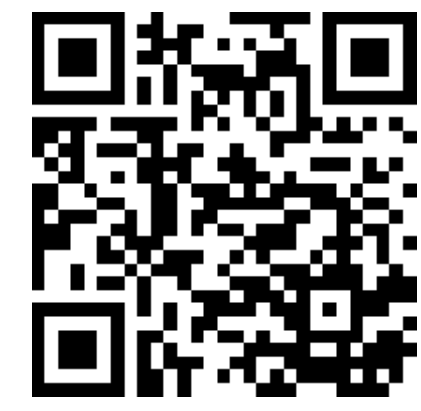


# Classification-Regression for Chart Comprehension



Matan Levy<sup>1</sup> Rami Ben-Ari<sup>2</sup> Dani Lischinski<sup>1</sup>

<sup>1</sup> The Hebrew University of Jerusalem, Israel <sup>2</sup> OriginAI, Israel

## The task: Chart Question Answering

Input: **Chart** (.png) + **Question** (text)  
Output: **Answer** (text / float)

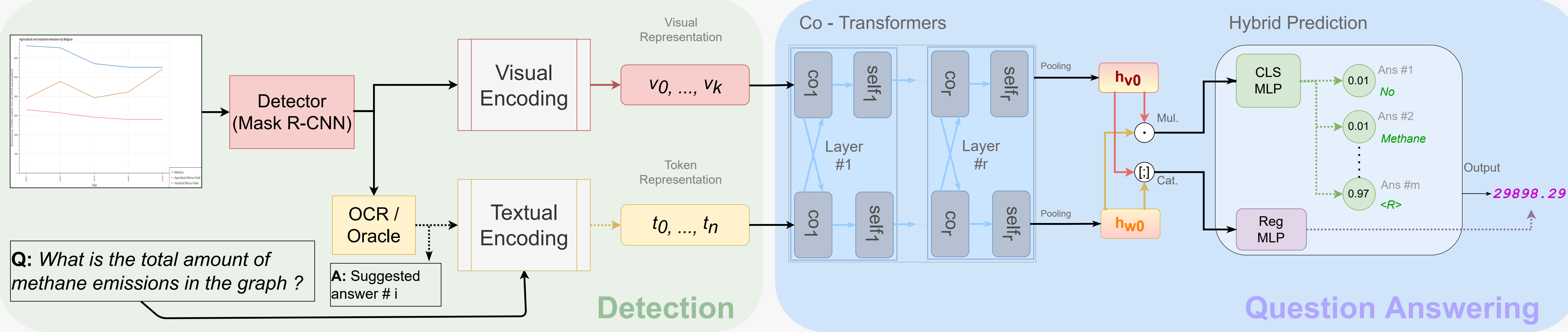
Figures and charts play a major role in data comprehension. Chart Question Answering (CQA) task is perceptually related to the task of Visual Question Answering (VQA) on natural images, yet exhibits significantly different characteristics. The VQA task is typically treated as classification; however, **answering questions about charts, often requires regression.**

Previous methods mostly focus on classification Q&As or handle classification and regression by two separate networks. **Our approach is the first to handle all types of Q&As jointly, leveraging multi-task learning regime.**

In light of the new comprehensive and realistic PlotQA dataset, we discuss the strong shortcomings of the previous benchmarks in CQA, show results on several datasets, while significantly outperforming previous results on PlotQA.

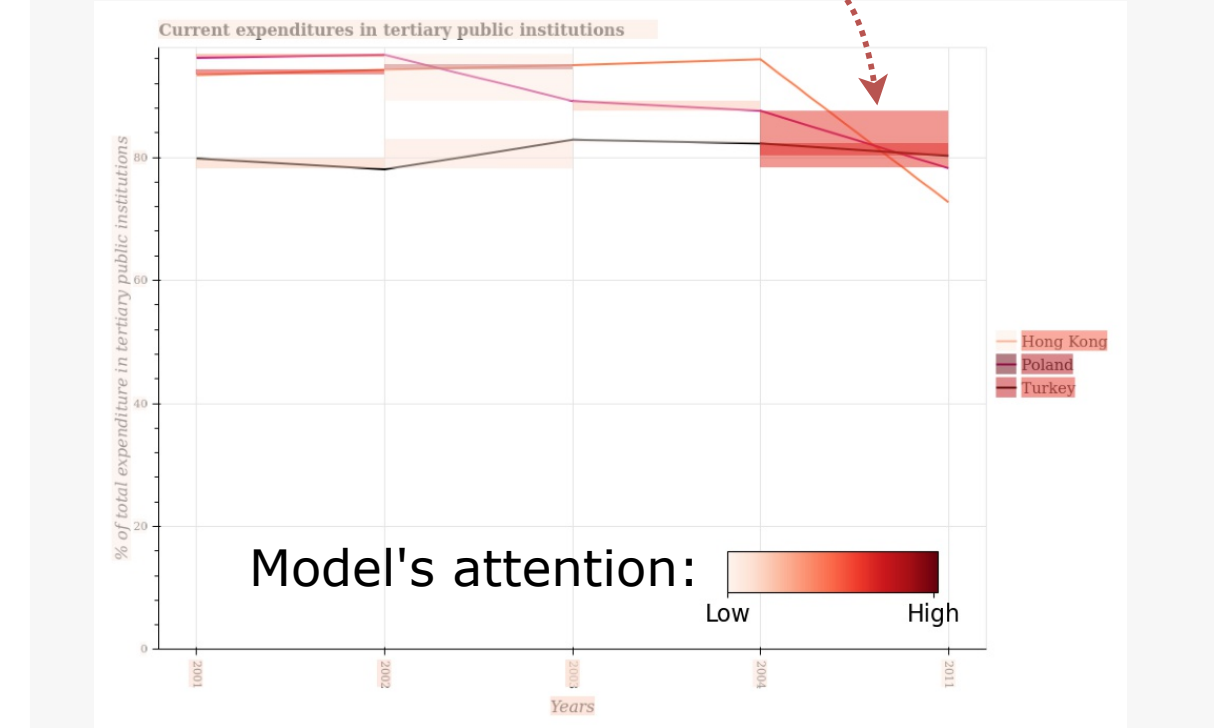
## Pipeline of Classification-Regression Chart Transformer (CRCT)

1. Detect visual/textual elements in the chart.
2. Apply OCR (where it is needed)
3. Build features for each element
4. Mix information (transformers)
5. Feed two parallel MLPs for hybrid prediction
6. Output answer from the relevant MLP

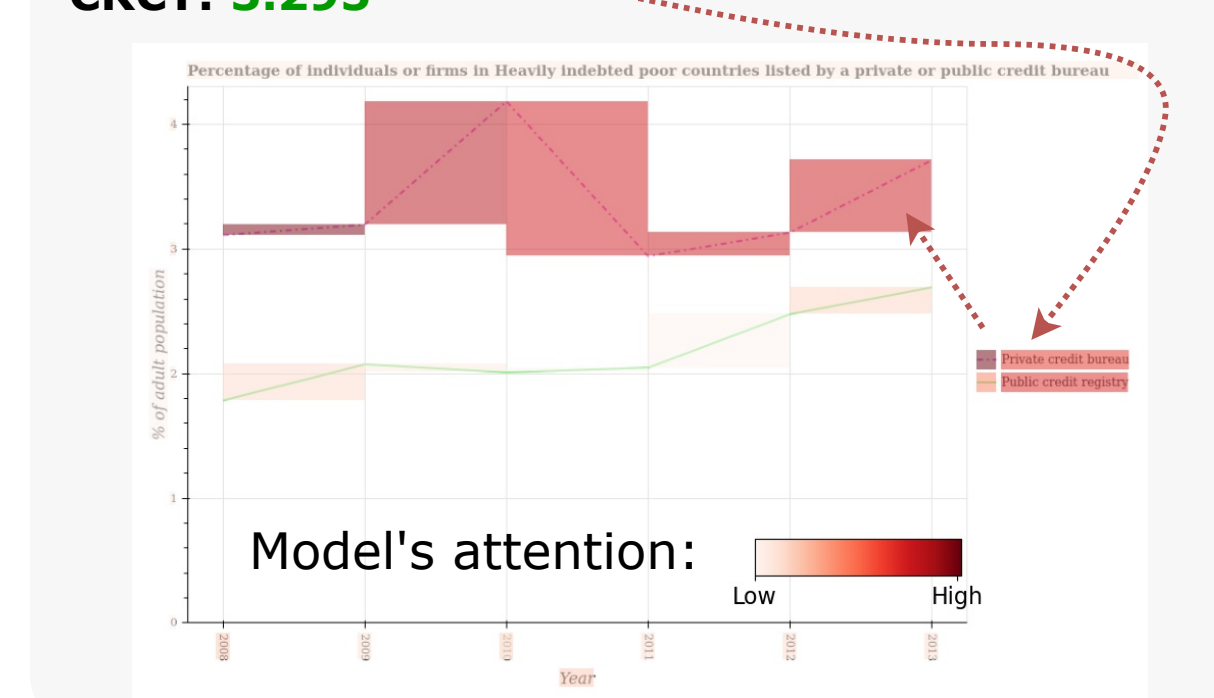


## Explainability

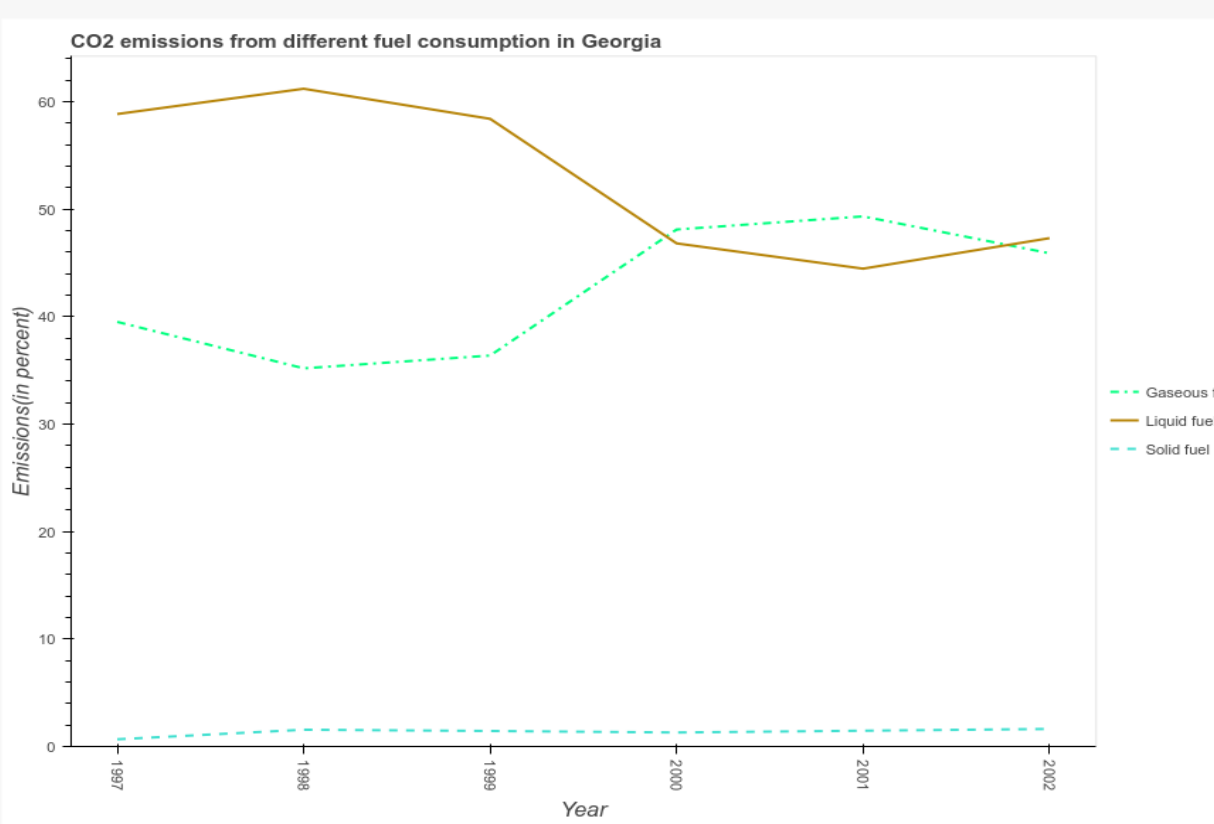
Q: How many lines intersect with each other?  
CRCT: 3



Q: What is the average percentage of firms listed by private credit bureau per year?  
CRCT: 3.295

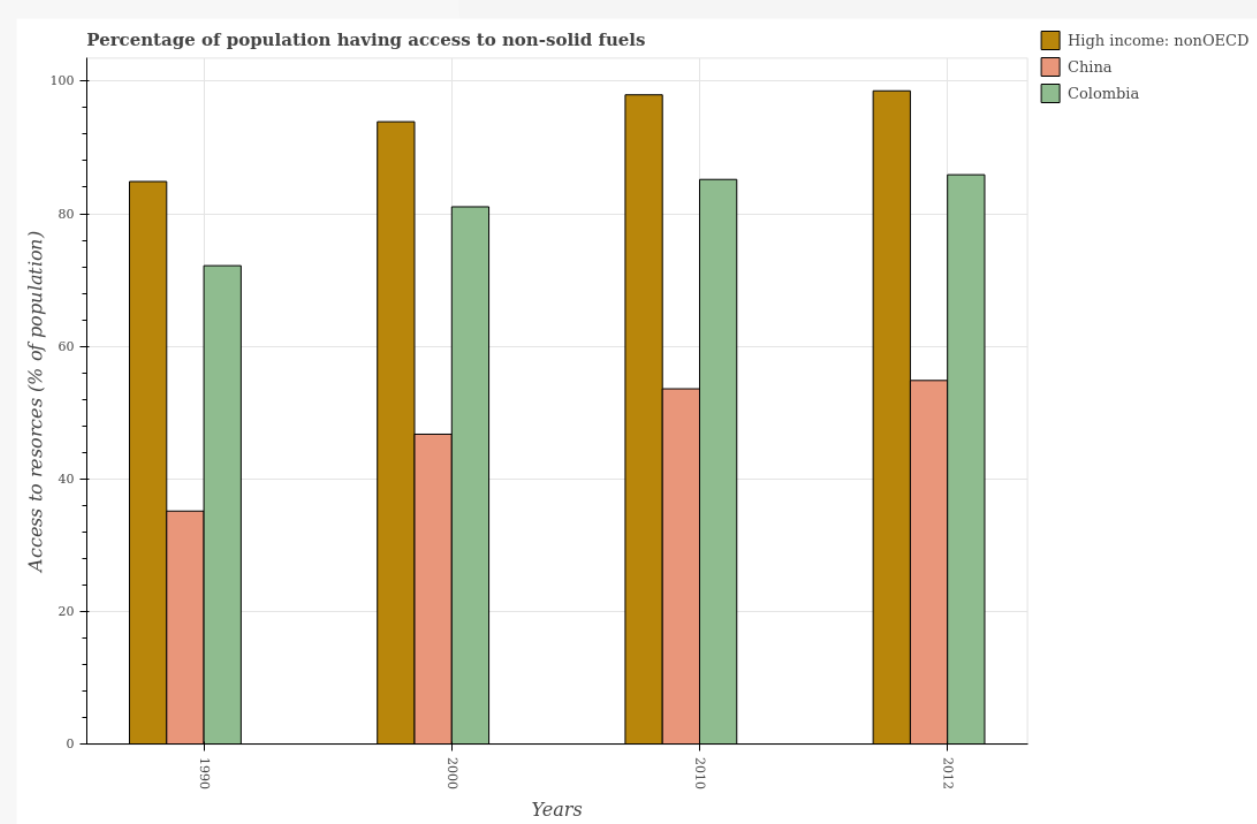


## Test Examples



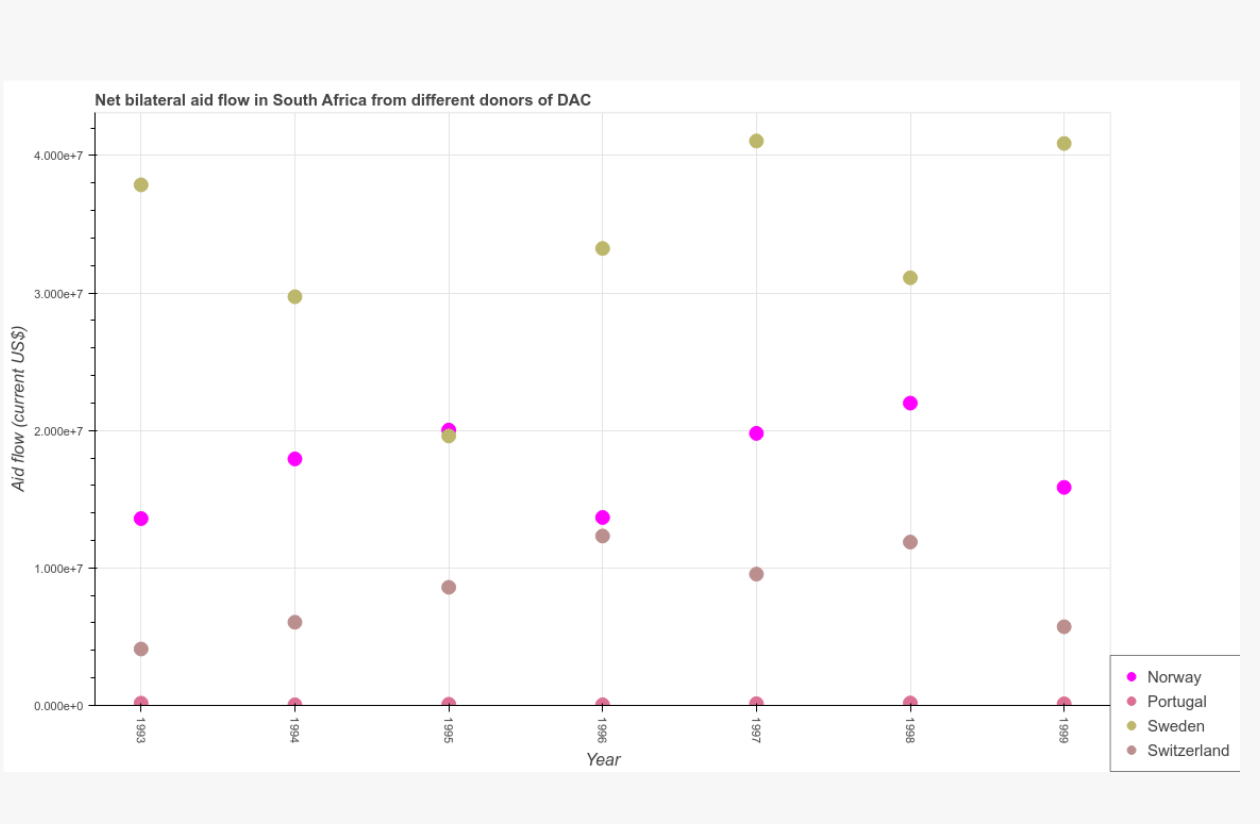
Q: Across all years, what is the minimum percentage of liquid fuel emission?  
CRCT: **45.02** (Error: +1.28%)

Q: What is the label or title of the Y-axis?  
CRCT: **Emissions(in percent)**



Q: What is the average percentage of population having access to non-solid fuels in High income: nonOECD per year?  
CRCT: **94.18** (Error: +0.4%)

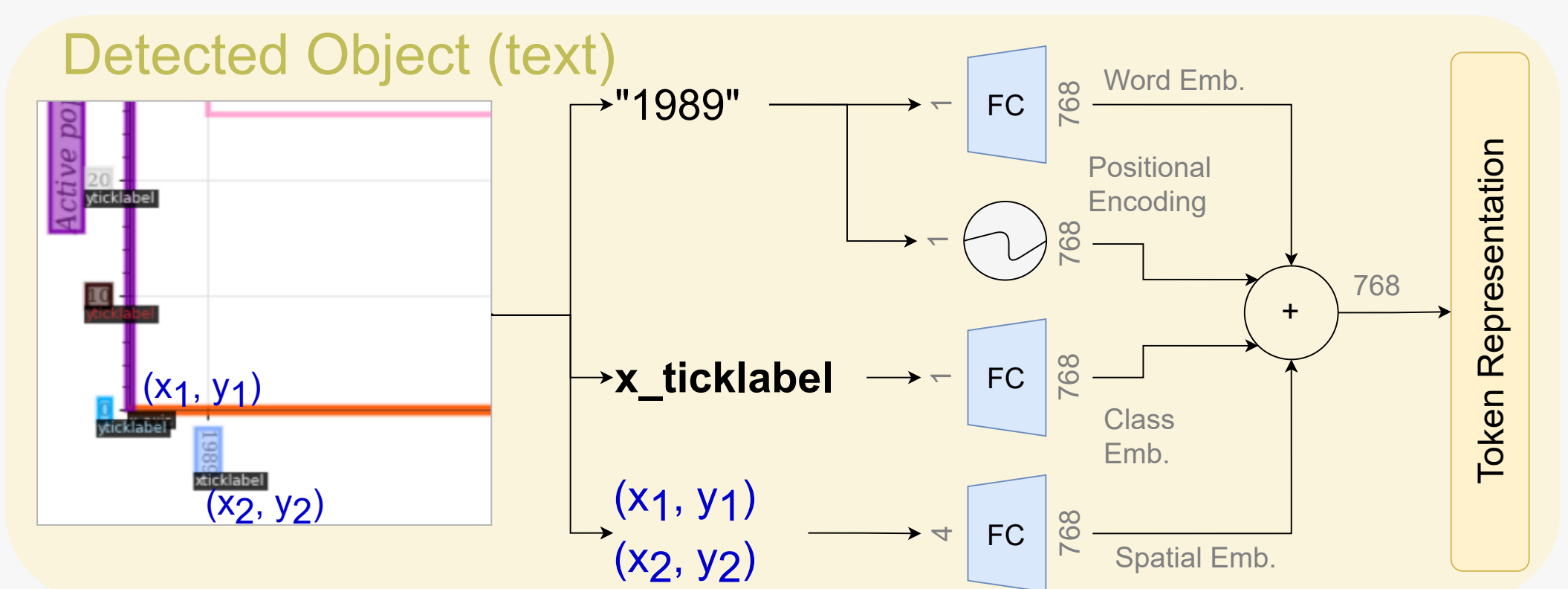
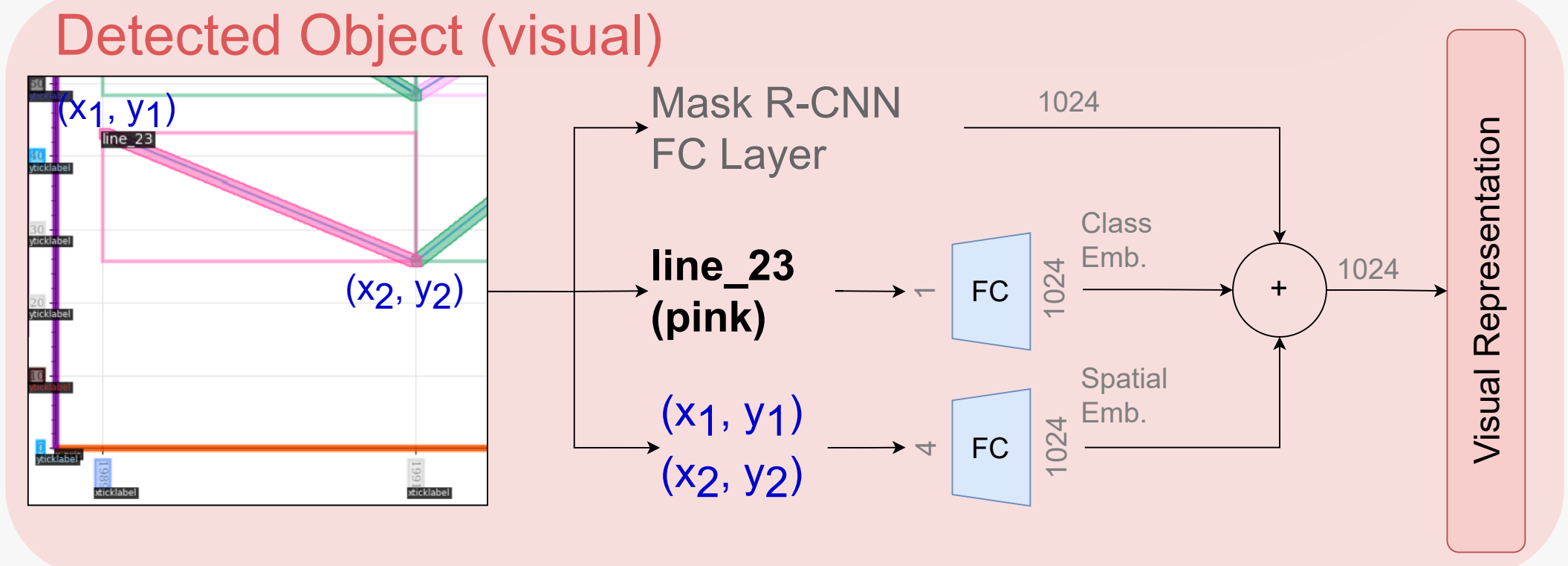
Q: Are the number of bars on each tick of the X-axis equal?  
CRCT: **Yes**



Q: What is the average amount of aid given by Norway per year?  
CRCT: **17859348.0** (Error: +1.75%)

Q: Does "Italy" appear as one of the legend labels in the graph?  
CRCT: **No**

## Feature Representation



## Results

Results on PlotQA dataset. **S**, **D** and **R** stand for **Structural**, **Data Retrieval** and **Reasoning** question categories, respectively. For results on other datasets, see the paper.

Method	PlotQA-D1 (29.86% regression Q&As)			
	S	D	R	Overall
PReFIL [14]	96.66	58.69	31.66	57.91
PlotQA-M [20]	86.31	45.68	31.2	53.96
CRCT (ours)	96.13	<b>94.52</b>	<b>54.87</b>	<b>76.94</b>
CRCT-10%	87.15	74.71	29.19	57.75

Method	PlotQA-D2 (88.84% regression Q&As)			
	S	D	R	Overall
PReFIL [14]	96.66	21.9	3.9	10.37
PlotQA-M [20]	75.99	58.94	15.77	22.52
CRCT (ours)	96.23	<b>66.65</b>	<b>25.81</b>	<b>34.44</b>

## Datasets

**Why PlotQA?**  
To the best of our knowledge, PlotQA is currently the most realistic benchmark, based on real-world data. It provides non-integer answers generally not seen in training, resulting over 5.7M of different answers.

Dataset	#Q&A pairs	#Unique answers	Regression answers	Publicly Available
FigureQA	2.4M	2	✗	✓
DVQA	3.5M	1.5k	✗	✓
LEAF-QA	1.9M	12k	✗	✗
LEAF-QA++	2.6M	25k	✗	✗
PlotQA-D1	8.2M	1M	(29.86%)	✓
PlotQA-D2	29M	5.7M	(88.84%)	✓

Code: <https://github.com/levymn/CQA-CRCT>  
Project Page: <https://www.vision.huji.ac.il/crct>