## PANDA

Shujun

July 15, 2020

## 1 Core architecture

Our core architecture is basically any pretrained vision model  $+\ 1$  transformer encoder layer.

Your text goes here.

## 1.1 A subsection

More text.