

ARを利用したスケッチ支援システムにおける ペン先を検出する物体検出モデルの作成

鈴木 伶奈[†] 菊地 真人[†] 大園 忠親[†]
名古屋工業大学情報工学科[†]

1 はじめに

「アタリ」とは、絵を描く時に行う大まかな位置決めのための下描きであり、この工程を「アタリをとる」という。スケッチの際の参考物と紙面間との視線移動回数を減らすために、絵やスケッチを描く際のアタリをとる工程に着目したスケッチ支援システムを開発している。本システムでは、AR技術により参考画像を紙面上に配置可能にする[1]。しかし、図1(左)のようにARで画像表示を行なった際に、参考画像が現実の筆記具のペンを遮るため描画位置が見えづらいという欠点があった。

本稿では、この欠点の解決のため、筆記具のペン先(以降単にペン先とする)の位置を物体検出により取得し、そこにポインタを表示する機能について述べる。ここで新たに筆記具のペン先を検出する物体検出モデルを実現した。汎用的な物体検出技術ではペン先のような微小な領域の特定が困難であるため、工夫が必要である。

2 ペン先検出モデルの構築

2.1 ペン先検出機能

本システムは、スケッチの対象である参考物と紙面間の視線移動を減らすための紙面への参考画像の重畳表示機能を持つ。本機能は、AR技術により、参考画像が紙面上に存在するかのように表示する機能である。

ここで、前述したように、筆記具のペン先の位置を強調表示したい。先行研究[1]では、図1(右)のように参考画像の表示部分における手との重複部分を非表示にすることで解決を図った。しかし、ペン先の位置の視認性が低い。そこで、物体検出によりペン先検出位置を推

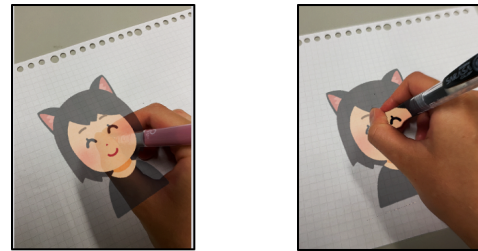


図1 参考画像表示例(左)、手のオクルージョンを適用した例(右)

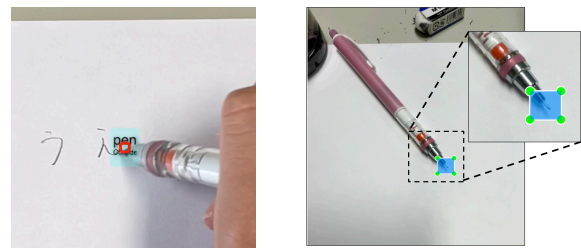


図2 ペン先検出モデルによる検出例(左)、アノテーション例(右)

定し、強調表示することとした。図2(左)は、2.2節で述べるペン先を検出する物体検出モデル(以降、ペン先検出モデル)を使用した検出例である。図2(左)の中央部分のように、ペン先を示すポインタを表示する。

2.2 ペン先検出モデルの作成

2.1節で述べたペン先検出機能のため、ペン先検出モデルを試作した。ペン先検出モデル作成に使用した学習画像群は、システム使用時を想定した画角でスケッチ中の場面を撮影し、動画から筆記具のペン先が写り込んでいるフレームを画像として収集した。まずはプロトタイプとして、筆記具は、三菱鉛筆株式会社のクルトガ^{*1}というシャープペンを使用した。アノテーションツールを使用し、図2(右)のようにアノテーションを付与した。事前の評価により、セマンティックセグメンテー

Creating pen-tip detection model for sketching support system based on augmented reality

[†]Rena SUZUKI, [†]Masato KIKUCHI and [†]Tadachika OZONO

[†]Department of Computer Science, Nagoya Institute of Technology.

^{*1} KURUTOGA official site.

<https://www.mpuni.co.jp/index.html>

表1 表1のモデルによる検出結果

モデル名	使用モデルの種類	mAP[%]	TP	FP	TN	FN	合計
M	YOLOv7	34.2	190	0	0	623	813
M-x	YOLOv7-x	20.4	72	1	0	740	813
M-tiny	YOLOv7-tiny	45.8	508	26	0	279	813

ションは、十分な検出速度が得られないため、より高速な物体検出を使用する。

ペン先検出モデル作成には、リアルタイム処理が可能な物体検出アルゴリズムであるYOLOv7[2]を用いるYOLOv7*²では、7種類の学習済みモデルが提供されている。本システムでは、事前評価の結果、YOLOv7、YOLOv7-x、YOLOv7-tinyの3種類を用いることにした。ペン先の位置の推定速度を向上させるため、検出速度と精度の高さを重視し、計算量の目安であるFLOPsが最低のYOLOv7-tiny、およびFLOPsは少し高いが検出の精度を表す指標APが高いYOLOv7、YOLOv7-xを選択した。

ペン先検出モデル作成時に、訓練用画像715枚、検証用画像120枚、バッチ数16、epoch数100回、学習時解像度640x640(YOLOv7-tiny使用時のみ320x320)で訓練し、実行環境はGoogle Colaboratoryを使用した。YOLOv7、YOLOv7-x、YOLOv7-tinyで訓練したペン先検出モデルを、それぞれ、M、M-x、M-tinyと呼ぶ。

3 評価実験と考察

学習用画像と同様の手段でテスト用画像813枚を収集した。表1に各モデルのmAPを示す。表1のペン先検出モデルを用いて、テスト用画像の物体検出を行なった。結果も表1に示す。表1において、TPはペン先を検出できた画像数、FPは誤検出した画像数、FNはなにも検出しなかった画像数とする。TPの多さの順は表1でのmAPの高さの順と同じとなった。M-tinyによるペン先検出が一番TPが多かったが、FPも多くなった。Mでは、TPはモデル3よりも検出数が少ないが、FPは0である。M-xでは、一番TPが低く、さらにFPも発生したという結果となった。Mに関して、学習時の条件の調整を行えばM-tinyと同等の性能まで向上可能かもしれないが、現段階でM-tinyは高い性能を持ち、実用性も高いといえる。

次に、ペン先検出が成功した画像における、ペン先位置の正しさを評価した。具体的には、バウンディング

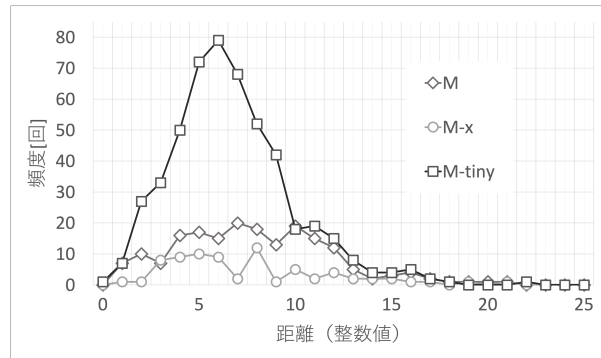


図3 ペン先とバウンディングボックスの中心との距離

ボックスの中心をペン先推定位置として、実際のペン先との距離を計測した。結果を図3に示す。なお、距離の単位はピクセル(px)である。検出ができれば、画像のうち84%は、ペン先とバウンディングボックスの中心との距離が10px以内に収まっていることがわかった。テスト用画像の大きさは1920×1080pxであり、10pxはペン先の確認において十分に小さいといえる。ペン先の位置はバウンディングボックスの中心とすればよいことがわかった。

図1(右)は、M-tinyによるペン先検出例である。検出されたバウンディングボックスを水色の四角で示し、バウンディングボックスの中心に半径10の赤点を描画している。M-tinyのFNが高いため、赤点の描画位置が不正解になる場合があるが、実用上はあまり問題にならないと考えている。

4 おわりに

本稿では、ARを利用したスケッチ支援システムの欠点を補うペン先検出機能を考案した。ペン先という極小領域を検出する物体検出モデルを試作した。本機能により、AR空間における参考画像が、ペン先位置の視認性を低くしているという現時点でのスケッチ支援システムの欠点を解決することを期待している。

謝辞 本研究の一部はJSPS科研費JP19K12266、JP22K18006の助成を受けたものです。

参考文献

- [1] 鈴木伶奈, 菊地真人, 大園忠親: ARと手の骨格認識を利用したスケッチ支援システムの試作, 電気・電子・情報関係学会東海支部連合大会, 2p, 2022.
- [2] C. Wang, A. Bochkovskiy, H. Liao: “YOLOv7: Trainable bag-of-freebies sets new state-of-the-art for real-time object detectors,” arXiv preprint arXiv:2207.02696, 2022.

*² WongKinYiu.2022.yolov7.
<https://github.com/WongKinYiu/yolov7>.