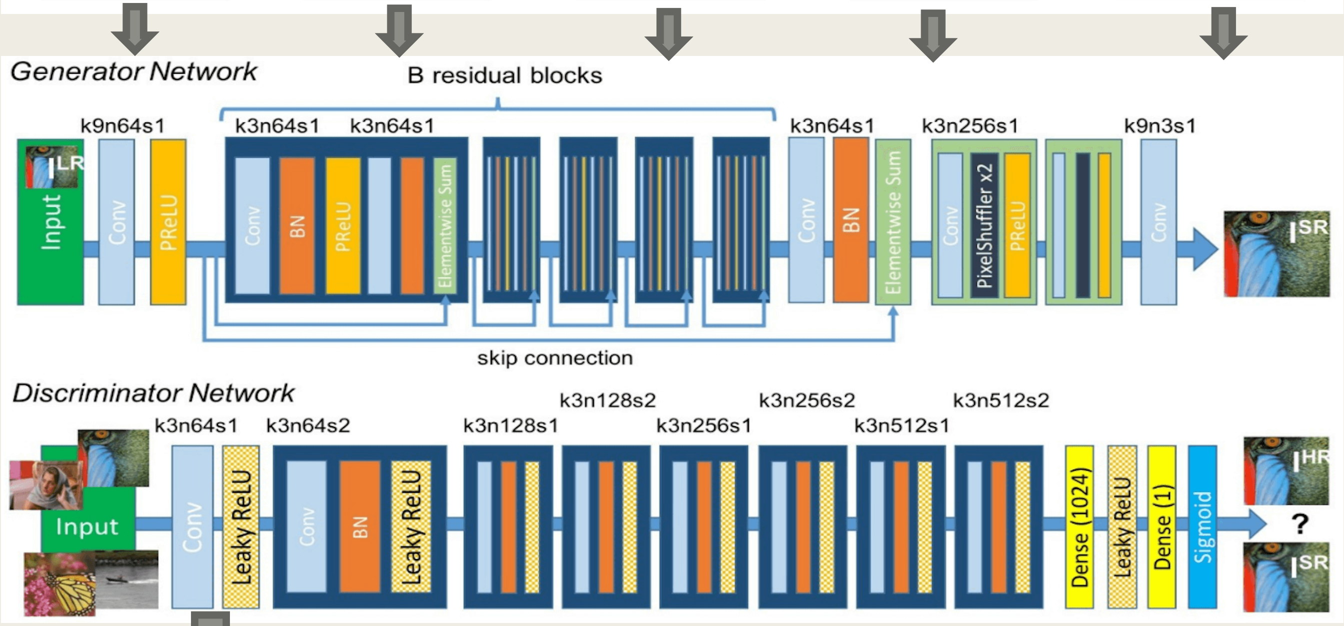
Experiment

- Planet Skysat Imagery Samples as training set
- SRGAN innovations in Super Resolution architecture
- Kaggle's Shipsnet competition and dataset
- Image Classification architecture
- Ultralytics Xview pretrained YoloV3 model

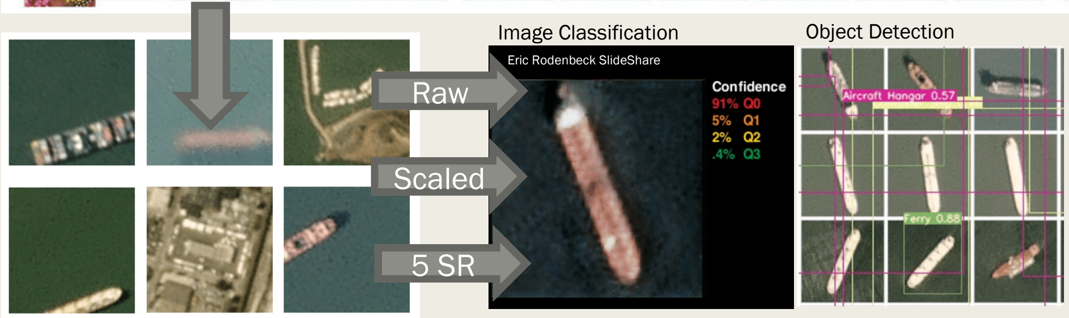
Skysat Training Set
Skysat sample imagery, 2019.



SRGAN
Ledig, Christian, et al. "Photo-realistic single image super-resolution using a generative adversarial network." *Proceedings of the IEEE conference on computer vision and pattern recognition*. 2017.



Shipsnet
Rhammell. "Ships in Satellite Imagery." Kaggle, 29 July 2018,

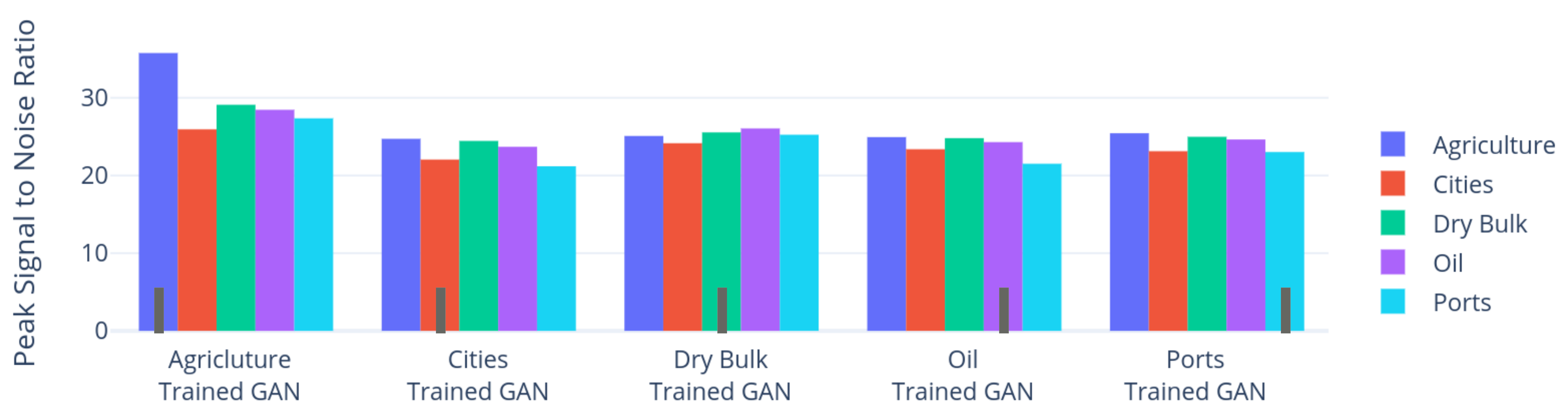


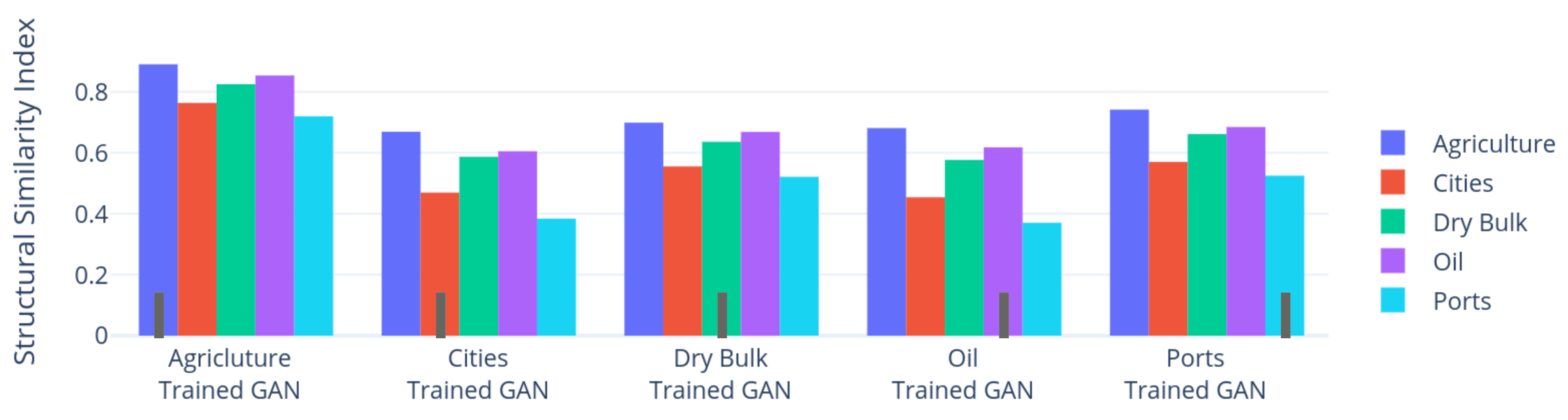
YoloV3
Ultralytics. ultralytics/xview-yolov3, Sep 2019.

Results

Super Resolution

- PSNR and SSIM metrics for each of the SR models on each of the training datasets
- SR models seem to have consistent relative strength across all training datasets
- GANs trained on their own datasets do not appear to perform better than others

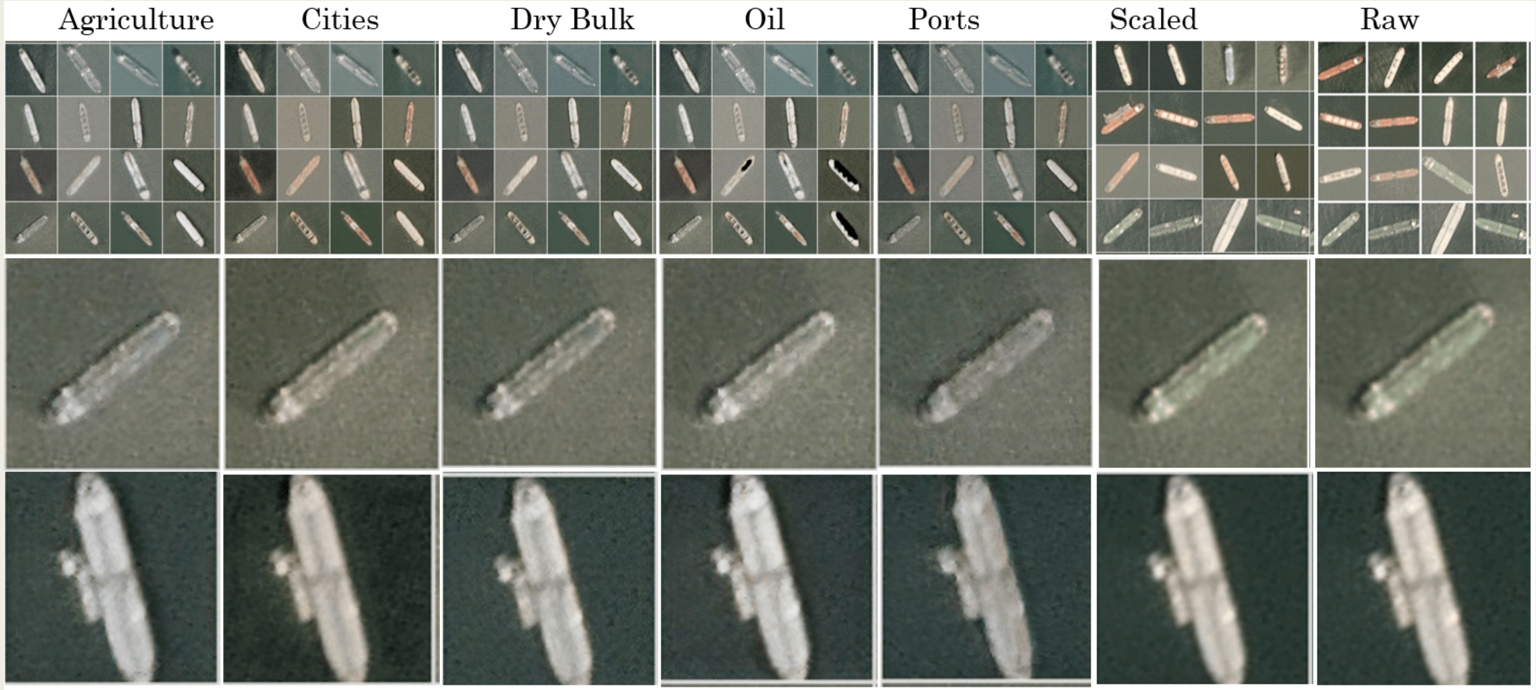


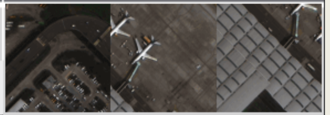
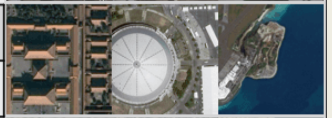
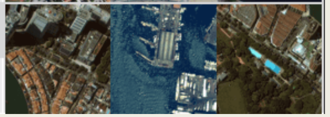


Results

Super Resolution

- Performance of a couple SR models during training
- PSNR/SSIM metrics for a few image samples of different satellites. Images are first down sampled than super resolved.
- SR on the Pleiades image samples seem to perform the best
- Agriculture and Dry Bulk SR model seem to achieve the highest metrics

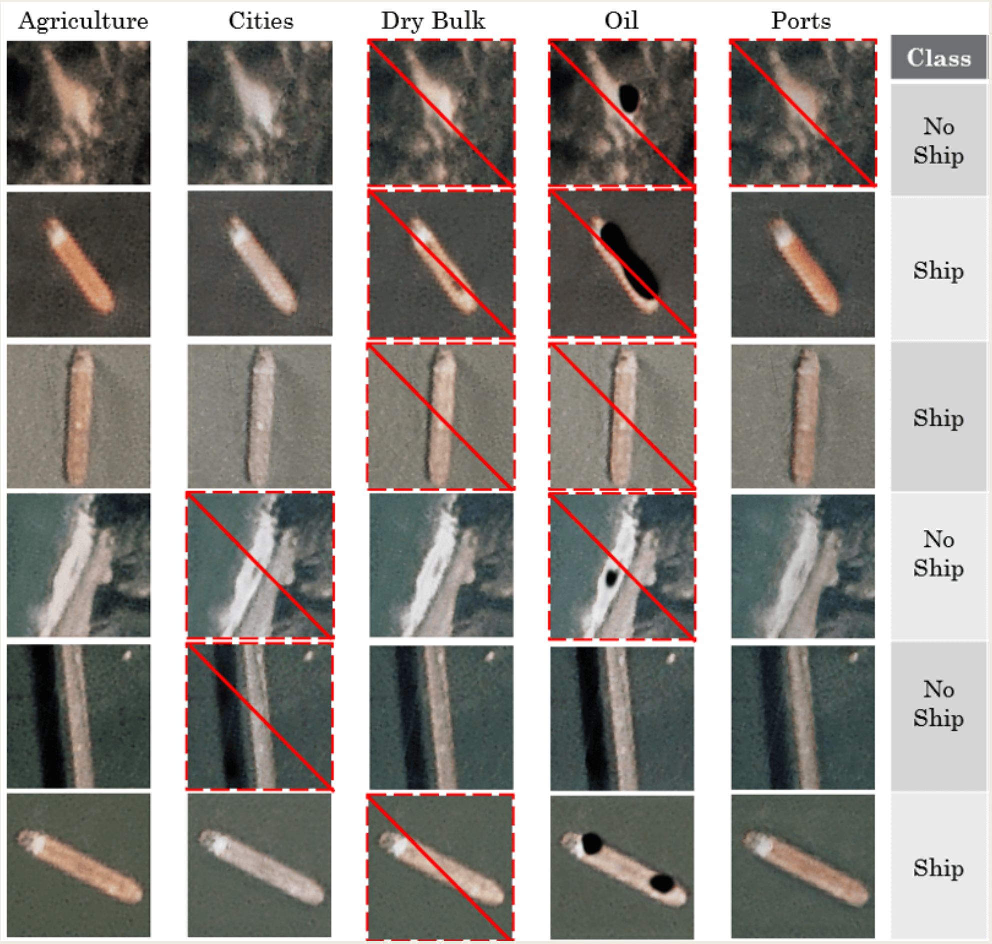


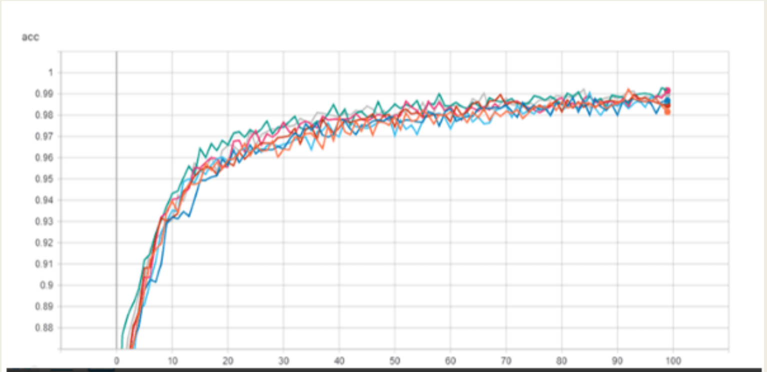
Dataset	Scoring	Agriculture	Cities	Dry bulk	Oil	Ports	Average	Example Images
Xview (.3m)	PSNR	17.011623	16.981220	17.307997	16.801039	17.162341	17.05284	
	SSIM	0.491469	0.477119	0.451492	0.483934	**0.477334**	0.47627	
Pleiades (.5m)	PSNR	**22.770306**	19.560364	**21.841543**	21.544979	18.182035	**20.77985**	
	SSIM	**0.625712**	0.413790	**0.541247**	**0.562504**	0.293370	**0.487325**	
Quickbird (.65m)	PSNR	22.087340	18.617457	18.647200	17.486228	17.244123	18.81647	
	SSIM	0.583851	0.371042	0.451841	0.463415	0.270718	0.428173	
Triplesat (.8m)	PSNR	20.925616	18.843711	20.599660	20.514242	17.674832	19.71161	
	SSIM	0.471851	0.423649	0.504717	0.509565	0.342879	0.450532	
Ikonos (1m)	PSNR	20.413195	**19.709312**	21.795880	**21.917841**	**19.604518**	20.68815	
	SSIM	0.483809	**0.449529**	0.518172	0.531595	0.394726	0.475566	

Results

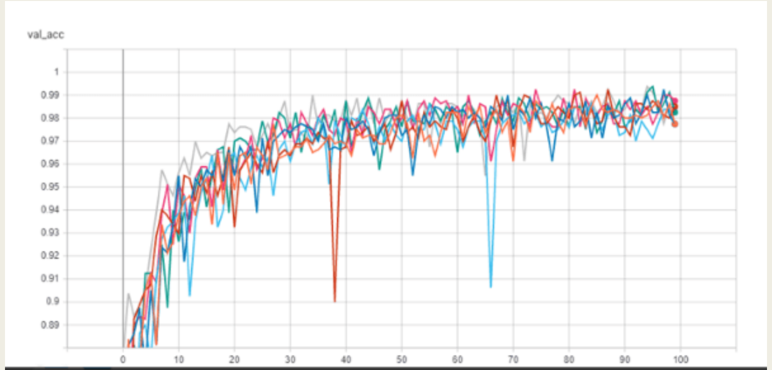
Image Classification

- Image classification examples with common misclassifications
- Training and validation accuracy for 80/20 split
- Most SR trained image models are marginally worse/better than raw scores
- This dataset is closer to being solved than other and shows a diminishing amount of returns





Name	Smoothed	Value	Step	Time	Relative
Logs\raw\sequential_6	0.9909	0.9916	99	Thu Sep 5, 02:11:43	12m 10s
Logs\ports\sequential_5	0.9893	0.9916	99	Thu Sep 5, 01:59:23	2h 23m 45s
Logs\scaled\sequential_7	0.9892	0.9897	99	Thu Sep 5, 04:29:04	2h 15m 47s
Logs\oil\sequential_4	0.9858	0.9856	99	Wed Sep 4, 23:34:00	2h 23m 28s
Logs\cities\sequential_2	0.9857	0.9869	99	Wed Sep 4, 18:44:21	2h 21m 21s
Logs\dry_bulk\sequential_3	0.9855	0.9847	99	Wed Sep 4, 21:08:54	2h 22m 55s
Logs\agriculture\sequential_1	0.9843	0.9816	99	Wed Sep 4, 16:21:24	2h 22m 14s



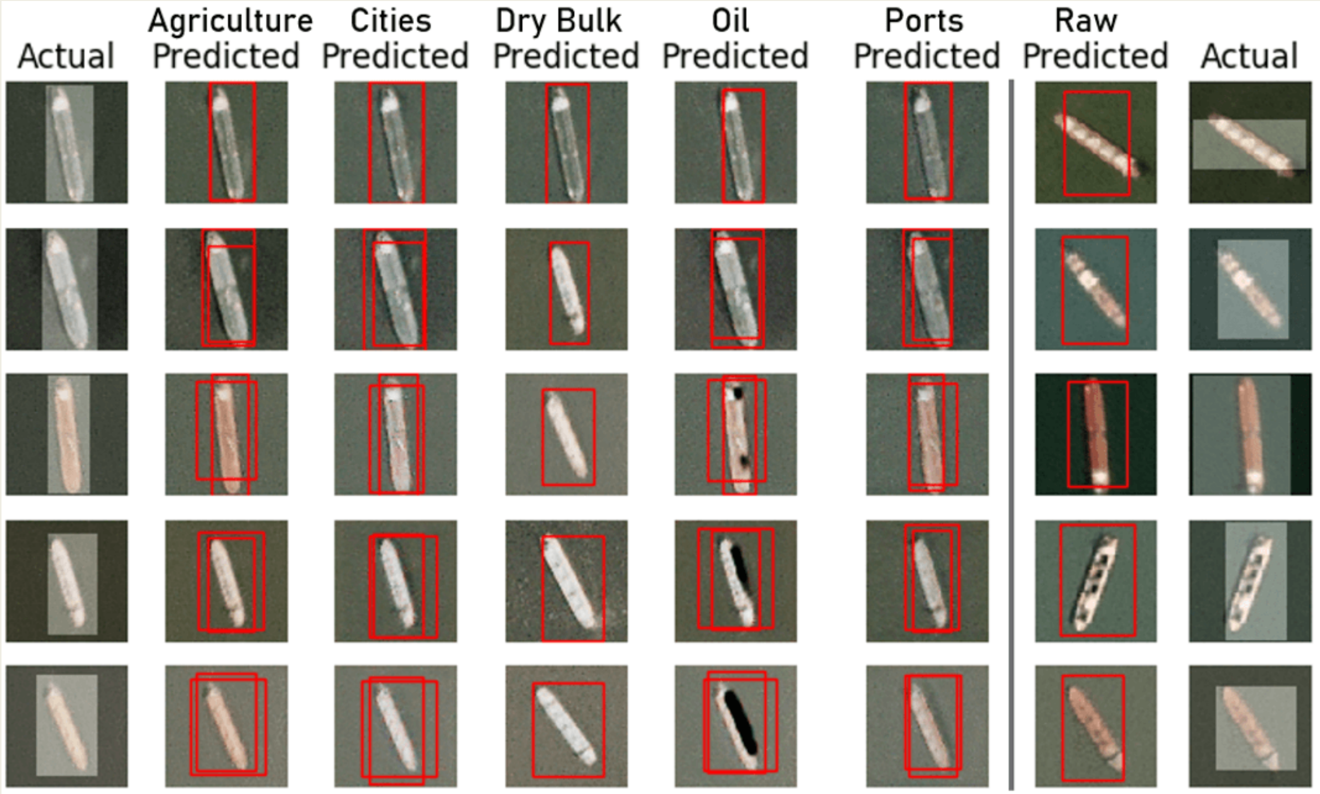
Name	Smoothed	Value	Step	Time	Relative
Logs\ports\sequential_5	0.9872	0.9875	99	Thu Sep 5, 01:59:23	2h 23m 45s
Logs\raw\sequential_6	0.9859	0.9825	99	Thu Sep 5, 02:11:43	12m 10s
Logs\scaled\sequential_7	0.9848	0.985	99	Thu Sep 5, 04:29:04	2h 15m 47s
Logs\dry_bulk\sequential_3	0.9832	0.985	99	Wed Sep 4, 21:08:54	2h 22m 55s
Logs\cities\sequential_2	0.9819	0.9775	99	Wed Sep 4, 18:44:21	2h 21m 21s
Logs\agriculture\sequential_1	0.9815	0.9775	99	Wed Sep 4, 16:21:24	2h 22m 14s
Logs\oil\sequential_4	0.9795	0.9775	99	Wed Sep 4, 23:34:00	2h 23m 28s

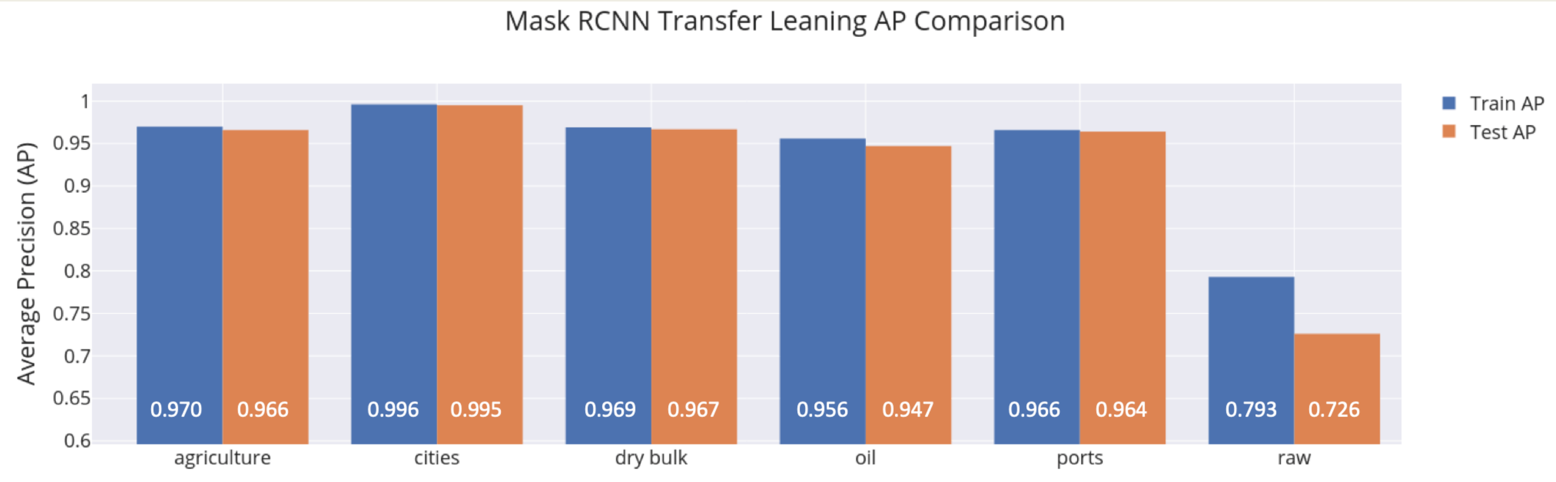
Results

Object Detection

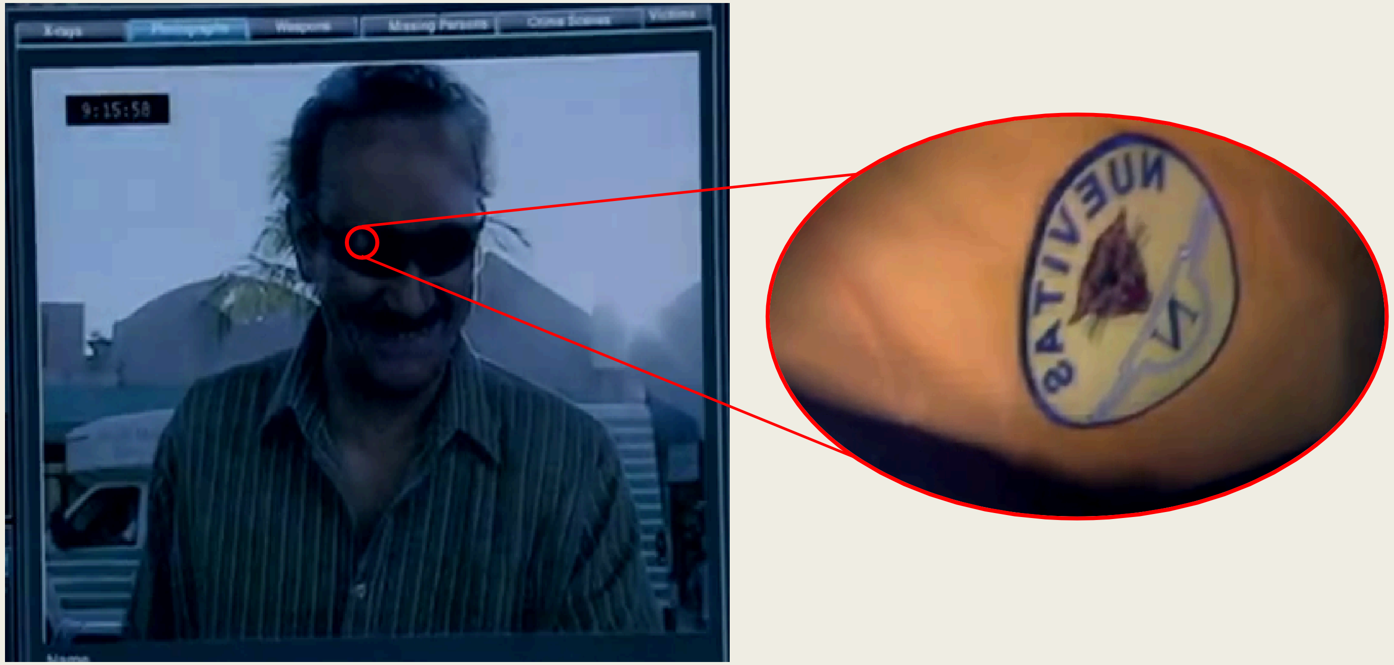
- Mask-RCNN Transfer learning show huge improvements in mAP for all Super-Resolved Test Sets (avg +18.4%)

- Cities trained Mask RCNN preforms best at 99.5% hinting at image diversity as a key factor of performance

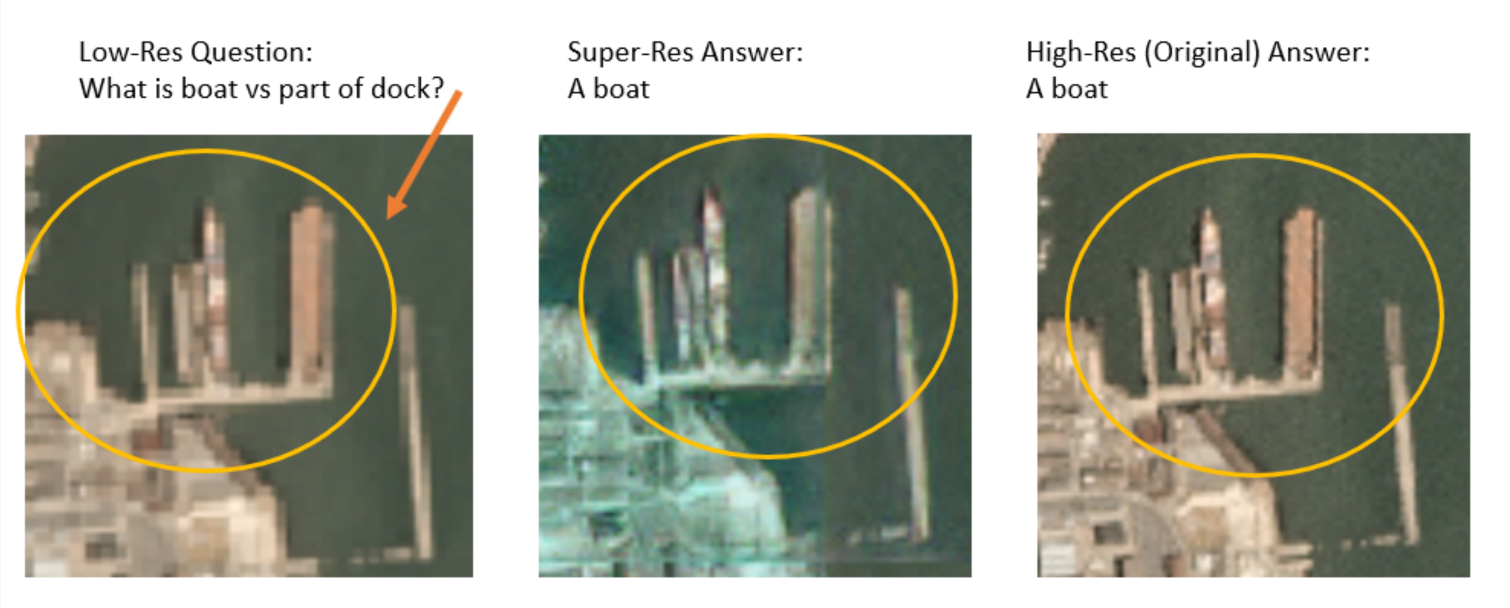




Fantasy vs Reality



- Unexperienced expectation of Super Resolution

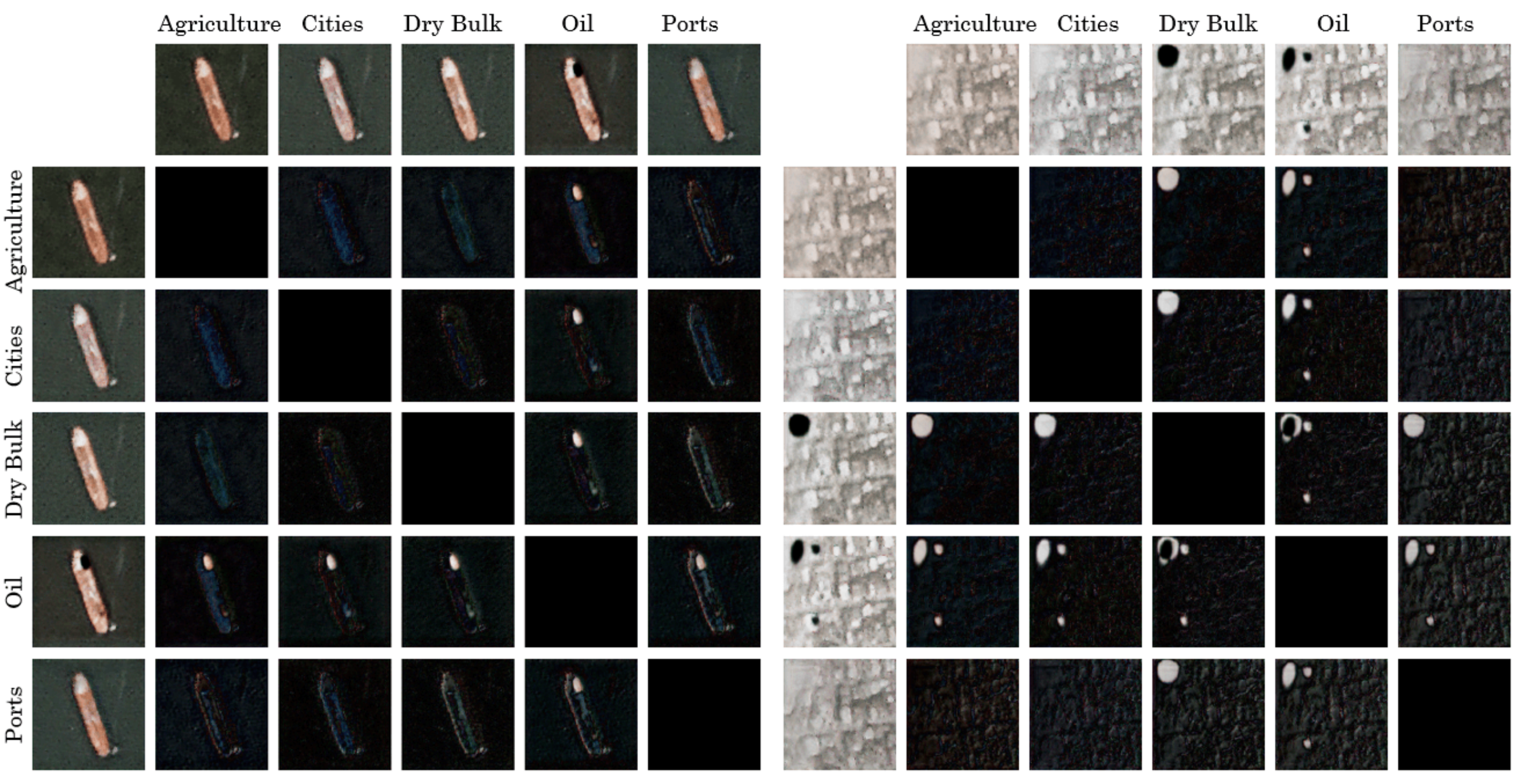


- Actual use case for Super Resolution

Key Takeaways

- Super resolution, when applied to satellite imagery, can increase the spatial resolution of your images
  - *Less money spent on developing bigger satellite camera systems (i.e. small sats)*
- Super resolution while beneficial, is not the always the answer for some datasets
  - *Curated training sets "Fine Tuned" for your use case is beneficial to performance*

Extras Absolute Difference in SR Models on an Image



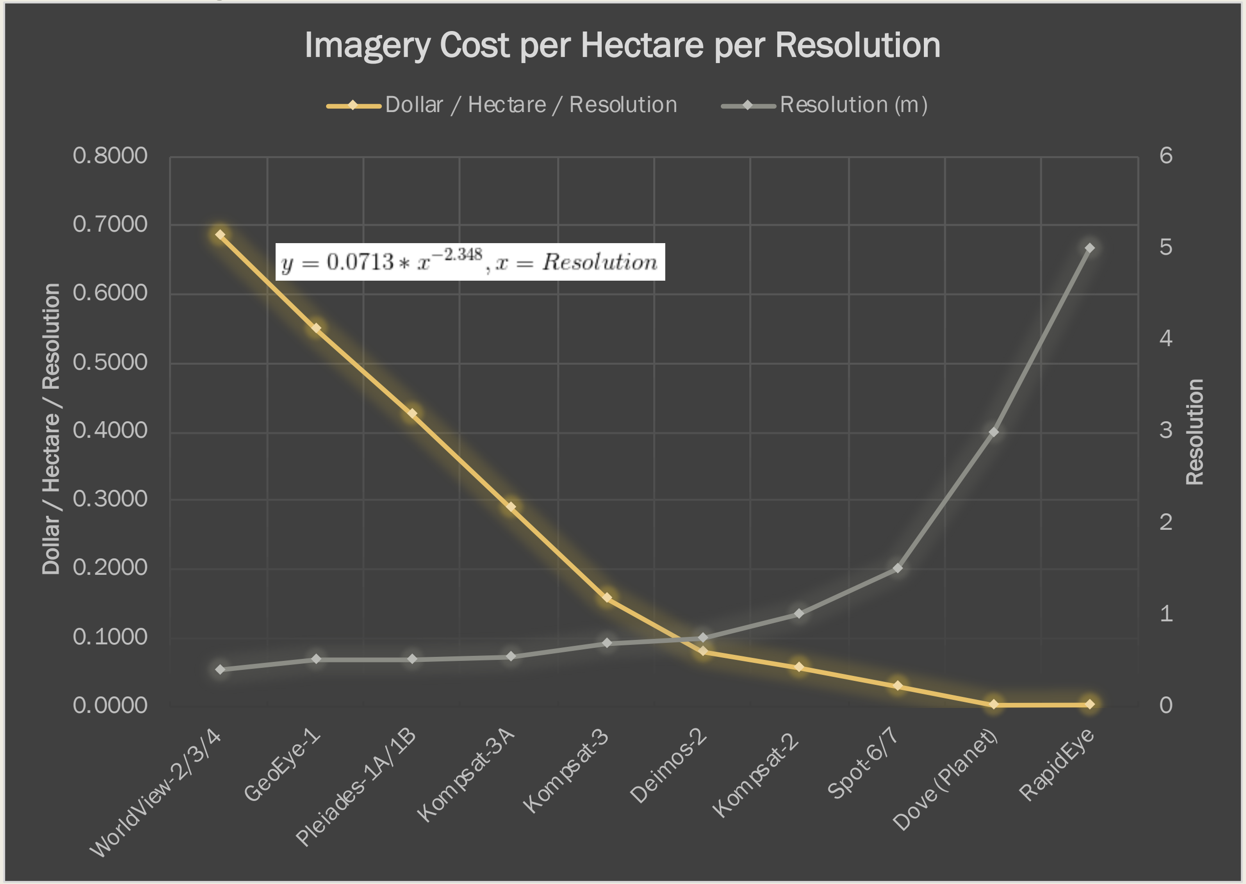
Satellite Image Cost Analysis

Table 1 Summary table of sensors features

	Spatial resolution *m*	Spectral resolution VIS/NIR *nm*	Radiometric resolution *bits pixel^{-1}*	Revisit time *dd*
Deimos-2	0.75	466-697/770-892	10	2
Dove (Planet)	3	420-700/770-900	16	1
GeoEye-1	0.5	450-690/780-920	11	3
Kompsat-2	1	450-690/760-900	14	6
Kompsat-3	0.7	450-690/760-900	14	3
Kompsat-3A	0.55	450-690/760-900	14	3
Landsat-7/8	15	450-690/770-900	8-12	8
Pleiades-1A/1B	0.5	430-720/750-950	12	1
RapidEye	5	440-685/690-850	12	5.5
Sentinel-2	10	458-680/785-900	16	5
Spot-6/7	1.5	455-695/760-890	12	1
WorldView-2/3/4	0.3-0.5	450-690/770-1040	11	1

	Minimum order area *ha*	Price per unit *\$ ha^{-1}*	Minimum area price* *\$*	Computational demand *KB ha^{-1}*
Deimos-2	10 000	0.060	700	50
Dove	10 000	0.012	218	8
GeoEye-1	10 000	0.275	2850	100
Kompsat-2	2 500	0.055	237.5	20
Kompsat-3	2 500	0.110	375	50
Kompsat-3A	2 500	0.160	500	100
Landsat-7/8	3 700 000 (one scene)	0	100	0.5
Pleiades-1A/1B	10 000	0.213	2225	100
Rapideye	10 000	0.012	218	4
Sentinel-2	1 200 000 (one scene)	0	100	0.63
Spot-6/7	10 000	0.045	550	8
WorldView-2/3/4	10 000	0.275	2850	130

*minimum area price is obtained by minimum order area times price per unit plus data processing cost



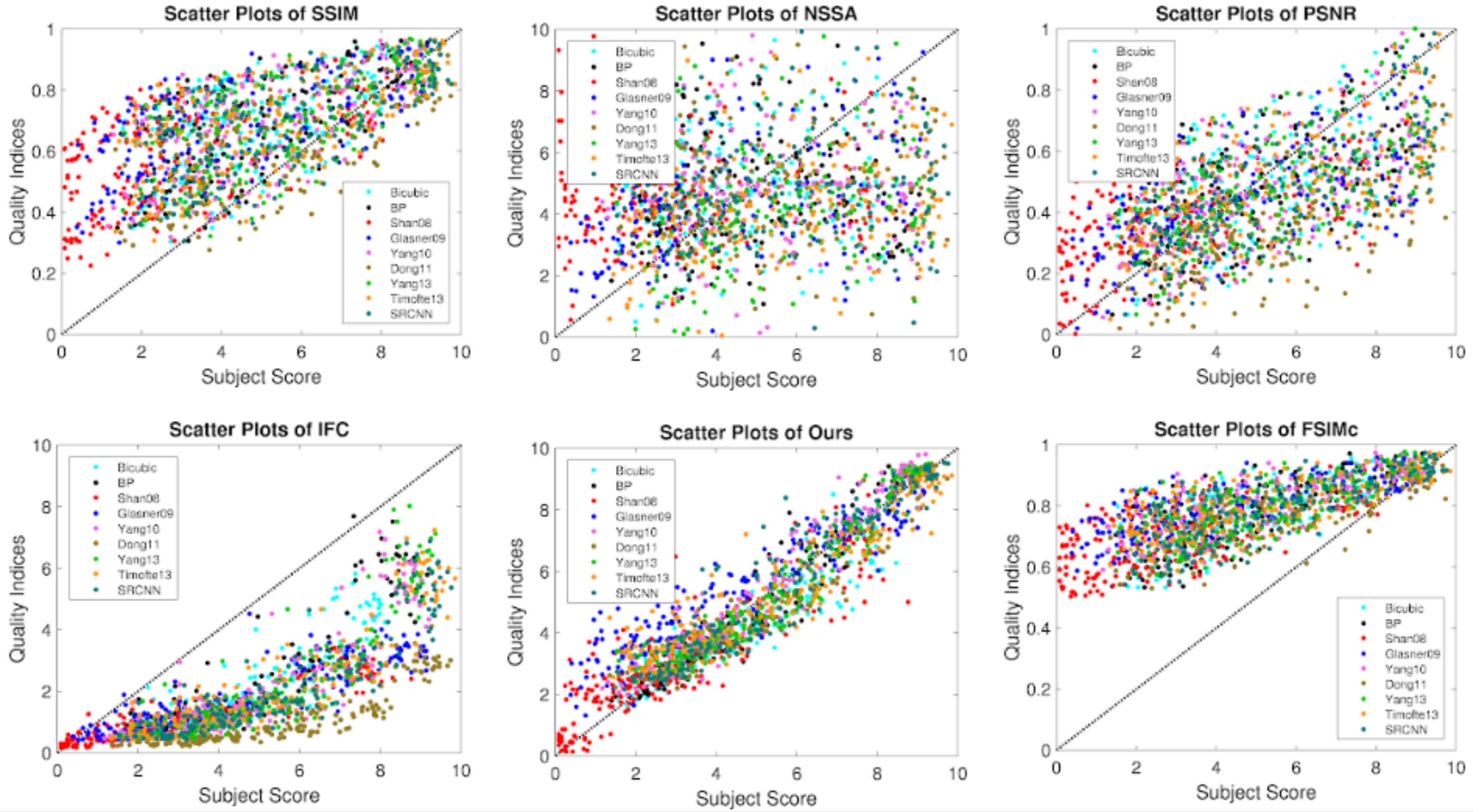
Sozzi, Marco, et al. "Benchmark of satellites image services for agricultural use." *Proceedings of the AgEng Conference, Wageningen, The Netherlands*. 2018.

- While research and develop costs are hard to find, costs eventually trickle down to the customer
- Once we have the data, we can calculate the price difference if the resolution was increased 2x.
  - On average, if resolution was increased 2x a 79.5% increase in price could be had

National Image Interpretability Rating Scales

NIRS Scale	GSD Approximate	Visible
1	>9.00m	Detect a medium-sized port facility and/or distinguish between taxi-ways and runways at a large airfield.
2	4.50 - 9.00m	Detect large buildings at a naval facility (e.g., warehouses, construction hall).
3	2.50 - 4.50m	Identify a large surface ship in port by type (e.g., cruiser, auxiliary ship, noncombatant/merchant).
4	1.20 - 2.50m	Detect an open missile silo door.
5	0.75 - 1.20m	Identify radar as vehicle-mounted or trailer-mounted.
6	0.40 - 0.75m	Identify the spare tire on a medium-sized truck.
7	0.20 - 0.40m	Identify individual rail ties.
8	0.10 - 0.20m	Identify windshield wipers on a vehicle.
9	<0.10m	Identify vehicle registration numbers (VRN) on trucks.

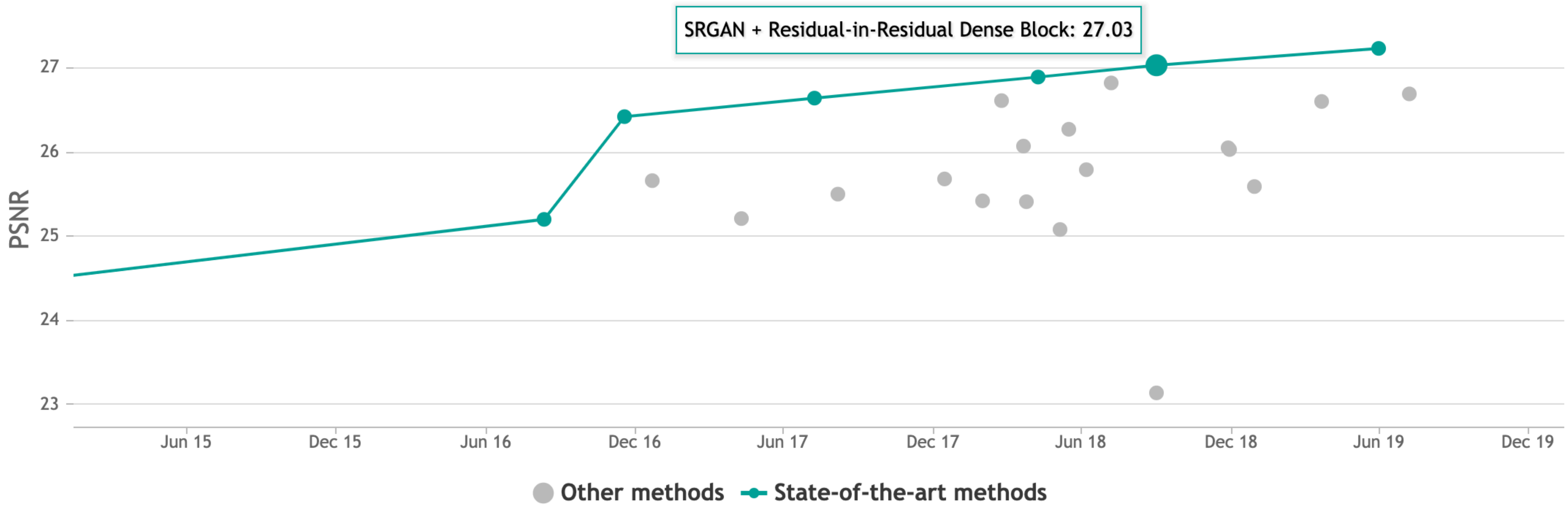
Extras Comparing Image Scores vs. Human Scores



Ma, Chao, et al. "Learning a no-reference quality metric for single-image super-resolution." Computer Vision and Image Understanding 158 (2017): 1-16.

State of Super Resolution

Image Super-Resolution on Urban100 - 4x upscaling



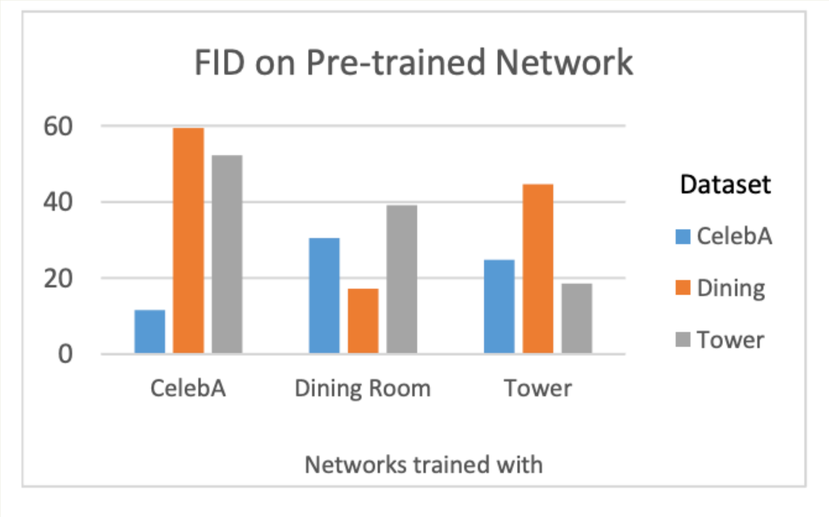
RANK	METHOD	PSNR	SSIM	PAPER TITLE	YEAR	PAPER	CODE
1	SAN	27.23	0.8169	Second-Order Attention Network for Single Image Super-Resolution	2019		
2	SRGAN + Residual-in-Residual Dense Block	27.03	0.8153	ESRGAN: Enhanced Super-Resolution Generative Adversarial Networks	2018		

Urban 100 Dataset

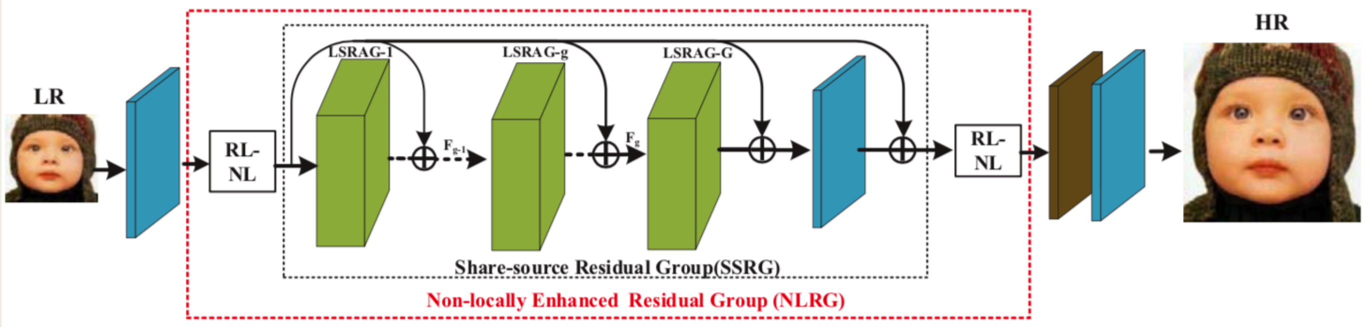
Next Steps

- Future work with testing on various datasets [11] will validate the increase in performance from SR networks.
- Using neural networks that allow varying image size inputs would allow one to use SR on images of arbitrary size unlike the fixed size of SRGAN.
- Future testing, with different neural network architectures [SAN], could yield better results

Nao Takano and Gita Alaghband. Srgan: Training dataset matters. arXiv preprint arXiv:1903.09922, 2019.



Li, Peihua, et al. "Is second-order information helpful for large-scale visual recognition?." *Proceedings of the IEEE International Conference on Computer Vision*. 2017.



Our Paper

- We explore the effect that different training sets have on Single Image Super Resolution (SISR) with the network, Super Resolution Generative Adversarial Network (SRGAN).
- We train 5 SRGANs on different land-use classes (e.g. agriculture, cities, dry bulk, oil, ports) and test them on the same unseen data set, Shipsnet.
- We attempt to find the qualitative and quantitative differences in SISR, binary classification, and object detection performance.
- We find that curated training sets that contain objects in the test ontology perform better in the computer vision tasks.
- However, SR might not be beneficial to certain problems and will see a diminishing amount of returns for data sets that are closer to being solved