

A literature review on the Visual Spatial Reasoning of Large Language Models

Abstract—Large Language Models has become a hot topic in the field of Natural Language Processing. The Visual Spatial Reasoning is a key ability of human intelligence. In this paper, we will review the recent research on the Visual Spatial Reasoning of Large Language Models.

I. Introduction

In recent years, Large Language Models(LLMs) has become a hot topic in the field of Natural Language Processing. In general research, Visual Spatatial Reasoning is mostly used on Visual Question Answering(VQA) and Visual Commonsense Reasoning(VCR), and Most of the research is testing on the Visual Language Models(VLMs),