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	2111434	提出日	令和 5年 7月 24日
学生氏名	VINCENT MICHAEL SUTANTO		
論文題目	Investigating the Effects of Pruning White Patches on Document-Related Tasks		

要旨

The Vision-Transformer (ViT) has emerged as a viable alternative to Convolutional Neural Networks (CNNs) and has also been applied in document-related tasks. However, ViT inherits Transformer's quadratic complexity, which will become a problem in current multi-modal settings. Researchers have explored various patch-pruning techniques, which involve discarding certain parts of inputs, to address the complexity. Even so, it is essential to be cautious when applying patch-pruning methods to document-related tasks, as indiscriminately dropping token patches can lead to information loss and degrade model performance. Nevertheless, the idea of pruning white parts of a document, which are typically devoid of meaningful information, suggests that White-patch Pruning (WPP) could lead to improved efficiency in addressing document-related tasks. This thesis investigates a computationally efficient ViT by leveraging WPP to solve document-related tasks: Document Image Classification, Document Entity Extraction, and Mathematical Equation Extraction. The results are threefold: Firstly, compared to the standard ViT, WPP-ViT had minimal to negligible impact on the model's performance across various evaluation metrics. Secondly, as the WPP-ViT is built upon the same architecture as ViT, it allows for the seamless transfer of pre-trained weights, ensuring the inherited knowledge is preserved. Lastly, the WPP-ViT benefits of higher throughput rate in 6 of 8 datasets used, thereby benefiting from improved efficiency in document-related tasks.