

先端科学技術研究科 修士論文要旨

所属研究室 (主指導教員)	自然言語処理学 (渡辺 太郎 (教授))		
学籍番号	2211141	提出日	令和 6年 1月 18日
学生氏名	四條 光		
論文題目	画像ベースとテキストベースのモデルを用いた表の構造解析の性能検証		
要旨			
<p>表は科学論文, Webサイト, 新聞など様々な媒体に現れるため, 表を解析することは膨大な文書を管理するために重要である. 表の構造解析を解くために, 深層学習ベースの画像エンコーダとテキストデコーダから構成されるモデルが考案され, 非常に高い精度を達成している. さらに近年エンコーダに画像とOCRから得られたテキスト, 座標を入力とするマルチモーダルエンコーダを使用するモデルが登場した. こうした背景から画像の特徴量やテキストの特徴量が表の構造解析に及ぼす影響を調査することは, 今後の表の構造解析の発展において必要である. 本研究では, 表の構造解析のエンコーダ・デコーダモデルを構築し, エンコーダに画像ベース, テキストベース, マルチモーダルの3つの異なるエンコーダを使用し, 表の構造解析のスコアを比較することで, どの特徴量を扱うモデルが優れているのかの比較を行った. 実験の結果, 画像ベースのアプローチが最も有効であると示唆された.</p>			