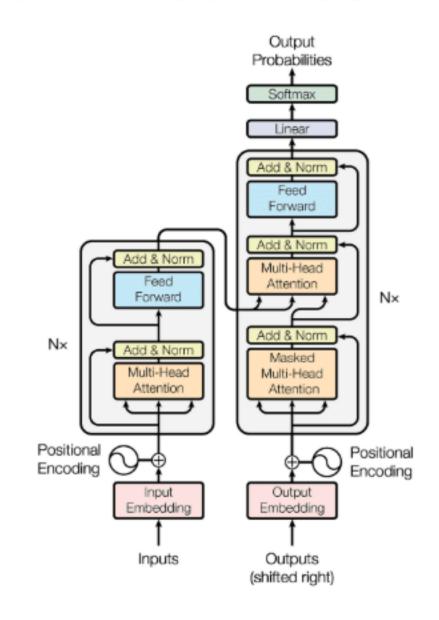
Compute-scalable architectures



Large-scale (open) datasets



Zero/Few-Shot Task Generalization









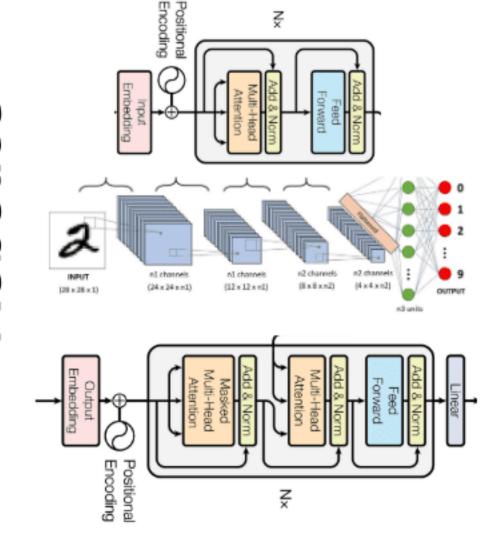
→ Image Segmentation

→ Q&A

→ Summarization

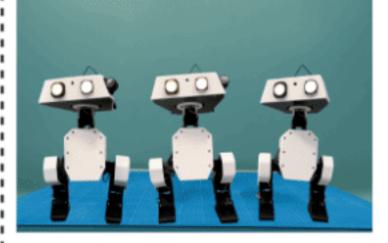
→ Image generation

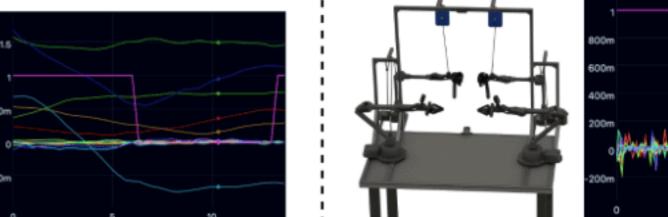
Small-size, crafted architectures

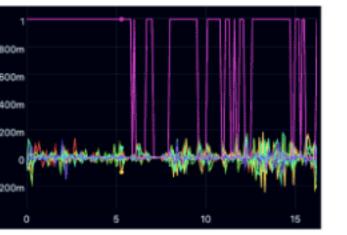


Task and embodiment-specific, siloed, small-scale datasets









Task-specific models

→ RAM Insertion

Quadrupedal Locomotion

•••

Bimanual manipulation

→ Bipedal balancing