Graph Convolutional Neural Network as a Chest X-Ray images classifier

Tan Ngoc Pham Dzung Tri Bui An Vo

University of Information Technology (UIT)
Vietnam National University Ho Chi Minh City (VNU-HCM)

Table of Contents

Table of Contents

Motivation

- In present days of medical diagnosis, **factor identification** for a specific sickness is crucial to give out appropriate and in-time cure.

Motivation

- In present days of medical diagnosis, **factor identification** for a specific sickness is crucial to give out appropriate and in-time cure.
- Manual diagnosis with CXR images is a **specialized-specific labour** requiring high specialty doctors from medical image analysis.

Motivation

- In present days of medical diagnosis, **factor identification** for a specific sickness is crucial to give out appropriate and in-time cure.
- Manual diagnosis with CXR images is a **specialized-specific labour** requiring high specialty doctors from medical image analysis.
- Recent advances in classifying and segmenting medical images have achieved **remarkable performance** on various tasks [1].

[1] Vo and Pham (2021). Detecting COVID-19 and Pneumonia with Chest X-Ray images using Deep Convolutional Neural Networks. In *Introduction to Computer Vision*. https://github.com/DTA-UIT /Detect-COVID19

Motivation

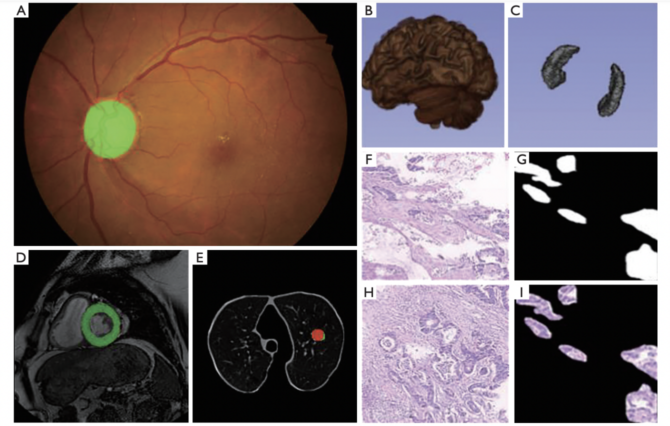


Figure: Deep learning application in medical image analysis[2].

[2] Cai, L., Gao, J., & Zhao, D. (2020). A review of the application of deep learning in medical image classification and segmentation. *Annals Of Translational Medicine*, *8*(11), 713. doi:10.21037/atm.2020.02.44

Motivation

- A crucial need on **visually explained** breakthroughs to practically deploy for computer-aided design (CAD) used in hospitals and medical centers[2].

[2] Lee, L., Kanthasamy, S., Ayyalaraju, R. S., & Ganatra, R. (2019). The Current State of Artificial Intelligence in Medical Imaging and Nuclear Medicine. *BJR open*, 1(1), 20190037. https://doi.org/10.1259/bjro.20190037

Motivation

- A crucial need on **visually explained** breakthroughs to practically deploy for computer-aided design (CAD) used in hospitals and medical centers[2].
- Medical images are **explicit**, especially annotated ones, due to patients' privacy and lack of expertised resources.

[2] Lee, L., Kanthasamy, S., Ayyalaraju, R. S., & Ganatra, R. (2019). The Current State of Artificial Intelligence in Medical Imaging and Nuclear Medicine. *BJR open*, 1(1), 20190037. https://doi.org/10.1259/bjro.20190037

Motivation

- A crucial need on **visually explained** breakthroughs to practically deploy for computer-aided design (CAD) used in hospitals and medical centers[2].
- Medical images are **explicit**, especially annotated ones, due to patients' privacy and lack of expertised resources.
- CNN- and Attention-based methods require tons of computational resources (GPUs) and data, i.e., DETR [3], ResNet[4].

---

[2] Lee, L., Kanthasamy, S., Ayyalaraju, R. S., & Ganatra, R. (2019). The Current State of Artificial Intelligence in Medical Imaging and Nuclear Medicine. *BJR open*, 1(1), 20190037. https://doi.org/10.1259/bjro.20190037

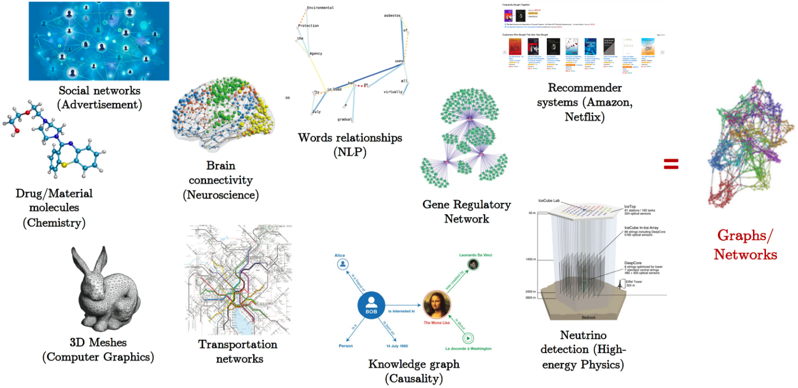
[3] Carion, N., Massa, F., Synnaeve, G., Usunier, N., Kirillov, A., & Zagoruyko, S. (2020, August). End-to-end object detection with transformers. In *European conference on computer vision* (pp. 213-229). Springer, Cham.

[4] K. He, X. Zhang, S. Ren and J. Sun, "Deep Residual Learning for Image Recognition," *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2016, pp. 770-778, doi: 10.1109/CVPR.2016.90.

Table of Contents

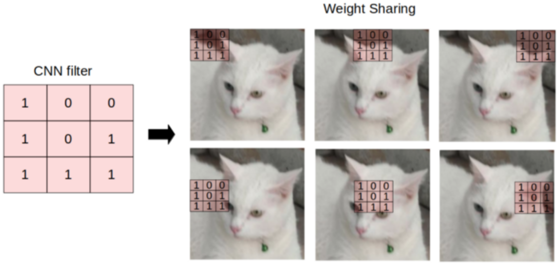
Graph Neural Network

- Previously, graph-based DL methods are used to solve problems that requires to understand **connections** and **contributions from entities** within the problem formation.



Graph Neural Network

- In Convolutional Neural Networks (CNNs), **kernels (filters)** are applied to synthesize neighboring pixels from original image to generalize image features.



Graph Neural Network

- In Convolutional Neural Networks (CNNs), **kernels (filters)** are applied to synthesize neighboring pixels from original image to generalize image features.
- In Graph Neural Networks (GNNs), neighboring pixels are also synthesized to generalize images by using **nodes and edges** instead of kernels.

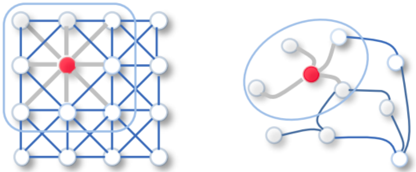


Figure: CNNs (left) versus GNNs (right) on neighboring pixels synthesis.

Assumption: nearer nodes often share more common features

Graph Neural Network

Strengths

- GNN is introduced to **overcome problems of overfitting and class imbalance**.
- GNN is usually **independent of GPUs** and performs faster but better compared to other DL-bases methods (based on evidence from experimental results, i.e., Xu et al. (2019)[5], Ma et al. (2021)[6]).

---

[5] Xu, P., Joshi, C. K., & Bresson, X. Multi-graph transformer for free-hand sketch recognition (2019). *arXiv preprint arXiv:1912.11258*

[6] Ma, C., Yang, F., Li, Y., Jia, H., Xie, X., & Gao, W. (2021). Deep human-interaction and association by graph-based learning for multiple object tracking in the wild. *International Journal of Computer Vision*, 129(6), 1993-2010.

Graph Neural Network

Strengths

- GNN is introduced to **overcome problems of overfitting and class imbalance**.
- GNN is usually **independent of GPUs** and performs faster but better compared to other DL-bases methods (based on evidence from experimental results, i.e., Xu et al. (2019)[5], Ma et al. (2021)[6]).

Weaknesses

- Graph construction is a **problem-variant** work and depends a lot on **expertised knowledge**.

---

[5] Xu, P., Joshi, C. K., & Bresson, X. Multi-graph transformer for free-hand sketch recognition (2019). *arXiv preprint arXiv:1912.11258*

[6] Ma, C., Yang, F., Li, Y., Jia, H., Xie, X., & Gao, W. (2021). Deep human-interaction and association by graph-based learning for multiple object tracking in the wild. *International Journal of Computer Vision*, 129(6), 1993-2010.

Graph Convolutional Network

- To apply GNN for CV problems, Graph Convolutional Network[7] (GCN) is proposed as a means to extract features from images.

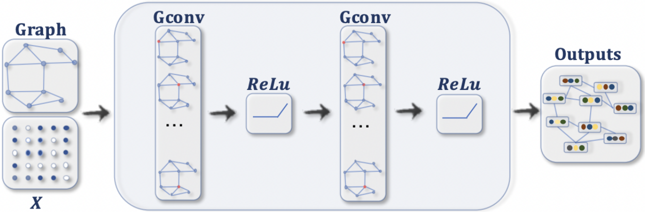


Figure: A typical Graph Convolutional Network architecture

[7] Kipf, T. N., & Welling, M. (2016). Semi-supervised classification with graph convolutional networks. *arXiv preprint arXiv:1609.02907*.

- To apply GNN for CV problems, Graph Convolutional Network[7] (GCN) is proposed as a means to extract features from images.
- In this work, GCN convolutional layers are designed to exchange information toward each other the same as the mechanism in *message-passing neural network* (MPNN)[8].

[7]Kipf, T. N., & Welling, M. (2016). Semi-supervised classification with graph convolutional networks. *arXiv preprint arXiv:1609.02907.*

[8]Gilmer, J., Schoenholz, S. S., Riley, P. F., Vinyals, O., & Dahl, G. E. (2017, July). Neural message passing for quantum chemistry. In *International conference on machine learning* (pp. 1263-1272). PMLR.

Graph Convolutional Network

MPNN comprises of 2 primary steps: **UPDATE** and **READOUT**.

MPNN comprises of 2 primary steps: **UPDATE** and **READOUT**.

1. Feature vectors $h(v)$ of node v , $\forall v \in V$ are sequentially updated via 2 minor phases:

MPNN comprises of 2 primary steps: **UPDATE** and **READOUT**.

1. Feature vectors $h(v)$ of node v , $\forall v \in V$ are sequentially updated via 2 minor phases:
   - Aggregation: neighboring nodes features are synthesized into one feature vector.
   - Update: aggregated feature vector would be assigned to node v.

Graph Convolutional Network

MPNN comprises of 2 primary steps: **UPDATE** and **READOUT**.

1. Feature vectors $h(v)$ of node v , $\forall v \in V$ are sequentially updated via 2 minor phases:
   - Aggregation: neighboring nodes features are synthesized into one feature vector.
   - Update: aggregated feature vector would be assigned to node v.

   $\rightarrow$ The update phase would be conducted for $\mathcal{T}$ iterations before moving to final step.

Graph Convolutional Network

MPNN comprises of 2 primary steps: **UPDATE** and **READOUT**.

1. Feature vectors $h(v)$ of node v , $\forall v \in V$ are sequentially updated via 2 minor phases:
   - Aggregation: neighboring nodes features are synthesized into one feature vector.
   - Update: aggregated feature vector would be assigned to node v.

   $\rightarrow$ The update phase would be conducted for $\mathcal{T}$ iterations before moving to final step.
2. The final graph is pooled via a Log-Softmax activation.

Architecture construction

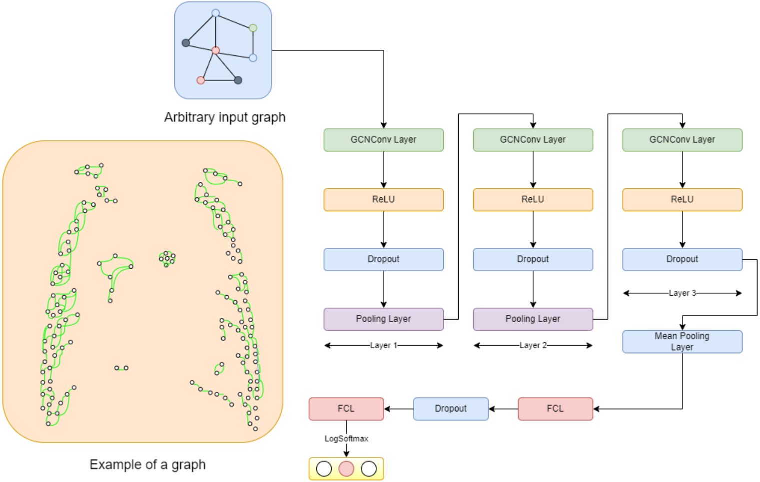


Figure: Our architecture used in this work

Table of Contents

Dataset

- To validate the applied method, we use "Labeled Optical Coherence Tomography (OCT) and Chest X-Ray Images for Classification dataset"[9] (6114 CXR images).

[9] Kermany, D., Zhang, K., & Goldbaum, M. (2018). Labeled optical coherence tomography (oct) and chest x-ray images for classification. *Mendeley data*, 2(2).

Dataset

- To validate the applied method, we use "Labeled Optical Coherence Tomography (OCT) and Chest X-Ray Images for Classification dataset"[9] (6114 CXR images).

Figure: Examples on dataset: Bacteria (left), Normal (middle), Virus (right).

[9] Kermany, D., Zhang, K., & Goldbaum, M. (2018). Labeled optical coherence tomography (oct) and chest x-ray images for classification. *Mendeley data*, 2(2).

Dataset

- To validate the applied method, we use "Labeled Optical Coherence Tomography (OCT) and Chest X-Ray Images for Classification dataset"[9] (6114 CXR images).

	Train	Val	Test	Total
Bacteria	2045	503	242	2790
Normal	1058	291	234	1583
Virus	1090	503	148	1741
Total	3193	1297	624	6114

Table: Distribution of labels on dataset

[9] Kermany, D., Zhang, K., & Goldbaum, M. (2018). Labeled optical coherence tomography (oct) and chest x-ray images for classification. *Mendeley data*, 2(2).

Methodological pipeline

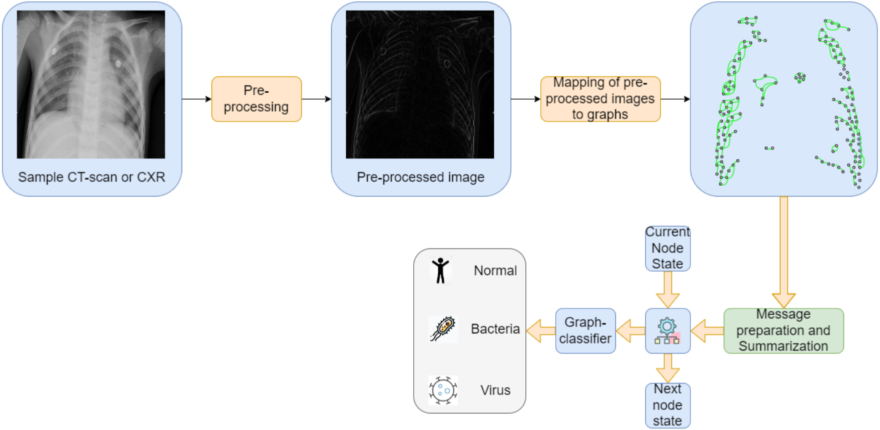


Figure: Experimental pipeline

Preprocessing

1. Original CXR images would be resized into 224x224 resolution.

Preprocessing

1. Original CXR images would be resized into 224x224 resolution.
2. Edge detection on resized CXR images.

Preprocessing

1. Original CXR images would be resized into 224x224 resolution.
2. Edge detection on resized CXR images.

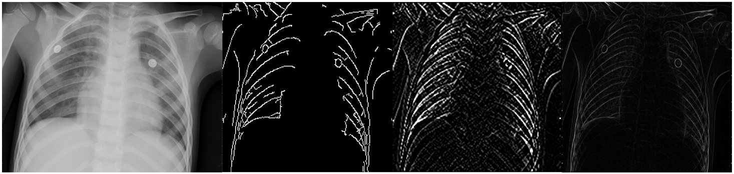


Figure: Comparison among edge detection methods. From left to right: Original image, Sobel image, Canny image, and Prewitt image.

Graph construction

Graphs from edge detected CXR images are constructed from 3 steps:

1. Pixels having grayscale intensity value of 128 is qualified as a node and feature of a node consists of the grayscale intensity of the corresponding pixels.

Graph construction

Graphs from edge detected CXR images are constructed from 3 steps:

1. Pixels having grayscale intensity value of 128 is qualified as a node and feature of a node consists of the grayscale intensity of the corresponding pixels.
2. Edge exists between 2 nodes representing neighboring pixels from original images.

Graph construction

Graphs from edge detected CXR images are constructed from 3 steps:

1. Pixels having grayscale intensity value of 128 is qualified as a node and feature of a node consists of the grayscale intensity of the corresponding pixels.
2. Edge exists between 2 nodes representing neighboring pixels from original images.
3. Node attributes, which are grayscale values, are normalized graph-wise.

Metrics

	Predicted 0	Predicted 1
Actual 0	TN	FP
Actual 1	FN	TP

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN}$$

$$PPV(Precision) = \frac{TP}{TP + FP}$$

$$Sensitivity(Recall) = \frac{TP}{TP + FN}$$

$$F1\ score = \frac{2 \times PPV \times Sensitivity}{PPV + Sensitivity}$$

Experimental configuration

- Graph Convolutional Network
  - Adam(learning rate = 1e-4)
  - batch size = 64
  - 50 epochs

Experimental configuration

- Graph Convolutional Network
  - Adam(learning rate = 1e-4)
  - batch size = 64
  - 50 epochs
- ResNet50, VGG16
  - pretrained on ImageNet
  - Adam (learning rate = 1e-3)
  - Reduce learning rate (factor = 0.1, patience = 3, min lr = 1e-6)
  - 50 epochs

Results comparison

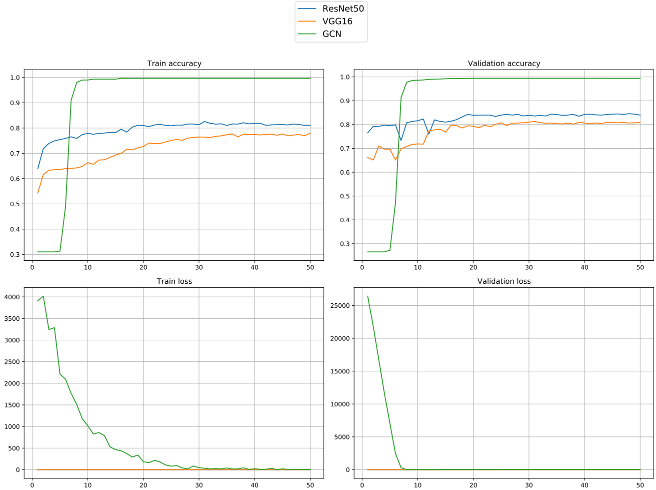


Figure: Training curve among Graph Convolutional Network(GCN), ResNet50, and VGG16.

Results comparison

	GCN	ResNet50	VGG16
Test accuracy	**96.75%**	83.81%	80.29%
Number of parameters	**37,955**	16,088,067	126,014,955
Training time (s)	9642 + 317	4715	5442

Table: Comparison among Graph Convolutional Network (GCN), ResNet50[10], and VGG16[11]

[10]He, K., Zhang, X., Ren, S., & Sun, J. (2016). Deep residual learning for image recognition. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 770-778).

[11]Simonyan, K., & Zisserman, A. (2014). Very deep convolutional networks for large-scale image recognition. *arXiv preprint arXiv:1409.1556*

Results comparison

		GCN	ResNet50	VGG16
PPV	Bacteria	**1.000**	0.875	0.765
	Normal	0.979	**0.982**	0.952
	Virus	**0.903**	0.656	0.712
Sensitivity	Bacteria	0.949	0.954	**0.958**
	Normal	**0.967**	0.717	0.683
	Virus	**1.000**	0.837	0.736
F1 score	Bacteria	**0.974**	0.913	0.851
	Normal	**0.973**	0.829	0.796
	Virus	**0.949**	0.735	0.724

Table: Comparison among Graph Convolutional Network(GCN), ResNet50, and VGG16.

Confusion matrix

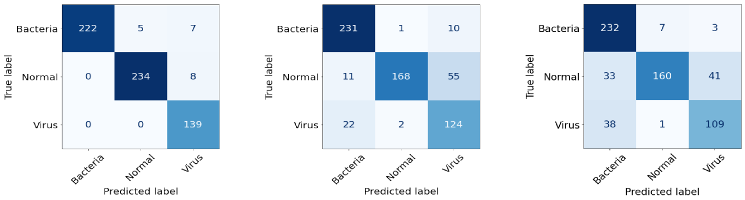


Figure: Confusion matrix among: GCN (left), ResNet50 (middle), VGG16 (right).

Table of Contents

Conclusion

- Graph Convolutional Network did achieve remarkable performance compared to other CNN-based methods on classifying CXR images for pneumonia.

Conclusion

- Graph Convolutional Network did achieve remarkable performance compared to other CNN-based methods on classifying CXR images for pneumonia.
- GCN overcame the problem of overfitting and class imbalance.

Conclusion

- Graph Convolutional Network did achieve remarkable performance compared to other CNN-based methods on classifying CXR images for pneumonia.
- GCN overcame the problem of overfitting and class imbalance.
- GCN is not an end-to-end method and requires specialised knowledge on constructing graphs for specific problems.

The end.