Enhancing Medical Image Quality Using Deep Learning Techniques

Nikita Patil, Sahil K. Shah*, Vidya Kumbhar, T. P. Singh

Symbiosis Institute of Geoinformatics, Symbiosis International (Deemed University), Pune, India

**sahilshahwnr@gmail.com*

2024 First International Conference for Women in Computing (InCoWoCo) | 979-8-3315-1894-3/24/$31.00 ©2024 IEEE | DOI: 10.1109/INCOWOCO64194.2024.10863327

***Abstract*—With the advent of advanced imaging technologies and the growth of digitization, the use of medical images for the precise diagnosis of patients' health conditions is increasing daily. However, medical images generated using traditional X-ray machines produce blurred or low-resolution images owing to poor lighting or machines that require maintenance. Medical practitioners face challenges while diagnosing diseases/injuries in certain body parts due to such low-resolution images. This can lead to flawed diagnoses and misleading treatment. To address these challenges, this study uses state-of-the-art deep learning techniques to convert low-resolution images to super-resolution (SR) images. This study focuses on the training and evaluation of two deep learning models, super-resolution convolutional neural network (SRCNN) and very deep super-resolution (VDSR), on a benchmark dataset of low-resolution X-ray images. Through rigorous experimentation, the hyperparameters pertaining to these two models were carefully optimized to achieve maximum performance. The performance of both models was assessed using metrics such as Peak Signal-to-Noise Ratio (PSNR) and Structural Similarity Index (SSIM). The experimental results show that the VDSR model outperformed SRCNN, achieving a PSNR of 98.3302 and SSIM of 1.0000, compared with SRCNN's PSNR of 43.5399 and SSIM of 0.9964. The higher value of PSNR in VDSR indicates the model's capability to effectively reduce noise and enhance the image quality. Compared with traditional image enhancement techniques, these deep learning models offer more precise and reliable results, enabling faster and more accurate diagnoses. This study provides insights into the strengths and weaknesses of each model, and aims to contribute to the development of better technologies for quicker and more accurate patient diagnosis.**

Keywords- Deep Learning; SRCNN; VDSR; PSNR; SSIM; Medical Imaging.

I. Introduction

This study focuses on improving the clarity of medical images such as chest X-rays, which are crucial for diagnosing diseases such as pneumonia, TB, and lung cancer. Clear images are essential for accurate diagnosis, but sometimes, the resolution of these images is not ideal, making it difficult to detect small details. Traditional methods for enhancing medical images, such as interpolation techniques, often lead to blurred and distorted results that lack the clarity required for an accurate diagnosis [1]. These conventional techniques struggle to preserve fine details and complex structures, making them inadequate for early detection of small abnormalities in medical images. Furthermore, many existing studies fail to utilize advanced deep learning models and lack a comprehensive evaluation using metrics such as PSNR and SSIM, which are crucial for assessing the overall image quality and structure. With the advent of deep learning architectures, images can be converted to a higher resolution with a focus on reducing noise and maintaining the structure. These techniques can significantly improve image clarity and detail, potentially causing significant differences in early disease detection [2]. This study presents a comparative evaluation and analysis of two well-known deep learning models, SRCNN and VDSR, to enhance low-resolution images. The SRCNN was one of the first models for image super-resolution, but VDSR is more advanced owing to its deeper architecture [3]. For model training and evaluation, a benchmark dataset of 1500 chest X-ray images from the National Institutes of Health (NIH) [4] was used. These images were further preprocessed and augmented to make them suitable for training the models. After training, the models were evaluated using PSNR and SSIM, which measure how well the image resolution is enhanced and how well the models preserve the important features of the original images. This study demonstrates how deep learning techniques can enhance medical images, leading to more accurate diagnosis. The main findings of the work undertaken are as follows:

- Medical image quality enhancement by training, experimentation and evaluation of deep learning models: SRCNN and VDSR on low-resolution X-ray images of chest
- Evaluation and comparative analysis of models' performance in enhancing medical image quality
- Metrics such as PSNR and SSIM are used to assess how well the models maintain the important image details and improve the overall image quality.
- Identification of strengths and weaknesses of super-resolution models in deep learning

Further sections present the existing body of knowledge in the literature review, followed by a methodological framework. The results and discussion section highlights the main findings and comparative evaluation of the implemented models, which follow the concluding remarks.

II. LITERATURE REVIEW

This study attempts to identify the challenges in existing techniques for image enhancement. In [5], multiple improved residual networks (MIRN) were used for medical image resolution enhancement. These contributions used public datasets named BSD500 and T91, which contain 591 images. The have marked result as PSNR and SSIM for different datasets were 36.72/0.955, 32.51/0.908, and 29.53/0.896, respectively. [5] introduced wavelet based medical image resolution using a cross-connected residual-in-dense grouped convolutional neural network. They have used 160CT images for training data and 75CT images for testing the model and it gather results for SRCNN are PSNR and SSIM values 40.83/0.9798 and for VDSR 44.31/0.9885. Authors of [6] introduced 'Image Super Resolution Using Deep Convolutional Networks,' they discussed about deep convolutional neural network (CNN) Called SRCNN which works in way that it learns end-to-end mapping from low-resolution to high resolution images. This method is important because it integrates all the steps of super-resolution into a single network, optimizing them together rather than one by one separately, as in traditional sparse coding-based methods. SRCNN shows state-of-the-art performance in image restoration quality and is sufficiently efficient for practical online applications with PSNR and SSIM For SRCNN is 30.29/0.8977, which is a simple and lightweight structure. The study in [3] proposed an enhanced super resolution network. This method eliminates the superfluous elements of the normalization layers from the conventional residual network design. In addition to performing better in single-scale super-resolution, the EDSR presents a multi-scale deep super-resolution system (MDSR) that can support many upscaling factors in a single model, which are examined and verified on the DIV2K datasets. The PSNR and SSIM values 35.12/0.9699[4]. The contributions in [7] used a deep residual feature distillation network with a channel attention mechanism to enhance image quality. By efficiently collecting edge and texture information, the model's residual-within-residual architecture increases computing efficiency and guarantees the integrity of the restored pictures. The results show that DRFDCAN performs better than other models, such as RFDN, in terms of both model simplicity and inference speed, which makes it a good choice for real-time medical applications with PSNR and SSIM for SRCNN as 30.45/0.9020 for SRCNN. The authors of [8] provide a model that uses Generative Adversarial Networks (GANs) to perform better when reconstructing high-resolution images from low-resolution inputs. The generator and discriminator networks in this model use an enhanced squeeze-and-excitation block to selectively enhance important visual characteristics, while suppressing less significant information. The provided GAN-based SR model is a reliable option for improving medical images in the real world, as it produced competitive performance in terms of PSNR and SSIM (43.04/0.996) metrics in addition to better visual effects. The findings of this review show that there is scope for enhancing the performance of the models. It is also observed that if training samples are carefully processed and given in larger sizes, the model's capability to generalize increase[9-11]. Considering these aspects, this study attempted to design and experiment with the SRCNN and VDSR models. This involved data augmentation, rigorous experimentation, and fine-tuning of hyperparameters.

III. METHODOLOGY

Fig. 1 shows the methodology followed for undertaking the experimentation, which involved dataset pre-processing, followed by model training and evaluations. Each phase is discussed below.

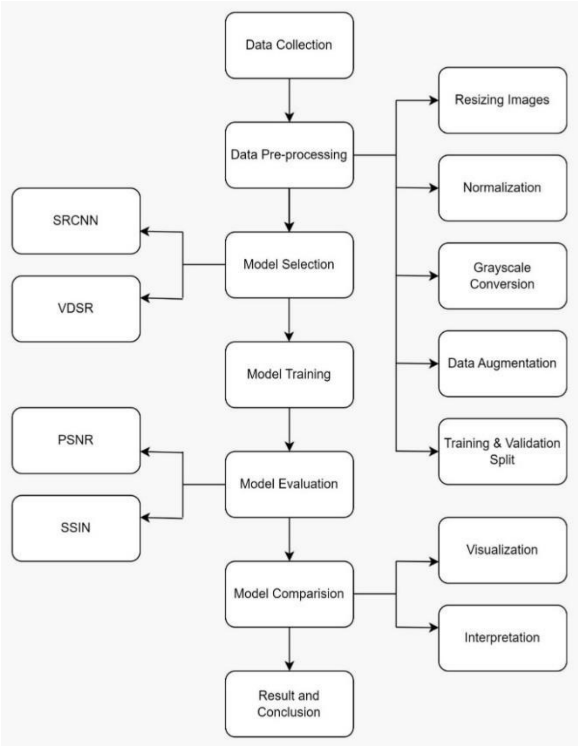

Fig. 1. Methodology

A. Data Collection and Pre-processing

A comprehensive benchmark dataset of over 1500 chest X-ray images from the NIH [4] was selected by considering its importance in diagnosing lung diseases. The images are augmented by using operations like resizing, cropping, zooming, rotation etc. This enhances the size of original dataset further. It ensures that sufficiently large number of consistent input samples are fed to the models while preserving sufficient detail during training phase. The pixel values were normalized to a range of 0 to 1, which helped improve the model's efficiency and stability during training. Effect of data augmentation techniques helped the model to learn by considering various conditions. Finally, the dataset was split into a ratio of 80:20, allowing for an accurate evaluation of the model's performance on new, unseen data.

B. Model Training

For model training, two state-of-the-art CNN architectures dealing with image quality enhancement were chosen. Models SRCNN and VDSR were trained using the pre-processed X-ray image dataset. These models were selected considering their proven effectiveness in previous studies and wide usage in image super-resolution tasks.

SRCNN

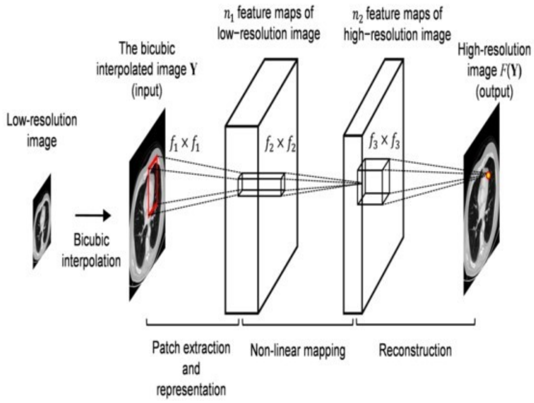

Fig. 2. Architecture of SRCNN

The architecture of SRCNN is relatively simple but effective, consisting of three convolutional layers that aim to reconstruct a high-resolution image from its low-resolution counterpart as shown in Fig.2 [1].

The first layer performs patch extraction and representation. It takes the low-resolution image as input and applies a convolutional operation using a set of filters (usually 64 filters of size 9X9). This layer extracts feature maps from the input image, capturing detailed information like edges and textures. The output is a set of feature maps that represent the key elements of the image. The second layer is responsible for non-linear mapping, which transforms the extracted features into a high-dimensional space. This layer applies 32 filters of size 5X5 to the feature maps obtained from the first layer. It learns the relationship between the low-resolution image patches and their corresponding high-resolution patches, allowing the model to predict the finer details needed for the higher resolution image. The final layer is the reconstruction layer, which aims to reconstruct the high-resolution image from the high-dimensional feature space. This layer uses a single filter of size 5X5 to combine the features learned in the previous layer and produce the output image. The final output is the high-resolution image, which ideally should have enhanced details and reduced blur compared to the input low-resolution image. SRCNN uses a mean squared error (MSE) loss function during training. This attempts to reduce the variance between the reconstituted high-resolution image and the actual high-resolution image(ground truth).

VDSR

VDSR builds on the idea that deeper networks can capture more complex features, leading to better performance in reconstructing high-resolution images from their low-resolution counterparts [2].

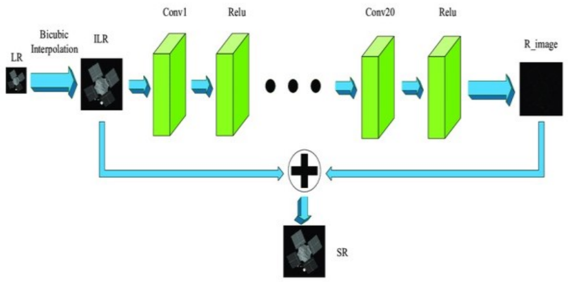

Fig. 3 Architecture of VDSR

VDSR uses 20 layers (Fig. 3), which allows it to capture finer details and more complex features than simpler models like SRCNN. Instead of predicting the high-resolution image directly, it predicts the difference between the low-resolution image and the high-resolution image, making the learning process more efficient and focused on fine details. This models' deep architecture and residual learning make it highly effective for improving image resolution.

C. Experimental Setup

All the experiments were conducted using python framework TensorFlow [12] and utilizing Tesla T4 GPUs for faster computation. The SRCNN model was trained from scratch using random weights. The process included feeding low-resolution images into the model, calculating the difference (loss) between the predicted and actual high-resolution images using MSE loss, and adjusting the model's weights through backpropagation using the Adam optimizer. This process was repeated over several epochs until the model's performance was stabilized. While choosing these hyperparameters, we have experimented using various combinations and over different epochs. VDSR model was also trained using the same approach and similar set of hyperparameters. It is observed that, more processing power and training time is required by VDSR due to its deeper architecture. Using the validation dataset, model's performance was continuously monitored during the training phase. To determine which version performed best the models were evaluated on the validation set following training. The model that obtained the maximum PSNR and SSIM on the validation set was selected as the final model for further transformations [3].

IV. Results and Discussion

For assessing the performance of both the models and check out their effectiveness in enhancing the image quality, following evaluation measures were utilized.

PSNR: It is a metric (Equation 01) which compares improved pictures produced by the model to the original high-resolution image, which is known as the ground truth to evaluate the quality of an image. A greater value of PSNR indicates that the model performed well since it increased the image's accuracy to the original high-resolution source [5].

$$PSNR = 10 \times log_{10}\left(\frac{MAX^2}{MSE}\right) \dots\dots..(1)$$

Where,

MAX - maximum possible value of the image (ideally 255 for 8-bit images)

MSE - mean squared error between the original and enhanced images.

A high PSNR shows models strong capability in precisely identifying and reducing the noise present in the image. This is crucial for medical imaging since even minute errors might compromise the accuracy of a diagnosis.

SSIM: It is based on a similarity between the original and enhanced images. It considers pixel-level differences like edges and textures with a focus on capturing changes in structure of data as shown in Equation 02. The range of SSIM values is -1 to 1. More similarity between the pictures shows the SSIM values closer to 1 [5].

$$SSIM(x,y) = \frac{(2\mu_x\mu_y + C_1)(\sigma_{xy} + C_2)}{({\mu_x}^2 + {\mu_y}^2 + C_1)(\sigma_x^2 + \sigma_y^2 + C_2)} \dots (2)$$

Where:
μ_x and μ_y are the mean pixel values of images x and y.
σ_x^2 $_{and}$ σ_y^2 are the variance of the images.
σ_{xy} is the covariance between the images.
C_1 and C_2 are constants to stabilize the division.

SSIM plays an important part in this attempt in determining how effectively the models captured the structural characteristics of the medical images. High SSIM values indicate models capability to preserve important characteristics necessary for accurate diagnosis, such as tissue borders and diseased structures. We considered both measures as they complement each other; PSNR assesses general image integrity, while SSIM captures important visual structures. Utilization of both the metrics allowed us to choose the best model that maintained both numerical accuracy and perceptual quality. Table-I shows comparative analysis of both the models in terms of PSNR and SSIM values.

TABLE I. COMPARATIVE ANALYSIS OF MODEL PERFORMANCE

Model	**PSNR**	**SSIM**
SRCNN	43.53	0.99
VDSR	98.33	1.00

It is inferred that VDSR outperforms SRCNN in both parameters i.e. PSNR and SSIM. VDSR model performs more effectively at improving the resolution of medical images than the SRCNN model. High score values indicate how well the VDSR model can decrease noise while maintaining image clarity, which makes it ideal for medical applications where accurate diagnosis depends on high-resolution images. VDSR model additionally shows its superior capacity to maintain picture structure by achieving a flawless SSIM score, marginally outperforming the SRCNN models. SRCNN is simpler, using less computing power and offering faster training, making it suitable for situations where quick processing is needed. However, VDSR has a deeper architecture that delivers better image quality, preserving details and structures crucial for medical diagnostics, though it requires more processing power and time. Fig. 4 shows results i.e. high-resolution images obtained by both the models along with input images.

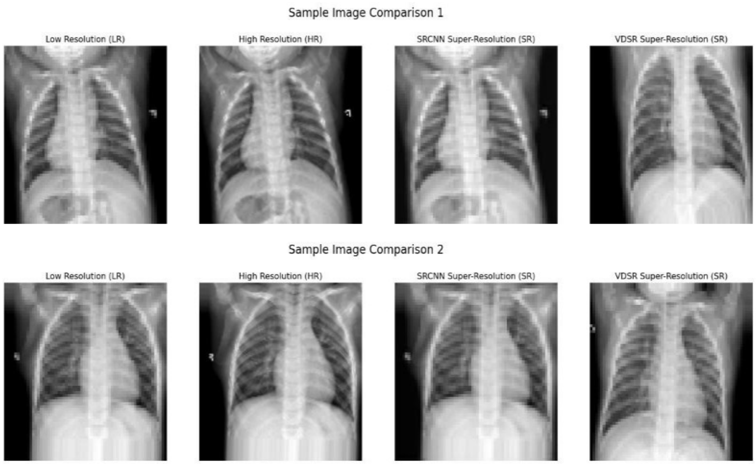

Fig. 4: Results obtained by models

It compares chest X-ray images processed by SRCNN and VDSR models. It shows how a blurry low-resolution image becomes sharper through these models. SRCNN improves the image, making details more visible, but VDSR goes further, producing an image that closely matches the original high-resolution quality. This demonstrates that while both models are effective, VDSR performs better in enhancing image clarity. In order to know the model performance further, error maps were generated as shown in Fig. 5. These maps highlight areas of large variation and the segments where models not performed well. The pixel-by-pixel actual difference between the ground truth and super-resolved pictures was computed in order to produce error maps. A heatmap was then used to display the differences, with warmer colours indicates larger error locations. These error maps help to identify patterns like edge blurring, loss of fine features generated during super-resolution etc. when compared with ground truth images. This helps to identify areas of improvement and knowing strengths and limitations of the model.

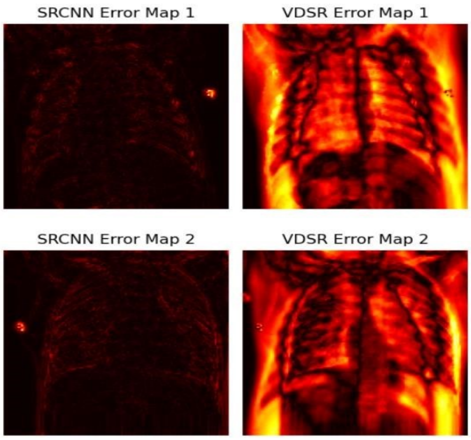

Fig. 5. Generated Error Maps

The maps show the differences between the predicted high-resolution images and the actual images. The SRCNN error maps (on the left) appear darker with fewer noticeable errors, while the VDSR maps (on the right) show more intense error regions, indicating larger errors in some areas. Both the trained models were stored along with their learned weights and biases using Keras. These models were later used for conversion of unseen low-resolution medical images into high-resolution super-resolved images The time required for conversion (inference time) for a batch of images was recorded and analyzed as shown in Fig. 6. The SRCNN model has a lower median inference time, with a range between approximately 0.11 to 0.14 seconds. In contrast, the VDSR model has a higher median inference time, ranging from about 0.14 to 0.17 seconds. This indicates that while VDSR might offer higher performance in terms of image enhancement, it requires more time to make predictions compared to SRCNN.

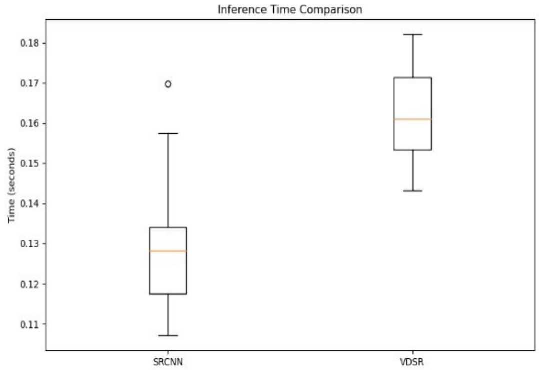

Fig. 6. Analysis of Model Inference Times

Discussion:

This study has attempted to evaluate and compare deep learning models: SRCNN and VDSR models for improving the image quality of chest X-ray images. It is observed that, SRCNN is efficient and fast, making it suitable for situations where resources are limited. VDSR is better for tasks requiring detailed and precise images, especially in critical diagnostics, but it demands more computational power. The choice between these models depends on the specific needs: SRCNN for speed and efficiency, and VDSR for superior image quality. VDSR is recommended for clinical use if computational resources allow, while SRCNN is a good option in resource-limited settings. The obtained results indicate the potential of this study to contribute in enhancing the image quality in medical imaging. The results are comparable with existing studies and it provides a viable solution which can be used by doctors while dealing with low-resolution images. This study can be further extended by implementing a hybrid approach that combines the strengths of both models.

V. CONCLUSION

This study demonstrates the use of deep learning techniques to improve the quality of low-resolution medical images. It shows an analysis of the performance of the deep learning models, SRCNN and VDSR, while handling low-resolution X-ray images. It was observed that both models effectively improved the resolution of the chest X-ray images. Considering the values of PSNR and SSIM, it further shows that VDSR outperforms SRCNN because its use of residual learning results in higher image quality. This is suitable for tasks in which details are crucial for diagnosis. Although the SRCNN is simpler and faster, VDSR is preferred when resources are available. This study highlights the potential of deep learning to enhance medical images, which could lead to better diagnostic accuracy. This study can be further extended by implementing a hybrid approach that combines the strengths of both models. Futuristic research can also focus on assessing the performance of these models using other medical images.

REFERENCES

[1] G. Amaranageswarao, S. Deivalakshmi, and S.-B. Ko, "Wavelet-based medical image super resolution using cross connected residual-in-dense grouped convolutional neural network," *Neural Networks*, 2020, doi: 10.1016/j.neunet.2020.04.012.

[2] S. P. Kannojia and G. Jaiswal, "Cascade CNN framework for low resolution image classification," *Journal of Advanced Information Resilience*, vol. 6, no. 1, 2020, doi: 10.37591/joaira.v6i1.1910.

[3] H. Zhang, P. Wang, C. Zhang, and Z. Jiang, "A comparable study of CNN-based single image super-resolution for space-based imaging sensors," *Sensors*, vol. 19, no. 14, p. 3234, 2019, doi: 10.3390/s19143234.

[4] S. Jaeger, S. Candemir, S. Antani, Y. X. J. Wáng, P. X. Lu, and G. Thoma, "Two public chest X-ray datasets for computer-aided screening of pulmonary diseases," *Quantitative Imaging in Medicine and Surgery*, vol. 4, no. 6, pp. 475-477, 2014.

[5] G. Qiu, L. Zheng, J. Zhu, and D. Huang, "Multiple improved residual networks for medical image superresolution," *Journal of Future Generation Computer Systems*, vol. 115, pp. 35-45, 2020, doi: 10.1016/j.future.2020.11.001.

[6] C. Dong, C. C. Loy, K. He, and X. Tang, "Image super-resolution using deep convolutional networks," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 38, no. 2, pp. 295-307, 2016, doi: 10.1109/TPAMI.2015.2439281.

[7] S. Umirzakova, S. Mardieva, S. Muksimova, S. Ahmad, and T. Whangbo, "Enhancing the super-resolution of medical images: Introducing the deep residual feature distillation channel attention network for optimized performance and efficiency," *Bioengineering*, vol. 10, no. 1332, 2023, doi: 10.3390/bioengineering10111332.

[8] X. Bing, W. Zhang, L. Zheng, and Y. Zhang, "Medical image super resolution using improved generative adversarial networks," *IEEE Access*, vol. 7, pp. 145766-145778, 2019, doi: 10.1109/ACCESS.2019.2944862.

[9] R. Ghosh, S. K. Shah, and V. Kumbhar, "Deep learning-based liver tumor segmentation: A comparative study of U-Net variants for medical imaging analysis," in *2023 Global Conference on Information Technologies and Communications (GCITC)*, 2023, pp. 1-7

[10] S. Zhang, G. Liang, S. Pan, and L. Zheng, "A fast medical image super resolution method based on deep learning network," *IEEE Access*, vol. 6, pp. 62021-62029, 2018, doi: 10.1109/ACCESS.2018.2871626.

[11] Y. Yu *et al.*, "A super-resolution network for medical imaging via transformation analysis of wavelet multi-resolution," *Neural Networks*, 2023, doi: 10.1016/j.neunet.2023.07.003.

[12] M. Abadi, "TensorFlow: learning functions at scale," in *Proceedings of the 21st ACM SIGPLAN International Conference on Functional Programming*, 2016, pp. 1-1.