1 Appendix

Network Architecture The network architecture of two cross-domain attention modules is shown in Figure 1.
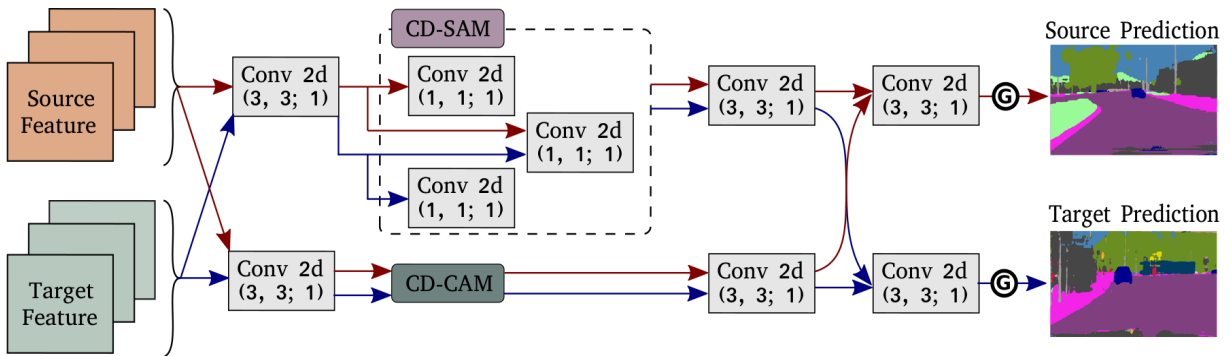

Figure 1: The network architecture of two cross-domain attention modules, where Conv 2d $(k, k; j)$ indicates the 2D convolution with kernel size $(k \times k)$ and stride j. The red and blue lines denote features from source and target domains, respectively.