[1]Automatic Pinhole Image Subtraction using Artificial Intelligence for Parathyroid Glands Localization

Ouassim Boukhennoufa ⋆, Laurent Comas ‡¶, Jean-Marc Nicod ⋆
Constantin Ungureanu ‡, *Noureddine Zerhouni* ⋆ and *Hatem Boulahdour* ‡¶
⋆ SUPMICROTECH, CNRS, institut FEMTO-ST, Besançon, France
‡ CHU Besançon, Médecine nucléaire, Besançon, France
¶ Université de Franche-Comté, LNIT, Besançon, France

Abstract

Primary hyperparathyroidism usually involves surgery. In this situation, pre-operative localization of aberrant Parathyroid Glands (PGs) is crucial. One of the existing approaches, Nuclear Medicine is the widely used procedure. It provides PINHOLE and Low Energy High-Resolution (LEHR) images through the injection of two radio-elements, technetium (^{99m}TC-MIBI) and iodine (^{123}I). After that, the physicians perform a manual normalization and subtraction of both images for PGs detection which is time-consuming. This paper proposes an automatic detection of the PGs by the combination of statistical normalization and artificial intelligence for the subtraction of ^{99m}TC-MIBI and iodine ^{123}I images. The proposed methodology is applied to PINHOLE images from 88 retrospective, single-center studies. The obtained results achieved a mean correlation of 0.95 compared to the physician results.

1 Introduction

Primary HyperParaThyroidism (PHPT) is a common endocrine disease that involves high levels of parathyroid hormone and calcium [1]. The typical size of Parathyroid Glands (PGs) is 3 – 4 mm, and most imaging techniques are unable to detect them [2]. The location of (PGs) is not generalized across all patients, they can be found anywhere between the upper mediastinum and the mandibular angle [2]. Neck ultrasound and dual-tracer or dual-phase parathyroid scintigraphy are complementary techniques used to diagnose parathyroid adenomas pre-operatively [3]. Nuclear medicine (NM) improves diagnosis accuracy for physicians [4].

NM for PGs detection is an imaging technique using dual tracer parathyroid scintigraphy which consists of injecting ^{123}I (iodine) which fixes the thyroid cells, and 3 hours later injecting ^{99m}Tc-MIBI which fixes both thyroid and parathyroid cells. Then, after 5 minutes, PINHOLE and LEHR images of both ^{123}I and ^{99m}Tc-MIBI injections are collected using a gamma camera system.

Once the images are stored, physicians start extracting the Region Of Interest (ROI), which is the thyroid, from the ^{123}I image [5]. After the ROI definition, a subtraction between the ^{99m}Tc-MIBI and ^{123}I images is performed which results in a unique final image containing only the PGs [5, 6]. In general,

[1]Published in the International Symposium on Biomedical Imaging

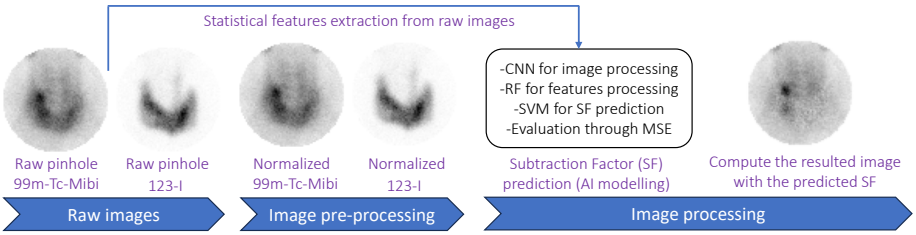

Figure 1: Proposed methodology.

the subtracted image is not optimal and is adjusted until the physician considers that the PG is detected. This process is 100% manual, and time-consuming. ROI and contours are usually obtained by thresholding at 50% of the maximum pixel from ^{123}I image. The obtained contour is applied to the ^{99m}Tc-MIBI image. Moreover, a normalization factor is computed to normalize the images. Finally, the subtraction is performed and adjusted [5, 6, 7].

In this paper, a fully automatic process for PINHOLE image subtraction using statistical normalization and Artificial Intelligence (AI) to predict a Subtraction Factor (SF) is proposed. The output of the methodology consists of a subtracted image provided to the physician for diagnosis.

The remainder of the paper is organized as follows: Section 2 details our contribution. Section 3 specifies the case study, presents and discusses the results. Section 4 is the conclusion of this work.

2 Proposed methodology

This section presents the steps of the proposed methodology for automatic PGs detection. First, image pre-processing for image normalization is presented in Section 2.1 and investigates three techniques. Then Section 2.2 details the introduction of AI for PINHOLE image subtraction.

2.1 Image pre-processing

The aim of this step is to normalize the images, it consists of adjusting the pixel values of an image to a standardized range. In our case, the range belongs to [0,1]. This helps ensure that the images have consistent intensity. We propose three different normalization techniques: local, global, and hybrid normalization. In each technique, we apply three image normalization methods: Z-Score, Power Law, and MinMax [8].

2.1.1 Local normalization

In this technique, the objective is to normalize every image locally. For each patient's image, the image is normalized according to its particular characteristics.

Equations 1, 2, 3 represent Z-Score, Power Law, and MinMax, respectively:

$$LNI_{Z-Score} = \frac{RI_i - \mu(RI_i)}{\sigma(RI_i)} \tag{1}$$

$$LNI_{PowerLaw} = \left(\frac{RI_i}{\max(RI_i)}\right)^{\gamma} \tag{2}$$

$$LNI_{MinMax} = \frac{RI_i - \min(RI_i)}{\max(RI_i) - \min(RI_i)} \tag{3}$$

where i either the ^{99m}Tc-MIBI or the ^{123}I image. LNI the Local Normalized Image, RI_i the patient's raw image, $\mu(RI_i)$, and $\sigma(RI_i)$ the average and the standard deviation of the raw image respectively. γ aims at improving contrast, its value 0.5, found empirically $\max(RI_i)$ and $\min(RI_i)$ the maximum and minimum pixel values of the raw image.

2.1.2 Global normalization

In this technique, the goal is to normalize every image globally, according to all patient's images. The equations of Z-Score, Power Law, and MinMax become as follows:

$$GNI_{Z-Score} = \frac{RI_i - \mu(RI_1, ...RI_n)}{\sigma(RI_1, ...RI_n)} \tag{4}$$

$$GNI_{PowerLaw} = \left(\frac{RI_i}{\max(RI_1, ...RI_n)}\right)^{\gamma} \tag{5}$$

$$GNI_{MinMax} = \frac{RI_i - \min(RI_1, ...RI_n)}{\max(RI_1, ...RI_n) - \min(RI_1, ...RI_n)} \tag{6}$$

where GNI the Global Normalized Image and n the number of patients.

2.1.3 Hybrid normalization

After applying local and global normalization, we use the normalized images to compute a hybrid normalized image. We attribute two weights: w_1 and w_2 to LNI and GNI respectively. $w_1 + w_2 = 1$. The higher weight is attributed to the method that performs better. Equation 7 summarizes the hybrid normalization.

$$HNI = w_1 \times LNI + w_2 \times GNI \tag{7}$$

where HNI the Hybrid Normalized Image, $w1$ the weight attributed to the local normalized image and $w2$ the one attributed to the global normalized image.

2.2 Image processing

In this step, we aim to apply AI to predict a Subtraction Factor (SF). We note that for each patient an SF is given by the senior physician who subtracted the images manually until now. The predicted SF is used to compute the final subtracted image as shown in Equation 8.

$$SI = GNI_{MinMax_j} - SF' \times GNI_{MinMax_k} \tag{8}$$

where SI the subtracted image, SF' the predicted subtraction factor, GNI the global normalized image with MinMax. The indices j and k reference respectively the ^{99m}Tc-MIBI and ^{123}I images.

From Equation 8, SF' is predicted using AI models, and for this purpose we build one AI model with different architectures. We propose first an approach based on the Convolutional Neural Network (CNN) and the Random Forest (RF) respectively that aim to automatically and adaptively learn hierarchical feature representations directly from pixel data and learn for statistical features respectively. We extract the needed features from the raw images, meaning that from each ^{99m}Tc-MIBI and ^{123}I image we compute kurtosis, entropy, peaktopeack, and sum of pixels. The aim of adding statistical features to the AI model is to differentiate between patient images. After image and statistical features processing, we combine the features extracted by CNN and RF to train a Support Vector Machine (SVM) for the regression task. The AI model architecture is represented in Table 1.

Model	Architecture
CNN	- Conv2D Layer with 32 filters of size 3 - MaxPooling2D Layer of size 2 - Conv2D Layer with 32 filters of size 3 - MaxPooling2D Layer of size 2 - Fully Connected Layer of 128 neurons
RF	Nbr of estimators = 100, criterion = MSE Max depth = None, Min samples split = 2 Bootstrap = True, Max features = 1
SVM	Kernel = rbf, degree = 3, C = 1

Table 1: AI model architecture for SF prediction.

The parameters in Table 1 are obtained after applying hyper-parameters tuning using **random searchcv** with total epochs of 200, a batch size of 5, and a learning rate of 0.01.

2.2.1 Train and test split

We split our data into 5 different splits: starting with 50% for both train and test; to 90% for train and 10% for test. We augment the size of the train by 10% in each split. This gives an idea of how the model is learning, what amount of data it needs to learn, and if there is a problem with the dataset.

3 Application and results

In this section, we describe the case study and present the obtained results with discussions in subsections 3.1, 3.2.

3.1 Case study

88 successive patients addressed for parathyroid adenoma were retrospectively included in this study (examinations between April 2022 and October 2023). 88 PINHOLE ^{99m}Tc-MIBI and 88 PINHOLE ^{123}I images. The image size is 128×128 with gray-scale coloring. Figure 2 shows an example of the images. The left part is a PINHOLE ^{123}I image. The right part is a PINHOLE ^{99m}Tc-MIBI image.

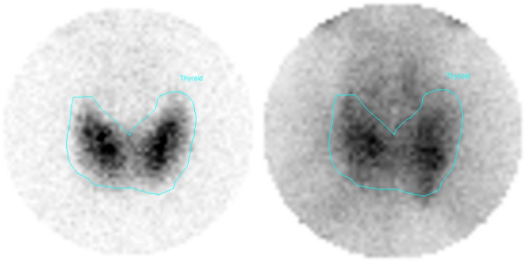

Figure 2: Example of images used in the study.

The mean patient age is 61 years [21-91 years]. 62 female patients [70.45% female, 29.55% male]. Table 2 shows information on the dataset that we used.

Patients number	Images number	Body part	Patients age	Patients weight
88	176	neck	28-91	49-125kg
Patients size	**Patients gender**	**Images type**	**Acquisition dates**	**Injections time**
1.48 - 1.86m	62 F 26 M	static planar	04/2022 10/2023	0h: ^{123}I 3h: ^{99m}Tc

Table 2: Metadata related to the dataset of the study.

3.2 Results and discussion

To evaluate the proposed normalizations, we use the Mean Squared Error (MSE) metric as shown in Equation 9. The objective is to minimize the MSE.

$$MSE = \mu((LNI_{^{99m}TC-MIBI} - GNI_{^{123}I})^2) \tag{9}$$

For hybrid normalization, an empirical test is applied by combining weights. The shown results (see Table 3) are the combination with $w_1 = 0.2$ and $w_2 = 0.8$ for hybrid normalization which gives the best (lowest) MSE.

Technique	Z-score	Power law	MinMax
Local	417	19	**9**
Global	72	15	**4**
Hybrid	74	17	**5**

Table 3: MSE of the three different normalization techniques.

We use the combination of global MinMax normalization since it gives the lowest MSE (see Table 3). To evaluate AI model performance we use the MSE metric, since the problem is a regression task. We apply it during training and test. SF is the actual subtraction factor made by the physician.

$$MSE_{test} = \frac{1}{n}\sum_{z=1}^{n}\left(SF_z - SF'_z\right)^2 \tag{10}$$

We also use the Correlation (Corr) metric. After computing the subtracted image with Equation 8, we measure the linear relationship of the corresponding pixels in SI with the corresponding pixel in the Manual Subtracted Image (MSI) by the physician.

$$Corr_n = \frac{\sum_{m=0}^{p}(SI_m - SI)(MSI_m - MSI)}{\sqrt{\sum_{m=0}^{p}(SI_m - SI)^2 \sum_{m=0}^{p}(MSI_m - MSI)^2}} \tag{11}$$

where p the number of pixels, m the position of the corresponding pixel in our subtracted image (SI) and the one generated by the physician (MSI). During AI model training, the MSE achieves 0%, our interest is in the MSE during the test phase. Figure 3 shows the MSE values with different train sets. As explained in subsection 2.2.1 we train our model with different train test split. We observe that the model is learning more by increasing the amount of data in the training phase, but also the test MSE achieves 8% with 90% (78 patients) for the training and 10% (10 patients) for the test.

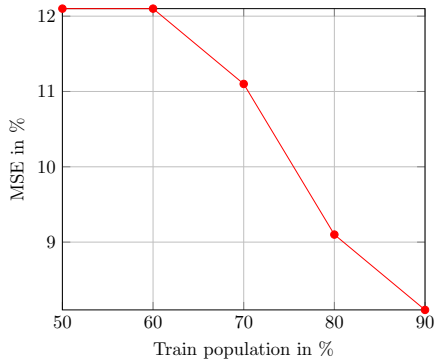

Figure 3: Test MSE values with different train populations.

To test the robustness of the model, we make sure to have different sets of data by making 5 shuffles with a split of 90% and 10% for train and test respectively. It ensures that in each split we get different sets that are fed to the model. The obtained MSE values in Table 4 confirm the model's robustness and its ability to predict the SF with different sets accurately.

Shuffle	1	2	3	4	5
Testset MSE (%)	8.06	8.11	8.16	8.46	8.24

Table 4: MSE of test set.

Table 5 highlights the Corr values of patients judged doubtful cases (the physicians couldn't make the diagnosis). The values are a comparison between our SI and the one made manually by the physicians (MSI). The values show that the model can predict accurate SF for difficult cases which generates identical SI and MSI. We note that the mean Corr of 20% and 10% test sets achieved 0.95 [min:0.836, max:0.994].

Patient	1	2	3	4	5
Correlation	0.992	0.911	0.981	0.923	0.885

Table 5: Correlation with respect to the physician subtraction.

Figure 4 shows an example of a manually subtracted image (left) and our subtracted image (right) with Corr = 0.973.

Figure 4: Example of image subtraction comparison.

4 Conclusion

This paper presents a fully automatic PINHOLE image subtraction process. First, PINHOLE images of ^{99m}Tc and ^{123}I are normalized using the global technique which gives the lowest MSE. For the processing step, the normalized images are injected into a CNN model whereas statistical features are fed to a RF model, and then an SVM to predict the subtraction factor. This latter is used to generate a Subtracted Image (SI). AI shows promising results, achieving an MSE of 8%. Finally, the obtained SI is compared to the one performed by senior physicians. The average correlation is 0.95. This automatic methodology gives good results PGs detection and can be provided to physicians for decision-making.

5 Compliance with ethical standards

This study was registered by the Clinical Research and Innovation Delegation of the University Hospital Center of Besançon under the number 2023/796.

6 Acknowledgment

This work has been supported by the EIPHI graduate school (contract "ANR-17-EURE-0002").

References

[1] Walker Marcella D et al., "Primary hyperparathyroidism," *NRE*, vol. 14, no. 2, pp. 115–125, 2018.

[2] Sung Jin Yong et al., "Parathyroid ultrasonography: the evolving role of the radiologist," *Ultrasonography*, vol. 34, no. 4, pp. 268, 2015.

[3] Lee Sang-Woo et al., "Direct comparison of preoperative imaging modalities for localization of primary hyperparathyroidism: a systematic review and network meta-analysis," *JAMA OHNS*, vol. 147, pp. 692–706, 2021.

[4] Nieciecki Michał et al., "The role of ultrasound and nuclear medicine methods in the preoperative diagnostics of primary hyperparathyroidism," *Journal of Ultrasonography*, vol. 15, no. 63, pp. 398–409, 2015.

[5] S. Hassler et al., "Scintigraphie des parathyroïdes double isotope 99mtc-mibi/123i dans l'hyperparathyroïdisme primaire : comparaison entre un protocole d'acquisition planaire et une temp de soustraction couplée à la tdm," *Médecine Nucléaire*, vol. 35, no. 3, pp. 105–116, 2011.

[6] S. Hassler et al., "Dual-isotope 99mtc-mibi/123i parathyroid scintigraphy in primary hyperparathyroidism: comparison of subtraction spect/ct and pinhole planar scan," *CNM*, vol. 39, no. 1, pp. 32–36, 2014.

[7] Petranović Ovčariček Petra et al., "The eanm practice guidelines for parathyroid imaging," *EJNMMI*, vol. 48, pp. 2801–2822, 2021.

[8] Patro SGOPAL et al., "Normalization: A preprocessing stage," *arXiv preprint arXiv:1503.06462*, 2015.