Project Report Draft

Part 1

* We wanted to test how the combination of film layer and cbn layer affects the performance of film model on the clever data set

Film is a general-purpose conditioning method that is highly effective for visual reasoning, however it makes some logical mistakes that humans won’t do, for example: (example from film paper).

We introduce a model that uses the film model but instead of batch normalization we use a cbn layer from (paper name/link). Our goal is to check if the new model we built can produce better results than film (add about fixing the logical mistakes film makes).

(add final results)

Part 2

* Briefly explain about the film model and cbn model

(take from film paper and cbn paper)

* Explain why we decided to combine these 2 models

Both models produce similar results but have different spatial reasoning for some of the relations(maybe change), for example: film keeps all query relations close however, cbn query color is far from the other query types, hence we believe that combining these models can create a model that benefits from both and producing better results.

* Problem domain – improve film model performance (accuracy wise)

Part 3

* Explain the implementation we added

1. Added cbn layer
2. Added new configuration flags
3. Added parallelism to the GRU model

* Explain why we changed some of the parameters

1. Changed data representation from int64 to int32
2. Changed GRU dimensions so it will return twice the amount of weights (used for cbn + film)
3. Changed dropout (for tests)

Part 4

* Describe the tests we ran to see the model performance (like film paper)

Part 5

* Represent the results in graphs and tables