International Journal of Scientific Research in Engineering and Management (IJSREM)
Volume: 07 Issue: 06 | June - 2023 SJIF Rating: 8.176 ISSN: 2582-3930

Gastric Carinoma Detection using Hybrid Model based TL

Turlapati Kavya sri[1], Satuluri Likitha Naga Vigneswari[1], Kommuri Vijay Manikanta[1], Manikonda V Srikar Janardhan Rao[1], Talluri V Lakshmi Bhavani Lalith[1], Venkat Sainadh Reddy[1], Mohammed Uzair Dastagir[2]

[1]School of Computer Science and Engineering, VIT-AP University, G-30, Inavolu, Beside AP Secretariat Amaravati, Andhra Pradesh 522237

[2]Department of Computer Science and Engineering, Amritha Viswa Vidyapetham, Amritapuri, India.

Abstract -*Recent advances in deep learning have led to significant improvements in single image super-resolution (SR) research. However, due to the amplification of noise during the upsampling steps, state-ofthe-art methods often fail at reconstructing high-resolution images from noisy versions of their low-resolution counterparts. However, this is especially important for images from unknown cameras with unseen types of image degradation. In this work, we propose to jointly perform denoising and super-resolution. To this end, we investigate two architectural designs: "in-network" combines both tasks at feature level, while "pre-network" first performs denoising and then super-resolution. Our experiments show that both variants have specific advantages: The in-network design obtains the strongest results when the type of image corruption is aligned in the training and testing dataset, for any choice of denoiser. The prenetwork design exhibits superior performance on unseen types of image corruption, which is a pathological failure case of existing super-resolution models. We hope that these findings help to enable super-resolution also in less constrained scenarios where source camera or imaging conditions are not well controlled.*

Key Words: Super-resolution, Denoising, Deep learning, Image enhancement

1.INTRODUCTION

Single image super-resolution (SR) aims at recovering a highresolution (HR) image from its low-resolution (LR) counterpart, in which high-frequency details have been lost due to degrading factors such as blur, hardware limitations, or decimation.

Early SR approaches were based on upsampling and interpolation techniques [1, 2]. However, these methods are limited in their representational power, and hence also limited in their ability to predict realistic high-resolution images. More complex methods construct mapping functions between low- and high-resolution images. Such a mapping function can be obtained from a variety of different techniques, including sparse coding [3, 4], random forests [5] 6] or embedding approaches [7 8]. Recently, deep learning methods for super-resolution lead to considerable performance improvements 19. 10]. ResNet-like architectures [11] obtain state-of-the-art results for SR tasks while maintaining low computational complexity [12, 13, 14]. Despite these successes, it is still challenging to prevent the amplification of noise during the upsampling steps, which often leads to loss of information and the emergence of artifacts.

Several approaches have been considered to jointly perform super-resolution and denoising. Image restoration can be formulated as an inverse problem. In this approach, the data term for the respective objectives is specific to the respective task. For the prior, a more generic function can be chosen that applies to multiple tasks. For example, using deep learning models [15], employing a denoiser as regularizer [16, 17] or a so-called plug-and-play prior [18]. Furthermore, deep learning approaches have also been considered to combine denoising with a SR model [19]. The authors in [20] propose cascading a denoiser with a SR model so that the output of the denoiser is fed to the SR network.

The pre-network architectural design in our experiments is similar in spirit, but instead of using a fixed convolutional neural network (CNN) denoiser, our design allows for further flexibility regarding the choice of the integrated denoiser. In contrast to previous works, we compare two architectures that allow for further flexibility regarding the choice of the integrated denoiser. This flexibility can be used to incorporate domain knowledge into the network by selecting a denoising technique optimized for the particular type of degradation. For scenarios where domain knowledge is missing, we investigate the generalization capability of different denoisers and the proposed architectural designs. Especially for low-resolution images from cameras in-the-wild, image degradation, such as unseen noise distributions, can lead to artifacts in the reconstructed high-resolution images.

2. DATASET

The experiments use the DIV2K dataset. In our experiments, we consider lowresolution images which have been downsampled by a factor of two. We use the 800 images from the training set for training or fine-tuning the models, and the 100 validation images for evaluation. Closely following [13], we feed to the model 96× 96 RGB image patches extracted from HR images, along with their noisy bicubic downsampled counterparts. We train our models using downsampled image patches corrupted with additive

Gaussian noise. The testing data is corrupted using additive Gaussian noise with the same distribution as during training.

Moreover, Poisson, speckle, and salt-and-pepper noise are used to degrade the testing data for the robustness analysis. As a baseline, we consider two versions of WDSR. First, the publicly available version of WDSR, pretrained on the DIV2K dataset, denoted as "No tuning". Second, WDSR fine-tuned on DIV2K images with added noise, denoted as "No denoiser". The model weights are initialized with the pretrained WDSR weights. Each model is fine-tuned using the mean absolute error (MAE) loss function and the ADAM update rule [27] with an initial learning rate of 10−4 . To avoid overfitting, the fine-tuning procedure is stopped after 100 epochs. Regarding the median and Wiener filter, we use square kernel sizes with side length of five pixels. For the denoising autoencoders, we construct a fully convolutional DAE composed of three convolutional layers with 64, 128, and 256 kernels of size 5×5 respectively. Each layer is followed by a ReLU activation function and max-pooling. DAEs are trained for 80 epochs using the MSE loss function and the ADAM update rule with an initial learning rate of 10−4 . For the integration into the WDSR model, the autoencoder parameters are fixed during the finetuning of the network.

2. METHOD

In this work, we employ the Wide Activation Super-Resolution (WDSR) model [13] as a building block to investigate architectures for joint denoising and superresolution. The original WDSR architecture is shown on top of Fig - 1 It consists of two paths. The main path is on top, consisting of a user-defined number B of residual blocks. Each block consists of two convolutional layers followed by weight normalization [22] and ReLU activation. The lower path is a residual connection. It provides low-level features from the input to the output, which is critically important for SR tasks [14]. Both paths contain a pixelshuffle layer [23], which performs the upsampling for image superresolution. WDSR is sensitive to input images that are corrupted by additive noise. However, we show in this work that it can be paired with a denoiser in two different ways, denoted as "pre-network" and "in-network", which both considerably improve the results.

Fig - 1 shows both architectures. Pre-network (abbreviated pre-net) is shown in the middle. It first passes the image through the denoiser prior to branching into main path and skip connection. This is conceptually similar to the denoiser and SR concatenation by Bei et al. [20]. One potential limitation of this approach is error propagation: if the denoiser removes information that is relevant to super-resolution, it cannot be recovered afterwards. In-network (abbreviated in-net) is shown on the bottom of Fig.1 Here, the denoiser integrates into the residual connection. Hence, the SR model can jointly combine low-level features from the denoised input and high-level features from the noisy input. Both designs are open for the choice of denoiser, which allows to choose a task-specific denoiser, i.e., that performs best on an expected noise distribution. In our experiments, we evaluate three popular denoisers of varying complexity: median filter [24], Wiener filter [25], and denoising autoencoders (DAE) [26].

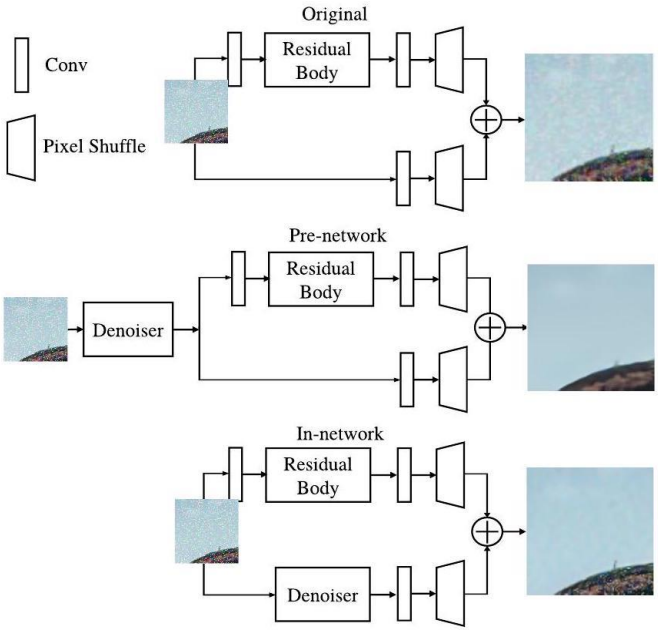

Fig -1: Original WDSR architecture (top) in comparison to the prenetwork architectural design (center) and the in-network architectural design (bottom).

The U-Net network structure is a commonly used image classification network in the processing of medical images. U-network Net's structure is broken into three sections: down-sampling, up-sampling, and skip connection. U-Net is an Encoder-Decoder structure, with the left structure representing the down-sampling process, and the right representing the up-sampling method, which is the Decoder structure. The encoder is in charge of feature extraction, which means that the picture size is decreased by downsampling and convolution to extract some features in shallow layers. Using upsampling and convolution, the decoder receives certain global characteristics in deep layers. The shallow feature map has more specific information (local features), whereas the deep feature map contains more additional context (global features).

3. EVALUATION

The noise in real images can often be approximated as additive Gaussian noise, which we model in this experiment as a fixed power σ 2 for training and testing. This evaluation assumes that noise distribution and strength are known during training. This is a strong assumption for practical cases, but this setup shows the fundamental capability of the models to capture non-ideal images. Table 1 shows the numerical performances for the peak signalto-noise ratio (PSNR). The experiment is performed for additive Gaussian

noise with 0.00 ≤ σ^ 2 ≤ 0.30. For low noise levels, the fine-tuned WDSR ("No denoiser") and in-network with Median filter perform best. In-network with the denoising autoencoder outperforms the competing methods for higher noise levels. A consistent decrease in PSNR can be observed with increasing strength of the noise. This is expected, since image restoration becomes increasingly difficult with increasing noise strength.

Pre-net and in-net are both affected by increasing noise, since the denoiser increasingly removes useful information from the image. In comparison, the original WDSR model ("No tuning") is unable to accurately recover the HR image from its noisy LR counterpart. However, fine-tuning WDSR ("No denoiser") considerably improves the PSNR, for example by about 10 dB for σ ^2 = 0.1. Thus, the fine-tuned WDSR without any integrated denoiser learns to suppress noise and is on par with the best denoiser methods (i.e., in-net with autoencoder and median filter). Hence, WDSR's 16 residual blocks and 32 convolutional filters apparently possess sufficient representational power to jointly learn denoising and SR. However, we will show in the next section that these findings do not generalize to distortions that were not seen during training.

Figure 2 shows a qualitative comparison of the reconstructions of an image patch for a noise strength of σ^2 = 0.1. The in-network architecture yields very similar results for all denoisers. It exhibits comparable results to the fine-tuned WDSR and overall outperforms pre-network. Upon closer examination, prenet suffers from the de- Table 1. Average peak signal-to-noise ratio (PSNR) for varying levels of additive Gaussian noise σ^2 on the low-resolution testing images. Baseline WDSR fine-tuned on the noisy data ("No denoiser") performs best at low noise powers, while the in-net with an autoencoder denoiser achieves best the PSNR values for σ ^ 2 ≥ 0.20. Baseline WDSR without fine-tuning ("No tuning") performs worst. The Average Classification Accuracy of the proposed model with the existing techniques is presented in Fig. 2.

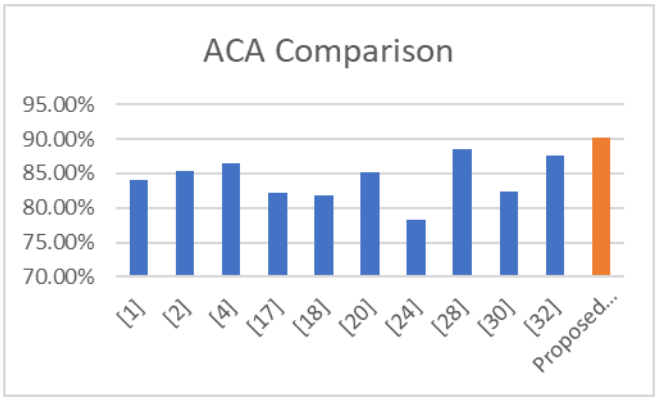

Fig -2:: The comparison of the ACA of proposed model with the existing techniques

For training the Convolution Neural Networks (CNNs), large number of training images are needed to achieve reliable learning. The annotated images for training are highly expensive. Alternatively, data augmentation is used to increase the examples in training dataset using appearance and geometric transformation. Five kinds of augmentation techniques: shift, zoom, rotation, flip and shear are used and applied twice on 9432 cancerous images and 10234 non-cancerous images of 224 × 224 pixels.

3. CONCLUSION

In this paper, we compare two architectures to jointly perform image denoising and single-image super-resolution. We combine the well-known WDSR model [13] with three denoisers that can be chosen depending on the type of degradation. Both networks have specific benefits. The "pre-network" architecture sequentially removes the noise first, and then recovers the high-resolution image. With a suitable denoiser, pre-network generalizes well to unseen noise distributions. However, details in the image are removed and the reconstructions appear slightly over-smoothed. The "in-network" architecture reconstructs the high-resolution image by combining lowlevel features from the denoiser with high-level features from the noisy input. This enables better structure preservation and sharper reconstructed images, but is more sensitive to unseen noise distributions and strength, independent of the chosen denoiser. We hope that these findings are useful toward enabling super-resolution inthe-wild, when camera and image conditions are not fully controlled.

This shows that the proposed network achieves note making results when compared with the existing models. The slice-based test is also performed to evaluate the model. The data of 224 cancerous slice images and 234 non-cancerous slices were taken with the same SCU. The proposed network has achieved the benchmarking results and attained 98.79% classification accuracy. The AUCROC score of the proposed model in comparison with the existing works. AUC scores are computed using the 10-cross validation technique and the graph.

3. REFERENCES

[1] Xin Li and Michael T Orchard, "New edge-directed interpolation," IEEE transactions on image processing, vol. 10, no. 10 , pp. 1521-1527, 2001. [2] Lei Zhang and Xiaolin Wu, "An edge-guided image interpolation algorithm via directional filtering and data fusion," IEEE transactions on Image Processing, vol. 15, no. 8, pp. 22262238, 2006. [3] Roman Zeyde, Michael Elad, and Matan Protter, "On single image scaleup using

sparse-representations," in International conference on curves and surfaces. Springer, 2010, pp. 711730. [4] Jianchao Yang, John Wright, Thomas S Huang, and Yi Ma, "Image superresolution via sparse representation," IEEE transactions on image processing, vol. 19, no. 11, pp. 28612873, 2010.
[5] Jordi Salvador and Eduardo Perez-Pellitero, "Naive bayes super-resolution forest," in Proceedings of the IEEE International conference on computer vision, 2015, pp. 325-333.
[6] Samuel Schulter, Christian Leistner, and Horst Bischof, "Fast and accurate image upscaling with super-resolution forests," in Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2015, pp. 3791-3799.
[7] Radu Timofte, Vincent De Smet, and Luc Van Gool, "Anchored neighborhood regression for fast example-based superresolution," in Proceedings of the IEEE international conference on computer vision, 2013, pp. 1920-1927.
[8] Radu Timofte, Vincent De Smet, and Luc Van Gool, "A+: Adjusted anchored neighborhood regression for fast superresolution," in Asian conference on computer vision. Springer, 2014, pp. 111-126.
[9] Chao Dong, Chen Change Loy, Kaiming He, and Xiaoou Tang, "Learning a deep convolutional network for image super-resolution," in European conference on computer vision. Springer, 2014, pp. 184-199.
[10] Jiwon Kim, Jung Kwon Lee, and Kyoung Mu Lee, "Deeplyrecursive convolutional network for image super-resolution," in Proceedings of the IEEE conference on computer vision and pattern recognition, 2016, pp. 1637-1645.
[11] Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun, "Deep residual learning for image recognition," in Proceedings of the IEEE conference on computer vision and pattern recognition, 2016, pp. 770-778.
[12] Bee Lim, Sanghyun Son, Heewon Kim, Seungjun Nah, and Kyoung Mu Lee, "Enhanced deep residual networks for single image super-resolution," in Proceedings of the IEEE conference on computer vision and pattern recognition workshops, 2017, pp. 136-144.
[13] Jiahui Yu, Yuchen Fan, Jianchao Yang, Ning Xu, Zhaowen Wang, Xinchao Wang, and Thomas Huang, "Wide activation for efficient and accurate image super-resolution," arXiv preprint arXiv:1808.08718, 2018.
[14] Yulun Zhang, Yapeng Tian, Yu Kong, Bineng Zhong, and Yun Fu, "Residual dense network for image super-resolution," in Proceedings of the IEEE conference on computer vision and pattern recognition, 2018, pp. 2472-2481.
[15] Tom Tirer and Raja Giryes, "Super-resolution via imageadapted denoising cnns: Incorporating external and internal learning," IEEE Signal Processing Letters, vol. 26, no. 7, pp. 1080 – 1084, 2019.
[16] Yaniv Romano, Michael Elad, and Peyman Milanfar, "The little engine that could: Regularization by denoising (red)," SIAM Journal on Imaging Sciences, vol. 10, no. 4, pp. 1804-1844, 2017.
[17] Franziska Schirrmacher, Christian Riess, and Thomas Köhler, "Adaptive quantile sparse image (aquasi) prior for inverse imaging problems," IEEE Transactions on Computational Imaging, vol. 6, pp. 503-517, 2020.
[18] Stanley H Chan, Xiran Wang, and Omar A Elgendy, "Plugand-play admm for image restoration: Fixed-point convergence and applications," IEEE Transactions on Computational Imaging, vol. 3, no. 1, pp. 84-98, 2016.
[19] Kai Zhang, Wangmeng Zuo, Shuhang Gu, and Lei Zhang, "Learning deep cnn denoiser prior for image restoration," in Proceedings of the IEEE conference on computer vision and pattern recognition, 2017, pp. 3929-3938.
[20] Yijie Bei, Alexandru Damian, Shijia Hu, Sachit Menon, Nikhil Ravi, and Cynthia Rudin, "New techniques for preserving global structure and denoising with low information loss in single-image super-resolution," in Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops, 2018, pp. 874-881.