PePe: ユーザと機械学習が相互に補い合うセグメンテーションのためのペイントツール

綾塚 祐二 *† 吉高 淳夫 †

概要. 機械学習の技術の発展により、画像から物体の認識や検出を高い精度で行えるようになってきた。特定の領域を検出するような機械学習を行うためにはまず対象物の領域に関するアノテーションが付加されたデータが多数必要となるが、医用画像を始めとする高度な専門知識を必要とする分野においては、正解となるアノテーションをつける作業が行える人の数も限られ、少数の人に大きな負荷がかかる。我々は、ユーザと機械学習が協調してアノテーション作業を進める手法を提案する。ユーザが部分的に領域を指定すると、それをシステムが軽量な機械学習で学習し即時に残りの部分を予測する。予測が不十分な部分のみをユーザが上書きすることでアノテーション作業を完成させていく。

1 はじめに

機械学習の技術が発展し、読影に高度な専門性を要するレントゲンなどの医用画像においても、専門家と同等以上の精度を達成する例が多く報告されている。学習のためのデータは人間が収集しアノテーションを行う必要があるが、医用画像など専門性の高い分野の場合にはこのアノテーション作業が可能な人員も限られ、機械学習の応用の研究や実用化への大きな障壁となっている。

アノテーション作業の負荷の高さの問題は広く認知されており、負荷を軽減するためのさまざまな研究が行われている [3, 4, 5]。しかし、これらは背景の中にある一般的な物体を識別するものを扱っており、判定に専門性を必要とし、識別性も低くまたその境界も曖昧であるような対象は扱われていない。

本稿では、画像中の抽出したいタイプの領域 (病変など) を検出するような場合を対象とし、ユーザと機械学習が協調的に作業を進めるシステムを提案する。ユーザが作業を進めていくと機械学習がまだアノテーションされていない部分を予測する。ユーザはあとは分類予測が正しくならない部分のみをアノテーションすればよく、効率的にアノテーション作業が進めることが可能である。

2 PePe

我々は、画像を数ピクセル〜数十ピクセル程度のブロックに分け、その中のピクセルの RGB 値の統計量を特徴量とし、サポートベクターマシン (SVM) を用いて予測を行う PePe (Predictive Elastic Painting Environment) と呼ぶ試作システムを実装した。PePe の機能は、ブロックの特徴量に基づき塗る (アノテーションする) 範囲が柔軟に変化する EMA (Elastic MArking) ブラシと、特徴量に基づき塗られていない部分を推測する QUIP (QUIck Prediction) キャンバスとで構成される (図 1)。実装は Windows 10 を OS とする PC 上で、Java (OpenJDK 11) を用い、SVM は、Java 版の libsvm (Version 3.23)[2] を利用した。

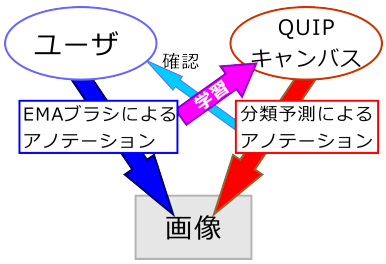


図 1. PePe の作業モデル

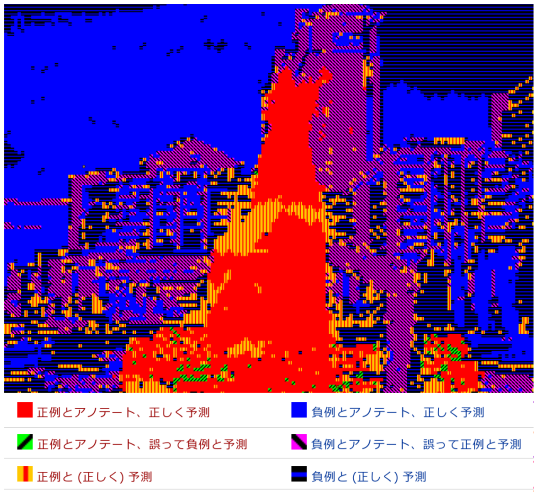
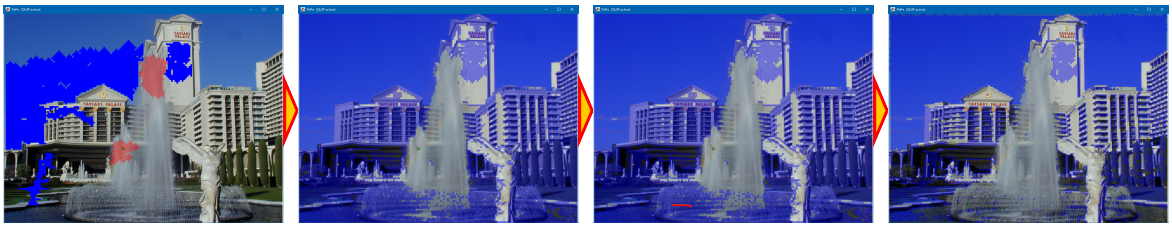


図 2. アノテーションと予測の関係

* 株式会社クレスコ 技術研究所
† 北陸先端科学技術大学院大学 先端科学技術研究科



(1) (2) (3) (4)

図 3. 作業の流れ: (1) 初期のアノテーション→ (2) 予測分類→ (3) 修正→ (4) 再予測

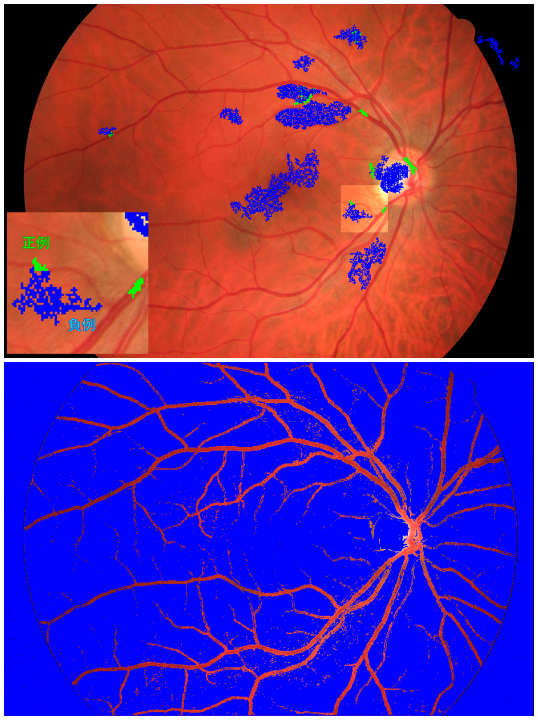


図 4. 眼底写真の血管をアノテーション

ブロックの特徴を表す量として現在は、ピクセル値の平均と分散、エッジ強調を施した画像でのピクセル値の分散を用いている。RGB それぞれの値に対しこれら三つの要素を計算し、これを長さ 1 に正規化して各ブロックの 9 次元の特徴ベクトルとする。EMA ブラシは、カーソル位置にある先端が指すブロックと類似度 (特徴ベクトルのコサイン類似度で計算する) の高い周囲のブロックを、ストロークの長さに応じた広さで塗る。

ユーザが正例・負例のアノテーションをある程度行ったところで (現在の実装では明示的にユーザが指定して) QUIP は予測を開始する。SVM はユーザがアノテーションした正例・負例のブロックの特徴ベクトルを学習データとして学習し、すべてのブロックに対する分類予測を行う。インタラクティブ性を保つため、学習にはアノテーション済の全データは用いず、学習に要する時間が 1 秒程度で済む量をサンプリングする。ユーザがアノテーションした分類と、予測した分類が食い違っているブロックはユーザのアノテーションを優先させる (図 2)。

ユーザから見ると、作業を進めていると途中で QUIP が勝手にアノテーションを追加したような状態になる。多くの場合、QUIP が予測を間違えている部分が存在するので、ユーザはその部分を何箇所か修正する。QUIP はアノテーションされた領域が増えたのを検知し、再学習と再予測を行う。この処理はバックグラウンドで行われ、ユーザが作業途中で「待つ」必要なく数秒ごとに更新される (図 3)。

医用画像に用いた例として眼底写真 (HRF Image Database[1] より) の血管を抽出した例を挙げる (図 4)。僅かな量 (図 4 上の画像の緑色/青色部) のユーザの手によるアノテーションで、全体の予測がうまく機能しているのが判る (図 4 下)。

3 まとめ

本稿では、専門知識を要するような画像のアノテーション作業の負荷軽減のために、ユーザと機械学習が協調して作業を進めるペイントツール、PePe を提案した。今後は、本システムが実際に専門家が必要とする支援を提供できているかを検証していく。

参考文献

[1] High-Resolution Fundus (HRF) Image Database. https://www5.cs.fau.de/research/data/fundus-images/

[2] LIBSVM – A Library for Support Vector Machines. https://www.csie.ntu.edu.tw/~cjlin/libsvm/

[3] Boykov and Kolmogorov. Computing geodesics and minimal surfaces via graph cuts. In *Proceedings Ninth IEEE International Conference on Computer Vision*, pp. 26–33 vol.1, Oct 2003.

[4] J. Fails and D. R. Olsen. Interactive Machine Learning. *International Conference on Intelligent User Interfaces, Proceedings IUI*, May 2003.

[5] L. Yan, B. Fan, S. Xiang, and C. Pan. Adversarial Domain Adaptation with a Domain Similarity Discriminator for Semantic Segmentation of Urban Areas. In *25th IEEE International Conference on Image Processing (ICIP)*, pp. 1583–1587, Oct 2018.