Large Residual Multiple View 3D CNN for False Positive Reduction in Pulmonary Nodule Detection

Thesis by

Dobrenkii Anton Nikolaevich

In Partial Fulfillment of the Requirements for the

degree of

Bachelor Degree

INNOPOLIS UNIVERSITY

Innopolis, Russian Federation

2017

Defended 08 May 2017

ACKNOWLEDGEMENTS

ABSTRACT

Pulmonary nodules detection plays a significant role in the early detection and treatment of lung cancer. And false positive reduction is the one of the major parts of pulmonary nodules detection systems. In this study was provided a framwork that solves following problems: lungs segmentation, left and right lung separation, nodule candidates detection and false positive reduction. Two novel methods were proposed. The first one aims at lungs separation and is made via dilation propagation approach, described in this work, gains better performance in term of time complexity in comparison with the State of the Art methods. The second algorithm which objective is to recognizing real pulmonary nodule among a large group of candidates was proposed. The latter method consists of three steps: appropriate receptive field selection, feature extraction and a strategy for high level feature fusion and classification. Receptive field's objective is to fit tradeoff between covering 3D nature of nodules appearance and parameters amount. Deep residual 3D CNN acing over such receptive fields prior to fusion part provide an opportunity for spatial merging. Multi-scale information was handled by dimensionality reduction step as part of feature extraction network. The dataset consists of 888 patient's chest volume low dose computer tomography (LDCT) scans, selected from publicly available LIDC-IDRI dataset. This dataset was marked by LUNA16 challenge organizers resulting in 1186 nodules. Trivial data augmentation and dropout were applied in order to avoid overfitting. Proposed method achieved high competition performance metric (CPM) of 0.735 and sensitivities of 78.8% and 83.9% at 1 and 4 false positives per scan, respectively. This study is also accompanied by detailed descriptions and results overview in comparison with the state of the art solutions.

PUBLISHED CONTENT AND CONTRIBUTIONS

Anton. "CT Framework (2017)". In: *URL https://github .com/vessemer/LungCancerDetection*.

TABLE OF CONTENTS

LIST OF ILLUSTRATIONS

LIST OF TABLES

Chapter 1

INTRODUCTION

Nowadays pulmonary nodules detection still requires high carefulness from a doctor in order to perform a proper examination and analysis of volumetric thoracic Computed Tomography (CT). Nodules are actually spots inside the lung volume caused by difference in density. They vary in size from 3 mm up to 3 cm in diameter. Nodules larger than 3 cm, are considered as a lung mass instead of nodules and is more likely to be cancerous. Relatively small size of nodules and the fact that the environment around them often varies explains the complexity of the early detection of lung cancer Firmino et al., 2014. For example chapter 1 presents pairs of true pulmonary nodules and false positives with similar morphology. Usually to become visible on frontal chest X-ray, nodules need to grow up to 1 cm. However, with recent advancement in technology, screening with low-dose spiral computed tomography (LDCT) can now detect nodules of up to 3 mm's in diameters. This leads to a reduction in lung cancer mortality by 20% in comparison with standard frontal chest X-ray. "American Cancer Society, 2016. Cancer facts and figures 2016." n.d.; "3rd Annual Congress of the Russian Breast Cancer Society." n.d.; "European Consensus Statement on Lung Cancer: Risk factors and prevention. Lung Cancer Panel." n.d.

For the last two decades researchers have been actively studying approaches that can automatically detect pulmonary nodules in computed tomography images. In general, problem of pulmonary nodules detection can be divided into two tasks: candidates detection and further false positive reduction. In candidates detection the number of candidates are extracted using different criteria e.g., intensity distribution or tissues separation via Hausfield Units thresholds. In false positive reduction discriminative methods are developed in order to lower the amount of candidates.

The major progress in dataset collection for comparative analysis of performance evaluation has been achieved due to the medical image analysis challenge **LUNA16** held in conjunction with ISBI 2016. It also provides an evaluation framework for automatic nodule detection. The detailed description of the dataset can be found in section II. Under the competitive ISBI LUNA16 programme, a number of novel methods have been proposed for both candidates detection and false positive
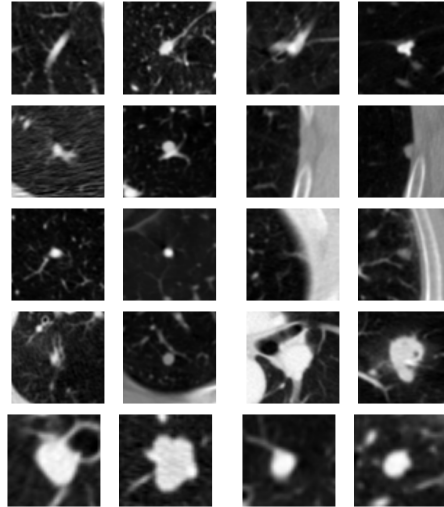
Figure 1.1: Here you can see pairs of horizontal views of pictures with pulmonary noodles on the right and false positives on the left provided by LUNA16.

reduction tasks. The most notable results are listed in the **LUNA16** report.

The candidates selection is often performed based on local intensity distribution due to the features of nodule appearance: it may be modeled as an ellipsoid with a radial intensity decreasing. How it was done by Tan et al., 2011 via merging clusters generated by the minima of divergence of normalized gradient taken over CT scan and by dot and line 3D selective enhancement filters which was also applied over CT scan Li, Sone, and Doi, 2003.

In the false positive reduction track two main approaches have been use. The one that took the first place is proposed by Dou et al., 2016. In this work three multi-level contextual 3D convolutional neural networks were employed, with an aim of handling different scale features of pulmonary nodules. Those networks were trained separately and then merged via a fusion function that was a weighted sum of predictors probability output.

Another approach was provided by A. A. A. Setio et al., 2016, which works by analyzing 2D planes and despite of the 3D nature of nodule appearance, this approach showed a comparative results. The researchers developed false positive reduction method constituting a combination of committee-fusion, late-fusion and mixed-fusion of 2D CNNs acting over 2D patches that were extracted using nine views of a volumetric object.

In the LUNA16 report, it was also shown that the combination of the multiple

algorithms showed better performance than the use of a single algorithm for both the false positive reduction and candidates detection tasks. However, the problem is still not fully solved, and needs further investigation. This was the motivator to implement novel false positive reduction algorithm with the method that combines the strengths of multiple views A. A. A. Setio et al., 2016 with the advantages of 3D multi-scale approach in order to add it to the ensemble of algorithms.

Chapter 2

METHODOLOGY

In this study we developed a novel neural network architecture aimed at combining benefits from multiple views approach with multi-scale one. Such effect was achieved by selecting appropriate receptive fields which cover multiple views and merging strategy due to which multi-scale information was also handled. To ensure robustness simple ensemble technique was employed. It takes the average from two predictions made from the same input, mirrored along profile view (due to relative symmetry between left and right lungs).

2.1 Lungs Segmentation

In low dose CT scans lungs segmentation performs relatevly easay in compare, with lung segmentation in chest radiography, for example in the work of Van Ginneken, Stegmann, and Loog, 2006 they showed that multi-resolution pixel classification method which employs a multi-scale filter bank of Gaussian derivatives and a k-nearest-neighbors classifier outperform both standart approaches such as active shape models and active appearance models. For the same task but in CT volume, simple sequential thresholding in Hounsfield scale Brooks, 1977 (the Hounsfield units of common substances was presented at Appendix B), performs well due to the natural contrast between the low-density lungs and the surrounding high-density chest wall. In a number of semi-automatic methods provided by Kalender et al., 1991, Hedlund et al., 1982, Keller, Edwards, and Rundle, 1981, where lungs were segmented at each transverse plane via edge tracking, manual interaction is still required in order to edit the resulting segmentation. The same intervention in automatic process, in a form of threshold values and seeds selection for 3D region growing, is also requared in the algorithms proposed by ERIC A Hoffman, Sinak, et al., 1983, ERIC A Hoffman and Ritman, 1985 and ERIC A Hoffman, 1985. The main reason for such interaction, is that approaches, listed above, have a lack of a shape inconsistency along lungs' outer border or lungs conglutination (discribed in details in the next section). In the work of Rikxoort et al., 2009, manual intervention has been avoided by a consistent region growing and a number of automatic errors detection tests, consist of statistically derived valid intervals for the following values: a left lung's volume, a right lung's volume, a ratio between lungs' volumes and an

analysis of lung border convexity. The latter and the major segmentation test consist of lung's convex hull reconstruction at each transverse slice and comparison with the result of segmentation.

2.2 Lungs Separation

Prior to the tests application, lungs must be separated in order to deal with the case of conglutination. This can be observed at the transvers slices where the anterior and posterior junctions between the lungs' borders may be very thin. In the lung segmentation approach of Rikxoort et al., 2009, lungs separation algorithm proposed by Hu, Eric A Hoffman, and Reinhardt, 2001 has been employed. This method relies on sequential 2D morphological operations in transverse slices instead of 3D morphology operations used in the work of Brown et al., 1997, due to the time complexity issue. More concrete, they apply at each slice, which consist of single connected component, a binary morphological erosion with a scaled four-connected binary structuring element prior to conditional dilation aimed at lung's shape reconstruction. In the following work an algorithm of dilation propagation was proposed 1, 2. The main idea of the algorithm is to select reference binary image (i.e. image in term of functions) of the lungs in a one slice which further will be propagated over all adjacent slices, that have not yet been visited, by comparison of its complements. This process will be repeated, taking adjacent slices from previous step, as new reference images untill lungs are not separated. This approach differs from the methods listed above in applying binary morphological erosion only once if it's needed. It allows the dilation propagation approach to gain beter performance in term of time complexity. The condition to once apply an erosion at the first stage is whether selected initial transverse slice consist of one connected component or not. In order to avoid even single errosion, selection of the initial slice may be repeated until one with two connected components are obtained. This is done due to the difference in the time complexity between component analysis with von Neumann neighborhood for which requires approximately $2hw$ operations and for erosion it's about kn, where k is amount of components in a binary structuring element and hw is number of components in a slice. Taking into account rare appearence of conglutination in slices ($\frac{1}{86}$, on average), number of sampling iterations m in 1 varies from 1 up to 6.

Algorithm 1: Separation Outline

Input $\ : \mathcal{P} : \mathbb{X} \rightarrow \{false,\ true\}$

Output : $\mathcal{D} : \mathbb{X} \rightarrow \{0,\ 1,\ 2\}$

1 $seeds \in \{z \mid z \in [\lfloor \frac{d}{2} \rfloor - m,\ \lfloor \frac{d}{2} \rfloor + m] \ \wedge \ \overline{mingled(\mathcal{P}_{[z]})}\}$

2 **if** $seeds = \emptyset$ **then**

3 $\quad\quad z \leftarrow \{\lfloor \frac{d}{2} \rfloor\}$

4 $\quad\quad$ **do**

5 $\quad\quad\quad\quad \mathcal{P}_{[z]} \leftarrow \mathcal{P}_{[z]} \ominus [1]_{7 \times 1}$

6 $\quad\quad$ **while** $mingled(\mathcal{P}_{[z]})$;

7 **else**

8 $\quad\quad z : z \in seeds$

9 **end**

10 $\mathcal{D}_{[z]} \leftarrow label(\mathcal{P}_{[z]})$

11 **for** $(i,\ j) \in \{(z,\ z-1),\ ...,\ (1,\ 0),\ (z,\ z+1),\ ...,\ (d-1,\ d)\}$ **do**

12 $\quad\quad D_{[j]} \leftarrow$ Dilation Propagation $(\mathcal{P}_{[j]},\ \mathcal{D}_{[i]},\ j)$

13 **end**

Algorithm 2: Dilation Propagation

Input $\ : \mathcal{I} : \mathbb{Y} \rightarrow \{false,\ true\}^{h \times w}, \quad \mathcal{D} : \mathbb{Y} \rightarrow \{0,\ 1,\ 2\}^{h \times w}, \quad iteration \in \mathbb{N}$

Output : $\mathcal{D} : \mathbb{Y} \rightarrow \{0,\ 1,\ 2\}^{h \times w}$

1 $amount_i(\varepsilon) \equiv |\{j \mid j \in \varepsilon,\ j = i\}|$

2 $dom_{\bar{\emptyset}}(f) \equiv \{x \mid x \in domain(f) \ \wedge \ f(x)\}, \quad \forall f : \mathbb{X} \rightarrow \{false,\ true\}$

3 $\mathcal{S} \leftarrow dom_{\bar{\emptyset}}(\mathcal{I}) \setminus dom_{\bar{\emptyset}}(\mathcal{D} \neq 0)$

4 **do**

5 $\quad\quad \mathcal{E} \leftarrow \{\{(D_{[p]},\ q) \mid p \in domain(\mathcal{D}),\ \| p - q \|_1 \leq 1\} \mid q \in \mathcal{S}\}$

6 $\quad\quad$ **for** $\varepsilon \in \mathcal{E}$ **do**

7 $\quad\quad\quad\quad left \leftarrow amount_1(\varepsilon_1)$

8 $\quad\quad\quad\quad right \leftarrow amount_2(\varepsilon_1)$

9 $\quad\quad\quad\quad$ **if** $left,\ right \in \mathbb{N}$ **then**

10 $\quad\quad\quad\quad\quad\quad left \leftarrow left + int(left = right) \cdot (-1)^{iteration}$

11 $\quad\quad\quad\quad\quad\quad \mathcal{D}_{[point]} \leftarrow 1 + int(left < right)$

12 $\quad\quad\quad\quad\quad\quad \mathcal{S} \leftarrow \mathcal{S} \setminus \{\varepsilon_2\}$

13 $\quad\quad\quad\quad$ **end**

14 $\quad\quad$ **end**

15 **while** $\mathcal{S} \neq \emptyset$;

In provided algorithms all expressions in curly brackets formed multisets. In our application, the domains $\mathbb{X}$, $\mathbb{Z}$ is represented via cartesian product of bounded intervals:

$$\mathbb{X} \equiv \mathbb{Z}^{[0,d]} \times \mathbb{Z}^{[0,h]} \times \mathbb{Z}^{[0,w]}, \tag{2.1a}$$

$$\mathbb{Y} \equiv \mathbb{Z}^{[0,h]} \times \mathbb{Z}^{[0,w]}, \tag{2.1b}$$

$$\mathbb{Z}^{[a,b]} \equiv \{x \mid x \in \mathbb{Z},\ a \leq x \leq b\},\ \forall a \leq b \in \mathbb{Z}, \tag{2.1c}$$

$$d,\ h,\ w\ \in \mathbb{N}:\ \exists N \in \mathbb{N}, max\{d,\ h,\ w\} < N. \tag{2.1d}$$

This is equivalent to the function discretization, and since intervals are bounded, functions can be represented numerically as tensors. In this way, in the line 5 of 1, the binary structuring element set to be $[1]_{7\times1} : \mathbb{Z}^{[0,6]} \times \mathbb{Z}^{[0,0]} \rightarrow 1$. The choice of such structuring element was caused by the small deviation of junction curve from the vertical line. The *label* function, which was used in algorithm 1 at the lines 10 and 15, can be defined as follow:

$$label(f)_{[i,\ j]} \equiv \begin{cases} 1, & (i,\ j) \in \text{ domain of the left connected component of } f \\ 2, & (i,\ j) \in \text{ domain of the right connected component of } f \\ 0, & otherwise \end{cases}$$

$$\forall f : \mathbb{Z}^{[0,a]} \times \mathbb{Z}^{[0,b]} \rightarrow \{false, true\},$$

$$domain(label(f)) = domain(f).$$

$$\tag{2.2}$$

The core update procedure in the algorithm 1 is the lines 13 and 15 to calculate a new reference image. If current slice are not mingled, then line 15 holds and it just set the new reference image to be labeled version of current slice. Else it going to apply the dilation propagation algorithm 2. The algorithm 2 takes as input a slice to be addressed, a number of iteration *iteration* and an adjacent reference image. At the line 2 of algorithm 2 $dom_{\bar{\emptyset}}$ has been defined, return the largest sub domain of binary function, over which it have only *true* values. In this way, at the line 3 of 2, $\mathcal{S}$ is set to be a relative complement of the non-zero areas of $\mathcal{D}$ in $\mathcal{I}$. The amount

of computaion operations needed for the complement finding is about hw.

The expression at the line 5 of the procedure 2 assigns to the $\mathcal{E}$ a set consisted of sets of vectors, the first component of which is a unit ball nighborhood of a point q from the set $\mathcal{S}$ in a corresponding metric and the seconnd component is the q itself. In general such nighborhood may be stated as follow:

$$\mathcal{B}^i_{\mathcal{D}, \epsilon}(point) \equiv \{D_{[p]} \mid p \in domain(\mathcal{D}), \; \| \, p - point \, \|_i \leq \epsilon\}, \qquad (2.3a)$$

$$\text{e.g., } \mathcal{B}^1_{\mathcal{D}, \epsilon}: domain(\mathcal{D}) \rightarrow \{0, \, 1, \, 2\}^{(2\epsilon+1)^2}. \qquad (2.3b)$$

Taking into account Equation 2.1a, the line 5 of the procedure 2 will take approximately $9|\mathcal{S}|$ operations, where $\mathcal{S}$ reduced at each iteration of while loop. First three lines $(7 - 9)$ of the for-loop of 2, can be computed by the cost of 20 operations at each while-loop itertion (since ε_1 consist of 9 components and two comparisen at the $9th$ line 9). Lines $10 - 12$ together for the whole execution will take $7|\mathcal{S}_{init}|$ operations, with assumtion that delete in $\mathcal{S}$ takes exactly one operation and with $\mathcal{S}_{init} \equiv dom_{\bar{\emptyset}}(\mathcal{I}) \setminus dom_{\bar{\emptyset}}(\mathcal{D} \neq 0)$. More, at the $10th$ line the point is considered to belong to the left lung at each even iteration in order to neutralize effect of lung's "drift".
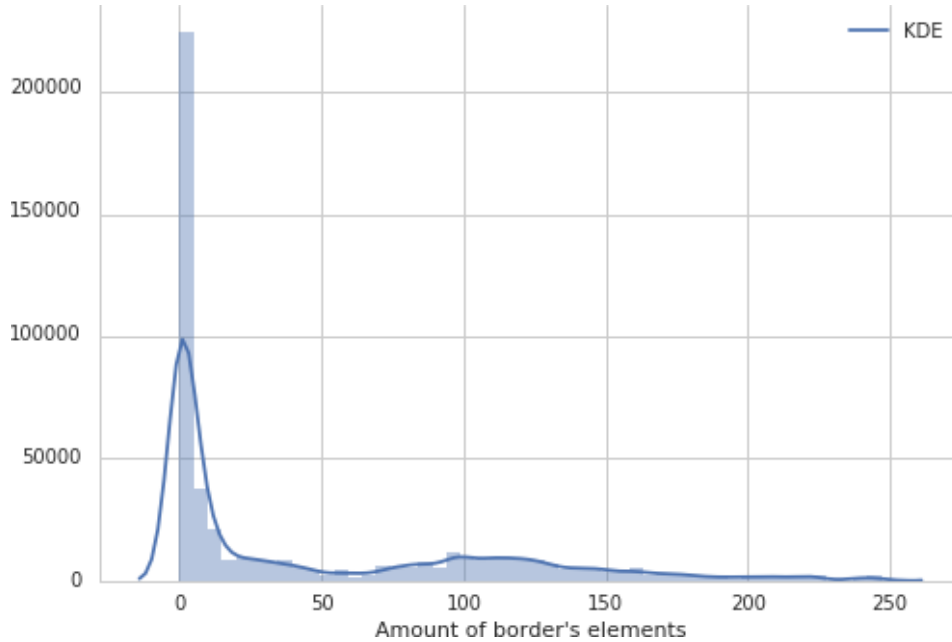
Figure 2.1: $|\mathcal{S}_{init}|$ values among 888 patients' low dose CT scans, reduced to the unit space, with on average 360 slice per patient. It has the mode equal to 8, and the maximum value of 247, while an amount of elements in the scan $hw \simeq 10^5$

Therefore, for the whole algorithm 2, in the worst case, when at each while-loop iteration only one element from $\mathcal{S}$ has been deleted, the amount of operations will be:

$$\mathcal{T}_2 = hw + 7|\mathcal{S}_{init}| + 29\frac{|\mathcal{S}_{init}|^2}{2}. \tag{2.4a}$$

A quadratic dependence on $|\mathcal{S}_{init}|$ has been obtained, since $\mathcal{S}_{init}$ mast satisfy $|\mathcal{S}_{init}| \ll n$, due to the task nature. Empirical evidence for that statement is provided in the Figure 2.1. Taking into account Equation 2.5, Therefore, for the whole algorithm 2, in the worst case, when at each while-loop iteration only one element from $\mathcal{S}$ has been deleted, the amount of operations will be:

$$\mathcal{T}_1 = z\mathcal{T}_1. \tag{2.5a}$$

Figure 2.2: Results of segmentation's stages performed over four patients. Conglutination appear ether at the first stage of sequential segmentation presented at the second column or after some morphological and regions growing operations aimed at both: shape consistency and hol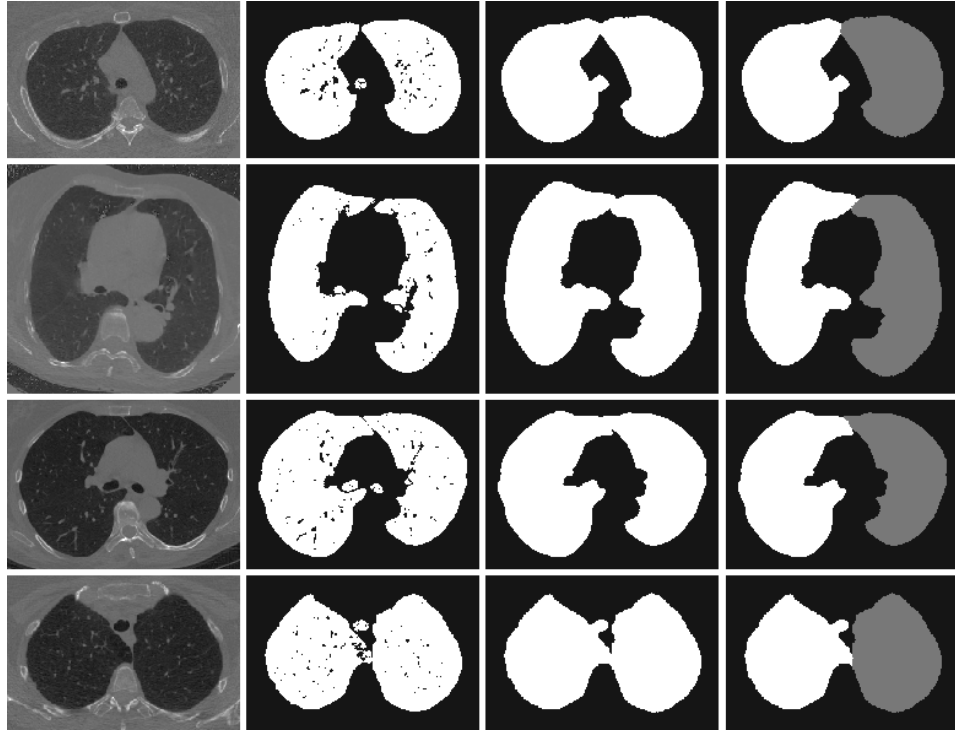es filling, at the third column. Output of lungs separation procedure applied to the latter stage of segmentation showed at the last column.

2.3 Candidates Detection

The candidates detection was performed over region of interests extracted via lung segmentation algorithm, described in previous section. The candidates selection algorithm proposed by Tan et al., 2011 was implemented as a part of the framework. Method consists of merging clusters generated by the minima of divergence of normalized gradient taken over CT scan and by dot and line 3D selective enhancement filters which was also applied over CT scan Li, Sone, and Doi, 2003. The main idea of applying dot enhanced filters in order to obtain nodules' clasters is to consider nodules and vessels to be modeled via intensity distrebuted in source's neigbourhood in conformity with Gaussian distribution.

Since such consideration takes place, condition on voxel to belong to dot center can be stated as:

$$H(d(x))_{i,j} = \frac{\partial^2 (e^{-\frac{1}{2}x^T \Sigma^{-1} x})}{\partial x_i \, \partial x_j}$$

$$= \begin{cases} \Sigma^{-1}, & i = j \\ 0, & otherwise \end{cases} \tag{2.6}$$

2.4 Convolutional Neural Networks

In this study technique of spatial information extraction and aggregation such as convolutional network has been employed, and actively used over volumetric data. The convolution is a binary operator, acting on f and φ as follow:

$$(f * \varphi)(t) = \int_{\Omega} f(\tau)\varphi(t - \tau)d\tau. \tag{2.7}$$

With this defenition basic convolutional neural network (CNN) with N layers and vector $l = (l_1, \quad ..., \quad l_N)$, where each i-th layer of CNN consists of l_i components, can be defined as $C_l \equiv \mathcal{L}_N$, where:

$$\mathcal{L}_{top} \colon \mathcal{L}_N \times \mathcal{F}(\mathcal{I}, \xi) \to \mathbb{R}, \tag{2.8}$$

$$\mathcal{L}_1 \equiv \mathcal{I}, \tag{2.9}$$

$$\mathcal{L}_{k+1} = \{\sigma(\sum_{j=1}^{|\mathcal{L}_k|} f_i * \mathcal{L}_{k_j} + b_i)\}_{i=1}^{l_{k+1}}. \tag{2.10}$$

Where $\mathcal{I}$ is the given input signal or function, defined over $Span(\Omega)$ and Ω is a compact. $\mathcal{F}(\mathcal{I}, \xi)$ is a target function (in general also dependendable on some stochastic variable ξ), which is going to be approximated via $\mathcal{L}_{|N|}$. The function $\mathcal{L}_{top}$ is the discrepancy measure of approximation and often called loss function. Some non-linear element-wise function σ, along with f_i and b_i functions should be selected in order to minimize the discrepancy.

2.5 Receptive Fields Selection

The importance of receptive fields is caused by the distinction made in the size and shape of a nodules. Influence of nodules with the minimum size of 3mm in diameter may be fully ignored by selecting huge receptive fields. On the other hand, nodules with diameter of 3 cm can't be fully covered by small receptive fields. Due to this the decision was made to use a combination made of three intersected parallelepipeds with it's largest base parallel to the frontal, horizontal and side views respectively. Parallelepiped's largest base was selected to cover all kinds of nodules in a proper 2D view. In this way, horizontal aligned parallelepiped receptive field base's side was set to 42 units. The same operation for frontal and profile aligned parallelepipeds provide base's side 36 units. Width of this parallelepipeds was selected to be 18
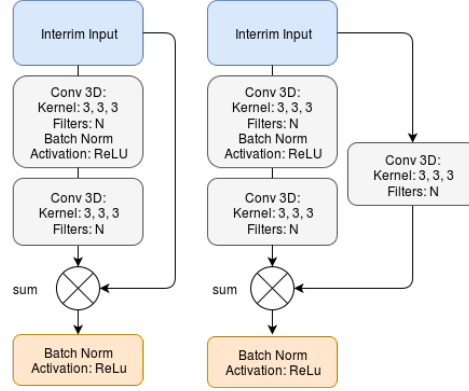
Figure 2.3: Possible configurations of residual units due to limitations described above. The first architecture means that W_s is an identical tensor (in term of convolution). The second configuration includes non-identical shortcut function $\mathcal{S}$.

units for horizontal view and 23 for frontal and profile views. Width values selection was based on volume of it's intersection which is chosen in order to cover 85% of nodules appearance. With respect to the unit space (0.9 mm, 0.7 mm, 0.7 mm). Unit space will be discussed further in Preprocessing section.

2.6 High Level Feature Extraction and Merging

Since multiple view ability for the model has been provided via multiple receptive fields, oriented along the three principal views, we needed a feature extraction method to combine information from them. For this purpose high level features Boureau et al., 2010 were extracted from each of the receptive fields and then fused together by weighted sum. In order to do this fusion, a deep neural network architecture was developed. To escape the lack of gradient vanishing in deep network, described by Glorot and Bengio, 2010 residual units have been employed. Residual units were proposed by Kaiming He et al. in deep residual learning He et al., 2015 and were successfully generalized over bunch of tasks including volumetric brain MRI segmentation, by **MRI3D** and 3D brain MRI classification, by H. Chen et al., 2016. The main idea of residual learning is based on the approximation of the residual functions instead of the original one, and generally may be written as:

$$\mathcal{Y} - \mathcal{F}(\mathcal{X}, \theta_s) = \mathcal{F}(\mathcal{X}, \theta) \tag{2.11}$$

Where $\mathcal{X}$ and $\mathcal{Y}$ are sampled random variables from some joint distribution $\mathbb{P}(\mathcal{Y}, \mathcal{X})$, and θ is a set of trainable parameters. $\mathcal{F}(\cdot), \mathcal{S}(\cdot)$ is the residual mapping and shortcut functions respectively, which is usually defined as stacked convolution operations alternating with non-linear activations. In our approach $\theta \equiv \{W, b, \mu_I, \sigma_I^2\}_i$ and
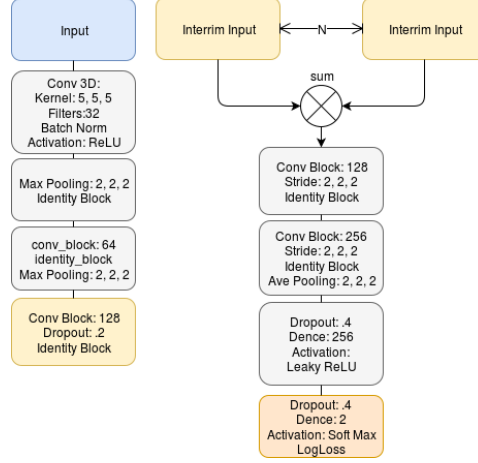
Figure 2.4: The architecture on the left represents the high level features extractor, which processes each of the receptive fields. The resulting interim output is fed as the input to the architecture on the right.

$\theta_s \equiv \{W_s, \ b_s\}$ also residual mapping function were defined as follow:

$$\mathcal{F}(X, \ \theta) = W_{|\theta|} * X_{|\theta|}, \tag{2.12a}$$

$$X_i = a[\mathcal{N}orm(W_i * X_{i-1} + b_i)], \tag{2.12b}$$

$$\mathcal{S}(X, \ \theta_s) = W_s * X. \tag{2.12c}$$

Where a is an activation function and $\mathcal{N}orm(\cdot)$ is:

$$\mathcal{N}orm(x) = \frac{x - \mu_{\mathcal{I}}}{\sqrt{\sigma_{\mathcal{I}}^2 + \epsilon}}. \tag{2.13}$$

Equation Equation 2.13 represent the mini-batch normalization technique provided by Sergey I. et al. and aimed at speed up training process of a deep networks by reducing internal covariate shift Ioffe and Szegedy, 2015, in which $\mu_{\mathcal{I}}$ and $\sigma_{\mathcal{I}}^2$ are mini-batch's mean and variance respectively. Also ϵ is a positive smoothing constant. As it was stated by He et al., 2015 in deep residual learning the use of non-identity tensor W_s in a shortcut only makes sense if it's a paired function $\mathcal{F} : (X, \ \theta) \mapsto X'$ leads to: $dom(X) \neq dom(X')$. Due to this statement and with empirically fixed $|\theta| = 2$ residual units result in two possible configuration presented on Figure 2.3

2.7 Preprocessing

In order to handle different slice thickness resampling to the unit space was performed. The unit space was selected as a tradeoff between minimization of data corruption due to resampling operation and maximization of nodule size. To satisfy that requirement new spacing was set to (0.9 mm, 0.7 mm, 0.7 mm) as a shape of voxel. To reduce impact of ribs, scan's intensity was clipped in range from -1000 up to 400 Housfield Unit and then normalized to the range of [0, 1].

Since the dataset has a significant skewness: 1 nodule against 460 false positives, the majority class under-sampling technique was applied with under-sampling co-efficient 80 Chawla, Japkowicz, and Kolcz, 2004. Additionally data augmentation in form of random center shifting and rotation on π radians along horizontal plane was performed.

2.8 Training Process

Optimization was performed via ADAM optimizer with learning rate set to 0.001, $\beta_1 = 0.9$, $\beta_2 = 0.999$, $\epsilon = 10^{-8}$ Kingma and Ba, 2014. Training was performed with batch size $bs = 64$ on the whole dataset, excluding the test part.

Prior to training the classifier, model's weights were set in accordance with the Xavier uniform initialisation algorithm Simonyan and Zisserman, 2014. Over the training set $\mathbb{D} = \{X_i, C_i\}$, where $X_i = \{I_{ix}, I_{iy}, I_{iz}\}$ and $C_i \in \{0, 1\}$ logarithmic loss was used and it was defined as follows:

$$\mathcal{L}(\theta, \mathbb{P}, \mathcal{I}) = -\frac{1}{|\mathcal{I}|} \sum_{\{X_i, C_i\} \in \mathcal{I}} \sum_{\mathcal{J} \in \{0,1\}} \mathbb{I}_{\mathcal{J} = C_i} \mathbb{P}(\mathcal{J} = C_i | X_i, \theta). \qquad (2.14)$$

Where $\mathcal{I} \subset \mathbb{D}$, $\mathbb{I}$ is an indicator function and $\mathbb{P}$ is, actually, the classifier which estimates classe probability over input X and with parameters θ.

It should also be noted that the whole NoduleResNet framework is implemented using Keras Chollet, n.d. with Tensorflow as a back-end Abadi et al., 2016. The computation time of the system over a scan with on average 330 candidates per scan is 2 second on a standard PC with a GPU NVIDIA Tesla K40. The training time for 1827 iterations with a batch size 32 take on average 3,465 seconds.

Chapter 3

EXPERIMENTS AND RESULTS

3.1 Dataset

The dataset, which has been used in this work was publicly provided by **LUNA16** challenge organizers. It consist of 888 patients' low dose CT scans selected from 1018 scans presented by **LIDC** The selection was performed based on facts of nodule's morphology appearence. In this way CTs with the inconsistent slice spacing or high slice thickness were omitted (CT scans with slice thickness greater than 3 mm were discarded from the final dataset). Remaining scans were divided into 10-folds with the objective to perform cross validation over them.

For the resulting dataset, nodules was annotated by experienced thoracic radiologists in a two-phase screening process. Due to the first stage each of four radiologist marked separately any suspicious on nodule spots, and then addressed them into one of the following categories: nodules with diameter $\geq$ 3 mm, nodules with diameter $>$ 3 mm, and non-nodule abnormality. In the next stage all nodules annotated only by one or two out of four radiologists, non-nodules, and nodules with diameter $>$ 3 mm were stated as irrelevant findings. Remaining nodules (marked by at least three out of four radiologists) were referred as the reference standard, which formed 1186 ground true nodules.

For the false positive reduction track nodule candidates were extracted via merging outputs from a banch of algorithms. The generic nodule candidates detection algorithm based on computing shape index and curvedness at every voxel of CT which is transfered isotropic resolution, ISICAD, developed by Murphy et al., XXXX. ISICAD has the major impact in term of sencitivity (it has recieve 85.6% sensitivity over proposed 888 patients' CT stans). Another method is SubsolidCAD, algorithm aimed at a detection of rare but more likely to be cancerous subsolid nodules. SubsolidCAD proposed by Jacobs et al., XXXX with the main is to thresholding by statistically inferred values in Hounsfield scale. The latter method, wich has been used in the banch, is LargeCAD, the purpose of which is to detect nodules with the diameter $\geq$ than 10 mm. LargeCAD was developed by Arnaud A. A. Setio et al., 2015. The main idea of LargeCAD is similar to SubsolidCAD but with different thresholds which is relevant to the density value of large nodules.

Table 3.1: results of the false positive reduction tasck in ISBI-LUNA16 challenge.

Team	0.125	0.25	0.5	1	2	4	8	CPM
arnaud.setio	0.636	0.727	0.792	0.844	0.876	0.905	0.916	0.814
subru1603	0.394	0.491	0.570	0.660	0.732	0.795	0.851	0.642
hirokinakano	0.113	0.165	0.265	0.465	0.596	0.695	0.785	0.440
mattdns100689	0.640	0.698	0.750	0.804	0.847	0.874	0.897	0.787
bim bam	0.453	0.535	0.591	0.635	0.696	0.741	0.797	0.635
ccanoespinosa	0.655	0.745	0.807	0.849	0.880	0.907	0.925	0.824
QiDou	0.677	0.737	0.815	0.848	0.879	0.907	0.922	0.827
Ours	0.517	0.602	0.720	0.788	0.822	0.839	0.856	0.735

The sensitivities of last two methods, computed over LUNA16 dataset is 36.1% and 31.8% respectively, and in banch with the first one, provide 551, 065 nodule candidates for the same dataset, with the sensitivity of 94.4%, which corresponds to 1120 ground true nodules out of 1186.

3.2 Evaluations Metrics

Results of cross validation has been evaluated using the Free-Response Receiver Operating Characteristic (FROC Radiation Units and Measurements, 2008) and competition performance metric (CPM Niemeijer et al., 2011). It computes average of the seven sensitivities measured at several false positives per scan (FPPS) thresholds, more concretely, at each FPPS $\in \{0.125, 0.25, 0.5, 1, 2, 4, 8\}$ true positive rate was computed. Mean of those values forms the CPM.

3.3 Results

The resulted FROC calculated over zero fold provided LUNA16 dataset and is shown in Figure 3.1. The sensitivity of 0.788 and 0.839 was reached at 1 and 4 false positives per scan respectively with corresponding CPM score 0.735. CPM score in comparison with different teams is provided at Table 3.1. As it can be seen the result of this implementation is competitive in comparison with other methods.
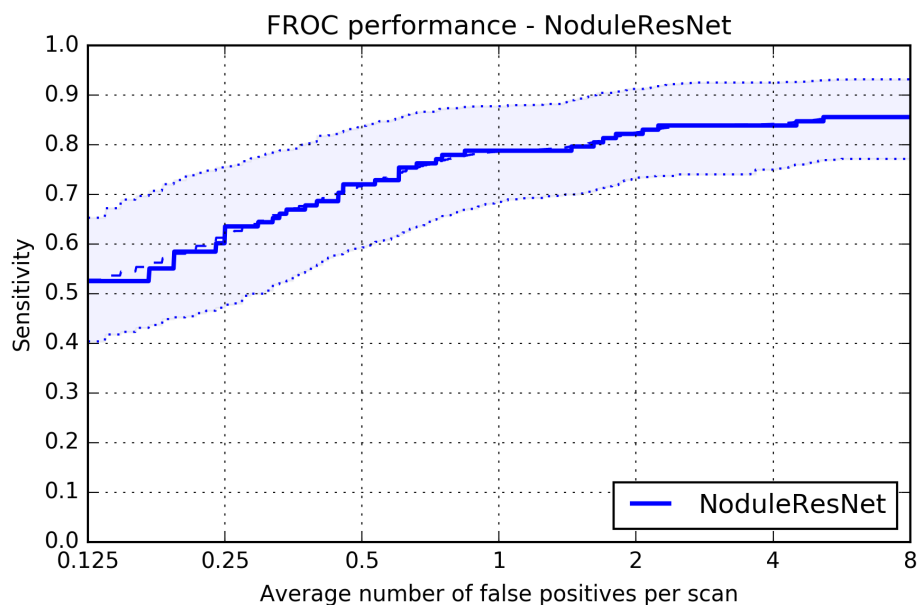
Figure 3.1: FROC curve calculated over zero fold via LUNA16 framework as a result of proposed method, dashed curves stands for 95% confidence intervals, evaluated using bootstrap (Efron and Tibshirani, 1994)

Figure 3.2: Left: Pulmonary nodules detected by our method method with high confidence. Each patch is a representative maximum through three sequential horizontal planes of one annotated nodule. P-values for all presented candidates is ≥ 0.9. Right: Non-nodule objects inside-lungs mistakenly detected by our method method with high confidence. Each patch is a representative maximum through three sequential horizontal planes. P-values for all presented candidates is ≥ 0.9.

Chapter 4
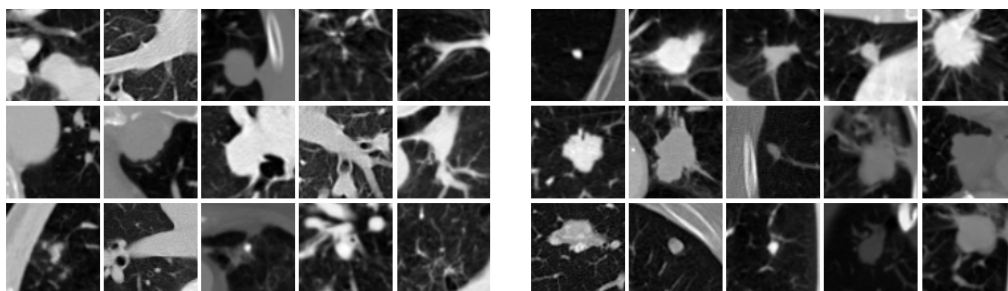
DISCUSSION

In this work a novel method aimed at false positive reduction in pulmonary nodule detection was described in detail along with results overview. Compared to the state of the art approaches our method show comparable or better performance.

The focus was made only on the false positive reduction instead of developing methods both for candidate detection and a false positive reduction. That means that this implementation is independent of how the candidate detection is done and can be used in combination with any candidate detection method. But it is also important to notice that the accuracy depends on candidate detection quality.

Due to the nature of CPM score it's sensitivities at the first four false positives per scan is highly susceptible to noise thus some ensemble techniques could make it less unstable and tend to converges to some $O(\text{FPRPS})$. Such behaviour of CPM score was shown by FROC curves in Figure 3.1.

As was showed in the study of Niemeijer et al., 2011 it is a good practice to combine different CAD systems in order to increase the performance because different CAD systems may have different strengths and weaknesses. So in the work of DIAG CON-VNET(arnaud.setio), iitm03(subru1603), LUNA16CAD(hirokinakano), LungNess(bim bam), despite of the benefits provided by fusing information extracted from different views and reducing the model's complexity, such fusion strategy does not take benefit of the nature of nodule appearance as a volumetric spatial information.

At the same time, in the works of LUNA16CAD(mattdns100689) 3D space is covered via receptive fields selected to be cubes with shape (42, 42, 42) acting over input taken at different scales. But due to the proprieties of the measure concentration in high dimension euclidean space (even in three dimension) the major part of the cube volume results in addressing the surroundings of a nodule instead of the nodule itself, e.g., in unit space, when nodule's diameter is less than 12mm ($> 80\%$ of the whole annotated nodules) the fraction between nodule and its surroundings is about $\frac{1}{100}$. Like LUNA16CAD(mattdns100689), CUMedVis(QiDou) uses pretty similar approach, with a slight difference in that they employed specific input's size and kernel's size for each dimension (this is caused by the difference in resolution

along dimensions) and they also use full ten-fold cross validation instead of the two-fold used by LUNA16CAD(mattdns100689). In contrast, our approach handles the problem of redundant contextual information by combining dimensionality reduction strategy with spatial high level features merging technique.

Figure 3.2 presents the high confidence examples of true positives detection made by our system. For these examples, their morphological appearance is highly similar to the false positives shown on the same figure. In many LUNA16 participants' works there were a lot of cases similar to ones described above. As mentioned earlier, the solution for dealing with such challenging cases may be lies in the combination of methods.

4.1 Conclusion

In this study a deep residual 3D convolutional neural network architecture aimed at feature extraction was described along with merging and dimensionality reduction strategies. The results show that proposed method is competitive and can be further proposed to use in combination with other methods to improve the detection performance and lower the risk of late cancer detection. Methods presented in this paper can be further extended to work with other kinds of 3D medical images in different tasks.

References

"3rd Annual Congress of the Russian Breast Cancer Society." In:

Abadi, Martın et al. (2016). "Tensorflow: Large-scale machine learning on heterogeneous distributed systems". In: *arXiv preprint arXiv:1603.04467*.

"American Cancer Society, 2016. Cancer facts and figures 2016." In:

Boureau, Y-Lan et al. (2010). "Learning mid-level features for recognition". In: *Computer Vision and Pattern Recognition (CVPR), 2010 IEEE Conference on*. IEEE, pp. 2559–2566.

Brooks, Rodney A (1977). "A quantitative theory of the Hounsfield unit and its application to dual energy scanning." In: *Journal of computer assisted tomography* 1.4, pp. 487–493.

Brown, Matthew S et al. (1997). "Method for segmenting chest CT image data using an anatomical model: preliminary results". In: *IEEE transactions on medical imaging* 16.6, pp. 828–839.

Chawla, N, Nathalie Japkowicz, and A Kolcz (2004). "Special issue on learning from imbalanced datasets, sigkdd explorations". In: *ACM SIGKDD*.

Chen, Hao et al. (2016). "VoxResNet: Deep Voxelwise Residual Networks for Volumetric Brain Segmentation". In: *CoRR* abs/1608.05895. URL: `http://arxiv.org/abs/1608.05895`.

Chollet, François. "Keras (2015)". In: *URL http://keras. io*.

Dou, Q. et al. (2016). "Multi-level Contextual 3D CNNs for False Positive Reduction in Pulmonary Nodule Detection". In: *IEEE Transactions on Biomedical Engineering* PP.99, pp. 1–1. ISSN: 0018-9294. DOI: `10.1109/TBME.2016.2613502`.

Efron, Bradley and Robert J Tibshirani (1994). *An introduction to the bootstrap*. CRC press.

"European Consensus Statement on Lung Cancer: Risk factors and prevention. Lung Cancer Panel." In:

Firmino, Macedo et al. (2014). "Computer-aided detection system for lung cancer in computed tomography scans: review and future prospects". In: *Biomedical engineering online* 13.1, p. 41.

Glorot, Xavier and Yoshua Bengio (2010). "Understanding the difficulty of training deep feedforward neural networks." In: *Aistats*. Vol. 9, pp. 249–256.

He, Kaiming et al. (2015). "Deep Residual Learning for Image Recognition". In: *CoRR* abs/1512.03385. URL: `http://arxiv.org/abs/1512.03385`.

Hedlund, LW et al. (1982). "Two methods for isolating the lung area of a CT scan for density information." In: *Radiology* 144.2, pp. 353–357.

Hoffman, ERIC A (1985). "Effect of body orientation on regional lung expansion: a computed tomographic approach". In: *Journal of Applied Physiology* 59.2, pp. 468–480.

Hoffman, ERIC A and Erik L Ritman (1985). "Effect of body orientation on regional lung expansion in dog and sloth". In: *Journal of Applied Physiology* 59.2, pp. 481–491.

Hoffman, ERIC A, LJ Sinak, et al. (1983). "Noninvasive quantitative imaging of shape and volume of lungs". In: *Journal of Applied Physiology* 54.5, pp. 1414–1421.

Hu, Shiying, Eric A Hoffman, and Joseph M Reinhardt (2001). "Automatic lung segmentation for accurate quantitation of volumetric X-ray CT images". In: *IEEE transactions on medical imaging* 20.6, pp. 490–498.

Ioffe, Sergey and Christian Szegedy (2015). "Batch normalization: Accelerating deep network training by reducing internal covariate shift". In: *arXiv preprint arXiv:1502.03167*.

Jacobs, Colin et al. (XXXX). "Automatic detection of subsolid pulmonary nodules in thoracic computed tomography images". In: *Medical Image Analysis* 18.2, pp. 374–384. ISSN: 1361-8415. DOI: 10.1016/j.media.2013.12.001. URL: http://dx.doi.org/10.1016/j.media.2013.12.001.

Kalender, Willi A et al. (1991). "Semiautomatic evaluation procedures for quantitative CT of the lung." In: *Journal of computer assisted tomography* 15.2, pp. 248–255.

Keller, James M, F Marc Edwards, and Robert Rundle (1981). "Automatic outlining of regions on CT scans." In: *Journal of Computer Assisted Tomography* 5.2, pp. 240–245.

Kingma, Diederik P. and Jimmy Ba (2014). "Adam: A Method for Stochastic Optimization". In: *CoRR* abs/1412.6980. URL: http://arxiv.org/abs/1412.6980.

Li, Qiang, Shusuke Sone, and Kunio Doi (2003). "Selective enhancement filters for nodules, vessels, and airway walls in two- and three-dimensional CT scans". In: *Medical Physics* 30.8, pp. 2040–2051. ISSN: 2473-4209. DOI: 10.1118/1.1581411. URL: http://dx.doi.org/10.1118/1.1581411.

Murphy, K. et al. (XXXX). "A large-scale evaluation of automatic pulmonary nodule detection in chest CT using local image features and k-nearest-neighbour classification". In: *Medical Image Analysis* 13.5, pp. 757–770. ISSN: 1361-8415. DOI: 10.1016/j.media.2009.07.001. URL: http://dx.doi.org/10.1016/j.media.2009.07.001.

Niemeijer, M. et al. (2011). "On Combining Computer-Aided Detection Systems". In: *IEEE Transactions on Medical Imaging* 30.2, pp. 215–223. ISSN: 0278-0062. DOI: 10.1109/TMI.2010.2072789.

Radiation Units, International Commission on and Measurements (2008). "Receiver operating characteristic analysis in medical imaging." In:

Rikxoort, Eva M van et al. (2009). "Automatic lung segmentation from thoracic computed tomography scans using a hybrid approach with error detection". In: *Medical physics* 36.7, pp. 2934–2947.

Setio, A. A. A. et al. (2016). "Pulmonary Nodule Detection in CT Images: False Positive Reduction Using Multi-View Convolutional Networks". In: *IEEE Transactions on Medical Imaging* 35.5, pp. 1160–1169. ISSN: 0278-0062. DOI: `10.1109/TMI.2016.2536809`.

Setio, Arnaud A. A. et al. (2015). "Automatic detection of large pulmonary solid nodules in thoracic CT images". In: *Medical Physics* 42.10, pp. 5642–5653. ISSN: 2473-4209. DOI: `10.1118/1.4929562`. URL: `http://dx.doi.org/10.1118/1.4929562`.

Simonyan, Karen and Andrew Zisserman (2014). "Very deep convolutional networks for large-scale image recognition". In: *arXiv preprint arXiv:1409.1556*.

Tan, Maxine et al. (2011). "A novel computer-aided lung nodule detection system for CT images". In: *Medical Physics* 38.10, pp. 5630–5645. ISSN: 2473-4209. DOI: `10.1118/1.3633941`. URL: `http://dx.doi.org/10.1118/1.3633941`.

Van Ginneken, Bram, Mikkel B Stegmann, and Marco Loog (2006). "Segmentation of anatomical structures in chest radiographs using supervised methods: a comparative study on a public database". In: *Medical image analysis* 10.1, pp. 19–40.

Appendix A

QUESTIONNAIRE

Appendix B

CONSENT FORM

[1]Endnotes are notes that you can use to explain text in a document.

POCKET MATERIAL: THE HU OF COMMON SUBSTANCES.

Substance	HU
Air	1000
Lung	500
Fat	120 to 90
Water	0
Urine	-5 to +15
Bile	-5 to +15
CSF	+15
Kidney	+20 to +45
Blood	+30 to +45
Muscle	+35 to +55
Grey matter	+37 to +45
White matter	+20 to +30
Liver	+40 to +60
Soft Tissue, Contrast	+100 to +300
Bone	+700 (cancellous bone) to +3000 (cortical bone)
Chyle	30