Patch-wise Contrastive Style Learning for Instagram Filter Removal

Furkan Kınlı[1] Barış Özcan[2] Furkan Kıraç[3]

Department of Computer Science, Özyeğin University, İstanbul, Türkiye

{[1]furkan.kinli, [3]furkan.kirac}@ozyegin.edu.tr, [2]baris.ozcan.10097@ozu.edu.tr

Problem Definition

- Image-level corruptions and perturbations degrade the performance of DL-based vision models.
- CNNs are not robust to the image-level corruptions and perturbations for the downstream vision tasks.
- Social media filters are one of the most common resources of various corruptions and perturbations.

Figure 1. Failure cases arising from filtering on detection and segmentation tasks in fashion domain. Retrieved from [1].

Approach

- These distractive factors can be alleviated by recovering the original images with their pure style.

  Assumption: Filters substantially inject additional style information to the images.

- Following the same assumption in the prior work, we formulate the problem statement as follows:

$$\sim\mathbf{X} = \mathbf{T}(\mathbf{X})$$

where **X** and ~**X** are RGB images (i.e., original and filtered), and **T** is the transformation function applies the filters. Since finding $\mathbf{T}^{-1}$ is **ill-posed**, we try to find **F**, the best approximation to **T**, as follows:

$$\mathbf{X} = \mathbf{F}(\sim\mathbf{X})$$

Isolated Patch Sampling Module for Style Distilling

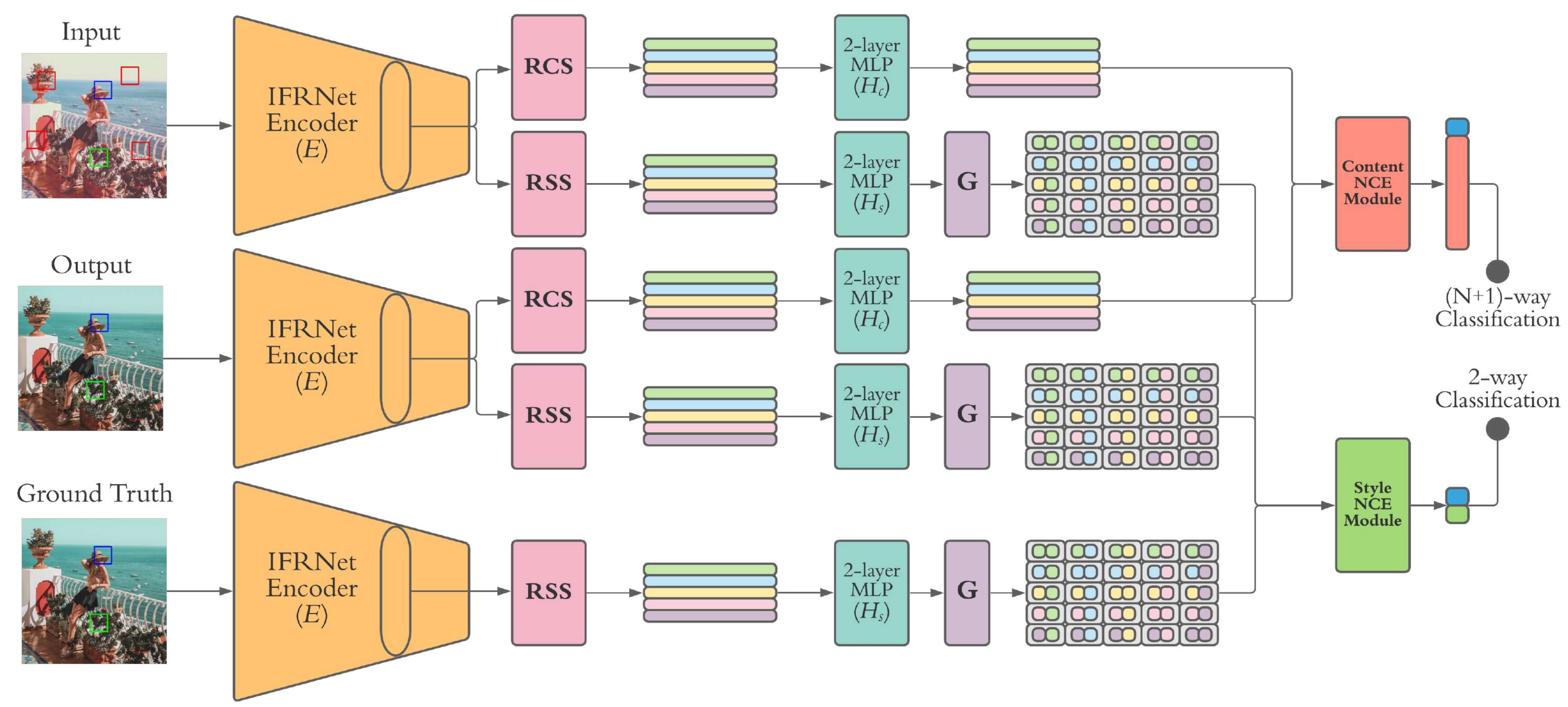


Figure 2. This figure shows the pipeline for only a single level features. The extracted feature maps by IFRNet Encoder (E) are first fed into the random sampling modules for the content (RCS) and style (RSS), separately. After encoding them by corresponding 2-layer MLP modules, the *content* patches for the input and the output are sent to Content NCE module, and content NCE loss is calculated as proposed in [38]. Moreover, the *style* patches are extracted by G for calculating the Gram matrix of the encoded features, and style NCE loss learns to select the patch with *pure* style over the filtered patch.

Qualitative Results

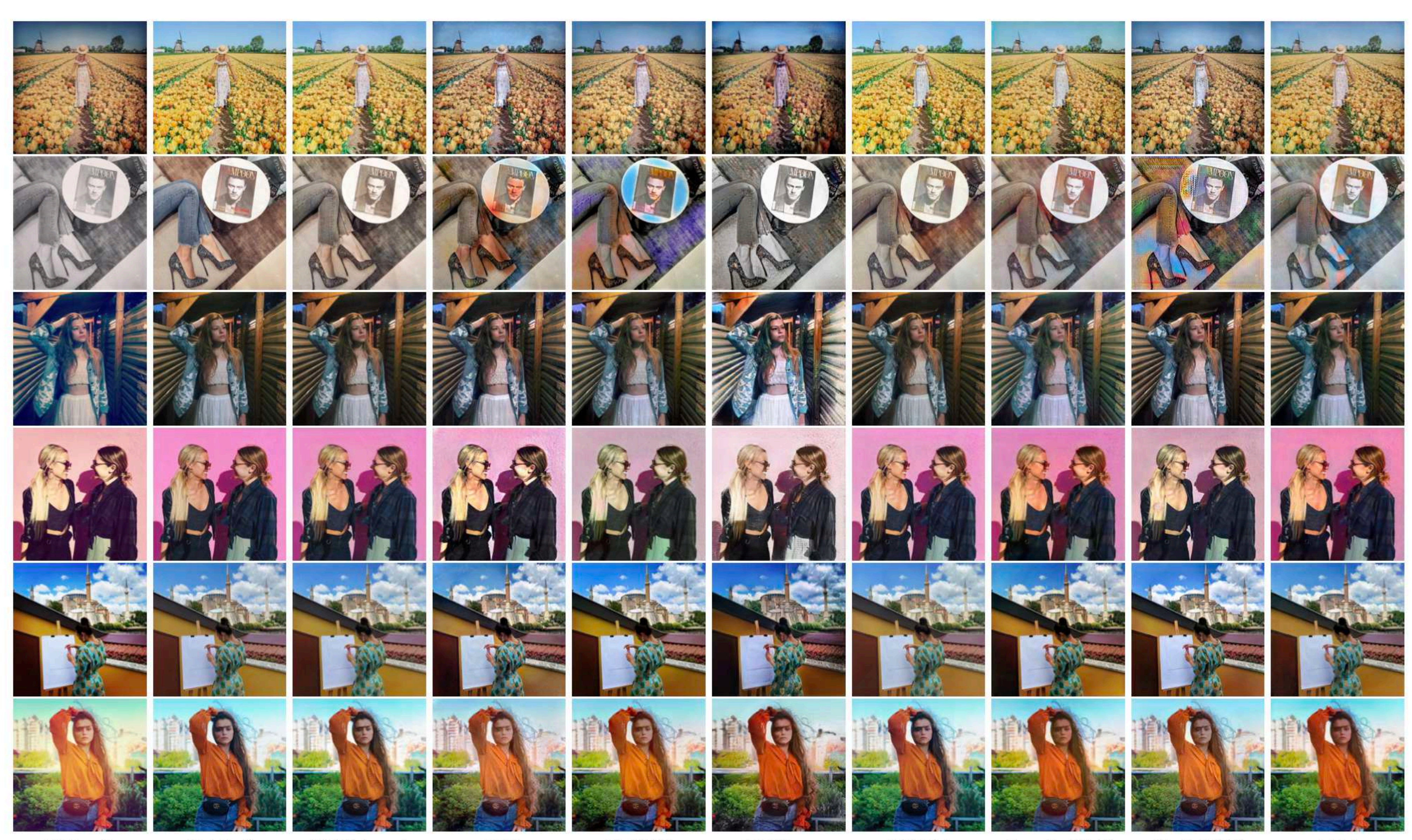


Figure 3. Comparison of the qualitative results of Instagram filter removal on IFFI dataset.

References

[1] Furkan Kinli, Baris Ozcan, and Furkan Kirac. Instagram filter removal on fashionable images. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR) Workshops, pages 736–745, June 2021.

Objective Function

$$\mathcal{L}_C(E,H,\mathbf{X},\tilde{\mathbf{X}}) = \mathbb{E}_{x\sim\mathbf{X},\tilde{x}\sim\tilde{\mathbf{X}}} \sum_{l=1}^{L}\sum_{t=1}^{T^l} \ell(\hat{\mathbf{z}}_c^{l,t}, \tilde{\mathbf{z}}_c^{l,t}, \tilde{\mathbf{z}}_c^{l,T\setminus t})$$

$$\mathcal{L}_S(E,H,\mathbf{X},\tilde{\mathbf{X}}) = \mathbb{E}_{x\sim\mathbf{X},\tilde{x}\sim\tilde{\mathbf{X}}} \sum_{l=1}^{L}\sum_{t=1}^{T^l} \ell(\hat{\mathbf{z}}_s^{l,t}, \mathbf{z}_s^{l,t}, \tilde{\mathbf{z}}_s^{l,t'})$$

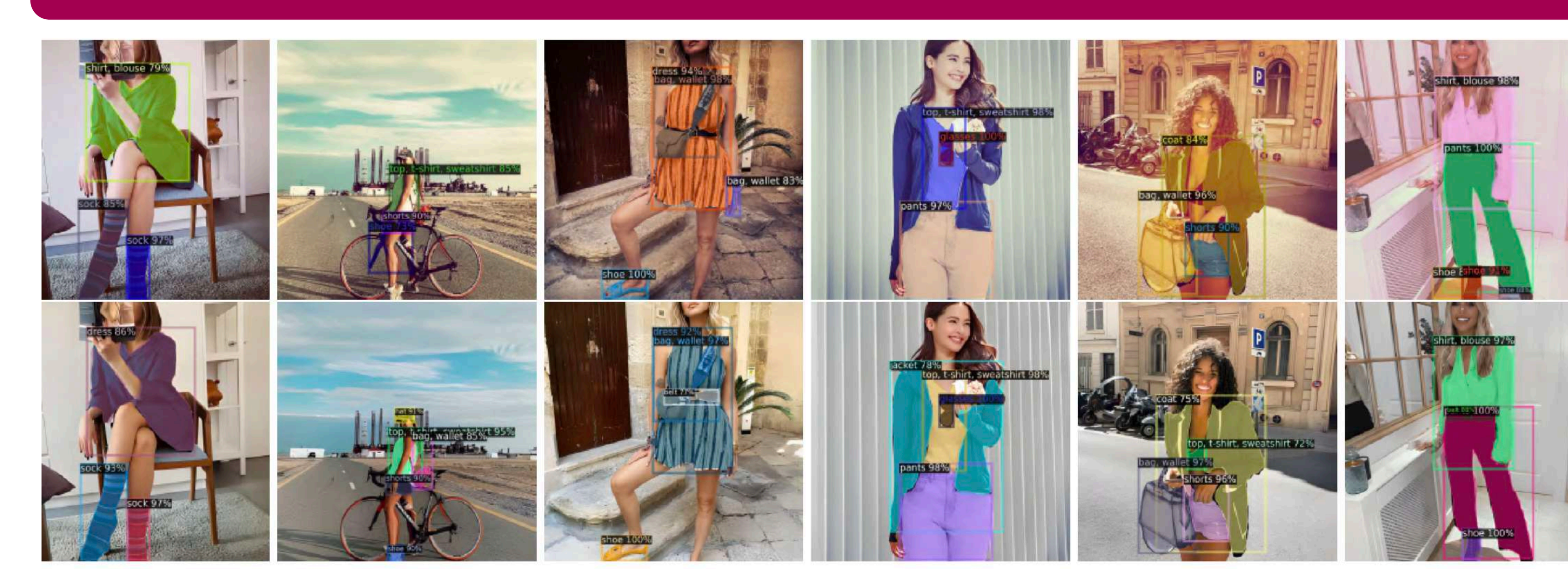
$$\mathcal{L}_{PatchNCE} = \gamma_c\mathcal{L}_C + \gamma_s\mathcal{L}_S$$

$$\mathcal{L}_G = \lambda_p\mathcal{L}_{PatchNCE} + \lambda_c\mathcal{L}_{Cons} + \lambda_a\mathcal{L}^G_{WGAN-GP}$$

Benchmark

Method	SSIM ↑	PSNR ↑	LPIPS ↓	CIE-ΔE ↓
PE [4]	0.748	23.41	0.069	39.55
pix2pix [22]	0.825	26.35	0.048	30.32
CycleGAN [55]	0.819	22.94	0.065	36.59
AngularGAN [44]	0.846	26.30	0.048	31.11
IFRNet [27]	0.864	**30.46**	0.025	20.72
DRIT++ [32]	0.626	16.23	0.162	47.95
GcGAN [11]	0.838	21.75	0.060	38.54
FastCUT [38]	0.763	20.08	0.083	39.86
CUT [38]	0.744	20.96	0.081	38.64
CIFR-no-pre-training	**0.888**	29.24	0.02441	20.65
CIFR-no-style-nce	0.859	28.13	0.03426	23.01
CIFR-no-id-reg	0.879	29.40	0.02528	19.82
CIFR-no-consistency	0.874	29.42	0.02708	21.23
CIFR	0.880	30.02	**0.02321**	**19.05**

Impact on Downstream Tasks



Filters		Localization (mAP)						Segmentation (mAP)					
		Top	Shirt	Pants	Dress	Shoe	Glasses	Top	Shirt	Pants	Dress	Shoe	Glasses
1977	Filtered	7.976	0.000	11.348	5.406	16.084	9.505	9.773	0.000	10.228	6.713	13.424	7.657
	R-IFRNet	12.653	6.931	13.871	11.042	24.318	**13.175**	11.521	7.178	12.815	11.978	**19.716**	9.769
	R-CIFR	**13.115**	**10.891**	**15.175**	**11.314**	**24.332**	10.297	**14.088**	**10.561**	**13.307**	**12.866**	19.004	**9.901**
Amaro	Filtered	11.269	2.970	14.132	7.525	21.051	7.525	10.414	3.960	13.508	10.179	15.323	7.525
	R-IFRNet	**13.035**	6.188	13.890	10.144	26.027	10.594	**11.658**	7.426	14.001	11.560	20.232	9.208
	R-CIFR	12.673	**8.168**	**16.006**	**11.083**	**28.626**	**10.693**	11.057	**8.911**	**14.598**	**11.644**	**22.275**	**9.901**
Brannan	Filtered	10.790	2.475	13.228	6.943	19.572	10.990	11.607	3.960	11.484	7.017	15.271	7.168
	R-IFRNet	13.673	6.931	12.895	**10.562**	**26.027**	8.911	13.359	6.436	12.665	**11.453**	**21.615**	8.020
	R-CIFR	**14.999**	**9.901**	**13.516**	9.537	25.709	**11.221**	**15.200**	**11.634**	**13.264**	10.911	20.977	**8.581**
Hudson	Filtered	13.294	5.941	13.512	9.285	24.554	13.861	13.818	5.941	12.437	11.243	18.329	10.693
	R-IFRNet	**15.093**	6.188	13.964	10.664	27.041	13.812	15.558	7.426	**14.420**	**11.863**	21.283	**11.023**
	R-CIFR	14.322	**10.297**	**14.844**	**10.673**	**29.872**	**11.287**	**15.815**	**11.337**	14.308	11.241	**21.654**	10.297
Nashville	Filtered	12.322	6.931	12.110	10.326	21.806	**11.089**	11.432	6.436	11.305	10.927	16.387	8.079
	R-IFRNet	13.707	6.931	14.645	**10.686**	24.811	7.525	14.546	7.426	12.643	10.485	19.994	6.733
	R-CIFR	**15.077**	**9.571**	**15.705**	9.884	**28.064**	10.108	**14.712**	**9.901**	**13.452**	**11.193**	**22.437**	**8.515**
Perpetua	Filtered	14.494	5.941	14.238	7.475	21.133	**14.072**	14.628	5.941	12.202	8.376	17.263	**12.208**
	R-IFRNet	15.407	6.188	13.768	**11.634**	26.154	12.541	15.879	6.931	12.932	**11.997**	19.264	10.693
	R-CIFR	**16.903**	**8.168**	**15.880**	11.541	**28.047**	13.333	**16.939**	**9.406**	**13.186**	11.861	**22.065**	10.033
Valencia	Filtered	12.481	6.188	12.105	9.010	23.083	9.901	12.558	7.426	10.671	9.036	18.131	7.683
	R-IFRNet	**14.490**	6.436	**15.904**	10.771	27.347	10.337	14.315	7.178	14.138	**12.291**	20.624	9.743
	R-CIFR	14.467	**9.901**	14.809	**11.735**	**30.238**	**10.693**	**14.932**	**9.653**	**14.464**	12.263	**23.358**	**10.198**
X-Pro II	Filtered	13.604	6.188	12.555	8.465	21.637	**12.752**	12.389	6.188	11.111	8.540	16.369	9.795
	R-IFRNet	**15.252**	8.168	13.746	10.815	25.722	11.116	15.751	8.911	12.605	12.546	19.757	9.003
	R-CIFR	15.189	**9.818**	**14.538**	**12.397**	**27.538**	12.488	**16.253**	**9.571**	**13.360**	**13.385**	**21.394**	**10.078**
Original	-	17.639	9.941	17.425	13.223	30.868	17.471	18.758	10.178	16.432	15.509	24.937	15.278