Lecture 3 : Research Paper – Part 2 (Full Paper)

Instructor: Greta Mok
Dept of Electrical and Computer Engineering
University of Macau
Sept 13, 2021

Manuscript types

- Original articles
- Short communications/Letters
- Review
- Traditional subscribe-based or open access?
- Requirement, e.g., word counts, number of figs./tables and general standard according to the journals

Introduction

- One of the most difficult parts to write (Discussion is the next)
- Background (depending on the readers of the journals)
- Literature review
- State-of-the-art
- Current problems
- Reason and goal of your work

Introduction

- Read related papers before starting (true for all parts)
- Define the theme for each paragraph (true for all parts)
- Introduce your former related works (if any) to show yr credentials
- Most refs located
  - Cite relevant and updated refs
    - Make friends with the potential reviewers
  - Cite precisely

Example 1

Tremendous progress has been made in image recognition, primarily due to the availability of large-scale annotated datasets (i.e., ImageNet [1], [2]) and the recent revival of deep convolutional neural networks (CNN) [3], [4]. For data-driven learning, large-scale well-annotated datasets with representative data distribution characteristics are crucial to learning more accurate or generalizable models [5], [4]. Unlike previous image datasets used in computer vision, ImageNet ~~[1]~~ offers a very comprehensive database of more than 1.2 million categorized natural images of 1000+ classes. The CNN models trained upon this database serve as the backbone for significantly improving many object detection and image segmentation problems using other datasets [6], [7], e.g., PASCAL [8] and medical image categorization [9]–[10][11][12]. However, there exists no large-scale annotated medical image dataset comparable to ImageNet, as data acquisition is difficult, and quality annotation is costly.

There are currently three major techniques that successfully employ CNNs to medical image classification: 1) training the "CNN from scratch" [13]–[14][15][16][17]; 2) using "off-the-shelf CNN" features (without retraining the CNN) as complementary information channels to existing hand-crafted image features, for chest X-rays [10] and CT lung nodule identification [9], [12]; and 3) performing unsupervised pre-training on natural or medical images and fine-tuning on medical target images using CNN or other types of deep learning models [18]–[19][20][21]. A decompositional 2.5D view resampling and an aggregation of random view classification scores are used to eliminate the "curse-of-dimensionality" issue in [22], in order to acquire a sufficient number of training image samples.

Example 1

Previous studies have analyzed three-dimensional patch creation for LN detection [23], [24], atlas creation from chest CT [25] and the extraction of multi-level image features [26], [27]. At present, there are several extensions or variations of the decompositional view representation introduced in [22], [28], such as: using a novel vessel-aligned multi-planar image representation for pulmonary embolism detection [29], fusing unregistered multiview for mammogram analysis [16] and classifying pulmonary peri-fissural nodules via an ensemble of 2D views [12].

Although natural images and medical images differ significantly, conventional image descriptors developed for object recognition in natural images, such as the scale-invariant feature transform (SIFT) [30] and the histogram of oriented gradients (HOG) [31], have been widely used for object detection and segmentation in medical image analysis. Recently, ImageNet pre-trained CNNs have been used for chest pathology identification and detection in X-ray and CT modalities [10], [9], [12]. They have yielded the best performance results by integrating low-level image features (e.g., GIST [32], bag of visual words (BoVW) and bag-of-frequency [12]). However, the fine-tuning of an ImageNet pre-trained CNN model on medical image datasets has not yet been exploited.

Example 1

In this paper, we exploit three important, but previously under-studied factors of employing deep convolutional neural networks to computer-aided detection problems. Particularly, we explore and evaluate different CNN architectures varying in width (ranging from 5 thousand to 160 million parameters) and depth (various numbers of layers), describe the effects of varying dataset scale and spatial image context on performance, and discuss when and why transfer learning from pre-trained ImageNet CNN models can be valuable. We further verify our hypothesis by inheriting and adapting rich hierarchical image features [5], [33] from the large-scale ImageNet dataset for computer aided diagnosis (CAD). We also explore CNN architectures of the most studied seven-layered "AlexNet-CNN" [4], a shallower "Cifar-CNN" [22], and a much deeper version of "GoogLeNet-CNN" [33] (with our modifications on CNN structures). This study is partially motivated by recent studies [34], [35] in computer vision. The thorough quantitative analysis and evaluation on deep CNN [34] or sparsity image coding methods [35] elucidate the emerging techniques of the time and provide useful suggestions for their future stages of development, respectively.

Two specific computer-aided detection (CADe) problems, namely thoraco-abdominal lymph node (LN) detection and interstitial lung disease (ILD) classification are studied in this work. On mediastinal LN detection, we surpass all currently reported results. We obtain 86% sensitivity on 3 false positives (FP) per patient, versus the prior state-of-art sensitivities of 78% [36] (stacked shallow learning) and 70% [22] (CNN), as prior state-of-the-art. For the first time, ILD classification results under the patient-level five-fold cross-validation protocol (CV5) are investigated and reported. The ILD dataset [37] contains 905 annotated image slices with 120 patients and 6 ILD labels. Such sparsely annotated datasets are generally difficult for CNN learning, due to the paucity of labeled instances.

Example 1

Evaluation protocols and details are critical to deriving significant empirical findings [34]. Our experimental results suggest that different CNN architectures and dataset re-sampling protocols are critical for the LN detection tasks where the amount of labeled training data is sufficient and spatial contexts are local. Since LN images are more flexible than ILD images with respect to resampling and reformatting, LN datasets may be more readily augmented by such image transformations. As a result, LN datasets contain more training and testing data instances (due to data augmentation) than ILD datasets. They nonetheless remain less comprehensive than natural image datasets, such as ImageNet. Fine-tuning ImageNet-trained models for ILD classification is clearly advantageous and yields early promising results, when the amount of labeled training data is highly insufficient and multi-class categorization is used, as opposed to the LN dataset's binary class categorization. Another significant finding is that CNNs trained from scratch or fine-tuned from ImageNet models consistently outperform CNNs that merely use off-the-shelf CNN features, in both the LN and ILD classification problems. We further analyze, via CNN activation visualizations, when and why transfer learning from non-medical to medical images in CADe problems can be valuable.

Example 2

Introduction

Conventional SPECT using a parallel-hole collimator lacks sufficient detection efficiency and resolution trade-off for imaging small animals such as rodents, especially their even smaller organs-of-interest. High-resolution SPECT using a pinhole collimator, especially the multi-pinhole (MPH) collimator, with an appropriate pinhole pattern arrangement, has been recognized to provide improved image quality as compared to that by conventional parallel-hole and single pinhole (SPH) collimators, anchoring its unique role for the molecular imaging of small animals (Beekman and Van der Have 2007, Jaszczak *et al* 1994, Meikle *et al* 2005, Nuyts *et al* 2009, Freed *et al* 2008, Metzler *et al* 2005, Strand *et al* 1994, DiFilippo 2008). Multiple research and commercial systems with different design parameters are now available (Van der Have *et al* 2009, Schramm *et al* 2003, Kim *et al* 2006, Lackas *et al* 2005, DiFilippo 2008). Although most of the design parameters, such as the number and placement of pinholes, collimator length, aperture size and degree of multiplexing, i.e. projection overlapping from neighboring apertures, have been studied, there is still no consensus about the optimized MPH design as it always depends on the detector size, intrinsic resolution and, ultimately, the specific application with specific object activity distribution.

Among all MPH SPECT system design parameters, multiplexing is the one that draws a lot of attention. Justifications that are being used for multiplexing in the projections, as once studied by the authors (Mok *et al* 2009), are to further increase detection efficiency by having a larger number of pinholes or allowing a shorter radius-of-rotation (ROR), or to improve spatial resolution by having a larger magnification factor (figure 1). However, studies have demonstrated that multiplexing poses ambiguous information for reconstructions which may lead to non-unique solutions for the inverse problem, because counts in the overlapping region cannot be traced back through a particular pinhole aperture (Vunckx *et al* 2008, Mok *et al* 2009) and significant longer image reconstruction times are needed to suppress multiplexing artifacts. These substantial adverse effects are actually intertwined with the pinhole pattern and are more prominent for certain non-optimized pinhole patterns (Mok *et al* 2005).

Example 2

Vunckx *et al* (2008) studied the effect of overlapping projections based on simulations with homogeneous sphere and rat brain phantoms. They conclude that once the detector area has been fully utilized, the extra detection efficiency can only compensate for the ambiguity posed by multiplexing, and removing overlap actually improves the contrast-to-noise ratio especially for the central field-of-view. Results from Mahmood *et al* (2010) also showed that multiplexing leads to image artifacts as compared to non-multiplexing designs; however, they found that spatial resolution can be maintained and noise can be reduced for high degrees of multiplexing given that non-multiplexed data were also incorporated into the reconstructions. The authors also showed that improved detection efficiency and resolution by increased multiplexing are offset by increased image NMSE (Mok *et al* 2009), and the common belief is that the tolerance for multiplexing depends on the specific imaging task.

The extreme case for multiplexing is coded aperture collimation with lots of pinholes ($\sim$100). It was originally proposed and works well for far-field applications, su[illegible] high-energy astronomical imaging with bright and isolated sources (Caroli *et al* 1987). A[illegible]si *et al* (2001) validated the relationship between SNR improvement and sensitivity gain with a coded aperture for point sources, and demonstrated its usage in imaging high contrast point-like objects with ^{125}I tracer when coupled to a hybrid pixel detector (Accorsi *et al* 2008). However, the majority of tracers for nuclear medicine imaging do not bind to the targets with high specificity, and as a result background activity always exists. The common diagnostic tasks in SPECT include detecting hot lesions, cold defects and quantifying tracer concentrations, and multiplexing may have a different impact on each of these application areas. This leads to the

Example 2

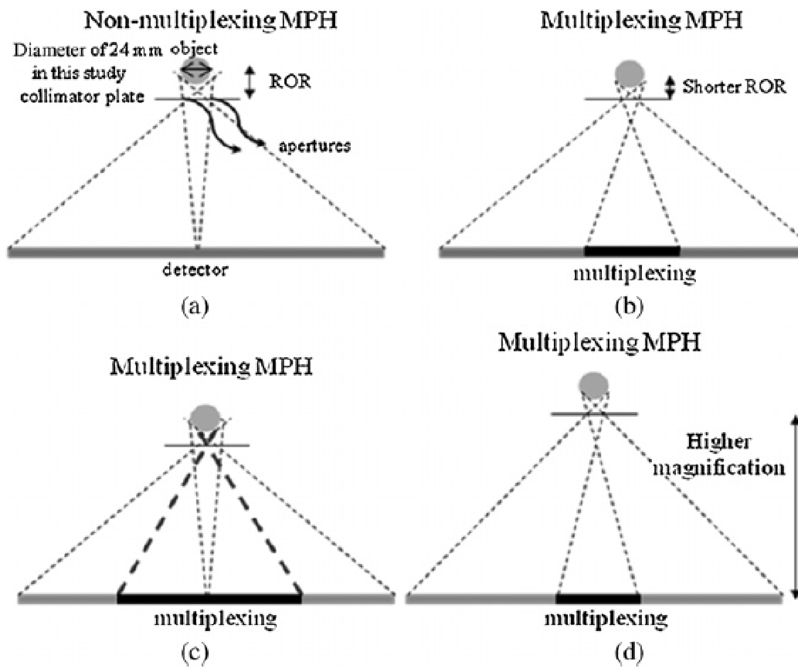

Figure 1. When compared to a non-multiplexing MPH design (a), which shows a simplified 2-pinhole configuration here, multiplexing can improve detection efficiency by allowing a shorter ROR (b), by packing a larger number of pinholes (3 pinholes here) (c) or by increasing the magnification for improved resolution in the projection domain as each projection contains more resolvable detector pixel elements (d).

英

question 'Is the effect of multiplexing dependent on the specific tracer distributions?' which was briefly discussed by Vunckx *et al* for two different phantoms (Vunckx *et al* 2008). This study aims to address this concern with simulations using hot and cold rod digitized phantoms and a digitized cold sphere phantom.

Methods and Materials

- Subsections to describe different parts of the methods
  - Theory
  - Data/subjects
  - Data generation
    - Simulations
    - Experiments
  - Systems
  - Evaluation
    - Data analysis
    - Statistical analysis
- Usually need to compare your work w/ existing methods

Methods and Results

- Figures
  - Label axes
  - Caption under figures
  - All lines and bars can be discriminated by different colors or patterns
  - Color printing fee for physical copies thus avoid color figs if possible
- Tables
  - Title on top of the tables
- Tables and figures not redundant with each other and can be self-explained without referring to the main content

Results

- Following the order of the methods
- Prepare corresponding figures and tables
- Figures and tables all mentioned and described in the text
- Mainly reporting the results, reasons for the results can be placed in the Discussions

Discussions

- Discuss the findings – why?
  - Showing the depth of your works and you have thought thoroughly for every relevant possible perspectives
- Set paragraphs to discuss different topics
- Results as compare to other existing related works
- Limitations & future works

Conclusions

- Summarizing the study – what is done?
- Key findings and take home messages
- No over- or under- statement

References

- Cite the updated and relevant papers
- Format according to the journals.
- Software
  - Endnote (free from library, https://library2.um.edu.mo/html/e_resources/EndNote.html)

Acknowledgement

- Funding info
- People/facilities helping your manuscript but not on the author list

Cover letter (optional)

- Letter to the editor
  - Briefly explaining the novelty of the work
  - Usually prepared by the corresponding author

New findings/key points (optional)

- Extract information further from the abstract

Abbreviation list (optional)

- Prepare a list of all abbreviations

Supplementary data (optional)

- Technical description/data relevant but not critical for the manuscript
- May be helpful to place all content when considering the word/figs/tables limit for the journal

Homework 2

- Write a 2-page paper without abstract and index terms