Jimut Bahan Pal

PW06681131

RESEARCH PROBLEM #2

Image Restoration

For **Prof. Suyash P. Awate**

IIT Bombay

Thursday, May 5, 2022

RESEARCH PROBLEM 2

Jimut Bahan Pal
ID: PW06681131 *[1]

[1]https://jimut123.github.io/

May 5, 2022

1 Task A

For this task, we have used an SR-GAN (2) as shown in Figure 1, to increase the resolution of satellite aerial imagery. There are two versions of the image, one is low resolution image, i.e., 64x64 in size, which is increased to 256x256 size via adversarial training. The proposed statistical model is the standard SR-GAN architecture. We can also use a U-Net type architecture, which will minimize the SNR ratio, i.e., Signal to Noise Ratio, but SR-GAN was a state-of-the-art for this task during 2017. We have used jimutmap (5), a geospatial API, to get the data of Kolkata region by providing a set of latitude and longitudes. Next, we created a pipeline by segregating the 12500 images into a training set comprising 10K images and the test set comprising 2.5K images. We have collected a sample of 30 images for showing how the SR-GAN performs. The samples from the test set of Kolkata data is shown in Figure 2.

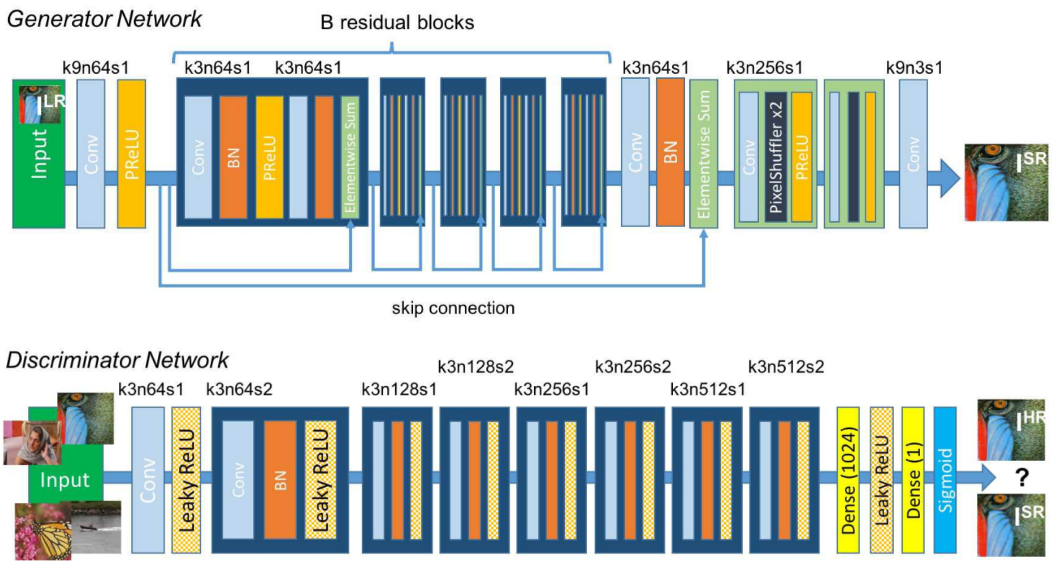


Figure 1: The architecture of SR-GAN (2). Picture is used with proper permission from the original authors.

The SR-GAN network was trained with a learning rate of 2e-04 and a batch size of 8. The Generator is trained for 1000 epochs initially and the images generated are shown in Figure 4 (Top). The corresponding loss is shown in Figure 5a. The minmax Game with Discriminator and Generator is played for 10K epochs and the images generated are shown in Figure 4 (Bottom). The corresponding loss is shown in Figure 5b. The test images were used, 64x64 px images were passed to the SR-GAN Generator and the corresponding results are shown in Figure 3.

2 Task B

We propose a modified U-Net architecture by using MobileNetV2 enocder trained on ImageNet dataset as shown in Figure 7. The results of the experiment are shown in Table 1. We have experimented with different loss functions

*For Prof. Suyash P. Awate

(a) Low Res (b) High Res (c) Low Res (d) High Res (e) Low Res (f) High Res

Figure 2: Samples from jimutmap API for the coordinates of Kolkata. The low resolution images are 64x64 px, and is down-sampled version of the high resolution 256x256px images.

(a) Low Res (b) Predicted (c) High Res (d) Low Res (e) Predicted (f) High Res

(g) Low Res (h) Predicted (i) High Res (j) Low Res (k) Predicted (l) High Res

Figure 3: Samples from the test set showing the low resolution, predicted and the high resolution version. The textures are preserved when we see visually, but some predicted images may have some noise as shown in Figure 3h.

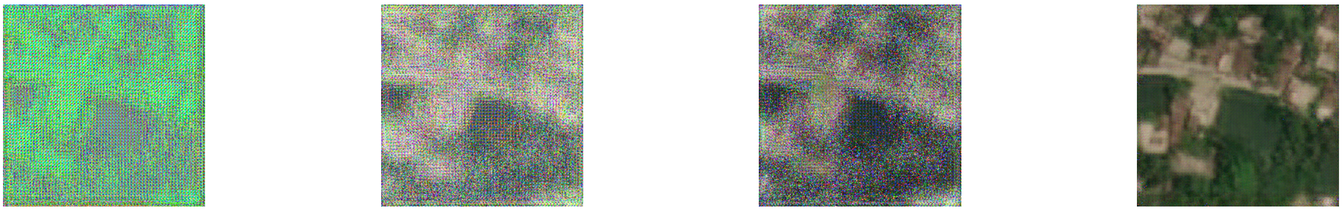
to test the prediction of 5-fold cross validation results. The model is able to successfully extract features for images pertaining to skin lesion from the test set as shown in Figure 6. For training the model we have used Adam optimizer with a learning rate of 1e-05 and we have trained the model for 500 epochs for each fold.

Loss Function	Dice	Jaccard	Precision
Binary Cross-Entropy (4)	92.55 ± 1.29	86.67 ± 1.99	88.31 ± 1.54
Dice Loss (6)	**92.82 ± 1.50**	**87.13 ± 2.36**	88.71 ± 2.00
Tversky Loss (4)	92.56 ± 1.59	86.69 ± 2.49	88.43 ± 2.06
Jacard Loss (4)	92.49 ± 1.70	86.63 ± 2.61	88.39 ± 1.93
BCE Dice Loss (4)	92.80 ± 1.47	87.11 ± 2.27	**88.74 ± 1.86**

Table 1: The various metrics got by applying different loss functions on the Skin Lesion dataset. The values are represented as $\mu \pm \sigma$ across the 5 folds.

3 Task A part - 2

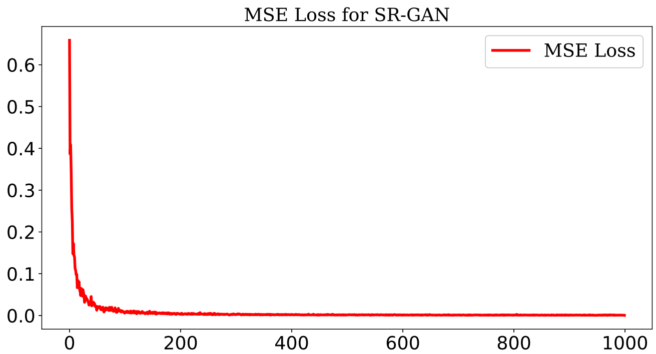
Looks like we need a Bayesian SR-GAN with multiple uncertainties for multiple outputs. We can probably work for a week after joining IIT B and maybe get results after a month. The architecture will be like the original SR-GAN and maybe we can experiment with Variational Auto Encoders in that period. I am not sure if there is such a thing in the current literature. We need to create a novel dataset for the purpose, which will have a single image of low resolution and multiple images on high resolution. This type of problems can be very useful in medical image compression and other allied research fields.



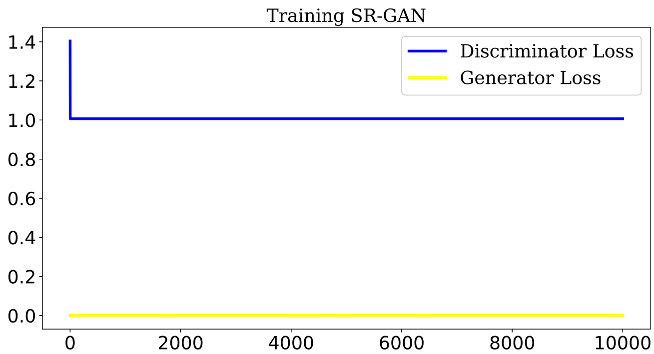
(a) init gen 0 (b) init gen 10 (c) init gen 30 (d) init gen 500

(e) train gen 0 (f) init gen 20 (g) init gen 50 (h) init gen 500 (i) init gen 9500

Figure 4: Top: Initializing the generator for 1000 epochs and shown after respective epochs. Bottom: Training the generator for 10K epochs and shown after respective epochs.



(a) MSE error, initial training.



(b) SR-GAN error, initial training.

Figure 5: Losses of the SR-GAN.

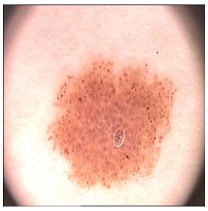
4 Task B part - 2

We can use a probablistic U-Net (1) which is standard for this task. The model achieves multiple segmentation delineations which is common in dealing with medical images where several expert annotators gives different rating. I am proposing something new which was accepted as a MICCAI workshop paper, but note, this might be not as good as probabilistic U-Net.

The proposed method is evaluated with MICCAI-QUBIQ-2020 challenge datasets. QUBIQ-2020 dataset can be got from the following website (`https://qubiq21.grand-challenge.org/participation/`). For this Kidney dataset, out method surpasses the previous state-of-the-art model (`https://qubiq21.grand-challenge.org/evaluation/665a46e2-68f0-45e5-ae89-239c024ac8e3/`). The Kidney have only 1 task unlike other datasets of the QUBIQ challenge where we had other tasks too. Nadam optimizer was used with a learning rate of 1e-05 for this task and Tversky loss was used for this task. The model uses multiple Encoder-Decoder as shown in Figure 9 and the overall model is shown in Figure 8. The outputs for the Kidney dataset is shown in Figure 10 and the results for the validation dataset is shown in Table 2. The main motivation of this model was to create an efficient network which can act as an individual annotator by just using the encoder and decoder pair of the network.

Table 2: The Dice Coefficient generated from validation dataset for 2D image Kidney dataset from MICCAI-QUBIQ-2020 Challenge. Our paper surpasses the previously state of the art for this task.

Model Name	**Dataset**	**Task1**	**Task2**	**Task3**
Proposed	**Kidney**	**95.20**	-	-
U-Net Combined (3)	**Kidney**	83.48	-	-



(a) Original Image

(b) Ground Truth

(c) Prediction Binary CE

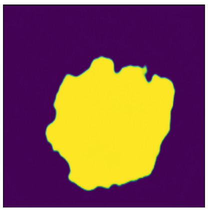
(d) Prediction Binary Dice

(e) Prediction Binary Tversky

(f) Prediction Binary Tversky

(g) Prediction Binary Tversky

(h) Probability Map Binary CE



(i) Probability Map Dice

(j) Probability Map Tversky

(k) Probability Map Tversky

(l) Probability Map Tversky

Figure 6: Two samples of the Skin Lesion dataset are shown, along with the probability map and the binary predictions of all the segmentation masks.

Acknowledgements

The author is grateful to Swathy Prabhu Mj, Ramakrishna Mission Vivekananda Educational and Research Institute, for arranging a machine with an Nvidia A100 GPU (80GB VRAM) and 256 GB RAM, to hasten the research.

References

[1] Simon Kohl, Bernardino Romera-Paredes, Clemens Meyer, Jeffrey De Fauw, Joseph R. Ledsam, Klaus H. Maier-Hein, S. M. Ali Eslami, Danilo Jimenez Rezende, and Olaf Ronneberger. A probabilistic u-net for segmentation of ambiguous images. In Samy Bengio, Hanna M. Wallach, Hugo Larochelle, Kristen Grauman, Nicolò Cesa-Bianchi, and Roman Garnett, editors, *Advances in Neural Information Processing Systems 31: Annual Conference on Neural Information Processing Systems 2018, NeurIPS 2018, December 3-8, 2018, Montréal, Canada*, pages 6965–6975, 2018. URL `https://proceedings.neurips.cc/paper/2018/hash/473447ac58e1cd7e96172575f48dca3b-Abstract.html`.

[2] Christian Ledig, Lucas Theis, Ferenc Huszár, Jose Caballero, Andrew Cunningham, Alejandro Acosta, Andrew Aitken, Alykhan Tejani, Johannes Totz, Zehan Wang, and Wenzhe Shi. Photo-realistic single image super-resolution using a generative adversarial network. In *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 105–114, 2017. doi: 10.1109/CVPR.2017.19.

[3] Jun Ma. Estimating segmentation uncertainties like radiologists, 2020. QUBIQ 2020 Method Description.

[4] Jun Ma, Jianan Chen, Matthew Ng, Rui Huang, Yu Li, Chen Li, Xiaoping Yang, and Anne L. Martel. Loss odyssey in medical image segmentation. *Medical Image Anal.*, 71:102035, 2021. doi: 10.1016/j.media.2021.102035. URL `https://doi.org/10.1016/j.media.2021.102035`.

[5] Jimut Bahan Pal. jimutmap. `https://github.com/Jimut123/jimutmap`, 2019.

[6] Carole H. Sudre, Wenqi Li, Tom Vercauteren, Sébastien Ourselin, and M. Jorge Cardoso. Generalised dice overlap as a deep learning loss function for highly unbalanced segmentations. In M. Jorge Cardoso, Tal Arbel, Gustavo Carneiro, Tanveer F. Syeda-Mahmood, João Manuel R. S. Tavares, Mehdi Moradi, Andrew P. Bradley, Hayit

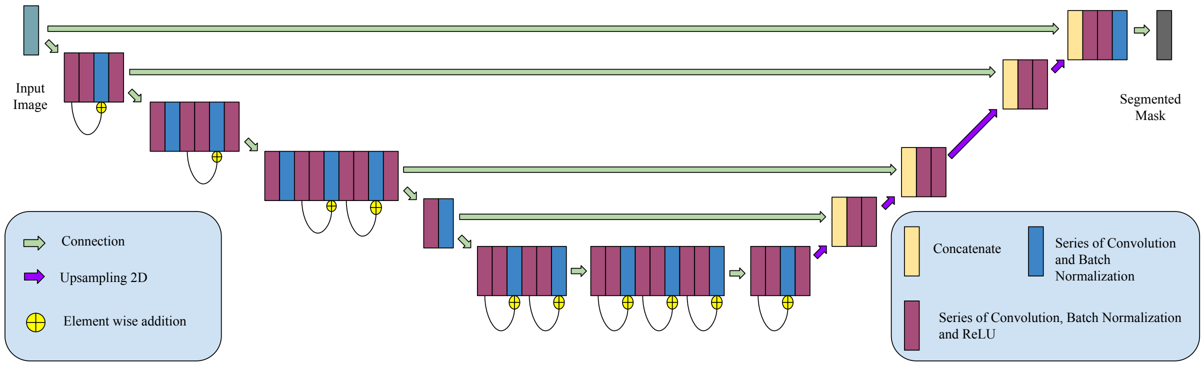


Figure 7: Modified U-Net architecture.

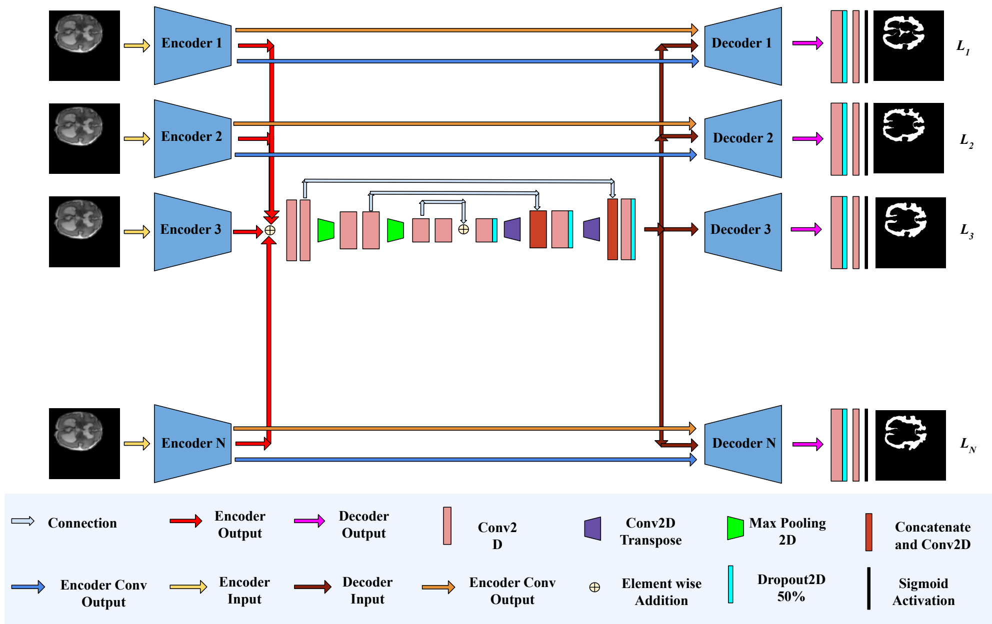


Figure 8: The overall architecture of the proposed model with multiple Encoders and Decoders to holistically segment medical images.

Greenspan, João Paulo Papa, Anant Madabhushi, Jacinto C. Nascimento, Jaime S. Cardoso, Vasileios Belagiannis, and Zhi Lu, editors, *Deep Learning in Medical Image Analysis and Multimodal Learning for Clinical Decision Support - Third International Workshop, DLMIA 2017, and 7th International Workshop, ML-CDS 2017, Held in Conjunction with MICCAI 2017, Québec City, QC, Canada, September 14, 2017, Proceedings*, volume 10553 of *Lecture Notes in Computer Science*, pages 240–248. Springer, 2017. doi: 10.1007/978-3-319-67558-9_28. URL `https://doi.org/10.1007/978-3-319-67558-9_28`.

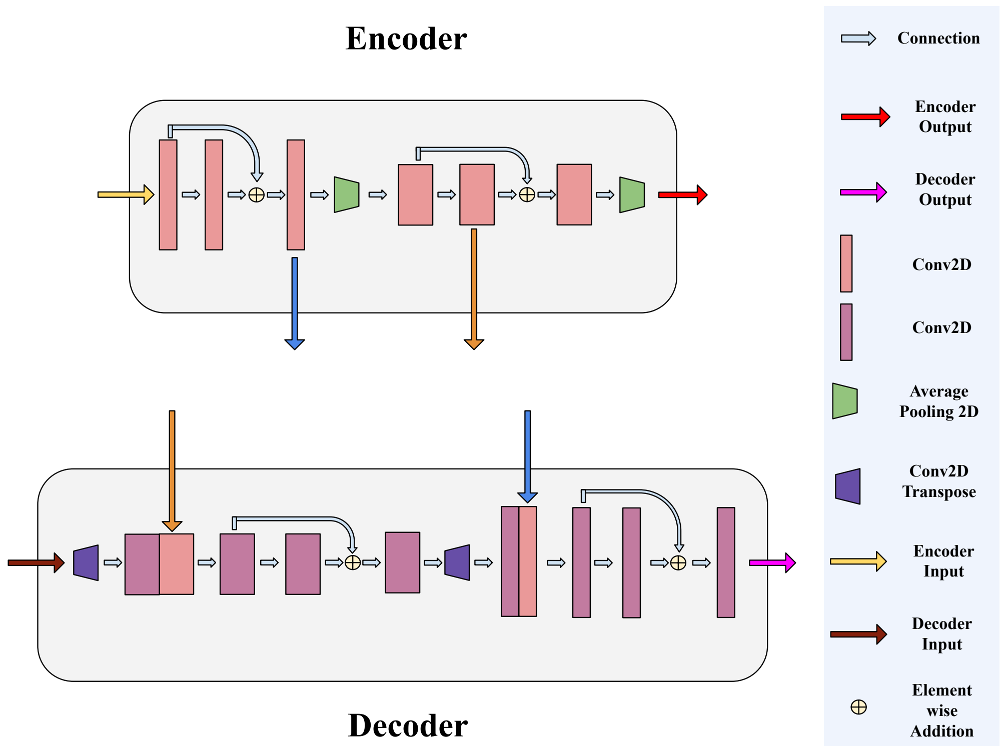


Figure 9: The structures of Encoder (shown above) and Decoder (shown below) block of the proposed architecture.

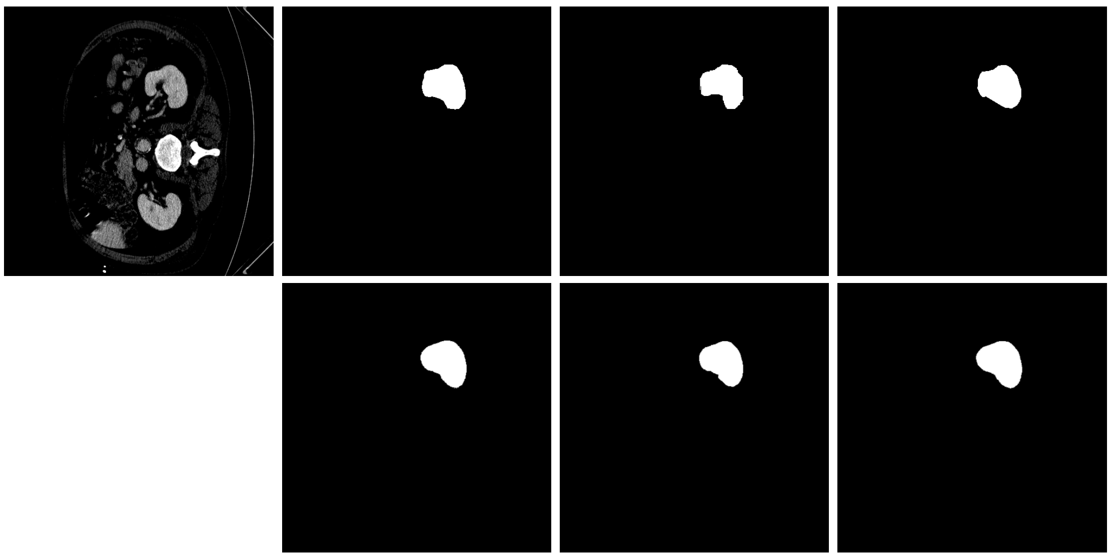


Figure 10: The uncertainty corresponding to the images of Kidney (top) with corresponding predictions (bottom) for individual task of case 23 from the validation dataset.