

知能情報総合実験 (データマイニング班)

G2:

225756E 天久ジェイデン玲, 225758B 山城りん,
225760C 喜舎場唯斗, 235706C 比嘉琢真,
235722E 安谷屋 紘

2025 年 8 月 6 日

1 テーマ「犬種の識別」について

本グループでは、日本犬 6 種 (秋田犬、柴犬、甲斐犬、紀州犬、四国犬、北海道犬) の画像分類を対象問題として設定した。日本犬とは、一般社団法人ジャパンケネルクラブ [5] によると、日本原産の 6 種の犬種を指し、それぞれ外見や体格に特徴を持つ。しかし、犬種間での視覚的な違いは微妙であり、人間でも判別が難しいことがある。本実験では、これらの犬種を画像から分類するモデルを構築することで、視覚的な類似性に対する機械学習モデルの識別性能を検証するとともに、画像前処理やデータ構築の重要性を学ぶことができる。また、データセットの独自構築を通して、実用的な画像収集・加工の技術も習得できるため、教育的意義が大きいと考えられる。

2 実験方法

2.1 実験目的

本グループでは、日本犬 6 種 (秋田犬、柴犬、甲斐犬、紀州犬、四国犬、北海道犬) の画像を対象とした犬種識別モデルの構築に取り組む。日本犬は外見的な違いが微妙であるため、人間による識別が難しい場合もある。そこで、機械学習を用いた自動分類が有効かどうかを検証したいと考えた。

本実験を通して、画像処理や深層学習 (CNN) の基礎的な理解を深めながら、前処理から学習・評価までの流れを自ら設計・実装する力を身につけることを目的とする。

2.2 データセット構築

本実験では、日本犬の犬種識別を目的とした画像データセットを構築した。対象とする犬種は、柴犬、秋田犬、紀州犬、北海道犬、甲斐犬、四国犬の 6 種類である。まず、各メンバーが 1 種類の犬種を担当し、Google 画像検索を用いて該当犬種の写真を収集した。今回は画像検索の結果には誤ったラベルが付与されている可能性もあるため、犬種ごとの典型的な見た目に基づき、目視による確認とラベルの再チェックを行い、重複画像や犬種が不明瞭な画像を除外しつつ、各犬種につきおよそ 50~100 枚程度の画像を収集することができた。また、機械なし学習は、日本の犬種の特徴が類似しているため正確性が非常に低く、正答率を上げることが困難と判断したため今回の実験では機械なし学習でのデータの分類は行わず実行することにした。

収集した画像は、それぞれの犬種名をフォルダ名としたディレクトリ (例:dataset/柴犬) に分類・保存した。次に、収集したオリジナル画像 (約 50~100 枚) を訓練用とテスト用に 8:2 の比率で分割し、以下のような構成で保存した。

- 訓練用：dataset/柴犬/柴犬 訓練用
- テスト用：dataset/柴犬/柴犬 テスト用

しかし、機械学習においては十分な学習性能を得るためには、より多くのデータが必要である。そこで本研究では、訓練用・テスト用それぞれの画像に対して以下のデータ拡張処理を行い、各犬種に対するデータセットの枚数を増やした。

- 画像の回転 1 (90° , 180° , 270° 回転を追加。)
- 画像反転
- 明るさの調整 (明るさの倍率 (元の明るさは 1.0) を 0.6, 0.8, 1.2, 1.4 を追加。)
- 画像の回転 2 (0° を除く 0~15° の微調整回転をランダムで追加。実験 3 以降適用し、詳しい詳細は実験 3 で記載。)
- 背景の削除 (実験 4 のみで適用し、詳しい詳細は実験 4 で記載。)

背景の有無が分類性能に及ぼす影響を調査するため、「背景削除あり」と「背景削除なし」の 2 パターンのデータセットを作成し、それぞれのモデル精度を比較することにした。

以上のようにして構築したデータセットを用いて、畳み込みニューラルネットワーク (CNN) による画像分類モデルの学習と評価を行う。

ただし、画像検索の結果には誤ったラベルが付与されている可能性もあるため、犬種ごとの典型的な見た目 (毛色、耳の形、尾の巻き方等) に基づき、目視による確認とラベルの再チェックを行った。それでもなお一部の誤ラベルが含まれる可能性は否定できず、今後は信頼性の高い公的データセットの導入や、専門家の協力を得た精査が望まれる。

2.3 モデル選定

本実験では、事前学習済みの畳み込みニューラルネットワーク (CNN) である MobileNetV2[1] をベースに転移学習を行った。MobileNetV2 は ImageNet データセット [2] で事前学習された軽量のモデルであり、少量データに対しても高い精度を発揮するという利点がある。

具体的には、TensorFlow の Keras API を用いて事前学習済みの MobileNetV2 を読み込み、最終の全結合層 (出力層) を日本犬 6 種の分類に適した形に置き換えた。その上で、新たに追加した出力層を含む全体を再学習する形で再学習を行った。

CNN は画像内の局所的な特徴を自動で抽出することができるため、犬の耳や顔、毛並みなどの視覚的な違いを識別するのに適している。本課題でも、犬種間の微妙な違いを捉えるために有効であると考えられる。また、全結合層のみのネットワークと比較してパラメータ数が少なく、過学習のリスクを抑えながら計算効率にも優れている点が、本モデルの採用理由の一つである。さらに、MobileNetV2 はこれまで多くの画像処理タスクで高い精度を示しており、信頼性のある基盤モデルとして本課題への応用が妥当であると判断した。

構築したモデルの構成は以下の通りである。

- 入力層 (128 × 128 ピクセル、RGB 画像)
- MobileNetV2 ('include_top=False' により出力層を除去、事前学習済み重みを保持)
- GlobalAveragePooling2D 層
- Dropout 層 (ドロップ率: 0.3)
- Dense 層 (128 ユニット、活性化関数: ReLU)
- 出力層: Dense 層 (6 ユニット、活性化関数: softmax)

2.4 パラメータ調整

入力画像のサイズは、情報量を保持しつつ計算コストやメモリ消費の増加を抑えるため、縦 128 ピクセル × 横 128 ピクセルに統一した。このサイズは、学習時間や計算負荷の軽減を目的として設定したものである。なお、より高い分類精度の可能性を検討するために、MobileNetV2 本来の入力サイズである 224 × 224 ピクセルでも一度学習を行った。結果として、128 × 128 では正答率が 73.00%、224 × 224 では 77.12% と精度の向上が見られた。しかし、計算資源とのバランスや他の条件との一貫性を考慮し、本実験では最終的に 128 × 128 ピクセルを採用することとした。

学習エポック数は 30 に設定した。これは標準的な回数であり、学習の進行状況や検証精度の推移から妥当であると判断した。バッチサイズはデフォルト値である 32 を採用し、学習の安定性とメモリ使用量のバランスに配慮した。最適化手法としては Adam を使用し、学習率はデフォルトの 0.001 とした。

損失関数には多クラス分類に適した categorical_crossentropy を使用し、評価指標には accuracy を設定した。また、過学習抑制のために Dropout[3] (0.3) を中間層に適用している。

これらの設定は、分類精度と計算効率のバランスを考慮して、実験的に調整した結果である。

3 実験結果

本チームは、大きく分けて 4 つの実験を行った。

3.1 実験 1

実験 1 では、転移学習を用いず、入力層から畳み込み層・プーリング層・全結合層までを独自に構築した CNN モデルを使用した。モデルは、Conv2D 層を 3 層重ね、それぞれの畳み込みの後に MaxPooling2D によるダウンサンプリングを行っている。最後に Flatten 層で平坦化し、全結合層 (Dense(128)) と出力層 (クラス数 6 の softmax) で分類を行う構成とした。

すべての重みは初期化された状態から訓練されており、事前学習済みモデルは使用していない。また、評価には分類性能の各指標 (図 1) と混同行列 (図 2) を用いて分析を行った。

3.1.1 実験 1 の実験結果

```

Classification Report:
      precision    recall  f1-score   support

   siba         0.73     0.61     0.66       128
  kisyuu        0.74     0.56     0.64        86
   sikoku       0.43     0.47     0.45       108
 hokkaidou     0.37     0.41     0.39        86
    kai        0.49     0.38     0.42       136
   akita       0.77     0.93     0.84       269

 accuracy              0.63       813
 macro avg           0.59     0.56     0.57       813
 weighted avg        0.63     0.63     0.62       813

混同行列を保存 : confusion_matrix_noweight.png
✓ Accuracy: 63.10% を保存しました。

```

図 1: 実験 1 の CNN モデルの分類性能

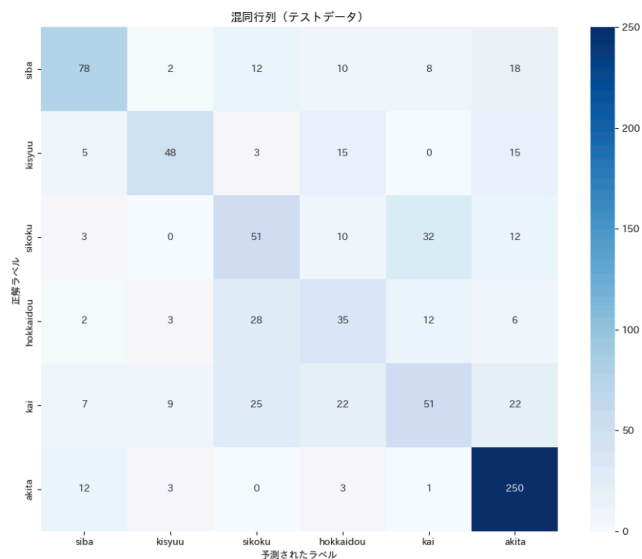


図 2: 実験 1 に対する混同行列

図 1 から実験 1 での全体の正解率は 63% であった。また各ラベルの適合率を見てみると、柴犬、紀州、秋田が 70 %を超えており、四国、甲斐が 40% ほど、北海道犬は 37% と一番低く、ばらつきがあった。図 2 の混同行列からは「甲斐犬と四国犬」、「北海道犬と四国犬」の誤分類が多く見ら

れた。

3.1.2 実験 1 の考察

本モデルにおける全体の正答率は 63.1% であり、一定の分類性能は示されたものの、高精度と
は言い難い結果となった。各クラスの適合率や再現率には顕著なばらつきが見られ、特に「北海道
犬」や「甲斐犬」など一部のクラスでは F1 スコアが 0.4 前後と低く、分類が困難であることが示
唆された。

このような結果から、限られたデータ量でゼロから学習を行う手法には性能の限界があると考え
た。そこで、より高い分類精度を達成するためのアプローチとして、「転移学習」の導入が有効で
はないかと考えた。

3.2 実験 2

実験 2 では、分類精度の向上を目的として、転移学習を取り入れたモデルの構築を行った。ベー
スモデルには、軽量かつ高精度で知られる MobileNetV2 を採用し、ImageNet[2] で事前学習された
重みを活用した。MobileNetV2 の出力に対してグローバル平均プーリング層、全結合層 (Dense)、
および Dropout 層を追加し、最終的に対象の 6 犬種に対応する出力層 (softmax) を接続した構成
とした。

評価には実験 1 と同様、分類性能の各指標 (図 3) と混同行列 (図 4) を用いて分析を行った。

3.2.1 実験 2 の実験結果

	precision	recall	f1-score	support
siba	0.76	0.86	0.81	128
kisyuu	0.87	0.48	0.62	86
sikoku	0.51	0.82	0.63	108
hokkaidou	0.31	0.19	0.23	86
kai	0.82	0.50	0.62	136
akita	0.85	0.99	0.91	269
accuracy			0.73	813
macro avg	0.69	0.64	0.64	813
weighted avg	0.73	0.73	0.71	813

📄 混同行列を保存 : confusion_matrix_noweight.png
✅ Accuracy: 72.57% を保存しました。

図 3: 実験 2 の分類性能

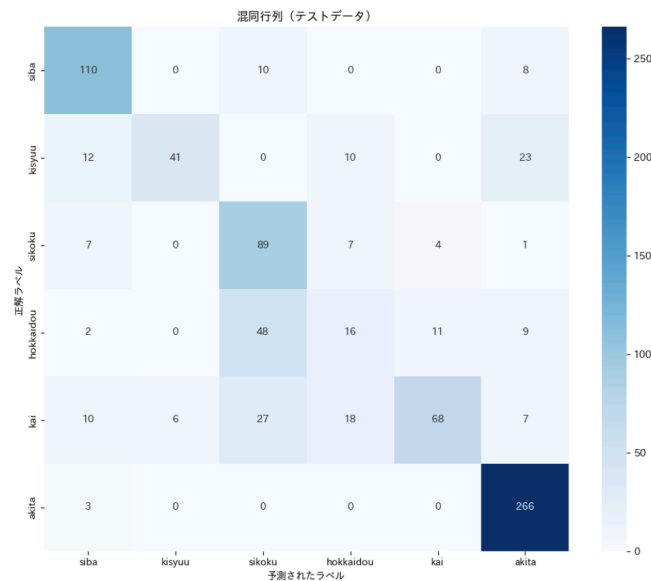


図 4: 実験 2 に対する混同行列

図 3 から全体の正解率は 72% であった。また各ラベルの適合率を見てみると、紀州、甲斐、秋田犬は 80% を超えており、柴犬は 76%、四国犬が 51%、北海道犬が 31 % となった。図 4 の混同行列から、実験 1 と同様「北海道と四国犬」の誤分類が目立っていた。

3.2.2 実験 2 の考察

実験 1 の考察どおり、MobileNetV2 を用いた転移学習により、正答率は 72% と実験 1 と比較すると大幅に向上した。これは、ImageNet[2] で事前学習されたモデルが画像の特徴を効果的に捉える能力を有しており、比較的少量のデータでも高い分類性能を発揮できたためと考えられる。特に、日本犬のように外見上の違いが微妙である分類問題において、転移学習の効果が顕著に現れたことが確認できた。

一方で、各ラベル（犬種）ごとの適合率にはばらつきが見られた。この要因の一つとして、使用した訓練データの偏りが挙げられる。本研究では、1 人 1 種の犬を担当し、元画像を個別に収集したうえでデータ拡張を行っており、その結果、犬種ごとの画像枚数に差が生じていた。そのため、このようなデータセット内の不均衡による分類性能のばらつきにつながったと考えた。

したがって、精度向上には、各犬種の画像枚数を均等に増やし、データの多様性を確保することが有効であると考えられる。

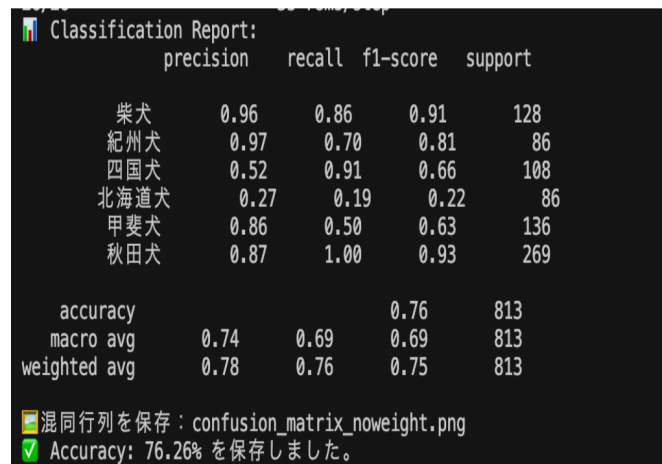
3.3 実験 3

実験 3 では、訓練用データの犬種ごとの枚数を揃えることで、この影響を抑えた上での精度改善を検証することとした。具体的には、実験 1 および実験 2 で使用した訓練用データセットに対して、以下の手順でデータ拡張による補完処理を実施した。

1. 各犬種について、訓練用画像の枚数を調査し、最も枚数が多かった犬種に合わせて他の犬種の画像を増やす。
2. 拡張処理は以下のルールで行った
 - 既存の画像に対し、左右反転を行うかどうかをランダムに決定
 - さらに、 0° を除く $0\sim 15^\circ$ の範囲でランダムな角度の回転を適用
3. この処理を、目的の枚数に到達するまで手順 2 を繰り返す

また、実験 3 では、実験 2 と同様のモデル構成およびコードを用いて学習・評価を行った。

3.3.1 実験 3 の実験結果



```

Classification Report:
precision    recall  f1-score   support

   柴犬      0.96    0.86    0.91     128
   紀州犬      0.97    0.70    0.81      86
   四国犬      0.52    0.91    0.66     108
  北海道犬      0.27    0.19    0.22      86
   甲斐犬      0.86    0.50    0.63     136
   秋田犬      0.87    1.00    0.93     269

 accuracy          0.76     813
  macro avg      0.74    0.69    0.69     813
 weighted avg    0.78    0.76    0.75     813

混同行列を保存：confusion_matrix_noweight.png
✓ Accuracy: 76.26% を保存しました。
  
```

図 5: 実験 3 の分類性能

表 1: 実験 2, 3 の適合率の比較

	実験 2	実験 3
柴犬	0.76	0.96
紀州犬	0.87	0.97
四国犬	0.51	0.52
北海道犬	0.31	0.27
甲斐犬	0.82	0.86
秋田犬	0.85	0.87

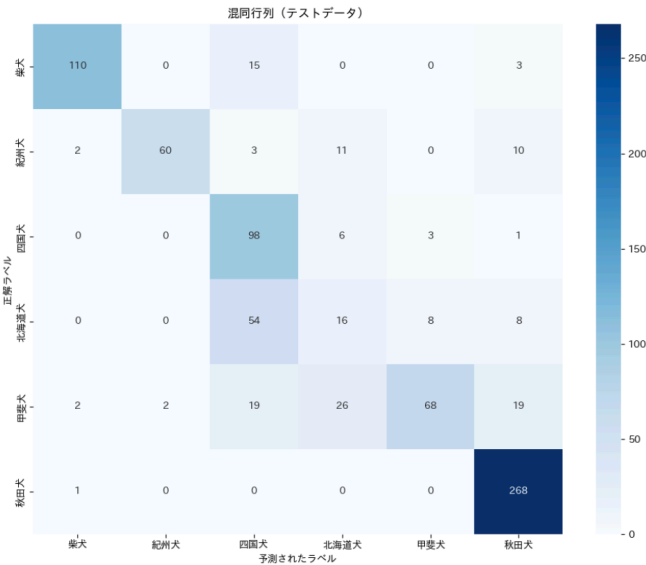


図 6: 実験 3 に対する混同行列

図 5 から全体の正解率は 76% で精度の向上が確認できた。また、各犬種ごとの分類性能の変化をより明確に把握するために、実験 2 および実験 3 における「適合率」を表にまとめた (表 1)。

3.3.2 実験 3 の考察

表 1 から多くの犬種で適合率の改善が見られたことがわかる。特に、柴犬・紀州犬・甲斐犬などでは順当に精度が向上しており、データ枚数の調整が分類性能の安定化に寄与していると考えられる。一方で、北海道犬のみは例外的に適合率が大きく低下しており、実験 2 では 0.31 とやや低めであったものの、実験 3 では 0.27 とさらに下がる結果となった。これは、他犬種との外見上の類似性や、学習データに含まれる画像の質・多様性の偏りなどが要因として考えられる。今後は、北海道犬に対して特化したデータ拡張や追加収集を行うことで、さらなる改善が期待されると考えた。

3.4 実験 4

実験 3 までの結果から、北海道犬の分類精度が他の犬種と比較して著しく低い傾向が見られた。これは、北海道犬が持つ毛色や模様のバリエーションが多いことに加え、インターネット上で取得可能な画像枚数が限られており、十分な訓練データを確保できなかったことが一因と考えられる。

しかし、データセットの再収集は、画像数の少なさと時間的制約から現実的ではないと判断した。そこで本実験では、分類の対象である犬本体にモデルの着目を集中させることを目的として、訓練用画像から背景を削除する処理を施すことにした。

背景除去には、Python ライブラリ「rembg」を用い、すべての訓練画像に対して前処理を行った。また、実験 4 では、実験 3 と同様のモデル構成およびコードを用いて学習・評価を行った。

3.4.1 実験 4 の実験結果

```

Classification Report:
precision    recall  f1-score   support

   柴犬      0.69    0.88    0.78     349
   紀州犬     0.95    0.55    0.69     349
   四国犬     0.46    0.86    0.60     349
  北海道犬    0.39    0.07    0.12     349
   甲斐犬     0.68    0.53    0.60     349
   秋田犬     0.63    0.82    0.71     349

 accuracy          0.62    2094
  macro avg      0.63    0.62    0.58    2094
 weighted avg    0.63    0.62    0.58    2094

📁 混同行列を保存 : confusion_matrix_noweight.png
✅ Accuracy: 61.89% を保存しました。

```

図 7: 実験 4 の分類性能

表 2: 実験 3, 4 の適合率の比較

	実験 3	実験 4
柴犬	0.96	0.69
紀州犬	0.97	0.95
四国犬	0.52	0.46
北海道犬	0.27	0.39
甲斐犬	0.86	0.68
秋田犬	0.87	0.63

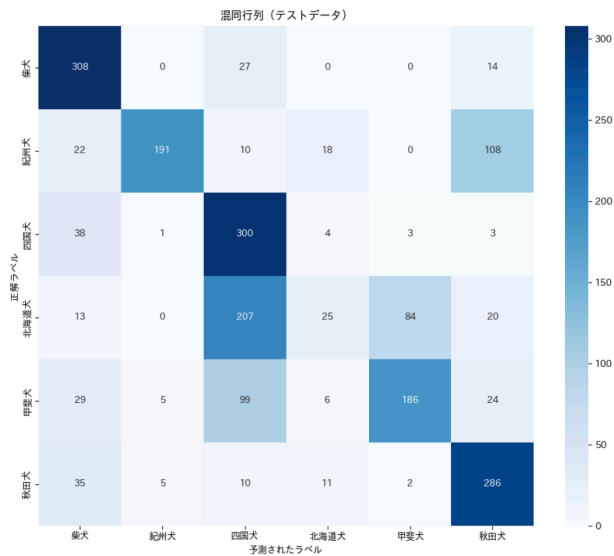


図 8: 実験 4 に対する混同行列

図 5 から全体の正解率は 61% と実験 3 より大幅に精度が下がる結果となってしまった。実験 3 同様、各犬種ごとの分類性能の変化をより明確に把握するために、実験 3 および実験 4 における「適合率」を表にまとめた（表 2）。

3.4.2 実験 4 の考察

背景除去を行った訓練用データを用いてモデルを再学習させた結果、全体の正答率は 61% と、実験 3 の 76% に比べて精度が低下する結果となった。加えて、表 2 から読み取れるように各ラベルにおける適合率も全体的に大幅に低下しており、北海道犬のみわずかに適合率が向上したが、統計的な誤差の範囲であると考えられる。

さらに、図 8 に示す混同行列の分析から、「紀州犬と秋田犬」、「北海道犬と四国犬」、「北海道犬と甲斐犬」、「甲斐犬と四国犬」の間での誤分類が多発していることが確認された。これまでの実験では顕著に見られなかった傾向であり、背景除去によってモデルが参照する画像の特徴量に大きな変化が生じた可能性がある。

この結果は、背景を取り除くことで本来期待していた「対象の犬に着目した学習」がうまく機能しなかったことを示唆している。むしろ、背景情報が補助的に機能していた可能性が考えられる。

3.5 失敗分析

実験 3 および実験 4 のモデルに対し Grad-CAM[4] による可視化を行い、実際にモデルが画像のどの部分に着目して分類を行っているのかを比較・分析することを試みた。

3.5.1 Grad-CAM の考察

Grad-CAM を用いて、モデルが分類時に注目している画像の領域を可視化した。図 9,10 のような「正しく分類された画像（成功例）」では、主に犬の顔や体全体に強く反応しているケースが多く見られ、モデルが対象物に適切に着目していることが確認できた。

一方で、図 11 14 に示す「誤分類された画像（失敗例）」では、犬本体にはほとんど注目せず、背景やその他の無関係な領域に強い反応が見られる場合が多かった。これは、モデルが犬本体ではなく背景の情報に依存して判断を行っていたことを示唆しており、そのことが誤分類の一因となったと考えられる。

また、犬本体に正しく注目していたにもかかわらず誤分類された事例も一部存在した。これらのケースについては、画像の画質のばらつき、構図の偏り、あるいは対象となる犬種同士の類似性など、データセットの質や内容に起因する要因が影響している可能性が高い。

これらの結果から、分類精度のさらなる向上には、画像の収集・選別段階における質の担保や、背景の取り扱いを含めた画像全体の構成に対する工夫が重要であることが示唆された。

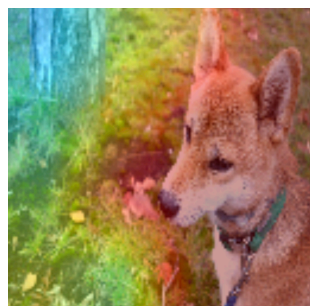


図 9: 実験 3 の成功事例 1

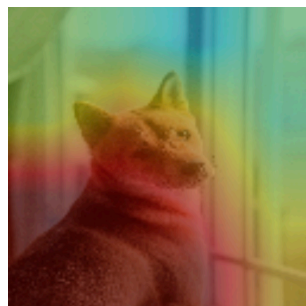


図 10: 実験 3 の成功事例 2

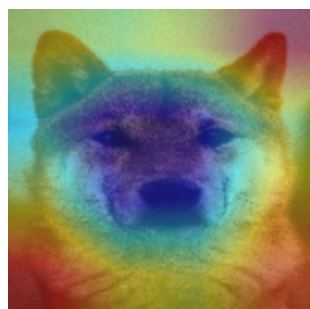


図 11: 実験 3 の失敗事例 1



図 12: 実験 3 の失敗事例 2



図 13: 実験 4 の失敗事例 1



図 14: 実験 4 の失敗事例 2

4 全体の考察

本プロジェクトでは、複数の手法や条件を比較しながら、日本犬の分類精度向上を目指して取り組んだ。転移学習による大幅な精度改善が得られた一方で、データの枚数調整や背景除去といった工夫では、必ずしも期待した効果が得られず、モデルが背景などの副次的情報にも依存していた可能性が示された。Grad-CAM による可視化を通じて、成功時には犬の本体に注目していた一方、失敗時には背景に注意が向いている例が多いことがわかり、誤分類の要因分析に有効であることを確認できた。また、正しく注目していても誤判定される例があり、データの画質・構図・類似犬種の難しさといった、データセットの質的課題も明らかとなった。今後は、分類性能のさらなる向上に向けて、単なるモデル改良だけでなく、データの質・構造・収集方針の見直しや、可視化分析を活用した失敗事例の改善が重要であると考えられる。

5 まとめ

本研究では、日本犬 6 種の分類モデルを構築し、転移学習・データ拡張・背景処理などの手法を比較した。その結果、転移学習を用いることで正答率が大きく向上することが確認された。また、犬種ごとに識別の難易度に差があることが分かり、データの質や視覚的特徴のばらつきが分類精度に大きく影響することが明らかになった。ResNet50 との比較や Grad-CAM による可視化といった補助的な分析を通じて、分類性能の背景にある要因を多角的に検討することができた。一方で、データの枚数を揃えたり背景を統一したりしても、精度の改善にはあまりつながらなかった。また、モデルが犬本体ではなく背景に注目して誤分類するケースも確認された。今後さらに精度を高めるためには、データの質の見直しやアノテーションの正確性向上、そして識別が難しい犬種に対する対策が重要であると考えられる。

6 今後の課題

- データセットの信頼性が低かったため、撮影条件の統一、専門家の監修、アノテーション基準の明確化により、正確で一貫したデータ収集を行う。
- 視覚的に似た犬種の識別が難しかったため、類似性を意識した特徴量設計や、判別性能を高める損失関数の導入を検討する。
- モデルが背景情報に依存し誤分類を引き起こしていたため、画像中で注目すべき領域（例：犬の顔や体）に焦点を当てるよう、前処理や学習データの工夫、もしくは注意機構を導入することで、判断根拠を誘導する手法を検討する。
- バッチサイズや学習率などのハイパーパラメータについては、十分な検証時間が確保できず、限られた設定での学習に留まった。今後は複数の設定を比較検討し、最適な組み合わせを探索することで、さらなる性能向上を図る。
- 評価や分析に十分な時間が取れなかったため、クロスバリデーションや別データセットを用いた汎化性能の客観的な評価を行う。
- モデルの比較が限定的だったため、EfficientNet などの軽量モデルやアンサンブル学習を取り入れ、より高精度な構成を検討する。

参考文献

- [1] Sandler, M., Howard, A., Zhu, M., Zhmoginov, A., & Chen, L. C. (2018). MobileNetV2: Inverted Residuals and Linear Bottlenecks. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR).
<https://arxiv.org/abs/1801.04381>
- [2] Deng, J., Dong, W., Socher, R., Li, L. J., Li, K., & Fei-Fei, L. (2009). ImageNet: A large-scale hierarchical image database. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR).
<https://doi.org/10.1109/CVPR.2009.5206848>
- [3] Srivastava, N., Hinton, G., Krizhevsky, A., Sutskever, I., & Salakhutdinov, R. (2014). Dropout: A simple way to prevent neural networks from overfitting. Journal of Machine Learning Research, 15(1), 1929 - 1958.
<http://jmlr.org/papers/v15/srivastava14a.html>
- [4] Selvaraju, R. R., Cogswell, M., Das, A., Vedantam, R., Parikh, D., & Batra, D. (2017). Grad-CAM: Visual Explanations from Deep Networks via Gradient-Based Localization.

Proceedings of the IEEE International Conference on Computer Vision (ICCV).

<https://arxiv.org/abs/1610.02391>

- [5] ” 一般社団法人 ジャパンケネルクラブ” , 2025/05/15 閲覧

<https://www.jkc.or.jp/>

- [6] ” 毒のある蛇かそうでないかを画像判別” , 2025/05/15 閲覧

<https://drive.google.com/drive/u/1/folders/1BxD-jnUHMfFm6gSKO1b MIcH- R7NdkG>