IEEE Access
Multidisciplinary ⋮ Rapid Review ⋮ Open Access Journal

Received November 25, 2021, accepted December 12, 2021, date of publication December 23, 2021,
date of current version December 30, 2021.

Digital Object Identifier 10.1109/ACCESS.2021.3137995

A Dual Model for Restoring Images Corrupted by Mixture of Additive and Multiplicative Noise

CUICUI ZHAO[1], JUN LIU[1], AND JIE ZHANG[2]
[1]Laboratory of Mathematics and Complex Systems (Ministry of Education of China), School of Mathematical Sciences, Beijing Normal University, Beijing 100875, China
[2]Electronics and Computer Science, University of Southampton, Southampton SO17 1BJ, U.K.

Corresponding author: Jun Liu (jliu@bnu.edu.cn)

This work was supported in part by the National Key Research and Development Program of China under Grant 2017YFA0604903, in part by the National Natural Science Foundation of China under Grant 11871035, and in part by the Leverhulme Trust Research Project Grant. The work of Jun Liu was supported by the Super Computing Center of Beijing Normal University.

ABSTRACT A critical challenge in image restoration is the presence of various types of noise. Meanwhile, noise detection is a crucial step in mixed noise removal. This paper tackles the challenge of restoring images corrupted by a mixture of additive Gaussian and multiplicative Gamma noise. In the proposed method, we integrate the noise detection process into a variational model using a dual formulation of a *maximum a posteriori* (MAP) estimator. The variational model consists of a novel adaptive fidelity term and a plugin-and-play regularization term. The fidelity term contains an adaptive weight that can automatically detect the noise types, levels, and pollution ways for each pixel. There is flexibility in choosing a plugin-and-play regularization term. For example, we can use a model-based regularizer or a deep learning-based regularizer. In addition, we present a splitting algorithm to minimize the proposed cost functional. This splitting technique enables us to transfer a mixed noise removing problem to several subproblems, including noise removal and detection. The noise detection process can be iteratively estimated by the proposed algorithm itself. Therefore, in the numerical experiments, the proposed model outperforms the existing Rudin-Osher-Fatemi (ROF), Aubert-Aujol (AA), BM3D, and deep learning-based single type denoiser. Experimental results show that the proposed model can remove noise more efficiently and better preserve details in images. Compared to the existing best-performing single type denoiser, on average, the improvements of PSNR values range from 0.33 dB to 0.81 dB under noise mixture ratios $\alpha = 0.4, 0.6$.

INDEX TERMS Image denoising, mixed noise, regularization, deep learning, dual algorithm.

I. INTRODUCTION

Image denoising is an important research topic in image processing. Its task is to remove as much noise as possible while preserving the original images information. Image denoising is challenging because noise removal is an ill-posed inverse problem. Over the years, image denoising has been widely applied in medical images [1], [2], synthetic aperture radar images [3], [4], and remote sensing images [5], [6]. In the literature, according to different ways of noise pollution, two noise models, the *additive* [7]–[13] and the *multiplicative* [14]–[17] models have been extensively studied.

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

The additive noise (AN) model is formulated as

$$f = u + n, \tag{1}$$

where f, u, n: $\Omega \rightarrow \mathbb{R}$ are mappings from the domain of the images to the intensities of the observed image, original image, and additive noise, respectively. Common additive noise includes white Gaussian noise, uniform noise, and impulse noise. Many methods have been developed to remove additive noise, including variational-based methods [7]–[9], boosting techniques for variational-based methods [10]–[13], wavelet-based methods [18], nonparametric estimation [19], chronological techniques [20], [21], nonlocal methods [22]–[24], and deep learning-based algorithms [25]–[27], to name a few. Amongst these methods, a prominent variational model is the Rudin-Osher-Fatemi (ROF) [7] model. It is worth noticing that there are many adaptive

approaches. Thanh *et al.* [10] proposed an adaptive method based on the high-order total variation. A fast and adaptive regularization was proposed in [11]. More related methods can refer to [12], [13]. He *et al.* [18] proposed an automatical estimation of the noise levels based on maximum a posterior (MAP) estimation and wavelet generic Gaussian. Khmag *et al.* combined the clustered batches of noisy images and hidden Markov models (HMMs) to remove AWGN in [20]. Furthermore, they proposed an adaptive denoising framework based on second-generation wavelet domain using hidden Markov models (SGWD-HMMs) in [21]. BM3D [22] is a well-known nonlocal method and can obtain a good performance to restore images corrupted by the additive Gaussian noise. Recently, CNN deep learning-based method (e.g. [27]) was designed to remove AWGN, and could significantly improve the quality of the reconstructed images.

The multiplicative noise (MN) model is formulated as

$$f = uv, \tag{2}$$

where $v : \Omega \to \mathbb{R}$ denotes multiplicative noise, which follows some standard distribution such as Poisson [28], [29], Gamma [14], and Nakagami distributions [30]. The Aubert-Aujol (AA) variational model [14] is well-known for the case where v follows a Gamma distribution. The AA model is conditionally convex. Although the existence of a minimizer can be proven by the variational method [14], many convex optimization algorithms can not be applicable to the AA model due to its non-convexity. Therefore, many studies have focused on the construct of a convex model. For example, using the exponential transformation through replacing u by e^u, Huang *et al.* [31] and Jin *et al.* [32] converted the AA model to a convex variational model. Moreover, Jin *et al.* [32] provided some theoretical analysis. Ullah *et al.* derived a new data term under the assumption that the noise followed Nakagami distribution instead of Gamma distribution in [30]. Huang *et al.* [33] applied higher-order curvature variation to a convex model, which was superior to others in image edge and corner preserving. More variational models for multiplicative noise removal can refer to TABLE 1. in [33].

However, in reality, the noise type is not necessarily either additive or multiplicative. Rather, it can be a mixture of the same type [34]–[37] or a mixture of these two types [38]–[43]. Tackling these mixture models is even more challenging, and the methods mentioned above cannot straightforwardly solve this problem. Recently, several methods have emerged for this purpose. For a mixture of additive noise (MoAN), new approaches were for reducing a mixture of Gaussian-Gaussian noise [34], [35], and a mixture of Gaussian-impulse noise [35]–[37]. Especially, Wang *et al.* [35] proposed an adaptive algorithm named EM-CNN based on CNN deep learning-based algorithm. The algorithm in [35] combined the previous fidelity term [34] and CNN deep learning-based algorithm [27], where the fidelity term identifies the two different additive noise. It can achieve a desirable quality of the reconstructional images. For a mixture of additive and multiplicative noise (MoAMN), the model is usually formulated as

$$f = u + k_0 n + k_1 uv, \tag{3}$$

where k_0, k_1 are fixed constants, n, v are components of additive noise and multiplicative noise, respectively. Several methods have been developed to deal with a mixture of Poisson-Gaussian noise [38]–[40], to name a few. Thanh *et al.* [40] proposed a model based on total variation and applied a linear combination of log-likelihood functions of Poisson and Gaussian distributions through manual settings. For the removal of mixed Gaussian-Gamma noise, some works have also been raised. Chumchob *et al.* [41] proposed a variational model based on total variation (TV) and a linear combination of the fidelity terms in [7] and [32]. In their model, the parameter that balances the fidelity terms of the additive and multiplicative noise contributions was chosen empirically. Ullah *et al.* [42] proposed a new model with a linear combination of fractional-order total variational (FOTV), FoE image prior, and the data fidelity term in [7]. The parameters that balance the above three terms were also chosen empirically. As a result, the performance of these two variational models depended strongly on manual intervention. A similar method is in [43]. There are other recent reports on mixed Gaussian and Gamma noise removal techniques in [44]–[46].

This paper is devoted to restoring images that are corrupted by a mixture of additive Gaussian and multiplicative Gamma noise. The proposed approach provides a unified framework for a mixture of additive and multiplicative noise (MoAMN). The first challenge is the automatic detection of the noise type, level, and pollution way at each pixel. This paper provides a unified variational model with statistical parameters to discriminate against the distribution and noise level. Operator splitting is adopted to tackle this problem. The corresponding algorithm includes several iterative steps: noise removal, parameter estimation for noise type and level, and denoising adjustment.

The contribution of this paper is threefold:

- This paper proposes a unified variational method to remove MoAMN. The weighting function in the proposed model is adaptive because it can be iteratively determined by the cost function itself. Additionally, the weighting function in our model plays the role of a detector for types and levels of noise. Moreover, this paper provides different interpretations of the weighting function with the EM and soft-max from statistics and convex optimization, respectively.
- The existence of a minimizer and the convergence of the algorithm are proved under some mild conditions.
- The proposed model can integrate TV, BM3D, and DCNN regularizers into its corresponding algorithm. Significantly, the proposed method with DCNN regularization combines the variational models and deep learning using the plugin-and-play method. The idea of the algorithm in this paper can be described as follows. The mixed noise removal problem is decomposed

into three steps: noise detection, noise removal, and image synthesis. The noise removal can be solved by the deep learning method (CNN regularization), and the noise detection and image synthesis are solved by the variational method.

Organization of the Paper: The related works are reviewed in Section 2. Section 3 provides a process of deriving the proposed model, including mathematical assumptions, fidelity construction, and theoretical analysis. An algorithm is designed for the proposed method using splitting methods in Section 4. Section 5 presents some convergence analyses of the proposed algorithm. In Section 6, we present the numerical results. Section 7 shows the conclusion.

II. RELATED WORK

A. METHODS OF REMOVING ADDITIVE NOISE (AN)

In image processing, there are some classical methods to remove additive noise (AN), such as ROF [7], BM3D [22], DCNN [27] and their variants. They all focused on removing additive Gaussian noise. Rudin *et al.* [7] proposed the ROF model

$$\min_u \left\{ \int_\Omega |Du|dx + \frac{\lambda}{2}\int_\Omega (f-u)^2 dx \right\}, \tag{4}$$

where $\int_\Omega |Du|dx$ is the total variation of u, and $\lambda > 0$ is a fixed parameter. The drawback of the ROF model is that it will create some artifacts. BM3D [22] is a nonlocal method that combines wavelet shrinkage and inter-patch correlation. However, the BM3D method may over-smooth the patches that do not have self-similarity properties in the image. Other nonlocal methods can refer to [23], [24]. To improve the quality of the reconstructed images, Zhang *et al.* [27] proposed the IRCNN algorithm that could deal with the model

$$\min \left\{ \ell(\Theta) = \frac{1}{2N}\sum_{i=1}^{N} \|f(\boldsymbol{y}_i;\Theta) - (\boldsymbol{y}_i - \boldsymbol{x}_i)\|_F^2 \right\}, \tag{5}$$

where $\boldsymbol{x}_i, \boldsymbol{y}_i$ are N noisy-clean patch pairs. For space reasons, other methods to remove AWGN are no longer described here.

B. METHODS OF REMOVING MULTIPLICATIVE NOISE (MN)

AA model [14] is a well-known variational method to remove multiplicative Gamma noise. The mathematical expression is

$$\min_u \left\{ \int_\Omega \left(\log u + \frac{f}{u} \right) dx + \lambda \int_\Omega |Du|dx \right\}, \tag{6}$$

where $\lambda > 0$ is a fixed parameter. To convert the AA model to be convex, Huang *et al.* [31] proposed the following model

$$\min_u \left\{ \int_\Omega (u + fe^{-u})dx + \lambda \int_\Omega |Du|dx \right\}. \tag{7}$$

Jin *et al.* [32] provided some theoretical analyses about the model (7). Under the hypothesis of Nakagami noise, Ullah *et al.* [30] changed the data term and proposed a new model as

$$\min_u \left\{ \frac{\alpha_1}{2}\int_\Omega (2u + f^2 e^{-2u})dx + \frac{\alpha_2}{2}\int_\Omega (e^{2u} - 2f^2 u)dx + \int_\Omega |Du|dx \right\}. \tag{8}$$

For page limitation, many other variational methods are omitted here, and one can refer to TABLE 1. in [33].

C. METHODS OF REMOVING A MIXTURE OF ADDITIVE NOISE

Wang *et al.* [35] proposed a model (EM-CNN) to remove a mixture of additive noise (MoAN), which can achieve better performance than the other existing methods like weighted K-SVD [34] and so on. The model combined the EM method and the IRCNN algorithm. The general model can be written as

$$\min_{u,v,\boldsymbol{\Theta},\boldsymbol{w}\in\mathbb{S}} \max_{\mu} \left\{ \mathcal{H}(v,\boldsymbol{\Theta},\boldsymbol{w}) + < \mu, v-u > + \frac{\eta}{2}\|u-v\|_2^2 + \lambda_1 J(u) + \lambda_2 TV(u) \right\}, \tag{9}$$

where $\mathbb{S} = \{\boldsymbol{w}(x) | 0 \le w_k(x) \le 1, \sum_{k=1}^{K} w_k(x) = 1, \forall x \in \Omega\}$; μ is Lagrange multiplier; and η is a penalty parameter. Then they applied the alternating minimization scheme to get iterative formulas. Especially, the iterative formula of u is a problem of removing AWGN. They chose the IRCNN [27] to update u^v. In addition, the expression of $\mathcal{H}$ for the mixture Gaussian model can be expressed as

$$\begin{aligned}\mathcal{H}(u,\boldsymbol{\Theta},\boldsymbol{w}) = {} & \frac{1}{2}\sum_{x\in\Omega}\sum_{k=1}^{K}\frac{(u(x)-f(x))^2}{\sigma_k^2}w_k(x) \\ & - \sum_{x\in\Omega}\sum_{k=1}^{K} w_k(x)\ln r_k + \frac{1}{2}\sum_{x\in\Omega}\sum_{k=1}^{K} w_k \ln\sigma_k^2 \\ & + \sum_{x\in\Omega}\sum_{k=1}^{K} w_k(x)\ln w_k(x),\end{aligned} \tag{10}$$

where $u(x), f(x)$ are clean and noisy images, respectively.

D. METHODS OF REMOVING A MIXTURE OF ADDITIVE AND MULTIPLICATIVE NOISE

Chumchob *et al.* [41] proposed a variational model to remove mixed additive Gaussian and multiplicative Gamma noise. Their mathematical model is

$$\min_{u\in BV(\Omega)} \left\{ \int_\Omega |Du|dx + \frac{\alpha_1}{2}\int_\Omega (f-u)^2 dx + \alpha_2 \int_\Omega (u + fe^{-u})dx \right\}, \tag{11}$$

where $\alpha_1, \alpha_2 > 0$ are two parameters used to balance the fidelity terms of additive and multiplicative noise. These parameters are often chosen empirically. Ullah *et al.* [42]

TABLE 1. Comparison of image denoising methods using different techniques.

Method	Main Methodology	Noise	Highlight	Limitation
ROF [7]	MLE and TV regularizer	AN	Smoothing pixels inside images while keeping the difference between edge pixels	Creating watercolor-like artifacts; The fidelity term is only suitable for removing AN
BM3D [22]	Wavelet shrinkage and inter-patch correlation	AN	Keeping internal images prior while removing as much noise as possible	Over-smoothing the patches that do not have self-similarity property in images
IRCNN [27]	CNN deep learning	AN	Learning external images prior by using CNN; High efficiency with GPU	Over-smoothing some pixels for other kinds of noise
AA [14]	MAP and TV regularizer	MN	A new fidelity for MN; Keeping the difference between edge pixels	Creating watercolor-like artifacts; The fidelity term is just for MN; Conditionally convex
HMW [31]	Replacing u by e^u on AA	MN	Convex and theoretical guarantees	Creating artifacts; Only suitable for MN
EM-CNN [35]	EM algorithm and IRCNN regularizer	MoAN	Detecting the noise type of MoAN adaptively; Learning images prior by using CNN	The fidelity term only for detecting the noise type of MoAN adaptively
CCB [41]	Combination of ROF [7] and HMW [31]; TV regularizer	MoAMN	A new variational model for MoAMN instead of AN and MN	Creating watercolor-like artifacts; Depending strongly on manual intervention
UCKS [42]	Fractional order total variation (FOTV); FoE-image prior; A combination of fidelity terms for AN and MN	MoAMN	Removing much more noise while keeping more images prior; Removing MoAMN instead of AN and MN	Depending strongly on the manual intervention; Three parameters need to be set manually.

proposed another variational model. The model is expressed as

$$\min_u \Big\{ E(u) = \int_\Omega |\nabla^\alpha u| dx + \mu \sum_{i=1}^{N} \theta_i \rho(k_i * u) + \frac{\alpha_1}{2} \int_\Omega (f-u)^2 dx + \frac{\alpha_2}{2} \int_\Omega \left(\log u + \frac{f}{u} \right) dx + \frac{\alpha_3}{2} \int_\Omega (u^2 - 2f^2 \log u) dx \Big\}, \tag{12}$$

Set $u = e^w$. (12) can be written as

$$\min_w \Big\{ E(w) = \int_\Omega |\nabla^\alpha e^w| dx + \mu \sum_{i=1}^{N} \theta_i \rho(k_i * e^w) + \frac{\alpha_1}{2} \int_\Omega (f - e^w)^2 dx + \frac{\alpha_2}{2} \int_\Omega (w + f e^{-w}) dx + \frac{\alpha_3}{2} \int_\Omega (e^{2w} - 2f^2 w) dx \Big\}, \tag{13}$$

where $\mu,\ \alpha,\ \alpha_1,\ \alpha_2,\ \alpha_3 > 0$ are parameters.

In order to see the comparison of methods more intuitively, TABLE 1 summarizes the main methods mentioned above.

In the above models, the parameters that balance the fidelity terms are fixed. Therefore, these models cannot detect the noise type and level. Thus, an inadequate prior to the mixed noise would subvert the very foundation of the models such that they are not able to produce desirable reconstructions. To improve this situation, we propose an adaptive MoAMN denoiser in the next section.

III. THE PROPOSED METHOD

According to the literature survey, the existing variational-based works on MoAMN removal depend strongly on manual intervention. To automatically update noise parameters, an adaptive variational model with selectable regularization is derived for restoring images corrupted by a mixture of additive Gaussian and multiplicative Gamma noise in this section. The adaptivity is accomplished by the weights in the fidelity term, which is derived by adopting a dual formulation of the MAP estimation. The regularization term can be chosen among TV, BM3D, and IRCNN regularizations.

A. THE NOISE MODEL

There are some assumptions in the following:

A1 The images are corrupted by mixed Gaussian-Gamma noise with a ratio of α. In particular, the additive noise n follows Gaussian distribution with mean 0 and variance σ^2. The multiplicative noise v follows Gamma distribution with mean 1 and shape parameter L. The parameters $\alpha,\ \sigma^2$, and L are unknown.

A2 The noise is the realization of a random variable.

A3 The original clean image follows a Gibbs prior distribution.

Throughout the paper, $\mathscr{U}$, $\mathscr{N}$, $\mathscr{V}$, $\mathscr{F}$ denote the random variable of the pixels in the clean image, additive noise, multiplicative noise, and the observed image, respectively. Meanwhile, denote $u,\ n,\ v,\ f$ as the corresponding samples of these random variables. $P_{\mathscr{X}}(x)$ and $p_{\mathscr{X}}(x)$ denote the cumulative distribution and probability density functions of random variable $\mathscr{X}$, respectively.

Let A be the event that the noise is additive and its complement $\bar{\text{A}}$ be the event that the noise is multiplicative. The unknown mixture ratio $\alpha = P(\text{A})$. Then, the forward problem in our model is represented as

$$\mathscr{F} = \begin{cases} \mathscr{U} + \mathscr{N} & \text{when event A occurs,} \\ \mathscr{U}\mathscr{V} & \text{otherwise.} \end{cases} \tag{14}$$

For a better understanding, FIGURE 1 shows three noisy images corrupted by additive Gaussian noise, multiplicative Gamma noise and mixed Gaussian-Gamma noise, separately. It is unknown for any pixel how likely the clean image is to be corrupted by an additive noise or a multiplicative noise. In other words, α is unknown.

(a) Clean (b) Noisy (AG) (c) Noisy (MG) (d) Noisy (MoAMN)

FIGURE 1. From left to right: (a). Clean image, (b). Noisy images with additive Gaussian (AG), (c). Multiplicative Gamma noise (MG) and (d). A mixture of these two kinds of noise (MoAMN).

By assumption A1, the probability density function of additive noise $\mathscr{N}$ at each pixel is

$$p_1(n) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{n^2}{2\sigma^2}}, \tag{15}$$

and the probability density function of multiplicative noise $\mathscr{V}$ at each pixel is

$$p_2(v) = \frac{L^L}{\Gamma(L)} v^{L-1} e^{-Lv} \mathbf{1}_{\{v\geq 0\}}. \tag{16}$$

Here Γ is the well-known Γ function parameterized by L. According to A3, the probability density function of the latent image u is

$$p_{\mathscr{U}}(u) = \frac{1}{T} e^{-\gamma\phi(u)}, \tag{17}$$

where T is the normalized parameter and ϕ is a given potential function. This Gibbs prior is widely used to model natural images, for example, in [14].

B. THE PROPOSED FIDELITY

In this subsection, an adaptive fidelity term is proposed to detect noise types and levels automatically for the mixed Gaussian-Gamma noise. There are two propositions regarding the probability density function of the additive and multiplicative noise in the MAP.

Proposition 3.1 (additive Noise): Assume $\mathscr{F} = \mathscr{U} + \mathscr{N}$ in which $\mathscr{U}$ and $\mathscr{N}$ are mutually independent. Let $p_1(n)$ be the probability density function of $\mathscr{N}$; then,

$$p_{\mathscr{F}|\mathscr{U}}(f|u) = p_1(f-u). \tag{18}$$

Proposition 3.2 (multiplicative Noise, [14]): Assume $\mathscr{F} = \mathscr{U}\mathscr{V}$ in which $\mathscr{U}$ and $\mathscr{V}$ are mutually independent. Let $p_2(v)$ be the probability density function of $\mathscr{V}$; then,

$$p_{\mathscr{F}|\mathscr{U}}(f|u) = \frac{1}{u} \cdot p_2\left(\frac{f}{u}\right). \tag{19}$$

Let m, n be the row and column numbers of a clean image, respectively. The total number of pixels is denoted by $N = m \times n$. The original image $\boldsymbol{u} = (u_0, u_1, \ldots, u_{N-1})$ and the observed image $\boldsymbol{f} = (f_0, f_2, \ldots, f_{N-1})$ are realizations of the random variables $\boldsymbol{\mathscr{U}} = (\mathscr{U}_0, \mathscr{U}_1, \ldots, \mathscr{U}_{N-1})$ and $\boldsymbol{\mathscr{F}} = (\mathscr{F}_0, \mathscr{F}_1, \ldots, \mathscr{F}_{N-1})$, respectively. The corresponding inverse problem is to maximize $P(\boldsymbol{\mathscr{U}}|\boldsymbol{\mathscr{F}})$ for a given $\boldsymbol{f}$ in terms of the MAP.

By the Bayes' law and logarithm operation, the MAP can be converted to the following maximization problem

$$\max \quad \ln P(\boldsymbol{\mathscr{F}}|\boldsymbol{\mathscr{U}}) + \ln P(\boldsymbol{\mathscr{U}}). \tag{20}$$

Because the noise values of the pixels are assumed to be mutually independent, we obtain

$$P(\boldsymbol{\mathscr{F}}|\boldsymbol{\mathscr{U}}) = \prod_{i=0}^{N-1} P(\mathscr{F}_i|\mathscr{U}_i) \tag{21}$$

Furthermore, the prior is

$$P(\boldsymbol{\mathscr{U}}) = \prod_{c\in\mathcal{C}} P(\mathscr{U}_c). \tag{22}$$

where $\mathcal{C}$ is the numbers of clicks in the graph representation of the prior. By plugging (21) and (22) into (20), the problem (20) is equal to the following minimization problem

$$\min_{\boldsymbol{u},\boldsymbol{\Theta}} \left\{ -\sum_{i=0}^{N-1} \ln p_{\mathscr{F}|\mathscr{U}}(f_i|u_i) - \sum_{c\in\mathcal{C}} \ln p_{\mathscr{U}}(u_c) \right\}, \tag{23}$$

where $\boldsymbol{\Theta} = (\alpha, \sigma^2, L)$ is a parameter vector of the distribution for noise. According to the law of total probability and Propositions 3.1 and 3.2, the problem (23) can be rewritten as

$$\min_{\boldsymbol{u},\boldsymbol{\Theta}} \Bigg\{ \mathcal{H}(\boldsymbol{u},\boldsymbol{\Theta}) = -\sum_{i=0}^{N-1} \ln \Big[\alpha p_1(f_i - u_i) + (1-\alpha)\frac{1}{u_i} p_2\left(\frac{f_i}{u_i}\right) \Big] - \sum_{c\in\mathcal{C}} \ln p_U(u_c) \Bigg\}. \tag{24}$$

Note that its first term is an ln-sum term, which is difficult to handle in the minimization problem. Next, we will expand it by employing the following result reported in [47], which is equivalent to the EM algorithm.

Proposition 3.3 [47]: Given a positive matrix with elements $\delta_{ij} > 0$, then

$$-\sum_{i=0}^{N-1} \ln \sum_{j=1}^{2} \delta_{ij} = \min_{w\in\Delta^+} \left\{ -\sum_{i=1}^{N-1}\sum_{j=1}^{2} w_{ij} \ln \delta_{ij} + \sum_{i=0}^{N-1}\sum_{j=1}^{2} w_{ij} \ln w_{ij} \right\}, \tag{25}$$

where $\Delta^+ = \{\boldsymbol{w}|0 \le w_{ij} \le 1, \sum_{j=1}^{2} w_{ij} = 1, \forall i = 0, 1, \cdots, N-1\}$.

By setting $\delta_{i1} = \alpha p_1(f_i - u_i)$ and $\delta_{i2} = (1-\alpha)\frac{1}{u_i} p_2\left(\frac{f_i}{u_i}\right)$ in Proposition 3.3, the problem (24) can be converted to the following form

$$\min_{\boldsymbol{u},\boldsymbol{\Theta},\boldsymbol{w}\in\Delta^+} \Big\{ \mathcal{J}(\boldsymbol{u},\boldsymbol{\Theta},\boldsymbol{w}) = -\sum_{i=0}^{N-1} \Big[w_{i1} \ln\left(\alpha p_1(f_i - u_i)\right) + w_{i2} \ln\left((1-\alpha)\frac{1}{u_i} p_2\left(\frac{f_i}{u_i}\right)\right)\Big] + \sum_{i=0}^{N-1}\sum_{j=1}^{2} w_{ij} \ln w_{ij} - \sum_{c\in\mathcal{C}} \ln p_U(u_c) \Big\}, \tag{26}$$

Plug the probability density functions (15), (16) and (17) into the above formula. For ease of notation, denote $w_{i1} = w_i$ and $w_{i2} = 1 - w_{i1} = 1 - w_i$. With some simplifications, the minimization problem can be written as

$$\begin{aligned}\min_{\boldsymbol{u},\boldsymbol{\Theta},\boldsymbol{w}\in\Delta^+} \Big\{ \mathcal{J}(\boldsymbol{u},\boldsymbol{\Theta},\boldsymbol{w}) = &\sum_{i=0}^{N-1} \Big[\frac{w_i (f_i - u_i)^2}{2\sigma^2} + (1-w_i)\big(L \ln u_i + \frac{L f_i}{u_i}\big)\Big] \\ &+ \sum_{i=0}^{N-1} w_i \Big(\frac{1}{2}\ln(2\pi) + \frac{1}{2}\ln\sigma^2 - \ln\alpha\Big) \\ &+ \sum_{i=0}^{N-1}(1-w_i)\Big[\ln\Gamma(L) - (L-1)\ln f_i \\ &- \ln(1-\alpha) - L\ln L\Big] + \sum_{i=0}^{N-1}[w_i \ln w_i \\ &+ (1-w_i)\ln(1-w_i)] - \sum_{c\in\mathcal{C}} \gamma\phi(u_c)\Big\}. \end{aligned} \tag{27}$$

It should be noted that there are some differences between the discrete model (27) in this paper and that in [35], which is recalled in part C of Section II. The fidelity terms in (27) and the formula (10) in [35] are different. The reason for the difference is that the mathematical problems are totally different. Specifically, the types of mixed noise in this paper are addition and multiplication, while those in [26] are only addition. Furthermore, the differences of those two models also make the corresponding algorithms different, shown in the next section.

The continuous version of the functional $\mathcal{J}$ in the problem (27) can be written as

$$\begin{aligned}\mathcal{J}(u,\boldsymbol{\Theta},w) = &\int_\Omega \frac{w(x)(f(x)-u(x))^2}{2\sigma^2}dx \\ &+ \int_\Omega \left(L\ln u(x) + \frac{L f(x)}{u(x)}\right)(1-w(x))dx \\ &+ \left(\frac{1}{2}\ln(2\pi) + \frac{1}{2}\ln\sigma^2 - \ln\alpha\right)\int_\Omega w(x)dx \\ &+ \int_\Omega \Big(\ln\Gamma(L) - \ln(1-\alpha) - L\ln L \\ &-(L-1)\ln f(x)\Big)(1-w(x))dx \\ &+ \int_\Omega (1-w(x))\ln(1-w(x))dx \\ &+ \int_\Omega w(x)\ln w(x)dx + \gamma\int_\Omega \phi(u(x))dx \end{aligned} \tag{28}$$

Therefore, the proposed model for removing the mixture of additive and multiplicative noise (MoAMN) can be written as

$$\min_{u\in\mathbb{X},\ \boldsymbol{\Theta}\in\mathbb{K},\ w\in\mathbb{W}} \mathcal{J}(u,\boldsymbol{\Theta},w), \tag{29}$$

where $\mathbb{K} = \{\boldsymbol{\Theta} = (\alpha,\sigma^2,L)|0 \le \alpha \le 1, 0 < \sigma^2_{min} \le \sigma^2, 0 < L_{min} \le L \le L_{max}\}$, $\mathbb{W} = \{w : \Omega \to \mathbb{R}|0 \le w(x) \le 1, \forall x \in \Omega\}$, and $\mathbb{X}$ is a function space.

C. SOFT-MAX INTERPRETATION OF THE PROPOSED FIDELITY

Note that the weight $\boldsymbol{w}$ in the cost of the problem (27) can adaptively detect the noise type and level according to the noise parameters $\boldsymbol{\Theta}$. This section is devoted to interpreting the weight $\boldsymbol{w}$ from the viewpoint of convex optimization. Although $\mathcal{J}$ is not convex with respect to $(\boldsymbol{u},\boldsymbol{\Theta},\boldsymbol{w})$, it is convex with respect to the single variable $\boldsymbol{u}$ or $\boldsymbol{w}$. We show that this non-convex problem can be derived by a dual method of convex functions. There are some definitions and propositions.

Definition 3.1 (Soft-Max): Given a vector $\boldsymbol{y} = (y_1, y_2, .. y_M)$, for all $\varepsilon > 0$, the soft-max operator is defined by

$$\max\nolimits_\varepsilon(\boldsymbol{y}) := \varepsilon \ln \sum_{j=1}^{M} e^{\frac{y_j}{\varepsilon}}. \tag{30}$$

It is easy to check that $\lim_{\varepsilon\to 0} \max_\varepsilon(\boldsymbol{y}) = \max\{\boldsymbol{y}\}$.

Proposition 3.4: Let $F_\varepsilon(\boldsymbol{y}) = \max_\varepsilon(\boldsymbol{y})$. Then for any fixed $\varepsilon > 0$, $F_\varepsilon(\boldsymbol{y})$ is convex with respect to $\boldsymbol{y}$.

The proof of Proposition 3.4 is deferred to Subsection S1 in the submitted supplementary material.

Definition 3.2 (Fenchel-Legendre Transformation): F^* is the Fenchel-Legendre transformation of F, defined by

$$F^*(\boldsymbol{w}) := \max_{\boldsymbol{y}}\{<\boldsymbol{y},\boldsymbol{w}> - F(\boldsymbol{y})\}. \tag{31}$$

Proposition 3.5: [48] A function $F : \mathbb{R}^M \to \mathbb{R}\cup\{+\infty\}$ is convex and lower semi-continuous if and only if $F = F^{**}$.

Proposition 3.6: For any fixed $\varepsilon > 0$, if F^*_ε is the Fenchel-Legendre transformation of the soft-max function F_ε, then

$$\begin{aligned} F^*_\varepsilon(\boldsymbol{w}) &= \max_{\boldsymbol{y}}\{<\boldsymbol{y},\boldsymbol{w}> - F_\varepsilon(\boldsymbol{y})\} \\ &= \begin{cases} \varepsilon\sum_{j=1}^{M} w_j \ln w_j, & \boldsymbol{w}\in\Delta^+, \\ +\infty, & else. \end{cases} \end{aligned} \tag{32}$$

where $\Delta^+ = \{\boldsymbol{w} = (w_1, w_2, \ldots, w_M) : 0 \le w_j \le 1, \sum_{j=1}^{M} w_j = 1\}$, and thus

$$F_\varepsilon(\boldsymbol{y}) = F_\varepsilon^{**}(\boldsymbol{y}) = \max_{\boldsymbol{w}\in\Delta^+} \left\{ < \boldsymbol{w}, \boldsymbol{y} > -\varepsilon \sum_{j=1}^{M} w_j \ln w_j \right\}, \tag{33}$$

where F_ε^{**} is the Fenchel-Legendre transformation of F_ε^*.

The proof of Proposition 3.6 is deferred to Subsection S2 in the submitted supplementary material.

The problem (26) can be derived from (24) by applying the Proposition 3.6. Set $M = 2$, $\boldsymbol{y} = \boldsymbol{y}_i = (y_{i1}, y_{i2})$ with $y_{i1} = \varepsilon \ln(\alpha p_1(f_i - u_i))$, $y_{i2} = \varepsilon \ln\left[(1-\alpha)\frac{1}{u_i} p_2(\frac{f_i}{u_i})\right]$, then

$$\begin{aligned}
(24) &\Leftrightarrow \min_{\boldsymbol{u},\boldsymbol{\Theta}} \left\{ -\sum_{i=0}^{N-1} \ln \sum_{j=1}^{2} e^{\frac{y_{ij}}{\varepsilon}} - \sum_{i=0}^{N-1} \ln p_U(u_i) \right\} \\
&\overset{(33)}{\Leftrightarrow} \min_{\boldsymbol{u},\boldsymbol{\Theta}} \Big\{ -\sum_{i=0}^{N-1} \frac{1}{\varepsilon} \max_{\boldsymbol{w}_i\in\Delta^+} \Big\{ < \boldsymbol{w}_i, \boldsymbol{y}_i > \\
&\quad -\varepsilon \sum_{j=1}^{2} w_{ij} \ln w_{ij} \Big\} - \sum_{i=0}^{N-1} \ln p_U(u_i) \Big\} \\
&\Leftrightarrow \min_{\boldsymbol{u},\boldsymbol{\Theta}} \Big\{ \sum_{i=0}^{N-1} \frac{1}{\varepsilon} \min_{\boldsymbol{w}_i\in\Delta^+} \Big\{ - < \boldsymbol{w}_i, \mathbf{y}_i > \\
&\quad +\varepsilon \sum_{j=1}^{2} w_{ij} \ln w_{ij} \Big\} - \sum_{i=0}^{N-1} \ln p_U(u_i) \Big\} \Leftrightarrow (26).
\end{aligned}$$

In this procedure, the weight $\boldsymbol{w}$ is a dual variable that can be used to classify noise.

D. REGULARIZATION

There are many choices for the regularization term $\phi(u)$. For example, we can use p-norm ($p \geqslant 1$), p-pseudo-norm ($0 < p < 1$) or 0 pseudo-norm [49] of the gradient of u, i.e. $\phi(u) = \int_\Omega |\nabla u(x)|^p dx$. To preserve textures, one may set $\phi(u)$ as detail-preserving [50] regularization or image patches-based nonlocal operators such as nonlocal TV [23], BM3D [22], and low rank [24]. Moreover, to use the nonlinear prior in nature images, CNN-based methods [27] may be integrated. This can be done by considering the variational of ϕ with respect to u as a nonlinear convolution neural network. In this paper, we will use TV [7], BM3D [22], and CNN [27] regularizations for comparison. For the model-based regularizers such as TV and BM3D, we will prove the existence of a minimizer for the proposed model. For the data-driving regularizer such as CNN, we present some numerical results to demonstrate its performance.

E. EXISTENCE OF A MINIMIZER

In this section, we prove the existence of a minimizer of (28) by setting $\phi(u) = TV(u)$. In this case, $\mathbb{X}$ is the well-known BV space. Without loss of generality, let $\gamma = 1$. Denote $\mathbb{X} = \{u \in BV(\Omega) \mid u(x) \geq 0 \text{ for any } x \in \Omega\}$.

Theorem 3.1: Assume that $f \in L^\infty(\Omega)$ and $0 < \inf_\Omega f, \sup_\Omega f < +\infty$, then there exists a minimizer of the problem (29) in $\mathbb{I} = \{(u, \boldsymbol{\Theta}, w) \mid u \in \mathbb{X}, \boldsymbol{\Theta} \in \mathbb{K}, w \in \mathbb{W}\}$.

We defer the proof of the theorem to Subsection S3 in the submitted supplementary material.

In conclusion, there are three points to highlight in the proposed model. Firstly, we propose a new adaptive fidelity term to detect the noise types and levels automatically. Secondly, compared with TV and FOTV regularizations in [41]–[43], the proposed model with DCNN regularization can take advantage of the CNN deep learning algorithm to eminently learn images prior. That is, the adaptive fidelity term and the DCNN regularization work together to remove as much noise as possible while retaining much images' prior information. Thirdly, we prove the existence of a minimizer in Theorem 3.1.

A block diagram in FIGURE 2 is shown in advance to demonstrate the main structure of the proposed model clearly. The related algorithm is introduced in the next section for more details.

IV. RELATED ALGORITHM

This section presents an algorithm for the variational model (29) using the splitting technique. The discretization schemes for the variational terms are standard. For integration, we use the rectangular formula. For the derivative that appeared in the TV regularization, we use a central difference scheme. For better representation, set $\boldsymbol{a} = \{a_0, a_1, \ldots, a_{N-1}\}$, $\boldsymbol{b} = \{b_0, b_1, \ldots, b_{N-1}\}$, $\boldsymbol{c} = \{c_0, c_1, \ldots, c_{N-1}\}$ be vectors in $\mathbb{R}^N$. Denote

$$\|\boldsymbol{a}\|_1 = \sum_{i=0}^{N-1} |a_i|, \quad \|\boldsymbol{a}\|_2^2 = \sum_{i=0}^{N-1} |a_i|^2, \quad \|\boldsymbol{a}\|_{2,\boldsymbol{c}}^2 = \sum_{i=0}^{N-1} c_i |a_i|^2,$$

$$< \boldsymbol{a}, \boldsymbol{b} >= \sum_{i=0}^{N-1} a_i b_i.$$

Using these notations, (29) can be rewritten as

$$\min_{\boldsymbol{u},\ \boldsymbol{\Theta},\ \boldsymbol{w}} \mathcal{J}(\boldsymbol{u},\ \boldsymbol{\Theta},\ \boldsymbol{w}), \tag{34}$$

where

$$\begin{aligned}
&\mathcal{J}(\boldsymbol{u}, \boldsymbol{\Theta}, \boldsymbol{w}) \\
&= \frac{1}{2\sigma^2} \|\boldsymbol{u} - \boldsymbol{f}\|_{2,\boldsymbol{w}}^2 + L \left\langle \ln \boldsymbol{u} + \frac{\boldsymbol{f}}{\boldsymbol{u}}, \mathbf{1} - \boldsymbol{w} \right\rangle \\
&\quad + \left(-\ln \alpha + \frac{1}{2} \ln \sigma^2 + \frac{1}{2} \ln(2\pi) \right) < \mathbf{1}, \boldsymbol{w} > \\
&\quad + [-\ln(1-\alpha) - L \ln L + \ln \Gamma(L)] < \mathbf{1}, \mathbf{1} - \boldsymbol{w} > \\
&\quad - (L-1) < \ln \boldsymbol{f}, \mathbf{1} - \boldsymbol{w} > + < \boldsymbol{w}, \ln \boldsymbol{w} > \\
&\quad + < \mathbf{1} - \boldsymbol{w}, \ln(\mathbf{1} - \boldsymbol{w}) > + \lambda \phi(\boldsymbol{u}).
\end{aligned}$$

The above problem can be decomposed into three subproblems according to the alternating minimization algorithm. Denote $\nu = 0, 1, \ldots$ as an iteration step.

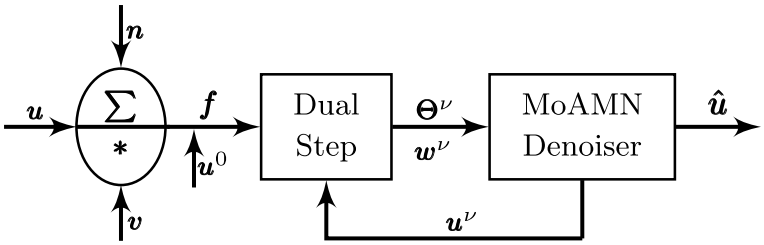


FIGURE 2. Block diagram of the proposed model, where ν is the iteration step. The final calculated image $\hat{\boldsymbol{u}}$ outputs until convergence.

Subproblem 1:

$$\begin{aligned}\boldsymbol{u}^{\nu+1} &= \arg\min_{\boldsymbol{u}} \mathcal{J}(\boldsymbol{u}, \boldsymbol{\Theta}^{\nu}, \boldsymbol{w}^{\nu}) \\ &= \arg\min_{\boldsymbol{u}} \left\{ \frac{1}{2(\sigma^2)^{\nu}} \|\boldsymbol{u} - \boldsymbol{f}\|^2_{2,\boldsymbol{w}^{\nu}} \right. \\ &\quad \left. + L^{\nu} \left\langle \ln \boldsymbol{u} + \frac{\boldsymbol{f}}{\boldsymbol{u}}, \mathbf{1} - \boldsymbol{w}^{\nu} \right\rangle + \lambda\phi(\boldsymbol{u}) \right\}. \end{aligned} \tag{35}$$

Subproblem 2:

$$\begin{aligned}&\boldsymbol{\Theta}^{\nu+1} \\ &= \arg\min_{\boldsymbol{\Theta}} \mathcal{J}(\boldsymbol{u}^{\nu+1}, \boldsymbol{\Theta}, \boldsymbol{w}^{\nu}) \\ &= \arg\min_{\boldsymbol{\Theta}} \left\{ \frac{1}{2\sigma^2} \|\boldsymbol{u}^{\nu+1} - \boldsymbol{f}\|^2_{2,\boldsymbol{w}^{\nu}} + L < \ln \boldsymbol{u}^{\nu+1} \right. \\ &\quad + \frac{\boldsymbol{f}}{\boldsymbol{u}^{\nu+1}}, \mathbf{1} - \boldsymbol{w}^{\nu} > + \left(- \ln \alpha + \frac{1}{2} \ln \sigma^2 + \frac{1}{2} \ln(2\pi) \right) \\ &\quad < \mathbf{1}, \boldsymbol{w}^{\nu} > + [-\ln(1-\alpha) - L \ln L + \ln \Gamma(L)] \\ &\quad \left. < \mathbf{1}, \mathbf{1} - \boldsymbol{w}^{\nu} > - (L-1) < \ln \boldsymbol{f}, \mathbf{1} - \boldsymbol{w}^{\nu} > \right\}. \end{aligned} \tag{36}$$

Subproblem 3:

$$\begin{aligned}&\boldsymbol{w}^{\nu+1} \\ &= \arg\min_{\boldsymbol{w}} \mathcal{J}(\boldsymbol{u}^{\nu+1}, \boldsymbol{\Theta}^{\nu+1}, \boldsymbol{w}) \\ &= \arg\min_{\boldsymbol{w}} \left\{ \frac{1}{2(\sigma^2)^{\nu+1}} \|\boldsymbol{u}^{\nu+1} - \boldsymbol{f}\|^2_{2,\boldsymbol{w}} \right. \\ &\quad - L^{\nu+1} \left\langle \ln \boldsymbol{u}^{\nu+1} + \frac{\boldsymbol{f}}{\boldsymbol{u}^{\nu+1}}, \boldsymbol{w} \right\rangle \\ &\quad + \left(- \ln \alpha^{\nu+1} + \frac{1}{2} \ln(\sigma^2)^{\nu+1} + \frac{1}{2} \ln(2\pi) \right) < \mathbf{1}, \boldsymbol{w} > \\ &\quad + [\ln(1-\alpha^{\nu+1}) + L^{\nu+1} \ln L^{\nu+1} - \ln \Gamma(L^{\nu+1})] \\ &\quad < \mathbf{1}, \boldsymbol{w} > + (L^{\nu+1} - 1) < \ln \boldsymbol{f}, \boldsymbol{w} > \\ &\quad \left. + < \boldsymbol{w}, \ln \boldsymbol{w} > + < \mathbf{1} - \boldsymbol{w}, \ln(\mathbf{1} - \boldsymbol{w}) > \right\}. \end{aligned} \tag{37}$$

Note that for giving $\boldsymbol{w}^{\nu}, \boldsymbol{\Theta}^{\nu}$, Subproblem 1 is for updating the clean image $\boldsymbol{u}^{\nu+1}$. Under this circumstance, it can be regarded as a MoAMN denoiser. By employing the augmented Lagrange method, Subproblem 1 can be transformed into the following problem

$$\begin{aligned}\min_{\boldsymbol{u},\boldsymbol{d}} \left\{ \frac{1}{2(\sigma^2)^{\nu}} \|\boldsymbol{d} - \boldsymbol{f}\|^2_{2,\boldsymbol{w}^{\nu}} + L^{\nu} \left\langle \ln \boldsymbol{d} + \frac{\boldsymbol{f}}{\boldsymbol{d}}, \mathbf{1} - \boldsymbol{w}^{\nu} \right\rangle \right. \\ \left. + \lambda\phi(\boldsymbol{u}) \right\} \quad s.t \ \ \boldsymbol{u} = \boldsymbol{d}. \end{aligned}$$

The corresponding augmented Lagrange function is

$$\begin{aligned}\mathcal{L}_r(\boldsymbol{u}, \boldsymbol{d}; \boldsymbol{p}) = &\frac{1}{2(\sigma^2)^{\nu}} \|\boldsymbol{d} - \boldsymbol{f}\|^2_{2,\boldsymbol{w}^{\nu}} \\ &+ L^{\nu} \left\langle \ln \boldsymbol{d} + \frac{\boldsymbol{f}}{\boldsymbol{d}}, \mathbf{1} - \boldsymbol{w}^{\nu} \right\rangle \\ &+ \lambda\phi(\boldsymbol{u}) + < \boldsymbol{p}, \boldsymbol{u} - \boldsymbol{d} > \\ &+ \frac{r}{2} \|\boldsymbol{u} - \boldsymbol{d}\|_2^2, \end{aligned} \tag{38}$$

where $\mathbf{p}$ is a Lagrange multiplier and $r > 0$ is a penalty parameter. By the ALM algorithm, problem (38) can be solved by the following inner iterative scheme

$$\begin{cases} (\boldsymbol{u}^{\nu_1+1}, \boldsymbol{d}^{\nu_1+1}) = \arg\min_{\boldsymbol{u},\boldsymbol{d}} \mathcal{L}_r(\boldsymbol{u}, \boldsymbol{d}; \boldsymbol{p}^{\nu_1}), \\ \boldsymbol{p}^{\nu_1+1} = \boldsymbol{p}^{\nu_1} + \tau(\boldsymbol{u}^{\nu_1+1} - \boldsymbol{d}^{\nu_1+1}). \end{cases} \tag{39}$$

where $\tau > 0$ is a constant, and $\nu_1 = 0, 1, 2 \cdots$ is an inner iteration number. By applying ADMM, (39) can be calculated by the following iterative scheme

$$\begin{cases} \boldsymbol{u}^{\nu_1+1} = \arg\min_{\boldsymbol{u}} \mathcal{L}_r(\boldsymbol{u}, \boldsymbol{d}^{\nu_1}; \boldsymbol{p}^{\nu_1}), \\ \boldsymbol{d}^{\nu_1+1} = \arg\min_{\boldsymbol{d}} \mathcal{L}_r(\boldsymbol{u}^{\nu_1+1}, \boldsymbol{d}; \boldsymbol{p}^{\nu_1}), \\ \boldsymbol{p}^{\nu_1+1} = \boldsymbol{p}^{\nu_1} + \tau(\boldsymbol{u}^{\nu_1+1} - \boldsymbol{d}^{\nu_1+1}). \end{cases} \tag{40}$$

For $\boldsymbol{u}^{\nu_1+1}$,

$$\begin{aligned}\boldsymbol{u}^{\nu_1+1} = \arg\min_{\boldsymbol{u}} \left\{ \lambda\phi(\boldsymbol{u}) + < \boldsymbol{p}^{\nu_1}, \boldsymbol{u} > \right. \\ \left. + \frac{r}{2} \|\boldsymbol{u} - \boldsymbol{d}^{\nu_1}\|_2^2 \right\}. \end{aligned} \tag{41}$$

By simple calculation, (41) can be rewritten as

$$\boldsymbol{u}^{\nu_1+1} = \arg\min_{\boldsymbol{u}} \left\{ \lambda\phi(\boldsymbol{u}) + \frac{r}{2} \|\boldsymbol{u} - \boldsymbol{g}^{\nu_1}\|_2^2 \right\}, \tag{42}$$

where $\boldsymbol{g}^{\nu_1} = \boldsymbol{d}^{\nu_1} - \dfrac{\boldsymbol{p}^{\nu_1}}{r}$. Note that this is a Gaussian denoiser so that we can apply one of ROF [7], BM3D [22], and

IRCNN [27] algorithms in the proposed method. For $\boldsymbol{d}^{\nu_1+1}$,

$$\begin{aligned}\boldsymbol{d}^{\nu_1+1} = \arg\min_{\boldsymbol{d}} \Big\{ & \frac{1}{2(\sigma^2)^{\nu}} \|\boldsymbol{d}-\boldsymbol{f}\|^2_{2,\boldsymbol{w}^{\nu}} \\ & + L^{\nu} \left\langle \ln \boldsymbol{d} + \frac{\boldsymbol{f}}{\boldsymbol{d}}, \mathbf{1}-\boldsymbol{w}^{\nu} \right\rangle \\ & + <\boldsymbol{p}, \boldsymbol{u}-\boldsymbol{d}> + \frac{r}{2}\|\boldsymbol{u}-\boldsymbol{d}\|_2^2 \Big\}. \end{aligned} \tag{43}$$

We only need to solve a cubic equation

$$\frac{(\boldsymbol{d}-\boldsymbol{f})\boldsymbol{w}^{\nu}}{(\sigma^2)^{\nu}} + L^{\nu}\left(\frac{\mathbf{1}}{\boldsymbol{d}} - \frac{\boldsymbol{f}}{\boldsymbol{d}^2}\right)(\mathbf{1}-\boldsymbol{w}^{\nu}) - \boldsymbol{p}^{\nu_1} + r(\boldsymbol{d}-\boldsymbol{u}^{\nu_1+1}) = 0. \tag{44}$$

That is,

$$\begin{aligned}\left(\frac{\boldsymbol{w}^{\nu}}{(\sigma^2)^{\nu}} + rI\right)\boldsymbol{d}^3 - \left(\frac{\boldsymbol{f}\boldsymbol{w}^{\nu}}{(\sigma^2)^{\nu}} + \boldsymbol{p}^{\nu_1} + r\boldsymbol{u}^{\nu_1+1}\right)\boldsymbol{d}^2 \\ + L^{\nu}(\mathbf{1}-\boldsymbol{w}^{\nu})\boldsymbol{d} - L^{\nu}(\mathbf{1}-\boldsymbol{w}^{\nu})\boldsymbol{f} = 0. \end{aligned} \tag{45}$$

Once there is a converged $\boldsymbol{u}^{\nu_1+1}$ with the inner iteration scheme (40), we set $\boldsymbol{u}^{\nu+1} = \boldsymbol{u}^{\nu_1+1}$ to get the solution to the $\boldsymbol{u}$-subproblem.

Updating $\boldsymbol{w}^{\nu+1}, \boldsymbol{\Theta}^{\nu+1}$ depends on the fidelity term, which is mainly derived by the dual method. So the process of updating $\boldsymbol{w}^{\nu+1}, \boldsymbol{\Theta}^{\nu+1}$ can be called a dual process.

Now, for Subproblem 2, by using the first-order optimization condition of α and σ^2, the closed-form solutions for α and σ^2 can be easily obtained as:

$$\alpha^{\nu+1} = \frac{<\mathbf{1}, \boldsymbol{w}^{\nu}>}{|\Omega|}, \tag{46}$$

$$(\sigma^2)^{\nu+1} = \frac{\|\boldsymbol{u}^{\nu+1}-\boldsymbol{f}\|^2_{2,\boldsymbol{w}^{\nu}}}{<\mathbf{1}, \boldsymbol{w}^{\nu}>}. \tag{47}$$

Generally, it is difficult to calculate the exact parameter L in the Γ distribution. However, Subproblem 2 only requires an estimation of the parameter in the Γ distribution by given $\boldsymbol{w}^{\nu}$. Thus, we can derive $L^{\nu+1}$ by employing the parameter estimation method for the Γ distribution [51].

$$\begin{aligned}L^{\nu+1} = \arg\min_{L} \Big\{ & L\left\langle \ln \boldsymbol{u}^{\nu+1} + \frac{\boldsymbol{f}}{\boldsymbol{u}^{\nu+1}}, \mathbf{1}-\boldsymbol{w}^{\nu} \right\rangle \\ & + (\ln\Gamma(L) - L\ln L) <\mathbf{1}, \mathbf{1}-\boldsymbol{w}^{\nu}> \\ & - (L-1) <\ln \boldsymbol{f}, \mathbf{1}-\boldsymbol{w}^{\nu}> \Big\}. \end{aligned}$$

By the first-order optimization condition of L, the above problem can be converted to

$$\begin{aligned}\left\langle \ln\frac{\boldsymbol{u}^{\nu+1}}{\boldsymbol{f}} + \frac{\boldsymbol{f}}{\boldsymbol{u}^{\nu+1}}, \mathbf{1}-\boldsymbol{w}^{\nu} \right\rangle \\ + (\psi(L) - \ln L - 1) <\mathbf{1}, \mathbf{1}-\boldsymbol{w}^{\nu}> = 0, \end{aligned}$$

where $\psi(L) := \frac{d}{dL}\ln\Gamma(L)$. This equation can be rewritten as

$$\ln L - \psi(L) = M^{\nu},$$

where

$$M^{\nu} := \frac{\left\langle \ln\frac{\boldsymbol{u}^{\nu+1}}{\boldsymbol{f}} + \frac{\boldsymbol{f}}{\boldsymbol{u}^{\nu+1}} - \mathbf{1}, \mathbf{1}-\boldsymbol{w}^{\nu} \right\rangle}{<\mathbf{1}, \mathbf{1}-\boldsymbol{w}^{\nu}>}.$$

According to [51], [52], we can approximate $L^{\nu+1}$ as follows

$$L^{\nu+1} \approx \frac{3 - M^{\nu} + \sqrt{(M^{\nu}-3)^2 + 24M^{\nu}}}{12M^{\nu}}, \tag{48}$$

which is within 1.5% confidence.

Subproblem 3 also has a closed-form solution

$$\boldsymbol{w}^{\nu+1} = \frac{\alpha^{\nu+1}p_1(\boldsymbol{v}^{\nu+1})}{\alpha^{\nu+1}p_1(\boldsymbol{v}^{\nu+1}) + (1-\alpha^{\nu+1})(\boldsymbol{u}^{\nu+1})^{-1}p_2(\boldsymbol{v}^{\nu+1})}, \tag{49}$$

where

$$p_1(\boldsymbol{v}^{\nu+1}) = \frac{1}{\sqrt{2\pi(\sigma^2)^{\nu+1}}} \exp\left(-\frac{(\boldsymbol{u}^{\nu+1}-\boldsymbol{f})^2}{2(\sigma^2)^{\nu+1}}\right),$$

$$p_2(\boldsymbol{v}^{\nu+1}) = \frac{(L^{\nu+1})^{L^{\nu+1}}}{\Gamma(L^{\nu+1})} \frac{\boldsymbol{f}^{L^{\nu+1}-1}}{(\boldsymbol{u}^{\nu+1})^{L^{\nu+1}}} \exp\left(-\frac{L^{\nu+1}\boldsymbol{f}}{\boldsymbol{u}^{\nu+1}}\right).$$

Above all, the proposed algorithm consists of a MoAMN denoiser and a dual process that can update noise parameters and weights with iteration steps. So the proposed method can be named an adaptive MoAMN denoiser.

Finally, we summarize the proposed method and present it as Algorithm 1 below. A flowchart of Algorithm 1 is shown in FIGURE 3.

Algorithm 1 An Adaptive MoAMN Denoiser

1: Set initial values. Given $\boldsymbol{u}^0 = \boldsymbol{f}$, $\boldsymbol{\theta}^0$. Calculate $\boldsymbol{w}^0$ through (49). Set $\nu = 0$.

2: Denoising step. Update $\boldsymbol{u}^{\nu+1}$ by calculating the iteration formulation (40) until convergence. This inner iteration mainly includes three sub-steps:

a. Smoothness: Gaussian denoiser (42), which is chosen directly among ROF [7], BM3D [22], and IRCNN [27].

b. Synthesis: choosing suitable fidelity by (44) based on the updated noise parameters and weight. The parameter r is set appropriately to adjust the level of correction of $\boldsymbol{u}$.

c. Balance: Updating $\boldsymbol{p}^{\nu+1}$ to balance $\boldsymbol{u}$ and $\boldsymbol{d}$ by comp- uting the second formulation of (39) with $\tau = r$. 3: Parameters estimation. Update $\alpha^{\nu+1}, (\sigma^2)^{\nu+1}, L^{\nu+1}$ by calculating (46), (47), (48).

4: Noise classification. Update noise detection function $\boldsymbol{w}^{\nu+1}$ by calculating (49).

5: Convergence checking. If $\frac{||\boldsymbol{u}^{\nu+1}-\boldsymbol{u}^{\nu}||^2}{||\boldsymbol{u}^{\nu}||^2} < \varepsilon$, stop; Else, return to step 2.

V. CONVERGENCE ANALYSIS

In this section, we show that Algorithm 1 is an energy descent. Moreover, the inner iteration of the $\boldsymbol{u}-$subproblem can converge to the minimizer of the subproblem (35).

Recalling the derivation of the proposed model in Section III, we use a dual formulation of the MAP estimation to derive the problem (24). Due to the existence of the ln-sum term, we use Proposition 3.3 to obtain the problem (26). Then, the problem of the minimization of $\mathcal{H}$ is replaced by the problem of optimizing $\mathcal{J}$. We have the following property of $\mathcal{H}$ and $\mathcal{J}$.

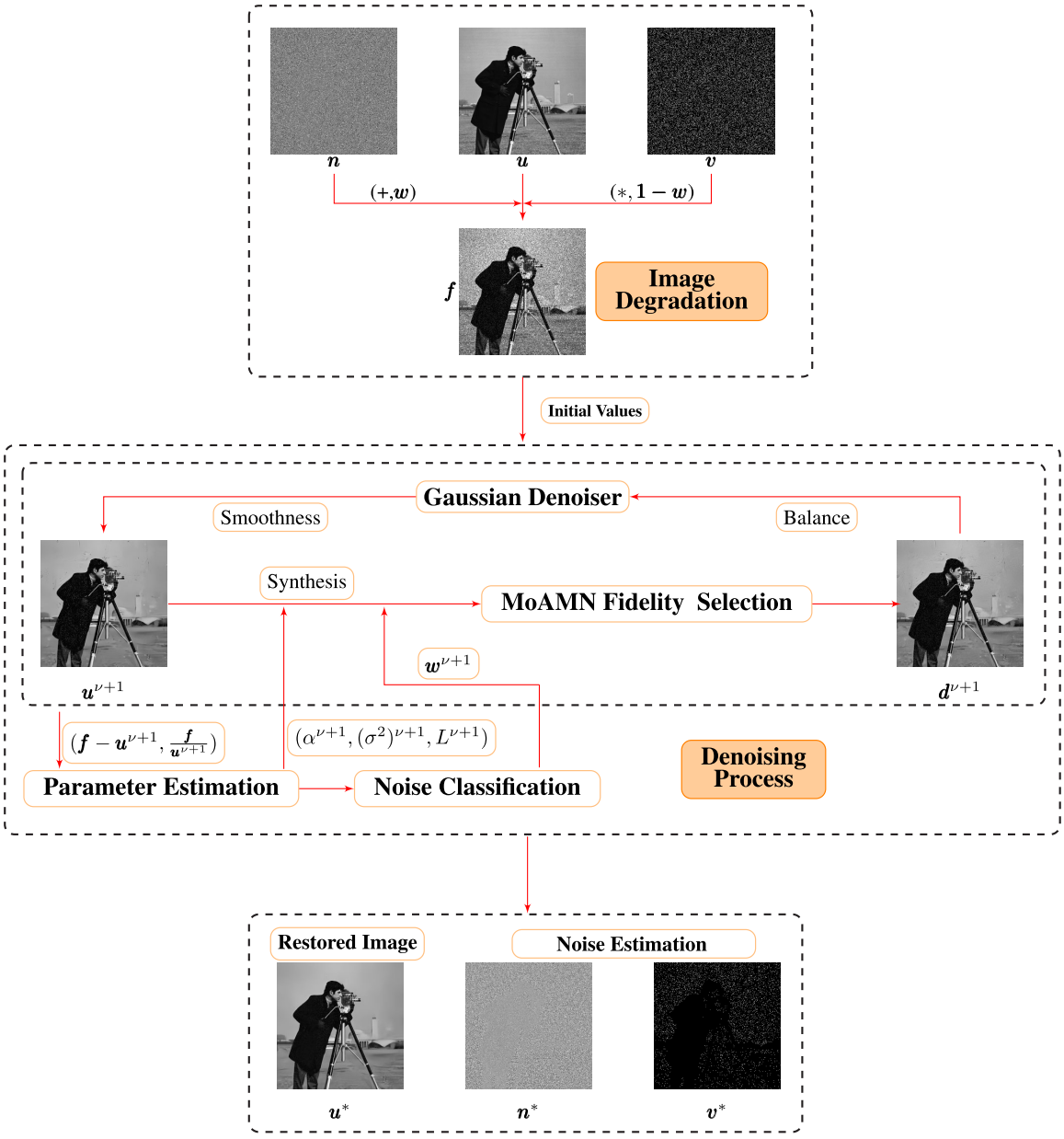


FIGURE 3. Flowchart of the proposed MoAMN denoising algorithm.

Theorem 5.1 [34]: The functionals $\mathcal{H}$ and $\mathcal{J}$ have the same global minimizer with respect to $(\boldsymbol{u}^*, \boldsymbol{\theta}^*)$.

Theorem 5.2 (Energy Descent): The sequence $(\boldsymbol{u}^\nu, \boldsymbol{\theta}^\nu)$ generated by the Algorithm 1 satisfies

$$\mathcal{H}(\boldsymbol{u}^{\nu+1}, \boldsymbol{\theta}^{\nu+1}) \leq \mathcal{H}(\boldsymbol{u}^\nu, \boldsymbol{\theta}^\nu). \tag{50}$$

Proof: The proof can be followed as Theorem 1 in [34].

The two following theorems are discussed under setting $\phi(\boldsymbol{u}) = TV(\boldsymbol{u})$ and $\boldsymbol{u}$ in the discrete BV space.

Theorem 5.3: In each ν-th outer inner, i.e., for fixed $(\sigma^2)^\nu$, $L^\nu, \boldsymbol{w}^\nu \in \Delta^+$, $\lambda > 0$, $r > 0$. Assume $\boldsymbol{f} \geq 0$, $\|\boldsymbol{f}\|_\infty < +\infty$. Let $\boldsymbol{u}^*$ be the minimizer of the problem (35). $0 < \boldsymbol{u} < 2\boldsymbol{f}$, $\forall\ 0 < \tau < 2r$; then, the sequence $\boldsymbol{u}^{\nu_1}$ generated by the iteration scheme (39) converges to $\boldsymbol{u}^*$, i.e., $\lim_{\nu_1 \to +\infty} \boldsymbol{u}^{\nu_1} = \boldsymbol{u}^*$.

Proof: The proof can be followed as Theorem 1 in [53].

Theorem 5.4: Assume that $(\sigma^2)^\nu$, L^ν, $\boldsymbol{w}^\nu$, λ, r, $\boldsymbol{f}$, $\boldsymbol{u}^*$, $\boldsymbol{u}$ satisfies the same conditions as Theorem 5.3, and let $\tau = r$; then, the sequence $\boldsymbol{u}^{\nu_1}$ generated by the iteration scheme (40) converges to $\boldsymbol{u}^*$, i.e., $\lim_{\nu_1 \to +\infty} \boldsymbol{u}^{\nu_1} = \boldsymbol{u}^*$.

Proof: The proof can be followed as Theorem 2 in [53].

VI. NUMERICAL RESULTS

In this section, we present the performance of the proposed model on the test images corrupted by a mixture of additive Gaussian and multiplicative Gamma noise. Algorithm 1 with

IRCNN regularization adopts a trained and plugin-and-play Gaussian denoiser in [27]. The training set is the same as [27] that includes 400 BSD images, 400 selected images from the validation set of the ImageNet dataset and 4744 images of Waterloo Exploration Dataset. Algorithm 1 is tested on a dataset that contains 10 test images: Cameraman (256 × 256), House (256 × 256), Peppers (256 × 256), Monarch (256 × 256), Lena (256 × 256), Man (512 × 512), Barbara (512 × 512), Couple (512 × 512), Boat (512 × 512), and Hill (512 × 512). Furthermore, it is also tested on two colour images: House(256 × 256×3), Lena(512 × 512×3), and three real medical images. The experiments are run on a computer with Inter(R) Core(TM) i7-8550U CPUs @1.8GHz (8 CPUs), 2.0GHz. The code is written by MATLAB and available on https://github.com/CuicuiZhao2018/MoAMN.

There are two evaluation indicators, the peak signal to noise ratio (PSNR) and structural similarity index (SSIM), to measure the quality of the restored images. The PSNR is defined by

$$PSNR(I, u) = 10 \log_{10} \left(\frac{mn}{\sum_{i',j'} (u_{i',j'} - I_{i',j'})^2} \right),$$

where I and u are the original image and restored image, respectively, and m and n are the row number and column number of the clean image's pixels, respectively. A larger PSNR value indicates a higher quality of the restored images.

SSIM [54] measures the structural similarity between two images. It is defined as

$$SSIM(I, u) = \frac{(2\mu_I \mu_u + c_1)(2\sigma_{Iu} + c_2)}{(\mu_I^2 \mu_u^2 + c_1)(\sigma_I^2 + \sigma_u^2 + c_2)}, \tag{51}$$

where μ_I, μ_u, σ_I^2, σ_u^2, and σ_{Iu} are the mean value of I, the mean value of u, the variance of I, the variance of u, and the covariance of I and u, respectively. The small constants c_1, $c_2 > 0$ make the computation of the SSIM stable.

We compare the results obtained by the proposed method with those by the existing methods, including the most closely related ROF [7], AA [14], BM3D [22], and IRCNN [27] algorithms. In the following experiments, the restored results are presented under the mixture of additive and multiplicative noise. The restored results of the proposed model are obtained using different regularization terms, including TV, BM3D, and DCNN. In the experiments, the test images are corrupted with different mixture ratios $\alpha \in \{0, 0.2, 0.4, 0.6, 0.8, 1\}$, different variances $\sigma^2 = 0.005, 0.01$ for Gaussian noise and different shape parameters $L \in \{10, 20, 30\}$ for Gamma noise.

In order to demonstrate the effectiveness of the proposed model, there is a clear visualization of the restored images in FIGUREs 4 and 5. FIGURE 4 shows the restored images of image Barbara under the true noise parameters $\alpha = 0.2$, $\sigma^2 = 0.01$, $L = 10$. In these figures, we zoom in for a part of the images. Firstly, the restored image by ROF method misses some details, as shown in FIGURE 4 (c). There are some speckles in the restored image by the AA method in FIGURE 4 (d). In comparison, the proposed method with TV regularization yields a better-restored image in FIGURE 4 (e), which not only removes noise more clearly but also remains much more information of the original image. Secondly, the restored image by the BM3D method in FIGURE 4 (f) is over-smoothing, but the proposed method with BM3D regularization in FIGURE 4 (g) retains more image information. Similar comparisons can be found between FIGURE 4 (h) and (i) for the IRCNN regularization. Similar results can be found in FIGURE 5 and more results (FIGUREs S1, S2, S3) are available as supplementary material. In short, with the same regularization term, it is found that the images restored by the proposed model have better texture details and fewer impurities than the existing methods. To further show the effectiveness of the proposed method, FIGURE 6 exhibits intermediate restored images and the corresponding PSNR values for image Cameraman under the true noise parameters $\alpha = 0.4$, $\sigma^2 = 0.005$, $L = 10$.

One step of the proposed approach is noise parameters estimation. TABLE 2 shows the estimated values of $\boldsymbol{\Theta} = (\alpha, \sigma^2, L)$ by the proposed method on image Cameraman. The true parameters are $\alpha = 0 : 0.2 : 1$, $\sigma = 0.005$, $L = 10$. There is just a little difference between the estimations of L and its true values except for $\alpha = 1$. Because the proposed model would degrade into a single additive denoiser when α equals 1, and there is no multiplicative noise in this case. It is a challenging task to precisely estimate all the parameters owing to unknown types and pollution ways of noise. The estimations of σ^2 and α are not very good. In fact, the restoration performance mainly depends on the estimation of the weight $\boldsymbol{w}$. The estimation of noise parameters can be rough since they are only used for noise classification.

The classification of noise is a key step of this approach. Thus, it is crucial to estimate the weight $\boldsymbol{w}$ accurately. Comparisons of the proposed method under known or unknown $\boldsymbol{w}$ are done on image Cameraman as shown in TABLE 3. Due to the page limitation, only parts of the experimental results are displayed in this table. For more similar results, please find them in the related table (TABLE S1) in the supplementary material. FIGURE 7 shows the PSNR curves for both known and unknown $\boldsymbol{w}$ in TABLE 3 as increasing $\alpha = 0 : 0.2 : 1$ among the proposed method and the two existing best performance algorithms under fixed true parameters $\sigma^2 = 0.005, L = 10$. FIGURE 8 shows the similar PSNR curves when $L = 10 : 10 : 30$, $\alpha = 0.4$, $\sigma^2 = 0.005$. Although the proposed method with unknown $\boldsymbol{w}$ provides worse results than those with known $\boldsymbol{w}$, it is still better than other existing algorithms. This fact indicates that the proposed method is applicable.

To show the statistical information, the average PSNR and SSIM values of the total 10 test images are listed in TABLE 4. It is obvious to see that the numerical comparison shows that most of the average PSNR and SSIM values produced by the proposed model are higher than those of the existing methods with the same regularizers. The

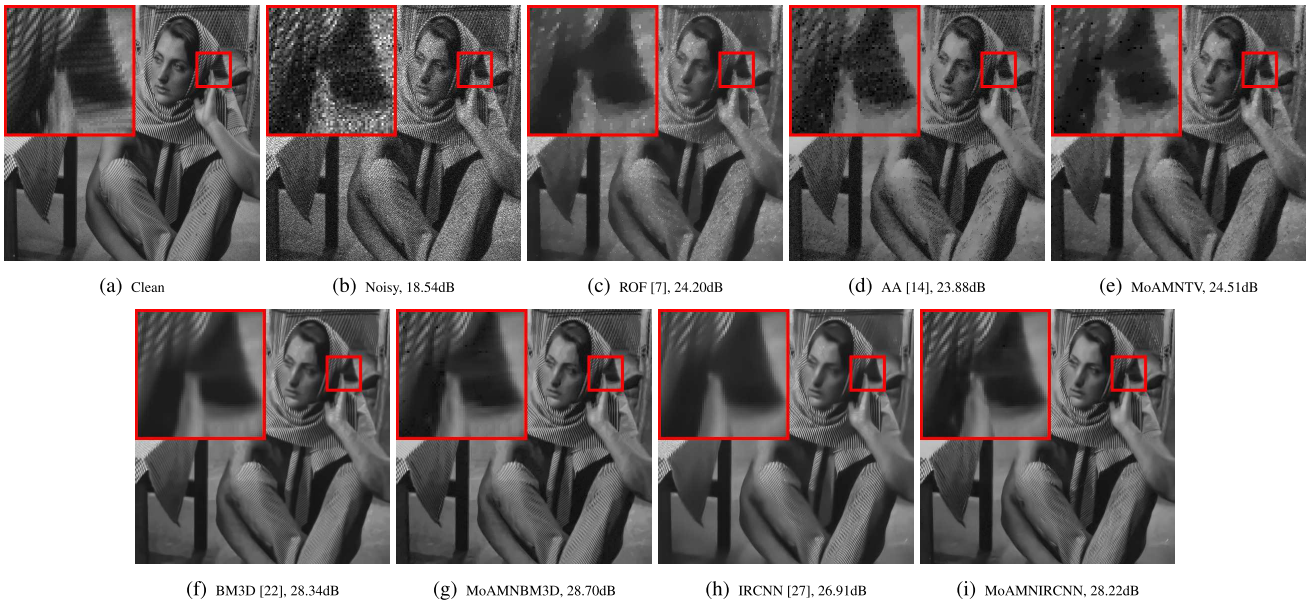


FIGURE 4. Denoising results and PSNR values of image Barbara. The true parameters of the noise are $\alpha = 0.2$, $\sigma^2 = 0.01$, and $L = 10$.

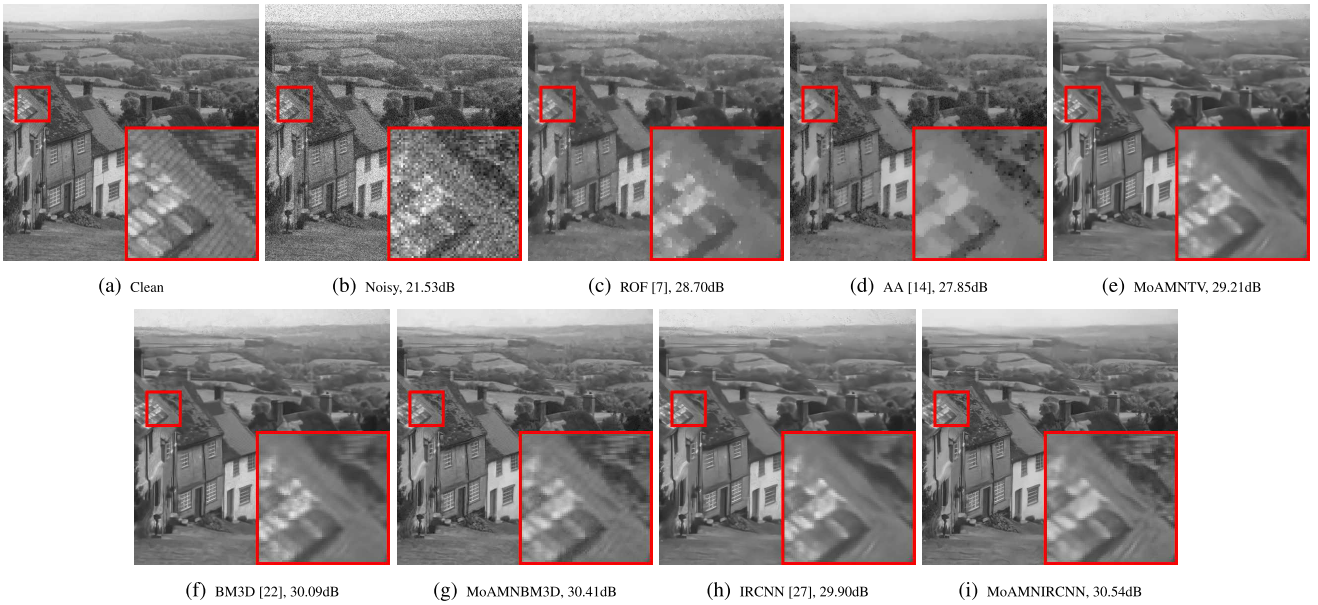


FIGURE 5. Denoising results and PSNR values of image Hill. The true parameters of noise are $\alpha = 0.6$, $\sigma^2 = 0.005$, and $L = 20$.

FIGURE 6. The intermediate restored images of the ν-th iteration steps for image Cameraman by using MoAMNIRCNN. The true parameters of noise are $\alpha = 0.4$, $\sigma^2 = 0.005$, and $L = 10$.

TABLE 2. Estimations of noise parameters α, σ^2 and L for denoised image Cameraman by the proposed method with TV, BM3D, and IRCNN regularizations under the true parameters $\alpha = 0 : 0.2 : 1$, $\sigma^2 = 0.005$, $L = 10$.

α	σ^2	L	Proposed Method								
			Regularizers								
			TV			BM3D			IRCNN		
$\downarrow$	$\downarrow$	$\downarrow$	$\hat{\alpha}$	$\hat{\sigma}^2$	$\hat{L}$	$\hat{\alpha}$	$\hat{\sigma}^2$	$\hat{L}$	$\hat{\alpha}$	$\hat{\sigma}^2$	$\hat{L}$
0	0.005	10	0	0	13.76	0	0	11.41	0.66	0.02	10.02
0.2	0.005	10	0.67	0.01	11.00	0.68	0.01	9.37	0.66	0.02	10.13
0.4	0.005	10	0.80	0.01	8.72	0.68	0.01	11.38	0.67	0.01	12.00
0.6	0.005	10	0.72	0.01	12.97	0.69	0.01	13.41	0.69	0.01	12.66
0.8	0.005	10	0.76	0.0038	16.13	0.77	0.0048	10.64	0.71	0.0044	16.05
1	0.005	10	1	0.0038	14.03	1	0.0042	24.04	1	0.0043	$2.7 * 10^{15}$

TABLE 3. Comparison of the PSNR/SSIM values for denoised image Cameraman under different levels by ROF [7], AA [14], BM3D [22], IRCNN [27], and MoAMN model with setting different initial values of w. (The highest PSNRs and SSIMs are shown in red, and the second-highest PSNRs and SSIMs are shown in blue).

α	σ^2	L	Existing Methods				Proposed Method					
							Regularizers					
							TV		BM3D		IRCNN	
$\downarrow$	$\downarrow$	$\downarrow$	ROF [7]	AA [14]	BM3D [22]	IRCNN [27]	Initial Parameter w		Initial Parameter w		Initial Parameter w	
							Unknown	Known	Unknown	Known	Unknown	Known
0	0.005	10	25.00	25.56	26.64	27.08	25.62	25.62	27.46	27.46	27.88	27.70
			0.7006	0.7926	0.7628	0.7886	0.8112	0.8112	0.8307	0.8307	0.8303	0.8300
0.2	0.005	10	25.45	25.78	27.29	27.57	26.16	26.21	27.84	28.18	28.29	28.69
			0.7115	0.7631	0.7805	0.7959	0.7875	0.8107	0.8225	0.8375	0.8265	0.8429
0.4	0.005	10	25.90	25.93	27.80	28.01	26.51	27.00	28.43	28.74	28.88	28.93
			0.7311	0.7466	0.7927	0.8115	0.7629	0.8148	0.8360	0.8531	0.8368	0.8636
0.6	0.005	10	26.53	26.19	28.57	28.71	27.45	27.74	29.02	29.31	29.41	29.68
			0.7517	0.7320	0.8160	0.8110	0.8118	0.8213	0.8381	0.8598	0.8425	0.8682
0.8	0.005	10	27.30	26.49	29.42	29.60	28.12	28.42	29.64	29.76	29.81	30.25
			0.7612	0.7229	0.8416	0.8388	0.8121	0.8273	0.8635	0.8701	0.8382	0.8770
1	0.005	10	29.37	27.49	30.99	31.42	29.55	29.55	30.79	30.79	31.37	31.37
			0.8254	0.7621	0.8834	0.8947	0.8494	0.8494	0.8720	0.8720	0.8910	0.8910

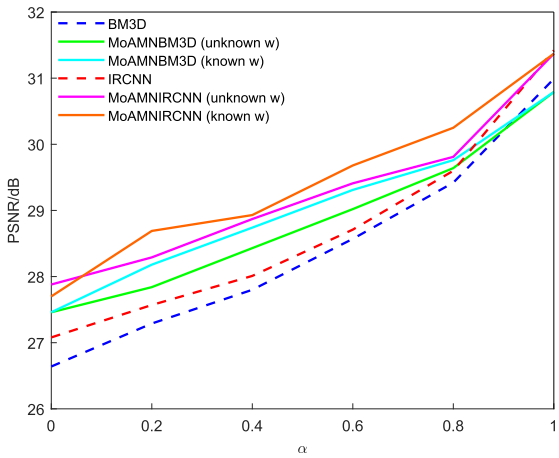


FIGURE 7. PSNR values for denoising image Cameraman in TABLE 3 under the mixture of additive and multiplicative noise with increasing ratio $\alpha = 0: 0.2: 1$. The parameters $\sigma^2 = 0.005$, $L = 10$ are fixed.

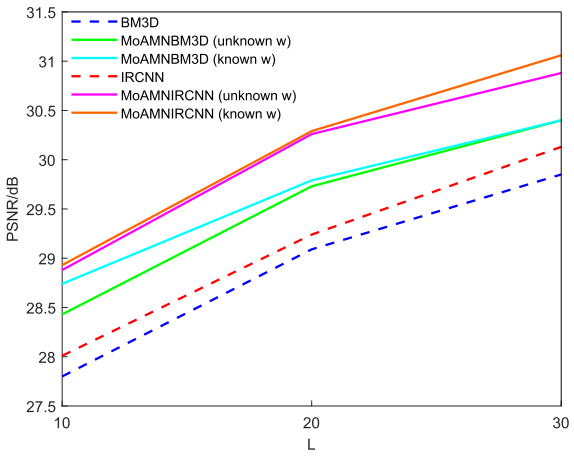


FIGURE 8. PSNR values for denoising image Cameraman in TABLE 3 under the mixture of additive and multiplicative noise with increasing ratio $L = 10: 10: 30$. The parameters $\alpha = 0.4$ and $\sigma^2 = 0.005$ are fixed.

best-performing method is the proposed method with IRCNN regularization. For the case $\alpha = 0.4$, the biggest average improvement is about 0.81 dB between the proposed method and the original IRCNN algorithm, while the smallest is 0.56 dB. For the case of $\alpha = 0.6$, the biggest average improvement is 0.62 dB while the smallest is 0.33 dB. To avoid repetition, more clear illustrations of TABLE 4 can refer to FIGUREs 9, 10, and their analysis in the following paragraphs.In addition, the comparison of PSNR/SSIM values for all test grey images are shown in the table and

TABLE 4. Average PSNR/SSIM values for all 10 test images under different levels for ROF [7], AA [14], BM3D [22], IRCNN [27], and MoAMN models with different regularizers. (The highest PSNRs are shown in red, and the second-highest PSNRs are shown in blue.

α	σ^2	L	Existing Methods				Proposed Method		
							Regularizers		
↓	↓	↓	ROF [7]	AA [14]	BM3D [22]	IRCNN [27]	TV	BM3D	IRCNN
0	0.005	10	25.57/0.7542	25.45/0.7834	27.32/0.8213	27.32/0.8124	25.94/0.8027	27.81/0.8503	28.14/0.8573
0.2	0.005	10	26.03/0.7667	25.72/0.7844	27.98/0.8359	27.93/0.8304	26.58/0.8091	28.41/0.8587	28.72/0.8652
0.4	0.005	10	26.49/0.7836	25.97/0.7857	28.55/0.8517	28.48/0.8449	27.05/0.8222	28.92/0.8684	29.29/0.8742
0.6	0.005	10	27.09/0.8025	26.28/0.7928	29.28/0.8691	29.22/0.8614	27.66/0.8369	29.54/0.8809	29.84/0.8834
0.8	0.005	10	27.89/0.8250	26.66/0.7990	30.23/0.8884	30.24/0.8846	28.35/0.8486	30.37/0.8964	30.49/0.8928
1	0.005	10	29.67/0.8745	27.40/0.8239	31.56/0.9147	31.88/0.9187	29.73/0.8798	31.51/0.9131	31.89/0.9184
0	0.005	20	27.10/0.8032	27.42/0.8376	29.05/0.8633	28.93/0.8534	27.71/0.8459	29.56/0.8833	30.04/0.8895
0.2	0.005	20	27.43/0.8104	27.41/0.8305	29.41/0.8712	29.36/0.8618	28.06/0.8459	29.93/0.8880	30.38/0.8957
0.4	0.005	20	27.79/0.8212	27.39/0.8243	29.81/0.8797	29.84/0.8752	28.34/0.8532	30.25/0.8941	30.61/0.8990
0.6	0.005	20	28.25/0.8360	27.39/0.8222	30.34/0.8910	30.49/0.8902	28.71/0.8611	30.59/0.9004	30.94/0.9041
0.8	0.005	20	28.87/0.8529	27.39/0.8210	30.98/0.9025	31.14/0.9030	29.11/0.8689	31.12/0.9077	31.32/0.9101
1	0.005	20	29.68/0.8742	27.41/0.8232	31.56/0.9148	31.88/0.9187	29.73/0.8798	31.51/0.9131	31.89/0.9184
0	0.005	30	28.00/0.8285	28.52/0.8636	29.92/0.8825	29.88/0.8735	28.70/0.8663	30.52/0.8997	30.97/0.9060
0.2	0.005	30	28.25/0.8346	28.26/0.8491	30.16/0.8854	30.21/0.8834	28.87/0.8638	30.73/0.9022	31.11/0.9081
0.4	0.005	30	28.58/0.8427	28.07/0.8399	30.52/0.8919	30.67/0.8929	29.07/0.8679	30.95/0.9061	31.25/0.9090
0.6	0.005	30	28.90/0.8529	27.84/0.8321	30.88/0.8994	31.06/0.9025	29.28/0.8720	31.07/0.9065	31.44/0.9127
0.8	0.005	30	29.24/0.8620	27.62/0.8267	31.22/0.9071	31.50/0.9108	29.47/0.8761	31.30/0.9105	31.55/0.9140
1	0.005	30	29.68/0.8742	27.41/0.8223	31.56/0.9148	31.88/0.9187	29.73/0.8798	31.51/0.9131	31.89/0.9184
0	0.01	10	25.57/0.7543	25.43/0.7836	27.32/0.8219	27.32/0.8124	25.94/0.8027	27.81/0.8503	28.14/0.8573
0.2	0.01	10	25.96/0.7669	25.44/0.7730	27.84/0.8302	27.84/0.8292	26.42/0.7998	28.21/0.8540	28.59/0.8627
0.4	0.01	10	26.33/0.7763	25.38/0.7663	28.31/0.8453	28.33/0.8417	26.73/0.8123	28.45/0.8572	28.89/0.8651
0.6	0.01	10	26.79/0.7932	25.32/0.7608	28.81/0.8578	28.86/0.8557	27.10/0.8205	28.88/0.8672	29.19/0.8718
0.8	0.01	10	27.31/0.8105	25.28/0.7580	29.38/0.8711	29.53/0.8723	27.51/0.8290	29.43/0.8766	29.58/0.8791
1	0.01	10	28.05/0.8345	25.23/0.7605	29.96/0.8852	30.28/0.8916	28.02/0.8385	29.89/0.8831	30.25/0.8903

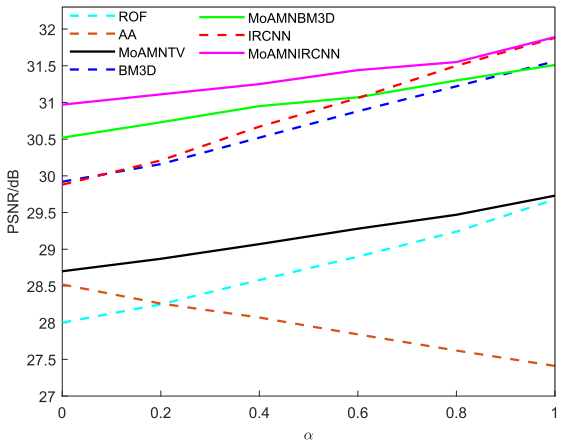


FIGURE 9. Average PSNR values in TABLE 4 under the mixture of additive and multiplicative noise with increasing ratio $\alpha = 0 : 0.2 : 1$. The parameters $\sigma^2 = 0.005$ and $L = 30$ are fixed.

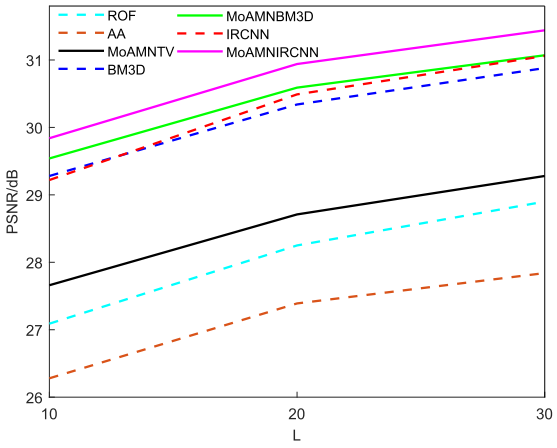


FIGURE 10. Average PSNR values in TABLE 4 under the mixture of additive and multiplicative noise with increasing ratio $L = 10 : 10 : 30$. The parameters $\alpha = 0.6$, $\sigma^2 = 0.005$ are fixed.

figure in the submitted supplementary materials (TABLE S2. FIGURE S4).

However, it is worth noticing that there are some exceptions in TABLE 3 and 4, such as when $\alpha = 0$ and $\alpha = 1$. The reason is that the noise becomes a single multiplicative noise or additive noise in these two cases, and there will be no improvement for the proposed fidelity term in this scenario. If we set all elements of the weight $\boldsymbol{w}$ as 0 or 1, then the proposed fidelity will be degraded to ROF's or *AA*'s, and the algorithm will produce similar results to those obtained by the existing methods. That is to say, to obtain results similar to those of the existing methods for removing single noise, it is necessary to choose appropriate initial values and model parameters.

To more clear visualization of the relationship between the PSNR values and the parameters, FIGUREs 9 and 10 show some PSNR curves in TABLE 4. Specifically, FIGURE 9 shows the relationship between the average PSNR values and the mixed ratio of additive and multiplicative noise in increasing order as $\alpha = 0 : 0.2 : 1$ and the fixed

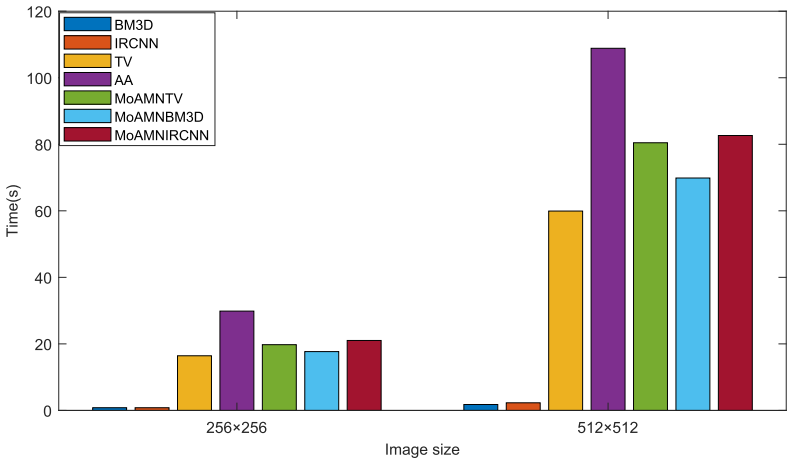


FIGURE 11. Comparison of computational cost for different methods under the true parameters $\alpha = 0.4$, $\sigma^2 = 0.005$, $L = 10$.

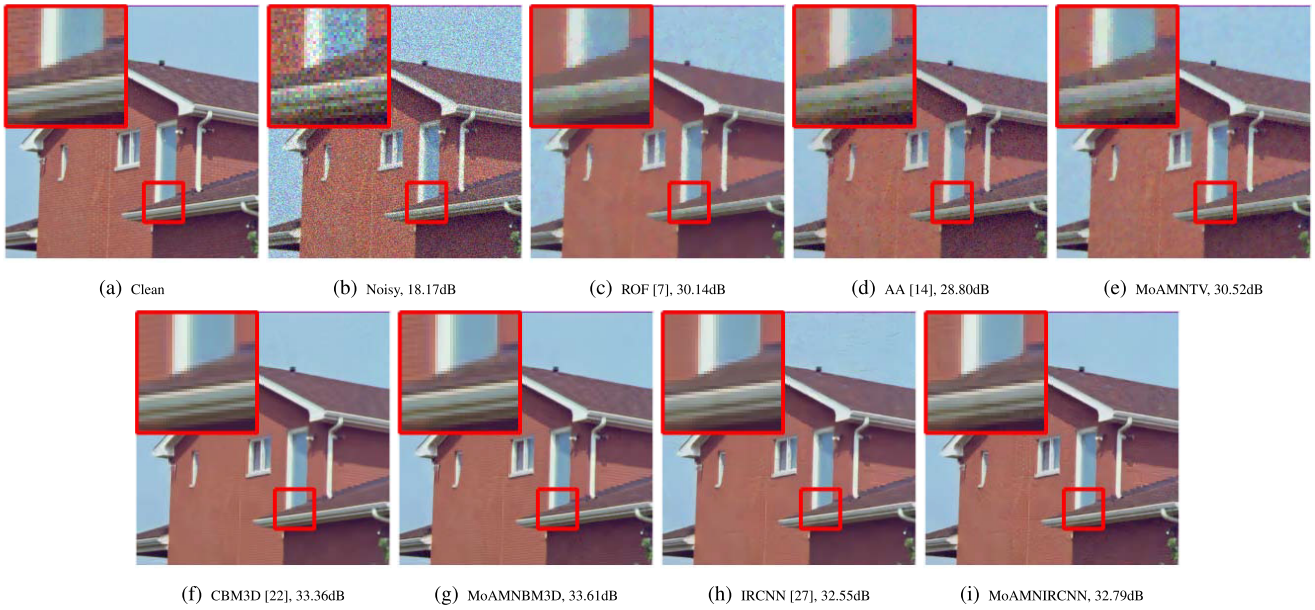


FIGURE 12. Denoising results and PSNR values of color image House($256 \times 256 \times 3$). The true parameters of the noise are $\alpha = 0.6$, $\sigma^2 = 0.005$, and $L = 30$.

noise parameters $\sigma^2 = 0.005$, $L = 30$. When $\alpha = 0$, the difference between the PSNR values obtained by the proposed model and the ROF model with the same regularizer reaches maximum magnitude. As the mixed ratio increases, the difference decreases. When $\alpha = 1$, the advantage of the proposed model almost disappears, and the ROF model was even slightly better than the proposed model. This outcome is reasonable because the mixed noise becomes a single additive noise with a Gaussian distribution in this case. Furthermore, FIGURE 10 shows the relationship between the average PSNR values and the difference in the two noise variances under increasing $L = 10 : 10 : 30$ while fixing the mixed ratio $\alpha = 0.6$ and variance of Gaussian distribution $\sigma^2 = 0.05$. One can see that as L increases, the advantage of the proposed model decreases, indicating that the difference between the variances in the Gaussian and Gamma distribution influences the validity of the proposed model.

According to the numerical experiments, the running time of the proposed method changes a little as the noise parameters change except for $\alpha = 1$ when the proposed model degrades into a Gaussian denoiser. FIGURE 11 shows a histogram of running time among the existing ROF, BM3D, DCNN algorithms, and the proposed method under the true parameters $\alpha = 0.4$, $\sigma^2 = 0.005$, $L = 10$. The histogram involves two image sizes, 256×256 and 512×512. The proposed method is time-consuming compared with other methods. The reason is that the proposed method needs to

TABLE 5. Comparison of the PSNR/SSIM values for color image House denoising. (The highest PSNRs are shown in red, and the second-highest PSNRs are shown in blue).

α	σ^2	L	Existing Methods				Proposed Method (MoAMN)		
							Regularizers		
↓	↓	↓	ROF [7]	AA [14]	CBM3D [22]	IRCNN [27]	TV	CBM3D	IRCNN
0	0.005	10	25.85/0.7340	25.06/0.7255	28.35/0.8018	27.82/0.7933	26.09/0.7208	28.48/0.8166	28.45/0.8185
0.2	0.005	10	26.59/0.7449	25.69/0.7372	29.22/0.8116	28.72/0.8052	26.83/0.7313	29.54/0.8221	29.44/0.8220
0.4	0.005	10	27.42/0.7488	26.36/0.7437	30.20/0.8149	29.70/0.8123	27.62/0.7409	30.76/0.8333	30.51/0.8291
0.6	0.005	10	28.24/0.7499	27.00/0.7459	31.36/0.8240	30.64/0.8161	28.44/0.7528	31.80/0.8401	31.42/0.8353
0.8	0.005	10	29.22/0.7693	27.79/0.7572	32.64/0.8448	31.71/0.8301	29.32/0.7656	32.73/0.8588	32.17/0.8485
1	0.005	10	31.21/0.8189	28.89/0.7758	34.15/0.8762	33.53/0.8640	30.72/0.7793	34.14/0.8713	33.39/0.8616
0	0.005	20	28.01/0.7595	27.60/0.7731	30.95/0.8207	30.23/0.8131	28.18/0.7488	31.37/0.8367	31.04/0.8350
0.2	0.005	20	28.52/0.7667	27.90/0.7746	31.57/0.8246	30.76/0.8137	28.64/0.7571	31.94/0.8440	31.61/0.8412
0.4	0.005	20	28.99/0.7709	28.17/0.7713	32.10/0.8336	31.23/0.8106	29.13/0.7653	32.50/0.8531	32.01/0.8482
0.6	0.005	20	29.58/0.7796	28.43/0.7693	32.73/0.8458	31.90/0.8224	29.89/0.7498	32.94/0.8616	32.37/0.8540
0.8	0.005	20	30.17/0.7895	28.60/0.7615	33.52/0.8594	32.68/0.8457	30.07/0.7818	33.47/0.8660	32.63/0.8575
1	0.005	20	31.21/0.8189	28.93/0.7988	34.15/0.8762	33.53/0.8640	30.72/0.7793	34.14/0.8713	33.39/0.8616
0	0.005	30	29.06/0.7734	28.92/0.7968	32.11/0.8338	31.20/0.8178	29.21/0.7650	32.54/0.8545	32.07/0.8489
0.2	0.005	30	29.35/0.7841	28.94/0.7914	32.54/0.8369	31.68/0.8243	29.51/0.7724	32.98/0.8577	32.36/0.8521
0.4	0.005	30	29.73/0.7872	28.98/0.7870	32.88/0.8430	32.07/0.8263	29.84/0.7782	33.27/0.8653	32.56/0.8574
0.6	0.005	30	30.14/0.7824	28.80/0.7650	33.36/0.8557	32.55/0.8373	30.52/0.7892	33.61/0.8715	32.79/0.8610
0.8	0.005	30	30.57/0.7996	28.89/0.7776	33.73/0.8644	33.06/0.8523	30.32/0.7857	33.56/0.8680	32.75/0.8578
1	0.005	30	31.21/0.8189	28.89/0.7758	34.15/0.8762	33.53/0.8640	30.72/0.7793	34.14/0.8713	33.37/0.8627
0	0.01	10	25.85/0.7340	25.06/0.7255	28.35/0.8018	27.82/0.7933	26.09/0.7208	28.48/0.8166	28.45/0.8185
0.2	0.01	10	26.54/0.7307	25.42/0.7307	29.16/0.8098	28.65/0.8020	26.68/0.7202	29.38/0.8171	29.31/0.8208
0.4	0.01	10	27.25/0.7423	25.81/0.7195	29.88/0.8064	29.54/0.8056	27.35/0.7335	30.35/0.8237	30.18/0.8258
0.6	0.01	10	27.94/0.7535	26.08/0.7172	30.96/0.8205	30.28/0.8174	27.81/0.7284	31.03/0.8222	30.80/0.8243
0.8	0.01	10	28.73/0.7682	26.35/0.7139	32.03/0.8363	31.19/0.8233	28.36/0.7423	31.65/0.8281	31.17/0.8259
1	0.01	10	29.75/0.7957	26.44/0.7225	32.93/0.8486	32.03/0.8420	28.89/0.7179	32.82/0.8415	32.13/0.8407

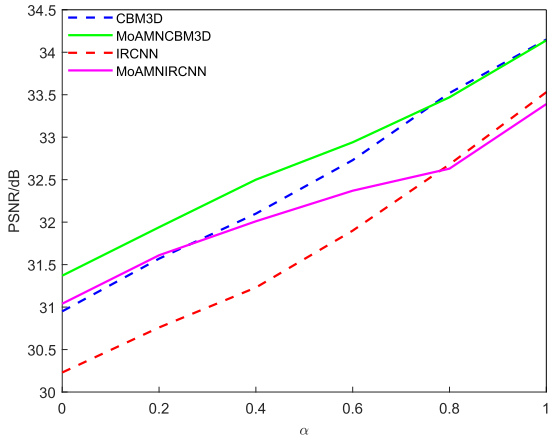


FIGURE 13. PSNR values for denoised colorful image House in TABLE 5 under the mixture of additive and multiplicative noise with increasing ratio $\alpha = 0{:}\ 0.2{:}\ 1$. The parameters $\sigma^2 = 0.005$, $L = 20$ are fixed.

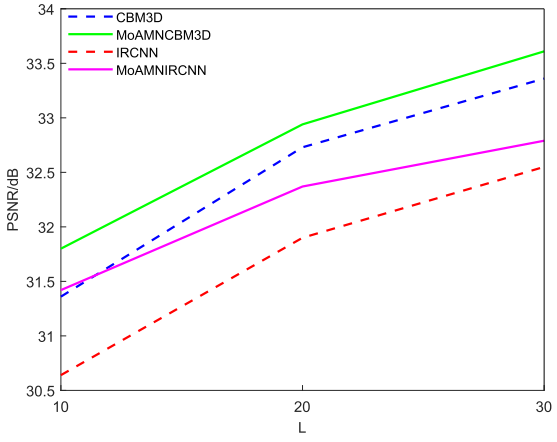


FIGURE 14. PSNR values for denoised colorful image House in TABLE 5 under the mixture of additive and multiplicative noise with increasing ratio $L = 10{:}\ 10{:}\ 30$. The parameters $\alpha = 0.6$, $\sigma^2 = 0.005$, are fixed.

solve cubic equations. Furthermore, the running time of the proposed method increases as image sizes become bigger. Hence, the main factors that affect the running time are image sizes and equipment.

To illustrate the effectiveness of the proposed model on colour images, the proposed method tests on Lena($512 \times 512 \times 3$) and House($256 \times 256 \times 3$). Specifically, each channel is dealt with for ROF, AA, and the proposed method based on the TV regularizer separately, and then the colour channels are combined to obtain the corresponding restored colour image. For CBM3D and IRCNN, we apply them directly to the proposed model as regularizers and compare them with the corresponding original method. FIGURE 12 shows the restored images of the test colour image House and enlarges parts of the restored images. One can find that there are similar analysis results for the colour image House as the previous grey images.

TABLE 5 shows PSNR/SSIM results for ROF, AA, CBM3D, IRCNN, and the proposed method with the corresponding regularizers for the colour image House. Most results in TABLE 5 show that the proposed method is superior

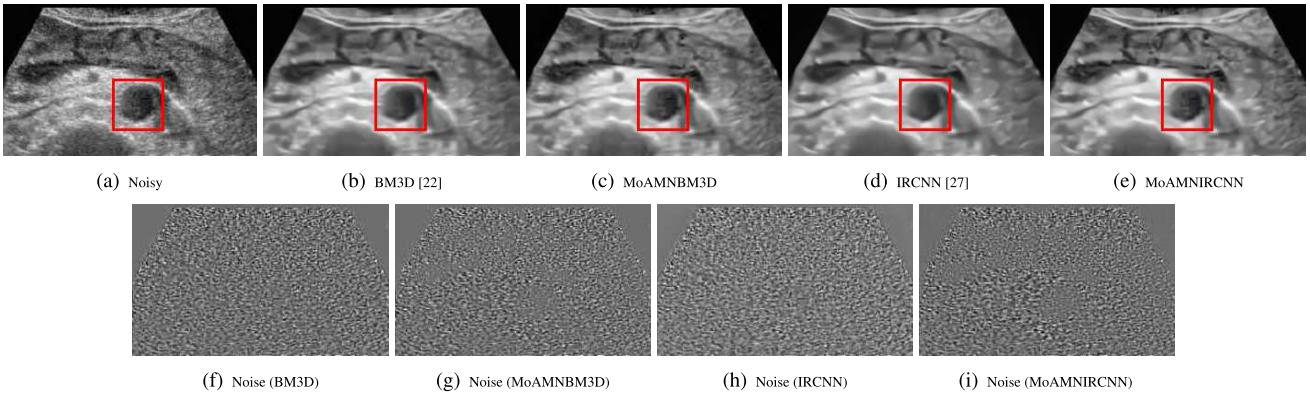


FIGURE 15. The denoising results of a real ultrasound image obtained by the proposed model compared with BM3D and IRCNN.

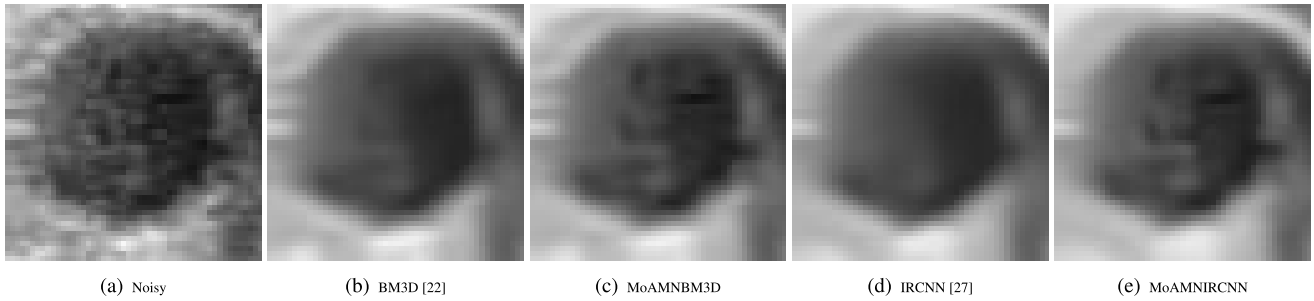


FIGURE 16. The enlarged parts of the red rectangle areas in FIGURE15.

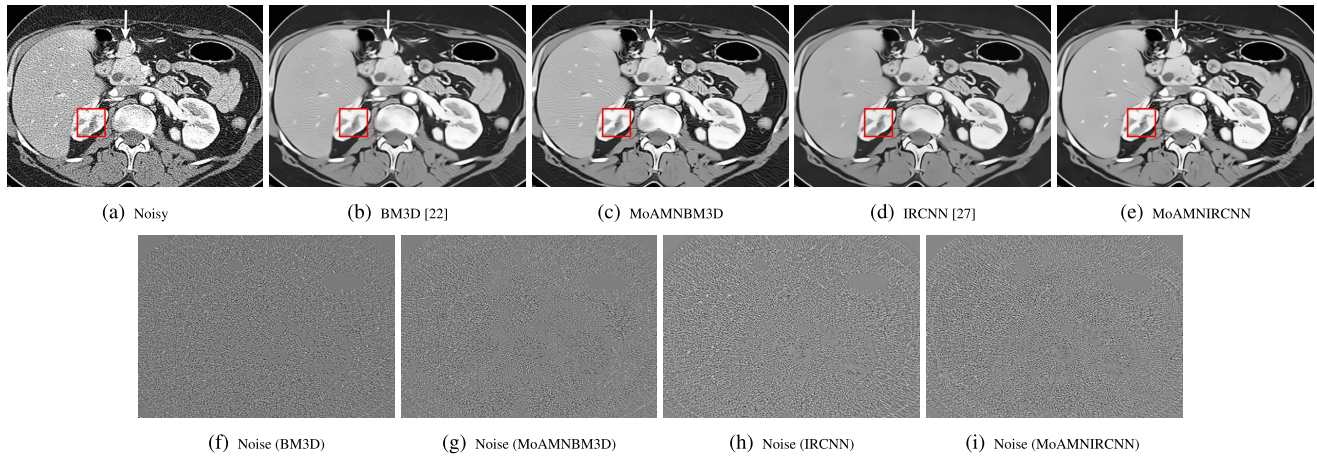


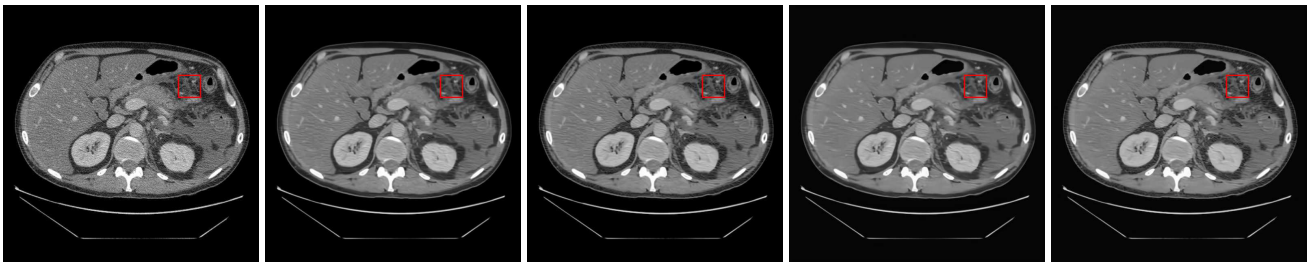
FIGURE 17. The denoising results of a real image obtained by the proposed model compared with BM3D and IRCNN.

to other methods, which coincides with the results of the above grey images. FIGURE 13 shows the curve of PSNR values in TABLE 5 as increasing $\alpha = 0 : 0.2 : 1$ under the true parameters $\sigma^2 = 0.005, L = 20$. FIGURE 14 shows the curve of PSNR values in TABLE 5 as increasing $L = 10 : 10 : 30$ under the true parameters $\alpha = 0.6, \sigma^2 = 0.005$. There are similar analysis results with the grey images. Furthermore, the restored images of the colour image Lena are shown in FIGURE S5 and PSNR/SSIM values are shown in TABLE S3 as supplementary material. It is easy to find that the proposed method can better preserve details and textures.

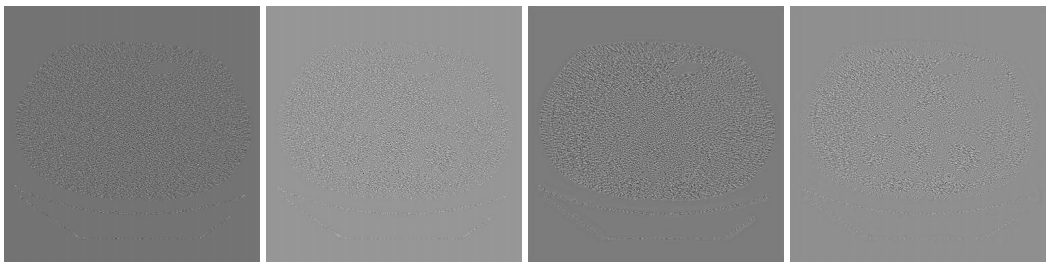
We apply the proposed method to real noisy data. FIGURE 15 shows a real noisy ultrasound image downloaded from the website http://www.lib.dmu.edu.cn/database/csyx.jsp. The real ultrasound image contains noise with unknown types and levels. FIGURE 15 shows the comparison of the restored images by different algorithms, which contains the proposed method, BM3D and IRCNN algorithms.

(a) Noisy (b) BM3D [22] (c) MoAMNBM3D (d) IRCNN [27] (e) MoAMNIRCNN

FIGURE 18. The enlarged parts of the red rectangle areas in FIGURE17.

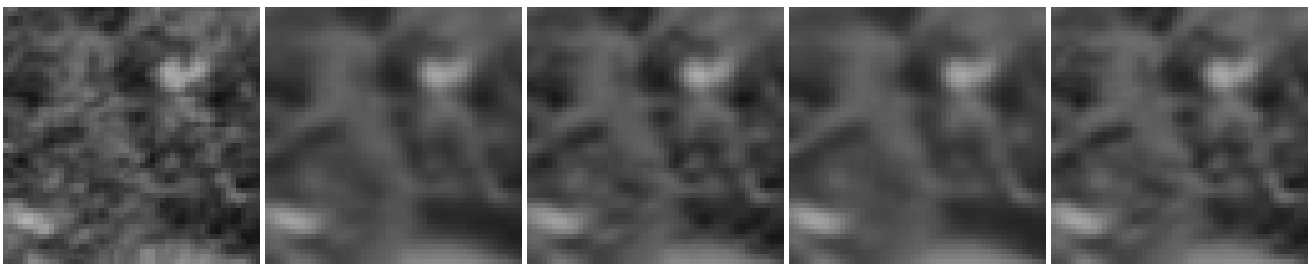


(a) Noisy (b) BM3D [22] (c) MoAMNBM3D (d) IRCNN [27] (e) MoAMNIRCNN



(f) Noise (BM3D) (g) Noise (MoAMNBM3D) (h) Noise (IRCNN) (i) Noise (MoAMNIRCNN)

FIGURE 19. The denoising results of a real CT image obtained by the proposed model compared with BM3D and IRCNN.



(a) Noisy (b) BM3D [22] (c) MoAMNBM3D (d) IRCNN [27] (e) MoAMNIRCNN

FIGURE 20. The enlarged parts of the red rectangle areas in FIGURE19.

As observed from this figure, the proposed method can preserve more image textures than BM3D and IRCNN while removing noise efficiently. To see more details, we choose the same rectangle area for each image marked in red and enlarge them in FIGURE 16. The structures of the tissues are observed to be preserved in the restorations produced by the proposed method, but not the restored images obtained by BM3D and IRCNN. Examination of the removed noise presented in FIGURE 15 shows that the noise removed by BM3D and IRCNN is almost uniform everywhere, while the noise removed by the proposed method can be inhomogeneous. This scenario is reasonable since the weighting function in the proposed algorithm can adaptively detect noise types and levels.

Similar results can be found in FIGUREs 17, 18, 19, and 20, which show the denoising results of real noisy CT images in [55].

VII. DISCUSSION AND CONCLUSION

In this paper, a novel model is proposed to remove mixed additive and multiplicative noise at different ratios. The introduced noise type detection function is novel and has a statistical interpretation. Moreover, The weighting function for noise detection is a function in dual space from the convex optimization perspective. We prove the existence of a minimizer for the proposed variational model with TV regularization. To design an efficient algorithm, we apply the popular splitting scheme to separate the proposed model into several easily solvable subproblems, enabling the application of many good regularizers such as BM3D and IRCNN. Significantly, the proposed approach with IRCNN regularization combined the variational method to detect noise parameters and CNN deep learning algorithm to learn images prior. The numerical experiments were done on ten test grey images, two colour images, and three real noisy medical images. According to the comparison of the restored images, the method in this paper can preserve more image details when removing mixed noise than the existing ROF, AA, BM3D, and IRCNN algorithms. For the most attractive cases $\alpha = 0.4$ and $\alpha = 0.6$, the biggest improvements of average PSNR values on 10 test grey images can reach about 0.81 dB and 0.62 dB respectively between the proposed approach with IRCNN regularization and the existing IRCNN algorithm.

Note that accurate detection of a mixture of additive and multiplicative noise is challenging. Firstly, in the proposed method, the minimizer may not be unique because the mathematical model is non-convex with respect to both $\boldsymbol{w}$ and $\boldsymbol{u}$. Secondly, the estimation of the noise parameters is not accurate enough. So the estimation influences the denoising performance.Thirdly, the update step on denoising in this paper involves a Gaussian denoiser and a fidelity selection. It may cause incomplete denoising. Fourthly, it is well-known that the EM algorithm is locally converging. Therefore, the performance of the proposed method partly depends on the initial parameters α, σ^2, and L. In the future, a convex mathematical model could be developed, which can refer to the transformation way of the AA model into convex. A new splitting way may need to make a perfect combination of an additive denoiser and a multiplicative denoiser for the algorithm. In this case, the existing CNN deep learning algorithm for denoising multiplicative noise can be integrated into the proposed model to improve the performance. Some other trendy single denoisers also can be integrated to improve the presented model. In addition, one can design a CNN architecture according to the proposed variational problem by unrolling technique, and this will let the denoising CNN have physical significance and interpretability. We leave these as future research.

REFERENCES

[1] L. Gondara, "Medical image denoising using convolutional denoising autoencoders," in *Proc. IEEE 16th Int. Conf. Data Mining Workshops (ICDMW)*, Dec. 2016, pp. 241–246.
[2] S. Li, L. Fang, and H. Yin, "An efficient dictionary learning algorithm and its application to 3-D medical image denoising," *IEEE Trans. Biomed. Eng.*, vol. 59, no. 2, pp. 417–427, Feb. 2012.
[3] X. Su, C. A. Deledalle, F. Tupin, and H. Sun, "Two-step multitemporal nonlocal means for synthetic aperture radar images," *IEEE Trans. Geosci. Remote Sens.*, vol. 52, no. 10, pp. 6181–6196, Oct. 2014.
[4] L. Xu, J. Li, Y. Shu, and J. Peng, "SAR image denoising via clustering-based principal component analysis," *IEEE Trans. Geosci. Remote Sens.*, vol. 52, no. 11, pp. 6858–6869, Nov. 2014.
[5] S. Xu, Y. Zhou, H. Xiang, and S. Li, "Remote sensing image denoising using patch grouping-based nonlocal means algorithm," *IEEE Geosci. Remote Sens. Lett.*, vol. 14, no. 12, pp. 2275–2279, Dec. 2017.
[6] E. S. Chang, C.-C. Hung, W. Liu, and J. Yina, "A denoising algorithm for remote sensing images with impulse noise," in *Proc. IEEE Int. Geosci. Remote Sens. Symp. (IGARSS)*, Jul. 2016, pp. 2905–2908.
[7] L. I. Rudin, S. Osher, and E. Fatemi, "Nonlinear total variation based noise removal algorithms," *Phys. D, Nonlinear Phenomena*, vol. 60, nos. 1–4, pp. 259–268, 1992.
[8] C. Wu, J. Zhang, and X.-C. Tai, "Augmented Lagrangian method for total variation restoration with non-quadratic fidelity," *Inverse Probl. Imag.*, vol. 5, no. 1, pp. 237–261, 2011.
[9] A. Rajwade, A. Rangarajan, and A. Banerjee, "Image denoising using the higher order singular value decomposition," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 35, no. 4, pp. 849–862, Apr. 2013.
[10] D. N. H. Thanh, V. B. S. Prasath, L. M. Hieu, and S. Dvoenko, "An adaptive method for image restoration based on high-order total variation and inverse gradient," *Signal, Image Video Process.*, vol. 14, no. 6, pp. 1189–1197, Sep. 2020.
[11] S. Wali, C. M. Li, A. Basit, A. Shakoor, R. A. Memon, S. Rahim, and S. Samina, "Fast and adaptive boosting techniques for variational based image restoration," *IEEE Access*, vol. 7, pp. 181491–181504, 2019.
[12] S. Wali, "A boosting procedure for variational-based image restoration," *Numer. Math., Theory, Methods Appl.*, vol. 11, no. 1, pp. 49–73, Jun. 2018.
[13] S. Wali, H. Zhang, H. Chang, and C. Wu, "A new adaptive boosting total generalized variation (TGV) technique for image denoising and inpainting," *J. Vis. Commun. Image Represent.*, vol. 59, pp. 39–51, Feb. 2019.
[14] G. Aubert and J.-F. Aujol, "A variational approach to removing multiplicative noise," *SIAM J. Appl. Math.*, vol. 68, no. 4, pp. 925–946, 2008.
[15] J. Shi and S. Osher, "A nonlinear inverse scale space method for a convex multiplicative noise model," *SIAM J. Imag. Sci.*, vol. 1, no. 3, pp. 294–321, 2008.
[16] Y. M. Huang, L. Moisan, M. K. Ng, and T. Zeng, "Multiplicative noise removal via a learned dictionary," *IEEE Trans. Image Process.*, vol. 21, no. 11, pp. 4534–4543, Nov. 2012.
[17] Z. Li, Y. Lou, and T. Zeng, "Variational multiplicative noise removal by DC programming," *J. Sci. Comput.*, vol. 68, no. 3, pp. 1200–1216, Sep. 2016.
[18] H. He, W.-J. Lee, D. Luo, and Y. Cao, "Insulator infrared image denoising method based on wavelet generic Gaussian distribution and MAP estimation," *IEEE Trans. Ind. Appl.*, vol. 53, no. 4, pp. 3279–3284, Jul. 2017.
[19] G. Chen, F. Zhu, and P. A. Heng, "An efficient statistical method for image noise level estimation," in *Proc. IEEE Int. Conf. Comput. Vis. (ICCV)*, Dec. 2015, pp. 477–485.
[20] A. Khmag, S. A. R. Al Haddad, R. A. Ramlee, N. Kamarudin, and F. L. Malallah, "Natural image noise removal using nonlocal means and hidden Markov models in transform domain," *Vis. Comput.*, vol. 34, no. 12, pp. 1661–1675, Dec. 2018.
[21] A. Khmag, A. R. Ramli, S. J. bin Hashim, and S. A. R. Al-Haddad, "Additive noise reduction in natural images using second-generation wavelet transform hidden Markov models," *IEEJ Trans. Electr. Electron. Eng.*, vol. 11, no. 3, pp. 339–347, May 2016.
[22] K. Dabov, A. Foi, V. Katkovnik, and K. Egiazarian, "Image denoising by sparse 3-D transform-domain collaborative filtering," *IEEE Trans. Image Process.*, vol. 16, no. 8, pp. 2080–2095, Aug. 2007.
[23] G. Gilboa and S. Osher, "Nonlocal linear image regularization and supervised segmentation," *Multiscale Model. Simul.*, vol. 6, no. 2, pp. 595–630, Jan. 2007.

[24] Y.-Q. Zhao and J. Yang, "Hyperspectral image denoising via sparse representation and low-rank constraint," *IEEE Trans. Geosci. Remote Sens.*, vol. 53, no. 1, pp. 296–308, Jan. 2015.
[25] K. Zhang, W. Zuo, Y. Chen, D. Meng, and L. Zhang, "Beyond a Gaussian Denoiser: Residual learning of deep CNN for image denoising," *IEEE Trans. Image Process.*, vol. 26, no. 7, pp. 3142–3155, Jul. 2017.
[26] K. Zhang, W. Zuo, and L. Zhang, "FFDNet: Toward a fast and flexible solution for CNN-based image denoising," *IEEE Trans. Image Process.*, vol. 27, no. 9, pp. 4608–4622, Sep. 2018.
[27] K. Zhang, W. Zuo, S. Gu, and L. Zhang, "Learning deep CNN denoiser prior for image restoration," in *Proc. Conf. Comput. Visi. Pattern Recognit. (CVPR)*, 2017, pp. 2808–2817.
[28] S. Palakkal and K. M. M. Prabhu, "Poisson image denoising using fast discrete curvelet transform and wave atom," *Signal Process.*, vol. 92, no. 9, pp. 2002–2017, Sep. 2012.
[29] K. Imamura, N. Kimura, F. Satou, S. Sanada, and Y. Matsuda, "Image denoising using non-local means for Poisson noise," in *Proc. Int. Symp. Intell. Signal Process. Commun. Syst. (ISPACS)*, Oct. 2016, pp. 1–6.
[30] A. Ullah, W. Chen, H. Sun, and M. A. Khan, "A modified multi-grid algorithm for a novel variational model to remove multiplicative noise," *J. Vis. Commun. Image Represent.*, vol. 40, pp. 485–501, Oct. 2016.
[31] Y.-M. Huang, M. K. Ng, and Y.-W. Wen, "A new total variation method for multiplicative noise removal," *SIAM J Imag. Sci.*, vol. 2, no. 1, pp. 20–40, 2009.
[32] Z. Jin and X. Yang, "Analysis of a new variational model for multiplicative noise removal," *J. Math. Anal. Appl.*, vol. 362, no. 2, pp. 415–426, 2010.
[33] B. Huang, Y. Mu, Z. Pan, L. Bai, H. Yang, and J. Duan, "Speckle noise removal convex method using higher-order curvature variation," *IEEE Access*, vol. 7, pp. 79825–79838, 2019.
[34] J. Liu, X. C. Tai, H. Huang, and Z. Huan, "A weighted dictionary learning model for denoising images corrupted by mixed noise," *IEEE Trans. Image Process.*, vol. 22, no. 3, pp. 1108–1120, Mar. 2013.
[35] F. Wang, H. Huang, and J. Liu, "Variational-based mixed noise removal with CNN deep learning regularization," *IEEE Trans. Image Process.*, vol. 29, no. 4, pp. 1246–1258, Sep. 2019.
[36] A. Langer, "Locally adaptive total variation for removing mixed Gaussian–impulse noise," *Int. J. Comput. Math.*, vol. 96, no. 2, pp. 298–316, Feb. 2019.
[37] H. Hu, B. Li, and Q. Liu, "Removing mixture of Gaussian and impulse noise by patch-based weighted means," *J. Sci. Comput.*, vol. 67, no. 1, pp. 103–129, Apr. 2016.
[38] A. K. Boyat and B. K. Joshi, "Image denoising using wavelet transform and Wiener filter based on log energy distribution over Poisson-Gaussian noise model," in *Proc. Int. Conf. Comput. Intell. Comput. Res.*, 2015, pp. 1–6.
[39] D. N. H. Thanh and S. D. Dvoenko, "A method of total variation to remove the mixed Poisson-Gaussian noise," *Pattern Recognit. Image Anal.*, vol. 26, no. 2, pp. 285–293, Apr. 2016.
[40] D. N. H. Thanh, D. D. Sergey, and D. V. Sang, "A mixed noise removal method based on total variation," *Inf.-J. Comput. Inf.*, vol. 40, no. 2, pp. 159–167, 2016.
[41] N. Chumchob, K. Chen, and C. Britoloeza, "A new variational model for removal of combined additive and multiplicative noise and a fast algorithm for its numerical approximation," *Int. J. Comput. Math.*, vol. 90, no. 1, pp. 140–161, 2013.
[42] A. Ullah, W. Chen, M. A. Khan, and H. Sun, "An efficient variational method for restoring images with combined additive and multiplicative noise," *Int. J. Appl. Comput. Math.*, vol. 3, no. 3, pp. 1999–2019, Sep. 2017.
[43] C. Li and Q. Fan, "A modified variational model for restoring blurred images with additive noise and multiplicative noise," *Circuits, Syst., Signal Process.*, vol. 37, no. 6, pp. 2511–2534, Jun. 2018.
[44] C. Luo, B. Zhang, Y. Xiang, and M. Qi, "Gaussian-gamma collaborative filtering: A hierarchical Bayesian model for recommender systems," *J. Comput. Syst. Sci.*, vol. 102, pp. 42–56, Jun. 2019.
[45] M. Mafi, S. Tabarestani, M. Cabrerizo, A. Barreto, and M. Adjouadi, "Denoising of ultrasound images affected by combined speckle and Gaussian noise," *IET Image Process.*, vol. 12, no. 12, pp. 2346–2351, Dec. 2018.
[46] Y. Zhan, J. Wu, M. Ding, and X. Zhang, "Nonlocal means image denoising with minimum MSE-based decay parameter adaptation," *IEEE Access*, vol. 7, pp. 130246–130261, 2019.
[47] J. Liu and X. Zheng, "A block nonlocal TV method for image restoration," *SIAM J. Imag. Sci.*, vol. 10, no. 2, pp. 920–941, 2017.
[48] H. Attouch, G. Buttazzo, and G. Michaille, *Variational Analysis in Sobolev and BV Spaces: Applications to PDEs and Optimization*. Philadelphia, PA, USA: SIAM, 2005.
[49] Y. Duan, H. Chang, W. Huang, J. Zhou, Z. Lu, and C. Wu, "The L_0 regularized mumford-shah model for bias correction and segmentation of medical images," *IEEE Trans. Image Process.*, vol. 24, no. 11, pp. 3927–3938, Jul. 2015.
[50] W. Guo, J. Qin, and W. Yin, "A new detail-preserving regularization scheme," *SIAM J. Imag. Sci.*, vol. 7, no. 2, pp. 1309–1334, 2014.
[51] S. Choi and R. Wette, "Maximum likelihood estimation of the parameters of the Gamma distribution and their bias," *Technometrics*, vol. 11, no. 4, pp. 683–690, 1969.
[52] J. H. Gerald and S. S. Samuel, *Statistical Models in Engineering*. Hoboken, NJ, USA: Wiley, 1994, p. 88.
[53] J. Liu and S. Osher, "Block matching local SVD operator based sparsity and TV regularization for image denoising," *J. Sci. Comput.*, vol. 78, no. 1, pp. 607–624, Jan. 2019.
[54] Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli, "Image quality assessment: From error visibility to structural similarity," *IEEE Trans. Image Process.*, vol. 13, no. 4, pp. 600–612, Apr. 2004.
[55] A. K. P. Shanbhogue, N. Fasih, V. R. Surabhi, G. P. Doherty, D. K. P. Shanbhogue, and S. K. Sethi, "A clinical and radiologic review of uncommon types and causes of pancreatitis," *RadioGraphics*, vol. 29, no. 4, pp. 1003–1026, Jul. 2009.

CUICUI ZHAO received the B.S. degree in mathematics from Capital Normal University, Beijing, China, in 2015, and the M.S. degree in computational mathematics from Beijing Normal University (BNU), Beijing, in 2018. She is currently pursuing the Ph.D. degree in applied mathematics with the Renmin University of China. Her research interests include variational methods and image processing.

JUN LIU received the B.S. degree in mathematics from Hunan Normal University, Changsha, China, in 2004, and the M.S. and Ph.D. degrees in computational mathematics from Beijing Normal University (BNU), Beijing, China, in 2008 and 2011, respectively. He is currently an Associate Professor with BNU. His research interests include variational, optimal transport, and deep learning-based image processing algorithms and their applications.

JIE ZHANG is currently a Lecturer with the School of Electronics and Computer Science, University of Southampton (UoS). Before joining UoS, he was a Research Associate with the Algorithms and Complexity Theory Group, Computer Science Department, University of Oxford. Before Oxford, he was a Postdoctoral Researcher with the Mathematical Computer Science Group, Aarhus University. His research interests include interface between computer science and economics, algorithmic game theory, and mechanism design.