

深層学習の転移学習を応用した 少数の高齢者画像を用いた顔表情認識システムの開発

美添健¹⁾, 田中正行¹⁾, 荻田美穂子²⁾, 月野木ルミ³⁾

- 1) 東京科学大学工学院システム制御系
- 2) 滋賀医科大学臨床看護学講座（老年）
- 3) 東京科学大学大学院保健衛生学研究科公衆衛生看護学分野

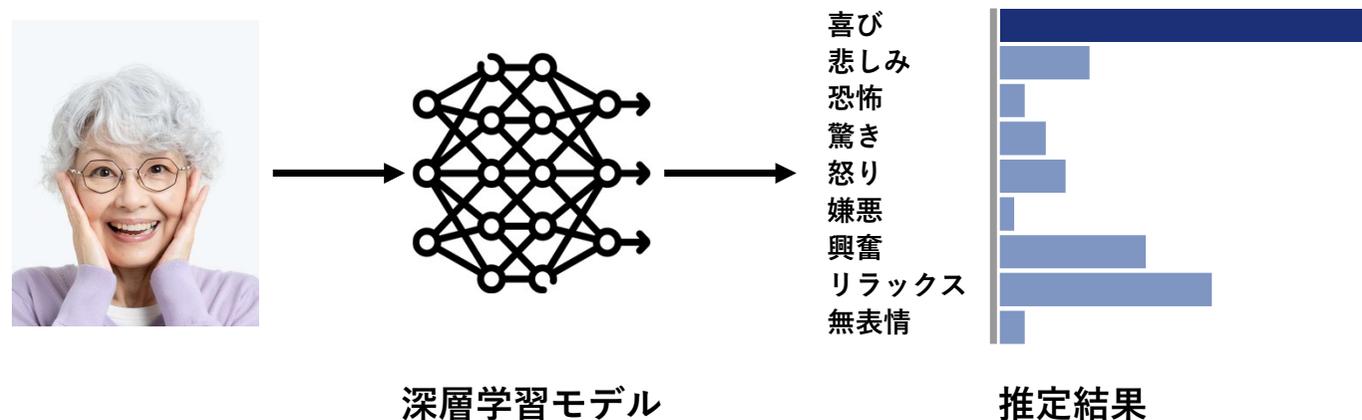
日本老年医学会 COI開示

発表者名：美添健，田中正行，荻田美穂子，月野木ルミ

演題発表に関連し、発表者らに開示すべき
COI関係にある企業などはありません。

概要

100人規模の日本人高齢者の顔画像に基づき、感情を推定する深層学習モデルを開発した。



日本人と外国人の間、高齢者と若年層の間には、表情の作り方や特徴に差異がある。[1][2]

→日本人高齢者特化のモデルの必要性

→大規模な日本人高齢者表情データセットを構築するのは収集コストの観点で難しい

→**小規模データセット**で高い精度のモデルを構築するため、**転移学習**を導入

[1] W. Sato, S. Hyniewska, K. Minemoto, and S. Yoshikawa, "Facial expressions of basic emotions in japanese laypeople," Frontiers in Psychology, vol.10, p.259, 2019.

[2] 中村真, and 益谷真. "高齢者の感情表出演技された表情の実証的検討." 感情心理学研究 7.2 (2001): 74-90.

転移学習

通常、深層学習で性能の高いモデルを構築するには、大量のデータが必要になる。

転移学習：

大規模なデータセットで学習したネットワークの一部をコピーして使うことで、学習した知識を別のタスクに転用する手法。

事前学習



スパコンで
丸一日以上

大規模データセット
(1万人程度 / 数十万枚以上)



特徴
抽出器

分類
器

推定結果

目的の顔特徴認識の学習



デスクトップPCで
1-2時間

小規模データセット
(100人程度 / 数千枚)



特徴
抽出器

分類
器

推定結果

初期値として利用

- ✓ 小データでも過学習しづらく、転移学習なしに比べて高い精度を出すモデルを構築できる
- ✓ 学習時間を短縮できる

事前学習

事前学習において、様々なアーキテクチャ・データセットを試し、最もターゲットタスクで精度が出る組み合わせを選定

ImageNet[1]
自然画像分類
128万枚



AffectNet[2]
顔表情分類
45万枚



CASIA-WebFace[3]
顔識別
49万枚



ネットワークアーキテクチャ

事前学習

	ResNet18	ResNet50	Face-ResNet18	Face-ResNet50
Scratch	0.2409	0.1432	0.2839	0.2800
ImageNet	0.3242	0.3464	0.3125	0.2982
AffectNet-8	0.3425	0.3294	0.3646	0.3203
CASIA+Softmax	0.2995	0.3190	0.3294	0.3568
CASIA+CosFace	0.2487	-(Pretraining Failed)	0.3372	0.3307
<u>CASIA+AdaFace</u>	0.2500	0.2083	0.3659	0.3451

目的の高齢者顔表情認識

データセット：名古屋大学 高齢者顔表情データベース [1]

このデータセットにおける表出感情を深層学習により推定し、人間による心理評定結果と比較

- 日本人高齢者111名（73.2±4.6歳，男性56名・女性55名）から取得された顔画像データセット
- 8感情（喜び・悲しみ・恐怖・驚き・怒り・嫌悪・興奮・リラックス）の表情+無表情をシナリオに基づく想起 or 写真の模倣で表出。
- 36名の人間（39.3±11.6歳，男性18名・女性18名）により、1枚あたり5名以上での心理評定が行われた。

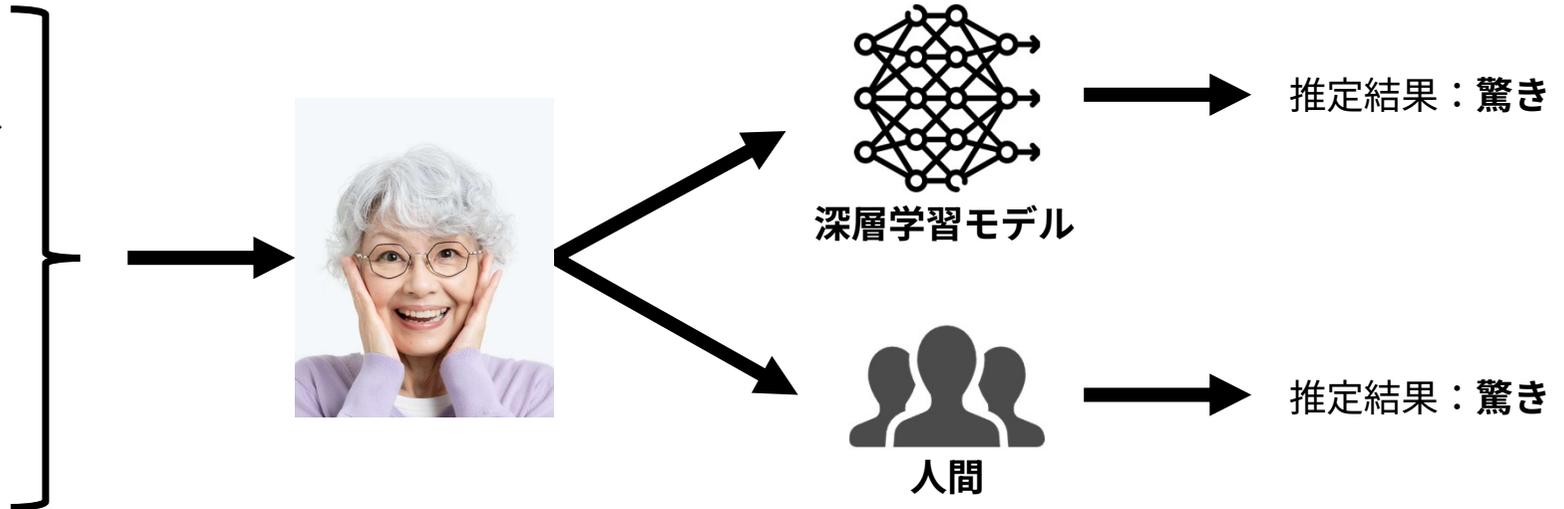
顔表情の表出の指示方法

シナリオに基づく想起

例) 「誕生日だと忘れていた日に、突然みんなにサプライズパーティーを開かれた。」

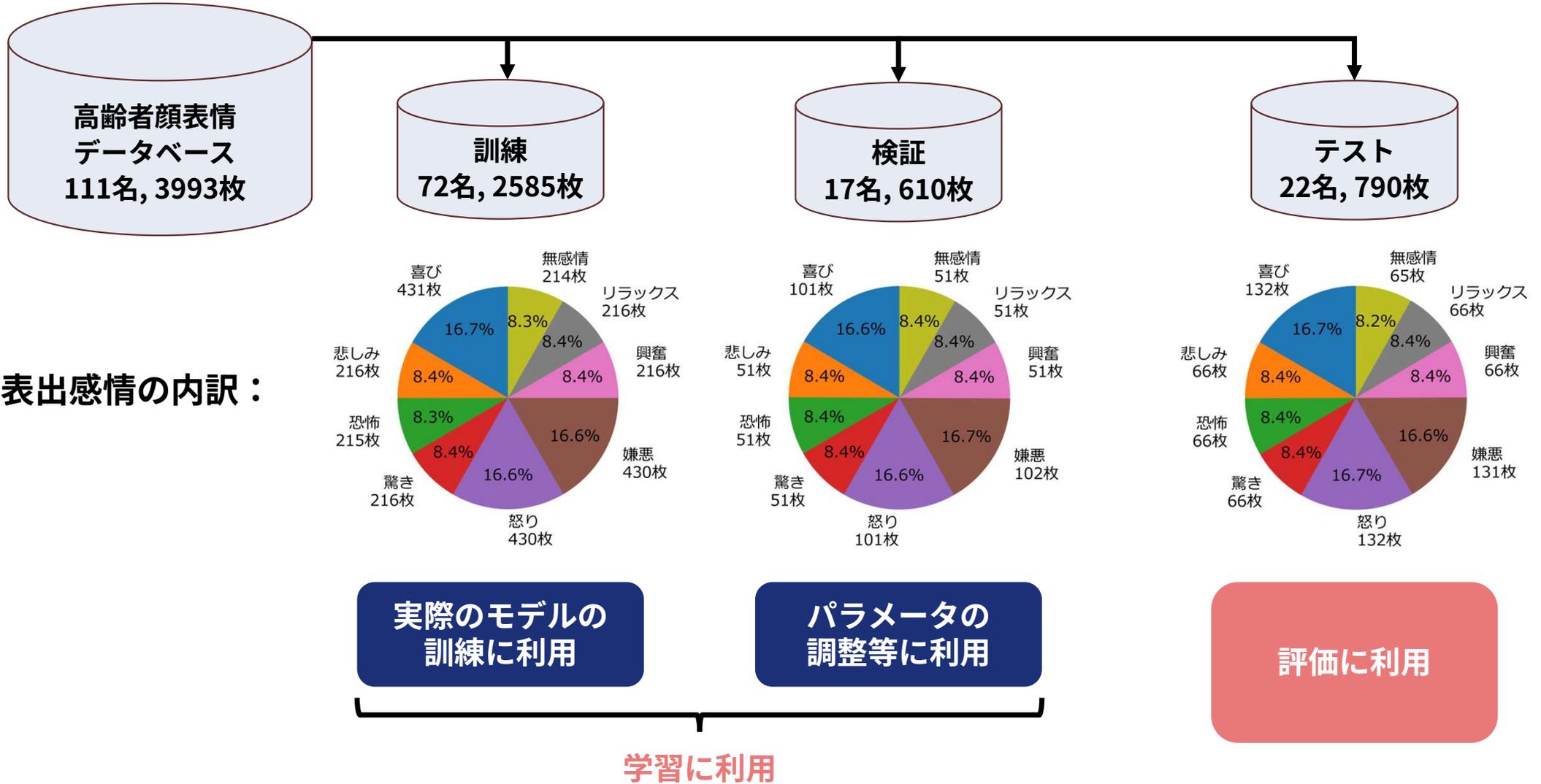
写真の模倣

例)



[1] 村上,吉田, et al. “日本人高齢者表情データベース構築のための感情ラベリングに基づく分析” HAI2022

高齢者顔表情データベースの分割

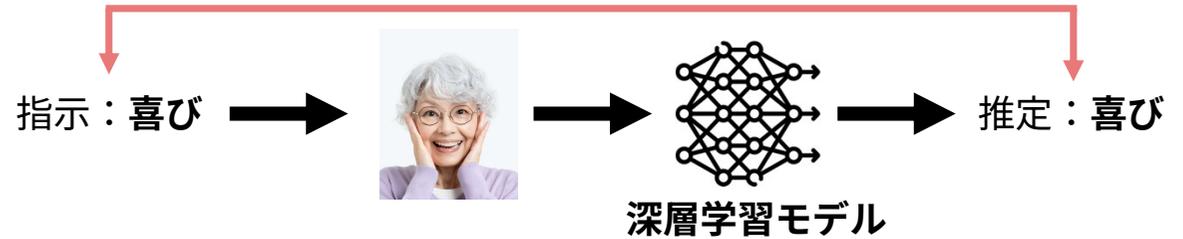


高齢者顔表情認識の評価

比較して評価

本研究：深層学習による結果

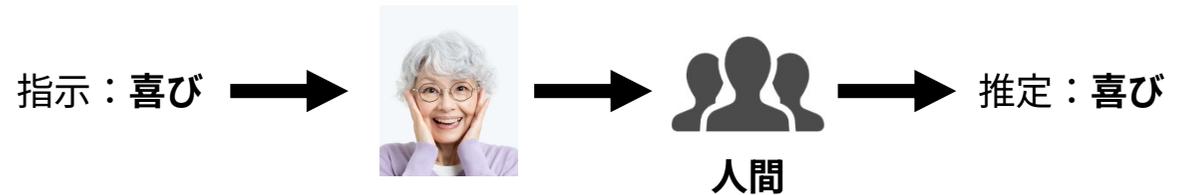
精度：36.6%



評価項目	マクロ平均	喜び	悲しみ	恐怖	驚き	怒り	嫌悪	興奮	リラックス	無感情
適合率	0.380	0.518	0.300	0.269	0.420	0.525	0.604	0.287	0.237	0.259
再現率	0.384	0.454	0.422	0.281	0.578	0.250	0.228	0.422	0.219	0.603
F値	0.355	0.484	0.351	0.275	0.487	0.339	0.331	0.342	0.228	0.362

人間による心理評定結果

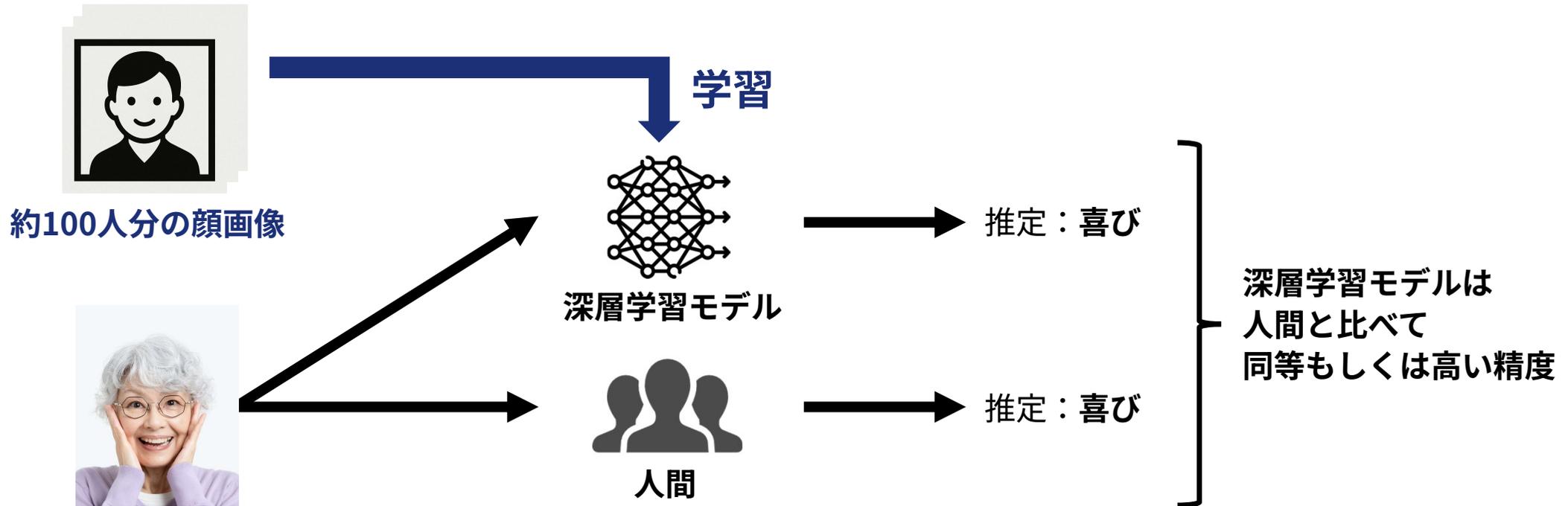
精度：34.1%



評価項目	マクロ平均	喜び	悲しみ	恐怖	驚き	怒り	嫌悪	興奮	リラックス	無感情
適合率	0.298	0.520	0.244	0.219	0.346	0.439	0.379	0.278	0.221	0.336
再現率	0.304	0.480	0.355	0.109	0.565	0.279	0.288	0.187	0.350	0.451
F値	0.290	0.499	0.282	0.146	0.429	0.341	0.327	0.223	0.271	0.385

[1] 村上,吉田, et al. “日本人高齢者表情データベース構築のための感情ラベリングに基づく分析” HAI2022

結論



今後の展望

同様の手法を用いた、顔表情推定以外のタスクのモデル構築・評価
関心を持っていただけた方、ご連絡いただくと幸いです。

Project Page: <https://www.vip.sc.eng.isct.ac.jp/proj/ffe/>

