## (b) Prior-guided fusion module (a) Overview of framework You can find them online, or your local bookseller, and I hope you enjoy them, but I still recommend, Pose-only do improve first, read about it second. RGB-pose Default **Pre-trained Large Language Model** Gater $\mathcal{F}_{i,t}^{p}$ Feature Concatenation Deformable Position init. Few and Low Vision Attention Resolution Clip Encoder **Temporal** Temporal **Temporal** K, VQ Encoder Encoder Encoder Score-aware Multi-head Sampling $\mathcal{F}_{i,t}^r$ Attention **PGF** Pose Pose Pose K, VModule **\** Encoder Encoder Encoder $J_{i,t}^{s}$ Divide Into Sub-pose Keypoints Sequence $\mathcal{F}_{i,t}^{p}$ $J_{i,t}^s$ $\mathcal{F}_{i,t}^{r}$ Sub-pose