Research Papers Details

Sl no: 1
Paper Title: Answering Questions about Charts and Generating Visual Explanations
Year: April 2020
Publication: CHI '20: Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems
Link: https://dhkim16.github.io/vis-qa/pdf/paper.pdf
Reader: Sneha/Soorya/Himanshu
Summary (In Points):
Dataset:
Architecture:
Future Scope / Shortcomings:

Sl no: 2
Paper Title: FigureNet: A Deep Learning model for Question-Answering on Scientific Plots
Year: July 2019
Publication: 2019 International Joint Conference on Neural Networks (IJCNN)
Link: https://arxiv.org/pdf/1806.04655.pdf
Reader: Sneha
Summary (In Points):
Uses CNN with depth-wise convolutions, LSTM and feed-forward NN.
Dataset: FigureQA
Architecture:
3 MODULES:
- Module 1: Identifying the plot elements and the color of each of the elements (max of 11 plot elements) - Uses a Convolutional Neural Network that includes depth-wise convolutions to retain color information, and not just get an aggregate. Output of the CNN is a compressed image representation that is passed to an LSTM with 2 layers.
- Module 2: Identifying the values associated with each of the plot elements: Similar to that of Module 1
- Module 3: Question encoding using LSTM and question color encoding

All the three modules are concatenated and passed on to a feed-forward NN to produce a binary yes/no output (using sigmoid activation)
Future Scope / Shortcomings:
- OCR
- Open-ended questions using OCR, finding bounded boxes.

Sl no: 3

Paper Title: ChartNet: Visual Reasoning over Statistical Charts using MAC-Networks

Year: July 2019

Publication: IJCNN 2019. International Joint Conference on Neural Networks.

Link: https://arxiv.org/pdf/1911.09375.pdf

Reader: Himanshu

Summary (In Points):

- The paper solves the problem of reasoning over statistical charts (only bar charts and pie charts) using MAC-Network (Memory, Attention, and Composition)
- The model is capable of answering open-ended questions and gives chart-specific answers
- The classification layer of MAC is replaced by the regression layer and constructs a bounding box around the text of the answer. OCR is used to read the text and display the answer.

Dataset: CLEVR dataset (Dataset Generation),

       10 qa for each image (Total = 3000 images used)

Architecture:

MAC Network: 3 Layers:

       Input Unit

       MAC Cell

       Output Unit

Future Scope / Shortcomings:

- Training the model over different charts and different orientation
- Answer questions that require numerical operation
- Textual summary of the statistical charts

Sl no: 4

Paper Title: PlotQA: Reasoning over Scientific Plots

Year:

Publication:

Link: https://arxiv.org/pdf/1909.00997.pdf

Reader:

Summary (In Points):

Dataset:

Architecture:

Future Scope / Shortcomings:

SI no: 5
Paper Title: DVQA: Understanding Data Visualizations via Question Answering
Year:
Publication:
Link: https://arxiv.org/pdf/1801.08163.pdf
Reader: Himanshu
Summary (In Points):
Dataset:
Architecture:
Future Scope / Shortcomings:

SI no: 6
Paper Title: Open-Ended Visual Question Answering by Multi-Modal Domain Adaptation
Year:
Publication:
Link: https://arxiv.org/pdf/1911.04058.pdf
Reader:
Summary (In Points):
Dataset:
Architecture:
Future Scope / Shortcomings:

SI no: 7
Paper Title: VQA: Visual Question Answering
Year:
Publication:
Link: https://arxiv.org/pdf/1505.00468.pdf
Reader: Soorya
Summary (In Points):
Dataset:
Architecture:
Future Scope / Shortcomings:

Additional (Not related to our topic)

SI no: 1
Paper Title: Visual Question Answering on 360∘ Images
Year:
Publication:
Link: https://faculty.ucmerced.edu/mhyang/papers/wacv2020_360_vqa.pdf
Reader:
Summary (In Points):
Dataset:
Architecture:
Future Scope / Shortcomings:

SI no: 2
Paper Title: CGMVQA: A New Classification and Generative Model for Medical Visual Question Answering
Year:
Publication:
Link: https://ieeexplore.ieee.org/abstract/document/9032109
Reader:
Summary (In Points):
Dataset:
Architecture:
Future Scope / Shortcomings:

Template

SI no:
Paper Title:
Year:
Publication:
Link:

Reader:

Summary (In Points):

Dataset:

Architecture:

Future Scope / Shortcomings: