

# Combined Internal and External Category-Specific Image Denoising

Saeed Anwar

[Saeed.Anwar@anu.edu.au](mailto:Saeed.Anwar@anu.edu.au)

Cong Phuoc Huynh

[Cong.Huynh@anu.edu.au](mailto:Cong.Huynh@anu.edu.au)

Fatih Porikli

[Fatih.Porikli@anu.edu.au](mailto:Fatih.Porikli@anu.edu.au)

The Australian National University,  
Canberra, Australia

## Abstract

In this paper, we present a category-specific image denoising algorithm that exploits patch similarity within the input image and between the input image and an external dataset. We rely on standard internal denoising for smooth regions while consulting external images in the same category as the input to denoise textured regions. The external denoising component estimates the latent patches using the statistics, *i.e.* means and covariance matrices, of external patches, subject to a low-rank constraint. In the final stage, we aggregate results of internal and external denoising using a weighting rule based on the patch SNR measure. Our experimental results on five datasets confirms that the proposed algorithm produces superior results compared with state-of-the-art denoising methods both qualitatively and quantitatively.

## 1 Introduction

Image denoising is a prevalent, well known, yet ill-posed problem in low-level vision, where the aim is to recover the clean image from its noisy version. Since the problem is under-constrained due to missing information, regularization assumptions on the noise model are taken into account including the noise is additive white Gaussian and stationary, and there is no correlation between contaminated pixels. Furthermore, the variance of the noise is usually assumed to be known.

During the last decade, many patch based algorithms [1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13] have been developed to improve the performance of noise removal. Nevertheless, their performance is often a marginal improvement to the BM3D method [10], which is still considered a widely accepted baseline even after a decade. According to Chatterjee *et al.* [1] and Levin *et al.* [13], BM3D achieves near optimal performance, close to theoretical limits on natural images. However, there is still a possibility in performance improvement of denoising using external images [1, 13, 14].

Internal image denoising with a single image is popular and usually has a low computational load. Earlier techniques focused on recovering noisy pixels from their neighboring noisy pixels *e.g.* Gaussian filtering, bilateral filtering, and total variation. Later algorithms

focused on re-occurrence of patches [13] in the noisy image to reconstruct the noise-free image, examples are non-local means [14], BM3D [15], WNNM [16], SAIST [17], SAPCA [18], and TSID [19]. These algorithms are effective for areas with repetitive texture, however, on the downside, they suffer when they attempt to find corresponding matches for infrequent patches *i.e.* the patches that are rarely present in the image. To overcome this issue, some methods [20, 21] proposed alternatives albeit with limited applicability. Moreover, when the noise is strong, internal denoising performance degrades drastically as it struggles to find correct reference patches.

To complement internal image information, other works[22, 23] resorted to external targeted datasets for image denoising. This strategy improves the denoising in specific situations but requires correlated image datasets, and thus fails when the dataset variation becomes high for the same object. Another problem with these algorithms is they involve an exhaustive search policy, which makes them computationally expensive. Alternatively, several learning methods, such as EPLL [24], PCLR [25], PGPD [26], KSVD [27], and MLP [28], were proposed to derive priors from natural noise-free images or the input image itself. However, these learned priors are generic for natural images and are not specific to any image category.

Since internal denoising and external denoising both have their own strengths, attempts have been made to incorporate both [29, 30, 31]. Mosseri *et al.* [32] modified the internal denoising to exploit external natural image patches for textured regions, however, it has the same shortcoming as of generic priors. Yue *et al.* [33] combined internal and external BM3D for denoising. Although it has shown promising results in scenarios where same external images having different scales and orientations are used, it fails to demonstrate such performance when the images are different, even if they belong to the same category.

In this paper, we propose a novel image denoising method that consolidates the strengths of internal and external denoising techniques in a complementary manner. We achieve this by exploiting category-specific information using high frequencies from external patches and low frequencies from the noisy image. For computational efficiency, we design a strategy to select patches of interest from the external images without excessive computational burden.

Our method is based on the intuition that, although internal denoising methods are capable of locating similar patches within smooth regions of the input image, it might be more effective to look up external clean images to denoise textured regions. This is because, when there is no noise, it is straightforward to gather a high number of similar patches. However, there would be a limited number of patches depicting a similar spatial pattern to a highly textured patch.

## 2 Denoising framework

Suppose that we are given a noisy image  $\mathbf{Y}$ , which is related to the original image  $\mathbf{X}$  as

$$\mathbf{Y} = \mathbf{X} + \mathbf{N}, \quad (1)$$

where  $\mathbf{N}$  is the noise component with a variance  $\sigma_n^2$ . Our aim is to recover  $\mathbf{X}$ , using the additional input from a dataset of images  $\{\mathbf{Z}_i : i = 1, \dots, K\}$ .

The process flow of our algorithm is illustrated in Fig. 1. It comprises four stages: candidate patch search, internal denoising, external denoising, and aggregation of the two denoising results.

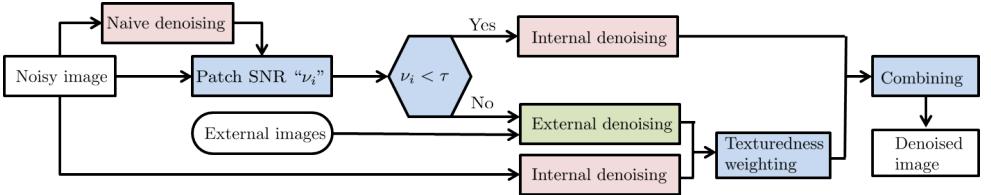


Figure 1: Framework of our denoising algorithm. For high noise levels, *i.e.*  $\sigma_n > 40$ , the whole procedure is iterated twice while it is applied once for low noise levels, *i.e.*  $\sigma_n < 40$ .

## 2.1 External patch search

In this section, we describe the process of selecting external patches that are similar to highly textured patches in the input image. Such patches are identified using a smoothness measure  $\nu_i$  as described in Section 2.3. The patch selection process occurs in two stages. Firstly, candidate external images are selected from a category-specific training dataset. In the second stage, we locate external patches that are similar to a given noisy input patch.

We specifically designed our algorithm to retrieve images from the external database. We select a preset number of external images that are structurally most similar to the input noisy image based on the structural similarity index (SSIM). This helps pruning the number of searches incurred during the denoising process.

Next, for each noisy patch  $y$ , we search for the most similar external patches. A number of algorithms such as [28, 36, 38] search for reference patches in external databases using KD-tree and graph-cut. However, these search methods are not sufficiently efficient for large databases. We opt for the PatchMatch algorithm [1] and modify it to retrieve the reference patches efficiently from external datasets in the two following steps.

**Initialization:** For faster convergence, we initialise the nearest neighbor fields with zero displacement and the scale equal to the current scale of the patch, this is based on the idea that good similar reference patches can be found in the neighborhood of the noisy patch.

**Propagation:** PatchMatch [1] provides the option to select multiple reference patches from a single external image for a noisy patch. However, this option is not feasible for our algorithm as searching for multiple reference patches is time consuming. Therefore, we present an alternative to increase the number of patches without searching multiple times through the image for a single noisy patch. Our idea is inspired by the works in [2, 10], which suggests that good reference patches can be found in the neighborhood. Hence, we extract eight neighbor patches which are one pixel apart in every direction from the current selected reference patch. This approach is more computationally and time efficient than the original patch match method, while providing sufficient external information for denoising. As a result, for each noisy patch  $y_i$ , we obtain top external patches from each of the closest external candidate images. Finally, we only retain  $T_i$  patches in each pool whose Euclidean distance from the input patch  $y_i$  is no more than a threshold  $\delta$ .

## 2.2 External denoising

Next, we detail the external denoising component with input from external images. To formulate the problem, we choose to express patches using the Discrete Cosine Transform (DCT) domain. The DCT space can be represented as a linear transformation with an orthonormal basis  $\Phi \in \mathbb{R}^{D \times N}$ , *i.e.*  $\Phi^T \Phi = \mathbb{I}$ . In our representation, the pixel intensity vector

$\mathbf{y}_i$  in the noisy patch, and that of the latent patch  $\mathbf{x}_i$ , are related to their respective DCT coefficients  $\beta_i$  and  $\alpha_i$  as  $\beta_i = \Phi\mathbf{y}_i$  and  $\alpha_i = \Phi\mathbf{x}_i$ . In addition, for each patch  $\mathbf{y}_i$  from the input image, we have collected a set of  $T_i$  external patches  $\{\mathbf{z}_{i,j} : j = 1, \dots, T_i\}$  similar to it. Let us denote  $\{\gamma_{i,j} : j = 1, \dots, T_i\}$  as the transform coefficients of  $\mathbf{z}_{i,j}$  and the mean and covariance matrix estimated from these transform coefficient vectors as  $\mu_i$  and  $\Sigma_i$ , respectively.

## 2.2.1 Objective function

**Fidelity Term:** First, we define the conditional likelihood of the input noisy image and the original noise-free image, which is given by

$$p(\mathbf{Y}|\mathbf{X}) \propto \exp\left(-\frac{\|\mathbf{Y} - \mathbf{X}\|_2^2}{\sigma_n^2}\right), \quad (2)$$

In Equation 2, the aim is to maximise the conditional log likelihood to recover the noise-free image  $\mathbf{Y}$ , which is same as minimizing the data fidelity term  $\|\mathbf{Y} - \mathbf{X}\|_2^2$ . Since each pixel value in the image is shared by approximately the same number of overlapping patches, the data fidelity term can be approximated as  $\|\mathbf{Y} - \mathbf{X}\|_2^2 \approx \frac{1}{N} \sum_{i=1}^N \|\mathbf{y}_i - \mathbf{x}_i\|_2^2$ , where  $N$  is the number of patches in the image. Moreover, due to the orthonormality property of the basis function, we obtain  $\|\alpha_i - \beta_i\|_2^2 = \|\Phi(\mathbf{y}_i - \mathbf{x}_i)\|_2^2 = \|\mathbf{y}_i - \mathbf{x}_i\|_2^2$ , and after simplification we get

$$\|\mathbf{Y} - \mathbf{X}\|_2^2 \approx \frac{1}{N} \sum_{i=1}^N \|\alpha_i - \beta_i\|_2^2. \quad (3)$$

**Patch group membership:** Here, we present another constraint which imposes the similarity between noisy patch and the set of reference patches. In order to recover the patch  $\mathbf{x}_i$  from the noisy patch  $\mathbf{y}_i$  in transform domain, we compute and rely on the statistics such as mean ( $\mu_i$ ) and covariance matrix ( $\Sigma_i$ ) of the reference patch group. Moreover, We consider that similar patches in the transform domain is part of Gaussian distribution and the most probable  $\mathbf{x}_i$  is one that maximises its likelihood of belonging to the reference patch group, i.e.  $p(\alpha_i|\mu_i, \Sigma_i) \propto \exp\left(-\frac{1}{2}(\alpha_i - \mu_i)^T \Sigma_i^{-1} (\alpha_i - \mu_i)\right)$ . This is similar to minimizing the log-likelihood in the transform domain

$$\log p(\alpha_i|\mu_i, \Sigma_i) \propto \frac{1}{2}(\alpha_i - \mu_i)^T \Sigma_i^{-1} (\alpha_i - \mu_i), \quad (4)$$

**Low-rank constraint:** We further formulate a low-rank constraint and the intuition behind this constraint is that the local structure of a patch can be sparsely represented by a basis with a low cardinality. Therefore, when similar patch vectors are stacked as columns of a matrix, the matrix should exhibit the low rank property. We formulate our problem similar to Candès and Recht [6]. We form a data matrix  $\mathbf{M}_i$  which contains the transform coefficients of each latent patch  $\mathbf{x}_i$  and its reference patches as its columns i.e.  $\mathbf{M}_i = [\alpha_i, \gamma_{i,1}, \dots, \gamma_{i,T_i}]$ . Our aim is to minimise the matrix nuclear norm  $\|\mathbf{M}_i\|_*$ , which is the sum of its singular values.

## 2.2.2 Optimisation

In this section, we present the optimization for external patch denoising. To obtain the objective function, we aggregate the terms in Equations 3, 4 and the nuclear norm constraint for  $i$ -th noisy patch. In addition, we introduce auxiliary term to substitute for the equality constraint  $\mathbf{M}_i = [\alpha_i, \gamma_{i,1}, \dots, \gamma_{i,T_i}]$ . The resulting objective function is formulated as

$$\begin{aligned} (\alpha_i^*, \mathbf{M}_i^*) = \operatorname{argmin}_{\alpha_i, \mathbf{M}_i} & \frac{1}{\sigma_n^2} \|\alpha_i - \beta_i\|_2^2 + \lambda_1 (\alpha_i - \mu_i)^T \Sigma_i^{-1} (\alpha_i - \mu_i) \\ & + \lambda_2 \|\mathbf{M}_i\|_* + \frac{\|\mathbf{M}_i - [\alpha_i, \gamma_{i,1}, \dots, \gamma_{i,T_i}]\|_F^2}{(T_i + 1) \sigma_n^2}. \end{aligned} \quad (5)$$

Here, the normalisation factor  $\frac{1}{(T_i+1)\sigma_n^2}$  takes into account the image noise and the number of reference patches. We apply coordinate descent to minimise the cost function in Equation 5, with respect to either  $\alpha_i$  or  $\mathbf{M}_i$ , while keeping the other variable fixed.

**Compute  $\alpha_i$  with fixed  $\mathbf{M}_i$ :** First we fix the value of  $\mathbf{M}_i^*$  and compute for  $\alpha_i$  by solving the following sub-problem

$$\alpha_i^* = \operatorname{argmin}_{\alpha_i} \frac{\|\alpha_i - \beta_i\|_2^2}{\sigma_n^2} + \frac{\|\alpha_i - \mathbf{M}_i^*(:, 1)\|_2^2}{(T_i + 1)\sigma_n^2} + \lambda_1(\alpha_i - \mu_i)^T \Sigma_i^{-1}(\alpha_i - \mu_i), \quad (6)$$

where  $\mathbf{M}_i^*(:, 1)$  denotes the first column of the matrix  $\mathbf{M}_i^*$ . taking its derivative w.r.t.  $\alpha_i$  leads to,

$$\alpha_i^* = \frac{\beta_i + \lambda_1 \sigma_n^2 \Sigma_i^{-1} \mu_i + \frac{\mathbf{M}_i^*(:, 1)}{T_i + 1}}{\left( \frac{T_i + 2}{T_i + 1} \mathbb{I} + \lambda_1 \sigma_n^2 \Sigma_i^{-1} \right)} \quad (7)$$

**Compute  $\mathbf{M}_i$  with fixed  $\alpha_i$ :** Now with  $\alpha_i^*$  values in hand, we form the data matrix  $\hat{\mathbf{M}}_i \triangleq [\alpha_i^*, \gamma_{i,1}, \dots, \gamma_{i,T_i}]$  for every patch. The sub-problem to be solved with respect to  $\mathbf{M}_i$  is then stated as

$$\mathbf{M}_i^* = \operatorname{argmin}_{\mathbf{M}_i} \|\mathbf{M}_i - \hat{\mathbf{M}}_i\|_F^2 + \zeta \|\mathbf{M}_i\|_*, \quad (8)$$

where  $\zeta = \lambda_2(T_i + 1)\sigma_n^2$ .

Let us consider that we have  $U\Lambda V^T$  as the singular value decomposition of  $\hat{\mathbf{M}}_i$ , with  $\Lambda_k$  being the  $k$ -th singular value. Cai *et al.* [5] derived the optimal solution to Equation 8, by soft-thresholding the singular values as

$$\mathbf{M}_i^* = U \mathcal{S}_\zeta(\Lambda) V^T, \quad (9)$$

where  $\mathcal{S}_\zeta(\Lambda) = \operatorname{diag}(\{(\Lambda_k - \zeta)_+\})$  is the soft-thresholding operator with  $(x)_+ = \max(x, 0)$ .

The final step of the external denoising component is to recover the full image from all the patches. Once we have the transform coefficients  $\alpha_i$ , we compute the patch intensity by an inverse transform as  $\mathbf{x}_i^{ext} = \Phi^T \alpha_i, \forall i = 1, \dots, M$ , where  $\Phi$  is the DCT basis.

## 2.3 Combined internal and external denoising

Now we combine the results of the internal and external denoising component. Here, we opt for BM3D [10] as the internal denoising component due to its efficiency and effectiveness for various benchmark datasets. To perform both internal and external denoising, we extract overlapping patches  $\mathbf{y}_i, i = 1, \dots, M$  of a fixed size  $m \times n$  from the input at a regular stride. Let  $\mathbf{x}_i^{int}$  denote the patch centred at pixel  $i$ , which has been denoised by BM3D. To obtain the final denoised image, we combine the internal denoising result  $\mathbf{x}_i^{int}$  with the external one  $\mathbf{x}_i^{ext}$  based on the level of smoothness  $v_i$ .

The noise variance can be estimated using an approximated filtered image  $\tilde{\mathbf{y}}$ , which is obtained by an existing denoising algorithm such as BM3D [10]. Let  $\operatorname{var}(\tilde{\mathbf{y}}_i)$  denote the signal variance of the pre-filter patch  $\tilde{\mathbf{y}}_i$  while  $\operatorname{var}(n_i)$  denotes its noise variance. The patch noise variance is then yielded as  $\operatorname{var}(n_i) \approx \operatorname{var}(\mathbf{y}_i - \tilde{\mathbf{y}}_i)$ .

To determine whether to apply internal or external denoising to a patch  $\mathbf{y}_i$ , we evaluate its level of smoothness as defined by the patch SNR metric presented in [28]  $v_i = \sqrt{\frac{\operatorname{var}(\tilde{\mathbf{y}}_i)}{\operatorname{var}(n_i)}}$ . Next, we qualify whether a patch is smooth or highly textured with a threshold  $\tau = \max(0.15, 0.5 - 0.5\sqrt{\frac{\sigma_n}{255}})$ . Specifically, if  $v_i \leq \tau$ , we consider the patch smooth and apply only internal denoising to it, *i.e.*  $\mathbf{x}_i = \mathbf{x}_i^{int}$ . Otherwise, we weigh  $\mathbf{x}_i^{ext}$  and  $\mathbf{x}_i^{int}$  according to  $\mathbf{x}_i = v_i \mathbf{x}_i^{ext} + (1 - v_i) \mathbf{x}_i^{int}$  to complement the low frequency details for the external denoised patches.

| Internal Only | External Only | Combined |
|---------------|---------------|----------|
| 24.94         | 25.58         | 26.23    |

Table 1: The denoising performance of individual components for  $\sigma_n = 70$  on the Gore dataset. Combining both internal and external components improves the PNSR metric.

To recover the full image, we translate the patches to their original locations and average the values of overlapping patches at shared pixels. Let  $\mathbf{Q}_i$  denote the patch extraction matrix at the  $i$ -th pixel of an image, *i.e.*  $\mathbf{x}_i = \mathbf{Q}_i \mathbf{X}$ . With the known matrices  $\mathbf{Q}_i$ 's, the latent image is the optimal solution to the problem

$$\mathbf{X}^* = \underset{\mathbf{X}}{\operatorname{argmin}} \lambda_0 \|\mathbf{X} - \mathbf{Y}\|_2^2 + \sum_{i=1}^M \|\mathbf{Q}_i \mathbf{X} - \mathbf{x}_i\|_2^2. \quad (10)$$

where  $\lambda_0$  is a positive constant. The least-squares solution to the above equation is

$$\mathbf{X}^* = \left( \lambda_0 \mathbb{I} + \sum_{i=1}^M \mathbf{Q}_i^T \mathbf{Q}_i \right)^{-1} \left( \lambda_0 \mathbf{Y} + \sum_{i=1}^M \mathbf{Q}_i^T \mathbf{x}_i \right) \quad (11)$$

## 3 Experiments

We first analyse different aspects of our proposed method and then compare the whole method with state-of-the-art algorithms, including BM3D [10], WNNM [19], SAIST [3], EPLL [39], PCLR [15], PGPD [31], TID [22] and CID [22]. Furthermore, we modify internal denoising methods, including NLM [9], SAPCA [10], TSID [32] and BM3D [10] to take extra input from external class-specific datasets and term denote as eNLM, eSAPCA, eTSID, and eBM3D, respectively.

We generate noisy input images from the original noise-free data using Gaussian noise with standard deviations of  $\sigma_n = 20, 25, 30, 40, 50, 70, 80, 100$ . We use PSNR as the metric for quantitative evaluation. In the following experiments, we set the parameters of our algorithm as  $k = 8$ ,  $\mathbf{Z}_i = 16$ ,  $\lambda_0 = 1$ ,  $\lambda_1 = 0.5$ , and  $\lambda_2 = 10$ .

**Datasets:** We experimented on five datasets to demonstrate performance of our algorithm. These datasets include face [4], text [12], bike [16] and buildings [33]. We also present ample qualitative results on the mentioned datasets as well as on another category Multi-view [20] in supplementary materials. In each dataset under study, we arbitrarily select between 10 and 20 images, from which noisy images are generated as input to denoising algorithms. For each noisy image, we take five other clean images from the same dataset as external input to our algorithm.

### 3.1 Ablation study

**Separate and combined internal & external denoising:** Here, we present the effect of individual denoising components and their combined effect. Table 1 present average PSNR values on 14 test images of Gore dataset for  $\sigma_n = 70$ . A higher PSNR value is obtained when both components are combined, confirming our motivation behind the proposed scheme.

**Number of patches:** Table 2 demonstrates the denoising performance of our algorithm with respect to the number of patches. We can observe that the denoising performance degrades by a little margin as the number of patches increases. This decrease may be due to incorporation of dissimilar patches for denoising.

**Number of external images:** In our algorithm, external images are important and here we show the impact of different number of external images on outcome of our algorithm. The last row of the table 2 shows the denoising results with respect to the increase in number of

| Effect of | No of Patches/Images |       |       |       |       |
|-----------|----------------------|-------|-------|-------|-------|
|           | 8                    | 16    | 24    | 32    | 40    |
| Patches   | 27.53                | 27.49 | 27.45 | 27.43 | 27.41 |
| Images    | 27.53                | 27.54 | 27.57 | 27.57 | 27.59 |

Table 2: Variation of PSNR for the Gore dataset and  $\sigma_n = 50$  with respect to the number of external patches and the number of external images.

|      |         | TRAIN        |              |              |              |              |         |
|------|---------|--------------|--------------|--------------|--------------|--------------|---------|
|      |         | Datasets     | Face         | Views        | Text         | Bike         | Natural |
| TEST | Face    | <b>27.66</b> | 25.55        | 18.45        | 25.83        | 25.78        |         |
|      | Views   | 25.79        | <b>29.02</b> | 22.91        | 26.56        | 26.57        |         |
|      | Text    | 15.99        | 18.44        | <b>24.87</b> | 20.11        | 19.44        |         |
|      | Bike    | 20.97        | 22.70        | 21.77        | <b>25.15</b> | 22.81        |         |
|      | Natural | 27.27        | 27.75        | 27.69        | 27.95        | <b>28.74</b> |         |

Table 3: The performance of our algorithm for  $\sigma_n = 50$ , with the external dataset from different categories. The PSNR is maximum along the diagonal, when the external category matches the test category.

external images. It is obvious that performance improves as the number of images increases, however, the improvement becomes stagnated after specific increase in number.

**Choice of external image category:** We demonstrate the impact of the correct external image category on the denoising performance. In table 3, the PSNR value degrades when denoising is assisted by a dissimilar category while its results reaches maximum when same category is used for denoising. This experiment further confirms the importance of category specific information for denoising.

**Running time:** We implemented our algorithm in MATLAB. For an image of  $455 \times 280$ , it takes about 10s to retrieve structurally similar external images from a category specific dataset of 810 images. The processing time for patch search, internal denoising and external denoising is about 0.6s, 6s and 106s. In short, the total processing time for our method is about 122.6s. On other hand, the denoising methods such as BM3D, PCLR, WNNM and TID takes 0.6s, 74s, 129s and 265s, respectively.

### 3.2 Face images

Face images are an important part of many computer vision tasks such as human surveillance, tracking and recognition. Face images are usually corrupted when captured in low light conditions, hence, denoising maybe required as an initial step for the mentioned tasks. Here, we show the capability of our algorithm to denoise human face images. To this end, we randomly select 14 images from the Gore [29] dataset as the test set while using others as the training set. The comparison between our method and others is shown in Figure 5(a). As observed, our algorithm outperforms others on all noise levels.

Figure 2 demonstrate the qualitative results on a sample image from Gore [29] dataset and illustrates that our algorithm restores aesthetically pleasing texture. In our results the image details are more similar to the ground-truth and numerically higher than others. Also, the denoising results of the modified algorithms perform comparable to their original implementations. This experiment illustrates that in the presence of a good dataset only, superior performance cannot be achieved, rather we need a laborious and carefully designed algorithm to exploit the external datasets.

### 3.3 Text images

We evaluate our algorithm for text denoising on the same dataset as reported in [27]. In Figure 5(b), we show the PSNR curve for each denoising method averaged over a range of noise

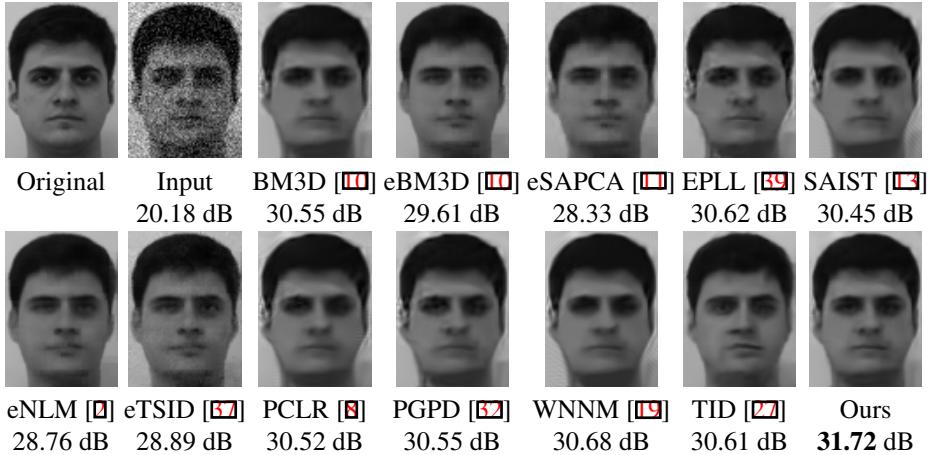


Figure 2: Comparison of denoising methods on face images from the dataset in [20] for  $\sigma_n = 25$ . Our method is able to recover sharp edges with less artifacts than the state-of-the-art.

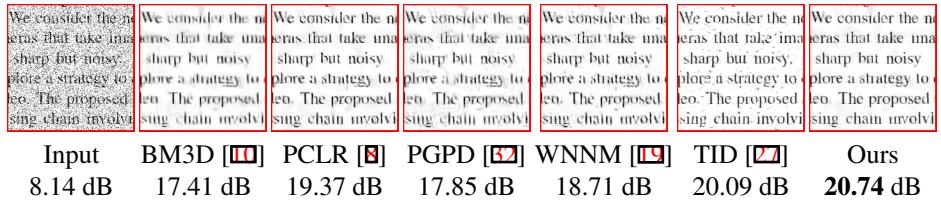


Figure 3: Comparison for  $\sigma_n = 100$  on a text image from the dataset in [20]. Our method is able to recover sharper text edges without artifacts compared with the state-of-the-art.



Figure 4: Denoising results on a bike image [20], for  $\sigma_n = 40$ . Our method is able to recover sharp boundaries around the bike.

levels. Our proposed algorithm performs at least 5 dB better than state-of-the-art BM3D. We also observe that TID [20] performs below average for low noise ( $\sigma_n < 50$ ) as it is unable to handle variation in the images of the text dataset. In addition, Figure 3 demonstrates the qualitative result for a high level of white noise, *i.e.*  $\sigma_n = 100$ . Our algorithm produces sharp edges and crisper text than the competing methods, which in most cases, exhibit artifacts in the form of distorted text.

### 3.4 Bike images

In Figure 5(c), we plot and compare the denoising results on Bike dataset [20]. This experiment also shows that other external denoising methods such as CID [35] and TID [20] fail to exploit the external dataset due to high variation in the images. TID [20] is unable to exploit the external images due to patch variety of the bike image dataset, and hence poor results. Similarly, CID [35] failed to do registration with the external images and produce results similar

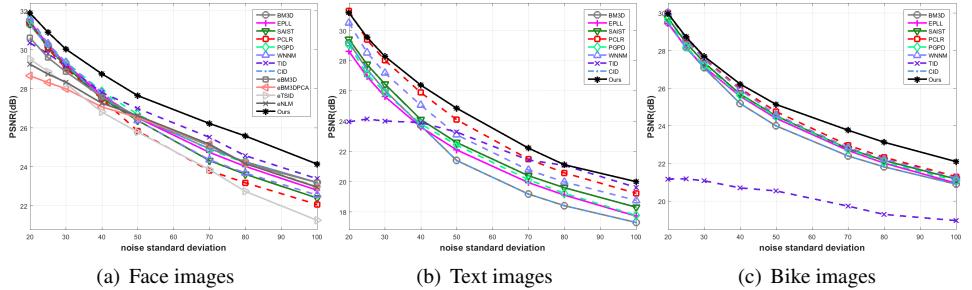


Figure 5: The PSNR of denoised images with respect to the noise standard deviation over: (a) 14 face images with a size of  $65 \times 90$ , (b) text images with a size of  $200 \times 200$ , (c) bike images with a size of  $320 \times 200$ .

| Images | $\sigma_n$ | Methods |       |       |       |       |       |       |       |              |
|--------|------------|---------|-------|-------|-------|-------|-------|-------|-------|--------------|
|        |            | BM3D    | EPLL  | SAIST | PCLR  | PGPD  | WNNM  | TID   | CID   | Ours         |
| d      | 50         | 27.56   | 27.23 | 27.42 | 27.54 | 27.64 | 27.79 | 23.23 | 26.80 | <b>28.18</b> |
|        | 70         | 26.22   | 25.85 | 26.20 | 26.22 | 26.32 | 26.48 | 22.48 | 26.03 | <b>27.01</b> |
|        | 90         | 25.23   | 24.88 | 25.26 | 25.29 | 25.32 | 25.51 | 21.59 | 25.25 | <b>26.06</b> |
| f      | 50         | 28.88   | 28.45 | 28.58 | 28.85 | 28.94 | 29.14 | 21.87 | 28.50 | <b>29.30</b> |
|        | 70         | 27.55   | 27.17 | 27.41 | 27.66 | 27.67 | 27.83 | 21.29 | 27.75 | <b>28.13</b> |
|        | 90         | 26.62   | 26.26 | 26.45 | 26.82 | 26.73 | 26.85 | 20.92 | 26.91 | <b>27.34</b> |

Table 4: Comparisons on the Building dataset [35]. The results for CID are reproduced from [35]. The best results are highlighted in bold.

to BM3D [10]. On the other hand, our scheme outperforms competing algorithms as noise level increases. This suggests that our method is able to retrieve information from external image. Figure 4 illustrates that the proposed scheme recovers more details on the tyres while SAIST and PGPD oversmooth them and BM3D, EPLL and TID generate artifacts.

### 3.5 Building images

Our algorithm also produce superior results on the building images introduced with the Combined Image Denoising (CID) method [25]. For this purpose, we select the images “d” and “f” and simulate their noisy versions at  $\sigma_n = 50, 70$  and  $90$  as reported in [25]. For each input image, we search for patches in five related images as its external dataset. Table 4 shows that our proposed scheme outperforms various state-of-the-art methods on these images. Specifically, our algorithm outperforms CID [25] by more than 1 dB. This experiment confirms that we are able to achieve better PSNR by incorporating patches from other related images.

## 4 Conclusion

Single image denoising algorithms are approaching performance limit. To overcome this deficiency, we propose a novel algorithm which apply internal denoising to smooth regions and combines internal and external denoising for textured regions. We also present a patch matching technique to improve the reference patch search performance. Extensive experimental results demonstrated that our algorithm significantly outperform competitive methods qualitatively and numerically for many noise levels.

## References

- [1] Connelly Barnes, Eli Shechtman, Adam Finkelstein, and Dan Goldman. Patchmatch: a randomized correspondence algorithm for structural image editing. *ACM Transactions on Graphics-TOG*, 28(3):24, 2009.
- [2] Antoni Buades, Bartomeu Coll, and Jean-Michel Morel. A non-local algorithm for image denoising. In *CVPR*, pages 60–65, 2005.
- [3] Harold Christopher Burger, Christian J Schuler, and Stefan Harmeling. Image denoising: Can plain neural networks compete with bm3d? In *CVPR*, pages 2392–2399, 2012.
- [4] Harold Christopher Burger, Christian Schuler, and Stefan Harmeling. Learning how to combine internal and external denoising methods. In *German Conference on Pattern Recognition*, pages 121–130. Springer, 2013.
- [5] Jian-Feng Cai, Emmanuel J. Candès, and Zuowei Shen. A Singular Value Thresholding Algorithm for Matrix Completion. *SIAM J. on Optimization*, pages 1956–1982, 2010.
- [6] Emmanuel J. Candès and Benjamin Recht. Exact matrix completion via convex optimization. *Foundations of Computational Mathematics*, pages 717–772, 2009.
- [7] P. Chatterjee and P. Milanfar. Is denoising dead? *Image Processing, IEEE Transactions on*, pages 895–911, 2010.
- [8] Fei Chen, Lei Zhang, and Huimin Yu. External patch prior guided internal clustering for image denoising. 2015.
- [9] Knaus Claude and Zwicker Matthias. Progressive image denoising. *Image Processing, IEEE Transactions on*, pages 3114–3125, 2014.
- [10] Kostadin Dabov, Alessandro Foi, Vladimir Katkovnik, and Karen Egiazarian. Image denoising by sparse 3-D transform-domain collaborative filtering. *Image Processing, IEEE Transactions on*, pages 2080–2095, 2007.
- [11] Kostadin Dabov, Alessandro Foi, Vladimir Katkovnik, and Karen Egiazarian. BM3D image denoising with shape-adaptive principal component analysis. *Signal Processing with Adaptive Sparse Structured Representations*, 2009.
- [12] Charles-Alban Deledalle, Joseph Salmon, Arnak S Dalalyan, and France Champs-sur Marne. Image denoising with patch based PCA: local versus global. In *BMVC*, pages 1–10, 2011.
- [13] Weisheng Dong, Guangming Shi, and Xin Li. Nonlocal image restoration with bilateral variance estimation: a low-rank approach. *Image Processing, IEEE Transactions on*, pages 700–711, 2013.
- [14] Michael Elad and Michal Aharon. Image denoising via sparse and redundant representations over learned dictionaries. *Image Processing, IEEE Transactions on*, pages 3736–3745, 2006.

- [15] L. Zhang F. Chen and H. Yu. External Patch Prior Guided Internal Clustering for Image Denoising. In *ICCV*, pages 1211–1218, 2015.
- [16] Robert Fergus, Pietro Perona, and Andrew Zisserman. Object class recognition by unsupervised scale-invariant learning. In *Computer Vision and Pattern Recognition, 2003. Proceedings. 2003 IEEE Computer Society Conference on*, volume 2, pages II–264. IEEE, 2003.
- [17] Alessandro Foi, Vladimir Katkovnik, and Karen Egiazarian. Pointwise shape-adaptive DCT for high-quality denoising and deblocking of grayscale and color images. *IEEE transactions on image processing*, pages 1395–1411, 2007.
- [18] Daniel Glasner, Shai Bagon, and Michal Irani. Super-resolution from a single image. In *ICCV*, 2009.
- [19] Shuhang Gu, Lei Zhang, Wangmeng Zuo, and Xiangchu Feng. Weighted nuclear norm minimization with application to image denoising. In *CVPR*, pages 2862–2869, 2014.
- [20] Heiko Hirschmüller and Daniel Scharstein. Evaluation of cost functions for stereo matching. In *CVPR*, pages 1–8, 2007.
- [21] Claude Knäus and Matthias Zwicker. Dual-domain image denoising. In *ICIP*, pages 440–444, 2013.
- [22] M Lebrun, Antoni Buades, and Jean-Michel Morel. A nonlocal bayesian image denoising algorithm. *SIAM Journal on Imaging Sciences*, pages 1665–1688, 2013.
- [23] A. Levin and B. Nadler. Natural image denoising: Optimality and inherent bounds. In *CVPR*, pages 2833–2840, 2011.
- [24] Anat Levin, Boaz Nadler, Fredo Durand, and William T. Freeman. Patch complexity, finite pixel correlations and optimal denoising. In *ECCV*, pages 73–86, 2012.
- [25] Yifei Lou, Paolo Favaro, Stefano Soatto, and Andrea Bertozzi. Nonlocal similarity image filtering. In *International Conference on Image Analysis and Processing*, pages 62–71. Springer, 2009.
- [26] Enming Luo, Stanley H Chan, and Truong Q Nguyen. Image denoising by targeted external databases. In *2014 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pages 2450–2454. IEEE, 2014.
- [27] Enming Luo, Stanley H Chan, and Truong Q Nguyen. Adaptive image denoising by targeted databases. *Image Processing, IEEE Transactions on*, pages 2167–2181, 2015.
- [28] Inbar Mosseri, Maria Zontak, and Michal Irani. Combining the power of internal and external denoising. In *ICCP*, pages 1–9, 2013.
- [29] Yigang Peng, Arvind Ganesh, John Wright, Wenli Xu, and Yi Ma. Rasl: Robust alignment by sparse and low-rank decomposition for linearly correlated images. *TPAMI*, pages 2233–2246, 2012.
- [30] J. Portilla, V. Strela, M.J. Wainwright, and E.P. Simoncelli. Image denoising using scale mixtures of gaussians in the wavelet domain. *Image Processing, IEEE Transactions on*, pages 1338–1351, 2003.

- [31] Jun Xu, Lei Zhang, Wangmeng Zuo, David Zhang, and Xiangchu Feng. Patch Group Based Nonlocal Self-Similarity Prior Learning for Image Denoising. In *ICCV*, pages 1211–1218, 2015.
- [32] L Xu, L Zhang, W Zuo, D Zhang, and X Feng. Patch group based nonlocal self-similarity prior learning for image denoising. 2015.
- [33] Ruomei Yan, Ling Shao, Sascha D Cvetkovic, and Jan Klijn. Improved nonlocal means based on pre-classification and invariant block matching. *Journal of display technology*, 8(4):212–218, 2012.
- [34] Guoshen Yu and Guillermo Sapiro. Dct image denoising: a simple and effective image denoising algorithm. *Image Processing On Line*, 2011.
- [35] H. Yue, X. Sun, J. Yang, and F. Wu. Cid: Combined image denoising in spatial and frequency domains using web images. In *CVPR*, pages 2933–2940, June 2014.
- [36] H. Yue, X. Sun, J. Yang, and F. Wu. Image denoising by exploring external and internal correlations. *TIP*, pages 1967–1982, 2015.
- [37] Lei Zhang, Weisheng Dong, David Zhang, and Guangming Shi. Two-stage image denoising by principal component analysis with local pixel grouping. *Pattern Recognition*, pages 1531–1549, 2010.
- [38] M. Zontak and M. Irani. Internal statistics of a single natural image. In *CVPR*, pages 977–984, 2011.
- [39] Daniel Zoran and Yair Weiss. From learning models of natural image patches to whole image restoration. In *ICCV*, pages 479–486, 2011.