

深層学習による複数種類の動物 に対する顔個体識別の特性比較

宮崎大学 大学院 工学研究科 工学専攻
機械・情報コース 情報システム分野

T2303394 長友祐磨

指導教員 椋木雅之

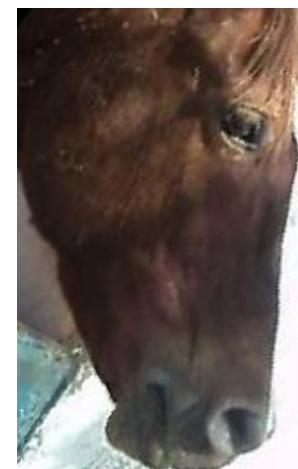
2025/1/30

研究背景

深層学習モデルを用いた人の顔個体識別は精度が高く実用化されている

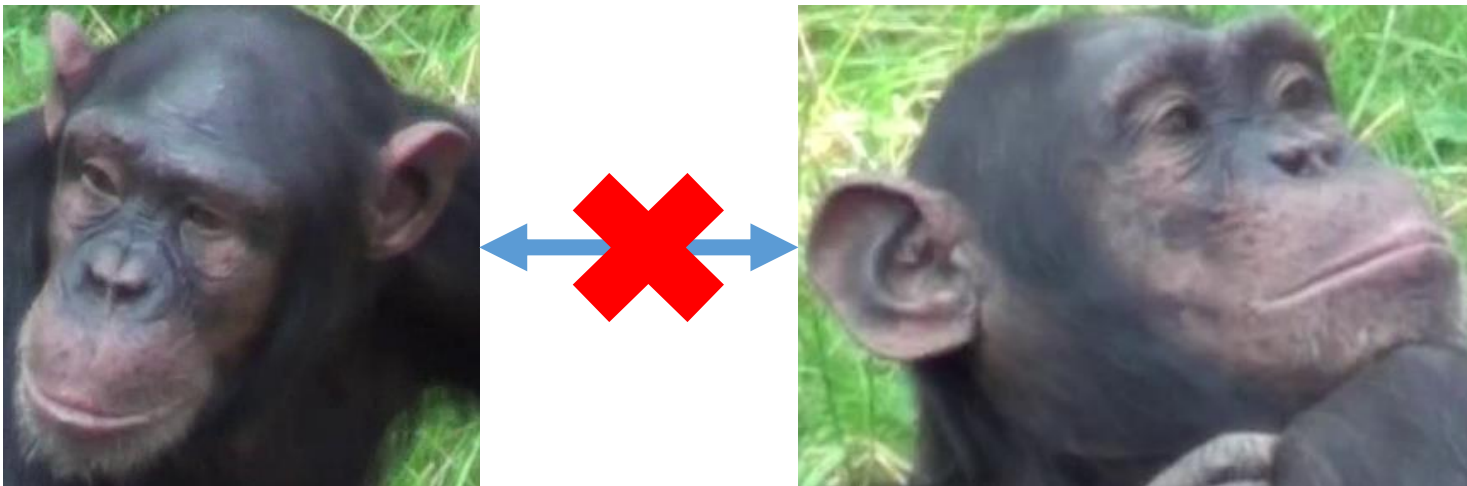
動物の顔個体識別の実用化も期待されている

例. 生態調査、動物保護、ペット管理、農業など



問題点

- 人のデータセットは大規模なため高精度を実現
- 動物の大規模なデータセットを作成することは難しい
- 特に識別したい動物のみでは十分に学習できない場合もある



研究目的

小規模データセットの動物の識別精度向上を目的とする

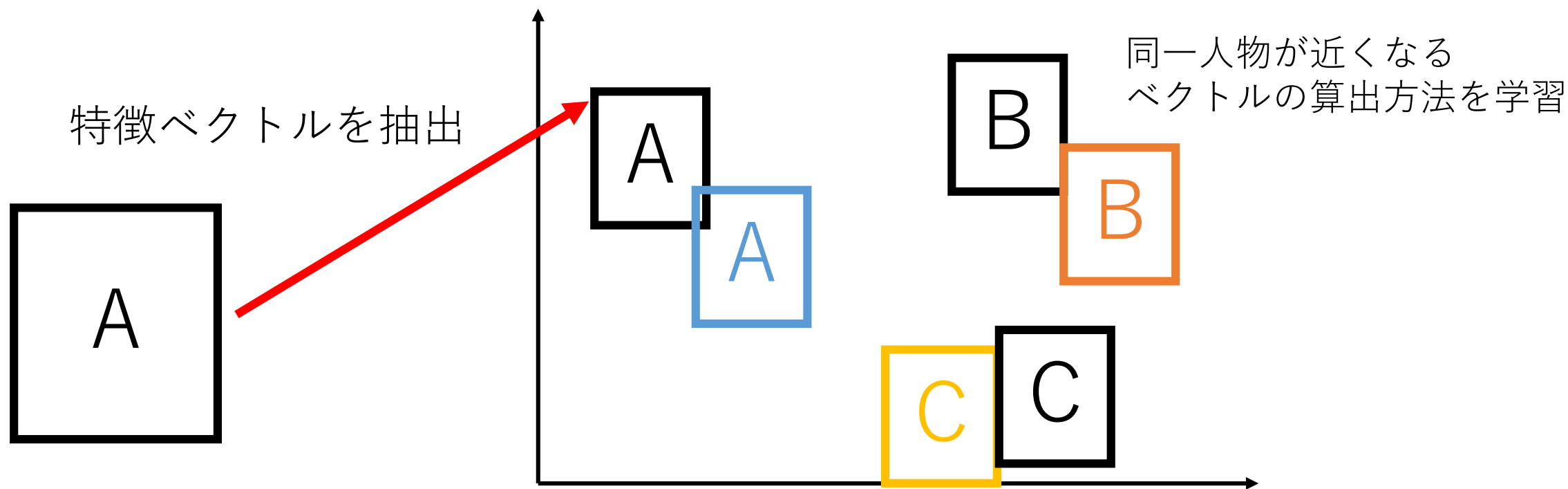
複数の動物のデータセットを組み合わせて
転移学習や**Joint Training**を行う



ArcFaceを用いた**識別精度**やGrad-CAMを用いた**ヒートマップ**で検証

人の顔認識技術

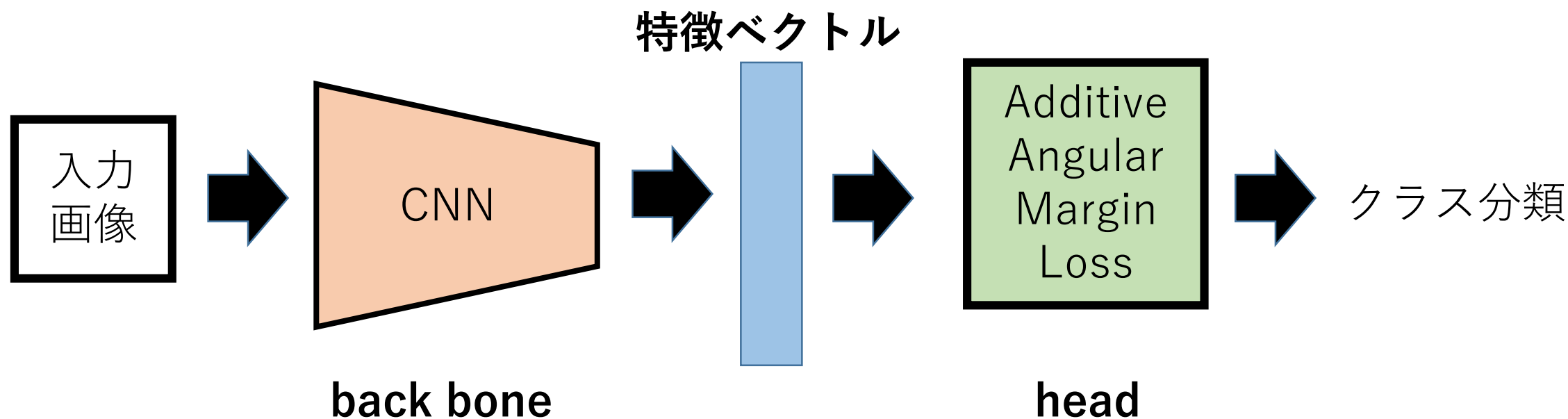
- 深層距離学習が主流
- 画像を埋め込み空間上で特徴ベクトルとして表現



ArcFace[1]

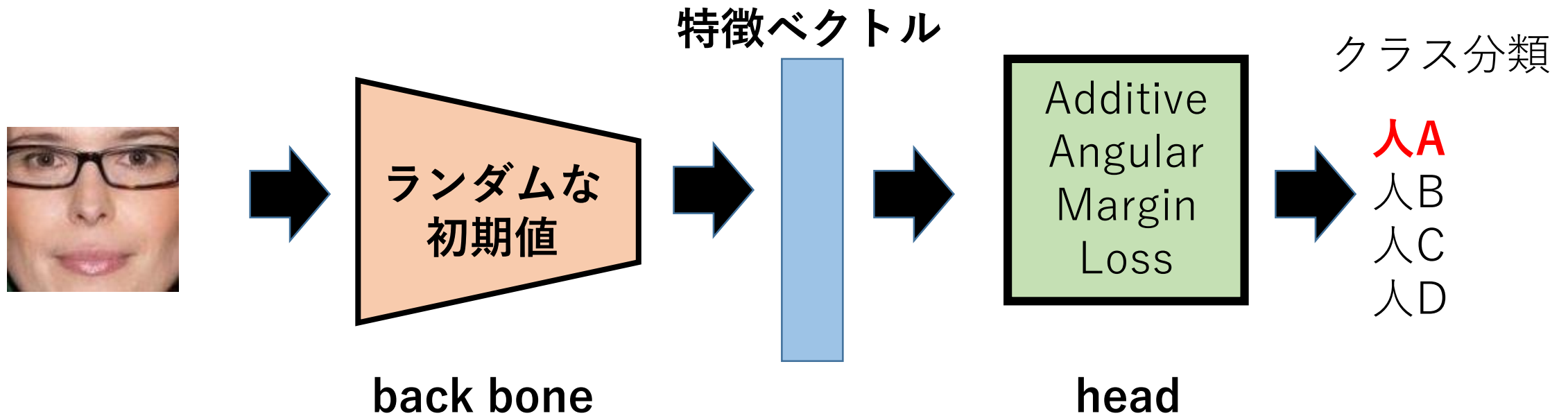
深層距離学習の代表的な手法

特徴ベクトル同士の角度(cos類似度)で識別を行う



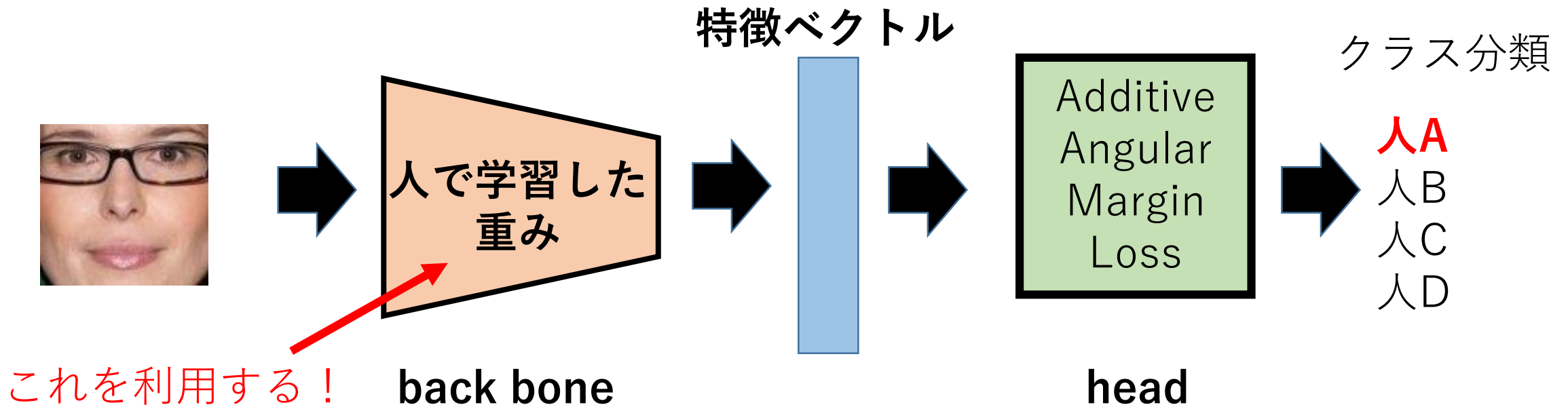
転移学習[2]

ある分野で学習されたモデルを別の分野の学習に適用
back boneの部分を利用する



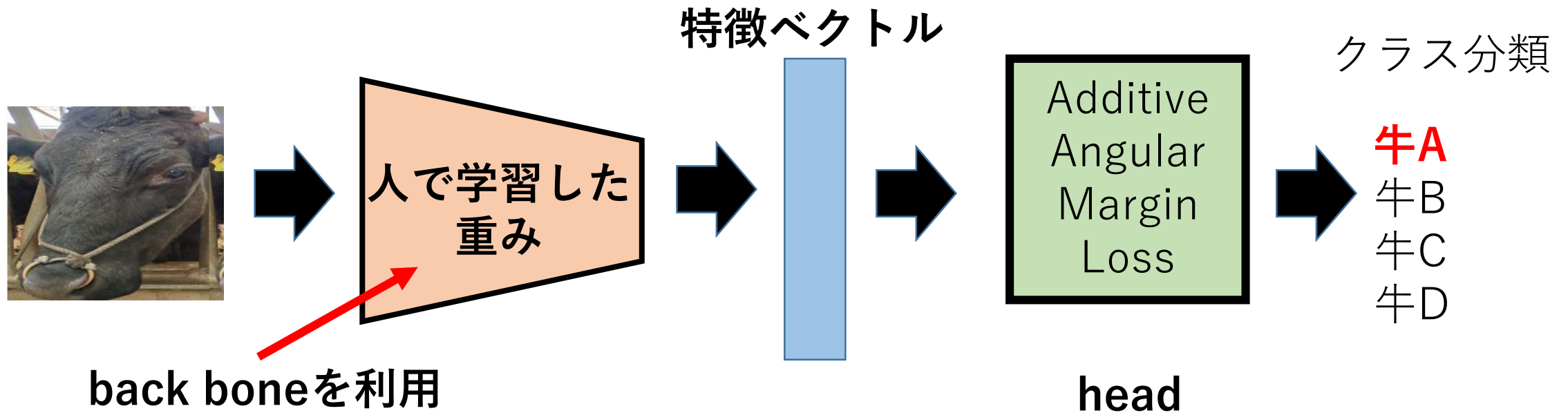
転移学習[2]

ある分野で学習されたモデルを別の分野の学習に適用
back boneの部分を利用する



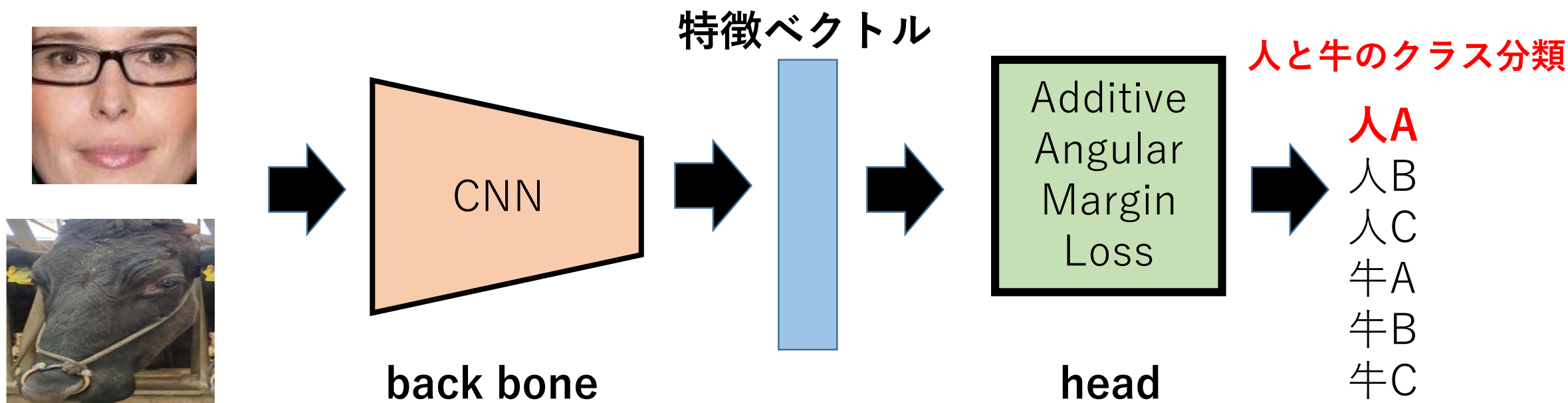
転移学習[2]

ある分野で学習されたモデルを別の分野の学習に適用
back boneの部分を利用する



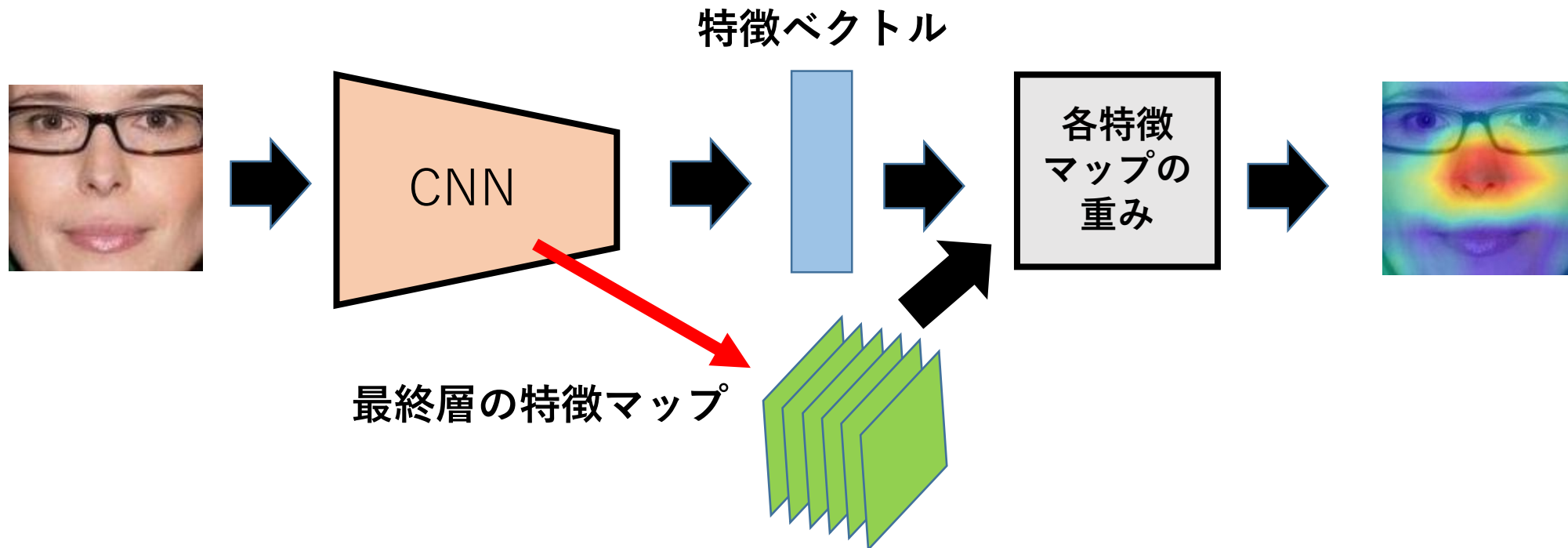
Joint Training[3]

複数の関連するタスクを同時に学習する手法
タスク間の相互関係を利用する



Grad-CAM[4]

ヒートマップで学習モデルの注目点を可視化
back boneの最終層の特徴マップの勾配からヒートマップを計算

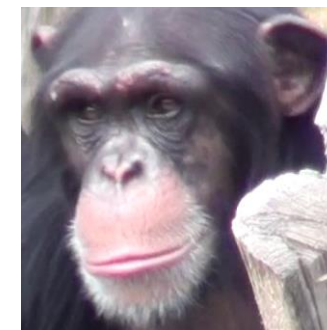


実験

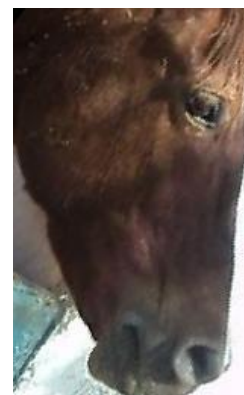
人、チンパンジー、牛、ヤギ、馬の5種類のデータセットを使用
様々な組み合わせで、転移学習やJoint Trainingを行う

- ArcFaceによる顔個体識別の精度
- Grad-CAMによるヒートマップの比較から有効性を検証

データセット



動物	個体数	画像枚数
人	9,131	3,308,040
チンパンジー	24	2,109
牛	475	491,581
ヤギ	10	1,291
馬	47	1,410



赤字：大規模データセット

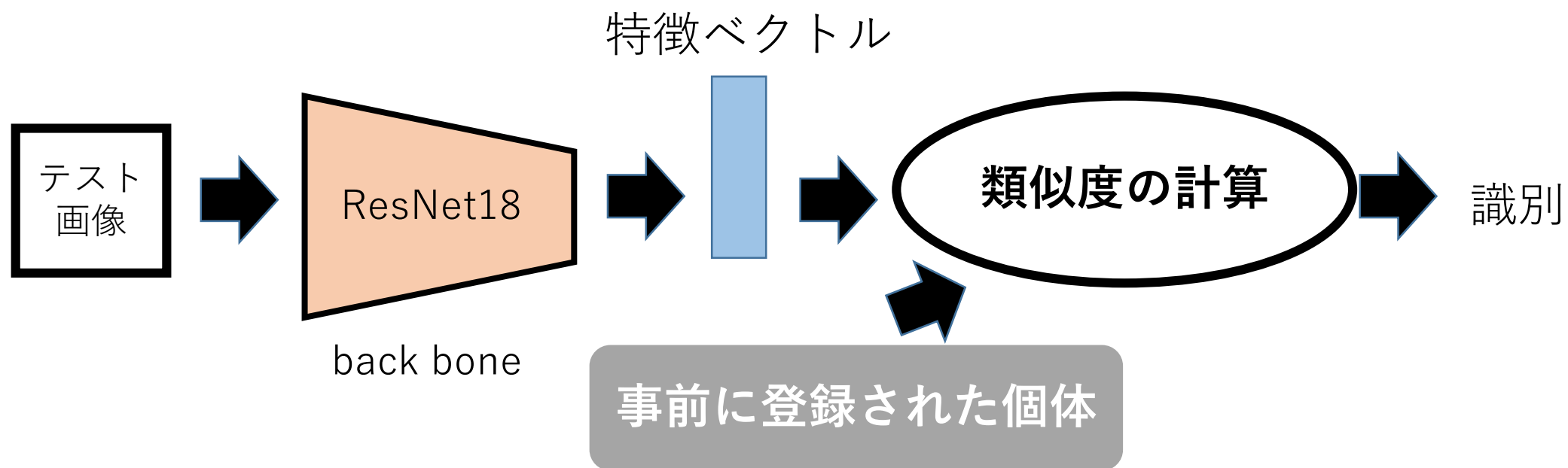
青字：小規模データセット

学習方法

- ArcFaceのback boneにはResNet18を利用
- すべての手法で100エポック
- 学習率は0.01から開始し、10エポックごとに0.5倍に減衰
- バッチサイズ300
- 各データセットは、訓練用とテスト用に分割し、訓練用とテスト用で異なる個体を使用

テスト時のArcfaceの構成

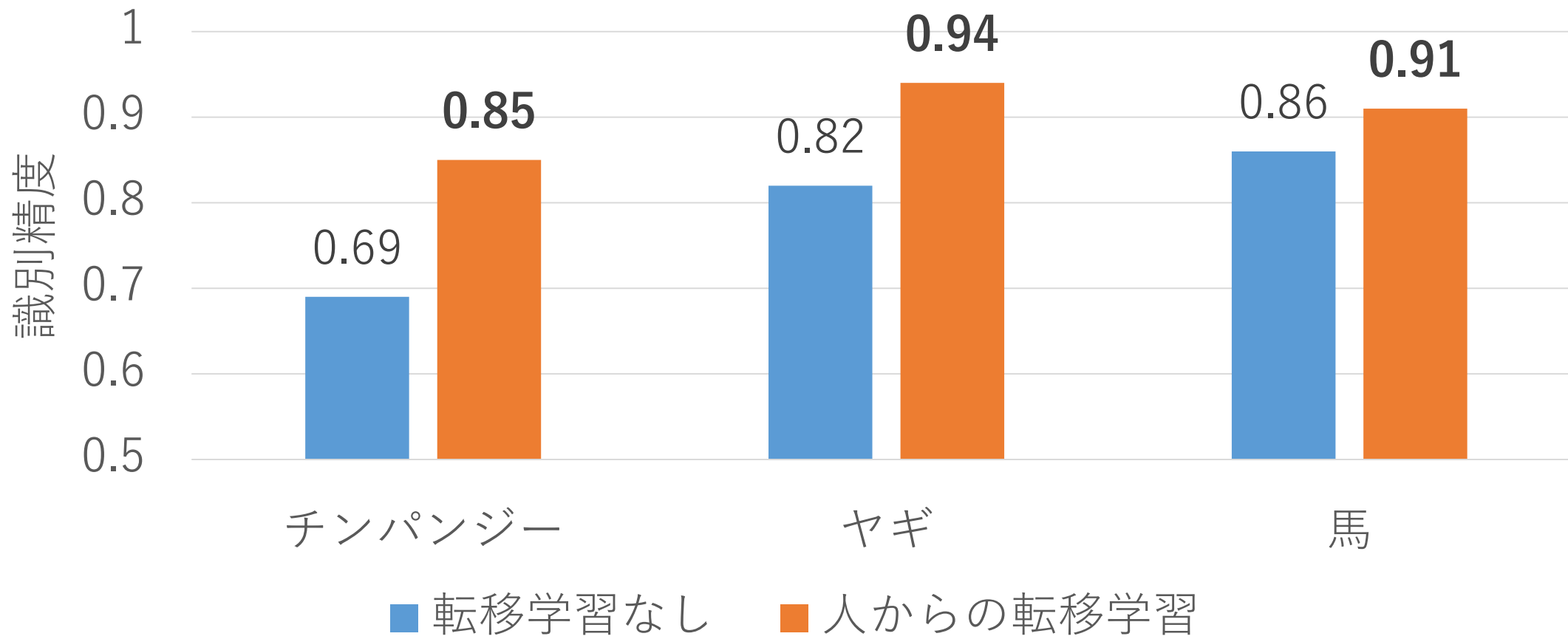
ArcFaceのheadを除き、特徴ベクトル同士のcos類似度で識別
テストデータの半分を登録し、残りのデータで識別精度を求める



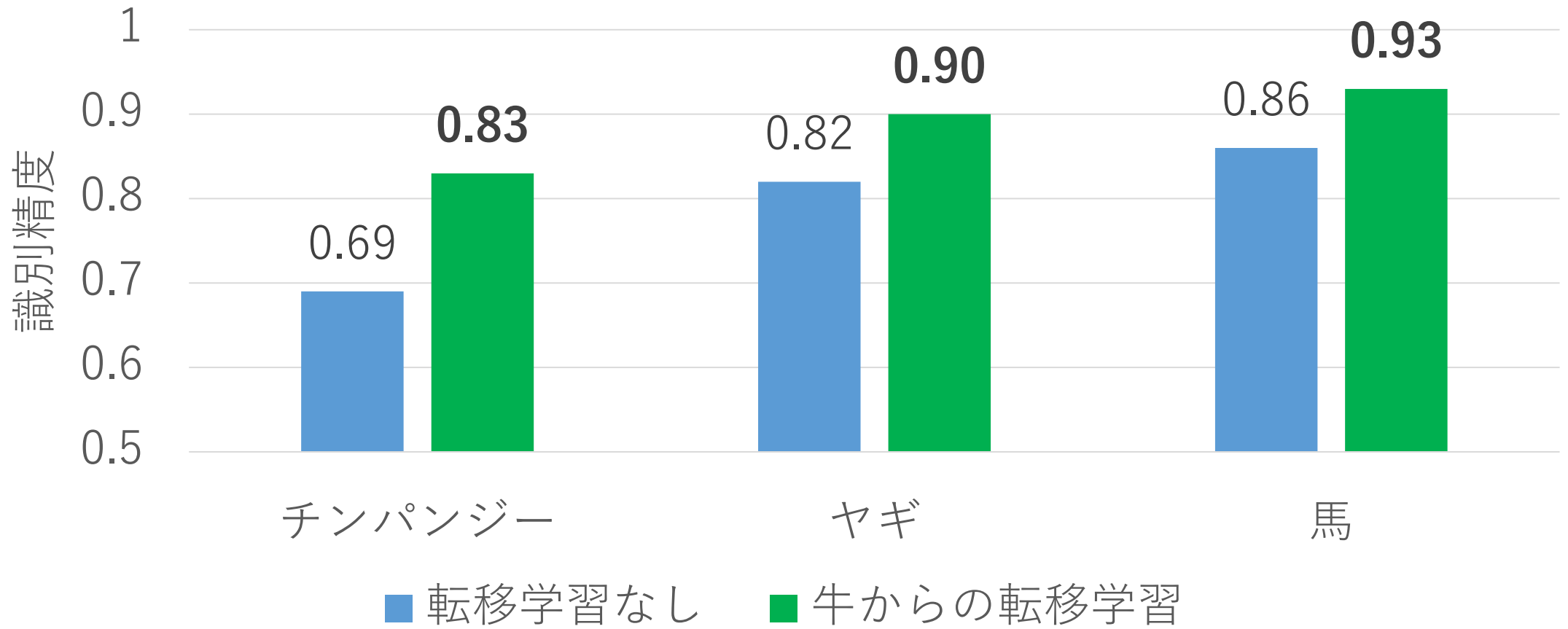
実験1：転移学習による比較

1. 大規模データセットからの転移学習
2. 小規模データセット同士での転移学習

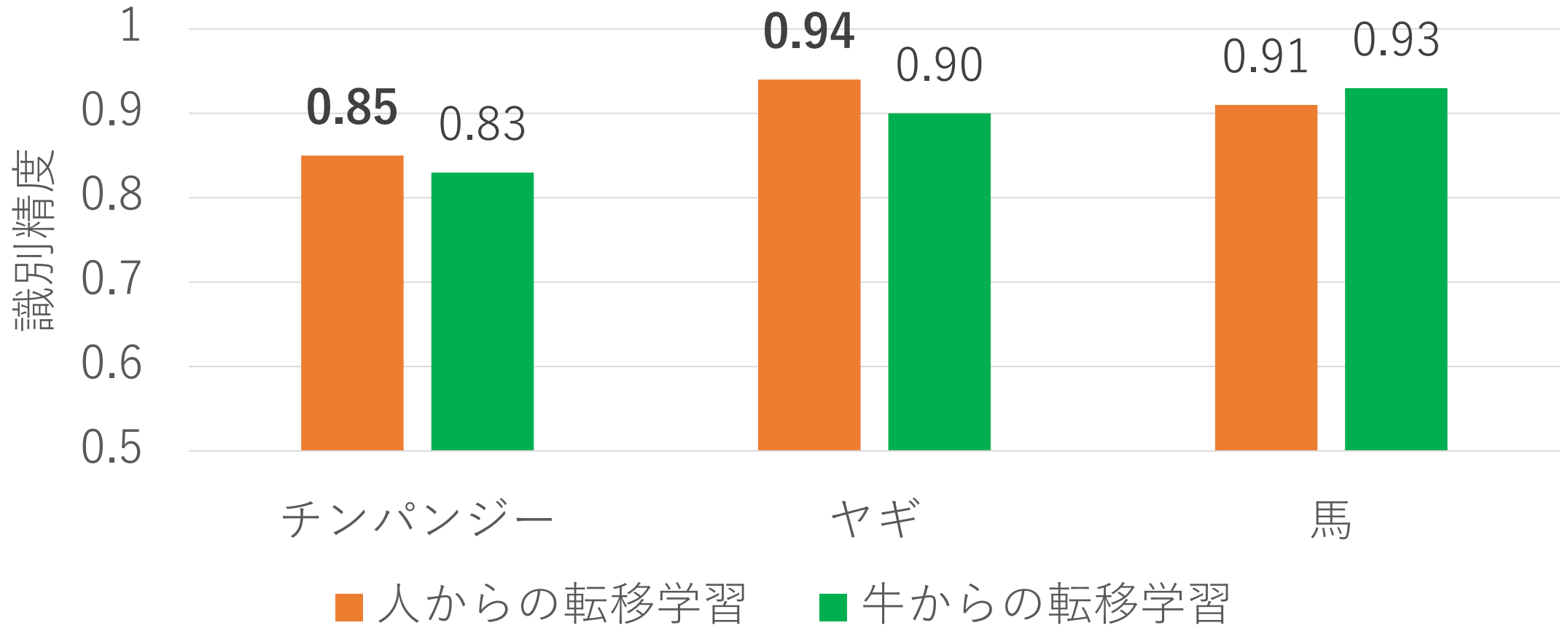
人からの転移学習



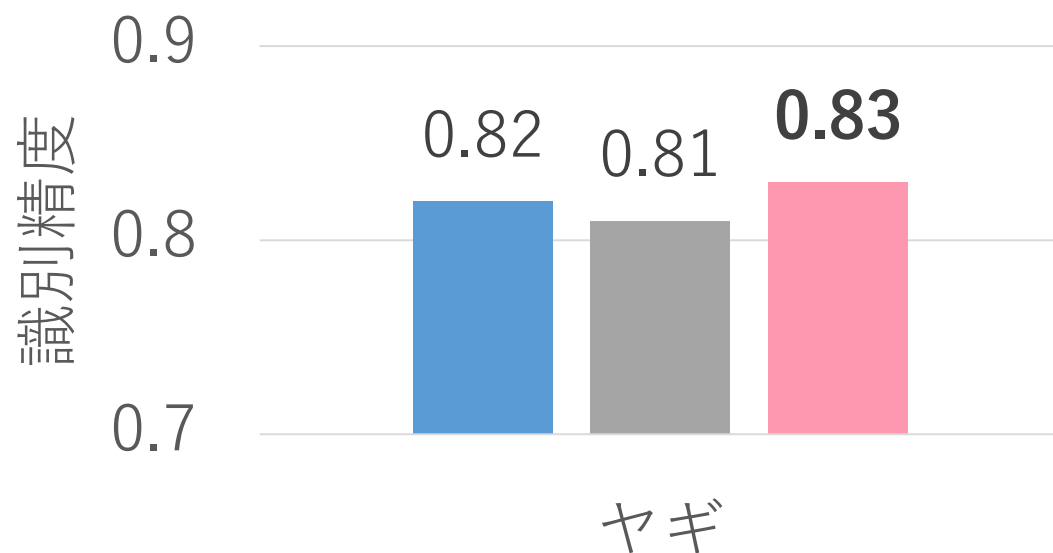
牛からの転移学習



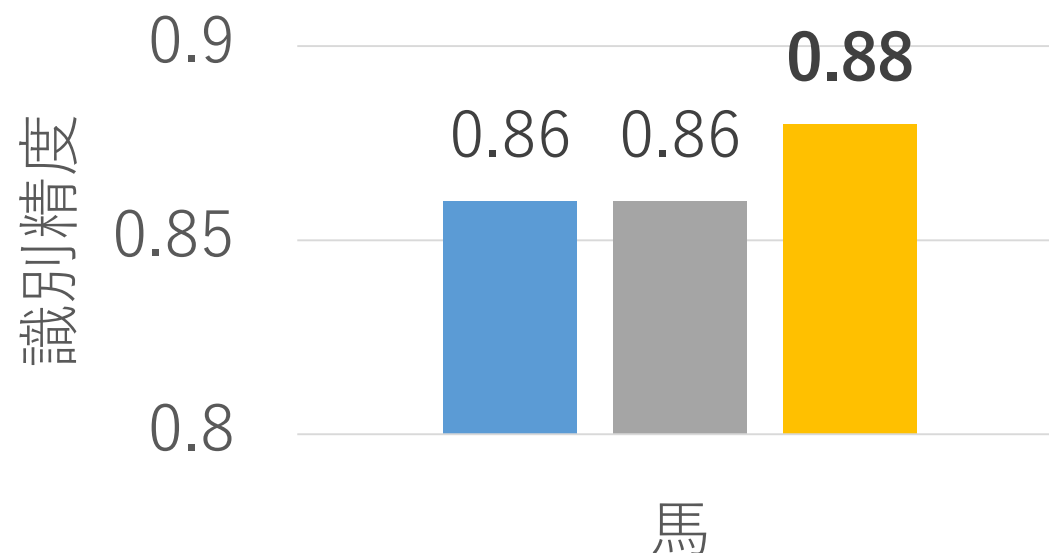
大規模データセットからの転移学習



小規模データセット同士の転移学習



- 転移学習なし
- チンパンジーからの転移学習
- 馬からの転移学習



- 転移学習なし
- チンパンジーからの転移学習
- ヤギからの転移学習

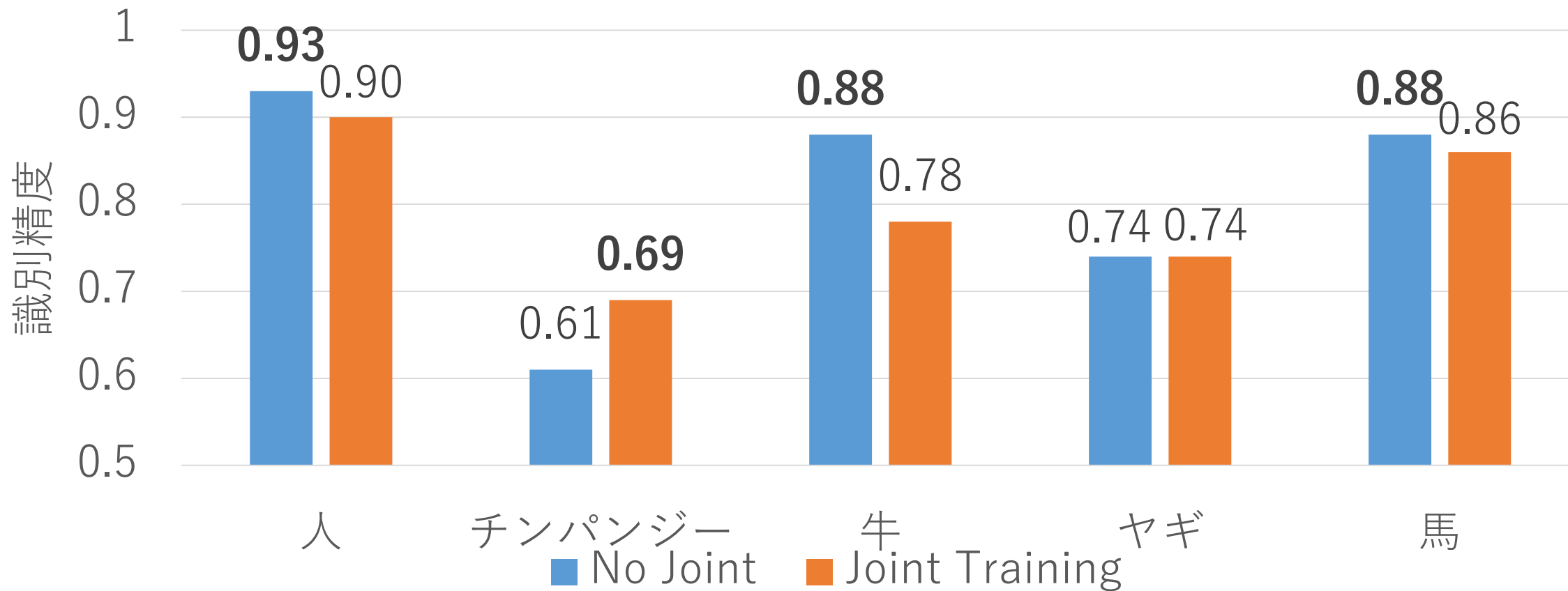
転移学習：まとめ

- 大規模データセットからの転移学習による識別精度の向上
- 似た顔同士の動物での転移学習が有効な傾向
例. チンパンジーと人や馬とヤギなど

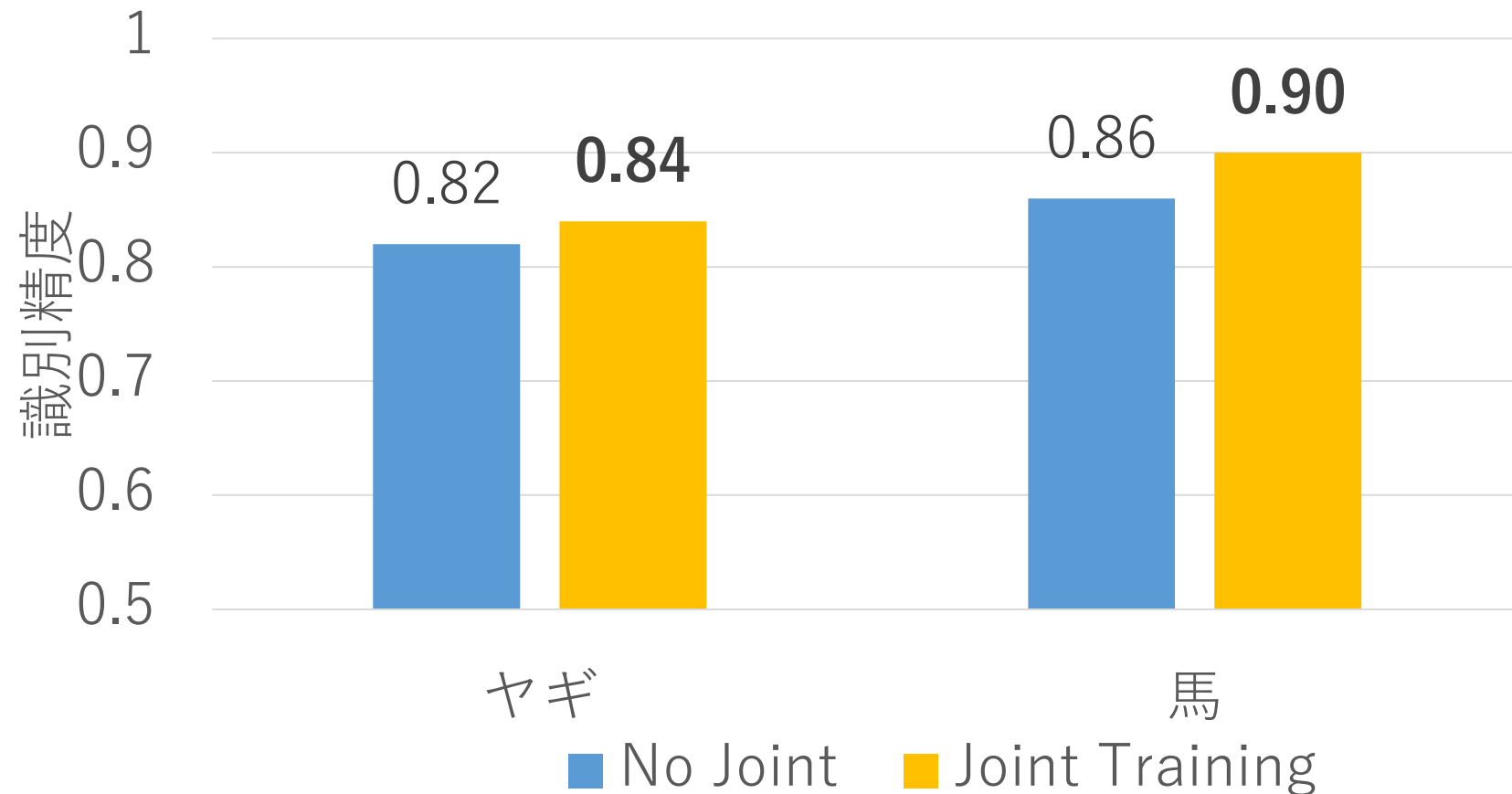
実験2：Joint Trainingによる比較

1. 5種類の動物でのJoint Training
2. 似た顔の小規模データセット同士(ヤギ・馬)でのJoint Training

5種類の動物でのJoint Training



ヤギ・馬のJoint Training



Joint Training : まとめ

- 5種類の動物を合わせたJoint Trainingはあまり効果なし
→人の割合が高すぎたため
- 似た顔の小規模データセット同士ではやや識別精度の向上

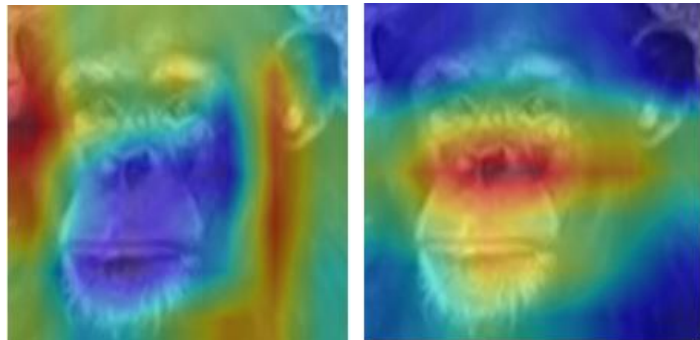
動物	個体数	画像枚数
人	9,131	3,308,040
チンパンジー	24	2,109
牛	475	491,581
ヤギ	10	1,291
馬	47	1,410

Grad-CAM：小規模データセットの転移学習

転移学習なしだと注目領域が不安定

転移学習によって顔を捉えられるようになっている

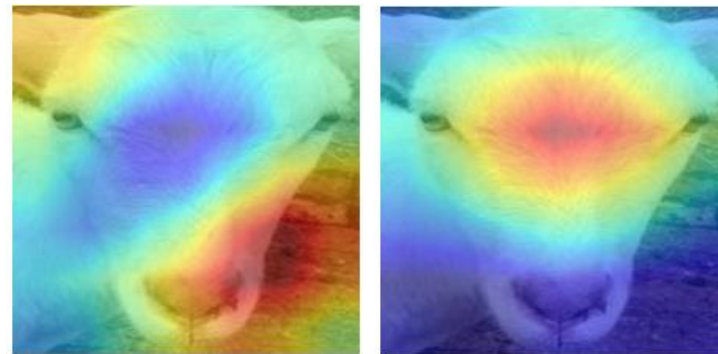
チンパンジー



転移学習なし

人からの
転移学習

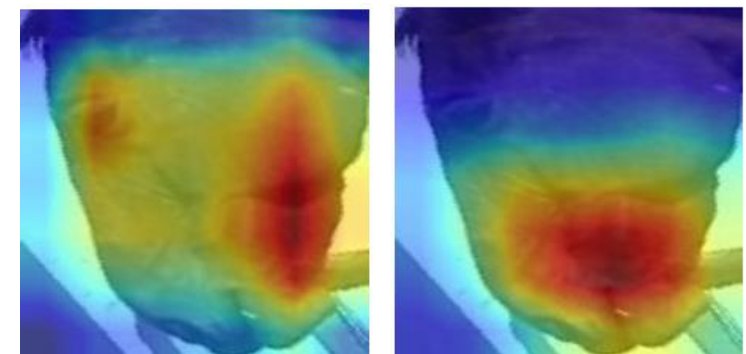
ヤギ



転移学習なし

人からの
転移学習

馬



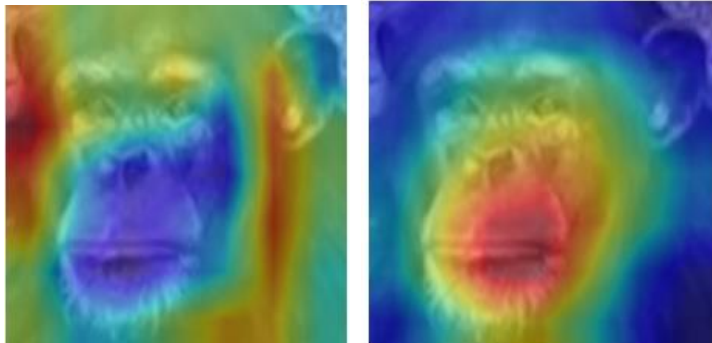
転移学習なし

牛からの
転移学習

Grad-CAM：小規模データセットのJoint Training

Joint trainingによって顔を捉えられるようになっている

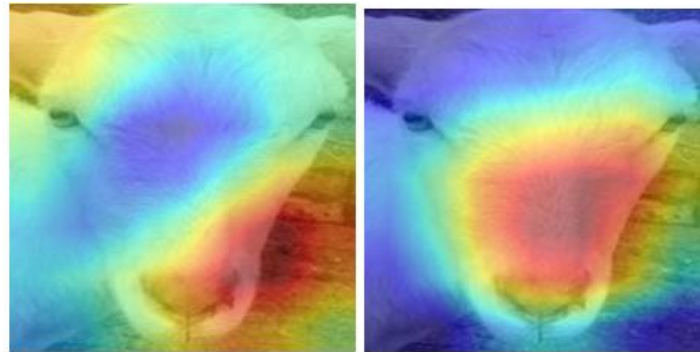
チンパンジー



No joint

joint

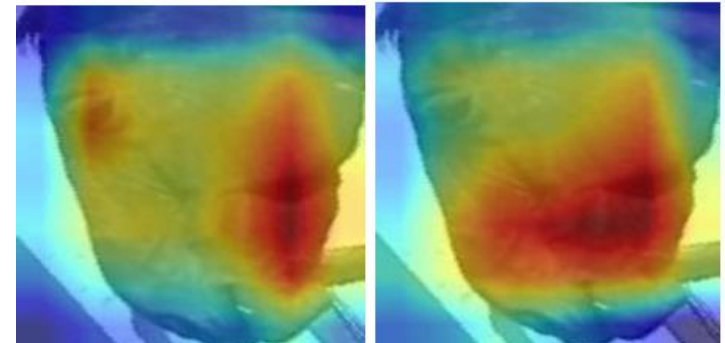
ヤギ



No joint

joint

馬



No joint

joint

Grad-CAM：まとめ

- 大規模データセットからの転移学習
→顔を捉えられるようになり、注目する場所が安定した
- Joint Training
→転移学習を行った学習モデルと同等に顔を捉えていた
→識別精度には表れていないが、学習モデルの性能は向上

まとめ

- 複数種類の動物を組み合わせた転移学習やJoint Trainingの有効性を検証
- 大規模データセットや似た顔同士の動物からの転移学習は有効
- 似た顔同士のJoint Trainingもやや有効

今後の課題

- より多くの動物での検証
- 野外環境など実際の使用を想定した評価

参考文献

- [1] Jiankang Deng, Jia Guo, Jing Yang, Niannan Xue, Irene Kotsia, and Stefanos Zafeiriou. Arcface: Additive angular margin loss for deep face recognition. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 44, No. 10, p. 5962–5979, October 2022
- [2] Lisa Torrey and Jude Shavlik. Transfer learning. In *Handbook of research on machine learning applications and trends: algorithms, methods, and techniques*, pp. 242–264. IGI global, 2010.
- [3] Jonathan Tompson, Arjun Jain, Yann LeCun, and Christoph Bregler. Joint training of a convolutional network and a graphical model for human pose estimation. In *Proceedings of the 28th International Conference on Neural Information Processing Systems- Volume 1, NIPS'14*, p. 1799–1807, Cambridge, MA, USA, 2014. MIT Press.
- [4] Ramprasaath R Selvaraju, Michael Cogswell, Abhishek Das, Ramakrishna Vedantam, Devi Parikh, and Dhruv Batra. Grad-cam: Visual explanations from deep networks via gradient-based localization. In *Proceedings of the IEEE international conference on computer vision*, pp. 618–626, 2017.