

FaceNetを用いた牛顔画像による個体識別の ハイパーパラメータチューニング

宮崎大学 工学部 情報システム工学科

67200351 戸高 海輝

指導教員 椋木 雅之

令和6年2月15日

牛個体識別の必要性

牛の健康管理や繁殖管理

繁殖計画の策定・基礎データの提供

→生産性の向上やコスト削減に寄与

牛トレーサビリティ制度

牛の個体識別情報の届出が義務化

→正確な牛の個体管理が求められる



牛個体識別の従来手法

耳標から読み取る

→汚れや牛の動きで読み取りにくい / 破損・欠落により利用できない

タグの埋め込み

→牛にストレスを与える / タグ読み取り装置のコストが高い

このような問題を解決するために、画像認識技術が用いられる

画像認識を利用した牛個体識別

深層学習を用いた牛個体識別

- 時任による研究

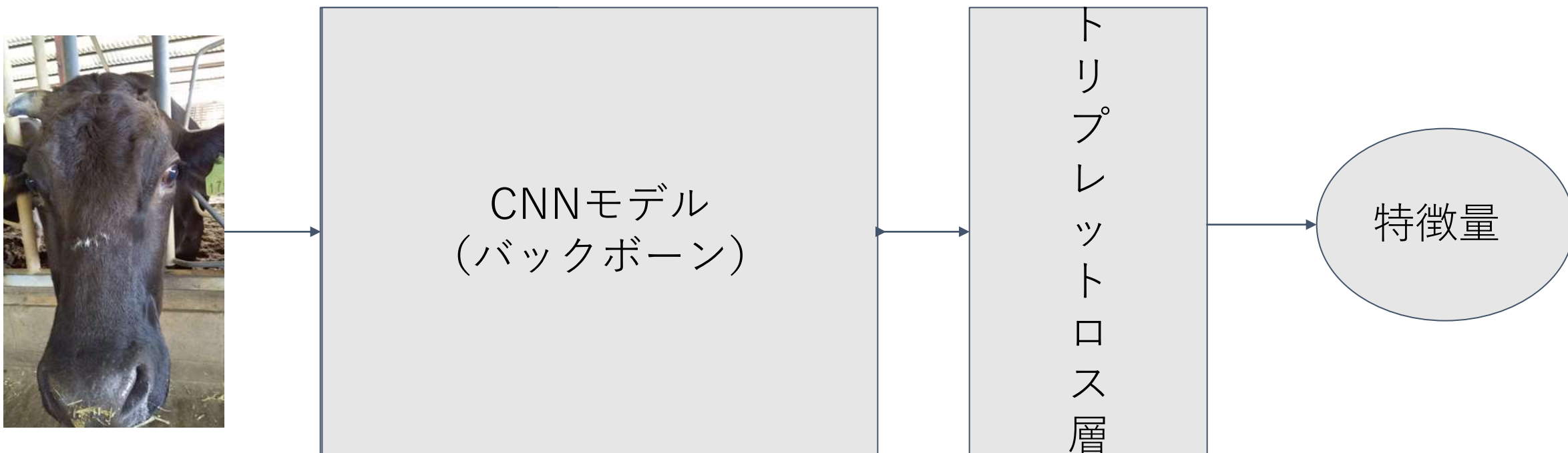
FaceNetの牛への適用可能性を調査した

→FaceNetの学習を1通りの設定でしか行っておらず、学習時のハイパーパラメータを調整すれば、より精度の高い個体識別が実現できると考えた

研究目的

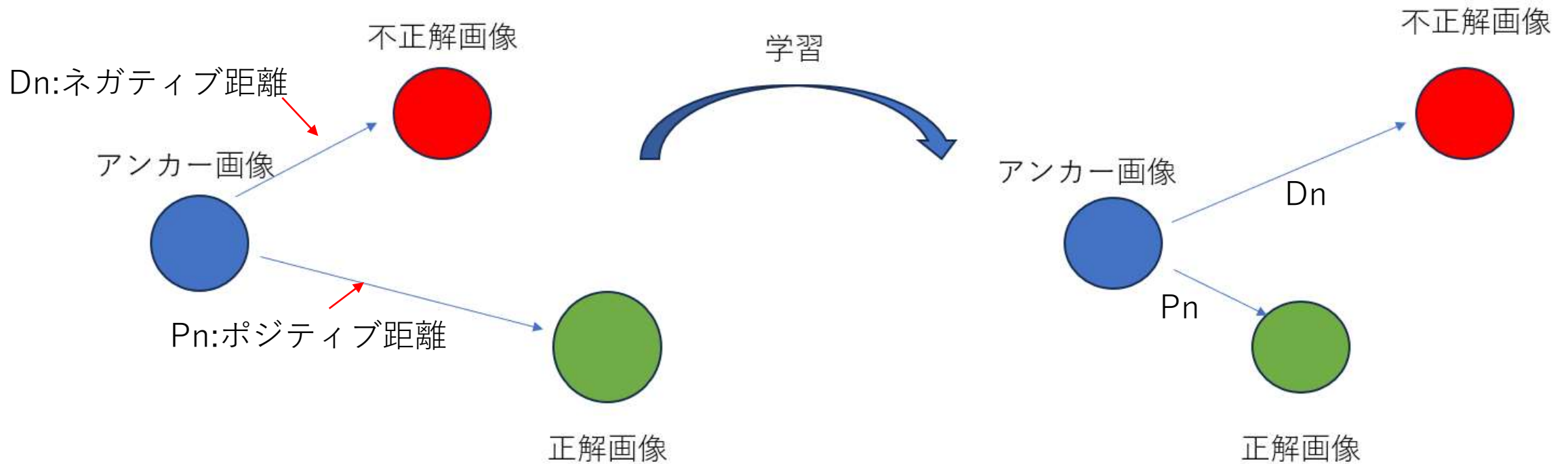
牛顔画像による個体識別の精度を上げるハイパーパラメータチューニングについて調査する

FaceNetによる個体識別



トリプレットロス層

類似した画像間の距離を近づけ、類似していない画像間の距離を離すことを目的とする。



ハイパーパラメータ

ハイパーパラメータとは機械学習アルゴリズムの挙動を設定するパラメータを指す。

本論文で調整するハイパーパラメータ

- CNNモデル
 - バックボーンとして利用する既存のモデル
- データ拡張
 - 学習用データを増やすために施す処理
- 学習率
 - 最適化手法において一度に修正する重みの量を定めるもの
- トリプレットの選択法
 - ハードネガティブとセミハードネガティブを比較

トリプレットの選択法

- ハードネガティブ
 - ネガティブ距離 - ポジティブ距離 < マージンを満たすトリプレットのみを学習に使用
- セミハードネガティブ
 - ハードネガティブの条件 + ポジティブ距離 < ネガティブ距離を満たすトリプレットのみを学習に使用

実験

使用したデータ

- 学習用データ
 - 牛349頭、112,952枚の顔画像が含まれている
- 評価用データ
 - 学習用データとは異なる牛81頭、49,302枚の顔画像が含まれている

評価指標

$$\text{Accuracy} = \frac{(\text{True_positive} + \text{True_negative})}{\text{すべてのデータ}}$$

True_positive = (本当は同じ牛に対して同じ牛と予測したケース) ←

True_negative = (本当は異なる牛に対して異なる牛と予測したケース) ←

評価用データ6000組のデータを10個に分割し、そのうちの9個を使い最適な閾値を定め、残り1個のAccuracyを求める。これを10通り行い、その平均の値を評価結果とする。

パラメータチューニングの基準値

GitHubで公開されているFaceNetのプログラムのハイパーパラメータのデフォルト値を使った結果を、チューニング前の基準値とする。

表 1: 本論文の基準となる値を求めるために使用したハイパーパラメータとその精度

ハイパーパラメータ	値
CNN モデル	ResNet18
データ拡張	左右反転、回転変換
学習率	0.005
エポック数	100
トリプレットの選択法	ハードネガティブ
Accuracy	0.92633

CNNモデルの評価

バックボーンとして「ResNet18」「ResNet34」
「ResNet50」「ResNet101」「MobileNetv2」を
それぞれ使用して学習

結果はResNet34が最良

CNN モデル	精度(Accuracy)
Resnet18	0.92633
Resnet34	0.92883
Resnet50	0.91800
Resnet101	0.92283
Mobilenetv2	0.91660

データ拡張

基準となるデータ拡張にぼかし、アフィン変換、色変換、CutOutをそれぞれ追加して学習をした。
その結果CutOutとぼかしで精度が向上

組合せとしてCutOutとぼかしを追加したとき精度が向上
しかしそこにさらにアフィン変換を追加したら逆に精度が落ちた

データ拡張	精度 (Accuracy)
左右反転、回轉變換	0.92633
+ぼかし	0.92766
+アフィン変換	0.92633
+色変換	0.88716
+ <u>CutOut</u>	0.92966
+ <u>CutOut</u> +ぼかし	0.93033
+ <u>CutOut</u> +ぼかし+アフィン変換	0.92616

学習率の評価

学習率を5種類使用して学習

結果は0.001の 때가最良であった。
学習率を上げてても下げてても精度
が落ちたため学習率は0.001が最
良

学習率↵	精度(Accuracy)↵
0.005↵	0.94416↵
0.001↵	0.95266↵
0.0005↵	0.95233↵
0.0001↵	0.94783↵
0.00005↵	0.94683↵

トリプレットの選択法の評価

基準であるハードネガティブとセミハードネガティブを比較した

結果ハードネガティブの方が精度が高かった。

ハードネガティブよりもセミハードネガティブの方が学習に使うトリプレットを減らし、学習が効率的になると考えていたが逆に精度が落ちてしまった。

表5：トリプレットの選択法による評価結果

トリプレットの選択法	精度(Accuracy)
ハードネガティブ	0.95266
セミハードネガティブ	0.94916

結果

ハイパーパラメータの最良の組み合わせ

基準となるAccuracyより0.026
ポイント精度が向上

ハイパーパラメータ	値
CNN モデル	ResNet34
データ拡張	左右反転 + 回転変換 + <u>CutOut</u> + ぼかし
学習率	0.001
エポック数	500
トリプレットの選択法	ハードネガティブ
精度 (Accuracy)	0.95266
基準となる Accuracy	0.92633

まとめ

牛顔画像による個体識別の精度を上げるハイパーパラメータチューニングについて調査

- ハイパーパラメータの組み合わせで最良だったのは、
CNNモデルにResNet34、
データ拡張にCutOutとぼかしを追加、
学習率に0.001、
トリプレットの選択法にハードネガティブ
を使用した場合の0.95266であった

今後の課題

- 調査したハイパーパラメータの種類（例えば、オプティマイザーやデータ拡張の種類や組合せなど）を増やす