



**HAL**  
open science

# Evaluation de l'apport d'une approche de localisation spatiale pour la détection par DEEP LEARNING des nodules thoraciques chez le chien et le chat

Eve Solan

► **To cite this version:**

Eve Solan. Evaluation de l'apport d'une approche de localisation spatiale pour la détection par DEEP LEARNING des nodules thoraciques chez le chien et le chat. Sciences du Vivant [q-bio]. 2024. dumas-04780980

**HAL Id: dumas-04780980**

**<https://dumas.ccsd.cnrs.fr/dumas-04780980v1>**

Submitted on 13 Nov 2024

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

**Oniris - École Nationale Vétérinaire, Agroalimentaire et de  
l'Alimentation**

ANNEE 2024 - Thèse n° 59

# **ÉVALUATION DE L'APPORT D'UNE APPROCHE DE LOCALISATION SPATIALE POUR LA DÉTECTION PAR DEEP LEARNING DES NODULES THORACIQUES CHEZ LE CHIEN ET LE CHAT**

## **THÈSE**

pour l'obtention du diplôme d'État de

## **DOCTEUR VÉTÉRINAIRE**

présentée et soutenue publiquement devant l'UFR de

Médecine de l'Université de Nantes

le 26/09/2024

par

**SOLAN EVE**

Sous la direction de

**GUATTEO Raphael**

Président du jury : Madame Marie-Anne COLLE, Professeur à l'École Nationale Vétérinaire,  
Agroalimentaire et de l'alimentation de Nantes - ONIRIS

Membres du jury : Monsieur Raphaël GUATTEO, Professeur à l'École Nationale Vétérinaire,  
Agroalimentaire et de l'alimentation de Nantes – ONIRIS

Madame Marion FUSELLIER, Professeur à l'École Nationale Vétérinaire, Agroalimentaire  
et de l'alimentation de Nantes - ONIRIS



**Oniris - École Nationale Vétérinaire, Agroalimentaire et de  
l'Alimentation**

ANNEE 2024 - Thèse n° 59

# **ÉVALUATION DE L'APPORT D'UNE APPROCHE DE LOCALISATION SPATIALE POUR LA DÉTECTION PAR DEEP LEARNING DES NODULES THORACIQUES CHEZ LE CHIEN ET LE CHAT**

## **THÈSE**

pour l'obtention du diplôme d'État de

## **DOCTEUR VÉTÉRINAIRE**

présentée et soutenue publiquement devant l'UFR de

Médecine de l'Université de Nantes

le 26/09/2024

par

**SOLAN EVE**

Sous la direction de

**GUATTEO Raphaël**

Président du jury : Madame Marie-Anne COLLE, Professeur à l'École Nationale Vétérinaire, Agroalimentaire et de l'alimentation de Nantes - ONIRIS

Membres du jury : Monsieur Raphael GUATTEO, Professeur à l'École Nationale Vétérinaire, Agroalimentaire et de l'alimentation de Nantes - ONIRIS

Madame Marion FUSELLIER, Professeur à l'École Nationale Vétérinaire, Agroalimentaire et de l'alimentation de Nantes - ONIRIS



## LISTE DES MEMBRES DU CORPS ENSEIGNANT

<b>Département <b>BPSA</b> Biologie, Pathologie et Sciences de l'Aliment</b>		
<b>Responsable : Emmanuel JAFFRES – Adjointe : Frédérique NGUYEN</b>		
Pharmacologie et Toxicologie	Jean-Claude DESFONTIS (Pr) Yassine MALLEM (Pr) Hervé POULIQUEN (Pr)	Antoine ROSTANG (MC) Meg-Anne MORICEAU (CERC) Martine KAMMERER (Pr émérite) Marc GOGNY (Pr émérite)
Physiologie fonctionnelle, cellulaire et moléculaire	Jean-Marie BACH (Pr) Lionel MARTIGNAT (Pr) Julie HERVE (MC HDR) Grégoire MIGNOT (MC)	Solenn GAVAUD (CERC)
Histologie et anatomie pathologique	Marie-Anne COLLE (Pr) Jérôme ABADIE (MC)	Laetitia JAILLARDON (MC) Frédérique NGUYEN (MC)
Biochimie alimentaire industrielle	Carole PROST (Pr) Joëlle GRUA (MC)	Clément CATANEO (MC) Alix KHALIL (MC) Laurent LE THUAUT (MC)
Microbiotech	Hervé PREVOST (Pr) Géraldine BOUE (MC) Nabila HADDAD (MC) Emmanuel JAFFRES (MC)	Mathilde MOSSER (MC) Boris MISERY (MC) Raouf TAREB (MC) Quentin PRUVOST (CEC)
PACENV = VET1	Eléonore BOUGUYON (PRAG) Nicolas BROSSAUD (PRAG)	Charlotte MOCQUARD (PRAG) Aurore CALVEL (PRAG)
<b>Département <b>SAESP</b> Santé des Animaux d'Élevage et Santé Publique</b>		
<b>Responsable : Raphaël GUATTEO – Adjoint : Jean-Michel CAPPELIER</b>		
Élevage, nutrition et santé des animaux domestiques	Nathalie BAREILLE (Pr) François BEAUDEAU (Pr) Christine FOURICHON (Pr)	Juan Manuel ARIZA CHACON (MC) Ségolène CALVEZ 5(Pr) Aurélien MADOUASSE (MC) Nora NAVARRO-GONZALES (MC)
Infectiologie	Alain CHAUVIN (Pr) Emmanuelle MOREAU (Pr) Nathalie RUVOEN-CLOUET (Pr) Pauline MAISONNASSE (CERC)	Albert AGOULON (MC) Suzanne BASTIAN (MC) Léa LOISEL (AERC) Kenny OBERLE (MC) Nadine RAVINET (MC)
Médecine des animaux d'élevage	Catherine BELLOC (Pr) Christophe CHARTIER (Pr émérite) Raphaël GUATTEO (Pr)	Sébastien ASSIE (MC) Isabelle BREYTON (MC) Mily LEBLANC MARIDOR (MC)

	Anne RELUN (MC)	Maud ROUAULT (AERC)
Hygiène et qualité des aliments	Jean-Michel CAPPELIER (Pr) Louis DELAUNAY (CERC) Bruno LE BIZEC (Pr) Marie-France PILET (Pr)	Sofia STRUBBIA (MC)

## Département **DSC** Sciences cliniques

Responsable : **Catherine IBISCH** – Adjoint : **Marion FUSELLIER**

Anatomie comparée	Eric BETTI (MC) Claude GUINTARD (MC) Margarida RIBEIRO DA SILVA NEUNLIST (MC Stagiaire)	
Pathologie chirurgicale et anesthésiologie	Eric AGUADO (Pr) Olivier GAUTHIER (Pr) Eric GOYENVALLE (MC HDR)	Pierre MAITRE (MC) Caroline TESSIER (MC) Gwénola TOUZOT-JOURDE (MC) Claire DEFOURMSTRAUX (MC)
Dermatologie, parasitologie des carnivores et des équidés, mycologie	Jacques GUILLOT (Pr) Emmanuel BENSIGNOR (Pr Ass)	Sabrina VIEU (AERC) Maria Dolores SANCHEZ (AERC)
Médecine interne, imagerie médicale et législation professionnelle vétérinaire	Anne COUROUCE (Pr) Jack-Yves DESCHAMPS (Pr) Françoise ROUX (Pr) Juan HERNANDEZ-RODRIGUEZ (Pr Ass) Nora BOUHSINA (MC)	Nicolas CHOUIN (MC) Amandine DRUT (MC) Marion FUSELLIER-TESSON (Pr) Catherine IBISCH (MC) Aurélia LEROUX (MC) Odile SENECAT (MC)
Biotechnologies et pathologies de la reproduction	Jean-François BRUYAS (Pr) François FIENI (Pr)	Djemil BENCHARIF (Pr) Lamia BRIAND (Pr)

## Département **GPA** Génie des procédés alimentaires

Responsable : **Sébastien CURET-PLOQUIN** – Adjointe : **Vanessa JURY**

Lionel BOILLEREAUX (Pr) Sébastien CURET-PLOQUIN (Pr) Marie DE LAMBALLERIE (Pr) Francine FAYOLLE (Pr) Michel HAVET (Pr)	Alain LEBAIL (Pr) Olivier ROUAUD (Pr) Kévin CROUVISIER-URION (MC) Vanessa JURY (Pr) Emilie KORBEL (MC)	Jean-Yves MONTEAU (MC HDR) Eve-Anne NORWOOD (MC) Raphaël PORYLES (MC) Laurence POTTIER (MC) Cyril TOUBLANC (MC)
PAC-ING	Cyril Gaillard (PCEA)	

## Département **MSC** Management, statistiques et communication

Responsable : **Samira ROUSSELIERE** – Adjointe : **Véronique CARIOU**

Mathématiques, statistiques, informatique	El Mostafa QANNARI (Pr émérite) Chantal THORIN (Pr Ag) Evelyne VIGNEAU (Pr) Jean-Michel GALHARRET(MC stagiaire)	Véronique CARIOU (Pr) Benjamin MAHIEU (MC) Michel SEMENOU (MC)
Économie, gestion, législation	Jean-Marc FERRANDI (Pr) Pascal BARILLOT (MC) Ibrahima BARRY (MC) Florence BEAUGRAND (MC)	Sibylle DUCHAINE (MC) Sonia MAHJOUR (MC) Samira ROUSSELIERE (MC) Christophe PAPINEAU (Ens. Cont.)
Langues et communication	Marc BRIDOU (PLPA) David GUYLER (Ens. Cont.) Nathalie GOODENOUGH (PCEA) Patricia JOSSE (Ens. Cont.)	Shaun MEEHAN (Ens. Cont.) Linda MORRIS (PCEA) Ian NICHOLSON (ENS. Cont.)

Pr Ag : Professeur Agrégé, Pr : Professeur, MC : Maître de Conférence, MCC : MC contractuel,  
 PLPA : Professeur Lycée Professionnel Agricole, PCEA : Professeur Certifié Enseignement Agricole,  
 HDR : Habilité à Diriger des Recherches, CERC : Chargé d'Enseignement et de Recherche Contractuel, Ens. Cont. : Enseignant Contractuel

La reproduction d'extraits de cette thèse est autorisée avec mention de la source. Toute reproduction partielle doit être fidèle au texte utilisé. Cette thèse devra donc être citée en incluant les éléments bibliographiques suivants :

- Nom et prénoms de l'auteur : SOLAN Eve
- Année de soutenance : 2024.
- Titre de la thèse : Évaluation de l'apport d'une approche de localisation spatiale pour la détection, par *Deep Learning*, des nodules thoraciques chez le chien et le chat
- Intitulé du diplôme : Thèse de doctorat vétérinaire
- Université de soutenance : Faculté de Médecine de Nantes
- École de soutenance : Oniris VetAgroBio
- Nombre de pages : 88p.

Le défaut de citation est considéré comme du plagiat. Ce dernier est puni par la loi française et passible de sanctions allant jusqu'à 3 ans d'emprisonnement et 300 000 € d'amende.

# Remerciements

**À Madame Marie-Anne Colle,**

Professeur à Oniris VetAgroBio - École Nationale Vétérinaire de Nantes,  
Pour m'avoir fait l'honneur d'accepter de présider le jury de cette thèse,  
Sincères remerciements.

**À Monsieur Raphaël Guatteo,**

Professeur à Oniris VetAgroBio - École Nationale Vétérinaire de Nantes,  
Pour m'avoir proposé ce sujet, pour avoir été disponible tout au long de ce projet, et  
pour tous les conseils que vous avez pu me donner,  
Sincères remerciements.

**À Madame Marion Fusellier,**

Professeur à Oniris VetAgroBio - École Nationale Vétérinaire de Nantes,  
Pour avoir accepté de faire partie de ce jury, pour votre enseignement et pour avoir  
accepté de m'accueillir en stage dans votre service,  
Sincères remerciements.

**À Monsieur Benjamin Toni,**

Développeur WEB pour PicoxIA,  
Pour m'avoir suivie dans ce projet, et pour m'avoir fait rencontrer l'équipe,  
Sincères remerciements.

**À Monsieur Guilhem Codron,**

Développeur WEB pour PicoxIA,  
Pour avoir été mon interlocuteur pendant tout ce projet, pour avoir pris le temps de  
résoudre tous les petits problèmes techniques, pour tes conseils, pour avoir codé cet  
algorithme,  
Sincères remerciements.

**À Monsieur Awadi Bedja,**

Développeur en apprentissage pour PicoxIA  
Pour avoir transformé mes idées en une réalité, pour avoir codé cet algorithme,  
Sincères remerciements.

**À Monsieur Dorian Piette,**

Chargé des affaires juridiques à Oniris VetAgroBio - École Nationale Vétérinaire de Nantes,  
Pour avoir rendu cette collaboration possible,  
Sincères remerciements.



# **Table des matières**

<b>REMERCIEMENTS.....</b>	<b>7</b>
<b>TABLE DES ANNEXES .....</b>	<b>11</b>
<b>TABLE DES ILLUSTRATIONS (FIGURES ET TABLEAUX) .....</b>	<b>12</b>
<b>TABLE DES ABREVIATIONS ET DES SIGLES .....</b>	<b>13</b>
<b>INTRODUCTION.....</b>	<b>15</b>
<b>PARTIE 1 – LES TECHNIQUES D’IMAGERIE UTILISEES DANS LE BILAN D’EXTENSION THORACIQUE.....</b>	<b>17</b>
<b>I) Rappels sur les principaux cancers des carnivores domestiques et leur potentiel métastatique.....</b>	<b>17</b>
a) Les mécanismes de la cancérisation .....	17
b) Prévalence métastatique des principaux cancers du chien et du chat .....	19
<b>II) La radiographie .....</b>	<b>21</b>
a) Principe de la radiographie .....	21
b) La radiographie thoracique .....	22
c) Avantages de la radiographie .....	23
d) Limites de la radiographie .....	23
<b>III) L’échographie.....</b>	<b>24</b>
a) Principe de l’échographie.....	24
b) L’échographie du thorax, excluant le cœur.....	25
c) Avantages de l’échographie.....	26
d) Limites de l’échographie .....	27
<b>IV) Le scanner .....</b>	<b>27</b>
a) Principe du scanner .....	27
b) Tomographie thoracique .....	28
c) Avantages du scanner .....	28
d) Limites du scanner.....	29
<b>V) Synthèse .....</b>	<b>30</b>
<b>PARTIE 2 – APPORTS ACTUELS DES INTELLIGENCES ARTIFICIELLES EN MEDECINE VETERINAIRE .....</b>	<b>31</b>
<b>I) Préalable à l’utilisation des IA.....</b>	<b>31</b>
a) Définitions et concepts.....	31
b) Organisation des réseaux de Deep Learning .....	32
c) Le DL appliqué au traitement d’images .....	35
d) Entraînement des algorithmes de Deep Learning .....	36
i. Les données nécessaires au fonctionnement de l’IA .....	36
ii. Les hyperparamètres.....	37
<b>II) Les grands secteurs d’utilisation des IA en médecine vétérinaire .....</b>	<b>37</b>
a) Épidémiologie .....	37
b) Comportement des animaux de compagnie .....	38
c) Cytologie et biochimie des carnivores domestiques .....	39
d) Suivi automatisé de l’activité électrique cardiaque et cérébrale .....	40
<b>III) Extension au domaine de l’imagerie médicale.....</b>	<b>41</b>
a) Historique.....	41
b) Applications actuelles .....	41
i. Imagerie osseuse .....	41
ii. Imagerie cardiaque.....	42
iii. Imagerie hépatique.....	42
iv. Imagerie du système nerveux (central et périphérique) .....	43
v. Imagerie pulmonaire (autres que les masses et nodules).....	44
c) Utilisation des IA dans le bilan d’extension thoracique : constats et perspectives .....	45

<b>PARTIE 3 – EXECUTION D’UN ALGORITHME DE DEEP LEARNING POUR LA LOCALISATION DE NODULES THORACIQUES .....</b>	<b>49</b>
<b>I) Contexte et enjeux de notre étude.....</b>	<b>49</b>
<b>II) Constitution de la base de données nécessaire à l’entraînement de l’algorithme de Deep Learning.....</b>	<b>50</b>
a) Estimation du nombre d’images nécessaires .....	50
b) Population d’étude .....	50
c) Critères d’inclusion.....	54
d) Récupération des clichés radiographiques et constitution de notre échantillon d’étude.....	56
<b>III) Annotations des radiographies en vue de la phase d’apprentissage de l’algorithme .....</b>	<b>58</b>
a) Logiciels utilisés pour l’annotation et l’apprentissage .....	58
i. Visual Geometry Group Image Annotator (VIA) .....	58
ii. Python.....	59
b) Annotations manuelles des radiographies en vue de la localisation automatisée des métastases thoraciques .....	59
c) Développement d’un algorithme de détection de nodules thoraciques .....	60
i. Suppression Non Maximale.....	60
ii. Annotation automatisée des images.....	61
iii. Récupération des coordonnées des Bboxes.....	62
iv. Fonctions d’exécution du programme.....	62
<b>IV) Paramètres nécessaires à l’exécution de l’algorithme de Deep Learning et à l’évaluation de ses performances. ....</b>	<b>63</b>
a) Choix des hyperparamètres.....	63
b) Métriques utilisées pour l’évaluation des performances du modèle de détection.....	64
i. Faux positifs, Faux négatifs, Vrais positifs, et rappel.....	64
ii. Précision et moyenne de la précision moyenne .....	65
iii. Score F1 .....	66
<b>V) Résultats obtenus à l’issue des différents apprentissages .....</b>	<b>66</b>
a) Sur la base de données de 711 images .....	66
b) Sur 7 lots distincts d’images issues de la banque de clichés .....	67
c) Sur une base de données obtenue d’après les meilleurs lots .....	68
<b>VI) Discussion.....</b>	<b>70</b>
a) Performances globales de notre algorithme de détection.....	70
b) Détection des nodules de petite taille .....	71
c) Comparaison des performances de l’IA sur les deux jeux de données .....	72
<b>CONCLUSION .....</b>	<b>75</b>
<b>BIBLIOGRAPHIE.....</b>	<b>77</b>
<b>ANNEXE I : SCRIPT PYTHON "PREDICT.PY" DANS SON INTEGRALITE.....</b>	<b>83</b>
<b>ANNEXE II : DETAILS SUR LA TECHNIQUE DE SUPER CONVERGENCE .....</b>	<b>87</b>

## **Table des annexes**

ANNEXE I : SCRIPT PYTHON "PREDICT.PY" DANS SON INTEGRALITE .....	83
ANNEXE II : DETAILS SUR LA TECHNIQUE DE SUPER CONVERGENCE .....	87

# Table des illustrations (figures et tableaux)

<b>FIGURE 1.</b> SCHEMA SIMPLIFIE DE LA FORMATION D'UNE TUMEUR ET DE SA DISSEMINATION ; EXEMPLE DES METASTASES PULMONAIRES .....	17
<b>FIGURE 2.</b> SCHEMA SIMPLIFIE DE LA FORMATION D'UNE TUMEUR ET DE SA DISSEMINATION (SUITE) ; EXEMPLE DES METASTASES PULMONAIRES.....	18
<b>FIGURE 3.</b> SCHEMA DE LA FORMATION D'UNE IMAGE RADIOGRAPHIQUE.....	21
<b>FIGURE 4.</b> RADIOGRAPHIES DE POUMONS SAINS AVEC DELIMITATION SCHEMATIQUE DES LOBES.....	22
<b>FIGURE 5.</b> SCHEMA DE LA FORMATION D'UNE IMAGE ECHOGRAPHIQUE.....	24
<b>FIGURE 6.</b> IMAGE ECHOGRAPHIQUE D'UN NODULE PULMONAIRE DE 8MM DE DIAMETRE CHEZ UN CHIEN.....	26
<b>FIGURE 7.</b> REPRESENTATION SCHEMATIQUE D'UN RESEAU DE NEURONES A 2 COUCHES CACHEES.....	32
<b>FIGURE 8.</b> ILLUSTRATION DU FONCTIONNEMENT D'UNE FONCTION DE CONVOLUTION: PARCOURS DU FILTRE SUR L'IMAGE.....	34
<b>FIGURE 9.</b> ILLUSTRATION DU FONCTIONNEMENT DE LA FONCTION MAX-POOLING .....	34
<b>FIGURE 10.</b> REPRESENTATION DES DIFFERENTES TACHES DU CV .....	35
<b>FIGURE 11.</b> REPARTITION DES AGES DES 352 CHIENS A LA PREMIERE RADIOGRAPHIE.....	51
<b>FIGURE 12.</b> REPARTITION DES AGES DES 161 CHATS A LA PREMIERE RADIOGRAPHIE.....	51
<b>FIGURE 13.</b> EXAMENS COMPLEMENTAIRES REALISES POUR CONFIRMER LA SUSPICION DE METASTASES ET LE NOMBRE DE CAS RECENSES POUR CHACUN SUR LES 162 ANIMAUX CONCERNES.....	54
<b>FIGURE 14.</b> DIAGRAMME DE FLUX RECAPITULANT LES CRITERES APPLIQUES SUR LA POPULATION D'ETUDE INITIALE DE 513 ANIMAUX ET QUI ONT CONDUIT A L'OBTENTION D'UN ECHANTILLON DE 393 CAS.....	55
<b>FIGURE 15.</b> FICHIER JSON DES IMAGES A ET B EXPORTER DEPUIS VIA.....	58
<b>FIGURE 16.</b> EXEMPLES DE DESSIN MANUEL DE BOUNDING BOXES .....	60
<b>FIGURE 17.</b> CODE PYTHON DE LA FONCTION "APPLY_NMS".....	60
<b>FIGURE 18.</b> ILLUSTRATION MATHEMATIQUE DE L'IOU.....	61
<b>FIGURE 19.</b> CODE PYTHON DE LA FONCTION "DISPLAY_PREDICTIONS".....	61
<b>FIGURE 20.</b> EXEMPLE D'IMAGES ANNOTEES EN SORTIE DE LA FONCTION "DISPLAY_PREDICTIONS".....	62
<b>FIGURE 21.</b> CODE PYTHON DE LA FONCTION "PREDICTIONS_TO_VIA_REGIONS".....	62
<b>FIGURE 22.</b> EXEMPLE D'UNE RADIOGRAPHIE UTILISEE POUR CALCULER LES PERFORMANCES DE L'ALGORITHME ..	65
<b>TABLEAU 1.</b> PRINCIPALES TUMEURS DU CHIEN ET DU CHAT, FREQUENCE ET PREVALENCE METASTATIQUE PULMONAIRE.....	20
<b>TABLEAU 2.</b> TABLEAU RECAPITULATIF POUR LE CHOIX DES EXAMENS D'IMAGERIE DANS LE BILAN D'EXTENSION THORACIQUE.....	30
<b>TABLEAU 3.</b> REPARTITION DES DIFFERENTES TUMEURS PRIMITIVES RENCONTREES DANS LES 2 ESPECES ETUDIEES ET PROPORTION DANS LA POPULATION D'ETUDE.....	53
<b>TABLEAU 4.</b> REPARTITION DES CAS POUR L'ESPECE CANINE ET FELINE EN FONCTION DU TYPE DE TUMEUR PULMONAIRE ET PROPORTION DANS L'ECHANTILLON DE 26 ANIMAUX ATTEINTS D'UNE TUMEUR PULMONAIRE PRIMITIVE.....	53
<b>TABLEAU 5.</b> REPARTITION DES RADIOGRAPHIES DE L'ETUDE EN FONCTION DE LA TAILLE DES MASSES.....	57
<b>TABLEAU 6.</b> REPARTITION DES CLICHES RADIOGRAPHIQUES DANS CHAQUE LOT EN FONCTION DE LA TAILLE DES METASTASES, EN POURCENTAGE.....	57
<b>TABLEAU 7.</b> PERFORMANCES METRIQUES DU MODELE SUR 711 CLICHES.....	66
<b>TABLEAU 8.</b> RESULTATS DES PERFORMANCES DU MODELE POUR CHAQUE LOT D'IMAGES.....	67
<b>TABLEAU 9.</b> PERFORMANCES DU MODELE DE DETECTION SUR LES MEILLEURS LOTS.....	68

# Table des abréviations et des sigles

BE : Bilan d'extension  
CAD: *Computer Aided Design*  
CADe : *Computer Aided Detection*  
CADx : *Computer Aided Diagnosis*  
*cf.* : *confer*  
CR : Compte rendu  
CV : *Computer Vision*  
Ddx : Diagnostic différentiel  
DICOM : *Digital Imaging and Communications in Medicine*  
DIV : Disque Intervertebral  
DL : *Deep Learning*  
DLD/DLG : Décubitus latéral droit/Décubitus latéral gauche  
ECG : Électrocardiogramme  
EEG : Électroencéphalogramme  
EPI : Équipement de protection individuelle  
FN : Faux négatif  
FP : Faux positif  
GPU: *Graphic Processing Unit*  
IA : Intelligence Artificielle  
*i.e.* : « *id est* », c'est-à-dire  
IoU : *Intersection over Union*  
IRM : Imagerie Résonance Magnétique  
JPEG : *Joint Photographic Expert Group*  
JSON : *JavaScript Object Notation*  
mAP : *mean average precision*  
MICI : Maladie inflammatoire chronique de l'intestin  
ML : *Machine Learning*  
MRC : Maladie rénale chronique  
NFS : Numération formule sanguine  
NMS : *Non-maximal suppression*  
RGBA : *Red Green Blue Alpha*  
RNC : Réseau de Neurones Convolutif  
ROI : *Region of interest*  
RVB : Rouge Vert Bleu  
Se : Sensibilité  
SGD : *Stochastic Gradient Descent*  
SNC : Système nerveux central  
TNM : Tumor, Node, Metastasis  
VD/DV : Ventro-Dorsal/Dorso-ventral  
VHS : *Vertebral Heart Scale*  
VIA : *Visual Geometry Group Image Annotation*  
VP : Vrai positif



# Introduction

Malgré la difficulté à donner une incidence et une prévalence exactes, la *Veterinary Cancer Society* estime qu'au moins 1 chien sur 4 et 1 chat sur 5 feront l'objet d'un diagnostic de cancer<sup>1</sup>. Il s'agirait de la première cause de mortalité chez le chien, et de la troisième chez le chat<sup>2</sup>, représentant la principale cause de décès chez 47% des chiens et chez 32% des chats. Il est clair que les propriétaires ont donc une attente croissante quant à l'expertise de leur vétérinaire dans le domaine de l'oncologie, autant pour la détection des tumeurs primitives que pour leur bilan d'extension.

Le propre d'un cancer est de se propager. Chaque processus tumoral a, en fonction de sa nature, un potentiel métastatique local, régional ou général donné. Sur la base de ses critères, la classification *Tumor Node Metastasis (TNM)* de l'Organisation Mondiale de la Santé a été adaptée à la médecine vétérinaire et est très largement utilisée. L'intérêt de cette méthode est de déterminer le stade cancéreux afin d'affiner le pronostic de la maladie et d'en adapter la thérapeutique<sup>3,4</sup>.

La découverte de métastases lors du bilan d'extension assombrit nettement le pronostic. En médecine humaine, c'est d'ailleurs la principale cause de mortalité chez les patients atteints d'un cancer<sup>3,5</sup>. Chez les carnivores domestiques, les poumons sont le site privilégié pour le développement des métastases, lors de dissémination par voie sanguine<sup>3</sup>. Même si tous les mécanismes de cancérisation ne sont pas élucidés, il est communément admis que l'installation des métastases est possible par l'achèvement de nombreuses étapes, permettant à une cellule tumorale de se détacher de la tumeur primitive et de s'implanter dans un tissu cible. Mais il est également probable que la progression de ces métastases, dans l'organisme ou au sein d'un même tissu, fasse suite à la répétition de ce processus entraînant le développement de métastases à partir de métastases<sup>6</sup>. Il convient donc de les diagnostiquer le plus précocement possible.

Dès la fin des années 1990, le domaine de l'imagerie médicale humaine s'aide des premières Intelligences Artificielles (IA) pour la lecture et l'interprétation de ses données<sup>7</sup>. Aujourd'hui, ce sont des centaines d'IA présentes sur le marché et qui assistent au quotidien les médecins dans la réalisation, l'analyse et l'interprétation de radiographies<sup>8</sup>. L'imagerie médicale vétérinaire suit aujourd'hui le même chemin. Au vu de l'importance d'une détection précoce des lésions pulmonaires, telles que les nodules thoraciques, des approches de classification de clichés radiographiques par des algorithmes intelligents ont déjà été étudiées. Ces derniers sont basés sur l'utilisation de réseaux de neurones entraînés à prédire la présence ou l'absence de métastases. Cependant ils ne produisent pas de performances satisfaisantes, contrairement à d'autres anomalies<sup>9,10</sup>. Ces résultats décevants ont pour origine une base de données d'entraînement trop petite, l'IA pouvant alors confondre ces nodules avec d'autres lésions voire avec d'autres structures semblables en forme (calcification des jonctions chondro-costales, mamelles, bronches et bronchioles, vaisseaux sanguins...). La diversité de leur taille et de leur localisation est également un obstacle<sup>11</sup>.

Le but de notre étude était de déterminer si une autre approche, telle que l'approche de localisation spatiale, produirait des performances significativement meilleures que l'approche par classification pour la détection des métastases pulmonaires. Ce projet a été mené en partenariat avec l'entreprise PicoxIA, une start-up française de la société FAS, spécialisée dans le développement d'algorithmes d'aide au diagnostic radiographique. L'étude bibliographique préalable rappelle les principaux cancers des carnivores domestiques et les techniques d'imagerie utiles à leur diagnostic, puis présente les principes fondamentaux nécessaires à la compréhension du fonctionnement des IA ainsi que leurs applications actuelles en santé animale et en médecine vétérinaire.

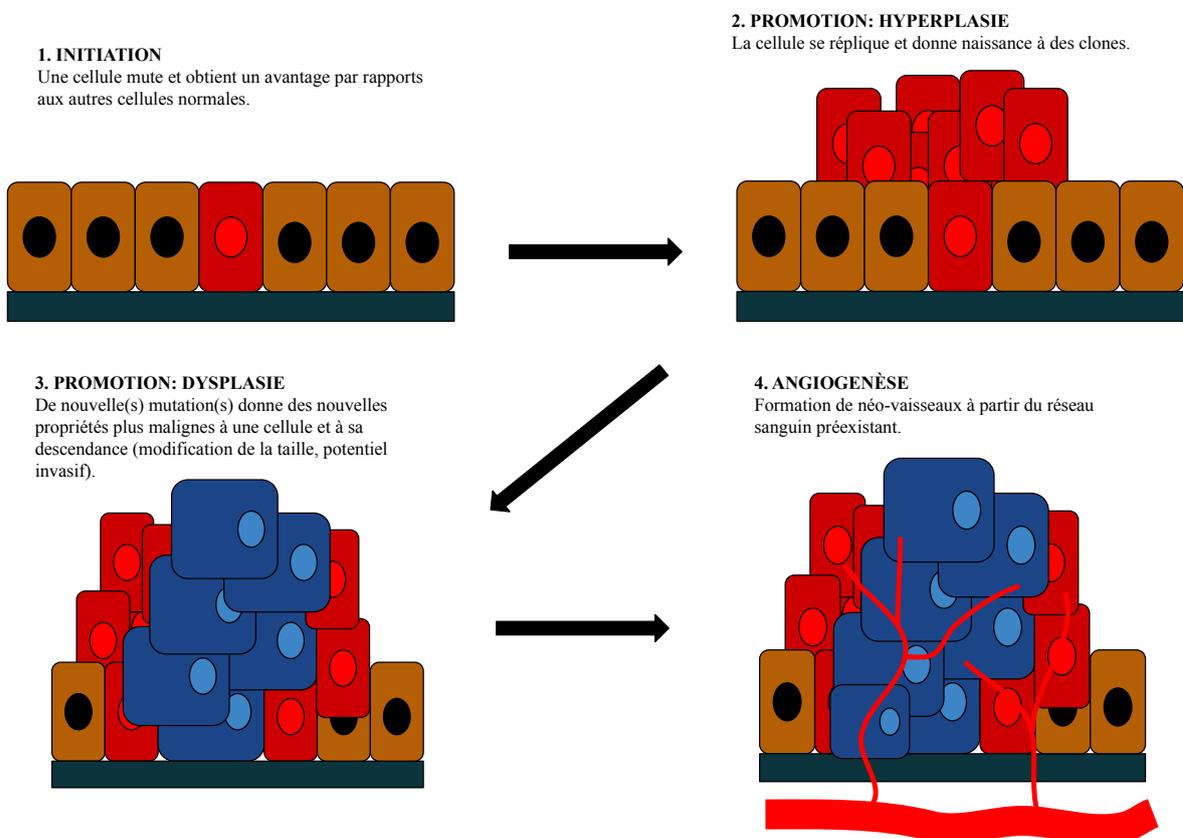
# Partie 1 – Les techniques d'imagerie utilisées dans le bilan d'extension thoracique

## l) Rappels sur les principaux cancers des carnivores domestiques et leur potentiel métastatique

### a) Les mécanismes de la cancérisation

Le terme « cancer » désigne un ensemble de maladies qui ont pour seuls points communs une croissance et une prolifération incontrôlables. Pour expliquer l'origine d'une tumeur cancéreuse, on utilise communément le modèle « initiation, promotion et généralisation »<sup>6</sup> (cf. **Figure 1**) :

- **Initiation** : une seule et unique cellule va subir des lésions irréversibles mais non létales de son ADN, par l'action d'agents extérieurs appelés agents cancérogènes (agents physiques, chimiques ou biologiques, traumatismes répétés...). Cette transformation va offrir à la cellule somatique un avantage par rapport aux cellules environnantes, comme la capacité de se répliquer de manière illimitée.
- **Promotion** : un autre événement va venir renforcer cette supériorité de la cellule mutée par rapport à ses voisines. Par multiplication, elle donne naissance à des clones de cellules transformées et c'est ainsi qu'elle prolifère localement pour donner une tumeur primitive.
- **Généralisation** : un troisième et dernier événement vient renforcer le potentiel malin de la cellule mutée. Elle a dorénavant la possibilité de détruire et d'envahir localement les tissus, mais aussi de métastaser à distance.



**Figure 1.** Schéma simplifié de la formation d'une tumeur et de sa dissémination ; exemple des métastases pulmonaires. *Adaptation personnelle de l'auteur du manuscrit, d'après Dobson et Lascelles, 2016.*

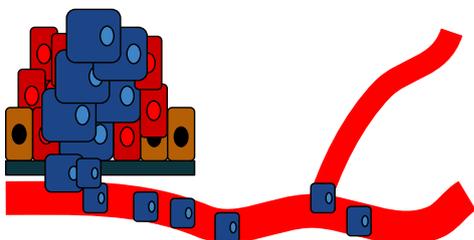
Cette théorie qui veut qu'une population de cellules cancéreuses soient toutes issues d'une même cellule cancéreuse initiale est appelée « théorie monoclonale »<sup>12</sup>. Attention, il est important de noter que ce modèle est très simplifié, et scientifiquement peu réaliste car les mutations sont des événements stochastiques qui ne surviennent pas étape par étape. Il permet néanmoins d'apprécier les grands événements de la cancérogenèse.

Les métastases sont définies comme la capacité d'une tumeur maligne à se disséminer vers des organes à distance de la masse primitive et d'y développer de nouvelles tumeurs<sup>13</sup>. Des études ont montré qu'au cours d'un cancer, une tumeur peut subir de multiples modifications. Les cellules de la masse primitive peuvent à leur tour subir des mutations génétiques qu'elles transmettront à leur descendance. C'est pourquoi toutes les cellules d'une même tumeur n'ont pas le même potentiel métastatique, la dissémination semble donc être un caractère acquis pendant l'évolution de la maladie<sup>5</sup>. Pour que ce processus puisse arriver, il faut que certaines cellules cancéreuses achèvent une série d'évènements<sup>6</sup> (cf. **Figure 2**) :

- Détachement des cellules métastatiques de la tumeur primitive.
- Pénétration dans le système sanguin ou lymphatique.
- Lors de leur circulation, les cellules doivent survivre à l'anoïkis (apoptose provoquée par la perte de contact entre une cellule et son environnement) et échapper au système immunitaire.
- Arrêt au niveau de l'organe cible, sortie du système circulatoire et pénétration du tissu.
- Survie dans le nouveau microenvironnement et formation d'une tumeur secondaire.

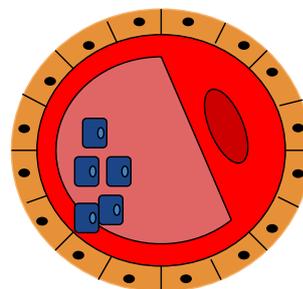
#### 5. GÉNÉRALISATION

Détachement des cellules de la tumeur primaire, pénétration dans le système lymphatique ou sanguin et mise en circulation



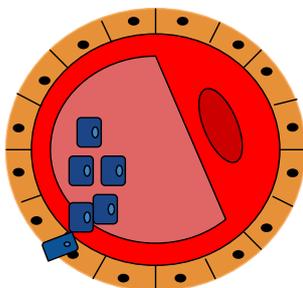
#### 6. GÉNÉRALISATION

Arrêt des cellules cancéreuses dans les vaisseaux de l'organe cible (ici les poumons) et adhérence à la paroi



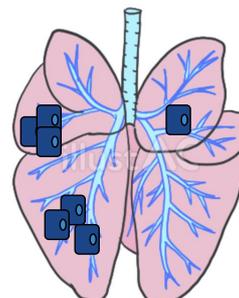
#### 7. GÉNÉRALISATION

Extravasation des cellules vers le tissu pulmonaire



#### 8. GÉNÉRALISATION

Prolifération des cellules tumorales dans leur nouvel environnement, angiogenèse et installation des métastases



**Figure 2.** Schéma simplifié de la formation d'une tumeur et de sa dissémination (suite) ; exemple des métastases pulmonaires. *Adaptation personnelle de l'auteur du manuscrit, d'après Withrow and MacEwen, 2007.*

Chacune de ces étapes peut être limitante et arrêter tout le processus. La réussite ou l'échec de chacune d'entre elles est autant dépendant de la tumeur initiale que de l'organisme de l'hôte.

La mutation d'une cellule somatique « saine » en une cellule tumorale est un processus complexe et long aboutissant à la formation d'une tumeur primitive localement envahissante et capable de disséminer plus ou moins rapidement vers un organe à distance du site de cancérisation primitif.

#### b) Prévalence métastatique des principaux cancers du chien et du chat

Le cancer est aujourd'hui une préoccupation majeure en santé des animaux de compagnie<sup>3</sup>. Les carnivores domestiques sont exposés à un nombre important de cancers, qui peuvent toucher tous les grands systèmes de leur organisme.

Chez le chien, les tumeurs les plus représentées sont :

- Les tumeurs cutanées et sous-cutanées qui représentent plus d'un tiers des tumeurs canines<sup>4,13,14</sup>.
- Les tumeurs mammaires qui touchent les femelles non stérilisées principalement<sup>13,14</sup>.
- Le lymphome multicentrique qui représente 7 à 24% des cancers diagnostiqués<sup>4</sup>.

Chez le chat, on retrouve globalement les mêmes affections, à quelques différences près :

- Le lymphome est beaucoup plus représenté, comptant pour 1/3 des tumeurs félines<sup>4</sup>.
- Les tumeurs cutanées représentent un quart des tumeurs du chat<sup>3,4</sup>.
- Les tumeurs mammaires sont également très fréquemment diagnostiquées, mais semblent plus agressives dans cette espèce<sup>3,4,13</sup>.

Comme il a été constaté précédemment, chaque tumeur a son propre potentiel métastatique, sa propre vitesse de généralisation et ses propres organes cibles. Le **Tableau 1** ci-dessous présentent les principaux cancers des carnivores domestiques cités ci-dessus et leur pouvoir métastatique pulmonaire<sup>13</sup>.

Tableau 1. Principales tumeurs du chien et du chat, fréquence et prévalence métastatique pulmonaire.

Fréquence des principales tumeurs	Fréquence par type tumorale	Prévalence des métastases pulmonaires Chez le chien et le chat
<b>Tumeur mammaire</b> 50 à 70% <sup>15</sup> - 20% <sup>3</sup>	<b>Adénocarcinome mammaire</b> >50% - La majorité	Fort pouvoir métastatique <sup>3</sup> avec les poumons comme site principal chez le chien comme le chat
	<b>Carcinome épidermoïde</b> 20% - 50%	Peu fréquentes, dissémination lente Possibles si récidives
<b>Tumeur cutanée</b>	<b>Mastocytome cutané</b> 16 à 21% - 2 à 15%	Très rarement, dépend du grade de la tumeur Rare
	<b>Mélanome cutané</b> Assez fréquent – 2 à 15%	30-75% de métastases à distance, poumons non cités
<b>Lymphome</b>	<b>Lymphome multicentrique</b> Incidence annuelle 114/100 000 <b>Digestif</b> 1/3 des cas	Dans plus de 50% des cas <sup>6</sup> : infiltration lymphomateuse pulmonaire Non

D'autres cancers, beaucoup moins fréquents que ceux cités précédemment ont un fort potentiel métastatique, avec les poumons comme site privilégié, on retrouve entre autres:

- Les ostéosarcomes<sup>16</sup>,
- Les tumeurs prostatiques<sup>13</sup>,
- Les carcinomes transitionnels de la vessie<sup>3</sup>,
- L'adénocarcinome thyroïdien<sup>17</sup>,
- Le carcinome corticosurrénalien<sup>6</sup>,
- Les tumeurs de l'orbite et du globe oculaire<sup>13</sup>,
- L'hémangiosarcome de l'atrium droit<sup>17</sup>,

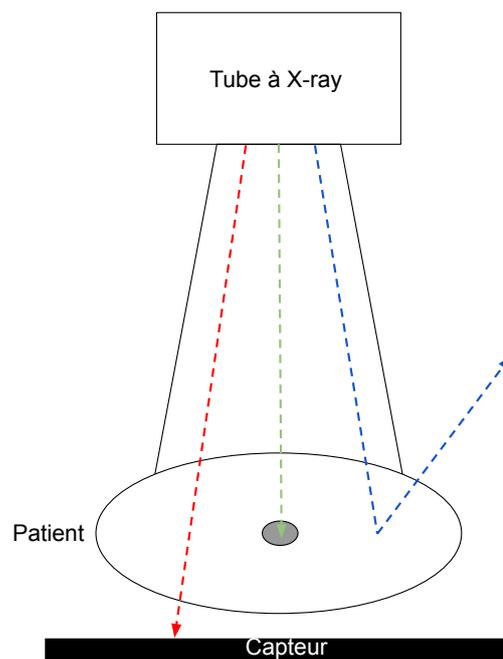
Au vu des informations précédentes, on remarque qu'une part significative des cancers des carnivores va métastaser plus ou moins rapidement à distance, et que l'un des organes privilégiés pour l'installation de ces tumeurs secondaires est le poumon. Malheureusement, il semblerait que malgré ces données, les métastases soient en général peu visibles lors du diagnostic et du bilan d'extension, notamment lorsque l'on utilise la radiographie comme examen<sup>16</sup>, du fait de leur petite taille difficilement discernable à l'œil nu. Cette problématique soulève l'intérêt de développer des technologies pour aider au diagnostic des métastases sur les radiographies pulmonaires.

## II) La radiographie

### a) Principe de la radiographie

Pour réaliser une radiographie, le patient est placé entre un tube à X-ray et la cassette de film radiographique (s'il s'agit d'une radiographie analogique) ou le capteur plan (s'il s'agit d'une radiographie numérique). Lorsqu'un rayon atteint le patient, trois interactions sont possibles<sup>18</sup> (cf. **Figure 3**) :

- Le rayon traverse le patient et sera enregistré par le film ou capteur plan (**rouge**),
- Le rayon est réfléchi (**bleu**),
- Le rayon est entièrement absorbé (**vert**).



*Figure 3. Schéma de la formation d'une image radiographique.*

La différence d'absorption des rayons X lorsque le faisceau primaire traverse les différents tissus du patient permet la formation de l'image. L'absorption dépend de l'épaisseur, de la densité et du nombre d'atomes composant le tissu de l'animal.

Les rayons qui ne sont pas absorbés arrivent jusqu'au capteur et permettent de former une image en nuance de gris.

Il est possible pour l'opérateur de régler l'intensité, la puissance et la durée d'émission des rayons X, pour optimiser au maximum la qualité de l'image. Ces variables dépendent de trois paramètres<sup>19</sup> :

- La tension en kiloVolts (kV) : elle détermine l'énergie des rayons X. Plus la tension est forte, plus les rayons X vont pénétrer les tissus, ce qui aura pour effet de diminuer le contraste entre deux structures de densités proches,
- L'intensité en milliAmpères (mA) : elle reflète le nombre de rayons X. Plus l'intensité est importante, plus il y aura de rayons X traversant les tissus, ce qui augmentera l'exposition,
- Le temps d'exposition en millisecondes (ms) : il représente la durée d'exposition du patient aux rayons. Plus l'exposition est longue, plus le nombre de rayons X atteignant le capteur est important, ce qui augmente l'exposition également. Cependant, on augmente également le risque de flou cinétique.

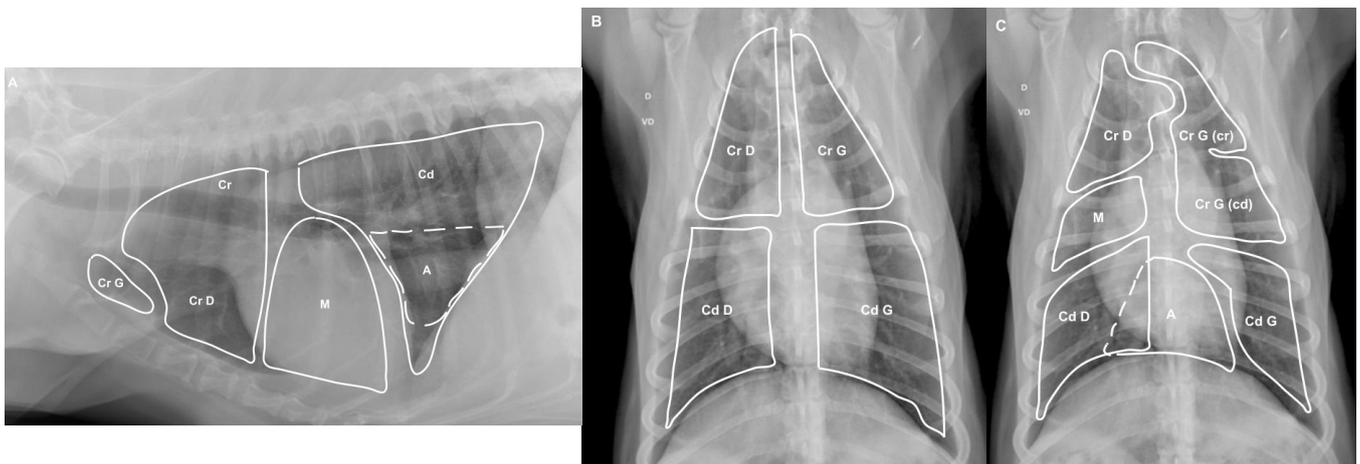
Les poumons étant remplis d'air, cette région présente naturellement un fort contraste, on privilégiera donc des réglages avec des hauts kV et des bas mAs. Pour prévenir le risque de flou cinétique naturel du thorax, lié aux mouvements cardiaques et respiratoires, la durée d'exposition doit être courte<sup>18</sup>.

## b) La radiographie thoracique

Le cliché radiographique doit être réalisé en fin d'inspiration pour maximiser le contraste pulmonaire. L'idéal est de prendre trois clichés avec trois incidences différentes : décubitus latéral droit (DLD), décubitus latéral gauche (DLG) et une incidence ventro-dorsale (VD) (ou dorso-ventrale (DV))<sup>18</sup>. En effet, une étude menée par *Lang et al* sur des radiographies de chiens a montré que l'association de ces trois vues permettait d'identifier correctement la présence d'une masse pulmonaire dans presque 100% des cas, lorsque les clichés étaient interprétés par des radiologues. Ils ont également obtenu de très bons résultats avec les associations suivantes : DLD + VD avec 97% de bonne identification en moyenne et 98,5% pour DLD + DLG<sup>20</sup>, montrant ainsi une amélioration de la spécificité de la radiographie.

Un cliché pulmonaire sain montre des poumons d'opacité aérienne homogène, avec parfois quelques modifications de l'opacité à relier à l'âge de l'animal. Du fait de son fort contraste, le thorax est la seule cavité dans laquelle on peut distinguer les vaisseaux sanguins. On visualise également les autres structures et organes thoraciques tels que le cœur et ses principaux vaisseaux (aorte et veine cave caudale), le médiastin, la trachée et le hile du poumon<sup>17,18</sup>.

La connaissance anatomique du poumon des carnivores domestiques permet une délimitation « empirique » des lobes des poumons droit et gauche (cf. **Figure 4**).



**Figure 4.** Radiographies de poumons sains avec délimitation schématique des lobes. *Adaptation personnelle par l'auteur du manuscrit d'après Owens et Biery, 1999.* (A : Radiographie en DLD, B : Radiographie en vue VD et délimitation des lobes ventraux à la trachée, C : Radiographie en vue VD et délimitation des lobes dorsaux à la trachée. Cr : Crânial ; Cd : Caudal ; A : Accessoire ; M : Moyen ; D : Droit ; G : Gauche)

Lors d'un bilan d'extension thoracique, si la radiographie révèle la présence de métastases, on pourra les reconnaître par la présence d'un effet masse d'opacité tissulaire relativement homogène (les nodules apparaîtront rarement, voire jamais,

cavitaires ou minéralisés)<sup>17,18</sup>. Ces masses peuvent se répartir partout dans le parenchyme pulmonaire et elles n'ont pas de localisation préférentielle. On peut également en retrouver au niveau du médiastin, de la paroi thoracique, du diaphragme, du cœur.

Généralement les métastases sont des masses bien définies, tout du moins pour le chien, à quelques exceptions près (métastases d'hémangiosarcome, de carcinome mammaire anaplasique...). Pour le chat, une étude de 1998 a montré que sur 12 clichés radiographiques de chattes atteintes d'une adénocarcinome mammaire, 10 d'entre elles présentaient des masses mal définies<sup>21</sup>.

Le fort contraste naturel du poumon permet de visualiser, en plus des organes intrathoraciques, les vaisseaux sanguins cardiaques et pulmonaires. Pour limiter les omissions et les images construites par la superposition, il est important de réaliser à minima 2 incidences perpendiculaires.

Les métastases apparaissent, dans la plupart des cas, comme des éléments arrondis plus ou moins homogènes, d'opacité proche des opacifications interstitielles ou alvéolaires.

#### c) Avantages de la radiographie

La radiographie a l'avantage d'être un examen complémentaire disponible dans la plupart des cliniques vétérinaires. La réalisation du cliché est également à la portée de tous les praticiens, généralistes ou spécialistes en imagerie médicale.

C'est un examen réalisable à moindre coût, qui ne nécessite pas nécessairement l'anesthésie de l'animal.

La radiographie thoracique permet d'évaluer toute la profondeur du tissu pulmonaire.

La radiographie constitue l'examen de première intention lors d'un bilan d'extension thoracique en médecine vétérinaire.

#### d) Limites de la radiographie

La prise d'un cliché radiographique nécessitant la contention de l'animal, les opérateurs sont exposés aux rayons X radioactifs, ils doivent obligatoirement porter des équipements de protection individuelle (EPI) tels qu'un tablier plombé, un protège thyroïde ou des lunettes de protection.

Aussi, de simples facteurs techniques (mauvaises constantes, mauvais positionnement du patient...) ou liés à l'animal (âge, pathologies antérieures...) peuvent conduire à des erreurs d'interprétation<sup>18</sup>. En ce qui concerne la détection des nodules thoraciques, il semble que même si les vétérinaires de tout niveau puissent les identifier, plus l'entraînement du lecteur est avancé, moins il a de risque d'en omettre<sup>22</sup>.

Les clichés radiographiques ne permettent pas de détecter des nodules de taille inférieure à 4-5mm<sup>17</sup>, avec une détection considérée comme fiable à partir d'une taille de 7 à 9mm<sup>23</sup>, démontrant un certain manque de sensibilité. Ce constat crée alors un délai entre la vraie présence d'une métastase et le moment de sa détection sur la radiographie, délai pouvant être plus important chez les praticiens moins expérimentés. Ce délai constitue alors un frein à la prise en charge du patient.

La sensibilité de la radiographie dans la détection des métastases pulmonaires est moyenne. Elle dépend notamment de la taille des nodules à détecter mais aussi de l'expérience du lecteur. Même si la spécificité peut être augmentée lors de la réalisation de 3 incidences radiographiques, la détection d'un nodule pulmonaire ne garantit pas toujours sa nature métastatique.

### III) L'échographie

#### a) Principe de l'échographie

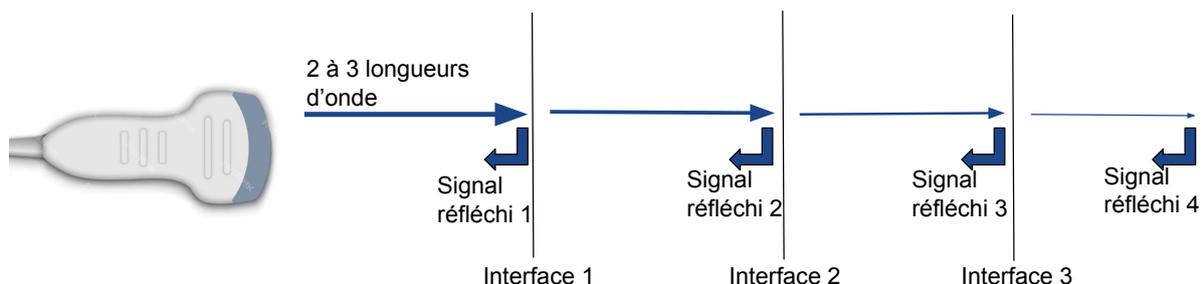
L'échographie est une technique d'imagerie basée sur les ondes sonores, ces dernières ont besoin d'un milieu pour se propager. Au sein des gaz, liquides ou éléments semi-solides, comme le corps, les ultrasons avancent longitudinalement. La sonde échographique est à la fois émettrice (1% du temps) et réceptrice (99% du temps) des ondes sonores. Chaque pulse envoyé est composé de deux à trois ondes de longueurs d'ondes différentes<sup>24,25</sup>.

Quand une onde traverse un tissu, quatre interactions sont possibles :

- La réflexion,
- La réfraction,
- La diffraction,
- L'atténuation.

La formation de l'image échographique est basée sur la différence d'impédance acoustique (en  $\text{kg/m}^2 \cdot \text{s}$ ) entre deux tissus, qui va déterminer la quantité de son réfléchi puis captée par le transducteur, et la quantité qui va les traverser (cf. **Figure 5**). L'impédance d'un tissu dépend de sa densité et de la vitesse du son en son sein<sup>25</sup>.

Il est possible de modifier la fréquence des ondes (en Hz) : plus elle est haute, plus la résolution en profondeur est bonne et plus la zone proche de la sonde apparaît longue à l'écran. Cependant, les hautes fréquences augmentent l'absorption des ondes par les tissus, ce qui ne convient pas aux tissus de haute densité.



**Figure 5.** Schéma de la formation d'une image échographique.

Le mode doppler est utilisé pour repérer et permettre d'évaluer la direction et la vitesse du flux sanguin.

L'effet Doppler est basé sur la différence de fréquence de l'onde sonore consécutive au déplacement de la source ou du récepteur. Ce sont les globules rouges qui agissent comme des réflecteurs en mouvement. La sonde, immobile va alors recevoir une onde de plus grande ou de plus petite fréquence, permettant ainsi d'évaluer et de caractériser le flux sanguin<sup>24</sup>.

L'air est un milieu avec une impédance très faible comparée à la plupart des autres milieux (446 kg/m<sup>2</sup>.s contre des valeurs de l'ordre de 10<sup>6</sup> kg/m<sup>2</sup>.s). Cette grande différence d'impédance entre l'air des poumons et celle des tissus cutanés, sous-cutanés et graisseux provoque une forte réflexion des ondes ultrasonores. Cette effet reflet empêche les ondes de pénétrer en profondeur. Afin de permettre la meilleure visualisation échographique possible des poumons, il est conseillé d'utiliser une sonde convexe de fréquence de 5 à 10 MHz<sup>24,26</sup>.

#### b) L'échographie du thorax, excluant le cœur

Pour l'échographie de la région du thorax, On utilise communément trois différentes fenêtres de visualisation <sup>24,25</sup>:

- Entrée du thorax,
- Intercostal,
- Parasternal.

Prenons la fenêtre intercostale par exemple. Sur un animal sain, on distingue au plus proche de la sonde, les tissus cutanés, le gras sous cutanée et le tissu musculaire, formant une alternance de lignes hypo et hyperéchogènes à l'écran.

La fenêtre de visualisation est délimitée par les côtes, que l'on reconnaît par des lignes curvilignes hyperéchogènes suivies par des cônes d'ombres (artéfacts) s'étendant sur toute la profondeur du thorax. La transition plèvre-poumon est caractérisée par une ligne hyperéchogène sur laquelle on peut voir des mouvements de glissements. Le reste du parenchyme pulmonaire, s'il est sain, n'est pas visualisable<sup>24,27</sup>.

Seules les masses pulmonaires les plus superficielles pourront être décelées<sup>24</sup>, sauf si un épanchement pleural permet le passage des ondes sonores plus loin dans le parenchyme (cf. **Figure 6**). Les masses apparaissent comme des éléments arrondis hypoéchogènes homogènes, aux bords échogènes et lisses, souvent bien délimitées. Elles peuvent éventuellement contenir des foyers de nécrose ou du gaz<sup>24,28</sup>, ce qui les rend parfois difficilement différenciables des abcès<sup>27</sup>.

L'échographie thoracique est plus généralement utilisée pour l'observation des métastases médiastinales, cardiaques, péricardiques ou des gros vaisseaux (aorte et veines caves).

En mode doppler, on peut observer la vascularisation de la masse, qui est généralement tortueuse, avec une vitesse très rapide contrairement au reste de la vascularisation pulmonaire.<sup>6</sup>



**Figure 6.** Image échographique d'un nodule pulmonaire de 8mm de diamètre chez un chien (flèches blanches) dans le lobe caudal droit. Image obtenue auprès du service d'imagerie de l'école nationale vétérinaire OnirisVetAgroBio..

Dans quelques cas particuliers, notamment lors de la présence d'un épanchement thoracique, l'échographie peut constituer un examen d'imagerie intéressant dans le cadre du bilan d'extension thoracique.

### c) Avantages de l'échographie

Tout comme la radiographie, l'échographie est un examen complémentaire disponible dans la plupart des cliniques. Il permet lui aussi, de réaliser des images sur un animal vigile.

Elle permet de s'affranchir d'anomalies telles que les effusions pleurales et les consolidations, car ces lésions permettent la formation de fenêtre de visualisation du parenchyme pulmonaire<sup>26-28</sup>.

L'examen échographique dans le bilan d'extension thoracique permet également la visualisation des nœuds lymphatiques intrathoraciques, ainsi que la réalisation de prélèvements échoguidés pour cytologie ou biopsie<sup>3</sup>, quand la localisation le permet. Enfin, en suivant les mouvements de la masse lors de la respiration, l'échographie va nous permettre d'identifier son origine : si elle concorde, elle est plutôt d'origine pulmonaire, si elle reste immobile, elle est soit pariétale soit médiastinale<sup>25</sup>.

Contrairement à la radiographie, l'échographie permet de s'affranchir d'anomalies pouvant masquer les métastases pulmonaires, telles que les épanchements thoraciques. L'échographie présente donc une meilleure sensibilité que la radiographie dans ce seul cas particulier. Cet examen permet également la réalisation de prélèvements afin de caractériser les masses pulmonaires.

#### d) Limites de l'échographie

L'échographie pulmonaire est une technique d'imagerie complexe qui nécessite une formation poussée de l'opérateur. La réalisation d'un tel examen, ainsi que l'interprétation des images ne sont donc pas à la portée de tous les praticiens.

L'échographie thoracique (hors base d'échocardiographie) ne se fait pas en routine. C'est en réalité une confirmation d'une suspicion radiographique. La radiographie permet de localiser la fenêtre de visualisation optimale pour éviter d'avoir à tondre tout le thorax de l'animal<sup>26,28</sup>. Sinon, elle est restreinte aux masses présentes dans ou proches de la paroi thoracique. Nous rappelons également que la visualisation des métastases est dépendante de la présence de fluides dans le thorax, réduisant encore le nombre de cas où cet examen devient pertinent.

Elle est plutôt utile pour des masses de taille importante, comme les tumeurs primitives du poumon (carcinomes pulmonaires, sarcomes...). Il n'y a cependant pas d'utilisation rapportée pour la détection des métastases, probablement du fait de leur localisation dans le parenchyme pulmonaire.

Le thorax étant une zone épaisse et non dépressible, l'exploration échographique exhaustive de cette région à la recherche d'une masse est presque impossible. En effet, du fait de l'atténuation progressive, et de par la conformation du thorax, cette technique d'imagerie ne permet pas de visualiser toute la profondeur de cette région.

Étant donné l'aspect des masses sur les images échographiques, il est difficile de déterminer si la nature des lésions est bénigne ou maligne. L'échographie a donc une mauvaise spécificité<sup>3,6</sup>.

Globalement, l'échographie a une mauvaise sensibilité liée à la conformation du thorax. Dans la majorité des cas, seules les métastases les plus superficielles sont visualisables.

Tout comme la radiographie, la spécificité de cet examen est moyenne. En effet, la seule image échographique ne permet pas d'assurer de la nature métastatique de la masse.

Enfin, la réalisation d'un tel examen nécessite une formation de l'opérateur. Sa pratique n'est pas à la portée de tous les vétérinaires et nécessitent une certaine expérience.

#### IV) Le scanner

##### a) Principe du scanner

Tout comme la radiographie, l'imagerie scanner se base sur la différence d'absorption des rayons X entre les tissus formant le corps de l'animal. Mais cette fois-ci, le tube à rayons X tourne autour du patient pour réaliser de nombreuses coupes fines d'une région anatomique donnée. Toutes les images de projection ainsi recueillies sont communiquées à un ordinateur qui va reconstituer la coupe du tissu étudié.

Pour faire une acquisition scanner, l'opérateur a la possibilité de modifier un grand nombre de paramètres, dont les kV et les mAs dont nous avons parlé précédemment. Le plus important de tous est l'épaisseur de tranche, qui joue sur la netteté de l'image et sur le bruit : plus l'épaisseur est faible, plus l'image sera nette, mais plus le bruit numérique sera important (l'image subit des fluctuations parasites ou des dégradations)<sup>29</sup>.

Pour réaliser une acquisition de la région des poumons chez les petits animaux domestiques, les paramètres communément employés sont une acquisition à haute résolution avec des hauts mAs (200 à 300 mAs) et kV (120kV), une épaisseur de tranche de 3 à 5mm (pouvant être réduite jusqu'à 1mm pour la détection des métastases pulmonaires) et un *pitch* (défini comme la relation entre le déplacement de la table patient et l'épaisseur de coupe choisie lors d'une rotation de l'anneau) nul ou faible (maximum 1 à 1,5)<sup>29,30</sup>.

## b) Tomographie thoracique

Le scanner est un excellent examen complémentaire pour l'imagerie du thorax, notamment pour identifier des lésions difficilement reconnaissables à la radiographie ou à l'échographie<sup>30</sup>, telles que les métastases pulmonaires de petite taille. C'est une technique utilisée notamment pour la détection de métastases thoraciques pour l'évaluation du stade cancéreux.

Un poumon sain apparaît comme un tissu relativement homogène avec des densités oscillant entre -846 et -713 Unité Hounsfield (UH) en fonction de la respiration. Les bronches et bronchioles, ainsi que les vaisseaux sanguins sont visibles en son sein, il n'y a pas de septum interlobaire. Les métastases quant à elles sont reconnaissables par leurs bords bien délimités (sauf exception pour les métastases d'hémangiosarcome et pour la majorité des métastases du chat) et leurs densités tissulaires<sup>29</sup>.

Les métastases pulmonaires apparaissent au scanner comme des éléments arrondis de densité tissulaire. Ils sont distincts des vaisseaux pulmonaires mais peuvent contenir des vaisseaux nourriciers en leur sein. Les métastases apparaissent très rarement minéralisées.

## c) Avantages du scanner

Au vu des principes de fonctionnement de la tomographie, son avantage, contrairement à la radiographie, est qu'elle s'affranchit des superpositions et offre un meilleur contraste, ce qui fait d'elle une technique relativement spécifique. En médecine humaine, c'est d'ailleurs le gold standard lors de bilans d'extension thoracique<sup>31</sup>.

Le scanner permet une meilleure identification des nodules de petites tailles, avec une sensibilité dès la taille de 1mm<sup>18</sup>. L'étude de *Nemanic et al* datant de 2006 a comparé l'efficacité de la radiographie et du scanner dans la détection de métastases thoraciques chez le chien. Les auteurs ont montré que dans 39% des cas, les nodules n'étaient détectés que par la tomographie<sup>23</sup>.

C'est également un passage presque indispensable pour la thérapeutique à suivre<sup>6</sup> car cet examen permet de connaître la localisation exacte de la lésion (utile pour la chirurgie ou la radiothérapie).

Le scanner est une technique d'imagerie très sensible permettant une détection des métastases pulmonaires dès la taille de 1mm.  
C'est également un examen spécifique qui permet de différencier la nature des masses dans la plupart des cas. Cependant, certains granulomes peuvent parfois être confondus avec des métastases.

#### d) Limites du scanner

Le scanner est une technique d'imagerie à rayons X qui émet des radiations ionisantes à des niveaux beaucoup plus élevés que la radiographie, il faut donc une excellente contention de l'animal pour que les opérateurs puissent quitter la pièce. Aussi, l'animal doit rester immobile pendant tout le temps d'acquisition (qui peut prendre quelques minutes) pour que les images soient d'une qualité suffisante pour permettre leur interprétation. C'est pour ces deux raisons que le patient doit obligatoirement subir cet examen sous anesthésie générale.

Ces appareils sont disponibles dans seulement une trentaine<sup>32</sup> de structures sur les 7 976 établissements répertoriés en France<sup>33</sup>. C'est un examen relativement coûteux pour les propriétaires, de l'ordre de 500€ minimum, contre 70€ en moyenne pour deux incidences radiographiques orthogonales.

Les paramètres utilisés pour la détection des métastases pulmonaires demandent un temps d'acquisition long, et l'interprétation des images nécessite l'avis d'un vétérinaire spécialisé en imagerie. L'obtention des résultats constitue de ce fait une tâche longue et fastidieuse.

Le scanner est une technique d'imagerie qui, même si elle se développe, est encore peu répandue en médecine vétérinaire. Elle représente aussi un coût important pour les propriétaires. La nécessité d'une anesthésie générale peut également freiner la réalisation de cet examen, par peur du risque anesthésique existant.  
La lecture et l'interprétation des images obtenues nécessitent l'expertise d'un vétérinaire spécialiste en imagerie médicale.

L'IRM et la médecine nucléaire sont deux autres techniques d'imagerie utilisées en oncologie. Elles ne seront pas développées ici car elles n'ont pas d'application dans le bilan d'extension (BE) thoracique.

## V) Synthèse

Le **Tableau 2** récapitule les avantages et les inconvénients de chaque examen d'imagerie, explicités dans cette partie, et disponibles afin de réaliser un bilan d'extension thoracique.

**Tableau 2.** Tableau récapitulatif pour le choix des examens d'imagerie dans le bilan d'extension thoracique.

Examen d'imagerie	Coût	Disponibilité	Principaux avantages	Principaux Inconvénients
Radiographie	50€ minimum	Élevée	<ul style="list-style-type: none"> <li>- Examen de choix pour le BE thoracique</li> <li>- Interprétation à la portée de tout vétérinaire généraliste</li> <li>- Coût faible</li> </ul>	<ul style="list-style-type: none"> <li>- Omission</li> <li>- Se relativement faible comparée à d'autres techniques d'imageries</li> </ul>
Échographie	80€ minimum	Modérée à élevée	<ul style="list-style-type: none"> <li>- Permet de s'affranchir d'anomalies pulmonaires masquant les lésions sur les radiographies</li> </ul>	<ul style="list-style-type: none"> <li>- Constitue seulement une aide à la détection</li> <li>- Nécessite une formation poussée de l'opérateur</li> </ul>
Scanner	500€ minimum	Faible à modérée	<ul style="list-style-type: none"> <li>- Meilleure sensibilité et spécificité</li> </ul>	<ul style="list-style-type: none"> <li>- Relativement peu utilisée en médecine vétérinaire (contrairement à la médecine humaine)</li> <li>- Coût élevé</li> </ul>

Par la suite, le choix sera fait d'utiliser des clichés radiographiques car il s'agit de l'examen le plus accessible pour les vétérinaires praticiens. De plus, du fait de son faible coût et de son interprétation quasi-immédiate, il constitue également l'examen le plus choisi par les propriétaires.

Afin de pallier les problèmes de sensibilité (Se) et de spécificité (Sp) de cette technique comparée aux autres examens disponibles, et pour assurer le meilleur diagnostic possible, même pour les vétérinaires les moins expérimentés, un algorithme de *Deep Learning* en vue de détecter automatiquement les métastases pulmonaires sera développé dans cette étude. La décision de l'utilisation d'un tel outil s'inscrit dans un contexte où les IA sont dorénavant au cœur du diagnostic médical.

## Partie 2 – Apports actuels des Intelligences Artificielles en médecine vétérinaire

### I) Préalable à l'utilisation des IA

#### a) Définitions et concepts

L'intelligence artificielle est définie comme l'ensemble de théories et de techniques mises en œuvre en vue de réaliser des machines capables de simuler l'intelligence humaine (raisonnement, apprentissage...) <sup>34</sup>. En informatique, les applications de tels algorithmes sont multiples : analyse d'images, *language processing*... <sup>35</sup>

À leurs débuts, les IA étaient utilisées pour résoudre des problèmes complexes pour l'homme, (du fait d'un grand nombre de données ou du temps nécessaire pour leur résolution) mais relativement facile à automatiser pour un ordinateur, tels que la résolution de problèmes mathématiques, ou le recadrage d'images. Avec l'avènement du *Machine Learning*, ou apprentissage automatique (ML), dans la seconde moitié du XX<sup>ème</sup> siècle, les systèmes d'IA ont maintenant la capacité d'extraire des informations depuis des données brutes et d'apprendre à la suite d'entraînements successifs <sup>36</sup>. Le ML utilise des algorithmes pour identifier des similitudes entre chacune des données lui permettant de construire un modèle de plus en plus juste au cours des expériences successives, permettant ainsi la réalisation de prédictions ou la prise de décision <sup>37</sup>. Il existe trois catégories de ML, dépendant de la méthode d'apprentissage :

- **L'apprentissage supervisé** : l'algorithme se réfère à une donnée standard dite « ground truth » qui lui est fournie. Au cours de son apprentissage, l'algorithme compare ses propres résultats (que ça soit de la classification, de la régression, de la localisation, de la segmentation...) avec la donnée supposée vraie et fait des ajustements pour s'approcher au plus près de cette vérité donnée <sup>38</sup>,
- **L'apprentissage non supervisé** : aussi appelé de façon imagée « l'apprentissage sans professeur » <sup>39</sup>. Seules les données brutes sont fournies à l'algorithme, qui doit lui-même identifier un modèle au cours de ses entraînements successifs <sup>38</sup>,
- **L'apprentissage par renforcement** : ou « apprentissage par critique ». Contrairement au professeur (utilisé pour l'apprentissage supervisé), la critique fait un retour sur une action passée mais ne donne pas l'objectif à atteindre. Un algorithme utilisant cet apprentissage va générer une valeur « interne » pour chaque état intermédiaire emprunté. Cette valeur informe à quel point l'intermédiaire emprunté était bon pour accéder au but. Une fois les meilleurs intermédiaires appris, il suffit de les maximiser pour que chaque donnée emprunte le meilleur chemin et arrive à l'objectif <sup>40</sup>.

La branche de *Machine Learning* la plus utilisée aujourd'hui est le *Deep Learning* (DL), ou « apprentissage profond », qui comme le ML peut utiliser une des trois méthodes d'apprentissage précédemment citées. C'est la discipline de l'IA la plus performante à ce jour. Le DL permet une meilleure généralisation des données au modèle, cette méthode est également plus rapide et plus facile à mettre en place. Aussi, elle offre de meilleures performances que le ML <sup>41</sup>. Le *Deep Learning* est capable d'analyser des données non structurées telles que les images, les textes, les vidéos.

L'intérêt pour cette branche du ML est grandissant depuis que ses algorithmes ont surpassé les performances humaines en 2015 lors du *ImageNet Large Scale Visual*

*Recognition challenge*. Cette compétition évaluait les performances d’algorithmes de classification d’images pour labéliser tous types d’images (représentant des animaux, des moyens de transport, des fleurs) disponibles sur la banque de données *open source ImageNet*<sup>36</sup>.

Les algorithmes de *Machine Learning* peuvent être développés selon trois méthodes d’apprentissage (supervisé, non supervisé, par renforcement). Ces techniques d’entraînement permettent de traiter une grande variété de données dans un important panel de disciplines. L’intérêt porté au ML, et plus particulièrement pour la branche appelée *Deep Learning*, s’est intensifié depuis que ces algorithmes égalent, voire surpassent, les performances humaines.

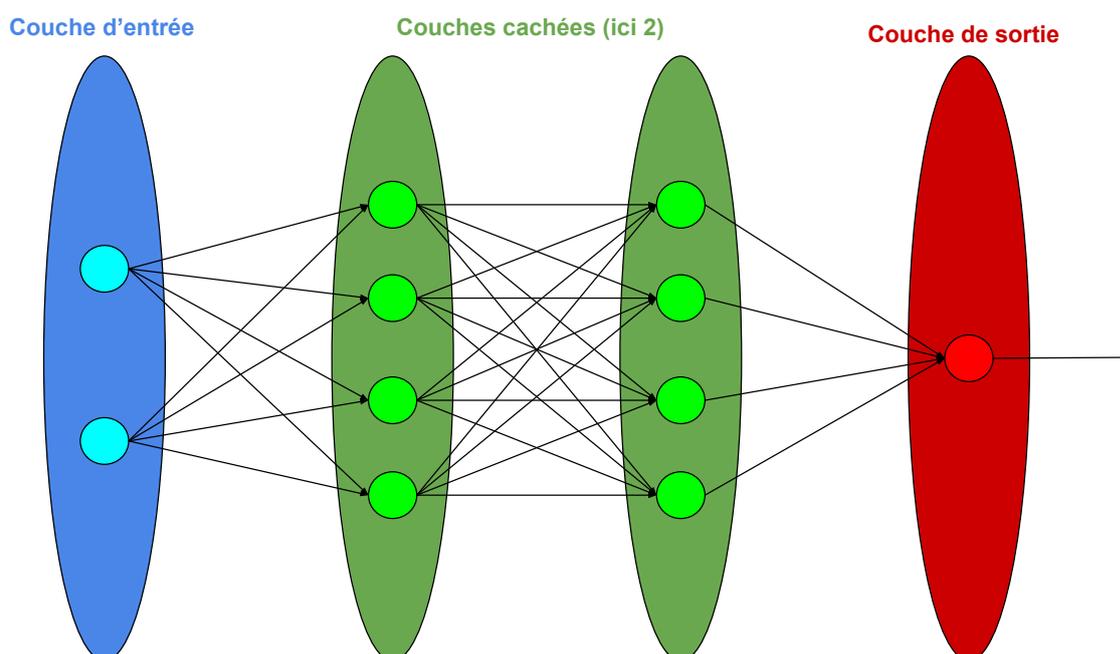
## b) Organisation des réseaux de Deep Learning

Le *Deep Learning* repose sur des algorithmes organisés en réseaux de neurones artificiels inspirés des neurones du cerveau humain. Des points de décisions semblables aux neurones, appelés nœuds sont interconnectés avec des nœuds voisins dans des couches successives. Les premiers nœuds, qui reçoivent les données brutes forment la « couche d’entrée » (*input layer*). Les derniers nœuds quant à eux, forment la « couche de sortie » (*output layer*), qui annonce les données de sortie. Les neurones intermédiaires entre l’entrée et la sortie forment les « couches cachées » qui pratiquent des opérations mathématiques transformant l’entrée en sortie<sup>42</sup> (cf. **Figure 7**).

Chaque réseau neuronal à un nombre défini de couches, pour chacune d’elles :

- Elles comportent un nombre fini de neurones, mais ce nombre peut varier d’une couche à une autre.
- Chaque nœud d’une même couche à la même fonction d’activation.

Les réseaux de neurones peuvent donc avoir à minima deux couches (la couche d’entrée et la couche de sortie). Les réseaux de *Deep Learning* quant à eux, ont un très grand nombre de couches (réseaux multicouches).



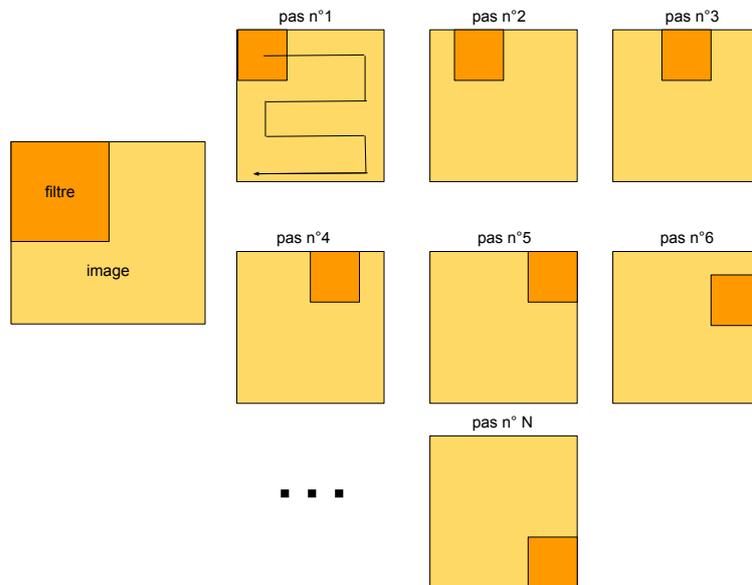
**Figure 7.** Représentation schématique d'un réseau de neurones à 2 couches cachées.

Il existe différents types de réseau de neurones. Ces différences résident dans les types de neurones utilisés (fonction d'activation, valeurs propres du nœud), le nombre de neurones, la complexité des connexions entre ses neurones mais aussi dans le type d'apprentissage<sup>43</sup> (cf. **Partie II- I a)**). On peut citer :

- Le **réseau à propagation avant** (ou *feed forward*) dans lequel l'information circule vers l'avant de l'entrée à la sortie du réseau.
- Le **réseau récurrent** dans lequel il existe au moins une rétroaction. Les sorties du modèle sont ainsi réinjectées sur lui-même. Un tel réseau permet à l'algorithme de modifier les poids de ses neurones dans le temps pour tendre vers la sortie attendue.
- Le **réseau convolutif**, le plus utilisé dans le traitement d'images<sup>44</sup>. Ce réseau a pour but d'extraire des *features* ou « caractéristiques » de plusieurs portions de l'image après application de filtres.

Les réseaux de neurones convolutifs (ou RNC) sont spécialisés dans le traitement de données en forme de grille. Les images, pouvant être représentées par une matrice de pixels, sont une des données d'entrée les plus fréquemment utilisées. Ce qui différencie les RNC des autres réseaux de neurones, c'est la partie de convolution : sur une image, on peut assimiler son effet à l'application de filtres successifs. Un réseau convolutif est généralement composé des couches suivantes :

- La couche de convolution, ou filtrage (cf. **Figure 8**), qui permet d'analyser des portions réduites de l'image et d'en extraire des *features* ou *caractéristiques*. Cette phase permet d'extraire les caractéristiques les plus pertinentes de la zone étudiée. Les filtres, qui ont une taille définie, peuvent permettre la détection des bords, de formes géométriques. En sortie, on obtient une carte d'activation ou *feature map* qui indique la position des caractéristiques sur l'image.
- Une couche ReLU (Rectified Linear Units) qui permet d'accélérer l'apprentissage du RNC en corrigeant les résultats des fonctions d'activation. Toutes les valeurs négatives sont remplacées par 0, rendant ainsi le modèle non linéaire et in fine plus complexe.
- La couche de *pooling*, que l'on retrouve la plupart du temps entre deux couches de convolution. Cette opération permet de réduire la taille de l'image, tout en préservant les caractéristiques importantes, permettant ainsi de réduire le nombre de calculs nécessaires au réseau. Pour les plus connues, on retrouve la fonction « *max-pooling* » qui pour chaque région filtrée prend le maximum<sup>45</sup> (cf. **Figure 9**) ou la fonction « *average pooling* » qui calcule la moyenne. Le processus est réitéré selon un pas de déplacement jusqu'à ce que toute l'image ait été balayée.
- La couche *fully connected*, qui constitue la dernière couche de n'importe quel réseau de neurones, n'est pas spécifique du RNC. Les dernières couches sont unidimensionnelles et connectées à toutes les activations de la couche précédente<sup>46</sup>.



**Figure 8.** Illustration du fonctionnement d'une fonction de convolution: parcours du filtre sur l'image  
Adaptation personnelle de l'auteur du manuscrit d'après Natural solutions, 2018



**Figure 9.** Illustration du fonctionnement de la fonction Max-pooling. En entrée une matrice 4x4 sur laquelle on applique un filtre de taille 2x2 et avançant d'un pas de 2. Pour chacune des 4 zones balayées, la fonction max-pooling prendra le maximum. Il est alors obtenu en sortie une matrice 2x2 où chaque élément correspond au maximum des régions rencontrées par le filtre. Adaptation personnelle par l'auteur du manuscrit d'après Germain P., 2018.

Les connexions entre les nœuds, assimilables à des synapses, vont moduler l'intensité de l'information transmise d'un neurone à l'autre via une valeur appelée le poids. Les neurones peuvent également moduler le signal reçu via la fonction d'activation et le biais. Ces deux valeurs (biais et poids) sont des variables de l'algorithme qui vont être mises à jour continuellement pour améliorer la précision du réseau (par le biais de la rétroaction).

Le poids et le biais peuvent donc être considérés comme des valeurs d'auto-apprentissage de nos réseaux de neurones.

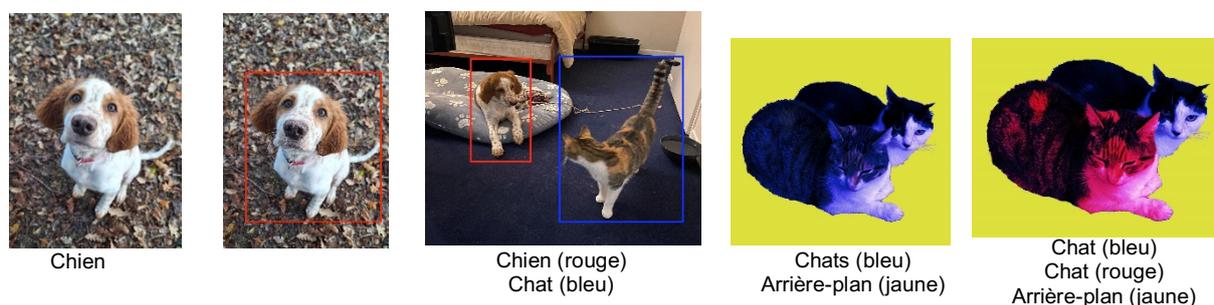
La spécificité des algorithmes de *Deep Learning* est qu'ils sont formés de nombreuses couches cachées. Parmi ces modèles informatiques, on retrouve les réseaux de neurones convolutifs, spécialisés dans le traitement de données en forme de grille, telles que les images. Une différence notable de ces algorithmes de DL est l'utilisation d'une fonction de convolution permettant de mettre en évidence des *features* c'est à dire des caractéristiques pertinentes de l'image (bords, formes géométriques...)

### c) Le DL appliqué au traitement d'images

L'utilisation des réseaux de neurones pour le traitement d'images s'est surtout démocratisée après l'avènement des *Graphic Processing Unit* (GPU). Ce processeur, équipant la carte graphique de l'ordinateur, a permis d'augmenter la taille des réseaux de *Deep Learning* et de la banque de données<sup>47</sup>. Il a surtout rendu possible le traitement de nombreuses ressources informatiques en un temps raisonnable.

La discipline du DL qui donne à l'IA la capacité de lire et d'analyser des images (et des vidéos) est appelée *Computer Vision*. Elle traite de tâches telles que<sup>48</sup> :

- La classification : chaque image est placée dans une catégorie prédéfinie.
- La localisation d'objets : la position d'un ou plusieurs objets présents dans l'image est donnée par une boîte englobante ou *Bounding box*. Cette boîte correspond au plus petit rectangle contenant l'objet.
- La détection d'objet : l'algorithme donne une liste de *Bounding Boxes* et catégorise les objets qu'elles contiennent. Cette tâche peut être assimilée à l'association des algorithmes de classification et de localisation. Les objets à détecter sont prédéfinis.
- La segmentation<sup>49</sup> : l'image est divisée en plusieurs régions ou segments pour faciliter son analyse, appelées *Region of interest (ROI)*. Cet algorithme permet de délimiter des frontières entre les différents objets présents sur l'image.
  - Sémantique : La segmentation sémantique consiste à diviser une image en régions sémantiquement significatives. Elle permet de reconnaître et d'attribuer des étiquettes à des objets. On ne peut faire de différences entre deux instances d'un même objet. Par exemple (cf. **Figure 10**), si on a deux chats sur une image, ce type de segmentation donnera le même label sur l'ensemble des pixels des deux chats, ils apparaîtront donc de la même couleur,
  - D'instance : La segmentation d'instance va attribuer un unique label sur chaque instance d'un même objet au sein de l'image. En reprenant l'exemple de **la Figure 10**, si nous avons à nouveau deux chats sur une image, chacun d'entre eux se verront attribuer une couleur différente.



**Figure 10.** Représentation des différentes tâches du CV. De gauche à droite : classification, localisation, détection, segmentation sémantique et segmentation d'instance. *Adaptation personnelle de l'auteur du manuscrit, d'après Casado-García et al, 2019.*

Les domaines d'applications du Computer Vision sont très larges, on retrouve comme domaines principaux : la surveillance et la sécurité, la vision industrielle, la conduite autonome, la reconnaissance faciale, le tri des déchets, la médecine<sup>50</sup> (Humaine et Vétérinaire).

Le *Deep Learning* ou apprentissage profond est une branche du Machine Learning très utilisée pour le traitement d'images. Elle permet la détection d'objets d'intérêt sur tous types d'images. Sa capacité à traiter un grand nombre d'informations a permis de démocratiser son utilisation dans de nombreux domaines, notamment en médecine humaine et vétérinaire.

#### d) Entraînement des algorithmes de Deep Learning

##### i. Les données nécessaires au fonctionnement de l'IA

Afin de parvenir à son objectif, le réseau de neurones doit auparavant s'entraîner sur une base de données, à faire ce pourquoi il a été construit. Dans notre étude, il s'agit de détecter des nodules pulmonaires sur des radiographies de thorax.

En amont, les nodules seront localisés manuellement sur des images, et cette banque de données servira de modèle à l'IA pour qu'elle puisse se corriger. L'objectif de cet apprentissage sera de placer une ou plusieurs *Bounding Box(es)* autour des métastases.

Pour apprendre, un lot d'images est défini et fourni à l'algorithme. Ce lot d'entraînement sert à modifier les poids et biais du réseau de neurones pour qu'il se rapproche le plus possible de l'objectif donné.

Pour tester l'efficacité de cet apprentissage, un deuxième lot appelé « lot de validation », permet de vérifier que l'algorithme de DL est capable d'appliquer ce qu'il a appris à de nouvelles images. Il est important que ces deux jeux de données soient strictement différents, afin de pouvoir généraliser le modèle.

Usuellement, sur une banque d'images, on utilise un ratio 80/20 pour le nombre d'images nécessaires à l'entraînement et à la validation respectivement.

On peut alors se demander combien d'images sont nécessaires pour entraîner un tel système. De nombreuses études se sont penchées sur cette question, sans pouvoir prédire le nombre exact de données numériques. Il semblerait que cela dépende :

- De la complexité de l'objectif à atteindre,
- De la complexité du réseau de neurones (sa structure),
- De la qualité des images fournies,
- De la sensibilité et de la spécificité que l'on souhaite atteindre.

Il semble finalement relativement difficile d'évaluer la quantité exacte d'images nécessaires pour l'apprentissage d'une IA, même si il semble assez logique que plus le nombre de données est important, plus l'on obtient de bons résultats, au risque d'un traitement plus long par l'algorithme<sup>51</sup>. L'évaluation du nombre minimum d'images nécessaires à notre étude (N minimum) sera abordée dans la **Partie 3 - II) b)**.

## ii. Les hyperparamètres

Les hyperparamètres contrairement aux paramètres tels que le biais et le poids (cf. **Partie II- I B**)), sont externes au processus d'entraînement. Ce sont des valeurs fixées par l'utilisateur avant la phase d'entraînement et qui vont configurer le réseau de neurones. Ces paramètres restent inchangés durant l'apprentissage.

Les hyperparamètres sont importants car ils vont significativement influencer les performances du modèle : avec des valeurs erronées, il est possible de sous-apprendre (*underfitting*) ou de surapprendre (*overfitting*) sur la base de données, ce qui empêchera la généralisation du modèle (ou tout du moins avec des performances décevantes). Parmi eux, on peut citer<sup>52</sup> :

- Le *batch size* : représente le nombre d'images qui sont utilisées par l'algorithme avant de modifier ses paramètres,
- L'*epoch* : représente le nombre de passages complets nécessaires pour que toutes les images de la base de données aient été utilisées pour l'entraînement du modèle,
- Le *learning rate* (ou pas d'apprentissage) : Ce paramètre régit la cadence à laquelle les paramètres de l'algorithme sont mis à jour.

Il existe de nombreuses méthodes permettant d'aider l'utilisateur à choisir les bons paramètres telles que : la recherche en grille, la recherche aléatoire, l'optimisation bayésienne, méthodes qui ne seront pas décrites dans cette étude.

Les 3 hyperparamètres explicités dans cette partie seront choisis par les développeurs du modèle de détection décrit dans la **Partie 3**. Dans cette étude, une optimisation par la méthode de descente de gradient stochastique sera utilisée pour choisir les paramètres.

## II) Les grands secteurs d'utilisation des IA en médecine vétérinaire

### a) Épidémiologie

L'étude des maladies au sein des populations est une discipline vétérinaire essentielle. Elle permet de comprendre les transmissions intra-espèces et inter-espèces, de prévoir les périodes à risque et ainsi de limiter les transmissions. Dans une optique *One Health*, l'utilisation des IA pourrait permettre de surmonter les obstacles que rencontrent aujourd'hui les épidémiologistes (comme par exemple le partage des données, l'impact du climat sur les schémas de transmission, transmission à des hôtes multiples...)<sup>53</sup>.

Par exemple, *Walsh et al* ont utilisé un algorithme d'arbre de décision (ML) pour estimer la probabilité d'isoler le virus de la grippe aviaire sur des oiseaux sauvages. Leur étude a été réalisée lors de la période de surveillance du virus de 2006 à 2011 par le centre National de protection de la faune sauvage des États-Unis, sur 24 243 échantillons d'écouvillons oropharyngés et cloacaux d'oiseaux chassés ou capturés vivants. Cette étude cherchait à savoir si la méthode classiquement utilisée sur les oiseaux d'élevage (valeur cut-off du Ct de la PCR) pouvait être utilisée sur les oiseaux sauvages. En plus de prouver que cette technique pouvait être extrapolée aux oiseaux migrateurs, *Walsh et al* ont démontré que 16% des échantillons déclarés négatifs avec la méthode classique

étaient en fait des faux-négatifs. En effet, en inoculant ces échantillons dans des embryons d'œufs pour culture virale, les auteurs se sont rendus compte que certains étaient en fait contaminés par le virus de l'influenza aviaire, alors que leurs *Cycle threshold* (Ct) étaient supérieurs à 35 (la valeur cut-off) sur l'échantillon<sup>54</sup>.

En 2019, une étude menée en Ontario au Canada a comparé quatre algorithmes de ML (*random forest*, réseau de neurone, arbre de classification et arbre de régression) pour prédire la future tendance du virus de la diarrhée épidémique du porc. Les données utilisées comme gold standard, telles que les résultats des tests PCR, ont été récoltées auprès d'une entreprise réalisant un projet d'épidémiologie et travaillant avec 151 sites d'élevage de truies (avec chacun environ 750 truies). Le risque d'une nouvelle émergence de la maladie dans l'élevage était caractérisé par les termes « nul », « faible », « modéré » et « élevé », catégories définies par le nombre de nouveaux cas détectés par PCR sur 4 semaines. Les modèles de ML devaient alors être capables, en fonction des cas détectés, de classer les élevages dans une des 4 catégories de risque. L'algorithme de *random forest* était le plus performant des quatre algorithmes, avec une exactitude de 68%. Les résultats de l'IA ont été vérifiés sur des données déjà récoltées grâce à des validations croisées<sup>55</sup>.

## b) Comportement des animaux de compagnie

L'évaluation du comportement du chien et du chat peut permettre de connaître leurs besoins, s'ils souffrent, s'ils sont atteints d'une maladie sous-jacente.

La plupart des études menées sur l'utilisation du ML pour analyser le comportement des animaux va tenter d'évaluer un score de douleur. Chez le chat par exemple, deux études ont développé et évalué les performances d'algorithmes de ML et de DL pour classer les chats comme étant « douloureux » ou « non douloureux » à partir de photographies. Pour mettre en place ces modèles, les auteurs se sont basés sur l'échelle de grimace féline élaborée par l'université de Montréal en 2019.

Pour le travail de *Feighelstein et al*, des photos de la face de 84 chats admis au service de médecine et de chirurgie de l'université de Hanovre en Allemagne ont été utilisées pour l'étude. Le score douleur considéré comme « vrai » (gold standard) a été réalisé grâce au score douleur de Glasgow. Dans cette étude, l'algorithme (qui utilise lui l'échelle de grimace féline) le plus performant est exact à 77%<sup>56</sup>.

*Steagall et al* ont pour leur part développé une application de téléphone qui identifie correctement la douleur chez le chat dans 95% des cas. L'algorithme a été entraîné et testé sur un total de 4635 images de face de chats préalablement scorées grâce à l'échelle de grimace féline. Ces images proviennent principalement des animaux admis à la faculté vétérinaire de Montréal et de la plateforme *Kaggle* qui fournit des jeux de données<sup>57</sup>.

Une étude de *Boneh-Shitrit et al* basée sur l'analyse de 248 vidéos de chiens cherchait à développer et évaluer des algorithmes de ML (arbre de décision) et de DL (Réseaux de neurones convolutifs) pour reconnaître deux états sur 29 Labradors : anticipation (état positif) et frustration (état négatif). Ces deux états ont été volontairement induits chez les chiens en donnant une friandise (anticipation) ou en la rendant inaccessible (frustration), dans des conditions expérimentales contrôlées. Le réseau de neurones a obtenu les meilleures performances avec une exactitude de 89%<sup>58</sup>.

*Chambers* et son équipe ont quant à eux cherché à monitorer différents comportements et activités de chiens, que ce soit en clinique ou à la maison. Le but de cette étude était d'évaluer les performances d'un algorithme de Deep Learning à classer les chiens suivant différentes catégories d'activités (prise de boisson, alimentation, grattage, boiterie...), à partir de 1200 vidéos recueillies au cours d'un projet lancé par

l'institut WALTHAM Petcare Science. Un tel outil pourrait permettre de connaître les habitudes et besoins d'un chien et de mettre en évidence des comportements anormaux pouvant signifier une maladie sous-jacente. Les comportements alimentaires et la prise de boissons étaient les deux plus correctement détectés avec une sensibilité de 99% et 95% respectivement. D'autres comportements tels que le léchage, les reniflements, les grattages, les frottements et les caresses étaient également relativement bien détectés (Se d'au moins 60%)<sup>59</sup>.

### c) Cytologie et biochimie des carnivores domestiques

L'utilisation des IA pour automatiser la lecture et l'analyse des examens de cytologie et de biochimie, deux examens réalisés en routine, pourrait permettre une interprétation plus rapide et à moindre coût de ces résultats. Ces outils sont également les précurseurs d'une détection plus précoce ou tout du moins d'une surveillance plus accrue de certaines maladies pour lesquelles la suspicion clinique est assez tardive. Dans le domaine de la cytologie, un des leaders mondiaux en santé animale, Zoetis, a par exemple déjà mis sur le marché un outil intelligent d'aide à la lecture de lames cytopathologiques (*VetScan Imagyst*<sup>®</sup>).

C'est ce qu'ont pensé *Biourge et al.* Leur étude avait pour but de construire un réseau de neurones capable de prédire précocement le risque qu'un chat développe une maladie rénale chronique (MRC). L'algorithme a été construit grâce au recueil des données de 218 chats de l'Académie vétérinaire de Londres (The Royal Veterinary College) et testé sur plus de 3500 données issues du registre du Banfield Hospital aux États-Unis (algorithme *Renal Detect*). L'algorithme a été capable de dire si un chat allait développer une MRC dans les 12 mois dans 87% des cas, et s'il allait rester en bonne santé dans 70% des cas. Les prédictions se sont basées sur un suivi dynamique de l'urémie, de la créatininémie et de la densité urinaire à chaque visite, avec au moins une nouvelle visite à 18 mois d'intervalle. Leur objectif n'était pas d'établir un diagnostic mais de prédire quels chats présentaient le plus de risque face à cette maladie, afin de leur recommander une surveillance plus accrue<sup>60</sup> et une alimentation adaptée.

D'autres maladies assez fréquentes chez le chat, comme les maladies inflammatoires chroniques de l'intestin (MICI) et le lymphome digestif ont suscité l'intérêt des chercheurs. Ces deux maladies sont assez difficiles à différencier cliniquement et lors de l'interprétation des résultats (imagerie ou biochimie par exemple), et nécessitent des examens beaucoup plus invasifs pour poser un diagnostic définitif. Une étude de *Awaysheh et al* en 2016 a cherché à évaluer si des algorithmes de ML pouvaient constituer un nouvel outil de diagnostic performant. Les données de 120 chats ont été récupérées à l'école vétérinaire de l'université du Maryland en Virginie. Les auteurs ont comparé 3 modèles (Bayes, arbre de décision, réseau de neurones) et évalué s'ils étaient capables, avec de bonnes performances, de prédire si le chat était atteint plutôt d'une MICI ou d'un lymphome à partir des résultats de la numération formule sanguine (NFS) et d'une biochimie plasmatique. Les algorithmes ont reçu comme gold standard les résultats des biopsies intestinales. L'algorithme de Bayes et le réseau de neurones ont obtenu les meilleures performances avec une Se de 70,8% et 69,2% respectivement<sup>61</sup>.

Chez le chien, c'est l'hypoadrénocorticisme, ou Maladie d'Addison, qui a intéressé *Reagan* et son équipe. Ils ont développé un algorithme d'arbre de décision qui devait classer 1041 chiens (admis à l'école vétérinaire de l'université de Davis en Californie) dans les catégories « atteint » et « non atteint » de la maladie d'Addison selon les valeurs de la NFS et des analyses biochimiques (ionogramme, créatinine, glucose, protéine totale, albumine...). Le modèle, qui a appris grâce aux résultats des tests de stimulation à l'ACTH, a correctement classé 202 chiens sur 208 individus, lors de la phase de test, soit une erreur de 2,9% seulement<sup>62</sup>.

Chez les carnivores domestiques, l'analyse de semences assistée par ordinateur (*Computer Aided Sperm Analysis* ou *CASA*) est une technologie déjà bien établie. Des automates ont même déjà été développés pour aider à la quantification du nombre de spermatozoïdes dans un éjaculat mais aussi pour évaluer leur morphologie et leur motilité. Cet outil est essentiel pour l'évaluation de la fertilité du chien et du chat<sup>63</sup>.

Aujourd'hui, l'utilisation de ces technologies s'étend à l'étude d'autres cellules des animaux domestiques. L'équipe de *Vinicki et al* par exemple ont utilisé un Réseau de Neurones Convolutif (RNC) pour compter le nombre de réticulocytes « agrégés » (matures) sur un frottis sanguin. Le RNC utilisait une technique de détection d'objets (*Bounding Boxes* et classification) pour différencier ces cellules des autres cellules sanguines (érythrocytes, réticulocytes ponctués non matures...) à partir d'images issues d'un microscope (photo de la caméra intégrée et d'un smartphone). Sur un total de 1046 photos collectées à partir des frottis de 2 chats, 800 ont été utilisées pour la phase d'entraînement et les 246 autres pour l'étape de vérification du réseau. Ce dernier a obtenu une exactitude de 98,7%<sup>64</sup>.

#### d) Suivi automatisé de l'activité électrique cardiaque et cérébrale

L'électrocardiogramme ou ECG est un examen qui permet d'obtenir une représentation graphique de l'activité électrique du cœur. En médecine vétérinaire, c'est un appareil utilisé en routine pour monitorer les animaux hospitalisés (surtout les individus critiques) ou lors d'examens ou chirurgies nécessitant une anesthésie générale.

En 2020, une étude menée par *Estrada et al* voulait évaluer les performances d'un algorithme de lecture d'ECG en comparant son interprétation de 399 enregistrements à celle de 7 vétérinaires cardiologues titulaires d'un board de spécialiste. L'analyseur automatisé détectait les arythmies sur les ECG avec une Se de 99,7% et une Sp de 99,5%<sup>65</sup>.

L'électroencéphalogramme quant à lui est un examen complémentaire qui permet de connaître l'activité électrique du cerveau. Rarement utilisé en médecine vétérinaire, on peut le retrouver dans le diagnostic de l'épilepsie.

Une étude a démontré que le développement de RCN pourrait à l'avenir constituer un bon monitoring pour prévoir la survenue d'une crise d'épilepsie. *Nejedly et al* a réalisé l'étude sur les relevés de 75 EEG de quatre chiens, et son algorithme s'est révélé plutôt performant, avec une sensibilité moyenne de 79%<sup>66</sup>.

Les applications du *Deep Learning* en médecine vétérinaire sont nombreuses. L'utilisation de tels algorithmes dans des domaines aussi variés que la cytologie, le comportement ou encore l'étude des activités électriques cardiaques ou cérébrales démontrent leur capacité à s'entraîner sur tous types de données. En effet, les études citées dans cette partie ont porté sur le développement d'algorithme de DL utilisant des résultats de test PCR, des vidéos, des photographies ou encore des enregistrements d'ECG ou d'EEG et permettant d'obtenir des performances globalement bonnes.

### III) Extension au domaine de l'imagerie médicale

#### a) Historique

Dès les années 60, les chercheurs ont voulu automatiser l'interprétation des radiographies, avec des résultats largement décevants. C'est dans les années 80 que la notion de détection et de diagnostic assistés par ordinateur fait son apparition (*Computer aided Detection and Diagnosis*, CAD). Le CAD est défini comme la détection et/ou le diagnostic réalisé par un radiologue ou généraliste qui a pris en compte, comme un « second avis », les résultats d'une IA<sup>67</sup>. Habituellement, on divise le CAD en deux catégories : *Computer Aided Diagnosis* (CADx) et *Computer Aided Detection* (CADE)<sup>68</sup>. Le CADx caractérise les lésions, ils les rangent dans une catégorie prédéfinie (assimilable à de la classification). Le CADE quant à lui, est plutôt assimilable à de la localisation d'objets.

Le but maintenant n'est plus de remplacer les radiologues, mais de les assister dans leurs interprétations. Ce domaine du *Computer vision* a été développé dans le but de limiter les variations intra et inter individuelles<sup>22</sup>, d'augmenter l'efficacité de la lecture et du diagnostic (gain de temps et réduction des coûts).

Le développement du CAD a conduit à l'émergence d'une nouvelle technologie : La Radiomique. Cette méthode consiste en l'extraction de nombreux paramètres quantitatifs (appelés « caractéristiques radiomiques ») à partir d'images médicales qui, combinées ou non à d'autres données médicales (par exemple l'anamnèse, des résultats d'examens complémentaires...), vont permettre le développement d'outils précis d'aide au diagnostic et de prise de décision<sup>69</sup>.

Le suffixe -omique rappelle d'ailleurs d'autres domaines scientifiques offrant des explications détaillées sur de nombreux processus biologiques comme la génomique, la protéomique, la transcriptomique et la métabolomique<sup>70</sup>.

En médecine humaine, la radiomique est aujourd'hui fréquemment utilisée en oncologie. En effet, l'extraction de données telles que la taille, la forme, le volume ou la texture permet d'obtenir des informations cruciales sur le phénotype d'une tumeur et sur son microenvironnement. En médecine vétérinaire en revanche, cette technologie n'en est qu'à ses prémices, c'est pourquoi les quelques applications recensées ne seront pas développées dans cette étude<sup>71</sup>.

#### b) Applications actuelles

L'imagerie médicale en médecine vétérinaire est utilisée pour visualiser tous les organes, ou presque, des carnivores domestiques. Ce paragraphe va détailler les grands organes ou systèmes pour lesquels des IA ont été développées et testées. Nous nous intéresserons dans la dernière partie à l'imagerie pulmonaire en général, avant de s'attarder sur la détection des métastases.

##### i. Imagerie osseuse

C'est l'un des premiers articles utilisant des algorithmes de Machine Learning en imagerie médicale. *Mc Evoy* et son équipe ont été les pionniers dans le développement des Intelligences Artificielles en imagerie vétérinaire. L'article relate la comparaison des performances entre un réseau de neurones non linéaire et un algorithme de régression des moindres carrés partiels pour classer 200 images comme étant des radiographies

« de hanches » (120 images de hanches saines) ou « pas de hanches » (80 images de n'importe quelle autre région) chez le chien. Les images avaient été préalablement lues par un radiologue qui, sur chacune des radiographies, avait construit une *Region of interest* ou ROI (articulation de la hanche, tubérosité ischiatique et bord latéral du foramen obturé) pour s'assurer qu'il s'agissait bien d'un bassin (gold-standard). Les deux algorithmes ont obtenu une erreur inférieure à 10%, avec une sensibilité supérieure à 80%<sup>72</sup>.

*Shim et al* ont quant à eux utilisé deux algorithmes de *Deep Learning* sur 2 382 radiographies de genoux de chien. Les images proviennent de l'école vétérinaire de l'université de Konkuk à Séoul et de quelques autres hôpitaux partenaires. Le premier RNC a pour but de détecter l'articulation du genou (grâce à des *Bounding Boxes*), et ses performances sont comparées à une zone prédéfinie par un vétérinaire radiologue. Toutes les articulations du genou ont été repérées par l'IA. Le deuxième RNC est celui qui nous intéresse : il est utilisé pour prédire si l'articulation est « normale » ou s'il y a « un signe du tiroir », « une déviation de la patelle » ou « la formation d'ostéophytes ». Le gold standard pour connaître l'affection est contenu dans les comptes rendus rédigés par des vétérinaires radiologues. L'algorithme a su reconnaître les différentes affections presque aussi bien que les vétérinaires, avec une exactitude de plus de 80% pour les 4 catégories<sup>73</sup>.

## ii. Imagerie cardiaque

La radiographie permet de détecter une cardiomégalie généralisée ou localisée (atrium gauche et/ou droit notamment).

Un article de 2020 évalue les performances de quatre réseaux de neurones convolutifs pour classer 1465 radiographies de thorax en DLD en deux catégories : « présence d'une cardiomégalie » ou « absence de cardiomégalie », en apprenant à partir des conclusions des interprétations radiographiques. Les images proviennent des archives de l'école vétérinaire de l'université de Padua en Italie. Les quatre algorithmes ont une aire sous la courbe supérieure à 0,9, montrant une très bonne exactitude<sup>74</sup>.

*Li et al* ont quant à eux développé un RNC pour détecter une dilatation atriale gauche sur 792 radiographies de thorax de chien prises à l'hôpital vétérinaire de l'université de Davis en Californie. Les résultats de l'algorithme ont été comparés à ceux de deux vétérinaires cardiologues. Les résultats concordent dans 85,19% des cas<sup>75</sup>.

Enfin, *Boissady et al* ont évalué les performances d'un RNC entraîné à calculer le *Vertebral Heart Scale* (VHS) ou indice de Buchanan sur des radiographies thoraciques de chiens et de chats. Le réseau composé de 121 couches de neurones a préalablement été entraîné sur des radiographies issues d'un centre de référentiel (nombre d'image et localisation du centre non précisés dans l'article). Ces images ont été annotées par un des auteurs de l'étude. Pour évaluer les performances de l'IA, 30 radiographies de chats et 30 radiographies de chiens en DLD ont été sélectionnées au hasard dans la banque d'images. Pour chaque cliché, le VHS a été calculé par deux vétérinaires spécialistes (un cardiologue et un radiologue). La moyenne de ces deux mesures a ensuite été utilisée pour comparer le VHS calculé par les vétérinaires (vérité terrain) avec le VHS calculé par l'IA. Le degré de concordance entre les vétérinaires et l'algorithme était excellent, avec un Indice de corrélation intraclass de 0,998<sup>76</sup>.

Cette dernière étude fait probablement suite à une Thèse pour l'obtention du diplôme d'état vétérinaire réalisée en partenariat avec PicoxIA et menée par *Quentin Berche*<sup>77</sup>.

## iii. Imagerie hépatique

La radiographie est relativement peu indicatrice de lésions hépatiques, elle permet

seulement d'informer l'opérateur d'une hépatomégalie. Les articles suivants relatent de l'utilisation du *Deep Learning* sur deux autres types d'images médicales : les images échographiques et de scanner.

Pour le premier, l'article de *Banzato et al* évalue les performances d'un réseau de DL pour classer 48 images échographiques de foie présentant des lésions dégénératives (inflammatoires, vasculaires, dégénératives) ou non. Ces images sont issues de cas admis à l'école vétérinaire de l'université de Padua en Italie. Les auteurs ont dans un premier temps évalué les performances de l'algorithme puis ont comparé la valeur diagnostique d'un tel outil à celle des méthodes classiques (biochimie : PAL et ALAT et cytologie). Le gold standard pour déterminer quelle(s) lésion(s) étaient présentes sur le foie est l'histologie. D'après les résultats présentés, le réseau de neurones est un bon outil diagnostique (Se : 100%, Sp : 82,8%) contrairement aux méthodes standards<sup>78</sup>.

L'étude de *Burti et al* en 2021 s'est concentrée sur le développement d'un arbre de décision (ML) pour prédire le type histologique d'une lésion focale du foie entre quatre catégories : carcinome hépatocellulaire, hyperplasie nodulaire, autres lésions bénignes, autres lésions malignes. Le gold standard pour déterminer quelle lésion était présente sur le foie est également l'histologie associée à la cytologie. En se basant sur des caractéristiques quantitatives et qualitatives des 69 images de scanner (issues de cas admis à l'école vétérinaire de l'université de Padua en Italie), l'algorithme a obtenu une exactitude de plus de 75% pour les quatre lésions. Cependant, les auteurs rappellent que même si cet outil permet une bonne prédiction de la lésion, le diagnostic de certitude sera toujours la cytologie ou l'histologie<sup>79</sup>.

#### iv. *Imagerie du système nerveux (central et périphérique)*

Concernant le système nerveux, l'IRM est l'examen d'imagerie de choix. C'est pourquoi *Banzato et al* ont testé les performances d'un RNC pour différencier les méningiomes et les gliomes (les tumeurs du système nerveux central (SNC) les plus fréquentes du chien) sur 80 images d'IRM recueillies auprès de deux centres de référentiel : un à Zola Predosa en Italie et un autre à Six Mile Bottom en Angleterre. L'algorithme a été testé sur des images avant et après produit de contraste en T1 et des images en T2, obtenant une exactitude de 94%, 91% et 90% respectivement<sup>80</sup>.

Une autre utilisation du DL pour détecter des maladies du SNC a été étudiée par *Biercher et al* en 2021. Les auteurs ont utilisé 2 693 images d'IRM prises au Royal Veterinary College de Londres, en T1 et T2 pour entraîner puis tester un RNC sur des images de moelle épinière afin de détecter une extrusion du disque intervertébral (DIV), une protrusion du DIV, un embole fibrocartilagineux, une syringomyélie ou une néoplasie. Le RNC a donné de bons résultats pour la plupart des maladies, avec des sensibilités supérieures à 90% (sauf pour l'embole fibrocartilagineux qui a obtenu une Se de 62%) et des spécificités supérieures à 90%. Seule la détection des néoplasies et de la syringomyélie a donné des résultats décevants<sup>81</sup>.

L'étude de *Spiteri et al* s'est centrée sur le développement et l'évaluation d'un algorithme de ML (Kernel Support Vector Machine) pour évaluer les marqueurs (ou *features*) visibles sur une image permettant de différencier un chien avec ou sans malformation de Chiari (associé ou non à une syringomyélie). Les 32 images d'IRM en T2 proviennent d'un *Digital Imaging and Communications in Medicine* (DICOM). Les DICOM sont des visualiseurs d'images médicales, accessibles en ligne ou sur un réseau local, et donc partagés entre les différentes structures et services les utilisant. Ils permettent une mise en commun d'images et de comptes rendus médicaux. L'algorithme a mis en évidence les bons biomarqueurs sur les images pour la maladie de Chiari et pour la syringomyélie avec une Se de 82% et 93% respectivement.

#### v. *Imagerie pulmonaire (autres que les masses et nodules)*

En médecine vétérinaire, la radiographie est l'examen de choix, en première intention, pour explorer le thorax des carnivores domestiques. Les articles qui vont suivre traitent donc tous d'utilisation des IA pour l'aide à la lecture et à l'interprétation de radiographies. Dans cette partie, seule les atteintes pulmonaires non cancéreuses seront abordées, les métastases feront l'objet d'un autre paragraphe (cf. **Partie II- III c)**).

*Banzato et al* ont comparé les performances de deux RNC pour classer 1 741 radiographies (1062 en DLG et 679 VD) de thorax de chats en 5 catégories : « pas d'anomalie », « opacification bronchique », « opacification alvéolaire », « opacification interstitielle », « effusion pleurale » et « pneumothorax ». Les images proviennent des archives de l'école vétérinaire de l'université de Padua en Italie. Deux des auteurs, des radiologues avec beaucoup d'expérience, ont préalablement annoté les images avec les anomalies visibles. Les deux algorithmes ont obtenu les mêmes résultats, avec des aires sous la courbe supérieures à 0,9 (sauf pour les labels « pas d'anomalie » et « pneumothorax » qui ont une aire sous la courbe de 0,8), montrant globalement une bonne exactitude<sup>11</sup>.

*Banzato et al* ont réalisé une étude identique sur des radiographies de 3 839 chiens cette fois-ci, elles aussi issues des archives de l'hôpital universitaire vétérinaire de Padua. Pour cette étude, les deux algorithmes ont également obtenu des performances assez semblables, avec une aire sous la courbe supérieure ou égale à 0,8<sup>82</sup>.

L'étude de *Kim et al* s'est intéressée aux performances d'un logiciel intelligent pour classer 481 radiographies de thorax de chiens comme étant « sans anomalie » ou présentant un « œdème pulmonaire cardiogénique ». Les images ont été récupérées à partir d'un DICOM et ont toutes été réalisées aux urgences du *Schwarzman Medical Animal Center* à New York. L'exactitude, la Se et la Sp de l'IA comparées au diagnostic des radiologues (gold standard), étaient de 92,3%, 91,3% et 92,4% respectivement<sup>83</sup>.

Dans le même esprit, *Muller et al* se sont penchés sur le développement d'un algorithme de classification d'images radiographiques de thorax présentant ou non une effusion pleurale. Le RNC a été préalablement entraîné sur 4000 images issues d'un DICOM et provenant de multiples cliniques dans le monde. Pour leur étude, les auteurs ont comparé la classification de 173 radiographies de thorax de chien par l'algorithme avec le diagnostic posé. L'algorithme a obtenu une exactitude de 88,7% avec une Se de 90,2% et Sp de 81,8%<sup>84</sup>.

*Boissady et al*, ont de leur côté comparé les performances de 3 algorithmes de DL pour classer 120 radiographies de thorax. Les RNC ont préalablement été entraînés sur plus de 22 000 images radiographiques de thorax de chien et de chat. Les 7 labels (opacification interstitielle, opacification bronchique, opacification alvéolaire, opacification vasculaire, pneumothorax, effusion pleurale, scissures interlobaires) présents sur les radiographies ont pour gold standard la conclusion des comptes rendus rédigés par des radiologues. Toutes ces données proviennent d'un institut vétérinaire. Une fois le meilleur RNC évalué, ses performances seules ont été comparées à celles de 10 vétérinaires avec plus ou moins d'expériences seules ainsi qu'avec les performances de ces mêmes vétérinaires assistés par le RNC. Le réseau de neurones seul a obtenu les meilleures performances globales, suivi par les vétérinaires assistés par l'algorithme et enfin les vétérinaires seuls, avec des taux d'erreur de 10,7%, 16,8% et 17,2% respectivement<sup>85</sup>. Dans la même démarche, *Boissady et al* ont comparé les performances de 4 RNC avec celles de 13 vétérinaires radiologues pour classer des radiographies thoraciques selon plusieurs catégories (on retrouve les opacifications interstitielles, bronchiques, alvéolaires et vasculaires, le pneumothorax, l'effusion pleurale et les scissures interlobaires). Ils ont réuni 55 780 radiographies d'un institut vétérinaire, possédant toutes un compte rendu permettant à l'algorithme d'apprendre. Le gold standard utilisé dans cette étude a été l'accord de 3 vétérinaires spécialistes (totalement différents des

13 autres spécialistes) sur la nature des lésions visibles sur les radiographies. L'évaluation des performances a été réalisée sur la phase de test de l'IA, pour laquelle 50 images de thorax de chats et de chiens ont été analysées par l'algorithme et également par 13 vétérinaires boardés en radiologie. Chacun des 13 vétérinaires a identifié les lésions présentes sur les 50 images de la phase de test, sans en informer les autres radiologues. Globalement, les vétérinaires ont commis moins d'erreurs que les 4 RNC (qui eux ont presque tous les mêmes performances) pour la plupart des labels<sup>9</sup>.

*Pattaramanee Arsomngern* et son équipe ont utilisé 2 862 images radiographiques pour développer leur algorithme de diagnostic assisté par ordinateur (*Computer-Aided Diagnosis*, cf. **Partie II- III) a)**) pour la détection de lésions pulmonaires chez le chien et le chat. Les images ont été recueillies dans les archives de l'hôpital universitaire de Kasetsart en Thaïlande. Ce système nommé « PET-X » va dans un premier temps localiser les lésions pulmonaires sur l'image grâce à des *Bounding boxes*, puis les classer suivant 3 patterns d'opacité : alvéolaire, bronchique et interstitiel. Les lésions présentes sur les images ont été récupérées des comptes rendus rédigés par les radiologues de la faculté. La sensibilité et la spécificité varient d'une lésion à l'autre mais sont en moyenne de 70% et 75% respectivement<sup>10</sup>.

Les technologies de *Computer Aided Diagnosis* ont fait leur apparition dans les années 80. Aujourd'hui l'utilisation de ces algorithmes d'aide au diagnostic s'est démocratisée en médecine vétérinaire. Ces modèles informatiques sont désormais disponibles sur le marché et aident les vétérinaires au quotidien pour la lecture et l'interprétation d'images échographiques, radiographiques, d'IRM et de scanner.

Pour la détection des lésions pulmonaires autres que les métastases, les études réalisées démontrent que les algorithmes de *Deep Learning* offrent de bonnes performances. Il convient alors d'étendre l'utilisation de ces modèles informatiques à des lésions plus difficiles à mettre en évidence, telles que les métastases pulmonaires.

### c) Utilisation des IA dans le bilan d'extension thoracique : constats et perspectives

Les paragraphes précédents ont illustré l'intérêt porté à l'utilisation des IA en imagerie médicale vétérinaire. Il s'avère même que la plupart des résultats obtenus lors de ces différentes études sont relativement prometteurs, laissant imaginer l'utilisation de tels outils en clinique.

Pour la lecture et l'interprétation des radiographies thoraciques des carnivores domestiques, la plupart des lésions identifiables lors d'un tel examen d'imagerie sont relativement bien détectées par les différents RNC. Cependant, comme nous l'avons mentionné plutôt, ces performances sont décevantes quand il s'agit d'identifier des nodules pulmonaires, lors de bilan d'extension. Reprenons les études vues précédemment (cf. **Partie II- III) b) v.)** et intéressons-nous aux performances obtenues pour chaque RNC lors de la classification des labels « masse » ou « nodule ».

Concernant les deux études de *Banzato et al*, que ça soit sur des radiographies thoraciques de chats<sup>11</sup> ou de chiens<sup>82</sup>, les RNC avaient des performances plutôt mauvaises dans la classification des masses, contrairement aux autres lésions, avec une aire sous la courbe inférieure à 0,77 à chaque fois. Pour ces études, la quantité d'images contenant des masses ne représentait pas plus de 4% de la base de données totale.

Les 4 RNC de l'étude de *Boissady et al* ont également montré de mauvaises performances. Le taux d'erreur pour le label « masse/nodule » était plus élevé que le taux d'erreur moyen global (c'est-à-dire pour tous les labels, qui était de 3%), s'étendant de 31,4% à 41,2%<sup>9</sup>.

La deuxième étude de *Boissady et al* va aussi dans ce sens. Le label « masse » est le seul pour

lequel l'IA a un taux d'erreur plus élevé que les vétérinaires, s'élevant à plus de 20% contre 14% respectivement<sup>85</sup>.

Dans les deux études précédentes, les images radiographiques possédant des nodules pulmonaires représentaient 10% à 12% de la base de données.

Une étude n'ayant pas encore été citée s'est entièrement intéressée aux performances d'une IA disponible à l'achat en ligne (Vetology AI<sup>86</sup>). Ce RNC est un lecteur et interpréteur d'image radiographique, le groupe à son initiative autorise les études à son sujet pour assurer une totale transparence. *Pomerantz et al* ont alors étudié ses performances dans la classification d'image comme étant « sans anomalie » ou alors s'il y avait présence de « nodules/masses ». Ainsi, 153 images (43 chiens ayant permis d'obtenir trois vues orthogonales, et 12 autres ayant permis d'obtenir deux vues orthogonales) ainsi que les comptes rendus d'interprétation issus des archives de l'école vétérinaire de l'université de Tufts dans le Massachusetts ont été utilisés. Le gold standard consistait en tout examen complémentaire (histologie, cytoponctions...) ayant permis de confirmer la présence de métastases (vues préalablement sur la radiographie). Le RNC a obtenu une exactitude de 69,3%, une sensibilité (Se) de 55,4% et une spécificité (Sp) de 93,75%<sup>87</sup>.

Devant des performances si décevantes, contrairement à tout ce que les IA semblent pouvoir nous offrir, de nombreuses questions se posent afin de savoir pourquoi la classification des masses donne de telles performances ?

- Comme souligné dans les études précédentes, le nombre d'images contenant des masses ou des nodules était trop faible.
- Il existe dans le thorax, ou environnant, des structures dites « mass-like », c'est-à-dire des éléments physiologiques ayant des formes similaires à des masses (jonctions chondro-costales, mamelles, coupe transversale de bronches ou de vaisseaux) qui rendent la tâche d'autant plus difficile pour l'IA<sup>30</sup>. Ces éléments mass-like faussent la spécificité de l'IA.
- De plus, les métastases se présentent sous de nombreuses formes, opacités et tailles qui, une fois de plus, rendent leur classification plus complexe<sup>17</sup>.

Il est également important de noter que toutes les études citées précédemment ont utilisé des algorithmes de classification pour détecter la présence de nodules ou masses pulmonaires. Une telle approche, rappelons-le, permet uniquement de classer une image dans une catégorie : « présence de masse » ou « absence de masse ». Et c'est probablement en cela que les performances des différents RNC ont déçu les divers auteurs. Il est difficile d'assurer la présence ou non d'un élément, qui de plus est inconstant en présentation radiographique, si on ne sait pas où le chercher. Les algorithmes de classification ont donc également des défauts de sensibilité. Mais ces données démontrent-elles l'incapacité des Intelligences Artificielles à mettre en évidence des nodules pulmonaires ?

Pour ce qui est de la classification, le constat en médecine humaine est le même que précédemment : les performances sont décevantes. C'est par exemple le cas pour l'équipe de *Wang et al* qui a utilisé 112 120 radiographies issues de la base de données en libre accès *ChestX-ray 14* pour entraîner un algorithme de classification. Les 15 labels, dont 2 concernant les « nodules » et les « masses » ont pour gold standard les comptes rendus médicaux des patients. L'aire sous la courbe pour ces deux catégories sont respectivement de 0,67 et 0,69, faisant ainsi partie des plus mauvais résultats avec les labels « pneumonie » et « infiltrations »<sup>88</sup>.

Dès lors que l'on s'intéresse à d'autres techniques de Computer Vision, les résultats offrent des perspectives plus optimistes. L'étude de *Nam, Park et al* par exemple a utilisé 43 292 radiographies thoraciques (dont 9 225 présentant des métastases), collectées entre janvier 2010 et décembre 2015 à l'hôpital de l'Université nationale de Séoul pour développer un algorithme de détection par DL. Le gold standard pour cette étude était une confirmation histologique de la nature de la masse. La localisation de la masse et ses contours ont préalablement été annotés

(annotation au pixel ou segmentation sémantique). Sur la phase de validation, pour les différentes bases de données, le taux d'erreur de l'algorithme varie de 2% à 34%. Les auteurs ont également comparé les performances de 18 médecins (dont 10 spécialistes en radiologie) avec et sans l'aide du programme de DL. Les statistiques montrent une amélioration globale de la performance des opérateurs lorsqu'ils sont assistés par l'IA, avec une différence significative pour 15 d'entre eux ( $p$ -value < 0,05)<sup>89</sup>.

*Behrendt et al* ont quant à eux développé un algorithme de localisation de masses pulmonaires entraîné et testé à partir de 4 882 radiographies pulmonaires issues de 4 bases de données accessibles au grand public (*ChestX-ray 14*, *PadChest*, *JRST* et *Open I*). Parmi elles, 1 134 contenaient au moins un nodule. Toutes les radiographies ont préalablement été vérifiées par des radiologues et les masses localisées manuellement par des *Bounding boxes* (*Bboxes*). Pour calculer les performances de l'IA, il est admis, dans cet article, que si la Bbox prédite par l'algorithme et celle dessinée par l'opérateur se chevauchent d'au moins 20%, alors la prédiction est un vrai positif. Sur l'étape de validation, le modèle développé par les auteurs (qui est en réalité l'association de 4 algorithmes préexistants : Faster-R-CNN, RetinaNet, EfficientDet-D2 and YoloV5) a obtenu une aire sous la courbe de 0,87<sup>90</sup>. Ces performances semblent plutôt prometteuses mais ont été obtenues avec un seuil de tolérance plutôt faible.

D'autres études sur l'utilisation des IA en médecine humaine ont obtenu de très bons résultats, mais cette fois-ci sur des images de scanner, technique d'imagerie beaucoup plus sensible que la radiographie et largement plus utilisée qu'en médecine vétérinaire. On citera par exemple l'étude de *Xu et al* qui évalue les performances d'un RNC pour détecter des nodules pulmonaires sur 1186 images de scanner issues de la banque de donnée publique *LUNA16*. Les images sont accompagnées d'un fichier texte donnant la localisation des masses. Les performances du RNC ont été évaluées sur l'algorithme de base et sur l'algorithme « amélioré » (6 améliorations au total avec entre autres l'amélioration de la fonction loss, l'utilisation d'une fonction de convolution déformable). L'algorithme le plus amélioré a obtenu une précision moyenne de 90,7%<sup>91</sup>.

Au vu des exemples décrits précédemment, il semble nécessaire de rechercher une nouvelle approche pour la détection des métastases pulmonaires, les performances des algorithmes actuellement développés n'étant pas satisfaisantes. Les études réalisées en médecine humaine, n'utilisant pas que des algorithmes de classification, pour la plupart d'entre elles, offrent des performances prometteuses qu'il conviendrait de transposer en médecine des animaux de compagnie.

En médecine des carnivores domestiques, une part non négligeable des cancers vont métastaser dans les poumons. De ce fait, les propriétaires ont des exigences croissantes quant à l'expertise du vétérinaire pour détecter au plus vite les nodules pulmonaires. Pour répondre à cette demande, le vétérinaire a la possibilité de choisir entre différents d'examen d'imagerie afin de réaliser un bilan d'extension. Même si le scanner semble être le choix le plus sûr, du fait de sa grande sensibilité, il ne constitue pas aujourd'hui l'examen de choix. L'utilisation de la radiographie est beaucoup plus justifiée car c'est un examen disponible dans un grand nombre d'établissement et dont l'utilisation est nettement plus démocratisée.

Malgré des performances relativement bonnes dans la plupart des domaines, les IA ont montré des résultats plutôt mauvais lorsqu'il s'agissait de détecter des métastases pulmonaires. Cependant, les recherches actuelles en médecine vétérinaire ne s'appuient que sur des algorithmes de classification d'images. Dès que l'on s'attarde sur d'autres techniques de *Computer Vision*, retrouvées en médecine humaine par exemple, les perspectives d'utilisation des IA dans le bilan d'extension thoracique sont plus prometteuses.

C'est dans ce contexte que nous avons décidé, dans cette étude, de développer un algorithme de *Deep Learning* innovant se basant sur une approche de localisation spatiale afin de détecter des métastases sur des radiographies thoraciques de chiens et de chats.

# Partie 3 – Exécution d’un algorithme de Deep Learning pour la localisation de nodules thoraciques

## I) Contexte et enjeux de notre étude

La présente étude a été réalisée en collaboration avec l’entreprise PicoxIA, une start-up française, créée en 2019, proposant aux vétérinaires des outils de lecture radiographique. Ces technologies, basées sur l’utilisation du *Deep Learning*, ont pour but d’offrir une deuxième lecture objective des clichés et d’aider les vétérinaires dans leur interprétation.

A l’heure actuelle, PicoxIA a déjà développé des algorithmes de classification capables d’analyser les trois régions anatomiques suivantes chez le chien et le chat :

### 1- Les hanches

L’algorithme de *Deep Learning* permet de détecter 5 lésions au niveau de l’articulation coxo-fémorale afin d’aider au diagnostic des dysplasies de la hanche. Il peut également tracer automatiquement les angles de Norberg-Olsson, donnant une indication sur le recouvrement de la tête fémorale par l’acétabulum. Un compte rendu peut alors être généré par le modèle afin de classer le stade de la dysplasie du chien selon la grille de classification établie par la Fédération Cynologique International (FCI).

### 2- L’abdomen

Cette fois-ci, l’algorithme peut détecter 13 lésions abdominales et proposer 3 hypothèses de diagnostic différentiel (Ddx) pour une lésion donnée. Un compte rendu est également généré automatiquement par le modèle.

### 3- Le thorax

Comme vu dans les 2 études de *Boissady et al* citées dans la **Partie 2- III) b) v.**, l’algorithme développé par PicoxIA peut également détecter 16 lésions thoraciques (opacifications alvéolaires, interstitielles, bronchiques, masses pulmonaires...), avec des performances plus ou moins satisfaisantes suivant la lésion et proposer 4 hypothèses de Ddx pour chaque anomalie détectée. Il propose également le tracé automatique du VHS (*cf. Partie 2 – III) b) ii.*) Encore une fois, un compte rendu est automatiquement généré par le modèle.

Pour la région du thorax, qui nous intéresse tout particulièrement, nous rappelons que les performances de l’algorithme proposé par PicoxIA pour classer le label « masse/nodule » sont aujourd’hui plutôt mauvaises comparées aux autres labels, avec des taux d’erreurs moyens globaux entre 31% et 42% pour les masses, contre environ 3% pour le reste des labels.

Dans ce contexte, l’entreprise souhaiterait optimiser les performances de détection de cette classe de lésions, notamment en utilisant une autre approche de *Computer Vision*.

Notre étude avait pour but de développer puis d’évaluer les performances d’un algorithme de *Deep Learning* pour la localisation de masses et de nodules sur des radiographies de thorax de chiens et de chats. L’intérêt de cette étude était de vérifier si une telle approche était au moins tout aussi voire plus performante que les algorithmes de classification déjà existants, et de pouvoir ainsi la proposer comme outil d’aide au diagnostic dans le cadre d’un bilan d’extension. Les étapes nécessaires au développement d’un tel outil sont détaillées dans la partie ci-dessous.

Ce travail est, comme mentionné plus haut, réalisé en partenariat entre l'auteur, l'École Vétérinaire Oniris VetAgroBio et l'entreprise PicoxIA. Il a été réalisé dans l'espoir d'optimiser les résultats précédemment obtenus par les modèles de PicoxIA et constitue une approche innovante encore non développée par l'entreprise.

## II) Constitution de la base de données nécessaire à l'entraînement de l'algorithme de Deep Learning

### a) Estimation du nombre d'images nécessaires

Comme explicité dans la précédemment, le choix du nombre de données nécessaires à l'apprentissage d'une IA est une tâche cruciale mais difficile. Pour notre étude, nous avons choisi empiriquement un N-minimum en s'inspirant d'autres travaux.

Voici un exemple pour illustrer notre propos : *Pattaramanee Arsomngern* et son équipe ont utilisé 2 862 images radiographiques pour développer leur algorithme de diagnostic assisté par ordinateur (*Computer-Aided Diagnosis*, cf **Partie 2- III a)**) pour la détection de lésions pulmonaires. L'algorithme nommé « PET-X » va dans un premier temps localiser les lésions pulmonaires sur l'image grâce à des *Bounding boxes*, puis les classer suivant 3 *patterns* d'opacité : alvéolaire (représentant 21% des clichés soit 608 radiographies) , bronchique (représentant 29% des clichés soit 842 radiographies) et interstitiel (636 images, représentant 22% des clichés). La sensibilité, la spécificité et l'exactitude varient d'une lésion à l'autre mais sont en moyenne de 81%, 64% et 72% respectivement<sup>10</sup>.

Comme notre étude portait sur la localisation d'une seule lésion, les nodules pulmonaires, on a pu supposer que, pour obtenir des performances au moins aussi correctes, le nombre d'images dans notre étude devait être sensiblement le même que dans l'étude citée. De plus, dans la majorité des cas, les métastases pulmonaires sont assimilables à une atteinte de type alvéolaire ou interstitielle et de forme arrondie. Ces informations laissaient envisager qu'un N minimum pouvait être fixé à 622 clichés, représentant le nombre moyen de radiographies utilisée dans l'étude citée pour les labels « opacification interstitielle » et « opacification alvéolaire ».

### b) Population d'étude

Afin de constituer notre population d'étude, nous avons cherché dans le logiciel Sirius, une plateforme commune aux quatre Écoles Nationales Vétérinaires de France, les animaux admis au service d'imagerie de l'école Oniris VetAgroBio pour lesquelles la première radiographie effectuée amenait à une suspicion de masses pulmonaires. La recherche a été réalisée à partir de la section Recherche avancée de l'onglet « Recherche Fiche de visite ». Les mots-clés suivants ont été utilisés : « métastase », « masse », « nodule », « bilan d'extension », « radiographie », « thorax ».

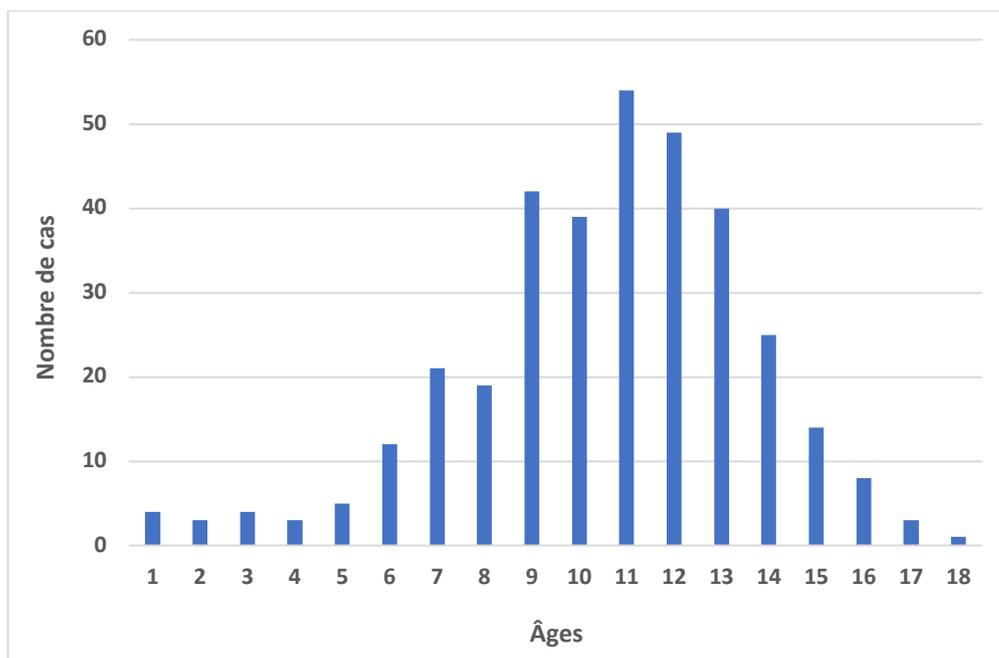
Cette recherche a permis de récupérer 513 cas admis à l'école vétérinaire de Nantes entre 2002 et 2024, réunissant 161 chats et 352 chiens. Pour 351 cas parmi ces 513 animaux, la première radiographie confirmait sans équivoque la présence d'un ou plusieurs nodules pulmonaires, d'après le compte rendu (CR) d'imagerie. Pour les 162 autres, la première radiographie ne permettait pas de conclure avec assurance, la présence d'une masse dans le thorax étant alors qualifiée de « douteuse ». Pour ces 162 derniers, ils étaient répartis comme suit entre les 2 espèces d'étude:

- 49 chats,
- 113 chiens.

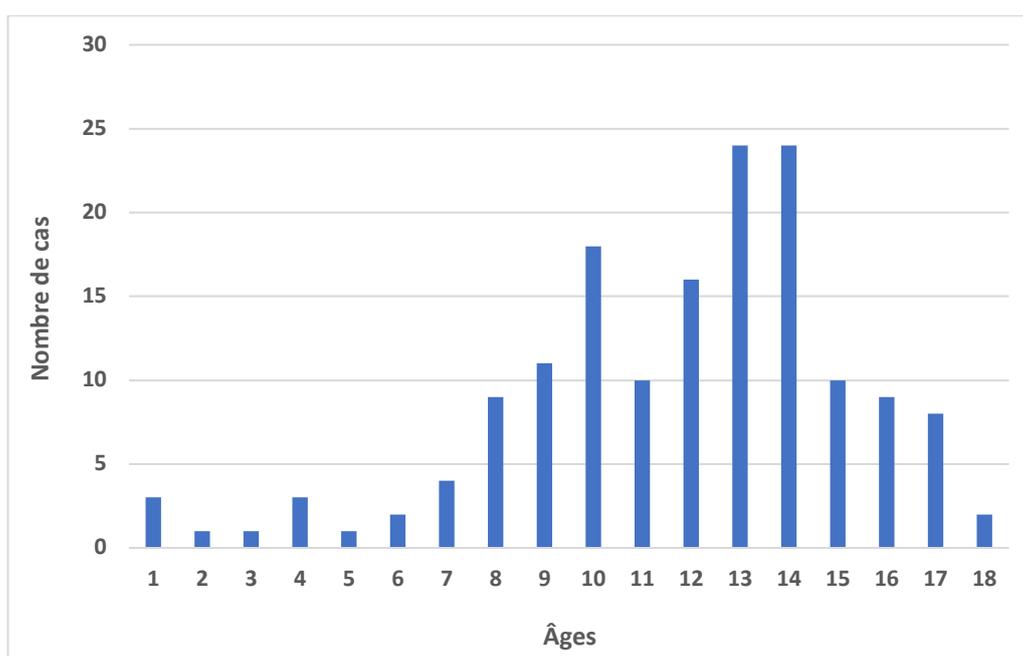
A noter qu'aucune image d'animaux « sains », c'est-à-dire n'ayant pas de métastases pulmonaires, n'a été recueillie. L'algorithme qui sera développé n'a pas besoin de ce type de données pour son apprentissage (cf. **Partie 3- IV b) i.**). De ce fait, seule la sensibilité pourra être utilisée pour évaluer les performances de notre modèle (cf. **Partie 3- IV b)**).

Les commémoratifs et l'anamnèse de la population d'étude ont également été récupérés.

L'âge médian lors de la réalisation de la première radiographie thoracique était de 11 ans pour les chiens (s'étendant de 1an à 18 ans) et de 12 ans pour les chats (entre 1 an et 18 ans également) . Les histogrammes des **Figures 11 et 12** indiquent la proportion de cas par âge pour les espèces canine et féline respectivement.



**Figure 11.** Répartition des âges des 352 chiens à la première radiographie.



**Figure 12.** Répartition des âges des 161 chats à la première radiographie.

Pour les chiens, on a observé que la majorité des cas se situait entre 9 et 13 ans, et qu'elle était plutôt entre 10 et 14 ans chez le chat. En effet, les animaux âgés sont plus à risque de développer des tumeurs, et *in fine* des métastases pulmonaires, que les individus de jeunes âges. Ce constat est en accord avec le principe de mutation d'une cellule somatique en une cellule tumorale explicité dans la **Partie 1- I) a)**.

A propos des races de chats, seuls 11 d'entre eux étaient des chats de race (7 Persans, 2 Chartreux, 1 Exotic Shorthair et 1 Sacré de Birmanie). Les 151 autres étaient de type Européen.

Pour le chien, on retrouve une plus grande diversité. Sur les 351 chiens, 51 étaient issus d'un croisement. Pour 69 patients canins, la race n'avait pas été renseignée dans le logiciel Sirius. Enfin, 231 chiens sur les 351 cas recueillis étaient des animaux de races. Parmi les plus représentées dans la population on retrouvait :

- Le Bouvier Bernois (17 cas)
- Le Boxer (12 cas)
- Le Caniche (15 cas)
- Le Cocker (11 cas)
- L'Épagneul Breton (15 cas)
- Le Labrador (21 cas)
- Le Setter (12 cas)

Les chiens de races moyennes à grandes étaient donc surreprésentés dans notre population d'étude.

Concernant l'anamnèse, sur ces 513 animaux, 26 animaux (6 chats et 20 chiens) présentaient une tumeur pulmonaire primitive, 369 ont fait l'objet de clichés radiographiques dans le cadre d'un bilan d'extension et les 120 autres ont consulté pour des raisons diverses et variées, hors motif oncologique. La mise en évidence d'une ou plusieurs masses pulmonaires constituait ainsi une découverte fortuite pour ces 120 derniers cas.

Comme remarqué précédemment, 369 animaux présentaient une tumeur primitive sur un organe autre que les poumons. La **Tableau 3** illustre les différentes tumeurs primitives rencontrées et le pourcentage de cas recueillis en fonction de l'espèce ainsi que dans la population d'étude. La catégorie nommée « Autres » dans le tableau rassemble toutes les tumeurs primitives représentant moins d'un pour cent des tumeurs diagnostiquées dans notre population de 513 animaux. On y retrouve entre autres le carcinome thyroïdien et l'hémangiosarcome. Pour les tumeurs pulmonaires primitives, le **Tableau 4** illustre la répartition des cas en fonction de la nature de la masse dans l'échantillon de 26 animaux atteints et pour chaque espèce d'étude. La mention « Non renseignée » désigne les masses pour lesquelles la nature de la tumeur primitive n'a pas été identifiée. Il est souligné que sur les 26 cas de tumeurs pulmonaires, 6 ont été présentés en consultation pour un syndrome de Cadiot-Ball (ou syndrome doigt-poumons), avec pour motif de consultation commun une boiterie ou une douleur à la manipulation d'un membre. .

Pour compléter notre base de données, l'entreprise PicoxIA a également fourni à l'auteur des radiographies thoraciques de chiens et de chats récupérées auprès de leurs cliniques partenaires. Ses images ont été envoyées à l'entreprise afin d'entraîner leur algorithme de classification (*cf.* **Partie 3- I)**). Ce jeu de données ayant été transmis totalement anonymisé, les commémoratifs et l'anamnèse des animaux ayant fait l'objet de ses radiographies n'ont pas pu être récupérés.

**Tableau 3.** Répartition des différentes tumeurs primitives rencontrées dans les 2 espèces étudiées et proportion dans la population d'étude.

Tumeur primitive	Proportion chez le chat	Proportion chez le chien	Proportion dans la population d'étude
Tumeur mammaire	40%	30%	33%
Tumeur osseuse	2%	4%	4%
Sarcome	1%	4%	3%
Lymphome	1%	3%	2%
Histiocytose maligne	0	2%	1,2%
Tumeur abdominale	2%	1%	1,2%
Tumeur buccale	2%	2%	1,3%
Tumeur cutanée	1%	2%	1,7%
Mastocytome	0	3%	1,7%
Tumeur pulmonaire	4%	6%	5%
Autres	19%	18%	18%

**Tableau 4.** Répartition des cas pour l'espèce canine et féline en fonction du type de tumeur pulmonaire et proportion dans l'échantillon de 26 animaux atteints d'une tumeur pulmonaire primitive.

Nature de la tumeur pulmonaire	Proportion chez le chat	Proportion chez le chien	Proportion dans l'échantillon total
Carcinome bronchique	1/6	2/20	3/26
Carcinome bronchoalvéolaire	1/6	1/20	2/26
Carcinome pulmonaire	1/6	2/20	3/26
Non renseignée	3/6	15/20	18/26

Les tumeurs mammaires étaient les tumeurs primitives les plus diagnostiquées dans notre population d'étude. Rappelons que ces tumeurs font partie des plus représentées chez le chien et le chat, et qu'elles possèdent un fort pouvoir métastatique pulmonaire.

Notre population d'étude était ainsi composée de 351 chiens et 162 chats pour un total de 513 animaux.

Les cas recueillis représentaient une tranche d'âge sénior pour la majorité.

La population féline était globalement de type européen. Les races de chiens étaient plus diversifiées avec une surreprésentation des races de grands et moyens formats.

### c) Critères d'inclusion

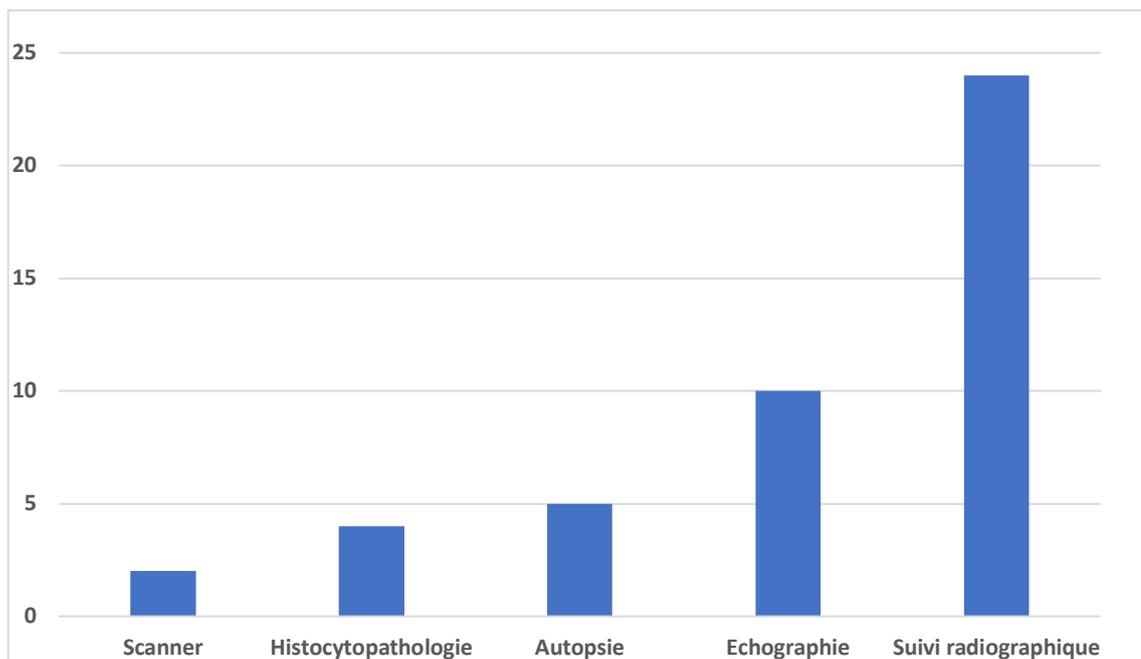
Nous rappelons que lors de la mise en évidence d'une masse ou d'un nodule pulmonaire, le diagnostic différentiel concernant la nature de ces lésions est le suivant :

- Les masses pulmonaires primitives,
- Les métastases,
- Les abcès,
- Les granulomes,
- Les hématomes,
- La métaplasie osseuse (lésions nodulaires millimétriques de densité plutôt minérale).

Nous rappelons aussi que pour 351 cas sur les 513 recueillis, la seule lecture des clichés radiographiques a permis d'affirmer la présence de nodules pulmonaires. Pour les 162 autres cas qualifiés de « douteux », les critères suivants ont été utilisés afin de s'assurer de la nature tumorale ou métastatique des masses pulmonaires présentes sur les radiographies thoraciques :

- L'augmentation de la taille ou du nombre de masses dans le cadre d'un suivi radiographique (avec ou sans mise en place de traitement),
- La réalisation d'un autre examen complémentaire plus sensible et/ou plus spécifique que la radiographie venant confirmer la suspicion.

L'histogramme de la **Figure 13** illustre la répartition des 162 cas douteux en fonction des examens complémentaires réalisés afin de confirmer la suspicion.



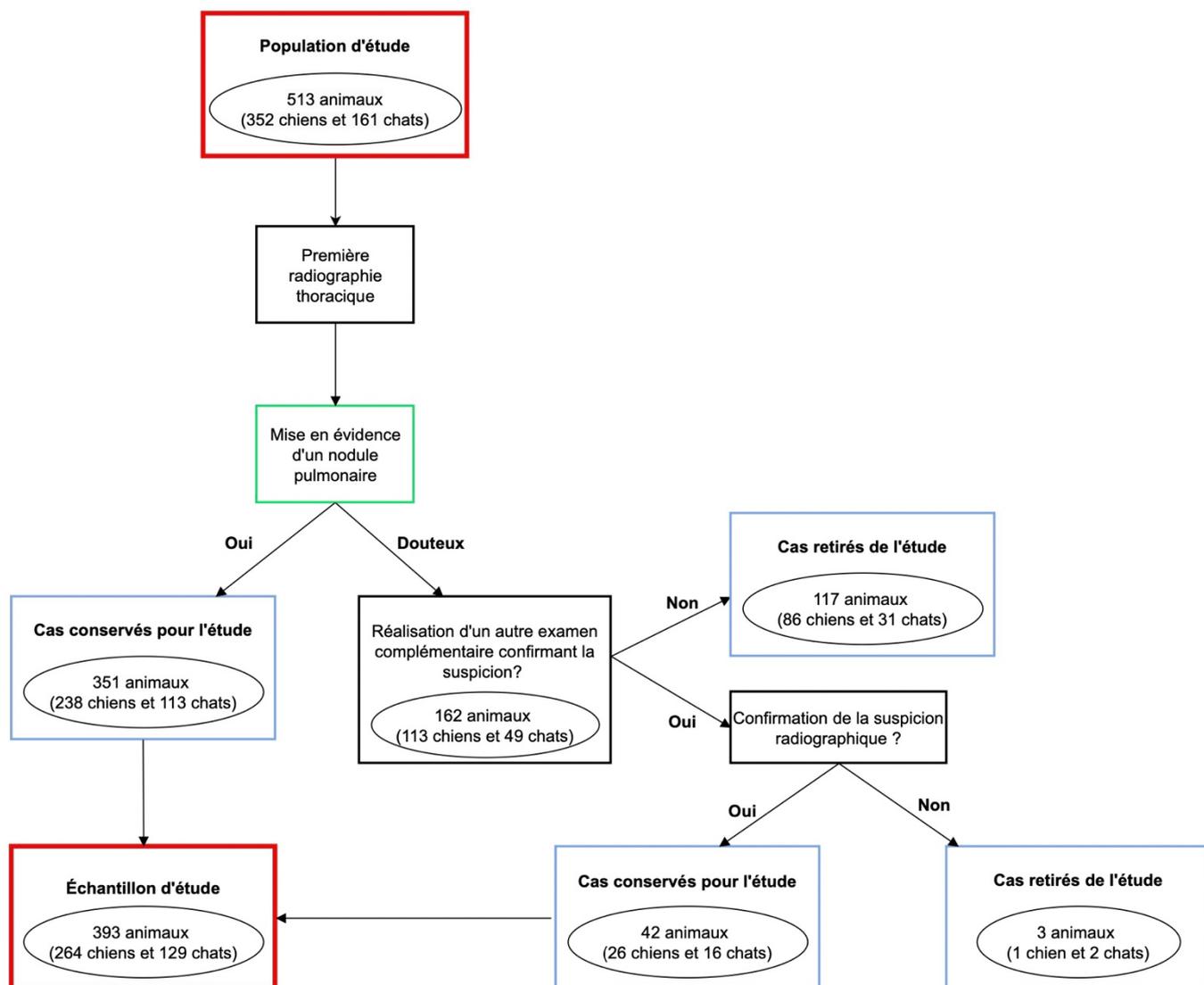
**Figure 13.** Examens complémentaires réalisés pour confirmer la suspicion de métastases et le nombre de cas recensés pour chacun sur les 162 animaux concernés.

Sur les 162 cas douteux, seuls 45 ont fait l'objet d'un autre examen complémentaire. Parmi eux, 3 animaux ont fait l'objet d'un examen complémentaire qui a infirmé la présence de métastases. Pour s'assurer que les annotations des nodules soient le plus proche de la réalité possible, les 117 résultats douteux à la première radiographie et n'ayant pas fait l'objet d'autres examens ont ainsi été retirés de la banque de données.

Enfin, pour limiter au maximum l'impact de la qualité des radiographies sur les performances de l'algorithme, un tri a été réalisé afin de constituer une banque d'images avec les critères d'inclusion suivants :

- Des radiographies thoraciques de chiens ou de chats en décubitus latéral droit et/ou gauche et/ou une incidence DV/VD,
- Une qualité radiographique correcte. Les rotations, défauts d'extension des membres et défauts de constantes (contraste et exposition) sont tolérés tant qu'ils ne gênent pas à la lecture de la radiographie,
- Des images où les métastases se présentent sous la forme de masses ou de nodules: les infiltrations métastatiques ne sont donc pas retenues.

Le diagramme de la **Figure 14** reprend les critères d'inclusion précédemment décrits. Cette représentation graphique permet de visualiser le passage de la banque de données initiales de 513 animaux à l'échantillon d'étude de 393 animaux.



**Figure 14.** Diagramme de flux récapitulant les critères appliqués sur la population d'étude initiale de 513 animaux et qui ont conduit à l'obtention d'un échantillon de 393 cas.

Sur les 513 cas constituant notre population d'étude, nous avons réalisé un tri afin de s'assurer que les nodules mis en évidence sur la première radiographie réalisée étaient bien de natures tumorale ou métastatique. Les critères instaurés par l'auteur ont permis de recueillir un échantillon de 393 cas pouvant être intégrés à notre étude.

d) Récupération des clichés radiographiques et constitution de notre échantillon d'étude

A partir de l'échantillon de 393 cas, les radiographies ont été téléchargées à partir du logiciel Synapse. Ce dernier est un DICOM regroupant tous les examens d'imagerie (radiographie, échographie, scanner...) réalisés dans l'enceinte de l'École Oniris VetAgroBio. Les radiographies ont été téléchargées puis anonymisées au format JPEG. L'anonymisation concerne :

- La présence d'informations relatives aux propriétaires présentes directement sur l'image (anonymisation des pixels),
- La présence d'informations sur le fichier visuel en lui-même (anonymisation des métadonnées).

Malheureusement, pour 185 des 393 cas, pour lesquelles les radiographies ont été réalisées entre 2002 et 2013, les images n'avaient pas été numérisées et n'étaient donc pas disponibles sur le DICOM. Aussi, pour les radiographies de 10 autres cas datant de 2015, une anomalie lors de leur numérisation sur le DICOM les rendaient inexploitable.

C'est finalement un échantillon de 211 animaux qui a pu constituer notre population d'étude.

Cet échantillon de 211 cas recueillis a permis d'obtenir au total 537 images radiographiques de thorax de chiens et de chats dans la base de données de l'école Oniris VetAgroBio. Chacun des 211 cas offraient au moins 2 radiographies, en raison de la réalisation d'au moins deux incidences (VD et DLD ou DLG), voir plus lorsqu'il s'agissait de suivi radiographique, ou lorsque trois incidences étaient réalisées (VD, DLG et DLD).

L'entreprise PicoxIA quant à elle a fourni à l'auteur un total de 2274 images radiographiques.

La base d'images ainsi récoltée était constituée de 897 clichés thoraciques.

Les images du jeu de données de PicoxIA n'ayant été fournies avec aucune information, la présence ou l'absence de métastases sur ces radiographies a été déterminée par l'auteur.

En ajoutant les critères d'inclusion décrits précédemment, nous avons finalement pu réunir l'ensemble de clichés suivants :

- 337 images radiographiques issues de l'échantillon d'étude de 211 cas. Cet ensemble constituera le jeu de données d'Oniris.
- 374 radiographies issues des 897 fournies par PicoxIA.

Le total était ainsi de 711 radiographies thoraciques de chiens et de chats.

Le **Tableau 5** illustre la répartition des 711 radiographies constituant la banque d'apprentissage en fonction de la taille des métastases, lorsqu'une seule métastase était présente sur la radiographie. Si elles étaient au nombre de 2 ou plus, la radiographie était rangée dans la catégorie « plusieurs masses ». Enfin, une séparation a été effectuée lorsqu'il s'agit d'une opacification milliaire. Nous rappelons qu'il n'y avait aucune image dans cette base de données qui ne contenait pas du tout de métastases.

**Tableau 5.** Répartition des radiographies de l'étude en fonction de la taille des masses.

	Petite	Moyenne	Grande	Opacification miliaire	Plusieurs masses
Nombre de cas	58	225	147	18	263
Pourcentage dans la population	8%	32%	21%	3%	36%

La taille des nodules a été définie comme suit :

- Petite taille : nodule < 1cm,
- Taille moyenne : 1cm < nodule < 3cm,
- Grande taille : nodule > 3cm, communément nommé « masse ».

La définition de ces tailles a été adaptée de la littérature<sup>17</sup> afin d'obtenir une répartition plus homogène des cas. Les tailles des nodules sur les radiographies thoraciques ont été mesurées par l'auteur lui-même.

Un dernier échantillonnage parmi les 711 radiographies a été réalisé. La base de données a été séparées en 5 lots d'images distincts. Les intérêts d'un rééchantillonnage étaient d'investiguer les hypothèses suivantes :

- Comparer les performances de l'algorithme de DL sur les deux jeux de données (PicoxIA et Oniris VetAgroBio),
- Comparer les performances de l'algorithme selon la taille des masses,
- Voir l'impact du nombre d'images de la banque de données sur les performances du modèle de détection.

Dans chaque lot, une importance a été accordée à l'origine des radiographies. C'est-à-dire que les deux jeux de données, l'un issu des radiographies provenant du service d'imagerie d'Oniris, et l'autre issu des images fournies par PicoxIA n'ont pas été mélangés dans les lots. Les lots sont constitués comme suit :

- Lot 1 : 148 radiographies issues du jeu de données d'Oniris,
- Lot 2 : 119 radiographies issues du jeu de données de PicoxIA,
- Lot 3 : 99 radiographies issues du jeu de données d'Oniris,
- Lot 4 : 95 radiographies issues du jeu de données de PicoxIA,
- Lot 5 : 250 radiographies issues du jeu de données de PicoxIA.

La répartition des clichés dans les lots 1 à 5 s'est appuyée sur la répartition de la taille et du nombre de métastases. Cette répartition est indiquée dans le **Tableau 6**.

**Tableau 6.** Répartition des clichés radiographiques dans chaque lot en fonction de la taille des métastases, en pourcentage.

	Petite	Moyenne	Grande	Opacification miliaire	Plusieurs masses
Lot 1	9%	43%	26%	2%	20%
Lot 2	2%	26%	26%	3%	43%
Lot 3	8%	26%	21%	6%	39%
Lot 4	16%	25%	16%	2%	41%
Lot 5	8%	32%	16%	1%	43%

On remarque que chaque lot se distinguait par une répartition différente du nombre de métastases présentes sur chaque image mais également par la taille des nodules présents.

Toutes les radiographies issues de l'échantillon de 393 cas n'ont malheureusement pas pu être téléchargées à partir du DICOM de l'école vétérinaire Oniris VetAgroBio. Ce problème s'explique par des défauts lors de la numérisation des clichés sur le logiciel Synapse, notamment pour les radiographies antérieures à 2013. Cette étape a permis de rassembler 337 radiographies de thorax de chiens et de chats. A cela sont venus s'ajouter 374 clichés issus du jeu de données de PicoxIA, pour un total de 711 radiographies exploitables.

### III) Annotations des radiographies en vue de la phase d'apprentissage de l'algorithme

#### a) Logiciels utilisés pour l'annotation et l'apprentissage

##### i. Visual Geometry Group Image Annotator (VIA)

VIA est un logiciel d'accès libre sur n'importe quel navigateur qui permet de lire et d'annoter des images puis d'exporter les annotations sous un fichier au format *JavaScript Object Notation* (json). Il s'agit d'un format de fichier texte standard ouvert pour le partage de données qui utilise du texte lisible par l'homme pour stocker et transmettre des données. Le texte renseigne pour chaque image annotée les coordonnées x et y de la Bbox ainsi que sa largeur et sa hauteur, informations essentielles pour pouvoir la replacer sur l'image.

Prenons l'exemple des images A et B de la **Figure 16** (cf. infra), les fichiers obtenus sont visibles en **Figure 15**. Pour l'image A, nommée « 14.jpeg », une seule métastase est présente d'où la présence d'une seule région « region\_attributes » avec pour coordonnées le 788<sup>ème</sup> pixel sur l'axe des abscisses (soit « x ») et le 332<sup>ème</sup> pixel sur l'axe des ordonnées (soit « y »), une largeur (width) de 130 pixels et une hauteur (height) de 120 pixels. Pour l'image B (ou « 342.jpeg »). Il y a cette fois-ci 15 Bboxes : on observe alors l'enchaînement de 15 « region\_attributes » avec chacune leurs propres coordonnées.

```
A masse}}, "via_img_metadata":{"14.jpg358209":{"filename":"14.jpg","size":358209,"regions":
[{"shape_attributes":{"name":"rect","x":788,"y":332,"width":130,"height":120},"region_attributes":
masses}}, "via_img_metadata":{"342.jpg355607":{"filename":"342.jpg","size":355607,"regions":
[{"shape_attributes":{"name":"rect","x":514,"y":177,"width":109,"height":108},"region_attributes":
{}}, {"shape_attributes":{"name":"rect","x":269,"y":611,"width":73,"height":73},"region_attributes":
{}}, {"shape_attributes":{"name":"rect","x":252,"y":558,"width":81,"height":75},"region_attributes":
{}}, {"shape_attributes":{"name":"rect","x":187,"y":613,"width":63,"height":73},"region_attributes":
{}}, {"shape_attributes":{"name":"rect","x":722,"y":562,"width":46,"height":54},"region_attributes":
{}}, {"shape_attributes":{"name":"rect","x":732,"y":612,"width":72,"height":70},"region_attributes":
{}}, {"shape_attributes":{"name":"rect","x":618,"y":663,"width":81,"height":87},"region_attributes":
{}}, {"shape_attributes":{"name":"rect","x":566,"y":626,"width":57,"height":58},"region_attributes":
{}}, {"shape_attributes":{"name":"rect","x":571,"y":447,"width":88,"height":79},"region_attributes":
{}}, {"shape_attributes":{"name":"rect","x":322,"y":469,"width":52,"height":50},"region_attributes":
{}}, {"shape_attributes":{"name":"rect","x":388,"y":598,"width":68,"height":75},"region_attributes":
{}}, {"shape_attributes":{"name":"rect","x":619,"y":529,"width":55,"height":45},"region_attributes":
{}}, {"shape_attributes":{"name":"rect","x":646,"y":440,"width":56,"height":57},"region_attributes":
{}}, {"shape_attributes":{"name":"rect","x":703,"y":704,"width":49,"height":46},"region_attributes":
{}}, {"shape_attributes":{"name":"rect","x":587,"y":559,"width":45,"height":33},"region_attributes":
```

Figure 15. Fichier json des images A et B exporter depuis VIA.

## ii. Python

Python est un langage de programmation largement utilisé dans le ML. Il sera utilisé pour développer notre algorithme de DL. Ce langage a été choisi car il s'agit de celui utilisé par l'entreprise PicoxIA pour coder ses algorithmes de *Deep Learning*. Aussi, ce langage informatique a déjà été rencontré dans le parcours scolaire de l'auteur de cette étude.

L'outil *pip* sera également utilisé pour installer et gérer les bibliothèques pythons nécessaires à notre IA. Cet outil de commande sert à développer un environnement virtuel *pipenv*, permettant d'isoler les dépendances spécifiques à un projet, évitant ainsi les conflits de versions.

Grâce à *pip*, nous pouvons également télécharger *Pytorch*, une bibliothèque logicielle permettant d'effectuer les calculs tensoriels nécessaires pour l'apprentissage profond.

*Pytorch* a été développé par Meta (anciennement connu sous le nom Facebook) et est disponible en open source.

Une autre bibliothèque d'intérêt est la bibliothèque *pillow* ou PIL spécialisée dans le traitement d'image.

L'algorithme de *Deep Learning* qui sera détaillé par la suite a été rédigé par des développeurs de PicoxIA suivant le protocole instauré par l'auteur de cette étude.

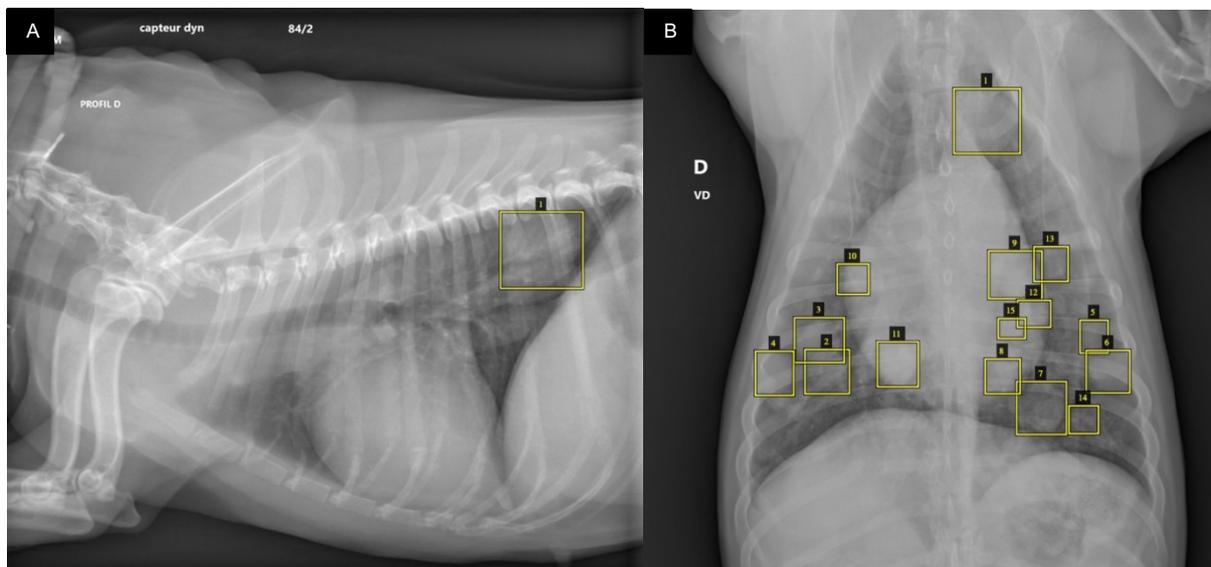
### b) Annotations manuelles des radiographies en vue de la localisation automatisée des métastases thoraciques

Afin de programmer un algorithme de DL capable de localiser les métastases sur une radiographie thoracique, il faut préalablement lui fournir une banque de données pré-annotées afin de l'entraîner. Pour cela, une Bbox a été dessinée manuellement par l'auteur autour de chaque métastase.

Pour les 337 clichés issus de la banque de données d'Oniris VetAgroBio, les annotations se sont appuyées sur les comptes rendus d'imagerie vérifiés par un vétérinaire radiologue. Chaque compte rendu (CR) relate le nombre, la localisation et la taille des métastases.

Concernant les 374 radiographies issues du jeu de données de PicoxIA, elles n'étaient pas accompagnées de CR permettant d'annoter les images. Les métastases ont donc été arbitrairement localisées et annotées par l'auteur. Ainsi, l'algorithme de détection sera testé distinctement sur les deux jeux de données issus des deux sources différentes (PicoxIA et Oniris) afin de comparer les performances obtenues (*cf. Partie 3- V*), en supposant que l'expérience de l'annotateur peut exercer une influence sur la qualité des annotations, et donc sur les performances du modèle.

La **Figure 16** illustre les *Bounding Boxes* obtenues par l'auteur via le logiciel d'annotation (*cf. Partie 3- I) b) i.*)



**Figure 16.** Exemples de dessin manuel de Bounding Boxes. Images de la banque de données issue du service d'imagerie de l'école vétérinaire Oniris VetAgroBio.

### c) Développement d'un algorithme de détection de nodules thoraciques

L'IA est alimentée par un programme python nommé « predict » composé de 4 fonctions détaillées ci-dessous.

#### i. Suppression Non Maximale

```
def apply_nms(predictions, iou_threshold=0.5, score_threshold=0.5):
    keep = predictions["scores"] > score_threshold
    predictions = {
        "boxes": predictions["boxes"][keep],
        "scores": predictions["scores"][keep],
        "labels": predictions["labels"][keep],
    }

    # Appliquer NMS pour éliminer les boîtes redondantes
    nms_keep_indices = nms(predictions["boxes"], predictions["scores"], iou_threshold)
    predictions = {
        "boxes": predictions["boxes"][nms_keep_indices],
        "scores": predictions["scores"][nms_keep_indices],
        "labels": predictions["labels"][nms_keep_indices],
    }

    return predictions
```

**Figure 17.** Code python de la fonction "apply\_nms".

La fonction « apply\_nms » (cf. **Figure 17**) applique la Suppression Non Maximale (ou *Non-Maximal Suppression*, NMS) qui a pour rôle de filtrer les prédictions effectuées par le modèle de détection. Elle permet d'éliminer les Bbox qui se chevauchent trop fortement, en conservant celles avec les plus hauts ratios de confiance.

En entrée, on retrouve 3 paramètres :

- « predictions » : une liste contenant les Bboxes et leur score de confiance. Ce dernier indique le seuil de confiance de l'IA en la boîte détectée,
- « iou\_threshold » : définit un seuil pour l'Intersection over Union (IoU), ici il est de 0,5. L'IoU est un outil comparant l'intersection entre la Bbox prédite par l'algorithme et celle annotée (Bbox « vraie » ou « vérité terrain ») à l'union de ces deux boîtes. La **Figure 18 illustre** cette métrique,

- « score\_threshold » : définit le seuil de confiance minimum pour lequel les Bboxes sont conservées, il est également fixé à la valeur 0,5.

La fonction donne en sortie la liste « predictions » actualisée.

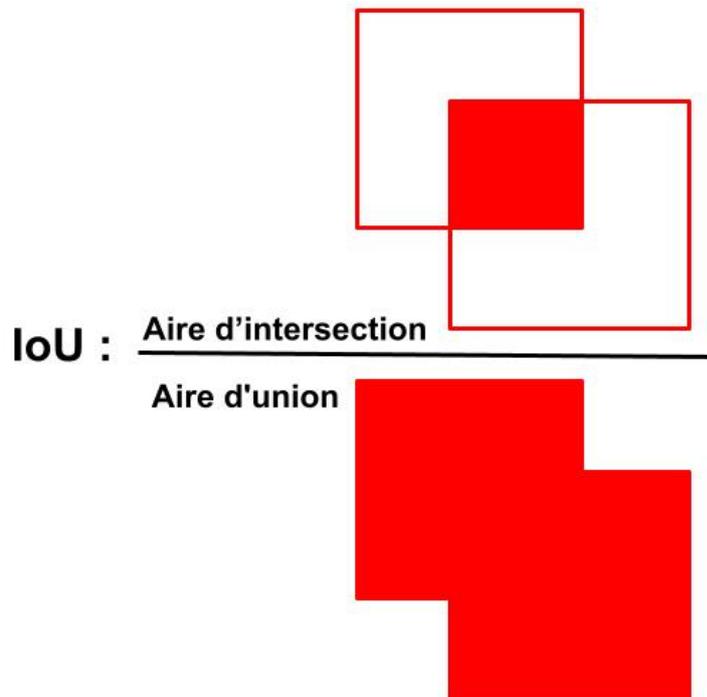


Figure 18. Illustration mathématique de l'IoU.

## ii. Annotation automatisée des images

```
def display_predictions(img, prediction):
    image = torchvision.transforms.functional.to_pil_image(img).convert("RGBA")
    draw = ImageDraw.Draw(image)

    if prediction and "boxes" in prediction:
        for index, box in enumerate(prediction["boxes"]):
            box = [int(coord) for coord in box.tolist()]
            score = prediction["scores"][index].item()
            label = prediction["labels"][index].item()

            text = f"Box{index+1} - Class: {label}, Score : {score:.2f}"
            color = colorsys.hsv_to_rgb(0.1 + 0.1 * label, 0.8, 0.8)
            color = tuple([int(c * 255) for c in color])

            draw.rectangle(box, outline=color, width=5)
            draw.text((box[0], box[1]), text, fill="black")

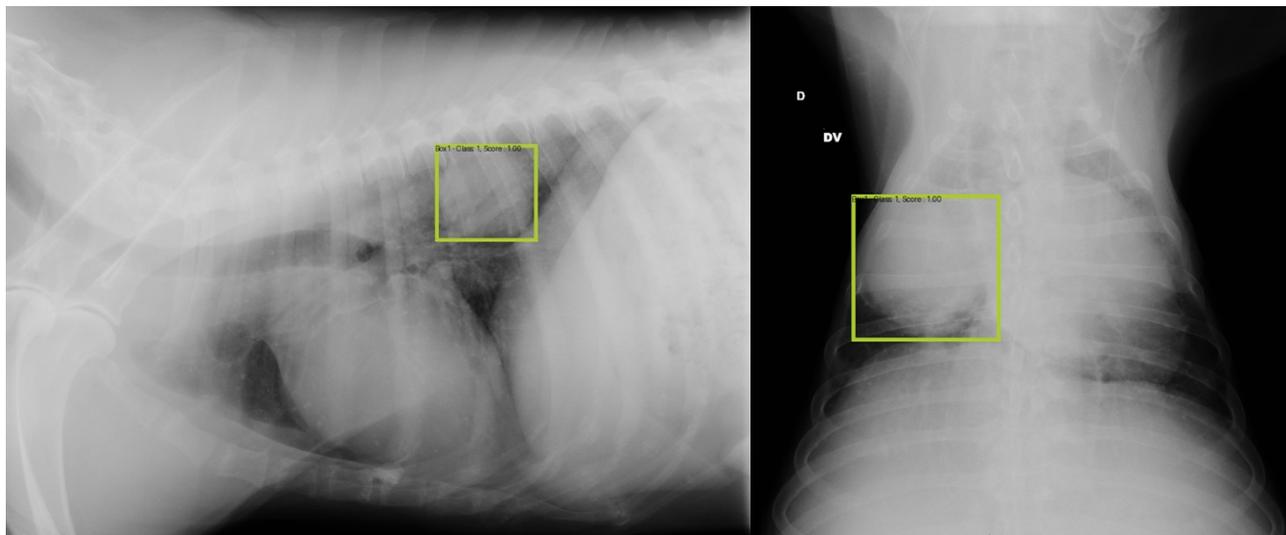
    Thread(target=lambda: ImageShow.show(image)).start()
```

Figure 19. Code python de la fonction "display\_predictions".

La fonction de la **Figure 19** a pour but d'afficher les précédentes prédictions sur les images sous forme de Bbox (entrées « predictions » et « img »). Pour ce faire, l'image est préalablement convertie au format « RGBA » afin que les Bboxes puissent y être dessinées par l'algorithme. Chaque boîte dessinée est

accompagnée de son « label » (sans importance ici car toutes les métastases sont labélisées sous le même nom « Class 1 ») et du score de confiance de l'IA.

En sortie la fonction affiche l'image avec la ou les Bboxes annotée(s) (cf. **Figure 20**).



**Figure 20.** Exemple d'images annotées en sortie de la fonction "display\_predictions".

### iii. Récupération des coordonnées des Bboxes

```
def predictions_to_via_regions(prediction):
    regions = []
    for box, label in zip(prediction["boxes"].tolist(), prediction["labels"].tolist()):
        regions.append(
            {
                "shape_attributes": {
                    "name": "rect",
                    "x": int(box[0]),
                    "y": int(box[1]),
                    "width": int(box[2] - box[0]),
                    "height": int(box[3] - box[1]),
                },
                "region_attributes": {"class": label},
            }
        )
    return regions
```

**Figure 21.** Code python de la fonction "predictions\_to\_via\_regions".

La fonction « predictions\_to\_via\_regions » (cf. **Figure 21**) convertit les prédictions en entrée, en un format compatible avec l'outil d'annotation VIA (cf. **Partie 3- I) b) i.**). On retrouve alors les coordonnées x, y, la largeur et la hauteur de la boîte annotée par l'algorithme.

### iv. Fonctions d'exécution du programme

On retrouve enfin la fonction « main » qui prend en entrée :

- « model\_path » : chemin vers le modèle à utiliser pour les prédictions (Faster R-CNN, cf. **Partie 3- II**)

- « imgs\_dir » : chemin vers le dossier contenant les images utilisées pour ces prédictions
- « output\_file » : chemin vers le fichier json de sortie de la fonction « predictions\_to\_via\_regions »

Cette fonction va ensuite charger le modèle, appliquer les fonctions sur les images du dossier indiqué, afficher les images avec les prédictions du modèle et générer le fichier json d'annotations.

Le bloc « if \_\_name\_\_ == "\_\_main\_\_" » permet à son tour d'exécuter le script lorsqu'il est lancé depuis la ligne de commande « python -m pipenv run python ./predict.py ./models/fasterRNCn\_resnet50\_fpn-masse\_thorax\_v1.ckpt ».

L'intégralité du code python est disponible en **Annexe I**.

#### IV) Paramètres nécessaires à l'exécution de l'algorithme de Deep Learning et à l'évaluation de ses performances.

Le modèle développé par les ingénieurs de PicoxIA et utilisé pour l'apprentissage est un *Faster Region-based Convolutional Neural Network* (Faster R-CNN), de 50 couches cachées.

Pour chaque phase d'entraînement et de validation de l'algorithme, le lot de données sera réparti, comme dans la plupart des études, en 2 : 80% des images seront utilisées pour la phase d'apprentissage et 20% pour la validation de l'algorithme.

##### a) Choix des hyperparamètres

Comme explicité dans la **Partie 2- I) D) ii.**, les hyperparamètres sont des valeurs fixées par l'opérateur pour paramétrer le modèle. Pour notre algorithme, les paramètres suivants ont été fixés :

- *Batch size* : 1. Cette valeur a été choisie afin d'optimiser l'utilisation de la mémoire et permettre une mise à jour plus fréquente des poids (cf. **Partie 2- I) b**),
- *Epochs* : 21. Ce nombre semble être le meilleur compromis entre performance et temps d'entraînement lors de l'étape de validation croisée. Cette méthode permet de tester les performances du modèle de DL, elle est assimilable à une procédure de rééchantillonnage,
- *Learning rate* : initié à 0,01 et ajusté par un modèle de type *Super Convergence*<sup>92</sup> à un minimum de 0,005 afin, d'accélérer la convergence.

La technique de Super Convergence est détaillée en **Annexe II**.

Les hyperparamètres ont été optimisés grâce à une méthode de Descente de gradient stochastique (*Stochastic Gradient Descent*, SGD). Cette dernière est un processus d'optimisation itératif : à chaque itération, un échantillon de données est sélectionné aléatoirement pour calculer le gradient et mettre à jour les paramètres du réseau<sup>93</sup>.

b) Métriques utilisées pour l'évaluation des performances du modèle de détection.

i. *Faux positifs, Faux négatifs, Vrais positifs, et rappel*

Rappelons la définition de ces indicateurs, adapté à notre algorithme de détection :

- Faux positif (FP) : comptage absolu des boîtes annotées par l'algorithme mais ne correspondant pas à une lésion sur l'image. C'est-à-dire que le modèle a entouré sur la radiographie un élément qui n'était pas une métastase.
- Faux négatifs (FN) : comptage absolu des boîtes manquées par l'algorithme. Dans ce cas, l'algorithme n'a pas entouré une métastase présente sur le cliché, alors que l'auteur l'avait annoté.
- Vrais positifs (VP) : comptage absolu des boîtes annotées par l'algorithme et correspondant bien à une lésion sur la radiographie. C'est donc le succès de détection d'une masse par le modèle.
- Rappel : score de performance, compris entre 0 et 1, indiquant la capacité du modèle à identifier toutes les instances réelles, c'est-à-dire la proportion de VP détectés par rapport à tous les cas réels. Le rappel se rapproche de ce que l'on nomme la sensibilité en statistique.

On remarquera l'absence des Vrais Négatifs (VN). En effet, en détection d'objets sur une image, Les VN ne sont pas comptabilisés explicitement car cela reviendrait à compter tous les « non-objets », c'est-à-dire les « non-masses » de l'image. Or, une radiographie peut contenir un nombre infini de « non-masses » rendant le décompte des VN impraticable et de ce fait non pertinent. De plus, comme aucune radiographie ne contenant pas de métastases n'a été intégrée dans l'échantillon d'étude, l'absence des VN dans la métrique utilisée est attendue. La spécificité (Sp) ne pourra donc pas être utilisée sous sa forme métrique pour évaluer les performances de l'algorithme. Le terme pourra cependant être employé pour interpréter les résultats obtenus mais n'aura pas de valeur statistique.

La valeur du rappel est obtenue grâce à la formule suivante :

$$\text{Rappel} = \frac{\text{Nombre de boîtes correctement annotées par l'algorithme}}{\text{Nombre total de boîtes annotées manuellement}}$$

Soit,

$$\text{Rappel} = \frac{VP}{VP + FN}$$

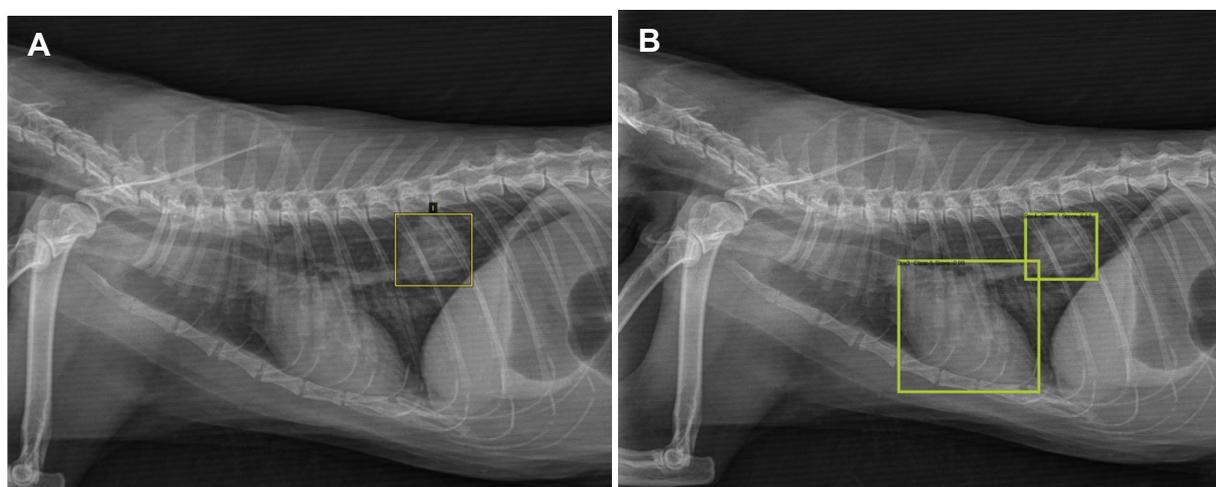
Est considéré comme VP toute boîte annotée avec un IoU > 0,5 et un seuil de confiance de détection > 0,5 (cf. **Partie 3- I) c) i.**)

On comprend donc que sur une seule image radiographique il est possible d'obtenir les valeurs de FP, FN et VP en même temps. Prenons l'exemple d'une radiographie annotée par l'auteur et présentant 2 *Bounding boxes* entourant chacune une métastase. Si le modèle de détection donne les prédictions suivantes :

- 1 boîte englobante correctement annotée, soit 1 VP,
- 1 boîte englobante manquée par l'algorithme, soit 1 FN,
- 1 boîte marquée par le modèle mais non présente en réalité sur le cliché, soit 1 FP.

La seule image ainsi traitée par l'algorithme présentera alors 1 VP, 1 FN et 1 FP. Lors de l'évaluation des performances de l'algorithme, toutes les valeurs de VP, FN et FP obtenues sont comptabilisées pour chaque image de la phase de validation. Les statistiques obtenues sont donc calculées à l'échelle du nodule, et non pas à l'échelle de l'animal. Prenons par exemple le cas d'un test de validation du modèle de détection avec une banque d'image de validation constitué d'une seule radiographie : l'image A de la **Figure 22**. Une seule masse a été entourée par l'auteur (vérité à atteindre par l'algorithme). L'image B de cette même figure montre l'image affichée par le modèle de détection avec les boîtes qu'il a annoté automatiquement. On peut voir un vrai positif (la masse) et un faux positif (l'algorithme a faussement entouré le cœur) et aucun faux négatif (il n'y avait pas d'autre masse à entourer). Les métriques de l'algorithme seront donc les suivantes :

- 1 VP
- 1 FP
- 0 FN
- Rappel = 1
- Précision = 0,5 (cf. infra)
- Score F1 = 0,67 (cf. infra)



**Figure 22.** Exemple d'une radiographie utilisée pour calculer les performances de l'algorithme. Image **A** : image annotée par l'auteur. Image **B** : image annotée par l'algorithme de détection

L'algorithme est donc bien évalué sur sa capacité à détecter correctement les nodules, d'où les métriques obtenues dans cet exemple. Il ne permet malheureusement pas d'informer l'auteur sur le résultat du bilan d'extension à l'échelle de l'animal.

## ii. Précision et moyenne de la précision moyenne

La précision mesure la proportion des détections correctes de l'IA par rapport à l'ensemble des détections effectuées par l'IA (qu'elles soient justes ou non), indiquant ainsi la qualité des prédictions positives. La valeur obtenue est comprise entre 0 et 1. Elle est obtenue comme suit :

$$\text{Précision} = \frac{\text{Nombre de boîtes correctement annotées par l'algorithme}}{\text{Nombre total de boites annotées par l'algorithme}}$$

Soit,

$$Précision = \frac{VP}{VP + FP}$$

En statistique, la précision est connue sous le nom de Valeur Prédicative positive.

La moyenne de la précision moyenne, ou *mean Average precision* (mAP) est une métrique très utilisée dans les algorithmes de détection et évaluant la précision à différents seuils de rappel, permettant d'offrir une vue globale de la performance du modèle sur toutes les classes.

Dans un premier temps, la précision moyenne est calculée comme la moyenne pondérée des précisions pour différents seuils de rappel (c'est-à-dire pour différent seuil d'IoU in fine) pour chaque classe. La mAP est en fait la moyenne des précisions moyennes de chaque classe.

### iii. Score F1

Ce score F1 est une moyenne harmonique de la précision et du rappel. Cette métrique est un type particulier de moyenne utilisée lorsque l'on souhaite calculer des moyennes de taux ou de ration. Cette métrique offre un équilibre entre ces 2 indicateurs et reflète à la fois la précision de l'algorithme et sa capacité à détecter toutes les instances pertinentes. C'est un bon indicateur de la performance du modèle, un score F1 proche de 1 indiquant une détection efficace par l'algorithme tout en minimisant les FP et les FN. Il est calculé selon la formule suivante :

$$Score\ F1 = 2x \frac{Précision \times Rappel}{Précision + rappel}$$

## V) Résultats obtenus à l'issue des différents apprentissages

### a) Sur la base de données de 711 images

Un premier essai du couple entraînement-validation a été réalisé en utilisant l'intégralité des 711 clichés radiographiques annotés. Les résultats obtenus sont rassemblés dans le **Tableau 7**. La validation a été effectuée sur 20% des images soit 142 clichés. Pour ces 142 images constituant l'échantillon de validation, 379 boîtes ont été annotées par l'auteur (correspondant à la somme des FN et des VP).

*Tableau 7. Performances métriques du modèle sur 711 clichés.*

FN	FP	VP	Rappel	Précision	mAP	mAP <sub>0,5</sub>	mAP <sub>0,75</sub>	Score F1
223	142	156	0,412	0,523	0,217	0,344	0,272	0,461

On remarque les éléments suivants :

- Un nombre très important de FN, montrant que le modèle rate un nombre important de nodules,

- Un nombre de FP sensiblement similaire au VP, indiquant que le modèle identifie correctement une partie des métastases mais que sa précision pourrait être améliorée. Il est possible que l'IA confonde certains éléments avec des nodules. En effet, la mAP est faible, confirmant une précision moyenne mauvaise,
- Un score F1 inférieur à 0,5, montrant que le modèle n'est ni très précis, ni très complet dans sa détection.

Globalement, ces résultats montrent que le modèle présente des lacunes significatives dans ses performances de détection, avec une performance globale faible. En effet, le modèle rate de nombreux nodules et bien que certaines de ses prédictions soient justes, presque autant sont incorrectes. Il n'est donc ni sensible, ni spécifique.

b) Sur 7 lots distincts d'images issues de la banque de clichés

Concernant la base de données composées pourtant de 711 images, les performances du modèle sont étonnamment faibles. Il a donc été décidé d'entraîner à nouveau le modèle sur 5 lots distincts d'images. Ainsi 20% des images constituant chaque lot ont été allouées à la phase de validation, les 80 autres pourcents étant utilisés pour la phase d'entraînement. Le nombre d'images de la phase de validation pour chaque lot était le suivant :

- Lot 1 : 30 images avec un total de 48 boîtes annotées par l'auteur,
- Lot 2 : 24 images avec un total de 76 boîtes annotées par l'auteur,
- Lot 3 : 20 images avec un total de 42 boîtes annotées par l'auteur,
- Lot 4 : 19 images avec un total de 53 boîtes annotées par l'auteur,
- Lot 5 : 50 images avec un total de 160 boîtes annotées par l'auteur.

Les performances de l'algorithme pour chaque lot sont regroupées dans le **Tableau 8**.

*Tableau 8. Résultats des performances du modèle pour chaque lot d'images.*

	FN	FP	VP	Rappel	Précision	mAP	mAP <sub>0,5</sub>	mAP <sub>0,75</sub>	Score F1
<b>Lot 1</b>	30	17	18	0,375	0,514	0,219	0,309	0,271	0,434
<b>Lot 2</b>	43	19	33	0,434	0,635	0,229	0,354	0,288	0,516
<b>Lot 3</b>	27	15	15	0,357	0,5	0,217	0,322	0,236	0,417
<b>Lot 4</b>	40	30	13	0,245	0,302	0,091	0,153	0,103	0,271
<b>Lot 5</b>	114	48	46	0,287	0,489	0,133	0,207	0,144	0,362

2 lots se distinguent particulièrement des autres :

• Le lot 4

Ce lot présente tout d'abord un nombre de FN et de FP nettement plus grand que le nombre de VP par rapport à la plupart des autres lots. On note également une mAP extrêmement basse à 0,091, bien en dessous des autres échantillons. Les valeurs de précision et de rappel sont également basses, impliquant un score F1 médiocre constituant la plus petite valeur de cette métrique parmi tous les lots.

### • Le lot 5

Toute proportion gardée, le lot 5 présente le nombre de FN le plus élevé, ce qui affecte la valeur de rappel (0,287) presque toute aussi basse que pour le lot 4.

Comme pour le lot précédent, mais dans une moindre mesure, la valeur de la mAP est également faible, avec des valeurs de  $mAP_{0,5}$  et de  $mAP_{0,75}$  discrètement plus élevées que pour le lot 4, mais inférieurs aux 3 autres lots.

Les lots 1, 2 et 3 quant à eux présentent des valeurs métriques améliorées par rapport au deux derniers échantillons cités. Le lot 2 présente les meilleures performances globales parmi tous les lots, avec une amélioration nette du score F1 (0,516) et de la précision (0,635), et modérée du rappel.

Ces différences entre les lots suggèrent que les échantillons 4 et 5 desservent les performances du modèle de détection de métastases. Les défauts de performance de l'algorithme sur ces deux lots peuvent être expliqués soit par la moindre qualité des images transmises par PicoXIA, soit par la qualité des annotations fournies par l'auteur seul. Le lot 4 est également celui présentant la plus grande proportion d'images avec des nodules de petites tailles. La sensibilité de l'algorithme est donc plus faible pour les métastases de moins d'un centimètre.

c) Sur une base de données obtenue d'après les meilleurs lots

Au vu des précédentes observations, un dernier entraînement a été réalisé avec les lots 1 à 3 semblants offrir les meilleures performances. L'association de ces 3 échantillons constituait une banque d'images de 366 radiographies thoraciques. La validation a donc été effectuée sur 74 clichés (20% de la base de données) pour un total de 166 boîtes à détecter.

Les résultats obtenus sont récapitulés dans le **Tableau 9**.

*Tableau 9. Performances du modèle de détection sur les meilleurs lots.*

FN	FP	VP	Rappel	Précision	mAP	$mAP_{0,5}$	$mAP_{0,75}$	Score F1
85	81	81	0,488	0,5	0,263	0,396	0,315	0,494

Pour cet entraînement, on remarque tout d'abord que le nombre de FN, FP et VP est plus équilibré que dans le premier entraînement (cf. **Partie 3- II) c) i.**). Il y a donc une meilleure balance entre les détections justes et fausses.

De plus, la mAP est également discrètement plus élevée que dans le Tableau 4 (0,263 contre 0,217), et il en est de même pour la  $mAP_{0,5}$  et la  $mAP_{0,75}$ . Comme pour tous les autres essais, la mAP diminue quand le seuil d'IoU devient plus strict ( $mAP_{0,5}$  vs  $mAP_{0,75}$ ). Enfin, le score F1 est également légèrement plus élevé pour la combinaison de ces lots (0,494) comparé au score F1 de la banque de données initiale (0,46).

Les performances du modèle sur ce jeu de données sont globalement meilleures que pour les précédents tests. L'amélioration des métriques tels que le rappel, le score F1 et la mAP souligne bien que notre modèle de détection automatique des métastases pulmonaires est plus performant lorsque les annotations s'appuient sur des comptes rendus de vétérinaires spécialistes. Aussi, ces résultats confirment le défaut de détection des nodules de petites tailles.

Cependant, l'étude menée laisse entendre que ce modèle de détection de nodules thoraciques offre des performances tout juste acceptables, et mériterait de trouver des axes d'amélioration quant à la réduction du nombre de FN et l'augmentation de sa précision globale.

## VI) Discussion

### a) Performances globales de notre algorithme de détection

L'algorithme de localisation développé dans cette étude a obtenu des performances globalement similaires à l'algorithme de classification actuel de PicoxIA, avec environ 50% des nodules correctement détectés. Ces performances sont en dessous des attentes de l'entreprise. Le manque de sensibilité de l'algorithme développé peut découler des éléments suivants :

- Échantillon d'étude : La taille de l'échantillon peut être trop petite pour la tâche demandée à l'algorithme. Aussi, même en augmentant la taille de l'échantillon, il faudra porter une attention particulière à la représentation des espèces et des races dans la base de données. À défaut, les performances pourraient ne pas être généralisées à l'ensemble des formats de thorax.
- Gold standard : le choix de s'appuyer uniquement sur les comptes rendus de radiographies peut également expliquer les mauvaises performances. Qui plus est, presque la moitié des images ont eu pour gold standard les seules annotations de l'auteur, dont l'expérience est limitée.
- Le modèle de détection : l'algorithme développé est l'un des premiers à réaliser une tâche de détection de nodules thoraciques par localisation en médecine vétérinaire. De ce fait, sa structure peut ne pas être suffisamment développée pour répondre à un tel objectif.

En médecine humaine, les algorithmes de détection basés sur des approches de localisation sont nettement plus performants. L'étude de *Behrendt et al* par exemple a obtenu de très bonnes performances dans la prédiction des *Bonding Boxes* (Bboxes). Sur une base de données de validation de 175 radiographies thoraciques avec au moins un nodule, l'aire sous la courbe était de 0,94<sup>90</sup>.

A titre de comparaison, encore en médecine humaine, la base de validation issue de l'université de San Francisco dans l'étude de *Park et al* comportait 89 radiographies thoraciques avec au moins un nodule pulmonaire. Au total, 104 nodules devaient être détectés par l'IA. La « sensibilité de détection des nodules par rapport à chaque nodule » est défini dans l'article comme le rapport des FP sur l'ensemble des nodules à détecter. La valeur obtenue dans l'étude de *Park et al* est de 75%, alors qu'elle est de 49% dans notre étude (soit, pour le modèle le plus performant : 81/166). Cet exemple montre une fois de plus que les performances obtenues avec notre algorithme de détection sont décevantes comparées à ce qui aurait pu être obtenu.

Afin d'augmenter les performances globales de notre algorithme, il pourrait être intéressant de trouver des solutions permettant à l'IA d'augmenter son score de confiance sur ses prédictions.

On pourrait par exemple simplifier la lecture des radiographies en isolant les poumons. La détection d'éléments non pulmonaires, tels que les côtes et le cœur par exemple, simplifierait la lecture si un algorithme pouvait ensuite les retirer de la radiographie. Le développement d'un tel algorithme de pré-traitement d'images, basé sur des techniques de segmentation d'instance, permettrait de mettre en lumière les lobes pulmonaires et de simplifier la tâche au modèle de détection. Ce travail complémentaire pourrait également limiter le nombre de FP en retirant du cliché une partie des éléments « mass-like ».

Aussi, il serait intéressant d'associer la lecture des différentes incidences (DLD, DLG ou DV/VD). Dans cette étude, il a été fait le choix d'annoter les radiographies de différentes incidences indépendamment les unes des autres. Une amélioration du modèle pourrait consister à donner à l'algorithme de détection un ensemble d'incidences,

pour un même cas, afin de renforcer sa confiance sur la localisation de la masse (augmentation du nombre de VP). Cette modification apportée à l'algorithme pourrait aussi lui permettre de diminuer le nombre de boîtes faussement annotées (FP) et de limiter les omissions (FN).

Cette dernière amélioration aurait pu être réalisée dès le développement de notre algorithme mais aurait à nouveau réduit la base de données totale. En effet, comme les radiographies fournies par PicoxIA n'étaient accompagnées d'aucune information concernant les patients, il aurait été impossible d'associer une radiographie en décubitus latéral droit avec la bonne radiographie en incidence ventro-dorsale.

## b) Détection des nodules de petites tailles

Notre algorithme de *Deep Learning* a notamment rencontré des difficultés à détecter les nodules de petites tailles (inférieurs à 1 centimètre de diamètre). En témoignent les performances du lot 4 (lot formé de 95 radiographies thoraciques de chiens et de chats issues du jeu de données de PicoxIA), lot présentant la plus grande proportion de nodules de petites tailles (16%, contre pas plus de 9% dans les autres lots) par rapport aux autres tailles. Nous rappelons que ce lot est celui pour lequel l'algorithme a obtenu les performances les plus médiocres.

Cependant, en termes de santé et de pronostic, il est tout aussi grave de manquer les petites métastases que les grandes. La détection précoce des métastases, notamment celles de petites tailles ou isolées, pourrait offrir des possibilités thérapeutiques et permettrait d'établir un pronostic plus précis. Un diagnostic précoce peut non seulement améliorer la qualité de vie de l'animal, mais aussi permettre de mieux préparer le propriétaire aux possibilités de traitement et à l'évolution de la maladie.

Ces résultats sont en accord avec les travaux de *Liang et al.* Dans cette étude, la réussite ou l'échec de détection des nodules des différents algorithmes développés (*Heat map*, *Abnormal probability*, *Nodule probability* et *Mass probability*) sont dépendants de la taille des nodules. Les résultats de cet étude montrent que les performances de détection baissent dès que la taille des nodules se retrouvent en dessous de 3 centimètres<sup>94</sup>.

Afin d'augmenter la sensibilité de notre algorithme pour la détection des nodules de petites tailles, il serait intéressant d'augmenter le nombre d'images dans la base de données, et notamment d'augmenter le nombre de radiographies présentant des petits nodules. Il pourrait s'avérer utile d'élargir la recherche de cas à plusieurs institutions, comme l'une des trois autres Écoles Vétérinaires de France, ou bien d'intégrer au projet d'autres cliniques vétérinaires sur le territoire français, afin de récolter leurs clichés radiographiques.

Au vu du travail fastidieux qu'une telle démarche représenterait, une autre solution pourrait être envisagée : La *Data Augmentation*<sup>95</sup>. Cette technique informatique a pour but d'augmenter artificiellement le nombre de données disponibles en effectuant des modifications plus ou moins importantes sur des données préexistantes. On retrouve par exemple l'application de filtres pour modifier le format RVB (sépia, noir et blanc...), modification du contraste, de l'exposition, rotation de l'image ou zoom sur l'image. Une technique aurait pu s'avérer intéressante pour notre étude : le *swapping*. Son principe est de déplacer une région de l'image sur une autre zone de l'image (par exemple dans notre cas, il s'agirait de déplacer une métastase d'un lobe vers un autre).

### c) Comparaison des performances de l'IA sur les deux jeux de données

Notre algorithme a été testé sur deux jeux de données, l'un constitué d'images récupérées auprès du service d'imagerie de l'école vétérinaire Oniris VetAgroBio et l'autre formé des images fournies par l'entreprise PicoxIA.

On remarque que les performances du modèle sur la base de données d'Oniris (lots 1 et 3) sont relativement semblables d'un lot à l'autre. Les échantillons 1 et 3 font également partis de ceux ayant obtenus les meilleures métriques. Cette stabilité dans les résultats peut s'expliquer des manières suivantes :

- Les clichés radiographiques ont tous été réalisés dans un service spécialisé en imagerie. L'origine des radiographies est gage de qualité et permet une standardisation de l'ensemble des images constituant la banque de données.
- Les annotations manuelles de l'auteur se sont basées sur des comptes rendus rédigés par des vétérinaires spécialistes en imagerie. La qualité des annotations obtenues offre une plus grande confiance dans la validité interne de cette étude.

Les résultats pour le jeu de données de PicoxIA sont plus hétérogènes, avec des performances pouvant être soit excellentes (lot 2) soit totalement médiocres (lots 4 et 5). Contrairement au jeu de données précédemment décrit, les images fournies par PicoxIA ne répondaient pas forcément aux standards de qualité obtenus auprès des spécialistes. Une telle diversité peut rendre la tâche plus difficile à l'algorithme et risque de ne pas toujours constituer un échantillon représentatif lorsque les radiographies sont de bonnes qualités.

De plus, aucun support n'a été fourni à l'auteur pour l'annotation de ses images, rendant critiquable la vérité terrain pour ces 374 radiographies, et démontrant également un effet de l'expérience du notateur.

Dans la littérature, une importance est accordée à la qualité des radiographies. La position de l'animal, les constantes choisies (mAS et kV), les incidences réalisées sont autant de détails qui peuvent influencer la qualité de la radiographie. Le non-respect des standards recommandés peut alors conduire à l'obtention d'un cliché de qualité médiocre. Ce défaut de qualité radiographique entraîne une mauvaise interprétation de nombreuses lésions<sup>96</sup>.

Afin d'homogénéiser les images constituant la banque de données tout en respectant la réalité terrain, il serait intéressant d'ajouter au modèle un algorithme de prétraitement des images capable de trier les radiographies en fonction de leur qualité, comme le propose *Banzato et al* dans leur étude<sup>97</sup>. Les différentes anomalies techniques pourraient également être rectifiées afin de limiter le nombre de clichés perdus.

Comme souligné plus haut, les annotations des images du jeu de données fourni par PicoxIA réalisées pour cette étude n'ont été faites que par un seul opérateur, l'auteur. Dans ce contexte, une importance pourrait être accordée à la standardisation des annotations constituant la vérité terrain.

Aussi, il serait également pertinent de donner à l'algorithme la possibilité de traiter le texte des comptes rendus (CR) afin d'obtenir des informations supplémentaires sur les différents cas, et notamment l'anamnèse qui a conduit à la réalisation de la radiographie thoracique. Aussi, l'utilisation d'une telle technologie dans l'analyse des résultats d'examen complémentaires venant confirmer la présence de métastases pourrait s'avérer intéressante. L'analyse de ces données demanderait, certes, plus de travail mais permettrait à l'IA d'augmenter son score de confiance sur ses prédictions mais aussi d'homogénéiser les performances sur les différents jeux de données. Ainsi, l'algorithme de détection pourrait voir ses performances globales améliorées (augmentation du

nombre de VP).

Afin de limiter les défauts de performances liés à l'expérience de l'annotateur, il pourrait être intéressant :

- Soit de n'utiliser dans la banque d'images que des radiographies accompagnées de CR rédigés par des vétérinaires spécialistes, afin de s'affranchir du biais lié à l'expérience.
- Soit de reprendre notre base de données de 711 radiographies et de faire annoter les clichés par un ou plusieurs vétérinaire(s) spécialiste(s) afin de comparer les performances obtenues avec celles de notre étude.

d) Choix des métriques utilisées pour évaluer les performances de notre modèle

La présente étude évalue les performances d'un algorithme de détection de nodules thoraciques chez le chien et le chat.

Les valeurs de faux positifs (FP), faux négatifs (FN) et vrais positifs (VP) sont comptabilisées pour chaque image de la phase de validation. Sur chaque radiographie thoracique il y a donc une quantité finie de VP (métastases annotées manuellement par l'auteur) et un nombre infini de vrais négatifs (tous les endroits du poumon où il n'y a pas de métastases). Il y a donc bien des zones dites « saines » et des zones avec des nodules sur un même cliché. Comme aucune radiographie d'animaux ne présentant pas de métastases n'a été intégrée dans cette étude, les FP sont comptés comme tout élément non métastatique confondu par l'algorithme. Il est donc important de souligner que la véritable valeur de FP donnée par l'algorithme est probablement mal appréciée. Afin de préciser les performances du modèle, il pourrait être intéressant d'ajouter à la banque de données un échantillon d'animaux « sains ». Cette amélioration permettrait également d'évaluer numériquement la spécificité de notre algorithme.

Dans ce contexte, il faudrait également modifier la définition des VP, FP, FN et VN. Au lieu de se situer à l'échelle d'une métastase, comme dans cette étude, il faudrait s'intéresser à l'échelle de l'animal.

Selon la même réflexion, nous avons vu que les métriques étaient comptabilisées à l'échelle d'un nodule, et non pas à l'échelle du patient. Les performances évaluées se concentrent donc uniquement sur la capacité de l'algorithme à détecter les métastases. Elles ne permettent en aucun cas d'évaluer la capacité de l'algorithme à prédire si le bilan d'extension de l'animal faisant l'objet de la radiographie est positif ou négatif. En vue de proposer une telle technologie aux vétérinaires, la perspective d'utilisation d'un tel algorithme nécessite d'intégrer des notions éthiques.

Il est important de souligner que, dans la pratique clinique, le fait d'avoir un seul VP sur une image contenant des métastases et de nombreux faux positifs FP n'est pas un problème en soi, tant que le diagnostic final est correct. Si l'algorithme identifie correctement une métastase, la présence de nombreux FP devient secondaire, car le patient sera pris en charge sur la base de la détection du VP.

Il convient tout de même de rappeler que la détection de nombreux FP pourrait dans certains cas entraîner des examens complémentaires inutiles, générant du stress pour le propriétaire et pour l'animal, ainsi que des coûts supplémentaires.

Un autre aspect essentiel est le choix de privilégier la sensibilité ou la spécificité de l'algorithme, en fonction de l'animal, du client, des considérations éthiques et des objectifs du diagnostic. Par exemple, dans certains cas, une sensibilité accrue pourrait être préférée pour minimiser les risques de passer à côté d'une métastase. Dans d'autres contextes, une spécificité plus élevée serait privilégiée pour éviter des traitements inutiles. Cette décision doit être adaptée à chaque situation, afin de garantir une prise en charge respectueuse et adaptée à chaque animal ainsi qu'à ses propriétaires.

#### e) Perspectives immédiates de PicoXIA

Les ingénieurs de PicoXIA travaillent encore aujourd'hui sur ce projet.

Un lot de 450 radiographies supplémentaires a été mis à leur disposition afin d'augmenter la base de données. Ils n'ont à ce jour pas communiqué qui sera l'opérateur responsable des annotations des clichés.

Au vu des résultats obtenus, il a été décidé d'utiliser des techniques d'augmentation de données (*Data Augmentation*) pour augmenter la proportion de nodules de tailles inférieures à 1 centimètre. La technique envisagée est une méthode d'insertion d'images permettant de déplacer un ensemble de pixels, ici un nodule, vers une autre localisation dans l'image.

Grâce à l'augmentation de données, il est également envisagé d'homogénéiser le format des thorax (dépend de l'espèce et de la race) afin de mieux généraliser les performances de l'algorithme.

L'étude que l'entreprise a réalisé avec l'auteur n'est que le début d'un processus visant à commercialiser l'algorithme. La mise sur le marché d'un tel outil nécessitera, comme il a été mentionné dans cette discussion, d'intégrer une démarche clinique répondant aux critères scientifiques et éthiques de la profession vétérinaire.

# Conclusion

L'utilisation de la radiographie dans le cadre du bilan d'extension pulmonaire est l'examen le plus utilisé en médecine vétérinaire . C'est une technique d'imagerie peu coûteuse, disponible dans la quasi-totalité des cliniques et pour laquelle l'interprétation des clichés est à la portée de tout généraliste. Afin d'optimiser la détection des métastases, et limiter les omissions, les vétérinaires se tournent vers des outils informatiques intelligents capables de les accompagner dans leur diagnostic. Cependant, les performances des premiers algorithmes de classification développés dans ce but sont décevantes à l'heure actuelle.

C'est pourquoi cette étude s'est focalisée sur le développement d'un nouveau genre d'algorithme de détection par *Deep Learning* en s'appuyant cette fois-ci non plus sur la classification, mais sur la localisation de ces métastases, à l'instar des technologies utilisées en médecine humaine.

Plusieurs entraînements ont été réalisés dans ce but, et l'évaluation des métriques de performance de notre modèle a permis de d'identifier parmi plusieurs un programme aux performances améliorées. Toutefois, ce modèle de détection est tout juste aussi performant que les précédents algorithmes de classification radiographique utilisés en médecine vétérinaire.

Notre étude se situant aux prémices de la révolution numérique en médecine des animaux de compagnie, l'avenir d'une telle technologie laisse présager de bonnes perspectives d'amélioration.

Entre autres, l'association de notre modèle à des algorithmes de pré-traitement des clichés radiographiques ou à des algorithmes de traitement de texte permettrait d'obtenir de meilleures performances globales quant à la détection des nodules pulmonaires.

Aussi, les résultats obtenus dans cette étude soulignent l'importance à accorder aux images collectées, qui doivent répondre à un certain standard de qualité à défaut de quoi les performances de l'algorithme développé peuvent être impactées négativement. Le développement de *Digital Imaging and Communication in Medicine* (DICOM) open source permettrait de mettre en commun des images radiographiques entre plusieurs centres référés vétérinaires mais aussi de partager les comptes rendus ou annotations associés. L'utilisation de ces outils en médecine vétérinaire permettrait d'élargir la base de données disponibles aux recherches en science du numérique tout en garantissant une certaine qualité.

Un tel outil informatique au service des vétérinaires promet une meilleure confiance en son diagnostic offrant ainsi une prise en charge plus juste et plus précoce des patients.



# Bibliographie

1. Pet Owners Ressources, 2023, *Veterinary Cancer Society* [en ligne], 2023. [Consulté le 20 septembre 2023]. Disponible à l'adresse: <https://vetcancersociety.org/resources/pet-owners/pet-owner-resources/>
2. Villalobos, A. & Hershey, B. Cancers in Dogs and Cats. in *Hospice and Palliative Care for Companion Animals* (eds. Shanan, A., Pierce, J. & Shearer, T.) 121–137 (Wiley, 2023).
3. Dobson, J. M. & Lascelles, B. D. X. *BSAVA Manual of Canine and Feline Oncology, Third Edition*. (2011).
4. Argyle, D. J., Brearley, M. J. & Turek, M. *Decision Making in Small Animal Oncology*. (Blackwell Pub. Professional, Ames, Iowa, 2008).
5. Fidler, I. J. The pathogenesis of cancer metastasis: the 'seed and soil' hypothesis revisited. *Nat Rev Cancer* **3**, 453–458 (2003).
6. Vail, D. M. *Withrow and MacEwen's Small Animal Clinical Oncology*. (Elsevier, Edinburgh ; New York, 2020).
7. Litjens, G. *et al.* A survey on deep learning in medical image analysis. *Medical Image Analysis* **42**, 60–88 (2017).
8. Health, C. for D. and R. Artificial Intelligence and Machine Learning (AI/ML)-Enabled Medical Devices. *FDA* (2023).
9. Adrien-Maxence, H. *et al.* Comparison of error rates between four pretrained DenseNet convolutional neural network models and 13 board-certified veterinary radiologists when evaluating 15 labels of canine thoracic radiographs. *Vet Radiology Ultrasound* **63**, 456–468 (2022).
10. Arsomngern, P. *et al.* Computer-Aided Diagnosis for Lung Lesion in Companion Animals from X-ray Images Using Deep Learning Techniques. in *2019 IEEE 10th International Conference on Awareness Science and Technology (iCAST)* 1–6 (IEEE, Morioka, Japan, 2019).
11. Banzato, T. *et al.* An AI-Based Algorithm for the Automatic Classification of Thoracic Radiographs in Cats. *Front. Vet. Sci.* **8**, 731936 (2021).
12. WYERS, Monique, COLLE, Marie-Anne, 2017. *Cancérologie générale*, Polycopié d'enseignement, UE N75, UV 73. Nantes: Oniris VetAgroBio, 171p.
13. North, S. M. & Banks, T. A. *Small Animal Oncology: An Introduction*. (Elsevier Saunders, Edinburgh, 2009).
14. De Biase, D. *et al.* Animal sentinels and cancer registries: state of the art and new perspectives. *Ann Res Oncol* **03**, 14 (2023).
15. Merlo, D. F. *et al.* Cancer Incidence in Pet Dogs: Findings of the Animal Tumor Registry of Genoa, Italy. *Veterinary Internal Medicine* **22**, 976–984 (2008).
16. Lamb, C. R., Whitlock, J. & Foster-Yeow, A. T. L. Prevalence of pulmonary nodules in dogs with malignant neoplasia as determined by CT. *Vet Radiology Ultrasound* **60**, 300–305 (2019).
17. Muhlbauer, M. C. & Kneller, S. K. *Radiography of the Dog and Cat*. (WILEY-BLACKWELL, 2013).
18. Thrall, D. E. *Textbook of Veterinary Diagnostic Radiology*. (Elsevier, St. Louis, 2018).
19. Manville, Sylvain, F., Marion. ÉTAPE 1: Réalisation d'un tableau de constantes radiographiques - Le Point Vétérinaire n° 308 du 01/09/2010. *Le Point Vétérinaire* **41**, 39–42 (2010).
20. Lang, J., Wortman, J. A., Glickman, L. T., Biery, D. N. & Rhodes, W. H. Sensitivity of radiographic detection of lung metastases in the dog. *Veterinary Radiology* **27**, 74–78 (1986).

21. Forrest, L. J. & Graybush, C. A. Radiographic patterns of pulmonary metastasis in 25 cats. *Vet Radiology Ultrasound* **39**, 4–8 (1998).
22. Reese, D. J. *et al.* Intra- and interobserver variability of board-certified veterinary radiologists and veterinary general practitioners for pulmonary nodule detection in standard and inverted display mode images of digital thoracic radiographs of dogs. *Javma* **238**, 998–1003 (2011).
23. Nemanic, S., London, C. A. & Wisner, E. R. Comparison of Thoracic Radiographs and Single Breath-Hold Helical CT for Detection of Pulmonary Nodules in Dogs with Metastatic Neoplasia. *Veterinary Internal Medicine* **20**, 508–515 (2006).
24. *BSAVA Manual of Canine and Feline Ultrasonography*. (British Small Animal Veterinary Association, Cheltenham, 2011).
25. Mattoon, J. S., Sellon, R. K. & Berry, C. R. *Small Animal Diagnostic Ultrasound*. (Elsevier, St. Louis, 2021).
26. Pires, S. T., Hage, M. C. F. N. S., Pinto, A. C. B. D. C. F. & Hagen, S. C. F. Comparative study between radiology and ultrasound in the evaluation of extracardiac thoracic diseases in dogs and cats. *Cienc. Rural* **45**, 2207–2213 (2015).
27. Tidwell, A. S. Ultrasonography of the Thorax (Excluding the Heart). *Veterinary Clinics of North America: Small Animal Practice* **28**, 993–1015 (1998).
28. Reichle, J. K. & Wisner, E. R. Non-cardiac thoracic ultrasound in 75 feline and canine patients. *Vet Radiology Ultrasound* **41**, 154–162 (2000).
29. *Veterinary Computed Tomography*. (Wiley-Blackwell, Chichester, West Sussex, UK, 2011).
30. Schwarz, T. & Johnson, V. S. *BSAVA Manual of Canine and Feline Thoracic Imaging*. (British small animal veterinary association, Quedgeley, 2008).
31. Eberle, N., Fork, M., Von Babo, V., Nolte, I. & Simon, D. Comparison of examination of thoracic radiographs and thoracic computed tomography in dogs with appendicular osteosarcoma. *Vet Comparative Oncology* **9**, 131–140 (2011).
32. Centre Hospitalier Vétérinaire Languedoc, 2019 Scanner / Imagerie spécialisée, *Centre Hospitalier Vétérinaire Languedoc* [en ligne], 2019. [Consulté le 20 mars 2024]. Disponible à l'adresse: <https://www.veterinaire-languedocia.com/services/scanner-imagerie/>.
33. Conseil national de l'Ordre des vétérinaires, *Atlas démographique de la profession vétérinaire 2021*, 6e éd., Observatoire national de la démographie de la profession vétérinaire, 2021
34. Larousse, 2008 , Définitions : intelligence, intelligences, *Dictionnaire Larousse* [en ligne], 2008. [Consulté le 15 octobre 2024]. Disponible à l'adresse : <https://www.larousse.fr/dictionnaires/francais/intelligence/43555>.
35. *Dictionnaire Hachette encyclopédique*. (Hachette, Paris, 2001).
36. McBee, M. P. *et al.* Deep Learning in Radiology. *Academic Radiology* **25**, 1472–1480 (2018).
37. Burt, J. R. *et al.* Deep learning beyond cats and dogs: recent advances in diagnosing breast cancer with deep neural networks. *BJR* 20170545 (2018).
38. Hennessey, E., DiFazio, M., Hennessey, R. & Cassel, N. Artificial intelligence in veterinary diagnostic imaging: A literature review. *Vet Radiology Ultrasound* **63**, 851–870 (2022).
39. Hastie, T., Tibshirani, R. & Friedman, J. H. *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*. (Springer, New York, NY, 2017).
40. Alpaydin, E. *Introduction to Machine Learning*. (MIT press, Cambridge, 2014).
41. Jiang, Y., Yang, M., Wang, S., Li, X. & Sun, Y. Emerging role of deep learning-based artificial intelligence in tumor pathology. *Cancer Communications* **40**, 154–166 (2020).

42. Vodrahalli, K. & Bhowmik, A. K. 3D computer vision based on machine learning with deep neural networks: A review. *J Soc Info Display* **25**, 676–694 (2017).
43. Gershenson, C. Artificial Neural Networks for Beginners. (2003).
44. Van Ginneken, B. Fifty years of computer analysis in chest imaging: rule-based, machine learning, deep learning. *Radiol Phys Technol* **10**, 23–32 (2017).
45. Chen, T., Xu, R., He, Y. & Wang, X. A Gloss Composition and Context Clustering Based Distributed Word Sense Representation Model. *Entropy* **17**, 6007–6024 (2015).
46. Bengio, Y. Learning Deep Architectures for AI. *FNT in Machine Learning* **2**, 1–127 (2009).
47. Tsai, H., Ambrogio, S., Narayanan, P., Shelby, R. M. & Burr, G. W. Recent progress in analog memory-based accelerators for deep learning. *J. Phys. D: Appl. Phys.* **51**, 283001 (2018).
48. Casado-García, Á. *et al.* CLoDSA: a tool for augmentation in classification, localization, detection, semantic segmentation and instance segmentation tasks. *BMC Bioinformatics* **20**, 323 (2019).
49. Maurice, B. Différences entre classification d'image, localisation/détection & segmentation d'objets, 2020, *Deeply Learning*[en ligne] 2020. [Consulté le 20 octobre 2024]. Disponible à l'adresse: <https://deeplylearning.fr/cours-theoriques-deep-learning/differences-entre-classification-dimage-localisation-detection-segmentation-dobjets/>
50. Marie-Laurence. Computer vision : tout connaître sur cette technique de l'IA!, 2023, *LePont*[en ligne], 2023. [Consulté le 20 octobre 2024]. Disponible à l'adresse: <https://www.lepont-learning.com/fr/computer-vision/> (2023).
51. Cho, J., Lee, K., Shin, E., Choy, G. & Do, S. How much data is needed to train a medical image deep learning system to achieve necessary high accuracy? Preprint at <http://arxiv.org/abs/1511.06348> (2016).
52. Kassel, R. Hyperparamètres : Qu'est-ce que c'est ? À quoi ça sert ? 2023, *DataScientest* [en ligne], 2023. [Consulté le 22 octobre 2024]. Disponible à l'adresse: <https://datascientest.com/hyperparametres-tout-savoir>
53. Ezanno, P. *et al.* Research perspectives on animal health in the era of artificial intelligence. *Vet Res* **52**, 40 (2021).
54. Walsh, D. P., Ma, T. F., Ip, H. S. & Zhu, J. Artificial intelligence and avian influenza: Using machine learning to enhance active surveillance for avian influenza viruses. *Transbound Emerg Dis* **66**, 2537–2545 (2019).
55. Ajayi, T., Dara, R. & Poljak, Z. Forecasting herd-level porcine epidemic diarrhea (PED) frequency trends in Ontario (Canada). *Preventive Veterinary Medicine* **164**, 15–22 (2019).
56. Feighelstein, M. *et al.* Explainable automated pain recognition in cats. *Sci Rep* **13**, 8973 (2023).
57. Steagall, P. V., Monteiro, B. P., Marangoni, S., Moussa, M. & Sautié, M. Fully automated deep learning models with smartphone applicability for prediction of pain using the Feline Grimace Scale. *Sci Rep* **13**, 21584 (2023).
58. Boneh-Shitrit, T. *et al.* Explainable automated recognition of emotional states from canine facial expressions: the case of positive anticipation and frustration. *Sci Rep* **12**, 22611 (2022).
59. Chambers, R. D. *et al.* Deep Learning Classification of Canine Behavior Using a Single Collar-Mounted Accelerometer: Real-World Validation. *Animals* **11**, 1549 (2021).
60. Biourge, V. *et al.* An artificial neural network-based model to predict chronic kidney disease in aged cats. *Veterinary Internal Medicine* **34**, 1920–1931 (2020).
61. Alwaysseh, A. *et al.* Evaluation of supervised machine-learning algorithms to distinguish between inflammatory bowel disease and alimentary lymphoma in cats. *J VET Diagn Invest* **28**, 679–687 (2016).
62. Reagan, K. L., Reagan, B. A. & Gilor, C. Machine learning algorithm as a diagnostic tool for hypoadrenocorticism in dogs. *Domestic Animal Endocrinology* **72**, 106396 (2020).

63. Van Der Horst, G. Computer Aided Sperm Analysis (CASA) in domestic animals: Current status, three D tracking and flagellar analysis. *Animal Reproduction Science* **220**, 106350 (2020).
64. Vinicki, K., Ferrari, P., Belić, M. & Turk, R. Using Convolutional Neural Networks for Determining Reticulocyte Percentage in Cats.
65. Estrada, A. H. *et al.* Diagnostic accuracy of computer aided electrocardiogram analysis in dogs. *J of Small Animal Practice* **62**, 145–149 (2021).
66. Nejedly, P. *et al.* Deep-learning for seizure forecasting in canines with epilepsy. *J. Neural Eng.* **16**, 036031 (2019).
67. Doi, K. Current status and future potential of computer-aided diagnosis in medical imaging. *BJR* **78**, s3–s19 (2005).
68. Suzuki, K. A review of computer-aided diagnosis in thoracic and colonic imaging. *Quantitative Imaging in Medicine and Surgery* **2**, (2012).
69. Mayerhoefer, M. E. *et al.* Introduction to Radiomics. *J Nucl Med* **61**, 488–495 (2020).
70. Gillies, R. J., Kinahan, P. E. & Hricak, H. Radiomics: Images Are More than Pictures, They Are Data. *Radiology* **278**, 563–577 (2016).
71. Basran, P. S. & Porter, I. Radiomics in veterinary medicine: Overview, methods, and applications. *Vet Radiology Ultrasound* **63**, 828–839 (2022).
72. McEvoy, F. J. & Amigo, J. M. Using machine learning to classify image features from canine pelvic radiographs: evaluation of partial least squares discriminant analysis and artificial neural network models. *Vet Radiology Ultrasound* **54**, 122–126 (2013).
73. Shim, H. *et al.* Deep learning-based diagnosis of stifle joint diseases in dogs. *Vet Radiology Ultrasound* **64**, 113–122 (2023).
74. Burti, S., Longhin Osti, V., Zotti, A. & Banzato, T. Use of deep learning to detect cardiomegaly on thoracic radiographs in dogs. *The Veterinary Journal* **262**, 105505 (2020).
75. Li, S., Wang, Z., Visser, L. C., Wisner, E. R. & Cheng, H. Pilot study: Application of artificial intelligence for detecting left atrial enlargement on canine thoracic radiographs. *Vet Radiology Ultrasound* **61**, 611–618 (2020).
76. Boissady, E., De La Comble, A., Zhu, X., Abbott, J. & Adrien-Maxence, H. Comparison of a Deep Learning Algorithm vs. Humans for Vertebral Heart Scale Measurements in Cats and Dogs Shows a High Degree of Agreement Among Readers. *Front. Vet. Sci.* **8**, 764570 (2021).
77. Berche, Q. Conception d'un programme d'Intelligence Artificielle pour calculer l'indice de Buchanan pour aider au diagnostic radiographique d'une cardiomégalie chez le chien. 85p. (2021).
78. Banzato, T. *et al.* Use of transfer learning to detect diffuse degenerative hepatic diseases from ultrasound images in dogs: A methodological study. *The Veterinary Journal* **233**, 35–40 (2018).
79. Burti, S., Zotti, A., Bonsembiante, F., Contiero, B. & Banzato, T. Diagnostic Accuracy of Delayed Phase Post Contrast Computed Tomographic Images in the Diagnosis of Focal Liver Lesions in Dogs: 69 Cases. *Front. Vet. Sci.* **8**, 611556 (2021).
80. Banzato, T., Bernardini, M., Cherubini, G. B. & Zotti, A. A methodological approach for deep learning to distinguish between meningiomas and gliomas on canine MR-images. *BMC Vet Res* **14**, 317 (2018).
81. Biercher, A. *et al.* Using Deep Learning to Detect Spinal Cord Diseases on Thoracolumbar Magnetic Resonance Images of Dogs. *Front. Vet. Sci.* **8**, 721167 (2021).
82. Banzato, T. *et al.* Automatic classification of canine thoracic radiographs using deep learning. *Sci Rep* **11**, 3964 (2021).
83. Kim, E., Fischetti, A. J., Sreetharan, P., Weltman, J. G. & Fox, P. R. Comparison of artificial intelligence to the veterinary radiologist's diagnosis of canine cardiogenic pulmonary edema. *Vet Radiology Ultrasound* **63**, 292–297 (2022).

84. Müller, T. R., Solano, M. & Tsunemi, M. H. Accuracy of artificial intelligence software for the detection of confirmed pleural effusion in thoracic radiographs in dogs. *Vet Radiology Ultrasound* **63**, 573–579 (2022).
85. Boissady, E., De La Comble, A., Zhu, X. & Hespel, A. Artificial intelligence evaluating primary thoracic lesions has an overall lower error rate compared to veterinarians or veterinarians in conjunction with the artificial intelligence. *Vet Radiology Ultrasound* **61**, 619–627 (2020).
86. Vetology | Veterinary Radiology and Artificial Intelligence. *Vetology* <https://vetology.ai/>.
87. Pomerantz, L. K., Solano, M. & Kalosa-Kenyon, E. Performance of a commercially available artificial intelligence software for the detection of confirmed pulmonary nodules and masses in canine thoracic radiography. *Vet Radiology Ultrasound* **64**, 881–889 (2023).
88. Wang, X., Peng, Y., Lu, L., Lu, Z. & Summers, R. M. TieNet: Text-Image Embedding Network for Common Thorax Disease Classification and Reporting in Chest X-rays. Preprint at <http://arxiv.org/abs/1801.04334> (2018).
89. Nam, J. G. *et al.* Development and Validation of Deep Learning–based Automatic Detection Algorithm for Malignant Pulmonary Nodules on Chest Radiographs. *Radiology* **290**, 218–228 (2019).
90. Behrendt, F. *et al.* A systematic approach to deep learning-based nodule detection in chest radiographs. *Sci Rep* **13**, 10120 (2023).
91. Xu, J., Ren, H., Cai, S. & Zhang, X. An improved faster R-CNN algorithm for assisted detection of lung nodules. *Computers in Biology and Medicine* **153**, 106470 (2023).
92. Smith, L. N. Cyclical Learning Rates for Training Neural Networks, (2017).
93. Association for Computational Linguistics. *ACL-IJCNLP 2009: Joint Conference of the 47th Annual Meeting of the Association for Computational Linguistics and 4th International Joint Conference on Natural Language Processing of the AFNLP, 2 - 7 August 2009, Suntec, Singapore. Vol. 1: Proceedings of the Conference.* vol. 1 (Singapore, 2009).
94. Liang, C.-H. *et al.* Identifying pulmonary nodules or masses on chest radiography using deep learning: external validation and strategies to improve clinical practice. *Clinical Radiology* **75**, 38–45 (2020).
95. Mumuni, A. & Mumuni, F. Data augmentation: A comprehensive survey of modern approaches. *Array* **16**, 100258 (2022).
96. Thrall, D. E. Principles of Radiographic Interpretation of the Thorax. in *Textbook of Veterinary Diagnostic Radiology* 568–582 (Elsevier, 2018).
97. Banzato, T. *et al.* An AI-based algorithm for the automatic evaluation of image quality in canine thoracic radiographs. *Sci Rep* **13**, 17024 (2023).



# Annexe I : Script python "predict.py" dans son intégralité

```
import argparse
import json
import os
from threading import Thread
import torch
from PIL import Image, ImageDraw, ImageShow
from torchvision import transforms as T
import torchvision
from torchvision.models import get_model
import torchvision.transforms.functional
from torchvision.ops import nms
import ssl

ssl._create_default_https_context = ssl._create_unverified_context

def apply_nms(predictions, iou_threshold=0.5, score_threshold=0.5):
    keep = predictions ["scores"] > score_threshold
    predictions = {
        "boxes": predictions ["boxes"] [keep] ,
        "scores" : predictions ["scores"] [keep]
        "labels" : predictions ["labels"] [keep]
    }

    nms_keep_indices = nms(predictions ["boxes"] , predictions ["scores"] , iou_threshold)
    prediction = {
        "boxes": predictions ["boxes"] [nms_keep_indices] ,
        "scores" : predictions ["scores"] [nms_keep_indices]
        "labels" : predictions ["labels"] [nms_keep_indices]
    }

    return predictions

def display_predictions(img, predictions):
    image = torchvision.transforms.functional.to_pil_image(img).convert("RGBA")
    draw = ImageDraw.Draw(image)

    if prediction and "boxes" in prediction:
        for index, box in enumerate(prediction["boxes"]):
            box = [int(coord) for coord in box.tolist()]
            score = prediction ["scores"] (index).item()
            label = prediction ["labels"] [index].item()

            text = f"Box{index+1} – Class: {label}, Score : {score: .2f}"
            color = colorsys.hsv_to_rgb(0.1 + 0.1 * label, 0.8, 0.8)
            color = tuple([int(c * 255) for c in color])

            draw.rectangle(box, outline=color, width=5)
            draw.text((box[0], box[1]), text, fill="black")

    Thread(target=lambda: ImageShow.show(image)).start()
```

```

def predicitions_to_via_regions(predictions) :
    regions = []
    for box, label in zip(prediction["boxes"].tolist(), prediction["labels"].tolist()):
        regions.append(
            {
                "shape_attributes": {
                    "name": "rect",
                    "x": int(box[0]),
                    "y": int(box[1]),
                    "width": int(box[2] - box[0]),
                    "height": int(box[3] - box[1]),
                },
                "region_attributes": {"class": label},
            }
        )
    Return regions

def main(model_path, img_dir, output_file) :
    basename = os.path.basename(model_path)
    model_type = basename.split("-")[0]
    model = get_model(name=model_type, weights="DEFAULT", num_classes=91)
    state_dict = torch.load(model_path, map_location=torch.device("cpu"))["state_dict"]
    weights = {
        k.replace("model.", "") : v
        for k, v in state_dict.items()
        if k.startswith("model.")
    }
    model.load_state_dict(weights)
    model.eval
    transform = T.Compose([T.ToTensor()])

    annotations = {}

    with torch.no_grad()
        for entry in os.scandir(imgs_dir) :
            if not entry.is_file()
                continue
            try:
                img = Image.open(entry.path).convert("RGB")
            except:
                print(f"Impossible de lire l'image {entry.path} »)
                continue
            img_input = transform(img)

            (predictions,) = model([img_input])
            predictions = apply_nms(predictions)
            regions = predicitions_to_via_regions(predictions)
            file_basename = os.path.basename(entry.path)
            annotations[file_basename] = {
                "filename" : file_basename,
                "size" = 0,
                "regions" = regions,
                "file_attributes" = {},
            }
            display_predictions(img_input, predicitions)
    json.dump(annotations, open(output_file, "w"))
    print(f"Annotations sauvegardes dans {output_file} »)

```

```
if __name__ == « __main__ » :
    parser = argparse.ArgumentParser(
        « Run predictions on a directory of images and output a VIA annotations file»
    )
    parser.add_argument(
        "model_file",
        help="Path to the model to use for predictions",
    )
    parser.add_argument(
        "--img_dir",
        default="./imgs",
        help="Directory where images are stored",
    )
    parser.add_argument(
        "--output_file",
        default="./annotations.json",
        help="Path to generated via annotations file",
    )
    )

args = parser.parse_ags()
main(args.model_file, args.img_dir, args.output_file)
```



# Annexe II : Détails sur la technique de Super Convergence

La Super Convergence est utilisée pour accélérer l'entraînement en appliquant des variations cycliques au *Learning rate* (LR). Dans cette étude, pour atteindre cette super convergence, 2 éléments ont été utilisés :

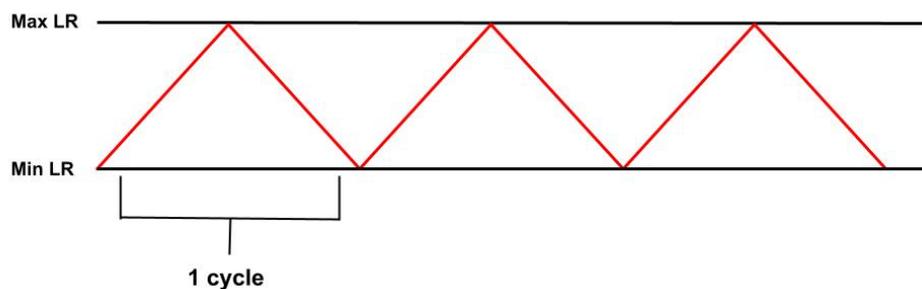
## 1- Le CyclicLR

Cet algorithme fait varier le taux d'apprentissage de manière cyclique, entre une borne inférieure et une borne supérieure. Cette stratégie aide à éviter les minimums locaux et peut améliorer la découverte de minimums globaux plus profonds. En effet :

- Un trop petit LR nécessite de très nombreuses modifications des poids rendant l'apprentissage extrêmement lent.
- Un LR trop élevé au contraire cause des changements trop drastiques des poids entraînant des comportements divergents.

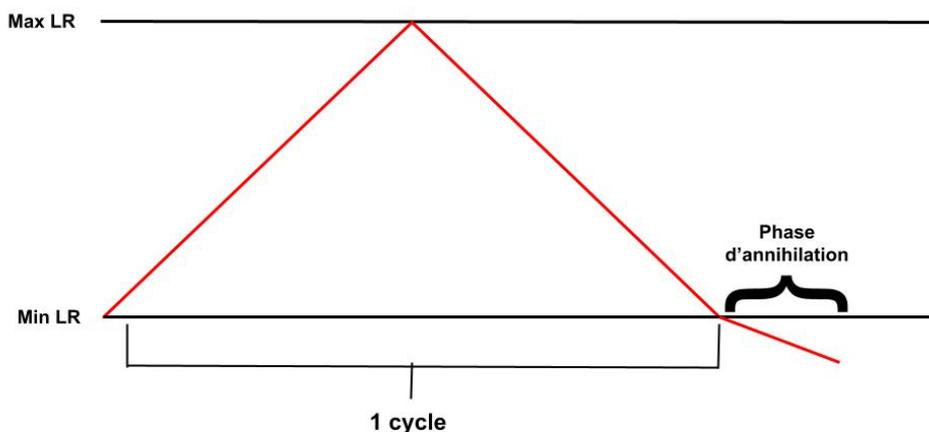
C'est là que le *CyclicLR* intervient, afin d'ajuster au mieux le LR.

Dans notre étude, le mode de variation suit une loi triangulaire, soit une simple variation linéaire.



Comme illustré sur le schéma ci-dessus, 3 cycles ont été réalisés avec le *CyclicLR*, chaque cycle étant représenté par un triangle rouge. Chaque cycle est composé de plusieurs itérations dans lesquelles le LR va atteindre la limite maximum puis redescendre jusqu'à sa valeur initiale minimum.

## 2- Le OneCycleLR



Cet algorithme reprend le principe du précédent avec cette fois-ci un unique cycle. À la dernière itération apparaît une phase d'annihilation, pour laquelle le LR est réduit en dessous de la valeur minimum fixée.

**Eve SOLAN**

**EVALUATION DE L'APPORT D'UNE APPROCHE DE LOCALISATION  
SPATIALE POUR LA DETECTION PAR DEEP LEARNING DES NODULES  
THORACIQUES CHEZ LE CHIEN ET LE CHAT**

**EVALUATION OF THE CONTRIBUTION OF A SPATIAL LOCALIZATION  
APPROACH FOR DEEP LEARNING-BASED DETECTION OF THORACIC  
NODULES IN DOGS AND CATS**

**Thèse d'État de Doctorat Vétérinaire : Nantes, le 26 septembre 2024**

**RESUME**

De nos jours, l'air du numérique s'installe à grande échelle et cherche aujourd'hui à répondre à des objectifs de diagnostic dans le domaine de la médecine vétérinaire. Dans ce contexte, l'étude que nous avons menée cherche à développer un nouvel algorithme de *Deep Learning* pour la détection des métastases pulmonaires, via une approche de localisation, afin de proposer une meilleure prise en charge aux patients et à leur propriétaire.

Dans ce but, cette étude a utilisé 711 radiographies thoraciques de chiens et de chats dont les métastases ont été préalablement annotées par l'auteur afin d'entraîner un modèle de détection automatisée. Les performances de ce modèle ont ensuite été évaluées sur 20% des 366 radiographies offrant la meilleure détection possible. Le score F1 obtenue lors de ce dernier test est de 0,494 avec une précision de 50%. L'algorithme de DL ainsi développé propose des performances moyennes à ce jour.

Ces performances sont insuffisantes pour permettre un usage immédiat, mais offrent des perspectives prometteuses.

**MOTS CLES :**

- CHIEN
- CHAT
- THORAX
- LOCALISATION ANATOMIQUE
- METASTASE TUMORALE
- RADIOGRAPHIE THORACIQUE
- AIDE AU DIAGNOSTIC
- INTELLIGENCE ARTIFICIELLE
- APPRENTISSAGE PROFOND

**DATE DE SOUTENANCE : 26 septembre 2024**