RELATIONAL REASONING

CS337: Artificial Intelligence and Machine Learning

KARAN GODARA (210050082)

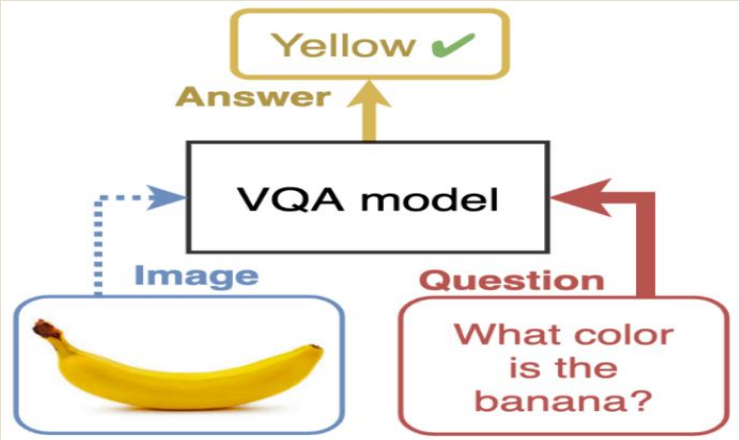
ISHA ARORA (210050070)

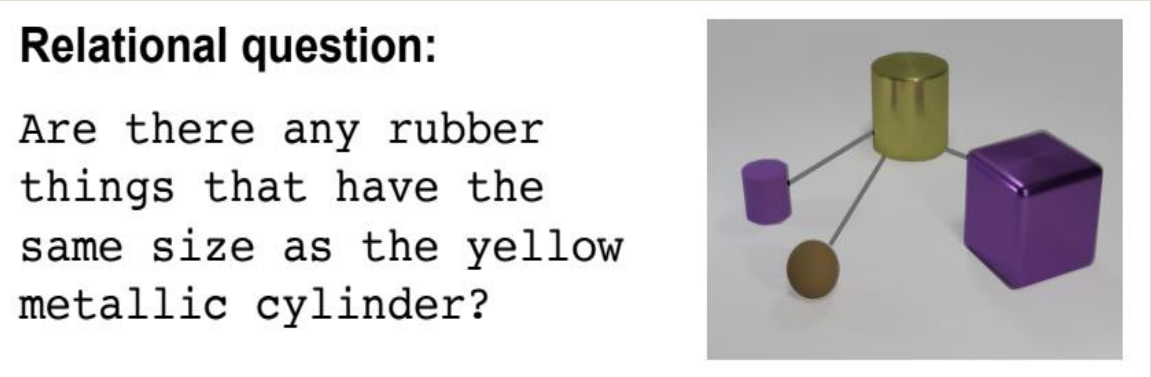
SHANTANU WELLING (210010076)

DHANANJAY RAMAN (210050044)

Project Overview

- ***Visual Question Answering (VQA)*** : ML task to understand natural language questions about an image and provide accurate answers.
- ***Relation Network (RN) Model*** : A general solution to flexible ***relational reasoning*** in neural networks





Project Overview

- ***Our Experiments*** :
  - Conducted experiments on both CLEVR and Sort-of-CLEVR datasets
  - Explored relation networks, focusing on their flexibility through the incorporation of various descriptions of 'objects'
  - Enhanced the model's performance on Sort-of-CLEVR dataset

Why previous approaches fail?

- Perform well on non-relational questions but perform below par in the relational question domain.
- Symbolic approach to AI is inherently relational and not robust to task and input variations.
- Relational question domain demands mechanisms to compute relations between a set of objects.
- ***CNN*** – capacity to reason about spatial, translation invariant properties
- ***RNN*** - capacity to reason about sequential dependencies
- Now let's explore RN architecture which has the capacity to compute relations baked into them.

Relation Network

- ***Design philosophy*** - constrain the functional form of a neural network so that it captures the core common properties of relational reasoning

$$\mathrm{RN}(O) = f_\phi\left(\sum_{i,j} g_\theta(o_i, o_j)\right)$$

- input is a set of "objects" $\boldsymbol{O = \{o_1, o_2, \ldots, o_n\}}$, $\boldsymbol{o_i \in R^m}$ is the i^{th} object, and $\boldsymbol{f_\varphi}$ and $\boldsymbol{g_\theta}$ are functions with parameters φ and θ, respectively.
- In our case, f_φ and g_θ are MLPs.

Strengths of RNs

1

Learn to infer relations: Considers the potential relations between all object pairs

2

Data efficient : Use a single function g_θ to compute each relation – greater generalisation

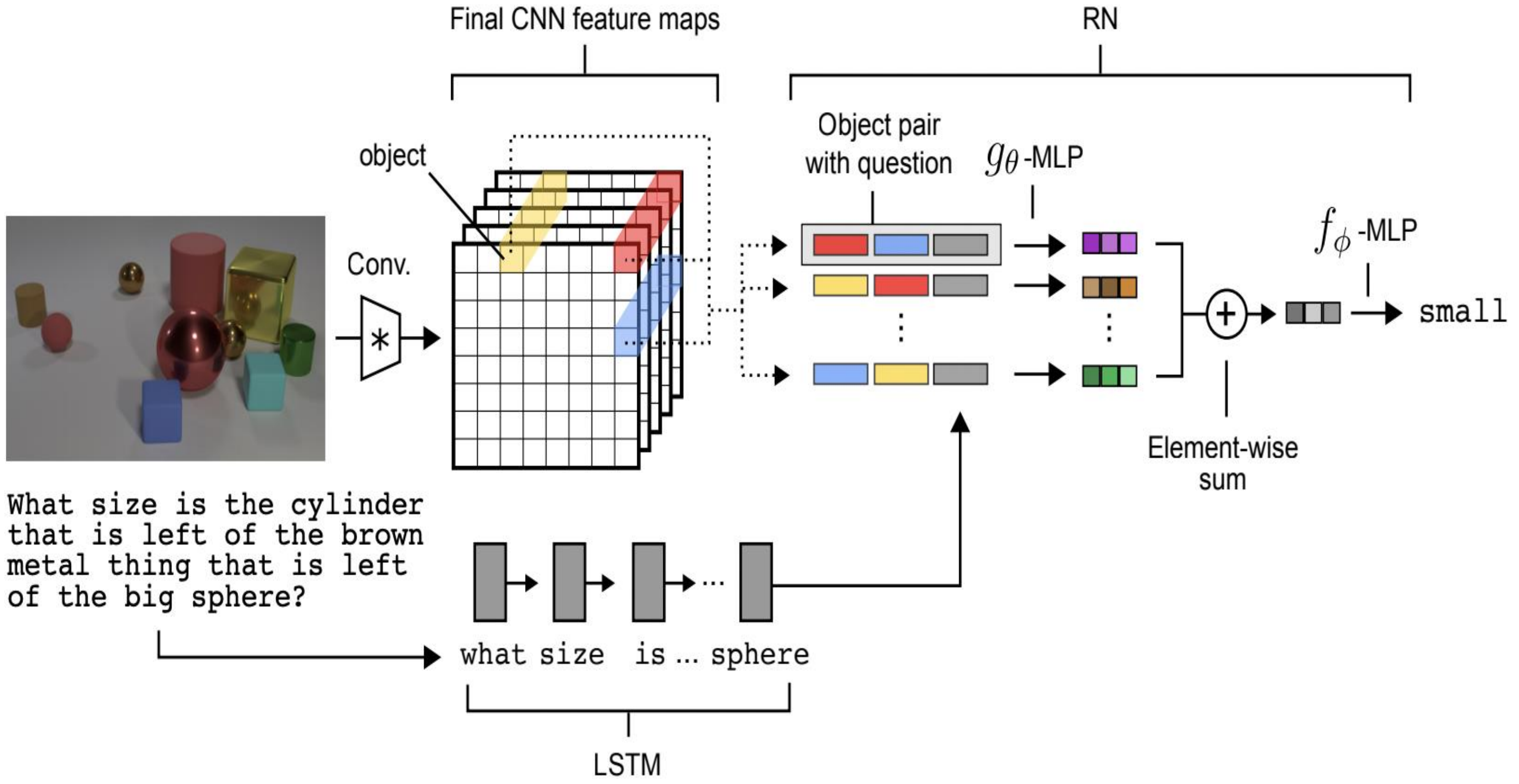
3

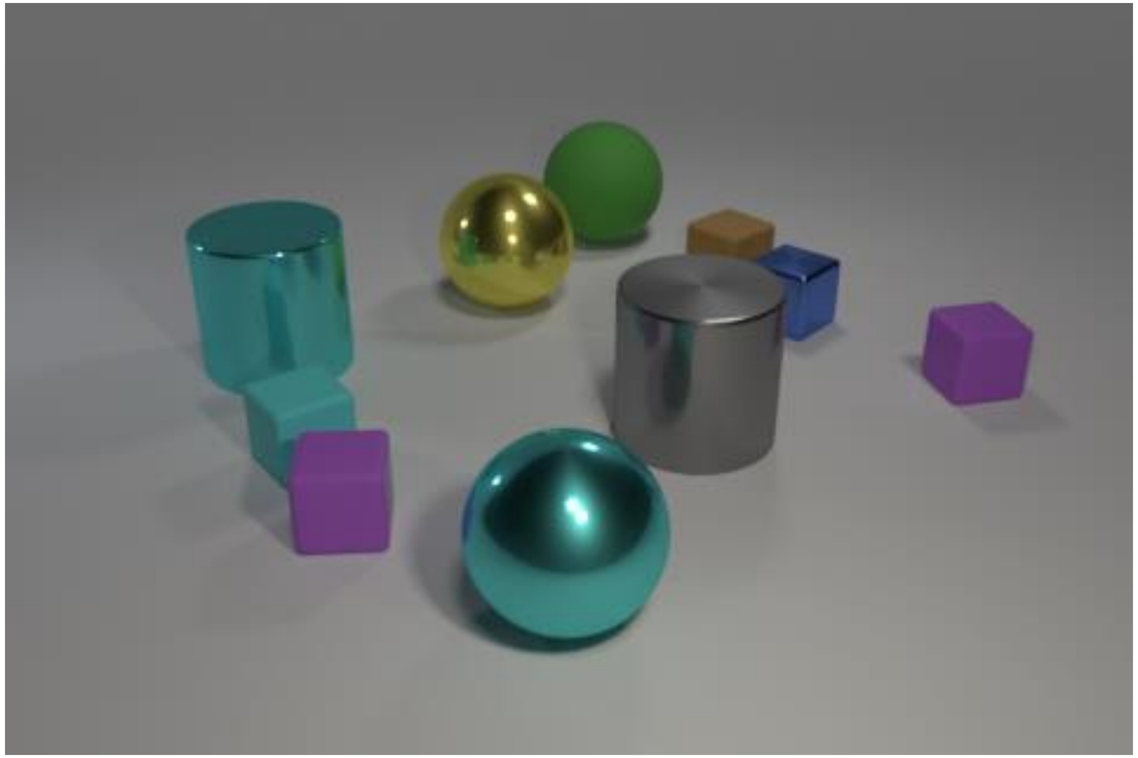
Operate on a set of objects : Invariant to the order of objects in the input

4

Flexible : Do not explicitly operate on images or natural language

Relation Network Architecture for Images





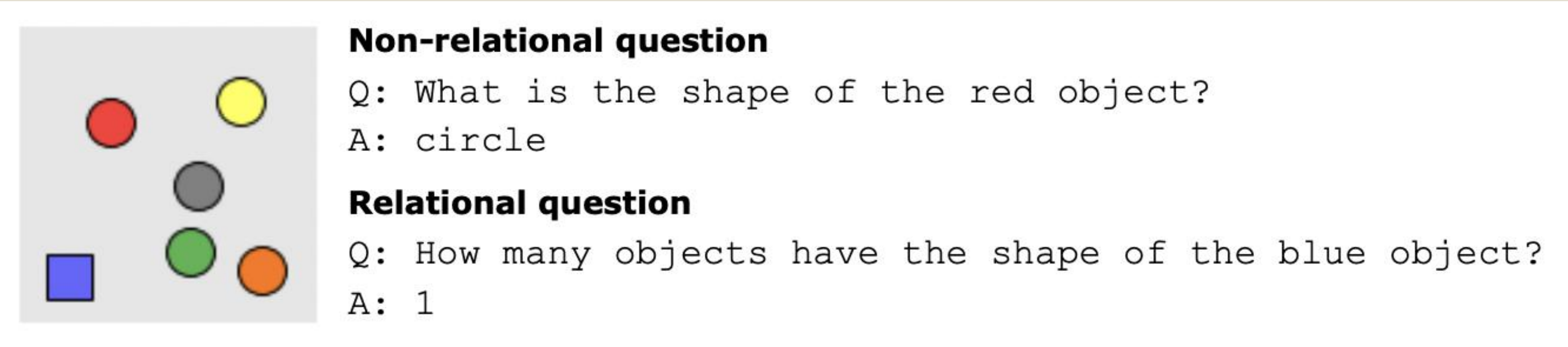
Datasets Employed

CLEVR (Compositional Language and Elementary Visual Reasoning) :

VQA dataset of images of 3D objects on which state-of-the-art approaches have struggled due to the demand for rich relational reasoning – spatial and otherwise

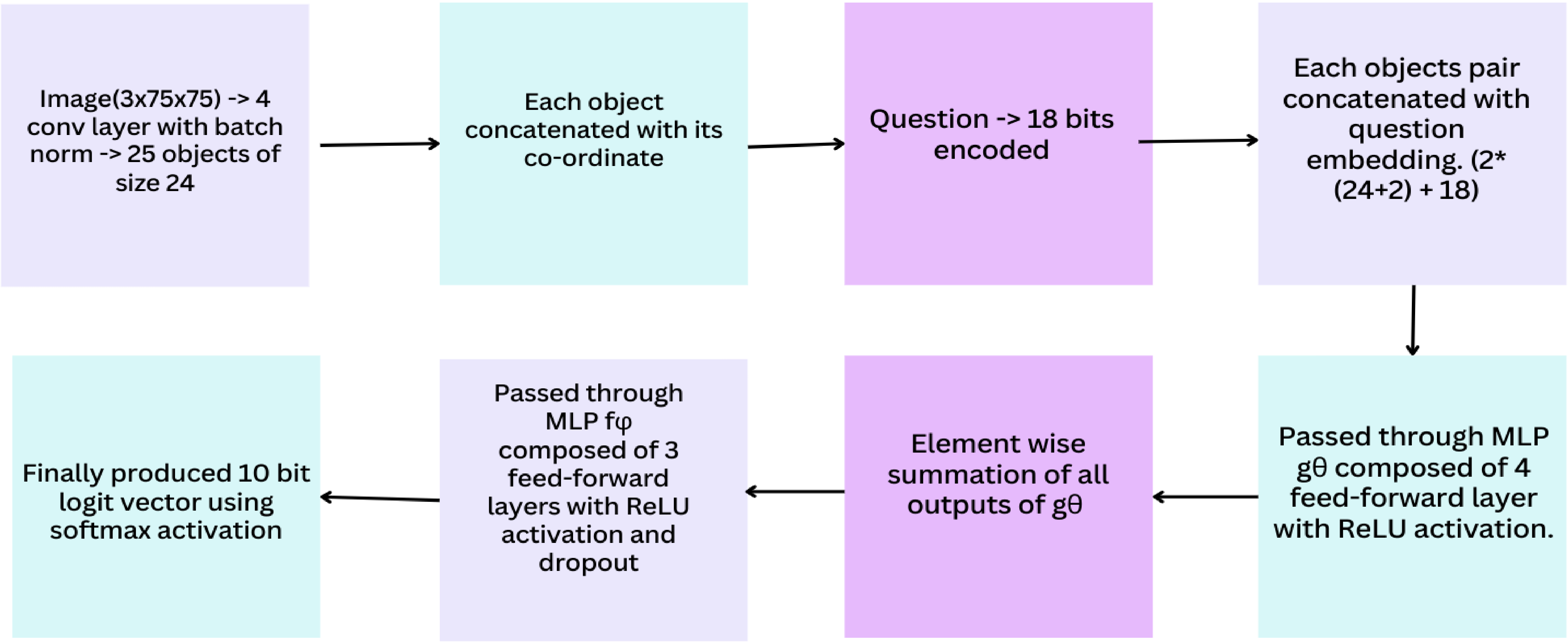
Datasets Employed

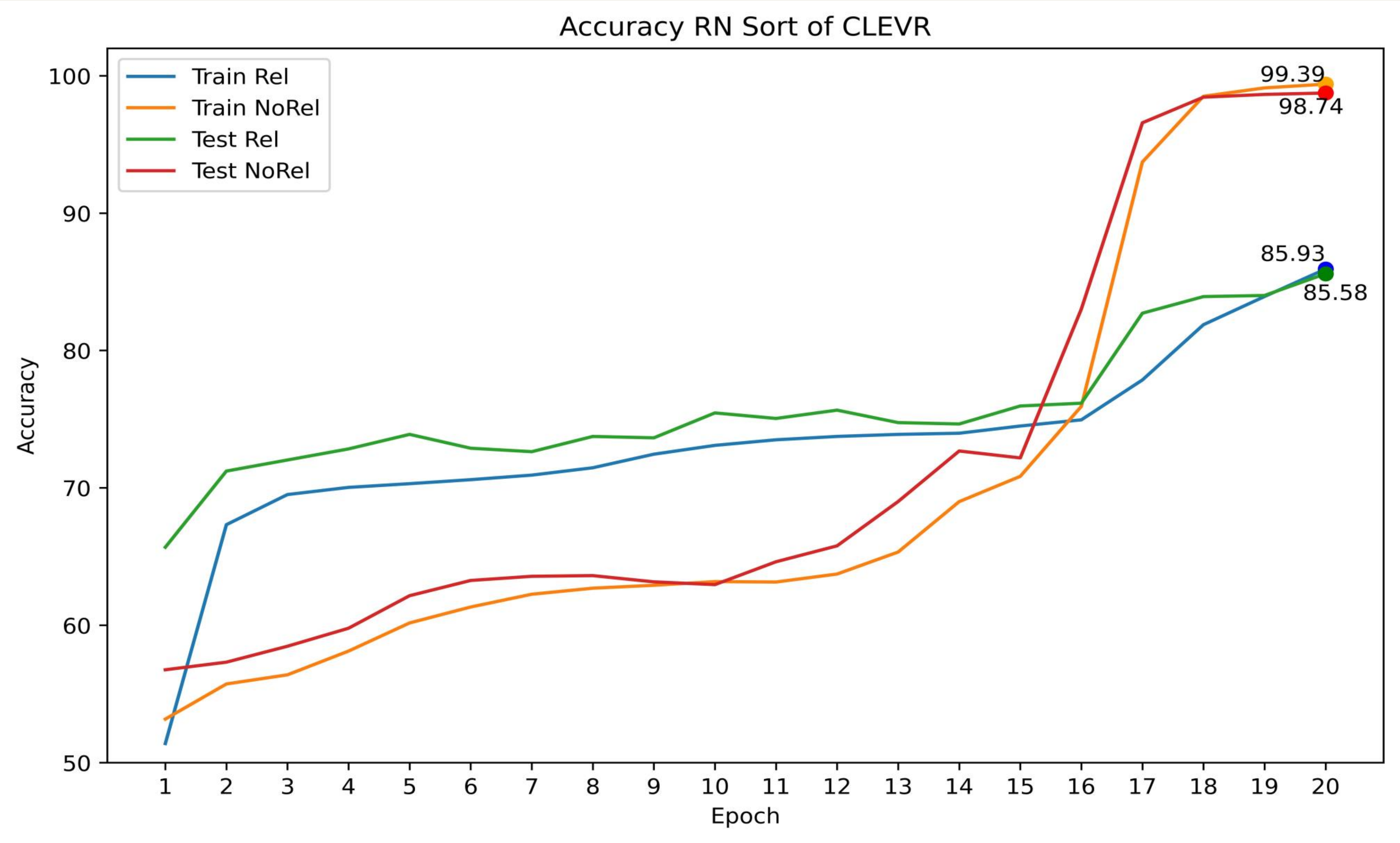
- **Sort-of-CLEVR** : Dataset similar to CLEVR but consists of only 2D color shapes that separates relational and non-relational questions



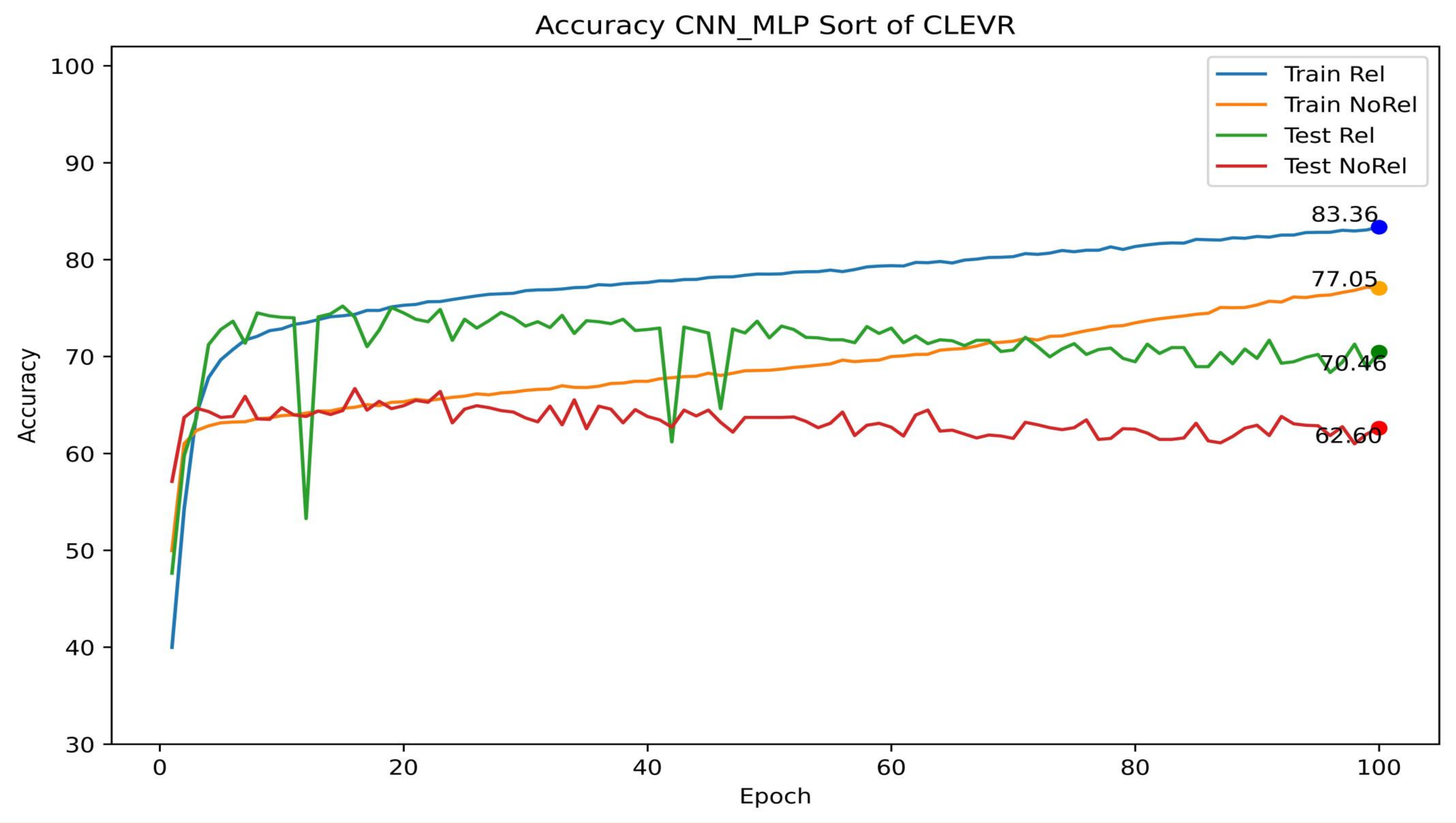
OUR EXPERIMENTS

Sort-of-CLEVR with RN





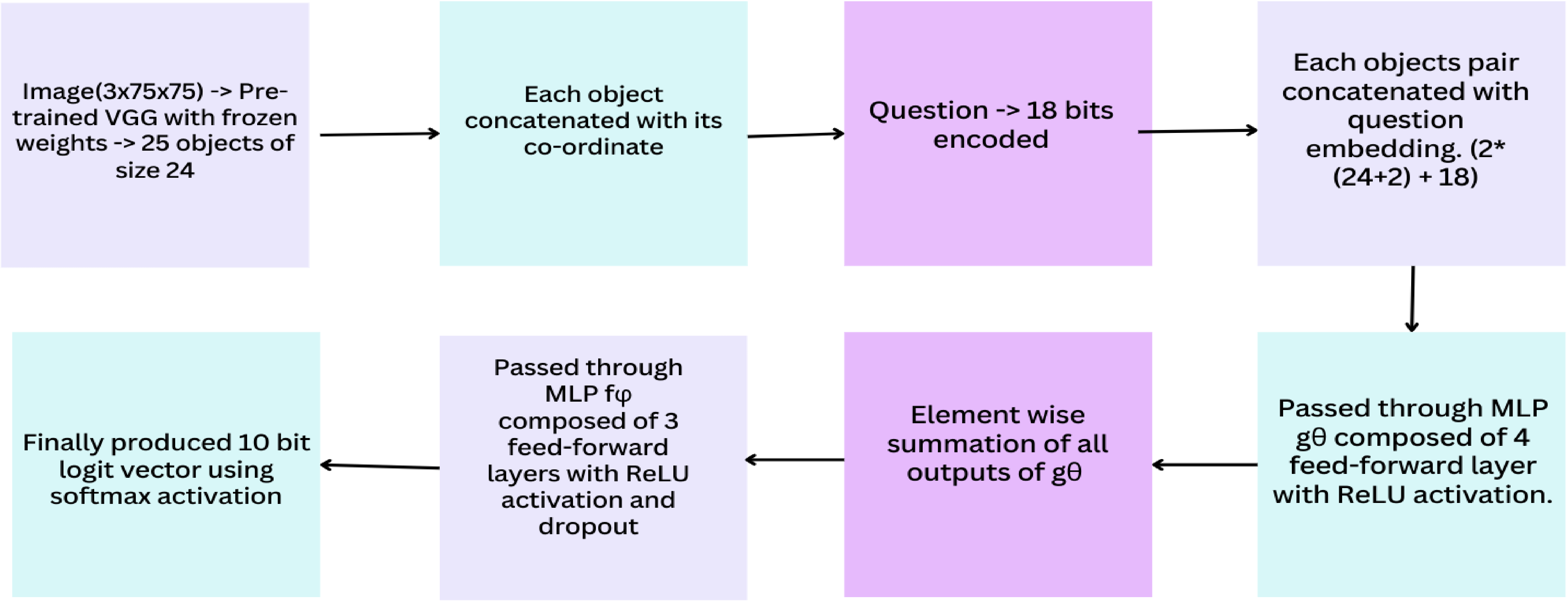
RN on Sort-of-CLEVR data

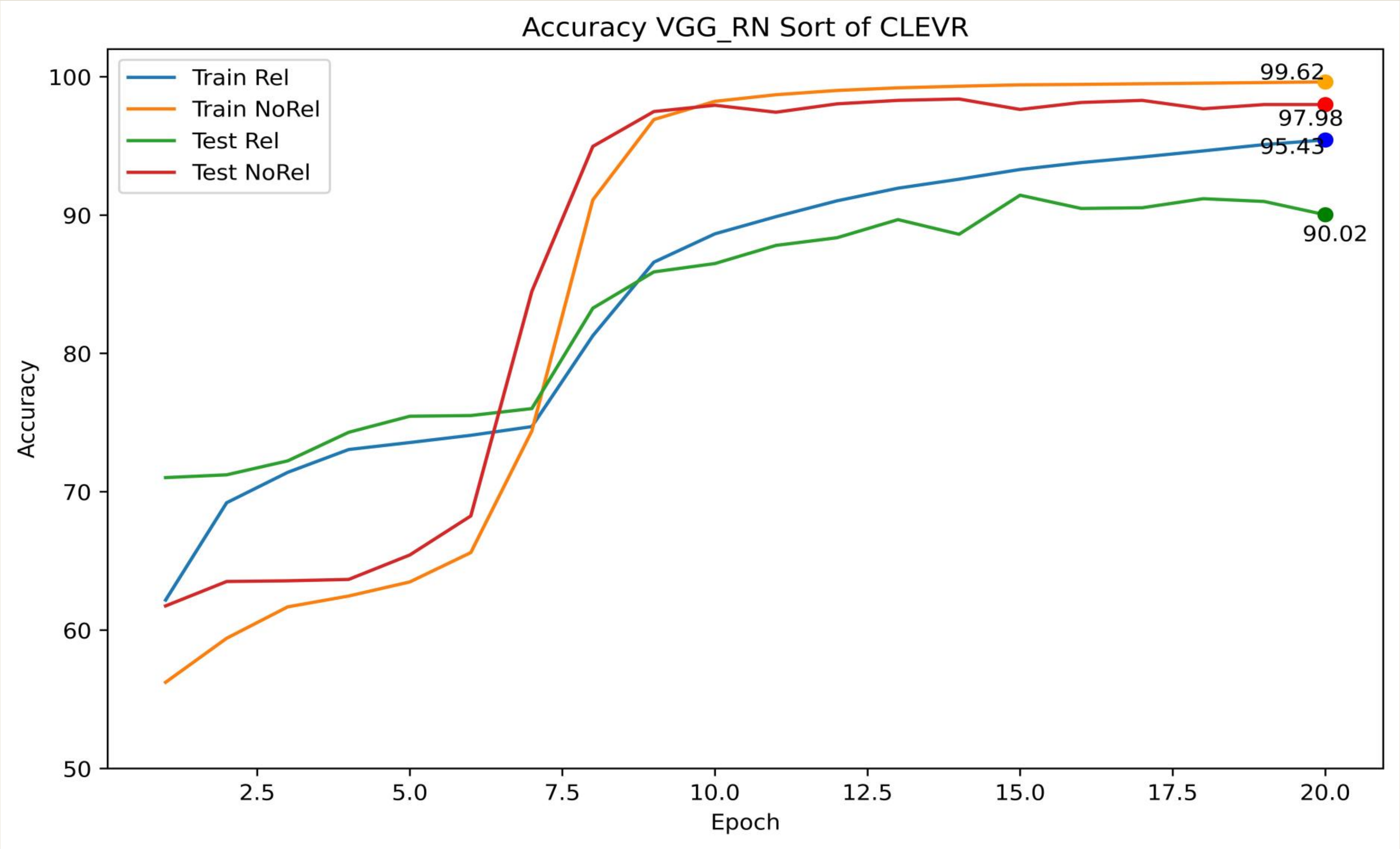


Relation Network Model outperforms the CNN MLP model

Changing object generator CNN

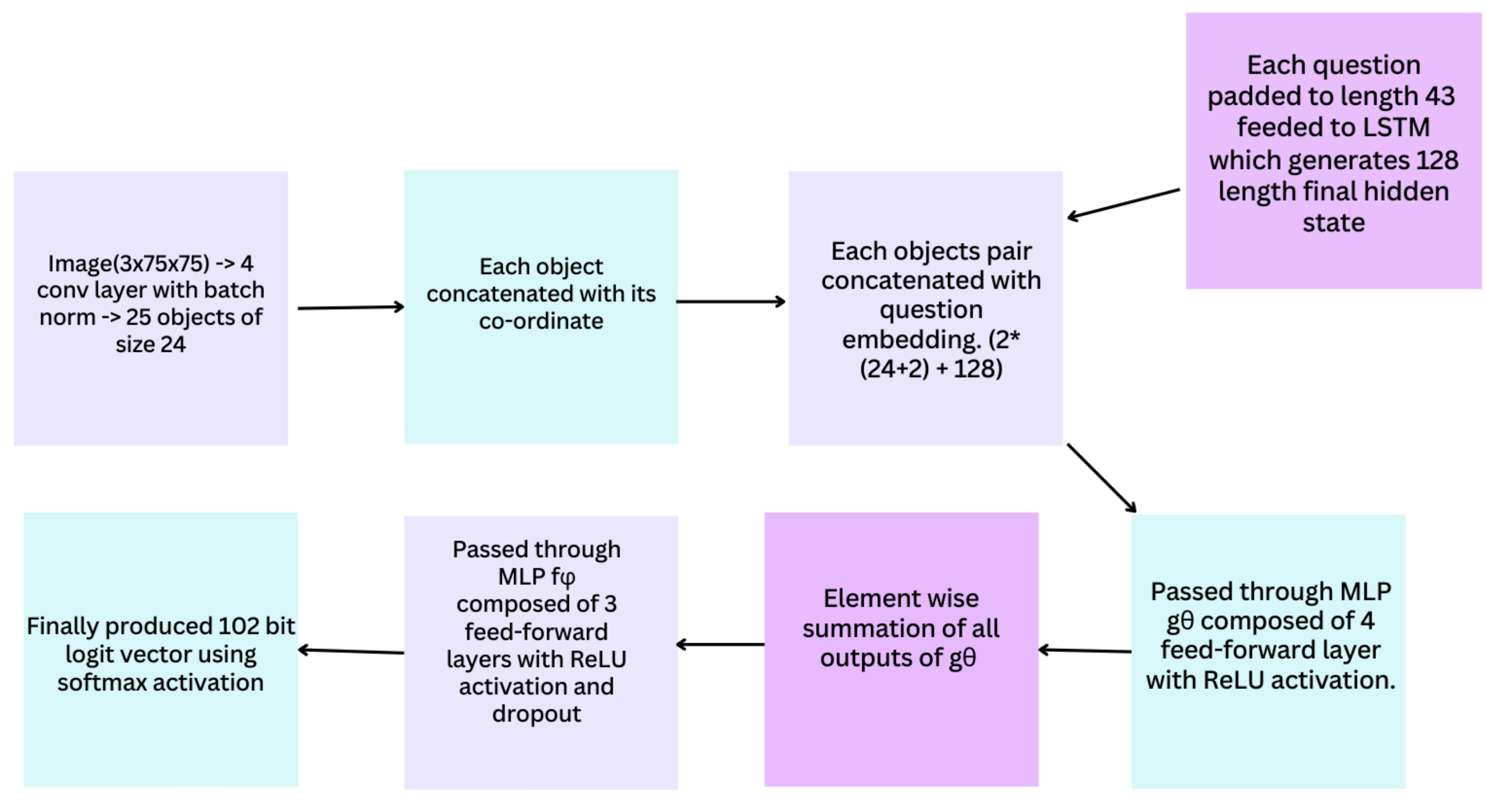
Sort-of CLEVR dataset with pre-trained

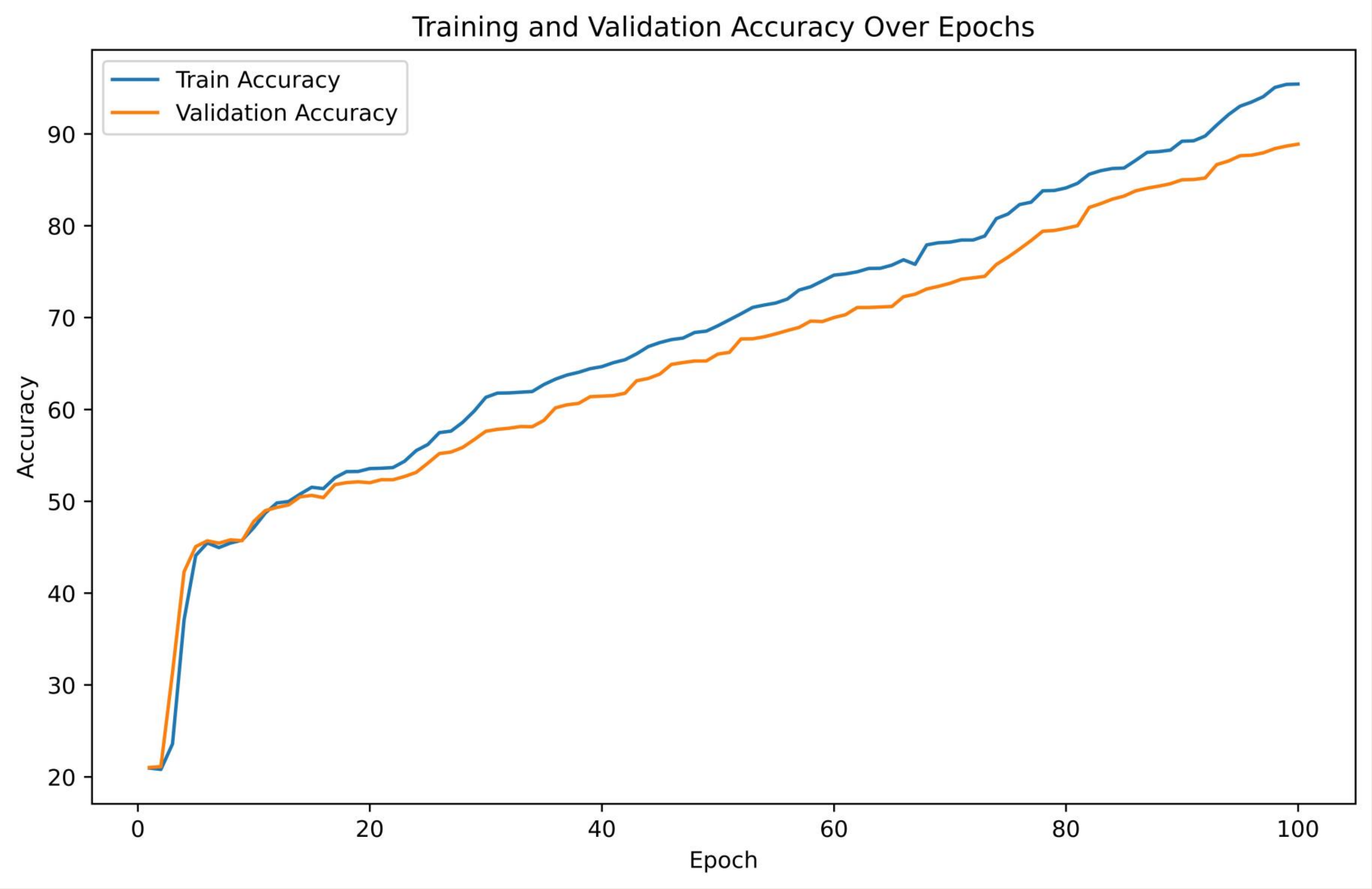




Improvement in performance by using pre-trained VGG based RN on Sort-of-CLEVR Data

CLEVR Dataset





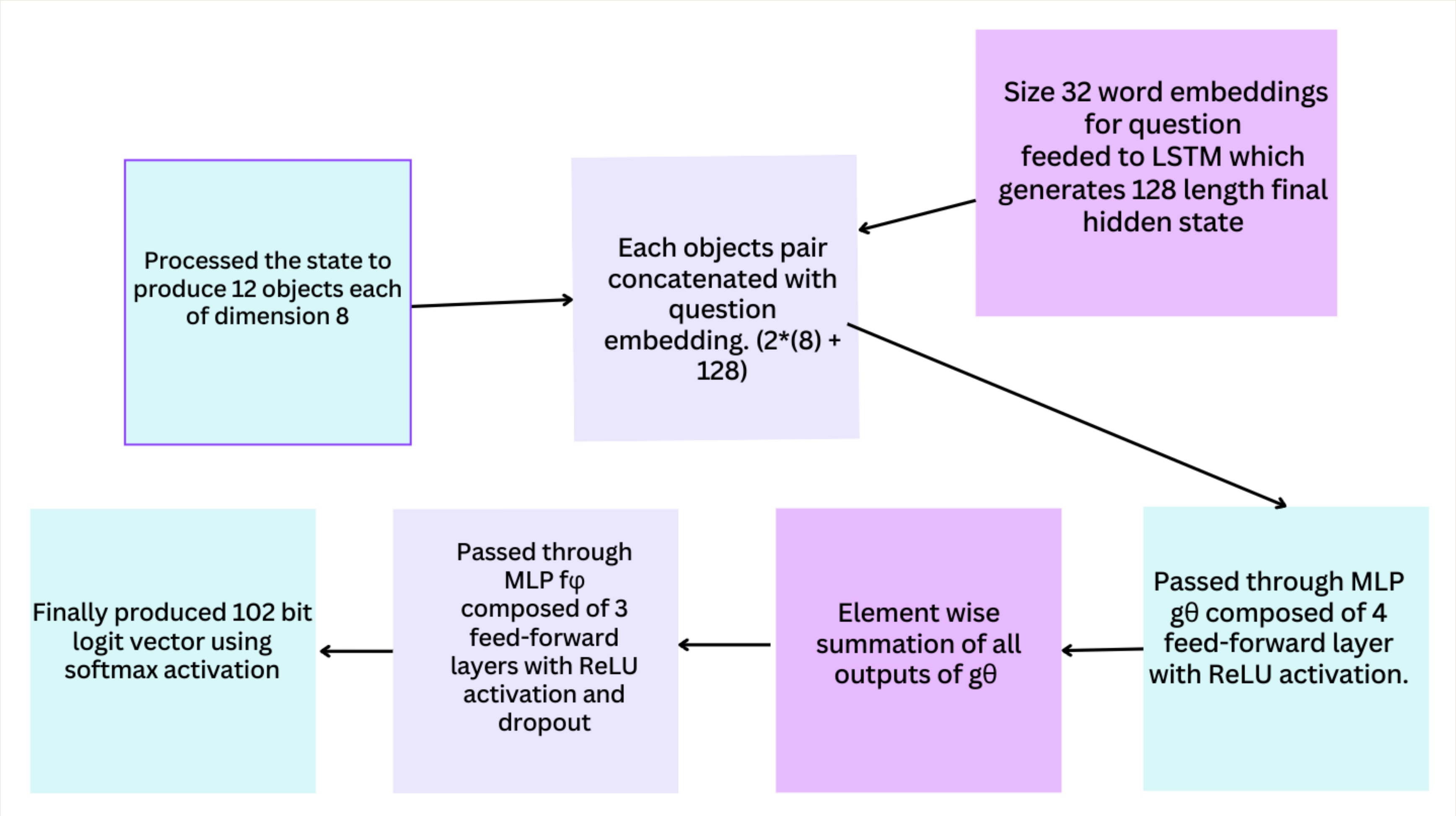
RN on CLEVR Dataset (Pixel Representation)

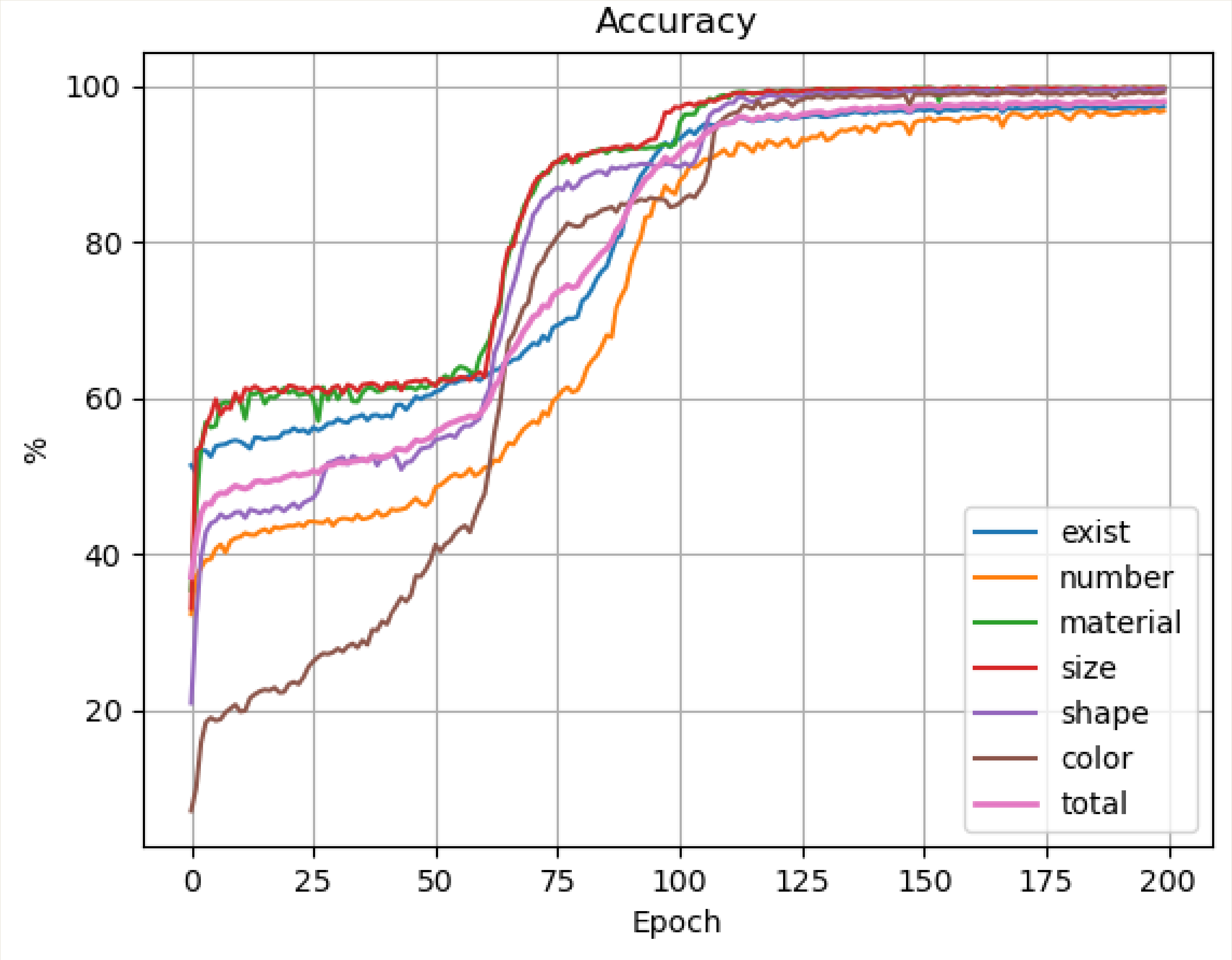
Versatility with regards to form of input

CLEVR Dataset with State Description

- Images explicitly represented by state description matrices containing factored object descriptions
- Each row in the matrix contains the features of a single object
  - 3D coordinates (x, y, z)
  - color (r, g, b)
  - shape (cube, cylinder, etc.)
  - material (rubber, metal, etc.)
  - size (small, large, etc.)

CLEVR Dataset with State Description





RN model on CLEVR Data with state description as input

Further Improvements

Inclusion of ternary relations

Employ deep CNNs with attention to improve object representation

RN definition can be adjusted to consider only some object pairs. RNs can take as input a list of only those pairs that should be considered, made available by some upstream mechanism

References

- https://github.com/kimhc6028/relational-networks
- https://github.com/mesnico/RelationNetworks-CLEVR/tree/master
- https://arxiv.org/pdf/1706.01427.pdf

Contributions

Each individual experiment's codebase was developed majorly by one individual with the assistance of others. Overall, everybody had the same contribution.

Thankyou

KARAN GODARA (210050082)

ISHA ARORA (210050070)

SHANTANU WELLING (210010076)

DHANANJAY RAMAN (210050044)