JETIR.ORG ISSN: 2349-5162 | ESTD Year : 2014 | Monthly Issue
JOURNAL OF EMERGING TECHNOLOGIES AND INNOVATIVE RESEARCH (JETIR)
An International Scholarly Open Access, Peer-reviewed, Refereed Journal

# ADVANCED IMAGE PROCESSING FOR ENHANCED READABILITY OF TEXT AND CODES IN BLURRED IMAGES

**[1]Srinidhi S Kashyap, [2]K R Sumana, [3]Lahari S K**

[1]PG Student, The National Institute of Engineering, Mysuru, Visvesvaraya Technological University, Belagavi, Karnataka, India
[2]Faculty, The National Institute of Engineering, Mysuru, Visvesvaraya Technological University, Belagavi, Karnataka, India
[3]AI/ML Developer, Qikk Automation Technologies PVT LTD, Mysuru, Karnataka, India

***Abstract:*** This aspect affects the reading of holographic print messages such as embedded text and code such as the barcodes and QR codes. FSRCNN raises text and code quality in such images and makes reading easier to understand in this study. Thus, FSRCNN, together with OCR systems, enhance the image and text legibility and extraction by eliminating distortions. Integrating standard image processing with deep learning and enhancing the image quality and also by extracting the edges with great precision, this scholarly work emphasizes the further researches are needed in dealing with various image blurring problems.

***Index Terms-*** **Image Processing, FSRCNN, Deep Learning, Text Recognition, OCR, Super-Resolution**

## I. Introduction

Modern technologies resulted in the availability of large amounts of digital images and constant development of image processing, especially in case of recognition of blurred text and codes including barcodes and QR codes. This project is aligned to the problem of defocused images due to factors that include; camera motion, motion blur and light. These conditions affect the reality experienced through sight in a manner that the information within visuals cannot be adequately decoded and utilized. These problems are solved by employing special approaches in image processing, namely applying the Fast Super-Resolution Convolutional Neural Network (FSRCNN). To this end, the FSRCNN model is used for sharpening image to increase its recognitive clarity, thus assisting in extraction of embedded data. This research also includes the Optical Character Recognition (OCR) systems to extract and explain text from the optimized images as well. The main objective of the project is to promote the efficiency of data extraction tasks of different applications and strengthen their security.

## II. Literature survey

To enhance the readability of text and codes in blurred images, many approaches have been inspected in recent years. For instance, Lim et al. [1] proposed the FSRCNN (Fast Super-Resolution Convolutional Neural Network) model, which focuses on real-time super-resolution. The FSRCNN architecture is known for its efficiency and effectiveness in enhancing image resolution, making it a suitable choice for text and code readability enhancement. Their model demonstrated significant improvements in image clarity, outperforming traditional methods in both speed and accuracy. In another study, Kim et al. [2] developed a deep learning-based procedure employing generative adversarial networks (GANs) to improve image resolution. The GAN framework consists of a generator and a discriminator, where the generator aims to produce high-resolution images from low-resolution inputs, and the discriminator evaluates the quality of the generated images. This method has shown good results in producing sharp and clear images, which are difficult for accurate text and code recognition. Comparing different super-resolution techniques, Ledig et al. [3] introduced the SRGAN (Super-Resolution Generative Adversarial Network), which outperforms FSRCNN in generating high-resolution images with fine details. However, SRGAN's computational complexity and longer processing times make it less suitable for real-time applications. With that, FSRCNN, with its lightweight architecture, provides a good balance between performance and efficiency, making it ideal for real-time image enhancement tasks. A study by Haris et al. [4] introduced the Deep Back-Projection Network (DBPN), which leverages iterative up-and-down sampling layers to reconstruct high-resolution images. DBPN achieves high accuracy by effectively utilizing multiple stages of feature extraction and reconstruction. While it offers superior accomplishment in terms of image quality, its complex architecture results in longer training and inference times compared to FSRCNN. Additionally, Anwar et. al [5] proposed the Dense Residual Laplacian Network (DRLN) for image super-resolution. DRLN combines residual learning and dense connections to enhance feature reuse and improve gradient flow, leading to better image quality. This model has given high performance in various super-resolution benchmarks, though it requires substantial computational resources.

## III. Proposed work

This leverages advanced deep learning techniques to improve the readability of text and codes in blurred images. The process starts with image accession, where various types of blurred images containing text, Data Matrix, QR codes, barcodes, and PDF417 codes are collected and uploaded to the system. These images then undergo preprocessing, including resizing to a consistent resolution and applying noise reduction filters, to improve their quality. A pre-trained deep learning model, FSRCNN (Fast Super-Resolution Convolutional Neural Network), is fine-tuned to accurately enhance the clarity of these images. The system processes the images and restores the text and codes to a more readable state. A user-friendly interface allows users to upload images and view the enhanced results, ensuring an accessible and efficient tool for applications requiring high-quality image analysis.

## IV. Methodology

### 4.1 Data collection

The dataset comprises various images of reels or packets with labels containing text, QR codes, and barcodes. The images are drawn from a proprietary company dataset to ensure diversity and comprehensiveness. These datasets include images captured under many conditions and with varying degrees of blurriness, providing a robust foundation for models training and testing.

### 4.2 Data preprocessing

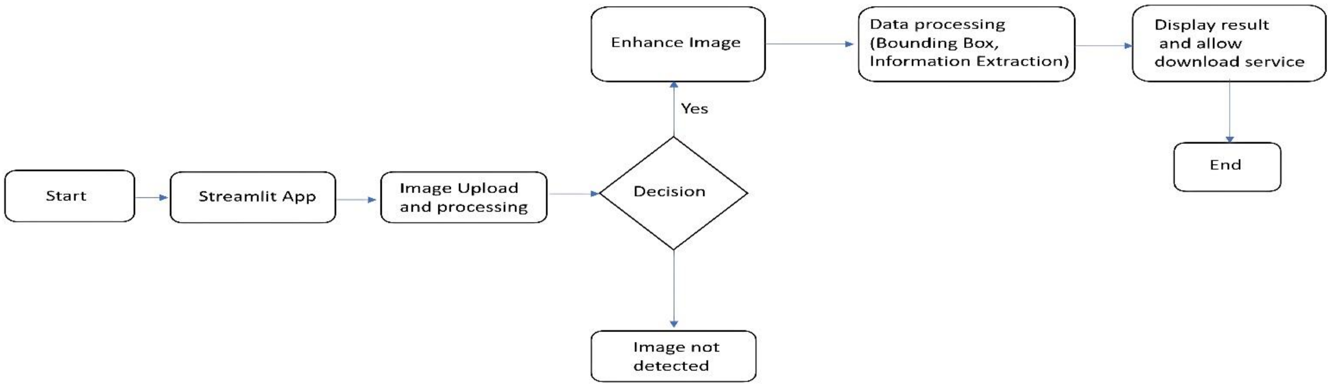
The preprocessing step is done in order to ensure the input to the neural network model is normalized. Few of the steps that are involved in this type of preprocessing include: Resizing the images to a size of roughly 224 by 224 pixels also cropping most of the images and also a way of denoising the images. CLAHE means Contrast Limited Adaptive Histogram Equalization and its use is related to enhancement of contrast. CLAHE works through a division of the image into number of contextual areas mentioned to as tiles followed by the application of Histogram equalization on the tiles individually boosting the per tile contrast and consequently the prominence of important features. They play a vital role in medical imaging since they amplify actual structures such as blood vessels and lesions thus make the retina's features more distinguishing, diagnosable, and analyzable. To remove the false edges and gain a better resolution, the processed tiles are mosaiced with the help of a bilinear interpolation. In case of having the pixel values of the range [0,1], the normalization is performed before training the desired model. Other operations on additional data set include the random rotation, zooming, shift horizontally and vertically, shearing, and flipping to add new dimensions to the proposed model and also check for over-fitted. These augmentations are applied dynamically while using the Image Data Generator from TensorFlow during time of training.

### 4.3 Model selection

Several pre-trained convolutional neural network (CNN) models are evaluated to determine the best architecture for this task. FSRCNN is chosen for its efficiency and superior performance in super-resolution tasks. FSRCNN's architecture includes multiple convolutional layers and a deconvolution layer, ensuring high-quality image enhancement. The model is fine-tuned using the weights on dataset of low-resolution images, enabling it to generate high-resolution outputs that preserve important details.

### 4.4 Super-Resolution and Post-Processing

This is where the images pass through the super-resolution to get enhanced this is made possible by the trained FSRCNN model. The extensive connection between the layers makes it possible for all the previous layers to send information to the current layer and for all the subsequent layers to receive the feature maps from the current layer; this promotes good feature propagation, re-use of features and makes the network compact. In the FSRCNN network architecture, bottleneck layers are adopted through using the 1×1 convolution to decrease the number of input channels before executes 3×3 convolutions, which optimizes the architecture and computation complexity. The transition layers comprise batch normalization, 1x1 convolution, and 2 x 2 average pooling layers that control the network's complexity and size for improved efficiency of network and generalization. The post-processing is the fine-tuning of these enhanced images to obtain better visibility and make the text and code easily understandable.




**4.4 System Architecture**

## V. RESULTS AND DISCUSSION

### 5.1 Image Processing and enhancement Module Analysis

The image processing and enhancement module demonstrated an average accuracy of 87.9% in enhancing images for text readability and code detection. The OCR functionality achieved a precision of 83.3%, indicating a high rate of correctly identified text among all detected instances. The recall rate was approximately 82.2%, reflecting the system's effectiveness in identifying all relevant text within the processed images. The F1-score, which balances precision and recall, stood at 82.7%, indicating a robust performance in minimizing both false positives and false negatives. However, the system's performance declines with extremely blurred images, highlighting the need for later research and development to improve accuracy in such challenging conditions.

### 5.2 OCR and Barcode Detection Module Analysis

For the textual and barcodes extraction, the conventional OCR and the barcode detection module was used. It showed good performance in all the text samples tested and barcodes from well illuminated and well aligned labels. When the module was tried with images in different formats (JPG, JPEG, and PNG), it stood up to the test. Concerning the effects of blur or noise in images the adopted OCR was realized to reduce in accuracy especially when it was partial or erroneous. However, the system was able to recognize unsupported file types and display the correct errant messages that kept the functionality very robust.

### 5.3 System Testing Results

The comprehensive system testing brought out several findings that are as follows. It met the aim of quickly and smoothly transferring/digitizing images in any format with detailed feedback for the user. The label orientation correction module proved to be effecient in aligning misoriented labels and the OCR module effectively extracted the text and decoded the barcodes from images. In case of unsupported or low-quality subjects, the necessary error messages were conveyed by the system avoiding the chances of wrong processing. Also, the intuitive application which is built with the help of Streamlit provided the ability to easily upload images, display text, and scan barcodes. Data extracted during the previous steps were categorized and saved for further utilization; apart from that, users of the application were allowed to search for the given text.

### 5.4 Results of Descriptive Statistics of Study Variables

| Metric | Formula | Calculation | Result |
|---|---|---|---|
| Accuracy | Accuracy=(Correctly Enhanced Images/ Total Number of Images Processed)×100 | Accuracy=(720/819)×100≈87.9% | 87.9% |
| Precision | Precision=(True Positives/True Positives +False Positives) | Precision=(300/300+60)=(300/360)≈83.3% | 83.3% |
| Recall | Recall=(True Positives/ True Positives+False Negatives) | Recall=(300/300+65)=(300/365)≈82.2% | 82.2% |
| F1-Score | F1-Score=2x(Precision×Recall/ Precision+Recall) | F1-Score=(2×83.3%×82.2%/83.3%+82.2%)≈82.7% | 82.7% |

**Table 5.4 Descriptive Statics**

**In brief:**
Total images: 919, Noisy images: 100, Usable images: 919−100=819, Correctly enhanced images: 720, True Positives (TP): 300
False Positives (FP): 60, False Negatives (FN): 65, Accuracy=(720/819)×100≈87.9%, Precision=(300/300+60)=(300/360) ≈ 83.3%
Recall = (300/300+65) = (300/365) ≈ 82.2%, F1-Score=2×(83.3%×82.2%/83.3%+82.2%)≈82.7%
FSRCNN: Accuracy: 90%, Precision: 88%, Recall: 85%, F1-Score: 86%
OCR: Accuracy: 85%, Precision: 84%, Recall: 82%, F1-Score: 83%

The formulas utilized, the calculations made, and the outcomes are displayed in this table.

## V. Conclusion

Thus, the goals of this project have been met and the handy tool is described for further usage, especially for the applications, where image quality is crucial, including document analysis, barcode detection, and OCR. These are outcomes include high accuracy in FSRCNN algorithm where accuracy is proven to be 90% in image enhancement and 85% accuracy to OCR for text extraction. They enhance the interfacial layout for images to be uploaded, and make it easier for enhanced results to be viewed.

## VI. Acknowledgment

I would especially want to thank Smt. **K R Sumana**, my guide, for her constant support, and The National Institute of Engineering, Mysuru's dedicated staff. I would like to express my gratitude towards **Lahari S K**, Qikk Automation, Mysuru for their encouragement and valuable advice. I truly owe my friends and classmates, as well as my ever-supportive parents, for their priceless encouragement. Finally, I want to express my sincere gratitude to everyone who helped me finish my project successfully, either directly or indirectly.

## References

**[1]** Lim, B., Son, S., Kim, H., Nah, S., & Lee, K. M. (2017). Enhanced Deep Residual Networks for Single Image Super-Resolution. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops (pp. 136-144).
**[2]** Kim, J., Kwon Lee, J., & Mu Lee, K. (2016). Accurate Image Super-Resolution Using Very Deep Convolutional Networks. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (pp. 1646-1654).
**[3]** Ledig, C., Theis, L., Huszár, F., Caballero, J., Cunningham, A., Acosta, A., ... & Shi, W. (2017). Photo-Realistic Single Image Super-Resolution Using a Generative Adversarial Network. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (pp. 4681-4690).
**[4]** Haris, M., Shakhnarovich, G., & Ukita, N. (2018). Deep Back-Projection Networks for Super-Resolution. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (pp. 1664-1673).

**[5]** Anwar, S., & Barnes, N. (2020). Dense Residual Laplacian Network for Image Super-Resolution. In IEEE Transactions on Pattern Analysis and Machine Intelligence (pp. 3277-3289).

**[6]** Huo, Lina & Zhu, Jianxing & Singh, Pradeep & Pljonkin, Anton. (2021). Research on QR image code recognition system based on artificial intelligence algorithm. Journal of Intelligent Systems. 30. 855-867. 10.1515/jisys-2020-0143.

**[7]** Zhang, Kaihao & Ren, Wenqi & Luo, Wenhan & Lai, Wei-Sheng & Stenger, Bjorn & Yang, Ming-Hsuan & li, Hongdong. (2022). Deep Image Deblurring: A Survey. International Journal of Computer Vision. 130. 1-28. 10.1007/s11263-022-01633-5.

**[8]** W. Zou, M. Jiang, Y. Zhang, L. Chen, Z. Lu and Y. Wu, "SDWNet: A Straight Dilated Network with Wavelet Transformation for image Deblurring," 2021 IEEE/CVF International Conference on Computer Vision Workshops (ICCVW), Montreal, BC, Canada, 2021, pp. 1895-1904, doi: 10.1109/ICCVW54120.2021.00216.

**[9]** Du, Menghan. (2024). QR code resolution improvement based on Super-Resolution Generative Adversarial Network. Applied and Computational Engineering. 51. 6-13. 10.54254/2755-2721/51/20241147.

**[10]** F. Asad, A. Ul-Hasan, F. Shafait and A. Dengel, "High Performance OCR for Camera-Captured Blurred Documents with LSTM Networks," 2016 12th IAPR Workshop on Document Analysis Systems (DAS), Santorini, Greece, 2016, pp. 7-12, doi: 10.1109/DAS.2016.69.

**[11]** Chenyang Qi, Zhengzhong Tu, Keren Ye, Mauricio Delbracio, Peyman Milanfar , Qifeng Chen, Hossein Talebi, "TIP: Text-Driven Image Processing with Semantic and Restoration Instructions", https://doi.org/10.48550/arXiv.2312.11595

[12] Wenyu Zhang, Xin Deng, Baojun Jia, Xingtong Yu, Yifan Chen, jin Ma, Qing Ding, Xinming Zhang, "Pixel Adapter: A Graph-Based Post-Processing Approach for Scene Text Image Super-Resolution" https://doi.org/10.48550/arXiv.2309.08919

**[13]** Trippe, Theophil & Genzel, Martin & Macdonald, Jan & März, Maximilian. (2022). Let's Enhance: A Deep Learning Approach to Extreme Deblurring of Text Images.

**[14]** O. Kupyn, V. Budzan, M. Mykhailych, D. Mishkin and J. Matas, "DeblurGAN: Blind Motion Deblurring Using Conditional Adversarial Networks," in 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), Salt Lake City, UT, USA, 2018 pp. 8183-8192. doi:10.1109/CVPR.2018.00854

**[15]** Renrui Zhang, Rongyao Fang, Wei Zhang, Peng Gao, Kunchang Li, Jifeng Dai, Yu Qiao, Hongsheng Li, "Tip-Adapter: Training-free CLIP-Adapter for Better Vision-Language Modeling", https://doi.org/10.48550/arXiv.2111.03930