

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	2211151	提出日	令和 6年 1月 18日
学生氏名	白井 尚登		
論文題目	Scalar Mixing Weightsを用いた視覚と言語の情報を事前学習したモデルの言語知識に関する分析		
要旨			
<p>視覚と言語の情報を大規模に学習した機械学習モデルが次々に提案され、画像情報に対する質問応答などのマルチモーダルタスクで成功を収めている。このようなVision and Language (V&L) モデルは大規模言語モデルを拡張したものが多く、一方、画像情報も追加で学習することで本来の言語知識を忘却する可能性が問題となる。そのため、本研究では事前学習済みの言語モデルとV&Lモデルを対象に言語に関する分類タスクと生成タスクの精度比較を行い、言語知識の忘却の有無を調査した。また、画像情報の追加学習によるエンコーダの獲得情報の変化を調査するため、Scalar Mixing Weightsと呼ばれる分析手法を採用した。このScalar Mixing Weightsを用いることで事前学習時に隠れ層で獲得した言語知識がタスクごとに如何に寄与するかを調査し、タスクごとのエンコーダの各層の重要度を定量化した。その結果、分類タスクの分析では画像情報も学習することによる精度の低下やモデル構造の違いによる言語知識への理解の差が明らかとなった。また、生成タスクにおいても画像情報も学習することが精度の悪化に繋がる結果となった。そして、Scalar Mixing Weightsによる調査からタスクごとの重要な層の特定や構造の近いモデル同士でのタスクに寄与する層の類似性が確認された。</p>			