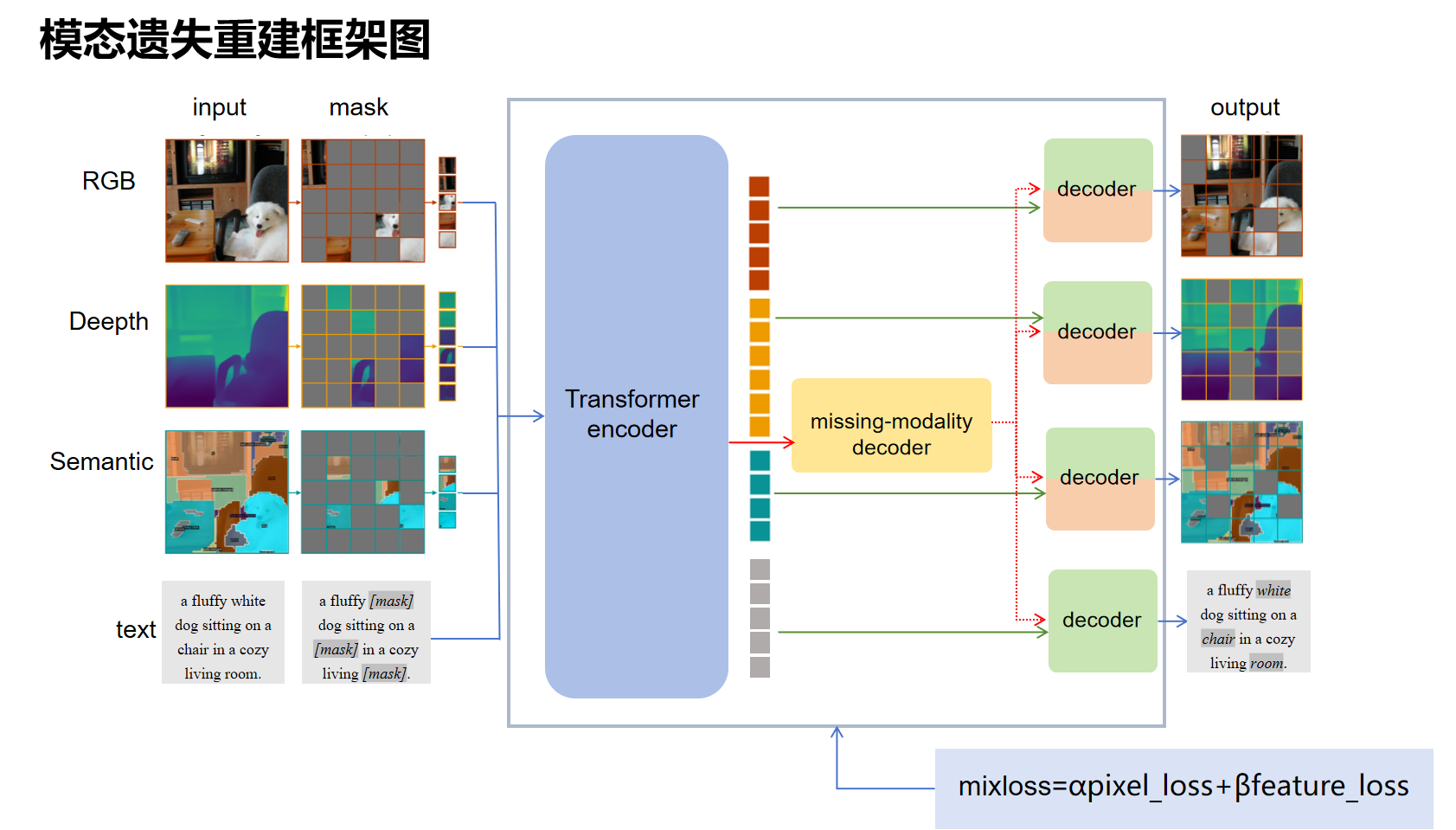
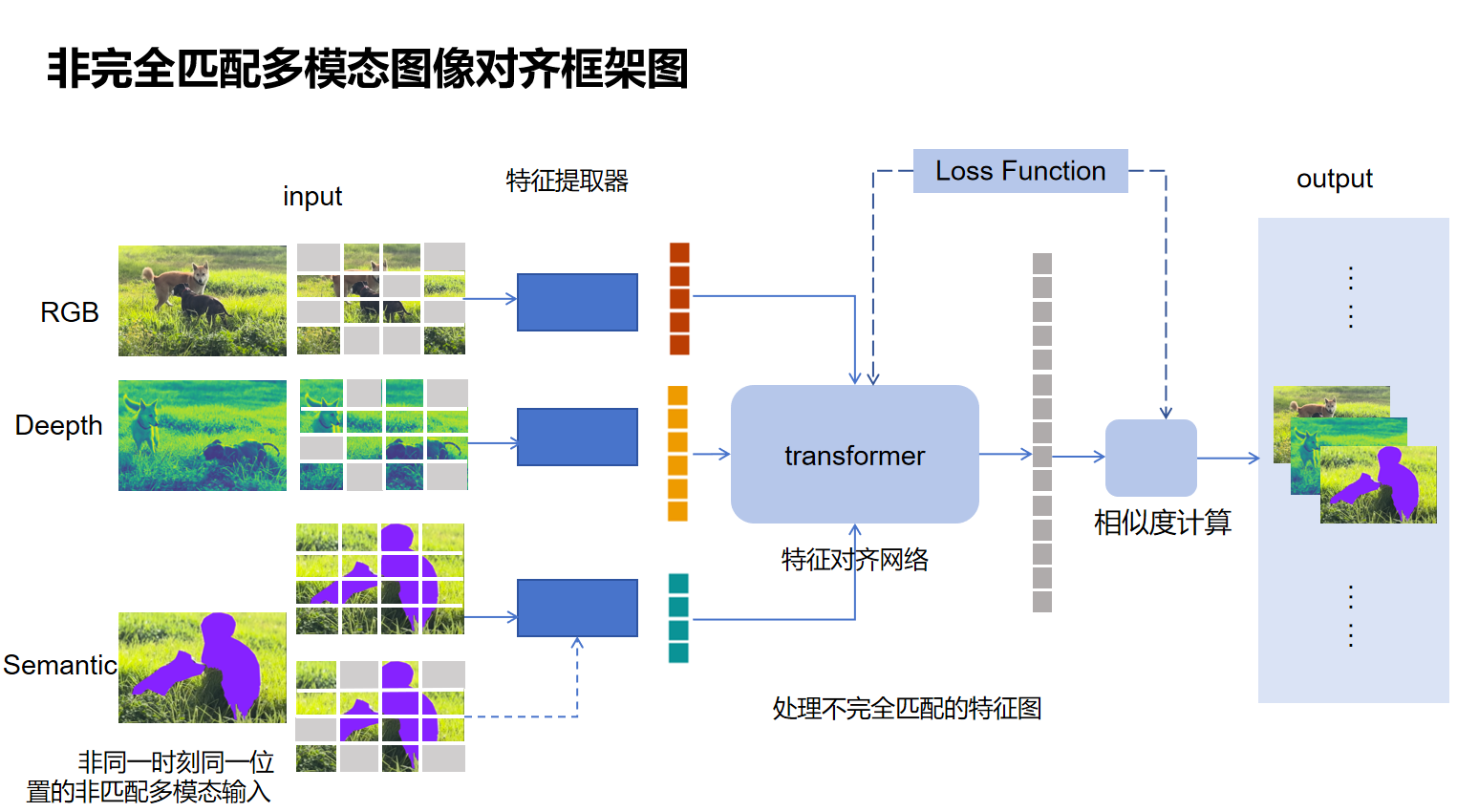
1. 读文献
2. NeurIPS-2023-siamese-masked-autoencoders-Paper-Conference
3. MultiMAE: Multi-modal Multi-task Masked Autoencoders
4. Multi-modal Learning with Missing Modality via Shared-Specific Feature Modelling（综述）
5. VLM See, Robot Do:Human Demo Video to Robot Action Plan via Vision Language Model
6. 整理上周讨论的两个思路图





待实验确定：

* 是否需要一种模态不做掩码？
* 非完全匹配多模态图像的对齐哪一种方式可行？

