

# Information Graphic Summarization using a Collection of Multimodal Deep Neural Networks

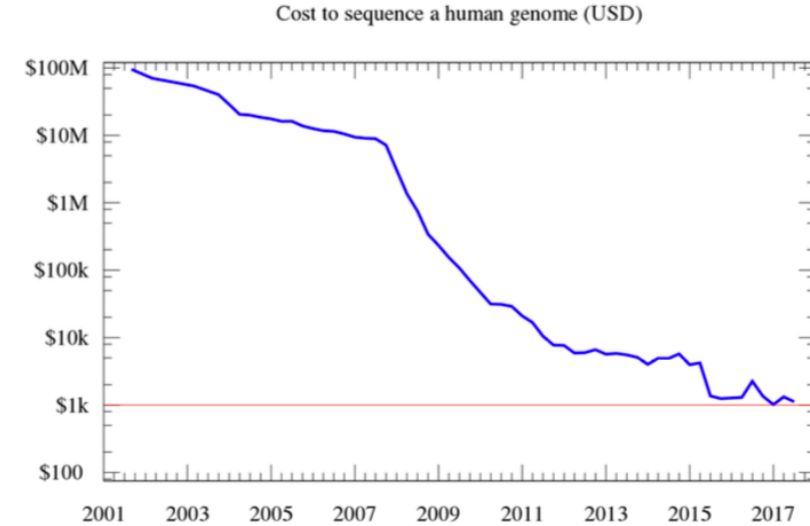
Edward Kim<sup>1</sup>, Connor Onweller<sup>2</sup>, Kathleen F. McCoy<sup>2</sup>

<sup>1</sup>Department of Computer Science, Drexel University, PA

<sup>2</sup>Computer and Information Sciences Department, University of  
Delaware, DE

# Background

- The goal of our work is to provide a useful textual summary of an online information graphic to someone who is visually impaired or otherwise unable to view the graphic. Our objective is to produce a summary containing the salient information content of the graphic.



This graph titled **cost to sequence a human genome (usd)** shows a **falling trend** in **millions** over **year**. The X-axis goes from **2001** to **2017**. The Y-axis goes from **\$100** to **\$100M**. The end of the graph **continues to fall** and the graph shows **few fluctuations**.

# Classification Categories

- For our dataset, we use the 1000 line graph images<sup>1</sup>. These line graphs are non-scientific and representative images that one might see in popular media articles.

<b>“B” Classes</b>	<b>Description of Graph</b>
<b>RT</b>	<b>Rising Trend</b> in the line graph.
<b>FT</b>	<b>Falling Trend</b> in the line graph.
<b>ST</b>	<b>Stable Trend</b> , generally no change in the line graph.
<b>CT</b>	<b>Changing Trend</b> in the graph, e.g. up and down, or down and up, possibly multiple times.
<b>BJ</b>	<b>Big sudden Jump</b> in the line graph.
<b>BF</b>	<b>Big sudden Fall</b> in the line graph.
<b>“T” Classes</b>	<b>Description of Graph</b>
<b>RTT</b>	<b>Rising Trend</b> at the line <b>Tail</b> .
<b>FTT</b>	<b>Falling Trend</b> at the line <b>Tail</b> .
<b>STT</b>	<b>Stablizing Trend</b> at the <b>End</b> of the line graph.
<b>“J” Classes</b>	<b>Description of Graph</b>
<b>V</b>	<b>Volatile</b> with many fluctuations.
<b>NV</b>	<b>Non-Volatile</b> graph.

<sup>1</sup>E. Kim and K. F. McCoy, “Multimodal deep learning using images and text for information graphic classification,” in Proceedings of the 20th International ACM SIGACCESS Conference on Computers and Accessibility, 2018, pp. 143–148.

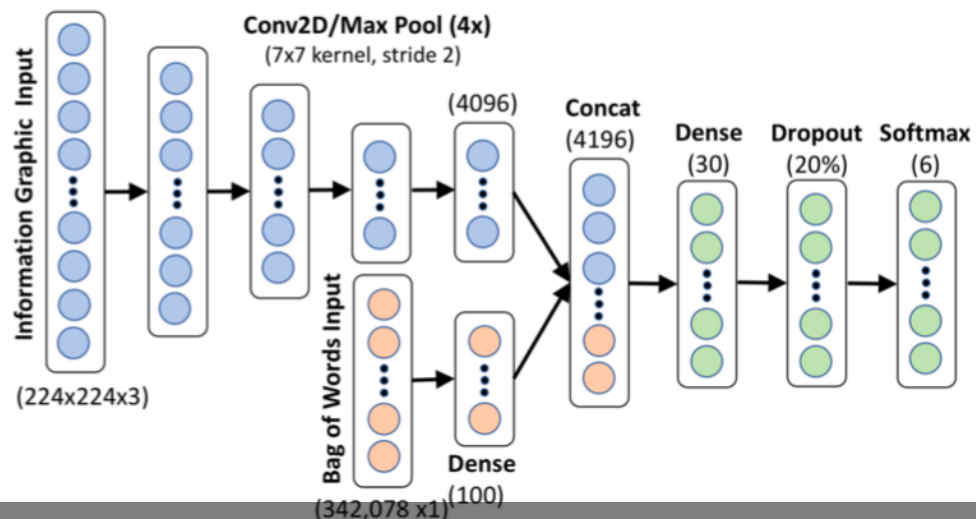


Image and Text Based Networks

## Network “I”(Tail) and “J” (Volatility)

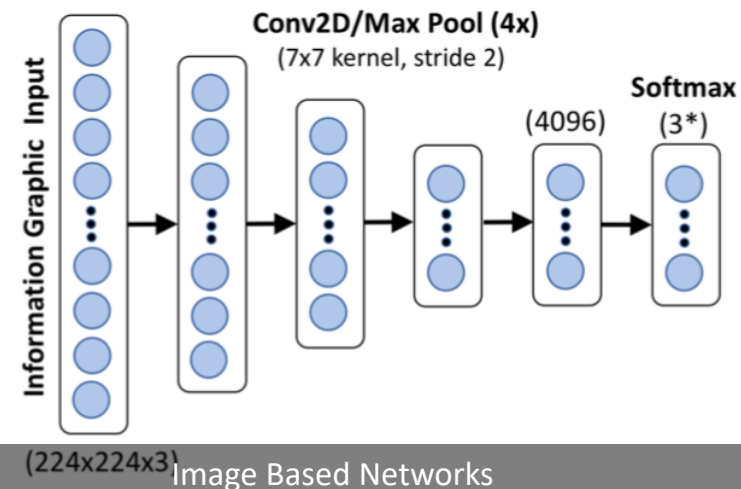
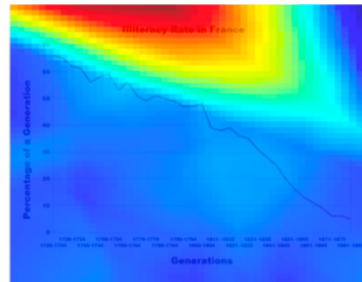


Image Based Networks

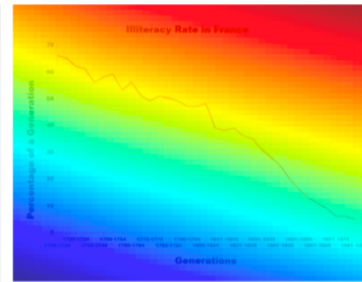
# Network Architectures

# Experiments

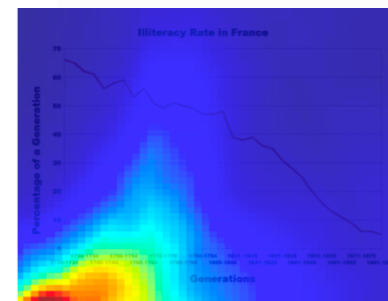
## Attribution



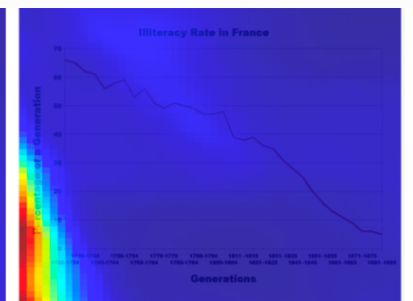
(a) Title



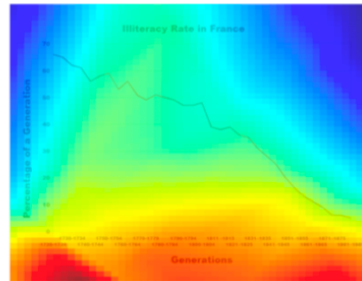
(b) Title (loss of positional sensitivity)



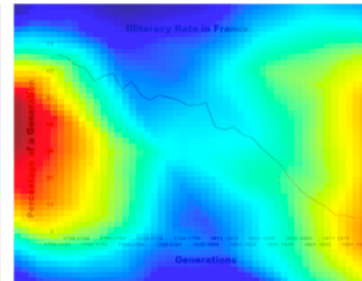
(a) X Domain (Low)



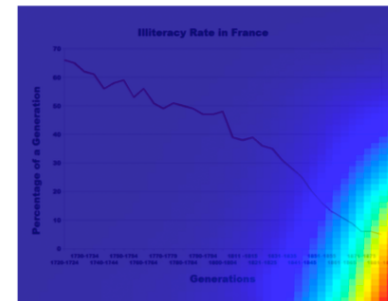
(b) Y Range (Low)



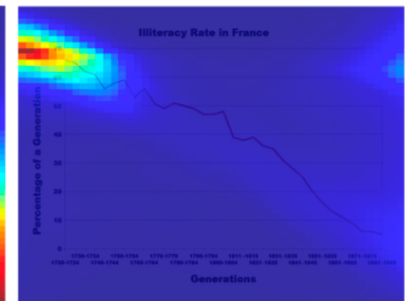
(c) X-Axis Label



(d) Y-Axis Label



(c) X Domain (High)



(d) Y Range (High)

# Experiments

## Summarization Preference



Fig. 8: Summarizations for this information graphic.

**Our method:** This graph titled the rate shows a changing trend in rate over year. The X-axis goes from 2005 to 2018. The Y-axis goes from 0 to 10%. The end of the graph is falling and the graph shows many fluctuations.

**Human 1:** This is a line graph showing a changing trend in the u.s unemployment rate.between 2005 to 2018,

**Human 2:** From 2005, unemployment rises to a recession high of 10.0% in October 2009, then falls to 4% in early 2018.

**Human 3:** A line graph which conveys a THE UNEMPLOYMENT RATE WILL BE INCREASED BETWEEN THE YEAR OF 2006-2018.

**Im2Caption:** a line graph which conveys a changing chart



Fig. 9: Summarizations for this information graphic.

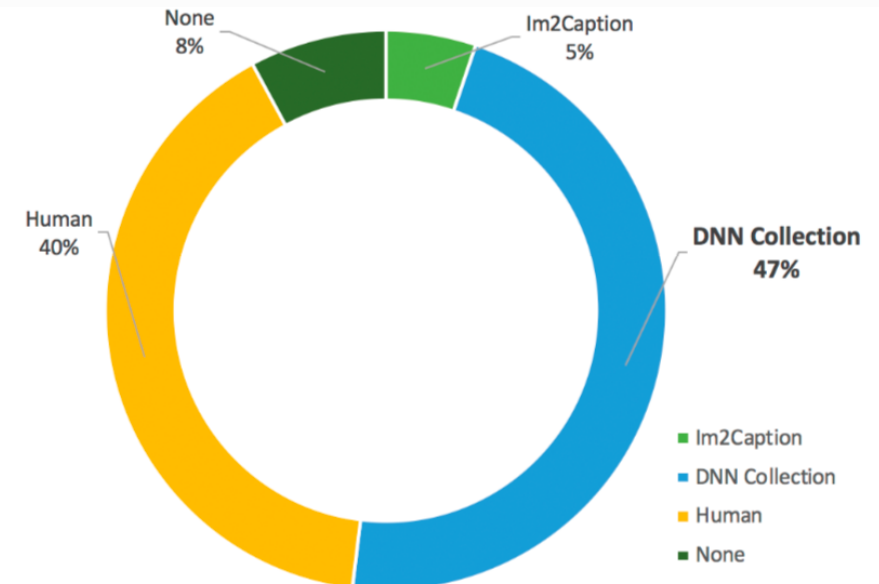
**Our method:** This graph titled homeownership rate (%) united states shows a rising trend in rate over year. The X-Axis goes from 1900 to 2015. The Y-Axis goes from 45 to 65. The end of the graph continues to rise and the graph shows few fluctuations.

**Human 1:** Between 19000 and 1940 the homeownership rate was relatively stable, with a large jump in 1940 into a steady incline through 2000.

**Human 2:** THE POPULATION OF DECREASE TO INCREASE.

**Human 3:** Line graph showing an upward trend in the percent of homeowners in the United States

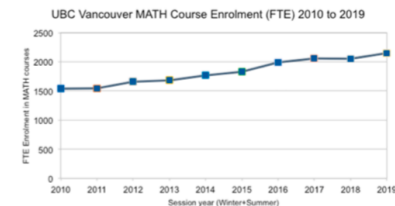
**Im2Caption:** looks stable trend of u s concern interest among from 2012 2012.



# Experiments

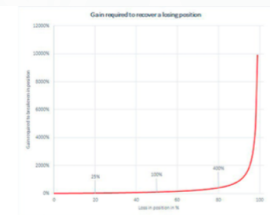
## Visually Impaired Users

Prompt	DNN	Microsoft	Twitter
The alt-text for the graphic improved my understanding of the graphic	<b>5.232</b>	4.493	4.362
The alt-text for the graphic improved my understanding of the Tweet	<b>5.29</b>	4.362	4.406
It seems that some of the information in the Tweet and some of the information in the alt-text don't match	<b>3.957</b>	4.652	4.638



**tweet:** I put this graph together for a short piece that will appear in the upcoming edition of the First Year Math and Stats in Canada newsletter. Even with hard caps on some of our courses, there is still an uptick for 2019. FTE faculty counts are almost constant over this time period.

**alt text (source DNN):** This graph titled ubc vancouver math course enrolment ( fte ) 2010 to 2019 2500 shows a rising trend in enrolment in math courses over year. The X-Axis goes from 2010 to 201 8. The Y-Axis goes to 2500. The end of the graph continues to rise and the graph show s few fluctuations.



**tweet:** The importance of stop losses.

Whether we invest fundamentally or quantitatively or by looking at charts it makes sense to imprint this graph in our minds.

**alt text (source MSCS):** A close up of text on a black background

# Conclusion

- In conclusion, we created an information graphic summarization framework that utilizes a collection of neural networks to formulate a summary.
- The framework utilizes the visual, textual, positional, and size characteristics within several multimodal neural networks.
- In our experimentation, we showed the effectiveness of our models and were able to generate summaries that were preferred even over some human provided summaries.
- Importantly, people with visual impairments rated our summaries as more helpful in understanding Tweets containing line graphs over other alternatives.