

# Supplement material for ICCVW CVAMD paper: Robust MSFM Learning Network for Classification and Weakly Supervised Localization

Komal Kumar<sup>1</sup>, Balakrishna Pailla<sup>2</sup>, Kalyan Tadepalli<sup>2,3</sup>, Sudipta Roy<sup>1,\*</sup>

<sup>1</sup>Artificial Intelligence & Data Science, Jio Institute, Navi Mumbai-410206, India

<sup>2</sup>Reliance Jio - Artificial Intelligence Centre of Excellence (AICoE), Hyderabad, India.

<sup>3</sup>Sir HN Reliance Foundation Hospital, Mumbai – 400004, India

Komal2.Kumar@jioinstitute.edu.in; Balakrishna.Pailla@ril.com;

Kalyan.Tadepalli@rfhospital.org; sudipta1.roy@jioinstitute.edu.in;

## 1. Additional Results

This provides additional results on the MURA dataset for fracture detection and the tumor dataset for tumor detection on X-ray images. The image visualization of the model in each step is shown in Figures 1 2 3. The Figure 1 consists of six different columns namely original image, feature map-based augmentation (FMA), area localization, and looking closer (LC) to it as discussed in Main-Branch, we have used six windows for informative parts (IP) based on our experiments shown as discriminative features, combine all IPs to produce final informative region (FIP) shown as look closer to see the tumor, and final column talks about the possible region in bounding box based on our weakly supervised methods. Similarly, Figures 2 3 consists of four columns namely Input, Look closer, discriminative features, and look closer to see fracture as fracture detection column. FMA removes the no informative region shown in the second column of 1. The final IPs were biased towards the artifacts but after FMA no box only consists of non-informative regions for each part. After FMA, we go to the first main branch which localized the region based on its levels. Which consists of the area of interest, for example, the region of fracture for the fractured image. Then cropped images from the main branch are shared with the object branch where all the six windows corresponding to the class label are proposed. These all windows are corresponding to the scale value which means that the higher the value, the higher the chance of informative region, we combine all the information parts by giving higher weightage to the window corresponding to the higher scale value. Finally, we pasted FIP in the original image based on the coordinates from the main branch and we get the possible informative area for that class. For fracture or tumor class we have reported the results in the column of fracture detection and tumor detection.

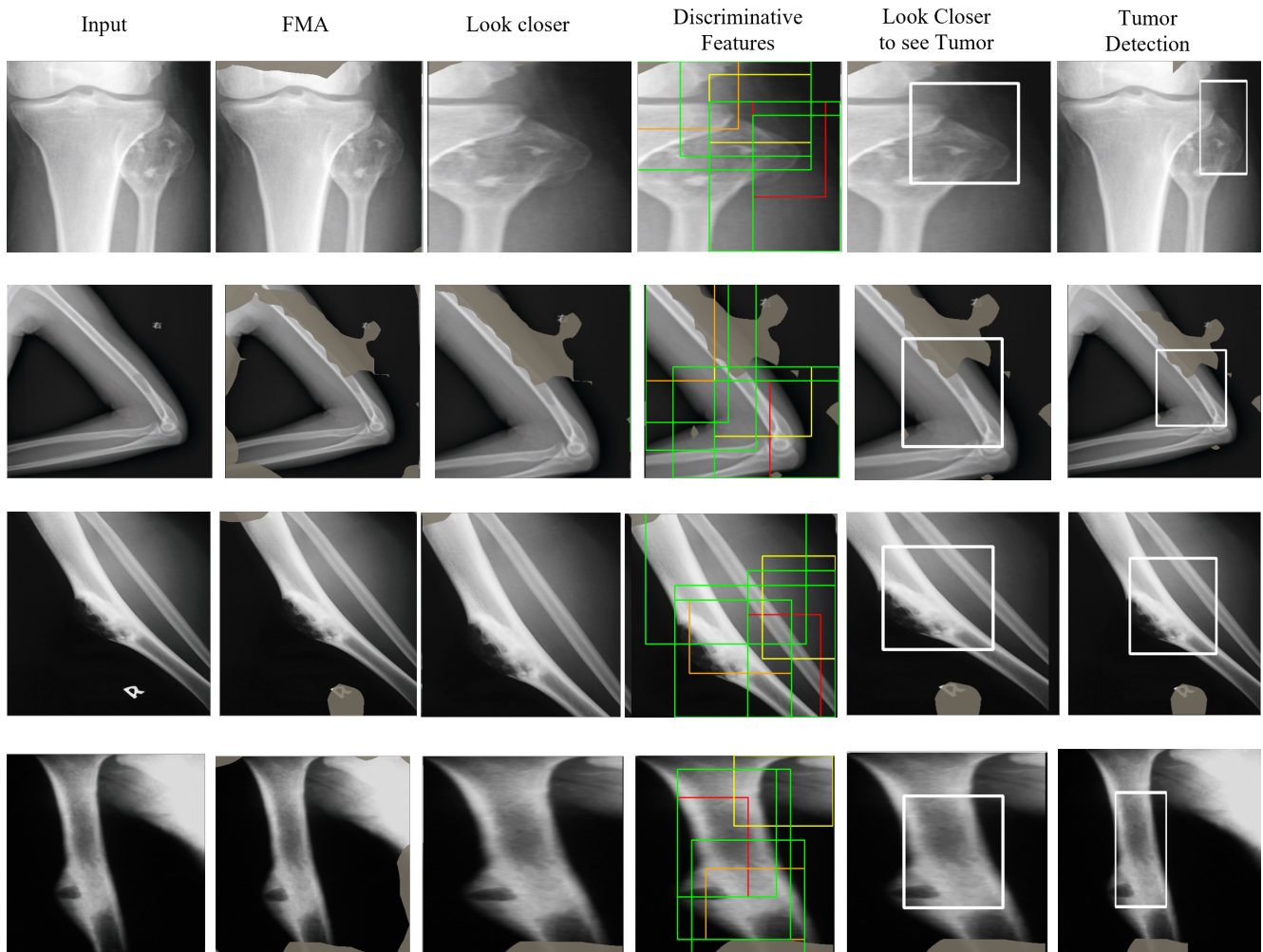
## 1.1. Tumor Detection

X-ray images show a well-defined radiolucent lesion with a sclerotic margin in the left femur, consistent with a bone tumor. The lesion measures approximately 4 cm in diameter and appears to involve the cortex of the bone. There is no evidence of pathological fracture or soft tissue invasion. This description provides information about the location of the tumor (left femur), and its appearance on the X-ray images (well-defined radiolucent lesion with a sclerotic margin) as shown in Figure 1. This information can help doctors determine the appropriate treatment for the tumor and monitor its progression over time.

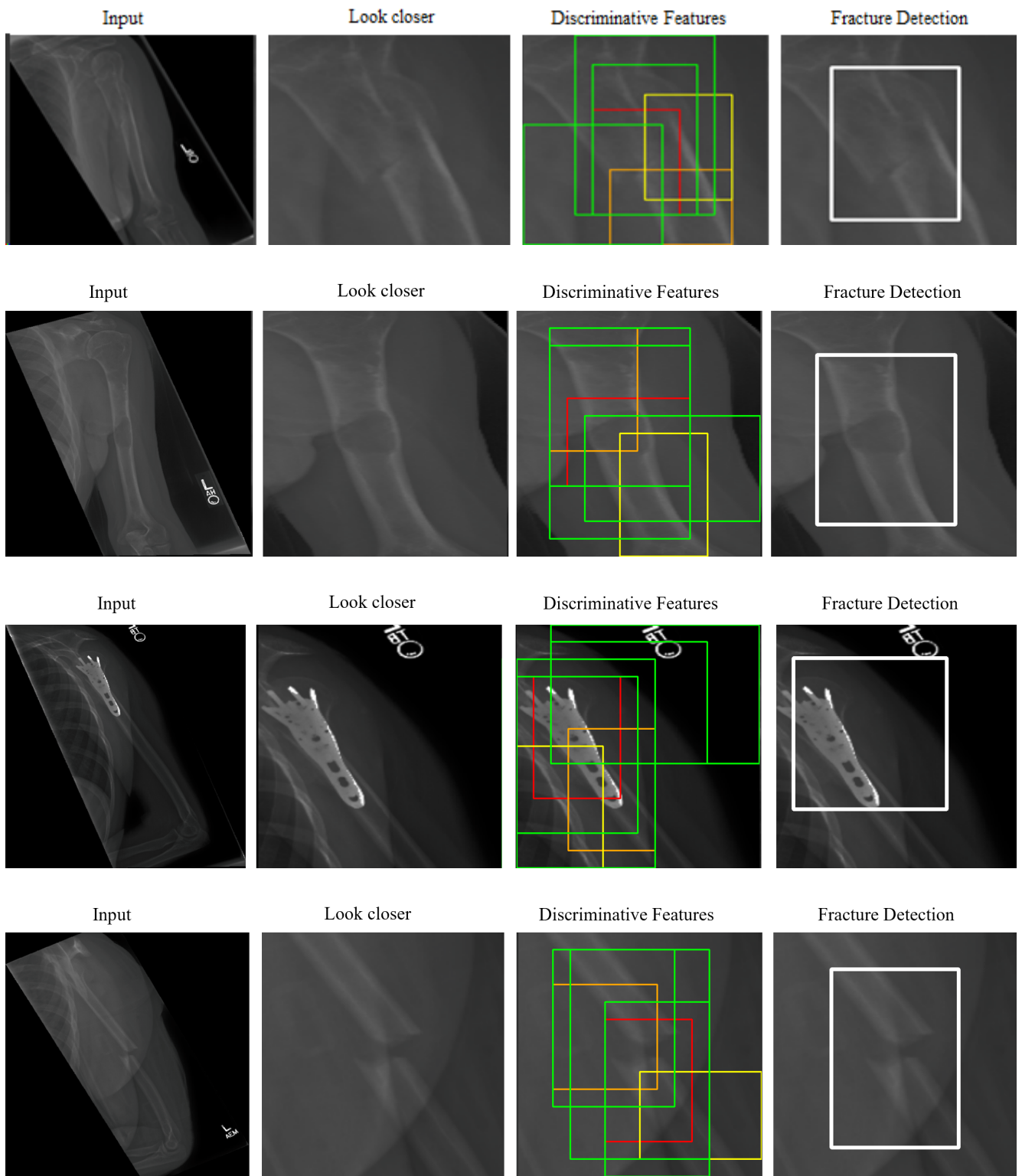
## 1.2. Fracture Detection

Fracture localization refers to identifying the specific location of a bone fracture in an X-ray image. When interpreting X-ray images of fractures, the radiologist or healthcare professional will carefully examine the image to identify the presence of any fractures, and then describe the location and type of fracture.

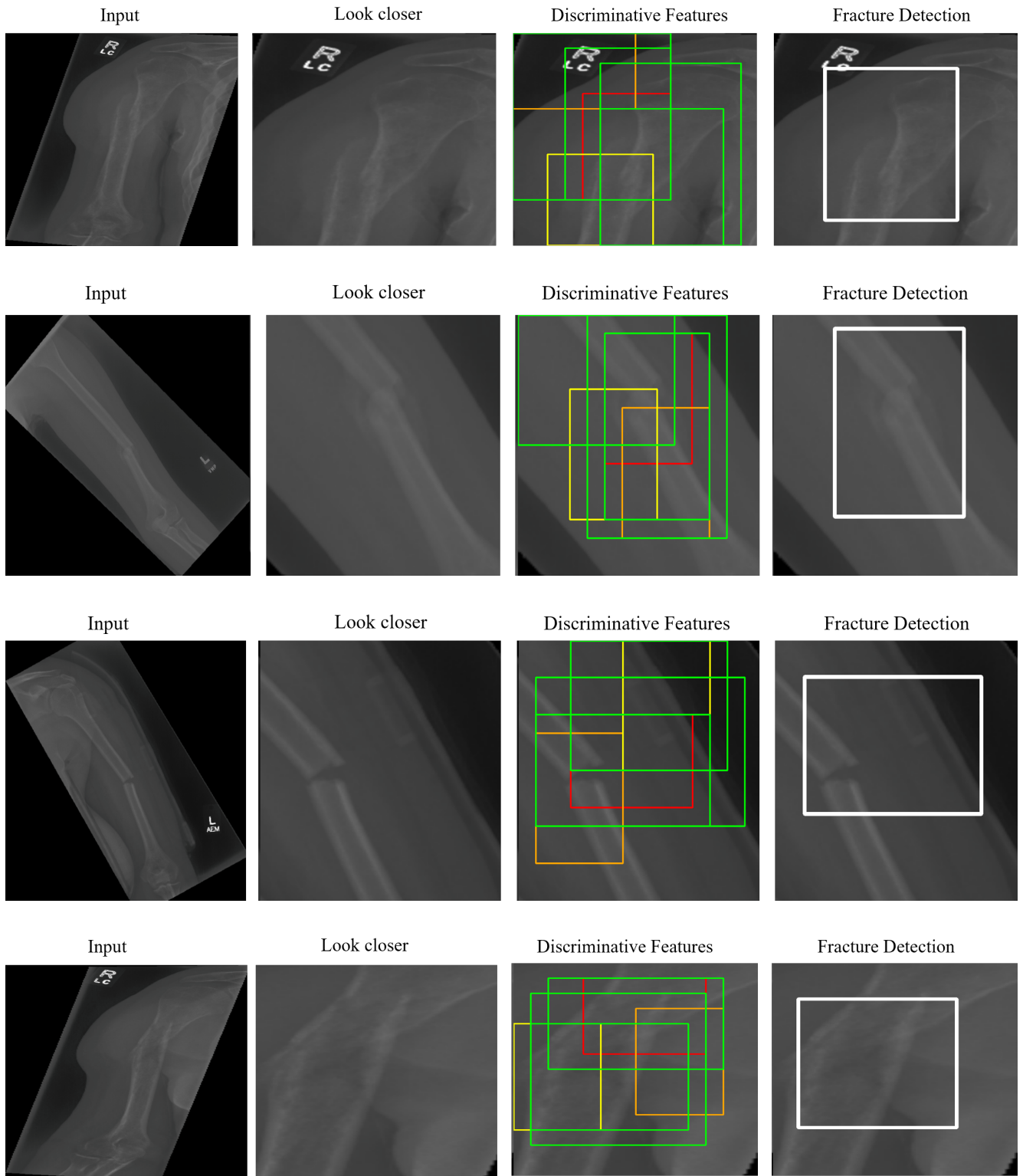
The description of a fracture in an X-ray image shown in Figures 2 3 can help guide the appropriate treatment plan, which may include immobilization, surgery, or other interventions to promote healing and restore normal function.



**Figure 1:** Visualization of the outputs from MSFM at each step of the model. The first column consists of a random sample original image from each part of the entire dataset followed by the columns of FMA, Look Closer in the original image, proposed informative part, tumor, and the area of interest in the original image as tumor detection. The final tumor detection in the white box is shown in the last two columns.



**Figure 2:** Visualization of the outputs from MSFM at each step for MURA dataset



**Figure 3:** Some more results on the visualization of the outputs from MSFM at each step for MURA dataset