Cross-Database Mammographic Image Analysis through Unsupervised Domain Adaptation

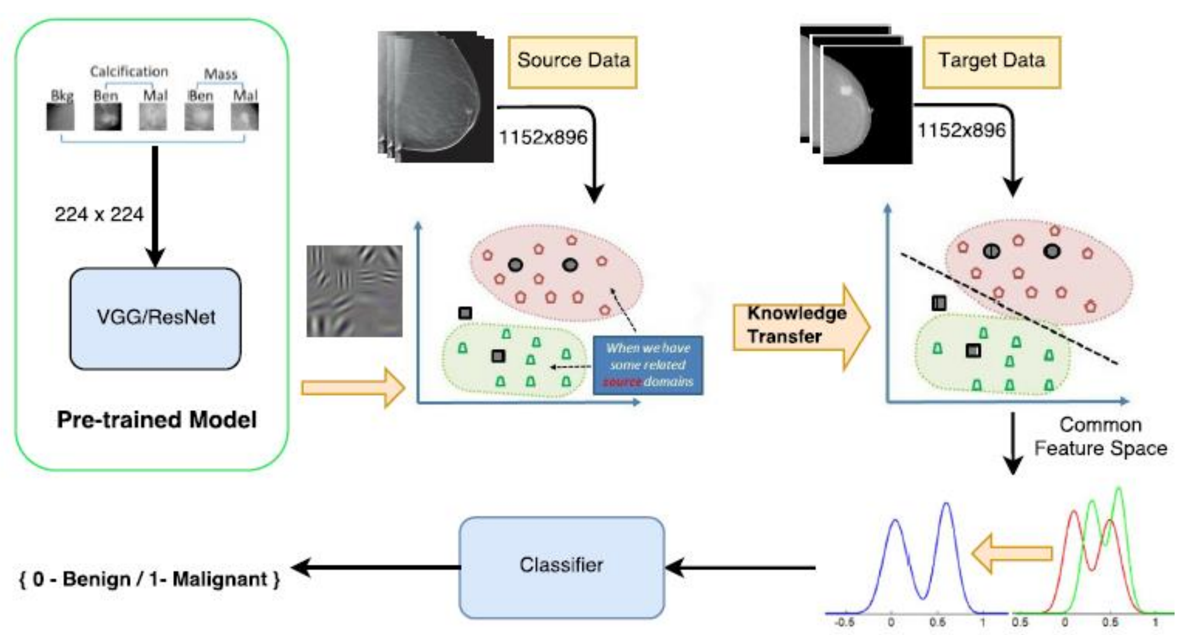
Deepak Kumar[1], Chetan Kumar[1], Ming Shao[2]

[1]Department of Data Science, University of Massachusetts , Dartmouth

[2]Department of Computer and Information Sciences, University of Massachusetts, Dartmouth

UMass | Dartmouth

Introduction



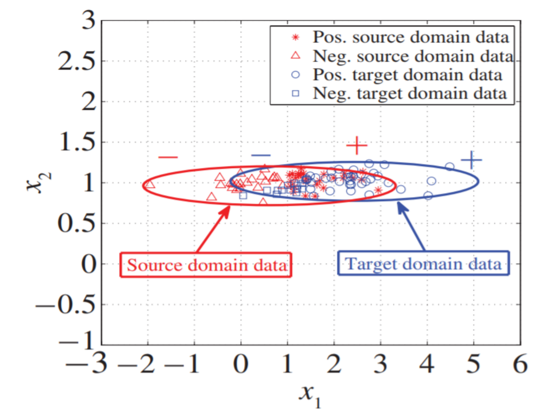
- Breast Cancer is common health issue
- 1.7 million cases were found and total 14.71% deaths occurred world wide
- Mammography is the best option available to diagnose breast cancer at early stage
- Radiologists expertise are required for interpolation of mammograms
- Computer Assisted Diagnosis (CAD) have caught the attention to automate the process of medical image classification
- CAD completely rely on the pre segmented portion of mammograms know as Region of Interest (ROI)
- These system rely on the assumption that training and future data has the same feature space and follows the same distribution
- It is expensive to acquire the sufficient new labeled data to rebuild the model
- Pre-trained End-to-End training models are used for feature extraction from publicly available databases
  - CBIS-DDSM
  - InBreast
  - MIAS
- Data has been transformed into same feature space using the following transfer learning methods:
  - TCA (Transfer Component Analysis)
  - BDA (Balance Distribution Adaption)
  - CoRAL (Correlation Alignment)

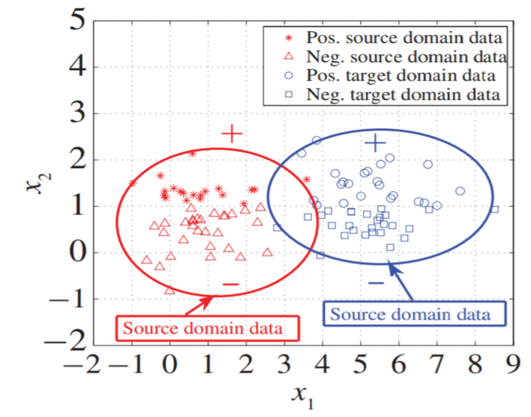
Transfer Learning

- Traditional machine learning methods learn models based on:
- Pre collected label data to predict new data
- New data should be in same distributed space
- It is difficult to get rich labeled data
- Transfer learning facilitates
- Different distribution for training and future data
- Transfer learning can be promising
  - › Characteristics of different sub domain are mapped into wide common domain

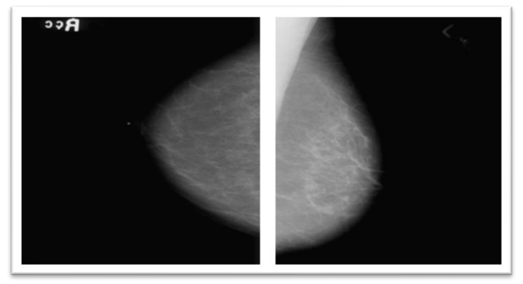
Methodology

- End-to-End training algorithm for whole-image breast cancer diagnosis is based on mammograms without the need of the ROI annotations.
  - In End-to-End three different convolutional network are used but first they trained on the patch classifier by changing the learning rate by freezing and unfreezing the layers.
  - They fixed the mammogram patch image size to 224 x 224 and whole image size 1152 x 896.
- Domain Adaption
  - TCA (Transfer Component Analysis)
    - › Learn the common Components
    - › Reduce the differences between source and target distribution by projecting on common subspace
  - BDA (Balance Distribution Adaption)
    - › Balance the marginal and conditional distribution adaption
      - For different data Marginal distribution is important
      - For similar data Conditional distribution is important
    - › Handles the data class imbalanced by adjusting each class weight
  - CoRAL (Correlation Alignment)
    - › CoRAL works on unsupervised domain adaption
    - › Map source domain distribution by recoloring the whitened features with second order statistics of target domain

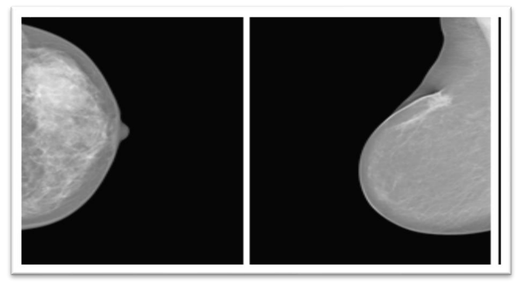




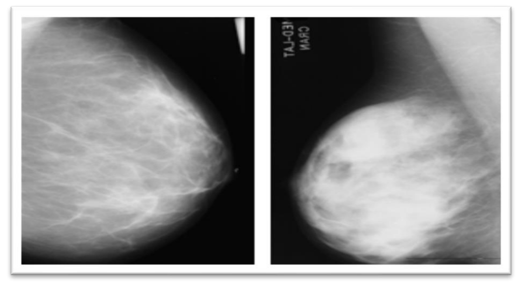
Datasets



CBIS- DDSM



InBreast

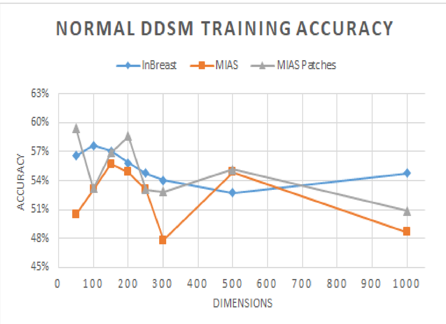


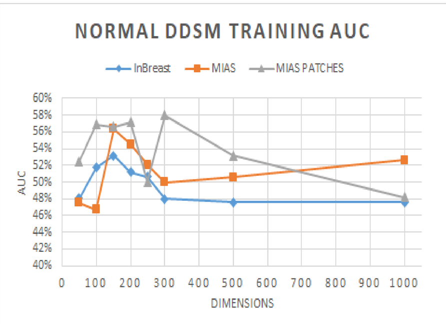
MIAS

- Three publicly available dataset are used
  - CBIS-DDSM
    - › 3103 scanned mammograms
    - › Data contain Benign and Malignant images
    - › Dataset contains Cranial Cardo (CC) and Media later oblique (MLO) views
  - InBreast
    - › Full field digital images contain different color profiles
    - › The dataset contains 410 mammograms from 115 patients including Cranial Cardo (CC) and Media Later Oblique (MLO) views
    - › Data contains BI-RAIDS readings
    - › Sample was divided in Benign and Malignant based on BI-RAIDS reading

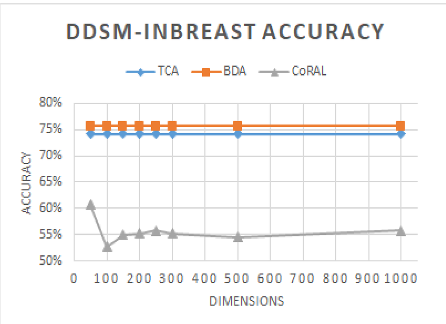
- MIAS
  - Dataset contains 322 MLO view mammograms
  - MIAS images are scanned copies of films
  - Only 109 images are labeled

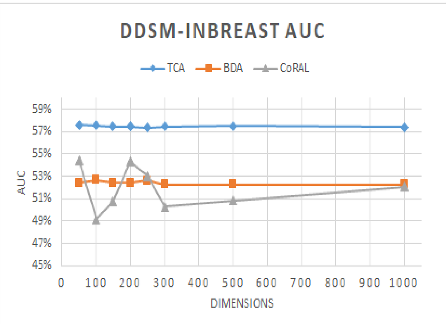
Experimental Results



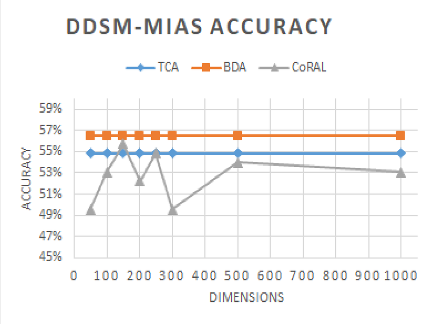


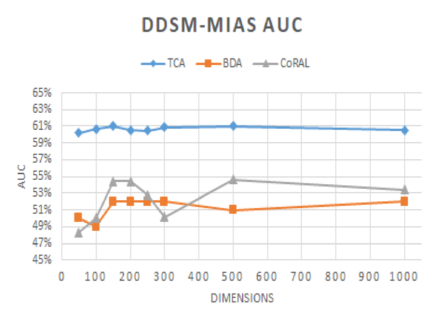
Dimension	InBreast		MIAS		MIAS PATCHES	
	Accuracy	AUC	Accuracy	AUC	Accuracy	AUC
50	56.58%	48.08%	50.44%	47.5%	59.48%	52.49%
100	57.62%	51.7%	53.09%	46.68%	53.15%	56.89%
150	57.1%	53.16%	55.75%	56.33%	56.89%	56.62%
200	55.81%	51.19%	54.86%	54.49%	58.62%	57.19%
250	54.78%	50.67%	53.09%	52.09%	53.09%	50%
300	54%	47.98%	47.78%	49.97%	52.58%	57.98%
500	52.71%	47.6%	54.86%	50.6%	55.17%	53.14%
100	54.78%	47.6%	48.67%	52.62%	50.86%	48.17%



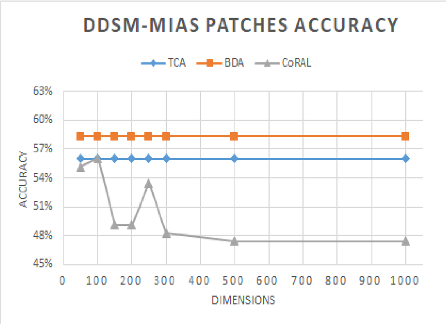


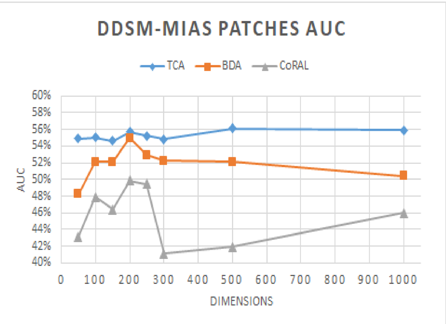
Dimension	TCA		BDA		CoRAL	
	Accuracy	AUC	Accuracy	AUC	Accuracy	AUC
50	74.16%	57.6%	75.64%	52.43%	60.72%	54.48%
100	74.16%	57.55%	75.64%	52.68%	52.71%	49.18%
150	74.16%	57.45%	75.64%	52.4%	55.03%	50.78%
200	74.16%	57.48%	75.64%	52.4%	55.29%	54.31%
250	74.16%	57.36%	75.64%	52.6%	55.81%	53.01%
300	74.16%	57.46%	75.64%	52.28%	55.29%	50.29%
500	74.16%	57.52%	75.64%	52.28%	54.52%	50.85%
1000	74.16%	57.43%	75.64%	52.28%	55.81%	52.08%





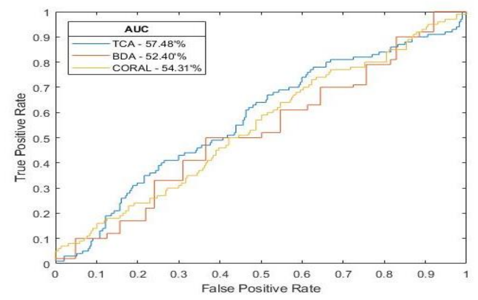
Dimension	TCA		BDA		CoRAL	
	Accuracy	AUC	Accuracy	AUC	Accuracy	AUC
50	56.03%	54.93%	58.33%	48.24%	55.17%	43.14%
100	56.03%	55.05%	58.33%	52.1%	56.03%	47.9%
150	56.03%	54.6%	58.33%	52.1%	49.13%	46.4%
200	56.03%	55.72%	58.33%	54.96%	49.13%	49.89%
250	56.03%	55.26%	58.33%	52.94%	53.44%	49.47%
300	56.03%	54.81%	58.33%	52.25%	48.27%	41.09%
500	56.03%	56.14%	58.33%	52.1%	47.41%	41.9%
1000	56.03%	55.93%	58.33%	50.44%	47.41%	45.97%



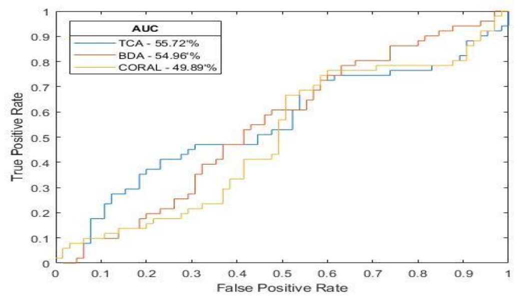


Dimension	TCA		BDA		CoRAL	
	Accuracy	AUC	Accuracy	AUC	Accuracy	AUC
50	54.86%	60.22%	56.52%	50.01%	49.55%	48.36%
100	54.86%	60.66%	56.52%	48.99%	53.1%	50.06%
150	54.86%	61.01%	56.52%	51.99%	55.75%	54.46%
200	54.86%	60.59%	56.52%	51.99%	52.21%	54.43%
250	54.86%	60.47%	56.52%	51.99%	54.86%	52.78%
300	54.86%	60.88%	56.52%	51.99%	49.55%	50.13%
500	54.86%	61.04%	56.52%	51.01%	53.98%	54.65%
1000	54.86%	60.53%	56.52%	51.99%	53.09%	53.45%

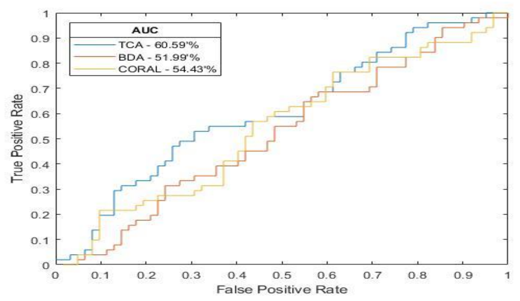
DDSM- INBREAST



DDSM- MIAS



DDSM- MIAS PATCHES



References

1. L. Shen, “End-to-end training for whole image breast cancer diagnosis using an all convolutional design,” arXiv preprint arXiv:1708.09427, 2017
2. S. J. Pan, I. W. Tsang, J. T. Kwok, and Q. Yang, “Domain adaptation via transfer component analysis,” IEEE Transactions on Neural Networks, vol. 22, no. 2, pp. 199–210, 2011.
3. B. Sun, J. Feng, and K. Saenko, “Return of frustratingly easy domain adaptation.” in AAAI, vol. 6, no. 7, 2016, p. 8.
4. J. Wang, Y. Chen, S. Hao, W. Feng, and Z. Shen, “Balance distribution adaption for transfer learning,” 2017.