

UNIVERSITE TOULOUSE III – PAUL SABATIER
FACULTE DE SANTE – DEPARTEMENT D'ODONTOLOGIE

ANNEE 2024

2024 TOU3 3038

THÈSE

POUR LE DIPLOME D'ETAT DE DOCTEUR EN CHIRURGIE DENTAIRE

Présentée et soutenue publiquement

par

Pierre STEFFENS

Le 31 mai 2024

**INTÉRÊTS DE LA MIRE COLORIMÉTRIQUE POUR
LE DIAGNOSTIC DE L'INFLAMMATION GINGIVALE
PAR RÉSEAUX NEURONAUX CONVOLUTIFS**

Directeurs de thèse : Pr Paul MONSARRAT

JURY

Président :	Pr Sara LAURENCIN-DALICIEUX
1 ^{er} assesseur :	Pr Paul MONSARRAT
2 ^{ème} assesseur :	Dr Antoine GALIBOURG
3 ^{ème} assesseur :	Dr Antoine DUBUC

→ **DIRECTION**

Doyen de la Faculté de Santé

M. Philippe POMAR

Vice Doyenne de la Faculté de Santé

Directrice du Département d'Odontologie

Mme Sara DALICIEUX-LAURENCIN

Directeurs Adjoints

Mme Sarah COUSTY

M. Florent DESTRUHAUT

Directrice Administrative

Mme Muriel VERDAGUER

Présidente du Comité Scientifique

Mme Cathy NABET

→ **HONORARIAT**

Doyens honoraires

M. Jean LAGARRIGUE +

M. Jean-Philippe LODTER +

M. Gérard PALOUDIER

M. Michel SIXOU

M. Henri SOULET

Chargés de mission

M. Karim NASR (*Innovation Pédagogique*)

M. Olivier HAMEL (*Maillage Territorial*)

M. Franck DIEMER (*Formation Continue*)

M. Philippe KEMOUN (*Stratégie Immobilière*)

M. Paul MONSARRAT (*Intelligence Artificielle*)

→ **PERSONNEL ENSEIGNANT**

Section CNU 56 : Développement, Croissance et Prévention

1. 56.01 ODONTOLOGIE PEDIATRIQUE et ORTHOPEDIE DENTO-FACIALE (Mme Isabelle BAILLEUL-FORESTIER)

2. ODONTOLOGIE PEDIATRIQUE

Professeurs d'Université : Mme Isabelle BAILLEUL-FORESTIER, M. Frédéric VAYSSE

Maîtres de Conférences : Mme Marie- Cécile VALERA, M. Mathieu MARTY

Assistants : Mme Anne GICQUEL, M. Robin BENETAH

Adjoints d'Enseignement : M. Sébastien DOMINE, M. Mathieu TESTE, M. Daniel BANDON

ORTHOPEDIE DENTO-FACIALE

Maîtres de Conférences : M. Pascal BARON, M. Maxime ROTENBERG

Assistants : Mme Carole VARGAS JOULIA, Mme Chahrazed BELAILI, Mme Véronique POINSOTTE

Adjoints d'Enseignement : Mme. Isabelle ARAGON, M. Vincent VIDAL-ROSSET

3. 56.02 PRÉVENTION, ÉPIDÉMIOLOGIE, ÉCONOMIE DE LA SANTÉ, ODONTOLOGIE LÉGALE (Mme NABET Catherine)

Professeurs d'Université : M. Michel SIXOU, Mme Catherine NABET, M. Olivier HAMEL, M. Jean-Noël VERGNES

Maîtres de Conférences : Mme Géromine FOURNIER

Adjoints d'Enseignement : M. Alain DURAND, Mlle. Sacha BARON, M. Romain LAGARD, M. Jean-Philippe GATIGNOL

Mme Carole KANJ, Mme Mylène VINCENT-BERTHOUMIEUX, M. Christophe BEDOS

Section CNU 57 : Chirurgie Orale, Parodontologie, Biologie Orale

4. 57.01 CHIRURGIE ORALE, PARODONTOLOGIE, BIOLOGIE ORALE (M. Philippe KEMOUN)

5. PARODONTOLOGIE

Professeurs d'Université : Mme Sara LAURENCIN- DALICIEUX,

Maîtres de Conférences : Mme Alexia VINEL, Mme. Charlotte THOMAS

Assistants : M. Antoine AL HALABI, M. Pierre JEHLE

Adjoints d'Enseignement : M. Loïc CALVO, M. Antoine SANCIER, M. Ronan BARRE, Mme Myriam KADDECH, M. Matthieu RIMBERT, M. Joffrey DURAN

6. CHIRURGIE ORALE

Professeur d'Université : Mme Sarah COUSTY
Maîtres de Conférences : M. Philippe CAMPAN, M. Bruno COURTOIS
Assistants : M. Antoine DUBUC
Adjoints d'Enseignement : M. Gabriel FAUXPOINT, M. Arnaud L'HOMME, Mme Marie-Pierre LABADIE, M. Jérôme SALEFRANQUE,
M. Clément CAMBRONNE

7. BIOLOGIE ORALE

Professeurs d'Université : M. Philippe KEMOUN, M. Vincent BLASCO-BAQUE
Maîtres de Conférences : M. Pierre-Pascal POULET, M. Matthieu MINTY
Assistants : Mme Chiara CECCHIN-ALBERTONI, M. Maxime LUIS, Mme Valentine BAYLET GALY-CASSIT,
Mme Sylvie LE
Adjoints d'Enseignement : M. Mathieu FRANC, M. Hugo BARRAGUE, Mme Inessa TIMOFEEVA-JOSSINET

Section CNU 58 : Réhabilitation Orale

8. 58.01 DENTISTERIE RESTAURATRICE, ENDODONTIE, PROTHESES, FONCTIONS-DYSFONCTIONS, IMAGERIE, BIOMATERIAUX (M. Franck DIEMER)

9. DENTISTERIE RESTAURATRICE, ENDODONTIE

Professeur d'Université : M. Franck DIEMER
Maîtres de Conférences : M. Philippe GUIGNES, Mme Marie GURGEL-GEORGELIN, Mme Delphine MARET-COMTESSE
Assistants : M. Nicolas ALAUX, M. Vincent SUAREZ, M. Lorris BOIVIN, M. Thibault DECAMPS, Mme Emma STURARO,
Mme Anouk FESQUET
Adjoints d'Enseignement : M. Eric BALGUERIE, M. Jean- Philippe MALLET, M. Rami HAMDAN, M. Romain DUCASSE,
Mme Lucie RAPP, Mme Marion CASTAING-FOURIER

PROTHÈSES

Professeurs d'Université : M. Philippe POMAR, M. Florent DESTRUHAUT,
Maîtres de Conférences : M. Antoine GALIBOURG, M. Julien DELRIEU
Assistants : Mme Coralie BATAILLE, Mme Mathilde HOURSET, Mme Constance CUNY
M. Anthony LEBON, M. Paul POULET
Adjoints d'Enseignement : M. Christophe GHRENASSIA, Mme Marie-Hélène LACOSTE-FERRE, M. Olivier LE GAC, M. Luc
RAYNALDI, M. Jean Claude COMBADAZOU, M. Bertrand ARCAUTE, M. Fabien LEMAGNER, M. Eric
SOLYOM,
M. Michel KNAFO, M. Victor EMONET-DENAND, M. Thierry DENIS, M. Thibault YAGUE,
M. Antonin HENNEQUIN, M. Bertrand CHAMPION

10. FONCTIONS-DYSFONCTIONS, IMAGERIE, BIOMATERIAUX

Professeur d'Université : Mr. Paul MONSARRAT
Maîtres de Conférences : Mme Sabine JONQOT, M. Karim NASR, M. Thibault CANCEILL,
Assistants : M. Olivier DENY, Mme Laura PASCALIN, Mme Alison PROSPER
Adjoints d'Enseignement : Mme Sylvie MAGNE, M. Thierry VERGÉ, M. Damien OSTROWSKI

Mise à jour pour le 01 Mai 2024

Remerciements

Cette thèse n'aurait pas été possible sans l'aide de nombreuses personnes. :

Aux professeurs, praticiens et étudiants du DU de parodontologie ; pour leur aide et intérêt dans mes travaux de recherche. Merci pour la confiance accordée, notamment pour le prêt de matériels photographique onéreux. Merci pour leur temps, en plus des données recueillis j'ai appris beaucoup de chose à les observer.

À Gaëlle ; pour ta relecture, tes nombreux encouragements ; ta bonne humeur quand les échecs et les difficultés m'empêchaient d'avancer. Tu as su dissiper mes doutes, et éclaircir des idées très noires, et sans toi il n'y aurait pas de thèse. Ensemble nous sommes plus forts.

Aux enseignants de la faculté d'odontologie ; vous êtes évidemment impliqué dans toutes les thèses soutenues. Mais vous m'avez appris à soigner, à lire des articles scientifiques, et je sais surtout la chance que j'ai eu d'étudier auprès de vous à Toulouse. Nous avons la chance d'avoir des enseignants bienveillants et impliqués dans l'andragogie.

À mes amis, dentistes français et européens ; trop nombreux à citer, mais qui demeurent une source d'inspiration tous à leur manière. Leurs conseils, leur rigueur et leurs encouragements ont grandement influencé ce travail. Merci pour ces souvenirs et cette joie.

À tous mes proches, famille et amis toulousains ou albigeois ; qui sans contribuer directement, ont par leurs encouragements et sollicitudes participé à dissiper ces idées noires.

À mes professeurs et camarades du master d'éthique de Toulouse ; provenant de milieux universitaires éloignés du mien. Ils ont été source de découvertes. Et m'ont permis de m'ouvrir à d'autres horizons sur mon travail et de le réfléchir sous un angle différent. J'espère poursuivre ces études une fois cette thèse terminée.

À tous les anonymes du net, programmeurs et chercheurs ; tous ceux qui participent à la science et à la diffuser. Internet est un outil formidable, entre les vidéos et les IA conversationnelles, j'ai tellement appris. Il est dommage que certains sites d'accès à la science soient interdits en France, que vive la science libre.

À nos patients ; et particulièrement ceux du DU de parodontologie, qui finalement mériteraient la première place. Sans eux, pas d'étude, pas de photographies ni de questionnaires. Merci pour votre patience et intérêt dans mon travail, merci pour votre bienveillance.

À notre présidente du jury,

Madame le Professeur Sara LAURENCIN-DALICIEUX

- Vice Doyenne de la Faculté de Santé de l'Université de Toulouse III Paul Sabatier
- Directrice Département d'Odontologie
- Professeur des Universités, Praticien Hospitalier d'Odontologie,
- Docteur en Chirurgie Dentaire,
- Docteur de l'Université Paul Sabatier,
- Diplôme Universitaire de Parodontologie
- Lauréate de l'université Paul Sabatier
- Diplôme Universitaire d'Injection d'Acide Hyaluronique en Odontologie
- Diplôme Universitaire Approches Innovantes en Recherche Biomédicale et en Méta-recherche
- Habilitation à Diriger des Recherches (H.D.R.).

Je vous remercie de présider ce jury de thèse.

*Je vous suis très reconnaissant de votre enseignement, et de la confiance accordée
notamment pour le prêt de matériel photo et m'avoir permis d'observer les séances
du DU de paro.*

*Votre investissement auprès de la faculté et pour les étudiants ainsi que votre
rigueur sont un exemple pour moi.*

Votre présence m'honore.

À notre jury de thèse,

Monsieur le Professeur Paul MONSARRAT

- Professeur des Universités, Praticien Hospitalier d'Odontologie,
- Docteur de l'Université Paul Sabatier - Spécialité Physiopathologie,
- Diplôme Universitaire d'Imagerie 3D maxillo-faciale,
- Diplôme universitaire de Recherche Clinique en Odontologie,
- Habilitation à Diriger les Recherches
- Lauréat de la faculté de Médecine Rangueil et de Chirurgie Dentaire de l'Université Paul Sabatier

Je suis très heureux d'avoir pu vous compter comme directeur de thèse. Je vous remercie de votre confiance, et de m'avoir laissé la pleine latitude pour traiter le sujet sous un angle bien personnel.

Je vous remercie également de votre enseignement et encadrement. Votre énergie et bienveillance marquent de nombreux souvenirs. Vos encouragements et soutiens ont accompagné mes études de l'anatomie premiers gestes chirurgicaux.

Votre présence est un honneur et plaisir.

À notre jury de thèse,

Monsieur le Docteur Antoine GALIBOURG

- Maître de conférence des Universités, Praticien Hospitalier d'Odontologie
- Docteur en Chirurgie Dentaire,
Docteur de l'Université Paul Sabatier
- Ingénieur de l'Institut Catholique des Arts et Métiers,
- Responsable du Diplôme d'Université d'Implantologie

*Je vous remercie de votre présence dans ce jury, et de votre intérêt pour ce
travail.*

Soyez assuré de ma considération et de mon profond respect.

À notre jury de thèse,

Monsieur le Docteur Antoine DUBUC

- Chef de Clinique des Universités – Assistant des Hôpitaux
- Docteur en Chirurgie Dentaire,
- Spécialiste qualifié en Chirurgie Orale
- Master 2 : Sciences chirurgicales et nouvelles technologies
- Lauréat de l'université Paul Sabatier

Je suis très heureux de ta présence dans ce jury.

J'ai été très fier d'avoir pu poursuivre à mon niveau les travaux que tu avais commencé. Ceux-ci ont été une source d'inspiration, et j'espère avoir pu contribuer pour que ce sujet soit poursuivi. Il faut également t'attribuer l'ébauche de la mire colorimétrique sans qui rien n'aurait été possible.

Sois assuré de ma considération et de mon respect. Avoir eu la possibilité de s'inspirer de tes travaux et ta présence sont déjà une grande fierté.

Glossaire

DU : Diplôme-Universitaire

CNN : Convolutional Neural Network, réseau neuronal convolutif

MLP : Multi Layer Perceptron, perceptron multi-couche

DF : DataFrame, tableau de données

Table des matières

Glossaire	10
1. Introduction	12
2. Cadre contextuel et conceptuel	15
2.1. L'intelligence artificielle	15
2.2. Le neurone artificiel	15
2.3. Le traitement d'image et les réseaux convolutifs	24
2.4. Les réseaux convolutifs en odontologie	29
2.5. Les mires colorimétriques.....	31
3. Matériels et méthode.....	34
3.1. Critères de sélection	34
3.2. Collecte des données	35
3.3. Développement du modèle	40
4. Résultats	51
5. Analyse et discussion.....	56
5.1. Rappel des problématique et objectifs.....	56
5.2. Analyse des biais	56
5.3. Points forts.....	63
5.4. Piste d'améliorations	64
6. Discussion	66
6.1. La place de la mire dans le diagnostic par réseaux convolutifs.....	66
6.2. La place de l'IA en parodontologie	66
6.3. La place de l'IA dans les soins	68
Conclusion	70
Table des illustrations	71
Figures.....	71
Tableaux.....	72
Bibliographie.....	73

1. Introduction

Les maladies parodontales représentent un ensemble de pathologies inflammatoires qui affectent les tissus de soutien des dents, notamment les gencives, le ligament parodontal et l'os alvéolaire. Elles sont largement répandues dans la population mondiale et constituent un problème de santé bucco-dentaire majeur, environ la moitié de la population après 50 ans (1). L'identification précoce et la prise en charge adéquate de ces affections revêtent une importance capitale pour prévenir la progression de la maladie et préserver la santé bucco-dentaire des individus.

Les maladies parodontales désignent un large spectre de pathologies, depuis des symptômes variés, une inflammation réversible des gencives à une affection plus sévère qui entraîne une destruction progressive des tissus de soutien des dents. Les facteurs de risque incluent entre autres une mauvaise hygiène bucco-dentaire, le tabagisme, le diabète, le stress, ainsi que certains facteurs génétiques et hormonaux. Si elles ne sont pas traitées à temps, les maladies parodontales peuvent conduire à la perte des dents et être associées à des problèmes de santé générale, tels que les maladies cardiovasculaires et le diabète (2).

La prise en charge précoce des maladies parodontales est cruciale pour prémunir le patient de la perte de ces tissus de soutien. Une détection précoce permet d'initier rapidement des mesures préventives et thérapeutiques appropriées, telles que le détartrage, le débridement radiculaire non chirurgical, voire la chirurgie parodontale dans les cas avancés (2). Le diagnostic de ces maladies parodontales repose sur l'évaluation clinique du praticien : la présence de poches parodontales, le saignement

des gencives, la récession gingivale et l'inflammation gingivale. Certains de ces symptômes sont évalués de manière visuelle par le praticien.

Les critères visuels d'inflammation gingivale, tels que la rougeur, et le saignement des gencives, peuvent indiquer la présence de sites actifs de maladie parodontale. Cependant, l'évaluation précise de ces signes inflammatoires peut être difficile et dépendante de l'expérience et de la formation du praticien. C'est ici que l'intégration des technologies de l'Intelligence Artificielle (**IA**) et des Réseaux Convolutifs Neuronaux (*Convolutif Neural Network*, **CNN**) peut apporter une contribution significative.

Dans le cadre de cette thèse d'exercice, nous nous intéressons spécifiquement à l'utilisation des réseaux de neurones convolutifs pour l'aide au diagnostic des lésions gingivales associées aux maladies parodontales. Les CNNs sont des modèles d'intelligence artificielle qui ont montré leur efficacité dans divers domaines de la science et de la médecine, notamment dans l'analyse d'images médicales. Les CNNs pourrait être entraînés à reconnaître et à interpréter les caractéristiques visuelles spécifiques des lésions gingivales et à les associer à des stades précoces de la maladie parodontale. Leur potentiel d'automatisation et de prise de décision assistée apporterait de nouveaux outils pour la dentisterie. Cette automatisation permettrait d'améliorer la détection précoce, d'objectiver le processus de diagnostic et de faciliter la prise de décision clinique.

Cependant, si les CNNs se montrent déjà efficaces quant au diagnostic de pathologies orales, cela se fait principalement à partir d'images standardisées (comme les clichés radiographiques). Si l'objectif du praticien est d'utiliser ces réseaux à partir de photographies intra-orales, il faut trouver un moyen de standardiser la prise de

cliché. Ce qui demeure un obstacle. Nous avons dès lors pensé à la calibration d'image via une mire colorimétrique utilisé notamment en dermatologie.

L'objectif principal de cette thèse d'exercice est donc d'explorer le potentiel d'une mire colorimétrique lors de la prise de cliché intra-oraux pour une exploitation par un CNN. Nous chercherons à développer deux types de modèles de CNNs, l'un sans information colorimétrique, l'autre avec les informations de la mire colorimétrique. Leurs architectures seront dès lors sensiblement similaires. Leurs entraînements s'effectueront à partir de sets de données distincts qui différeront principalement par l'utilisation d'une mire colorimétrique lors de la prise des clichés. Et nous compareront leurs capacités de prédiction de l'inflammation gingivale à partir de ces clichés, selon des métriques adaptées afin de répondre à la question suivante :

L'utilisation d'une mire colorimétrique lors de la prise de clichés intra-oraux, améliore-t-elle la capacité de prédiction de l'index d'inflammation gingivale d'un réseau neuronal convolutif ?

2. Cadre contextuel et conceptuel

2.1. L'intelligence artificielle

L'idée d'intelligence artificielle a accompagné le développement de l'informatique. Alan Turing pose la question dès 1950 « Can machines think ? »(3), il y développe l'idée d'un test qui permettrait de considérer une machine intelligente lorsque celle-ci soutient une conversation avec un humain sans que celui-ci ne le décèle.

Le terme « d'intelligence artificielle » est introduit lors de la conférence de Dartmouth (Dartmouth Summer Research Project on Artificial Intelligence) durant l'été 1956. Dès lors, cette discipline devient un champ de recherche à part entière. Elle va se développer concomitamment au développement des neurosciences, dont elle s'inspirera pour théoriser des modèles décisionnels artificiels à partir des découvertes sur le vivant.

2.2. Le neurone artificiel

L'apparition d'un modèle mathématique représentant le neurone biologique est dû à McCulloch et Pitts en 1943 (4). Ils développent la programmation de neurone artificiels à partir du fonctionnement des neurones biologiques.

Si en biologie le neurone est une cellule excitatrice composée de dendrites et d'un axone, dans le modèle de « *linear threshold function* » des auteurs, elles y sont imitées par des entrées (inputs) et une sortie (output).

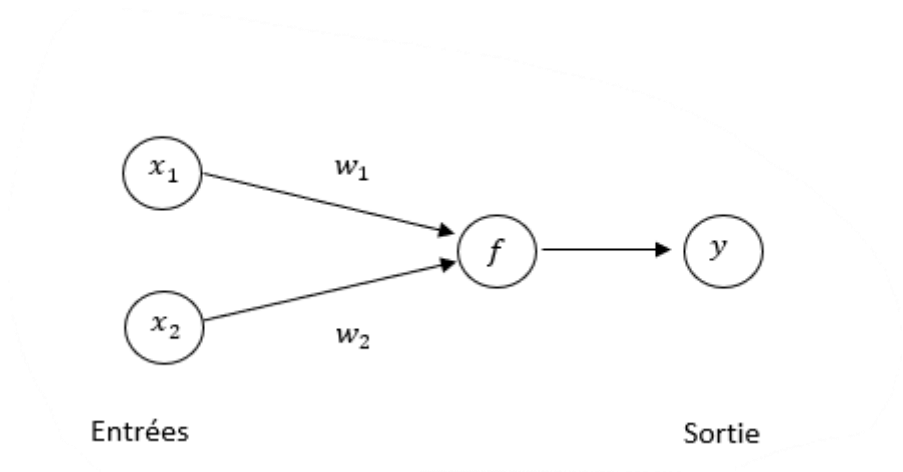


Figure 1, Schéma de neurone artificiel

f est une fonction de combinaison, qui calcule la somme des valeurs d'entrées x pondérées par des poids w (coefficient synaptiques) :

$$f = x_1 w_1 + x_2 w_2$$

Cette fonction émettra un signal de sortie auquel on appliquera une fonction seuil, une fonction d'activation, ici heaviside, qui permet d'obtenir une sortie y qui vaudra 0 ou 1 :

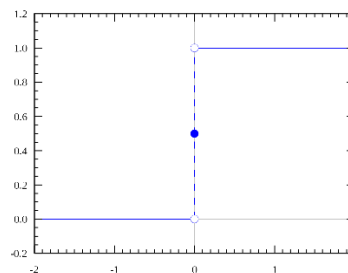


Figure 2, Fonction d'activation

$$y = 0 \forall f < 0 ; y = 1 \forall f > 0$$

Le neurone artificiel qu'ils développent est donc un outil qui permet à partir de deux entrées, de distinguer deux classes selon une sortie qui vaut 1 ou 0. Cependant les valeurs de pondération sont à adapter par l'utilisateur. Le neurone ne possède ici pas de capacité d'apprentissage.

1.1.1. Le perceptron

En 1958, le psychologue Rosenblatt développe le perceptron (5), à partir des théories développés en neurosciences sur le fonctionnement des neurones du vivant notamment par Hebb (1949). Selon ce dernier, lorsque deux neurones biologiques sont excités conjointement, ils renforcent leur lien synaptique.

Le perceptron est un modèle mathématique de neurone artificiel à n entrées auxquelles on applique des poids w_n ; à la somme de ces produits pourra être ajouté un biais noté b .

$$f = x_1w_1 + x_2w_2 + \dots + x_nw_n + b$$

A partir du résultat de f on appliquera une fonction d'activation qui donnera un résultat y qui vaudra 0 ou 1.

L'avancée par rapport au neurone McCulloch-Pitts est la capacité du perceptron à évoluer, à savoir s'améliorer grâce à un algorithme d'apprentissage. Il fonctionne sur un ensemble de données connus x et y , en comparant la sortie obtenue y de la sortie attendue $\hat{y}$ il est possible ajuster les valeurs des poids w_n . A chaque itération k les poids w seront donc ajustés pour améliorer la précision du modèle.

$$w_n(k + 1) = w_n(k) + (y - \hat{y})x_n\eta$$

Où $w_n(t)$ représente une valeur de poids w_n au moment d'une itération t et donc $w_n(t + 1)$ représente cette nouvelle valeur après une itération supplémentaire ;

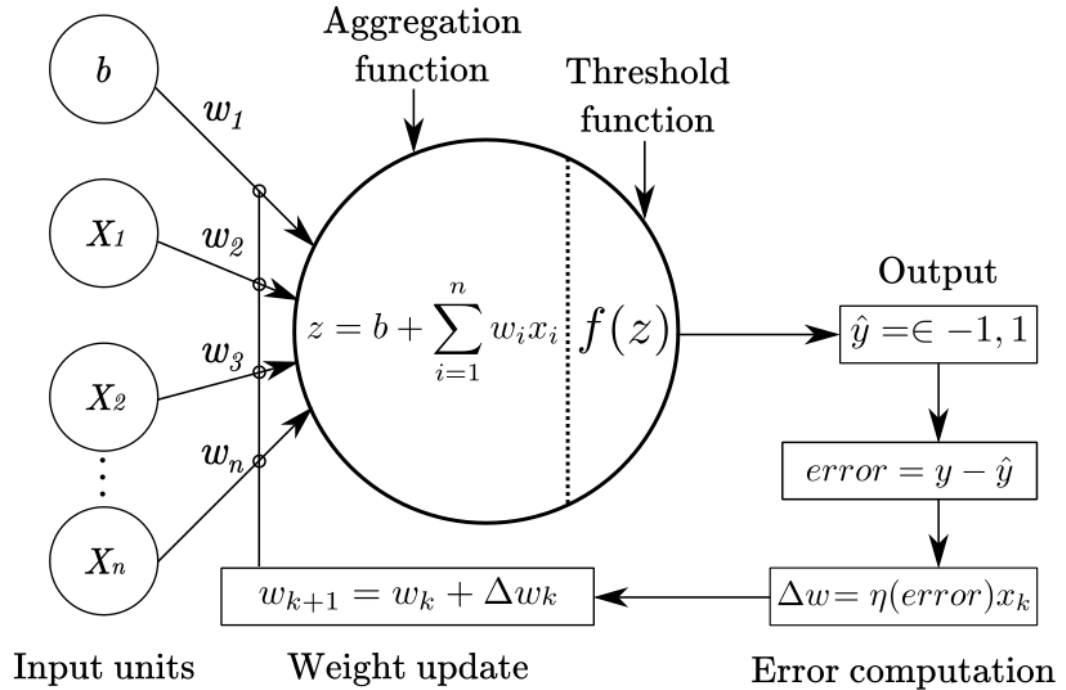


Figure 3, Schéma du perceptron, Pablo Caceres, <https://pabloinsente.github.io/the-perceptron>

La formule $(y - \hat{y})$ représente l'erreur, soit la sortie attendue du neurone, à laquelle on soustrait la sortie du neurone de cette itération. Cette formule vaudra 0 si la valeur de la sortie est identique à celle attendue, égale à 1 si la sortie est supérieure à la valeur attendue, égale à -1 si la sortie est inférieure à la valeur attendue.

Ainsi on modifiera les pondérations selon les cas suivant :

$$w_n(k+1) = w_n(k) \quad ; \text{ dans le cas où } (y - \hat{y}) = 0 \rightarrow \hat{y} = y$$

$$w_n(k+1) = w_n(k) + x_n \quad ; \text{ dans le cas où } (y - \hat{y}) = 1 \rightarrow \hat{y} = 0, y = 1$$

$$w_n(k+1) = w_n(k) - x_n \quad ; \text{ dans le cas où } (y - \hat{y}) = -1 \rightarrow \hat{y} = 1, y = 0$$

Enfin le symbole η représente le learning rate (ou pas d'apprentissage), un paramètre défini au préalable (en générale 0.01 ou 0.001) qui permet de contrôler la rapidité à laquelle les poids sont mis à jour. Enfin d'éviter qu'un paramètre plus

impactant ne fasse converger la pondération associée en négligeant les pondérations des autres entrées.

On voit donc que l'algorithme d'apprentissage permet l'ajustement de chacune des pondérations de chaque entrée après chaque itération, afin que la sortie soit celle attendue.

A partir d'un ensemble de données sur lequel le perceptron « s'entraînera », il sera possible après un certain nombre d'itération de réaliser des « prédictions » face à de nouvelles données. Le perceptron est donc capable de prédire l'appartenance à une catégorie A ($y = 1$) ou B ($y = 0$), d'où le terme d'algorithme de classification binaire.

Cependant le perceptron de Rosenblatt se limite à la classification de modèle linéaire : c'est-à-dire deux catégories qui sont distinguées en deux ensembles, pour lesquels $f \geq 0$ ou $f < 0$. La frontière de décision pourra donc être schématisée par une courbe :

$$w_1x_1 + w_2x_2 + b = 0.$$

Le perceptron de Rosenblatt est donc inadapté pour différencier deux sets de données, dès lors qu'ils ne sont pas séparables de manière linéaire.

Imaginons deux ensembles de patients adultes du même sexe que nous classerions selon leur IMC (ici l'exemple est présenté pour la compréhension plutôt que pour sa pertinence médical)

Ensemble 1 : Individus considérés au moins en surpoids

Ensemble 2 : Individus considérés au plus ayant un poids idéal

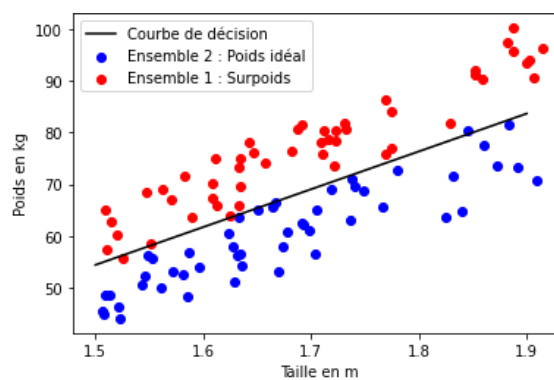


Figure 4, Exemple de frontière de décision linéaire

Ici les 2 ensembles semblent être séparables linéairement, et le perceptron de Rosenblatt pourrait être adapté pour distinguer ces sets de données. Ici le perceptron donne une courbe de décision $f = 72,96x - 55$ (avec x qui correspond à la taille en m). En réalité la frontière de classification de l'IMC n'est pas linéaire, en élargissant les sets de données (qui sont certes exceptionnels dans la population) le perceptron commet plus d'erreur, la courbe n'est plus adaptée :

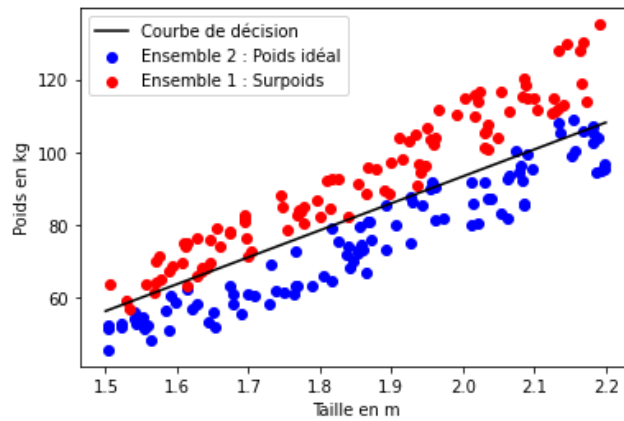


Figure 5, Exemple de frontière de décision non linéaire

En effet, la courbe de décision la plus adaptée serait plutôt $f = 24x^2$

1.1.2. Le perceptron multicouche

Pour résoudre les problèmes posés par le perceptron, Rumelhart et al. (6) proposent de rajouter aux entrées et sorties des unités cachées, qu'on organisera en différentes couches (la couche d'entrée étant numéroté couche 1) :

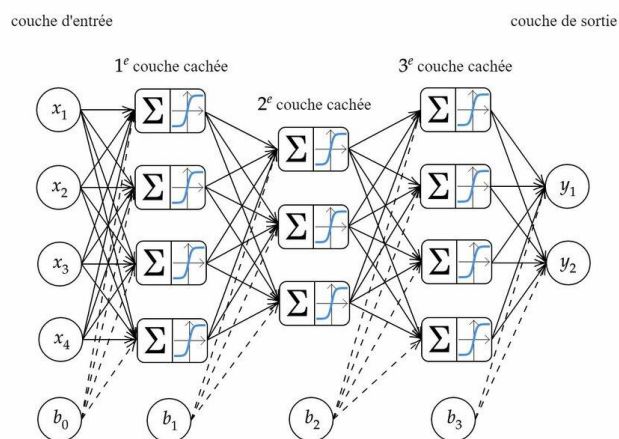


Figure 6, Exemple d'un perceptron multicouche avec 5 couches. M. Beaulieu, Analyse de la réduction du chatoiement sur les images, Thèse de l'Université de Montréal

Les unités cachées étant rajoutées en couche, on parle de perceptron multicouche (ou multi-linear-perceptron, MLP). Suivant le principe du perceptron de Rosenblatt (ou

perceptron monocouche), chaque unité verra sa valeur pondérée par un poids w . On parlera de Forward propagation (Propagation vers l'avant) pour désigner le remplissage des unités suivantes. Elle se fait par la somme, de la pondération des unités précédentes par leur poids respectives :

$$x_{ij} = \sum_{i=1}^m x_{i(j-1)} w_{i(j-1)} + b$$

Avec x représentant les valeurs des unités numérotés de i à m au sein de la couche j , et w les poids qui pondèrent chaque unité et b un biais aléatoire.

Une fois les différentes unités du modèle remplies, sont appliqués aux unités de sortie des fonctions d'activation qui donneront une valeur entre 0 et 1, ici les auteurs utilisent une fonction sigmoïde qui a l'avantage d'être linéaire par rapport à la fonction heaviside, et définie sur 0 :

$f(x) = \frac{1}{1+e^{-x}}$: avec x la valeur de l'unité de sortie

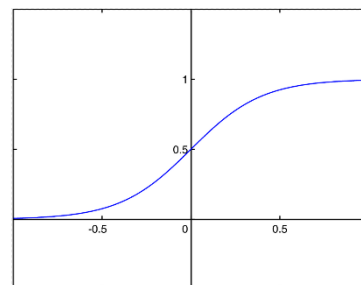


Figure 7, Fonction d'activation sigmoïde

Suite à la Forward propagation, il sera nécessaire de mesurer la performance du modèle. Pour ce faire on utilisera une Loss function (Fonction de perte) : une fonction qui évaluera la qualité du modèle. Notamment en quantifiant l'écart entre les valeurs de(s) l'unité(s) de sortie et les valeurs attendues. Dans leur article, Rumelhart et al. ont utilisé une fonction d'erreur quadratique moyenne, qui mesure l'écart moyen au carré entre les valeurs obtenus et les valeurs attendues.

$$E = \frac{1}{2} \sum_{i=1}^m (f(x) - y)^2 ;$$

E est l'erreur du modèle (ou fonction coût) ; $f(x)$ est la sortie de la fonction d'activation de chaque unité de sortie, y la sortie attendue de chaque unité de sortie

Le but du MLP sera donc de minimiser cette erreur E. Dès lors, il sera possible de prévoir l'évolution de cette loss function, en calculant sa dérivée.

A nouveau le learning rate η vient contrôler la vitesse à laquelle les pondérations vont converger vers une pondération optimale, par son produit à la dérivée de la fonction coût:

$$w_{ij+1} = w_{ij} - \eta \frac{\partial E}{\partial w} ;$$

w représente les différents poids des unités, $\frac{\partial E}{\partial w}$ la dérivée de la Loss function

Rumelhart et al. vont donc pouvoir modifier les poids de chacune des unités : ils réalisent une Descente de gradient : en comparant les sorties de chaque unités avec les sorties des unités précédentes, ils ajustent les poids et biais de chaque unités (4) grâce à la dérivée de chaque unité en aval. Ils parlent de Back-propagation, puisque l'erreur E, est diffusée couche par couche, des dernières couches vers les premières. Une fois les poids et les biais des unités remis à jour, un nouveau cycle de Forward propagation peut avoir lieu, qui servira à calculer une erreur E et à nouveau réajuster les différents poids et biais.

Ainsi à chaque itération, les pondérations et les biais de chaque unité se rapprocheront petit à petit d'une valeur idéale. La progression vers ces valeurs idéales

se fera selon le learning rate : d'où l'importance de bien le choisir. Si celui-ci est trop grand, les valeurs des unités oscilleront sans jamais atteindre la valeur idéale. Tandis que s'il est trop petit, le nombre d'itérations (ou « epoch ») nécessaires pour atteindre l'erreur minimum tendra vers l'infini.

2.3. Le traitement d'image et les réseaux convolutifs

Les images numériques sont composées de pixels d'intensité et de couleurs variables. On peut donc imaginer que ce sont autant de valeurs numériques qui pourront être traitées par des réseaux de neurones :

Les pixels de l'image pourront être représentés selon un tableau, avec des valeurs de 255 (blanc) à 0 (noir). Voici par exemple le moyen de représenter une image de 3x3 en noir et blanc d'une croix :

255	0	255
0	0	0
255	0	255

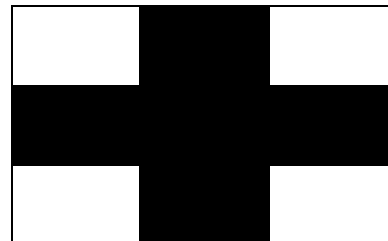


Figure 8, Schéma d'une image en niveau de gris

Par simplification pour le traitement de ces valeurs on représente les valeurs de ces pixels dans une matrice. Dans le cas d'images en couleurs, on rajoutera une dimension de profondeur à cette matrice afin de représenter les valeurs de ces pixels selon les différentes couleurs rouge, vert ou bleu. On obtiendra alors la superposition de trois matrices de dimensions 3840 x 2160. Dès lors, si cette image est fournie à un réseau de

neurone, on fournira en entrée une matrice avec 24 883 200 coefficients de 0 à 255, soit autant d'unités d'entrée.

On comprend à présent la nécessité de diminuer ces ordres de grandeur, si nos unités d'entrées sont connectées à un réseau de neurones, le nombre de connexions devient exponentiel. Il n'est donc pas possible de rajouter des couches et des neurones à l'infini pour augmenter les performances de notre algorithme sans subir des délais de traitement inadaptés.

Une manière de résoudre ce problème a été développée par Fukushima (7) en 1982, en s'inspirant des travaux de Hubel et Wiesel (8) (prix nobel 1981) sur le traitement de l'information dans le système visuel. Il imagine une série de couches de neurones qui effectuent des opérations de convolution et de pooling (regroupement) afin d'extraire progressivement des caractéristiques dans les images d'entrée. La convolution consiste à l'application d'un « filtre » à une région de l'image via un kernel (noyau). Le kernel correspond en une matrice de taille réduite (3x3 à 7x7) qui va se placer sur une zone de l'image de même taille et effectuer la multiplication des pixels par les coefficients de sa matrice. Le résultat permet d'obtenir un nouveau pixel dans une nouvelle image, le processus est répété à l'ensemble des pixels de l'image :

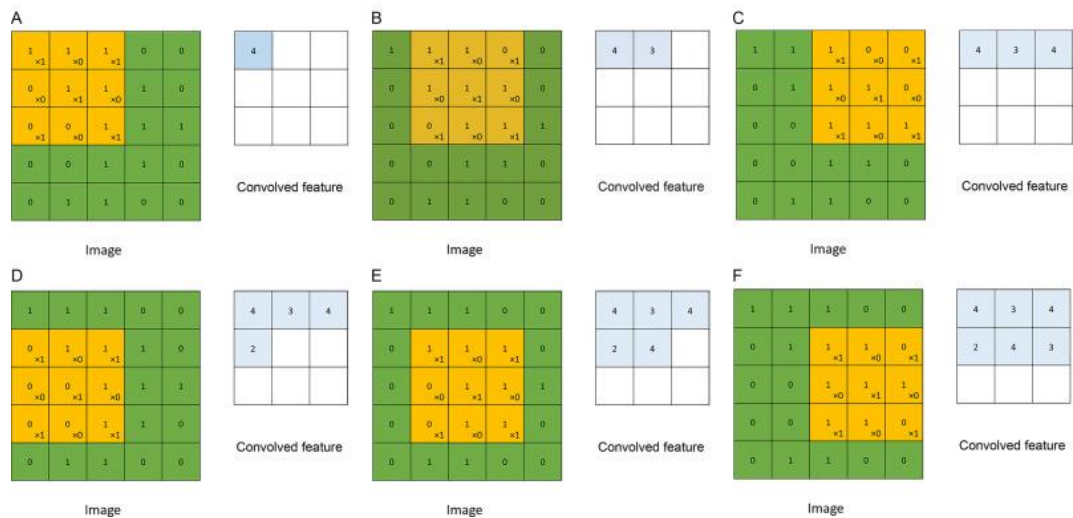


Figure 9, Exemple d'opération de convolution, Ehsan Fathi, Babak Maleki Shoja, Handbook of Statistics, 2018

Cette nouvelle image appelée « carte de caractéristiques » contient des informations qui pourront par la suite être utilisées dans des couches de réseaux de neurones pour en extraire des caractéristiques comme des contours dans l'image etc.

Le pooling, qui est généralement effectué après une opération de convolution, consiste à réduire la taille des données tout en conservant les valeurs importantes de la carte de caractéristique. On appliquera un filtre (en général 2x2) qui parcourra la carte de caractéristiques et remplacera les pixels, soit par une moyenne (average pooling) soit par un maximum (maximum pooling).

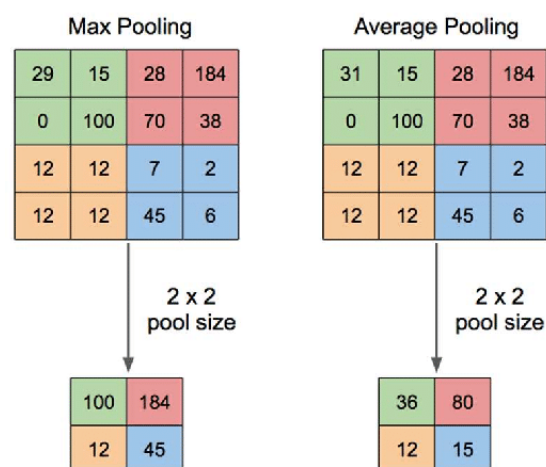


Figure 10, Exemple de pooling, Muhamad Yani et al 2019 J. Phys. : Conf. Ser. 1201 012052

Ces deux types d'opération ont été utilisés pour entraîner des réseaux de neurone à reconnaître des numéros manuscrits (9) ou des caractères malgré l'application de déformations simples ((agrandissement, rotation, changement de contraste/luminosité etc) (10). On distingue ainsi une organisation des réseaux de neurones, plus précisément une organisation des couches de notre réseau. Les différentes couches peuvent comporter plusieurs types de neurones différents. Elles peuvent être, de manière non exhaustive des couches de convolution, de pooling, des couches de neurones récurrents (11), ou des couches totalement connectées (9,10). On parle également d'architecture du réseau(12).

L'architecture d'un réseau de neurones se réfère à la structure et à l'organisation des neurones et des couches qui composent le réseau. Elle est déterminée par la tâche à accomplir et par la nature des données en entrée. (13,14)

Ces architectures permettent de classer plusieurs types de réseaux de neurones, tels que les perceptrons multicouches déjà évoqués (6), les réseaux de neurones convolutifs adaptés notamment pour la reconnaissance et la classification d'image (9,10,15,16), les réseaux de neurones récurrents (11,17), les réseaux de neurones auto-encodeurs (18), les réseaux antagonistes génératifs (19). Ces architectures ont chacune des caractéristiques différentes et sont utilisées pour différentes tâches d'apprentissage et de types de données.

Au sein d'une même classe de réseaux de neurone comme les réseaux convolutifs ou CNNs (Convolutional Neural Network), l'architecture peut être très variable : le réseau LeNet – 5, développé en 1998 par LeCun et al. comporte 7 couches (convolution,

pooling et dense) (15) et permettait de lire et reconnaître l'écriture sur des chèques bancaires. Le réseau AlexNet développé en 2012 par Krizhevsky et al. afin de remporter une compétition de reconnaissance d'image par classification (ImageNet 2012 Challenge) comporte 8 couches de neurones parallèles (20).

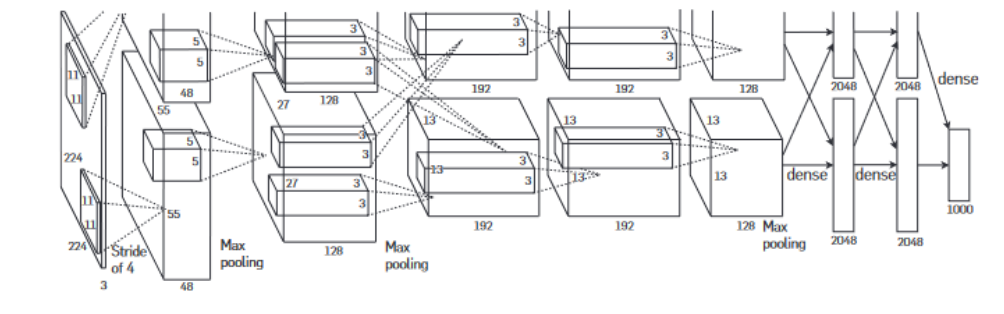


Figure 11, Schéma d'illustration du réseau de Krizhevsky et al. (20)

Enfin dans le domaine médical Ronneberger et al. développent un réseau convolutif pour la segmentation d'image biomédicales (21).

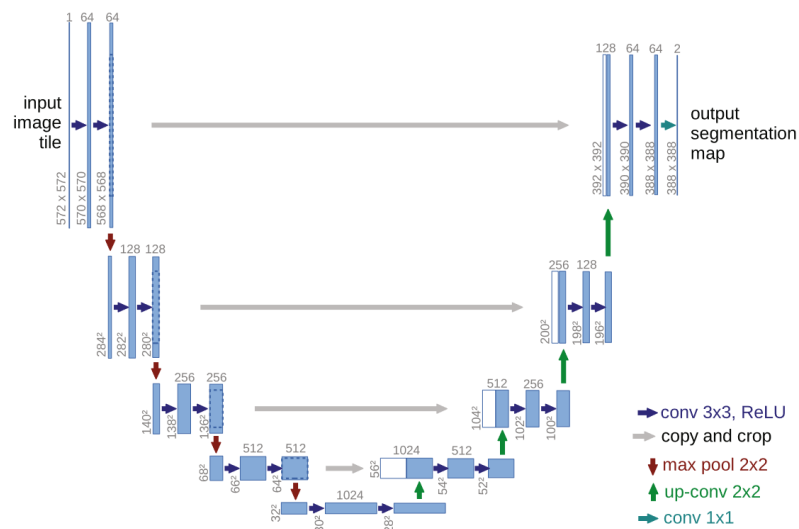


Figure 12, Schéma de l'architecture du réseau U-Net, Ronneberg et al. (21)

La configuration des couches, leur nombre et les traitements spécifiques appliqués à chacune d'entre elles sont des facteurs déterminants dans les performances des différents réseaux neuronaux. Par conséquent, le choix d'un réseau dépendra des

performances recherchées. On observe ainsi l'émergence d'une organisation des réseaux neuronaux qui permet de spécialiser un réseau pour accomplir des tâches spécifiques en fonction des types de connexions entre les neurones. Cela s'applique également aux réseaux neuronaux convolutifs, où l'agencement des couches et les opérations effectuées par celles-ci permettent de spécialiser le réseau dans une tâche particulière plutôt qu'une autre (comme la segmentation ou la classification d'image).

2.4. Les réseaux convolutifs en odontologie

En odontologie, ces réseaux convolutifs sont également le sujet d'articles recensés dans des revues systématiques (22–24). Ils ont des applications variées :

- Ils peuvent être utilisés pour classifier les édentements à partir d'arcades (25) ;

- Ces réseaux convolutifs peuvent également se révéler performants à partir d'imagerie volumétrique (Cone beam) pour la classification de la densité osseuse alvéolaire (26) ; ou encore pour la classification des tissus cancéreux à partir d'image d'endomicroscopie confocale (27).

- A partir d'orthopantomogrammes : ils ont été utilisés avec succès pour estimer l'âge dentaire d'individus grâce au développement des 3^{ème} molaires mandibulaires (28), pour la segmentation (individualisation) des différentes dents (29) ou encore la détection des dents surnuméraires incluses dans la région frontale du maxillaire (30). A partir d'images segmentées de dents, les CNNs se montrent efficaces pour détecter des lésions péri-apicales (31). Enfin en parodontologie, ils s'avèrent pertinent pour la prédiction de la sévérité de l'atteinte parodontale des dents : la détection de dents

compromises au niveau de l'atteinte parodontale(32), et ils réalisent des performances quant à la détection de la perte osseuse parodontale au moins équivalentes (33,34).

-Concernant la détection de lésion carieuse, les réseaux convolutifs vont avoir des applications à partir de radiographies rétro-coronaires, rétro-alvéolaires ou encore à partir de photographies intra-orales ou d'images de transillumination du proche infrarouge (35). Leurs capacités de détection sont alors encourageantes et peuvent surpasser les capacités d'examineurs humains, notamment sur les lésions précoces (35–37).

-Enfin à partir de clichés oraux : ces réseaux ont pu se montrer efficaces notamment sur le diagnostic de la gingivite des incisives centrales supérieures (38), ou encore sous fluorescence pour la segmentation d'inflammation gingivale (39).

Ainsi, l'application des réseaux neuronaux convolutifs en odontologie présente une perspective prometteuse. Cependant, il est important de noter la nécessité de standardiser les images utilisées par ces réseaux. La plupart des études se concentrent sur l'utilisation d'images radiographiques, tandis que les photographies intra-orales sont principalement limitées à un seul appareil avec un protocole défini et un contrôle spécifique de l'éclairage, par exemple par fluorescence, ou sont restreintes à une zone très précise de l'arcade dentaire.

Pourtant, dans le cas des maladies parodontales, l'examen initial se base principalement sur une évaluation visuelle des tissus mous de la bouche, tels que la couleur et le gonflement des gencives. Les photographies intra-orales pourraient constituer un atout considérable pour les praticiens, d'autant plus qu'elles sont

largement disponibles grâce à la généralisation des smartphones et une qualité d'image améliorée.

Toutefois la prise de clichés photographiques intra-oraux est une compétence complexe qui requiert une formation spécifique ainsi qu'un équipement spécialisé. Les dentistes doivent investir dans un équipement adéquat et mettre en place une procédure. Il serait donc bénéfique de développer une méthode qui facilite l'acquisition d'images intra-orales afin d'en extraire certaines caractéristiques permettant d'aider, voire d'automatiser, le diagnostic des symptômes initiaux des maladies parodontales, telles que les gingivites. Voire un outil qui pourrait être utilisé par le patient.

2.5. Les mires colorimétriques

En photographie numérique, une méthode sur lequel travailler pour standardiser sa photographie et notamment la balance des blancs. Il s'agit d'indiquer au logiciel de conversion de l'appareil le type d'éclairage de la scène, afin de retranscrire le plus fidèlement possible les couleurs (ou au contraire de faire varier ce réglage pour jouer avec la créativité). Les appareils modernes disposent majoritairement d'un réglage automatique de la balance des blancs en fonction de l'éclairage, l'utilisateur peut également le renseigner manuellement avant la prise (éclairage néon, ampoule tungstène, flash etc), il se mesure en Kelvin. Cependant, on peut affirmer que ces systèmes automatiques ne sont pas infaillibles, que les utilisateurs ne règlent pas toujours rigoureusement la balance des blancs ou que le sujet est rarement exposé à un illuminant D65 standard.

Ce sont tout autant de facteurs qui peuvent biaiser un diagnostic visuel comme en parodontologie. Si le cortex visuel du praticien est capable de calibrer ce qu'il voit en fonction de l'éclairage, ce n'est pas le cas des réseaux neuronaux. Il est donc nécessaire de leur fournir une colorimétrie fiable, proche de la réalité et que ceux-ci comprennent la conversion des couleurs effectuée par l'appareil numérique. Le changement d'espace colorimétrique peut permettre également de minimiser l'impact de l'environnement sur la prise de vue finale (RVB vers Teinte Saturation Luminosité), bien que ce changement ne permette pas de résoudre le problème dans son intégralité.

C'est dans ce contexte que nous avons eu l'idée d'introduire une mire d'étalonnage colorimétrique (colorchecker). Inspirée de son utilisation courante en photographie, photocopie, impression ou encore pour le réglage des écrans et moniteurs pour standardiser les couleurs. Cette mire colorimétrique offre la possibilité de normaliser les conditions d'éclairage, de couleur et de contraste des images intra-orales. En intégrant cette approche à notre étude, nous cherchons à pallier les limitations actuelles en matière de standardisation des clichés et que l'algorithme exploite au mieux les informations visuelles pour le diagnostic des gingivites.

Une mire d'étalonnage colorimétrique a pu être utilisée en dermatologie humaine avec la transformation par un algorithme, d'images numériques de la peau prises avec des conditions d'expositions, de zoom et d'appareils différents (40). Également à travers un système de télédermoscopie vétérinaire avec une calibration couleur pour évaluer la couleur de la peau du chien, permettant des mesures cliniques rapides et précises en dermatologie vétérinaire (41). Conjointement à un réseau de neurones convolutifs elle permet une classification en temps réel des données

d'imagerie hyper-spectrales obtenues lors de l'endoscopie, ouvrant des perspectives pour améliorer la détection précoce du cancer (42). Enfin au niveau de la cavité orale elle a été utilisée lors de l'examen de la langue en médecine chinoise, pour une correction des couleurs sur les images prises avec des smartphones sur la base de Support Vector Machine (46), et une mire dédiée à ce type d'examen avec un espace colorimétrique adapté à celui de la face antérieure de la langue a été développée (43).

C'est dans ce cadre que nous commençons une étude clinique pour étudier les atouts de la mire colorimétrique dans le diagnostic de l'inflammation gingivale. Nous supposons ainsi que l'introduction de la mire colorimétrique dans le processus de capture d'images intra-orales permette de renforcer l'objectivité et la fiabilité des données visuelles utilisées pour l'analyse par réseau convolutifs. En utilisant cette technique, nous visons à obtenir des images intra-orales cohérentes et comparables, quel que soit l'équipement ou le protocole de capture utilisé. Cette approche pourrait ouvrir la voie à une meilleure exploitation des informations visuelles dans le diagnostic précoce et la prise en charge des maladies parodontales, en offrant aux praticiens un outil précieux et facilement accessible pour évaluer les signes cliniques initiaux et suivre l'évolution des pathologies parodontales.

3. Matériels et méthode

3.1. Critères de sélection

Nous avons inclus dans cette étude des patients consultant au service de parodontologie de l'hôpital universitaire de Toulouse-Rangueil, ainsi que des étudiants externes du service d'odontologie de Toulouse-Rangueil entre Janvier 2022 et Juin 2022.

Les patients remplissant ces critères ont pu être inclus :

-Patients consultant au service de parodontologie de l'hôpital universitaire de Toulouse Rangueil

-Patients dont la prise en charge incluait la prise de cliché photographique lors de l'examen de sondage parodontal

-Sujets âgés de plus de 18 ans au moment de la consultation

-Sujets capables de comprendre le français, de recevoir les informations relatives à l'étude

-Sujets qui ont donné leur consentement oral pour participer au recueil de données

-Les patients sous tutelle, curatelle ou sauvegarde de justice n'ont pu être inclus dans l'études

L'inclusion des étudiants s'est faite sur la base du volontariat en dehors d'un besoin de consultation, selon les critères suivants :

- Sujets acceptant la prise de cliché photographique, et examen clinique des tissus mous de la bouche

- Sujets âgés de plus de 18 ans au moment de la consultation

- Sujets capables de comprendre le français, de recevoir les informations relatives à l'étude

- Sujets qui ont donné leur consentement pour participer à l'étude

Tous les sujets inclus dans cette étude ont donné leur consentement pour la collection de données médicales et leur utilisation.

Cette étude a été conclue en accord avec le code de la santé publique et les lois de bioéthiques qui régissent la recherche en France (45). Cette étude est rétrospective et n'inclus pas la personne humaine. Ainsi, elle n'est pas soumise à un comité de bioéthique mais conforme à la méthodologie de référence (MR-004 de la Commission Nationale de l'Informatique et des Libertés (CNIL) (numéro 2206723 v0)) pour laquelle l'Hôpital Universitaire de Toulouse a signé un engagement de conformité.

3.2. Collecte des données

3.2.1. Collecte des photographies

Pour chaque sujet, nous avons réalisé deux ensembles de photographies intra-buccales haute résolution, en utilisant des écarteurs monobras et des miroirs selon les

besoins. Différents appareils numériques ont été utilisés lors de la capture des images sur les patients, selon la disponibilité du matériel (prêté par les enseignants ou les praticiens du DU de parodontologie). Chacun de ces ensembles comprend au moins trois prises de vue : une vue frontale couvrant l'ensemble des arcades dentaires, une vue latérale gauche, et une vue latérale droite des arcades.

Le premier ensemble de photographies a été réalisé sans mire colorimétrique (Fig. 13). Les photos en exemple (Fig. 12 et 13, de qualité discutables) illustrent la complexité d'obtenir un ensemble de photographie de qualité satisfaisante sur un même patient sans nuire à la séance de soin prévu (voir 4.2.3.1, Biais liés à la complexité de la photographie intra-orales)



Figure 13, Photos intra-orales, de la gauche vers la droite : Vue latérale gauche, Vue frontale, Vue latérale droite

De manière spécifique, le deuxième ensemble de photographie était accompagné d'une mire colorimétrique, maintenue par un assistant, positionnée au même plan que les tissus mous d'intérêt. Cette mire était présente lors des trois prises de vue. Les arcades étaient photographiées en occlusion ou entre-ouvertes en fonction de l'encombrement généré par la mire, assurant ainsi une variabilité dans les conditions de prise de vue (Fig. 13).

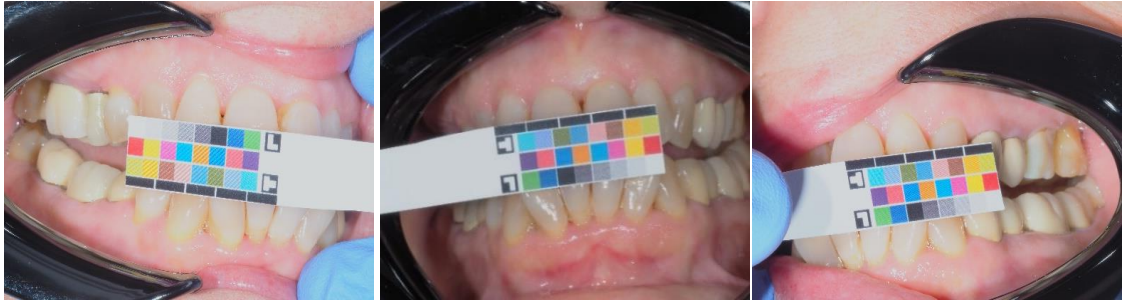


Figure 14, Photos intra-orales avec mire, de la gauche vers la droite : Vue latérale droite, Vue frontale, Vue latérale gauche

Une mire colorimétrique a été spécialement conçue pour l'intra-buccale (Fig. 14) en modifiant la mire de Macbeth (46) afin de répondre aux exigences spécifiques de la photographie. Elle a été conçue en papier couché mat, 350g/m² et mesure 25x8mm.



Figure 15, Photographie intra-orale recadrée d'une mire colorimétrique

À partir de chaque photographie, un investigateur effectue un recadrage précis pour générer plusieurs images dentaires unitaires centrées sur la région cervicale des dents 1 à 5 (correspondant aux secteurs 1 à 4). Ces images englobent la région parodontale vestibulaire, s'étendant jusqu'aux papilles gingivales mésiales et distales, et sont ainsi qualifiées de photos vestibulo-cervico-centrées. Ceci dans la limite de leur exploitabilité : Lorsque la prise de vue ne permet pas d'exploiter certaines dents, celles-ci ne sont pas incluses (par exemple dans le cas d'une prise de vue frontale, au-delà des canines).

Les images fournies aux deux algorithmes sont donc des photographies unitaires vestibulo-cervico-centrées, toutes issues des divers recadrages réalisés par l'investigateur à partir des photographies intra-buccales. Dans le cas de l'algorithme opérant à partir des mires colorimétriques, les photographies unitaires sont dérivées des photographies intra-buccales comportant également une mire colorimétrique positionnée dans le même plan. L'investigateur procède également à un recadrage spécifique autour de la mire colorimétrique pour inclure cette image de la mire avec les photographies unitaires vestibulo-cervico-centrées correspondantes. Cette démarche vise à exploiter les informations colorimétriques de la mire, avec l'espoir que les couleurs capturées soient interprétées par l'algorithme.

3.2.2. Collecte des données patients

Un questionnaire est renseigné par un investigateur à la suite d'un examen clinique dento-parodontal. Cet examen peut être effectué par le même investigateur ou par un interne ou praticien participant au diplôme universitaire de parodontologie de Toulouse. L'objectif du questionnaire est de recueillir des informations :

- Le numéro d'anonymat du patient
- L'âge du patient
- Le sexe du patient
- Le poids du patient
- La taille du patient
- Antécédents tabagiques du patient (ancien fumeur/actuel fumeur/fumeur passif/non-fumeur)

- La moyenne de cigarette par jour
- Le nombre d'année passée à fumée
- L'antériorité du tabagisme si ancien fumeur (le délai qui sépare la dernière imprégnation tabagique)
- La présence d'une pathologie générale en lien avec une complication parodontale
- L'indice de plaque gingivale selon Loë et Silness(47) pour le parodonte vestibulaire en rapport des dents 1 à 5 des secteurs 1 à 4 (0 : absence de plaque, 1 : film fin le long de la marge gingivale à la sonde, 2 : plaque modérée visible à l'œil nu, 3 : plaque importante avec espaces interdentaires remplis)
- La présence ou absence de tartre sur les faces vestibulaires de ces dents
- L'indice gingivale selon Loë et Silness (48) sur le parodonte vestibulaire en rapport de ces dents (0 : gencive normale, 1 : légère inflammation, 2 :inflammation modérée et saignement au sondage, 3 : inflammation importante, saignements spontanés)
- La présence ou l'absence de pigmentation gingivale sur le parodonte vestibulaire en rapport à dents
- La présence ou l'absence de lésion ou ulcération gingivale sur le parodonte vestibulaire en rapport à ces dents

Ces données patientes seront jointes aux photographies unitaires et formeront deux ensembles de données ou (datasets), avec et sans mire.

3.3. Développement du modèle

Deux algorithmes sont intégrés dans un même notebook Python avec un framework Keras/Tensorflow. L'exécution du code relatif à l'algorithme lié à l'utilisation de la mire est gérée par un paramètre nommé « mire = True ». En revanche, l'attribution de la valeur « False » à ce paramètre déclenche l'exécution de l'algorithme fonctionnant sans mire. (L'intégralité du code est disponible sur plateforme externe Github). Les deux modèles fonctionnent de manière similaire et indépendantes, à partir de datasets distincts : les patients et données associées sont les mêmes, les photographies utilisées sont différentes. Nous n'avons pas réutilisé les photos prises avec une mire pour entraîner l'algorithme fonctionnant sans.

3.3.1. Pre-processing des données

Pour démarrer, nos algorithmes prennent en charge un fichier CSV contenant les données des patients (Tableau 1). Chaque ligne de ce document représente une dent spécifique étudiée chez un patient, accompagnée des informations définies dans la section 9.2.3.

NA	Localisation	Sexe	Age	Poids	Taille	Pathologie	Plaque	Tartre	Gencive	Pigmentation gingivale	Lésion gingivale	Tabac	Moyenne cig/j	Nombre d'années
BA	15	H	72	60	170	0	0	0	1	0	0	Actuel fumeur	10	20
BA	14	H	72	60	170	0	1	0	1	0	0	Actuel fumeur	10	20
BA	13	H	72	60	170	0	0	1	3	0	1	Actuel fumeur	10	20
BA	12	H	72	60	170	0	0	1	0	0	0	Actuel fumeur	10	20

Tableau 1, Données patients

À partir de chaque ligne, chaque algorithme explore parmi le dataset correspondant (avec ou sans mire selon un paramètre dédié) les différentes images disponibles de chaque dent. Ces images sont issues des recadrages développés dans la partie 2.2.1. En conséquence, l'algorithme crée un tableau de données (DataFrame ; DF) : Pour chaque image, une nouvelle ligne est créée, comprenant les données du patient ainsi que les colonnes spécifiant le chemin du fichier image et le chemin du fichier mire le cas échéant. En utilisant les informations collectées à partir des questionnaires des patients, deux DF distincts pour chaque modèle.

Le premier, appelé X, regroupe les features (caractéristiques) associées à ces images, tandis que le second, appelé Y, contient les labels (étiquettes) correspondantes. Le DF X compile les données relatives à chaque image, incluant des éléments tels que l'appartenance à certaines catégories (comme le sexe du patient, le numéro de la dent étudiée, l'indice de plaque, la présence de tartre, les antécédents tabagiques, les pigmentations et les lésions/ulcérations). Et constituera les éléments qui seront fournis

en entrée aux modèles après convolution des images. Le DF Y constitue les données attendues, ce sont les valeurs d'inflammation gingivales considérées comme vraies, qui constituent la « solution » vers laquelle doit tendre nos modèles. Ces informations sont label-encodées pour faciliter le traitement ultérieur :

Il s'agit de transformer les informations multi-catégories en tout autant de catégories binaires. Le sexe du patient n'est plus traité comme masculin ou féminin, mais dans l'exemple d'une femme : 0 à la catégorie homme et 1 à la catégorie femme. La localisation d'une dent n'est plus traitée comme valant 21, mais 1 à la catégorie incisive supérieure centrale gauche et 0 pour toutes les autres catégories de localisations.

Parmi les photographies analysées, il existe un déséquilibre de répartition des différents indices d'inflammation gingivales (voir Tableau 2). Ce déséquilibre nous a poussé à développer des stratégies pour tenter de réduire son impact :

-Premièrement, nous avons réunis toutes les classes d'indices d'inflammation supérieures ou égales à 1. De sorte que ne demeurent pour les modèles, que des photographies présentant une inflammation (légère, modérée ou importante selon Loë et Silness) et des photographies présentant une gencive normale. Les modèles discriminent entre « absence d'inflammation » ou « présence d'inflammation » sans tenir compte du degré de celle-ci.

-Deuxièmement, nous avons réalisé une pondération des classes à partir de leur distribution dans le DF Y. Malgré la fusion des classes d'indices d'inflammation supérieures ou égales à 1, les photographies représentant des gencives normales restent plus nombreuses que celles présentant une inflammation. La pondération des

classes consiste à ajuster la loss des prédictions de l'algorithme de manière inversement proportionnelle à leur distribution dans l'échantillon ; afin de pallier les différences de nombre d'images par classe.

	Avec mire	Sans mire
Nombre de patients	14	
Nombre de photos		
Intra-orales	71	63
Localisation des dents		
11	42	39
12	38	39
13	26	28
14	14	20
15	7	12
21	39	35
22	40	36
23	37	23
24	19	14
25	11	10
31	35	44
32	35	41
33	32	31
34	28	13
35	25	10
41	33	46
42	44	48
43	49	38
44	19	21
45	21	15
Totaux	594	563
Indice d'inflammation		
Aucune inflammation	385	361
Légère inflammation	122	96
Inflammation modérée	67	73
Inflammation importante	20	33
Total	594	563
Aucune inflammation	385	361
Présence d'inflammation	209	202

$p \approx 1$

$p \approx 0,99$

* $p \approx 1$

* $p \approx 0,99$

Tableau 2, Répartition des différents types de clichés, * homogénéité des échantillons entre eux, p-value calculée avec test du khi deux

Un déséquilibre dans la répartition des classes peut engendrer un biais dans les algorithmes. Ce biais pourrait les conduire à prédire systématiquement une inflammation correspondant à la classe sur-représentée dans l'échantillon, afin de réduire la loss de leurs prédictions. Pour illustrer ce biais, considérons une représentation à 80% de la classe d'index gingival 0 : dans ce scénario, les algorithmes sans pondération de classe prédiraient de manière systématique une appartenance à la classe 0, ce qui résulterait en un taux de bonne réponse correspondant à la représentation de la classe majoritaire (80% dans cet exemple). Cela signifierait que l'algorithme ne ferait aucune discrimination.

Les deux ensembles de données (DF X et Y) sont divisés de manière équivalente et aléatoire en ensembles d'entraînement (Train set) et de test (Test set), selon un ratio de 80:20, une augmentation artificielle des données images du Train set est réalisée en utilisant un générateur d'images (Image Data Generator). Ce générateur crée des variations aléatoires telles que des rotations, des zooms, des ajustements de luminosité, des cisaillements, des retournements verticaux et horizontaux, ainsi que des décalages des canaux de couleur. Sur l'ensemble de photos comportant une mire, ces variations aléatoires sont appliquées de manière uniforme aussi bien aux images des mires colorimétriques qu'aux photographies dentaires correspondantes. Ces augmentations permettent d'aider les algorithmes à généraliser leurs prédictions en leur présentant artificiellement davantage d'images différentes lors de leur entraînement(49).

Paramètre Image Data Generator :

rotation_range=24, width_shift_range=0.25, height_shift_range=0.25,
shear_range=0.19, zoom_range=0.2, horizontal_flip=False, vertical_flip=False,
brightness_range=[0.5, 1.5], channel_shift_range=24, fill_mode='nearest',

3.3.2. Data-modelling

Le modèle des algorithmes repose sur l'utilisation d'un CNN de type VGG16. Ce réseau est pré-entraîné en utilisant les poids issus du concours 'Imagenet' (50). L'objectif de ce réseau convolutif est d'extraire les caractéristiques des images. Ces caractéristiques sont ensuite utilisées par des réseaux neuronaux denses pour effectuer la classification de l'indice gingival. Les couches du VGG16 ne sont pas entraînées.

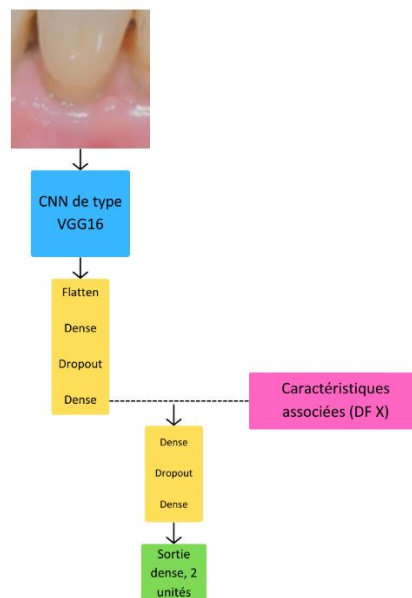


Figure 16, Schéma des algorithmes, une mire est concaténée à l'image en entrée pour les modèles avec mire

À la sortie du VGG16, plusieurs couches supplémentaires sont ajoutées successivement : une couche 'Flatten' (aplati), une couche 'Dense', une couche 'Dropout', puis à nouveau une couche 'Dense' (voir Fig. 16 pour le schéma des algorithmes).

La couche de flatten consiste à aplatir une matrice de neurone, en une couche unidimensionnelle. La couche dense désigne une couche totalement connectée, chaque neurone qui la compose est relié à chacune des unités précédentes et effectue un produit scalaire des unités précédentes avec les pondérations tels que définies en 1.1.2. La couche de dropout (abandon ou extinction) est une extinction aléatoire de certaines unités (produit par 0) afin de limiter que certain neurone ne soit surspécialisés et prédominant sur les autres. Cela permet de favoriser l'apprentissage de l'ensemble des neurones, ceux-ci demeurent robustes et gardent leurs capacités à généraliser leurs prédictions.

Est ensuite concaténé un MLP, qui comporte deux couches denses. Enfin, en sortie des modèles concaténés, d'autres couches sont ajoutées successivement : 'Dense' – 'Dropout' – 'Dense' – 'Dense'. La dernière couche Dense, comporte 2 neurones, correspondants à la prédiction de présenter ou non de l'inflammation.

Les couches du CNN n'étant pas entraînées, elles ne distinguent pas l'inflammation à travers les photographies présentées, mais servent uniquement à extraire des caractéristiques d'intérêt de l'image. C'est la raison pour laquelle nous utilisons un VGG16 pré-entraîné, avec des poids issus du concours Imagenet (« ImageNet Large Scale Visual Recognition Challenge », ou Compétition ImageNet de Reconnaissance Visuelle à Grande Échelle ; qui est compétition logicielle de reconnaissance d'image).

Les réseaux pré-entraînés ont des performances largement suffisantes et permettent une économie de puissance de calcul. L'entraînement des couches du CNN n'est pas pertinent par rapport aux ressources réclamées (28h pour 1 entraînement sur notre matériel).

On comprend dès lors que ce sont les couches denses à la suite du CNN dans nos modèles qui vont déterminer l'inflammation sur les photographies de leur dataset respectif.

Les photographies individuelles, centrées sur la zone cervicale et issues du recadrage des photographies intra-orales prises avec une mire, sont concaténées avec les mires colorimétriques correspondantes. Ces éléments fournissent les entrées pour la partie CNN du modèle. Pour le modèle sans mire, seules les photographies centrées sur la zone cervicale issues du recadrage des photographies intra-orales prises sans mire sont utilisées en entrée pour la partie CNN du modèle.

La partie MLP du modèle utilise en entrée les caractéristiques associées aux images (développées en 2.2.1.) à partir du DataFrame X (sexe du patient, l'indice de plaque, la présence de tartre, les antécédents tabagiques, etc).

Matériel informatique : Les modèles ont été entraînés sur un ordinateur sous système d'exploitation Windows 10.

Processeur : Intel Core i5 6300HQ CPU 2.30 GHz

Carte graphique : NVidia GeForce GTX950M

Ram : 8Go en DDR3

3.3.3. Performance des modèles

Considérant le faible nombre d'échantillon, une validation croisée en 10 plis est effectuée afin d'améliorer les performances des modèles : Les modèles séparent leur Train set en 10 sous-ensembles ou plis égaux, s'entraînant sur 9 d'entre eux et s'évaluant sur le 10^{ème} qui sera appelé Validation set. Lorsque les photographies sont utilisées pendant leur entraînement, elles présentent des modifications d'augmentation artificielle des données (49). Lorsqu'elles sont utilisées pour évaluation (Validation set ou Test set) aucune modification n'est effectuée. Le processus est répété 10 fois afin que chaque pli soit utilisé comme validation au moins une fois.

Enfin durant chacun des 10 plis, chaque modèle s'entraîne et s'évalue sur les mêmes Train set et Validation set lors de 10 epochs. Ces valeurs de 10 epochs par pli, et de 10 plis par entraînement sont des valeurs arbitraires correspondant à un compromis entre les performances des modèles et les ressources nécessaires aux calculs (5heures et 29 minutes nécessaires à l'entraînement du modèle avec mire, 2heures et 40 minutes pour le modèle sans mire).

Suite à chaque entraînement, nous avons mesuré les performances de chaque modèle sur des données neuves : issues du Test set. Les modèles n'ont jamais été entraînés sur ces données, et ne modifient pas les paramètres de leurs couches suite à leur évaluation sur le Test set. Leur entraînement sert donc à généraliser leurs prédictions aux données du Test set.

Leurs performances sont étudiées au moyen de métriques, que sont :
-L'exactitude ou « accuracy », qui est le rapport de prédictions correctes sur l'ensemble des prédictions faites. Cette métrique est inadaptée dans le cas de classes

déséquilibrées (si le data set contient 95% de photographies de gencives saine, le modèle peut atteindre 95% d'accuracy en prédisant toutes les images comme saines), mais utile pour l'entraînement des algorithmes et évaluer leurs performances dans le temps (« over/underfitting » ou sous/surapprentissage (51)).

-La loss, qui est une métrique qui évalue les erreurs faites par chacun des modèles. C'est une moyenne des loss function (tel que décrite en 1.1.2.) de chacune des classifications faites. Cette métrique sera également intéressante pour évaluer l'under ou l'overfitting. Nous avons ici utilisé une fonction « sparse categorical crossentropy ».

-Le F1_score, qui est la moyenne harmonique de la précision et du rappel. La précision est la proportion d'observations correctement classées comme positives parmi toutes les observations prédites comme positives, une précision élevée indique que le modèle a moins tendance à classer à tort des exemples négatifs comme positifs. Le rappel est la proportion d'observations correctement classées comme positives parmi toutes les observations réellement positives. Un rappel élevé indique que le modèle a moins tendance à manquer des exemples positifs réels.

$$Précision = \frac{Vrais\ Positifs + Vrais\ Négatifs}{Vrais\ Positifs + Faux\ Positifs} ; Rappel = \frac{Vrais\ Positifs}{Vrais\ Positifs + Faux\ Négatifs}$$

L'utilisation du F1_score vise à trouver un équilibre entre ces deux métriques, et est plus adaptée pour notre répartition de classes déséquilibrée, car prenant en compte à la fois les Faux Positifs et Faux Négatifs. Il est compris entre 0 et 1, avec 1 étant la valeur à atteindre pour un modèle parfait (52,53).

$$F1_{score} = 2 \frac{Précision * Rappel}{Précision + Rappel}$$

C'est suite aux prédictions faites par les modèles sur les Test set que nous calculerons ces métriques. Enfin l'évolution de ces métriques à travers les différents entraînements nous permettra de les analyser et de les comparer.

4. Résultats

Après 10 entraînements (effectués sur nos trainset), nous obtenons 2 ensembles (avec et sans mire) de 10 modèles différents. Nous traçons les performances de ces modèles dans des graphiques (évalués sur nos testset), afin d'en comparer les tendances :

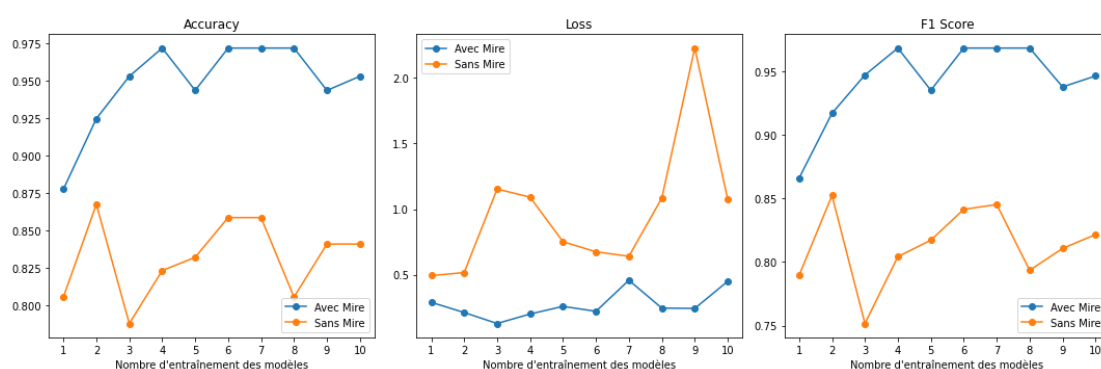


Figure 17, Variation des performances des modèles selon le nombre d'entraînement

Nous pouvons observer que ces performances varient, croissant ou décroissant selon le nombre d'entraînements qu'a subi le modèle étudié. Les courbes correspondantes à l'accuracy et au F1_score semblent suivre les mêmes tendances, ce qui s'explique par la formule du F1_score (2.3.). Concernant la courbe loss, elle permet de synthétiser les erreurs faites par les modèles. On observe ainsi, une diminution des erreurs effectuée par les modèles avec mire jusqu'à 3 entraînements, avant une augmentation de celles-ci. Ce qui peut s'interpréter par un phénomène d'overfitting : les modèles se sur-spécialisent sur les photos d'entraînement et perdent leur capacité à généraliser. Concernant la courbe Loss des modèles sans mire, elle semble indiquer que le nombre d'entraînement des modèles ne diminuent pas les erreurs faites par ceux-ci.

Selon les métriques utilisées, les algorithmes sans mire ne font jamais mieux que les algorithmes avec mires. Aussi nous pouvons à priori supposer une supériorité de ces derniers. Pour le vérifier nous avons décidé de comparer les performances de ces modèles avec l'ajout de métriques supplémentaires (Tableau 2 et 3). Enfin plutôt que d'évaluer graphiquement la supériorité d'une méthode sur l'autre ; nous avons décidé de comparer les moyennes des performances selon chaque métrique mesurée. Conséquemment aux phénomènes d'overfitting ou d'underfitting, les meilleures performances de chaque catégorie ne sont pas regroupées sur un seul modèle.

modèle	Type	nement	d' entrai	Nombre	Accuracy	Loss	F1_score	Sensibilité	Spécificité	VPP	VPN
Mire				1	0,877	0,290	0,866	0,886	0,873	0,775	0,939
				2	0,925	0,212	0,917	0,943	0,915	0,846	0,970
				3	0,953	0,128	0,947	0,943	0,958	0,917	0,971
				4	0,972	0,201	0,968	0,971	0,972	0,944	0,986
				5	0,943	0,259	0,935	0,886	0,972	0,939	0,945
				6	0,972	0,221	0,968	0,971	0,972	0,944	0,986
				7	0,972	0,458	0,968	0,971	0,972	0,944	0,986
				8	0,972	0,246	0,968	0,971	0,972	0,944	0,986
				9	0,943	0,243	0,938	0,971	0,930	0,872	0,985
				10	0,953	0,450	0,946	0,914	0,972	0,941	0,958
Moyenne					0,948	0,271	0,942	0,943	0,951	0,907	0,971

Tableau 3, Métriques des différents modèles avec mire

Sans mire				1	0,805	0,493	0,789	0,732	0,847	0,732	0,847
				2	0,867	0,516	0,852	0,756	0,931	0,861	0,870
				3	0,788	1,153	0,752	0,561	0,917	0,793	0,786
				4	0,823	1,091	0,804	0,707	0,889	0,784	0,842
				5	0,832	0,751	0,817	0,756	0,875	0,775	0,863
				6	0,858	0,674	0,841	0,732	0,931	0,857	0,859
				7	0,858	0,640	0,845	0,780	0,903	0,821	0,878
				8	0,805	1,084	0,794	0,780	0,819	0,711	0,868
				9	0,841	2,227	0,811	0,610	0,972	0,926	0,814
				10	0,841	1,075	0,822	0,707	0,917	0,829	0,846
Moyenne					0,832	0,970	0,813	0,712	0,900	0,809	0,847

Tableau 4, Métriques des différents modèles sans mire

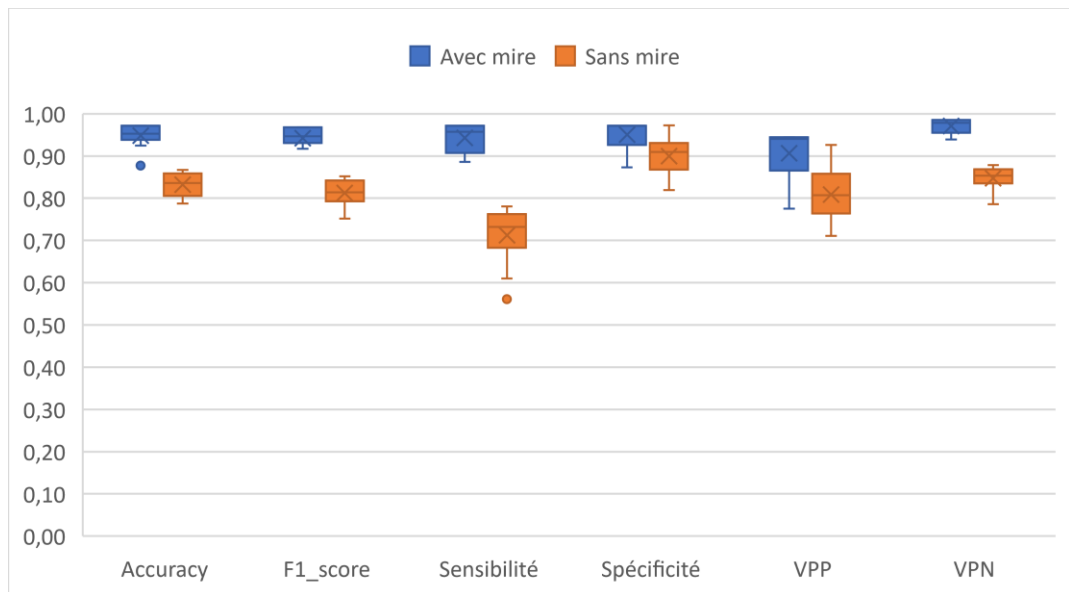


Figure 18, Répartition des performances des différents modèles

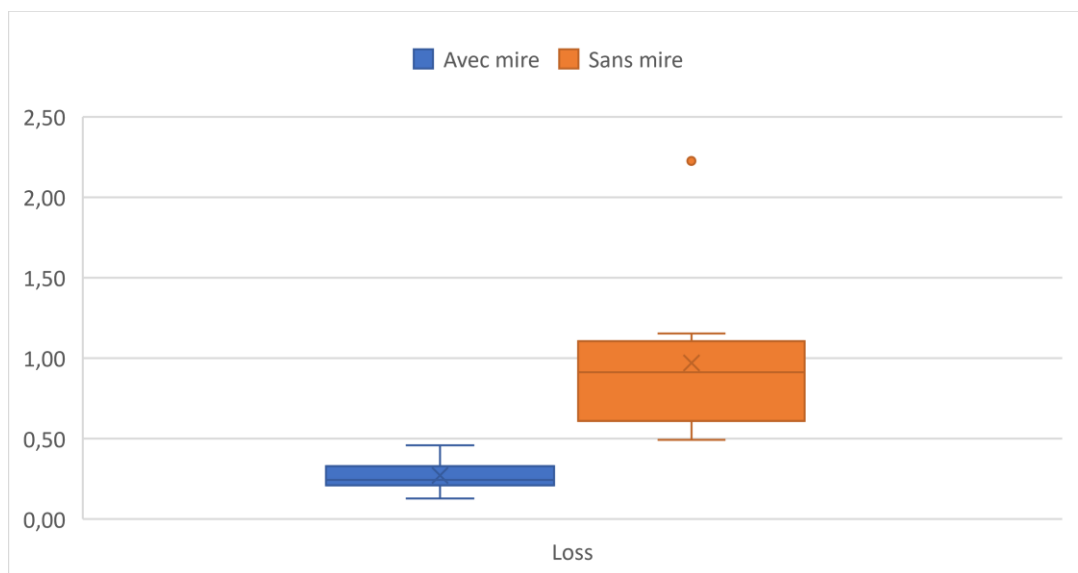


Figure 19, Répartition de la Loss des différents modèles

Pour comparer la répartition des performances de ces modèles selon les différentes métriques mesurées, nous effectuons un test de Student à deux échantillons indépendants. Ce test peut s'effectuer sur des variables indépendantes, ce qui est le cas ici : les performances sont évaluées après prédiction des modèles sur un test set qui n'influence pas son entraînement, celui-ci s'effectuant par validation croisée sur un

train set et un validation set. Les observations ne modifient pas les observations ultérieures.

Nous établissons une hypothèse nulle est qu'il n'y a pas de différence entre les performances de ces modèles. Nous établissons une hypothèse primaire bilatérale, qu'il y a une différence de performances entre les modèles avec mire et sans mire. Tandis que la loss est inférieure dans les modèles avec mire à ceux sans mire. Nous acceptons un risque primaire α (rejeter l'hypothèse nulle alors qu'elle est vraie) de 5%. Les *p-value* mesurées selon tests de Student bilatérales appariés sont les suivantes :

	P-value
Accuracy	<0.01
F1_score	<0.01
Sensibilité	<0.01
Spécificité	≈ 0.011
VPP	<0.01
VPN	<0.01
Loss	<0.01

Tableau 5, Significativité des différences observées

Il est important de rappeler que la comparaison des modèles ne peut se faire qu'au regard de la moyenne de leur performance : leurs performances se mesurant sur leur capacité à classifier l'inflammation, ils peuvent correctement classer les photos en faisant malgré tout davantage d'erreurs (Loss augmentant alors que le F1_score augmente).

5. Analyse et discussion

5.1. Rappel des problématique et objectifs

Nous posons l'hypothèse initiale suivante :

L'utilisation d'une mire colorimétrique améliore les performances de prédiction d'appartenance à une classe d'indice gingivale à partir d'images unitaires centrées sur la limite cervicale de dents par un réseau neurone convolutif.

La comparaison des performances deux modèles n'est jamais absolue. Elle ne peut se faire qu'au regard d'une métrique que nous devons choisir à priori, en l'espèce le F1_score. Le F1_score des modèles avec mire n'étant jamais inférieur à celui d'un modèle sans mire, nous pouvions conclure sans faire de test de Student. Cependant, nous avons décidé d'augmenter les métriques étudiées pour comparer les performances de ces types de modèles sur d'autres aspects. Les métriques choisies sont parmi les plus courantes pour décrire les performances de tests diagnostic. Au regard de celles-ci, les modèles avec mire font en moyenne mieux que les modèles sans mire, avec une p-value inférieure au risque α déterminé à priori.

Notre hypothèse est vérifiée.

5.2. Analyse des biais

Malgré tout, notre étude rencontre un certain nombre de biais qu'il convient de souligner.

5.2.1. Biais du questionnaire

Dans un premier temps, plusieurs biais existent au niveau du recueil de données via le questionnaire patient :

5.2.1.1. Tabagisme des patients

Les patients éprouvent des difficultés à se remémorer avec précision leur historique tabagique. Certains fournissent une date approximative de début de consommation de cigarettes, et leur habitude de tabagisme peut varier au fil du temps, rendant complexe l'établissement d'une mesure représentative. Par exemple, il peut s'agir d'un arrêt du tabagisme suivi de reprises ultérieures ou bien d'une transition vers l'utilisation de la cigarette électronique après avoir cessé de fumer. Dans ce contexte, il est légitime de se questionner sur la pertinence de notre relevé quant au nombre de cigarettes consommées par les patients en tant que représentation fidèle de leur imprégnation tabagique. De plus, il est crucial d'évaluer dans quelle mesure cette variable peut introduire des biais dans les prédictions des algorithmes.

5.2.1.2. Sites inflammatoires

Le nombre restreint de sites évalués soulève des préoccupations. Lors des bilans de sondage effectués au service de Rangueil, les faces vestibulaires des dents sont segmentées en trois sites distincts (mésial, central et distal). Il y a alors une perte d'information à les regrouper en une unique mesure. En se limitant à considérer la face vestibulaire dans son ensemble, on risque de biaiser le diagnostic de l'algorithme. Une lésion inflammatoire pourrait n'être présente que sur l'un de ces sites et l'angle de prise de vue peut partiellement masquer ces sites. Ainsi, une face vestibulaire pourrait être considérée comme inflammatoire, alors que les deux sites visibles sur la photographie

sont sains, tandis le site effectivement inflammatoire peut être partiellement masqué. Cette situation peut entraîner des biais tant dans l'entraînement que dans les prédictions des algorithmes.

5.2.1.3. Présence tartre ou plaque

La quantification de l'indice de plaque et de la présence de tartre est également inconstante. La présence de plaque est notée par un investigateur lors d'un bilan de sondage (patients issus du DU de parodontologie) ou examen clinique (dans le cas des étudiants volontaires du service de Rangueil). Dans le cas des patients du DU, les investigateurs ne notent pas la présence de tartre ni l'indice de plaque (seulement la présence de celle-ci), et ces valeurs sont remplis par un unique investigateur sur le recueil de donnée. De plus dans certains cas, l'investigateur n'a pu participer au bilan de sondage, et doit remplir le recueil de donnée à posteriori à partir de photo, ce qui présente un important biais.

5.2.1.4. Examen clinique effectué par différents dentistes

Outre le biais concernant la présence de tartre, il est important de noter que le bilan de sondage est effectué par différents dentistes pour les patients issus du DU, ce qui peut entraîner des variations dans l'évaluation de la présence de saignement. Aucun score de Kappa n'a été calculé pour évaluer la cohérence inter-individuelle quant à la détection du saignement. De plus, comme mentionné précédemment, les séances de soins pouvaient nécessiter la prise de photographie et le remplissage du recueil de donnée à la suite du bilan de sondage. Introduisant ainsi des biais potentiels concernant la présence de saignement, notamment la spontanéité de ceux-ci.

5.2.2. Problèmes liés à la sélection du patient

L'étude a été confrontée à de nombreuses limitations concernant le recrutement de patients. L'utilisation d'un appareil photo avec un objectif et un flash dédié représente un coût. Bien que ce type d'appareil ait pu être emprunté, leur utilisation a dû se restreindre au service d'odontologie de Rangueil.

Cependant, les créneaux de soins étant limités, il n'est possible d'inclure que des patients dont la prise en charge prévoit un examen clinique du parodonte suivi de prises de clichés intra-oraux. Il est également compliqué d'interrompre un soin prévu pour effectuer des actes non planifiés. Malgré la volonté du patient, les externes peuvent manquer de temps pour assurer les soins déjà prévus. Ce sont les raisons pour lesquelles il fut possible d'inclure uniquement des patients du DU de parodontologie (dont la prise en charge prévoit un bilan de sondage et des clichés intra-oraux) ou encore des étudiants externes non occupés pendant leurs vacances.

5.2.2.1. Sujets non représentatifs

Cependant, il est légitime de se questionner sur la représentativité des sujets examinés par rapport à la population générale. Les patients provenant du DU sont des individus confrontés à des problèmes parodontaux, ce qui signifie que, lors d'un examen initial, ils peuvent présenter un nombre de sites inflammatoires supérieur à la moyenne. Dans le cadre d'un réexamen, même si la présence de poches parodontales persistantes est constatée, le nombre de sites inflammatoires visibles pourrait avoir diminué.

De la même manière, les étudiants du service d'odontologie sont sensibilisés à l'importance de l'hygiène bucco-dentaire et peuvent également présenter un nombre de sites inflammatoires inférieur à la moyenne.

5.2.3. Biais photos

La prise de prise de clichés intra-oraux présente également certains biais :

5.2.3.1. Complexité photos intra-orales

Comme souligné en introduction, la compétence requise pour la prise de clichés intra-oraux est une véritable aptitude nécessitant un apprentissage. Nous avons observé une variation de qualité entre différentes photographies ; selon que celles-ci soient issus de clichés pris par des investigateurs possédant cette compétence (notamment dentistes du DU) ou qu'ils soient capturées par le même enquêteur à plusieurs semaines ou mois d'intervalle, avec une courbe de progression différente.

Par ailleurs, même si l'étude ne prend pas en compte les dents au-delà des deuxièmes prémolaires, une différence de qualité est perceptible entre les dents antérieures et les prémolaires en particulier. Les dents du bloc incisivo-canin sont ainsi surreprésentées et mieux définies sur les photos, tandis que l'image peut être plus floue et moins nette pour les prémolaires. Enfin aucune stratégie ou protocole n'a été mis en place pour assurer une reproductibilité et similarité des clichés en termes de qualité.

5.2.3.2. Complexité de la mire

La prise de cliché intra-oral, déjà complexe en soi, devient encore plus délicate avec l'ajout de la mire. Cette dernière peut masquer le parodonte des dents étudiées ou projeter une ombre sur celles-ci. Malgré son papier mat, elle peut réfléchir la lumière du scialytique, devenant ininterprétable, et rendant la photographie inutilisable. De plus, elle peut représenter un obstacle à la mise au point de l'appareil. Étant souvent

positionnée sur un plan différent de celui des dents étudiées ; il peut être difficile d'obtenir un cliché qui soit net à la fois sur la mire et sur les dents examinées.

Enfin il est important de souligner que si les prises de vue s'inscrivent généralement dans le cadre d'une séance de soins, qui n'inclue pas l'utilisation d'une mire à priori. L'investigateur a pu ne pas suffisamment insister pour garantir la qualité et la quantité de photographies prise avec une mire ; afin de ne pas perturber davantage le dentiste et le patient. Il existe donc manifestement un biais entre les clichés avec et sans mire, soulevant des questions sur leur comparabilité.

5.2.3.3. Différence de technique (avec et sans miroir)

La prise de clichés intra-oraux exige l'utilisation d'accessoires complémentaires tels que des écarteurs et des miroirs occlusaux. Ceux-ci facilitent la prise de cliché et améliorent leur qualité. Cependant l'utilisation de ces accessoires fut inconstante : parfois absents ou encore mal utilisés. Ce problème concerne à la fois les photographies avec mire et sans mire, dans une proportion non mesurée.

5.2.4. Biais de code

Concernant l'algorithme en soit, un certain nombre de biais découle de son utilisation :

5.2.4.1. Temps d'optimisation

Les deux codes utilisent un CNN issu d'une compétition (imagenet, compétition de reconnaissance d'image), les différentes couches convolutives sont donc non entraînaibles, car celles-ci sont déjà optimisées pour la catégorisation d'image.

Cependant un certain nombre de couches succèdent aux couches convolutives, notamment des couches denses en association avec les entrées patients X. Ces couches sont configurées avec différents hyperparamètres. De même, l'augmentation de données artificielle par ImageDataGenerator est accompagnée de plusieurs hyperparamètres.

Ces hyperparamètres sont différents réglages, internes au code, qui vont faire varier le fonctionnement de l'algorithme : le nombre de neurone par couche, la taille des images, le nombre de plis, les valeurs de modifications appliquées aux images lors de l'entraînement (cisaillement, rotation, luminosité voire paramètres ImageData, 3.3.1.). En variant ces hyperparamètres, on fait varier les performances du modèle, on procède alors à un ajustement (*fine tuning*). C'est la comparaison des performances des modèles associés à ces variations qui permet l'ajustement des hyperparamètres. Cependant le temps nécessaire à l'entraînement est important, et la moindre modification d'un de ces hyperparamètres peut mener à un délai considérable.

Bien que nous ayons utilisé Optuna pour optimiser notre algorithme sans mire, il est fort probable qu'il n'ait pas atteint les associations d'hyperparamètres optimales, entraînant ainsi des valeurs de performance qui pourraient ne pas être totalement représentatives. Nous avons utilisé les mêmes hyperparamètres pour l'algorithme avec mire, il prend plus de temps à s'entraîner et nous n'avons pas pu nous permettre de faire une optimisation de l'un différent de l'autre.

5.2.4.2. Architecture du code

Corollairement à ces biais liés aux hyperparamètres, des biais liés à l'architecture du code existent. L'architecture désigne la manière globale dont est agencée l'algorithme.

Si des contraintes de temps existent quant à l'optimisation des hyperparamètres, elles sont encore plus importantes pour tester différentes architectures. Ainsi le réseau de couches convolutionnelles peuvent être issus de plusieurs réseaux existants, notamment présentés en introduction (LeNet, VGG16, ResNet). Le choix du VGG16 s'est fait à priori, dans un souci de compromis entre performance et simplicité des calculs. Pourtant il n'est pas impossible qu'un autre réseau convolutionnel associé à d'autres hyperparamètres présente des performances supérieures.

De même, la manière dont la machine associe les images et les mires associées présente diverses solutions. Nous avons décidé de concaténer les photographies de mires avec les photos unitaires associées, puis de passer l'ensemble à travers des couches convolutionnelles. Cependant nous avons également éprouvé d'autres méthodes, notamment de passer les photos de mire et de dents à travers deux réseaux convolutionnels distincts puis de réunir leurs sorties respectives à travers un réseau dense avant d'insuffler les données des patients du dataframe X ; ce qui n'a pas semblé présenté d'améliorations.

Par conséquent, nous ne pouvons pas affirmer avec certitude que la méthode de traitement des images par les algorithmes soit la plus optimisée. Il est envisageable d'explorer d'autres méthodes en recourant à différentes architectures.

5.3. Points forts

Notre étude présente l'avantage notable de fournir des photographies intra-orales provenant de différents appareils photo. Ainsi, cela nous permet de mettre à l'épreuve la résilience de nos algorithmes face à des images issues de matériels variés.

5.4. Piste d'améliorations

Pour pallier aux différents biais et poursuivre l'étude dans de meilleures conditions, les choses suivantes pourraient être mis en place.

5.4.1. Remaniement du questionnaire

La qualification de l'inflammation gingivale est éminemment clinique, la détermination d'index gingivale par Loë et Silness peut sembler caduc dans la prise en charge de la maladie parodontale par analyse d'image via machine learning. En effet, nous avons déjà évoqué la nécessité d'évaluation du saignement par sondage parodontale pour l'évaluation de l'indice d'inflammation gingivale. Mais il faut souligner également que ; d'après les dernières recommandations concernant la prise en charge de la maladie parodontale ; le diagnostic de celle-ci repose principalement sur la présence de saignement au sondage, la perte d'attache clinique (évaluée en sondant) voire la radiolyse osseuse. Tout autant de critères inaccessibles à notre algorithme qui analyse exclusivement des images.

Afin de poursuivre l'évaluation de l'intérêt de la mire colorimétrique pour les réseaux convolutifs en parodontologie, il serait judicieux de modifier le critère d'évaluation. On pourrait envisager une évaluation exclusivement visuelle, comme l'appréciation de la présence de plaque ou encore l'évaluation de la récession gingivale.

Ces critères pourraient être évalués par le même examinateur ou par plusieurs examinateurs, avec une évaluation du score de cohérence Kappa pour assurer la fiabilité des résultats. De plus, il serait plus pertinent d'évaluer ces critères sur des sites spécifiques plutôt que sur l'ensemble de la face vestibulaire, en considérant trois sites distincts (mésial, central et distal).

5.4.2. Recrutement des patients

Avec l'objectif d'être plus représentatif de la population générale, l'étude ne devrait pas se limiter aux patients du DU de parodontologie et éviter les photographies issues d'étudiants ou soignants qui sont déjà sensibilisés à l'hygiène bucco-dentaire.

5.4.3. Méthodologie photographie

L'analyse des données par une méthode via réseau convolutif implique des attentes concernant la qualité des images fournies aux modèles. Si on peut imaginer une application laissée à l'usage des patients, avec une qualité moindre des clichés, il faudrait d'abord dans un premier temps s'assurer que les photographies soient comparables lors de l'entraînement des modèles. Dans un second temps seulement, ces modèles pourraient être entraînés sur des images de qualités variables. Ainsi il est nécessaire d'adopter un protocole rigoureux pour la prise des clichés. Dans notre étude, nous remarquons sur de nombreuses photos, différents défauts non pris en charge : salive qui gêne la visibilité du parodonte, utilisation des miroirs inconstante, sujets non centrés, ligne d'occlusion en biais... Ces différents défauts auraient pu être évités si un protocole strict avait été rédigé en amont, afin de préserver une qualité comparable des clichés sans empiéter sur le temps dédié à la séance de soin.

6. Discussion

6.1. La place de la mire dans le diagnostic par réseaux convolutifs

Si les résultats sont encourageants, nous avons pu montrer la quantité de biais qu'il existe dans notre étude. Si nous ne pouvons pas rejeter notre hypothèse, il faut précaution gardée, retenir que cette étude reste modeste et sur un faible nombre d'échantillons. Si la mire semble permettre à nos modèles d'extraire les caractéristiques des photographies étudiées plus précocement au fil des entraînements, et aider à la généralisation du processus de prédiction ; ils n'ont pu le faire que sur une catégorisation binaire. Nous n'avons pas entraîné les modèles à quantifier le degré d'inflammation. Cette analyse est encore trop sommaire. Mais la mire colorimétrique présente de bonnes perspectives pour aider les diagnostics par réseau convolutifs, on peut supposer qu'elle améliore la prédiction sur des zones où les conditions d'éclairage sont limitées et variables comme dans la cavité orale. Malgré tout nous n'avons pu lire qu'un faible nombre d'article relatant son utilisation ; il serait intéressant de voir ses apports sur d'autres réseaux convolutifs, qu'ils soient à visée médicales ou non.

6.2. La place de l'IA en parodontologie

Nous pouvons questionner les apports de notre étude pour la prise en charge parodontale : les modèles développés dans notre étude ne sont pas pertinents pour évaluer la qualité de l'inflammation gingivale (inflammation modérée et importante sous représentées dans notre échantillon de photographies) et se limitent à prédire sa présence. Ils auraient alors un apport très limité dans la prise en charge du patient atteint de maladie parodontale.

Enfin nous pouvons nous questionner sur ce que nous mesurons dans notre étude : si l'indice d'inflammation gingivale semblait initialement pertinent comme mesure pour un CNN, il apparaît dès lors que ce critère est inadapté. En effet le critère d'inflammation gingivale n'est pas uniquement visuel, mais également clinique avec la nécessité de sonder les poches parodontales pour y évaluer le saignement. Cela demeure pourtant encore insuffisant : à la lumière de la dernière Classification des maladies et affections parodontales du World Workshop de 2017 (54), la détermination de l'affection en grade et stade s'effectue notamment par de multiples facteurs. Notamment la mesure de la perte d'attache clinique, l'alvéolyse radiographique, la profondeur de sondage(55,56). Tout autant de facteurs qui ne peuvent être déterminé par une simple analyse visuelle. En somme si notre étude est encourageante, elle ne prend pas en compte la totalité des critères cliniques, qui nécessitent des mesures effectuées par un chirurgien-dentiste.

Dès lors que le diagnostic de la maladie parodontale ne peut reposer uniquement sur un examen visuel, il demeure peu probable d'espérer que des réseaux convolutifs reposant sur l'analyse d'images cliniques en soient capables. Cependant l'application des CNNs en parodontologie n'est à rejeter complètement. Certains facteurs de risques évalués visuellement par un chirurgien-dentiste demeurent envisageable. Ainsi l'analyse et le suivi de la présence d'un biofilm supra-gingivale sur des patients par des outils de réseaux convolutifs pourraient être évalués dans une étude ultérieure.

Enfin il est important de questionner l'impact de ces outils dans la prise en charge de nos patients. Que ce soit lors des premières étapes du traitement parodontales ou des étapes de maintenances parodontales, les actes prodigués par le chirurgien-

dentiste doivent être accompagnés d'interventions éducatives et préventives des facteurs de risques locaux et systémiques (56). Ainsi il est important de se questionner sur la place des outils d'intelligence artificielle dans nos stratégies motivationnelles pour améliorer l'hygiène orale ou diminuer les facteurs de risque systémique (tabagisme, diabète). Les outils de diagnostics et préventions autonomes de l'hygiène orale (vidéos, manuels avec instruments prêtés aux patients) semblent avoir une efficacité limitée, notamment dans le temps (57). Cependant les études sur le sujet sont anciennes (plus de trente ans), et n'incluent pas d'outils numériques tels qu'ils peuvent être répandus dans la société moderne. Les stratégies comportementales de la prévention de la maladie parodontale insistent sur l'efficacité d'entrevues avec le patient menées par des soignants formés à la psychologie motivationnelle (57–59).

Il y apparaît dès lors l'importance pour la motivation du patient de définir ensemble les objectifs, la planification et l'autosurveillance (Goal setting, Planning and Self-monitoring ; GPS) de son hygiène orale. Ainsi des outils numériques et de communication pourraient avoir un impact non négligeable sur ces points. Ceux-ci pourraient être affinés et personnalisables via des outils de deep learning pour une efficacité accrue, il faudra dès lors se demander à quel point l'intrusion de ces outils dans nos vies est acceptable : notification de rappel, analyse des habitudes par le dentiste, algorithme de profil du patient et de recommandations (tels que les plateformes de streaming multi-média et réseaux sociaux).

6.3. La place de l'IA dans les soins

Les outils de deep learning peuvent constituer une aide dans notre prise en charge en parodontologie. Cependant leurs applications dans l'aide au diagnostic demeurent

pour l'instant limité : alors que les CNNs reposent sur une multitude de traitement de l'image, la prise en charge de la maladie parodontale reste éminemment clinique. Il serait possible de développer une application dans le domaine de la prévention de l'hygiène orale, mais les dernières recommandations ne permettent pas d'imaginer en l'état actuel des connaissances un diagnostic sans manipulation au fauteuil. Sans négliger non plus toutes les explications et stratégies motivationnelles à mettre en place auprès du patient, et donc à priori favorisant le contact humain avec un soignant.

Enfin même à considérer que soit développé un outil permettant de mesurer la présence de biofilm oral, la question de l'acceptance de cet outil par le patient et le praticien demeure. Son impact sur la prise en charge parodontal reste à déterminer. Il est important de retenir que l'IA reste un domaine mystérieux pour le grand public. Celle-ci n'étant pas capable d'expliquer le procédé menant à son diagnostic, on peut comprendre les mécanismes à l'origine d'une certaine méfiance vis-à-vis de ces outils.

De même, la démocratisation de tels outils devra accompagner une réflexion éthique et législative sur leur utilisation : en cas de mésusage ou d'erreur diagnostic, qui de l'industriel fabriquant le dispositif médical ou de l'utilisateur devient responsable ? Est-ce que la machine n'induit pas un biais chez le soignant à conformer son diagnostic vers celui de la machine ?

De toutes ces questions émergentes, un équilibre reste à déterminer afin de ne pas manquer une révolution technique sans pour autant sacrifier l'humain sur la route du progrès.

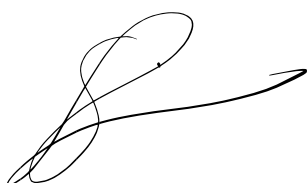
Conclusion

À travers cette thèse nous avons tenté d'évaluer les apports d'une mire colorimétrique sur les performances de réseaux neuronaux convolutifs pour le diagnostic de l'inflammation gingivale. Si nous avons pu montrer que l'analyse d'images intra-orales n'est pas suffisante pour qualifier une inflammation gingivale ; les réseaux neuronaux qui fonctionnent à partir d'images intra-orales combinées à une mire colorimétrique présentent des performances supérieures aux réseaux neuronaux similaires fonctionnant à partir d'images intra-orales sans mires colorimétriques. Cette supériorité s'est vérifiée sur l'ensemble des métriques.

Bien que nous ayons défini tout un ensemble de biais, propres à l'étude ou aux limites de l'investigateur principal, ces résultats demeurent encourageants. Si ce n'est pour diagnostiquer la maladie parodontale, car la méthodologie nécessite une intervention manuelle d'un praticien, cela pourrait éventuellement être une aide vers une détection des premiers signes de cette maladie.

Dans le futur, il serait intéressant d'approfondir l'étude du potentiel et l'application de ces mires colorimétriques dans le diagnostic par réseaux convolutifs. Leur utilisation pourrait s'avérer utile en odontologie et à travers d'autres domaines de la santé, sans s'y limiter, leur utilité dans les réseaux convolutifs en général reste à approfondir.

Vu la Présidente du Jury,
Pr S. Laurencin-Dalieux, 21/05/24



Vu le Directeur de thèse,
Pr P. Monsarrat, 16/05/24



Table des illustrations

Figures

Figure 1, Schéma de neurone artificiel	16
Figure 2, Fonction d'activation	16
Figure 3, Schéma du perceptron, Pablo Caceres,	18
Figure 4, Exemple de frontière de décision linéaire	20
Figure 5, Exemple de frontière de décision non linéaire.....	21
Figure 6, Exemple d'un perceptron multicouche avec 5 couches	21
Figure 7, Fonction d'activation sigmoïde.....	22
Figure 8, Schéma d'une image en niveau de gris	24
Figure 9, Exemple d'opération de convolution	26
Figure 10, Exemple de pooling	26
Figure 11, Schéma d'illustration du réseau de Krizhevsky et al.	28
Figure 12, Schéma de l'architecture du réseau U-Net, Ronneberg et al.....	28
Figure 13, Photos intra-orales sans mire	36
Figure 14, Photos intra-orales avec mire.....	37
Figure 15, Photographie intra-orale recadrée d'une mire colorimétrique	37
Figure 16, Schéma des algorithmes	45
Figure 17, Variation des performances des modèles	51
Figure 18, Répartition des performances des différents modèles	54
Figure 19, Répartition de la Loss des différents modèles	54

Tableaux

Tableau 1, Données patients	41
Tableau 2, Répartition des différents types de clichés.....	43
Tableau 3, Métriques des différents modèles avec mire	53
Tableau 4, Métriques des différents modèles sans mire	53
Tableau 5, Significativité des différences observées	55

Bibliographie

1. World Health Organization. Santé bucco-dentaire [Internet]. Disponible sur <https://www.who.int/fr/news-room/fact-sheets/detail/oral-health>
2. Agence Nationale d'Accréditation et d'Évaluation en Santé (ANAES). PARODONTOPATHIES : DIAGNOSTIC ET TRAITEMENTS [Internet]. 2002. Disponible sur: https://www.has-sante.fr/upload/docs/application/pdf/Parodontopathies_recos.pdf
3. Turing A. I.—Computing machinery and intelligence. *Mind*; 1950;**59**(236):.
4. McCulloch WS, Pitts W. A logical calculus of the ideas immanent in nervous activity. *Bull Math Biophys.* déc 1943;**5**(4):115-33.
5. Rosenblatt F. The perceptron: A probabilistic model for information storage and organization in the brain. *Psychol Rev.* 1958;**65**(6):386-408.
6. Rumelhart DE, Hinton GE, Williams RJ. Learning Internal Representations by Error Propagation. In: *Readings in Cognitive Science*. Elsevier; 1988. p. 399-421.
7. Fukushima K, Miyake S. Neocognitron: A Self-Organizing Neural Network Model for a Mechanism of Visual Pattern Recognition. In: Amari S ichi, Arbib MA, éditeurs. *Competition and Cooperation in Neural Nets*. Berlin, Heidelberg: Springer Berlin Heidelberg; 1982. p. 267-85. (Levin S. *Lecture Notes in Biomathematics*; vol. 45).
8. Hubel DH, Wiesel TN. Receptive fields and functional architecture of monkey striate cortex. *J Physiol.* 1 mars 1968;**195**(1):215-43.
9. LeCun Y, Boser B, Denker JS, Henderson D, Howard RE, Hubbard W, et al. Backpropagation Applied to Handwritten Zip Code Recognition. *Neural Comput.* 1989;**1**(4):541-51.
10. Zhang W, Itoh K, Tanida J, Ichioka Y. Parallel distributed processing model with local space-invariant interconnections and its optical architecture. *Appl Opt.* 1990;**29**(32):4790.
11. Hopfield J. Neural networks and physical systems with emergent collective computational abilities. 1982;**79**:2554-8.
12. Hecht-Nielsen. Theory of the backpropagation neural network. *International Joint Conference on Neural Networks*. Washington, DC, USA: IEEE; 1989;**1**: 593-605.
13. Jain AK, Jianchang Mao, Mohiuddin KM. Artificial neural networks: a tutorial. *Computer.* 1996;**29**(3):31-44.

14. Bengio Y. Learning Deep Architectures for AI. Found Trends® Mach Learn. 2009;**2**(1):1-127.
15. Lecun Y, Bottou L, Bengio Y, Haffner P. Gradient-based learning applied to document recognition. Proc IEEE. 1998;**86**(11):2278-324.
16. Gu J, Wang Z, Kuen J, Ma L, Shahroudy A, Shuai B, et al. Recent advances in convolutional neural networks. Pattern Recognit. 2018;**77**:354-77.
17. Hochreiter S, Schmidhuber J. Long Short-Term Memory. Neural Comput. 1997;**9**(8):1735-80.
18. Hinton GE, Zemel RS. Autoencoders, Minimum Description Length and Helmholtz Free Energy. 1983;3-10.
19. Goodfellow I, Pouget-Abadie J, Mirza M, Xu B, Warde-Farley D, Ozair S, et al. Generative adversarial networks. Commun ACM. 2020;**63**(11):139-44.
20. Krizhevsky A, Sutskever I, Hinton GE. ImageNet classification with deep convolutional neural networks. Commun ACM. 2017;**60**(6):84-90.
21. Ronneberger O, Fischer P, Brox T. U-Net: Convolutional Networks for Biomedical Image Segmentation. arXiv; 2015.
22. Khanagar SB, Al-eHaideb A, Maganur PC, Vishwanathaiah S, Patil S, Baeshen HA, et al. Developments, application, and performance of artificial intelligence in dentistry – A systematic review. J Dent Sci. 2021;**16**(1):508-22.
23. Ahmed N, Abbasi MS, Zuberi F, Qamar W, Halim MSB, Maqsood A, et al. Artificial Intelligence Techniques: Analysis, Application, and Outcome in Dentistry—A Systematic Review. BioMed Res Int. 2021;**2021**:e9751564.
24. Revilla-León M, Gómez-Polo M, Barmak AB, Inam W, Kan JYK, Kois JC, et al. Artificial intelligence models for diagnosing gingivitis and periodontal disease: A systematic review. J Prosthet Dent. 2022.
25. Takahashi T, Nozaki K, Gonda T, Ikebe K. A system for designing removable partial dentures using artificial intelligence. Part 1. Classification of partially edentulous arches using a convolutional neural network. J Prosthodont Res. 2021;**65**(1):115-8.
26. Sorkhabi MM, Saadat Khajeh M. Classification of alveolar bone density using 3-D deep convolutional neural network in the cone-beam CT images: A 6-month clinical study. Measurement. 2019;**148**:106945.
27. Aubreville M, Knipfer C, Oetter N, Jaremenko C, Rodner E, Denzler J, et al. Automatic Classification of Cancerous Tissue in Laserendomicroscopy Images of the Oral Cavity using Deep Learning. Sci Rep. 2017;**7**:11979.

28. De Tobel J, Radesh P, Vandermeulen D, Thevissen PW. An automated technique to stage lower third molar development on panoramic radiographs for age estimation: a pilot study. *J Forensic Odontostomatol*. 2017;**35**(2):42-54.
29. Lee JH, Han SS, Kim YH, Lee C, Kim I. Application of a fully deep convolutional neural network to the automation of tooth segmentation on panoramic radiographs. *Oral Surg Oral Med Oral Pathol Oral Radiol*. 2020;**129**(6):635-42.
30. Kuwada C, Ariji Y, Fukuda M, Kise Y, Fujita H, Katsumata A, et al. Deep learning systems for detecting and classifying the presence of impacted supernumerary teeth in the maxillary incisor region on panoramic radiographs. *Oral Surg Oral Med Oral Pathol Oral Radiol*. 2020;**130**(4):464-9.
31. Ekert T, Krois J, Meinhold L, Elhennawy K, Emara R, Golla T, et al. Deep Learning for the Radiographic Detection of Apical Lesions. *J Endod*. 2019;**45**(7):917-922.e5.
32. Lee JH, Kim D hyung, Jeong SN, Choi SH. Diagnosis and prediction of periodontally compromised teeth using a deep learning-based convolutional neural network algorithm. *J Periodontal Implant Sci*. 2018;**48**(2):114.
33. Krois J, Ekert T, Meinhold L, Golla T, Kharbot B, Wittemeier A, et al. Deep Learning for the Radiographic Detection of Periodontal Bone Loss. *Sci Rep*. 2019;**9**(1):8495.
34. Kim J, Lee HS, Song IS, Jung KH. DeNTNet: Deep Neural Transfer Network for the detection of periodontal bone loss using panoramic dental radiographs. *Sci Rep*. 2019;**9**(1):17615.
35. Mohammad-Rahimi H, Motamedian SR, Rohban MH, Krois J, Uribe SE, Mahmoudinia E, et al. Deep learning for caries detection: A systematic review. *J Dent*. 2022;**122**:104115.
36. Schwendicke F, Elhennawy K, Paris S, Friebertshäuser P, Krois J. Deep learning for caries lesion detection in near-infrared light transillumination images: A pilot study. *J Dent*. 2020;**92**:103260.
37. Cantu AG, Gehrung S, Krois J, Chaurasia A, Rossi JG, Gaudin R, et al. Detecting caries lesions of different radiographic extension on bitewings using deep learning. *J Dent*. 2020;**100**:103425.
38. Alalharith DM, Alharthi HM, Alghamdi WM, Alsenbel YM, Aslam N, Khan IU, et al. A Deep Learning-Based Approach for the Detection of Early Signs of Gingivitis in Orthodontic Patients Using Faster Region-Based Convolutional Neural Networks. *Int J Environ Res Public Health*. 2020;**17**(22):8447.
39. Rana A, Yauney G, Wong LC, Gupta O, Muftu A, Shah P. Automated segmentation of gingival diseases from oral images. In: 2017 IEEE Healthcare Innovations and Point of Care Technologies (HI-POCT). Bethesda, MD: IEEE. 2017:44-7.

40. Vander Haeghen Y, Naeyaert JM. Consistent Cutaneous Imaging With Commercial Digital Cameras. *Arch Dermatol*. 2006;**142**(1):42-6.
41. Cugmas B, Štruc E. Accuracy of an Affordable Smartphone-Based Teledermoscopy System for Color Measurements in Canine Skin. *Sensors*. 2020;**20**(21):6234.
42. Grigoriu A, Yoon J, Bohndiek SE. Deep learning applied to hyperspectral endoscopy for online spectral classification. *Sci Rep*. 2020;**10**:3947.
43. Wang X, Zhang D. A New Tongue Colorchecker Design by Space Representation for Precise Correction. *IEEE J Biomed Health Inform*. 2013;**17**(2):381-91.
44. Hu MC, Lan KC, Fang WC, Huang YC, Ho TJ, Lin CP, et al. Automated tongue diagnosis on the smartphone and its applications. *Comput Methods Programs Biomed*. 2019;**174**:51-64.
45. Gorphe P, Jannin C. Regulatory aspects of prospective and retrospective clinical research in France in 2018. *Eur Ann Otorhinolaryngol Head Neck Dis*. 2019;**136**(2):103-8.
46. Pascale D. RGB coordinates of the Macbeth ColorChecker. 2005.
47. Silness J, Loe H. Periodontal Disease in Pregnancy II. Correlation Between Oral Hygiene and Periodontal Condition. *Acta Odontol Scand*. janv 1964;**22**(1):121-35.
48. Loe H, Silness J. Periodontal Disease in Pregnancy I. Prevalence and Severity. *Acta Odontol Scand*. 1963;**21**(6):533-51.
49. Nalepa J, Marcinkiewicz M, Kawulok M. Data Augmentation for Brain-Tumor Segmentation: A Review. *Front Comput Neurosci*. 2019;**13**.
50. Simonyan K, Zisserman A. Very Deep Convolutional Networks for Large-Scale Image Recognition. *arXiv*; 2015.
51. Ying X. An Overview of Overfitting and its Solutions. *J Phys Conf Ser*. 2019;**1168**:022022.
52. M H, M.N S. A Review on Evaluation Metrics for Data Classification Evaluations. *Int J Data Min Knowl Manag Process*. 2015;**5**(2):01-11.
53. Taha AA, Hanbury A. Metrics for evaluating 3D medical image segmentation: analysis, selection, and tool. *BMC Med Imaging*. 2015;**15**(1):29.
54. Caton JG, Armitage G, Berglundh T, Chapple ILC, Jepsen S, Kornman KS, et al. A new classification scheme for periodontal and peri-implant diseases and conditions – Introduction and key changes from the 1999 classification. *J Clin Periodontol*. 2018;**45**(S20):S1-8.
55. Needleman I, Garcia R, Gkraniias N, Kirkwood KL, Kocher T, Iorio AD, et al. Mean annual attachment, bone level, and tooth loss: A systematic review. *J Clin Periodontol*. 2018;**45**(S20):S112-29.

56. Sanz M, Herrera D, Kebschull M, Chapple I, Jepsen S, Berglundh T, et al. Treatment of stage I–III periodontitis—The EFP S3 level clinical practice guideline. *J Clin Periodontol*. 2020;**47**(S22):4-60.
57. Carra MC, Detzen L, Kitzmann J, Woelber JP, Ramseier CA, Bouchard P. Promoting behavioural changes to improve oral hygiene in patients with periodontal diseases: A systematic review. *J Clin Periodontol*. 2020;**47**(S22):72-89.
58. Tonetti MS, Eickholz P, Loos BG, Papapanou P, van der Velden U, Armitage G, et al. Principles in prevention of periodontal diseases. *J Clin Periodontol*. 2015;**42**(S16):S5-11.
59. Newton JT, Asimakopoulou K. Managing oral hygiene as a risk factor for periodontal disease: a systematic review of psychological approaches to behaviour change for improved plaque control in periodontal management. *J Clin Periodontol*. 2015;**42**(S16):S36-46.

INTÉRÊTS DE LA MIRE COLORIMÉTRIQUE POUR LE DIAGNOSTIC DE L'INFLAMMATION GINGIVALE PAR RÉSEAUX NEURONAUX CONVOLUTIFS

RÉSUMÉ :

Le progrès récent dans le domaine de l'intelligence artificielle a ouvert de nouvelles perspectives en santé, notamment en chirurgie dentaire. Les réseaux neuronaux convolutifs (CNNs), s'inspirant du fonctionnement du cortex visuel, sont capables d'extraire des informations pertinentes à partir d'images. Afin d'explorer les opportunités d'amélioration de la capacité prédictive de ces réseaux, nous nous sommes tournés vers un outil utilisé dans le traitement numérique des photographies, la mire colorimétrique. Notre étude vise ainsi à comparer les performances d'un CNN dans la prédiction de l'inflammation gingivale à partir de photographies intra-orales, en fonction de l'utilisation ou non d'une mire colorimétrique.

ADVANTAGES OF COLORCHECKER FOR DIAGNOSING GINGIVITIS USING CONVOLUTIONAL NEURAL NETWORKS

SUMMARY :

The recent advancements in artificial intelligence have paved the way for various applications in healthcare, notably in dentistry. Convolutional Neural Networks (CNNs), drawing inspiration from the visual cortex, demonstrate the ability to extract meaningful information from presented images. In exploring avenues to enhance the predictive capabilities of these networks, we turned our attention to a tool commonly employed in digital photo processing—the colourchecker. Our study aims to compare the performance of a CNN in predicting gingival inflammation from intra-oral photographs, based on the presence or absence of a colourchecker.

DISCIPLINE ADMINISTRATIVE : Chirurgie dentaire

MOTS-CLÉS : mire colorimétrique, parodonte, inflammation gingivale, deep learning, réseaux convolutifs, CNN, chirurgien- dentiste.

INTITULE ET ADRESSE DE L'UFR

Université Toulouse III-Paul Sabatier

Faculté de Santé - Département d'Odontologie

3 chemin des Maraîchers 31062 Toulouse Cedex 09

DIRECTEURS DE THÈSE : Pr Paul Monsarrat