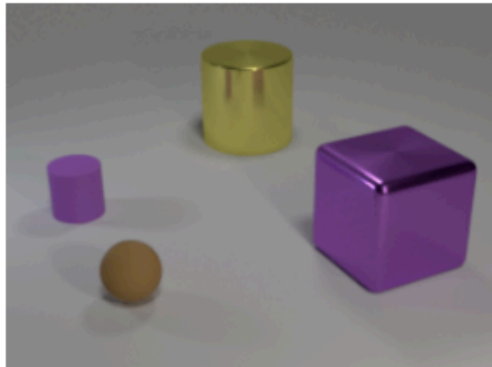


A simple neural network module for relational reasoning



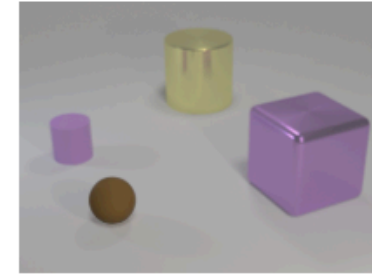
Visual Question Answering

Original Image:



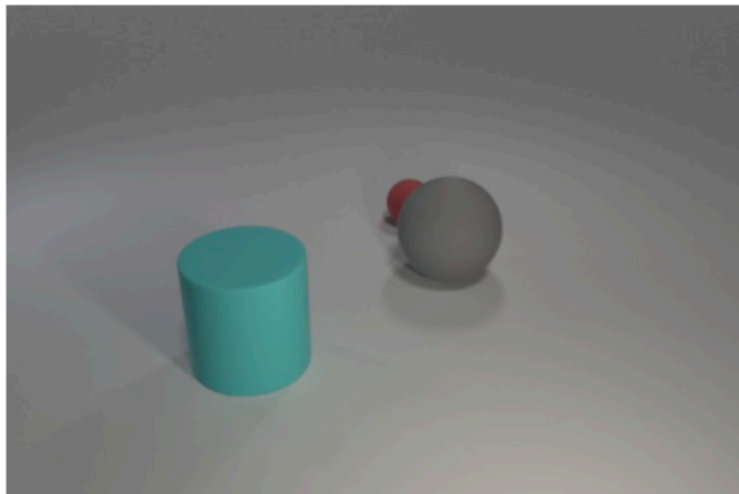
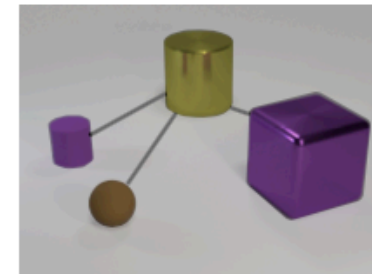
Non-relational question:

What is the size of the brown sphere?

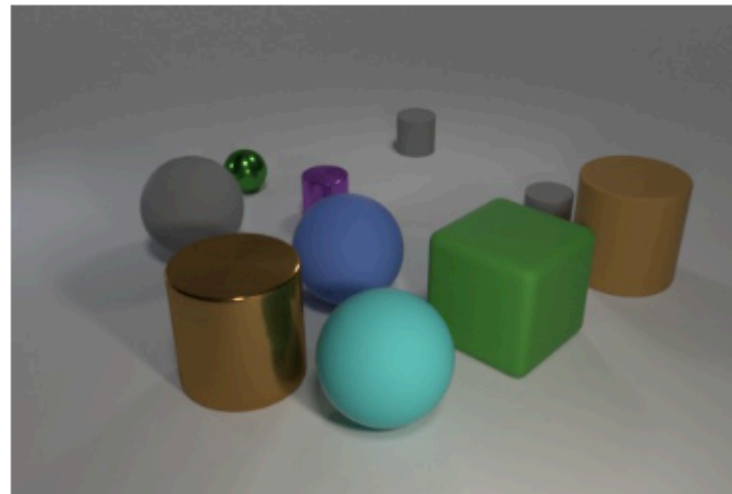


Relational question:

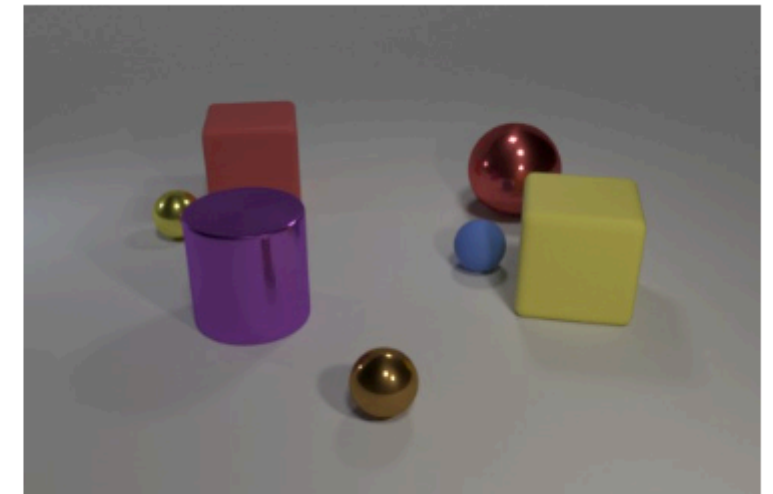
Are there any rubber things that have the same size as the yellow metallic cylinder?



Is the shape of the small red object the same as the large matte object that is right of the small rubber ball?



How many gray objects are in front of the tiny green shiny ball and right of the big blue matte thing?



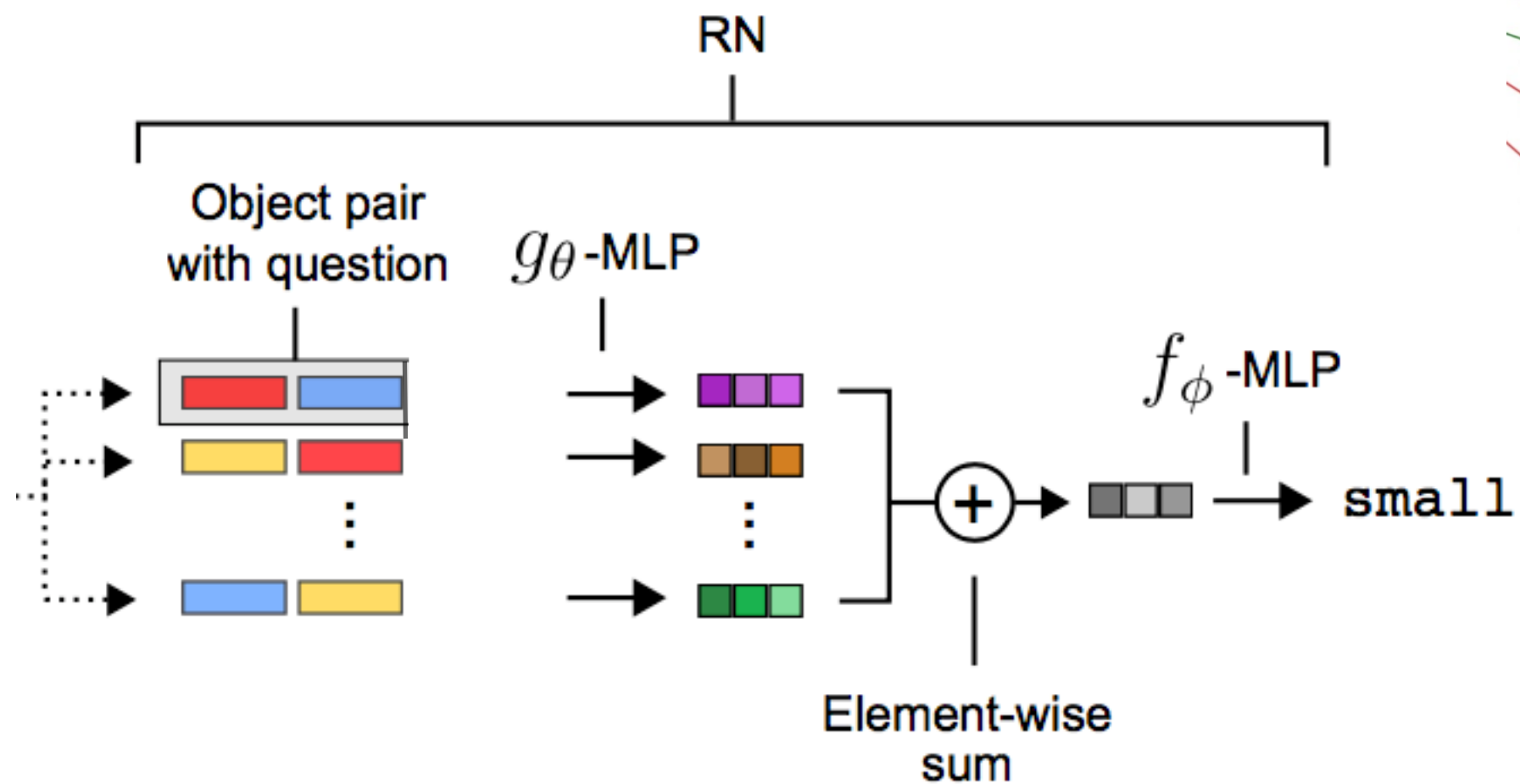
What number of objects are big red matte cubes or things on the right side of the large red matte block?

Results on CLEVR from pixels

Model	Overall	Count	Exist	Compare Numbers	Query Attribute	Compare Attribute
Human	92.6	86.7	96.6	86.5	95.0	96.0
Q-type baseline	41.8	34.6	50.2	51.0	36.0	51.3
LSTM	46.8	41.7	61.1	69.8	36.8	51.8
CNN+LSTM	52.3	43.7	65.2	67.1	49.3	53.0
CNN+LSTM+SA	68.5	52.2	71.1	73.5	85.3	52.3
CNN+LSTM+SA*	76.6	64.4	82.7	77.4	82.6	75.4
CNN+LSTM+RN	95.5	90.1	97.8	93.6	97.9	97.1

Relation Networks

Focus on relational reasoning



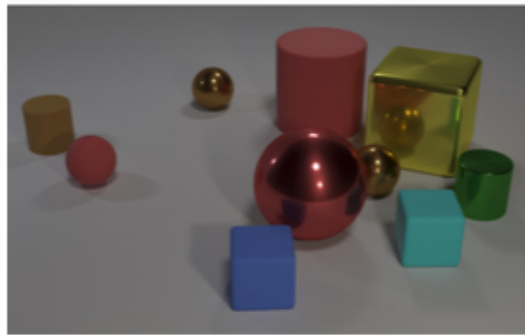
Relation Networks

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

RN has three notable strengths:

- Learn to infer relations
 - 모든 object 쌍을 고려하기 때문에 선행 지식이 필요 없음
- Data efficient
 - Object간 relation을 계산하는데 단 하나의 함수를 사용
- Operate on a set of objects
 - Input과 output의 순서가 상관 없음

Relation Networks



What size is the cylinder
that is left of the brown
metal thing that is left
of the big sphere?

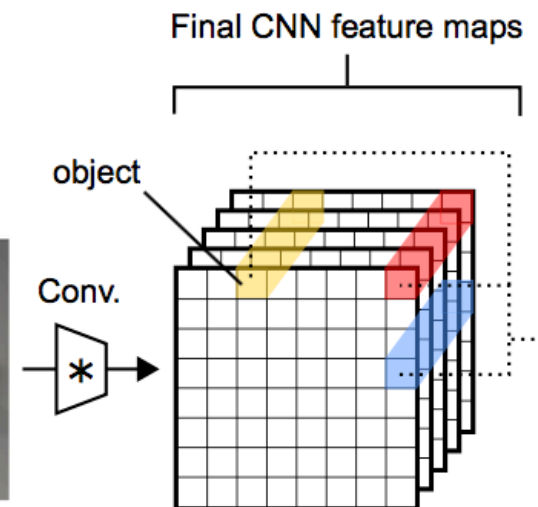
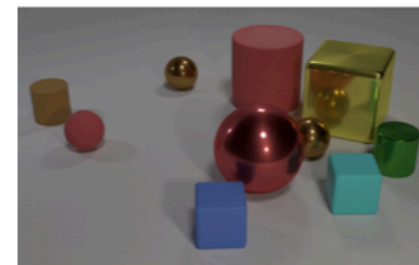
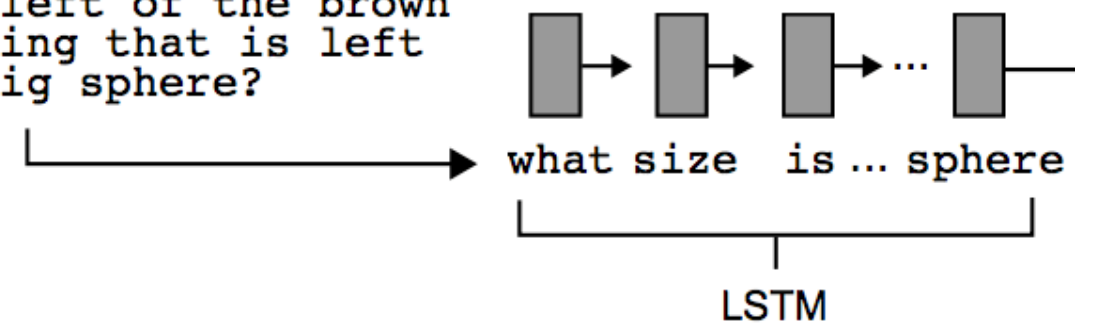


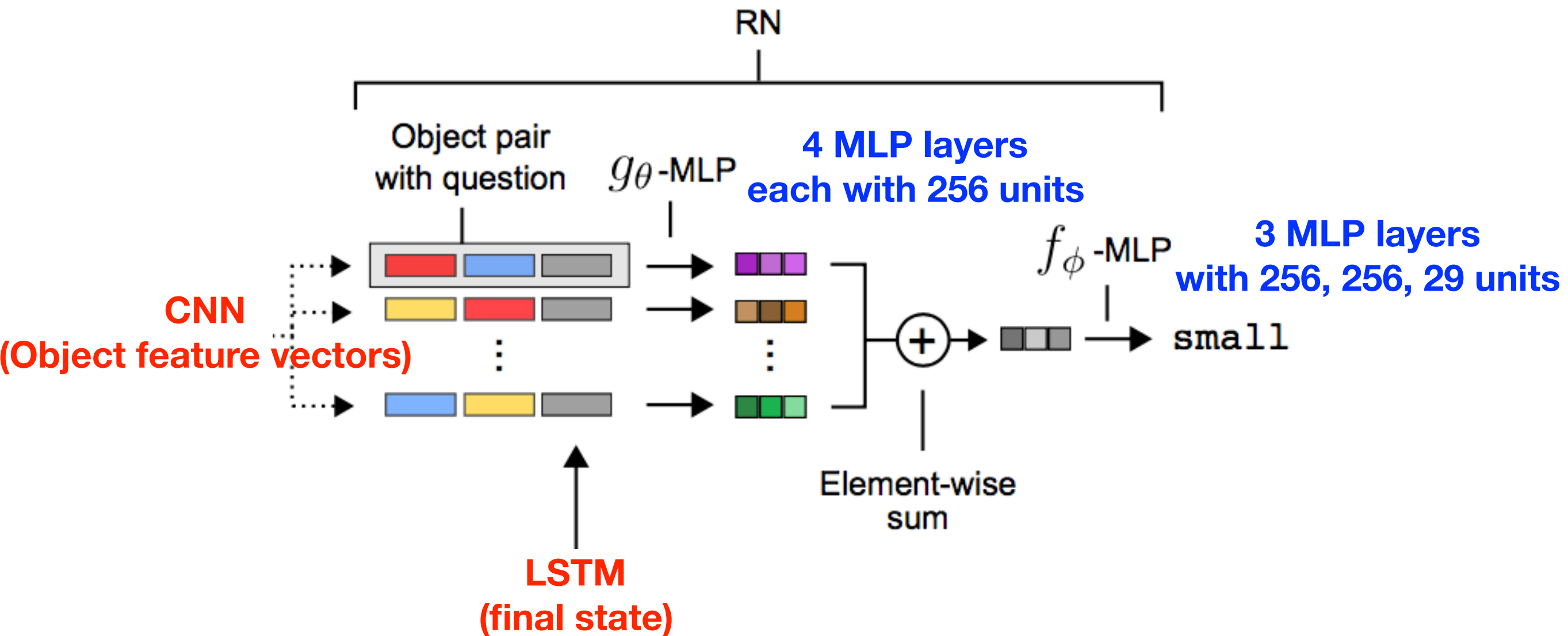
Image - CNN

What size is the cylinder
that is left of the brown
metal thing that is left
of the big sphere?



Text - LSTM

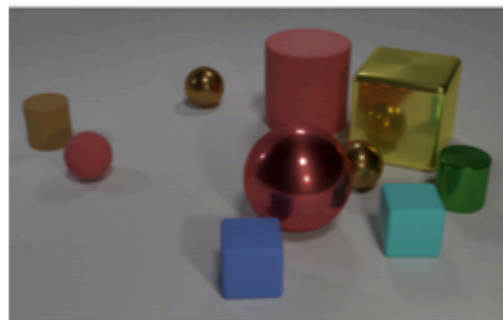
Relation Networks



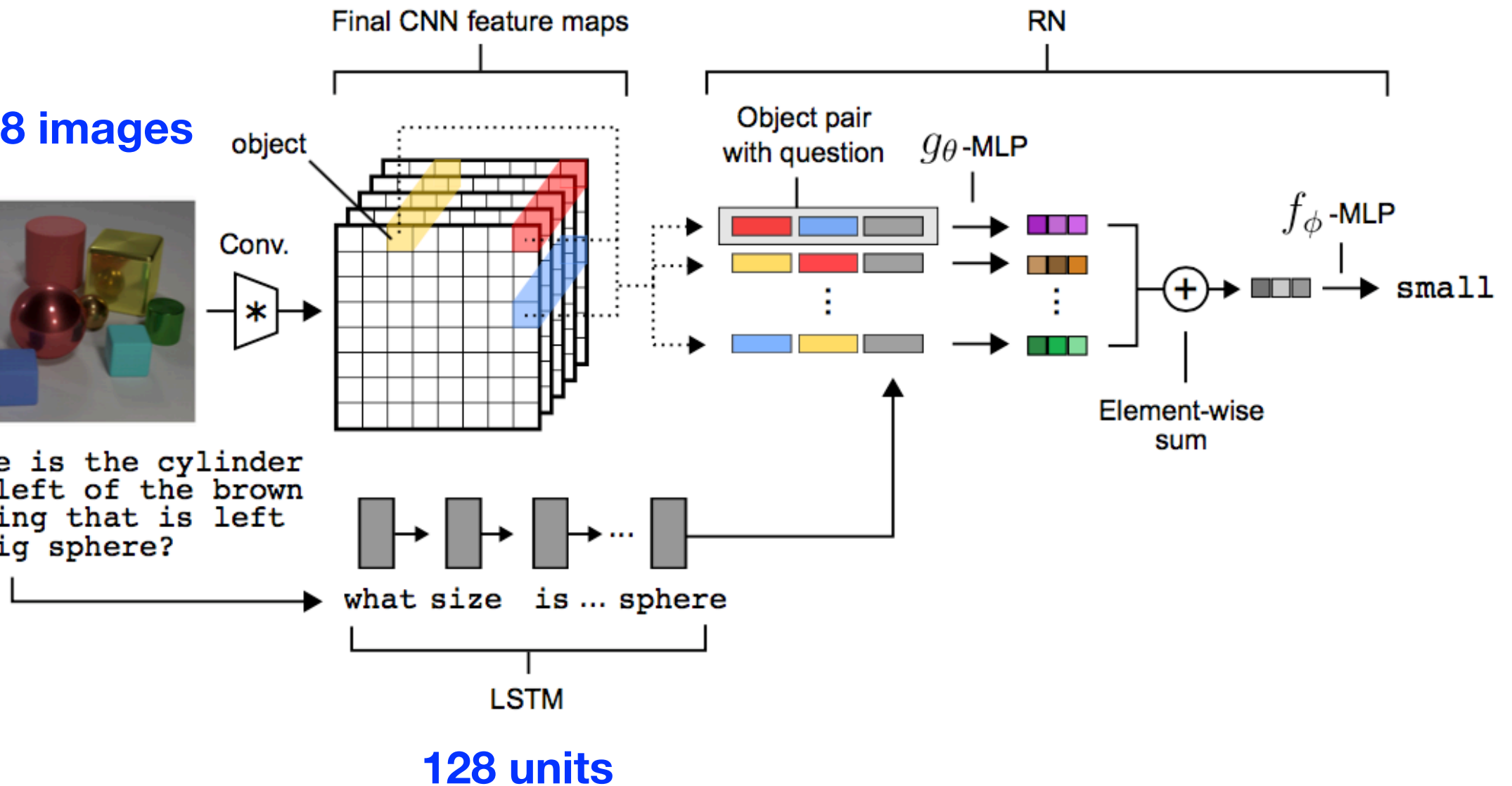
Visual QA architecture(1)

4 conv. layers
each with 3x3x24 kernels, stride 2

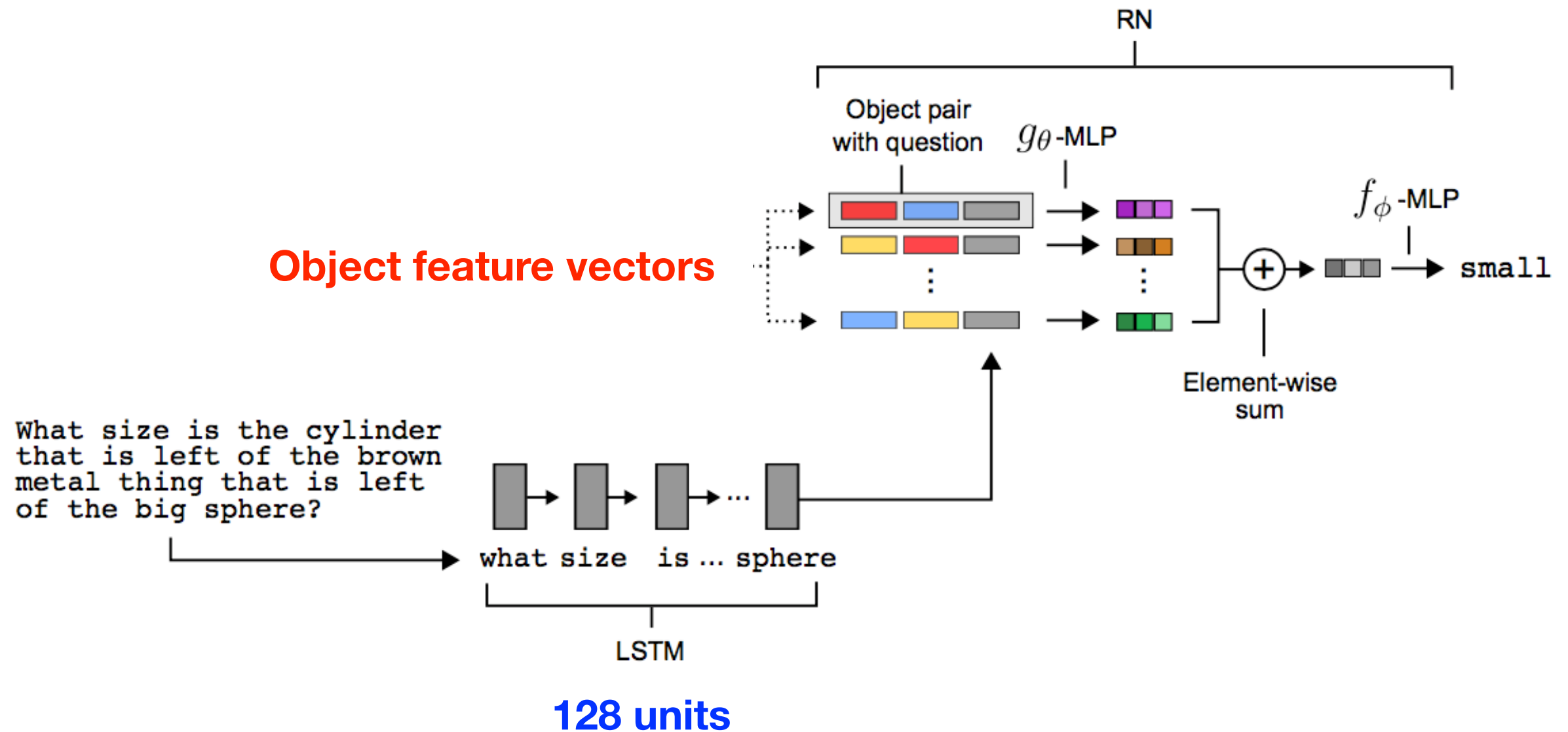
128x128 images



What size is the cylinder
that is left of the brown
metal thing that is left
of the big sphere?



Visual QA architecture(2)

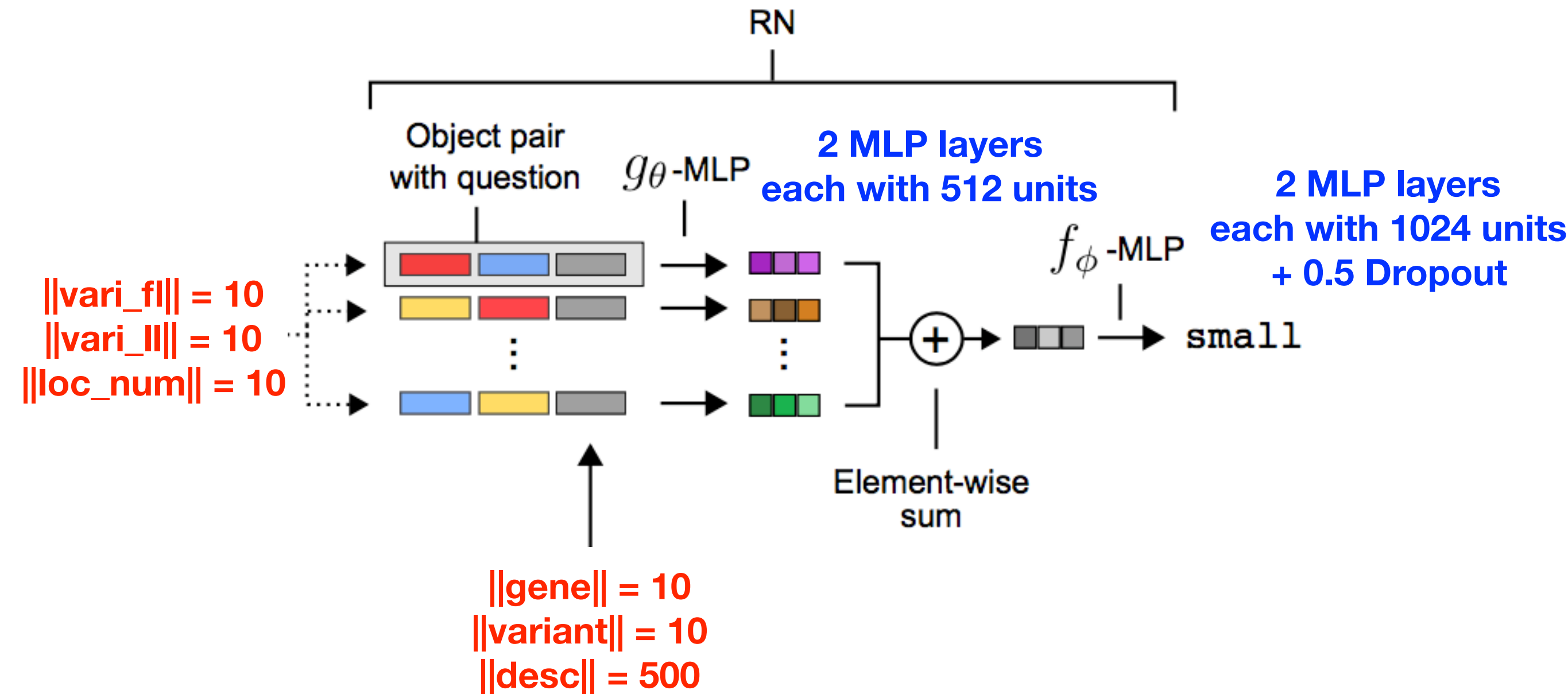


Kaggle

Gene, variant, description

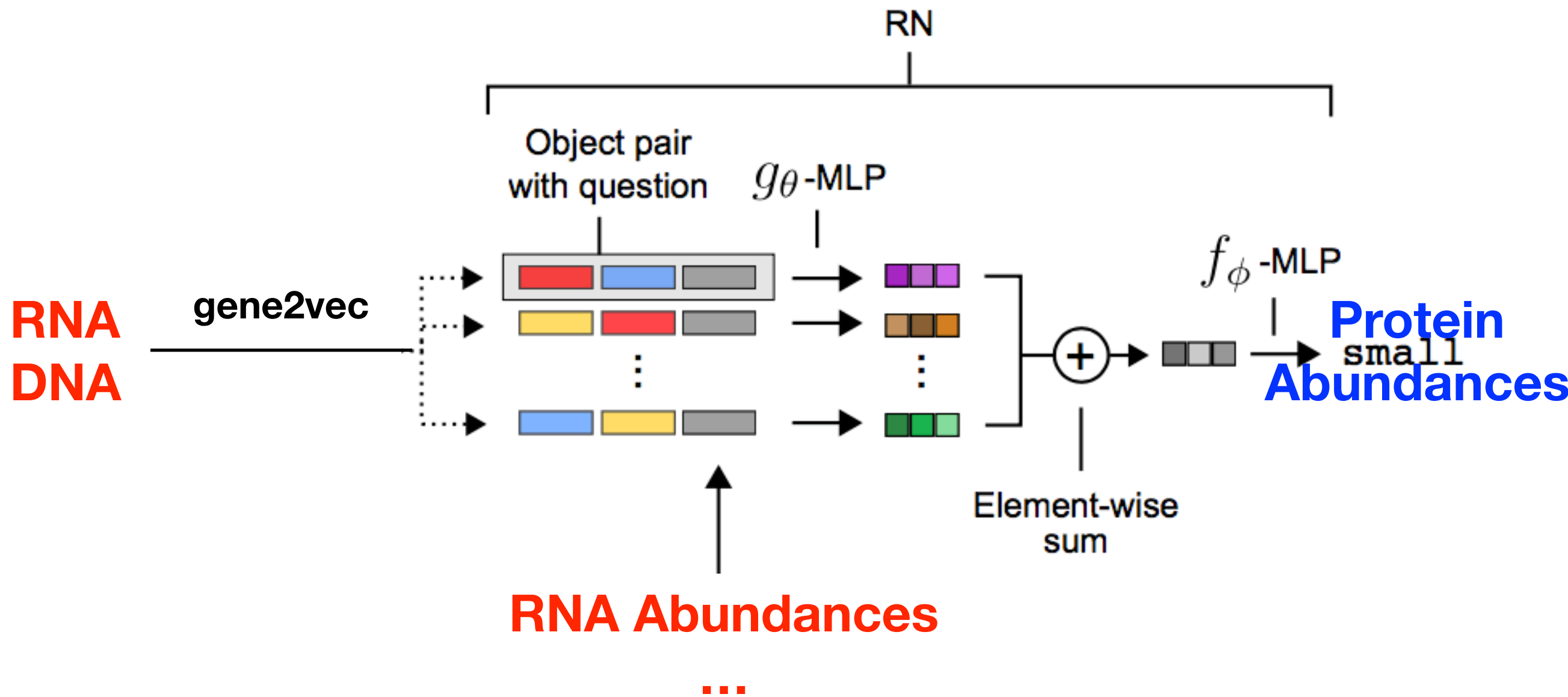


Genetic mutation class



Dream challenge

DNA, RNA $\xrightarrow{\text{predict}}$ Protein Abundances



Question

- 3개 이상의 object pair relation에 대해서도 가능한가
- 몇개의 Object까지 가능한가

Q & A