Medical Image Retrieval: A Multimodal Approach

B S Suhas - 15IT110
A Aditya - 15IT201
M M Vikram - 15IT217

Agenda

Introduction

Goal of medical image retrieval - to find the most clinically relevant images in response to specific information needs represented as search queries.

Paper Overview

- Probabilistic Latent Semantic Analysis (PLSA) model to integrate the visual and textual information from medical images to bridge the semantic gap.

- Deep Boltzmann machine-based multimodal learning model to learn the joint density model from multimodal information in order to derive the missing modality.

Literature Review

Authors	Methodology	Advantages	Limitations
Greenspan and Pinhas	continuous and probabilistic image representation scheme using Gaussian mixture modeling (GMM) along with information-theoretic image matching via the Kullback-Leibler (KL) measure.	1. A classification rate of 97.5% was achieved. 2. Precision versus recall curves indicate a strong retrieval result as compared with other state-of-the-art retrieval techniques	1. GMM is a very crude (and lossy) image representation 2. the number of Gaussians k is defined and kept constant per image, which affects the classification accuracy
Rahman et al	unified medical image retrieval framework integrating visual and textual keywords using a novel multimodal query expansion	1. a different QE technique in a unified multi-modal framework explored, which is based on the correlation analysis of both visual and text keywords and relies only on the local feedback information.	
Quellec et al	content-based heterogeneous information retrieval framework and proposed a Bayesian network to recover missing infor-mation.	1. retrieval model is trained separately for each attribute, which is useful for incomplete documents	1. only one type of image feature [25] has been included in the retrieval 2. the size of the dataset has an influence on the correctness of the generated Bayesian networks.

Outcome of Literature Review

- The different methods and models available for CBIR understood
- Multimodal approaches are better
- Many-a-times there are a lot of missing modalities, which needs to be handled effectively, for a good accuracy

Research Gap Analysis

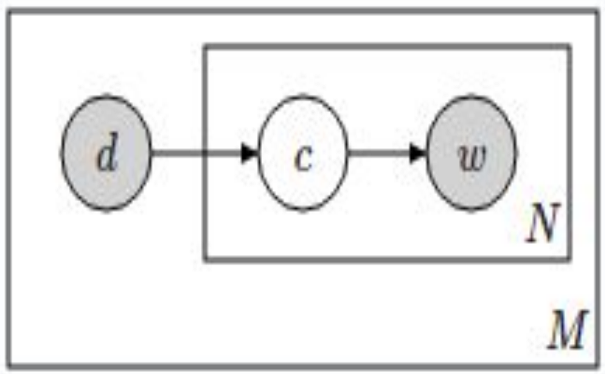
Drawbacks of PLSA (Probabilistic Latent Semantic Analysis)

1. It is generative model of the documents in the collection it is estimated on, it is not a generative model of new documents. Hence it is prone to overfitting.

2. Also the number of parameter in the model grows linearly with the size of the corpus.

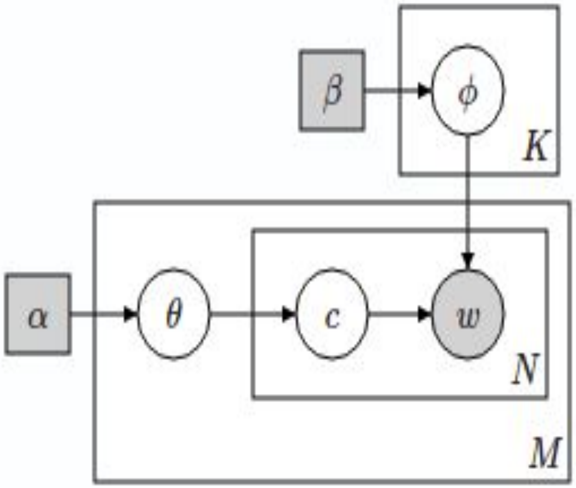
- Since it is crucial to have a model flexible enough to properly handle text that has not been seen before, **Latent Dirichlet allocation** is used – intuition : documents cover only a small set of topics and that topics use only a small set of words frequently.

PLSA



$$P(w \mid d) = P(d) \sum_{c} P(c \mid d)\, P(w \mid c)$$

LDA



Issues and Challenges

- Semantic gap exists between the low-level features (eg, low-level visual and textual features) and the high-level medical concepts.
- The real-world data are very noisy and some modality information (eg, text annotation) may be missing from input.

Motivation

- As medical imaging is becoming an essential component for cancer care and research, many departments of cancer care and research would benefit directly from research efforts on multimodal image retrieval.

- Tremendous amounts of medical image data, in the last few years, are captured and recorded in a digital format during the daily clinical practice, medical research, and education.

- Text-based information retrieval methods being both mature and well researched, they are limited by the quality of image annotations and are affected by noise.

Problem Statement

Medical Image retrieval using multimodal (visual and textual) features.

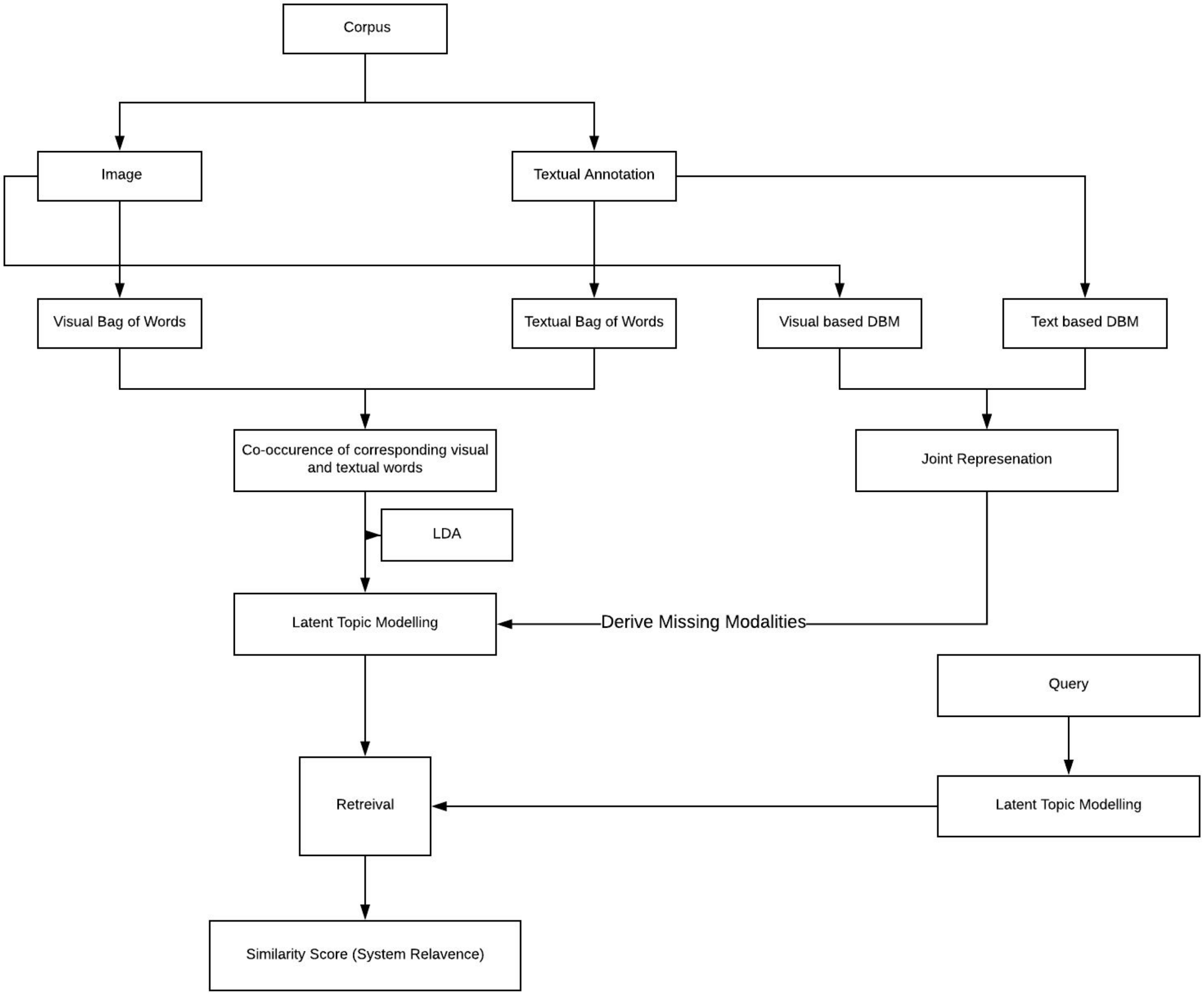
Research Objectives

- Integrate the visual and textual information from medical images to bridge the semantic gap between features.

- Develop a new deep Boltzmann machine (DBM)-based multimodal learning model to learn the joint density model from multimodal information.

- Integrate previous two steps and develop a retrieval model for medical images.

Proposed Model



Methodology

- Fusing the multimodal information:

1. The goal in this step is to generate a latent topic representation for each image, with its textual annotation.

2. Visual Bag of Words used (VBoW) to represent images and textual bag of words(BoW) model for textual annotations.

3. We use LDA (latent dirichlet allocation) model to encode visual and textual features.

Methodology

- Deriving Missing Modalities:

1. In real-world clinical applications, sometimes modalities are missing or noisy.
2. DBM (Deep Boltzmann machine) used learn a joint probability density model from the visual and textual information with the capacity of filling in missing modalities.

Methodology

- Retrieval:

1. The visual and textual features are extracted from query image.
2. A similarity score is calculated between the database images and query.
3. The most similar images are retrieved .

Work Done

- Literature survey
- Implementation
  - Generated sift keypoints and and their corresponding sift descriptors for all images.
  - Clustered the sift descriptors into 3000 clusters - the respective 3000 centroids are the visual words. Thus, each document now consists of visual words from the vocabulary of 3000 visual words.

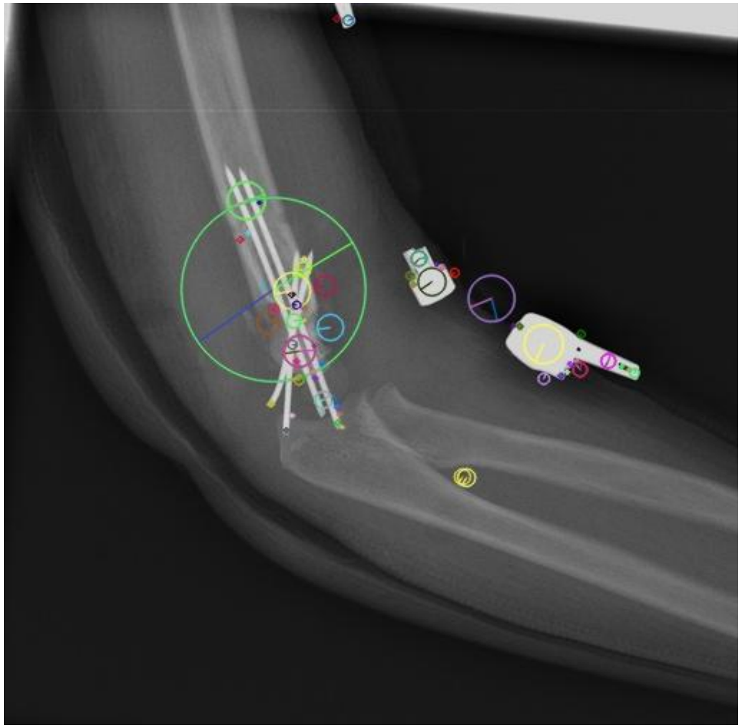
Work Done

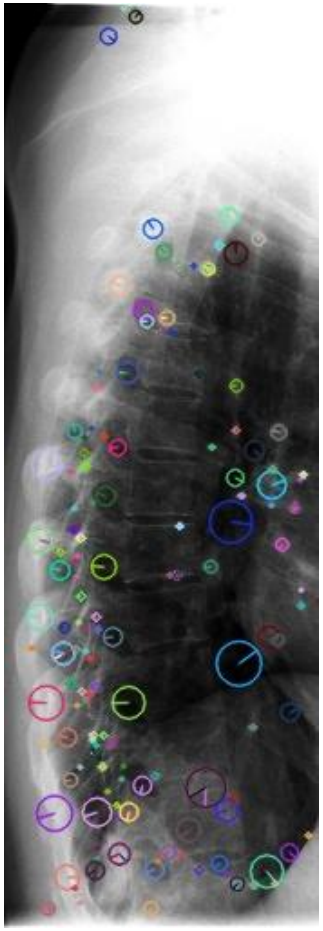
- Used LDA to obtain 100 latent topics in the documents. Now, each document is modeled as a combination of these 100 latent topics.

- Can query an image, the sift features of this image will be generated and clustered as before. The combination of latent topics for this document is used to retrieve 4 most relevant documents from the corpus(used euclidean distance).

Results and Analysis

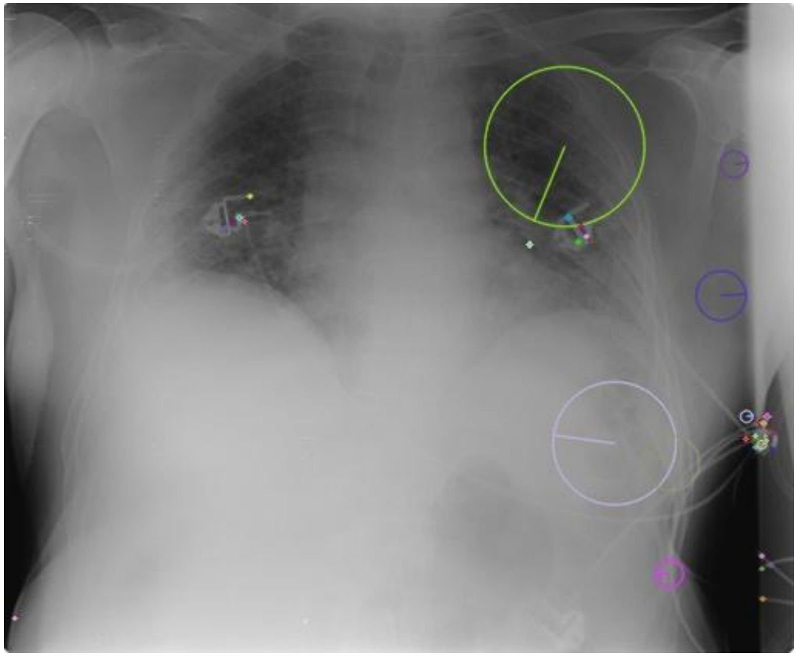
SIFT KEY POINTS

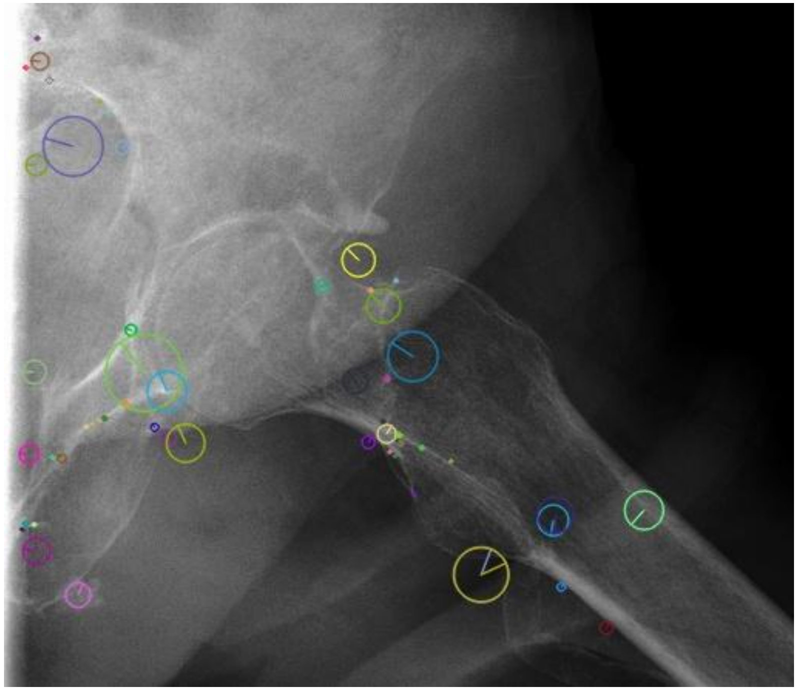




Results and Analysis

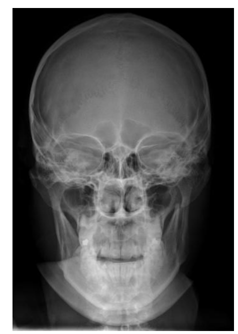
SIFT KEY POINTS





Results and Analysis

Query



Retrieved documents in order of relevance

Time Line of Project

26 Feb
Identify research gaps, methodology and feasibility study

05 Mar
Step 1 implementation

12 Mar
Step 2 implementation

19 Mar
Step 3 implementation

26 Mar
Improve, evaluate and refine Model

Individual Contribution

Suhas BS : LDA Model, Similarity score for query retrieval, Joint representation DBM layer, 2 evaluation metrics

A Aditya : Textual Bag of Words (BoW) Model, k-means clustering, Text based DBM, 2 evaluation metrics

M M Vikram : SIFT-based point detection, k-means clustering : Visual Bag of Words, Visual based DBM, 2 evaluation metrics

References

Selected Base Paper

Medical Image Retrieval: A Multimodal Approach

(Yu Cao, Shawn Steffey, Jianbiao He , Degui Xiao, Cui Tao , Ping Chen and Henning Müller)

(Other Related References)

1. Quellec G, Lamard M, Cazuguel G, Roux C, Cochener B. Case retrieval in medical databases by fusing heterogeneous information. IEEE Trans Med Imaging. 2011;30(1):108–18.
2. Srivastava N, Salakhutdinov R. Multimodal learning with deep Boltzmannmachines. Paper presented at: Advances in Neural Information Processing Systems 25; 2012.
3. David M. Blei, Andrew Y. Ng, Michael I. Jordan; Latent Dirichlet Allocation;3(Jan):993-1022, 2003.