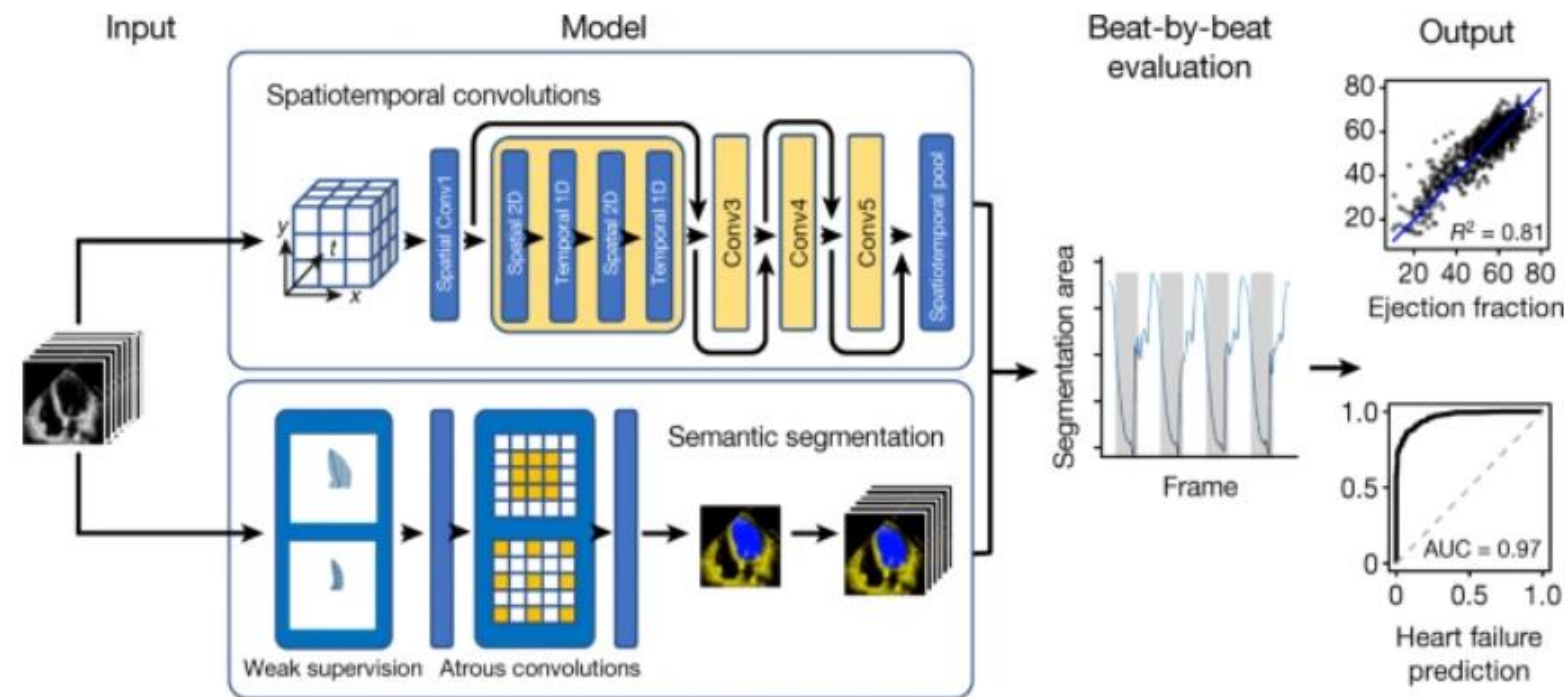


Basudha Pal, Arushi Sinha, Kunal Kotkar

Problem Statement

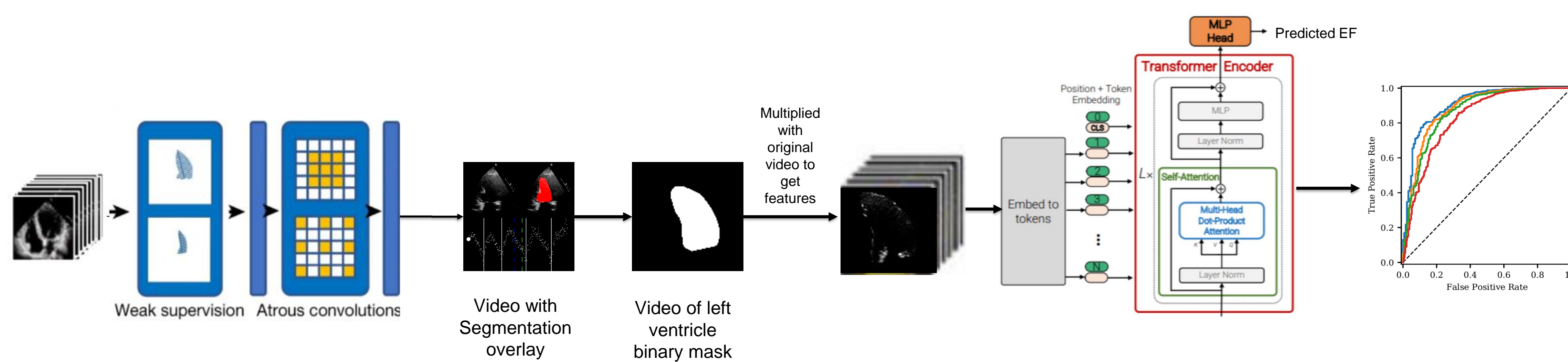
EchoNet is one of the most popular techniques which uses spatiotemporal convolutions and semantic segmentation to predict EF. While this technique provides respectable results, the use of transformers in predicting the ejection fraction from echocardiographic videos has not been explored in great detail. We explore the usage of video vision transformers guided by segmentation results.



Overall Approach

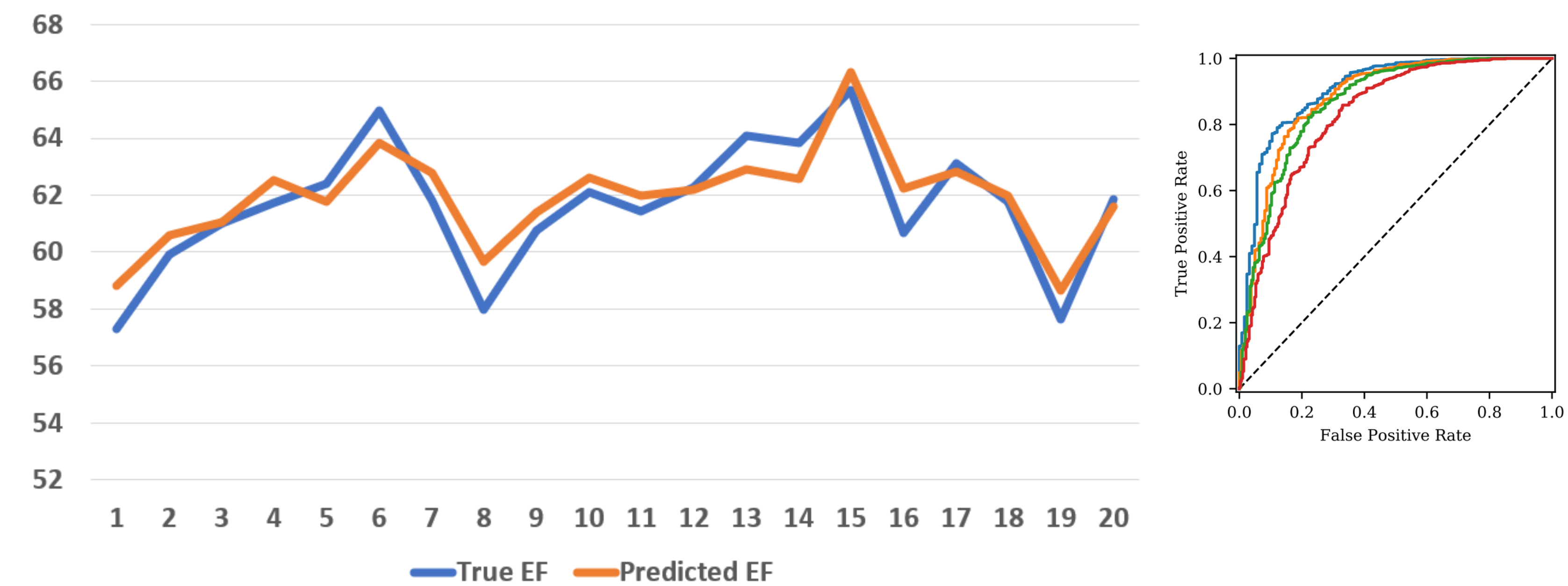
- The videos from the EchoNet Database were fed into a DeepLabV3 based segmentation network to create a binary mask of the left ventricle.
- These masks were multiplied with the original videos to obtain the video information of only the left ventricle.
- This was the input to the video vision transformer which was then trained to predict Ejection Fractions “attending” to the Left Ventricular Video information.
- This left ventricular segmentation guided transformer for EF prediction is a novel approach in our work.

Solution



Results

Few Random Comparisons of EF



Threshold: 35 AUC Score: 0.91 MAE: 7.7%

