

機械学習による人物抽出と奥行きある画像生成

2016SS057 岡田宗馬

指導教員：小市俊悟

1 はじめに

奥行きある画像生成のプログラムはすでに身近な場面で活用されている。例として、iPhone に搭載されたカメラ機能の一つであるポートレートという機能がある。これは、人物を主な被写体とした写真撮影機能であり、背景がぼやけて撮影されるため、人物を際立たせるような写真の撮影が可能となる。

Instagram や Twitter などの多くの SNS が普及している世の中、普段の生活の中でポートレートを使う機会が多くなっていると感じる。現在若者の間で流行っている「インスタ映え」というのは、Instagram に写真を投稿した際に、見栄えが良く映えているという意味であるが、そのように評価される画像にはポートレートを使用した画像が多いと言われている。そういった世の中において今よりも楽しく SNS を活用するために、普通のカメラで撮影した画像をポートレート風の画像にする機能に興味を持つに至った。

ポートレートは人物を対象とした機能であるが、このような特定の物体のみに焦点を当てた画像を生成する技術は、防犯を目的とした監視機能を補助することなどにも応用が可能であると考え、このような機能を既存技術の組み合わせにより実現する方法を研究することにした。

2 手法の説明

人物に関する画像処理技術が進んでいることもあり、本研究ではポートレート風の画像生成を目指す。したがって、使用する画像には人物が含まれることを前提にする。

本研究では物体認識と物体検出の二つの既存技術を使用する。本研究では物体認識と物体検出を次の意味で用いる。物体認識とは、画像や動画内の物体が何であるかを識別する手法であり、物体検出とは、画像や動画内に含まれている対象の物体を自動的に背景から区別して、その範囲を特定する手法である。

2.1 本研究の流れ

本研究では図 1 に示す手順に従ってポートレート風の画像を作成する。

2.2 物体検出の手順

本研究では物体検出の手法として、Semantic Soft Segmentation[1] と呼ばれる手法を利用した。Semantic Soft Segmentation では、はじめに深層学習を用いて、物体をおおまかに分ける。図 2 から図 3 のようになっているように、それが色で表現されている。この時点では物体の境界が不鮮明である。続いて、Semantic Soft Segmen-

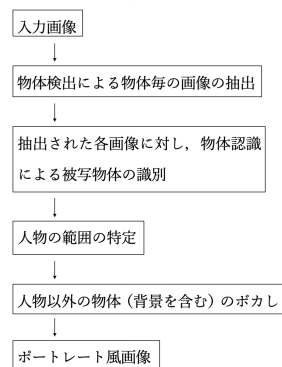


図 1 ポートレート風画像の作成手順

tion では画像データを行列とみなして、元の画像とおおまかに色分けされた画像から固有値分解などを応用して、物体の境界を明確に分ける。図 4 ではそれが色によって表現され、同一色がひとつの物体である。物体ごとの画像抽出は色分けされた画像を元に行うことができる。



図 2 元画像

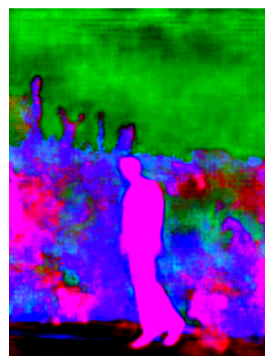


図 3 物体ごとにおおまかに色分けされた画像



図4 物体ごとに明確に色分けされた画像

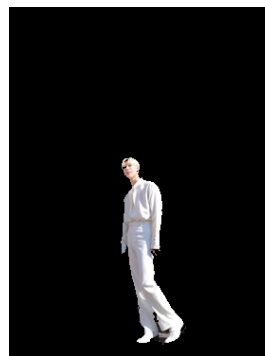


図6 人物以外を黒くした画像

2.3 物体認識の手順

本研究では、物体認識に Single Shot MultiBox Detector[2] と呼ばれる手法を用いる。Single Shot Multi-Box Detector も深層学習を利用した機械学習モデルである。これを利用すると、図5のように、画像中の物体について、物体を囲む矩形領域と、その物体が何であるかを確率とともに表示される。ただし、識別できる物体は、学習した物体のみに限られる。図5に写る物体は、人(Person)である確率が、1.00 である。このような情報が得られれば、物体検出によって切り出された物体が何であるかを識別できる。



図5 Single Shot MultiBox Detector を行なった結果

物体検出によって得られた人物領域の各候補に対して、その領域を元画像から切り出せば、領域に対応する物体の画像が得られる。その際、領域外の部分はすべて黒色に塗りつぶすことにした。図6がこの処理により得られる画像の例である。

2.4 人物以外の領域のボカし

人物と特定された各領域について、それ以外の領域をボカすために、人物領域外の各ピクセルについて、そのピクセルの値と上下左右数ピクセルの値を係数の値の和が1になるように足し合わせることを行った。

ボカす方法を変えることで、異なる印象を受ける画像を生成することが可能であると考えが、本研究の成果はここまでである。

3 結果

結果として、図7のような画像を作成した。

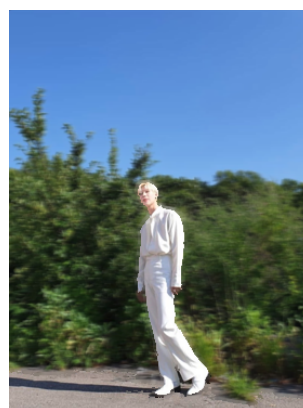


図7 最終的な生成画像

この画像を見てわかるように人物に焦点が合っており、人物以外の領域はボカされていることがわかる。しかし、完成した画像の画質が低いので、全体としてボケているようにも見えている。

4 おわりに

既存技術である物体検出・物体認識を実際に用いて、画像を作成したところ問題があることがわかった。それは、物体ごとに色分けされるべきなのに人物と判定される部分に背景が含まれてしまっていることや作成された画像の画質が極端に落ちていることである。この問題の原因として、人物が認識しにくい形をしていることや、背景の明度が極端に強すぎたり弱すぎたりしている、元画像の画質が低いことが考えられる。

さらに改善するには、これらの問題を解決する必要がある。

参考文献

- [1] Y. Aksoy et al. : Semantic Soft Segmentation, ACM Tran. on Graphs, Vol.37(4), Article 72, 2018.
- [2] W. Liu et al. : SSD: Single Shot MultiBox Detector, arXiv:1512.02325v5, 2016.