VisImages: A Fine-Grained Expert-Annotated Visualization Dataset

Dazhen Deng, Yihong Wu, Xinhuan Shu, Jiang Wu, Siwei Fu, Weiwei Cui, Yingcai Wu

Group 32 - 111550037 嚴偉哲, 111550127 郭穎達

Motivation – Importance of Images

\# **Image** is a crucial element to visualization
    - Showcase innovative visual designs and representation
    - Contain many information other than the dataset itself
      (graph types/combinations, keywords, etc.)

\# Can expand the understanding of this field

\# Provide new ways to research **AI for visualization**

2

Motivation – Flaw in Existing Studies

Existing image datasets mostly focus on **basic chart types**

   - Classification might fail when dealing with complex visualizations

VisImages aims to address this problem

   - Collect and provide a comprehensive dataset

   - Serve as a **benchmark** for future classification tests

3

Related Works – Image Datasets

Existing datasets focus on basic charts with **simple layouts**

- D3, Vega-Lite, etc.

Datasets are categorized by basic chart types

- e.g., ChartSense (10 types), MassVis (12 types)

Some datasets also collect **novel designs** from experts but without detailed information
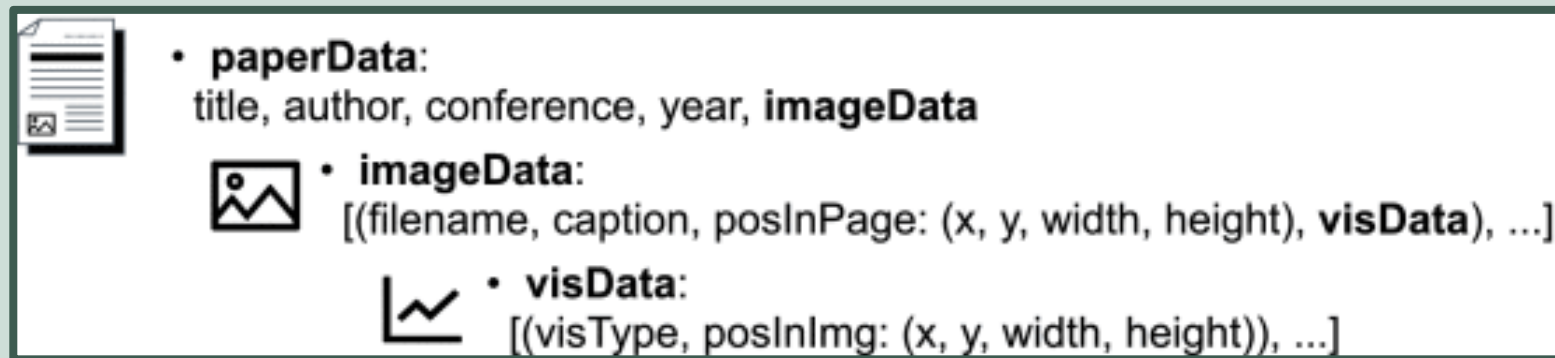
- e.g., VIS30K (4 types)

4

Related Works – Comparison

1) **Unique Data Source:** VisImages focuses on fresh, complex designs from visualization researchers

2) **Comprehensive Annotations:** Fine-grained taxonomy with visualization types and positions

3) **Design Complexity & Quantity:** Includes multi-chart systems, 100x the visualizations of MV Dataset

5

Dataset – Structure

Three levels: **paper**, **image**, and **visualization** data

Annotated with **captions**, **positions**, and **visualization types**

- **paperData**:
  title, author, conference, year, **imageData**
  - **imageData**:
    [(filename, caption, posInPage: (x, y, width, height), **visData**), ...]
    - **visData**:
      [(visType, posInImg: (x, y, width, height)), ...]

6

Dataset – Data Source

\# Collect images from visualization conferences, **VAST** and **InfoVis**

    - 2D static visualizations only

\# Extract all **images** with their **captions** from papers

    - 12267 images, 12057 captions

Dataset – Taxonomy

 # Image Type: 13 categories and 34 sub-types based on the proposal of Borkin et al.

Categories	Sub-types
Area	area chart, proportional area chart (PAC)
Bar	bar chart
Circle	donut chart, pie chart
Diagram	flow diagram, chord diagram, Sankey diagram, Venn diagram
Statistic	box plot, error bar, stripe graph
Table	table
Line	contour graph, line chart, storyline, polar plot, parallel coordinate (PCP), surface graph, vector graph
Map	map
Point	scatter plot
Grid & Matrix	heatmap, matrix
Text	phrase net, word cloud, word tree
Graph & Tree	graph, tree, treemap, hierarchical edge bundling (HEB), sunburst/icicle plot
Special	glyph-based visualization, unit visualization

Annotation – Image Type

 # Annotator: Experts and students with experience in data visualization

 # Procedure:

- **Training Session:** Introduction to taxonomy and annotation tasks, followed by a test

- **Formal Study:** Participants annotated images independently

Annotation – Image Type

 # Quality Control:

a. Gold Standards | Simple images for participants to classify to ensure **focus** and **accuracy**

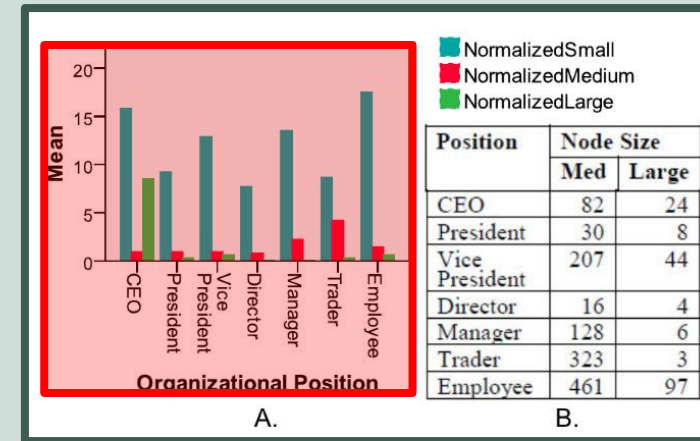b. Majority Voting | 3 participants per image; 2 votes required for acceptance

 # Results:

- Accuracy: 96.4% on gold standards

- Intercoder Reliability: 76.8%

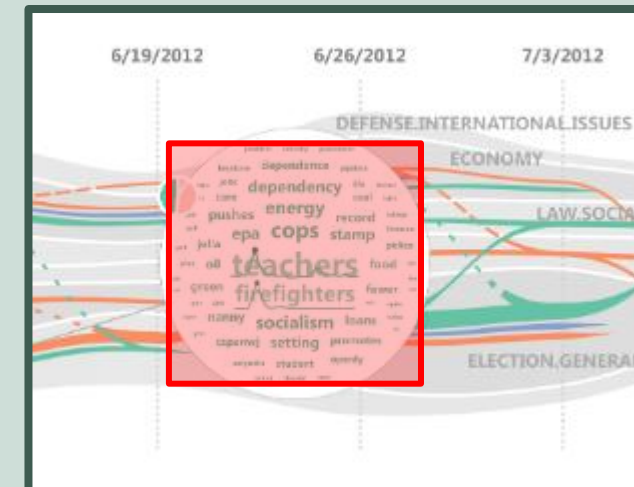- 10289 images labeled; rest marked as "others"

10

Annotation – Bounding Box

1) Coordinated Graph
   - Include axis names, legends, etc.

2) Uncoordinated Graph
   - Focus on the desired content
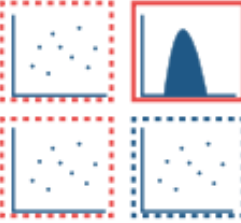

11

Annotation – Bounding Box

 # Annotator: Crowd workers from professional data annotation company

 # Quality Control:

    a. Bounding Boxes | IoU rate *(> 0.9)*

$$IoU = \frac{\text{area of intersection}}{\text{area of union}}$$

    b. Task Quality | F1 score *(> 0.95)*

$$recall = \frac{\#\text{true positive}}{\#\text{conditional positive}}$$

$$precision = \frac{\#\text{true positive}}{\#\text{annotated positive}}$$

$$F_1 \text{ score} = \frac{2 \cdot precision \cdot recall}{precision + recall}$$

Result – Overview

Sub-type	#bbox	#img	Sub-type	#bbox	#img.
bar chart	5053	2058	pie chart	371	153
scatterplot	4269	1754	PAC	288	130
graph	3722	1615	box plot	277	147
heatmap	3202	1187	unit visualization	275	107
line chart	3004	1300	sunburst/icicle	260	120
table	2172	1676	sankey diagram	260	147
map	2106	986	stripe graph	239	123
matrix	1611	656	HEB	185	61
tree	1292	667	chord diagram	128	72
area chart	1125	527	polar plot	123	56
flow diagram	1118	873	storyline	46	25
PCP	975	541	contour graph	16	12
error bar	709	342	surface graph	13	7
treemap	554	268	word tree	9	9
glyph-based	523	259	phrase net	7	7
word cloud	392	184	Venn Diagram	4	4
donut chart	376	143	vector graph	4	2
others	-	1978			

12267 images; 35096 bounding boxes

 # Bounding Boxes / Images Ratio:

a. Bar Chart / Scatterplot (≥2)
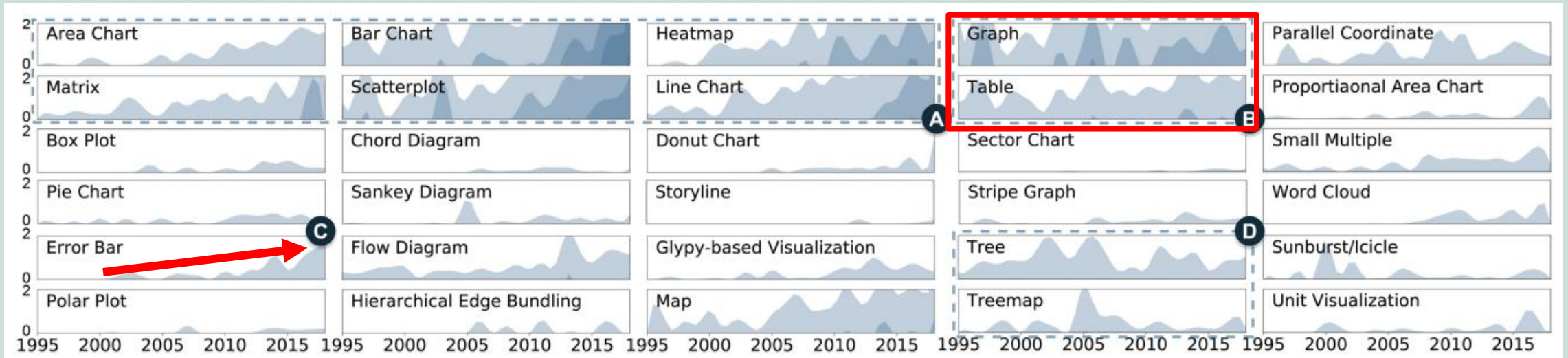
- Commonly serve as unit of small multiples

b. Table / Flow Chart (≈1)

- Independently show the result / framework

13

Result – Overview

 # Image Type Trend

Result – Comparison

 # Compare with MassVis

Source	VisImages	MassVis			
		Scientific	Infographics	News	Government
Area	4.0%	1.9%	4.4%	4.4%	3.5%
Bar	12.1%	6.4%	5.9%	40.2%	36.9%
Circle	1.8%	0.3%	4.7%	1.3%	6.6%
Diag.	6.3%	27.4%	30.6%	7.2%	5.0%
Stat.	3.7%	3.2%	0.3%	0.3%	1.3%
Table	9.8%	8.3%	32.8%	8.2%	21.5%
Line	11.2%	19.1%	1.6%	19.1%	12.9%
Map	6.0%	9.2%	9.1%	13.5%	7.3%
Point	10.6%	16.6%	2.8%	5.0%	0.5%
Grid	7.2%	2.5%	1.9%	0%	0%
Text	1.1%	0%	0%	0.5%	0%
Graph	16.7%	5.1%	5.9%	0.3%	0%
Special	9.5%				
#Vis.	10,289	348	490	704	528

Use Case – VisImages Explorer

\# A website for searching the dataset

\# Useful for **paper research** and **visualization design inspiration**

Use Case – Image Classification

\# Use **ResNet** & **VGG** (both are **CNN**) for the classification models

\# Use **Beagle** as the baseline dataset

The Top-1 and Top-3 Accuracies (%) of Different Models on Visualization Classification Under Different Situations

Training Set	(A) Beagles		(B) VisImages	
Test Set	Beagle	VisImages (Acc↓)	VisImages	Beagle (Acc↓)
ResNet-50	79.3/99.0	32.6/38.8	78.3/92.6	**10.6/9.5**
ResNet-101	80.6/98.9	26.0/33.2	78.9/94.4	**11.5/9.4**
VGG-16	79.9/98.7	36.7/43.8	77.5/92.7	**8.7/7.1**
VGG-19	80.1/98.7	34.8/40.4	76.9/92.2	**8.3/6.0**

The underlined numbers denote the highest accuracies among the models.

Use Case – Image Classification

\# Self test: Beagle is more accurate

 - Images are more **similar** in Beagle

\# Cross test: VisImages is more accurate

 - Images are more **generalized** in VisImages

The Top-1 and Top-3 Accuracies (%) of Different Models on Visualization Classification Under Different Situations

Training Set	(A) Beagles		(B) VisImages	
Test Set	Beagle	VisImages (Acc↓)	VisImages	Beagle (Acc↓)
ResNet-50	79.3/99.0	32.6/38.8	78.3/92.6	**10.6/9.5**
ResNet-101	80.6/98.9	26.0/33.2	78.9/94.4	**11.5/9.4**
VGG-16	79.9/98.7	36.7/43.8	77.5/92.7	**8.7/7.1**
VGG-19	80.1/98.7	34.8/40.4	76.9/92.2	**8.3/6.0**

The underlined numbers denote the highest accuracies among the models.

Use Case – Visualization Localization

\# Use the **bounding box** of the images to train the model

\# Use **Faster R-CNN** for the model
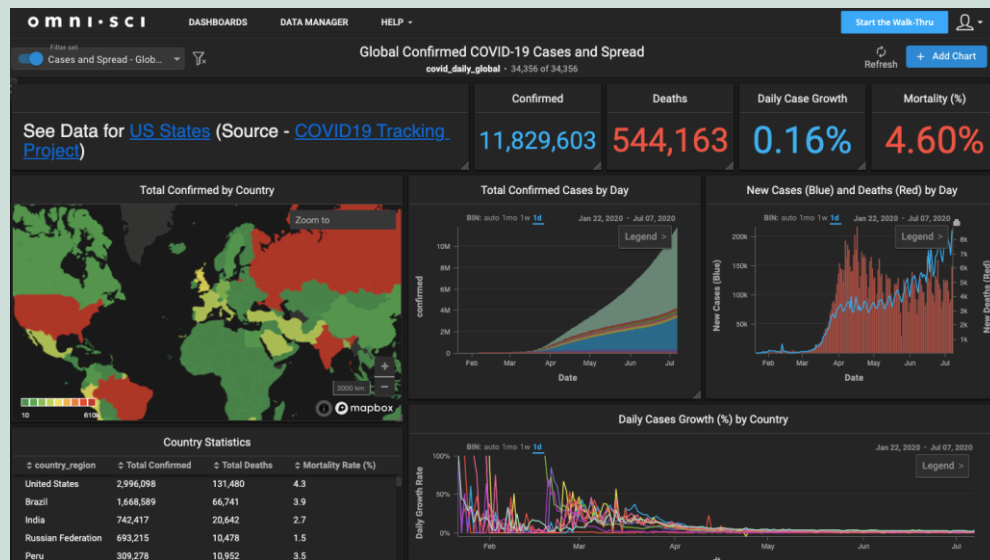
Example of visual analytics system

APs Under Different IoU Thresholds

Sub-type	$AP_{IoU=0.50}$	$AP_{IoU=0.75}$	$AP_{IoU=0.90}$
graph	**0.96**	**0.96**	**0.70**
table	0.95	0.90	0.69
scatterplot	0.83	0.77	0.29
line chart	0.82	0.71	0.23
heatmap	0.80	0.80	0.40
flow diagram	0.75	0.70	0.46
bar chart	0.69	0.58	0.10
map	0.68	0.64	0.54
parallel coordinate	0.67	0.52	0.28
mAP	0.78	0.78	0.52

Use Case – Visualization Localization

Distribution of each visualizations in an interface

20
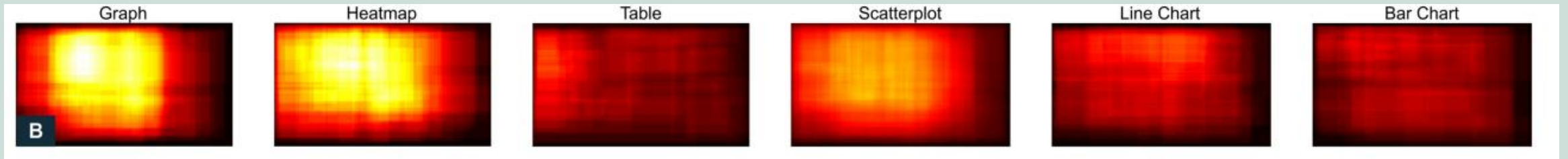
Limitations

Many features of the graph (axis titles, marks, etc.) are not included in the dataset

Only compared the accuracy with one other dataset

     - May not be sufficient to determine its generality

- Thank You For Listening -