An Image is Worth 16x16 Words: Transformers for Image Recognition at Scale.

Paper review done by Sankalp K. Jajee

Paper Summary:

The paper introduces the Vision Transformer (ViT), a transformative model for image recognition, challenging the conventional use of Convolutional Neural Networks (CNNs). ViT uses a unique approach, utilizing the Transformer architecture for image handling, demonstrating superior performance in visual recognition.

Key contributions:

As described in the slides, the key contributions of the paper are:

Transformer Architecture for Images

Image to Sequence conversion

Scaling with data

Empirical Results

Computational Efficiency

Inductive Bias Analysis

Potential for Self-Supervision

Challenges and future work

Those keys overall highlight a special tool with pros and cons. Showing an alternative to Convolutional Neural Networks is great as to open more methods of approach for future explorations of the system.

As for my critiques, one argument used often in the paper was the efficiency of ViT on large datasets. Since this argument is brought out multiple times, defining the term “larger” would be insightful. For example, in the sentence: “While CNNs inherently have biases like translation invariance and locality, ViT needs to learn these from data, which it does effectively given sufficient training data.”. When comparing/showing how ViT can outperform other AI, using the argument of “it needs more data” feels stretched since most AI model is learning based, meaning most of them can outperform other models with enough data compared.

Finally, as said, it’s dependency on larger database and lack of inductive bias makes ViT more of a specialized tool. This information provides context on ViT's applicability and potential use cases for larger organizations.

Towards the ultimate brain: Exploring scientific discovery with ChatGPT AI

By Erin McCormack.

Paper Summary:

This paper tests the uses of AI chatbot chat GPT, in science. That is, how well can Chat GPT assist in creating new theories, observations, and calculations. For this paper, a combination of AI GPT and physics GPT were added together to make “GPT^4”. With this chatbot, various tests/prompts were asked to experiment with this chatbot, and the results, despite being useful, we not reliable enough to allow GPT^4 to come up with scientific innovations on its own.

Key contributions:

The key contributions of this paper, are:

* Creating it’s a unique environment:   
  By prompting GPT to imagine a world where the rules are made and selected by the AI, this section of the paper starts with a strong request that seem impossible to achieve. But this also showcase the sort of challenging questions that scientists may have in certain fields.
* Coming up with theories:  
  In this portion of the paper, GPT simply defined theories instead of designing some. So, this does show some level of inaccuracy in the response since GPT will not always be reliable in its answers.
* Testing criteria and completing them.

This part mostly showcases GPT^4 and its ability to reply to scientific criteria with a professional reply.

Your critique:

As mentioned on the slides, since only the successful prompts were shown, we are missing an important part of the paper, reliability. One of the more important aspects of using AI in science is to make sure that the output is meaningful. So, not having a section to discuss hurts the overall idea of the paper.