Ruben: A multi-modal AI music composition system

David Korcak, Frantisek Kmjec

February 8, 2026

Abstract

This report presents Ruben, a multi-modal AI music composition system that integrates text, audio, and visual inputs to generate original musical pieces. We describe the architecture of Ruben, its working principles, and the key infrastructural and tooling innovations it introduces.

1 Introduction

2 Methods

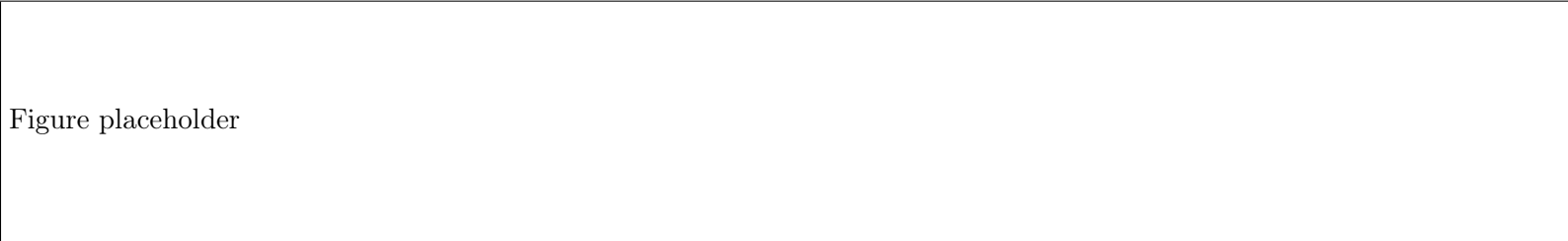Describe the approach and evaluation setup.

3 Results

Present key results.

Figure 1: Example figure.

4 Conclusion

Summary and future work.

5 References

References

[1] J. Smith, *Example Reference*, 2019.