

GIA TĂNG DỮ LIỆU CHO SELF-ATTENTION, CÁCH NÀO LÀ CÓ ÍCH CHO HIỆU QUẢ CỦA MÔ HÌNH?

Nguyễn Văn Dũng - 19521400

Lâm Thành Tín - 19521020

Huỳnh Phương Như - 19520208

Tóm tắt

- Lớp: CS519.M11.KHCL
- Link Github của nhóm: <https://github.com/lialic/CS519.M11.KHCL>
- Thành viên nhóm:



Lâm Thành Tín



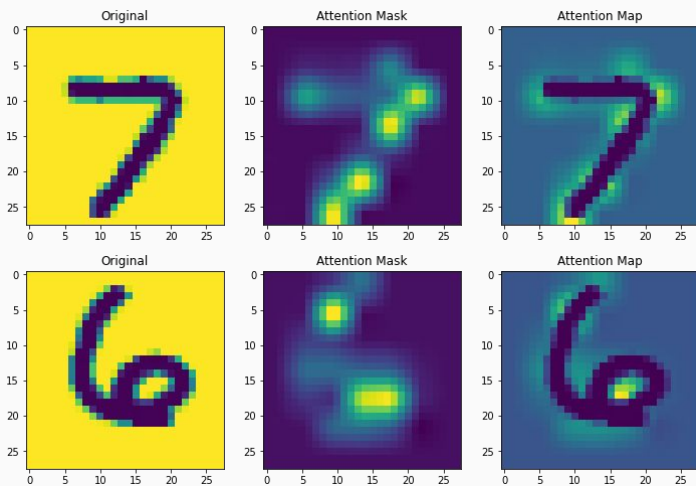
Huỳnh Phương Như



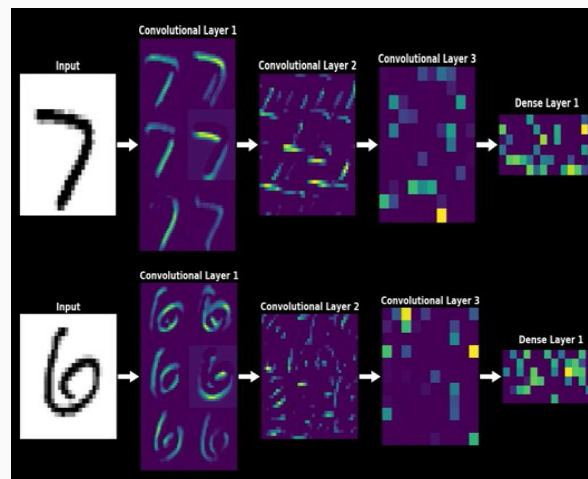
Nguyễn Văn Dũng

Giới thiệu

- Tăng cường dữ liệu là các phương pháp quen thuộc để tăng hiệu quả của mô hình, nhưng khi áp dụng cùng một phương pháp tăng cường từ CNN lên mô hình có self-attention thì có còn phù hợp?



MNIST qua Vision Transformer



MNIST qua CNN

Mục tiêu

- Đề xuất các kỹ thuật gia tăng dữ liệu và kết hợp các kỹ thuật phù hợp cho những bộ dữ liệu mang tính đặc trưng.
- Giúp gia tăng hiệu quả mô hình self-attention khi sử dụng bộ dữ liệu vừa và nhỏ.
- Giúp phát triển mô hình sử dụng self-attention thành SOTA và phổ biến hơn CNN cho tất cả kích thước tập dữ liệu.

Nội dung và Phương pháp

Nội dung:

- Nghiên cứu các phương pháp tăng cường dữ liệu hiện có.
- Nghiên cứu các dạng phân bố vật thể phổ biến và tìm bộ dữ liệu đặc trưng đại diện cho từng dạng.
- Nghiên cứu và thực nghiệm các phương pháp tăng cường dữ liệu lên các bộ dữ liệu đã nghiên cứu.
- So sánh hiệu quả và phân tích sự khác nhau giữa các kỹ thuật gia tăng dữ liệu trên các dữ liệu có phân bố vật thể khác nhau.
- Sử dụng phương pháp tăng cường tốt nhất cho từng bộ dữ liệu và thực hiện việc huấn luyện mô hình.

Nội dung và Phương pháp

Phương pháp:

- Tìm hiểu các phương pháp gia tăng dữ liệu.
- Tìm hiểu các cách đánh giá kỹ thuật gia tăng dữ liệu.
- So sánh các kỹ thuật dựa trên thước đo đánh giá và thực nghiệm.
- Tìm hiểu các dạng phân bố vật thể trong các bộ dữ liệu phổ biến.
- Tìm hiểu các mô hình sử dụng cơ chế self-attention.
- Lập bảng tổ hợp các kỹ thuật gia tăng dữ liệu thích hợp.
- So sánh đánh giá các tổ hợp kỹ thuật tăng cường dữ liệu bằng thực nghiệm.

Kết quả dự kiến

- Báo cáo lý thuyết và thực nghiệm các phương pháp tăng cường dữ liệu.
- Báo cáo lý thuyết và thực nghiệm các cách kết hợp các phương pháp tăng cường dữ liệu phù hợp với từng bộ dữ liệu và phân bố vật thể của chúng.
- Thu được kết quả thực nghiệm, so sánh đánh giá các kỹ thuật.
- Tìm được mối liên hệ giữa các kỹ thuật với phân bố vật thể của dữ liệu.
- Tìm được các phương pháp tăng cường dữ liệu phù hợp và tối ưu cho mô hình sử dụng self-attention.
- Công bố bài báo khoa học.

Tài liệu tham khảo

- [1]. Dai, Zihang, et al. “[2106.04803] CoAtNet: Marrying Convolution and Attention for All Data Sizes.” *arXiv*, 9 June 2021, <https://arxiv.org/abs/2106.04803>.
- [2]. Dosovitskiy, Alexey, et al. “[2010.11929] An Image is Worth 16x16 Words: Transformers for Image Recognition at Scale.” *arXiv*, 22 October 2020, <https://arxiv.org/abs/2010.11929>.
- [3]. Raghu, Maithra, et al. “[2108.08810] Do Vision Transformers See Like Convolutional Neural Networks?” *arXiv*, 19 August 2021, <https://arxiv.org/abs/2108.08810>.
- [4]. Steiner, Andreas, et al. “[2106.10270] How to train your ViT? Data, Augmentation, and Regularization in Vision Transformers.” *arXiv*, 18 June 2021, <https://arxiv.org/abs/2106.10270>.
- [5]. Zhai, Xiaohua, et al. “[2106.04560] Scaling Vision Transformers.” *arXiv*, 8 June 2021, <https://arxiv.org/abs/2106.04560>.