Cardiac Landmark Detection using Generative Adversarial Networks from Cardiac MR Images

Aparna Kanakatte, Divya Bhatia, Pavan Reddy, J Gubbi and Avik Ghose

TCS Research, Bengaluru, India, aparna.kg, bhatia.divya, pavank.reddy, jay.gubbi, avik.ghose@tcs.com

Abstract and Its Importance

- Cardiac MR is considered the gold standard for the non-invasive characterization of cardiac function, primarily due to its high spatial resolution and 3D capabilities.
- Reliable anatomical landmark detection is an important first step for many medical imaging algorithms.
- A landmark or local feature is a specific image location that serves as a fixed reference. Local features can be corners, edges, or image regions.
- Particularly in medical imaging, these landmark points act as individual anchor points that help in interpreting the image and understanding the location of one anatomical structure in relation to another.
- These landmarks can be used in registration, motion tracking, segmentation, building 3D models, and other applications.
- In cardiology, these landmarks facilitate robust and precise functional and structural analysis of the heart also helps in accurate surgical pre-planning.
- A recent study showed a detailed manual analysis and annotation by an expert can take 9 to 19 minutes.
- The need for a robust, reliable, and accurate system motivated us to explore GAN for landmark detection on both long and short-axes imaging views with great consistency.
- This is STACOM LV landmark detection challenge dataset with 100 patient images acquired in both the long and short-axis views.
- The dataset had 6 distinct landmark annotations; 2 from Mitral valve, 2 from RV insert points and one each from base and apex central axis points. These landmark points are necessary to build a 3D left ventricle model.

Equations

Generator

$$G_L = A_L + \beta \times L_L$$

$$A_L = MSE(I, GP_h)$$

$$L_L = Huber(GT_h, GP_h, \delta = 0.4)$$

Discriminator

$$SDL = 0.5\times(DRL+DFL)$$

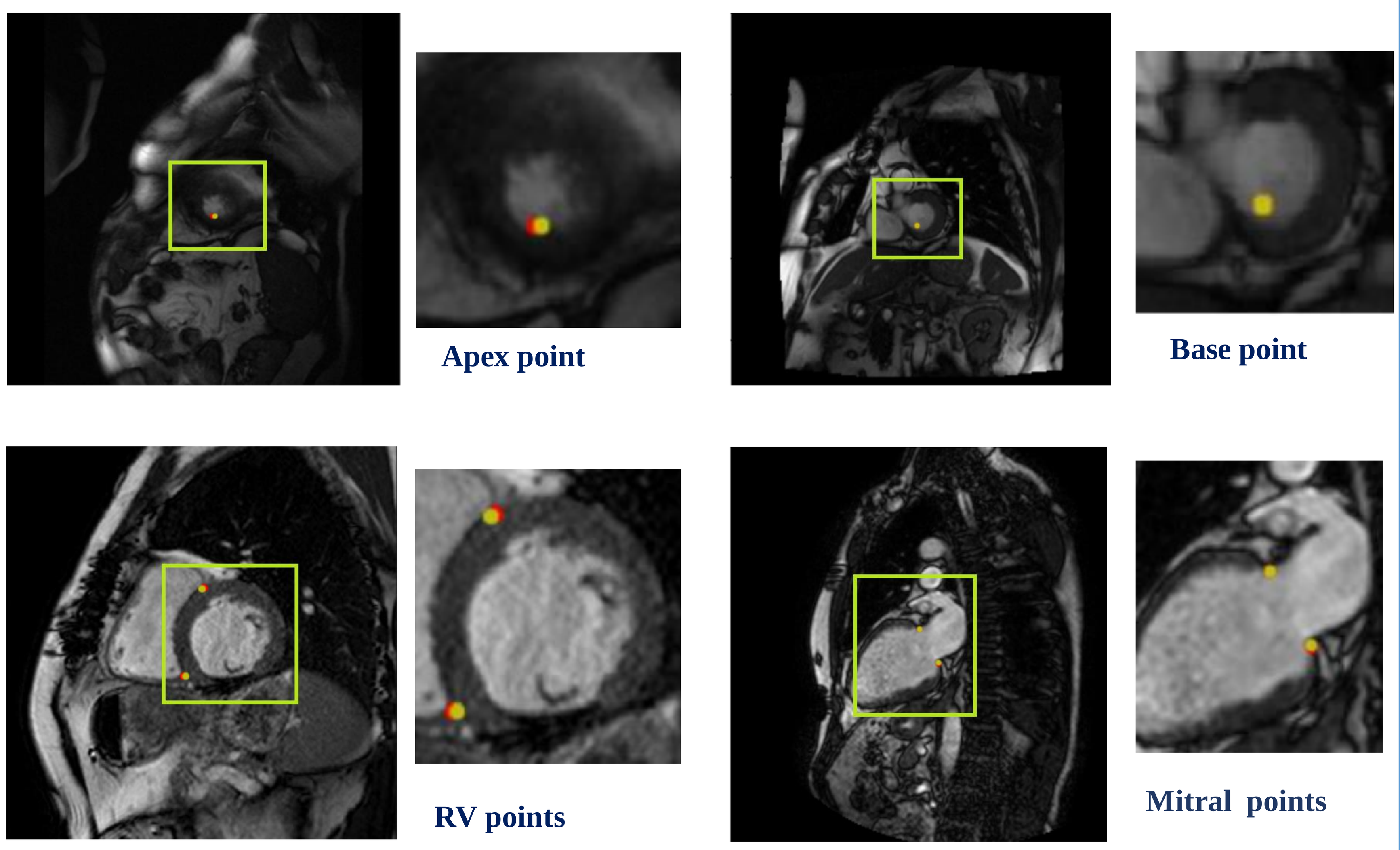
$$DRL = MSE(I, GT_h) \qquad DFL = MSE(I, GP_h)$$

$$FPL = \alpha \times MSE(P_1,P_2) + MAE(P_1,P_2)$$

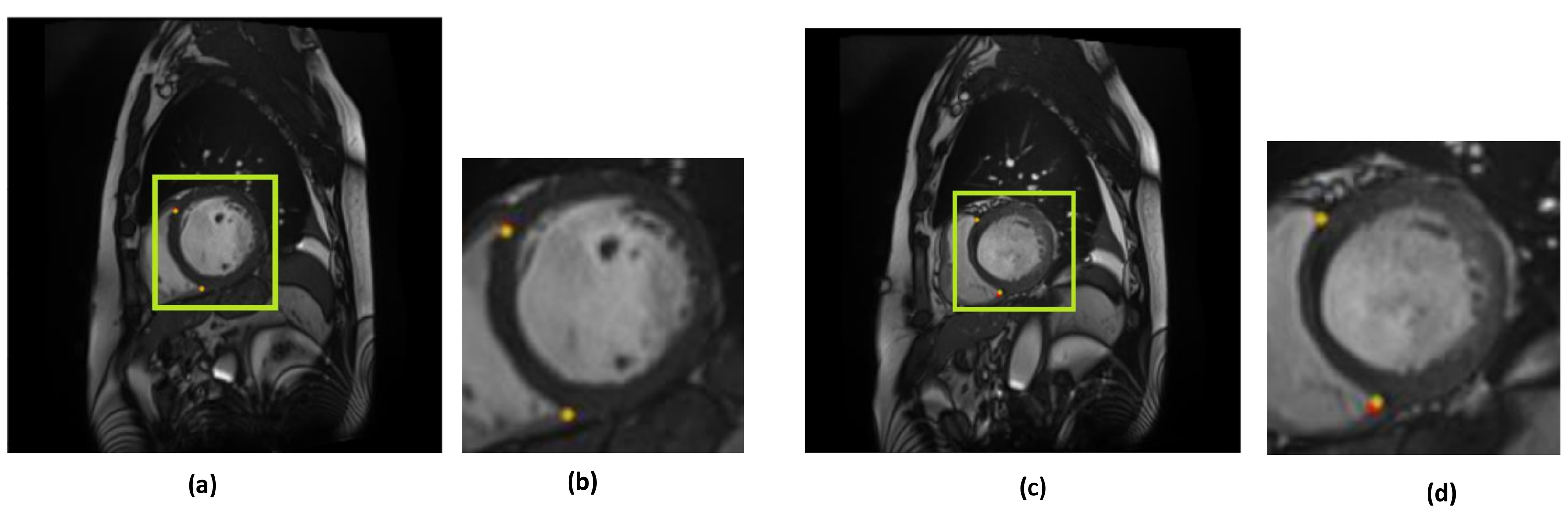
$$MDL = SDL+FPL$$

G_L is **Generator Loss**, A_L is **Adversarial Loss**, L_L is **Learning Loss**, β is **Intensity Parameter**, I is **Input Image**, GP_h is **Generator Predicted Heatmap**, GT_h is **Ground Truth Heatmap**, SDL is **Standard Discriminator Loss**, DRL is **Discriminator Real Loss**, DFL is **Discriminator Fake Loss**, FPL is **Foreground Pixel Loss**, α is **Regularizer** and MDL is **Modified Discriminator Loss.**

Performance Analysis

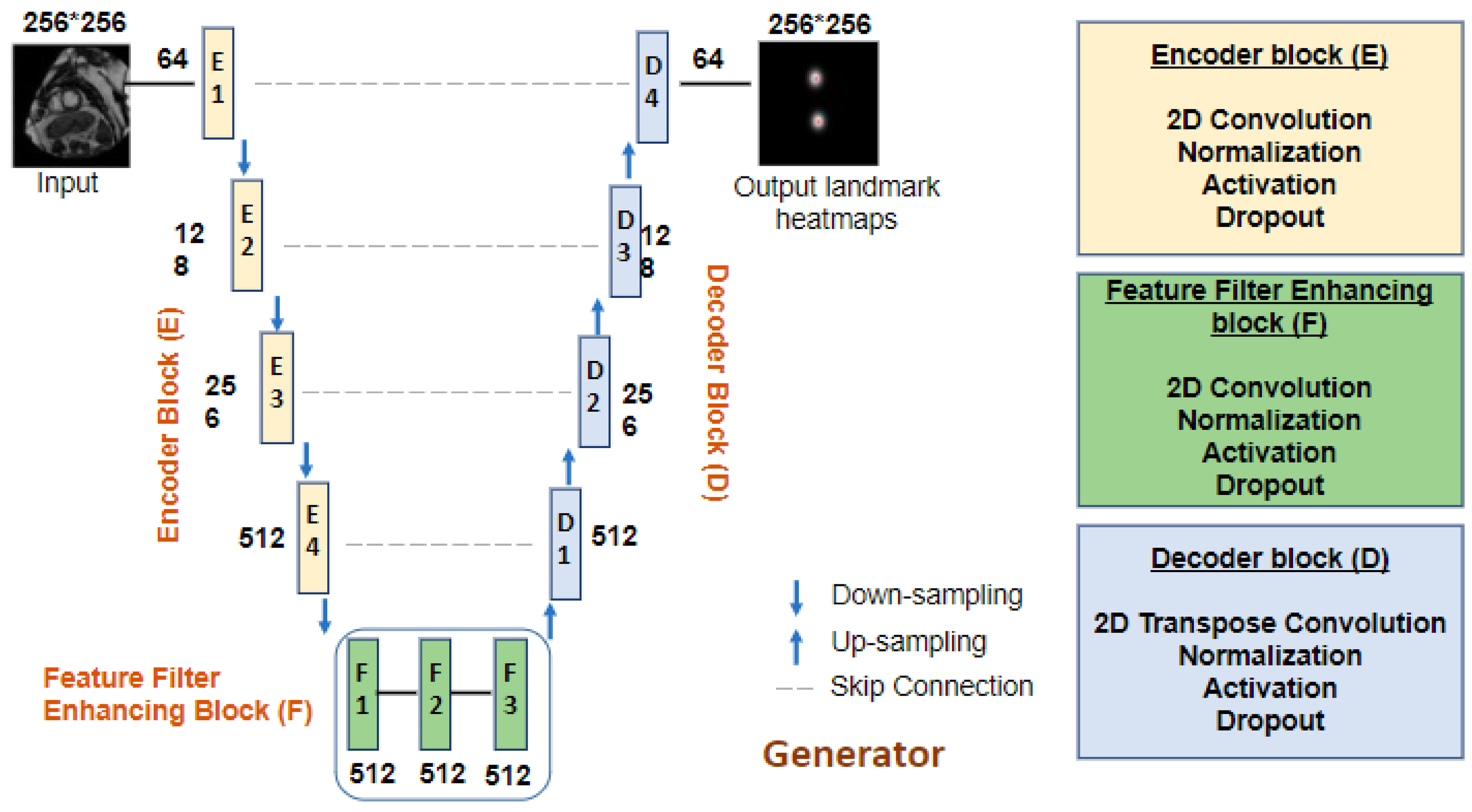


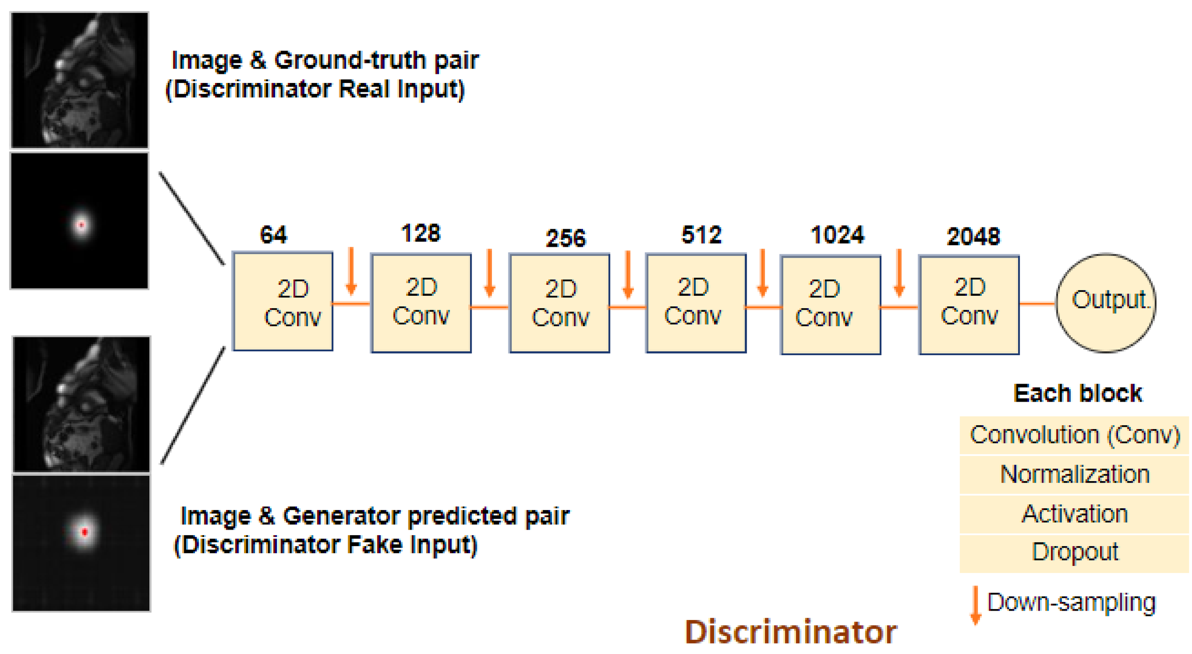
Results of **landmark detection.**



Results of RVI blind-testing with **ground-truth (red)** and **predicted point (yellow)**.The zoomed region are displayed next to the image for better visualization of landmarks. Both have a 1-pixel error

Architecture





Flow chart of the proposed **Detection-Gan Block**

Loss Function and Results

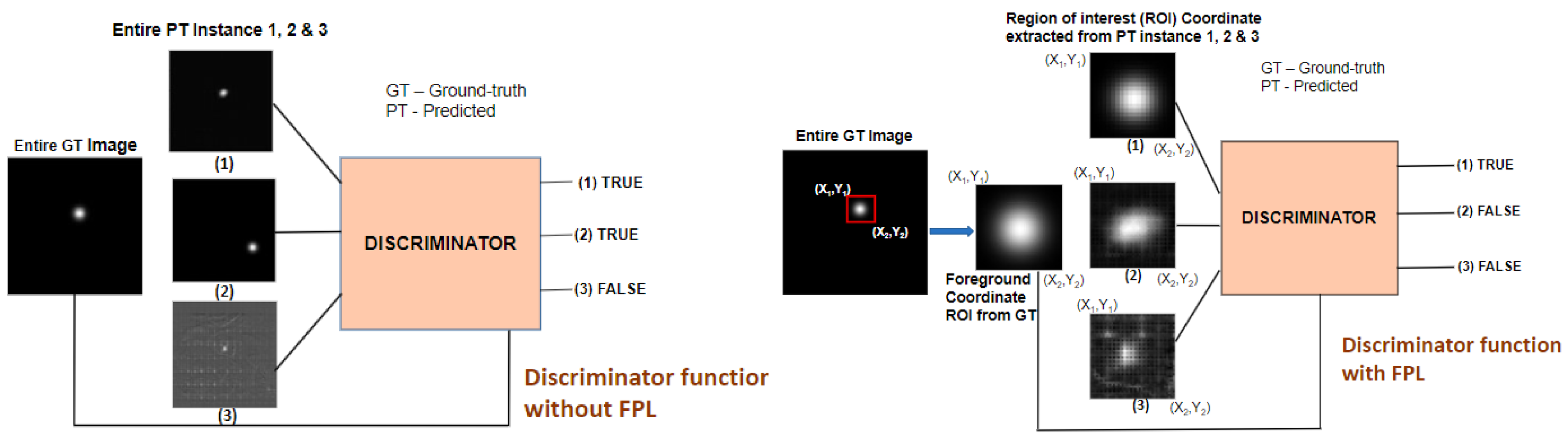


Illustration of **Foreground Pixel Loss (FPL)** Function

Table1: Average error measures (in pixels) for landmark detection on training datasets. Figures indicate **mean and standard deviation.**

	ACA	BCA	MV	RVI
Mahapatra	2.2±1.2	3.0±1.6	9.3±2.5	7.4±2.6
Lu	-	6.2±4.0	3.5±5.6	7.9±11.5
Proposed	**1.8±1.2**	**1.6±1.5**	**3.0±1.4**	**2.8±1.5**

- We have also tested our algorithm on blind tested data of **ACDC dataset**
- We observed **mean error of 2.3 pixels** with a **standard deviation of 1.8 pixels** when tested on 1000 images with varied pathologies.

Table 2: Average **successful detection rate in %** for each landmark.

	ACA	BCA	MV	RVI
≤ 2 Pixel error	70	75	62	72
≤ 3 Pixel error	85	85	88	86
≤ 5 Pixel error	95	90	98	96

Conclusion

- By incorporating Foreground pixel loss, we have developed a unique encoder-decoder architecture and a generative mechanism for image translation that can be applied to any small object detection problem.
- Uses faster and efficient Detection-Gan and FPL loss to integrate both landmark localization and global context.
- Since our design is minimalist and simple, our network generates a single heatmap image containing N heatmaps if multiple landmarks are present in the image instead of generating separate heatmap image for every landmark.
- The method's robustness is shown by achieving the reduced mean error when blind-tested on the ACDC dataset.
- The future work of our research includes extending this to detect 3D landmarks.