Machine Learning Applications in Medical Diagnosis, case study bone metastasis

Marwa Afnouch

Thèse de doctorat

Pour obtenir le grade de Docteur de

l'UNIVERSITÉ POLYTECHNIQUE HAUTS-DE-FRANCE

et de l'INSA HAUTS-DE-FRANCE

et de l'UNIVERSITE DE SFAX-TUNISIE

et de l'ECOLE NATIONALE D'INGENIEURS DE SFAX-TUNISIE

Discipline, spécialité selon la liste des spécialités pour lesquelles l'École Doctorale est accréditée :
Sciences et technologies de l'information et de la communication

Présentée et soutenue par Marwa AFNOUCH.

Le 18/12/2023, à Valenciennes

École doctorale :
École Doctorale Polytechnique Hauts-de-France (ED PHF n°635)

Unité de recherche :
Institut d'Electronique de Microélectronique et de Nanotechnologie – Site de Valenciennes (IEMN – UMR CNRS 8520)
Laboratoire de Recherche « Computer & Embedded Systems » Ces_Lab

Application de Machine Learning dans le domaine de diagnostic médical, cas d'utilisation : métastases osseuses

JURY

Président du jury
- Chausse, Fréderic. Professeur. Université Clermont Auvergne.

Rapporteurs
- Chausse, Fréderic. Professeur. Université Clermont Auvergne.
- Khalifa, Nawres. Professeur. Université de Tunis El Manar.

Examinateurs
- Boonaert, Jacques. HdR. IMT Nord Europe.
- Sellami, Dorra. Professeur. ENIS, Université de Sfax, Tunisie.

Co-directeur de thèse : TALEB-AHMED, Abdelmalik. Professeur. Université polytechnique Hauts-de-France.
Co-directeur de thèse : ABID, Mohamed. Professeur. ENIS, Université de Sfax, Tunisie.

Co-encadrant : Gaddour, Olfa. Docteur. ENIS, Université de Sfax, Tunisie.

INSA INSTITUT NATIONAL DES SCIENCES APPLIQUÉES HAUTS-DE-FRANCE

Université Polytechnique HAUTS-DE-FRANCE

PhD Thesis

Submitted for the degree of Doctor of Philosophy from

UNIVERSITÉ POLYTECHNIQUE HAUTS-DE-FRANCE

And INSA HAUTS-DE-FRANCE

And University of Sfax

And National Engineering School of Sfax

Subject:

Information and communication sciences and technologies

Presented and defended by Marwa AFNOUCH.

On 18/12/2023, Valenciennes

Doctoral school:

Doctoral School Polytechnique Hauts-de-France (ED PHF n°635)

Research unit:

Institute of Electronics Microelectronics and Nanotechnology – Valenciennes site (IEMN – UMR CNRS 8520)
Research Laboratory Computers & Embedded Systems, Tunisia (Ces_Lab)

Machine Learning Application in Medical Diagnosis: Case Study: Bone Metastases

JURY

President of jury

- Chausse, Fréderic. Professor. University of Clermont Auvergne, France

Reviewers

- Chausse, Fréderic. Professor. University of Clermont Auvergne, France
- Khalifa, Nawres. Professor. University of Tunis El Manar, Tunisia

Examiners

Boonaert, Jacques. HdR. IMT Nord Europe.

- Sellami, Dorra. Professor. University of Sfax, Tunisia.

Thesis co-director: TALEB-AHMED, Abdelmalik. Professor. Université polytechnique Hauts-de-France.

Thesis co-director: ABID, Mohamed. Professor. ENIS, Université of Sfax, Tunisia.

Co-supervisor: Gaddour, Olfa. PhD. ENIS, Université of Sfax, Tunisia.

Résumé substantiel en langue française

Introduction

La technologie et l'innovation humaine ne cessent de surprendre et d'étonner. La médecine a toujours été l'un des domaines les plus dynamiques de l'innovation scientifique et l'un de ceux où les avantages de cette innovation sont le plus clairement visibles. Il y a cent ans, consulter un radiologue pour diagnostiquer une métastase osseuse aurait probablement entraîné une intervention chirurgicale extrêmement défigurant et un taux de survie à cinq ans de seulement 40%. Aujourd'hui, les taux de survie à cinq ans dépassent les 90% et une bien meilleure compréhension de la maladie permet d'adapter le traitement à chaque cas. Ce rythme de changement et de développement se poursuit au XXIe siècle. La période actuelle a été qualifiée de quatrième révolution industrielle, car des technologies plus avancées et l'utilisation de grandes quantités de données transforment la fabrication et les processus industriels. En médecine, l'intelligence artificielle (IA) et les techniques basées sur les données promettent de changer fondamentalement la façon dont la médecine clinique est menée. Il s'agit d'un domaine de recherche naissant, volatile et fluide, un creuset de différentes professions, avec de nombreuses questions sans réponse. Ma motivation pour entreprendre ce doctorat était d'apporter mon expertise scientifique et mon enthousiasme à ce domaine, de faire partie de cette nouvelle révolution et d'y contribuer.

Description du Sujet

Les métastases sont un groupe de cellules anormales qui se développent en dehors de leurs limites normales et se propagent à d'autres organes. En particulier, les métastases osseuses sont des cancers qui prennent naissance dans les organes du corps, tels que le sein, le poumon ou la prostate, et qui se propagent dans les os. Dans 90% des décès dus au cancer, les métastases ont été considérées comme un facteur contributif. Les os occupent la troisième place, derrière le foie et les poumons, parmi les sites de métastases les plus courants [16]. Les statistiques montrent que plus de 70% des patients atteints d'un cancer du sein ou de la prostate présentent des métastases osseuses [21]. À cet égard, la détection précoce des cancers des os est essentielle pour prendre la bonne décision [6]. Le suivi des métastases osseuses nécessite l'utilisation de

l'imagerie médicale pour diagnostiquer les maladies et déterminer l'efficacité des traitements. Plusieurs modalités d'imagerie médicale ont été utilisées, notamment la tomodensitométrie (CT), l'IRM (imagerie par résonance magnétique), la scintigraphie osseuse et la TEP, chaque modalité d'imagerie ayant ses avantages et ses inconvénients [27]. En particulier, les tomodensitogrammes ont une résolution spatiale élevée, ce qui permet de détecter les grandes lésions osseuses ainsi que les lésions légères et de petite taille [21]. Ils offrent également une évaluation simultanée des os primaires et des os métastatiques. En outre, les tomodensitogrammes sont couramment disponibles à un coût relativement faible [9]. En clinique, la tomodensitométrie est la modalité d'imagerie la plus couramment utilisée pour les tests de starification de base et la surveillance en série des patients atteints de cancer [11]. Cependant, la détection des métastases osseuses à partir des tomodensitogrammes est à la fois difficile et longue pour les raisons suivantes : (i) les os étant présents dans tout le corps, les radiologues doivent examiner toutes les coupes [19] ; (ii) l'aspect radiologique de la MO dépend des types de lésions [11]. Par conséquent, aucun paramètre de fenêtre unique ne peut représenter correctement toutes les métastases osseuses ; et (iii) les imitations bénignes telles que les fractures, les îlots osseux et les changements dégénératifs peuvent être confondus avec des lésions de la moelle osseuse, ce qui complique le diagnostic [28]. Le diagnostic assisté par ordinateur (CAO) peut aider les radiologues à trouver de petites lésions qui pourraient autrement passer inaperçues [12]. Par conséquent, il existe une forte demande de systèmes de diagnostic assisté par ordinateur pour les métastases osseuses sur les tomodensitogrammes. Les méthodes d'apprentissage profond ont récemment fait l'objet d'une grande attention pour résoudre plusieurs tâches liées à la vision [15, 4]. En particulier, plusieurs solutions sont proposées pour résoudre le problème de la détection des métastases osseuses à l'aide de l'apprentissage automatique. Ce travail vise à créer un ensemble de données de référence pour la segmentation et la classification des os à partir de tomodensitogrammes. En outre, nous proposons un nouveau système CAD pour la segmentation et la classification des métastases osseuses afin d'améliorer les performances dans le domaine de la détection des métastases osseuses.

Objectifs

Les principaux objectifs de ce projet de doctorat sont les suivants :

- Examen des techniques de pointe actuelles en matière de diagnostic médical et des défis existants en matière de diagnostic des métastases osseuses.

- Exploration de diverses méthodologies d'apprentissage automatique, y compris l'apprentissage profond et les méthodes d'ensemble, afin d'identifier leurs applications potentielles dans le diagnostic des métastases osseuses.
- Collecte de nouveaux ensembles de données sur les métastases osseuses afin de faciliter la segmentation et l'identification des métastases osseuses. Cet ensemble de données vise à aider la communauté des chercheurs à améliorer leurs évaluations des métastases osseuses.
- Proposition d'une approche innovante d'apprentissage profond adaptée pour optimiser la segmentation des métastases osseuses.
- Exploration de l'utilisation de méthodologies avancées d'apprentissage profond pour améliorer la précision et l'efficacité de la segmentation et la classification des métastases osseuses.
- Création d'un cadre d'évaluation standardisé pour la classification automatique des métastases osseuses et déploiement d'une approche innovante pour classifier les lésions métastatiques.
- Évaluation des modèles proposés en comparaison avec les approches existantes et étude de leur capacité à fournir un aperçu des caractéristiques et modèles sous-jacents qui contribuent au diagnostic des métastases osseuses.
- Discussion des implications cliniques et de l'applicabilité des modèles proposés dans le monde réel, en tenant compte des défis et des opportunités potentielles pour le développement de nouveaux modèles.

Contributions

Pour répondre à ces objectifs, les principales contributions de ce travail sont résumées comme suit :

- Nous présentons tout d'abord une revue des techniques utilisées pour classer et segmenter les métastases osseuses. Dans cette revue, nous donnons un aperçu des métastases osseuses, y compris l'importance de l'oncologie médicale et des modalités d'imagerie médicale. Nous présentons ensuite l'état de l'art des méthodes actuelles de classification et de segmentation des métastases osseuses. Nous nous concentrons sur la comparaison de ces méthodes en termes de divers paramètres tels que l'évaluation des performances et les ensembles de données.

- Nous proposons un ensemble de données de référence pour l'analyse des métastases osseuses. Cet ensemble de données a été créé à partir de tomodensitogrammes de plus de 57 patients de l'hôpital Hedi Chaker de Sfax, en Tunisie, avec plus de 102 614 images. À notre connaissance, il s'agit du premier ensemble de données accessibles au public dans le domaine du diagnostic des métastases osseuses qui inclut la tomodensitométrie comme modalité d'imagerie. Ainsi, des avancées potentielles dans le domaine des métastases osseuses pourraient être réalisées grâce à notre base de données.

- Nous proposons une nouvelle approche de segmentation pour détecter les métastases osseuses. C'est une architecture hybride AttUnet++ qui contient des décodeurs doubles et la methode de fusion. Nous évaluons notre technique avec la nouvelle base de données proposée. Les résultats obtenus montrent une meilleure performance dans la segmentation des lésions par rapport aux techniques traditionnelles.

- La dernière contribution comprend une comparaison complète des principales architectures CNN telles que InceptionV3, EfficientNet, ResNext50 et DenseNet161 avec des architectures de transformers, à savoir ViT et DeiT pour la classification des métastases osseuses. Ceci s'ajoute à l'introduction d'une approche innovante afin de fournir un cadre d'évaluation standardisé pour la classification automatique des métastases osseuses.

- Pour assurer la reproductibilité des résultats, nous avons mis notre ensemble de données et nos codes à la disposition du public à l'adresse suivante GitHub - Marwa-Afnouch/EH-AttUnetplus

Etat de l'art dans le domaine

Ces dernières années, l'évaluation automatique de la moelle osseuse à partir de différentes modalités d'imagerie médicale telles que les scintigraphies osseuses, les tomodensitométries et les tomographies par émission monophotonique (SPECT) a suscité un intérêt considérable. Les travaux les plus récents se concentrent sur deux tâches principales : (i) la classification des scanners de la moelle osseuse en cas normaux et anormaux et (ii) la segmentation des lésions. Récemment, plusieurs travaux de recherche ont traité de la classification des BM en utilisant des méthodes d'apprentissage profond [24, 2, 17, 25, 13, 8, 1]. D'autre part, la segmentation automatisée des lésions métastatiques n'en est qu'à ses débuts. [26] a proposé un système d'interprétation d'images pour la segmentation du squelette et l'extraction des points chauds d'un os métastatique à partir d'un scintigramme osseux du corps entier basé sur l'apprentissage

profond. [29] a développé un algorithme de segmentation basé sur UNet avec un mécanisme d'attention pour la segmentation des os SPECT, qui peut automatiquement identifier l'emplacement de la BM. [18] a comparé différentes approches des os et des lésions métastatiques osseuses dans la segmentation du cancer du sein. Deux méthodes d'apprentissage profond basées sur UNet ont été développées et entraînées pour segmenter soit les os et les lésions osseuses, soit les lésions osseuses seules sur les images PET/CT. Plus tard, [7] a proposé un algorithme d'apprentissage profond pour les métastases spinales du cancer du poumon. Ils ont proposé un modèle de réseau en U convolutionnel dilaté (DC-U-Net) pour segmenter les images CT énergétiques/spectrales. Plus récemment, [20] a utilisé des images de tomodensitométrie et a proposé une approche de segmentation basée sur trois réseaux neuronaux convolutionnels (CNN) : un réseau 2D basé sur UNet pour segmenter les os, un réseau 3D basé sur UNet pour segmenter les régions candidates, et un réseau 3D basé sur ResNet pour réduire les résultats faussement positifs. Bien que les méthodes de pointe aient obtenu des résultats prometteurs, leur impact est limité par l'utilisation d'ensembles de données privés, ce qui empêche la communauté d'exploiter les données et de s'appuyer sur leurs résultats. En outre, les travaux existants ont utilisé plusieurs modalités d'imagerie, notamment l'IRM, la scintigraphie osseuse et la TEP, tandis que peu de travaux ont utilisé la tomodensitométrie malgré ses avantages distincts. Le tableau 1 présente une comparaison des ensembles de données BM existants. D'après le tableau 1, la principale limitation des ensembles de données existants est leur petite taille, qui peut conduire à un sur ajustement. Pour surmonter ces problèmes, nous avons créé un nouvel ensemble de données de segmentation BM et l'avons mis à la disposition de la communauté des chercheurs.

Tableau 1 : Comparaison des ensembles de données BM existants

Literature	Data Modality	Dataset description	Task	Availability
[24]	Bone scan	586 scans[368 BM;18 Normal] 507 patients	Classification	Private
[22]	Bone scan	778 scans [328 BM; 271 degenerative; 179 normal] 817 patients	Classification	Private
[23]	Bone scan	408 scans[221 BM; 187 Normal; 382] patient	Classification	Private
[25]	Bone scan	15,474 images 13.811 patients [9595 benign; 5879 malignant]	Classification	Private
[1]	Bone scan	1088 labled chest ; 18.560 unlabeled 9.280 patients	Segmentation Classification	Private
[26]	Bone scan	246 bone scans	segmentation	Private
[5]	Bone scan	205 scans[100 BM; 105 Normal]	Classification	Private
[2]	Bone scan	130 images [30 Normal;100 BM] 60 patients	Classification	Private
[10]	Bone scan	9133 images[2991 BM; 6142 Normal] 5342 patients	Classification	Private
[17]	CT	2.880 scans 114 patients [41 BM]	Classification	Private
[20]	CT	269 BM scans 169 patients	Segmentation	Private
[18]	PET/CT	24 patients	Segmentation	Private
[3]	SPECT	260 scans 130 patients	Segmentation	Private
[14]	SPECT	76 patients 152 scans	Classification	Private
The proposed dataset	**CT**	**23 patients** **1517 images**	**Segmentation**	**Public**

References

[1] Terapap Apiparakoon et al. “MaligNet: semisupervised learning for bone lesion instance segmentation using bone scintigraphy”. In: Ieee Access 8 (2020), pp. 27047–27066.

[2] Ali Aslantas et al. “CADBOSS: A computer-aided diagnosis system for whole-body bone scintigraphy scans”. In: Journal of Cancer Research and Therapeutics 12.2 (2016),p.787.

[3] Yongchun Cao et al. “Segmentation of lung cancer-caused metastatic lesions in bone scan images using self-defined model with deep supervision”. In: Biomedical Signal Processing and Control 79 (June 2023), p. 104068.

[4] Shouvik Chakraborty and Kalyani Mali. “SuFMoFPA: A superpixel and meta-heuristic based fuzzy image segmentation approach to explicate COVID-19 radiological images”. In: Expert Systems with Applications 167 (2021), p. 114142.

[5] Da-Chuan Cheng et al. “Bone Metastasis Detection in the Chest and Pelvis from a Whole-Body Bone Scan Using Deep Learning and a Small Dataset”. In: Electronics 10.10(May2021),p.1201.

[6] Robert E Coleman, Janet Brown, and Ingunn Holen. “Bone metastases”. In: Abeloff’s Clinical Oncology (2020), pp. 809–830.

[7] Xiaojie Fan et al. “Deep learning-based identification of spinal Metastasis in lung cancer using spectral CT images”. In: Scientific Programming 2021 (2021).

[8] Yanru Guo et al. “Automated detection of lung cancer-caused metastasis by classifying scintigraphic images using convolutional neural network with residual connection and hybrid attention mechanism”. In: Insights into Imaging 13.1 (2022), pp. 1–13.

[9] Matthias Hammon et al. “Automatic detection of lytic and blastic thoracolumbar spine metastases on computed tomography”. In: European Radiology 23.7 (Feb. 2013),

pp. 1862–1870.

[10] Sangwon Han, Jungsu S. Oh, and Jong Jin Lee. “Diagnostic performance of deep learning models for detecting bone metastasis on whole-body bone scan in prostate cancer”. In: European Journal of Nuclear Medicine and Molecular Imaging 49.2 (Aug. 2021), pp. 585–595.

[11] Walter Heindel et al. “The diagnostic imaging of bone metastases”. In: Deutsches Ärzteblatt International 111.44 (2014), p. 741.

[12] Shruti Jadon. “A survey of loss functions for semantic segmentation”. In: 2020 IEEE Conference on Computational Intelligence in Bioinformatics and Computational Biology (CIBCB). IEEE. 2020, pp. 1–7.

[13] Chen Li et al. “Attention unet++: A nested attention-aware u-net for liver ct image segmentation”. In: 2020 IEEE International Conference on Image Processing (ICIP).

IEEE. 2020, pp. 345–349.

[14] Qiang Lin et al. “Deep learning based automatic segmentation of metastasis hotspots in thorax bone SPECT images”. In: PLOS ONE 15.12 (Dec. 2020).

[15] Geert Litjens et al. “A survey on deep learning in medical image analysis”. In: Medical

image analysis 42 (2017), pp. 60–88.

[16] Filipa Macedo et al. “Bone metastases: an overview”. In: Oncology reviews 11.1 (2017).

[17] Samira Masoudi et al. “Deep learning based staging of bone lesions from computed tomography scans”. In: IEEE Access 9 (2021), pp. 87531–87542.

[18] Noémie Moreau et al. “Deep learning approaches for bone and bone lesion segmentation on 18FDG PET/CT imaging in the context of metastatic breast cancer”. In: 2020 42nd Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC). IEEE. 2020, pp. 1532–1535.

[19] Shunjiro Noguchi et al. “Bone segmentation on whole-body CT using convolutional neural network with novel data augmentation techniques”. In: Computers in Biology and Medicine 121 (June 2020), p. 103767.

[20] Shunjiro Noguchi et al. “Deep learning–based algorithm improved radiologists’ performance in bone metastases detection on CT”. In: European Radiology (2022), pp. 1– 12.

[21] Gerard J O’Sullivan. “Imaging of bone metastasis: An update”. In: World Journal of Radiology 7.8 (2015), p. 202.

[22] Nikolaos Papandrianos, Elpiniki I. Papageorgiou, and Athanasios Anagnostis. “Development of Convolutional Neural Networks to identify bone metastasis for prostate cancer patients in bone scintigraphy”. In: Annals of Nuclear Medicine 34.11 (Aug. 2020),

pp. 824–832.

[23] Nikolaos Papandrianos et al. “A Deep-Learning Approach for Diagnosis of Metastatic Breast Cancer in Bones from Whole-Body Scans”. In: Applied Sciences 10.3 (Feb. 2020), p. 997.

[24] Nikolaos Papandrianos et al. “Bone metastasis classification using whole body images from prostate cancer patients based on convolutional neural networks application”. In: PloS one 15.8 (2020), e0237213.

[25] Yong Pi et al. “Automated diagnosis of bone metastasis based on multi-view bone scans using attention-augmented deep neural networks”. In: Medical Image Analysis 65 (Oct. 2020), p. 101784.

[26] Akinobu Shimizu et al. “Automated measurement of bone scan index from a wholebody bone scintigram”. In: International Journal of Computer Assisted Radiology and Surgery 15.3 (2020), pp. 389–400.

[27] Anthony Turpin et al. “Imaging for metastasis in prostate cancer: a review of the literature”. In: Frontiers in Oncology 10 (2020), p. 55.

[28] Robert M. Vandemark, Elizabeth J. Shpall, and Mary Lou Affronti. “Bone Metastases from Breast Cancer”. In: Journal of Computer Assisted Tomography 16.4 (July 1992), pp. 608–614.

[29] Jingyi Zhang et al. “Bone metastasis segmentation based on Improved U-NET algorithm”. In: Journal of Physics: Conference Series. Vol. 1848. 1. IOP Publishing. 2021, p. 012027.

ABSTRACT

Metastases are a group of abnormal cells that develop outside the original organ boundaries and spread to other organs. Bone metastases in particular originate in one organ of the body, e.g. the breast, lung or prostate, and spread to the bone. Although this disease was discovered more than a century ago, it is still not well defined and existing treatments are only weakly effective, possibly because it is difficult and time-consuming to detect. To help physicians, new deep learning technologies promise to improve overall accuracy. This dissertation aims to assist radiologists in the routine detection of bone metastases using deep learning algorithms. The discovery of methodological biases in studies on the diagnosis of bone metastases and the lack of consensus on the interpretability of deep learning has shifted the focus of this dissertation. It now focuses primarily on data collection and overcoming the challenges of validation and interpretability of deep learning. To properly assess the ability of deep learning to detect bone metastases, three main contributions were made. First, we collected the BM-Seg dataset, the first benchmark dataset for the segmentation and classification of bone metastases using CT-scans images with 102,614 images from 57 patients. This novel open-source dataset was used to improve the reproducibility of deep learning experiments. Second, we propose one of these deep learning experiments, a novel segmentation approach based on the attUnet++ architecture with dual decoders for bone lesion localization called Hybrid-AttUnet++. In addition, we use an ensemble of trained hybrid attUnet++ models, which we call EH-AttUnet++, to optimize the segmentation performance.

Finally, we compared several pre-trained CNN models with pre-trained Transformers models for the classification of bone metastases. The last study highlights the lack of robustness of classification using deep learning methods and proposes a method to improve accuracy based on both CNN and Transformer approaches.

The results of the various preliminary studies are encouraging and promising.

Keywords: Deep Learning, Bone metastasis, Segmentation, Classification, Medical Imaging, Dataset, Computer Tomography, CNN, Transformer, Unet.

RESUMÉ

Les métastases sont un groupe de cellules anormales qui se développent en dehors des limites de l'organe d'origine et se propagent à d'autres organes. Les métastases osseuses, en particulier, proviennent d'un organe du corps, par exemple le sein, le poumon ou la prostate, et se propagent à l'os. Bien que cette maladie ait été découverte il y a plus d'un siècle, elle n'est toujours pas bien définie et les traitements existants ne sont que faiblement efficaces, probablement parce qu'elle est difficile et longue à détecter. Les nouvelles technologies d'apprentissage profond promettent d'améliorer la précision globale pour aider les médecins. Cette thèse vise à aider les radiologues dans la détection de routine des métastases osseuses à l'aide d'algorithmes d'apprentissage profond. La découverte de biais méthodologiques dans les études sur le diagnostic des métastases osseuses et l'absence de consensus sur l'interprétabilité de l'apprentissage profond ont déplacé l'objectif de cette thèse. Elle se concentre désormais principalement sur la collecte de données et sur la résolution des problèmes de validation et d'interprétabilité de l'apprentissage profond. Pour évaluer correctement la capacité de l'apprentissage profond à détecter les métastases osseuses, trois contributions principales ont été apportées. Premièrement, nous avons collecté l'ensemble de données BM-Seg, le premier ensemble de données de référence pour la segmentation et la classification des métastases osseuses à l'aide d'images de tomodensitométrie, avec 102 614 images provenant de 57 patients. Ce nouvel ensemble de données open-source a été utilisé pour améliorer la reproductibilité des expériences d'apprentissage profond.

Ensuite, nous proposons l'une de ces expériences d'apprentissage profond, une nouvelle approche de segmentation basée sur l'architecture attUnet++ avec des décodeurs doubles pour la localisation des lésions osseuses, appelée Hybrid-AttUnet++. En outre, nous utilisons un ensemble de modèles attUnet++ hybrides entraînés, que nous appelons EH-AttUnet++, pour optimiser les performances de segmentation.

Enfin, nous avons comparé plusieurs modèles CNN pré-entraînés avec des modèles Transformers pré-entraînés pour la classification des métastases osseuses. La dernière étude met en évidence le manque de robustesse de la classification à l'aide de méthodes d'apprentissage profond et propose une méthode pour améliorer la précision basée à la fois sur les approches CNN et Transformer.

Les résultats des différentes études préliminaires sont encourageants et prometteurs.

Mots-clés : Apprentissage profond, métastase osseuse, segmentation, classification, base de données, imagerie médicale, tomodensitométrie, CNN, Transformer, Unet.

PUBLICATIONS

- Afnouch, M., Gaddour, O., Hentati, Y., Bougourzi, F., Abid, M., Alouani, I., Ahmed, A. T. (2023). BM-Seg: A new bone metastases segmentation dataset and ensemble of CNN-based segmentation approach. Expert Systems with Applications, 228, 120376. https://doi.org/10.1016/j.eswa.2023.120376, **IF=8.5, SJR=Q1.**

- Afnouch, M., Gaddour,O., Bougourzi, F., Hentati, Y., Ahmed, A. T, Abid, M.. Automatic Bone Metastasis Classification: An in-depth Comparison of CNN and Transformer Architectures. **INISTA 2023, Class=C.**

- Application de machine Learning dans le domaine du diagnostic médical, Cas d'utilisation:“ Métastases osseuses ”, **Best Poster Award during Mardi des Chercheurs Journey, UPHF, 2023.**

- Afnouch, M., Gaddour,O., Bougourzi, F., Hentati, Y., Ahmed, A. T, Abid, M.. Machine Learning Applications for Bone Metastasis: Overview of Methods, Findings, and Datasets. **Submitted to ACM Computing Surveys.**

DEDICATION

I would like to acknowledge everyone who played a role in accomplishing this thesis.

To my beloved family that always supported me with their love and understanding. Without them, I could never have reached this current level of success.

I would like to give special thanks to my friends who have always encouraged me.

ACKNOWLEDGEMENT

First of all, I would like to express my deep and respectful gratitude to my PhD supervisors, **Pr. Mohamed ABID**, Professor at the National Engineering School of Sfax, and **Pr. Abdelmalik TALEB-AHMED**, Professor at Université Polytechnique Hauts-de-France, for the trust they have placed in me by agreeing to supervise my doctoral work. May they find in this work the expression of my deep gratitude.

I would also like to thank my co-supervisors, **Dr. Olfa GADDOUR**, assistant professor at the National Engineering School of Sfax and **Dr. Fares BOUGOURZI**, assistant professor at University Paris-Est Creteil for their outstanding supervision, advice, support, patience, and exceptional seriousness. I am also thankful for their assistance in keeping my progress on schedule.

I would like to acknowledge **Dr. Yosr HENTATI**, Expert radiologist in the Department of Radiology Hedi Chaker University Hospital, Sfax, and **Ahmed BEN SALEM**, radiologist in the Department of Radiology Hedi Chaker University Hospital, Sfax. I had the pleasure of working with them. I would be remiss in not mentioning the Department of Radiology Staff for the chance to do my research in collaboration between an academic laboratory and an honored hospital.

My deepest thanks go to the members of my defense committee who agreed to judge my thesis. My sincere gratitude goes to **Pr. Dorra SELLAMI**, Professor at the National Engineering School of Sfax, for giving me the honor of being the chair of the jury. Furthermore, I deeply thank **Pr. Nawres KHALIFA**, Professor at Tunis Al-Manar University, **Pr. Frederic CHAUSSE** professor at University of Clermont Auvergne, France, and **Dr. Jacques BOONAERT**, Associate professor at MT Nord Europe, France for accepting to be the reviewers of this thesis report and for giving me the honor to judge this work with great interest.

No words can express my gratitude to my colleagues in CES and IEMN laboratories for their exchanges and administration help, in addition to the wonderful moments we shared.

Contents

LIST OF FIGURES

LIST OF TABLES

ACRONYMS

AI	Artificial Intelligence
AUC	Area Under ROC Curve
Acc	Accuracy
BM	Bone Metastasis
BM-Seg	Bone Metastasis Segmentation
CNN	Convolutional Neural Network
CT	Computed Tomography
CAD	Computer-Aided Diagnosis
DL	Deep Learnning
DSC	Dice Score Similarity
DICOM	Digital Imaging and Communications in Medicine
FN	False Negative
FP	False Positive
FCN	Fully Convolutional Network
GAN	Generative Adversarial Network
GPU	Graphics Processing Unit
IoU	Intersection over Union
ISBI	International Symposium on Biomedical Imaging
KNN	k-Nearest Neighbors

LSTM	Long Short-Term Memory
LR	Logistic Regression
ML	Machine Learning
MM	Multiple Myeloma
MRI	Magnetic Resonance Imaging
MESCC	epidural spinal cord compression
PET	Positron Emission Tomography
Prec	Precision
RAM	Random Access Memory
ResNet	Residential Network
ROI	Regions Of Interest
ROIs	Regions Of Interests
R-CNN	Region-based ConvNet
RF	Random Forest
ROC	Receiver Operating Characteristic
SPECT	Single Photon Emission Computed Tomography
Sens	Sensitivity
Spec	Specificity
SVM	Support Vector Machine
TP	True Positive
YOLO	You Only Look Once
2D	2 dimensional
3D	3 dimensional

1

INTRODUCTION

Contents

In recent years, machine learning and artificial intelligence have had a revolutionary impact on healthcare, particularly in the field of medical image analysis. These advances have the potential to significantly improve diagnostic accuracy, which is a critical factor in patient outcomes and overall quality of care. In this work, we explore the profound capabilities of machine learning algorithms in medical diagnosis, focusing on the important area of bone metastases in cancer treatment. Accurate identification and characterization of bone metastases is a major challenge in this field. This phenomenon, in which cancer cells spread to the bone, is a major medical problem due to its frequency, complex manifestations, and significant impact on patient prognosis and quality of life. Addressing this challenge requires the use of sophisticated computational methods that can be seamlessly integrated into clinical workflows and provide reliable diagnostic support.

1.1 THESIS CONTEXT

The advent of Deep Learning has ushered in a new era in medical imaging. It offers the ability to uncover intricate patterns and relationships in medical data that were previously beyond the reach of conventional methods. In this landscape, metastases refer to a group of abnormal cells that develop outside their normal boundaries and spread to other organs. In particular, bone metastases are cancers that originate in organs of the body such as the breast, lung, or prostate, and spread to the bone. Metastases have been identified as a contributing cause in 90 % of cancer deaths [2, 3]. Bones are the third most common site of metastasis after liver and lung [4]. Statistics show that more than 70 % of patients with breast and prostate cancer have bone metastases [5]. Bone metastases represent a major medical problem because of their widespread prevalence, complex manifestations, and profound impact on patient prognosis and well-being. Timely detection and accurate diagnosis of bone metastases play a critical role in tailoring effective treatment approaches, tracking disease progression, and increasing patient survival rates.

The tasks of segmentation and classification of bone metastases, which are critical for accurate diagnosis, treatment formulation, and continuous monitoring, have undergone a profound revolution with the application of deep learning techniques. Conventional diagnostic techniques, while having their merits, can be hampered by subjectivity, variability, and time-consuming procedures. These challenges are exacerbated by the varying manifestations of lesions and complicated anatomical structures. Therefore, the integration of state-of-the-art technologies, such as machine learning, is urgently needed to improve current diagnostic methods and skillfully overcome these obstacles.

1.2 MOTIVATIONS

Bone metastasis tracking requires medical imaging to diagnose disease and determine treatment efficacy. Several medical imaging modalities have been used, including computed tomography (CT), MRI (Magnetic Resonance Imaging), bone scans, and PET scans, each imaging modality having its advantages and disadvantages [6]. In particular, CT-scans have a high spatial resolution, allowing the detection of both large and mild and small bone lesions [5]. They also provide simultaneous assessment of primary bone and bone metastatic lesions. In addition, CT-scans are usually available at relatively low cost [7]. In clinical practice, CT-scans are the most commonly used imaging modality for both baseline staging tests and serial monitoring of cancer patients [8].

However, detection of bone metastases using CT-scans is both difficult and time-consuming for the following reasons: (i) because bones are located throughout the body, radiologists must examine all slices [9]; (ii) the radiological appearance of bone metastases depends on lesion types [8]. Therefore, no single window parameter can adequately represent all bone metastases; and (iii) benign mimics such as fractures, bone islands, and degenerative changes can be confused with bone metastatic lesions, making diagnosis difficult [10]. Computer-aided diagnosis (CAD) can help radiologists find small lesions that might otherwise be missed [11]. Consequently, there is a great demand for CAD systems for bone metastases on CT-scans.

Recently, Deep Learning methods have attracted much attention to solve various visual tasks [12, 13, 14, 15]. In particular, Convolutional Neural Networks (CNNs) achieved top performance on several types of image recognition problems, including the diagnosis of bone metastases.

The motivation for this work lies in the potential of machine learning algorithms to provide faster, more accurate, and more objective diagnoses of bone metastases. By harnessing the power of these algorithms, physicians can make more informed decisions that lead to timely interventions and better patient outcomes. In addition, integrating machine learning into the medical diagnostic process has the potential to reduce the workload of physicians and improve the overall efficiency of the healthcare system, allowing more patients to receive timely and appropriate care.

1.3 OBJECTIVES

The primary objectives of this PhD project are as follows:

- Review of the current state-of-the-art techniques in medical diagnosis and the existing challenges in diagnosing bone metastasis.
- Exploration of various machine learning methodologies, including deep learning, and ensemble methods, to identify their potential applications in diagnosing bone metastasis.
- Collecting new bone metastasis datasets to facilitate the segmentation and identification of bone metastasis. This dataset aims to assist the research community in enhancing their evaluations of bone metastasis.
- Proposing an innovative deep learning approach tailored to optimize bone metastasis segmentation.
- Exploration of the use of advanced deep learning methodologies to enhance the accuracy and efficiency of bone metastasis classification in medical imaging.
- Creation of a standardized evaluation framework for the automatic classification of bone metastases and deployment of an innovative approach to classify metastatic lesions.
- Evaluation of the proposed models in comparison to existing diagnostic approaches and investigation of their ability to provide insight into the underlying features and patterns that contribute to the diagnosis of bone metastases.
- Discussion of the clinical implications and real-world applicability of the proposed models, considering potential challenges and opportunities for integration into clinical practice.

1.4 CONTRIBUTIONS

To address these objectives, the main contributions of this work are summarized as follows:

- We first provide a comprehensive overview of the techniques used to classify and segment bone metastases. In this review, we provide an overview of bone metastases, including the importance of medical oncology and medical imaging modalities. We then present the state-of-the-art in current methods for classification and segmentation of bone metastases. We focus on the comparison of these methods in terms of various parameters such as performance evaluation and datasets.

- We propose a benchmark dataset for bone metastasis analysis. This dataset was created using CT-scans from more than 57 patients from Hedi Chaker Hospital in Sfax, Tunisia, with more than 102.614 images. To our knowledge, this is the first publicly available dataset in the field of bone metastasis diagnosis that includes CT-scans as an imaging modality. Thus, potential advances in the field of bone metastases could be made with our dataset.

- We propose a new segmentation approach to detect bone metastasis where we propose a hybrid AttUnet++ architecture, Dual-decoders, and ensemble approach. We evaluate our technique with the newly proposed dataset. The obtained results show higher performance in lesion segmentation compared to the state-of-the-art techniques.

- The last contribution includes a comprehensive comparison of leading CNN architectures such as InceptionV3, EfficientNet, ResNext50, and DenseNet161 with transformer architectures, namely ViT and DeiT, for the purpose of bone metastasis classification. This is in addition to the introduction of an innovative approach using a benchmark dataset to provide a standardized evaluation framework for the automatic classification of bone metastases.

- For results reproducibility, we made our dataset and codes publicly available at https://github.com/Marwa-Afnouch/EH-AttUnetplus

1.5 THESIS OUTLINE

This thesis is organized as follows.

In Chapter 2, the current state of the art is presented by first placing the topic of bone metastases in the context of machine learning. In addition, an overview of bone metastases and the relatively most important challenges in this area is provided. Ma-

chine learning applications in bone metastasis segmentation and classification are then presented.

Chapter 3 presents the collected dataset. This chapter details the implementation in terms of the imaging modality used, the acquisition and labeling process, and the respective challenges.

Chapter 4 discusses in more detail the development of a CAD system for bone metastases. In this chapter, we present a new architecture derived from Unet++ that combines attention gates and dual decoders. It efficiently segments both bone metastases and bone regions. To improve the segmentation of metastatic bone, we propose a *ensemble* approach. We perform extensive experiments on our dataset and obtain good results compared to other CNN-based segmentation architectures.

In Chapter 5, we perform a comparative study using state-of-the-art techniques for bone metastasis classification. We then present a novel classification approach that incorporates dual deep-learning models. The obtained results are discussed in detail in this chapter.

Finally, the conclusions of this work and some perspectives are presented in Chapter 6.

2

MACHINE LEARNING IN BONE METASTASIS: STATE OF THE ART

Contents

2.1 INTRODUCTION

In this chapter, we first present the current state of the science regarding bone metastases, including the importance of medical oncology and its clinical impact on patients' quality of life. We then provide a detailed overview of machine learning methods for bone metastases, including classification and segmentation methods.

This chapter is organized as follows. Section 2.2 presents an overview of bone metastasis that includes definition, primary cancer sites, and the clinical implications surroundings bone metastasis. The ML methods in the bone metastasis field are deeply detailed in Section 2.4. Finally, section 2.5 refers to the conclusion of this chapter.

2.2 OVERVIEW OF BONE METASTASIS

2.2.1 *Bone Metastases in Medical Oncology*

Bone metastases (BM), also referred to as bone mets or bone metastatic disease, are a medical condition in which cancer cells from a primary tumor spread to the bone. Although they can occur in any type of cancer, they are commonly seen in breast and prostate cancer [16]. The invasion of cancer cells into the bones can cause significant damage and fragility, resulting in severe pain, fractures, and other complications. In addition, cancer cells can disrupt normal bone function, resulting in decreased bone strength and an impaired ability to heal itself. In addition, BM can attack the nerves and spinal cord, causing numbness, tingling and even paralysis [17]. This is a serious complication of cancer that is a major health problem and can be difficult to manage. BM is a major cause of morbidity in people with cancer.

Skeletal metastases tend to occur in body regions where red bone marrow is most abundant, suggesting that the greater blood supply to red bone marrow compared with yellow bone marrow is an important factor. The most commonly affected sites include the vertebrae, favoring the lumbar spine over the thoracic and cervical spine, as well as the pelvis, proximal femur, ribs, and scapula, as shown in Figure 4.11. On the other hand, metastases to distal limb sites, such as the elbow and knee, are rare. In summary, the distribution of BM in the body is not random and tends to occur in areas with a high concentration of red bone marrow. This underscores the importance of understanding the

underlying mechanisms that drive the development and spread of cancer cells to these areas.

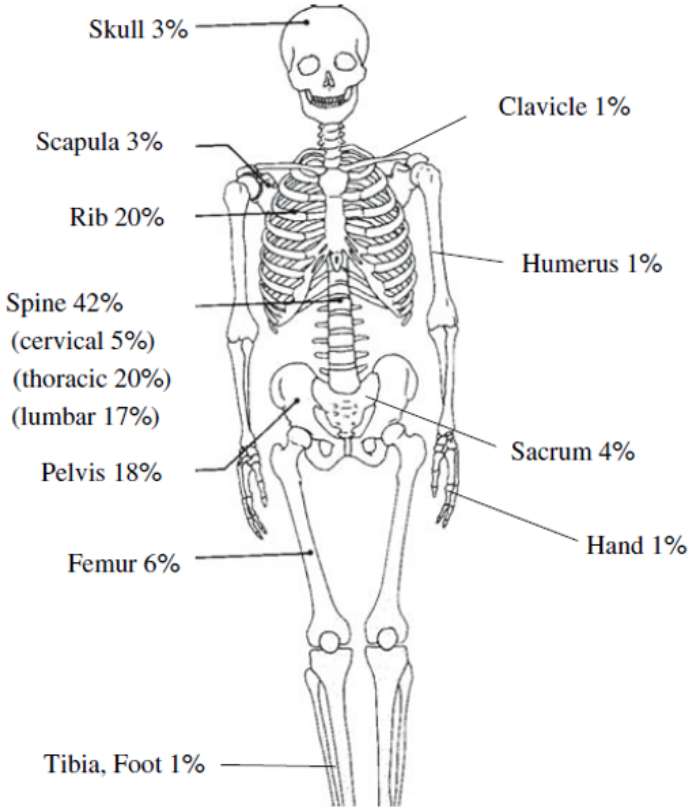

Figure 2.1: Anatomic localization of skeletal metastases from lung cancer [18]

BM may occur by hematogenous or lymphatic routes or by direct spread of tumors to the bone. Regardless of the route of spread, BM results in both bone loss and bone formation, leading to different patterns of bone destruction and remodeling. BM can adopt one of three dominant patterns, including lytic (osteolytic), sclerotic (osteoblastic), or mixed lytic and sclerotic metastases. In addition, BM may also have different morphological features, including diffuse, focal, or expansive. While lytic metastases cause bone destruction, sclerotic metastases promote bone formation. Figure 2.1 illustrates lytic and sclerotic bone lesions.

Diagnosis of BM can sometimes be difficult, especially in elderly patients with degenerative diseases and osteoporosis. Imaging tests and serum tumor markers are critical for diagnosis; in some cases, a bone biopsy may be required. Treatment of BM is aimed at relieving symptoms, as a cure is rarely possible. Treatment options include external beam radiation therapy, endocrine treatments, chemotherapy, targeted therapies, radioisotopes,

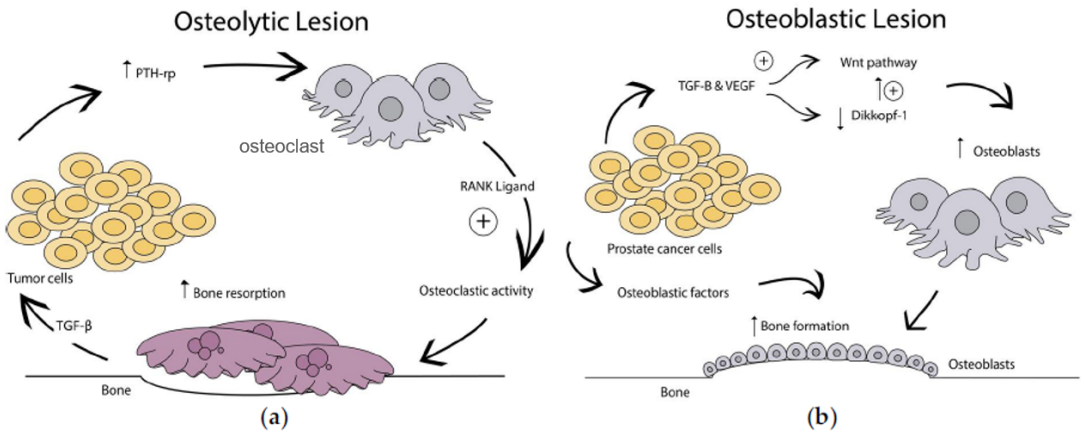

Figure 2.2: Metastatic Bone Lesions [19]

and orthopedic surgery for structural complications. The choice of treatment depends on the extent of bone disease, the presence of metastases outside the skeleton, and the underlying malignancy. As the disease progresses, resistance to systemic treatments may develop, necessitating a change in therapy. Artificial intelligence can help us improve the detection of bone metastases and potentially diagnose them earlier, leading to better treatment outcomes.

2.2.2 *Common Primary Cancers Leading to Bone Metastases*

BM happens when cancer cells migrate to the bones from other parts of the body. According to [20], breast, prostate, and lung cancers account for about 80% of all primary cancers that metastasize to the skeleton. In addition, bone metastases from thyroid and kidney cancers are also conceivable. Bone is the third most common site for metastatic disease after the lung and liver. According to the [4] study, breast and prostate cancers can cause up to 70% of skeletal metastases. Lung cancer, renal cell carcinoma, and prostate cancer are the top three cancers that most commonly cause bone metastases. When the tumor and microenvironment are susceptible, bone metastases occur more frequently. Skeletal problems associated with advanced cancer are common in patients [21]. Indeed, the presence of bone metastases significantly shortens patient survival and makes it challenging to live beyond a relatively short period of time, as seen in Table 2.1.

Table 2.1: Incidence of BM in cancer (Data in the table based on [4] and [22])

Primary cancer type	Relative incidence in bone	Median survival from diagnosis
Breast	65%–75%	19–25 months
Prostate	65%–75%	12–53 months
Lung	30%–40%	6 months
Thyroid	40%–60%	48 months
Bladder	40%	6–9 months
Renal	20%–25%	12 months
Melanoma	14%–45%	6 months

2.2.3 *Clinical Impact of Bone Metastases on Patient Quality of Life*

Clinically, the presence of bone metastases can have a significant impact on a patient's prognosis and overall quality of life. Moreover, bone metastases can give rise to malaise, a decline in performance status, and a lower quality of life [23]. Pathologic fractures, spinal cord compression, and hypercalcemia are among the common skeletal-related events (SREs) associated with bone metastases. These events necessitate radiation therapy and/or surgical intervention targeting the affected bone. As the prevalence, incidence, and survival rates of patients with bone metastases continue to rise, the consideration of quality of life has become integral to their management. The impairment of quality of life is frequently attributed to heightened bone pain at the time of diagnosis and the presence of bone metastases [24]. Metastatic bone disease in individuals with advanced cancer is typically associated with skeletal problems and can result in debilitating pain. The general quality of life and emotional functioning of patients with metastatic bone disease deteriorated significantly during the COVID-19 pandemic due to disruptions in healthcare services, increased isolation, limited access to support systems, and fear of infection [25]. It is widely accepted that bone metastasis is one of the most important factors that affect the quality of life and survival in patients with HCC [26]. Therefore, it is important to consider the clinical implications of bone metastases when developing treatment plans for patients with cancer.

2.3 IMAGING MODALITIES FOR BONE METASTASIS DIAGNOSIS

There are several types of medical imaging modalities that have been used to diagnose, detect, and segment BM. The main imaging modality used for the detection and evaluation of BM is a combination of positron emission tomography (PET) and computed tomography (CT)). In addition to PET/CT, other imaging modalities such as magnetic resonance imaging (MRI) and bone scans are also used in the evaluation of bone metastases. It is important to note that the choice of imaging modality depends on various factors, including the primary cancer site, the suspected extent of metastatic disease, and the specific clinical scenario. Each imaging modality has its own strengths and weaknesses, and combining different modalities can increase sensitivity and specificity, leading to improved diagnostic accuracy. Therefore, a multidisciplinary approach involving radiologists, oncologists, and other healthcare professionals is crucial in determining the most appropriate imaging strategy for each individual patient. In the following sections, we will detail deeply some of the imaging modalities that are used in BM diagnosis.

2.3.1 *Bone Scintigraphy*

Bone Scintigraphy, also called Bone Scan, is an imaging modality that involves injecting a radioactive substance into a patient's body. It is used to examine the various bones of the skeleton. Two gamma cameras are placed in front and behind the patient to detect radiation emitted from the injected radioactive substance. The two resulting images will show hotspots with high intensities. As the bone scan is widely available and less expensive, it continues to be the most widely used radionuclide technique to evaluate skeletal metastases [27]. The main advantage of bone scintigraphy is its ability to visualize the whole skeleton. This is important given that bone metastasis is often found in regions of the appendicular skeleton that are neglected during routine skeletal examinations [28]. Bone scan has the disadvantage of making it difficult to identify metastatic lesions from other possible hotspots which can lead to a false positive diagnosis. A further limitation is that the image resolution of nuclear medicine images may not be as high as MRI or CT.

2.3.2 *Computed Tomography (CT)*

CT-scans use X-rays and computer technology to create detailed images of the body. CT-scans have a high spatial resolution, which allows the detection of bone metastases by showing detailed images of the bone structure, including changes in bone density and the presence of tumors [5]. They are also useful in guiding needle biopsy, especially in spinal lesions. A further advantage of CT-scans is that they can detect bone marrow metastases before they destroy bone, resulting in an earlier diagnosis and an improved prognosis [29]. Besides, CT-scans are commonly available at a relatively low-cost [7]. However, CT-scans are relatively insensitive in showing small lesions, and it has the disadvantage of limited skeletal coverage.

2.3.3 *Magnetic Resonance Imaging (MRI)*

MRI is a radiation-free, noninvasive imaging modality that uses a magnetic field and radio waves to create precise images of the inside of the body and provide 3D visualization of tissues. With excellent soft tissue resolution, MRI allows the direct visualization of bone marrow and is useful for determining the extent of local disease before surgery or radiation [30]. It is also possible to detect metastatic lesions using MRI before changes in bone metabolism become detectable on bone scans. Although the whole-body MRI is an accurate technique, it is currently limited by its high cost and takes more than 30 minutes which makes image acquisition more difficult mainly for restless people.

2.3.4 *Positron Emission Tomography (PET)*

PET is a nuclear medicine imaging technique that produces 3D anatomic information or map of functional processes in the body. A small amount of radioactive material is injected into the patient's bloodstream, and the gamma rays emitted by the radioactive material uses detected by the PET scanner. Because bone metastases often have increased blood flow, PET can be used to detect the abnormal metabolic activity of cancer cells in multiple locations throughout the whole body in a single scan. The applicability and utility of PET images lie in the radiopharmaceutical used, among other reasons, however, PET images have a higher resolution compared with conventional planar and SPECT techniques [27]. PET scan has proved to be a noninvasive modality, which means that they do not involve any incisions or radiation exposure [31]. Despite its high

functional assessment, PET scan is available in only very few centers and is a high-cost, sophisticated test that requires special expertise. It's important to note that despite these disadvantages, PET scans can provide important information for the diagnosis and management of bone metastases, and may be used in conjunction with other imaging modalities for better results.

2.3.5 *Single Photon Emission Computed Tomography (SPECT)*

SPECT is one of the most commonly used techniques that uses trace concentrations of radioactively-labeled compounds to provide insight into physiological processes. In SPECT examination, imaging equipment captures the emitted gamma rays from radionuclides that were injected into a patient's body in advance to generate a map of the inside of a body [8]. Unlike conventional skeletal scintigraphy where planar imaging is limited by the superimposition of structures, SPECT images can show axial slices through the body, which is more precise in localizing abnormal radionuclide uptake [32]. Moreover, SPECT scan is a widely used tool due to the low-cost equipment. It can also provide an accurate assessment of disease stage and severity. However, SPECT imaging is characterized by inferior spatial resolution and low signal-to-noise ratio. Furthermore, manual analysis of SPECT scan findings by radiologists is time-consuming and subjective [33].

2.3.6 *Hybrid imaging methods*

Molecular and hybrid imaging has an increasing role in the early detection of metastatic bone and in monitoring response at early time points [34]. In this sense, functional imaging such as SPECT/CT, PET/CT, and PET/MRI provides a standardized uptake value and allow the fusion of anatomic data from cross-sectional imaging with functional information from nuclear medicine studies. As a result, the radiologist can determine if focal radiotracer uptake on a nuclear medicine study corresponds to a discrete bone lesion. Similarly, diagnostic confidence increases when an osseous lesion suspicious for metastasis on cross-sectional imaging avidly accumulates radiotracer. Hybrid modalities allow imaging of the skeleton with high contrast and spatial resolution. Compared to conventional imaging methods, these modalities offer enhanced diagnostic precision for evaluating stage and response, as they can quantify the biological processes affecting the bone environment and tumor cells [35]. A potential weakness of hybrid imaging methods compared with standard imaging methods is that Additionally, hybrid imaging

Table 2.2: Imaging modality comparison

Imaging Modality	Advantages	Limitations
Bone Scan	- Sensitive to early bone changes - Whole-body coverage - Detects multiple lesions simultaneously	-Low specificity - Cannot provide precise anatomical detail -Requires radiotracer injection
CT Scan	- High spatial resolution - Detailed visualization of bone structures -Rapid scanning	- Ionizing radiation exposure - Limited soft tissue contrast - May miss small lesions
MRI	- Excellent soft tissue contrast - Multiplanar imaging capability - No ionizing radiation exposure	- Expensive and time-consuming - Not widely available - Motion artifacts may affect images
Hybrid Methods	- Combined advantages of different modalities - Improved sensitivity and specificity - Enhanced anatomical and functional data	- Costly - Limited availability - Longer scan times

modalities may also be associated with higher radiation exposure compared to standard modalities.

2.4 MACHINE LEARNING TECHNIQUES IN BONE METASTASIS

Machine learning has been used in various ways to predict and differentiate bone metastases and to classify bone tumors. In this section, we will detail some examples of machine learning applications to bone metastases. Figure 2.3 shows a literature-based taxonomy for bone metastasis analysis.

While there are numerous applications of machine learning in bone metastases, including detection, prediction, and classification, we will specifically address the classification and segmentation of bone metastases. Classification is about developing models that can accurately distinguish between benign and malignant bone tumors or classify different types of bone metastases based on imaging data. Segmentation, on the other hand, is about accurately delineating the boundaries of bone metastases in medical images to enable precise localization and measurement of lesions. By focusing on these two aspects, we aim to contribute to the evaluation of bone metastases and provide insights to improve diagnosis and treatment planning for patients.

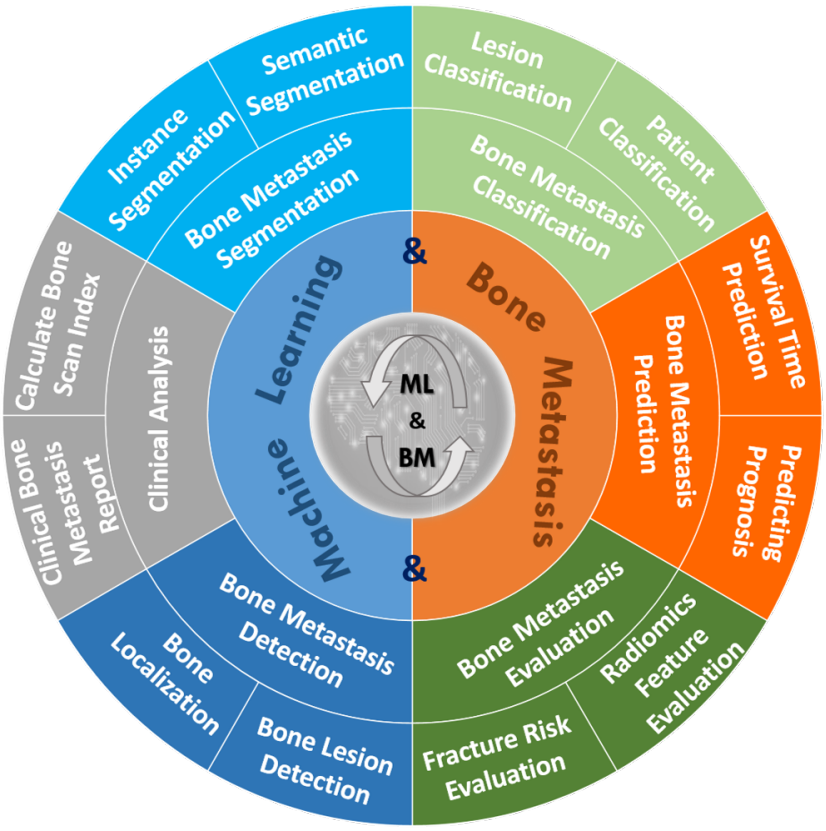

Figure 2.3: Literature Taxonomy of Bone Metastasis Using Deep Learning

2.4.1 *Existing Bone Metastasis Classification Methods*

The classification of BM using machine learning methods has been the subject of several recent research works. In the following, we describe papers about BM classification based on imaging modalities.

a. Bone Scan

One of the first CAD systems to classify bone scans was BONENAVI® [36], a software tool that can automatically detect hot spots in bone scans and classify the hot spots as BM or non-malignant. The BONENAVI software achieved a sensitivity of 90% [36]. Later, Tokuda et al. [37] investigated the diagnostic capability of the fully automated CAD system "BONENAVI version 1" by focusing on two different patterns. The first pattern was dedicated to the detection of metastases per region, and the second to the detection of metastases per patient. The results obtained showed that the BONENAVI system

is able to reduce the number of false positives, which depends on the primary cancer lesion. Similarly, Mitsuru Koizumi et al. [38] used BONENAVI version 2 to evaluate the accuracy of BM diagnosis using a ANN model with a threshold of 0.5. The sensitivity and specificity were 85% and 82%, respectively, while the highest accuracy was 72% for BM of prostate cancer. In addition, Kikushima et al. [39] showed that bone scans analyzed with BN2-B accurately detected the presence of BM in cancer patients. They reported that the sensitivity and specificity for BN2 were 94% and 88% in male patients and 86% and 85% in female patients, respectively. BONENAVI® has a prognostic effect in prostate cancer patients [40], although the sensitivity of BONENAVI is not as high in cases with disseminated skeletal metastasis [41].

Another automatic diagnostic system for identifying possible metastases in cancer patients was proposed by Aslantas et al. [42] "CADBOSS". CADBOSS used image rasterization to extract features from whole-body bone scans and then employed a ANN classifier at the lesion level. The calculated specificity, sensitivity, and accuracy of CADBOSS were 87%, 94%, and 92%, respectively.

Recently, a new parallelepiped classification method (PC) was proposed to automatically detect BM in bone scan images [43]. The method PC involved mapping the radionuclide distribution in the images and generating color maps of the scintigraphic images. It was evaluated using images from 12 patients with BM. The results showed that the PC method could distinguish metastatic bone from normal tissue with an accuracy of 87.58% and a K coefficient of 0.8367.

The use of CNNs in classifying BM has shown considerable potential in several studies. For example, Belcher [44] used CNNs to classify hot spots as benign or malignant, achieving an AUC of 0.9739. Hot spots were hand-extracted from 2164 prostate cancer patients, and 10,428 hot spots from the lower spines were included because they were considered easiest to classify. In the same way, Dang et al. [45] used an ensemble patch-based CNN model to detect metastatic hot spots on bone scans with 89% accuracy. Later, [46] introduced a CNN method for analyzing both anterior and posterior views on bone scan examinations to detect BM in patients with different types of cancer. The authors employed a spatial attention feature aggregation operator to improve the extraction of spatial location information. Their model was trained on a large dataset of 15,474 trials from 13,811 patients. Results showed excellent performance metrics, including an F1 score of 0.933, accuracy of 95.00%, sensitivity of 93.17%, and specificity of 96.60%.

In [47], the authors propose a new lightweight CNN architecture for BM classification in nuclear medicine, specifically for prostate cancer patients. The proposed method achieved 97.41% accuracy with a dataset of 778 images. Papandrianos et al. performed

two studies to classify bone scintigraphy images of prostate cancer patients into three classes: normal, malignant, and degenerative. In the first study [48, 49], they developed a CNN architecture to diagnose BM and achieved an accuracy of 91.61% ± 2.46% and an F1 score of 0.938. In the second study [50], they used VGG16 and DenseNet to classify bone scintigraphy images. They found an accuracy of 92.08% ± 2.81% for Densenet and 92.14% ± 2.91% for VGG16. These results are consistent with two similar studies by the same authors [51, 52]in which they performed a dual classification: Metastases and non-metastases. The approach in [51] used CNNs to classify 908 images from breast cancer patients and achieved 92.50% accuracy. In addition, the CNN model in [52] achieved 97.38% accuracy for prostate cancer patients.

In a study by Zhao et al. [53], the authors collected bone scans from 12,222 patients with 40 cancers, with 42.1% of the scans viewed at BM to detect malignant bone disease using the ResNet-50 model. Their results showed an overall AUC of 96.4% for diagnosing various cancers. In another study, Hajianfar et al. [54] used a dataset of 2253 patients with suspected bone disease and metastases and two assessment strategies, one for detecting normal vs. abnormal bone and another for distinguishing metabolic bone disease from BM. Six CNN models, including ResNet50, Inception-V3, and Inception ResNet50-V2, were used to analyze the images. The model with the best performance achieved an AUC of 0.68 for discrimination between normal and abnormal and 0.65 for discrimination between metabolic and metastatic.

Similarly, Liu et al. [55] constructed an AI-assisted identification of suspicious bone metastases from bone scan images using CNNs. The proposed CNN model achieved an accuracy of 81.23%. Similar to most studies, the results showed that complicated models such as DenseNet did not perform better, which may be due to the relatively simple nature of bone scan images. In [56], the authors presented an object-oriented classification for the interpretation of bone scan images. The method used the spatial and textural information of pixel groups to classify images with higher accuracy. Han et al. [57] proposed two different CNN architectures for analyzing bone scans of patients with prostate cancer: whole-body-based (WB) and global-local uniformly emphasized (GLUE). Both were trained on limited data to analyze 9113 images. The achieved accuracy was 88.9% for the WB model and 90% for the GLUE model . Recently, [58] developed a DL algorithm to classify areas of increased uptake on bone scans to improve the diagnosis of metastatic bone disease. The DL algorithm showed high performance in detecting BM with a specificity of 0.80 and sensitivity of 0.82 on the external test set. The DL algorithm outperformed the nuclear medicine physicians in terms of time efficiency. It took only 2.5 minutes to classify 134 BS scans compared to 30 minutes for the physicians.

Tables 2.3, 2.4 and 2.5 present a summary of all papers in the literature dealing with bone metastasis classification.

b. SPECT

Using SPECT images, Lin et al. have proposed different methods for BM classification. In [61], they presented a deep network called dSPIC for automated multi-disease and multi-lesion diagnosis. The network was designed to extract optimal features from images and classify them into classes such as metastasis, arthritis, and normal. The evaluated network achieved a value of 77.47% for accuracy, 78.83% for precision, 78.63% for sensitivity, 88.20% for specificity, and 78.60% for F1 score on the testing samples. Then, they proposed a self-defined CNN called Dscint [62], which used a hybrid attention mechanism to classify whole-body scintigraphic images into different disease categories. They included AlexNet, ResNet, DenseNet, VGGNet and Inception-v4 as backbones to make multi-class classifications. The accuracy of Dscint outperformed several classical deep classification networks with a value of 98.01%. Based on popular deep networks such as ResNet, VGG and DenseNet, they also developed deep classifiers [63] to automatically diagnose metastases in 251 thoracic SPECT bone images. They achieved an AUC of 0.98. More recently, three distinct two-class classifiers are proposed in [64] based on the VGG 16 model. These classifiers autonomously identify whether a SPECT image contained lesions or not. Experimental outcomes showcased an accuracy of 83.80%, a precision of 92.90%, a sensitivity of 96.60%, an F1-score of 90.80%, and an AUC value of 0.875.

In another study [66], the authors proposed a CNN-based classification model for accurate diagnosis of bone metastases. They adopted the standard VGG model to develop a classifier for SPECT images. The authors evaluated the created model using a set of 624 SPECT images and found an accuracy of 99.6%. In [65], an automated classification model based on the VGG model was proposed to determine whether an image contains lesions or not. Experimental evaluation of a series of 642 SPECT images showed that the VGG13 model used achieved 99.20% accuracy. Liu et al. [67] presented a method for classifying multiple diseases including normal cases, bone metastases, arthritis, and thyroid cancer using a customized CNN model called SPNT9. The SPNT9 model achieved a precision of 78.06%, a sensitivity of 79.04%, and an F1 score of 78.99%. More recently, [68] focused on thorax image classification using a self-defined CNN called Att-ResNet24 with a hybrid attention mechanism. Experimental results showed that Att-ResNet24 achieved an accuracy of 73.70%.

c. CT-scan

Table 2.3: Summary of existing Bone Metastasis classification works (part 1)

Ref.	Method	Imaging Modality	Dataset	Data Split	Evaluation Metric	Limits
[36]	BONENAVI(ANN)	Bone Scan	795 bone scans 904 bone scans	Test Set =257 patients	Sens=90% Acc=83%	Incomplete database details Limited comparison scope
[37]	BONENAVI(ANN)	Bone Scan	406 patients	External val=406	AUC=0.82	Biased data(prostate cancer)
[39]	BONENAVI BN2-B(ANN)	Bone Scan	399 patients	External Val= 399	Concordance Test Results= 65% Mismatch Test Results=35%	Retrospective and Single-center Study
[38]	BONENAVI(ANN)	Bone scan	394 patients	External Val=394	Sens=85% Spec=82%	Retrospective and Single-center Study
[40]	BONENAVI(ANN)	Bone Scan	226 patients	External val=226	Sens= 82% Spec=83%	Gender-biased data (prostate cancer)
[41]	BONENAVI(ANN)	Bone Scan	54 patients	External val=54	Sens= 76%	Low results
[45]	CNN	Bone Scan	2164 patients	Train=40% Val=30% Test=30%	Acc=89%	Biased data (prostate cancer only)
[59]	CADBOSS (ANN)	Bone Scan	60 patients 130 images	10-fold cross-validation	Spec=87%, Sens=94% Acc=92%	Small Dataset Insufficient capability due to handcrafted features
[44]	CNN	Bone Scan	2146 patients 10.428 hotspots	Train/Val=80% Test=20%	AUC=97.39%	Biased data (prostate cancer)
[43]	parallelepiped classification (PC)	Bone Scan	12 patients	N/A	Acc=87.58% K coefficient=0.8367	Small Dataset Insufficient capability due to handcrafted features
[47]	LB-FCN	Bone Scan	778 images	10-fold stratified cross-validation	Acc=97.41%	Moderate dataset size Limited to prostate cancer patients
[51]	CNN (3 conv layer,1 FC layers)	Bone Scan	908 images	10-fold stratified cross-validation Train=817, Test=90	Acc= 92.50%	Biased data (breast cancer)
[49]	RGB CNN (4 conv,2 FC layers)	Bone Scan	778 patients	Train=505 Val=156, Test=117	Acc=91.42%	Biased data (prostate cancer)
[52]	CNN (3 conv –pooling layers, 1 dense layer)	Bone Scan	586 images	Train=68% Val 17%, Test=15%	Acc= 97.38% Sens= 95.8%	Biased data (prostate cancer) Small Dataset
[50]	VGG16 DensNet	Bone Scan	778 scans	Train=85% Test=15%	Densenet Acc =92.08% ± 2.81% VGG16 Acc= 92.14% ± 2.91%	Biased data (prostate cancer) Small Dataset

Table 2.4: Summary of existing Bone Metastasis classification works (part 2)

Ref.	Method	Imaging Modality	Dataset	Data Split	Evaluation Metric	Limits
[48]	CNN (4 conv, 2 FC layer)	Bone Scan	778 images	Train=85% Test=15%	Acc=91.61% ±2.46%	Biased data (prostate cancer) Small Dataset
[53]	CNN	Bone Scan	12,222 patients	Train=9776 Val=1223, Test= 1223	AUC= 95.5%- 98.8%	No comparative analysis with existing methods high in contrast to real-world norms
[46]	Inception-V3	Bone Scan	15,474 images 13,811 patients	Train= 12,274 images Val=1600, Test=1600	F1= 0.933, Acc=95.00% Sens=93.17%, Spec=96.60%	No comparative analysis with existing methods Lack of details about the diversity of the dataset
[55]	CNN (ResNet50)	Bone Scan	3352 patients	5 fold cross validation	Acc=81.23% Sens=81.30%, Spec=81.14%	Biased data (specific regions: pelvis, sacrum, iliac joint Specific primary cancer: lung, prostate, breast)
[56]	SVM and KNN	Bone Scan	9 images	Train=46%, Test=54%	KNN Acc=86.62% SVM Acc= 86.81%	Small sample size No comparison with existing automatic interpretation systems
[54]	CNN pre-trained models	Bone Scan	2253 patients	Train=80% Test=20%	AUC=68%,Sens=66% Spec=71%	Modest outcomes
[57]	2 CNNs: WB model GLUE model	Bone Scan	9113 bone scans 5342 prostate cancer patients		Acc=88.9% for WB model Acc=90% for GLUE model	Biased data (prostate cancer) 2D CNNs might not fully capture data complexity
[60]	CNN	Bone Scan	2185 patients	Train=70% Test=30%	Acc=73.92%, Prec=75.92% Sens= 72.42%,F1= 72.92%	Lack of External Validation
[58]	VGG16	Bone Scan	2365 images	Train=1203, Val=164 External Val=998	Spec=80% Sens=82%	Quantitative assessment lacking Need for prospective validation
[61]	dSCIP (CNN)	SPECT	768 SPECT images 384 patients	Train= 70% Test=30%	Acc=77.47%, Prec=78.83% Sens=78.63%,Spec=88.20%, F1 =78.60%	Not distinguish between primary cancers
[62]	Dscint(CNN)	SPECT	600 patients 1078 images	Train/Test=70/30	Acc=98.01%, Prec=97.95%, Sens=97.91% Spec=99.33% F1=97.92%	Sampling bias
[63]	VGG ResNet DenseNet	SPECT	251 Thoracic SPECT images	Train=70% Test=30%	AUC=0.98	Biased data(thorax) Risk of overfitting
[64]	3 classifiers based on VGG16	SPECT	3831 patients	Train=3271 Test=560	Auc=83.8%, Prec=92.9%, Sens= 96.6% F1=90.8%, AUC= 87.5%	Imbalanced Data skewed classification accuracy
[65]	VGGNET	SPECT	642 images	Train=70% Test=30%	Acc=99.20%, Prec=99% Sens=99%, F1=95%	Low results on original data

Table 2.5: Summary of existing Bone Metastasis classification works (part 3)

Ref.	Method	Imaging Modality	Dataset	Data Split	Evaluation Metric	Limits
[66]	VGGNET	SPECT	642 images	Train/Test= 80/20	Acc=91.4%,Prec=88.4% Sens=95.3%,F1=99.5%	Small Dataset
[67]	CNN	SPECT	Train=80% Test=20%	615 images	Acc=88.14%, Prec=78.06% Sens= 79%, F1 =78.99%	Limited Dataset Size Clinical Validation
[68]	Att-ResNet24	SPECT	1668 images	Train/Test=70/30	Acc=73.7%, Prec=74.4 Sens=73.6, F1= 73.5	Biased data (Thoracic region, lung cancer)
[69]	Watershed algorithm Graph cut	CT	22 clinical cases	12 cases for training 10 cases for testing	Sens=71.2%	Small Test Dataset Biased data(sclerotic BM)
[70]	CNN	CT	59 patients	5 fold cross validation	Sens= 60-80% AUC=83.4%	Focus on for sclerotic metastasis Small dataset Varying slice thickness compromises detail
[71]	ML classifiers best algo: KNN	CT	58 patients	Train=41, Test=17	Sens=82.4% Spec=81.5% Acc=82.4%, AUC=0.861	Small dataset Biased data(lytic lesion)
[72]	CNN (2D ResNet50 3D ResNet182D)	CT	2.880 CT scans 114 patients	Train=75% Val=12%, Test=13%	Acc= 92.2% F1=92.3%	Biased data (prostate cancer) No comparative analysis with other classification models
[73]	Radiomics ML classifiers	CT	359 patients	Train=70% Test=30%	AUC=0.97 for GPR (best classifier)	Small Dataset Biased data(Lung cancer)
[74]	Radiomics +NLP ML classifiers	CT	176 patients	Train=70% Test= 30%	Acc=0.82, Sens=0.59 Spec=0.85, AUC=0.83	Small Dataset Resilient Low Sensitivity
[75]	ML Classifiers	PET/CT	75 patients	10 fold cross validation	AUC=0.76 for best classifier: weighted KNN algorithm	Biased data (prostate cancer)
[76]	ML classifiers	MRI	107 patients 60 MM lesions; 118 metastatic lesions	10-fold cross-validation	MCC = 0.605; Acc= 81.50% Sens=87.90%; Spec=79%	Modest results
[77]	Radiomics ML classifiers	MRI	146 patients	Train=100, Test=46 External test=35	2-label classification: Acc=86% 3-label classification: Acc=69%	Small sample size focused on spinal lesions
[78]	LR	MRI	103 MM patients 138 patients with metastases	5-fold cross-validation	AUC=85%	Small Dataset Evaluation only on AUC metric
[79]	DL based on ResNet 50	MRI	164 patients 215 thoracic MRI spine	Train/Val= 82% Test=18%	Kappas = 0.94–0.95, $p < 0.001$	Difficulty in interpreting the learned features Region and sequence limitations

AUC= Area under the ROC Curve, Acc=Accuracy, Sens=Sensitivity, Spec=Specificity,Prec=Precision, F1=F1-score, MM=Multiple Myeloma, conv= convolution layer, FC= Fully Connected, LR= Logistic Regression, Val=Validation

The authors in [69] presented a CAD system for detecting sclerotic BM in the spine. The system used a watershed algorithm and graph cut to detect lesions and an SVM classifier to classify them. Similarly, Roth et al. [70] proposed a two-stage coarse-to-fine cascade [80] to detect sclerotic spinal metastases. Using CNN classifiers, difficult false positives were eliminated by an additional selection process. Compared with the high sensitivity results of 60%, 70%, and 80% in their test set, the false-positive rate decreased significantly from 4 to 1.2, 7 to 3, and from 12 to 9.5, respectively, while an AUC of 0.83 was achieved. Later, Mutlu et al. [71] used eight ML algorithms to classify osteolytic bone lesions from 58 patients as malignant or benign. They found that KNN had the best predictive performance with values for accuracy, AUC, sensitivity, and specificity of 82.4%, 86.1%, 82.4%, and 81.5%, respectively. Also, a DL model [72] using 2,880 annotated CT scans from 114 patients with prostate cancer was able to detect bone lesions and classify them as malignant or benign with an accuracy of 92.2%.

A CT -radiomics model that distinguishes BM from normal bone marrow was published in [73]. After manually locating the lesions, they performed automatic segmentation to collect sufficient information. In the study, BM was successfully distinguished from healthy bone regions using various radiomic features and ML classifiers. More recently, Naseri et al. [74] presented an approach combining NLP and radiomics to discriminate between painful and painless BM lesions in simulated CT images of cancer patients. By automatically extracting pain scores from clinical notes using NLP and identifying BM central points on CT images, the study extracted radiomics features from these areas. The best model NNet achieved an accuracy of 82%, a sensitivity of 59%, a specificity of 85%, and an AUC of 83%.

Using PET /CT images, Acar et al. [75] focused on the use of CT texture analysis in combination with ML methods to distinguish between metastatic bone lesions and fully addressed sclerotic areas in prostate cancer patients with BM. With an AUC of 0.76, they found that the weighted KNN was the best algorithm.

d. MRI

The authors in [77] evaluated the diagnostic performance of ML classifiers in the differential diagnosis of spinal lesions using radiomics data. Lesions were classified into two groups: benign and malignant (2-label classification) or three groups: benign, primary malignant, and metastatic (3-label classification). For 2-label classification, D-ANN, the best model, achieved 94% accuracy on internal test data and 86% on external data. For 3-label classification, the best model, the Iteratively Optimized Logit Boost decision stump tree (IOLB), provided an accuracy of 80% for the internal test data and 69% for the external test data.

In the same context, Liu et al. [78] and Xiong X et al. [76] conventional T1-weighted (T1W) and fat-suppressed T2-weighted (T2W) MRI sequences to discriminate between spinal metastases and multiple myeloma(MM). [76] combined the radiomics models with a variety of ML algorithms to predict the probability of spinal metastases. The accuracy, sensitivity, and specificity of the radiomics model developed by Xiong X et al were 81.5%, 87.9%, and 79.0%, respectively, in their validation cohort. On the basis of logistic regression [78], the model of Liu et al performed well with 10-EPV (events per independent variable) in discriminating MM from spinal metastases with an AUC of 0.85. Recently, Hallinan et al. [79] developed an automated classification system for metastatic epidural spinal cord compression (MESCC) on MRI using Bilsky classification. Compared with specialized readers on internal and external datasets, the model showed near perfect agreement with kappas of 0.92–0.98.

2.4.2 *Segmentation methods for metastatic bone lesions*

Automatic segmentation of metastatic lesions is still in its infancy. A variety of imaging modalities have been explored for BM segmentation using machine learning approaches. Various methods, including deep learning models, segmentation algorithms, and combination techniques, have been used to address the challenges of automatically identifying and delineating bone metastases. Below, we describe the current state of the art in bone metastasis segmentation using the most commonly used imaging modalities, including bone scans, CTs, MRIs, and PET /CTs.

a. Bone Scan

Among the various imaging modalities that have been explored in the field of BM segmentation, the most commonly used modality is bone scanning. As can be seen in Table 2.6, subsequent work has utilized bone scanning imaging modalities to advance the field of BM segmentation. In [81], an automated method for segmentation of BM was proposed that takes into account the characteristics of different anatomical regions and has comparable sensitivity and significantly improved overall specificity. In [82], the authors proposed a method for diagnosing BM using fuzzy C-Means clustering and image processing techniques. The method was evaluated using bone scintigraphy from twelve patients with breast and prostate cancer and was able to determine the locations and areas of bone involvement as well as abnormalities of disease along the times. [83] used a Sparse autoencoder and CNNs to train an image-level classifier that classified input images as normal or suspicious, resulting in a 95 percent accuracy rate. Next, an image-level classifier was trained to produce a probability map of the hot spots. Finally,

level set segmentation was applied to the probability map to identify hot spots. The test data set contained 68 suspicious thoracic images with 572 hot spots and a Jaccard index of 0.8051.

Later, [84] proposed an image interpretation system for skeletal segmentation and extraction of hotspots of a metastatic bone from a whole-body bone scintigram based on Deep Learning. [85] provided a literature review on the Constrained Local Model (CLM) and its applications in medical image segmentation. CLM is a technique for modeling a class with a particular set of corresponding features. The proposed approach, which employs CLM with regularized landmark mean shift (RLMS), effectively mitigates ambiguity and outperforms the general CLM-based method, as shown by experimental results. [86] proposed a novel neural network model called MaligNet for bone lesion instance segmentation and bone cancer metastasis classification in breast using bone scintigraphy. The authors in [87] proposed a network to evaluate the bilateral difference of a whole-body bone scintigram and then integrate it with their previous network extracting the hotspot from a pair of anterior and posterior images. [88] developed an automatic fine-grained skeletal segmentation method for whole-body bone scintigraphy (WBS) using atlas-based registration. The proposed method includes four steps: denoising, restoration, standardization, and registration. Experimental results show that the proposed method outperforms the traditional registration method, with a decrease in mean square error, an increase in signal-to-noise ratio, and an increase in mean structural similarity. [89] presented a comprehensive approach to automate the analysis of BM on bone scintigrams by applying deep learning technology. The study includes several models, including a classification model to identify BMs, a segmentation model to automatically delineate lesion regions, an assessment model to quantify tumor burden, and a model to generate diagnostic reports. The approach is evaluated on a dataset of patients with and without BMs and demonstrates high sensitivity, specificity, and accuracy in classifying BMs, outperforming less experienced physicians. The segmentation model achieves remarkable accuracy in identifying the spatial extent of BMs. In addition, the study introduces a novel Bone Scan Tumor Burden Index (BSTBI) to quantitatively assess the severity of BMs.

The paper [90] presents a new end-to-end multi-task deep learning model for automatic lesion detection and anatomical localization in whole-body bone scintigraphy. The authors conducted several experiments to evaluate the effectiveness of the proposed method on a dataset with 62 test volumes. They also used a 7:3 ratio for 5-fold cross-validation to evaluate the degree of overfitting. The experimental results show that the proposed architecture achieves the highest precision in the finest bone segmentation in both anterior and posterior whole-body scintigraphy images. The proposed approach

reliably balances unsupervised label generation with supervised learning and provides useful insights for weak label image analysis.

Recently, [91] proposed a Deep Learning-based segmentation method for imaging bone lesions using Swin Transformer as a decoder to extract feature information in the image. The authors conducted a literature review and found that Swin Transformer has demonstrated excellent remote information acquisition capability in various domains, including natural language processing and computer vision. The authors also compared the proposed algorithm with other segmentation algorithms in the test set under the same experimental conditions. The proposed method achieved 97.65%, 77.81%, and 35.59% of accuracy, dice coefficient, and IOU, respectively. [92] presented a simulation of the post-market fine-tuning of a computerized bone scintigram detection system and its performance analysis. The study identifies the factors affecting the performance changes and provides useful information for deriving an effective design scheme for continuous learning in artificial intelligence systems. [93] developed a CNN-based diagnostic system for automatic segmentation of BM regions in bone scintigraphy. The system was developed using a dataset of 100 breast cancer and 100 prostate cancer patients. The Double U-Net model was adapted by incorporating background removal, addition of negative samples, and transfer learning methods for BM detection. The performance of the system was examined using 10-fold cross-validation and calculated on a pixel-wise scale. The best model achieved an accuracy of 63.08%, a sensitivity of 70.82%, and an F1-score of 66.72%.

b. SPECT

In addition to bone scans, single-photon emission computed tomography (SPECT) images are also widely used for segmentation of metastatic bone lesions. The authors in [94] introduced three different algorithms, namely the K-Means clustering method, the region growth method, and the C-V model, for segmenting lesions in SPECT bone scan images. This study focuses on the region of BM lesions in SPECT imaging. The authors experimentally determined different parameters for the above three algorithms, resulting in three different segmentation algorithms. The study concludes that the proposed methods are suitable for lesion segmentation in SPECT imaging. The three methods achieve a value of 0.7307, 0.7768, and 0.8076 for the Tanimoto similarity coefficient, respectively. The segmentation method based on the C-V model is able to provide more helpful information to oncologists in the diagnosis of tumors and other related diseases. The U-Net method has been shown to be able to segment bone metastases.

In an attempt to automatically assess metastases, [95] used two deep learning models to segment hotspots in bone SPECT images based on U-Nets and Mask R-CNNs. The results

show that the segmentation models created achieve a value of 0.9920 for PA, 0.7721 for CPA, 0.6788 for Rec, and 0.6103 for IoU. [96] applied improved U- NET algorithms to segment BM on SPECT images. By adding an attention mechanism between jump links, the algorithm improves data feature selection, prevents feature redundancy, and makes the model more suitable for training. The proposed approach achieved 0.633 from MIOU and 0.571 from DSC. In addition, [97] uses U-Net as the base model while conducting research to optimize the model performance. The attention mechanism is integrated to segment BM in the pelvic region. Through a series of comprehensive experimental validations, it is shown that the synergistic fusion of the U-Net network with the attention mechanism significantly improves the segmentation of BM in the pelvic region when applied to SPECT images. Experimental results show that the integrated model exhibits significant improvements in key metrics, specifically achieving values of 0.6045, 0.7214, and 0.7564 for IoU, DSC, and Precision, respectively. Later, [98] presented a study on automatic segmentation of lung cancer bone metastases from SPECT images using UNet based models. The authors performed experiments with real clinical SPECT nuclear medicine data and proposed an improved U-net model that takes into account the residual structure and the attention mechanism. The obtained pixel accuracy PA, average pixel accuracy mPA, and average cross-merge ratio MIoU were 0.9955, 0.7824, and 0.7291, respectively.

Lin et al [99] have proposed a custom semi-supervised segmentation model for identifying and delineating lesions of skeletal metastases on augmented SPECT images. The model includes a feature extraction subtask and a pixel classification subtask. The feature extraction stage uses cascaded layers, including dilated residual convolution, inception connection, and feature aggregation, to learn hierarchical representations from low-resolution SPECT images. The pixel classification stage classifies each pixel into categories in a semi-supervised manner and delineates pixels belonging to an individual lesion using a closed curve. A DSC score of 0.6920 can be achieved when the model is trained with 37% of labeled samples.

Recently, [100] proposed a self-defined five-layer segmentation network based on the U-Net++ model for spine lytic lesion detection in CT scans. They also developed a view aggregation method to enhance the high uptake area in the image while leaving the background area unchanged. The experiment was evaluated on 260 images from SPECT, and the results showed mean values of 0.6556, 0.6257, and 0.6885 in terms of DSC, CPA, and sensitivity, respectively. [101] developed an automated segmentation model based on the U-Net++ model to identify and segment BM lesions in bone SPECT images. The authors used feature fusion and attention mechanisms to improve feature learning in important regions. The performance of the proposed model was evaluated using clinical

data with 306 SPECT images, obtaining DSC, CPA, and sensitivity values of 0.6221, 0.6612, and 0.5878, respectively. A comparative analysis with classical segmentation models highlights the superiority of the proposed approach and highlights the advantages of integrating attention gates and feature fusion to improve the overall segmentation performance.

c. CT-scan

[102] noted that noisy CT images led to misclassification of BM and some lesions could not be detected in some areas. The authors suggested that developing a segmentation method to identify the contour and location of lesion areas of BM in CT images by modifying the HED network could eliminate these problems. The proposed method achieved a TP rate of 79.8% and a IoU rate of 69.2%. [103] demonstrated the ability of the U-Net in segmenting spine sclerosis BM on CT images. The proposed CNN demonstrated high sensitivity (95% global, 92% local) and positive predictive value (97% local) with a Dice score of 0.83. This work illustrated the potential of CNN to aid in lesion detection while acknowledging the need for continued refinement and broader validation. Later, [**noguchi_2022deep**] developed and evaluated a deep learning-based algorithm (DLA) for automatic detection of BM at CT. They proposed a segmentation approach based on three convolutional neural networks (CNNs): a 2D UNet-based network for bone segmentation, a 3D UNet-based network for candidate region segmentation, and a 3D ResNet-based network for reducing false positives. The clinical efficacy of the DLA was evaluated in an observational study with board-certified radiologists. The study found that the algorithm improved radiologists' overall performance in detecting BM while decreasing interpretation time.

Recently, [104] discussed the development, training, and testing of a CNN for automatic detection of lytic lesions of the spine in chest, abdominal, and pelvic images CT. The authors manually segmented the images into three classes: lesion, normal bone, and background. The trained model yielded promising results with a mean Dice score of 0.61 for lytic lesions, 0.95 for normal bone, and 0.99 for background. Using a U-Net architecture implemented in Keras/TensorFlow, the model showed a mean global sensitivity of 90.6% for detecting lesions on a single image, 74.0% local sensitivity, 78.3% positive predictive value for detecting all lesions, and a global specificity of 63.3% for false-positive rates in nonpathologic bone. The study concluded that the CNN-based approach has high sensitivity for detecting lytic lesions of the spine in axial CT images. [105] presented a retrospective study on the development of an automated deep learning method for bone tumor segmentation and classification using CT imaging. The study used a dataset of 84 femoral scans CT with final histologic confirmation of bone le-

sions and incorporated a deep-learning architecture that predicted a segmentation mask over the estimated tumor region and a corresponding class as benign or malignant. The approach demonstrated comparable specificity (75%) and sensitivity (79%) despite unbalanced datasets, along with an average Dice score of 56% for segmentation and up to 80% for individual image slices. The study highlights the potential of machine learning tools to improve the accuracy of musculoskeletal tumor diagnosis by helping physicians decide whether a biopsy is necessary. Similarly, spectral CT (dual-energy) images have been used to identify and segment spinal metastases in lung cancer using a deep learning DC-U-Net model [106]. The energy spectral CT scan, also known as dual-energy CT or spectral CT, is a type of CT scan that uses two separate X-ray energy spectra to create two sets of data. This technique allows the study of materials with different attenuation properties at different energies.

d. PET/CT

Using a hybrid imaging modality, [107] compared the effectiveness of two approaches for segmenting bone lesions in metastatic breast cancer using the nnUnet [108] architecture. The first method used lesion annotations with images from PET and CT as two-channel input, while the second method used both reference bones and lesion masks as ground truth. The inclusion of bone masks resulted in higher precision and a slight improvement in the Dice score for bone lesion segmentation. In addition,in [107] two nnUnet segmentation models were introduced that aimed to calculate imaging biomarkers to assess treatment response based on baseline and follow-up images. Because experts typically assess treatment response using baseline and follow-up images, the models incorporated baseline images PET and lesion segmentation on baseline PET as input channels to the follow-up network. The resulting calculation of four imaging biomarkers, from both manual and automated segmentations, showed promising results in predicting response to treatment.

[109] discussed the use of deep learning algorithms for automated PET /CT lesion segmentation in oncology. They also mentioned the challenges of manually annotating tumors in whole-body scans FDG-PET and the potential of DL-based automated tumor segmentation to solve this problem. The authors used the AutoPET Challenge 2022 platform to develop and test their deeplearning models for automatic segmentation of PET lesions. They applied 5-fold cross-validation on residual UNETs to automatically segment lesions and used the results of an adaptive ensemble of highly contributing models as the final segmentation. Their method achieved a dice score of 0.5541 in the test data set (N=150 studies).

e. MRI

In a routine radiological examination to detect suspected PCa,[110] identified and segmented pelvic bone metastases using dual 3D UNET DLAs based on T1-weighted and diffusion-weighted images. After two rounds of evaluation, the DSC value for pelvic bone segmentation was above 0.85, while the AUC value for metastasis detection was 0.85, indicating high accuracy in pelvic bone metastasis detection and segmentation. In addition, [111] found that automatic segmentation of vertebral metastases on MRI was almost as accurate as expert annotation using deep convolutional neural networks (U-net-like architecture). Their proposed segmentation solution achieved a DSC of up to 0.78 and a mean sensitivity of up to 78.9% with a DSC of 0.79 for inter-reader variability. Recently, [112] focused on improving the accuracy of spinal metastasis segmentation by contrast-enhanced magnetic resonance imaging (CE-MRI) using an improved U-Net and Inception-ResNet architecture. They investigated the combination of CE-MRI, radiomics, and deep learning techniques to distinguish between spinal metastases and primary malignant bone tumors of the spine. The proposed approach combined a region-growing algorithm for lesion segmentation with improved U-Net and Inception-ResNet models. The results showed that the diagnostic accuracy of Radiomics reached 0.74, while the improved U-Net achieved an impressive average diagnostic accuracy of 0.98. Notably, the proposed model exhibited a precision accuracy of 98.001%, highlighting its potential to support the differential diagnosis of spinal metastases and primary malignant bone tumors of the spine.

Tables 2.6, 2.7 and 2.8 present a summary of all papers in the literature dealing with bone metastasis segmentation.

2.4.3 *Others ML Tasks for Bone Metastasis Diagnosis*

In addition to classification and segmentation, many works have attempted to develop new approaches for various tasks such as detection, prediction, identification, prognosis, and so on. To detect sclerotic, lytic, and mixed spinal lesions, Yao et al. [113] used an SVM classifier with sensitivities (with a false positive rate per patient) of 81% (2.1), 81% (1.3), and 76% (2.1), respectively. The same group achieved a sensitivity of 75% with a false-positive rate of 5.6 for identifying sclerotic rib lesions by using a different SVM classification system [114].

Cheng et al. [115, 116] used YOLO models to detect chest and pelvis bone metastasis lesions in scintigraphic images of prostate and breast cancer patients. The CNN-based classification models they developed achieved a mean precision of 90% for classification

Table 2.6: Summary of existing Bone Metastasis segmentation works(Part 1)

Ref.	Method	Imaging Modality	Dataset	Data Split	Evaluation Metric	Limits
[81]	Atlas-based segmentation Intensity normalization Anatomy-specific thresholding	Bone Scan	40 2-D bone scans	Leave-one-out cross-validation	Sens= 95.5% Spec = 93.9%	Small dataset Limited to metastatic prostate cancer patients
[82]	Fuzzy C-Means	Bone Scan	12 patients	4 clusters for the C-Means	high sensitivity	Lack of Standardization in Images
[83]	CNN and Autoencoder	Bone Scan	1030 images	5-fold cross-validation Test = 68 thoracic images 572 hot spots	Jaccard= 80.51% DSC= 88.87%	Conducted patient-wise classification Not carry out skeleton segmentation Not perform hotspot detection
[84]	BtrflyNet	Bone Scan	246 patients	3-fold cross-validation:Train=164 Validation=41; Test=41	Dice coefficient = 0.8756	misclassify hot spots of bone limited amount of training data
[85]	Constrained Local Model (CLM)	Bone Scan	20 patients; 20 images	5 fold cross validation	error=0.0.283	overfitting due to the small dataset
[86]	MaligNet	Bone Scan	1088 labeled (chest images) 18,560 unlabeled images 9,280 patients	Train=15.527 train Val=4005 Test=116	Prec= 0.852 Sens =0.856 F1 =0.848	Detection of metastasis only in chest area
[87]	Ensemble ResBtrflyNet(AP) ResBtrflyNet(LR)	Bone Scan	665 patients 330 scans	3 fold cross-validation	False positive regions= 2.13 for anterior 2.62 for posterior scintigrams	Biased data(prostate cancer)
[88]	atlas-based registration	Bone Scan	600 images	N/A	Mean square error = 0,0021 Peak signal to noise ratio = 26.92 Mean structural similarity = 0.9998	Use of specific dataset
[89]	Resnet34 Unet	Bone Scan	12 patients	Train=80% Test=20%	DSC = 0.7387	Small dataset with potential biases Missing clinical workflow integration discussion
[90]	End-to-end multi-task DL model	Bone Scan	617 patients	5 fold cross validation	DSC=89.20 Prec = 81.32 Sens= 90.67	Small Test Dataset Limited Generalizability Analysis
[91]	Swin Transformer	Bone Scan	242 images	Train=230 Test=12	Acc=97.65% dice coefficient=77.81% IoU= 35.59%	Limited Dataset Computational Resource Demand

Table 2.7: Summary of existing Bone Metastasis segmentation works(Part 2)

Ref.	Method	Imaging Modality	Dataset	Data Split	Evaluation Metric	Limits
[92]	ResBtrflyNet	Bone Scan	1032 patients from 5 hospitals	Train = 692 Val= 170 Test= 172	GFI= 0.862	findings not be applicable to other medical images
[93]	Double Unet	bone scan	200 patients 100 breast cancer 100 prostate cancer	Train=80% Val=10% Test=10%	Prec = 63.08%, Sens= 70.82% F1=66.72%	Lesion vs. Normal Discrimination Challenge Limited Dataset Size and Generalization
[110]	3D UNET	MRI	859 patients	Train=683 Val=88, Test=88	DSC and VS values above 0.85 the HD values were <15 mm	dataset specificity small external evaluation potential labeling variability
[111]	Unet	MRI	40 patients	8-fold cross-validation	DSC= 77.6% Sens= 78.9%	Limited dataset size and Single epoch evaluation Narrow focus on spinal metastases Lack of Comparative Analysis
[112]	Improved U-Net Inception-ResNet	MRI	81 patients 81 scans	10 fold cross validation	PA=98.001% IoU=96.819% DSC =98.384%	Small Dataset Lack of diversity of samples
[94]	K-means region growth C-V model	SPECT	120 images	N/A	Tanimoto= 0.7307 for the K-means Tanimoto=0.7768 for region growth Tanimoto=0.8076 for C-V model	Specific Lesion Focus Missing Computational Analysis Limited Generalizability Information
[95]	U-Net Mask R-CNN	SPECT	112 images (2.280 augmented) 76 patients	Train=1830 Test= 450	Acc= 0.9920,CPA= 0.7721 Sens= 0.6788, IoU=0.6103	Only Thorax Bone region
[96]	Unet -attention	SPECT	125 images	Train= 80% Test=20%	DSC=0.571	Only images of thoracic bone
[97]	Att-U-NET	SPECT	121 images 2390 sheets(data augmentation)	Train=2112 Test= 275	IoU=0.6045 DSC= 0.7214 Prec= 0.7564	Segmentation Focus on Pelvic Bone Metastases
[98]	UNet	SPECT	306 diagnostic records 176 patients	214 for Training 92 for Testing	mPA = 0.7824 PA = 0.9955 MIoU = 0.7291	Small Dataset segment thoracic region only lack of comparison with other segmentation models
[99]	semi-supervised model	SPECT	724 whole-body images 362 patients	112 used images Train =70% Test=30%	DSC= 0.692	small dataset

Table 2.8: Summary of existing Bone Metastasis segmentation works (Part 3)

Ref.	Method	Imaging Modality	Dataset	Data Split	Evaluation Metric	Limits
[100]	model based on UNET++	SPECT	130 patients 260 images	10-fold validation Train=70% Test=30%	DSC=0.6556 Sens= 0.6257 CPA= 0.6885	Focused on segmenting metastatic lesions caused by lung cancer
[101]	AG Unet++	SPECT	306 images	Train set (n = 214, 70%) Test set (n = 92,30%)	DSC=0.6221 CPA = 0.6612 Sens= 0.5878	Small Dataset Omitted comparison with other segmentation models
[102]	modified HED network	CT	21 patients 250 images	5-cross folding	TP rate=79.8% IoU= 69.2%	Limited Dataset Comparative Analysis Gap
[103]	Deep CNN	CT	600 images	Train=90 Test=10	DSC=0.83 for lesion DSC=0.96 for non-pathologic bone and 0.99 for background	small dataset
[**noguchi_2022deep**]	deep learning–based algorithm (DLA)	CT	732 patients	Train= 632 Val=50 Test=50	Sens = 82.7% False positives= 0.617 Mean interpretation time = 85 s	Small Dataset
[104]	UNET	CT	600 images	540 train 60 Test	DSC = 0.61 for lytic lesion DSC =0.95 for normal bone DSC = 0.99 for background	Model tested on healthy controls focused on lytic spinal lesions
[105]	Masked-attention Mask Transformer (Mask2Former)	CT	84 femur CT	5-fold cross-validation Train= 80% Test= 20%	DSC= 56 %	Focus only on femur CT scans Small Dataset
[106]	DC-U-Net	Energy/spectral CT	180 images 36 patients	Train= 135 Val=45	Rate detection= 81.41% Rate detection= 66.03%	Targeted Patient Group Limited Generalizability Restricted Comparative Analysis
[107]	nnUnet	PET/CT	24 patients	3-fold cross-validation	DSC= 0.61	Small dataset
[109]	residual 3D Unet	PET/CT	900 subjects	5-fold cross-validation	DSC = 0.554	Mismatch between predicted and segmented regions

DSC= Dice similarity coefficient, Acc= Accuracy, Sens=Sensitivity, Spec= Specificity,Prec=Precision F1= F1-score, N/A=not applicable,GFI=The goodness of fit index, Tanimoto= Tanimoto coefficient similarity, Val= validation, MIoU= average cross-merge ratio, PA= Pixel Precision, CPA= Class Pixel Accuracy, Jaccard= Jaccard Similarity Coefficient

of detected lesions in the chest [115] and a precision of 70% (81%) for classification of detected lesions in the chest (pelvis)[116].

The authors in [117] have proposed an algorithm to reduce false positives in automatically detected hotspots on bone scintigraphy images. The algorithm is designed to work in a semi-supervised manner, as the construction of a fully annotated database is challenging. The classification sensitivity, specificity, and false-negative rates achieved were 63%, 58%, and 37%, respectively.

In a study by Fan et al. [118] on the differential diagnosis of spinal metastases using 18F-FDG PET /CT, there were 51 texture parameters that meaningfully differentiated between benign and malignant lesions, four of which had AUC higher than SUVmax. The texture parameters were used to build a classification model with logistic regression, support vector machine, and decision tree. The accuracy of manual diagnosis was 84.27%.

2.4.4 *Challenges and Limitations in Previous Works*

Over the past decade, numerous articles have been published on the topic of bone metastases, focusing on lesion detection, segmentation, and classification in various imaging modalities. However, there are still some open research areas with few studies. In these open research areas, there are several challenges that should be addressed in future research.

- **(1) Absence of common dataset**

  One of these challenges is benchmark datasets. The results of previous studies are based on their own datasets. However, we believe that it is not straightforward to evaluate and numerically compare different studies based solely on their reported results because they use different datasets, evaluation methods, and performance metrics. To compare studies numerically, it is definitely necessary to develop benchmark datasets. These datasets should consist of samples obtained from a large number of patients and annotated by different radiologists. This would allow a numerical comparison of the results of the different studies and identify the distinguishing features.

  In ML, models are trained to make predictions based on known datasets, from which the machine then “learns”. After developing a model, it applies its knowledge

to perform diagnostic tasks on unknown datasets [119]. Application of ML requires collection of data labeled by experts (usually radiologists) or direct data extraction through various computational methods.

As far as we know, we have found only one open-source dataset: BS80K [120]. The BS -80K benchmark dataset consists of bone scan images that can be used to evaluate the performance of new advanced computer vision methods. The data includes anterior view (ANT) and posterior view (POST), along with 13 region-wise slices for vulnerable body parts. Tasks include detection, classification, and tracking. This dataset illustrates that working with common professional digital image databases allows us to overcome any bias in the tested images and enable accurate publication. We all still have a long way to go to achieve clinically acceptable results with the best results, because for images of BS -80K they retained only 0.2484 of the average detection accuracy of about 3247 original candidates.

- **(2) Small scale**

  The datasets used in existing studies are generally not large enough, barely exceeding 1000 images [107, 106, 105, 101, 93, 87, 74, 52]. Obviously, such a small amount of data is far from sufficient to build a generalized and robust ML model.

  One of the major challenges of using such small datasets is the high risk of overfitting. Overfitting occurs when a model is fitted too closely to the specific features and noise in the training data, such that it memorizes those features instead of learning the underlying patterns. This can result in a model that performs exceptionally well on the limited training data, but fails to generalize to unseen data, rendering it virtually useless in real-world applications. To address the problem of overfitting and increase the amount of data used, several solutions have been proposed, such as using data augmentation techniques [97, 91] or transfer learning [121, 122]. While data augmentation and transfer learning are valuable tools in the ML toolbox, dealing with overfitting often requires a holistic approach that combines multiple techniques, data quality improvements, and ongoing model maintenance to achieve robust and reliable AI systems.

  In addition, small datasets also limit the diversity and quality of training data. ML models thrive best when confronted with a variety of examples that include different scenarios, conditions, and variations. A small dataset cannot adequately represent the full range of real-world complexity, which is critical for developing robust models that can handle multiple situations and variations.

- **(3) Non generalization**

  Most of the current BM with ML works suffer from the nongeneral problem, i.e., they contain only images derived from a specific cancer type, a specific body region, or a specific metastatic bone lesion. For example, the datasets used in [63, 98] contain only bone scan images of the thoracic spine. In [97], the dataset used contains bone scans of the pelvis with annotations indicating only the presence or absence of metastases, which is obviously insufficient to provide an accurate and comprehensive diagnostic aid. In addition, the study dataset in [68] is limited to patients with lung cancer and thoracic bone metastases, which may limit the generalizability of the model to broader cancer types or anatomic regions. The authors in [104] focused exclusively on lytic spinal metastases and did not consider other types of lesions or organs, which could limit the applicability of the proposed method to other medical imaging tasks.

  Other problems related to standardization and experimental methods include (i) different scanners used for image acquisition having different image resolution and pixel size, (ii) different staining characteristics, (iii) different illumination conditions, (iv) different magnification levels, (v) different number and size of images (frames, whole slide images, and regions of interest are examples of the types of images used by different authors).

  In addition, several works have used radiomics as a feature extraction technique from medical imaging [78, 73, 77, 74]. Although radiomics has demonstrated its potential in several applications, it encounters several limitations [123]. First, radiomic feature values are affected by patient variability, imaging artifacts, and image reconstruction algorithms. Poor data quality can lead to inaccurate radiomic features, which can affect the reliability and validity of the analysis. Therefore, most studies lack external validation and require further evaluation. Second, radiographic studies are often not reproducible due to lack of standardization, inadequate reporting, and limited open-source code. The lack of adequate validation and the possibility of false-positive results also hinders clinical implementation. In addition, a large number of radiomic features can complicate interpretation of results for specific clinical questions. Further research is also needed to better understand the biological mechanisms underlying the observed radiomic changes.

  Another limitation of the works on classification, segmentation, and detection of BM is that they have not been compared with other state-of-the-art methods. This means that the results of these works may not be as reliable as they could be because they were not compared to other methods that may have been more

accurate. Comparing the results of different methods can help identify the strengths and weaknesses of each method and provide a more comprehensive understanding of the problem. Therefore, future studies should consider comparing their methods to other modern methods to ensure the accuracy and reliability of their results.

In response to these limitations, we have taken the initiative to create a new dataset for BM and make it available to the broader research community. This proactive step aims to foster collaboration, improve transparency, and facilitate the development of a more robust and universally applicable computer-aided diagnostic system for BM.

However, the challenges in ensuring the generalizability of these CAD systems to broader patient populations and diverse medical centers, as well as the lack of transparency in the decision-making process, raise important questions about their broader clinical applicability. Further research and validation efforts must be undertaken to address these limitations and ensure the reliability and practicality of the proposed automated analysis framework in a clinical setting.

To improve the performance and generalizability of these systems, collaboration with other medical centers is urgently needed. This collaboration will allow more diverse and comprehensive data to be collected, ultimately leading to more effective CAD systems.

To advance the field and facilitate the adoption of these cutting-edge technologies into clinical practice, future research efforts should prioritize the inclusion of comprehensive comparative assessments. This approach will contribute to a deeper understanding of the current state of the art and promote advances in the field of bone metastasis diagnosis and treatment.

2.5 CONCLUSION

In this chapter, we have presented the current state-of-the-art classification, segmentation, and detection techniques for BM, highlighting both their promising results and existing limitations. Our discussion highlighted the importance of addressing challenges related to dataset availability, model generalizability, and transparency in clinical applicability. In the next chapter, we will explore the data aspect of this area in more details by examining datasets commonly used in research from BM. In addition, we are pleased to present our own carefully developed BM dataset tailored to improve research in this area. We will then provide an overview of the datasets used in this research area and introduce our BM developed dataset. Then we will present a BM segmentation system.

3

BONE METASTASIS PROPOSED DATASET

Contents

3.1 INTRODUCTION

As we mentioned in the previous chapter, most of the machine-learning approaches used in the BM field are based on supervised learning. Accordingly, datasets including labeled ground-truth are needed in order to train such methods, so, to develop bone metastasis detectors. This chapter summarizes numerous published real-world datasets regarding imaging modality, their size, and labels. Consequently, we detail our proposed BM dataset and the challenges related to databases in the field of BM.

Section 3.2 refers to an exhaustive overview of the database concept in the medical imaging field. In section 2.3, the most commonly used imaging modalities in the field of BM are discussed. After that, available private and public BM databases are detailed in section3.3. Finally, our proposed dataset for BM is introduced in section 3.4.

3.2 BACKGROUND

Public datasets play a vital role in driving advancements in autonomous medical diagnosis, particularly in the field of medical imaging. Over the years, numerous datasets have been generated, enabling the development and evaluation of computer-aided medical diagnosis systems (CAD). While the use of datasets for radiological decision-making can be traced back to the 1960s [124], recent breakthroughs in algorithm development and computational resources have paved the way for the application of artificial intelligence (AI) and deep learning techniques in this domain.

In the realm of medical imaging, datasets have been instrumental in addressing a wide range of diseases and conditions. For instance, in the field of Alzheimer's disease research, datasets containing neuroimaging scans have been essential for studying the progression and early detection of the disease [125]. These datasets enable researchers to analyze patterns and biomarkers associated with Alzheimer's, facilitating the development of more accurate diagnostic tools.

Breast cancer, another significant area of study, has also benefitted from the availability of diverse datasets [126, 127]. These datasets include mammograms, ultrasound images, and histopathological data, enabling researchers to investigate different aspects of breast cancer detection, classification, and treatment. Such datasets have helped in the development of CADx systems that assist radiologists in accurately identifying and diagnosing breast cancer, leading to improved patient outcomes. Similarly, datasets in

the field of COVID-19 have played a crucial role in advancing radiological diagnosis [128, 129]. CT scans and radiographs, combined with clinical and demographic information, have formed the basis for developing computational models that aid in COVID-19 detection, lesion classification, and treatment planning. These datasets have significantly contributed to the development of automated systems that assist radiologists in making accurate and timely diagnoses, ultimately enhancing patient care.

While several medical sectors have benefited from the availability of robust datasets, it is important to acknowledge that some fields of medicine still face challenges due to the scarcity of data. One such example is bone metastasis [120, 130], which involves the spread of cancer from primary tumors to the bone and suffers from a lack of large-scale, diverse datasets. The scarcity of data poses obstacles to developing accurate imaging-based tools for detecting and characterizing bone metastasis, impacting patient management and outcomes.

The availability of public datasets has been instrumental in driving advancements in autonomous medical diagnosis, particularly in the field of medical imaging [131, 132]. These datasets help in improving the automatic detection of diseases, creating study protocols, improving image quality, decreasing radiation dose, decreasing MRI scanner time, and optimizing staffing and scanner utilization, thereby reducing costs, and offering the possibility of performing expensive and time-consuming screening programs in countries that otherwise cannot afford them [133]. Access to big data of medical images is needed to provide training material to AI devices so that they can learn to recognize imaging abnormalities. The absence of adequately annotated extensive datasets for the education of artificial intelligence algorithms in domains such as bone metastasis introduces obstacles that necessitate resolution through cooperative endeavors and initiatives for data exchange.

3.3 EXISTING BONE METASTASIS DATABASES

Bone metastasis is a common condition characterized by the spread of cancer cells from primary tumors to the bones. It is very necessary to diagnose it at a very early stage, therefore, the research community has introduced various datasets and methods based on deep learning for the identification of bone metastasis. These datasets can help in the early identification and characterization of BM by extracting significant features. Using these extracted features, different CADs can be performed to improve the accuracy and efficiency of diagnosing bone metastases, leading to rapid interventions and better patient

outcomes. In this section, we review various BM databases by detailing and observing the characteristics of each one. Table 3.1 exhibits an overview of various BM datasets.

Papapandrios and his team have conducted a series of studies aimed at proposing different datasets related to bone metastasis. In their first work [52], they utilized a dataset consisting of 586 consecutive whole-body scintigraphy images of men, sourced from 507 different prostate cancer patients. These images were carefully selected and diagnosed by a nuclear medicine specialist. Out of the 586 bone scan images, 368 were from male patients with bone metastasis, while 218 were from male patients without bone metastasis. The nuclear medicine physician categorized all the patient cases into two categories: 1) metastasis absent and 2) metastasis present. Building upon their initial work, in their second study [49], Papapandrios and his team aimed to expand the dataset by including additional bone scan images. They retrospectively reviewed a total of 778 planar bone scan images from patients with known prostate cancer. Of these, 328 bone scans were from patients with bone metastasis, 271 were benign cases with degenerative changes, and 179 were from normal patients without bone metastasis. The nuclear medicine physician classified these cases into three categories: healthy, degenerative, and malignant. In another endeavor to gather more data on bone metastasis, Papapandrios, and his team [51]collected information specifically related to breast cancer as the primary tumor. They conducted a retrospective review of 422 consecutive whole-body scintigraphy images obtained from 382 different breast cancer patients (women). Out of the total 408 bone scan images the nuclear medicine specialist selected, 221 were categorized as malignant and 187 as benign, without bone metastasis.

In [46], Pi et al. utilized a vast dataset comprising more than 15,000 bone scan images from 13,811 patients to investigate the identification of bone metastasis. The dataset included 9595 benign diagnoses and 5879 malignant cases. The patients in the dataset consisted of 6699 males and 7112 females. However, currently, only the validation subset with limited access is available.

The dataset used in [86] consisted of medical data from a total of 9,824 patients. The dataset was utilized for two main purposes: chest detection and lesion instance segmentation. For chest detection, 680 whole-body images were used for training, with 200 for validation and 240 for testing. In the case of lesion instance segmentation, the dataset consisted of 19,648 chest images, of which 1,088 were labeled and the remaining 18,560 were unlabeled.

Cheng et al. [115] conducted a study on BM identification using 524 whole-body bone scans collected from China Medical University Hospital. The dataset included 194

scans from prostate cancer patients and 371 scans from breast cancer patients. However, the study suffered from insufficient data as only 524 bone scan images were utilized.

The datasets conducted by Shimizu et al. [84], Cheng et al. [115], and Aslantas [42] were characterized by their relatively small size with 130,246,524 bone scans respectively.

Using the same imaging modality as previous works, bone scan, BS-80K [120] emerged as a significant breakthrough in the field of datasets, addressing the lack of publicly available resources until November 2022. This dataset consists of a vast collection of 82,544 bone scan images obtained from 3,247 patients at West China Hospital. Each patient contributes two whole bone scan images, representing the anterior view (ANT) and posterior view (POST), along with 13 region-wise slices for susceptible body parts. Expert specialists meticulously annotated the images using an authorized labeling criterion, providing accurate labels for bone metastasis presence. Furthermore, multiple bounding boxes containing suspectable hot spots and their corresponding annotations are included within each whole body image. The availability of BS-80K has greatly facilitated research in bone metastasis, offering a comprehensive dataset that aids in algorithm development and analysis for improved detection and understanding of this medical condition.

In addition to bone scans, the previous works in the field incorporated various imaging modalities including CT, SPECT, and PET/CT. Using CT scans, Samira et al. conducted a study in which they utilized 2,880 annotated bone lesions from CT scans of 114 patients diagnosed with prostate cancer. Among these patients, 41 had histopathologically confirmed metastatic bone lesions. Additionally, another study [1] collected a total of 269 positive CT scans, consisting of 1,375 bone metastases, and 463 negative scans for lesion segmentation purposes. Despite the limited number of patients and scanners available, their study's authors in [95, 100] used SPECT, a medical imaging technique, to detect bone metastases. While in [107], authors used PET/CT images from 24 breast cancer patients to detect and segment automatically bones and metastatic bones.

The analysis of existing work in the field of computer-aided diagnosis highlights several challenges regarding BM datasets that hinder further research progress. These challenges can be summarized as follows:

- Lack of a standardized dataset: Despite the promising results achieved by individual studies using local datasets, the absence of common datasets creates obstacles in terms of communication, mutual learning, and data sharing among researchers.

Table 3.1: Summary of existing Bone Metastasis Datasets

Dataset	Data Modality	Dataset description	Availability
[52]	Bone Scan	586 scans [368 BM;18 Normal] 507 patients	Private
[49]	Bone Scan	778 scans [328 BM; 271 degenerative; 179 normal] 817 patients	Private
[51]	Bone Scan	408 scans [221 BM; 187 Normal] 382 patient	Private
[46]	Bone Scan	15,474 images 13.811 patients [9595 benign; 5879 malignant]	Partly available with restricted access
[86]	Bone Scan	1088 labeled chest ; 18.560 unlabeled 9.280 patients	Private
[57]	Bone Scan	9133 images [2991 BM; 6142 Normal] 5342 patients	Private
[42]	Bone Scan	130 images [30 Normal;100 BM] 60 patients	Private
[84]	Bone Scan	246 bone scans	Private
[115]	Bone Scan	524 scans 524 patients	Private
[120]	Bone Scan	82544 images 3247 patients	Open Access
[72]	CT	2.880 scans 114 patients [41 BM]	Private
[1]	CT	269 BM scans 169 patients	Private
[107]	PET/CT	24 patients	Private
[100]	SPECT	260 scans 130 patients	Private
[95]	SPECT	76 patients 112 scans	Private
[130]	**CT**	**102,614 images** **57 patients**	**Open Access**

This limitation also reduces the reliability of research contributions, potentially leading to a stagnation of advancements in this particular area.

- Limited dataset scale: The datasets employed in current studies are typically small in scale, with the number of images hardly exceeding 1000 [95, 100, 107, 1, 115, 84]. Such a small amount of data is inadequate for building generalized and robust machine-learning models that can effectively address the complexities of BM detection.

- Lack of generalization: Many existing BM image datasets suffer from a lack of generalization, as they primarily consist of whole-body or specific-body region images or data from a specific primary tumor. For example, the dataset used in [86] only includes bone scan images of the thoracic spine. In [46], the dataset comprises solely whole-body bone scan images, and their annotations merely indicate whether metastasis is present or not. Such limited information hampers the accuracy and comprehensiveness of diagnostic assistance.

To address these challenges and advance the field, we took the initiative to develop a new BM dataset. This dataset was meticulously curated to include a comprehensive range of imaging modalities, including CT-scans, to leverage their distinct advantages. Furthermore, we have made this dataset publicly available to the research community, ensuring that researchers from diverse backgrounds can access and utilize this resource to foster advancements in the BM field. By providing this new dataset, we aim to facilitate the development of more accurate and robust models while promoting collaboration and knowledge sharing within the research community. With sufficient and manifold data, we believe the open-accessed BM-Seg dataset will galvanize the wide research about computer-aided CT-scan image analysis in the field of bone metastasis.

3.4 COLLECTED DATASET

As mentioned in the previous section, the extent of research efforts in the BM field emphasizes the need for datasets that explore complementary images. Consequently, in this section, we introduce our proposed dataset and provide comprehensive information about its implementation, along with the associated challenges and opportunities.

3.4.1 *Imaging modality*

As previously mentioned in section 2.3, several imaging modalities are being used in Bone metastasis detection such as bone scan, MRI, SPECT, and CT. CT-scan was selected as the imaging modality for this study due to the following reasons:

- CT-Scan is a global modality and the first examination physicians need to get a clear overview of primary and metastatic cancers.
- Because radiologists know the specific types of primary cancers, they can focus their examinations on specific areas of the bone to detect lesions.
- CT-Scan can provide the right tradeoff between cost, time, and resolution.
- CT-Scan allows radiologists to visualize different types of metastases in a single scan.

Considering the above factors and the fact that images from CT-scan are the most accessible modality at Hedi Chaker Hospital, Sfax, Tunisia, we have chosen this modality in mutual agreement with the radiologists.

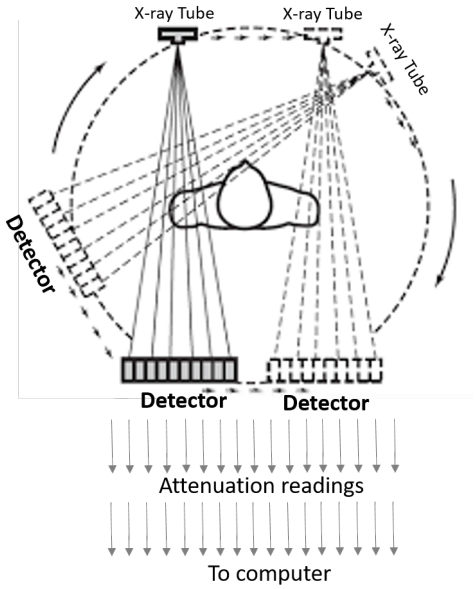

Figure 3.1: Computed tomography imaging acquisition process

CT-Scan is a medical imaging technique for obtaining reconstructed cross-sectional images based on measurements of X-ray beam attenuation coefficients in the volume

under consideration [134]. It uses X-rays and detectors to measure how different tissues block the X-ray beams. By gathering these measurements from various angles, the elemental surfaces of the image are reconstructed from the data projected onto a reconstruction matrix [135]. With the help of a calculator, a more or less important tonality is assigned to these elemental surfaces according to attenuation coefficients. The X-ray scanner analyzes how X-rays are absorbed as they pass through the body section being examined [136].

3.4.2 *Data Aquisition*

A thoracic-abdominal-pelvic CT scan was systematically performed, including the entire spine, pelvis, and proximal extremities of all four limbs. Patients underwent CT scans using a Siemens 64-bar unit in the Radiology Department of Hôpital Hedi-Chaker Sfax.

3.4.3 *Labeling Process*

The labeling process for our dataset begins with the collection of CT scans from each patient at the Hedi Chaker Hospital in Sfax, Tunisia. These CT scans serve as the primary source of medical imaging data for our dataset.

After obtaining a series of high-quality CT scan images from each patient, we first transform the DICOM images into the JEPG format for two primary reasons. Firstly, the DICOM image always contains sensitive patient-related information, necessitating a crucial de-identification process through transformation. Secondly, compared with the DICOM format, the standard JEPG is more convenient as it offers enhanced convenience for compatibility with various image-processing libraries including OpenCV.

All CT-scans were collected retrospectively from UHC (University Hospital Center) Hedi Chaker, Sfax, Tunisia. The data had been collected from November 2020 to June 2023. Each CT-scan is reviewed by three expert radiologists using the Radiant Dicom Viewer tool [137]. Based on the location of bone pain, the patient's health, and any prior trauma history, two doctors annotated the slices of each CT-scan as infected or not infected. Then, an expert radiologist (more than 20 years of diagnostic imaging experience) supervised the annotated slices that had been originally evaluated by the two other doctors. Finally, the radiologists selected the regions that correspond to BM lesions in each CT-scan infected slice. This process involves a meticulous review of the

CT images, slice by slice, to identify any signs of bone metastasis or other abnormalities. The examination was excluded if there was no clear diagnosis agreement between the three radiologists. Using the annotations we receive from doctors, we further organize the CT scan images based on that information. For each CT scan image, its label is determined by whether there is a metastatic lesion within it or not. The image is classified as abnormal if it contains at least one lesion, otherwise, it is classified as normal.

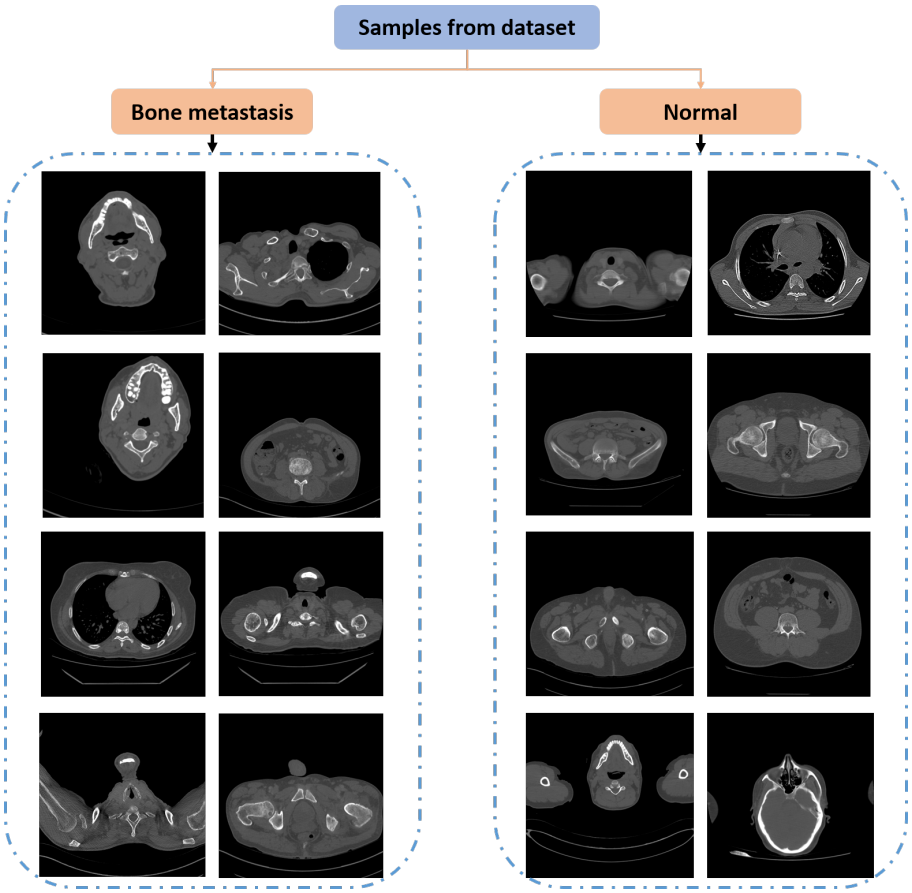

Figure 3.2: Some of the images from the proposed dataset

Once the labeling process is completed, the labeled CT scan patients are organized and categorized into two main folders. The first folder contains CT scan images of healthy patients who do not show any signs of lesions in any of their slices. These patients are classified as "Normal". The second folder contains patients with metastatic lesions that belong to an abnormal class (bone metastasis). There are two subfolders in this folder, one for normal slices and one for metastatic slices. Through this subdivision, it is possible to distinguish CT scan images of normal bones from those that depict bone metastases clearly. Some images from our dataset are presented in section 3.2.

3.4.4 *Dataset Description and Statistics*

The primary objective of our dataset is to train and evaluate a model for metastatic bone detection specifically to differentiate between normal and metastatic bone based on the CT-scan images. The proposed dataset consists of CT scan images obtained from 46 patients. More specifically, except in a few cases where relevant information is uncertain, the dataset includes 32 males with a mean age of 59.03 ± 15.19 and 25 females with a mean age of 53.55 ± 13.06. The age and gender distributions of the patients are visually represented through intuitive histograms and statistical data in Figure 3.3.

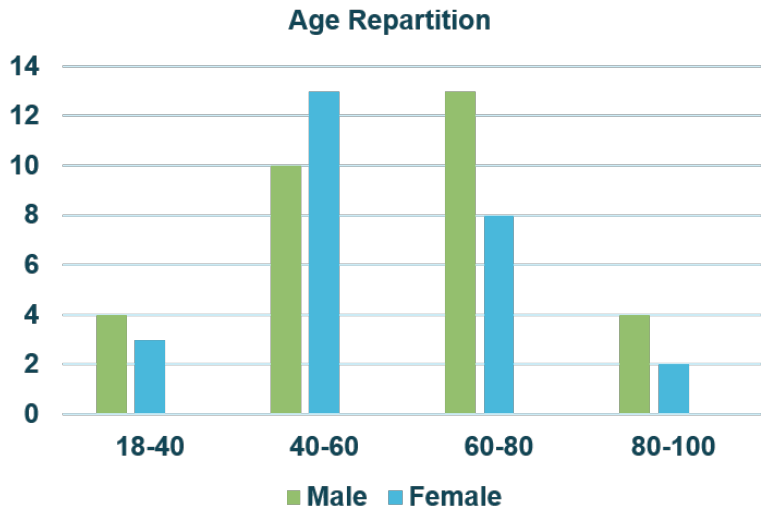

Figure 3.3: Gender and age distribution in our proposed dataset

Our dataset, available for analysis, encompasses two distinct groups: 26 healthy individuals and 31 patients with metastatic bone conditions. The healthy group contributes a total of 40,257 slices to the dataset, while the metastatic bone group provides a substantial set of 62,357 CT-scan slices. In total, our dataset comprises 102,614 CT-scan slices. It is noteworthy that this dataset has the potential for easy augmentation with additional slices from healthy patients, further enhancing its scope and utility. It's worth noting that the patients in this dataset were treated as separate and distinct entities during the analysis process. This means that the data from each patient was kept independent of the data from other patients, ensuring that the scans from each patient were considered unique instances.

As we previously discussed in Chapter 2, the bone lesion can be the result of multiple primary tumors such as lung, breast, prostate, and liver. Furthermore, cancerous cells

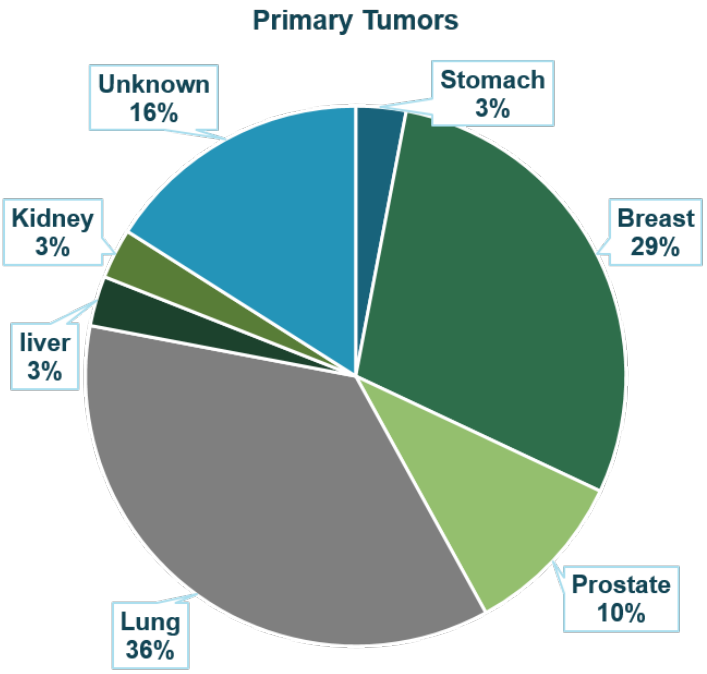

Figure 3.4: Primary tumors distribution in our proposed dataset

can propagate in various sites in the bone such as the pelvis, spine, or ribs. To gain deeper insights into the characteristics of metastatic bone patients in our dataset, we compiled a summary of primary cancer types in figure 3.4. Although the majority of patients have lung tumors as their primary cancer, the vertebrae represent the bone area most commonly affected by lung tumors, making spinal sections the most common. In addition, our dataset followed a real distribution without excluding cases, unlike previous studies that excluded misleading examples and focused on only one type of primary cancer. Therefore, it is expected that a trained system using our dataset would be more suitable for routine clinical applications.

Currently, we are still working on collecting more data to enlarge our dataset. Part of our dataset, which we have already collected, has been published as a benchmark in the field of bone metastasis segmentation under the name BM-Seg. This benchmark dataset serves as a standardized reference for evaluating and comparing different algorithms and methods for accurately segmenting bone metastases. It includes a diverse range of CT scan images annotated with ground truth labels for the presence and location of bone metastases. In Section 3.4.6, we will provide a more detailed description of the BM-Seg benchmark dataset, including its composition, annotation guidelines, and evaluation

protocol. We will also discuss the significance of this benchmark in advancing the field of bone metastasis segmentation and its potential applications in clinical practice.

Finally, We are planning to publish a second version of our dataset that will cover more types of patients and can be used for many tasks. Consequently, this dataset will provide a substantial amount of CT-scan data from both healthy individuals and patients with bone metastasis, enabling researchers and practitioners to develop and test algorithms for automatic bone abnormality segmentation, detection, and classification.

3.4.5 *Dataset challenges*

An expert radiologist must look at clinical and anatomical contexts to evaluate bone metastasis from CT-scan. The most accurate assessment comes from comparing and contrasting CT-scan with other medical imaging modalities such as MRI and SPECT scans. The main challenges that face the radiologist when examining the CT-scans are:

- Bone lesions can occur in various regions and sites of the body, including the ribs, pelvis, and spine. In order to accurately identify and distinguish these bone lesions, healthcare professionals need to thoroughly examine CT scans. This comprehensive analysis ensures a meticulous examination and enables the detection of lesions that may be present in different areas and locations of the skeletal structure.

- BM lesions are classified as osteoblastic, osteolytic, and mixed [138]. While osteolytic lesions are responsible for bone resorption and show a lower density than the normal cancellous bone, osteoblastic lesions cause too many bone cells to form, which makes the bone very dense. According to [7], it is challenging and time-consuming to detect bone lesions at an early stage on CT-scan images especially when a variety of benign bone lesions with an osteolytic appearance are present. As a result, each lesion type has a different appearance, which complicates the annotation process and makes the segmentation task more challenging.

- Benign processes such as inflammation, degenerative activity, fracture, and injury will show hotspots. In addition, primary malignant and benign bone tumors will show similar appearances, i.e. increase in radiotracer uptake.

To overcome the above challenges, annotation by different radiologists is essential to obtain redundant information or complementary data that can compensate for the challenges present in each CT-scan. Figures 3.5 and 3.6 highlight the most introduced challenges in

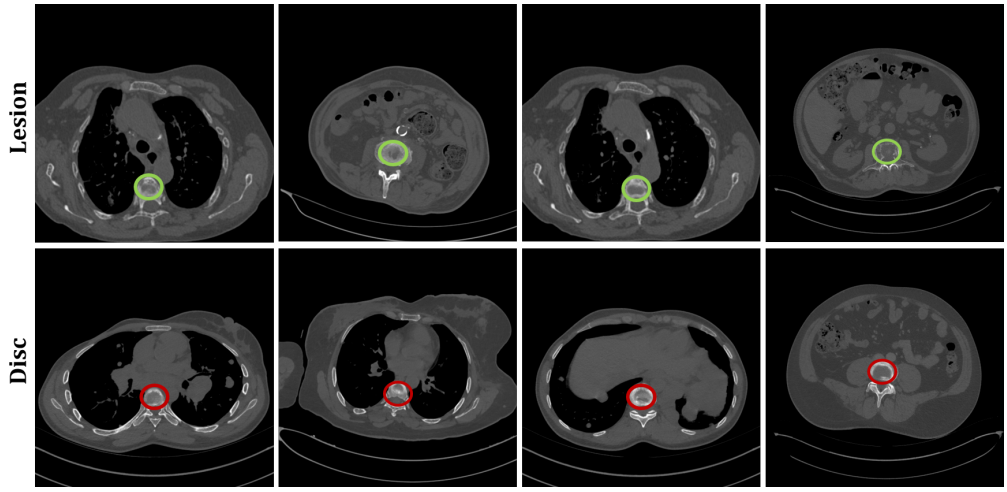

Figure 3.5: Osteolytic lesion appearance (circled in green) vs Disc appearance (circled in red) in our dataset

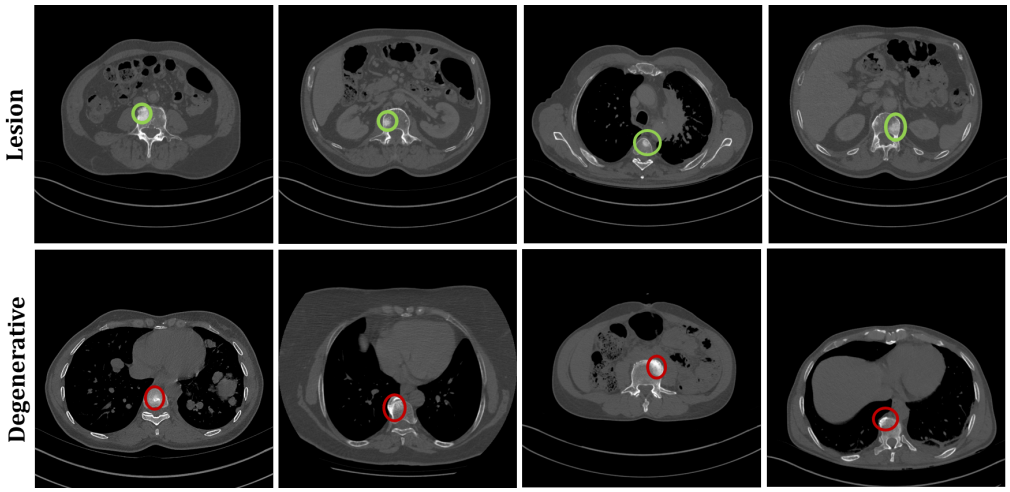

Figure 3.6: Osteoblastic lesion appearance (circled in green) vs Degenerative bone appearance (circled in red) in our dataset.

our dataset. As shown in Figure 3.5, it is difficult to distinguish between disc in CT-scan and bone lesion, especially osteolytic lesions where radiologists have to spend considerable time verifying one patient scan. Figure 3.6 shows an additional issue in metastatic bone labeling which is the difficulty to differentiate between osteoblastic lesions and degenerative bones as they have similar appearances.

3.4.6 *BM-Seg Dataset*

In the previous section, we mentioned that the BM-Seg is part of our dataset dedicated to segmentation purposes. Therefore, in this section, we present our proposed dataset BM-Seg and we detail its label processing and data preparation process.

3.4.6.1 *BM-Seg Description*

The BM-Seg benchmark dataset [130] is a valuable resource for researchers and practitioners in the field of medical image analysis. It allows them to assess the performance of their segmentation algorithms and techniques against a common set of evaluation metrics. By using this benchmark, researchers can identify their approaches' strengths and weaknesses and contribute to advancing bone metastasis segmentation methods.

As a part of our main dataset, the BM-Seg dataset consists of CT-scans from 23 metastatic bone patients with 9 female and 14 male subjects, ranging from 18 to 83 years old. All CT-scans were collected retrospectively from UHC (University Hospital Center) Hedi Chaker, Sfax, Tunisia. The data had been collected from November 2020 to January 2023.

Table 3.2: Types and incidences of primary tumors among CT scan examinations in our dataset

Primary Tumors	Type	Lung	Breast	Kidney	Prostate	Stomach	Liver	Others
	Nb	8	6	1	2	1	1	4
Lesion location	Type	Full Bone		Vertebrae	Pelvis	Rib	Humerus	Foot
	Nb	490		639	128	112	175	78

It's important to note that this dataset includes detailed patient demographics, such as gender and age, providing valuable information for potential analyses. Moreover, the retrospective data collection from a reputable healthcare center enhances the dataset's reliability and potential clinical relevance. The majority of patients included in the BM-Seg dataset presented with lung and breast primary cancer as shown in table 3.2. Other common primary cancers that can spread to the bone include prostate, kidney, liver, and stomach cancer are also present in BM-Seg. This dataset provides valuable CT-scans of patients with these types of metastatic bone cancer, allowing for further research and analysis in the BM segmentation field.

3.4.6.2 *Data Collection and Labelling*

In the same way as the main dataset, the labeling of BM-Seg was carried out using a similar approach, with additional steps to obtain masks for each image. The steps performed are summarized in Figure 3.7.

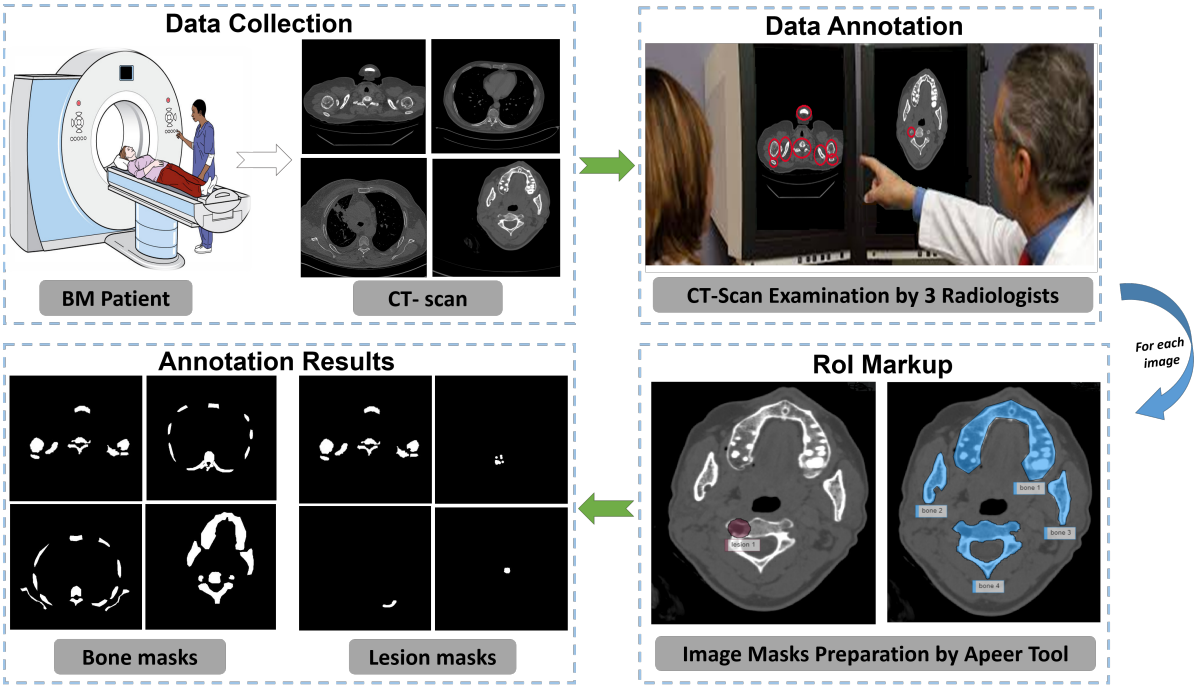

Figure 3.7: Process of building BM-Seg Dataset for Segmentation

Initially, a contrast substance is injected into the patient's veins to highlight any bone lesions. Once a patient exits the CT scanner, the CT images are sent to three radiologists for examination. After interpreting data by doctors, the selected slices were converted to JPEG format. The regions of interest (RoIs) that represent the bone and lesion regions for each CT-scan slice are manually marked using Apeer software to extract ground truth (GT) masks.

Figure 5.3 shows four extracted CT-scan images from BM-seg with their corresponding ground truth masks. The selected images in the figure belong to different sites in the human skeleton and have various BM lesion sizes. Since the labeling process is time-consuming, 70 infected slices were randomly selected from each CT-scans. On the other hand, in the CT-scans that have less than 70 infected slices, all infected slices are selected. In total, 1517 slices were annotated by creating the bone metastasis and bone masks.

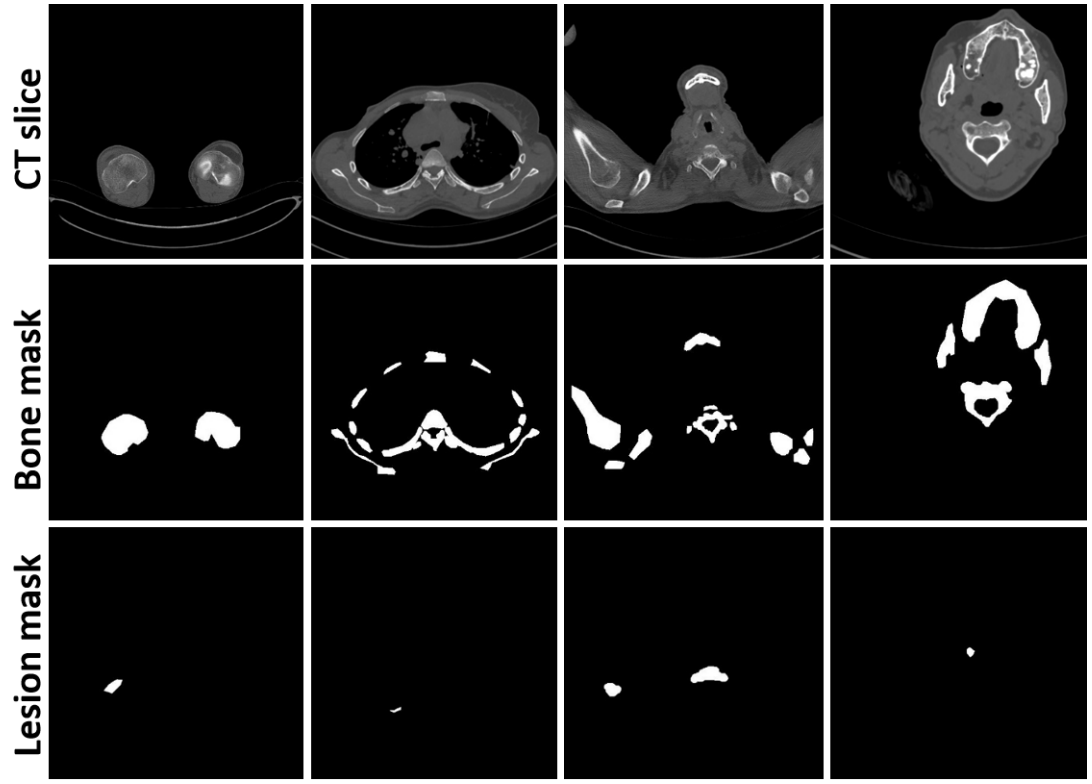

Figure 3.8: Examples of CT images with various BM sizes and different positions in the bone skeleton.

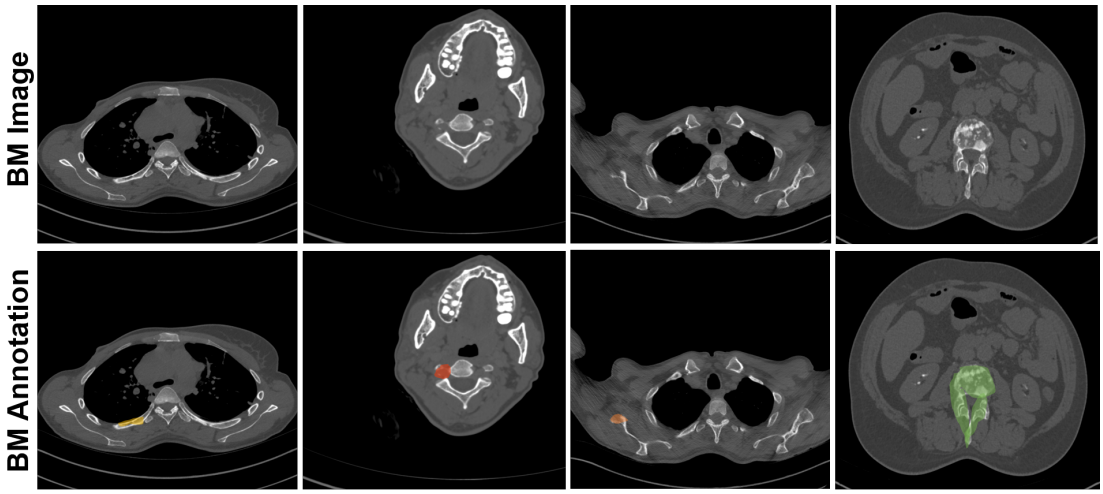

Figure 3.9: BM examples from BM-Seg. The first row contains BM images, and the second row contains an annotation of each example (colors refer to lesion locations).

At the end of the process of building the BM-Seg dataset, we obtain both bone and lesion masks for each BM image. Figure 5.2 shows examples of BM-Seg images with their corresponding annotated lesions.

3.5 CONCLUSION

In this chapter, we have presented state-of-the-art of available datasets that were introduced for bone metastasis. By analyzing the characteristics of each dataset, we have concluded that there is no published dataset that employs CT-scans. Following, we represented our proposed dataset by detailing its challenges, its implementation, opportunities, etc. Then, we presented our published dataset, denoted as BM-Seg, a part of our dataset dedicated to the segmentation task. To validate our dataset, we conducted various experiments using machine learning algorithms that will be detailed in the next chapter. Within, we will focus on segmenting BM using Unet models and their variants. Then, we will detail the classification of BM using CNN and Transformer algorithms.

4

BONE METASTASIS SEGMENTATION

Contents

4.1 INTRODUCTION

Accurate identification and delineation of BM on medical imaging is of paramount importance for both diagnosis and treatment planning. This chapter presents a breakthrough approach to BM segmentation using advanced techniques and methods. Conventional methods for segmenting BM are often inadequate to capture the full extent and subtleties of these lesions. In response to this pressing need, we present an innovative system that leverages the power of DL and CT imaging modality. Our approach improves the precision and efficiency of BM segmentation and opens new avenues for improved clinical decision-making and patient care.

In this chapter, we will explain the intricacies of our proposed methodology and discuss the underlying principles, data preprocessing steps, model architecture, and validation techniques. We will also present the promising results obtained through extensive experimentation to demonstrate the potential impact of our novel approach on the medical imaging field. In section 4.2, an overview of existing traditional and deep learning-based methods for medical imaging segmentation is presented. In section 4.3, a novel AttUnet++-based method is proposed for BM segmentation. Following, we exhibit a comparison of our results with the state-of-the-art techniques in section 4.4. Afterward, We discuss our results and their potential impact in section 4.5. We conclude the chapter in section 4.6.

4.2 IMAGE SEGMENTATION: AN OVERVIEW

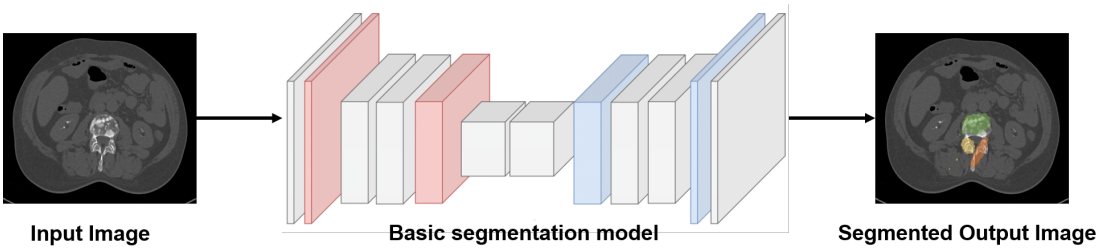

Figure 4.1: A segmentation system in which CT scans of the vertebrae bone are fed into a deep CNN architecture that not only classifies and localizes the tumor region but also highlights the anatomical structure.

In the medical field, there are numerous cases where it is difficult to distinguish between two different lesions because they present little differences. In order to effectively treat

these lesions, they must be recognized as separate entities. Such subtle differences between lesions can be challenging, but they are not impossible to detect. An image is classified pixel-by-pixel and localized, providing an outline of the object. With the help of this precise process, known as image segmentation, we can make these vital distinctions, which ensures that the right treatment strategy is followed. For example, when a CT scan image is input into the model, the result will classify the tumor type and highlight its anatomical structure, as shown in Figure 4.1. There are various applications for image segmentation, e.g., blood vessels [139], lung cancer [140], cardiovascular structures [141], etc. Further, as illustrated in Figure 4.2, multiple types and techniques are used for image segmentation. These types and techniques will be discussed in depth in the following sections.

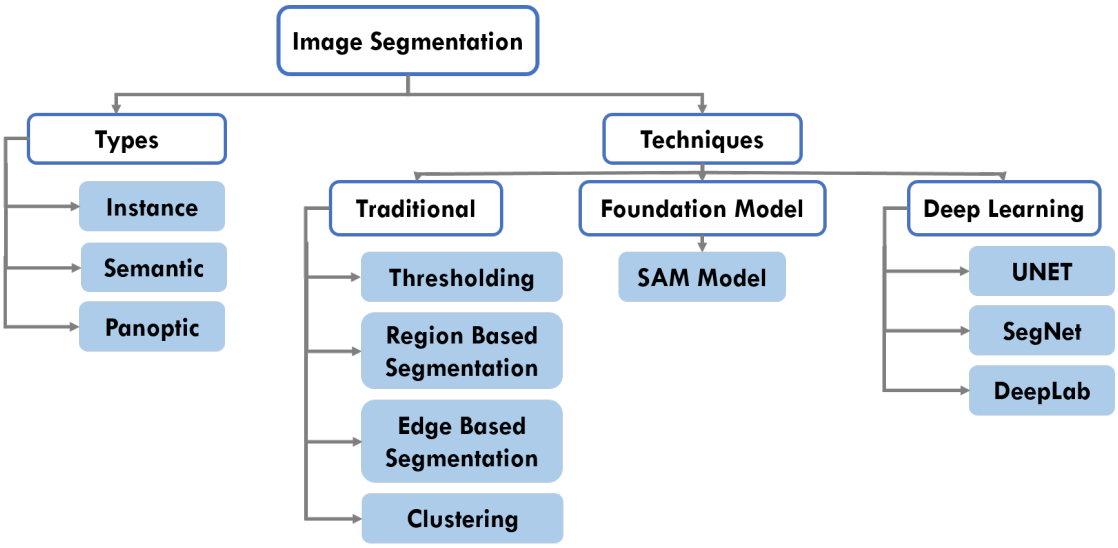

Figure 4.2: Overview of types and techniques of image segmentation [142]

4.2.1 *Image Segmentation Types*

Image segmentation ways are divided into three categories, depending on the amount and type of information to be extracted from the image: semantic, instance and panoptic [143].

a. Semantic Segmentation

In semantic segmentation, each pixel in an image is assigned a corresponding class label without considering any other information or context. The goal is to assign a label

to each pixel in the image, resulting in a dense labeling of the image. The algorithm takes an image as input and generates a segmentation map in which the pixel values (0,1,...255) of the image are converted into class labels (0,1,...n). It is useful for applications where the identification of the different classes of objects on the road is important [144].

b. Instance Segmentation

Instance segmentation is similar to semantic segmentation, but it goes a step further by distinguishing between different instances of the same object class. As with object detection, this method is also concerned with segmenting the boundaries of an object. The algorithm has no knowledge of the region's class, but it separates overlapping objects. This technique is useful in various applications such as object tracking, autonomous driving, and robotics [145].

c. Panoptic Segmentation

Panoptic segmentation is a combination of semantic and instance segmentation. It involves assigning a class label to each pixel and identifying each object in the image. This type of image segmentation provides maximum high-quality granular information from ML algorithms. It is useful for applications where the computer vision model needs to recognize and interact with different objects in its environment, such as an autonomous robot [146].

4.2.2 *Image Segmentation Techniques*

Fine-grained segmentation is a critical step in image-guided treatment and computer-aided diagnosis. The widely used architectures for image segmentation are U-Net [147], Masked R-CNN [148], DeepLab [149], etc. The use of these segmentation models depends on the problem to be solved. For example, for objects with multiple scales in the image, DeepLab and its various structures are a wise choice. Another problem in image segmentation is the lack of labeled data, so researchers are considering more unsupervised approaches, but they are still under development [150]. Generally, image segmentation techniques are classified into three categories: traditional, DL, and foundation model techniques.

a. Traditional Methods

Computer vision has long used traditional segmentation techniques to extract meaningful information from images. These techniques rely on mathematical models and

algorithms that identify regions of an image with common features such as color, texture, or brightness. They are usually computationally efficient and relatively easy to implement. Traditional techniques are commonly used for applications such as recognition, tracking, and object detection that require fast and accurate segmentation of images. In the following, we will examine some of the most common techniques.

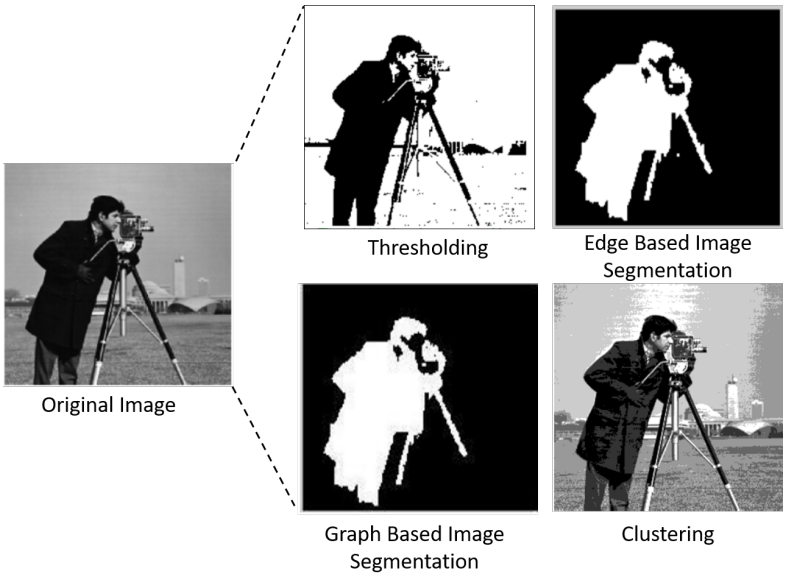

Figure 4.3: Traditional Techniques for Image Segmentation

- Thresholding is a basic image segmentation method that divides pixels into two classes based on their intensity relative to a fixed threshold. Global thresholding assigns pixels above the threshold to the foreground and those below to the background. Adaptive thresholding, on the other hand, adjusts the threshold locally for smaller regions and is therefore suitable for images with uneven illumination or contrast variations, which are often used in tasks such as document scanning and image binarization.

- Region-based segmentation in image processing divides images into regions based on similarity criteria such as color, texture, or intensity. Split and merge segmentation recursively divides the image into smaller regions and merges similar regions but may have problems with complex, irregular regions. Graphical segmentation represents the image as a graph where the nodes are pixels and the edges represent similarity. It divides the graph into regions by minimizing a cost function.

- Edge-based segmentation is an image processing technique that focuses on identifying and delineating object boundaries from the background by detecting abrupt

changes in intensity or color values. Canny edge detection is a widely used method that employs multi-level processing, including Gaussian smoothing, gradient calculation, non-maximum suppression, and hysteresis thresholding, to accurately detect edges. Sobel edge detection, on the other hand, uses a gradient-based approach that employs the Sobel operator to compute horizontal and vertical edge information. Finally, Laplacian of Gaussian (LoG) edge detection combines Gaussian smoothing and the Laplacian operator for robust edge detection, but can be very computationally intensive and may not be suitable for complex edge structures.

- Clustering is a widely used image segmentation technique in which pixels with similar features are grouped into clusters or segments. Various clustering algorithms are used for this purpose, including K-means clustering and mean shift clustering. K-means assigns pixels to K clusters based on their similarity while mean shift shifts pixels toward local density maxima. These techniques are fast and memory efficient but may require adjustments and are best suited for simpler segmentation tasks, with limited accuracy for complex scenes.

b. Deep Learning Methods

Neural networks also provide solutions for image segmentation by training them to recognize which features are important in an image, rather than relying on matched features as in traditional algorithms. Neural networks that perform the task of segmentation typically use an encoder-decoder structure. The encoder extracts features of an image through narrower and deeper filters. When the encoder has been pre-trained for a task such as image or face recognition, it uses this knowledge to extract features for segmentation (transfer learning). The decoder then uses a series of layers to inflate the encoder's output into a segmentation mask that resembles the pixel resolution of the input image. Several models that are inspired by FCNs and encoder-decoder networks have been originally developed for medical image segmentation, but are now used for image segmentation outside the medical domain as well.

U-Net [147] proposed by Ronneberger et al (Figure 4.4) is one of the most used algorithms for efficient segmentation of biological microscopy images. The U-Net architecture consists of a contracting path to capture the context and a symmetric expanding path that enables precise localization. The U-Net training strategy relies on the use of data expansion to effectively learn from very few annotated images. It was trained on 30 transmitted light microscopy images and won the 2015 ISBI Cell Tracking Challenge by a large margin. Several extensions of U-Net have been developed for different types of images and problem domains. For example, Zhou et al [151] developed a nested

U-Net architecture and Zhang et al [152] developed a U-Net-based algorithm for road segmentation.

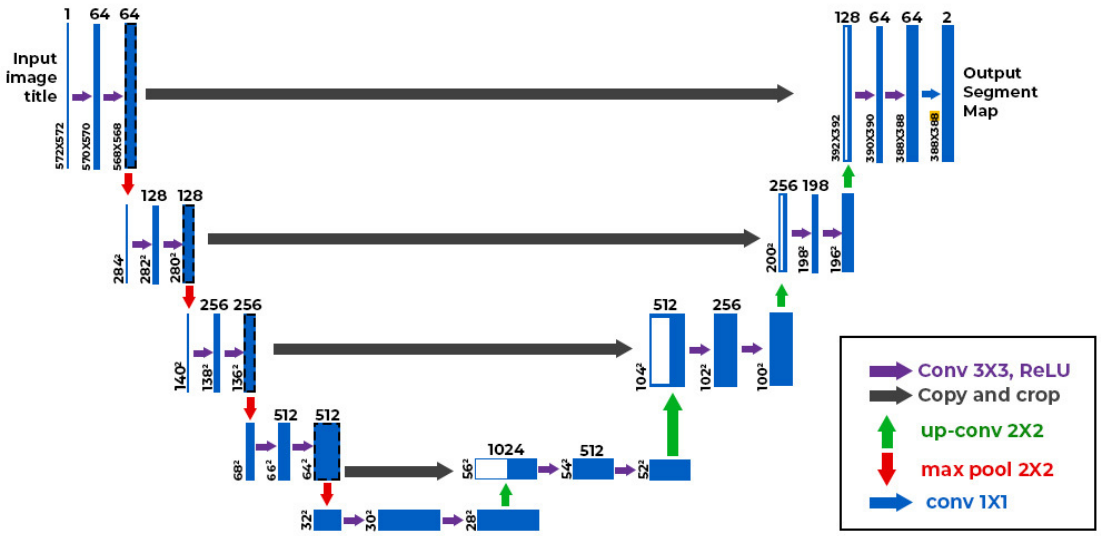

Figure 4.4: U-Net Architecture [147]

While U-Net is a widely used algorithm in the field of segmentation in medical imaging, other algorithms have gained popularity. One such algorithm is V-Net, a fully convolutional neural network based model proposed by Milletari et al [153]. for segmentation of 3D medical images. V-Net uses a new loss function based on the Dice coefficient to deal with situations where there is a strong imbalance between the number of voxels in the foreground and background. Another popular algorithm is DeepLab [149], a semantic image segmentation model that uses atrous convolution to increase the field of view of the filters without increasing the number of parameters. DeepLab has been used for a variety of medical imaging tasks, including brain tumor and retinal vessel segmentation. These algorithms, along with U-Net, are widely used in the field of medical image segmentation and continue to be the subject of ongoing research and development.

c. Foundation Model Techniques

Foundation models have also been used for image segmentation, which divides an image into distinct regions or segments. Unlike language models, which are typically based on transformer architectures, foundation models for image segmentation often use CNNs designed to handle image data.

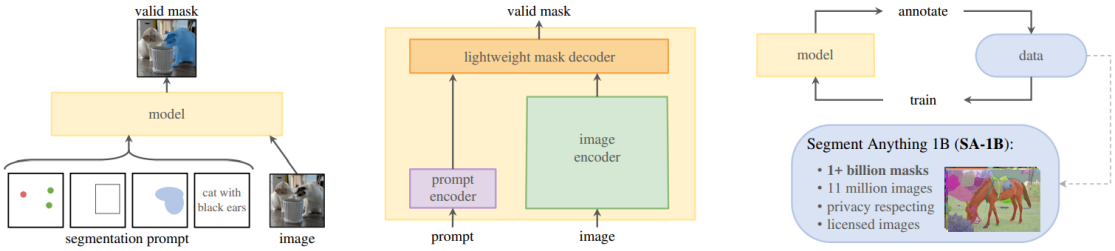

Figure 4.5: SAM Model [154]

The Segment Anything Model (SAM) [154] is widely regarded as the inaugural foundational model designed specifically for the purpose of image segmentation. SAM is built on the largest segmentation dataset to date, with over 1 billion segmentation masks. It is trained to return a valid segmentation mask for any prompt, where a prompt can be foreground/background points, a rough box or mask, or general information indicating what to segment in an image. Under the hood, an image encoder produces a one-time embedding for the image, while a lightweight encoder converts any prompt into an embedding vector in real-time. These two information sources are combined in a lightweight decoder that predicts segmentation masks.

4.3 PROPOSED APPROACH

Our proposed approach consists of two main contributions. First, we propose Hybrid-AttUnet++ architecture built from the following blocks: BCBlocks, Attention Gates, and Dual Decoders. In addition to the Hybrid-AttUnet++, the Ensemble approach is used to improve the performance and robustness of our system. The Ensemble approach is referred to as *EH-AttUnet++*. In the following, we'd like to describe the details of our proposed approach.

4.3.1 *Hybrid-AttUnet++*

Our proposed Hybrid-AttUnet++ is an extension version of the AttUNet++ architecture [155]. As described in [155], AttUnet++ consists of an encoder, nested layers, and a decoder that share the same Basic Convolution Block (BCBlock). Although AttUnet++ has improved performance in medical imaging tasks compared to many Unet variants [155], it does not show this superiority in BM segmentation, as shown in Section

4.5. In our proposed Hybrid-AttUnet++, two significant changes are made to improve the performance of AttUnet++. First, the main BCBlock is modified by adding batch normalization and residual skip connection. Second, we proposed dual decoders to segment BM and bone regions together. The proposed Hybrid-AttUnet++ architecture consists of an encoder, nested layers, and two decoders. The overall architecture of the proposed Hybrid-AttUnet++ model is shown in Figure 4.6 and described in detail below.

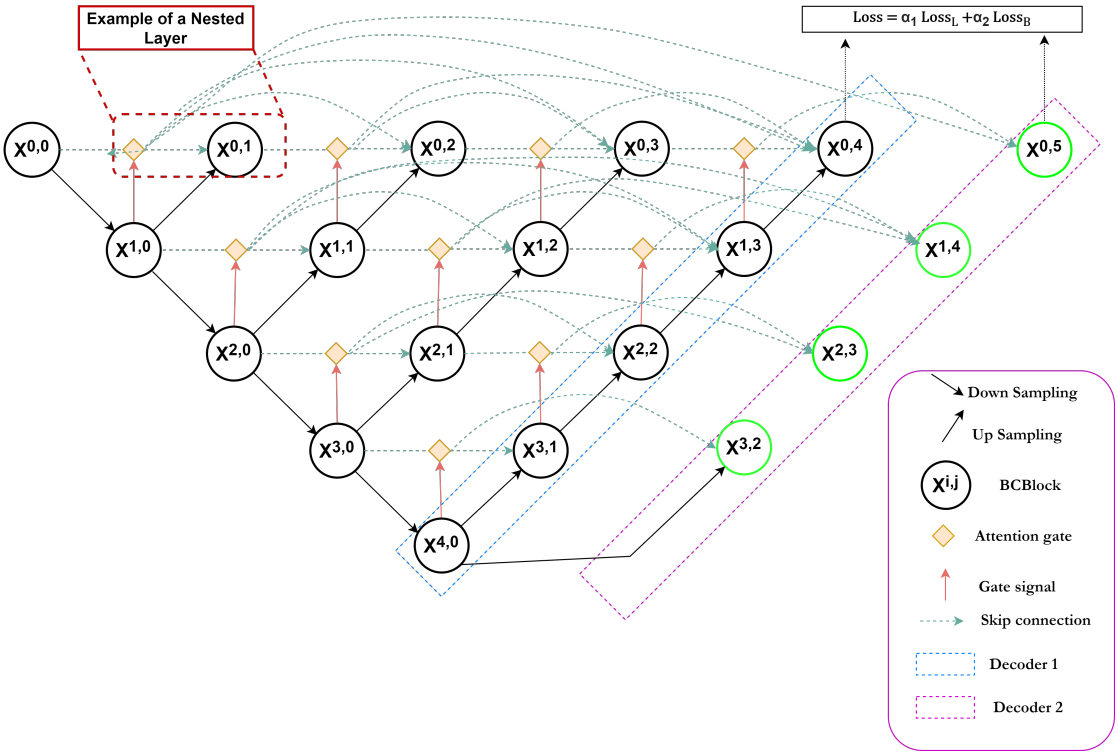

Figure 4.6: Hybrid-AttUnet++ Architecture

4.3.1.1 *BCBlocks*

The BCBlock is a residual convolutional block that converts shallower, lower, and coarse-grained features into deeper, semantic, and fine-grained ones. For $x^{i,j}$ features representing the output of the BCBlock $X^{i,j}$, where i refers to the feature depth in the encoder and j refers to the convolutional layer depth in the nested block, we define the extracted feature map of the convolutional layer $x^{i,j}$ as follows:

$$x^{i,j} = \begin{cases} H[x^{i-1,j}] & j = 0 \\ H[\int_{n=0}^{j-1} Ag(x^{i,n}), Up(x^{i+1,j-1})] & j > 0 \end{cases} \tag{4.1}$$

where *H[.]* is a cascaded operation that includes convolution, BN, ReLU activation function, and residual skip connection. *Ag ()* denotes the attention gate function, and *Up()* is a linear up-sampling.

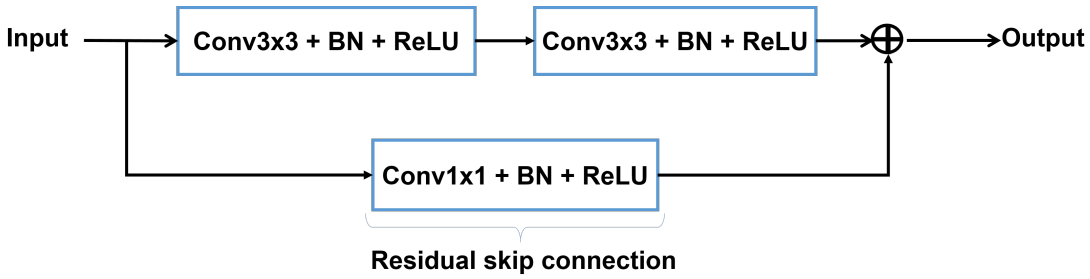

Figure 4.7: The used BCBlock structure of the proposed Hybrid-AttUnet++.

In our Hybrid-AttUnet++ architecture, each BCBlock is modified by adding a Batch Normalization (BN) and a residual skip connection, as shown in Figure 4.7. The residual skip connection is a 1-by-1 convolution layer followed by a ReLU activation function and BN, to convert the input feature map dimension to the desired output feature map dimension. Experimentally, the modified BCBlock has significantly improved performance over the BCBlock presented in AttUnet++ [155], as shown in Section 4.5.

4.3.1.2 *Encoder*

The encoder consists of a series of BCBlocks and pooling layers that increase the number of channels and downsample the spatial dimensionality of the input image, respectively. The encoder is constructed of four BCBlocks, each followed by a 2 by 2 max pooling layer, except for the last BCBlock. The BCBlock learns higher features from the previous layers and expands the feature maps with 32, 64, 128, 256, and 512 filters, respectively. This doubles the number of feature maps after each BCBlock. The structure of the encoder allows higher-level features to be learned and then transferred to the decoders of the corresponding layers. We used nested layers, which allow more efficient extraction of hierarchical features.

4.3.1.3 *Nested layers*

The feature maps extracted by the encoder pass through a number of nested layers depending on the pyramid level. For example, the skip path between nodes $X^{1,0}$ and $X^{1,3}$ consists of two nested layers. Each nested layer consists of two main components: a nested convolutional block (a BCBlock) and an attention gate (AG). As shown in Figure 4.7, the nested convolutional block combines features from its corresponding AG and from previous layers, depending on its position in the network. For example, the $X^{0,1}$ input features are the concatenation of the AG output, up-sampled features of the lower hop path $X^{1,0}$. For the nested convolutional block $X^{2,1}$, it obtains the concatenation of two deep features coming from the previous AG and the up-sampling of $X^{3,0}$. In addition, up-sampling and skip connections are used in nested layers to adjust the input features and allow the combination of features from different layers.

The AG is integrated between BCBlocks to find the most relevant spatial parts from the encoding layers before transferring them to the decoding part. Each AG combines the output of the previous BCBlock with the output of the corresponding high-sampled low-density block. The architecture of AG is shown in Figure 4.8.

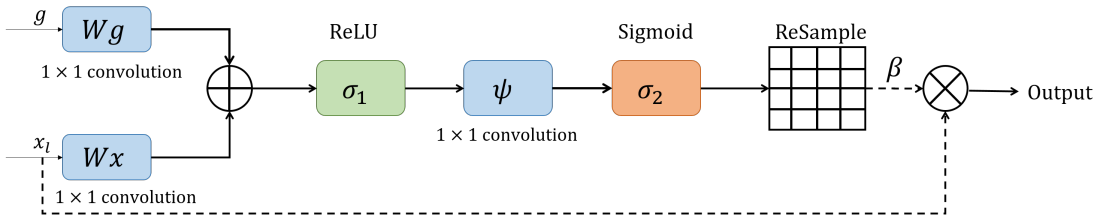

Figure 4.8: Schematic representation of the Attention Gate (AG). The g and x_l denote the gating signal vector and the feature map of layer l, respectively. W_x, W_g and ψ represent linear transformations. The β refers to the attention coefficient.

Let x_l be the feature map of the l^{th} layer and g be the gating signal. Denoting β as the attention coefficient in the range 0 to 1, the output of $Ag()$ is calculated from x_l and β as follows:

$$x_{out} = x_l.\beta \tag{4.2}$$

The attention coefficient can be calculated as:

$$\beta = \sigma_2[\psi^T \sigma_1(W_x^T x_l + W_g^T g + b_g) + b_\psi] \tag{4.3}$$

where σ_2 denotes Sigmoid function, σ_1 denotes ReLU activation function. W_x, W_g and ψ are linear transformations and b_g and b_ψ are bias terms.

4.3.1.4 *Dual decoders*

Our proposed Hybrid-AttUnet++ contains two decoders. The feature maps of the encoder are transmitted to the decoders through the nested layers. These features are used by the first decoder to segment the BM regions, while the second decoder uses them to segment the bone regions. The two decoders in the proposed model are based on BCBlocks. In each decoder, the number of channels is reduced from 512 to 256, 128, 64, 32, and 1, respectively, using five BCBlocks. Unlike the pooling layer in the encoder, each decoder uses an upsampling layer to up-sample the extracted deep features. The upsampling layer reshapes the feature maps of adjacent dimensions into a feature map with higher resolution. In this way, the network can reconstruct the spatial dimensions to obtain the same dimensions as the input images for the segmentation head.

Hybrid Loss Function: To make Hybrid-AttUnet++ pay more attention to lesion segmentation than bone segmentation, we design a hybrid loss function that combines bone and lesion losses based on the cross-entropy loss function (BCE) to calculate each loss part [11]. BCE is known to be suitable for classification and semantic segmentation and is expressed as follows:

$$L_{BCE}(y, \hat{y}) = -(y\ log(\hat{y}) + (1 - y)\ log(1 - \hat{y})) \tag{4.4}$$

where $y \in [0, 1]$ is the prediction, and $\hat{y} \in [0, 1]$ is the ground truth.

Testing different weighting values for the segmentation tasks BM and Bone, we found that $\alpha_1 = 0.7$ and $\alpha_2 = 0.3$ are the most appropriate weights for lesion loss ($Loss_L$) and bone loss ($Loss_B$), respectively. The insight behind these values is that assigning a larger weight than 0.3 for bone segmentation results in discarding the segmentation of bone metastasis lesions (main task) since they have a smaller number of pixels compared to the bone class. On the other hand, selecting a smaller weight for bone segmentation will cause the secondary task (bone segmentation) to be overlooked, rendering the proposed dual decoders useless. The proposed hybrid loss function is defined as follow:

$$Loss = \alpha_1 \cdot Loss_L + \alpha_2 \cdot Loss_B \tag{4.5}$$

$$Loss_L = L_{BCE}(y_L, \hat{y_L}) \tag{4.6}$$

$$Loss_B = L_{BCE}(y_B, \hat{y_B}) \tag{4.7}$$

where y_L and y_B are the bone metastasis and bone predictions, respectively. $\hat{y_L}$ is the lesion mask and $\hat{y_B}$ is the bone mask.

4.3.2 *EH-AttUnet++*

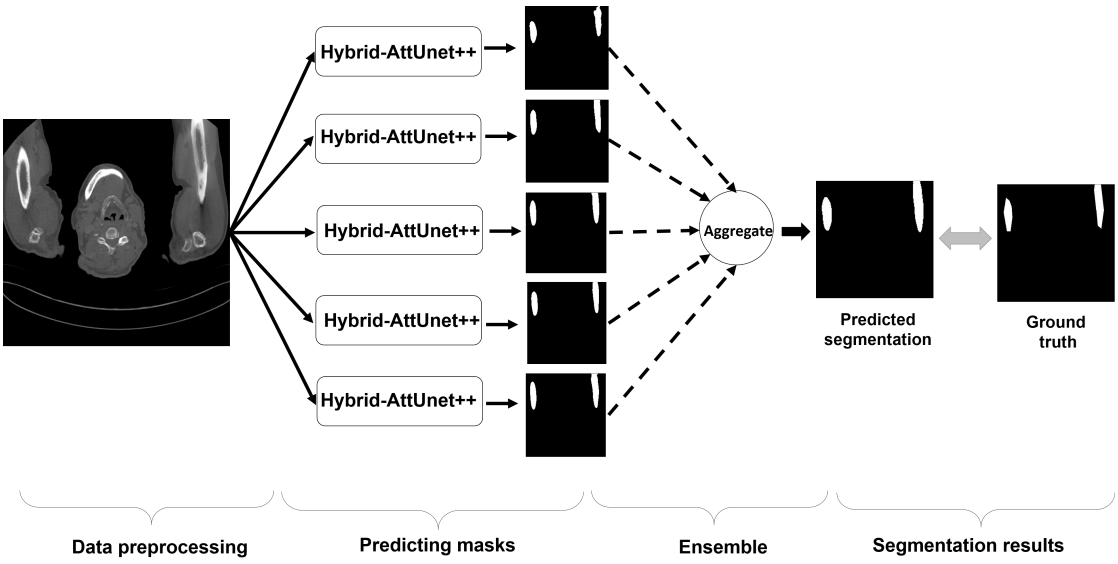

Figure 4.9: The proposed ensemble framework for BM segmentation.

Ensemble models are created in ML by combining the predictions of different independent models to improve the overall predictions. The goal of the ensemble method is to reduce generalization errors in ML algorithms and improve the segmentation result by taking lessons from the results of the individual models. To improve the performance and robustness of our system, we propose an ensemble method consisting of five Hybrid-AttUnet++ models as shown in Figure 4.9. To ensure model diversity, which is critical for ensemble performance, we used a k-fold cross-validation technique, where k, the number of folds, was set to 5. To create the partitions, we first randomly divided the

dataset into five folds. Of these folds, one was used for validation, and the remaining four for training. Then we created five splits with different training/validation subsets. For each of these five splits, Hybrid-AttUnet++ was trained from scratch. After training, the models were used to make individual predictions. Then, an ensemble approach was used by averaging the predictions of the five trained models.

4.4 EXPERIMENTS AND RESULTS

4.4.1 *Methodology*

4.4.1.1 *Experimental Setup*

The model is implemented using the PyTorch framework with NVIDIA GPU Device Quadro RTX 5000 with 16 GB from RAM. A total of 60 epochs are used to train the models with an initial learning rate of 0.01 for the first 20 epochs, 0.005 for the next 20 epochs, and 0.001 for the last 20 epochs. A batch size of 8 is used in the training phase. The loss function used is the hybrid loss function for the dual decoder approach, while the BCE loss is used for the other architectures.

4.4.1.2 *Evaluation Metrics*

Various evaluation metrics are used to assess the performance of the proposed BM segmentation model on BM-Seg. The evaluation metrics include F1-score (Eq. 5.2), Accuracy (Eq. 5.3), Precision(Prec) (Eq. 5.4), Specificity (Spec) Eq. 4.11), Sensitivity (Sens) (Eq. 5.5), Intersection-Over-Union (IoU)(Eq. 4.13) and Dice Coefficient (Dice)(Eq. 4.14).

In bone metastasis segmentation, TP (True Positive) indicates that the image is malignant and has been correctly classified. FP (False Positive) indicates that the image is benign but is classified as malignant. TN (True Negative) indicates that the image is benign and is classified as such. FN (False Negative) means that the image is classified as benign even though the label is malignant.

$$\text{F1-score} = \frac{2 \times TP}{(2 \times TP + FP + FN)} \tag{4.8}$$

$$\text{Accuracy} = \frac{TP + TN}{(TP + TN + FP + FN)} \tag{4.9}$$

$$\text{Precision} = \frac{TP}{(TP + FP)} \tag{4.10}$$

$$\text{Specificity} = \frac{TN}{(TN + FP)} \tag{4.11}$$

$$\text{Sensitivity} = \frac{TP}{(TP + FN)} \tag{4.12}$$

$$\text{IoU} = \frac{(Target \cap Predicted)}{(Target \cup Predicted)} \tag{4.13}$$

$$\text{Dice-score} = \frac{1}{N} \sum_{i=1}^{N} 2 \times \frac{TP_i}{(2 \times TP_i + FP_i + FN_i)} \tag{4.14}$$

Where N is the number of testing images.

4.4.1.3 *Comparison with the State-of-the-art Approaches*

The comparison with the state of the art includes two groups of approaches to investigate the performance of the model from different angles. First, we compare EH-AttUnet++ with Unet [147] and three of its variants, including AttUnet [156], Unet++ [151], and AttUnet++[155]. Furthermore, we compare the performance of our approach with the sequential scenario [1], where bone regions are segmented first and then BM is segmented from bone regions. For this purpose, three sequential combinations are tested, denoted *Sequential (Unet, AttUnet)*, *Sequential (Unet, Unet++)*, and *Sequential (Unet, AttUnet++)*, with the first architecture for bone segmentation and the second for bone metastasis segmentation.

4.4.2 *Experimental results*

To evaluate our approach, we perform 5-fold cross-validation using 80% of the combined dataset as training data and the rest as test data. The average results of the five-fold cross-validation are then calculated and summarized in Table 4.1. In it, the quantitative experimental performance of Hybrid-AttUnet++ and EH-AttUnet++ is compared with the state-of-the-art models on BM-Seg. The results show that our proposed Hybrid-AttUnet++ performs better than the four Unet variants in terms of F1-score, Dice and IoU, which are commonly used to evaluate segmentation tasks. More specifically, F1-score was improved by 2.81%, 2.86%, 2.53% and 13.49% compared to Unet, AttUnet, Unet++ and AttUnet++, respectively. Similarly, Dice scores improved 3.44%, 3.94%, 3.71%, and 16.96% compared to Unet, AttUnet, Unet++, and AttUnet++, respectively. Hybrid-AttUnet++ outperforms Unet, AttUnet, Unet++, and AttUnet++ by 3.96%, 4.03%, 3.58%, 2.83% for the IoU metric. In terms of accuracy, specificity, sensitivity and precision metrics, Hybrid-AttUnet++ is still the most effective. Furthermore, the results in Table 4.1 show that EH-AttUnet++ further improves the performance of Hybrid-AttUnet++, improving the F1 score, Dice and IoU results by 1.4%, 1.35% and 2.03%, respectively.

Table 4.1: Performance comparison of our proposed approach with state-of-the-art approaches (Unet, AttUnet, Unet++, AttUnet++) using the average results of the five cross-validations in terms of F1 score, Dice, IoU, accuracy, specificity, sensitivity and precision.

Model	**F1-score**	**Dice**	**IoU**	**Acc**	**Spec**	**Sens**	**Prec**
Unet [147]	79.46	72.26	65.93	98.91	78.75	99.46	80.22
AttUnet [156]	79.41	71.76	65.86	98.90	79.20	99.44	79.64
Unet++ [151]	79.74	71.99	66.31	98.92	79.41	99.46	80.16
AttUnet++ [155]	68.78	58.74	52.47	98.29	99.06	70.23	67.36
Hybrid-AttUnet++(Ours)	**82.27**	**75.70**	**69.89**	**99.05**	**81.88**	**99.53**	**82.68**
EH-AttUnet++ (Ours)	**83.67**	**77.05**	**71.92**	**99.13**	**99.59**	**82.68**	**84.68**

Table 4.2 shows a comparison between EH-AttUnet++ and the three sequential combinations. It can be seen that EH-AttUnet++ performs better than *Sequential (Unet, AttUnet)* in terms of F1-score by 6.25%, Dice by 10.16% and IoU by 8.76%. Moreover, our approach outperforms *Sequential (Unet,Unet++)* by 6.13%, 9.29%, and 8.6% for F1-score, Dice, and IoU, respectively. An improvement of 5.94%, 8.44%, and 8.34% for F1-score, Dice, and IoU relative to EH-AttUnet++ is observed compared to *Sequential*

(Unet, AttUnet++). Therefore, EH-AttUnet++ achieves higher performance compared to the sequential approaches.

Table 4.2: Performance comparison between our proposed approach (EH-AttUnet++) and the sequential approach similar to that proposed in [1](‡). The present results are the average of the five cross-validations in terms of F1-score, Dice, IoU, accuracy, specificity, sensitivity, and precision.

Model	F1-score	Dice	IoU	Acc	Spec	Sens	Prec
Sequential (Unet, AttUnet) ‡	77.41	66.88	63.16	98.83	74.54	99.50	80.63
Sequential (Unet, Unet++) ‡	77.54	67.75	63.32	98.83	75.14	99.49	80.12
Sequential (Unet, AttUnet++) ‡	77.73	68.60	63.58	98.83	76.33	99.45	79.24
EH-AttUnet++ (Ours)	**83.67**	**77.05**	**71.92**	**99.13**	**99.59**	**82.68**	**84.68**

To compare the performance of our approach with state-of-the-art approaches at the level of individual fold performance, Table 4.3 summarizes the F1-score results obtained. These experiments show that EH-AttUnet++ not only has a higher mean F1-score than the baseline models but also has a higher F1-score and high reliability for each fold.

Table 4.3: Performance comparison between our proposed approach and the state-of-the-art approaches using the first fold with F1- score indicator.

Model	**Fold 1**	**Fold 2**	**Fold 3**	**Fold 4**	**Fold 5**	**Mean**
Unet [147]	80.41	77.52	79.27	80.65	79.44	79,46
AttUnet [156]	79.17	77.79	79.04	80.63	80.39	79,41
Unet++ [151]	80.12	80.21	78.44	80.27	79.67	79.74
AttUnet++ [155]	70.34	65.99	68.65	72.59	66.32	68.78
Sequential (Unet, AttUnet)	78.23	76.81	75.86	78.67	77.49	77.41
Sequential (Unet, Unet++)	76.49	77.58	78.93	76.84	77.84	77.54
Sequential (Unet, AttUnet++)	78.19	78.26	76.04	77.87	78.29	77.73
EH-AttUnet++ (Ours)	**84.19**	**83.49**	**83**	**83.96**	**83.69**	**83.67**

4.5 DISCUSSION AND ABLATION STUDY

BM detection on the CT-scan is a common but difficult task for radiologists. Moreover, the diagnosis of BM by imaging is further complicated by the rarity of data sets available

at BM. In this study, we created our BM-Seg dataset and used a deep-learning algorithm combining dual decoders and an ensemble method to solve these problems. As mentioned in Table 4.2, the F1-score, Dice, and IoU of the segmentation of BM were 83.67%, 77.05%, and 71.92%, respectively, which were higher than the results of the baseline models. The visual comparison of EH-AttUnet++ with state-of-the-art algorithms is shown in Figure 4.10. The images from left to right show the raw infested CT-scan images and the segmentation results of Unet, AttUnet, Unet++, AttUnet++, our proposed EH-AttUnet++, and the GT-masks. The comparison of the masks clearly shows that EH-AttUnet++ has fewer false detection pixels compared to the baseline algorithms and is thus more successful in segmenting BM. Moreover, EH-AttUnet++ effectively suppresses background noise and preserves lesion details. In Figure 4.10, row 1 shows that EH-AttUnet++ is less affected by noise in the bone region because it removes noise in irrelevant locations. In addition, lines 2 and 3 show that the Unet, AttUnet, Unet++, and AttUnet++ methods cannot detect small lesions on complex BM-Seg images. Although details of small lesions can be easily lost, these details were better preserved with EH-AttUnet++. Similarly, EH-AttUnet++ shows a lower misinterpretation rate on images where the bone is completely infected compared to the reference models, as shown in lines 4 and 5.

In Figure 4.11, the results of five runs of Hybrid-AttUnet++ are compared with the results of the ensemble method. It is clear that EH-AttUnet++ consistently gives the best result in terms of F1-score, regardless of lesion location and size. For example, in the second line, EH-AttUnet++ achieves 77.44% of the F1-score, while Hybrid-AttUnet++ did not exceed 75% in the five runs (M1-M5). In particular, for small metastases, as shown in the third row, our model is able to extract more lesion features with a higher F1-score, demonstrating its ability to learn less redundant information and integrate global context. Consequently, metastatic bone segmentation using EH-AttUnet++ can be significantly improved by extracting more comprehensive background features, providing more detailed information, and obtaining a more comprehensive semantic representation.

For a comprehensive analysis, we perform ablation experiments to evaluate the efficiency of the modified BCBlock, dual decoders, and ensembling approach. The AttUNet++ architecture [155] is considered as the baseline approach. The ablation results are shown in Table 4.4, which indicates that the modified BCBlock (MB), the Dual-Decoder structure (DD), and the ensembling approach (EN) significantly improve the F1-score, Dice, and IoU metrics. In particular, the modified BCBlock can improve the F1-score, Dice, and IoU by 11.49%, 13.89%, and 14.59%, respectively, compared to the original AttUnet++. This shows that the BN and the residual skip connections are very important to prevent overfitting since the dataset is small and AttUnet++ has a higher

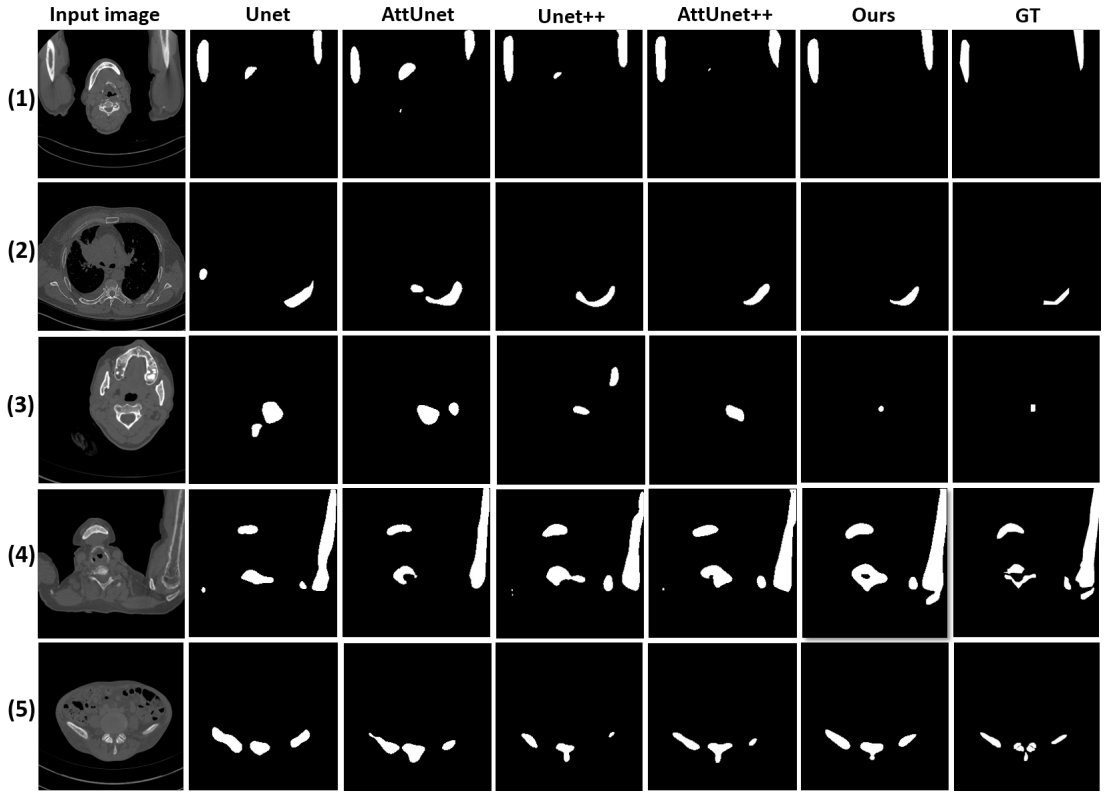

Figure 4.10: Visual comparison of BM segmentation results.

number of parameters to be optimized compared to the other variants of Unet [155]. Similarly, the dual decoder improves the F1-score, the Dice, and the IoU by 2.0%, 3.0%, and 2.0%, respectively. This demonstrates the importance of the proposed dual decoder to focus on the object of interest (BM regions), where the second task allows the network to focus attention on the bone regions and discard the non-bone regions. Finally, the ensemble approach improves the F1-score, Dice, and IoU results by 1.0%, 2.0%, and 2.0%, respectively, compared to Hybrid-AttUnet++. This proves that the ensemble approach leads to a better and more stable decision by coping with outliers in the prediction of individual models.

Both the visualization results and the test metrics show that the proposed approach can capture semantic information from both bone and metastatic bone regions. Moreover, the model outperforms other state-of-the-art models in complex scenarios where lesions of different sizes coexist and small lesions can be easily overlooked due to their similar appearance. Although our research has shown the benefits and prospects of using deep learning to segment BM from the CT-scan, more work needs to be done to analyze BM.

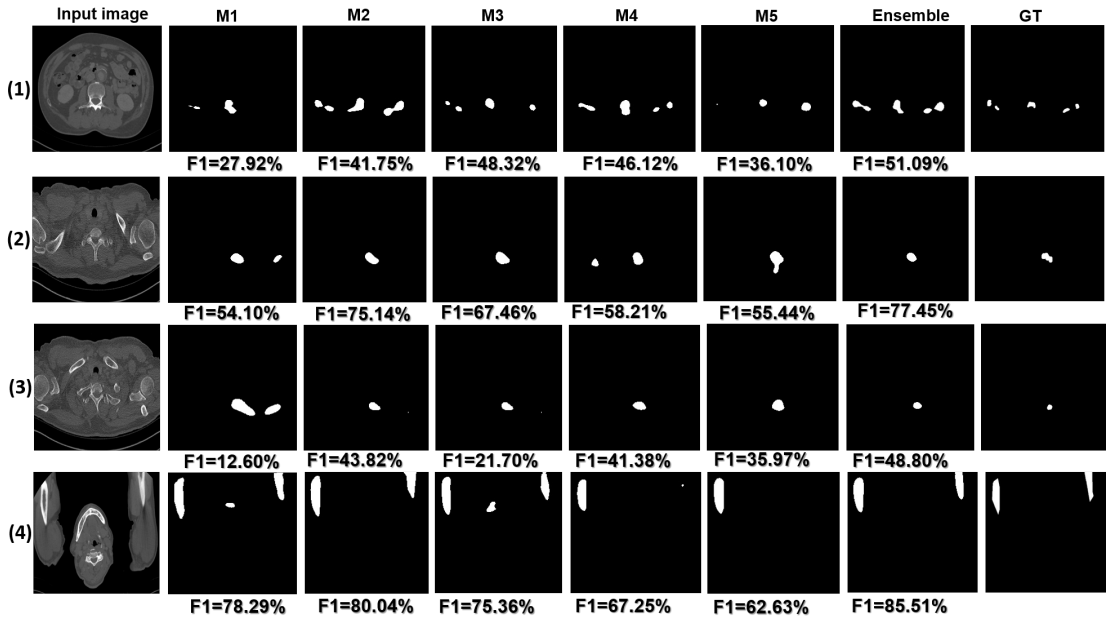

Figure 4.11: Comparison of segmentation performance of five Hybrid-AttUnet++ models (M1-M5) and EH-AttUnet++ model on BM-Seg (F1=F1-score, GT =Ground Truth).

Table 4.4: Ablation study of our proposed approach components on BM-Seg dataset (MB= Modified BCBlock, DD= Dual-Decoders, EN= Ensemble method)

	Ablation			**Dataset**		
Architecture	**MB**	**DD**	**EN**	**F1-score**	**Dice**	**IoU**
AttUnet++ (Baseline)	✗	✗	✗	68.78	58.74	52.47
AttUnet++	✓	✗	✗	80.27	72.63	67.06
Hybrid-AttUnet++	✓	✓	✗	82.27	75.70	69.89
EH-AttUnet++	✓	✓	✓	**83.67**	**77.05**	**71.92**

4.6 CONCLUSION

In this chapter, we focused on BM segmentation. Thus, we reviewed the existing methods to detect BM using Deep learning. In addition, we proposed an ensemble method (EH-AttUnet++) that highlighted areas of high uptake in CT-scan images to improve segmentation performance. The proposed approach outperformed the popular cutting-edge

models evaluated in our experiments and significantly improves the F1-score. In the next chapter, we will discuss deeply the methods used for BM classification.

5

BONE METASTASIS CLASSIFICATION

Contents

5.1 INTRODUCTION

Automatic classification of BM is a major challenge that is receiving increasing attention from the research community. One of the major challenges is the accurate classification of medical images, especially the distinction between benign and malignant images, which can greatly help physicians in decision making. Recently, several deep-learning techniques have been proposed for medical image classification. However, their performance is affected by both the dataset and the imaging modality. In this chapter, we investigate the performance of several state-of-the-art CNN architectures, namely InceptionV3, EfficientNet, ResNext50, and DenseNet161, as well as the ViT and DeiT transformer architectures. We trained and tested these algorithms on our dataset. We then present a combined approach based on DL models to improve the BM classification results.

In section 5.2, an overview of existing shallow and deep learning methods for classifying BM is given. In section 5.3, we show a comparison between CNN and Transformer models for classifying BM. Then, in section 5.4, we add a complementary experimental study. We conclude the chapter in section 5.5.

5.2 IMAGE CLASSIFICATION: AN OVERVIEW

Physicians make their diagnosis by analyzing medical images to determine the presence and nature of the disease. This traditional way of diagnosis can be supported by ML techniques. Through these techniques, the ambiguity of diagnosis between different physicians can be mitigated and the results will be more accurate. Thus, the results obtained with CNNs are not only time-saving, but can also help medical professionals. When using CNN models, an input image is fed into the network, which is then evaluated, analyzed, and interpreted to determine the target and object of the different modes [157]. This process is called image classification. For example, in Figure 5.1, CT scan images are given as input and the model classifies them into normal and BM images. Currently, there are various applications for image classification in the medical field, such as skin cancer [158], tuberculosis [159], etc.

The importance of image classification using the algorithms of ML is evident in the above applications. Considerable results have been achieved in medical applications using these ML algorithms and DL networks, but resource improvements are still needed. ML Algorithms for medical image classification can be divided into shallow and deep

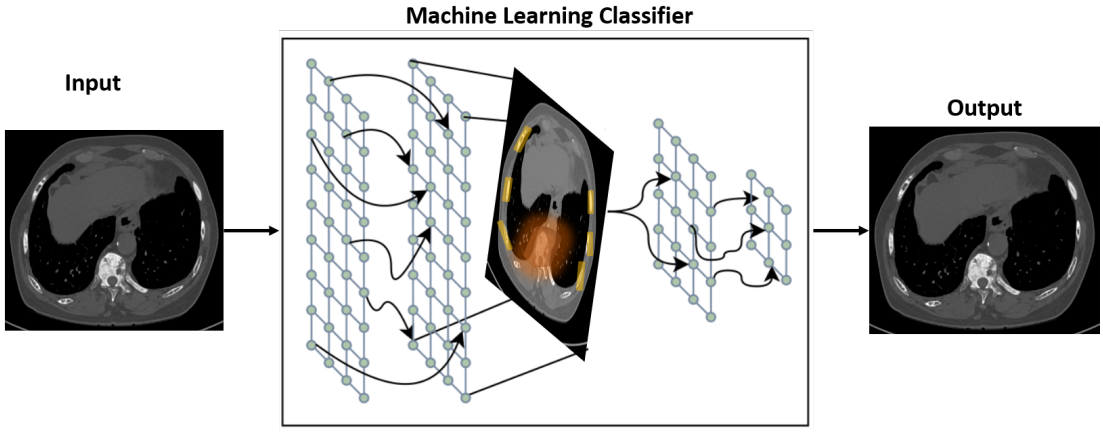

Figure 5.1: Medical image classification pipeline in which CT-scan images are fed into a Deep CNN architecture, which assigns each image a class, such as BM patient or healthy patient.

learning methods. The development history and main algorithms of these two methods are discussed in more detail in the next subsections.

5.2.1 *Shallow Methods*

Shallow learning describes all machine learning algorithms that do not use layered neural networks or multilayer perceptrons. Despite the fact that almost any technique can be tested under the guise of 'shallow learning', we chose to focus on k-nearest neighbors (KNN), logistic regression, SVM, random forest, and ANNs. The patterns in the data are almost entirely responsible for the success of these algorithms.

- **KNN (K-Nearest Neighbors):** a classification algorithm used to classify data points based on their proximity to their neighbors [160]. KNN does not assume an underlying data distribution and is therefore referred to as non-parametric. KNN is a simple technique to implement. It is a flexible and suitable classification scheme for many modalities. However, classifying unknown datasets using k-nearest neighbors can be very computationally intensive as it involves computing the distance between the unknown dataset and its k-nearest neighbors in the training data. KNN is considered a lazy learner since it does not generalize the training data and keeps all the data. Nevertheless, KNN can handle large datasets efficiently and perform computations without significant slowdown.

- **Artificial Neural Networks (ANNs):** ANNs are computer programs inspired by the way the human brain processes information. They learn from experience by recognizing patterns and relationships in data, not by programming. ANNs are made up of many individual units called artificial neurons, or processing elements, connected by weights and organized into layers. Each neuron has weighted inputs, a transfer function, and an output. During training, the connections between units are optimized until the error in predictions is minimized. ANNs can be used for classification, prediction, and modeling and have many applications in the pharmaceutical sciences.

- **Random Forest (RF):** RF is an ensemble learning method that aims to reduce noise and overfitting by creating multiple decision trees based on random, independent samples of variables and data [161]. In this technique, known as bootstrapping, random samples with replacement are selected from a learning sample of size n and the trees are matched to these samples using a modified tree learning algorithm. At each candidate split in the learning process, a random subset of predictors is selected to decorrelate each tree and reduce the intercorrelation between them. The final estimate is obtained by aggregating the information from each tree. The use of randomized models helps to reduce generalization error and reduce variance, making Random Forest a powerful tool for classification and regression problems. By mixing information from randomized models, Random Forest achieves better results than a single non-randomized model. Overall, Random Forest is an effective ensemble classifier that uses randomization to improve its predictive power and overcome the limitations of traditional decision tree algorithms.

- **Support Vector Machines (SVM):** SVM can handle classification and regression problems [162]. It uses a hyperplane as a decision boundary to separate objects belonging to different classes. SVM can handle structured and semi-structured data and scales with high-dimensional data. However, SVM's performance can degrade with large datasets and noisy data. SVM does not provide likelihood estimates, and understanding the final model can be difficult. SVM is used in cancer diagnosis, fraud detection, handwriting recognition, face recognition, and text classification. When choosing between logistic regression, decision trees (random forests), and SVM, logistic regression is tried first, followed by decision trees, and finally SVM for high-dimensional data.

- **Logistic Regression:** Logistic regression is a statistical algorithm used for binary or multinomial classification problems where the response variable is categorical and has two or more possible outcomes [163]. It provides the probability of an

event occurring based on the values of the input variables. For example, it can predict whether a tumor is malignant or benign, or classify e-mail as spam or not. Logistic regression has several advantages, including ease of implementation, computational efficiency, ease of regularization, and the fact that no scaling is required for the input features. This algorithm is widely used to solve real-world problems, such as predicting the risk of developing a certain disease, diagnosing cancer, or predicting the mortality of injured patients. However, it has some limitations, such as the inability to solve nonlinear problems, the tendency to overfit, and the need to identify all independent variables.

5.2.2 *Deep Learning Methods*

Deep Learning (DL) is gaining popularity in scientific research due to its ability to learn from context and adapt to different types of data in different domains [164, 165]. It is widely used in applications such as image recognition[166], prediction problems[167], object recognition [168, 169], smart homes [170], and self-driving cars [171, 172, 173] . Deep Learning mimics the human brain by filtering information through multiple layers, which contributes to accurate decision-making. These multilayered filters act as feedback loops, similar to neural networks in the brain, with each layer providing feedback to the next. The process continues until the exact output is achieved. To accomplish this, weights are assigned to each layer and adjusted during training to achieve accurate results. In the techniques of DL, "deep" refers to the transformation of data over a large number of layers.

Convolutional Neural Network (CNN) is a widely used DL algorithm for image classification, object recognition, and other computer vision tasks [174]. CNNs are specifically designed to analyze image data by extracting features and patterns from images through a series of convolutional layers that apply learned filters to the input images. These filters enable the network to recognize and understand complex visual patterns such as edges, shapes, and textures.

A CNN usually consists of three layers: a convolutional layer, a pooling layer, and a fully linked layer. The convolutional layer has a high weighting, while the pooling layer performs subsampling to produce the output of the convolutional layer and reduce the data rates of the underlying layers. The output of the pooling layer is used as input to multiple fully connected layers. Convolutional features, also known as feature maps,

are the result of a convolutional filter applied to the dataset. Below is an equation that describes the convolution process:

$$y(i,j) = (W * X)(i,j) = \sum_{k=1}^{K} \sum_{l=1}^{L} W(k,l)X(i-k, j-l) \tag{5.1}$$

Where W represents the input image, X represents the kernel or filter used in the convolution, k represents a series of images, and l represents the column of images.

The use of convolutional layers in CNNs has greatly improved the accuracy and efficiency of image classification tasks, making them a popular choice for many computer vision applications. There are several variants of CNN such as VGG [175], AlexNet [176], Xception [177], Inception [178], ResNet [179] and others. Depending on their learning capabilities, these models can be used in different application domains. As for image segmentation, UNet, a variant of CNN, is the most popular architecture with variants such as Unet++, AttUnet, etc. used for medical image analysis.

DL has enormous potential in various fields, such as smart cities [180] and healthcare [181], although its evaluation depends on the amount of data to be processed and the computational power. In healthcare, DL helps in faster analysis of complex medical images and comprehensive data interpretation, which supports early diagnosis of diseases and reduces manual workload [182].

5.3 COMPARISON OF CNN AND TRANSFORMER ARCHITECTURES FOR BM CLASSIFICATION

In this section, we review the differences between CNN and Transformers algorithms, more specifically, the use of these algorithms for the BM classification problem. Figure 5.2 shows the structure of the proposed deep learning framework for classifying BM using CT -scan images.

5.3.1 *Background*

5.3.1.1 *Pre-trained CNN Models for BM Classification*

Over the past decade, CNN has demonstrated its effectiveness for a variety of computer vision tasks such as detection, recognition, and segmentation [183, 184]. CNN's strength

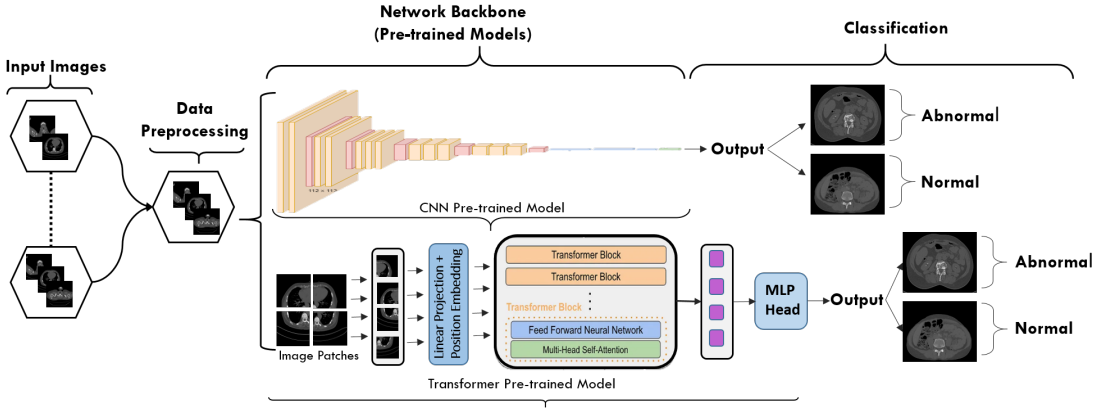

Figure 5.2: The abstract organizational structure of the proposed deep learning framework for classification of bone metastases from CT-scan images. The pre-trained CNN and Transformers models serve as deep feature extractors while the fully connected layer is tuned.

lies in its ability to automatically identify relevant features from raw data without complicated and expensive feature engineering. By using convolutional layers with small filters, CNNs can capture local patterns and gradually learn more global features. The spatial hierarchy of CNN is particularly advantageous for medical imaging tasks where spatial relationships between structures are important [183]. In addition, using transfer learning with pre-trained CNNs is particularly useful when there is limited labeled medical data and allows for more accurate classification.

The selected CNN pre-trained models are :

- **DenseNet** [185] is a CNN architecture that ensures the flow of information between the layers of the network by connecting each layer to every other layer in a feed-forward manner. For each layer, the feature maps of all previous layers are used as inputs, and their own feature maps are used as inputs for all subsequent layers [121]. DenseNet201 is one of the variants of the DenseNet architecture, which consists of several design variants such as DenseNet121, DenseNet169, DenseNet201, and DenseNet264.

- **InceptionV3** [186] is the third edition of the GoogLeNet [178] architecture, originally introduced during the ImageNet Recognition Challenge. The InceptionV3 architecture consists of 48 layers and was designed to allow for deeper networks while keeping the number of parameters from becoming too large. The inclusion of factorized convolutions was a key strategy to improve computational efficiency.

InceptionV3 was trained on ImageNet and compared to other current models, achieving the lowest error rates compared to its competitors.

- **EfficientNet** [122] is a CNN that achieves high performance through clever scaling in depth, width, and resolution. By carefully balancing model size and computational cost, EfficientNet achieves the highest accuracy on various image classification tasks while requiring fewer resources. The architecture includes a novel combination of depth-separable convolutions and squeeze-and-excite blocks that significantly reduce the number of parameters and computations without compromising accuracy.

- **ResNext50** [187] is an extension of the ResNet [179] architecture. In this algorithm, the residue blocks are modified and parallel convolutional layers with fewer filters are added. The output of each filter is combined by summation and then used as input to the next residual filter. This allows ResNext50 to learn more diverse and sophisticated features than ResNet while keeping the number of parameters similar.

5.3.1.2 *Transformer Pre-trained Models*

Transformers are a new attention-driven building block for machine translation introduced by Vaswani et al. [188] These attention blocks are layers of neural networks that combine information from the entire input sequence [189]. In recent years, these models have shown excellent performance on various natural language processing (NLP) tasks, making them the first choice over recurrent models. They have also shown promising results in medical image classification. Transformers have recently been introduced and have been shown to be more powerful and robust, taking into account both spatial pixel correlation and distance relationships for more accurate and context-aware classification. In addition, transformers provide a flexible architecture that can handle inputs of different sizes and allow processing of images with different resolutions [190].

The selected Transformers pre-trained models are :

- **ViT (Vision Transformer)** [191] is based on the vanilla transformer model [188] and cascades multiple transformer layers to capture the global context of an input image. This architecture consists of a series of self-observation layers followed by a feed-forward network. The image is interpreted as a series of patches that are then encoded by a transformer encoder, similar to natural language processing. Due to the availability of large datasets combined with improved computational

capabilities, these ViT models continue the trend of removing hand-crafted visual features and inductive biases from models.

- **DeiT (Data efficient image Transformers)** [192] is a novel native distillation approach for Transformers that uses a CNN as a teacher model to train the Transformer model on medium-sized datasets. DeiT introduces a distillation token that uses the teacher model's attention maps to provide contextual guidance to the student model to learn from both labeled data and distilled knowledge, leading to better performance. By leveraging knowledge distillation and distillation token, DeiT achieves impressive results on image classification tasks while requiring significantly less computational resources and labeled data compared to traditional vision transformers. It demonstrates the potential for more data-efficient training of large-scale transform models, making them accessible to a wider range of applications and domains.

5.3.2 *Experimental setup*

The proposed deep learning models were trained, validated, and evaluated using our dataset, which is described in detail in Chapter 3. Our dataset consists of 46 patients, each of whom underwent CT-scans. The dataset includes 23 healthy individuals with 33,424 slices and 23 patients with metastases with 46,109 slices. The dataset contains CT -scan images representing a binary classification scenario with the two classes normal vs. metastasis (Abnormal Bone). A total of 79,533 slices are available for analysis. The patients in this dataset were split independently to ensure that each patient's scans were treated as separate and distinct entities for analysis. The dataset is split into two separate parts: 80 % of the data is used to train each model, while 20 % is assigned to the test set. Using this test set, we can assess the model's ability to generalize and make predictions for unseen data. The CT -scan images were resized to 224×224 pixels prior to the training process to allow for magnification of the data. Each internally used pre-trained model is designed to resize the input images to fit its specific structure. Random rotation within a 35-degree boundary is applied to allow the model to learn from images with different orientations. Horizontal and vertical flipping creates mirrored versions of the images and increases the diversity of the dataset.

The model was implemented using the PyTorch framework with the NVIDIA GPU Quadro RTX 5000 device with 16 GB from RAM. To effectively train the CNN models, a total of 60 epochs are used to train the models with an initial learning rate of 0.01 for the first 20 epochs, 0.005 for the next 20 epochs, and 0.001 for the last 20 epochs. It

was experimentally found that the Transformers were more difficult to train and needed to be trained for more epochs and used a different learning schedule. To improve the performance of the Transformer model, the learning rate was adjusted during training. The initial setting of 0.001 for ten epochs provided stable and accurate early training. After 10 epochs, the learning rate was increased to 0.01 for faster updates and convergence. At 20 epochs, it was set back to 0.001 to avoid glitches. After 30 epochs, it was reduced to 10^{-3} to refine the parameters. Finally, at 70 epochs, it was reduced to 10^{-5} for fine tuning. In the training phase, a stack size of 8 is used for all models. In the study, a focal loss function was used as described by Mushava et al [193]. This particular loss function effectively handles the class imbalance problem by giving more weight to difficult or misclassified examples during the training process.

To evaluate and compare the performance of the algorithms used in this study, four evaluation metrics were used: Precision, recall, F1-score, and Accuracy. These metrics are defined below:

$$\text{F1-score} = \frac{2 \times TP}{(2 \times TP + FP + FN)} \tag{5.2}$$

$$\text{Accuracy} = \frac{TP + TN}{(TP + TN + FP + FN)} \tag{5.3}$$

$$\text{Precision} = \frac{TP}{(TP + FP)} \tag{5.4}$$

$$\text{Recall} = \frac{TP}{(TP + FN)} \tag{5.5}$$

Where TP (True Positives) indicates the number of metastases successfully detected by the algorithm. TN (True Negatives) indicates the number of non-metastases correctly classified as normal. FP (False Positives) indicates the number of non-metastases that were incorrectly detected as metastases. FN (False Negatives) indicates the number of metastases that the algorithm did not detect as metastases.

5.3.3 *Results and discussion*

The comparison between the performance of the CNN and the Transformer architectures is summarized in Table 5.1. Within the context of ViT, the variations between "Tiny"

and "Small" versions pertain to architectural compactness and computational efficiency. Similarly, the distinctions between DeiT Small and Tiny models revolve around their scale and efficiency.

From this table, it can be seen that ViT Tiny performs the best on all evaluation metrics. Specifically, the ViT Tiny architecture outperformed the best-performing CNN architecture (ResNeXt50) by 5.13% and 6.49% in terms of accuracy and F1-score metrics, respectively. On the other hand, it is observed that the four CNN architectures achieved similar results in terms of accuracy. In terms of F1-score, it is observed that DenseNet201 and EfficientNet have lower scores than ResNext50 and Inception-V3 by about 3%. This shows that both DenseNet201 and EfficientNet are more sensitive to data imbalances. In contrast, Transformer architectures performed differently on all evaluation metrics. As can be seen in Table 5.1, ViT Tiny achieved the best performance, outperforming the two Deit architectures by a large margin. This shows the suitability of ViT approaches for the BM classification task. Interestingly, despite having a deeper architecture compared to ViT Tiny, the ViT Small model achieves lower performance. This indicates that the dataset used to train the ViT Small architecture is relatively small, which in turn leads to over-fitting.

Table 5.1: Performance comparison between Pre-trained CNN approaches and Pre-trained Transformer approaches using accuracy, F1-score, recall, and precision.

Model	Accuracy %	F1-score %	Recall %	Precision %
ResNext50	81.81	80.08	92.32	70.69
DenseNet201	80.18	77.60	93.35	66.39
EfficientNet	79.90	77.07	93.92	65.34
Inception V3	80.96	79.34	90.37	45.17
Vit Small	79.36	78.47	85.16	72.76
Vit Tiny	86.94	86.57	92.44	81.40
Deit Small	73.07	67.82	88.75	54.88
Deit Tiny	70.09	62.79	88.03	48.80

For transformer-based models, it is difficult to determine the best model because the behavior of the model may vary depending on the training parameters. However, we found that ViT Tiny performs very well under different training settings, with accuracy greater than 86%. ViT Tiny is the best model for classifying BM, but only if the learning rate is well determined.

Figure 5.3 shows the confusion matrix used to evaluate ViT Tiny, the best pre-trained model for classifying BM. From the confusion matrix, it can be seen that ViT Tiny has a commendable ability to accurately detect and classify normal bones. With a high percentage of 92.88% for normal images, the model shows its ability to correctly identify them. However, when it comes to abnormal images, the model encounters some challenges and achieves a lower accuracy of 81.40%.

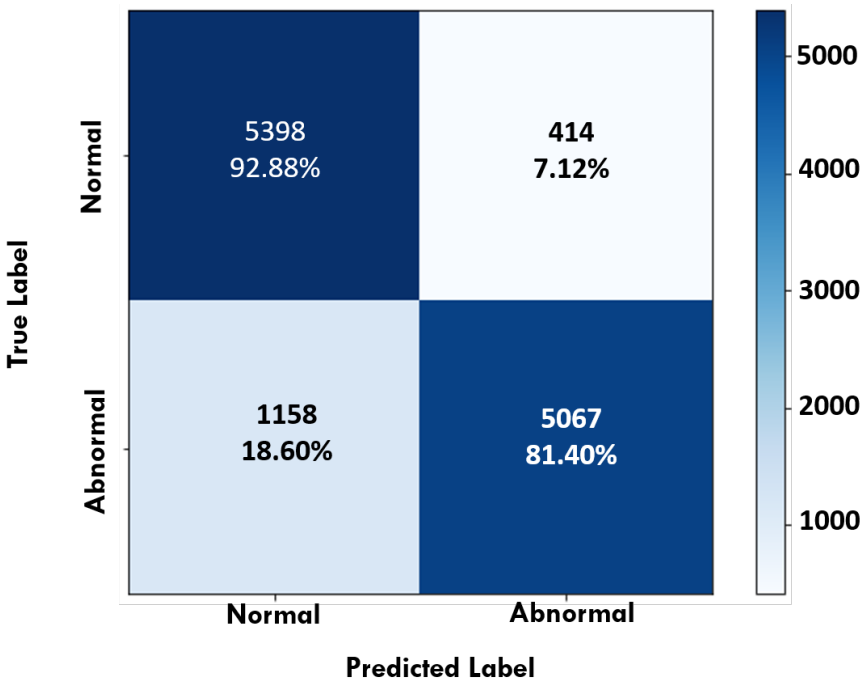

Figure 5.3: ViT Tiny Confusion matrix for BM Classification

Figure 5.4 is a visual representation to enhance understanding of how the ViT Tiny model classifies BM lesions by presenting both misclassified and correctly classified scans. The figure illustrates that when lesions are sporadically distributed across the bone (as in example (a)) or are very small (as in example (b)), the model has difficulty accurately categorizing such lesions. In images with intervertebral disks, as in example (c), or degenerative changes, as in example (d), ViT Tiny often fails to classify them correctly. This is because the presence of intervertebral disks or degenerative changes in the bone can introduce additional complexity and variation into the image that can confuse the model's classification process. On the other hand, when the lesions are larger, the model tends to achieve more accurate classifications, as seen in examples (e) and (f).

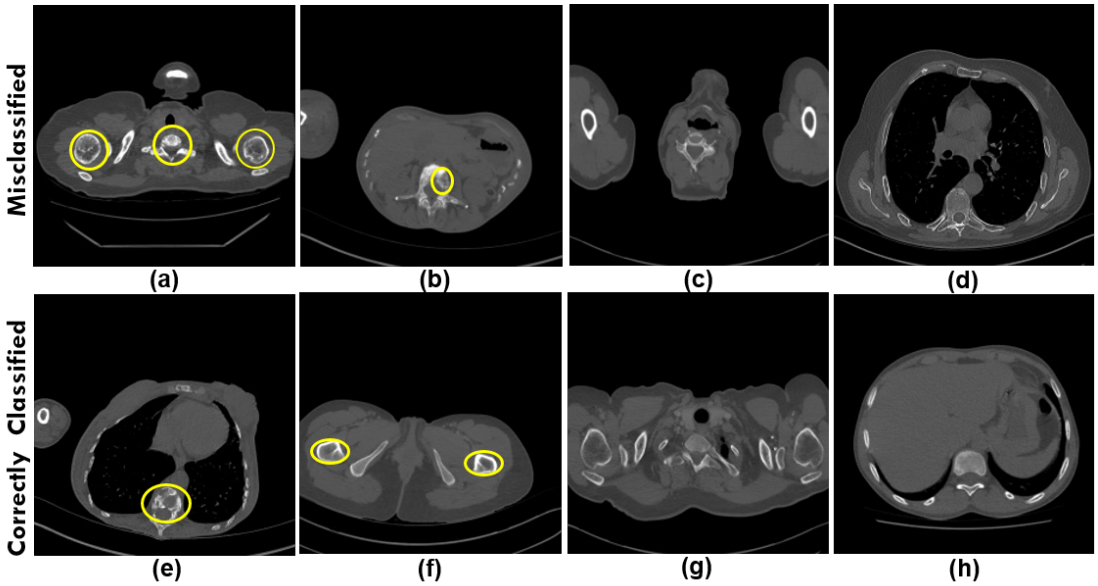

Figure 5.4: Misclassfied vs. Correctly Classified images for ViT Tiny

In addition, the model also performs well when classifying normal images, as can be seen in examples (g) and (h).

These results suggest that the model has difficulty correctly identifying and classifying abnormal instances. Thus, further improvements are needed to increase the performance of the ViT Tiny model in detecting and accurately classifying abnormal instances.

When comparing recall rates, the results in Figure 5.5 show that *EfficientNet* significantly outperforms the other pre-trained models (93.92%). Recall measures the ability of the algorithm to detect all instances of bone lesions in the image. Unlike the other pre-trained models, *EfficientNet* was able to more accurately extract all instances of lesions in an image. This is due to the low rate of FN (false negatives), which is 4.53%.

5.4 COMPLEMENTARY EXPERIMENTAL STUDY

The previous comparison of the CNN and Transformer algorithms showed that both algorithms have their strengths and weaknesses. While the four pre-trained CNN models are good at capturing local features and have proven to be effective in BM classification tasks, Transformers have shown superior performance by capturing global dependencies. Therefore, combining the best CNN algorithms with the best Transformers algorithms

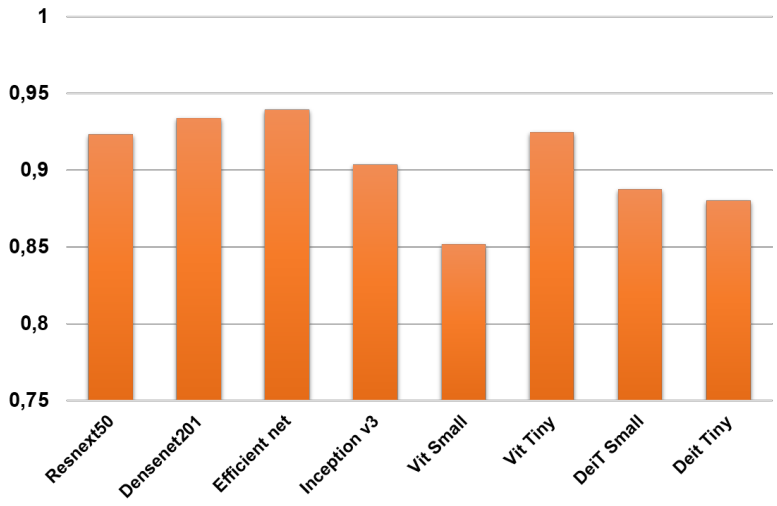

Figure 5.5: Recall Comparative Results of Pre-trained Deep Learning Architectures for BM Classification

can lead to a more powerful and effective model that can achieve top results in different domains.

5.4.1 *Architecture*

In this section, we present an architecture that leverages the strengths of two different architectures, ViT-Tiny and ResNeXt-50, to achieve improved performance in image classification. This combination aims to leverage the unique features of the two architectures, ultimately resulting in a more robust and accurate model. ViT-Tiny excels at capturing global context and spatial relationships through self-observation mechanisms, while ResNeXt-50's deep convolutional layers are capable of extracting complex features from images. By leveraging these complementary strengths, we seek to create a model that performs exceptionally well on a wide range of images.

The procedure for combining the two models ViT Tiny and ResNext50 is shown graphically in Figure 5.6.

Combining features from two models, ResNeXt and ViT Tiny, can result in a more powerful image classification model. ResNeXt is a variant of ResNet that uses a split transform merge strategy to improve the representational power of the network. ViT

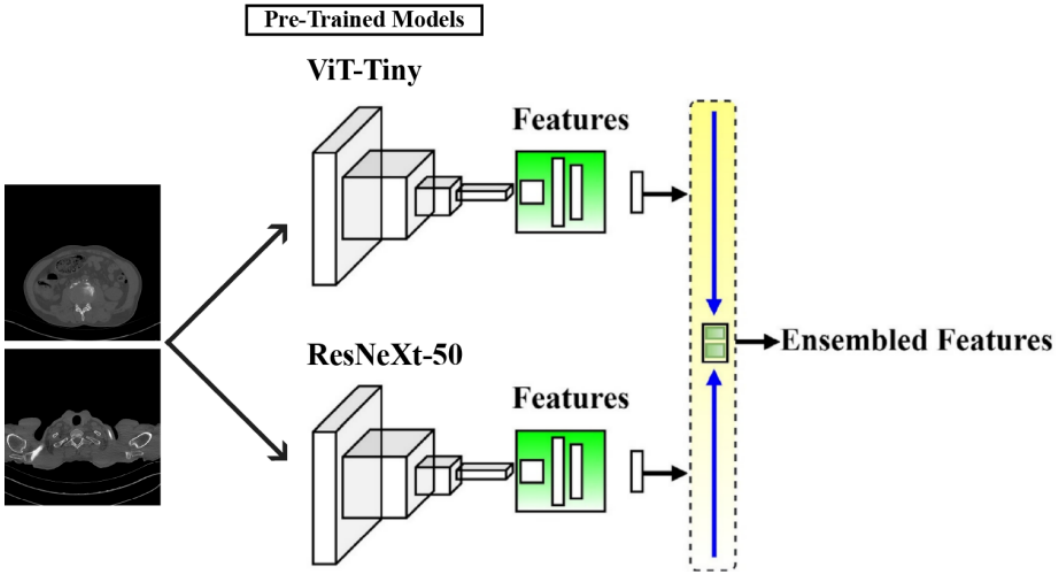

Figure 5.6: The flow diagram of the architecture combining ViT Tiny and ResNext-50

Tiny, on the other hand, is a variant of ViT that uses a smaller architecture and achieves comparable performance to larger models. By combining the features of these two models, we can leverage the strengths of both architectures and create a more robust model.

To combine the features of ResNeXt and ViT Tiny, we used a concatenation operation. Let x be the input image and f_RN and f_ViT be the feature maps obtained from ResNeXt50 and ViT Tiny, respectively. We concatenated these feature maps along the channel dimension to obtain a combined feature map f_comb, as shown in equation 5.6:

$$f_{comb} = [f_{RN}; f_{ViT}] \tag{5.6}$$

Where $[;]$ stands for the concatenation operation. We can then apply a global average pooling operation to obtain a feature vector h, as shown in equation5.7:

$$h = \frac{1}{N} \sum_{i=1}^{N} f_{comb,i} \tag{5.7}$$

Where N is the total number of spatial locations in the feature map. Finally, we can pass this feature vector through a fully connected layer with softmax activation to obtain the class probabilities as follows:

$$p(y|x) = softmax(W_{fc}h + b_{fc}) \tag{5.8}$$

Where W_fc and b_fc are the weight matrix and bias vector of the fully connected layer, respectively.

5.4.2 *Results and discussion*

To ensure a fair and meaningful comparison between models, we adopted a consistent approach to several critical aspects. These included using the same dataset for both the classification portion of BM and the evaluation of our combined model, ensuring consistent hardware settings, such as using the same GPU, and maintaining consistency in hyperparameters, including learning rate, batch size, and training epochs. In addition, we used a consistent set of evaluation metrics, including accuracy, precision, recognition, and F1 score. By adhering to this dataset, hardware, hyperparameters, and evaluation standards, we have created a robust and unbiased framework that minimizes potential sources of divergence and allows us to effectively evaluate inherent architectural differences and model performance.

In Table 5.2 we present a performance comparison between ResNext50, Vit-Tiny, and the proposed hybrid approach, evaluated using accuracy, F1 score, recall, and precision as evaluation metrics.

Table 5.2: Performance comparison between Pre-trained CNN approaches and Pre-trained Transformer approaches using accuracy, F1-score, recall, and precision.

Model	Accuracy %	F1-score %	Recall %	Precision %
ResNext50	81.81	80.08	92.32	70.69
Vit Tiny	86.94	86.57	92.44	81.40
ViT Tiny + ResNext50	**88.41**	**87.65**	79.55	**97.59**

The comparison table shows that the combined model outperforms both individual approaches in terms of accuracy, F1 score, recall, and precision. Specifically, the combined approach achieves an accuracy of 88%, which is higher than the accuracy of ViT Tiny (86%) and ResNeXt (81%). Thus, we can conclude that the proposed model is more effective than the individual approaches in classifying images. It also outperforms the

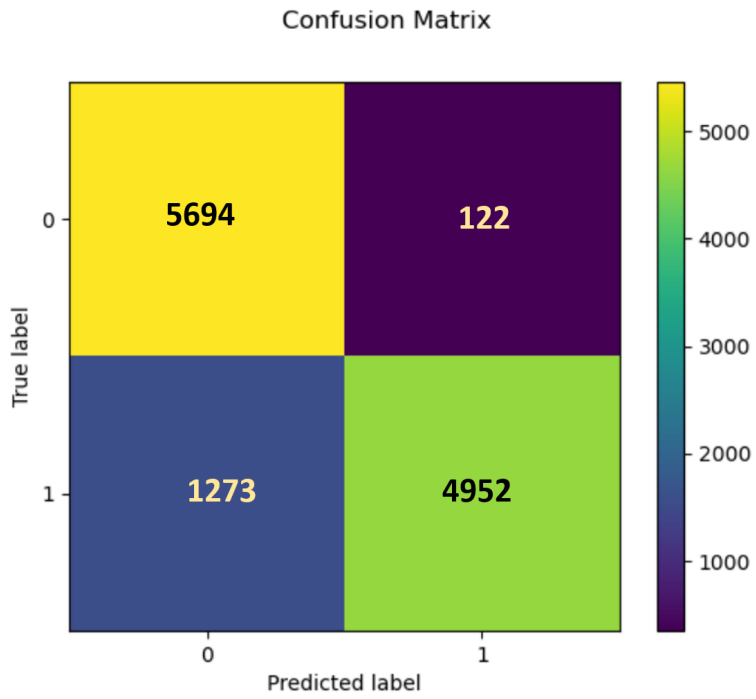

Figure 5.7: Confusion matrix of combined approach (0 refers to normal image and 1 refers to abnormal image

individual models in terms of F1 score. The F1 score of our approach is 1.08% higher than that of ViT Tiny and 7.57% higher than that of ResNeXt. Accordingly, our method achieved a better balance between precision and recall than the individual methods. In addition, the combined approach also achieved higher precision than the individual approaches. This indicates that our approach is better at detecting true positives and avoiding false positives.

The proposed approach aimed to improve accuracy by exploiting the key features of both models. To illustrate the effectiveness of this approach, we show Figure 5.7, which displays a confusion matrix. This matrix provides a comprehensive view of the model's performance in discriminating between different classes. The combined model correctly predicted 5694 negative cases and 4952 positive cases. The high number of correct positive and negative cases indicates that the proposed approach correctly identifies both classes.

This suggests that while the model can identify both negative and positive instances, there is room for improvement in reducing false positives and false negatives, which may be critical in BM classification where accurate identification is critical.

5.5 CONCLUSION

In this chapter, we have compared different pre-trained models belonging to CNN and Transformer architectures. We have also conducted a series of experiments to analyze whether current advanced Deep Learning architectures can support the BM classification task. The experimental results show that ViT Tiny outperforms the CNN-based solutions. The performance of different architectures was evaluated for the BM classification task and it was found that CNN architectures performed consistently well across all tested models, while Transformer architectures showed more variable performance. This highlights the stability and reliability of CNNs compared to Transformers. In addition, Transformers often require more extensive exploration of training hyperparameters to achieve optimal results.

To further improve the classification of BM, we proposed to combine ViT Tiny and ResNext50. The results showed that the combined model outperformed other models such as single models like ResNext50 and ViT Tiny. This approach leveraged the strengths of both architectures and achieved higher accuracy, F1 score, recall, and precision, making it a promising approach for various image classification tasks. The success of this approach encourages us to experiment with further improvements that can increase the accuracy of the classification system of BM.

6

CONCLUSIONS AND PERSPECTIVES

Contents

6.1 CONCLUSIONS AND CONTRIBUTIONS

In this work, we conducted a broad study of machine learning-based algorithms for the bone metastasis scenario to facilitate the clinical workflow by creating useful and relevant guidance for users. The work in this manuscript can be divided into two halves. In the first half, we were particularly interested in the problem of lack of data and annotations and thus focused on collecting data. The second half focused on using deep learning methods to automate medical imaging problems. In Chapter 4, we succeeded in creating a model for automatic localization and segmentation of metastatic lesions that performs comparably to a physician. We hope that in the not-too-distant future, a similar model will help in the segmentation of bone lesions in real clinical situations as we perform more data and testing. In Chapter 5, we developed a deep-learning model to distinguish bone metastases from normal bone in CT scans. The method showed potential and we hope to explore it further with more data. These results and a cursory glance at current medical journals show the trend toward AI-based solutions. It is clear that automated analysis is the future of medical imaging.

Our work resulted in multiple contributions:

- We conducted a comprehensive review of bone metastasis detection, classification, and segmentation techniques, including a survey of AI algorithms in bone metastasis diagnosis.
- We propose a benchmark dataset for bone metastasis analysis containing over 102,500 CT scan images, facilitating potential advancements in the field.
- We proposed a new segmentation approach based on Hybrid-AttUnet++ architecture, dual-decoders, and ensemble approach, achieving superior bone metastasis lesion segmentation performance.
- We conducted a comprehensive comparison of state-of-the-art CNN and Transformer architectures for bone metastasis classification, establishing a standardized evaluation framework. To further improve BM classification, we proposed a new approach combining ViT Tiny and ResNext50.
- To enhance the reproducibility of our findings, we have openly shared our dataset and code at https://github.com/Marwa-Afnouch/EH-AttUnetplus.

Our research highlights the potential of machine learning to enhance the diagnosis of bone metastasis. These advancements not only offer improved diagnostic accuracy

but also have the potential to reduce the workload of medical practitioners and enhance overall healthcare efficiency.

The integration of machine learning into clinical workflows holds great promise for improving patient outcomes and the quality of care. However, it is essential to consider the clinical implications and real-world applicability of these technologies, taking into account potential challenges and opportunities for integration into clinical practice.

6.2 PERSPECTIVES

Despite these advances made by our contributions, the field of bone metastases remains an open research topic, and various improvements are conceivable. The main limitations and perspectives to be considered are presented below.

Collecting data: The most important factor determining the success or failure of an ML model is probably the amount of data available for training. While in other fields networks are trained on millions of images, data sets for medical imaging are usually on the order of hundreds of images. This is not because large data sets are not available - hospital computer systems are filled with millions of images - but because these images are not annotated. As a solution, we have created a benchmark for bone metastases analysis named "BM-Seg" that would greatly improve access to BM data and results in medical imaging. We will continue to collect more CT-scan images from patients with different types of primary cancers and from different medical centers in order to publish a new version of BM-Seg dataset. Another solution is to develop models that can directly use unlabeled data (e.g., self-supervised or semi-supervised models) or even synthesize data using GANs. We would like to see future work thoroughly test the effectiveness of these methods.

Reproducibility of Publication Results:

Reproducibility of machine learning results is a major challenge, especially in the context of medical imaging research, as data privacy limits data availability and important hyperparameters for training cannot always be disclosed. This lack of reproducibility means that it is difficult to compare different methods, and this hinders overall scientific progress. In addition, articles too often report results without error estimates (without which the results are meaningless). One possible solution is to encourage the use of publicly available datasets with real medical data that allow independent validation. In addition, comprehensive publication along with open release of the associated code

would improve the quality of scientific publications and help other researchers. Such measures would increase transparency and enable more robust, reproducible research results in this important area.

Bridging the Gap of CAD Systems

The results of our experiments in Chapters 4 and 5 on the segmentation and classification of BM are very promising and encouraging, especially when it comes to computer-aided diagnosis. The superior performance of the proposed models makes them even more useful for routine clinical use, as they can help improve the accuracy and efficiency of the diagnostic process. However, there is still room for improvement, and further research could focus on exploring hybrid methods to improve the overall performance of the BM CAD system. By combining segmentation and classification aspects, we can leverage the strengths of the different architectures and improve the accuracy and robustness of the system. This approach can also help to address the performance variations observed in the individual models and provide a more reliable and consistent diagnosis.

In summary, this work contributes to ongoing efforts to leverage the power of machine learning for the benefit of patients and healthcare providers in the diagnosis and treatment of bone metastases, ultimately improving the quality of care and patient outcomes in the field of oncology.

BIBLIOGRAPHY

[1] Shunjiro Noguchi, Mizuho Nishio, Ryo Sakamoto, Masahiro Yakami, Koji Fujimoto, Yutaka Emoto, Takeshi Kubo, Yoshio Iizuka, Keita Nakagomi, Kazuhiro Miyasa, et al. "Deep learning–based algorithm improved radiologists' performance in bone metastases detection on CT." In: *European Radiology* (2022), pp. 1–12.

[2] Patricia S Steeg. "Tumor metastasis: mechanistic insights and clinical challenges." In: *Nature medicine* 12.8 (2006), pp. 895–904.

[3] Thomas N Seyfried and Leanne C Huysentruyt. "On the origin of cancer metastasis." In: *Critical Reviews™ in Oncogenesis* 18.1-2 (2013).

[4] Filipa Macedo, Katia Ladeira, Filipa Pinho, Nadine Saraiva, Nuno Bonito, Luisa Pinto, and Francisco Gonçalves. "Bone metastases: an overview." In: *Oncology reviews* 11.1 (2017).

[5] Gerard J O'Sullivan. "Imaging of bone metastasis: An update." In: *World Journal of Radiology* 7.8 (2015), p. 202. DOI: 10.4329/wjr.v7.i8.202. URL: https://doi.org/10.4329\%2Fwjr.v7.i8.202.

[6] Anthony Turpin, Edwina Girard, Clio Baillet, David Pasquier, Jonathan Olivier, Arnauld Villers, Philippe Puech, and Nicolas Penel. "Imaging for metastasis in prostate cancer: a review of the literature." In: *Frontiers in Oncology* 10 (2020), p. 55.

[7] Matthias Hammon, Peter Dankerl, Alexey Tsymbal, Michael Wels, Michael Kelm, Matthias May, Michael Suehling, Michael Uder, and Alexander Cavallaro. "Automatic detection of lytic and blastic thoracolumbar spine metastases on computed tomography." In: *European Radiology* 23.7 (2013), pp. 1862–1870. DOI: 10.1007/s00330-013-2774-5. URL: https://doi.org/10.1007\%2Fs00330-013-2774-5.

[8] Walter Heindel, Raphael Gübitz, Volker Vieth, Matthias Weckesser, Otmar Schober, and Michael Schäfers. "The diagnostic imaging of bone metastases." In: *Deutsches Ärzteblatt International* 111.44 (2014), p. 741.

[9] Shunjiro Noguchi, Mizuho Nishio, Masahiro Yakami, Keita Nakagomi, and Kaori Togashi. "Bone segmentation on whole-body CT using convolutional neural network with novel data augmentation techniques." In: *Computers in Biology and Medicine* 121 (2020), p. 103767. DOI: 10.1016/j.compbiomed.2020.103767. URL: https://doi.org/10.1016\%2Fj.compbiomed.2020.103767.

[10] Robert M. Vandemark, Elizabeth J. Shpall, and Mary Lou Affronti. "Bone Metastases from Breast Cancer." In: *Journal of Computer Assisted Tomography* 16.4 (1992), pp. 608–614. DOI: 10.1097/00004728-199207000-00022. URL: https://doi.org/10.1097\%2F00004728-199207000-00022.

[11] Shruti Jadon. "A survey of loss functions for semantic segmentation." In: *2020 IEEE Conference on Computational Intelligence in Bioinformatics and Computational Biology (CIBCB)*. IEEE. 2020, pp. 1–7.

[12] Fares Bougourzi, Fadi Dornaika, Karim Mokrani, Abdelmalik Taleb-Ahmed, and Yassine Ruichek. "Fusing Transformed Deep and Shallow features (FTDS) for image-based facial expression recognition." In: *Expert Systems with Applications* 156 (2020), p. 113459.

[13] Geert Litjens, Thijs Kooi, Babak Ehteshami Bejnordi, Arnaud Arindra Adiyoso Setio, Francesco Ciompi, Mohsen Ghafoorian, Jeroen Awm Van Der Laak, Bram Van Ginneken, and Clara I Sánchez. "A survey on deep learning in medical image analysis." In: *Medical image analysis* 42 (2017), pp. 60–88.

[14] Shouvik Chakraborty and Kalyani Mali. "SuFMoFPA: A superpixel and meta-heuristic based fuzzy image segmentation approach to explicate COVID-19 radiological images." In: *Expert Systems with Applications* 167 (2021), p. 114142.

[15] Fares Bougourzi, Fadi Dornaika, Nagore Barrena, Cosimo Distante, and Abdelmalik Taleb-Ahmed. "CNN based facial aesthetics analysis through dynamic robust losses and ensemble regression." en. In: *Applied Intelligence* (Aug. 2022). ISSN: 1573-7497. DOI: 10.1007/s10489-022-03943-0. URL: https://doi.org/10.1007/s10489-022-03943-0 (visited on 08/27/2022).

[16] Filipa Macedo. "Bone Metastases: An Overview." In: *Cancer therapy &amp Oncology International Journal* 4.3 (2017). DOI: 10.19080/ctoij.2017.04.555637. URL: https://doi.org/10.19080%2Fctoij.2017.04.555637.

[17] Alison Shupp, Alexus Kolb, Dimpi Mukhopadhyay, and Karen Bussard. "Cancer Metastases to Bone: Concepts, Mechanisms, and Interactions with Bone Osteoblasts." In: *Cancers* 10.6 (2018), p. 182. DOI: 10.3390/cancers10060182. URL: https://doi.org/10.3390%2Fcancers10060182.

[18] Hideshi Sugiura, Kenji Yamada, Takahiko Sugiura, Toyoaki Hida, and Tetsuya Mitsudomi. "Predictors of survival in patients with bone metastasis of lung cancer." In: *Clinical orthopaedics and related research* 466 (2008), pp. 729–736.

[19] Alexander Jinnah, Benjamin Zacks, Chukwuweike Gwam, and Bethany Kerr. "Emerging and Established Models of Bone Metastasis." In: *Cancers* 10.6 (2018), p. 176. DOI: 10.3390/cancers10060176. URL: https://doi.org/10.3390%2Fcancers10060176.

[20] Kendall M Masada, Sarah R Blumenthal, and Cara A Cipriano. "Fixation Principles for Pathologic Fractures in Metasatic Disease." In: *Orthopedic Clinics* 54.1 (2023), pp. 47–57.

[21] Roger von Moos, Luis Costa, Carla Ida Ripamonti, Daniela Niepel, and Daniele Santini. "Improving quality of life in patients with advanced cancer: targeting metastatic bone pain." In: *European journal of cancer* 71 (2017), pp. 80–94.

[22] Larry J Suva, Charity Washam, Richard W Nicholas, and Robert J Griffin. "Bone metastasis: mechanisms and therapeutic opportunities." In: *Nature Reviews Endocrinology* 7.4 (2011), pp. 208–218.

[23] Ursa Brown-Glaberman and Alison T Stopeck. "Role of denosumab in the management of skeletal complications in patients with bone metastases from solid tumors." In: *Biologics: Targets and Therapy* (2012), pp. 89–99.

[24] Sukirtha Tharmalingam, Edward Chow, Kristin Harris, Amanda Hird, and Emily Sinclair. "Quality of life measurement in bone metastases: A literature review." In: *Journal of pain research* (2008), pp. 49–58.

[25] MMTJ Bartels, R Gal, JM van der Velden, JJC Verhoeff, JJ Verlaan, and HM Verkooijen. "Impact of the COVID-19 pandemic on quality of life and emotional wellbeing in patients with bone metastases treated with radiotherapy: a prospective cohort study." In: *Clinical & experimental metastasis* 38 (2021), pp. 209–217.

[26] Muhammet Ozer, Suleyman Yasin Goksu, Jesus C Fabregas, Brian Hemendra Ramnaraign, Sherise C Rogers, Thomas J George, and Ilyas Sahin. *Influence of clinical and tumor characteristics on survival in patients with hepatocellular carcinoma with bone metastasis.* 2023.

[27] J Orcajo-Rincon, J Muñoz-Langa, JM Sepúlveda-Sánchez, GC Fernández-Pérez, M Martínez, E Noriega-Álvarez, S Sanz-Viedma, Joan Carles Vilanova, and A Luna. "Review of imaging techniques for evaluating morphological and functional responses to the treatment of bone metastases in prostate and breast cancer." In: *Clinical and Translational Oncology* 24.7 (2022), pp. 1290–1310.

[28] Mohammed M Nasef, Fatma T Eid, and Amr M Sauber. "Skeletal scintigraphy image enhancement based neutrosophic sets and salp swarm algorithm." In: *Artificial Intelligence in Medicine* 109 (2020), p. 101953.

[29] Tobias Bäuerle and Wolfhard Semmler. "Imaging response to systemic therapy for bone metastases." In: *European radiology* 19.10 (2009), pp. 2495–2507.

[30] Mateusz C Florkow, Koen Willemsen, Vasco V Mascarenhas, Edwin HG Oei, Marijn van Stralen, and Peter R Seevinck. "Magnetic Resonance Imaging Versus Computed Tomography for Three-Dimensional Bone Imaging of Musculoskeletal Pathologies: A Review." In: *Journal of Magnetic Resonance Imaging* (2022).

[31] Yan-Ran Wang, Lucia Baratto, K Elizabeth Hawk, Ashok J Theruvath, Allison Pribnow, Avnesh S Thakor, Sergios Gatidis, Rong Lu, Santosh E Gummidipundi, Jordi Garcia-Diaz, et al. "Artificial intelligence enables whole-body positron emission tomography scans with minimal radiation exposure." In: *European journal of nuclear medicine and molecular imaging* 48 (2021), pp. 2771–2781.

[32] Adelaide Greco, Leonardo Meomartino, Giacomo Gnudi, Arturo Brunetti, and Mauro Di Giancamillo. "Imaging techniques in veterinary medicine. Part II: Computed tomography, magnetic resonance imaging, nuclear medicine." In: *European Journal of Radiology Open* 10 (2023), p. 100467.

[33] Jay A Gottfried. "Structural and functional imaging of the human olfactory system." In: *Handbook of olfaction and gustation* (2015), pp. 279–304.

[34] Gary JR Cook and Vicky Goh. "Molecular imaging of bone metastases and their response to therapy." In: *Journal of Nuclear Medicine* 61.6 (2020), pp. 799–806.

[35] Gary JR Cook and Vicky Goh. "Functional and hybrid imaging of bone metastases." In: *Journal of Bone and Mineral Research* 33.6 (2018), pp. 961–972.

[36] Hiroyuki Horikoshi, Akihiro Kikuchi, Masahisa Onoguchi, Karl Sjöstrand, and Lars Edenbrandt. "Computer-aided diagnosis system for bone scintigrams from Japanese patients: importance of training database." In: *Annals of Nuclear Medicine* 26.8 (2012), pp. 622–626. DOI: 10.1007/s12149-012-0620-5. URL: https://doi.org/10.1007%2Fs12149-012-0620-5.

[37] Osamu Tokuda, Yuko Harada, Yona Ohishi, Naofumi Matsunaga, and Lars Edenbrandt. "Investigation of computer-aided diagnosis system for bone scans: a retrospective analysis in 406 patients." In: *Annals of Nuclear Medicine* 28.4 (2014), pp. 329–339. DOI: 10.1007/s12149-014-0819-8. URL: https://doi.org/10.1007%2Fs12149-014-0819-8.

[38] Mitsuru Koizumi, Kei Wagatsuma, Noriaki Miyaji, Taisuke Murata, Kenta Miwa, Tomohiro Takiguchi, Tomoko Makino, and Masamichi Koyama. "Evaluation of a computer-assisted diagnosis system, BONENAVI version 2, for bone scintigraphy in cancer patients in a routine clinical setting." In: *Annals of Nuclear Medicine* 29.2 (2014), pp. 138–148. DOI: 10.1007/s12149-014-0921-y. URL: https://doi.org/10.1007%2Fs12149-014-0921-y.

[39] Shoichi Kikushima, Noboru Hanawa, and Fumio Kotake. "Diagnostic performance of bone scintigraphy analyzed by three artificial neural network systems." In: *Annals of Nuclear Medicine* 29.2 (2014), pp. 125–131. DOI: 10.1007/s12149-014-0919-5. URL: https://doi.org/10.1007%2Fs12149-014-0919-5.

[40] Mitsuru Koizumi, Kazuki Motegi, Masamichi Koyama, Takashi Terauchi, Takeshi Yuasa, and Junji Yonese. "Diagnostic performance of a computer-assisted diagnosis system for bone scintigraphy of newly developed skeletal metastasis in prostate cancer patients: search for low-sensitivity subgroups." In: *Annals of Nuclear Medicine* 31.7 (2017), pp. 521–528. DOI: 10.1007/s12149-017-1175-2. URL: https://doi.org/10.1007%2Fs12149-017-1175-2.

[41] Mitsuru Koizumi, Kazuki Motegi, Masamichi Koyama, Mitsutomi Ishiyama, Takashi Togawa, Tomoko Makino, Yukiko Arisaka, and Takashi Terauchi. "Diagnostic performance of a computer-assisted diagnostic system: sensitivity of BONENAVI for bone scintigraphy in patients with disseminated skeletal metastasis is not so high." In: *Annals of Nuclear Medicine* 34.3 (2020), pp. 200–211. DOI: 10.1007/s12149-020-01435-0. URL: https://doi.org/10.1007%2Fs12149-020-01435-0.

[42] Ali Aslantas, Emre Dandil, Semahat Sağlam, and Murat Çakiroğlu. "CADBOSS: A computer-aided diagnosis system for whole-body bone scintigraphy scans." In: *Journal of cancer research and therapeutics* 12.2 (2016), pp. 787–792.

[43] Florina-Gianina Elfarra, Mihaela Antonina Calin, and Sorin Viorel Parasca. "Computer-aided detection of bone metastasis in bone scintigraphy images using parallelepiped classification method." In: *Annals of Nuclear Medicine* 33.11 (2019), pp. 866–874. DOI: 10.1007/s12149-019-01399-w. URL: https://doi.org/10.1007%2Fs12149-019-01399-w.

[44] Lewis Belcher. "Convolutional neural networks for classification of prostate cancer metastases using bone scan images." In: (2017).

[45] Johnny Dang. "Classification in bone scintigraphy images using convolutional neural networks." In: *Master's Theses in Mathematical Sciences* (2016).

[46] Yong Pi, Zhen Zhao, Yongzhao Xiang, Yuhao Li, Huawei Cai, and Zhang Yi. "Automated diagnosis of bone metastasis based on multi-view bone scans using attention-augmented deep neural networks." In: *Medical Image Analysis* 65 (2020), p. 101784. DOI: 10.1016/j.media.2020.101784. URL: https://doi.org/10.1016\%2Fj.media.2020.101784.

[47] Charis Ntakolia, Dimitrios E. Diamantis, Nikolaos Papandrianos, Serafeim Moustakidis, and Elpiniki I. Papageorgiou. "A Lightweight Convolutional Neural Network Architecture Applied for Bone Metastasis Classification in Nuclear Medicine: A Case Study on Prostate Cancer Patients." In: *Healthcare* 8.4 (2020), p. 493. DOI: 10.3390/healthcare8040493. URL: https://doi.org/10.3390%2Fhealthcare8040493.

[48] Nikolaos Papandrianos, Elpiniki Papageorgiou, Athanasios Anagnostis, and Konstantinos Papageorgiou. "Efficient Bone Metastasis Diagnosis in Bone Scintigraphy Using a Fast Convolutional Neural Network Architecture." In: *Diagnostics* 10.8 (2020), p. 532. DOI: 10.3390/diagnostics10080532. URL: https://doi.org/10.3390%2Fdiagnostics10080532.

[49] Nikolaos Papandrianos, Elpiniki I. Papageorgiou, and Athanasios Anagnostis. "Development of Convolutional Neural Networks to identify bone metastasis for prostate cancer patients in bone scintigraphy." In: *Annals of Nuclear Medicine* 34.11 (2020), pp. 824–832. DOI: 10.1007/s12149-020-01510-6. URL: https://doi.org/10.1007%2Fs12149-020-01510-6.

[50] Nikolaos I. Papandrianos, Elpiniki I. Papageorgiou, Athanasios Anagnostis, Konstantinos Papageorgiou, Anna Feleki, and Dionysis Bochtis. "Development of Convolutional Neural Networkbased models for bone metastasis classification in nuclear medicine." In: *2020 11th International Conference on Information, Intelligence, Systems and Applications (IISA*. IEEE, 2020. DOI: 10.1109/iisa50023.2020.9284370. URL: https://doi.org/10.1109%2Fiisa50023.2020.9284370.

[51] Nikolaos Papandrianos, Elpiniki Papageorgiou, Athanasios Anagnostis, and Anna Feleki. "A Deep-Learning Approach for Diagnosis of Metastatic Breast Cancer in Bones from Whole-Body Scans." In: *Applied Sciences* 10.3 (2020), p. 997. DOI: 10.3390/app10030997. URL: https://doi.org/10.3390%2Fapp10030997.

[52] Nikolaos Papandrianos, Elpiniki Papageorgiou, Athanasios Anagnostis, and Konstantinos Papageorgiou. "Bone metastasis classification using whole body images from prostate cancer patients based on convolutional neural networks application." In: *PLOS ONE* 15.8 (2020). Ed. by Jeonghwan Gwak, e0237213. DOI: 10.1371/journal.pone.0237213. URL: https://doi.org/10.1371%2Fjournal.pone.0237213.

[53] Zhen Zhao et al. "Deep Neural Network Based Artificial Intelligence Assisted Diagnosis of Bone Scintigraphy for Cancer Bone Metastasis." In: (2020). DOI: 10.21203/rs.3.rs-28656/v1. URL: https://doi.org/10.21203%2Frs.3.rs-28656%2Fv1.

[54] Ghasem Hajianfar, Maziar Sabouri, Soroush Bagheri, Yazdan Salimi, Mehrdad Oveisi, Isaac Shiri, and Habib Zaidi. "Dual Input Scintigraphy Image-based Fused Deep Neural Networks for Bone Abnormalities Detection and Differentiation." In: *2021 IEEE Nuclear Science Symposium and Medical Imaging Conference (NSS/MIC)*. IEEE, 2021. DOI: 10.1109/nss/mic44867.2021.9875765. URL: https://doi.org/10.1109%2Fnss%2Fmic44867.2021.9875765.

[55] Yemei Liu et al. "Automatic identification of suspicious bone metastatic lesions in bone scintigraphy using convolutional neural network." In: *BMC Medical Imaging* 21.1 (2021). DOI: 10.1186/s12880-021-00662-9. URL: https://doi.org/10.1186%2Fs12880-021-00662-9.

[56] Mihaela Antonina Calin, Florina-Gianina Elfarra, and Sorin Viorel Parasca. "Object-oriented classification approach for bone metastasis mapping from whole-body bone scintigraphy." In: *Physica Medica* 84 (2021), pp. 141–148. DOI: 10.1016/j.ejmp.2021.03.040. URL: https://doi.org/10.1016%2Fj.ejmp.2021.03.040.

[57] Sangwon Han, Jungsu S. Oh, and Jong Jin Lee. "Diagnostic performance of deep learning models for detecting bone metastasis on whole-body bone scan in prostate cancer." In: *European Journal of Nuclear Medicine and Molecular Imaging* 49.2 (2021), pp. 585–595. DOI: 10.1007/s00259-021-05481-2. URL: https://doi.org/10.1007\%2Fs00259-021-05481-2.

[58] Abdalla Ibrahim et al. "Deep learning based identification of bone scintigraphies containing metastatic bone disease foci." In: *Cancer Imaging* 23.1 (2023). DOI: 10.1186/s40644-023-00524-3. URL: https://doi.org/10.1186%2Fs40644-023-00524-3.

[59] Ali Aslantas, Emre Dandil, Semahat Sağlam, and Murat Çakiroğlu. "CADBOSS: A computer-aided diagnosis system for whole-body bone scintigraphy scans." In: *Journal of Cancer Research and Therapeutics* 12.2 (2016), p. 787. DOI: 10.4103/0973-1482.150422. URL: https://doi.org/10.4103%2F0973-1482.150422.

[60] Tongtong Li, Qiang Lin, Yanru Guo, Shaofang Zhao, Xianwu Zeng, Zhengxing Man, Yongchun Cao, and Yonghua Hu. "Automated detection of skeletal metastasis of lung cancer with bone scans using convolutional nuclear network." In: *Physics in Medicine &amp Biology* 67.1 (2022), p. 015004. DOI: 10.1088/1361-6560/ac4565. URL: https://doi.org/10.1088%2F1361-6560%2Fac4565.

[61] Qiang Lin, Chuangui Cao, Tongtong Li, Zhengxing Man, Yongchun Cao, and Haijun Wang. "dSPIC: a deep SPECT image classification network for automated multi-disease, multi-lesion diagnosis." In: *BMC Medical Imaging* 21.1 (2021). DOI: 10.1186/s12880-021-00653-w. URL: https://doi.org/10.1186%2Fs12880-021-00653-w.

[62] Qiang Lin, Chuangui Cao, Tongtong Li, Yongchun Cao, Zhengxing Man, and Haijun Wang. "Multiclass classification of whole-body scintigraphic images using a self-defined convolutional neural network with attention modules." In: *Medical Physics* 48.10 (2021), pp. 5782–5793. DOI: 10.1002/mp.15196. URL: https://doi.org/10.1002%2Fmp.15196.

[63] Qiang Lin, Tongtong Li, Chuangui Cao, Yongchun Cao, Zhengxing Man, and Haijun Wang. "Deep learning based automated diagnosis of bone metastases with SPECT thoracic bone images." In: *Scientific Reports* 11.1 (2021). DOI: 10.1038/s41598-021-83083-6. URL: https://doi.org/10.1038%2Fs41598-021-83083-6.

[64] Qiang Lin, Zhengxing Man, Yongchun Cao, and Haijun Wang. "Automated Classification of Whole-Body SPECT Bone Scan Images with VGG-Based Deep Networks." In: *The International Arab Journal of Information Technology* 20.1 (2023). DOI: 10.34028/iajit/20/1/1. URL: https://doi.org/10.34028%2Fiajit%2F20%2F1%2F1.

[65] Xiaoyan Chen, Ziwen Zheng, Liangxia Liu, Weili Wang, and Qiang Lin. "Classifying Bone Metastasized SPECT Images in RGB Format with VGGNets." In: *2021 International Conference on Communications, Information System and Computer Engineering (CISCE)*. IEEE, 2021. DOI: 10.1109/cisce52179.2021.9446035. URL: https://doi.org/10.1109%2Fcisce52179.2021.9446035.

[66] Shaofang Zhao, Xiaoyan Chen, Ziwen zheng, and Qang Lin. "Classifying SPECT Bone Metastasis Images in Grayscale Format with VGGNets." In: *2021 International Conference on Communications, Information System and Computer Engineering (CISCE)*. IEEE, 2021. DOI: 10.1109/cisce52179.2021.9446029. URL: https://doi.org/10.1109%2Fcisce52179.2021.9446029.

[67] Liangxia Liu, Yongchun Cao, Qiang Lin, Zhengxing Man, and Weiqiong Wang. "Multi-disease classification of whole-body scintigraphy images based on deep learning." In: *2022 7th International Conference on Intelligent Computing and Signal Processing (ICSP)*. IEEE, 2022. DOI: 10.1109/icsp54964.2022.9778791. URL: https://doi.org/10.1109%2Ficsp54964.2022.9778791.

[68] Xu Cao, Yongchun Cao, Qiang Lin, Zhengxing Man, and Yubo Wang. "Classification of thoracic bone scintigraphic images using ResNet with attention modules." In: *2nd International Conference on Signal Image Processing and Communication (ICSIPC 2022)*. Ed. by Deqiang Cheng and Omer Deperlioglu. SPIE, 2022. DOI: 10.1117/12.2643937. URL: https://doi.org/10.1117%2F12.2643937.

[69] Tatjana Wiese, Jianhua Yao, Joseph E. Burns, and Ronald M. Summers. "Detection of sclerotic bone metastases in the spine using watershed algorithm and graph cut." In: *SPIE Proceedings*. Ed. by Bram van Ginneken and Carol L. Novak. SPIE, 2012. DOI: 10.1117/12.911700. URL: https://doi.org/10.1117%2F12.911700.

[70] Holger R. Roth, Jianhua Yao, Le Lu, James Stieger, Joseph E. Burns, and Ronald M. Summers. "Detection of Sclerotic Spine Metastases via Random Aggregation of Deep Convolutional Neural Network Classifications." In: *Recent Advances in Computational Methods and Clinical Applications for Spine Imaging*. Springer International Publishing, 2015, pp. 3–12. DOI: 10.1007/978-3-319-14148-0_1. URL: https://doi.org/10.1007%2F978-3-319-14148-0_1.

[71] İlhan Nahit Mutlu, Burak Koçak, Ece Ateş Kuş, Melis Baykara Ulusan, and Özgür Kılıçkesmez. "Machine Learning-Based Computed Tomography Texture Analysis of Lytic Bone Lesions Needing Biopsy: A Preliminary Study." In: *Istanbul Medical Journal* 22.3 (2021), pp. 223–231. DOI: 10.4274/imj.galenos.2021.87528. URL: https://doi.org/10.4274%2Fimj.galenos.2021.87528.

[72] Samira Masoudi et al. "Deep Learning Based Staging of Bone Lesions From Computed Tomography Scans." In: *IEEE Access* 9 (2021), pp. 87531–87542. DOI: 10.1109/access.2021.3074051. URL: https://doi.org/10.1109%2Faccess.2021.3074051.

[73] Hossein Naseri, Sonia Skamene, Marwan Tolba, Mame Daro Faye, Paul Ramia, Julia Khriguian, Haley Patrick, Aixa X Andrade, Marc David, and John Kildea. "A radiomics-based machine learning pipeline to distinguish between metastatic and healthy bone using lesion-center-based geometric regions of interest." In: (2022). DOI: 10.21203/rs.3.rs-1446196/v1. URL: https://doi.org/10.21203%2Frs.3.rs-1446196%2Fv1.

[74] Hossein Naseri, Sonia Skamene, Marwan Tolba, Mame Daro Faye, Paul Ramia, Julia Khriguian, Marc David, John Kildea, et al. "A Scalable Radiomics-and Natural Language Processing–Based Machine Learning Pipeline to Distinguish Between Painful and Painless Thoracic Spinal Bone Metastases: Retrospective Algorithm Development and Validation Study." In: *JMIR AI* 2.1 (2023), e44779.

[75] Emine Acar, Asım Leblebici, Berat Ender Ellidokuz, Yasemin Başbınar, and Gamze Çapa Kaya. "Machine learning for differentiating metastatic and completely responded sclerotic bone lesion in prostate cancer: a retrospective radiomics study." In: *The British Journal of Radiology* 92.1101 (2019), p. 20190286. DOI: 10.1259/bjr.20190286. URL: https://doi.org/10.1259%2Fbjr.20190286.

[76] Xing Xiong, Jia Wang, Su Hu, Yao Dai, Yu Zhang, and Chunhong Hu. "Differentiating Between Multiple Myeloma and Metastasis Subtypes of Lumbar Vertebra Lesions Using Machine Learning–Based Radiomics." In: *Frontiers in Oncology* 11 (2021). DOI: 10.3389/fonc.2021.601699. URL: https://doi.org/10.3389%2Ffonc.2021.601699.

[77] Vito Chianca et al. "Radiomic Machine Learning Classifiers in Spine Bone Tumors: A Multi-Software, Multi-Scanner Study." In: *European Journal of Radiology* 137 (2021), p. 109586. DOI: 10.1016/j.ejrad.2021.109586. URL: https://doi.org/10.1016%2Fj.ejrad.2021.109586.

[78] Jianfang Liu, Wei Guo, Piaoe Zeng, Yayuan Geng, Yan Liu, Hanqiang Ouyang, Ning Lang, and Huishu Yuan. "Vertebral MRI-based radiomics model to differentiate multiple myeloma from metastases: influence of features number on logistic regression model performance." In: *European Radiology* 32.1 (2021), pp. 572–581. DOI: 10.1007/s00330-021-08150-y. URL: https://doi.org/10.1007%2Fs00330-021-08150-y.

[79] James Thomas Patrick Decourcy Hallinan et al. "Deep Learning Model for Classifying Metastatic Epidural Spinal Cord Compression on MRI." In: *Frontiers in Oncology* 12 (2022). DOI: 10.3389/fonc.2022.849447. URL: https://doi.org/10.3389%2Ffonc.2022.849447.

[80] Le Lu, Meizhu Liu, Xiaojing Ye, Shipeng Yu, and Heng Huang. "Coarse-to-fine classification via parametric and nonparametric models for computer-aided diagnosis." In: *Proceedings of the 20th ACM international conference on Information and knowledge management*. 2011, pp. 2509–2512.

[81] Gregory H Chu, Pechin Lo, Hyun J Kim, Peiyun Lu, Bharath Ramakrishna, David Gjertson, Cheryce Poon, Martin Auerbach, Jonathan Goldin, and Matthew S Brown. "Automated segmentation of tumors on bone scans using anatomy-specific thresholding." In: *Medical Imaging 2012: Computer-Aided Diagnosis*. Vol. 8315. SPIE. 2012, pp. 137–144.

[82] A. Aslantaş. "Detection of Bone Metastases Using FCM and Edge Detection Algorithm." In: *International Journal of Information and Electronics Engineering* 4.6 (2014). DOI: 10.7763/ijiee.2014.v4.477. URL: https://doi.org/10.7763%2Fijiee.2014.v4.477.

[83] Shijie Geng, Shaoyong Jia, Yu Qiao, Jie Yang, and Zhenhong Jia. "Combining CNN and MIL to Assist Hotspot Segmentation in Bone Scintigraphy." In: *Neural Information Processing*. Springer International Publishing, 2015, pp. 445–452. DOI: 10.1007/978-3-319-26561-2_53. URL: https://doi.org/10.1007%2F978-3-319-26561-2_53.

[84] Akinobu Shimizu, Hayato Wakabayashi, Takumi Kanamori, Atsushi Saito, Kazuhiro Nishikawa, Hiromitsu Daisaki, Shigeaki Higashiyama, and Joji Kawabe. "Automated measurement of bone scan index from a whole-body bone scintigram." In: *International Journal of Computer Assisted Radiology and Surgery* 15.3 (2020), pp. 389–400.

[85] Ema Rachmawati, Jondri Jondri, Kurniawan Nur Ramadhani, Achmad Hussein Sundawa Kartamihardja, Arifudin Achmad, and Rini Shintawati. "Automatic whole-body bone scan image segmentation based on constrained local model." In: *Bulletin of Electrical Engineering and Informatics* 9.6 (2020), pp. 2526–2537. DOI: 10.11591/eei.v9i6.2631. URL: https://doi.org/10.11591%2Feei.v9i6.2631.

[86] Terapap Apiparakoon, Nutthaphol Rakratchatakul, Maythinee Chantadisai, Usanee Vutrapongwatana, Kanaungnit Kingpetch, Sasitorn Sirisalipoch, Yothin Rakvongthai, Tawatchai Chaiwatanarat, and Ekapol Chuangsuwanich. "MaligNet: semisupervised learning for bone lesion instance segmentation using bone scintigraphy." In: *Ieee Access* 8 (2020), pp. 27047–27066.

[87] Atsushi Saito, Hayato Wakabayashi, Hiromitsu Daisaki, Atsushi Yoshida, Shigeaki Higashiyama, Joji Kawabe, and Akinobu Shimizu. "Extraction of metastasis hotspots in a whole-body bone scintigram based on bilateral asymmetry." In: *International Journal of Computer Assisted Radiology and Surgery* 16.12 (2021), pp. 2251–2260. DOI: 10.1007/s11548-021-02488-w. URL: https://doi.org/10.1007%2Fs11548-021-02488-w.

[88] Jianan Wei, Huawei Cai, Yong Pi, Zhen Zhao, and Zhang Yi. "An automatic fine-grained skeleton segmentation method for whole-body bone scintigraphy using atlas-based registration." In: *International Journal of Computer Assisted Radiology and Surgery* 17.4 (2022), pp. 673–681. DOI: 10.1007/s11548-022-02579-2. URL: https://doi.org/10.1007%2Fs11548-022-02579-2.

[89] Simin Liu, Ming Feng, Tingting Qiao, Haidong Cai, Kele Xu, Xiaqing Yu, Wen Jiang, Zhongwei Lv, Yin Wang, and Dan Li. "Deep Learning for the Automatic Diagnosis and Analysis of Bone Metastasis on Bone Scintigrams." In: *Cancer Management and Research* Volume 14 (2022), pp. 51–65. DOI: 10.2147/cmar.s340114. URL: https://doi.org/10.2147%2Fcmar.s340114.

[90] Kaibin Huang, Shengyun Huang, Guojing Chen, Xue Li, Shawn Li, Ying Liang, and Yi Gao. "An end-to-end multi-task system of automatic lesion detection and anatomical localization in whole-body bone scintigraphy by deep learning." In: *Bioinformatics* 39.1 (2022). Ed. by Hanchuan Peng. DOI: 10.1093/bioinformatics/btac753. URL: https://doi.org/10.1093%2Fbioinformatics%2Fbtac753.

[91] Tao Wu, Renze Luo, Hongyu Lin, Hong Yu, Qingsong Wang, and Heng Liu. "Research on focal segmentation of bone scan based on Swin Transformer." In: *2023 4th International Conference on Computer Vision, Image and Deep Learning (CVIDL)*. IEEE, 2023. DOI: 10.1109/cvidl58838.2023.10166150. URL: https://doi.org/10.1109%2Fcvidl58838.2023.10166150.

[92] Kaho Shimada, Hiromitsu Daisaki, Shigeaki Higashiyama, Joji Kawabe, Ryusuke Nakaoka, and Akinobu Shimizu. "Simulation of Postmarket Fine-tuning of a Computer-aided Detection System for Bone Scintigrams and Its Performance analysis." In: *Advanced Biomedical Engineering* 12.0 (2023), pp. 51–63. DOI: 10.14326/abe.12.51. URL: https://doi.org/10.14326%2Fabe.12.51.

[93] Yi-You Chen, Po-Nien Yu, Yung-Chi Lai, Te-Chun Hsieh, and Da-Chuan Cheng. "Breast Cancer Bone Metastasis Lesion Segmentation on Bone Scintigraphy." In: (2023). DOI: 10.20944/preprints202307.0674.v1. URL: https://doi.org/10.20944%2Fpreprints202307.0674.v1.

[94] Ruiting Gao, Qiang Lin, Mingyang Luo, Zhengxing Man, Yongchun Cao, and Haijun Wang. "Automatic Segmentation of Metastasis Lesions in SPECT Bone Scan Images." In: *2020 International Conference on Intelligent Computing, Automation and Systems (ICICAS)*. IEEE, 2020. DOI: 10.1109/icicas51530.2020.00051. URL: https://doi.org/10.1109%2Ficicas51530.2020.00051.

[95] Qiang Lin, Mingyang Luo, Ruiting Gao, Tongtong Li, Zhengxing Man, Yongchun Cao, and Haijun Wang. "Deep learning based automatic segmentation of metastasis hotspots in thorax bone SPECT images." In: *PLOS ONE* 15.12 (2020). Ed. by Xiaodi Huang, e0243253. DOI: 10.1371/journal.pone.0243253. URL: https://doi.org/10.1371\%2Fjournal.pone.0243253.

[96] Jingyi Zhang, Mengge Huang, Tao Deng, Yongchun Cao, and Qiang Lin. "Bone metastasis segmentation based on Improved U-NET algorithm." In: *Journal of Physics: Conference Series*. Vol. 1848. 1. IOP Publishing. 2021, p. 012027.

[97] Guoyi Che, Yongchun Cao, Ao Zhu, Qiang Lin, Zhengxing Man, and Haijun Wang. "Segmentation of bone metastases based on attention mechanism." In: *2021 IEEE International Conference on Power Electronics, Computer Applications (ICPECA)*. IEEE, 2021. DOI: 10.1109/icpeca51329.2021.9362531. URL: https://doi.org/10.1109%2Ficpeca51329.2021.9362531.

[98] Runxia Gao, Qiang Lin, Zhengxing Man, and Yongchun Cao. "Automatic lesion segmentation of metastases in SPECT images using U-Net-based model." In: *2nd International Conference on Signal Image Processing and Communication (ICSIPC 2022)*. Ed. by Deqiang Cheng and Omer Deperlioglu. SPIE, 2022. DOI: 10.1117/12.2643560. URL: https://doi.org/10.1117%2F12.2643560.

[99] Qiang Lin, Runxia Gao, Mingyang Luo, Haijun Wang, Yongchun Cao, Zhengxing Man, and Rong Wang. "Semi-supervised segmentation of metastasis lesions in bone scan images." In: *Frontiers in Molecular Biosciences* 9 (2022). DOI: 10.3389/fmolb.2022.956720. URL: https://doi.org/10.3389%2Ffmolb.2022.956720.

[100] Yongchun Cao, Liangxia Liu, Xiaoyan Chen, Zhengxing Man, Qiang Lin, Xianwu Zeng, and Xiaodi Huang. "Segmentation of lung cancer-caused metastatic lesions in bone scan images using self-defined model with deep supervision." In: *Biomedical Signal Processing and Control* 79 (2023), p. 104068. DOI: 10.1016/j.bspc.2022.104068. URL: https://doi.org/10.1016%2Fj.bspc.2022.104068.

[101] An Xie, Yang He, Qiang Lin, Yongchun Cao, and Zhengxing Man. "Segmentation of Metastasized Lesions in Bone Scintigrams Using U-Net++ with Attention Gate." In: *2023 3rd International Conference on Neural Networks, Information and Communication Engineering (NNICE)*. IEEE. 2023, pp. 239–245.

[102] Yuchan Song, Huimin Lu, Hyoungseop Kim, Seiichi Murakami, Midori Ueno, Takashi Terasawa, and Takatoshi Aoki. "Segmentation of Bone Metastasis in CT Images Based on Modified HED." In: *2019 19th International Conference on Control, Automation and Systems (ICCAS)*. IEEE, 2019. DOI: 10.23919/iccas47443.2019.8971539. URL: https://doi.org/10.23919%2Ficcas47443.2019.8971539.

[103] Connie Y. Chang, Colleen Buckless, Kaitlyn J. Yeh, and Martin Torriani. "Automated detection and segmentation of sclerotic spinal lesions on body CTs using a deep convolutional neural network." In: *Skeletal Radiology* 51.2 (2021), pp. 391–399. DOI: 10.1007/s00256-021-03873-x. URL: https://doi.org/10.1007%2Fs00256-021-03873-x.

[104] Connie Y. Chang, Florian A. Huber, Kaitlyn J. Yeh, Colleen Buckless, and Martin Torriani. "Original research: utilization of a convolutional neural network for automated detection of lytic spinal lesions on body CTs." In: *Skeletal Radiology* (2023). DOI: 10.1007/s00256-023-04283-x. URL: https://doi.org/10.1007%2Fs00256-023-04283-x.

[105] Ilkay Yildiz Potter, Diana Yeritsyan, Sarah Mahar, Jim Wu, Ara Nazarian, Aidin Vaziri, and Ashkan Vaziri. "Automated Bone Tumor Segmentation and Classification as Benign or Malignant Using Computed Tomographic Imaging." In: *Journal of Digital Imaging* (2023). DOI: 10.1007/s10278-022-00771-z. URL: https://doi.org/10.1007%2Fs10278-022-00771-z.

[106] Xiaojie Fan, Xiaoyu Zhang, Zibo Zhang, and Yifang Jiang. "Deep Learning-Based Identification of Spinal Metastasis in Lung Cancer Using Spectral CT Images." In: *Scientific Programming* 2021 (2021). Ed. by Gustavo Ramirez, pp. 1–7. DOI: 10.1155/2021/2779390. URL: https://doi.org/10.1155%2F2021%2F2779390.

[107] Noémie Moreau, Caroline Rousseau, Constance Fourcade, Gianmarco Santini, Ludovic Ferrer, Marie Lacombe, Camille Guillerminet, Mario Campone, Mathilde Colombié, Mathieu Rubeaux, et al. "Deep learning approaches for bone and bone lesion segmentation on 18FDG PET/CT imaging in the context of metastatic breast cancer." In: *2020 42nd Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC)*. IEEE. 2020, pp. 1532–1535.

[108] Fabian Isensee, Paul F. Jaeger, Simon A. A. Kohl, Jens Petersen, and Klaus H. Maier-Hein. "nnU-Net: a self-configuring method for deep learning-based biomedical image segmentation." In: *Nature Methods* 18.2 (2020), pp. 203–211. DOI: 10.1038/s41592-020-01008-z. URL: https://doi.org/10.1038%2Fs41592-020-01008-z.

[109] Gowtham Krishnan Murugesan, Diana McCrumb, Eric Brunner, Jithendra Kumar, Rahul Soni, Vasily Grigorash, Anthony Chang, Jeff VanOss, and Stephen Moore. "Automatic Whole Body FDG PET/CT Lesion Segmentation using Residual UNet and Adaptive Ensemble." In: (2023). DOI: 10.1101/2023.02.06.525233. URL: https://doi.org/10.1101%2F2023.02.06.525233.

[110] Xiang Liu, Chao Han, Yingpu Cui, Tingting Xie, Xiaodong Zhang, and Xiaoying Wang. "Detection and Segmentation of Pelvic Bones Metastases in MRI Images for Patients With Prostate Cancer Based on Deep Learning." In: *Frontiers in Oncology* 11 (2021). DOI: 10.3389/fonc.2021.773299. URL: https://doi.org/10.3389%2Ffonc.2021.773299.

[111] Georg Hille, Johannes Steffen, Max Dünnwald, Mathias Becker, Sylvia Saalfeld, and Klaus Tönnies. "Spinal metastases segmentation in MR imaging using deep convolutional neural networks." In: *arXiv preprint arXiv:2001.05834* (2020).

[112] Hai Wang, Shaohua Xu, Kai-bin Fang, Zhang-Sheng Dai, Guo-Zhen Wei, and Lu-Feng Chen. "Contrast-enhanced magnetic resonance image segmentation based on improved U-Net and Inception-ResNet in the diagnosis of spinal metastases." In: *Journal of Bone Oncology* (2023), p. 100498.

[113] Jianhua Yao, Joseph E. Burns, Vic Sanoria, and Ronald M. Summers. "Mixed spine metastasis detection through positron emission tomography/computed tomography synthesis and multiclassifier." In: *Journal of Medical Imaging* 4.2 (2017), p. 024504. DOI: 10.1117/1.jmi.4.2.024504. URL: https://doi.org/10.1117%2F1.jmi.4.2.024504.

[114] Jianhua Yao, Joseph E Burns, and Ronald M Summers. "Sclerotic rib metastases detection on routine CT images." In: *2012 9th IEEE International Symposium on Biomedical Imaging (ISBI)*. IEEE. 2012, pp. 1767–1770.

[115] Da-Chuan Cheng, Te-Chun Hsieh, Kuo-Yang Yen, and Chia-Hung Kao. "Lesion-based bone metastasis detection in chest bone scintigraphy images of prostate cancer patients using pre-train, negative mining, and deep learning." In: *Diagnostics* 11.3 (2021), p. 518.

[116] Da-Chuan Cheng, Chia-Chuan Liu, Te-Chun Hsieh, Kuo-Yang Yen, and Chia-Hung Kao. "Bone Metastasis Detection in the Chest and Pelvis from a Whole-Body Bone Scan Using Deep Learning and a Small Dataset." In: *Electronics* 10.10 (2021), p. 1201. DOI: 10.3390/electronics10101201. URL: https://doi.org/10.3390\%2Felectronics10101201.

[117] Laura Providência, Inês Domingues, and João Santos. "An Iterative Algorithm for Semisupervised Classification of Hotspots on Bone Scintigraphies of Patients with Prostate Cancer." In: *Journal of Imaging* 7.8 (2021), p. 148. DOI: 10.3390/jimaging7080148. URL: https://doi.org/10.3390%2Fjimaging7080148.

[118] Xin Fan, Han Zhang, Yuzhen Yin, Jiajia Zhang, Mengdie Yang, Shanshan Qin, Xiaoying Zhang, and Fei Yu. "Texture analysis of 18F-FDG PET/CT for differential diagnosis spinal metastases." In: *Frontiers in Medicine* 7 (2021), p. 605746.

[119] Bradley J Erickson, Panagiotis Korfiatis, Zeynettin Akkus, and Timothy L Kline. "Machine learning for medical imaging." In: *Radiographics* 37.2 (2017), pp. 505–515.

[120] Zongmo Huang, Xiaorong Pu, Gongshun Tang, Ming Ping, Guo Jiang, Mengjie Wang, Xiaoyu Wei, and Yazhou Ren. "BS-80K: The first large open-access dataset of bone scan images." In: *Computers in Biology and Medicine* 151 (2022), p. 106221.

[121] Shivaji D Pawar, Kamal K Sharma, Suhas G Sapate, Geetanjali Y Yadav, Roobaea Alroobaea, Sabah M Alzahrani, and Mustapha Hedabou. "Multichannel DenseNet Architecture for Classification of Mammographic Breast Density for Breast Cancer Detection." In: *Frontiers in Public Health* 10 (2022).

[122] Mingxing Tan and Quoc Le. "Efficientnet: Rethinking model scaling for convolutional neural networks." In: *International conference on machine learning*. PMLR. 2019, pp. 6105–6114.

[123] Ralph TH Leijenaar, Sara Carvalho, Emmanuel Rios Velazquez, Wouter JC Van Elmpt, Chintan Parmar, Otto S Hoekstra, Corneline J Hoekstra, Ronald Boellaard, André LAJ Dekker, Robert J Gillies, et al. "Stability of FDG-PET Radiomics features: an integrated analysis of test-retest and inter-observer variability." In: *Acta oncologica* 52.7 (2013), pp. 1391–1397.

[124] Yann LeCun, Yoshua Bengio, and Geoffrey Hinton. "Deep learning." In: *nature* 521.7553 (2015), pp. 436–444.

[125] Pamela J LaMontagne, Tammie LS Benzinger, John C Morris, Sarah Keefe, Russ Hornbeck, Chengjie Xiong, Elizabeth Grant, Jason Hassenstab, Krista Moulder, Andrei G Vlassenko, et al. "OASIS-3: longitudinal neuroimaging, clinical, and cognitive dataset for normal aging and Alzheimer disease." In: *MedRxiv* (2019), pp. 2019–12.

[126] Fabio A Spanhol, Luiz S Oliveira, Caroline Petitjean, and Laurent Heutte. "A dataset for breast cancer histopathological image classification." In: *Ieee transactions on biomedical engineering* 63.7 (2015), pp. 1455–1462.

[127] Walid Al-Dhabyani, Mohammed Gomaa, Hussien Khaled, and Aly Fahmy. "Dataset of breast ultrasound images." In: *Data in brief* 28 (2020), p. 104863.

[128] Jannis Born, Gabriel Brändle, Manuel Cossio, Marion Disdier, Julie Goulet, Jérémie Roulin, and Nina Wiedemann. "POCOVID-Net: automatic detection of COVID-19 from a new lung ultrasound imaging dataset (POCUS)." In: *arXiv preprint arXiv:2004.12084* (2020).

[129] Jinyu Zhao, Yichen Zhang, Xuehai He, and Pengtao Xie. "Covid-ct-dataset: a ct scan dataset about covid-19." In: *arXiv preprint arXiv:2003.13865* 490.10.48550 (2020).

[130] Marwa Afnouch, Olfa Gaddour, Yosr Hentati, Fares Bougourzi, Mohamed Abid, Ihsen Alouani, and Abdelmalik Taleb Ahmed. "BM-Seg: A new bone metastases segmentation dataset and ensemble of CNN-based segmentation approach." In: *Expert Systems with Applications* (2023), p. 120376.

[131] Andreas S Panayides, Amir Amini, Nenad D Filipovic, Ashish Sharma, Sotirios A Tsaftaris, Alistair Young, David Foran, Nhan Do, Spyretta Golemati, Tahsin Kurc, et al. "AI in medical imaging informatics: current challenges and future directions." In: *IEEE journal of biomedical and health informatics* 24.7 (2020), pp. 1837–1857.

[132] Curtis P Langlotz, Bibb Allen, Bradley J Erickson, Jayashree Kalpathy-Cramer, Keith Bigelow, Tessa S Cook, Adam E Flanders, Matthew P Lungren, David S Mendelson, Jeffrey D Rudie, et al. "A roadmap for foundational research on artificial intelligence in medical imaging: from the 2018 NIH/RSNA/ACR/The Academy Workshop." In: *Radiology* 291.3 (2019), pp. 781–791.

[133] Filippo Pesapane, Caterina Volonté, Marina Codari, and Francesco Sardanelli. "Artificial intelligence as a medical device in radiology: ethical and regulatory issues in Europe and the United States." In: *Insights into imaging* 9 (2018), pp. 745–753.

[134] Gunnar Brix, Ursula Lechel, Gerhard Glatting, Sibylle I Ziegler, Wolfgang Münzing, Stefan P Müller, and Thomas Beyer. "Radiation exposure of patients undergoing whole-body dual-modality 18F-FDG PET/CT examinations." In: *Journal of Nuclear Medicine* 46.4 (2005), pp. 608–613.

[135] Philip J Withers, Charles Bouman, Simone Carmignato, Veerle Cnudde, David Grimaldi, Charlotte K Hagen, Eric Maire, Marena Manley, Anton Du Plessis, and Stuart R Stock. "X-ray computed tomography." In: *Nature Reviews Methods Primers* 1.1 (2021), p. 18.

[136] Richard A Robb. "The dynamic spatial reconstructor: an x-ray video-fluoroscopic CT scanner for dynamic volume imaging of moving organs." In: *IEEE transactions on medical imaging* 1.1 (1982), pp. 22–33.

[137] Medixant. *Radiant Dicom Viewer*. https://www.radiantviewer.com/. 2009.

[138] Guray Gurkan, Ismet Sarikaya, and Ali Sarikaya. "Semiquantitative assessment of osteoblastic, osteolytic, and mixed lytic-sclerotic bone lesions on fluorodeoxyglucose positron emission tomography/computed tomography and bone scintigraphy." In: *World Journal of Nuclear Medicine* 18.02 (2019), pp. 132–136.

[139] Sufian A Badawi and Muhammad Moazam Fraz. "Optimizing the trainable B-COSFIRE filter for retinal blood vessel segmentation." In: *PeerJ* 6 (2018), e5855.

[140] Jue Jiang, Yu-Chi Hu, Chia-Ju Liu, Darragh Halpenny, Matthew D Hellmann, Joseph O Deasy, Gig Mageras, and Harini Veeraraghavan. "Multiple resolution residually connected feature streams for automatic lung tumor segmentation from CT images." In: *IEEE transactions on medical imaging* 38.1 (2018), pp. 134–144.

[141] Lohendran Baskaran, Subhi J Al'Aref, Gabriel Maliakal, Benjamin C Lee, Zhuoran Xu, Jeong W Choi, Sang-Eun Lee, Ji Min Sung, Fay Y Lin, Simon Dunham, et al. "Automatic segmentation of multiple cardiovascular structures from cardiac computed tomography angiography images using deep learning." In: *PloS one* 15.5 (2020), e0232573.

[142] Pooja Bagad. *Classification of Image Segmentation Techniques*. 2021. URL: https://pooja-bagad18.medium.com/classification-of-image-segmentation-techniques-cc6a031b75fc.

[143] Nikhil R Pal and Sankar K Pal. "A review on image segmentation techniques." In: *Pattern recognition* 26.9 (1993), pp. 1277–1294.

[144] Alberto Garcia-Garcia, Sergio Orts-Escolano, Sergiu Oprea, Victor Villena-Martinez, and Jose Garcia-Rodriguez. "A review on deep learning techniques applied to semantic segmentation." In: *arXiv preprint arXiv:1704.06857* (2017).

[145] Abdul Mueed Hafiz and Ghulam Mohiuddin Bhat. "A survey on instance segmentation: state of the art." In: *International journal of multimedia information retrieval* 9.3 (2020), pp. 171–189.

[146] Alexander Kirillov, Kaiming He, Ross Girshick, Carsten Rother, and Piotr Dollár. "Panoptic segmentation." In: *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*. 2019, pp. 9404–9413.

[147] Olaf Ronneberger, Philipp Fischer, and Thomas Brox. "U-net: Convolutional networks for biomedical image segmentation." In: *International Conference on Medical image computing and computer-assisted intervention*. Springer. 2015, pp. 234–241.

[148] Kaiming He, Georgia Gkioxari, Piotr Dollár, and Ross Girshick. "Mask r-cnn." In: *Proceedings of the IEEE international conference on computer vision*. 2017, pp. 2961–2969.

[149] Liang-Chieh Chen, George Papandreou, Iasonas Kokkinos, Kevin Murphy, and Alan L Yuille. "Deeplab: Semantic image segmentation with deep convolutional nets, atrous convolution, and fully connected crfs." In: *IEEE transactions on pattern analysis and machine intelligence* 40.4 (2017), pp. 834–848.

[150] Ruixin Yang and Yingyan Yu. "Artificial convolutional neural network in object detection and semantic segmentation for medical imaging analysis." In: *Frontiers in oncology* 11 (2021), p. 638182.

[151] Zongwei Zhou, Md Mahfuzur Rahman Siddiquee, Nima Tajbakhsh, and Jianming Liang. "Unet++: A nested u-net architecture for medical image segmentation." In: *Deep learning in medical image analysis and multimodal learning for clinical decision support*. Springer, 2018, pp. 3–11.

[152] Zhengxin Zhang, Qingjie Liu, and Yunhong Wang. "Road extraction by deep residual u-net." In: *IEEE Geoscience and Remote Sensing Letters* 15.5 (2018), pp. 749–753.

[153] Fausto Milletari, Nassir Navab, and Seyed-Ahmad Ahmadi. "V-net: Fully convolutional neural networks for volumetric medical image segmentation." In: *2016 fourth international conference on 3D vision (3DV)*. Ieee. 2016, pp. 565–571.

[154] Alexander Kirillov, Eric Mintun, Nikhila Ravi, Hanzi Mao, Chloe Rolland, Laura Gustafson, Tete Xiao, Spencer Whitehead, Alexander C Berg, Wan-Yen Lo, et al. "Segment anything." In: *arXiv preprint arXiv:2304.02643* (2023).

[155] Chen Li, Yusong Tan, Wei Chen, Xin Luo, Yuanming Gao, Xiaogang Jia, and Zhiying Wang. "Attention unet++: A nested attention-aware u-net for liver ct image segmentation." In: *2020 IEEE International Conference on Image Processing (ICIP)*. IEEE. 2020, pp. 345–349.

[156] Ozan Oktay, Jo Schlemper, Loic Le Folgoc, Matthew Lee, Mattias Heinrich, Kazunari Misawa, Kensaku Mori, Steven McDonagh, Nils Y Hammerla, Bernhard Kainz, et al. "Attention u-net: Learning where to look for the pancreas." In: *arXiv preprint arXiv:1804.03999* (2018).

[157] Kenji Suzuki. "Overview of deep learning in medical imaging." In: *Radiological physics and technology* 10.3 (2017), pp. 257–273.

[158] Md Shahin Ali, Md Sipon Miah, Jahurul Haque, Md Mahbubur Rahman, and Md Khairul Islam. "An enhanced technique of skin cancer classification using deep convolutional neural network with transfer learning models." In: *Machine Learning with Applications* 5 (2021), p. 100036.

[159] Vinayakumar Ravi, Harini Narasimhan, and Tuan D Pham. "EfficientNet-based convolutional neural networks for tuberculosis classification." In: *Advances in Artificial Intelligence, Computation, and Data Science: For Medicine and Life Science* (2021), pp. 227–244.

[160] Leif E Peterson. "K-nearest neighbor." In: *Scholarpedia* 4.2 (2009), p. 1883.

[161] Leo Breiman. "Random forests." In: *Machine learning* 45 (2001), pp. 5–32.

[162] Marti A. Hearst, Susan T Dumais, Edgar Osuna, John Platt, and Bernhard Scholkopf. "Support vector machines." In: *IEEE Intelligent Systems and their applications* 13.4 (1998), pp. 18–28.

[163] David G Kleinbaum, K Dietz, M Gail, Mitchel Klein, and Mitchell Klein. *Logistic regression*. Springer, 2002.

[164] AVLN Sujith, Guna Sekhar Sajja, V Mahalakshmi, Shibili Nuhmani, and B Prasanalakshmi. "Systematic review of smart health monitoring using deep learning and Artificial intelligence." In: *Neuroscience Informatics* 2.3 (2022), p. 100028.

[165] Daoliang Li and Ling Du. "Recent advances of deep learning algorithms for aquacultural machine vision systems with emphasis on fish." In: *Artificial Intelligence Review* (2022), pp. 1–40.

[166] Priya Rani, Shallu Kotwal, Jatinder Manhas, Vinod Sharma, and Sparsh Sharma. "Machine learning and deep learning based computational approaches in automatic microorganisms image recognition: methodologies, challenges, and developments." In: *Archives of Computational Methods in Engineering* 29.3 (2022), pp. 1801–1837.

[167] Fares Bougourzi, Fadi Dornaika, and Abdelmalik Taleb-Ahmed. "Deep learning based face beauty prediction via dynamic robust losses and ensemble regression." In: *Knowledge-Based Systems* 242 (2022), p. 108246.

[168] Zhe Li, Josue Ortega Caro, Evgenia Rusak, Wieland Brendel, Matthias Bethge, Fabio Anselmi, Ankit B Patel, Andreas S Tolias, and Xaq Pitkow. "Robust deep learning object recognition models rely on low frequency information in natural images." In: *PLOS Computational Biology* 19.3 (2023), e1010932.

[169] D Balamurugan, SS Aravinth, P Chandra Shaker Reddy, Ajay Rupani, and A Manikandan. "Multiview objects recognition using deep learning-based wrap-CNN with voting scheme." In: *Neural Processing Letters* 54.3 (2022), pp. 1495–1521.

[170] Sweta Bhattacharya, Siva Rama Krishnan Somayaji, Thippa Reddy Gadekallu, Mamoun Alazab, and Praveen Kumar Reddy Maddikunta. "A review on deep learning for future smart cities." In: *Internet Technology Letters* 5.1 (2022), e187.

[171] Abhishek Gupta, Alagan Anpalagan, Ling Guan, and Ahmed Shaharyar Khwaja. "Deep learning for object detection and scene perception in self-driving cars: Survey, challenges, and open issues." In: *Array* 10 (2021), p. 100057. ISSN: 2590-0056. DOI: https://doi.org/10.1016/j.array.2021.100057. URL: https://www.sciencedirect.com/science/article/pii/S2590005621000059.

[172] Ke Wang, Yong Wang, Bingjun Liu, and Junlan Chen. "Quantification of Uncertainty and Its Applications to Complex Domain for Autonomous Vehicles Perception System." In: *IEEE Transactions on Instrumentation and Measurement* 72 (2023), pp. 1–17. DOI: 10.1109/TIM.2023.3256459.

[173] Ze Liu, Yingfeng Cai, Hai Wang, Long Chen, Hongbo Gao, Yunyi Jia, and Yicheng Li. "Robust Target Recognition and Tracking of Self-Driving Cars With Radar and Camera Information Fusion Under Severe Weather Conditions." In: *IEEE Transactions on Intelligent Transportation Systems* 23.7 (2022), pp. 6640–6653. DOI: 10.1109/TITS.2021.3059674.

[174] M Kalpana Devi and M Mary Shanthi Rani. "Classification of Diabetic Retinopathy Using Ensemble of Machine Learning Classifiers with IDRiD Dataset." In: *Evolutionary Computing and Mobile Sustainable Networks: Proceedings of ICECMSN 2021*. Springer, 2022, pp. 291–303.

[175] Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. "Spatial pyramid pooling in deep convolutional networks for visual recognition." In: *IEEE transactions on pattern analysis and machine intelligence* 37.9 (2015), pp. 1904–1916.

[176] Alex Krizhevsky, Ilya Sutskever, and Geoffrey E Hinton. "Imagenet classification with deep convolutional neural networks." In: *Communications of the ACM* 60.6 (2017), pp. 84–90.

[177] François Chollet. "Xception: Deep learning with depthwise separable convolutions." In: *Proceedings of the IEEE conference on computer vision and pattern recognition*. 2017, pp. 1251–1258.

[178] Christian Szegedy, Wei Liu, Yangqing Jia, Pierre Sermanet, Scott Reed, Dragomir Anguelov, Dumitru Erhan, Vincent Vanhoucke, and Andrew Rabinovich. "Going deeper with convolutions." In: *Proceedings of the IEEE conference on computer vision and pattern recognition*. 2015, pp. 1–9.

[179] Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. "Deep residual learning for image recognition." In: *Proceedings of the IEEE conference on computer vision and pattern recognition*. 2016, pp. 770–778.

[180] Arash Heidari, Nima Jafari Navimipour, and Mehmet Unal. "Applications of ML/DL in the management of smart cities and societies based on new trends in information technologies: A systematic literature review." In: *Sustainable Cities and Society* 85 (2022), p. 104089. ISSN: 2210-6707. DOI: https://doi.org/10.1016/j.scs.2022.104089. URL: https://www.sciencedirect.com/science/article/pii/S2210670722004061.

[181] Zhihan Lv, Zengchen Yu, Shuxuan Xie, and Atif Alamri. "Deep learning-based smart predictive evaluation for interactive multimedia-enabled smart healthcare." In: *ACM Transactions on Multimedia Computing, Communications, and Applications (TOMM)* 18.1s (2022), pp. 1–20.

[182] Xuxin Chen, Ximin Wang, Ke Zhang, Kar-Ming Fung, Theresa C Thai, Kathleen Moore, Robert S Mannel, Hong Liu, Bin Zheng, and Yuchen Qiu. "Recent advances and clinical applications of deep learning in medical image analysis." In: *Medical Image Analysis* (2022), p. 102444.

[183] Fares Bougourzi, Cosimo Distante, Fadi Dornaika, and Abdelmalik Taleb-Ahmed. "PDAtt-Unet: Pyramid Dual-Decoder Attention Unet for Covid-19 infection segmentation from CT-scans." In: *Medical Image Analysis* 86 (2023), p. 102797.

[184] Fares Bougourzi, Cosimo Distante, Fadi Dornaika, and Abdelmalik Taleb-Ahmed. "Cnr-iemn-cd and cnr-iemn-csd approaches for covid-19 detection and covid-19 severity detection from 3d ct-scans." In: *Computer Vision–ECCV 2022 Workshops: Tel Aviv, Israel, October 23–27, 2022, Proceedings, Part VII*. Springer. 2023, pp. 593–604.

[185] Gao Huang, Zhuang Liu, Laurens Van Der Maaten, and Kilian Q Weinberger. "Densely connected convolutional networks." In: *Proceedings of the IEEE conference on computer vision and pattern recognition*. 2017, pp. 4700–4708.

[186] Christian Szegedy, Vincent Vanhoucke, Sergey Ioffe, Jon Shlens, and Zbigniew Wojna. "Rethinking the inception architecture for computer vision." In: *Proceedings of the IEEE conference on computer vision and pattern recognition*. 2016, pp. 2818–2826.

[187] Saining Xie, Ross Girshick, Piotr Dollár, Zhuowen Tu, and Kaiming He. "Aggregated residual transformations for deep neural networks." In: *Proceedings of the IEEE conference on computer vision and pattern recognition*. 2017, pp. 1492–1500.

[188] Ashish Vaswani, Noam Shazeer, Niki Parmar, Jakob Uszkoreit, Llion Jones, Aidan N Gomez, Łukasz Kaiser, and Illia Polosukhin. "Attention is all you need." In: *Advances in neural information processing systems* 30 (2017).

[189] Dzmitry Bahdanau, Kyunghyun Cho, and Yoshua Bengio. "Neural machine translation by jointly learning to align and translate." In: *arXiv preprint arXiv:1409.0473* (2014).

[190] Fahad Shamshad, Salman Khan, Syed Waqas Zamir, Muhammad Haris Khan, Munawar Hayat, Fahad Shahbaz Khan, and Huazhu Fu. "Transformers in medical imaging: A survey." In: *Medical Image Analysis* (2023), p. 102802.

[191] Alexey Dosovitskiy, Lucas Beyer, Alexander Kolesnikov, Dirk Weissenborn, Xiaohua Zhai, Thomas Unterthiner, Mostafa Dehghani, Matthias Minderer, Georg Heigold, Sylvain Gelly, et al. "An image is worth 16x16 words: Transformers for image recognition at scale." In: *arXiv preprint arXiv:2010.11929* (2020).

[192] Hugo Touvron, Matthieu Cord, Matthijs Douze, Francisco Massa, Alexandre Sablayrolles, and Hervé Jégou. "Training data-efficient image transformers & distillation through attention." In: *International conference on machine learning*. PMLR. 2021, pp. 10347–10357.

[193] Jonah Mushava and Michael Murray. “A novel XGBoost extension for credit scoring class-imbalanced data combining a generalized extreme value link and a modified focal loss function.” In: *Expert Systems with Applications* 202 (2022), p. 117233.

****