

Small Vision-Language Model (SVLM) for Dense Video Captioning

Author
Marija Brkić

Abstract

Write here

1 Introduction

The reference apperas only when you cite it like
so: [1]

2 Related Work

3 Contribution and Methodology

4 Experiments

5 Results and Evaluation

6 Conclusion

References

- [1] Andrés Marafioti et al. “SmolVLM: Redefining Small and Efficient Multimodal Models”. In: *arXiv preprint arXiv:2504.05299* (2025). URL: <https://arxiv.org/abs/2504.05299> (cit. on p. 1).