Graph Convolutional Networks for Multiple Concept Detection from Medical Radiology Images

Erin Brown[1] and Yuxi Ke[1]
[1] Stanford University, CA 94305

The 2019 ImageCLEF Challenge

ROCO_CLEF_01527

- C3539923 adcor
- C0021156 dens incisivus
- C1962945 radiogr
- C1548003 radiograph
- C0026367 dens molaris
- C1561543 year
- C0024687 inferior maxillary bone
- C0043299 x-ray procedure

Fig. 1 An example from the dataset. [1]

- 56,629 training and 14,157 validation images from the PubMed Open Access subset of the Radiology Objects in Context (ROCO) dataset
- For a given image, predict multiple concepts from a list of 5,528 strings

Related Work

- CNN-RNN: Multi-label image classification with a learnt joint image-label embedding [2]
- GCN: a variant of CNN that operates directly on graphs and learns hidden layer representations that encode both graph structure and node features [3]
- ML-GCN: GCN structure adapted for multi-label image recognition [4]

Model Architecture

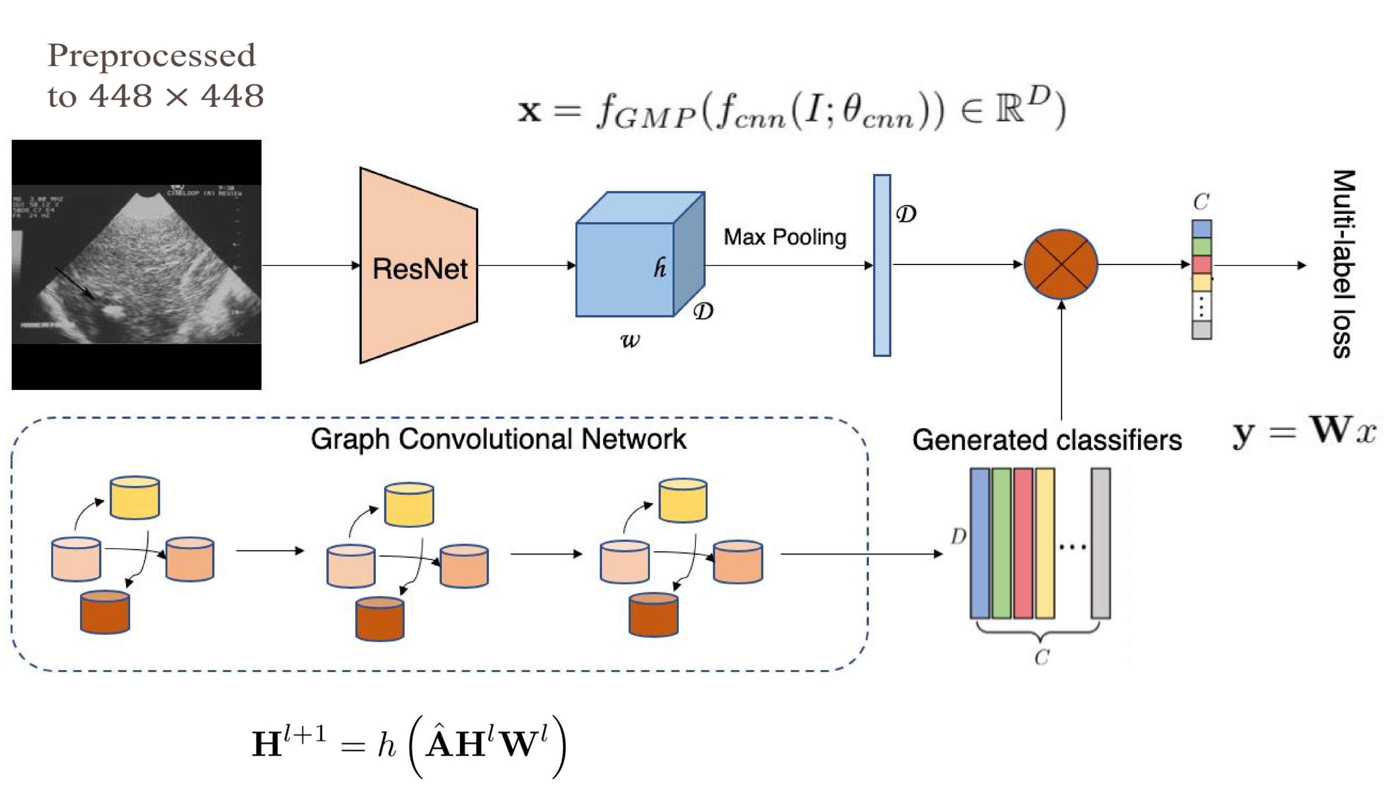


Fig. 2 The overall architecture of our model. Reproduced with minor modifications from [4]

Data Preprocessing

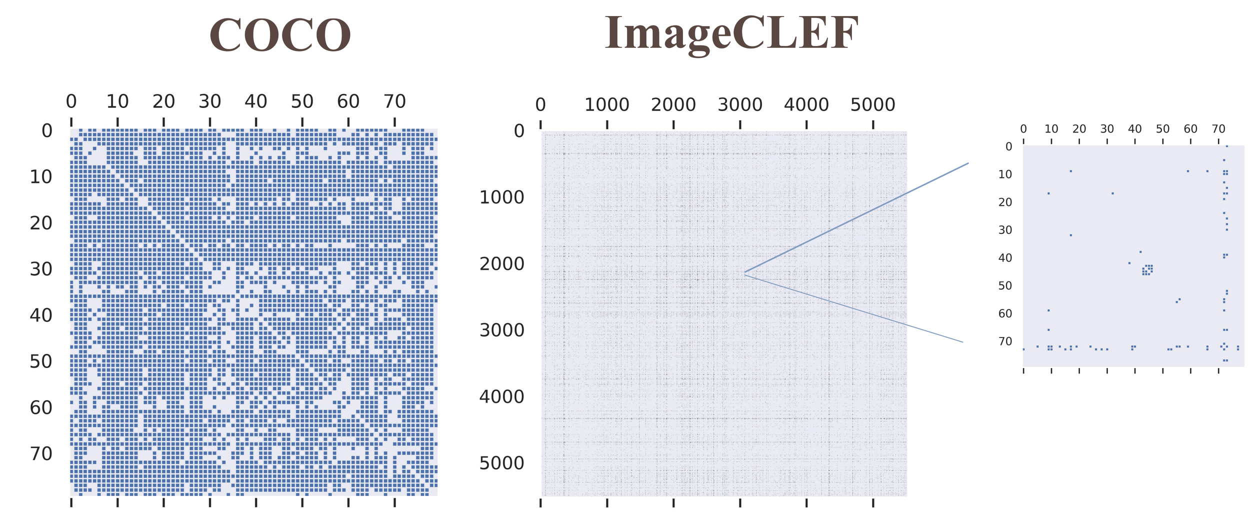


Fig. 3 Adjacency matrices from the COCO and ImageCLEF dataset, with a representative zoom in on the right.

- Concepts are represented by a 300-dimensional random vector drawn from uniform distributions, statistics comparable to GloVe.
- Adjacency matrix is computed from the co-occurrence frequencies of concepts followed by binarization and reweighting.
- Co-occurrence between the ImageCLEF concepts are sparse compared to MS-COCO.
- Rare-occurring concepts are prevalent.

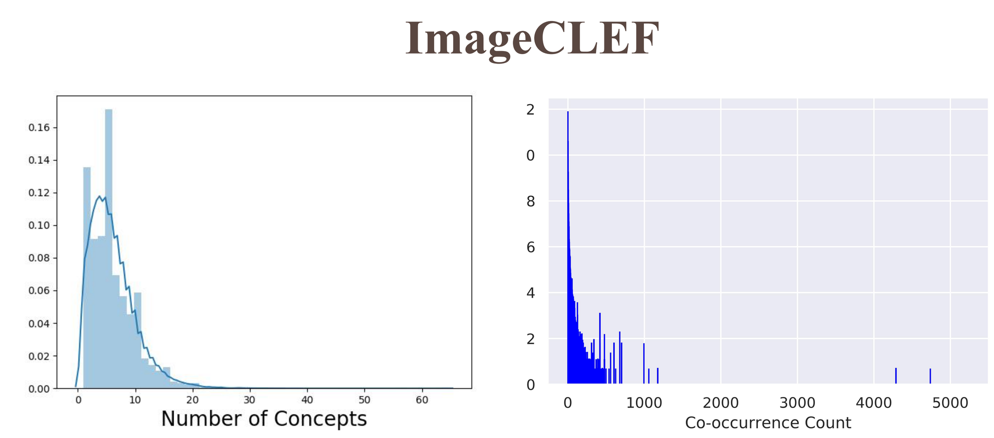


Fig. 4 Histograms showing rare occurring and low co-occurrence concepts from the ImageCLEF dataset.

Baseline

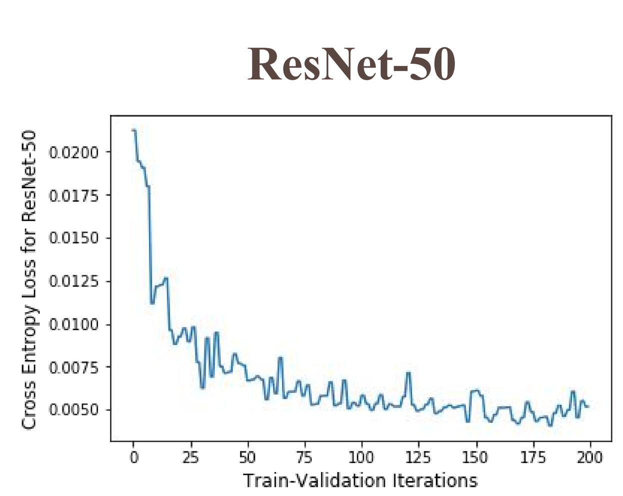


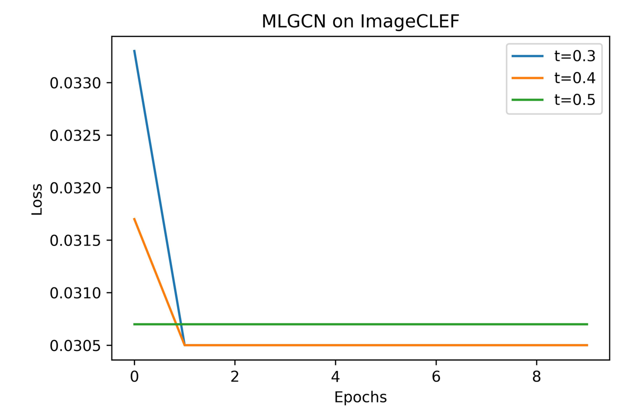
Fig. 5 Cross Entropy loss over training the baseline model.

- Pretrained a 50-layer ResNet with a binary flag label matrix.
- Multi-label loss

$$\mathcal{L} = \sum_{c=1}^{C} y^c \log\left(\sigma\left(\hat{y}^c\right)\right) + \left(1 - y^c\right) \log\left(1 - \sigma\left(\hat{y}^c\right)\right)$$

- Loss decreased from 0.025 to 0.005 over the train-validate steps.

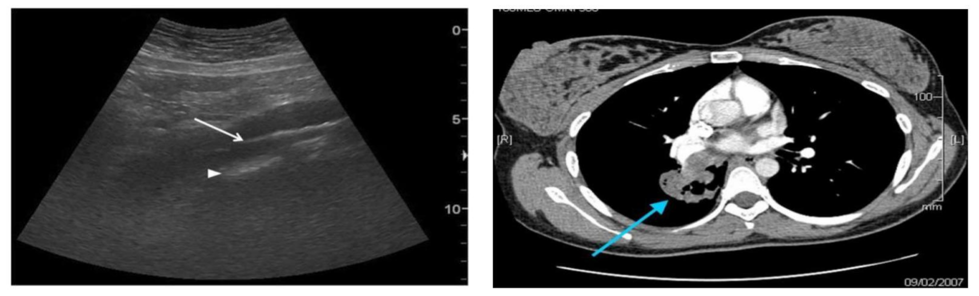
Training Process



- The GCN model was verified on the MS-COCO dataset.
- Difficulty in training: encountering loss plateuau
- The best mAP achieved is 0.185

Fig. 6 Multi-label loss over training ML-GCN in ImageCLEF with different adjacency matrix binarization thresholds.

Examples of Output



C0003483 trunk of systemic arterial tree
C0038925 flap created in surgery
C0549207 bone tissue of vertebra
C0012737 dissection procedures
C0000726 abdominopelvis
C1321564 strips

C1440080 alveolar
C3554497 prominent trabeculations
C0729203 omnipaque
C4084203 it improved
C1548915 sialogram

C0002978 x-ray of the blood vessel
C0040395 tomogr
C4255126 angiograms
C0034579 pantomogr
C1548816 angiogram
C3891555 traditional angiogram
C0024109 lungs pair
C0040405 x-ray computer assisted tomography
C3539779 sdtm-procedur

C0038128 diagnostic coloring material
C4284399 dehydration technique
C0038128 diagnostic coloring material
C0039991 incision of thorax
C0039990 thoracostomies
C0132840 norisen

Fig. 7 Representative model outputs.

Future Plans

- Optimize the adjacency matrix construction method for large and sparse labels.
- Expore mechanisms for concept representation via word phrase embeddings.

Acknowledgement and References

We thank Lu Yang, who is not enrolled in CS231n but contributed in discussion. Lu Yang is a Bioengineering PhD student at Stanford.

[1] Imageclef 2019 caption. https://www.imageclef. org/2019/medical/caption.
[2] J. Wang, Y. Yang, J. Mao, Z. Huang, C. Huang, and W. Xu. Cnn-rnn: A unified framework for multi-label image classi- fication. In The IEEE Conference on Computer Vision and Pattern Recognition (CVPR), June 2016.
[3] T. N. Kipf and M. Welling. Semi-supervised classification with graph convolutional networks. CoRR, abs/1609.02907, 2016.
[4] Z. Chen, X. Wei, P. Wang, and Y. Guo. Multi-label im_x0002_age recognition with graph convolutional networks. CoRR, abs/1904.03582, 2019.