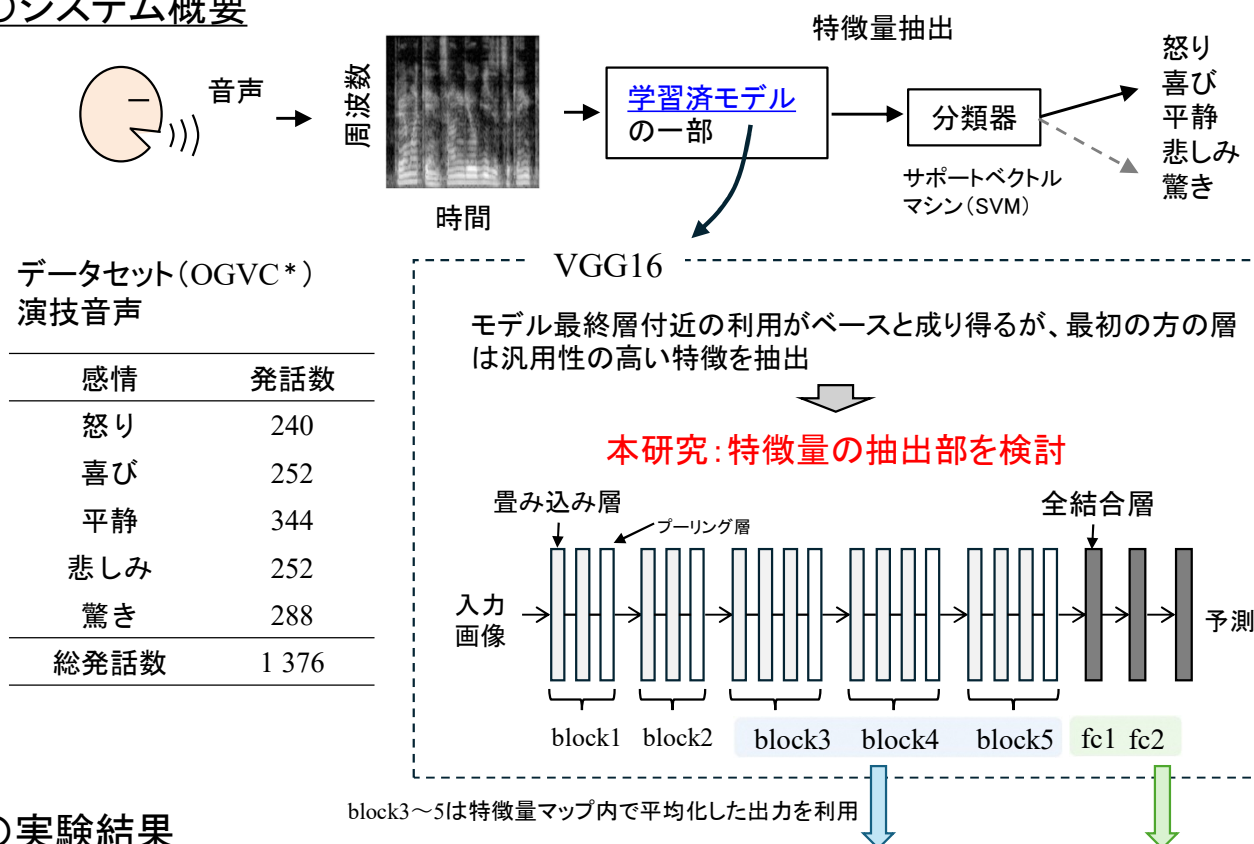


機械学習を用いた人の状態認識システムの構築

音声データと機械学習を用いて人の感情や健康状態を推定する技術があり、人と機械との対話や疾病の簡易自動検知への応用が期待される。感情認識では、正解感情が定められた音声データベースの作成にコストがかかるため、学習用の音声データが少量となり、これが感情認識を困難とする一要因となっている。対応策のひとつとして、大量のデータを用いて事前に学習された深層学習モデルの一部を感情認識モデルに転用する手法がある。本研究では、**少量の音声学習データによる人の状態認識の基礎**として、**事前学習済モデル**を用いた音声感情認識を検討した。

○システム概要



○実験結果

話者3名分の音声データを用いてSVMを学習させ、残りの1名の音声データに対して評価

感情	再現率[%]
怒り	56
喜び	63
平静	72
悲しみ	69
驚き	64

特徴量抽出の出力部	平均再現率[%]
block3	63
block4	65
block5	56

特徴量抽出の出力部	平均再現率[%]
fc1	59
fc2	56

モデルの最終層付近よりも中間付近のblock4出力を用いた方が高い再現率が得られた

今後の課題: 他のデータセットや手法による結果との比較検証、再現率向上のため他のモデルの適用等

*国立情報学研究所 音声資源コンソーシアムから提供を受けた「感情評定値付きオンラインゲーム音声チャットコーパス (OGVC)」を利用した。