

UNIVERSITE Mohamed El-Bachir El-Ibrahimi Bordj Bou Arréridj

Faculté Des Mathématique et Informatique

Département De L'informatique



Mémoire en traitement partiel des exigences du diplôme de master

DOMAINE : INFORMATIQUE

Spécialité : réseaux et multimédias

Classification du son contre la toux fondée sur des méthodes explicables :

étude de cas Covid 19

Présenté par :

- M. Hamza Abdessamed
- M.Yaiche Imad Eddine

devant les jurys :

□ Dr. Naili Makhlouf	Président	UMBI Bordj Bou Arreridj
□ Dr. Benabid Sonia	Examineur	UMBI Bordj Bou Arreridj
□ Dr. Attia Abdelouahab	Directeur de Mémoire	UMBI Bordj Bou Arreridj

Année Universitaire : 2023/2024

Abstract

This project aims to develop a computer-aided diagnostic tool for the early detection of COVID-19 through voice analysis. The proposed system operates in two main stages: sound feature extraction and classification. For feature extraction, we utilized Mel-Frequency Cepstral Coefficients (MFCCs), a common technique in voice-based disease detection. The disease classification task employs three supervised machine learning algorithms: Support Vector Machine (SVM), K-Nearest Neighbors (KNN), and Decision Tree (DT). We will evaluate our proposed system using a publicly available dataset (TOS). The performance of the system will be measured using metrics such as accuracy, sensitivity, specificity, F1 score, and Receiver Operating Characteristic (ROC) curves. These metrics provide insights into the system's ability to correctly identify positive and negative cases.

Keywords— : voice, COVID-19, features extraction, classification.

ملخص

يهدف هذا المشروع إلى تطوير أداة تشخيص بمساعدة الكمبيوتر للكشف المبكر عن COVID-19 من خلال تحليل الصوت.

يعمل النظام المقترح على مرحلتين رئيسيتين: استخراج ميزات الصوت والتصنيف. لاستخراج الميزات، استخدمنا معاملات الكيبسروم في

ترددات ميل (MFCCs)، وهي تقنية شائعة في اكتشاف الأمراض المعتمدة على الصوت. تستخدم مهمة تصنيف الأمراض ثلاثة خوارزميات

تعلم آلي خاضعة للإشراف: آلة المتجهات الداعمة (SVM)، أقرب الجيران (KNN)، وشجرة القرار (DT). سنقوم بتقييم النظام المقترح

باستخدام مجموعة بيانات متاحة للجمهور (TOS). سيتم قياس أداء النظام باستخدام مقاييس مثل الدقة، الحساسية، الخصوصية، درجة

F1، ومنحنيات خاصة التشغيل (ROC). توفر هذه المقاييس رؤى حول قدرة النظام على تحديد الحالات الإيجابية والسلبية بشكل صحيح.

الكلمات المفتاحية:--- الكلمات المفتاحية: الصوت ، COVID-19 ، استخراج الميزات ، تعلم آلي ، .

Résumé

Ce projet a pour objectif de développer un outil de diagnostic assisté par ordinateur pour la détection précoce du COVID-19 grâce à l'analyse vocale. Le système proposé fonctionne en deux étapes principales : l'extraction des caractéristiques sonores et la classification. Pour l'extraction des caractéristiques, nous avons utilisé les coefficients cepstraux en fréquences mel (MFCCs), une technique courante dans la détection de maladies basées sur la voix. La tâche de classification des maladies utilise trois algorithmes d'apprentissage supervisé : la machine à vecteurs de support (SVM), les k-plus proches voisins (KNN) et l'arbre de décision (DT). Nous évaluerons notre système proposé à l'aide d'un jeu de données disponible publiquement (TOS). Les performances du système seront mesurées à l'aide de métriques telles que la précision, la sensibilité, la spécificité, le score F1 et les courbes ROC (Receiver Operating Characteristic). Ces métriques fournissent des indications sur la capacité du système à identifier correctement les cas positifs et négatifs.

- **Mots-clés:** voix, COVID-19, extraction de fonctionnalités, classificateurs, simulation.

REMERCIEMENTS

Tout d'abord, nous remercions ALLAH et nous Le remercions de nous avoir permis d'atteindre ce moment et de nous avoir soutenus tout au long de notre vie d'étude. Nous tenons à exprimer notre gratitude à notre superviseur, **M. Attia Abdelouahab**, pour c'est efforts et sa patience qui nous ont permis de mener à bien cette recherche et contribué au succès de ce projet et à l'obtention de ces bons résultats . Nous remercions également notre chere frere **Ibrahim** pour ses efforts, et aussi a toute la famille db

1001 .



DEDICATIONS

We dedicate this modest work as a testimony of affection, of admiration :

je dédie ce travail :

À ma mère : Tu m'as donné la vie, le courage de réussir. Merci pour tes sacrifices.

À mon père : ma force, la personne qui mérite toute mon appréciation et mon respect.

À mes sœurs et mes frère : 'Mohamed, Khalil, Zaki, Amina, Imane'.

À toute ma famille, mes proches et amis, en particulier 'Lapic Prod' , et mes chère amis
Ibrahim , imad , aymen , ayoub , fouzi , habib , youcef , bilal , afif , .

À tous mes enseignants, du premier jour d'école jusqu'à aujourd'hui.

À celui qui nous a aidés et guidés malgré toutes les circonstances.



DEDICATIONS

We dedicate this modest work as a testimony of affection, of admiration :

Avec tout respect et amour je dédie ce travail :

A mes chers parents : pour tous les efforts consentis pour m'assurer
une bonne éducation.

A mes frères : pour tout leur soutien moral et leur amour et
affection.

A tous mes amis notamment : en Souvenir des plus beaux instants
qu'on a passé ensemble.

Aussi a Ayoub , Aymen , Khelifa , Ala et Lopic Prod
Aussi bien à tous ceux qui m'ont aidé..

Yaiche Imad Eddine

La Liste d'abbreviation

AI:	Artificial Intelligence
ANN:	Artificial Neural Networks
AUC:	Area Under Curve
CNN:	Convolutional neural network
Covid:	Corona virus
CPU:	central processing unit
CWT:	Continuous Wavelet Transform
DBN:	Deep belief network
DCT:	Discrete cosine transform
DNN:	Deep neural network
DSS:	Decision Support Systems
DT:	Decision Tree
F0:	Fundamental Frequency
FFT:	Fast Fourier Transform
FN:	False negative
FP:	False positive
GMM:	Gaussian mixture model
GPU:	Graphical processing unit
GRNN:	eneralized regression neural network
HC:	Healthy Controls

HMM:	Hidden Markov model
HNR:	Harmonic to Noise Ratio
KELM:	Kernel-based Extreme Learning Machine
KNN:	K-Nearest Neighbors
LDA:	Linear didcriminant analysis
LPC:	Linear Predictive Coding
LPCC:	Linear Predictive Cepstral Coefficients
MFCC:	Mel Frequency Cepstral Coefficients
ML:	Machine Learning
PCA:	principal component analysis
PLP:	The Perceptual Linear Prediction
RAM:	Random access memory
RAT:	Rapid Antigen Test
RASTA:	Relative Spectral Transform
RBF:	Radial Basis Function
ROC:	Receiver Operating Characteristic
SVM:	Support Vector Machine
TEO:	Teager energy operator
TN:	True Negative
TP:	True Positive
UBM:	Universal background model
WHO:	World Health Organization
ZCR:	Zero-crossing rate

Contenu

Résumé	i
Remerciements	iv
Dédication	v
Introduction Général	1
Introduction	2
Généralités vocale	3
I.1 introduction	4
I.2 Définition et Anatomie du Tractus Vocal	4
I.3 Paramètres vocaux	7
I.3.1 Intensité	8
I.3.2 Hauteur tonale (Pitch)	8
I.3.3 Le Timbre	9
I.4 Trouble de la voix	9
I.4.1 Organique	9
I.4.2 Fonctionnel	10
I.4.3 Psychogène	11
I.4.4 Coronavirus	11
I.5 Covid-19	12
I.6 Conclusion	15
Antelligence Artificielle, Apprentissage Automatique, Extraction de Caractéristiques et Classification	16

II.1 Introduction	17
II.2 Intelligence Artificielle	17
II.3 Machine learning.....	18
II.3.1 Apprentissage automatique supervisé.....	18
II.3.2 Apprentissage automatique non supervisé	19
II.3.3 Apprentissage par renforcement	20
II.4 Extraction de Caractéristiques:	20
II.4.1 Les coefficients cepstraux de fréquence Mel (MFCC)	23
II.4.2 Spectrogramme	26
II.5 Méthodes de Classification	26
II.5.1 Machine à vecteurs de support (SVM).....	28
II.5.2 K-nearest Neighbor(KNN).....	31
II.5.3 Arbre de Décision.....	33
II.6 Conclusion:.....	34
Mise en œuvre et Évaluation	35
III.1 Introduction.....	36
III.2 Détection du Covid-19 à l'aide de la voix	36
III.3 État de l'art.....	38
III.3.1 N. Sharma 2020.....	38
III.3.2 C. Brown et al. 2021	38
III.3.3 V. Espotovic 2021.....	39
III.3.4 R.Islam 2022	39
III.3.5 Rahman.T 2022	39
III.4 Conception du système	40
III.5 Base de données expérimentale	41
III.6 Configuration expérimentale	42
III.6.1 Matériel de développement.....	42
III.6.2 Logiciel de développement.....	42
III.6.3 Configuration de la méthode.....	42
III.8 Métriques expérimentales	48
III.9 Résultats et discussions	49
III.10 Courbes ROC	50

III.11 Conclusion:	51
Conclusion Générale	53
Bibliography	54

liste des tableaux

1	Résultats des femmes avec des caractéristiques MFCC	49
2	Résultats des hommes avec des caractéristiques MFCC	50

liste des figures

1	Anatomie des cordes vocales	5
2	A natomie du pharynx	6
3	Anatomie du larynx	7
4	Intensité de la voix	8
5	Hauteur tonale (Pitch)	8
6	La représentation schématique du coronavirus [1].	11
7	Le schéma des voies de transmission du covid-19 [2].	13
8	Épidémiologie et du covid-19 dans le monde [3].	14
9	Types de Machine Learning.	20
10	Évolution de l'extraction des caractéristiques audio	21
11	Conversion A/N	24
12	Spectrogrammes de toux Covid-19.	27
13	Spectrogrammes de toux Covid-19.	27
14	Concept de SVM [4].	30
15	Concept de k-NN.	32
16	Conception système	40
17	Conception de la configuration de la base de données	41
18	Un signal sonore typique de toux	44
19	Comparaison des sons	44
20	Signal segmenté	45
21	Les blocs de MFCC.	46
22	La banque de filtres Mel	46
23	La courbe ROC pour la classification des femmes en utilisant KNN	51
24	La courbe ROC pour la classification des hommes en utilisant SVM	51

Introduction Général

Introduction

Les coronavirus appartiennent à une vaste famille de virus pouvant déclencher des maladies chez les êtres humains et les animaux. Chez les humains, ces infections vont du simple rhume à des affections plus graves telles que le syndrome respiratoire aigu sévère (SRAS) et le syndrome respiratoire du Moyen-Orient (MERS-CoV). Le SARS-CoV-2, découvert récemment, est responsable de la COVID-19, apparue à la fin de 2019.

La COVID-19 est devenue l'un des problèmes auxquels l'humanité doit faire face. Par conséquent, il est nécessaire de découvrir des moyens rapides et faciles de la détecter, afin qu'elle n'affecte pas le déroulement quotidien de la vie humaine. La détection de la Covid-19 par des signaux sonores est considérée comme une méthode innovante. Dans cette recherche, nous nous concentrons sur les effets auditifs que les symptômes de la COVID-19 peuvent avoir sur les individus, car les changements peuvent être détectés en analysant les enregistrements audio de la toux d'un individu. Les personnes sans symptômes peuvent différer des personnes en bonne santé dans la façon dont elles toussent. Ces différences ne peuvent être déchiffrées en adoptant l'oreille humaine, mais les différences entre elles peuvent être capturées par des techniques d'intelligence artificielle utilisant des méthodologies d'apprentissage automatique et d'apprentissage profond entraînées sur des échantillons de toux de personnes en bonne santé et malades. Cela indique la progression de la maladie et l'efficacité du traitement grâce à un dépistage rapide, un diagnostic à distance précoce et la limitation de sa propagation en formant un classificateur basé sur la prédiction de la COVID-19 à partir de caractéristiques basées sur la toux.

Nous commencerons cette recherche dans :

Premier chapitre : en présentant l'anatomie de la production vocale ainsi que les troubles de la voix, leurs causes et leurs différents types où nous présenterons la maladie COVID-19. Suivi des caractéristiques vocales

Deuxième chapitre : est consacré à la méthode proposée.

Troisième chapitre : dédié aux résultats expérimentaux et à la discussion.

Généralités vocale

I.1 introduction

La voix est l'une des caractéristiques les plus importantes d'un être humain, donc les troubles de la voix peuvent affecter n'importe qui, mais ne sont plus courants chez les personnes qui utilisent beaucoup leur voix dans leur vie quotidienne ou professionnelle. Par exemple, les enseignants ou les orateurs comme les journalistes. Cela soulève la question de savoir si ces maladies peuvent être détectées par l'intelligence artificielle, et la dépendance au son ouvre de nombreux horizons pour le diagnostic de nombreuses maladies qui affectent directement ou indirectement la voix humaine. Dans ce chapitre, en commençant par la première partie, nous essayons d'élucider les mécanismes de production du son et de déterminer leur structure. Dans la deuxième partie, nous explorerons les types de troubles de la parole à travers quelques exemples. Enfin, nous examinons les paramètres vocaux.

I.2 Définition et Anatomie du Tractus Vocal

Notre tractus vocal, bien qu'une partie relativement petite de l'anatomie humaine, est une partie étonnante de notre corps. Le tractus vocal est la cavité présente chez les humains qui sont responsable de la production de sons, sans laquelle nous ne pourrions pas parler. De nombreux scientifiques pensent que notre capacité à communiquer de manière aussi sophistiquée nous distingue nettement des autres mammifères. Alors, comment fonctionne le tractus vocal ? Cependant, sa fonction est fournie par un schéma pour vous aider à comprendre comment nos corps produisent un processus fondamental : la parole humaine. De plus, identifier les principales parties du tractus vocal. Chez les humains, cela signifie la cavité buccale, la cavité nasale, le larynx et le pharynx. Chacun de ces quatre composants est composé de composants plus petits à l'intérieur, et nous avons besoin de toutes ces pièces pour produire du son. La partie la plus facile de notre leçon d'anatomie, et aussi la moins compliquée. La cavité buccale est votre bouche, vos lèvres, vos dents et vos joues. La cavité nasale est l'espace derrière votre nez. Les détails complexes de ces cavités dépassent le cadre de cette leçon mais souvenez-vous simplement de ces bases. Ensuite, parlons de certains des éléments d'anatomie plus compliqués. [11].

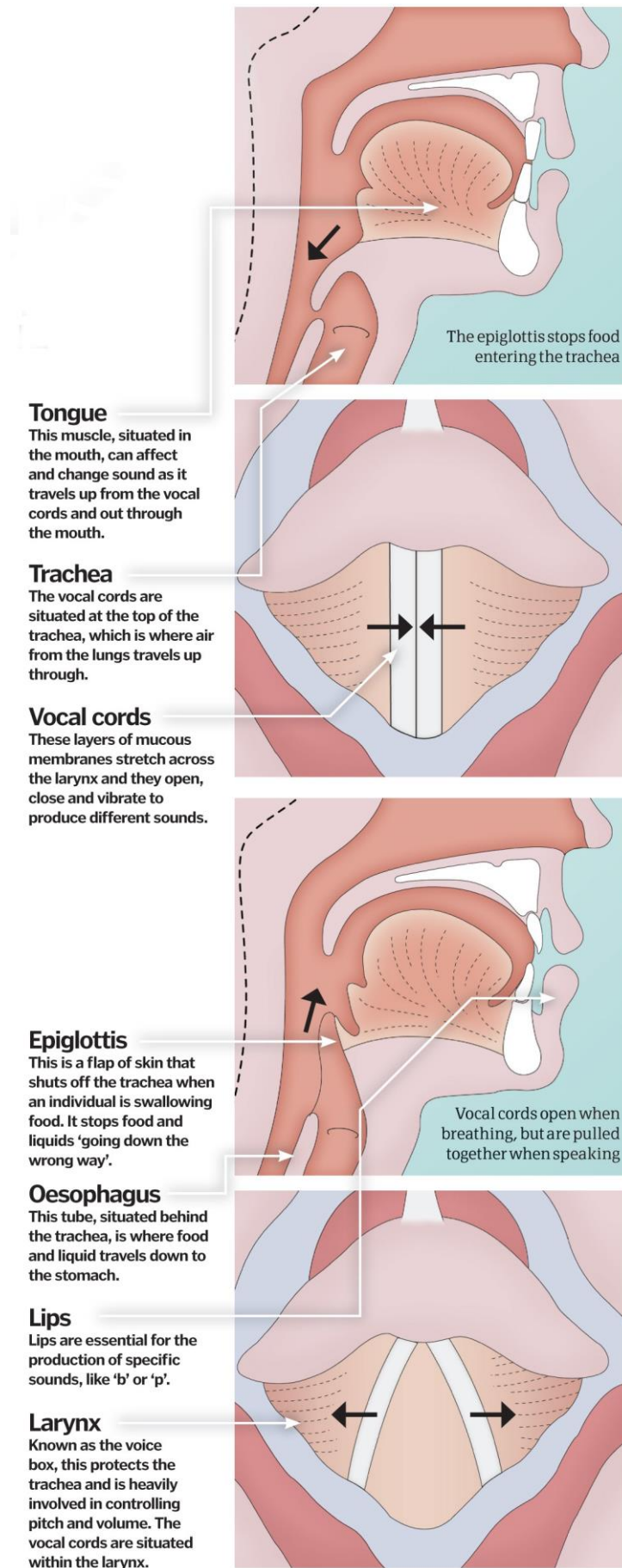


Figure 1: Anatomie des cordes vocales

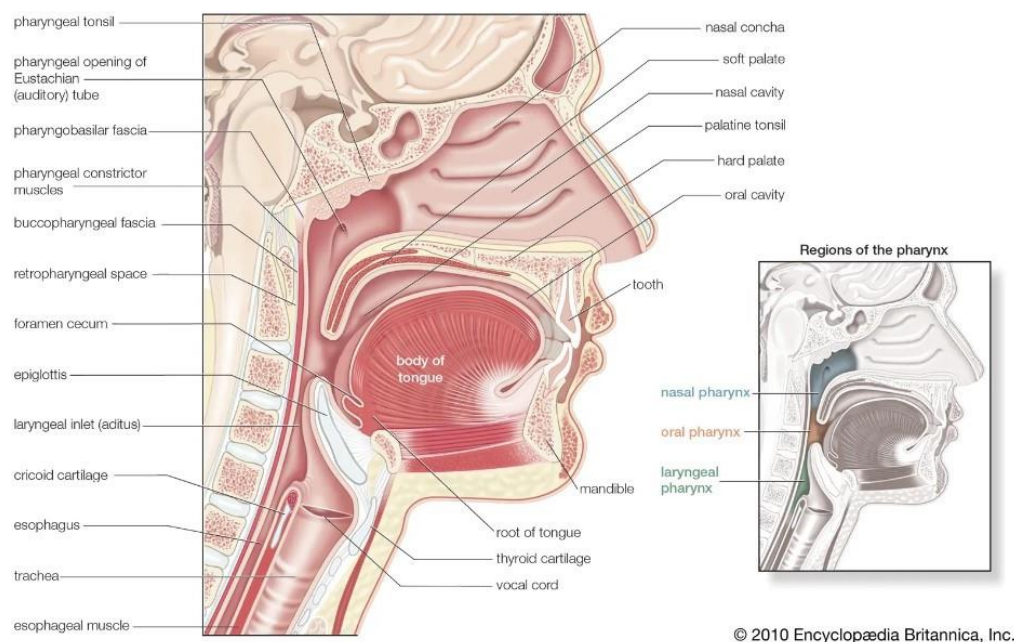
I.2.a Le Pharynx:

Le pharynx est une manière imagée de désigner la partie supérieure de votre gorge. Il s'agit d'un espace creux qui relie les cavités nasales et buccales au larynx, dont nous parlerons plus en détail ensuite. Comme vous le constaterez, le larynx est ce qui produit le son sous forme de vibration, mais vous pouvez envisager le pharynx comme amplifiant ou résonnant ce son. La forme de la gorge, des cavités buccales et nasales modifie les sons de vibration produits par le larynx en sons que les gens comprennent. Le pharynx se divise en deux parties :

L'œsophage, qui descend vers votre estomac.

La trachée, qui descend vers vos poumons.

De ce fait, le pharynx est très important dans la digestion et la respiration [12].



© 2010 Encyclopædia Britannica, Inc.

Figure 2: Anatomie du pharynx

I.2.b Le larynx:

Le larynx, ou comme vous l'avez peut-être entendu, la « boîte vocale », est un organe musculaire qui abrite vos cordes vocales. Il sert également à façonner un passage d'air jusqu'à vos poumons. Le larynx est composé d'os et de cartilage et se situe au sommet de votre trachée (votre trachée-artère), qui descend jusqu'à vos poumons. À l'intérieur du larynx se trouvent également vos cordes vocales, aussi appelées plis vocaux en raison de la façon dont elles peuvent être formées. Ce sont elles qui sont responsables de la

production de votre voix. Lorsque ces lambeaux de tissu vibrent l'un contre l'autre, le son est produit ! Avez-vous déjà eu des difficultés à parler à cause d'une laryngite ? Cela est dû à une irritation de vos cordes vocales, qui rend la parole difficile et douloureuse. [13].

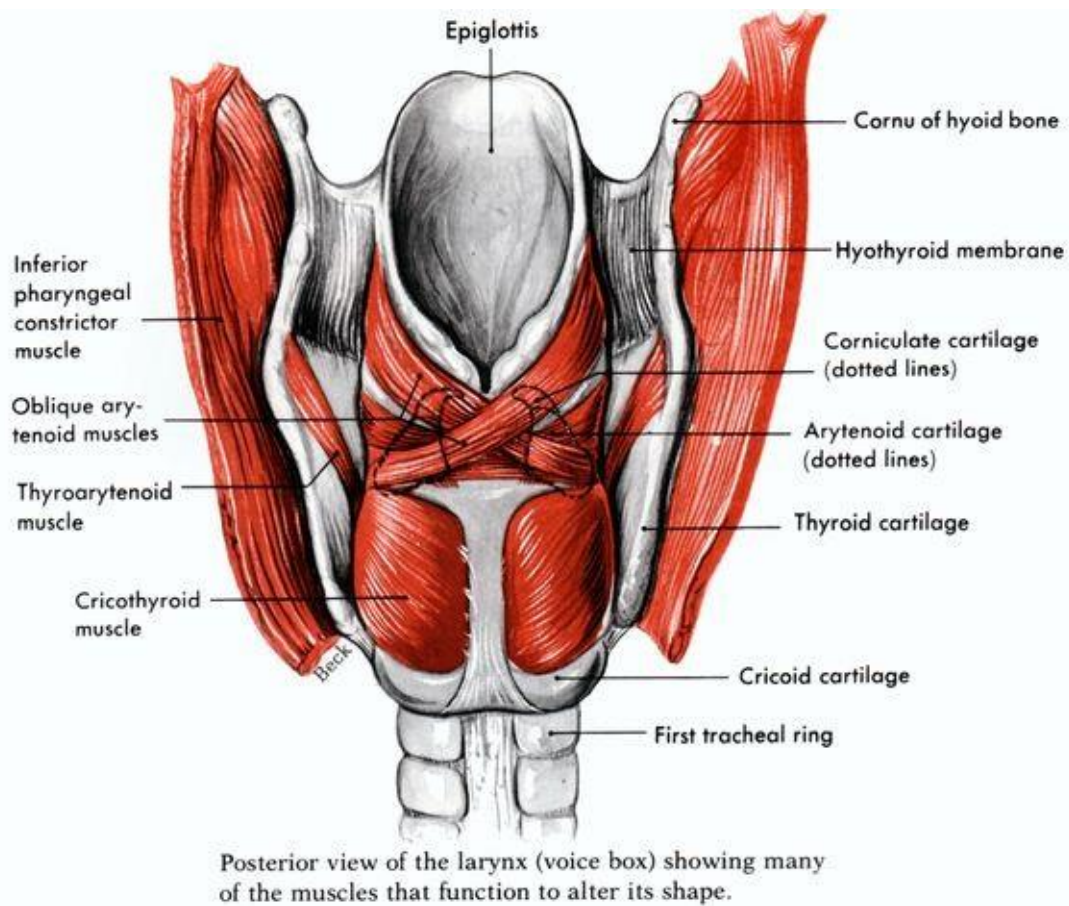


Figure 3: Anatomie du larynx

I.3 Paramètres vocaux

L'analyse acoustique vient en complément de l'évaluation perceptive de la fonction vocale et permet une approche diagnostique plus précise. En outre, cette méthode, qui évalue la parole de manière non invasive, quantifie la fonction vocale. Elle inclut l'examen des paramètres acoustiques suivants:

I.3.1 Intensité

L'intensité est la sensation auditive qui nous permet de distinguer un son fort d'un son faible. On l'appelle aussi parfois "force" du son. Elle est exprimée en décibels (dB) et varie également dans la parole. L'intensité est liée à la pression de l'air sous la glotte (fente située entre les cordes vocales) [14].

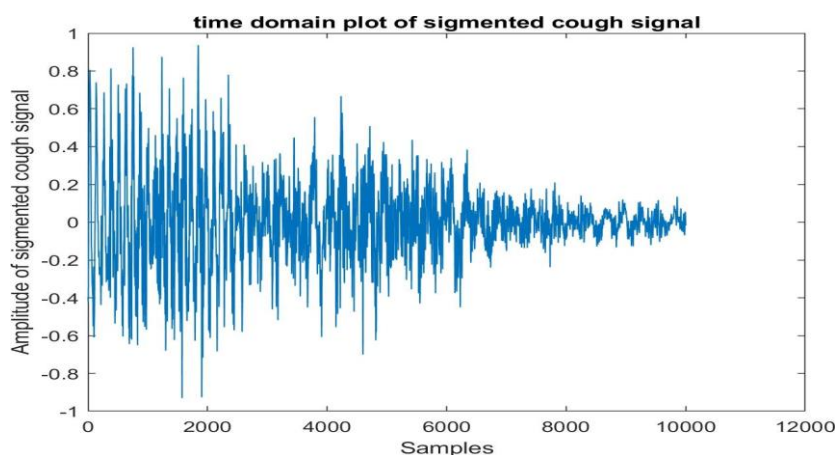


Figure 4: Intensité de la voix

I.3.2 Hauteur tonale (Pitch)

La hauteur tonale, aussi appelée fréquence ou tonalité, est la sensation auditive liée à la fréquence de vibration d'un son. Elle est exprimée en Hertz (Hz). La fréquence fondamentale correspond au nombre de cycles d'ouverture et de fermeture des cordes vocales par seconde, ce qui donne naturellement une voix grave, moyenne ou aiguë. Dans la parole, la hauteur tonale de la voix varie en permanence. C'est l'intonation qui, lorsqu'elle est riche, confère des qualités esthétiques et communicationnelles à la voix[14].

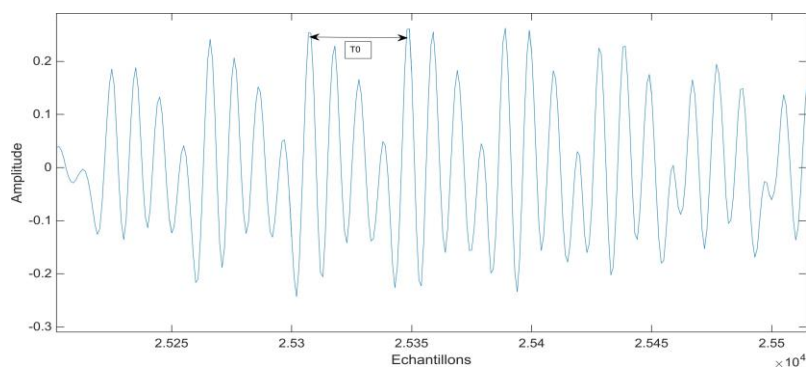


Figure 5: Hauteur tonale (Pitch)

I.3.3 Le Timbre

Le timbre est un indice pour identifier une voix et la reconnaître. La mobilité du palais et de la langue, ainsi que le ton des lèvres et des joues, jouent un rôle important dans sa richesse, tout comme la qualité de la membrane muqueuse qui tapisse les cavités de résonance[14].

I.4 Trouble de la voix

Le langage est un processus complexe. Chez les humains, vous poussez l'air hors des poumons et à travers deux couches de tissu (cordes vocales) dans le larynx (voix). L'air provoque des vibrations dans les cordes vocales qui produisent des sons lorsqu'elles sont touchées. S'il y a un problème avec la hauteur, le ton ou le volume de votre voix, cela signifie que vos cordes vocales ne vibrent pas correctement - on appelle cela un trouble vocal. Selon la cause, cela peut être temporaire ou permanent [15] [16].

I.4.1 Organique

I.4.1.a structurel

Les troubles structurels résultent de certains dommages au larynx (anomalies physiques).

- Cystes.
- Granulome.
- Hémorragie.
- Hyperkératose.
- Laryngite.
- Leucoplasie.
- Nodules .
- Papillome.
- Polypes.
- Traumatisme.
- Croissances diverses.
- **Covid-19..**

I.4.1.b Neurologique

Les troubles neurologiques de la parole sont causés par certains problèmes du système nerveux lorsqu'il interagit avec le larynx. En termes simples, deux nerfs partent du cerveau vers le larynx et contrôlent le mouvement du larynx. Le nerf laryngé récurrent, le plus important des deux nerfs, sort du cerveau, entoure l'aorte, remonte et se fixe au larynx gauche. Comme le nerf laryngé récurrent est situé dans le cou, il est susceptible d'être endommagé lors d'interventions chirurgicales cardiaques, pulmonaires, de la colonne vertébrale et de la thyroïde. Les lésions du nerf provoquent une paralysie (faiblesse) ou une parésie (absence de mouvement complet) des cordes vocales du côté de la lésion. D'autres troubles neurologiques de la voix sont associés à différents types de problèmes du système nerveux central. -Paralysie/Parésie.-Dysphonie spasmodique (Dystonie laryngée).-Tremblement (Tremblement essentiel bénin).-Parkinson.-Problème vocal causé par un autre trouble neurologique (par exemple, maladie de Parkinson, myasthénie grave, SLA/maladie de Lou Gehrig).

I.4.2 Fonctionnel

Les problèmes fonctionnels sont dus à un dysfonctionnement musculaire. Tous les problèmes utiles relèvent de la catégorie de la dysphonie musculaire due à l'anxiété. Les problèmes inhabituels énumérés ici indiquent des types inhabituels de tension musculaire. Rappelez-vous qu'il peut y avoir une maladie où il n'y a pas d'autre choix que de produire un son, mais l'utilisation du son peut entraîner une surcharge, un inconfort ou une fatigue supplémentaire [18].

- Dysphonie due à la tension musculaire.
- Constriction antéro-postérieure.
- Hyperabduction.
- Constriction pharyngée.
- Phonation ventriculaire.
- Inclinaison des plis vocaux.

I.4.3 Psychogène

Les troubles psychogènes existent parce que la voix peut être altérée pour des raisons psychologiques. Dans ce cas, il n'y a pas de raison structurelle pour le trouble de la parole, et il peut y avoir ou non un schéma de tension musculaire. Un composant psychologique ou émotionnel est très courant dans les troubles de la parole, mais les troubles de la parole dus à des troubles psychologiques sont relativement rares. Les deux types les plus courants de troubles psychogènes sont :

- Dysphonie ou aphonie de conversion.
- Pubertophonie (faux-voix mutationnelle).

I.4.4 Coronavirus

Les coronavirus sont des virus à ARN enveloppés. Ces virus à ARN sont divisés en cinq branches et les coronavirus sont très probablement liés à la branche 2.

Les coronavirus peuvent également être classés comme des coronavirus animaux et humains en ce qui concerne leurs cibles hôtes, et les virus de cette sous-famille des Coronavirinae sont divisés en quatre genres, dont le bêta-coronavirus.

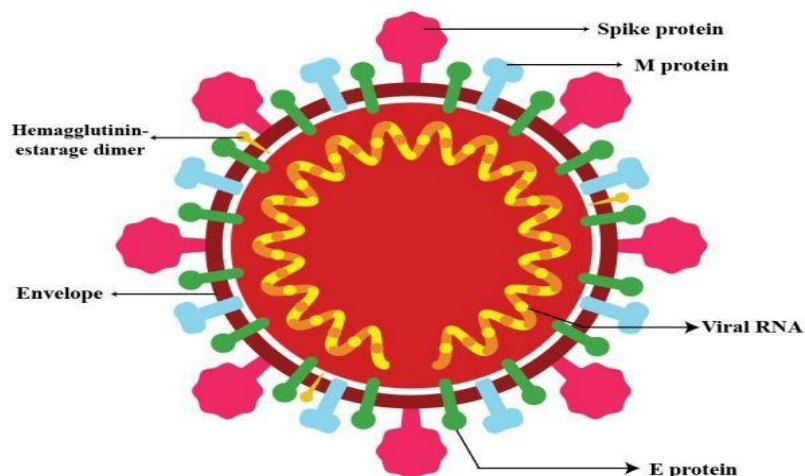


Figure 6: La représentation schématique du coronavirus [1].

Il existe quatre lignées différentes de bêta-coronavirus : 2a, 2b, 2c et 2d. Le SARS-CoV et le SARS-CoV-2 appartiennent à la lignée 2b et le MERS-CoV appartient à la lignée 2c. Le processus d'infection des humains par les animaux par le SARS-CoV, le MERS-CoV et le SARS-CoV-2 est encore inconnu. Comme 96,2 % du génome du SARS-CoV-2 est analogue au génome des chauves-souris [19] [20] [21] [22] [23].

I.5 Covid-19

Le coronavirus est l'une des pathologies qui a causé la mort d'êtres humains au cours des trois dernières années. Dans ce travail, nous nous concentrons sur la recherche d'une solution pour le détecter avec le coût le plus bas possible et le temps de diagnostic le plus rapide. Selon le gouvernement chinois, le premier cas a été découvert le 17 novembre 2019 chez un patient âgé dans la province de Hubei. En outre, quatre hommes et cinq femmes infectés ont été découverts le même mois. Les cas de coronavirus dans le Hubei ont augmenté régulièrement et ont dépassé les deux cents cas, certains avec des décès, d'ici la fin de l'année 2019.

Dans la dernière semaine de décembre 2019, une épidémie de pneumonie étrange caractérisée par quelques symptômes inhabituels tels qu'une toux sèche, de la fatigue, de la fièvre et des symptômes gastro-intestinaux peu fréquents s'est produite sur le marché de gros de fruits de mer de Huanan dans la ville de Wuhan, province de Hubei, en Chine. Le marché a été fermé le premier jour de l'année 2020 après la proclamation d'une alerte épidémiologique par l'autorité voisine.

À ce moment-là, l'Hôpital central de Wuhan a envoyé un échantillon de liquide de lavage broncho-alvéolaire (BALF) issu d'un cas clinique inconnu à une société de séquençage, et après avoir terminé le travail de séquençage, cette institution a informé l'hôpital de la découverte d'un nouveau coronavirus dans le test.

Dans la dernière semaine de l'année précédente, un résultat de test adressé à cet hôpital indiquait un résultat positif pour le SRAS, ce qui a conduit certains médecins de l'hôpital à informer leur personnel et les autorités hospitalières concernées de ce résultat.

Après cela, la Commission de santé municipale de Wuhan a donné avis à certains établissements médicaux sur le processus de traitement de la pneumonie d'origine inconnue. Au cours des premières étapes de l'épidémie, le nombre de cas doublait environ tous les sept jours et demi et la propagation de l'infection était presque entièrement due à la transmission humaine. À ce moment-là, l'Hôpital central de Wuhan a envoyé un échantillon de liquide de lavage broncho-alvéolaire (BALF) issu d'un cas clinique inconnu à une société de séquençage, et après avoir terminé le travail de séquençage, cette institution a informé l'hôpital de la découverte d'un nouveau coronavirus dans le test. Dans la dernière semaine de l'année précédente, un résultat de

test adressé à cet hôpital indiquait un résultat positif pour le SRAS, ce qui a conduit certains médecins de l'hôpital à informer leur personnel et les autorités hospitalières concernées de ce résultat. Après cela, la Commission de santé municipale de Wuhan a donné avis à certains établissements médicaux sur le processus de traitement de la pneumonie d'origine inconnue. Au cours des premières étapes de l'épidémie, le nombre de cas doublait environ tous les sept jours et demi et la propagation de l'infection était presque entièrement due à la transmission humaine .

Début et mi-janvier de l'année en cours, le virus s'est propagé à d'autres provinces chinoises en dehors du Hubei. En Chine, plus d'une centaine de nouveaux cas étaient détectés chaque jour après la mi-janvier, dont deux à Pékin et un à Shenzhen. De nombreuses personnes présentaient les mêmes signes et symptômes, bien que le nombre de tests ne soit pas suffisant pour voir la véritable image. Au cours du même mois, l'OMS a déclaré le coronavirus comme une préoccupation internationale et à ce moment-là, l'épidémie s'était propagée par un facteur de cent à deux cents.

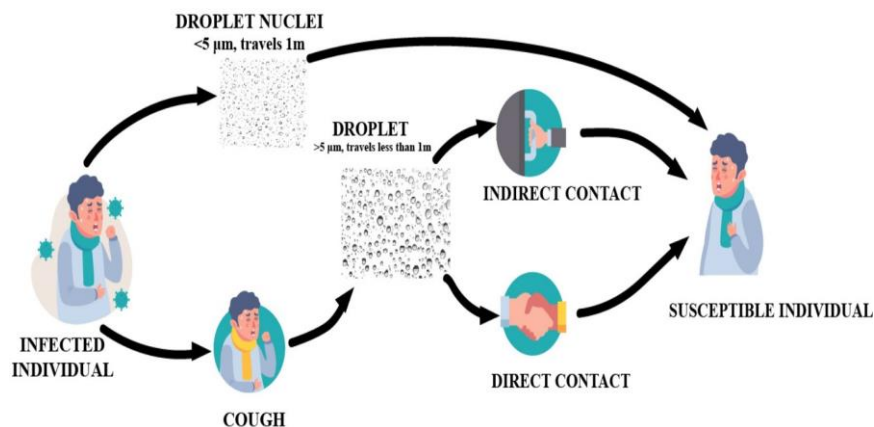


Figure 7: Le schéma des voies de transmission du covid-19 [2].

Le pathogène de l'épidémie a ensuite été identifié comme un nouveau bêta-coronavirus par plusieurs études et a été nommé coronavirus nouveau de 2019 (2019-nCoV), ce qui rappelle aux scientifiques la pandémie passée (SRAS-2003, causée par un autre bêta-coronavirus) qui s'est produite dix-sept ans auparavant. De plus, la maladie s'est propagée dans d'autres pays avant le début de février 2020 et avant la fin de la première semaine de ce mois, le monde a enregistré plus de cinq cents décès impliquant au moins vingt-cinq pays, selon les données de l'OMS. Le dernier jour de janvier, l'Italie a enregistré ses

premiers cas confirmés, qui étaient deux touristes en provenance de Chine. Avant la mi-mars, l'OMS considérait l'Europe comme un centre actif de la pandémie, et le 19 mars,

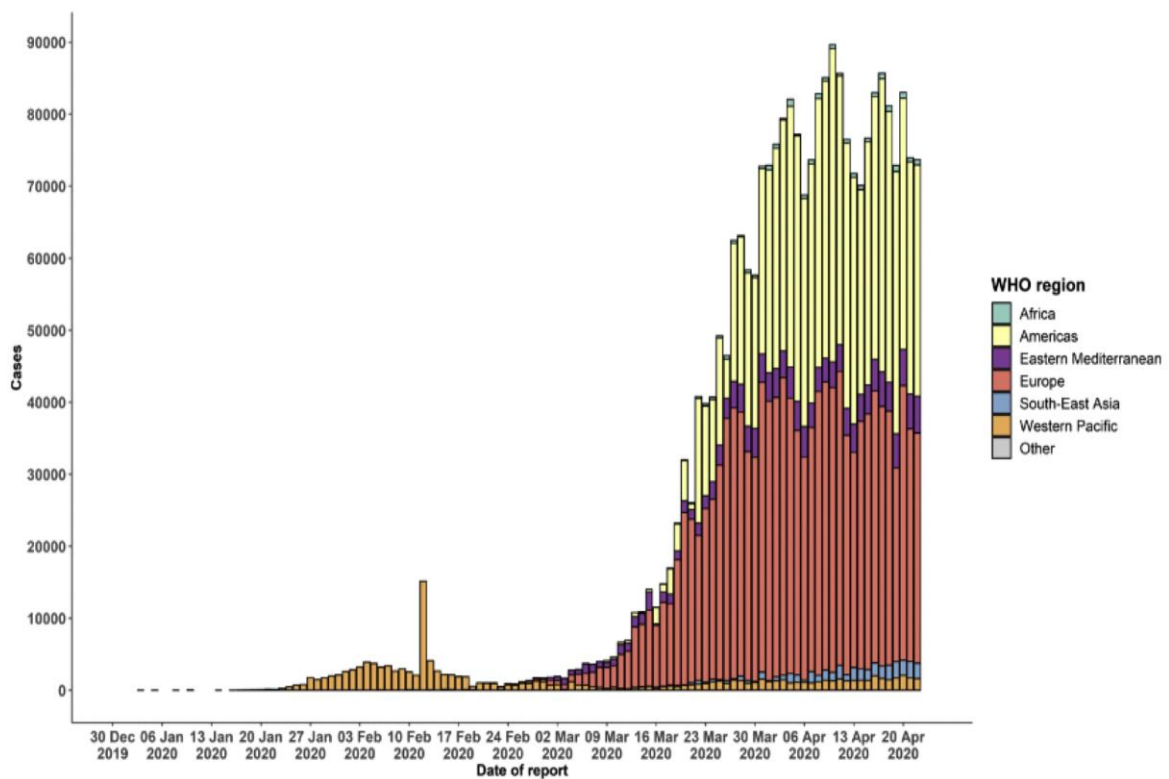
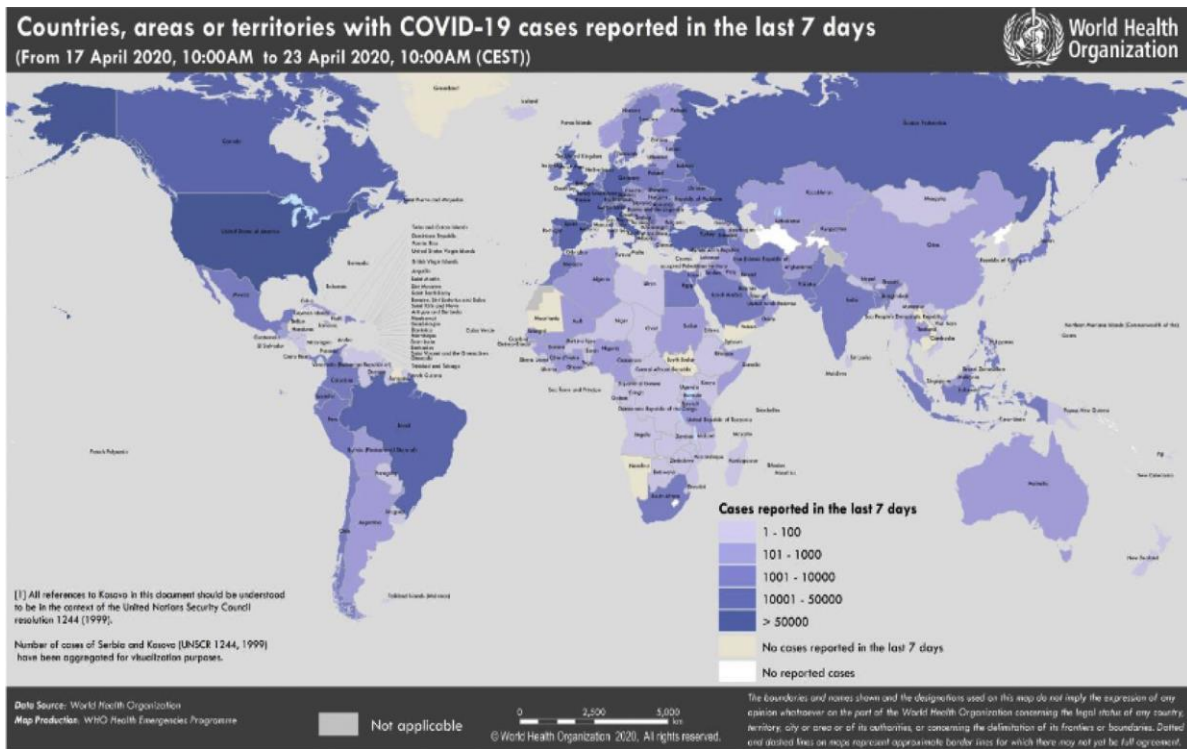


Figure 8: Épidémiologie et du covid-19 dans le monde [3].

l'Italie a dépassé la Chine comme le pays avec le plus grand nombre de décès. Avant la

fin du même mois, les États-Unis ont dépassé la Chine et l'Italie avec le plus grand nombre de cas confirmés dans le monde. La majorité des cas de COVID-19 dans la ville de New York provenaient de voyageurs en provenance d'Europe, plutôt que directement de Chine ou de tout autre pays asiatique, et en France, le retest des échantillons antérieurs de l'année précédente a révélé une personne porteuse du virus. Au Bangladesh, la transmission se poursuit. Dans certaines régions, la deuxième vague de cette attaque virale est observée. Selon les dernières mises à jour, jusqu'au 21 novembre, 58 283 585 cas ont été confirmés, avec 1 383 651 décès et 40 364 210 cas de guérison. Le virus Corona est transmis de plusieurs façons, ce qui a grandement facilité sa propagation dans le monde entier [24][25][26][27] [28] [29].

I.6 Conclusion

Dans ce chapitre, nous avons mentionné la définition et l'anatomie du tractus vocal ainsi que les troubles de la voix et leurs types, puis nous avons discuté de la maladie du covid-19, de son épidémiologie et de ses modes de transmission. Ensuite, nous avons établi une étude des paramètres acoustiques de la voix. Dans le chapitre suivant, nous avons abordé l'extraction des caractéristiques et les classificateurs utilisés dans notre étude pour la détection précoce de la maladie covid-19.

**Intelligence Artificielle, Apprentissage
Automatique, Extraction de
Caractéristiques et Classification.**

II.1 Introduction

Les caractéristiques de détection ont été largement utilisées dans les systèmes de détection médicale car elles sont plus fiables et rapides par rapport aux méthodes conventionnelles. Les troubles de la voix sont des maladies courantes ; la plupart des gens en ont fait l'expérience dans leur vie. Les personnes souffrant de troubles de la voix ne consultent généralement pas un médecin en raison du temps nécessaire et des coûts médicaux élevés. Récemment, les chercheurs ont proposé divers algorithmes d'apprentissage automatique pour la détection rapide des troubles de la voix basée sur l'analyse de la voix humaine. Le reste du chapitre est organisé comme suit : la Section 2 décrit l'intelligence artificielle et l'apprentissage automatique. La Section 3 décrit la méthode d'extraction de caractéristiques et les classificateurs. Enfin, nous concluons ce chapitre par une conclusion [30].

II.2 Intelligence Artificielle

Comme l'apprentissage automatique et le deep learning, l'IA n'est pas "nouvelle", mais elle connaît certainement une sorte de renaissance. La manière dont les gens utilisent le terme change également, ce qui déplaît beaucoup aux traditionalistes. Lorsque Turing a développé son test, le terme intelligence artificielle désignait principalement une technologie capable de reproduire largement l'intelligence humaine. Ainsi, le voyage dans le temps est aujourd'hui une chose lointaine et futuriste pour nous. (Il a fallu 60 ans, mais en 2014, un ordinateur a enfin passé le test de Turing.) Les chatbots comme le "Tay" de Microsoft sont synonymes de nombreuses technologies d'IA. Aujourd'hui, le terme "intelligence artificielle" ou simplement "IA" est souvent utilisé pour désigner n'importe quel type de programme d'apprentissage automatique. À cet égard, il commence à remplacer "big data" et ses adeptes, "analyse avancée" et "analyse prédictive". Pour ceux d'entre vous qui détestent le terme "big data", cela pourrait être une bonne chose. Mais certains préfèrent garder le terme IA comme quelque chose de défini de manière étroite qui reproduit de nombreux aspects de l'intelligence humaine et devient une entité distincte. Nous n'en sommes pas encore là, et nous n'y serons probablement jamais, même si ce n'est probablement pas le pari que vous voulez faire. Il y a un an, le PDG de Facebook, Mark Zuckerberg, a prédit que nous étions à cinq à dix ans du développement d'une "intelligence artificielle capable de comprendre vraiment le sens du contenu"[31].

II.3 Machine learning

Le machine learning désigne tout programme informatique capable d'"apprendre" de manière autonome sans nécessiter une programmation explicite par un humain. Cette idée, et la phrase qui l'exprime, remontent à plusieurs décennies, jusqu'au document fondateur d'Alan Turing de 1950 intitulé "Ordinateurs et Intelligence", qui présentait une section sur sa célèbre "machine d'apprentissage" capable de tromper les humains en leur faisant croire qu'elle est réelle. Aujourd'hui, le terme "apprentissage automatique" est largement utilisé pour englober divers types de programmes rencontrés dans l'analyse de données volumineuses et l'extraction de données. Fondamentalement, les "cervelles" qui alimentent la plupart des programmes prédictifs, y compris les filtres anti-spam, les recommandations de produits et les détecteurs de fraude, sont des algorithmes d'apprentissage automatique.

Les data scientists doivent être familiers avec la différence entre l'apprentissage automatique supervisé et non supervisé, ainsi qu'avec la modélisation ensembliste utilisant une combinaison de techniques de méthode et l'apprentissage semi-supervisé combinant des méthodes supervisées et non supervisées. Il existe trois types d'apprentissage automatique :

II.3.1 Apprentissage automatique supervisé

Dans l'apprentissage supervisé, la machine est enseignée par l'exemple. L'opérateur fournit à l'algorithme d'apprentissage automatique un ensemble de données connu comprenant des entrées et des sorties souhaitées, et l'algorithme doit trouver une méthode pour déterminer comment parvenir à ces entrées et sorties. Bien que l'opérateur connaisse les réponses correctes au problème, l'algorithme identifie des motifs dans les données, apprend des observations et fait des prédictions. L'algorithme fait des prédictions et est corrigé par l'opérateur - et ce processus continue jusqu'à ce que l'algorithme atteigne un haut niveau de précision/de performance. Sous l'ombrelle de l'apprentissage supervisé, on trouve : la Classification, la Régression et la Préviation.

II.3.1.a Classification: Dans les tâches de classification, le programme d'apprentissage automatique doit tirer une conclusion à partir des valeurs observées et déterminer à quelle catégorie appartiennent de nouvelles observations. Par exemple, lors du filtrage des e-mails en "spam" ou "non spam", le programme doit examiner les données observationnelles existantes et filtrer les e-mails en conséquence.

II.3.1.b Régression:

Dans les tâches de régression, le programme d'apprentissage automatique doit estimer - et comprendre - les relations entre les variables. L'analyse de régression se concentre sur une variable dépendante et une série d'autres variables changeantes, ce qui la rend particulièrement utile pour la prédiction et la prévision.

II.3.1.c Prévision:

La prévision est le processus de faire des prédictions sur l'avenir basées sur les données passées et présentes, et est couramment utilisée pour analyser les tendances [33][34].

II.3.2 Apprentissage automatique non supervisé

Ici, l'algorithme d'apprentissage automatique étudie les données pour identifier des motifs. Il n'y a pas de clé de réponse ou d'opérateur humain pour fournir des instructions. Au lieu de cela, la machine détermine les corrélations et les relations en analysant les données disponibles. Dans un processus d'apprentissage non supervisé, l'algorithme d'apprentissage automatique est chargé d'interpréter de grands ensembles de données et de traiter ces données en conséquence. L'algorithme tente d'organiser ces données d'une manière ou d'une autre pour décrire sa structure. Cela peut signifier regrouper les données en clusters ou les organiser de manière plus organisée. À mesure qu'il évalue davantage de données, sa capacité à prendre des décisions sur ces données s'améliore progressivement et devient plus affinée. Sous l'ombrelle de l'apprentissage non supervisé, on trouve :

II.3.2.a Regroupement:

Le regroupement implique de regrouper des ensembles de données similaires (sur la base de critères définis). Il est utile pour segmenter les données en plusieurs groupes et effectuer des analyses sur chaque ensemble de données pour trouver des motifs.

II.3.2.b Réduction de dimension:

La réduction de dimension réduit le nombre de variables prises en compte pour trouver les informations exactes nécessaires [33][34].

II.3.3 Apprentissage par renforcement

L'apprentissage par renforcement se concentre sur des processus d'apprentissage régimentés, où un algorithme d'apprentissage automatique est fourni avec un ensemble d'actions, de paramètres et de valeurs finales. En définissant les règles, l'algorithme d'apprentissage automatique tente ensuite d'explorer différentes options et possibilités, en surveillant et en évaluant chaque résultat pour déterminer lequel est optimal. L'apprentissage par renforcement enseigne à la machine par essais et erreurs. Elle apprend des expériences passées et commence à adapter son approche en réponse à la situation pour obtenir le meilleur résultat possible [33] [34].

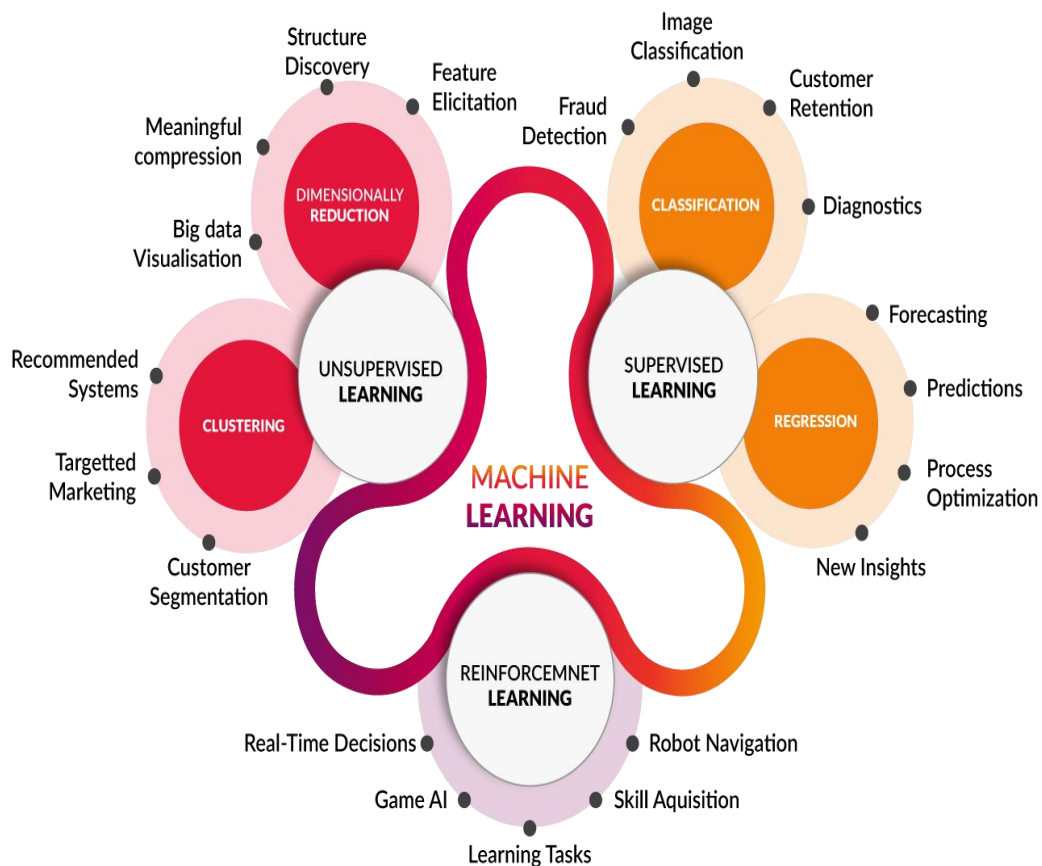


Figure 9: Types de Machine Learning.

II.4 Extraction de Caractéristiques:

L'évolution des caractéristiques audio : En termes simples, l'extraction de caractéristiques met en évidence les propriétés des processus les plus dominantes et discriminantes du signal. Les propriétés d'imitation de caractéristiques appropriées rendent un signal très

compact. L'évolution des signaux audio est décrite dans la figure ci-dessous. Le développement des fonctions audio peut être divisé en domaine temporel, domaine fréquentiel, domaine temporel-fréquentiel conjoint et en caractéristiques de profondeur. Les caractéristiques les plus anciennes et les plus simples sont extraites du domaine temporel.

Les caractéristiques du domaine temporel ont été développées jusqu'à la fin des années 1950.

Jusqu'à présent, les caractéristiques du domaine temporel jouent un rôle important dans l'analyse et la classification audio. Diverses fonctions sont disponibles pour analyser le spectre de fréquence d'un signal audio. Par exemple, la hauteur, les formants, etc., sont développés à partir du domaine fréquentiel et sont utilisés dans diverses applications jusqu'à aujourd'hui. Le développement des caractéristiques du domaine fréquentiel remonte aux années 1950-1960.

À la fin des années 1960, les fréquences temporelles communes sont devenues plus répandues. L'extraction de caractéristiques est nécessaire pour améliorer les performances de classification des signaux biomédicaux. L'objectif de l'extraction de caractéristiques est de trouver l'ensemble de caractéristiques le plus compact et informatif (motifs distinguables) pour améliorer l'efficacité du classificateur. De plus, l'extraction de caractéristiques est utilisée pour extraire des caractéristiques du signal original afin d'obtenir une classification fiable. L'extraction de caractéristiques est la partie la plus importante de la classification des signaux biomédicaux car les performances de classification peuvent être compromises si les caractéristiques ne sont pas sélectionnées correctement[35],[36],[37][38],[39][40],[41][42].



Figure 10: Évolution de l'extraction des caractéristiques audio

Nous mentionnons ici les algorithmes les plus célèbres développés pour l'extraction de caractéristiques :

- Spectrogramme.
- Formant .
- Analyse en ondelettes.
- Le codage prédictif linéaire (LPC).
- La prédiction linéaire perceptuelle (PLP) .
- RASTA - Prédiction linéaire perceptuelle.
- Jitter .
- Shimmer .
- NNE, GNR, HNR et CHNR.
- Le taux de passage par zéro (ZCR).
- Le coefficient cepstral fréquentiel linéaire (LFCC).
- L'opérateur d'énergie de Teager (TEO).
- **Cepstraux de fréquence Mel (MFCC).**

II.4.1 Les coefficients cepstraux de fréquence Mel (MFCC)

L'extraction de représentations paramétriques optimales des signaux acoustiques est une tâche importante dans la création de meilleures performances d'identification des représentations paramétriques. L'efficacité de cette étape est importante pour l'étape suivante car elle affecte son comportement. Les coefficients cepstraux de fréquence Mel (MFCC) sont basés sur la perception auditive humaine et ne peuvent pas percevoir les fréquences au-dessus de 1 kHz. En d'autres termes, les MFCC sont basés sur la variation connue de la bande passante critique de l'oreille humaine avec la fréquence. Il existe deux types de MFCC : le filtre est linéaire à basses fréquences en dessous de 1000 Hz et espacé logarithmiquement au-dessus de 1000 Hz. La gamme de fréquences Mel capture des caractéristiques importantes de la parole. Les coefficients cepstraux de fréquence Mel (MFCC) sont très populaires dans le domaine du traitement de la parole. Ils représentent le signal de manière cepstrale, où les bandes de fréquence sont distribuées selon l'échelle Mel, au lieu de l'approche espacée linéairement. Pour extraire les MFCC à partir d'une trame, on calcule les coefficients cepstraux de fréquence Mel (MFCC) à partir du spectre d'une petite fenêtre du signal vocal obtenu par transformée de Fourier rapide (FFT) du signal. On obtient ainsi une approximation de la perception auditive humaine des fréquences sonores et du spectre, soumise à une série de filtres triangulaires uniformément répartis sur la plage de fréquences Mel. Ensuite, une transformée en cosinus discrète est appliquée à la sortie du filtre, et les coefficients sont déterminés dans les fréquences Mel [43].

II.4.1.a Conversion A/N:

À cette étape, nous convertirons notre signal audio du format analogique au format numérique avec une fréquence d'échantillonnage.

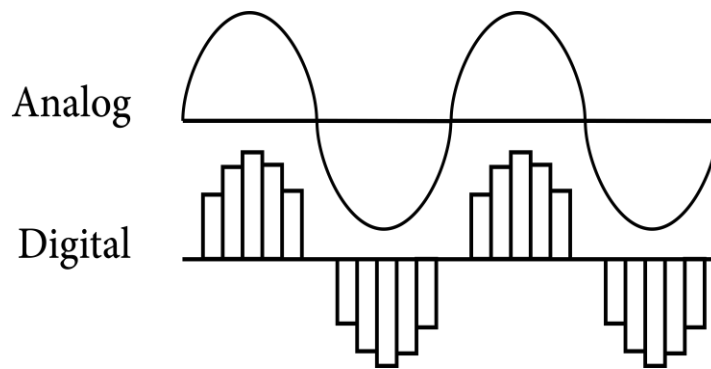


Figure 11: Conversion A/N

II.4.1.b Préaccentuation

La préaccentuation fait référence à un processus systématique conçu pour augmenter la magnitude des fréquences dans la bande passante (généralement plus élevées) par rapport aux autres (généralement plus basses) afin d'améliorer le rapport signal sur bruit global. Ainsi, cette étape consiste à faire passer le signal à travers un filtre qui accentue les fréquences plus élevées. Ce processus ajoute de l'énergie au signal de fréquence plus élevée[44].

II.4.1.c Cadre

Le processus de division des échantillons vocaux reçus de l'ADC en petits cadres de longueur contenue dans une plage de 20 à 40 millisecondes. Le signal vocal est divisé en cadres de N échantillons. Les cadres adjacents seront séparés par M ($M < N$) [44].

II.4.1.d Windowing

La technologie MFCC est conçue pour développer des caractéristiques des signaux audio pouvant être utilisées pour capturer les sons de la parole. Mais il y a de nombreux appels dans un signal audio donné, donc nous divisons le signal audio en différents segments. Nous utilisons la fenêtre de Hamming comme forme de fenêtre en regardant le bloc suivant dans la chaîne de traitement de l'extraction des caractéristiques et intégrons toutes les lignes de fréquence les plus proches. La fenêtre de Hamming ressemble à ceci :

$$W(n), 0 \leq n \leq N - 1 \text{ ou}$$

N = est le nombre d'échantillons dans chaque trame

$Y[n]$ = est le signal de sortie

$X(n)$ = est le signal d'entrée

$W(n)$ = la fenêtre de Hamming, alors le résultat du signal de fenêtrage est indiqué ci-dessous [44]:

$$Y[n] = X(n) * W(n) \quad (1)$$

II.4.1.e DFT (Transformée de Fourier Discrète):

Nous allons convertir le signal du domaine temporel au domaine fréquentiel en appliquant la transformée DFT. Pour les signaux audio, l'analyse dans le domaine fréquentiel est plus facile que dans le domaine temporel. FFT est utilisée pour convertir chaque trame de N échantillons du domaine temporel au domaine fréquentiel. La transformée de Fourier est utilisée pour transformer la convolution de l'impulsion glottique $U[n]$ et de la réponse impulsionnelle du tractus vocal $H[n]$ dans le domaine temporel. L'énoncé prend en charge ce qui suit [44] :

$$Y(w) = \text{FFT}[h(t)*X(t)] = H(w)*X(w) \quad (2)$$

II.4.1.f Mel-Filter Bank:

La plage de fréquences dans le spectre FFT est très large et le signal vocal ne suit pas l'échelle linéaire. La réponse en fréquence de magnitude de chaque filtre est de forme triangulaire et égale à l'unité à la fréquence centrale, diminuant linéairement à zéro à la fréquence centrale de deux filtres adjacents. Ensuite, la sortie de chaque filtre est la somme de ses composantes spectrales filtrées. Ensuite, l'équation suivante, comme indiqué ci-dessous, est utilisée pour calculer le Mel pour une fréquence donnée f en Hz [44] :

$$m(f) = 2595 \log_{10} \left(1 + \frac{f}{700} \right) \quad (3)$$

II.4.1.g Énergie delta et spectre delta

Le signal vocal et les trames changent, tout comme la pente d'un formant à ses transitions. Par conséquent, il est nécessaire d'ajouter des caractéristiques liées au changement des caractéristiques cepstrales au fil du temps. 13 caractéristiques delta ou de vitesse (12 caractéristiques cepstrales plus l'énergie), et 39 caractéristiques delta double ou de double accélération sont ajoutées. L'énergie dans une trame pour un signal x dans une fenêtre allant de l'échantillon temporel t_1 à l'échantillon temporel t_2 , est représentée comme suit [44].

$$E = \int_{t_1}^{t_2} X(t) dt \quad (4)$$

II.4.1.h Caractéristiques dynamiques

En plus de ces 13 caractéristiques, la technique MFCC considère également les première et deuxième dérivées des caractéristiques, ce qui donne 26 autres caractéristiques. Les dérivées sont calculées en prenant la différence de ces coefficients entre les échantillons du signal audio et cela aide à comprendre comment se produit la transformation.

Où $X[t]$ = signal, chacune des 13 caractéristiques delta représente le changement entre les trames correspondant à une caractéristique cepstrale ou d'énergie, tandis que chacune des 39 caractéristiques delta double représente le changement entre les trames dans les caractéristiques delta correspondantes [44].

II.4.2 Spectrogramme

Une forme d'onde vocale se compose d'une série d'événements différents qui changent avec le temps. Correspondant à cette propriété variable dans le temps, les propriétés spectrales varient considérablement dans le temps. Enfin, une seule transformée de Fourier ne peut pas capturer cette vitesse des signaux variables dans le temps, et donc des STFT sont utilisés à la place [45].

Les STFT consistent en une seule transformée de Fourier pour une partie de la forme d'onde sous une fenêtre glissante. Le spectrogramme du signal vocal est ensuite dérivé des STFT.

$$S(\omega) = |X(m, \omega_k)|^2 \quad (5)$$

Les spectrogrammes peuvent être la distribution de la densité de puissance dans le temps et la fréquence, comme le montre la Figure 3. La distribution de la densité de puissance du signal vocal varie considérablement avec le temps et la fréquence et peut être utilisée pour distinguer la voix normale de la voix malade. On peut également voir sur la figure que la distribution de puissance de la parole normale est uniforme dans le temps et la fréquence. Cependant, elle n'est pas non plus uniforme [46].

II.5 Méthodes de Classification

L'extraction de caractéristiques est la première étape de tout système de détection des troubles de la voix. À cette étape, les signaux vocaux donnés sont convertis en caractéristiques acoustiques représentatives à l'aide de diverses techniques de traitement numérique du signal. Nous discuterons maintenant des techniques les plus couramment utilisées pour l'analyse acoustique et l'extraction de caractéristiques dans le domaine concerné et les décrirons.

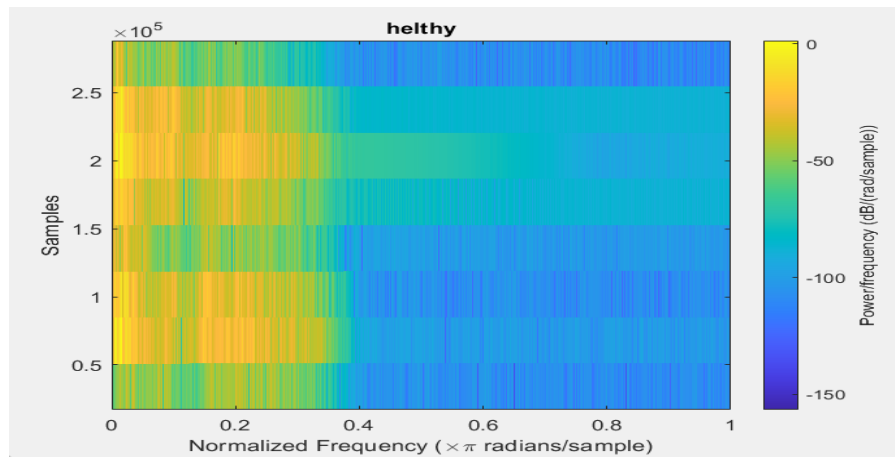


Figure 12: Spectrogrammes de toux Covid-19.

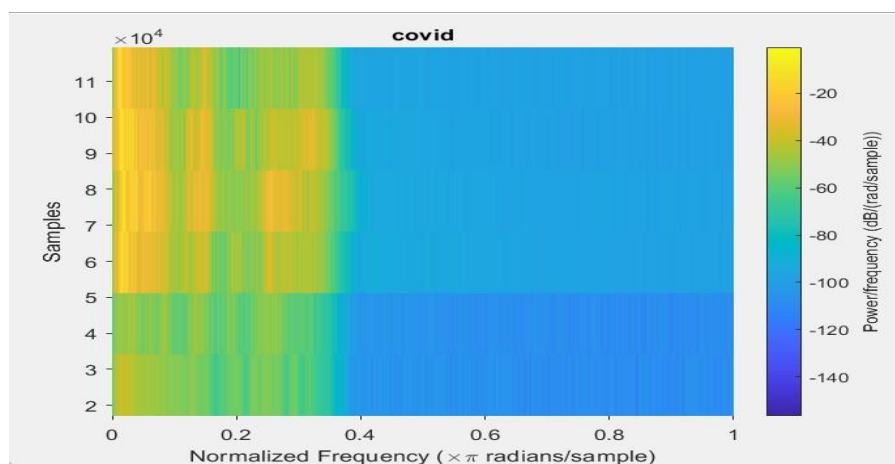


Figure 13: Spectrogrammes de toux Covid-19.

- **Machine à vecteurs de support (SVM).**
- **K plus proches voisins (KNN).**
- **Arbre de decision .**
- Modèle de mélange gaussien (GMM).
- Modèle de fond universel avec mélange gaussien (GMM-UBM).
- Réseau neuronal artificiel (ANN).
- Modèle de Markov caché (HMM).
- Réseau neuronal profond (DNN).
- Réseau neuronal convolutif (CNN).
- Réseau neuronal probabiliste (PNN).
- Réseau de croyances profondes (DBN).
- Réseau neuronal de régression généralisée (GRNN).
- Classifieur bayésien.
- Le clustering k-means.

- Algorithme de l'arbre de décision.
- Analyse discriminante linéaire (LDA).

Récemment, plusieurs approches ont été utilisées dans le domaine, dans notre système de détection nous utilisons quatre types différents (Machine à vecteurs de support (SVM) avec fonction de base radiale (RBF), K plus proches voisins (KNN), Arbre de décision (DT)).

II.5.1 Machine à vecteurs de support (SVM)

Le SVM est un algorithme d'apprentissage automatique qui peut résoudre des problèmes tels que la classification, la régression et la détection. Une machine à vecteurs de support est une technique discriminante, une méthode d'apprentissage supervisée utilisée pour la classification et la régression. Elle consiste à séparer deux ensembles ou plus de points par un hyperplan. Dépend de la situation et de la configuration des points. L'idée originale du SVM était basée sur l'utilisation d'une fonction noyau, permettant une séparation optimale des points planifiés en différentes classes.

La méthode utilise un ensemble de données d'entraînement. Cela permet à l'hyperplan de séparer au mieux les points. Nous utilisons un SVM multi-classe.

Plus formellement, les SVM construisent un hyperplan ou un ensemble d'hyperplans dans un espace de grande ou infinie dimension qui peut être utilisé pour la classification, la régression ou d'autres tâches telles que la détection d'outliers. Intuitivement, une bonne séparation est obtenue par l'hyperplan avec la plus grande distance par rapport au point de données d'entraînement le plus proche de n'importe quelle classe (le soi-disant marge fonctionnelle), puisque en général, plus la marge est grande, meilleur est l'erreur de généralisation du classifieur. Alors que le problème original peut être formulé dans un espace de dimension finie, il arrive souvent que les ensembles à distinguer ne soient pas séparables linéairement dans cet espace. Par conséquent, il a été proposé de mapper l'espace de dimension finie original dans un espace de dimension supérieure, ce qui peut aider à la séparation dans cet espace. Afin de maintenir la charge de calcul raisonnable, le mappage utilisé par le schéma SVM est conçu pour garantir que le produit scalaire de paires de vecteurs de données d'entrée peut être facilement calculé à partir des variables dans l'espace d'origine en choisissant une fonction de noyau $k(x,y)$ pour correspondre au problème. Un hyperplan dans un espace de dimension supérieure est défini comme un

ensemble de points dont le produit scalaire avec un vecteur dans cet espace est constant, un tel ensemble de vecteurs est l'ensemble des vecteurs orthogonaux (et donc minimaux) qui définissent l'hyperplan.[47][48]

La définition du plan hyperplan peut être choisie comme une combinaison linéaire avec le paramètre α_i de l'image du vecteur caractéristique x_i apparaissant dans la base de données. En choisissant un hyperplan, le point x dans l'espace des caractéristiques qui se mappe sur l'hyperplan est défini par la relation:

$$\sum_i \alpha_i k(x_i, x) = \text{constant} \quad (6)$$

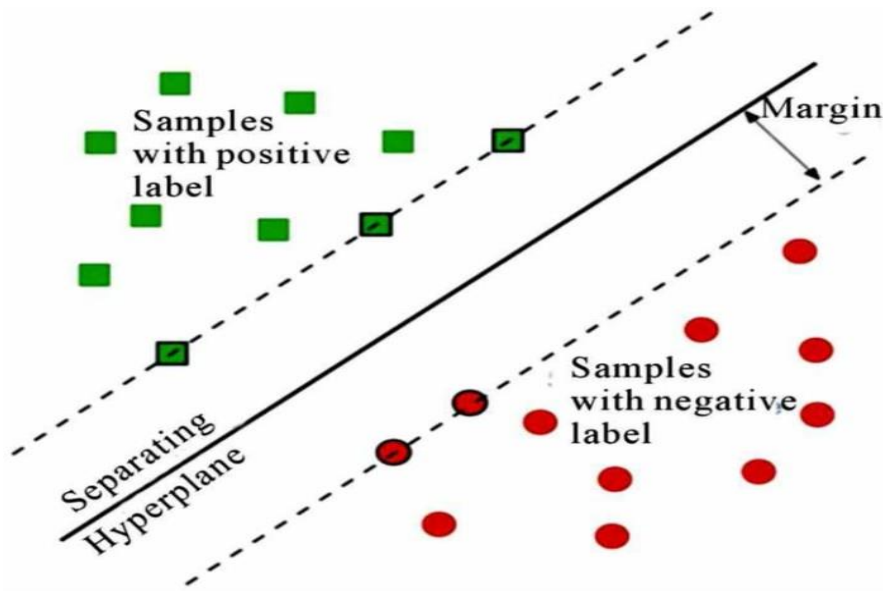


Figure 14: Concept de SVM (Machine à Vecteurs de Support) [4].

Nous avons deux types de classificateurs SVM:

II.5.1.a SVM Linéaire

est un classificateur utilisé pour des données linéairement séparables, ce qui implique que si un ensemble de données peut être classifié en deux classes en utilisant une seule ligne droite, il est appelé données linéairement séparables et le classificateur utilisé est appelé SVM Linéaire.

II.5.1.b SVM Non-Linéaire

est utilisé pour des données non-linéairement séparées, ce qui implique que si un ensemble de données ne peut pas être classifié en utilisant une ligne droite, il s'agit de données non-linéaires, et le classificateur utilisé est appelé SVM Non-linéaire.

II.5.2 K-nearest Neighbor(KNN)

Le K-plus proches voisins est une méthode traditionnelle de reconnaissance de motifs statistiques supervisée qui classe les images en comparant la valeur "K" des données d'entraînement avec les données de test pour trouver la proximité avec l'image ou les données de test. La valeur "K" est estimée à partir de l'extraction de caractéristiques effectuée pendant l'entraînement. Il utilise le principe de l'équation euclidienne dans le classificateur ANN pour identifier la similarité. Il est utilisé pour la classification et la régression. Dans les deux cas, l'entrée est constituée des k échantillons d'entraînement les plus proches dans l'ensemble de données. La sortie dépend de l'utilisation de KNN pour la classification ou la régression :

- Avec la classification KNN, la sortie est un membre de classe. Les objets sont classés par des votes multiples de leurs voisins, et les objets se voient attribuer la classe la plus fréquente parmi leurs k voisins les plus proches (K est un entier positif, généralement petit). Si $k = 1$, alors l'objet est simplement attribué à la classe de ce seul voisin le plus proche. Pour la régression KNN, la sortie est la valeur de l'attribut de l'objet. Cette valeur est la moyenne des valeurs des k voisins les plus proches. KNN est un type de classification où la fonction est seulement approximée localement et toutes les calculs sont différés jusqu'à l'évaluation de la fonction. Parce que l'algorithme repose sur la distance pour la classification lorsque les caractéristiques représentent différentes entités physiques ou sont à des échelles différentes, normaliser les données d'entraînement peut considérablement améliorer sa précision.
- Pour la classification et la régression, il peut être une technique utile d'attribuer des poids aux contributions des voisins afin que les voisins plus proches contribuent davantage à la moyenne que les voisins plus éloignés. Par exemple, un schéma de pondération courant est de donner à chaque voisin un poids de $1/d$, où d est la distance au voisin. Les voisins sont pris dans un ensemble d'objets dont les classes (pour la classification KNN) ou les valeurs d'attribut d'objet (pour la régression KNN) sont connues. Cela peut être vu comme un ensemble d'entraînement pour l'algorithme, bien que aucune étape d'entraînement explicite ne soit requise. Une caractéristique de l'algorithme KNN est qu'il est sensible à la structure locale des données, comme les fruits et légumes. Les céréales peuvent être mesurées par leur croustillance et leur douceur (voir figure ci-dessous). À des fins d'affichage, il n'y a que deux caractéristiques sur le tracé en 2D utilisé. En fait, il peut y avoir un nombre quelconque de prédicteurs, et l'exemple peut être étendu pour inclure un nombre quelconque de caractéristiques. En général, le fruit est plus sucré qu'un légume. Les grains ne sont ni croquants ni sucrés. Notre tâche

consiste à déterminer dans quelle catégorie tombe la patate douce. Dans cet exemple, nous avons sélectionné quatre types de nourriture suivants : des pommes, des haricots verts, de la laitue et du maïs. Comme les légumes ont obtenu le plus de votes, les patates douces ont remporté le prix dans la classe des légumes. Vous pouvez voir que les concepts clés du KNN sont faciles à comprendre. Dans l'exemple ci-dessus, il y a deux concepts importants. L'un d'eux est une manière de calculer la distance entre les patates douces et les autres aliments. Par défaut, la fonction KNN utilise la distance euclidienne, qui peut être calculée avec l'équation suivante [49],[50] :

$$D_m(\mathbf{x}_i, \mathbf{x}_j) = \left(\sum_{k=1}^n |x_{ik} - x_{jk}|^m \right)^{1/m} \quad (7)$$

ou p et q sont les éléments à comparer avec n caractéristiques. Il existe d'autres

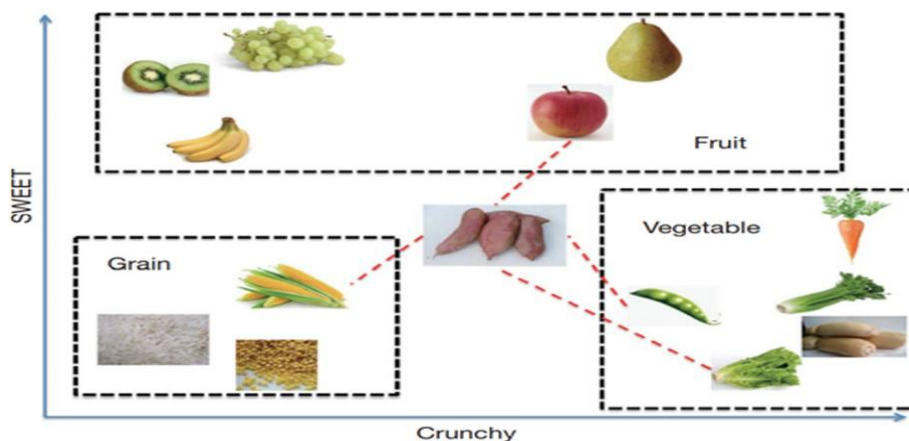


Figure 15: Concept de k-NN (k plus proches voisins).

Les méthodes de calcul de la distance incluent la distance de Manhattan [51],[52].

Un autre concept est de déterminer le paramètre K , qui indique combien de voisins sont sélectionnés par l'algorithme KNN. Le choix approprié de K a un impact significatif sur les performances diagnostiques de l'algorithme KNN. Un grand K réduit l'effet de la variance causée par les erreurs aléatoires, mais au risque d'ignorer de petits mais importants motifs. La clé pour choisir la bonne valeur de K est de trouver un équilibre entre le surajustement et le sous-ajustement. Certains auteurs recommandent de fixer K égal à la racine carrée de la quantité d'observations dans l'ensemble de données d'entraînement. [53] [54].

II.5.3 Arbre de D'écision

Un arbre de décision est un classificateur représenté sous forme d'une partition récursive de l'espace des instances. Un arbre de décision se compose de nœuds formant un arbre raciné, ce qui signifie que c'est un arbre dirigé avec un nœud appelé "Racine" et aucun bord d'entrée. Tous les autres nœuds n'ont qu'un seul bord entrant. Un nœud avec des bords sortants est appelé nœud interne ou nœud de test. Tous les autres nœuds sont appelés feuilles (également appelées nœuds terminaux ou nœuds de décision). Chaque nœud interne dans un arbre de décision divise l'espace des instances en deux ou plusieurs sous-espaces en fonction de conditions données et entre une fonction discrète des valeurs d'attribut. Dans le cas le plus simple et le plus courant, chaque test examine une propriété, l'espace des instances par valeur d'attribut. Pour les propriétés numériques, une condition fait référence à une plage. Chaque feuille est assignée à une classe qui représente la valeur cible la plus appropriée [55],[56].

II.6 Conclusion:

Dans ce chapitre, nous avons abordé l'extraction des caractéristiques et leurs types, en citant les caractéristiques les plus utilisées dans ce domaine. Ensuite, nous avons discuté des classificateurs utilisés dans ce travail et des concepts associés. Par la suite, nous avons établi une explication sur l'apprentissage automatique et ses types, et mentionné l'intelligence artificielle. Dans le prochain chapitre, nous évoquerons d'autres travaux similaires et discuterons de la base de données expérimentale ainsi que de la conception de notre système et de la configuration expérimentale pour construire notre système. Ensuite, nous aborderons les résultats.

Mise en œuvre et Évaluation

III.1 Introduction

Les données de toux sont collectées pour les patients symptomatiques et asymptomatiques. Cependant, comme mentionné précédemment, le son de la toux peut être fiable pour différencier les patients atteints de COVID-19 des groupes en bonne santé. Il est également possible de constater que la scène la plus puissante est de stratifier les patients asymptomatiques avec le son de toux de COVID-19. Les patients asymptomatiques sont ceux qui propagent le virus sans le savoir. Notre réseau entraîné peut bien les identifier grâce à leurs toux. C'est pourquoi ce cadre de dépistage du COVID-19 peut aider de manière significative à dépister la population des cas suspects et à réduire le risque de transmission. Dans ce chapitre, nous discuterons des résultats de l'application de la méthode proposée pour la détection précoce du Covid-19. La base de données utilisée est TOS COVID. Pour la détection du Covid-19, les caractéristiques utilisées pour la correspondance sont les MFCC (Mel-frequency cepstral coefficients). En fait, nous les utilisons car ils simulent l'audition humaine. La première partie de notre travail consiste à segmenter le signal puis à extraire les attributs. La deuxième partie du travail utilise les vecteurs d'attributs pour construire le modèle de prédiction du classificateur. Nous évaluerons la méthode proposée en calculant : la précision, la sensibilité, la spécificité, le score F1 et la courbe ROC. Enfin, nous interpréterons les résultats.

III.2 Détection du Covid-19 à l'aide de la voix

L'utilisation de la voix pour détecter les symptômes du COVID-19 est particulièrement attrayante dans la situation actuelle alors que de nombreux pays font face à des pénuries de kits de dépistage du COVID-19. Il est donc impératif d'utiliser judicieusement les kits de dépistage disponibles. Comme il existe une possibilité de détecter les symptômes à un stade précoce à l'aide de la méthode proposée, les individus peuvent immédiatement s'isoler par mesure de précaution afin de réduire la propagation avant d'être testés et de bénéficier d'un soutien médical supplémentaire. Une fois développée, il n'y aura pas besoin de dispositifs médicaux spécialisés. La parole peut être recueillie à l'aide d'un microphone disponible sur chaque appareil mobile et peut être analysée à l'aide d'une application. Si des symptômes sont détectés, l'application peut être configurée pour alerter l'utilisateur ainsi que les autorités sanitaires. Le dépistage initial des symptômes suspects du COVID-19 peut être effectué à distance, car les signaux de toux peuvent être

collectés par téléphone. Les anomalies dans les paramètres biologiques peuvent être détectées bien avant que les symptômes du COVID-19 ne deviennent apparents. Le dépistage à grande échelle sera possible à faible coût, et sans recourir à des dispositifs médicaux complexes ou coûteux. La prolifération et l'omniprésence des smartphones permettent de déployer rapidement l'application de dépistage, une fois que les modèles d'IA ont été entraînés et incorporés dans une application, rendant le dépistage en temps réel accessible à toute personne possédant un smartphone.

III.3 État de l'art

Ce travail présente une approche novatrice en matière d'apprentissage automatique pour la détection automatique du COVID-19 en utilisant les sons de toux pour les patients symptomatiques et asymptomatiques. Comme le montre ci-dessous, l'approche innovante proposée a fourni les meilleures performances par rapport à quelques autres études. En revanche, certaines ont indiqué des solutions pour la détection des troubles liés au Covid-19 basées sur une analyse de la voix.

III.3.1 N. Sharma 2020

Le premier système proposé basé sur le son pour détecter la COVID-19 a été présenté par "N.Sharma" en 2020, où il a mené son étude basée sur une base de données contenant (941 individus sains et positifs à la COVID-19). Il a étudié les phénomènes suivants : toux, respiration, voyelle et comptage (de 1 à 20). La méthode présentée par "N.Sharma" était un système basé sur un classificateur de forêt aléatoire utilisant le contraste spectral, les MFCC (coefficients de cepstre de fréquence mél), le roll-off spectral, le centroïde spectral, l'énergie quadratique moyenne, l'ajustement polynomial, le taux de passage par zéro, la largeur de bande spectrale et la platitude spectrale. Ce système a atteint une précision de 76,74 (%). [57]

III.3.2 C. Brown et al. 2021

"C.Brown" a présenté son système proposé pour la détection de l'asthme et de la COVID ensemble, et son étude a été réalisée en 2021. Il a basé son étude sur une base de données contenant (141 positifs à la COVID-19, 298 non-COVID, 54 positifs à la COVID-19 avec toux, 32 non-COVID avec toux, 20 asthmatiques non-COVID). Ce système a étudié les phénomènes suivants : toux et respiration, et il est basé sur une approche utilisant des réseaux de neurones convolutifs (CNN) avec des spectrogrammes, le centroïde spectral et les MFCC. Ce système a atteint une précision de 80 (%). [58]

III.3.3 V. Espotovic 2021

La même année, "V. Espotovic" est venu présenter une nouvelle étude basée sur la détection du COVID-19 à partir de la toux et de la respiration. Il a utilisé une base de données contenant (COVID-19-Positif : 84, COVID-19-Négatif : 419). La méthode présentée par "V. Espotovic" était un système basé sur une approche d'ensemble boostée utilisant un spectrogramme et une ondelette. Ce système a atteint une précision de 88,52%. [59]

III.3.4 R.Islam 2022

"R. Islam" est venu présenter une nouvelle étude basée sur la détection du Covid uniquement à partir de la toux, où il a utilisé une base de données contenant (COVID-19-Positif : 50, Sain : 50). La méthode proposée par "R. Islam" était un système basé sur une approche basée sur les CNN utilisant le taux de passage par zéro, l'énergie, l'entropie de l'énergie, le centroïde spectral, l'entropie spectrale, le flux spectral, les rouleaux spectraux, MFCC. Ce système a atteint une précision de 88,52 %. [60]

III.3.5 Rahman.T 2022

"Rahman.T" a présenté son système de détection du Covid-19, où il a basé son étude sur une base de données contenant (COVID-19-Positif : 237, Sain : 827). Ce système a étudié les phénomènes suivants (toux, respiration) et était basé sur une approche basée sur les CNN empilés utilisant des spectrogrammes. Ce système a atteint une précision de 96,5 % pour les symptomatiques et de 98,85 % pour les asymptomatiques. [61]

III.4 Conception du système

La procédure générale de l'approche proposée pour la détection du COVID-19 est décrite dans le diagramme de flux de la figure 16. Dans la première partie, nous restructurons et organisons la base de données, puis nous échantillons le signal. La méthode d'analyse utilise un algorithme pour estimer le signal de toux à partir du signal d'enregistrement. Cela signifie que l'algorithme extraira des échantillons. Dans la deuxième partie, nous passons à l'étape suivante, où les paramètres sonores spéciaux sont extraits de chaque échantillon du signal de toux prélevé. Ainsi, chaque son a ses propres paramètres. Après avoir extrait les caractéristiques audio, nous passons à la dernière étape de cette partie, qui consiste à choisir l'un des classificateurs suivants (KNN, SVM, DT). Dans la troisième et dernière partie, le classificateur sélectionné construit son propre modèle de classification, puis classe les signaux de toux en deux classes : positif et négatif.

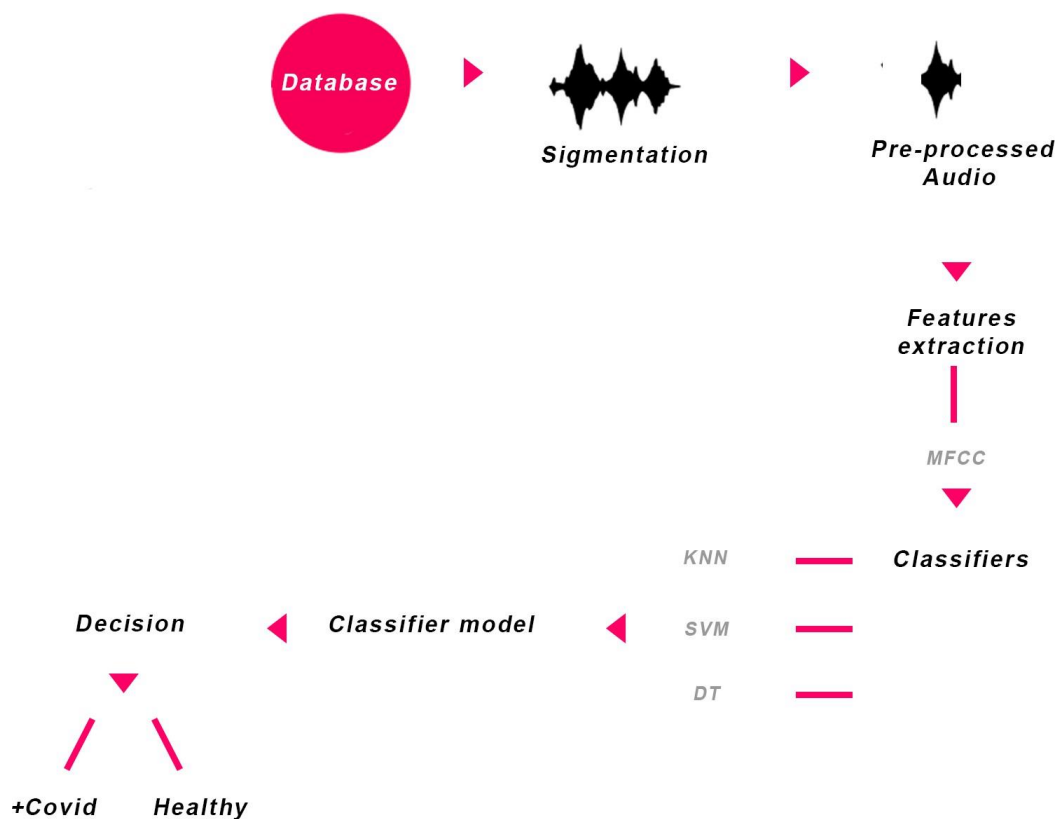


Figure 16: Conception système

III.5 Base de données expérimentale

Les échantillons audio collectés pour cette étude proviennent d'individus retirés de tous les établissements de test publics et d'un établissement privé dans la ville de Buenos Aires. Ces institutions effectuent des études RT-PCR sur des patients suspectés de COVID, et 14 unités d'isolement extra-hospitalières pour des patients atteints de cas de COVID légers confirmés. Ces données audio ont été collectées via un chatbot WhatsApp spécialement conçu dans la ville de Buenos Aires pour répondre aux questions des citoyens liées à la pandémie de coronavirus (maladie à coronavirus). Le premier ensemble de données collectées correspond à 2821 personnes testées à Buenos Aires du 11 août au 2 décembre 2020. Les personnes ont été réparties en 1409 personnes testées positives au COVID-19 et 1412 personnes testées négatives au COVID-19. Dans ce groupe d'échantillons, 52,6 % des sujets étaient des femmes et 47,4 % étaient des hommes. Sur la base de leurs résultats de test PCR par écouvillon, les fichiers audio de chaque personne ont été classés comme "positifs" et "négatifs". Dans ce travail, nous utilisons 2771 fichiers audio, qui ont été classés en 1378 détectables et 1393 non détectables.

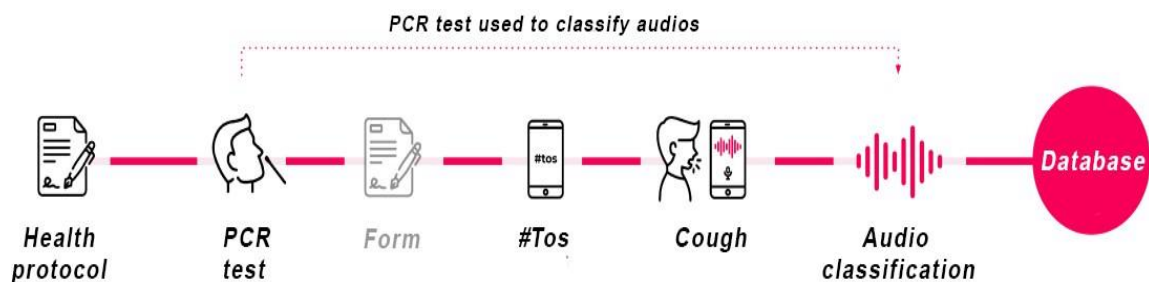


Figure 17: Conception de la configuration de la base de données

III.6 Configuration expérimentale

III.6.1 Matériel de développement

On a utilisés PyCharm Community Edition 2024.1 sur une plateforme expérimentale sous forme d'un PC Portable (HP ProBOOK G4), avec un système d'exploitation Microsoft Windows 10 64 bits, équipé de :

- *CPU* : Intel i5-6200U
- *RAM* : 8 GB
- *GPU*: (Inet HD Graphique)

III.6.2 Logiciel de développement

Le programme utilisé : PyCharm Community Edition 2024.1

PyCharm est un environnement de développement intégré (IDE) populaire spécifiquement conçu pour le développement en Python. Il offre des fonctionnalités telles que l'auto-complétion du code, des outils de débogage et une intégration de contrôle de version, ce qui en fait un outil puissant pour les programmeurs Python. PyCharm existe en versions gratuite et payante, cette dernière offrant des fonctionnalités supplémentaires pour le développement web et le calcul scientifique. Avec son interface conviviale et ses fonctionnalités étendues, PyCharm est un excellent choix tant pour les débutants que pour les développeurs Python expérimentés. Si vous recherchez un moyen puissant et efficace d'écrire du code Python, PyCharm vaut vraiment le détour.

III.7 Configuration de la méthode

III.7.1 Séparation des Données

Dans la première étape, nous effectuons le prétraitement des données, en classant les fichiers ID en quatre catégories : 'Homme', 'Femme', 'détectable', 'non détectable' pour chacune d'entre elles.

III.7.2 Segmentation des données

Nous segmentons deux échantillons de sons de toux pour comparer les différences entre les toux des sujets COVID-négatifs (c'est-à-dire sains/contrôlés) et les patients COVID-positifs. Les échantillons de sons de toux provenant de sujets sains sont montrés

dans la Figure 14, tandis que ceux d'un patient COVID-positif sont représentés dans la Figure 18. Cela illustre qu'un échantillon sain ressemble à une toux humaine typique. Cependant, les échantillons de sons de toux des patients atteints de COVID-19 présentent des différences significatives par rapport aux toux humaines typiques. Par exemple, le stade médian et le stade vocalisé durent tous deux plus longtemps chez les patients COVID-positifs que chez les sujets sains.

De plus, pendant la phase voisée, l'amplitude du signal est plus élevée chez les patients positifs au COVID que chez les sujets sains. Les deux toux sonnent également différemment pendant la phase d'explosion, comme illustré dans la Fig. 16. Les différences mentionnées ci-dessus indiquent que le son de la toux peut être utilisé comme un outil précieux pour discriminer les patients positifs au COVID des sujets sains. Le signal de toux enregistré est en format numérique et est pré-traité.

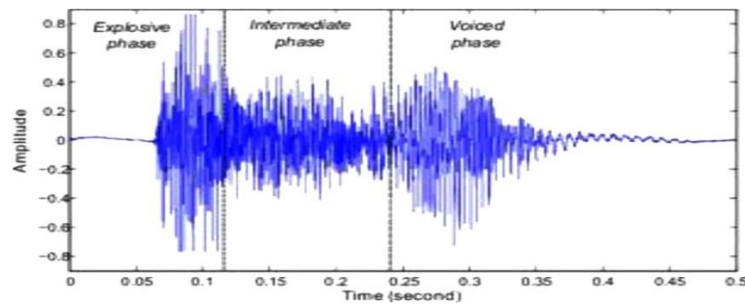


Figure 18: Un signal sonore typique de toux

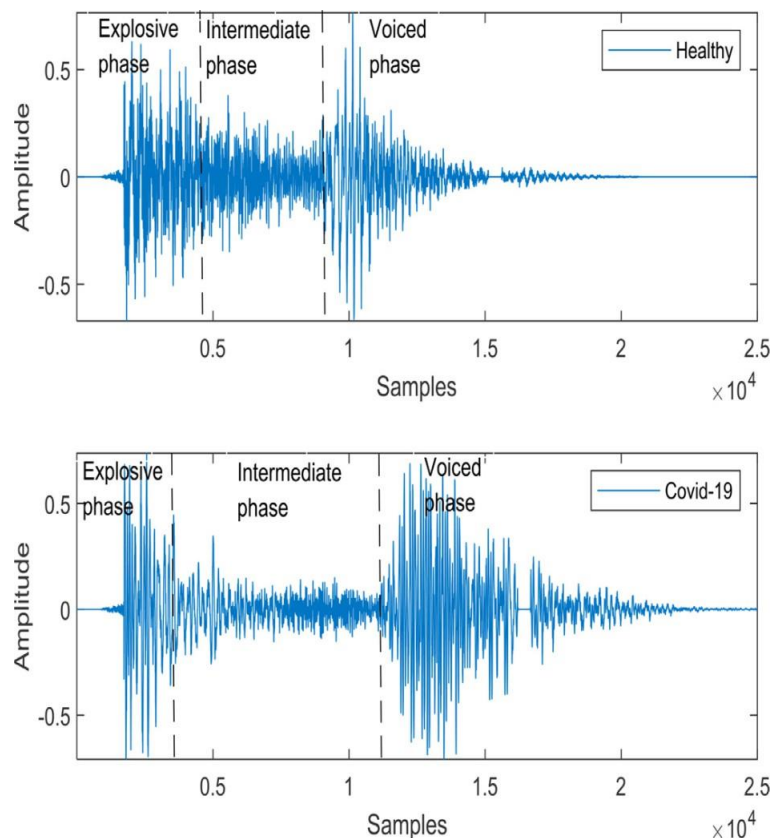


Figure 19: Comparaison des sons de toux pour un sujet en bonne santé et un sujet atteint de la COVID-19

Pour supprimer les composants indésirables, le bruit de fond et même les intervalles de silence/pause à l'aide de la technique de traitement du signal. Le signal de toux était un signal aléatoire que nous avons normalisé, puis nous avons fixé le seuil (0,9) de

l'amplitude du signal pour localiser la toux dans le signal enregistré. Ensuite, nous avons appliqué une convolution entre le signal normalisé et une fenêtre rectangulaire de taille (2084 ms) pour découper la durée de la première et de la deuxième toux.

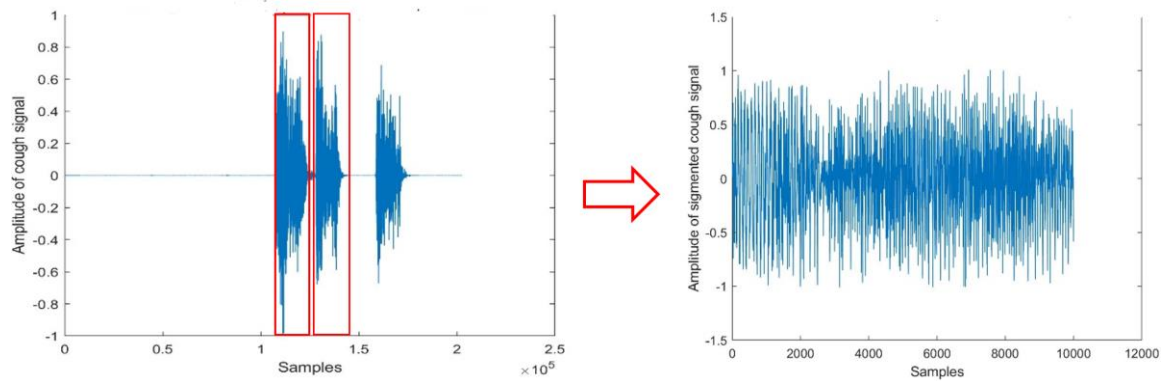


Figure 20: Signal segmenté

III.7.3 Extraction des caractéristiques

Ensuite, le signal de toux est davantage traité à l'aide d'un algorithme d'extraction de caractéristiques pour extraire des "caractéristiques" qui caractérisent le signal de toux. Cet algorithme de caractéristiques était les coefficients cepstraux de fréquence MEL (MFCC). Les MFCC sont une simulation de l'oreille humaine, qui peut distinguer les sons pour les deux sexes, et donc ses paramètres diffèrent entre les sexes, ainsi que les sons entre eux.

III.7.3.a Algorithme de coefficients spectraux de fréquence MEL

Les MFCC sont des caractéristiques populaires extraites des signaux vocaux pour les tâches de reconnaissance. Dans le modèle source-filtre de la parole, les coefficients cepstraux sont compris pour représenter le filtre (le tractus vocal). La réponse en fréquence du tractus vocal est relativement lisse, et la source du son vocalisé peut être modélisée comme un train d'impulsions. Par conséquent, le tractus vocal peut être estimé à partir de l'enveloppe spectrale du segment de parole. L'idée derrière les coefficients cepstraux de fréquence MEL est de compresser les informations sur le tractus vocal (spectre lisse) en un petit nombre de coefficients, sur la base d'une compréhension de la cochlée. Bien qu'il n'y ait pas de norme fixe pour le calcul des coefficients, la figure décrit les étapes de base.

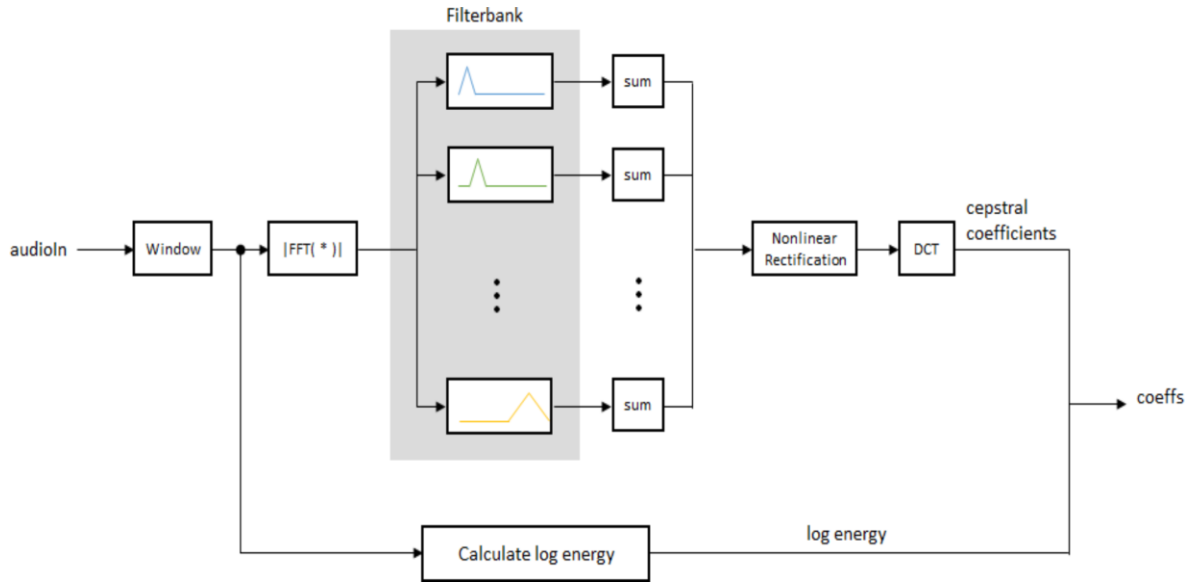


Figure 21: Les blocs de MFCC

Le banc de filtres Mel par défaut espace linéairement les 10 premiers filtres triangulaires et espace logarithmiquement les filtres restants. Les informations contenues dans les coefficients cepstraux de la fréquence Mel zéro

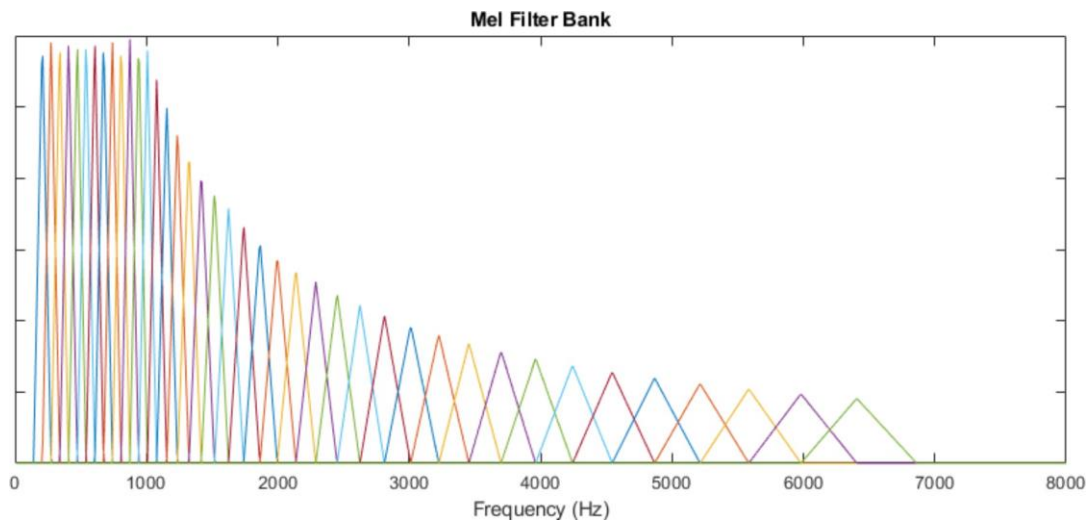


Figure 22: La banque de filtres Mel

sont généralement complétées ou remplacées par une énergie logarithmique. Le calcul de l'énergie logarithmique dépend du domaine d'entrée. Si l'entrée (audioIn) est un signal temporel, l'énergie logarithmique est calculée en utilisant l'équation suivante :

$$\log E = \log \left(\sum x^2 \right) \tag{8}$$

Si l'entrée (audioIn) est un signal dans le domaine fréquentiel, l'énergie logarithmique est calculée à l'aide de l'équation suivante :

$$\log E = \log (\text{sum } (|x|^2) / \text{FFTLenght}) \quad (9)$$

III.7.3.b Classification

La classification, qui est la dernière étape de notre système proposé. La sélection d'un des classificateurs suivants avec ses paramètres : SVM, CNN, DT.

III.8 Métriques expérimentales

Dans cette étude, les performances du système de détection des pathologies vocales ont été évaluées par les classificateurs KNN, SVM et DT, en utilisant les MFCC pour les deux sexes. Les résultats de classification sont fournis en termes d'exactitude, de rappel, de précision, de spécificité, de score F1 et de ROC. Ces indicateurs de performance sont largement utilisés pour évaluer l'efficacité des différents systèmes de décision médicale.

- **Accuracy:**

$$\text{Accuracy}(\%) = \frac{TP + TN}{TP + TN + FP + FN} \times 100 \quad (10)$$

- **Sensitivity:**

$$\text{Recall (Sensitivity)} (\%) = \frac{TP}{TP + FN} \times 100 \quad (11)$$

- **Specificity:**

$$\text{Specificity}(\%) = \frac{TN}{TN + FP} \times 100 \quad (12)$$

- **Precision:**

$$\text{Precision}(\%) = \frac{TP}{TP + FP} \times 100 \quad (13)$$

- **F1 Score:**

$$F_1\text{Score}(\%) = \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \times 100 \quad (14)$$

- **Vrai Positif (TP):** la nature pathogène de l'échantillon vocal est reconnue par le marqueur.
- **Vrai Négatif (TN):** l'échantillon vocal est sain, et le marqueur le détecte.
- **Faux Positif (FP):** l'échantillon vocal est normal, mais le marqueur le détecte comme pathologique.

- **Faux Négatif (FN)**: l'échantillon vocal malade est reconnu par le marqueur comme étant sain.

III.9 Résultats et discussions

Comme nous le savons, l'orientation vers un système de détection médicale est considérée comme l'une des meilleures solutions pour améliorer les performances. Dans ce sens, lors de notre étude expérimentale basée sur un système proposé, nous avons procédé à tous les entraînements et tests. Dans la première étape, nous avons testé notre système avec des signaux de toux féminins puis masculins. De la même manière, dans la deuxième étape, nous avons mesuré les performances de notre système. Le tableau ci-dessous résume les résultats de l'évaluation des femmes en utilisant les MFCC avec ces paramètres.

Pour les femmes :

- Longueur de la fenêtre: 1984
- Longueur de chevauchement : 90

Table 1: Résultats des femmes avec des caractéristiques MFCC

Modele	Accuracy%	Sensitivity%	specificity%	F1 score%
SVM	50.93	14.61	14.61	21
KNN	55.27	48.8	48.8	49
DT	45.54	27.85	27.85	31

Le tableau précédent montre les performances du système proposé pour détecter le Covid-19 et son efficacité chez les femmes. L'algorithme KNN a montré la meilleure précision de détection, atteignant 49 % pour les individus infectés et 48.8 % pour les individus non infectés.

La table ci-dessous résume les résultats de l'évaluation des hommes en utilisant les MFCC avec les paramètres suivants:

- Longueur de la fenêtre: 1952
- Longueur de chevauchement 512

Table 2: Résultats des hommes avec des caractéristiques MFCC

Modele	Accuracy%	Sensitivity%	specificity%	F1 score%
SVM	67	59.07	59.07	64
KNN	58	61.00	61.00	59
DT	62	63.32	63.32	62

Le tableau précédent montre les performances du système proposé pour détecter le Covid-19 et son efficacité chez les hommes. L'algorithme SVM a montré la meilleure précision de détection, atteignant 64 % pour les individus infectés et 59.07 % pour les individus non infectés.

III.10 Comparaison des Courbes ROC

Les courbes ROC (Receiver Operating Characteristic) sont utilisées pour évaluer la performance des modèles de classification. Elles tracent le taux de vrais positifs (True Positive Rate, TPR) en fonction du taux de faux positifs (False Positive Rate, FPR) à différents seuils de classification. La surface sous la courbe (AUC) est une mesure de la capacité du modèle à distinguer entre les classes positives et négatives.

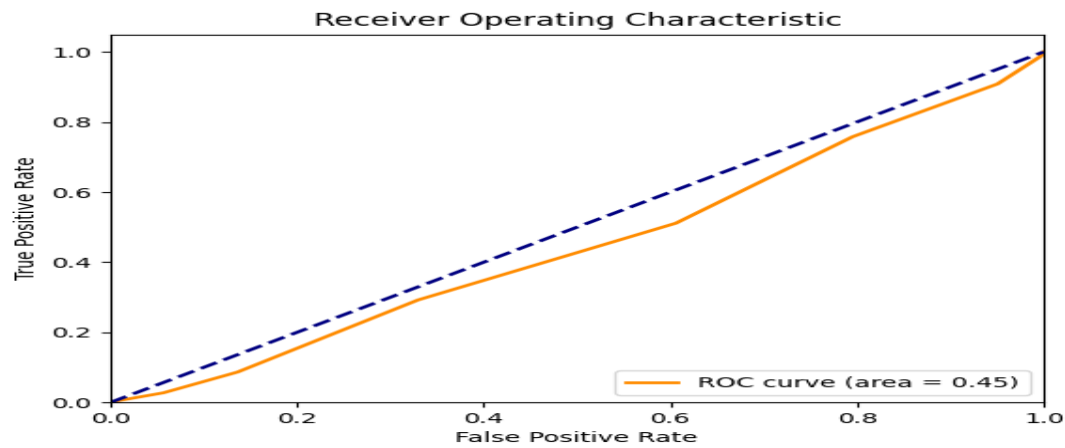


Figure 23: La courbe ROC pour la classification des femmes en utilisant KNN

La courbe est presque une ligne droite qui suit la diagonale, avec une AUC de 0.45.

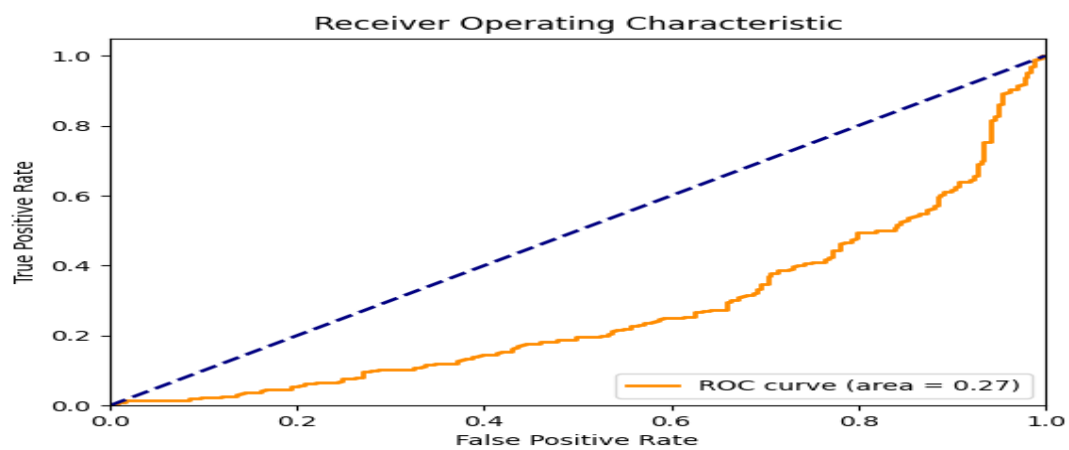


Figure 24: La courbe ROC pour la classification des hommes en utilisant SVM

La courbe montre une légère amélioration par rapport à la diagonale, mais reste très proche de celle-ci, avec une AUC de 0.27.

III.11 Conclusion:

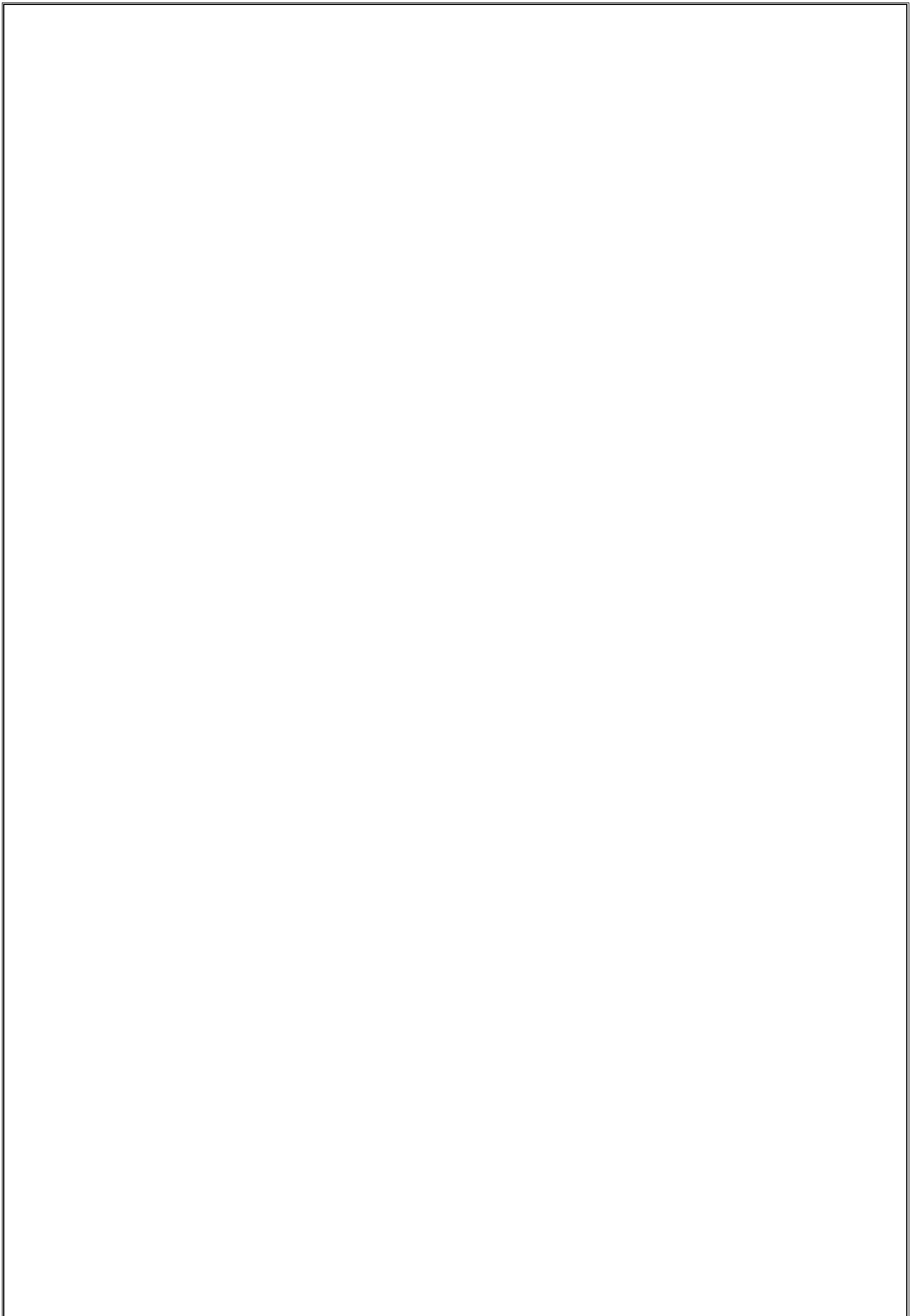
Dans ce chapitre, nous avons analysé et évalué les techniques et architectures les plus avancées proposées dans les recherches récentes pour le diagnostic du Covid-19 à partir de données vocales médicales en utilisant l'apprentissage automatique. La majorité des architectures étudiées atteignent des niveaux de précision relativement élevés. Cependant, il demeure crucial de poursuivre les recherches sur des ensembles de données plus larges et nouveaux afin de perfectionner les méthodes actuelles et obtenir des résultats encore plus efficaces.

Conclusion Générale

Les usages médicaux évoluent constamment avec la technologie et sont largement utilisés dans de nombreux domaines pour la détection des pathologies. L'objectif fondamental des techniques de détection médicale est de détecter la COVID-19 chez la personne testée. Cette méthode de dépistage de la COVID est une alternative aux méthodes traditionnelles des analyses médicales. Tout le monde essaie de se protéger pour assurer leur vie. Ainsi, lorsque les gens prennent un selfie dans le cadre des avantages incroyables pour la société. Toute innovation est bénéfique pour le bien-être des gens et l'amélioration dynamique. Une sorte de développement mécanique ultérieur réside dans le domaine de la connaissance manufacturière et ses domaines associés. De cette manière, il vise à aider à freiner cette propagation mondiale. Avec cela en tête, ce sont des approches prospectives et elles se sont avérées efficaces en termes de précision. Notre travail dans ce domaine a des limites en nombre de tests utilisés pour construire des connaissances préalables. Cela peut bien sûr être utilisé comme confirmation, fournissant le concept d'opérabilité de la machine et des méthodes d'apprentissage bien fondées pour préparer le diagnostic de la COVID.

Les principaux objectifs de notre étude:

- Nous avons essayé de proposer la réalisation d'un système de détection basé sur les MFCC utilisés (KNN, SVM, DT).
- Nous avons utilisé une base de données prétraitée et appliqué l'apprentissage supervisé de type machine learning.
- Comparaison des classificateurs (SVM, KNN, DT).
- Les résultats obtenus dans ce travail montrent une excellente performance pour la pathologie et les personnes en bonne santé. Nous avons finalement présenté les résultats expérimentaux, ces résultats sont présentés à partir des métriques et des courbes ROC de l'algorithme SVM. Pendant ce travail, ces résultats sont très encourageants et montrent l'efficacité pour les deux sexes, afin de développer un système de détection efficace.



Bibliography

- [1] P. Zhou, X.-L. Yang, X.-G. Wang, B. Hu, L. Zhang, W. Zhang, H.-R. Si, Y. Zhu, B. Li, C.-L. Huang *et al.*, “A pneumonia outbreak associated with a new coronavirus of probable bat origin,” *nature*, vol. 579, no. 7798, pp. 270–273, 2020.
- [2] B. Chen *et al.*, “Overview of human lethal coronaviruses,” *Signal Transduction and Targeted Therapy*, vol. 5, p. 89, 2020.
- [3] S. Su, G. Wong, W. Shi, J. Liu, A. C. Lai, J. Zhou, W. Liu, Y. Bi, and G. F. Gao, “Epidemiology, genetic recombination, and pathogenesis of coronaviruses,” *Trends in microbiology*, vol. 24, no. 6, pp. 490–502, 2016.
- [4] E. Hadjaidji, M. C. A. Korba, and K. Khelil, “Spasmodic dysphonia detection using machine learning classifiers,” in *2021 International Conference on Recent Advances in Mathematics and Informatics (ICRAMI)*. IEEE, 2021, pp. 1–5.
- [5] J. M. Sanders, M. L. Monogue, T. Z. Jodlowski, and J. B. Cutrell, “Pharmacologic treatments for coronavirus disease 2019 (covid-19): a review,” *Jama*, vol. 323, no. 18, pp. 1824–1836, 2020.
- [6] S. A. Kujawski, K. K. Wong, J. P. Collins, L. Epstein, M. E. Killerby, C. M. Midgley, and G. R. Abedi, “Clinical and virologic characteristics of the first 12 patients with coronavirus disease 2019 (covid-19) in the united states,” *Nature medicine*, vol. 26, no. 6, pp. 861–869, 2020.
- [7] D. Chang, M. Lin, L. Wei, L. Xie, G. Zhu, C. S. D. Cruz, and L. Sharma, “Epidemiologic and clinical characteristics of novel coronavirus infections involving 13 patients outside wuhan, china,” *Jama*, vol. 323, no. 11, pp. 1092–1093, 2020.
- [8] P. Sun, X. Lu, C. Xu, W. Sun, and B. Pan, “Understanding of covid-19 based on current evidence,” *Journal of medical virology*, vol. 92, no. 6, pp. 548–551, 2020.
- [9] M. Asiaee, A. Vahedian-Azimi, S. S. Atashi, A. Keramatfar, and M. Nourbakhsh, “Voice quality evaluation in patients with covid-19: An acoustic analysis,” *Journal*

of Voice, 2020.

- [10] A. Scohy, A. Anantharajah, M. Bodéus, B. Kabamba-Mukadi, A. Verroken, and H. Rodriguez-Villalobos, “Low performance of rapid antigen detection test as front-line testing for covid-19 diagnosis,” *Journal of Clinical Virology*, vol. 129, p. 104455, 2020.
- [11] R. D. Kent, “Vocal tract acoustics,” *Journal of Voice*, vol. 7, no. 2, pp. 97–117, 1993.
- [12] J. Schroeter and M. M. Sondhi, “Techniques for estimating vocal-tract shapes from the speech signal,” *IEEE Transactions on Speech and Audio Processing*, vol. 2, no. 1, pp. 133–150, 1994.
- [13] J. L. Flanagan, K. Ishizaka, and K. L. Shipley, “Synthesis of speech from a dynamic model of the vocal cords and vocal tract,” *Bell System Technical Journal*, vol. 54, no. 3, pp. 485–506, 1975.
- [14] L. F. Tracy, R. K. Segina, M. D. Cadiz, and C. E. Stepp, “The impact of communication modality on voice production,” *Journal of Speech, Language, and Hearing Research*, vol. 63, no. 9, pp. 2913–2920, 2020.
- [15] K. E. Bainbridge, N. Roy, K. G. Losonczy, H. J. Hoffman, and S. M. Cohen, “Voice disorders and associated risk markers among young adults in the united states,” *The Laryngoscope*, vol. 127, no. 9, pp. 2093–2099, 2017.
- [16] D. R. Boone, S. C. McFarlane, S. L. Von Berg, and R. I. Zraick, “The voice and voice therapy,” 2005.
- [17] S. G. Adams and A. E. Lang, “Can the lombard effect be used to improve low voice intensity in parkinson’s disease?” *European Journal of Disorders of Communication*, vol. 27, no. 2, pp. 121–127, 1992.
- [18] J. K. Casper and R. Leonard, *Understanding voice problems: A physiological perspective for diagnosis and treatment*. Lippincott Williams & Wilkins, 2006.
- [19] A. Deslandes, V. Berti, Y. Tandjaoui-Lambotte, C. Alloui, E. Carbonnelle, J. Zahar, S. Bricler, and Y. Cohen, “Sars-cov-2 was already spreading in france in late december 2019,” *International journal of antimicrobial agents*, vol. 55, no. 6, p. 106006, 2020.
- [20] D. S. W. Ting, L. Carin, V. Dzau, and T. Y. Wong, “Digital technology and covid-19,” *Nature medicine*, vol. 26, no. 4, pp. 459–461, 2020.
- [21] L. K. Tanno, T. Casale, and P. Demoly, “Coronavirus disease (covid)-19:

- World health organization definitions and coding to support the allergy community and health professionals,” *The Journal of Allergy and Clinical Immunology: In Practice*, vol. 8, no. 7, pp. 2144–2148, 2020. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/S2213219820304700>
- [22] K. Erles, C. Toomey, H. W. Brooks, and J. Brownlie, “Detection of a group 2 coronavirus in dogs with canine infectious respiratory disease,” *Virology*, vol. 310, no. 2, pp. 216–223, 2003.
- [23] G. Liya, W. Yuguang, L. Jian, Y. Huaiping, H. Xue, H. Jianwei, M. Jiaju, L. Youran, M. Chen, and J. Yiqing, “Studies on viral pneumonia related to novel coronavirus sars-cov-2, sars-cov, and mers-cov: a literature review,” *Apmis*, vol. 128, no. 6, pp. 423–432, 2020.
- [24] F. Di Gennaro, D. Pizzol, C. Marotta, M. Antunes, V. Racalbutto, N. Veronese, and L. Smith, “Coronavirus diseases (covid-19) current status and future perspectives: a narrative review,” *International journal of environmental research and public health*, vol. 17, no. 8, p. 2690, 2020.
- [25] J. E. Kim, J. H. Lee, H. Lee, S. J. Moon, and E. W. Nam, “Covid-19 screening center models in south korea,” *Journal of Public Health Policy*, vol. 42, no. 1, pp. 15–26, 2021.
- [26] W. Scientific, D. Heymann, N. Shindo *et al.*, “Covid-19: What is next for public health?” *Lancet*, vol. 395, pp. 542–5, 2020.
- [27] Y.-C. Wu, C.-S. Chen, and Y.-J. Chan, “The outbreak of covid-19: An overview,” *Journal of the Chinese medical association*, vol. 83, no. 3, p. 217, 2020.
- [28] I. Futterman, M. Toaff, L. Navi, and C. A. Clare, “Covid-19 and hellp: overlapping clinical pictures in two gravid patients,” *American Journal of Perinatology Reports*, vol. 10, no. 02, pp. e179–e182, 2020.
- [29] X. Yang, Y. Yu, J. Xu, H. Shu, H. Liu, Y. Wu, L. Zhang, Z. Yu, M. Fang, T. Yu *et al.*, “Clinical course and outcomes of critically ill patients with sars-cov-2 pneumonia in wuhan, china: a single-centered, retrospective, observational study,” *The Lancet Respiratory Medicine*, vol. 8, no. 5, pp. 475–481, 2020.
- [30] M. Cascella, M. Rajnik, A. Aleem, S. C. Dulebohn, and R. Di Napoli, “Features, evaluation, and treatment of coronavirus (covid-19),” *Statpearls [internet]*, 2022.
- [31] H.-D. Wehle, “Machine learning, deep learning, and ai: What’s the difference?” 07

- 2017.
- [32] C. A. U. Hassan, M. S. Khan, and M. A. Shah, "Comparison of machine learning algorithms in data classification," in *2018 24th International Conference on Automation and Computing (ICAC)*. IEEE, 2018, pp. 1–6.
- [33] K. Wakefield, "A guide to the types of machine learning algorithms and their applications," URL: https://www.sas.com/en_gb/insights/articles/analytics/machine-learning-algorithms.html [Accessed on 10 February 2021], 2021.
- [34] T. M. Mitchell *et al.*, "Machine learning," 1997.
- [35] G. Sharma, K. Umamathy, and S. Krishnan, "Trends in audio signal feature extraction methods," *Applied Acoustics*, vol. 158, p. 107020, 2020.
- [36] F. Goldman-Eisler, "Speech analysis and mental processes," *Language and speech*, vol. 1, no. 1, pp. 59–75, 1958.
- [37] K. Stevens, "Autocorrelation analysis of speech sounds," *The Journal of the Acoustical Society of America*, vol. 22, no. 6, pp. 769–771, 1950.
- [38] C. G. Howard, "Speech analysis-synthesis scheme using continuous parameters," *The Journal of the Acoustical Society of America*, vol. 28, no. 6, pp. 1091–1098, 1956.
- [39] R. K. Potter and J. C. Steinberg, "Toward the specification of speech," *The Journal of the Acoustical Society of America*, vol. 22, no. 6, pp. 807–820, 1950.
- [40] G. Gambardella, "A contribution to the theory of short-time spectral analysis with nonuniform bandwidth filters," *IEEE Transactions on Circuit Theory*, vol. 18, no. 4, pp. 455–460, 1971.
- [41] S. K. Card, T. P. Moran, and A. Newell, *The psychology of human-computer interaction*. Crc Press, 2018.
- [42] G. Gambardella, "Time scaling and short-time spectral analysis," *The Journal of the Acoustical Society of America*, vol. 44, no. 6, pp. 1745–1747, 1968.
- [43] S. Shetty, S. Hegde, T. Dodderi *et al.*, "Classification of healthy and pathological voices using mfcc and ann," in *2018 Second International Conference on Advances in Electronics, Computers and Communications (ICAECC)*. IEEE, 2018, pp. 1–5.
- [44] A. Bala, A. Kumar, and N. Birla, "Voice command recognition system based on mfcc and dtw," *International Journal of Engineering Science and Technology*, vol. 2, no. 12, pp. 7335–7342, 2010.
- [45] L. Sriratana, S. Poochong, and K. Bisalyaputra, "International journal of basic sci-

- ences and applied computing (ijbsac) international journal of basic sciences and applied computing (ijbsac).”
- [46] R. Islam, M. Tarique, and E. Abdel-Raheem, “A survey on signal processing based pathological voice detection techniques,” *IEEE Access*, vol. 8, pp. 66 749–66 776, 2020.
- [47] Y. Ninja, “Support vector machines in python-svm in python 2019.”
- [48] —, “Support vector machines in python-svm in python 2019.”
- [49] A. Subasi, *Practical guide for biomedical signals analysis using machine learning techniques: A MATLAB based approach*. Academic Press, 2019.
- [50] R. Short and K. Fukunaga, “The optimal distance measure for nearest neighbor classification,” *IEEE transactions on Information Theory*, vol. 27, no. 5, pp. 622–627, 1981.
- [51] K. Q. Weinberger and L. K. Saul, “Distance metric learning for large margin nearest neighbor classification.” *Journal of machine learning research*, vol. 10, no. 2, 2009.
- [52] S. Cost and S. Salzberg, “A weighted nearest neighbor algorithm for learning with symbolic features,” *Machine learning*, vol. 10, no. 1, pp. 57–78, 1993.
- [53] L. Breiman, “Random forests,” *Machine learning*, vol. 45, no. 1, pp. 5–32, 2001.
- [54] Z. Zhang, “Introduction to machine learning: k-nearest neighbors,” *Annals of translational medicine*, vol. 4, no. 11, 2016.
- [55] H. Almuallim, “An efficient algorithm for optimal pruning of decision trees,” *Artificial Intelligence*, vol. 83, no. 2, pp. 347–362, 1996.
- [56] L. Breiman, J. H. Friedman, R. A. Olshen, and C. J. Stone, *Classification and regression trees*. Routledge, 2017.
- [57] N. Sharma, P. Krishnan, R. Kumar, S. Ramoji, S. R. Chetupalli, P. K. Ghosh, S. Ganapathy *et al.*, “Coswara—a database of breathing, cough, and voice sounds for covid-19 diagnosis,” *arXiv preprint arXiv:2005.10548*, 2020.
- [58] C. Brown, J. Chauhan, A. Grammenos, J. Han, A. Hasthanasombat, D. Spathis, T. Xia, P. Cicuta, and C. Mascolo, “Exploring automatic diagnosis of covid-19 from crowdsourced respiratory sound data,” *arXiv preprint arXiv:2006.05919*, 2020.
- [59] V. Despotovic, M. Ismael, M. Cornil, R. Mc Call, and G. Fagherazzi, “Detection of covid-19 from voice, cough and breathing patterns: Dataset and preliminary results,” *Computers in Biology and Medicine*, vol. 138, p. 104944, 2021.

- [60] R. Islam, E. Abdel-Raheem, and M. Tarique, "A study of using cough sounds and deep neural networks for the early detection of covid-19," *Biomedical Engineering Advances*, p. 100025, 2022.
- [61] T. Rahman, N. Ibtehaz, A. Khandakar, M. S. A. Hossain, Y. M. S. Mekki, M. Ezedin, E. H. Bhuiyan, M. A. Ayari, A. Tahir, Y. Qiblawey *et al.*, "Qucoughscope: An intelligent application to detect covid-19 patients using cough and breath sounds," *Diagnostics*, vol. 12, no. 4, p. 920, 2022.