

小学生の書道基本姿勢定着に向けた点画検出および採点手法

Character Strokes Detection and Scoring for Increasing Children's Calligraphy Training time

柿野 優衣 *1 河野 慎 *2 米澤 拓郎 *2 中澤 仁 *1
Yui Kakino Makoto Kawano Takuro Yonezawa Jin Nakazawa

*1 慶應義塾大学環境情報学部

Faculty of Environment and Information Studies, Keio University

*2 慶應義塾大学大学院政策・メディア研究科

Graduate School of Media and Governance, Keio University

In order to encourage elementary school students unfamiliar with Japanese calligraphy to have an assertive attitude towards acquiring basic calligraphy skills, it is imperative to have an application that provides an environment which allows them to practice calligraphy regardless of the time and place. Due to the fact that this application requires an evaluation system that does not need a calligraphy teacher's assistance, this research proposes a method of automatically evaluating a work of calligraphy. All detected strokes within one sheet will be regarded as an object detection problem, therefore, the entire work's evaluation will depend on the detected strokes. Strokes are a fundamental technique in calligraphy, and are a vital component in creating a work of calligraphy.

1. はじめに

書道に対して基本姿勢を身につけることは書作品の向上において重要である。書道基本技術である点画を意識して筆を動かすことは文字の構成を整え、作品全体を美しくする。特に習い始めは、点画の練習において筆の角度、圧力などからの線質の変化を学ぶことで基本を身につけるため、練習量の増加に伴わない書道技術が向上していく。ゆえに初めて書道に触れる小学生からこれらを行っていくことは、書作品の向上において重要であるといえる。小学生は、小学校の授業内や教室に通うなどして指導者がいる下で練習を行う。しかし、書道指導者は複数の小学生に対し1人であることが多い。そのため小学生の希望時間に沿った練習を行うことは困難であり、限られた時間内での練習量になってしまう。

この問題の解決策として小学生が希望時間に練習環境を得られるスマートフォンアプリの導入が挙げられる。例えばスタディサプリー *1 ではスマートフォン上で予備校の授業を受けることが可能となっており、小・中学校導入による成果も報告されている [株式会社 17]。書道練習に同様なアプリケーションを導入するには、書道作品の評価を自動で行えることが重要だと考える。そこで本研究では画像認識技術を用いて小学生の書作品に含まれる点画を検出し、作品の点数化を行う手法を提案する。図 1 に全体の流れを示す。まず点画検出には物体検出手法の Single Shot MultiBox Detector (SSD) [Liu 16] を採用する。データセットには書道教室に通う小学生 1 - 6 年生の実際の書作品を用いて、5 種類の点画検出を行う。さらに、5 種類の点画を美しさに関する 2 種類のラベルを用いて評価し、10 種類の点画検出も行う。実験では学習に用いるデータセット数の増加に伴う平均適合率 (AP) の全カテゴリの平均 (mAP) を用いた精度比較および、書作品の点数化について定性的に検証する。実験の結果、データセット数が増加するにつれて、点画検出の精度の向上および点数に関する妥当性を確認



図 1: 全体の流れ。1. 作品を書く。2. 書作品をカメラやスキャナなどで画像化する。3. 5 種類の点画およびそれぞれの美しさを検出し、その結果にもとづいて点数を算出する。

することができ、提案手法の有効性を示した。

2. 書道作品における点画検出

小学校の授業では伝統的な言語文化に親しむことを目的として 1, 2 年生では硬筆, 3 年生からは毛筆を学び始める。小学生学習指導要領解説, 国語編に「点画の書き方や文字の形に注意しながら、筆順に従って丁寧に書くこと」[gak 17] とある。基本技術である点画を意識して書くことは文字を構成し書作品を美しくする上で重要なためである。また、点画の筆の運びを練習することは筆をうまく扱うことにつながり、線筆の上達を促す。線筆は書道作品にとって重要な要素であり、点画を練習することは線筆の向上にも役立つ。これより、点画を意識した筆の運びが基礎を身につける上で大切であり、点画の美しさが書作品を評価すると考える。そのため点画評価を用いて作品の点数化を行う。

書作品内の様々な箇所に見られる点画を検出及び採点をする必要がある。そこで本研究では、点画検出を物体検出問題として取り組む。カメラやスキャナなどを使用して画像化された書作品が入力として与えられた時、その画像に含まれる点画を検出する。本稿で点画の種類を図 2 に示す。本研究では全 8 種類の点画のうち書作品構成で重要な点画である点、止めの始筆、そして左払い、右払い、跳ねの終筆の 5 つを対象とする。

連絡先: 柿野優衣, 慶應義塾大学環境情報学部

〒 252-0882 神奈川県藤沢市遠藤 5322 デルタ棟 S213,
yuui@ht.sfc.keio.ac.jp

*1 <https://studysapuri.jp>

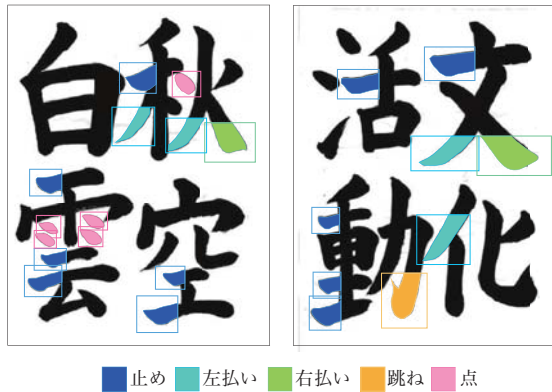


図 2: データセット例. 小学生楷書作品を対象とする. 止め, 左払い, 右払い, 跳ね, 点の 5 種類の点画を示す. 図に示すように点画を四角で囲みアノテーションを作成する.

3. 提案手法

本章では, 書作品における点画の検出手法と採点手法について説明する.

3.1 点画検出

本研究では, 点画の検出問題を物体検出問題とみなし, 深層学習を用いた物体検出技術を用いる. 物体検出技術では, Faster R-CNN [Girshick 15] や YOLO [Redmon 16a, Redmon 16b] なども挙げられるが, 本研究では検出速度でなく検出精度に重点を置くため, single shot multibox detector (SSD) [Liu 16] を採用する. SSD は, ニューラルネットワークの構造に VGG16 [Simonyan 14] を採用しており, 非常に高い精度での物体検出が可能である. VGG16 によって抽出された特徴マップを 6 種類のスケールに畳み込み, それぞれのスケールでの枠線に関する出力をする. 出力では, あらかじめ定義されている枠線の初期値からのオフセットと, 検出対象となる各点画に対する確信度を算出する. この 6 種類のスケールを用意することで, 様々な大きさの枠線に対応することができ, 点のように小さな点画から, 払いのような大きい点画まで対応することが可能である. SSD に図 2 で示している点, 止めの始筆, そして左払い, 右払い, 跳ねの終筆の 5 種類を学習させる. 学習では, バッチサイズは 1, 最適化アルゴリズムには Adam [Kingma 14] を利用し, 学習率は 0.0003 とした.

3.2 採点手法

点画の評価は, 美しさを定量的に図ることは難しいと考えることから, 数段階の評価で美しさを判断するのではなく美しいか美しくないかの 2 段階で評価を行う. 5 種類の各点画に対し「美しい」「美しくない」の 2 種類のラベル付けをさらに行う. 「美しい点」「美しくない跳ね」などの合計 10 種類の物体検出問題とする. 作品の点数は検出された点画の合計に対し「美しい点画」と検出された点画の割合を 100 倍することで算出する.

$$score = \frac{N_{\text{美しい点画}}}{\sum N_{\text{検出された点画}}} \times 100 \quad (1)$$

図 5 に, 実際に実験結果から点数を算出する様子を示す.

4. 実験

実験では, はじめに SSD による点画の検出が可能であるか mAP を用いて検証する. その際に学習に用いるデータ数を変

表 1: データセットアノテーション数

点画種類	美しい	美しくない	合計
点	148	65	213
跳ね	18	80	98
左払い	56	164	220
右払い	146	88	234
止め	403	260	663

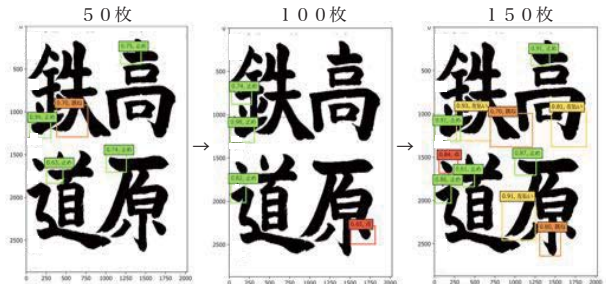


図 3: 点画検出数変移. データセット数が 50 枚, 100 枚時の検出総数と比較して 150 枚での検出総数が増加している.

化させ, 検出精度の変化を観察する. 次に, 各点画の評価を含めた 10 種類に拡張し, 検出精度の検証と書作品の点数化の定性的な評価を行う.

4.1 データセット

本研究では, 書道教室に通う小学 3 年生から 6 年生の楷書作品, 合計 180 枚をスキャナを利用し画像にしたものをデータセットとして用いる. アノテーション作成は図 2 に示すように行う. 全作品にアノテーションした結果を表 1 に示す. 点画評価の際のラベル付けは書道師範取得者が行う. 表 1 には「美しい点画」「美しくない点画」のアノテーション個数を示す.

4.2 結果・考察

実験の結果を表 2 に示す. データセット数が増加するにつれて mAP を比較すると精度が向上することがわかった. これは, 図 3 に示すようにデータセットが 50 枚, 100 枚時点での点画の検出総数に対し, 150 枚での結果の方が検出総数が多いことからわかる. このことからデータセット数を増やすことでモデルの精度が向上する可能性を持つことが考えられる. 次に表 2 に示すデータセット数 150 枚での学習結果に着目し, 各点画で mAP の値を比較すると, 「止め」はアノテーション個数が最も多かったに対し AP の値は「点」について 2 番目に低い. 一方で, 「跳ね」のアノテーション個数は全体と比較して少ないが, AP の値は最も優れていることがわかる. これより, 点画の形によって検出のしやすさが異なることがわかった. これは点画それぞれの特徴が要因であると考えられる. 図 4a には「止め」の始筆の例, 図 4b 「跳ね」の終筆の例を示す. 「止め」の始筆は「跳ね」の終筆に比べ特徴が少なく形が多様であるのに対し, 「跳ね」は輪郭が特徴的であることが他の点画に比べ AP が最も高い要因ではないかと考察する.

図 7 に点画の誤検出例を示す. 1 つの「右払い」に対し, 2 つの「右払い」が検出されている. このように 1 つの点画に対して複数の点画が検出されるような誤検出があった. これは Non-Maximum Suppression 適用時における IoU の閾値が適していないことが考えられる.

図 5 に検出結果の例を示し, 各点画の AP と mAP を表 3 に示す. 図 5 の作品の場合, 検出された検出された運筆数が

表 2: データセット数増加に伴う mAP の変化

データ数	mAP	点	跳ね	左払い	右払い	止め
50 枚	27.1	9.1	4.5	34.2	31.9	55.7
100 枚	49.9	34.7	57.9	44.1	45.1	67.9
150 枚	62.2	45.5	72.0	68.4	59.5	65.9



(a) 止め



(b) 跳ね

図 4: 点画の例

10個に対し美しいと判断された運筆が4個であるため作品の点数は40点である。表2の各点画のAPを見ると150枚のデータセットで学習した際の値と比較して、ラベル付けを行なった各点画のAPが低いことがわかる。これはデータセットを「美しい点画」「美しい点画」に分類したことで表2に示す通りデータセットが減ってしまったことが原因であると考えられる。また、上記に述べたように一つの点画に複数の点画が検出されることで、この検出された点画が「美しい点画」であった場合、作品の点数が高くなってしまいう問題が生じた。

図6に評価のラベル付けを行なった点画の検出結果の成功例と失敗例を示す。表内が青く示されているものはラベルと検出結果が一致したもので、桃色のものはラベルと検出結果が異なったものである。表内の左払いに着目すると左払いを書く筆の運びの途中に線がかすれていなものを美しいと判断しているのではないかと考える。次に止めに着目すると、始筆の筆を置いて下に向かって反りながら筆を運ぶことでできる線を美しいと判断していると推測する。美しい止めが美しくないと判断される要因として始筆から山なりに筆を運んでいることが挙げられる。書作品点数化を師範者が定性的に評価を行なった結果、書作品に対する点数は妥当であるものが多いことがわかった。

一方で本研究では1人の師範者がラベル付け、定性的評価を行なったため主観的な点数化であることが考えられる。そのため今後の課題として、複数人の師範者を評価に取り入れることでより書作品点数化に客観性を持たせる必要がある。今後アプリケーションを実装する際、書作品の取得にはカメラを用いることが想定される。提案手法で用いている深層学習には同一の物体であっても撮影状況によって異なる結果を出力してしまうドメインバイアスが報告されている[Ganin 16]。ゆえに本提案手法がカメラで撮影した書作品の点画を正しく検出できるか検証する必要がある。今後の課題としてあげられる。

5. 関連研究

書作品の評価。 本研究と同様にChin-Chuan Hanらの研究では書道作品の点数化を行なっている[Han 08]。彼らの研究では主に文字の中心、大きさ、手本の線との差異で点数をつけている。文字の中心や大きさは基本的な技術を身につけた後の評価対象となるため、書道の基本技術である点画をもとに点数化をすることは初心者に対し有用であると考えられる。また久保田ら

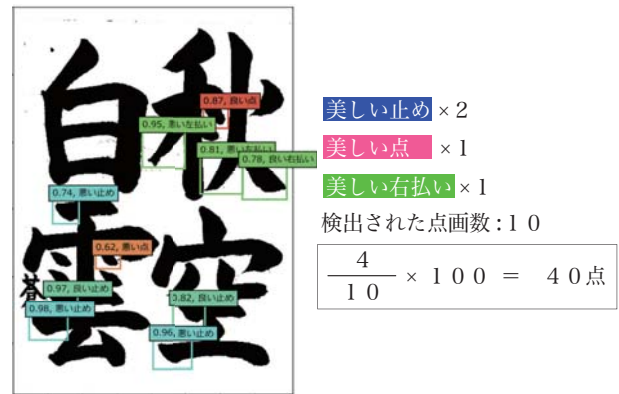


図 5: 作品点数化。点画を評価し採点手法に基づいて作品の点数を算出する。美しいと判断された点画は全部で4個であり、検出された点画は全部で10個である。したがって、この書作品の点数は $\frac{4}{10} \times 100 = 40$ 点となる。

		モデルの出力結果	
		美しい	美しい
正解	美しい		
	美しい		

図 6: 評価後の点画検出の成功例・失敗例。青色は検出の成功例、桃色は失敗例となっている。これらの結果から、左払いを書く筆の運びの途中に線がかすれていなものを美しいと判断し、止めにおいては、始筆の筆を置いて下に向かって反りながら筆を運ぶことでできる線を美しいと判断している。

のように、ユーザに字を書く技術が向上していると錯覚させ、練習への意欲を向上させる研究もある[久保田 16]。

深層学習による物体検出。 本研究では、物体検出手法として、SSD [Liu 16] を用いたが、SSDと同様に物体検出するアルゴリズムとしては、YOLO [Redmon 16a, Redmon 16b] や Faster R-CNN [Girshick 15] などが挙げられる。これらの手法は、SSDと比較しても性能は十分に高い。したがって、今後これらの手法でも点画の検出および評価を行い、検証していく必要があるといえる。また、今回は点画を四角の枠線で囲み、物体検出タスクとして取り組んだ。しかし、点画を四角で囲む範囲は、アノテーションする人によって異なってしまうため、実験における検出精度に影響があると考えられる。そこで、Heらの研究[He 17]のようなインスタンス分割手法の適用も有用であると考えられる。点画の始筆から終筆までを一つのインスタンスとみなして検出することで、より正確な点画の検出および評価が可能になると考えられる。

アプリによる行動支援。 本研究では、書道練習をアプリケーションで行うことを想定している。同じように、スマートフォンのアプリケーションを用いて、教育を行おうとしている研究は数多く存在する。Kangらは、外国語を学ぶ子どもに本を積極的に読ませるためにアプリケーションを実装している[Kang 16]。同様にGonzalezらは物理学を学ぶためのアプリケーションを提案しており[González 17]、Briz-Ponceらは医学部

表 3: 評価した点画の検出結果

点画種類	美しい	美しくない
点	34.4	1.8
跳ね	100	50.9
左払い	16.3	33.4
右払い	28.7	18.1
止め	37.3	46.3

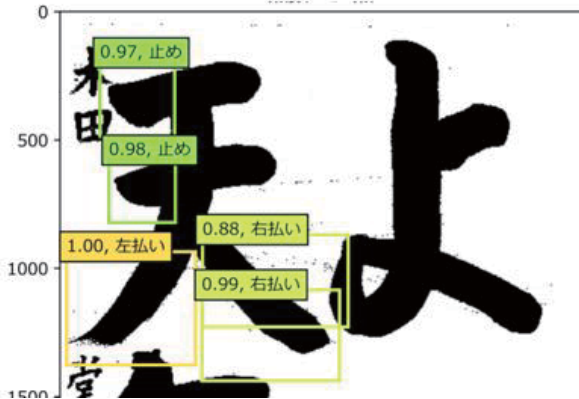


図 7: 作品誤検出. 1つの右払いに対し, 2つの右払いが検出されている. これは, 重複した枠線を取り除く NMS の閾値が高かったためと考えられる.

の学生や先生に対して様々な健康に関する知識を提供するアプリケーションを実装している [Briz-Ponce 16]. このように, スマートフォンのアプリケーションを用いて, 子どもや大人の教育を試みる研究は数多くあり, 本研究での書道アプリケーションも有効であることが考えられる.

6. まとめ

本研究では, 小学生の書道基本姿勢定着に向けたアプリケーションの実現のため, 書作品の点数化を行った. その際に物体検出問題とみなし, 物体検出手法の SSD を適用した. 実験の結果, 書作品に含まれる点画の検出が可能であることを確認した. また, データセット数の増加により mAP が向上したことで, 今後データセットを増加させることでより正確な点画検出が可能であることを示した. また, 「美しい点画」「美しくない点画」のラベルづけを行い提案した採点手法により, 書作品の評価を物体検出を通して行えることの可能性について述べた. 今後, 提案手法を組み込んだアプリケーションを実現させることで小学生の書道技術の向上させ, 書道の発展に貢献していきたい.

参考文献

[Briz-Ponce 16] Briz-Ponce, L., Juanes-Méndez, J. A., and García-Peñalvo, F. J.: Exploring Mobile Learning Apps for Medical Students and Health Care Professionals, in *Proceedings of the Fourth International Conference on Technological Ecosystems for Enhancing Multiculturality*, TEEM '16, pp. 461–465, New York, NY, USA (2016), ACM

[gak 17] 小学校学習指導要領解説 国語編, 文部科学省 (2017)

[Ganin 16] Ganin, Y., Ustinova, E., Ajakan, H., Germain, P., Larochelle, H., Laviolette, F., Marchand, M., and Lempitsky, V.: Domain-adversarial training of neural networks, *The Journal of Machine Learning Research*, Vol. 17, No. 1, pp. 2096–2030 (2016)

[Girshick 15] Girshick, R.: Fast r-cnn, *arXiv preprint arXiv:1504.08083* (2015)

[González 17] González, M. Á., González, M. Á., Martín, M. E., Llamas, C., Martínez, Ó., Vegas, J., Herguedas, M., and Hernández, C.: Teaching and learning physics with smartphones, in *Blended Learning: Concepts, Methodologies, Tools, and Applications*, pp. 866–885, IGI Global (2017)

[Han 08] Han, C.-C., Chou, C.-H., and Wu, C.-S.: An interactive grading and learning system for chinese calligraphy, *Machine Vision and Applications*, Vol. 19, No. 1, pp. 43–55 (2008)

[He 17] He, K., Gkioxari, G., Dollár, P., and Girshick, R.: Mask r-cnn, in *Computer Vision (ICCV), 2017 IEEE International Conference on*, pp. 2980–2988 IEEE (2017)

[Kang 16] Kang, D. and Yoon, K.: A Study on Developing the Interface of Mobile E-learning Application for Children's Foreign Language Education, in *Proceedings of the International Working Conference on Advanced Visual Interfaces*, pp. 320–321 ACM (2016)

[Kingma 14] Kingma, D. P. and Ba, J.: Adam: A method for stochastic optimization, *arXiv preprint arXiv:1412.6980* (2014)

[Liu 16] Liu, W., Anguelov, D., Erhan, D., Szegedy, C., Reed, S., Fu, C.-Y., and Berg, A. C.: Ssd: Single shot multibox detector, in *European conference on computer vision*, pp. 21–37 Springer (2016)

[Redmon 16a] Redmon, J., Divvala, S., Girshick, R., and Farhadi, A.: You only look once: Unified, real-time object detection, in *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 779–788 (2016)

[Redmon 16b] Redmon, J. and Farhadi, A.: YOLO9000: better, faster, stronger, *arXiv preprint*, Vol. 1612, (2016)

[Simonyan 14] Simonyan, K. and Zisserman, A.: Very deep convolutional networks for large-scale image recognition, *arXiv preprint arXiv:1409.1556* (2014)

[株式会社 17] 株式会社リクルートマーケティングパートナーズ: スタディサプリを活用した東京都渋谷区立小・中学校でのアダプティブラーニング実証実験結果のご報告, プレスリリース (2017)

[久保田 16] 久保田 夏美, 新納 真次郎, 中村 聡史, 鈴木 正明: Mojivator: 手書き文字の融合により書きたくなる練習システム, in *Workshop on Interactive Systems and Software* (2016)