A-11　回路図画像に対する視覚言語モデルの構造抽出性能の評価

磯辺彰哉*，嶋田和孝**
（*九州工業大学 情報工学部，**九州工業大学大学院 情報工学研究院）

1．はじめに

現在，生成 AI の回路設計への応用が進んでいる[1]．その一環として，視覚言語モデル（VLM）による回路図の構造抽出が試みられている．しかし，VLM が元々持つ事前知識や，欠陥のない正しい回路図（以降，整合データと呼ぶ）のみで学習を行った場合，モデルが，回路は正常に接続されているという前提に強く引きずられる．その結果，断線などが生じた回路図（以降，不整合データと呼ぶ）が入力されると，不整合箇所を誤って補完し，存在しない接続として推論してしまう問題がある．

本研究では，整合データの抽出精度を維持しつつ，この不整合な回路図に対しても，不整合な状態として正確に構造抽出することを目的とする．そのために，不整合データを擬似的に生成し，それを用いて VLM をファインチューニングする．不整合データを含まない学習モデルと比較することで，不整合データを学習に用いることの有効性を検証する．

2．回路図データセットの構築と不整合データの定義

本研究の構造抽出では，回路図における接続関係や素子情報をテキスト形式で表現する出力形式として，作図記法である Mermaid コードを採用する．回路図画像および対応する Mermaid コードの自動生成には，Python の描画ライブラリである schemdraw を用いる．回路図画像の生成は，直列，並列，ブリッジ構造を左から右にランダムに繋ぎ，素子もランダムに割り当てることで，ベースとなる回路である整合データを構築する．さらに，整合データに対して，以下の 2 種類の処理を加えることで不整合データを生成する．

・断線（BROKEN_WIRE）：導線が途切れる
・ラベル欠落（NO_LABEL）：素子記号のみで素子名無し

図 1 は不整合データの回路図の例であり，図 2 は図 1 とペアで生成された Mermaid コードの抜粋である．図 1 の色枠の箇所は，図 2 の同色の記述に対応する．この手順で，整合・不整合を各 1,000 件，計 2,000 件作成した．

3．学習手法と評価手法

VLM として Qwen2-VL-7B-Instruct[1] を採用し，入力を回路図画像，出力を対応する Mermaid コードとするタスクにおいて，4-bit 量子化と LoRA を用いてファインチューニングする．断線等の不整合箇所を誤って補完してしまう問題の抑制効果を検証するため，整合データのみで学習した不整合なしモデルと，不整合データを含めて学習した不整合ありモデルの 2 種類を比較する．

評価には乱数シード値を変えた 5 回の反復評価を行う．各反復では，全 2,000 件からテストデータとして整合・不整合データを各 100 件ランダムに抽出する．残りのデータから，不整合なしモデル用に整合データ 800 件，不整合ありモデル用に両データを各 400 件ランダムに抽出し，それぞれ学習させる．評価指標には，出力と正解の Mermaid コードから空白や改行を除外した完全一致率を用い，全 5 回の平均値を最終評価とする．

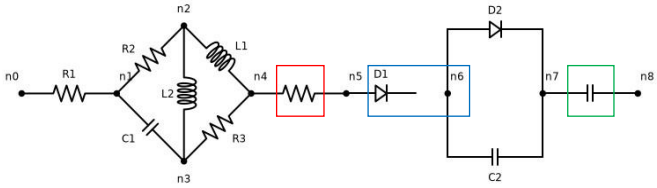

図 1．断線やラベル欠落の不整合を含む回路図画像例

```
graph TD
  %% ---［中略：ブリッジ回路部（整合認識）］---
  **UNKNOWN_0["Resistor  (NO_LABEL) "]**
  **UNKNOWN_0 -- "p" --- n4**
  **UNKNOWN_0 -- "n" --- n5**
  **D1["Diode  (BROKEN_WIRE) "]**
  **D1 -- "p" --- n5**
  **D1 -- "n" --- open0**
  %% ---［中略：孤立した回路ブロック］---
  **UNKNOWN_1["Capacitor  (NO_LABEL) "]**
```

図 2．回路図とペアで生成された Mermaid コードの抜粋

表 1. 全 5 回における反復評価での完全一致率の平均値

	不整合なしモデル		不整合ありモデル	
	整合	不整合	整合	不整合
平均	85.6%	22.8%	85.6%	89.8%

4．結果と考察

表 1 に全 5 回の試行での完全一致率の平均値を示す．不整合なしモデルは，整合データに対して 85.6% を記録したが，不整合データに対しては 22.8% に留まった．これは，未知の不整合箇所に対して，モデルが回路は正常であるという前提に強く引きずられたためである．その結果，実際の描画状態を無視し，存在しない接続やラベルを誤って補完してしまったと考えられる．

一方，不整合ありモデルは，整合データに対して不整合なしモデルと同等の 85.6% の精度を維持しつつ，不整合データに対しては 89.8% の精度を達成した．この結果から，モデルは整合データを読み取る能力を失わず，回路図画像に描かれた状態を正確に構造抽出する能力を獲得したことが示唆される．

5．まとめ

本研究では，不整合データを含めた学習の有効性を明らかにするため，不整合なしモデルとの比較を行った．反復評価の結果，不整合ありモデルは，整合データの高精度な抽出を維持しつつ，断線やラベル欠落等の不整合を含む回路図に対しても，高い構造抽出能力を得た．

参考文献

[1] Y. Lai, et al., "AnalogCoder: Analog Circuit Design via Training-Free Code Generation," Proc. AAAI, 2025.

[1] https://huggingface.co/Qwen/Qwen2-VL-7B-Instruct