

# 开题报告

毕业设计题目： 基于深度学习的高效半色调  
处理技术研究



# 基于深度学习的高效半色调处理技术研究

## 开题报告

### 1 选题意义与可行性分析

#### 1.1 选题意义

在数字印刷、墨水屏显示等领域，半色调处理技术是核心关键技术之一，其核心任务是将连续调图像转换为像素仅含黑白二值的半色调图像，以满足打印机等成像设备的离散性输出要求<sup>[1]</sup>。传统半色调处理方法要么通过人工设计的启发式策略牺牲图像质量换取效率，要么通过迭代式搜索优化算法保证质量却导致计算复杂度激增，难以在图像质量与计算效率之间实现良好均衡<sup>[2]</sup>。

随着深度学习技术在图像处理领域的蓬勃发展，为半色调处理技术的革新提供了新的思路<sup>[3]</sup>。本研究聚焦于基于深度学习的高效半色调处理技术，旨在解决传统方法的固有缺陷，通过构建适配半色调处理特性的深度学习框架、优化网络模型结构及设计动态调度算法，实现高质量与高效率的半色调图像生成。该研究成果可直接应用于嵌入式打印主控芯片等资源受限设备，提升打印图像处理的实时性与成像效果，对推动数字印刷技术的智能化升级、降低硬件设计复杂度具有重要的实际应用价值；同时，针对半色调处理离散性特点的深度学习范式探索，也为底层图像处理领域的相关研究提供了新的理论参考<sup>[4]</sup>。

#### 1.2 可行性分析

**理论可行性：**半色调处理技术已形成有序抖动法、误差扩散法、搜索算法等传统技术体系<sup>[5]</sup>，为深度学习方法的对比研究提供了坚实基础。近年来，深度神经网络在图像复原、超分辨率等底层视觉任务中取得显著突破<sup>[6]</sup>，为其应用于半色调处理提供了成熟的理论支撑。本研究提出的深度强化学习框架、误差反馈网络模型及动态调度算法等核心思路，均基于现有机器学习与图像处理理论<sup>[7]</sup>，具备充分的理论可行性。

**技术可行性：**现有研究已验证深度学习在逆半色调处理等相关任务中的有效性<sup>[8]</sup>，为正向半色调处理提供了技术借鉴。本研究可参考 ResNet、UNet 等成熟网络结构进行定制化设计<sup>[9]</sup>，借助 PyTorch、TensorFlow 等深度学习框架快速实现模型搭建与训练<sup>[10]</sup>。同时，强化学习中的策略梯度方法<sup>[11]</sup>、误差扩散法的核心机制<sup>[12]</sup>等技术，可直接用于解决半色调处理的离散性与模型适配性问题，技术路线清晰可行。

**设备可行性：**本研究可利用 NVIDIA GeForce RTX 5080 等 GPU 设备进行模型训练与实验验证<sup>[13]</sup>，满足深度学习任务的算力需求；实验验证可依托通用 CPU 平台完成模型性能测试，硬件设备能够支撑研究的全流程开展。

**数据可行性：**可通过公开数据集（如 VOC2012、DIV2K）获取大量连续调图像<sup>[14]</sup>，结合传统半色调算法生成半色调图像数据集，或直接使用连续调图像进行无监督 / 强化学习训练<sup>[15]</sup>，数据获取与制备方式灵活，能够满足模型训练与测试的需求。

## 2 国内外研究现状

### 2.1 半色调处理技术研究进展

传统半色调处理技术主要分为三类：有序抖动法通过预设抖动矩阵实现快速阈值分割，计算效率高但图像质量有限，典型代表为 Void-And-Cluster 方法<sup>[16]</sup>；误差扩散法通过扩散量化误差维持局部灰度一致性，图像质量优于有序抖动法，但并行性较差，如 Floyd-Steinberg 算法及后续的结构感知误差扩散法<sup>[17]</sup>；搜索算法将半色调处理视为优化问题，通过迭代搜索生成高质量图像，如直接二值搜索法（DBS）<sup>[9]</sup>，但计算复杂度极高，难以实时应用。

近年来，深度学习方法开始应用于半色调处理：部分工作利用生成对抗网络（GAN）从现有半色调数据集中学习分布特征<sup>[18]</sup>，如基于条件 GAN 的半色调处理方法<sup>[5]</sup>；另有研究通过自编码器结构尝试生成可逆半色调图像，并采用直通估计器（STE）解决二值化操作的梯度消失问题<sup>[19]</sup>；但现有深度半色调处理方法仍存在训练范式不清晰、网络模型未针对半色调特性优化、片外访存竞争等问题<sup>[4]</sup>，尚未实现质量与效率的最优均衡。

### 2.2 基于强化学习的底层图像处理研究进展

强化学习已成功应用于部分底层图像处理任务，通过将像素级操作建模为智能体的动作选择，实现对离散优化目标的有效优化<sup>[20]</sup>。PixelRL 框架提出在每个像素位置设置虚拟智能体，共享全卷积网络策略，通过多轮迭代优化图像质量<sup>[20]</sup>；在半色调处理领域，强化学习的离散动作输出特性与半色调的二值化要求高度契合<sup>[11]</sup>，为解决监督学习与无监督学习的适配性问题提供了新途径，但目前尚未形成成熟的深度强化学习半色调处理框架<sup>[4]</sup>。

## 2.3 半色调评价指标与蓝噪声优化研究进展

现有半色调评价指标中，SSIM 指标在优化时会导致图像平坦区域出现空洞现象<sup>[8]</sup>，需通过对对比度加权等方式修正。蓝噪声特性作为半色调图像的关键要求，现有方法多通过频域滤波或损失函数惩罚低频分量实现<sup>[2]</sup>，但如何在强化学习框架中高效融合蓝噪声约束仍需探索<sup>[21]</sup>。

## 3 研究的基本内容与拟解决的主要问题

### 3.1 研究的主要内容

#### (1) 半色调处理数据集构建与预处理

收集 VOC2012、DIV2K 等公开连续调图像数据集<sup>[20]</sup>，整理并预处理数据；针对模型训练需求，生成恒定灰度图像数据集用于蓝噪声特性优化<sup>[22]</sup>；对数据进行随机裁剪、缩放等增强操作，提升模型泛化能力<sup>[4]</sup>。

#### (2) 基于深度强化学习的半色调处理框架设计

将半色调处理建模为多智能体强化学习问题，每个像素位置对应一个虚拟智能体，输出二值动作<sup>[8]</sup>；设计低方差的策略梯度估计器，优化学习目标<sup>[4]</sup>；提出频域上的各向异性抑制损失函数，保证半色调图像的蓝噪声特性<sup>[8]</sup>；修正现有结构相似度（SSIM）指标，提出对比度加权的结构相似度（CSSIM）<sup>[23]</sup>，优化图像结构细节保留效果。

#### (3) 网络模型与训练策略优化

采用轻量级全卷积网络作为策略网络，输入连续调图像与高斯噪声图以生成非周期半色调图案<sup>[4]</sup>；设计包含 MSE、CSSIM 的复合奖励函数，兼顾色调一致性与结构细节<sup>[8]</sup>；优化训练超参数（学习率、折扣因子等），提升框架收敛速度<sup>[23]</sup>。

#### (4) 模型训练与验证

在 GPU 上完成模型训练，对比传统方法与现有深度学习方法的图像质量（PSNR、SSIM/CSSIM、蓝噪声特性）与计算效率<sup>[9]</sup>；分析梯度估计器、损失函数等组件的有效性，验证框架在多色调处理等场景的可拓展性<sup>[4]</sup>。

### 3.2 拟解决的主要问题

#### (1) 半色调处理的离散性与学习范式适配问题

传统学习范式难以处理二值离散特性，导致模型训练困难。通过多智能体强化学习建模，将像素二值化视为智能体离散动作，利用策略梯度直接优化目标指标期望<sup>[11]</sup>，解决梯度消失与范式不匹配问题。

### (2) 策略梯度估计方差与训练稳定性问题

直接应用 REINFORCE 算法会导致梯度估计方差过大。引入反事实基准与局部期望梯度估计器，利用半色调指标的局部性特点，降低方差并提升训练稳定性<sup>[22]</sup>。

### (3) 蓝噪声特性与结构细节协同优化问题

现有方法难以同时保证蓝噪声特性（频域各向同性）与结构细节完整性。通过频域各向异性抑制损失函数优化蓝噪声特性，结合 CSSIM 指标增强结构细节保留<sup>[4]</sup>。

### (4) 模型计算效率问题

现有深度学习方法参数量大、计算复杂。采用轻量级全卷积网络作为共享策略，利用 GPU 并行计算优势，确保推理速度满足实时处理需求<sup>[6]</sup>。

## 4 总体研究思路及预期研究成果

### 4.1 总体流程图

本研究将从半色调处理相关数据集的收集与预处理起步，数据集是深度强化学习模型训练的基础支撑；接着深入研究现有半色调处理技术及深度强化学习核心理论，为模型设计提供理论依据；随后针对半色调处理的离散性、蓝噪声特性等核心需求，设计基于深度强化学习的半色调处理优化框架，包括策略梯度估计器、损失函数等关键组件；通过迭代训练优化模型参数，最终在硬件平台上完成模型测试与验证，形成完整的技术流程。总体流程图如图 1 所示。

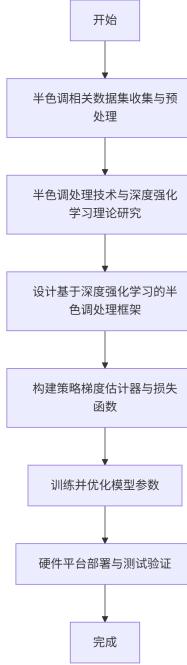


图 1

## 4.2 研究思路

### (1) 数据集的收集与预处理

收集连续调图像数据集作为模型训练与测试的基础数据，选用 VOC2012 数据集、DIV2K 数据集等包含丰富场景与纹理的公开数据集<sup>[14-15]</sup>，涵盖自然图像、纹理图像等多种类型，确保数据的多样性与代表性。

对收集的连续调图像进行预处理操作：首先将图像统一转换为灰度图像，去除颜色通道干扰；然后进行尺寸标准化，裁剪为  $64 \times 64$ 、 $512 \times 512$  等固定尺寸，适配模型输入要求；最后通过随机翻转、灰度扰动等数据增强手段扩充数据集规模，提升模型的泛化能力<sup>[5]</sup>。

预处理过程中，需计算图像的关键指标以评估数据质量，包括用于衡量半色调图像与连续调图像相似度的峰值信噪比（PSNR）、结构相似度（SSIM）等，其计算公式参考现有研究<sup>[8]</sup>：

均方误差：

$$MSE = \frac{1}{mn} \sum_{i=0}^{m-1} \sum_{j=0}^{n-1} [I(i, j) - K(i, j)]^2 \quad (1)$$

峰值信噪比：

$$PSNR = 10 \cdot \log_{10} \left( \frac{\text{MAX}_I^2}{MSE} \right) \quad (2)$$

结构相似度指标：

$$\text{SSIM}(h, c) = l(h, c) \cdot c(h, c) \cdot s(h, c) \quad (3)$$

其中， $l(h, c)$ 、 $c(h, c)$ 、 $s(h, c)$  分别代表亮度、对比度和结构对比系数，综合评估图像的结构相似性。

## (2) 半色调处理技术与深度强化学习理论研究

系统梳理现有半色调处理技术，包括有序抖动法、误差扩散法、搜索算法及现有深度学习方法<sup>[1,4,9]</sup>，分析各类方法的优缺点：有序抖动法计算高效但图像质量有限，误差扩散法能保持局部灰度但存在人工痕迹，搜索算法图像质量高但计算复杂度高，现有深度学习方法存在训练范式不清晰、模型效率低等问题<sup>[3,5]</sup>。

深入研究深度强化学习核心理论，包括马尔科夫决策过程、策略梯度算法、多智能体强化学习等<sup>[11,24]</sup>。重点分析强化学习在离散动作输出、目标函数直接优化等方面的优势，解决半色调处理中离散像素二值化导致的梯度消失问题，为后续模型设计提供理论支撑<sup>[20]</sup>。

## (3) 基于深度强化学习的半色调处理框架设计

### ① 问题建模

将半色调处理建模为多智能体强化学习问题，每个像素位置对应一个虚拟智能体，智能体的动作空间为 0,1（分别对应黑色、白色像素）<sup>[4]</sup>。所有智能体共享同一个卷积神经网络（CNN）作为策略网络，确保模型参数量可控，同时保证并行处理能力<sup>[20]</sup>。

环境状态定义为连续调图像与高斯噪声图的拼接，即

$$s = \text{Concatenate}(c, z) \quad (4)$$

，其中  $z$  为高斯噪声图，用于解耦各输出离散像素，使智能体能够条件独立地进行决策<sup>[4]</sup>。奖励函数设计结合半色调图像质量指标，定义为

$$R(h, c) = -\text{MSE}(\text{HVS}(h), \text{HVS}(c)) + w_s \cdot \text{CSSIM}(h, c) \quad (5)$$

，其中 HVS 为人类视觉系统模型，CSSIM 为对比度加权的结构相似度，兼顾色调一致性与结构相似性<sup>[25-26]</sup>。

## ② 策略梯度估计器设计

针对半色调指标的计算特点，设计低方差的策略梯度估计器。基于 REINFORCE 算法，引入反事实基准（COMA）减少梯度估计方差，进一步推导局部期望梯度估计器 ( $\hat{g}_{LE}$ )，通过遍历智能体动作空间提高梯度估计的准确性<sup>[22,27]</sup>。局部期望梯度估计器公式如下：

$$\hat{g}_{LE} = \sum_a \sum_{h'_a} \nabla_\theta \pi_a(h'_a | c, z; \theta) R(\{h'_a, h_{-a}\}, c) \quad (6)$$

其中， $\hat{g}_{LE}$  为局部期望梯度估计器； $\pi_a(\cdot)$  表示第  $a$  个像素智能体的策略概率； $h'_a$  为智能体  $a$  的候选动作（0/1，对应黑白像素）； $h_{-a}$  为除第  $a$  个智能体外其余所有智能体的动作集合； $R(\cdot)$  为半色调处理的奖励函数， $c$  为连续调图像， $z$  为高斯噪声图。该估计器通过充分利用动作空间的局部信息，实现对目标函数的高效优化<sup>[4]</sup>。

## ③ 损失函数优化

为赋予半色调图像蓝噪声特性，提出频域上的各向异性抑制损失函数 ( $\mathcal{L}_{AS}$ )，通过抑制恒定灰度图像输出的各向异性，确保图像高频段成分均匀分布<sup>[2,4]</sup>。损失函数公式如下：

$$\mathcal{L}_{AS} = \mathbb{E}_{c,z} \left[ \sum_{f_p} \sum_{f \in r(f_p)} \left( \hat{P}_\theta(f) - P_\theta(f_p) \right)^2 \right] \quad (7)$$

其中， $\hat{P}_\theta(f)$  为功率谱估计， $P_\theta(f_p)$  为径向平均功率谱密度，通过该损失函数消除图像中的周期性痕迹与方向偏好<sup>[4]</sup>。

同时，针对原始 SSIM 指标在半色调图像评价中存在的缺陷，提出对比度加权的结构相似度（CSSIM）指标，公式为

$$CSSIM(h, c) = \sigma_c \cdot SSIM(h, c) + (1 - \sigma_c) \cdot 1 \quad (8)$$

其中  $\sigma_c$  为连续调图像的对比度图，避免在图像平坦区域出现空洞现象<sup>[8]</sup>。

## ④ 网络结构选择

选用轻量级全卷积 ResNet 作为策略网络的主干结构，包含 16 个残差块、33 个卷积层，每个卷积核设置 32 个通道，确保模型在保持高性能的同时，具备高效的推理速度<sup>[4,6]</sup>。网络末端设置 Sigmoid 层，输出每个智能体选择“白色动作”的概率，满足伯努利分布要求<sup>[4]</sup>。

#### (4) 模型训练与优化

采用 Adam 优化器对模型进行训练，学习率按余弦退火方案从 3e-4 下降至 1e-5<sup>[28]</sup>。训练过程中，Mini-batch 大小设置为 64，通过随机裁剪图像块进行训练，节省显存资源<sup>[4]</sup>。

训练过程分为两个阶段：第一阶段优化策略梯度与各向异性抑制损失函数，确保模型生成具备蓝噪声特性的半色调图像；第二阶段引入 CSSIM 指标优化，进一步提升图像的结构细节保留能力<sup>[4]</sup>。训练过程中实时监控 PSNR、SSIM、CSSIM 等指标，通过调整超参数  $w_a$ （各向异性抑制损失权重）、 $w_s$ （CSSIM 权重）等，实现图像质量与计算效率的平衡<sup>[4]</sup>。

#### (5) 模型训练与验证

在 VOC2012 测试集、DIV2K 测试集上对训练好的模型进行测试，评估指标包括 PSNR、SSIM、CSSIM 及处理速度<sup>[14-15]</sup>。将测试结果与现有半色调处理方法（如有序抖动法 VAC、误差扩散法 SAED、搜索算法 DBS、基于 GAN 的 cGAN 方法等）进行对比，验证模型的优越性<sup>[5,9,16-17]</sup>。在硬件平台上进行部署验证，选用 CPU、GPU 作为测试硬件，评估模型在不同硬件环境下的推理速度、片外访存量等性能指标，验证模型的工程实用性<sup>[4,29]</sup>。

### 4.3 预期研究成果

1. 提出一套基于深度强化学习的半色调处理优化框架，解决现有深度学习方法在半色调处理中训练范式不清晰、离散动作优化困难等问题，实现对蓝噪声半色调图像的高效生成。
2. 设计的轻量级模型在保持高质量图像输出的同时，处理速度达到现有结构感知半色调处理算法（如 SAH）的 15 倍以上，参数量控制在 300K 以内，满足嵌入式设备实时处理需求<sup>[4,30]</sup>。
3. 模型生成的半色调图像在 PSNR（Näsänen HVS 模型）指标上不低于 31.6dB，CSSIM 指标不低于 0.923，具备良好的蓝噪声特性与结构细节保留能力，各向异性抑制在 -10dB 附近<sup>[4,25]</sup>。
4. 形成相关技术报告与实验数据集，为后续半色调处理技术的研究与应用提供参考，可拓展至多色调处理、彩色半色调处理等相关领域<sup>[31-32]</sup>。

## 5 研究工作计划

起止时间	课题实施计划
2025年11月14日—2025年12月14日	查阅文献、撰写文献综述、外文翻译
2025年12月15日—2026年01月09日	开题报告的撰写和答辩
2026年01月10日—2026年01月31日	熟悉机器学习、深度学习相关知识和算法
2026年02月01日—2026年03月24日	查找对应数据集，数据预处理、模型训练
2026年03月25日—2026年03月25日	中期检查
2026年03月26日—2026年04月10日	程序编写，参数调优，结果与指标的测评与展示
2026年04月11日—2026年05月07日	撰写毕业论文
2026年05月08日—2026年05月17日	评阅老师评阅
2026年05月18日—2026年05月20日	毕业设计答辩

## 参考文献

- [1] LAU D L, ARCE G R. Modern Digital Halftoning[M]. CRC Press, 2008.
- [2] ULICHNEY R. Dithering with Blue Noise[J]. Proceedings of the IEEE, 1988, 76(1): 56-79.
- [3] KIM T H, PARK S I. Deep Context-aware Descreening and Rescreening of Halftone Images[J]. ACM Transactions on Graphics, 2018, 37(4).
- [4] XIA M, HU W, LIU X, et al. Deep Halftoning with Reversible Binary Pattern[C]//IEEE/CVF International Conference on Computer Vision. 2021.
- [5] CHOI B, ALLEBACH J P. Mimicking DBS Halftoning via a Deep Learning Approach[C]// Electronic Imaging. 2022.
- [6] HE K, ZHANG X, REN S, et al. Deep Residual Learning for Image Recognition[C]//IEEE Conference on Computer Vision and Pattern Recognition. 2016.
- [7] RONNEBERGER O, FISCHER P, BROX T. U-Net: Convolutional Networks for Biomedical Image Segmentation[C]//International Conference on Medical Image Computing and Computer Assisted Intervention. 2015.
- [8] WANG Z, BOVIK A, SHEIKH H, et al. Image Quality Assessment: from Error Visibility to Structural Similarity[J]. IEEE Transactions on Image Processing, 2004, 13(4): 600-612.

- [9] ANALOUI M, ALLEBACH J P. Model-based Halftoning Using Direct Binary Search[C]// Symposium on Electronic Imaging: Science and Technology. 1992.
- [10] CHEN T, MOREAU T, JIANG Z, et al. TVM: an Automated End-to-end Optimizing Compiler for Deep Learning[C]//USENIX Symposium on Operating Systems Design and Implementation. 2018.
- [11] SUTTON R S, BARTO A G. Reinforcement Learning: an Introduction[M]. MIT Press, 2018.
- [12] FLOYD R W, STEINBERG L. An Adaptive Algorithm for Spatial Grayscale[C]//Proceedings of the Society for Information Display. 1976.
- [13] NVIDIA. Deep Learning Super Sampling (DLSS) Technology[Z]. <https://www.nvidia.com/en-us/geforce/technologies/dlss/>. 2024.
- [14] EVERINGHAM M, VAN GOOL L, WILLIAMS C K I, et al. The Pascal Visual Object Classes Challenge 2012 (VOC2012) Results[Z]. <http://host.robots.ox.ac.uk/pascal/VOC/voc2012/>. 2012.
- [15] AGUSTSSON E, TIMOFTE R. NTIRE 2017 Challenge on Single Image Super-Resolution: Dataset and Study[C]//IEEE Conference on Computer Vision and Pattern Recognition Workshops. 2017.
- [16] ULICHNEY R A. Void-and-cluster Method for Dither Array Generation[C]//Human Vision, Visual Processing, and Digital Display IV. 1993.
- [17] CHANG J, ALAIN B, OSTROMOUKHOV V. Structure-Aware Error Diffusion[J]. ACM Transactions on Graphics, 2009, 28(5): 1-8.
- [18] GOODFELLOW I, POUGET-ABADIE J, MIRZA M, et al. Generative Adversarial Nets[C]// Advances in Neural Information Processing Systems. 2014.
- [19] BENGIO Y, LÉONARD N, COURVILLE A. Estimating or Propagating Gradients through Stochastic Neurons for Conditional Computation[Z]. <https://arxiv.org/abs/1308.3432>. 2013.
- [20] FURUTA R, INOUE N, YAMASAKI T. PixelRL: Fully Convolutional Network with Reinforcement Learning for Image Processing[J]. IEEE Transactions on Multimedia, 2020, 22(7): 1704-1719.
- [21] XIA M, WONG T T. Deep Inverse Halftoning via Progressively Residual Learning[C]//Asian Conference on Computer Vision. 2019.
- [22] FOERSTER J, FARQUHAR G, AFOURAS T, et al. Counterfactual Multi-Agent Policy Gradients [C]//AAAI Conference on Artificial Intelligence. 2018.
- [23] KINGMA D P, BA J. Adam: a Method for Stochastic Optimization[C]//International Conference on Learning Representations. 2015.

- [24] SUTTON R S, MCALLESTER D, SINGH S, et al. Policy Gradient Methods for Reinforcement Learning with Function Approximation[C]//Advances in Neural Information Processing Systems. 1999.
- [25] NÄSÄNEN R. Visibility of Halftone Dot Textures[J]. IEEE Transactions on Systems, Man, and Cybernetics, 1984, SMC-14(6): 920-924.
- [26] KIM S H, ALLEBACH J. Impact of HVS Models on Model-based Halftoning[J]. IEEE Transactions on Image Processing, 2002, 11(3): 258-269.
- [27] AUEB T R, LÁZARO-GREDILLA M, et al. Local Expectation Gradients for Black Box Variational Inference[C]//Advances in Neural Information Processing Systems. 2015.
- [28] LOSHCHILOV I, HUTTER F. SGDR: Stochastic Gradient Descent with Warm Restarts[C]// International Conference on Learning Representations. 2017.
- [29] HUNTER T M, DENISENKO D, KANNAN S, et al. FPGA Acceleration of Multifunction Printer Image Processing Using OpenCL[R/OL]. 2014. <https://www.intel.la/content/dam/www/programmable/us/en/pdfs/literature/wp/wp-01223-ieee-fpga-acceleration-of-multifunction-printer-image-processing.pdf>.
- [30] PANG W M, QU Y, WONG T T, et al. Structure-aware Halftoning[J]. ACM Transactions on Graphics, 2008, 27(3): 1-8.
- [31] BACCA RODRIGUEZ J, ARCE G R, LAU D L. Blue-noise Multitone Dithering[J]. IEEE Transactions on Image Processing, 2008, 17(8): 1368-1382.
- [32] GUO J M, CHANG J Y, LIU Y F, et al. Tone-replacement Error Diffusion for Multitoning[J]. IEEE Transactions on Image Processing, 2015, 24(11): 4312-4321.