

卒業論文概要書

Summary of Bachelor's Thesis

Date of submission: 02/07/2020

学科名 Department	情報通信	氏名 Name	沈 竹笑	指 導 教 員 Advisor	渡辺 裕 ㊞
研究指導名 Research guidance	オーディオビジュアル 情報処理研究	学籍番号 Student ID number	1w153097-3 ^{CD}		
研究題目 Title	オブジェクト細分化を適用した動物スタイル変換 Object subdivision applying animal style transformation				

1. まえがき

ディープラーニングの技術は日々進化している。新たな研究が発表されると、すぐに実装コードが公開されたり、応用研究が進められたり、ビジネスに適用されたりする。2014年から敵対的生成ネットワークである Generative Adversarial Networks (GAN)が注目されている。GANは生成モデルの一種であり、データから特徴を学習することで、実在しないデータを生成したり、存在するデータの特徴に沿って変換できる[1]。Cycle-Consistent Adversarial Networks(CycleGAN)は、GANでスタイル変換を行う手法のひとつである。スタイル変換は、データの外見的特徴の変換を行うことである。

しかし、CycleGANで画像を変換するとき、全胴体のスタイルが変換されてしまい、重なった物胴体の変換がうまくいかない場合がある。重なった物胴体をうまく変換するためには、物胴体の複数のクラスの集合として認識する必要がある。

2. Creative Adversarial Networks

Generative Adversarial Networks (GAN)は、GeneratorとDiscriminatorの二つのネットワークから構成される生成モデルの一種であり、データから特徴を学習することで、実際に存在しないデータを生成し、存在するデータの特徴に沿って変換できる。

3. Cycle-Consistent Adversarial Networks

Cycle-Consistent Adversarial Networks (CycleGAN)とはGANを複数活用した機械学習の一種で、入力データに任意の特徴を付加したデータを出力できるというものである。

4. Gradient-weighted Class Activation Mapping

Gradient-weighted Class Activation Mapping (GradCAM)は、CNNが分類のために注視している

範囲をカラーマップで表示する。GradCAMを用いればCNNで何を根拠にその識別をしたのかなどの情報を目に見える形で表現することが可能である[2]。

5. 提案手法

本研究では、各動物は2クラスのオブジェクト、頭と胴体で構成される前提とする。VGG16を使って8クラスの分類をする。具体的には馬の胴体、馬の頭、シマウマの胴体、シマウマの頭、猫の頭、猫の胴体、犬の頭、犬の胴体の8クラスである。学習したCNNモデルを用いることで、入力画像のCAMマップを得る。入力画像をVGG16に通したとき、予測確率の上位2クラスを用いる。この予測クラスに基づいて、CycleGANを用いて入力画像を変換する。たまに、入力画像をVGG16に通したときに得られるCAMマップをマスク画像に変換する。最後にCycleGANによる変換画像に対してマスク処理をする。以上は、一つのクラスに対する画像変換の流れである。この処理を頭と胴体の2クラスに適用し、元の画像に対して二枚の変換画像を重ね合わせることで、最終的な合成画像を得る。

生成された合成画像を学習したCNNに入力し、馬、シマウマ、猫および犬を正しく認識できるかどうかを評価する。合成したそれぞれ種類の画像とCNNで正しく認識された画像の割合を求めて、正解率を算出する。

6. 実験

4種類動物の胴体と頭画像は、イメージネット上の合計1344枚のデータセットを用いて、CNNを学習した。学習したCNNモデルを用いることで、入力画像のCAMマップを得る。結果例として、馬の入力画像を図1に示す、馬の頭のCAMマップを図2に示す、馬の胴体のCAMマップを図3に示す。また、VGG16で8クラスを分類した結果を表1に示す。

CycleGANを用いて、二枚の画像変換を重ね合わ

せた変換結果例を図 4, 図 5 に示す.

最後, CycleGAN によって生成されたそれぞれの画像を CNN によって評価し, 正解率を算出した. 正解率を表 2 に示す.



図 1 馬の入力画像



図 2 馬の頭の CAM マップ 図 3 馬の胴体の CAM マップ



図 4 馬からシマウマの変換 図 5 犬から猫の変換

表 1 各クラスの精度

	馬の胴体	馬の頭	シマウマの胴体	シマウマの頭
正しい予測	50	43	30	31
合計	50	47	29	32
精度[%]	100	91.5	96.7	96.9
	猫の胴体	猫の頭	犬の胴体	犬の頭
正しい予測	32	69	17	50
合計	43	70	24	52
精度[%]	74.4	98.6	70.8	96.2

表 2 CNN による正解率

	正解率[%]
馬	77.7
シマウマ	78.5
犬	40.2
猫	46.8

7. 考察

VGG16 で 8 クラスを分類した結果から見ると, 馬の胴体と頭, シマウマの胴体と頭, 犬と猫の頭のクラス精度が高くて, 90%以上になったが, 今回実験で用いた犬と猫のデータは同じ白い品種で, 犬と猫の胴体のクラス精度が低い. 最後, CycleGAN の変換, 重ね合わせた結果から比較すると, 馬からシマウマへの変更結果がより良い結果になった. 犬から猫への変更は, 胴体の部分が全然変換されない.

一方, CNN の評価結果から比較すると, 馬とシマウマの場合は 70%ぐらい, 犬と猫は 40%ぐらいである. 全体的には馬とシマウマの変換が理想的となる.

8. まとめ

CNN による評価結果と最後の合成結果から見ると, 馬とシマウマの変換は成功したといえる. 猫と犬は品種の違いによって, 結果が理想ではない.

本実験の結果は, CAM の領域分割精度と CycleGAN の変換結果によるとなる. 両方の結果を改善する方法は今後の課題となる. また, 本研究では, 動物だけを研究対象とする. 今後, もっと一般的なものを対象として, 適切なオブジェクト細分化を検討すべきである方法と考えられる.

参考文献

- [1] AL SCHOLAR, "実在しないアイドル画像も瞬時に生成する AI 「GAN」が賢すぎる!", <<https://gendai.ismedia.jp/articles/-/63615>>(参照 2020 年 1 月 18 日)
- [2] Ramprasaath R. Selvaraju · Michael Cogswell · Abhishek Das · Ramakrishna Vedantam · Devi Parikh · Dhruv Batra, Grad-CAM: Visual Explanations from Deep Networks via Gradient-based Localization, arXiv:1610.02391, 2016.

2019 年度 卒業論文

オブジェクト細分化を適用した動物スタイル変換

Object subdivision applying animal style transformation

指導教員 渡辺 裕 教授

早稲田大学 基幹理工学部

情報通信学科

1W153097-3

沈 竹笑

目次

第1章	序論.....	1
1.1	研究の背景.....	1
1.2	本研究の目的.....	1
1.3	本論文の構成.....	1
第2章	CNN.....	2
2.1	まえがき.....	2
2.2	ニューラルネットワーク.....	2
2.3	Convolutional Neural Network.....	3
2.4	VGG16.....	3
2.5	むすび.....	4
第3章	GAN.....	5
3.1	まえがき.....	5
3.2	Generative Adversarial Networks.....	5
3.3	Cycle-Consistent Adversarial Networks.....	5
3.4	Class Activation Mapping.....	6
3.5	Gradient-weighted Class Activation Mapping.....	7
3.6	むすび.....	7
第4章	提案手法.....	8
4.1	まえがき.....	8
4.2	提案手法.....	8
4.2.1	分類.....	8
4.2.2	画像変換.....	8
4.2.3	評価.....	8
4.3	むすび.....	11
第5章	実験結果.....	12
5.1	まえがき.....	12
5.2	データセット.....	12
5.3	VGG16の分類結果.....	13
5.4	GradCAMの実行結果.....	13
5.5	CycleGANの実行結果.....	15
5.6	CNNによる評価.....	15
5.7	考察.....	166
5.8	むすび.....	166

第6章 結論と今後の課題.....	17
6.1 結論.....	17
6.2 今後の課題.....	17
謝辞.....	18
参考文献.....	19
図一覧.....	20
表一覧.....	211

第1章 序論

1.1 研究の背景

ディープラーニングの技術は日々進化している。新たな研究が発表されると、すぐに実装コードが公開されたり、応用研究が進められたり、ビジネスに適用されたりする。2014年から敵対的生成ネットワークである Generative Adversarial Networks (GAN)が注目されている。GANは生成モデルの一種であり、データから特徴を学習することで、実在しないデータを生成し、存在するデータの特徴に沿って変換できる[1]。Cycle-Consistent Adversarial Networks(CycleGAN)は、GANでスタイル変換を行う手法のひとつである。スタイル変換は、データの外見的特徴の変換を行うことである。

しかし、CycleGANで画像を変換するとき、全胴体のスタイルが変換されてしまい、重なった物胴体の変換がうまくいかない場合がある。重なった物胴体をうまく変換するためには、物胴体の複数のクラスの集合として認識する必要がある。

1.2 本研究の目的

本研究では、重なった物胴体の画像変換の問題を解決することを目的とする。そこで、一つの物胴体を複数のクラスの集合として扱い、画像変換する手法を提案する。なお、本研究では動物画像を研究対象とする。従来、動物画像全胴体に CycleGAN を適用していたことに対して、本研究では動物を身胴体と頭の 2 クラスに分割して、それぞれに対して CycleGAN を適用する方法を提案する。

1.3 本論文の構成

本論文の構成を以下に示す。

第1章は本研究の背景と目的について述べる。

第2章はニューラルネットワークと Convolutional Neural Network(CNN)の基本概念、構造について述べる。また、実験で用いたモデル VGG16 について述べる。

第3章では Generative Adversarial Networks (GAN) と Class Activation Mapping(CAM)の基本概念について述べる。また本研究で用いた CycleGAN と Gradient-weighted Class Activation Mapping(GradCAM)について述べる。

第4章では CNN と GradCAM を用いて CycleGAN で重なった動物の画像の領域を分割し、画像変換する方法を提案する。

第5章は VGG16 で作った 8 クラスの精度とそれぞれのクラスに対する CAM の結果と CycleGAN の変換結果とその CNN による評価について述べる。

第6章は本研究の結論と今後の課題について述べる。

第2章 CNN

2.1 まえがき

本章では，基本概念であるニューラルネットワークと Convolutional neural network(CNN)について述べる．また，VGG16 について述べる．

2.2 ニューラルネットワーク

ニューラルネットワークとは，シナプスの結合によりネットワークを形成した人工ニューロン（ノード）が，学習によってシナプスの結合強度を変化させ，問題解決能力を持つようなモデル全般を指す[2]．最もシンプルなニューラルネットワークには，入力層，出力層，および，その間の中間層が，それぞれ一つずつある．各層は，複数のノードがエッジで結ばれる構造となっている．この中間層は複数の層を持つことができ，特に深い中間層を持つものをディープニューラルネットワークと呼ぶ．各層は活性化関数と呼ばれる関数を持ち，エッジは重みを持つことができる．そして各ノードの値は，そのノードと接続する前の層のノードの値から計算する．前の層のノードの値，接続エッジの重みの値，そして層が持つ活性化関数から計算する．ニューラルネットワークの構造例を図 2.1 に示す．

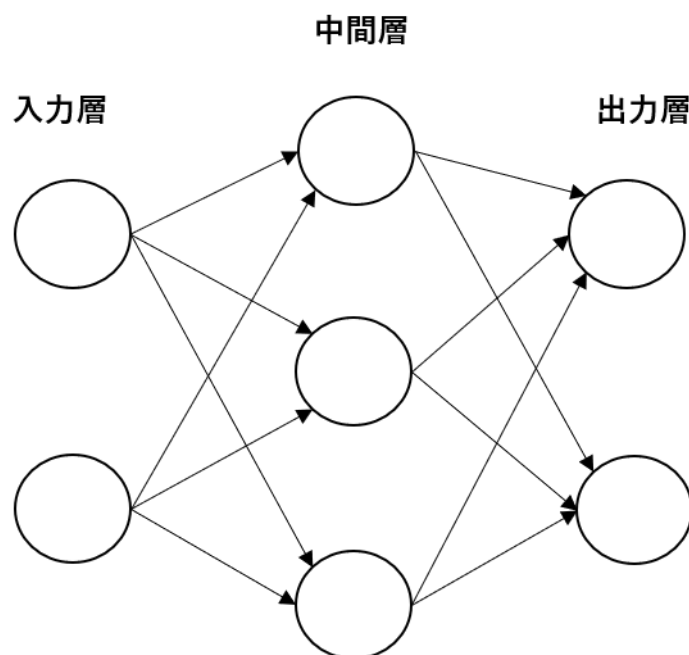


図 2.1 ニューラルネットワークの構造例

2.3 Convolutional Neural Network

畳み込みニューラルネットワーク Convolutional neural network(CNN)は、何段もの深い層を持つニューラルネットワークで、特に画像認識の分野で優れた性能を発揮しているネットワークである。畳み込みとは、画像から特徴を抽出する操作のことである。画像処理で CNN の役割は、予測に必要な重要な特徴を失うことなく、画像のイメージを処理しやすい形式に変換することである。CNN は入力層、畳み込み層(Convolutional layer)、プーリング層(Pooling layer)、全結合層(Fully Connected layer)と出力層(Output layer)から構成され、畳み込み層とプーリング層を複数繰り返して、最後に、全結合層に続く仕組みである。CNN の構造例は図 2.2 に示す。

畳み込み層は、元の画像にフィルタをかけて特徴マップを出力する。プーリング層は、物胴体の位置が変動しても同一の物胴体であるとみなす。これらの層を組み合わせることによって、画像から特徴マップを抽出する働きを担っている。全結合層は、畳み込み層とプーリング層を通して、特徴マップが取り出された画像データを一つのノードに結合し、活性化関数によって変換された値を出力する。出力層は全結合層からの出力である特徴マップを元に、ソフトマックス関数を用いて確率に変換し、それぞれのクラスに正しく分類される確率を最大化することによって分類する[4]。

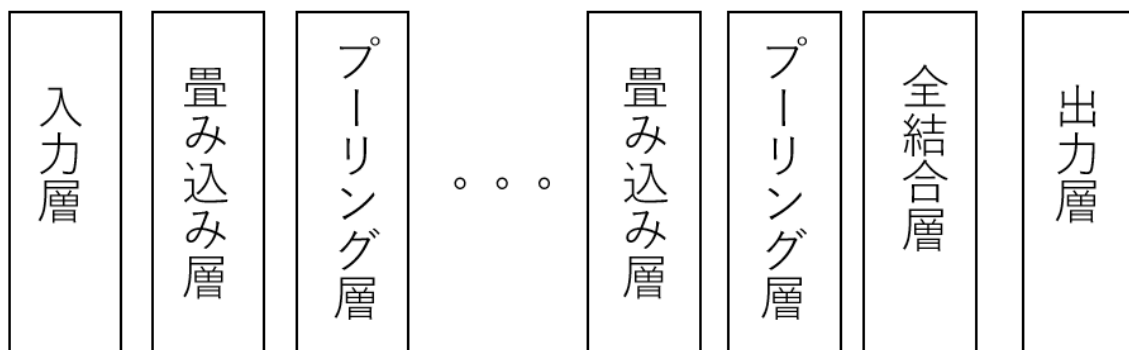


図 2.2 CNN の構造例

2.4 VGG16

VGG16 は、ImageNet データベースの 100 万枚を超える画像で学習済みの畳み込みニューラルネットワークである。このネットワークは、深さが 16 層であり、画像を 1000 個のクラスに分類できる[5]。VGG16 のネットワーク構造例は図 2.3 に示す。

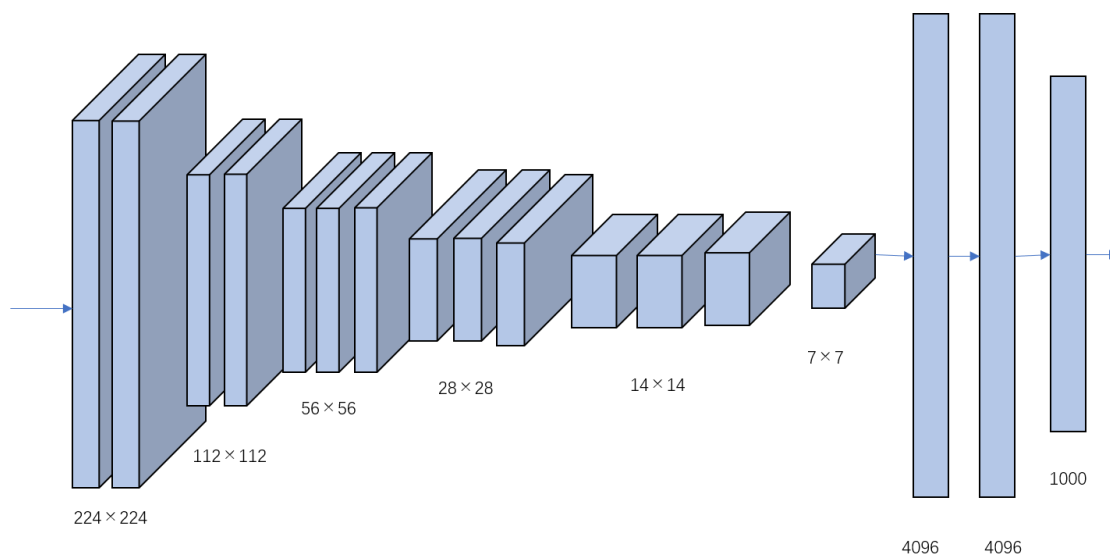


図 2.3 VGG16 のネットワーク構造例

2.5 むすび

本章では、本研究の基本概念であるニューラルネットワーク，CNN，と VGG16 について述べた。

第3章 GAN

3.1 まえがき

本章では，GAN について述べる．また本研究で用いた CycleGAN と GradCAM について述べる．

3.2 Generative Adversarial Networks

Generative Adversarial Networks (GAN)は，Generator と Discriminator の二つのネットワークから構成される生成モデルの一種であり，データから特徴を学習することで，実際に存在しないデータを生成し，存在するデータの特徴に沿って変換できる．GAN の基本構造は図 3.1 に示す．

まず Generator は，潜在変数を入力値として受け取り，画像データを出力する．通常，潜在変数は 100 次元ほどの各要素が 0 から 1 までの値をとる変数で，一様分布や正規分布からランダムサンプリングさる．

Discriminator は画像データを入力値として取り，そのデータが本物のデータか，それとも Generator から生成されたデータかを出力値として識別する．出力値は，本物である場合は 1，偽物である場合は 0 として，その確率を連続値として返す[6]．

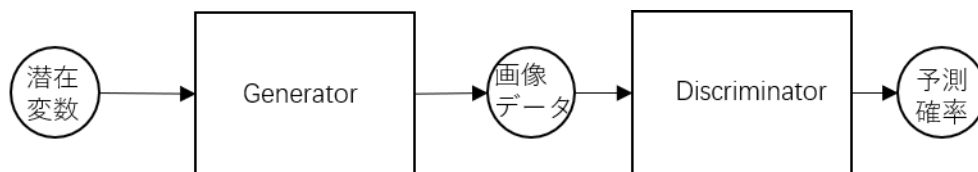


図 3.1 GAN の基本構造

3.3 Cycle-Consistent Adversarial Networks

Cycle-Consistent Adversarial Networks(CycleGAN)とは GAN を複数活用した機械学習の一種で，入力データに任意の特徴を付加したデータを出力できるというものである．CycleGAN の構造は図 3.2 に示す．

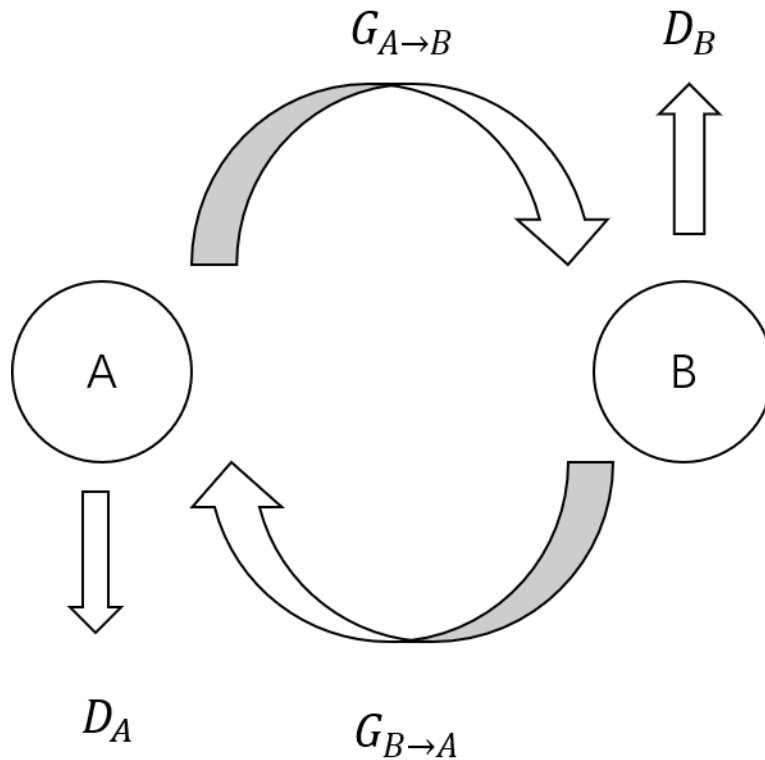


図 3.2 CycleGAN の構造

CycleGAN はドメイン A の画像集合とドメイン B の画像集合を相互に変換する。 $G_{A \rightarrow B}$ は A の画像から B に変換する Generator で、 $G_{B \rightarrow A}$ は B の画像から A に変換する Generator である。 D_A はオリジナルの A と $G_{B \rightarrow A}$ によって生成される偽物の A を見分ける Discriminator である。 D_B はオリジナルの B と $G_{A \rightarrow B}$ によって生成される偽物の B を見分ける Discriminator である。 $G_{A \rightarrow B}$ と $G_{B \rightarrow A}$ は、 D_A と D_B が本物のデータと認識するようなデータを生成しようとする。一方、 D_A と D_B は、 $G_{A \rightarrow B}$ と $G_{B \rightarrow A}$ が作成した偽物のデータと本物のデータを区別できるように学習する[7].

3.4 Class Activation Mapping

Class Activation Mapping(CAM)という手法は、分類に用いられたモデルアーキテクチャに手を加えることで、入力画像に対する予測の可視化を行っているものである。

3.5 Gradient-weighted Class Activation Mapping.

Gradient-weighted Class Activation Mapping(GradCAM)は, CNN が分類のために注視している範囲をカラーマップで表示する. GradCAM を用いれば CNN で何を根拠にその識別をしたのかなどの情報を目に見える形で表現することが可能である[8].

CAM と比べて, ネットワークを変更せずに可視化ができることが特徴である. GradCAM の例を図 3.3 に示す. 学習済みモデルが馬と予測するとき, 反応したピクセルほど赤く示し, 反応しなかったピクセルを青く示している.



図 3.3 GradCAM の例

3.6 むすび

本章では GAN の構造と CAM の概念について述べた.

第4章 提案手法

4.1 まえがき

本章では、CNN と GradCAM を用いて、各物胴体の各クラスの領域をカラーマップで得ることで、それぞれの領域に対して CycleGAN を用いて画像変換する方法を提案する。

4.2 提案手法

4.2.1 分類

本研究では、VGG16 を使って 8 クラスの分類をする。具胴体的には馬の胴体、馬の頭、シマウマの胴体、シマウマの頭、猫の頭、猫の胴体、犬の頭、犬の胴体の 8 クラスである。学習した CNN モデルを用いることで、入力画像の CAM マップを得ることができる。本研究では、各動物は 2 クラスのオブジェクトで構成されるものとする。2 クラスのオブジェクトとは、頭と胴体である。そこで、入力画像を VGG16 に通したとき、予測確率の上位 2 クラスを用いる。

4.2.2 画像変換

まず、VGG16 の予測クラスに基づいて、CycleGAN を用いて入力画像を変換する。次に、入力画像を VGG16 に通したときに得られる CAM マップを、マスク画像に変換する。最後に CycleGAN による変換画像に対してマスク処理をする。以上は、一つのクラスに対する画像変換の流れである。この処理を頭と胴体の 2 クラスに適用する。元の画像に対して二枚の変換画像を重ね合わせることで、最終的な合成画像を得る。全胴体の流れは図 4.1 に示す。画像変換の流れは図 4.2 に示す。マスクの説明は図 4.3 に示す。

4.2.3 評価

生成された合成画像を学習した CNN に入力し、馬、シマウマ、猫および犬を正しく認識できるかどうかを評価する。合成したそれぞれ種類の画像と CNN で正しく認識された画像の割合を求めて、正解率を算出する。

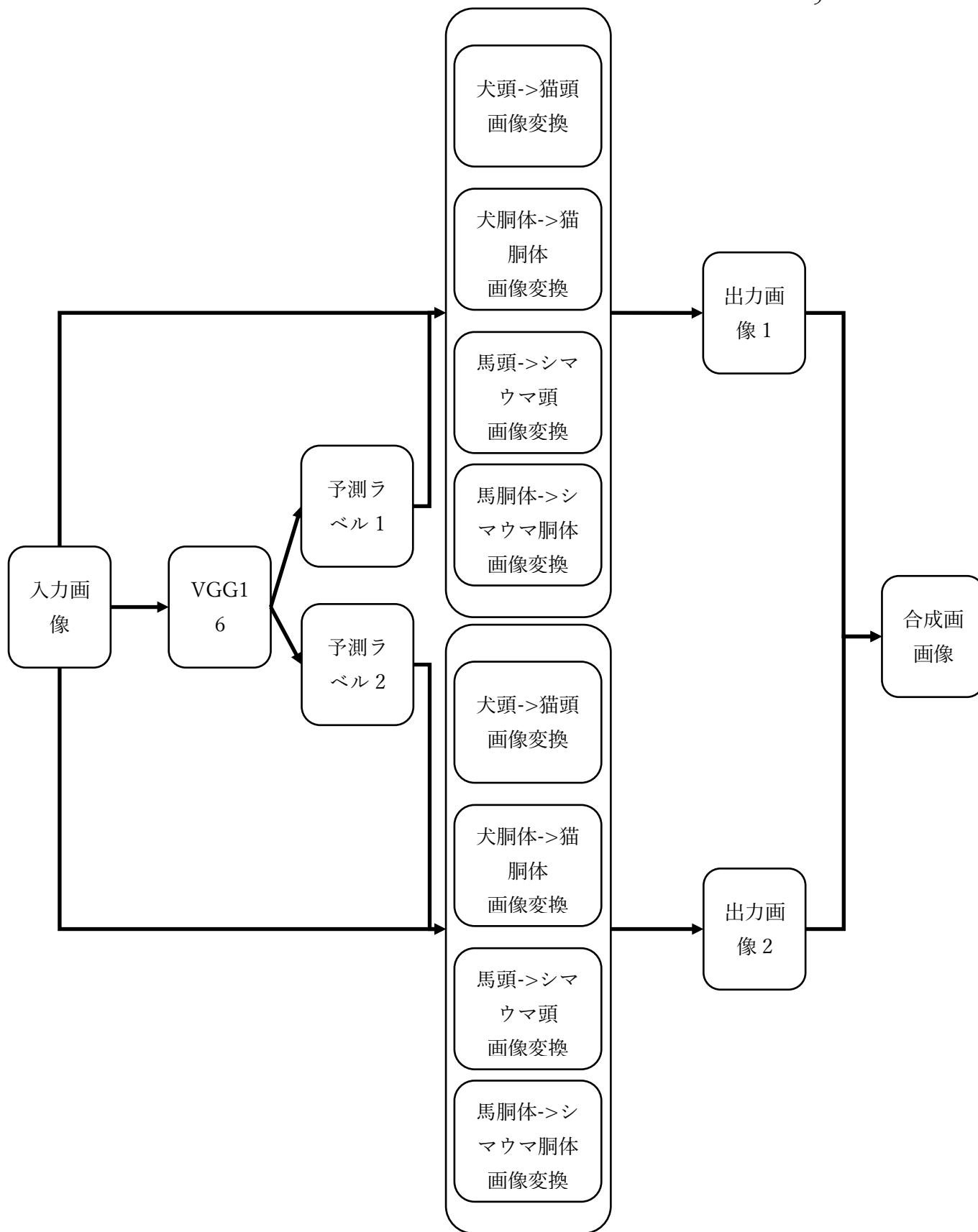


図 4.1 全胴体の流れ

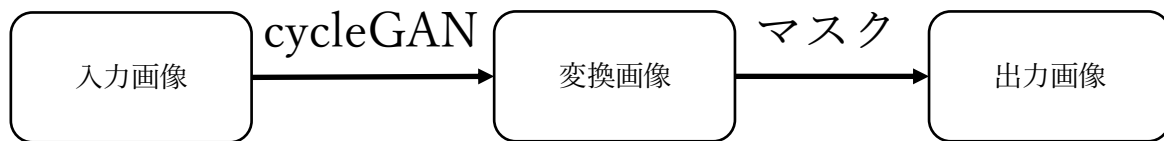


図 4.2 画像変換の流れ

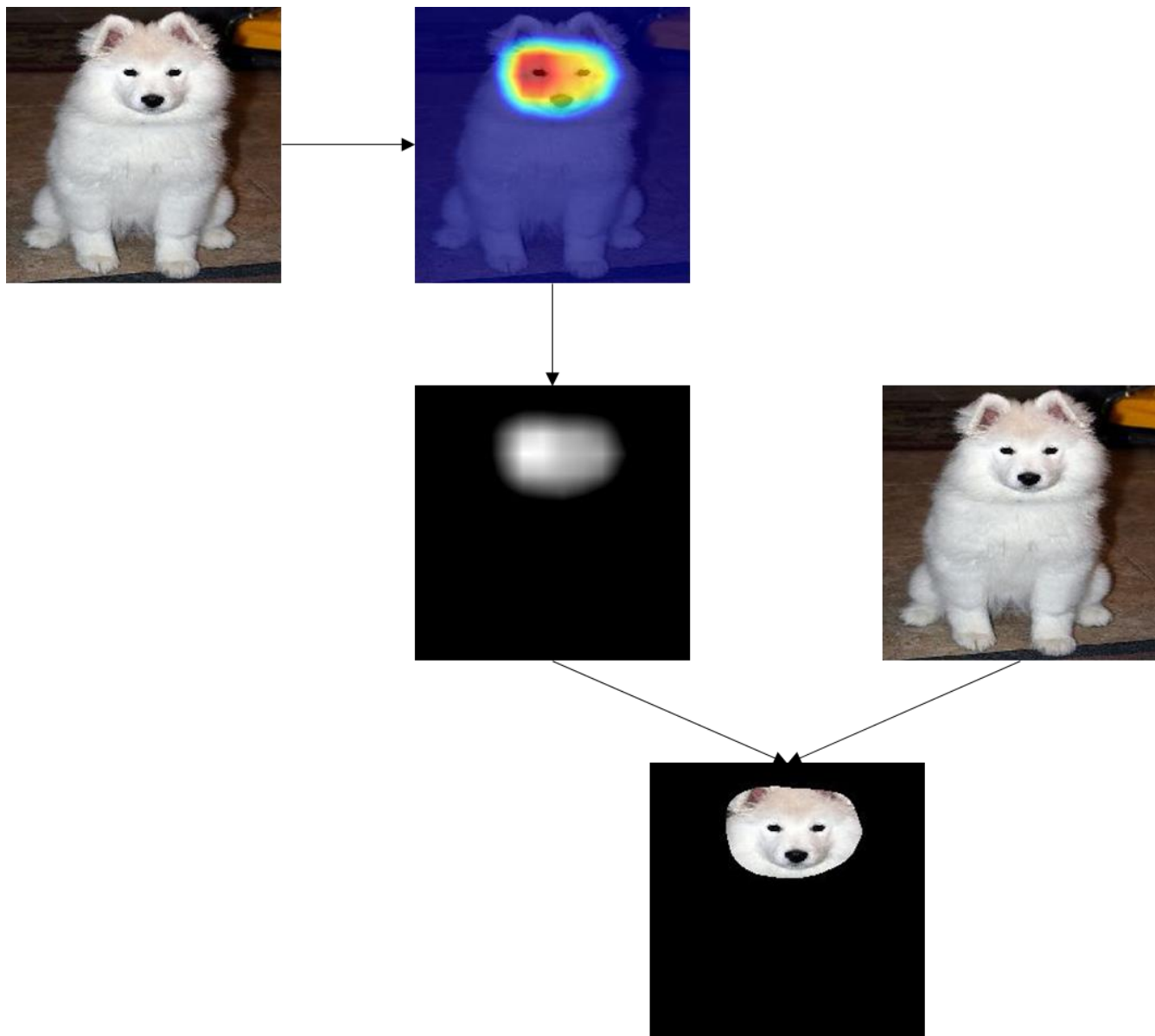


図 4.3 マスクの説明図

4.3 むすび

本章では CycleGAN による重なった動物の生成画像の評価手法についての提案手法を述べた。

第5章 実験結果

5.1 まえがき

本章では, VGG16 で作った 8 クラス分類の精度, それぞれのクラスに対する CAM の結果, CycleGAN による変換結果, CNN による変換画像の評価について述べる.

5.2 データセット

本研究で使用する画像データセットは, イメージネット上で馬, シマウマ, 猫と犬の画像を取得し, 頭だけを切り取って, 残りの胴体を一部でデータセットとして作成した[9]. 猫はラグドール, 犬はサモエドという品種を使用する. CNN と CycleGAN の学習用とテスト用のデータセットは同じものとする. 各クラスのデータセットの枚数を表 5.1 に示す.

表 5.1 データセットの枚数

クラス	画像枚数
馬の胴体	191
馬の頭	202
シマウマの胴体	106
シマウマの頭	168
猫の胴体	150
猫の頭	240
犬の胴体	124
犬の頭	163

5.3 VGG16 の分類結果

データセットを用いて、VGG16 で 8 クラスを分類した。学習時の画像サイズは 224×224[pixel]、ミニバッチサイズは 48、エポック数は 50 とした。VGG16 の分類結果及び各クラスの精度は表 5.2 に示す。

表 5.2 各クラスの精度

	馬の胴体	馬の頭	シマウマの胴体	シマウマの頭
正しい予測	50	43	30	31
合計	50	47	29	32
精度[%]	100	91.5	96.7	96.9
	猫の胴体	猫の頭	犬の胴体	犬の頭
正しい予測	32	69	17	50
合計	43	70	24	52
精度[%]	74.4	98.6	70.8	96.2

5.4 GradCAM の実行結果

学習した CNN モデルを用いることで、入力画像の動物頭と胴体の CAM マップを得ることができる。結果の一部を示す。馬の CAM マップを図 5.1 に示す。シマウマの CAM マップを図 5.2 に示す。



図 5.1 馬の CAM マップ

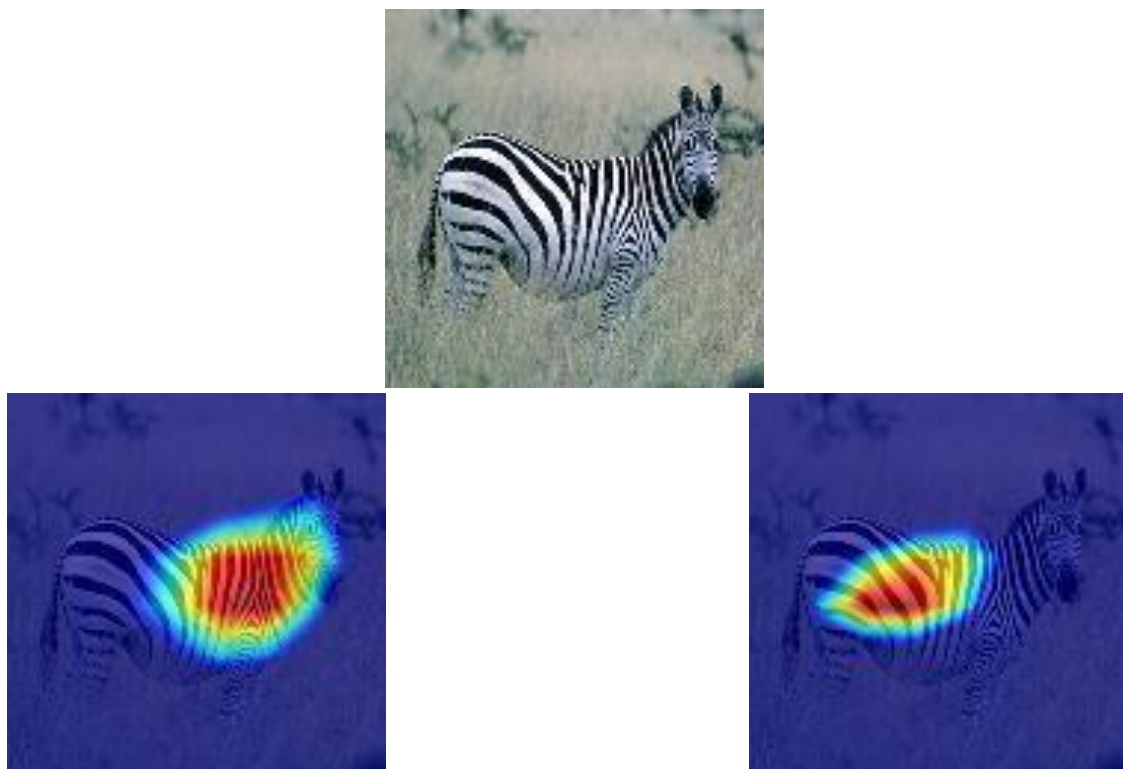


図 5.2 シマウマの CAM マップ

5.5 CycleGAN の実行結果

CycleGAN による変換画像に対してマスク処理をする。この処理を頭と胴体の 2 クラスに適用する。元の画像に対して二枚の変換画像を重ね合わせることで、最終的な合成画像を得る。馬からシマウマへの変換結果例を図 5.3 に示す。犬から猫への変換結果例を図 5.4 に示す。



図 5.3 馬からシマウマの変更結果例

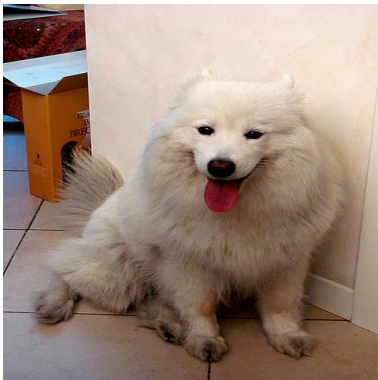


図 5.4 犬から猫の変更結果例

5.6 CNN による評価

生成された合成画像を学習した CNN に入力し、馬、シマウマ、猫および犬を正しく認識できるかどうかを評価する。合成したそれぞれ種類の画像と CNN で正しく認識された画像の割合を求めて、正解率を算出する。今回はシマウマと猫の生成を例として、評価する。CNN による正解率は表 5.3 に示す。

表 5.3 CNNによる正解率

	正解率[%]
馬	77.7
シマウマ	78.5
犬	40.2
猫	46.8

5.7 考察

VGG16で8クラスを分類した結果から見ると、馬の胴体と頭、シマウマの胴体と頭、犬と猫の頭のラス精度が高く、90%以上になったが、今回実験で用いた犬と猫のデータは同じ白い品種で、犬と猫の胴体のクラス精度が低い。また、CycleGANの変換、最後の重ね合わせた結果から比較すると、馬からシマウマへの変更結果がより良い結果になった。犬から猫への変更は、胴体の部分が全然変換されてない。最後、CNNの評価結果から比較すると、馬とシマウマの場合は70%ぐらい、犬と猫は40%ぐらいである。全体的には馬とシマウマの変換が理想的となる。

5.8 むすび

本章では、4種類の動物それぞれ胴体と頭がGradCAMを通して生成された結果とCycleGANにより変換され、重ね合わせた結果とCNNによる結果評価について述べた。

第6章 結論と今後の課題

6.1 結論

本研究では、4種類の動物をそれぞれ2クラスのオブジェクトとして8クラスに分類し、CNNを学習する。学習済みのCNNをGradCAMによりCAMマップを得る。VGG16を通して、予測確率上位2位のクラスに基づいて、CycleGANをそれぞれの適切な処理をする。最後は二つの出力画像を重ね合わせることで、一つの生成画像とする。生成された画像がCNNモデルで正解率を計算される。

CNNによる評価結果と最後の合成結果から見ると、馬とシマウマの変換は成功したといえる。猫と犬は品種の違いによって、結果が理想ではない。

6.2 今後の課題

本実験の結果は、CAMの領域分割精度とCycleGANの変換結果によるとなる。両方の結果を改善する方法は今後の課題となる。また、本研究では、動物だけを研究対象とする。今後は、もっと一般的なものを対象として、適切なオブジェクト細分化を検討すべきである方法と考えられる。

謝辞

研究の実験環境を整えてくださり，問題点を提起して下さったりと適切な指導を賜った渡辺裕教授に心より感謝申し上げます。

また，日頃からご意見や問題解決をくださった研究室の皆様に御礼申し上げます。

最後に，私をここまで育ててくれ，支えてくれた家族に感謝します。

参考文献

- [1] AL SCHOLAR,"実在しないアイドル画像も瞬時に生成する AI「GAN」が賢すぎる！,"< <https://gendai.ismedia.jp/articles/-/63615>>(参照 2020 年 1 月 18 日)
- [2] 斎藤康毅, ゼロから作る Deep Learning - Python で学ぶディープラーニングの理論と実装, オライリージャパン,2016.
- [3] 谷田 部卓, よくわかるディープラーニングの仕組み,2016.
- [4] 岡谷 貴之, ディープラーニングと画像認識 : 基礎と最近の動向,2015.
- [5] Simonyan, Karen, and Andrew Zisserman. "Very deep convolutional networks for large-scale image recognition." arXiv preprint arXiv:1409.1556 (2014).
- [6] Ian J. Goodfellow, Jean Pouget-Abadie* , Mehdi Mirza, Bing Xu, David Warde-Farley, Sherjil Ozair† , Aaron Courville, Yoshua Bengio, Generative Adversarial Nets, arXiv:1406.2661,2014.
- [7] Jun-Yan Zhu, Taesung Park, Phillip Isola, Alexei A. Efros, Unpaired Image-to-Image Translation using Cycle-Consistent Adversarial Networks, arXiv preprint arxiv:1703.10593, 2017.
- [8] Ramprasaath R. Selvaraju · Michael Cogswell · Abhishek Das · Ramakrishna Vedantam · Devi Parikh · Dhruv Batra, Grad-CAM: Visual Explanations from Deep Networks via Gradient-based Localization, arXiv:1610.02391, 2016.
- [9] Jia Deng, Wei Dong, Richard Socher, Li-Jia Li, Kai Li and Li Fei-Fei Dept. of Computer Science, Princeton University, USA, ImageNet: A Large-Scale Hierarchical Image Database, June 2009.

図一覧

図 2.1	ニューラルネットワークの構造例	2
図 2.2	CNN の構造例	3
図 2.3	VGG16 のネットワーク構造例	4
図 3.1	GAN の基本構造	5
図 3.2	CycleGAN の構造	6
図 3.3	GradCAM の例	7
図 4.1	全胴体の流れ	9
図 4.2	画像変換の流れ	10
図 4.3	マスクの説明図	10
図 5.1	馬の CAM マップ	14
図 5.2	シマウマの CAM マップ	14
図 5.3	馬からシマウマの変更結果	15
図 5.4	犬から猫の変更結果	15

表一覧

表 5.1	データセットの枚数.....	12
表 5.2	各クラスの精度.....	13
表 5.3	CNNによる正解率.....	16