

スマートフォン向けにカスタマイズが可能な サイレントスピーチインタフェース

—音声入力の自在化を目指して—

蘇子雄 (東京大学大学院)

方詩涛 (東京大学大学院)



音声入力便利すぎて
毎日使いたいけど、
外で使うのが恥ずかしい...



わかる！スマホに向かって
話すと恥ずかしいよね
個人情報の漏洩も問題

いつでも、どこでも使えるように 音声入力を自在化する

どうやって？

- ◆ 口元の動きから発話を読み取る
「サイレントスピーチシステム」

どこがすごい？

- ◆ どんな環境でも高い精度で認識
- ◆ 一発で自分のコマンドを登録



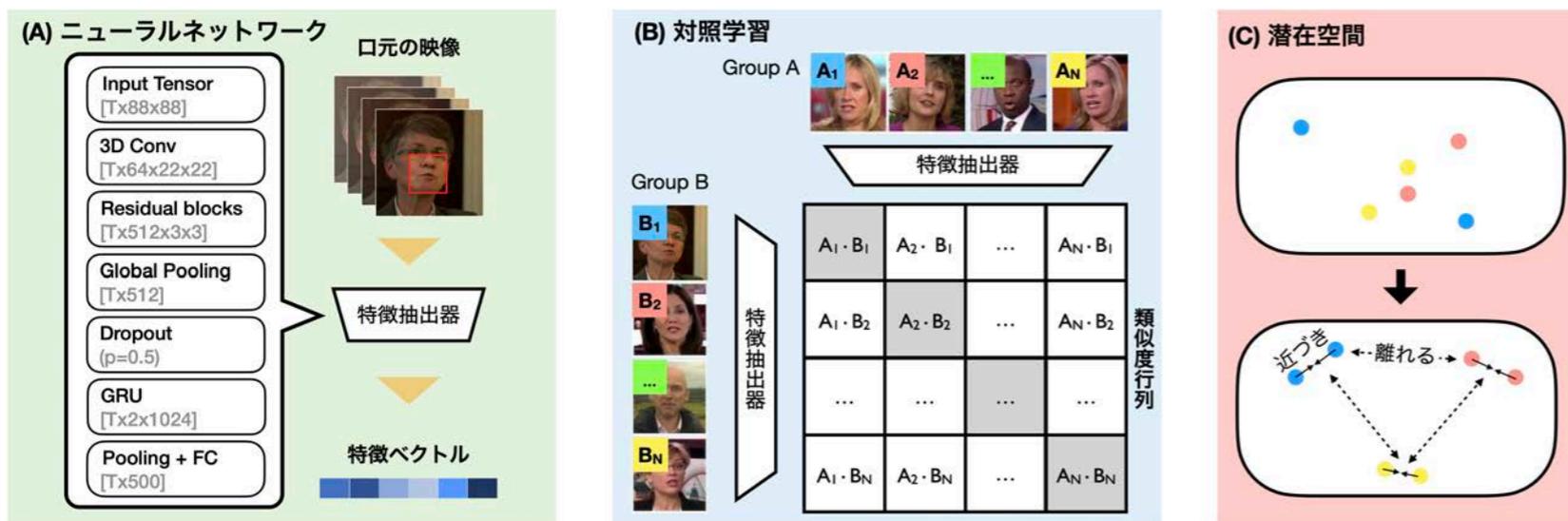
1. 音声入力コマンド登録



2. 発声しづらい場所でサイレントスピーチ

対照学習パイプライン

自己教師あり学習で有用な表現を抽出する、新しい深層学習手法



どんな環境でも正確に認識

照明環境や手ブレにロバスト、30コマンドの認識率98%!



サイレントで起動

世界初の Silent wake-up

キーワード コマンド

Hello Alexa 電気を消して



Stand-by Mode

Wake-up keywordが発話されるのを待機する

振動フィードバック



Awake Mode

口元の画像を記録する



Recognition Mode

コマンドを読唇モデルで認識する