Towards Better Analysis of Deep Convolutional Neural Networks

Mengchen Liu, Jiaxin Shi, Zhen Li, Chongxuan Li, Jun Zhu, Shixia Liu

1

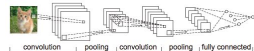## Problem: Understanding deep CNNs

Fig. 2. The typical architecture of a CNN.

- Tens or hundreds of **layers**, thousands of **neurons**, millions of **connections**
- Many functional components whose values and roles are not well understood either as individuals or as a whole
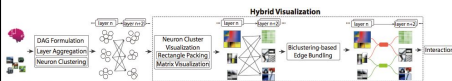
2

Solution: Overview CNNVis
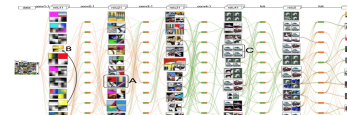

Fig. 4. CNNVis pipeline.

Fig. 1. CNNVis, a visual analytics toolkit that helps experts understand, diagnose, and refine deep CNNs.

3

Design Process: Overview

- Collaborated with 6 domain experts (3 of them are co-authors)
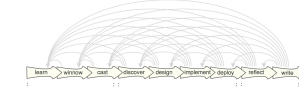- Iterative process over 12 months

[Fig 2. Sedlmair, Michael, Miriah Meyer, and Tamara Munzner. "Design study methodology: Reflections from the trenches and the stacks." IEEE transactions on visualization and computer graphics 18.12 (2012): 2431-2440.]

4

Design Process-Discover: Requirements

- R1 - Providing an overview of the learned features of neurons.
- R2 - Interactively modifying the neuron clustering results.
- R3 - Exploring multiple facets of neurons.
- R4 - Revealing how low-level features are aggregated into high- level features.
- R5 - Examining the debugging information.

5

Design: Fulfilling the requirements

- R1 - Providing an overview of the learned features of neurons.
- R4 - Revealing how low-level features are aggregated into high-level features.

- **DAG formulation to address R1 & R4**
  - Nodes = Neurons
  - Edge = Connection between neurons
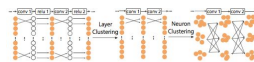
Fig. 5. Illustration of the DAG formulation.

6

Design: Fulfilling the requirements

- R3 - Exploring multiple facets of neurons.

- **Neuron cluster visualization module (R3)**
  - Learned features vis via Rectangle packing
  - Activations as matrix vis

7

Design: Fulfilling the requirements

- R3 - Exploring multiple facets of neurons.

- **Neuron cluster visualization module (R3)**
  - **Learned features vis via Rectangle packing**
  - Activations as matrix vis

8
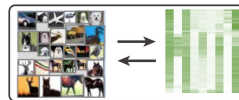
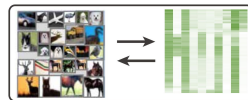
Neuron cluster visualization module

- **Learned features vis via Rectangle packing**
  - Each neuron represented as a rectangle
  - Computationally heavy for 100s - 1000s neurons

Fig. 7. Illustration of hierarchical rectangle packing.

9

Design: Fulfilling the requirements

- R3 - Exploring multiple facets of neurons.

- **Neuron cluster visualization module (R3)**
  - Learned features vis via Rectangle packing
  - Activations as matrix vis

10

Design: Fulfilling the requirements

- R3 - Exploring multiple facets of neurons.

- **Neuron cluster visualization module (R3)**
  - Learned features vis via Rectangle packing
  - **Activations as matrix vis**

11
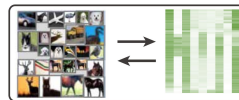
Neuron cluster visualization module

- **Activation as Matrix Vis**
  - Average activation of each neuron encoded
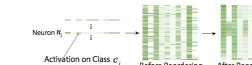 - Reordered based on cosine similarity

Fig. 8. Illustration of matrix reordering: (a) before reordering; (b) after reordering.

12

Design: Fulfilling the requirements

- R4 - Revealing how low-level features are aggregated into high-level features.

- **Biclustering-based edge bundling R4**
  - Avoid visual clutter by aggregation
  - + and - weights separated by color

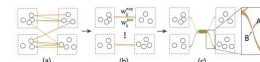Fig. 9. Illustration of biclustering-based edge bundling.

13

Interaction

14

Interaction

No slides for this part because

"A picture is worth a thousand words"

15

Interaction

No slides for this part because

If "A picture is worth a thousand words"

A video with minimum 24FPS is minimum

   Yx24x1000 words

where Y is the length of the the video in seconds

Video time…

16

Evaluation

- No Formal Eval
- 2 Case studies with the two of the experts conducted
- They both found the tool extremely useful
- Interesting findings through case study

Activation Space of a Layer
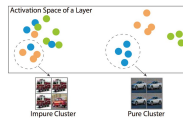
Impure Cluster   Pure Cluster

Fig. 12. Illustration of an "impure" cluster and a "pure" cluster.

17

Strength

- Very detailed and comprehensive approach
- Interesting choices such as bi-clustering, rectangle packing, matrix vis
- A lot of functionalities
- Very intuitive images clarifying difficult concepts to absorb by just reading

18

Weakness

- No formal evaluation
- Vague use of the term "Expert"
- 3/6 of the experts are also co-authors and participated in the case studies
- "we use two colored regions (green and red) to ….between the number of positive edges and of negative edges" ???

19

Weakness

- Clustering is debatable since it is not perfect, might lead to misrepresentation
- Some ideas were not clarified in the paper, had to contact the author...

Dear Mahdi,

1. Is CNNVis's source code, publicly available? Or is there any other version of this tool available to demo? The reason I ask this question is that I was unable to do some of the interactions mentioned in the paper and shown on the video using the web demo version. Some of the features I was unable to use were, dragging a neuron from one cluster to another, exploring more neurons in a cluster, or for instance the control panel is missing and some more.
Sorry, currently no…

2. I have noticed the edges for some of the clusters are missing. I read in your paper, you mention that the biclustering method of edges emphasize on the edges with high absolute values. Is that why we see no input or output edge for the first cluster under "relu5" in figure 6 of the paper?
Yes, we filter out some edges. Filters can be found in the video (right in the screen)

3. Each cluster is represented by 4 rectangle/square in the overview. Are those the top 4 neurons based on the activation score?
No, they are selected by their distance to the cluster center (less is better)

Best,
Mengchen

20

Weakness

- Videos and available online tool do not match
- The letter "C" is used to refer to the classes on page 5, and to the clusters on page 6

21

zbohom   القلبي   пока   xudafiz

Adiós   Ciao   **Good bye!**   حفلف   再见

farvel   الوداع   Au revoir   नमस्कार   từ biệt

Wiedersehen   さようなら   বিদায়   mějte se

22