

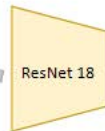
(a) Wearing SpeeChin camera



Pre-processing



(b) Captured image sequence



padding



FC Layer

(c) SpeechNet model

"Seven"
 "Answer"
 "Paste"
 ...
 "Hey Siri"
 "Right"
 ...
 "返回"
 "搜索"
 "打开相机"
 ...
 "复制"
 "挂断"

(d) Silent speech commands