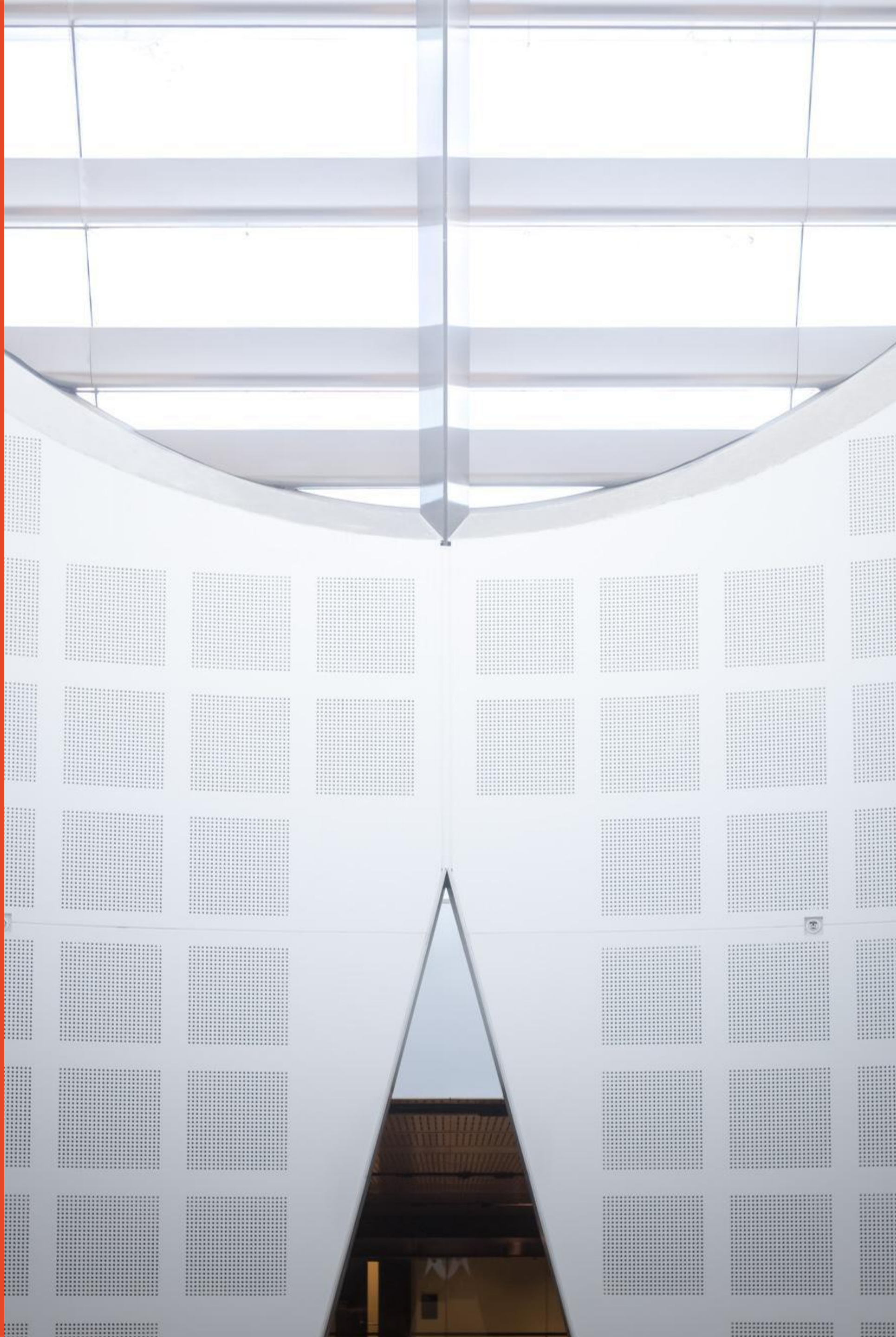
Deep Feature In-painting for Unsupervised Anomaly Detection in Radiography Images

Presented by
Tiange Xiang
For fulfilment of BCST (Adv) (Hons)
Supervisor: A/Prof. Weidong Cai

THE UNIVERSITY OF SYDNEY

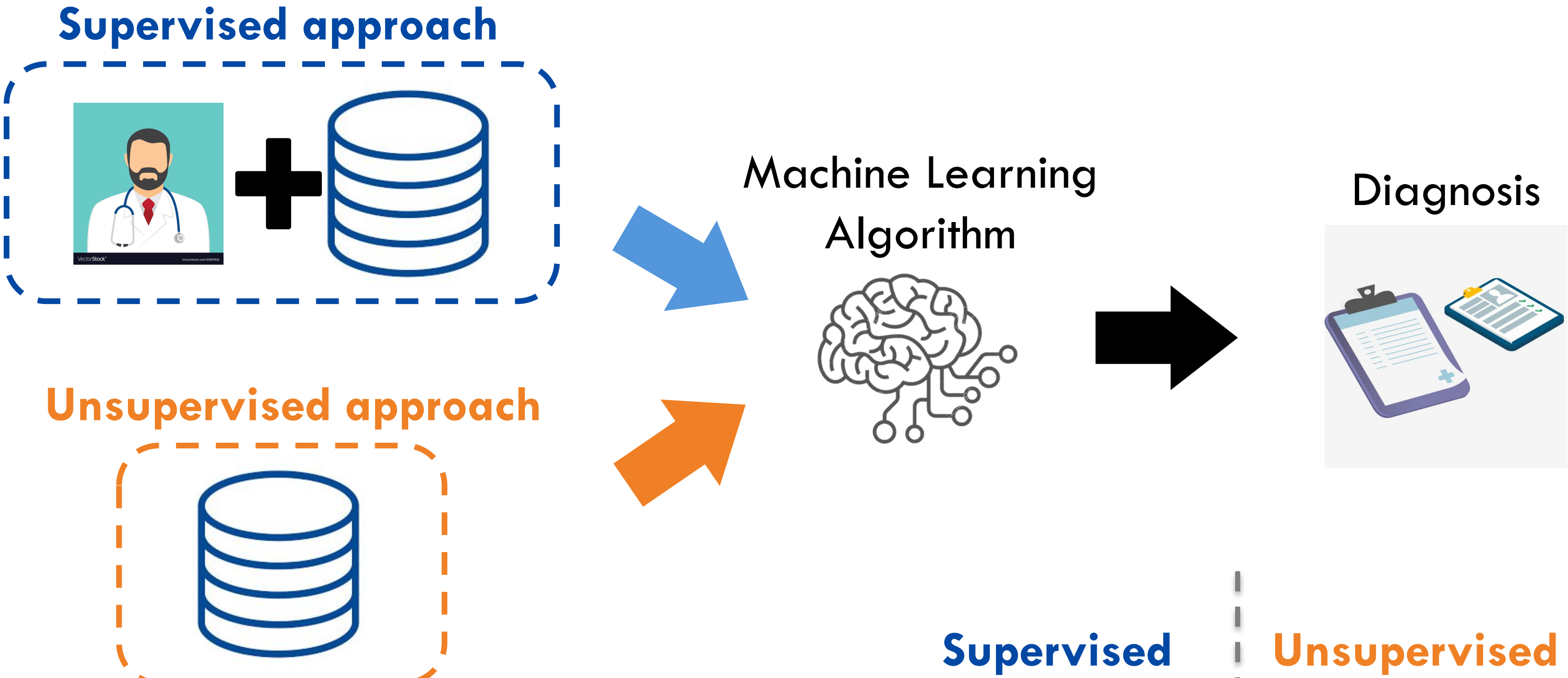


Outline

Motivation: Why ML for Medical Imaging?

- **Faster diagnosis/treatment.**
- **Less human intervention.**
- **Saves more lives.**

Motivation: Why Unsupervised Learning?



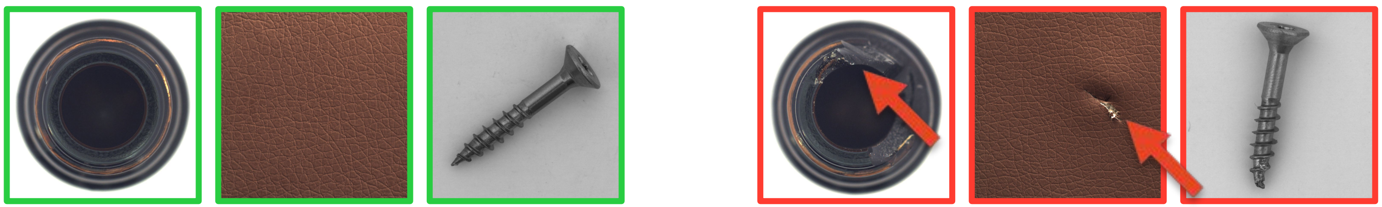
	Supervised	Unsupervised	
✗	Paired ground truth	**ZERO** ground truth needed	✓
✗	Human intelligence + Machine intelligence	Machine intelligence **ONLY**	✓

Motivation: Anomaly in Chest X-rays

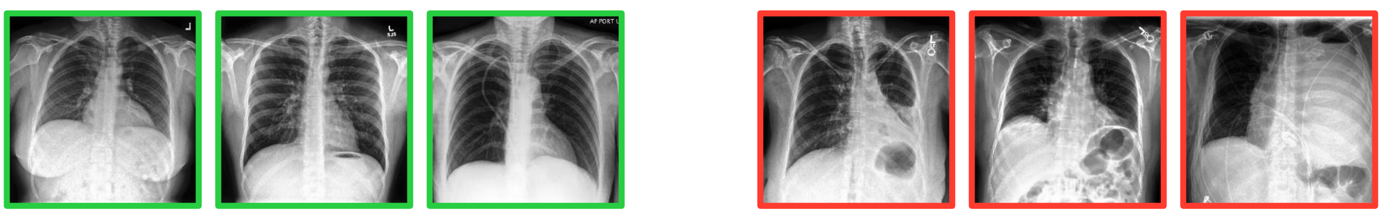
Normal

Abnormal

Anomaly Detection in Crowded Scenes
(photography images)



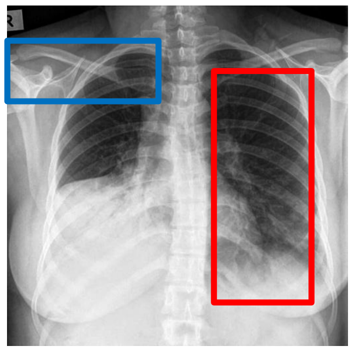
Anomaly Detection in Textures and Objects
(photography images)

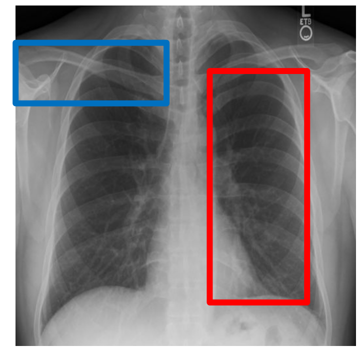


Anomaly Detection in Chest Anatomy
(**radiography images**)

Motivation: Unique Characteristics for Chest X-rays

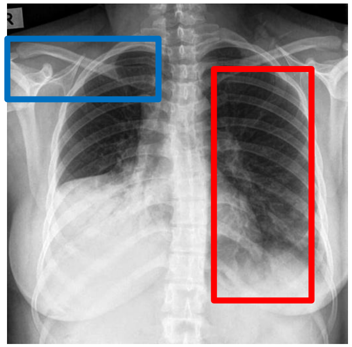
Radiography images

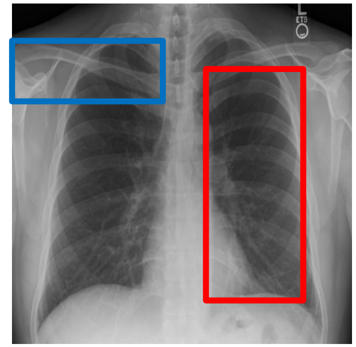




Motivation: Unique Characteristics for Chest X-rays

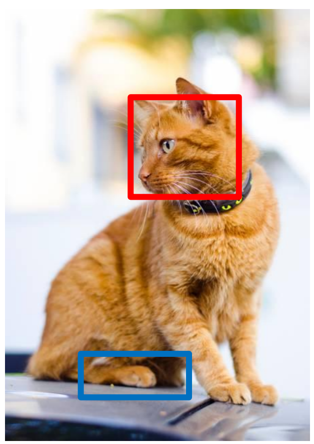
Radiography images

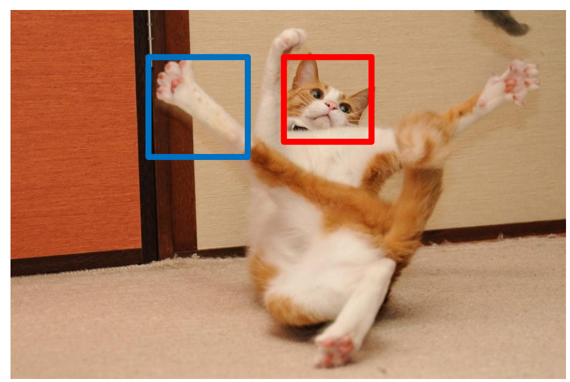




Consistent shapes/appearances and **fixed poses.**

Photography images

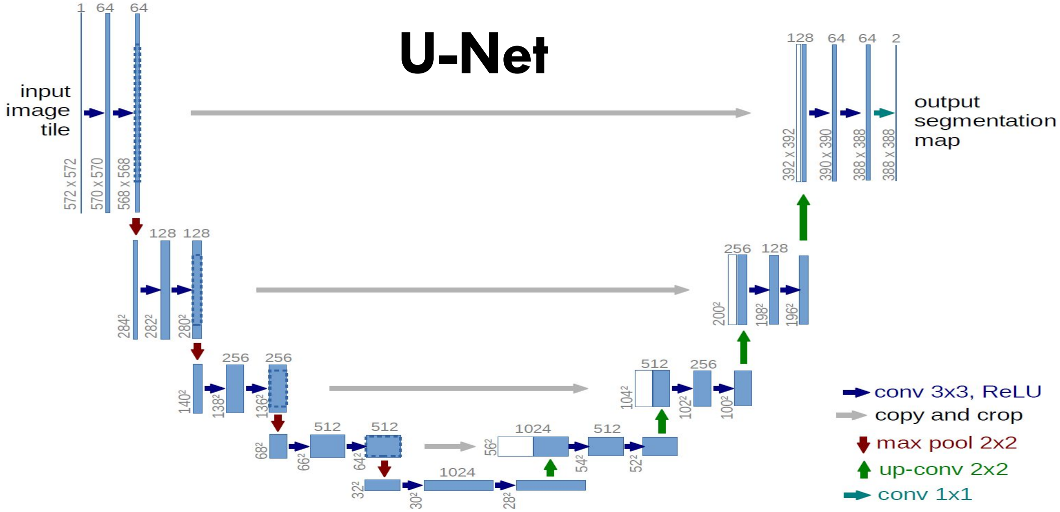




Outline

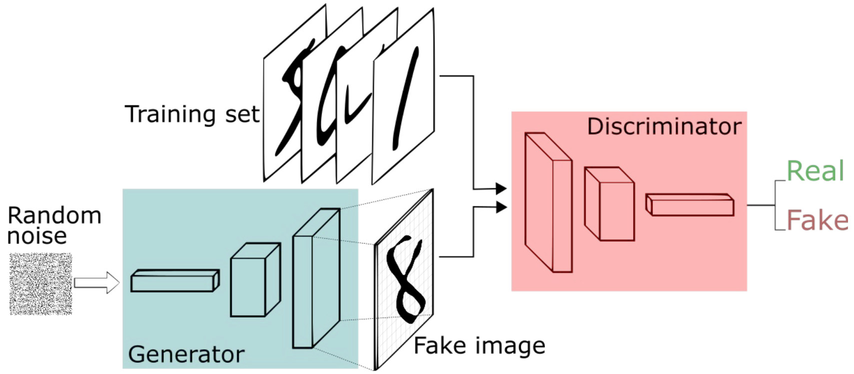
Literature: Preliminaries

U-Net



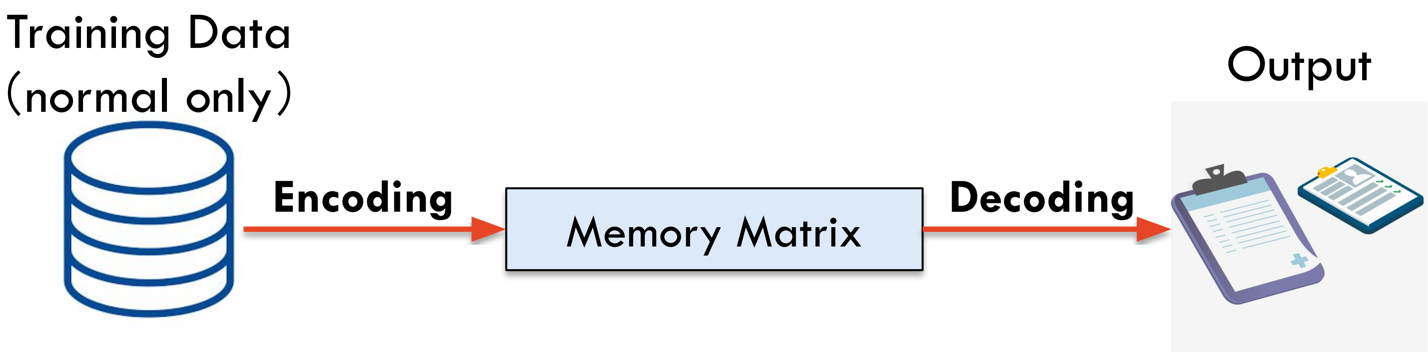
- Encoder + Decoder + Skip connections
- Pixel-to-pixel mapping

GAN



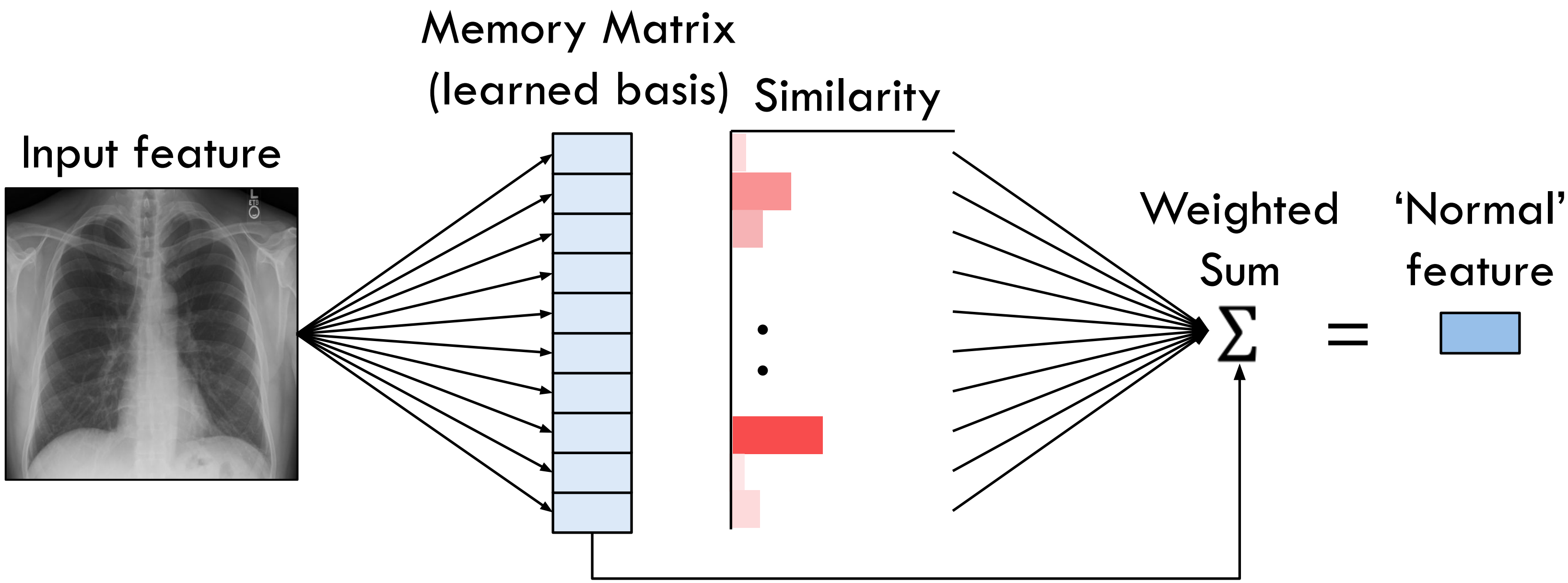
- Generator + Discriminator + Adversarial Learning
- Content generation

Literature: Baseline – MemAE (Gong *etal.*, ICCV 2019)



Literature: Baseline – MemAE (Gong *etal.*, ICCV 2019)

Feature Augmentation



Outline

1. Motivation
2. Background & Literature
3. **Problem Definition**
4. Methodology
5. Results
6. Discussion
7. Conclusion

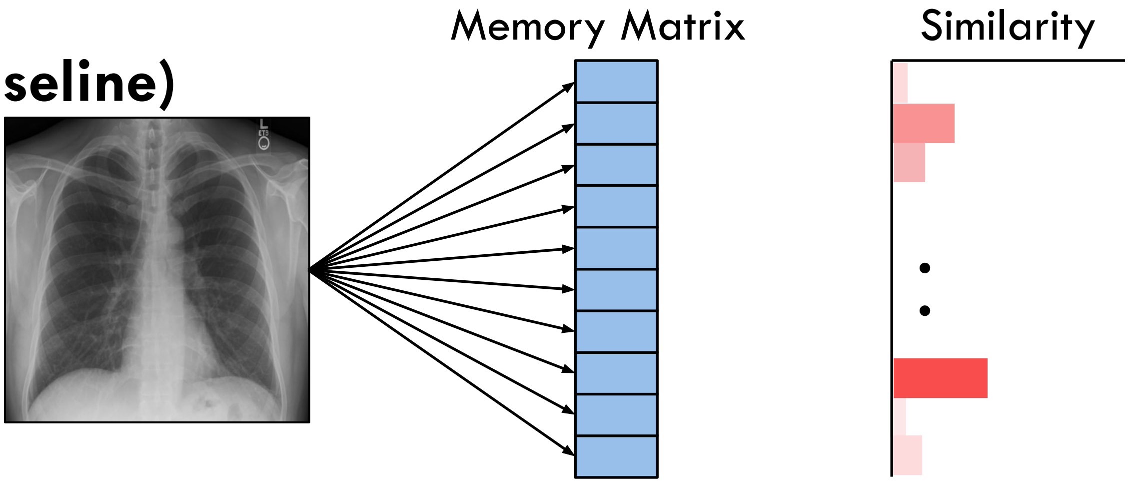
Problem Definition & Objectives

Algorithm Robustness	Robust to pixel distortions, mixed training dataset.
Methodology novelty	Proposal of multiple new techniques and strategies
Methodology Interpretability	Creation of an intuitive dataset to better interpretate ideas
Performance superiority	SOTA performances on public and challenging benchmarks
Evaluation correctness	Evaluation under the TRUE UAD protocol with the best results.

Outline

Methodology: Space-aware Memory

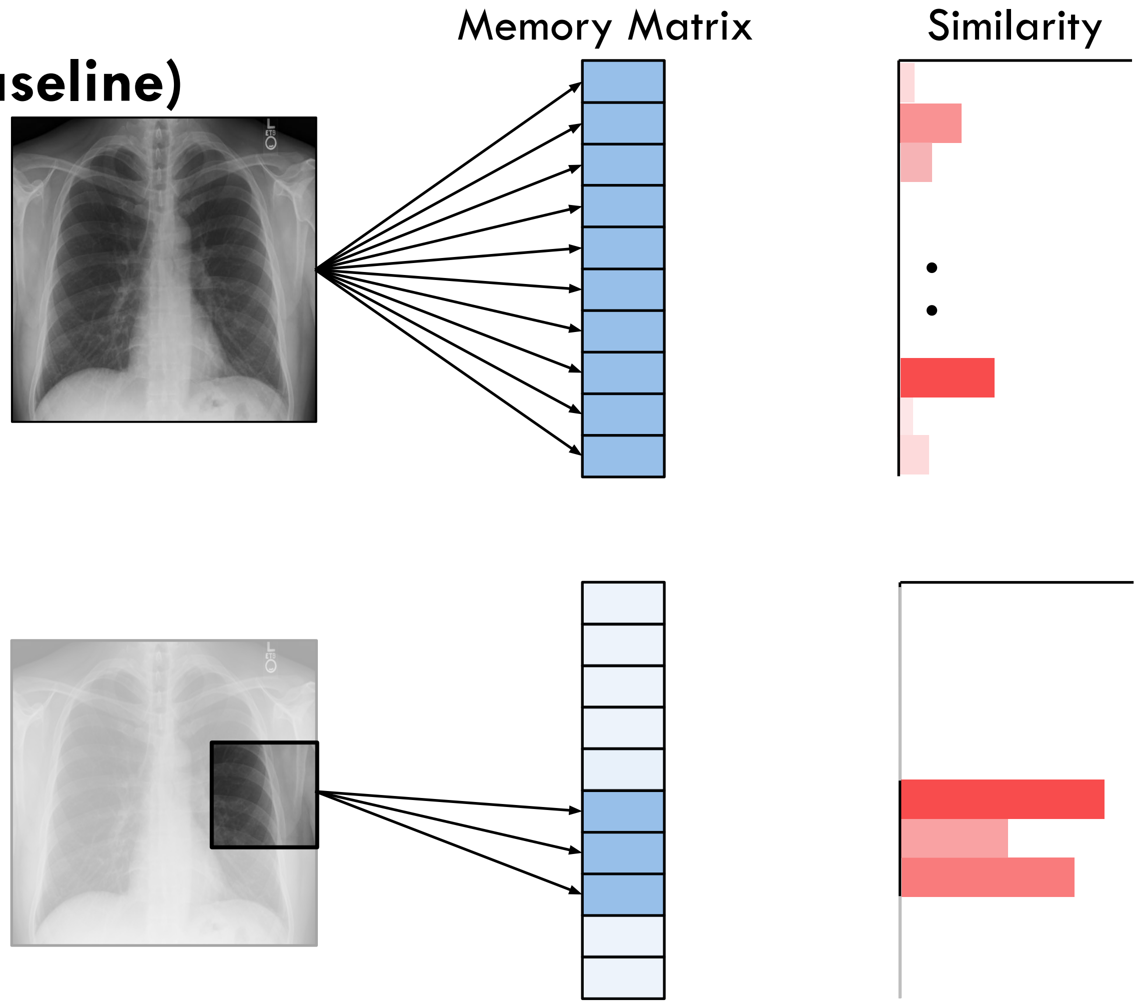
MemAE (Baseline)



Methodology: Space-aware Memory

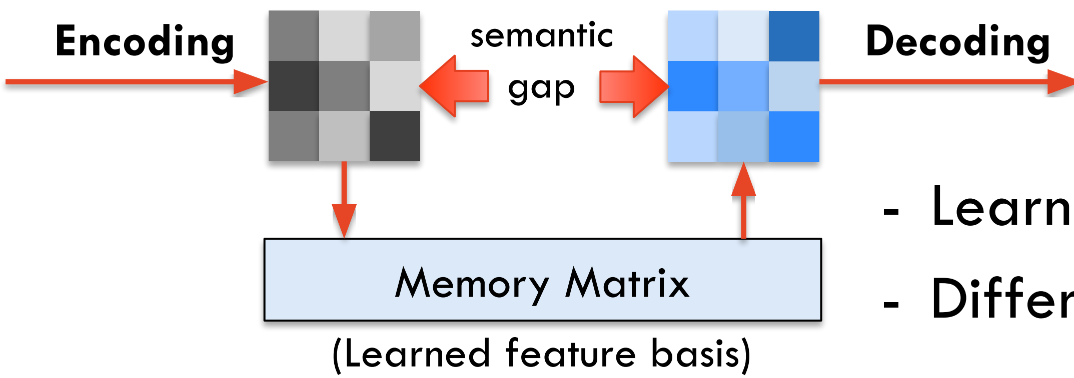
MemAE (Baseline)

Ours



Methodology: Memory Queue

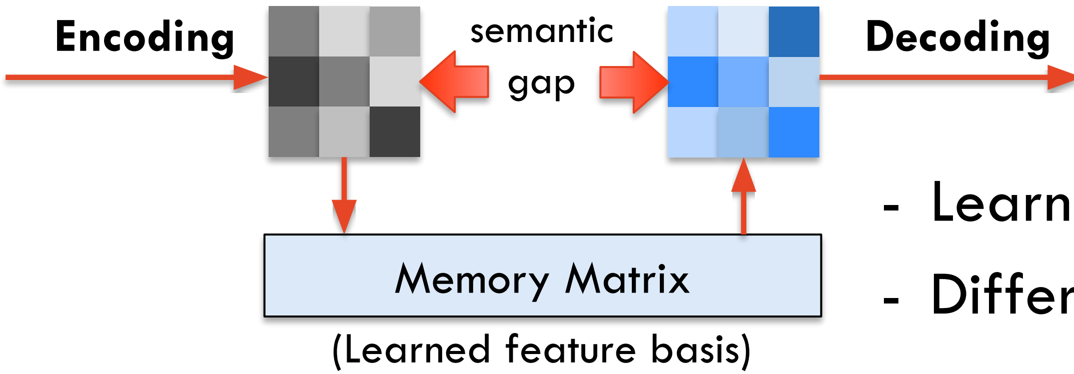
MemAE (Baseline)



- Learnable Matrix
- Different feature dist.

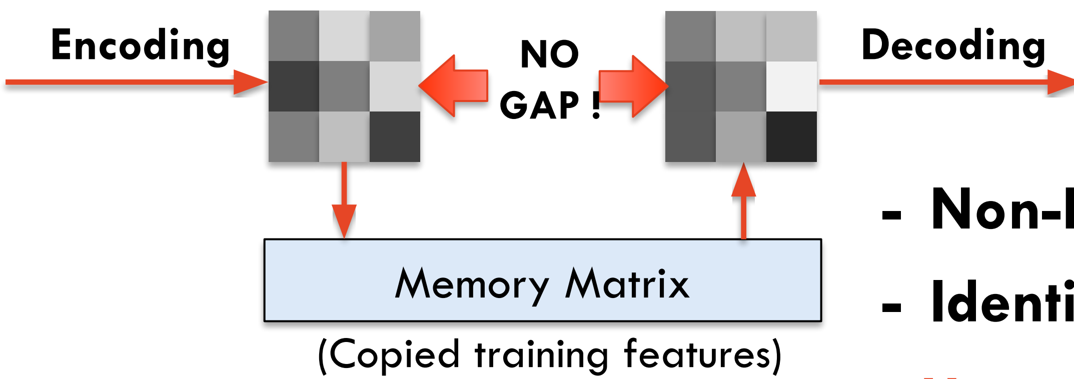
Methodology: Memory Queue

MemAE (Baseline)



- Learnable Matrix
- Different feature dist.

Ours

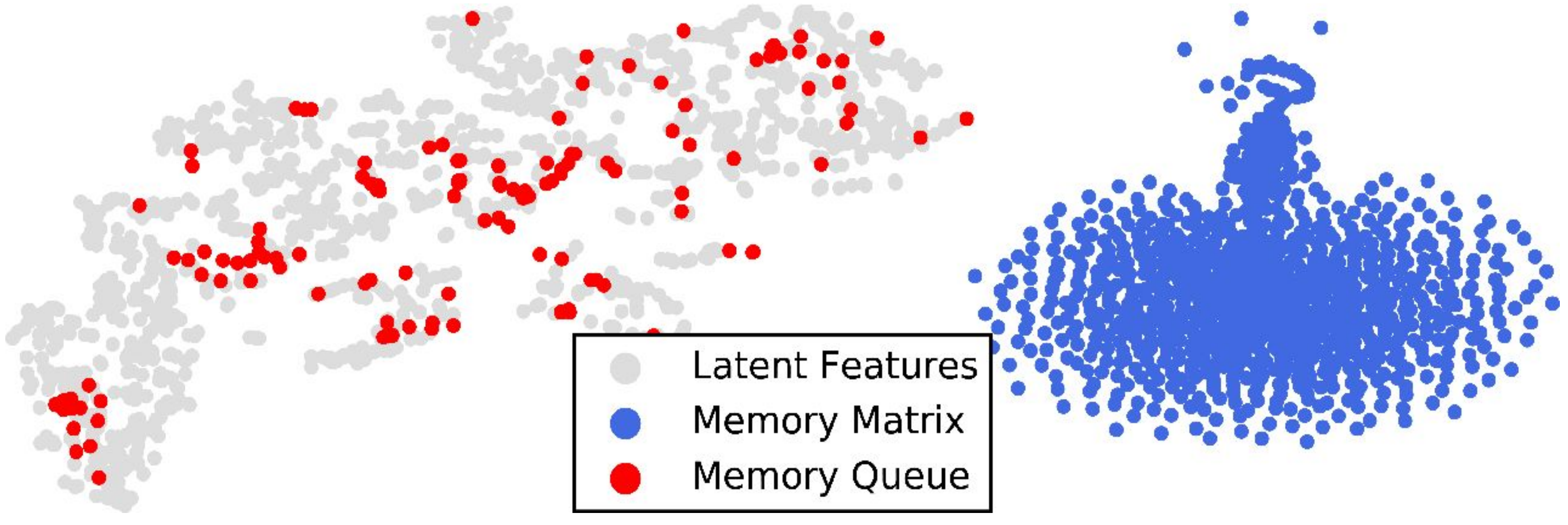


- **Non-Learnable** Matrix
- **Identical** feature dist.

How to prove?

Methodology: Memory Queue

t-SNE feature visualizations



Methodology: Memory Queue

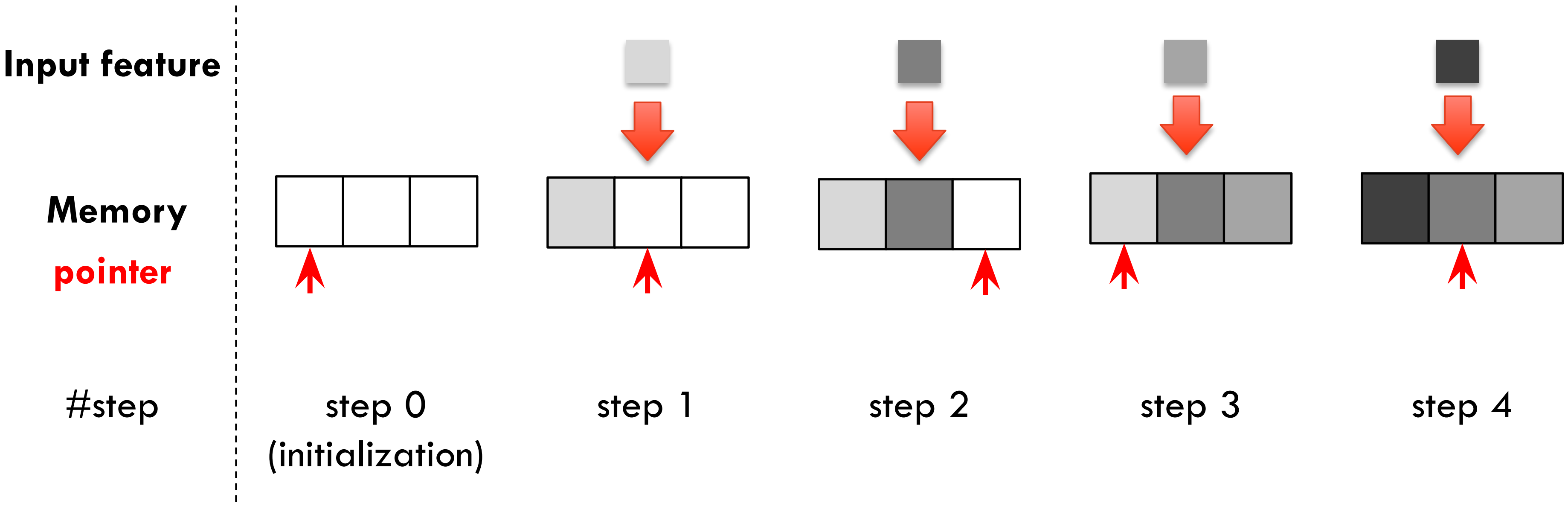
How to copy and paste?

- Memory matrix needs to be updated with most recent features.

- Refresh the entire matrix at every training step is inefficient.

Methodology: Memory Queue

Memory Queue Processing

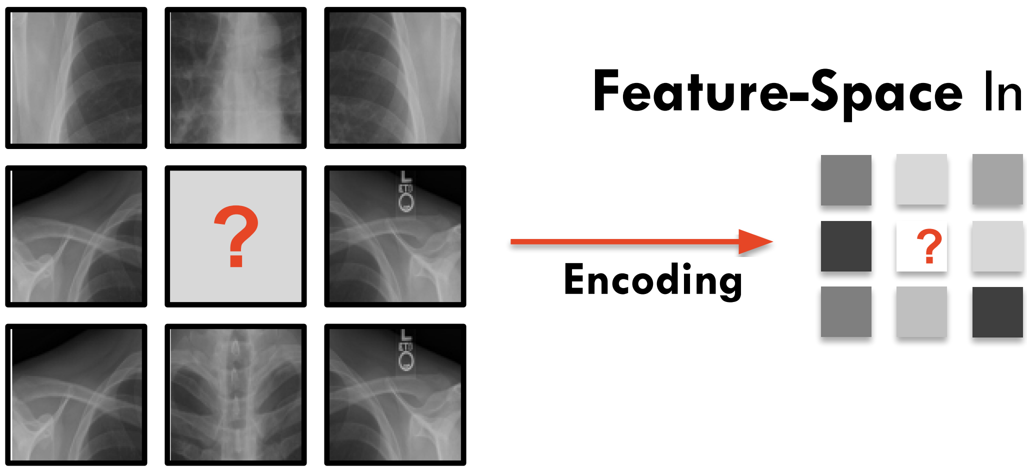


- First-in-first-out updating rule.
- Small learning rate helps.

Methodology: UAD as Feature-Space In-painting

Pixel-Space In-painting

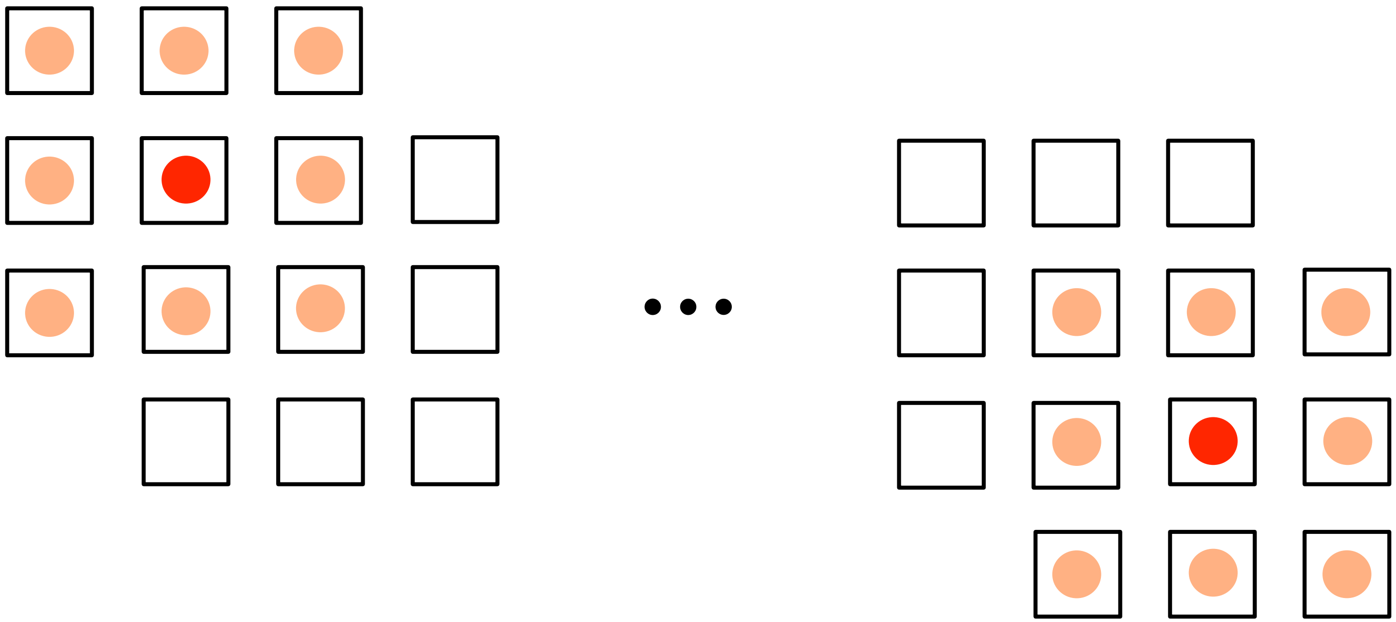
Feature-Space In-painting



Given the contextual information

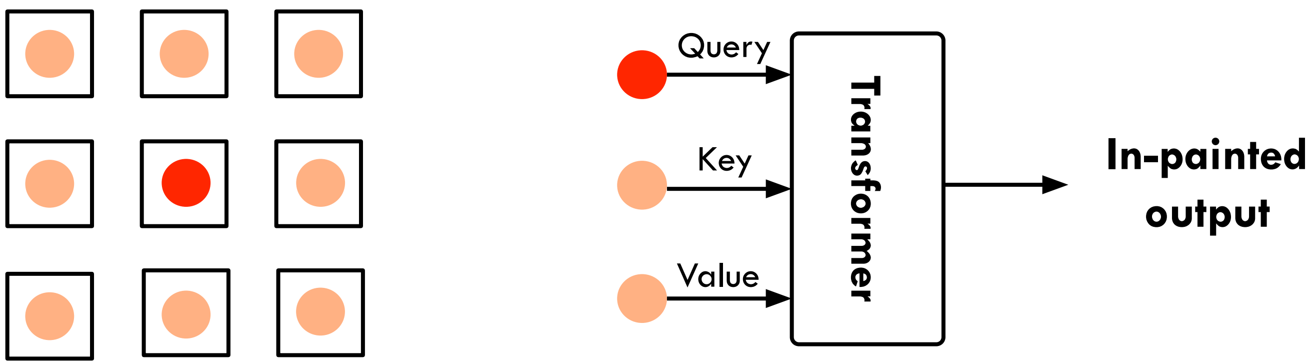
What does a normal patch look like?

Methodology: UAD as Feature-Space In-painting

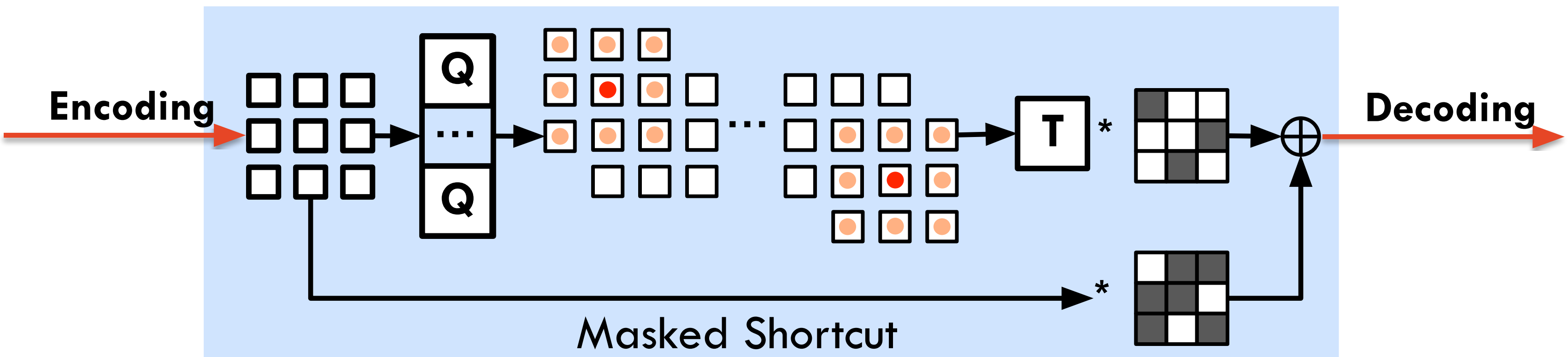


- Sliding window to traverse all patches
- Zero padding for out-of-range patches

Methodology: UAD as Feature-Space In-painting



Methodology: In-painting Block

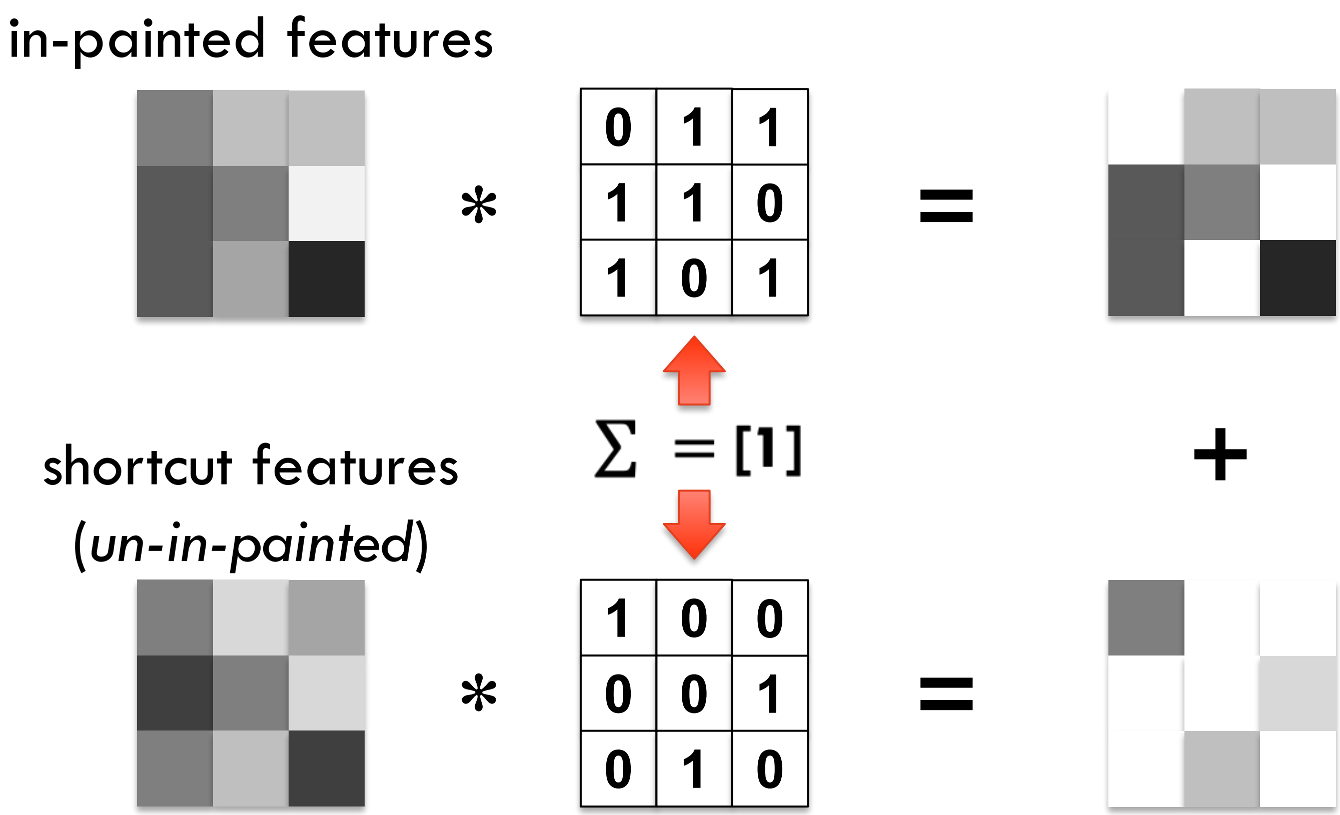


Q Memory Queue

T Transformer Layer

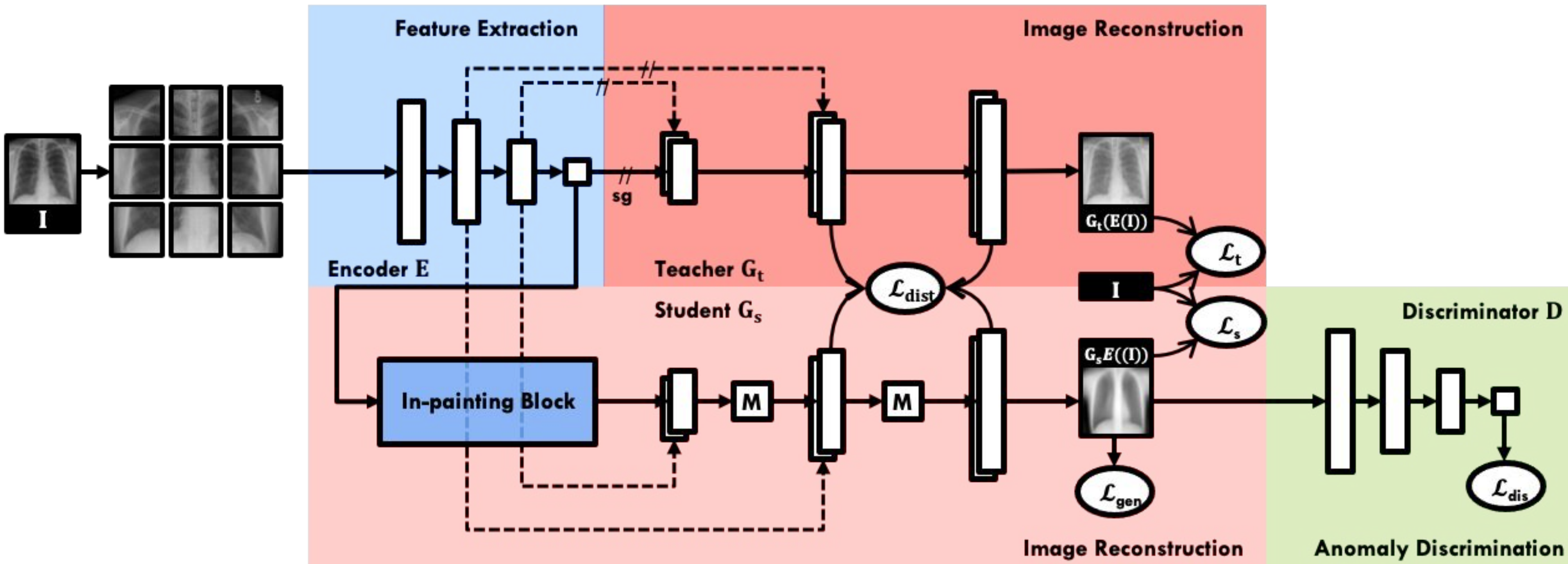
- Use shortcut for gradient preservation and better feature aggregation.
- Naive identity shortcut leads to degenerations.

Methodology: Masked Shortcut



Methodology: 🦑 SQUID 🦑

Space-aware memory **QU**eue for **I**n-painting and **D**etecting

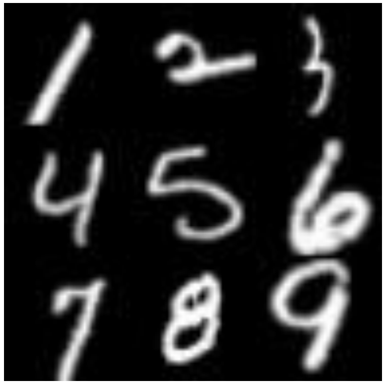


- GAN for adversarial learning
- Knowledge Distillation for regularization
- Generation quality as anomaly score
- Supervised by self-reconstruction losses.

Methodology: Creation of DigitAnatomy

Chest Anatomy

Digit Anatomy



Characteristics

- Consistent shape
- Fixed pose

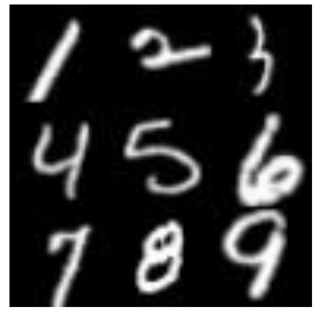
Benefits

- Intuitive demos ✓
- Easy development/debug ✓

Methodology: Creation of DigitAnatomy

Normal

(1-9 in order)

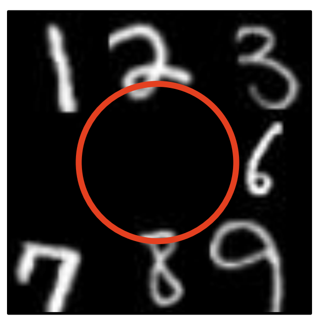


Abnormal

(novel digits)

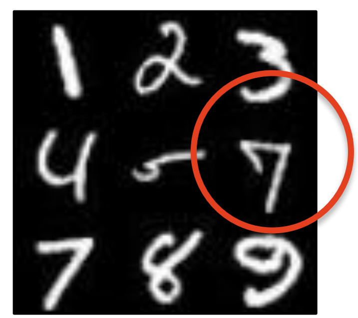
Abnormal

(missing digits)



Abnormal

(disorder digits)

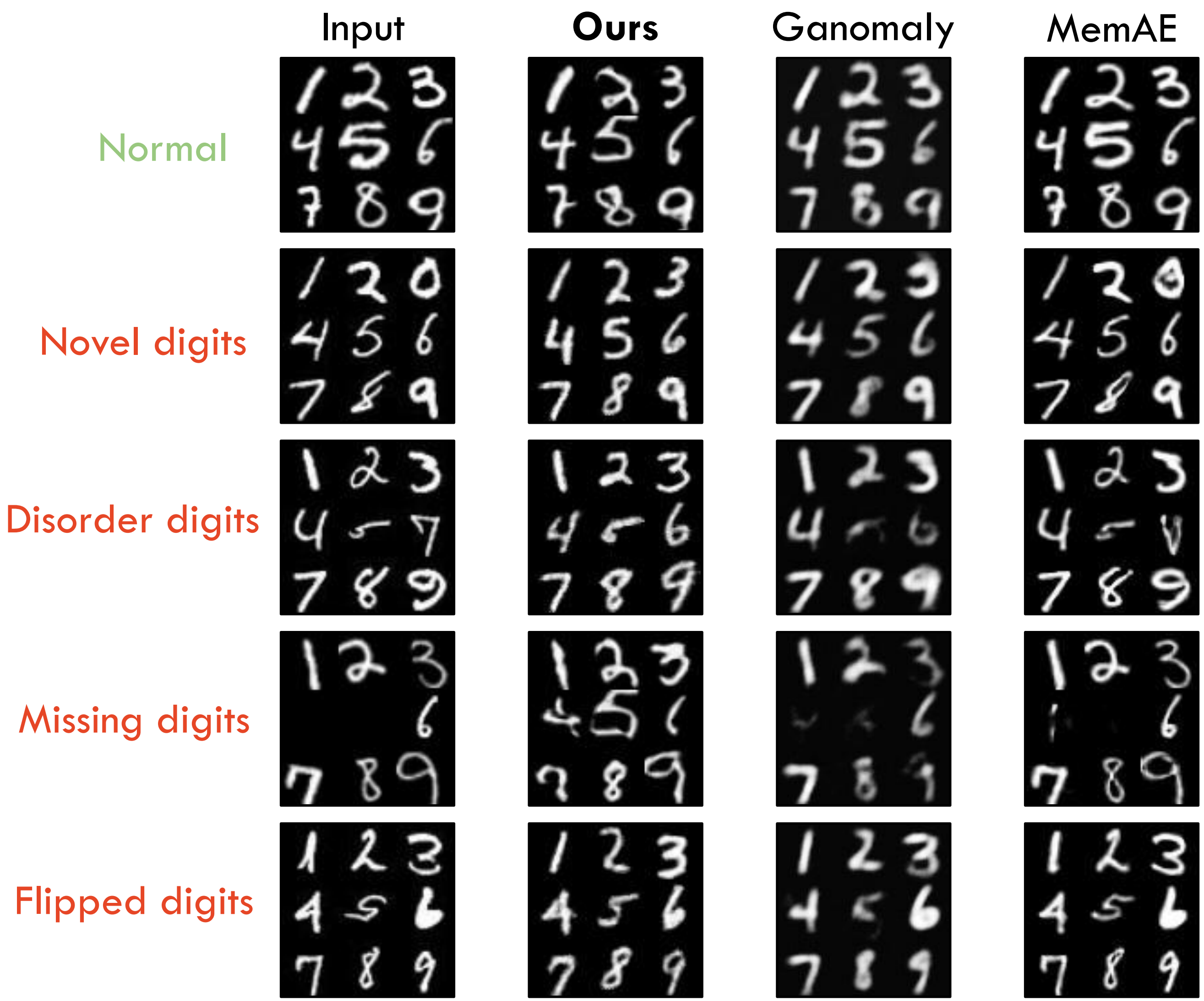


Abnormal

(flipped digits)

Outline

Results: Interpretations on DigitAnatomy



Results: Public Benchmarks

Zhang Lab Chest X-ray Dataset:

Disease include: *Pneumonia.*

Stanford CheXpert:

Disease include: *Cardiomegaly, Enlarged Cardiomegaly, Lung Lesion, Lung Opacity, Edema, Consolidation, Pneumonia, Atelectasis, Pneumothorax, Pleural Effusion, Pleural Other, Fracture.*

COVIDx:

Disease include: *Regular pneumonia, Covid19 infections.*

	ZhangLab	CheXpert	COVIDx
Train set (pos./neg.)	3783 / 1249	21171 / 4999	20802 / 7985
Val. set (pos./neg.)	100 / 100	19 / 14	300 / 100
Test set (pos./neg.)	390 / 234	250 / 250	300 / 100
#Anomalies	1	12	2
Difficulty level	⭐	⭐⭐	⭐⭐⭐

Results: Public Benchmarks

Quantitative Eval.

- AUC, Acc, F1 as metrics.
- Results of 3+ independent runs.
- >*5%* *AUC imp.* on ZhangLab.
- >*9%* *AUC imp.* on CheXpert.
- >*2%* *AUC imp.* on COVIDx.

TABLE 4.1. Results on the test sets of the ZhangLab dataset. Both average results and standard deviations are reported. † denotes the results taken from other literature.

ZhangLab	Ref & Year	AUC (%)	Acc (%)	F1 (%)
Auto-Encoder†	-	59.9	63.4	77.2
VAE† [36]	Arxiv'13	61.8	64.0	77.4
Ganomaly† [1]	ACCV'18	78.0	70.0	79.0
f-AnoGAN† [64]	MIA'19	75.5	74.0	81.0
MemAE [15]	ICCV'19	77.8±1.4	56.5±1.1	82.6±0.9
MNAD [55]	CVPR'20	77.3±0.9	73.6±0.7	79.3±1.1
SALAD† [90]	TMI'21	82.7±0.8	75.9±0.9	82.1±0.3
CutPaste [43]	CVPR'21	73.6±3.9	64.0±6.5	72.3±8.9
PANDA [58]	CVPR'21	65.7±1.3	65.4±1.9	66.3±1.2
M-KD [62]	CVPR'21	74.1±2.6	69.1±0.2	62.3±8.4
IF 2D [52]	MICCAI'21	81.0±2.8	76.4±0.2	82.2±2.7
PaDiM [10]	ICPR'21	71.4±3.4	72.9±2.4	80.7±1.2
IGD [8]	AAAI'22	73.4±1.9	74.0±2.2	80.9±1.3
SQUID	This work	**87.6±1.5**	**80.3±1.3**	**84.7±0.8**

TABLE 4.2. Results on the test sets of the CheXpert dataset. Both average results and standard deviations are reported.

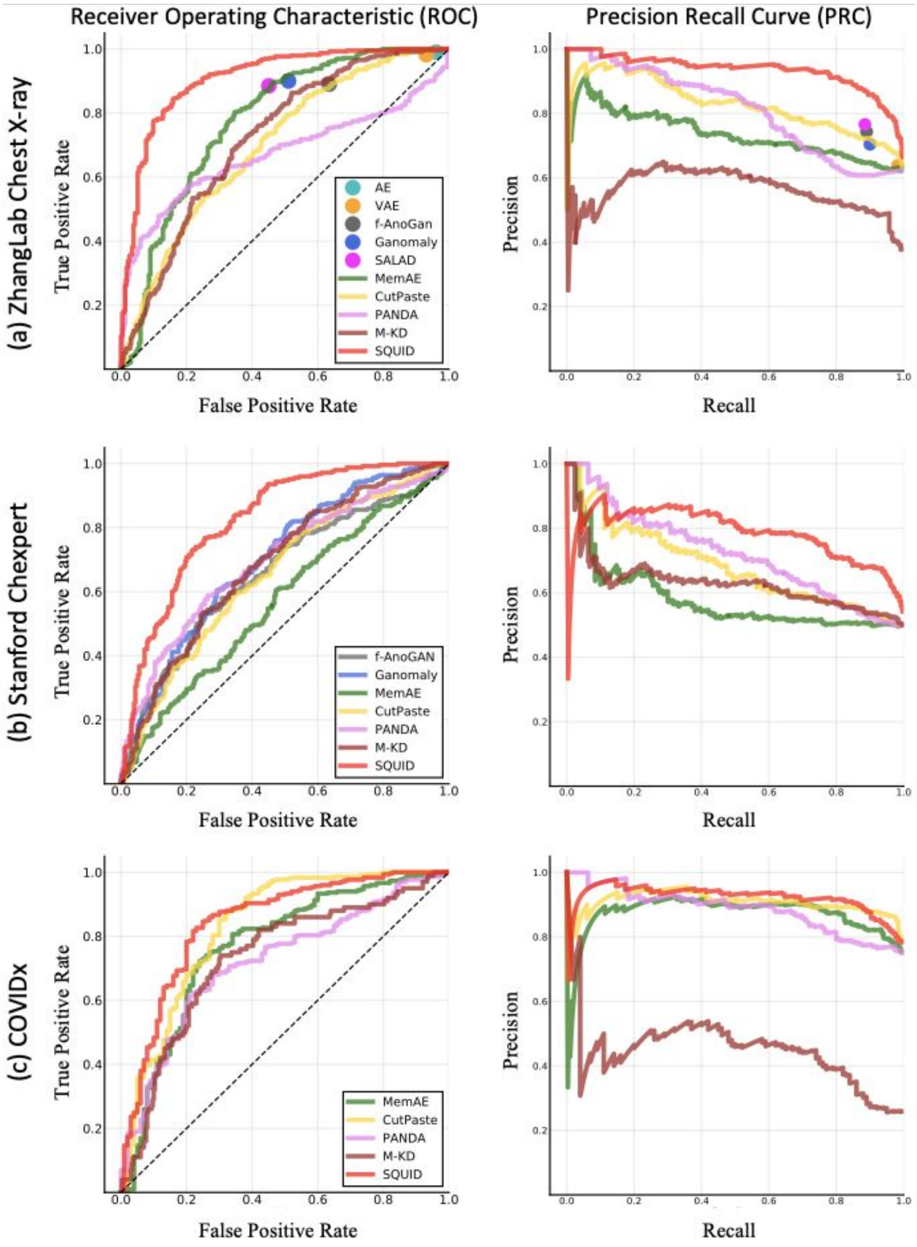
CheXpert	Ref & Year	AUC (%)	Acc (%)	F1 (%)
Ganomaly [1]	ACCV'18	68.9±1.4	65.7±0.2	65.1±1.9
f-AnoGAN [64]	MIA'19	65.8±3.3	63.7±1.8	59.4±3.8
MemAE [15]	ICCV'19	54.3±4.0	55.6±1.4	53.3±7.0
CutPaste [43]	CVPR'21	65.5±2.2	62.7±2.0	60.3±4.6
PANDA [58]	CVPR'21	68.6±0.9	66.4±2.8	65.3±1.5
M-KD [62]	CVPR'21	69.8±1.6	66.0±2.5	63.6±5.7
SQUID	This work	**78.1±5.1**	**71.9±3.8**	**75.9±5.7**

TABLE 4.3. Results on the test sets of the COVIDx dataset. Both average results and standard deviations are reported. † denotes the results are taken from [69]. ‡ denotes the results are taken from [78].

COVIDx	Ref & Year	AUC (%)	Acc (%)	F1 (%)
PaDiM† [10]	ICPR'21	54.0	-	-
Ganomaly† [1]	ACCV'18	58.4	-	-
f-AnoGAN‡ [64]	MIA'19	66.9	-	-
MemAE [15]	ICCV'19	71.8±3.6	77.1±2.1	86.4±0.8
PANDA [58]	CVPR'21	72.3±1.0	**76.9±0.8**	**86.4±0.4**
M-KD [62]	CVPR'21	71.7±1.1	69.7±4.5	55.6±2.5
SQUID	This work	**74.7±0.9**	76.8±0.1	86.0±0.2

Results: Public Benchmarks

Quantitative Eval.

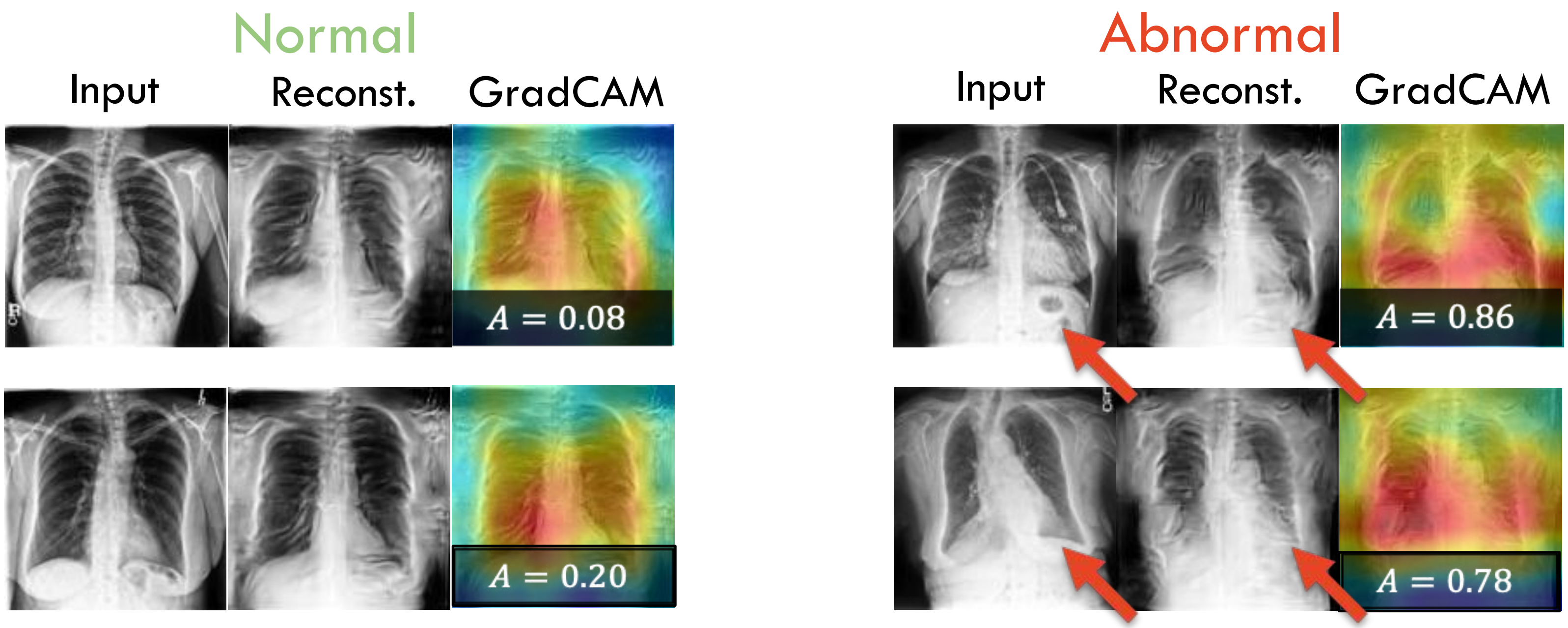


- ROC, PRC as metrics.

- SQUID (red plots) yields much better AUC on all datasets.

Results: Public Benchmarks

Qualitative Eval.



- Reconstructed normal images seem normal. ✓
- Reconstructed abnormal images seem normal. ✓
- Reconstructed normal/abnormal images have clear quality diff. ✓
- High anomaly score (*A*) for abnormal images, low for normal images. ✓

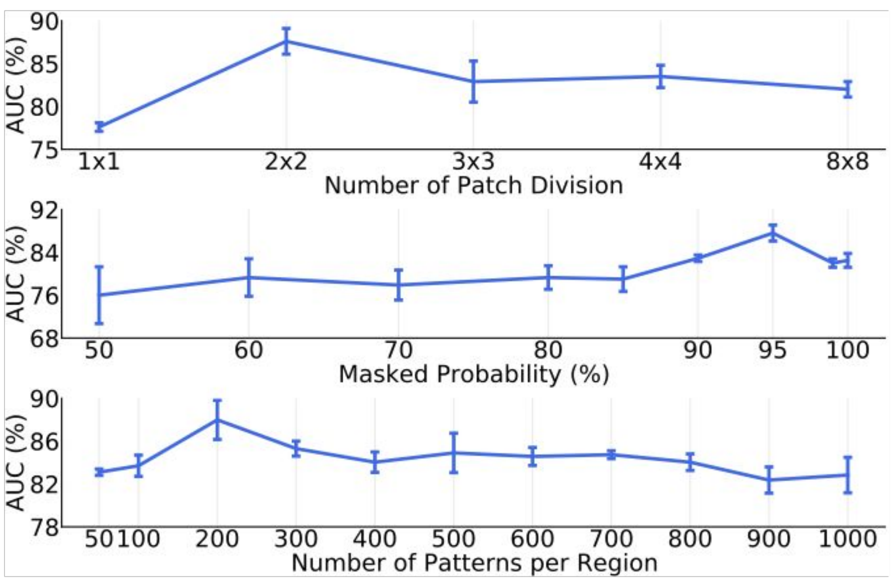
Results: Ablation Studies

Component Studies

TABLE 4.4. Component studies indicate that the overall performance benefits from all of the components in SQUID. The ablation study is conducted on the ZhangLab dataset.

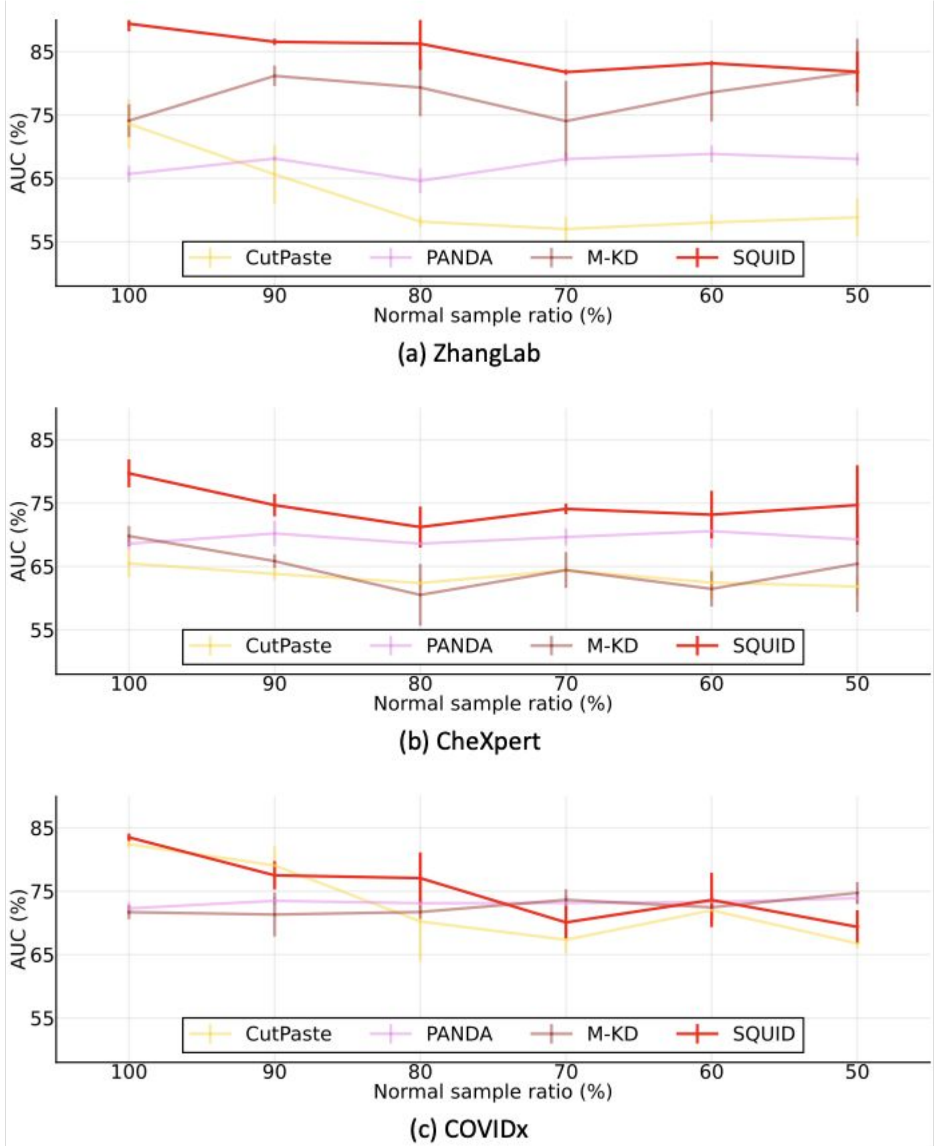
Method	AUC(%)	Acc(%)	F1(%)
w/o Space-aware Memory	77.6±0.5	75.5±0.5	82.5±0.6
w/o In-painting Block	80.9±2.1	75.8±1.5	81.6±1.3
w/o Skip Connection	79.5±1.6	73.0±1.4	78.8±0.5
w/o Hierarchical Memory	82.9±1.2	77.4±1.1	81.2±0.5
w/o Knowledge Distillation	85.4±0.8	79.5±0.7	83.5±0.8
w/o Stop Gradient	85.0±4.3	77.6±2.8	79.8±1.6
w/o Gumbel Shrinkage	86.2±3.3	80.5±3.2	85.4±2.1
Full SQUID	**87.6±1.5**	**80.3±1.3**	**84.7±0.8**

Hyper-param. Studies



Results: True UAD Training

- Training dataset contains unknown data (normal/abnormal mixture).
- UAD algorithms should be robust to the mixed training.
- SQUID (red plots) yields the best robustness when the normal sample ratio >=60%.



(a) ZhangLab

(b) CheXpert

(c) COVIDx

Outline

Discussion

Limitations & Future Work

- **Complex framework.**
  - Reuse networks/layers.
  - Better skip connections.
- **Inefficient inference.**
  - Lighter-weight backbone/operators.
  - Network pruning/quantization/compression.
- **Inaccurate pixel-wise anomaly detection.**
  - Feature-space residual.
  - In-painting + data augmentation.

Outline

1. Motivation
2. Background & Literature
3. Problem Definition
4. Methodology
5. Results
6. Discussion
7. **Conclusion**

Conclusion

Contributions

- Reformulated UAD as **feature-space in-painting**.
- Proposed **Space-aware Memory Queue** that caters to the unique characteristics of chest radiography.
- Designed multiple functional modules: **Gumbel Shrinkage, Masked Shortcut, Anomaly discrimination** that have never been explored in the UAD domain.
- Created the **DigitAnatomy** dataset to assist algorithm design in this domain.
- Achieved **SOTA performances** on three public benchmarks.
- Evaluated methods under the **real UAD training** settings for the first time.

Research outputs

- Part of this thesis has been submitted to **ECCV2022** and **Medical Image Analysis**.

References

- Dong Gong, Lingqiao Liu, Vuong Le, Budhaditya Saha, Moussa Reda Mansour, Svetha Venkatesh, and Anton van den Hengel. 2019. Memorizing normality to detect anomaly: Memory-augmented deep autoencoder for unsupervised anomaly detection. In *Proceedings of the IEEE/CVF International Conference on Computer Vision*, pages 1705–1714.
- Olaf Ronneberger, Philipp Fischer, and Thomas Brox. 2015. U-net: Convolutional networks for biomedical image segmentation. In *International Conference on Medical Image Computing and Computer-Assisted Intervention*, pages 234–241. Springer.
- Ian Goodfellow, Jean Pouget-Abadie, Mehdi Mirza, Bing Xu, David Warde-Farley, Sherjil Ozair, Aaron Courville, and Yoshua Bengio. 2014. Generative adversarial nets. *Advances in neural information processing systems*, 27.
- Mohammadreza Salehi, Niousha Sadjadi, Soroosh Baselizadeh, Mohammad Hossein Rohban, and Hamid R Rabiee. 2020. Multiresolution knowledge distillation for anomaly detection.
- Tiange Xiang, Chaoyi Zhang, Yang Song, Siqi Liu, Hongliang Yuan, and Weidong Cai. 2021. Partial graph reasoning for neural network regularization. *arXiv preprint arXiv:2106.01805*.
- Chun-Liang Li, Kihyuk Sohn, Jinsung Yoon, and Tomas Pfister. 2021. Cutpaste: Self-supervised learning for anomaly detection and localization. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 9664–9674.
- He Zhao, Yuexiang Li, Nanjun He, Kai Ma, Leyuan Fang, Huiqi Li, and Yefeng Zheng. 2021. Anomaly detection for medical images using self-supervised and translation-consistent features. *IEEE Transactions on Medical Imaging*.
- Samet Akcay, Amir Atapour-Abarghouei, and Toby P Breckon. 2018. Ganomaly: Semi-supervised anomaly detection via adversarial training. In *Asian conference on computer vision*, pages 622– 637. Springer.
- Thomas Schlegl, Philipp Seeböck, Sebastian M Waldstein, Georg Langs, and Ursula Schmidt- Erfurth. 2019. f-anogan: Fast unsupervised anomaly detection with generative adversarial net- works. *Medical Image Analysis*.
- Thomas Schlegl, Philipp Seeböck, Sebastian M Waldstein, Ursula Schmidt-Erfurth, and Georg Langs. 2017. Unsupervised anomaly detection with generative adversarial networks to guide marker discovery. In *International conference on information processing in medical imaging*, pages 146–157. Springer.
- Tiange Xiang, Yixiao Zhang, Yongyi Lu, Alan Yuille, Chaoyi Zhang, Weidong Cai, Zongwei Zhou, “In-painting Radiography Images for Unsupervised Anomaly Detection", Submitted to *The 17th European Conference on Computer Vision (ECCV 2022)*, 2022. (Under Review)
- Tiange Xiang, Yixiao Zhang, Yongyi Lu, Alan Yuille, Chaoyi Zhang, Weidong Cai, Zongwei Zhou, “Feature-level In-painting for Unsupervised Anomaly Detection in Radiography Images", Sub- mitted to *Medical Image Analysis*, 2022. (Under Review)

Thank you for your time and patience !

Any questions ?