

الجمهورية الجزائرية الديمقراطية الشعبية

DEMOCRATIC AND POPULAR ALGERIAN REPUBLIC

وزارة التعليم العالي والبحث العلمي

Ministry of Higher Education and Scientific Research

جامعة ابو بكر بلقايد - تلمسان

University Abou Bakr Belkaïd Tlemcen -Faculty of TECHNOLOGY

Department of Biomedical Engineering



Final Year Project Thesis

To Obtain The Master's Degree

IN BIOMEDICAL ENGINEERING

Specialty: Medical Imaging

Presented By: Ahmed Benchaib Bouchra

Subject

**Artificial Intelligence-Based Diagnostic Assistance System for
Thyroid Nodules in Ultrasound Imaging**

Supported The 10/06/2026 Before The Jury

M.	Hadj Slimane Zine Eddine	Professor	University of Tlemcen	President
M.	Djebbari Abdelghani	Professor	University of Tlemcen	Examiner
M.	BENGANA Abdelfatih	MCA	University of Tlemcen	Supervisor
M.	BOUKLI HACENE Ismail	Professor	University of Tlemcen	Co-supervisor

Academic year 2025-2026

FOREWORD

In recent years, advances in medical imaging have significantly improved the detection and diagnosis of diseases. Ultrasound imaging, in particular, has become an essential, non-invasive tool for the evaluation of thyroid nodules, whose increasing prevalence requires accurate and early diagnosis.

However, ultrasound interpretation remains highly dependent on clinician expertise and may be subject to variability. This limitation highlights the need for intelligent systems capable of providing more consistent and reliable diagnostic support.

In this context, artificial intelligence, especially machine learning and deep learning, offers promising solutions for enhancing medical image analysis. This project, carried out under the guidance of Mr. BENGANA Abdelfatih and Mr. BOUKLI HACENE Ismail, aims to develop a computer-aided diagnosis system for the segmentation and classification of thyroid nodules in ultrasound images.

By combining advanced image processing techniques with AI models, this work seeks to improve diagnostic performance and contribute to more effective clinical decision-making, while strengthening the link between engineering and medicine.

ACKNOWLEDGMENTS

First of all, I would like to thank Allah, who granted me the health, willpower, and patience necessary to complete this work.

*I wish to express my deepest gratitude to Mr. **BENGANA Abdelfatih**, my supervisor, for his constant guidance, patience, and invaluable support throughout this project. His encouragement and confidence were essential in helping me successfully complete this work.*

*I am also sincerely grateful to Mr. **BOUKLI HACENE Ismail**, my co-supervisor, for his availability, valuable advice, and constructive feedback, which greatly contributed to the improvement of this thesis.*

*I would like to sincerely thank the members of the jury for the time and attention they dedicated to evaluating this work. I express my deepest gratitude to **Professor Zine Eddine Hadj Slimane**, who has honored me by presiding over this jury, and to **Professor Abdelghani Djebbari**, who agreed to examine this work with care and rigor.*

I would also like to thank all the teachers and staff of the Faculty of Technology at Abou Bakr Belkaid University of Tlemcen for their support and everything they have taught me throughout my academic journey.

*Most importantly, I would like to express my sincerest gratitude to my beloved parents, **Djamel and Houria**. Their unconditional love, sacrifices, patience, and constant prayers have been my greatest source of strength. This success would not have been possible without their support, encouragement, and belief in me. I will always be deeply grateful for everything they have done for me.*

DEDICATION

(وَأَخِرُ دَعْوَاهُمْ أَنْ الْحَمْدُ لِلَّهِ رَبِّ الْعَالَمِينَ)

الحمد لله الذي ما تم جهد ولا خُتِم سعي إلا بفضلِهِ، الحمد لله على البلوغ ثم الحمد لله على التمام

AT THE END OF THIS ACADEMIC JOURNEY, I DEDICATE THIS THESIS TO THOSE WHO ILLUMINATED MY DARKEST HOURS, STRENGTHENED MY RESOLVE IN MOMENTS OF UNCERTAINTY, AND SERVED AS MY COMPASS IN TIMES OF DOUBT.

I DEDICATE THIS WORK, FIRST AND FOREMOST, TO MYSELF, **AHMED BENCHAI B BOUCHRA**, IN RECOGNITION OF MY PERSEVERANCE, UNWAVERING COMMITMENT, AND STEADFAST DESIRE TO SEE THIS SCIENTIFIC ADVENTURE THROUGH TO THE END.

TO MY BELOVED PARENTS, **HOURLIA AND DJAMEL**, WHOSE UNCONDITIONAL LOVE, SILENT SACRIFICES, AND CONSTANT SUPPORT HAVE FORMED THE UNSHAKEABLE FOUNDATION OF MY JOURNEY. YOUR PRAYERS, YOUR TRUST, AND YOUR PRESENCE HAVE BEEN MY MOST PRECIOUS STRENGTH. THIS WORK BELONGS TO YOU AS MUCH AS IT DOES TO ME.

TO MY DEAR BROTHERS, **ABDE NOUR AND ABDEL KADER**, FOR THEIR SINCERE AFFECTION, THEIR COMFORTING PRESENCE, AND THEIR ENCOURAGEMENT THAT HAVE ACCOMPANIED ME THROUGHOUT THIS JOURNEY. YOUR BROTHERLY SUPPORT HAS BEEN AN INEXHAUSTIBLE SOURCE OF STRENGTH AND MOTIVATION.

TO MY EXTENDED FAMILY ON MY MOTHER'S SIDE, TO EACH AND EVERY ONE OF ITS MEMBERS, FOR THEIR WARMTH, THEIR KINDNESS, AND THEIR PRAYERS THAT HAVE SURROUNDED EACH STEP OF MY JOURNEY. YOUR FAMILIAL LOVE HAS ALWAYS BEEN A PRECIOUS REFUGE.

TO **MR. BENGANA ABDELFATH**, MY SUPERVISOR, AND **MR. BOUKLI HACENE ISMAIL**, MY CO-SUPERVISOR, FOR THEIR INVALUABLE SUPPORT, WISE ADVICE, AND EXEMPLARY AVAILABILITY. THEIR EXPERTISE AND COMMITMENT WERE CRUCIAL TO THE COMPLETION OF THIS PROJECT.

FINALLY, TO ALL THE KIND SOULS WHO CROSSED MY PATH AND SOWED SEEDS OF HOPE AND ENCOURAGEMENT, I EXPRESS MY DEEPEST GRATITUDE. EVERY GESTURE, EVERY WORD CONTRIBUTED TO THIS ACHIEVEMENT.

ABSTRACT

Thyroid cancer is one of the most common malignant endocrine diseases, and its early diagnosis remains a major challenge in medical imaging. Although thyroid ultrasound is the reference examination due to its non-invasive and accessible nature, its interpretation is affected by low image contrast, speckle noise, nodule variability, and operator dependency. To address these limitations, this work proposes an automatic thyroid nodule analysis system based on **Artificial Intelligence (AI)** and integrated into a **Computer-Aided Diagnosis (CAD)** framework. The proposed approach combines nodule segmentation and benign/malignant classification from ultrasound images. Experiments were conducted on public thyroid ultrasound datasets using preprocessing, data augmentation, and optimization techniques. The obtained results achieved a Dice Score of 85.44% and an IoU of 75.45% for segmentation, while classification reached an accuracy of 97.03% and an AUC of 99.05%. These results demonstrate the effectiveness of the proposed system and its potential to support clinical decision-making in thyroid cancer diagnosis.

Keywords: Thyroid nodules, ultrasound imaging, , segmentation, classification, deep learning.

RESUME

Le cancer de la thyroïde constitue l'une des pathologies endocriniennes malignes les plus fréquentes, dont le diagnostic précoce représente un enjeu majeur en imagerie médicale. Bien que l'échographie thyroïdienne soit l'examen de référence grâce à son caractère non invasif et accessible, son interprétation reste limitée par le faible contraste des images, le bruit speckle, la variabilité des nodules et la subjectivité des opérateurs. Pour répondre à ces limites, ce travail propose un système automatique d'analyse des nodules thyroïdiens basé sur l'**Intelligence Artificielle (IA)** et intégré dans un cadre d'**Aide au Diagnostic Assistée par Ordinateur (Computer-Aided Diagnosis, CAD)**. L'approche développée combine la segmentation des nodules et leur classification en bénins ou malins à partir d'images échographiques. Les expérimentations menées sur des bases publiques, avec des techniques de prétraitement, d'augmentation de données et d'optimisation, ont permis d'obtenir un Dice Score de 85,44 % et une IoU de 75,45 % en segmentation, ainsi qu'une accuracy de 97,03 % et une AUC de 99,05 % en classification. Ces résultats confirment l'efficacité du système proposé et son potentiel pour l'assistance à la décision clinique dans le diagnostic du cancer thyroïdien.

Mots-clés : Nodules thyroïdiens, échographie, segmentation, classification, apprentissage profond.

الملخص

يُعد سرطان الغدة الدرقية أحد أكثر الأمراض الخبيثة في الجهاز الغدد الصماء شيوعًا، ويُعد التشخيص المبكر له تحديًا كبيرًا في مجال التصوير الطبي. على الرغم من أن الموجات فوق الصوتية للغدة الدرقية هي الفحص المعياري بفضل طبيعتها غير الغازية وسهولة الوصول إليها، إلا أن تفسيرها يظل محدودًا بسبب ضعف تباين الصور، وضوضاء البقع، وتباين العقد، وذاتية القائمين على الفحص. للتغلب على هذه القيود، يقدم هذا العمل نظامًا آليًا لتحليل العقيدات الدرقية يعتمد على الذكاء الاصطناعي ويجمع النهج المطور بين تجزئة العقيدات وتصنيفها إلى (CAD) ويندرج ضمن إطار التشخيص بمساعدة الكمبيوتر (AI) حميدة أو خبيثة استنادًا إلى صور الموجات فوق الصوتية. وقد سمحت التجارب التي أجريت على قواعد بيانات عامة، بنسبة 85,44% ومعدل DICE مما أدى إلى الحصول على معدل باستخدام تقنيات المعالجة المسبقة وزيادة البيانات والتحسين، بنسبة 99,05% (AUC) بنسبة 75,45% في عملية التجزئة، بالإضافة إلى دقة بنسبة 97,03% ومنطقة تحت المنحنى IoU بنسبة 99,05% في عملية التصنيف. تؤكد هذه النتائج فعالية النظام المقترح وإمكاناته في دعم اتخاذ القرار السريري لتشخيص سرطان الغدة الدرقية.

الكلمات المفتاحية: عقيدات الغدة الدرقية، الموجات فوق الصوتية، التجزئة، التصنيف، التعلم العميق.

SUMMARY:

FOREWORD	i
ACKNOWLEDGMENTS.....	ii
DEDICATION.....	iii
ABSTRACT	iv
RESUME	v
الملخص	vi
LIST OF FIGURES	vii
LIST OF TABLES	viii
LIST OF ABBREVIATIONS	ix
GENERAL INTRODUCTION.....	1
CHAPTER I: General information about the thyroid gland and thyroid nodules.....	4
I.1 Introduction:	5
I.2 Anatomy and physiology of the thyroid gland:	5
I.2.1 Gross Anatomy and Vascularization:.....	5
I.2.2 Functional histology (follicles, C cells):	9
I.2.3 Endocrine physiology: T3, T4, calcitonin:	10
I.2.4 Lymphatic system and innervation:	11
I.3 Thyroid nodules:	13
I.3.1 Definition, epidemiology, and risk factors:.....	13
I.3.2 Types of nodules: benign, malignant, solitary vs. multinodular:	17
I.3.2.1 Solitary vs. multinodular nodules:	17
I.3.2.2 Benign nodules (cysts, adenomas):	17
I.3.2.3 Malignant nodules (thyroid cancer):.....	18
I.3.3 TI-RADS classification (EU TIRADS, ACR TI RADS):.....	20
I.3.4 Clinical appearance of thyroid nodules:	23
I.3.4.1 Medical history taking:	23

I.3.4.2 Signs of thyroid disorder:.....	23
I.3.4.3 Signs of local compression:	27
I.4 Diagnosis and management of thyroid nodules:	28
I.4.1 Additional exams: biology, ultrasound, cytology, scintigraphy	30
I.4.1.1 Biological balance:	30
I.4.1.2 Thyroid Ultrasound:	32
I.4.1.3 Cytology/Fine Needle Aspiration (FNA):.....	33
I.4.1.4 Scintigraphy:	34
I.4.1.5 Other Exams:	35
I.4.2 Operative indications:	37
I.4.2.1 Malignant nodules:	37
I.4.2.2 Symptomatic/compressive nodules:	38
I.4.2.3 Hyperfunctional nodules resistant to medical treatment:.....	39
I.4.2.4 Other specific indications:.....	39
I.4.3 Surgical treatment:	40
I.4.3.1 Types of surgical procedures:	40
I.4.3.2 Conduct of the operative act:.....	45
I.4.3.3 Peri-operative and post-operative complications:	46
I.4.3.4 Post-operative care:	49
I.4.3.5 Objectives of the surgical treatment:	50
I.4.4 Conclusion of chapter I :	51
CHAPTER II : ULTRASOUND IMAGING AND COMPUTER-AIDED DIAGNOSIS.....	52
II.1 Introduction:	53
II.2 Basics of ultrasound imaging:	53
II.2.1 Physical principles of ultrasound and imaging modes:	54
II.2.1.1 Physical principles:	54
II.2.1.2 Imaging modes:	56
II.2.2 Echographic semiology of thyroid nodules:.....	60
II.2.2.1 Equipment required:.....	60
II.2.2.2 Conduct of the review:	61
II.2.2.3 Ultrasound Risk of Malignancy Prediction Scores:	65

II.2.3 Limits of human interpretation and interest of standardized systems:	66
II.2.3.1 Limits of human interpretation:	66
II.2.3.2 Interest of standardized stratification systems:.....	66
II.2.3.3 Towards the integration of artificial intelligence:	67
II.3 Diagnostic Assistance Systems (CAD):	67
II.3.1 CAD Definition & General Architecture:	68
II.3.2 Classic image processing methods:	71
II.3.2.1 Image enhancement:	72
II.3.2.2 Image segmentation:	72
II.3.2.3 Classification and extraction of characteristics:	75
II.3.3 Artificial intelligence approaches: machine learning and deep learning:.....	75
II.3.3.1 Machine Learning (ML):.....	75
II.3.3.2 Deep Learning (DL):.....	76
III.3.3.3 Specific applications in thyroid ultrasound imaging:.....	77
II.3.4 state of the art :.....	78
II.3.4.1 Work based on classical methods:	78
II.3.4.2 Deep Learning-based works:	78
II.3.4.3 Comparison of existing approaches:	79
II.3.4.4 Limites des travaux existants :.....	79
II.4 Segmentation and classification of nodules:	80
II.4.1 Segmentation: classical techniques and deep learning (U Net):	80
II.4.1.1 Classic segmentation techniques:.....	80
II.4.1.2 Deep learning segmentation:	81
II.4.2 Extraction of characteristics (morphological, textural, deep features):	85
II.4.2.1 morphological characteristics:	85
II.4.2.2 Textural characteristics:	86
II.4.2.3 Frequency characteristics:	87
II.4.2.4 Deep Features:	87
II.4.2.5 Hybrid approaches: merging of features:	88
II.4.3 Classification: classical classifiers (SVM, Random Forest) and neural networks (CNN):	88
II.4.3.1 Classical classifiers:.....	89

II.4.3.2 convolutional neural networks (CNN):	90
II.4.4 Evaluation metrics and conclusion of chapter 2:	92
II.4.4.1 Evaluation metrics in classification:	92
II.4.4.2 Evaluation metrics in segmentation:	93
II.4.4.3 experimental validation:	94
II.4.4.4 Conclusion of Chapter II:	94
CHAPTER III : DESIGN, DEVELOPMENT AND EVALUATION OF THE PROPOSED SYSTEM	96
III.1 Introduction:	97
III.2 Hardware and data:	97
III.2.1 Description of the ultrasound image database:	98
III.2.1.1 The TN3K database:	98
III.2.1.2 The TN5000 database:	98
III.2.2 Nodule annotation and TI-RADS classification:	99
III.2.2.1 Annotation of thyroid nodules in TN3K and TN5000:	99
III.2.2.2 TI-RADS classification system:	100
III.2.2.3 Correspondence between annotation and TI-RADS in datasets:	101
III.2.3 Pretreatments and data augmentation:	101
III.2.3.1 TN3K:	101
III.2.3.2 TN5000:	102
III.2.3.3 Development environment and implementation tools :	103
III.3 Methodology:	104
III.3.1 Segmentation of nodules: chosen architecture and training:	105
III.3.2 Feature extraction and classification: tested models:	107
III.3.2.1. General classification pipeline:	108
III.3.2.2 Preprocessing applied before classification:	108
III.3.2.3 Deep Learning models tested:	109
III.3.2.4 Classical methods of machine learning :	113
III.3.3 Evaluation protocol and metrics:	115
III.3.3.1 Experimental protocol:	115
III.3.3.2. Evaluation metrics for segmentation:	116
III.3.3.3. Evaluation metrics for classification :	117

III.4 Results and discussion:	118
III.4.1. Segmentation performance (Dice, IoU):	118
4.1.1 Quantitative segmentation results:	119
III.4.1.1 Learning curve analysis:	119
III.4.1.2 Qualitative analysis of segmentation results:	120
III.4.2 Classification performance (accuracy, AUC, sensitivity, specificity):	122
III.4.2.1 Results of Deep Learning Models:	122
III.4.2.2 Analysis of confusion matrices and OCR curves:	123
III.4.2.3 Results of classical machine learning methods :	125
III.4.3 Comparison with the state of the art and analysis of errors:	127
III.4.3.1. Comparison of segmentation performance:	128
III.4.3.2. Comparison of classification performance :	129
III.4.3.3 Segmentation error analysis:	130
III.4.3.4. Analysis of classification errors:	130
III.4.3.5 Interpretability of Model Decisions Using Grad-CAM:	131
III.4.4 Graphical interface of the proposed system:	133
III.4.4.1 Segmentation Application — TN3K Clinical Segmentation System :	134
III.4.4.2 Classification Application — TN5000 Thyroid Classification System.....	139
III.4.5 Limits, perspectives and conclusion of chapter III:	143
III.4.5.1. Limitations of the study :	143
III.4.5.2 Conclusion to chapter III:	144
GENERAL CONCLUSION AND PERSPECTIVES	145
Bibliographic reference:	148
Appendix	156
Appendix A: Examples of annotated images	157
Appendix B: Code Snippets :	158
Appendix C: Detailed tables of results	159

LIST OF FIGURES

Figure I.1: Diagram of the thyroid gland with vertebral landmarks (adapted from Journal des Femmes Santé, 2025). [54]

Figure I.2: Superior thyroid artery – anatomical diagram (adapted from Vasković, 2023). [55]

Figure I.3: Thyroid gland (adapted from Khodja, 2023). [2]

Figure I.4: Arterial blood supply of the thyroid gland (adapted from Anatomy.app, n.d.). [56]

Figure I.5: Histology of the thyroid gland (H&E staining) (adapted from Khodja, n.d.). [3]

Figure I.6: Photomicrograph showing thyroid follicles (adapted from Achouri & Adili, 2018). [57]

Figure I.7: Synthesis of thyroid hormones (adapted from Bessaguet, Suteau & Desmoulière, 2023). [58]

Figure I.8: Arterial vascularization of the thyroid (anterior view) (adapted from Chenafa, 2026). [59]

Figure I.9: Lymphatic drainage (adapted from Médecine Intégrée, n.d.). [60]

Figure I.10: Transverse ultrasound image showing a thyroid nodule in the right lobe of the thyroid and adjacent anatomical structures (adapted from Abdolali et al., 2020). [76]

Figure I.11: Prevalence of thyroid nodules according to age (adapted from Balice & Theintz, 2007). [61]

Figure I.12: Ultrasound of differentiated thyroid carcinomas (adapted from Marcy et al., 2024). [62]

Figure I.13: Histological differences between a benign nodule, papillary carcinoma, and medullary carcinoma. (original work, 2026).

Figure I.14: TI-RADS classification (adapted from American College of Radiology, 2017). [13]

Figure I.15: Fine-needle aspiration (FNA) (adapted from Alamy, n.d.). [77]

Figure I.16: Signs of hyperthyroidism (adapted from Traynard, 2010). [63]

Figure I.17: Clinical signs of hypothyroidism (adapted from Physiopharmlab, 2021). [64]

Figure I.18: Autoimmune thyroiditis (adapted from Journal des Femmes Santé, 2025). [65]

Figure I.19: Palpation of the thyroid region (adapted from Journal des Femmes Santé, 2026). [66]

Figure I.20: Laboratory assessment of the thyroid (adapted from Elsevier Masson, 2009). [67]

Figure I.21: Thyroid ultrasound (adapted from iSonic – SonoSchool, n.d.). [68]

Figure I.22: Thyroid scintigraphy (adapted from Laradioactivite.com, n.d.). [69]

Figure I.23: Thyroid MRI (adapted from Le Monde – Réalités Biomédicales, 2024). [70]

Figure I.24: Thyroid elastography (adapted from Radiologie Croisette, n.d.). [71]

Figure I.25: Thyroid scan (adapted from ORL Pôle Azur Chirurgie – ORPAC Grasse, n.d.). [72]

Figure I.26: Malignant nodule (original work, 2026).

Figure I.27: Robotic transaxillary thyroid lobectomy surgical technique (adapted from Agus et al., 2018). [73]

Figure I.28: Total thyroidectomy (adapted from Depositphotos, n.d.). [80]

Figure I.29: Lymph node dissection in thyroid surgery (adapted from Chirurgie Thyroïde Paris, n.d.). [74]

Figure I.30: Anatomy of the laryngeal nerve (adapted from Faculty of Medicine of Oran, n.d.). [81]

Figure I.31: Thyroid hematoma secondary to blunt cervical trauma (adapted from R. Ait Laabdia et al.).[82]

Figure II.1: Physical phenomena underlying ultrasound image formation (adapted from Ouchtati, n.d.). [28]

Figure II.2: B-mode ultrasound in a 128 grayscale scale — transverse section of the left thyroid lobe (adapted from Haute Autorité de Santé, 2021). [27]

Figure II.3: 3D ultrasound (adapted from Ouchtati, n.d.). [28]

Figure II.4: Color Doppler ultrasound — longitudinal section of a lobe (adapted from Haute Autorité de Santé, 2021). [27]

Figure II.5: Calculation of thyroid lobe volume (adapted from Haute Autorité de Santé, 2021). [27]

Figure II.6: Calculation of isthmus volume (adapted from Haute Autorité de Santé, 2021). [27]

Figure II.7: Parenchymal–muscular gradient (adapted from Haute Autorité de Santé, 2021). [27]

Figure II.8: Nodule localization diagram (adapted from Haute Autorité de Santé, 2021). [27]

Figure II.9: Diagram showing nodule shape according to its location within the lobe (adapted from Haute Autorité de Santé, 2021). [27]

Figure II.10: Diagram showing nodule margins — only micro-lobulated and spiculated margins are highly suspicious signs (adapted from Haute Autorité de Santé, 2021). [27]

Figure II.11: Continuous peripheral macrocalcification “eggshell pattern” and “broken eggshell pattern,” more suspicious (adapted from Haute Autorité de Santé, 2021). [27] **Figure II.12:** The three types of nodule vascularization (adapted from Haute Autorité de Santé, 2021). [23]

Figure II.13: Overall architecture of a CAD system: segmentation, feature extraction, classification, and final decision (adapted from Sun, Yu & Xu, 2020). [75]

Figure II.14: Examples of nodule segmentation results using the ResUNet model — (A) original image, (B) mean of physicians 1, 2, 3, and (C) ResUNet segmentation (adapted from Litjens et al., 2017). [23]

Figure II.15: Architecture of a hybrid CAD system with deep feature extraction and pre-weighting (original work, 2026).

Figure II.16: Basic U-Net architecture showing encoder, decoder, bottleneck, and skip connections — encoding layers (contraction phase): feature extraction; decoding layers (expansion phase): shape reconstruction (adapted from Ronneberger, Fischer & Brox, 2015). [35]

Figure II.17: Feature extraction of thyroid nodules (original work, 2026).

Figure III.1: Distribution of benign and malignant classes in the TN5000 dataset after preprocessing (generated with Python in Google Colab, 2026).

Figure III.2: Architectures of the models used for the segmentation of thyroid nodules (UNET++, EfficientNetB3) thyroid (original work, 2026).

Figure III.3: EfficientNetB0 convolutional neural network architecture used for image classification (personal development, 2026).

Figure III.4: DenseNet121 convolutional neural network architecture used for image classification (personal development, 2026).

Figure III.5: ResNet-18 convolutional neural network architecture used for image classification (personal development, 2026).

Figure III.6: Vision Transformer (ViT-B/16) convolutional neural network architecture used for image classification (personal development, 2026).

Figure III.7: Training monitoring: evolution of loss function and performance metrics (Accuracy, Dice, IoU) (generated with Python in Google Colab, 2026).

Figure III.8: Illustration of the thyroid nodule segmentation process: comparison between original images, preprocessed images, ground truth masks, and predicted masks (generated with Python in Google Colab, 2026).

Figure III.9: Example of thyroid nodule segmentation: (A) original ultrasound image, (B) ground truth mask, (C) predicted segmentation result (generated with Python in Google Colab, 2026).

Figure III.10: ROC curve and normalized confusion matrix of the EfficientNetB0 model

(generated with Python in Google Colab, 2026).

Figure III.11: ROC curve and normalized confusion matrix of the DenseNet121 model (generated with Python in Google Colab, 2026).

Figure III.12: Normalized confusion matrix of the ResNet18 model (generated with Python in Google Colab, 2026).

Figure III.13: ROC curve and normalized confusion matrix of the ViT-B/16 model (generated with Python in Google Colab, 2026).

Figure III.14: Normalized confusion matrix of the XGBoost model (generated with Python in Google Colab, 2026).

Figure III.15: Normalized confusion matrix of the Random Forest model (generated with Python in Google Colab, 2026).

Figure III.16: ROC curve and normalized confusion matrix of the SVM model (generated with Python in Google Colab, 2026).

Figure III.17: Visualization of benign and malignant nodule contours after enhancement of the TN5000 dataset (generated with Python in Google Colab, 2026).

Figure III.18: Example of segmentation failure due to low contrast and ill-defined borders (generated with Python in Google Colab, 2026).

Figure III.19: Example of a false positive with Grad-CAM visualization of the DenseNet121 model. The heatmap highlights regions contributing to the prediction, with high activations shown in red and low contributions in blue (generated with Python in Google Colab, 2026).

Figure III.20: Example of a true positive with Grad-CAM visualization of the DenseNet121 model (generated with Python in Google Colab, 2026).

Figure III.21: Example of a true negative with Grad-CAM visualization of the DenseNet121 model (generated with Python in Google Colab, 2026).

Figure III.22: Global architecture and functional pipeline of thyroid diagnostic support applications (personal work, 2026).

Figure III.23: Academic header of the TN3K Clinical Segmentation System application (generated using the Streamlit segmentation application, 2026).

Figure III.24: Quadruple visualization — original image, predicted mask, probability map, and clinical overlay (generated using the Streamlit segmentation application, 2026).

Figure III.25: Performance table and mode selection interface (generated using the Streamlit segmentation application, 2026).

Figure III.23: Clinical analysis — diagnosis, calibrated confidence, and TI-RADS estimation

(TR2, 99.98% confidence) (generated using the Streamlit segmentation application, 2026).

Figure III.26: Comparison between the original image and the annotated ground truth mask from the TN3K dataset (generated using the Streamlit segmentation application, 2026).

Figure III.27: Result download interface and technical information section (generated using the Streamlit segmentation application, 2026).

Figure III.28: Academic header of the TN5000 Thyroid Classification System application (generated using the Streamlit classification application, 2026).

Figure III.29: Performance table, mode selection, and image upload interface (generated using the Streamlit classification application, 2026).

Figure III.30: Malignant case — example ultrasound image (004476.jpg) (generated using the Streamlit classification application, 2026).

Figure III.31: Benign case — example ultrasound image (004076.jpg) (generated using the Streamlit classification application, 2026).

Figure III.32: Malignant case — clinical result with high confidence (93.84%) (generated using the Streamlit classification application, 2026).

Figure III.33: Ambiguous benign case — clinical result with low confidence (55.74%) (generated using the Streamlit classification application, 2026).

Figure III.34: Technical section and application execution commands (generated using the Streamlit classification application, 2026).

Appendix:

Figure A.1: Example of an annotated ultrasound image from TN3K (adapted from Jenny, 2023). [52]

Figure A.2: Comparison between ground truth mask and predicted segmentation (generated with Python in Google Colab, 2026).

Figure A.3: Examples of benign and malignant nodules from TN5000 (generated with Python in Google Colab, 2026)

LIST OF TABLES

Table I.1: Comparison of thyroid nodules (adapted from Ameli, 2025). [78]

Table I.2: Comparison between thyroid cysts and adenomas (adapted from Santé sur le Net, n.d.). [79]

Table I.3: The ACR TI-RADS system (adapted from American College of Radiology, 2017). [13]

Table III.1: Architecture of the proposed segmentation model — EfficientNet-B3 + U-Net++ (original work, 2026).

Table III.2: Training hyperparameter configuration of the model (original work, 2026).

Table III.3: Performance results of the segmentation model (original work, 2026).

Table III.4: Comparison of deep learning model performances (original work, 2026).

Table III.5: Comparison of classical machine learning methods (original work, 2026).

Table III.6: Comparison of segmentation performance with state of the art (adapted from CNN and Transformer-based methods [35]-[89]-[90]-[91]-[93] and related works).

Table III.7: Comparison of classification performance with state of the art (adapted from CNN and Transformer-based methods [85]-[87]-[92] and related works).

Table III.8: Comparison and complementarity of the two system applications. (original work, 2026).

Appendix:

Table C.1: Detailed performance results of classification models evaluated on the TN5000 dataset (original work, 2026).

Table C.2: Detailed performance results of the segmentation model evaluated on the TN3K dataset (original work, 2026).

LIST OF ABBREVIATIONS

FNA Fine Needle Aspiration.

TI-RADS Thyroid Imaging Reporting and Data System — Ultrasound classification system for thyroid nodules.

C5 Category 5 of the cytological classification (strong suspicion of malignancy/malignant).

T3 Triiodothyronine — Active thyroid hormone.

T4 Thyroxine — Main thyroid hormone secreted by the thyroid gland.

TSH Thyroid Stimulating Hormone — Thyroid-stimulating hormone (secreted by the pituitary gland).

H&E Hematoxylin and Eosin — Standard histological stain.

APUD Amine Precursor Uptake and Decarboxylation — Diffuse endocrine cell system.

TSH-R Thyroid Stimulating Hormone Receptor — TSH receptor.

NIS Sodium Iodide Symporter — Sodium/iodine transporter.

Na⁺ /K⁺ Sodium/potassium pump.

ATP Adenosine Triphosphate — Cellular energy molecule.

TPO Thyroid Peroxidase — Key enzyme in thyroid hormone synthesis.

MIT Monoiodotyrosine — Simple iodinated residue.

DIT Diiodotyrosine — Résidu iodé double.

EU-TIRADS European Thyroid Imaging Reporting and Data System – Système européen de classification échographique des nodules thyroïdiens.

ACR TI-RADS American College of Radiology Thyroid Imaging Reporting and Data System – Système américain basé sur un score pour évaluer le risque de malignité.

TR1 TI-RA DIT Diiodotyrosine — Double iodine residue.

EU-TIRADS European Thyroid Imaging Reporting and Data System — European system for classifying thyroid nodules on ultrasound.

ACR TI-RADS American College of Radiology Thyroid Imaging Reporting and Data System – American scoring system for assessing the risk of malignancy.

TR1 TI-RADS 1 – Benign nodule, very low risk.

TR2 TI-RADS 2 – Nodule probably benign.

TR3 TI-RADS 3 – Nodule of low suspicion.

TR4 TI-RADS 4 – Moderately suspicious nodule.

TR5 TI-RADS 5 – Highly suspicious nodule, high risk of malignancy.

TR ab TSH Receptor Antibodies – Anti-TSH receptor antibodies (indicator of Graves' disease).

CMT Calcitonin – Calcitonin (hormone secreted by the C cells of the thyroid gland, marker of medullary cancer).

Tg Thyroglobulin – Thyroglobulin (tumor marker for differentiated thyroid cancers).

anti-Tg Anti-thyroglobulin antibodies – Anti-thyroglobulin antibodies.

anti-TPO Anti-thyroid peroxidase antibodies – Anti-thyroid peroxidase antibodies.

SVC Superior vena cava – Vascular structure, sometimes invaded by thyroid tumors.

ATS American Thyroid Society.

MEN2 Multiple Endocrine Neoplasia type 2 (genetic syndrome).

RET Proto-oncogene REarranged during Transfection – Mutation associated with medullary cancers and MEN2.

MEN2A Variant of MEN2 with pheochromocytoma, hyperparathyroidism, and medullary thyroid cancer.

MEN2B Aggressive variant of MEN2 with pheochromocytoma, mucositis, and early medullary cancer.

ATA American Thyroid Association – Issues clinical recommendations.

PTEN Phosphatase and Tensin Homolog – Tumor suppressor gene, mutation in certain syndromes.

DICER1 Gene involved in microRNA processing, mutation associated with certain pediatric thyroid cancers.

RAI Radioactive Iodine – Used for the treatment and diagnosis of thyroid cancer.

CAD Computer-Aided Diagnosis.

Mode A Amplitude mode – Mode for displaying an ultrasound signal in amplitude (single line).

B Mode Brightness mode – Ultrasound mode in “grayscale” (two-dimensional image).

M Mode Motion mode – Display of the movement of structures on an ultrasound line (often used for the heart).

TM Mode Time-Motion mode – Variation of M mode, temporal tracking of movements.

2D Two-dimensional – Standard two-dimensional imaging

3D Three-dimensional – Reconstructed three-dimensional imaging.

MHz Megahertz – Frequency of the ultrasound waves used by the probe.

PRF Pulse Repetition Frequency – Frequency of repetition of ultrasound pulses.

ROI Region of Interest.

U-Net U-shaped Convolutional Neural Network – Used for medical image segmentation.

CNN Convolutional Neural Network – Used for image analysis and segmentation.

SVM Support Vector Machine – Support vector machine, algorithm

KNN K-Nearest Neighbors – K-nearest neighbors algorithm, classification method based on the proximity of data points.

AI Artificial Intelligence – Set of methods enabling a machine to

FCM Fuzzy C-Means – A fuzzy classification method often used for medical image segmentation.

ML Machine Learning – A set of methods that enable computers to learn from data.

DL Deep Learning – A subfield of machine learning that uses deep neural networks to analyze complex data.

ResNet Residual Network – Deep neural network using residual connections to improve learning and avoid performance degradation in very deep networks.

FCN Fully Convolutional Network – A network used for pixel-by-pixel image segmentation.

R-CNN Region-based Convolutional Neural Network – A convolutional network used for detecting and locating objects in an image.

SAM Segment Anything Model – Image segmentation model capable of automatically segmenting objects from minimal indications (points, boxes).

CRF Conditional Random Field – Probabilistic model used to refine segmentation results by taking spatial context into account.

RBF Radial Basis Function – Radial-based function, kernel frequently used in SVM models to handle nonlinear decision boundaries.

ROC Receiver Operating Characteristic – Curve representing the performance of a classifier based on the rate of true positives (sensitivity) and the rate of false positives (1 – specificity).

AUC Area Under the Curve – Area Under the Curve ROC, global measure of a model's performance.

GENERAL INTRODUCTION

GENERAL INTRODUCTION

Thyroid cancer represents the most widespread malignant endocrine pathology worldwide, with an incidence that has been steadily increasing over recent decades. While the vast majority of detected thyroid nodules are benign in nature, the early and reliable identification of malignant forms remains a major clinical imperative, directly conditioning the vital prognosis of patients and the effectiveness of their therapeutic management.

Ultrasound imaging currently constitutes the first-line modality for exploring the thyroid gland. Non-invasive, rapid, and widely accessible, it enables detailed morphological evaluation of thyroid nodules and guides clinical decisions according to standardized stratification systems such as the EU-TIRADS and ACR TI-RADS classifications. Yet, despite its undeniable advantages, the interpretation of ultrasound images remains a complex task, heavily dependent on the radiologist's expertise. The limiting factors are numerous: the inherently low contrast of ultrasound images, the presence of speckle noise, the high morphological variability of nodules, and inter-operator subjectivity all constitute sources of diagnostic uncertainty.

It is in this context that artificial intelligence techniques, and more particularly Deep Learning, are emerging as a promising avenue for the automation and objectivization of medical image analysis. Convolutional neural network architectures, combined with attention mechanisms and pre-trained encoders, have demonstrated remarkable performance in complex medical image analysis tasks, paving the way for a new generation of Computer-Aided Diagnosis (CAD) systems.

Within this scientific and clinical framework, the present final-year project is situated at the intersection of biomedical engineering, signal processing, and artificial intelligence, contributing to the growing field of intelligent medical imaging. It aims to design, develop, and evaluate an automatic two-stage system for the analysis of thyroid nodules from ultrasound images: a first segmentation stage of nodular regions, carried out using a U-Net++ architecture coupled with an EfficientNet-B3 encoder pre-trained on ImageNet, and a second stage of binary classification of nodules — benign or malignant — leveraging several Deep Learning and Machine Learning models.

Experiments were conducted on the public databases TN3K and TN5000, widely recognized in the scientific community for benchmarking thyroid nodule segmentation and classification approaches. Rigorous preprocessing, data augmentation, and hyperparameter optimization techniques were implemented to improve the robustness and generalization capacity of the proposed models.

GENERAL INTRODUCTION

This dissertation is organized into three complementary chapters. The **first chapter** lays the medical foundations of the work by presenting the anatomy and physiology of the thyroid gland, the semiology of thyroid nodules, and the modalities of clinical diagnosis and management. The **second chapter** exposes the theoretical basis of ultrasound imaging, CAD systems, and AI-based image processing approaches, ranging from classical Machine Learning to the most advanced deep learning architectures. The **third chapter** details the proposed methodology, the architectural choices made, the datasets employed, and a thorough analysis of the experimental results obtained, including a comparison with the state of the art. A general conclusion synthesizes the major contributions of this work and outlines future research and clinical integration perspectives for AI-based diagnostic support systems.

**CHAPTER I: General information about the
thyroid gland and thyroid nodules**

I.1 Introduction:

The presence of thyroid nodules is a common clinical problem, particularly with the widespread use of imaging techniques such as ultrasound, which has significantly increased the rate of incidental detection. Although the majority of these nodules are benign, a small percentage may be malignant, making accurate diagnosis essential for appropriate patient management.

In clinical practice, the evaluation of thyroid nodules relies heavily on ultrasound imaging due to its non-invasive nature, accessibility, and effectiveness in characterizing nodule features. However, the interpretation of ultrasound images remains highly dependent on the radiologist's expertise and may be subject to inter-observer variability.

To address these limitations, several standardized risk stratification systems, such as the ACR TI-RADS, have been developed to guide clinical decision-making. Nevertheless, these systems still rely on manual assessment, which can be time-consuming and prone to subjectivity.

In this context, the integration of computer-aided diagnosis (CAD) systems based on artificial intelligence, particularly deep learning techniques, has emerged as a promising solution to improve the accuracy and consistency of thyroid nodule classification.

Therefore, this work aims to contribute to the development of an automated system for the segmentation and classification of thyroid nodules in ultrasound images, in accordance with established clinical guidelines [1].

I.2 Anatomy and physiology of the thyroid gland:

I.2.1 Gross Anatomy and Vascularization:

The thyroid gland is a single, butterfly-shaped endocrine organ located in the anterior neck. (see [1]), in front of the trachea, below the thyroid cartilage (Figure I.3), extending approximately to the level of the C5 vertebral segments (the fifth cervical vertebra (of the neck region)) to T1 (the first thoracic vertebra (beginning of the thoracic region, at shoulder level/top of the thorax)). See (Figure I.1)

On a macroscopic scale, the thyroid gland consists of two lateral divisions, specifically the right lobe and the left lobe. These two parts are joined by a transverse structure, called the isthmus, which extends over the anterior surface of the trachea (Figure I.2) and inferior to the thyroid cartilage. It establishes anatomical relationships with the subhyoid muscles, the trachea, the esophagus, as well as with the carotid arteries and jugular veins.

In adults, the average measurements are a length ranging from 4 to 6 centimeters, a width of 2 to 3 centimeters, and a thickness varying from 1 to 2 centimeters. The total weight is generally between 15 and 25 grams.[1]

GLANDE THYROÏDE

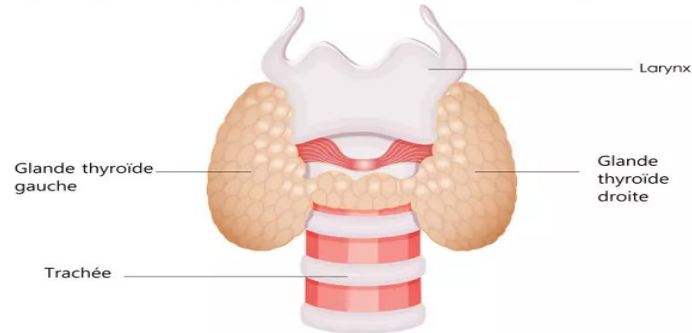


Figure I.1:Diagram of the thyroid gland with vertebral landmarks .[54]

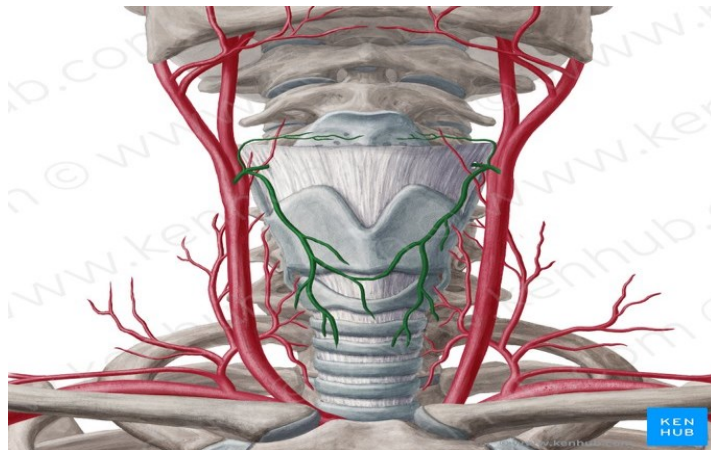


Figure I.2: Superior thyroid artery – anatomical diagram .[55]

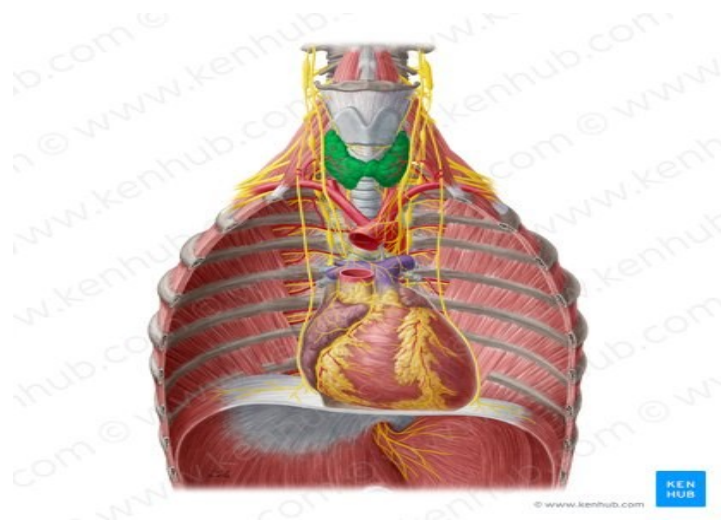


Figure I.3: Thyroid gland .[2]

The thyroid gland, as an endocrine gland, is characterized by abundant vascularization, which facilitates the rapid and effective diffusion of its hormones within the circulatory system. Its blood supply is mainly provided by the superior and inferior thyroid arteries, which have dense branching and numerous anastomoses, thus ensuring an adequate blood supply.

The superior thyroid artery branches into an anterior branch, specifically destined for the ventral surface of the gland, and a posterior branch, which supplies blood to the internal and external surfaces.

The inferior thyroid artery subdivides into two branches, one superior and one inferior, which supply the caudal and dorsal surfaces of the gland, respectively.

An ima thyroid artery, whose presence is variable, may occasionally originate from the brachiocephalic trunk or directly from the aorta, in which case, if present, it serves as a collateral circulation. See (Figure I.4) [1]



© 2006 The Authors
Journal compilation © 2006 Blackwell Publishing Ltd

1

0	0	1	3	1
---	---	---	---	---

Figure 1

I.2.2 Functional histology (follicles, C cells):

The thyroid gland has a lobular architecture, with the lobes themselves subdivided into lobules by thin connective tissue septa. Each lobule is intrinsically composed of follicles, which represent the primary morphofunctional units of this endocrine organ. See (Figure I.5)

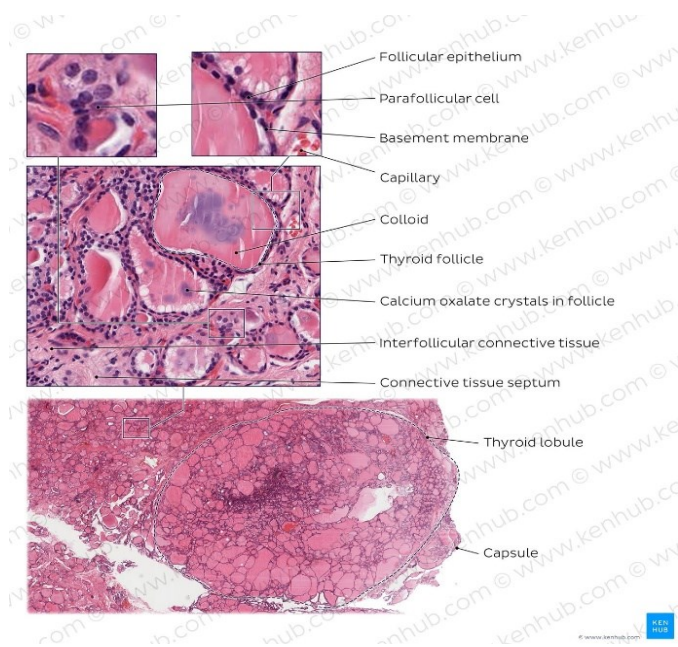


Figure I.5: Histology of the thyroid gland. [3]

The follicular units of the thyroid gland are embedded in a richly vascularized connective stroma. This surrounding environment is characterized by a dense network of fenestrated capillaries, lymphatic vessels, and sympathetic nerve endings. This histological organization promotes optimal transport of thyroid hormones to the circulatory system.

► Follicular cells (thyrocytes):

- **Morphology:** Thyrocytes have a variable morphology, ranging from cuboidal to flattened cylindrical during metabolic activity, and evolving into a squamous (pavement-like) shape during phases of functional rest. See (Figure I.6)
- **Function:** These cells are responsible for the biosynthesis of thyroglobulin, a glycoprotein that is the precursor of thyroid hormones. Thyroglobulin is stored as a colloid in the follicular lumen.
- **State of activity:**

Resting follicles: The follicular lumen is enlarged and the colloid is dense.

Active follicles: The lumen is narrowed and the amount of colloid is more limited.

- Mechanism of action: Thyrocytes internalize amino acids and iodine through their basolateral membrane. They then release triiodothyronine (T3) and thyroxine (T4) into the bloodstream via their basal pole.

► Parafollicular cells (C cells):

- Location: Parafollicular cells, also known as C cells, are located at the periphery of the thyroid follicles or scattered within the interfollicular connective tissue.

- Histological appearance: Histologically, their cytoplasm appears pale after conventional hematoxylin-eosin (H&E) staining, indicating limited dye reactivity. See (Figure I.6)

- Function: These cells are involved in the synthesis and secretion of calcitonin, a peptide hormone essential for the regulation of calcium homeostasis. Its main action is exerted by:

- Inhibiting osteoclastic bone resorption activity.

- Reducing tubular reabsorption of calcium in the renal nephrons.

- Nature: Parafollicular cells (C cells) are classified as neuroendocrine cells and are an integral part of the APUD (Amine Precursor Uptake and Decarboxylation) system. [3].

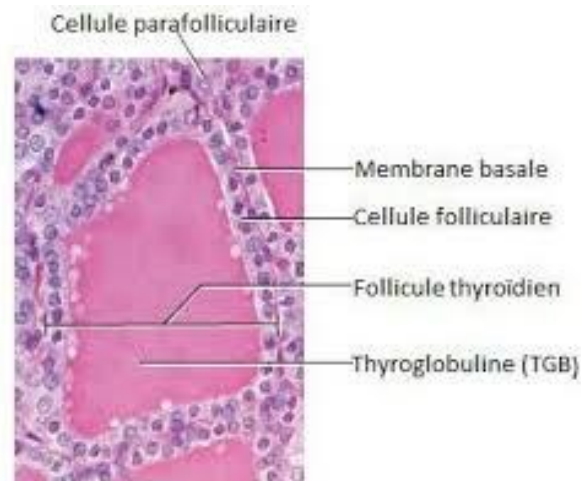


Figure I.6: Photomicrograph showing thyroid follicles .[57]

I.2.3 Endocrine physiology: T3, T4, calcitonin:

The biosynthesis of thyroid hormones is initiated by the binding of thyroid-stimulating hormone (TSH), a glycoprotein produced by the pituitary gland, to its specific membrane receptor (TSH-R), located on the basolateral membrane of thyrocytes (thyroid follicular cells). The internalization of iodide into the cell cytoplasm is mediated by the sodium/iodide symporter

(NIS), a membrane transporter that ensures the co-transport of sodium and iodide. The electrochemical sodium gradient, which is essential for this process, is maintained by the activity of Na^+/K^+ -ATPase, an ion pump that consumes ATP to expel sodium and import potassium.

The essential step in iodide organification is catalyzed by thyroperoxidase (TPO), a key enzyme that oxidizes iodide and covalently incorporates it into the tyrosyl residues of thyroglobulin. This reaction results in the formation of iodinated precursors: monoiodotyrosine (MIT), a monoiodinated tyrosyl residue, and diiodotyrosine (DIT), a diiodinated tyrosyl residue.

TPO is also responsible for coupling iodotyrosines, allowing the condensation of MIT and DIT to generate triiodothyronine (T3), the biologically active form, and the condensation of two DITs to form thyroxine (T4), the most abundant hormone, which acts as a prohormone by being subsequently converted to T3 in peripheral tissues.

The thyroid hormones produced in this way are stored in the colloid, a gelatinous matrix occupying the lumen of the thyroid follicles. They remain attached to iodinated thyroglobulin, a large thyroid protein that is the main storage matrix for MIT, DIT, T3, and T4. Following stimulation by TSH, iodinated thyroglobulin is endocytosed by thyrocytes, where it undergoes proteolysis by lysosomal enzymes. This process leads to the release of T3 and T4 into the bloodstream. [4] See (Figure I.7)

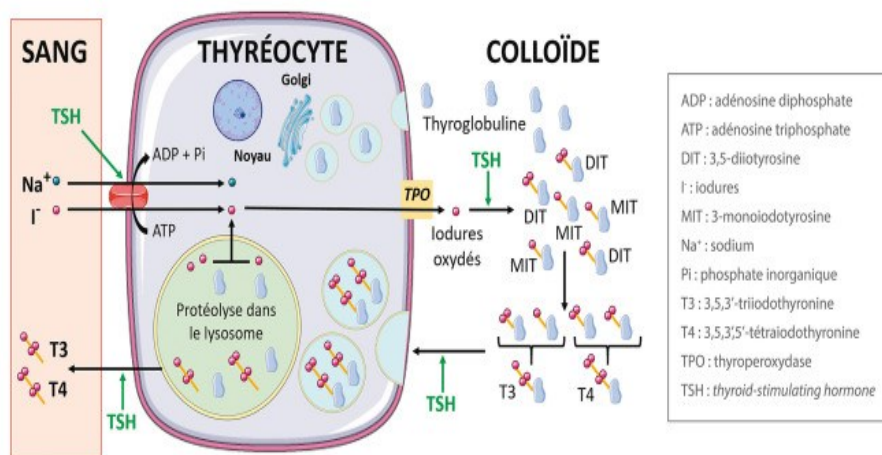


Figure I.7: Synthesis of thyroid hormones. [58]

I.2.4 Lymphatic system and innervation:

Lymphatic drainage: initially, lymphatic flow from the thyroid gland is directed to the pre-laryngeal lymph nodes (also known as Delphic nodes), before spreading to the pre-tracheal,

paratracheal, or deep cervical lymph nodes (particularly the jugular nodes). Additionally, anastomotic connections can direct lymph to the thoracic duct or mediastinal nodes. This specific anatomical architecture highlights the need for a comprehensive evaluation of the lymph nodes (both lateral and central compartments) when detecting nodular formations or neoplastic processes in the thyroid. See (Figure I.8)

The sympathetic innervation of the thyroid gland originates from a network emanating from the cervical ganglia (superior, middle, and inferior). These nerve fibers travel along the thyroid arteries to innervate the glandular parenchyma. A parasympathetic contingent originates from the vagus nerve (Xth cranial nerve), carried by the superior laryngeal nerve and the recurrent nerve. Although this innervation has a modulatory influence on the vascularization and trophicity of the gland, it is important to note that the primary regulation of hormone production is mainly controlled by TSH (thyroid-stimulating hormone). See (Figure I.9) [5]

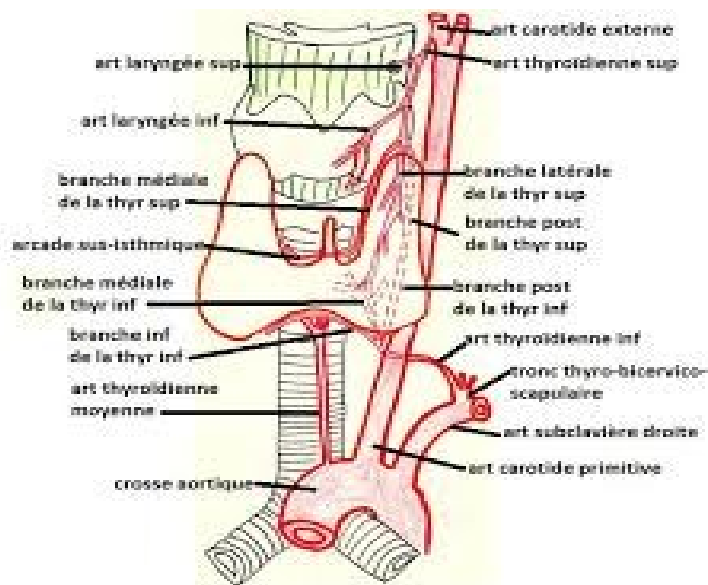


Figure I.8: Arterial vascularization of the thyroid (anterior view). [59]

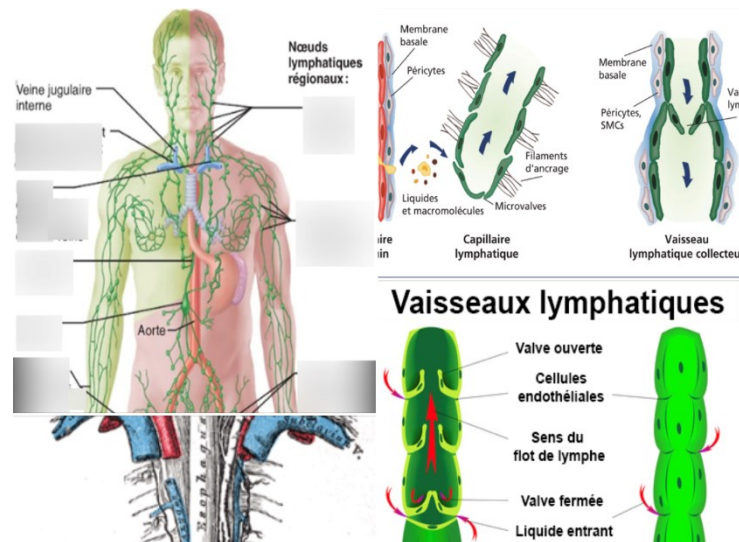


Figure I.9 :Lymphatic drainage .[60]

The anatomical organization of the lymphatic drainage and innervation of the thyroid gland has important clinical implications. In particular, the rich lymphatic network explains why thyroid cancers, especially papillary carcinoma, frequently metastasize to cervical lymph nodes. This characteristic spread pattern makes the evaluation of cervical lymph nodes an essential component of thyroid ultrasound examination.

From an imaging perspective, ultrasound plays a key role not only in the assessment of thyroid nodules but also in the detection of suspicious lymph nodes, which may present features such as microcalcifications, cystic changes, or abnormal vascularization.

Furthermore, understanding this anatomical context is crucial for the development of computer-aided diagnosis systems, as it highlights the importance of incorporating both nodule characteristics and surrounding tissue information in automated analysis. This is particularly relevant for deep learning models, which can exploit spatial and contextual features to improve diagnostic performance.

I.3 Thyroid nodules:

I.3.1 Definition, epidemiology, and risk factors:

A thyroid nodule is defined as a focal abnormality within the thyroid gland that has characteristics distinct from the surrounding glandular parenchyma. These formations can be detected during a clinical palpation examination or incidentally revealed by medical imaging techniques such as ultrasound or computed tomography. The prevalence of thyroid nodules is

considerable, with contemporary epidemiological data indicating that a substantial proportion of the adult population, exceeding 50%, is likely to have at least one of these lesions, detectable by ultrasound. However, although most of these formations are benign in nature, the emergence of a suspicion of malignancy requires a comprehensive diagnostic workup, including advanced imaging studies, fine-needle aspiration cytology, and, potentially, surgery. A thorough understanding of the pathophysiology of thyroid nodules is fundamental to the optimal management of thyroid health.

The table below shows the different ways of diagnosing:

Type	Characteristics	Implications
Solid nodule	Dense tissue	Requires medical evaluation
Cystic nodule	Filled with fluid	Often benign, may require drainage
Mixed nodule	Both solid and fluid	Requires further evaluation
Hot Nodule	Produces excess hormones	May cause hyperthyroidism
Cold Nodule	Inactive, does not produce hormones	Often benign, but must be checked for cancer

Table I.1:Comparison of thyroid nodules .[78]

Nodules can sometimes affect how your thyroid works. A “hot” nodule, for example, could produce extra thyroid hormones, leading to hyperthyroidism. On the other hand, a “cold” nodule does not produce hormones and could potentially be malignant. Regular checkups and early detection are essential for effectively managing thyroid nodules. [6]

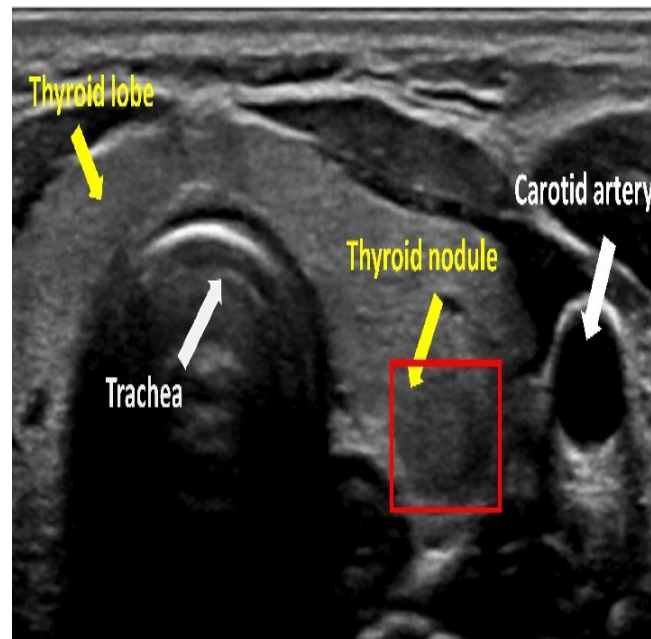


Figure I.10 :Transverse ultrasound image showing a thyroid nodule in the right lobe of the thyroid and adjacent anatomical structures .[76]

Causes and Risk Factors:

Thyroid nodules are formations within the thyroid parenchyma resulting from localized cell proliferation with multiple causes:

- **Iodine deficiency:** Prolonged iodine deficiency is associated with glandular hypertrophy (goiter) and the emergence of nodular structures.
- **Chronic inflammation:** Persistent inflammatory processes in the thyroid gland, such as Hashimoto's thyroiditis, can induce the development of nodules.
- **Benign cystic formations:** The presence of cysts, characterized by their fluid content, is also an etiological factor in thyroid nodules.
- **Genetic predisposition:** A family history of nodular or carcinomatous thyroid pathologies increases individual susceptibility.
- **Hormonal fluctuations:** Physiological (pregnancy, menopause) or pathological hormonal variations can modulate their appearance.
- **Chronic TSH stimulation:** A persistent elevation of thyroid-stimulating hormone (TSH) is likely to cause glandular hypertrophy.

Several factors must be considered when assessing the potential for malignancy:

- Age and sex: Male sex and pediatric age are associated with an increased risk of malignant transformation of thyroid nodules.
- Exposure to radiation: Previous exposure to ionizing radiation, whether therapeutic (radiotherapy) or accidental (nuclear events), is a significant risk factor.
- Family history of thyroid carcinoma: A family history of thyroid cancer, particularly in the presence of specific genetic mutations or hereditary syndromes (e.g., familial medullary carcinoma), confers increased susceptibility.

In terms of prevalence:

- Clinical examination: Cervical palpation can identify nodules in approximately 4-7% of women and 1-2% of men.
- Ultrasound imaging: Thyroid ultrasound reveals a significantly higher prevalence, reaching up to 20% of the general population and exceeding 50% in specific cohorts (e.g., women over 50). See (Figure I.11)

It should be noted that the observed increase in incidental detections (incidentalomas) of nodules is directly correlated with the increased use of diagnostic imaging modalities, which alters the perception of their true prevalence in the population. [7]

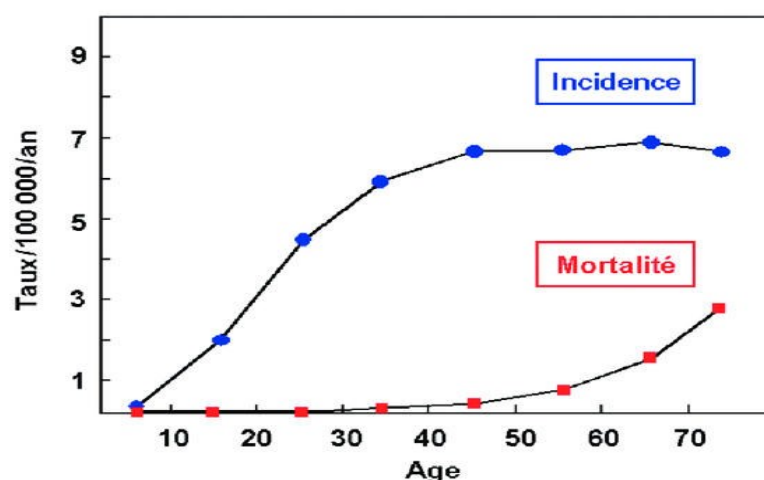


Figure I.11: Prevalence of thyroid nodules according to age. [61]

I.3.2 Types of nodules: benign, malignant, solitary vs. multinodular:

The characterization of thyroid nodular formations is based on two main criteria: on the one hand, their multiplicity (distinguishing them as solitary lesions or multinodular goiters) and, on the other hand, their histopathological profile (benign or malignant). This taxonomy is fundamental to guiding the diagnostic process and developing a therapeutic management strategy.

I.3.2.1 Solitary vs. multinodular nodules:

A solitary thyroid nodule is defined as an individualized formation confined within the thyroid gland. It is often detected incidentally or during a clinical examination. Its size varies, and it may remain asymptomatic or manifest as a palpable or visible increase in volume.

Multinodular goiter, on the other hand, is characterized by the presence of multiple nodules. Its incidence is higher in older individuals and in populations living in areas where iodine deficiency is endemic. It causes enlargement of the thyroid gland, which can lead to compressive symptoms such as difficulty swallowing, speaking, or breathing. Although the majority of these nodular formations, whether solitary or multiple, are benign in nature, their multiplicity can complicate the diagnostic process. An ultrasound evaluation and, if necessary, a fine needle aspiration (FNA) biopsy are fundamental diagnostic steps in order to accurately characterize their nature.

I.3.2.2 Benign nodules (cysts, adenomas):

Benign thyroid nodules, including cysts and adenomas, often cause no symptoms but still require careful monitoring. These nodules are usually noncancerous and can vary in size. A cyst is a fluid-filled sac, while an adenoma is a solid growth. Even though they are benign, it is important to keep an eye on them to ensure they do not grow larger or cause complications.

Here is a quick comparison to help you understand:

Type	Description
Cysts	Fluid-filled sacs
Adenomas	Solid growths
Symptoms	Generally asymptomatic
Monitoring	Regular ultrasound scans and checkups

Table I.2: Comparison between thyroid cysts and adenomas. [79]

Although benign, these nodules can sometimes cause discomfort or compressive symptoms depending on their size and location (Gharib et al., 2016)[8] . Regular checkups and ultrasounds will allow your doctor to monitor changes in size or appearance. If necessary, they may recommend a fine needle biopsy to rule out malignancy. In some cases, treatment options such as surgery or radiofrequency ablation may be considered if the nodule causes symptoms or cosmetic concerns. However, most benign nodules do not require treatment, just periodic observation. Always remain vigilant and follow your doctor's advice for the best results.[8]

I.3.2.3 Malignant nodules (thyroid cancer):

Although most thyroid nodules are benign, it is essential to be aware of malignant nodules, which can indicate thyroid cancer. Thyroid cancer is not common, but recognizing the types of malignant nodules can make a significant difference in early diagnosis and treatment. There are several types of thyroid cancer that you should be aware of.

- Papillary Carcinoma: The most common form of thyroid cancer, originating in follicular cells. See (Figure I.12)
- Follicular Carcinoma: Another type of thyroid cancer originating in follicular cells.
- Medullary Carcinoma: Thyroid cancer originating in parafollicular cells (C cells).

- Anaplastic Carcinoma: An aggressive and poorly differentiated form of thyroid cancer.

Ultrasound Indicators of Suspected Malignancy:

- “Taller than wider” shape (vertical orientation).
- Irregular or poorly defined edges.
- Presence of microcalcifications.
- Hypoechogenicity (nodule darker than normal thyroid tissue).

Histological origin:

The majority of thyroid nodules, whether benign or malignant, originate from follicular cells. See (Figure I.13)

Parafollicular cells (C cells), responsible for the production of calcitonin, are less frequently the cause of thyroid nodules. [8]

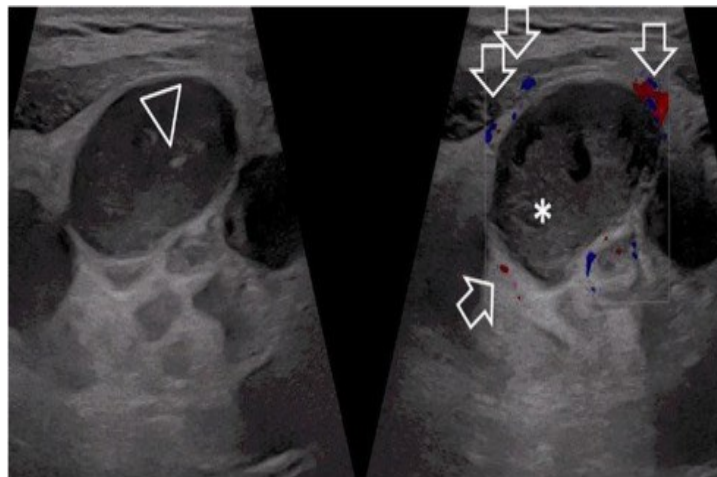


Figure I.12: Ultrasound of differentiated thyroid carcinomas. [62]

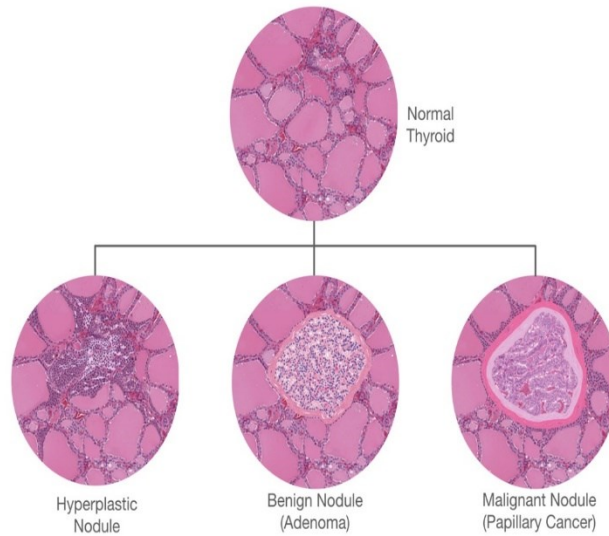


Figure I.13: Histological differences between a benign nodule, papillary carcinoma, and medullary carcinoma.

I.3.3 TI-RADS classification (EU TIRADS, ACR TI RADS):

The TI-RADS system is a standardized methodology for assessing the malignant potential of thyroid nodules based on the examination of their ultrasound attributes. This classification establishes a typology of nodules according to a progressive gradation of risk, ranging from TR1 (indicating the absence of lesion or benign morphology) to TR5 (indicating a high risk of malignant neoplasia). See (Figure I.14). The fundamental objectives of TI-RADS are to standardize the analysis of ultrasound examinations, guide the decision-making process regarding the indication for fine needle aspiration (FNA) cytology, and reduce inter-observer variability in the assessment of these formations.

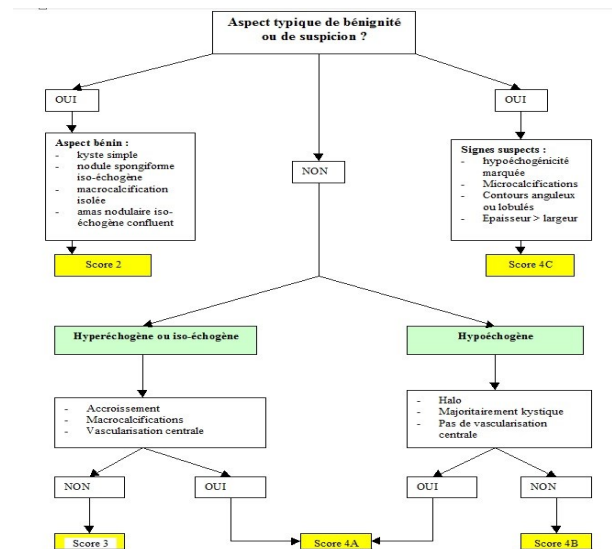


Figure I.14: TI-RADS classification. [13]

Complementary approaches coexist, such as the American College of Radiology's ACR-TI-RADS and the European Thyroid Association's EU-TIRADS. These systems share fundamental evaluation criteria, including internal composition, echogenicity, morphology, margin characteristics, and calcification detection.

► EU-TIRADS (European Thyroid Association) classification:

EU-TIRADS categorizes thyroid nodules into five distinct groups based on their ultrasound characteristics:

- EU-TIRADS 1 – Normal: No nodular lesion.
- EU-TIRADS 2 – Benign (risk of malignancy ~0%): Characterized by predominantly cystic or spongiform nodules.
- EU-TIRADS 3 – Low risk of malignancy (~2–4%): Includes isoechoic or hyperechoic nodules with well-defined contours and no suspicious ultrasound features.
- EU-TIRADS 4 – Moderate risk of malignancy (~6–17%): Identifies hypoechoic nodules that are ovoid in shape and have regular margins.
- EU-TIRADS 5 – High risk of malignancy (~26–87%): Defined by the presence of at least one highly suspicious ultrasound sign, such as:
 - microcalcifications,
 - irregular or spiculated margins,

- a “taller-than-wide” morphology,
- marked hypoechogenicity.

This classification system facilitates the harmonization of malignancy risk assessment and guides FNA indications in correlation with the assigned risk category and nodule dimensions.

[9] . See (Figure I.15)

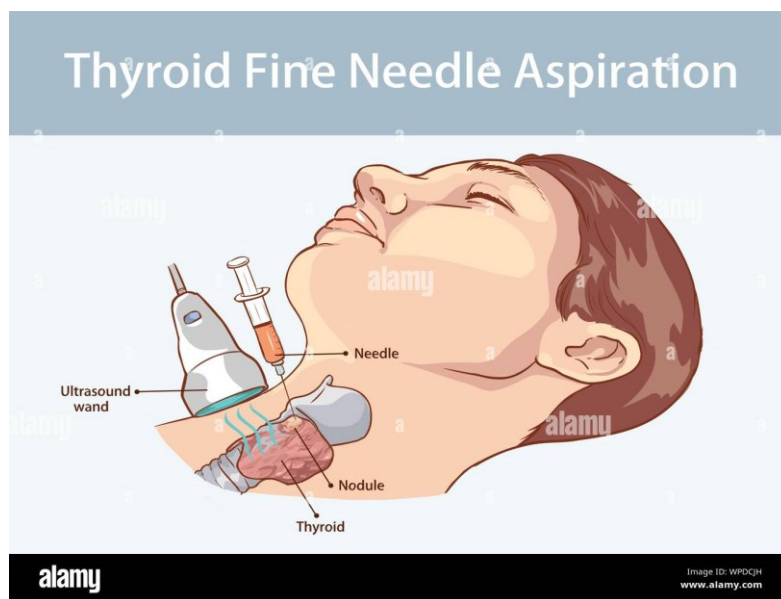


Figure I.15: Fine-needle aspiration (FNA). [77]

► ACR-TI-RADS classification (American College of Radiology):

Each nodule is evaluated according to five distinct ultrasound parameters: its structural composition, intrinsic echogenicity, morphology, peripheral boundaries, and the presence of hyperechoic foci. An overall score is then calculated, allowing the nodule to be assigned to an appropriate risk category.

Category	Significance	Risk	Management
TR1	Benign	0 %	No follow-up
TR2	Probably benign	< 2 %	Surveillance
TR3	Low suspicion	≈ 5 %	Biopsy if ≥ 2.5 cm

TR4	Moderate suspicion	5–20 %	Biopsy if ≥ 1.5 cm
TR5	High suspicion	> 20 %	Biopsy if ≥ 1 cm

Table I.3: The ACR TI-RADS. [13]**I.3.4 Clinical appearance of thyroid nodules:**

Thyroid nodules are distinct localized lesions in the thyroid tissue, identifiable by clinical examination or medical imaging techniques. The objectives of clinical evaluation, based on medical history and physical examination, are to:

- Determine the presence of thyroid dysfunction (hyperthyroidism, hypothyroidism),
- Identify symptoms related to compression of surrounding structures,
- Identify factors suggesting a risk of malignant transformation.

I.3.4.1 Medical history taking:

Collecting the patient's medical history is a fundamental step in the clinical examination, and its purpose is to accurately identify:

- The chronology of onset and growth kinetics of the nodular lesion, as well as any recurrence following therapeutic intervention.
- The patient's personal and family history, including the presence of thyroid neoplastic diseases, exposure to radiation in the cervical region, or the existence of autoimmune disorders.
- Clinical symptoms, including signs of thyroid dysfunction or manifestations related to compression.
- Risk factors for malignancy, such as previous exposure to radiation, a family history of cancer, age under 20 or over 60, and male gender.

I.3.4.2 Signs of thyroid disorder:

- Hyperthyroidism:

Hyperthyroidism is characterized by excessive production of thyroid hormones (triiodothyronine [T3] and thyroxine [T4]), leading to a generalized acceleration of fundamental metabolic processes. Possible causes include:

- An autonomous thyroid adenoma (hyperfunctioning nodule).
- Graves' disease (an autoimmune disorder characterized by hyperthyroidism, resulting from the presence of autoantibodies [TRAb] that stimulate the TSH receptor, leading to diffuse goiter, systemic manifestations, and orbitopathy). See (Figure I.16)
- Subacute or postpartum thyroiditis (a transient inflammation of the thyroid gland that often causes an initial phase of hyperthyroidism followed by a phase of hypothyroidism).

Clinical manifestations may include:

- Neuropsychiatric disorders: agitation, irritability, sleep disturbances, distal tremors, brisk deep tendon reflexes.
- Cardiovascular manifestations: increased heart rate, abnormal heartbeat perceptions, atrial arrhythmia, decreased exercise tolerance.
- Metabolic changes: weight loss despite maintaining or increasing food intake, increased appetite, elevated blood glucose levels.
- Digestive disorders: increased stool frequency, accelerated intestinal peristalsis, occasional fat malabsorption.
- Thermoregulation abnormalities: increased sensitivity to heat, excessive sweating, excessive moisture in the extremities.
- Musculoskeletal involvement: feeling of muscle fatigue, weakness of the proximal muscles, reduction in bone density during prolonged progression.

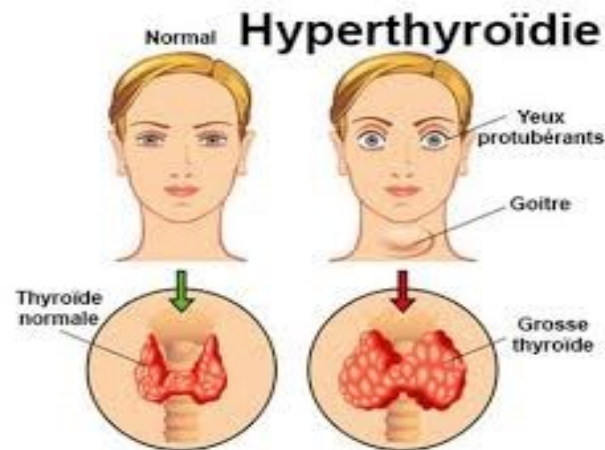


Figure I.16: Signs of hyperthyroidism .[63]

- Hypothyroidism:

Hypothyroidism is characterized by insufficient production of thyroid hormones, leading to decreased metabolic activity. The most common causes include:

- Chronic autoimmune thyroiditis (Hashimoto's thyroiditis), an autoimmune disease characterized by the progressive destruction of thyroid tissue, often leading to hypothyroidism. See (Figure I.17)

- Surgical removal of the thyroid gland or irradiation of the gland.

- Iodine deficiency.

The clinical symptoms manifest as follows:

- General symptoms: Asthenia, hypersomnia, and slowed psychomotor functions.
- Skin symptoms: Xerosis, cold skin, myxedema, dry and brittle hair, and ciliary alopecia.
- Cardiovascular manifestations: Bradycardia and decreased blood pressure.
- Gastrointestinal manifestations: Constipation.
- Thermoregulation disorders: Increased sensitivity to cold and intolerance to low temperatures.
- Other manifestations: Dysphonia, facial edema, and swelling of the extremities.

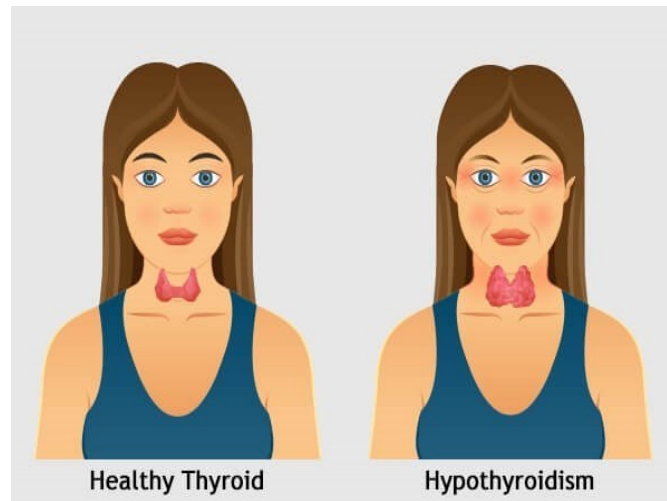


Figure I.17: Clinical signs of hypothyroidism .[64]

- **Thyroiditis:**

Thyroiditis, a condition characterized by an inflammatory reaction in the thyroid gland, can result from autoimmune causes, following an infection, or be subacute in nature. Its progression may manifest as a phase of temporary thyroid hyperactivity, potentially preceding thyroid insufficiency. See (Figure I.18)

The main forms include:

- Hashimoto's thyroiditis: An immune-mediated condition, frequently associated with painless thyroid enlargement and leading progressively to hypothyroidism.
- De Quervain's thyroiditis: A subacute form characterized by neck pain, fever, and initial hyperthyroidism.
- Postpartum thyroiditis: A generally transient autoimmune disorder.

Clinically, thyroiditis may manifest as:

- Pain localized in the anterior region of the neck, potentially radiating to the mandible.
- A moderate increase in body temperature and a feeling of general malaise.
- An increase in the volume of the thyroid gland, which may be firm or sensitive to palpation.
- Fluctuations in hormone concentrations, varying according to the etiology of the thyroiditis.

[10]

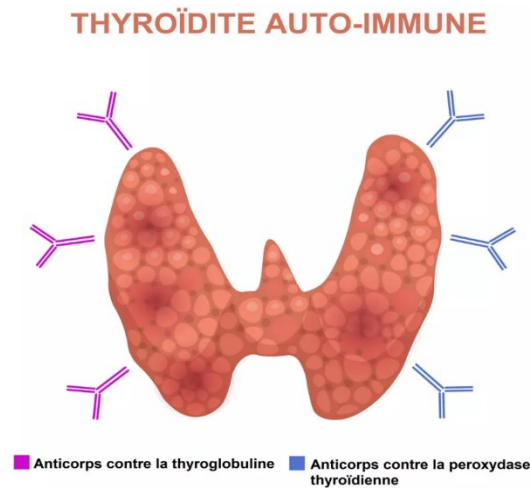


Figure I.18: Autoimmune thyroiditis .[65]

Despite the diversity of clinical manifestations associated with thyroid disorders, these symptoms are often non-specific and may overlap between different conditions. As a result, clinical evaluation alone is insufficient to establish a definitive diagnosis.

In practice, the detection of a palpable nodule or visible enlargement of the thyroid gland often leads to further investigation using ultrasound imaging. Ultrasound plays a crucial role in assessing morphological features such as nodule size, echogenicity, margins, and vascularization, which are essential for distinguishing between benign and malignant lesions.

Therefore, while clinical signs provide an initial indication of thyroid dysfunction, ultrasound imaging remains the cornerstone of diagnosis and risk stratification in thyroid diseases.

I.3.4.3 Signs of local compression:

A large nodule may exert pressure on nearby cervical anatomical structures, including the trachea, esophagus, and inferior laryngeal nerve. Associated clinical manifestations may include:

- **Dysphagia:** Difficulty swallowing, resulting from compression of the esophagus.
- **Dyspnea or feeling of respiratory obstruction:** Respiratory discomfort resulting from tracheal compression, potentially exacerbated when lying down.
- **Chronic or irritative cough:** A persistent cough or one triggered by irritation.

- **Dysphonia:** Voice alteration attributable to disruption of laryngeal nerve function.
- **Cervical pain or limited cervical range of motion:** Pain in the cervical region or restricted neck mobility.

Clinical significance: The presence of these symptoms suggests a symptomatic nodule, often requiring surgical intervention or careful monitoring. [7]

I.4 Diagnosis and management of thyroid nodules:

To understand the nature of small lumps (nodules) in the thyroid, several things must be considered at the same time: what the doctor observes when examining the patient, the results of blood tests (laboratory tests), and images obtained by machines (such as ultrasound, which is the most commonly used).

How these nodules are treated depends on several factors: the likelihood that they are serious (cancerous), the problems experienced by the patient, and the physical appearance of the lump (its shape, size, etc.).

• Clinical examination:

The physical examination, in addition to the medical history, provides a morphological and functional assessment of the thyroid nodule.

• Palpatory examination of the thyroid gland

The initial evaluation of a thyroid lesion relies on a thorough clinical examination including palpation. The patient's position is essential: they should sit comfortably with a slight cervical extension.

Systematic palpation of the thyroid region allows the dimensions, morphology and texture of the nodule to be characterised, as well as any algic reaction or sensitivity to pressure. These elements can indeed indicate an inflammatory or infectious process. During this examination, particular attention will be paid to the following characteristics:

- **consistence:** a firm or hard texture may suggest malignancy, while a soft texture is more frequently associated with a benign cause.

- mobility in relation to adjacent structures during swallowing, with fixation or adhesion to the surrounding tissues being a cause for concern.

- Dimensions and symmetry: estimation of size and comparison with the contralateral lobe are crucial, as any significant asymmetry could potentially signal a more advanced pathology.

In addition, exploration of the cervical lymph node areas is imperative to detect any adenopathies, the presence of which could indicate metastatic dissemination. A thorough patient history must also be taken to identify concomitant symptoms, such as dysphagia, dysphonia, or a recent change in nodular volume.

These initial clinical data are crucial for guiding the subsequent diagnostic strategy and form the basis of an exhaustive evaluation of the thyroid lesion.



Figure I.19: Palpation of the thyroid region .[66]

general examination

An investigation of any signs indicating excessive or deficient thyroid function was conducted.

A careful palpation of the cervical lymph nodes was performed.

The evaluation of the anatomical configuration of the neck focused on the potential presence of visible thyroid hypertrophy (goiter) or morphological irregularity.

Indicators of an inflammatory process, suggesting thyroiditis (e.g., localized erythema), were identified.

Imaging techniques:

You should start with an ultrasound to evaluate thyroid nodules, as it is the main imaging technique used.

The TI-RADS classification system will help you categorize nodules based on their echography characteristics.

This systematic approach guides the need for additional evaluation or biopsy. [11]

I.4.1 Additional exams: biology, ultrasound, cytology, scintigraphy ...

The paraclinical investigation is indispensable for:

- Confirm or refute the presence of a thyroid anomaly.
- Determine whether a nodular formation is benign or malignant.
- Guide the therapeutic strategy (observation, pharmacological treatment, or surgical intervention).

These investigations include biological analyses, morphological examinations, and functional evaluations.

Overall, these limitations demonstrate that no single diagnostic modality is sufficient on its own, thereby justifying the integration of advanced computational methods for improved accuracy and decision support.

I.4.1.1 Biological balance:

- Determine the activity of the thyroid gland.
- Identify the presence of an autoimmune etiology.
- Search for particular tumor markers.
- Exams included in the analysis:

Ultrasensitive dosage of thyrotropin (TSH):

- Precise quantification of the thyroid stimulating hormone.
- Initial examination for the detection of hyperthyroidism (decreased TSH) or hypothyroidism (increased TSH). See Figure I.20.

Dosage of triiodothyronine (T3) and thyroxine (T4) (free and total fractions):

Confirmation of a hormonal function disturbance.

Hyperthyroidism: increased T3/T4 levels and decreased TSH levels. Hypothyroidism: decreased T3/T4 levels and increased TSH levels.

Calcitonin:

It is a specific marker of medullary thyroid carcinoma (MTC).

Recommended for individuals with a hereditary risk or cytological suspicion.

Thyroglobulin (Tg):

Glycoprotein synthesized by thyroid follicular cells.

It is primarily used for postoperative follow-up after surgical removal of the thyroid to detect a recurrence of differentiated carcinoma.

Anti-thyroglobulin (anti-Tg) and anti-thyroperoxidase (anti-TPO) antibodies.

Auto-immunity indicators (Hashimoto's thyroiditis, postpartum thyroiditis, Graves' disease).

Their dosage completes the biological evaluation and can clarify hormonal anomalies.[12]

	Exploration Initiale	Enquête étiologique	Suivi
Goitre simple	TSH		TSH
Nodule isolé	TSH, (CT)		
Hypothyroïdies	TSH, (T4L)	anti-TPO, (Iodurie)	TSH
Hyperthyroïdies	TSH, (T3, T4 libres)	anti-TPO, anti-RTSH (Iodurie, Tg, HCG)	T4L, TSH (anti-RTSH)
Thyroïdites	TSH, VS, CRP	anti-TPO (sérodiagnostics viraux)	TSH, CRP
Cancers différenciés de la thyroïde	TSH		TSH, anti-Tg, Tg
Cancer médullaire de la thyroïde	TSH, CT		TSH, CT, ACE

Figure I.20: Laboratory assessment of the thyroid. [67]

I.4.1.2 Thyroid Ultrasound:

Determine the dimensions, quantity, structure, and echographic attributes of the thyroid nodule(s).

Identify nodules that are not palpable and assist with fine needle aspiration (FNA) biopsies.

Parameters studied:

Nature of the nodule: predominant solid, cystic, or mixed.

Echogenicity: hypoechoic, isoechoic, or hyperechoic.

Contours: clear, blurred, or lobulated (irregularity is a suspicious criterion).

Microcalcifications: Small calcifications, less than 2 mm, are indicative of papillary carcinoma.

Vascularization: Increased internal vascularization is a suspicious sign.

Use the size-to-volume ratio for follow-up.

Clinical relevance:

It categorizes nodules according to the TIRADS (Thyroid Imaging Reporting and Data System) system, which determines the need for a cytoponction.

It ensures the follow-up of benign nodules without recourse to puncture. See Figure I.21.

Disadvantages: Performance depends on the examiner and does not allow for the definitive differentiation of certain benign and malignant follicular lesions [13].

Consequently, this limitation highlights the importance of advanced image analysis techniques, particularly deep learning-based segmentation models such as U-Net, which will be explored in the following chapters.

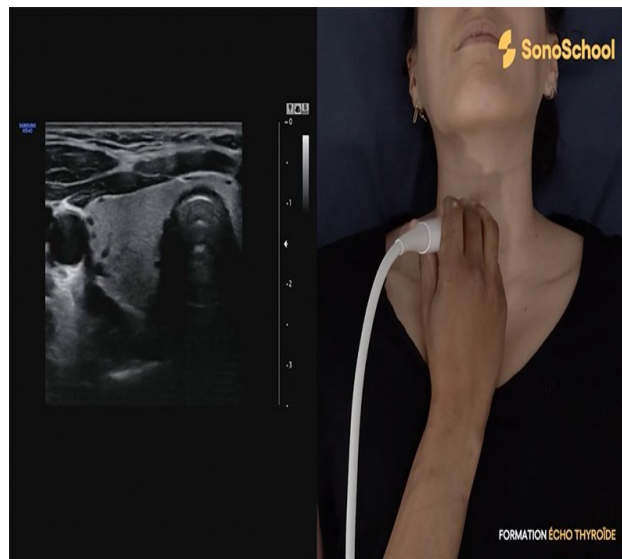


Figure I.21: Thyroid ultrasound .[68]

I.4.1.3 Cytology/Fine Needle Aspiration (FNA):

Obtain a histopathological characterization using a minimally invasive intervention method.

Objective: Distinguish benign, malignant, or uncertain nodular masses.

Method: The procedure is performed under ultrasound control for exact localization.

The specimens are analyzed and classified according to the Bethesda system:

I: Specimen inadequate for evaluation

II: Absence of malignancy

III: Cellular abnormality of uncertain significance

IV: Suspicious follicular neoplasm

V: High probability of malignancy

VI: Presence of malignancy

Indications:

Nodule mesurant au moins 1 cm qui présente des signes échographiques préoccupants.

Rapid increase in size or non-functioning nodules in scintigraphic examination.

Family history of thyroid neoplasia or prior exposure to cervical radiation.

Restrictions: Identification of follicular malignancy is often impossible without surgery [14].

This diagnostic limitation reinforces the need for complementary imaging-based approaches for improved characterization.

I.4.1.4 Scintigraphy:

Thyroid scintigraphy allows one to determine the functional activity of thyroid nodules by classifying them as hypercaptant (hot) or hypocaptant (cold). This contributes to the diagnostic evaluation of hyperthyroidism associated with nodular pathology. See Figure I.22.

The principle involves administering a radiopharmaceutical (technetium-99m or iodine-123) and subsequently imaging its distribution in the thyroid gland. Nodules with low isotopic uptake raise a higher suspicion of malignancy. Conversely, hypercapturing nodules are typically benign and associated with increased secretory activity.

However, this technique is used less frequently to explore euthyroid nodules, particularly when cytological analysis by fine-needle aspiration (FNA) is an option. [7]

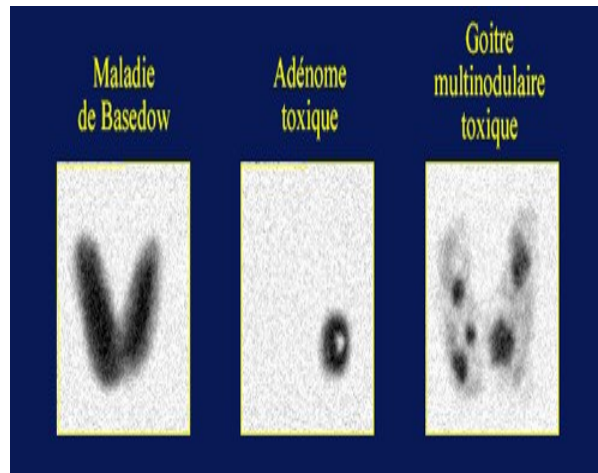


Figure I.22: Thyroid scintigraphy .[69]

I.4.1.5 Other Exams:

Cervical MRI:

Cervical magnetic resonance imaging (MRI) for evaluating nodular masses of significant size or located in a retrosternal position. This technique allows for evaluation of propagation towards the mediastinum and/or compression of surrounding anatomical elements. (See Figure I.23).

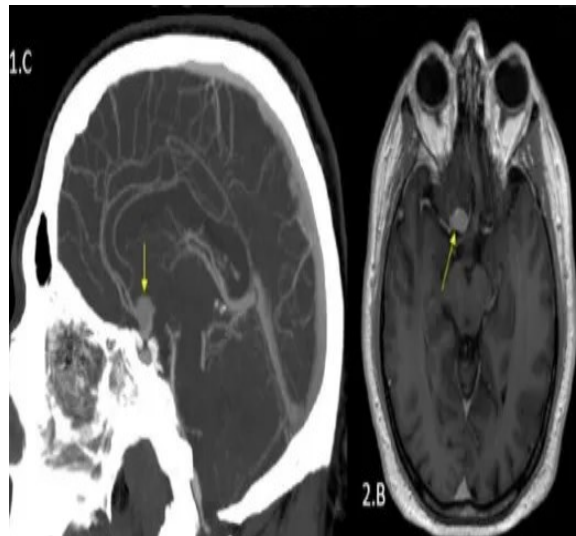


Figure I.23: Thyroid MRI .[70]

Ultrasound elastography:

Ultrasound elastography, a medical imaging modality, allows for the evaluation of the hardness or rigidity of a nodular lesion. There is a correlation between tissue stiffness and malignancy probability: nodules that are more resistant to deformation are generally considered more suspicious, while compressible nodules are more frequently associated with benign pathologies.

This technique is an additional diagnostic tool that complements conventional ultrasound. (See Figure I.24.)

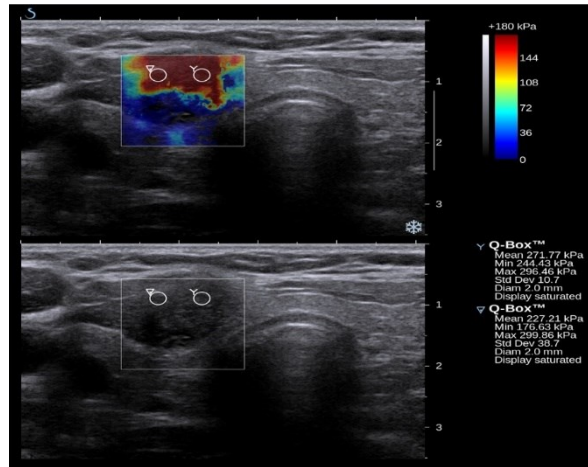


Figure I.24: Thyroid elastography .[71]

Ultrasound elastography enables quantitative assessment of tissue stiffness by analyzing shear wave propagation through the thyroid parenchyma according to Equation (I.1). Tissue stiffness is estimated from shear wave velocity using the following relationship:

$$E = 3\rho v^2 \quad (\text{I.1})$$

where:

- E represents the tissue elastic modulus (Pa or kPa),
- ρ represents tissue density (kg/m^3),
- v represents the shear wave propagation velocity (m/s).

Malignant tissues generally exhibit higher stiffness compared to benign tissues. In clinical practice, stiffness values below 20 kPa are often associated with benign nodules, whereas values above 35 kPa may be considered suspicious for malignancy. For example, a stiffness value of $E = 18 \text{ kPa}$ is typically consistent with a probably benign nodule. [15]

Scanner (CT scan):

Indicated sporadically, mainly for:

Large nodules

Goiters exerting a compression

In preparation for surgery, for interventional planning purposes. [15] See (Figure I.25)

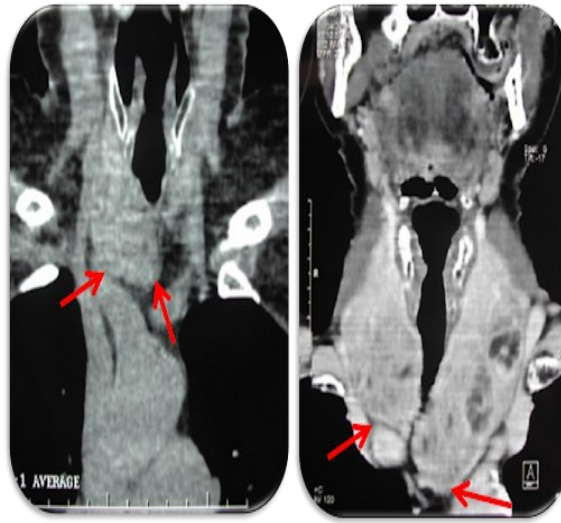


Figure I.25: Thyroid scan .[72]

I.4.2 Operative indications:

Surgical thyroid excision is indicated in the presence of thyroid nodules with a high probability of malignant transformation, significant compressive manifestations or endocrine dysregulation refractory to pharmacological management. The selection of the appropriate surgical technique is based on the dimensions of the nodule, its anatomical positioning, its isolated or multiple nature, as well as on the conclusions from additional diagnostic investigations.

I.4.2.1 Malignant nodules:

Indications for surgical removal of a thyroid nodule include: cytopathological findings according to the Bethesda V or VI classification (corresponding to high suspicion or confirmation of malignancy), ultrasound features suggesting high risk (microcalcifications, irregular margins, hypoechogenicity, intra-nodular vascularization), cold nodules on thyroid scan associated with increased risk of malignancy, and significant increase in the size of a previously stable nodule. The rationale for surgical intervention is to obtain a definitive histological assessment, limit neoplastic progression and minimize the possibility of disease recurrence. [14] See (Figure I.26)

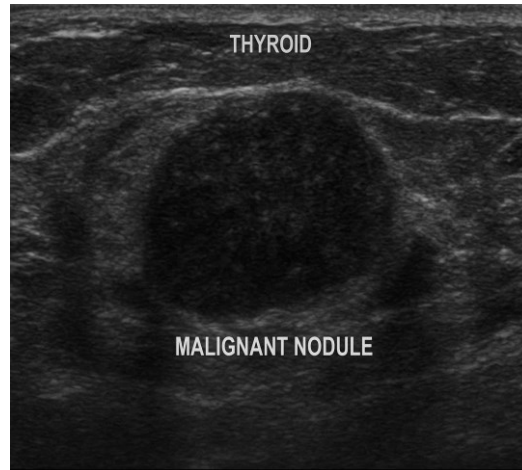


Figure I.26: Malignant nodule .

I.4.2.2 Symptomatic/compressive nodules:

Symptomatic or compressive thyroid nodules appear when they are large enough to exert pressure on the surrounding cervical or mediastinal structures. Surgery is the treatment of choice, as medical approaches fail to reduce volume significantly to lift compression.

- **Clinical signs requiring surgical intervention:**

Difficulties in swallowing (dysphagia) in the event of compression of the esophagus.

Respiratory problems (dyspnoea, stridor, choking sensation) due to tracheal compression, requiring urgent surgery in case of stridor or severe compression (>60%).

Voice alterations (dysphonia) by recurrent laryngeal nerve damage.

Significant neck pain or disabling mechanical discomfort.

Retro-sternal nodules or plunging goiters, extending behind the sternum, capable of compressing the trachea, esophagus, or superior vena cava (HSV syndrome). Surgical removal is often essential.

- **Objectives of the surgery:** is to relieve these symptoms, prevent serious complications (such as asphyxia or superior vena cava syndrome) and improve the patient's quality of life (comfort, aesthetics, reduction of anxiety). [8]

I.4.2.3 Hyperfunctional nodules resistant to medical treatment:

Overactive thyroid nodules (hot nodules or toxic adenomas) can cause chronic hyperthyroidism. Surgery is considered if drug or radioactive iodine treatments fail. When to operate?

Persistent hyperthyroidism despite ATS: The nodule continues to produce hormones autonomously, leading to risks of tachycardia, weight loss, exercise intolerance and heart rhythm disorders (e.g. atrial fibrillation).

Recurrence after discontinuation of medications: After stabilization by ATS, hyperthyroidism often recurs when stopped, because the drugs do not suppress the nodule. Surgery offers a lasting solution.

Serious cardiovascular or metabolic complications: Hyperthyroidism can lead to hypertension, arrhythmias (atrial fibrillation), heart failure, osteoporosis, or major weight loss. Surgery then quickly corrects thyrotoxicosis.

Purpose of the surgery: Permanently eliminate hormonal overproduction to normalize thyroid function, avoid cardiovascular/metabolic complications, and prevent severe thyrotoxic seizures. [10]

I.4.2.4 Other specific indications:

► High suspicion of MTC (medullary thyroid cancer):

Elevated calcitonin levels ($> 100\text{-}150\text{ pg/mL}$) strongly indicate CMT. Cytology may be uncertain. Surgery is necessary even if the cytopunction is not clear, especially with a very high calcitonin. A post-puncture calcitonin assay (wash-out) can help reinforce suspicion. The procedure aims to remove the entire thyroid and often the central lymph nodes.

► Nodules related to heredity or genetic syndromes:

Surgery is recommended for high-risk individuals to prevent cancer progression.

MEN2 Syndrome (RET mutation): CMT is almost certain and often aggressive (especially MEN2B). Preventive surgery can be advised from childhood for high-risk RET mutations.

Family history of thyroid cancer: Increased risk of malignancy, requiring careful monitoring and earlier surgical decision.

Other genetic syndromes (e.g. Cowden, DICER1, Carney): Surgery is indicated earlier to avoid malignant transformation.

► Patient preferences or impossibility of follow-up:

Anxiety or refusal of risk: Patients who are anxious or refuse any risk of cancer, even minimal.

Monitoring impossible: Large benign nodules requiring regular monitoring which is impossible (geographical constraints, access to care, social situation).

Refusal of biopsies: Patients refusing repeated biopsies (e.g. uncertain results, recurrent cysts).

The objective is to offer a definitive solution if the follow-up is difficult or not desired, and to avoid complications due to a large nodule or chronic anxiety. [16]

I .4.3 Surgical treatment:

Surgical intervention represents a significant therapeutic modality for thyroid nodules in the following situations: high suspicion of cancer, presence of notable compression symptoms or thyroid dysfunction induced by an overactive nodule. The selection of the appropriate surgical method is based on a multifactorial evaluation integrating ultrasound, cytological, clinical and functional criteria, as well as the patient's genetic profile and choices.

I.4.3.1 Types of surgical procedures:

I.4.3.1.1 Thyroid lobectomy:

Thyroid lobectomy is a surgical operation that involves removing a single lobe (part) of the thyroid gland as well as the isthmus (central part), through a small incision in the neck. This procedure protects the vocal nerve on the operated side and the parathyroid glands on the unaffected side. Faster than a complete thyroid removal, it aims to reduce complications while effectively treating unilateral thyroid problems. See (Figure I.27)

Main indications:

1. Suspicious or benign troublesome nodules: To obtain an accurate diagnosis without removing all the gland unnecessarily, or to relieve symptoms of compression.[7]
2. Localized small papillary cancers (1 cm): Often sufficient treatment if the tumor is confined and at low risk, thus avoiding other treatments such as radioactive iodine.[7]

3. Unilateral compressive nodules: In case of discomfort, pain or difficulty swallowing caused by a single-sided nodule.[7]

4. Isolated hyperactive nodule: To treat hyperthyroidism due to a single nodule, especially in young people or as an alternative to radioactive iodine.

Benefits:

- Maintenance of thyroid function: The majority of patients do not require lifelong hormonal treatment.
- Low risk of hypoparathyroidism: The risk of calcium problems is minimal.
- Less risk of severe vocal problems: Only one vocal nerve is exposed, reducing the risk of bilateral paralysis and severe breathing difficulties.
- Quick recovery: Shorter surgery and hospitalization, faster return to activity and less pain.
- Diagnostic aid: Allows confirmation of the nature of a suspicious nodule, with possible additional surgery if necessary.[7]

Limits:

- A second operation to remove the rest of the thyroid may be necessary if aggressive cancer is discovered.
- Only suitable for strictly unilateral conditions.
- Minimal risk of recurrence of nodules in the remaining lobe. [7]

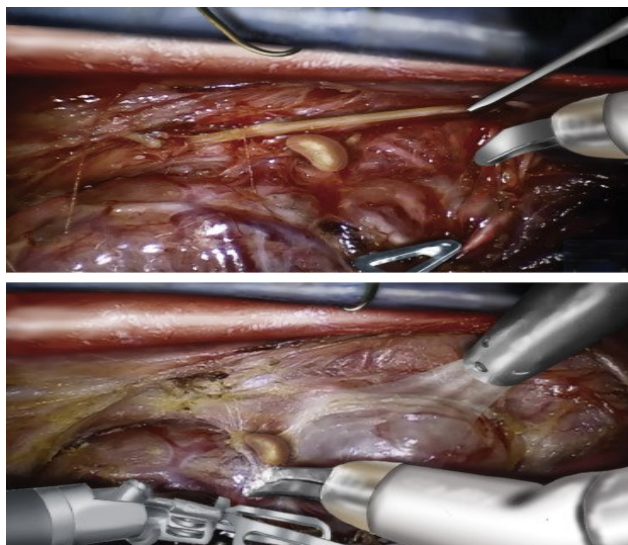


Figure I.27: Robotic transaxillary thyroid lobectomy surgical. [73]

I.4.3.1.2 Total thyroidectomy:

Total thyroidectomy is the complete removal of the thyroid gland (both lobes and the isthmus). It is the main treatment for bilateral, malignant thyroid diseases, or before radioactive iodine treatment. See (Figure I.28)

Main indications:

- 1- Thyroid cancers (papillary, follicular >1cm or aggressive, medullary, some anaplastic). It optimizes radioactive iodine treatment (RAI) and thyroglobulin monitoring.
- 2- Bulky bilateral multinodular goiters, especially if they cause compression (trachea, esophagus) or aesthetic problems, and if there is a significant risk of recurrence.
- 3- Severe or treatment-resistant hyperthyroidism, particularly Graves' disease or toxic nodules. It allows for quick and definitive control.
- 4- Radioactive Iodine Preparation (RAI): It prepares and optimizes the effectiveness of RAI to target remaining cancer cells, and facilitates the interpretation of follow-up examinations.

Surgical technique:

It involves a neck incision, careful dissection to remove both lobes and the isthmus, while identifying and preserving recurrent laryngeal nerves and parathyroid glands. Nervous monitoring is frequent.

Benefits:

Maximum control of the disease, a reduction in the risk of recurrence, more reliable post-operative follow-up (thyroglobulin), better effectiveness of IRA, and definitive resolution of compressive goiters.

Disadvantages and risks:

Higher than lobectomy, they include hypoparathyroidism (temporary or permanent), recurrent laryngeal nerve damage (rare but severe), lifelong hypothyroidism requiring hormonal treatment, and the risk of bleeding. These risks are lower in specialized centers.. [17]



Figure I.28: Total thyroidectomy .[80]

I.4.3.1.3 Thyroidectomy with lymph node dissection:

Thyroid surgery is often associated with removal of the lymph nodes in the neck (lymph node dissection) if metastatic involvement is confirmed by imaging or cytology, or strongly clinically suspected. This procedure is essential in the treatment of thyroid cancers, particularly papillary and medullary carcinomas, which have a high risk of spreading to the lymph nodes. See (Figure I.29)

Principles of Cervical Lymph Node Dissection:

Lymph node dissection aims to remove potentially affected nodes in order to eradicate existing metastases, prevent local recurrences, refine the diagnosis (staging) and optimize post-operative management (biological monitoring). For papillary cancers, 30 to 70% of patients have unapparent metastases in the central compartment; this figure reaches 70 to 90% for bone marrow carcinomas, which justifies a systematic cure in the latter case.

Types of Curages:

- 1- Central Flushing (Compartment VI): Concerns the ganglia located around the trachea and larynx. It is indicated for papillary cancer with confirmed central metastases, voluminous papillary tumors, and systematically for medullary cancer. It reduces the risk of local recurrence, but carries risks of hypoparathyroidism or vocal nerve injury.
- 2- Lateral Flushing (Compartments II-V): Involves the removal of ganglia along the vessels of the neck. It is performed only if lateral metastases are clearly identified and confirmed (e.g. by ultrasound and puncture). Unlike central flushing, it is never performed preventively for papillary cancer. Its aim is to eradicate extensive metastases and improve disease control.

Detection and Indications:

Ultrasound of the neck is crucial for identifying suspected lymph nodes (e.g. disappearance of the hilum, microcalcifications, spherical shape, abnormal vascularization). A puncture (cytopunction) with thyroglobulin analysis in the washing liquid allows to confirm the presence of cancer cells.

Specific indications vary:

Papillary Carcinoma: Central removal if central metastases are confirmed, lateral removal only if lateral metastases are confirmed. No preventive lateral cleaning.

Medullary Carcinoma: Quasi-systematic central flushing. A bilateral lateral dissection is considered if the calcitonin level is very high, and a more extensive lateral dissection if metastatic ganglia are found.

Anaplastic Carcinoma: Discharging is rarely indicated and the decision is multidisciplinary.

Benefits and Risks: the benefits of curage include a significant reduction in local recurrences, better control of the disease, more reliable interpretation of post-operative tumor markers, and improved efficacy of radiotherapy with I-131. These risks are carefully weighed against the oncological benefits, particularly in high-risk cancers. Possible complications include hypocalcemia (due to injury or devascularization of the parathyroid glands), recurrent laryngeal nerve injury, hemorrhage, and infection, as well as, more rarely, the development of a chyle leak or cervical lymphocele following lateral neck dissection. [18]

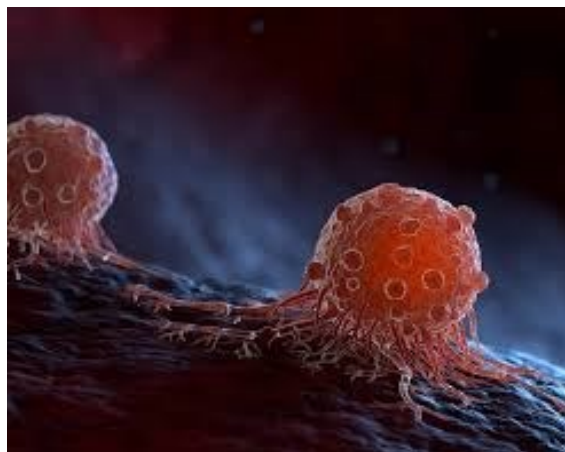


Figure I.29: Lymph node dissection in thyroid surgery. [74]

I.4.3.2 Conduct of the operative act:

Thyroid surgery (lobectomy or total thyroidectomy) is performed under general anesthesia to ensure the immobility and comfort of the patient. It follows precise standardized steps to safely remove the gland while preserving the vital cervical structures [19].

- 1- Positioning and preparation: The patient is lying on his back, head slightly back to expose the thyroid. The neck is disinfected and covered with sterile sheets. This positioning facilitates vision, dissection, and reduces the risk of injury. [19].
- 2- Incision and exposure: A 4-6 cm horizontal incision is made in a natural neck fold (Kocher incision). It offers good access to the thyroid, is discreet and can be enlarged if nodes need to be removed. After the skin, the platysma muscle is incised. The deep muscles are spread to reach the thyroid. This step is delicate to avoid bleeding and trauma, which could cause post-operative pain. [19].
- 3- Dissection of the gland: Dissection begins at the top, where the blood vessels of the thyroid are located, ligated and cut. The gland is then gradually detached. [19].
- 4- Identification and protection of essential structures[19].

Critical anatomical structures are carefully identified and preserved:

The recurrent laryngeal nerve: Essential for the voice, it is located, dissected and sometimes monitored to avoid any lesion that could affect the voice or breathing.

Parathyroid glands: Located behind the thyroid, they are protected to maintain blood calcium levels. If their irrigation is compromised, they are autografted into a muscle.

- 5- Thyroid removal: Depending on the procedure (removal of a single lobe or the entire gland), the thyroid is carefully removed to limit bleeding and protect surrounding tissues. [19].
- 6- Bleeding control and closure: The bleeding is meticulously controlled to avoid a hematoma. The muscles are put back in place, then the skin is tightened aesthetically. A drain may be placed if the risk of bleeding is significant, but it is less and less frequent. [19].
- 7- Duration of the operation: The duration of the operation varies: 45 to 60 minutes for a lobectomy, 60 to 120 minutes for a total thyroidectomy, with an additional 30 to 60 minutes if lymph nodes are removed. [19].
- 8- Hospitalization and follow-up: Hospitalization is short (24 to 48 hours without complications). Post-operative follow-up includes monitoring of calcium, voice,

monitoring for a possible hematoma, assessment of pain and scar. The patient can resume his usual activities quickly. [19]

I.4.3.3 Peri-operative and post-operative complications:

Thyroid surgery is now a relatively safe procedure, with a low incidence of complications in expert facilities. Nevertheless, various particular complications must be brought to the attention of the patient, described in detail, and carefully monitored after the operation.

I.4.3.3.1 Recurrent laryngeal nerve injury:

The recurrent laryngeal nerve is crucial for vocal cord function; its involvement can cause voice disorders. The causes of this impairment include stretching, compression, thermal damage or accidental (often surgical) section. See (Figure I.30)

The risk is increased during complete thyroid ablation, in the presence of infiltrating cancer, after repeated surgeries (scars), or during a cervical lymph node dissection.

Symptoms include frequent dysphonia (hoarse voice, faint, tiring), often temporary. Transient unilateral weakness is observed in 2-5% of cases. A rare bilateral involvement can cause respiratory distress.

Management relies on voice monitoring. Orthophonic rehabilitation is indicated in cases of slow recovery, and rarely secondary surgery may be considered.

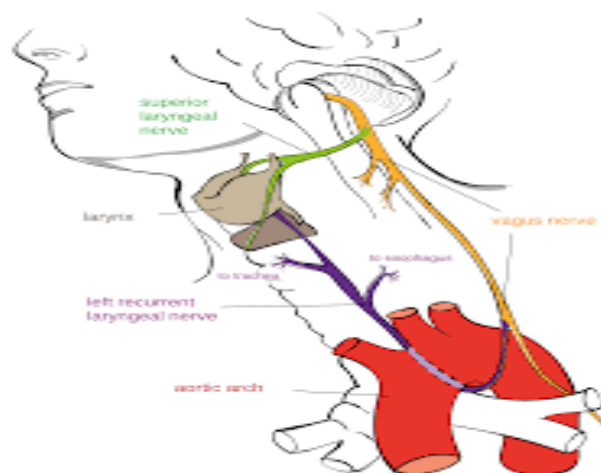


Figure I.30: Anatomy of the laryngeal nerve .[81]

I.4.3.3.2 Hypocalcemia / hypoparathyroidism:

Hypoparathyroidism after cervical surgery can result from:

- Damage to the vessels of the parathyroid glands during the operation.
- Their accidental removal.
- A temporary disruption of their operation.

Types:

- Transient: The most common form (20-30% of cases).
- Permanent: Rare (<2% in specialized centers).

Events:

- Tingling around the mouth, muscle spasms, fasciculations and cramps.
- Possibility of signs of Trousseau and Chvostek on the exam.

Management:

- Calcium supplementation, sometimes with vitamin D.
- Close monitoring of blood calcium levels during the 48 hours post-surgery.
- Continuous follow-up if symptoms persist.

Compressive hematoma:

Compressive hematoma constitutes an unusual but potentially severe complication in the immediate postoperative period. See (Figure I.31)

It results from a hemorrhage following an inadequate or defective vascular ligature. Its clinical manifestation generally occurs within the first few hours following surgery.

Clinical Signs Precursors

Warning signs include: a feeling of tension in the operated area, cervical edema, sudden onset of dyspnea, and dysphonia.

Conduct to be Held: Medical Imperative

Due to the risk of upper airway obstruction, compressive hematoma requires immediate management. In case of respiratory distress, an opening of the wound at the patient's bedside is imperative, followed by emergency surgical reintervention to control the bleeding.



Figure I.31: Thyroid hematoma secondary to blunt cervical trauma.[82]

I.4.3.3.3 Infection:

Surgical site infection following thyroidectomy is an uncommon complication (incidence less than 1%).

Associated risk factors:

- Preoperative glycemic imbalance.
- Pre-existing cervical hematoma.
- Contamination of the wound, although this situation is unusual given the peculiar nature of the surgical procedure.

Clinical manifestations:

- Erythema, increased local pain, exudation.
- Moderate increase in body temperature.

Conduct to be taken:

- Antibiotic therapy oriented according to microbiological data.
- Wound care.
- Surgical evacuation if presence of a fluid collection. [20]

I.4.3.4 Post-operative care:

Post-surgical management is contingent on the nature of the intervention (lobectomy compared to complete thyroidectomy) and the justification of the procedure. The primary objectives are the promotion of efficient convalescence, the minimization of long-term sequelae and the establishment of appropriate endocrine monitoring.

Endocrine monitoring:

Following the excision of a single thyroid lobe, the contralateral lobe usually compensates by ensuring adequate hormone production. A measure of thyrotropin-stimulating hormone (TSH) is taken within 6 to 8 weeks post-surgery to assess remaining thyroid function.

TSH level within normal range No hormone substitution is required.

High TSH level (thyroid hypofunction) Initiation of low dose levothyroxine therapy. In some situations, temporary hypothyroidism may be observed; reassessment is advised after a few weeks.

Structural (or morphological) monitoring

Ultrasound imaging of the neck may be warranted in the presence of:

Residual nodule tissue on the opposite side of the excision,

Unusual cytological features,

Family history with an increased risk of thyroid pathologies.

I.4.3.4.1 After a total thyroidectomy:

Post-Chirurgical Management of the Thyroid: Hormonal Treatment and Monitoring

Administration of levothyroxine begins immediately after surgery. The treatment goal (TSH level) is adjusted according to the patient's clinical situation.

Benign conditions: the goal is to maintain TSH within normal values (approximately 0.5–2.5 mIU/L) to restore balanced thyroid function without causing hyperactivity.

Differentiated thyroid cancers (papillary or follicular): according to the guidelines, TSH must be lowered or suppressed in order to limit tumor growth.

Low risk: TSH around 0.1–0.5 mUI/L

Intermediate risk: TSH less than 0.1 mUI/L

High risk: TSH very low (< 0.01 mUI/L)

Specific Post-Cancer Monitoring:

- Differentiated tumors: followed by thyroglobulin (Tg) and its antibodies, and cervical ultrasound scans (initially 6-12 months, then regularly depending on the risk). After a total thyroidectomy, the Tg must be undetectable.
- Medullary carcinoma: measurement of calcitonin and ACE. A persistent increase in these markers suggests the presence of residual or metastatic disease.

Other aspects of monitoring:

- Calcium control and vitamin D intake if transient hypoparathyroidism occurs.
- Assessment of vocal function if laryngeal nerve has been affected.
- Examination of the evolution of the scar.

I.4.3.5 Objectives of the surgical treatment:

Thyroid surgery is a therapeutic option (curative or palliative) whose precise objectives vary according to the nature of the problem.

- 1- Treatment of Malignant or Suspicious Lesions: It aims to cure thyroid cancers or highly suspect nodules, offering a complete diagnosis and reducing the risk of recurrence.
- 2- Compression relief: It helps to resolve breathing difficulties, swallowing or nerve irritations due to voluminous nodules.
- 3- Refractory Hyperthyroidism Correction: It restores normal thyroid function when hyperactivity is persistent and does not respond to other treatments, thus preventing heart complications.
- 4- Genetic Cancer Prevention: It is practiced preventively in people with a high genetic risk of developing thyroid cancer.
- 5- Permanent Solution in Difficult Follow-up Cases: It offers a lasting response to patients with regular follow-up difficulties or whose benign nodules cause significant anxiety. [21]

I.4.4 Conclusion of chapter I :

This chapter provided a comprehensive review of the thyroid gland and thyroid nodules, addressing their structure, function, clinical manifestations, diagnostic investigations and surgical interventions. Ultrasound, integrated into the TI-RADS system, remains the predominant imaging modality for diagnosis, supported by fine needle cytopathology and biochemical analyses for accurate risk categorization. Despite the advances made, conventional approaches present restrictions, such as subjectivity of interpretation and indeterminate diagnoses, highlighting the need for innovative methods. Artificial intelligence is emerging as a promising avenue to automate ultrasound analysis, mitigate human biases, and increase diagnostic accuracy. The next chapter will focus on the fundamentals of ultrasound imaging and computer-aided diagnostic systems applied to the study of thyroid nodules.

CHAPTER II : ULTRASOUND IMAGING AND COMPUTER-AIDED DIAGNOSIS

II.1 Introduction:

Ultrasound is the first imaging test used to examine thyroid nodules. It is privileged because it is easily accessible, without danger, and offers very precise images to observe the structures close to the surface of the neck. It allows to describe in detail the appearance of nodules, by looking at their "color" on the image (echogenicity), their edges, their shape, the presence of small calcium deposits and their blood supply (vascularization). All these details are then used in risk classification systems, such as EU-TIRADS and ACR TI-RADS. [7] [9] [13]

However, although ultrasound is very useful for diagnosis, the interpretation of images depends a lot on the experience of the person doing the examination. This means that the results may vary from one doctor to another, or even for the same doctor at different times. Nodules with characteristics that are neither clearly benign nor clearly suspect are difficult to assess. This can lead either to too many unnecessary biopsies (cytopuncations) or, conversely, to delaying a necessary treatment. [22]

It is in this context that computer-aided diagnosis (CAD) systems have appeared. These are additional tools designed to make ultrasound assessment more consistent and accurate. At first, these systems used traditional image processing techniques, but they have evolved to integrate methods based on artificial intelligence (AI), notably machine learning and deep learning. [23] [24] [25] [26]

The purpose of this chapter is to present the physical basis of ultrasound imaging, the signs to look for on thyroid nodules, and the functioning and structure of CAD systems. We will particularly focus on methods of segmentation (to delimit nodules), extraction of characteristics (to extract useful information) and classification (to categorize them). These are the essential steps to design a diagnostic assistance system based on artificial intelligence.

II.2 Basics of ultrasound imaging:

Ultrasound is an imaging technique that uses very fine sound waves. A small device, called a probe, sends them into the body. These waves travel through tissues and are partially reflected when they encounter an organ or a different area. The part of the wave that returns to the probe (which captures it) is then processed to create the image of the body part examined. See (Figure II.1)

For the evaluation of small formations or growths within the thyroid gland, commonly called thyroid nodules, this technique is positioned as the preferred diagnostic approach. Its pre-eminence is explained by several major assets: it offers a very high image accuracy, remarkable clarity of details, it is completely risk-free for the patient because non-invasive, and its cost remains relatively accessible, making it widely available.

It is therefore fundamental to have a thorough understanding of the physical laws governing the behavior of ultrasound, to master the various methods of acquisition and presentation of echographic images, as well as to accurately identify the characteristic visual features of thyroid nodules. This comprehensive understanding is essential for identifying the imperfections and potential biases of the analysis performed by the human operator and, consequently, to fully legitimize the integration and use of diagnostic assistance tools based on advanced technologies. [27]

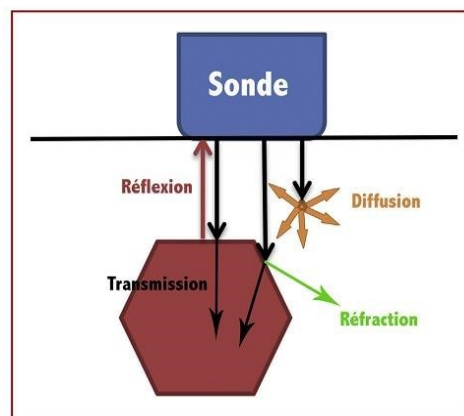


Figure II.1: Physical phenomena underlying ultrasound image. [28]

II.2.1 Physical principles of ultrasound and imaging modes:

II.2.1.1 Physical principles:

Ultrasound is a gentle medical examination method that does not require opening the body and does not hurt. It is used to observe the internal organs in humans and animals. This technique works by sending very high frequency sound waves, called ultrasound. The "height" of these waves (their frequency, measured in megahertz) changes depending on the part of the body being examined (for example, from 1 to 20 megahertz for most organs, and up to 50 megahertz for the eye).

The ultrasound imaging system consists essentially of a transductive probe, a computer processing unit and a display screen. The probe generates ultrasound waves in the anatomical

region of interest, whose frequencies are specifically selected according to the diagnostic objective:

- 1.5 – 4.5 MHz: These low frequencies are used for the investigation of deep organs, such as those in the abdomen, allowing a millimetre-scale spatial resolution.
- 5 MHz: Medium frequencies are dedicated to the visualization of intermediate depth structures, such as the core, offering a submillimetric resolution.
- 7 MHz: Higher frequencies are reserved for the evaluation of superficial anatomical structures, including blood vessels (veins and arteries).
- 10 – 18 MHz: This frequency range is suitable for skin imaging, small parts or the study of small organisms.
- Up to 50 MHz: Very high frequencies are specifically applied to ophthalmic ultrasound.

Prior to data acquisition, an acoustic coupling agent (gel) is applied to the epidermis in order to ensure an optimal acoustic interface between the transducer and the body surface, minimizing reflection and attenuation of the ultrasonic beam at this interface. The transducer, through direct contact with the body surface, facilitates real-time visualization of ultrasound images on the display unit, thus providing a reliable and non-invasive diagnostic method.

Ultrasound imaging relies on the exploitation of several fundamental properties of ultrasound:

- Celerity: The propagation speed (celerity) of ultrasound is intrinsically linked to the properties of the medium being crossed. It is about 340 m/s in air, approximately 1540 m/s in aqueous biological media (soft tissues, water), and significantly higher in bone structures.
- Absorption: The attenuation of the ultrasonic beam by absorption is directly proportional to the square of its frequency ($\alpha = K f^2$). Thus, high frequencies lead to increased absorption, which restricts beam penetration and consequently the exploration of deep structures.
- Partial reflection: When an ultrasound interacts with an interface between two media exhibiting a difference in acoustic impedance, it produces a partial reflection of the incident wave, while the rest is transmitted.

The transducer operates sequentially in transmitter and receiver mode over extremely short time cycles. The echograph then exploits two crucial parameters for image reconstruction:

- Echo propagation delay: This parameter determines the depth of the reflective interface, assuming a constant average celerity of ultrasound in soft tissues (about 1540 m/s).

Echo amplitude: It is proportional to the difference in acoustic impedance between adjacent tissues. Interfaces with high impedance contrast (e.g. soft tissue/bone, soft tissue/air) generate echoes of large amplitude, manifesting on the display through hyperechoic areas (clear or bright).

The electrical signal resulting from receiving echoes is then amplified and digitized for display on current digital displays. By integrating the propagation time, amplitude and velocity of echoes, ultrasound provides accurate information about the intrinsic characteristics, spatial location and dimensions of the tissues examined. [28]

II.2.1.2 Imaging modes:

The ultrasound imaging modalities directly reflect the technological advances observed in modern instrumentation. Current transducers operate universally in real-time and incorporate the ability to merge Temporal-Motion (TM) and Doppler modes. Nevertheless, understanding the previous techniques remains relevant to understand the fundamental principles of ultrasound image generation. [27]

II.2.1.2.1 Mode A (Amplitude):

A-mode was the first imaging method used with single (single element) probes. It gives an image in a single dimension. Each sound that comes back is shown by a tip (a peak): the height of this tip indicates the strength of the sound, and its location on the horizontal line corresponds to the depth where the tissue boundary lies. [27]

II.2.1.2.2 Mode B (Brightness):

In B-mode ultrasound, instead of seeing peaks, the device shows dots. The stronger the echo, the brighter the dot. This is how the shades of gray that are found in modern ultrasound scans are created. In the image, the depth is indicated from top to bottom: the elements near the probe are at the top, and those further away are at the bottom. This offers a representation of the organs much more visual and easy to understand than mode A. [27]

II.2.1.2.3 M and TM Mode (Time Motion):

These methods allow to see how the different parts move with time, which is very convenient to understand the heartbeat. The device (probe) remains at a fixed location, and changes in the distance of signals it returns show the movement of observed surfaces. In the image, the vertical axis (the one that goes up and down) indicates the depth. [27]

II.2.1.2.4 Bidimensional imaging (B-Mode or BD):

Two-dimensional (2D) imaging was a great progress, as it allowed to see the interior of an organ in "slices" See (Figure II.2) . For this, the probe (the device that scans), which had only one sensor at first, was moved by hand to explore the chosen area. The returned signals are then assembled to create the final image. In this image, the depth is shown from top to bottom, and the position of the sweep from left to right. [27]



Figure II.2: B-mode ultrasound in a 128 grayscale scale — transverse section of the left thyroid lobe. [27]

II.2.1.3.5 Real time:

Nowadays, all ultrasound scans are done live, thanks to special probes that can explore a complete area without having to move them by hand. Thus, one can see at the same time the shape and movement of the organs, which makes the examination much more accurate and fast.

II.2.1.2.6 Three-dimensional imaging:

Today's probes have many small sensors, spread over a larger area. This allows creating a three-dimensional image (a volume) instead of a simple flat image. This method is mainly used during pregnancy to examine the baby. But this clearly shows that ultrasound can give very fine and detailed 3D images. See (Figure II.3) [28]



Figure II.3: 3D Ultrasound.[28]

II.2.1.2.7 The Doppler echo:

In the field of diagnostic imaging, a flow of ultrasonic waves, emitted by a transducer, penetrates biological tissue structures and encounters various acoustic discontinuities. When the intercepted structure is static, the frequency of the returned waves remains unchanged compared to the initial emission frequency.

On the other hand, if the entity presents kinetics, like blood corpuscles in intravascular circulation, the frequency of the reverberated signal undergoes an alteration.

The difference between the transmission frequency (F_e) and the reception frequency (F_r), referred to as the Doppler frequency (Δf), provides the ability to quantify the velocity and orientation of the hematic current. The fundamental equation indicates that the Doppler frequency variation (Δf) is dependent on several factors:

- the frequency of the emitted ultrasound (F_e),
- the celerity of the figured components of blood (V),
- the angle of incidence (θ) formed between the ultrasound beam vector and the longitudinal axis of the vessel,
- the velocity of propagation of ultrasonic waves within tissue matrices (approximately 1540 m/s).

The angle θ is a crucial factor:

- When this angle is 90 degrees, the Doppler effect is nonexistent ($\cos\theta = 0$).
- At 0 degrees, the Doppler shift amplitude reaches its optimum ($\cos\theta = 1$).

Therefore, a rigorous assessment of blood velocity requires an exact determination of this angle.

► Continuous Doppler: uses a two-crystal configuration: one ensures continuous transmission, while the second is dedicated to receiving the reflected signal.

This method allows the analysis of flows at extremely high velocities, without presenting intrinsic measurement limits. It is also characterized by a pronounced sensitivity, proving particularly effective for slow flows.

Nevertheless, its main limitation lies in the lack of precise spatial location of the examined structure.

► Pulsed Doppler: the transducer generates pulse sequences and proceeds to analyze return signals during inter-pulse intervals. The pulse repetition frequency (PRF) is a function of the depth of investigation:

A weak FRP is preferred for in-depth investigation and identification of low velocity flows.

High PRF is required for the examination of high velocity flows, contributing to the reduction of aliasing phenomenon.

This modality confers a precise spatial determination of the examined vascular structure and is frequently associated with B-mode imaging, thus constituting a duplex system.

The main limitation of this approach is characterized by a lower sensitivity for flows of extremely low velocity, compared to the continuous Doppler technique.

► Color Doppler technology integrates two-dimensional visualization with real-time dynamic blood flow mapping. In accordance with the established conventions:

- The flows approaching the transducer are represented by a red coloration.
- The flows moving away from it are visualized by a blue coloration.
- Increased velocities result in brighter shades of color.

This modality requires high-velocity data processing for a large volume, in order to ensure immediate spectral analysis. See (Figure II.4) [27]

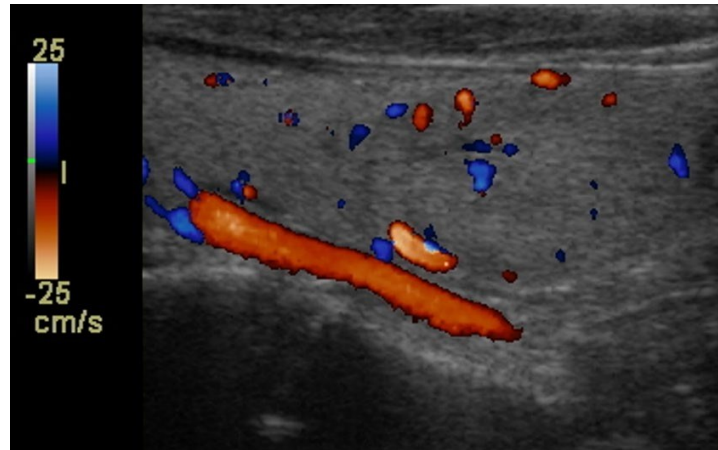


Figure II.4: color ultrasound. Longitudinal section of a lobe. The color coding indicates the direction of blood flow: in this example, blue for flows directed upwards, red for those directed downwards. The lighter the color, the faster the flow (left scale in the image).[27]

II.2.2 Echographic semiology of thyroid nodules:

The thyroid, due to its superficial anatomical location, is easily accessible for graphic ultrasound examination. Therefore, ultrasound represents the first-line diagnostic modality for the evaluation of thyroid nodular conditions. [27]

II.2.2.1 Equipment required:

The conduct of this review requires:

An ultrasound imaging system equipped with a high frequency linear transducer (greater than 12 MHz), ensuring optimal spatial definition.

A color Doppler module, intended for vascularization analysis.

A module of elastography (optional), likely to optimize the assessment of malignant character, although its use requires specialized training and generates additional cost.

The linear transducer constitutes the reference instrument, offering a rectangular-shaped field of exploration. Its dimensions must be adjusted to the cervical morphology of the examined subject. An undersized transducer restricts complete observation of the gland.

A micro-convex type transducer (similar to vascular probes) may be relevant for the investigation of the sub thyroid region, due to its triangular shape acquisition field.

II.2.2.2 Conduct of the review:

The evaluation is carried out on an individual positioned in dorsal decubitus, with the cervical region placed in hyperextension. This consists of a series of methodical phases:

Step 1: Determination of thyroid volume

The total glandular volume is calculated by summing the volumes of the two lobes and the isthmus.

Each lobe is modeled as an ellipsoid of revolution, whose volume (V) is obtained by the equation (II.1): See (Figure II.5)

$$V=L \times l \times e \times \pi/6 \quad (II.1)$$

where:

L = length

l = width

e = thickness

$\pi/6 \approx 0,52-0,56$



Figure II.5: volume calculation of a thyroid lobe.[27]

Contemporary ultrasound equipment incorporates automated calculation functionality of this volume.

The volume should not exceed 18 cm³ in the patient.

The volume should not exceed 20 cm³ in the patient.

The physiological volume of the isthmus is less than 1 cm³. See (Figure II.6)

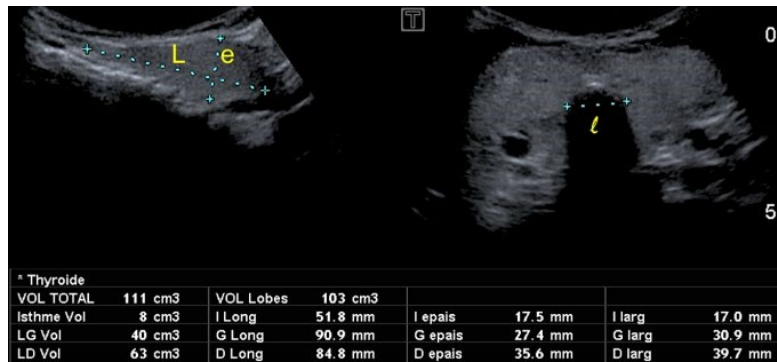


Figure II.6: calculation of the volume of the isthmus: the echograph automatically calculates the volume of the two lobes and the isthmus and automatically sums the three to obtain the total thyroid volume, here very enlarged to more than 110 cm³[27]

Step 2: Analysis of the echogenicity of the parenchyma[35]

The echogenicity of thyroid tissue is assessed by comparing it to adjacent pre-thyroid muscle structures.

Normal thyroid: It exhibits higher echogenicity than muscles, indicating a positive parenchyma-muscle gradient. See (Figure II.7)

Diffuse hypo-echogenicity: This sign is suggestive of thyroiditis or Graves' disease.

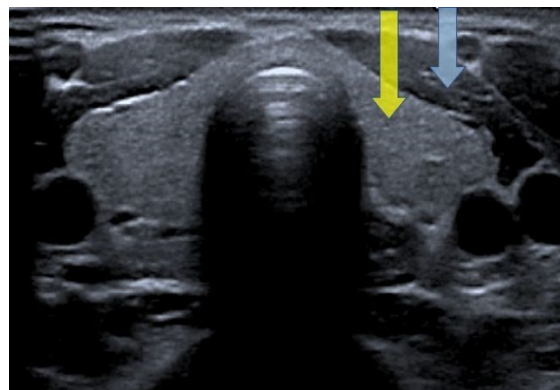


Figure II.7: parenchyma-muscle gradient: the parenchyma (yellow arrow) is more echogenic than the prethyroid muscles (blue arrow) [27]

Step 3: Texture echo assessment

This step consists of assessing the homogeneous or heterogeneous nature of the texture of the parenchyma. Heterogeneity is frequently observed in chronic thyroiditis.

Step 4: Examination of the glandular contours

The contours are usually regular in physiological condition. Wavy, festooned or irregular contours can be observed in the presence of an inflammatory process or a tumor syndrome.

Step 5: Appreciation of vascularization

Vasculature is analyzed using color Doppler imaging.

Normal state: Moderate intra-thyroid vasculature is typical.

The interpretation of vasculature is dependent on the adjustment parameters of the device (pulse repetition frequency (PRF), filters, color gain). It is recommended to use a standardized adjustment profile in order to ensure the reproducibility of the examinations.

Step 6: Systematic search for nodular formations

This investigation is based on a methodical exploration of the lobes and the isthmus, carried out according to transverse and longitudinal sections, encompassing the sus- and subthyroid areas. Each nodular lesion identified requires evaluation according to the following five criteria:

- 1- Detection: Systematic identification of any focal formation.
- 2- Location: The precise location must be specified, including: See (Figure II.8)
 - The laterality
 - The division into thirds (upper, middle, lower)
 - The anteroposterior position
 - The medio-lateral position

A three-dimensional schematic representation optimizes evolutionary tracking.

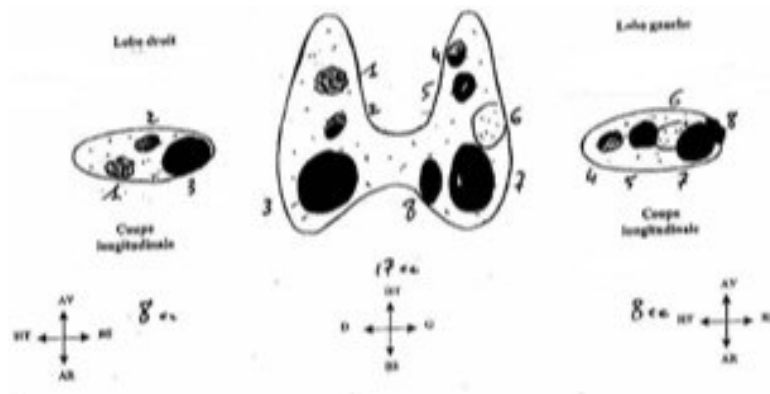


Figure II.8: nodular location diagram. Each nodule is perfectly located in space. [27]

1- Measurement: The determination of three orthogonal diameters and volume is required. The reproducibility of these measurements is fundamental for proper monitoring.

2- Morphological description: The morphological characteristics to be described include:

Liquid (cystic), tissue (solid) or composite (mixed) nature.

- Echogenicity: Hyperechoic, isoechoic, hypoechoic, or very hypoechoic.

- Shape: A profile wider than high is generally considered benign, while a profile higher than wide is an indicator of suspicion. See (Figure II.9)

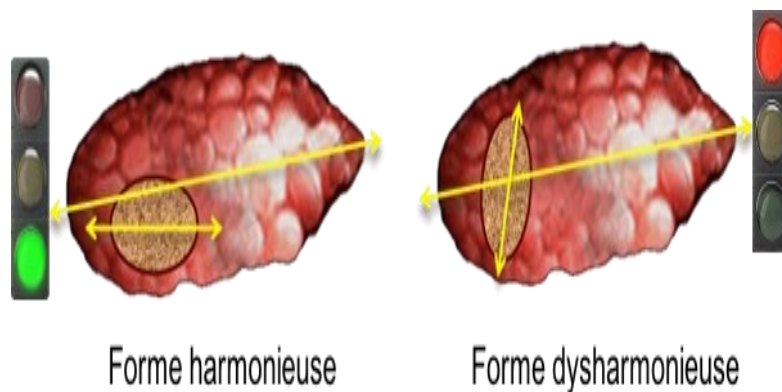


Figure II.9: diagram of the shape of a nodule according to the implantation of the nodule in the lobe. [27]

- Contours: Regular, indistinct, micro lobulate, or spiculate. See (Figure II.10)

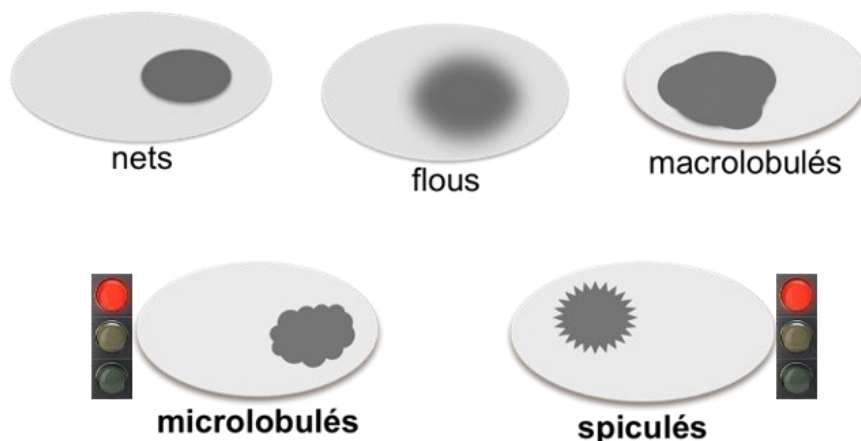


Figure II.10: pattern of nodular contours: only micro-lobulated and spiculated contours are signs of strong suspicion. [27]

- Calcifications: The presence of microcalcifications confers a high degree of suspicion. Continuous and peripheral macro-calcifications are reassuring signs. See (Figure II.11)

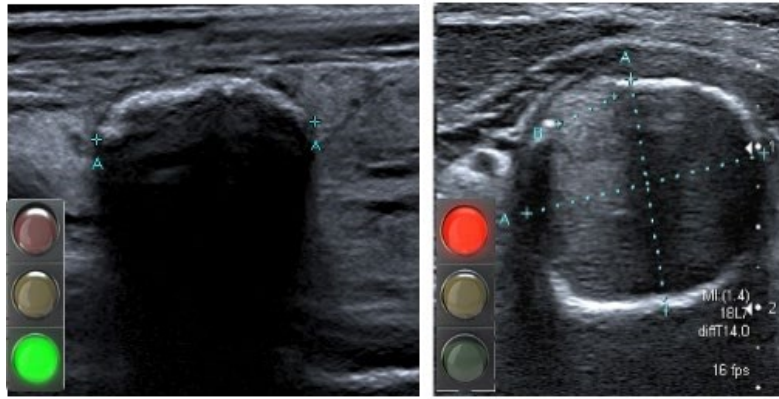


Figure II.11: left image: continuous peripheral "eggshell" macrocalcification– right image: "broken eggshell," more suspicious aspect .. [27]

- Vascularization: Appreciated at the Doppler. See (Figure II.12)

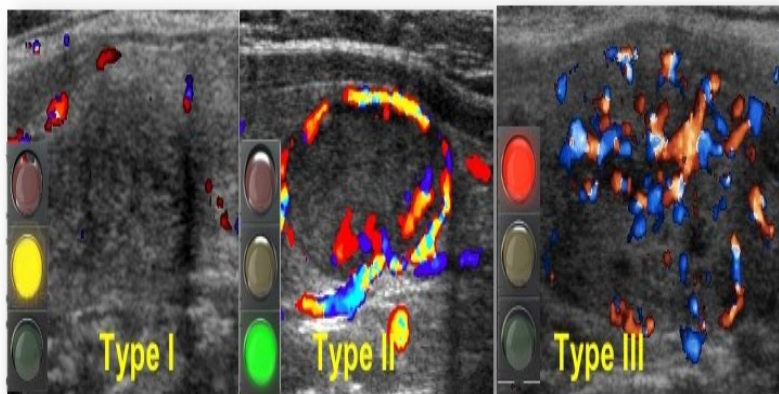


Figure II.12: the three types of nodular vasculature. [27]

- Hardness (assessed by elastography): An increased hardness of the nodule is a criterion of suspicion.

5- Assessment of the risk of malignancy: This assessment is based on the application of TIRADS classification systems.

II.2.2.3 Ultrasound Risk of Malignancy Prediction Scores:

The assessment of the risk of malignancy is based on the analysis of previously exposed ultrasound morphological characteristics.

Risk classification is carried out using standardized classification frameworks, including the American College of Radiology TI-RADS and the European Thyroid Association EU-TIRADS, as detailed in Chapter I.

These rating scales allow to standardize the clinical management and guide the indication of fine needle cytopunction (FNA) according to the degree of suspicion identified by the ultrasound examination.

II.2.3 Limits of human interpretation and interest of standardized systems:

II.2.3.1 Limits of human interpretation:

The ultrasound examination of balls (nodules) in the thyroid is based on the observation of several visual characteristics (their clarity, their edges, their shape, the presence of calcium deposits, their content and their blood supply). Even if there are clear signs, interpretation of the results depends in part on who is conducting the examination.

Here are the main difficulties:

- **Differences between observers:** Doctors do not always interpret things in the same way.
- **Differences in the same observer:** The same doctor may have slightly different opinions on the same examination at different times.
- **The role of experience:** The more experienced the specialist is in thyroid ultrasound, the better his diagnoses.
- **Shared characteristics:** Some balls, whether harmless (benign) or dangerous (malignant), can resemble ultrasound, making it difficult to clearly know what type it is.

Because of these problems, we can take samples (cytopunctions) that are not necessary, or, on the contrary, delay the discovery of a disease. That is why the Haute Autorité de Santé (in 2021) insisted on the importance of having clear and uniform rules for performing thyroid ultrasound scans, in order to better decide when a complementary examination is really useful.

II.2.3.2 Interest of standardized stratification systems:

In order to make assessments less subjective (less based on personal opinion) and for review reports to be more consistent, risk ranking methods have been developed. These systems provide an organized way of assessing nodules using clear and established rules. Among the most well-known and recognized systems, we find:

- EU-TIRADS , created by the European Thyroid Association (in 2017),[9]
- ACR TI-RADS , developed by the American College of Radiology (in 2017). [13]

These classifications have several advantages:

- A common language and uniform exam reports (everyone speaks the same language).
- Fewer differences in interpretation between the different doctors.
- A better choice of nodules that really need sampling.
- Fewer invasive procedures (such as swabs) that are not necessary.

The ACR TI-RADS system assigns points to different features seen on ultrasound, while the EU-TIRADS classifies nodules into categories according to their overall appearance.

II.2.3.3 Towards the integration of artificial intelligence:

Even though classification systems like EU-TIRADS and ACR TI-RADS have helped to standardize the analysis of small thyroid masses, there is still a measure of personal judgment when looking at ultrasound scans. To reduce these differences of appreciation and make diagnostics more reliable, software for computer-based diagnostics (called CAD) has been developed. These programs work thanks to artificial intelligence techniques, such as 'machine learning' and 'deep learning'. They are able to identify details about the shape and texture of nodules directly on ultrasound images. Contrary to what the human eye can see, these tools can detect complex clues and hidden "signals" in images that we do not perceive directly. Several studies have shown that these artificial intelligences can be as effective for diagnosis, if not better, than less experienced radiologists. In addition, they allow to reduce disagreements between different specialists. These tools are particularly valuable in places where there are fewer cases to analyze or less experts in thyroid ultrasound.

In summary, artificial intelligence does not replace the doctor. It rather complements existing ranking methods, thus helping to make fairer, more objective and more coherent decisions. [29]

II.3 Diagnostic Assistance Systems (CAD):

Computer Assisted Diagnostic (CAD) software is a computer program designed to assist health professionals in the analysis of medical images. Concerning thyroid diseases, their main goal is to increase the accuracy of diagnosis, reduce the differences in interpretation between specialists and optimize therapeutic choices.

Thyroid nodules are very common and are frequently detected in a significant proportion of adults. Although most of these nodules are non-cancerous, a significant proportion require further diagnostic examination to ensure there is no malignancy. Ultrasound is the primary imaging method used for this assessment, as it allows non-invasive observation of physical characteristics of nodules, such as their ability to reflect ultrasound, their edges, the presence of calcium deposits and their blood supply. However, the interpretation of ultrasound images depends on the experience of the professional and may vary from one specialist to another, especially for nodules whose signs are unclear or ambiguous. To overcome these limitations, computer-based diagnostic tools, known as Computer-Aided Diagnosis (CAD) systems, have been developed to provide radiologists and physicians with objective and consistent support for decision making.

A CAD system dedicated to thyroid nodules aims to automate the analysis of characteristics observed on ultrasound, identify areas of potential concern and propose diagnostic categories (for example, benign or suspect, or according to a ranking system like the TI-RADS). It can thus contribute to reducing diagnostic uncertainty, reducing the number of false negative results and improving medical management for patients with thyroid diseases.

II.3.1 CAD Definition & General Architecture:

A CAD system is software designed for the automatic analysis of medical images. Its role is to identify relevant elements for the diagnosis and provide an interpretation that complements that of the doctor. Specifically for thyroid nodules, a CAD system based on ultrasound uses imaging data. It evaluates shape and texture characteristics in order to predict the risk of malignancy or to guide decisions regarding cytopunction.

thyroid CAD systems automate several stages of image processing. These include segmentation, feature extraction and classification. This automation makes it possible to standardize the analysis and reduce the number of errors due to subjectivity. [30]

General architecture of a CAD system for thyroid nodules:

A computer-aided diagnosis (CAD) system typically adopts an organized processing methodology, which transforms raw ultrasound data into an automated diagnostic assessment. See (Figure II.13)

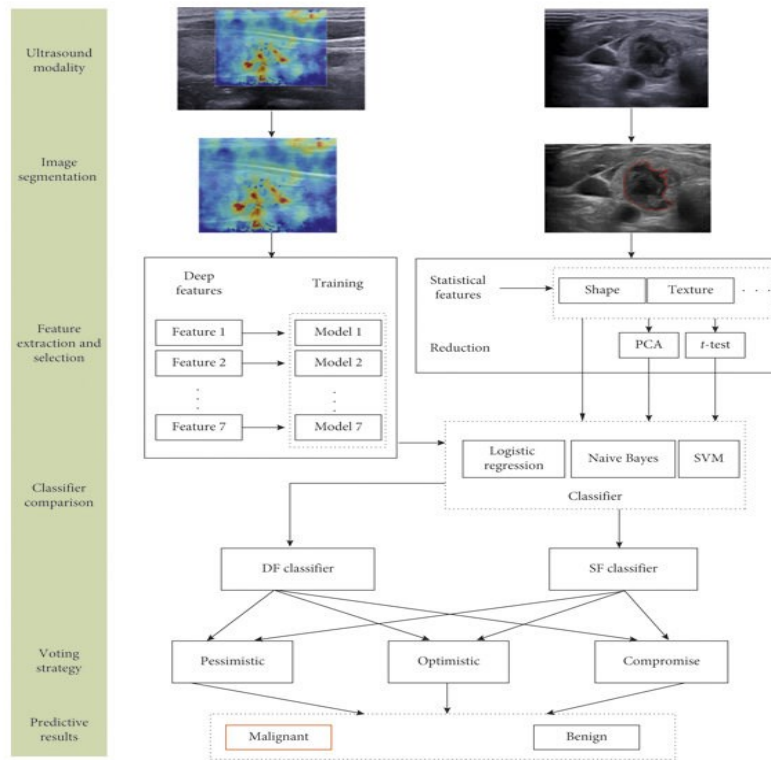


Figure II.13: Global architecture of a CAD system: segmentation, feature extraction, classification and final decision.[75]

1. Acquisition and pre-processing

Before any analysis, ultrasound images are obtained, whether in B mode or Doppler. Pretreatment is used to reduce the speckle noise inherent in ultrasound imaging. a characteristic inherent in ultrasound, which manifests itself as disturbances such as attenuation and artifacts. Commonly used approaches, such as intensity standardization, filtering (median, low-pass) and contrast enhancement, are applied to optimize image quality before any subsequent processing operation.

The objective of this step is to eliminate unwanted disparities in order to facilitate the segmentation process.

2. Automatic segmentation

Segmentation aims to circumscribe the boundaries of the thyroid nodule within the image, thus establishing the Region of Interest (ROI) for subsequent analyses. Neural architectures, such as U-Net or their specialized variants, are commonly employed for this deep learning task, in order to accurately distinguish nodular formations of the echographic background. See (Figure II.14)

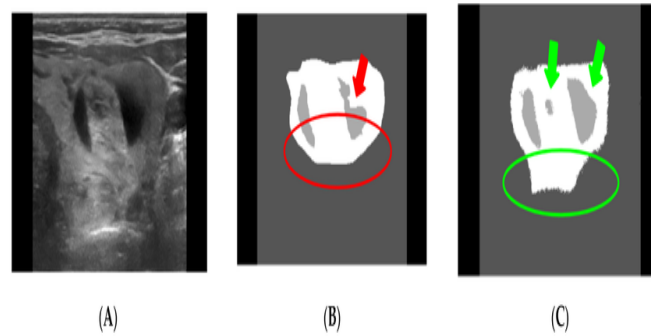


Figure II.14: Exemples de segmentation nodulaire (nodule detection) results obtained using the ResUNet model. (A) Original image, (B) average of Physicians 1, 2, 3, and (C) ResUNet segmentation.[23]

3. feature extraction

Once the nodule is delineated, indicators are collected to define its morphology, texture and other key properties:

- Shape criteria: surface, contour, and degree of roundness.
- Texture criteria: contrast, intensity and uniformity.
- Automatic deep features: numerical representations obtained via convolutional neural networks (CNN).
- Hybrid approaches: methods that combine statistical measures and data from deep learning to improve the relevance of diagnosis.

4. Classification and decision making

After their extraction, the identified attributes are then routed to a classification module. The latter's mission is to determine the nature of the nodule, characterizing it as benign or potentially malignant. Approaches employed include: Traditional machine learning algorithms, such as SVM, Random Forests and k-NNs. Deep convolutional neural networks (CNNs), with a stratified architecture.

Composite systems, incorporating prior weighting and specific characteristics.

An illustration of such a multi-layered architecture, which associates automatically learned attributes with morphological descriptors perfected by preliminary algorithmic weightings, in order to optimize the distinction between categories. See (Figure II.15)

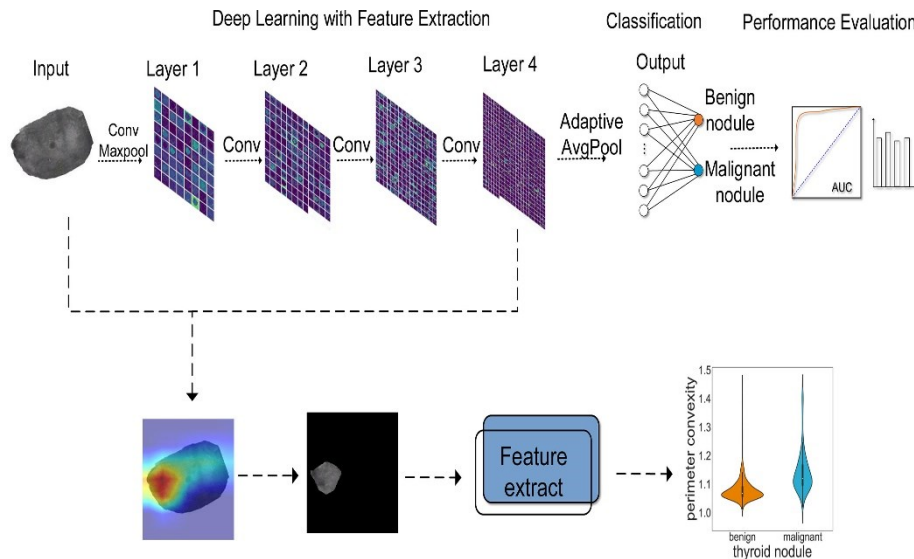


Figure II.15: Diagram of a hybrid CAD architecture with deep extraction + pre-weighting.

5. Diagnostic output and clinical recommendation

The conclusive phase involves formulating a diagnostic assessment based on a classification. This may include, for example, an estimate of the likelihood of malignancy, a TI-RADS classification, or a suggestion for monitoring or fine needle aspiration (FNA) biopsy.

This evaluation can take various forms:

- A quantified probability value (score).
- A specific classification, such as the category TI-RADS.
- A comprehensive clinical guideline (e.g., monitoring plan, biopsy procedure, or other relevant interventions).

II.3.2 Classic image processing methods:

After having presented the general architecture of a CAD system, it is advisable to examine the classic image processing methods historically used in each step of the pipeline before the rise of deep learning.

Image processing is an essential pillar of computing, whose goal is to explore, transform and give meaning to digital images. It is found everywhere today, playing an indispensable role in various fields such as medicine (for imaging and diagnostics), security (to identify faces), industry (to check the quality of products), not to mention robotics, the media and even augmented reality.

Its purpose is simple: either make the images more beautiful and clearer, or find relevant information. For this, he uses common actions such as filtering or resizing, but also more complex technical prowess, for example to spot contours, isolate specific elements or even identify shapes. [31]

II.3.2.1 Image enhancement:

When we talk about improving an image, it's a bit like giving it a facelift to make it easier to look at or understand. The goal is to better distinguish certain elements and facilitate their automatic analysis. Here is how we go about it, with some essential techniques:

- Histogram equalization: It is a method that involves better distributing the shades of gray in the image. Imagine that all the colors are a little 'stuck'; this method 'spreads' them out so that the general contrast is more striking. It's great for medical images like ultrasound, where it's often difficult to distinguish a small mass from the rest, because everything looks a bit the same.

- Spatial filtering:

Averaging filter: It attenuates small random disturbances that can spoil the image, making it smoother.

Median filter: This is a champion to remove the unwanted "blackheads" or "whites" (the impulse noise), and it has the advantage of not blurring the contours of objects, which remain well defined.

Edge improvement: Here, the idea is to highlight areas where colors change at once. It's great for making the boundaries of elements, like a nodule, much more distinct and easy to spot.

Pseudo-colours and multi-bands: Sometimes, certain values (pseudo-colours) are assigned "false" colours or different "layers" of information (multi-bands) are combined. This allows to better see certain details or help machines do their analysis work more efficiently.

II.3.2.2 Image segmentation:

Basically, segmentation is like cutting an image into several parts where things look similar inside. The idea is to be able to easily identify and isolate what really interests us. It's a really crucial step, for example, when studying masses in the thyroid.

1. Segmentation by contours:

Derivative methods: It is a way to find edges or sudden changes by measuring the evolution of data.

Analytical methods: Canny is a bit like a very meticulous detective to find the sharp edges of an image. It proceeds in several steps, a bit like a well-crafted investigation:

1. Erase the Blur: First, it softens the image a bit to mitigate small defects (noise) that could distract it. He doesn't want to be deceived by tiny color variations.
2. Spot the Changes: Then, it looks for places where the color or brightness changes abruptly. These are where the contours are most likely to be found. He notes the "strength" of this change.
3. Refine the Traits: Then, it purifies these raw tracings. If an outline is a bit thick or blurry, it makes it fine and clear by eliminating unnecessary duplicates or streaks. He only keeps the "heart" of the edge.
4. Connect the Pieces: To finish, he connects the edges that he has found. He uses a trick: if a small piece of contour is close to a large and solid one, he considers that they are part of the same set. This allows to obtain continuous lines and to ignore small isolated points that are not real edges.

Canny's goal is to obtain clear, precise contours without too many false positives, as if someone had drawn them in pencil.

Deriche's method: The Clever Engineer

Deriche, he is a bit of an astute engineer who has found a faster and more elegant way to do the first steps of contour detection (smoothing and tracking changes).

Instead of processing the image in one go to make it softer (as Canny would do with a "big" filter), Deriche uses a more fluid and economical technique. He "scans" the image several times in a row with a simpler calculation at each pass. It's a bit like passing through a light sponge several times rather than a big brush once.

Deformable Methods (Snakes/Active Contours): They model a path so that it respects the contours of the object.

2. Segmentation by regions:

Region segmentation groups neighboring pixels according to their homogeneity. It includes several methods:

► Region Growing:

This image segmentation approach is characterized by a progressive propagation starting from priming points (or initial seeds). It consists of iteratively aggregating adjacent pixels that share similar properties.

Homogeneity criteria used for this aggregation may include, but are not limited to, pixel intensity or textural characteristics.

► Division/Fusion (Split & Merge) – Quadtree:

Iterative decomposition of the image domain into elementary spatial entities, guided by an intrinsic uniformity criterion, such as the statistical dispersion of values or their radiometric mean.

Aggregation of contiguous areas with similar characteristics, aiming to mitigate the phenomenon of over-segmentation. The frequent use of a tree structure of type Quadtree is observed for the management of the spatial hierarchy of the resulting entities.

► Region classification (Clustering):

Each pixel of an image is defined by a vector of intrinsic properties.

The K-Means algorithm performs an exclusive assignment of each entity to the aggregate whose centroid is closest.

Conversely, the C-fuzzy averages (FCM) algorithm performs a non-exclusive assignment of entities to several aggregates, this assignment being determined by a degree of belonging coefficient.

3.Segmentation by threshold:

Segmentation by thresholding subdivides the image plane based on a criterion of light intensity:

The global threshold approach: it applies an invariant threshold to the entire imaged domain (for example, the algorithm of Otsu, 1979).

Adaptive thresholding: it uses a modulated threshold value, determined by the local pixel context, which is particularly relevant for images with heterogeneous illuminance distributions.

II.3.2.3 Classification and extraction of characteristics:

At the end of the segmentation, the delimited entities (pixels or regions) can be characterized and then classified according to their intrinsic properties:

Morphological features: Include dimensions (size), spatial configuration (shape) and compactness (for example, of a nodule).

Textural features: Include indicators such as variance, contrast, and co-occurrence measures.

Supervised classification: Among which we find the Support Vector Machines (SVM), the Random Forests (Random Forest) and the K closest neighbors (KNN).

Unsupervised classification: Such as the K-means algorithm and the Fuzzy C-means.

These techniques are essential precursors to artificial intelligence approaches that will be examined in more detail in the subsequent section.

II.3.3 Artificial intelligence approaches: machine learning and deep learning:

As there is an increase in ultrasound images and they are increasingly difficult to understand, the usual techniques are no longer enough. Artificial Intelligence (AI) offers tools that can learn on their own from many data. These tools are capable of performing tasks such as classifying (sorting) or delimiting areas, with an accuracy equal to or even superior to that of human specialists. [23]

II.3.3.1 Machine Learning (ML):

Machine learning is a method where a computer learns to make decisions by studying examples that have already been labeled or classified. In CAD diagnostic software for thyroid ultrasound, older machine learning systems use precise information (such as shape, texture or numbers) that experts have chosen by hand. This data comes from key areas of the image that have been identified and delimited beforehand.

Typical processing flow

1. Manual feature extraction:

Morphology: Aspects such as size (area), contour, or proportion (for example, if it is rather long or wide) are examined.

Texture: We study how the surface presents itself, such as differences in brightness (contrast), whether it is uniform or not, and its complexity (if the small elements are well-ordered or rather scattered).

Statistics: average, variance of intensities.

2. Selection and standardization of characteristics: Important elements are selected and organized. This includes reducing the amount of detail or rearranging it to better understand it.

3. Supervised classification: Computer programs are trained to connect this prepared information to already known labels or diagnostics.

Classical ML algorithms

Support Vector Machines (SVM): Creates the widest possible separation between different groups of data.

k Nearest Neighbors (k NN): Classifies items based on what the closest neighbors.

Random Forest: Combines several "decision trees" (small choice guides) to better learn and avoid focusing too much on the details of the starting data.

Naïve Bayes: A simple prediction system based on probabilities, assuming that each piece of information or trait is independent from the others.

Artificial intelligence (AI) is very useful in many medical fields. However, it has limits: its proper functioning depends heavily on the quality of the basic information given to it, and how well the elements are separated or cut at the very beginning.

II.3.3.2 Deep Learning (DL):

Deep Learning is a branch of machine learning that uses large neural networks, i.e. computer systems with many layers of information stacked. These systems learn directly from images as they are, eliminating the need to manually identify and select important details.

Main features of DL approaches

1. Automatic feature extraction:

These computer tools learn to identify specific visual patterns in ultrasound images. They then build a progressive understanding of these images, starting with simple elements like edges, and moving up to more complex structures like nodules (or masses).

2. Integrated segmentation:

Systems like U-Net allow for the complete identification and isolation of important parts before classification. This helps to locate small elements (such as nodules) more accurately, and the information collected to identify them is of better quality.

3. Deep classification:

CNNs (a type of specialized computer program) can directly analyze specific areas of an image and determine their category for diagnosis (e.g., whether it is "benign" or "malignant").

Often, evolutions of these systems that are more complex (more "deep"), that integrate particular design elements (such as ResNet blocks), or that reuse already learned knowledge (so-called learning transfer) give better results.

Positive Points (Benefits):

- He needs less to be told what to look for or what the important features are.
- He can find very complicated and subtle links in the information.
- It is often very efficient, especially when given a lot of clear and labeled data.

Weak Points (Limits):

- It often needs a lot, a lot of clear and labeled data to function well.
- It needs more computing power (high-performance computers).
- It is more difficult to understand why he makes certain decisions, unlike older methods.

III.3.3.3 Specific applications in thyroid ultrasound imaging:

Recent studies have successfully used deep learning methods (DL or advanced AI) to identify and delineate thyroid nodules.

For example, an article published in the journal BMC Cancer in 2025 describes a technique where already trained AI systems are adapted to classify thyroid nodules from ultrasound images. The results showed that this approach can improve diagnostic accuracy, even when there is little data available.

In addition, AI tools capable of cutting out or delineating regions on neck ultrasound scans have proven highly effective in helping to automatically detect whether papillary thyroid cancer has spread to the neck ganglia. This confirms the clinical interest of advanced AI for tasks that require both locating areas and classifying them. These examples clearly demonstrate that deep learning methods are capable of solving complex medical problems in automated thyroid nodule diagnosis.

II.3.4 state of the art :

The automated assessment of thyroid nodular formations by ultrasound imaging has undergone a significant transformation, evolving from a reliance on classical feature-based approaches to contemporary paradigms focused on deep learning. This section is dedicated to the review of the major contributions identified, their respective effectiveness, as well as the limitations associated with them.

II.3.4.1 Work based on classical methods:

The initial methodologies in computer-aided diagnosis were based on the manual derivation of specific, expertly designed attributes from ultrasonographic data. These properties included morphology, textural patterns, delineation and echogenicity of nodular formations.

In 2013, **Gopinath and Sivaram** introduced a methodological approach based on texture descriptors using the grey level co-occurrence matrix (GLCM), combined with a support vector machine (SVM) classifier, for the categorization of thyroid nodules from **cytological images**. Their methodology demonstrated notable performance, achieving an accuracy rate of approximately 85%. However, this approach is inherently dependent on the quality of the preliminary segmentation of the images [32].

At the same time, Acharya et al. (2016) exploited morphological and statistical attributes, combined with a Random Forest-type classifier, to discriminate between benign nodular lesions and malignant ones, obtaining a sensitivity rate of 87%[33].

Although these methodologies are effective, they remain circumscribed by their reliance on manual extraction of features and heuristic selection of descriptors.

II.3.4.2 Deep Learning-based works:

The advent of deep neural network architectures has led to significant performance improvements in many research areas.

Li et al. (2019) developed a convolutional neural network (CNN)-based model for the automated identification of thyroid nodules from ultrasound images. This model achieved 91.6% accuracy on an annotated clinical dataset [29].

Zhou et al. (2020) exploited a pre-trained ResNet50 architecture to optimize the benign/malignant binary classification, resulting in an AUC value of 0.94. This performance demonstrates the superiority of deep models over classical approaches [34].

In terms of segmentation, the U-Net architecture, introduced by Ronneberger et al., has quickly become a reference in medical imaging. It allows precise delineation of anatomical structures, even with small data sets [35].

More advanced variants, such as Attention U-Net and U-Net++, have further improved the accuracy of thyroid nodule segmentation, including in noisy ultrasound environments.

II.3.4.3 Comparison of existing approaches:

Traditional methodologies and approaches based on deep learning are fundamentally different in several ways:

Whereas conventional paradigms rely on the engineering of manually defined features, deep learning autonomously acquires data representations. Classical techniques offer a greater intelligibility of their process, but generally display lower performance. In addition, convolutional neural networks (CNNs) confer a higher generalization ability, although they require large labeled datasets for training.

In terms of operational performance, deep learning approaches usually outperform conventional methods, with accuracy improvements of 5-10%, and increased resilience to variations in ultrasound images.

II.3.4.4 Limites des travaux existants :

Despite the considerable progress made by machine learning and deep learning methodologies in characterizing thyroid nodules, the existing literature reveals several persistent gaps. Most of the proposed approaches are based on small and frequently proprietary datasets, which significantly limits their ability to be generalized to diverse patient cohorts.

Moreover, the robustness of the models is likely to vary considerably depending on the intrinsic quality of the ultrasound images and the acquisition devices used, introducing a high sensitivity to the data acquisition conditions. While some conventional approaches are still constrained by

the need for manual extraction of relevant features, deep learning architectures, although more efficient, have a deficiency in terms of clinical interpretability.

Finally, few studies propose an exhaustive integration of the entire treatment chain (including segmentation, classification and clinical decision support).

II.4 Segmentation and classification of nodules:

Image segmentation is a truly crucial step for computer systems that help diagnose thyroid nodules during an ultrasound. This consists of clearly delineating the exact area where the nodule is located (the "region of interest"). This precision is essential in order to then be able to extract relevant information from it or for artificial intelligence programs to analyze it on their own.

However, on thyroid ultrasound scans, this task of delineation is particularly complicated due to several factors:

- The "noise speckle", these small imperfections or the natural grain of the image related to the ultrasound technique.
- The low contrast between the nodule and the rest of normal thyroid tissue, making the contours difficult to distinguish.
- The edges of nodules that may be irregular or slightly sharp.
- And finally, the "acoustic artifacts", kinds of false images or disturbances caused by ultrasound itself.

Two main approaches are used:

II.4.1 Segmentation: classical techniques and deep learning (U Net):

II.4.1.1 Classic segmentation techniques:

Classical methods are based on concrete information, such as brightness, patterns or the regularity of lines that delimit shapes.

1. Segmentation by threshold: Threshold segmentation The most common way to segment an image is by using a threshold technique; this means that pixels can be classified according to their intensity level.

Global thresholding technique: a single threshold value is applied to all pixels in the image (e.g. Otsu's method).

Adaptive threshold: the threshold varies depending on the characteristics of the local area (i.e. neighboring areas in the image).

The use of thresholding in ultrasound thyroid images is limited because nodules often exhibit intensities similar to those of surrounding areas. [36]

2. Segmentation by contour: There are several methods to detect intensity discontinuities:

gradient detection (for example, Sobel and Canny), active contour models (snakes) and region segmentation methods that group pixels according to homogeneity criteria.

Region growth is a subset of segmentation methods.

Active contour modeling uses curves that evolve towards the contours of the object by minimizing internal and external energy functions.

The main limitation of active contour models is their high sensitivity to speckle noise. [37]

3 . The segmentation of regions: consists of grouping pixels based on their similarity or homogeneity.

- a) Region growth, for example, generates groups of pixels from starting points based on the degree of intensity or texture similarity between pixel values.
- b) Division/Fusion (Quadtree): By performing a recursive cut of the image, it is divided into smaller sections until each section reaches a level of homogeneity.
- c) Classification (K-Means, Fuzzy C-Means): Pixels can be classified into different categories in order to group similar pixels based on their common characteristics.

II.4.1.2 Deep learning segmentation:

Image segmentation has been profoundly transformed by deep learning, allowing models to automatically generate hierarchical representations from vast databases. The traditional method used rule-based and manually developed descriptors, while convolutional neural networks (CNNs) learn directly the discriminant features at the pixel level. This machine learning capability has significantly increased the robustness, adaptability and accuracy of segmentation

systems.. The U-Net architecture among the most efficient for biomedical segmentation tasks. [38]

1. Fully convolutional networks (FCN)

FNCs are among the first networks designed exclusively for semantic segmentation. Their architecture is based on fully convolutional layers; in other words, the fully connected layers of a traditional CNN are replaced by convolutional layers. This allows to capture and preserve the spatial dimension of the image. The architecture consists of two distinct parts: the encoder extracts deep features from the input image; and the decoder performs oversampling to generate an output image that is the same size as the original input image, thus allowing a pixel-by-pixel classification.

Advantages: Ability to process an image of various dimensions. Creates a segmentation at the pixel level with high density and precision. Works great with larger datasets. The FCN have contributed to the development of most of the new architectures used today.

2. U-Net Architecture:

an encoder/decoder model specifically designed for medical imaging, uses a unique method of direct connection between the corresponding encoding and decoding layers (which allows to preserve fine details during reconstruction). Initially designed for the segmentation of micrographs, U-Net stands out as a standard model in medical imaging for the segmentation of lesions/tumors and organ segmentation. The advantages of using U-Net in medical imaging are:

Fine details are perfectly preserved. High performance on high-resolution medical images. Works efficiently with small datasets. Data processing takes time; in addition, short-distance data is processed much faster than long-distance ones. U-Net should perform well for ultrasonic segmentation as it tolerates irregular contours and speckle noise. See (Figure II.16)

- **Contracting path (encoder)**

Successions of convolution + Max Pooling to gradually extract higher-level features.

- **Expansive path (decoder)**

Upsampling + convolutions to reconstruct the segmentation at original resolution.

- **Skip connections**

Transfer feature maps from the contracting side to the expansive side to preserve fine spatial details.

- **Mathematical functioning**

Let x the input ultrasound image, and y the expected segmentation. U-Net learns a function f_θ such that equation (II.2):

$$\hat{y} = f_\theta(x) \quad (II.2)$$

where:

(y) is the probability map of segmentation,

θ represents the parameters (weight) of the network.

The parameters are adjusted by minimizing a loss function (Loss), often based on the Dice coefficient equation (II.3):

$$DiceLoss = 1 - \frac{2 \times |Y \cap \hat{Y}|}{|Y| + |\hat{Y}|} \quad (II.3)$$

This choice of loss is particularly suitable for situations where the classes are unbalanced (small nodular surface vs large background).

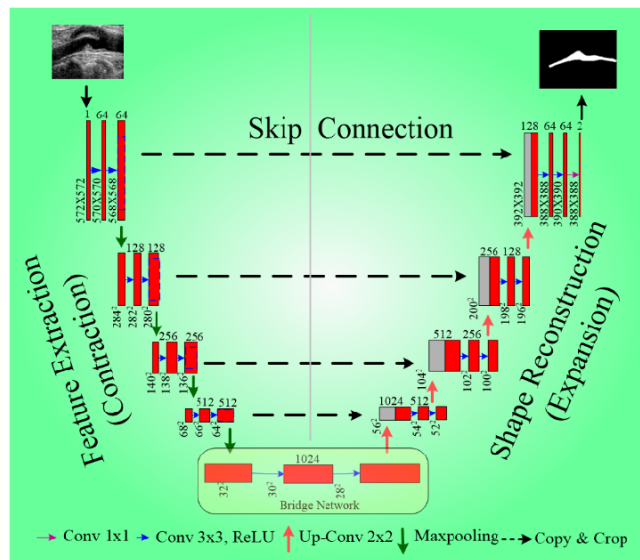


Figure II.16: Basic UNet architecture showing encoder, decoder, bridge network and hop connection. Encoding layers (contraction phase): feature extraction; decoding layers (expansion phase): shape reconstruction.[35]

3. Mask R-CNN:

The Mask R-CNN is an evolution of the Faster R-CNN. It adds a special feature to draw the exact shapes (or masks) of each element it identifies. But unlike the "semantic" segmentation which groups everything by categories, Mask R-CNN is able to differentiate each individual object (this is the "instance segmentation"), even if they are all of the same type.

Its strengths:

- High precision to isolate and understand each object.
- He manages very well even when the objects overlap or partially hide.
- It is very adaptable and efficient, even with very varied and complex image databases.

4. DeepLab:

The DeepLab series is distinguished by the use of special filters, called "dilated convolutions", which allow it to perceive a wider area in the image without making the model heavier or more complex. This greatly helps him understand the overall context, no matter the size of the elements he observes.

With DeepLabv3+, the developers have integrated an even more efficient "encoder-decoder" system, which makes the borders of segmented objects particularly fine and precise.

Moreover, some variants of DeepLab even go so far as to use post-processing methods, such as "conditional random fields" (CRF), to fine-tune the separations between already identified objects.

Benefits:

- He knows very well how to adapt to all sizes of objects, from the smallest to the largest.
- He cuts the images with great accuracy, and the edges are impeccably sharp.
- He is comfortable even in visually complex and charged scenes.

5. Segment Anything Model (SAM):

The Segment Anything Model, or SAM, created by Meta AI, represents a big step forward for image clipping tools capable of adapting to all kinds of situations. Unlike the old systems that required dedicated learning for each new job, SAM is a basic model capable of

recognizing and cutting out any object just by giving it prompts ("prompts"), and this, without needing to be retrained. One can use it to cut elements in an interactive way, by guiding it with a simple point, by drawing a frame around the object, or even by providing it with a mask beginning.

Its strengths:

- Very great adaptability: It works wonderfully in many different fields.
- Less data preparation: It greatly reduces the need to provide him with marked examples.
- Easily usable: It can be integrated into various applications with very few modifications.

II.4.2 Extraction of characteristics (morphological, textural, deep features):

Identifying and retrieving key information is a really crucial step when using the computer to help diagnose (a CAD system). Once a suspicious area, such as a nodule, has been delineated, the next task is to accurately measure all relevant visual information of that 'region of interest' (or ROI). The goal is to make them understandable and usable by specialized software, called classifiers.

II.4.2.1 morphological characteristics:

Main characteristics: To know what a nodule looks like, we examine its shape, size and arrangement. These details stem from the spatial organization of the region that has been highlighted. In addition, it is based on indicators that doctors commonly use for their diagnoses (such as the ACR TI-RADS).

- **Area (A)**

Area occupied by the nodule (number of pixels in the ROI). It is related to the size of the nodule.

- **Perimeter (P)**

Contour length. Useful for measuring shape irregularity.

- **Compactness (C)**

Measure of roundness equation (II.4):

$$C = \frac{P^2}{4\pi A} \quad (\text{II.4})$$

High compactness may indicate a more irregular shape, often associated with a suspicion of malignancy.

- **Height/width ratio**

Defined as equation (II.5):

$$R = \frac{H}{W} \quad (II.5)$$

A 'taller-than-wide' nodule ($R > 1$) is clinically considered suspicious.

- **Eccentricity**

Ratio between the main axes of an ellipse adjusted to the ROI. [39]

II.4.2.2 Textural characteristics:

The aspect of texture informs us about the organization of hues and patterns inside a mass. This is very useful during ultrasound scans, as texture can be an important clue to distinguish healthy tissue from suspect tissue.

1 Cooccurrence matrix of gray levels (GLCM)

Basically, the GLCM (Gray Level Co-occurrence Matrix) is a table or a kind of register that helps us to count how many times certain pairs of shades of grey appear together in an image. She examines these duets taking into account the distance that separates them and the direction in which they are observed.

Excerpts from the GLCM:

- **Contrast** equation (II.6):

$$Contrast = \sum_{i,j} (i - j)^2 p(i, j) \quad (II.6)$$

Measures the intensity of changes in the image.

- **Energy (Angular Second Moment)** equation (II.7):

$$Energy = \sum_{i,j} p(i, j)^2 \quad (II.7)$$

Indicates textural homogeneity.

- **Homogeneity** equation (II.8) :

$$Homogeneity = \sum_{i,j} \frac{p(i,j)}{1+|i-j|} \quad (II.8)$$

- **Correlation**

Quantifies the linear dependence between neighboring pixels. [40]

2 Histograms and statistical characteristics:

Grayscale histogram: global distribution of intensities in the ROI.

Statistical moments: mean, variance, skewness, kurtosis.

These measures allow to globally evaluate the intensity and internal variation of the nodule.

II.4.2.3 Frequency characteristics:

In fact, it is as if we were "listening" to the image rather than simply looking at it. We break it down to hear its different 'frequencies' or 'vibrations'. This helps us to flush out what is repeating or the very small details.

- **Discrete wavelet transform (DWT):** decomposes the image into sub-bands (LL, LH, HL, HH).
- **Fourier transform:** global analysis of spatial frequencies.

And these finds are super useful for spotting tiny structures within small balls, like tiny calcifications. [41]

II.4.2.4 Deep Features:

Thanks to Deep Learning, an advanced approach of artificial intelligence, convolutional neural networks (or CNNs) are able to learn by themselves to recognize and distinguish elements in images, without human intervention.

The idea behind all this is as follows:

In a CNN, it's a bit like if each analysis step (called 'layer') developed increasingly elaborate detection tools.

At the beginning, the first steps identify very simple details, such as the contours or textures of an object. Then, the further you go in the network, the more the following steps learn to understand much more complex and significant patterns, those that are really useful for identifying or classifying what we observe.

The result of this analysis, a sort of 'digest' of information from the last stages of the network, can then be considered as a 'deep feature'. It is a very rich and unique description of the nodule in question. [39]

II.4.2.5 Hybrid approaches: merging of features:

Some methods combine classic visual characteristics (shape, texture) with more sophisticated information extracted by an artificial intelligence that has already 'learned' elsewhere. The objective is to provide a better description of the elements and greatly improve the capacity of sorting or recognition tools. See (Figure II.17)

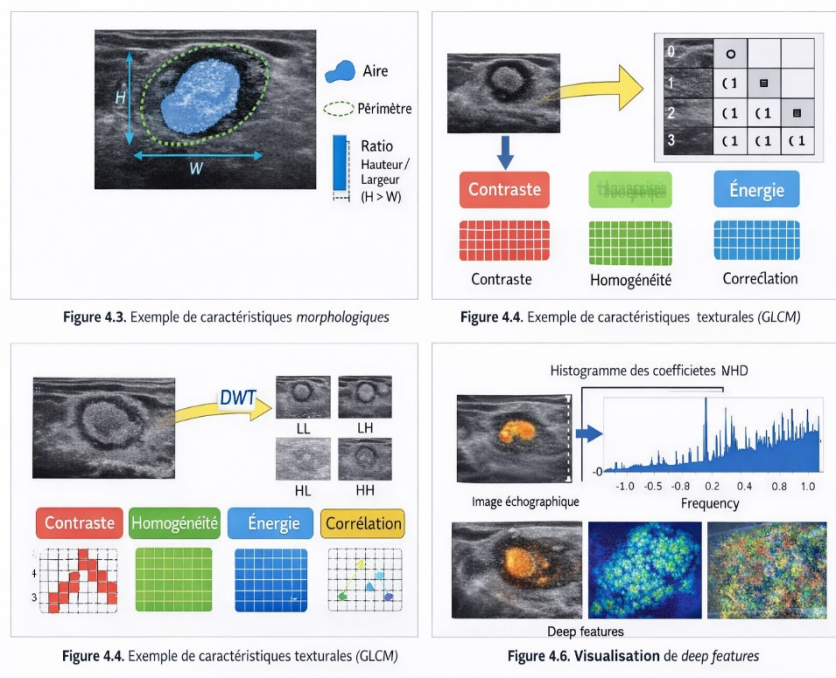


Figure II.17: Extraction of thyroid nodule features.

II.4.3 Classification: classical classifiers (SVM, Random Forest) and neural networks (CNN):

After segmentation and feature extraction, the classification step aims to automatically label the thyroid nodule (benign or malignant). This phase constitutes the core of the diagnostic

assistance system (CAD). The approaches used are divided into two main categories: classical machine learning classifiers and deep neural networks (Deep Learning).

II.4.3.1 Classical classifiers:

Les méthodes classiques reposent sur des caractéristiques préalablement extraites (morphologiques, texturales, fréquentielles). Elles nécessitent donc une étape d'ingénierie des caractéristiques (*feature engineering*).

A. Support Vector Machine (SVM)

Support Vector Machine (SVM) is a supervised classifier that seeks to determine the optimal hyperplane separating two classes with maximum margin.

Mathematical principle

Be a set of learning equation (II.9):

$$(x_i, y_i), y_i \in \{-1, +1\} \quad (II.9)$$

The SVM solves the optimization problem equation (II.10):

$$\min \frac{1}{2} \|w\|^2 \quad (II.10)$$

under duress equation (II.11):

$$y_i(w \cdot x_i + b) \geq 1 \quad (II.11)$$

When data is not linearly separable, a kernel is used:

Linear

Polynomial

Radial Basis Function (RBF)

Benefits:

Good performance on small datasets

Robust against large-scale data

Good generalization

Limits:

Sensitive to the choice of kernel and parameters

Performances dependent on the quality of the extracted characteristics.[42]

B. Random Forest:

The Random Forest (RF) is a set of decision trees built on random subsets of data and variables.

Principle

Each tree is trained on a bootstrap sample.

The final decision is obtained by majority vote.

$$\hat{y} = \text{mode}(T_1(x), T_2(x), \dots, T_n(x)) \quad (\text{II.12})$$

where T_i represents a decision tree.

Benefits:

Reduction of overfitting

Good management of noisy data

Importance of measurable variables

Limits:

Less efficient than CNNs on raw images

More complex global interpretation.[43]

II.4.3.2 convolutional neural networks (CNN):

Unlike classical methods, CNNs do not require manual feature extraction. They automatically learn the relevant representations directly from the ultrasound images.

A. General architecture

A CNN usually includes:

- Convolution layers: Extraction of local features using filters.
- Activation function (ReLU): Introduces non-linearity.
- Pooling (MaxPooling): Dimension reduction and spatial invariance.
- Fully connected layers: Final decision (classification).

B. Mathematical formulation

The output of a convolution layer is given by equation (II.13):

$$Y = f(W * X + b) \quad (\text{II.13})$$

where:

X: input image

W: convolutional filter

b: bias

f: activation function

Binary classification usually uses a Softmax or Sigmoid function (II.14):

$$P(y = 1 | x) = \frac{1}{1+e^{-z}} \quad (\text{II.14})$$

Benefits of CNNs:

Machine learning of features Excellent performance on medical images

Less dependent on manual pre-processing

Limits:

Need large annotated databases

High computational cost

Less interpretable than classical models.[44]

A comparative analysis between traditional machine learning and deep learning approaches highlights key differences in their applicability to thyroid nodule analysis. While SVM classifiers are effective for small-scale datasets with carefully selected features [37], they struggle to generalize when dealing with heterogeneous ultrasound images.

On the other hand, CNN-based architectures such as U-Net [42] overcome this limitation by learning high-level and spatial features directly from the data. This capability makes them particularly suitable for segmenting nodules with irregular shapes and low contrast, which are common challenges in thyroid ultrasound imaging.

Nevertheless, the performance of deep learning models is strongly dependent on the availability of large annotated datasets and computational power. Therefore, the choice between these approaches should consider the trade-off between data availability, model complexity, and required accuracy.

II.4.4 Evaluation metrics and conclusion of chapter 2:

Performance evaluation is an essential step in the validation of a computer-aided diagnosis (CAD) support system. In the context of thyroid nodule segmentation and classification, different quantitative metrics are used to measure the accuracy, robustness and generalizability of the developed models.

The metrics vary depending on the task:

- Spatial overlap metric segmentation
- Metric classification from the confusion matrix

II.4.4.1 Evaluation metrics in classification:

The performance of a binary classifier (benign / malignant) is calculated from the confusion matrix:

	Malignant prediction	Benign prediction
Real Malignant	TP (True Positive)	FN (False Negative)
Benin real	FP (False Positive)	TN (True Negative)

1. Accuracy equation (II.15):

$$\text{Accuracy} = \frac{TP+TN}{FP+FN+TP+TN} \quad (\text{II.15})$$

Measures the overall proportion of correct predictions.

2. Sensitivity (Recall or True Positive Rate) equation (II.16):

$$\text{Sensitivity} = \frac{TP}{FN+TP} \quad (\text{II.16})$$

Measures the model's ability to correctly detect malignant nodules.

In medicine, this metric is particularly important to avoid false negatives.

3. Specificity (True Negative Rate) equation (II.17):

$$\text{Specificity} = \frac{TN}{FP+TN} \quad (\text{II.17})$$

Assesses the ability to correctly identify benign nodules.

4. Precision equation (II.18):

$$\text{Precision} = \text{TP} / (\text{TP} + \text{FP}) \quad (\text{II.18})$$

Indicates the proportion of predicted nodules that are actually malignant.

5. F1-Score equation (II.19):

$$\text{F1} = 2 \times (\text{Precision} \times \text{Sensitivity} / (\text{Precision} + \text{Sensitivity})) \quad (\text{II.19})$$

Harmonizes precision and sensitivity, useful when classes are unbalanced.

6. ROC and AUC curve:

The ROC curve (Receiver Operating Characteristic) represents the sensitivity as a function of the rate of false positives equation (II.20):

$$\text{FPR} = \text{FP} / \text{TN} + \text{FP} \quad (\text{II.20})$$

The AUC (Area Under Curve) measures the overall ability of the model to discriminate between the two classes.

AUC = 0.5 random classification

AUC 1 excellent performance

II.4.4.2 Evaluation metrics in segmentation:

In the case of nodule segmentation, performance is evaluated by comparing predicted segmentation to reference segmentation (expert annotation).

1. Dice coefficient equation (II.21):

$$\text{Dice} = \frac{2|A \cap B|}{|A| + |B|} \quad (\text{II.21})$$

where:

A = predicted segmentation

B = reference segmentation

Value between 0 and 1.

The closer Dice is to 1, the better the segmentation.

2. Intersection over the Union (IoU) equation (II.22):

$$IoU = \frac{|A \cap B|}{|A \cup B|} \quad (II.22)$$

Measure the spatial overlap between the two regions.

Relationship between Dice and IoU:

$$Dice = \frac{2 \times IoU}{1 + IoU} \quad (II.23)$$

3. Hausdorff Distance:

Measure the maximum distance between predicted and actual contours equation (II.24):

$$H(A, B) = \max \left\{ \sup_{a \in A} \inf_{b \in B} d(a, b), \sup_{b \in B} \inf_{a \in A} d(a, b) \right\} \quad (II.24)$$

It allows to evaluate the accuracy of segmented boundaries.

II.4.4.3 experimental validation:

In order to ensure the robustness of the results, several validation strategies can be used:

- Hold-out validation
- Cross validation (k-fold cross validation)
- Stratified cross-validation

Cross-validation is particularly recommended when the database is limited, a common situation in medical imaging.

II.4.4.4 Conclusion of Chapter II:

In this chapter, we presented the basics of ultrasound imaging as well as the different stages of the diagnostic aid system applied to thyroid nodules. We first addressed segmentation methods, ranging from classical approaches (thresholding, regions, clustering) to deep learning architectures such as U-Net, recognized for their robustness and accuracy.

Then, we studied the extraction of morphological, textural and deep characteristics, essential elements for the quantitative representation of nodules. Finally, we analyzed classification methods, including classical models (SVM, Random Forest) and convolutional networks (CNN), as well as evaluation metrics allowing to objectively measure their performances.

The set of these techniques constitutes the methodological foundation of the system proposed in this thesis. The following chapter will be devoted to the adopted methodology, to the description of the database used as well as to the experimental implementation of the developed model.

CHAPTER III : DESIGN, DEVELOPMENT AND EVALUATION OF THE PROPOSED SYSTEM

III.1 Introduction:

Ultrasound imaging is a reference modality for the examination of thyroid nodules due to its non-invasive nature, rapidity, and wide availability [8]. However, the interpretation of ultrasound images remains challenging, particularly because of speckle noise, low intrinsic contrast, and the high morphological variability of nodules [45]. Moreover, diagnosis strongly depends on the radiologist's expertise, which may lead to inter- and intra-observer variability in clinical interpretation [46].

With the rise of artificial intelligence, deep learning methods have demonstrated remarkable performance in automated medical image analysis [44]. Convolutional neural network architectures, such as U-Net for segmentation and ResNet or EfficientNet for classification, are now widely used in thyroid nodule diagnostic support systems [35][47].

This chapter presents the development of the proposed system dedicated to the automatic segmentation and classification of thyroid nodules from ultrasound images. The methods employed are based on the TN3K [52] and TN5000 [48] datasets, as well as various preprocessing, data augmentation, and deep learning techniques, aiming to optimize diagnostic performance.

The obtained results are then presented and compared with the state of the art, along with a detailed analysis and interpretation of the system's performance.

Finally, this work is integrated into an interactive application with a graphical user interface developed using Streamlit, enabling practical use of the model in a simulated clinical environment and paving the way toward more advanced medical decision-support systems.

III.2 Hardware and data:

This section presents the hardware configuration, the datasets of ultrasound images used in this study, as well as the preprocessing procedures applied prior to the training of the model. Adequate data preparation is essential to ensure stability, reproducibility and suitability of the learning process for segmentation and classification purposes.

III.2.1 Description of the ultrasound image database:

In this study, two collections of thyroid ultrasound images, namely TN3K and TN5000, were used. Their goal was to design and validate the proposed approach for segmentation and classification of thyroid nodules.

The TN3K database is specifically designed for nodule segmentation, while TN5000 represents a more up-to-date and broader dataset, particularly relevant for identification and categorization tasks.

The joint exploitation of these two datasets allows to achieve two complementary objectives: precisely locate the nodule within the image, and then determine its diagnostic nature.

III.2.1.1 The TN3K database:

The TN3K (Thyroid Nodule Region Segmentation Dataset) is a public database of thyroid ultrasound images, designed for automated nodule segmentation. It includes 3,493 images from 2,421 patients, collected between January 2016 and August 2020 at Southern Medical University's Zhujiang Hospital. These images were manually annotated at the pixel level by specialists.

The selection of images was rigorous: they were chosen from a much larger set of examinations, retaining only those that contained at least one usable nodule. Lymph node images, overly colored images and duplicates were excluded to ensure clinical consistency and quality annotations.

Thus, TN3K is an essential resource for the evaluation of segmentation methods in thyroid ultrasound. It faithfully represents the variability of acquisitions, the diversity of equipment used and the challenges inherent to this imaging technique, such as low contrast and granular noise.[52]

III.2.1.2 The TN5000 database:

TN5000 (An Ultrasound Image Dataset for Thyroid Nodule Detection and Classification) is a public collection of B-mode ultrasound images dedicated to the identification and categorization of thyroid nodules. It contains 5,000 ultrasound images, marked with precise contours of the nodules and validated by histopathological confirmations obtained by fine needle aspiration (FNA), thus ensuring high clinical accuracy of annotations. According to its creators, TN5000 represents both the largest dataset of thyroid ultrasound images freely

available with annotations made by specialists, and the first dataset developed to jointly address the issues of identification and categorization of thyroid nodules. The images were standardized in a format optimized for computer vision algorithms, with a structured architecture intended to simplify comparative assessments between different deep learning models. The dataset also provides a predefined segmentation for training, validation and testing, thus facilitating a stricter and reproducible evaluation of algorithmic methods. TN5000 is therefore particularly relevant for the creation of medical diagnostic support systems, thanks to its combination of a considerable volume and markup by specialists, histopathological confirmation and a significant ability to drive reliable classification models.[53]

III.2.2 Nodule annotation and TI-RADS classification:

The evaluation and characterization of thyroid nodular formations, carried out by means of ultrasound imaging, require a fundamental phase of structured labeling of abnormalities. This process is then followed by their categorization according to standardized risk stratification systems, the most prominent of which is the TI-RADS (Thyroid Imaging Reporting and Data System). Within deep learning methodologies, the fidelity of this annotation turns out to be a crucial parameter to ensure the robustness of segmentation and classification models.[49]

III.2.2.1 Annotation of thyroid nodules in TN3K and TN5000:

The TN3K and TN5000 databases are two important collections that serve as a reference to help computers automatically analyze thyroid nodules visible on ultrasound scans. They consist of ultrasound images that have been labeled (annotated) according to well-defined medical rules.

The TN3K dataset is mainly designed to precisely delineate nodules. Its annotations are very detailed, either by marking each pixel of the nodule, or by drawing bounding boxes around it. The images come from clinical ultrasound and are provided with "masks" (drawings) that show the exact location of the problematic area. Experienced radiologists create these annotations, which are then formatted in a standard way (such as COCO or PASCAL VOC) for use by programs. These formats include the coordinates of boxes (xmin, ymin, xmax, ymax) or binary masks.[52]

The TN5000 dataset focuses more on the classification of thyroid nodules. It brings together ultrasound images with their diagnosis, taken from the medical records of patients or tissue analyses (histopathological). Sometimes, the annotation is "mixed": it gives both the position of the nodule and its label of malignancy (benign or malignant). To facilitate analysis by the

programs, delimitation boxes are extracted from specific annotation files (XML), which allows each nodule to be automatically framed for classification models.[13]

In both types of databases, the annotation process generally takes place in three stages:

1. A radiologist locates the nodule on the ultrasound image.
2. It draws the contours, either with a box or by marking each pixel that makes up it.
3. It combines clinical information, often based on standard ultrasound criteria.

This methodical organization is essential so that the artificial intelligence systems that analyze thyroid images can be reliable and give reproducible results.

III.2.2.2 TI-RADS classification system:

The TI-RADS system is a common method used for doctors to assess thyroid nodules (small balls) in a more uniform way and to better estimate if one of them could be dangerous.

The most common version is the ACR TI-RADS. It is based on the analysis of several characteristics of the nodule visible on ultrasound. Each nodule is examined according to five main points:

- Its composition (if it is filled with liquid, a mixture of both, or completely solid)
- His appearance on ultrasound (if he is lighter, darker, or the same as the rest of the thyroid)
- Its shape (whether it is higher than wide or not)
- Its edges (if they are smooth, wavy or irregular)
- The presence of very shiny areas inside (such as small or large calcifications)

Each point receives a certain value. By adding these values, the nodule can be classified into a risk category, ranging from TR1 (probably safe) to TR5 (highly suspected of cancer). See (Table I.3: The TI-RADS ACR System).

This system is widely clinically validated and is now an international reference for the decision of fine needle aspiration (FNA).

III.2.2.3 Correspondence between annotation and TI-RADS in datasets:

In the TN3K and TN5000 datasets (or image collections), the TI-RADS score is not always provided directly as a simple category. It is often deduced or estimated from:

- reports written by the radiologists,
- descriptions of the nodule structure, or simple clinical labels (such as "harmless" or "dangerous").

In deep-learning artificial intelligence (AI) systems, this information is used to transform the problem into:

- location of the nodule (for TN3K): find the exact location of the nodule, or classification (for TN5000): predict whether the nodule is at risk or if it is potentially serious.

Thus, marking the nodules is the essential basis for linking what the computer "sees" to the clinical rules of TI-RADS. This helps to integrate AI tools into how doctors make decisions in thyroid radiology.

III.2.3 Pretreatments and data augmentation:

III.2.3.1 TN3K:

For the TN3K dataset, dedicated to the segmentation of thyroid nodules, several pre-processing steps were applied in order to improve the quality of ultrasound images and optimize the training of the U-Net++ model based on an EfficientNet-B3 backbone.

First, the grayscale images were normalized and converted to a standardized format. A local contrast enhancement was then performed using the CLAHE method (Contrast Limited Adaptive Histogram Equalization), allowing better highlighting of anatomical structures and nodule contours while limiting noise amplification, particularly present in ultrasound images [83].

According to the information provided by the TRFE-Net project and the documentation of the TN3K dataset on Kaggle, the images have variable dimensions ranging from 216×217pixels to 1463×771pixels, with a median size of about 412×337pixels. In order to ensure data homogeneity and reduce computational cost, all images as well as segmentation masks were resized to a fixed resolution of 320×320pixels [35][52]. An adapted interpolation was used to preserve the integrity of the binary annotations. The images also underwent statistical

normalization, followed by channel replication to make them compatible with the pre-trained EfficientNet-B3 backbone on RGB images.

Finally, several data augmentation techniques, notably rotations, affine transformations, horizontal reversals as well as brightness and contrast variations, were applied simultaneously to the associated images and masks in order to improve the generalization capacity of the model and to limit overfitting.

III.2.3.2 TN5000:

As part of the TN5000 dataset, several pre-processing steps were applied to improve the quality of the ultrasound images and standardize the data before training the models. An exploratory analysis highlighted the presence of 126 duplicate images out of a total of 5000 images ($\approx 2.5\%$), as well as a high variability in resolutions and intensities, related to acquisition conditions and differences in ultrasound equipment.

According to the documentation of the TN5000 dataset available on Kaggle, the images have heterogeneous resolutions, sometimes exceeding 700×700 pixels for certain acquisitions, while other images are smaller in size, which introduces significant variability in the input data [53]. In order to standardize inputs and ensure compatibility with pre-trained convolutional neural network architectures (such as ResNet, DenseNet and EfficientNet), all images have been resized to a fixed size of 224×224 pixels. This resizing allows to significantly reduce the computational cost (memory and training time) while respecting the input standards of deep models commonly used in computer vision.

Furthermore, the images have been standardized in order to reduce variations in intensity and stabilize the learning process. The XML annotations have been converted to a structured format that can be used in supervised learning, followed by a consistency check between images and labels to eliminate missing or invalid data. A data separation based on patient groups was also applied to prevent any data leakage between the training and validation sets. The dataset was then divided into 80% for training and 20% for validation.

In order to improve the robustness of the model and reduce overfitting, several data augmentation techniques were used, including rotations, geometric transformations, horizontal reversals as well as brightness and contrast variations. These transformations make it possible

to simulate different ultrasound acquisition conditions and artificially increase the diversity of data.

Finally, in order to address the class imbalance observed in the TN5000 dataset (71.5% of malignant nodules versus 28.5% of benign nodules), a class-sensitive learning strategy was adopted. This is based on an inverse weighting of classes in the loss function, supplemented by resampling techniques such as over-sampling of the minority class. In addition, loss functions adapted to unbalanced data, such as Focal Loss, were used to better take into account difficult samples and reduce the impact of the majority class. This combination of strategies improves the model's sensitivity to benign nodules while maintaining good overall performance.

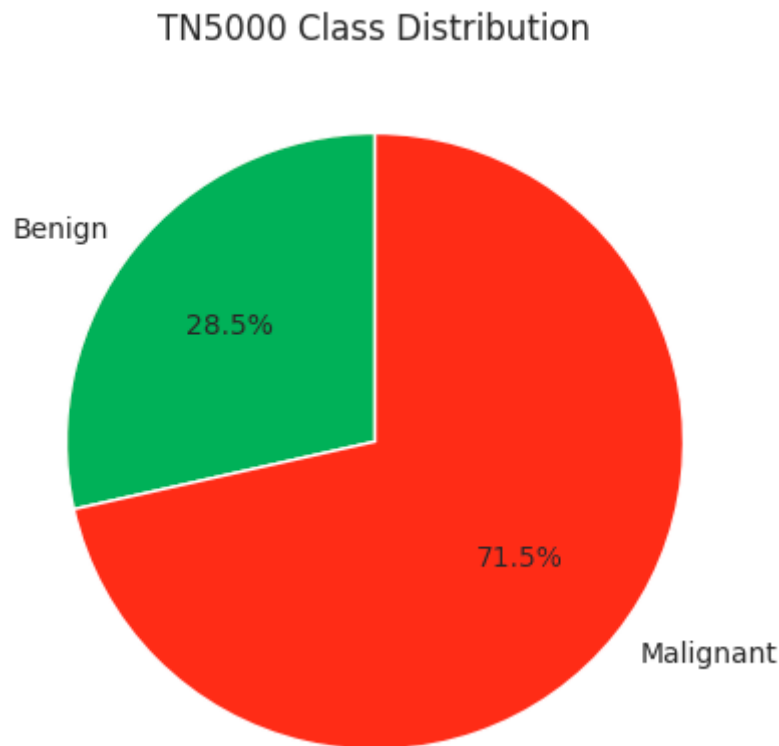


Figure III.1: Distribution of benign and malignant classes in the TN5000 dataset used after the pretreatment step.

III.2.3.3 Development environment and implementation tools :

The design and training of the system in question were carried out within the Google Colaboratory (Google Colab) environment, using the Python 3 programming language. This platform provides an interactive framework based on the Jupyter notebooks. which allows the

execution of machine learning models on powerful and remotely accessible hardware computing capabilities.[50]

In this study, the experimental calculations were optimized using an NVIDIA Tesla T4 GPU. The latter is specifically designed for performance-intensive applications in artificial intelligence, deep learning and data analysis. Based on the NVIDIA Turing architecture, this graphics processor offers a substantial acceleration of the training and prediction phases of deep neural networks. This also had the effect of significantly reducing the processing time of the segmentation and classification models used for this research.[51]

To ensure the reproducibility of the experiments, they were conducted in the Google Colab environment, which provides access to GPU computing resources. The programming language used was Python (version 3.10).

he main libraries used and their versions are as follows:

- PyTorch (version 2.x) for implementing deep learning models
- Torchvision for pre-trained architectures and image transformations
- scikit-learn for evaluation metrics and classic machine learning models
- pandas for data management and analysis
- numpy for numerical operations
- matplotlib for visualizing the results
- segmentation-models-pytorch for designing segmentation architectures (U-Net++)

Regarding hardware resources, the training was performed on NVIDIA Tesla T4 GPUs provided by Google Colab. As an example, training the U-Net++ segmentation model on the TN3K basis took approximately 3 hours and 40 minutes, while the classification models had varying training times depending on their complexity, ranging from a few minutes to several hours for the most complex architectures.

III.3 Methodology:

The presented methodology is based on an approach that combines delineation and classification of thyroid nodules from ultrasound images.

Initially, a U-Net++ architecture based on EfficientNet-B3 was used to isolate nodular areas.

Subsequently, deep learning and machine learning algorithms were implemented for attribute extraction and nodule classification.

Finally, various numerical indicators were used to assess the effectiveness of the proposed system.

III.3.1 Segmentation of nodules: chosen architecture and training:

In this study, the segmentation of thyroid nodules is approached as a binary classification task at each pixel level, aiming to identify and automatically separate the nodular area on ultrasound images. This operation is particularly difficult in medical imaging, notably because of the reduced contrast, speckle noise and the great heterogeneity of nodular morphologies. In order to overcome these challenges, we used a U-Net++ architecture combined with an EfficientNet-B3 encoder, pre-trained on ImageNet. This model is widely recognized for its effectiveness in the field of medical image segmentation.

- **Segmentation model architecture:**

The model is based on an improved encoder–decoder structure: This combination improves the network’s ability to capture both global structures and local details essential for nodule segmentation.

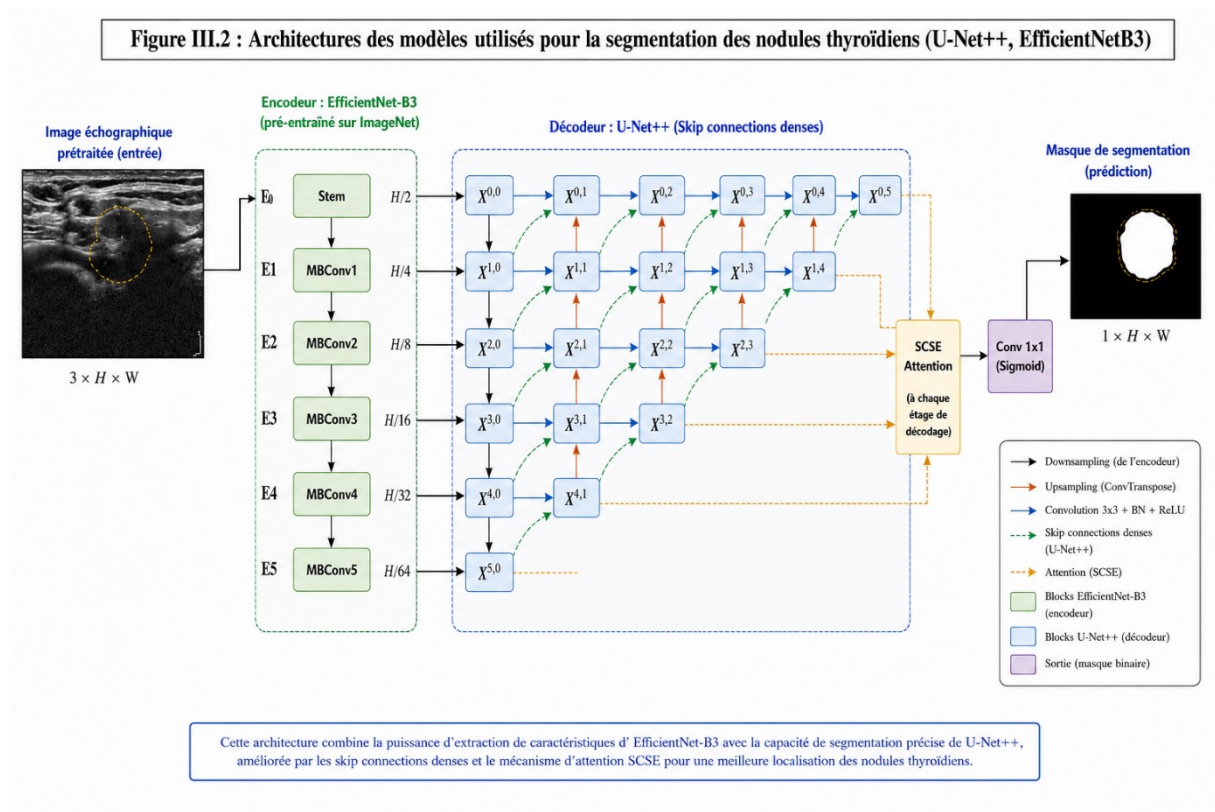


Figure III.2: Architectures of the models used for the segmentation of thyroid nodules (UNET++, EfficientNetB3) thyroid (personal development, 2026).

Component	Role
EfficientNet-B3 (encoder)	Multi-level deep feature extraction.
U-Net++ (decoder)	Gradual and precise reconstruction of masks
Skip dense connections	Preserving fine details and improving contours
SCSE attention	Spatial and channel recalibration to better locate nodules

Table III.1: Architecture of the proposed segmentation model (EfficientNet-B3 + U-Net++).

• Learning Strategy:

The model training is carried out in a supervised manner on the TN3K dataset according to the following configuration: for adaptive reduction of the learning rate is used to improve convergence and avoid oscillations at the end of training.

Parameters	Values
Optimizer	AdamW
Learning rate	1e-4
Batch size	4
Number of epochs	70
Split	80% training / 20% validation
Mixed precision	Enabled (AMP)

Table III.2: Configuration of the model's training hyperparameters

- **loss function:**

To effectively address the imbalance between background and target pixels (nodules), a composite cost function is used:

- **The BCE With Logits Loss:** contributes to the stability of learning at the pixel level.
- **The Dice Loss:** is directly oriented towards optimizing similarity or overlap between model predictions and reference annotations.

the final loss function is thus established as a balanced integration of these two components, allowing both a robust convergence of the optimization process and a notable improvement in the quality of the segmentation results.

- **Optimization and robustness techniques:**

Several strategies have been integrated to strengthen the robustness of the model:

- **Test-Time Augmentation (TTA):** averaging of predictions on different transformations (horizontal and vertical flip) to stabilize results.
- **AdamW optimizer:** improves weight regularization.
- **Mixed Precision Training (AMP):** speeds up training and reduces GPU memory consumption.
- **Adaptive scheduler:** automatically adjusts the learning rate according to validation performance.

III.3.2 Feature extraction and classification: tested models:

Following the segmentation phases as well as the pre-processing and data augmentation operations, as detailed in section 2.3, images from the TN5000 database were used for the classification step.

The main objective of this phase is to automatically assign a category to each thyroid image. This is achieved based on the visual features extracted by various deep learning and machine learning models.

In order to evaluate several classification strategies, two complementary methodologies were explored:

- Deep learning models, notably convolutional neural networks (CNNs) and Transformer architectures, which have been trained via learning transfer.
- Traditional machine learning classifiers, applied to high-level features extracted in an automated manner.

Specific models tested include: EfficientNetB0, DenseNet121, ResNet18, Vision Transformer (ViT-B/16), Support Vector Machine (SVM), XGBoost and Random Forest.

III.3.2.1. General classification pipeline:

The overall pipeline adopted in this study is illustrated by the following steps:

1. Loading images from the TN5000 database;
2. Applying Data Transformations and Augments;
3. Resizing images to a fixed size of 224×224 pixels;
4. Standardization of images according to ImageNet statistics;
5. Automatic extraction of deep features;
6. Binary classification using the tested models;
7. Evaluate performance using classification metrics.

The data was divided into two sets:

80% for training;

20% for validation.

III.3.2.2 Preprocessing applied before classification:

Before training the models, the images underwent several treatments:

- Scaling to a uniform resolution of 224×224 pixels;
- The conversion of data to PyTorch tensor format;

- The standardization of pixel intensity values;
- Data enrichment (augmentation) including:
 - random rotations;
 - horizontal symmetries.

The objective of these operations is to:

- strengthen the robustness of the model;
- reduce the risk of overfitting;
- and to increase the diversity of all training samples.

Specifically for architectures that were previously trained on the ImageNet dataset, a standardized standardization was applied, using predefined means and standard deviations:

$$\mu=[0.485,0.456,0.406]$$

$$\sigma=[0.229,0.224,0.225]$$

These values are used by default for models pre-trained on ImageNet to ensure compatibility with the distribution of the original training data. They are not specific to the dataset used, but represent a widely adopted convention in transfer learning.[84]

III.3.2.3 Deep Learning models tested:

➤ *EfficientNetB0* :

EfficientNetB0 is a convolutional neural network architecture optimized to achieve a good compromise between accuracy and computational cost.[85]

This model is based on a compound scaling strategy that simultaneously adjusts:

- the depth of the network;
- the width of the layers;
- the resolution of the images.

In this study, the pre-trained model on ImageNet was used via the PyTorch library.

The last classification layer of the model has been replaced by a linear layer suitable for binary classification:

model.classifier[1] = nn. Linear(..., 2)

The model has been trained for 50 eras with:

- the optimizer Adam;
- a learning rate of 10^{-4} ;
- a batch size of 32.

Loss Function

A loss function like Focal Loss was used:

$$FL(pt) = -\alpha(1-pt)^{\gamma} * \log(pt) \quad (III.1)$$

with:

$$\alpha=0.25$$

$$\gamma=2$$

This function reduces the impact of easy examples and improves the management of class imbalances.

After training, an automatic adjustment of the classification threshold was carried out in order to optimize the F1 score and reduce false positives.

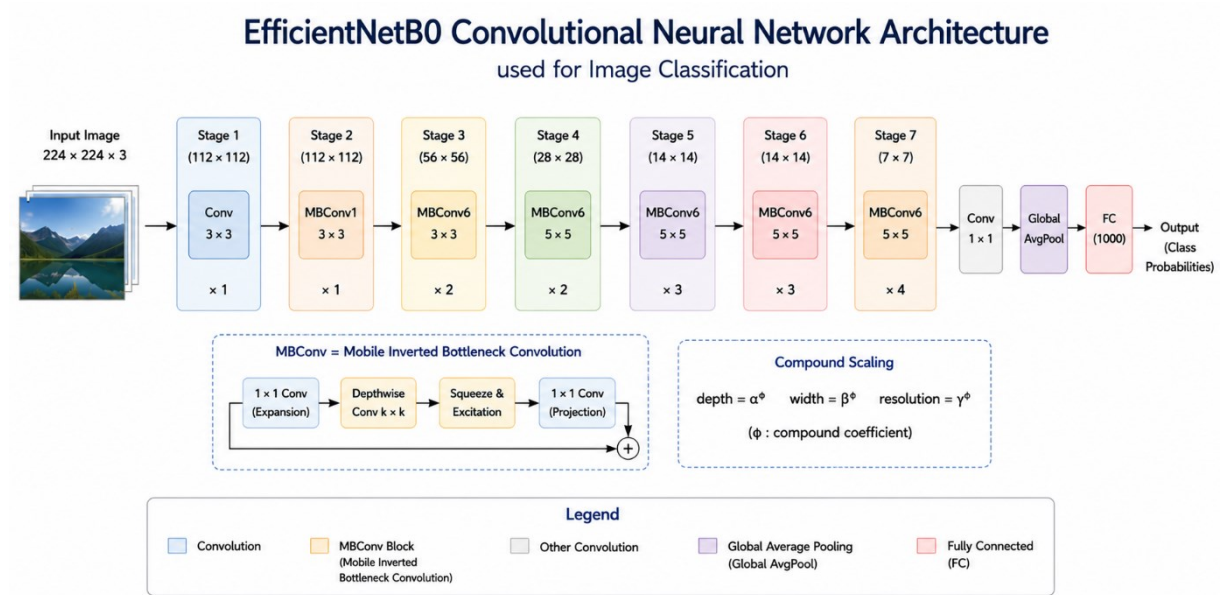


Figure III.3: EfficientNetB0 convolutional neural network architecture used for image classification (personal development, 2026).

➤ *B. DenseNet121 :*

DenseNet121 is a convolutional architecture based on dense connections between layers. Each layer receives the feature maps from all previous layers, which improves:

- the gradient propagation;
- the reuse of characteristics;
- the stability of learning.

Focal Loss was also used to improve class discrimination.[86]

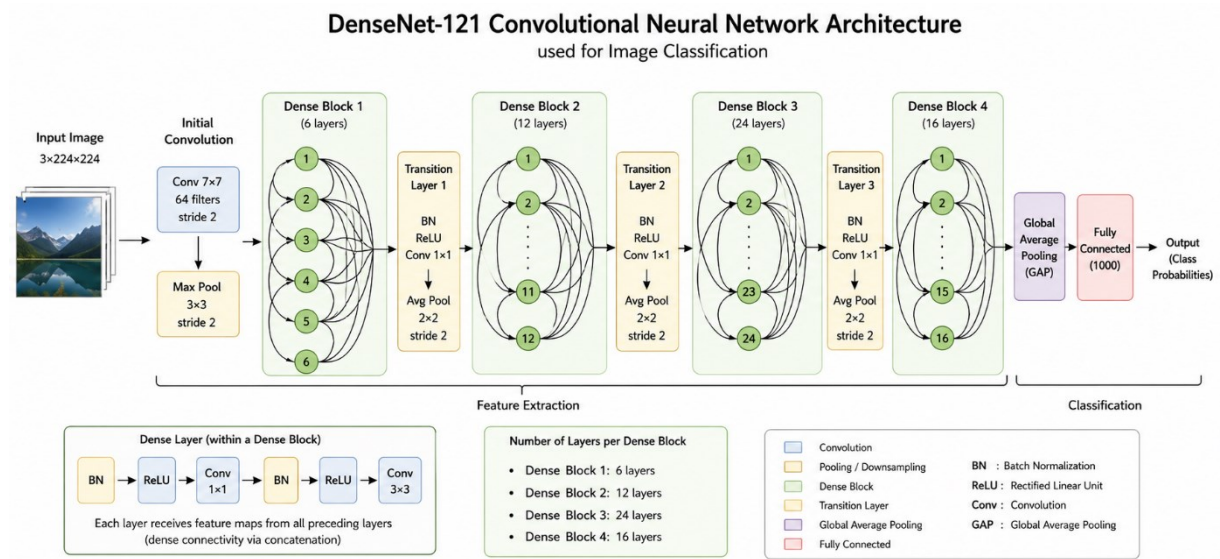


Figure III.4: DenseNet121 convolutional neural network architecture used for image classification (personal development, 2026).

➤ *C. ResNet18 :*

ResNet18 is a deep convolutional network based on the residual connections mechanism. These connections help avoid the problem of gradient disappearance in deep networks.[87]

The architecture has 18 main layers and offers a good compromise between:

- speed of training;
- complexity ;
- performances.

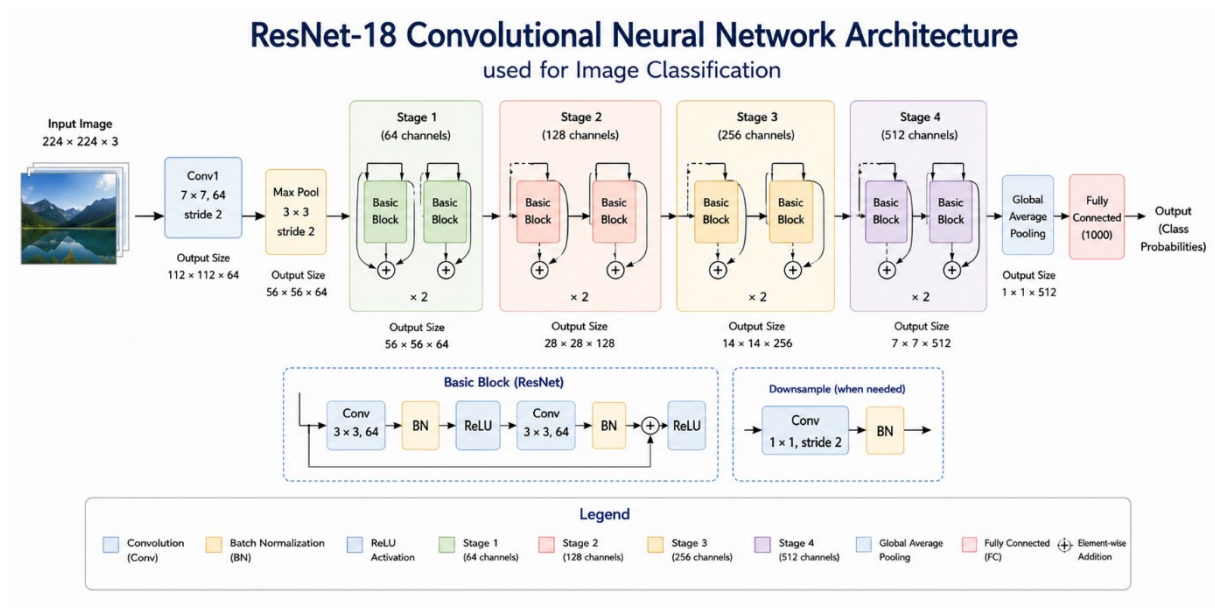


Figure III.5: ResNet-18 convolutional neural network architecture used for image classification (personal development, 2026).

➤ *Vision Transformer (ViT-B/16)* :

The Vision Transformer model applies the Transformer architecture to the field of computer vision.

Unlike traditional CNNs, ViT divides the image into small patches and then treats these patches as sequences.[44]

This approach allows:

- to model the overall relationships between regions of the image;
- to exploit the mechanisms of attention.

This comparison highlights the structural differences between classical convolutional architectures, networks with dense connections, residual networks as well as architectures based on Transformer-type attention mechanisms.

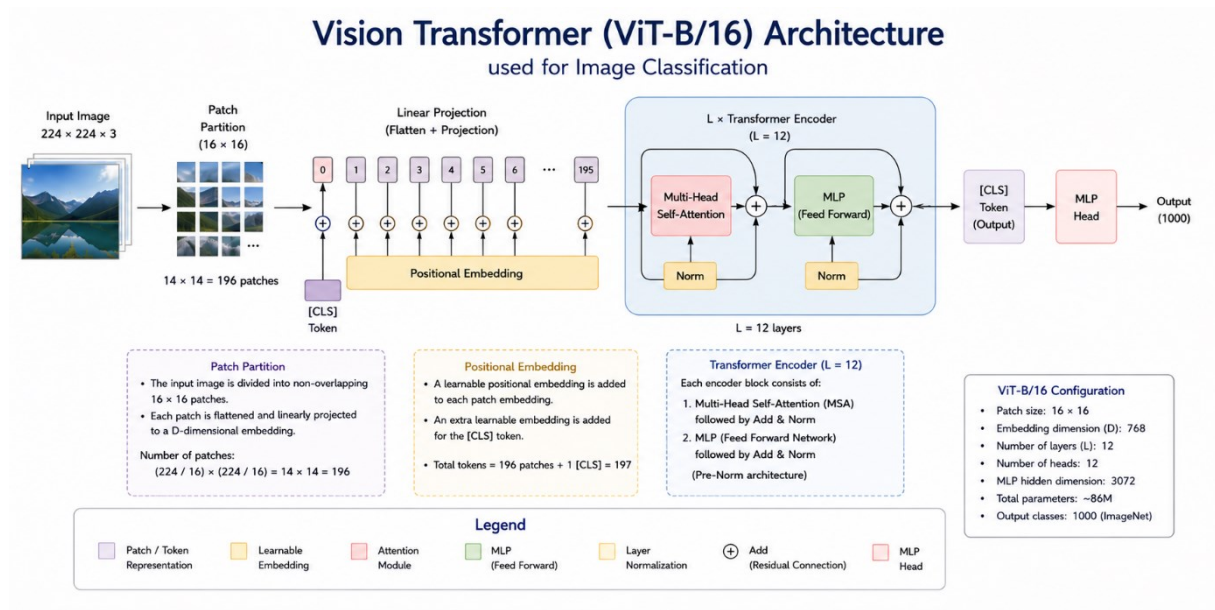


Figure III.6: Vision Transformer (ViT-B/16) convolutional neural network architecture used for image classification (personal development, 2026).

III.3.2.4 Classical methods of machine learning :

In addition to Deep Learning models, several classical classifiers have been evaluated on automatically extracted deep features.

In this approach:

1. DenseNet121 is used only as a feature extractor;
2. The feature maps are compressed by Global Average Pooling;
3. the obtained vectors are normalized with StandardScaler;
4. Classical classifiers are trained on these vectors.

➤ A. Support Vector Machine (SVM) :

The SVM model seeks an optimal hyperplane separating classes with a maximum margin. [42]

In this study:

- an RBF kernel was used;
- the classification probability has been activated (probability=True);
- the deep features extracted by DenseNet121 served as input.

Settings used:

SVC(kernel='rbf', C=1.0, gamma='scale')

➤ **B. XGBoost :**

XGBoost is a boosting method based on sequentially optimized decision trees.[88]

This model is known for:

- its robustness;
- its speed;
- its ability to manage complex data.

Settings Used

n_estimators = 300

max_depth = 6

learning_rate = 0.05

subsample = 0.8

colsample_bytree = 0.8

➤ **C. Random Forest :**

Random Forest is a set of decision trees that are randomly trained and then combined by majority vote.[43]

This method allows:

- to reduce overfitting;
- to improve the stability of predictions.

Settings Used

n_estimators = 300

max_depth = None

The detailed quantitative performance of the different models will be presented in section 4.2.

III.3.3 Evaluation protocol and metrics:

To accurately assess the effectiveness of the proposed method, a uniform experimental procedure was implemented regarding the segmentation and classification of thyroid nodules. The main objective is to measure the ability of algorithms to correctly identify nodules and distinguish non-cancerous from cancerous nodules.

III.3.3.1 Experimental protocol:

The experiments were carried out on the TN3K database for segmentation and TN5000 for classification.

The data was divided into two sets:

- 80% of the images for training;
- 20% of the images for validation and evaluation.

In this work, a simple split of the data into training (80%) and test (20%) was adopted for training and evaluating the models. Although k-fold cross-validation is a more robust approach for estimating model generalization, it was not chosen due to the high computational cost associated with training the deep architectures used (DenseNet121, EfficientNet, Vision Transformer).

Furthermore, the large size of the TN3K and TN5000 datasets allows for a reliable performance estimate with a simple split, thus limiting the risk of high variance in the evaluation.

During the training, the models were evaluated at each time point in order to monitor performance changes and select the best network weights according to validation metrics.

To limit overfitting and improve the generalization of models, several techniques have been applied:

- increase in data (rotations, horizontal flips);
- image normalization;
- transfer of learning from pre-trained models to ImageNet;
- regularization via AdamW and Focal Loss;
- use of mixed precision (AMP).

The final performance was analyzed using:

- learning curves;

- confusion matrices;
- OCR curves;
- as well as several quantitative metrics adapted to the tasks of segmentation and classification.

III.3.3.2. Evaluation metrics for segmentation:

The evaluation of nodule segmentation is based on metrics measuring the degree of similarity between the masks predicted by the model and the reference annotations (Ground Truth).

➤ *A. Dice Similarity Coefficient (DSC):*

The Dice coefficient measures the overlap between the prediction and the actual mask:

$$Dice = \frac{2TP}{2TP+FP+FN} \quad (III.2)$$

where:

TP represents the real positives;

FP the false positives;

FN the false negatives.

A value close to 1 indicates excellent segmentation.

➤ *B. Intersection over Union (IoU):*

The IoU, also known as the Jaccard index, measures the ratio between the intersection and union of segmented regions:

$$IoU = \frac{TP}{TP+FP+FN} \quad (III.3)$$

This metric is particularly used to assess the spatial accuracy of segmented contours.

➤ *C. Accuracy:*

Accuracy represents the overall proportion of correctly classified pixels:

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \quad (III.4)$$

where TN corresponds to the real negatives.

➤ **D. Precision:**

Precision measures the proportion of predicted positive pixels that are actually positive:

$$Precision = \frac{TP}{TP+FP} \quad (III.5)$$

➤ **E. Recall :**

The recall evaluates the model's ability to detect pixels that actually belong to the nodule:

$$Recall = \frac{TP}{TP+FN} \quad (III.6)$$

➤ **F. Specificity :**

Specificity measures the model's ability to correctly recognize background pixels:

$$Specificity = \frac{TN}{TN+FP} \quad (III.7)$$

III.3.3.3. Evaluation metrics for classification :

For the binary classification task of thyroid nodules, several metrics were used to analyze the performance of different deep learning and machine learning models.

➤ **A. Confusion matrix:**

The confusion matrix allows you to visualize the model's correct and incorrect predictions in four categories:

true positives (TP) ;

true negatives (TN);

false positives (FP);

false negatives (FN).

This representation facilitates the analysis of classification errors.

➤ **B. Accuracy:**

Accuracy measures the total proportion of correct predictions:

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \quad (III.8)$$

➤ **C. F1 Score:**

The F1 score represents the harmonic mean between precision and recall:

$$F1 = \frac{2 \times Precision \times Recall}{Precision + Recall} \quad (III.9)$$

This metric is particularly important in medical problems with class imbalance.

➤ **D. OCR and AUC curve:**

The ROC (Receiver Operating Characteristic) curve illustrates the relationship between the rate of true positives and the rate of false positives for different decision thresholds.

The area under the curve (AUC – Area Under Curve) is used to measure the overall ability of the model to discriminate between benign and malignant classes:

- an AUC close to 1 indicates excellent discrimination;
- an AUC close to 0.5 corresponds to a random classification.

The different studied configurations were evaluated on the basis of their measurable efficiency, their ability to process new data, their consistency in the face of ultrasound image variations and their computing resources requirements. The objective was to determine the most appropriate systems for automated detection of thyroid nodules, in support of medical diagnosis.

III.4 Results and discussion:

III.4.1. Segmentation performance (Dice, IoU):

This section presents the experimental evaluation of the developed segmentation model, which uses the U-Net++ architecture in conjunction with the EfficientNet-B3 encoder. The effectiveness of this model was analyzed by using the quantitative measures detailed above,

including the Dice similarity coefficient (DSC), the intersection ratio on union (IoU), accuracy, precision, sensitivity (or recall) and specificity.

4.1.1 Quantitative segmentation results:

The following table summarizes the performance obtained on the TN3K basis after training the model.

Metrics	Results
Dice (best)	0.8544 (85.44%)
IoU	0.7545
Accuracy	≈ 0.96
Precision	≈ 0.85
Recall	≈ 0.86
Specificity	≈ 0.989

Table III.3: Performance results of the segmentation model

The results obtained show that the model reaches a Dice coefficient above 85%, reflecting a significant overlap between the predicted masks and the reference annotations. The high value of the IoU also confirms the good spatial accuracy of the segmented contours.

Moreover, the very high specificity (~0.989) indicates an excellent ability of the model to distinguish nodular regions from background tissues, while limiting false positives. The observed balance between accuracy and recall shows that the model manages to efficiently detect nodules without generating an excessive number of segmentation errors.

III.4.1.1 Learning curve analysis:

The analysis of performance indicators observed during the learning phase reveals a convergence of the model that is both progressive and stable. The loss function shows a constant decrease, while the Dice and IoU scores increase rapidly during the first periods before reaching a stabilization phase.

This development reflects:

- a good learning ability on the part of the model;
- the absence of divergence during training;

- an adequate generalization on the validation data.

Consistent performance can be attributed to the combined integration of:

- transfer learning;
- the use of a hybrid cost function (ECB + Dice Loss);
- methods of increasing data;
- as well as mixed precision (MPA).

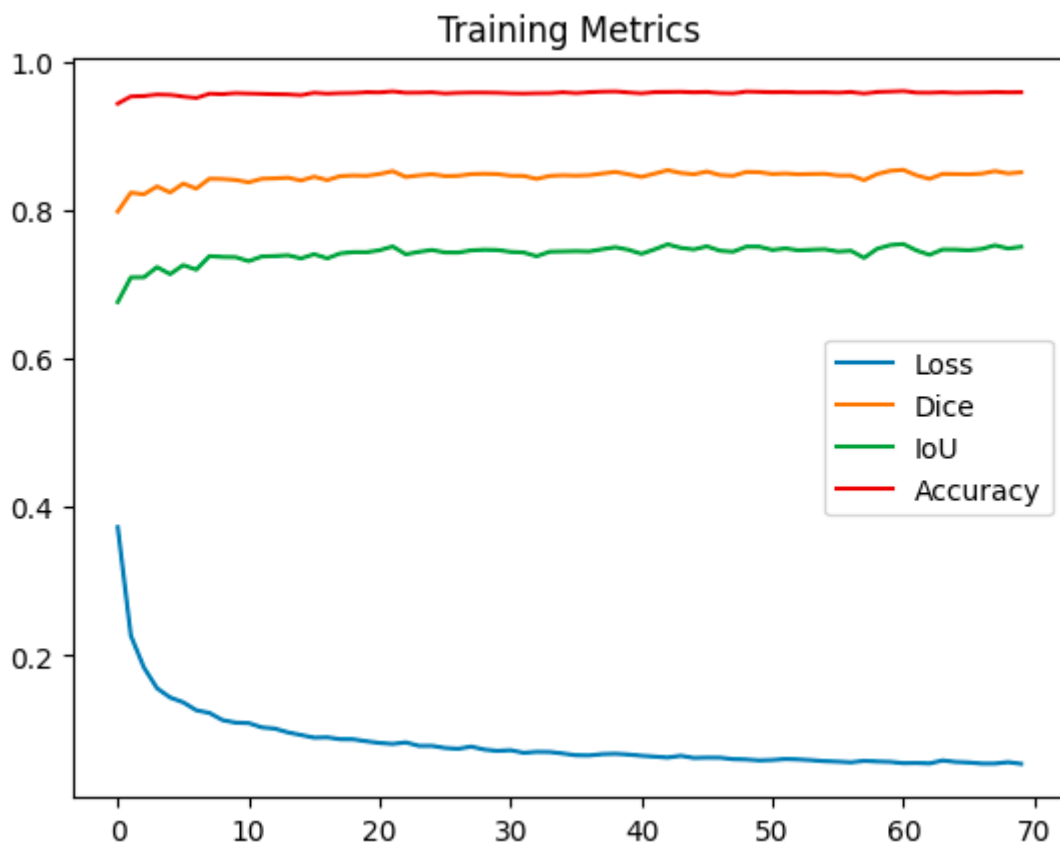


Figure III.7: Monitoring model learning: evolution of the loss function and performance metrics (Accuracy, Dice, IoU)

III.4.1.2 Qualitative analysis of segmentation results:

Visual analysis of the results shows that the model successfully locates the majority of thyroid nodules despite:

the low contrast of ultrasound images;

the presence of speckle noise;

and the heterogeneity of nodular forms.

The predicted contours remain broadly consistent with reference annotations, including for irregularly shaped nodules.

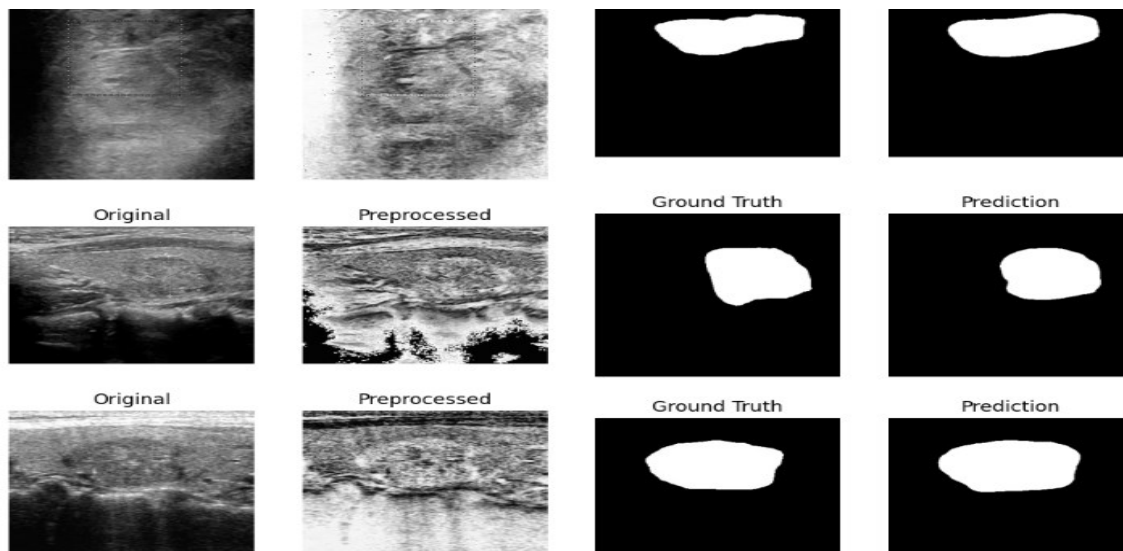


Figure III.8: Illustration of the segmentation process of thyroid nodules: comparison between original images, pre-processed images, real masks (ground truth) and masks predicted by the model.

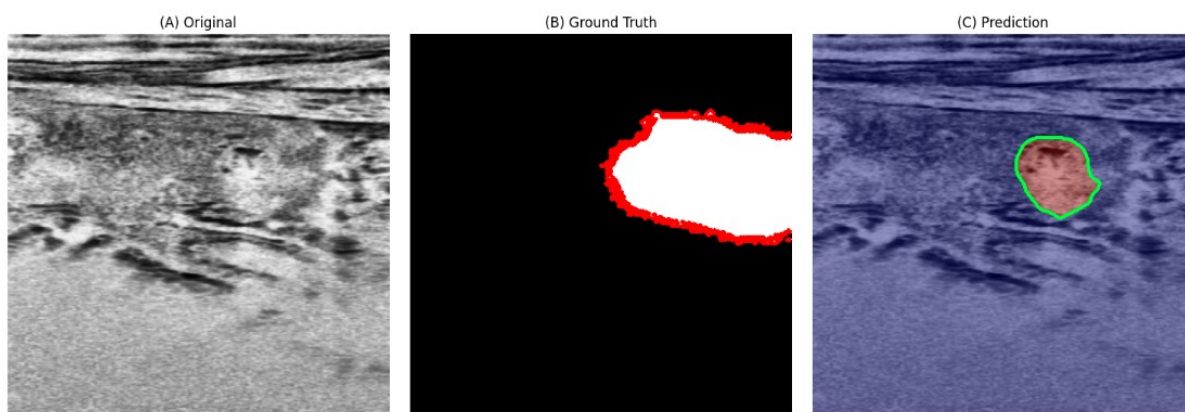


Figure III.9: Example of segmentation of a thyroid nodule: (A) original ultrasound image, (B) ground truth mask, (C) segmentation result predicted by the model.

III.4.2 Classification performance (accuracy, AUC, sensitivity, specificity):

This section discusses the comparative analysis of various classification models used with the TN5000 dataset, in order to distinguish between benign and malignant nodules. The effectiveness of these models was measured using the following indicators:

the accuracy,

the F1 score,

the sensitivity (recall),

the specificity,

the AUC-ROC.

III.4.2.1 Results of Deep Learning Models:

The following table summarizes the overall performance of the tested deep learning architectures.

Model	Accuracy	Precision	Recall	F1-score	AUC
DenseNet121	97.03%	97.18%	98.61%	97.89%	99.05%
EfficientNetB0	96.56%	96.62%	98.53%	97.57%	98.59%
ResNet ResNet18	95.85%	97.35%	96.70%	97.02%	98.72%
ViT (Vision Transformer)	93.08%	93.93%	96.34%	95.12%	97.09%

Table III.4: Performance comparison of deep learning models.

Results show that DenseNet121 and EfficientNetB0 have the best overall performance. DenseNet121 particularly benefits from its dense connections allowing better reuse of features and more efficient propagation of the gradient.

EfficientNetB0 also has excellent performance while maintaining a relatively low computational cost, which makes it particularly interesting in a clinical deployment context.

ResNet18 offers an interesting trade-off between speed of execution and accuracy, while the ViT-B/16 model remains more sensitive to the limited size of the dataset despite its ability to model global dependencies.

III.4.2.2 Analysis of confusion matrices and OCR curves:

Confusion matrices show that deep convolutional models produce a relatively low number of false positives and false negatives.

The obtained OCR curves also confirm the good discriminating capacity of the models, with high AUC values for DenseNet121 and EfficientNetB0.

The most stable performance was observed for:

EfficientNetB0;

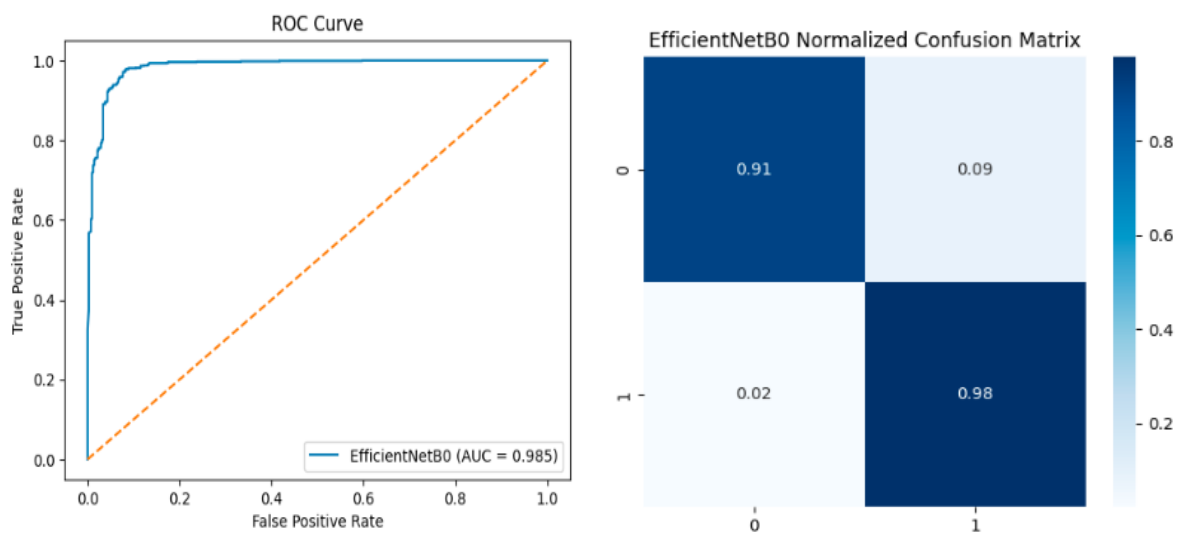


Figure III.10: ROC curve and normalized confusion matrix of the EfficientNetB0 model

DenseNet121;

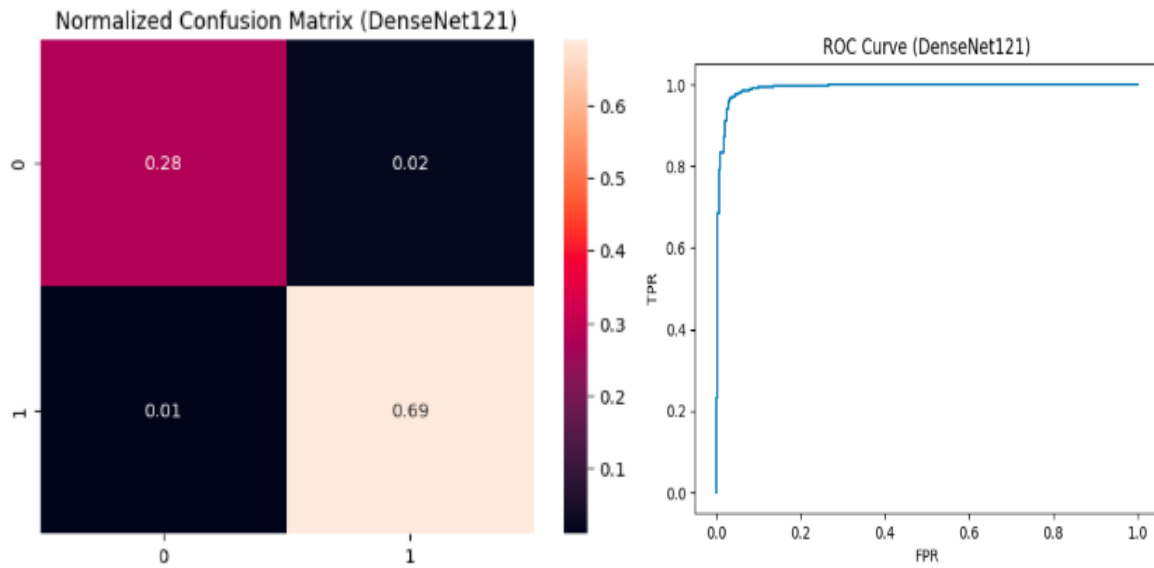


Figure III.11: ROC curve and Normalized Confusion Matrix of the DenseNet121 model.
then ResNet18;

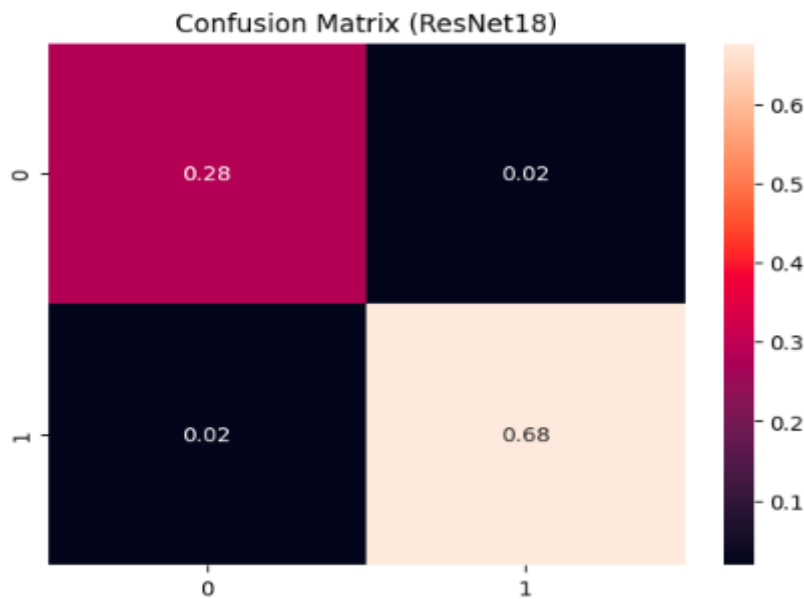


Figure III.12: Normalized confusion matrix of the ResNet18 model.

The ViT-B/16 model shows greater variability in predictions, probably due to:

- the relatively limited number of images;
- to the high variability of ultrasound images;
- as well as the Transformers' dependence on large volumes of data.

Confusion Matrix:

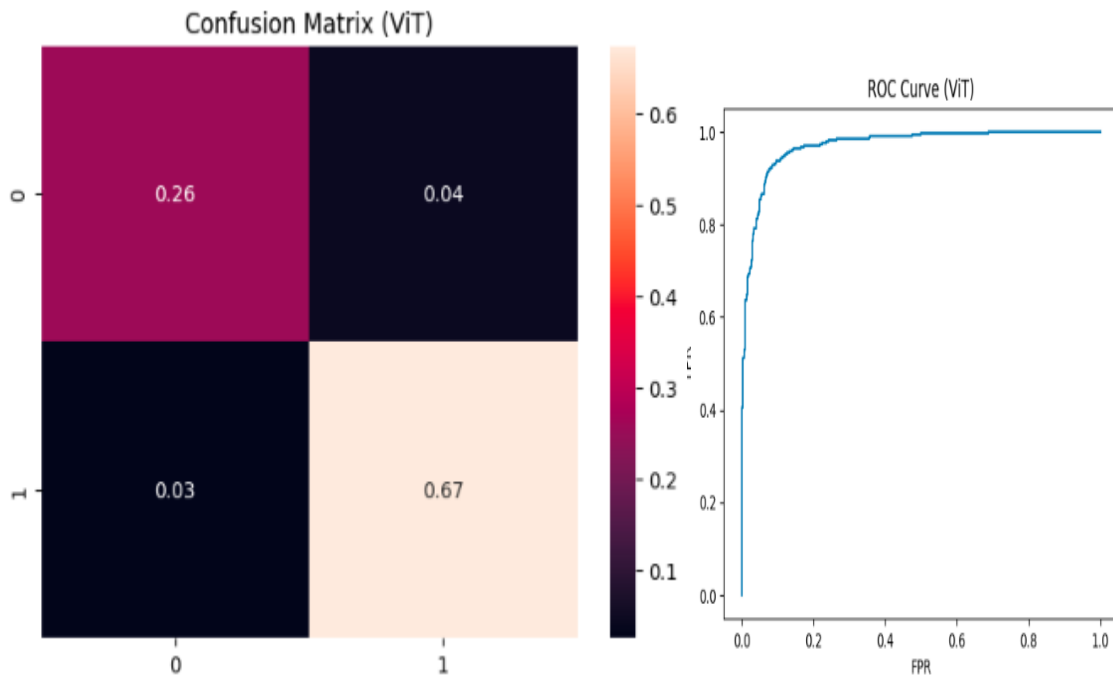


Figure III.13: ROC curve and Normalized Standardized Confusion Matrix of the ViT-B/16 model.

III.4.2.3 Results of classical machine learning methods :

Classical machine learning methods were also evaluated from the deep features extracted by DenseNet121.

Model	Accuracy	Precision	Recall	F1-score	AUC
XGBoost	94.56%	93.75%	98.83%	96.22%	97.58%
Random Forest	94.36%	92.77%	99.71%	96.12%	98.78%
Support Vector Machine SVM	88%	87.18%	97.14%	91.89%	94.08%

Table III.5: Comparison of classical machine learning methods.

The results show that hybrid approaches combining deep feature extraction and traditional classifiers can produce competitive performances.

XGBoost gets the best results among classical methods thanks to its ability to model complex relationships between extracted features.

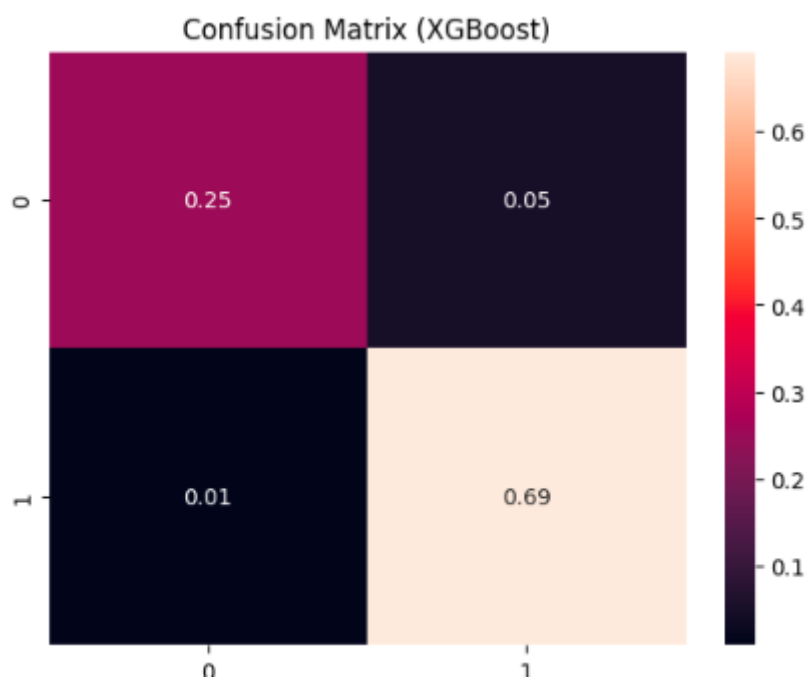


Figure III.14: Normalized confusion matrix of the XGBoost model.

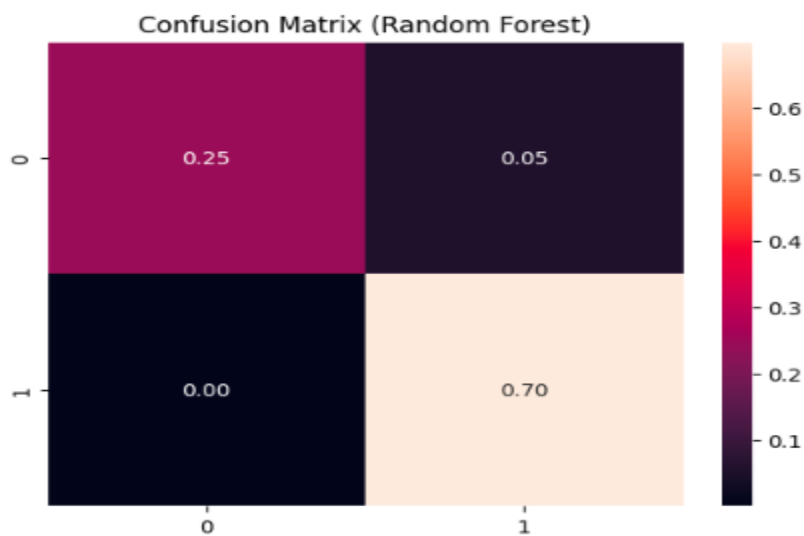


Figure III.15: Normalized confusion matrix of the Random Forest model.

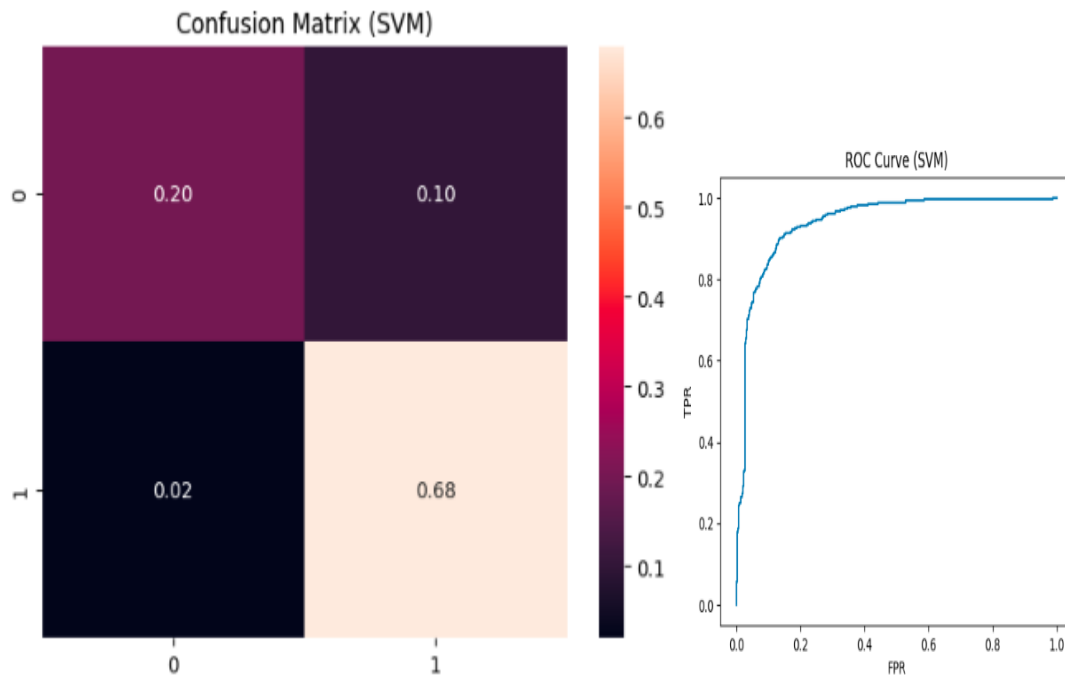


Figure III.16: ROC curve and Normalized Confusion Matrix of the SVM model.

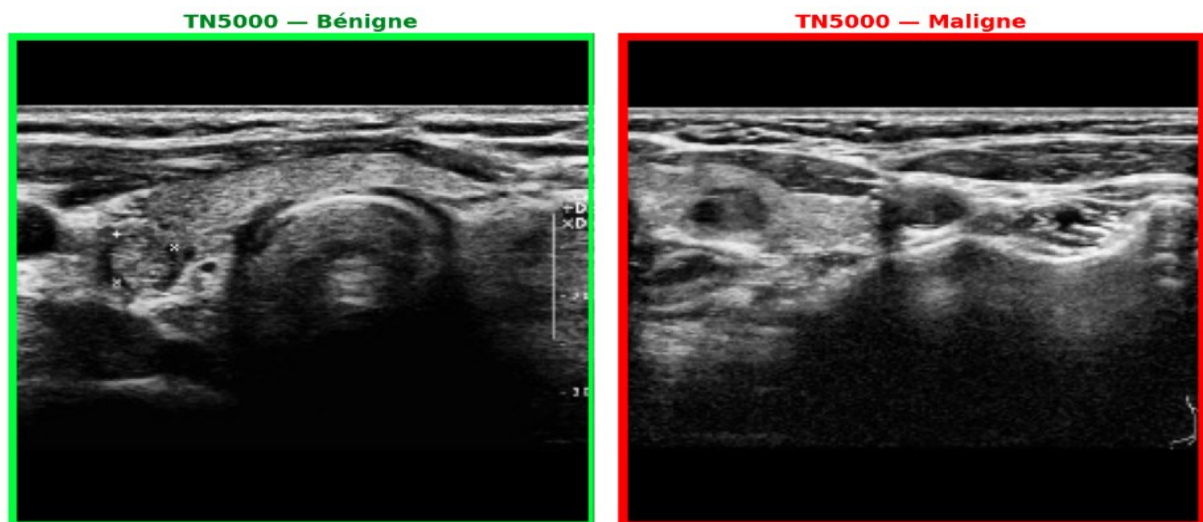


Figure III.17: Visualization of the contours of benign and malignant nodules after preprocessing of the TN5000 database.

II.4.3 Comparison with the state of the art and analysis of errors:

This section presents a comparative analysis of the performance obtained by the proposed models compared to recent literature work applied to TN3K and TN5000 bases. The objective is to assess the relevance of the approaches developed in this study for both segmentation and classification of thyroid nodules. An analysis of the main errors observed is also presented in order to identify the limits of the proposed system.

III.4.3.1. Comparison of segmentation performance:

The results obtained by the segmentation model based on U-Net++ with an EfficientNet-B3 encoder were compared to the main recent architectures reported in the literature on the TN3K basis.

Authors	Method	Main idea	Reported performance	Commentary
Ronneberger et al. (2015) [35]	U-Net	Basic biomedical segmentation	Dice $\approx$ 0.70–0.75	limited on the blurry outlines
Zhang et al. (2018) [90]	ResUNet	Residual connections to improve learning	Dice $\approx$ 0.80–0.84	Better capture of nodular structures
Okta et al. (2018) [89]	U-Net with attention	Highlighting the relevant regions	Dice $\approx$ 0.82–0.85	Reduce Localization Errors
Zhou et al. (2018) [93]	U-Net++	Local and global information merging	Dice $\approx$ 0.85–0.86	Adapted to variations in size and shape
Cao et al. (2021) [91]	Swin-UNet / hybrides Transformer	Modeling the global context	Dice comparable or slightly higher	Very efficient, but more expensive in calculation
Our approach	(U-Net++ + EfficientNet-B3)	Encodeur profond + skip connections denses + attention SCSE	Dice = 0.8544	Good compromise between precision and robustness

Table III.6: Comparison of segmentation performance with the state of the art .

The results obtained show that the proposed model achieves performances comparable to the most recent advanced architectures. The integration of a pre-trained EfficientNet-B3 encoder and dense U-Net++ connections allows for better multi-scale feature extraction and more accurate reconstruction of nodular contours.

Furthermore, the addition of the SCSE attention mechanism improves the localization of relevant regions in ultrasound images, especially in areas with low contrast.

Unlike hybrid Transformer architectures, the proposed model also maintains a more moderate computational cost while maintaining a high level of accuracy.

III.4.3.2. Comparison of classification performance :

The performance of the different classification models was also compared with recent work applied to the TN5000 database.

Authors	Method	Task	Accuracy
He et al. (2016) [87]	ResNet-50	Classification	89–91%
He et al. (2016) [87]	ResNet-101	Classification	92–94%
Tan & Le (2019) [85]	CNN-Transformer hybride	Classification	93–96%
Liu et al. (2021) [92]	Swin Transformer	Classification	90–93%
Our approach	DenseNet121 (suggested)	Classification	97.03%
Our approach	EfficientNetB0 (proposed)	Classification	96.56%

Table III.7: Comparison of classification performance with the state of the art.

The DenseNet121 and EfficientNetB0 models developed in this study have overall performance consistent with those reported in recent literature.

DenseNet121 was particularly distinguished thanks to:

- to its dense connections;
- to better gradient propagation;
- and to an effective reuse of the deep characteristics.

EfficientNetB0 also offers an excellent trade-off between diagnostic accuracy and computational cost, making it suitable for medical diagnostic support systems that require fast execution.

Transformers-based models, such as ViT-B/16, show a good global image modeling capacity, but remain more dependent on the size of datasets and available hardware resources.

III.4.3.3 Segmentation error analysis:

Despite the good performance achieved, some segmentation errors persist in complex cases, particularly for:

small, faintly visible nodules;

irregular contours;

images heavily noisy due to speckle; and

low contrast between the nodule and surrounding tissues.

False negatives occur mainly with small, low-contrast nodules, while some false positives are due to ultrasound artifacts or hyperechoic structures mistaken for nodular regions. In some cases, the model produces incomplete contours or partial segmentations.

These limitations are primarily due to the variability of ultrasound images, the heterogeneity of medical annotations, and the inherent complexity of ultrasound segmentation.

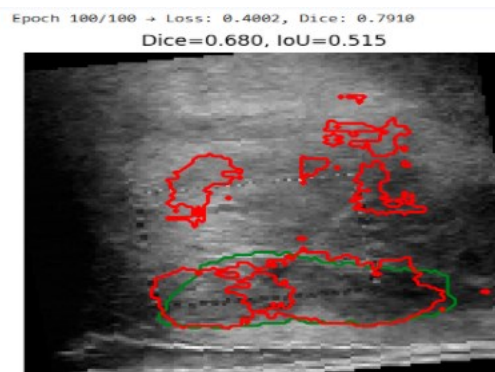


Figure III.18: Example of segmentation failure due to low contrast and ill-defined borders

III.4.3.4. Analysis of classification errors:

The main classification errors are related to:

the visual similarity between certain benign and malignant nodules;

class imbalances;

variations in acquisition devices; and

and morphological differences between patients.

False positives mainly concern benign nodules with suspicious characteristics such as irregular borders or macrocalcifications. False negatives occur primarily with small or poorly contrasted malignant nodules.

Deep learning models generally show better generalization capabilities than traditional machine learning methods, but more complex architectures, such as transformers, require more data and computational resources.

III.4.3.5 Interpretability of Model Decisions Using Grad-CAM:

To increase clinicians' confidence in the model's predictions and identify the regions on which it bases its decisions, we applied the Grad-CAM (Gradient-weighted Class Activation Mapping) method to the DenseNet121 model, the best-performing model in our study.

Grad-CAM generates a heatmap overlaid on the original ultrasound image, highlighting the areas that contributed most to the neural network's final decision. Red/yellow regions indicate high importance, while blue/green areas indicate low contribution.

a) True positive case (malignant nodule correctly identified)

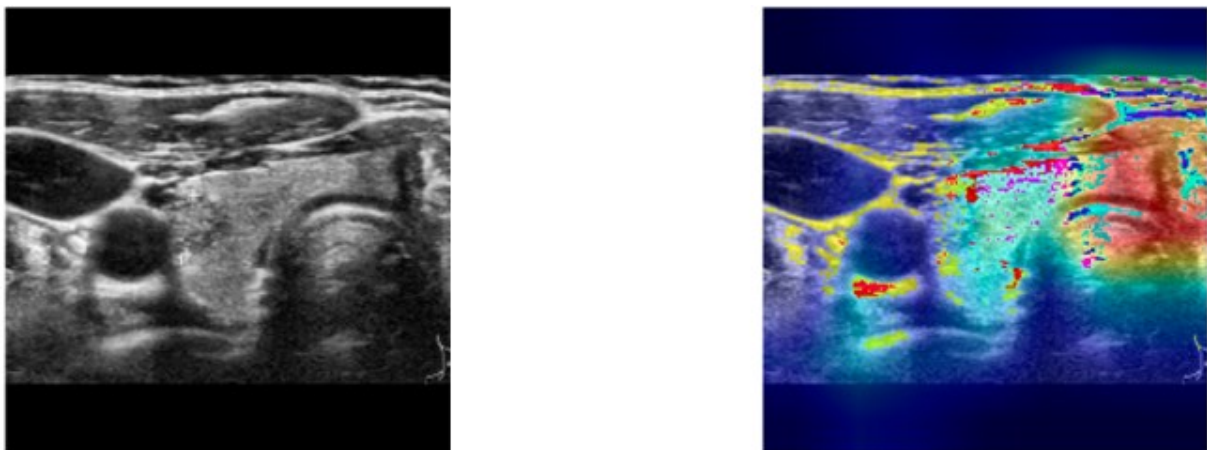


Figure III.19: Example of a true positive with Grad-CAM visualization of the DenseNet121 model.

Analysis of Figure III.16 shows that the model focuses its attention primarily on the nodular area itself, with an intense (red) heatmap centered on the nodule. This behavior is expected and reassuring: the model learns to look at the correct anatomical regions.

b) Case of a false positive (benign nodule mistakenly classified as malignant)

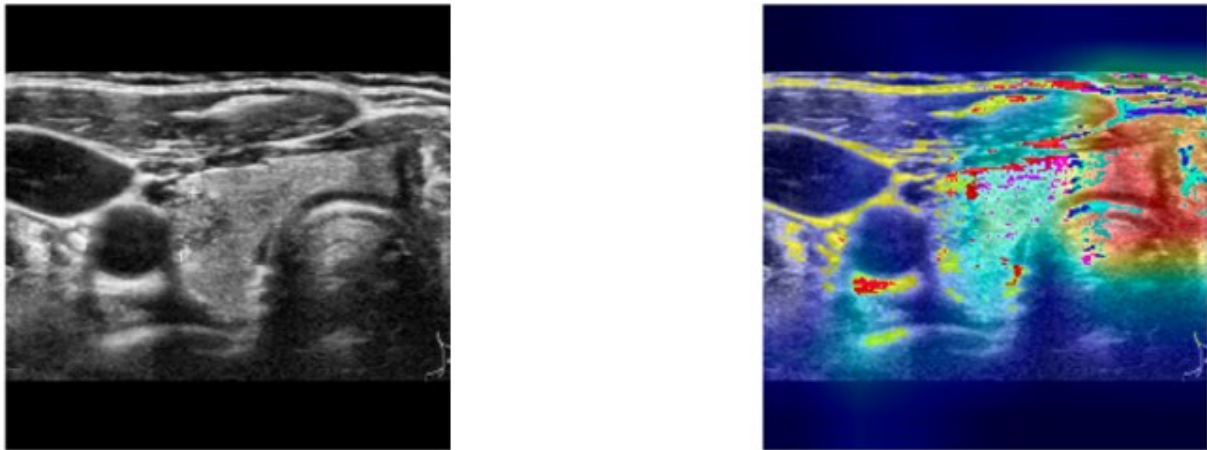


Figure III.20: Example of a false positive accompanied by a Grad-CAM visualization of the DenseNet121 model.

In this case of an error, the heatmap reveals several areas of concern:

- Moderate activation in the benign nodule (green arrow),
- But also significant activation in a peritumoral area (red arrow), likely an ultrasound artifact or a blood vessel.

This observation suggests that the model was misled by abnormal surrounding structures, mistakenly interpreted as signs of malignancy (e.g., hypervascularization, microcalcifications).

c) True negative case (benign nodule correctly identified)

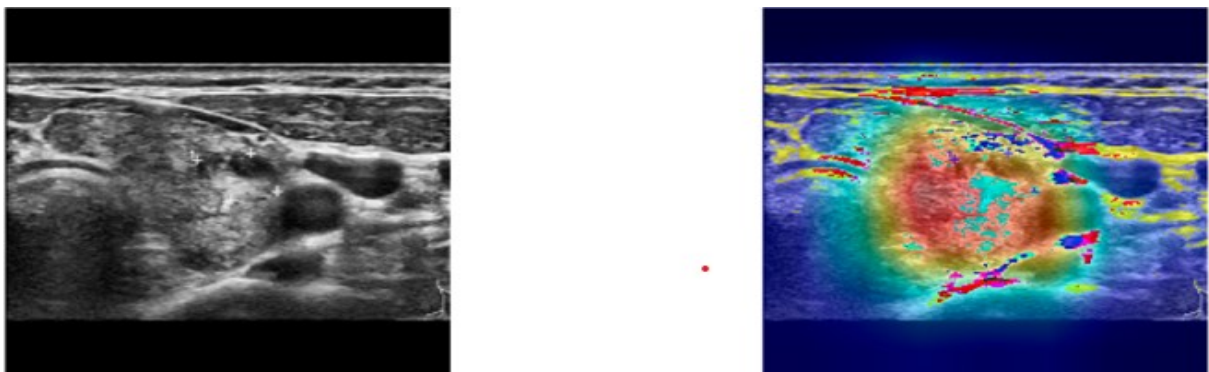


Figure III.21: Example of a true negative with Grad-CAM visualization of the DenseNet121 model.

For this well-classified benign nodule, we observe that:

- Activation is low and diffuse,
- The model does not show any marked peaks of attention,
- No suspicious regions are overactive.

This behavior is consistent with a model that has not detected any alarming features.

d) Clinical interpretation and value of Grad-CAM

Grad-CAM analysis provides several useful insights:

1/ Key observations:

- 1.1/ The model looks at the right areas in true positives.
- 1.2/ False positives are often associated with artifacts or vessels.
- 1-3/ The absence of clear activation corresponds to true negatives.

2/ Implications:

- 2-1/ Validation of the training
- 2-2/ Area for improvement: better preprocessing of these cases.
- 2-3/ Consistency with the absence of malignancy

This analysis confirms the value of explainable AI (XAI) methods for the clinical validation of diagnostic assistance models.

III.4.4 Graphical interface of the proposed system:

As part of this project, two interactive web applications were developed using the Streamlit framework, transforming trained artificial intelligence models into diagnostic support tools directly accessible to healthcare professionals. The design of these interfaces addresses a dual requirement: scientific rigor—by displaying performance metrics, calibrated confidence levels, and TI-RADS estimations—and usability, through a linear and guided workflow. Streamlit was

chosen for its ability to deploy Python applications in real time without requiring separate front-end development.

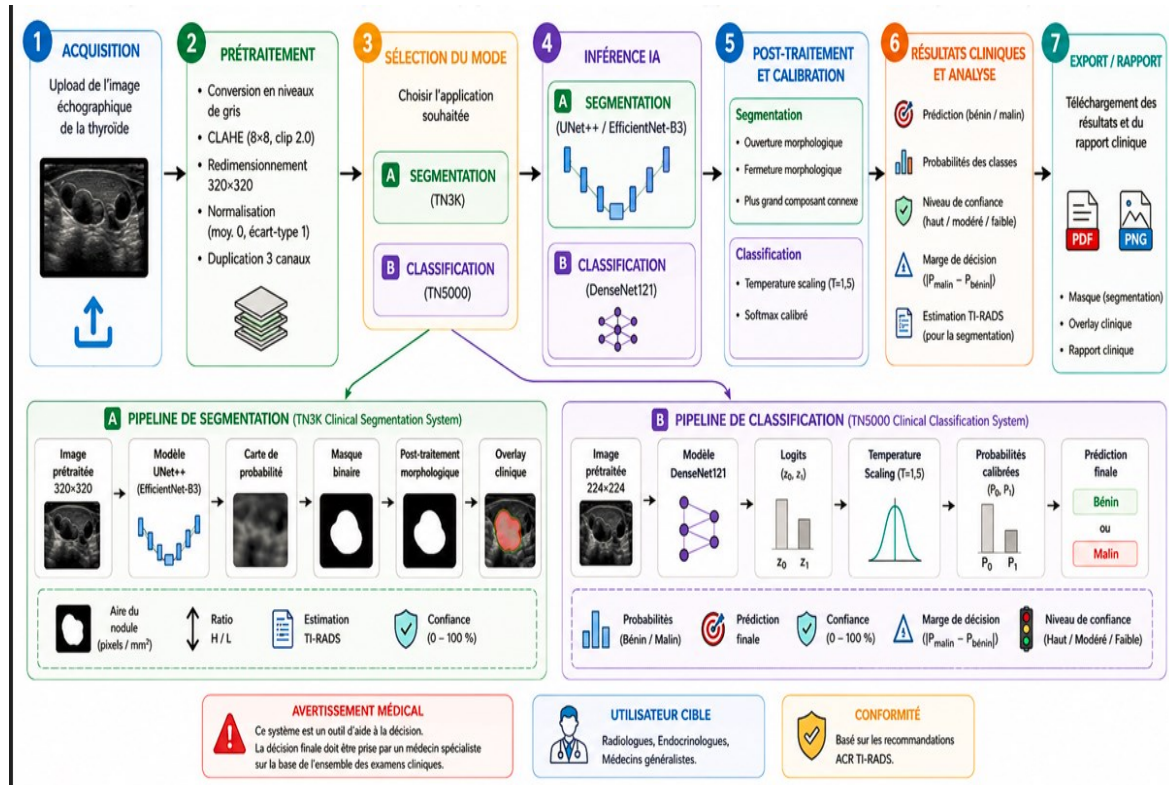


Figure III.22: Global architecture and functional pipeline of thyroid diagnostic support applications.

III.4.4.1 Segmentation Application — TN3K Clinical Segmentation System :

1/ Objective of the application

The objective of this application is to automate the segmentation of thyroid nodules in ultrasound medical images using deep learning, and to provide a clinical visualization and structured analysis of the result. The application is based on the following elements:

- Dataset used: TN3K (Thyroid Nodule 3K images)
- Main task: Binary semantic segmentation
- Framework: PyTorch + Streamlit
- Architecture: UNet++ with EfficientNet-B3 encoder

2/ Academic header and scientific contextualization

CHAPTER III : DESIGN, DEVELOPMENT AND EVALUATION OF THE PROPOSED SYSTEM

The interface header acts as a scientific identity card of the system: it defines the academic context (Master 2026–2027), specifies the architecture used (UNet++ / EfficientNet-B3), and indicates compliance with ACR TI-RADS recommendations.



Figure III.23: Academic header of the TN3K Clinical Segmentation System application.

3/ Image preprocessing

The ultrasound image undergoes a five-step preprocessing pipeline: grayscale conversion, local contrast enhancement using CLAHE (8×8 tiles, factor 2.0), resizing to 320×320 pixels, statistical normalization (zero mean, unit variance), and channel duplication to match the EfficientNet-B3 encoder input requirements. This pipeline ensures that every input image is processed consistently with training data, which is essential for reliable and reproducible predictions.

4/ Segmentation using the UNet++ model

The preprocessed image is passed to the loaded UNet++ model, which analyzes each pixel of the ultrasound image and generates a probability map indicating the likelihood of each pixel belonging to a thyroid nodule.

5/ Morphological post-processing

After prediction, a three-step post-processing stage refines the binary mask: morphological opening (erosion followed by dilation) removes small noisy regions, morphological closing (dilation followed by erosion) fills internal holes, and finally only the largest connected component is retained, corresponding to the main detected nodule.

6/ Clinical overlay construction

The clinical overlay is generated by fusing the original image with the segmented mask: a semi-transparent red mask (35% opacity) highlights the nodule region over the original image (preserved at 75% intensity, with the remaining 10% corresponding to edge interpolation), and a green contour with 2-pixel thickness is drawn to enable precise visual localization for radiologists.

7/ Quadruple visualization — real-time pipeline

The interface simultaneously displays four side-by-side columns: the original uploaded image, the predicted binary mask (white pixels for nodule, black background), the grayscale probability heatmap representing pixel-wise model confidence, and the clinical overlay combining the red mask and green contour over the original image—providing an immediate comparative interpretation of the model's reasoning.

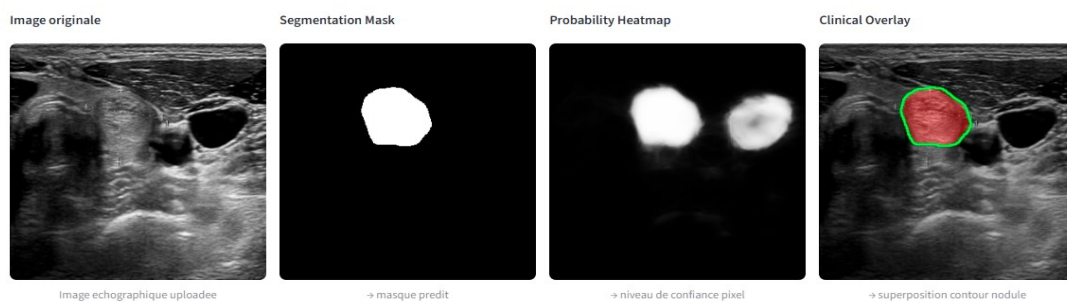
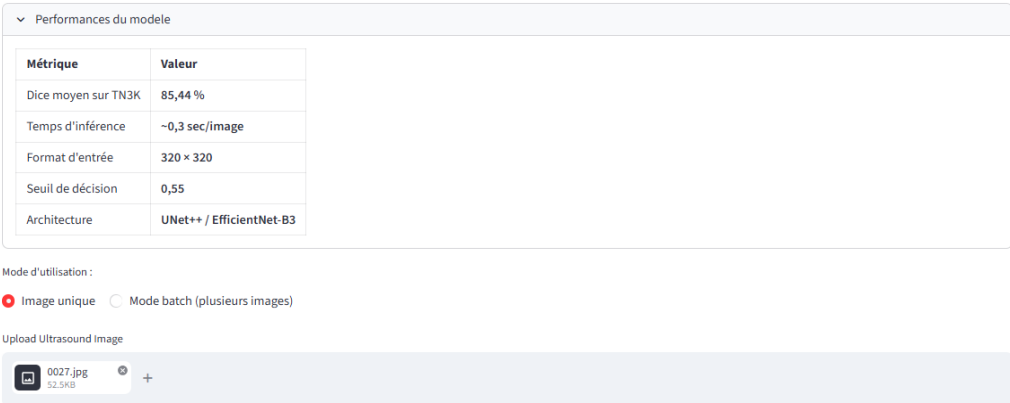


Figure III.24: Quadruple visualization — original image, predicted mask, probability heatmap, and clinical overlay.

This system is designed as a diagnostic support tool and should not be interpreted as an autonomous medical decision system. Final interpretation of results remains the sole responsibility of the physician.

8/ Performance table and medical disclaimer



The screenshot shows a web interface for a medical image analysis model. At the top, there is a section titled 'Performances du modele' which contains a table of performance metrics. Below this table, there are radio buttons for selecting the mode of use: 'Image unique' (selected) and 'Mode batch (plusieurs images)'. At the bottom, there is a file upload section labeled 'Upload Ultrasound Image' with a file named '0027.jpg' (52.5KB) already uploaded.

Métrique	Valeur
Dice moyen sur TN3K	85,44 %
Temps d'inférence	~0,3 sec/image
Format d'entrée	320 × 320
Seuil de décision	0,55
Architecture	UNet++ / EfficientNet-B3

Mode d'utilisation :

☒ Image unique ☐ Mode batch (plusieurs images)

Upload Ultrasound Image

0027.jpg
52.5KB

Figure III.25: Performance table and mode selection interface.

The Dice score (85.44%) is used as a global evaluation metric of the model on the TN3K dataset, while the confidence displayed in the interface is a local measure computed at the level of each prediction.

9/ Clinical analysis, calibrated confidence, and TI-RADS estimation

The clinical analysis section provides a structured report in three complementary levels. The main diagnosis automatically classifies the nodule into four categories based on segmented area (No Significant Nodule, Small, Thyroid Nodule Detected, Large). The confidence level is expressed as a percentage and categorized into three calibrated thresholds: High confidence ($S \geq 0.90$), indicating a reliable prediction; Moderate confidence ($S \in [0.70; 0.90[$), recommending expert verification; and Low confidence ($S < 0.70$), indicating an uncertain result.

Finally, a simplified TI-RADS estimation module is provided, based on the nodule area in pixels and the height-to-width ratio. For the tested image, the estimated score is TR2 (Not suspicious), with an H/W ratio of 0.91 and a confidence of 99.98% classified as High confidence.

Clinical Analysis

Diagnosis : Thyroid Nodule Detected

Detected Region Area : 5447 pixels

Model Confidence : 99.98%

Niveau : High confidence ($S \geq 0,90$)

Interprétation : Diagnostic fiable, confiance élevée

Estimation TI-RADS

Score estimé : TR2 — TR2 — Non suspect

Ratio H/L : 0.91

• Taille intermédiaire

⚠ Estimation indicative — ne remplace pas une évaluation radiologique complète.

Segmentation completed successfully.

Figure III.26: Clinical analysis — diagnosis, calibrated confidence, and TI-RADS estimation (TR2 score, 99.98% confidence).

10/ Validation using the TN3K dataset

The system's reliability is validated by direct comparison between the automatically generated mask and the expert-annotated ground truth masks from the TN3K dataset. The observed agreement in terms of localization, shape, and size confirms that the application faithfully reproduces the trained model's behavior, achieving a mean Dice coefficient of 85.44% across the dataset.

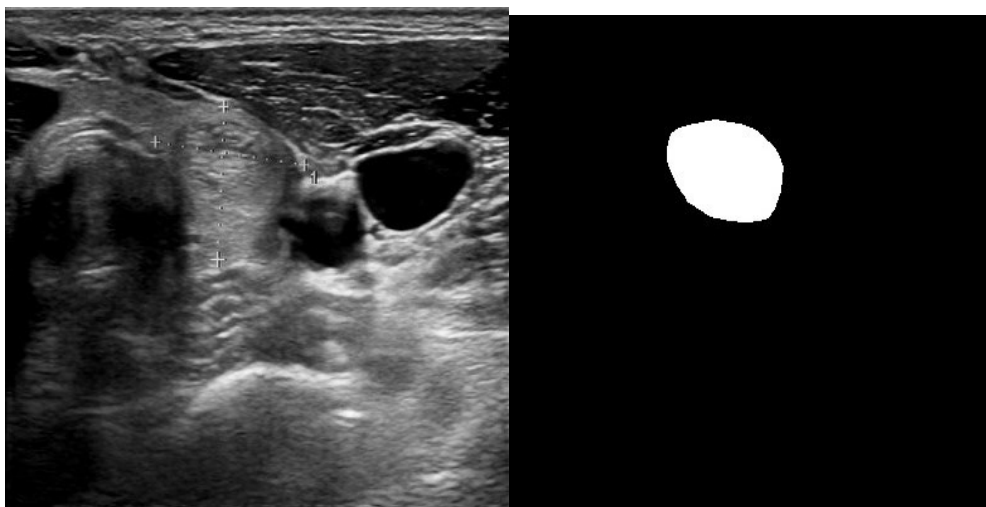


Figure III.27: Original image and annotated mask from TN3K dataset.

11/ Result export and technical information

The interface includes two export buttons allowing the user to save the segmented mask and the clinical overlay in PNG format, facilitating archiving and integration into real radiological workflows. A technical section summarizes model parameters and provides installation and execution commands, making the system fully reproducible and independently deployable.

The application was designed to operate with standard image formats such as PNG and JPG, commonly used in AI research datasets. However, it does not yet support the DICOM format, which is the standard for medical image interoperability in hospital environments. Native DICOM integration (including patient metadata, ultrasound series handling, and PACS compatibility) remains a future development perspective essential for real clinical deployment.

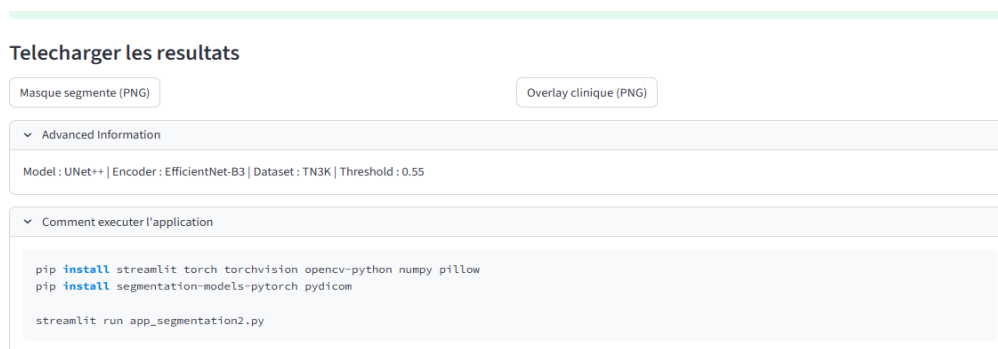


Figure III.28: Download buttons and technical information section.

III.4.4.2 Classification Application — TN5000 Thyroid Classification System

1/ Objective of the application

The objective of this application is to automatically classify thyroid nodules as benign or malignant from ultrasound images using a DenseNet121 model trained on the TN5000 dataset.

Complementary to the segmentation application, it addresses a different clinical question: not “where is the nodule?”, but “what is its nature?”

- Dataset used: TN5000 (Thyroid Nodule 5000 images)
- Main task: Binary classification (benign / malignant)
- Framework: PyTorch + Streamlit
- Architecture: DenseNet121 with temperature scaling ($T = 1.5$)

2/ Academic header and performance indicators

CHAPTER III : DESIGN, DEVELOPMENT AND EVALUATION OF THE PROPOSED SYSTEM

The classification application shares the same institutional structure as the segmentation system, ensuring visual and conceptual consistency between both tools.

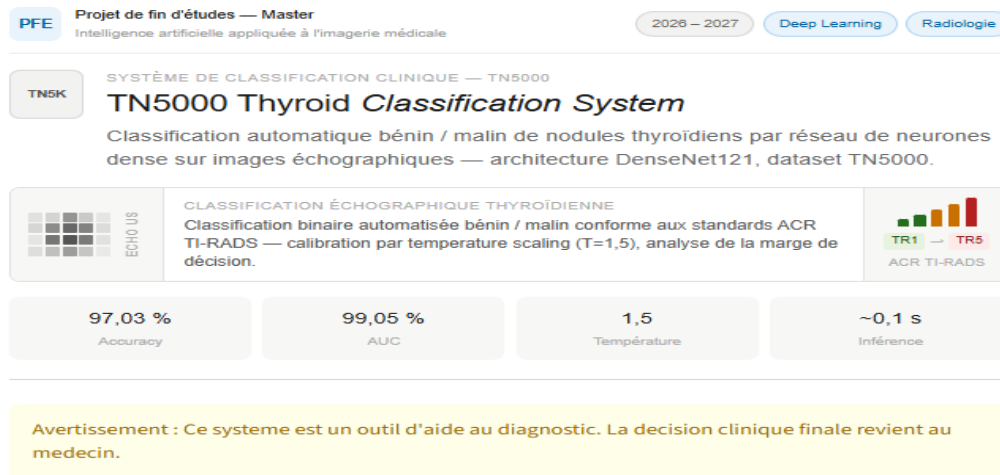


Figure III.28: Academic header of the TN5000 Thyroid Classification System application.

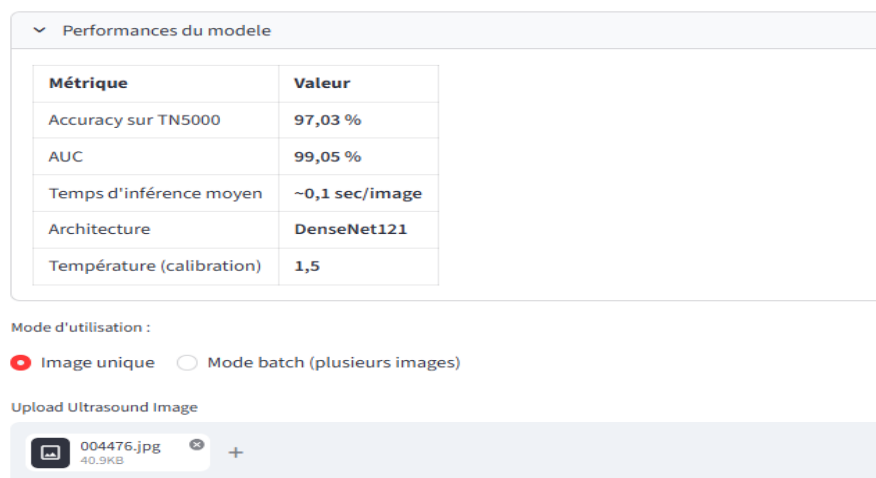


Figure III.29: Performance table, mode selection, and image upload interface.

3/ Prediction pipeline and temperature scaling

The uploaded image is processed using a standard ImageNet normalization pipeline and then passed through the DenseNet121 model. Temperature scaling ($T = 1.5$) is applied to logits to calibrate output probabilities and reduce model overconfidence, which is essential for medical applications.

4/ Image display and clinical results

CHAPTER III : DESIGN, DEVELOPMENT AND EVALUATION OF THE PROPOSED SYSTEM

The application first displays the analyzed ultrasound image in full width before presenting predictions, allowing clinicians to form an independent judgment prior to viewing model output. This approach promotes critical and responsible usage.



Figure III.30 : Malignant case — image 004476.jpg

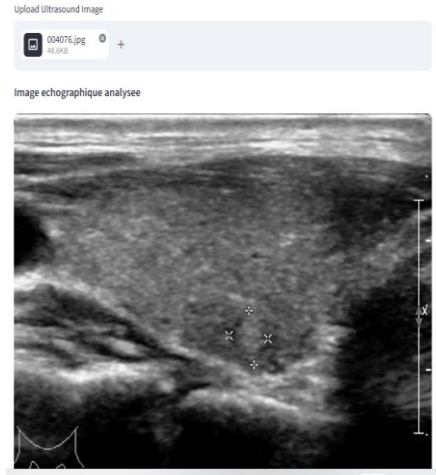


Figure III.31 : Benign case — image 004076.jpg

5/ Clinical results and decision margin analysis

The Clinical Result section provides, for each image: binary prediction (benign/malignant), posterior probabilities for both classes, calibrated confidence level, and decision margin (difference between class probabilities). This margin is a key indicator: a low margin (< 0.2) indicates a borderline case requiring additional expert review.

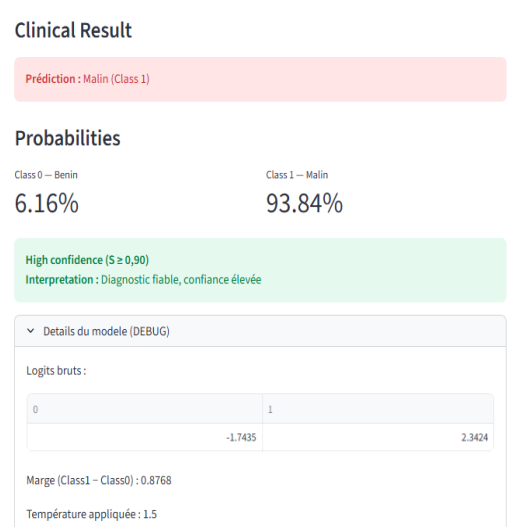


Figure III.32: Malignant case — High confidence (93.84%)

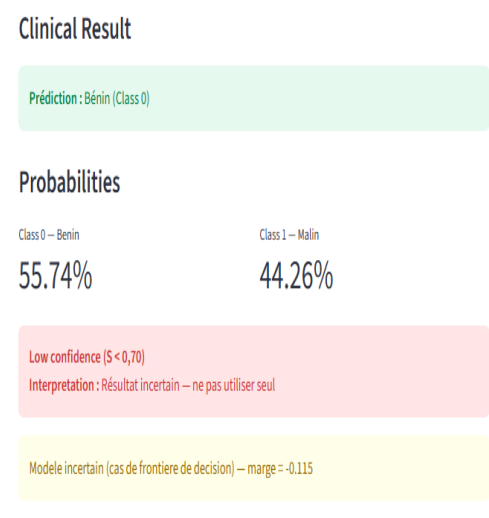


Figure III.33 : Benign ambiguous case — Low confidence (55.74%)

This system is designed as a diagnostic support tool and should not be interpreted as an autonomous decision-making system. Final interpretation of results remains the sole responsibility of the physician.

6/ Technical section and execution

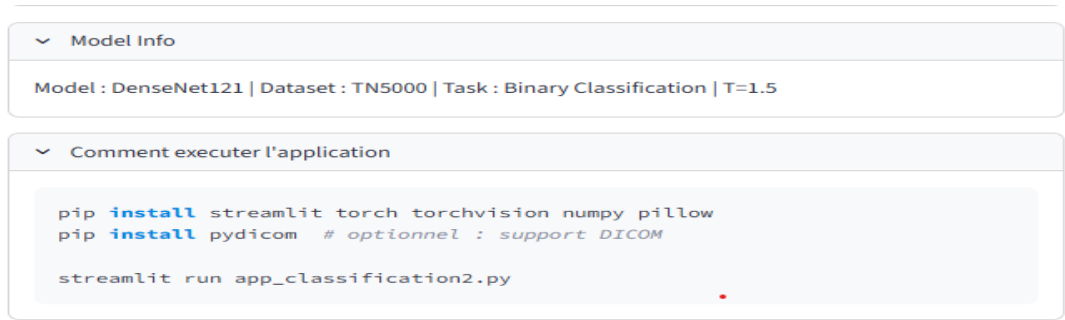


Figure III.34: Technical information and application execution commands.

III.4.4.3 Comparison and Complementarity of the Two Applications

The two applications form an integrated and complementary thyroid diagnostic support system. The following table summarizes their respective characteristics:

Criterion	Segmentation (TN3K)	Classification (TN5000)
Clinical task	Nodule localization	Benign vs malignant nature
Architecture	UNet++ / EfficientNet-B3	DenseNet121
Dataset	TN3K	TN5000
Main metric	Dice : 85,44 %	Accuracy 97,03 % / AUC 99,05 %
Output	Binary mask + overlay	Probabilities + prediction
Calibrated confidence	3 levels $S \geq 0.90 / 0.70 / 0.70$	Same + decision margin
TI-RADS	Yes (area + H/W ratio)	No (direct classification)
Export	PNG mask + overlay	CSV batch report
Batch mode	Yes (ZIP compression)	Yes (CSV report)

Table III.8: Comparison and complementarity of the two system applications.

III.4.5 Limits, perspectives and conclusion of chapter III:

III.4.5.1. Limitations of the study :

Despite the satisfactory results obtained, some limits must be considered. Thyroid ultrasound images show high variability related to acquisition devices, imaging parameters, and anatomical differences between patients, which can reduce the generalization of models, especially for small nodules or low-contrast regions. Moreover, the presence of speckle noise and ultrasound artifacts complicates automatic segmentation and can cause inaccuracies in the predicted contours.

Performance also remains dependent on the quality of medical annotations, the volume of available data, and the imbalance between benign and malignant classes in the TN5000 database. Moreover, some advanced architectures, especially the Transformers-based models, require significant computational resources. Finally, this study was conducted only on public databases; a multicenter clinical validation remains necessary before real use in medical practice.

III.4.5.2 Conclusion to chapter III:

This chapter presented the design, implementation, and evaluation of an automated system for the segmentation and classification of thyroid nodules from ultrasound images. The proposed approach is based on deep learning methods, notably a U-Net++ architecture with an EfficientNet-B3 encoder for segmentation, as well as several classification models (DenseNet121, EfficientNetB0, ResNet18, Vision Transformer), in addition to classical machine learning techniques.

The experiments conducted on the TN3K and TN5000 datasets demonstrate generally high performance, with strong segmentation accuracy and excellent classification capability, particularly for DenseNet121 and EfficientNetB0. The overall results confirm the effectiveness of transfer learning techniques, preprocessing strategies, and data augmentation methods in improving model robustness.

Finally, the integration of the system into an interactive Streamlit application enables the developed models to be transformed into a usable diagnostic support tool. However, some limitations remain, particularly the variability of ultrasound images, class imbalance, and the need for multicenter clinical validation.

Thus, this work represents an effective contribution to the development of thyroid diagnostic support systems, while opening perspectives for further improvements toward more reliable and widely applicable clinical use.

GENERAL CONCLUSION AND PERSPECTIVES

This work focused on the design and development of an automatic diagnostic support system for thyroid nodules from ultrasound images, leveraging recent advances in artificial intelligence and deep learning. Given the inherent limitations of manual ultrasound image interpretation — inter-operator variability, speckle noise, and low contrast — the automation of the diagnostic process presents itself as both a clinical and technological necessity.

The proposed approach is built upon a two-stage complementary pipeline. The first stage, dedicated to the segmentation of nodular regions, was carried out using a U-Net++ architecture integrating an EfficientNet-B3 encoder pre-trained on ImageNet. This combination enabled the efficient extraction of deep, multi-scale features while preserving the accuracy of nodular contours. The results obtained on the TN3K and TN5000 databases confirmed the robustness of this architecture, achieving a Dice Score of 85.44% and an IoU of 75.45%, reflecting faithful and precise nodule localization.

The second stage, devoted to the binary classification of nodules as benign or malignant, placed several Deep Learning and Machine Learning models in competition. Among all evaluated architectures, DenseNet121 distinguished itself with superior performance, reaching an accuracy of 97.03% and an AUC of 99.05%, demonstrating its capacity to capture discriminative representations from the segmented regions.

These results, benchmarked against the state of the art, confirm the relevance and competitiveness of the developed approach. They demonstrate that the integration of modern deep learning architectures, combined with rigorous preprocessing and data augmentation strategies, can achieve a particularly high level of diagnostic performance, capable of providing genuine support to clinical decision-making.

Nevertheless, this work also highlighted certain limitations that deserve to be acknowledged. The intrinsic variability of ultrasound images, class imbalances within the datasets used, and the absence of prospective clinical validation represent obstacles to the immediate generalization of the proposed system.

Looking ahead, several avenues for improvement and extension can be considered. The integration of attention mechanisms and Transformer-based architectures into the segmentation pipeline could further enhance the delineation accuracy of complex nodules. Leveraging multi-domain transfer learning techniques and weakly supervised learning methods would help address the scarcity of annotated data. Furthermore, the development of a user-friendly clinical

GENERAL CONCLUSION AND PERSPECTIVES

interface and the evaluation of the system under real-world conditions of use would constitute essential steps toward its effective integration into medical practice.

Ultimately, this final-year project represents a concrete and significant contribution to the application of artificial intelligence in thyroid medical imaging. It illustrates the considerable potential of deep learning-based CAD systems to assist clinicians, reduce diagnostic subjectivity, and ultimately improve the management of patients with thyroid nodules.

Bibliographic reference:

- [1] Kenhub. (2022). *Thyroid gland: anatomy*. Retrieved from <https://www.kenhub.com/en/library/anatomy/thyroid-gland>
- [2] Kenhub. (2025). *Thyroid gland — Gross Anatomy and Clinical Relevance*. Retrieved from <https://www.kenhub.com/en/library/anatomy/thyroid-gland>
- [3] Kenhub. (2025). *Thyroid gland — Histology: Follicles and C Cells*. Retrieved from <https://www.kenhub.com/en/library/anatomy/thyroid-gland>
- [4] Kenhub. (2025). *Thyroid gland — Endocrine Physiology: T3, T4 and Calcitonin*. Retrieved from <https://www.kenhub.com/en/library/anatomy/thyroid-gland>
- [5] Arrangoiz, R., Fauza, D., Godoy, A., & Janjua, R. (2018). Comprehensive review of thyroid embryology, anatomy, histology, and physiology for surgeons. *International Journal of Surgery*, 52, 39–46.
- [6] French Society of Endocrinology. (2025). *Thyroid nodules*. Retrieved from <https://www.sfendocrino.org/en/item-241-thyroid-nodules> (English version)
- [7] Haugen, B. R., Alexander, E. K., Bible, K. C., et al. (2016). 2015 American Thyroid Association management guidelines for adult patients with thyroid nodules and differentiated thyroid cancer. *Thyroid*, 26(1), 1–133.
- [8] Gharib, H., Papini, E., Garber, J. R., et al. (2016). American Association of Clinical Endocrinologists, American College of Endocrinology, and Associazione Medici Endocrinologi medical guidelines for clinical practice for the diagnosis and management of thyroid nodules. *Endocrine Practice*, 22(5), 622–639.
- [9] Russ, G., Bonnema, S. J., Erdogan, M. F., et al. (2017). European Thyroid Association guidelines for ultrasound malignancy risk stratification of thyroid nodules in adults: The EU-TIRADS. *European Thyroid Journal*, 6(5), 225–237.
- [10] Ross, D. S., et al. (2016). Hyperthyroidism. *Lancet*, 388, 906–918.
- [11] Cooper, D. S. (2005). Thyroid nodules. *New England Journal of Medicine*, 352, 1760–1769.

Bibliographic reference

- [12] Garber, J. R., et al. (2012). Clinical practice guidelines for hypothyroidism in adults. *Thyroid*, 22, 1200–1235.
- [13] Tessler, F. N., Middleton, W. D., Grant, E. G., Hoang, J. K., Berland, L. L., Teefey, S. A., & Cronan, J. J. (2017). ACR thyroid imaging, reporting and data system (TI-RADS): White paper of the ACR TI-RADS committee. *Journal of the American College of Radiology*, 14(5), 587–595. <https://doi.org/10.1016/j.jacr.2017.01.046>
- [14] Cibas, E. S., & Ali, S. Z. (2009). The Bethesda system for reporting thyroid cytopathology. *American Journal of Clinical Pathology*, 132, 658–665.
- [15] Russ, G., Royer, B., Bigorgne, C., Rouxel, A., Bienvenu-Perrard, M., & Leenhardt, L. (2013). Prospective evaluation of thyroid imaging reporting and data system on 4550 nodules with and without elastography. *European Journal of Endocrinology*, 168(5), 649–655. <https://doi.org/10.1530/EJE-12-0936>
- [16] Wells, S. A. Jr., et al. (2015). Medullary thyroid carcinoma: Management guidelines of the American Thyroid Association. *Thyroid*, 25, 567–610.
- [17] Filetti, S., et al. (2019). Thyroid cancer: ESMO clinical practice guidelines. *Annals of Oncology*, 30(12), 1856–1883.
- [18] Dionigi, G., et al. (2017). Complications in thyroid surgery. *Updates in Surgery*, 69(2), 151–158.
- [19] Dralle, H., et al. (2017). Complications of thyroid surgery: The European perspective. *Annals of Surgery*, 265(3), 457–464.
- [20] European Thyroid Association. (2019). *Guidelines for thyroid nodule management*. European Thyroid Journal.
- [21] European Society of Endocrine Surgeons. (2016). *Consensus statement on thyroid surgery*. ESES Publications.
- [22] Park, S. H., Kim, S. J., Kim, E. K., Kim, M. J., Son, E. J., Kwak, J. Y., & Oh, K. K. (2009). Interobserver agreement in assessing the sonographic and elastographic features of malignant thyroid nodules. *AJR American Journal of Roentgenology*, 193(5), W416–W423. <https://doi.org/10.2214/AJR.09.2385>

Bibliographic reference

- [23] Litjens, G., Kooi, T., Bejnordi, B. E., Setio, A. A. A., Ciompi, F., Ghafoorian, M., & Sánchez, C. I. (2017). A survey on deep learning in medical image analysis. *Medical Image Analysis*, 42, 60–88. <https://doi.org/10.1016/j.media.2017.07.005>
- [24] Ginat, D. T., & Schwartz, T. H. (2013). Computer-aided diagnosis in ultrasound imaging of the thyroid. *Ultrasound Clinics*, 8(2), 239–248. <https://doi.org/10.1016/j.cult.2013.01.004>
- [25] Zhang, L., Han, X., & Gu, X. (2020). Deep learning-based computer-aided diagnosis for thyroid nodules: A survey. *Frontiers in Endocrinology*, 11, 600943. <https://doi.org/10.3389/fendo.2020.600943>
- [26] Moon, H. J., Sung, J. Y., Kim, E. K., & Kwak, J. Y. (2015). Thyroid nodule classification in ultrasound: Comparison of CAD and radiologist performance. *Ultrasound in Medicine & Biology*, 41(6), 1501–1509. <https://doi.org/10.1016/j.ultrasmedbio.2015.01.018>
- [27] French National Authority for Health (HAS). (2021). *Workup of thyroid conditions in adults: Relevance and quality criteria for ultrasound*. Retrieved from https://www.has-sante.fr/jcms/p_3281401/en/thyroid-nodules-management
- [28] University of Constantine. (2012). *Physical principles of ultrasound* (Lecture notes). Retrieved from https://fac.umc.edu.dz/vet/telechargement/cours/cours_imagerie_medicale_1a.pdf
- [29] Li, X., Zhang, S., Zhang, Q., et al. (2019). Diagnosis of thyroid cancer using deep convolutional neural network models applied to sonographic images. *The Lancet Oncology*, 20(2), 193–201.
- [30] Jin, Z., Zhu, Y., Zhang, S., Xie, F., et al. (2020). Ultrasound computer aided diagnosis (CAD) based on TI-RADS to distinguish benign from malignant thyroid nodules. *Medical Science Monitor*, 26, e920802.
- [31] Canny, J. (1986). A computational approach to edge detection. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 8(6), 679–698.
- [32] Gopinath, B., & Sivaram, N. M. (2013). Computer-aided diagnosis system for classifying benign and malignant thyroid nodules in multi-stained FNAB cytological images. In *2013 IEEE International Conference on Emerging Trends in Computing, Communication and Nanotechnology (ICECCN)* (pp. 617–622). doi: 10.1109/ICECCN.2013.6628475

Bibliographic reference

- [33] Acharya, U. R., Sree, S. V., Krishnan, M. M. R., Molinari, F., Garberoglio, R., & Suri, J. S. (2016). Computer-aided diagnostic system for classification of thyroid nodules on ultrasound using Gabor transform and machine learning. In *2016 IEEE International Conference on Systems, Man, and Cybernetics (SMC)* (pp. 001234–001239). doi: 10.1109/SMC.2016.7844378
- [34] Zhou, H., Jin, Y., Dai, L., et al. (2020). Differential diagnosis of benign and malignant thyroid nodules using deep learning radiomics of thyroid ultrasound images. *European Journal of Radiology*, 127, 108992. doi: 10.1016/j.ejrad.2020.108992
- [35] Ronneberger, O., Fischer, P., & Brox, T. (2015). U-Net: Convolutional networks for biomedical image segmentation. In **Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015** (pp. 234–241). Springer. doi: 10.1007/978-3-319-24574-4_28
- [36] Otsu, N. (1979). A threshold selection method from gray-level histograms. *IEEE Transactions on Systems, Man, and Cybernetics*, 9(1), 62–66.
- [37] Kass, M., Witkin, A., & Terzopoulos, D. (1987). Snakes: Active contour models. *International Journal of Computer Vision*, 1, 321–331.
- [38] FlyPix AI. (2025). *Deep learning-based image segmentation: A comprehensive guide*. Retrieved from <https://flypix.ai/deep-learning-segmentation/>
- [39] Zhang, X., et al. (2022). Morphological characterization of thyroid nodules in ultrasound images. *Journal of Medical Imaging*, 9(4), 041202.
- [40] Haralick, R. M., Shanmugam, K., & Dinstein, I. (1973). Textural features for image classification. *IEEE Transactions on Systems, Man, and Cybernetics*, SMC-3(6), 610–621.
- [41] Mallat, S. (1999). *A wavelet tour of signal processing*. Academic Press.
- [42] Cortes, C., & Vapnik, V. (1995). Support-vector networks. *Machine Learning*, 20(3), 273–297.
- [43] Breiman, L. (2001). Random forests. *Machine Learning*, 45(1), 5–32.
- [44] LeCun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. *Nature*, 521, 436–444.
- [45] Acharya, S., et al. (2021). Classification of thyroid lesions in ultrasound images using deep learning approaches. *Applied Soft Computing*, 98, 106876.

Bibliographic reference

- [46] Yoo, Y. J., Ha, E. J., Cho, Y. J., et al. (2018). Computer-aided diagnosis of thyroid nodules via ultrasonography: Initial clinical experience. *Korean Journal of Radiology*, 19(4), 665–672. doi: 10.3348/kjr.2018.19.4.665
- [47] Dosovitskiy, A., et al. (2021). An image is worth 16×16 words: Transformers for image recognition at scale. In *International Conference on Learning Representations (ICLR)*. Retrieved from <https://arxiv.org/abs/2010.11929>
- [48] Zhang, H., et al. (2025). TN5000: An ultrasound image dataset for thyroid nodule detection and classification. *Scientific Data*, 12(1), 1437. <https://doi.org/10.1038/s41597-025-05757-4>
- [49] Horvath, E., et al. (2009). An ultrasound reporting system for thyroid nodules. *Radiology*, 252(1), 178–186.
- [50] Bisong, E. (2019). Google Colaboratory. In *Building Machine Learning and Deep Learning Models on Google Cloud Platform* (pp. 59–64). Apress. https://doi.org/10.1007/978-1-4842-4470-8_7
- [51] NVIDIA. (2018). *NVIDIA T4 Tensor Core GPU Architecture*. Retrieved from <https://www.nvidia.com/en-us/data-center/tesla-t4/>
- [52] Jenny, T. J. (2023). *TN3K: Thyroid nodule region segmentation dataset*. Kaggle. Retrieved from <https://www.kaggle.com/datasets/tjahan/tn3k-thyroid-nodule-region-segmentation-dataset>
- [53] ElAfifi, A. (2025). *TN5000 thyroid nodule dataset*. Kaggle. Retrieved from <https://www.kaggle.com/datasets/abdullahelafifi/main-data>
- [54] Journal des Femmes Santé. (2025). *Thyroid gland: role, anatomy, location, conditions*. Retrieved from <https://sante.journaldesfemmes.fr/en/thyroid-gland>
- [55] Vasković, J. (2023). *Superior thyroid artery*. Kenhub. Retrieved from <https://www.kenhub.com/en/library/anatomy/superior-thyroid-artery>
- [56] Anatomy.app. (n.d.). *Arterial blood supply of thyroid gland*. Retrieved from <https://anatomy.app/media/arterial-blood-supply-of-thyroid-gland-4-10446>
- [57] Achouri, & Adili. (2018). *Study of thyroid nodules* (Master's thesis). University of 8 May 1945 Guelma. Retrieved from <https://dspace.univ-guelma.dz/jspui/handle/123456789/1047>

Bibliographic reference

- [58] Bessaguet, C., Suteau, V., & Desmoulière, A. (2023). The thyroid gland. *Actualités Pharmaceutiques*. Elsevier Masson SAS. Retrieved from <https://www.em-consulte.com/article/1636546/the-thyroid-gland>
- [59] Chenafa, S. (2026). *Vascularization of the thyroid* (Lecture notes). Faculty of Medicine, University of Oran. Retrieved from https://www.facmed-univ-oran.dz/ressources/fichiers_produits/fileer_produit_2213.pdf
- [60] Médecine Intégrée. (n.d.). *Lymphatic drainage of the thyroid*. Retrieved from <https://www.medecineintegree.fr>
- [61] Balice, G., & Theintz, G. (2007). Goiters and thyroid nodules in children and adolescents. *Revue Médicale Suisse*. Retrieved from <https://www.revmed.ch/revue-medicale-suisse-107/goitres-et-nodules-thyroidiens-chez-l-enfant-et-l-adolescent>
- [62] Marcy, P.-Y., Sanglier, T., Bizeau, A., & Ghanassia, E. (2024). Ultrasound of differentiated thyroid carcinomas. *ScienceDirect*. [https://doi.org/10.1016/S2543-3431\(24\)00004-6](https://doi.org/10.1016/S2543-3431(24)00004-6)
- [63] Traynard, P.-Y. (2010). *Signs of hyperthyroidism*. Docteurcllic. Retrieved from <https://www.docteurcllic.com/en/hyperthyroidism.aspx>
- [64] Physiopharmlab. (2021). *Hypothyroidism – Clinical signs*. Retrieved from <https://physiopharmlab.com/hypothyroidism/>
- [65] Journal des Femmes Santé. (2025). *Autoimmune thyroiditis*. Retrieved from <https://sante.journaldesfemmes.fr/en/autoimmune-thyroiditis>
- [66] Journal des Femmes Santé. (2026). *Palpation of the thyroid region*. Retrieved from <https://sante.journaldesfemmes.fr/en/thyroid-palpation>
- [67] Carayon, P. (2009). Biological workup of the thyroid. *Annales d'Endocrinologie*, 70(2), 111–121. [https://doi.org/10.1016/S0003-4266\(09\)72563-8](https://doi.org/10.1016/S0003-4266(09)72563-8)
- [68] iSonic – SonoSchool. (n.d.). *Thyroid ultrasound*. Retrieved from <https://sonoschool.com>
- [69] [Laradioactivite.com](https://www.laradioactivite.com). (n.d.). *Thyroid scintigraphy*. Retrieved from <https://www.laradioactivite.com/en/thyroid-scintigraphy>
- [70] Le Monde – Réalités Biomédicales. (2024). *Thyroid MRI*. Retrieved from <https://www.lemonde.fr/en/thyroid-mri>

Bibliographic reference

- [71] Radiologie Croisette. (n.d.). *Thyroid elastography*. Retrieved from <https://www.radiologiecroisette.fr/en/thyroid-elastography>
- [72] ORL Pôle Azur Chirurgie – ORPAC Grasse. (n.d.). *Thyroid CT scan*. Retrieved from <https://www.orpacgrasse.fr/en/thyroid-ct>
- [73] Agus, B., Nomine-Criqui, C., Demarquet, L., Klein, M., & Brunaud, L. (2018). Surgical technique: Transaxillary robotic thyroid lobectomy. *Journal of Visceral Surgery*. <https://doi.org/10.1016/j.jviscsurg.2018.06.003>
- [74] Chirurgie Thyroïde Paris. (n.d.). *Lymph node dissection in thyroid surgery*. Retrieved from <https://chirurgiethyroideparis.fr/en/lymph-node-dissection>
- [75] Sun, J., Yu, J., & Xu, S. (2020). Computer-aided diagnosis system for thyroid nodules based on ultrasound images. *Biomedical Signal Processing and Control*. <https://doi.org/10.1016/j.bspc.2020.102314>
- [76] Abdolali, F., Kapur, J., Jaremko, J., & Punithakumar, K. (2020). Automated thyroid nodule detection and segmentation in ultrasound images. *arXiv*. Retrieved from <https://arxiv.org/abs/2006.05861>
- [77] Alamy. (n.d.). *Thyroid fine needle aspiration* [Image]. Retrieved from <https://www.alamy.com/stock-photo/thyroid-fine-needle-aspiration.html>
- [78] Ameli. (2025). *Thyroid nodule: workup*. French Health Insurance. Retrieved from <https://www.ameli.fr/en/thyroid-nodule>
- [79] Santé sur le Net. (n.d.). *Thyroid nodule*. Retrieved from <https://www.sante-sur-le-net.com/en/thyroid-nodule>
- [80] Depositphotos. (n.d.). *Total thyroidectomy* [Image]. Retrieved from <https://depositphotos.com/stock-photos/total-thyroidectomy.html>
- [81] Faculty of Medicine, University of Oran. (n.d.). *Anatomy of the laryngeal nerve* (Lecture notes). Retrieved from <https://www.facmed-univ-oran.dz>
- [82] Ait Laabdia, R., Lakhdar, Y., Elbouderkou, M., et al. (2020). Thyroid hematoma secondary to closed cervical trauma. *PAMJ Clinical Medicine*, 2, 131. <https://doi.org/10.11604/pamj-cm.2020.2.131.21204>

Bibliographic reference

- [83] Zuiderveld, K. (1994). Contrast Limited Adaptive Histogram Equalization. In *Graphics Gems IV* (pp. 474–485). Academic Press.
- [84] Deng, J., Dong, W., Socher, R., Li, L.-J., Li, K., & Fei-Fei, L. (2009). ImageNet: A large-scale hierarchical image database. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. Retrieved from https://www.image-net.org/static_files/papers/imagenet_cvpr09.pdf
- [85] Tan, M., & Le, Q. V. (2019). EfficientNet: Rethinking model scaling for convolutional neural networks. In *International Conference on Machine Learning (ICML)*. Retrieved from <https://arxiv.org/abs/1905.11946>
- [86] Huang, G., Liu, Z., Van Der Maaten, L., & Weinberger, K. Q. (2017). Densely connected convolutional networks. In *CVPR*. Retrieved from <https://arxiv.org/abs/1608.06993>
- [87] He, K., Zhang, X., Ren, S., & Sun, J. (2016). Deep residual learning for image recognition. In *CVPR*. Retrieved from <https://arxiv.org/abs/1512.03385>
- [88] Chen, T., & Guestrin, C. (2016). XGBoost: A scalable tree boosting system. In *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*. Retrieved from <https://arxiv.org/abs/1603.02754>
- [89] Oktay, O., Schlemper, J., Le Folgoc, L., et al. (2018). Attention U-Net: Learning Where to Look for the Pancreas. *arXiv preprint arXiv:1804.03999*.
- [90] Zhang, Z., Liu, Q., & Wang, Y. (2018). Road Extraction by Deep Residual U-Net. *IEEE Geoscience and Remote Sensing Letters*, 15(5), 749–753.
- [91] Cao, H., Wang, Y., Chen, J., Jiang, D., Zhang, X., Tian, Q., & Wang, M. (2021). Swin-Unet: Unet-like Pure Transformer for Medical Image Segmentation. *arXiv preprint arXiv:2105.05537*.
- [92] Liu, Z., Lin, Y., Cao, Y., et al. (2021). Swin Transformer: Hierarchical Vision Transformer using Shifted Windows. In *Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV)*, pp. 10012–10022.
- [93] Zhou, Z., Siddiquee, M. M. R., Tajbakhsh, N., & Liang, J. (2018). UNet++: A Nested U-Net Architecture for Medical Image Segmentation. *arXiv preprint*.

Appendix

Appendix A: Examples of annotated images

This appendix presents several examples of ultrasound images from the TN3K and TN5000 databases used in this study. The figures illustrate:

- the original ultrasound images;
- the annotations made by the medical experts;
- the segmentation masks (Ground Truth);
- as well as the examples of predictions generated by the proposed model.

These illustrations help visualize the variability of thyroid nodules in terms of size, shape, contrast and location.



Figure A.1: Example of an annotated ultrasound image from TN3K.[52]

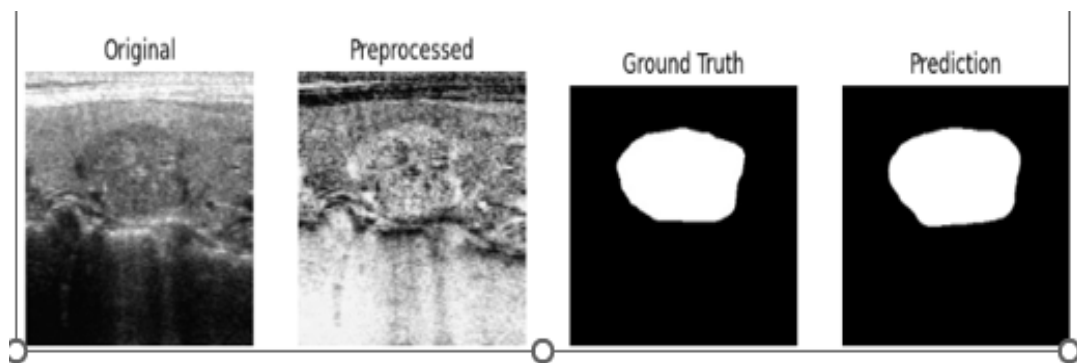


Figure A.2: Comparison between the actual mask and the predicted segmentation.

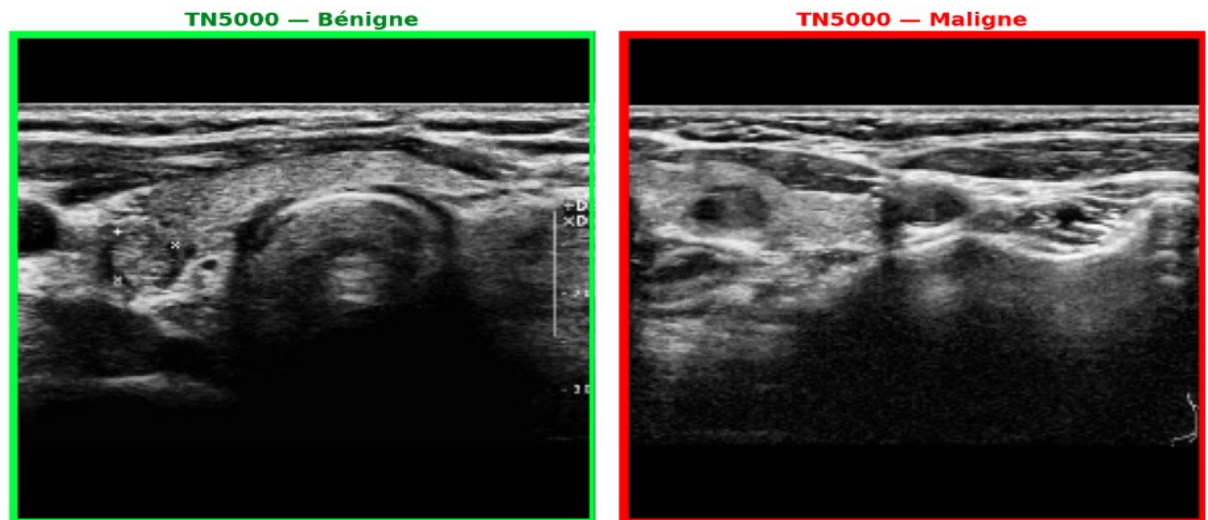


Figure A.3: Examples of benign and malignant nodules from TN5000.

Appendix B: Code Snippets :

This appendix gathers the main code extracts used for the implementation of segmentation and classification models. The scripts were developed in Python using the PyTorch, `segmentation_models_pytorch`, `scikit-learn` and `timm` libraries.

The excerpts presented mainly concern:

- the loading and pre-processing of data;
- the definition of network architectures;
- the loss functions;
- training procedures;
- the evaluation steps;
- as well as the calculation of performance metrics.

➤ B.1 Loading and pre-processing of images

```
transform = transforms.Compose([  
    transforms.Resize((224,224)),  
    transforms.RandomHorizontalFlip(),  
    transforms.ToTensor(),  
])
```

➤ B.2 U-Net++ Architecture with EfficientNet-B3

```
model = smp.UnetPlusPlus(
    encoder_name="efficientnet-b3",
    encoder_weights="imagenet",
    classes=1,
    activation=None
)
```

➤ B.3 Combined loss function

```
loss = bce_loss + dice_loss
```

➤ B.4 Example of training

```
optimizer = torch.optim.AdamW(model.parameters(), lr=1e-4)
```

Appendix C: Detailed tables of results

Model	Accuracy	Precision	Recall	F1-score	AUC
DenseNet121	97.03%	97.18%	98.61%	97.89%	99.05%
EfficientNetB0	96.56%	96.62%	98.53%	97.57%	98.59%
ResNet ResNet18	95.85%	97.35%	96.70%	97.02%	98.72%
ViT (Vision Transformer)	93.08%	93.93%	96.34%	95.12%	97.09%
XGBoost	94.56%	93.75%	98.83%	96.22%	97.58%
XGBoost					
Random Forest	94.36%	92.77%	99.71%	96.12%	98.78%
Support Vector Machine SVM	88%	87.18%	97.14%	91.89%	94.08%

Table C.1: Detailed performance results of classification models evaluated on the basis of TN5000

Table C.1 presents the detailed quantitative performance of the different deep learning and machine learning models evaluated in this study for the classification of thyroid nodules. The results show that the DenseNet121 model has the best overall performance in terms of

Accuracy, F1-score and AUC, confirming its effectiveness for discrimination between benign and malignant nodules.

segmentation model	Dice	IoU	Accuracy	Precision	Recall	Specificity
U-Net++ + EfficientNet-B3	85.44 %	75.45 %	96.00 %	85.00 %	86.00 %	98.90 %

Table C.2: Detailed performance results of the segmentation model evaluated on the basis of TN3K

Table C.2 presents the quantitative performance obtained by the proposed segmentation model based on TN3K data. The results show good segmentation quality of thyroid nodules, with a high Dice coefficient and high specificity, reflecting the model's ability to distinguish nodular regions from surrounding tissue while limiting false positives.