Supplementary Material

1 Illustration of Gaze-directed Graph Construction layer

Fig. S1 illustrates the steps of the first GDGC in GDC. The i-th row of the three distances in the figure represents the distance between the i-th node and the other 3136 nodes. The i-th row of the edge index represents the indices of the 9 nearest neighbor nodes of the i-th node.
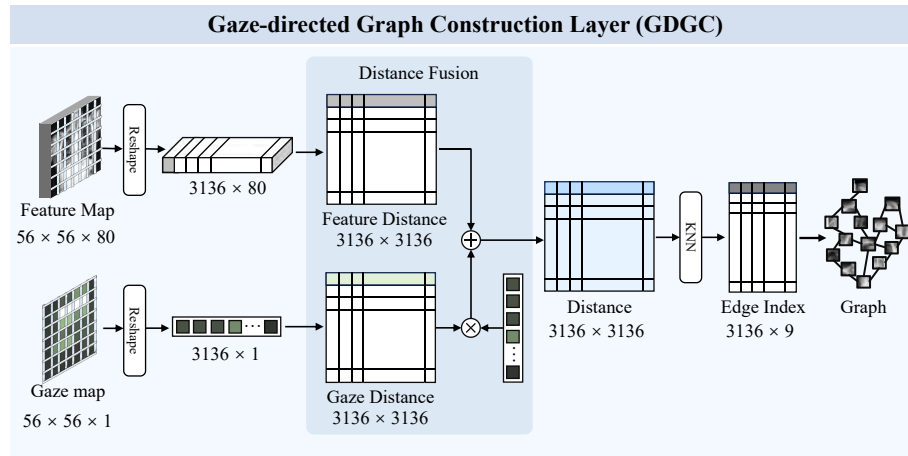

Fig. S1. Overview of the first gaze-directed graph construction layer in GDC.

2 Qualitative Analysis of Gaze Maps

In this section, we provide our qualitative results on gaze map generation. Fig. S2 compares the gaze maps generated by our proposed GD-ViG and other methods. Blue arrows indicate regions where the generation is close to the real gaze, while red arrows indicate regions where the generation fails. The results show that our generated gaze maps have better quality and are closer to the real gaze.

3 Qualitative Visualization on EGD-CXR

Fig. S3 visualizes the attention maps of different methods on the EGD-CXR dataset. It can be seen that our proposed GD-ViG focuses on the foreground region of the image without shortcut learning, and has more refined attention maps than other methods.
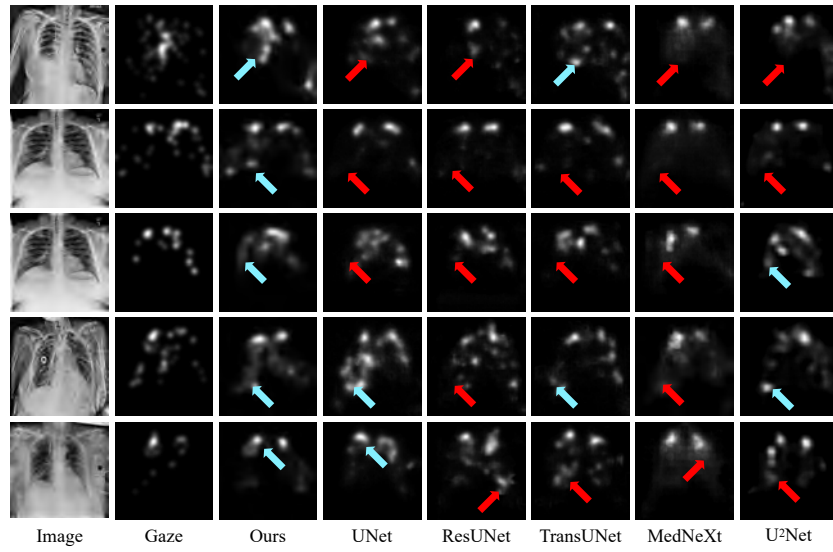
| Image | Gaze | Ours | UNet | ResUNet | TransUNet | MedNeXt | U²Net |

Fig. S2. Comparison of gaze maps generated by different methods. Blue arrows and red arrows indicate the regions where the generation is correct and failed, respectively.
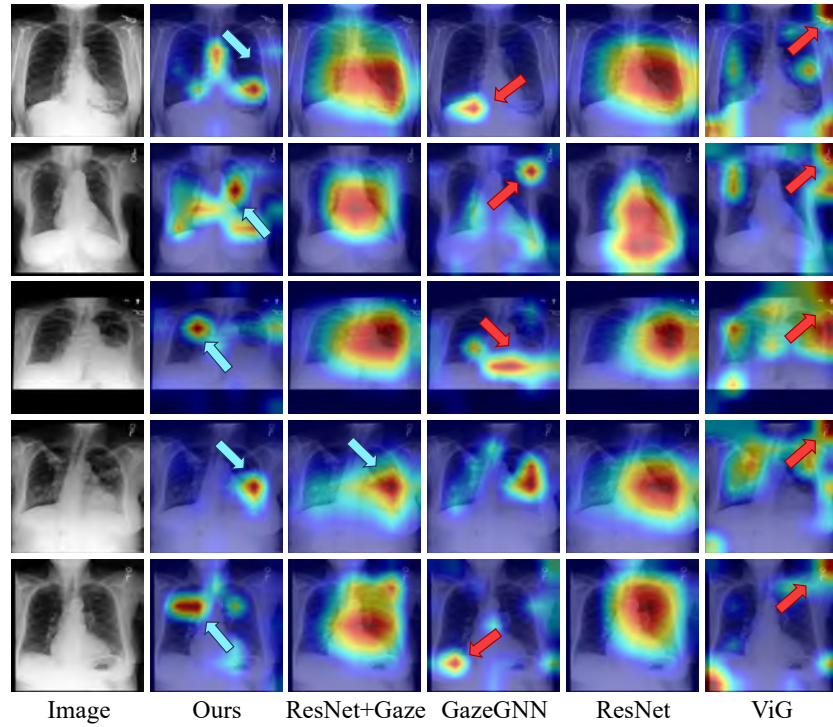

| Image | Ours | ResNet+Gaze | GazeGNN | ResNet | ViG |

Fig. S3. Visualizes the attention maps of different methods on the EGD-CXR dataset.