2019-01-01

Image Analysis For Lung Cancer Diagnosis And Superresolution Localization Microscopy

Xia Huang
University of Texas at El Paso

IMAGE ANALYSIS FOR LUNG CANCER DIAGNOSIS AND SUPERRESOLUTION LOCALIZATION MICROSCOPY

XIA HUANG

Doctoral Program in Biomedical Engineering

APPROVED:

Wei Qian, Ph.D., Chair

Tzu-Liang (Bill) Tseng, Ph.D., Co-chair

Thomas Boland, Ph.D.

Chuan (River) Xiao, Ph.D.

Thompson Sarkodie-Gyan, Ph.D.

Stephen L. Crites, Jr., Ph.D.
Dean of the Graduate School

IMAGE ANALYSIS FOR LUNG CANCER DIAGNOSIS AND SUPERRESOLUTION LOCALIZATION MICROSCOPY

by

XIA HUANG, M.S.

DISSERTATION

Presented to the Faculty of the Graduate School of

The University of Texas at El Paso

in Partial Fulfillment

of the Requirements

for the Degree of

DOCTOR OF PHILOSOPHY

Department of Metallurgical, Materials, and Biomedical Engineering

THE UNIVERSITY OF TEXAS AT EL PASO

August 2019

Acknowledgements

The path toward success is always circuitous, riffed with not only troubles and challenges, but also supports and fortune. I consider myself fortunate indeed to have the opportunity to pursue the Ph.D. degree, one of the milestones in my entire life. Its completion is thanks in large part to all the professors, family members, and friends who encouraged me when I was frustrated and congratulated me when I was achieved some success. Without you, I cannot be myself.

To my committee members, especially for my advisors Dr. Wei Qian and Chunqiang Li, thanks for all your supports, resources, and opportunities. Your instruction and guidance will be secured in my memory forever.

To my parents, thanks for your unending sacrifices, the video chats are my eternal sources of motivations.

To my wife, Liting, I am tremendously fortunate to have you sleep on my side. Your spirits inside the emaciated body is the most beautiful sunshine I've ever seen.

To my two little ones, Emily and Emile, thanks for stepping in my life. Your smiles are the sharpest sword to conquer any predicaments.

To my friend, Dani, I am toughed beyond words. Your enthusiasm inspired me all the time. I hope you have the brightest future. Keep in touch.

Abstract

With the development of imaging analysis techniques, flurry of applications associated with it has been hatched. In this paper, image analysis on both organ and cellular levels will be demonstrated. For organ level images, a deep learning based computer aided lung cancer diagnosis based on computer tomography images is studied. Deep learning techniques have been extensively used in computerized pulmonary nodule analysis recent years. Many reported studies still utilized hybrid methods for diagnosis, in which convolutional neural networks (CNNs) are used only as one part of the pipeline, and the whole system still needs either traditional image processing modules or human intervention to obtain the final results. In this paper, we introduced a fast and fully-automated end-to-end system that can efficiently segment the precise lung nodule contours from the raw thoracic CT scans. Our proposed system has three major modules: candidate nodule detection with Faster regional-CNN (R-CNN), candidate merging and false positive (FP) reduction with CNN, nodule segmentation with customized fully convolutional neural network (FCN). The entire system has no human interaction or database specific design. The average runtime is about 16 seconds per scan on a standard workstation. The nodule detection accuracy is 91.4% and 94.6% with an average of 1 and 4 false positives (FPs) per scan. The average dice coefficient of nodule segmentation compared to the groundtruth is 0.793.

For cellular level images, we studied localization algorithms on supperresolution localization microscopy to improve resolutions. Localization algorithms play a significant role in determining the accuracy in super resolution fluorescence imaging. A primary challenge is that choosing the right algorithm depends on users' prior knowledge about their specific imaging system. We introduce a Deep Matching method that combines convolutional neural networks to process raw images together with several conventional localization algorithms to calculate fluorophore positions. This method not only improves the localization accuracy, but also removes the dependence of accuracy on the algorithm chosen by the user. Our results also indicate the possibility to overcome the practical limit of the Cramér-Rao lower bound in the low signal-to-

noise ratio regime with Deep Matching processed images. Furthermore, inspired by the nature of the point spread function (PSF) in defocused images that have ring structures, they can be used to localize the 3D position of single particles by calculating the ring center (x & y) and radius (z). Since there is no well-developed mathematical model for a defocused PSF, it is difficult to perform fitting based algorithm in such images. A new particle localization algorithm based on radial symmetry and ellipse fitting is developed to localize the centers and radii of defocused PSFs. Our method can localize the 3D position of a fluorophore within 20 nm precision in three dimensions in a range of 40 μm in z dimension from defocused 2D images.

Table of Contents

List of Tables

List of Figures

Figure 3.7: Comparison of several particle localization algorithms and Deep Matching. (a & b) Localization accuracy for various particle localization algorithms applied to simulated particle images at SNR = 5 after Deep Matching. The x (a) or y (b) components of the difference between the algorithm-determined value and true particle position value is plotted as a function of the particle position value. RS: radial symmetry; GNLLS: Gaussian fitting using non-linear least-squares minimization; GMLE: Gaussian fitting with maximum-likelihood estimation. (c) The algorithm error from simulated particle images over a range of SNR, from SNR = 2 (37 photons detected) to SNR = 20 (3700 photons detected). Each point denotes the average of 1,000 tests at that SNR level. Markers of red circle, cyan triangle, and black diamond represent localization errors on original simulated images through RS, GNLLS, and GMLE, respectively. Markers of green square, purple asterisk, and orange cross denote localization errors on corresponding CNN+RS, CNN+GNLLS, and CNN+GMLE respectively. The red solid line (CRB1) indicates the fundamental Cramér-Rao bound of localization accuracy, the green solid line (CRB2) indicates the CRB of pixelated images, the blue solid line (CRB3) indicates the practical CRB of pixelated images with noise. (d) The algorithm errors of RS (RS0), GNLLS (GNLLS0), and GMLE (GMLE0) on simulated particle images with zero background Poisson noise. (e) The algorithm errors of RS (CNN+RS0), GNLLS (CNN+GNLLS0), and GMLE (CNN+GMLE0) on simulated particle images with zero background Poisson noise after CNN processing..... 57

Figure 3.8: Localization algorithm accuracy in the presence of an adjacent particle. Test images were constructed at SNr = 5, consisting of two particles at a given center-to-center distance and a random angular direction. Each point represent the average particle localization error of 1000 tests. Inset: an example image with two particles separated by 4 pixels and oriented at 135 degrees with respect to x axis. The yellow box with size of 7×7 pixels indicates the segmented region for employing CNN localization algorithms. 58

Figure 3.9: The accuracy of particle localization for simulated images of asymmetric PSFs. (a) The localization error from simulated particle images over a range of from SNR = 2 (37 photons detected) to SNR = 20 (3700 photons detected). Each point denotes the average of 1,000 tests at

Abbreviations

ANNs: Artificial neural networks

AUC: Area under the curve

CAD: Computer aided detection

CNNs: Convolutional neural networks

CPM: Competition performance metric

CRB: Cramér-Rao bound

CT: Computerized tomography

DNG: Divergence of the normalized gradient

DSC: Dice coefficient

FCNs: Fully convolutional neural networks

FN: False negative

FP: False positive

FROC: Free receiver operating characteristic

GLCM: Grey level co-occurrence matrix

GMLE: Gaussian maximum-likelihood estimation

GNLLS: Gaussian non-linear least-squares

IoU: Intersection over union

LBP: Local binary pattern

LDA: Linear discriminant analysis

LIDC-IDRI: Lung Image Database Consortium image collection

LUNA16: Lung Nodule Analysis 2016

MDS: Multidimensional scaling

MRF: Markov random field

MRI: Magnetic resonance imaging

MSE: Mean squared error

ODN: Object detection network

OHSS: Online hard sample selection

PALM: Photoactivated localization microscopy

PSF: Point spread function

R-CNN: Regional convolutional neural networks

ReLu: Rectified linear units

RNN: regression neural network

ROC: Receiver operating characteristic

ROI: Region of interest

RPN: Region proposal network

RS: Radial symmetry

RUS: Random under-sampling

SG: Savitzky-Golay

SGD: Stochastic gradient descent

SIFT: Scale-invariant feature transform

SMLM: Single Molecule Localization Microscopy Challenge

SNR: Signal-to-noise ratio

SPC: Spatial pooling and cropping

STORM: Stochastic optical reconstruction microscopy

SVMs: Support vector machines

TFM: Two-photon microscopy

TP: True positive

TPFM: Two-photon fluorescence microscopy

WT: wavelet transform

Chapter 1: Introduction

Image, a visual format to represent, measure, and reproduce a large variety of objects from different levels, is strongly associate with majority of scientific subjects. In terms of biomedical field, the development of methodologies and techniques of imaging has been revolutionarily expanding the horizon of discoveries. In the field of medical diagnosis, with the development of x-rays, computer tomography (CT), magnetic resonance imaging (MRI), positron emission tomography, precise clinical diagnosis and intervention for organ level of interior of body can be conducted. Over the several decades, medical imaging have been expanded in numerous branches, from the hardware to software. Image analysis, as a tip of iceberg, plays more and more significant role. The overall goal is to help radiologists to acquire direct and indirect imaging biomarkers so that more efficient and precise diagnostic decisions can be made. As for the field of microscopy, the development of some microscopy techniques such as two-photon fluorescence microscopy (TPFM), temporal focusing two-photon microscopy (TFM), and other superresolution microscopies, can explore specimens at cellular level with image resolution as high as nanometers. Image analysis on those microscopy images can improve both image quality and spatial resolution, which helps to retrieve much deeper and more detailed information from the raw scanned images. In this paper, we will discuss the applications of image analysis on both organ and cellular levels of images. For organ level images, fully-automated lung cancer diagnosis based on CT images will be demonstrated. For cellular level images, we will study image analysis algorithms on super resolution localization microscopy to improve the localization resolution from noisy raw images. The background and some state-of-the-art approaches from literatures will be discussed as follows.

1.1 Computer Aided Lung Cancer Diagnosis

Worldwide, lung cancer has been having the leading mortality rate of cancer deaths in both males and females for decades with 1.2 million global deaths a year (Fitzmaurice et al., 2017). Due to the inconspicuous symptoms, the majority of lung cancer cases are diagnosed at distant stages with only 4% five-year survival rate (Siegel et al., 2018). Early detection of suspicious pulmonary

nodules is crucial to improve the life quality of lung cancer patients. Currently, computed tomography (CT) is considered the best and most widely used imaging modality for early detection and analysis of lung nodules. However, because of the complicated morphological and anatomical appearance of nodules, the nodule identification would be largely dependent on the skill, experience, and vigor of the radiologists (Winkels and Cohen, 2018). After identifying the nodule, precise segmentation is significant for clinical measurements (such as diameter and volume), which objectively provides repeatability of diagnosis and consistency of image interpretation (Liu et al., 2018). Therefore, a fast and fully-automated computer aided detection (CAD) system on nodule detection and segmentation with limited number of false positives (FPs) will dramatically decrease the workload of radiologists as well as the cost of treatment.

1.1.1 Nodule Detection

For nodule detection many published works proposed a two-stage system, which includes a candidate screening step to rapidly extract nodule regions from pulmonary parenchyma and remove other structures, and a false positive (FP) reduction step to massively eliminate FP candidates from the detected ones until reaching clinically acceptable performance. Ge et al. (2005) used adaptive weighted k-means clustering to segment suspicious candidates, and reduced FPs by combing 3D gradient field and ellipsoid features with a linear discriminant analysis (LDA) classifier. Li et al. (2008) added a multiscale selective filter to enhance nodule and simultaneously suppress normal structures, and then a rule-based classifier was used to reduce FPs based on six shape features and twelve intensity features extracted from the enhanced images. Tan et al. (2011) improved the performance of candidate screening by introducing a maxima of the divergence of the normalized gradient (DNG) to find centers of nodule candidates with a merging stage to remove duplicates and further reduce the number of FPs. Forty-five invariant features, defined on a gauge coordinates system, are used to differentiate nodules from large amount of FPs. Finally, a novel feature-selective classifier based on genetic algorithms and artificial neural networks (FD-NEAT) was first implemented to improve the flexibility and adaptability of a classifier. Its

performance was compared with that of two other classifiers based on support vector machines (SVMs) and fixed-topology artificial neural networks (ANNs). Even though these traditional machine learning algorithms achieved remarkable accuracy in nodule detection, some disadvantages including but not limited to arduous human interventions, slow computation time, and mediocre representation capability of hand-crafted features obstruct the further development of traditional CAD system to deal with the large variations of lung nodules from real clinical CT scans. Recently, deep learning techniques in particular convolutional neural networks (CNNs) motivate flurry of researchers to develop powerful and robust algorithms on pulmonary nodule detection, which outperform many traditional machine learning approaches (Anirudh et al., 2016; Tajbakhsh and Suzuki, 2017; Sun et al., 2017a; Ypsilantis and Montana, 2016; Fu et al., 2017; Hamidian et al., 2017; Winkels and Cohen, 2018; Zhu et al., 2018). Setio et al. (2016) delineated multi-view 2D CNNs by taking advantage of nine symmetric planes of nodule cubes without increasing the network complexity. It is fed with nodule candidates obtained by three individual candidate screening algorithms that are exclusively designed for solid, subsolid, and large nodules respectively. The best detection accuracy is achieved by applying a dedicated mix-fusion method. Inspired by the 3D nature of pulmonary nodules, Huang et al. (2017) exploited a single-scale 3D CNNs to encode much richer and more comprehensive spatial contextual information compared with conventional 2D CNNs. Dou et al. (2017a) significantly boosted the detection accuracy through a multi-level 3D CNNs. In order to cover the large variations of nodules with different sizes of receptive fields, three independent 3D CNNs were involved to learn discriminative features for small, medium, and large size of nodules, respectively. It is a generic 3D CNNs framework that can in principle transfer to other applications to extract targets from variety of complicated mimics. Dou et al. (2017b) proposed a novel two-stage 3D CNN for end-to-end nodule detection with a 3D FCN based nodule candidates screening and a 3D hybrid-loss residual learning based FP reduction. They first tackled the severe imbalance problem of hard and easy samples by employing an online sample filtering scheme. This dynamic scheme naturally splits hard and easy samples based on the loss of each forward propagation of training, so that the training

convergence can be fastened. Ding et al. (2017) provided a combination of a 2D Faster R-CNN for the initial nodule candidate detection and a 3D CNN for FP reduction. The Faster R-CNN based nodule detection optimally ensured the sensitivity while maintaining low number of FPs. Jin et al. (2018) constructed a 27-layer 3D residual CNNs, which is much deeper and more effective than the traditional 3D CNNs. A spatial pooling and cropping (SPC) layer ensures the capability of learning multi-level contextual information using a single-scale 3D CNNs architecture. Such design overcomes the restriction of tedious parameter tuning while dealing with model fusion, and drastically accelerates the training and testing process. Moreover, an online hard sample selection (OHSS) unlocks the potential of network to detect extreme nodules with complex morphological characteristics. Table 1.1 lists some representative pulmonary nodule detection approaches. Even though flurry of powerful contributions have been proposed regarding pulmonary nodule, however, due to the nature of the prevalent two-stage nodule detection framework, some unignorable drawbacks still occur. First, the candidate screening, which should ideally detect all the suspicious nodule candidates, determines the upper-bound sensitivity of the entire CAD system. But morphological difference of nodules makes it impossible to achieve the optimal performance based on single or multiple hand-crafted mathematical models (Murphy et al., 2009; Jacobs et al., 2014; Setio et al., 2015), and the tedious experiment-based parameter adjustment restricts the applications onto real clinical trials. Moreover, the suboptimal segmentation of lung parenchyma negatively impacts the candidate screening especially for juxtapleural nodules (Dai et al., 2015). Second, because of the serial algorithm structure with many subcomponents, the long computation time stands out as another demerit. Therefore, a simpler and more independent nodule detection framework is urgently desired.

Table 1.1: A select listing of pulmonary nodule detection algorithms

Author	Dataset	Method	Automated ?	Nodule Types
Ge et al. (2005)	82 scans 116 nodules	K-means clustering	No	Circumscribed, juxtapleural, spiculated
Ginneken (2006)	23 scans 18 nodules	Learning based approach combining region grow, morphological operation and non-linear regression	No	Solid and nonsolid
Ozekes (2007)	12 scans 153 nodules	Rules and template matching trained by Genetic algorithm	Yes	-
Lee et al. (2008)	23 scans 1203 nodules	Two-stage segmentation based on decision tree classification	No	-
Li et al. (2008)	117 scans 153 nodules	Nodule enhancement filter with an automated rule-based classifier	Yes	Solid, pure and mixed GGO
Ozekes et al. (2008)	16 scans 425 slices	3D template matching with fuzzy rule based thresholding	Yes	-
Moltz et al. (2009)	28 scans 101 nodules	Smart opening algorithm	No	Juxtapleural
Messay et al. (2010)	84 scans	Classification based on 3D computed features	Yes	Isolated, juxtapleural, and juxtavascular
Tan et al. (2011)	125 scans	Nodule enhancement with feature-selective classifier based on ANNs and genetic algorithms	Yes	All types of nodules with diameter >=3mm
Kubota et al. (2011)	105 nodules	Morphological operations and convexity models	No	Solid, non-solid, part-solid, solitary, vascularized, and juxtapleural
Zhu et al. (2012)	41 scans	Markov random field based algorithm	Yes	Ground glass opacities (GGO)
Farag et al. (2013)	742 nodules	Level sets with shape prior	No	Juxtapleural, vascularized, well-circumscribed, and pleural-tail
Badura et al. (2014)	551 nodules	Fuzzy connectedness (FC) and evolutionary computation	No	-
Liu et al. (2015)	115 scans 363 nodules	Adaptive fuzzy c-means (FCM)	Yes	Solid, vascularized, juxtapleural, and GGO
Cavalcanti et al. (2016)	350 scans	Background estimation	No	-
Setio et al. (2016)	888 scans	Multi-view 2D convolutional neural networks (CNNs)	Yes	Solid, subsolid, and large
Dou et al. (2016)	888 scans	3D convolutional neural networks (CNNs)	Yes	-
Li et al. (2016)	1013 scans	2D convolutional neural networks (CNNs)	No	-

1.1.2 Nodule Segmentation

Because of the critical clinical value in nodule segmentation, a growing numbers of pulmonary nodule segmentation algorithms have been proposed in literature. They can be roughly categorized into four types: 1) Threshold based methods (Reeves et al., 2006; Magalhães Barros

Netto et al., 2012; Tachibana et al., 2006; Xia et al., 2016). For instance, Tachibana et al. (2006) designed a coarse-to-fine scheme that consists of a rough segmentation step using multiple fixed thresholds to roughly identify the nodule regions and a precise segmentation step using a watershed-based algorithm to remove the unnecessary structures attached to them. 2) Morphology based methods (Kubota et al., 2011; Dehmeshki et al., 2008; Vijaya Kishore et al., 2013; Lassen et al., 2015). For example, Dehmeshki et al. (2008) presented an efficient sphericity-oriented region growing algorithm applied on the fuzzy connectivity mask created by a connectivity region growing technique with only one single seed point provided by the user. 3) Statistical model based method (Wang et al., 2009; Dong et al., 2014; Tan. et al., 2013; Mao et al., 2018). Tan et al. (2013) utilized a hybrid algorithm combining marker-controlled watershed, geometric active contours as well as Markov random field (MRF). Similar to the method Tachibana et al. proposed, they imposed watershed method to generate an initial surface of nodule, followed by the refinement of active contours. And MRF optimally estimates the texture distribution of ground glass opacity, so that it improves the segmentation accuracy for this portion. 4) Clustering methods based on traditional machine learning (Van Ginneken, 2006; Tuinstra, 2008; Messay et al., 2015). Messay et al. (2015) proposed a selective regression neural network (RNN) based algorithm with both fully-automated and semi-automated options. The feature learning process using RNN can automate the parameter setting for each nodule based on the learned features. Table 1.2 lists some state-of-the-art methods. However, the majority of the aforementioned methods perform well only on specific type of nodules (e.g. solitary pulmonary nodule) or on relatively small size of dataset, which cannot satisfy the variety and complexity of pulmonary nodules. In addition, most of the methods still need human interventions, which largely undermine the purpose of CAD systems. Finally, in order to achieve optimal performance, most of the techniques require massive iterations and parameter tunings, which to a large extent, slow down the overall computation process. Recently, the success of semantic segmentation in computer vision field based on fully convolutional neural networks (FCNs) (Long et al., 2015; Wang et al., 2017; Chen et al., 2018; Lekić et al., 2018; Yu et al., 2018) attracts some researchers to concentrate on the application of

pulmonary nodule segmentation. Wu et al. (2018) firstly deployed an interpretable and multi-task CNNs model to segment and classify pulmonary nodules by feeding 3D patches and achieved the state-of-the-art performance. However, most of current deep learning based algorithms still rely on several preprocessing steps such as lung parenchyma segmentation, which decrease the level of automation.

Table 1.2: A select listing of pulmonary nodule segmentation algorithms

Methods	Nodule amount	IoU
Tachibana et al. 2006	23	50.7 ± 21.9%
Wang et al. 2009	64	58%
Messay et al. 2010	68	63 ± 16%
Kubota et al. 2011	23	69 ± 18%
	82	59 ± 19%
Tan et al. 2013	23	65%
Lassen et al. 2015	19	52 ± 7%
	40	50 ± 14%
Messay et al. 2015	66	71.70 ± 19.89%
	77	69.23 ± 13.82%
Wang et al. 2017	493	71.16 ± 12.22%
Ours	**223**	**70.24 ± 12.04%**

1.1.3 Nodule Malignancy Prediction

Fully-automated nodule malignancy prediction is demanding to alleviate non-invasive cancer treatment. In this paper, we introduced both traditional machine learning based and deep learning based approaches for nodule malignancy prediction. A single CT examination can generate up to 700 axial images creating a challenging task for image interpretation. From the reported literature, most research groups are mainly analyzing lung images via two-dimensional (2D) features either from only one single representative slice or from multiple slices. In addition, three-dimensional (3D) features are more descriptive than 2D features because they provide not only the complete information on every slice but also the connections between adjacent slices. A few researchers investigated 3D texture features to distinguish benign and malignant lung nodules.

In this study, we analyzed application of 3D texture features to classify lung nodules into different malignancy levels. To the best of our knowledge, no other research group has reported usage of 3D texture features to differentiate lung nodule malignancy levels. In addition, convolutional neural network (CNN) has been a trending technique for image analysis and computer vision in recent years. Some preliminary studies on medical images using deep learning algorithms have shown promising performances. This deep structured architecture gives computer the possibility to automatically learn features at multiple levels of abstraction without using human-crafted features. Each hierarchy layer of features is weighted combinations of lower level features. The structure of hundreds of neurons in each layer represents human brain's perception. However, applying deep learning algorithms to medical images remains many obstacles: First, deep learning algorithms usually contains a huge number of parameters, and the fine-tuning process requires large amount of training data. Many other computer vision tasks using deep learning algorithms have more than 1 million data, but for medical images, the data is difficult and expensive to collect and label. Second, convolutional neural network requires the input image of the same size, but the region of interest (ROI) usually has different sizes according to the actual size of suspicious area. Third, compared to other computer vision tasks the images usually are visually easily classified by human, the medical images are hard to diagnosis and read, even the well-trained radiologists can have very different diagnosis results with each other. The complexity character of medical images makes it more difficult to train a deep learning algorithm. In this study, we designed a novel scheme using CNN on lung CT images nodule diagnosis, and this scheme considered the three factors mentioned above. By sharing the experiences and tricks we used for lung nodule diagnosis, we hope it can be helpful for other medical image related deep learning study as well.

1.2 Superresolution Localization Microscopy

Super-resolution localization microscopy, such as photoactivated localization microscopy (PALM) (Betzig et al., 2006) and stochastic optical reconstruction microscopy (STORM) (Rust et al., 2006), has revolutionized the field of fluorescence microscopy. Localization microscopy

reaches resolution in the nanometer range by combining photoswitching mechanisms to sequentially image different fluorophores and algorithms to calculate their positions precisely. The development of localization algorithms for super resolution has been a key aspect in this research, and the imaging accuracy largely depends on the algorithm being used (Deschout et al., 2014; Small et al., 2014; Sage et al., 2015). Most current algorithms are based on either fitting the image of a single particle with a known point spread function (PSF) model or non-fitting approaches such as calculating the centroid of a particle PSF (Babcock et al., 2017). In addition, since localization microscopy is a two-dimensional wide-field imaging modality using CCD camera as the detector, acquiring three-dimensional (3D) super resolution image needs sophisticated PSF engineering to represent the depth information in the distinct characteristics of the microscope's PSFs, such as astigmatism (Huang et al., 2008), double helix (Pavani et al., 2009), Airy function (Jia et al., 2014), saddle-point, tetrapod (Shechtman et al., 2015; Shechtman et al., 2016), etc., which require appropriate fitting algorithms to retrieve such information. Choosing the right localization algorithm depends on the users' prior knowledge of the imaging system, such as the PSF, noise, and fluorophore properties. Such prior knowledge cannot be obtained perfectly in most real life experiments. Also the signal-to-ratio (SNR) of the raw images in super resolution experiments is typically low due to the requirement that only a few fluorophores can be turned on stochastically within the acquisition time of one frame of image. Therefore, learning noise model directly from image itself instead of building complicated mathematical model that optimizes specific type of noise and with good accuracy under low SNR conditions can significantly improve the imaging speed and localization accuracy. Recently, a number of researches conducted deep learning based localization algorithms (Ouyang et al., 2018; Aristov et al., 2018; Schnitzbauer et al., 2018; Nehme et al., 2019; Hershko et al, 2019; Sage et al., 2019). However, such methods are sensitive to many constraints, such as the shape of PSF0 and the appearance of reconstructed image. An independent preprocessing step that is compatible to any localization-based methods is desired.

Another approach to obtain depth information from a 2D image is based on defocused imaging. The PSF in a defocused image is made of a central spot and multiple concentric rings

due to Fresnel diffraction (Speidel et al., 2003). The radii of the rings are correlated to the defocus level, thus the depth (z) information. By locating the centroid and measuring the radius of the outermost ring the 3D position of a single fluorophore can be determined. Such defocused imaging has been used to track single particles with Ångström accuracy (Huhle et al., 2015). Recently we have developed a temporal focusing two-photon microscope that can track single fluorophores at nanometer precision with a depth range of 100 μm (Ding et al., 2016). The temporal focusing two-photon microscope is a wide-field 2D imaging modality. It can achieve z-sectioning capability (axial resolution on the order a few micrometers) by stretching the femtosecond laser pulse in temporal domain before the objective lens and compressing it to its shortest temporal width at the focal plane of the objective (Oron et al., 2005; Zhu et al., 2005). When a fluorophore is out of the focal plane, it forms a defocused image on the CCD camera. We have implemented the same calculation method to obtain its 3D position at 50 nm precision from such defocused images.

There are two obstacles preventing the wide adoption of defocused imaging for 3D localization. First, to our best knowledge, there is no well-developed mathematical model for a defocused PSF (Gibson et al., 1989), which makes it difficult to develop localization algorithms based on fitting a known PSF. Second, the signal-to-noise ratio (SNR) in a defocused image is substantially lower than that in an in-focus image, since the emitted light is spread out and detected by many pixels of the CCD camera instead of only a few. Furthermore, the excitation efficiency is usually lower than that in in-focus imaging. Therefore, developing novel localization algorithms will greatly improve the capability of defocusing imaging in 3D localization of particles.

Recently deep learning has achieved stellar performance in pattern recognition and decision making. Central to this success is the convolutional neural network (CNN) architecture with deep layered structure (LeCun et al., 2015; Shen et al., 2017; Wang et al., 2018). The layers in CNNs are restricted to perform convolutions, which greatly reduces the number of parameters to be learned. In image processing CNNs have been applied to image and video super-resolution where a high-resolution image is obtained from the input of only a single frame low-resolution image (Dong et al., 2016; Kappeler et al., 2016). In biomedical imaging research the same concept

has been applied on improving the resolution of optical microscopic images (Rivenson et al., 2017). Most recently CNNs were used to obtain super resolution PALM/STORM images from a sequence of raw images without using the conventional single molecule detection and localization processes (Nehme et al., 2018; Boyd et al., 2018; Wang et al., 2018; Ouyang et al., 2018). The obtained image resolution is comparable to the resolution obtained with conventional algorithms.

Chapter 2: Computer Aided Lung Cancer Diagnosis

In this section, we propose a CNNs based algorithm to automatically detect and segment pulmonary nodules with very limited number of FPs. Comparing with methods using the original nodule patch only, our method improves the area under the curve (AUC) malignancy prediction by 4% with the combination of the original nodule patch, the segmented nodule patch highlighting the shape, and the gradient of nodule patch highlighting the texture with accurate boundary information provided (Sun et al., 2017b). To our best knowledge, this is one of the first works that exploit pure 2D CNNs based algorithm for pulmonary nodule segmentation from the raw CT scan without any manual settings. The output is the corresponding nodule masks with high overlapping ratio compared to radiologists' markings. This study is the follow up study of our previous researches (Qian et al., 1993; DeVore et al., 1995; Sun et al., 2004; Zhang et al., 2007; Ye et al., 2013; Sun et al., 2017a, 2017b, 2016, 2017c, 2017d, 2017e). The reason of using 2D axial slice instead of 3D volume in our system is three-fold. First, a powerful deep learning model relies on large amount of training data. However, due to the stressful workload of manual nodule identification, the available thoracic CT scans with high quality groundtruth annotations are still insufficient to train a robust and discriminative deep learning network. Since a nodule may appear in several neighboring 2D slices of a single CT scan, using 2D instead of 3D slices will naturally augment the size of training set without manual data augmentation process. Second, 3D deep learning models will exponentially add computation complexity compare to 2D models. Therefore, some models require image down sampling or cropping (Jin et al., 2018) to compensate the massive memory consumption. This requirement, to some extent, restricts the feasibility of 3D models on common workstation with limited GPU resources. Third, transfer learning (Weiss et al., 2016; Long et al., 2016) can be performed based on some fine-tuned 2D models such as VGG16 (Simonyan and Zisserman, 2015), ResNet-50, and ResNet-101 (He et al., 2016), significantly accelerates the convergence of training, boosts the performance, and cuts down the computation complexity.

2.0 Hypothesis

It is hypothesized that, with the implementation of traditional image processing techniques and machine learning algorithms, pulmonary nodules from CT images can be precisely detected, segmented, and predicted. And such achievement can be served as an objective assistant for radiologists to make better diagnosis and treatment plans.

2.1 Dataset

Similar to most of aforementioned literatures, we train and evaluate our fully-automated system on a large publicly available dataset, organized by Lung Image Database Consortium image collection (LIDC-IDRI) (Armato III et al., 2015, Armato et al., 2011, Clark et al., 2013) with 1,018 scans from seven academic centers and eight medical imaging companies. The slice thickness of these scans is ranging from 0.6 mm to 5.0 mm. For each scan, four experienced radiologists performed two-phase nodule assessment and recorded the detailed nodule information such as boundary coordinates, malignancy level etc. into an XML file. In this data set three types of lesions are included: non-nodule, small nodule (< 3 mm), and large nodule ($\geq$ 3 mm). According to some clinical recommendations (Aberle et al., 2011), we only consider large nodules ($\geq$ 3 mm) in our study. The pre-filtering strategy proposed in the Lung Nodule Analysis 2016 (LUNA16) challenge only selected nodules with the consensus of at least three out of four radiologists with slice thickness no more than 2.5 mm. As a consequence, 888 CT scans with 1,186 nodules are involved.

2.2 Methodology

The framework of our proposed nodule segmentation algorithm is shown in Figure 2.1. It can be simply divided into four main components: 1) 2D Faster R-CNN based candidate detection to rapidly locate pulmonary nodule patches; 2) merging overlapping candidates by combining 2D patches with close Euclidean distances; 3) traditional three-layer 2D CNN based FP reduction to further eliminate FPs; and 4) modified FCNs based nodule segmentation to precisely segment the initial nodule mask. The geometric centers of detected nodules will guide the system to refine the

segmentation and output the final nodule segmentation result. The more detailed explanation of each component is provided below.

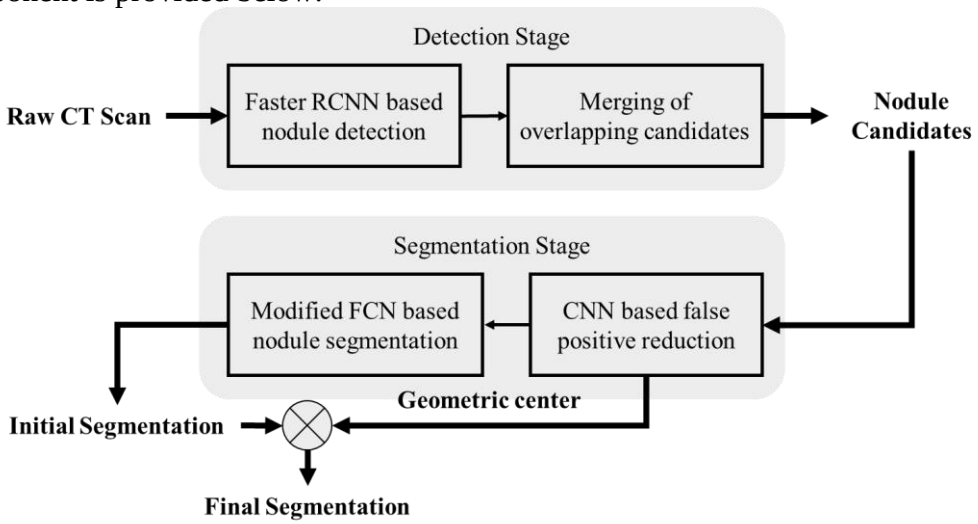

Figure 2.1: Top level framework of CNNs based nodule detection and segmentation system.

2.2.1 Preprocessing

In order to reduce memory consumption, we assigned -1000 HU (air) as lower bound and +3000 HU (bone) as upper bound, then applied a linear mapping to convert the original signed 16-bit CT volume scans into 8-bit intensity values in the range of 0-255. This is the only preprocessing work in our proposed approach.

2.2.2 Faster R-CNN based nodule detection

The Faster R-CNN model evolved from Fast R-CNN (Girshick, 2015), which is mainly composed of a region proposal network (RPN) to propose potential regions of objects and an object detection network (ODN) to classify the region proposals from RPN. The new contribution of Faster R-CNN is that RPN and ODN can share the same convolutional layers, which enabled a more unified system to run at near real-time frame rates on natural images without performance loss (Ren et al., 2017). Based on our experiments, it is fully convertible to be applied on nodule detection after conducting several modifications. The network architecture is shown in Figure 2.2.

To accelerate training convergence and save computation workload, we reused the weights of first five groups of convolutional and pooling layers with a total of thirteen consecutive convolutional (Conv) and five pooling layers from a pretrained VGG16 model, and the calculated feature maps are shared with both RPN and ODN training. For RPN, we utilized a small network by sliding a 3×3 window at a time over the shared feature space to convert this feature map to a 512-dimension feature vector followed by a ReLU layer (Nair and Hinton, 2010). We implement the aforementioned small network by a 3×3 convolutional layer. After two sibling 1×1 convolutional layers, an RPN regression (rpn_reg) layer outputs the bounding box coordinates of each proposal, and an RPN classification layer (rpn_cls) estimates the probability of the proposal being a nodule. The design of anchors ensures the capability to parallelly predict multiple nodule proposals at each sliding window location. Different from common objects in natural images with big sizes and elongated shapes, the nodules are with relatively small and square boundaries. Therefore, we used a fixed scale ratio (1:1) and removed the other scale ratios in original RPN design, and implemented seven anchors with ascending common differences: 4×4, 6×6, 10×10, 16×16, 24×24, 36×36, 52×52 to fit the size variations of nodules. Because of the sparse distribution of nodules, we also adopt non-maximum suppression (NMS) based on the scores of rpn_cls with intersection over union (IoU) threshold of 0.7 between groundtruth and RPN proposals. NMS massively reduces the number of proposals and also potentially improves the detection accuracy, since high-density proposals may cause RPN to identify many surrounding regions that partially overlapped with true nodules (please see Figure 2.3 as an example). The learnt proposals from RPN are fed into ODN for further classifications.

Taking the proposals predicted by RPN, the ODN is involved to serve as a binary classifier to determine nodule regions using the architecture of traditional CNNs. A ROI pooling layer is imposed to map each proposal to a smaller feature map by implementing max pooling operation of the values in a fixed 7×7 sub-window. Then two 4096-way fully-connected layers (FC1 and FC2) are conducted to produce a lower dimension feature vector, followed by two independent Softmax layers to output bounding boxes and probability scores of predicted nodules (*box_reg* and

box_cls). In Faster R-CNN, RPN and ODN are mutually finetuned by adopting a pragmatic four-step alternating training procedure (Ren et al., 2017). As such, we achieved a unified network with sharing convolutional layers for RPN and ODN and the loss function for a single batch of *N* images is defined as follows.

$$\mathcal{L}_t = \frac{1}{N}\left[\frac{1}{\mathrm{N}_{rr}}\sum_i \mathcal{L}_r(t_i, t_i^*) + \frac{1}{\mathrm{N}_{rc}}\sum_i \mathcal{L}_c(p_i, p_i^*) + \frac{1}{\mathrm{N}_{br}}\sum_j \mathcal{L}_r(t_j, t_j^*) + \frac{1}{\mathrm{N}_{bc}}\sum_j \mathcal{L}_c(p_j, p_j^*)\right] \quad (1)$$

$$\mathcal{L}_r(t, t^*) = \mathrm{R}(t - t^*) \quad (2)$$

$$\mathcal{L}_c(p, p^*) = -log\,[pp^* + (1-p)(1-p^*)] \quad (3)$$

where N_{rr}, N_{rc}, N_{br}, and N_{bc} are numbers of inputs in *rpn_reg*, *rpn_cls*, *box_reg*, *box_cls* layers respectively, $\mathcal{L}_\mathrm{r}$ and $\mathcal{L}_\mathrm{c}$ represent loss associated the regression and classification layers, t_i represents the four coordinates of predicted nodule proposal, t_i^* is the coordinates of the corresponding groundtruth nodule, p_i and p_i^* denote the predicted and true probability of current anchor to be a nodule in RPN. Similarly, t_j, t_j^*, p_j, p_j^* implicit the same concepts in ODN. Besides, R is a robust loss function (smooth L_1 loss) explained in (Girshick, 2015).

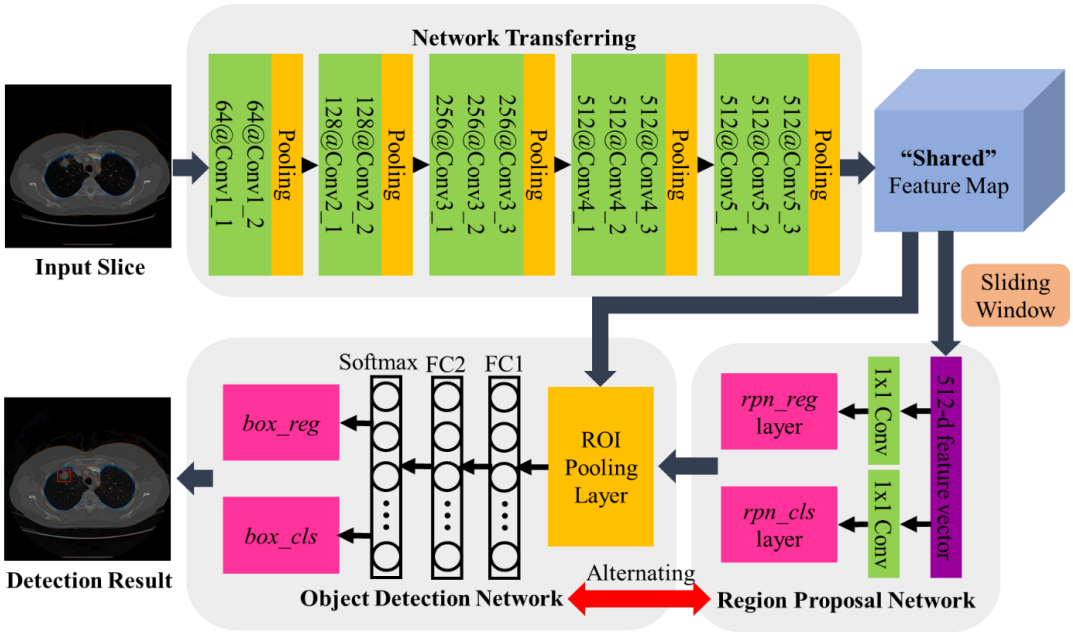

Figure 2.2: The architecture of proposed Faster R-CNN based nodule detection

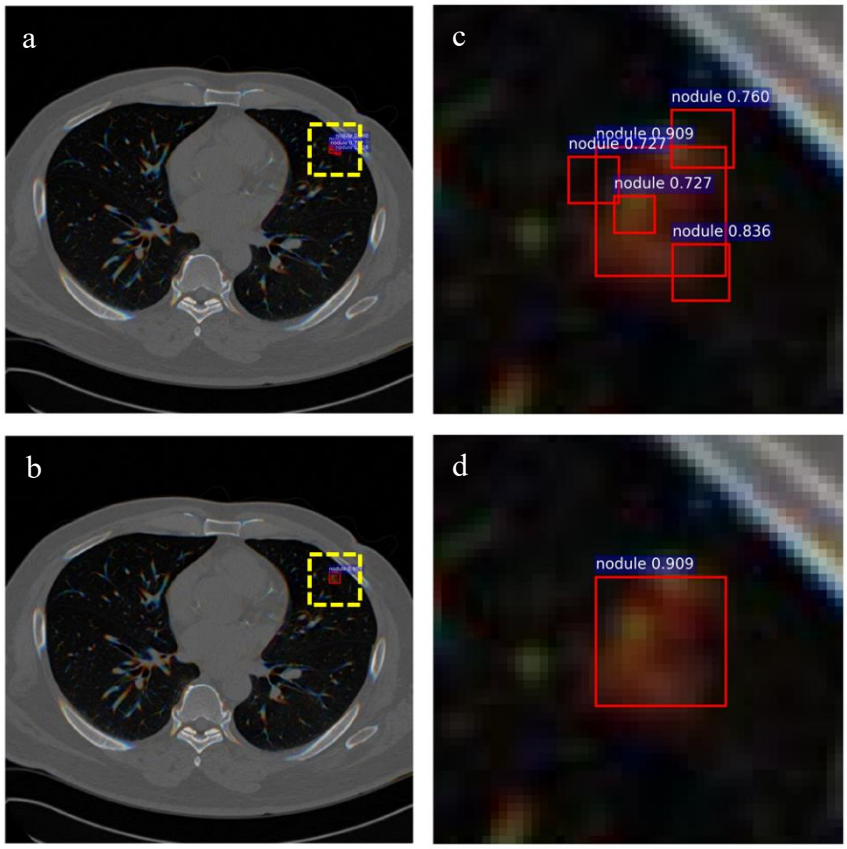

Figure 2.3: Comparison of nodule candidate detection with and without non-maximum suppression (NMS) operation. (a) Nodule candidates without NMS operation. (b) Nodule candidates with NMS operation. (c) and (d) are zoom-in view of the region marked by yellow dashed square in (a) and (b). The predicted nodule candidates by Faster R-CNN are marked by red solid squares with the classification probability on top of them. Note that only the candidates with classification probability of nodules larger than 0.7 are displayed in this example. With NMS operation, our system successfully detects a ground glass opacity (GGO) nodule without partially overlapped duplicates.

2.2.3 Merging of overlapping nodule candidates

After Faster R-CNN, most nodule candidates can be detected. However, a single nodule may appear in several slices, therefore it may have more than one candidate representing the same nodule especially for some nodules with blurry edges. Based on some intuitive observations, these

candidates will lie in close proximity to each other. Thus, a simple and computation-efficient merging operation is implemented by recursively combining candidates within five voxels of each other until no further merge is needed. This merging procedure ensures that a single nodule is identified within a single 2D slice rather than multiple slices alongside each other, dramatically decreases the number of unnecessary detections, and fastens the processing speed of the following FP reduction since the base number of candidates are lowered.

2.2.4 CNN based false positive reduction

With the nodule candidates extracted by Faster R-CNN, true nodule regions are successfully identified with small amount of FPs. However, the existence of FPs still prohibits the use of CAD system in clinical practice. Considering the requirement of computation time and the advantages of Faster R-CNN component, a simple 2D CNN based classifier is sufficient to handle FP reduction task. Based on our previous work (Sun et al., 2017a) with the addition of multi-view nodule patches obtained from nine symmetrical planes presented in a previous work (Setio et al., 2016), an individual CNN with three pairs of convolution-pooling layers and one fully-connected layer is implemented. The loss is calculated by using the cross-entropy error, and weights are updated using mini-batches of 128 images. Then the testing is incorporated based on the Faster R-CNN initial candidate detection results. The kernel size for each convolutional layer is 5×5, 3×3, and 3×3 and the numbers of filters are 24, 48, 64, respectively. The input image size is 64×64 for both training and testing sets. After FP reduction step, the occurrence of FPs is largely decreased while maintaining a high detection sensitivity. Since the FP reduction network will be executed after Faster R-CNN initial candidate detection, different initial candidate FP levels with corresponding sensitivity may cause variations on the overall performance. In order to achieve the best performance, we empirically set three initial candidate FP levels (small, medium, large) and individually employ FP reduction on each FP level.

2.2.5 Modified FCN based nodule segmentation

FP reduction eliminated the most unlikely nodule candidates detected by Faster R-CNN. Among the remaining candidates, the 100 by 100 patches containing the nodule detection bounding boxes are created. Compared to FP reduction network that conducts a single imagewise two-class classification, more precise pixelwise classification is performed for segmentation purpose. Therefore, larger receptive field with richer background texture is needed. The modified deconvolutional neural network (Long et al., 2015) is used to generate the detailed nodule segmentation contour. In this experiment, the VGG16 is imposed as the backbone, and the weights of all the convolutional layers are initiated by ImageNet VGG16 pretrained model, while the weights of the later deconvolutional layers are randomized. In VGG16, there are five groups of convolutional layers altogether (Conv1 to Conv5), and each group contains a few consecutive convolutional layers. The original FCN used three convolutional layers (Conv3, Conv4 and Conv7) to generate the segmentation results. Since the lower convolutional layers have higher resolution, incorporating these layers should help with segmentation precision. In this study, two extra rounds of training are added to FCN training procedure to incorporate the first two convolutional layers (Conv1 and Conv2) to deconvolutional networks as well. The architecture of our modified FCN is shown in Figure 2.4.

To initiate the FCN, the ImageNet VGG16 weights are downloaded and a classification FCN is trained based on it. Each convolutional layer provides the local neighborhood information of the image, and each pooling layer down-samples the image by the stride of 2. To fine tune the VGG16, the third fully connected layer (fc8) is removed, and the output node is set to 2 since our dataset has two subsets: nodule and non-nodule. Then all the fully connected layers are converted to convolutional layers, and all the weights are preserved in the transformed model.

Then the fine-tuned VGG16 weights are used to initiate FCN32s, and an extra deconvolutional layer connected to the last convolutional layer is initialized with random noise attached to the end. Then we up sample the deconvolutional layer result by 2 to generate the FCN32s segmentation mask. The added deconvolutional layer and up sampling layer can be

considered as a block, after FCN32s being trained, another block is attached to the end to train FCN16s. Because there is another up sample layer in FCN16s compared to FCN32s, the original deconvolutional results are twice as large. Different deconvolutional results are up sampled to the identical size as the original image, and concatenated together to generate the segmentation output. The whole process is repeated 5 times so the FCN2s output has the same size and resolution of the original image. All the deconvolutional layers are up sampled at the same size and concatenated together. Compared to original FCN8s, our model utilizes the lower level convolution results thus yielding higher resolutions and preserved the details of the nodule. Since segmentation is indeed a pixel wise classification task, we still use cross entropy loss as the loss function during the FCN training.

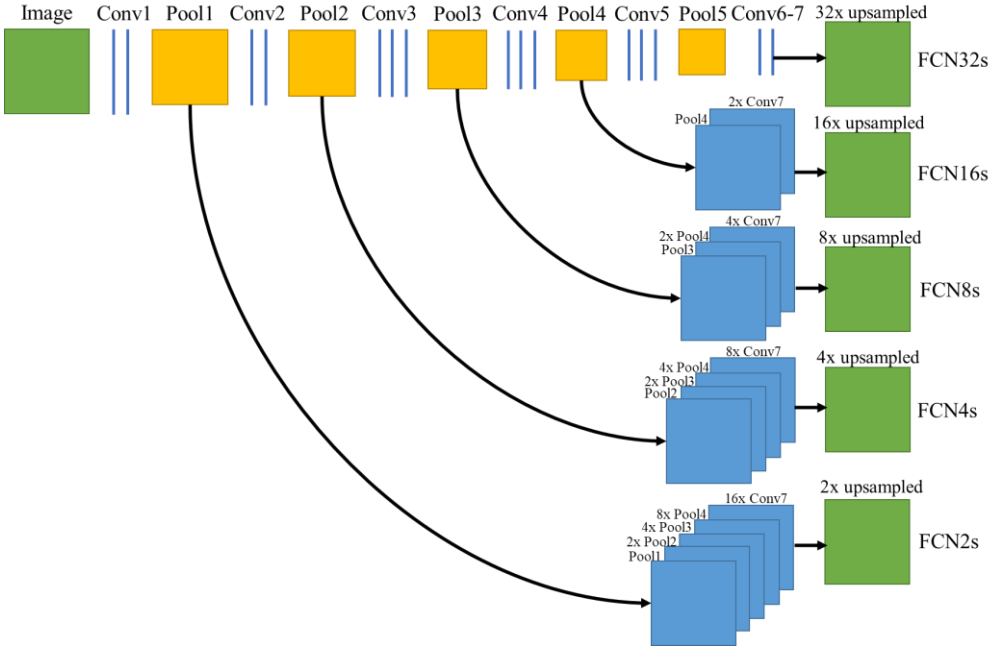

Figure 2.4: The architecture of our modified FCN

2.2.6 Postprocessing

The final segmentation results are obtained after the fusion of detected nodule centers and initial segmentation masks to remove non-nodule segmentations. This fusion operation only selects the corresponding segmented object with shortest Euclidean distances between detected nodule center and object centers in the initial segmentation mask.

2.2.7 Malignancy Level Analysis

In this section, we developed both machine learning based and deep learning based classification for malignancy levels. For machine learning based classification, we not only distinguish benign and malignant nodules but also distinguish different malignancy level cases individually. For CNNs based classification, we selected manual markings from the consensus of four radiologists and distinguish benign and malignant nodules.

. For majority of the nodules, four experienced thoracic radiologists diagnosed them and marked the boundary of each nodule with the greatest in-plane dimension larger than 3 mm. Besides the nodule boundary information, other nodule characteristics such as subtlety, internal texture, and likelihood of malignancy were also included in the annotation files. All the selected nodules were marked by all four radiologists. Because of the inter-observer variation in defining nodule boundaries, we chose the union region from the markings of all the radiologists. Because of the different CT scanning protocols across different vendors, the volume resolution varies across the dataset with the range from 0.5 mm to 3 mm. To avoid the partial volume effects, bi-cubic interpolation method was used to normalize CT volumes and boundaries marked by the all four radiologists resulting in isotropic resolution. Each radiologist gave a malignancy likelihood rating score in the range of 1 to 5, with 1 representing highly unlikely for cancer and 5 representing highly suspicious for cancer. Figure 2.5 shows an example of five different nodules with malignancy level 1 to 5.

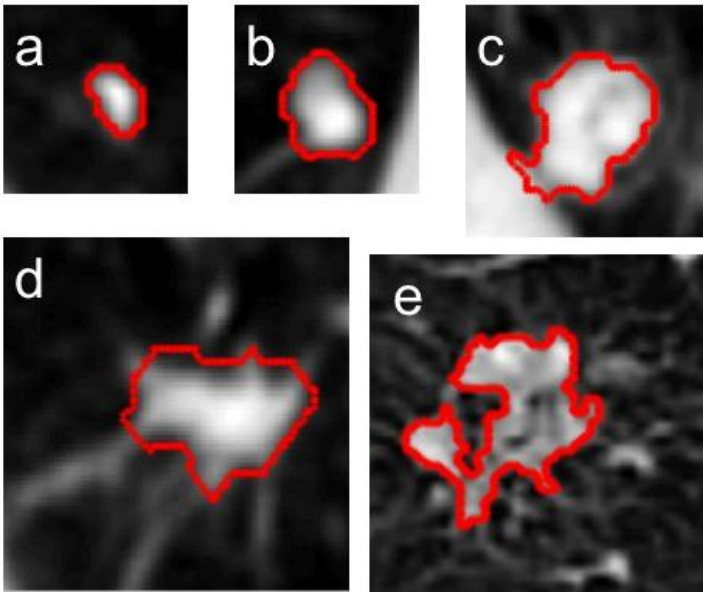

Figure 2.5: An example of nodules with different malignancy levels. Figure a - e represent malignancy level from level 1 to level 5 respectively.

We modified and redesigned the 2D features to 3D features. For each cubic ROI, five different types of 3D texture features were extracted: grey level co-occurrence matrix (GLCM) feature, local binary pattern (LBP) feature, scale-invariant feature transform (SIFT) feature, steerable feature, and wavelet feature. The number of features for each feature group is summarized in Table 2.1.

The first group of feature is 3D GLCM feature. In grey level co-occurrence matrix, the number of rows and columns is equal to the number of gray levels in original image, and each element represent the relative frequency of two pixels with given intensity separated by a pixel distance. In 3D cases, we calculated GLCM in 13 directions, including (0, 1, 0), (-1, 1, 0), (-1, 0, 0), (-1, -1, 0), (0, 1, -1), (0, 0, -1), (0, -1, -1), (-1, 0, -1), (1, 0, -1), (-1, 1, -1), (1, -1, -1), (-1, -1, -1), (1, 1, -1). Then the mean and standard deviation were computed for every matrix.

The second group of feature is 3D LBP feature. From the equally divided blocks of the original image, the extracted LBP features can describe the local and global textures. We computed the LBP features by comparing each pixel with the 26 neighbor pixels, and it returns the 26 digits

binary code for each pixel. Then the local binary pattern feature histogram was calculated from the coded image.

In the third feature group, we computed 3D SIFT features. This scale invariant feature transform is invariant to uniform scaling, orientation, and it is also partially invariant to affine distortion and illumination changes. Because of this property, it is a classical algorithm in object matching and action recognition. When creating descriptors, the sub-region was considered in our algorithm instead of every pixel, thus it is more adaptive to the image noises and subtle distortions. In our experiment, we took 8 by 8 by 8 neighbors of the key point descriptor and divided them into eight blocks. Each block we measure the 27 bins orientation histogram, and altogether we get 216 bins histogram to describe each key point. Mean and variance were measured for all the key points in one image.

In group four, we tested 3D steerable features. The steerable filters are linear combination of a set of basis filters with arbitrary orientations. Similar to 2D classical steerable feature extraction, in our 3D procedure we applied non-maximum suppression to the filtered images. From the response images, we calculated the mean, variance, skewness, entropy, and energy.

The feature group five contains 3D wavelet features, and they provide the spatial and frequency representation of the CT images. In 3D wavelet transform, high-pass and low-pass filters are used in all the three dimensions. For each scale, one input image can generate eight sub-band images: HHH, HHL, HLH, HLL, LHH, LHL, LLH, and LLL, where H represent high frequency band and L represent low frequency band. In this study, we used two different scales, and thus generating 15 different images. For each image, the mean and variance were calculated, so 30 wavelet features were generated in this group.

Table 2.1: Number of calculated features in each feature group

Feature group	Number of features
1	26
2	256
3	532
4	5
5	30
Total	849

Because we have a large number of computational features compared to the dataset, dimension reduction techniques were used. In this study we used multidimensional scaling (MDS), which is an algorithm placing each object in N dimensional space and at the same time preserving the between object distances. Since the malignancy ratings are not evenly distributed in our dataset, we used a boosting method called RUSBoost as our classification algorithm. It is a hybrid method that combines random under-sampling (RUS) and the standard boosting procedure AdaBoost resulting in efficient modeling of the minority class by removing majority class samples. To fully take advantage of the limited dataset, we used leave-one-case-out method for evaluation, so every data will be used once for as testing data.

For CNN based method, in order to generate a good number of ROIs to train CNN, every slice from one CT nodule volume was used except the top slice and bottom slice. After the nodule was located, a rectangular box with an extra 5 pixels surrounding areas of the nodule region was extracted. Each ROI slice was rotated at 90, 180 and 270 degrees, and we treat each radiologist's notation as an individual sample. The input of the CNN requires the image patches of same size, so we resized the ROIs to 52 by 52 pixels. All the ROIs larger than 52 by 52 pixels were downsampled to 52 by 52 pixels and all the ROIs less than 52 by 52 pixels were placed in center of the 52 by 52 matrix, and more surrounding regions were filled into the matrix if needed. Among all the 174412 obtained ROIs, 54880 samples are benign and 59848 are malignant.

From each original ROI, we obtained three other ROIs of the same size. By removing the surrounding areas, we generated the ROI with nodule pixels only; by removing the nodule pixels,

we got the ROI only contains surrounding region; by calculating the gradient on the original ROI, the gradient ROI was obtained. For the surrounding region ROI, we only want to obtain the textures of the surrounding regions, so we extracted the mean value of the surrounding region on each of the pixels. From these four one channel ROIs, we generated two multichannel ROIs. Each multichannel image contains three gray images, so it visually looks like color images. But instead of red, green, blue channels, the multichannel image contains three different ROIs and with each as an independent channel. The first multichannel ROI is generated by incorporating original ROI, segmented nodule ROI and surrounding region ROI; and the second multichannel ROI is the combination of original ROI, segmented nodule ROI and gradient ROI. Figure 2.6 shows the examples of the ROIs mentioned above.

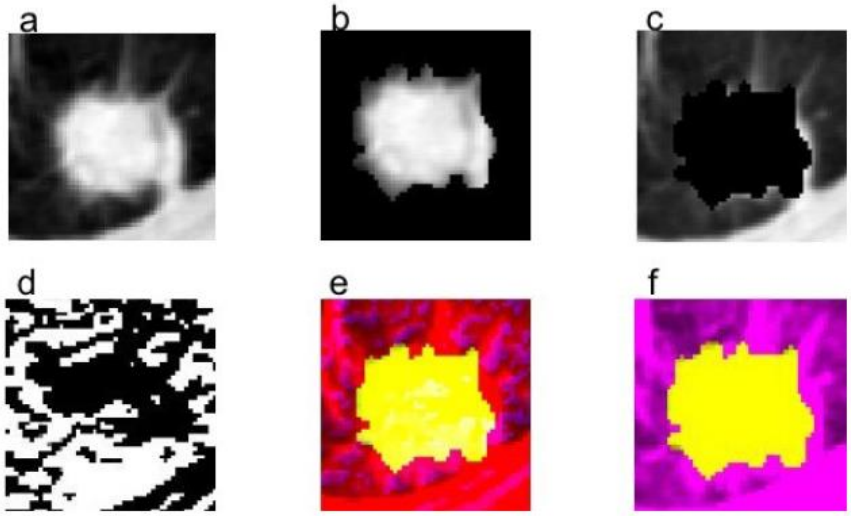

Figure 2.6: Examples of ROIs channels. a) nodule with surrounding ROI; b) nodule only without surrounding ROI; c) surrounding only ROI; d) gradient ROI; e) three channel ROI combined a, b, and c ROIs; and f) three channel ROIs combined a, b, and d ROIs.

The CNNs we used contains three pairs of convolutional layers and two fully connected layers. The max pooling layers were used to downsample the size of patches to prevent over-fitting. The downsample rate was set at 2. The number of kennels for each convolutional layer was 12, 8, and 6, and the kernel size was 5 by 5. The output neuron for the first fully connected layer

was 50 and the second fully connected layer was 2. We tested all six different ROIs mentioned above and 5-fold cross validation was employed to evaluate the performance. All the slices, rotations, segmentations from one case were kept in single fold.

2.3 Computational Results

2.3.1 Experimental Design

For the Faster R-CNN based nodule detection, we take the 888 filtered CT scans with 1,186 high confidence pulmonary nodules from LUNA16 dataset and retrieve their original annotations with precise nodule boundary coordinates recorded in XML format from LIDC-IDRI. To create the bounding square for each nodule, we calculate the smallest bounding rectangle and treat the longer side as the side of the result bounding square and record the four edge coordinates in groundtruth file. Because of the 2D network design, a total of 7,909 axial nodule slices are generated. Leveraging from the ten patient-level subsets by LUNA16, we include nodule slices of subset0 to subset5 as training, subset6 to subset7 as validation, and subset8 to subset9 as testing. As a consequence, 534 scans with 5,040 nodule slices, 178 scans with 1,411 nodule slices, and 176 scans with 1,458 nodules slices are assigned to training, validation, and testing set, respectively. For the training step, we adapt backpropagation and stochastic gradient descent (SGD) (LeCun et al., 1989) to train our end-to-end network. The minibatch size of images per iteration is set as 2 and the minibatch size of anchors per image is set as 10. For each image, based on the number of positive anchors (nodule), we randomly crop equal number of negative anchors (background) so that the sampled positive and negative anchors have the ratio of 1:1. Except for the first five groups of convolutional layers from VGG16, we randomly initialize all the other layers by drawing weights from a zero-mean Gaussian distribution with standard deviation 0.01. The learning rate is 0.001 for the first 10 epochs, and 0.0001 for the remaining with the momentum of 0.9 and a weight decay of 0.0005. Based on the aforementioned settings, the loss is stable after 20 epochs training. The total training time for Faster R-CNN is about 2.5 hours. For the modified FCN based nodule segmentation, the learning rates for the five deconvolutional training are set to

10-12, 10-14, 10-16, 10-14, 10-15, respectively. The momentum is set to 0.99, weight decay is 0.0005, the iteration is set to 50,000 for each round of training. Besides the initialization round, we train FCN with five rounds manner, adding one fully connected layer each time. Each round of FCN training is about 5 hours, and total training time for FCN is about 30 hours.

All the training and testing of CNNs are implemented using Tensorflow (Abadi et al., 2015) on a standard PC with a single 6GB memory GPU GeForce GTX 980Ti.

2.3.2 Evaluation Metrics

To evaluate nodule detection accuracy we calculate the detection sensitivity and the corresponding average FPs per scan for every prediction probability level, and the Free Receiver Operating Characteristic (FROC) is imposed to visualize the sensitivity versus FPs relationship. A predicted nodule candidate will be counted as a true positive (TP) if it stays within five voxels to the real nodule center. The competition performance metric (CPM) score (Niemeijer et al., 2011) is calculated by taking the average detection sensitivity of seven predefined FPs per scan: 0.125, 0.25, 0.5, 1, 2, 4, and 8.

With regards to the nodule segmentation, we evaluate the segmentation similarity between automated segmentation and groundtruth mask through the calculation of the dice coefficients (Sørensen, 1948), which is defined by the following equation.

$$DSC = \frac{2TP}{2TP+FP+FN} \tag{4}$$

2.3.3 Nodule detection

After nodule candidate detection by Faster R-CNN, we apply a merging step to combine the candidates that are geometrically close to each other. The number of candidates before and after the merging process is shown in Figure 2.7. The number of candidates decrease by 70.9%, 70.9%, 70.7%, and 70.3% at cut-off prediction probability threshold of 0.2, 0.4, 0.6, and 0.8, respectively, which verifies the superiority of using the merging operation as a simple postprocessing step in 2D nodule candidate detection. The FROC curve of our Faster R-CNN based nodule candidate detection is shown in Figure 2.8a. When the threshold of prediction

probability is set as 0.5, we achieve the sensitivity of 95.2% with an average of 19.8 FPs per scan, which outperforms the traditional nodule detection algorithm in LUNA16 with an overall sensitivity of 94.4% with an average of 620.6 FPs per scan. Our proposed merging method massively reduces the burden of FP reduction task. The FROC curves after 2D CNN based FP reduction are shown in Figure 2.8b-d by taking initial candidate FP level of 10 (L10), 15 (L15), 20 (L20) from Faster R-CNN results. The CPM score for L10, L15, and L20 are 0.866, 0.875, and 0.880, respectively. The best CPM is achieved in L20 with sensitivity of 91.4% and 94.6% at an average of 1 and 4 FPs per scan, respectively. The number of trainable parameters of Faster R-CNN, FP reduction, and modified FCN networks are 2.4×106, 7.1×104, and 1.3×108. The 3D U-net-based architecture in CASED (Jensen et al., 2017), which has identical CPM value compared to our approach, has approximated 6.6 × 106 learning parameters. Our Faster R-CNN + CNN model for detection has about 36% parameters of CASED model. By investigating different initial candidate FP levels from detecting network, these results show that a less conservative initial candidate selection threshold yields stronger performance (better CPM value), when FP reduction is employed. Moreover, a more progressive initial candidate selection threshold only performs better sensitivity at smaller FP levels (less than 0.25 FPs per scan in our case) in FROC curve.

Comparing with the top 15 models under the nodule detection track in LUNA16 challenge that employed 3D CNNs, our 2D nodule detection method successfully outperforms some of the 3D methods (Row 10 to 15 in Table 2.2) while keeping the computation and memory usage efficiency. Different from the methods (Row 1, 2, 3, and 5) using multiple powerful GPUs, our 2D model shows the capability of using cost-efficient GPUs. For instance, compared with Zhu et al. 2018 that employed a 3D Faster R-CNN network with 3D dual-path blocks and U-net-like encoder-decoder structure to compactly and effectively exploit features, our light and straightforward approach has much smaller number of learning parameters, thus the whole 2D slices that contain the entire natural contextual information can be fed into the training without considering the GPU memory limitation. They tried to save memory consumption by cropping 3D

patches with pixel size 96×96×96 from original scans, which may also add computation burden at the testing phase.

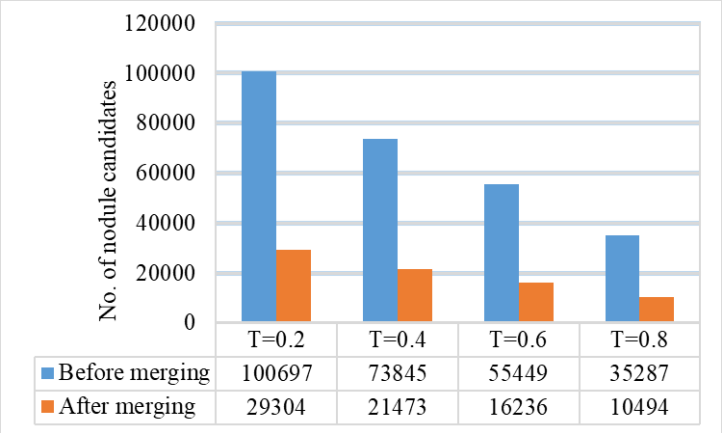

Figure 2.7: Number of detected nodule candidates before and after candidate merging process with different cut-off thresholds of prediction probability for all 888 LUNA16 scans.

Nodules that are successfully detected by candidate detection network but eliminated by FP reduction network are shown in Figure 2.9. About 4% of correctly detected nodules in our testing set are falsely removed. These nodules present either irregular shapes or ambiguous boundaries. Bringing more data targeting these nodule representatives might potentially boost the sensitivity performance.

2.3.4 Nodule segmentation

Not considering FPs, the mean and standard deviation of dice coefficients (DSC) regarding the nodule segmentation using FCN2s, FCN4s, and FCN8s are presented in Table 2.3, which quantitatively proves that FCN2s with lower level convolution will potentially yield higher resolution and preserve the boundary information of the nodule. Therefore, higher DSC is obtained with the comparison to groundtruth mask. This proves our assumption that using lower level convolutional layers can help improve the segmentation accuracy. The higher level

deconvolutional results provide the rough nodule locations and shape, while the lower level deconvolutional results provide the nodule boundary details.

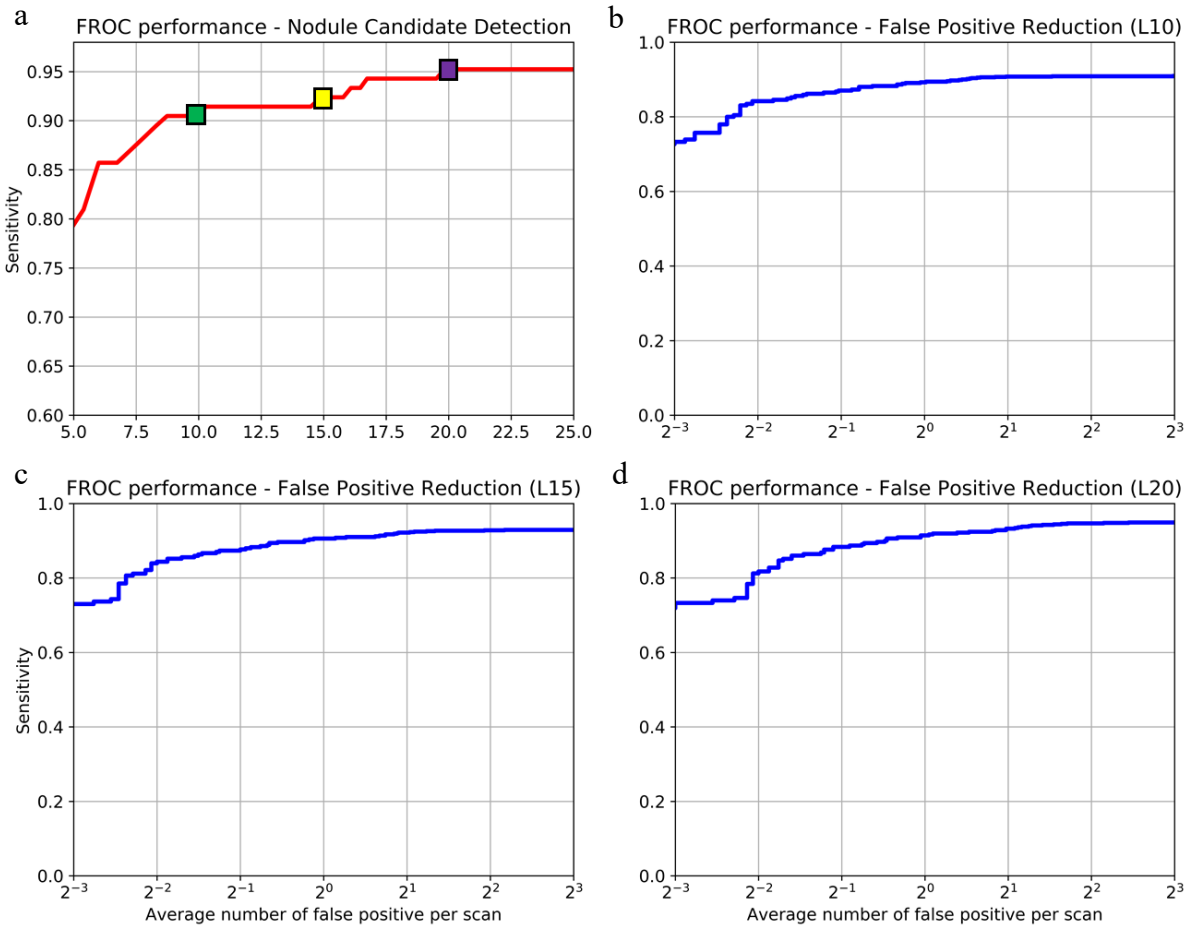

Figure 2.8: The FROC curve of our nodule detection results. (a) FROC curve of Faster R-CNN based nodule candidate detection without FP reduction. (b) FROC curve of FP reduction by taking initial candidate FP level of 10 in (a) as initial candidates marked by green square. (c) FROC curve of FP reduction by taking initial candidate FP level of 15 in (a) as initial candidates marked by blue square. (d) FROC curve of FP reduction by taking initial candidate FP level of 20 in (a) as initial candidates marked by purple square.

Table 2.4 shows the comparisons between our proposal segmentation result and other researchers' results based on the LIDC dataset. Since these methods use intersection over union (IoU) as the measurement metric, we report our results based on the same measurement in this table. Among all the listed approaches, our method achieves high performance on a relatively large amount of testing dataset.

Table 2.2: The system performance and CPM score comparison of the proposed method and other state-of-the-art approaches. Note that "online" means models with online descriptions available on LUNA16 competition website: https://luna16.grand-challenge.org/Results/, "*" represents models with limited details provided.

Team	GPU	CNN Dimension	CPM
PAtech (PA_tech)[online]	Tesla K80 × 4	3D	0.951
JianPeiCAD (weiyixie)[online]	NVIDIA Titan X × 2	3D	0.950
LUNA16FONOVACAD (zxp774747)[online]	NVIDIA M40 × 2	3D	0.947
iFLYTEK-MIG (yinbaocai)*	-	-	0.941
zhongliu_xie (zhongliu.xie)[online]	NVIDIA K80 × 2	3D	0.922
iDST-VC (chenjx1005)*	-	-	0.897
qfpxfd (qfpxfd)[Ding et al., 2017]	-	3D	0.891
CASED (CASED) [Jesson et al., 2017]	-	3D	0.887
3DCNN_NDET (lishaxue3)*	-	3D	0.882
Aidence (mjharte)*	-	-	0.871
junxuan20170516 (chenjx1005)*	-	3D	0.865
MEDICAI (bharadwaj)*	-	3D	0.862
Ethan20161221 (ethanhwang2012)*	-	-	0.856
resnet (QiDou)[Dou et al., 2017b]	NVIDIA Titan X × 1	3D	0.839
CCELargeCubeCnn (Intel_wuhui)[online]	CPU	3D	0.833
Ours	NVIDIA 980 Ti × 1	2D	**0.880**

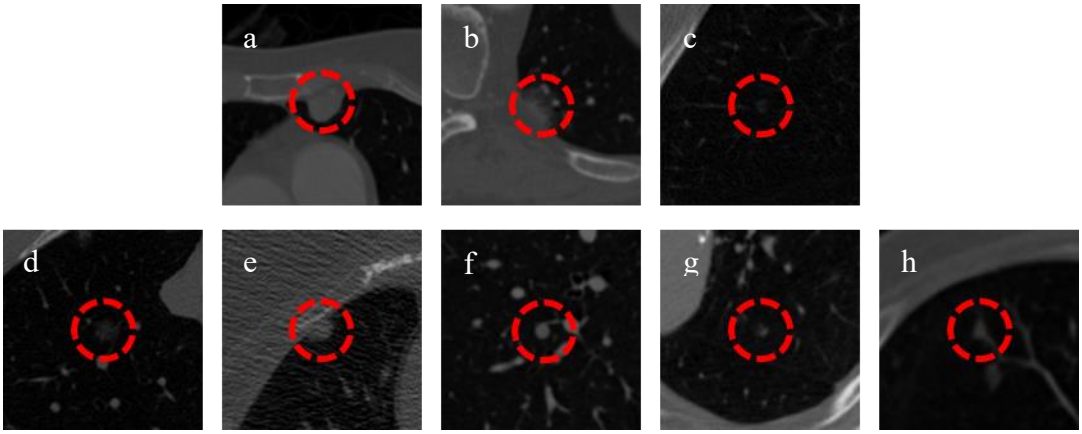

Figure 2.9: The FROC curve of our nodule detection results. (a) FROC curve of Faster R-CNN based nodule candidate detection without FP reduction. (b) FROC curve of FP reduction by taking initial candidate FP level of 10 in (a) as initial candidates marked by green square. (c) FROC curve of FP reduction by taking initial candidate FP level of 15 in (a) as initial candidates marked by blue square. (d) FROC curve of FP reduction by taking initial candidate FP level of 20 in (a) as initial candidates marked by purple square.

Table 2.3: The mean and standard deviation of segmentation dice coefficients among FCN2s, FCN4s, and FCN8s.

	Mean	Std
FCN2s	**0.793**	**0.082**
FCN4s	0.723	0.149
FCN8s	0.674	0.172

In clinical practice, nodules have various types of clinical characteristics. The capability of segmenting large variety of nodules is necessary. Nine nodule characteristics are given in LIDC groundtruth markings, reflecting nodules' calcification, malignancy level and so on. We chose four representative characteristics and separate our testing set into different groups according to the corresponding characteristic scores. Based on the quantitative results in Table 2.5, our segmentation algorithm possesses the robustness of processing various types of nodules with similar performance.

Table 2.4: Performance of the proposed method and other state-of-the-art approaches.

Methods	Nodule amount	IoU
Tachibana et al. 2006	23	50.7 ± 21.9%
Wang et al. 2009	64	58%
Messay et al. 2010	68	63 ± 16%
Kubota et al. 2011	23	69 ± 18%
	82	59 ± 19%
Tan et al. 2013	23	65%
Lassen et al. 2015	19	52 ± 7%
	40	50 ± 14%
Messay et al. 2015	66	71.70 ± 19.89%
	77	69.23 ± 13.82%
Wang et al. 2017	493	71.16 ± 12.22%
Ours	**223**	**70.24 ± 12.04%**

Visual comparisons between groundtruth, FCN2s, FCN4s, and FCN8s are shown in Figure 2.10 by displaying several representative nodules. Our nodule segmentation algorithm can precisely segment large variety of nodules with different anatomical characteristics. Due to the complicated boundary pattern, the DSC for poor-circumscribed (juxta-pleural and juxta-vascular) and fuzzy-boundary (ground glass opacity) nodules is a little worse than well-circumscribed nodules. But since the main components of such nodules are successfully segmented, the quantitative performance decay will not impact the robustness in terms of visualization.

Because we are conducting an end-to-end nodule segmentation directly from raw CT scans, the combined dice coefficient including FPs can reflect the overall performance. The false positives (FPs) can be involved in both detection and segmentation steps. For the segmentation algorithm, the FPs can occur when the algorithm falsely detects other structures (vessels, airway walls, etc.) while avoiding the true nodules, which brings the dice coefficient to zero on FPs (Type I). For the detection algorithm, the FPs are apparently added when the algorithm fails to separate FPs from detected candidates (Type II). Leveraging from our full CAD system (detection + segmentation), the Type I FPs are eliminated by detection network since the center coordinates far from detected candidates will be removed after the postprocessing step illustrated in Section 2.2.6. Therefore, only Type II FPs are remained. We can simply calculate the combined dice score by

considering all positive detected regions (TPs and FPs). For the given testing set, we choose sensitivity of 91.4% with 1 FPs per scan in the detection step and feed all the candidates in the segmentation network, the end-to-end dice score will be decreased to 0.426 (0.793 if we exclude the Type II FPs). Therefore, the combined dice coefficient provides the convention to be compared to other end-to-end nodule segmentation algorithms.

Table 2.5: Dice coefficients on different nodule groups based on clinical characteristics. Note that nodules in the testing set are grouped based on their clinical characteristic scores. The numbers in square brackets represent the number of nodules in the corresponding group. We average the characteristic scores from four radiologists.

Characteristics	Scores					
	1	2	3	4	5	6
Calcification	-	-	0.794	0.817	0.788	0.795
	-	-	[10]	[4]	[1]	[208]
Sphericity	0.784	0.740	0.823	0.781	0.798	-
	[1]	[19]	[72]	[76]	[55]	-
Spiculation	0.783	0.824	0.779	0.874	0.858	-
	[157]	[36]	[14]	[6]	[10]	-
Malignancy	0.737	0.780	0.807	0.812	0.800	-
	[14]	[62]	[71]	[43]	[33]	-

2.3.5 Nodule Malignancy Prediction

In this section, we performed 3D texture feature based classification and CNNs based classification separately to distinguish benign and malignant nodules from LIDC database. And then compare the performances.

For 3D texture feature based method, we acquired 172 nodules from LIDC dataset with confident malignancy likelihood ratings from four radiologists. We defined the confident ratings as at least three out of four radiologists give the same score to one nodule, and the distribution of the confident ratings are listed in Figure 2.11. The cases with ratings 1 or 2 are likely benign, the cases with ratings 4 or 5 are likely malignant, and the score 3 represent suspicious cases. To test

the possibility of using our proposed 3D features to differentiate the high and low likelihood of malignancy cases, we used each feature group to classify level 4 and level 5 cases. The results are shown in Table 2.6. Then we tested the performance of our system by combining all features together. The MDS dimension reduction method was applied to each feature group, and then we incorporated them to the feature pool. The reported accuracy is 0.85. When combining all the features together, all the experiments are based on the features after MDS procedure, because the total number of features far exceeds the total number of cases. To have a better understanding of the ability of these 3D features to distinguish nodules with different malignancy levels, we calculated and compared the AUCs using nodules with different malignancy levels and the relevant AUCs are shown in Table 2.7.

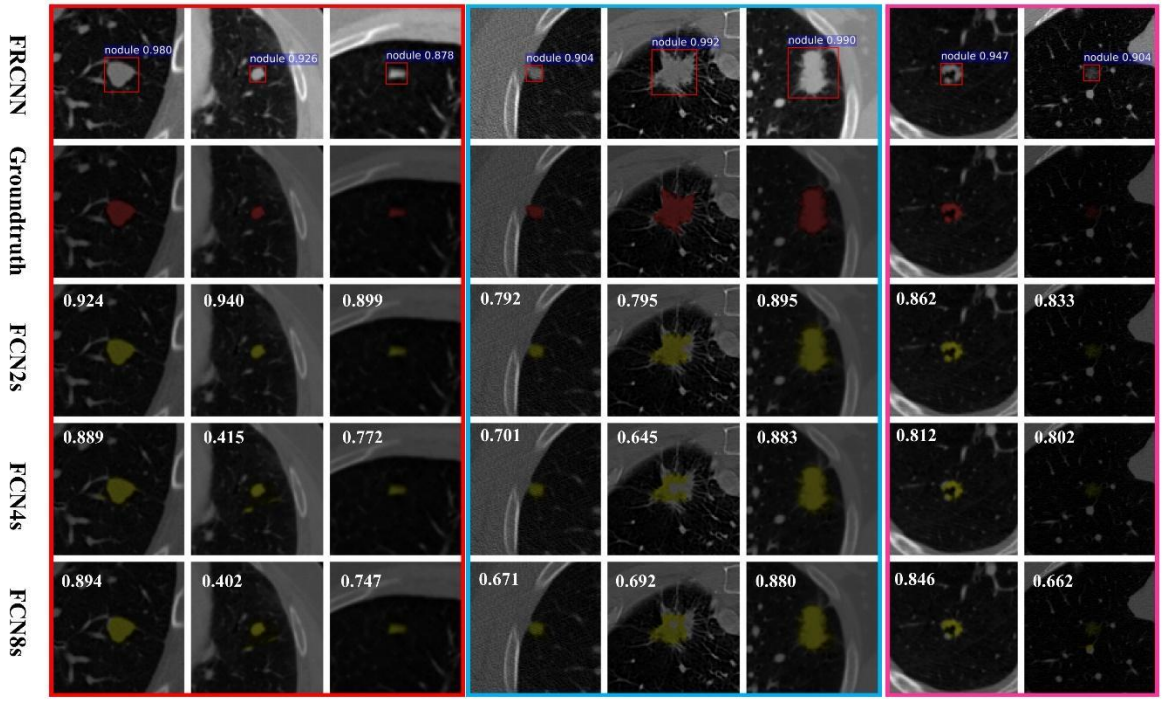

Figure 2.10: Visualization results of our proposed nodule segmentation system with various anatomical characteristics. The columns 1-3 marked by red rectangle represent three isolated nodules, the columns 4-6 marked by light blue rectangle denote one juxta-pleural (column 4) and two juxta-vascular (column 5 and 6) nodules, and the columns 7 and 8 marked by light purple rectangle show one subsolid nodule with center excavation and one ground glass opacity (GGO) nodule. The first row represents original nodule patches after Faster R-CNN detection and FP reduction with predicted bounding boxes marked by solid red squares as well as the classification probabilities in light blue background. The second row represents the corresponding annotations by radiologists. The manual segmentations are emphasized by red masks. The third to fifth rows denote nodule segmentation results by FCN2s, FCN4s, and FCN8s, respectively. The fully-automated segmentations are emphasized by yellow masks. The white decimals implicit the dice coefficients for each segmentation compared to groundtruth markings.

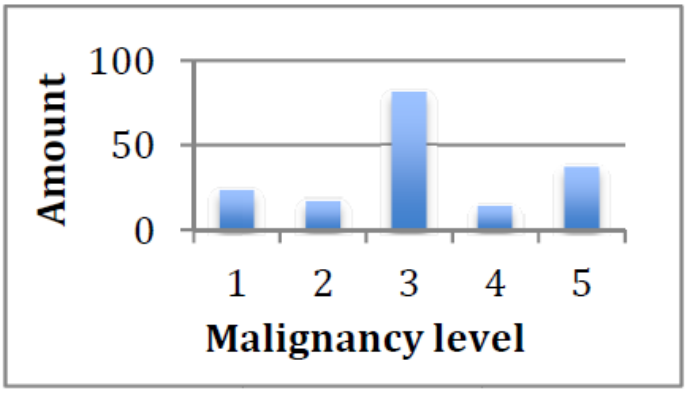

Figure 2.11: The malignancy level distribution of confident data.

Table 2.6: The accuracy and AUC value of differentiating level 4 and 5 cases using different feature groups

Feature group	Accuracy		AUC	
	Without MDS	With MDS	Without MDS	With MDS
3D GLCM	0.7510	0.8158	0.7010	0.7358
3D LBP	0.7950	0.8226	0.7150	0.7526
3D SIFT	0.7845	0.8126	0.7245	0.7426
3D Steerable	0.7703	0.8065	0.7003	0.7365
3D Wavelet	0.7438	0.7681	0.6938	0.7181
All	N/A	0.8365	N/A	0.7659

Table 2.7: Comparison of AUC on classifying different malignancy level group cases

	Level 2	Level 3	Level 4	Level 5
Level 1	0.6118	0.6518	0.9073	0.9827
Level 2		0.5599	0.8550	0.9641
Level 3			0.8047	0.9575
Level 4				0.7659

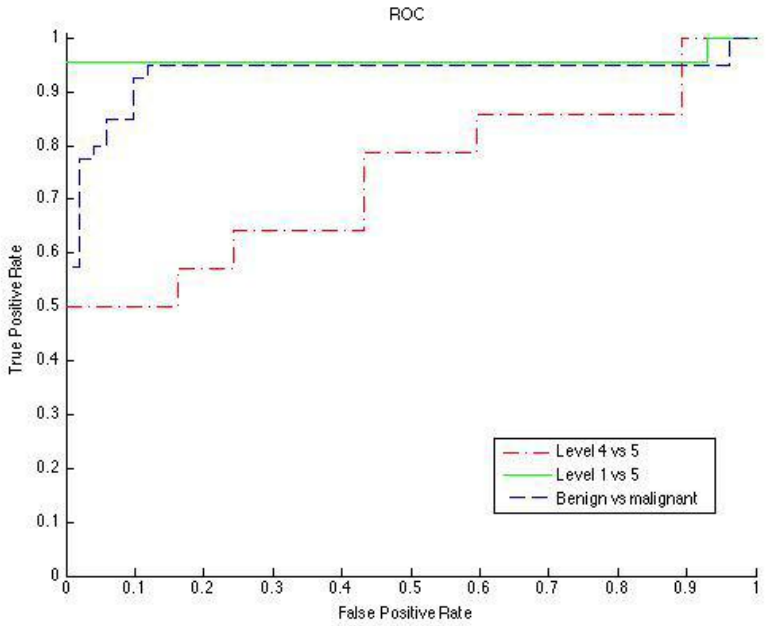

Figure 2.12: Three ROC curves: malignancy level 4 and level 5 cases; level 1 and level 5 cases; benign and malignant cases.

Figure 2.12 shows three ROC curves for classifying malignancy level 4 and level 5 cases, level 1 and level 5 cases, and benign (level 1 & 2) and malignant (level 4 & 5) cases. The accuracy and AUC for classifying level 1 and level 5 cases are 0.98 and 0.96; for benign and malignant cases are 0.94 and 0.89 respectively.

To measure and compare the performance of different types of ROIs, ROI based AUC and lesion based AUC were calculated (See Table 2.8). For the lesion based AUC, the predicted score of each nodule is the average score of all slices from one nodule. From Table 2.8, it can be noted that Multichannel ROI II generated the best performance on our proposed CNNs. Figure 2.13 shows ROC curves using our proposed CNNs for Original ROI, Nodule ROI, and Multichannel ROI I and II.

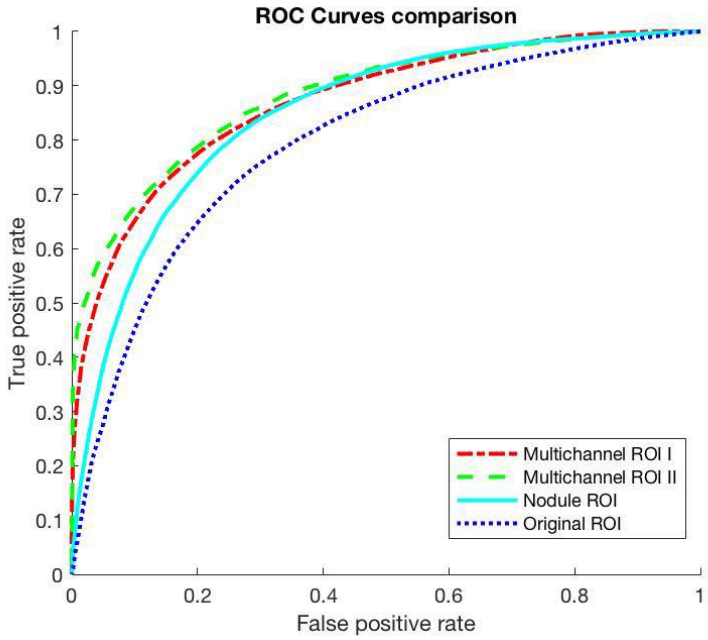

Figure 2.13: ROC curves of proposed CNNs using Original ROI, Nodule ROI, and Multichannel ROI I & II.

Table 2.8: The CNNs performance comparisons using different types of ROIs

	ROI based AUC	Lesion based AUC
(a) Original ROI	0.8484±0.0204	0.8381±0.0226
(b) Nodule ROI	0.8504±0.0153	0.8475±0.0173
(c) Surrounding ROI	0.8184±0.0178	0.8111±0.0148
(d) Gradient ROI	0.8211±0.0201	0.8167±0.0258
(e) Multichannel ROI I(a, b, c)	0.8705±0.0183	0.8662±0.0192
(f) Multichannel ROI II(a, b, d)	0.8823±0.0177	0.8793±0.0210

2.3.6 Execution Performance

As for execution perspective, we demonstrate the efficiency of the proposed method by running our automated nodule segmentation system for all 888 LUNA16 scans. Figure 2.14 shows statistical box plots regarding nodule detection (Faster R-CNN + CNN) and nodule segmentation (FCN2s). Note that the execution time is scan wised, so there are some variations due to the variations of slice numbers. The mean values and standard deviations are indicated in Table 2.9. Our end-to-end fully-automated nodule segmentation system achieves an average of 16 seconds per scan without any human interventions.

Table 2.9: The mean and standard deviation of execution time for our proposed nodule detection and segmentation algorithm.

	Mean (secs)	Std
Detection (Faster RCNN + CNN)	15.319	8.126
Segmentation (FCN2s)	0.917	0.835

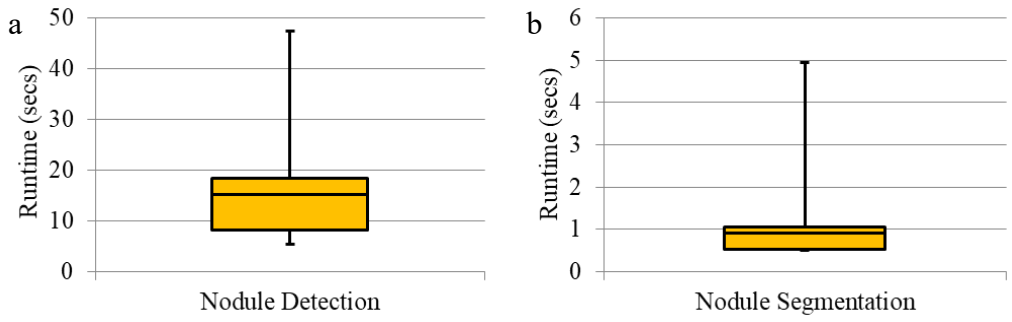

Figure 2.14: Box plot of execution time for our proposed (a) nodule detection algorithm and (b) nodule segmentation algorithm.

Chapter 3: Deep Matching for 2D Supperresolution Microscopy

Here, we developed a method named Deep Matching, a feed-forward CNN structure to improve the localization accuracy in super resolution fluorescence microscopic imaging. The idea is to use CNNs to perform denoising on the raw images after learning the statistics of the noise or signal, which are random variables (Zhang et al., 2017). CNNs can learn features of the signal or features of noise. Although learning features of signal is a common approach, residual learning that learns the noise features is proposed to overcome the training accuracy degradation problem where accuracy becomes saturated then degrades rapidly with increasing network depth (He et al., 2016). Further adding more convolution layers may even cause higher training error. Another advantage of residual learning for localization algorithms is that many current 3D localization microscopy techniques use altered PSF shapes to obtain 3D positions of particles (Li et al., 2018). By learning the noise our Deep Matching method can be applied to many types of PSFs used in 3D super resolution imaging experiments. Since this Deep Matching process effectively separates signal from noise, particle positions can be calculated with many current localization algorithms and compared for performance evaluation. This Deep Matching process is similar to a 2D matched filter that maps the detected noisy signal to an expected signal (template), which has a single convolution layer. It has been proven that this matched filter is the optimal filter that maximizes the SNR of the detected signal (Turin et al., 1960). However, the computing time for the matched filter approach is much more demanding than CNNs, and deep neural network based filters have achieved comparable performance with ideal matched filters as demonstrated in 1D gravitational wave signal processing (Gabbard et al., 2018; George et al., 2018).

3.0 Hypothesis

It is hypothesized that the end-to-end CNN network can massively extract signal from noisy PSFs so that the localization accuracy can be optimally improved. And the framework can be used in both simulated datasets and real experiments with very limited pre-knowledge.

3.1 Dataset

3.1.1 Particle Image Simulation for Theoretical Evaluation

In order to theoretically evaluate the localization algorithm performance and benefits of Deep Matching, we followed an established method with details in (Parthasarathy, 2012) to generate simulated particle images. Here, we only briefly describe this procedure. Ideally if a fluorophore can be accurately detected, namely, the particle ground truth position is inside the center pixel of the detected PSF image, the range of true particle position shifting is [-0.5, +0.5] pixels (±50 nm) in each dimension. Therefore, the particle ground truth positions in these simulated images are randomly shifted in the range of [-0.5, +0.5] pixels in each dimension for generality.

Single symmetric fluorescent particle images are modeled by convolving a point source at (x0, y0) with a symmetric point-spread function (PSF): $[\frac{2J_1(v)}{v}]^2$. J_1 represents the first kind of Bessel function with order 1 and $v = (2\pi \text{ NA } r)/(\lambda\, n_w)$, where NA = 1.3 denotes the numerical aperture of the objective lens, λ = 530 nm is the wavelength of light, n_w = 1.33 implicates the index of refraction of water, and r is the radial coordinate. Figure 3.5a shows an example of high resolution images of a single symmetric particle PSF with pixel size 1 nm centered at (x0, y0) indicated by the red circle, and Figure 3.5b is the intensity integrated image over Figure 3.5a, which shows a coarser pixelated image with pixel size of 100 nm and image size of 11×11 pixels. In order to generate noisy images at desired SNR level, we scale the pixel intensity such that the peak intensity equals to SNR2; then treat each pixel as an independent random variable with Poisson distribution; and at last replace each pixel intensity I by a random value generated from a Poisson distribution $e^{-\mu}\frac{\mu^k}{k!}$ with μ=I, and k as a non-negative integer. Also background intensity of 10 photons with Poisson random processes are added to each pixel. Figure 1c shows an example of simulated particle image at SNR = 5. By following such criteria, the total number of detected photons (the summation of all pixel intensities) is proportional to SNR^2 (Parthasarathy, 2012). We have also generated clean PSF images without background noise. In this case, each pixel intensity is created with the same Poisson random variable method mentioned above except that no background level of 10 photons is added. Note that in the real experiment, the background photon

number varies. However, benefiting from residual learning that only focuses on noise model, such variation remains insensitive.

Single asymmetric fluorescent particle images are generated in a similar way to single symmetric fluorescent particle images with expanding or contracting both x and y directions by a random factor ranging from 0.7 to 1.5. Then these noise-free images are added with Poisson noises at different SNR levels mentioned above.

For theoretical analysis, 18,000 noisy particle images with SNR ranging from 2 to 20 are created in both symmetric and asymmetric cases with 1,000 images at each SNR level. All testing images are blindly fed into CNN without any prior knowledge.

3.1.2 Particle Image Simulation for Experimental Data

In order to train a robust CNN model for practical use, it is critical to simulate particle images whose model PSF needs to be as realistic as possible. Therefore, instead of using a PSF based on theoretical mathematical models, we obtained the in-focus experimentally measured 2D PSF with maximal exposure time and averaging over several frames for high signal-to-noise ratio from Single Molecule Localization Microscopy Challenge 2016 (SMLM 2016) (Single-Molecule Localization Microscopy Software Benchmarking). After pixilation, the noise free particle image is achieved with image size of 11×11 pixels and pixel size of 100 nm. In order to generate noisy images at desired SNR level, we scale the pixel intensity such that the peak intensity equals to SNR^2. Based on such criteria, the total number of detected photons (the summation of all pixel intensities) is proportional to SNR^2 as shown in Figure 3.1. Following the Methods in Challenge 2016 Poisson shot and Gaussian noises are added by treating the input photon number of each pixel as an independent variable. A baseline of 10 photons are also added to each pixel. In reality the true particle position can locate in the neighboring pixels of the center pixel in an 11×11-pixels image. Therefore, in these simulated images the true particle position are randomly shifted in the range of [-1.5, +1.5] pixels (±150 nm) in each dimension.

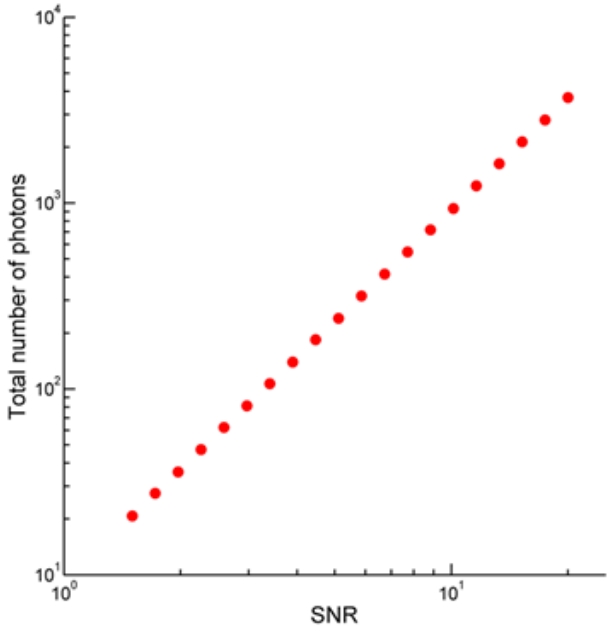

Figure 3.1: The total number of photons detected in the CCD plane as a function of the peak SNr for the simulated images with size of 11×11 in this paper. Note that the number of photons detected at the center brightest pixel is SNr^2 and the summation of photons numbers for all pixels contributes the total number of photons.

3.2 Methodology

Figure 3.2 shows the overall framework of our algorithm. The first step in a typical localization algorithm is particle detection, which identifies potential candidates for further localization calculation. Since this work is focused on localization accuracy improvement, ThunderSTORM (Ovesný et al., 2014), an open source software with top detection rate (high Jaccard index) in low SNR scenarios, is implemented for detection. Then a five-layer residual learning convolutional neural network is designed to learn the comprehensive features of the noise model and indirectly recover the signal from noisy images. Three sources of data are generated to feed into the CNN training phase: theoretical symmetric PSF, theoretical asymmetric, and experimental symmetric PSF images. Each type of training data is independently trained using the same CNN architecture. The corresponding testing data is created to evaluate the learning

performance. For localization algorithms, we applied radial symmetry (RS) (Parthasarathy, 2012), Gaussian non-linear least-squares fitting (GNLLS), and Gaussian maximum-likelihood estimation fitting (GMLE). The localization accuracy is evaluated using the software from the single-molecule localization microscopy software benchmarking Challenge 2016 (Single-Molecule Localization Microscopy Software Benchmarking).

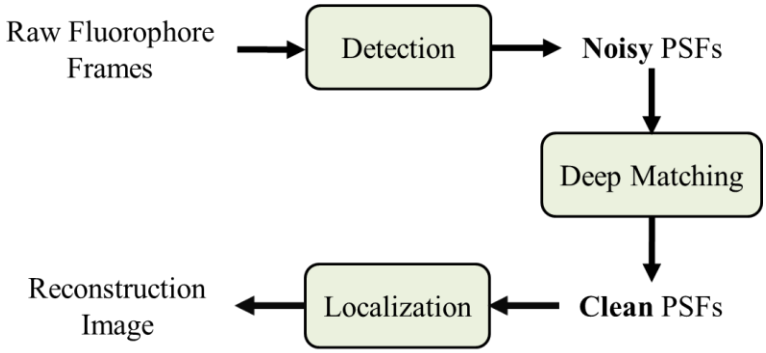

Figure 3.2: Overall structure of our proposed algorithm

3.2.1 Particle Detection

Practically, the detection is necessary to coarsely identify PSFs from original frames. In this paper, we adopt the detection algorithm from the well-developed software ThunderSTORM (Ovesný et al., 2014). First, the difference of Gaussian band-pass filter is involved to preprocess images and enhance image features. The two Gaussian kernels σ1 and σ2 are set as 1.46 and 6.0, respectively. Then we approximate the locations of particles by finding local maxima peaks with threshold of 0.95 times the standard deviation of the 1st wavelet level. Maximum likelihood Gaussian fitting with multi-emitter fitting analysis are calculated to identify the sub-pixel locations of fluorophores. The fitting radius is 3 pixels and initial σ for Gaussian fitting is 1.3 pixels.

3.2.2 Deep Matching

A noisy observation y can be decomposed as a clean image x with additive noise component n, which follows the relationship of $y = x + n$ [3]. The goal of Deep Matching is to recover x from y with minimal bias and error, which is equivalent to calculating a mapping

function F between y and x, i.e. F(y) = x. Rather than directly learning F, our CNN method learns a residual mapping of R(y) = n, and x can be predicted with y – R(y) (He et al., 2016). Implementing such a residual learning strategy makes CNN more robust, and it is effective to learn the features of the hidden noise model. It has been shown that deep CNNs with small-size convolutional kernels make decision function more discriminative and efficient than superficial CNNs with large-size convolutional kernels (Simonyan et al., 2015). Therefore, we set the size of convolutional kernels to be 3×3, and remove all pooling layers. The receptive field of CNN with depth d is (2d+1)×(2d+1) by using these 3×3 kernels. Because the size of input images is 11×11, the CNN depth is set at 5 to use all the context information. The architecture of the developed CNN is shown in Figure 3.3a. The input to the first layer is the observed image y, and 128 convolution kernels of size 3×3×1 are implemented to calculate 128 features, at last a rectified linear units (ReLU) activation function is used to truncate the negative neurons and add nonlinearity. For layers 2, 3, and 4, similar to the first layer, 128 convolution kernels of size 3×3×128 are used to generate hierarchical feature maps, followed by a ReLU activation function. For the last layer, one convolution kernel of size 3×3×128 is used to construct the output residual image R(y). In addition, zeros are padded into feature maps before convolution to ensure that each intermediate feature map as well as output image have the same size as the input image. Ideally, R(y) is expected to be similar to the ground truth n, and the predicted clean image $\hat{\boldsymbol{x}}$ is the subtraction between y and R(y). The error is measured by calculating the mean squared error (MSE) between R(y) and n, which can be adopted as the loss function L(Θ) to guide CNN to learn more discriminative features:

$$L(\Theta) = \frac{1}{N}\sum_{i=1}^{N} w_i |n_i - \mathrm{R}(y_i;\Theta)|^2 \tag{5}$$

where $\{n_i\}_{i=1}^{N}$ denotes N desired residual images, $w_i = \frac{I_i}{\max(I)}$ is the normalized weight that is proportional to the pixel intensity I_i, and Θ represents all trainable parameters in CNN. The feature calculation for layer k is expressed as an operation F_k:

$$F_k = \max(0, W_k * F_{k-1} + B_k) \tag{6}$$

where W_k represents the convolution kernel with size of 3×3×128 and, and B_k indicates a 128-dimensional bias vector, and * denotes the convolution operation between current convolution kernel and features map F_{k-1} calculated from previous layer, $\max(0, \boldsymbol{x})$ is the ReLU operation on the convolution response to prevent negative outputs.

3.2.3 Particle Localization

Localization algorithms such as RS, GNLLS, and GMLE are downloaded from a well-established package (Parthasarathy, 2012) . These algorithms are implemented on images with size of 7×7 pixels, centrally cropped from original or CNN processed images with size of 11×11 pixels. Peripheral pixels contain little information related to PSF. Hence, excluding peripheral pixels improves the localization accuracy for both original and CNN processed images.

We used an easy-to-use software FandPLimitTool to calculate CRB bounds (*FandPLimitTool*). The fundamental CRB is calculated by setting inputs as numerical aperture = 1.3, wavelength = 530 nm, photon detection rate = 2×10^6 photons/s. For pixelated noise-free CRB calculation, parameter are set as pixel size = 100×100 nanometers, ROI size = 7×7 pixels, and background level = 1×10^{-10} photons/pixel/s. For pixelated image with noise CRB calculation, same inputs are chosen as before with Poisson noise of background level = 10 photons/pixel.

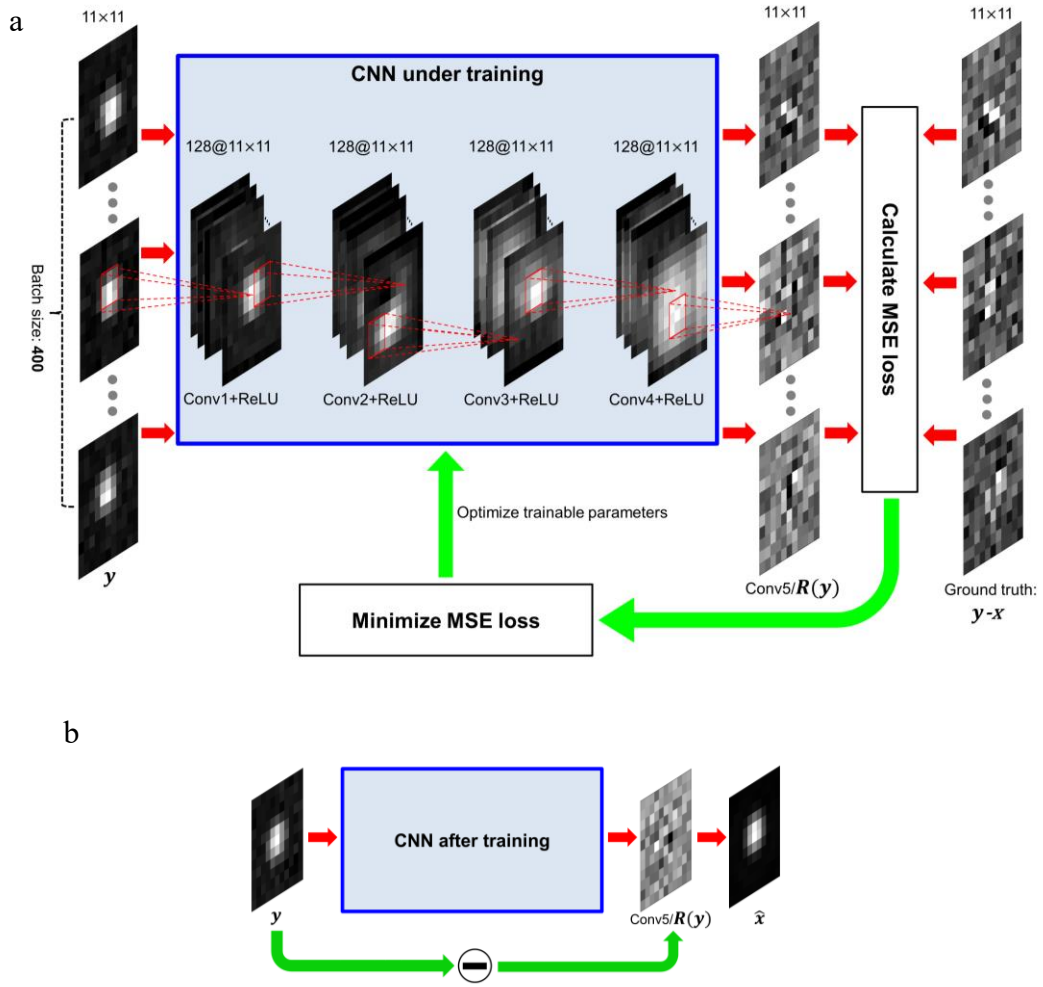

Figure 3.3: The architecture of developed convolutional neural network for localization algorithm. (a) The input image is $\boldsymbol{y}$. Five convolution layers are implemented. The output of the last convolution layer is R($\boldsymbol{y}$), the residual image of $\boldsymbol{y}$. This residual image is compared with the ground truth ($\boldsymbol{y}$-$\boldsymbol{x}$) to minimize mean squared error (MSE) for optimizing parameters in this CNN. (b) After training a test image is blindly fed into this CNN. The output R($\boldsymbol{y}$) is subtracted from y to obtain the final output $\hat{\boldsymbol{x}}$, an estimator of the ideal PSF.

3.3 Computational Results

3.3.1 Experimental Design

This end-to-end CNN model requires very few parameters setting, and training for both symmetric and asymmetric PSF images are using the same 5-layer CNN architecture. The initial convolution kernels are estimated using a Gaussian distribution of non-zero values and biases B are set as zero. The batch size, namely the number of images fed into CNN at a time, is set as 400. We apply Adam optimizer (Kingma et al., 2015) with initial learning rate 10-4, and the training takes about 500 epochs until the loss $L(\Theta)$ starts to fluctuate (Figure 3.4). The training phase typically takes less than 40 minutes for symmetric PSF images and 90 minutes for asymmetric images to achieve optimal convergence. The testing phase computation time is averagely less than 1.4 s for all 18,000 images (~78 μs each image).

3.3.2 CNN Training

There are two degrees of freedom in symmetric PSF images training: particle position shifting and SNR. For theoretical dataset, a total of 200,000 noisy PSF images with corresponding ground truth with 20 different SNR levels ranging from 1.5 (20 photons detected) to 20 (3700 photons detected) are generated. At each SNR level, 10,000 images with randomly shifted particle position (±50 nm) are generated to ensure the diversity. For practical dataset, since the searching area is 9 times larger than theoretical images, we therefore augment the training set by 10. At each SNR level, 100,000 images with randomly shifted particle position in the range of ±150 nm are generated to ensure the diversity. The CNN training data set is composed of 1,000,000 noisy images generated from the procedure mentioned above. And the corresponding noise-free particle images are stored for training accuracy evaluation. These images (size 11×11 pixels) have 20 different SNR levels ranging from 3.5 (110 photons detected) to 10 (920 photons detected). Using training images with larger size (11×11 pixels) improves CNN learning capability than smaller training images (7×7 pixels) (Figure 3.4). This can be explained as large images contain more information of noise which helps the training. The MSE loss of training images with size of 11×11

pixels (red curve) is much lower than that of training images with size of 7×7 pixels (blue curve) (Figure 3.4). Also training with larger images fastens the speed of convergence. In addition the loss of training images with smaller size starts to rise at around 500 epochs, while the loss of training images with larger size remains stable, which indicates the bottleneck of smaller size training images and further training will lead to overfitting problem. The minimal loss values are 13.8 and 1.9 for training images with size of 7×7 and 11×11 respectively.

The configuration of experimental desktop for training and testing purposes includes 8GB RAM, Intel E8600 processor, and a NVIDIA GTX 1080 Ti GPU.

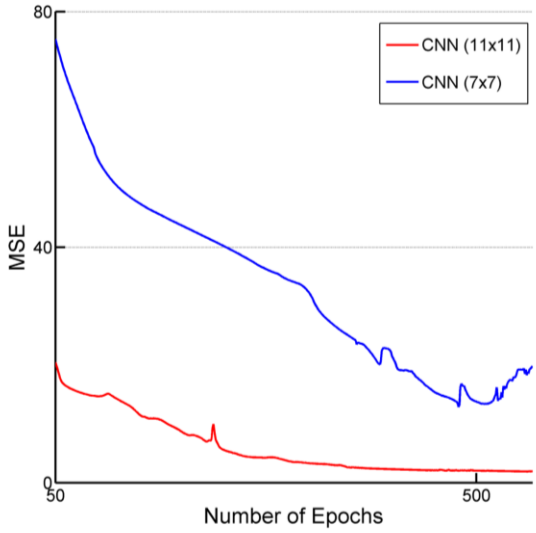

Figure 3.4: Training performance of CNN with larger images (11×11 pixels) and smaller images (7×7 pixels). MSE: mean squared error.

3.3.2 Results for Simulated Data with Symmetric PSFs

The performance of this Deep Matching method and the improvement of SNR on simulated data with symmetric PSFs are illustrated in Figure 3.5. One simulated single particle PSF image with true particle localization marked as a red circle (Figure 3.5a) is pixelated (Figure 3.5b), and then is added with Poisson noise at an SNR of 5, equivalent to 240 photons at the detector plane, as shown in Figure 3.5c. Employing a radial symmetry based particle localization algorithm, the

calculated particle position is marked as a green triangle (Figure 3.5c). After training, Figure 3.5c is blindly fed into this Deep Matching network, which predicts a residual image (Figure 3.5d). After subtracting this residual image from the input image, this CNN recovers an optimal restored image (Figure 3.5e). This restored image shows great qualitative similarity with the pixelated ideal PSF image without noise (Figure 3.5b). After applying the same radial symmetry based particle localization algorithm, the calculated particle position (blue cross) is closer to its ground truth value (red circle). The peak SNR (PSNR) value in optimized images, on average, improves by 10.29 dB in a total of 1000 experiments (Figure 3.5f). Similar improvement of PSNR is achieved at other SNR levels (Figure 3.6).

Quantitative analysis of the Deep Matching algorithm performance at SNR = 5 is shown in Figure 3.7a & 3.7b. In Figure 3.7a, the x coordinate shows the x components of true particle positions in 1000 test images with randomly shifting the x and y components between -0.5 and 0.5 pixels in both dimensions, and y coordinates are the errors of x components' position estimation of these 1000 images. The errors for y components are similar (Figure 3.7b). Before applying the CNN process the average particle localization errors using radial symmetry (RS), Gaussian fitting using non-linear least-squares (GNLLS), and Gaussian fitting using maximum-likelihood estimation (GMLE) are 0.179, 0.188, and 0.179 pixels, respectively (Figure 3.7c). The error is defined as the L2 distance between groundtruth and estimated centers as stated in (Parthasarathy, 2012). After applying the Deep Matching process, the average errors decrease to 0.163, 0.163, and 0.163 pixels respectively by using the aforementioned algorithms (Figure 3.7c). The mean errors of various particle localization algorithms at different levels of SNR ranging from 2 (37 photons count) to 20 (3700 photons count) are plotted in Figure 3.7c. The theoretical lower limits of algorithm error, Cramér-Rao bound (CRB), under different conditions are also plotted (Ober et al., 2004). CRB1 is the fundamental limit of localization accuracy for a quantum emitter, CRB2 is the limit for pixelated images without background noise, and CRB3 marks the practical limit for pixelated images with Poisson noise at different SNR levels.

There are three significant improvements after the Deep Matching process. First, the errors are reduced and closer to the CRB limit with the same particle localization algorithms (CNN+RS, CNN+GNLLS, and CNN+GMLE vs. RS, GNLLS, and GMLE). This improvement is significant especially in the low SNR range (<5). For instance, at SNR=2, the error is about 0.45 pixel with CNN algorithms, which is comparable to the errors achieved by conventional algorithms (RS, GNLLS, and GMLE) at SNR=3 (Figure 3.7c). Since the number of detected photons is proportional to SNR^2, this indicates that only 37 photons (SNR=2) are needed for CNN algorithms to achieve the same performance as the conventional algorithms which need 83 photons (SNR=3). This improvement has the potential to significantly reduce the data acquisition time in current super resolution imaging experiments. Second, the errors of all Deep Matching based algorithms (CNN+RS, CNN+GNLLS, and CNN+GMLE) are almost identical (Figure 3.7c). This implies that CNN processing significantly reduces the noise in images, which makes the later algorithms work under close to ideal conditions and leads to almost identical results. This property of Deep Matching almost removes the dependence of calculation errors on the algorithm being used. There is still a small performance gap between CNN+GMLE and CNN+GNLLS, because the MLE method is generally more precise than NLLS (Smith et al., 2010). Third, when the SNR level is very low (≤ 3), the errors of CNN algorithms overcome the practical limit set by CRB3. All CRBs are calculated based on the Fisher information matrix, which calculates the amount of information that an observed random variable y carries about an unknown parameter θ (Ober et al., 2004). The noise information is always present in CRB calculations, which limits our ability to predict the ground truth position from the information contained in signal. However, Deep Matching has efficiently separated the signal information from the noise information. Hence, this optimized signal information can determine the particle position with much better accuracy. A brief explanation based on mutual information and Fisher information theory is provided in the section of Appendix. However, in current real experiments when SNR is very low (≤ 3), it is impossible to detect such low signal emitters for further localization (Tang et al., 2015; Burgess et al., 1999).

Also in this theoretical calculation, it is assumed that the ground truth position is located within the center pixel. Such assumption cannot be guaranteed in a real experiment.

When the SNR is moderate (5-10) to high (>10), Deep Matching does not overcome the CRB3 limit. To further test their performance, pixelated particle images without background Poisson noise were simulated. This particle PSF is still affected by the Poisson process during fluorescence emission. The localization accuracy from conventional algorithms only approaches CRB2 in the low SNR range (<6), and is over CRB3 when SNR is moderate to high (>6) (Figure 3.7d). This CNN was retrained with these images, and retested. Since these images do not contain background noise, the CNN can learn the Poisson random processes contained in the PSF. The CNN algorithms localization accuracy approaches CRB2 (Figure 3.7e). When SNR is high (>10), the improvement when using Deep Matching is not significant. The possible reason is that the variance of Poisson process increases as its expectation value increases. Therefore, at high SNR levels the training and test errors in CNN take a bigger role in the overall algorithm error.

In real imaging experiments the presence of overlapped particles composed of two or more single PSFs is common. Apparently, this scenario will degrade the localization accuracy. In order to evaluate the performance of this CNN algorithm under such circumstances, we make a simple model by adding a second particle at a given center-to-center distance and random angular direction to the first one. The two particle PSFs are either manually or automatically segmented to two regions before applying CNN and localization algorithms. Each region contains only one particle PSF but with varying degree of influence from the other region. The total error as a function of the center-to-center separation distances of two particles at SNR = 5 is shown in Figure 3.8.

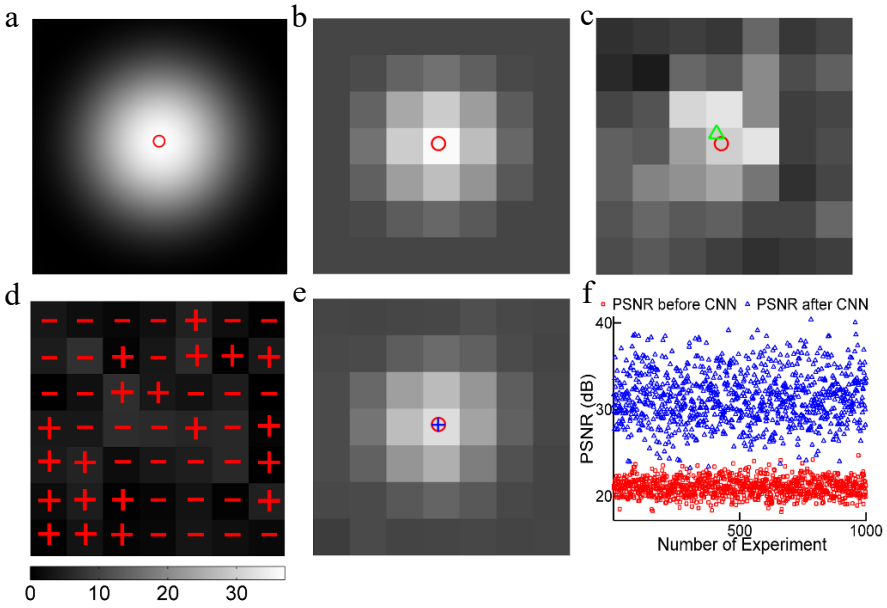

Figure 3.5: Illustration of the Deep Matching performance. (a) A simulated noise-free high resolution CCD image of a fluorophore, its true position is indicated by a red circle. (b) The pixelated result from the high resolution simulated image (a). (c) Shot noises at SNR=5 are added to (b). The green triangle marks the particle position calculated with a conventional radial symmetry based algorithm. (d) Residual image (noise) calculated by CNN. "+" indicates a positive value, "-" indicates a negative value. (e) Output optimized image after Deep Matching. The blue cross marks the particle position calculated with a conventional radial symmetry based algorithm. (f) PSNR values of noisy (c) and optimized (e) images. Red squares represent PSNR values of 1000 noisy PSF images; blue triangles represent values of corresponding clean PSF images after Deep Matching processing. The averaged PSNR improves by 10.29 dB.

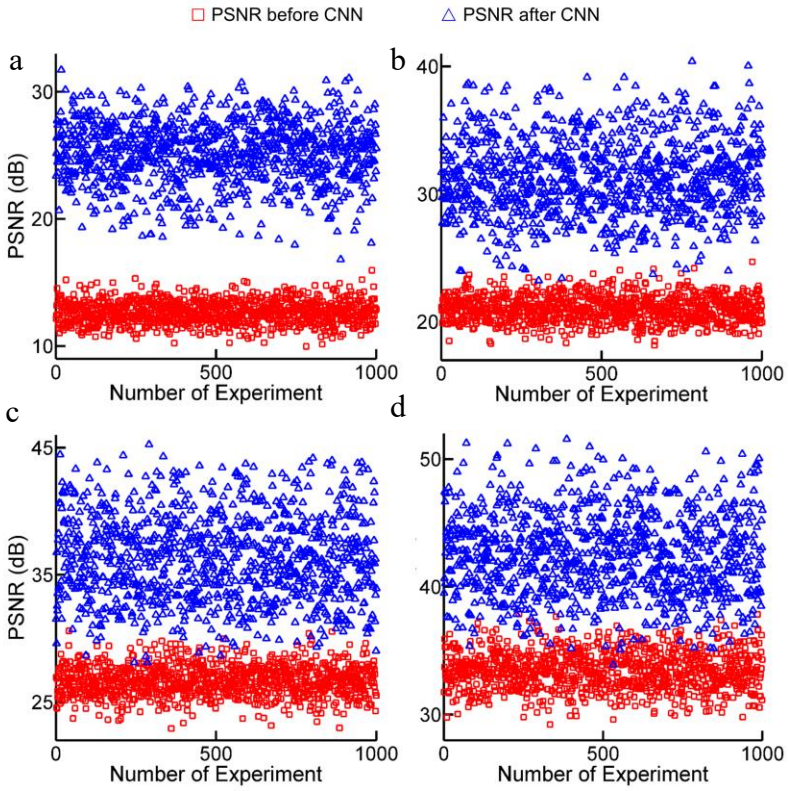

Figure 3.6: PSNR values of noisy and clean images at different SNR levels. Red squares represent PSNR values of 1000 noisy PSF images, and blue triangles represent the corresponding PSNR values of 1000 clean PSF images after Deep Matching from corresponding noisy images. (a) SNR=3, the averaged PSNR improves by 12.50 dB. (b) SNR=5, the averaged PSNR improves by 10.29 dB. (c) SNR=10, the averaged PSNR improves by 9.64 dB. (d) SNR=20, the averaged PSNR improves by 9.02 dB.

Generally, the error becomes smaller as the center-to-center distance between two simulated particles increases. For original noisy images radial symmetry based algorithm without prior assumption significantly outperforms Gaussian fitting based algorithms, when the particle separation is smaller than 4.5 pixels (about 85% wavelength). After implementing CNN algorithm, targeted single particle regions originally degraded by overlaps are optimally recovered similar to Figure 3.5c and 3.5e. And all three algorithms achieve better performances after this step. The

most significant improvement comes from Gaussian fitting based algorithms, where the CNN processing effectively removes the peripheral pixels that are heavily influenced by the other particle's PSF when the two particles center-to-center distance is smaller than 4.5 pixels. All three state-of-the-art algorithms output similar results. This CNN algorithm overcomes the performance bottleneck of classical Gaussian fitting algorithms while dealing with adjacent particles.

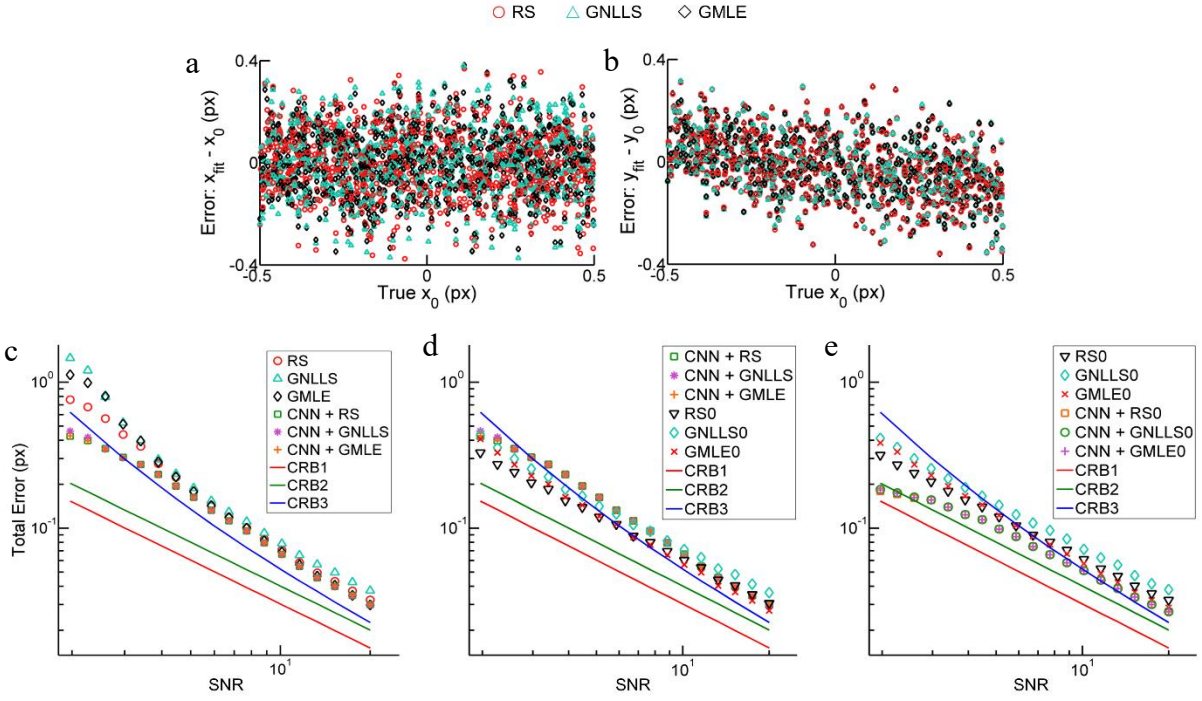

Figure 3.7: Comparison of several particle localization algorithms and Deep Matching. (a & b) Localization accuracy for various particle localization algorithms applied to simulated particle images at SNR = 5 after Deep Matching. The x (a) or y (b) components of the difference between the algorithm-determined value and true particle position value is plotted as a function of the particle position value. RS: radial symmetry; GNLLS: Gaussian fitting using non-linear least-squares minimization; GMLE: Gaussian fitting with maximum-likelihood estimation. (c) The algorithm error from simulated particle images over a range of SNR, from SNR = 2 (37 photons detected) to SNR = 20 (3700 photons detected). Each point denotes the average of 1,000 tests at that SNR level. Markers of red circle, cyan triangle, and black diamond represent localization errors on original simulated images through RS, GNLLS, and GMLE, respectively. Markers of green square, purple asterisk, and orange cross denote localization errors on corresponding CNN+RS, CNN+GNLLS, and CNN+GMLE respectively. The red solid line (CRB1) indicates the fundamental Cramér-Rao bound of localization accuracy, the green solid line (CRB2) indicates the CRB of pixelated images, the blue solid line (CRB3) indicates the practical CRB of pixelated images with noise. (d) The algorithm errors of RS (RS0), GNLLS (GNLLS0), and GMLE (GMLE0) on simulated particle images with zero background Poisson noise. (e) The algorithm errors of RS (CNN+RS0), GNLLS (CNN+GNLLS0), and GMLE (CNN+GMLE0) on simulated particle images with zero background Poisson noise after CNN processing.

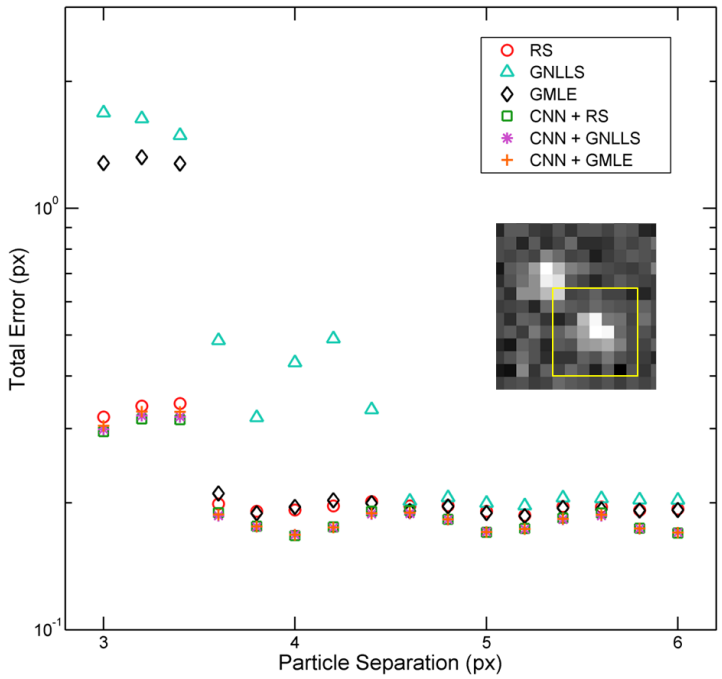

Figure 3.8: Localization algorithm accuracy in the presence of an adjacent particle. Test images were constructed at SNr = 5, consisting of two particles at a given center-to-center distance and a random angular direction. Each point represent the average particle localization error of 1000 tests. Inset: an example image with two particles separated by 4 pixels and oriented at 135 degrees with respect to x axis. The yellow box with size of 7×7 pixels indicates the segmented region for employing CNN localization algorithms.

3.3.2 Results for Simulated Data with Asymmetric PSFs

In 3D localization microscopy the PSF is engineered with different shapes to represent depth information. The first example is using intentionally introduced astigmatism to generate elliptical PSFs, and measuring long and short axes to determine the z position (Huang et al., 2008). Also in 2D imaging experiments the dipole orientation of molecules can induce asymmetric PSF shapes (Mortensen et al., 2010). Therefore, whether this Deep Matching method improves localization accuracy under such circumstances is important. The intuitive thinking is that the CNN's performance should not depend on the specific shape of the PSFs as long as it is trained

with correct training images. We examined the elliptical PSFs as a simple demonstration. In this test, the PSFs were expanded or contracted by a random factor ranging from 0.7 to 1.5 in both x and y directions respectively, meanwhile the true particle positions are randomly distributed over ±0.5 pixels in x and y directions. This CNN is further trained with these asymmetric PSF images. After the training phase, the CNN is blindly tested with randomly generated images.

The performance improvement for asymmetric PSF images after deep matching is similar to that from symmetric PSF images. Figure 3.9a shows the total errors of several algorithms for both original test images and clean images after Deep Matching at different levels of SNR (2-20). The performance improvement is similar to that from symmetric PSF images. The total errors for radial symmetry based and Gaussian fitting based algorithms are significantly reduced after the CNN processing at very low SNRs. Figure 3.9b and 3.9c show three examples of original images and clean images after CNN processing. Because the SNR in these images are significantly improved after CNN processing (similar to Figure 3.5f), the accuracy for determining the length of long and short axes will also be significantly improved. Hence, this Deep Matching process also helps to improve the depth (z) accuracy.

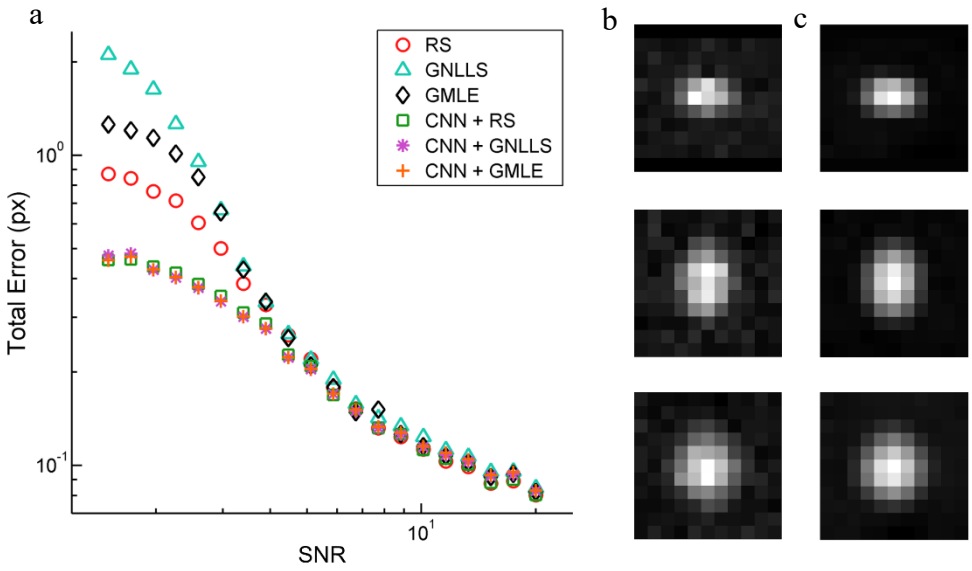

Figure 3.9: The accuracy of particle localization for simulated images of asymmetric PSFs. (a) The localization error from simulated particle images over a range of from SNR = 2 (37 photons detected) to SNR = 20 (3700 photons detected). Each point denotes the average of 1,000 tests at each SNR level. Markers of red circle, cyan triangle, and black diamond represent centroid localization error on original simulated images through Radial Symmetry, Gaussian fitting using nonlinear least-squares minimization and maximum-likelihood estimation, respectively. While markers of green square, purple asterisk, and red crosses denote localization error on corresponding clean images after CNN calculation using three aforementioned algorithms. The asymmetric PSFs are constructed by random scaling with a factor ranging from 0.7 to 1.5 in both x and y directions, respectively. (b) Three simulated asymmetric PSF images with random scaling in x and y axis at SNR = 10. (c) The corresponding clean images after CNN processing from (b).

3.3.3 Results for Real Experiment Data

The potential of using this Deep Matching process in real imaging experiments is demonstrated by applying it to analyze the online data from single-molecule localization microscopy software benchmarking Challenge 2016 (Parthasarathy, 2012). Since the ground truth positions are provided for these data, it is possible to quantitatively evaluate the performance of Deep Matching. The first step in a typical localization algorithm is particle detection, which identifies potential candidates for further localization calculation. Since this work is focused on localization accuracy improvement, ThunderSTORM (Ovesný et al., 2014), an open source software with top detection rate (high Jaccard index) in low SNR scenarios, is implemented for detection. Figure 3.10 shows the reconstructed localization results on a 2D low SNR training data with mixed noises such as Poisson noise and Gaussian noise (MT0.N2.LD.2D). The ground truth image is shown in Figure 3.10a. The ThunderSTORM reconstructed image with a radial symmetry algorithm is shown in Figure 3.10b. Out of total 22665 activations (after excluding 300 nm border fluorophores) 15149 activations were detected with Jaccard index of 65.8% (sensitivity 66.8%). The CNN was retrained with simulated PSFs with background noise and signal levels matching

this image sequence. After Deep Matching and radial symmetry localization, the reconstructed image is shown in Figure 3.10c. The enlarged views of the ground truth, ThunderSTORM result, and Deep Matching result are shown in Figure 3.10d-3.10f respectively. Usually after the detection step, a local image surrounding the candidate is segmented for localization calculation. The ground truth position may not fall in the center pixel of the segmented image, although typically the center pixel has the highest intensity value (Figure 3.10g). If the CNN is trained with ground truth position variations of ±1.5 pixels from the center of the image, Deep Matching recognized this case and the corrected image shows the upper neighbor of the center pixel has the highest intensity value (Figure 3.10h). The difference between localization errors for all 15149 fluorophores with Deep Matching and conventional methods is shown in Figure 3.10i. Overall 61.6% fluorophores achieved better localization accuracy. For these detected activations, the radial symmetry algorithm accuracy is 45.7 nm, after Deep Matching the radial symmetry accuracy improves to 40.9 nm (Table 3.1). GNLLS and GMLE achieve similar accuracy with 40.9 nm and 40.8 nm respectively. For data with a high density of fluorophores (MT0.N2.HD.2D), Deep Matching achieves similar improvement (Table 3.1). Out of total 28780 activations (after excluding 300 nm border fluorophores) 15543 activations were detected with Jaccard index of 52.9% (sensitivity 54.0%).

Table 3.1: Accuracy improvement (nm)

Algorithm		**Low Density**	**High Density**
Conventional	RS	45.7	59.1
	GNLLS	44.5	58.5
	GMLE	43.5	58.4
Deep Matching	RS	40.9	55.8
	GNLLS	40.9	55.9
	GMLE	40.8	55.9

The possible maximum improvement with Deep Matching is evaluated under ideal conditions. In this case, we assume all 22665 activations are segmented with the ground truth position falling in the center pixel of the cropped image. Deep Matching achieves a significant improvement in accuracy (Table 3.2). This demonstrates the great potential of further improving localization accuracy.

Table 3.2: Potential accuracy improvement (nm), known ground truth center pixel.

Algorithm		Low Density	High Density
Conventional	RS	58.0	69.0
	GNLLS	74.6	81.0
	GMLE	65.5	74.0
Deep Matching	RS	32.6	43.9
	GNLLS	32.6	43.8
	GMLE	32.5	43.8

We further applied our Deep Matching approach to real super resolution experiments. Figure 3.11a is the average intensity plot of a series of raw images of a fixed cell stained with mouse anti-alpha-tubulin primary antibody and Alexa647 secondary antibody with pixel size of 100 nm (TubulinAF647) (Parthasarathy, 2012). The Octane (Parthasarathy, 2012; Niu et al., 2008), ThunderSTORM, and Deep Matching reconstruction results are shown in Figure 3.11b-3.11d, and 3.11e-3.11g respectively. ThunderSTORM has a higher detection rate. Since currently Deep Matching uses the detection step in ThunderSTORM, the localization results are qualitatively similar for both methods as shown in Figure 3.11f and 3.11g.

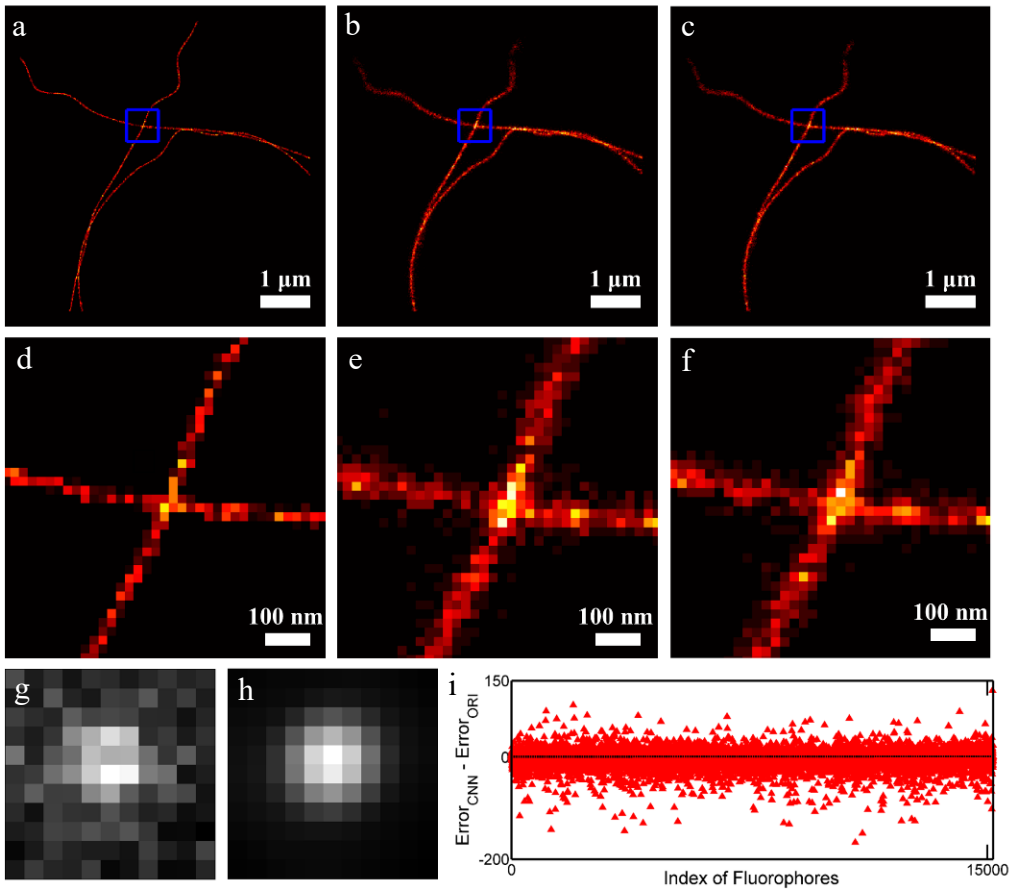

Figure 3.10: Localization algorithm performance on single-molecule localization microscopy software benchmarking data. (a) Ground truth image of the online data MT0.N2.LD.2D. (b) ThunderSTORM reconstructed image using radial symmetry method. (c) Reconstructed image after deep matching using ThunderSTORM for particle detection. (d)-(f) Enlarged view of the blue boxes in (a)-(c) respectively. (g) One example of a segmented 11×11 pixel image with the ground truth position actually in the pixel above the center pixel, although the center pixel has the highest intensity in the raw data. (h) Deep matching recovered image showing the pixel above the center pixel with the highest intensity. (i) The difference between localization errors in original noisy PSFs (ErrorORI) and CNN processed PSFs (ErrorCNN), unit: nm. 61.6 % fluorophores achieved better localization accuracy after Deep Matching.

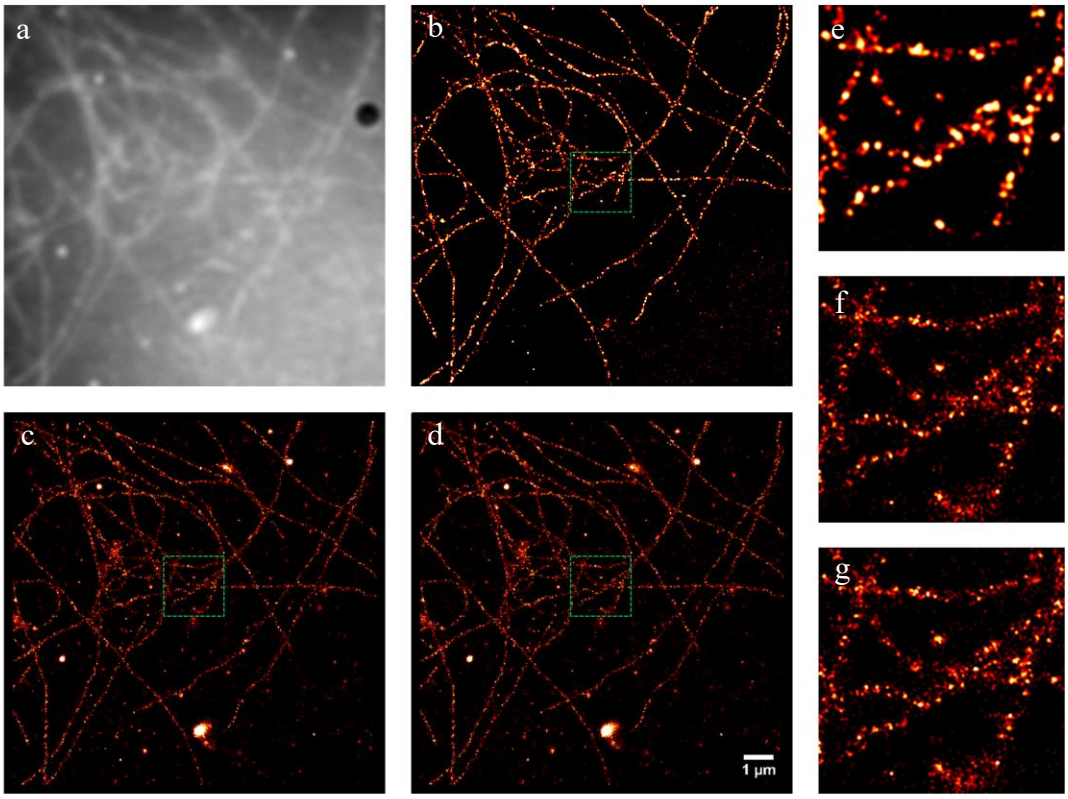

Figure 3.11: Localization results of STORM experiments. (a) Average intensity plot of a 9,990-frame video. Pixel size 100 nm. (b) Octane reconstructed super resolution image. Each pixel intensity level depicts the number of fluorophores located in 10 nm × 10 nm bins. (c) ThunderSTORM reconstruction result. (d) Deep Matching reconstruction result. (e)-(g) Enlarged views of the 2 µm × 2 µm green dash boxes in (b)-(d) respectively.

Chapter 4: Hybrid Scheme for 3D Supperresolution Microscopy

A practical consideration in developing algorithms is that diffraction rings can exhibit an elliptical shape due to optical aberration such as astigmatism. In this section, we developed a hybrid particle localization algorithm based on radial symmetry and ellipse fitting to localize the center of defocused PSF as well as both horizontal and vertical radii of the ellipse. The method can localize the 3D position of a fluorophore particle within 20 nm precision in three dimensions with a range of 40 μm in z dimension using a single defocused 2D image.

4.0 Hypothesis

It is hypothesized that the accurate ellipse fitting algorithm can optimally estimate the horizontal and vertical radius of ellipse rings from defocused PSFs. Given the relationship between radius and z depth, the 3D location of the particle can be calculated through a 2D defocused PSFs.

4.1 Dataset

The temporal focusing two-photon microscope setup is described in details in (Ding et al., 2016). The laser source is a femtosecond chirped pulse amplifier (Solstice ACE, 35 fs, 5 kHz, 6 W, Newport Corp., Mountain View, California, USA). Fluorescent images were acquired by an EMCCD camera (iXon Ultra 897, Andor Technology, Belfast, UK) working at -80 °C with TEC cooling system.

Two groups of images were acquired for particle localization experiments. In group #1, the z position of specimen (1 μm diameter, FluoSpheres carboxylate-modified, F8821, Thermo Fisher Scientific, Waltham, Massachusetts, USA) is scanned from 15.00 μm to 53.50 μm at 100 nm steps with an automatic nano-positioning sample stage, and the in-focus plane is at z = 0 μm. In group #2, the z position of specimen is scanned from 39.00 μm to 40.00 μm at 10 nm steps. At each z position, 100 images are acquired, and then the averaged intensity images are calculated. The signal to noise ratio (SNR) of raw images is defined as the ratio of the averaged intensity inside the ring P_s and the averaged intensity outside the ring P_n. For all images collected through this temporal focusing two-photon microscope, the integration time of EMCCD is set at 3 ms, and the

number of photons received on each pixel can be calculated by multiplying the pixel intensity level and photon count ratio. The pixel level is a 16-bit digital value recorded by the camera software and can be read out by image processing software such as ImageJ. The photon count ratio is a parameter that can be tuned in the camera software when recording video. This photon count ratio is set at 500 when taking the experimental images, and the maximum number of photons detected by one pixel in the ring PSF is about 25×10^6.

4.2 Methodology

The flowchart of our algorithm is shown in Figure 4.1. The approach is to first preprocess the raw image to retain the region of interest (outermost ring) and crop out other parts. Manual segmentation of large number of images is neither accurate nor practical. Two subroutines are developed for automatic preprocessing as shown in the first box (rough center localization) and the second box (rough radius estimation) of Figure 4.1. After the first two steps, ellipse fitting subroutine is executed to calculate the lengths of the long and short axes shown as in the third box.

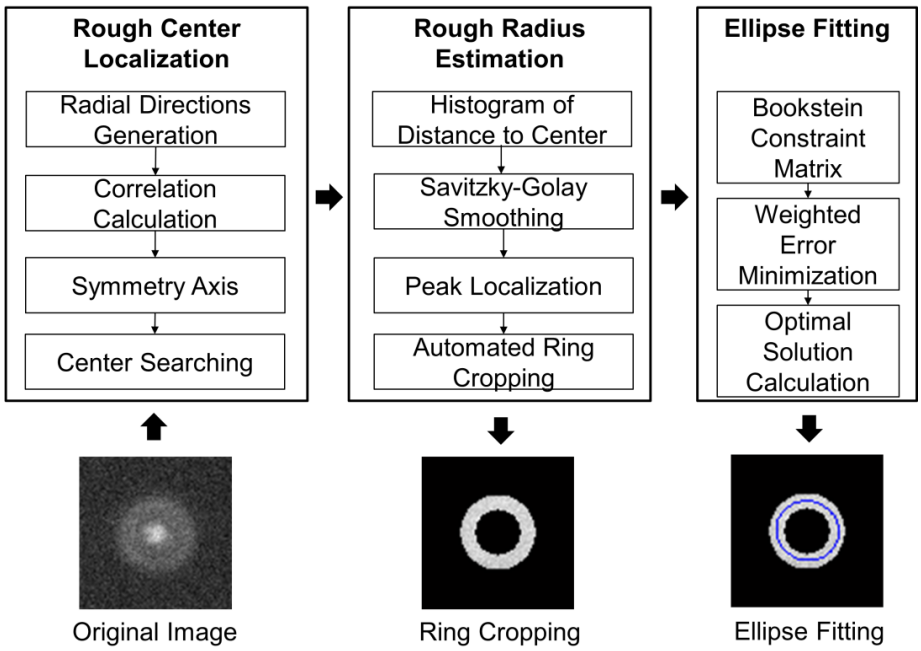

Figure 4.1: Flowchart of proposed particle localization algorithm

4.2.1 Rough Center Localization

To utilize the two-fold rotational symmetry in an ellipse, the first step in rough center estimation is to determine the symmetrical axis in four directions by calculating the correlation between the original image and a new image obtained from the original image after 180° rotation. For a given raw image I, four equally spaced radial directions with polar angles at 0°, 45°, 90°, 135° respectively are chosen. For each direction a series of parallel lines $l_{1,2,3,\dots,n}$ are tested as the centrosymmetric axis by calculating symmetric correlation coefficients $c_i(i = 1 \dots n)$ as:

$$c_i = \frac{\sum_{x=1}^{H}\sum_{y=1}^{W}(I(x,y)-\bar{I})\times(I_r(x,y)-\bar{I_r})}{\sqrt{[\sum_{x=1}^{H}\sum_{y=1}^{W}(I(x,y)-\bar{I})]^{2}\times[\sum_{x=1}^{H}\sum_{y=1}^{W}(I_r(x,y)-\bar{I_r})]^{2}}} \quad (7)$$

where I and I_r denote pixel intensities in the original image and the new image after 180° rotation, $\bar{I}$ is the average intensity, H and W are the height and width of images.

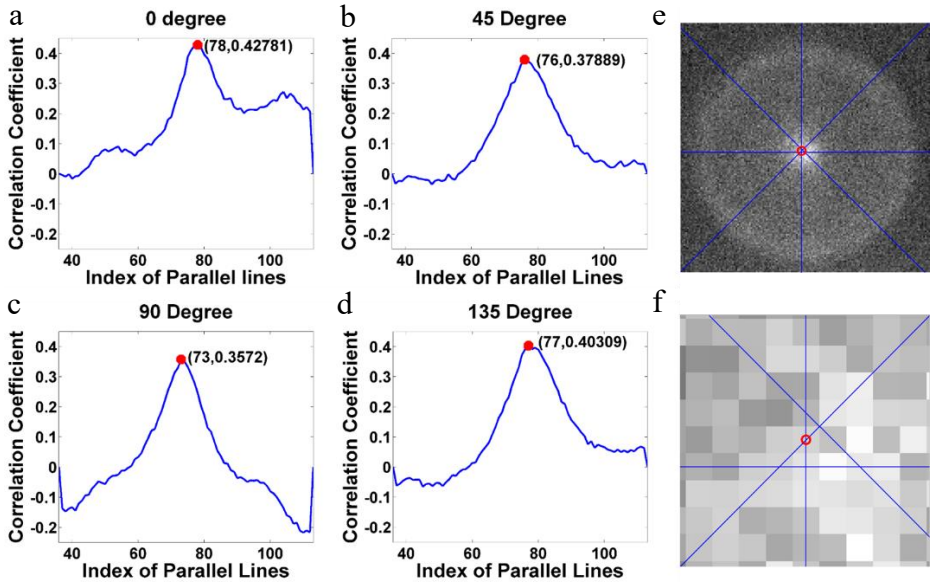

Figure 4.2: Rough center localization. (a)~(d) shows the distribution of correlation in four directions, 0, 45, 90, and 135 degrees, respectively. Horizontal axis is the index of parallel line. The red dot in each figure marks the peak of c_i, and the numbers in brackets show index i value and the maximal correlation coefficient. (e) The original image with four blue lines indicating the symmetrical axes in four directions. (f) The magnified view of (e). The red circle indicates the estimated center.

The variation of c_i versus i in all four directions are shown Figure 4.2a-d, respectively, and the line with maximum correlation coefficient is the symmetry axis in that direction (red points in Fig. 2(a)-(d)). The symmetry axis in each direction is represented as:

$$y = K_d x + C_d, \quad d = 0, 45, 90, 135 \tag{8}$$

Ideally these lines intersect exactly at the center. However, due to noise this does not happen as shown in the zoomed in image near center (Figure 4.2e and 4.2f). We determine the center based on least squares model, which searches for a pixel (x, y) that has the minimum sum of Euclidean distances D_{sum} to all four lines by the following equation:

$$D_{sum}(x, y) = \sum_d \frac{|K_d x - y + C_d|}{\sqrt{1 + K_d^2}} \tag{9}$$

Apparently, D_{sum} achieves minimum at (x_s, y_s) when the following two criteria are met simultaneously:

$$\begin{cases} \frac{\partial D_{sum}(x_s, y_s)}{\partial x_s} = 0 \\ \frac{\partial D_{sum}(x_s, y_s)}{\partial y_s} = 0 \end{cases} \tag{10}$$

Therefore, (x_s, y_s) is the rough center coordinates (red circles in Figure 4.2e and 4.2f).

4.2.2 Rough Radius Estimation

Based on the center found above, a histogram based radius estimation algorithm is applied. Figure 4.3a gives a schematic demonstration of calculating the distance-to-center histogram. For all the pixels that have equal Euclidean distance (D_0) to the center (x_s, y_s) marked with a red

cross, the average pixel intensity I_0 was calculated and plotted as a function of D_0 in Fig. 4.3b. If the noise level is low, this histogram curve should have a peak as the average pixel intensity reaches its maximum value when the Euclidean distance to center equals the radius. However, high-frequency noises appear in this histogram and this peak is not identifiable. In order to remove these noises without greatly distorting the signal, Savitzky-Golay (SG) filter is applied to smooth the histogram by a 597-point quadratic polynomial (Savitzky et al., 1964). Compared to other methods such as derivative calculation and wavelet transform (WT), SG filter performs a fast lest-square-fit convolution procedure by setting only two parameters: the width of the smoothing window and the degree of the smoothing polynomial, which significantly saves computation costs with fewer manual parameter settings (Li et al., 2015). Rough radius R_0 is calculated by finding the local maximum of the smoothed curve (red dot in Figure 4.3b).

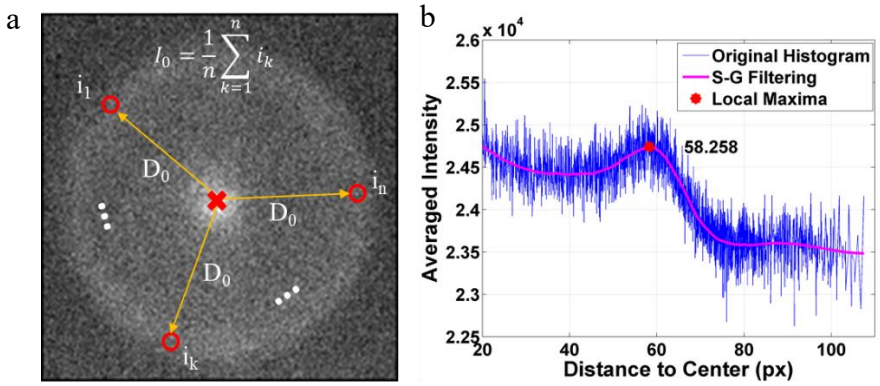

Figure 4.3: Rough radius estimation. (a) Schematic graph interprets the histogram. Given a distance (D_0) to center (x_s, y_s), the corresponding averaged intensity is calculated by the average intensities of all pixels whose Euclidean distances to center equal to D_0. (b) Blue and magenta curves represent the original histogram and S-G filtering result. The horizontal coordinate of red solid dot indicates the radius.

4.2.3 Ellipse Fitting

With the estimated center (x_s, y_s) and radius R_0 known, most of the background region can be excluded from calculation except the region of interest (the ring) which was determined by Equation 11:

$$(x - x_s)^2 + (y - y_s)^2 \in [(R_0 - b)^2, (R_0 + b)^2] \quad (11)$$

The parameter b is the half width of the ring, and is empirically set as 5 pixels in this experiment. The equation of an ellipse is represented by an implicit second order polynomial:

$$E(P; X) = P^T \cdot X = ax^2 + bxy + cy^2 + dx + ey + f = 0 \quad (12)$$

where $P = [a, b, c, d, e, f]^T$ and $X = [x^2, xy, y^2, x, y, 1]^T$. In a 2D image each pixel coordinate is represented as (x_k, y_k), and the pixel intensity is I_k (k=1…N, N is the total number of data points). $E(P; X_k)$ represents the algebraic distance between point X_k and the ellipse $E(P; X) = 0$. According to least-square criterion, the optimal ellipse fitting can be achieved by minimizing the sum of squared algebraic distance of all N data points to the ellipse (Equation 13), and the optimal coefficient vector P_0 is achieved when such error is minimal as shown in Equation 14.

$$D_a(P) = \sum_{k=1}^{N} E(P; X_k)^2 \quad (13)$$

$$P_0 = \underset{P}{argmin}\, D_a(P) \quad (14)$$

After the background is cropped out, there are still noises inside the rings. Intuitively pixels with higher intensity are more likely to be the real signal and pixels with lower intensity tend to be noise. Therefore, a weight factor w_k is implemented to enhance signal and suppress noise:

$$w_k = norm({I_k}^c) \quad (15)$$

where c is a positive parameter to control the weight among pixels with low (noise) and high (signal) intensities. We empirically set c=10. Now the squared algebraic distance with weight factor from the ellipse to N data points is represented as:

$$D_a(P) = \sum_{k=1}^{N} w_k * E(P; X_k)^2 \quad (16)$$

While finding the polynomial coefficient vector P by minimizing $D_a(P)$, some constraint is needed in order to avoid the trivial solution $P = \mathbf{0_6}$ and several possible solutions representing the same ellipse. In general, constraints can be expressed in the matrix form of $P^T C P = 1$, where

$\mathbf{C}$ is a 6×6 constraint matrix. Here, Bookstein constraint, one type of Euclidean-invariant constraints, is implemented by the following matrix equation (Kesaniemi et al., 2017):

$$\mathrm{P}^{\mathrm{T}}\begin{bmatrix} 1 & 0 & 0 & \\ 0 & 1/2 & 0 & \mathbf{0_{3\times3}} \\ 0 & 0 & 1 & \\ & \mathbf{0_{3\times3}} & & \mathbf{0_{3\times3}} \end{bmatrix}\mathrm{P} = 1 \tag{17}$$

Now the minimization of Equation 16 can be solved by considering rank deficient generalized eigenvalue system with a Lagrange multiplier λ:

$$\mathrm{S}^{\mathrm{T}}\mathrm{SP} = \lambda\mathrm{CP} \tag{18}$$

where $\mathrm{S} = [\mathrm{X}_1, \mathrm{X}_2, \ldots, \mathrm{X}_\mathrm{N}]^{\mathrm{T}}$. P is solved by calculating generalized eigenvectors of Equation 18 under the constraint of Equation 17 (Fitzgibbon et al., 1999). Fig. 4.4 shows an example of the fitting with P = [0.0036, 0.0004, 0.0035, -0.4796, 0.4329, 30.3565]$^{\mathrm{T}}$.

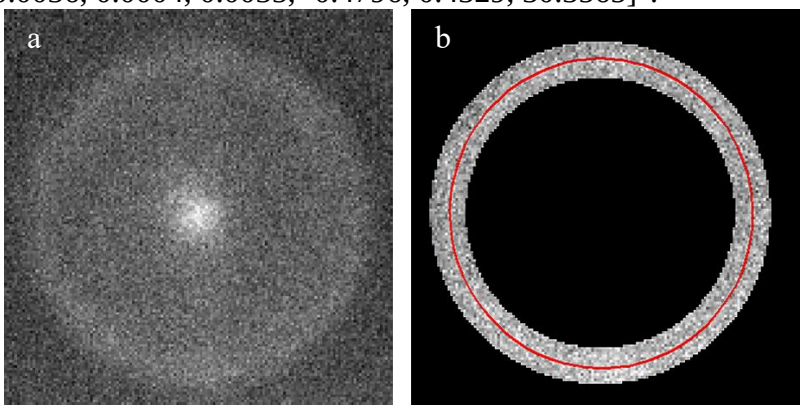

Figure 4.4: Weighted least square ellipse fitting. (a) Original image. (b) Ellipse fitting based on ring cropping of (a). The red solid curve in (b) is the optimal solution.

4.3 Computational Results

The performance of this algorithm is first tested on simulated noisy images, and the algorithmic error is evaluated with a method developed in Ref. 6. Simulated rings with intensity obeying Gaussian distribution are considered as ideal noise-free images ($\mathrm{I}_{\mathrm{ideal}}$), the long and short axes are L and S with center at $(\mathrm{x}_\mathrm{s}, \mathrm{y}_\mathrm{s})$. Simulated noisy images ($\mathrm{I}_{\mathrm{noisy}}$) with artificially added Poisson noises to $\mathrm{I}_{\mathrm{ideal}}$ spanning a range of SNR are used to evaluate this algorithm. Figure 4.5 shows three simulated noisy images with SNR at 1.05, 1.5, and 2 respectively. The algorithmic

error is defined as the difference between algorithm outputs from I_{noisy} and real ring parameters from I_{ideal}. For purpose of brevity, only the algorithmic error of horizontal axis (ΔR) and center coordinate ($\Delta x_s, \Delta y_s$) are displayed in figures.

In the first test, different levels of Poisson noises ranging from SNR = 1.01 to 2 at interval of 0.1 are added. At each SNR level 1000 test images are generated. The mean of absolute value of algorithmic error (red circle dots) and standard deviation (blue plus signs) of algorithmic error at different SNR levels are plotted in Figure 4.6a. Furthermore, since the SNR of experimental raw images is around 1.05, we calculate the mean of absolute value of algorithmic error (red circle dots) and standard deviation (blue plus signs) of error in a fine range from SNR = 1.01 to 1.1 with interval 0.01 (Figure 4.6b). The data of such 1000 images at SNR level of 1.05 is shown in Figure 4.6c. The errors, ΔR, Δx_s, and Δy_s in these 1000 experiments are randomly distributed with standard deviation of 9.69 nm, 7.62 nm, and 7.51 nm with corresponding parameters in I_{ideal} as 0. These tests show that under the current image SNR around 1.05, we have approximately 95% confidence to expect the algorithmic error for R, x_s, and y_s to be around 20 nm, 15 nm, and 15 nm respectively.

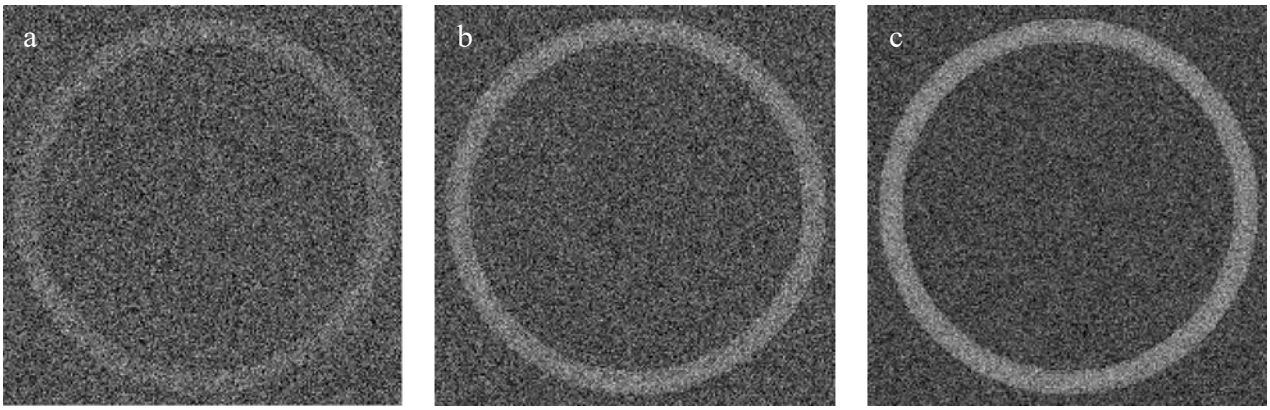

Figure 4.5: Simulated noisy images at different levels of SNR. (a) SNR = 1.05. (b) SNR = 1.5. (c) SNR = 2.

This algorithm is further applied on real microscopic images. Figure 4.7a-c show raw experimental images at depth 20 μm, 35 μm, and 50 μm, respectively and Figure 4.7d and 4.7e show the distribution of SNR for all raw images in group #1 and group #2, respectively. The mean of SNR is 1.045 with standard deviation 0.032 in group #1 and the mean of SNR is 1.051 with standard deviation 0.008 in group #2. The elliptical long and short radii for both two experimental groups are displayed in Figure 4.8a and 4.8b respectively. In a long range the relationship of radius vs. z is not linear, however, in a short range it shows better linearity. The ratio of long axis over short axis is about 1.02, indicating the aspect ratio of ellipse. Apparently the fluctuations of the calculated radii of rings are much larger than the means and standard deviations in test images (Figure 4.6). The most possible reason is the mechanical instability of the microscope, since it is not equipped with feed-back control.

To estimate the mechanical instability, the same specimen was kept stationary and imaged for 20 minutes with a total of 20 images taken at 1 minute interval. The fluctuation of calculated horizontal radius over time after subtracting average is shown in Figure 4.9. Since at SNR=1.05 the algorithmic error for R, x_s, and y_s are around 20 nm, 15 nm, and 15 nm (Figure 4.6c), this shows that the microscope has a fluctuation for R, x_s, and y_s in round 50 nm. The mean and standard deviation of Figure 4.6c and Figure 4.9 are summarized in Table 4.1. These tests demonstrate that mechanical instability is the limiting factor in demining particle position precision in our current experiments.

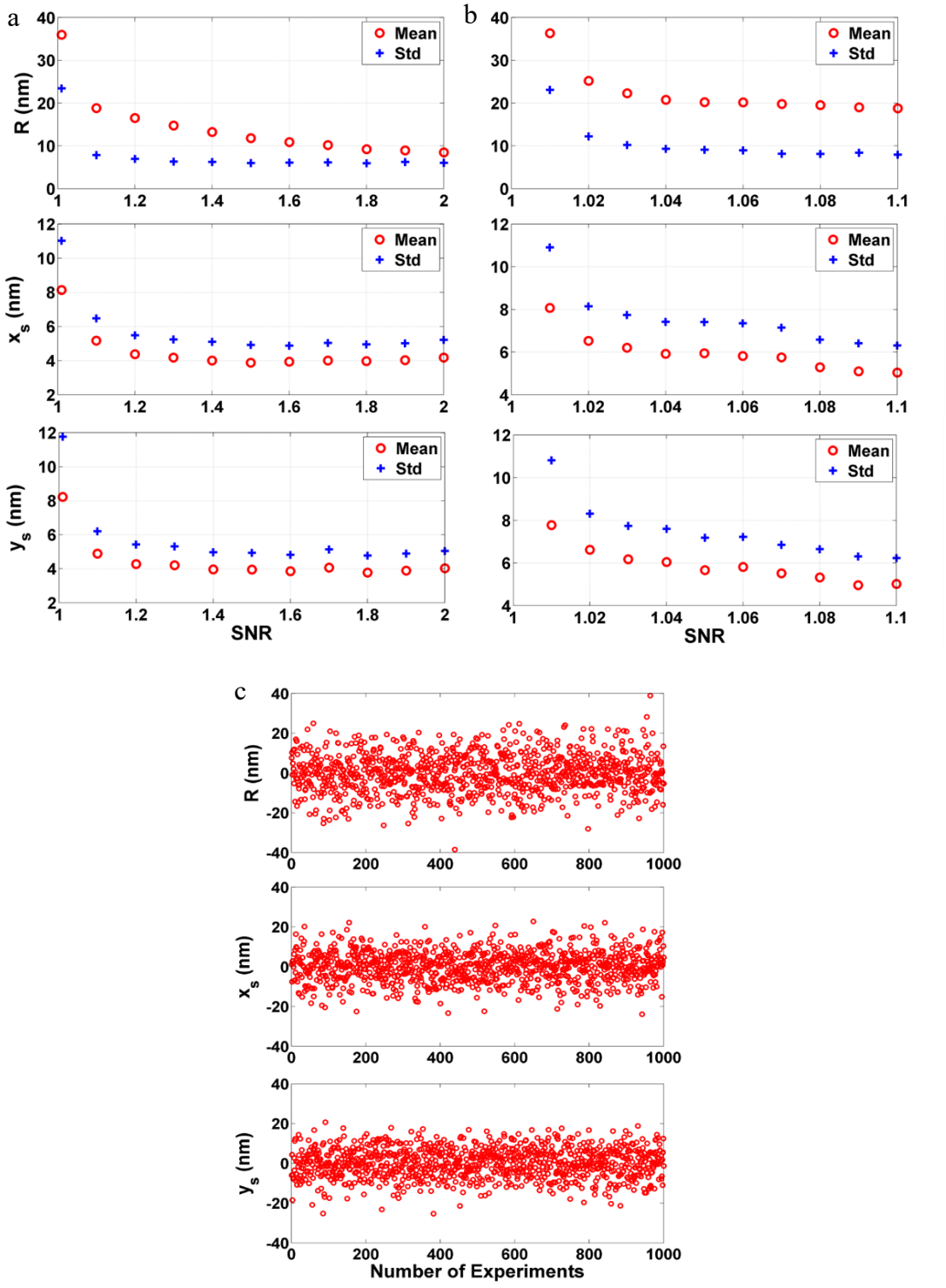
a
b
c
Mean
Std
R (nm)
x_s (nm)
y_s (nm)
SNR
Number of Experiments

Figure 4.6: Evaluation of algorithmic error. (a) Mean and standard deviation of ΔR, Δx_s, and Δy_s in large SNR range (1.01 to 2 at 0.1 interval). (b) Mean and standard deviation of ΔR, Δx_s, and Δy_s in short SNR range (1.01 to 1.1 at 0.01 interval). Note that at each SNR level in (a) and (b), 1000 test images were conducted. (c) 1000 test results of ΔR, Δx_s, and Δy_s at SNR level of 1.05. The mean SNR of real imaging experimental raw data is 1.05.

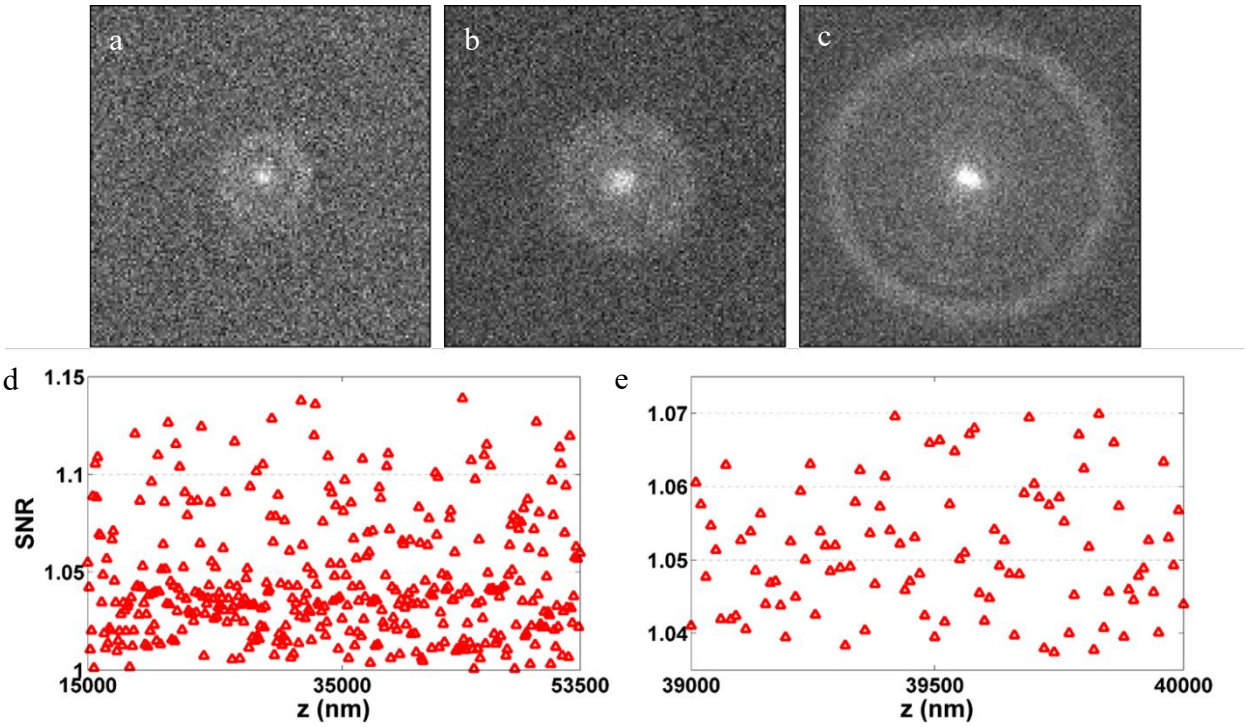

Figure 4.7: Data analysis of experimental raw images

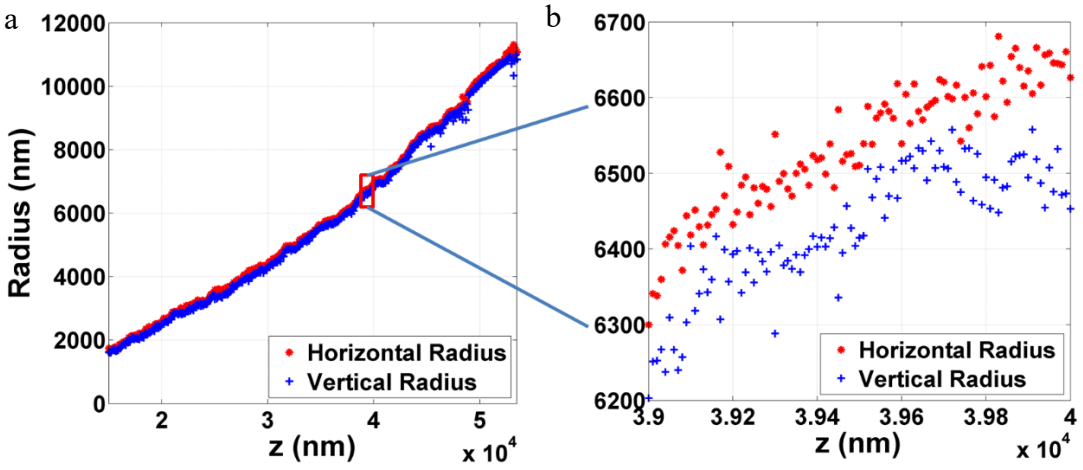

Figure 4.8: Radius vs. depth for (a) group #1 (large range), and (b) group #2 (small range). Red: long axis, blue: short axis.

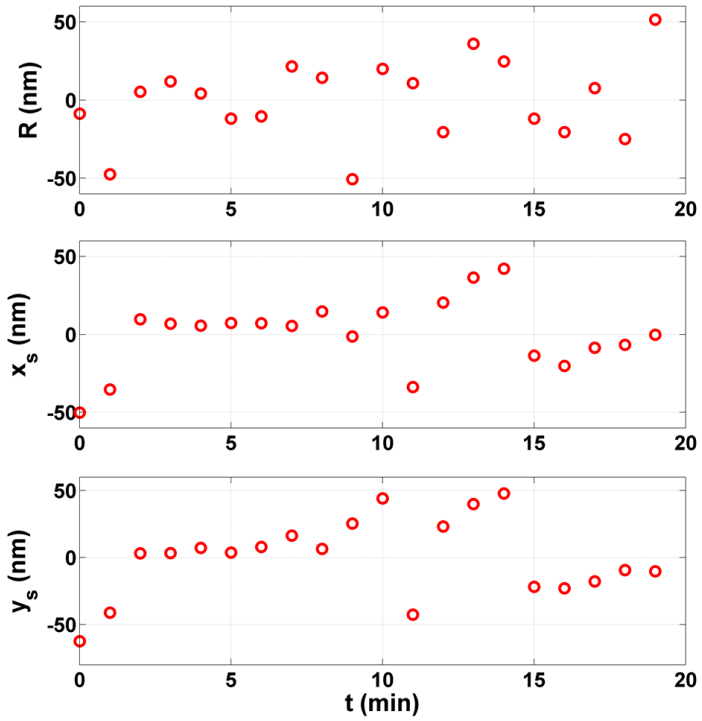

Figure 4.9: Mechanical stability measurement for R, x_s, and y_s.

All calculations were run on a Lenovo^{TM} ideapad 320 laptop computer with an AMD A12-9720P processor. The average execution time of our proposed localization algorithm for a single image is shorter than 1.3 s. Please refer to Table 4.2 for more details.

Table 4.1: Mean and standard deviation of ΔR, Δx_s, and Δy_s for simulated images (Figure 4.6c) and raw experimental images (Figure 4.8)

	Mean (nm)			**Std** (nm)		
	ΔR	Δx_s	Δy_s	ΔR	Δx_s	Δy_s
Simulated Images	-0.7	0.2	0.2	9.7	7.6	7.5
Experimental Images	/	/	/	25.9	22.8	29.3

Table 4.2: Mean and standard deviation of computation time for Rough Center Localization (S1), Rough Radius Estimation (S2), and Ellipse Fitting (S3)

	Mean (ms)	**Std** (ms)
S_1	1097.0	48.8
S_2	160.9	9.4
S_3	12.8	1.4

Chapter 5: Conclusion

In this paper, some of image analysis techniques have been developed to be used on lung cancer diagnosis and superresolution localization microscopy. Both of them achieved advanced outcomes compare to state-of-the-art methods. For computer aided lung cancer diagnosis, we developed a four-step pure CNNs based pulmonary CAD algorithm on thoracic CT scans, which can automatically and efficiently segmented lung nodules with reasonable amount of FPs. Even though there is still a performance gap compared to some state-of-the-art methods that involved much denser 3D convolutional layers with much more complex designs, our results and evaluations demonstrated the capability of using pure 2D CNNs on a standard workstation to detect and segment pulmonary nodules with high performance. Such automation and efficiency significantly facilitate the translation from scientific researches to real applications on the computerized lung nodule segmentation trend. For the future outlook, further evaluation based on independent testing dataset is desired to perform a more comprehensive comparison. Besides, for any 2D detection scheme implemented on volumetric imaging, the model performance is also critical when being applied across an entire volume. We will assess it in the future. Finally, it is attractive to implement more advanced 2D CNNs architectures (such as generative adversarial network in Goodfellow et al., 2014) to further boost the performance.

As for superresolution localization microscopy, we have developed a Deep Matching method based on convolutional neural networks to improve super resolution localization images. This method not only significantly improves the localization accuracy in the low SNR range, and can potentially overcome the practical Cramér-Rao bound of localization accuracy by separating the signal information from noise. Also it reduces the localization performance variations among common algorithms by significantly improving the SNR of processed images. In the testing phase, this CNN architecture is a feed-forward calculation without iteration. The CNN computing time on a desktop computer is about 78 μs for processing a single image. This CNN process can be easily adopted by current localization software for improved performance. For real imaging

experiments this CNN can be trained with raw images as training data, and ground truth PSFs can be obtained with strong fluorophores e.g. nanoparticles at high SNR levels. The current bottleneck of most localization algorithms is the detection step which misses almost half of the activated fluorophores in the low SNR range (Sage et al., 2015). The Jaccard index of current deep learning based detection is slightly higher than previous methods (Boyd et al., 2018). CNNs could possibly improve the detection with a generative adversarial network (GAN) structure (Ouyang et al., 2018; Su et al., 2018). Such improved detection could have a large impact on localization accuracy. One trend in current localization microscopy imaging is to turn on as many as possible fluorescent molecules in one imaging frame in order to reduce the imaging time (Huang et al., 2011; Zhu et al., 2012). Such a high density fluorophore scenario will decrease localization accuracy. Therefore, implementing CNNs on high density data is attractive. Moreover, from Fig 3.10i, only 61% of the fluorophores achieve better localization accuracy. The reason is the Deep matching may encounter over fitting problem. In the future, the hybrid loss function that considers localization error is desired. In addition, we have developed an advanced algorithm to optimize the precision of localization in defocused imaging super resolution microscopy. A previous particle tracking work based on calculation of radial symmetry centers from in-focus super resolution images achieved near theoretical limits with orders-of-magnitude faster execution time over Gaussian fitting approach (Parthasarathy, 2012). With improvement of mechanical stability we expect that this method could achieve nanometer level resolution under current SNR. The advantage of defocused imaging is the capability of imaging multiple fluorophores at different depth simultaneously. Such capability allows us to achieve 3D super resolution imaging without PSD engineering. The challenge is when multiple particles are imaged simultaneously, their PSFs overlap. Further development is needed for segmentation of overlapping rings in such scenario.

References

Abadi, M., Agarwal, A., Barham, P., Brevdo, E., Chen, Z., Citro, C., Corrado, G., Davis, A., Dean, J., Devin, M., Ghemawat, S., Goodfellow, I., Harp, A., Irving, G., Isard, M., Jia, Y., Kaiser, L., Kudlur, M., Levenberg, J., Man, D., Monga, R., Moore, S., Murray, D., Shlens, J., Steiner, B., Sutskever, I., Tucker, P., Vanhoucke, V., Vasudevan, V., Vinyals, O., Warden, P., Wicke, M., Yu, Y., Zheng, X., 2015. TensorFlow: Large-Scale Machine Learning on Heterogeneous Distributed Systems. None.

Aberle, D.R., Adams, A.M., Berg, C.D., Black, W.C., Clapp, J.D., Fagerstrom, R.M., Gareen, I.F., Gatsonis, C., Marcus, P.M., Sicks, J.D., National Lung Screening Trial Research Team, 2011. Reduced Lung-Cancer Mortality with Low-Dose Computed Tomographic Screening. New England Journal of Medicine 365, 395–409.

Anirudh, R., Thiagarajan, J.J., Bremer, T., Kim, H., 2016. Lung nodule detection using 3D convolutional neural networks trained on weakly labeled data, in: Tourassi, G.D., Armato, S.G. (Eds.), Medical Imaging 2016: Computer-Aided Diagnosis. Presented at the SPIE Medical Imaging, SPIE.

Aristov, A., Lelandais, B., Rensen, E., Zimmer, C., 2018. ZOLA-3D allows flexible 3D localization microscopy over an adjustable axial range. Nature Communications 9.

Armato III, S.G., McLennan, G., Bidaut, L., McNitt-Gray, M.F., Meyer, C.R., Reeves, A.P., Clarke, L.P., 2015. Data From LIDC-IDRI [WWW Document]. Cancer Imaging Arch.

Armato, S.G., III, McLennan, G., Bidaut, L., McNitt-Gray, M.F., Meyer, C.R., Reeves, A.P., Zhao, B., Aberle, D.R., Henschke, C.I., Hoffman, E.A., Kazerooni, E.A., MacMahon, H., van Beek, E.J.R., Yankelevitz, D., Biancardi, A.M., Bland, P.H., Brown, M.S., Engelmann, R.M., Laderach, G.E., Max, D., Pais, R.C., Qing, D.P.-Y., Roberts, R.Y., Smith, A.R., Starkey, A., Batra, P., Caligiuri, P., Farooqi, A., Gladish, G.W., Jude, C.M., Munden, R.F., Petkovska, I., Quint, L.E., Schwartz, L.H., Sundaram, B., Dodd, L.E., Fenimore, C., Gur, D., Petrick, N., Freymann, J., Kirby, J., Hughes, B., Vande Casteele, A., Gupte, S., Sallam, M., Heath, M.D., Kuhn, M.H., Dharaiya,

E., Burns, R., Fryd, D.S., Salganicoff, M., Anand, V., Shreter, U., Vastagh, S., Croft, B.Y., Clarke, L.P., 2011. The Lung Image Database Consortium (LIDC) and Image Database Resource Initiative (IDRI): A Completed Reference Database of Lung Nodules on CT Scans. Medical Physics 38, 915–931.

Babcock, H.P., Zhuang, X., 2017. Analyzing Single Molecule Localization Microscopy Data Using Cubic Splines. Scientific Reports 7.

Betzig, E., Patterson, G.H., Sougrat, R., Lindwasser, O.W., Olenych, S., Bonifacino, J.S., Davidson, M.W., Lippincott-Schwartz, J., Hess, H.F., 2006. Imaging Intracellular Fluorescent Proteins at Nanometer Resolution. Science 313, 1642–1645.

Boyd, N., Jonas, E., Babcock, H., and Recht, B., 2018. DeepLoco: Fast 3D Localization Microscopy Using Neural Networks. bioRxiv.

Brunel, N., Nadal, J.-P., 1998. Mutual Information, Fisher Information, and Population Coding. Neural Computation 10, 1731–1757.

Burgess, A.E., 1999. The Rose model, revisited. Journal of the Optical Society of America A 16, 633.

Chen, K., Fu, K., Yan, M., Gao, X., Sun, X., Wei, X., 2018. Semantic Segmentation of Aerial Images With Shuffling Convolutional Neural Networks. Geoscience and Remote Sensing Letters IEEE.

Clark, K., Vendt, B., Smith, K., Freymann, J., Kirby, J., Koppel, P., Moore, S., Phillips, S., Maffitt, D., Pringle, M., Tarbox, L., Prior, F., 2013. The Cancer Imaging Archive (TCIA): Maintaining and Operating a Public Information Repository. Journal of Digital Imaging 26, 1045–1057.

Cover, T.M., Thomas, J.A, 2006. Elements of Information Theory. 2nd edition. Wiley-Interscience.

Dai, S., Lu, K., Dong, J., Zhang, Y., Chen, Y., 2015. A novel approach of lung segmentation on chest CT images using graph cuts. Neurocomputing 168, 799–807.

Dehmeshki, J., Amin, H., Valdivieso, M., Xujiong Ye, 2008. Segmentation of Pulmonary Nodules in Thoracic CT Scans: A Region Growing Approach. IEEE Transactions on Medical Imaging 27, 467–480.

Deschout, H., Zanacchi, F.C., Mlodzianoski, M., Diaspro, A., Bewersdorf, J., Hess, S.T., Braeckmans, K., 2014. Precisely and accurately localizing single emitters in fluorescence microscopy. Nature Methods 11, 253–266.

DeVore, R.A., Yang, Z., Kallergi, M., Qian, W., Clark, R.A., Clarke, L.P., Lucier, B.J., 1995. Effect of wavelet bases on compressing digital mammograms. IEEE Engineering in Medicine and Biology Magazine 14, 570–577.

Ding, J., Li, A., Hu, Z., Wang, L., 2017. Accurate pulmonary nodule detection in computed tomography images using deep convolutional neural networks, in: Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics).

Ding, Y., Li, C., 2016. Dual-color multiple-particle tracking at 50-nm localization and over 100-μm range in 3D with temporal focusing two-photon microscopy. Biomedical Optics Express 7, 4187.

Dong, C., Loy, C.C., He, K., Tang, X., 2016. Image Super-Resolution Using Deep Convolutional Networks. IEEE Transactions on Pattern Analysis and Machine Intelligence 38, 295–307.

Dong, F., Peng, J., 2014. Brain MR image segmentation based on local Gaussian mixture model and nonlocal spatial regularization. Journal of Visual Communication and Image Representation 25, 827–839.

Dou, Q., Chen, H., Yu, L., Qin, J., Heng, P.-A., 2017a. Multilevel Contextual 3-D CNNs for False Positive Reduction in Pulmonary Nodule Detection. IEEE Transactions on Biomedical Engineering 64, 1558–1567.

Dou, Q., Chen, H., Jin, Y., Lin, H., Qin, J., Heng, P.A., 2017b. Automated pulmonary nodule detection via 3D convnets with online sample filtering and hybrid-loss residual learning,

in: Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics).

FandPLimitTool. http://www.wardoberlab.com/software/fandplimittool/.

Fitzgibbon, A., Pilu, M., Fisher, R.B., 1999. Direct least square fitting of ellipses. IEEE Transactions on Pattern Analysis and Machine Intelligence 21, 476–480.

Fitzmaurice, C., Allen, C., Barber, R.M., Barregard, L., Bhutta, Z.A., Brenner, H., Dicker, D.J., Chimed-Orchir, O., Dandona, R., Dandona, L., Fleming, T., Forouzanfar, M.H., Hancock, J., Hay, R.J., Hunter-Merrill, R., Huynh, C., Hosgood, H.D., Johnson, C.O., Jonas, J.B., Khubchandani, J., Kumar, G.A., Kutz, M., Lan, Q., Larson, H.J., Liang, X., Lim, S.S., Lopez, A.D., MacIntyre, M.F., Marczak, L., Marquez, N., Mokdad, A.H., Pinho, C., Pourmalek, F., Salomon, J.A., Sanabria, J.R., Sandar, L., Sartorius, B., Schwartz, S.M., Shackelford, K.A., Shibuya, K., Stanaway, J., Steiner, C., Sun, J., Takahashi, K., Vollset, S.E., Vos, T., Wagner, J.A., Wang, H., Westerman, R., Zeeb, H., Zoeckler, L., Abd-Allah, F., Ahmed, M.B., Alabed, S., Alam, N.K., Aldhahri, S.F., Alem, G., Alemayohu, M.A., Ali, R., Al-Raddadi, R., Amare, A., Amoako, Y., Artaman, A., Asayesh, H., Atnafu, N., Awasthi, A., Saleem, H.B., Barac, A., Bedi, N., Bensenor, I., Berhane, A., Bernabé, E., Betsu, B., Binagwaho, A., Boneya, D., Campos-Nonato, I., Castañeda-Orjuela, C., Catalá-López, F., Chiang, P., Chibueze, C., Chitheer, A., Choi, J.-Y., Cowie, B., Damtew, S., das Neves, J., Dey, S., Dharmaratne, S., Dhillon, P., Ding, E., Driscoll, T., Ekwueme, D., Endries, A.Y., Farvid, M., Farzadfar, F., Fernandes, J., Fischer, F., G/hiwot, T.T., Gebru, A., Gopalani, S., Hailu, A., Horino, M., Horita, N., Husseini, A., Huybrechts, I., Inoue, M., Islami, F., Jakovljevic, M., James, S., Javanbakht, M., Jee, S.H., Kasaeian, A., Kedir, M.S., Khader, Y.S., Khang, Y.-H., Kim, D., Leigh, J., Linn, S., Lunevicius, R., El Razek, H.M.A., Malekzadeh, R., Malta, D.C., Marcenes, W., Markos, D., Melaku, Y.A., Meles, K.G., Mendoza, W., Mengiste, D.T., Meretoja, T.J., Miller, T.R., Mohammad, K.A., Mohammadi, A., Mohammed, S., Moradi-Lakeh, M., Nagel, G., Nand, D., Le Nguyen, Q., Nolte, S., Ogbo, F.A., Oladimeji, K.E., Oren, E., Pa, M., Park, E.-K., Pereira, D.M., Plass, D., Qorbani, M., Radfar, A., Rafay, A., Rahman, M., Rana, S.M., Søreide, K., Satpathy, M., Sawhney, M., Sepanlou, S.G., Shaikh, M.A.,

She, J., Shiue, I., Shore, H.R., Shrime, M.G., So, S., Soneji, S., Stathopoulou, V., Stroumpoulis, K., Sufiyan, M.B., Sykes, B.L., Tabarés-Seisdedos, R., Tadese, F., Tedla, B.A., Tessema, G.A., Thakur, J.S., Tran, B.X., Ukwaja, K.N., Uzochukwu, B.S.C., Vlassov, V.V., Weiderpass, E., Wubshet Terefe, M., Yebyo, H.G., Yimam, H.H., Yonemoto, N., Younis, M.Z., Yu, C., Zaidi, Z., Zaki, M.E.S., Zenebe, Z.M., Murray, C.J.L., Naghavi, M., 2017. Global, Regional, and National Cancer Incidence, Mortality, Years of Life Lost, Years Lived With Disability, and Disability-Adjusted Life-years for 32 Cancer Groups, 1990 to 2015. JAMA Oncology 3, 524.

Fu, L., Ma, J., Ren, Y., Han, Y.S., Zhao, J., 2017. Automatic detection of lung nodules: false positive reduction using convolution neural networks and handcrafted features, in: Armato, S.G., Petrick, N.A. (Eds.), Medical Imaging 2017: Computer-Aided Diagnosis. Presented at the SPIE Medical Imaging, SPIE.

Gabbard, H., Williams, M., Hayes, F., Messenger, C., 2018. Matching Matched Filtering with Deep Networks for Gravitational-Wave Astronomy. Physical Review Letters 120.

Ge, Z., Sahiner, B., Chan, H.-P., Hadjiiski, L.M., Cascade, P.N., Bogot, N., Kazerooni, E.A., Wei, J., Zhou, C., 2005. Computer-aided detection of lung nodules: False positive reduction using a 3D gradient field method and 3D ellipsoid fitting. Medical Physics 32, 2443–2454.

George, D., Huerta, E.A., 2018. Deep neural networks to enable real-time multimessenger astrophysics. Physical Review D 97.

Gibson, S.F., Lanni, F., 1989. Diffraction by a circular aperture as a model for three-dimensional optical microscopy. Journal of the Optical Society of America A 6, 1357.

Girshick, R., 2015. Fast R-CNN, in: Proceedings of the IEEE International Conference on Computer Vision.

Goodfellow, I., Pouget-Abadie, J., Mirza, M., 2014. Generative Adversarial Networks. arXiv Prepr. arXiv.

Guo, D., Shamai, S., Verdu, S., 2008. Mutual Information and Conditional Mean Estimation in Poisson Channels. IEEE Transactions on Information Theory 54, 1837–1849.

Guo, D., Shamai, S., Verdu, S., 2005a. Mutual Information and Minimum Mean-Square Error in Gaussian Channels. IEEE Transactions on Information Theory 51, 1261–1282.

Guo, D.; Shamai, S.; Verdu, S. Additive Non-Gaussian Noise Channels: Mutual Information and Conditional Mean Estimation. Proceedings. International Symposium on Information Theory 2005b, Adelaide, Austrilia.

Hamidian, S., Sahiner, B., Petrick, N., Pezeshk, A., 2017. 3D convolutional neural network for automatic detection of lung nodules in chest CT, in: Armato, S.G., Petrick, N.A. (Eds.), Medical Imaging 2017: Computer-Aided Diagnosis. Presented at the SPIE Medical Imaging, SPIE.

He, K., Zhang, X., Ren, S., Sun, J., 2016. Deep Residual Learning for Image Recognition, in: 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Presented at the 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), IEEE.

Hershko, E., Weiss, L.E., Michaeli, T., Shechtman, Y., 2019. Multicolor localization microscopy and point-spread-function engineering by deep learning. Optics Express 27, 6158.

Huang, B., Wang, W., Bates, M., Zhuang, X., 2008. Three-Dimensional Super-Resolution Imaging by Stochastic Optical Reconstruction Microscopy. Science 319, 810–813.

Huang, F., Schwartz, S.L., Byars, J.M., Lidke, K.A., 2011. Simultaneous multiple-emitter fitting for single molecule super-resolution imaging. Biomedical Optics Express 2, 1377.

Huang, X., Shan, J., Vaidya, V., 2017. Lung nodule detection in CT using 3D convolutional neural networks, in: 2017 IEEE 14th International Symposium on Biomedical Imaging (ISBI 2017).

Huhle, A., Klaue, D., Brutzer, H., Daldrop, P., Joo, S., Otto, O., Keyser, U.F., Seidel, R., 2015. Camera-based three-dimensional real-time particle tracking at kHz rates and Ångström accuracy. Nature Communications 6.

Jacobs, C., van Rikxoort, E.M., Twellmann, T., Scholten, E.T., de Jong, P.A., Kuhnigk, J.-M., Oudkerk, M., de Koning, H.J., Prokop, M., Schaefer-Prokop, C., van Ginneken, B., 2014.

Automatic detection of subsolid pulmonary nodules in thoracic computed tomography images. Medical Image Analysis 18, 374–384.

Jesson, A., Guizard, N., Ghalehjegh, S.H., Goblot, D., Soudan, F., Chapados, N., 2017. CASED: Curriculum adaptive sampling for extreme data imbalance, in: Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics).

Jia, S., Vaughan, J.C., Zhuang, X., 2014. Isotropic three-dimensional super-resolution imaging with a self-bending point spread function. Nature Photonics 8, 302–306.

Jin, H., Li, Z., Tong, R., Lin, L., 2018. A deep 3D residual CNN for false-positive reduction in pulmonary nodule detection. Medical Physics 45, 2097–2107.

Kappeler, A., Yoo, S., Dai, Q., Katsaggelos, A.K., 2016. Video Super-Resolution With Convolutional Neural Networks. IEEE Transactions on Computational Imaging 2, 109–122.

Kay, S.M, 1993. Fundamentals of Statistical Signal Processing: Estimation Theory. Prentice Hall.

Kesaniemi, M., Virtanen, K., 2018. Direct Least Square Fitting of Hyperellipsoids. IEEE Transactions on Pattern Analysis and Machine Intelligence 40, 63–76.

Kingma, D.P; Ba, J.L. ADAM: A Method for Stochastic Optimization. International Conference on Learning Representations (ICLR) 2015, 1-15.

Kishore, V.V., Satyanarayana, R.V.S., 2013. Performance evaluation of edge detectors - morphology based ROI segmentation and nodule detection from DICOM lung images in the noisy environment, in: 2013 3rd IEEE International Advance Computing Conference (IACC). Presented at the 2013 3rd IEEE International Advanced Computing Conference (IACC 2013), IEEE.

Kubota, T., Jerebko, A.K., Dewan, M., Salganicoff, M., Krishnan, A., 2011. Segmentation of pulmonary nodules of various densities with morphological approaches and convexity models. Medical Image Analysis 15, 133–154.

Lassen, B.C., Jacobs, C., Kuhnigk, J.-M., van Ginneken, B., van Rikxoort, E.M., 2015. Robust semi-automatic segmentation of pulmonary subsolid nodules in chest computed tomography scans. Physics in Medicine and Biology 60, 1307–1323.

LeCun, Y., Bengio, Y., Hinton, G., 2015. Deep learning. Nature 521, 436–444.

LeCun, Y., Boser, B., Denker, J.S., Henderson, D., Howard, R.E., Hubbard, W., Jackel, L.D., 1989. Backpropagation Applied to Handwritten Zip Code Recognition. Neural Computation 1, 541–551.

Lekić, V., Babić, Z., 2018. Using GANs to Enable Semantic Segmentation of Ranging Sensor Data. Zooming Innovation in Consumer Technologies Conference (ZINC).

Li, J., Deng, H., Li, P., Yu, B., 2015. Real-time infrared gas detection based on an adaptive Savitzky–Golay algorithm. Applied Physics B 120, 207–216.

Li, Q., Li, F., Doi, K., 2008. Computerized Detection of Lung Nodules in Thin-Section CT Images by Use of Selective Enhancement Filters and an Automated Rule-Based Classifier. Academic Radiology 15, 165–175.

Li, Y., Mund, M., Hoess, P., Deschamps, J., Matti, U., Nijmeijer, B., Sabinina, V.J., Ellenberg, J., Schoen, I., Ries, J., 2018. Real-time 3D single-molecule localization using experimental point spread functions. Nature Methods 15, 367–369.

Liu, H., Geng, F., Guo, Q., Zhang, C., Zhang, C., 2017. A fast weak-supervised pulmonary nodule segmentation method based on modified self-adaptive FCM algorithm. Soft Computing 22, 3983–3995.

Long, J., Shelhamer, E., Darrell, T., 2015. Fully convolutional networks for semantic segmentation, in: Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition.

Long, M., Zhu, H., Wang, J., Jordan, M.I., 2016. Deep transfer learning with joint adaptation networks. arXiv Prepr. arXiv.

Magalhães Barros Netto, S., Corrêa Silva, A., Acatauassú Nunes, R., Gattass, M., 2012. Automatic segmentation of lung nodules with growing neural gas and support vector machine. Computers in Biology and Medicine 42, 1110–1121.

Mao, Q., Zhao, S., Gong, T., Zheng, Q., 2018. An Effective Hybrid Windowed Fourier Filtering and Fuzzy C-Mean for Pulmonary Nodule Segmentation. Journal of Medical Imaging and Health Informatics 8, 72–77.

Messay, T., Hardie, R.C., Tuinstra, T.R., 2015. Segmentation of pulmonary nodules in computed tomography using a regression neural network approach and its application to the Lung Image Database Consortium and Image Database Resource Initiative dataset. Medical Image Analysis 22, 48–62.

Mortensen, K.I., Churchman, L.S., Spudich, J.A., Flyvbjerg, H., 2010. Optimized localization analysis for single-molecule tracking and super-resolution microscopy. Nature Methods 7, 377–381.

Murphy, K., van Ginneken, B., Schilham, A.M.R., de Hoop, B.J., Gietema, H.A., Prokop, M., 2009. A large-scale evaluation of automatic pulmonary nodule detection in chest CT using local image features and k-nearest-neighbour classification. Medical Image Analysis 13, 757–770.

Nair, Vinod, E. Hinton, Geoffrey. (2010). Rectified Linear Units Improve Restricted Boltzmann Machines Vinod Nair. Proceedings of ICML. 27. 807-814.

Nehme, E., Weiss, L.E., Michaeli, T., Shechtman, Y., 2018. Deep-STORM: super-resolution single-molecule microscopy by deep learning. Optica 5, 458.

Nehme, E., Hershko, E., Weiss, L.E., Michaeli, T., Shechtman, Y., 2019. Deep learning for dense and multicolor localization microscopy (Conference Presentation), in: Gregor, I., Gryczynski, Z.K., Koberling, F. (Eds.), Single Molecule Spectroscopy and Superresolution Imaging XII. Presented at the Single Molecule Spectroscopy and Superresolution Imaging XII, SPIE.

Niemeijer, M., Loog, M., Abràmoff, M.D., Viergever, M.A., Prokop, M., van Ginneken, B., 2011. On Combining Computer-Aided Detection Systems. IEEE Transactions on Medical Imaging 30, 215–223.

Niu, L., Yu, J., 2008. Investigating Intracellular Dynamics of FtsZ Cytoskeleton with Photoactivation Single-Molecule Tracking. Biophysical Journal 95, 2009–2016.

Ober, R.J., Ram, S., Ward, E.S., 2004. Localization Accuracy in Single-Molecule Microscopy. Biophysical Journal 86, 1185–1200.

Oron, D., Tal, E., Silberberg, Y., 2005. Scanningless depth-resolved microscopy. Optics Express 13, 1468.

Ouyang, W., Aristov, A., Lelek, M., Hao, X., Zimmer, C., 2018. Deep learning massively accelerates super-resolution localization microscopy. Nature Biotechnology 36, 460–468.

Ovesný, M., Křížek, P., Borkovec, J., Švindrych, Z., Hagen, G.M., 2014. ThunderSTORM: a comprehensive ImageJ plug-in for PALM and STORM data analysis and super-resolution imaging. Bioinformatics 30, 2389–2390.

Palomar, D.P., Verdu, S., 2007. Representation of Mutual Information Via Input Estimates. IEEE Transactions on Information Theory 53, 453–470.

Parthasarathy, R., 2012. Rapid, accurate particle tracking by calculation of radial symmetry centers. Nature Methods 9, 724–726.

Pavani, S.R.P., Thompson, M.A., Biteen, J.S., Lord, S.J., Liu, N., Twieg, R.J., Piestun, R., Moerner, W.E., 2009. Three-dimensional, single-molecule fluorescence imaging beyond the diffraction limit by using a double-helix point spread function. Proceedings of the National Academy of Sciences 106, 2995–2999.

Qian, W., Kallergi, M., Clarke, L.P., 1993. Order statistic-neural network hybrid filters for gamma camera-bremsstrahlung image restoration. IEEE Transactions on Medical Imaging 12, 58–64.

Reeves, A.P., Chan, A.B., Yankelevitz, D.F., Henschke, C.I., Kressler, B., Kostis, W.J., 2006. On measuring the change in size of pulmonary nodules. IEEE Transactions on Medical Imaging 25, 435–450.

Ren, S., He, K., Girshick, R., Sun, J., 2017. Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks. IEEE Transactions on Pattern Analysis and Machine Intelligence 39, 1137–1149.

Rivenson, Y., Göröcs, Z., Günaydin, H., Zhang, Y., Wang, H., Ozcan, A., 2017. Deep learning microscopy. Optica 4, 1437.

Rust, M.J., Bates, M., Zhuang, X., 2006. Sub-diffraction-limit imaging by stochastic optical reconstruction microscopy (STORM). Nature Methods 3, 793–796.

Sage, D., Kirshner, H., Pengo, T., Stuurman, N., Min, J., Manley, S., Unser, M., 2015. Quantitative evaluation of software packages for single-molecule localization microscopy. Nature Methods 12, 717–724.

Sage, D., Pham, T.-A., Babcock, H., Lukes, T., Pengo, T., Chao, J., Velmurugan, R., Herbert, A., Agrawal, A., Colabrese, S., Wheeler, A., Archetti, A., Rieger, B., Ober, R., Hagen, G.M., Sibarita, J.-B., Ries, J., Henriques, R., Unser, M., Holden, S., 2019. Super-resolution fight club: assessment of 2D and 3D single-molecule localization microscopy software. Nature Methods 16, 387–395.

Savitzky, A., Golay, M.J.E., 1964. Smoothing and Differentiation of Data by Simplified Least Squares Procedures. Analytical Chemistry 36, 1627–1639.

Schnitzbauer, J., Wang, Y., Zhao, S., Bakalar, M., Nuwal, T., Chen, B., Huang, B., 2018. Correlation analysis framework for localization-based superresolution microscopy. Proceedings of the National Academy of Sciences 115, 3219–3224.

Setio, A.A.A., Ciompi, F., Litjens, G., Gerke, P., Jacobs, C., van Riel, S.J., Wille, M.M.W., Naqibullah, M., Sanchez, C.I., van Ginneken, B., 2016. Pulmonary Nodule Detection in CT Images: False Positive Reduction Using Multi-View Convolutional Networks. IEEE Transactions on Medical Imaging 35, 1160–1169.

Setio, A.A.A., Jacobs, C., Gelderblom, J., van Ginneken, B., 2015. Automatic detection of large pulmonary solid nodules in thoracic CT images. Medical Physics 42, 5642–5653.

Shechtman, Y., Weiss, L.E., Backer, A.S., Lee, M.Y., Moerner, W.E., 2016. Multicolour localization microscopy by point-spread-function engineering. Nature Photonics 10, 590–594.

Shechtman, Y., Weiss, L.E., Backer, A.S., Sahl, S.J., Moerner, W.E., 2015. Precise Three-Dimensional Scan-Free Multiple-Particle Tracking over Large Axial Ranges with Tetrapod Point Spread Functions. Nano Letters 15, 4194–4199.

Shen, D., Wu, G., Suk, H.-I., 2017. Deep Learning in Medical Image Analysis. Annual Review of Biomedical Engineering 19, 221–248.

Siegel, R.L., Miller, K.D., Jemal, A., 2018. Cancer statistics, 2018. CA: A Cancer Journal for Clinicians 68, 7–30.

Simonyan, K.; Zisserman, A., 2015. Very Deep Convolutional Networks for Large-scale Image Recognition. International Conference on Learning Representations (ICLR) 2015, 1-14.

Single-Molecule Localization Microscopy Software Benchmarking. http://bigwww.epfl.ch/smlm

Small, A., Stahlheber, S., 2014. Fluorophore localization algorithms for super-resolution microscopy. Nature Methods 11, 267–279.

Smith, C.S., Joseph, N., Rieger, B., Lidke, K.A., 2010. Fast, single-molecule localization that achieves theoretically minimum uncertainty. Nature Methods 7, 373–375.

Sørensen, T.J., 1948. A method of establishing groups of equal amplitude in plant sociology based on similarity of species and its application to analyses of the vegetation on Danish commons. Kongelige Danske Videnskabernes Selskab.

Su, M., Zhang, H., Schawinski, K., Zhang, C., and Cianfrocco, M.A., 2018. Generative adversarial networks as a tool to recover structural information from cryo-electron microscopy data. bioRxiv.

Sun, W., Huang, X., Tseng, T.-L.B., Qian, W., 2017a. Automatic lung nodule graph cuts segmentation with deep learning false positive reduction, in: Armato, S.G., Petrick, N.A. (Eds.),

Medical Imaging 2017: Computer-Aided Diagnosis. Presented at the SPIE Medical Imaging, SPIE.

Sun, W., Zheng, B., Huang, X., Qian, W., 2017b. Balance the nodule shape and surroundings: a new multichannel image based convolutional neural network scheme on lung nodule diagnosis, in: Armato, S.G., Petrick, N.A. (Eds.), Medical Imaging 2017: Computer-Aided Diagnosis. Presented at the SPIE Medical Imaging, SPIE.

Sun, W., Tseng, T.-L. (Bill), Zhang, J., Qian, W., 2017c. Enhancing deep convolutional neural network scheme for breast cancer diagnosis with unlabeled data. Computerized Medical Imaging and Graphics 57, 4–9.

Sun, W., Zheng, B., Qian, W., 2017d. Automatic feature learning using multichannel ROI based on deep structured algorithms for computerized lung cancer diagnosis. Computers in Biology and Medicine 89, 530–539.

Sun., W., 2017e. Deep Learning Method vs. Hand-Crafted Features for Lung Cancer Diagnosis and Breast Cancer Risk Analysis, Ph.D. Thesis. The University of Texas at El Paso.

Sun, W., Zheng, B., Qian, W., 2016. Computer aided lung cancer diagnosis with deep learning algorithms, in: Tourassi, G.D., Armato, S.G. (Eds.), Medical Imaging 2016: Computer-Aided Diagnosis. Presented at the SPIE Medical Imaging, SPIE.

Sun, X., Qian, W., Song, D., 2004. Ipsilateral-mammogram computer-aided detection of breast cancer. Computerized Medical Imaging and Graphics 28, 151–158.

Tachibana, R., Kido, S., 2006. Automatic segmentation of pulmonary nodules on CT images by use of NCI lung image database consortium, in: Reinhardt, J.M., Pluim, J.P.W. (Eds.), Medical Imaging 2006: Image Processing. Presented at the Medical Imaging, SPIE.

Tajbakhsh, N., Suzuki, K., 2017. Comparing two classes of end-to-end machine-learning models in lung nodule detection and classification: MTANNs vs. CNNs. Pattern Recognition 63, 476–486.

Tan, M., Deklerck, R., Jansen, B., Bister, M., Cornelis, J., 2011. A novel computer-aided lung nodule detection system for CT images. Medical Physics 38, 5630–5645.

Tan, Y., Schwartz, L.H., Zhao, B., 2013. Segmentation of lung lesions on CT scans using watershed, active contours, and Markov random field. Medical Physics 40, 43502.

Tang, Y., Dai, L., Zhang, X., Li, J., Hendriks, J., Fan, X., Gruteser, N., Meisenberg, A., Baumann, A., Katranidis, A., Gensch, T., 2015. SNSMIL, a real-time single molecule identification and localization algorithm for super-resolution fluorescence microscopy. Scientific Reports 5.

Tuinstra, T.R., 2008. Automatic Segmentation of Small Pulmonary Nodules in Computed Tomography Data Using a Radial Basis Function Neural Network with Application to Volume Estimation. Ph.D. Thesis. University of Dayton.

Turin, G., 1960. An introduction to matched filters. IEEE Transactions on Information Theory 6, 311–329.

van Ginneken, B., 2006. Supervised Probabilistic Segmentation of Pulmonary Nodules in CT Scans, in: Medical Image Computing and Computer-Assisted Intervention – MICCAI 2006. Springer Berlin Heidelberg, pp. 912–919.

Wang, G., Ye, J.C., Mueller, K., Fessler, J.A., 2018. Image Reconstruction is a New Frontier of Machine Learning. IEEE Transactions on Medical Imaging 37, 1289–1296.

Wang, H., Rivenson, Y., Jin, Y., Wei, Z., Gao, R., Günaydın, H., Bentolila, L.A., Kural, C., Ozcan, A., 2018. Deep learning enables cross-modality super-resolution in fluorescence microscopy. Nature Methods 16, 103–110.

Wang, Q., Song, E., Jin, R., Han, P., Wang, X., Zhou, Y., Zeng, J., 2009. Segmentation of Lung Nodules in Computed Tomography Images Using Dynamic Programming and Multidirection Fusion Techniques1. Academic Radiology 16, 678–688.

Wang, S., Zhou, M., Liu, Z., Liu, Z., Gu, D., Zang, Y., Dong, D., Gevaert, O., Tian, J., 2017. Central focused convolutional neural networks: Developing a data-driven model for lung nodule segmentation. Medical Image Analysis 40, 172–183.

Weiss, K., Khoshgoftaar, T.M., Wang, D., 2016. A survey of transfer learning. Journal of Big Data 3.

Winkels, M., Cohen, T.S., 2018. 3D G-CNNs for Pulmonary Nodule Detection. arXiv.

Wu, B., Zhou, Z., Wang, J., Wang, Y., 2018. Joint learning for pulmonary nodule segmentation, attributes and malignancy prediction, in: 2018 IEEE 15th International Symposium on Biomedical Imaging (ISBI 2018). Presented at the 2018 IEEE 15th International Symposium on Biomedical Imaging (ISBI 2018), IEEE.

Xia, Z., Wang, X., Sun, X., Liu, Q., Xiong, N., 2014. Steganalysis of LSB matching using differences between nonadjacent pixels. Multimedia Tools and Applications 75, 1947–1962.

Ye, H., Sun, C., Ren, P., Dai, L., Peng, B., Wang, K., Qian, W., Zhang, J., 2012. Mini-array of multiple tumor-associated antigens (TAAs) in the immunodiagnosis of breast cancer. Oncology Letters 5, 663–668.

Ypsilantis, P.-P., Montana, G., 2016. Recurrent Convolutional Networks for Pulmonary Nodule Detection in CT Imaging. arXiv.

Yu, B., Yang, L., Chen, F., 2018. Semantic Segmentation for High Spatial Resolution Remote Sensing Images Based on Convolution Neural Network and Pyramid Pooling Module. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing 11, 3252–3261.

Zhang, K., Zuo, W., Chen, Y., Meng, D., Zhang, L., 2017. Beyond a Gaussian Denoiser: Residual Learning of Deep CNN for Image Denoising. IEEE Transactions on Image Processing 26, 3142–3155.

Zhang, Y., Sankar, R., Qian, W., 2007. Boundary delineation in transrectal ultrasound image for prostate cancer. Computers in Biology and Medicine 37, 1591–1599.

Zhu, G., van Howe, J., Durst, M., Zipfel, W., Xu, C., 2005. Simultaneous spatial and temporal focusing of femtosecond pulses. Optics Express 13, 2153.

Zhu, L., Zhang, W., Elnatan, D., Huang, B., 2012. Faster STORM using compressed sensing. Nature Methods 9, 721–723.

Zhu, W., Liu, C., Fan, W., Xie, X., 2018. DeepLung: Deep 3D Dual Path Nets for Automated Pulmonary Nodule Detection and Classification, in: 2018 IEEE Winter Conference on

Applications of Computer Vision (WACV). Presented at the 2018 IEEE Winter Conference on Applications of Computer Vision (WACV), IEEE.

Zhu, Y., Tan, Y., Hua, Y., Zhang, G., Zhang, J., 2011. Automatic Segmentation of Ground-Glass Opacities in Lung CT Images by Using Markov Random Field-Based Algorithms. Journal of Digital Imaging 25, 409–422.

Appendix: Mutual Information, Fisher Information and Localization Accuracy

Here, we briefly discuss the information theory on estimation of particle position from acquired images. In a microscope system, the input θ is the fluorophore position, the output ($\boldsymbol{y_1}$) is a pixelated image of the optical system PSF ($\boldsymbol{x}$) with added Poisson noise ($\boldsymbol{n}$) (Figure A.1a). This noisy image ($\boldsymbol{y_1}$) is fed into the CNN for processing to obtain a clean image ($\boldsymbol{y_2}$) that is composed of $\boldsymbol{x}$, and error ($\boldsymbol{e}$) from network training and testing. This error $\boldsymbol{e}$ is much smaller than the original noise $\boldsymbol{n}$ as shown in Figure 3.3. All θ, $\boldsymbol{x}$, $\boldsymbol{n}$, and $\boldsymbol{e}$ are random variables. Localization algorithms estimate θ from the measured images $\boldsymbol{y_1}$ or $\boldsymbol{y_2}$. The Cramér-Rao bound (CRB) of localization accuracy is based on the Fisher information of estimation theory which states that there is a lower bound for the variance of any unbiased estimator $\widehat{\boldsymbol{\theta}}$ (Ober et al., 2004), i.e.

$$\mathrm{var}\left(\widehat{\boldsymbol{\theta}}\right) \geq \mathcal{J}^{-1}(\boldsymbol{\theta}) \tag{19}$$

The general expression for the Fisher information $\boldsymbol{\mathcal{J}(\theta)}$ of estimate is (Kay et al., 2007)

$$\mathcal{J}(\boldsymbol{\theta}) = \int \left(\frac{\partial \ln p(\boldsymbol{y}|\boldsymbol{\theta})}{\partial \boldsymbol{\theta}}\right)^2 p(\boldsymbol{y}|\boldsymbol{\theta}) \tag{20}$$

where $p(\boldsymbol{y}|\boldsymbol{\theta})$is the conditional probability density function of $\boldsymbol{y}$ given θ.

Studying the link between estimation theory and other theoretic information quantities (such as entropy, and mutual information) is going on in two fields, communication channels (Palomar et al., 2007) and neural computation (Brunel et al., 1998). Mutual information is a measure of the mutual dependence between two random variables. In a communication channel model, it measures how much information can be transmitted through a noisy channel reliably given a certain input signal, whereas the mean squared error (MSE) statistically determines how accurately input sample can be estimated with the corresponding channel output (Guo et al., 2005). For an unbiased estimator

$$\mathrm{MSE}\left(\widehat{\boldsymbol{\theta}}\right) = E\left[\left(\widehat{\boldsymbol{\theta}} - \boldsymbol{\theta}\right)^{\mathbf{2}}\right] = \mathrm{var}\left(\widehat{\boldsymbol{\theta}}\right) \tag{21}$$

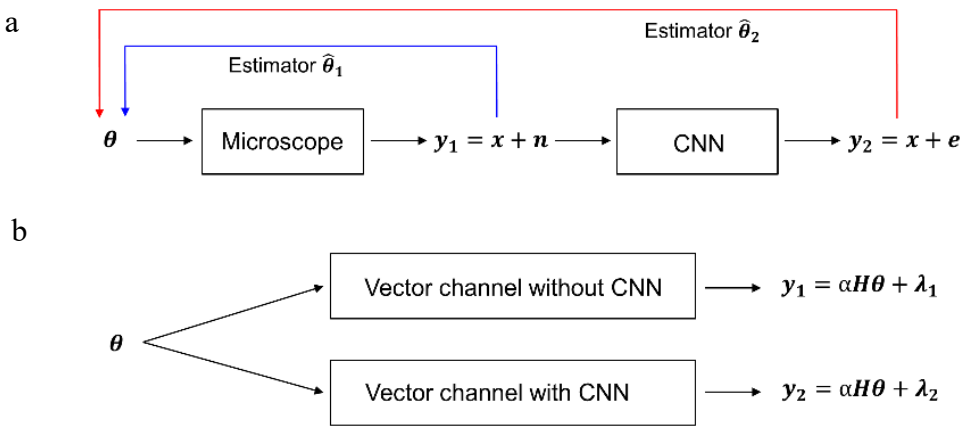

Figure A.1: Schematic representation of the microscope and CNN system. (a) The noisy output image from microscope y1 is used to obtain an estimator $\hat{\boldsymbol{\theta}}_1$ of the fluorophore position θ, and the CNN processed image y2 is used to obtain another estimator $\hat{\boldsymbol{\theta}}_2$. (b) Communication channel model of (a).

To simplify the analysis, we transform the microscope and CNN system into an abstract model of vector communication channels (Figure A.1b). The input θ is fed into the first noisy channel (without CNN), and the output is an image $\boldsymbol{y_1}$. Or this input θ passes through the clean channel (with CNN) to obtain output $\boldsymbol{y_2}$. H is a matrix representing channel transformation, α is the channel gain coefficient, λ is the channel noise and λ1>λ2. It has been shown that when λ decreases, mutual information I[θ, y] is a monotonically increasing function for channels with Gaussian, Poisson and other types of noises (Guo et al., 2005a; Guo et al., 2005b; Guo et al., 2008). Therefore,

$$I[\boldsymbol{\theta}, \boldsymbol{y_2}] > I[\boldsymbol{\theta}, \boldsymbol{y_1}] \tag{22}$$

For simple demonstration purpose we calculated the mutual information between y and the intermediate variable x at SNR = 5 for 1000 images (Figure A.2). The mean and standard deviation for $I[\boldsymbol{x'}, \boldsymbol{y_1}]$ are 1.007 and 0.219, and the mean and standard deviation for $I[\boldsymbol{x'}, \boldsymbol{y_2}]$ are 1.875 and 0.208. This numerical calculation empirically demonstrates Equation 22.

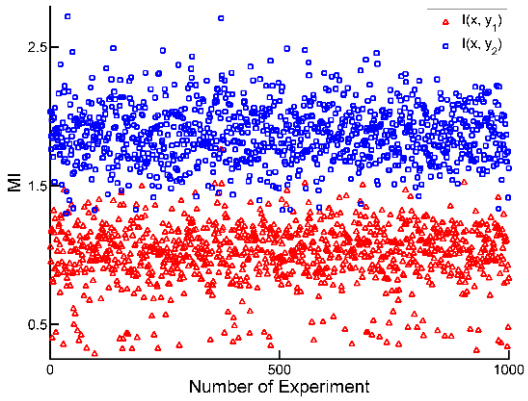

Figure A.2: Mutual information (MI) between the ideal PSF and the noisy image $I[\boldsymbol{\theta}, \boldsymbol{y_1}]$, and the CNN processed clean image $I[\boldsymbol{\theta}, \boldsymbol{y_2}]$.

For Gaussian communication channels, there is a direct link between mutual information and minimum MSE (MMSE) (Guo et al., 2005a):

$$\frac{\mathrm{d}}{\mathrm{d}\,SNR} I\left(\boldsymbol{\theta}, \boldsymbol{y} = \sqrt{SNR}\boldsymbol{H\theta} + \boldsymbol{n}\right) = \frac{1}{2}\mathrm{MMSE}(SNR) \tag{23}$$

where n is the vector noise. MMSE is the lower bound of MSE, which is equivalent to the inverse of Fisher information (Cover et al., 2006). Since MMSE is always positive, $\boldsymbol{I}\left(\boldsymbol{\theta}, \sqrt{SNR}\boldsymbol{H\theta} + \boldsymbol{n}\right)$ is a monotonically increasing function with SNR. For Poisson channels, although no such direct relationship is obtained, it shows a similar trend (Guo et al., 2008). If the second order derivative $\frac{d^2}{\mathrm{d}\,SNR^2} I(\boldsymbol{\theta}, \boldsymbol{y})$ is always negative, MMSE(SNR) is a monotonically decreasing function with SNR, which means when SNR increases the MMSE is decreasing. This has been shown in the Gaussian channel case. For Poisson channels, numerical calculations show that when the channel gain $\alpha = \sqrt{SNR}$ is not small (>1, with noise level set as 1), the second order derivative $\frac{d^2}{\mathrm{d}\,SNR^2} I(\boldsymbol{\theta}, \boldsymbol{y})$ always has negative values. This result shows that when SNR increases, the accuracy of localization algorithm with $\boldsymbol{y2}$ can overcome the practical CRB limit (CRB3) which is calculated with $\boldsymbol{y1}$.

Vita

Xia Huang was born in Wuhu, China. He graduated with a bachelor's in Biomedical Engineering from Northeastern University at Shenyang, China with the focus on biomedical hardware device design for detecting heart rate through measuring ballistocardiography signal. Later, he pursued Biomedical Engineering M.S. from Northeastern University at Shenyang, China with concentrated in medical imaging analysis. Then he continued to pursue his Ph.D. under mutual guidance of Dr. Wei Qian and Dr. Chunqiang Li from UTEP Biomedical Engineering program.

During the Ph.D. career, Xia actively published seven quality publications. These include two peer reviewed first-author journal papers, five peer reviewed conferences, and two technical oral presentations at international conferences. He also worked as an algorithm scientist intern and at VIDA Diagnostics Inc. started from May 2017.

After graduation, he plans on pursuing a career either in academia or industry in the fields of medical image analysis and deep learning.

Contact Information : xhuang3@miners.utep.edu