

深層学習による平面図形での定理適用可能部の提示

西大和学園高等学校2年 煙岡知樹 吉田伊織

1 背景

現在図形問題を解くAIが多く開発されているが、代数的な処理に頼るものが多く、チェバの定理などの複雑な定理を用いて図形の形状に注目するものはない。

→画像認識技術を用いてチェバの定理等の一定の規則を持つ形状を抽出する

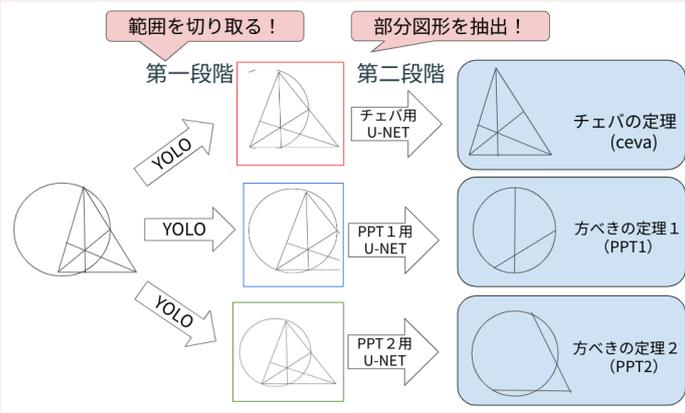
画像認識がよく応用される現実風景の写真とは異なり、白黒で明るさの変化も少ない単調な画像のため、特徴をうまく捉え抽出できるかは明らかではなかったが、応用できると仮説を立て、検証を行った。

2 研究方法

<対象とする定理>

→チェバの定理(以下ceva)、方べきの定理1(以下PPT1)、方べきの定理2(以下PPT2)

本実験では、チェバの定理・方べき定理を対象とした。これらは下図の右側の形状の図形に対して適用できる定理であるが、方べきの定理は2直線の交点が円の内部にある場合を1、外部にある場合を2というように位置関係の異なるものを別の定理として実験を行った。



<第一段階>

YOLOと呼ばれる物体検出技術を用いて、元の図形から定理を含む部分図形をトリミングし、その定理の分類も合わせて行った。(※これは1つの学習モデル)

↓ 図形を切り出すだけでは定理に関係のない線も残るので、...

<第二段階>

U-NETと呼ばれる、元画像から抽出した特徴をもとに新たな画像を生成する技術を用いて、不要な線を消した定理部分のみを抽出した。(※ここでは定理ごとに1つの学習モデルを作成し、第一段階で分類した定理名をもとに、どのモデルを用いるかを決定する。)

※第一段階と第二段階を接続する部分はOpenCVなどの既存の画像処理で容易に実現可能と考えたため実装は行わず、2つのモデルの作成のみに焦点をおいた。よって第二段階の入力画像は入力として予測される画像を作成したものである。

3 モデルの学習

○学習環境、使用ライブラリ → google colabouratory上でkerasを用いて行った。

○図形画像→OpenCVを用いて条件を満たす画像をランダムに作成した。

※画像サイズは128px×128pxとした。

2つのモデルのそれぞれの情報は以下の左の表の通りである。

※今回学習したU-NETのモデルの構造は以下の右図の通りで、画像サイズを128px×128pxから32px×32pxに落として特徴を抽出し、元の画像サイズの新しい画像を生成している。

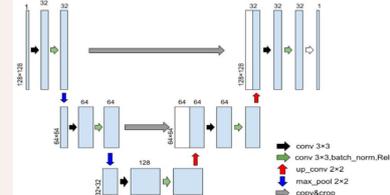
<YOLO>

データ数	訓練:600枚 検証:150枚 テスト:150枚
エポック数	100エポック

<U-NET(3つ全て)>

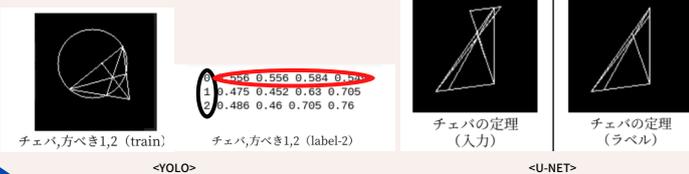
データ数	訓練:3000枚 検証:750枚 テスト:750枚
エポック数	100エポック
損失関数	binary_crossentropy
最適化関数	adam

<U-NETのモデル図>



2つのモデルの学習に用いた教師データの一例をそれぞれ以下に示す。

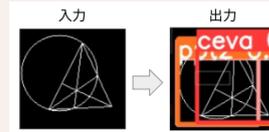
※YOLOのラベルについて、この例では3つの全ての定理が含まれるが、下図の黒い楕円で囲った数字は定理の識別番号(0→ceva, 1→PPT1, 2→PPT2)を表し、赤の楕円で囲った数字は切り取るべき長方形(バウンディングボックス)の場所と大きさを示している。



4 結果

モデルの評価に関して、損失関数(binary_crossentropy)では不十分なので(うまく抽出できなかった場合にも損失が、値としてはやや低く出ることが分かった。)、学習後のモデルに検証データの画像を入力した際の出力画像を目視で判断した。YOLOでは「定理を全て囲っている。かつ、定理の名称が合っている」、U-NETでは「余分な線と認知できるものがない」というのを抽出に成功した画像の判断基準とした。結果、YOLOでは150枚全てで、U-NETでは3種類750枚ずつのうち9割以上で抽出に成功した。実際の出力を以下に示す。

<YOLO>



<U-NET>

モデルの種類	入力画像	実際の出力画像
U-NET (Ceva用)		
U-NET (PPT1用)		
U-NET (PPT2用)		

○ 消えるべき部分 (一部のみ示した) ○ 消えるべきが残った部分

5 考察

<第一段階>

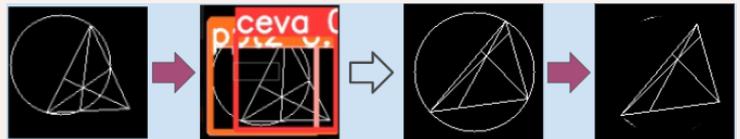
第一段階のYOLOでは複数の定理を名称、範囲ともうまく囲うことができたと言える。ただし、今回の入力画像は全て3つの定理のうち何らかの定理画像が含まれるものだったので、3つとも含まない図形に対してうまく抽出できるかはこの実験からだけでは判断できない。このような図形に対応するには「定理なし」というクラスを作れば可能か検証できるのではないかと考えた。

<第二段階>

第二段階のU-NETでは、それぞれに特化したモデルを作れば定理のみの部分図形を取り出すことは可能であると考えられる。一部余分な線が少し残ってしまうのが見られたが、些細なもので、定理の部分の判別するのに支障をきたすほどのものは少なかった。余分な線が残る画像の1つの傾向として、その画像に近い位置関係の教師データの画像が少ないことがやや感じられた。また、辺の本数を大幅に増やした入力画像に対してはうまく抽出できなかった。これらの点を踏まえて、教師データの数をより増やし、位置関係、辺の本数の種類等も増やしていけばより精度は上がるのではないかと考えた。

今回は3つの定理のみを対象としたが、他の定理に関しても行うことは大きく変わらないので同じ手法が適用可能なのではないかと考えた。

6 展望



<第一段階>

<第二段階>

5の考察を踏まえ、教師データを増やすことでより多くの図形に対応できるモデルを作り、その上で、今回実装しなかった第一段階と第二段階を接続する部分も実装して1つのモデルとして作成したい。この際接続部分では、第二段階のU-NETに入力するために画像サイズを128px×128pxに直すが、①足りない部分を付け足す。②長い方の一辺が正方形の一辺になるように足した上で拡大する。の2つの方法を考えており、どちらのほうが精度がいいのか等も検証していきたい。

7 参考文献

- 1.岩根 秀直, 松崎 拓也, 穴井 宏和, 新井 紀子. 「東口ぼくん, 代ゼミ模試に挑戦!」数学チームの結果について (数式処理とその周辺分野の研究). 2014.
- 2.Minjoon Seo, Hannaneh Hajishirzi, Ali Farhadi, Oren Etzioni, and Clint Malcolm. Solving geometry problems: combining text and diagram interpretation. 2015.
- 3.Pan Lu1,Ran Gong1, Shibiao Jiang2, Liang Qiu1, Siyuan Huang1, Xiaodan Liang3, Song-Chun Zhu. Inter-GPS: Interpretable Geometry Problem Solving with Formal Language and Symbolic Reasoning. 2021
- 4.鈴木 慎人, 加藤 祐悟, 全 邦釘 ディープラーニングを用いた地中レーダ探査による埋設物の位置推定2021
- 5.https://github.com/ultralytics/yolov5
- 6.Olaf Ronneberger, Philipp Fischer, and Thomas Brox. U-Net: Convolutional Networks for Biomedical Image Segmentation. 2015
- 7.https://github.com/Microsoft/VoT/releases