

---

# Recursive Self-Improvement for Camera Image and Signal Processing Pipeline

---

Chandrajit Bajaj

Yi Wang

Yunhao Yang

Yuhan Zheng

## Abstract

Current camera image and signal processing pipelines (ISPs), including deep trained versions, tend to apply a single filter that is uniformly applied to the entire image. This despite the fact that most acquired camera images have spatially heterogeneous artifacts. This spatial heterogeneity manifests itself across the image space as varied Moiré ringing, motion-blur, color-bleaching or lens based projection distortions. Moreover, combinations of these image artifacts can be present in small or large pixel neighborhoods, within an acquired image. Here, we present a deep reinforcement learning model that works in learned latent subspaces, recursively improves camera image quality through a patch-based spatially adaptive artifact filtering and image enhancement. Our RSE-RL model views the identification and correction of artifacts as a recursive self-learning and self-improvement exercise and consists of two major sub-modules: (i) The latent feature sub-space clustering/grouping obtained through an equivariant variational auto-encoder enabling rapid identification of the correspondence and discrepancy between noisy and clean image patches. (ii) The adaptive learned transformation controlled by a trust-region soft actor-critic agent that progressively filters and enhances the noisy patches using its closest feature distance neighbors of clean patches. Artificial artifacts that may be introduced in a patch-based ISP, are also removed through a reward based de-blocking recovery and image enhancement. We demonstrate the self-improvement feature of our model by recursively training and testing on images, wherein the enhanced images resulting from each epoch provide a natural data augmentation and robustness to the RSE-RL training-filtering pipeline.

## 1 Introduction

Digital camera’s Image and Signal Processing(ISP) pipeline commonly relies on specialized digital signal processors for image processing. They are used in converting RAW acquired images, captured by the camera’s digital sensors into conventional RGB or JPEG images. Camera manufacturers have been pursuing and requiring the development of sophisticated filters as part of their camera ISP to resolve diverse image artifacts (distortions) during the conversion process. These image filters include methods for demosaicing(35; 18), deblurring(7), white balancing(47), color correction(34), etc.. Today’s digital camera ISPs however need to be even more sophisticated. With increased image resolutions, image artifacts are naturally heterogeneous, as there is a mixing of distortions caused by sensors, lenses, motion etc. For instance, Fixed-pattern Noise (FPN) is a known issue which commonly refers to as Dark Signal Non-Uniformity (DSNU) and Photo Response Non-Uniformity (PRNU)(6). Another example would be *Bayer Filter artifacts* which occur when demosaicing Color-Filter Array(CFA). The two typical artifacts from Bayer Filtering are false color, an abrupt color shift along edges that prevents good interpolation, and *Moiré ringing*.

Deep learning approaches have progressively replaced the image and signal processing applied in conventional computational photography tasks. For instance, with low-level details and hierarchi-

cal structures of neural network implemented, one could achieve superior performance for image deblurring (e.g. (3; 51; 45; 46)) and deblurring (e.g. (33; 44)) tasks. Nevertheless, most trained and deep learned prior solutions rely on the assumption that there is a single image artifact per RAW image which needs to be diagnosed and filtered. Heterogeneous camera image artifacts, especially mixtures, are insensitive to pixel locations. The assumption is widely accepted with image signal processing tasks, and particularly agreed within image artifact removal problems, including many of the deblurring tasks. For a camera ISP, however, this assumption fails to capture spatial heterogeneous acquired corruption caused by a mismatch of acquisition settings and environmental lighting in realistic scenes. Moreover, one can hardly tell if the artifacts come from sensor limitations, environmental changes (day v.s. night) or post processing such as lossy compression. Such varied real world image processing issues motivates us to adapt a recursive self-improving machine learning approach for the next generation of camera ISPs.

In this paper, we present a deep reinforcement learning model that works in learned latent subspaces, recursively improves camera image quality through a patch-based spatially adaptive artifact filtering and image enhancement. Our RSE-RL model (section 3) views the identification and correction of artifacts as a recursive self-learning and self-improvement exercise and consists of two major sub-modules: (i) The latent feature sub-space clustering/grouping obtained through an equivariant variational auto-encoder enabling rapid identification of the correspondence and discrepancy between noisy and clean image patches. (ii) The adaptive learned transformation controlled by a trust-region soft actor-critic agent that progressively filters and enhances the noisy patches using its closest feature distance neighbors of clean patches. Artificial artifacts that may be introduced in a patch-based ISP, are also removed through a reward based de-blocking recovery and image enhancement. We demonstrate (section 4) the performance of our RSE-RL model including the self-improvement features by recursively training and testing on images, wherein the enhanced images resulting from each epoch provide a natural data augmentation and robustness to the RSE-RL training- filtering pipeline.

## 2 Related Work

### 2.1 Source of Camera Image Artifacts

- **Optical Aberrations:** The optical instrument has its own limitation. For instance, distortion and blur occur when the lens has a spherical aberration. Chromatic aberration is a failure of a lens to focus all colors to the same point. Vignetting is a reduction in illumination and saturation toward the periphery compared to the image center. Lens flare occurs in photography when rays from a very bright light source have internal reflections and scatter in the lens system of a camera, overlaying the captured image with artifacts such as blown out starbursts, colored shapes, rain-bow patterns and haze.
- **Light and Sensor Capturing Issues:** Noise, contrast and atmospheric haze in environments definitely affect the result of RAW image capturing. Fixed-pattern Noise (FPN) is a known issue which commonly refers to Dark Signal Non-Uniformity (DSNU) and Photo Response Non-Uniformity (PRNU)(6).
- **Bayer Filter Artifacts:** These artifacts occur when demosaicing Color-Filter Array(CFA). Two typical artifacts are false color, which is abrupt color shift along edges that prevents a good interpolation, and Moiré Artifact, which is caused by discretization of continuous signals and yields repeated patterns.

There have been dedicated efforts to calibrate instruments and to invent more sophisticated camera ISP algorithms for inevitable artifacts removal (e.g. (37)). With the advent of powerful computing units, deep learning algorithms are applicable in more and more aspects of research, so does in computational photography (c.f. (53)).

**Image Filtering and Enhancement** The ubiquity of noise in digital photos leads to a fast-growing image denoising problem. In the early years, there are traditional methods that apply Gaussian blurring, TV regularization (41), or a coefficient transform in Fourier domain (43). However, it is the idea of non-local mean denoising from *Buades* et al. (11) that truly made a gigantic leap in denoising performance. The non-local mean method is built upon self-similarity and redundant information over realistic images. Later on, another non-local denoising approach, referred to as

BM3D(13) uses the sample idea and exploits the sparsity further. There are also discussions with respect to patch-based scheme (19). As the deep learning era arrives, more and more works are using Convolutional Neural Network(CNN) or Generative Adversarial Network(GAN) that beats most of the classical, sophisticated methods (55; 21; 50; 38; 24). However, due to the expensive capturing procedure, as noise is diverse in large-scale realistic photos, most of the prior works fall into the study in the synthesized domain only. For instance, the most common AWGN noise model cannot effectively remove noise from real images, as discussed in earlier benchmarks (39). How to fill in the gap between the synthesized noise model and real-world noisy images remains an open question.

**Camera ISP** There are numerous works related to camera image signal processing, where different types of articulated modelings are applied to different vision tasks. Among all of them, some related works such as color demosaicing(e.g. (27)), image denoising(e.g. (45)), auto white balance correction(e.g. (4)), removing lossy compression artifacts (e.g. (17)), *etc.*, have been separately discussed under deep learning settings. Within the scope of image denoising, there is a novel work, referred to as CycleISP(54), which develops a generative model to generate synthesized realistic image data that is both forward and reversed. Moreover, solving multi-task image enhancement is possible: (42) is an early work that focuses on jointly solving demosaicing and denoising problems by applying deep learning schemes from the source sampling of sensors. Later, a recent work (25) presented a single deep learning model containing 5 parallel learning levels to replace the entire camera ISP pipeline, referred to as PyNET. The input RAW image from a cellphone is aligned with a DSLR camera output as the supervised training data and PyNET outputs a visually high-quality sRGB image.

**Latent Subspace Learning** One of our major contributions is to learn a latent encoding for maximizing the usage of self-similarity in natural images via probing subspaces. There are at least two different branches of workflows. One branch is focusing on a direct disentanglement of latent space clustering under variational. For example, in (20; 8; 15), the hierarchical structure of latent space is explored to obtain richer representations compared to a single prior. Another branch of works is a direct fusion of deep learning models and machine learning methods. They developed algorithms that perform clustering or learning a mixture model within the latent space encoding, including Hard K-means Clustering (52), Soft K-means Clustering (26), Gaussian Mixture Model VAE(14), and direct subspace clustering VAE (31). All of these exemplary models manage to learn from visual recognition and classification tasks, yet their performance in a more realistic image denoising task needs to be testified.

### 3 Recursive Self-Enhancing Camera ISP

In this section, we describe our reinforcement learning model that recursively improves in spatially adaptive, heterogeneous image artifact filtering and image enhancement. We target the problem of resolving image artifacts, specifically on image denoising, but our approach can be extended to solve other comprehensive tasks, such as generating RAW to sRGB images.

In the paper, we assume the observed image  $I_{obs}$  is obtained via the following mixture modeling:

$$f(I_{obs}) = I_{gt} + \sum_{s=1}^S \Sigma_s \odot M_s. \quad (1)$$

The function  $f$  is an identity function when we are performing sRGB to sRGB image denoising tasks. The noise  $\Sigma_s$  and the mask matrix  $M_s$  are independent and blind to the model ( $\odot$  refers to element-wise product). Moreover, we do not rely on the underlying distribution  $\Sigma_s$ . Except for the synthesized dataset, we cannot obtain the accurate number of artifact types  $S$ , and  $S$  is a hyper-parameter in most scenarios. We would like to present and discuss how to disentangle and filter from (1) using our RSE-RL.

**Overall Pipeline of RSE-RL** Our recursive self-improving camera ISP in Figure 1 is a multiple latent subspaces variational autoencoder. For every input image  $I_{obs} \in \mathbb{R}^{H \times W \times C}$ , we first divide the input image into  $D$  by  $D$  patches  $P \in \mathbb{R}^{D \times D \times C}$  with overlaps allowed. For every  $D$  by  $D$  patch  $P$  in the image, we denote its location mask  $H_n$  in the original image domain.  $P$  relates to the observed

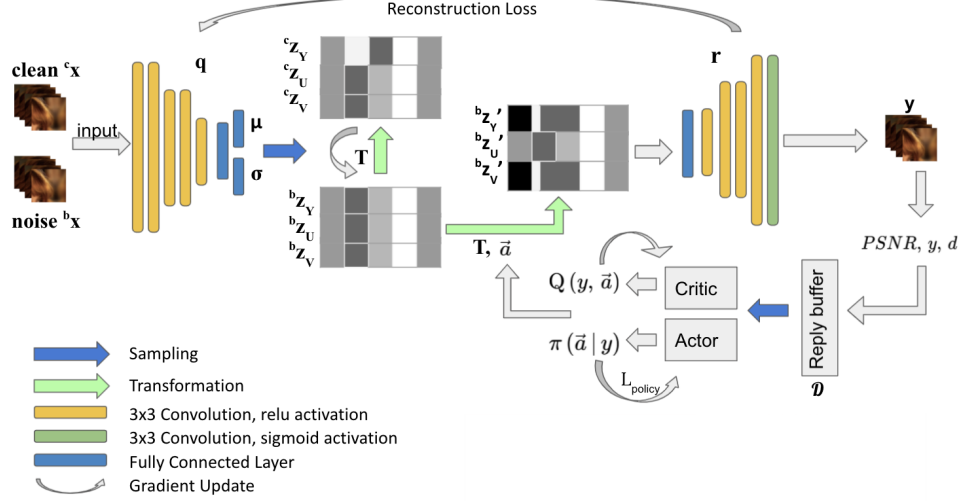


Figure 1: The overall pipeline of our RSE-RL: For each given observed image, we split the image into local patches and feed every patch as a stack into the encoding network. The latent space is divided into three subspaces, the encoder projects the YUV features of the patches onto three latent subspaces  $Z_y$ ,  $Z_u$ , and  $Z_v$ . Both the clean patches and noisy patches are projected onto the three spaces. A set of transformations  $T$  are learned to transform the latent representation of the noisy patches to a corresponding representation of the clean patches in all three subspaces. The transformed noisy representations are sent to the decoder for image reconstructions. After the denoised images are constructed, a PSNR is calculated and used to obtain the reward for a soft-actor-critic reinforcement learning model. The RL model uses the distance from the target PSNR and actual PSNR as the reward to adjust the trainable weights in the transformation  $T$ . Hence we have a self-enhancing image denoising network.

image in the form  $P \stackrel{def}{=} I_{obs} \odot H_n$ . Even though we cannot probe the magnitude of the artifacts, it can be approximated within patch with a single dominant modeling, namely:

$$f(P) = f(I_{obs} \odot H_n) = I_{gt} \odot H_n + \sum_{s=1} \Sigma_s \odot M_s \odot H_n, \approx (I_{gt} + \Sigma_i \odot M_i) \odot H_n. \quad (2)$$

Our RSE-RL network is learning from image patches  $P$  and  $I_{gt} \odot H_n$ . Both clean patches  $P^c$  and noisy patches  $P^b$  are fed into the network. Each batch of image patches is fed into our encoder  $q(P|\vec{\theta})$ , with parameter  $\vec{\theta}$ . The encoder will generate three latent vectors  $\vec{z}_y$ ,  $\vec{z}_u$ , and  $\vec{z}_v$  on three subspaces that we defined. The three subspaces preserve the features on three channels of the patches. The three channels are one luma component Y and two chrominance components, U (blue projection) and V (red projection) respectively.

On each subspace, this is a transformation function  $T$  that learns to transform a noisy projection to a clean projection:  $\vec{z}_1^c = T(\vec{z}_1^b)$ . The transformation  $T$  is trained so that the transformed projection approaches the clean patch representation  $T(\vec{z}^b) \rightarrow \vec{z}^c$ .

Then, based on the transformed latent space projections, we set up the decoder  $r(T_y(\vec{z}_y^c), T_u(\vec{z}_u^c), T_v(\vec{z}_v^c)|\vec{\psi}_i)$  with parameter  $\vec{\psi}$ . Each patch's latent encoding representation is fed into the decoder to decode independently. The decoded latent vectors are integrated to reconstruct the patches. The output is the reconstructed blocks that approximate  $I_{gt} \odot H_n$ .

**Loss Function** In practical implementation, our training process optimizes different loss functions. For training the network, our training loss  $\mathcal{L}_{vae}$  consists of two parts: the evidence lower-bound(ELBO) (30), including the data fitting term and the KL loss, and the regularization term. The gradient computed from  $\mathcal{L}_{vae}$  is used to update all the parameters in the network, including the encoder, decoder, and the transformation functions.

$$\mathcal{L}_{vae} = \mathcal{L}_{MSE} + \mathcal{L}_{KL} + \mathcal{L}_{reg}. \quad (3)$$

where

$$\mathcal{L}_{MSE} = \|r(q(P^c|\vec{\theta})|\vec{\psi}) - P^c\|_F^2 + \|r(T(q(P^b|\vec{\theta}))|\vec{\psi}) - P^c\|_F^2, \quad (4)$$

$$\mathcal{L}_{KL} = -\frac{1}{2} \sum_i (1 + \hat{\sigma}_i - \exp(\hat{\sigma}_i) - \hat{\mu}_i^2), \quad (5)$$

$$\mathcal{L}_{reg} = \lambda_{reg} (\|\vec{\theta}\|_2^2 + \|\vec{\psi}\|_2^2). \quad (6)$$

The transformation functions  $T_y$ ,  $T_u$ , and  $T_v$  are trained under the identical terms. On each latent subspace,  $\mathcal{L}_{tran}$  is computed from the noisy patch projection and clean patch projection. Three transformation functions are optimized separately, where the loss function on each latent subspace is only computed by the projections in the corresponding latent subspace. Hence, three losses  $\mathcal{L}_{tran_y}$ ,  $\mathcal{L}_{tran_u}$ , and  $\mathcal{L}_{tran_v}$  are computed. The transformation functions are optimized by the corresponding loss:  $T_y$  is optimized by  $\mathcal{L}_{tran_y}$ .

$$\mathcal{L}_{tran} = \|q(P^c|\vec{\theta}) - q(P^b|\vec{\theta})\|_F^2 \quad (7)$$

**Patch Assembly and Block Artifacts** Upon getting the reconstructed blocks from decoders, we merge these blocks back to the original image. If one naively concatenates two adjacent patches without any tolerance of the overlaps generated when creating them, one could observe block artifacts if these two patches pass different decoders. We apply post processing to remove block artifacts generated from concatenation. For overlapping regions, we average the pixel output based on the distance between the pixel and the overlapped patch centers via linear interpolation. To further remove block artifacts, an reinforcement learning algorithm can be applied to learn a better size of overlapping region.

**Reinforcement Learning** The Soft Actor-Critic (SAC) reinforcement learning algorithm (22; 23) is utilized for self-enhancing image denoising. The RL model learns the action  $a$ , which is the concatenation of the weight vectors  $a_y$ ,  $a_u$ , and  $a_v$  that used to adjust the transformation functions  $T_y$ ,  $T_u$ , and  $T_v$ . These trainable weights are multiplied with the weight vectors, hence we define the action as:

$$T_s = \vec{a}_s \odot T'_s \quad s \in \{y, u, v\} \quad (8)$$

where  $T_i$  is the updated  $i$ th weight and  $T'_i$  is the previous weight. The algorithm continuously updates these trainable weights to enhance the transformation from noisy to clean patch representations, which consequently improves the final denoising performance. Alternatively, the RL model can learn to reduce the blocking artifacts.

A central feature of SAC is entropy regularization. This policy is trained to maximize a trade-off between expected return and entropy, a measure of randomness in the policy. This is connected to the exploration-exploitation trade-off: increasing entropy results in more exploration, which can accelerate learning speed. It can also prevent the policy from prematurely converging to a bad local optimum. In SAC, an entropy bonus is reflected in  $Q^\pi$ :

$$Q^\pi \approx r + \gamma(Q^\pi(I', \tilde{a}') - \alpha \log \pi(\tilde{a}'|I')), \tilde{a}' \sim \pi(\cdot|I') \quad (9)$$

where  $I$  is an input image and  $I'$  is the resulting image when the transformation in the network has the set of weight vectors  $a$ ,  $\alpha > 0$  is the trade-off coefficient and  $\log \pi(\tilde{a}'|I')$  is the defined entropy.

SAC learns a policy  $\pi$  and two Q functions  $Q$  concurrently. In particular, the policy is learned by the Critic network which maximizes  $V^\pi(I)$ , and the Q functions is learned by the Actor network which minimizes a sample-approximated MSBE  $L(\Phi_i, D)$ .

$$V^\pi(I) = E_{a \sim \pi}[Q^\pi(s, a)] + \alpha H(\pi(\cdot|I)) = E_{a \sim \pi}[Q^\pi(I, a)] - \alpha \log \pi(\cdot|I) \quad (10)$$

SAC sets up the MSBE loss for each of the two Q-functions:

$$L(\Phi_i, D) = E [((Q(I, a) - f(r, I', d))^2)], \quad (I, a, r, I', d) \sim D \quad (11)$$

where  $d$  is the done signal to set a terminating state, the target is given by

$$f(r, I', d) = r + \gamma(1 - d)(\min_{j=1,2}(Q_j(I', \tilde{a}') - \alpha \log \pi_\theta(\tilde{a}'|I'))), \quad \tilde{a}' \sim \pi(\cdot|I') \quad (12)$$

In this RL algorithm, the states are the images, and the reward  $r$  is given by computing the distance between the actual PSNR and the target PSNR, namely

$$r = k \cdot (PSNR - PSNR_t) + c \quad (13)$$

where  $k$  and  $c$  are constants for adjusting the scale of the reward and  $PSNR$  and  $PSNR_t$  are the averaged actual PSNR from testing and the target PSNR of the entire patch-assembled image, respectively. The RL model learns to enhance the transformation and reduce the blocking artifacts. Thus, to maximize the reward, the model is optimized so that the average PSNR of the testing images approaches the target PSNR.

## 4 Experiments

To justify the performance of our RSE-RL algorithm, two datasets are considered in our experiments: Synthesized Noisy CelebFaces Attributes (CelebA) Dataset(36) and Smartphone Image Denoising Dataset (SIDD)(2). Gaussian noise is applied to images in the synthesized noisy CelebA dataset, while the SIDD dataset contains realistic artifacts generated by smartphone cameras. An objective of our experiments is to exhibit a performance enhancement by utilizing the RSE-RL network structure. For comparison, an ordinary Variational Autoencoder is trained under identical experimental settings in each experiment, referred to as *Single Decoder VAE* for comparison. For more details in training, we refer to the supplementary materials.

### 4.1 Our Networks with CelebA Patches

**Dataset Construction** Large-scale CelebFaces Attributes (CelebA) Dataset is adapted for our experiment. CelebA\_HQ/256 dataset, which consists of images with size  $256 \times 256$  pixels, are selected from the CelebA dataset. These images are then sampled into two sub-datasets: a training set with 2250 images and a validation set with 11250 images. Heterogeneous artifacts generator is applied to the training set to generate noisy images from celebA, referred to as the synthesized noisy CelebA dataset. Gaussian noise is generated on these images utilizing OpenCV(9). Each image, both noisy and ground truth, in the training set is divided into  $16 \times 16$  pixels patches, with 4 pixels overlap with the surrounding patches.



Figure 2: CelebA Denoising Results: images on the top row are images that contain the synthesized artifacts (Gaussian Noise). The images on the bottom row are the denoising result from our RSE-RL.

**Experimental Setup** We test the denoising performance of our RSE-RL over synthesized noisy CelebA dataset. Our encoder projects the patch-based images into the latent space using 5 convolutional layers and 2 fully-connected layers in a subsequent order, while the decoder is a similar structure with 2 upsampling layers and 5 transposed convolutional layers. The model is implemented in Keras(12) and Tensorflow(1), and we use a single 12GB NVIDIA Tesla K80 GPU for training and testing on the synthesized noisy CelebA dataset. One batch has 128 patches that are trained simultaneously. The parameters are optimized in 50 epochs using Adam algorithm(29) with  $\beta_1 = 0.9$  and  $\beta_2 = 0.999$ . The learning rate is set by an exponential decay learning rate scheduler that has an initial rate of 0.001, decay factor of 0.95, and decay step 1000.

The soft actor-critic algorithm for self-enhancing is implemented by using Stable-Baseline3(40) from OpenAI(10). We use OpenAI Gym for setting up the environment. The reward in the environment is computed as stated in Equation 13. The model is trying to maximize the reward by optimizing the

actions that adjusting the trainable weights in the three transformation functions, as stated in Equation 8. The action space is a set of weight vectors within the bound (0.999, 1.001) for minor adjustments. The observation space is the actual PSNR score we achieve. In the synthesized CelebA dataset, the learning rate for the model is set to 0.001. The results before and after the recursive self-enhancing procedure are indicated in Table 1, denoted  $RSE-RL(before)$  and  $RSE-RL$ , respectively.

Table 1: CelebA Results Comparison

	PSNR	SSIM	UQI
Image with Artifacts	16.64	0.5835	0.7327
N2V(32)	21.66	0.7242	0.9249
N2N(38)	26.60	/	/
Single Decoder VAE	26.81	0.7621	0.9604
4-Decoder PS-VAE(53)	28.04	0.8266	0.9704
RSE-RL(before)	28.83	0.8322	0.9721
RSE-RL	<b>29.03</b>	<b>0.8339</b>	<b>0.9731</b>
RSE-RL(S)	28.79	0.8317	0.9717

**Our denoising result on CelebA dataset with synthesized artifacts** We evaluate our result using PSNR, SSIM(49), and UQI(48) scores. In addition, baseline denoising methods, Noise2Void(N2V), Noise2Noise(N2N) and PS-VAE, are applied to the synthesized artifacts dataset, and its denoising results are compared with our networks. For N2V, the default model for 2D RGB images is trained with 400 noisy images for 50 epochs and tested with 1575 noisy images. In the default model, each image is divided into 128  $16 \times 16 \times 3$  patches, so a total of 51,200 patches are fed into the network. For N2N, the pre-trained N2N model for Gaussian noise is tested on 1575 noisy images.

Table 1 provides the results obtained from using 2250 training images and 11250 testing images. The training images are divided into 1 million  $16 \times 16 \times 3$  patches (441 patches per image) and feed into the network. The result of the baseline method is compared to the results of our RSE-RL. We can observe a significant improvement regarding all the quality metrics by utilizing RSE-RL. A visualization of our denoised image results after self-enhancement can be found in Figure 2.

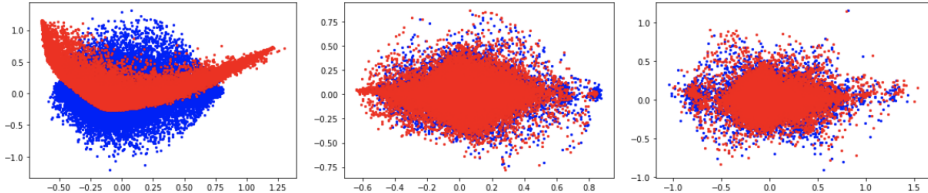


Figure 3: YUV Latent subspaces: the figure shows the YUV latent subspaces from the network. From left to right, it shows the spaces  $Z_Y$ ,  $Z_U$ , and  $Z_V$  respectively. Blue points are noise projections and red points are clean projections. Principle Component Analysis is applied to reduce the latent subspaces dimension into 2 for visualization.

Among the results in Table 1, the RSE-RL achieves the best performance regarding all three metrics. We can observe small enhancements on all three quality metrics after applying the RL model for self-improvement. This demonstrates that our self-enhancing model is able to continue being optimized during the testing stage.

In addition, the results indicate that learning the transformations on the latent spaces is effective. By comparing the three latent subspaces  $Z_Y$ ,  $Z_U$ , and  $Z_V$ , we can observe that the noise has the largest impact on the Y space  $Z_Y$ , which represents the luminance (brightness) of the image. There is no significant difference between the noisy and clean patch representations on the other two subspaces. This indicates that Gaussian noise has the most significant impact on brightness, compare to chrominance (represented by U and V). The visualized latent subspaces can be seen on Figure 3.

Another observation is that RSE-RL’s denoising performance is not significantly affected by the training data size. In order to demonstrate this feature, an experiment using 450 training images (0.2 million patches) is conducted. This network is also tested under 11250 testing images, and the result is shown in Table 1, denoted as *RSE-RL(S)*. This result can be compared with *RSE-RL(before)* to observe how the training data size affect the performance. This observation provides an effective way of utilizing this network to largely reduce training time.

## 4.2 SIDD Denoising Result

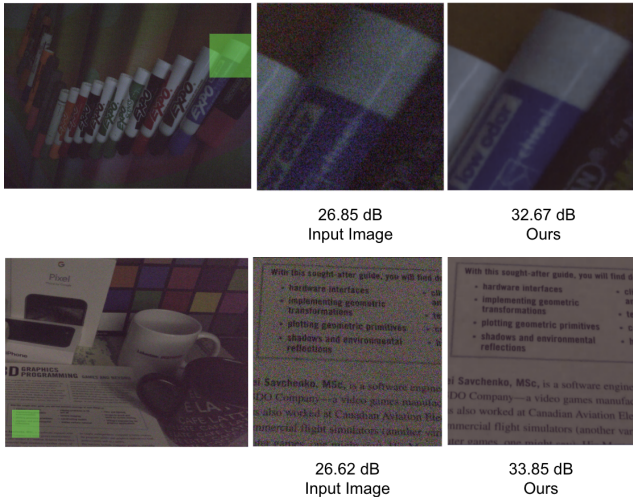


Figure 4: SIDD Denoising Result: to better visualize the denoising result, we zoom in at the area covered by the green rectangle in each figure on the left column. The figures from the left column to the right column are: original noisy images from SIDD, zoomed noisy patches, and zoomed denoising patches, respectively.

**Dataset Construction** Smartphone Image Denoising Dataset (SIDD) is a large-scale high-quality image dataset collected from five representative smartphone cameras under 10 different scenes. Each noisy image collected from smartphone cameras consists of artifacts via different ISO levels, illumination, and lighting conditions, as well as signal-dependent noise. This dataset is a benchmark for denoising algorithms, as its noise is generated under realistic scenarios.

In the experiment, 320 sRGB images are selected for training the network, and the SIDD Benchmark Data is used to evaluate the performance of our network. The SIDD Benchmark Data contains 40 noisy sRGB images and their ground truth. 32 patches are selected from each benchmark image for evaluation. For training the network, each noisy image and its ground truth are divided into  $24 \times 24 \times 3$  patches, with 8 pixels overlapping. 11.19 million patches build up the training data.

**Experiment Setup** The encoder is composed of five convolutional layers and two fully connected layers, while decoders have the inverse structure as the encoder. Each encoder has approximately 2.2 million parameters, and each decoder has 1.6 million parameters. The model is trained on a single 12GB NVIDIA Tesla K80 GPU with a batch size of 128. The parameters are optimized in 20 epochs using the same optimizer as the Synthesized Noisy CelebA dataset defined in Section 4.1.

The reinforcement learning model is identical to the model implemented in 4.1, which adjusts the trainable weights in the three transformation functions. We want to show the improvement by using the self-enhancing RL technique. The results before and after self-enhancement are shown in Table 2, denoted as *RSE-RL(before)* and *RSE-RL*, respectively.

**Our denoising result on SIDD dataset** The results show that our self-enhancing RL model contributes a small enhancement on PSNR, which demonstrates that our RL model is able to improve the



Table 2: SIDD sRGB to sRGB Results (Small Scale)

	PSNR	SSIM		PSNR	SSIM
Noisy Image	31.18	0.831	BM3D(13)	25.65	0.685
NLM(11)	26.75	0.699	KSVD(5)	26.88	0.842
DANet(16)	39.25	0.955	RDB-Net(56)	38.11	0.945
Single Decoder VAE	31.89	0.874	4-Decoder PS-VAE(53)	32.17	0.881
RSE-RL(before)	32.38	<b>0.891</b>	RSE-RL	<b>32.53</b>	0.887

denoising results. Since we are only involving PSNR in the reward function, we can only observe some improvements in terms of PSNR.

Table 2 also lists a set of benchmark denoising methods and deep learning methods which are used to compare against our network. In the table, noisy images are the images before denoising procedures; BM3D, NLM, and KSVD are the benchmark non-DL results; DANet and RDB-Net are two of the state-of-art deep learning methods used for comparison with our method. The performance of our model is significantly better than traditional methods. The visualized results of self-enhanced denoised images can be seen in figure 4.2. BM3D, NLM, KSVD, DANet, and RDB-Net are the benchmark results. As for efficiency comparison, our RSE-RL only contains 2.5 million parameters in total, whereas DANet contains  $\sim 60$  million parameters, leading our network train much faster compared to the state-of-the-art structure.

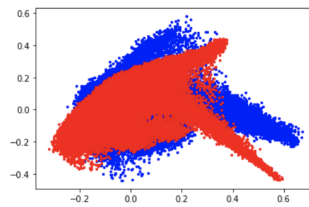


Figure 5: SIDD Y Latent subspace: the figure shows the Y latent subspace  $Z_Y$  from the network. Blue points are noise projections and red points are clean projections. Principle Component Analysis is applied to reduce the dimension of the latent subspace into 2 for visualization. We can observe a rotation on the latent space between the noisy and clean patch projections.

The images from SIDD consist of the same set of realistic artifacts generated by smartphone cameras. This leads to the same transformation on the latent space for every patch since each patch consists of the same types of noises. From figure 5 we can observe a rotation between the noisy patch projections and the clean patch projections on the latent space which is the transformation on the latent space.

## 5 Conclusions and Discussions of Broader Impact

**Overview of our work** We have presented our RSE-RL model, a self-improving camera ISP built upon policy learning. The patch-based transformation is trained in decomposed subspace to identify and rectify different types of artifacts progressively and respectively. We define the action and reward for a self-enhancement framework and further discussed its potential in real-world image data. Nonetheless, our work is an early-stage exploration, where the transformation is mostly in the linear components. We are moving toward considering patch ordering or other complicated environmental settings to further strengthen our work.

**Discussions of Broad Impact** The method we proposed is an RL-based solution to low-level vision tasks. There might exist an accuse of collecting private photo information from users when the our research extends to application side. But our current research don't envision any broad ethical issues surrounding our largely mathematical learning technique. We would of course welcome any issues the reviewers might raise, and promise to address them.

*Acknowledgement:* This research was supported in part by a grant from NIH - R01GM117594, by the Peter O'Donnell Foundation and in part from a grant from the Army Research Office accom-

plished under Cooperative Agreement Number W911NF-19-2-0333. The views and conclusions contained in this document are those of the authors and should not be interpreted as representing the official policies, either expressed or implied, of the Army Research Office or the U.S. Government. The U.S. Government is authorized to reproduce and distribute reprints for Government purposes notwithstanding any copyright notation herein.

## References

- [1] Martín Abadi et al. TensorFlow: Large-scale machine learning on heterogeneous systems, 2015. Software available from tensorflow.org. 6
- [2] Abdelrahman Abdelhamed, Stephen Lin, and Michael S. Brown. A high-quality denoising dataset for smartphone cameras. In *2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 1692–1700, 2018. 6
- [3] Abdelrahman Abdelhamed, Radu Timofte, and Michael S Brown. Ntire 2019 challenge on real image denoising: Methods and results. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops*, pages 0–0, 2019. 2
- [4] Mahmoud Afifi and Michael S Brown. Deep white-balance editing. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 1397–1406, 2020. 3
- [5] Michal Aharon, Michael Elad, and Alfred Bruckstein. K-svd: An algorithm for designing overcomplete dictionaries for sparse representation. *IEEE Transactions on Signal Processing*, 54(11):4311–4322, 2006. 9
- [6] European Machine Vision Association et al. Standard for characterization of image sensors and cameras. *EMVA Standard*, 1288, 2010. 1, 2
- [7] Yeul-Min Baek, Dong-Chan Cho, Jin-Aeon Lee, and Whoi-Yul Kim. Noise reduction for image signal processor in digital cameras. In *2008 International Conference on Convergence and Hybrid Information Technology*, pages 474–481. IEEE, 2008. 1
- [8] Diane Bouchacourt, Ryota Tomioka, and Sebastian Nowozin. Multi-level variational autoencoder: Learning disentangled representations from grouped observations. In Sheila A. McIlraith and Kilian Q. Weinberger, editors, *Proceedings of the Thirty-Second AAAI Conference on Artificial Intelligence, (AAAI-18), the 30th innovative Applications of Artificial Intelligence (IAAI-18), and the 8th AAAI Symposium on Educational Advances in Artificial Intelligence (EAAI-18), New Orleans, Louisiana, USA, February 2-7, 2018*, pages 2095–2102. AAAI Press, 2018. 3
- [9] G. Bradski. The OpenCV Library. *Dr. Dobb’s Journal of Software Tools*, 2000. 6
- [10] Greg Brockman, Vicki Cheung, Ludwig Pettersson, Jonas Schneider, John Schulman, Jie Tang, and Wojciech Zaremba. Openai gym. *arXiv preprint arXiv:1606.01540*, 2016. 6
- [11] Antoni Buades, Bartomeu Coll, and J-M Morel. A non-local algorithm for image denoising. In *2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR’05)*, volume 2, pages 60–65. IEEE, 2005. 2, 9
- [12] François Chollet et al. Keras. <https://keras.io>, 2015. 6
- [13] Kostadin Dabov, Alessandro Foi, Vladimir Katkovnik, and Karen Egiazarian. Image denoising by sparse 3-D transform-domain collaborative filtering. *IEEE Transactions on image processing*, 16(8):2080–2095, 2007. 3, 9
- [14] Nat Dilokthanakul, Pedro AM Mediano, Marta Garnelo, Matthew CH Lee, Hugh Salimbeni, Kai Arulkumaran, and Murray Shanahan. Deep unsupervised clustering with gaussian mixture variational autoencoders. *arXiv preprint arXiv:1611.02648*, 2016. 3
- [15] Fei Ding, Feng Luo, and Yin Yang. Clustering by directly disentangling latent space, 2019. 3
- [16] Jun Fu, Jing Liu, Haijie Tian, Yong Li, Yongjun Bao, Zhiwei Fang, and Hanqing Lu. Dual attention network for scene segmentation, 2019. 9
- [17] Leonardo Galteri, Lorenzo Seidenari, Marco Bertini, and Alberto Del Bimbo. Deep generative adversarial compression artifact removal. In *Proceedings of the IEEE International Conference on Computer Vision*, pages 4826–4835, 2017. 3
- [18] Michaël Gharbi, Gaurav Chaurasia, Sylvain Paris, and Frédo Durand. Deep joint demosaicking and denoising. *ACM Transactions on Graphics (TOG)*, 35(6):1–12, 2016. 1
- [19] Gabriela Ghimpeanu, Thomas Batard, Tamara Seybold, and Marcelo Bertalmío. Local denoising applied to raw images may outperform non-local patch-based methods applied to the camera output. *Electronic Imaging*, 2016(18):1–8, 2016. 3
- [20] Prasoon Goyal, Zhiting Hu, Xiaodan Liang, Chenyu Wang, and Eric P Xing. Nonparametric variational auto-encoders for hierarchical representation learning. In *Proceedings of the IEEE International Conference on Computer Vision*, pages 5094–5102, 2017. 3

- [21] Shi Guo, Zifei Yan, Kai Zhang, Wangmeng Zuo, and Lei Zhang. Toward convolutional blind denoising of real photographs. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 1712–1722, 2019. 3
- [22] Tuomas Haarnoja, Aurick Zhou, Pieter Abbeel, and Sergey Levine. Soft actor-critic: Off-policy maximum entropy deep reinforcement learning with a stochastic actor. In *International Conference on Machine Learning*, pages 1861–1870. PMLR, 2018. 5
- [23] Tuomas Haarnoja, Aurick Zhou, Kristian Hartikainen, George Tucker, Sehoon Ha, Jie Tan, Vikash Kumar, Henry Zhu, Abhishek Gupta, Pieter Abbeel, et al. Soft actor-critic algorithms and applications. *arXiv preprint arXiv:1812.05905*, 2018. 5
- [24] Yingkun Hou, Jun Xu, Mingxia Liu, Guanghai Liu, Li Liu, Fan Zhu, and Ling Shao. NLH: A blind pixel-level non-local method for real-world image denoising. *IEEE Transactions on Image Processing*, 29:5121–5135, 2020. 3
- [25] Andrey Ignatov, Luc Van Gool, and Radu Timofte. Replacing mobile camera isp with a single deep learning model. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops*, pages 536–537, 2020. 3
- [26] Mohammed Jabi, Marco Pedersoli, Amar Mitiche, and Ismail Ben Ayed. Deep clustering: On the link between discriminative models and k-means. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2019. 3
- [27] Daniel Khashabi, Sebastian Nowozin, Jeremy Jancsary, and Andrew W Fitzgibbon. Joint demosaicing and denoising via learned nonparametric random fields. *IEEE Transactions on Image Processing*, 23(12):4968–4981, 2014. 3
- [28] Sung Deuk Kim, Jaeyoun Yi, Hyun Mun Kim, and Jong Beom Ra. A deblocking filter with two separate modes in block-based video coding. *IEEE transactions on circuits and systems for video technology*, 9(1):156–160, 1999. 14, 17
- [29] Diederik P. Kingma and Jimmy Ba. Adam: A method for stochastic optimization. In *ICLR (Poster)*, 2015. 6
- [30] Diederik P. Kingma and Max Welling. Auto-Encoding Variational Bayes. In *2nd International Conference on Learning Representations, ICLR 2014, Banff, AB, Canada, April 14-16, 2014, Conference Track Proceedings*, 2014. 4
- [31] Jack Klys, Jake Snell, and Richard Zemel. Learning latent subspaces in variational autoencoders. In *Advances in Neural Information Processing Systems*, pages 6444–6454, 2018. 3
- [32] Alexander Krull, Tim-Oliver Buchholz, and Florian Jug. Noise2void-learning denoising from single noisy images. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 2129–2137, 2019. 7
- [33] Orest Kupyn, Tetiana Martyniuk, Junru Wu, and Zhangyang Wang. Deblurgan-v2: Deblurring (orders-of-magnitude) faster and better. In *Proceedings of the IEEE International Conference on Computer Vision*, pages 8878–8887, 2019. 2
- [34] Ngai Ming Kwok, HY Shi, Quang Phuc Ha, Gu Fang, SY Chen, and Xiuping Jia. Simultaneous image color correction and enhancement using particle swarm optimization. *Engineering Applications of Artificial Intelligence*, 26(10):2356–2371, 2013. 1
- [35] Xin Li, Bahadır Gunturk, and Lei Zhang. Image demosaicing: A systematic survey. In *Visual Communications and Image Processing 2008*, volume 6822, page 68221J. International Society for Optics and Photonics, 2008. 1
- [36] Ziwei Liu, Ping Luo, Xiaogang Wang, and Xiaoou Tang. Deep learning face attributes in the wild. In *Proceedings of International Conference on Computer Vision (ICCV)*, December 2015. 6
- [37] Henrique S Malvar, Li-wei He, and Ross Cutler. High-quality linear interpolation for demosaicing of bayer-patterned color images. In *2004 IEEE International Conference on Acoustics, Speech, and Signal Processing*, volume 3, pages iii–485. IEEE, 2004. 2
- [38] Nick Moran, Dan Schmidt, Yu Zhong, and Patrick Coady. Noisier2noise: Learning to denoise from unpaired noisy data. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 12064–12072, 2020. 3, 7
- [39] Tobias Plotz and Stefan Roth. Benchmarking denoising algorithms with real photographs. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 1586–1595, 2017. 3
- [40] Antonin Raffin, Ashley Hill, Maximilian Ernestus, Adam Gleave, Anssi Kanervisto, and Noah Dormann. Stable baselines3. <https://github.com/DLR-RM/stable-baselines3>, 2019. 6
- [41] Leonid I Rudin, Stanley Osher, and Emad Fatemi. Nonlinear total variation based noise removal algorithms. *Physica D: nonlinear phenomena*, 60(1-4):259–268, 1992. 2

- [42] Eli Schwartz, Raja Giryes, and Alex M Bronstein. Deepisp: Toward learning an end-to-end image processing pipeline. *IEEE Transactions on Image Processing*, 28(2):912–923, 2018. 3
- [43] Eero P Simoncelli and Edward H Adelson. Noise removal via bayesian wavelet coring. In *Proceedings of 3rd IEEE International Conference on Image Processing*, volume 1, pages 379–382. IEEE, 1996. 2
- [44] Maitreya Suin, Kuldeep Purohit, and AN Rajagopalan. Spatially-attentive patch-hierarchical network for adaptive motion deblurring. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 3606–3615, 2020. 2
- [45] Chunwei Tian, Lunke Fei, Wenxian Zheng, Yong Xu, Wangmeng Zuo, and Chia-Wen Lin. Deep learning on image denoising: An overview. *Neural Networks*, 2020. 2, 3
- [46] Diego Valsesia, Giulia Fracastoro, and Enrico Magli. Deep graph-convolutional image denoising. *IEEE Transactions on Image Processing*, 29:8226–8237, 2020. 2
- [47] Joost Van De Weijer, Theo Gevers, and Arjan Gijsenij. Edge-based color constancy. *IEEE Transactions on image processing*, 16(9):2207–2214, 2007. 1
- [48] Zhou Wang and Alan C Bovik. A universal image quality index. *IEEE signal processing letters*, 9(3):81–84, 2002. 7
- [49] Zhou Wang, Alan C Bovik, Hamid R Sheikh, and Eero P Simoncelli. Image quality assessment: from error visibility to structural similarity. *IEEE transactions on image processing*, 13(4):600–612, 2004. 7
- [50] Jun Xu, Yuan Huang, Ming-Ming Cheng, Li Liu, Fan Zhu, Zhou Xu, and Ling Shao. Noisy-as-clean: Learning self-supervised denoising from corrupted image. *IEEE Transactions on Image Processing*, 29:9316–9329, 2020. 3
- [51] Xiangyu Xu, Muchen Li, and Wenxiu Sun. Learning deformable kernels for image and video denoising. *arXiv preprint arXiv:1904.06903*, 2019. 2
- [52] Bo Yang, Xiao Fu, Nicholas D Sidiropoulos, and Mingyi Hong. Towards k-means-friendly spaces: Simultaneous deep learning and clustering. In *international conference on machine learning*, pages 3861–3870. PMLR, 2017. 3
- [53] Yunhao Yang, Yuhan Zheng, Yi Wang, and Chandrajit Bajaj. Learning deep latent subspaces for image denoising, 2021. 2, 7, 9
- [54] Syed Waqas Zamir, Aditya Arora, Salman Khan, Munawar Hayat, Fahad Shahbaz Khan, Ming-Hsuan Yang, and Ling Shao. CycleISP: Real image restoration via improved data synthesis. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 2696–2705, 2020. 3
- [55] Kai Zhang, Wangmeng Zuo, Yunjin Chen, Deyu Meng, and Lei Zhang. Beyond a gaussian denoiser: Residual learning of deep cnn for image denoising. *IEEE Transactions on Image Processing*, 26(7):3142–3155, 2017. 3
- [56] Yulun Zhang, Yapeng Tian, Yu Kong, Bineng Zhong, and Yun Fu. Residual dense network for image super-resolution, 2018. 9

## A Detailed Setup of Our RSE-RL Framework

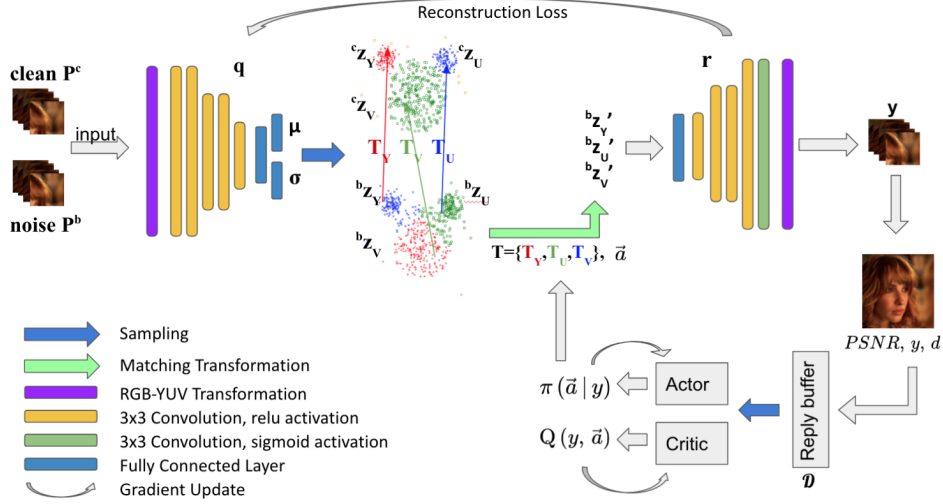


Figure A1: For each given observed image, we split the image into local patches and feed every patch as a stack into the encoding network. For both clean  $P^c$  and noisy patches  $P^b$ , the encoder transforms the channels of patches from RGB to YUV. The latent space is divided into three subspaces, the encoder projects the YUV features of the patches onto three latent subspaces  $Z_y$ ,  $Z_u$ , and  $Z_v$ . Both the clean patches and noisy patches are projected onto the three spaces. A set of transformations  $T = \{T_y, T_u, T_v\}$  are learned to match the latent representation of the noisy patches to a corresponding representation of the clean patches in all three subspaces. For instance,  $T_y(z_y^b) = z_y^c$  for a noisy patch  $z_y^b$  and a clean patch  $z_y^c$  in subspace  $Z_y$ . The transformed noisy representations are sent to the decoder for image reconstructions. The decoder reconstructs YUV channels from the latent spaces representations and transform the channels from YUV back to RGB, hence we get the denoised images.

**Architecture of Patch Transformation - Correspondence Network** In our network architecture, the image patches are transformed from RGB to YUV channels prior to the encoding procedure. The RGB-YUV transformation is defined as

$$\begin{bmatrix} Y \\ U \\ V \end{bmatrix} = \begin{bmatrix} 0.299 & 0.587 & 0.114 \\ -0.147 & -0.289 & 0.436 \\ 0.615 & -0.515 & -1.000 \end{bmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix}$$

A set of encoders  $q_y$ ,  $q_u$ , and  $q_v$  encode the YUV channels respectively and project the patch information on three latent subspaces  $Z_y$ ,  $Z_u$ , and  $Z_v$ . The dimension of one subspaces is set to 72 for both sets of experiments, hence the latent space dimension is 216. In each of the latent subspaces, both clean and noisy patch representations are projected and we want to learn a transformation that matches noisy patches to clean patches representations. The transformations  $T_y, T_u, T_v$  are defined and operated in their corresponding latent subspaces. Each of the transformation  $T_s$  ( $s \in \{y, u, v\}$ ) is a three-layer MLP, with identical dimension layers and ReLU activation. Each transformation is trained to match from a noisy patch representation  $z_s^b$  to a clean patch representation  $z_s^c$  within its latent subspaces, the loss function is defined in Equation (7).

**Our reinforcement learning setting** In the reinforcement learning settings, we start with our pre-trained VAE and transformations and try to enhance the transformations recursively. The RL model starts with the trainable weights in our pre-trained models and learns to minor adjust the trainable weights in the transformations for obtaining higher PSNR. The reason for starting with our pre-trained models is that random initialization on the trainable parameters leads to extreme difficulty on convergence as well as getting better performance.

In the soft actor-critic implementation, we use the default parameter settings for the Q-function (9), where the entropy regularization coefficient  $\alpha = 0.2$  and the discount factor  $\gamma = 0.99$ . The reward  $r$  for both sets of experiments are defined as below:

$$r = 1.25 \times (PSNR_t - PSNR) + 5$$

where the specifications are identical to Equation (13). The target PSNR is set to  $PSNR_t = 30.0$  for the experiments on CelebA dataset and  $PSNR_t = 34.0$  for SIDD dataset.

The action space  $a = \{a_y, a_u, a_v\}$  is a set of weight vectors whose dimension is identical to the hidden dimension of  $T_s$ . The action function is defined in Equation (8), which the dot products of the weights in  $T_s$  and its corresponding weight vector  $a_s, s \in \{y, u, v\}$  are assigned as the new weights in  $T_s$ . More specifically, we have

$$T_{s_j} = a_s \cdot T'_{s_j} \text{ for } s \in \{y, u, v\}, j = \{1, 2, 3\}$$

since  $T_s$  is a three-layer MLP. The state space is a set of trainable weights in the three transformations. The states are denoted  $T_{s_j}$  in the equation above.

## B Additional Experimental Results of Our RSE-RL framework

**Decomposed Subspace Visualization** The following sets of figures show the denoising results on each of the Y, U, and V spaces, which further demonstrate how the noises are removed on each space. The figures A2 and A3 present the examples on both our synthesized CelebA dataset and SIDD dataset. Furthermore, we show some patch-wise matching in Figure A4. The figure gives several specific patches as examples for demonstrating how the noisy patches map to the clean patches. It also shows the noises on the patches specifically, and we can observe the noises on Y, U, and V spaces.

**Iterative Image Enhancement Improvement Framework using our RSE-RL network** We also present how the images are recursively self-enhanced in the reinforcement learning framework in Figure A5. The resulting images over CelebA dataset has shown us a improvement using RL backbone. Without a RL training, our VAE performance yield at a local minimum while the weight updating under feedback control gives us a closer to optimal result. We would also justify that, based on our observation, random initialization of the entire scheme would yield a much more slower convergence rate and an extreme low-PSNR local minimum, and thereby the pretrained network parameter initialization is crucial for generating high-quality enhanced images.

**Justification of Removing Block Artifacts** There might be the case where our filter generates patch-based enhancement result locally while ignoring the neighboring patches. The one-to-one correspondence from noisy patches to clean patches might cause additional block artifacts, as stated earlier in Section 3. We propose the post processing using overlapping patch smoothing, or additional deblocking algorithm to correct the newly introduced artifacts. Below we show an ablation study under the influence of overlapping patch selections and the use of deblocking artifacts.

In general testing, we compare the qualities between the non-overlapping patches and overlapping patches, as well as the qualities before and after using the deblocking method (28). The average PSNR for images composed by non-overlapping patches is 27.8214. And we can observe obvious blocking artifacts on the edge of the patches (in Figure A6). When we apply the deblocking method, the average PSNR is slightly reduced to 27.8212 and the block artifacts can still be visualized.

By comparison, after we apply the overlapping patches, there is a smooth transition on each of the edge between two blocks. The average PSNR for images composed by overlapping patches is 28.84, which is a significant enhancement. We can also observe the enhancement in the figure A6. However, we applied the deblocking method to the images with overlapping patches and there is no observable improvement, which the average PSNR stays the same.



Figure A2: CelebA Denoising Visualization in YUV Spaces: images on the top row are images that contain the synthesized artifacts (Gaussian Noise). The images on the second row are the denoising result from our RSE-RL. And the images on the bottom row show the difference between the first two rows, which are the expected noises we removed by the network. The images are scaled to  $[0, 255]$  for all channels. Columns from left to right show the images on RGB channels, Y space, U space, and V space, respectively. Our method reveals and remove the noise decomposed in three channels respectively.

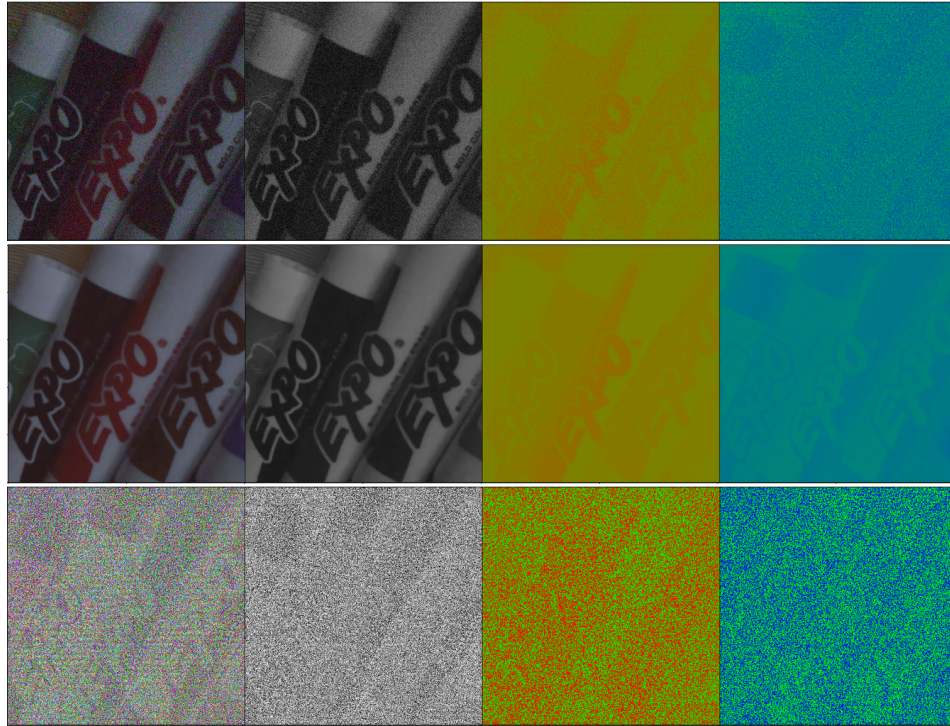


Figure A3: SIDD Denoising Visualization in YUV Spaces: specifications are identical to Figure A2. The figure demonstrates a noise removal over channels and show our patch based method can apply to large-scale, realistic image as well.

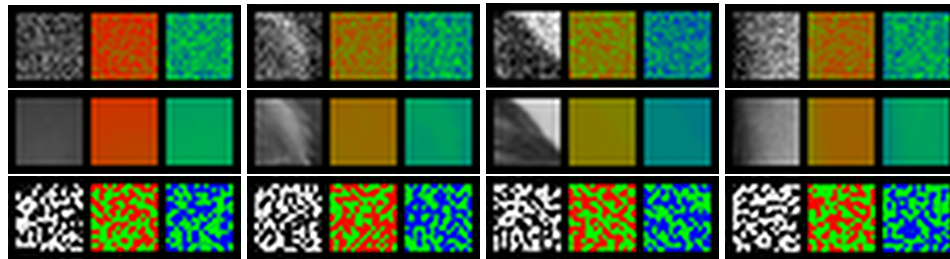


Figure A4: Patch based matching results on YUV spaces: the first row are the noisy patches, the second row are the clean patches that match to the noisy patches, the third are the contrast, representing the noise we are removing. Columns from left to right- three columns as a group- are the images on YUV spaces respectively. For presentation, the patches are scaled to  $[0, 255]$  for all the channels. This justifies the correctness of our patch transferring scheme within each patch locally. The detailed is preserved, too. Hence, our method is totally amenable to any size of images.





Figure A5: Recursive Self-enhancing RL Visualization: the figure shows how the test images are recursively enhanced in a 50-epoch RL training. We observe a performance boost when iterating the weights of decomposed transformations under three latent subspaces. With the reinforcement learning agent, the network converges to a better result compared to the case where only a solo VAE framework can achieve. The last line also specify the difference between the starting image we fed into RL agent and the final result after recursive learning.

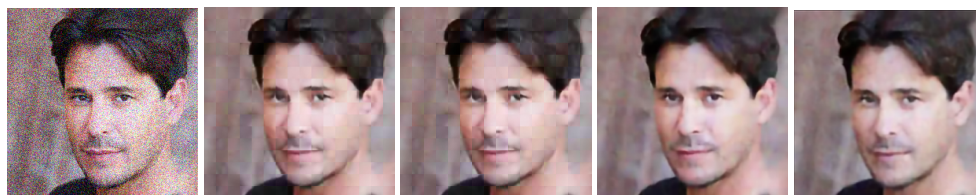


Figure A6: Deblocking Results: this figure shows the results of a deblocking method(28), as well as our overlapping patch smoothing alternative. It shows that our overlapping patch smoothing method can remove the block artifacts that may be created by our patch-based scheme. The columns from left to right show the noisy image, image composed by patches without overlapping, non-overlapping patches with deblocking enhancement, image with overlapping patch smoothing, and image with overlapping patch smoothing + deblocking enhancement. The PSNR for these images from left to right are 19.89, 29.51, 29.50, 30.31, and 30.31.