

# 11-777 Final Report

**First Last 1\***    **First Last 2\***    **First Last 3\***    **First Last 4\***  
{ID1, ID2, ID3, ID4}@andrew.cmu.edu

## Abstract

In this project we introduce ... to tackle... we  
show that ...

## 1 Introduction and Problem Definition (1-1.25 pages)

This will be about 1 page, might include a figure  
on the top right and lays out the contributions.

---

\*Everyone Contributed Equally – Alphabetical order

## **2 Related Work and Background (5 papers per person)**

**Related Datasets**

**Unimodal Baselines**

**Prior Work**

**Relevant techniques**

### **3 Task Setup and Data**

Integrate your previous report here

## **4 Baselines**

Integrate your previous report here

## **5 Proposed Model (>1 page)**

This is not a copy and paste from R2, it should be updated to reflect the final proposed model. Explain if there are multiple versions (e.g. your model has possible extensions, different amounts of exemplars, etc). Please also include an easy to read model diagram here.

### **5.1 Loss functions**

### **5.2 Changes to training data**

### **5.3 Hyperparameters and their effects**

Methods	Dev	
	Accuracy $\uparrow$	$L_2$ Error $\downarrow$
Unimodal 1 ()		
Unimodal 2 ()		
Unimodal 3 ()		
Simple Multimodal 1 ()		
Simple Multimodal 2 ()		
Simple Multimodal 3 ()		
Previous Approach 1 ()		
Previous Approach 2 ()		
Previous Approach 3 ()		
Proposed Method (v1)		
Proposed Method (v2)		
Proposed Method (final)		

## 6 Results (1 page)

Replace columns with the correct metrics for your task (extrinsic). Include multiple versions of your final model. You do not need to run on the test set but are encouraged to try if you have nice results on Dev.

Methods	Dev	
	Detections $\uparrow$	Invalid Actions $\downarrow$
Unimodal 1 ()		
Unimodal 2 ()		
Unimodal 3 ()		
Simple Multimodal 1 ()		
Simple Multimodal 2 ()		
Simple Multimodal 3 ()		
Previous Approach 1 ()		
Previous Approach 2 ()		
Previous Approach 3 ()		
Proposed Method		

## 7 Analysis (2 pages)

This section should include plots. For example, how key metrics vary with a specific hyperparameter, task complexity, etc.

### 7.1 Intrinsic Metrics

These are not the task itself (they might overlap with auxiliary losses) but are skills the model should have.

## **7.2 Qualitative Analysis and Examples (full page tables – multiple pages for most projects)**

A full page table that includes: Example input, comparison of model outputs, theory about reason for failure. Every model presented should have at least one failure condition. Please visualize the data (include images, etc)



## **8 Future work and Limitations (1 page)**

Please thoroughly discuss the limitations of the proposed model. Where are all the places it does not work, phenomena it does not capture, places that ideally it can be improved. As guidance, there should be more things listed here than members of the team.

## **9 Ethical Concerns and Considerations (unintentional, malicious, and dual-use)**

## **10 Team member contributions**

**Member 1** contributed ...

**Member 2** contributed ...

... contributed ...