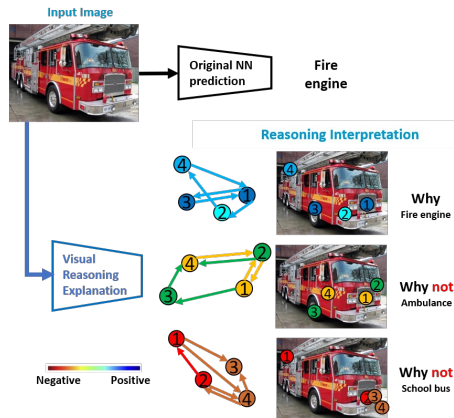




Overall

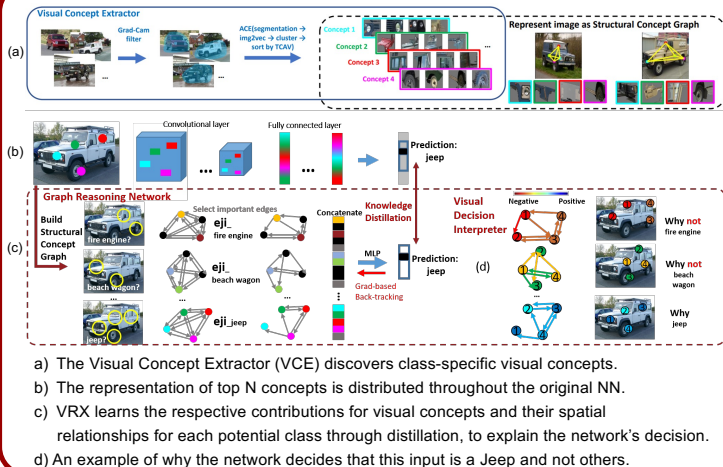
- Most of the existing explanations are limited to low-level relationships and are insufficient to provide in-depth reasoning.
- They do not offer guidance on how to correct mistakes made by the original model.



Our Contribution:

- Reveal reasoning logic with the structural visual concept.
 - Answer why and why not given a prediction
 - Visual concepts + relationship between concepts
 - Logical concept-level explanations
- We propose a GNN-based Graph Reasoning Network (GRN) that imitates the original NN's decision-making process with knowledge transfer and distillation.
- Our method takes a step towards diagnosing reasons for any incorrect predictions and guide improving the performance of original network.

Visual Reasoning Explanation Framework

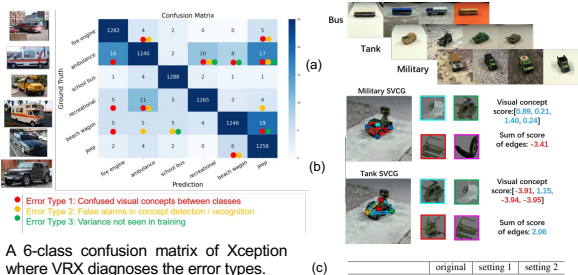


Sensitive and Diagnosis Experiments

(a) Visual sensitive Experiments



(b) Structure sensitive Experiments



A 6-class confusion matrix of Xception where VRX diagnoses the error types.

Error type	Cause of error			
	total	concept	structure	both
Before correction	119	5	6	108
Substitute with Random patches	117	5	6	106
Change good concepts	115	5	6	104
VRX guided correction	5	1	2	2

Out of 119 images initially misclassified by Xception, only 5 remain misclassified after VRX-guided image editing.

- Pose biased training dataset.
- VRX diagnosis finds the negative contribution of concept relationships caused most incorrect predictions while visual concepts contribute positively.
- VRX diagnosis can guide performance improvement.