IEEE Access
Multidisciplinary ⋮ Rapid Review ⋮ Open Access Journal

Received 27 July 2024, accepted 17 August 2024, date of publication 23 August 2024, date of current version 2 September 2024.

Digital Object Identifier 10.1109/ACCESS.2024.3448451

RESEARCH ARTICLE

Optimized Color Filter Array for Denoising Diffusion Null-Space Model-Based Demosaicing

INDRA IMANUEL[1], HYOSEON YANG[2], AND SUK-HO LEE[1]
[1]Department of Computer Engineering, Dongseo University, Busan 47011, South Korea
[2]Department of Mathematics, Kyung Hee University, Seoul 02447, South Korea

Corresponding author: Suk-Ho Lee (petra@g.dongseo.ac.kr)

This work was supported by the Basic Science Research Program through the National Research Foundation of Korea under Grant NRF-2022R1I1A3065211 and Grant NRF-2022R1F1A1066389.

ABSTRACT Recently, deep learning-based demosaicing methods have shown promising results. However, there has been little research on designing CFAs that are well-suited for specific deep learning-based demosaicing methods. This is because it is challenging to establish a relationship between deep learning-based demosaicing methods and the CFAs they employ. This contrasts with traditional CFA design methods, which targeted fixed, non-deep learning demosaicing methods that did not depend on data learning, making signal processing theory applicable to the design. In this paper, we propose an optimized color filter array(CFA) tailored for the Denoising Diffusion Null-space Model (DDNM) based demosaicing. We begin by demonstrating the application of the DDNM to the demosaicing problem and establish the conditions under which the DDNM can accurately recover the true colors from CFA images containing colored pixels. Based on this analysis, we propose a CFA pattern that significantly improves the likelihood of accurate color reconstruction using the DDNM-based demosaicing method. Then, we outline the training process for obtaining the optimal filter coefficient values for the proposed CFA pattern. Experimental findings demonstrate that the proposed CFA yields favorable results when paired with the DDNM-based demosaicing technique which surpass those achieved by other CFA patterns.

INDEX TERMS Image demosaicing, deep learning, denoising diffusion model, color filter array, range-null space decomposition.

I. INTRODUCTION

Due to limited computing resources and design constraints in digital cameras, especially in smartphones, a single-channel sensor is typically used with a color filter array(CFA) to capture a scene. Demosaicing refers to the process of reconstructing a full color image from the single-channel sensor data, also known as the CFA image, captured by the digital camera. The aim of demosaicing is to reconstruct a high-quality color image that closely resemble the original scene. This aim can be achieved through the design of an appropriate CFA pattern and the application of sophisticated demosaicing algorithms tailored to that pattern. By optimizing both the CFA design and the corresponding demosaicing process, the accuracy and quality of the reconstructed color images can be significantly improved.

With traditional, non-deep learning demosaicing methods, the CFA design was based on analysis rooted in signal processing theory. For example, the authors of [2] pose the problem of color filter array design formally as one of simultaneously maximizing the spectral radii of luminance and chrominance channels subject to perfect reconstruction. Hao, et al. further extend the method in [2] to formulate an optimization problem for selecting an optimal tri-color CFA [3]. In [4], an analysis on the modulation transfer function of arbitrary pixel shapes is performed and utilized for the pixel design of the CFA.

Meanwhile, advancements in deep learning-based image demosaicing techniques have demonstrated promising results when applied to Bayer CFA (color filter array) images [5],

The associate editor coordinating the review of this manuscript and approving it for publication was Joewono Widjaja.

[6] [7], [8]. However, demosaicing for non-Bayer patterned CFAs remains still challenging, especially when the CFA includes colored pixels. This challenge stems from the requirement to solve a linear system rather than simply interpolating data. Nonetheless, CFAs incorporating colored pixels are desired due to their potential to enhance image quality by improving the SNR (signal-to-noise) ratio. This enhancement is achievable as colored pixels allow more light to reach the sensor, thereby increasing the signal energy. Furthermore, a colored filter array can enable improved spectral sensitivity, which can be useful for specialized or scientific applications. Therefore, in this paper, we aim to design a CFA pattern that includes colored pixels and determine the CFA coefficient values so that the CFA works effectively with a deep learning-based generative model, i.e., the denoising diffusion null-space model (DDNM) for demosaicing.

The challenge with CFA design for deep learning-based demosaicing is that the design of an appropriate CFA must be considered in conjunction with a learning-based model. Since the demosaicing method relies on a learned model derived from training data, establishing a rigorous relationship between the CFA and the demosaicing method becomes difficult. This difficulty has resulted in a lack of research on CFA design for deep learning-based demosaicing methods. With deep learning-based demosaicing methods, the main issue in determining the CFA pattern and the CFA coefficient values is ensuring that the CFA can guide the demosaicing network to find the optimal solution.

Meanwhile, with the advancement of diffusion-based generative models [11], [12], [15], image restoration methods have begun to incorporate pre-trained diffusion models [9], [10], [13], [14]. Among the methods, the authors of [14] introduce a highly efficient zero-shot framework known as the denoising diffusion null-space Model(DDNM), which can be applied to various linear image restoration problems.

In this paper, we first demonstrate how the DDNM model can be applied to the demosaicing problem by constructing the appropriate degradation matrix. However, the DDNM model cannot successfully demosaic every CFA image. For instance, when using an All-White CFA image, experimental results reveal the reconstruction of false colors. Additionally, the performance of the DDNM model fluctuates across different CFA images. Therefore, we provide a brief analysis of the conditions under which a CFA image with colored pixels can yield a unique solution, i.e., a uniquely reconstructed color image using the DDNM model. Leveraging the result of this analysis, we propose a colored CFA pattern that offers a high likelihood of accurately restoring the original color components. Furthermore, we propose a training method that can find the optimal filter coefficients that yields high-quality demosaicing outcomes when employing the DDNM as the demosaicing technique.

The main contributions of this work can be summarized as follows:

- We conduct an analysis to determine the conditions under which the DDNM can generate a unique solution.
- Based on the analysis, we propose a CFA pattern that offers a high level of assurance in recovering the true colors from the CFA image.
- We present a method for training the CFA filter coefficients to achieve optimal demosaicing results when employing the DDNM as the demosaicing method.

II. PRELIMINARIES

In this section, we outline the preliminaries required to understand the proposed method.

A. THE DEMOSAICKING PROBLEM

Contemporary digital imaging systems frequently comprise monochrome image sensors overlaid with color filter arrays (CFAs) to capture color information. The Bayer CFA, which is the most prevailing choice among all CFAs, is composed of Red, Green, and Blue color filters. This arrangement allows for the acquisition of only one color component at each pixel, as illustrated in Fig. 1. Consider Ω as the two-dimensional spatial domain of the image. Let $\mathbf{I}_{orig}[\mathbf{k}] = [R_{orig}[\mathbf{k}], G_{orig}[\mathbf{k}], B_{orig}[\mathbf{k}]]^T$ represent the true color image at $\mathbf{k}$, while $\mathbf{C_k} = [c_R[\mathbf{k}], c_G[\mathbf{k}], c_B[\mathbf{k}]]$ is a 1×3 matrix that signifies the effect of the CFA. Subsequently, the intensity value $I_s[\mathbf{k}]$ of the Red (or Green/Blue) component detected at position $\mathbf{k} \in \Omega$ within the Bayer CFA can be mathematically represented as the inner product between $\mathbf{C_k}$ and $\mathbf{I}_{orig}[\mathbf{k}]$:

$$I_s[\mathbf{k}] = \mathbf{C_k}\mathbf{I}_{orig}[\mathbf{k}]. \tag{1}$$

For the Bayer CFA, the components of $\mathbf{C_k}$ are defined as

$$\begin{aligned} c_R[\mathbf{k}] = 1, c_G[\mathbf{k}] = 0, c_B[\mathbf{k}] = 0 \quad & \text{if } \mathbf{k} \in S_R \\ c_R[\mathbf{k}] = 0, c_G[\mathbf{k}] = 1, c_B[\mathbf{k}] = 0 \quad & \text{if } \mathbf{k} \in S_G \\ c_R[\mathbf{k}] = 0, c_G[\mathbf{k}] = 0, c_B[\mathbf{k}] = 1 \quad & \text{if } \mathbf{k} \in S_B. \end{aligned} \tag{2}$$

Here, S_R, S_G, and S_B correspond to the sets of pixels containing only the Red, Green, and Blue components, respectively. The aim of demosaicing is to reconstruct $\mathbf{I}_{orig}$ from I_s. This is an underdetermined inverse problem since a three-channel image $\mathbf{I}_{orig}$ has to be reconstructed using the sensed image I_s which has only one channel. Consequently, solving the demosaicking problem requires the imposition of supplementary constraints. One common approach is to introduce the assumption that neighboring pixels exhibit comparable colors. In conventional demosaicing techniques, this assumption is employed to interpolate the absent two color components based on the spatially neighboring CFA data.

Meanwhile, due to their potential to enhance image quality by reducing noise, many colored CFAs, including RGBW CFAs, have been proposed. This is because colored pixels, including white pixels, allow more light to reach the sensor.

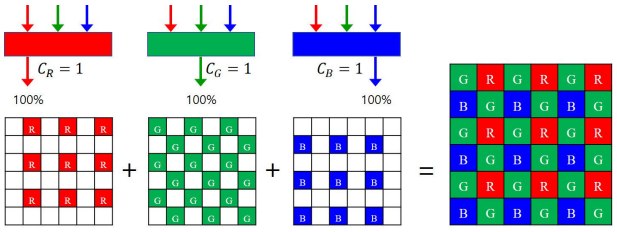


FIGURE 1. Partial cut of a Bayer color filter array.

For colored pixels, the CFA coefficient values are normally non-zeros for all color components:

$$c_R[\mathbf{k}] = \alpha_R, c_G[\mathbf{k}] = \alpha_G, c_B[\mathbf{k}] = \alpha_B, \tag{3}$$

where $\alpha_R \neq 0$, $\alpha_G \neq 0$, and $\alpha_B \neq 0$. For example, for a white pixel as used in [26], α_R, α_G, and α_B becomes $\alpha_R = 0.2936$, $\alpha_G = 0.4905$, and $\alpha_B = 0.2159$. The demosaicing of colored pixels is more complex than the demosaicing of R, G, and B pixels, because it involves solving a linear system of equations rather than simply performing interpolation.

B. DENOISING DIFFUSION NULL-SPACE MODEL

The Denoising Diffusion Null-Space Model (DDNM) method [14] tries to solve the noise-free image restoration problem

$$\mathbf{y} = \mathbf{A}\mathbf{x}, \tag{4}$$

where $\mathbf{x} \in \mathbb{R}^{D\times 1}$, $\mathbf{A} \in \mathbb{R}^{d\times D}$, $\mathbf{y} \in \mathbb{R}^{d\times 1}$, denote the solution image, the linear degradation operator, and the degraded image, respectively. Here, D and d denote the lengths of $\mathbf{x}$ and $\mathbf{y}$, respectively. The DDNM is based on the Denoising Diffusion Implicit Model(DDIM) which starts from a noise image $\mathbf{x}_{t=T}$, where T is the total number of time steps in the forward diffusion, and converges to a clean generated image $\mathbf{x}_0$ as it gots denoised by the following reverse diffusion equation,

$$\mathbf{x}_{t-1} = \frac{1}{\sqrt{\alpha_t}}\left(\mathbf{x}_t - \boldsymbol{\epsilon}_\theta(\mathbf{x}_t, t)\frac{1-\alpha_t}{\sqrt{1-\bar{\alpha}_t}}\right), \tag{5}$$

where $\boldsymbol{\epsilon}_\theta(\mathbf{x}_t, t)$ denotes a neural network trained to estimate the noise in $\mathbf{x}_t$ at t, α_t is a predefined scale factor at time t, and $\bar{\alpha}_t = \prod_{i=0}^{t} \alpha_i$. If we recursively apply (5), we can establish a direct relationship between $\mathbf{x}_t$ and $\mathbf{x}_0$ as follows:

$$\mathbf{x}_0 = \frac{1}{\sqrt{\bar{\alpha}_t}}\left(\mathbf{x}_t - \boldsymbol{\epsilon}_\theta(\mathbf{x}_t, t)\sqrt{1-\bar{\alpha}_t}\right). \tag{6}$$

Here, $\mathbf{x}_0$ should be denoted as $\mathbf{x}_{0|t}$ because it is not the ideal clean image, but rather an estimate of the clean image $\mathbf{x}_0$ at time t, so (6) becomes:

$$\mathbf{x}_{0|t} = \frac{1}{\sqrt{\bar{\alpha}_t}}\left(\mathbf{x}_t - \boldsymbol{\epsilon}_\theta(\mathbf{x}_t, t)\sqrt{1-\bar{\alpha}_t}\right). \tag{7}$$

As $t \to T$, $\mathbf{x}_{0|t}$ approaches the desired clean image. However, $\mathbf{x}_{0|t}$ is not the restored version of $\mathbf{y}$, as $\mathbf{x}_{0|t}$ includes no information of $\mathbf{y}$. Therefore, to obtain a restored version of $\mathbf{y}$, the DDNM method decomposes $\mathbf{x}_{0|t}$ into its range-space part and null-space part as $\mathbf{x}_{0|t} = \mathbf{A}^\dagger\mathbf{A}\mathbf{x}_{0|t} + (\mathbf{I} - \mathbf{A}^\dagger\mathbf{A})\mathbf{x}_{0|t}$ and then replaces the range-space part $\mathbf{A}^\dagger\mathbf{A}\mathbf{x}_{0|t}$ by $\mathbf{A}^\dagger\mathbf{y}$ at every time step t to obtain a rectified version of $\mathbf{x}_{0|t}$:

$$\hat{\mathbf{x}}_{0|t} = \mathbf{A}^\dagger\mathbf{y} + (\mathbf{I} - \mathbf{A}^\dagger\mathbf{A})\mathbf{x}_{0|t}. \tag{8}$$

Now, as $t \to 0$, $\hat{\mathbf{x}}_{0|t}$ converges to the restored version of $\mathbf{y}$ as $\hat{\mathbf{x}}_{0|t}$ obeys the constraint $\mathbf{A}\hat{\mathbf{x}}_{0|t} = \mathbf{y}$. This is due to the fact that the range-space part when multiplied by $\mathbf{A}$ becomes $\mathbf{y}$, i.e., $\mathbf{A}\mathbf{A}^\dagger\mathbf{y} = \mathbf{y}$, while the null-space part when multiplied by $\mathbf{A}$ becomes $\mathbf{0}$, i.e., $\mathbf{A}(\mathbf{I} - \mathbf{A}^\dagger\mathbf{A})\mathbf{x}_{0|t} = \mathbf{0}$. The role of the null-space part is that it adds realness to $\hat{\mathbf{x}}_{0|t}$.

III. PROPOSED METHOD

In this section, we first demonstrate how to apply the DDNM to the demosaicing problem. Then, we provide a brief analysis of the conditions under which the DDNM can recover the original colors, and based on this analysis, we propose a CFA pattern. Lastly, we outline the method for obtaining the optimal filter coefficients for the proposed CFA pattern.

A. APPLYING THE DDNM TO THE DEMOSAICING PROBLEM

The DDNM can be used to solve the demosaicking problem by letting the linear degradation matrix $\mathbf{A}$ to be a pixel-wise operator that converts the RGB channel pixels into CFA image pixels. That is, for every pixel $\mathbf{k}$, we can define the pixel-wise operator $\mathbf{A}_\mathbf{k}$ to be the action of the CFA, i.e., $\mathbf{A}_\mathbf{k} = [c_R[\mathbf{k}], c_G[\mathbf{k}], c_B[\mathbf{k}]]$ where $c_R[\mathbf{k}]$, $c_G[\mathbf{k}]$, and $c_B[\mathbf{k}]$ are as defined in (5). For example, for the Red pixel we have $\mathbf{A}_\mathbf{k} = \begin{bmatrix}1 & 0 & 0\end{bmatrix}$. For a colored pixel we have $\mathbf{A}_\mathbf{k} = \begin{bmatrix}c_R[\mathbf{k}] & c_G[\mathbf{k}] & c_B[\mathbf{k}]\end{bmatrix} = \begin{bmatrix}\alpha_R[\mathbf{k}] & \alpha_G[\mathbf{k}] & \alpha_B[\mathbf{k}]\end{bmatrix}$, according to (3).

The problem of demosaicing for a colored pixel is to recover the original color from the given sensed (mosaiced) intensity $I_s[\mathbf{k}] = \alpha_R[\mathbf{k}]R_{orig}[\mathbf{k}] + \alpha_G[\mathbf{k}]G_{orig}[\mathbf{k}] + \alpha_B[\mathbf{k}]B_{orig}[\mathbf{k}]$ for every pixel $\mathbf{k}$ in Ω. To apply (8) to the demosaicking problem, $\mathbf{y}$ becomes $I_s[\mathbf{k}]$ which is now a single value. Then, the demosaiced image can be point-wisely reconstructed by using the following equation instead of (8) in the DDNM process:

$$\hat{\mathbf{x}}^\mathbf{k}_{0|t} = \mathbf{A}^\dagger_\mathbf{k} I_s[\mathbf{k}] + (\mathbf{I} - \mathbf{A}^\dagger_\mathbf{k}\mathbf{A}_\mathbf{k})\mathbf{x}^\mathbf{k}_{0|t}. \tag{9}$$

As $t \to 0$, the image containing all the pixel values $\hat{\mathbf{x}}^\mathbf{k}_{0|t}$ converges to the demosaiced image.

B. PROPOSED CFA PATTERN GUARANTEEING UNIQUE SOLUTION

In this section, a CFA pattern that incorporates colored pixels is introduced, resulting in favorable demosaicing outcomes when the DDNM method is employed. For clarity, $I_s[\mathbf{k}]$ is denoted as I_s, $\alpha_R[\mathbf{k}]$, $\alpha_G[\mathbf{k}]$, and $\alpha_B[\mathbf{k}]$ are denoted as α_R, α_G, and α_B, respectively, and $R_{orig}[\mathbf{k}]$, $G_{orig}[\mathbf{k}]$, and $B_{orig}[\mathbf{k}]$ are

denoted as R, G, and B. The term "colored pixels" refers to pixels with non-zero filter coefficients for all color channels, i.e., $\alpha_R \neq 0$, $\alpha_G \neq 0$, and $\alpha_B \neq 0$, implying that even a white pixel can be considered a colored pixel. It should be noted that, unlike other restoration applications, the degradation operator $\mathbf{A}$ is a pixel-wise operator, which varies depending on the colored pixel in the CFA. The conditions under which the DDNM can accurately reconstruct the original colors in the CFA are first identified. Using the simplified notation, (9) is rewritten as:

$$\hat{\mathbf{x}}_{0|t} = \mathbf{A}^{\dagger} I_s + (\mathbf{I} - \mathbf{A}^{\dagger}\mathbf{A})\mathbf{x}_{0|t}, \tag{10}$$

where $\mathbf{A} = \begin{bmatrix} \alpha_R & \alpha_G & \alpha_B \end{bmatrix}$, and $\mathbf{A}^{\dagger}$, the pseudo-inverse of $\mathbf{A}$, becomes:

$$\mathbf{A}^{\dagger} = \frac{1}{\alpha_R^2 + \alpha_G^2 + \alpha_B^2} \begin{bmatrix} \alpha_R \\ \alpha_G \\ \alpha_B \end{bmatrix}. \tag{11}$$

It should be noted that $\hat{\mathbf{x}}_{0|t}$ represents now a pixel-wise solution, i.e., the color vector of a single pixel, rather than the entire solution image. At convergence, $\mathbf{x}_{0|t}$ converges to $\hat{\mathbf{x}}_{0|t}$, and therefore, (10) becomes

$$\mathbf{A}^{\dagger} I_s - \mathbf{A}^{\dagger}\mathbf{A}\mathbf{x}_{0|t} = \mathbf{0}. \tag{12}$$

Simplifying $\mathbf{x}_{0|t}$ as $\mathbf{x}$, and letting $\kappa = \alpha_R^2 + \alpha_G^2 + \alpha_B^2$ and $I_s = \alpha_R R + \alpha_G G + \alpha_B B$, we get

$$\begin{aligned} &\mathbf{A}^{\dagger} I_s - \mathbf{A}^{\dagger}\mathbf{A}\mathbf{x} \\ &= \frac{1}{\kappa} \begin{bmatrix} \alpha_R^2 & \alpha_R\alpha_G & \alpha_R\alpha_B \\ \alpha_G\alpha_R & \alpha_G^2 & \alpha_G\alpha_B \\ \alpha_B\alpha_R & \alpha_B\alpha_G & \alpha_B^2 \end{bmatrix} \begin{bmatrix} R - x_1 \\ G - x_2 \\ B - x_3 \end{bmatrix} \\ &= \mathbf{B}\mathbf{n} = \mathbf{0}, \end{aligned} \tag{13}$$

where we let

$$\mathbf{B} = \frac{1}{\kappa} \begin{bmatrix} \alpha_R^2 & \alpha_R\alpha_G & \alpha_R\alpha_B \\ \alpha_G\alpha_R & \alpha_G^2 & \alpha_G\alpha_B \\ \alpha_B\alpha_R & \alpha_B\alpha_G & \alpha_B^2 \end{bmatrix}, \quad \mathbf{n} = \begin{bmatrix} R - x_1 \\ G - x_2 \\ B - x_3 \end{bmatrix}. \tag{14}$$

If the null space solution $\mathbf{n}$ is the zero vector, this implies that $x_1 = R$, $x_2 = G$, and $x_3 = B$, indicating that the solution identifies the original color components. However, achieving this relies on the invertibility of $\mathbf{B}$, which is not the case, as $\mathbf{B}$ has a rank of 1, since each row is a multiple of the others.

To derive a condition for $\mathbf{x}$ to ensure that the solution represents the original color components, we introduce an additional assumption, i.e., the assumption that neighboring pixels have same color components. Based on the assumption, we aim to obtain an identical solution $\mathbf{x}$ for neighboring pixels. Furthermore, we assume that the filter coefficients of the CFA exhibit a cyclic property, meaning that three neighboring pixels in the CFA have filter coefficients of $\boldsymbol{\alpha}_a = [\alpha_1, \alpha_2, \alpha_3]$, $\boldsymbol{\alpha}_b = [\alpha_3, \alpha_1, \alpha_2]$, and $\boldsymbol{\alpha}_c = [\alpha_2, \alpha_3, \alpha_1]$. The Bayer CFA is a special case exhibiting this cyclic property, wherein the R, G, and B pixels have filter coefficients of $[\alpha_1, \alpha_2, \alpha_3]$, $[\alpha_3, \alpha_1, \alpha_2]$, and $[\alpha_2, \alpha_3, \alpha_1]$, with $\alpha_1 = 1$, $\alpha_2 = 0$, and $\alpha_3 = 0$. The sensed values for three different neighboring pixels, i.e., I_s^a, I_s^b, and I_s^c will become:

$$\begin{cases} I_s^a = \alpha_1 R + \alpha_2 G + \alpha_3 B \\ I_s^b = \alpha_3 R + \alpha_1 G + \alpha_2 B \\ I_s^c = \alpha_2 R + \alpha_3 G + \alpha_1 B. \end{cases} \tag{15}$$

The problem (12) becomes identical to the following problem, under the condition that α_R, α_G, and α_B are all non-zero:

$$I_s - \mathbf{A}\mathbf{x} = 0. \tag{16}$$

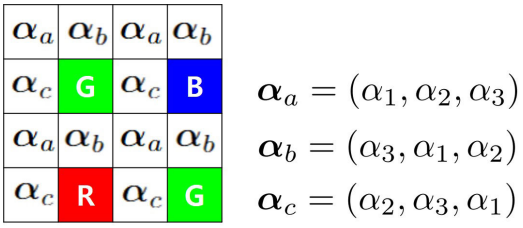


FIGURE 2. Proposed CFA pattern composed of colored pixels α_a, α_b, and α_c exhibiting cyclic properties, along with R, G, and B pixels.

Applying (16) to our three neighboring pixels, and replacing α_R, α_G, and α_B by α_1, α_2, and α_3, and define $\mathbf{A}_a = \begin{bmatrix} \alpha_1 & \alpha_2 & \alpha_3 \end{bmatrix}$, $\mathbf{A}_b = \begin{bmatrix} \alpha_3 & \alpha_1 & \alpha_2 \end{bmatrix}$, and $\mathbf{A}_c = \begin{bmatrix} \alpha_2 & \alpha_3 & \alpha_1 \end{bmatrix}$, respectively, we get

$$\begin{cases} I_s^a - \mathbf{A}_a\mathbf{x} = 0 \\ I_s^b - \mathbf{A}_b\mathbf{x} = 0 \\ I_s^c - \mathbf{A}_c\mathbf{x} = 0. \end{cases} \tag{17}$$

This can be re-written in matrix form as

$$\begin{bmatrix} \alpha_1 & \alpha_2 & \alpha_3 \\ \alpha_3 & \alpha_1 & \alpha_2 \\ \alpha_2 & \alpha_3 & \alpha_1 \end{bmatrix} \begin{bmatrix} R - x_1 \\ G - x_2 \\ B - x_3 \end{bmatrix} = \mathbf{0}. \tag{18}$$

The condition for the solution $\mathbf{x}$ to become the sought-after solution $\mathbf{x} = \begin{bmatrix} R & G & B \end{bmatrix}^T$, i.e., the original color components, is outlined in the following proposition:

Proposition 1: If $\alpha_1 \neq 0$, $\alpha_2 \neq 0$, $\alpha_3 \neq 0$, and if $\alpha_1^3 + \alpha_2^3 + \alpha_3^3 \neq 3\alpha_1\alpha_2\alpha_3$, the solution of (18) becomes $\mathbf{x} = \begin{bmatrix} R & G & B \end{bmatrix}^T$.

Proposition 1 can be easily proved by establishing the condition that (12) and (16) become identical problems and the condition that the determinant of the matrix in (18) is non-zero. Based on the analysis, we propose the CFA pattern shown in Fig. 2 to be used together with the DDNM as the demosaicing method. The filter coefficients exhibit a cyclic property for three neighboring pixels. Since the DDNM prefers natural images, an inherent constraint arises, i.e., the condition that neighboring pixels should have similar or identical colors. This constraint satisfies the condition for Proposition 1 to hold and contributes to the restoration of the original colors.

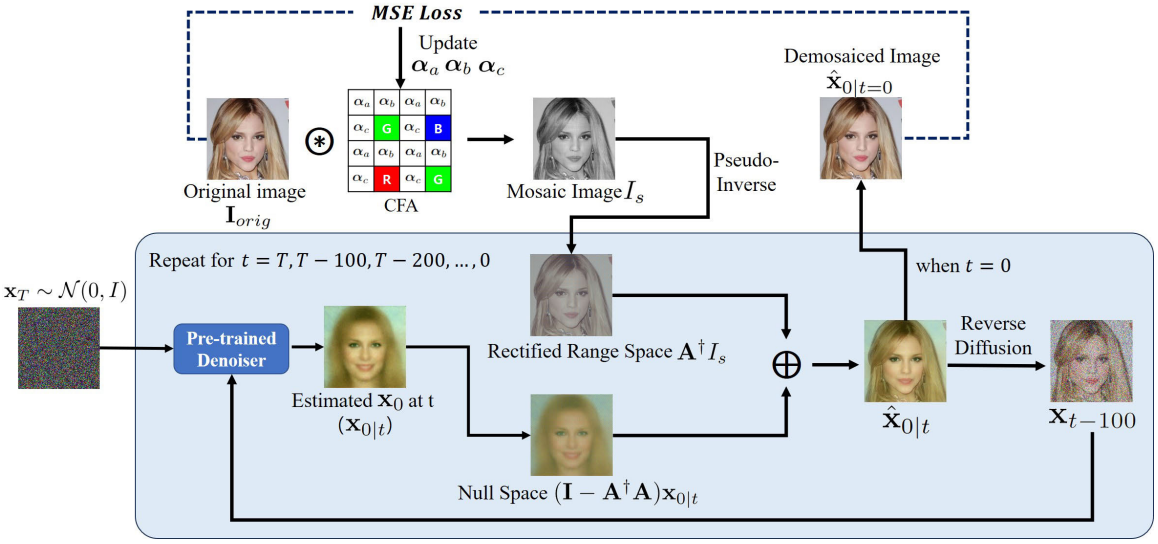


FIGURE 3. The overall training workflow for optimizing the color filter coefficients.

In addition to the colored pixels, we allocate 25% of the pixels as R, G, and B pixels. The R (or G, B) pixels serve as anchor points that retain the true R (or G, B) color components of the sensed image, thereby aiding in the discrimination of the true color components.

C. OPTIMIZING THE CFA COEFFICIENTS

In this section, we demonstrate the process of obtaining the optimal filter coefficient values for the CFA pattern proposed in the previous section. We optimize the filter coefficient values α_1, α_2, and α_3 with a gradient descent based optimization technique, the ADAM optimizer. The coefficient values are all initiated with a value of 0.333, and then updated by taking a gradient step on the mean squared error (MSE) loss between the original image $\mathbf{I}^i_{orig}$ and the demosaicked image $\hat{\mathbf{x}}_{0|t=0}$ in each iteration i of the outer loop in Algorithm 1. That is, for every iteration i, we minimize the following loss function with respect to the coefficient values:

$$\mathcal{L} = ||\mathbf{I}^i_{orig} - \hat{\mathbf{x}}_{0|t=0}||_2^2, \; i = 1, 2, \ldots, N. \tag{19}$$

Additionally, the softmax function is applied to α_1, α_2, and α_3 to constrain them within a range of 0 to 1. Figure 3 provides an overview of the training workflow, while Algorithm 1 presents the detailed pseudocode.

It should be noted that the parameters of the pretrained network $\boldsymbol{\epsilon}_\theta(\cdot)$ are not updated during the optimization process. Instead, the gradient values of $\boldsymbol{\epsilon}_\theta(\cdot)$ are detached during the backpropagation as shown in Fig. 4. That is, we are not fine-tuning $\boldsymbol{\epsilon}_\theta(\cdot)$. This approach reduces memory consumption and simplifies the training process, especially considering the large number of parameters in $\boldsymbol{\epsilon}_\theta(\cdot)$.

By detaching the pre-trained network during the training process, the denoising process is excluded from the backpropagation. This approach is taken because we are not interested in identifying the noise from which the generated image originates. Our primary focus is on determining the CFA coefficients that can produce an optimal null space and range space, which ensures the high quality of the reconstructed image. Even though the pre-trained network is detached during training, the gradient can still flow through the null space and range space images. Since both the null space and range space images are functions of $\mathbf{A}$, the gradient flow can update $\mathbf{A}$, thereby updating the coefficients in the CFA.

Since we are not tracking the denoising process during backpropagation, it is unnecessary to monitor the gradients at all the time steps in the reverse diffusion. This is because the gradient with respect to $\hat{\mathbf{x}}_{0|t=0}$ remains unchanged during backpropagation. However, we still need the denoised reverse diffusion result during the forward propagation. Therefore, instead of running the full T time steps originally required by the pre-trained diffusion model, we run only one-hundredth of the time steps during training, i.e., run for $t = T, T - 100, T - 200, \ldots, 0$, where $T = 1000$. This greatly reduces the training time. During testing, we run one-tenth of the whole time steps, i.e., run for $t = T, T - 10, T - 20, \ldots, 0$.

Following the optimization process, the filter coefficients typically converge to $\alpha_1 = 0.20504$, $\alpha_2 = 0.05198$, and $\alpha_3 = 0.74298$. Occasionally, the values may rotate, for example, resulting in $\alpha_2 = 0.20504$, $\alpha_3 = 0.05198$, and $\alpha_1 = 0.74298$. However, the relative ratios between the coefficients remain the same, ensuring consistent performance. This rotation does not affect the overall effectiveness of the filter, as the proportional relationship between the coefficients is maintained, which is the key factor in achieving optimal results.

The optimized values result from a struggle to strike a balance between robustness and accurate color reproduction. A well-conditioned matrix in (18) enhances the likelihood

Algorithm 1 Training Workflow for $\alpha_1, \alpha_2, \alpha_3$

Initialize: $\alpha_1, \alpha_2, \alpha_3 \leftarrow 0.333, 0.333, 0.333, e \leftarrow 10^{-8}$,
$N \leftarrow 1000, T \leftarrow 1000, \beta_1, \beta_2 \leftarrow 0.9, 0.999$,
$\gamma \leftarrow 0.001, \lambda \leftarrow 0.001, m_1, m_2, m_3 \leftarrow 0, 0, 0$ (first moment), $v_1, v_2, v_3, \leftarrow 0, 0, 0$ (second moment)

for $i = 1, 2, \ldots,$N **do**
  New input image: $\mathbf{I}^i_{orig}$
  $(\alpha_1, \alpha_2, \alpha_3) \leftarrow \text{Softmax}(\alpha_1, \alpha_2, \alpha_3)$
  $\mathbf{A} \leftarrow \text{CreateCFA}(\alpha_1, \alpha_2, \alpha_3)$ ▷ Mosaicing
  Compute $I_s = \mathbf{A}\mathbf{I}^i_{orig}, \mathbf{A}^\dagger I_s, \mathbf{A}^\dagger\mathbf{A}$
  $\boldsymbol{\epsilon}_\theta(\mathbf{x}_t, t) \leftarrow$ Pre-trained noise predicting UNet

  DDNM based Demosaicing
  $\mathbf{x}_T \sim \mathcal{N}(0, I)$
  for $t = T, T-100, T-200, \ldots, 0$ **do**
    $\boldsymbol{\epsilon}_t \leftarrow \boldsymbol{\epsilon}_\theta(\mathbf{x}_t, t)$ ▷ No gradient calculated here
    $\mathbf{x}_{0|t} \leftarrow \frac{1}{\sqrt{\bar{a}}}(\mathbf{x}_t - \boldsymbol{\epsilon}_t\sqrt{1-\bar{a}_t})$ ▷ $\boldsymbol{\epsilon}_t$ is detached
    $\hat{\mathbf{x}}_{0|t} \leftarrow \mathbf{A}^\dagger I_s + (\mathbf{I} - \mathbf{A}^\dagger\mathbf{A})\mathbf{x}_{0|t}$
    $\mathbf{x}_{t-100} = \frac{\sqrt{\bar{\alpha}_{t-1}}\beta_t}{1-\bar{\alpha}_t}\hat{\mathbf{x}}_{0|t} + \frac{\sqrt{\alpha_t}(1-\bar{\alpha}_{t-1})}{1-\bar{\alpha}_t}\mathbf{x}_t + \sigma_t\boldsymbol{\epsilon}$
  end for
  Intermediate Output: $\hat{\mathbf{x}}_{0|t=0}$

  Update $\alpha_1, \alpha_2, \alpha_3$ **with Adam Optimizer**
  $g_k \leftarrow \lambda\frac{\partial}{\partial\alpha_k}||\mathbf{I}^i_{orig} - \hat{\mathbf{x}}_{0|t=0}||_2^2$, k=1,2,3
  $m_k \leftarrow \beta_1 m_k + (1-\beta_1)g_k, \hat{m}_k \leftarrow \frac{m_k}{1-\beta_1}$, k=1,2,3
  $v_k \leftarrow \beta_2 v_k + (1-\beta_2)g_k^2, \hat{v}_k \leftarrow \frac{v_k}{1-\beta_2}$, k=1,2,3
  $\alpha_k \leftarrow \alpha_k - \gamma\frac{\hat{m}_k}{\sqrt{\hat{v}_k}+e}$, k=1,2,3
end for
Final Output: Optimized $\alpha_1, \alpha_2, \alpha_3$

of achieving a robust numerical solution for the null-space. Such a matrix is obtained when the matrix in (18) deviates from uniformity, since this deviation increases the condition number. Meanwhile, having all non-zero values in the filter coefficient ensures the consideration of each color channel's contribution and provides more color information per pixel. This results in cleaner and more accurate color reproduction. This roughly explains why the resulting filter coefficients exhibit non-uniformity, while still preserving information about all color components.

IV. EXPERIMENTAL RESULTS

We used the CelebA-HQ dataset [24] for our training and testing experiments. This dataset comprises 30,000 face images, each of size 256 × 256. The dataset is divided into a training set and a testing set with a ratio of 9:1. During training, we employed a batch size of 8, while for testing, we used a batch size of 16. The training process exclusively focuses on determining the optimal filter coefficients, while the parameters for the neural network remain fixed and are not updated during training. Each batch was trained for only

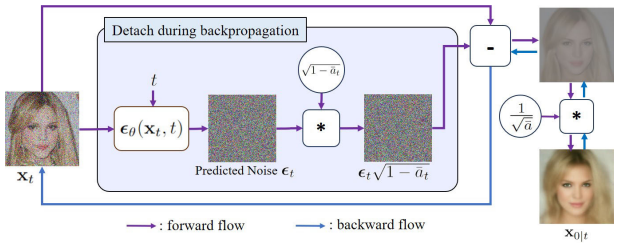


FIGURE 4. Detaching of the pre-trained network during the optimization process.

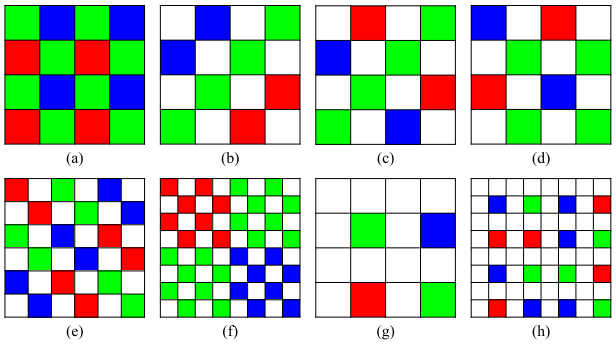


FIGURE 5. Different kinds of CFA: (a) Bayer [16] (b) Kodak [17] (c) Sony [18] (d) Yamagami [19] (e) Kaizu [20] (f) Hamilton [21] (g) Honda [22] (h) Random [23].

1 epoch. We utilized the Adam optimizer with a learning rate(λ) of 0.001, and set the hyper-parameters (β_1, β_2, etc.) as shown in Algorithm 1.

We conducted qualitative and quantitative evaluations to compare the demosaicing results with other CFAs, using the DDNM as the demosaicing method for all CFAs. The CFAs we used for comparison are depicted in Fig. 5, showing various patterns with different combinations of white pixels and R, G, and B pixels. For the white pixels, we employed the coefficient values of $\alpha_R = 0.2936, \alpha_G = 0.4905, \alpha_B = 0.2159$ as used in VEML6040 sensors from Vishay company [25], [26].

TABLE 1. Average PSNR & SSIM values of the demosaiced images using DDNM-based demosaicing with different CFA patterns.

CFA Pattern	PSNR	SSIM
Bayer [16]	41.0252	0.9776
Kodak [17]	41.9344	0.9882
Sony [18]	42.2027	0.9886
Yamagami [19]	42.1959	0.9887
Kaizu [20]	43.1673	0.9889
Hamilton [21]	41.2943	0.9873
Honda [22]	38.6331	0.9836
Random [23]	39.8758	0.9858
All-white	22.9997	0.9377
Proposed	**44.2341**	**0.9903**

The CFA coefficient values have an influence on both the range space and the null space. The null space is formed by

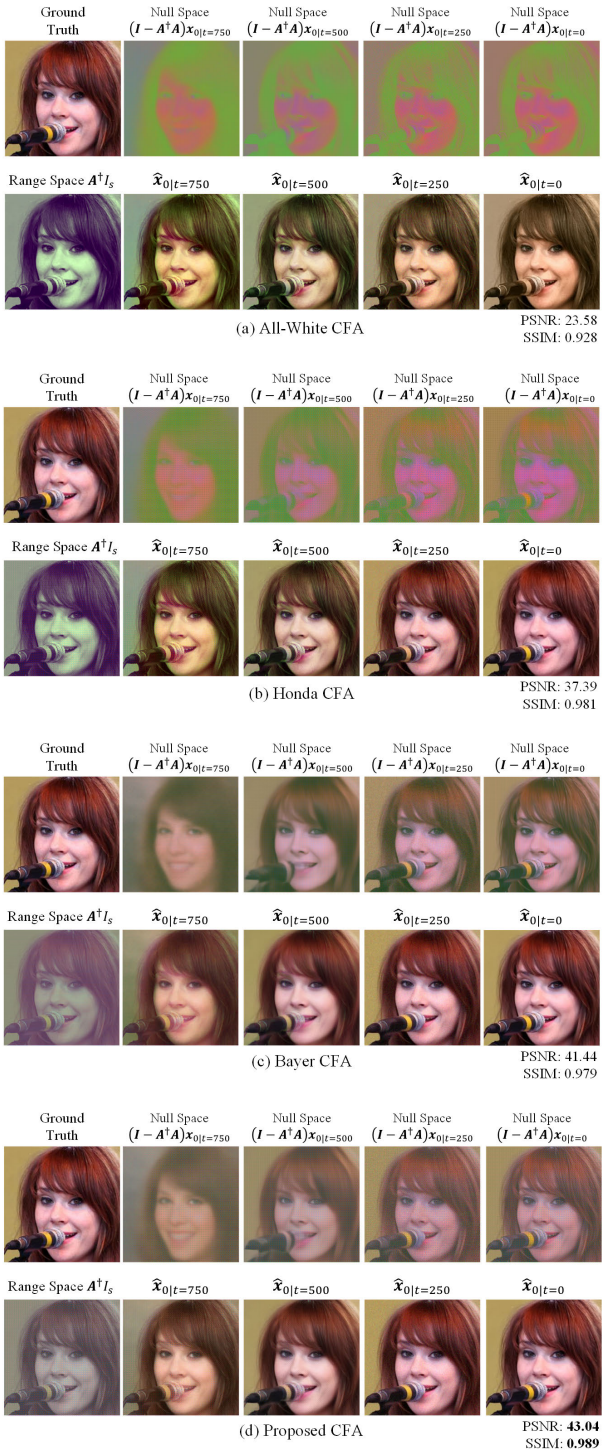


FIGURE 6. Displaying the range space, the null space, and the demosaiced image obtained from various CFAs using the DDNM for demosaicing. The range space remains constant for each CFA, while the null space evolves over time. We show the null spaces for $t = 750$ and $t = 0$, and the corresponding reconstructed images. It should be noted that $\hat{\mathbf{x}}_{0|t=0}$ actually corresponds to only a single pixel value within the reconstructed image, but here we use it to represent all the pixels in the image.

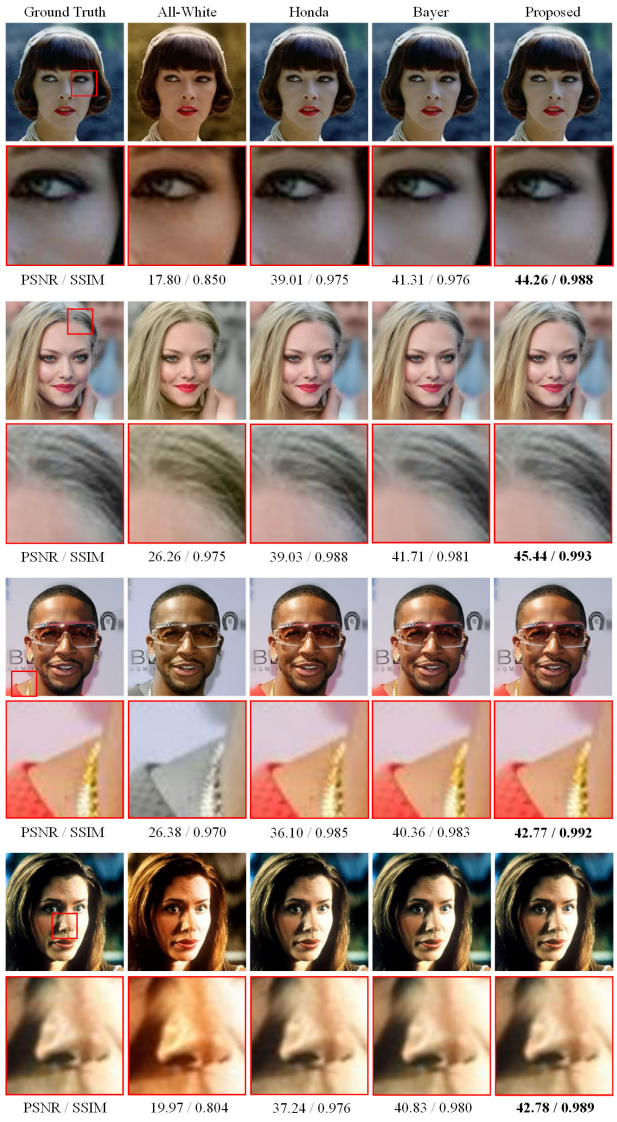


FIGURE 7. Comparison of demosaiced results on the CelebA-HQ dataset with different CFAs. The All-white CFA produces false colors, while the Honda CFA occasionally generates accurate colors with fewer color artifacts compared to Bayer CFA, but results in noisier and rougher textures. Additionally, the Honda CFA fails to produce accurate colors when tested on the ImageNet dataset, as shown in Fig. 8. The Bayer CFA overly smooths the image and introduces false color artifacts. The proposed CFA offers a balanced result with smooth textures as with the Bayer CFA and fewer false color artifacts as with the Honda CFA. This results in better image quality and the highest PSNR/SSIM values.

the multiplication of $\mathbf{I} - \mathbf{A}^{\dagger}\mathbf{A}$ and $\mathbf{x}_{0|t}$. As $\mathbf{x}_{0|t}$ is generated by the reverse diffusion, which favors natural images, it is easier for the DDNM to generate a null space solution that closely resembles a natural image rather than an unnatural one. However, the null space is also restricted by the range space. If the range space limits the null space from producing a natural image, the resulting combination of null space and range space may not yield a high-quality image. Therefore, it is important for the CFA to produce a range space that allows the null space to closely resemble a natural image.

Figure 6(a) and Fig. 6(b) reveal that the null space associated with CFA patterns containing non-varying colored pixels

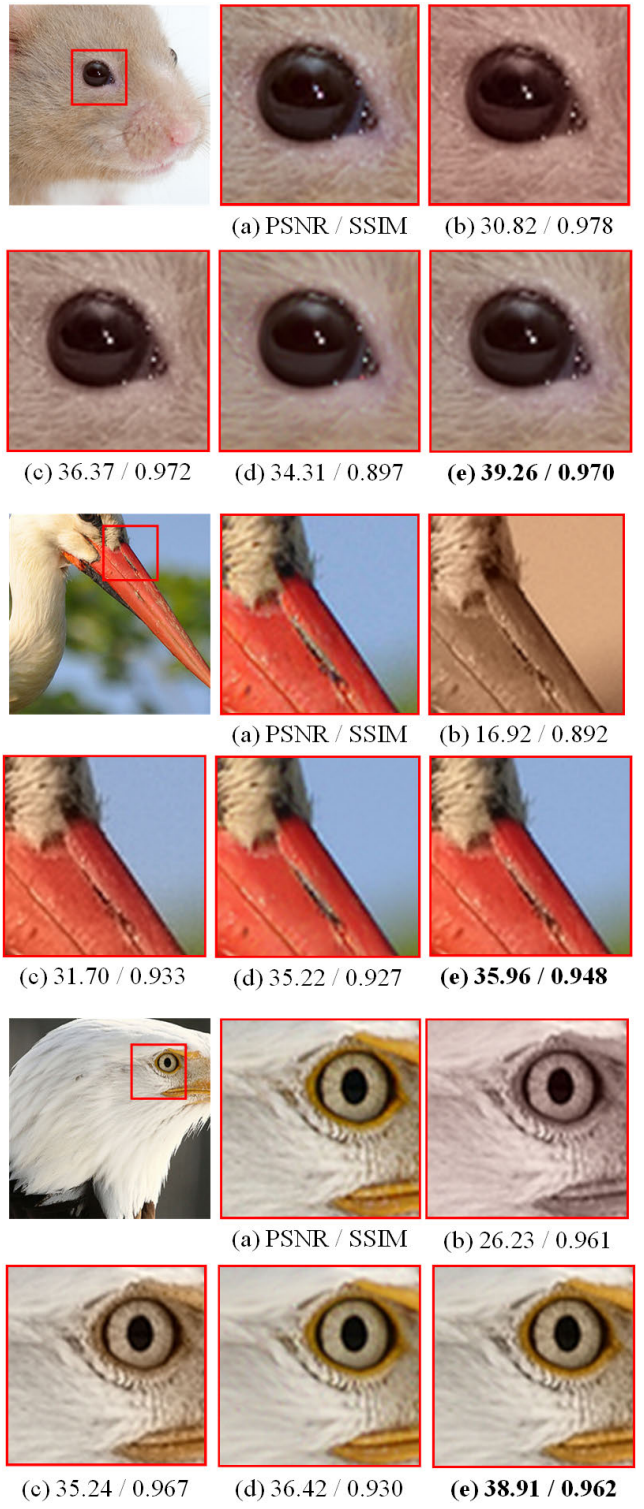


FIGURE 8. Showing the demosaicing results with the proposed CFA trained on the CelebA-HQ dataset and tested on the ImageNet dataset. All CFAs use the DDNM as the demosaicing method. (a) Original (b) Results of the All-white CFA: fails to reconstruct the true colors of the original image (c) Results of the Honda CFA: reconstructs the colors, but does not accurately reflect the true colors. (d) Results of the Bayer CFA: mostly reconstructs colors similar to the original image, but also produces false color artifacts (zoom in to see better, or refer to Fig. 9). (e) Results of the Proposed CFA: successfully reconstructs the original image's colors without false color artifacts, achieving the highest PSNR and SSIM values.

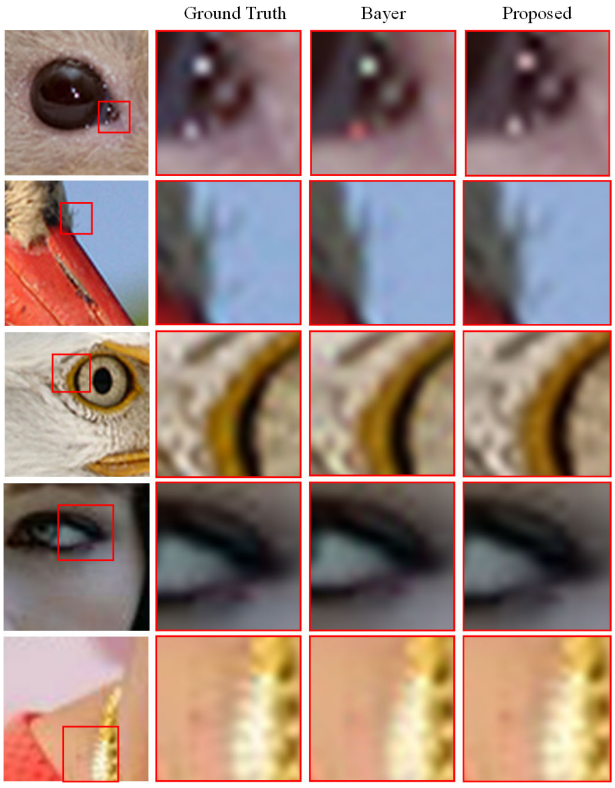


FIGURE 9. Close-up view of the demosaiced results of the images from the ImageNet and CelebA-HQ datasets. The results with the Bayer CFA exhibit some false color artifacts not present in the original image. With the proposed CFA, the details in the original images are recovered.

(white pixels) tends to exhibit non-natural characteristics, attributing to the unnatural appearance of the range space. In contrast, as can be observed in Fig. 6(c) and Fig. 6(d), the null space corresponding to the Bayer CFA and the proposed optimized CFA resembles a more natural image. This is due to the fact that the cyclic pattern of the filter coefficients can better discriminate colors, resulting in a range space that already exhibits diverse colors. Consequently, the null space also showcases diverse colors, contributing to a more natural appearance. Compared with the Bayer CFA, the proposed optimized CFA contains more information of the color components, resulting in a more natural appearance of the range space. After 25 iterations ($t = 750$), the reconstructed image already appears quite natural with the proposed CFA, and the final reconstructed image is superior to the one reconstructed from the Bayer CFA.

Table 1 presents the quantitative results, comparing the demosaicing performance of the proposed CFA with other CFAs using the DDNM method as the demosaicing method. We utilized the PSNR (Peak Signal-to-Noise Ratio) and the SSIM (Structural Similarity Index Measure) metrics for the evaluation. It can be seen in Table 1 that the proposed CFA shows the best PSNR and SSIM values.

Figure 7 compares the demosaiced results on the CelebA-HQ dataset using different CFAs. The All-white

CFA produces false colors. The Honda CFA generates more accurate colors with fewer color artifacts than the Bayer CFA, but results in noisier and rougher textures. Furthermore, the Honda CFA fails to produce accurate colors when tested on the ImageNet dataset, as shown in Fig. 8. The Bayer CFA overly smooths the image and introduces false color artifacts. In comparison, the proposed CFA offers a balanced result, combining the smooth textures of the results using the Bayer CFA and reduced false color artifacts of the results using the Honda CFA. This results in better image quality and the highest PSNR/SSIM values.

To validate that the CFA coefficients optimized for a certain dataset also perform well with other datasets, we conducted additional experiments. We used the same CFA pattern and coefficient values, which were trained on the CelebA-HQ dataset, to test the demosaicing results on the ImageNet dataset. Figure 8 displays the demosaicing results on images from the ImageNet dataset along with the PSNR/SSIM values. The PSNR and SSIM values are slightly lower when tested on the ImageNet dataset compared to the CelebA-HQ dataset because the CFA coefficients were not specifically trained on ImageNet. If the CFA coefficients were trained on the ImageNet dataset, the PSNR/SSIM values would be higher. Nonetheless, as shown in Fig. 8, the CFA coefficients trained on one dataset still produce good results when applied to another dataset. Figure 8(b) shows the results with the All-white CFA which fails to reconstruct the true colors of the original image. The Honda CFA manages to reconstruct the colors, but they appear muted and do not accurately reflect the true colors as can be observed in Fig. 8(c). The Bayer CFA reconstruct colors similar to the original image but also produce false color artifacts as can be seen in Fig. 8(d) and in Fig. 9. In comparison, the proposed successfully reconstructs the original image's colors without false color artifacts, achieving the highest PSNR and SSIM values.

Figure 9 presents further zoomed-in images to better observe the artifacts produced by the Bayer CFA. It can be seen that the Bayer CFA sometimes introduces false colors or removes small structures, whereas the proposed CFA reproduces even the small details in the original image.

V. CONCLUSION

In this paper, we first demonstrated how the Denoising Diffusion Null-space Method (DDNM) can be applied to the demosaicing problem. We then analyzed the conditions under which DDNM-based demosaicing can recover the original color components. Based on this analysis, we proposed a CFA pattern tailored for DDNM-based demosaicing and outlined how to determine the optimal coefficient values for that pattern. Experimental results showed that the proposed CFA achieves superior demosaicing results with DDNM compared to other CFAs. Typically, deep learning-based demosaicing methods focus on optimizing the neural network parameters for a fixed CFA. However, we contend that the CFA itself should also be subject to adaptation in conjunction with the development of new demosaicing methodologies. This work represents the first attempt at designing CFAs for deep learning-based demosaicing methods, particularly those utilizing generative deep learning, and can serve as a reference for future research in this area. Future work for this research could explore how the proposed CFA pattern performs with other generative deep learning models beyond DDNM. Another potential direction is to develop the CFA as an output of a neural network, enabling it to adapt dynamically to changes in scene content or specific image characteristics.

REFERENCES

[1] R. Lukac and K. N. Plataniotis, "Color filter arrays: Design and performance analysis," *IEEE Trans. Consum. Electron.*, vol. 51, no. 4, pp. 1260–1267, Nov. 2005.

[2] K. Hirakawa and P. J. Wolfe, "Spatio-spectral color filter array design for optimal image recovery," *IEEE Trans. Image Process.*, vol. 17, no. 10, pp. 1876–1890, Oct. 2008.

[3] P. Hao, Y. Li, Z. Lin, and E. Dubois, "A geometric method for optimal design of color filter arrays," *IEEE Trans. Image Process.*, vol. 20, no. 3, pp. 709–722, Mar. 2011.

[4] O. Yadid-Pecht, "Geometrical modulation transfer function for dierent pixel active area shapes," *Opt. Eng.*, vol. 39, pp. 859–865, 2000.

[5] D. Wu, Z. Xin, and C. Zhang, "A joint multi-gradient algorithm for demosaicing Bayer images," in *Proc. 8th Int. Conf. Commun., Image Signal Process. (CCISP)*, Chengdu, China, Nov. 2023, pp. 340–346.

[6] L. Liu, X. Jia, J. Liu, and Q. Tian, "Joint demosaicing and denoising with self guidance," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2020, pp. 2237–2246.

[7] T. Li, A. Lahiri, Y. Dai, and O. Mayer, "Joint demosaicing and denoising with double deep image priors," in *Proc. IEEE Int. Conf. Acoust., Speech Signal Process. (ICASSP)*, Seoul, South Korea, Apr. 2024, pp. 4005–4009.

[8] Y. Becker, R. Z. Nossek, and T. Peleg, "SDAT: Sub-dataset alternation training for improved image demosaicing," *IEEE Open J. Signal Process.*, vol. 5, pp. 611–620, 2024.

[9] B. Fei, Z. Lyu, L. Pan, J. Zhang, W. Yang, T. Luo, B. Zhang, and B. Dai, "Generative diffusion prior for unified image restoration and enhancement," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Los Alamitos, CA, USA, Jun. 2023, pp. 9935–9946.

[10] H. Tang, T. Xie, A. Feng, H. Wang, C. Zhang, and Y. Bai, "Solving noisy inverse problems via posterior sampling: A policy gradient view-point," presented at the Symbiosis Deep Learn. Differ. Equ., New Orleans, LA, USA, Dec. 2023.

[11] J. Ho, A. Jain, and P. Abbeel, "Denoising diffusion probabilistic models," in *Proc. NIPS*, 2020, pp. 6840–6851.

[12] J. Song, C. Meng, and S. Ermon, "Denoising diffusion implicit models," in *Proc. 9th Int. Conf. Learn. Represent.*, 2021, pp. 6840–6851.

[13] J. Song, C. Meng, and S. Ermon, "Progressive deblurring of diffusion models for coarse-to-fine image synthesis," presented at the Workshop Score-Based Methods (NeurIPS Workshop), New Orleans, LA, USA, Nov. 2021.

[14] Y. Wang, J. Yu, and J. Zhang, "Zero-shot image restoration using denoising diffusion null-space model," in *Proc. 11th Int. Conf. Learn. Represent. (ICLR)*, Kigali, Rwanda, 2023, pp. 1–33.

[15] J. Choi, S. Kim, Y. Jeong, Y. Gwon, and S. Yoon, "ILVR: Conditioning method for denoising diffusion probabilistic models," in *Proc. IEEE/CVF Int. Conf. Comput. Vis. (ICCV)*, Oct. 2021, pp. 14347–14356.

[16] B. E. Bayer, "Color imaging array," U.S. Patent 3 971 065, Jul. 20, 1976.

[17] J. T. Compton and J. F. Hamilton, "Image sensor with improved light sensitivity," U.S. Patent 8 139 130, Mar. 20, 2012.

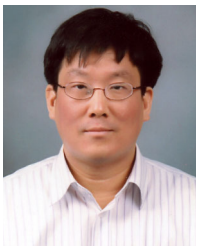
[18] M. Tachi, "Image processing device, image processing method, and program pertaining to image correction," U.S. Patent 8 314 863, Nov. 20, 2012.

[19] T. Yamagami, T. Sasaki, and A. Suga, "Image signal processing apparatus having a color filter with offset luminance filter elements," U.S. Patent 5 323 233, Jun. 21, 1994.

[20] S. Kaizu, "Image processing apparatus, imaging device, image processing method, and program for reducing noise or false colors in an image," U.S. Patent 9 699 429, Jul. 4, 2017.
[21] J. T. Compton and J. F. Hamilton, "Processing color and panchromatic pixels," U.S. Patent 8 274 715, Sep. 25, 2012.
[22] H. Honda, Y. Iida, Y. Egawa, and H. Seki, "A color CMOS imager with 4 × 4 white-RGB color filter array for increased low-illumination signal-to-noise ratio," *IEEE Trans. Electron Devices*, vol. 56, no. 11, pp. 2398–2402, Nov. 2009.
[23] P. Oh, S. Lee, and M. Kang, "Colorization-based RGB-white color interpolation using color filter array with randomly sampled pattern," *Sensors*, vol. 17, no. 7, p. 1523, Jun. 2017.
[24] T. Karras, T. Aila, S. Laine, and J. Lehtinen, "Progressive growing of GANs for improved quality, stability, and variation," in *Proc. Int. Conf. Learn. Represent.*, Vancouver, BC, Canada, 2018, pp. 1–26.
[25] G. O. Young, "Demosaicking algorithms for RGBW color filter arrays," in *Proc. IS&T Int. Symp. Electron. Imag.*, San Francisco, CA, USA, Feb. 2016, pp. 1–6.
[26] *VEML6040 RGBW Color Sensor With I2C Interface*, Vishay Intertechnology, Inc., Malvern, PA, USA, 2016.

INDRA IMANUEL was born in Surabaya, Indonesia. He received the B.Comp. degree in computer engineering from the University of Surabaya, Surabaya, in 2019, and the M.S. degree in computer engineering from Dongseo University, Busan, South Korea, in 2021, where he is currently pursuing the Ph.D. degree in computer engineering. His research interests include deep learning, computer vision, generative model, and image processing. He was a recipient of the Excellent Research Award from Dongseo University, in 2022.

HYOSEON YANG received the B.S., M.S., and Ph.D. degrees in mathematics from Ewha Womans University, Seoul, South Korea, in 2011, 2013, and 2017, respectively. She was a Postdoctoral Research Associate at the Department of Computational Mathematics, Science and Engineering, Michigan State University, East Lansing, MI, USA, from 2018 to 2021. She is currently working as an Assistant Professor with the Department of Mathematics, Kyung Hee University, Seoul. Her research interests include approximation theory, image processing, scientific computation, and numerical PDEs. She was a recipient of the Excellence Award in the field of mathematics at the 7th S-Oil Outstanding Thesis Awards, in 2018.



SUK-HO LEE received the B.S., M.S., and Ph.D. degrees in electronics engineering from Yonsei University, Seoul, South Korea, in 1993, 1998, and 2003, respectively. He was a Researcher at the Impedance Imaging Research Center, from 2003 to 2006, and an Assistant Professor at Yonsei University, from 2006 to 2008. He has been with the Department of Computer Engineering, Dongseo University, Busan, South Korea, since 2008, where he is currently working as a Professor. He has undertaken more than ten government projects related to artificial intelligence (deep learning) over the past five years. His current research interests include deep learning, image processing, image and video filtering based on PDEs, and computer vision. He was awarded the Minister of Education Award in the field of academic promotion, in 2016.