

Table of Contents

- 1 Background
- 2 Introduction
- 3 Vision language tasks and variants
- 4 Current approach to VLM
- 5 SOTA models
- 6 Real world applications
- 7 Future research directions
- 8 Conclusion