# 2D/3D Registration

## 1. Results from Evaluation of Trained Patient-Specific Model

As part of our system hardware upgrades, we aligned the training configuration with the default parameters: batch size of **116**, **100** batches per epoch, and a total of **1000** epochs.

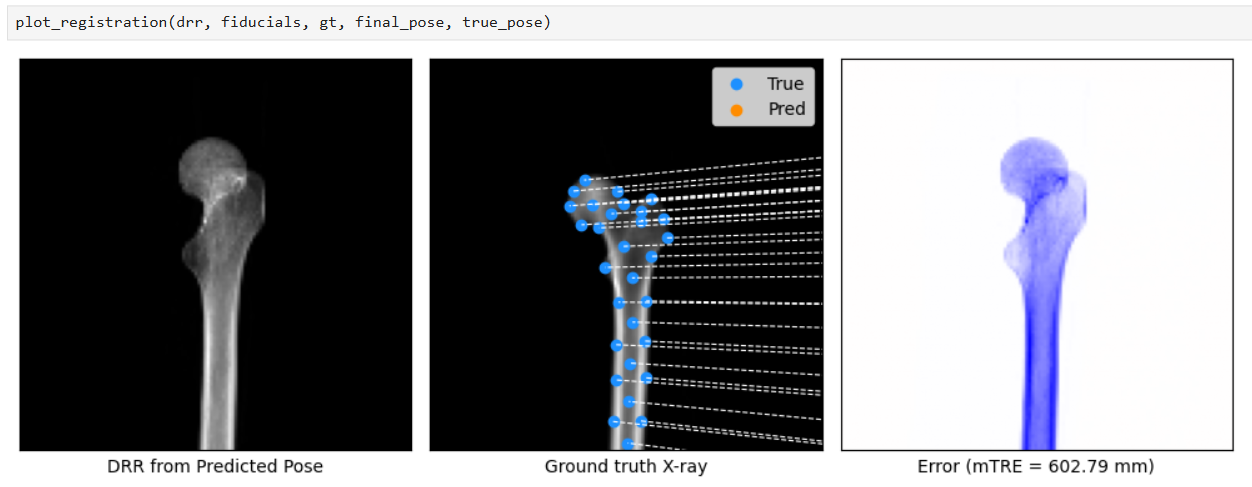
We initialized the Source-to-Detector Distance (SDD) based on the C-arm machine operator’s manual, setting it to **1000 mm**. The delx (pixel spacing) value was selected after reviewing numerous X-ray images to ensure it closely matched real C-arm generated X-rays.

To maintain a constant **Field of View (FOV)**, calculated as **pixel width × pixel spacing**, we followed the recommended approach of adjusting the relationship between pixel spacing (delx) and image resolution. For evaluation purposes, we experimented with increasing the image resolution while proportionally decreasing the pixel spacing.

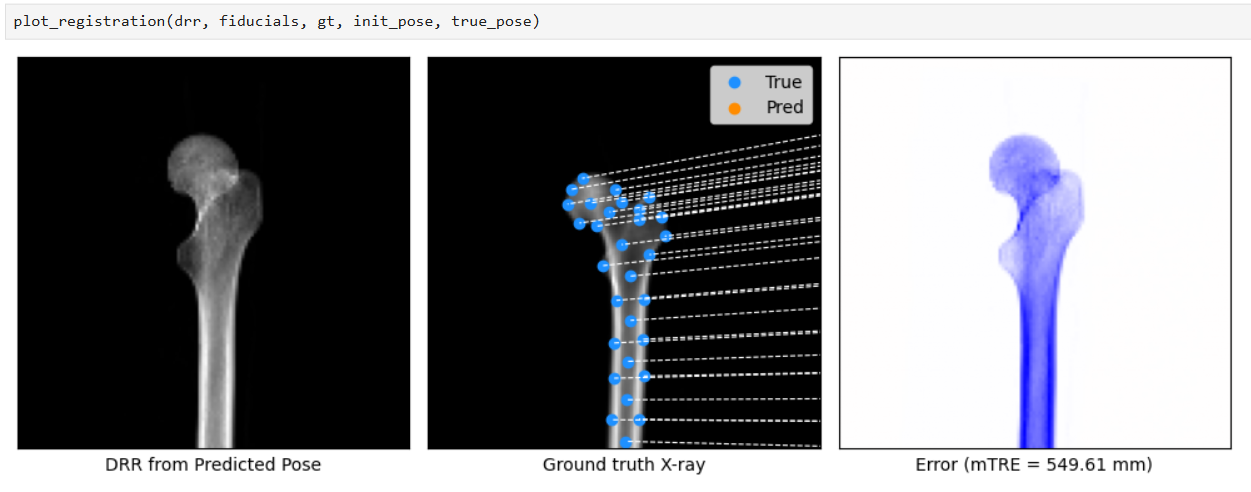
The table below summarizes the combinations of image resolution and pixel spacing that we tested:

| **Resolution (pixel width) in mm** | **Pixel Spacing (delx) in mm** | **FOV in mm** |
| --- | --- | --- |
| 128x128 | 2.42 | 309.76 |
| 256x256 | 1.21 | 309.76 |
| 512x512 | 0.605 | 309.76 |
| 980x980 | 0.31608 | 309.76 |
| 1024x1024 | 0.3025 | 309.76 |

We saved the best-performing model, based on the lowest dGEO value, as *RXVR1.1\_best.pth*. Below is an example output generated by the trained model using a resolution of 1024×1024 and a pixel spacing of 0.3025:



**Fig-1:** *Final\_Pose Error (mTRE) visualization for the* ***Patient-Specific trained model*** *with a resolution 1024x1024 and delx = 0.3025*

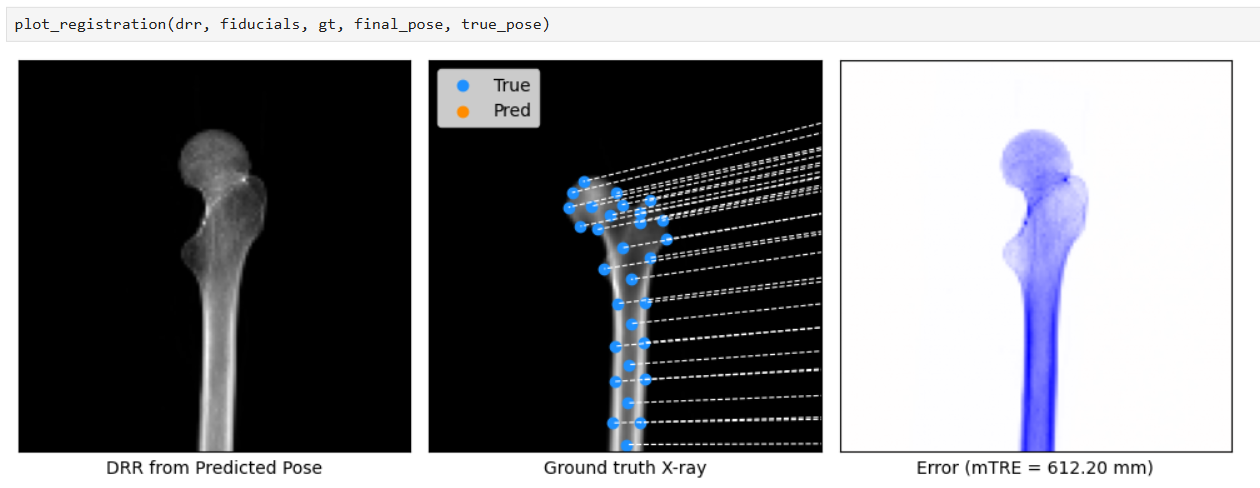


**Fig-2:** Initial\_Pose *Error (mTRE) visualization for the* ***Patient-Specific trained model*** *with a resolution 1024x1024 and delx = 0.3025*

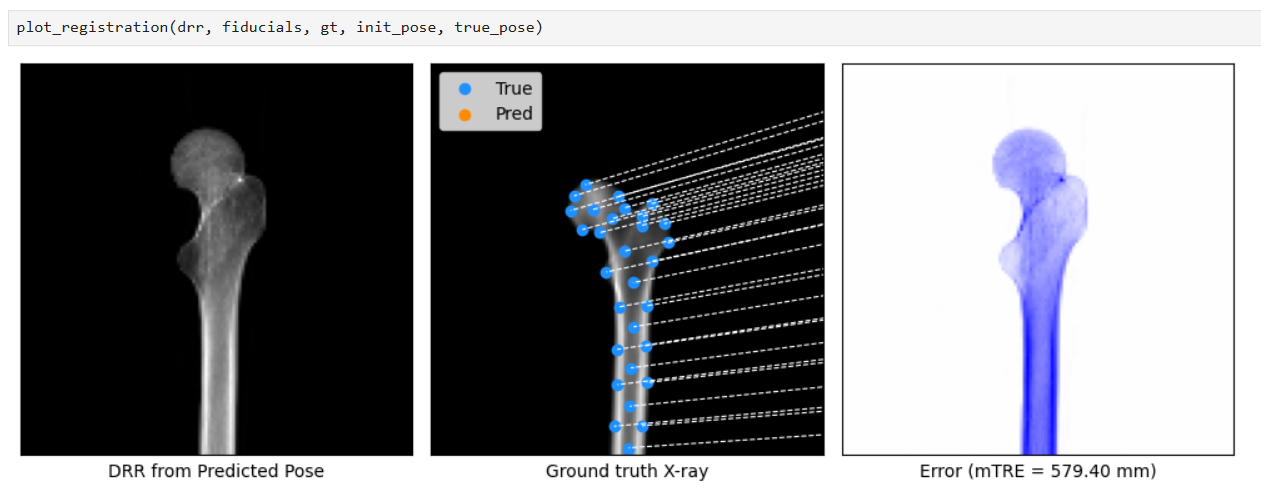
*\*\*All the resolution results have been uploaded to:* [*https://github.com/robossis/Robossis-xvr/tree/main/Robossis-Results*](https://github.com/robossis/Robossis-xvr/tree/main/Robossis-Results)

## 2. Results from Evaluation of Finetuned Trained Patient-Specific Model

The trained model was fine-tuned using the default parameters, and the resulting outputs were comparable to those from patient-specific models. Below is an example output generated by the fine-tuned model at a resolution of 1024×1024 with a pixel spacing of 0.3025:



**Fig-3:** Final\_Pose *Error (mTRE) visualization for the* ***Finetuned Patient-Specific trained model*** *with a resolution 1024x1024 and delx = 0.3025*

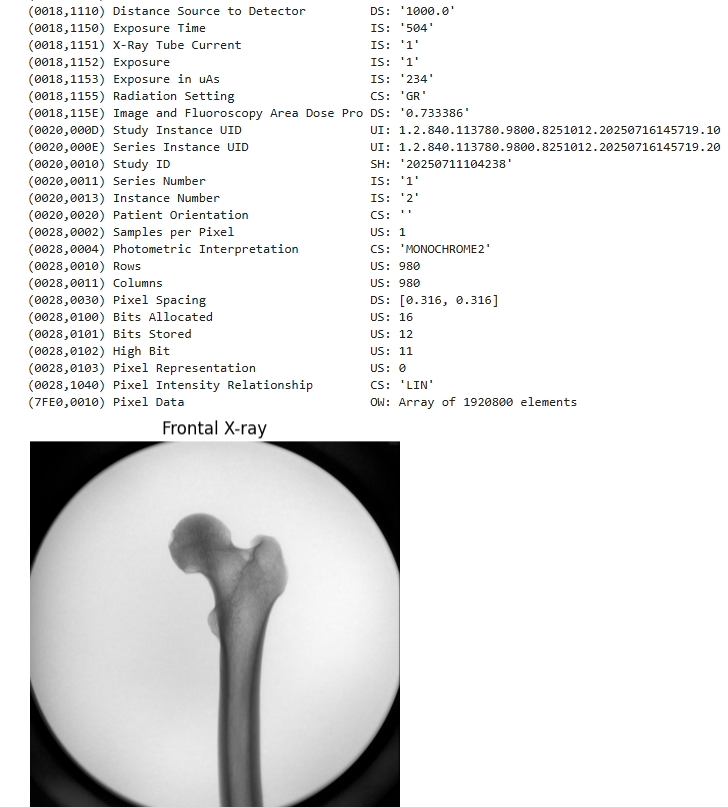


**Fig-4:** Initial\_Pose *Error (mTRE) visualization for the* ***Finetuned******Patient-Specific trained model*** *with a resolution 1024x1024 and delx = 0.3025*

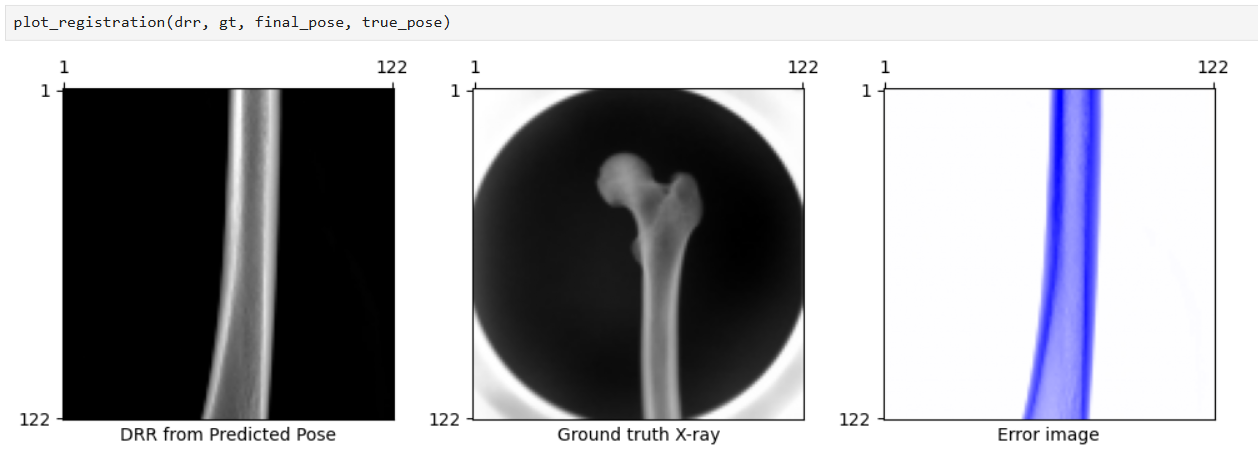
## 3. Results of real X-ray images from C-arm:

To evaluate the performance of the trained patient-specific model, we fed a real X-ray image captured from the C-arm as input, with a resolution of 980×980 and a pixel spacing of 0.316 mm. The model failed to predict the true pose value and deviated by a large margin. Instead of showing the proximal part, it predicted the shaft of the femur.

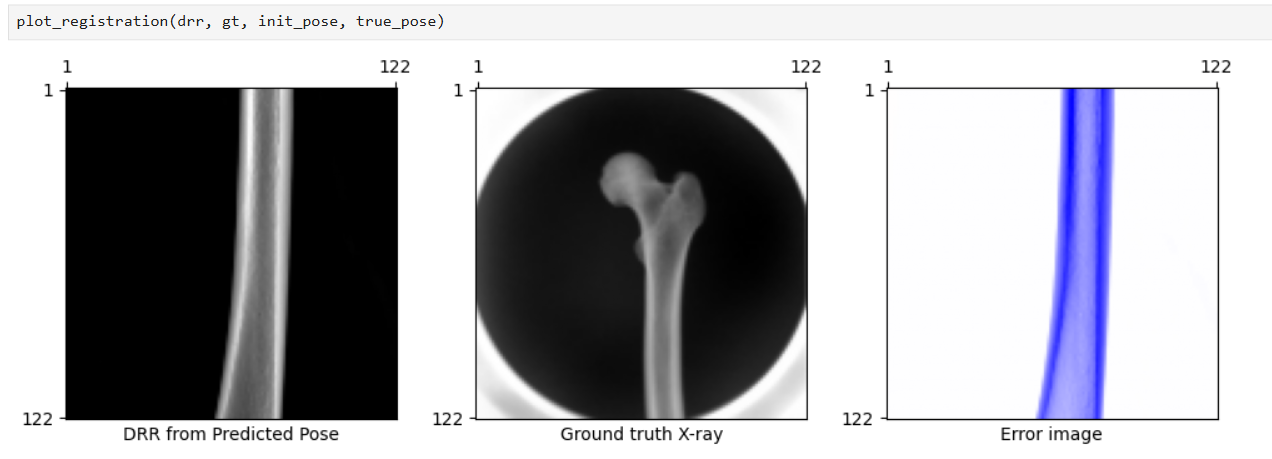




**Fig-5:** *Metadata from the real X-ray captured from the C-arm. Since our C-arm machine doesn’t save any values for* ***‘SDD’*** *and* ***‘Pixel Spacing’*** *we manually fed the values as 1000 and 0.316, respectively.*



**Fig-6:** Final\_Pose *Error (mTRE) visualization for the* ***Patient-Specific trained model*** *for the real X-ray captured using C-arm with a resolution of 980x980 and delx = 0.316.*



**Fig-7:** Initial\_Pose *Error (mTRE) visualization for the* ***Patient-Specific trained model*** *for the real X-ray captured using C-arm with a resolution of 980x980 and delx = 0.316.*

## 

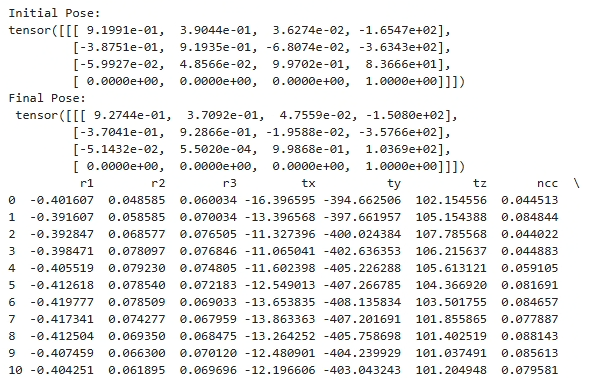
## 4. Results of the Registration model



**Fig-8:** *a.) final\_img b.) ground truth (gt) c.) init\_img*

In addition to the previous step, where we evaluated the performance of the trained model using a real C-arm X-ray as input, we also attempted to register real X-ray images using a pretrained model, followed by iterative pose refinement with differentiable rendering. For this, we used the xvr register module. The inputs included the CT volume, the path to the real X-ray image (used as the target model), and the patient-specific trained model (RXVR1.1\_best.pth).

The output consisted of the final, ground truth, and initial projections. However, the registration was inaccurate, showing a significant downward shift in the predicted result, likely due to a large negative offset along the Y-axis.



**Fig-9:** *Sample output from parameters.pt file*