

Fig. 1: Label distribution of different datasets in our setup.

Table 1: Comparison with upper and lower bound values in homogeneous setup: Lowerbound refers to a model trained only with public data. The Upperbound refers to when all data is stored in a central server including public data. mFedAvg*P*-NM refers to mFedAvg*P* in the case where there are no missing modalities in clients. Both mFedAvg*P* and CAR-MFL values are of the extreme setting of 8 image-only clients (8:0:2).

| | Upperbound | Lowerbound | $\mathbf{mFedAvg}{P}\textbf{-}\mathbf{NM}$ | mFedAvgP | CAR-MFL |
|-----|------------|------------|--|----------|---------|
| AUC | 91.67 | 83.11 | 90.17 | 81.95 | 87.31 |

Table 2: No. of data samples at various patient counts in public data.

| No. of Patients | 1000 | 750 | 500 | 250 |
|---------------------|------|------|------|-----|
| No. of Data Samples | 2701 | 1888 | 1210 | 602 |

Table 3: No. of data samples across clients in *homogeneous* and *heterogeneous* setups. In a heterogeneous setup, clients 8 and 9 contain multimodal data from NIH Open-I, and the rest are image-only clients with images from CheXpert.

| Client ID | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 |
|---------------|------|------|------|------|------|------|------|------|------|------|
| Homogeneous | 2343 | 2123 | 2171 | 2107 | 2195 | 2127 | 2164 | 2188 | 2528 | 2086 |
| Heterogeneous | 2245 | 2154 | 2113 | 2359 | 2133 | 2003 | 2189 | 2205 | 1116 | 1116 |

Table 4: Validation AUC across the various α for homogeneous 4:0:6 setting.

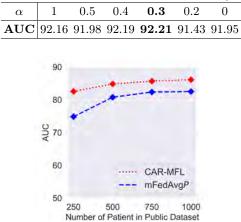


Fig. 2: Model AUC on varying patient size in public data for *heterogeneous* setup.

| Original Pair | Retrieved Complementary Modality | | | | | | |
|---|--|--|--|--|--|--|--|
| Original Pan | Round 1 | Round 15 | Round 30 | | | | |
| | there is no parenchymal consolidation, the cardiomediastinal silhouette is unchanged . an arygos fissure is re- demonstrated a normal variant as seen on chest ct dated . bony structures are notable for mid thoracie dextroscoliosis . | pa and lateral views of the chest provide ed. there is no focal consolidation effu- sion or pneumothorax. The cardiomedi astinal silhouette is normal imaged os seous structures are intact. Ino free air below the right hemidiaphragm is seen | The cardiase mediastinal and hiar contours are n ormal. Jungs are clear, pulmonary vasculature (a) no pleural efflusion or pneumothorax is seen : no acute osseous abnormalities are vi sualized. | | | | |
| heart size is normal. the mediastin al and hilar contours are normal. I he pulmonary vasculature is norma 1. lungs are clear, no pleural effusi on or pneumothorax is seen. there are no acute osseous abnormalitie s | PS | | 25 | | | | |
| P | in comparison with the study of there is progressive decrease in the right b asilar consolidation , retrocardiac opa cification is essentially unchanged . co ontinued enlargement of the cardiac s lihouette with pacer extending to the apex of the right ventricle . Here may be mild elevation of pulmonary venou s pressure | et tube ends above the clavicles and c ould be advanced by approximately 3 c in for more secure positioning, enteric tube with the tip in the stomach but the side hole above the level of the ge junc tion. Left subclavian central venous cat heter in appropriate position, right low er lung and retrocardiac opacities mos tikely represent aspiration, impressio n posted to the ed dashboard by dr at 0780 on min after review. | as compared to the recent radiograph from seve al hr earlier and endotracheal tube remains in pl ce terminating 4 cm above the carina and the lef internal jugular catheter continues to abut the ey pected location of the lateral wall of the superior vena cava near the junction with the brachicoeg alic vein. bildteral asymmetrical airspace opacit es have slightly improved but remain most confl uent in the right upper lobe - additionally note is made of worsening opacification in the left retro ardiac region | | | | |
| lines and tubes and cardiomediastin al silhouette stable . faint opacities r emain visible but have improved co mpared with at 2005 pm . | 11 | R.F. | | | | | |

Fig. 3: Qualitative Analysis of retrieved samples across different training rounds. Column 1 contains two paired image text samples. Row 1 and 3 display text reports from the public dataset when the text modality is missing. Row 2 and 4 display retrieved images when the image modality is missing.