

Supplementary Material

Abnormality	Base w/ CT-Net			CT2Rep (Ours)			Test Set
	P	R	F1	P	R	F1	
Medical material	0.117	0.400	0.180	0.736	0.752	0.308	0.114
Arterial wall calcification	0.795	0.635	0.706	0.928	0.356	0.675	0.267
Cardiomegaly	0.088	0.336	0.139	0.657	0.686	0.555	0.108
Pericardial effusion	0.996	0.820	0.899	0.769	0.920	0.786	0.074
Coronary artery wall calcification	0.173	0.163	0.168	0.405	0.705	0.223	0.244
Hiatal hernia	0.110	0.666	0.188	0.998	0.214	0.890	0.134
Lymphadenopathy	0.932	0.380	0.540	0.641	0.645	0.555	0.266
Emphysema	0.660	0.870	0.750	0.436	0.864	0.310	0.195
Atelectasis	0.052	0.122	0.073	0.998	0.638	0.778	0.232
Lung nodule	0.713	0.381	0.496	0.818	0.502	0.596	0.425
Lung opacity	0.994	0.588	0.741	0.579	0.234	0.320	0.374
Pulmonary fibrotic sequela	0.948	0.465	0.624	0.569	0.669	0.443	0.267
Pleural effusion	0.137	0.649	0.226	0.570	0.396	0.443	0.126
Mosaic attenuation pattern	0.289	0.230	0.256	0.984	0.242	0.386	0.078
Peribronchial thickening	0.194	0.749	0.308	0.888	0.660	0.624	0.098
Consolidation	0.293	0.795	0.428	0.804	0.417	0.526	0.172
Bronchiectasis	0.782	0.502	0.601	0.756	0.425	0.516	0.099
Interlobular septal thickening	0.969	0.811	0.883	0.953	0.547	0.684	0.071
Mean	0.513	0.531	0.456	0.749	0.548	0.534	0.186

Table 1: Abnormality-based clinical efficacy metrics, including precision (P), recall (R), and F1 score, are showcased for generated reports by the CT-Net-based baseline and our CT2Rep method. CT2Rep’s superior performance underscores the benefits of utilizing our novel auto-regressive causal transformer for 3D feature extraction, coupled with relational memory and memory-driven conditional layer normalization, to generate clinically accurate reports for 3D chest CT volumes. Additionally, the ratios of abnormalities in the test set are provided.

Abnormality	Baseline			CT2RepLong			Test Set
	P	R	F1	P	R	F1	
Medical material	0.490	0.593	0.492	0.910	0.617	0.555	0.233
Arterial wall calcification	0.390	0.417	0.389	0.613	0.476	0.472	0.330
Cardiomegaly	0.806	0.120	0.478	0.993	0.690	0.886	0.153
Pericardial effusion	0.851	0.718	0.448	0.711	0.467	0.553	0.145
Coronary artery wall calcification	0.947	0.250	0.569	0.750	0.657	0.664	0.296
Hiatal hernia	0.486	0.316	0.472	0.943	0.805	0.472	0.148
Lymphadenopathy	0.683	0.414	0.498	0.926	0.857	0.435	0.333
Emphysema	0.658	0.396	0.567	1.000	0.635	0.777	0.300
Atelectasis	0.577	0.416	0.572	0.454	0.287	0.457	0.258
Lung nodule	0.864	0.347	0.740	0.703	0.348	0.585	0.453
Lung opacity	0.563	0.894	0.490	0.480	0.153	0.315	0.535
Pulmonary fibrotic sequela	0.640	0.570	0.539	0.541	0.766	0.548	0.295
Pleural effusion	0.502	0.667	0.509	0.698	0.618	0.391	0.286
Mosaic attenuation pattern	0.655	0.375	0.527	0.729	0.501	0.574	0.024
Peribronchial thickening	0.695	0.536	0.569	0.974	0.955	0.958	0.168
Consolidation	0.445	0.569	0.439	0.413	0.181	0.304	0.378
Bronchiectasis	0.986	0.211	0.562	0.514	0.713	0.440	0.165
Interlobular septal thickening	0.740	0.570	0.605	0.737	0.098	0.266	0.117
Mean	0.666	0.465	0.525	0.727	0.511	0.536	0.257

Table 2: Clinical efficacy metrics (precision, recall, and F1) for each abnormality are showcased for reports generated by the baseline and our enhanced CT2RepLong method. Augmenting CT2Rep, CT2RepLong integrates a cross-attention multi-modal fusion module and longitudinal memory, effectively leveraging historical reports and volumes from previous visits. This method’s superior performance over the baseline underlines the benefits of employing longitudinal multimodal data in producing clinically precise radiology reports for 3D chest CT volumes. The distribution of abnormalities in the test set is also detailed.